

Article

SellaMorph-Net: A Novel Machine Learning Approach for Precise Segmentation of Sella Turcica Complex Structures in Full Lateral Cephalometric Images

Kaushlesh Singh Shakya^{1,2,3} , Manojkumar Jaiswal⁴ , Julie Porteous³ , Priti K^{1,2}, Vinay Kumar⁴ , Azadeh Alavi^{3,*}  and Amit Laddi²

¹ Academy of Scientific & Innovative Research (AcSIR), Ghaziabad 201002, India; kaushlesh.csio19a@acsir.res.in (K.S.S.); priti.csio20j@acsir.res.in (P.K.)

² CSIR—Central Scientific Instruments Organisation, Chandigarh 160030, India; amittladdi@csio.res.in

³ School of Computing Technologies, RMIT University, Melbourne, VIC 3000, Australia; julie.porteous@rmit.edu.au

⁴ Oral Health Sciences Centre, Post Graduate Institute of Medical Education & Research (PGIMER), Chandigarh 160012, India; drmanojjaiswal@yahoo.in (M.J.); sonidrvinay16@gmail.com (V.K.)

* Correspondence: azadeh.alavi@rmit.edu.au

Abstract: Background: The Sella Turcica is a critical structure from an orthodontic perspective, and its morphological characteristics can help in understanding craniofacial deformities. However, accurately extracting Sella Turcica shapes can be challenging due to the indistinct edges and indefinite boundaries present in X-ray images. This study aimed to develop and validate an automated Sella Morphology Network (SellaMorph-Net), a hybrid deep learning pipeline for segmenting Sella Turcica structure and extracting different morphological types; Methods: The SellaMorph-Net model proposed in this study combined attention-gating and recurrent residual convolutional layers (AGM and RrCL) to enhance the encoder's abilities. The model's output was then passed through a squeeze-and-excitation (SE) module to improve the network's robustness. In addition, dropout layers were added to the end of each convolution block to prevent overfitting. A Zero-shot classifier was employed for multiple classifications, and the model's output layer used five colour codes to represent different morphological types. The model's performance was evaluated using various quantitative metrics, such as global accuracy and mean pixel-wise Intersection over Union (IoU) and dice coefficient, based on qualitative results; Results: The study collected 1653 radiographic images and categorised them into four classes based on the predefined shape of Sella Turcica. These classes were further divided into three subgroups based on the complexity of the Sella structures. The proposed SellaMorph-Net model achieved a global accuracy of 97.570, mean pixel-wise IoU scores of 0.7129, and a dice coefficient of 0.7324, significantly outperforming the VGG-19 and InceptionV3 models. The publicly available IEEE ISBI 2015 challenge dataset and our dataset were used to evaluate the test performance between the state-of-the-art and proposed models. The proposed model provided higher testing results, which were 0.7314 IoU and 0.7768 dice for our dataset and 0.7864 IoU and 0.8313 dice for the challenge dataset; Conclusions: The proposed hybrid SellaMorph-Net model provides an accurate and reliable pipeline for detecting morphological types of Sella Turcica using full lateral cephalometric images. Future work will focus on further improvement and utilisation of the developed model as a prognostic tool for predicting anomalies related to Sella structures.

Keywords: Sella Turcica; lateral cephalogram; morphological types; Sella segmentation; SellaMorph-Net



Citation: Shakya, K.S.; Jaiswal, M.; Porteous, J.; K, P.; Kumar, V.; Alavi, A.; Laddi, A. SellaMorph-Net: A Novel Machine Learning Approach for Precise Segmentation of Sella Turcica Complex Structures in Full Lateral Cephalometric Images. *Appl. Sci.* **2023**, *13*, 9114. <https://doi.org/10.3390/app13169114>

Academic Editors: Mitsutaka Nemoto and Katsuhiko Mikami

Received: 27 June 2023

Revised: 3 August 2023

Accepted: 7 August 2023

Published: 10 August 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The lateral cephalogram (lat ceph) is an essential tool for orthodontists when planning treatment for craniofacial deformities. Accurate traceability and identification of different

cephalometric landmark points on radiographs allow orthodontists to diagnose and evaluate orthodontic treatment. A landmark point (Sella point) marked at the centre of Sella Turcica (represented with ST) is widely used in orthodontics. Sella Turcica is an important structure in understanding craniofacial deformities by analysing its characteristics (shape, size, and volume) [1,2]. The pituitary gland is enclosed within the ST, which is a bony structure with a saddle shape. It has two clinoid projections, anterior and posterior, that extend across the pituitary fossa [3,4]. The size of these projections can vary, and when they converge, it is known as Sella Turcica Bridging (STB). This condition can lead to dental deformities and disrupt pituitary hormone secretion [5,6]. Camp has identified three types of standard Sella shapes: circular/round, flat, and oval, as depicted in Figure 1. The oval shape is the most common, while the flat shape is the least common [7–9]. In addition to these shapes, a condition called bridging occurs when the anterior and posterior clinoid lobes converge, as shown in Figure 1. This condition is known to be associated with specific syndromes and malformations that can affect dental and skeletal health, as documented in sources [10–13].

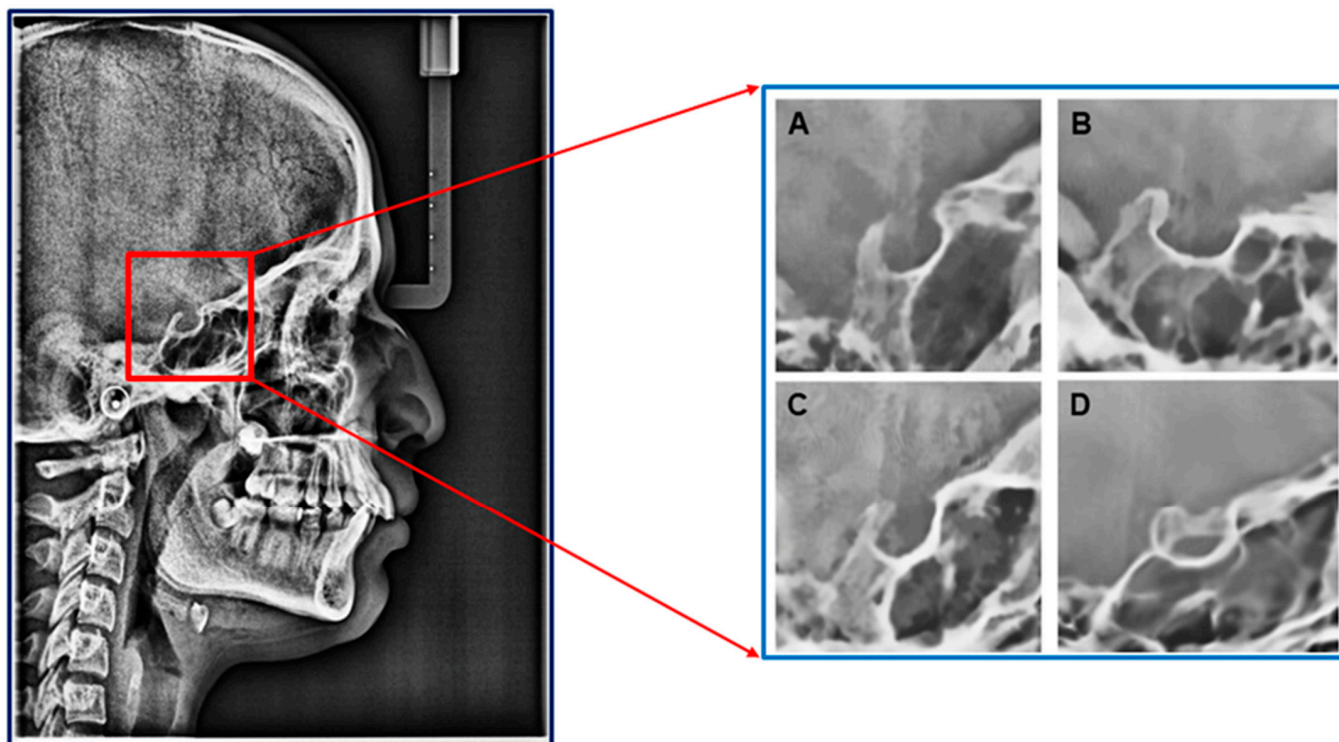


Figure 1. Pre-defined types of ST: (A) Circular, (B) Flat, (C) Oval, and (D) Bridging (additional).

Studies have used various statistical methods to analyse the morphological characteristics of the ST structure in radiographs. One commonly used method for measuring the length of ST involves calculating the distance between two points, i.e., the tuberculum sellæ (T_S) and the dorsum sellæ (D_S) projection [6,14–22]. To measure the area of ST, some researchers multiply the length and breadth of the structure, as illustrated in Figure 2. Another manual method involves sketching the outline of the ST on transparent graph paper and then placing it over a calibrated graph. Then the dental experts or practitioners count the number of squares within the outline. Finally, many studies of ST size have employed Silverman’s method, which involves measuring three variables (length, depth, and diameter) [23].

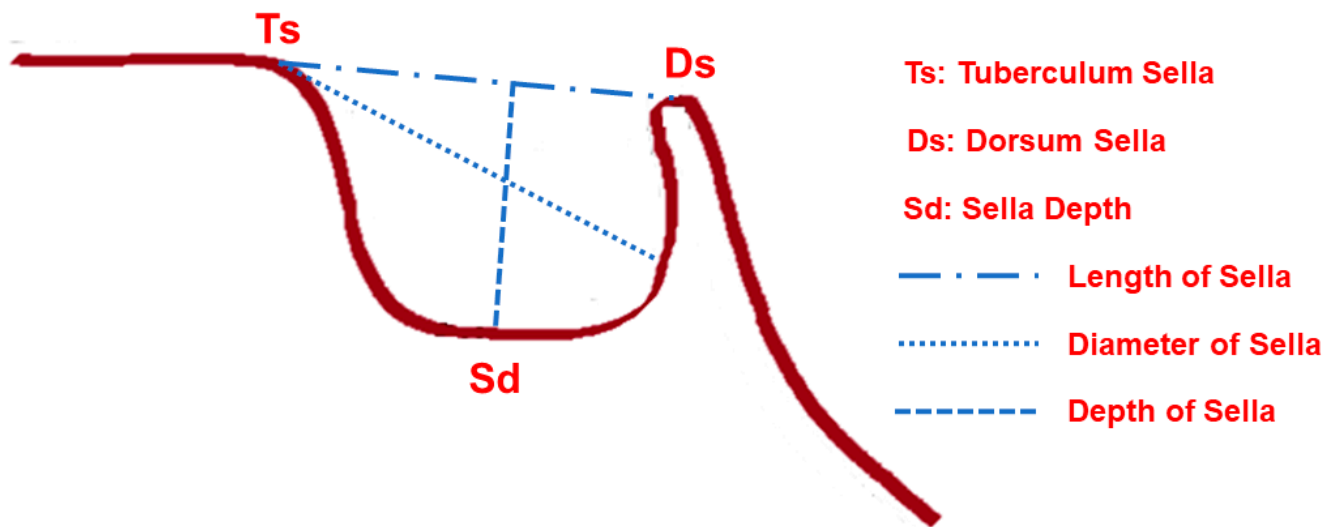


Figure 2. Manual quantification of ST size using reference lines.

Subsequently, studies have been conducted to evaluate the length, area, volume, and forms of ST in cleft patients through the use of CBCT images. These studies have confirmed that ST length is indeed shorter in cleft patients [15,16]. To locate the Sella point, 3D maxillofacial software is commonly used to convert 3D CBCT data into 2D ceph datasets. Another method of identifying the Sella point involves the use of 3D models based on a new reference system, which can be more effective and reliable than converting 3D image datasets into 2D [17,18]. However, using Cone-Beam-CT and digital volume tomography to measure Sella size may not be practical for routine use.

In recent years, there has been a growing interest in the use of Artificial Intelligence (AI) in the biomedical field due to its ability to mimic cognitive processes like learning and decision-making, which are similar to those of humans. One popular type of AI model is the Convolutional Neural Network (CNN), which can detect important image features with high computational efficiency, closely resembling biological visual processing. However, in orthodontics, manually inspecting ceph X-ray images for morphological features of ST or using 3D models based on Cone-Beam-CT scans can be time-consuming, require expert intervention, and expose patients to radiation. Therefore, in addition to clinicians, AI can be used as an assistive tool to identify subtle details that may not be visible to the naked eye.

Therefore, we identified the gap and were the first to work on AI-based learning of ST features using radiographic images (easily accessible) [24,25]. In reference to our study, a recent study was conducted on identifying ST features automatically using CBCT images [26]. The challenge of our study is working on radiographs themselves, as the characteristics of ST vary from patient to patient in the considered dataset; therefore, in our study, ST differs from regular to complex altered structures. Thus, the proposed study has grouped the dataset into three subgroups based on the complexity of the ST features: regular ST, moderately altered ST, and complex altered ST, which require the extraction of minute details.

This study presents a novel hybrid segmentation pipeline for identifying the different morphological features of the Sella Turcica (ST) in lateral cephalometric radiographs. The method uses a hybrid encoder-decoder convolution structure with densely connected layers as the encoder to predict the initial mask. This predicted mask is combined with the input image and the encoder-decoder network, consisting of recurrent residual convolutional layers (RrCLs), squeeze-and-excite blocks (SEs), and an attention gate module (AGM), to provide the final improved segmentation mask by eliminating undesirable features. Additionally, a zero-shot linear classifier (ZsLC) is integrated for the accurate and distinctive classification of ST characteristics.

The introduced method leverages the novel combination of SE, AGM, and ZsLC components, not previously studied in this context, making the segmentation process more robust, accurate, and adaptable. The hybrid model effectively captures intricate patterns, focuses on relevant regions, and seamlessly addresses the complexities present in the pre-defined ST classes, which is difficult through conventional methods. The model works on deep layers through extracting minute features and distinct between different Sella types automatically. The unique combination of different modules allows the model to achieve state-of-the-art segmentation results and shows potential for a wide range of segmentation tasks in the future. The proposed method is evaluated on two standard datasets. The first is a self-collected dataset comprising four pre-defined classes (circular, flattened, oval, and bridging ST types) and three sub-classes (regular, moderately altered, and complex altered ST). The second dataset is the publicly available IEEE ISBI 2015 challenge dataset for four-class segmentation. The proposed method outperforms the existing state-of-the-art CNN architectures in segmenting the morphological characteristics of lateral cephalometric radiographs.

The study's significant contributions are as follows:

- We introduced Sella Morphology Network (SellaMorph-Net), a novel hybrid pipeline for accurately segmenting different Sella Turcica (ST) morphological types with precise delineation of fine edges.
- The approach utilised a complete CNN framework with efficient training and exceptional precision, eliminating the need for complex heuristics to enhance the network's sensitivity for focused pixels.
- Attention Gating Modules (AGMs) were incorporated into our proposed method, allowing the network to concentrate on specific regions of interest (RoI) while preserving spatial accuracy and improving the feature map's quality.
- The results demonstrate meaningful differentiation of ST structures using a Zero-shot linear classifier, along with an additional colour-coded layer depicting ST structures in lateral cephalometric (lat ceph) radiographic images.

2. Materials and Methods

2.1. Materials

This analytical study was approved by the Institute Ethical Committee (IEC-09/2021-2119) of the Postgraduate Institute of Medical Education and Research, Chandigarh, India. The research team collected 1653 radiographic images of 670 dentofacial patients and 983 healthy individuals for the study. The images were categorised based on pre-defined morphological shapes of ST: circular, flat, oval, and bridging. The 450 images were considered for the circular, oval, and bridging shapes, and 303 were flattened. Further, based on the complexity of these structures, we sub-grouped the Sella shapes into regular ST, moderately altered ST, and complexly altered ST to study the morphological characteristics of this important structure.

To ensure privacy and anonymity for participants, radiographs were randomly chosen without considering any personal information, such as age and gender. The cephalometric radiographs were captured using the Carestream Panorex and followed standard procedures.

Pre-Processing

This study utilised image processing techniques to improve the quality of cephalometric radiographs and acquire more comprehensive information on the Sella Turcica (ST). The process involved several steps to enhance grayscale medical images. Firstly, the image resolution was adjusted to 512 px, and a weighted moving averaging filter was applied to eliminate noise [27]. This filter used a 3×3 weighted grid size combined with the original radiographs. To further enhance the image's details, we modified the contrast ratio using the transformed intensity of the image [28]. Lastly, the Sobel operator enhances the edges of the image by calculating the gradient along the 'y' and 'x' dimensions [29,30]. We then obtained a sharper image with increased edge clarity by computing convolution

using the original image [31]. Lastly, we performed a negative transformation using the enhanced image as a reference for manual labelling (annotation). These pre-processing techniques resulted in reduced training complexity and more accurate analysis of cephalometric radiographs.

A team consisting of a radiographic specialist and two dental clinicians utilised a cloud-hosted artificial intelligence platform named Apeer to meticulously annotate images pixel-by-pixel, as shown in Figure 3 [24]. This platform allows for the precise storage and annotation of images with high resolution. The process involves saving segmented portions with high accuracy following rough calculations. Then, a precise binary mask is generated by assigning sub-pixel values to the chosen boundary pixels and applying pixel-level morphological operations through border modification. Apeer's workflow involves re-annotating unmarked pixels, filling the image background with a region-filling technique, clustering pixels with specific labels using a connected component labelling approach, and ultimately saving the mask images for training.

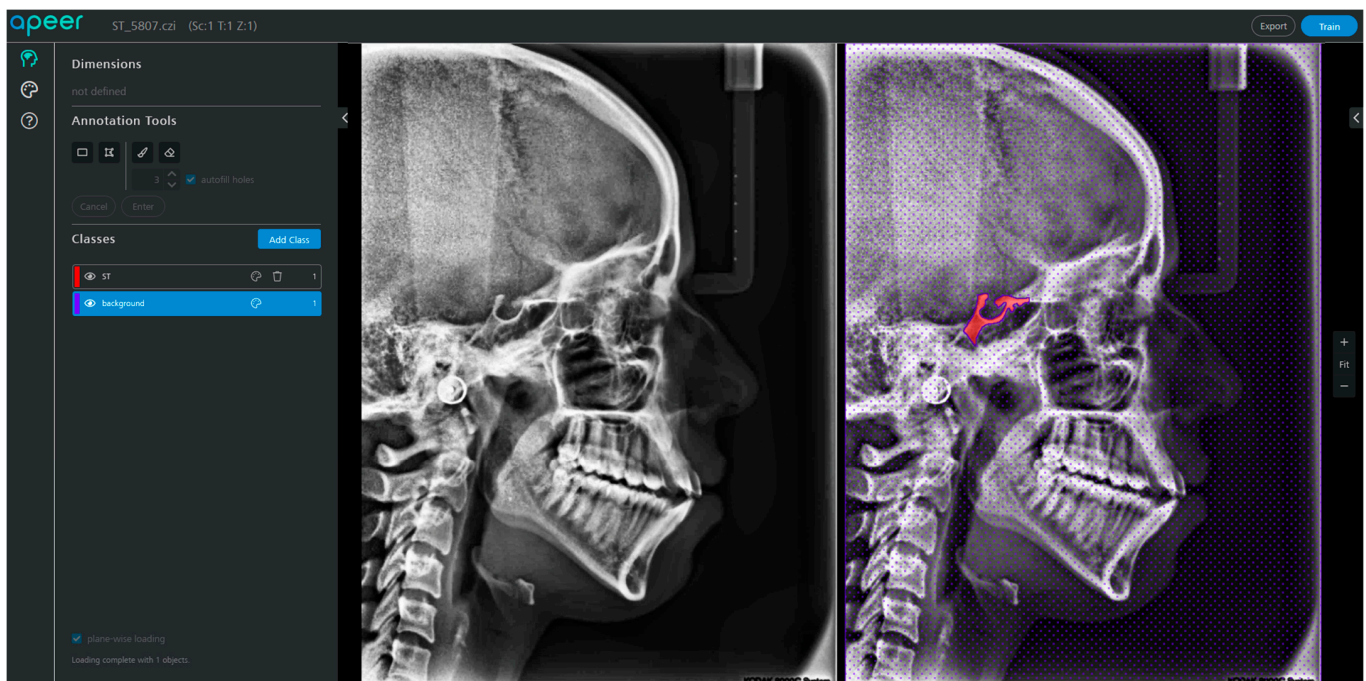


Figure 3. Screenshot of the cloud-based Apeer annotation platform user interface, showing a raw and annotated Cephalometric X-ray image.

2.2. Methods

In this section, we begin by presenting an overview of the current state-of-the-art methods utilised for performance comparison with the proposed method, as discussed in Section 2.3. Subsequently, in Section 2.4, we outline the proposed pipeline in detail, providing a comprehensive explanation of its components and workflow. Finally, in Section 2.5, we describe the evaluation metric employed to assess the performance and effectiveness of the proposed method, elucidating the specific criteria and measurements used in the evaluation process.

2.3. Deep Learning Methods

Deep learning (DL) approaches utilise hierarchical and dynamic feature representations that leverage multiple levels of abstraction instead of shallow learning techniques. However, despite the widespread success of DL compared with shallow learning in numerous domains, the scarcity of labelled data poses challenges to the practical implementation of DL. However, in the medical field, Convolutional Neural Networks (CNNs) have been

successfully used to segment and classify the Sella structure in cephalometric radiographs, despite the challenge of limited labelled data for training. Recently, researchers have employed few techniques for segmenting and classifying the Sella Turcica in medical images [26]. Shakya et al. used a U-Net model with a different pre-trained model as the encoder for cephalometric radiographs, while Duman et al. adopted the Inception V3 model to segment and classify the structure using CBCT images, with reference to Shakya's work [24–26]. Additionally, Feng et al. utilised a U-Net-based approach that incorporated manual measurements to evaluate the length, diameter, and depth of the Sella Turcica [32].

The published studies have developed different methods to segment and classify the Sella Turcica part from cephalometric ROI images instead of considering full cephalometric images, as suggested by Shakya, to identify dentofacial anomalies related to the Sella Turcica in future work. It is important to note that a complete cephalometric analysis with accurate feature extraction of the Sella Turcica is necessary to evaluate dentofacial anomalies comprehensively. Recently, Kok et al. proposed traditional machine learning algorithms such as SVM, KNN, Naive Bayes, Logistic Regression, ANN, and Random Forest for analysing and determining growth development in cephalometric radiographs [33]. However, this approach is slow in performance and complex in terms of comprehending the structure of the algorithm. Furthermore, Palanivel et al. and Asiri et al. have suggested a training technique that uses a genetic algorithm. The method involves training a network on input cephalometric X-rays in the first attempt, followed by multiple attempts with the same input images to determine the optimal solution [34,35]. However, these techniques are costly in terms of computational resources and require extra post-processing steps for better results. On the other hand, our study proposes a hybrid encoder-decoder method that can be trained without post-processing, making it a more efficient solution. The results of the proposed SellaMorph-Net model are further compared with other popular state-of-the-art models, e.g., VGG-19 and InceptionV3, in terms of mean IoU, dice coefficient, training accuracy, and validation accuracy, with corresponding loss values defined by ST structure segmentation and classification.

VGG-19 and InceptionV3

The VGG-19 model uses a combination of convolutional and fully connected layers to improve the process of extracting features [24,25]. In addition, it uses Max-pooling, rather than average pooling, for down-sampling and the SoftMax activation function for classification [36–38].

The InceptionV3 model contains a classification element and a learned convolutional base [24,26,39]. The classification component comprises a global average pooling layer, a fully connected layer that uses softmax activation, and dropout regularisation to prevent overfitting [40,41]. To extract features from the convolutional layers, 3×3 filters are used for down-sampling with max-pooling, followed by softmax activation and max-pooling. The classification element consists of a classifier that is fully connected, along with layers incorporating dropout. The global average pooling layer is used instead of a flattened layer to maintain spatial information and reduce the number of parameters in the model [42].

2.4. Proposed Method

The proposed framework for the segmentation task includes a hybrid encoder and decoder, as illustrated in the figure. The encoder comprises convolutional and dense layers to enhance the extraction of characteristics from the input images. After extracting details, they are passed through a squeeze-and-excite (SE) block, which only sends the required details to the decoder for creating the preliminary segmentation mask. The SE block promotes feature representation by capturing inter-channel dependencies. As a result, the model can prioritise crucial features while suppressing less informative ones. Integrating softmax and ReLU activation functions signifies the utilisation of non-linear mapping, while incorporating the recurrent residual convolutional layer (RrCL) helps capture deeper layer characteristics by enhancing the RrCL block by incorporating two RCL units. The

RCL expansion strategy involves expanding the RCL to two-time steps ($T = 2$) to extract valuable features further.

Additionally, the features of the encoder are improved using Attention Gating Modules (AGMs) that focus on specific areas of interest without losing the spatial resolution of the feature maps. The resulting feature maps are then up-sampled and sent to the decoder, along with a zero-shot linear classifier (ZsLC) that generates the final classified segmentation output with improved accuracy. The flow diagram in Figure 4 illustrates the process of the proposed methodology.

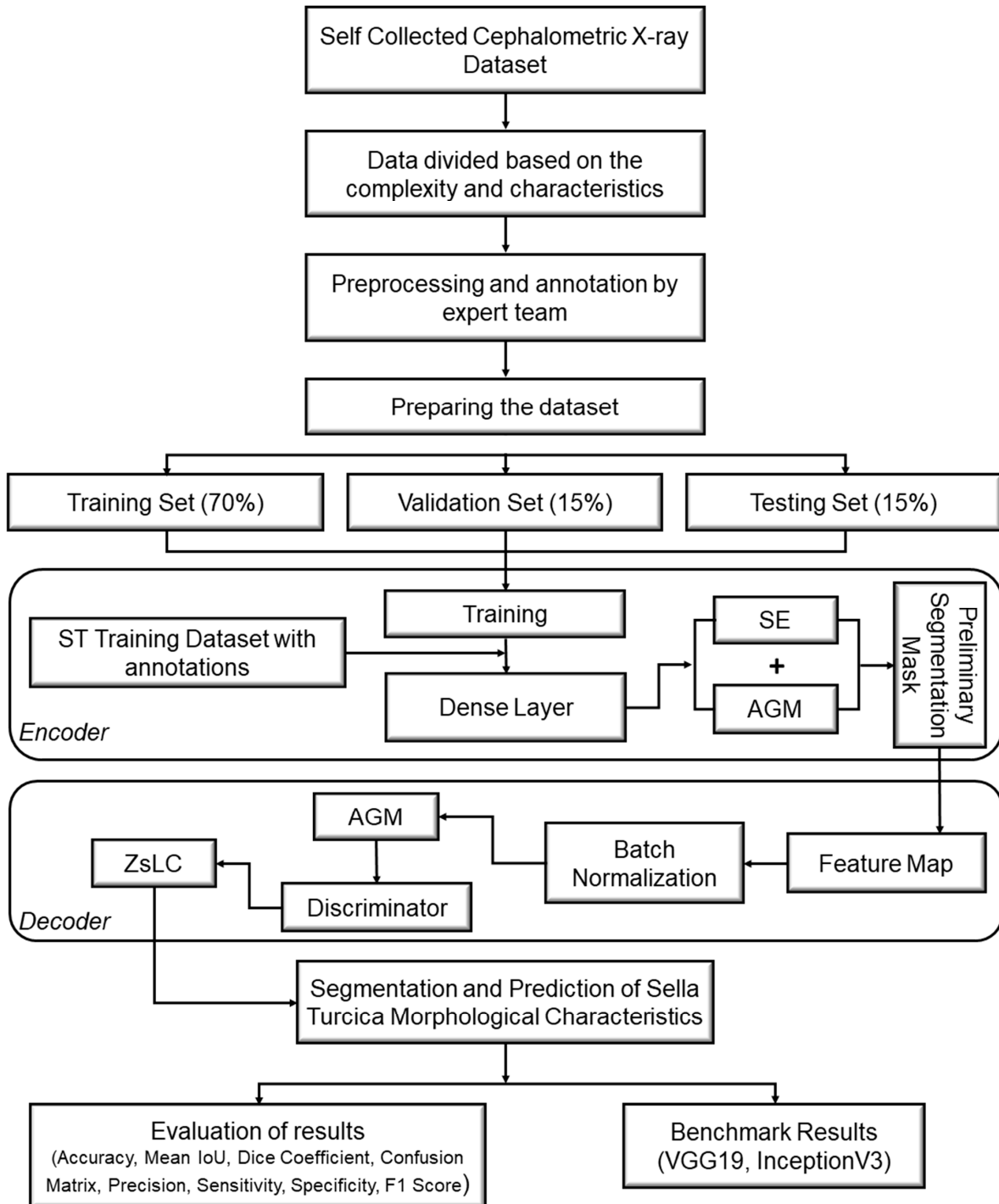


Figure 4. Process flow diagram of the proposed SellaMorph-Net model.

2.4.1. Encoder

The present study constructed a DL encoder from scratch using a series of conv and max-pooling layers, repeated four times with the same number of filters. To prevent overfitting, we added a dropout layer after each max-pooling layer. Instead of a dense layer, we used a single conv layer that acted as a bottleneck in the network, separating the encoding and decoding networks. The Rectified Linear Unit (ReLU) activation function is employed to represent the non-linear mapping into the model [43,44]. The SE block, which is exhibited in Figure 5, is then involved in enhancing the feature maps and learning accurate boundaries of Sella structures in cephalometric radiographic images [44,45]. In the squeezing step $Z_{squeeze} \in R^c$ reduces the spatial dimensions of feature maps to a 1×1 representation using global average pooling, while in the exciting step $Z_{excite} \in R^c$, learnable weights are applied to the squeezed feature maps to recalibrate channel-wise information. The mathematical equation below illustrates the SE combination:

$$SE = \gamma_2 \left(W_2 \times \gamma_1 \left(W_1 \times \frac{1}{H \times W} \sum_i \sum_j X_{ijc} \right) \right) \otimes X, \tag{1}$$

where γ_2 denotes the softmax activation function, γ_1 denotes the ReLU activation function, W_1 and W_2 are learnable weight matrices, H and W are spatial dimensions, and \times denotes matrix multiplication.

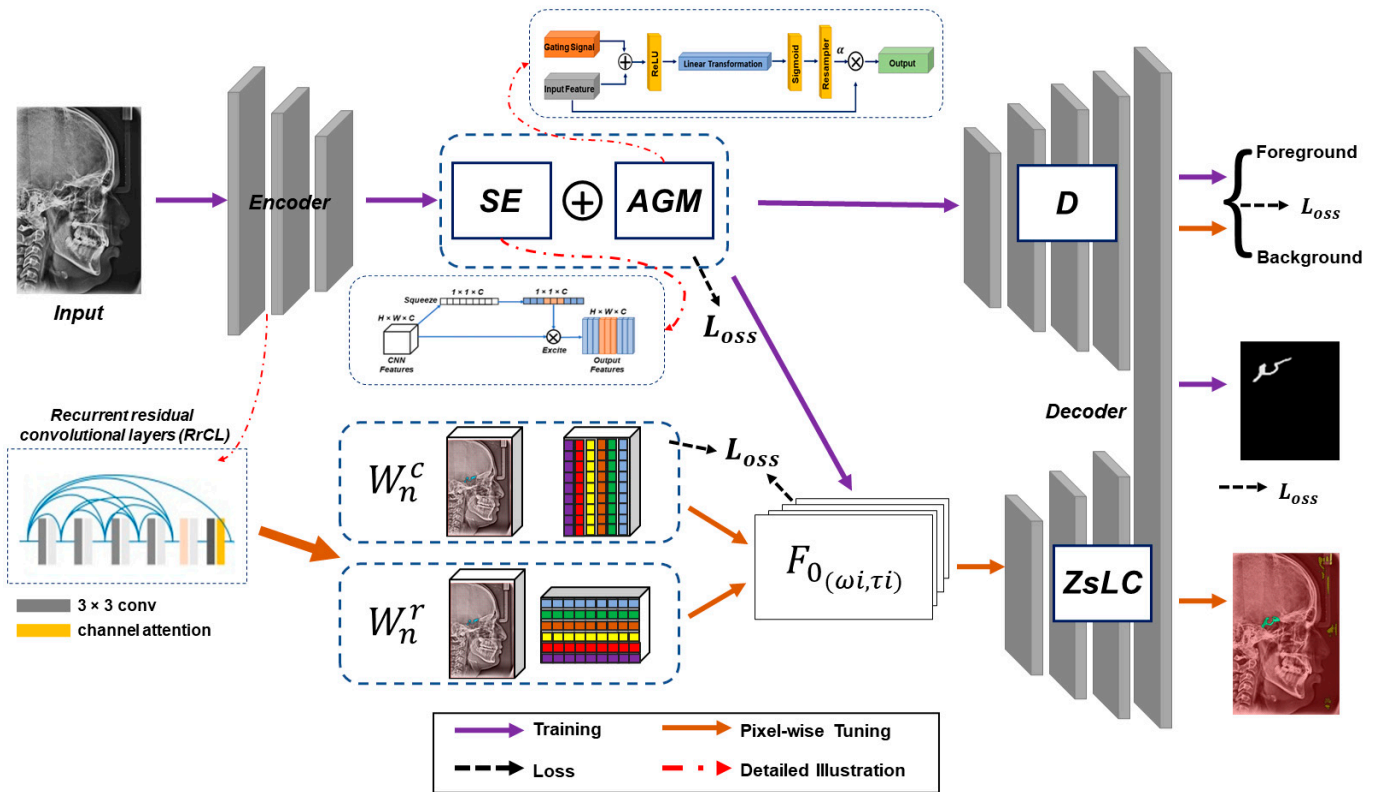


Figure 5. Outline of the proposed method. The proposed model contains an encoder part incorporated with recurrent residual convolutional layer RrCL, squeeze-and-excite SE block, attention gating modules AGM, discriminator D, and zero-shot linear classifier ZsLC.

Moreover, to enhance the encoder’s ability to integrate context information, a recurrent residual convolutional layer (RrCL) is incorporated in individual steps [44,46]. The RrCL consists of 3×3 convolutions, which allows for a deeper model and better accumulation of feature representations across different time steps. However, the use of RrCL leads to an increase in the number of feature maps and a reduction in size by approximately

half. In addition, batch normalisation (BN) is applied after each convolution operation to regularise the model and reduce internal covariate shifts [44,47]. In the RrCL block, the network output O_i at time step t_s can be mathematically described by the equation below, assuming ω_i as the input of the i th layer and (p, q) as the localised pixel in the input on the n th feature map, as illustrated in the accompanying figure.

$$O_{i(p,q,t_s)} = (W_n^c)^T \times \omega_{i(p,q,t_s)} + (W_n^r)^T \times \omega_{i(p,q,t_{s-1})} + b_n, \tag{2}$$

where O_i is the output at position (p, q) in the n th feature map of the i th layer at time step t_s , with $(W_n^c)^T$ and $(W_n^r)^T$ as weight matrices, ω_i as the input at time step t_s , $\omega_{i(p,q,t_{s-1})}$ as the input at time step t_{s-1} , and b_n as the bias term for the n th feature map. The recurrent convolutional unit's output $\omega(i + 1)$ is then sent to the residual unit, which is represented by the following equation:

$$\omega(i + 1) = \omega_i + F_{O(\omega_i, \tau_i)} \tag{3}$$

The RrCL takes inputs ω_i and produces outputs $F_{O(\omega_i, \tau_i)}$, which are utilised in the encoder and decoder layers of the proposed network's down- and up-sampling layers. The subsequent sub-sampling or up-sampling layers rely on the final output $\omega(i + 1)$ as their input.

2.4.2. Decoder

The decoder blocks in the DL network use transposed convolution for up-sampling the feature maps, effectively increasing their size. However, spatial resolution is preserved, and output feature map quality is improved through fully connected connections between the encoder's and the decoder's output feature maps. Each decoding step involves up-sampling the previous layer's output using the Rr-Unit, which halves the feature map count and doubles the size. The final decoder layer restores the feature map size to the original input image size. Batch normalisation (BN) is employed during up-sampling for stability and convergence acceleration during training, and the BN output is then passed to the Attention Gating Modules (AGMs) [44,48]. Additionally, a zero-shot linear classifier is incorporated with AGMs to further enhance the model's classification capabilities [49].

The network utilises AGMs to enhance the down-sampling output and combine it with the equivalent features from the decoder, which prioritises high-quality features and helps to preserve spatial resolution, ultimately enhancing the quality of the feature maps. In addition, a zero-shot linear classifier is incorporated to enhance classification performance further. The combination of AGMs and the linear classifier improves feature representation and classification accuracy, making the approach well-suited for multi-class classification tasks. First, the attention values are calculated for each individual pixel of all the input features ω_p^i . Then, the gating vector δ_p is applied to each pixel to capture pixel-wise information for fine-grained details and local contextual information, resulting in accurate and discriminative classification ST_{m+1} outcomes. The additive equation is as follows:

$$\beta_p^i = S_{max} \left(ST_{m+1} \left(\gamma_2 \left(\varphi^T \left(\gamma_1 \left(W_{\omega}^T \omega_p^i + W_{\delta}^T \times \delta_p + B_{\delta} \right) \right) + B_{\varphi} \right) \right) \right), \tag{4}$$

where the activation functions ReLU (γ_1) and sigmoid (γ_2) are utilised in the Attention Gating Modules (AGMs), along with linear transformation weights (W_{ω}) and (W_{δ}) and biases (B_{δ}) and (B_{φ}). The introduction of ST_{m+1} enables a tailored transformation of attention values for multi-class classification, thereby improving the model's accuracy and performance in handling multi-class tasks. Prior to applying the activation functions γ_1 and γ_2 , attention values are transformed using the transpose of variable φ , identified as φ^T . The final prediction output is generated using a SoftMax activation function.

2.5. Performance Evaluation Metrics

This study used six image operations that included different transformations such as scaling, flipping 360 degrees, Gaussian blur addition, histogram colour adjustment, Gaussian noise addition, and elastic 2D modification. In addition, the magnitude and scale of the elastic deformation were adjusted using Alpha and Sigma, and the study was divided into seven values for easy magnitude range location. The experiments were conducted on a professional Windows operating system with a Ryzen computer that had 32 GB of memory and a 16 GB RTX graphic processor. The network could predict segmented images of 512 px and accept tests on images of any size. After testing various combinations, it was determined that a batch size of 4, along with learning rates of 0.001, beta values of 0.9 and 0.999, and 2170 iterations for 70 epochs, yielded satisfactory results [50]. The neural network framework Keras, which is open-source and written in Python, was used alongside TensorFlow.

To measure the similarity between the predicted and true segmentations, the mean Intersection Over Union (IoU) was used, which is a comparison metric between sets [51]. It assesses the accuracy of segmentations predicted by different models by comparing their outputs to the ground truth. Specifically, the IoU for two sets, M and N, can be defined as follows:

$$J_{index} = \frac{|M \cap N|}{|M \cup N|} = \frac{|M \cap N|}{|M| + |N| - |M \cap N|}, \quad (5)$$

where the images are composed of pixels. As a result, the final equation can be adjusted to represent discrete objects in this manner:

$$J_{index} = \frac{1}{num} \sum_{C=1}^4 \omega_C \sum_{P_i=1}^{num} \left(\frac{l_{P_i}^C \hat{l}_{P_i}^C}{l_{P_i}^C + \hat{l}_{P_i}^C - l_{P_i}^C \hat{l}_{P_i}^C} \right) \quad (6)$$

To address the class imbalance, the study incorporated class weights ω_C into the network. For this problem, $\omega_C = 1$ was assigned to classes $C \in [1, \dots, 4]$ for simplicity. To enhance robustness and evaluate the loss performance, the study employed the Binary Logistic loss ($BLog_{loss}$), which is shown in Equation (7). This formula includes C_T as the actual mask image, C_{Ts} as a single element of that mask, C_P as the projection of the output image, and C_{Ps} as a single feature of that projection. In this equation, $l_{P_i}^C$ represents the binary value (label), and $\hat{l}_{P_i}^C$ represents the estimated probability for the pixel P_i of class C .

$$BLog_{loss} = \sum_s -(C_{Ts} \log C_{Ps} + (1 - C_{Ts}) \log(1 - C_{Ps})) \quad (7)$$

when C_{Ps} has a value of 1 or 0, the result of $\log(0)$ is undefined. To avoid this, the values of C_P are restricted to a range of $[\delta, 1 - \delta]$. The Keras framework handles this by setting δ to 1×10^{-7} .

In order to further evaluate the performance of segmentation models, this study employed the dice similarity coefficient, a statistical metric. The dice coefficient, also known as the Sørensen–Dice coefficient or F1 score, is calculated by multiplying the overlapping area of actual and predicted images by two and then dividing it by the sum of the pixels in those images [52]. By measuring the spatial overlap, the dice coefficient provides a measure of accuracy and predictability for the segmentation model:

$$DSC = \frac{2TP}{(2TP + FN + FP)} \quad (8)$$

Unlike the dice coefficient, which overlooks many instances of improper segmentation, the IoU (Intersection over Union) only takes into account one occurrence of imprecise segmentation. This means that an algorithm with occasional errors will have a lower IoU

score than the dice coefficient. However, the dice coefficient is a better indicator of average performance when compared with each individual.

After analysing the performance metrics, the morphological types of the ST were classified using a colour-coded system for each of the defined classes, including the background. Figure 6 illustrates this colour-coding representation for the ST types.

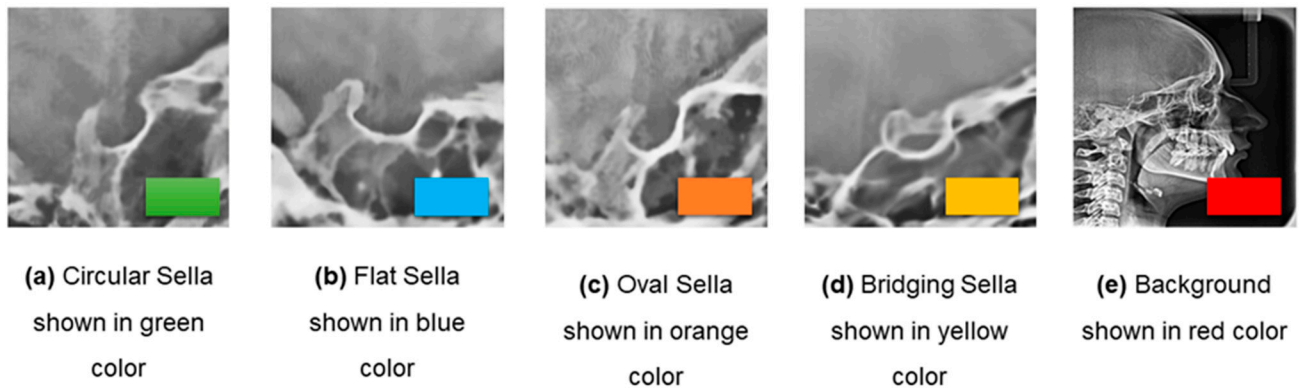


Figure 6. Illustration of different morphological shapes of ST assigned with Colour-Coded scheme.

3. Results

This section accounts for the results obtained from the proposed method's quantitative and qualitative evaluation, utilising three subgroups of the collected cephalometric dataset. The cephalometric dataset was divided into three subgroups based on the complexity of the ST, ranging from regular to complex altered structures. These subgroups include regular ST, moderately altered ST, and complex altered ST, which require more detailed extraction. To obtain optimal results, different techniques of data pre-processing, augmentation, and hyper-parameter fine-tuning were implemented. The collected dataset consists of 1653 images, which is an odd number. Therefore, a split ratio of 70:15:15 was applied for model evaluation, and a k-Fold cross validation method, where k size is five, was used for equal size distribution. The quantitative results compare how well the proposed method can identify different anatomical structures in datasets with varying levels of complexity in Sella features. Meanwhile, the qualitative analysis provides an observable illustration of the performance of the proposed method.

The proposed hybrid SellaMorph-Net model's quantitative results are presented in Tables 1–3. This model was specifically designed to identify and classify different types of ST. The model reduces the input size of the images and then extracts features from the dataset to identify Sella types. The model was created from scratch using a consistent batch size of 4. It was optimised using AGM with a momentum value of 0.99. The learning rate for training was initially set at 0.001 and gradually decreased with each training epoch until it reached a specific number, which is 0.00001. The evaluation results of the proposed hybrid model on the regular ST subgroup are presented in Table 1. Table 2 shows the results for the moderately altered ST subgroup, and Table 3 presents the results for the complexly altered ST subgroup using the cephalometric radiographic dataset employed in this study.

First, we trained the proposed SellaMorph-Net model for the divided subgroups of the cephalometric radiographic dataset with augmentation and 2170 iterations and indicated this in Tables 1–3. At the end of the 70th epoch and 2170th iteration, the time elapsed for the regular ST subgroup was 02 h 25 min 44 s, while the time elapsed for the moderately altered ST subgroup and complex altered ST subgroup was 02 h 29 min 07 s and 02 h 32 min 17 s, respectively. Similarly, the mini-batch accuracy values of the regular ST subgroup were equal to 99.57%, while these values were 99.51% and 99.49% for the moderately altered ST and complexly altered ST subgroups, respectively.

Table 1. Evaluation metrics of the proposed SellaMorph-Net model for regular ST subgroup of the dataset.

Epoch	Iteration	Time Elapsed (hh:mm:ss)	Mini-Batch Accuracy	Base LEARNING Rate
1	1	00:03:27	88.45%	0.001
10	310	00:20:43	99.13%	0.001
20	620	00:41:26	99.19%	0.001
30	930	01:02:09	99.23%	0.001
40	1240	01:22:52	99.31%	0.001
50	1550	01:43:35	99.36%	0.0001
60	1860	02:04:18	99.43%	0.0001
70	2170	02:25:44	99.57%	0.00001

Table 2. Evaluation metrics of the proposed SellaMorph-Net model for moderately altered ST subgroup of the dataset.

Epoch	Iteration	Time Elapsed (hh:mm:ss)	Mini-Batch Accuracy	Base Learning Rate
1	1	00:02:57	86.67%	0.001
10	310	00:20:07	98.93%	0.001
20	620	00:42:29	99.07%	0.001
30	930	01:04:59	99.13%	0.001
40	1240	01:25:13	99.29%	0.0001
50	1550	01:46:03	99.34%	0.0001
60	1860	02:10:57	99.44%	0.0001
70	2170	02:29:07	99.51%	0.00001

Table 3. Evaluation metrics of the proposed SellaMorph-Net model for complex altered ST subgroup of the dataset.

Epoch	Iteration	Time Elapsed (hh:mm:ss)	Mini-Batch Accuracy	Base Learning Rate
1	1	00:02:42	84.27%	0.001
10	310	00:21:27	98.64%	0.001
20	620	00:43:57	98.73%	0.001
30	930	01:07:48	98.97%	0.0001
40	1240	01:27:03	99.15%	0.0001
50	1550	01:50:00	99.23%	0.0001
60	1860	02:13:53	99.37%	0.00001
70	2170	02:32:17	99.49%	0.00001

The experiments were conducted to evaluate the performance of the proposed model through a comparative study using two state-of-the-art pre-trained models, e.g., VGG-19 and InceptionV3, for identifying different morphological shapes of ST. According to Table 4, the proposed CNN model efficiently identifies various shapes of the Sella, ranging from regular to complex altered subgroups. This is achieved by distinguishing between bridging, circular, oval, and flat classes.

Table 4. Class based training and validation IoU results of VGG-19, InceptionV3, and SellaMorph-Net (proposed CNN approach).

Class	Bridging		Circular		Flat		Oval	
	Training IoU	Validation IoU	Training IoU	Validation IoU	Training IoU	Validation IoU	Training IoU	Validation IoU
VGG-19	0.4372 ± 0.07	0.4526 ± 0.05	0.3668 ± 0.04	0.3694 ± 0.03	0.40 ± 0.03	0.4327 ± 0.03	0.2433 ± 0.06	0.2447 ± 0.03
InceptionV3	0.4357 ± 0.07	0.4381 ± 0.06	0.3847 ± 0.03	0.40 ± 0.03	0.4325 ± 0.03	0.3793 ± 0.07	0.2976 ± 0.03	0.2484 ± 0.07
SellaMorph-Net	0.6539 ± 0.03	0.6493 ± 0.02	0.6370 ± 0.01	0.5933 ± 0.01	0.6517 ± 0.01	0.5876 ± 0.03	0.5663 ± 0.02	0.5718 ± 0.03

Furthermore, the study discovered that it was possible to achieve stable and high-score graphs for identifying single non-linear shapes in a specific subgroup of cephalometric radiographic data. However, the state-of-the-art models VGG-19 and InceptionV3 were less efficient on the cephalometric dataset, sometimes identifying outliers instead of actual morphological types and leading to false classification.

In addition, the results demonstrate that SellaMorph-Net has a faster convergence of the loss function and significant mean pixel-wise IoU efficacy than the state-of-the-art models under the same training conditions. Furthermore, the proposed model's training and validation score variation is less, between ± 0.03 and ± 0.01 , compared with the state-of-the-art VGG-19 and InceptionV3 (± 0.07 and ± 0.03) models. These outcomes show that SellaMorph-Net is more reliable in identifying ST types in cephalometric images based on the specified colour scheme and class provided, which is crucial in accurately detecting different morphological types of ST.

Table 5 presents a summary of the performance results achieved by three models on the cephalometric image dataset. The performance metrics included training and validation accuracy, mean/average IoU, and dice coefficient scores. Among the three models, SellaMorph-Net had the highest mean IoU at 71.29% and a dice coefficient of 73.24% in Sella-type segmentation, higher than VGG-19 and InceptionV3. The Dice Coefficient score measures how similar the predicted values are to the ground truth values and is identical to the IoU score. Additionally, SellaMorph-Net is a hybrid model that is hyperparameter-tuned on the cephalometric X-ray image subgroup datasets, which resulted in a higher IoU and dice score compared with the other models.

Table 5. Performance results of VGG-19, InceptionV3, and proposed models SellaMorph-Net.

Model	Accuracy (%)	Pixel-Wise IoU		Mean IoU	Dice Coefficient
		Train IoU	Validation IoU		
VGG-19	84.479	0.5637	0.5719	0.5974	0.6188
InceptionV3	87.283	0.5654	0.5903	0.6029	0.6324
SellaMorph-Net	97.570	0.7329	0.6973	0.7129	0.7324

We assessed the proposed SellaMorph-Net model performance through a comparative analysis with VGG19 and Inception V3, which are state-of-the-art models. The following performance matrix evaluates all the considered models (Precision (P_s), Sensitivity (S_n), Specificity (S_p), and F1-score ($F1_s$)) shown in Figure 7 and Table 6.

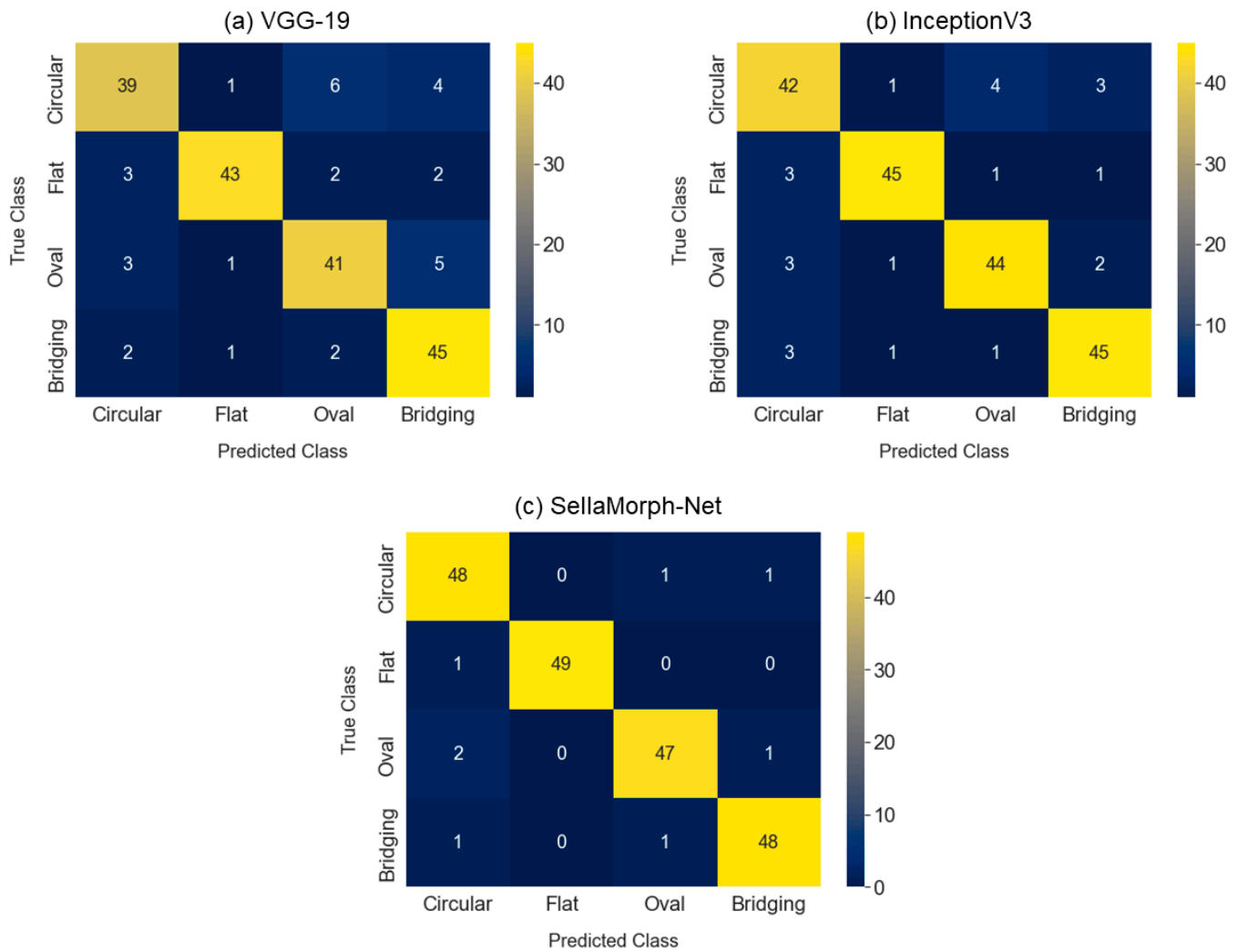


Figure 7. Confusion Matrix for Four Classes of Lateral Cephalometric X-ray Images using (a) VGG19, (b) InceptionV3, and (c) Proposed SellaMorph-Net: Predicted vs. True Class.

Table 6. Performance Matrix for Predictive Analysis on Proposed Model, VGG19, and InceptionV3 model.

Model	P_s	Sn_s	Sp_s	$F1_s$
VGG-19	76.163	82.096	90.419	73.448
InceptionV3	86.154	90.667	89.922	90.090
SellaMorph-Net	96.832	98.214	97.427	97.314

The findings demonstrate that SellaMorph-Net achieved significantly higher accuracy than the state-of-the-art models. The proposed model outperformed in correctly identifying positive instances (Sensitivity), accurately identifying negative instances (Specificity), and achieving a balanced performance between precision and recall (F1-score). The confusion matrix and performance graphs suggest that the SellaMorph-Net model provides more accurate characterizations ($P_s = 96.832$, $Sn_s = 98.214$, $Sp_s = 97.427$, $F1_s = 97.314$) than the compared models VGG19 ($P_s = 76.163$, $Sn_s = 82.096$, $Sp_s = 90.419$, $F1_s = 73.448$) and InceptionV3 ($P_s = 86.154$, $Sn_s = 90.667$, $Sp_s = 89.922$, $F1_s = 90.090$). Thus, the proposed model seems more accurate and reliable for predicting the Sella structure.

In order to assess the qualitative performance and efficacy of the suggested method, the study presents visual outcomes for the segmentation of regular, moderately altered, and

complexly altered ST structures in Figures 8–10. These structures were selected based on the cephalometric dataset’s complexity. In addition, the study compared the accuracy of our hybrid SellaMorph-Net model against VGG-19 and InceptionV3. The results demonstrated that while VGG-19 and InceptionV3 were able to provide an approximate ST type, they were unable to predict the detailed morphological structures of ST.

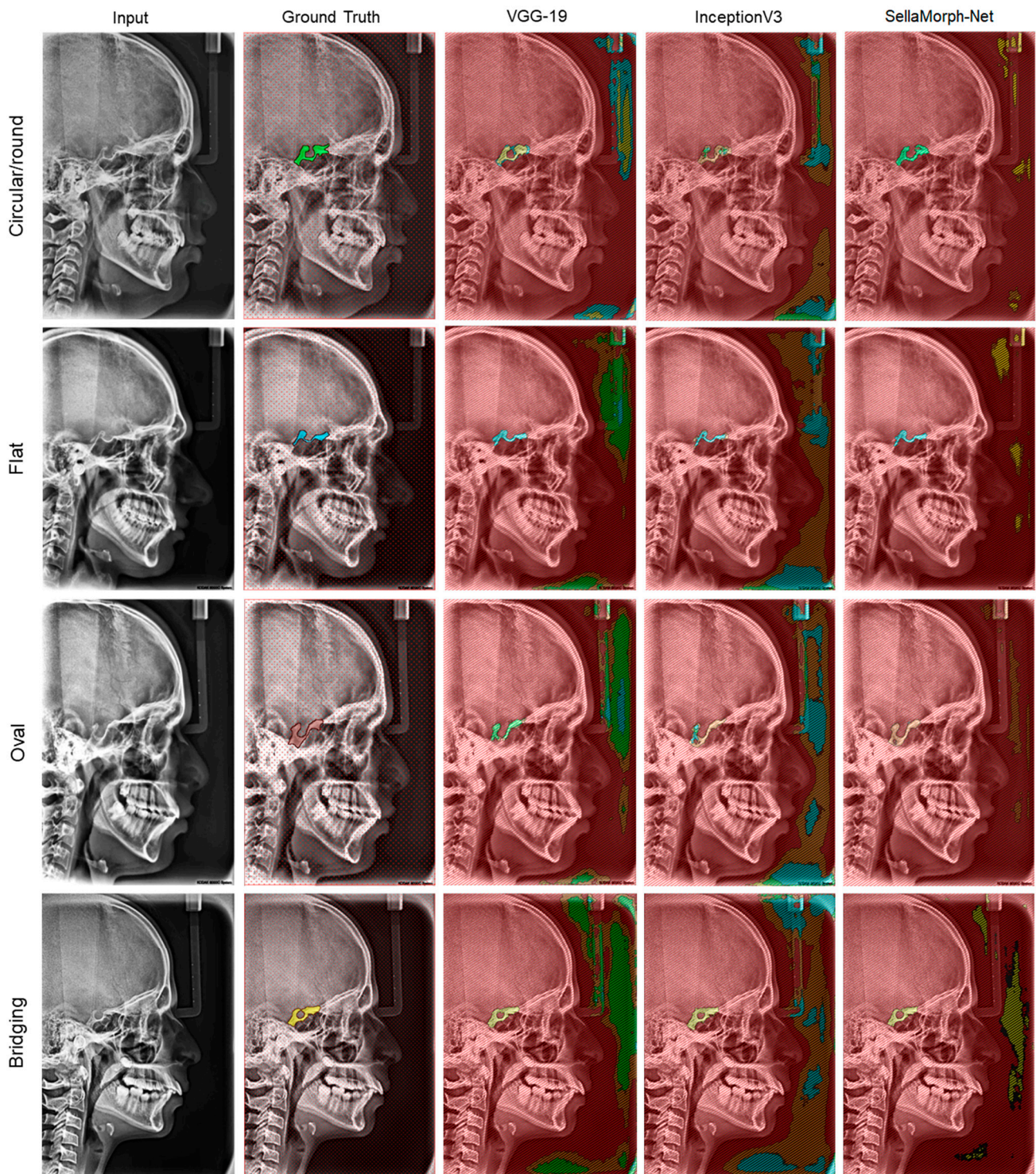


Figure 8. Qualitative results of VGG-19, InceptionV3, and proposed hybrid SellaMorph-Net on regular Sella Turcica subgroup dataset.

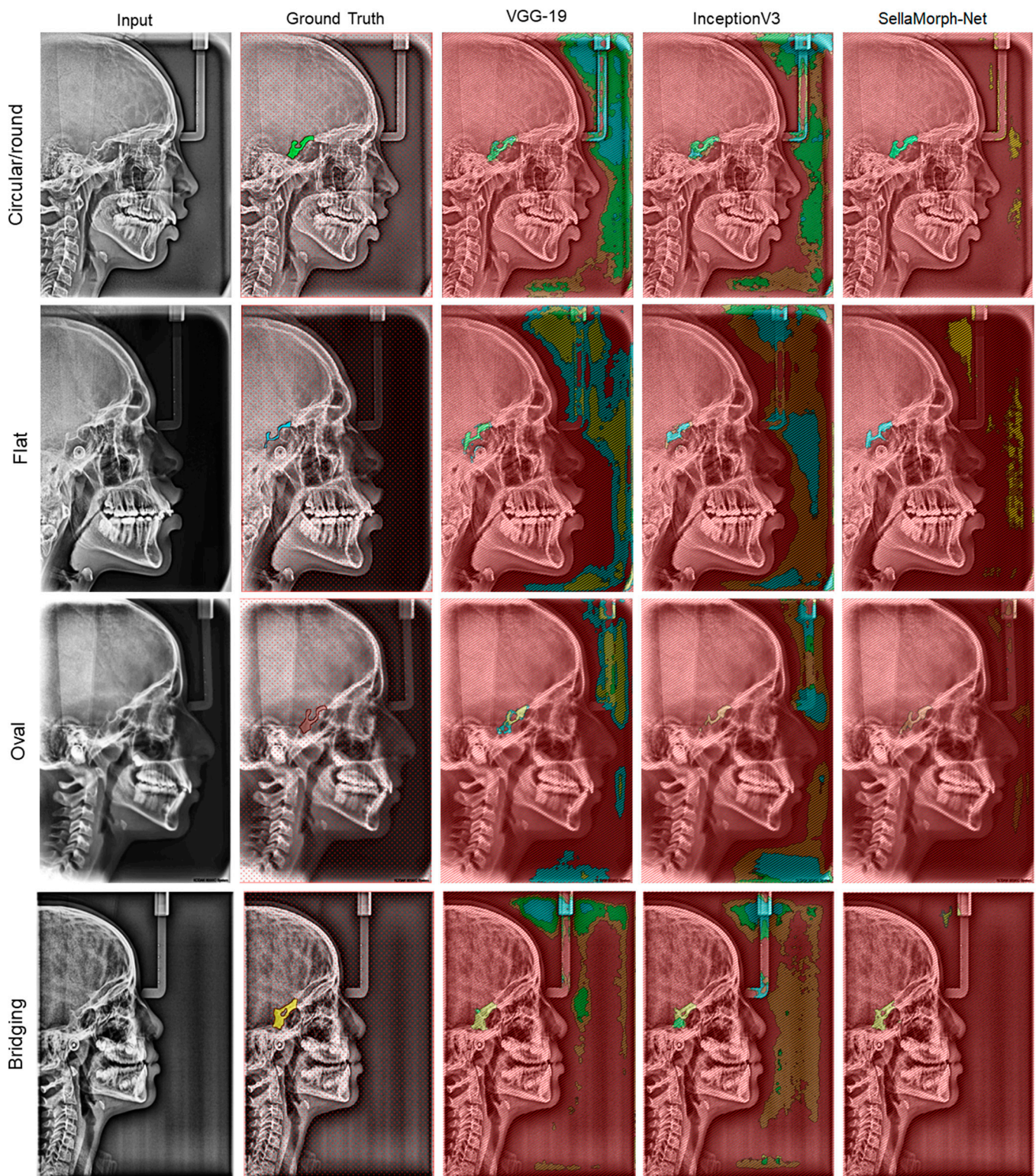


Figure 9. Qualitative results of VGG-19, InceptionV3, and proposed hybrid SellaMorph-Net on moderately altered Sella Turcica subgroup dataset.

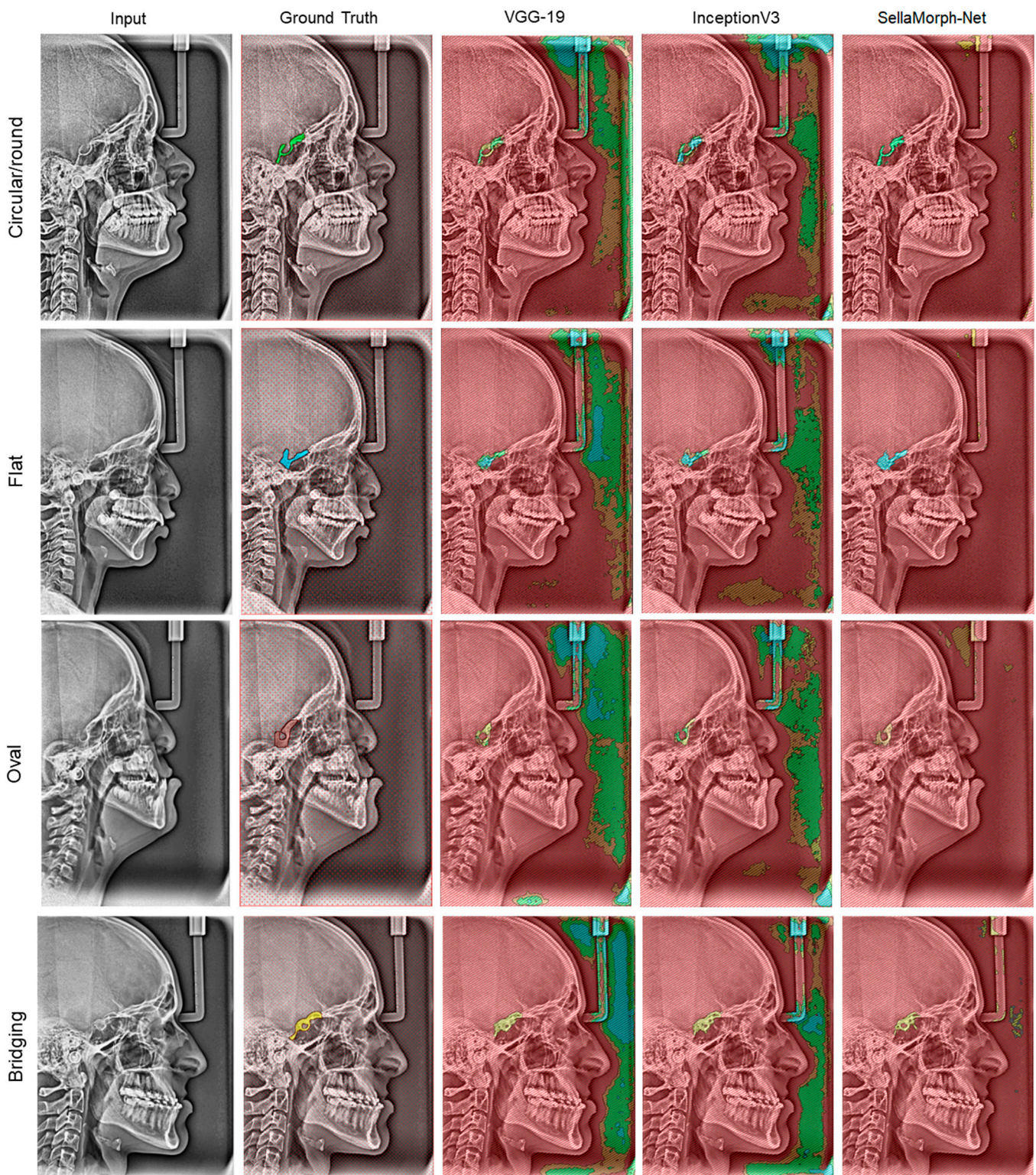


Figure 10. Qualitative results of VGG-19, InceptionV3, and proposed hybrid SellaMorph-Net on complex altered Sella Turcica subgroup dataset.

On the other hand, SellaMorph-Net predicted the ST types with high precision for all three dataset subgroups. Furthermore, the complex altered ST structure in Figure 10 was evaluated, and it was found that SellaMorph-Net predicted the edges of these complex altered categories with greater accuracy and smoothness than the other two models. Further,

the proposed model accuracy is validated using 400 cephalometric dataset images available publicly, and the results show the significant performance of the proposed hybrid model in the discussion section (Table 7).

Table 7. Comparative testing evaluation of state-of-the-art models and proposed approach on our dataset and publicly available dataset.

Method	Train Dataset	Test Dataset	IoU	Dice Coefficient
VGG19	Our dataset	Our dataset	0.5703	0.5925
		IEEE ISBI dataset	0.5753	0.5940
InceptionV3	Our dataset	Our dataset	0.5913	0.6265
		IEEE ISBI dataset	0.5897	0.6108
Proposed SellaMorph-Net	Our dataset	Our dataset	0.7314	0.7768
		IEEE ISBI dataset	0.7864	0.8313

4. Discussion

The extended usage of computer-aided diagnostics (CAD) in healthcare highlights the need for a reliable and consistent system for cephalometric radiographs that can be continuously evaluated and applicable to potential medical diagnoses. This research presents a new hybrid approach to autonomous Sella structure segmentation in cephalometric radiographs. The proposed method employs an encoder-decoder CNN architecture with attention-gating modules (AGMs) and recurrent residual convolutional (RrCL) blocks, replacing traditional convolutional blocks. Extensive experimental evaluations show that the proposed hybrid encoder-decoder CNN architecture surpasses existing methods for non-linear structural segmentation.

The figures presented in this study demonstrate the efficacy of the proposed method in segmenting complex non-linear structures of the ST in lat ceph radiographs. The segmentation was successful for various classes of ST, including bridging, circular, flat, and oval in severely anomalous structures, indicating its potential for clinical prognostics. Additionally, the segmentation of the Sella was conducted on a dataset that included regular and complex subgroups, allowing for the evaluation of dataset combinations. The study utilised statistical methods such as mean IoU (Jaccard Index) and Dice Coefficient (Equations (5)–(8)) to evaluate the performance of the proposed and pre-trained models. Furthermore, the accuracy and pixel-wise IoU of each architecture were compared with determine which performed better.

Tables 1–3 present the mean/mini-batch accuracy and time elapsed for three different subgroups of the SellaMorph-Net model: regular, moderately altered, and complex altered ST, respectively. The reported time elapsed refers to the inference time required for the model to predict each subset. Table 4 presents the training and validation IoU results for the state-of-the-art VGG-19 and InceptionV3 models and the proposed SellaMorph-Net model based on morphological classes of the Sella. The table also indicates the variation in training and validation accuracy using a pulse-minus (\pm) sign.

The results of our study on the SellaMorph-Net model are presented in Table 5. It includes the scores for global accuracy, pixel-wise intersection-over-union (IoU) for training and validation, mean IoU, and dice coefficient. The findings were also compared with the results obtained under 5-fold k-cross-dataset settings. Our proposed method outperformed the pre-trained models VGG-19 and InceptionV3 in terms of mean IoU and dice coefficient scores, as presented in Table 5. In order to evaluate the effectiveness of the proposed model against other advanced models, we utilised performance matrices and created confusion matrix graphs, which helped us determine the model's robustness, as shown in Table 6 and Figure 7. The qualitative results from Figures 8–10 also support the quantitative findings, especially when evaluating the model in cross-dataset settings.

The hybrid SellaMorph-Net model has several advantages over the InceptionV3 [26] and U-Net [32] models. It offers an end-to-end solution and has demonstrated remarkable

accuracy in analysing full lat ceph radiographs. Previous studies using InceptionV3 and U-Net models only considered the ST RoI in CBCT and X-ray images, respectively, and excluded the complex and crucial bridging structures, focusing only on regular circular, flat, and oval classes. Unlike these models, the proposed SellaMorph-Net model requires no significant image post-treatment, substantially reducing performance time. Additionally, the AGM and SE blocks are used in the network to enhance the input images' feature quality and the segmentation map to direct the network. The zero-shoot classification approach and the colour-coded layer added to the model effectively reduce outliers and misclassification.

Further, to evaluate the performance of the proposed hybrid model, we applied testing on a public dataset of the IEEE ISBI 2015 Challenge [53] for cephalometric landmark detections. The dataset consists of 400 cephalometric radiograph images, of which 150 are for training, 150 for validation, and 100 for testing.

Table 7 provides the comparative results of the proposed method's performance for our and publicly available datasets.

Although the proposed model has demonstrated outstanding performance in various settings, we have identified some limitations that must be considered. Firstly, the proposed system only achieves segmentation with colour-coded class differentiation. Secondly, this approach is computationally expensive due to the use of advanced feature enhancement techniques and the encoder brought in. To provide a more comprehensive CAD pipeline for lat ceph radiographs, we plan to incorporate additional downstream tasks such as ST classification and identification of dentofacial anomalies in future research. Additionally, we aim to increase the model's speed by using pruning techniques to reduce the number of model parameters. Table 8 presents an overview of studies that have reported Sella Turcica segmentation. Upon examining the literature, it becomes apparent that the ST is usually segmented as round, oval, or flat. However, the proposed model can accurately segment the Sella Turcica not only from regular ST but also from complex altered ST, including circular, flat, oval, and bridging classes.

Table 8. Detailed overview of studies reporting Sella Turcica segmentation and classification.

Methods	Year	Method Description	Datasets	Performance
Shakya et al. [24]	2022	Manual annotations performed by dental experts; U-Net architecture; backbone: VGG19, ResNet34, InceptionV3, and ResNeXt50	Consider randomly selected full lat ceph radiographs	IoU: 0.7651, 0.7241; dice coefficient: 0.7794, 0.7487
Duman et al. [26]	2022	Polygonal box annotation method performed by radiologists; U-Net architecture; Google InceptionV3	Consider Sella Turcica roi in CBCT dataset	Sensitivity: 1.0; Precision: 1.0; and F-measure values: 1.0
Feng et al. [32]	2023	Labelme software used for annotation; U-Net architecture	Consider Sella Turcica roi in X-ray dataset	Dice coefficients: 92.84%
SellaMorph-Net (proposed model)	-	Annotation performed by radiologists, dental experts, and researchers; a hybrid "AGM + RrCL + SE + Zero-shot classifier" pipeline developed from scratch	Full lat ceph radiographic images are considered; subdivide dataset into four groups based on the morphological characteristics of Sella; dataset group divided into subgroups based on the structure alteration complexity	Data subgroups mini-batch accuracy: 99.57%, 99.51%, and 99.49%; class-based IoU score: 0.6539 ± 0.03 , 0.6370 ± 0.01 , 0.6517 ± 0.01 , and 0.5663 ± 0.02 ; global accuracy: 97.570%; mean IoU: 0.7129; dice coefficient: 0.7324

5. Conclusions

The study presents a hybrid framework that leverages deep learning to effectively identify various Sella Turcica (ST) morphologies in full lateral cephalometric (lat ceph) radiographic images, including bridging, circular, oval, and flat. The framework employs an encoder-decoder network to iteratively refine the network output by merging it with the input data and passing it through it to Attention Gating Modules (AGM) and Squeeze-and-Excitation (SE) blocks. This process allows for the identification of anatomical characteristics within the radiographic images. To enhance the feature maps and prioritise regions of interest, the proposed framework utilised RrCL and AGMs instead of relying solely on conventional convolutional layers.

By accurately extracting data from the encoder-decoder network, the proposed framework facilitates precise segmentation of anatomical characteristics. This assists healthcare professionals in predicting various dentofacial anomalies associated with ST structures.

Moving forwards, enhancement to the anatomical feature segmentation pipeline could involve using synthetically generated samples, produced via generative adversarial networks, to enhance performance. Further improvements may include downstream tasks such as ST structure classification and the identification of dentofacial anomalies related to the defined ST structures. This approach will contribute to a more comprehensive computer-aided diagnosis (CAD) for cephalometric radiographs. Moreover, to bridge the gap between high performance and interpretability, our future work will consider incorporating explainable AI techniques. This approach can help elucidate the decision-making process of our deep learning model, making its predictions more transparent and interpretable to healthcare professionals. Additionally, a more advanced segmentation and classification method will be investigated to achieve more accuracy and reliability in anatomical structure segmentation and dentofacial anomaly classification.

Author Contributions: Conceptualization, K.S.S., M.J., P.K., A.A. and A.L.; methodology, K.S.S.; validation, M.J., V.K. and A.L.; formal analysis, V.K. and J.P.; investigation, K.S.S., M.J. and A.L.; resources, M.J., V.K. and A.L.; data curation, K.S.S., M.J. and V.K.; writing—original draft preparation, K.S.S. and P.K.; writing—review and editing, K.S.S., P.K., J.P., A.A. and A.L.; supervision, M.J., J.P., V.K., A.A. and A.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data sharing not applicable. The data are not publicly available due to privacy regulation.

Acknowledgments: The authors are grateful to the Oral Health Sciences Centre, Postgraduate Institute of Medical Education and Research (PGIMER), Chandigarh, for providing cephalometric data under the ethical clearance number IEC-09/2021-2119, and the faculty of the School of Science, RMIT University, Melbourne, for their valuable assistance.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Alkofide, E.A. The shape and size of the sella turcica in skeletal Class I, Class II, and Class III Saudi subjects. *Eur. J. Orthod.* **2007**, *29*, 457–463. [[CrossRef](#)]
2. Sinha, S.; Shetty, A.; Nayak, K. The morphology of Sella Turcica in individuals with different skeletal malocclusions—A cephalometric study. *Transl. Res. Anat.* **2020**, *18*, 100054. [[CrossRef](#)]
3. Ghadimi, M.H.; Amini, F.; Hamed, S.; Rakhshan, V. Associations among sella turcica bridging, atlas arcuate foramen (ponticulus posticus) development, atlas posterior arch deficiency, and the occurrence of palatally displaced canine impaction. *Am. J. Orthod. Dentofacial. Orthop.* **2017**, *151*, 513–520. [[CrossRef](#)]
4. Meyer-Marcotty, P.; Reuther, T.; Stellzig-Eisenhauer, A. Bridging of the sella turcica in skeletal Class III subjects. *Eur. J. Orthod.* **2010**, *32*, 148–153. [[CrossRef](#)]

5. Back, S.A. Perinatal white matter injury: The changing spectrum of pathology and emerging insights into pathogenetic mechanisms. *Ment. Retard. Dev. Disabil. Res. Rev.* **2006**, *12*, 129–140. [[CrossRef](#)]
6. Teal, J. Radiology of the adult sella turcica. *Bull. LA Neurol. Soc.* **1977**, *42*, 111–174.
7. Amar, A.P.; Weiss, M.H. Pituitary anatomy and physiology. *Neurosurg. Clin. N. Am.* **2003**, *14*, 11–23. [[CrossRef](#)]
8. Leeds, N.E.; Kieffer, S.A. Evolution of diagnostic neuroradiology from 1904 to 1999. *Radiology* **2000**, *217*, 309–318. [[CrossRef](#)]
9. Senior, B.A.; Ebert, C.S.; Bednarski, K.K.; Bassim, M.K.; Younes, M.; Sigounas, D.; Ewend, M.G. Minimally invasive pituitary surgery. *Laryngoscope* **2008**, *118*, 1842–1855. [[CrossRef](#)]
10. Honkanen, R.A.; Nishimura, D.Y.; Swiderski, R.E.; Bennett, S.R.; Hong, S.; Kwon, Y.H.; Stone, E.M.; Sheffield, V.C.; Alward, W.L. A family with Axenfeld–Rieger syndrome and Peters Anomaly caused by a point mutation (Phe112Ser) in the FOXC1 gene. *Am. J. Ophthalmol.* **2003**, *135*, 368–375. [[CrossRef](#)]
11. Becktor, J.P.; Einersen, S.; Kjær, I. A sella turcica bridge in subjects with severe craniofacial deviations. *Eur. J. Orthod.* **2000**, *22*, 69–74. [[CrossRef](#)]
12. Townsend, G.; Harris, E.F.; Lesot, H.; Clauss, F.; Brook, A. Morphogenetic fields within the human dentition: A new, clinically relevant synthesis of an old concept. *Arch. Oral Biol.* **2009**, *54*, S34–S44. [[CrossRef](#)]
13. Souzeau, E.; Siggs, O.M.; Zhou, T.; Galanopoulos, A.; Hodson, T.; Taranath, D.; Mills, R.A.; Landers, J.; Pater, J.; Smith, J.E. Glaucoma spectrum and age-related prevalence of individuals with FOXC1 and PITX2 variants. *Eur. J. Hum. Genet.* **2017**, *25*, 839–847. [[CrossRef](#)]
14. Axelsson, S.; Storhaug, K.; Kjær, I. Post-natal size and morphology of the sella turcica. Longitudinal cephalometric standards for Norwegians between 6 and 21 years of age. *Eur. J. Orthod.* **2004**, *26*, 597–604. [[CrossRef](#)]
15. Andredaki, M.; Koumantanou, A.; Dorotheou, D.; Halazonetis, D. A cephalometric morphometric study of the sella turcica. *Eur. J. Orthod.* **2007**, *29*, 449–456. [[CrossRef](#)]
16. Sotos, J.F.; Dodge, P.R.; Muirhead, D.; Crawford, J.D.; Talbot, N.B. Cerebral gigantism in childhood: A syndrome of excessively rapid growth with acromegalic features and a nonprogressive neurologic disorder. *N. Engl. J. Med.* **1964**, *271*, 109–116. [[CrossRef](#)]
17. Bambha, J.K. Longitudinal cephalometric roentgenographic study of face and cranium in relation to body height. *J. Am. Dent. Assoc.* **1961**, *63*, 776–799. [[CrossRef](#)]
18. Haas, L. The size of the sella turcica by age and sex. *AJR Am. J. Roentgenol.* **1954**, *72*, 754–761.
19. Chilton, L.; Dorst, J.; Garn, S. The volume of the sella turcica in children: New standards. *AJR Am. J. Roentgenol.* **1983**, *140*, 797–801. [[CrossRef](#)]
20. Di Chiro, G.; Nelson, K. The volume of the sella turcica. *AJR Am. J. Roentgenol.* **1962**, *87*, 989–1008.
21. McLachlan, M.; Williams, E.; Fortt, R.; Doyle, F. Estimation of pituitary gland dimensions from radiographs of the sella turcica. *Br. J. Radiol.* **1968**, *41*, 323–330. [[CrossRef](#)]
22. Underwood, L.E.; Radcliffe, W.B.; Guinto, F.C. New standards for the assessment of sella turcica volume in children. *Radiology* **1976**, *119*, 651–654. [[CrossRef](#)] [[PubMed](#)]
23. Silverman, F.N. Roentgen standards for size of the pituitary fossa from infancy through adolescence. *AJR Am. J. Roentgenol.* **1957**, *78*, 451–460.
24. Shakya, K.S.; Laddi, A.; Jaiswal, M. Automated methods for sella turcica segmentation on cephalometric radiographic data using deep learning (CNN) techniques. *Oral Radiol.* **2023**, *39*, 248–265. [[CrossRef](#)]
25. Shakya, K.S.; Priti, K.; Jaiswal, M.; Laddi, A. Segmentation of Sella Turcica in X-ray Image based on U-Net Architecture. *Procedia Comput. Sci.* **2023**, *218*, 828–835. [[CrossRef](#)]
26. Duman, Ş.B.; Syed, A.Z.; Celik Ozen, D.; Bayrakdar, İ.Ş.; Salehi, H.S.; Abdelkarim, A.; Celik, Ö.; Eser, G.; Altun, O.; Orhan, K. Convolutional Neural Network Performance for Sella Turcica Segmentation and Classification Using CBCT Images. *Diagnostics* **2022**, *12*, 2244. [[CrossRef](#)]
27. Hosseini, H.; Hesar, F.; Marvasti, F. Real-time impulse noise suppression from images using an efficient weighted-average filtering. *IEEE Signal Process. Lett.* **2014**, *22*, 1050–1054. [[CrossRef](#)]
28. Lee, E.; Kim, S.; Kang, W.; Seo, D.; Paik, J. Contrast enhancement using dominant brightness level analysis and adaptive intensity transformation for remote sensing images. *IEEE Geosci. Remote Sens. Lett.* **2012**, *10*, 62–66. [[CrossRef](#)]
29. Poobathy, D.; Chezian, R.M. Edge detection operators: Peak signal to noise ratio based comparison. *Int. J. Image Graph. Signal Process.* **2014**, *10*, 55–61. [[CrossRef](#)]
30. Suzuki, K.; Horiba, I.; Sugie, N. Neural edge enhancer for supervised edge enhancement from noisy images. *IEEE Trans. Pattern Anal. Mach. Intell.* **2003**, *25*, 1582–1596. [[CrossRef](#)]
31. Kim, S.H.; Allebach, J.P. Optimal unsharp mask for image sharpening and noise removal. *J. Electron. Imaging* **2005**, *14*, 023005.
32. Feng, Q.; Liu, S.; Peng, J.-X.; Yan, T.; Zhu, H.; Zheng, Z.-J.; Feng, H.-C. Deep learning-based automatic sella turcica segmentation and morphology measurement in X-ray images. *BMC Med. Imaging* **2023**, *23*, 41. [[CrossRef](#)] [[PubMed](#)]
33. Kök, H.; Acilar, A.M.; İzgi, M.S. Usage and comparison of artificial intelligence algorithms for determination of growth and development by cervical vertebrae stages in orthodontics. *Prog. Orthod.* **2019**, *20*, 41. [[CrossRef](#)] [[PubMed](#)]
34. Palanivel, J.; Davis, D.; Srinivasan, D.; Nc, S.C.; Kalidass, P.; Kishore, S.; Suvetha, S. Artificial Intelligence—Creating the Future in Orthodontics—A Review. *J. Evol. Med. Dent. Sci.* **2021**, *10*, 2108–2114. [[CrossRef](#)]
35. Asiri, S.N.; Tadlock, L.P.; Schneiderman, E.; Buschang, P.H. Applications of artificial intelligence and machine learning in orthodontics. *APOS Trends Orthod.* **2020**, *10*, 17–24. [[CrossRef](#)]

36. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.
37. Pang, G.; Shen, C.; Cao, L.; Hengel, A.V.D. Deep learning for anomaly detection: A review. *ACM Comput. Surv.* **2021**, *54*, 1–38. [[CrossRef](#)]
38. Meedeniya, D.; Kumarasinghe, H.; Kolonne, S.; Fernando, C.; De la Torre Díez, I.; Marques, G. Chest X-ray analysis empowered with deep learning: A systematic review. *Appl. Soft Comput.* **2022**, *126*, 109319. [[CrossRef](#)]
39. Fernando, C.; Kolonne, S.; Kumarasinghe, H.; Meedeniya, D. Chest radiographs classification using multi-model deep learning: A comparative study. In Proceedings of the 2022 2nd International Conference on Advanced Research in Computing (ICARC), Sabaragamuwa, Sri Lanka, 23–24 February 2022; pp. 165–170.
40. Mateen, M.; Wen, J.; Song, S.; Huang, Z. Fundus image classification using VGG-19 architecture with PCA and SVD. *Symmetry* **2018**, *11*, 1. [[CrossRef](#)]
41. Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. Dropout: A simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* **2014**, *15*, 1929–1958.
42. Alzubaidi, L.; Zhang, J.; Humaidi, A.J.; Al-Dujaili, A.; Duan, Y.; Al-Shamma, O.; Santamaría, J.; Fadhel, M.A.; Al-Amidie, M.; Farhan, L. Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *J. Big Data* **2021**, *8*, 53. [[CrossRef](#)]
43. Arora, R.; Basu, A.; Mianjy, P.; Mukherjee, A. Understanding deep neural networks with rectified linear units. *arXiv* **2016**, arXiv:1611.01491.
44. Ullah, I.; Ali, F.; Shah, B.; El-Sappagh, S.; Abuhmed, T.; Park, S.H. A deep learning based dual encoder–decoder framework for anatomical structure segmentation in chest X-ray images. *Sci. Rep.* **2023**, *13*, 791. [[CrossRef](#)]
45. Hu, J.; Shen, L.; Sun, G. Squeeze-and-excitation networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2018, Salt Lake City, UT, USA, 18–22 June 2018; pp. 7132–7141.
46. Alom, M.Z.; Hasan, M.; Yakopcic, C.; Taha, T.M.; Asari, V.K. Improved inception-residual convolutional neural network for object recognition. *Neural Comput. Appl.* **2020**, *32*, 279–293. [[CrossRef](#)]
47. Ioffe, S.; Szegedy, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In Proceedings of the International Conference on Machine Learning, Lille, France, 7–9 July 2015; pp. 448–456.
48. Schlemper, J.; Oktay, O.; Schaap, M.; Heinrich, M.; Kainz, B.; Glocker, B.; Rueckert, D. Attention gated networks: Learning to leverage salient regions in medical images. *Med. Image Anal.* **2019**, *53*, 197–207. [[CrossRef](#)] [[PubMed](#)]
49. Xian, Y.; Lampert, C.H.; Schiele, B.; Akata, Z. Zero-shot learning—A comprehensive evaluation of the good, the bad and the ugly. *IEEE Trans. Pattern Anal. Mach. Intell.* **2018**, *41*, 2251–2265. [[CrossRef](#)] [[PubMed](#)]
50. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.
51. Jaccard, P. The distribution of the flora in the alpine zone. *New Phytol.* **1912**, *11*, 37–50. [[CrossRef](#)]
52. Sorensen, T.A. A method of establishing groups of equal amplitude in plant sociology based on similarity of species content and its application to analyses of the vegetation on Danish commons. *Biol. Skar.* **1948**, *5*, 1–34.
53. Wang, C.-W.; Huang, C.-T.; Lee, J.-H.; Li, C.-H.; Chang, S.-W.; Siao, M.-J.; Lai, T.-M.; Ibragimov, B.; Vrtovec, T.; Ronneberger, O. A benchmark for comparison of dental radiography analysis algorithms. *Med. Image Anal.* **2016**, *31*, 63–76. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.