



An Automatic Framework for Nasal Esthetic Assessment by ResNet Convolutional Neural Network

Maryam Ashoori¹ · Reza A. Zoroofi¹ · Mohammad Sadeghi²

Received: 20 August 2023 / Revised: 30 November 2023 / Accepted: 4 December 2023 / Published online: 29 January 2024
© The Author(s) 2024

Abstract

Nasal base aesthetics is an interesting and challenging issue that attracts the attention of researchers in recent years. With that insight, in this study, we propose a novel automatic framework (AF) for evaluating the nasal base which can be useful to improve the symmetry in rhinoplasty and reconstruction. The introduced AF includes a hybrid model for nasal base landmarks recognition and a combined model for predicting nasal base symmetry. The proposed state-of-the-art nasal base landmark detection model is trained on the nasal base images for comprehensive qualitative and quantitative assessments. Then, the deep convolutional neural networks (CNN) and multi-layer perceptron neural network (MLP) models are integrated by concatenating their last hidden layer to evaluate the nasal base symmetry based on geometry features and tiled images of the nasal base. This study explores the concept of data augmentation by applying the methods motivated via commonly used image augmentation techniques. According to the experimental findings, the results of the AF are closely related to the otolaryngologists' ratings and are useful for preoperative planning, intraoperative decision-making, and postoperative assessment. Furthermore, the visualization indicates that the proposed AF is capable of predicting the nasal base symmetry and capturing asymmetry areas to facilitate semantic predictions. The codes are accessible at <https://github.com/AshooriMaryam/Nasal-Aesthetic-Assessment-Deep-learning>.

Keywords Nasal base · Symmetry · Landmarks · Automatic framework · Hybrid model · Combined model

Introduction

The nose is a vital element of a person's esthetic appearance of the face that affects the overall appearance, esthetics, and attractiveness of the facial [1, 2]. The appearance and balance of the nose as important features of rhinoplasty deeply affected by the nasal base [3]. The nasal base so-called alar-columellar complex provides an ideal starting point to develop an analytic approach to quantitative analysis of nasal

shape [4] that can be useful to distinguish the face characteristics (for example [5]). This area of nose is a common source of patient dissatisfaction and neglecting of it would lead to some revision rhinoplasty [6]. Some of the complications in rhinoplasty that significantly affect facial attractiveness [2] are asymmetry and deformation appearance [7]. Symmetry can be considered a major factor in nasal base esthetics [8] especially, by using selfies in the digital age, the rate of requests for a more symmetrical nose is increased [9, 10].

Recent developments in artificial intelligence, machine learning, and deep learning techniques have opened new avenues for efficient knowledge discovery from healthcare data [11] which can be trained to carry out tasks that are either challenging or time-consuming for surgeons [12, 13]. Deep learning algorithms are a subset of machine learning algorithms that have led to the construction of several novel deep neural network architectures that are able to illuminate patterns and features that are not always visible to the human eye [11].

Several researchers studied rhinoplasty based on computer technology such as introducing simulation or prediction models for the nasal shape that esthetically matches the patient's face

✉ Maryam Ashoori
Maryam.ashoori@ut.ac.ir; Maryam.ashoori@gmail.com

Reza A. Zoroofi
zoroofi@ut.ac.ir

Mohammad Sadeghi
sadeghih@tums.ac.ir

¹ Control and Intelligent Processing Center of Excellence, School of Electrical and Computer Engineering, College of Engineering, University of Tehran, Tehran, Iran

² Tehran University of Medical Sciences, Imam Khomeini Hospital Complex, Tehran, Iran

[14, 15], constructing three-dimensional (3D) facial images from two-dimensional images, and producing 3D simulation models to revolutionize the practice of functional and esthetic rhinoplasty [16, 17]; the others are listed in [18]. Also, A parametric model (PM) is used to describe objectively nasal base shape [4], and a classification system is created by evaluating and comparing the PM with the categorization by surgeons [19]. Machine learning was used to simulate rhinoplasty results according to the criteria of the doctors surveyed [13]. Finally, deep learning was used to predict rhinoplasty status accurately [20] and patient's age before and after rhinoplasty [21] and to find which geometric facial features that affect attractiveness in order to considered them within the rhinoplasty procedures [22].

In the latest years, deep learning-based object detection algorithms have played an important role in reducing human efforts in the processing of modern approaches. Object detection algorithms based on deep learning such as region-based convolutional neural networks (R-CNN) [23], Fast R-CNN [24], and Faster R-CNN [25] are characterized by the bounding boxes and categories probabilities for each object. Faster RCNN as a deep object detection algorithm utilizes region proposal networks (RPNs) to generate image regions that provide better performance and more speed for object detection. Object detection to recognize landmarks [26] is one of the most state-of-the-art methods that solves the problem of facial landmark detection [27, 28]. In the most recent research works of the latest years [29], two-stage object detector methods have excellent performance in object recognition and localization accuracy [30]. The advantages of the R-CNN family as a two-stage method, rather than the one-stage detectors are as follows: (1) utilizing the sampling heuristics to deal with class imbalance; (2) regressing the object box parameters by two-step cascade; (3) describing the objects according to two-stage features [31]. However, the above discussed previous works related to the nasal base have some limitations that addressed as follows:

1. In the previous works of the researchers, the symmetry of the nasal base was considered only for evaluating the results of rhinoplasty on cleft palate patients [32–36].
2. In these works, extracting the nasal base landmarks has been taken by manual methods using either direct or indirect anthropometry [32, 35–37]. The important key is the lack of an accurate and rapid automated method to detect the landmarks of the nasal base.
3. In other related research, the symmetry of the nasal base has been studied using some quantitative methods to survey the geometry features that are limited to the statistical methods [4, 8, 32, 34, 37–39]. Barnes and et. Al [4] utilized the lateral deviation (symmetry) of the nasal base as a parameter of a polar function without calculating the value of the symmetry. [8] Presented a clinical

technique to improve the symmetry of the columella and nostrils. Then used the χ^2 test to compare the results of pre and post-operative based on patients' opinion [34]. Applied descriptive statistics to compare the nasal symmetry of infant with unilateral cleft lip with or without cleft palate between time points from frontal, lateral, and submental views [37]. Utilized Student's *t* test to analyze narsi symmetry of the patients were treated by using the Hotz plate. [38] used analysis of variance and equality of two proportions tests to compare the symmetry after fat grafting in paranasal and midface groups based on manual extracted measurements and evaluator rating. In [32], Pietruski et al. conducted a validation study to develop a computer system as a tool for objective anthropometric analysis of the nasolabial region. In addition, the practical application of the system was further confirmed through a comparative objective analysis of nasolabial morphology and symmetry in the both healthy individuals and the cleft subjects [39]. It is important to note that in the both of the last works the specified number of landmarks was set by the user. However, we can't find deep learning algorithms for analyzing the nasal base symmetry. Moreover, it is crucial to highlight that none of the abovementioned studies included a scoring system for evaluating the symmetry of the nasal base.

Therefore, until now there is no unique AF which able to evaluate the symmetry of the nasal base based on deep learning algorithms. So, it is important to propose an exact AF to assess the symmetry value of nasal base before and after rhinoplasty. Also, this paper pays attention to adopting the Faster R-CNN technique to detect the nasal base inside the image and then recognize the nasal base landmarks (Appendix 1). The main motivation of this research is to suggest an AF that is compatible with human opinion to reduce the role of human factors to evaluate the symmetry of the nasal base. The remainder of the paper is as follows: The “[Materials and Methods](#)” section illustrates the materials and methods; the “[Results and Discussion](#)” section describes the experimental results. Finally, the paper is concluded in the “[Conclusions](#)” section.

Materials and Methods

Figure 1 shows the block diagram of the proposed method which involves three main steps: data preparation, processing, and comparison. In the initial step, the septorhinoplasty dataset (SRD) consisting of the preoperative and one-year postoperative photographs of 438 primary rhinoplasty patients (comprising 740 women and 136 men, totaling 876 image data) was randomly divided into a training dataset (80%, 720 images) and a testing dataset (20%, 156 images).

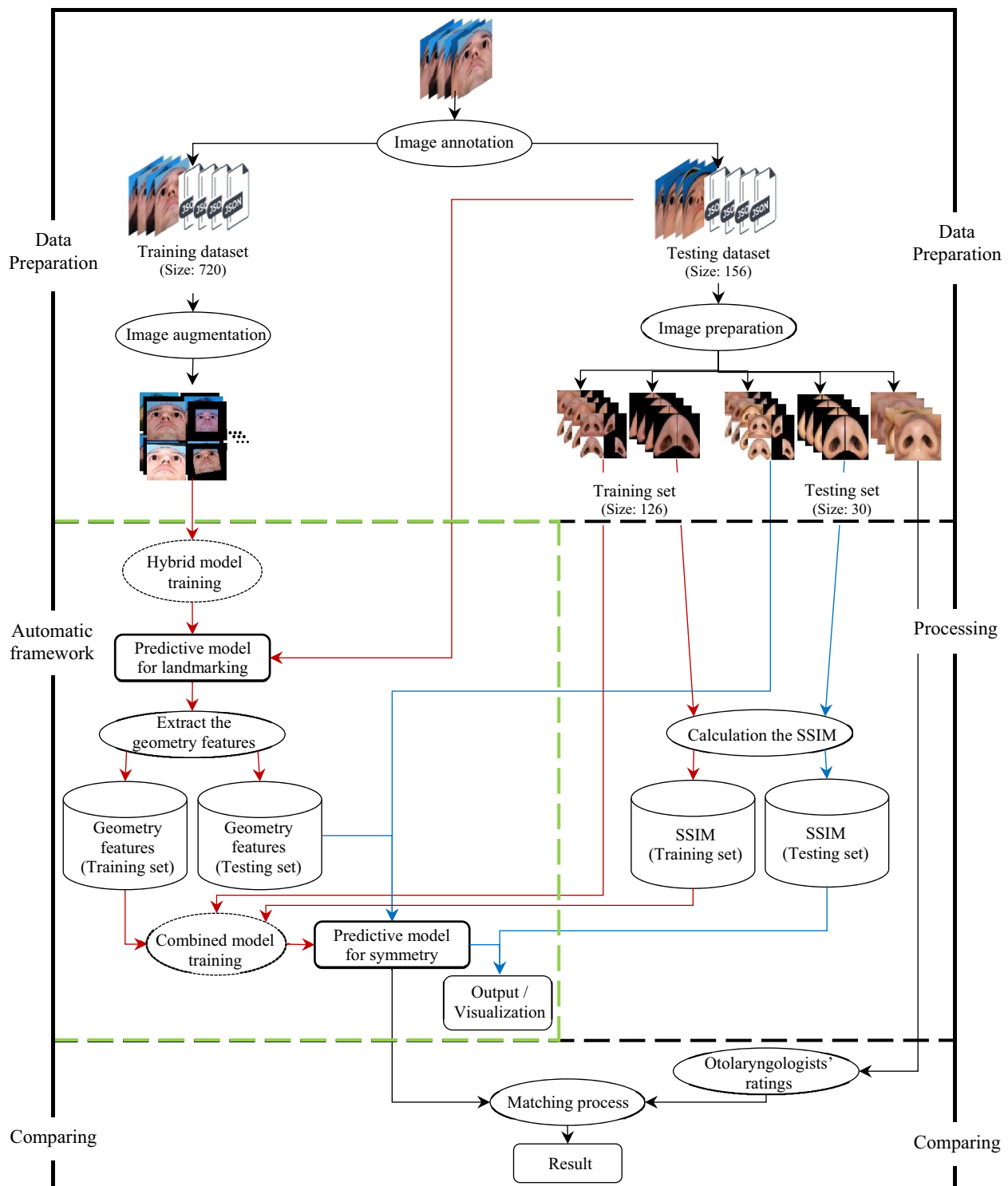


Fig. 1 Detailed block diagram of the proposed method to predict the symmetry of the nasal base. The red, blue, and black arrows correspond to the training, testing, and other phases, respectively. The enclosed area between the green dashed lines shows the AF

Image annotation was applied to all images in the SRD. Subsequently, image augmentation (25 augmented images per input image) was performed on each image in the training dataset, and the nasal base region was cropped based

on the columellar axis. Additionally, tiled image versions of the testing dataset were generated. In the second step, the augmented annotated training dataset was used to train the hybrid model for nasal base landmarks recognition. The

original images of the testing dataset (156 images) were fed to the predictive model for landmarking during the testing phase to extract the geometry features. The tiled images and geometry features of the training set (126 images) were utilized as input, and structural similarity index measure (SSIM) [40] was used as the output for the combined model during the training phase. The predictive model for symmetry utilized the testing set (30 images) in conjunction with geometry features to predict the symmetry value of the nasal base. Finally, in the last step, the matching process was investigated to compare the results of the predictive model for symmetry and otolaryngologists' ratings.

Dataset Description and Preparation

The color profile photographs of 438 primary rhinoplasty patients (370 women and 68 men) are selected from the database of all patients who had been referred to the otolaryngology office of the third author from 2010 to 2019, where almost all of these patients were elective for cosmetic procedures. While their initial motivation for seeking treatment was cosmetic enhancement, in a substantial portion of subjects (potentially around half), concurrent medical treatment was also provided alongside the cosmetic procedure. The mean age and standard deviation of the patients at the time of surgery was 30.17 ± 5.01 years (range 13–75 years). The skin of the subjects included the types II, III, and IV of skin tone categories in the Fitzpatrick scale [41].

Preoperative and 1-year postoperative photographs were taken from frontal, lateral, and basal views with a Canon 60 D camera in high resolution according to the standard guidelines [42] for clinical photography. The resolution of the images was 72 dpi, and photography was conducted without the use of flash due to enough ambient lighting. The image quality was not compromised by refraining from applying any compression during the image acquisition. As the surgery was septorhinoplasty and there was sensitivity, structured illumination was used to capture images of the patient. The photographs were taken in specialized studios approved by the surgeon and were not color-calibrated. Additionally, if the imaging conditions are not met, the surgeon repeats the image. All patient's photographs were included in the study after giving written informed consent. Full ethical approval was granted by the University of Tehran research ethics committee. The preparation process of the SRD for training the deep learning models includes collecting data, data cleaning, and format conversion, bounding box annotation and labeling, dataset partition, and data augmentation.

Furthermore, in order to assess the robustness of the hybrid model, we construct a new multi-ethnicity rhinoplasty dataset (MERD) [43, 44] containing paired facial images of 100 rhinoplasty patients of different races and nations extracted

from publicly available websites. The MERD includes categories of Ethnic Rhinoplasty, Middle Eastern Rhinoplasty, Latino Rhinoplasty, Asian rhinoplasty, and African American Rhinoplasty. The demographic composition of the MERD consists of 74 women and 26 men (with ages approximately ranging from 18 to 70+ years). The selection criteria for images from these websites included (1) relevance to the scope of this research, (2) a minimum image resolution equal to that of the SRD, and (3) the availability of six views of the face (frontal, lateral (right and left), three-quarter oblique (right and left), and basal). We gathered images depicting frontal, lateral, three-quarter oblique, and basal views of rhinoplasty subjects, in accordance with the available categories of the rhinoplasty images on the websites. Subsequently, we manually annotated them with 18 nasal base landmarks. In terms of size and resolution, the images of SRD and MERD have been resized to equal size, but the resolutions of MERD images were either 72 or 96 dpi.

Proposed Automatic Framework

The Proposed Hybrid Model

In this subsection, we designed a novel hybrid model combining different architectures including a mixture of Faster R-CNN and CNN models to detect nasal base and to predict nasal base landmarks coordinates automatically. Firstly, the augmented data which was obtained from the training dataset (80% for training: 508 women and 72 men and 20% for validation: 114 women and 26 men) was used to train the hybrid model. Labelme [45, 46] is used to annotate objects, and json file is created for each image to comprise nasal base landmark annotations. There are several landmarks of the nasal base in surgical books (Appendix 1) that produce geometry features including nine scale values $\mathcal{F} = [f_{g-1}, f_{g-2}, f_{g-3}, f_{g-4}, f_{g-5}, f_{g-6}, f_{g-7}, f_{g-8}, f_{g-9}]$ that are calculated via Eqs. (1–2) and are shown in Fig. 2.

$$f_{g-1} = a_{(g-1)L} / a_{(g-1)R} \quad (1)$$

$$f_{g-i} = d_{(g-i)L} / d_{(g-i)R} \quad i = 2, 3, \dots, 8, 9 \quad (2)$$

where f_{g-1} represents the angular ratio of the angulation of the long axis of the left nostril ($a_{(g-1)L}$) and the angulation of the long axis of the right nostril ($a_{(g-1)R}$) and f_{g-i} represents the i -th geometry ratio feature, $d_{(g-i)L}, d_{(g-i)R}$ are the i -th distances of left and right sides of nasal base, left/right midalar widths ($i = 2$), horizontal distance between the left/right subalare and the facial midline ($i = 3$), horizontal distance between the left/right alare and the columellar axis ($i = 4$), left/right midcolumellar apex width ($i = 5$), left/right midcolumellar base width ($i = 6$), left/right nostril width ($i = 7$), left/right nostril height ($i = 8$), and left/right thickness of ala ($i = 9$), respectively.

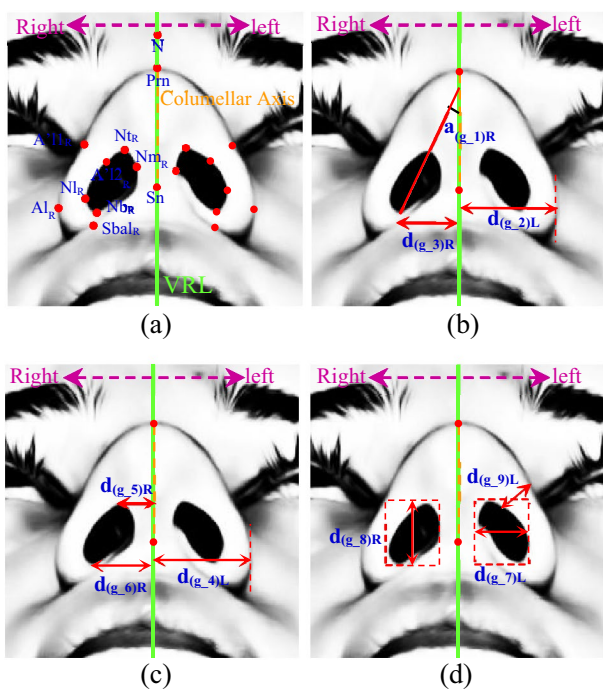


Fig. 2 Landmarks (a) and geometry features (b)–(d) of the nasal base. The vertical reference line (VRL) corresponding to the facial midline was set as a perpendicular line crossing Nasion, Subnasale, and Gnathion:

We use Faster R-CNN with VGG16 as the backbone network for the object detection task to generate a bounding box around the nasal base inside the input image (detection module). Then, the input is cropped based on this bounding box to have a cropped nasal base. After that, we set again the landmark annotations based on the corresponding bounding region coordinates to build a predictive model that takes a set of cropped nasal base boxes and object annotations as input. Then a customized pre-trained ResNet152V2 model is designed and trained to predict the nasal base landmarks coordinates (prediction module) (Fig. 3). In this method, the landmark predictive model is trained on the bounding box region of the input image, and it can predict the coordinates of landmarks more accurately rather than using the whole image for prediction.

Implementation Details In the detection module, the ground-truth box of the nasal base, which is obtained according to the pronasal and alars coordinates of object annotations, is used to determine the bounding box region in the images. It is trained using Adam optimizer [47] with a learning rate of 1e-5. We selected the small initial learning rate since the pre-trained VGG16 model is not proper for intense changes. In the prediction module, we utilized the Adam optimizer and “LearningRateScheduler” callback with initial alpha equal to 1e-3 to calculate the learning rate depending on the current training epoch. The activation function of all convolution layers is the LeakyReLU [48] function. The value of the batch size of the prediction module is 32.

Data Augmentation Image augmentation techniques are used to enhance the efficiency and result of the model. The flipping, rotating, and color transformation data augmentation was used to improve the quality and generalization of images [49] so the hybrid model is trained by random augmentations: color jitter, flipping, shifting, rotating, and scaling. The number of generated augmented samples from each data sample is 25.

Training Loss The task of detecting the object contains both classification (object recognition) and regression (localization of the object) learning problems. The loss function of Faster R-CNN as a two-stage detector can be unified as in (3):

$$\mathcal{L}(\{p_i\}, \{t_i\}) = \frac{1}{N_{cls}} \sum_i \mathcal{L}_{cls}(p_i, p_i^*) + \lambda \frac{1}{N_{reg}} \sum_i p_i^* \mathcal{L}_{reg}(t_i, t_i^*) \tag{3}$$

$$\mathcal{L}_{cls}(p_i, p_i^*) = -\log[p_i^* p_i + (1 - p_i^*)(1 - p_i)] \tag{4}$$

$$\mathcal{L}_{reg}(t_i, t_i^*) = smooth_{L1}(t_i - t_i^*) = \begin{cases} 0.5(t_i - t_i^*)^2 & \text{if } |t_i - t_i^*| < 1 \\ |t_i - t_i^*| - 0.5 & \text{otherwise} \end{cases} \tag{5}$$

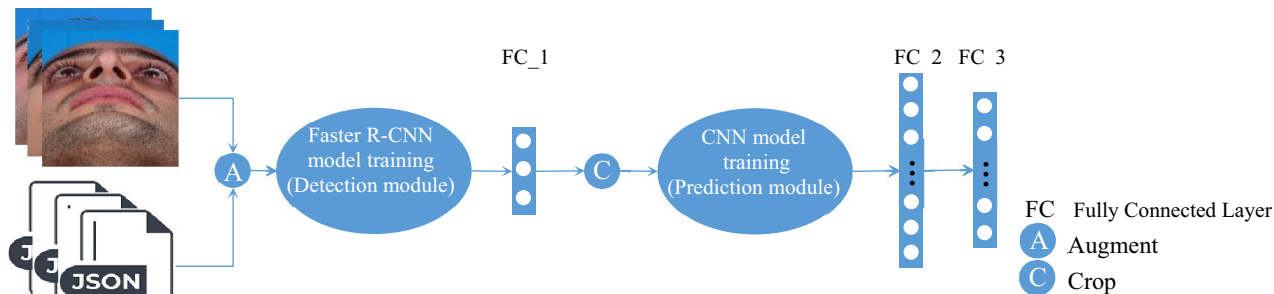


Fig. 3 The hybrid model combined the Faster R-CNN and CNN models

where i is the index of an anchor in a mini-batch, p_i is the predicted probability of the anchor i being an object, p_i^* is the ground-truth label that is equal to one if the anchor is positive, and is zero if the anchor is negative. The vectors t_i and t_i^* represent 4 parameter coordinates of the predicted box and ground-truth box associated with a positive anchor, respectively, and λ is the weight balance parameter also, and N_{cls} and N_{reg} are the mini-batch sizes and the number of anchor locations, correspondingly. The smooth function calculates the loss function between the ground-truth and the predicted box. We formulate the task of predicting the nasal base landmark coordinates as a regression problem. The loss function of the model is the mean absolute error (MAE).

Evaluation Metrics Since in a given image, the aim of object detection is to find out where objects are located and each object belongs to which category; therefore, the well-known metrics, such as precision and recall, are not sufficient for this task. To effectively evaluate the performance of the detection module, the precision (P), the recall rate (R), F1 score (F1), and the mean average precision (mAP) are selected to evaluate the detection ability of the model. The metrics of MAE and the normalized mean errors (NME) [50] are used to measure the performance of the prediction module are given as follows:

$$NME = \frac{1}{n} \sum_j \left(\frac{1}{d_j} \sum_k^m \|\tilde{Y}_j(:,k) - Y_j(:,k)\|_2 \right) \tag{6}$$

where n and m are the numbers of the samples (images) and landmarks, respectively, Y_j is the ground-truth landmarks, \tilde{Y}_j is the corresponding estimated landmarks, $Y_j(:,k)$ is the k th column of Y_j , and d_j is $\sqrt{width * height}$ that is the square root of the ground-truth bounding box.

In constructing the hybrid model, various experiments are performed to fine-tune the architecture and the learning process, including changing the pre-trained model, batch size, normalization scheme, learning rate, and the number of augmentation and layers. It should be noted that we

examined several schedules of “LearningRateScheduler” as the standard, linear, step-based, polynomial learning rate and exponential decay for training the prediction module then, the best result of them is selected as the result of the models.

Combined Model Based on the Geometry Features and Tiled Images

After constructing the hybrid model, the testing part of the data (156 images) is partitioned into two parts: the training set (126 images) for the symmetry predicting model and the testing set (30 images). The 156 accumulated nasal base images included 118 women and 38 men patients. To prepare the images, first, the goal area of the nasal base is cropped from the original image after standardization (resizing the images to a uniform size for annotation of the pupils). Additionally, the rotation of the patient’s head is removed to ensure that the inferior right pupil connects to the inferior left pupil, which should be parallel to a line of the Frankfort plane. Then, the image of the nasal base is masked with white color to separate the nasal base region of the image from the background. Typically, white color background is more efficient for masking irrelevant parts of the image [51] and is used to extract the region of interest (ROI). After that, the nasal base is divided into two regions by the line connecting pronasale and subnasale Fig. 4. Finally, a tiled image is constructed according to a vector of length 4 (the total number of tiles in the tiled image). The tiled normalized version of images and geometry features (\mathcal{F}) of the training set (80% for training and 20% for validation) is utilized to train the network to find the regression model. The tiled image is passed through the CNN, and the symmetry value is used as the target predicted value. This approach permits the CNN to learn from all images rather than trying to pass the images one at a time, it also enables the CNN to learn discriminative filters from all images at once. The z -score method is used to obtain the normalized pixel value (NPV) as in (7):

$$NPV(x, y, z) = \frac{PV(x, y, z) - \mu(PV(z))}{\sigma(PV(z))} \tag{7}$$

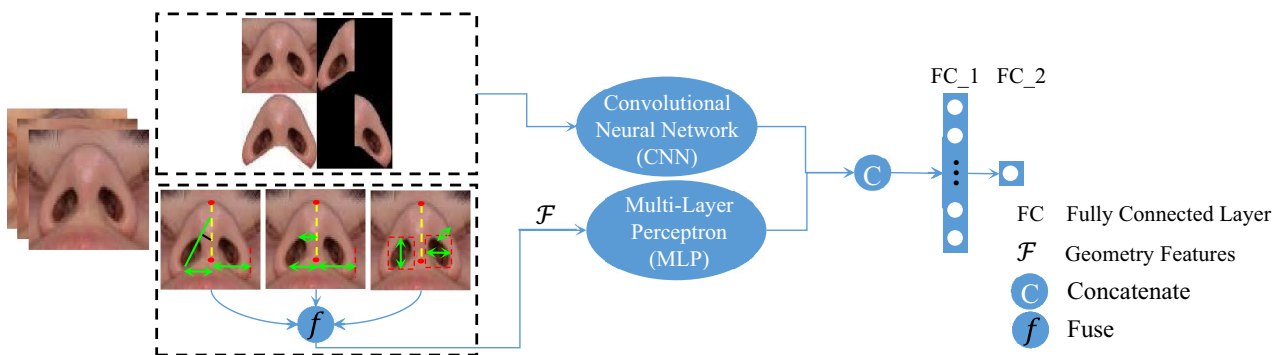


Fig. 4 The combined model (CNN and MLP) to predict the nasal base symmetry

Table 1 Performance of the detection module for the Faster and the Fast R-CNN methods

	Precision rate	Recall rate	F1 score	mAP
Faster R-CNN	0.78	1	0.88	0.92
Fast R-CNN	0.69	1	0.81	0.87
<i>P</i> -value	0.00	0.40	0.88	0.01

where $PV(x,y,z)$ is the pixel value of the coordinate (x,y) and channel (z) . $\mu(PV(z))$ and $\sigma(PV(z))$ are the mean and standard deviation values of pixels in channel z of the training set. Also, the geometry features are transformed into a common range $[0,1]$; then, the min–max normalization method is used to calculate the normalized geometry feature (NGF_i) for each geometry feature i by (8):

$$NGF_i = \frac{GF_i - GF_{min}}{GF_{max} - GF_{min}} \quad (8)$$

Here, GF_{max} and GF_{min} are the maximum and minimum values of the i -th geometry feature (\mathcal{F}), respectively. MLP is trained with geometry features, and CNN is trained with tiled images; then, CNN and MLP are combined to estimate nasal base symmetry. Then the obtained ground-truth symmetry by SSIM was used as a combined model target. The combined model was compiled using Adam optimizer with the learning rate value set at $1e-3$ and batch size equal to 8. The loss function and performance metric of the combined model are MAE and the root mean square error (RMSE), respectively. Also, the Pearson correlation coefficient (PC) is used to measure the performance of the model. Several experiments were examined to fine-tune hyperparameter settings to reach optimal performance. The final architecture follows the combination of MLP and CNN models depicted in Fig. 4. The detailed description of the structure of the parameter setup of the combined model components is given in Appendix 2.

Comparing the Outcomes

In this stage, the results generated by the automatic framework and the otolaryngologists' ratings were analyzed. The testing

set (30 images) was given to otolaryngologists (2 men and 2 women) through a visual questionnaire and asked them to rate the perceived symmetry (completely symmetric, very symmetric, slightly symmetric, asymmetric, and completely asymmetric). It should be noted that all the images are assigned to the patient without any defect in the nasolabial region.

Results and Discussion

In this section, the performance of the proposed symmetry AF is evaluated. Firstly, the efficiency of the introduced hybrid model is explained; then, the performance of the combined models is analyzed. The transfer learning technique is applied in the hybrid model based on the CNN approach such that Faster R-CNN with pre-trained VGG16 is used to extract bounding box information and a deep learning model with ResNet152V2 as the backbone is exploited to detect landmarks coordinates. Then, the deep CNN and multi-layer perceptron neural network models are integrated by concatenating their last hidden layer to evaluate the nasal base symmetry based on geometry features and tiled images of the nasal base. Finally, the process of matching is executed between the AF results and otolaryngologists' ratings. All experiments were performed on a Windows machine pre-installed with a 64-bit Win 10 Pro. It has a GTX 1080 Ti 22 GB GPU, 32 GB of RAM, and an Intel(R) Xeon(R) CPU E5-2699 v4 @ 2.20 GHz. Also, the AF is developed by using the Python programming language.

Evaluating the Hybrid Model

We obtained the results of the precision rate, recall rate, F1 score, and mAP value of the “nasal_base” class for the Faster RCNN (Table 1). These reported values of the results are achieved according to the testing dataset (156 images). In addition, the plots of the PR curve, loss, and accuracy metric plots are shown to further evaluate the efficiency of the Faster R-CNN (Fig. 5) such that the minimum total loss of 0.0085 and maximum accuracy of 0.9996 acquired

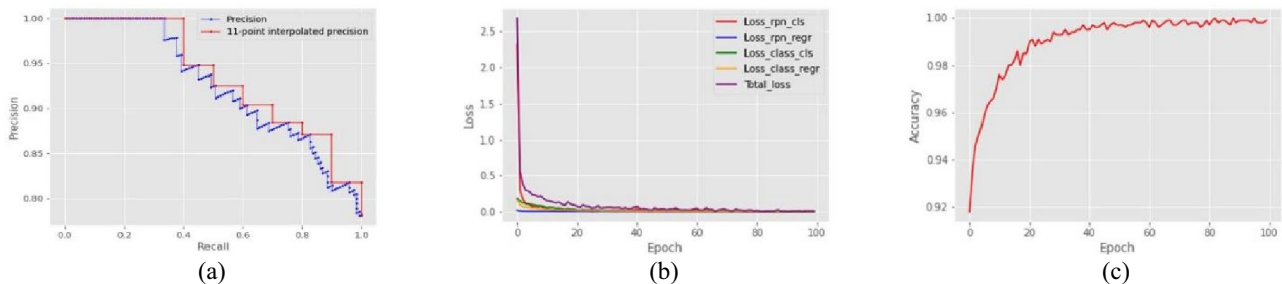


Fig. 5 PR curve (precision at 11 recall levels) (a), training and validation loss for the Faster RCNN (b), training and validation accuracy for the Faster RCNN (c)

Table 2 The performance metrics (also their corresponding values of the mean and Std in all epoch) of alignment results of the nasal base

Partition	MAE	Mean (Std)	NME	Mean (Std)	RMSE	Mean (Std)	PC	Mean (Std)
Validation	0.0442	0.0542 (0.0536)	0.0162	0.0201 (0.0170)	0.1210	0.1245 (0.0607)	0.8607	0.8490 (0.0730)
Test	0.0359	0.0673 (0.0849)	0.0137	0.0268 (0.0359)	0.0953	0.0965 (0.1291)	0.9011	0.7181 (0.5238)
<i>P</i> -value	0.4079		0.4347		0.4341		0.3801	

during the process of model training in epoch 96. Besides, we applied the Fast R-CNN with VGG16 for the nasal base detection and the performance-based comparisons in the object detection methods indicate that the Faster R-CNN model outperforms the Fast R-CNN model (Table 1). The paired *t*-test analysis with a 0.05 significant level was used to compare the performance metrics of Faster R-CNN and Fast R-CNN. The findings revealed a significant difference in the precision rate and mAP metrics as indicated by the last row of Table 1.

The prediction module is trained to detect the nasal base landmarks based on the predicted bounding box extracted by the Faster R-CNN model. The network weights of the module were updated after improving the validation loss from previous epochs. At the end of the training, we pick the model that gives the best performance of validation loss in terms of MAE and NME to evaluate the validation and test datasets and to use later for the landmark predicting process. Table 2 demonstrates the MAE, NME, RMSE, PC, Mean, and Std of them for the prediction module on the SRD. The validation set is used to tune the parameters of the regressor, and the test set is used only to assess the performance of the fully specified regression model. Evaluating the validation and the test sets using the trained prediction module shows that MAE, NME, and RMSE are low for the two sets. The independent *t*-test was conducted at a significance level of 0.05 to compare the performance metrics of the validation and test sets. The results indicated no significant difference among the four metrics of the two sets, suggesting that the prediction module is equally effective at nasal base landmarking in both the validation and test sets. Figure 6 shows the qualitative results of the hybrid model.

It should be pointed out that we examined two pre-trained VGG-16 and Resnet50 architectures for the detection module

and utilized Resnet152V2 and Xception backbones for the prediction module. The hybrid model integrates both detection and prediction techniques in order to provide accurate predictions because each technique is responsible for a different task. Also, the hybrid model decreases the human role in landmark detection which facilitates using it.

Performance of the Combined Model

The combined model consisted of the CNN and MLP trained to estimate the nasal base symmetry via the tiled images and the geometry features. To do this, we randomly selected 126 images of the testing dataset (156 images) for training the combined model. The weights of the network were set to update only when the validation loss function improved from previous epochs. After training for 100 epochs, we roll back to the model with the lowest validation loss in terms of the MAE based on the training dataset. The model performance metrics (MAE, RMSE, PC) for the validation set (31 images) are 0.1363, 0.0665, and 0.7669, and the values of these metrics for the test set (30 images) are 0.1585, 0.0863, and 0.8394, respectively.

In this experiment, we examined the AF to measure the prediction results of the nasal base symmetry. Some sample images of the testing set and their heatmap visualization with the symmetry predicted results of the nasal base images by AF beside ground-truth (or the SSIM value) are shown in Fig. 7. The heatmap is generated through an intersection of the heatmaps derived from both the CNN and the geometric features. The CNN heatmap highlights the tiles used by the network as important features in tiled images, while the geometric feature heatmap is generated from features with a large distance. The resulting heatmap identifies the region of the nasal base that disrupts symmetry.

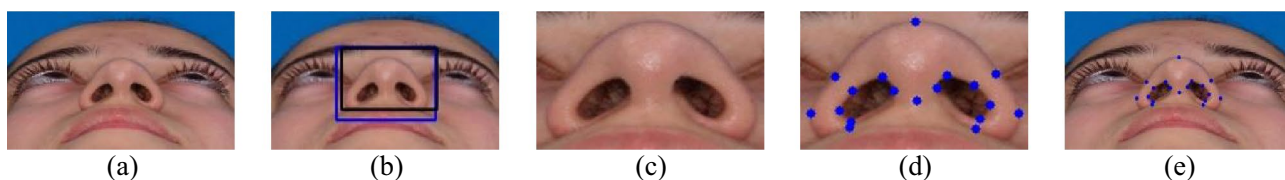


Fig. 6 Qualitative result of the hybrid model. Input image (a), the output of the detection module (b), cropped region (c), output of the prediction module (d), and the returned landmarks on the input image

(e). The black and blue bounding boxes are detected and ground-truth bounding boxes (b)

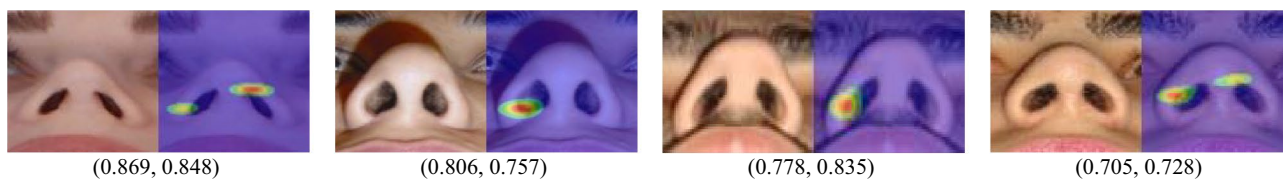


Fig. 7 The original image and the heatmap visualization of the cropped nasal base region. The pair values show (the ground-truth value, the predicted values by AF)

Therefore, after data preparation, the clinical application of AF is according to this process that, when a new nasal base image comes in, the landmarks are detected by using the hybrid model. Then, based on the detected landmarks, the nine paired geometry features are calculated. Then the derived geometry features and the tiled image are fed to the combined model. Finally, the predicted symmetry value obtained from the AF and the heatmap visualization of the nasal base is shown to users.

Matching Process of the AF and Otolaryngologists’ Results

The matching process of the results of the AF for the testing set (30 images) and otolaryngologists’ ratings reports that in 19 cases of 30 patients, the combined model predicts the symmetry value according to the opinions of specialist doctors. The results of the matching process showed that the exact and fine matchings of AF with the otolaryngologists’ ratings are 0.6333 and 0.7666. The considered range for transforming quantitative ranges to qualitative ranks is based on the otolaryngologists’ opinions. The important factors that cause gaps between the otolaryngologists’ ratings and the results of the combined model are as follows: there are no standard angles of images, errors in human vision, and errors in the combined model.

Ablation Study

In order to investigate the effectiveness of the detection module, an extensive ablation study was conducted with the goal of surveying the impact of the detection module to decrease the error of the prediction module’s output. The trained prediction module, after 100 epochs, is evaluated by MAE, the mean average pixel error (MAPE) [52] (9), RMSE, and PC metrics to compute the landmark estimation errors.

$$MAPE = \frac{1}{n} \sum_j^n \sum_k^m \|\tilde{Y}_j(:, k) - Y_j(:, k)\|_2 \tag{9}$$

Module Robustness Analysis

The proposed AF uses the principle of object detection before landmarks recognition. Extensive experiments on augmented data are conducted. To assess the efficacy of the detection module, two models are formed: Model I, prediction module without detection module based on the whole image and Model II, prediction module with detection module based on the cropped nasal base region. Note that Model I and Model II are trained only with the SRD training dataset (Fig. 1). Assessing the outputs of Model I and Model II is done in two methods: the output of Model I is transformed to the scale of the output of Model II (Algorithm 1) and on the contrary (Algorithm 2).

Algorithm 1 Transformation to zoom-in for the nasal base region (maximize)

```

Input = { Images: I, Landmarks_coordinate: < x, y >,
bbox_coordinate: < x_bbox1, x_bbox2, y_bbox1, y_bbox2 >,
new_scale: < x_scale, y_scale > }

Output: New coordinates of landmarks < X, Y >
1: K = number of landmarks
2: foreach Image, I, do
3: for j = 1 to K do
4: [x'_{j1} y'_{j1}] = [x_{j1} y_{j1}] - [x_bbox11 y_bbox11]
5: [X'_{j1} Y'_{j1}] = [x'_{j1} y'_{j1}] * [ x_scale / (x_bbox21 - x_bbox11)
y_scale / (y_bbox21 - y_bbox11) ]
6: end for
    
```

Algorithm 2 Transformation to zoom-out for the nasal base region (minimize)

```

Input = { Images: I, Landmarks_coordinate: < x, y >,
bbox_coordinate: < x_bbox1, x_bbox2, y_bbox1, y_bbox2 >,
new_scale: < x_scale, y_scale > }

Output: New coordinates of landmarks < X, Y >
1: K = number of landmarks
2: foreach Image, I, do
3: for j = 1 to K do
4: [x'_{j1} y'_{j1}] = [x_{j1} y_{j1}] * [ (x_bbox21 - x_bbox11) / x_scale
(y_bbox21 - y_bbox11) / y_scale ]
5: [X'_{j1} Y'_{j1}] = [x'_{j1} y'_{j1}] + [x_bbox11 y_bbox11]
6: end for
7: end foreach
    
```

To compare the performance of the two models, the paired *t*-test at a significance level of 0.05 is implemented. The *P*-values of Table 3 show that there is a significant difference between the evaluation metrics (MAE, MAPE,

Table 3 Performance of the prediction module for Model I and Model II

Dataset	Transformation method	Model	MAE	MAPE	RMSE	PC
SRD	Algorithm 1	I	0.0375	0.2870	0.0478	0.9447
		II	0.0329	0.2577	0.0429	0.9529
		P-value	0.0000	0.0000	0.0000	0.0002
	Algorithm 2	I	0.0077	0.0604	0.0101	0.9763
		II	0.0079	0.0604	0.0101	0.9817
		P-value	0.0004	0.0021	0.0021	0.1401
MERD	Algorithm 1	I	0.0796	0.6404	0.1067	0.7632
		II	0.0533	0.4313	0.0719	0.8813
		P-value	0.0000	0.0000	0.0000	0.0004
	Algorithm 2	I	0.0257	0.2150	0.0358	0.8135
		II	0.0140	0.1154	0.0192	0.9262
		P-value	0.0000	0.0000	0.0000	0.0000

RMSE, and PC) of the Model I and Model II for both datasets. Since the values of the MAE, MAPE, and RMSE of the Model II are less or equal to the Model I, Model II has a better performance, independent of algorithms 1 and 2. This means that the learning on the cropped region of the image can be more effective for image feature extraction (Fig. 8). It should be noted that the MAE, MAPE, and RMSE values are higher on MERD, which indicates the error of the models in testing these images. This error can be attributed to the diverse nature of the images and, at some times, the lack of training of the model on various types of nasal bases.

The Impact of Landmarks

A study was conducted to analyze the impact of incorrect landmarks on the performance of the prediction module. This involved randomly shifting the position of correct 3–7 landmarks on two datasets by 1–5 pixel points in different directions (top, bottom, right, and left). The results of paired *t*-tests, as shown in Table 4, indicate that there is a significant difference between the Model I and Model \hat{I} , except for the algorithms 2 for MERD dataset. However, no significant difference was observed between Model II and Model \hat{II} , suggesting that Model II is more robust in the presence of incorrect landmarks compared to Model I. Additionally, the *P*-value between Model \hat{I} and Model \hat{II} declares a significant difference in the evaluation metrics (MAE, MAPE, and RMSE) for both datasets. Furthermore, the model evaluation metrics (except PC) with incorrect landmarks are slightly greater than those with correct landmarks (PC: 0.9169 vs. 0.9565, 0.9479 vs. 0.9817, 0.7669 vs. 0.8813, 0.8632 vs. 0.9262).

The Impact of Noise in Images

We utilized Gaussian noise, as observed in recent studies on assessing robustness [53–56], to study its impact on the

images. Subsequently, we compared the outcomes of the noise dataset with those of the non-noise dataset (Table 5). The results of the paired *t*-test to investigate the effect of noise on the Model I and Model II indicate a significant difference between the noisy (\hat{I}) and noiseless (I) modes, while for the Model II, there is no significant difference between the noise (\hat{II}) and noiseless (II) modes. Therefore, the robustness of Model II, which is based on object detection, is confirmed. Moreover, the *P*-value comparison between Model \hat{I} and Model \hat{II} suggests a significant disparity in the evaluation metrics (MAE, MAPE, and RMSE) for both datasets.

The limitations of prior studies include the use of statistical analysis to examine medical interventions in surgery [8, 34, 38], reliance on user participation until the result is obtained [4, 32, 37], and the inability of the presented methods to provide a quantitative measure of symmetry [4, 32, 37, 39]. These studies have advantages in addressing the issues of a specific range of patients and offering computer-based approaches. In our research, we have addressed the limitations of previous studies by employing an automated method with repeatability and high accuracy to predict the quantitative value of symmetry.

The issue of imprecise bounding box localization presents a significant area for future research. In an attempt to minimize this problem in this research, after accurate annotations for training through manual and labor-intensive labeling by motivation of [57], we initiated with a larger bounding box. However, our team has thoroughly investigated this topic and plans to address it in an upcoming article, utilizing insights from recent publications [58–60]. As this matter exceeds the scope of the current work, we have conducted a comprehensive examination and analysis to be incorporated into the forthcoming article. The authors have openly shared the model, including its code, and have made it available to all interested researchers, while excluding the dataset.

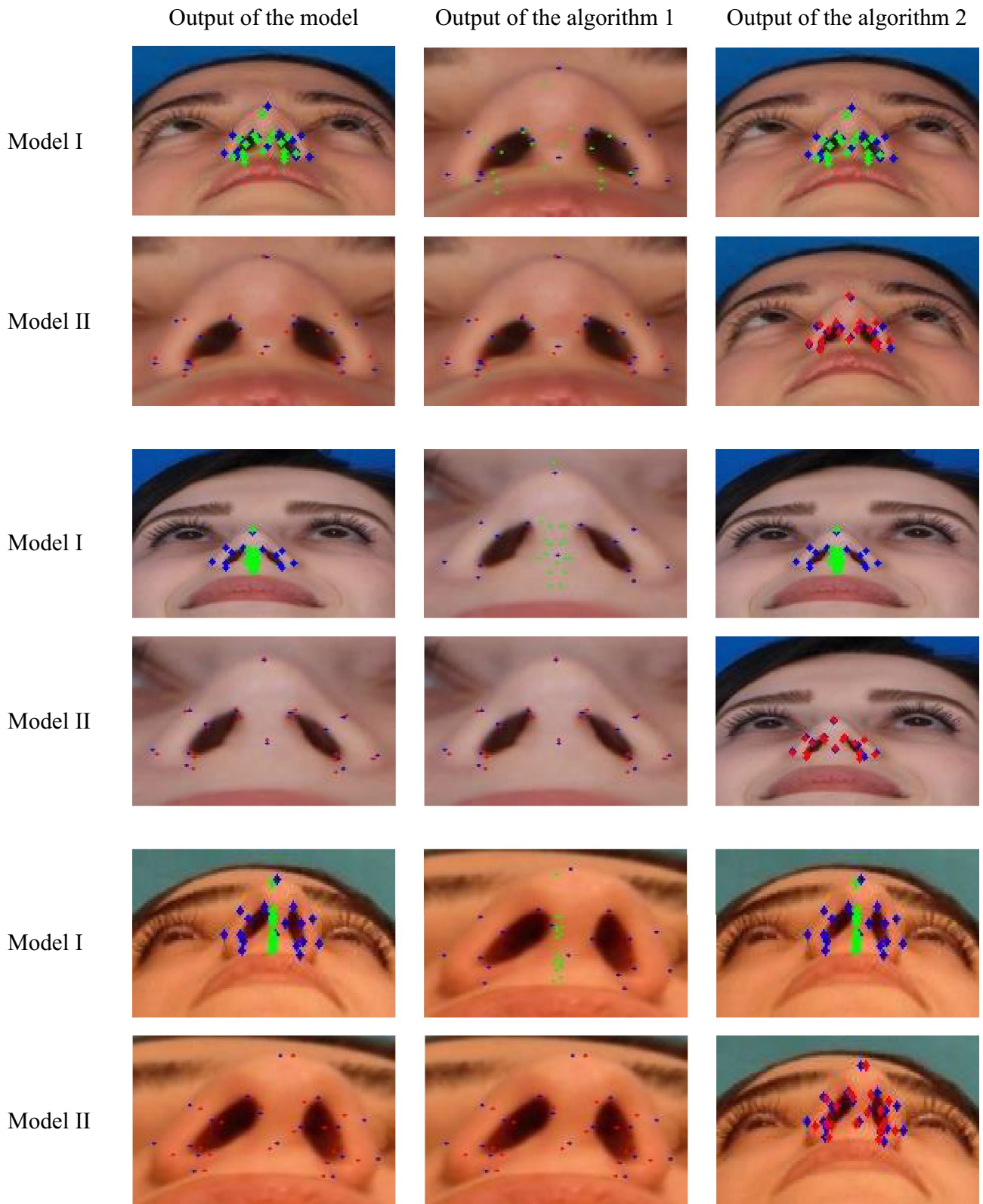


Fig. 8 Qualitative results of the AF on SRD. Blue: ground-truth landmarks. Green: detected landmarks by Model I. Red: detected landmarks by Model II. The first column is the model output, the second

column is the results of zoom-in using algorithm 1, and the third column is the results of zoom-out by algorithm 2

Table 4 Impact of landmarks on the prediction module for two datasets according to Model I and Model II ($\hat{\cdot}$ represents the model with the incorrect landmarks. P -value* is between \hat{I} and \hat{II} models)

Dataset	Transformation method	Model	MAE	P -value	MAPE	P -value	RMSE	P -value	PC	P -value	
SRD	Algorithm 1	\hat{I}	0.0464	0.0048	0.3583	0.0109	0.0597	0.0109	0.9165	0.1670	
		I	0.0375		0.2870		0.0478		0.9447		
		\hat{II}	0.0323	0.8530	0.2652	0.7850	0.0442	0.7850	0.9169	0.0486	
		II	0.0329		0.2577		0.0429		0.9529		
		P -value*	0.0083		0.0499		0.0499		0.9992		
	Algorithm 2	\hat{I}	0.0094	0.0181	0.0740	0.0389	0.0123	0.0389	0.9525	0.1304	
		I	0.0077		0.0604		0.0101		0.9763		
		\hat{II}	0.0080	0.9879	0.0671	0.5357	0.0112	0.5357	0.9479	0.0806	
		II	0.0079		0.0604		0.0101		0.9817		
		P -value*	0.0000		0.0000		0.0000		0.3296		
	MERD	Algorithm 1	\hat{I}	0.0944	0.0000	0.7332	0.0000	0.1222	0.0000	0.7511	0.2822
			I	0.0796		0.6404		0.1067		0.7632	
\hat{II}			0.0717	0.6054	0.6011	0.9168	0.1002	0.9168	0.7669	0.0000	
II			0.0533		0.4313		0.0719		0.8813		
P -value*			0.0000		0.0000		0.0000		0.0004		
Algorithm 2		\hat{I}	0.0274	0.2915	0.2311	0.2958	0.0385	0.2958	0.7719	0.1180	
		I	0.0257		0.2150		0.0358		0.8135		
		\hat{II}	0.0198	0.0554	0.1693	0.0631	0.0282	0.0631	0.8632	0.1053	
		II	0.0140		0.1154		0.0192		0.9262		
		P -value*	0.0003		0.0014		0.0014		0.0020		

Table 5 Performance of the noise and non-noise on two datasets applying Model I and Model II ($\check{\cdot}$ and $\ddot{\cdot}$ represent the models with the noise. P -value* is between \check{I} and \ddot{II} models)

Dataset	Transformation method	Model	MAE	P -value	MAPE	P -value	RMSE	P -value	PC	P -value	
SRD	Algorithm 1	\check{I}	0.0579	0.0006	0.4651	0.0143	0.0775	0.0143	0.8382	0.0001	
		I	0.0375		0.2870		0.0478		0.9447		
		\check{II}	0.0428	0.7120	0.3535	0.5982	0.0589	0.5892	0.8703	0.1283	
		II	0.0329		0.2577		0.0429		0.9529		
		P -value	0.0101		0.0289		0.0289		0.2755		
	Algorithm 2	\check{I}	0.0122	0.0000	0.1001	0.0000	0.0167	0.0000	0.8872	0.0617	
		I	0.0077		0.0604		0.0101		0.9763		
		\check{II}	0.0105	0.0393	0.0889	0.2320	0.0148	0.2320	0.9032	0.0312	
		II	0.0079		0.0604		0.0101		0.9817		
		P -value	0.0000		0.0003		0.0003		0.3187		
	MERD	Algorithm 1	\check{I}	0.1739	0.0313	1.4912	0.0005	0.2485	0.0005	0.1641	0.0000
			I	0.0796		0.6404		0.1067		0.7632	
\check{II}			0.0822	0.5719	0.7035	0.1740	0.1172	0.1740	0.6868	0.0000	
II			0.0533		0.4313		0.0719		0.8813		
P -value			0.0383		0.0019		0.0019		0.0069		
Algorithm 2		\check{I}	0.0486	0.0000	0.4273	0.0000	0.0712	0.0000	0.4980	0.0000	
		I	0.0257		0.2150		0.0358		0.8135		
		\check{II}	0.0229	0.0777	0.1993	0.0640	0.0332	0.0640	0.8053	0.2870	
		II	0.0140		0.1154		0.0192		0.9262		
		P -value	0.0000		0.0000		0.0000		0.0000		

Conclusions

This study presented a novel automatic framework (AF) based on deep learning algorithms for evaluating nasal base symmetry. The goal of this framework is to improve the previous symmetry methods based on quantitative and qualitative assessments. The suggested AF is compatible with otolaryngologists' ratings and reduces the role of human factors in evaluating symmetry, thereby aiding surgeons in planning expedited operations and improving surgical outcomes. In the proposed framework, two detection and prediction modules are combined into an integrated hybrid model to address the problem of nasal base landmark recognition. The standard data augmentation techniques and Adam optimizer are implemented to achieve the optimum value of the loss function during the training process. The evaluated results of the trained prediction module show that learning on the cropped region of the image can be more effective for image feature

extraction. A combined model was developed and trained using tiled images and the geometric features of the nasal base to predict nasal base symmetry. The results of the matching process confirmed that the AF was consistent at 0.7666 with the otolaryngologists' ratings which demonstrates the efficiency of the proposed framework. Also, the capability of the proposed AF can be seen in decreasing the rhinoplasty reconstructive surgery, especially, in cleft palate subjects that require a precise description of the symmetry and a visual representation of its changes. To sum up, the proposed AF is capable of sensing the nasal base symmetry and capturing the asymmetry regions to demonstrate a heatmap visualization for patients and otolaryngologists. Consequently, our future research will concentrate on studying and designing unsupervised learning-based object detectors to develop our hybrid model and construct a real-time symmetry AF. Furthermore, we suggest that future studies of the introduced hybrid model can be utilized in various application domains.

Appendix 1

List of landmarks of the nasal base [32, 61]

Facial landmark	Abbreviation	Description
Alare _{L/R}	Al _{L/R}	The most lateral point in the curved of each ala
Subalare _{L/R}	Sbal _{L/R}	The point at the lower limit of each alar base, where the alar base disappears into the skin of the upper lip
Al _{L/R}	Al _{L/R}	The marking level at the midportion of the alae
Nostril tip _{L/R}	Nt _{L/R}	The most cranial point of the inner border of the nostril
Nostril base _{L/R}	Nb _{L/R}	The most caudal point of the inner border of the nostril
Nostril mediale _{L/R}	Nm _{L/R}	The most medial point of the inner border of the nostril
Nostril laterale _{L/R}	Nl _{L/R}	The most lateral point of the inner border of the nostril
Pronasale	Prn	The most protruding point on nasal tip
Subnasale	Sn	The midline point at the junction of the inferior margin of the columella and the upper lip skin

Appendix 2

General parameters of the structure utilized to setup the combined model.

	Layers	Input size	Kernel size	Strides	Activation function	Padding	Output size	
MLP	Dense_1	9	-	-	ReLU	-	18	
	Dense_2	18	-	-	ReLU	-	9	
CNN	Block_1	Conv2D_1	64×64×3	3×3×32	1,1	-	same	64×64×32
		BatchNormalization_1	64×64×32	-	-	-	-	64×64×32
		Activation_1	64×64×32	-	-	ReLU	-	64×64×32
		MaxPooling2D_1	64×64×32	2×2	-	-	-	32×32×32
	Block_2	Conv2D_2	32×32×32	3×3×64	1,1	-	same	32×32×64
		BatchNormalization_2	32×32×64	-	-	-	-	32×32×64
		Activation_2	32×32×64	-	-	ReLU	-	32×32×64
		MaxPooling2D_2	32×32×64	2×2	-	-	-	16×16×64
	Block_3	Conv2D_3	16×16×64	3×3×128	1,1	-	same	16×16×128
		BatchNormalization_3	16×16×128	-	-	-	-	16×16×128
		Activation_3	16×16×128	-	-	ReLU	-	16×16×128
		MaxPooling2D_3	16×16×128	2×2	-	-	-	8×8×128
	Block_4	Conv2D_4	8×8×128	3×3×256	1,1	-	same	8×8×256
		BatchNormalization_4	8×8×256	-	-	-	-	8×8×256
		Activation_4	8×8×256	-	-	ReLU	-	8×8×256
		MaxPooling2D_4	8×8×256	2×2	-	-	-	4×4×256
	Block_5	Conv2D_5	4×4×256	3×3×512	1,1	-	same	4×4×512
		BatchNormalization_5	4×4×512	-	-	-	-	4×4×512
		Activation_5	4×4×512	-	-	ReLU	-	4×4×512
		MaxPooling2D_5	4×4×512	2×2	-	-	-	2×2×512
	Flatten	2×2×512	-	-	-	-	2048	
	Dense_3	2048	-	-	-	-	36	
	Activation_6	36	-	-	ReLU	-	36	
	BatchNormalization_6	36	-	-	-	-	36	
	Dropout_1	36	=	=	=	=	36	
	Dense_4	36	-	-	-	-	9	
	Activation_7	9	-	-	ReLU	-	9	
Combined	Concatenate_1	[9]	=	=	=	=	18	
	Dense_1	18	-	-	ReLU	-	9	
	Dense_2	9	-	-	Linear	-	1	

Acknowledgements We would like to extend our sincere gratitude to the anonymous reviewers for their meticulous evaluation and constructive feedback, which significantly enhanced the quality and rigor of this paper. Their expertise and thoughtful comments have been invaluable in shaping the final version of this work. The authors also would like to thank Dr. Amin Amali for preparing and evaluating the results.

Author Contribution All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Maryam Ashoori, Reza A. Zoroofi, and Mohammad Sadeghi. The first draft of the manuscript was written by Maryam Ashoori, and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Declarations

Informed Consent Written informed consent was obtained from all individual participants included in the study.

Competing Interest The authors declare no competing interests.

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/>.

References

- Dong Y, Zhao Y, Bai S, Wu G, Wang B: Three-dimensional anthropometric analysis of the Chinese nose. *J Plast Reconstr Aesthet Surg*, 63 (11): 1832-1839, 2010
- Roxbury C, Ishii M, Godoy A, Papel I, Byrne P. J, Boahene K. D. O, Ishii L. E: Impact of Crooked Nose Rhinoplasty on Observer Perceptions Of Attractiveness. *Laryngoscope*, 122 (4): 773-778, 2012
- Choi J.Y: Alar base reduction and alar-columellar relationship. *Facial Plast Surg Clin North Am*, 26 (3): 367–375, 2018
- Barnes C.H, Chen H, Chen J.J, Su E, Moy W.J, Wong B.J: Quantitative Analysis and Classification of the Nasal Base Using a Parametric Model. *JAMA Facial Plast. Surg*, 20 (2): 160–165, 2018
- Chandaliya P.K, Nain N: ChildGAN: Face aging and rejuvenation to find missing children. *Pattern Recognit*, 129, 2022
- Cohn J.E, Shokri T, Othman S, Sokoya M, Ducic Y: Surgical Techniques to Improve the Soft Tissue Triangle in Rhinoplasty: A Systematic Review. *Facial Plast Surg*, 36 (1): 120-128, 2020
- Heilbronn C, Cragun D, Won B.J.F: Complications in Rhinoplasty: A Literature Review and Comparison with a Survey of Consent Forms. *Facial plast. surg. aesthet. med*, 22 (1): 50–56, 2020
- Cabbarzade C: Rhinoplasty Technique for Improving Nasal Base Aesthetics: Lateral Columellar Grafting. *Ann Plast Surg*, 90 (5):419-424, 2023.
- Eggerstedt M, Schumacher J, Urban M.J, Smith R.M, Revenaugh P.C: The Selfie View: Perioperative Photography in the Digital Age. *Aesthetic Plast Surg*, 44 (3): 1066–1070, 2020
- Cabbarzade C: Septal bony paste graft: a life-saving material in rhinoplasty camouflage. *Aesthetic Plast Surg*, 47(5):1967-1974, 2023
- Thomas J, Raj E.D: Deep Learning and Multimodal Artificial Neural Network Architectures for Disease Diagnosis and Clinical Applications: Machine Learning and Deep Learning in Efficacy Improvement of Healthcare Systems. 1st edition, Boca Raton: CRC, 2022, p. 27
- Hidaka T, Kurita M, Ogawa K, Tomioka Y, Okazaki M: Application of Artificial Intelligence for Real-Time Facial Asymmetry Analysis. *Plast. Reconstr. Surg*, 146 (2): 243e-245e, 2020
- Chinski H, Lerch R, Tournour D, Chinski L, Caruso D: An Artificial Intelligence Tool for Image Simulation in Rhinoplasty. *Facial Plast Surg*, 38 (02): 201-206, 2022
- Lee T.-Y, Lin C.-H, Lin H.-Y: Computer-Aided Prototype System for Nose Surgery. *IEEE Trans. Inf. Technol. Biomed*, 5 (4): 271–278, 2001
- Bashiri-Bawil M, Rahavi-Ezabadi S, Sadeghi M, Zoroofi R. A, Amali A: Preoperative Computer Simulation in Rhinoplasty Using Previous Postoperative Images. *Facial Plast Surg Aesthet Med*, 22 (6): 406-411, 2020
- Mao Z, Siebert J. P, Cockshott W. P , Ayoub A.F: Constructing Dense Correspondences to Analyze 3D Facial Change. InProceedings of the 17th International Conference on Pattern Recognition, 2004.
- Bottino A, Simone M.D, Laurentini A, Sforza C: A New 3-D Tool for Planning Plastic Surgery. *IEEE Trans. Biomed. Eng*, 59 (12): 3439-3449, 2012
- Eldaly A.S, Avila F.R, Torres-Guzman R.A, Maita K, Garcia J.P, Serrano L.P, Forte A.J: Simulation and Artificial Intelligence in Rhinoplasty: A Systematic Review. *Aesthetic Plast. Surg*, 46: 2368–2377, 2022
- Zhukhovitskaya A, Cragun D, Su E, Barnes C.H, Wong B.J.F: Categorization and Analysis of Nasal Base Shapes Using a Parametric Model. *JAMA Facial Plast. Surg*, 21 (5): 440–445, 2019
- Borsting E, DeSimone R, Ascha M, Ascha M: Applied Deep Learning in Plastic Surgery: Classifying Rhinoplasty With a Mobile App. *J Craniofac Surg*, 31 (1): 102-106, 2020
- Dorfman R, Chang I, Saadat S, Roostaeian J: Making the Subjective Objective: Machine Learning and Rhinoplasty. *Aesthet. Surg. J*, 40 (5): 493-498, 2020
- Štěpánek L, Kasal P, Měščík J: Machine-Learning and R in Plastic Surgery – Evaluation of Facial Attractiveness and Classification of Facial Emotions. InInternational Conference on Information Systems Architecture and Technology, 2019. Cham: Springer International Publishing.
- Girshick R, Donahue J, Darrell T, Malik J: Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. InProceedings of the IEEE conference on computer vision and pattern recognition, 2014.
- Girshick R: Fast R-CNN. InProceedings of the IEEE international conference on computer vision, 2015.
- S. Ren, K. He, R. Girshick and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *Advances in Neural Information Processing Systems* 28, 2015.
- Hoang V.-T, Huang D.-S, Jo K.-H: 3-D Facial Landmarks Detection for Intelligent Video Systems. *IEEE Trans. Industr. Inform*, 17 (1): 578–586, 2021
- Wu Y, Ji Q: Facial Landmark Detection: A Literature Survey. *Int. J. Comput. Vis*, 127: 115-142, 2019
- Agarkar S, Hande K: Real-Time Markerless Facial Landmark Detection Using Deep Learning. InICT Systems and Sustainability: Proceedings of ICT4SD, 2021. Singapore: Springer Nature Singapore
- Vo X.-T, Jo K.-H: A review on anchor assignment and sampling heuristics in deep learning-based object detection. *Neurocomputing*, 506: 96–116, 2022
- Dhiraj, Jain D. K: An evaluation of deep learning based object detection strategies for threat object detection in baggage security imagery. *Pattern Recognit Lett*, 120: 112–119, 2019

31. Mittal P, Singh R, Sharma A: Deep learning-based object detection in low-altitude UAV datasets: A survey. *Image Vis. Comput*, 104, 2020.
32. Pietruski P, Majak M, Debski T, Antoszewski B: A novel computer system for the evaluation of nasolabial morphology, symmetry and aesthetics after cleft lip and palate treatment. Part 1: General concept and validation. *J Craniomaxillofac Surg*, 45 (4): 491–504, 2016
33. Linden O. E, Taylor H. O, Vasudavan S, Byrne M. E, Deutsch C. K, Mulliken J. B, Sullivan S. R: Three-Dimensional Analysis of Nasal Symmetry Following Primary Correction of Unilateral Cleft Lip Nasal Deformity. *Cleft Palate Craniofac J*, 54 (6): 715-719, 2017
34. Mancini L, Gibson T. L, Grayson B. H, Flores R. L, Staffenberg D, Shetye P. R: Three-Dimensional Soft Tissue Nasal Changes After Nasoalveolar Molding and Primary Cheilorhinoplasty in Infants With Unilateral Cleft Lip and Palate. *Cleft Palate Craniofac J*, 56 (1): 31-38, 2019
35. Mercan E, Morrison C. S, Stuhaug E, Shapiro L. G, Tse R. W: Novel Computer Vision Analysis of Nasal Shape in Children with Unilateral Cleft Lip. *J Craniomaxillofac Surg*, 46 (1): 35-43, 2018
36. Morselli P. G, Pinto V, Negosanti L, Firinu A, Fabbri E: Early Correction of Septum JJ Deformity in Unilateral Cleft Lip–Cleft Palate. *J Plast Reconstr Surg*, 130 (3): 434e-441e, 2012
37. Karube R, Sasaki H, Togashi S, Yanagawa T, Nakane S, Ishibashi N, Yamagata K, Onizawa K, Adachi K, Tabuchi K, Sekido M, Bukawa H: A novel method for evaluating postsurgical results of unilateral cleft lip and palate with the use of Hausdorff distance: presurgical orthopedic treatment improves nasal symmetry after primary cheiloplasty. *Oral Surg. Oral Med. Oral Radiol*, 114 (6): 704-711, 2012
38. Denadai R, Raposo-Amaral C. A, Buzzo C. L, Raposo-Amaral C. E: Paranasal Fat Grafting Improves the Nasal Symmetry in Patients With Parry-Romberg Syndrome. *J Craniomaxillofac Surg*, 30 (3): 958-960, 2019
39. Pietruski P, Majak M, Pawlowska E, Skiba A, Antoszewski B: A novel computer system for the evaluation of nasolabial morphology, symmetry and aesthetics after cleft lip and palate treatment. Part 2: Comparative anthropometric analysis of patients with repaired unilateral complete cleft lip and palate and healthy individual. *J Craniomaxillofac Surg*, 45 (5): 505–514, 2017
40. Wang Z, Bovik A. C, Sheikh H. R, Simoncelli E. P: Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process*, 13 (4): 600-612, 2004
41. Sachdeva S: Fitzpatrick skin typing: Applications in dermatology. *Indian J. Dermatol. Venereol. Leprol*, 75 (1): 93-96, 2009
42. Rhinoplasty and Septorhinoplasty Photography. *J. Vis. Commun. Med*, 30 (3): 135-139, 2007
43. Dr Sedgh, facial Plastic Surgery. Available at <https://www.sedghplasticsurgery.com>. Accessed 16 July 2023.
44. Quaba Edinburgh Ltd. Available at <https://www.quaba.co.uk>. Accessed 19 July 2023.
45. Russell B. C, Torralba A, Murphy K. P, Freeman W. T: LabelMe: A Database and Web-Based Tool for Image Annotation. *Int J Comput Vis*, 77: 157-173, 2008
46. Torralba A, Russell B. C, Yuen J: LabelMe: Online Image Annotation and Applications. *Proc. IEEE*, 98 (8): 1467-1484, 2010
47. Kingma D. P, Ba J: Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.
48. Maas A. L, Hannun A. Y, Ng A. Y: Rectifier Nonlinearities Improve Neural Network Acoustic Models. In *Proceedings of the 30 th International Conference on Machine Learning*, 2013
49. Parvathi S, Selvi S: Detection of maturity stages of coconuts in complex background using Faster R-CNN model. *Biosyst. Eng*, 202: 119-132, 2021
50. Jourabloo A, Liu X: Pose-invariant 3D face alignment. In *Proceedings of the IEEE international conference on computer vision*, 2015
51. Kocejko T, Rumiński J, Mazur-Milecka M, Romanowska-Kocejko M, Chlebus K, Jo K.-H: Using convolutional neural networks for corneal arcus detection towards familial hypercholesterolemia screening. *J. King Saud Univ. - Comput. Inf*, 34 (9): 7225–7235, 2022
52. Yu X, Huang J, Zhang S, Yan W, Metaxas D. N: Pose-Free Facial Landmark Fitting via Optimized Part Mixtures and Cascaded Deformable Shape Model. In *Proceedings of the IEEE international conference on computer vision*, 2013
53. Hong S, Kang M, Kim J, Baek J: Sequential application of denoising autoencoder and long-short recurrent convolutional network for noise-robust remaining-useful-life prediction framework of lithium-ion batteries. *Comput Ind En*, 179: 2023.
54. LIU Y, HUANG Y.-X, ZHANG X, QI W, GUO J, HU Y, ZHANG L, SU H: Deep C-LSTM Neural Network for Epileptic Seizure and Tumor Detection Using High-Dimension EEG Signals. *IEEE Access*, 8: 37495–37504, 2020
55. Wang Z, Zhang S, Zhang C, Wang B: RPFNet: Recurrent Pyramid Frequency Feature Fusion Network for Instance Segmentation in Side-Scan Sonar Images. *IEEE J Sel Top Appl Earth Obs Remote Sens*, 2023.
56. Ghafari S, Ghobadi Tarnik M, Sadoghi Yazdi H: Robustness of convolutional neural network models in hyperspectral noisy datasets with loss functions. *Comput. Electr. Eng*, 90, 2021
57. Skadins A, Ivanovs M, Rava R, Nesenbergs K: Edge pre-processing of traffic surveillance video for bandwidth and privacy optimization in smart cities. In *2020 17th Biennial Baltic Electronics Conference (BEC)*, 2020
58. Xu Y, Zhu L, Yang Y, Wu F: Training Robust Object Detectors From Noisy Category Labels and Imprecise Bounding Boxes. *IEEE Trans Image Process*, 30: 5782 - 5792, 2021
59. Wu D, Chen P, Yu X, Li G, Han Z, Jiao J: Spatial Self-Distillation for Object Detection with Inaccurate Bounding Boxes. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2023
60. Liu C, Wang K, Lu H, Cao Z, Zhang Z: Robust Object Detection With Inaccurate Bounding Boxes. In *European Conference on Computer Vision*, 2022
61. Farkas L. G: *Anthropometry of the head and face*, 2nd edition, USA: Raven Press, 1994.

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