scientific reports

Check for updates

OPEN

Automatic segmentation and classification of frontal sinuses for sex determination from CBCT scans using a two-stage anatomy-guided attention network

Renan Lucio Berbel da Silva[1,6], Su Yang[2,6], DaEl Kim[3,6], Jun Ho Kim[1], Sang-Heon Lim[3], Jiyong Han[3], Jun-Min Kim[4], Jo-Eun Kim[5], Kyung-Hoe Huh[5], Sam-Sun Lee[5], Min-Suk Heo[5✉] & Won-Jin Yi[2,5✉]

Sex determination is essential for identifying unidentified individuals, particularly in forensic contexts. Traditional methods for sex determination involve manual measurements of skeletal features on CBCT scans. However, these manual measurements are labor-intensive, time-consuming, and error-prone. The purpose of this study was to automatically and accurately determine sex on a CBCT scan using a two-stage anatomy-guided attention network (SDetNet). SDetNet consisted of a 2D frontal sinus segmentation network (FSNet) and a 3D anatomy-guided attention network (SDNet). FSNet segmented frontal sinus regions in the CBCT images and extracted regions of interest (ROIs) near them. Then, the ROIs were fed into SDNet to predict sex accurately. To improve sex determination performance, we proposed multi-channel inputs (MSIs) and an anatomy-guided attention module (AGAM), which encouraged SDetNet to learn differences in the anatomical context of the frontal sinus between males and females. SDetNet showed superior sex determination performance in the area under the receiver operating characteristic curve, accuracy, Brier score, and specificity compared with the other 3D CNNs. Moreover, the results of ablation studies showed a notable improvement in sex determination with the embedding of both MSI and AGAM. Consequently, SDetNet demonstrated automatic and accurate sex determination by learning the anatomical context information of the frontal sinus on CBCT scans.

Keywords Deep learning, Sex determination, Frontal sinus, CBCT, Anatomy-guided attention

Sex determination is fundamental in establishing the biological profile of unidentified humans[1]. Accurate individual identification is important to society, particularly sexual dimorphism, as it has significant implications in criminal investigations[1,2]. Skeletal remains of the human body can preserve their shapes and structures even in harsh environments and disaster events, making them suitable for sex determination[3]. Manual morphological measurement and analysis of skeletal remains such as the hand bone[4], hip bone[5], teeth[6], mandible[7], skull[7], and paranasal sinus[8,9] are widely used to determine the sex of individuals. These methods are based on the presence of significant differences in the sizes and shapes of the teeth, pelvis, and skull between males and females[10]. Skull structures, including the paranasal sinuses, teeth, and cranial suture patterns can potentially be used to identify an individual.

[1]Discipline of Oral Radiology, Department of Stomatology, School of Dentistry, University of São Paulo, São Paulo, SP, Brazil. [2]Department of Applied Bioengineering, Graduate School of Convergence Science and Technology, Seoul National University, Seoul 08826, South Korea. [3]Interdisciplinary Program in Bioengineering, Graduate School of Engineering, Seoul National University, Seoul 08826, South Korea. [4]Department of Electronics and Information Engineering, Hansung University, Seoul 02876, South Korea. [5]Department of Oral and Maxillofacial Radiology and Dental Research Institute, School of Dentistry, Seoul National University, Seoul 03080, South Korea. [6]These authors contributed equally: Renan Lucio Berbel da Silva, Su Yang and DaEl Kim. ✉email: hmslsh@snu.ac.kr; wjyi@snu.ac.kr

The frontal sinuses, which are part of the paranasal sinuses of the head, are cavities located inside the frontal bone and can be used as an indicator of sex due to their unique sizes, shapes, and patterns in males and females[10–13]. The frontal sinus completes growth by around the 20th year and remains relatively unchanged throughout adulthood, making it ideal for postmortem identification as well. The unique and stable morphology of the frontal sinuses makes them an important and reliable tool for forensic identification purposes. Frontal sinus imaging can be done using a variety of techniques such as X-ray, computed tomography (CT), cone-beam computed tomography (CBCT), and magnetic resonance imaging (MRI)[10–15]. Several studies have reported success using frontal sinus imaging for individual identification, notwithstanding the challenge of establishing a universally accepted and objective standardized method. Identification of sexual dimorphism in the frontal sinuses has made sex determination based on this anatomical structure a valuable tool to reduce the range of possibilities to be considered during individual identification, hence aiding in the creation of a more dependable biological profile of human remains[10]. CBCT is widely used in the field of dentistry as it provides accurate and detailed three-dimensional imaging of the maxillofacial region. Its advantages over other imaging techniques have made it a valuable tool in various dental specialties, including orthodontics, implant dentistry, endodontics, and even forensic dentistry[16–18]. Several studies have demonstrated the effectiveness of manual analysis of the paranasal sinuses (such as frontal sinus and maxillary sinus) from CBCT images for sex determination, with a reported accuracy of 80.0% for manual identification[19,20].

In recent years, deep learning has been applied to medical image analysis tasks including image classification, detection, segmentation, denoising, and synthesis[21,22]. Several studies have reported methods based on deep learning for sex determination from CT or CBCT images. Bewes et al. proposed a sex determination method based on an artificial neural network on CT images[23]. This artificial neural network was trained on a dataset containing 900 skulls reconstructed from CT images and showed 95% accuracy for sex determination. Baban et al. reported a machine learning-based method that used morphometric measurements of the mandible on CBCT images as input[24]. This method showed 90% accuracy using Gaussian Naive Bayes. Senol et al. also proposed a machine learning-based method for sex determination using dental parameters of the maxillary molar and canine teeth obtained from CBCT images[25]. The method achieved 81% accuracy in sex determination with the ADA Boost Classifier algorithm. Although several studies have reported machine learning-based methods for sex determination, these approaches require several steps including manual segmentation of the skull, manual measurement of dental parameters, feature extraction, and classification; each of these steps is labor-intensive, time-consuming, and error-prone[26]. Moreover, existing methods face find it challenging to discriminate subtle differences in shapes and sizes of the frontal sinus on CBCT scans between males and females. Therefore, an automatic and accurate method for sex determination from CBCT scans is required. To the best of our knowledge, no previous study has performed fully automated segmentation and classification of the frontal sinuses for sex determination on CBCT scans using deep learning.

The purpose of this study was to automatically and accurately determine sex from a CBCT scan using a two-stage anatomy-guided attention network (SDetNet). Our main contributions are as follows: (1) The proposed SDetNet was designed to automatically and accurately segment the frontal sinus and predict the sex from a CBCT scan. The first stage deep learning model was a 2D frontal sinus segmentation network (FSNet) that segmented the frontal sinus on CBCT images and extracted regions of interest (ROIs) near it. The second stage was a 3D anatomy-guided attention network (SDNet) for accurate sex determination from a CBCT scan. (2) We introduced multi-channel inputs (MCI) and an anatomy-guided attention module (AGAM) to improve the performance of sex determination. The proposed MCI and AGAM encouraged SDetNet to learn the anatomical context of the frontal sinus for accurate and robust prediction of sex from a CBCT scan. In addition, we demonstrated the effectiveness of the AGAM and MCI by an experimental ablation study.

Materials and methods
Data acquisition and preparation
We collected a total of 310 CBCT scans acquired from 310 patients (mean age: 26.81 ± 11.36, 155 males and 155 females) who underwent Seoul National University Dental Hospital from 2020 to 2022. This study was performed with approval from the institutional review board of Seoul National University Dental Hospital (ERI123041). The ethics committee waived informed consent because this was a retrospective study. The study was performed following the Declarations of Helsinki. CBCT scans were acquired using a CS9300 (CS 9300, Carestream Health, Rochester, USA) with voxel sizes of $0.3 \times 0.3 \times 0.3$ mm^3, dimensions of $640 \times 670 \times 670$ pixels, and 16-bit depth under conditions of 80 or 90 kVp and 8 or 10 mA. All CBCT scans were anonymized and exported in DICOM format. The inclusion criterion was patients aged from 4 to 86 years (Supplementary Fig. S1), while exclusion criteria were patients with visible trauma, previous surgery, or pathological conditions in the frontal region of the skull.

Among the 310 CBCT scans, we split into 50 and 260 CBCT scans for the frontal sinus segmentation task and the sex determination task, respectively (Table 1). The 50 CBCT scans only used for frontal sinus segmentation were split into 30, 10, and 10 scans for training, validation, and test sets, respectively, and each set had the same sex distribution. Training, validation, and test sets comprised 19,200, 6400, and 6400 CBCT images, respectively. We observed a difference in volume (Supplementary Fig. S2a), length of the major axis (Supplementary Fig. S2b), and length of the minor axis (Supplementary Fig. S2c) between the frontal sinuses of males and females in our dataset. A region of interest (ROI) on a CBCT scan was cropped to $122 \times 128 \times 128$ pixels with the center at the frontal sinus region segmented by FSNet. The 260 CBCT scans only used for sex determination were split into 120, 40, and 100 CBCT scans for training, validation, and test sets, respectively, and each set had the same sex distribution. To generate the ground truth of segmentation masks (Fig. 1a,b), frontal sinus regions on CBCT images were labeled by a radiologist with over five years of experience using 3D Slicer software (www.slicer.org).

Dataset	Number of CBCT images	Number of patients	Number of males	Number of females	Mean age
Frontal sinus segmentation					
Training	19,200	30	15	15	30.2 ± 12.9
Validation	6,400	10	5	5	25.4 ± 2.8
Testing	6,400	10	5	5	32.0 ± 15.2
Sex determination					
Training	76,800	120	60	60	25.5 ± 10.6
Validation	25,600	40	20	20	28.2 ± 11.2
Testing	64,000	100	50	50	27.9 ± 11.5

Table 1. Data configuration for frontal sinus segmentation and sex determination tasks.

(a)　　　　　　　　　　　　　　　　**(b)**

Original image　　Ground truth　　3D visualization　　　　Original image　　Ground truth　　3D visualization
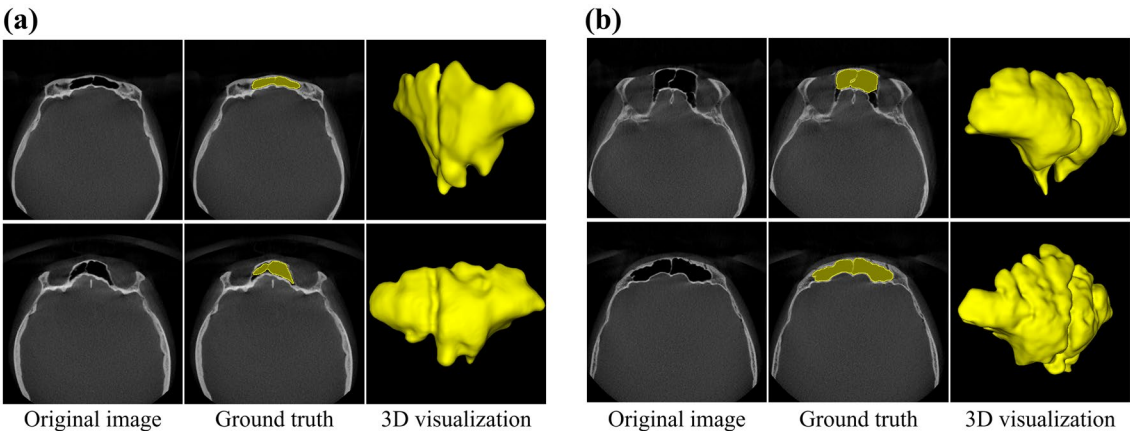
Figure 1. (**a**, **b**) CBCT images with label masks of the frontal sinus acquired from a female and male, respectively.

We estimated the minimum required sample size to detect significant differences in the accuracy of SDetNet and that of other networks when both assessed the same subjects (CBCT scans). We designed the study to capture a mean accuracy difference of 0.05 and a standard deviation of 0.10 between SDetNet and the other networks. Based on an effect size of 0.5, a significance level of 0.05, and a statistical power of 0.80, we calculated a required sample size of N = 128 (G* Power for Windows 10, Version 3.1.9.7; Universität Düsseldorf, Germany). Finally, we split the dataset of CBCT scans into 120, 40, and 100 scans for training, validation, and test sets, respectively.

The architecture of a two-stage anatomy-guided attention network

We proposed a two-stage anatomy-guided attention network (SDetNet) for automatic and accurate sex determination from a CBCT scan (Fig. 2). SDetNet consisted of a 2D frontal sinus segmentation network (FSNet) and a 3D anatomy-guided attention network (SDNet). The first stage was 2D frontal sinus segmentation using FSNet, which automatically segmented the frontal sinus regions on CBCT images. Next, 3D sex classification was performed using SDNet, which used the anatomy-guided information from the frontal sinus segmentation to automatically determine the sex of a patient on a CBCT scan. For frontal sinus segmentation on CBCT images, we used FSNet which had a U-shape encoder-decoder architecture with transfer learning. Five popular backbones, namely VGG16[27], ResNet101[28], DenseNet201[29], Inception V3[30], and EfficientNet-B5[31] were used as encoders in FSNet. The decoder part had five levels of layers with 2D convolution blocks and a 2D transposed convolution layer for 2D up-sampling. The 2D convolution block consisted of a 3 × 3 convolution layer, batch normalization (BN), and rectified linear unit (ReLU) activation. The final output layer in FSNet was a 1 × 1 convolution layer with a Sigmoid activation function.

After automatic segmentation of the frontal sinus on CBCT images by FSNet, the CBCT scan with corresponding prediction masks cropped at the centroid of the segmentation results of the frontal sinus were used as multi-channel inputs of the SDNet designed for automatic sex determination (Fig. 2a). SDNet had 3D convolutional blocks (ConvBlocks), an anatomy-guided attention module (AGAM), 3D max-pooling (MP), and 3D global average pooling (GAP). The ConvBlock consisted of a 3 × 3 × 3 convolution layer, BN, and ReLU. The MP was used for the down-sampling of 3D feature maps. We employed 3D GAP to average each 3D feature map. Final feature vectors by a 3D GAP were fed into the output layer with the Sigmoid activation function for sex prediction. The feature maps at each level of layers were gradually increased from 16 to 32, 64, and 128 in SDNet.

For accurate sex determination from a CBCT scan, deep learning models need to capture anatomical context information related to variations in the shape and size of the frontal sinuses between males and females (Supplementary Fig. S2a–c). Attention mechanisms in deep learning are inspired by the human visual cognition system,
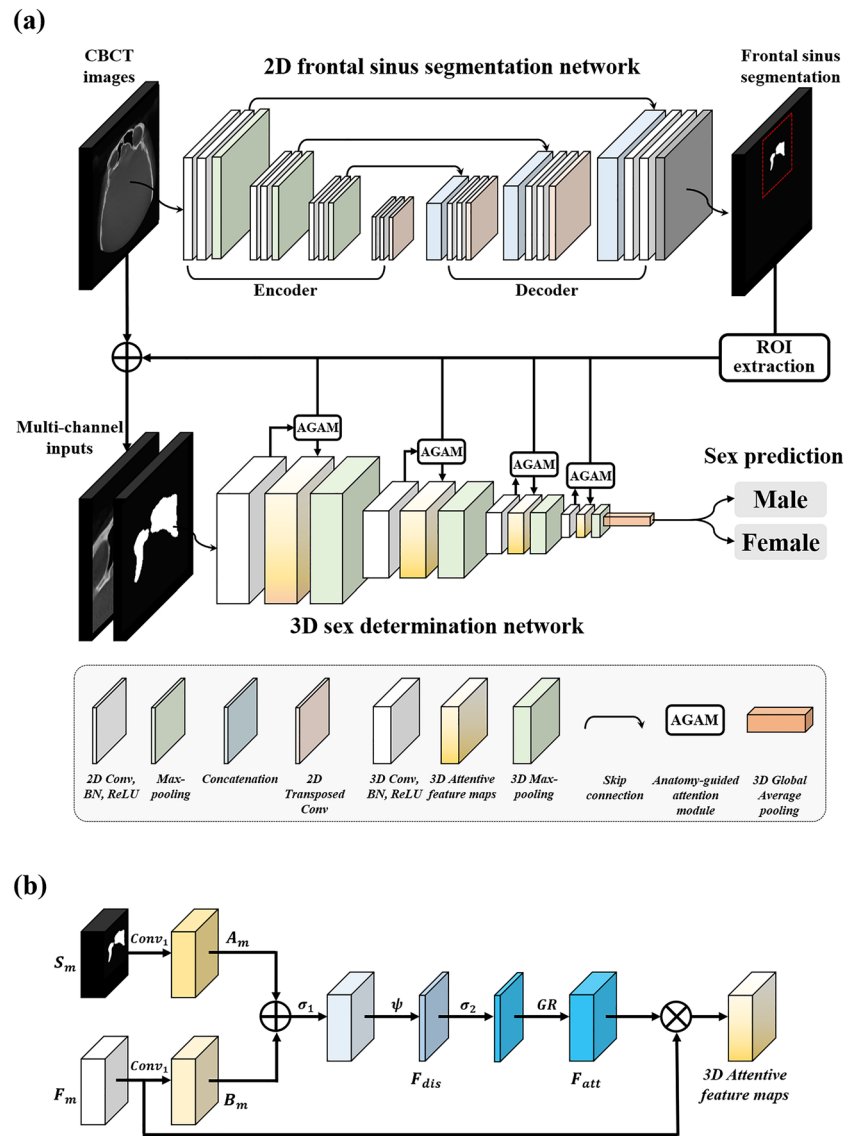
Figure 2. Overview of the proposed two-stage anatomy-guided attention network (SDetNet). (**a**) SDetNet consists of the 2D frontal sinus segmentation network (FSNet) and 3D sex determination network (SDNet). (**b**) The architecture of anatomy-guided attention module (AGAM).

and these are used to encourage deep learning models to focus more on the most relevant regions and ignore the background[32]. Based on this observation, we proposed AGAM in SDNet to encourage the deep learning model to focus more on the frontal sinus regions on a CBCT scan than on the background regions (Fig. 2b). AGAM was embedded in each level of the layer of SDNet to learn the anatomical context of the frontal sinus hierarchically. Frontal sinus segmentation maps $S_m \in R^{H \times W \times D \times 1}$ are obtained by the ROI extraction process after the inference of FDNet, where H, W, and D indicate the height, width, and depth of the mask maps, respectively. The 3D feature maps $F_m \in R^{H \times W \times D \times C}$ are acquired by a ConvBlock, where C indicates the number of channels in the feature maps. Then, we employ a $1 \times 1 \times 1$ convolution layer $Conv_1$ for each F_m and to obtain anatomy-guided feature maps $A_m \in R^{H \times W \times D \times C}$ and bottleneck feature maps $B_m \in R^{H \times W \times D \times C}$ as follows:

$$A_m = Conv_1(S_m), B_m = Conv_1(F_m) \tag{1}$$

To extract discriminative features F_{dis} between A_m and B_m, 3D attention maps ($F_{att} \in R^{H \times W \times D \times C}$) are acquired as follows:

$$F_{dis} = \psi(\sigma_1(A_m + B_m)) \tag{2}$$

$$F_{att} = GR(\sigma_2(F_{dis})) \tag{3}$$

where ψ is a $1 \times 1 \times 1$ convolution layer to extract the discriminative feature map $F_{dis} \in R^{H \times W \times D \times 1}$ and σ_1 and σ_2 are ReLU and Sigmoid activation functions, respectively. GR denotes a grid resampling operation to restore the dimensions of the discriminative feature map to the same as that of F_m using trilinear interpolation. Finally, 3D attentive feature maps F_n are acquired by elemental-wise multiplying F_m and F_{att} as follows:

$$F_n = F_{att} \otimes F_m \tag{4}$$

where $\otimes$ indicates elemental-wise multiplying. $F_{att} \in [0, 1]$, which are saliency maps, identified important regions in the feature maps and pruned the feature response to retain the activations relevant to the foreground, suppressing the background.

We used the Dice similarity coefficient (DL) and binary cross-entropy (BL) losses to train FSNet and SDNet, respectively. DL measured the overlap between the ground truth and segmentation results for the frontal sinus. DL is defined as:

$$DL(y, \widehat{y}) = 1 - 2 \frac{\sum_i^n (y_i \times \widehat{y}_i) + \epsilon}{(\sum_i^n y_i + \sum_i^n \widehat{y}_i) + \epsilon} \tag{5}$$

where y and $\widehat{y}$ are ground truth and segmentation results for the frontal sinus, respectively, and n is the number of pixels on CBCT images. ϵ provided numerical stability to prevent division by zero, with ϵ set to 10^{-3}. BL measured the average probability error between the ground truth (actual sex) and the sex predictions. BL is defined as:

$$BL(p, \widehat{p}) = 1 - \sum_i^N (p_i log \widehat{p}_i) \tag{6}$$

where p and $\widehat{p}$ are the ground truth and probability of sex prediction, respectively. N is the number of CBCT scans.

FSNet was trained for 200 epochs with a mini-batch size of 16. Data augmentation was performed with rotation ($-10°$ to $10°$), Gaussian blur (-10% to 10%), and brightness ($-10°$ to $10°$). Adam optimizer with a learning rate of 10^{-3} was used as the initial setting, and the learning rate was reduced by half up to 10^{-6} when the validation loss saturated for 20 epochs. SDetNet was trained for 100 epochs with a mini-batch size of 1. Adam optimizer was used with $\beta_0 = 0.9$ and $\beta_1 = 0.999$, and the learning rate was initially set to 10^{-4}, which was reduced by half up to 10^{-7} when the validation loss saturated for 25 epochs. Deep learning models were implemented with Python3 and Keras with a TensorFlow backend based on an Intel i9-7900X CPU 3.3 GHz, 256 RAM, and an NVIDIA RTX A6000 GPU 48 GB.

Performance evaluation

We used precision (PR), recall (RC), Jaccard index (JI), and F1-score (F1) to evaluate the segmentation performance of deep learning models for the frontal sinus, and area under the receiver operating characteristic curve (AUC), Brier score (BR), accuracy (ACC), specificity (SPE), sensitivity (SEN), and the polygon area metric (PAM) to evaluate its performance for sex determination. PR is calculated as the number of true positives (TP) divided by the sum of the TP and false positives (FP): $PR = \frac{TP}{TP+FP}$. RC is calculated as the number of TPs divided by the sum of the TPs and false negatives (FNs) as follows: $RC = \frac{TP}{TP+FN}$. JI is calculated as the intersection of the predicted segmentation and ground truth divided by the union of the two: $JI = \frac{TP}{TF+FP+FN}$. F1 is calculated as the harmonic mean of the PR and RC: $F1 = \frac{2 \times PR \times RC}{PR+RC}$. ACC is defined as the ratio of the number of correct sex predictions to the total number of input samples as follows: $ACC = \frac{TP+TN}{TF+TN+FP+FN}$, where TN indicates true negatives. SPE is a metric that measures a model's ability to predict negative cases correctly and is defined as $SPE = \frac{TN}{TN+FP}$. SEN, similar to RC, is a metric that measures a model's ability to correctly predict positive cases. BR is calculated as the mean squared difference between the predicted probabilities and the actual outcomes: $BR = \frac{1}{N} \sum_i^N (y_i - p_i)^2$, where N is the number of CBCT scans and y_i and p_i are the ground truth and prediction probability, respectively. AUC is calculated as the area under the receiver operating characteristic (ROC) curve, which is a plot of the true positive rate versus the false positive rate. PAM is calculated using the area of the polygon including ACC, SEN, SPE, AUC, Jaccard index (JI), and F-measure (FM) points generated in a regular hexagon[33]. The PAM is defined as:

$$PAM = \frac{PA}{2.59807} \tag{7}$$

where, the PA denotes the area of the polygon. To normalize the PAM into the [0, 1], the PA is divided by 2.59807. In terms of sex determination results, SDetNet outputs a probability within the range of 0.0 to 1.0, where females and males are classified based on the 0.5 threshold as 0 and 1, respectively. Therefore, SPE reflected the ability of the deep learning algorithm to correctly predict females and SEN the ability of the algorithm to correctly predict males. An analysis of variance (one-way ANOVA) with Scheffé post hoc tests was performed using IBM SPSS Statistics (IBM SPSS Statistics for Windows 10, Version 26.0; IBM, Armonk, New York, USA), and statistical significance (p-value) was set to 0.05.

Ethics declarations

This study was performed with approval from the Institutional Review Board (IRB) of Seoul National University Dental Hospital (ERI123041). The IRB of Seoul National University Dental Hospital approved the waiver

for informed consent because this was a retrospective study. The study was performed in accordance with the Declaration of Helsinki.

Results

We compared the performances of the VGG16, ResNet101, Inception V3, EfficientNet-B5, and DenseNet201 backbones in FSNet for frontal sinus segmentation using JI, F1, PR, and RC. After frontal sinus segmentation, the sex determination performance of SDNet was compared with that of 3D ResNet, 3D DenseNet, 3D MobileNet, and 3D EfficientNet-B0 using ACU, BR, ACC, SPE, SEN, and PAM. To ensure a fair comparison, all deep learning models were run in the same computational environment.

Backbones	JI	F1	PR	RC
VGG16	0.850 ± 0.064	0.917 ± 0.038	0.937 ± 0.043	0.902 ± 0.059
Inception V3	0.865 ± 0.059	0.927 ± 0.035	0.945 ± 0.031	0.912 ± 0.067
ResNet101	0.868 ± 0.058	0.928 ± 0.035	0.955 ± 0.018	0.906 ± 0.067
EfficientNet-B5	0.868 ± 0.048	0.929 ± 0.028	0.946 ± 0.022	0.914 ± 0.056
DenseNet201	0.878 ± 0.042	0.935 ± 0.024	0.941 ± 0.035	0.930 ± 0.038

Table 2. Comparison of the segmentation performance of different backbones in FSNet. Segmentation performance is presented as mean ± standard deviation.
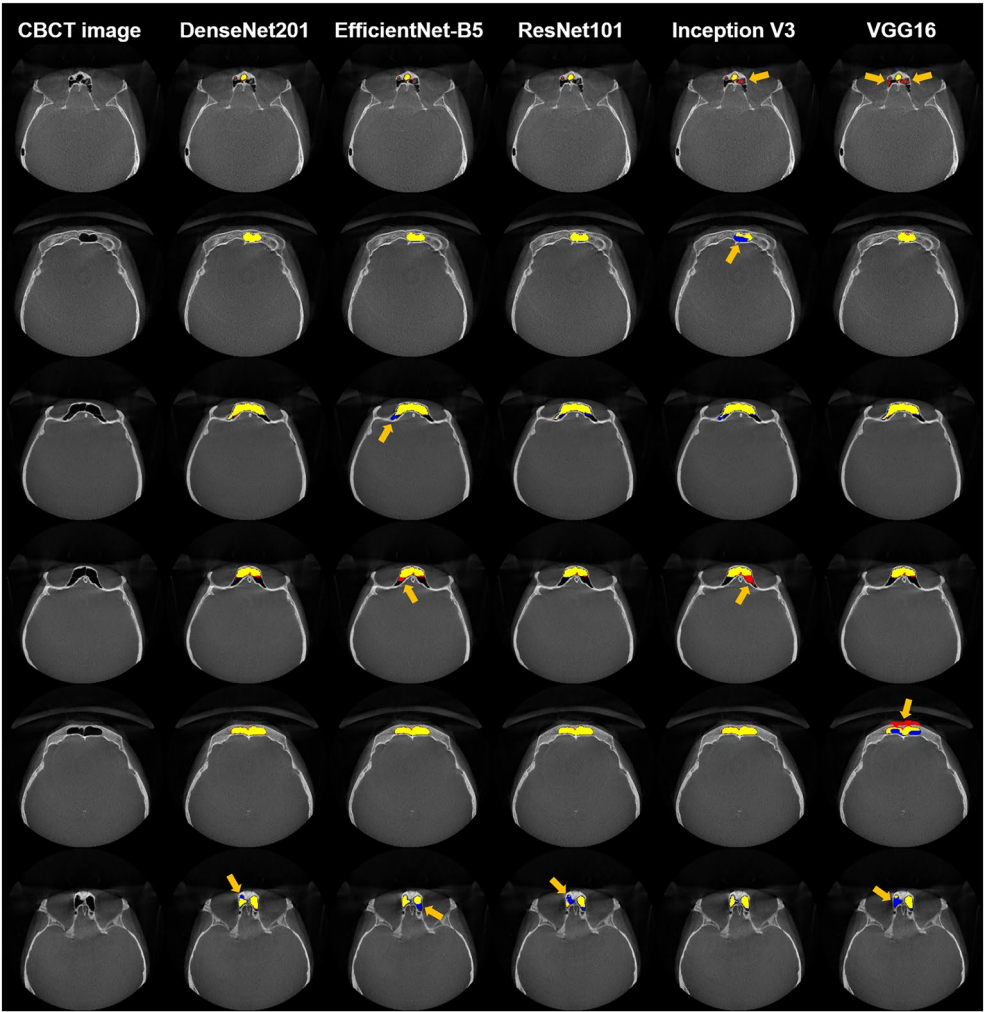

Figure 3. 2D segmentation results from the different backbones of DenseNet201, EfficientNet-B5, ResNet101, Inception V3, and VGG16 in FSNet. Yellow, blue, and red areas present true positives, false positives, and false negatives for frontal sinus segmentation, respectively. The orange arrow indicates segmentation errors.

As shown in Table 2, all backbones achieved high segmentation performance of F1 values of over 0.900. for frontal sinuses on CBCT images. DenseNet201 achieved the highest JI, F1, and RC values of 0.878 ± 0.042, 0.935 ± 0.024, and 0.930 ± 0.038, respectively, indicating its superior segmentation performance for frontal sinuses. Representative segmentation results of the frontal sinus from different backbones are shown in Fig. 3. DenseNet201 exhibited the most accurate frontal sinus segmentation with more true positives (yellow), fewer false positives (red), and fewer false negatives (blue) than the other backbones. In the 3D reconstruction of frontal sinus segmentation from different backbones, DenseNet201 exhibited fewer false negatives (blue circles) and false positives (red circles) than the other backbones, as shown in Fig. 4. Boxplots of the segmentation performance of the frontal sinuses from different backbones are shown in Fig. 5. ROC and BR curves for the sex determination performance of SDetNet according to segmentation results generated by different backbones in FSNet are shown in Fig. 9a and d, respectively.

From the quantitative results of the sex determination according to segmentation results from different backbones in FSNet, SDetNet with ROIs extracted by DenseNet201 achieved superior AUC, ACC, BR, SPE, SEN, and PAM values of 0.979, 0.920, 0.063, 0.960, 0.880, and 0.828, respectively, for sex determination (Table 3). Polygon area graphs are shown in Supplementary Fig. S3. Although SDetNet using ROIs extracted by the other backbones in FSNet showed comparable performance in sex determination, it achieved slightly lower AUC and ACC values than SDetNet using DenseNet201 (Table 3). Confusion matrices for the sex determination performance of SDetNet according to segmentation results generated by different backbones in FSNet are shown in Fig. 6.

The sex determination performance of SDetNet was compared quantitatively with those of the other 3D CNNs (Table 4), with the ROI of the frontal sinus extracted using DenseNet201. SDetNet outperformed the other 3D CNNs by obtaining the highest AUC, ACC, BR, SPE, and PAM values of 0.979, 0.920, 0.063, 0.960, and 0.828, respectively. Compared with the second-highest performing backbone, the AUC, ACC, BR, and PAM values of SDetNet were enhanced by 0.002, 0.020, 0.032, and 0.065 better, respectively. Polygon area graphs are shown in Supplementary Fig. S4. The confusion matrix for the sex determination performance of the different 3D CNNs is shown in Fig. 7. The ROC and BR curves for the sex determination performance of the different 3D CNNs according to segmentation results generated by DenseNet201 are shown in Fig. 9b and e, respectively.

Ablation studies were performed to demonstrate the effectiveness of MSI and AGAM in SDetNet (Table 5). For sex determination, SDetNet without MSI (without mask images) and AGAM obtained lower ACC, BR, SEN, and PAM values of 0.770, 0.155, 0.540, and PAM, respectively, than SDetNet using both CBCT scans and mask images. In addition, sex determination performance was further improved by embedding AGAM in SDetNet as evidenced by AUC, ACC, BR, SPE, and PAM values of 0.964, 0.920, 0.098, 0.900, and 0.785 to 0.979, 0.920, 0.063, 0.960, and 0.828, respectively. Polygon area graphs are shown in Supplementary Fig. S5. Figure 8 shows the confusion matrices for the sex determination performance of each component in SDetNet. The ROC and BR curves for the sex determination performance of each component in SDetNet are shown in Fig. 9c and f, respectively.
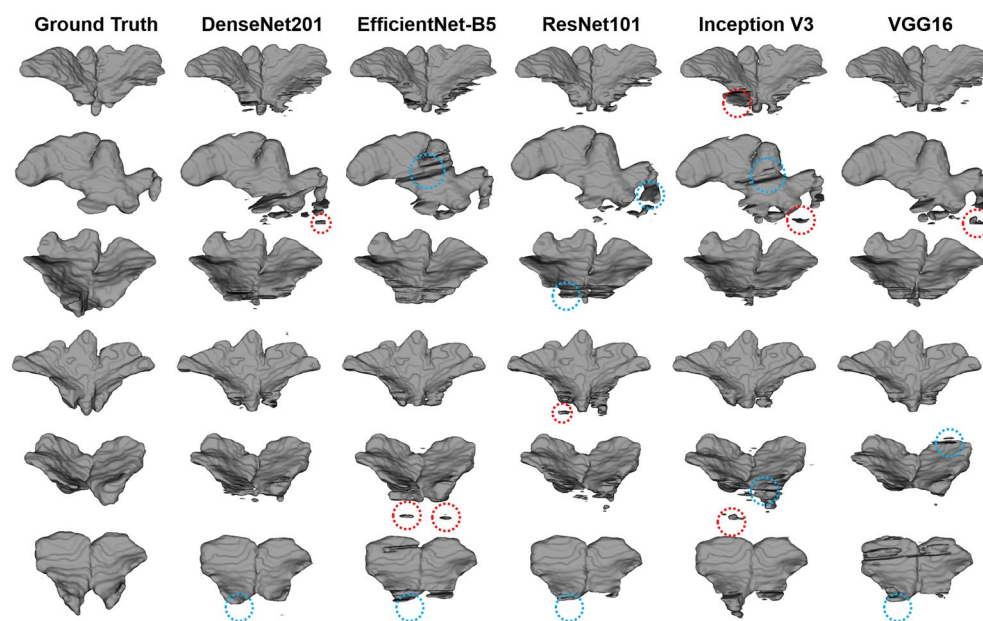

Figure 4. 3D reconstruction of the segmentation results of the frontal sinus from different backbones including DenseNet201, EfficientNet-B5, ResNet101, Inception V3, and VGG16 in FSNet. DenseNet201 shows fewer false negatives (blue dot circles) and false positives (red dot circles) than the other backbones.
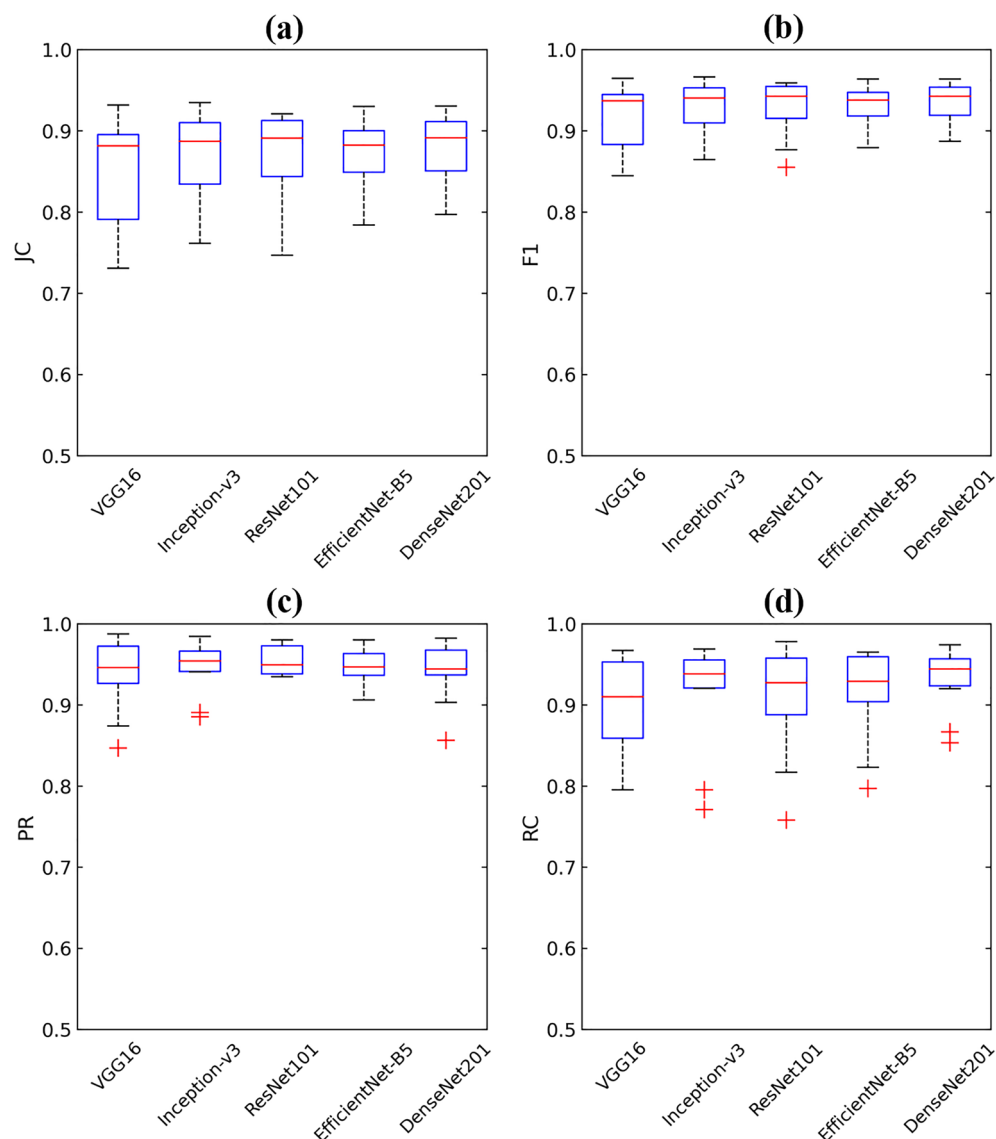
Figure 5. Boxplots of the segmentation performance of the frontal sinus from different backbones in FSNet. Each boxplot contains the first and third quartiles of data. Medians are located inside the boxes and are represented as red lines. Whiskers that extend above and below each box are ± 1.5 times the interquartile range (IQR), and outliers are indicated as red crosses (IQR values 1.5 or greater away from the box).

Backbones	AUC	ACC	BR	SPE	SEN	PAM
VGG16	0.946	0.860	0.098	0.780	0.940	0.716
Inception V3	0.968	0.850	0.093	0.940	0.760	0.680
ResNet101	0.981	0.890	0.063	0.920	0.860	0.718
EfficientNet-B5*	0.969	0.860	0.105	0.720	1.000	0.763
DenseNet201	0.979	0.920	0.063	0.960	0.880	0.828

Table 3. Sex determination performance of SDetNet according to segmentation results generated by different backbones in FSNet. *Significant difference for predicted probability between ResNet101 and EfficientNet-B5 (p-value < 0.05).
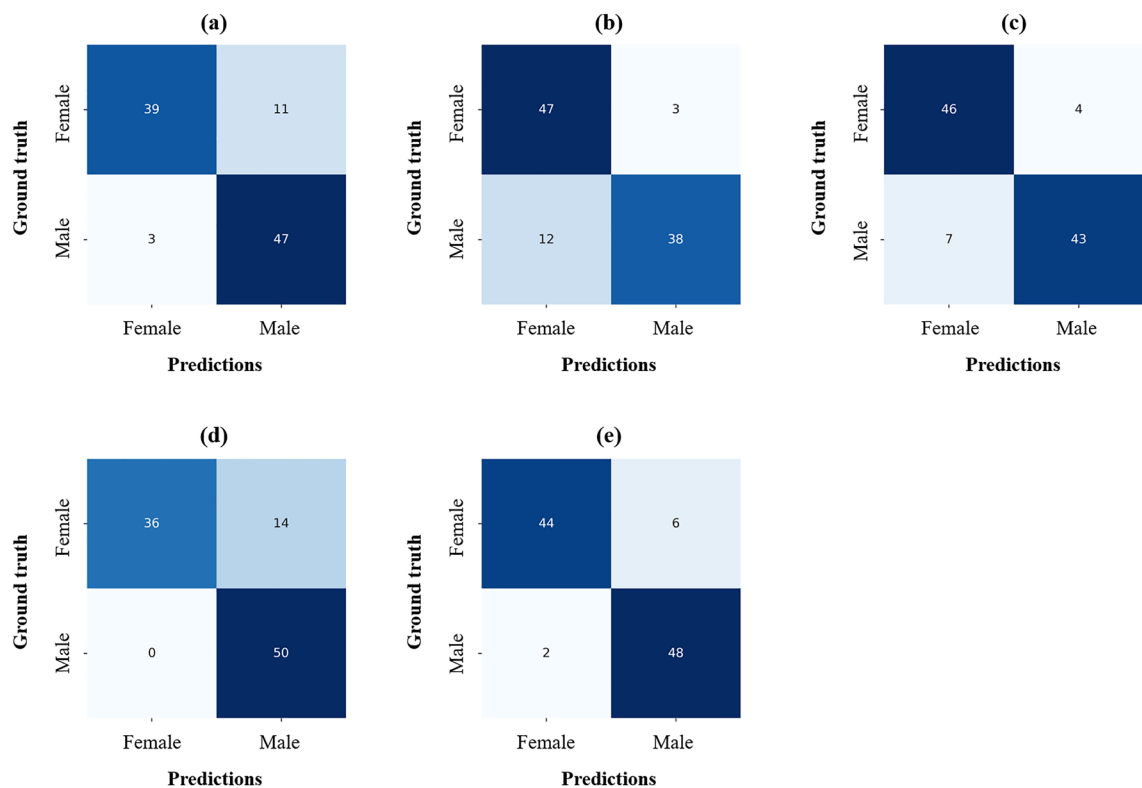
Figure 6. Confusion matrices for sex determination performance of SDetNet according to segmentation results generated by the different backbones in FSNet. (**a–e**) Results for VGG16, Inception V3, ResNet101, EfficientNet-B5, and DenseNet201, respectively.

Backbones*	AUC	ACC	BR	SPE	SEN	PAM
Simple 3D CNN	0.964	0.900	0.098	0.900	0.900	0.785
3D InceptionNet	0.950	0.870	0.105	0.880	0.860	0.727
3D DenseNet	0.977	0.870	0.091	0.960	0.780	0.717
3D ResNet	0.934	0.860	0.121	0.900	0.820	0.705
SDetNet	0.979	0.920	0.063	0.960	0.880	0.828

Table 4. Sex determination performance of different 3D CNNs according to segmentation results generated by DenseNet201. *No significant difference for predicted probability between backbones.

Discussion

Sex determination based on skeletal remains is essential in forensic investigations and human identification after mass disasters, homicides, and accidents[1,2]. Manual morphological measurement and analysis of the skeletal remains are widely used to determine the sex of individuals[4–9,26]. In particular, the unique patterns of the frontal sinuses in males and females on CBCT scans make these sinuses an important and reliable tool for sex determination[16,20,34]. Recently, deep learning-based methods have been applied for forensic investigation to predict sex from CT or CBCT scans[20,24,25,34]. However, previous methods required several steps including manual segmentation of the skull, manual measurement of dental parameters, feature extraction, and classification, which are all labor-intensive, time-consuming, and error-prone steps. In this study, we proposed SDetNet to automatically and accurately determine sex from a CBCT scan by capturing subtle differences in shapes and sizes of the frontal sinus on CBCT scans between males and females.

The segmentation performance of backbones such as VGG16, ResNet101, Inception V3, EfficientNet-B5, and DenseNet201 in FSNet was compared. DenseNet201 outperformed the other backbones for frontal sinus segmentation on CBCT images (Table 2), with much higher RC values than the other backbones. As shown in Figs. 3 and 4, segmentation models exhibited false positives of segmentation of the frontal sinus with invading ethmoid cells. The ethmoid sinuses are the only paranasal sinuses not formed by a single cavity, making them more complex. The anterior cranial fossa and the frontal bone limit the ethmoid cells superiorly, and thus, ambiguous structures between the ethmoid cells and the frontal sinus yield false positives. Additionally, inflammation of the frontal sinus can manifest as a thickening of the mucous membrane. Mucosal thickening may obscure clear borders, making automatic segmentation of the frontal sinus more challenging on CBCT
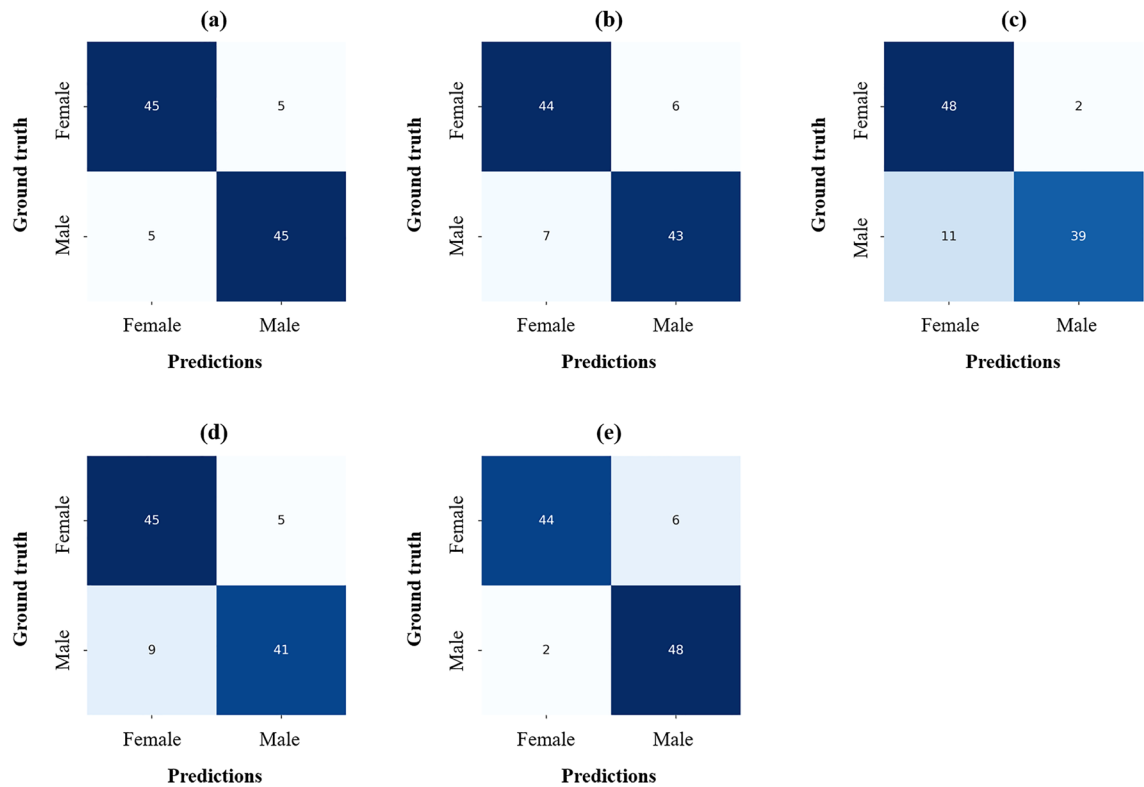
Figure 7. Confusion matrices for sex determination performance of different 3D CNNs according to segmentation results generated by DenseNet201. (**a**–**e**) Results of Simple 3D CNN, 3D InceptionNet, 3D DenseNet, 3D ResNet, and SDetNet, respectively.

CBCT scan	Mask images	AGAM	AUC	ACC	BR	SPE	SEN	PAM
✓			0.963	0.770	0.155	1.000	0.540	0.510
	✓		0.956	0.850	0.093	0.880	0.820	0.688
✓*	✓		0.964	0.900	0.098	0.900	0.900	0.785
✓†	✓	✓	0.979	0.920	0.063	0.960	0.880	0.828

Table 5. Ablation experimental results of each component of SDetNet on the test dataset. *Significant difference for predicted probability between only used CBCT scan and CBCT scan + Mask images (p-value < 0.05); †Significant difference for predicted probability between only used CBCT scan and CBCT scan + Mask images + AGAM (p-value < 0.05).

images. Compared with backbones in FSNet, SDetNet using segmentation masks generated by DenseNet201 achieved superior sex determination performance (Table 3). The sex determination performance of SDetNet was affected by the segmentation quality of the frontal sinus on CBCT images.

The rationale for choosing the five backbones can be summarized as follows: (1) The selected backbones have been extensively validated on benchmark datasets and have shown remarkable performance in various tasks, including medical image classification, object detection, and segmentation[35,36]. Their proven performance provides a strong foundation for FSNet, potentially enhancing its segmentation performance and reliability. (2) Each backbone has different architectural designs and principles that can affect the prediction performance. VGG16 increases the depth of an architecture using 3×3 convolution layers. ResNet101 is designed based on residual learning that facilitates deep learning without degradation. DenseNet201 adopts the reuse of feature maps to enhance information flow. Inception V3 uses multi-scale convolution layers that can capture information at various resolutions. EfficientNet leverages a compound scaling method, balancing the depth, width, and resolution of convolution layers, showcasing parameter usage and computational performance efficiency.

Compared with different 3D CNNs, SDetNet achieved the highest sex determination performance (Table 4). SDetNet with DenseNet201 achieved a superior performance owing to three key factors. First, ROIs including frontal sinus regions on CBCT images were automatically extracted by FSNet and used as the input volume in SDetNet. Using ROIs as input volume can help a deep learning model focus on the frontal sinus regions on a CBCT scan and determine sex well, without having to consider larger anatomical structures. Second, DenseNet201 in FSNet alleviated the vanishing gradient problem by connecting each convolutional layer to every other layer. This allowed the deep learning model to learn more complex features and improved its segmentation
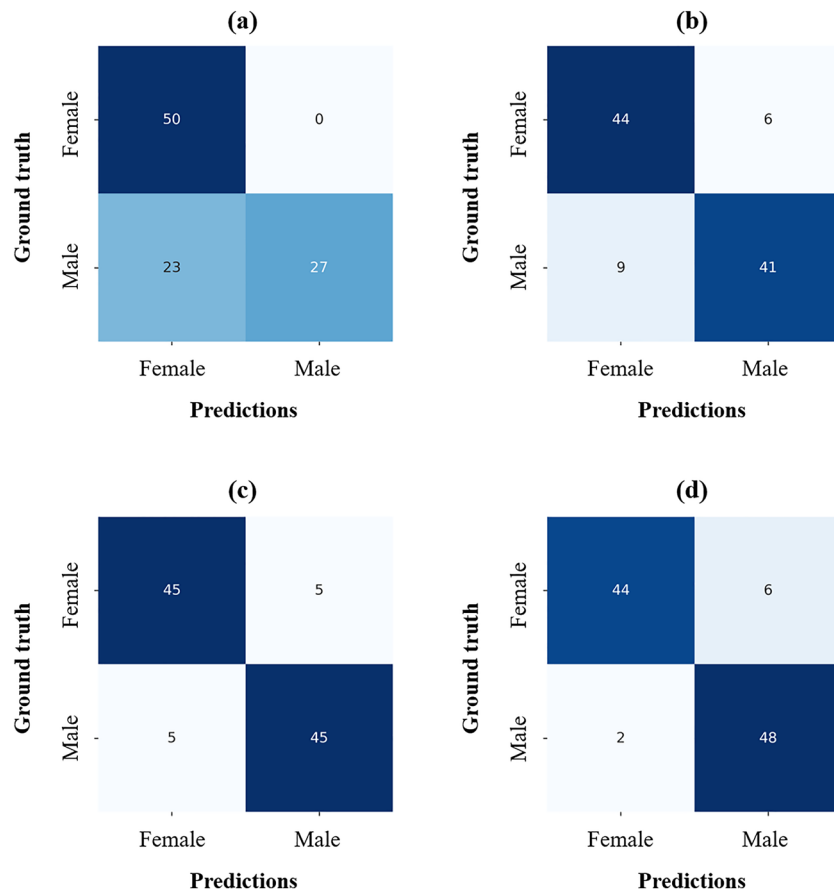
Figure 8. Confusion matrices for sex determination performance of each component in SDetNet. (**a**–**d**) Results of CBCT scan, Mask images, CBCT scan + Mask images, and CBCT scan + Mask images + AGAM, respectively.

performance. Finally, AGAM embedded in SDetNet was designed to focus on anatomical features of the frontal sinus for sex determination from a CBCT scan. The proposed AGAM and MSI improved the sex determination performance, and their effectiveness was demonstrated by an ablation study (Table 5). The primary reason for this improvement was that the shape and size of the frontal sinuses differ slightly between males and females, and SDetNet could learn the anatomical context information about subtle differences in shape and size of the frontal sinuses between males and females using AGAM and MSI.

Previous studies have used 2D slices from multiplanar reconstructions to measure the volume of the frontal sinus cavity[11]. However, this approach can be challenging due to the high variability in size, shape, and asymmetry of the cavity. Other studies segmented and reconstructed the frontal sinuses in 3D, and performed calculations on the reconstructed volume after exporting different views (frontal, lateral, basal)[19,20]. This approach avoids the loss of information and allows for more accurate measurements. Nevertheless, the accuracy of previous studies at predicting sex correctly based on analysis of images of the frontal sinuses ranged between 60 and 80%[13,19,20,34,37]. Our proposed SDetNet is a fully automatic and accurate sex determination method that consists of FSNet and SDNet, achieving an AUC, ACC, BR, and PAM of 0.979, 0.920, 0.063, and 0.828. SDetNet does not require any additional processes such as manual segmentation or analysis, dental parameters, or feature selection.

The following issues will be addressed in future studies to improve the sex determination performance of SDetNet. First, our dataset was built using CBCT scans from a patient group with a non-uniform age distribution. To improve the accuracy of our SDetNet for sex determination at all ages, we need to collect additional datasets with a uniform age distribution. Second, it would be valuable to assess the model's performance on a more diverse and larger dataset to validate its generalizability. This study relied on a CBCT dataset from a single organization in South Korea, which may not be generalizable to other populations or organizations. Therefore, further research is needed to train and evaluate SDetNet using CBCT datasets acquired from individuals of diverse ethnicities using various devices at multiple organizations. Finally, we applied several exclusion criteria when selecting CBCT scans. In future studies, we plan to improve the generalizability and clinical efficacy of SDetNet using large-scale panoramic radiographs from individuals of all ages with fewer exclusion criteria.

Conclusions

In this study, we proposed SDetNet for automatic and accurate sex determination from a CBCT scan. SDetNet was designed as a two-stage network to learn the anatomical context information of the frontal sinuses between males and females by embedding MSI and AGAM in an end-to-end manner. The experimental results showed
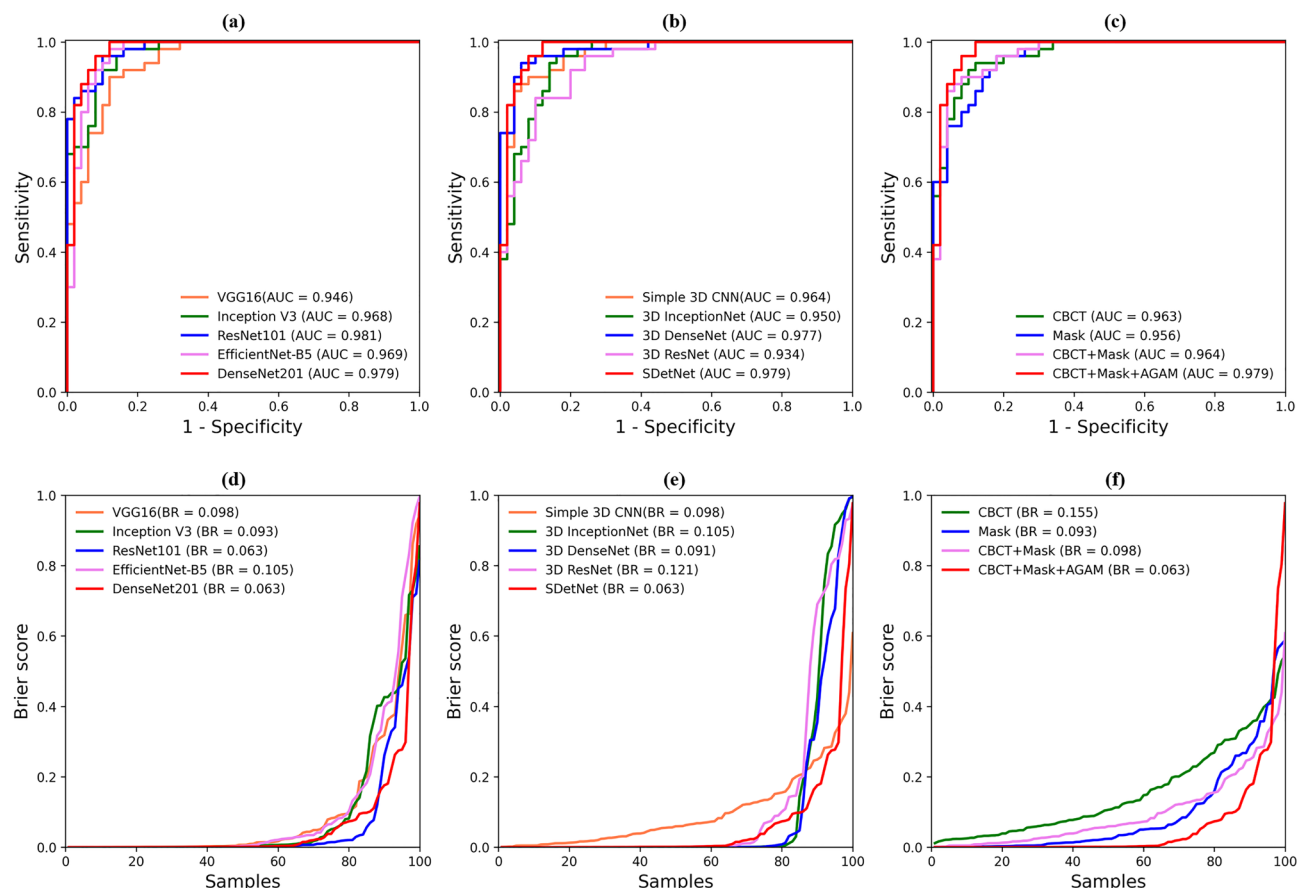
Figure 9. Receiver operating characteristic (ROC) and Brier score (BR) curves for sex determination performance. (**a**, **d**) are ROC and BR curves showing the sex determination performance of SDetNet according to segmentation results generated by different backbones in FSNet, respectively. (**b**, **e**) are ROC and BR curves for the sex determination performance of different 3D CNNs according to segmentation results generated by DenseNet201, respectively. (**c**, **f**) are ROC and BR curves for the sex determination performance of each component of SDetNet.

the SDetNet outperformed existing 3D CNNs for sex determination from CBCT scans. Furthermore, we demonstrated the effectiveness of MSI and AGAM of SDetNet by an ablation study, which substantially improved sex determination from CBCT scans. SDetNet is a fully automatic and accurate sex determination method, that will likely improve the workflow of forensic investigations and individual identification in clinical settings. In future studies, we plan to improve the generalizability and clinical efficacy of SDetNet by using CBCT scans of the frontal sinuses of individuals of varied ethnicities from diverse populations collected by multiple organizations using various devices.

Data availability

The datasets generated and/or analyzed during the current study are not publicly available due to restrictions of the Institutional Review Board (IRB) of Seoul National University Dental Hospital to protect patients' privacy but are available from the corresponding author on reasonable request.

References

1. Cameriere, R., Ferrante, L., Molleson, T. & Brown, B. Frontal sinus accuracy in identification as measured by false positives in kin groups. *J. Forensic Sci.* **53**, 1280–1282 (2008).
2. Nagare, S. P., Chaudhari, R. S., Birangane, R. S. & Parkarwar, P. C. Sex determination in forensic identification: A review. *J. Forensic Dent. Sci.* **10**, 61 (2018).
3. Kozerska, M., Skrzat, J. & Szczepanek, A. Application of the temporal bone for sex determination from the skeletal remains. *Folia Med. Cracov.* (2015).
4. Koot, M. G., Sauer, N. J. & Fenton, T. W. Radiographic human identification using bones of the hand: A validation study. *J. Forensic Sci.* **50**, 2004229 (2005).
5. Rajasekhar, S., Vasudha, T. & Aravindhan, K. Sex determination by biometry of anterior features of human hip bones in South Indian population. *J. Clin. Diagn. Res.* **11**, 13 (2017).

6. Sharma, N. & Dhillon, S. Identification through dental age estimation in skeletal remains of a child. *J. Forensic Dent. Sci.* **11**, 48 (2019).
7. Vinay, G. & Anbalagan, J. Sex determination of human mandible using metrical parameters. *J. Clin. Diagn. Res.* **7**, 2671 (2013).
8. Quatrehomme, G. *et al.* Identification by frontal sinus pattern in forensic anthropology. *Forensic Sci. Int.* **83**, 147–153 (1996).
9. Tuzoff, D. V. *et al.* Tooth detection and numbering in panoramic radiographs using convolutional neural networks. *Dentomaxillofac. Radiol.* **48**, 20180051 (2019).
10. Verma, S., Mahima, V. & Patil, K. Radiomorphometric analysis of frontal sinus for sex determination. *J. Forensic Dent. Sci.* **6**, 177 (2014).
11. Tatlisumak, E. *et al.* Identification of unknown bodies by using CT images of frontal sinus. *Forensic Sci. Int.* **166**, 42–48 (2007).
12. Uthman, A. T., Al-Rawi, N. H., Al-Naaimi, A. S., Tawfeeq, A. S. & Suhail, E. H. Evaluation of frontal sinus and skull measurements using spiral CT scanning: An aid in unknown person identification. *Forensic Sci. Int.* **197**(124), e121-124 (2010).
13. Goyal, M., Acharya, A. B., Sattur, A. P. & Naikmasur, V. G. Are frontal sinuses useful indicators of sex?. *J. Forensic Leg. Med.* **20**, 91–94 (2013).
14. Pereira, J. G. D., Santos, J. B. S., Sousa, S. P. D., Franco, A. & Silva, R. H. A. Frontal sinuses as tools for human identification: A systematic review of imaging methods. *Dentomaxillofac. Radiol.* **50**, 20200599 (2021).
15. Yoshino, M., Miyasaka, S., Sato, H. & Seta, S. Classification system of frontal sinus patterns by radiography: Its application to identification of unknown skeletal remains. *Forensic Sci. Int.* **34**, 289–299 (1987).
16. Issrani, R. *et al.* Cone-beam computed tomography: A new tool on the horizon for forensic dentistry. *Int. J. Environ. Res. Public Health.* **19**, 5352 (2022).
17. Jaju, P. P. & Jaju, S. P. Clinical utility of dental cone-beam computed tomography: Current perspectives. *Clin. Cosmet. Investig. Dent.* **1**, 29–43 (2014).
18. American Dental Association Council on Scientific Affairs. The use of cone-beam computed tomography in dentistry: An advisory statement from the American Dental Association Council on Scientific Affairs. *J. Am. Dent. Assoc.* **143**, 899–902 (2012).
19. Luz, J. *et al.* 3D-evaluation of the maxillary sinus in cone-beam computed tomography. *Int. J. Implant. Dent.* **4**, 1–7 (2018).
20. Choi, I. G., Duailibi-Neto, E. F., Beaini, T. L., da Silva, R. L. & Chilvarquer, I. The frontal sinus cavity exhibits sexual dimorphism in 3D cone-beam CT images and can be used for sex determination. *J. Forensic Sci.* **63**, 692–698 (2018).
21. Kang, S.-R. *et al.* Structure-preserving quality improvement of cone beam CT images using contrastive learning. *Comput. Biol. Med.* **158**, 106803 (2023).
22. Jeoun, B.-S. *et al.* Canal-Net for automatic and robust 3D segmentation of mandibular canals in CBCT images using a continuity-aware contextual network. *Sci. Rep.* **12**, 13460 (2022).
23. Bewes, J., Low, A., Morphett, A., Pate, F. D. & Henneberg, M. Artificial intelligence for sex determination of skeletal remains: Application of a deep learning artificial neural network to human skulls. *J. Forensic Leg. Med.* **62**, 40–43 (2019).
24. Baban, M. T. A. & Mohammad, D. N. The accuracy of sex identification using CBCT morphometric measurements of the mandible, with different machine-learning algorithms: A retrospective study. *Diagnostics* **13**, 2342 (2023).
25. Senol, D., Secgin, Y., Duman, B. S., Toy, S. & Oner, Z. Sex and age estimation with machine learning algorithms with parameters obtained from cone beam computed tomography images of maxillary first molar and canine teeth. *Egypt J. Forensic Sci.* **13**, 1–9 (2023).
26. Capitaneanu, C., Willems, G. & Thevissen, P. A systematic review of odontological sex estimation methods. *J. Forensic Odontostomatol.* **35**, 1 (2017).
27. Simonyan, K. & Zisserman, A. Very deep convolutional networks for large-scale image recognition. arXiv:1409.1556 (2014).
28. Eliades, A., Papadeli, C. & Tsirlis, A. Mandibular canal, foramina of the mandible and their variations: Part II: The clinical relevance of the preoperative radiographic evaluation and report of five cases. *Oral Surg.* **9**, 85–93 (2016).
29. Huang, G., Liu, Z., Van Der Maaten, L. & Weinberger, K. Q. Densely Connected Convolutional Networks. in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 4700–4708. https://doi.org/10.1109/CVPR.2017.243 (2017).
30. Duman, ŞB. *et al.* Convolutional neural network performance for Sella turcica segmentation and classification using CBCT images. *Diagnostics* **12**, 2244 (2022).
31. Novikov, A. A., Major, D., Wimmer, M., Lenis, D. & Buhler, K. Deep sequential segmentation of organs in volumetric medical scans. *IEEE Trans. Med. Imaging* **38**, 1207–1215 (2019).
32. Oktay, O. *et al.* Attention u-net: Learning where to look for the pancreas. arXiv:1804.03999 (2018).
33. Aydemir, O. A new performance evaluation metric for classifiers: polygon area metric. *J. Classif.* **38**, 16–26 (2021).
34. Michel, J. *et al.* Determination of sex: Interest of frontal sinus 3 D reconstructions. *J. Forensic Sci.* **60**, 269–273 (2015).
35. Elharrouss, O., Akbari, Y., Almaadeed, N. & Al-Maadeed, S. Backbones-review: Feature extraction networks for deep learning and deep reinforcement learning approaches. arXiv:2206.08016 (2022).
36. Mall, P. K. *et al.* A comprehensive review of deep neural networks for medical image processing: Recent developments and future opportunities. *Healthc. Anal.* **4**, 100216 (2023).
37. Luo, H., Wang, J., Zhang, S. & Mi, C. The application of frontal sinus index and frontal sinus area in sex estimation based on lateral cephalograms among Han nationality adults in Xinjiang. *J. Forensic Leg. Med.* **56**, 1–4 (2018).

Acknowledgements

Author contributions

Renan Lucio Berbel da Silva: Contributed to conception and design, data acquisition, analysis, and interpretation, and drafted and critically revised the manuscript. Su Yang: Contributed to study conception and design, data analysis and interpretation, and drafted and critically revised the manuscript. DaEl Kim: Contributed to study conception and design, data analysis and interpretation, and drafted and critically revised the manuscript. Jun Ho Kim: Contributed to data analysis and interpretation, and drafted the manuscript. Sang-Heon Lim: Contributed to conception and design, data interpretation, and drafted the manuscript. Jiyong Han: Contributed to conception and design, data interpretation, and drafted the manuscript. Jun-Min Kim: Contributed to conception and design, data interpretation, and drafted the manuscript. Jo-Eun Kim: Contributed to conception and design, data interpretation, and drafted the manuscript. Kyung-Hoe Huh: Contributed to conception and design, data interpretation, and drafted the manuscript. Sam-Sun Lee: Contributed to conception and design, data interpretation, and drafted the manuscript. Min-Suk Heo: Contributed to conception and design, data acquisition, analysis,

and interpretation, and drafted and critically revised the manuscript. Won-Jin Yi: Contributed to conception and design, data acquisition, analysis, and interpretation, and drafted and critically revised the manuscript. All authors have given their final approval and agreed to be accountable for all aspects of the work.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-62211-y.

Correspondence and requests for materials should be addressed to M.-S.H. or W.-J.Y.

Reprints and permissions information is available at www.nature.com/reprints.

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