scientific reports

OPEN

AI-based automatic segmentation of craniomaxillofacial anatomy from CBCT scans for automatic detection of pharyngeal airway evaluations in OSA patients

Kaan Orhan[1,2,3✉], Mamat Shamshiev[4], Matvey Ezhov[4], Alexander Plaksin[4], Aida Kurbanova[5], Gürkan Ünsal[5,6], Maxim Gusarev[4], Maria Golitsyna[4], Seçil Aksoy[5], Melis Mısırlı[5], Finn Rasmussen[7,8], Eugene Shumilov[4] & Alex Sanders[4]

This study aims to generate and also validate an automatic detection algorithm for pharyngeal airway on CBCT data using an AI software (Diagnocat) which will procure a measurement method. The second aim is to validate the newly developed artificial intelligence system in comparison to commercially available software for 3D CBCT evaluation. A Convolutional Neural Network-based machine learning algorithm was used for the segmentation of the pharyngeal airways in OSA and non-OSA patients. Radiologists used semi-automatic software to manually determine the airway and their measurements were compared with the AI. OSA patients were classified as minimal, mild, moderate, and severe groups, and the mean airway volumes of the groups were compared. The narrowest points of the airway (mm), the field of the airway (mm^2), and volume of the airway (cc) of both OSA and non-OSA patients were also compared. There was no statistically significant difference between the manual technique and Diagnocat measurements in all groups ($p > 0.05$). Inter-class correlation coefficients were 0.954 for manual and automatic segmentation, 0.956 for Diagnocat and automatic segmentation, 0.972 for Diagnocat and manual segmentation. Although there was no statistically significant difference in total airway volume measurements between the manual measurements, automatic measurements, and DC measurements in non-OSA and OSA patients, we evaluated the output images to understand why the mean value for the total airway was higher in DC measurement. It was seen that the DC algorithm also measures the epiglottis volume and the posterior nasal aperture volume due to the low soft-tissue contrast in CBCT images and that leads to higher values in airway volume measurement.

Obstructive sleep apnea (OSA) is identified by periods of partial or complete upper airway disruption during sleep. OSA patients can breathe normally when they are awake but the disruptions occur since those patients cannot preserve the pharyngeal airway space when they sleep[1,2]. OSA patients who do not receive any treatment may have hypertension, heart failure, stroke, and premature death[3]. OSA patients are unavailable to preserve the pharyngeal airway space when they sleep, however, they can breathe normally when they are awake[2,3].

Inferior displacement of the hyoid bone, mandibular insufficiency, and increased soft palate and tongue volume are reported in the etiology of OSA in the literature[4]. Frequent reasons for collapsing of the upper airway are described as; the competence of the airway by reflexes, pharyngeal inspiratory muscle activity, and anatomic contraction of the upper airway[5,6]. Since the diagnosis of OSA requires a multidisciplinary approach, a dentist,

[1]Department of Dentomaxillofacial Radiology, Faculty of Dentistry, Ankara University, Ankara, Turkey. [2]Medical Design Application, and Research Center (MEDITAM), Ankara University, Ankara, Turkey. [3]Department of Dental and Maxillofacial Radiodiagnostics, Medical University of Lublin, Lublin, Poland. [4]Diagnocat Inc., San Francisco, CA, USA. [5]Department of Dentomaxillofacial Radiology, Faculty of Dentistry, Near East University, Nicosia, Cyprus. [6]Research Center of Experimental Health Science (DESAM), Near East University, Nicosia, Cyprus. [7]Internal Medicine Department Lunge Section, SVS Esbjerg, Esbjerg, Denmark. [8]Life Lung Health Center, Nicosia, Cyprus. ✉email: call53@yahoo.com

a neurologist, a cardiologist, an otorhinolaryngologist, and a pulmonary medicine specialist should be involved in the diagnosis and the treatment process[7].

CBCT is a 3-Dimensional radiographic diagnostic unit that can scan a region of interest with superior hard-tissue contrast and this provides a thorough analysis of the bony structures which are crucial in OSA diagnosis[8,9]. Thanks to its lower dose, lower cost, and higher image quality, CBCT is preferred over other advanced imaging methods such as Multi-Detector CT in dentistry, especially for the evaluation of the craniofacial structures[10].

Several articles are present in the literature with specific deep learning models for automatic segmentation of the maxillofacial structures, mandibular canal, cephalometric landmarks, cervical vertebras, and maxillofacial defects such as cleft palate. The majority of these models had U-Net architecture with a high (90%-95%) Dice Similarity Coefficient[11–26].

Anatomical structures such as the craniofacial skeleton and soft tissues which surround the muscles and pharynx have an important role in the configuration of the upper airway. Pharynx morphology is known to be one of the major factors that may cause OSA. Air-flow obstruction in children is also thought to occur due to skeletal deficiency since the contraction in the anterior–posterior aspect of the airway ensues from the positioning of the mandible and maxilla[27–29].

Numerous software is available to analyze the CBCT data with semi-automatic or manual volumetric measurement process[30], however, most of that software is laborious, time-consuming and a completely automatic airway detection algorithm is limited[12]. Thus, this study aimed to generate and also validate an automatic detection algorithm for pharyngeal airway on OSA patients' CBCT data using an artificial intelligence software of Diagnocat (DC).

Materials and methods

The research protocol was performed following the principles of the Declaration of Helsinki and was approved by the non-interventional Institutional Review Board (IRB) of Near East University Health Sciences Ethics Committee (YDU/2022/87-1251). Written informed consent was obtained from all patients before their radiographic examinations and anonymization was performed in compliance with the Information Commissioner's Anonymization: managing data protection risk code of practice (https://ico.org.uk/media/1061/anonymisation-code.pdf). The study data was created only from the deidentified anonymized data.

Anonymized DICOM files of the CBCT images which were taken by 3 different CBCT units were used in this study. The CBCT units were Pax-i3D Smart PHT-30LFO0 (Vatech, South Korea), Carestream Health CS 8100 3D (Kodak, USA), and Orthophos XG 3D (Sirona, Germany). All mentioned CBCT units have isotropic voxels which differ between 0.1 and 0.2 mm^3.

This study aimed to generate an AI algorithm for segmentation of the craniomaxillofacial anatomy and to test this algorithm for automatic detection algorithm for pharyngeal airway both for OSA and control patients. Thus, this study has two notable parts dataset preparation for the evaluation and to test the practicability of the system to enhance the diagnostic capabilities.

CBCT anatomy localization generated with an AI model.   *Approach.*   To handle large volume sizes on a reasonably fine scale, we approach this task with a coarse-to-fine approach. In general, a coarse-to-fine framework performs an inference at successively finer scales. The approach uses the results from the previous coarser stages to guide and speed up inference at the finer stages. A coarse-to-fine framework allows to achieve high-quality segmentation masks while being efficient during inference.

In this model, we use a two-stage coarse-to-fine approach. Both stages are defined as semantic segmentation tasks but at different voxel scales. At the first (coarse) stage, the whole volume is analyzed at once in a single forward pass through the neural network. During this stage, the model operates in coarse resolution of scale 1 mm. The goal of this stage is to perform a coarse segmentation of anatomical structures in a computationally efficient manner.

Next, we pass the results of the first stage as an input to the second (fine) stage. The fine stage allows us to achieve accurate segmentation masks by refining the outputs of the coarse stage. In this work, the fine stage is implemented as a patch-based semantic segmentation. The main idea of a patch-based approach is to train a neural network on small parts of the original images (and not the whole images) which leads to substantially reduced required computational resources. During inference, we extract the patches from the original image with an overlap and pass them through the model one by one. The results are then aggregated to form the final segmentation masks. At this stage, the training and inference are performed on a fine voxel scale of 0.25 mm.

Our approach to training the system consists of 4 main steps which are described in detail in the following sections: preprocessing of incoming volumetric image; coarse model training; coarse hint generation; patch-based training in fine resolution with a hint from the coarse model.

Data.   We use a simple min–max normalization within a fixed window. We clip the intensities to be inside the [− 1000, 2000] interval, then subtract a minimum intensity value and divide by a maximum one. Different methods have also been examined. According to our experiments, the training procedure is not sensitive to the choice of preprocessing and all methods lead to approximately the same results. The data is split into training, development, and test sets. We use 90% of the data for training, 5% for the development set, and 5% for the test set.

For the Coarse step, we rescale the image to have a 1.0 mm isotropic voxel resolution using linear interpolation. To provide the Coarse model with more information, we obtain soft coarse segmentation ground truth labels by the following procedure. First, we encode the original semantic segmentation mask of shape DxHxW with a one-hot encoding scheme which results in a tensor of shape tensor CxDxHxW, where C represents the number of classes and D, H, and W are the spatial dimensions of the original volume. Next, we use linear interpolation

to rescale this tensor to have a 1.0 mm resolution. The resulting tensor consists of the probability distributions over classes for each spatial position and is referred to as soft targets.

For the Fine step, the target voxel spacing of the model is $0.25 \times 0.25 \times 0.25$ mm which is also achieved with linear interpolation of the image. For this step, we obtain the ground truth labels via a simple nearest-neighbor interpolation of original semantic segmentation masks. During training, we randomly sample patches of size $144 \times 144 \times 144$ voxels.

Model. We formulate both Coarse and Fine steps as a semantic segmentation task, where the background and each anatomical element are interpreted as a separate class. For both Coarse and fine steps, we use 3D U-Net[31] which is a standard, widely known, and well-studied fully convolutional neural network architecture[14,32–37]. Our implementation follows the architecture which is described in detail in the original paper[31].

Since the Fine model is trained using a patch-based approach, it's crucial to provide the model with global information. We achieve it by utilizing a coarse hint. A coarse hint is the Coarse model output which is interpolated to the Fine-scale and passed to the Fine model as additional input channels. To prevent possible data leakage, we train the Coarse model and prepare coarse hints via three-fold cross-validation. Therefore, the only difference between the Coarse and Fine model architectures is the number of input channels: for the Coarse step it equals 1, and for the Fine step, it equals the number of classes plus 1.

The class imbalance is known to be a challenging problem in medical semantic segmentation tasks. We approach this issue by using a sum of a standard cross-entropy loss and soft multiclass Jaccard loss. To prevent overfitting and enhance the performance of the model we also utilize a large variety of data augmentations. For the Coarse step the following augmentations are used during training: random blur, random noise, random rotations, random scaling, random crops, random elastic deformation, and random anisotropy[38]. For the Fine step, we used the same set of augmentations except for random elastic deformation and random anisotropy since these transformations are computationally expensive when applied to reasonably large images.

Training. To sum it up, our training procedure consists of the following steps. First, we train the Coarse model on the coarse training dataset with soft targets. This checkpoint is used during the testing. We also perform three-fold cross-validation and use the obtained checkpoints later to generate coarse hints for the Fine step. For both cross-validation and full data training, we follow the same procedure. We train for a total of 100 epochs using an Adam optimizer with a one-cycle scheduling policy with a maximum learning rate equal to 1e−3, minimum learning rate equal to 1e−6, warmup fraction of 0.05, and a batch size of 1.

Next, we prepare coarse hints for the Fine model. We utilize the checkpoints received via cross-validation and make out-of-folds predictions, then linearly interpolate the output probability maps to the Fine model voxel spacing and concatenate them with the original intensity value channel. Finally, we train the Fine model for a total of 40 epochs, using the Adam optimizer and the same learning rate scheduling policy, as in the Coarse step.

To train the Fine model we use a patch-based approach. At the beginning of the training epoch, we iterate over the images, randomly sample 20 patches per volume and store them in a queue of size 180. Once the queue has reached a specified maximum length we start to retrieve the random patches from it and pass them to the network while simultaneously preparing new patches and adding them to the queue. For evaluation, we use the checkpoint with the lowest recorded validation loss for both Coarse and Fine models.

Implementation. Our algorithm was based on the Python implementation of U-net. All training and experiments were done using NVIDIA GeForce RTX A100 GPU. Adam optimizer was used for the network training.

Inference. At test time the patch-based approach is known to produce the predictions of a worse quality near the borders of the output patch. To alleviate this issue, we perform inference in overlapping patches and aggregate the predictions with weights which make the center voxel of an output patch contribute more to the final result than its borders. We set the patches' overlap to 16 (Fig. 1).

Patient test dataset. To estimate the generalizability of our model, a retrospective patient CBCT dataset from Dentomaxillofacial Radiology Department at Near East University was used. A power analysis was conducted with a statistical power of 90%, a significance level of 0.05 α, and a probability of type II error of 0.2 β. A minimum number of 82 CBCT images for both control and OSA groups were required according to the power analysis.

Hence, our study was conducted with randomly selected artifact-free 100 OSA and 100 control CBCT images existing in our faculty's database. All patients provided their informed consent before irradiation, and the consent forms were reviewed and approved by the institutional review board of the faculty. Exclusion criteria were evident skeletal asymmetries, cleft palate, cleft lip, current ongoing orthodontic treatment, and any teeth that overlie the apical region of the incisors.

The dataset of a previous study[39] of ours is used in this study "CBCT records of 200 patients (100 images of OSA patients and 100 images of the control group) were retrospectively collected and analyzed along with the polysomnography records and body mass index (BMI) of OSA patients at the Department of Allergy, Sleep and Respiratory Diseases. AHI is the number of apnea + hypopnea seen each hour during sleep. Sleep apnea severity was evaluated in 4 different subtypes minimal, mild, moderate, and severe. Patients with Apnea–Hypopnea Index (AHI) value lower than 5 were classified as a minimal group while patients with AHI values between 5–15, 15–30, and more than 30 were classified as mild, moderate, and severe, respectively. 100 OSA patients had symptoms of this disease and evaluation of these patients was accomplished by a standardized program at the Department of Allergy, Sleep and Respiratory Diseases, which also consists of anthropometric measurements,
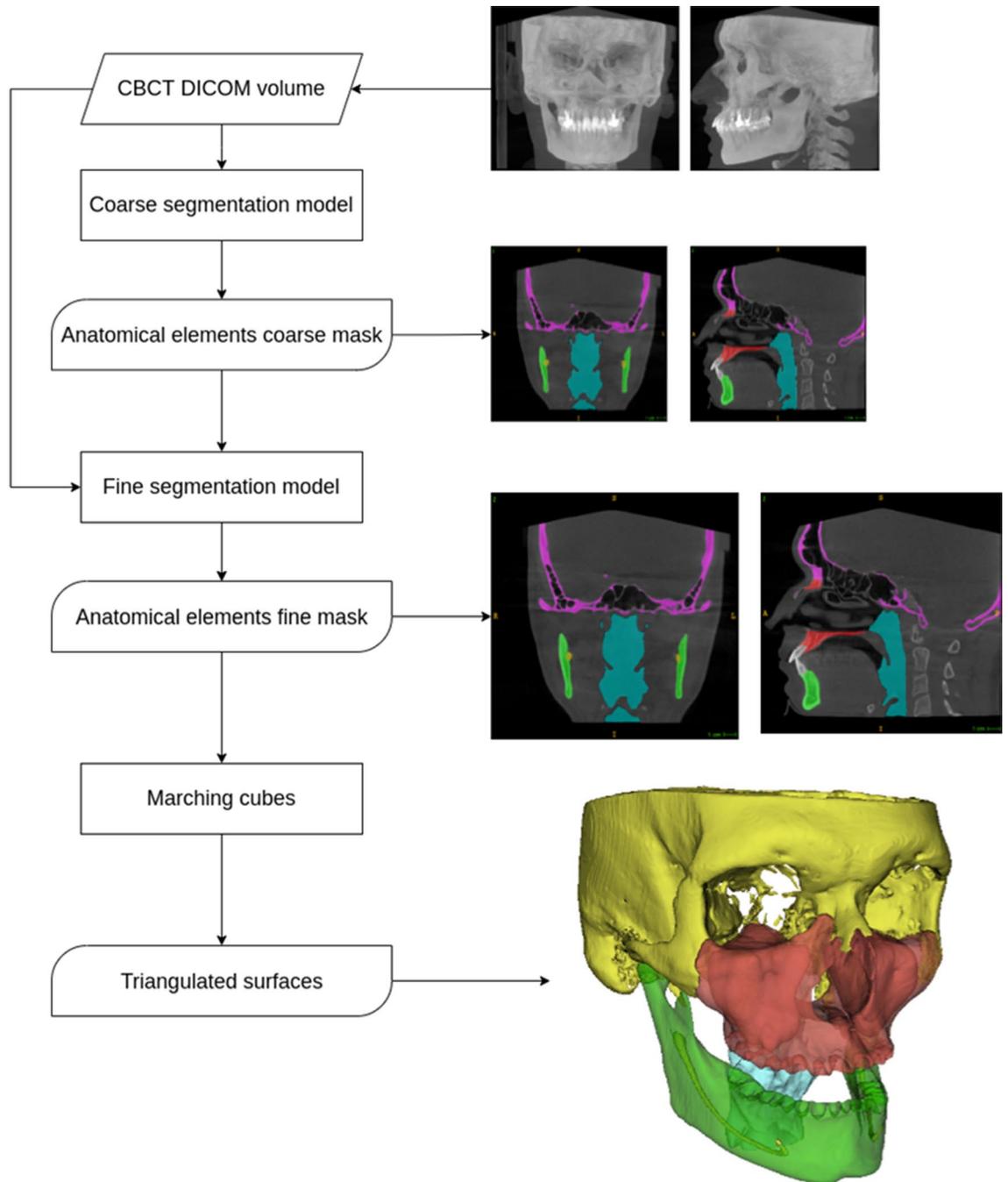
Figure 1. Inference algorithm of our study.

dental examination, CBCT, and polysomnography. Polysomnography uses various methods like electroencephalography, electromyogram, electro-oculography, respiratory effort measurement, airflow measurement, and snoring[29]. Control (non-OSA) patients had none of the clinical findings of the OSA patients such as snoring, dyspnea, witnessed apnea, coughing, or daytime sleepiness. So their images were used as a control group. The mean age for OSA patients was 53.2 years and for non-OSA patients was 46.4 years. Principles characterized in the Declaration of Helsinki were applied during the protocol of study along with modifications and revisions.

CBCT images of the test group were obtained by NewTom 3 G Quantitive Radiology s.r.l., (NewTom, Verona, Italy). CBCT records for non-OSA patients had been taken for implant planning, evaluation of impacted teeth, and prosthodontic and orthodontic purposes. Patients with osteoporosis, skeletal asymmetries, and medication-related bony alterations were excluded from the study.

Ground truth segmentation process. All CBCT data were exported as DICOM files and then anonymized. The axial, coronal, and sagittal slices were oriented to ensure a proper evaluation. The axial slices

4

OSA severity	The technique (subgroup)	Mean	Median	Min	Max	SD	Mann Whitney U		
							Mean rank	U	p
Airway volume (cc)									
Minimal OSA	Manual	21.18	21.03	11.81	34.87	6.49	31.67	252	0.052
	Diagnocat	17.73	16.20	7.60	29.60	6.16	23.33		
	Total	19.45	19.66	7.60	34.87	6.51			
Mild OSA	Manual	18.32	17.72	7.64	28.46	6.23	19.63	178	0.942
	Diagnocat	18.11	17.60	7.70	29.50	6.18	19.37		
	Total	18.22	17.66	7.64	29.50	6.12			
Moderate OSA	Manual	22.42	22.55	9.69	35.53	7.45	22.38	202	0.642
	Diagnocat	21.42	21.20	6.60	35.30	6.96	20.62		
	Total	21.92	21.92	6.60	35.53	7.14			
Severe OSA	Manual	19.21	17.85	7.34	34.97	7.54	36.48	446	0.207
	Diagnocat	16.79	15.20	5.60	29.60	6.73	30.52		
	Total	18.00	15.72	5.60	34.97	7.20			

Table 1. Comparison of the airway volume measurements of Diagnocat and manual technique in patients with different OSA severities.

were aligned with maintaining the palate line and the ground perpendicular to each other. Coronal slices were oriented by aligning the both orbits and midline of the head parallel to the ground and the sagittal slices were aligned with the linear orientation of the ANS and PNS.

All CBCT images had been segmented before our study to be used for diagnosis, pharyngeal airway evaluations, and surgical planning using InVivo 5.1.2 (*Anatomage Inc., San Jose, CA, USA*). DICOM files of the axial CBCT images were exported with a 512×512 matrix and were imported to InVivo 5.1.2. In this software, the evaluation of the pharyngeal airway can be measured by both automatic thresholding and manual tracing with semiautomatic thresholding.

The pharyngeal airway is originated from the nasopharynx and the oropharynx. To assess the borders of the oropharyngeal airway volume, the ANS-PNS plane which extends to the wall of the pharynx was determined as the superior border and the most inferior-anterior point of the 2nd cervical vertebrae which is parallel to the superior border was determined as the lower border of the oropharyngeal airway. Since the superior border of the oropharyngeal airway is also the lower border of the nasopharyngeal airway, a line perpendicular from the PNS to the palatal plane is drawn to form the anterior border of the nasopharyngeal airway. The Sum of the nasopharyngeal airway and oropharyngeal airway is calculated with both manual tracing with semi-automatic thresholding and automatic thresholding in InVivo 5.1.2. viewer. S.A. and A.K. observed the CBCT images twice with a week interval to avoid any intra-observer disagreement for ground truth measurement.

For automatic thresholding, the software itself detects the pharyngeal airway volume, area narrow point area, and measures the narrow point automatically.

The manual tracing with semiautomatic thresholding was done by cropping the airway using the "edit masks" feature and the connection with the outer air was cropped in each slice with the segmentation tools. The "region growing" tool was used to split the segmentation produced by thresholding into several objects and to remove floating pixels and the pharyngeal airway volume and area were calculated using the "calculate 3D" tool feature of the software.

3D U-net architecture framework (AI model). Our approach is automatic segmentation focusing on the regions of interest: the external surface of the bones, teeth, and airways. This process results in 5 segmentation masks the upper skull, the mandible, maxillary teeth, mandibular teeth, and the airways. We performed a series of trials to choose the best training configuration. Following, the generated STL files were downloaded and imported to 3rd party software for volumetric pharyngeal airway measurements (*3-Matic Version 15, Materialise*).

Statistical analysis. Statistical analysis was performed using SPSS 22.0 software (*SPSS Inc., Chicago, IL, USA*). Due to the non-normal distribution of the data, the Mann–Whitney U test was used for comparisons between paired groups, and the Kruskall Wallis H test was used for comparisons between three or more groups. The significance level was set as 0.05 and it was stated that there was a significant difference in the case of $p < 0.05$, and no significant difference in the case of $p > 0.05$. Interclass correlation coefficient (ICC) analysis with a two-way mixed model was performed. It was assumed that ICC values greater than 0.75 would guarantee good reliability and ICC values greater than 0.90 would guarantee excellent reliability between observers.

Results

There was no statistically significant difference in airway volume (cc) measurement difference between the manual measurement and DC in any of the OSA severity subtypes ($p > 0.05$). p values were 0.052, 0.942, 0.642, and 0.207 for the minimal, mild, moderate, and severe OSA groups, respectively (Table 1). Statistical analysis

	Non-OSA patients									
	n	Mean	Median	Minimum	Maximum	SD	Mean Rank	U	p	Test
Narrowest points (mm)										
Manual	100	5.96	5.58	1.89	13.30	2.07	104.36	4614.5	0.346	Mann–Whitney U
Diagnocat	100	5.70	5.41	1.69	14.56	2.10	96.65			
Total	200	5.83	5.46	1.69	14.56	2.08				
Airway area (mm²)										
Manual	100	883.41	856.73	437.65	1576.88	212.92	93.97	4347	0.111	Mann–Whitney U
Diagnocat	100	930.02	909.00	597.33	1694.00	201.18	107.03			
Total	200	906.71	895.67	437.65	1694.00	207.93				
Volume (cc)										
Manual	100	17.95	17.70	4.90	34.10	5.45	146.12	0.811	0.667	Kruskall-Wallis H
Diagnocat	100	18.50	18.40	5.50	35.20	5.63	156.71			
Automatic	100	17.96	18.20	4.80	32.80	5.41	148.68			
Total	300	18.14	18.00	4.80	35.20	5.48				

Table 2. Comparison of airway volume, airway area, and narrowest line of the airway measurements of Diagnocat, manual technique, and automatic technique in patients without OSA.

	OSA patients									
	n	Mean	Median	Minimum	Maximum	SD	Mean Rank	U	p	Test
Narrowest points (mm)										
Manual	100	6.31	5.86	1.48	23.08	3.42	100.14	4964	0.931	Mann–Whitney U
Diagnocat	100	6.10	5.76	1.01	19.90	2.50	100.86			
Total	200	6.20	5.78	1.01	23.08	2.99				
Airway area (mm²)										
Manual	100	1057.59	1033.07	598.85	1731.64	244.38	104.70	4580	0.305	Mann–Whitney U
Diagnocat	100	1013.90	989.81	466.52	1670.22	256.57	96.30			
Total	200	1035.74	1001.13	466.52	1731.64	250.88				
Volume (cc)										
Manual	100	19.63	19.05	7.40	35.30	6.90	153.05	3.9	0.139	Kruskall-Wallis H
Diagnocat	100	20.25	19.83	7.34	35.53	7.08	161.21			
Automatic	100	18.27	17.50	5.60	35.30	6.65	137.24			
Total	300	19.38	18.55	5.60	35.53	6.91				

Table 3. Comparison of airway volume, airway area, and narrowest line of the airway measurements of Diagnocat, manual technique, and automatic technique in patients with OSA.

showed excellent ICC (ICC > 0.90) for all inter-evaluator assessments. ICC values were 0.954 for manual and automatic segmentation, 0.956 for DC and automatic segmentation, 0.972 for DC and manual segmentation.

Measurements for non-OSA patients. There was no statistically significant difference in narrowest points (mm), airway area (mm²), and total airway volume (cc) measurements between the manual measurements and DC measurements in non-OSA patients. p values were 0.346, 0.111 and 0.667, respectively. The mean value for the narrowest distance was found 5.96 mm with the manual measurement and 5.70 mm with DC. The mean value for the airway area was found 883.41 mm² with the manual measurement and 930.02 mm² with DC. The mean value for the total airway volume was found 17.95 cc with the manual measurement, 17.96 cc with the automatic technique, and 18.50 cc with DC (Table 2).

Measurements for OSA patients. There was, also, no statistically significant difference in narrowest points (mm), airway area (mm²), and total airway volume (cc) measurements between the manual measurements and DC measurements in OSA patients. p values were 0.931, 0.305 and 0.139, respectively. The mean value for the narrowest distance was found 6.31 mm with the manual measurement and 6.10 mm with DC. The mean value for the airway area was found 1057.59 mm² with the manual measurement and 1013.90 mm² with DC. The mean value for the total airway volume was found 19.63 cc with the manual measurement, 18.27 cc with the automatic measurement, and 20.25 cc with DC (Table 3).

Authors	Year	Title	Modality	Softwares	Interclass correlation coefficient	Intersection over union (IoU)
Zhang et al.	2019	A new segmentation algorithm for measuring CBCT images of nasal airway: a pilot study	CBCT	Airway Segmentor, MIMICS 19.0, InVivo 5	0.899	
Leonardi et al.	2020	Fully automatic segmentation of sinonasal cavity and pharyngeal airway based on convolutional neural networks	CBCT	Own model	0.977	
Sin et al.	2021	A deep learning algorithm proposal to automatic pharyngeal airway detection and segmentation on CBCT images	CBCT	Own model	0.985	
Park et al.	2021	Deep learning based airway segmentation using key point prediction	CBCT	Own model	0.986	
Shujaat et al.	2021	Automatic segmentation of the pharyngeal airway space with convolutional neural network	CBCT, MSCT	Own model	–	0.93
This study	2022	AI-based automatic segmentation of craniomaxillofacial anatomy from CBCT scans for automatic detection of pharyngeal airway evaluations in OSA patients	CBCT	Diagnocat, InVivo 5	Manual-Automatic 0.954; DC-Automatic 0.956; DC-Manual 0.972	

Table 4. Comparative table of the studies using deep learning algorithms for pharyngeal airway volume segmentation and measurement.

Discussion

According to our review of the literature, this study is the first study that automatically measured the pharyngeal airway in OSA patients. However, manual measurements of the pharyngeal airway in OSA patients and automatic measurements of the pharyngeal airway in non-OSA patients are present in the literature. Since the studies demonstrated that the pharyngeal airway volume is significantly lower and the morphology is dissimilar in oral breathers than in nasal breathers, deep learning algorithms which concentrate on airway volume measurements should be trained and tested with various data. In orthodontics, airway volume and the underlying factors play a crucial role before orthognathic surgery planning, analyzing the airway volume is indispensable to understanding the oral and pharyngeal adjustments to respiratory conditions[12,40–42].

Although there was no statistically significant difference in total airway volume measurements between the manual measurements, automatic measurements, and DC measurements in non-OSA and OSA patients, we evaluated the output images to understand why the mean value for the total airway was higher in DC measurement. It was seen that the DC algorithm also measures the epiglottis volume and the posterior nasal aperture volume due to the low soft-tissue contrast in CBCT images and that leads to higher values in airway volume measurement.

The mean total airway volume difference between automatic measurement and manual measurement in non-OSA patients was just 0.01 cc, however, it was 1.36 cc in OSA patients. Output images were again compared and it was seen that there were voxel loss sites at the posterior nasal aperture border in the automatic measurement group.

Various authors measured the airway in adults and according to their findings, the difference between their manual measurements has occurred due to the "human factor" and different software that is used for the measurements[12,43,44]. Following an extensive literature research, we have given the ICC values of 5 studies that compared the segmentation of AI and ground truth in Table 4. ICC values were reported as 0.899 by Zhang et al.[45], 0.977 by Leonardi et al.[46], 0.985 by Sin et al.[12], 0.986 by Park et al.[47]. Shujaat et al.[48] provided precision, recall, accuracy, dice, intersection over union values in their study as 0.97 ± 0.01, 0.96 ± 0.03 1.00 ± 0.00 0.97 ± 0.02 and 0.93 ± 0.03, respectively. In our study, the ICC value between the ground truth and DC was 0.972 which indicates that an excellent reliability was present. As the previous studies which aimed to segment and measure the pharyngeal airway volume, we also achieved a high ICC value which evidently shows that well-trained deep learning algorithms can successfully segment the pharyngeal airway.

In our study, we had an imaging modality-related limitation, which was caused by the limited soft-tissue contrast of the CBCT units. Arranging a precise Hounsfield Unit (HU) threshold in the segmentation process is not possible with CBCT units since HUs are not applicable for them[49]. This limitation might affect the airway volume measurements as they have affected our segmentation process. The inconsistent head position of the patients, tongue, and breathing positions also cause errors in volumetric measurement, thus, scannings with controlling these possible limitations are required[47].

AI is widely known for its functions in image recognition, computer-aided diagnosis, and decision-making algorithms. Given that 90% of the clinical data is medical images, AI can collaborate with the Internet of Things (IoT) to make health care more advanced with the remote diagnosis that can accelerate the diagnosis and treatment phases[50–52]. Activating this potential collaboration for OSA patients would significantly reduce the effort and time required for the initial diagnosis and follow-up of these patients.

References

1. Dempsey, J. A., Veasey, S. C., Morgan, B. J. & O'Donnell, C. P. Pathophysiology of sleep apnea. *Physiol. Rev.* **90**, 47–112. https://doi.org/10.1152/physrev.00043.2008 (2010).

2. Salles, C., Terse-Ramos, R., Souza-Machado, A. & Cruz, A. A. Obstructive sleep apnea and asthma. *J. Bras. Pneumol.* **39**, 604–612. https://doi.org/10.1590/S1806-37132013000500011 (2013).

3. Ogna, A. *et al.* Obstructive sleep apnea severity and overnight body fluid shift before and after hemodialysis. *Clin. J. Am. Soc. Nephrol.* **10**, 1002–1010. https://doi.org/10.2215/CJN.08760914 (2015).

4. Drakatos, P. *et al.* Computed tomography cephalometric and upper airway measurements in patients with OSA and erectile dysfunction. *Sleep Breath* **20**, 769–776. https://doi.org/10.1007/s11325-015-1297-5 (2016).

5. Hudgel, D. W. Mechanisms of obstructive sleep apnea. *Chest* **101**, 541–549. https://doi.org/10.1378/chest.101.2.541 (1992).

6. White, D. P. & Younes, M. K. Obstructive sleep apnea. *Compr. Physiol.* **2**, 2541–2594. https://doi.org/10.1002/cphy.c110064 (2012).

7. Oz, U. *et al.* Association between pterygoid hamulus length and apnea hypopnea index in patients with obstructive sleep apnea: A combined three-dimensional cone beam computed tomography and polysomnographic study. *Oral Surg. Oral Med. Oral Pathol. Oral Radiol.* **121**, 330–339. https://doi.org/10.1016/j.oooo.2015.10.032 (2016).

8. Neelapu, B. C. *et al.* Craniofacial and upper airway morphology in adult obstructive sleep apnea patients: A systematic review and meta-analysis of cephalometric studies. *Sleep Med. Rev.* **31**, 79–90. https://doi.org/10.1016/j.smrv.2016.01.007 (2017).

9. von Arx, T., Matter, D., Buser, D. & Bornstein, M. M. Evaluation of location and dimensions of lingual foramina using limited cone-beam computed tomography. *J. Oral Maxillofac. Surg.* **69**, 2777–2785. https://doi.org/10.1016/j.joms.2011.06.198 (2011).

10. Sheikhi, M., Mosavat, F. & Ahmadi, A. Assessing the anatomical variations of lingual foramen and its bony canals with CBCT taken from 102 patients in Isfahan. *Dent. Res. J. (Isfahan)* **9**, S45-51 (2012).

11. Gupta, A., Kharbanda, O. P., Sardana, V., Balachandran, R. & Sardana, H. K. A knowledge-based algorithm for automatic detection of cephalometric landmarks on CBCT images. *Int. J. Comput. Assist. Radiol. Surg.* **10**, 1737–1752. https://doi.org/10.1007/s11548-015-1173-6 (2015).

12. Sin, C., Akkaya, N., Aksoy, S., Orhan, K. & Oz, U. A deep learning algorithm proposal to automatic pharyngeal airway detection and segmentation on CBCT images. *Orthod. Craniofac. Res.* **24**(Suppl 2), 117–123. https://doi.org/10.1111/ocr.12480 (2021).

13. Amasya, H., Cesur, E., Yildirim, D. & Orhan, K. Validation of cervical vertebral maturation stages: Artificial intelligence vs human observer visual analysis. *Am. J. Orthod. Dentofacial. Orthop.* **158**, e173–e179. https://doi.org/10.1016/j.ajodo.2020.08.014 (2020).

14. Wang, F., Jiang, R., Zheng, L., Meng, C. & Biswal, B. *International MICCAI Brainlesion Workshop* 131–141 (Springer, 2019).

15. Wang, H. *et al.* Multiclass CBCT image segmentation for orthodontics with deep learning. *J. Dent. Res.* **100**, 943–949. https://doi.org/10.1177/00220345211005338 (2021).

16. Zhang, J. *et al.* Context-guided fully convolutional networks for joint craniomaxillofacial bone segmentation and landmark digitization. *Med. Image Anal.* **60**, 101621. https://doi.org/10.1016/j.media.2019.101621 (2020).

17. Torosdagli, N. *et al.* Deep geodesic learning for segmentation and anatomical landmarking. *IEEE Trans. Med. Imaging* **38**, 919–931. https://doi.org/10.1109/TMI.2018.2875814 (2019).

18. Qiu, B. *et al.* Automatic segmentation of the mandible from computed tomography scans for 3D virtual surgical planning using the convolutional neural network. *Phys. Med. Biol.* **64**, 175020. https://doi.org/10.1088/1361-6560/ab2c95 (2019).

19. Qiu, B. *et al.* Recurrent convolutional neural networks for mandible segmentation from computed tomography. *J. Person. Med.* **11**, 492. https://doi.org/10.3390/jpm11060492 (2020).

20. Minnema, J. *et al.* Segmentation of dental cone-beam CT scans affected by metal artifacts using a mixed-scale dense convolutional neural network. *Med. Phys.* **46**, 5027–5035. https://doi.org/10.1002/mp.13793 (2019).

21. Lian, C. *et al.* Multi-task dynamic transformer network for concurrent bone segmentation and large-scale landmark localization with dental CBCT. *Med. Image Comput. Comput. Assist. Interv.* **12264**, 807–816. https://doi.org/10.1007/978-3-030-59719-1_78 (2020).

22. Jaskari, J. *et al.* Deep learning method for mandibular canal segmentation in dental cone beam computed tomography volumes. *Sci. Rep.* **10**, 5842. https://doi.org/10.1038/s41598-020-62321-3 (2020).

23. Kwak, G. H. *et al.* Automatic mandibular canal detection using a deep convolutional neural network. *Sci. Rep.* **10**, 5711. https://doi.org/10.1038/s41598-020-62586-8 (2020).

24. Cui, Z., Li, C. & Wang, W. in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* 6361–6370 (2019).

25. Chung, M. *et al.* Pose-aware instance segmentation framework from cone beam CT images for tooth segmentation. *Comput. Biol. Med.* **120**, 103720. https://doi.org/10.1016/j.compbiomed.2020.103720 (2020).

26. Dot, G., Schouman, T., Dubois, G., Rouch, P. & Gajny, L. Fully automatic segmentation of craniomaxillofacial CT scans for computer-assisted orthognathic surgery planning using the nnU-Net framework. *Eur. Radiol.* https://doi.org/10.1101/2021.07.22.21260825 (2021).

27. Stellzig-Eisenhauer, A. & Meyer-Marcotty, P. Interaction between otorhinolaryngology and orthodontics: Correlation between the nasopharyngeal airway and the craniofacial complex. *GMS Curr. Top. Otorhinolaryngol. Head Neck Surg.* **9**, Doc04. https://doi.org/10.3205/cto000068 (2010).

28. Avci, S., Lakadamyali, H., Lakadamyali, H., Aydin, E. & Tekindal, M. A. Relationships among retropalatal airway, pharyngeal length, and craniofacial structures determined by magnetic resonance imaging in patients with obstructive sleep apnea. *Sleep Breath* **23**, 103–115. https://doi.org/10.1007/s11325-018-1667-x (2019).

29. Rundo, J. V. & Downey, R. 3rd. Polysomnography. *Handb. Clin. Neurol.* **160**, 381–392. https://doi.org/10.1016/B978-0-444-64032-1.00025-4 (2019).

30. Park, C. W. *et al.* Volumetric accuracy of cone-beam computed tomography. *Imaging Sci. Dent.* **47**, 165–174. https://doi.org/10.5624/isd.2017.47.3.165 (2017).

31. Çiçek, Ö., Abdulkadir, A., Lienkamp, S. S., Brox, T. & Ronneberger, O. in *Medical Image Computing and Computer-Assisted Intervention: MICCAI 2016* Vol. 9901 424–432 (2016).

32. Chen, W., Liu, B., Peng, S., Sun, J. & Qiao, X. in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries Lecture Notes in Computer Science* Ch. Chapter 32, 358–368 (2019).

33. Zeng, G. *et al. From Large to Small Organ Segmentation in CT Using Regional Context* (Springer, 2017).

34. Mehta, R. & Arbel, T. *International MICCAI Brainlesion Workshop* 254–266 (Springer, 2018).

35. Baid, U. *et al.* A novel approach for fully automatic intra-tumor segmentation with 3D U-net architecture for gliomas. *Front Comput. Neurosci.* **14**, 10. https://doi.org/10.3389/fncom.2020.00010 (2020).

36. Müller, D., Soto-Rey, I. & Kramer, F. Robust chest CT image segmentation of COVID-19 lung infection based on limited data. *Inform. Med. Unlocked.* https://doi.org/10.1016/j.imu.2021.100681 (2021).

37. Park, J. *et al.* Fully automated lung lobe segmentation in volumetric chest CT with 3D U-Net: Validation with intra- and extra-datasets. *J. Digit. Imaging* **33**, 221–230. https://doi.org/10.1007/s10278-019-00223-1 (2020).

38. Perez-Garcia, F., Sparks, R. & Ourselin, S. TorchIO: A Python library for efficient loading, preprocessing, augmentation and patch-based sampling of medical images in deep learning. *Comput. Methods Prog. Biomed.* **208**, 106236. https://doi.org/10.1016/j.cmpb.2021.106236 (2021).
39. Firincioglulari, M., Aksoy, S., Orhan, K., Oz, U. & Rasmussen, F. Comparison of anterior mandible anatomical characteristics between obstructive sleep apnea patients and healthy individuals: A combined cone beam computed tomography and polysomnographic study. *Eur. Arch. Otorhinolaryngol.* **277**, 1427–1436. https://doi.org/10.1007/s00405-020-05805-2 (2020).
40. Alves, M. Jr., Baratieri, C., Nojima, L. I., Nojima, M. C. & Ruellas, A. C. Three-dimensional assessment of pharyngeal airway in nasal- and mouth-breathing children. *Int. J. Pediatr. Otorhinolaryngol.* **75**, 1195–1199. https://doi.org/10.1016/j.ijporl.2011.06.019 (2011).
41. Pinheiro de MagalhaesBertoz, A. *et al.* Three-dimensional airway changes after adenotonsillectomy in children with obstructive apnea: Do expectations meet reality?. *Am. J. Orthod. Dentofac. Orthop.* **155**, 791–800. https://doi.org/10.1016/j.ajodo.2018.06.019 (2019).
42. Cuccia, A. M., Lotti, M. & Caradonna, D. Oral breathing and head posture. *Angle Orthod.* **78**, 77–82. https://doi.org/10.2319/011507-18.1 (2008).
43. Hong, J. S., Oh, K. M., Kim, B. R., Kim, Y. J. & Park, Y. H. Three-dimensional analysis of pharyngeal airway volume in adults with anterior position of the mandible. *Am. J. Orthod. Dentofacial. Orthop.* **140**, e161-169. https://doi.org/10.1016/j.ajodo.2011.04.020 (2011).
44. Grauer, D., Cevidanes, L. S., Styner, M. A., Ackerman, J. L. & Proffit, W. R. Pharyngeal airway volume and shape from cone-beam computed tomography: Relationship to facial morphology. *Am. J. Orthod. Dentofac. Orthop.* **136**, 805–814. https://doi.org/10.1016/j.ajodo.2008.01.020 (2009).
45. Zhang, C. *et al.* A new segmentation algorithm for measuring CBCT images of nasal airway: A pilot study. *PeerJ* **7**, e6246. https://doi.org/10.7717/peerj.6246 (2019).
46. Leonardi, R. *et al.* Fully automatic segmentation of sinonasal cavity and pharyngeal airway based on convolutional neural networks. *Am. J. Orthod. Dentofac. Orthop.* **159**, 824–835. https://doi.org/10.1016/j.ajodo.2020.05.017 (2021).
47 Park, J. *et al.* Deep learning based airway segmentation using key point prediction. *Appl. Sci.* https://doi.org/10.3390/app11083501 (2021).
48. Shujaat, S. *et al.* Automatic segmentation of the pharyngeal airway space with convolutional neural network. *J. Dent.* **111**, 103705. https://doi.org/10.1016/j.jdent.2021.103705 (2021).
49. Pauwels, R., Jacobs, R., Singer, S. R. & Mupparapu, M. CBCT-based bone quality assessment: Are Hounsfield units applicable?. *Dentomaxillofac. Radiol.* **44**, 20140238. https://doi.org/10.1259/dmfr.20140238 (2015).
50. Lu, Z. X. *et al.* Application of AI and IoT in clinical medicine: Summary and challenges. *Curr. Med. Sci.* **41**, 1134–1150. https://doi.org/10.1007/s11596-021-2486-z (2021).
51. Qiu, J. *et al.* A Survey on access control in the age of internet of things. *IEEE Internet Things J.* **7**, 4682–4696. https://doi.org/10.1109/jiot.2020.2969326 (2020).
52. Qiao, C., Brown, K. N., Zhang, F. & Tian, Z. Federated adaptive asynchronous clustering algorithm for wireless mesh networks. *IEEE Trans. Knowl. Data Eng.* https://doi.org/10.1109/tkde.2021.3119550 (2021).

Acknowledgements

Author contributions

K.O. designed the work and interpreted the data. A.K, G.Ü., S.A., F.R., and M.M. interpreted the data and were responsible for the data acquisition and analysis. M.E., M.G., M.G., E.S., and A.S., created new software for this study. K.O. and G.Ü. drafted the work and all authors contributed or revised it. All authors agreed both to be personally accountable for the author's contributions and to ensure that questions related to the accuracy or integrity of any part of the work, even ones in which the author was not personally involved, are appropriately investigated, resolved, and the resolution documented in the literature.

Competing interests

Financial support was received by Diagnocat Co. Ltd., San Francisco CA. Matvey Ezhov, Maxim Gusarev, Alexander Plaksin, Mamat Shamshiev, Maria Golitsyna, Eugene Shumilov, and Alex Sanders are employees of Diagnocat Co. Ltd. Kaan Orhan is a scientific research advisor for the Diagnocat Co. Ltd., Aida Kurbanova, Melis Misirli, Gürkan Ünsal, Secil Aksoy, Finn Rasmussen have no potential competing interests.

Additional information

Correspondence and requests for materials should be addressed to K.O.

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.