

Deep Learning in Analysing Paranasal Sinuses

Serkan Ozbay^{1,*}, Orhan Tunc²

¹*Department of Electrical and Electronics Engineering, University of Gaziantep,
27310 Gaziantep, Turkey*

²*Department of Otorhinolaryngology, University of Gaziantep,
27310 Gaziantep, Turkey
sozbay@gantep.edu.tr*

Abstract—Deep neural network-based diagnostic tools have gained state-of-the-art performance in the medical field in recent years. Diagnostic accuracy has become very critical for medical treatments. This paper proposes a simple and novel deep learning-based system for the analysis of paranasal sinuses conditions. In this work, we focus on analysing the paranasal sinuses on CT images automatically, providing physicians with high-accuracy diagnosis. The proposed system enables one to reduce the number of images to be searched in a CT scan for a patient automatically, and also it provides automatic segmentation for marking and cropping the paranasal sinuses region. Thus, the proposed system significantly decreases the data required in the training phase with a gain in computational efficiency while maintaining high-accuracy performance. The proposed algorithm also makes the required segmentation automatically without manual cropping and yields outstanding performance on detecting abnormalities in the sinuses. The proposed approach has been tested on real CT images and achieved an accuracy rate of 98.52 % with a sensitivity of 100 %.

Index Terms—Convolutional neural network; Deep learning; Medical imaging; Paranasal sinus.

I. INTRODUCTION

In the healthcare industry, many medical images are produced and then used for diagnosis. Medical imaging technologies enable professionals to view the inner structures of the body for diagnosis, monitoring, or treating medical conditions. There is an ever-increasing amount of medical image data and the rate of growth itself is increasing. In recent years, most of the primary data produced in medical imaging is now digital. Therefore, the need for effective digital medical image analysis has increased in the medical community.

The new technologies such as deep learning algorithms have remarkable potential in the field of medical imaging. Deep learning models can provide better analyses of digital medical images and thus may serve as a tool for detection of abnormalities or diagnosis of diseases. Recent applications of deep learning in the field of medical imaging have been demonstrated, such as diagnosis of diabetic retinopathy, age-related macular degeneration, and glaucoma in the field of ophthalmology, diagnosis of lung nodules or lung cancer in the field of respiratory imaging, and diagnosis of breast cancer in breast imaging [1].

Paranasal sinuses are air-filled spaces that surround the

nasal cavity. There are four air-filled spaces named as “ethmoidal”, “sphenoidal”, “frontal”, and “maxillary” sinuses. The sinuses may become infected when blocked and filled with fluid. In the case of infection and inflammation of the mucosal lining of the paranasal sinuses, the condition is called “sinusitis” or “rhinosinusitis” and this condition is considered a common ailment [2]. Sinusitis can be acute or chronic. The Centers for Disease Control and Prevention (CDC) reported that the total number of adults with diagnosed sinusitis is 28.9 million in 2018 and the percent of adults with diagnosed sinusitis is 11.6 % [3]. Chronic rhinosinusitis affects around 10 %–12 % of the European population [4].

Technological advances in medical imaging techniques such as high-resolution computed tomography (CT) and magnetic resonance imaging (MRI) scans have improved the analysis and evaluation of anatomy and pathology of the paranasal sinuses. CT scans allow the details of the osseous anatomy to be seen, and MRI scans give the details of the soft tissues. Although both provide essential information for the interpretation and treatment of various paranasal sinus pathologies, CT scan should always be considered the first choice in sinonasal imaging, and MRI scan is treated as supplementary to CT scan [5].

A simple and efficient system that automatically analyses the conditions of the paranasal sinuses on CT images and that supports physicians in otolaryngology with a high-accuracy diagnosis is still lacking. The article presents the original deep learning-based system for the automatic analysis of paranasal sinuses conditions. The proposed system is the first fully automated algorithm that utilizes a deep learning approach for the analysis of the conditions of the paranasal sinuses to the best of our knowledge. The proposed algorithm primarily decreases the necessary and crucial data during the training phase and maintains a simple automatic segmentation with a considerable improvement in computational cost and complexity. As a result, by reducing the number of images to be analysed on a CT scan for a patient and by providing automatic segmentation for marking and cropping the paranasal sinuses region, the proposed system performs the required segmentation automatically without manual cropping and gives remarkable accuracy in detecting abnormalities in the paranasal sinuses.

In this paper, the purpose of our study is to develop a deep learning-based algorithm for the analyses of the

paranasal sinuses on CT images. The proposed system consists of two main stages: the first stage involves an image processing algorithm for automated segmentation of paranasal sinuses on CT images, and then the second stage is for determining the abnormalities on the sinuses. In the second stage, we use convolutional neural networks (CNN) to train the system.

The rest of the paper is organized as follows. Section II provides an overview of the literature on recent related works. The data preparation, including data collection and pre-processing stages, is given in Section III. The segmentation of the paranasal sinuses is presented in Section IV. In Section V, we introduce a description of the proposed CNN architecture. The experimental results are provided in Section VI. The proposed system performance is discussed in Section VII. Finally, the last section, Section VIII, concludes the work.

II. RELATED WORKS

In medical imaging, one of the most encouraging and attractive hot topics today is deep learning. Recent challenges in deep learning showed that it has the capacity to revolutionize medical diagnostics [1], [6], [7]. Deep learning is likely to play a critical role in image-heavy specialties such as otolaryngology for the detection and classification of disease. Recent studies in the related field of otolaryngology are summarized below.

Deep learning algorithms have been proposed to identify details of bone structures from CT scans [8], [9]. Heutink *et al.* [8] proposed a deep learning framework to automatically segment and measure the human cochlea in CT images, and stated that the algorithm provided accurate measurements of cochlear anatomy. Zhang, Wang, Noble, and Dawant [9] presented a deep CNN-based algorithm for the localization of multiple landmarks in head CTs and to classify CT images in terms of their content. They reached 99.5 % classification accuracy.

Deep neural networks have also been used to investigate the anatomy of nasal cavities and paranasal sinuses. Darknet-19 and You Only Look Once (YOLO) had been used for automatic detection of nasal cavities and paranasal sinuses in [10]. An automated paranasal sinus segmentation method based on a fully convolutional network (FCN) with a probability atlas [11] has been proposed. The segmentation accuracy (Dice coefficient) was found to be about 0.83.

Xu, Wang, Zhou, Liu, Jiang, and Chen [12] proposed an algorithm for automatic segmentation of the maxillary sinus (MS) by combining the Visual Geometry Group (VGG) network and the improved V-Net. The VGG network was used to label CT slices containing the MS region, and the improved V-Net functioned as a segmentation unit. As a classifier unit that decides whether CT slices contain MS region or not, the VGG network had a classification accuracy of 97.04 ± 2.03 %. In the segmentation unit, the segmentation accuracy (Dice coefficient) was 94.40 ± 2.07 %, the Iou (intersection over union) was 90.05 ± 3.26 % and the precision was 94.72 ± 2.64 %.

Ren, Li, Tian, and Li [13] have applied a deep learning architecture for the automatic recognition of inverted papilloma (IP) and nasal polyp (NP) on CT images. The proposed end-to-end deep learning model consists of two

parts: first is for pre-classification, and the second separate networks for differentiating IP and CP. They achieved 89.30 % accuracy in classification.

Jung, Lim, Lee, Cho, and Song [14] have developed an active learning framework for maxillary sinus segmentation. They used a customized 3D nnU-Net on cone-beam computed tomography (CBCT) to segment maxillary sinus into the maxillary bone, air, and lesion. Humphries *et al.* [15] presented a CNN method for fully automatic assessment of paranasal sinus opacification on CT images with truly objective volumetric quantitation of sinonasal inflammation. They expressed that the proposed method provided volumetric opacification scores that are consistent with Lund-Mackay (LM) visual scoring on test images that involve various degrees of sinonasal inflammation.

Parmar *et al.* have used a CNN algorithm to identify middle turbinate pneumatization on coronal sinus CT images [16]. They used Inception-V3 model transfer learning and that re-trained the classification layer. They found a diagnostic accuracy of 81 % (95 % confidence interval: 73.0 % to 89.0 %) with an area under the curve of 0.93. Huang *et al.* [17] performed a study to differentiate the location of the anterior ethmoidal artery adhered to the skull base or within a “mesentery” of the bone on sinus CT scans using the Inception-V3 CNN model. They achieved a total accuracy of 82.7 % (95 % confidence interval = 77.7–87.8), a kappa statistic of 0.62 and an area under the curve of 0.86.

III. DATA PREPARATION

A. Data Collection

The study protocol was approved by the ethics committee of our institution, Gaziantep University. All CT scans were acquired at our university hospital. Experiments were conducted with coronal CT scans from 140 patients. The CT scans were reviewed and interpreted by the physicians, and then the CT scans were classified into two parts: normal sinuses and abnormal sinuses conditions. Totally 72 patients CT scans were used for the training phase and 68 patients CT scans were used for the testing phase in our study.

B. Pre-Processing

In the initial phase, the training dataset was prepared. For this purpose, the coronal CT scan images for the training phase were selected for each patient. In the row CT scans for the patients, there were different amounts of images for each. Five images for each patient were chosen. To identify abnormalities in the paranasal sinuses and differentiate them from normal ones, the position of the head on the images needs to be properly chosen.

The pivot image is selected first. The pivot image refers to the full view from the front of the head. Coronal CT scans contain a variable number of images for each patient. Histograms of the images are computed for each image in order in a CT scan then the differences between the successive images based on histogram evaluations are obtained. It is expected that the histograms are similar in the neighborhood of the pivot image. Then the candidate images for pivot selection are labeled. Next, the pixel intensities averages and standard deviations are evaluated for the candidate images. Those feature vectors are used for precise estimation of the pivot image. Then two before and two after

are selected from consecutive images. Figure 1 shows a representation of selected images for a patient obtained after the pre-processing stage. The middle image appearing in the figure is the pivot image.

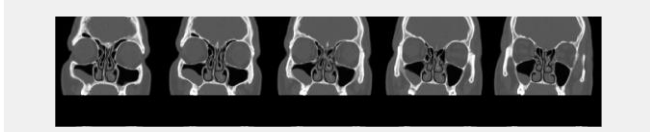


Fig. 1. A representation of selected CT images for a patient in the preparation stage of the training dataset. The middle image is the pivot image.

IV. SEGMENTATION

To format and make row RGB CT images more consistent in the training phase, paranasal sinus segmentation is the next stage after obtaining 5 selected images for each patient. In segmentation for the training dataset, images are first converted to binary image format. In this conversion, a luminance threshold is estimated first. All row CT images are analysed, and corresponding global thresholding values are evaluated. The Otsu's method [18] is used to find a global threshold for each CT image.

The Otsu's method is an iterative algorithm to find the global threshold value where the sum of the foreground and background spreads is at its minimum. In the proposed segmentation algorithm, the universal threshold value was found by averaging the global threshold values. The universal threshold value was computed and finally set to 0.2.

The resulting thresholded images are then filtered by the average filter to remove small-sized areas and noises. The images are then processed along vertical and horizontal runs to locate the paranasal sinus region. The limit pixel coordinate points to label the rectangular area containing paranasal sinuses are selected, and this region is cropped. Finally, the size of the images is adjusted to a fixed value. Bicubic interpolation is used to resizing the images to the fixed value. In the resized image, the output pixel value is a weighted average of pixels in the nearest 4-by-4 neighborhood. It was specifically set to 225×300.

The segmentation operation is illustrated in Fig. 2. The image on the left is one of the selected row CT images, the middle is thresholded and filtered, and the image on the right is the cropped image.



Fig. 2. Segmentation illustration for the preparation of training dataset: (a) row CT image, (b) thresholded and filtered form, and (c) cropped image.

In normal operations (testing phase), the segmentation algorithm is the same as in the training phase. All the defined operations are performed automatically.

V. THE PROPOSED CNN ARCHITECTURE

The proposed algorithm is intended to analyse and classify the conditions of the paranasal sinuses into two parts: normal sinus conditions and abnormal sinus

conditions. Therefore, the problem reduces to a classification problem. We propose a simple CNN architecture for deep learning classification.

The Convolutional Neural Network (CNN) is a type of deep neural network inspired by the biological structure of a visual cortex of the brain and that it tries to imitate how the visual cortex of the brain processes and recognizes the visual data. Thus, CNN architectures are commonly and specifically applied to analyse visual imagery. In deep neural network architectures, involving fully connected neurons, the number of adjustable parameters probably increases quickly if the size of the input is becoming larger. CNN architecture, as a candidate solution, can reduce the number of adjustable parameters with the reduced number of connections, shared weights, and downsampling. A CNN can do this job by the help of multiple layers, including convolutional layers, pooling layers, and fully connected layers. The proposed CNN architecture is summarized in Table I.

TABLE I. PROPOSED CNN ARCHITECTURE.

CNN Layers with specifications:
InputLayer (Image size is 225×300 with grayscale image format)
ConvolutionalLayer (Filter size is 3×3 and the number of filters is 8)
BatchNormalizationLayer
ReLULayer
MaxPoolingLayer (Pool size is 2×2)
ConvolutionalLayer (Filter size is 3×3 and the number of filters is 16)
BatchNormalizationLayer
ReLULayer
MaxPoolingLayer (Pool size is 2×2)
ConvolutionalLayer (Filter size is 3×3 and the number of filters is 32)
BatchNormalizationLayer
ReLULayer
FullyConnectedLayer (Output size is 2)
SoftmaxLayer
ClassificationLayer

The first layer in the proposed architecture is the image input layer which feeds the network with 2-D images and makes data normalization. Then, a 2-D convolutional layer is placed for applying sliding convolutional filters to the input. To speed up training, a batch normalization layer is used between the convolutional layer and the ReLU layer. The next activation layer, the ReLU layer, performs a threshold operation on each element of the input, where any value less than zero is set to zero. Then a max pooling layer is set to perform downsampling and computing the maximum of each region. Then the convolutional layer with doubling the filters, the normalization layer, the ReLU layer, and the max pooling layer are followed. Next, again a convolutional layer with tripling filters, normalization layer, and ReLU layer are placed. Then a fully connected layer is inserted to connect the neurons between the layers. This layer consists of weights and biases and that the layer performs multiplication of the inputs by the weights and then addition of biases. Next, a softmax layer is placed which applies a softmax function to the input. Finally, a classification layer is placed and it computes the cross entropy loss for the two-class classification problem. Hyperparameters optimization is done manually. Specifically, in both convolutional layers and pooling layers, filter sizes, number of filters, and pool sizes are adjusted with the random search. The best training and validation accuracy values are obtained with the proposed CNN

architecture.

VI. PERFORMANCE EVALUATION

A. Performance Evaluation Measures

The output of the proposed system is set to a decision or prediction that is labeled as either a “normal sinus condition” or an “abnormal sinus condition”, so this process is viewed as a binary classification problem. Thus, in this study, the performance of the system is represented in a confusion matrix attached with accuracy, precision, sensitivity (recall), specificity, and F_1 scores.

The confusion matrix is a way to visualize and summarize the performance of a classification problem. Figure 3 demonstrates the confusion matrix. In this study, TP is the test result that shows the number of test samples that have been correctly labeled as normal sinus condition, FP is the test result that determines the number of test samples that have been erroneously identified as normal sinus condition, FN is the test result that indicates falsely missed normal sinus conditions, and TN represents the number of test samples estimated as abnormal sinus condition where they are actually abnormal sinus conditions.

		Predicted	
		Negative	Positive
Actual	Negative	True Negatives (TN)	False Positives (FP) Type I Error
	Positive	False Negatives (FN) Type II Error	True Positives (TP)

Fig. 3. Confusion matrix.

The related performance metrics, including accuracy, precision, sensitivity (recall), specificity, and F_1 scores, are calculated as shown below:

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN}, \quad (1)$$

$$precision = \frac{TP}{TP + FP}, \quad (2)$$

$$sensitivity = \frac{TP}{TP + FN}, \quad (3)$$

$$specificity = \frac{TN}{TN + FP}, \quad (4)$$

$$F_1 = \frac{2 \times sensitivity \times precision}{sensitivity + precision}. \quad (5)$$

B. Experimental Results

To evaluate the performance of the proposed system, a series of experiments was conducted using a variety of CT images. All experiments were performed using MATLAB R2020b and a computer with an Intel Pentium processor at 1.60 GHz and 4 GB of memory. The experiments include both the training and the testing phases and were conducted with coronal CT scans from a total of 140 patients, 72 for the training phase and 68 for the testing phase.

For the training phase, the CT scans of 72 patients were first labeled as “normal sinuses conditions” and “abnormal sinuses conditions” by experienced physicians and then the

CT scans were classified as normal and abnormal. As described in the pre-processing stage section, five images, a pivot and four more images in its neighborhood, for each patient, were chosen. Thus, 360 images were used for training. Figure 4 shows an example of training dataset members labeled as “normal” and similarly, Fig. 5 illustrates an example of training dataset members labeled as “abnormal” obtained after pre-processing stage. The images in the middle positions are the selected pivot images and the others are images in the right and left neighborhoods of the pivot images.



Fig. 4. An example of members of the training dataset labeled as “normal”. The image in the middle is the pivot image, and the others are two before and two after images in the left and right neighborhoods of the pivot image.



Fig. 5. An example of members of the training dataset labeled as “abnormal”. The image in the middle is the pivot image, and the others are two before and two after images in the left and right neighborhoods of the pivot image.

After the pre-processing stage, segmentation was done to obtain and then crop the paranasal sinuses region on each selected image. In the final stage, the sizes of the cropped images were adjusted to a fixed value, which is 225×300. Some examples of images after segmentation stage are given in Fig. 6 and Fig. 7. Image samples representing the region of interest (i.e., paranasal sinuses region) labeled as “normal” are presented in Fig. 6. and image samples labeled as “abnormal” are represented in Fig. 7, respectively. Segmentation contains thresholding for RGB to binary conversion, filtering by the average filter to remove small-sized areas and noises and cropping the paranasal sinuses region.



Fig. 6. Image samples after the segmentation stage representing the region of interest (i.e., the paranasal sinus region) labeled as “normal”.



Fig. 7. Image samples after the segmentation stage representing the region of interest (i.e., the paranasal sinus region) labeled as “abnormal”.

The stochastic gradient descent with momentum (SGDM) optimizer with a 0.01 initial learning rate was used to update the network parameters in the training stage of the proposed architecture.

For testing the proposed system, coronal CT scans belonging to 68 patients were used. The pre-processing stage was applied to the test CT scans and five CT images were selected for each. The selected images are then

automatically segmented to produce more appropriate test images to analyse the conditions of the paranasal sinuses. The test CT scans were also reviewed and interpreted by experienced physicians to compare the results obtained from the proposed system. 32 patient images were labeled as “normal” and the rest 36 images of 68 were labeled as “abnormal”.

In the first phase of the experiments, the system was tested for all images selected for each patient individually. Therefore, $68 \times 5 = 340$ images were tested and their prediction results were recorded. The corresponding outcome results and performance indexes were listed in Table II and Table III, respectively.

TABLE II. OUTCOME RESULTS OF PREDICTION WHEN IMAGES ARE TAKEN INDIVIDUALLY.

	TP	TN	FN	FP
160 images labeled as “normal”	148		12	
180 images labeled as “abnormal”		150		30
Totally 340 images	148	150	12	30

TABLE III. EVALUATION INDEXES WHEN IMAGES ARE TAKEN INDIVIDUALLY.

Accuracy	87.64 %
Precision	83.14 %
Sensitivity	92.50 %
Specificity	83.33 %
F ₁ Score	87.57 %

In the second phase of the experiments, the performance of the system was measured by combining the individual prediction results of every selected image of each patient. The majority decision rule was applied for decision fusion. Within five individual predictions, the majority vote is taken as the common decision.

The experimental results and the values of the performance indexes with combining the individual decisions by the majority rule are given in Table IV and Table V, respectively.

TABLE IV. OUTCOME RESULTS OF PREDICTION WHEN MULTIPLE IMAGES DECISIONS ARE FUSED.

	TP	TN	FN	FP
32 patients CT images labeled as “normal”	32		0	
36 patients CT images labeled as “abnormal”		35		1
Totally 68 patient CT images	32	35	0	1

TABLE V. EVALUATION INDEXES WHEN MULTIPLE IMAGES DECISIONS ARE FUSED.

Accuracy	98.52 %
Precision	96.96 %
Sensitivity	100 %
Specificity	97.22 %
F ₁ Score	98.45 %

VII. DISCUSSION

In the first stage of the experiments, when the system was tested for each selected image of each patient individually, it was observed that Type II Error (False Negatives) was comparatively low, while Type I Error (False Positives) was high.

When the majority rule was applied in the second stage of the experiments, it was demonstrated that Type II Error was

eliminated and that Type I Error was diminished.

When comparing Table III and Table V, it was seen that when the selected images individual decisions were fused by the majority rule, all performance metrics became higher.

To see the comparative performance of the proposed method, previously studied approaches were investigated but it was seen that there was no direct related study in the literature. On the other hand, somewhat similar studies are mentioned here.

In [12], automatic segmentation of maxillary sinus (MS) was proposed by combining the VGG network and the improved V-Net. The proposed algorithm reached a classification accuracy of 97.04 ± 2.03 % and a segmentation accuracy (Dice coefficient) of 94.40 ± 2.07 %. Comparing our approach with this work, our proposed method makes automatic segmentation of the paranasal sinus region with perfect accuracy.

In [13], a deep learning framework has been introduced for automatic recognition of inverted papilloma (IP) and nasal polyp (NP) on CT images. It was declared to reach 89.30 % accuracy in classification. When compared our method with this study, our method outperforms with an accuracy of 98.52 % in classifying the conditions of the paranasal sinuses.

Another study [19] related to automatic recognition and volume calculation for the inferior turbinate and maxillary sinus using image processing techniques has been presented. It was found that the accuracy and sensitivity results on the recognition stage for the inferior turbinate and the maxillary sinus were 96.3 % and 95.1 %, respectively. Our method achieved perfect recognition of CT images related to the paranasal sinus.

VIII. CONCLUSIONS

In this study, we propose an approach for the analyses of the paranasal sinuses conditions on CT images. The proposed approach automatically reduces the number of images in a CT scan and then makes automatic segmentation for cropping the required region of interest (i.e., paranasal sinuses region). The proposed system significantly reduces the data in the training phase while maintaining high-accuracy performance.

The experiments conducted have demonstrated that the proposed system yields outstanding performance in detecting abnormalities in the sinuses. The proposed approach has been tested on real CT images, and it is shown that with the majority rule on the final decision stage, we achieved an accuracy rate of 98.52 % and a sensitivity of 100 %. The method proposed also delivered 96.96 % precision, 97.22 % specificity, and 98.45 % F₁ score. As a future study, we will try to adapt the proposed approach to the diagnostic investigations of some specific diseases.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

REFERENCES

- [1] R. Aggarwal, V. Sounderajah, G. Martin *et al.*, “Diagnostic accuracy of deep learning in medical imaging: A systematic review and meta-analysis”, *npi Digital Medicine*, vol. 4, art. no. 65, 2021. DOI: 10.1038/s41746-021-00438-z.

- [2] L. J. Fagnan, "Acute sinusitis: A cost-effective approach to diagnosis and treatment", *American Family Physician*, vol. 58, no. 8, pp. 1795–1802, 1805–1806, 1998.
- [3] "Summary Health Statistics", National Health Interview Survey, U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics, 2018, Tables A-2b, A-2c.
- [4] D. Hastan *et al.*, "Chronic rhinosinusitis in Europe – An underestimated disease. A GA²LEN study", *Allergy*, vol. 66, no. 9, pp. 1216–1223, 2011. DOI: 10.1111/j.1398-9995.2011.02646.x.
- [5] H. B. Eggesbø, "Imaging in sinonasal disorders", in *Sino-Nasal and Olfactory System Disorders*. IntechOpen, 2020. DOI: 10.5772/intechopen.90773.
- [6] Y. Y. M. Aung, D. C. S. Wong, and D. S. W. Ting, "The promise of artificial intelligence: A review of the opportunities and challenges of artificial intelligence in healthcare", *British Medical Bulletin*, vol. 139, no. 1, pp. 4–15, 2021. DOI: 10.1093/bmb/ldab016.
- [7] X. Liu, L. Faes, A. U. Kale *et al.*, "A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: A systematic review and meta-analysis", *Lancet Digital Health*, vol. 1, no. 6, pp. E271–E297, 2019. DOI: 10.1016/S2589-7500(19)30123-2.
- [8] F. Heutink, V. Koch, B. Verbist, W. J. van der Woude, E. Mylanus, W. Huinck, I. Sechopoulos, and M. Caballo, "Multi-Scale deep learning framework for cochlea localization, segmentation and analysis on clinical ultra-high-resolution CT images", *Computer Methods and Programs in Biomedicine*, vol. 191, art. 105387, 2020. DOI: 10.1016/j.cmpb.2020.105387.
- [9] D. Zhang, J. Wang, J. H. Noble, and B. M. Dawant, "HeadLocNet: Deep convolutional neural networks for accurate classification and multi-landmark localization of head CTs", *Medical Image Analysis*, vol. 61, art. 101659, 2020. DOI: 10.1016/j.media.2020.101659.
- [10] C. O. Laura, P. Hofmann, K. Drechsler, and S. Wesarg, "Automatic detection of the nasal cavities and paranasal sinuses using deep neural networks", in *Proc. of IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, 2019, pp. 1154–1157. DOI: 10.1109/ISBI.2019.8759481.
- [11] Y. Iwamoto *et al.*, "Automatic segmentation of the paranasal sinus from computer tomography images using a probabilistic atlas and a fully convolutional network", in *Proc. of 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2019, pp. 2789–2792. DOI: 10.1109/EMBC.2019.8856703.
- [12] J. Xu, S. Wang, Z. Zhou, J. Liu, X. Jiang, and X. Chen, "Automatic CT image segmentation of maxillary sinus based on VGG network and improved V-Net", *International Journal of Computer Assisted Radiology and Surgery*, vol. 15, pp. 1457–1465, 2020. DOI: 10.1007/s11548-020-02228-6.
- [13] T. Ren, X. Li, Y. Tian, and W. Li, "Deep learning framework for preoperative recognition of inverted papilloma and nasal polyp", *IEEE Access*, vol. 9, pp. 120502–120511, 2021. DOI: 10.1109/ACCESS.2021.3099687.
- [14] S.-K. Jung, H.-K. Lim, S. Lee, Y. Cho, and I.-S. Song, "Deep active learning for automatic segmentation of maxillary sinus lesions using a convolutional neural network", *Diagnostics*, vol. 11, no. 4, p. 688, 2021. DOI: 10.3390/diagnostics11040688.
- [15] S. M. Humphries, J. P. Centeno, A. M. Notary *et al.*, "Volumetric assessment of paranasal sinus opacification on computed tomography can be automated using a convolutional neural network", *International Forum of Allergy & Rhinology*, vol. 10, no. 11, pp. 1218–1225, 2020. DOI: 10.1002/alr.22588.
- [16] P. Parmar, A.-R. Habib, D. Mendis, A. Daniel, M. Duvnjak *et al.*, "An artificial intelligence algorithm that identifies middle turbinate pneumatization (concha bullosa) on sinus computed tomography scans", *The Journal of Laryngology & Otology*, vol. 134, no. 4, pp. 328–331, 2020. DOI: 10.1017/S0022215120000444.
- [17] J. Huang, A.-R. Habib, D. Mendis, J. Chong, M. Smith *et al.*, "An artificial intelligence algorithm that differentiates anterior ethmoidal artery location on sinus computed tomography scans", *The Journal of Laryngology & Otology*, vol. 134, no. 1, pp. 52–55, 2020. DOI: 10.1017/S0022215119002536.
- [18] N. Otsu, "A threshold selection method from gray-level histograms", *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979. DOI: 10.1109/TSMC.1979.4310076.
- [19] C.-F. J. Kuo, Y.-S. Leu, D.-J. Hu, C.-C. Huang, J.-J. Siao, and K. B. P. Leon, "Application of intelligent automatic segmentation and 3D reconstruction of inferior turbinate and maxillary sinus from computed tomography and analyze the relationship between volume and nasal lesion", *Biomedical Signal Processing and Control*, vol. 57, art. 101660, 2020. DOI: 10.1016/j.bspc.2019.101660.



This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution 4.0 (CC BY 4.0) license (<http://creativecommons.org/licenses/by/4.0/>).