

Differentiation of eosinophilic and non-eosinophilic chronic rhinosinusitis on preoperative computed tomography using deep learning

Hongli Hua¹, Song Li¹, Yu Xu¹, Shiming Chen¹, Yonggang Kong¹, Rui Yang¹, Yu-Qin Deng¹, and Zezhang Tao¹

¹Renmin Hospital of Wuhan University

April 12, 2022

Abstract

Objective: This study aimed to develop deep learning (DL) models for differentiating between eosinophilic chronic rhinosinusitis (ECRS) and non-eosinophilic chronic rhinosinusitis (NECRS) on preoperative computed tomography (CT). **Methods:** A total of 878 chronic rhinosinusitis (CRS) patients undergoing nasal endoscopic surgery were included. Axial spiral CT images were pre-processed and used to build the dataset. Two semantic segmentation models based on U-net and Deeplabv3 were trained to segment sinus area in CT images. All patient images were segmented using the better-performing segmentation model and used for training and validation of the transferred efficientnet_b0, resnet50, inception_resnet_v2, and Xception neural networks. Additionally, we evaluated the performances of the models trained using each image and each patient as a unit. The precision of each model was assessed based on the receiver operating characteristic curve. Further, we analyzed the confusion matrix, accuracy, and interpretability of each model. **Results:** The Dice coefficients of U-net and Deeplabv3 were 0.953 and 0.961, respectively. The average area under the curve and mean accuracy values of the four networks were 0.848 and 0.762 for models trained using a single image as a unit, while the corresponding values for models trained using each patient as a unit were 0.853 and 0.893, respectively. The generated Grad-Cams showed good interpretability. **Conclusion:** Combining semantic segmentation with classification networks could effectively distinguish between patients with ECRS and NECRS based on preoperative sinus CT images. Furthermore, labeling each patient to build a dataset for classification may be more reliable than labeling each medical image.

Abstract

Objective : This study aimed to develop deep learning (DL) models for differentiating between eosinophilic chronic rhinosinusitis (ECRS) and non-eosinophilic chronic rhinosinusitis (NECRS) on preoperative computed tomography (CT).

Methods : A total of 878 chronic rhinosinusitis (CRS) patients undergoing nasal endoscopic surgery were included. Axial spiral CT images were pre-processed and used to build the dataset. Two semantic segmentation models based on U-net and Deeplabv3 were trained to segment sinus area in CT images. All patient images were segmented using the better-performing segmentation model and used for training and validation of the transferred efficientnet_b0, resnet50, inception_resnet_v2, and Xception neural networks. Additionally, we evaluated the performances of the models trained using each image and each patient as a unit. The precision of each model was assessed based on the receiver operating characteristic curve. Further, we analyzed the confusion matrix, accuracy, and interpretability of each model.

Results : The Dice coefficients of U-net and Deeplabv3 were 0.953 and 0.961, respectively. The average area under the curve and mean accuracy values of the four networks were 0.848 and 0.762 for models trained

using a single image as a unit, while the corresponding values for models trained using each patient as a unit were 0.853 and 0.893, respectively. The generated Grad-Cams showed good interpretability.

Conclusion : Combining semantic segmentation with classification networks could effectively distinguish between patients with ECRS and NECRS based on preoperative sinus CT images. Furthermore, labeling each patient to build a dataset for classification may be more reliable than labeling each medical image.

Keywords: deep learning; eosinophil; computed tomography; rhinosinusitis; differentiation

Key points:

1. Accurate preoperative prediction of the CRS type is necessary for predicting postoperative outcomes and administering personalized treatment
2. We combined semantic segmentation with classification network for sinus region segmentation and differentiating between ECRS and NECRS on preoperative CT
3. The segmentation model could accurately segment the sinus region in CT images
4. The classification model could effectively distinguish between patients with ECRS and NECRS based on preoperative sinus CT images
5. Labeling each patient to build a dataset for classification may be more reliable than labeling each medical image.

1. Introduction

Chronic rhinosinusitis (CRS), which is characterized by inflammatory dysregulation of the nasal and paranasal mucosa for ≥ 12 consecutive weeks, is among the most common otolaryngological diseases, with a prevalence rate of 8% in the general population [1]. Currently, CRS is classified into two types based on the presence or absence of nasal polyps; further, it can be classified as eosinophilic sinusitis (ECRS) and non-eosinophilic sinusitis (NECRS) based on the eosinophilic infiltration level in the nasal mucosa or polyp. There has been recent progress in the elucidation of the pathogenesis of CRS. There are several limitations in the traditional phenotypic classification of CRS into CRSwNP and CRSsNP. First, this classification does not account for pathophysiological differences; moreover, immune system cells are crucially involved in inflammatory mechanisms [2-4]. Many patients with extensive tissue eosinophil infiltration do not present polyp-like degeneration, with the incidence rate being approximately 27.5% [5]. Further, even patients with polypoid changes may predominantly present non-eosinophilic inflammation [6]. Classification of CRS into ECRS and NECRS is more likely to reflect the underlying inflammatory process. Additionally, the level of eosinophil infiltration in CRS lesions is strongly associated with postoperative outcomes [7, 8].

ECRS has a recurrence rate as high as 98.5%; additionally, it is the main cause of refractory sinusitis recurrence [9]. The gold standard for ECSR diagnosis is histopathological examination; however, it is invasive. Recently, the main treatments for ECSR and NECRS are contoured endoscopic nasal surgery and functional endoscopic nasal surgery, respectively [10, 11]. Given the high postoperative recurrence rate of ECRS and differences in the surgical methods for both CRS types, accurate preoperative prediction of the CRS type is necessary for predicting postoperative outcomes and administering personalized treatment. Eosinophil levels in the peripheral blood are associated with the eosinophil infiltration degree in the nasal sinus mucosa [12]. However, allergies, autoimmune diseases, drug reactions, parasitic infections, and corticosteroid treatment can alter circulating eosinophil levels. Additionally, increased eosinophil levels in peripheral blood do not necessarily reflect an increase in tissue eosinophils; moreover, the predictive utility of blood eosinophil levels for ECRS remains limited [13, 14]. There has been extensive research on the preoperative predictive utility of exhaled nitric oxide levels, serum total immunoglobulin E (IgE), specific IgE, and skin prick tests for ECRS and NECRS; however, they have low sensitivity and specificity [15, 16]. There are several differences in sinus computed tomography (CT) findings between patients with ECRS and NECRS. Early-stage ECRS often presents as ethmoid sinus lesions and mild maxillary sinus lesions, while NECRS often presents with maxillary sinus lesions. However, these characteristics do not effectively distinguish ECRS from NECRS [8, 17].

Considering the strong feature extraction and screening ability of artificial intelligence, applying artificial intelligence technology to sinus CT-assisted ECRS diagnosis may allow accurate preoperative prediction of ECRS. CRS lesions are limited to the nasal cavity and paranasal sinus areas in each disease type; moreover, the sinus area comprises a small part of the sinus CT images, with the surrounding tissue structure being complex. Therefore, noise information in the whole image increases the required training data. Manually delimiting the regions of interest (ROIs) is inconsistent with the original intention of artificial intelligence; further, its huge sketching workload greatly reduces its clinical application value. Therefore, we aimed to develop a segmentation model that allowed automatic segmentation of the sinus region. Moreover, since different networks have certain data "preferences," we used four common classification networks to train the segmented images.

2. Methods

Patients

We included 878 patients with CRS who underwent nasal endoscopic surgery in the Department of Otolaryngology, Head and Neck Surgery of our hospital, from October 2016 to June 2021. The study protocol was approved by the institutional review board, which waived the requirement for informed consent. The inclusion criteria were as follows: diagnosis of chronic rhinosinusitis based on the European Position Paper on Rhinosinusitis and Nasal Polyps 2020 and having undergone a sinus CT scan within 2 preoperative weeks. The exclusion criteria were as follows: fungal sinusitis identified on pathological examination; sinus cystic fibrosis; unclear CT images; posterior nostril obstruction; a history of radiotherapy for the head and neck.

2.2 Histological examination

Intraoperative nasal polyps and pathological nasal mucosal tissues were selected for fixation, embedding, and sectioning. Sections underwent conventional hematoxylin-eosin staining, followed by observation under a high-magnification field (HPF) of $400\times$. In each field, ten fields were randomly selected for observation and the eosinophils and inflammatory cells were counted. The mean counts of eosinophils and inflammatory cells were calculated for each field. ECRS was diagnosed when the eosinophil to inflammatory cell count ratio (Eos%) was $\geq 10\%$, and NECRS was diagnosed when the Eos% was $<10\%$ [18].

2.3 Image collection and Pre-processing

The patients underwent sinus scanning with 64-slice spiral CT using the following parameters: tube voltage, 120 kV and tube current, 200 mA. There was a soft tissue window (window width: 300–350 HU, window position: 30–50 HU) and bone window (window width: 1000–2000 HU, window position: 300–350 HU). Exported CT images were saved in the DICOM format and converted to the PNG format for segmentation and classification model training; additionally, the slices of axial CT images with lesions were used to build the dataset.

To establish segmentation dataset, 1,365 images were randomly selected from patients with ECRS and NECRS; additionally, the nasal cavity and sinus regions in the image were marked using the ITK-SNAP software. We defined the nasal cavity and sinus regions as follows: the inferior side is the plane where the bilateral inferior turbinates begin to appear, while the superior side is the plane where the straight gyrus and orbital gyrus of the frontal lobe begin to appear. At the maxillary level, the anterior side is the anterior nostril or nasal bone, the lateral sides are the anterior lateral wall and posterior wall of the maxillary or lateral wall of the sphenoid, and the posterior side is the posterior nostril or posterior wall of the sphenoid. At the ethmoid level, the anterior side is the nasal or frontal bone, the lateral sides are the lateral walls of the ethmoid and sphenoid, and the posterior side is the posterior wall of the posterior ethmoid or sphenoid (Figure 1).

The dataset of the classification model comprised 56,892 images, including ECRS (343 patients and 22,671 images) and NECRS (535 patients and 34,221 images). Patients in each category were allocated to the training and validation cohorts at a ratio of 4:1. Since all sinus CT slices did not show between-disease differences and we sought to achieve accurate classification, we constructed two datasets using each patient

and each image as a unit. When using individual images, each image was labeled and input into the classification network for learning. When using patients as units, we labeled each patient, set the average probability value of all the images obtained from each patient as the patient’s probability value, and input it into the model for learning.

2.4 Network Architecture

Our compiling platform was based on the Pytorch library (version 1.9.0) with CUDA (version 10.0) for GPU (NVIDIA T4) acceleration on a Windows operating system (Server 2019 data center version 64 bit). We transformed the U-Net and Deeplabv3 networks to build semantic segmentation models. Additionally, 1,365 images were used to construct datasets, which were randomly divided into the training and validation cohorts at a ratio of 4:1. The model was trained using the RMSprop optimizer, with the batch size and initial learning rate set at 32 and 0.001, respectively. Both semantic segmentation models were trained for 20 epochs. We selected the model with the best performance and used a rectangular segmentation method to segment the nasal cavity and sinus areas on the CT images (Figure 2).

Since different neural networks may have different preferences for the data distribution, type, and dataset size, we used four common pre-trained classification networks for model building, including efficientnet_b0, resnet50, inception_resnet_v2, and Xception neural networks, to avoid model inclination. These networks were trained using the SGD optimizer with a batch size of 32; furthermore, each model was trained for 40 epochs.

Statistical Analyses

Statistical analyses were performed using SPSS22.0 statistical software. Normally distributed measurement data are expressed as $(\bar{X} \pm S)$ and were analyzed using an independent sample t-test. Counting data are expressed as frequencies and were analyzed using the chi-square test. Statistical significance was set at $P < 0.05$. Segmentation model performance was evaluated using Dice similarity coefficients, and classification model was evaluated using the ROC curve, accuracy, and confusion matrix; moreover, Grad-Cams were generated by extracting feature maps from the final convolutional layers to verify the reliability of the model.

3. Results

3.1 Clinical Characteristics of the Study Population

After screening, we enrolled 878 eligible patients in the training ($n = 702$) and validation ($n = 176$) cohorts of the classification model. There were no significant between-cohort differences in the general information.

3.2 Results of the Semantic Segmentation Models

After training for 20 epochs, the performance of U-net and Deeplabv3 quickly stabilized; moreover, they had Dice coefficients of 0.953 and 0.961, respectively (Figure 4). Deeplabv3 performed slightly better than U-net. We used the Deeplabv3 network trained to segment the sinus region in CT images.

3.3 Performance of the Classification Models

When trained using a single image as a unit, the areas under the curve (AUCs) of the efficientnet_b0, resnet50, inception_resnet_v2, and Xception networks were 0.84, 0.86, 0.83, and 0.86, respectively. When trained using each patient as a unit, the AUCs of the four neural networks increased to 0.89, 0.90, 0.88, and 0.90, respectively (Figure 5). The accuracy of the training process reflects the overall classification performance of the DL models (Figure: 6).

We incorporated a confusion matrix for class-wise comparison to evaluate whether the model could reliably detect and classify objects. As shown in Figures 7 and 8, the specificity and sensitivity of the four networks were higher when using each patient, rather than single images, as units.

The presentation of Grad-Cams allows elucidation of how the DL network captured image features for prediction and resolves doubts regarding the network’s ability to learn in the appropriate direction. Yellow areas shown in the Grad-Cams had the strongest correlation with the classification. For patients with ECRS, the yellow areas represent characteristics associated with an increased risk of ECRS. The efficientnet.b0 network was used as an example. Figures 9, 10, 11, and 12 show Grad-Cams for both the bone window and soft tissue window images of the patients. For both diseases, yellow areas were concentrated in areas with lesions, which is consistent with our medical experience.

4. Discussion

The recent advances in computer vision technology have allowed rapid development of artificial intelligence technology for image processing, automatic recognition, classification, and segmentation. This has led to more efficient extraction of large amounts of feature information from medical images. Medical image evaluation is not limited to qualitative disease diagnosis; rather, it also includes the acquisition and analysis of rich quantitative information to provide data regarding disease severity, optimal treatment options, and patient outcomes. In our study, we introduced semantic segmentation and classification networks to achieve effective classification of ECRS and NECRS based on preoperative sinus CT images. In contrast to patients with NECRS, patients with ECRS require repeated administration of corticosteroid therapy and multiple revision surgeries to achieve disease control [19]. Specifically, the therapeutic strategy involves local treatment through high-volume corticosteroid irrigation in a widely open surgical cavity [20, 21]. Given the high postoperative recurrence rate and drug resistance of ECRS, targeted biotherapeutics targeting the TH2 inflammatory mediators interleukin (IL)-5, IL-4, and IgE have been recently developed as a potential therapeutic approach [22, 23]. Increased eosinophil infiltration in the nasal polyps is an important biomarker for asthma development after nasal endoscopy [24]. Given the differences in surgical modalities and therapeutic agents between ECRS and NECRS, as well as the risk of postoperative asthma and recurrence, accurate preoperative diagnosis of ECRS is crucial for determining the optimal treatment plan. Our model showed satisfactory performance and could provide valuable information for accurate diagnosis and treatment.

We introduced a semantic segmentation model that could automatically segment the sinus and nasal area from a complex image for classification model learning. Medical image reading requires systematic anatomical knowledge. Accordingly, an excellent radiologist should master the anatomical and disease characteristics, including hidden prior knowledge regarding medical images. Traditional machine learning classification models, including the classic cat and dog recognition model, label images and input them into the network for training. Here, pixels representing the animal can appear anywhere in an image; further, the disease distribution is usually located in the corresponding anatomical region. Identifying the corresponding region for model learning can eliminate surrounding interference factors and prevent failure resulting from the lack of prior knowledge. We previously confirmed that a training method based on anatomical partitioning could effectively improve model performance and interpretability when the dataset was reduced [25]. Additionally, rectangular segmentation was used as the segmentation method. Here, irregular images should be filled with "0" pixels around them and they visually appear as black. Compared with irregular segmentation, rectangular segmentation retains the structure around the sinus, which is more consistent with the real situation.

Previous studies on DL application in the medical field have mostly labeled single images and input them into the network for learning. However, medical images are unique. Specifically, for a patient with a certain classification feature, not all image slices contain information for classification. For example, since tumors are heterogeneous, the patient’s prognosis or risk of metastasis cannot be attributed to each lesion slice. Similarly, the features in some CT image slices might not show differences between patients with ECRS and NECRS. The classification results of a single image may have insufficient predictive utility. Each patient has multiple images; moreover, image classifications may differ for the same patient, which affects the outcomes. We used the traditional learning method of labeling individual images and attempted to label each patient. Our findings demonstrated that the dataset composed of each patient as a unit allowed significantly better model performance than the dataset composed of a single image as a unit. This confirms the hypothesis that not all slices show between-disease differences; moreover, it demonstrates the correctness and reliability of

constructing datasets for model learning based on each patient as a unit.

5. Limitations

First, since we could only obtain information regarding the extent of tissue eosinophil infiltration from postoperative patient specimens, the data only included patients who underwent nasal endoscopic surgery, while patients with well-managed CRS without surgical interventions were excluded. Second, although our internal validation results showed excellent model performance, we did not perform external validation to confirm the universality of the model.

6. Conclusion

Our findings demonstrated that combining semantic segmentation with classification networks could effectively distinguish between patients with ECRS and NECRS based on preoperative sinus CT images. This may facilitate preoperative prediction of postoperative recurrence and selection of optimal treatment strategies. Moreover, labeling each patient to build a dataset for classification may be more reliable than labeling individual images.

Reference

- [1] Z. Liu, J. Chen, L. Cheng, et al. Chinese Society of Allergy and Chinese Society of Otorhinolaryngology-Head and Neck Surgery Guideline for Chronic Rhinosinusitis, *Allergy Asthma Immunol Res* 12(2) (2020) 176-237.
- [2] W.J. Fokkens, V.J. Lund, C. Hopkins, et al. European Position Paper on Rhinosinusitis and Nasal Polyps 2020 Executive summary of EPOS 2020 including integrated care pathways, *Rhinology* 58(Suppl S29) (2020) 1-464.
- [3] E. De Corso, S. Baroni, M. Battista, et al. Nasal fluid release of eotaxin-3 and eotaxin-2 in persistent sinonasal eosinophilic inflammation, *Int Forum Allergy Rhinol* 4(8) (2014) 617-24.
- [4] E. De Corso, S. Baroni, D. Lucidi, et al. Nasal lavage levels of granulocyte-macrophage colony-stimulating factor and chronic nasal hypereosinophilia, *Int Forum Allergy Rhinol* 5(6) (2015) 557-62.
- [5] K. Snidvongs, D. Chin, R. Sacks, et al. Eosinophilic rhinosinusitis is not a disease of ostiomeatal occlusion, *Laryngoscope* 123(5) (2013) 1070-4.
- [6] G.M. Oakley, J.M. Christensen, R. Sacks, et al. Characteristics of macrolide responders in persistent post-surgical rhinosinusitis, *Rhinology* 56(2) (2018) 111-117.
- [7] S. Vlaminc, T. Vauterin, P.W. Hellings, et al. The importance of local eosinophilia in the surgical outcome of chronic rhinosinusitis: a 3-year prospective observational study, *Am J Rhinol Allergy* 28(3) (2014) 260-4.
- [8] J. Ishitoya, Y. Sakuma, M. Tsukuda, Eosinophilic chronic rhinosinusitis in Japan, *Allergol Int* 59(3) (2010) 239-245.
- [9] *Rhinology/Allergy, Otolaryngol Head Neck Surg* 155(1-suppl) (2016) P28-P31.
- [10] C. Bachert, D.L. Hamilos, Are Antibiotics Useful for Chronic Rhinosinusitis?, *J Allergy Clin Immunol Pract* 4(4) (2016) 629-38.
- [11] M.E. Cornet, C. Georgalas, S.M. Reinartz, et al. Long-term results of functional endoscopic sinus surgery in children with chronic rhinosinusitis with nasal polyps, *Rhinology* 51(4) (2013) 328-34.
- [12] J. Ho, A.W. Hamizan, R. Alvarado, et al. Systemic Predictors of Eosinophilic Chronic Rhinosinusitis, *Am J Rhinol Allergy* 32(4) (2018) 252-257.
- [13] A. Ganti, H.N. Kuhar, M. Eggerstedt, et al. The Association of Serum Eosinophilia with Structured Histopathology in Chronic Rhinosinusitis, *Ann Otol Rhinol Laryngol* 129(5) (2020) 512-516.

- [14] S.A. Gitomer, C.R. Fountain, T.T. Kingdom, et al. Clinical Examination of Tissue Eosinophilia in Patients with Chronic Rhinosinusitis and Nasal Polyposis, *Otolaryngol Head Neck Surg* 155(1) (2016) 173-8.
- [15] K. Yoshida, T. Takabayashi, Y. Imoto, et al. Reduced nasal nitric oxide levels in patients with eosinophilic chronic rhinosinusitis, *Allergol Int* 68(2) (2019) 225-232.
- [16] J. Ho, P. Earls, R.J. Harvey, Systemic biomarkers of eosinophilic chronic rhinosinusitis, *Curr Opin Allergy Clin Immunol* 20(1) (2020) 23-29.
- [17] E.T. Wang, Y. Zheng, P.F. Liu, et al. Eosinophilic chronic rhinosinusitis in East Asians, *World J Clin Cases* 2(12) (2014) 873-82.
- [18] P.P. Cao, H.B. Li, B.F. Wang, et al. Distinct immunopathologic characteristics of various types of chronic rhinosinusitis in adult Chinese, *J Allergy Clin Immunol* 124(3) (2009) 478-84, 484 e1-2.
- [19] C.L. Wu, T.J. Lee, C.C. Huang, et al. Clinical predictors of revision surgery for chronic rhinosinusitis with nasal polyposis within 5-year follow-up, *Am J Otolaryngol* 41(6) (2020) 102654.
- [20] R.J. Harvey, K. Snidvongs, L.H. Kalish, et al. Corticosteroid nasal irrigations are more effective than simple sprays in a randomized double-blinded placebo-controlled trial for chronic rhinosinusitis after sinus surgery, *Int Forum Allergy Rhinol* 8(4) (2018) 461-470.
- [21] D. Chin, R.J. Harvey, Nasal polyposis: an inflammatory condition requiring effective anti-inflammatory treatment, *Curr Opin Otolaryngol Head Neck Surg* 21(1) (2013) 23-30.
- [22] K.A. Smith, A. Pulsipher, D.A. Gabrielsen, et al. Biologics in Chronic Rhinosinusitis: An Update and Thoughts for Future Directions, *Am J Rhinol Allergy* 32(5) (2018) 412-423.
- [23] K. Lam, R.C. Kern, A. Luong, Is there a future for biologics in the management of chronic rhinosinusitis?, *Int Forum Allergy Rhinol* 6(9) (2016) 935-42.
- [24] R. Kurokawa, Y. Kanemitsu, K. Fukumitsu, et al. Nasal polyp eosinophilia and FeNO may predict asthma symptoms development after endoscopic sinus surgery in CRS patients without asthma, *J Asthma* (2021) 1-9.
- [25] S. Li, H.L. Hua, F. Li, et al. Anatomical Partition-Based Deep Learning: An Automatic Nasopharyngeal MRI Recognition Scheme, *J Magn Reson Imaging* (2022).

Hosted file

table.docx available at <https://authorea.com/users/475749/articles/564963-differentiation-of-eosinophilic-and-non-eosinophilic-chronic-rhinosinusitis-on-preoperative-computed-tomography-using-deep-learning>

Hosted file

figure.docx available at <https://authorea.com/users/475749/articles/564963-differentiation-of-eosinophilic-and-non-eosinophilic-chronic-rhinosinusitis-on-preoperative-computed-tomography-using-deep-learning>