Original Article

Temporal bone CT-based deep learning models for differential diagnosis of primary ciliary dyskinesia related otitis media and simple otitis media with effusion

Bo Duan1*, Zhuoyao Guo2*, Lili Pan3, Zhengmin Xu1, Wenxia Chen1

1Department of Otolaryngology-Head and Neck Surgery, Children’s Hospital of Fudan University, Shanghai 201102, China; 2Department of Respirology, Children’s Hospital of Fudan University, Shanghai 201102, China; 3Department of Radiology, Children’s Hospital of Fudan University, Shanghai 201102, China. *Equal contributors.

Received April 12, 2022; Accepted May 29, 2022; Epub July 15, 2022; Published July 30, 2022

Abstract: Objective: To investigate the diagnostic value of deep learning (DL) in differentiating otitis media (OM) caused by otitis media with effusion (OME) and primary ciliary dyskinesia (PCD), so as to provide reference for early intervention. Methods: From January 2010 to January 2021, 31 patients with PCD who had temporal bone computed tomography (TBCT) in the Children’s Hospital of Fudan University were retrospectively analyzed. Another 30 age-matched cases of OME with TBCT were collected as the control group. The CT imaging signatures of children were observed. Besides, a variety of DL neural network training models were established based on PyTorch, and the optimal models were trained and selected for PCD screening. Results: The google net-trained model worked best, with an accuracy of 0.99. Vgg16_bn, vgg19_bn, resnet18, and resnet34; having neural networks with fewer layers, better model effects, with an accuracy rate of 0.86, 0.9, 0.86, and 0.86, respectively. Resnet50 and other neural networks with more layers had relatively poor results. Conclusion: DL-based CT radiomics can accurately distinguish OM caused by OME from that induced by PCD, which can be used for screening the PCD.

Keywords: PyTorch, deep learning, primary ciliary dyskinesia, otitis media

Introduction

Primary ciliary dyskinesia (PCD) manifests as congenital dysfunction of mucociliary clearance (MCC) [1], resulting in an inability to perform rhythmic movements, abnormal rhythmic movement, or complete loss of cilia. The disease is an autosomal recessive hereditary disease, with an equal prevalence of approximately 1/30,000-1/10,000 in both men and women [2]. PCD can involve multiple organs, as well as the central nervous system, immune function, fertility, great vessel transposition etc., but with lung involvement being the most common. Mild respiratory distress is common in neonates, and columnar or cystic bronchiectasis may occur in childhood. Besides, frequent runny nose and chronic nasal congestion begin to appear in early childhood. Chronic secretory otitis media (OM) with recurrent attacks of acute OM may occur in the ear, leading to conductive hearing loss (HL) [3]. The diagnosis is based on transmission electron microscopy (TEM) showing specific defects, such as loss of dynamin arms or radial spokes, and the lack or the presence of extra microtubule assembly. High-speed video microscopy analysis (HSVA) combined with ciliary rhythmic motion frequency testing can determine whether the ciliary coordination, rhythmic motion frequency and rhythmic motion patterns are normal [4]. But these tests require nasal or bronchial biopsy specimens and can only be performed in specialized centers. For patients with consistent clinical features, PCD can be diagnosed by genetic testing to determine whether it is caused by homozygous or compound heterozygous pathogenic variants, or the coexistence of different pathogenic genotypes of different PCD genes [5].
Otitis media with effusion (OME) refers to the accumulation of fluid in the middle ear without signs of acute infection [6]. If the effusion is prolonged, it may become thick and gel-like, which may lead to hearing loss. At present, the pathogenesis of OME is not fully understood. Studies on children with OME showed that in the same age group under 3 years old, the hearing threshold of children with bilateral middle ear effusion was about 20 dB higher than that of children without [7]. OME may lead to HL or delayed speech and language development in children, and the long course may trigger tympanic membrane sclerosis, middle ear cholesteatoma, etc. In addition, it’s necessary to conduct more in-depth studies in OM patients that not benefit from drug therapy and who still have recurrent OM episodes after myringotomy and grommet insertion (MGI). OME occurs in 50-90% of patients with PCD [8]. It is necessary to screen out PCD patients from OM patients. However, the diagnostic approaches are cumbersome, time-consuming and costly.

The concept of radiomics was first put forward by Dutch scholars in 2012 [9], which emphasizes extracting a large amount of information from images (MRI, CT, PET, etc.) with high throughput, realizing tumor signature extraction, segmentation, model building and prediction to assist clinicians in making more accurate diagnoses. In recent years, the organic fusion of medical image-aided diagnosis and big data technology has produced a new method, radiomics, which has important clinical value by extracting important features from images to quantify tumors and other major diseases [10]. With the introduction of this new research method, more and more attempts have been made to comprehensively evaluate various phenotypes of tumors using data extracted from radiomics. Radiomics signature analysis is relatively rare in otorhinolaryngology and head and neck surgery. This study is the first to use deep learning (DL) on temporal bone (TB) CT radiomics to differentially diagnose common OME from OM complicated with PCD.

Data and methods

General information

In this retrospective analysis, the clinical data of cases collected from the Children’s Hospital of Fudan University from January 2010 to January 2021 were analyzed.

Inclusion criteria for the case group: 1) Children who met more than two of the following criteria according to the 2018 American PCD Diagnostic Guidelines [11]: early-onset of chronic cough or chronic nasal obstruction within 6 months of age; unexplained neonatal respiratory distress; abnormal visceral arrangement. 2) Children suspected of having PCD and who were followed for at least 1 year. 3) Children who were diagnosed based on TEM, showing specific defects, such as the loss of dynein arms or radial spokes, the lack or presence of extra microtubule assembly, or genetic variations in PCD, including mutations of axonemal assembly of outer (DNAH5, DNAH9, DNAH12, DNA1, ARMC4 and CCDC103) and inner (DNAI1) dynein arms as well as assembly proteins (DNAF2).

TBCT examination revealed 31 patients with OM and mastoiditis, including 19 males and 11 females, aged 1-17 years old, with a mean of (7.4±4.3) years old.

Inclusion criteria for the control group: Thirty patients with chronic OME treated by MGI in our department were collected, and their TBCT data were collected. Children with adenoid atrophy, tonsil hypertrophy, cystic fibrosis, and no OM recurrence within one year of follow-up were excluded.

The male to female ratio in the control group was 19:11, and the age range was 3-15 years old (mean: 8.5±2.8).

The age and gender composition differed insignificantly between the two cohorts (P>0.05), suggesting group comparability.

This study was approved by the Ethics Committee of the Children’s Hospital of Fudan University.

Methods

The processing flow of radiomics is summarized as follows: (1) Acquisition of image data: the original data of TBCT in DICOM format were collected; (2) Layers involving the middle ear were selected, as shown in Figure 1; (3) Labelme software was used to mark the CT of patients.
Diagnosis of primary ciliary dyskinesia and exudative otitis media

Establishment of image database; There were 917 pictures of OM in the PCD group and 955 pictures in the OME group. (4) DL classification and prediction. Python and PyTorch were used to build DL neural networks with learning rates of 0.01 and epoch training of 200. The classification and prediction were carried out by vgg11_bn, vgg16_bn, vgg19_bn, resnet18, resnet34, resnet50, and googlenet networks. The performance of each classifier was evaluated by assessing the accuracy, specificity, sensitivity, and AUC.

Statistical processing

SPSS 22.0 was used to analyze the data. Measurement data were represented by mean ± standard deviation (x ± s), and the inter-group difference was identified using an independent samples t test. The categorical data were recorded as percentage [n (%)] and the inter-group difference was analyzed using a chi-square test. P<0.05 was regarded as the significance level.

Results

CT radiological features of the two groups

The CT findings of the two groups are demonstrated in Figure 1. As can be seen from the figure, bilateral OM and mastoiditis were presented in both PCD patients (Figure 1A) and OME children (Figure 1B). The model can predict the TBCT images with OM and calculate the predicted probability values. There are two examples as shown in Figure 2, DL calculates the classification probability value of input PCD images to be 0.99 and that of OME images to be 1.0 according to pre-training weights.

Classification and Prediction using DL

Vgg11_bn, vgg16_bn, and vgg19_bn, neural networks with fewer layers, were found to have higher accuracy. However, with the increase of neural network layers in Resnet18, Resnet34, and Resnet50, the recall rose and then declined, as shown in Table 1 and Figure 3. The googlenet-trained model worked best, with an accuracy rate of over 99%, as shown in Figure 3. The ROC curve is shown in Figure 4.

Discussion

The majority of PCD patients may be seen by medical professionals more than 50 times before being diagnosed, and the average age at diagnosis is about 10.9 to 14.4 years old [12]. Symptoms of PCD are nonspecific, and there is limited guidance on which population should be referred for appropriate specialized testing. Structural and functional ciliary abnormalities in PCD can cause a wide range of systemic diseases, including expectoration, bronchiectasis, sinusitis, nasal polyps, OM and infertility. Therefore, children with PCD must be followed up regularly, and nursing care should be strengthened. It is also very necessary to screen out PCD patients as early as possible. PCD patients present with congenital mucociliary clearance dysfunction. The ciliated cells of the Eustachian tube are similar to the respiratory ciliated cells, with motility, multiplicity, and typical structure of 9+2 microtubules [13]. There is a mucociliary transport system in the middle ear cavity of healthy people, which is
composed of middle ear mucosa and a mucous blanket [14]. Under normal conditions, middle ear cilia can clear mucus from the middle ear cavity. Disruption of this system can lead to a buildup of fluid in the middle ear.

This study is the first to apply PyTorch to build neural networks for DL. The accuracy rate of differential diagnosis of OM caused by OME and PCD based on TBCT can reach up to 98%, which can provide reference for early intervention of clinicians.

Most previous studies have focused on evaluating the otoscientific characteristics of PCD by means of otoscopic tympanum findings, audiological findings, and clinical course, which is critical to otological diagnosis and requires experience and proficiency [15]. It is currently not clear whether OME in PCD patients will self-heal with age [16], which warrants further research. Nikolaus et al. studied 31 OME patients with PCD, of which 20 received MGI, and 11 received follow-up strategy [17]. Their experimental data support MGI for PCD patients with mild to moderate HL and OME. The management of PCD patients with OM remains controversial. If MGI is required, a T-tube is recommended, because it remains in place for a long time and does not easily fall out.

The current inspection methods have their own disadvantages. Genetic testing is costly with a long reporting period. Moreover, the interpretation of pathogenicity of some genetic variation loci is questionable and cannot be completely determined. Electron microscopy requires invasive examination. While TBCT is a routine examination that takes a short time.

In previous studies, artificial intelligence (AI) and radiomics mostly focused on oncology [18]. In recent years, AI has developed multi-layer network architectures that allows for interpretation of complex data in a highly accurate manner. Radiomics is the result of AI application in the field of medical imaging, which

Table 1. Classification report

<table>
<thead>
<tr>
<th></th>
<th>Accuracy</th>
<th>Precision</th>
<th>Recall</th>
<th>F1_score</th>
</tr>
</thead>
<tbody>
<tr>
<td>googlenet</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>vgg11_bn</td>
<td>0.89</td>
<td>0.92</td>
<td>0.84</td>
<td>0.88</td>
</tr>
<tr>
<td>vgg16_bn</td>
<td>0.86</td>
<td>0.90</td>
<td>0.80</td>
<td>0.85</td>
</tr>
<tr>
<td>vgg19_bn</td>
<td>0.90</td>
<td>0.94</td>
<td>0.86</td>
<td>0.90</td>
</tr>
<tr>
<td>resnet18</td>
<td>0.86</td>
<td>0.88</td>
<td>0.84</td>
<td>0.86</td>
</tr>
<tr>
<td>resnet34</td>
<td>0.86</td>
<td>0.86</td>
<td>0.87</td>
<td>0.86</td>
</tr>
<tr>
<td>resnet50</td>
<td>0.71</td>
<td>0.79</td>
<td>0.56</td>
<td>0.66</td>
</tr>
</tbody>
</table>

The googlenet-trained model has the highest accuracy; the network model within 30 layers has a better effect, which is above 0.85, while the network model above 50 layers has a poor effect.
Diagnosis of primary ciliary dyskinesia and exudative otitis media

Figure 3. Confusion matrixes. A: resnet18 network model; B: resnet34 network model; C: resnet50 network model; D: vgg11 bn network model; E: vgg16 bn network model; F: vgg19 bn network model; G: googlenet-trained model; The googlenet-trained model works best and resnet50 worked the worst; Among vgg, resnet18 and resnet34 network models, vgg16 bn has the highest missed diagnosis rate (0.2).

Figure 4. ROC curves of various models; The AUC of the googlenet-trained model is the highest, and that of VGG, resnet18 and resnet34 network models is higher than 0.89, with the AUC of resnet50 being the lowest.

can even indirectly reflect the changes of genes or proteins of microscopic level at the macro-image level.

In this study, the DL method was used to train and analyze children's TBCT images and make predictions. It showed that the googlenet-trained model worked best with an accuracy rate of 99%, suggesting a high accuracy of this model for early differential diagnosis of PCD and OME.

VGG, proposed in 2014 by the Visual Geometry Group [19], in the Department of Science and Engineering, University of Oxford, has two structures,
namely VGG16 and VGG19, which have no essential difference except the network depth. Compared with other networks, VGG adopts smaller convolution kernels and deeper networks to improve parameter efficiency, enabling the network to learn more complex models and increasing network nonlinearity. VGG-Net has good general performance, and is often used in image feature extraction, target detection and candidate frame generation. The defect lies in the number of parameters. VGG-19 is basically the convolution network architecture with the largest number of parameters. The model effect of the VGG network in this study is better, with an accuracy of above 0.85.

Proposed by He Kaiming et al., from Microsoft Research Institute [20], the Residual Neural Network (ResNet) won the championship in ILSVRC (ImageNet Large Scale Visual Recognition Challenge) in 2015. The main contribution of ResNet is the discovery of “Degradation” and the invention of “Shortcut connection” in response to the Degradation, which greatly eliminates the training difficulty of some neural networks with excessive depth. The “depth” of neural networks broke through 100 layers for the first time, with the largest neural network even exceeding 1000 layers. However, for medical data, like the pictures of small objects in this study, the network is not suitable. In this study, the accuracy rates of both resnet18 and resnet34 were both 0.86, and the recall of the latter was relatively high (0.87). The resnet50 model performed the worst. It may be that there are too many layers of neural network and the receptive field is too large, which affects the model effect.

Inception (also known as GoogLeNet) is a brand-new DL architecture proposed by Christian Szegedy in 2014 [21]. VGG and other structures obtain better training results by increasing the number of layers of the network, but the increase of layers will cause gradient disappearance and gradient explosion. Instead, Inception uses 1×1 convolution for lifting dimensions and convolution reaggregation at multiple dimensions at the same time to make more efficient use of computing resources and extract more features for the same amount of computation to improve training results. In this study, the googlenet-trained neural network model achieved the best results, with its accuracy, precision, recall and f1_score all reaching above 0.99.

The object of this study is the secretions in the middle ear cavity, whose area accounts for a small proportion of the whole image. Generally speaking, the more layers of the DL neural network, the greater the deep receptive field. Deeper neural networks are not suitable for the case study of small lesions. For color picture classification, as many as one million pictures are generally required to have a better model effect. But it is difficult to accumulate such an order of magnitude of data in medicine, which usually can collect only thousands of images, let alone rare diseases. Transfer learning [22] is a machine learning method that focuses on storing solutions to existing problems and leveraging them to other different but related problems, without the need to re-collect and calibrate huge new data sets (and sometimes may be unavailable at all) at great cost. For emerging fields, it can be transferred and applied quickly, reflecting the advantages of timeliness. Therefore, if there is less training data and the training model of DL is less effective, transfer learning can be considered.

The disadvantage of this study is that the number of cases is small, so it is necessary to expand the sample size for further study to improve the diagnostic accuracy. Second, the CT data studied were all from the same CT scanning instrument in our hospital, without data from other hospitals for verification. Third, the CT layer thickness of TB was 0.65 mm, while images of another layer thickness were not studied. Therefore, the robustness of the model in this study needs to be further determined.

As a rare disease, PCD in children is difficult to diagnose in the absence of the characteristic manifestation of visceral inversion. Traditional diagnostic methods are very cumbersome, invasive and extremely expensive [23]. Therefore, neonatologists, respiratory physicians, and otolaryngologists should be familiar with the clinical characteristics of PCD, collect medical history comprehensively, and pay attention to full-term children with intractable OM accompanied by recurrent lower respiratory tract infection or previous neonatal respiratory distress to consider the possibility of PCD.

The vast majority of children with PCD have chronic OME. Compared with common chronic OM, the imaging features of OM in PCD children are obviously different. Future research can combine gene, TB CT and MR of other sequenc-
Diagnosis of primary ciliary dyskinesia and exudative otitis media


Gillies RJ, Kinahan PE and Hricak H. Radiomics: images are more than pictures, they are data. Radiology 2016; 278: 563-577.


Gillies RJ, Kinahan PE and Hricak H. Radiomics: images are more than pictures, they are data. Radiology 2016; 278: 563-577.


Gillies RJ, Kinahan PE and Hricak H. Radiomics: images are more than pictures, they are data. Radiology 2016; 278: 563-577.


Gillies RJ, Kinahan PE and Hricak H. Radiomics: images are more than pictures, they are data. Radiology 2016; 278: 563-577.
Diagnosis of primary ciliary dyskinesia and exudative otitis media


