

doi: 10.12052/gdutxb.230050

基于切片关联信息的慢性阻塞性肺疾病CT诊断

梁宇辰¹, 蔡念¹, 欧阳文生¹, 谢依颖¹, 王平²

(1. 广东工业大学 信息工程学院, 广东 广州 510006; 2. 广州医科大学附属第一医院 肝胆外科, 广东 广州 510120)

摘要: 慢性阻塞性肺疾病(Chronic Obstructive Pulmonary Disease, COPD)是一种常见的全球呼吸系统疾病, 需要耗费医生大量的时间和精力对CT图像进行初步评估诊断。为了提高阅片效率, 提出一种基于CT图像切片关联信息的深度网络, 辅助诊断慢性阻塞性肺疾病。提出一种分组方式将网络分成若干个网络分支, 每个网络分支能够提取局部CT图像切片内部关联信息, 结合双向LSTM技术整合各网络分支信息以提取CT图像全局切片关联信息。为了进一步提升网络分支的局部特征提取能力, 融入ConvNeXt提出增强的多头卷积注意力模块。对比实验结果表明, 所提出的深度网络能够更好地对CT图像进行分类, 辅助COPD诊断, 其准确率达到92.15%, 敏感度达到94.17%, 特异性达到91.17%, AUC达到95.33%。

关键词: 慢性阻塞性肺疾病; 深度学习; 多头卷积注意力

中图分类号: TP391

文献标志码: A

文章编号: 1007-7162(2024)01-0027-07

CT Diagnosis of Chronic Obstructive Pulmonary Disease Based on Slice Correlation Information

Liang Yu-chen¹, Cai Nian¹, Ouyang Wen-sheng¹, Xie Yi-ying¹, Wang Ping²

(1. School of Information Engineering, Guangdong University of Technology, Guangzhou 510006, China; 2. Department of Hepatobiliary Surgery, the First Affiliated Hospital of Guangzhou Medical University, Guangzhou 510120, China)

Abstract: Chronic obstructive pulmonary disease (COPD) is a common respiratory disease of the world, and the doctors need a lot of time to read the abdominal CT images for COPD pre-evaluation. To improve the pre-evaluation efficiency, a deep network based on slice correlation information was proposed for COPD auxiliary diagnosis. First, by using a grouping approach, the architecture of the deep network is divided into several network branches, each of which aims to extract the local intra-slice association information of the CT images. Then, the outputs from multiple network branches are integrated via a BiLSTM to extract the global inter-slice association information between the adjacent CT slices. To further improve the ability of local feature extraction for each network branch, the enhanced multi-headed convolutional attention is designed by embedding the ConvNeXt into the existing multi-headed convolutional attention. Experimental results show that the proposed deep network achieves promising effectiveness for CT image classification on auxiliary diagnosis of COPD, and the accuracy, sensitivity, specificity and AUC of the proposed network reach to approximately 92.15%, 94.17%, 91.17% and 95.33%, respectively.

Key words: chronic obstructive pulmonary disease (COPD); deep learning; multi-head convolutional attention

慢性阻塞性肺疾病(Chronic Obstructive Pulmonary Disease, COPD)是一种常见的全球性呼吸疾病, 已成为全球的第三大死因^[1]。近些年, 我国的COPD患病率也呈现不断上升趋势, 且死亡率高于全球平均

水平^[2-3]。COPD患者通常会出现气道严重阻塞现象, 进一步导致呼吸困难, 甚至存在肺心病、呼吸衰竭等风险^[4]。目前, 主要采用CT图像评估^[5-6]鉴别COPD肺部异常, 这需要医生根据COPD患者的上百张CT序

收稿日期: 2023-03-16

基金项目: 国家自然科学基金资助项目 (82172019); 广州市科技计划项目 (202102010251)

作者简介: 梁宇辰 (1998-), 男, 硕士研究生, 主要研究方向为深度学习、医学图像处理等

通信作者: 蔡念 (1976-), 男, 教授, 博士生导师, 主要研究方向为机器视觉、机器学习、数字信号处理等, E-mail: cainian@gdut.edu.cn

列图片中的病灶气泡占比情况进行主观分析,耗费医生大量精力和时间,也给医生带来视觉疲劳并影响主观评估^[7]。

近几年,随着深度学习尤其是卷积神经网络(Convolutional Neural Network, CNN)的高速发展, CNN已经被广泛应用于CT图像处理,以辅助肺部疾病诊断。Shah等^[8]使用VGG-19对新型冠状病毒肺炎(Corona Virus Disease 2019, COVID-19)进行诊断。Polat等^[9]在Inceptionv3模型上使用迁移学习,从患者的单张胸部CT图像获取病灶信息识别COPD的严重程度。虽然上述方法在肺部疾病辅助诊断上取得不错的效果,但是这些方法只考虑了部分CT切片图像内的局部病变肺泡特征,没有考虑CT切片图像之间的关联信息,这将不利于肺部病灶区病变肺泡识别^[10]。

Xu等^[11]从胸部CT序列中随机选出8张CT图像,采用AlexNet网络对每张CT图像进行特征提取,然后采用SVM进行分类实现COPD诊断。可是,该方法只对单张CT图像进行单独分析,虽然SVM对8张CT图像进行了分类,但是本质上仍未考虑这8张CT切片之间的关联信息。Ahmed等^[12]将VoxResNet拓展到3D卷积结构,通过多个残差3D卷积结构保留尽可能多的空间信息对CT序列图像进行处理来诊断COPD。Varchagall等^[13]使用3D ResNet提取CT图像病灶空间特征并使用迁移学习进行肺癌诊断。Kienzle等^[14]将ConvNeXt网络拓展到3维结构构成3D ConvNeXt网络对COVID-19进行诊断。这些网络都是采用3D卷积方法对CT序列图像进行处理,但是3D卷积难以提取病灶区细微肺泡特征,影响网络的分类效果。Wu等^[15]从CT序列图像中提取气管树图像和9张3D肺部图像等其他模态信息作为ResNet26的网络输入,从而实现COPD辅助诊断。可是,该方法受制于气管树3D形态提取的精准性,缺乏COPD辅助诊断的便捷性。Kollias等^[16]使用CNN和RNN分别获取局部病变肺泡的图像特征和全局切片间的关联信息,构建了一种MIA-COVID-19网络对COVID-19进行诊断。Humphries等^[17]结合CNN与长短期记忆(Long-Short Term Memory, LSTM)对COPD进行诊断。上述两个网络可以同时学习病灶图像局部特征和切片之间全局关联信息。但是,它们直接对上百张CT切片图像进行特征关联信息提取,忽视了不同距离的CT切片图像之间的关联性是不同的。

本文提出一种基于CT切片图像关联信息的深度网络,辅助诊断COPD。将COPD患者胸部CT切片

序列分为若干组作为整个网络的网络分支的输入,同时提取组内切片之间的局部关联信息和各切片内的病灶图像局部特征。为了提高各网络分支的病灶图像局部特征提取能力,融入ConvNeXt提出一种增强的多头卷积注意力模块。

1 方法

1.1 网络结构

基于切片关联信息的COPD诊断网络输入COPD患者的胸部CT图像序列,输出诊断结果。网络由多个局部切片关联信息提取分支和一个全局的切片关联信息提取阶段构成。图1展示了3个分支组成的网络结构示意图。局部切片关联信息提取分支由Conv STEM模块、En-MHCA模块、Down Sample模块和Ef-Transformer模块构成,主要提取组内CT切片之间的局部关联信息和组内CT切片病灶区的局部图像特征。将每个分支的特征信息以 N 切片数的维度进行二次拼接组成局部特征序列输入到全局切片关联信息提取阶段,该阶段主要由BiLSTM^[18]构成,提取各分支之间的切片之间的全局关联信息。最后,级联一个多层感知器(Multi-Layer Perceptron, MLP)进行COPD诊断。

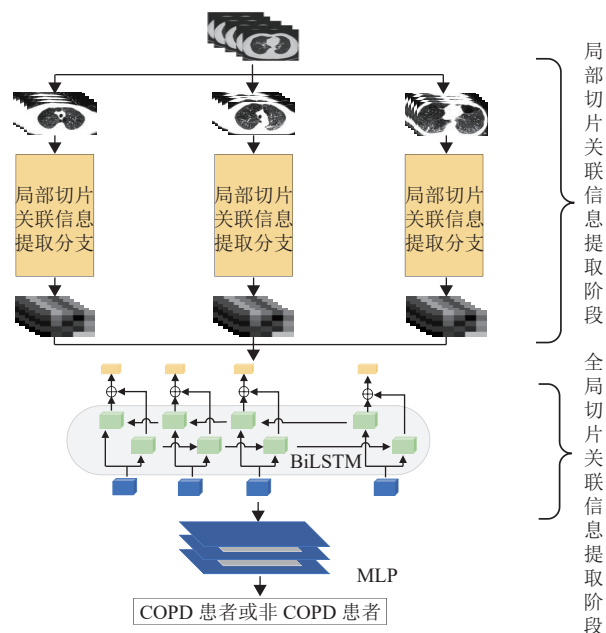


图1 本文提出的深度网络架构

Fig.1 The framework of the proposed deep network

1.2 局部切片关联信息提取分支

肺部病变肺泡是COPD患者的典型病理特征,是COPD诊断重要依据。根据合作医生提供的先验知识^[19], COPD病变肺泡通常在3至4张CT切片中就能

观察到完整的病变肺泡结构;在一些切片数量较多的情况中,病变肺泡可能分布在10张左右CT切片中。因此,本文提出对COPD患者胸部CT序列图像进行合理分组,各分组可以提取组内CT切片之间的局部关联信息和组内CT切片病灶区的局部图像特征。

将COPD患者CT序列数据划分为若干组,每组内有 N 张切片,每张CT切片图像维度为 (C, W, H) ,其中 C 代表特征通道数, W 代表图像宽度, H 代表图像高度。本文中, N 经验选择为10,后续消融实验将验证此选择。因此,分组后的序列图像组数由患者CT序列的切片总数决定,如患者CT序列为400张CT切片,则分组数为40组。

将各组CT切片作为各局部切片关联信息提取分支的输入,其结构如图2所示。其中,Conv STEM模块用于低维特征提取,2个En-MHCA模块和2个Down Sample下采样模块用于提取高维局部病灶特征。提取到的高维特征按 N 切片数的维度拼接为特征序列输入到Ef-Transformer模块。图2中仅展示了3张图像的特征拼接情况。在Ef-Transformer模块中,特征序列通过自注意力^[20]获取组内切片之间的局部关联信息和局部病灶图像特征。

1.3 Conv STEM模块

Conv STEM模块的结构如图3所示,包含一组由Depth Wise卷积层和Point Wise卷积层构成的深度可分离卷积层、Batch Normal归一化层、Max Pool层和ReLU层,数学表示为

$$\text{STEM}(x_1) = \text{M}(\text{R}(\text{BN}(\text{PWC}(\text{DWC}(x_1)))) \quad (1)$$

式中: x_1 为输入特征, M 代表Max Pool层, R 代表ReLU激活函数层, BN 代表Batch Normal归一化层, PWC 代表Point Wise卷积层, DWC 代表Depth Wise卷积层。

1.4 Down Sample模块

Down Sample模块的结构如图4所示,主要由1个Convolution层和1个Max Pool层构成,其数学表示为

$$\text{Down Sample}(x_2) = \text{M}(\text{Conv}(x_2)) \quad (2)$$

式中: x_2 为该模块的输入, $\text{Conv}()$ 代表卷积层操作。在图4中, H 、 W 和 C 分别为特征映射图的高、宽和通道数, N 为切片数。

1.5 En-MHCA模块

受ConvNeXt^[21]的启发,提出了一个增强的多头卷积注意力模块(Enhanced Multi Head Convolutional Attention, En-MHCA)。与传统的多头卷积注意力模块(Multi Head Convolutional Attention, MHCA)^[22]不

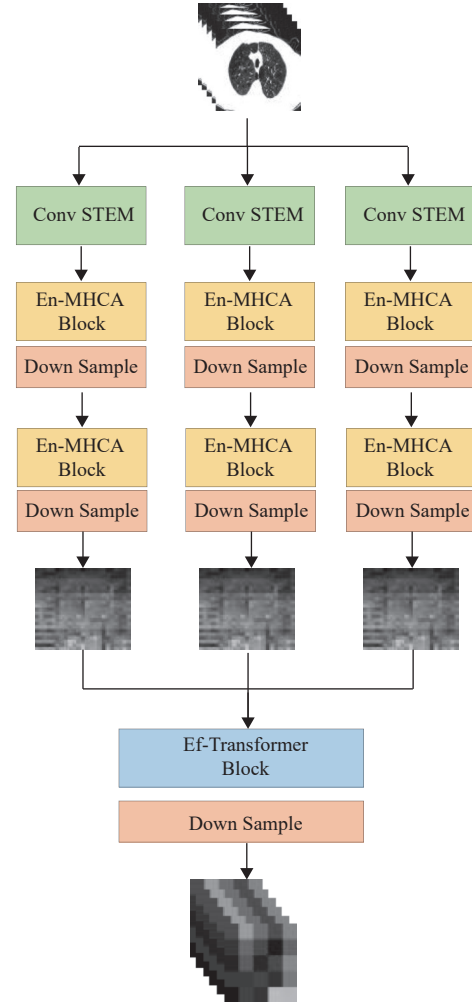


图2 局部切片关联信息提取分支的结构

Fig.2 The structure of the branch for local slice correlated information extraction

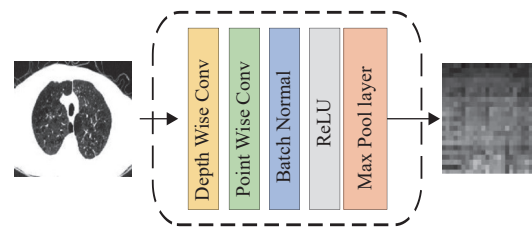


图3 Conv STEM的结构

Fig.3 The structure of the Conv STEM

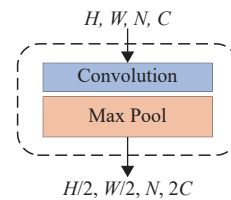


图4 下采样模块

Fig.4 Down sampling module

同的是,En-MHCA融合ConvNeXt进行特征提取和非线性拟合,去除了只有非线性拟合能力的MLP结构,

因此En-MHCA本质上是一个倒置的多头卷积注意力结构。传统MHCA仅使用MHCA层进行特征学习,而且MLP结构仅有非线性拟合能力。相比之下, En-MHCA包含了2个拥有特征学习能力的残差结构(即ConvNeXt和MHCA),具有更强的特征提取能力。

如图5所示,ConvNeXt由1个Depth Wise卷积层、2个Point Wise卷积层以及残差结构构成。Depth Wise卷积层是一个大小为3、步长为1且分组数和通道数一致的卷积层。Point Wise卷积层是一个大小为1、步长为1的通道转换卷积层。经过1个Depth Wise卷积层、2个Point Wise卷积层后输出的特征映射图和ConvNeXt输入特征映射图进行一次残差结构相加。

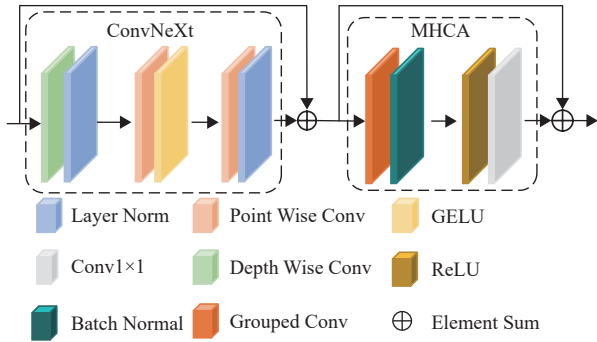


图5 增强的多头卷积注意力模块

Fig.5 Enhanced multi-head convolutional attention module

ConvNeXt输出的特征映射图作为MHCA输入。MHCA主要由1个 3×3 分组卷积层、1个 1×1 卷积层、1个Batch Normal层和1个ReLU层构成,其数学表示为

$$R_1 = \text{PWC}(\text{PWC}(\text{DWC}(x_3))) + x_3 \quad (3)$$

$$R_2 = \text{MHCA}(R_1) + R_1 \quad (4)$$

式中: x_3 为输入特征映射图, R_1 为ConvNeXt输出特征映射图, R_2 为MHCA输出特征映射图。

1.6 Ef-Transformer模块

为了更好地获取组内切片间的局部关联信息,在分支中采用Ef-Transformer模块^[23]提取特征。如图6所示,特征映射图经过1个Layer Norm层,再输入到E-MHSA(Efficient Multi-head Self-attention)。E-MHSA输出的特征映射图和Ef-Transformer输入的初始特征映射图通过一个残差连接进行相加得到组内切片间的关联信息 S_1 。随后, S_1 经过1个Layer Norm层和1个MLP后,再与 S_1 执行一次残差连接相加。Ef-Transformer模块的计算过程可以数学表示为

$$S_1 = \text{EMHSA}(\text{LN}(x_4)) + x_4 \quad (5)$$

$$S_2 = \text{MLP}(\text{LN}(S_1)) + S_1 \quad (6)$$

式中: x_4 为输入特征映射图,EMHSA代表多头自注

意力模块操作,LN代表LayerNorm正则化层操作, S_1 为一次残差连接的输出特征, S_2 为第二次残差连接的输出特征。

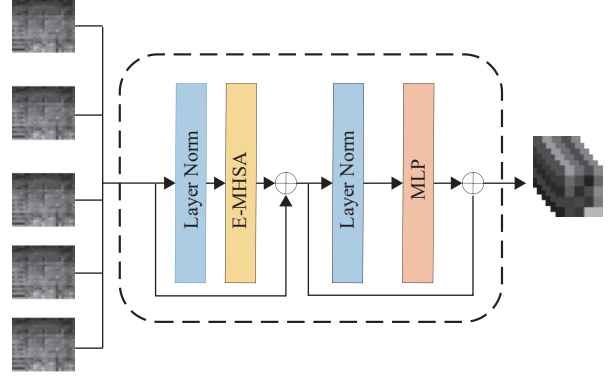


图6 Ef-Transformer模块

Fig.6 The Ef-Transformer module

1.7 BiLSTM模块

因为不同COPD患者的CT切片数是不一致的,所以各患者CT切片序列的分组数也有所不同。故采用BiLSTM多单元模块来提取各组之间的切片全局关联信息。BiLSTM结构如图7所示,可以数学表示为

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i) \quad (7)$$

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f) \quad (8)$$

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o) \quad (9)$$

$$C_t = C_{t-1} \cdot f_t + \tanh(W_C x_t + U_C h_{t-1} + b_C) \cdot i_t \quad (10)$$

$$h_t = o_t \cdot \tanh(C_t) \quad (11)$$

$$\vec{h}_t = (\vec{h}_t, \overleftarrow{h}_t) \quad (12)$$

式中: i_t 为输入门的输出特征, f_t 为遗忘门的输出特征、 o_t 为输出门的输出特征, b 为各门偏置项, W 和 U 分别为输入序列特征 x_t 和隐藏状态 h_{t-1} 中各门的网络参数。 x_t 、 h_t 分别为特征在切片 t 中的输入和隐藏状

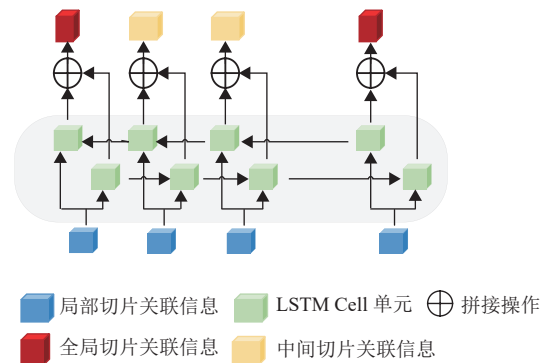


图7 BiLSTM模块

Fig.7 The BiLSTM module

态, h_{t-1} 为切片 $t-1$ 的隐藏状态, C_t 为记忆单元状态, $\vec{h}_t$ 和 $\overleftarrow{h}_t$ 分别为LSTM的前向输出和后向输出, $\tanh$ 为双曲正切函数, $\vec{h}_t$ 表示BiLSTM的输出, $\cdot$ 为hadamard乘积。

1.8 损失函数

采用交叉熵损失函数优化提出的深度网络参数, 数学表达为

$$L_{CE} = -\frac{1}{M} \sum_{i=1}^M y_i \ln(p_i) \quad (13)$$

式中: y_i 为COPD真实诊断标签, p_i 为模型根据数据诊断的结果, M 表示训练集的样本数, i 表示当前样本序号。

2 实验

2.1 数据集及训练环境配置

本文实验数据均由广州医科大学第一附属医院提供, 共收集了161例COPD患者CT序列, 每例数据包含130到400之间的CT切片图像。因此, 数据集总共包含43 140张CT切片图像。每张切片图像的分辨率是512×512, 像素平均间距为0.68 mm, 切片平均厚度为1.143 mm。由于患者隐私等协议条款, 本文不能公开这些影像数据。

训练集由110例共30 190张CT切片构成, 测试集由51例共12 950张CT切片构成, 数据标注均由经验丰富的医生完成。

本文实验皆在一台配置为NVIDIA RTX A6000 48 GB GPU和Inter Xeon(R) Gold 5218R 2.10GHz CPU的工作站上完成。模型训练和测试代码都是基于PyTorch 1.6(python3.8)深度学习框架编程实现, 其训练初始化参数如表1所示。

表1 模型训练初始化参数

Table 1 The initialization parameters of the model in training

参数	数值
模型训练的迭代次数	200
数据批次大小	1
学习率	0.000 03
训练时的优化器	AdamW
指数衰减率 β_1	0.9
指数衰减率 β_2	0.999
权重衰减率	0.05

评估标准有准确率(Accuracy, ACC)、灵敏度(Sensitivity, SEN)、特异性(Specificity, SPE)和受试者工作特征(Receiver Operating Characteristic, ROC)曲线下的面积值(Area Under Curve, AUC)。其中, AUC

的值由ROC曲线与坐标围成的区域面积所得。ACC、SEN、SPE的计算公式为

$$ACC = \frac{2TP}{2TP + FN + FP} \quad (14)$$

$$SEN = \frac{TP}{TP + FN} \quad (15)$$

$$SPE = \frac{TN}{TN + FP} \quad (16)$$

式中: TP表示真阳性, 真实标签为COPD患者, 模型预测为COPD患者; FP表示假阳性, 真实标签为COPD患者, 模型预测为非COPD患者; TN表示真阴性, 真实标签为非COPD患者, 模型预测为非COPD患者; FN表示假阴性, 真实标签为非COPD患者, 模型预测为COPD患者。

2.2 消融实验

(1) 不同模块对网络性能的影响。

本文通过相关模块的消融实验说明网络中相关模块对网络诊断性能的影响, Ef-Transformer缩写为Ef-Trans, 数据如表2所示。

表2 不同特征提取模块的对比

Table 2 Comparisons of different feature extraction modules

模块				指标			
BiLSTM	Ef-Trans	MHCA	En-MHCA	ACC/%	SEN/%	SPE/%	AUC/%
√	×	√	×	88.23	76.47	94.11	89.96
×	√	√	×	80.39	58.82	91.17	86.16
√	√	√	×	90.16	76.27	97.05	93.25
√	√	×	√	92.15	94.17	91.17	95.32

从表2可以看到, 当只使用Ef-Trans时, 网络只获取了组内切片间的局部关联信息, 而忽视了组间切片的全局关联信息, 从而导致网络诊断性能最差。而当只采用BiLSTM时, 网络的诊断准确率等性能要比只使用Ef-Trans有大幅度的提升, 这说明组间的长程关联性对COPD诊断具有重要指导价值。当同时使用BiLSTM和Ef-Trans时, 网络性能相比于只使用BiLSTM时的网络略微有所提升, 这说明组间的切片全局关联信息和组内的切片局部关联信息的有机融合能提升COPD患者的CT序列分类效果。从最后一行数据来看, 提出的增强多头卷积注意力模块与普通的卷积主力模块相比, 网络的诊断准确率等性能有所提升。

(2) 不同分组切片参数对网络性能的影响。

为了探索不同的分组切片数对网络性能的影响

响,本文做了对比实验。如表3所示,当分组切片数为10时,网络的AUC指标表现更好。这是因为过少的分组切片数会造成网络无法获取完整的结构关联信息,而过多的分组切片数提供了冗余的切片关联信息。

表3 不同分组切片参数的对比实验
Table 3 Comparative experiments with different grouping of slicing parameters

<i>N</i>	ACC/%	SEN/%	SPE/%	AUC/%
5	88.23	82.35	91.17	90.48
10	92.15	94.17	91.17	95.32
15	92.15	82.35	97.05	94.46
20	90.19	76.47	97.05	92.56

2.3 对比实验

为了验证提出模型的效果,将其与现有的一些基于CT图像的肺部疾病深度学习方法进行对比,这些方法分别是VGG-19^[8]、3D VoxResNet^[12]、DCT-MIL^[11]、MIA-COV19D^[16]、3D ResNet^[13]、CNN+LSTM^[17]和3D ConvNeXt^[14],结果如表4所示。

表4 不同深度学习方法的对比
Table 4 Comparisons among different deep learning methods

模型	ACC/%	SEN/%	SPE/%	AUC/%
VGG-19 ^[8]	72.54	41.17	88.23	63.15
3D VoxResNet ^[12]	78.43	88.23	73.53	81.49
DCT-MIL ^[11]	82.33	70.58	88.54	82.53
MIA-COV19D ^[16]	82.35	64.71	91.17	87.37
CNN+LSTM ^[17]	84.31	70.58	91.17	83.22
3D ResNet ^[13]	76.47	35.29	97.05	60.21
3D ConvNeXt ^[14]	88.23	70.58	97.05	92.91
本文方法	92.15	94.17	91.17	95.33

从表4可以看出,VGG-19网络在提取CT图像特征过程中,网络深度的增加导致病变肺泡特征信息丢失,且VGG-19网络也未考虑切片之间的关联信息,因此获得非常差的COPD诊断性能。

3D ResNet和3D VoxResNet都通过残差结构缓解因网络深度不断加深而丢失特征的现象,因此获得比VGG-19网络更好的COPD诊断性能。3D ConvNeXt网络使用的3D ConvNeXt模块具有很强的空间提取能力,因此获得了高达92.91%的AUC值。虽然这3种3D网络都较好地实现了COPD辅助诊断,但是它们仅使用CT图像内的局部空间特征,忽视了切片间的关联信息,因此难以获取病变肺泡的整体形貌信息,不同程度地影响网络诊断性能。MIA-COV19D和CNN+LSTM采用CNN提取切片内的局部

细节信息,分别采用RNN和LSTM提取CT序列的全局关联信息,因此分别获得了87.37%和83.22%的AUC值。但是,这两种方法在提取CT序列的全局关联信息时未考虑不同距离切片之间的关联程度。DCT-MIL采用AlexNet迁移特征,并采用多实例学习策略分析随机选取的8张连续CT切片图像,并考虑了邻近切片之间的关联性,因此也获得了82.53%AUC值。但是,DCT-MIL忽视了整个CT序列的长程关联关系,导致其诊断性能仍然较低。本文所提出的深度网络结合了组间切片的全局关联信息和组内切片的局部关联信息,同时融合了切片图像病灶区的局部图像特征,因此获得了最佳的COPD辅助诊断效果,准确率达到92.15%,敏感度达到94.17%,特异性达到91.17%,AUC达到95.33%。图8为不同深度学习方法的ROC曲线图。如图8所示,本文网络的ROC曲线包裹面积最大。

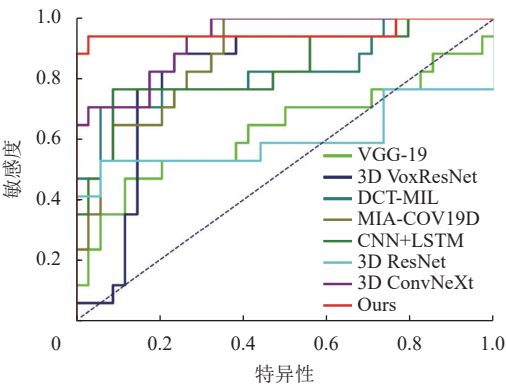


图8 ROC曲线图
Fig.8 The ROC curves

3 结论

本文提出一种基于切片关联信息的深度网络,应用于COPD辅助诊断。采用分组方式,融合ConvNeXt和MHCA提出增强的多头卷积注意力模块(En-MHCA),提取组内CT切片间的局部关联信息和CT切片病灶区的局部图像特征。采用BiLSTM提取组间CT切片的全局关联信息。实验结果表明,提出的深度网络比多个现有深度网络具有更好的诊断性能,准确率、敏感度、特异性和AUC分别达到92.15%、94.17%、91.17%和95.33%。

参考文献:

[1] MARKELIĆ I, HLAČIĆ I, ČERI A, et al. Activation of NLRP3 inflammasome in stable chronic obstructive pulmonary disease [J]. *Scientific Reports*, 2022, 12(1): 7544.
[2] VESTBO J, HURD S S, AGUSTÍ A G, et al. Global strategy

- for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary [J]. *American Journal of Respiratory and Critical Care Medicine*, 2013, 187(4): 347-365.
- [3] MATHERS C D, LONCAR D. Projections of global mortality and burden of disease from 2002 to 2030 [J]. *PLoS Medicine*, 2006, 3(11): e442.
- [4] 周祖东. 慢性阻塞性肺疾病急性加重期的临床治疗分析[J]. *中外医疗*, 2020, 39(36): 67-69.
- ZHOU Z D. Analysis of clinical treatment in acute exacerbation of chronic obstructive pulmonary disease [J]. *China & Foreign Medical Treatment*, 2020, 39(36): 67-69.
- [5] LYNCH D A, AUSTIN J H M, HOGG J C, *et al.* CT-definable subtypes of chronic obstructive pulmonary disease: a statement of the Fleischner society [J]. *Radiology*, 2015, 277(1): 192-205.
- [6] GALBÁN C J, HAN M K, BOES J L, *et al.* Computed tomography-based biomarker provides unique signature for diagnosis of COPD phenotypes and disease progression [J]. *Nature Medicine*, 2012, 18(11): 1711-1715.
- [7] LYNCH D A, MOORE C M, WILSON C, *et al.* CT-based visual classification of emphysema: association with mortality in the COPD Gene study [J]. *Radiology*, 2018, 288(3): 859-866.
- [8] SHAH V, KENIYA R, SHRIDHARANI A, *et al.* Diagnosis of COVID-19 using CT scan images and deep learning techniques [J]. *Emergency Radiology*, 2021, 28(3): 497-505.
- [9] POLAT Ö, ŞALK İ, DOĞAN Ö T. Determination of COPD severity from chest CT images using deep transfer learning network [J]. *Multimedia Tools and Applications*, 2022, 81(15): 21903-21917.
- [10] MONDOÑEDO J R, SATO S, OGUMA T, *et al.* CT imaging-based low-attenuation super clusters in three dimensions and the progression of emphysema [J]. *Chest*, 2019, 155(1): 79-87.
- [11] XU C, QI S, FENG J, *et al.* DCT-MIL: deep CNN transferred multiple instance learning for COPD identification using CT images [J]. *Physics in Medicine & Biology*, 2020, 65(14): 145011.
- [12] AHMED J, VESAL S, DURLAK F, *et al.* COPD classification in CT images using a 3D convolutional neural network[C]// *Bildverarbeitung für die Medizin* 2020. Berlin: Springer, 2020: 39-45.
- [13] VARCHAGALL M, NETHRAVATHI N P, CHANDRAMMA R, *et al.* Using deep learning techniques to evaluate lung cancer using CT images [J]. *SN Computer Science*, 2023, 4(2): 173.
- [14] KIENZLE D, LORENZ J, SCHÖN R, *et al.* COVID detection and severity prediction with 3D-ConvNeXt and custom pretrainings[C]// *Computer Vision—ECCV*. Tel Aviv: Springer, 2022: 500-516.
- [15] WU Y, DU R, FENG J, *et al.* Deep CNN for COPD identification by multi-view snapshot integration of 3D airway tree and lung field [J]. *Biomedical Signal Processing and Control*, 2023, 79: 104162.
- [16] KOLLIAS D, ARSENOS A, SOUKISSIAN L, *et al.* MIA-COV19D: COVID-19 detection through 3-D chest CT image analysis[C]// *International Conference on Computer Vision*. Montreal: IEEE/CVF, 2021: 537-544.
- [17] HUMPHRIES S M, NOTARY A M, CENTENO J P, *et al.* Deep learning enables automatic classification of emphysema pattern at CT [J]. *Radiology*, 2020, 294(2): 434-444.
- [18] LIN H, ZHANG S, LI Q, *et al.* A new method for heart rate prediction based on LSTM-BiLSTM-Att [J]. *Measurement*, 2023, 207: 112384.
- [19] FRIEDMAN P J. Imaging studies in emphysema [J]. *Proceedings of the American Thoracic Society*, 2008, 5(4): 494-500.
- [20] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, *et al.* An image is worth 16×16 words: Transformers for image recognition at scale [EB/OL]. arXiv: 2010.11929 (2021-01-03) [2023-03-15]. <https://doi.org/10.48550/arXiv.2010.11929>.
- [21] LIU Z, MAO H, WU C Y, *et al.* A convnet for the 2020s[C]// *Conference on Computer Vision and Pattern Recognition*. New Orleans: IEEE/CVF, 2022: 11976-11986.
- [22] LI J, XIA X, LI W, *et al.* Next-ViT: next generation vision transformer for efficient deployment in realistic industrial scenarios [EB/OL]. arXiv: 2207.05501 (2022-08-16) [2023-03-15]. <https://doi.org/10.48550/arXiv.2207.05501>.
- [23] WANG W, XIE E, LI X, *et al.* Pyramid vision transformer: a versatile backbone for dense prediction without convolutions[C]// *International Conference on Computer Vision*. Montreal: IEEE/CVF, 2021: 568-578.

(责任编辑: 赵少飞 英文审核: 费伦科)