

DOI:10.3969/j.issn.1005-202X.2024.11.007

医学影像物理

## 高效注意力金字塔网络在肺结节检测的应用

张琼<sup>1,2</sup>, 杭益柳<sup>1,2</sup>, 邱建林<sup>1,2,3</sup>, 吴芳<sup>1,2</sup>

1.南通理工学院计算机与信息工程学院, 江苏南通 226000; 2.南通市虚拟现实与云计算重点实验室, 江苏南通 226000; 3.南通大学信息科学技术学院, 江苏南通 226000

**【摘要】**针对肺部CT图像结节特征不明确和检测困难的问题,提出一种高效注意力特征金字塔网络。首先,该网络以多尺度特征融合的特征金字塔为主干网络,保留丰富的低层特征和高层特征,同时采用深度可分离卷积神经网络提取特征信息;然后,将注意力机制融入主干网络中,对重要特征信息进行加权处理;最后,将所提算法应用在Lung-PET-CT-Dx和Luna16数据集中。实验结果表明,本文算法的精度、召回率和mAP值均优于现有对比算法,证明本文算法在肺结节检测的优越性。

**【关键词】**肺结节;深度可分离卷积神经网络;注意力机制;特征金字塔;目标检测

**【中图分类号】**R318;TP391.4

**【文献标志码】**A

**【文章编号】**1005-202X(2024)11-1361-09

## Efficient attention feature pyramid network for pulmonary nodule detection

ZHANG Qiong<sup>1,2</sup>, HANG Yiliu<sup>1,2</sup>, QIU Jianlin<sup>1,2,3</sup>, WU Fang<sup>1,2</sup>

1. School of Computer and Information Engineering, Nantong Institute of Technology, Nantong 226000, China; 2. Nantong Key Laboratory of Virtual Reality and Cloud Computing, Nantong 226000, China; 3. School of Information Science and Technology, Nantong University, Nantong 226000, China

**Abstract:** To address the challenge of unclear features and difficulties in pulmonary nodule CT image detection, an efficient attention feature pyramid network is proposed. The network firstly employs a feature pyramid of multi-scale feature fusion as the backbone network for effectively preserving both low- and high-level features, and uses the depthwise separable convolutional neural network to extract feature information. Then, the attention mechanism is integrated into the backbone network for assigning weights to salient feature information. Finally, the proposed algorithm is applied to Lung-PET-CT-Dx dataset and Luna16 dataset, and the experimental results demonstrate that the proposed algorithm has higher precision, recall rate and mAP value than the existing comparative algorithms, substantiating its superiority in pulmonary nodule detection.

**Keywords:** pulmonary nodule; depthwise separable convolutional neural network; attention mechanism; feature pyramid; object detection

### 前言

肺癌作为所有癌症中发病率和致死率最高的疾病之一,严重影响着民众的生命健康<sup>[1]</sup>。早发现、早

诊断、早治疗是现代医学防控肺癌的重要手段之一,并且随着新冠肺炎的大流行,肺部CT图像诊断已经成为肺科医生最繁重的工作<sup>[2]</sup>。因此,如何借助计算机技术辅助医生对肺部CT的诊断是一项重要的研究任务。随着深度学习的发展,目标检测技术取得很大的进展,目前基于深度学习的目标检测技术主要分为两类:一阶段检测和两阶段检测。一阶段检测主要是输入到输出的过程,包含的主要算法有单次检测(Single Shot Detection, SSD)算法<sup>[3]</sup>、RetinaNet算法<sup>[4]</sup>、YOLO系列算法<sup>[5-10]</sup>等。两阶段检测主要是先生成建议区域,然后根据建议区域检测目标位置的过程,包含的主要算法有RCNN算法<sup>[11]</sup>、SPPNet算法<sup>[12]</sup>、Fast RCNN算法<sup>[13]</sup>、Faster RCNN算法<sup>[14]</sup>以及

**【收稿日期】**2024-06-24

**【基金项目】**江苏省自然科学基金(BK20231337);江苏省高校重大自然科学基金(21KJA510004);南通市科技局基础科学研究项目(JCZ2022108, MSZ2022161);南通理工学院中青年骨干教师项目(ZQNGGJS202237, ZQNGGJS202234);南通理工学院科研项目(2022XK(Z)19)

**【作者简介】**张琼,博士研究生,讲师,研究方向:深度学习、迁移学习, E-mail: zhangq@ntit.edu.cn

**【通信作者】**杭益柳,硕士,讲师,研究方向:深度学习、数字图像处理, E-mail: 18862928527@163.com

FPN 算法<sup>[15]</sup>等。针对这两类目标检测算法,研究者结合肺部CT图像进行深入研究。Su等<sup>[16]</sup>提出一种基于Faster RCNN框架的肺结节检测方法,该方法利用Faster RCNN在肺结节目标检测过程中对数据的处理能力和网络的鲁棒性以及过拟合能力来提高对肺结节目标的检测准确度,但是该方法并未给出肺结节的位置信息。Loraksa等<sup>[17]</sup>提出一种基于SSD-VGG16的肺结节检测方法,该方法采用SSD框架,结合VGG16骨干网络,获取每个骨肉瘤结节的边界框位置和大小,并给出置信度评分,通过该方法能检测和定位结节的位置,但是该方法的准确度相对较低。Mahum等<sup>[18]</sup>提出一种多尺度特征融合的RetinaNet方法,该方法引入基于多尺度特征融合模块对各网络层进行聚合,同时增加浅层预测层的语义信息,并根据上下文模块采用扩展轻量级算法,将上下文信息与网络各阶段层相结合,定位微小肿瘤特征信息,但是该方法对微小肿瘤特征提取准确度不高。Yan等<sup>[19]</sup>提出一种基于YOLOv5的肺结节自动检测和定位算法,该算法通过iResNet改进YOLOv5的主干网络去提高特征提取能力,并通过BiFPN改进特征融合网络,提高对肺结节的检测能力,该方法同样存在微小特征丢失,目标检测不准确的问题。秦源源等<sup>[20]</sup>提出基于注意力特征金字塔网络的肺结节检测算法,该算法以双路径网络为骨干网络,并结合特征金字塔网络进行多尺度预测,同时嵌入全局注意力机制,提高算法的敏感度,但是该算法提高敏感度的同时需要消耗更多计算机资源。张福玲等<sup>[21]</sup>提出一种融合注意力机制和特征金字塔网络的CT图像肺结节检测算法,该算法以ResNet为骨干网络的特征提取网络中设计通道-空间注意力机制来尽可能同时获取上下文语义及位置空间特征信息,但是ResNet作为骨干网络将产生更多的网络层,网络模型参数较为庞大,运行效率较慢。王国刚等<sup>[22]</sup>提出一种基于注意力机制和多空间金字塔池化的实时目标检测算法,该算法采用多空间金字塔池化,提取局部特征和全局特征,通过压缩激励通道注意力机制,自适应调整特征图各个通道的权重,并在检测头网络中使用深度可分离卷积,减少网络参数量,该算法虽然采用深度可分离卷积网络,但仅在检测头网络中,因此并未能有效地降低模型参数。针对上述文献在肺部CT图像检测中存在的问题,本文提出一种高效注意力特征金字塔网络模型。本文主要工作:(1)采用深度可分离卷积网络作为特征提出器;(2)将注意力机制融入特征金字塔网络结构中,使网络模型既保留深度可分离卷积网络的低消耗,又具有注意力机制特征金字塔网络的高准确性。

## 1 相关技术

### 1.1 特征金字塔

特征金字塔作为一种多尺度特征融合方式,可以同时保留高层的语义特征信息和低层的位置信息,因此在语义分割<sup>[23]</sup>、目标检测<sup>[24]</sup>、行为识别<sup>[25]</sup>等领域有着广泛的应用。根据特征金字塔的使用,可分为自底向上的特征提取、自顶向下的语义填充和横向连接的特征融合。通过特征金字塔可以解决卷积过程中低层特征丢失的问题。特征金字塔结构如图1所示。

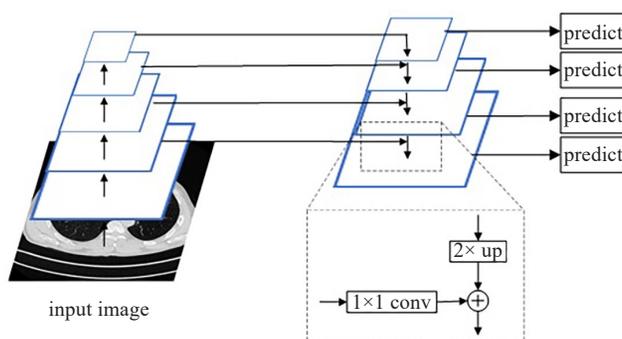


图1 特征金字塔网络  
Figure 1 Feature pyramid network

### 1.2 注意力机制

注意力机制作为一种特征权值分配机制,可以有效突显重要特征信息并抑制背景特征信息,其核心思想是根据特征重要程度分配不同的权值。如图2所示的注意力机制门<sup>[26]</sup>,根据输入特征 $g$ 和 $p$ ,通过 $1 \times 1 \times 1$ 卷积后,进行对应位置特征相加,然后经过ReLU函数抑制负值特征,接着经过Sigmoid函数对特征进行归一化,将处理后的结果与初始特征 $p$ 进行逐通道相乘得到分配权值后的新特征 $p'$ 。

## 2 深度可分离注意力特征金字塔网络

针对网络在特征提取过程中带来计算量增大的问题,本文采用一种深度可分离卷积网络来完成特征提取任务,同时由于网络在特征提取过程中导致特征丢失的问题,本文采用注意力机制的方式,加强重要特征的保护,为重要特征分配更高的权值。网络整体流程:输入图像,通过深度可分离卷积网络,从底向上完成特征提取,将顶层的特征按照从顶向下进行反卷积,并将反卷积特征与同层初始特征经过注意力机制门对初始特征分配不同的权值,最后将重新分配权值的特征进行融合,并经过批处理归一化(BN)操作后输出。网络整体结构如图3所示。

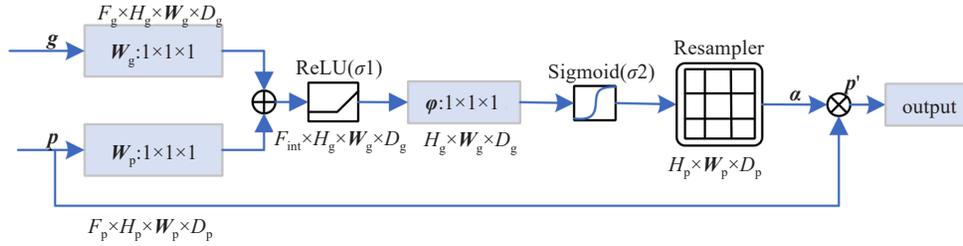


图 2 注意力机制门结构

Figure 2 Attention mechanism gate structure

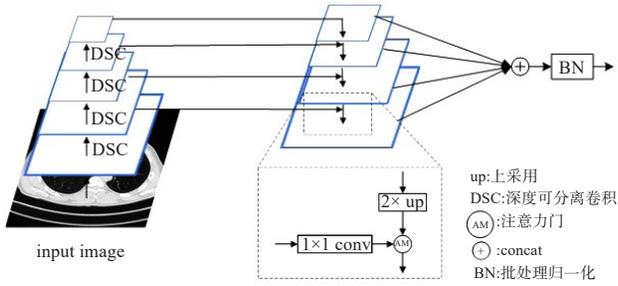


图 3 网络整体结构

Figure 3 Overall network structure

### 2.1 基于深度可分离卷积的网络结构

在神经网络训练过程中,特征提取往往需要消耗更多的计算资源,因此如何设计高效的特征提取网络,对算法的性能起着至关重要的作用。本文采用一种深度可分离卷积网络结构,其主要分为两部分:一是深度卷积,二是逐点卷积。深度卷积是一个卷积核只处理一个通道特征,将每个卷积核处理后的特征图堆叠在一起,因此输入和输出特征图的通

道数相同,如图 4 所示。因为深度卷积只处理长宽方向的空间信息,需要通过逐点卷积处理跨通道方向的信息,即使用  $1 \times 1$  卷积对跨通道维度处理。如对于一个  $[C, N, N]$  的输入数据,卷积核为  $[C, K, K]$  包含  $B$  个普通的卷积操作,需要执行  $N \times N \times K \times K \times C \times B$  次计算,采用深度可分离卷积操作需要执行  $N \times N \times K \times K \times C + N \times N \times C \times B$  次,所以深度可分离卷积操作计算量为普通卷积计算量的  $\frac{1}{B} + \frac{1}{K^2}$ 。因此通过该处理方式能

大大减少算法参数量,提高计算效率。本文将深度可分离卷积作为主干网络的构建块,用于特征提取可以极大地降低网络参数,提高算法的运行效率。图 4 展示了深度可分离卷积网络的执行过程,其中  $M$  由式(1)计算得出:

$$M = \frac{N - K + 2P}{S} + 1 \quad (1)$$

其中,  $N$  为输入数据长度,  $K$  为卷积核长度,  $P$  为填充大小,  $S$  为移动步长。

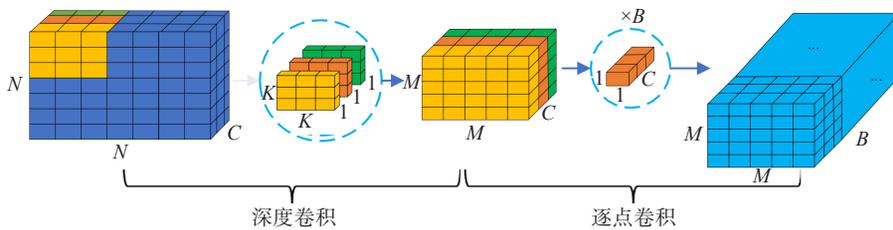


图 4 深度可分离卷积网络

Figure 4 Depthwise separable convolutional neural network

### 2.2 基于注意力特征金字塔网络

在特征提取过程中,会导致特征丢失问题,为加强对重要特征的保护,本文采用注意力机制对重要特征信息分配更大的权重防止在采样过程中出现重要特征丢失的问题。融合注意力机制的特征金字塔网络结构如图 5 所示,输入图片通过深度可分离网络进行特征的逐层提取,对应的特征值分别为  $p_2, p_3, p_4,$

$p_5, p_6$ , 然后再按照自顶向下进行反卷积,对应的特征值分别为  $g_6, g_5, g_4, g_3, g_2$ , 其中  $g_4$  是  $p_5$  和  $g_3$  经过注意力机制门后反卷积得到,  $g_3$  和  $g_2$  也由同样的原理得出。

AM 为注意力机制门,其网络结构如图 6 所示,对于输入特征矩阵  $g$  和  $p$ , 做  $1 \times 1 \times 1$  的卷积操作和 BN 操作后,将输出的结果按位相加,凸显重要特征,接着将相加后的结果经过 ReLU 函数,过滤背景特征,再次通过  $1 \times 1 \times 1$

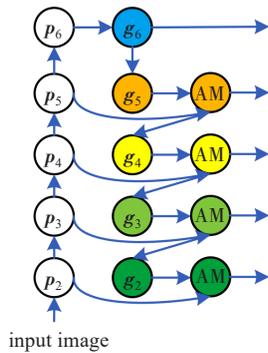


图5 融合注意力机制的特征金字塔网络

Figure 5 Feature pyramid network fused with attention mechanism

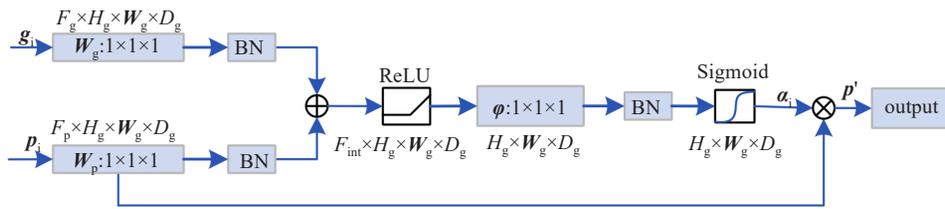


图6 改进的注意力机制门结构

Figure 6 Improved attention mechanism gate structure

### 3 实验与分析

#### 3.1 数据集

数据集1<sup>[27]</sup>为TCIA数据库中的Lung-PET-CT-Dx肺部癌症诊断数据集,该数据集格式为DICOM,大小为127.2 G,标注文件为XML格式,包含4分类:字母“A”为腺癌,字母“B”为小细胞癌,字母“E”为大细胞癌,字母“G”为鳞状细胞癌。同时可用于肺癌目标分类、目标分割和目标检测等任务。由于样本数据存在不均衡问题,大细胞癌和鳞状细胞癌样本偏少,因此在数据处理时,只保留腺癌和小细胞癌数据。经过整理后的腺癌和小细胞癌图片为1311张512×512大小的JPG格式图片,并随机地将腺癌和小细胞癌这两类数据按照70%、20%、10%的比例分成训练集、验证集和测试集。

数据集2<sup>[28]</sup>为Luna16数据集,该数据集是最大公用肺结节数据集LIDC-IDRI的子集,该数据集由美国国家癌症研究所部署,广泛应用于检测和分类早期肺癌。它包含1186个医学影像数据,来自888例病人的低剂量肺部CT影像,数据格式为mhd。经过数据整理和分析后包含1186张295×295大小的PNG格式图片,并随机地将Luna16数据集按照70%、20%、10%的比例分成训练集、验证集和测试集。

#### 3.2 实验环境

本研究的配置环境为Windows11系统的工作

站,运行在AMD Ryzen 7 5800H with Radeon Graphics 3.20 GHz, RAM 16 GB和NVIDIA GeForce RTX 3050 Laptop GPU 4 GB,使用Pycharm 2022编译环境,Pytorch1.12.1框架。在模型训练过程中,采用分类损失、定位损失和置信度多重损失函数,批处理大小batchsize=2,模型初始学习率为0.001,选择Adam随机梯度下降法作为优化器对模型的参数进行更新。

最后将权重 $\alpha_i$ 与特征矩阵 $p_i$ 逐通道相乘,重新调整特征矩阵 $p_i$ 的参数大小。其数学公式如式(2)~式(4)所示:

$$q_{att} = \varphi^T (\sigma_1 (W_p^T p_i + W_g^T g_i + b_g)) + b_\varphi \quad (2)$$

$$\alpha_i = \sigma_2 (q_{att} (p_i, g_i; \theta_{att})) \quad (3)$$

$$p'_i = p_i \cdot \alpha_i \quad (4)$$

其中, $\sigma_1$ 为ReLU函数, $\sigma_2$ 为Sigmoid激活函数,BN为批处理归一化函数, $W_p \in R^{F_p \times F_{int}}$ , $W_g \in R^{F_g \times F_{int}}$ , $\varphi \in R^{F_{int} \times 1}$ 为单位矩阵, $b_g \in R^{F_{int}}$ , $b_\varphi \in R$ 为偏移量, $R$ 为实数,T为矩阵转置。

站,运行在AMD Ryzen 7 5800H with Radeon Graphics 3.20 GHz, RAM 16 GB和NVIDIA GeForce RTX 3050 Laptop GPU 4 GB,使用Pycharm 2022编译环境,Pytorch1.12.1框架。在模型训练过程中,采用分类损失、定位损失和置信度多重损失函数,批处理大小batchsize=2,模型初始学习率为0.001,选择Adam随机梯度下降法作为优化器对模型的参数进行更新。

#### 3.3 评价指标

为了验证本文提出方法的有效性,采用定量评价指标,包括精度(Precision)、召回率(Recall)和mAP值,计算公式如式(5)~式(9)所示:

$$\text{Precision} = \frac{N_{TP}}{N_{TP} + N_{FP}} \quad (5)$$

$$\text{Recall} = \frac{N_{TP}}{N_{TP} + N_{FN}} \quad (6)$$

其中, $N_{TP}$ 表示真阳性,即将阳性样本预测为阳性的个数; $N_{FP}$ 表示假阳性,即将阴性样本预测为阳性的个数; $N_{FN}$ 表示假阴性,即将阳性样本预测为阴性的个数; $N_{TN}$ 表示真阴性,即将阴性样本预测为阴性的个数。

mAP值在目标检测中作为预测目标位置和类别的重要度量方法,物理含义为Precision-Recall曲线围成的平均面积。在计算Pinterp时采用的一种插值方法:即对于某个Recall值 $r$ ,Pinterp值为取所有

Recall> $r$ 中的最大Precision值。

$$\text{Pinterp}(r) = \max_{\tilde{r} > r} (p(\tilde{r})) \quad (7)$$

其中, $p$ 为Precision值, $\tilde{r}$ 为Recall值。

$$\text{AP}_i = \sum_{i=1}^{n-1} (r_{i+1} - r_i) \text{Pinterp}(r_i) \quad (8)$$

其中, $r_1, r_2, \dots, r_n$ 是按升序排列的Precision插值段第一个插值处对应的Recall值。mAP则是 $\text{AP}_i$ 的平均值:

$$\text{mAP} = \frac{\sum_{i=1}^k \text{AP}_i}{k} \quad (9)$$

### 3.4 交叉验证

为验证模型在数据集上的泛化能力和降低模型在数据样本上过拟合风险,采用10折交叉验证的方式,即将数据分成10份,每一份单独的数据被保留作

为验证模型的数据,其他9份样本用来训练。通过10倍的数据进行训练来降低模型的过拟合风险,同时根据每次的训练结果来验证模型的泛化能力。本次交叉验证将在数据集Lung-PET-CT-Dx和数据集Luna16上进行。通过表1可知,经过10折交叉验证,在Lung-PET-CT-Dx数据集上评价指标mAP@0.5,10折交叉验证的标准差为0.01,评价指标mAP@0.5:0.95,10折交叉验证的标准差为0.02。在Luna16数据集上评价指标mAP@0.5,10折交叉验证的标准差为0.01,评价指标mAP@0.5:0.95,10折交叉验证的标准差为0.01。该标准差值较小,可以认为本文算法在训练集以外的数据集上仍然具有较好的检测能力,证明本文算法具有较好的泛化能力。

表1 在数据集上10折交叉验证的结果

Table 1 Results of 10-fold cross-validation on the dataset

| 验证方法    | Lung-PET-CT-Dx |              | Luna16  |              |
|---------|----------------|--------------|---------|--------------|
|         | mAP@0.5        | mAP@0.5:0.95 | mAP@0.5 | mAP@0.5:0.95 |
| 1折交叉验证  | 0.9715         | 0.5997       | 0.8993  | 0.4968       |
| 2折交叉验证  | 0.9462         | 0.5927       | 0.8899  | 0.4776       |
| 3折交叉验证  | 0.9560         | 0.5882       | 0.8739  | 0.4936       |
| 4折交叉验证  | 0.9544         | 0.5991       | 0.8827  | 0.4842       |
| 5折交叉验证  | 0.9761         | 0.5906       | 0.8831  | 0.4824       |
| 6折交叉验证  | 0.9761         | 0.5987       | 0.8920  | 0.5061       |
| 7折交叉验证  | 0.9654         | 0.5859       | 0.8844  | 0.4729       |
| 8折交叉验证  | 0.9594         | 0.5906       | 0.8938  | 0.4705       |
| 9折交叉验证  | 0.9492         | 0.6191       | 0.8702  | 0.4694       |
| 10折交叉验证 | 0.9517         | 0.6624       | 0.8987  | 0.4925       |
| 标准差     | 0.01           | 0.02         | 0.01    | 0.01         |

### 3.5 消融实验

为证明本文算法的有效性,将网络的改进部分进行消融实验,实验结果在Lung-PET-CT-Dx数据集上完成,其中特征金字塔网络作为主干网络,残差卷积网络(ResNet50)和深度可分离卷积网络分别作为特征提取器,AM为注意力机制模块。通过表2可知,算法1特征金字塔网络和残差卷积网络的组合在网络参数大小上要远大于算法2特征金字塔网络和深度可分离卷积网络的组合,证明深度可分离卷积网络能有效降低网络参数。算法3在算法1的基础上添加注意力机制,虽然网络参数大小有所提升,但是在mAP值上相比算法1有较大的提高,证明注意力机制可以有效提高算法的检测准确性。本文算法采

用深度可分离卷积网络加上注意力机制的特征金字塔网络算法,该算法相比前面的算法在单项性能上并不是最优的,但是兼顾网络参数大小和mAP值两方面却是最好的算法。因此,通过消融实验可以证明本文算法相比初始算法性能有很好的提升。

### 3.6 对比实验

为进一步比较本文算法相比其它算法的优越性,将在Lung-PET-CT-Dx和Luna16数据集上与现有的算法<sup>[16-19]</sup>进行比较,实验结果如表3和表4所示。表3为不同算法在Lung-PET-CT-Dx数据集上训练500轮的实验结果,通过比较可知本文算法的精度、召回率、mAP@0.5和mAP@0.5:0.95值均为最优,相比文献[16]算法的精度提升27.51%,召回率提升

表2 不同算法的性能比较

Table 2 Performance comparison among different algorithms

| 算法   | 特征金字塔网络 | ResNet50 | 深度可分离卷积网络 | 注意力机制 | 参数大小/Mb | mAP@0.5 | mAP@0.5:0.95 |
|------|---------|----------|-----------|-------|---------|---------|--------------|
| 算法1  | √       | √        | ×         | ×     | 39.8209 | 0.8721  | 0.4912       |
| 算法2  | √       | ×        | √         | ×     | 18.4883 | 0.8651  | 0.4821       |
| 算法3  | √       | √        | ×         | √     | 41.6873 | 0.9703  | 0.6119       |
| 本文算法 | √       | ×        | √         | √     | 19.8776 | 0.9626  | 0.6027       |

8.43%, mAP@0.5 提升 12.17%, mAP@0.5:0.95 提升 13.43%。相比文献[17]算法的精度提升 14.32%, 召回率提升 5.73%, mAP@0.5 提升 8.86%, mAP@0.5:0.95 提升 10.39%。相比文献[18]算法的精度提升 8.34%, 召回率提升 2.53%, mAP@0.5 提升 4.31%, mAP@0.5:0.95 提升 4.82%。相比文献[19]算法的精度提升 5.92%, 召回率提升 1.63%, mAP@0.5 提升 2.23%, mAP@0.5:0.95 提升 3.22%。

表3 在 Lung-PET-CT-Dx 数据集上的实验结果

Table 3 Experimental results on Lung-PET-CT-Dx dataset

| 算法     | 精度     | 召回率    | mAP@0.5 | mAP@0.5:0.95 |
|--------|--------|--------|---------|--------------|
| 文献[16] | 0.6335 | 0.8306 | 0.8409  | 0.4684       |
| 文献[17] | 0.7654 | 0.8576 | 0.8740  | 0.4988       |
| 文献[18] | 0.8252 | 0.8896 | 0.9195  | 0.5545       |
| 文献[19] | 0.8494 | 0.8986 | 0.9403  | 0.5705       |
| 本文算法   | 0.9086 | 0.9149 | 0.9626  | 0.6027       |

表4 在 Luna16 数据集上的实验结果

Table 4 Experimental results on Luna16 dataset

| 算法     | 精度     | 召回率    | mAP@0.5 | mAP@0.5:0.95 |
|--------|--------|--------|---------|--------------|
| 文献[16] | 0.7960 | 0.7796 | 0.7931  | 0.4002       |
| 文献[17] | 0.8041 | 0.7967 | 0.8116  | 0.4028       |
| 文献[18] | 0.8059 | 0.7930 | 0.8251  | 0.4278       |
| 文献[19] | 0.8294 | 0.8778 | 0.8697  | 0.4707       |
| 本文算法   | 0.8846 | 0.9032 | 0.8868  | 0.4996       |

表4为不同算法在Luna16数据集上训练1000轮的实验结果,本文算法在评价指标中表现出更好的效果,相比文献[16]算法的精度提升8.86%,召回率提升12.36%, mAP@0.5 提升9.37%, mAP@0.5:0.95 提升9.94%;相比文献[17]算法的精度提升8.05%,召回率提升10.65%, mAP@0.5 提升7.52%, mAP@0.5:0.95 提升9.68%;相比文献[18]算法的精

度提升7.87%,召回率提升11.02%, mAP@0.5 提升6.17%, mAP@0.5:0.95 提升7.18%;相比文献[18]算法的精度提升5.52%,召回率提升2.54%, mAP@0.5 提升1.71%, mAP@0.5:0.95 提升2.89%。

在算法原理上,相比文献[16]中的算法,本文算法采用多重损失函数,对小目标检测效果更优;相比文献[17]中的算法,本文算法采用多尺度特征融合的方式,充分利用上下文语义特征信息提高检测准确性;相比文献[18]中的算法,本文算法采用更加灵活的多尺度融合网络,并将注意力机制融入,提高算法对小目标检测精度;相比文献[19]中的算法,本文算法将注意力机制融入特征金字塔网络结构中,对重要语义特征信息分配更高权重,凸显在网络中的作用,因此可以证明本文算法在肺结节目标检测中具有更高的检测性能。

### 3.7 检测结果

对比实验通过验证集证明本文算法的优越性,检测结果将通过测试集来进一步证明算法的优越性,如通过不同算法在测试集数据上进行目标检测,实验结果如图7和图8所示。通过图7可知,本文算法在目标特征不明确的基础上以较高的准确度检测出肺部CT图像的结节所在位置,并且与其它算法相比,本文算法在检测准确度方面要优于其它算法,其检测结果准确度基本达到90%以上,可以作为医生在肺癌诊断上的参考结果。图8是在阈值为0.5的基础上进行的肺结节检测结果,低于该阈值将不显示检测效果。通过图8可知,本文算法在Luna16数据集中,肺结节特征偏小的基础上,可以有效地检测到目标所在位置,相比其它算法准确性整体更高,在肺结节目标检测过程中,能辅助医生进行诊断,具有更好的临床价值。

## 4 结语

本文提出一种高效的注意力机制特征金字塔网络,该网络通过深度可分离网络作为特征提取器,同时将注意力机制算法融入特征金字塔网络中。相比

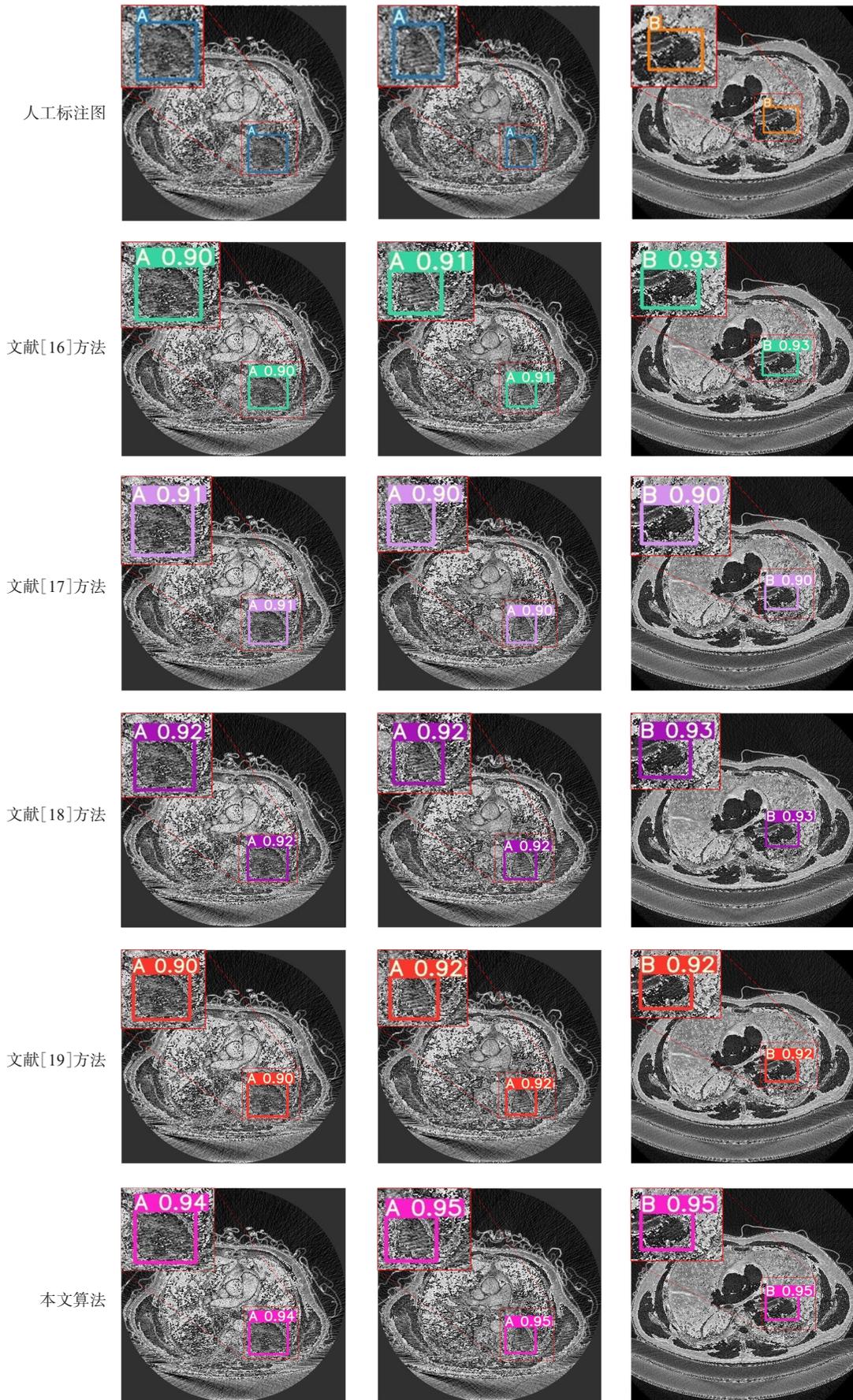


图7 在 Lung-PET-CT-Dx 数据集上的检测结果(A 数值表示检测为腺癌的精度,B 数值表示检测为小细胞癌的精度)  
Figure 7 Detection results on Lung-PET-CT-Dx dataset (A values indicate the precision of detecting as adenocarcinoma detection, B values indicate the precision of detecting as small cell carcinoma)

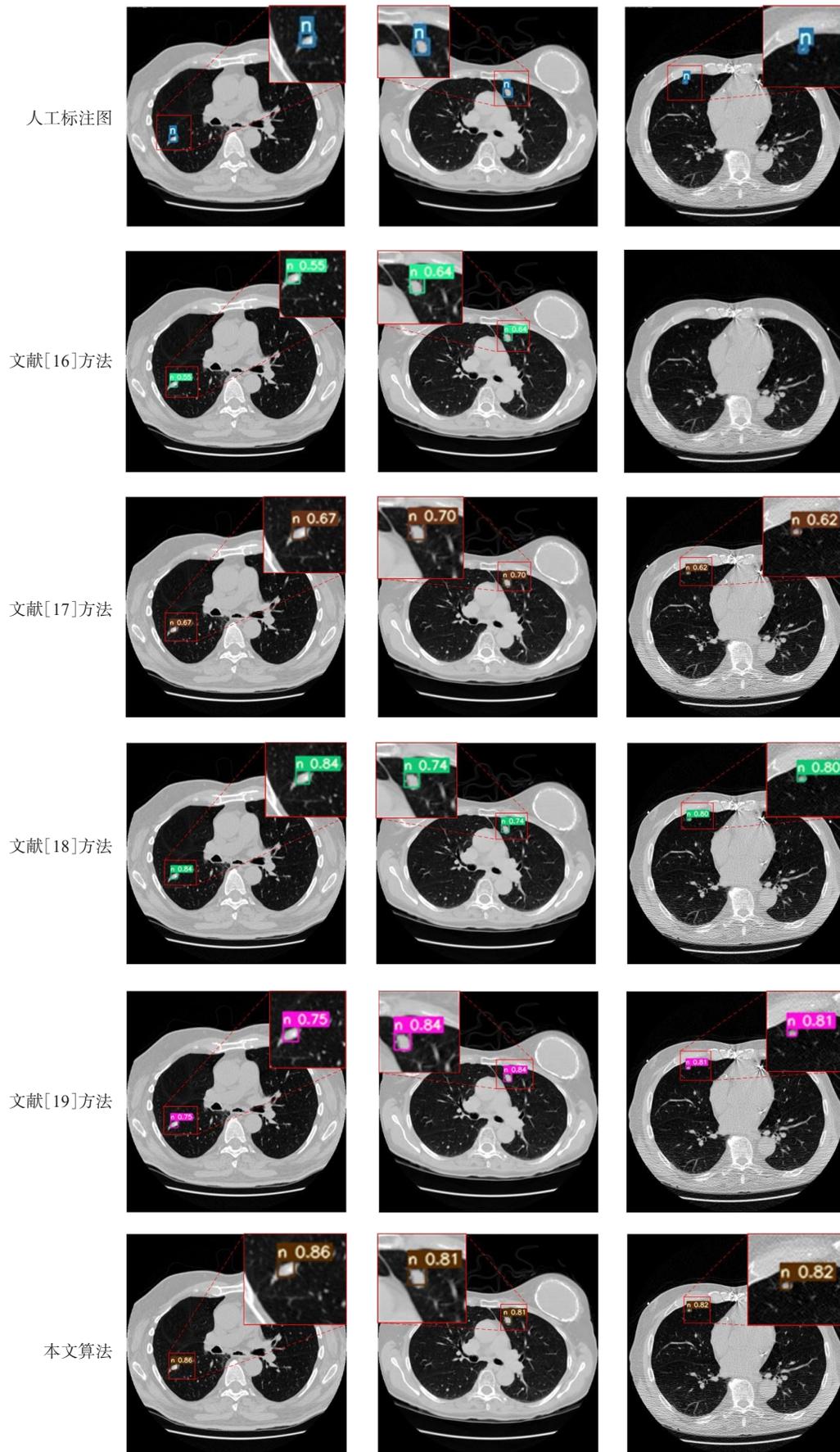


图8 在Luna16数据集上的检测结果(图中数值表示检测为肺结节的精度)

Figure 8 Detection results on Luna16 dataset (where the value indicates the precision of detecting as lung nodules)

其它算法,本文算法能有效解决目标检测精度低、消耗计算资源过多的问题。通过实验结果可知,本文算法在计算资源消耗和目标检测的准确性上综合评价要比其它算法性能更优越,有效证明本文算法的优越性。在临床上,可以作为辅助医生诊断的有效参考,提升医生的诊断效率,降低出现漏诊、误诊的概率。但是本文在目标检测过程中存在数据样本不均衡的问题,导致网络无法学习小样本数据,这也是下一步研究的主要方向。

## 【参考文献】

- [1] Leiter A, Veluswamy RR, Wisnivesky JP. The global burden of lung cancer: current status and future trends[J]. *Nat Rev Clin Oncol*, 2023, 20(9): 624-639.
- [2] Zhong J, Luo Z, Sun XF, et al. Psychological impact and workload of COVID-19 on healthcare workers in China during the early time of the pandemic: a cross-sectional study[J]. *Disaster Med Public Health Prep*, 2023, 17: e170.
- [3] Liu W, Anguelov D, Erhan D, et al. SSD: single shot MultiBox detector [C]//Computer Vision-ECCV 2016. Cham: Springer International Publishing, 2016: 21-37.
- [4] Lin TY, Goyal P, Girshick R, et al. Focal loss for dense object detection [C]//2017 IEEE International Conference on Computer Vision (ICCV). Piscataway, NJ, USA: IEEE, 2017: 2999-3007.
- [5] Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real-time object detection[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2016: 779-788.
- [6] Redmon J, Farhadi A. YOLO9000: better, faster, stronger[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2017: 6517-6525.
- [7] Redmon J, Farhadi A. YOLOv3: an incremental improvement[EB/OL]. (2018-04-08). <https://arxiv.org/abs/1804.02767>.
- [8] Bochkovskiy A, Wang CY, Liao HY. YOLOv4: optimal speed and accuracy of object detection[EB/OL]. (2020-04-23). <https://arxiv.org/abs/2004.10934>.
- [9] Li CY, Li LL, Jiang HL, et al. YOLOv6: a single-stage object detection framework for industrial applications[EB/OL]. (2022-09-07). <https://arxiv.org/abs/2209.02976>.
- [10] Wang CY, Bochkovskiy A, Liao HY. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2023: 7464-7475.
- [11] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//2014 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE, 2014: 580-587.
- [12] He KM, Zhang XY, Ren SQ, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. *IEEE Trans Pattern Anal Mach Intell*, 2015, 37(9): 1904-1916.
- [13] Girshick R. Fast R-CNN[C]//2015 IEEE International Conference on Computer Vision (ICCV). Piscataway, NJ, USA: IEEE, 2015: 1440-1448.
- [14] Ren SQ, He KM, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[C]//Proceedings of the 28th International Conference on Neural Information Processing Systems. Cambridge, MA, USA: MIT Press, 2015: 91-99.
- [15] Lin TY, Dollár P, Girshick R, et al. Feature pyramid networks for object detection[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2017: 936-944.
- [16] Su Y, Li D, Chen XD. Lung nodule detection based on faster R-CNN framework [J]. *Comput Methods Programs Biomed*, 2021, 200: 105866.
- [17] Loraksa C, Mongkolsomlit S, Nimsuk N, et al. Development of the osteosarcoma lung nodules detection model based on SSD-VGG16 and competency comparing with traditional method[J]. *IEEE Access*, 2022, 10: 65496-65506.
- [18] Mahum R, Al-Salman AS. Lung-RetinaNet: lung cancer detection using a RetinaNet with multi-scale feature fusion and context module [J]. *IEEE Access*, 2023, 11: 53850-53861.
- [19] Yan CM, Wang C. Automatic detection and localization of pulmonary nodules in CT images based on YOLOv5[J]. *J Comput*, 2022, 33(3): 113-123.
- [20] 秦源源, 张鸿. 基于注意力特征金字塔网络的肺结节检测算法[J]. *计算机应用*, 2023, 43(7): 2311-2318.  
Qin YY, Zhang H. Pulmonary nodule detection algorithm based on attention feature pyramid networks [J]. *Journal of Computer Applications*, 2023, 43(7): 2311-2318.
- [21] 张福玲, 张少敏, 支力佳, 等. 融合注意力机制和特征金字塔网络的CT图像肺结节检测[J]. *中国图象图形学报*, 2021, 26(9): 2156-2170.  
Zhang FL, Zhang SM, Zhi LJ, et al. Detection of pulmonary nodules in CT images by combining an attention mechanism and a feature pyramid network[J]. *Journal of Image and Graphics*, 2021, 26(9): 2156-2170.
- [22] 王国刚, 李泽欣, 董志豪. 基于注意力机制和多空间金字塔池化的实时目标检测算法[J]. *计算机测量与控制*, 2024, 32(2): 56-64.  
Wang GG, Li ZX, Dong ZH. Real-time object detection algorithm based on attention mechanism and multi-spatial pyramid pooling[J]. *Computer Measurement & Control*, 2024, 32(2): 56-64.
- [23] 杨建华. 基于FPN的高分辨率建筑物图像的语义分割[J]. *遥感信息*, 2021, 36(4): 133-141.  
Yang JH. Building extraction from high-resolution imagery based on FPN[J]. *Remote Sensing Information*, 2021, 36(4): 133-141.
- [24] Luo YH, Cao X, Zhang JT, et al. CE-FPN: enhancing channel information for object detection[J]. *Multimed Tools Appl*, 2022, 81(21): 30685-30704.
- [25] 何嘉宇, 雷军, 李国辉. 特征金字塔结构的时序行为识别网络[J]. *中国图象图形学报*, 2021, 26(7): 1637-1647.  
He JY, Lei J, Li GH. Temporal action detection based on feature pyramid hierarchies[J]. *Journal of Image and Graphics*, 2021, 26(7): 1637-1647.
- [26] Oktay O, Schlemper J, Le Folgoc L, et al. Attention U-net: learning where to look for the pancreas[EB/OL]. (2018-05-20). <https://arxiv.org/abs/1804.03999>.
- [27] Clark K, Vendt B, Smith K, et al. The cancer imaging archive (TCIA): maintaining and operating a public information repository[J]. *J Digit Imaging*, 2013, 26(6): 1045-1057.
- [28] Setio AA, Traverso A, de Bel T, et al. Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: the LUNA16 challenge[J]. *Med Image Anal*, 2017, 42: 1-13.

(编辑:陈丽霞)