

基于多尺度卷积神经网络的 X 光图像中肺炎病灶检测

张物华, 李强*, 关欣

天津大学微电子学院, 天津 300072

摘要 肺炎检测在医学图像处理中具有重要的研究意义, 针对当前经典检测算法对肺炎病灶检测精度较低的问题, 本文提出一种基于多尺度卷积神经网络的 X 光图像中肺炎病灶检测算法。在基础特征提取网络中加入特征通道注意力模块, 突出特征图中含有大量肺炎病灶信息的特征通道, 抑制不含病灶信息或者含有大量无用信息的特征通道, 形成高质量特征图; 然后通过统计分析, 使用聚类算法设计了一系列不同宽高比以及缩放尺度的候选框以适用于肺炎病灶检测。同时, 在含有肺炎病灶的胸部 X 光图像数据集上进行了单模型和多模型检测实验, 其中单模型下检测精度为 82.52%, 多模型融合下检测精度为 89.08%。通过对比实验与结果分析, 验证了本文算法在检测精度方面优于当前其他检测算法, 适用于 X 光图像中肺炎病灶检测。

关键词 图像处理; 肺炎病灶; 特征通道注意力模块; 聚类算法; 卷积神经网络

中图分类号 TP391.4

文献标志码 A

doi: 10.3788/LOP57.081020

Detection of Pneumonia Lesions in X-Ray Images Based on Multi-Scale Convolutional Neural Networks

Zhang Wuhua, Li Qiang*, Guan Xin

School of Microelectronics, Tianjin University, Tianjin 300072, China

Abstract Pneumonia detection has important research significance in medical image processing. For the problem that the current classical detection algorithms has low accuracy in detecting pneumonia lesions. This paper presents an algorithm for detecting pneumonia lesions in X-ray images based on multi-scale convolutional neural networks. The feature channel attention module is added to the basic feature extraction network to highlight the channel containing useful information in the feature map, and to suppress the feature channel without lesion information or containing a large amount of useless information to form a high-quality feature map. Then through statistical analysis, a series of candidate frames with different aspect ratios and scaling scales are designed using clustering algorithm to be suitable for pneumonia lesion detection. In this paper, the single-model and multi-model detection experiments are performed on chest X-ray datasets containing pneumonia lesions. The detection accuracy is 82.52% in the case of single model and 89.08% in the case of multi-model fusion. Through comparison experiments and results analysis, the proposed algorithm is superior to other detection algorithms in pneumonia lesion detection and is suitable for pneumonia lesion detection in X-ray images.

Key words image processing; pneumonia lesions; feature channel attention module; clustering algorithm; convolutional neural network

OCIS codes 100.2960; 170.3880; 100.4996

1 引言

肺炎是一种影响肺部功能的急性呼吸道感染疾

病, 肺炎会使肺泡发炎导致呼吸道充满脓液, 使得人体呼吸困难^[1]。世界卫生组织数据显示, 2015 年, 肺炎估计造成全世界 92 万名 5 岁以下儿童死亡, 该

收稿日期: 2019-08-08; 修回日期: 2019-09-11; 录用日期: 2019-09-17

基金项目: 国家自然科学基金(61471263)、天津市自然科学基金(16JCYDJC31100)

* E-mail: liqiang@tju.edu.cn

死亡人数占5岁以下儿童死亡人数的16%，其中发展中国家比例较大^[2]。目前，主要通过胸部X光检查来诊断肺炎，肺炎在胸部X光图像中表现为不透明度增加的区域^[3]。由于非专业人员难以区分胸部X光图像中肺炎病灶区域和正常区域，因此需要具备较高医疗图像判断能力的医生对胸部X光图像进行对比观察，最后诊断肺炎病灶位置。然而，这是一项具有挑战性的工作，诊断过程需要大量的时间和精力，且容易受主观因素影响导致误诊。

近年来，随着卷积神经网络(CNN)为代表的深度学习算法在图像处理领域的快速发展，许多研究人员将深度学习引入计算机辅助诊断(CAD)^[4]系统中，广泛应用于医疗影像学领域。在胸部X光图像领域，国内外学者纷纷展开了研究。Wang等^[5]在Chest X-ray 8数据集上使用分类网络输出诊断结果，同时将阈值法应用于每个类的可视化特征图中并导出每个胸部疾病的边界框，但是，结果显示模型生成的预测标注框通常比真实标注框大，因此精度并不高；Li等^[6]将弱监督学习方法和多实例学习方法结合，提出端到端神经网络模型，在只包含少量标注框的数据集中进行训练，实现了多种胸部疾病的检测，但是由于标注信息量过少，模型在肺炎病灶检测方面的精度较低；Ma等^[7]提出多注意力机制网络用于胸部疾病分类与检测，在Chest X-ray 14数据集上进行实验，但是由于使用热力图作为病灶位置信息参考，因此病灶检测结果并不精确。

利用深度学习中目标检测技术实现胸部X光图像的肺炎检测具有实用价值，准确检测肺炎病灶位置信息可以为医生诊断提供强有力的临床证据。为此，本文提出基于多尺度卷积神经网络的X光图像中肺炎病灶检测算法——SE-RetinaNet，该网络主要包括两方面改进：1)在多尺度卷积神经网络检测算法RetinaNet^[8]的基础上加入特征通道注意力机制模块，形成SE-RetinaNet检测网络，主要作用是突出特征图组中含有病灶信息的特征通道，抑制特征图组中不含病灶信息或者含有大量无用信息的特征通道；2)根据胸部X光图像数据集中肺炎病灶区域的统计信息，利用K-means聚类算法设计了一系列具有代表性的候选框宽高比(Anchor ratios)和尺度系数用于SE-RetinaNet检测算法中。

本文算法在北美放射协会(Radiological Society of North America, RSNA[®])的胸部X光图像数据集上进行了分析实验。实验结果表明，相较于其他检测算法，SE-RetinaNet在肺炎检测任务中

具有更高的检测精度。

2 肺炎检测网络框架

2018年，Lin等^[8]提出了RetinaNet目标检测算法，RetinaNet使用残差网络(ResNet)^[9]作为特征提取基础框架，融合了特征金字塔结构(FPN)^[10]以及全连接神经网络(FCN)^[11]。RetinaNet网络在交叉熵损失函数中加入类别权重因子以及可调聚焦参数构成Focal Loss损失函数，作用是通过降低容易分类的样本的损失权重，有效解决单阶段目标检测网络训练过程中正负样本区域极不平衡而导致准确率和精度较低的问题。为了提高肺炎检测精度，本文以RetinaNet作为肺炎检测基础网络，在ResNet中加入具有通道注意力作用的压缩-激发(SE)模块^[12]，最终构成SE-RetinaNet肺炎检测网络。

2.1 SE-ResNet 模块结构

SE模块原理如图1所示，假设给定任意卷积 $F_{tr}: \mathbf{X} \mapsto \mathbf{U}$, $\mathbf{X} \in \mathbb{R}^{H \times W \times C'}$, $\mathbf{U} \in \mathbb{R}^{H \times W \times C}$ ，其中 \mathbf{X} 为输入特征图， \mathbf{U} 为卷积后输出的多通道输出特征图，其中特征图 \mathbf{U} 中包含 C 层通道。

首先，为了使网络充分学习各通道的全局特征信息，对 \mathbf{U} 进行特征压缩 F_{sq} 处理，通过在空间维度($H \times W$)上使用全局平均池化层对特征图各通道进行压缩生成通道描述符，该通道描述符空间维度为(1×1)。其中压缩操作表达式为

$$z_c = F_{sq}(u_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W u_c(i, j), \quad (1)$$

式中： $u_c(c=1, 2, \dots, C)$ 代表特征图 \mathbf{U} 中第 c 通道的特征向量； z_c 代表经过压缩后第 c 通道特征向量的通道描述符。

其次，为了使网络充分学习通道间的依赖性以及通道间的非线性相互作用，同时提高网络表达能力，对通道描述符进行两次非线性激发 F_{ex} 操作。其中激发操作表达式为

$$\mathbf{S} = F_{ex}(\mathbf{z}, \mathbf{W}) = \sigma[g(\mathbf{z}, \mathbf{W})] = \sigma[\mathbf{W}_2 \delta(\mathbf{W}_1 \mathbf{z})], \quad (2)$$

式中： σ 代表Sigmoid门函数； δ 代表ReLU激活函数； $\mathbf{W}_1 \in \mathbb{R}^{\frac{C}{r} \times C}$ 代表维度缩减参数， $\mathbf{W}_2 \in \mathbb{R}^{C \times \frac{C}{r}}$ 代表维度增长参数， r 表示缩减因子； $\mathbf{S} = [s_1, s_2, \dots, s_C]$ 代表通道描述符向量； $\mathbf{z} = [z_1, z_2, \dots, z_C]$ 代表经过激发后输出的权重向量。

最后，通过全局信息有选择地调整卷积层输出 \mathbf{U} 每个通道的权重大小，表达式为

$$\tilde{\mathbf{X}} = F_{scale}(\mathbf{u}_c, s_c) = s_c \cdot \mathbf{u}_c, \quad (3)$$

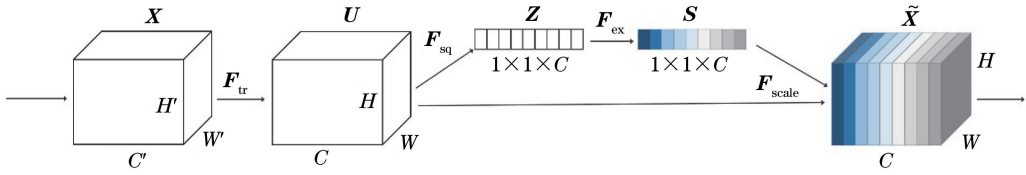


图1 压缩-激发模块

Fig. 1 Block of squeeze and excitation

式中： F_{scale} 代表第 c 通道的特征 $u_c \in \mathbb{R}^{H \times W}$ 与对应权重因子 s_c 相乘操作， $\tilde{\mathbf{X}} = [\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_C]$ 代表特征图加权后的输出。

相比原始残差块 (ResNet)，压缩激发残差块 (SE-ResNet) 在结构上没有额外增加卷积或下采样操作，残差块架构图如图 2 所示，输入向量依次经过 $1 \times 1, 3 \times 3, 1 \times 1$ 三层卷积操作后，输出多通道特征图。首先对多通道特征图使用全局平均池化层进行压缩处理，生成尺寸为 $1 \times 1 \times C$ 的通道描述符。其次为了构建通道描述符之间的非线性依赖关系，在全局池化层后使用全连接层以及 ReLU 非线性激活函数层作为激发处理。然后，通过 Sigmoid 门将通道描述符进行归一化处理。最后对原始残差结构输出的多通道特征图与对应的通道描述符进行加权操作。通过对胸部 X 光图像特征图中各层通道进行压缩-激发处理，有效突出含有肺炎病灶相关信息的特征通道，同时抑制含有大量无关信息的特征通道。

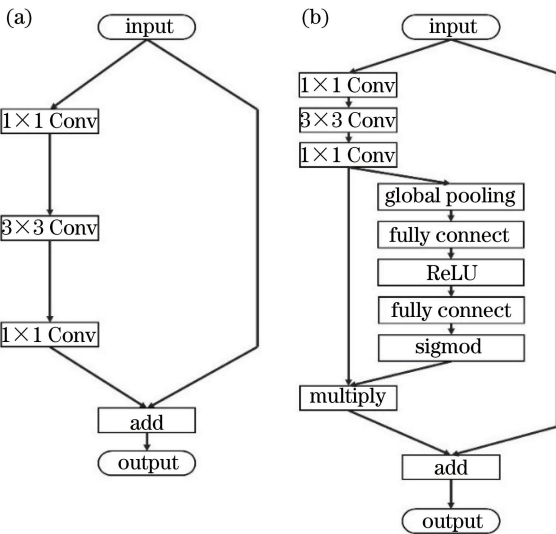


图2 残差块架构图。(a)原始残差块；(b)压缩-激发残差块
Fig. 2 Structure diagram of residual block. (a) Original residual block; (b) SE residual block

2.2 SE-RetinaNet 网络结构

以 SE-ResNet 作为特征提取骨干网络，加入

FPN 和分类、回归子网络，构成 SE-RetinaNet 肺炎检测网络，网络架构如图 3 所示。首先对输入图像进行多尺度特征抽取，构建 5 个不同尺度的特征图 $[P_3, P_4, P_5, P_6, P_7]$ ，不同尺度特征图的每个像素所包含的空间感受野各不相同， $P_3 \sim P_7$ 像素空间感受野逐渐扩大，同时特征图语义信息逐渐增强，浅层特征图有利于检测小物体，而深层特征图则有利于检测大物体。然后根据设定的 Anchor 进行候选框提取，最后将候选框输入到分类子网络与回归子网络进行类别分类与边框回归。SE-RetinaNet 与 RetinaNet 相比，一方面，在没有额外增加网络层数的情况下，对每一个特征通道进行加权处理，使得包含重要信息的通道在训练过程中重点突出；另一方面，虽然在每一个残差块里添加了 SE 模块，但并没有额外增加过多的训练参数。SE-RetinaNet 与 RetinaNet 网络参数量比较如图 4 所示，图中纵坐标表示网络的参数量，横坐标中 50 代表采用 ResNet50 作为基础框架，而 101 则代表采用 ResNet101 作为基础框架。

2.3 肺炎区域候选框设计

候选框的形状和大小会直接影响模型的检测性能，本文对肺炎病灶标注信息进行统计，根据图 5 肺炎病灶区域宽高比特性可知，宽高比通常处于 $0.3 \sim 1.5$ 之间，其中长方形占了 77.3%。根据肺炎区域面积开平方可知，肺炎区域的最小面积为 10×10 ，最大面积为 133×133 ，均值为 58×58 。因此，本文借鉴 YOLO-v2^[13] 目标检测网络中候选框 Anchor Box 生成方法，使用 K-means 聚类算法获取适合肺炎病灶数据分布的候选框宽高比 (Anchor ratios)。具体算法步骤如下：

1) 根据病灶标注信息对所有肺炎病灶区域的宽度和高度进行数值归一化处理，将所有病灶实例的宽度和高度都归一化在 $0 \sim 1$ 范围中，假设肺炎病灶标注信息中包含 N 例病灶实例。

2) 在归一化处理后的标注信息中，随机选取 K 例病灶标注的宽度和高度 (W_j, H_j) 作为聚类中心位置，其中 $j = 1, 2, \dots, K$ 。

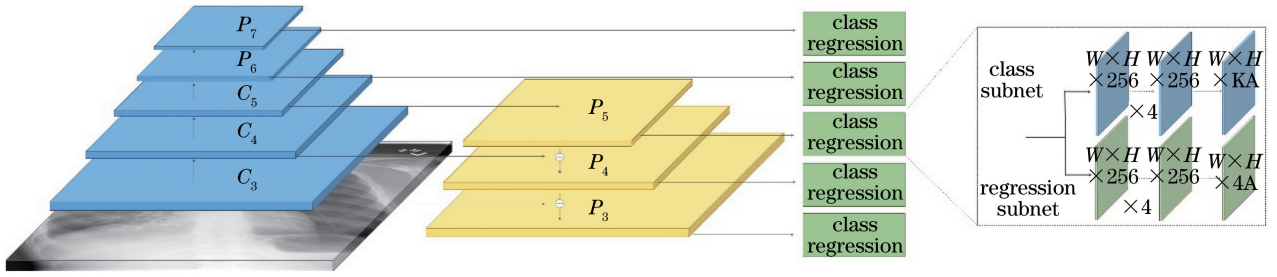


图 3 SE-RetinaNet 网络架构

Fig. 3 Structure diagram of SE-RetinaNet

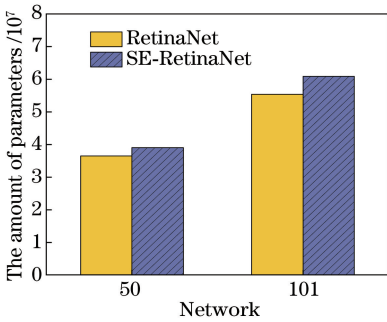


图 4 SE-RetinaNet 与 RetinaNet 网络参数量对比图
Fig. 4 Comparison of SE-RetinaNet and RetinaNet network parameters

3) 计算剩余病灶标注框与各个聚类中心的距离, 将每个标注框视为与其距离最小的聚类中心的簇的元素。第 i 例病灶标注框 (w_i, h_i) 与第 j 聚类中心位置 (W_j, H_j) 距离为 $d_{ij} = 1 - \frac{(w_i, h_i) \cap (W_j, H_j)}{(w_i, h_i) \cup (W_j, H_j)}$, $(i \neq j), i = 1, 2, \dots, N, j = 1, 2, \dots, K$, 其中 $(w_i, h_i) \cap (W_j, H_j)$ 表示第 i 例病灶标注框与以第 j 聚类中心为宽高标注框的重叠面积, $(w_i, h_i) \cup (W_j, H_j)$ 为第 i 例病灶标注框与以第 j 聚类中心为宽高的标注框的合并面积。

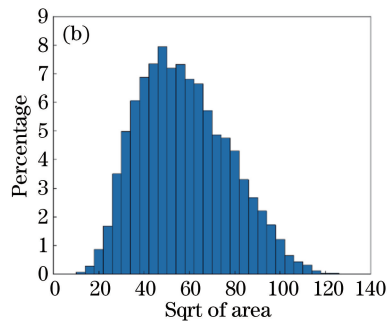
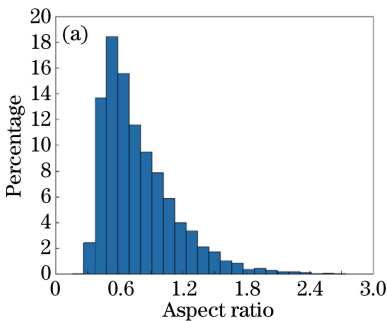


图 5 肺炎病灶数据分布信息统计

Fig. 5 Statistics of pneumonia lesion data distribution information

2.4 损失函数

检测网络在胸部 X 光图像生成候选框的过程

4) 对每个簇中元素求平均, 重新计算各个簇的

聚类中心位置, 计算方法为

$$\begin{cases} W'_j = \frac{1}{N_j} \sum_{z=1}^{N_j} w_{jz} \\ H'_j = \frac{1}{N_j} \sum_{z=1}^{N_j} h_{jz} \end{cases}, j = 1, 2, \dots, K.$$

其中 (W'_j, H'_j) 为重新计算后的聚类中心位置, N_j 为第 j 簇中元素数量, (w_{jz}, h_{jz}) 为第 j 簇中第 z 例病灶标注框的宽度和高度。

5) 计算所有聚类中心位置变化量, 计算方法为

$$\Delta c = \sum_{j=1}^K \left[1 - \frac{(W_j, H_j) \cap (W'_j, H'_j)}{(W_j, H_j) \cup (W'_j, H'_j)} \right], (W_j, H_j) \text{ 与 } (W'_j, H'_j) \text{ 分别为更新前后的聚类中心位置。}$$

当聚类中心位置没有发生改变时重叠面积与合并面积相等, 这时该聚类中心位置变化量为 0。

6) 重复执行步骤 3~5, 当所有聚类中心位置变化量 Δc 小于预设定的最小变化量时, 聚类任务已经完成, 停止循环并返回最新的聚类中心, 同时计算各个聚类中心的宽高比。

最终, 得到的候选框的宽高比 Anchor ratios 为 0.5、0.7、1.0、1.25。此外将候选框尺度系数 Anchor scales 设置为 0.5、0.75、1.0、1.25、1.5、1.7、2.0。

中, 所生成的正负样本候选框的数量严重不平衡, 因此简单而大量的负样本会主导整个网络的训练, 对

肺炎检测网络产生大量无用的学习信号导致模型退化,这会严重影响网络对肺炎病灶部的检测性能。为此,本文使用 Focal Loss 作为本文网络的分类损失函数进行监督训练。

肺炎病灶检测网络损失函数由两部分组成,即图像分类损失函数和预测边框回归损失函数。损失函数定义为

$$L(p_i, p_i^*, t_i, t_i^*) = \frac{1}{N_{\text{cls}}} \sum_i L_{\text{cls}}(p_i, p_i^*) + \lambda \frac{1}{N_{\text{reg}}} \sum_i p_i^* L_{\text{reg}}(t_i, t_i^*), \quad (4)$$

式中: p_i 和 p_i^* 分别是目标预测概率和真实标签(肺炎部位为 1,非肺炎部位为 0); $t_i = \{t_{ix}, t_{iy}, t_{iw}, t_{ih}\}$ 表示肺炎病灶预测框位置信息集合; $t_i^* = \{t_{ix}^*, t_{iy}^*, t_{iw}^*, t_{ih}^*\}$ 表示真实肺炎病灶标注框位置信息集合。 L_{cls} 是 Focal Loss 分类损失函数,定义为

$$L_{\text{cls}}(p_i, p_i^*) = -\alpha (1 - p_i)^\gamma p_i^* \ln(p_i), \quad (5)$$

式中: $\gamma \in [0, 5]$ 为可调节参数, $(1 - p_i)^\gamma$ 为调制因子。训练过程中,当 $\gamma > 0$ 时,降低易于分类的样本对网络模型总损失减少量的贡献,使网络模型更加专注于学习难于分类的样本的特征, $\alpha \in [0, 1]$ 为类别权重因子,通过调节 α 使得避免网络模型对大量的负样本进行学习,从而提高模型对肺炎检测的精度。而 L_{reg} 是边框回归损失函数,定义为

$$L_{\text{reg}}(t_i, t_i^*) = \sum_{v \in \{x, y, w, h\}} f_{\text{smooth } L_1}(t_{iv} - t_{iv}^*), \quad (6)$$

式中: $f_{\text{smooth } L_1}$ 表达式为

$$f_{\text{smooth } L_1} = \begin{cases} 0.5x^2, & |x| < 1 \\ |x| - 0.5, & |x| \geq 1 \end{cases} \quad (7)$$

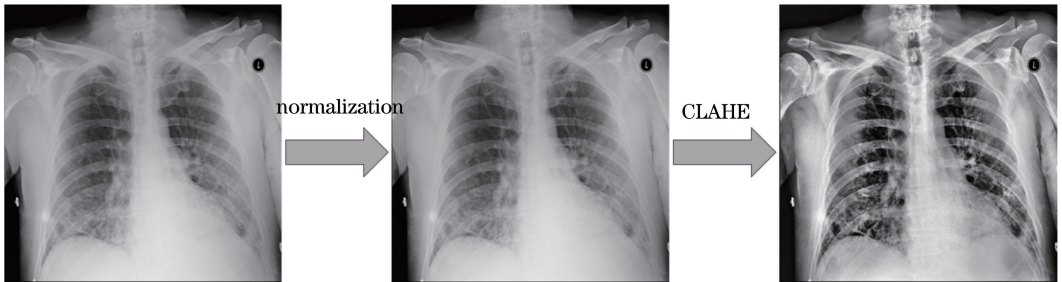


图 6 胸部 X 光图像预处理过程

Fig. 6 Pre-processing of chest X-ray images

3.2 参数设置与训练

为了提高检测网络学习效率,本文将 CXR 图像分辨率调整为 224×224 ,另外将数据集分成训练集、验证集和测试集,数据集信息如表 2 所示。在网络训练前设置使用 He 正态分布方法初始化网络权重参数^[16],使用随机梯度下降优化算法对损失函数

3 肺炎检测网络框架

实验环境如下:CPU Intel® Core i7-6800K 3.5 GHz,GPU Nvidia GTX1080Ti(11GB)×2,Ubuntu 16.04 位操作系统,使用以 TensorFlow 作为后端的 Keras 开源深度学习框架。

3.1 实验数据与预处理

实验使用 RSNA 胸部 X 光 (Chest X-ray, CXR) 图像数据集,其中包含了 5600 张具有肺炎症状的 X 光图像以及其对应真实肺炎病灶位置信息。每一张图像分辨率大小为 1024×1024 ,真实肺炎病灶位置信息由多名放射科领域专家或者医生手动标注,最后取各个标注的交集作为真实标签。

图像预处理过程如图 6 所示。由于胸部 X 光图像存在亮度低、对比度低等特点,为了更好检测到肺炎病灶部位,本文对胸部 X 光图像进行如下预处理:首先进行图像归一化处理;然后将图像转化成灰度图像,在灰度图像上使用限制对比度的自适应直方图均衡化 (CLAHE) 算法提高胸部 X 光图像的对比度^[14];最后将图像进行水平翻转以增强数据,防止训练时由于训练样本过少导致网络模型过拟合。实验使用的 CLAHE 算法相较于直方图均衡化 (HE) 算法在灰度变换的范围从全局缩小到某一局部^[15],因此有利于保留局部微弱病灶信息。预处理前后胸部 X 光图像灰度直方图分布如图 7 所示,由图 7 可知,经过胸部 X 光图像 CLAHE 算法处理后,图像灰度分布均匀,去除强噪声的同时对局部细节信息进行了增强。

进行优化,在参数设置上使用通用参数进行设置,其中学习率初始值为 0.001,动量设置为 0.9,正则化系数设置为 0.0005,批处理量 (Batch size) 大小设置与实验设备显存有关,为了充分利用显存,加快训练速度,因此将 Batch size 设置为 32,进行 100 轮迭代。在训练过程中,观察验证集的损失函数值的变化,当

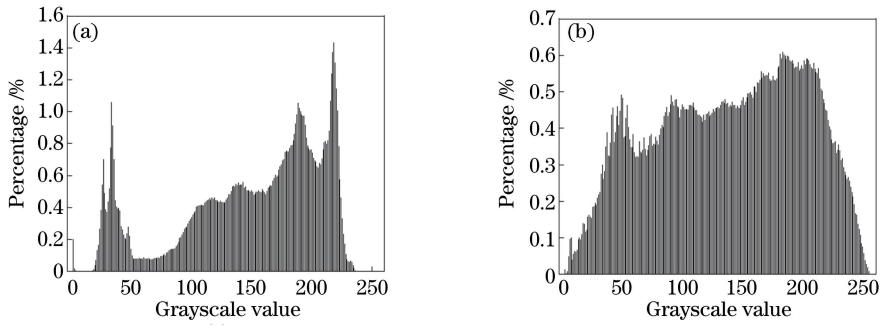


图7 胸部X光图像灰度直方图。(a)原图;(b)使用CLAHE后

Fig. 7 Chest X-ray image gray histogram. (a) Original image; (b) after using CLAHE method

指定训练次数内验证集的损失值不发生下降时,减小0.2学习率。通过缩小学习率使得模型跳出局部最优解,继续寻找全局最优解。

表2 数据集信息统计

Table 2 Statistics of data set information

Data set	Image	Size	Instance
Training	10260	224×224	16242
Validation	530	224×224	868
Testing	617	224×224	1000

3.3 性能评价指标

为了分析和检验网络的性能和检测结果,本文采用平均检测精度(AP)作为肺炎病灶检测结果的评价指标,AP指标被广泛应用于目标检测领域评估指标中^[17-19]。AP∈[0,1]表示模型精确率-召回率(P-R)曲线下的面积,P-R曲线以召回率作为横坐标,精确率作为纵坐标,反映了随着召回率的增加,模型精确率的变化情况。AP值越接近于1代表模型检测性能越好。因此,AP值和P-R曲线能够直观地反映出肺炎病灶检测网络的检测性能。其中,准确率和召回率定义为

$$V_{\text{precision}} = \frac{E_{\text{TP}}}{E_{\text{TP}} + E_{\text{FP}}}, \quad (8)$$

$$V_{\text{recall}} = \frac{E_{\text{TP}}}{E_{\text{TP}} + E_{\text{FN}}}, \quad (9)$$

式中: E_{TP} 代表真正例, E_{FP} 代表假正例, E_{FN} 代表假负例。准确率表示正确预测正样本占有所有预测正样本的比例,召回率表示正确预测正样本占有所有实际正样本的比例。AP值根据预测标注框与真实标注框面积之间的交并比(IoU)阈值进行评估。通常IoU阈值越高,说明评估模型的条件越严格,因此AP值会随之下降。

3.4 结果与分析

本文设计的肺炎病灶检测实验均在RSNA肺炎病灶数据集上进行训练和测试,对比实验检测结

果如表3、4所示,肺炎病灶检测结果的P-R曲线图如图8所示。

由表3检测结果可知,相较于HOG+SVM、SSD^[20]、YOLO^[21]、Faster R-CNN^[22]以及RetinaNet,本文所提的SE-RetinaNet在检测精度方面均优于其他算法。特别是基于SE-ResNet-50的SE-RetinaNet网络的检测精度为82.52%,与基于ResNet-101的RetinaNet网络相比精度提高了1.63%。由表3检测速度可知,SE-RetinaNet相比RetinaNet增加了参数量,但SE-RetinaNet并没有显著增加计算时间。实验表明,本文所提的SE-RetinaNet在肺炎检测方面性能优异,能够更加准确地检测出X光图像中肺炎病灶所在的位置。

为分析模型融合对肺炎病灶检测性能的影响,本文设计了多组模型融合分析实验。如表4所示,实验将RetinaNet50、RetinaNet101、SE-RetinaNet50、SE-RetinaNet101四个检测网络进行两两任意组合,并测试各个组合模型在肺炎病灶测试集上对应的AP值。通过对比表3和表4发现,模型融合的检测效果比任意子模型检测效果都有提高。其中SE-RetinaNet101与RetinaNet101的融合

表3 不同目标检测网络在肺炎病灶检测性能对比

Table 3 Performance comparison of different detection networks in the pneumonia detection

Network	Backbone	AP / % @0.25	AP / % @0.5	Speed / s
HOG+SVM	-	49.33	31.11	0.1123
SSD	VGG16	73.88	36.42	0.0312
YOLOv3	DarkNet	74.30	38.64	0.0238
Faster R-CNN	ResNet-50	79.12	45.91	0.0588
Faster R-CNN	ResNet-101	78.41	43.10	0.0833
RetinaNet	ResNet-50	78.73	48.87	0.0526
RetinaNet	ResNet-101	80.89	46.54	0.0769
SE-RetinaNet	SE-ResNet-50	82.52	49.96	0.0588
SE-RetinaNet	SE-ResNet-101	81.51	48.88	0.1000

表 4 不同组合模型在肺炎病灶检测性能对比

Table 4 Performance comparison of different combined models in the pneumonia detection

Network_1	Network_2	AP / %		Speed / s
		@0.25	@0.5	
RetinaNet50	RetinaNet101	88.20	57.21	0.1295
SE- RetinaNet50	SE- RetinaNet101	87.16	54.50	0.1588
RetinaNet50	SE-RetinaNet50	86.59	54.46	0.1114
Retina Net101	SE- RetinaNet101	89.08	57.64	0.1769

模型检测精度为 89.08%，与基于 SE-ResNet-50 的 SE-RetinaNet 网络相比，当 IoU 阈值为 0.25 时融合模型检测精度提高了 6.56%，另一方面当 IoU 阈值为 0.5 时模型融合检测精度提高了 7.68%。实验结果表明，本文所使用的融合模型可有效提高肺炎病灶的检测精度。

4 结 论

针对目前检测算法对肺炎病灶检测精度较低的问题，本文提出一种基于多尺度卷积神经网络的 X 光图像中肺炎病灶检测算法，目的在于提高胸部 X 光图像中对肺炎病灶位置的检测精度。该算法主要有两方面创新，一方面在 RetinaNet 网络的基础上加入特征通道注意力模块——SE 模块，另一方面通过统计分析并使用 K-means 聚类算法设计出一系列合适的候选框尺寸以及对应宽高比。通过两方面改进使得检测网络相较于其他常见检测算法对肺炎病灶的检测性能得到明显提高。实验结果表明，本文所提 SE-RetinaNet 检测算法在增加较小参数量的情况下，能够准确地检测出胸部 X 光图像中的肺炎病灶位置信息，为医生提供强有力的诊断参考。

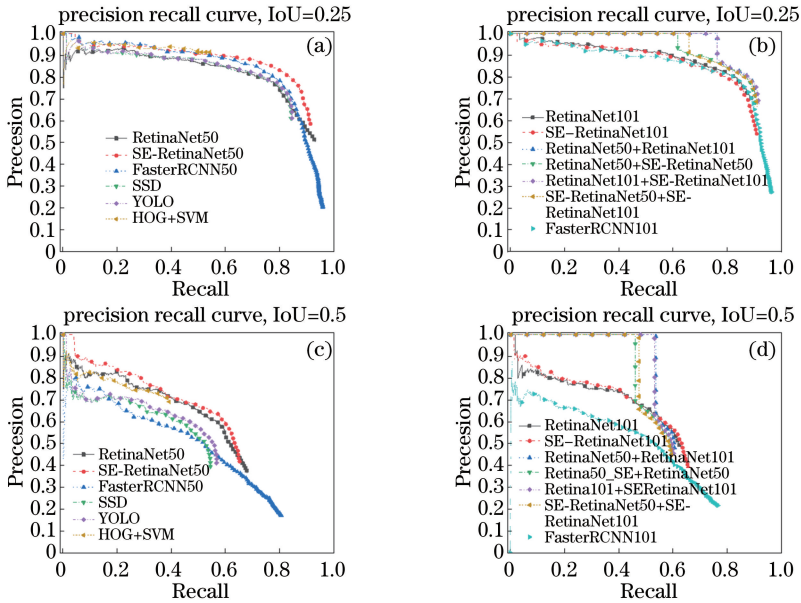


图 8 不同检测算法在肺炎数据集上的 P-R 曲线。(a)(b) IoU 阈值为 0.25 情况下 P-R 曲线；(c)(d) IoU 阈值为 0.5 情况下 P-R 曲线

Fig. 8 P-R curves of different algorithms on the pneumonia dataset. (a)(b) P-R curves with IoU threshold of 0.25. (c)(d) P-R curves with IoU threshold of 0.5

参 考 文 献

[1] McLuckie A. Respiratory disease and its management [M]. Berlin: Springer Science & Business Media, 2009.

[2] Cherry D K, Woodwell D A, Rechsteiner E A. National ambulatory medical care survey: 2005 summary[J]. Advance Data, 2007, 328(387): 1.

[3] Franquet T. Imaging of community-acquired pneumonia[J]. Journal of Thoracic Imaging, 2018,

33(5): 282-294.

[4] Baker J A, Rosen E L, Lo J Y, et al. Computer-aided detection (CAD) in screening mammography: sensitivity of commercial CAD systems for detecting architectural distortion [J]. American Journal of Roentgenology, 2003, 181(4): 1083-1088.

[5] Wang X, Peng Y, Lu L, et al. Chest X-ray8: hospital-scale chest X-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases [C] // Proceedings of the

- IEEE Conference on Computer Vision and Pattern Recognition, 2017: 2097-2106.
- [6] Li Z, Wang C, Han M, et al. Thoracic disease identification and localization with limited supervision [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 8290-8299.
- [7] Ma Y, Zhou Q, Chen X, et al. Multi-attention network for thoracic disease classification and localization [C] // ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2019: 1378-1382.
- [8] Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection[C] // Proceedings of the IEEE International Conference on Computer Vision, 2017: 2980-2988.
- [9] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.
- [10] Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 2117-2125.
- [11] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015: 3431-3440.
- [12] Hu J, Shen L, Sun G. Squeeze-and-excitation networks[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 7132-7141.
- [13] Redmon J, Farhadi A. YOLO9000: better, faster, stronger[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 7263-7271.
- [14] Qin C, Yao D, Shi Y, et al. Computer-aided detection in chest radiography based on artificial intelligence: a survey [J]. Biomedical Engineering Online, 2018, 17(1): 113.
- [15] Xiang J K, Wu M, Wang Z, et al. Application of image enhancement in X-ray photography of cultural relics[J]. Laser & Optoelectronics Progress, 2019, 56(6): 063402.
- 相建凯, 吴萌, 王展, 等. 图像增强在文物 X 光图像中的应用[J]. 激光与光电子学进展, 2019, 56(6): 063402.
- [16] He K, Zhang X, Ren S, et al. Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification[C] // Proceedings of the IEEE International Conference on Computer Vision, 2015: 1026-1034.
- [17] Guo J X, Liu L B, Xu F et al. Airport scene aircraft detection method based on YOLO v3 [J]. Laser & Optoelectronics Progress, 2019, 56(19): 191003.
- 郭进祥, 刘立波, 徐峰, 等. 基于 YOLO v3 的机场场面飞机检测方法[J]. 激光与光电子学进展, 2019, 56(19): 191003.
- [18] Liu W J, Gao M Y, Que H C, et al. Light-weight multi-object detection network based on inverted residual structure [J]. Laser & Optoelectronics Progress, 2019, 56(22): 221003.
- 刘万军, 高明月, 曲海成, 等. 一种基于反残差结构的轻量级多目标检测网络[J]. 激光与光电子学进展, 2019, 56(22): 221003.
- [19] Chen Y T, Li Y Y, Yao T T. A novel method for ship detection in remote sensing image under complex sea conditions [J]. Laser & Optoelectronics Progress, 2019, 56(18): 181007.
- 陈彦彤, 李雨阳, 姚婷婷. 复杂海况下遥感图像舰船目标检测方法研究[J]. 激光与光电子学进展, 2019, 56(18): 181007.
- [20] Liu W, Anguelov D, Erhan D, et al. SSD: single shot multibox detector[C] // European Conference on Computer Vision, Springer, Cham, 2016: 21-37.
- [21] Redmon J, Farhadi A. Yolo v3: an incremental improvement [J]. arXiv preprint arXiv: 1804.02767, 2018.
- [22] Ren S, He K, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2015, 39(6): 1137-1149.