用于肺结节分类的3D稠密倒残差潜在均衡 注意力网络

王文举¹, 叶 芳¹, 殷淑雅¹, 李嘉琪¹, 朱 琳², 于 红²

¹上海理工大学出版学院,上海 ²上海市胸科医院放射科,上海

收稿日期: 2024年12月20日; 录用日期: 2025年1月13日; 发布日期: 2025年1月17日

摘要

目前基于深度学习的高性能的肺结节良恶性分类网络通常由于结构复杂,伴随着大量的参数量和计算量 需求。为此本文提出了3D稠密倒残差潜在均衡注意力网络(3D Dense Inverted Residuals Latent Equilibrium Attention Network,简称3D DIRLEAN),以实现在进一步提高肺结节分类精度的同时减少模型 的参数量节省实际应用内存。所提出的3D DIRLEAN网络框架仅需4个阶段,主要由3D稠密倒残差LEA模 块(3D DIRL)和潜在均衡注意力模块(LEA)组成。我们受生物神经元的响应机制启发构建了LEA。它通过 我们设计的潜在均衡能量函数直接估计三维权重,获取更具区分性的特征;运算过程0参数并且可以有效 减少网络层之间的响应滞后,加速模型收敛。此外,我们构建的3D DIRL结合了倒残差和稠密连接的优 势,将其同时用于特征处理。它使用稠密连接获取不同网络层的特征信息,结合倒残差实现特征重用, 最大限度利用特征信息加强网络的判断能力;同时,标准卷积的分解为网络降低了参数量和计算量。这 两个模块共同作用使3D DIRLEAN能够以最少的参数量和计算量实现最高的分类精度。所提出的方法应 用于LUNA16数据集。实验结果表明,3D DIRLEAN的分类精度达到了93.33%,并且参数量和计算量分 别降低到了2.85M、0.63G,总体上优于先进的同类方法。

关键词

潜在均衡注意力,3D稠密倒残差LEA模块,3D稠密倒残差潜在均衡注意力网络,肺结节良恶性分类

3D Dense Inverted Residuals Latent Equilibrium Attention Network for Pulmonary Nodules Classification

Wenju Wang¹, Fang Ye¹, Shuya Yin¹, Jiaqi Li¹, Lin Zhu², Hong Yu²

¹College of Communication and Art Design, University of Shanghai for Science and Technology, Shanghai ²Department of Radiology, Shanghai Chest Hospital, Shanghai

Received: Dec. 20th, 2024; accepted: Jan. 13th, 2025; published: Jan. 17th, 2025

Abstract

Presently, the development of advanced classification networks for accurate differentiation between benign and malignant pulmonary nodules, utilizing deep learning techniques, often entails a multitude of parameters and substantial computational demands. Consequently, this research proposes the innovative 3D Dense Inverted Residuals Latent Equilibrium Attention Network (3D DIR-LEAN) to improve the accuracy of pulmonary nodule classification and to optimize the number of parameters in the model, thereby saving the memory of practical applications. 3D DIRLEAN mainly includes the 3D Dense Inverted Residual LEA Module (3D DIRL) and the Latent Equilibrium Attention (LEA) mechanism. The LEA module directly estimates three-dimensional weights through a bespoke latent equilibrium energy function, thereby capturing more distinct features. The operation process employs zero parameters, effectively curtailing response lags between network layers and hastening model convergence. Furthermore, the constructed 3D DIRL component seamlessly integrates the advantages of inverted residuals and dense connections for concurrent feature processing. Capitalizing on dense connections, it acquires feature information across diverse network layers, complemented by inverted residuals to enable feature recycling, thereby optimizing the utilization of feature data. Simultaneously, the deconstruction of standard convolutions serves to curtail the model's parameter count and computational overhead. The experimental accuracy of 3D DIRLEAN on LUNA16 reached 93.33%, and the parameters and FLOPs were reduced to 2.85M and 0.63G respectively. Clearly, LEA and 3D DIRL work together to enable the 3D DIRLEAN framework to achieve the highest classification accuracy over similar advanced technologies, while minimizing parameter counts and floating-point operations (FLOPs).

Keywords

Latent Equilibrium Attention, 3D Dense Inverted Residual LEA Module, 3D Dense Inverted Residuals Latent Equilibrium Attention Network, Pulmonary Nodule Classification

Copyright © 2025 by author(s) and Hans Publishers Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0). <u>http://creativecommons.org/licenses/by/4.0/</u>

CC O Open Access

1. 引言

肺癌是目前世界上发病率最高的癌症之一[1]。肺癌一旦处于晚期,治愈的可能性很小。但如果在早期及时诊断,切除病灶,大多数的患者可以治愈。医生根据 CT 扫描肺结节的密度、形状、有无钙化等来判断肺结节的良恶性。基于医生的人工方法对于较大的结节的诊断准确性很高,但对于较小的结节的诊断结果并不理想。因此,实现肺小结节早期的自动精确分类[2]具有重要意义。

肺结节的分类方法分为传统的机器学习方法和深度学习方法两大类[3]。

(1) 传统机器学习方法

传统的机器学习方法由人工对模型的特征进行编码,没有特定的硬件要求,模型训练时间较短。如 Yang Li 等人[4]提出的成本敏感的多核学习支持向量机(CS-MKL-SVM)算法,解决了正负样本不平衡的 问题,实现了数据集较少情况下的肺结节分类。Zhehao He 等人[5]使用微软的 Custom Vision AutoML (Automated Machine Learning)系统,通过学习肺结节的薄层 CT 影像数据和相应的病理诊断来训练模型,也

实现了较少数据集的肺结节分类。但传统的机器学习方法在进行特征提取时依赖人工预先设定的数据特征,缺乏自动提取特征的能力,并且泛化能力较弱。

(2) 深度学习方法

深度学习模型相比于传统的机器学习模型具有更加复杂的网络结构,在模型训练时自动学习特征,可以处理大规模的数据集且分类精度一般高于传统的机器学习方法。深度学习技术已经在行为识别[6]、自动驾驶[7]、语音识别[8]等其他领域取得了不错的成果,近些年也被引入到肺小结节分类的研究中。

在肺结节分类中,一般将高精度的肺小结节的良恶性分类作为算法的首要目标。Hongtao Xie 等人[9] 提出了一种基于 2D CNN 的计算机辅助肺结节自动检测方法,实现了对潜伏性肺结节的准确检测,但只能对 2D 肺结节图像进行检测和分类。Ruisheng Su 等人[10]介绍了一种 2.75D CNN 策略用于胸部 CT 肺结节分类。即使用螺旋扫描技术将 3D 图像表示为 2D 特征,再通过 2D CNN 实现分类。但该方法的学习能力和分类精度还有待提高。Yi Wang 等人[11]将一种采用了多路径特征提取方案和多尺度卷积层的新的卷积神经网络(ConvNet)架构用于肺结节的分类,提高了分类精度。但该方法需要人工裁剪结节区域。Enoumayri Elhoussaine 等人[12]提出了基于 3D CNN、3D AlexNet [13]模型和 3D ResNet [14]模型的集成学习方法,以及 Ge Zhang 等人[15]开发的一种 3D DenseNet 架构。两种方法对肺结节的分类精度分别达到了 84.66%和 92.4%,但都需要涉及到部分人工操作。Ioannis D. Apostolopoulos 等人[16]则将迁移学习技术用于 PET/CT 成像中肺结节的自动分类,取得了较高的精度。然而该技术调查和利用的预训练网络种类较少,精度提升空间有限。

提高正确分类肺结节的精度的同时,避免出现假阳性检测判定是必须面对的重要问题。Lingma Sun 等人[17]为了减少肺结节检测中的假阳性,提出了一种注意力嵌入式互补流卷积神经网络,提高了网络在 减少误报方面的性能。然而该网络复杂性较高,模型的训练时间较长,并且在低假正类(FP)率下的灵敏度 较低。Dandan Zhao 等人[18]提出了一种多尺度卷积神经网络,在减少误报的同时,还大大减少了网络模 型的训练时间。但该方法对假阳性的判定精度不够精准。Zhan Wu 等人[19]构建了一种基于深度卷积神经 网络(DCNN)的多维结节检测网络(MD-NDNet),提高了对假阳性结节的判定精度。但该框架模型尚不能 对良性结节和恶性结节进行进一步分类。而 Shiwei Li 等人[20]介绍的基于改进的迁移学习策略的肺结节 分类方法实现了对肺结节的进一步分类。相较于传统的将肺癌图像分为正常和癌性两类,该算法可以将 肺癌的 CT 图像分为四类:肺腺癌、肺鳞状细胞癌、转移性肺癌和正常类型的肺癌,但迁移学习方法存在 模型参数不易收敛的问题。上述方法都是对同一大小尺度的肺结节进行检测,无法对大小尺度不同的肺 结节进行检测。

对肺结节进行良恶性分类的过程中,还需考虑算法网络模型对大小不同尺度的肺小结节的适用性。 Wangxia Zuo 等人[21]为了减少肺结节检测中的误报,使用了一种包含嵌入式多分支的 3D 卷积神经网络。 该框架模型可以对不同大小和形状的肺结节进行有效的识别。但由于不同大小结节的训练数据量不同, 产生的预测结果的置信度也不同。Haiying Yuan 等人[22]提出了一种基于 ResNet [14]的可变形卷积网络 用于减少假阳性结节,可以对大小、形状有较大变化的肺结节进行更为精准的特征提取。Haiying Yuan 等 人[23]设计了一种 3-D 计算机辅助系统用于肺结节的分类检测。该方法包含 3-D 上下文引导模块、应用 通道注意机制和多任务学习的多分支分类网络,可以有效检测多尺度不同形状的肺结节,但不能提供其 他的肺结节信息。Yu Han 等人[24]介绍了一个以改进的 U-Net 和 3D Region Proposal Network (RPN)为骨 干网络的计算机辅助系统。该方法系统可以提供检测到的结节的详细指标,并且可以成功地检测出体检 获得的 CT 图像中小结节。尽管如此,这类方法对于一些微小结节仍然难以被检测。

为了进一步提高对微小结节的检测分类精度,一些研究方法在判断肺结节良恶性时考虑更多的相关因素,例如结节本身的一些肿瘤特征。Su Chen [25]提出了由卷积神经网络和长短期记忆递归神经网络相

融合得到的 3D U-net 模型。该模型在预测癌症的过程中考虑了结节的密度,但并没有考虑其他肿瘤特征。 Lihua Shen 等人[26]引入密集卷积网络从九个不同方向的二维视图中提取更多的输入多尺度肿瘤特征信息,优化了特征提取和利用。Xianfang Hu 等人[27]开发了一种计算机辅助诊断(CADx)方案用于对毛玻璃结节进行良恶性分类。该方法融合了深度学习和放射组学特征,可以很好的解码部分肺肿瘤周围特征以及内部特征。但该方法使用的数据集有限且不平衡,可能会导致欠拟合问题。为解决该问题,Seba Susan等人[28]设计了类别加权卷积自动编码器(CW-CAE),能够有效实现在不平衡的数据集中检测肺结节。但该方法目前仅实现了对肺结节的检测,还无法判断检测到的结节癌变程度。这点不足在 José Lucas Leite Calheiros 等人[29]提出的 SVM-RBF 方法中得到了解决。该方法评估了在实性肺结节表征中包括结节周围区域的影响以及对恶性肿瘤可能性的进一步分类,但也排除了一些可能是恶性肿瘤的额外指标的重要结构。

上述基于多因素的方法确实达到了较高的精度。但这些方法实现过程复杂且都以黑盒的形式工作, 对分类推理过程的可解释性不足。Xiangbo Zhang 等人[30]提出了一种用于肺结节分类的新型属性自引导 图卷积 V 形网络(AS-GCVN)。该框架模型引入了 Attribute Self-guided Feature Enhancement (ASFE)模块, 提高了图卷积神经网络的语义推理能力,但在医学图像计算的可解释方面还不够完善。Shiwen Shen 等人 [31]为了预测在计算机断层扫描上检测到的不确定结节是否为恶性,构建了一种新的可解释的深度层次 语义卷积神经网络(HSCNN)。该框架模型实现了对结节恶性程度的预测的同时,还有助于诊断预测的解 释。然而该框架模型中存在较多的低级子任务,使得模型参数量大难以收敛。Hanliang Jiang 等人[32]使 用神经结构搜索(NAS)自动搜索三维网络和卷积块注意模块(CBAM)建立了一个部分可解释的分类模型, 提高了分类精度。然而该模型共有 6 个阶段,单个阶段的网络深度最高可达到 8 层。

综上所述,我们发现当前对于肺小结节的良恶性分类算法,其精度有待于进一步的提高。同时其网络结构复杂、训练的参数量较多,这进一步导致网络模型占用大量存储空间。为此我们提出了 3D DIR-LEAN 网络结构,该算法的主要贡献有:

(1) 我们提出了 3D 稠密倒残差潜在均衡注意力网络(3D DIRLEAN)的框架用于高精度的肺小结节良 恶性分类。该网络结构简单,仅由 4 个阶段构成,其中所用的关键技术是 3D 稠密倒残差 LEA 模块(3D DIRL)和潜在均衡注意力(LEA)。我们将 3D DIRLEAN 应用在 LUNA16 数据集上,与已有的先进算法相 比获得了最先进的分类精度,同时还具有最小的模型参数量和计算量。

(2) 3D DIRLEAN 在提出潜在均衡能量函数的基础上构建了潜在均衡注意力机制(LEA)。潜在均衡能 量函数反映了生物神经元的响应机制,用来减少网络层之间的响应滞后,加速模型收敛。构建的 LEA 对 生物注意力机制进行仿生从而细化特征,在 0 参数的条件下直接估计三维权重得到更具区分性的特征, 从而解决了其他注意力机制参数量大的问题,整体提高网络的判定准确性。

(3) 3D DIRLEAN 设计了以 3D 稠密 LEA 模块为核心的 3D DIRL。该模块将残差和密集连接同时用 于特征处理。3D DIRL 把部分标准卷积分解,解决了传统卷积堆叠导致的参数量大的问题,有效降低了 模型存储空间。同时,它实现了特征重用,获取并结合运用不同网络层的丰富特征信息,最大化最有用 的特征。

本文的其余部分安排如下。第 2 节介绍了 3D DIRLEAN 网络框架。第 3 节介绍了数据集和实验环境,并对本文提出的模型进行了详细的实验和结果分析。第 4 部分总结了本文并提出了未来的工作。

2. 本文方法

本文提出了 3D 稠密倒残差潜在均衡注意力网络(3D DIRLEAN)框架,模型总体框架如图 1 所示。在 该网络模型中将肺小结节 CT 图像作为输入,先后经历 4 个阶段模块的处理即可实现对肺小结节的良恶 性的预测判定。阶段 1-3 中分别有 1、N = 2、L = 4 个 3D DIRL 与一个 LEA。阶段 4 中则是由一个 3D DIRL、最大池化层和全连接层构成。其中阶段 1 使用的是 3D DIRL-CE,而其他阶段使用的是 3D DIRL-CNE (二者的区别见 2.2 节)。



Figure 1. 3D dense inverted residuals latent equilibrium attention network framework 图 1. 3D 稠密倒残差潜在均衡注意力网络框架

由图 1 可见,所提 3D DIRLEAN 网络框架结构中起决定性作用的是 LEA 和 3D DIRL。我们将在本 节当中的 2.1、2.2 分别给予详细介绍。

2.1. 潜在均衡注意力模块



近年来,注意力机制广泛应用于图像处理、语音识别等领域。它通过为输入数据的不同部分赋予不同的权重,从而抽取关键信息,使模型做出更准确的判断。现存的注意力机制,如 SE [33]、CBAM [34]、 SRM [35]等,都是把空间和通道分开考虑或只考虑其中之一来获取特征,生成一维或二维权重。这些注意力机制通常都运算复杂且需要大量参数。这会导致神经网络各层之间存在一定的响应滞后,使整个神经网络的推理速度随着深度的增加而降低,影响网络的学习。生物神经元中也存在类似的问题。生物神经元可以使其相对于膜电位的输出相位提前,从而避免响应滞后。本文受此启发,基于潜在均衡理论[36]构建了潜在均衡能量函数并以此为基础提出了一种潜在均衡注意力(Latent Equilibrium Attention,简称LEA)(具体结构见图 2)来提高网络层之间的响应速度,加快模型收敛。 LEA 同时考虑空间和通道维度,直接从当前神经元中推断出三维权重,使网络学习到的神经元更有 区分性。除此之外,该模块不需要任何额外的参数。相比其他运算复杂且需要大量参数的注意力机制, LEA 不管是在推理速度还是性能方面都有更好的表现。LEA 的主体能量函数定义如下:

$$Y = \left(X_l - W \cdot \sigma_{\text{sigmoid}}\left(X_l\right) - B\right)^2 + \frac{1}{2}W$$
(1)

这里 X_l 是上一层网络输入的特征, $X_l \in \mathbb{R}^{C \times H \times W \times D}$ 。 $\sigma_{sigmoid}$ 是 sigmoid 激活函数。W 和 B 是分别是权重矩阵和偏置矩阵。为了使不同的神经元之间具有更强的线性可分性,我们定义了公式(2)来计算权重。同时用公式(3)计算得到偏置以控制神经元的激活状态并且加速网络的拟合。

$$W = \frac{2(X_{l} - Mean)}{(X_{l} - Mean)^{2} + 2Variance + 2\lambda}$$
(2)

$$B = \frac{1}{2} \left(X_l + Mean \right) \cdot W \tag{3}$$

以通道为单位,输入特征 X_i 的方差 $Variance = \frac{1}{N} \sum_{i=1}^{N} (x_i - Mean)^2$ 和均值 $Mean = \frac{1}{N} \sum_{i=1}^{N} x_i$ 被计算。其中 $N = H \times W \times D$ 是一个通道上的神经元数量, x_i ($i = 1, 2, \dots, N$)是输入特征 X_i 中的单个神经元。

将特征 X_i 通过能量函数处理后,得到与特征 X_i 尺寸相同的 Y。LEA 再将 Y 通过 Sigmoid 细化处理 直接用于估计三维权重 Y',见公式(4)所示。

$$Y' = \operatorname{sigmoid}(Y) \tag{4}$$

最后将原始的 X₁和得到的三维权重 Y'进行相乘操作得到更具区分性的特征 X₁₊₁,见公式(5)。

$$X_{l+1} = X_l \odot Y' \tag{5}$$



2.2. 3D 稠密倒残差 LEA 模块

Figure 3. Two 3D Dense Inverted Residual LEA Module structures 图 3. 2 种 3D 稠密倒残差 LEA 模块结构 本文提出了结合 LEA 的 3D 稠密倒残差模块用于阶段 1~4,如图 3 所示。该残差模块分为 3D DIRL-Channel Equal (用于阶段 1)和 3D DIRL-Channel Not Equal (用于阶段 2~4)两种结构。如图 3 所示,当特征 输入通道与输出通道相等时,我们直接对原始输入特征进行跳跃连接(3D DIRL-Channel Equal)。若不相 等,则会经过一个额外的卷积处理(3D DIRL-Channel Not Equal)。

3DDIRL 的两种结构的核心都是 3D 稠密 LEA 模块。该模块网络先经过 1×1 的逐点卷积扩张肺小结 节特征图的通道,丰富特征数量;再将标准卷积分解为深度卷积核进行逐点卷积,显著地降低了参数量 和计算量;在深度可分离卷积后,LEA 被加入来提高网络层之间的响应速度;最后再经过一个 1×1 卷 积压缩特征维度。我们还在模块结构中引入了稠密连接,减轻了梯度消失问题,同时使得特征和梯度的 传递更加有效。

3D 稠密倒残差 LEA 模块的实现过程如下:

(1) 输入特征并初始化参数:上一层网络输出的特征 X_m作为本模块的输入,并将 X_m备份。输入通道数 *inp* 可依据 X_m进行获取。隐藏层通道数 *hidden* 根据公式(6)计算得出。

$$hidden = t \times inp \tag{6}$$

其中 t 为扩展因子,我们通过实验确定为 4。

(2) 扩张特征图通道(升维):特征维度经过 ReLU 和 1×1 的 3D 卷积从 *inp* 维增加到 *hidden* 维。如公式(7)所示。

$$X_{1} = Conv3D_{1\times 1}\left(\sigma_{relu}\left(X_{m}\right)\right) \tag{7}$$

其中 σ_{relu} 表示 ReLU 激活函数, Conv3D_{1×1}表示输入通道为 inp 输出通道为 hidden 的 1×1卷积, X₁为输出特征, 维度是 hidden。

(3) 特征拼接并规范特征通道数: *X_m*和 *X*₁在通道维度上进行特征拼接(当 *X_m*和 *X*₁尺寸不匹配时, *X_m*需要额外通过 1 × 1 的卷积操作,如公式(8)所示,使原始输入特征 *X_m*可以直接传递到深度卷积层, 再通过逐点 1 × 1 卷积规范特征通道数。该过程可表征为公式(9)。

$$X_m = Conv3D_{1\times 1}\left(X_m\right) \tag{8}$$

这里的 Conv3D_{1×1}要求输入通道和输出通道都为 inp,步长为 2 的 1 × 1 卷积。

$$X_{1}' = PointConv_{1\times 1}\left(\sigma_{concat}\left(X_{m}, X_{1}\right)\right)$$
(9)

其中 *PointConv*_{1×1} 是输入通道为 *inp* + *hidden* 其输出通道为 *hidden* 的 1 × 1 逐点卷积, σ_{concat} 表示 $X_m = X_1$ 的合并。

(4) 特征提取: 3×3的 Dwise3D 用来提取特征, 见公式(10)。

$$X_2 = Dwise3D_{3\times3}\left(\sigma_{relu}\left(X_1'\right)\right) \tag{10}$$

其中 $Dwise3D_{3x3}$ 表示输入通道和输出通道维数都为 hidden 的 3×3 卷积, X_2 为输出特征其维度是 hidden。

(5) 特征拼接并规范特征通道数: *X_m*、*X*₁[']和*X*₂在通道维度上进行特征拼接,同样经过逐点 1×1卷 积规范特征通道数见公式(11)。

$$X_{2}' = PointConv_{1\times 1} \left(\sigma_{concat} \left(X_{m}, X_{1}', X_{2} \right) \right)$$

$$\tag{11}$$

(6) 注意力模块估计三维权重: LEA 直接对步骤(5)中所得到的 X₂ 估计三维权重,得到更具区分性的特征,如公式(12)所示。

$$X_3 = LEA(X_2') \tag{12}$$

这里 LEA 代表 Latent Equilibrium Attention,具体细节见 2.1。X₃为输出特征,维度是 hidden。

(7) 特征拼接并规范特征通道数: *X_m*、*X'*₁、*X'*₂和*X*₃在通道维度上进行特征拼接,并经过逐点1×1卷积规范特征通道数,如公式(13)所示。

$$X'_{3} = PointConv_{1\times 1} \left(\sigma_{concat} \left(X_{m}, X'_{1}, X'_{2}, X_{3} \right) \right)$$
(13)

(8) 压缩特征图(降维): X₃ 通过 1×1 的 Conv3D 实现降维,维度从 *hidden* 降低到 *oup* 维,得到处理 后的特征 X₄。该过程可以表征为公式(14)。

$$X_4 = Conv3D_{1\times 1}\left(X_3'\right) \tag{14}$$

经过步骤(1)~(8), 3D 稠密倒残差 LEA 模块基本被构建,但还需进行最后的合并权重。

当 X_m 和 X_4 尺寸或通道数不匹配时,需要额外通过如公式(15)所示的 1×1 卷积和 ReLU 操作得到 $X_{mbackup}$ 。 $X_{mbackup}$ 再和步骤(8)处理得到的特征 X_4 进行相加操作,得到输出特征 X_{m+1} ,见公式(16)。

$$X_{mbackup} = \sigma_{relu} \left(Conv3D_{1\times 1} \left(X_m \right) \right) \tag{15}$$

$$X_{m+1} = X_{mbackup} + X_4 \tag{16}$$

当 X_m 和 X_4 尺寸和通道数都匹配时,则直接相加得到输出特征 X_{m+1} ,见公式(17)。

$$X_{m+1} = X_m + X_4 \tag{17}$$

3. 实验

3.1. 环境配置

本实验在 Nvidia GeForce RXT 3080 显卡、32GB 内存的硬件环境下进行。操作系统为 Ubuntu 22.04。 使用的软件为 PyTorch1.13 和 CUDA11.7。实验设置的初始学习率为 0.0002, 训练和测试周期都设置为 500 个 epochs。衰减周期设置为 70, 即学习率在最后的 70 个 epochs 会进行如公式(18)所示的衰减。批量 归一化被插入在网络层中来减少过拟合,并且使用 A-Softmax 损失函数和动量参数 $\beta_1 = 0.5$ 、 $\beta_2 = 0.999$ 的 Adam 优化器来进一步优化网络。

$$lr = lr - (lr/decaynum) \tag{18}$$

3.2. 数据集

我们在实验中使用 LUNA16 数据集进行训练和测试。LUNA16 数据集是最大公用肺结节数据集 LIDC-IDRI 的子集,是 LIDC-IDRI 中切片厚度 <= 3 mm 且肺结节 >= 3 mm 的部分。共包括 888 个低剂 量肺部 CT 影像(mhd 格式)数据。原始图像为三维图像,由不同数量的二维图像组成,是一系列胸腔的多 个轴向切片。

为提高分类预测精度,训练数据都进行了相应的数据增强处理。即在训练过程中,先将尺寸大小为 32×32×32 的原始肺小结节 CT 图像填充为 36×36×36;接着使用水平翻转、y 轴翻转和 z 轴翻转进行 数据增强;最后再将增强后的图像随机裁剪为 32×32×32 的大小。

3.3. 评价指标

实验使用精确度、查全率、查准率和 F1-score 来评估网络模型的肺结节分类性能。 精确率是在所有预测中,正确预测结果占的比例:

$$Accuracy = \frac{TN + TP}{TN + FN + TP + FP}$$
(19)

查全率是在所有阳性事件中,正确预测结果占的比例:

$$\operatorname{Recall} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}$$
(20)

查准率是在全部阳性预测中,正确预测结果占的比例:

$$Precision = \frac{TP}{TP + FP}$$
(21)

F1-score 是查全率和查准率的调和均值:

$$F_1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$
(22)

这些标准的值越大表示模型性能越好。其中 TP、FP、TN、FN 分别表示真阳性、假阳性、真阴性、 假阴性。

除此之外,文中还使用 Params 和 FLOPs 来衡量模型的参数量和计算量。模型的参数量越多,意味 着模型需要更多的计算资源和数据来训练和推断。FLOPs 指运行一次网络模型需要进行浮点运算的次数。 如果模型的计算量过大,则会导致推断时间过长,不适于实际应用。

3.4. 实验结果与分析

3.4.1. 与先进方法的性能比较

Table	e 1. Po	erformance	e comparison	of classifica	tion algorithr	ns based o	on LUNA16	dataset
表 1.	基于	LUNA16	数据集的分	类算法的性	能比较			

Method	Accuracy (%)	Recall (%)	Precision (%)	F1-score
Multi-crop CNN [37]	87.14			
Nodule-level 2D CNN [38]	87.30	88.50		87.23
Vanilla 3D CNN [38]	87.40	<u>89.40</u>	_	87.25
DeepLung [39]	90.44	81.42	_	_
AE-DPN [40]	90.24	92.04		<u>90.45</u>
NASLung [32]	89.56	76.19	<u>94.11</u>	84.21
3D DIRLEAN (ours)	93.33	87.5	100	93.33

(每列中的最佳结果和次佳结果分别以粗体和下划线格式显示。)

为了验证本文提出的方法在分类精度上的优越性,我们将 3D DIRLEAN 与最先进的肺小结节分类方 法在 LUNA16 数据集上进行了比较。这些方法包括 Multi-crop CNN [37], Nodule-level 2D CNN [38], Vanilla 3D CNN [38], DeepLung [39], AE-DPN [40]和 NASLung [32]。从表 1 中可以看出,我们的 3D DIRLEAN 的 Accuracy、Recall、Precision 和 F1 分数分别为 93.33%、87.5%、100%和 93.33%,在对比的 所有方法中,取得了最高的分类精度、Precision 和 F1 分数。精度第二高的是 DeepLung,为 90.44%,比 我们的 3D DIRLEAN 低 2.89%。DeepLung 网络仅仅通过堆叠的卷积层来提取特征,而本文使用了 LEA 注意力机制使网络更加关注结节从而学习更精细的特征,使网络模型做出更准确的判断。除此之外, DeepLung 采取双路径连接,将部分特征图进行密集连接,其余用于残差学习。但这种方法并不能完全将 二者的优势相结合。然而,我们提出的 3D DIRL 将残差和密集连接同时用于特征处理,在减少网络中梯 度消失的同时实现特征重用获取特征图上最丰富的特征信息从而有助提高了分类精度。

3.4.2. 算法参数量与计算量比较

Method	Params (M)	FLOPs (G)
DeepLung [39]	141.57	_
AE-DPN [40]	678.69	_
NASLung [32]	<u>3.05</u>	<u>1.08</u>
3D DIRLEAN (ours)	2.85	0.63

 Table 2. Comparison between Params and FLOPs of the network model

 表 2. 网络模型参数量和计算量比较

(每列中的最佳结果和次佳结果分别以粗体和下划线格式显示。)

为了验证本文提出的方法可以在提高分类精度的同时有效降低参数量和计算量,我们与 DeepLung [39],AE-DPN [40]和 NASLung [32] 3 种高精度的网络进行了参数量和计算量的对比。如表 2 所示,我们 的 3D DIRLEAN 所需的参数量最少,为 2.85 M,约是 DeepLung 的 1/50。NASLung 的参数量仅比 3D DIRLEAN 高 0.2 M,但 3D DIRLEAN 的分类精度比 NASLung 高 3.77%,计算量也仅是 NASLung 的 1/60。 这表明我们的模型实现了出色的性能/成本权衡。3D DIRLEAN 的整体网络架构只有简单的 4 个阶段,尽量避免了参数信息的冗余。此外,其中 3D DIRL 使用倒残差结构将标准卷积分解,降低了模型的参数量和计算量。LEA 的整个模块中没有复杂的卷积运算,只通过能量函数进行简单的算数运算。并且它的特征推理过程无需任何参数,大大减少了实际应用中模型所需的内存和计算量。

3.4.3. 迭代推理速度性能比较



Figure 4. Comparison of accuracy and inference speed of the network model 图 4. 网络模型的准确性和推理速度对比

为了验证本文提出的方法在提高推理和收敛速度上的优势。我们将 3D DIRLEAN 和 NASLung [32] 前 300 次迭代的推理速度以及取得的分类精度绘制在图 4 中并进行了比较。公平起见,我们将开源的 NASLung 算法在同等硬件和软件实验环境中进行了复现。由图 4 可见,我们的 3D DIRLEAN 模型能够

在平均单次迭代推理速度 0.63 sec 的条件下快速收敛,达到 90%的精度。而 NASLung 的单次迭代推理速度约 1 sec,是 3D DIRLEAN 的 1.5 倍;同时收敛速度缓慢,取得的最高精度仅为 73%。这验证了我们的 网络在提高网络推理和收敛速度上的优越性。NASLung 中使用的是 CBAM 注意力机制。CBAM 将通道 注意力和空间注意力结合起来共同用于特征推理。而我们的网络中使用的 LEA 直接通过能量函数估计特征信息,因此推理过程要快于 CBAM。并且 LEA 的主体能量函数受潜在均衡理论启发,可以缓解网络层 之间的响应滞后。我们还将 LEA 插入在了 3DDIRL 中,以最大程度提高了网络层之间的响应速度,加快 模型收敛。

3.4.4. 消融实验

Table	3. Comparison of different frame structures
表 3.	不同框架结构的比较

	Accuracy (%)	Recall (%)	Precision (%)	F1-score
CNN (base)	90	87.5	93.33	90.32
CNN + LEA	91.11	83.33	100	90.90
CNN + 3D DIRL	92.22	91.66	93.61	92.63
3D DIRLEAN (our)	93.33	87.5	100	93.33

为了更好地验证本文提出的 3D 稠密倒残差模块和 Latent Equilibrium Attention (LEA)的有效性,我们 进行了如表 3 所示的消融实验。实验将只有 CNN 的网络作为基础架构,再分别用 LEA 和 3D DIRL 替换 base 中的部分层。我们将这三种网络框架组合与本文最终提出的 3D DIRLEAN (LEA + 3D DIRL)在 LUNA16 数据集上进行分类性能对比。3 个框架组合取得的 Accuracy、Recall、Precision 和 F1 分数如表 3 所示。基础 CNN 取得了 90%的 Accuracy 和 90.32 的 F1 分数。使用 LEA 或 3D DIRL 分别将精度提高 了 1.11%和 2.22%, F1 分数提高了约 0.6 和 2.3。两者同时使用,精度提高了 3.33%, F1 分数提高了 3.01。 这些结果表明 LEA 和 3D DIRL 都对提高模型的分类准确性和鲁棒性有贡献。并且 LEA 和 3D DIRL 对模 型性能的提升是可叠加的。

4. 结论

在本文中,我们提出了 3D DIRLEAN 框架用于实现少参数量和计算量的高精度肺结节良恶性分类。 整体框架仅由 4 个阶段组成,其中起决定性作用的是我们设计的 LEA 和 3DDIRL。LEA 通过我们构建的 潜在均衡能量函数在无需参数的条件下计算特征权重加强特征的区分性,同时缓解网络层之间的响应滞 后,加快模型收敛速度。而 3D DIRL 将标准卷积分解并加入稠密连接,在有效减少参数量的情况下最大 化利用多层不同特征,丰富网络获取的信息,以提升模型的分类精度。3D DIRLEAN 在 LUNA16 数据集 上的实验结果表明,相较于多种先进的肺结节分类方法,3D DIRLEAN 在分类精度、精确率和 F1 分数上 均表现出最优的性能。同时,其参数量和计算量也大幅减少,展现了出色的性能/成本比。总之,3D DIR-LEAN 通过引入 LEA 和 3D DIRL 有效提高了肺结节分类的精度,并解决了现有方法在模型复杂度和计 算资源上的问题。在未来的研究中,我们将进一步改进 3D DIRLEAN 框架,提升分类的精度和鲁棒性。 此外,我们还将对模型的可解释性进行研究。

基金项目

国家自然科学基金(8207070786);国家自然科学基金青年科学家基金(82302188);上海市浦江计划

(22PJD069); 上海市卫生人才研究基金(2022YQ060); 上海市国家科学技术委员会(22Y11911100); 上海市 卫生委员会(20204Y0201); 上海理工大学医学工程交叉项目(1021309706)。

参考文献

- Siegel, R.L., Miller, K.D., Fuchs, H.E. and Jemal, A. (2022) Cancer Statistics, 2022. CA: A Cancer Journal for Clinicians, 72, 7-33. <u>https://doi.org/10.3322/caac.21708</u>
- [2] Gomez-Bravo, D., Garcia, A., Vigueras, G., Rios-Sanchez, B., Otero, B., Hernandez, R., et al. (2022). Subgroup Discovery Analysis of Treatment Patterns in Lung Cancer Patients. 2022 IEEE 35th International Symposium on Computer-Based Medical Systems (CBMS), Shenzhen, 21-23 July 2022, 1-7. <u>https://doi.org/10.1109/cbms55023.2022.00082</u>
- [3] Liu, B., Chi, W., Li, X., Li, P., Liang, W., Liu, H., et al. (2019) Evolving the Pulmonary Nodules Diagnosis from Classical Approaches to Deep Learning-Aided Decision Support: Three Decades' Development Course and Future Prospect. Journal of Cancer Research and Clinical Oncology, 146, 153-185. <u>https://doi.org/10.1007/s00432-019-03098-5</u>
- [4] Li, Y., Chang, J. and Tian, Y. (2022) Improved Cost-Sensitive Multikernel Learning Support Vector Machine Algorithm Based on Particle Swarm Optimization in Pulmonary Nodule Recognition. *Soft Computing*, 26, 3369-3383. <u>https://doi.org/10.1007/s00500-021-06718-w</u>
- [5] He, Z., Lv, W. and Hu, J. (2020) A Simple Method to Train the AI Diagnosis Model of Pulmonary Nodules. *Computa*tional and Mathematical Methods in Medicine, 2020, 1-6. <u>https://doi.org/10.1155/2020/2812874</u>
- [6] Hu, K., Jin, J., Zheng, F., Weng, L. and Ding, Y. (2022) Overview of Behavior Recognition Based on Deep Learning. *Artificial Intelligence Review*, 56, 1833-1865. <u>https://doi.org/10.1007/s10462-022-10210-8</u>
- [7] Ge, M. and Zhang, Y. (2020) Visual Autopilot Decision System Based on Deep Learning. Proceedings of the 2nd International Conference on 3D Imaging Technologies-Multidimensional Signal Processing and Deep Learning, 3DIT-MSPandDL 2020, Kunming, 11-13 December 2020, 329-335.
- [8] Xie, Y. and Zhang, Y. (2021) Design of Speech Emotion Recognition Algorithm Based on Deep Learning. Proceedings of the 4th IEEE International Conference on Automation, Electronics and Electrical Engineering, AUTEEE 2021, Shenyang, 19-21 November 2021, 734-737.
- [9] Xie, H., Yang, D., Sun, N., Chen, Z. and Zhang, Y. (2019) Automated Pulmonary Nodule Detection in CT Images Using Deep Convolutional Neural Networks. *Pattern Recognition*, 85, 109-119. <u>https://doi.org/10.1016/j.patcog.2018.07.031</u>
- [10] Su, R., Xie, W. and Tan, T. (2020) 2.75D Convolutional Neural Network for Pulmonary Nodule Classification in Chest CT.
- [11] Wang, Y., Zhang, H., Chae, K.J., Choi, Y., Jin, G.Y. and Ko, S. (2020) Novel Convolutional Neural Network Architecture for Improved Pulmonary Nodule Classification on Computed Tomography. *Multidimensional Systems and Signal Processing*, **31**, 1163-1183. <u>https://doi.org/10.1007/s11045-020-00703-6</u>
- [12] Elhoussaine, E. and Salwa, B. (2021) Pulmonary Nodule Classification Based on Three Convolutional Neural Networks Models. In: *Lecture Notes on Data Engineering and Communications Technologies*, Springer International Publishing, 122-128. <u>https://doi.org/10.1007/978-3-030-70713-2_13</u>
- [13] Krizhevsky, A., Sutskever, I. and Hinton, G.E. (2017) Imagenet Classification with Deep Convolutional Neural Networks. *Communications of the ACM*, 60, 84-90. <u>https://doi.org/10.1145/3065386</u>
- [14] He, K., Zhang, X., Ren, S. and Sun, J. (2016) Deep Residual Learning for Image Recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, 27-30 June 2016, 770-778. https://doi.org/10.1109/cvpr.2016.90
- [15] Zhang, G., Lin, L. and Wang, J. (2021) Lung Nodule Classification in CT Images Using 3D Densenet. *Journal of Physics: Conference Series*, 1827, Article 012155. <u>https://doi.org/10.1088/1742-6596/1827/1/012155</u>
- [16] Apostolopoulos, I.D., Pintelas, E.G., Livieris, I.E., Apostolopoulos, D.J., Papathanasiou, N.D., Pintelas, P.E., et al. (2021) Automatic Classification of Solitary Pulmonary Nodules in PET/CT Imaging Employing Transfer Learning Techniques. Medical & Biological Engineering & Computing, 59, 1299-1310. <u>https://doi.org/10.1007/s11517-021-02378-y</u>
- [17] Sun, L., Wang, Z., Pu, H., Yuan, G., Guo, L., Pu, T., *et al.* (2021) Attention-Embedded Complementary-Stream CNN for False Positive Reduction in Pulmonary Nodule Detection. *Computers in Biology and Medicine*, **133**, Article 104357. <u>https://doi.org/10.1016/j.compbiomed.2021.104357</u>
- [18] Zhao, D., Liu, Y., Yin, H. and Wang, Z. (2022) A Novel Multi-Scale CNNs for False Positive Reduction in Pulmonary Nodule Detection. *Expert Systems with Applications*, **207**, Article 117652. <u>https://doi.org/10.1016/j.eswa.2022.117652</u>
- [19] Wu, Z., Ge, R., Shi, G., Zhang, L., Chen, Y., Luo, L., et al. (2020) MD-NDNet: A Multi-Dimensional Convolutional Neural Network for False-Positive Reduction in Pulmonary Nodule Detection. *Physics in Medicine & Biology*, 65, Article 235053. <u>https://doi.org/10.1088/1361-6560/aba87c</u>

- [20] li, S. and Liu, D. (2021) Automated Classification of Solitary Pulmonary Nodules Using Convolutional Neural Network Based on Transfer Learning Strategy. *Journal of Mechanics in Medicine and Biology*, 21, Article 2140002. https://doi.org/10.1142/s0219519421400029
- [21] Zuo, W., Zhou, F. and He, Y. (2020) An Embedded Multi-Branch 3D Convolution Neural Network for False Positive Reduction in Lung Nodule Detection. *Journal of Digital Imaging*, 33, 846-857. https://doi.org/10.1007/s10278-020-00326-0
- [22] Haiying, Y., Zhongwei, F., Ding, D. and Zengyang, S. (2021) False-positive Reduction of Pulmonary Nodule Detection Based on Deformable Convolutional Neural Networks. 2021 *IEEE 9th International Conference on Bioinformatics and Computational Biology (ICBCB)*, Taiyuan, 25-27 May 2021, 130-134. https://doi.org/10.1109/icbcb52223.2021.9459209
- [23] Yuan, H., Wu, Y., Cheng, J., Fan, Z. and Zeng, Z. (2022) Pulmonary Nodule Detection Using 3-D Residual U-Net Oriented Context-Guided Attention and Multi-Branch Classification Network. *IEEE Access*, 10, 82-98. https://doi.org/10.1109/access.2021.3137317
- [24] Han, Y., Qi, H., Wang, L., Chen, C., Miao, J., Xu, H., et al. (2022) Pulmonary Nodules Detection Assistant Platform: An Effective Computer Aided System for Early Pulmonary Nodules Detection in Physical Examination. Computer Methods and Programs in Biomedicine, 217, Article 106680. <u>https://doi.org/10.1016/j.cmpb.2022.106680</u>
- [25] Chen, S. (2022) Models of Artificial Intelligence-Assisted Diagnosis of Lung Cancer Pathology Based on Deep Learning Algorithms. *Journal of Healthcare Engineering*, **2022**, 1-12. <u>https://doi.org/10.1155/2022/3972298</u>
- [26] Shen, L.-H., Wang, X.-H., Gao, M.-X., et al. (2021) Classification of Benign-Malignant Pulmonary Nodules Based on Multi-View Improved Dense Network. Proceedings of the 17th International Conference on Intelligent Computing, ICIC 2021, Shenzhen, 12-15 August 2021, 582-593.
- [27] Hu, X., Gong, J., Zhou, W., Li, H., Wang, S., Wei, M., et al. (2021) Computer-Aided Diagnosis of Ground Glass Pulmonary Nodule by Fusing Deep Learning and Radiomics Features. *Physics in Medicine & Biology*, 66, Article 065015. <u>https://doi.org/10.1088/1361-6560/abe735</u>
- [28] Susan, S., Sethi, D. and Arora, K. (2020) CW-CAE: Pulmonary Nodule Detection from Imbalanced Dataset Using Class-Weighted Convolutional Autoencoder. *Proceedings of the 3rd International Conference on Innovative Computing and Communication*, ICICC 2020, Delhi, 21-23 February 2020, 825-833.
- [29] Calheiros, J.L.L., de Amorim, L.B.V., de Lima, L.L., de Lima Filho, A.F., Ferreira Júnior, J.R. and de Oliveira, M.C. (2021) The Effects of Perinodular Features on Solid Lung Nodule Classification. *Journal of Digital Imaging*, 34, 798-810. <u>https://doi.org/10.1007/s10278-021-00453-2</u>
- [30] Zhang, X., Wang, K., Zhang, X., et al. (2021) Pulmonary Nodule Classification of CT Images with Attribute Self-Guided Graph Convolutional V-Shape Networks. Proceedings of the 18th Pacific Rim International Conference on Artificial Intelligence, PRICAI 2021, Online, 8-12 November 2021, 280-292.
- [31] Shen, S., Han, S.X., Aberle, D.R., Bui, A.A. and Hsu, W. (2019) An Interpretable Deep Hierarchical Semantic Convolutional Neural Network for Lung Nodule Malignancy Classification. *Expert Systems with Applications*, **128**, 84-95. https://doi.org/10.1016/j.eswa.2019.01.048
- [32] Jiang, H., Shen, F., Gao, F. and Han, W. (2021) Learning Efficient, Explainable and Discriminative Representations for Pulmonary Nodules Classification. *Pattern Recognition*, **113**, Article 107825. https://doi.org/10.1016/j.patcog.2021.107825
- [33] Hu, J., Shen, L., Albanie, S., Sun, G. and Wu, E. (2020) Squeeze-and-Excitation Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42, 2011-2023. <u>https://doi.org/10.1109/tpami.2019.2913372</u>
- [34] Woo, S., Park, J., Lee, J. and Kweon, I.S. (2018) CBAM: Convolutional Block Attention Module. In: Lecture Notes in Computer Science, Springer, 3-19. <u>https://doi.org/10.1007/978-3-030-01234-2_1</u>
- [35] Lee, H., Kim, H. and Nam, H. (2019) SRM: A Style-Based Recalibration Module for Convolutional Neural Networks. 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, 27 October-2 November 2019, 1854-1862. <u>https://doi.org/10.1109/iccv.2019.00194</u>
- [36] Haider, P., Ellenberger, B., Kriener, L., et al. (2021) Latent Equilibrium: A Unified Learning Theory for Arbitrarily Fast Computation with Arbitrarily Slow Neurons. Advances in Neural Information Processing Systems, 22, 17839-17851.
- [37] Shen, W., Zhou, M., Yang, F., Yu, D., Dong, D., Yang, C., et al. (2017) Multi-Crop Convolutional Neural Networks for Lung Nodule Malignancy Suspiciousness Classification. Pattern Recognition, 61, 663-673. <u>https://doi.org/10.1016/j.patcog.2016.05.029</u>
- [38] Yan, X., Pang, J., Qi, H., Zhu, Y., Bai, C., Geng, X., et al. (2017) Classification of Lung Nodule Malignancy Risk on Computed Tomography Images Using Convolutional Neural Network: A Comparison between 2D and 3D Strategies. In: Lecture Notes in Computer Science, Springer, 91-101. <u>https://doi.org/10.1007/978-3-319-54526-4_7</u>
- [39] Zhu, W., Liu, C., Fan, W. and Xie, X. (2018) Deeplung: Deep 3D Dual Path Nets for Automated Pulmonary Nodule

Detection and Classification. 2018 *IEEE Winter Conference on Applications of Computer Vision (WACV)*, Lake Tahoe, 12-15 March 2018, 673-681. <u>https://doi.org/10.1109/wacv.2018.00079</u>

[40] Jiang, H., Gao, F., Xu, X., Huang, F. and Zhu, S. (2020) Attentive and Ensemble 3D Dual Path Networks for Pulmonary Nodules Classification. *Neurocomputing*, **398**, 422-430. <u>https://doi.org/10.1016/j.neucom.2019.03.103</u>