

融合深度残差网络和字典学习的肺炎检测^①

朱之强¹, 卞维新^{1,2}, 接标^{1,2}, 黄宜¹, 李文虎¹

¹(安徽师范大学 计算机与信息学院, 芜湖 241002)

²(网络与信息安全安徽省重点实验室, 芜湖 241002)

通信作者: 卞维新, E-mail: bwx2353@ahnu.edu.cn



摘要: 由于空气污染与吸烟等原因, 肺炎已成为人类死亡率最高的疾病之一。随着机器学习与深度学习技术在医疗图像检测上的应用, 为临床专家诊断各类疾病提供了帮助。但由于缺少有效的配对肺部 X 射线数据集, 以及现有针对肺炎检测的方法均采用不是针对肺炎任务的普遍分类模型, 难以发现肺炎图像与正常图像的细微差别, 导致识别失败。为此, 本文通过数据裁剪、旋转等方式扩充数据集中的正常图像, 再使用 50 层深度残差网络对胸部 X 射线中的浅层肺炎特征进行学习; 然后, 通过两层字典对残差网络学习到的肺炎特征进行更深入的抽象和学习, 发现不同肺部图像之间的微小差别; 最后, 融合残差网络和字典学习提取到的多级肺炎特征, 构建肺炎检测模型。为了验证算法的有效性, 在 Chest X-ray 肺炎数据集上评估肺炎检测模型的性能。根据测试结果, 本文提出模型的检测准确率为 97.12%; 指标测试中, 精度与召回率之间的调和平均数上的得分为 97.73%。与现有方法相比, 获得了更高的识别精度。

关键词: 肺炎; 扩充数据集; 深度学习; 深度残差网络; 两层字典学习

引用格式: 朱之强, 卞维新, 接标, 黄宜, 李文虎. 融合深度残差网络和字典学习的肺炎检测. 计算机系统应用, 2024, 33(3):95-102. <http://www.c-s-a.org.cn/1003-3254/9436.html>

Pneumonia Detection Based on Deep Residual Network and Dictionary Learning

ZHU Zhi-Qiang¹, BIAN Wei-Xin^{1,2}, JIE Biao^{1,2}, HUANG Yi¹, LI Wen-Hu¹

¹(School of Computer and Information, Anhui Normal University, Wuhu 241002, China)

²(Anhui Provincial Key Laboratory of Network and Information Security, Wuhu 241002, China)

Abstract: Due to factors such as air pollution and smoking, pneumonia has become one of the diseases with the highest mortality rates in humans. The application of machine learning and deep learning technology in medical image detection has provided assistance for clinical experts in diagnosing various diseases. However, there is a lack of effective paired lung X-ray datasets, and existing methods for pneumonia detection use universal classification models that are not specific to pneumonia tasks. As a result, it is difficult to detect subtle differences between pneumonia images and normal images, resulting in recognition failure. Therefore, this study expands the normal images in the dataset through data cropping, rotation, and other methods and uses a 50-layer deep residual network to learn the shallow pneumonia features in chest X-rays. Then, through a two-layer dictionary, the pneumonia features learned by the residual network are further abstracted and learned, and subtle differences between different lung images are discovered. Finally, a pneumonia detection model is constructed by fusing the multi-level pneumonia features extracted from residual networks and dictionary learning. To validate the effectiveness of the algorithm, the performance of the pneumonia detection model is evaluated on the chest X-ray pneumonia dataset. According to the test results, the proposed model has a detection accuracy of 97.12%. In the indicator test, the score on the harmonic mean between accuracy and recall is 97.73%. Compared with existing methods, it has

① 基金项目: 安徽省自然科学基金 (2108085MF206); 国家自然科学基金 (61976006)

收稿时间: 2023-09-02; 修改时间: 2023-10-08, 2023-10-25; 采用时间: 2023-11-03; csa 在线出版时间: 2024-01-09

CNKI 网络首发时间: 2024-01-10

achieved higher recognition accuracy.

Key words: pneumonia; extended dataset; deep learning; deep residual network; two-layer dictionary learning

肺炎是指终末气道、肺泡和肺间质的炎症。常见的肺炎类型主要有两种:细菌性肺炎与病毒性肺炎。其中,细菌性肺炎是由于人体受到细菌感染引起的,是当人体免疫力下降时,受到外来细菌感染,造成的肺泡支气管粘膜水肿。病毒性肺炎是由上呼吸道病毒性感染引起的,随着病毒的蔓延,最终会危害到肺部。对于存在咳嗽和呼吸困难,或可能伴有发热的5岁以下儿童来说,当出现呼吸加快或下胸壁内收等情况,即吸气时胸部内收或收缩(健康的人吸气时,胸部会扩张),就可做出肺炎诊断。喘息较常见于病毒性感染。肺炎可以使用抗生素治疗,但大多数患者不愿使用昂贵的口服抗生素,所以提前预防肺炎就成为降低其危害的关键方法。现有研究表明,借助计算机技术检测早期肺炎疾病,降低这类疾病的恶化概率是可行的。针对肺炎疾病目前有效的检测手段之一是通过肺部X射线图像来诊断^[1]。在得到患者的胸部X射线图像后,由相关的医生进行判断。

深度学习与字典学习近年来被广泛研究,在图像处理领域也取得了巨大成功,对图像增强(徐德琴等人^[2])、分类(Singhal等人^[3])、去噪(Zheng等人^[4])等任务都有优异的表现。许多研究者开始使用这类技术诊断生物医学图像。相比于专业的医生而言,其能在短时间内得到相同的预测结果,有时可能会获得更好的准确性。不过该类技术只是医生诊断相应疾病的一种辅助手段,为医生提供有价值的参考,并不能在临床诊断中取代医生。

本文基于迁移学习和深度字典学习,提出了一种可以有效分辨患者是否患有肺炎的模型。先通过ResNet-50(50层的深度残差网络)学习肺炎与非肺炎图像的内生特征;再通过深度字典学习,提取图像的深层特征;最终将两者学习到的特征进行融合,用于完成肺部X射线图像的肺炎检测任务。

相比已有的工作,本文主要贡献点如下。

1) 通过将深度残差网络与字典学习相结合的方式训练肺炎检测模型,与现有主流深度学习框架相比,进一步提升了肺炎检测的精度。

2) 通过数据增强来扩充原始训练集中的非肺炎图像数量,弥补数据集之间不平衡的缺陷,进一步提升模

型的泛化能力。

1 相关工作

深度学习技术为医学图像的检测与诊断提供了有效帮助。黎英等人^[5]提出了基于模型的迁移学习应用。Hashmi等人^[6]采用深度迁移学习对胸部X射线图像中的肺炎问题进行检测,通过加权分类器将多个深度学习模型学得的权重组合。Wang等人^[7]对迁移学习在医学图像及其子领域上的应用进行了论述。Lee等人^[8]基于深度学习对胸部X光图像中的区域特征分割的性能进行比较,指出提取到合理特征对最终结果的影响。Albahli等人^[9]基于迁移学习对胸部X射线中的14种疾病检测,实验结果表明迁移学习有助于提高胸部X射线图像上疾病分类精度。Murugan等人^[10]将深度学习用于视网膜图像的异常检测。Bajpai等人^[11]使用不同的深度学习模型对脑电图进行病理检测,他们通过从光谱图像中提取特征来训练模型。Coudray等人^[12]使用深度学习对非小细胞肺癌组织图像进行分类和突变预测,通过基因组图谱训练网络,获得了和病理学专家相当的准确度。Fang等人^[13]将基于深度学习的多模态图像检测,用于阿尔茨海默病诊断。Hu等人^[14]基于多核深度卷积,学习胸部X射线图像特征。多核卷积是指一个深度卷积层中包含具有不同滤波器大小的深度卷积核,能够适用大小不同的医学图像诊断任务。Oh等人^[15]通过卷积神经网络联合学习多种疾病的相似性。An等人^[16]提出了一种基于多层边界感知的自注意力深度学习模型,用于医学图像分割,引入了一种新的交互式自注意力模块,可以获得全局特征信息。

上述方法的着重点在于使用单一的深度学习模型或其组合来构建最终输出,对医学图像中的疾病进行检测。

由于字典学习基于稀疏模型的理论,可在消除图像冗余信息的同时保留其关键特征。

Zheng等人^[17]将字典学习用于人脸识别,他们通过基于局部约束和标签嵌入的词典算法进行学习,降低噪声对图像识别精度的影响。Zheng等人^[18]将字典学习与迁移学习结合,提出了一种基于词典学习的多

任务迁移学习方法. Zhang 等人^[19]通过代价敏感联合特征和字典学习进行人脸识别, 特征与字典同时学习, 以实现最小的整体损失. Zhang 等人^[20]提出了图像分类的最优判别特征和字典学习方法, 学习特征映射矩阵和字典, 以便在映射的特征空间中最大化数据的类间稀疏重建误差, 最小化类内稀疏重建误差. 另外为了提高字典学习的性能, Liu 等人^[21]提出了一类学习的多任务分析判别字典学习, 他们将分析判别字典学习合并到一类学习中, 提高了分类的准确性. Rodriguez-Dominguez 等人^[22]提出了一种新的分层判别式深度字典学习方法, 将深度字典学习嵌入到具有图像分类目标的神经网络中. Montazeri 等人^[23]提出了一种深度字典学习的新方法, 将多层奇异值分解用于图像分类. Wu 等人^[24]提出了一种在线深度迁移字典学习的方法用于图像分类. Tang 等人^[25]将深度字典学习与编码网络用于有限数据的图像识别, 用深度字典学习与稀疏编码取代深度学习中的卷积层.

最近的这些研究表明, 在单层字典学习的基础上, 构建更深层次的字典, 可以有效解决训练样本不平衡, 以及深层特征难以提取的问题.

目前提出的基于深度学习进行胸部肺炎疾病检测模型具有如下问题.

1) 用于胸部肺炎检测的深度学习模型设计过于简单, 或是在复杂模型的基础上仅通过迁移学习的方式, 添加新的特征提取层, 可能并不适用于肺炎检测任务.

2) 通过多核卷积或不同方法的权重叠加, 进行胸部肺炎检测, 可能会将重点放在相似或重复的浅层特征上, 无法有效提高识别精度.

针对上述问题, 本文结合深度字典学习, 提出了一种融合残差网络和字典学习的肺炎检测模型. DRDLNET (deep residual dictionary learning network). 在公开数据集上的实验结果表明, 对比目前最先进的肺炎检测方法, DRDLNET 取得了最好的结果.

2 胸部肺炎预测模型

2.1 字典学习

字典学习也被称为稀疏表示, 常用的字典学习方法是 K-SVD. 其学习目的是对给定数据样本进行矩阵分解, 保留数据中主要的特征, 去除冗余信息. 通过学习到的字典和稀疏编码将原本稠密的样本转为稀疏表现形式. 在字典学习之前, 先从给定数据样本中随机挑

选部分样本构建初始字典矩阵, 其中对于稀疏编码的求解可通过正交匹配追踪算法 OMP 进行. 设输入样本为 X_i , 字典 $D = \{d_1, d_2, \dots, d_k\}$, 该样本的稀疏编码 $Z_i = \{Z_i^1, Z_i^2, \dots, Z_i^k\}$, 通过 OMP 求出字典 D 中距离 X_i 最近的一列 d_i , 根据 $X_i = Z_i^i d_i$ 可求出对应的 Z_i^i 的具体值; 再通过式 (1) 计算 X_i' , 当其小于固定阈值时计算结束, 否则按照相同步骤继续求解新的 Z_i^j ; 然后在式 (1) 的基础上减去 $d_j Z_j$, 重复上述过程, 直到 X_i' 小于指定阈值. 通过该步骤, 即可求出所有样本的稀疏编码. 根据 K 个样本求得的稀疏编码 Z 来逐列更新字典, 以字典中第 k 列 d_k 为例, 取出稀疏编码 Z 中第 k 个行向量中不为零的索引集合, 根据式 (2) 计算误差矩阵 E_k , 并取出其中对应稀疏编码 Z 中不为零的列, 设为 E_k' , 对 E_k' 通过 K-SVD 进行奇异值分解, 根据分解结果更新字典中的第 k 列. 字典与稀疏编码的更新交替进行, 直到满足指定阈值. 字典学习网络结构如图 1 所示. 具体过程见式 (3):

$$X_i' = X_i - d_i z_i \quad (1)$$

$$E_k = X - \sum_{j \neq k} d_j z_j \quad (2)$$

$$\min_{D, Z} \|X - DZ\|_F^2 \quad (3)$$

其中, X 是字典学习的输入数据, 本文中将 ResNet50 提取到的多级特征作为输入. D 是学习到的字典, Z 是稀疏编码. 为了满足稀疏编码的稀疏性, 需要约定:

$$\|Z\|_0 \leq L \quad (4)$$

其中, L 是约束项, 指稀疏编码中非零元素的个数.

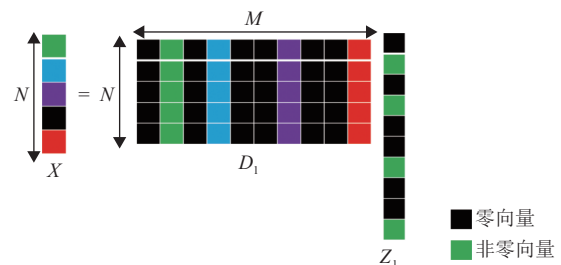


图 1 字典学习网络结构图

2.2 深度字典学习

由于单层字典学习很难学习到图像中的深层特征, 以及有限的数据集可能会导致训练样本的不平衡, 进而导致在实际应用时达不到预期效果. 随着深度学习

在图像处理领域取得的成功,这得益于其提取的多层特征更能表示图像中的复杂信息.基于单层字典学习的思想,已有学者提出了深度字典学习的概念,主要为在学习到单层字典与稀疏编码的基础上,再次学习新的字典,提取更深层次的特征.深度字典学习伪代码如算法1所示.

算法1.深度字典学习

- 1) 输入样本: X
- 2) 初始化字典: $D_i, i=1, \dots, N$
- 3) 进行第1层字典学习:

$$D_1 = \min_{D_1} \|X - D_1 Z_1\|_F^2$$

$$Z_1 = \min_{Z_1} \|X - D_1 Z_1\|_F^2$$
- 4) 进行第2层字典学习:

$$D_2 = \min_{D_2} \|D_1 - D_2 Z_2\|_F^2$$

$$Z_2 = \min_{Z_2} \|D_1 - D_2 Z_2\|_F^2$$
- 5) 进行第 N 层字典学习:

$$D_N = \min_{D_N} \|D_{N-1} - D_N Z_N\|_F^2$$

$$Z_N = \min_{Z_N} \|D_{N-1} - D_N Z_N\|_F^2$$
- 6) 输出: D_N, Z_N

在算法1中 D_N 表示第 N 层字典, Z_N 表示第 N 层字典的稀疏编码.他们可以由第 $N-1$ 层字典学习得到.由此可见,更深层次的字典是建立在前一层字典基础上的单层字典学习.这样就可以保证各层字典之间不会相互重叠,而且更深层的字典可以和前一层字典之间保持内在关联,形成特征融合.同样对 Z_N 也应满足稀疏性,即:

$$\|Z_N\|_0 \leq L \tag{5}$$

2.3 深度残差字典学习网络

深度学习技术在医学图像上的应用也面临着许多挑战.一方面缺少公开可用的医学数据集,另一方面是图像中的细节特征难以学习. Lu 等人^[26]基于 AlexNet 和迁移学习进行病理性脑检测,在预训练的 AlexNet 结构的基础之上,用随机权重替换最后3层的参数作

为初始网络模型进行训练.为了学习图像可靠的内在特征,往往需要很深的网络层数,但过深的网络层数,通常会导致梯度消失与过度拟合等问题.为此,有学者提出了深度残差网络的概念.在残差网络中通过“跳过连接”的思想,以及不使用删除神经元的方式来对抗过拟合,使其在分类等任务中取得了优异的结果.

本文结合深度残差网络与字典学习构建了一种深度残差字典学习网络,图2为构建的网络结构图,从左到右依次为预处理数据集、ResNet50层、第1层字典学习、第2层字典学习、全连接层、输出层、预测结果标签.首先将预处理后的胸部X射线图像通过迁移学习的方式输入 ResNet50,提取初始特征,在 ResNet50 中包含 49 个卷积层和 1 个全连接层,其开始部分通过卷积层提取图像浅层特征,后面随着网络层数的加深,可能会导致梯度消失或准确率下降的问题.结合具有跳跃连接的残差块来避免梯度消失,同时实现输出特征之间的跨层传递,学习更深层次的胸部肺炎图像信息.其次,将 ResNet50 提取的特征作为第1层字典学习的输入,利用浅层字典学习强大的特征筛选功能对其进行优化,使得最终输出中冗余的特征尽可能减少,提升识别效率.然后,将第1层字典学习的输出作为第2层字典学习的输入,提取更深层次特征,本文学习了两层字典,由于 ResNet50 已经对图像的浅层与深层特征进行了提取,字典学习是在其基础上的完善,通过实验比较发现更深层次的字典无法获得更好的识别准确率.最后,将 ResNet50 提取的特征与两层字典提取到的特征进行融合,仅通过两层字典学习或 ResNet50 作为模型输出,会出现部分特征丢失,ResNet50 中包含了图像的多级特征,字典学习对其优化后获得了更深层次的信息,将其结合起来会获得更好的效果.最终通过全连接层与输出层产生预测结果.

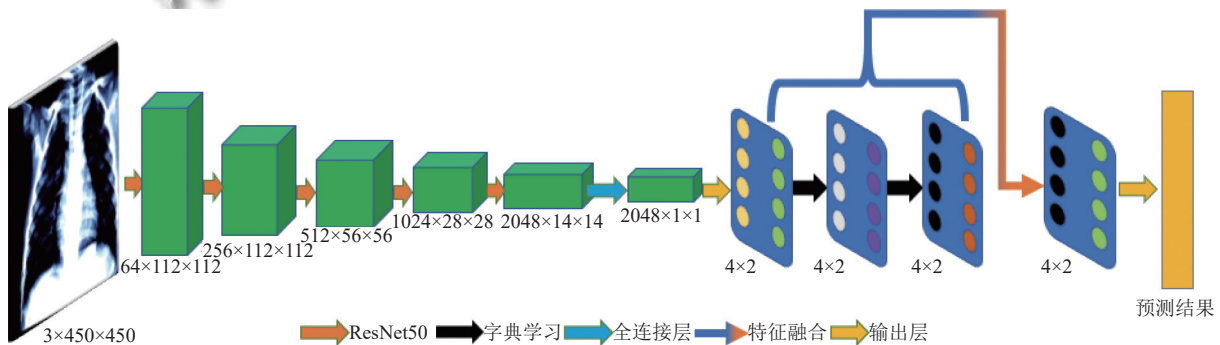


图2 深度残差字典学习网络结构图

具体学习过程为首先由深度残差网络进行第1阶段的特征学习,再由深度字典进行第2阶段的学习。

在对模型训练之前,需要对数据集进行预处理,以满足训练要求,处理过程如式(6)、式(7)所示:

$$Y = (\text{Normal}, \text{Pneumonia}) \quad (6)$$

$$C = \lambda(Y) \quad (7)$$

其中, Y 表示原始训练样本集,它包含两类训练样本,分别是代表正常肺部图像的 Normal, 以及肺炎图像 Pneumonia. C 是对原始数据集进行预处理后的结果. λ 表示预处理操作. 本文对原始数据进行的预处理包括中心裁剪、概率水平翻转、随机旋转、概率转灰度图、仿射变换以及标准化,相关参数将在第3节给出。

训练样本准备完毕后,将其送入 DRDLNET 模型中进行学习. 输入数据先经过残差网络 ResNet50 进行第1阶段的学习,如式(8)所示. 再对 ResNet50 学习的多级特征 X , 通过两层字典进一步优化, 过滤掉冗余信息, 获得更深层次的特征. 字典学习的过程如式(9)、式(10)所示. 其中 C 是预处理数据集, X 是经过 ResNet50 学习后得到的第1部分的输出. D_1 、 D_2 分别是第1层、第2层字典学习的输出。

$$X = \text{ResNet50}(C) \quad (8)$$

$$D_1 = \min_{D_1} \|X - D_1 Z_1\|_F^2 \quad (9)$$

$$D_2 = \min_{D_2} \|D_1 - D_2 Z_2\|_F^2 \quad (10)$$

两层字典学习结束后,将学习的深层特征与 ResNet50 学习的多级特征融合,使得最终输出包含两者共同学习的结果. 这将有助于区分肺炎图像与正常肺部图像之间的细微差别. 如式(11)所示,其中 τ 表示将学习到的特征进行融合,融合方式是将 X 的第1列特征与 D_2 的第2列特征作为最终输出. 组合结果 W 是最终的分类输出。

$$W = \tau(X, D_2) \quad (11)$$

得到输出结果后,通过 Softmax 分类器结合输出权重对输入数据进行预测. 如式(12)、式(13)所示: 其中 $P(y|x)$ 是将某一类肺部 X 射线图像的输出映射为 0-1 之间的概率值, x 表示某一类的输入, y 表示将输入 x 预测为正常图像以及肺炎图像的概率值. $W_{y..x}$ 表示模型针对单一肺部图像的预测结果. $W_{c..x}$ 是所有预

测结果之和,其中 $n=2$, 以此来判断模型的输出与真实值之间的差异。

$$P(y|x) = \frac{\exp(W_{y..x})}{\sum_{c=1}^n \exp(W_{c..x})} \quad (12)$$

$$W_{y..x} = \sum_{i=1}^d W_{y_i..x_i} \quad (13)$$

最终,再用交叉熵损失函数来计算反向传播过程中的损失,对模型参数进行优化. 如式(14)所示:

$$H(p, q) = - \sum_{i=1}^n p(x_i) \log(q(x_i)) \quad (14)$$

其中, $p(x_i)$ 表示真实标签, $q(x_i)$ 表示的是预测结果,将所有预测值与真实值的对数乘积进行求和,作为最终损失。

3 实验结果分析

3.1 数据集

实验采用的数据集包含 5840 张胸部 X 射线图像,其中 5216 张用于训练,624 张用于测试. 训练集与测试集的具体划分如表1所示. 在训练时,由于原数据集中包含 1341 张正常胸部 X 射线图像与 3875 张含有肺炎的图像,故通过裁剪等数据增强手段来扩充数据集. 扩充结果如图3所示. 最终训练集中有 3042 张正常胸部 X 射线图像,以及 3875 张含有肺炎的胸部 X 射线图像。

表1 数据集划分

数据集	总数	比例 (%)
训练集	6917	91.73
测试集	624	8.27

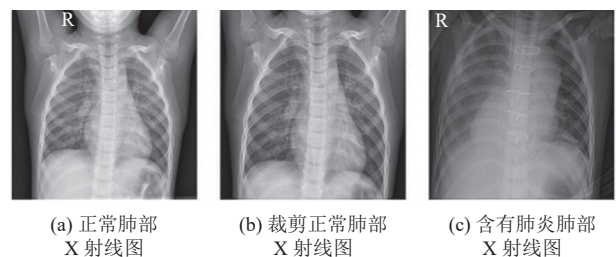


图3 数据扩充

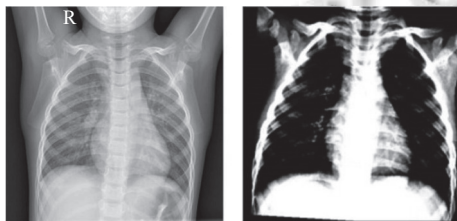
3.2 数据预处理

在使用 ResNet50 训练模型之前,需要对原始数据进行预处理. 使用预处理后的数据集训练模型可以使

模型学到更多的数据特征,从而提升模型的泛化能力.需要说明的是对于测试数据的预处理,本文只将其大小调整为和训练模型时的大小相同(本文中训练与测试样本的输入大小为450×450.并且对于测试样本的标准化也和训练样本一致.其他对于训练集使用的预处理方法并未对测试集使用.本文中用到的预处理技术的各项参数设置如表2所示.图4给出了一个样本预处理的结果示例.

表2 数据预处理

方法	参数
中心裁剪	450
概率水平翻转	0.5
随机旋转	10
概率转灰度图	0.1
仿射变换	(0.05, 0.05)



(a) 原始肺部 X 射线图 (b) 预处理肺部 X 射线图

图4 数据预处理示例

3.3 超参数设置

各项实验均采用相同的超参数,根据实验比较,得出了超参数的设定最优值.学习率设置为1E-5,优化器采用Adam.所有训练都建立在迁移学习的基础上,迭代次数在30次之后精度与损失曲线趋于平稳,达到拟合.

3.4 指标

在测试阶段通过准确率、精度、召回率以及F1得分作为模型的评价指标.其中准确率是在测试集中预测值与实际值的接近程度;精度是模型正确预测

肺炎图像的能力;召回率是正确预测为肺炎的个数相对于正确预测为肺炎的个数加上错误预测为正常图像的总个数的比值;F1得分是精度与召回率之间的调和平均数.在评价这些指标时,需要用到的参数有TP(肺炎图像预测正确的个数),TN(正常肺部图像预测正确的个数),FP(正常肺部图像预测失败的个数),FN(肺炎图像预测失败的个数).为了方便表示,这里用Acc表示准确率, Precision表示精度, Recall表示召回率.计算过程如下:

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \quad (15)$$

$$Precision = \frac{TP}{TP + FP} \quad (16)$$

$$Recall = \frac{TP}{TP + FN} \quad (17)$$

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (18)$$

3.5 实验结果

实验中将DRDLNET模型与现在主流的深度学习网络架构ResNet50、18层深度残差网络ResNet18、16层视觉几何群网络VGG16、121层稠密卷积网络DenseNet121、以及GoogLeNet网络、AlexNet网络、InceptionV3网络、MobileNetV2网络等进行了对比.同时还对比了最新的深度残差网络模型ResNet101和ResNet152.所有实验都在同一台计算机上完成,该计算机RAM为16GB,处理器为AMD Ryzen 7 5800H with Radeon Graphics,显卡为NVIDIA GeForce RTX 3060 6GB GPU.对比实验结果表明,DRDLNET模型在相同的数据集与训练环境下,获得了最好的结果.图5给出了不同方法在Chest X-ray数据集上的测试精度与损失比较.

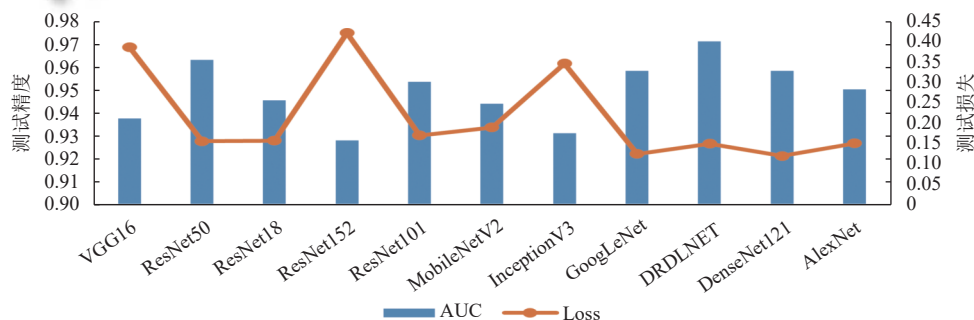


图5 Chest X-ray数据集上的测试精度与损失对比

通过图5、表3和表4中的混淆矩阵参数可以看出,在相同的测试集上,本文提出的方法要优于其他模型。DRDLNET在非肺炎图像上的预测结果最好,同时在肺炎图像的预测上只有本文提出的方法和InceptionV3模型取得了最好的结果。需要指出的是更深层次的深度残差网络并未在同样的数据集上取得更好的结果,这表明一味地叠加网络层数,已经无法在本数据集中学习到更深层次的数据特征,因此需要探索新的方法。

表3 测试准确率、精度、召回率、F1得分结果比较(%)

方法	Acc	Precision	Recall	F1
DRDLNET	97.11	96.03	99.49	97.73
ResNet50	96.31	95.53	98.72	97.10
ResNet18	94.55	93.63	97.95	95.40
VGG16	93.75	91.89	98.72	95.18
AlexNet	95.03	93.89	98.46	96.12
DenseNet121	95.83	96.67	96.67	96.67
GoogLeNet	95.83	94.83	98.72	96.74
InceptionV3	93.11	90.44	99.49	94.75
MobileNetV2	94.39	93.40	97.95	95.62
ResNet101	95.03	93.92	98.97	96.38
ResNet152	92.79	90/40	98.97	94.49

表4 混淆矩阵结果比较

方法	Normal	Pneumonia	F-Normal	F-Pneumonia
DRDLNET	218	388	2	16
ResNet50	216	385	5	18
ResNet18	208	382	8	26
VGG16	200	385	5	34
AlexNet	209	384	6	25
DenseNet121	221	377	13	13
GoogLeNet	213	285	5	21
InceptionV3	193	388	2	41
MobileNetV2	207	382	8	27
ResNet101	209	386	4	25
ResNet152	207	382	8	27

每种方法在测试集上的测试结果所对应的Acc、Precision、Recall和F1得分在表3中给出。

下面给出不同方法在相同测试集上测试的结果,用混淆矩阵来表示每种方法的对比结果。在测试集中,阴性图像有234张,肺炎图像有390张。在表4中Normal表示将阴性肺部图像判断为阴性,Pneumonia表示将肺炎图像判断为阳性,F-Normal表示将肺炎图像判断为阴性,F-Pneumonia表示将阴性肺部图像判断为肺炎图像。

表5中给出了每个模型在相同测试集上的测试精度对比,这里用Acc表示每个模型的测试精度。通过测试精度的比较可以看出DRDLNET模型取得了97.12%

的准确率,相比于ResNet50模型提升了约0.81%,这将有助于肺炎的早期预防与诊断工作。

表5 测试准确率对比

方法	Acc (%)
DRDLNET	97.12
ResNet50	96.31
ResNet18	94.55
VGG16	93.75
AlexNet	95.03
DenseNet121	95.83
GoogLeNet	95.83
InceptionV3	93.11
MobileNetV2	94.39
ResNet101	95.35
ResNet152	92.79

4 结语

肺炎疾病给全世界人民带来了严重的危害,尽早发现与诊断这一疾病意义重大。深度学习技术和基于图像稀疏表示的字典学习技术在图像识别领域展示了强大的潜力。结合二者的特征学习能力和识别能力,本文提出了一种深度残差字典学习网络,用于胸部X射线图像中的肺炎检测。提出的模型能够在有限的数据集中学到更为可靠和准确的特征,从而提高模型的识别精度。在Chest X-ray肺炎数据集上将提出的模型和多个已知的主流深度学习模型进行了评价和对比,实验结果表明,DRDLNET的特征融合可以有效提高胸部肺炎疾病的识别准确率。在未来针对医疗图像的研究中希望能够用更加有效的数据扩充方法来弥补数据集之间的不平衡问题,以及探索更多网络模型与深度字典学习相结合的方式有效处理医学图像相关问题。

参考文献

- Çınar A, Yıldırım M, Eroğlu Y. Classification of pneumonia cell images using improved ResNet50 model. *Traitement du Signal*, 2021, 38(1): 165–173. [doi: 10.18280/ts.380117]
- 徐德琴, 卞维新, 丁新涛, 等. 指纹图像多尺度分类字典稀疏增强. *中国图象图形学报*, 2018, 23(7): 1014–1023. [doi: 10.11834/jig.170632]
- Singhal V, Aggarwal HK, Tariyal S, et al. Discriminative robust deep dictionary learning for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 2017, 55(9): 5274–5283. [doi: 10.1109/TGRS.2017.2704590]

- 4 Zheng HY, Yong HW, Zhang L. Deep convolutional dictionary learning for image denoising. Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021. 630–641.
- 5 黎英, 宋佩华. 迁移学习在医学图像分类中的研究进展. 中国图象图形学报, 2022, 27(3): 672–686. [doi: [10.11834/jig.210814](https://doi.org/10.11834/jig.210814)]
- 6 Hashmi MF, Katiyar S, Keskar AG, *et al.* Efficient pneumonia detection in chest X-ray images using deep transfer learning. Diagnostics, 2020, 10(6): 417. [doi: [10.3390/diagnostics10060417](https://doi.org/10.3390/diagnostics10060417)]
- 7 Wang J, Zhu HD, Wang SH, *et al.* A review of deep learning on medical image analysis. Mobile Networks and Applications, 2021, 26(1): 351–380. [doi: [10.1007/s11036-020-01672-7](https://doi.org/10.1007/s11036-020-01672-7)]
- 8 Lee HM, Kim YJ, Kim KG. Segmentation performance comparison considering regional characteristics in chest X-ray using deep learning. Sensors, 2022, 22(9): 3143. [doi: [10.3390/s22093143](https://doi.org/10.3390/s22093143)]
- 9 Albahli S, Ahmad Hassan Yar GN. AI-driven deep convolutional neural networks for chest X-ray pathology identification. Journal of X-ray Science and Technology, 2022, 30(2): 365–376. [doi: [10.3233/XST-211082](https://doi.org/10.3233/XST-211082)]
- 10 Murugan R, Roy P, Singh U. An abnormality detection of retinal fundus images by deep convolutional neural networks. Multimedia Tools and Applications, 2020, 79(33): 24949–24967.
- 11 Bajpai R, Yuvaraj R, Prince AA. Automated EEG pathology detection based on different convolutional neural network models: Deep learning approach. Computers in Biology and Medicine, 2021, 133: 104434. [doi: [10.1016/j.combiomed.2021.104434](https://doi.org/10.1016/j.combiomed.2021.104434)]
- 12 Coudray N, Ocampo PS, Sakellaropoulos T, *et al.* Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. Nature Medicine, 2018, 24(10): 1559–1567. [doi: [10.1038/s41591-018-0177-5](https://doi.org/10.1038/s41591-018-0177-5)]
- 13 Fang XS, Liu ZB, Xu MC. Ensemble of deep convolutional neural networks based multi-modality images for Alzheimer's disease diagnosis. IET Image Processing, 2020, 14(2): 318–326. [doi: [10.1049/iet-ipr.2019.0617](https://doi.org/10.1049/iet-ipr.2019.0617)]
- 14 Hu MJ, Lin HZ, Fan ZM, *et al.* Learning to recognize chest X-ray images faster and more efficiently based on multi-kernel depthwise convolution. IEEE Access, 2020, 8: 37265–37274. [doi: [10.1109/ACCESS.2020.2974242](https://doi.org/10.1109/ACCESS.2020.2974242)]
- 15 Oh SH, Back S, Park J. Measuring patient similarity on multiple diseases by joint learning via a convolutional neural network. Sensors, 2022, 22(1): 131.
- 16 An FP, Liu JE. Medical image segmentation algorithm based on multilayer boundary perception-self attention deep learning model. Multimedia Tools and Applications, 2021, 80(10): 15017–15039. [doi: [10.1007/s11042-021-10515-w](https://doi.org/10.1007/s11042-021-10515-w)]
- 17 Zheng SJ, Zhang YJ, Liu WJ, *et al.* A dictionary learning algorithm based on dictionary reconstruction and its application in face recognition. Mathematical Problems in Engineering, 2020, 2020: 8964321.
- 18 Zheng X, Lin LY, Liu B, *et al.* A multi-task transfer learning method with dictionary learning. Knowledge-based Systems, 2020, 191: 105233. [doi: [10.1016/j.knsys.2019.105233](https://doi.org/10.1016/j.knsys.2019.105233)]
- 19 Zhang GQ, Porikli F, Sun HJ, *et al.* Cost-sensitive joint feature and dictionary learning for face recognition. Neurocomputing, 2020, 391: 177–188. [doi: [10.1016/j.neucom.2020.01.101](https://doi.org/10.1016/j.neucom.2020.01.101)]
- 20 Zhang GQ, Yang JC, Zheng YH, *et al.* Optimal discriminative feature and dictionary learning for image set classification. Information Sciences, 2021, 547: 498–513. [doi: [10.1016/j.ins.2020.08.066](https://doi.org/10.1016/j.ins.2020.08.066)]
- 21 Liu B, Xie HX, Xiao YS. Multi-task analysis discriminative dictionary learning for one-class learning. Knowledge-based Systems, 2021, 227: 107195. [doi: [10.1016/j.knsys.2021.107195](https://doi.org/10.1016/j.knsys.2021.107195)]
- 22 Rodríguez-Domínguez U, Dalmau O. Hierarchical discriminative deep dictionary learning. IEEE Access, 2020, 8: 142680–142690. [doi: [10.1109/ACCESS.2020.3008841](https://doi.org/10.1109/ACCESS.2020.3008841)]
- 23 Montazeri A, Shamsi M, Dianat R. MLK-SVD, the new approach in deep dictionary learning. The Visual Computer, 2021, 37(4): 707–715. [doi: [10.1007/s00371-020-01970-x](https://doi.org/10.1007/s00371-020-01970-x)]
- 24 Wu S, Wu AC, Zheng WS. Online deep transferable dictionary learning. Pattern Recognition, 2021, 118: 108007. [doi: [10.1016/j.patcog.2021.108007](https://doi.org/10.1016/j.patcog.2021.108007)]
- 25 Tang H, Liu H, Xiao W, *et al.* When dictionary learning meets deep learning: Deep dictionary learning and coding network for image recognition with limited data. IEEE Transactions on Neural Networks and Learning Systems, 2021, 32(5): 2129–2141. [doi: [10.1109/TNNLS.2020.2997289](https://doi.org/10.1109/TNNLS.2020.2997289)]
- 26 Lu SY, Lu ZH, Zhang YD. Pathological brain detection based on AlexNet and transfer learning. Journal of Computational Science, 2019, 30: 41–47. [doi: [10.1016/j.jocs.2018.11.008](https://doi.org/10.1016/j.jocs.2018.11.008)]

(校对责编: 牛欣悦)