



基于数据融合的智能医疗辅助诊断方法

张桃红 范素丽 郭徐徐 李倩倩

Intelligent medical assistant diagnosis method based on data fusion

ZHANG Tao-hong, FAN Su-li, GUO Xu-xu, LI Qian-qian

引用本文:

张桃红, 范素丽, 郭徐徐, 李倩倩. 基于数据融合的智能医疗辅助诊断方法[J]. *工程科学学报*, 2021, 43(9): 1197–1205. doi: 10.13374/j.issn2095-9389.2021.01.12.003

ZHANG Tao-hong, FAN Su-li, GUO Xu-xu, LI Qian-qian. Intelligent medical assistant diagnosis method based on data fusion[J]. *Chinese Journal of Engineering*, 2021, 43(9): 1197–1205. doi: 10.13374/j.issn2095-9389.2021.01.12.003

在线阅读 View online: <https://doi.org/10.13374/j.issn2095-9389.2021.01.12.003>

您可能感兴趣的其他文章

Articles you may be interested in

基于卷积神经网络的反无人机系统声音识别方法

Sound recognition method of an anti-UAV system based on a convolutional neural network

工程科学学报. 2020, 42(11): 1516 <https://doi.org/10.13374/j.issn2095-9389.2020.06.30.008>

深度神经网络模型压缩综述

A survey of model compression for deep neural networks

工程科学学报. 2019, 41(10): 1229 <https://doi.org/10.13374/j.issn2095-9389.2019.03.27.002>

基于集成神经网络的剩余寿命预测

Remaining useful life prediction based on an integrated neural network

工程科学学报. 2020, 42(10): 1372 <https://doi.org/10.13374/j.issn2095-9389.2019.10.10.005>

基于深度卷积神经网络的地磁导航方向适配性分析

Direction-matching-suitability analysis for geomagnetic navigation based on convolutional neural networks

工程科学学报. 2017, 39(10): 1584 <https://doi.org/10.13374/j.issn2095-9389.2017.10.018>

油气资源开发的大数据智能平台及应用分析

Big data intelligent platform and application analysis for oil and gas resource development

工程科学学报. 2021, 43(2): 179 <https://doi.org/10.13374/j.issn2095-9389.2020.07.21.001>

基于数据融合的智能医疗辅助诊断方法

张桃红^{1,2)}✉, 范素丽^{1,2)}, 郭徐徐^{1,2)}, 李倩倩^{1,2)}

1) 北京科技大学计算机通信与工程学院, 北京 100083 2) 材料领域知识工程北京市重点实验室, 北京 100083

✉通信作者, E-mail: zth_ustb@163.com

摘要 医生诊断需要结合临床症状、影像检查等各种数据, 基于此, 提出了一种可以进行数据融合的医疗辅助诊断方法. 将患者的影像信息(如 CT 图像)和数值数据(如临床诊断信息)相结合, 利用结合的信息自动预测患者的病情, 进而提出了基于深度学习的医疗辅助诊断模型. 模型以卷积神经网络为基础进行搭建, 图像和数值数据作为输入, 输出病人的患病情况. 该医疗辅助诊断方法能够利用更加全面的信息, 有助于提高自动诊断准确率、降低诊断误差; 另外, 仅使用提出的医疗辅助诊断模型就可以一次性处理多种类型的数据, 能够在一定程度上节省诊断时间. 在两个数据集上验证了所提出方法的有效性, 实验结果表明, 该方法是有效的, 它可以提高辅助诊断的准确性.

关键词 图像分类; 卷积神经网络; 特征融合; 医疗诊断; 深度学习

分类号 TG142.71

Intelligent medical assistant diagnosis method based on data fusion

ZHANG Tao-hong^{1,2)}✉, FAN Su-li^{1,2)}, GUO Xu-xu^{1,2)}, LI Qian-qian^{1,2)}

1) School of Computer and Communication Engineering, University of Science and Technology Beijing, Beijing 100083, China

2) Beijing Key Laboratory of Knowledge Engineering for Materials Science., Beijing 100083, China

✉ Corresponding author, E-mail: zth_ustb@163.com

ABSTRACT In the field of medicine, in order to diagnose a patient's condition more efficiently and conveniently, image classification has been widely leveraged. It is well established that when doctors diagnose a patient's condition, they not only observe the patient's image information (such as CT image) but also make final decisions incorporating the patient's clinical diagnostic information. However, current medical image classification only puts the image into a convolution neural network to obtain the diagnostic result and does not use the clinical diagnosis information. In intelligent auxiliary diagnosis, it is necessary to combine clinical symptoms with other imaging data for comprehensive diagnosis. This paper presented a new method of assistant diagnosis for the medical field. This method combined information from patients' imaging with numerical data (such as clinical diagnosis information) and used the combined information to automatically predict the patient's condition. Based on this method, a medical assistant diagnosis model based on deep learning was proposed. The model takes images and numerical data as input and outputs the patient's condition. Thus, this method is comprehensive and helps improve the accuracy of automatic diagnosis and reduce diagnostic error. Moreover, the proposed model can simultaneously process multiple types of data, thus saving diagnosis time. The effectiveness of the proposed method was verified in two groups of experiments designed in this paper. The first group of experiments shows that if the unrelated data are fused for classification, the proposed method cannot enhance the classification ability of the model, although it is able to predict multiple diseases at one time. The second group of experiments show that the proposed method could significantly improve classification results if the interrelated data are fused.

KEY WORDS image classification; convolution neural network; feature fusion; medical diagnosis; deep learning

收稿日期: 2021-01-12

基金项目: 中央高校基本科研业务费专项资金资助项目(FRF-GF-20-16B)

为了对图像分类, 研究人员提出了卷积神经网络 (CNN)。CNN 通常用于分析视觉图像, 它将图像的每个像素作为特征, 是一种类似于人工神经网络的深度学习模型或多层感知器。第一代 CNN 是由 LeCun 于 1998 年提出的 LeNet^[1]。LeNet 是为了解决手写数字的识别任务而提出的, 是早期 CNN 中最具代表性的结构之一。此后, CNN 最基本的体系结构被确定为卷积层、池化层和全连接层。在 2012 年, Krizhevsky 提出了 AlexNet^[2] 卷积神经网络, 并提出新的激活函数 (ReLU)、局部响应归一化 (LRN)、DropOut 和数据增强方法以提高网络的泛化能力。AlexNet 赢得了 ILSVRC2012^[3] 的第一名, 从此, CNN 受到了研究界的广泛关注。在 AlexNet 之后, 出现了许多优秀的 CNN 模型, 这些 CNN 主要有三个发展方向: (a) 更深: 网络层更深, 代表网络是 VggNet^[4]、ResNet^[5]; (b) 模块化: 模块化网络结构, 代表网络是 GoogleNet^[6]、Inceptionv2^[7]、Inceptionv3^[8] 和 Inceptionv4^[9]; (c) 更快: 适用于移动设备的轻量级网络模型, 代表网络是 SqueezeNet^[10]、MobileNet^[11]、ShuffleNet^[12]、MobileNetv2^[13]、ShuffleNetv2^[14] 和 MobileNetv3^[15]。

在医学领域, 为了更高效、更方便地诊断患者的病情, 图像分类已经得到了广泛关注。Li 等提出了一种基于注意的卷积神经网络 (AG-CNN) 用于青光眼检测^[16]。Yang 等提出了一种用于乳腺癌组织病理学图像分类的注意力引导卷积神经网络^[17]。Xu 等提出了一种用于检测肺结节性恶性肿瘤的卷积神经网络 (MSCS-DeepLN)^[18]。Mobiny 等提出了一种有效的胶囊网络变体 (Caps net) 作为 CNN 的替代^[19]。Zhou 等基于先验知识, 提取相应的特征对白内障进行分类^[20]。Wang 等提供一种创新的 3D 卷积网络, 用于自动乳腺超声检测癌症, 以加快检查速度, 同时获得较高的检测准确率^[21]。在 Liu 等的研究中, 提出了一种新的基于深度学习的 CAD 系统, 以特定任务的先验知识为指导, 用于超声图像中结节的自动检测和分类^[22]。Yao 等使用深度学习方法对宫颈细胞进行异常检测^[23]。众所周知, 当医生诊断患者的病情时, 不仅观察患者的图像信息 (如 CT 图像), 而且会结合患者的临床诊断信息做出最终决定。但是, 目前医学领域的图像分类问题仅将图像放入卷积神经网络中来获得诊断结果, 并没有使用到临床诊断信息。为了提高医学图像分类准确率, 一些研究引进特征融合技术。Zeng 等基于空间特征融合, 实现了颅内动脉瘤的自动诊断^[24]。Wang 等提出了一种基于注意

的特征聚合网络, 在一次检查中自动整合从多个图像中提取的特征, 利用结节的不同视图来提高超声图像中对恶性结节的识别性能^[25]。为了提高多模态 PET-CT 中互补信息的融合, Kumar 等提出了一种新的监督卷积神经网络来学习融合互补信息, 用于多模态医学图像分析^[26], 它通过融合两幅图像以实现更好的检测和分割。Joyseeree 等提出了一种基于 Riesz 和深度学习特征融合的新方法, 对肺部 CT 图像的病变类型进行识别^[27]。虽然, 这些基于特征融合的方法显示出了在医学图像分类、检测和分割任务上的优越性, 但是, 这些方法仍然没有利用到除医疗图像外的其他信息, 例如 Wang 等^[25] 只是融合多张图像特征, Joyseeree 等^[27] 是将卷积神经网络提取的图像特征与基于图像设计的特征融合。

而智能辅助诊断中经常需要结合临床症状与其他影像检查数据来进行综合诊断, 为此, 本文提出了一种医疗辅助诊断方法, 将图像特征与数值特征 (如临床诊断信息) 相结合, 利用融合的特征来确定患者的病情。本文提出的方法不仅可以充分利用患者的病例信息, 从而能够实现更准确的分类; 还可以同时处理不同类型的数据, 从而达到同时判断多种疾病的效果, 在一定程度上节省了医学分析的时间, 提高医学诊断的效率。该方法为基于深度学习的自动医疗辅助诊断方法提出了新的研究思路。

1 方法

为了降低医学领域中自动医学辅助诊断系统的误判概率, 本文提出了一种新的诊断方法。该方法将从图像中提取的特征与无法从图像中获取的其他特征相结合, 利用融合的数据来确定患者是否患有某种疾病。实验表明, 提出的方法可以减少医学分析的时间, 提高诊断准确率。基于提出的方法, 本文设计了一个数据融合深度学习模型, 该模型由数据融合层和分类层两部分组成, 模型的结构如图 1 所示。

1.1 数据融合层

数据融合层由特征提取模块和特征融合模块两部分组成。特征提取模块的作用是利用卷积神经网络提取图像特征, 将输入的图像转换为特征向量。特征提取模型可以基于任何 CNN 网络 (如 ResNet, VGG, ShuffleNet 等), 由于 ShuffleNetv2 属于轻量化网络, 能够达到速度和准确度的均衡, 因此本文把 ShuffleNetv2 作为特征提取模块的骨

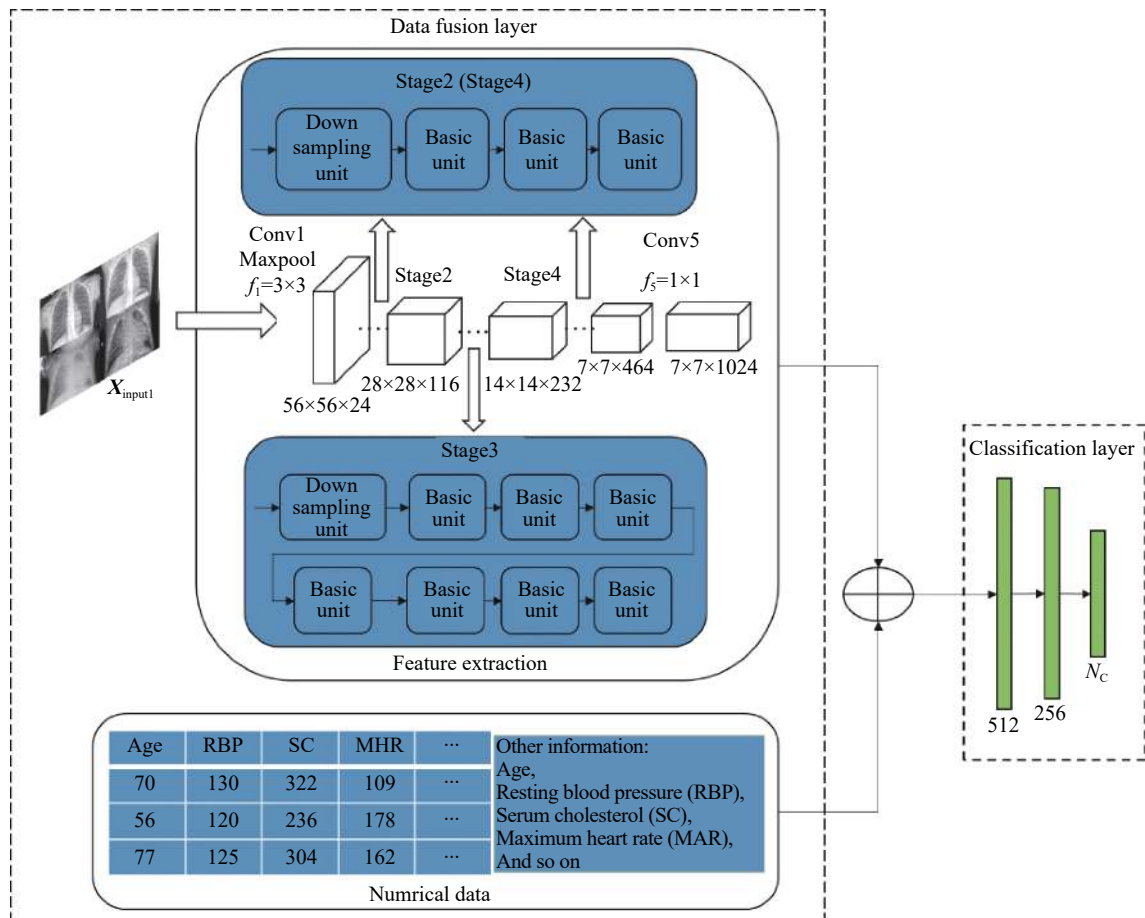


图1 基于提出的方法构建的模型结构

Fig.1 Diagram of the model structure based on the proposed method

干网络, 使用了 ShuffleNetv2 从输入层到平均池化层的网络层, 包括输入层、卷积层(conv1)、最大池化层(MaxPool)、stage2、stage3、stage4、卷积层(conv5), 下面对其网络结构进行具体解释. conv1 和 conv5 使用的卷积核 (f) 的大小和数目 (n) 分别是 $f_1 = 3 \times 3$, $n_1 = 24$, $f_5 = 1 \times 1$, $n_5 = 1024$. stage2、stage3 和 stage4 都是由下采样单元(Down sampling unit)和基本单元(Basic unit)连接而成, Down sampling unit 和 Basic unit 的结构如图 2 所示. 在图 2 中, 通道划分(Channel divide)的作用是将输入的特征图的通道数平均划分, 输出两个通道数相等的特征图; PWConv 是指使用 1×1 的卷积核进行卷积; DWConv 是指按通道对输入特征图进行分组, 每组包含一个通道, 卷积操作在每组内独立执行; 通道混洗(Channel shuffle)操作将 c 个通道分成 g 组, 形成大小为 $g \times (c/g)$ 的矩阵, 接下来将矩阵转置、平铺, 这样可以在通道之间对信息进行置乱和混合. Channel divide、PWConv 和 DWConv 的作用都是减少模型参数, 降低模型计算量. Channel shuffle 的作用是在不增加计算量的情况

下, 使通道充分融合, 进行不同分组的特征之间的信息流动, 以提高性能. 如果输入特征图的大小是 $w \times h \times c$, Basic unit 输出的大小是 $w \times h \times c$, Down sampling unit 输出的大小是 $(w/2) \times (h/2) \times (2c)$. 经过特征提取模块后, 每个输入图像都被转化一个为 1024 维的特征向量.

特征融合模块的作用是将图像中提取的特征 (X_A) 与无法从图像中获取的特征 (X_B) 进行融合. X_B (如临床诊断信息) 被组织成向量的形式输入到模型中, 然后特征融合模块将 X_A 和 X_B 融合, 形成一个新的特征向量 (X_f). 特征融合的实现方式如公式(1)所示, 特征融合模块是本文提出的方法的核心, 它将多种特征融合, 使融合后的特征更具代表性和更加全面, 这对分类任务更有益.

$$\begin{aligned} X_f &= \text{Fuse}(X_A, X_B) = \{x_{A1}, x_{A2}, \dots, x_{AM}, x_{B1}, x_{B2}, \dots, x_{BN}\} \\ X_A &= \{x_{Ai}\}, X_B = \{x_{Bj}\} \quad i = 1, 2, \dots, M; j = 1, 2, \dots, N \end{aligned} \quad (1)$$

1.2 分类层

分类层的作用是利用数据融合层输出的融合特征向量 (X_f), 输出分类结果. 分类层由两个全连

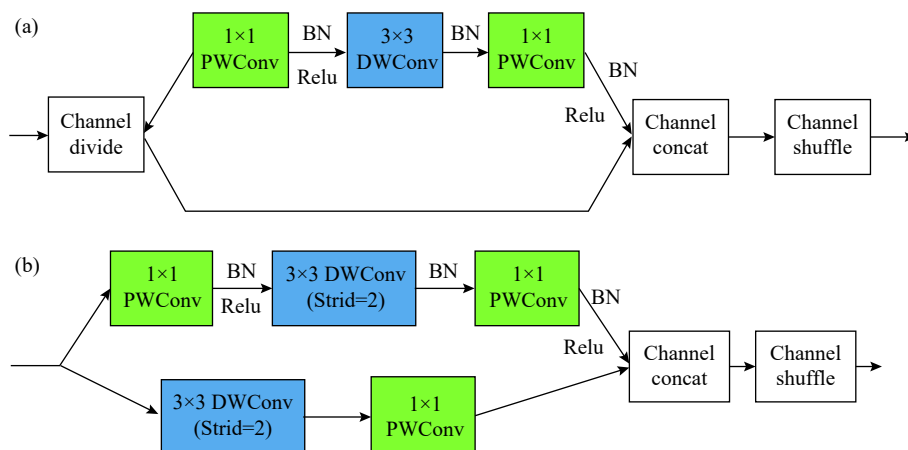


图 2 基本单元和下采样单元的结构。(a)基本单元的结构; (b)下采样单元的结构

Fig.2 Structure of the basic unit and down sampling unit: (a) structure of the basic unit; (b) structure of the down sampling unit

接层和一个输出层构成, 它们分别对应 512、256 和 N_C 个神经元 (N_C 表示类的总数量)。在输出层后面连接 Softmax 函数, 用于输出医疗诊断结果 (分类结果): $Y_{pre} = \{y_{pre_1}, y_{pre_2}, \dots, y_{pre_N_c}\}$, y_{pre_i} 表示输入数据的类是第 i 类的概率。样本标签用 $Y = \{y_1, y_2, \dots, y_{N_c}\}$ 表示, 如果输入的样本属于第 k 类, 那么 $y_k = 1, y_{lk} = 0$ 。训练模型时使用交叉熵损失函数来优化模型参数, 损失函数的计算方式如下所示:

$$\text{Loss} = - \sum_{i=1}^{N_c} y_i \log y_{pre_i} \quad (2)$$

2 实验

本文分别在两个数据集上进行了实验: 第一个实验预测肺炎和心脏病, 第二个实验预测新冠肺炎 (COVID-19)。在训练过程中, 模型参数通过批量迭代更新, 批量大小设置为 32, 初始学习率为 0.001, 使用的优化器是 Adam^[28]。实验细节分别在下面的实验 I 和实验 II 中展示。

2.1 实验 I

2.1.1 数据集

实验 I 中使用的数据包括一个肺部 CT 图像的数据集和一个心脏病的数值数据集, 它们都是从 Kaggle^[29] 收集的。我们将这两个公开数据集组合在一起, 生成一个同时包含图像和数值数据的数据集。肺部 CT 图像用于判断病人是否患有肺炎, 每张图像都被标注为是否有肺炎。数值数据用于判断病人是否患有心脏病, 每条数据包括 13 个属性: 年龄 (Age)、性别 (Sex)、胸痛类型 (CPT)、静息血压 (RBP)、血清胆固醇含量 (SC)、空腹血糖含量 (FBS)、静息心电图 (RER)、最大心率 (MHR)、运动





性心绞痛 (EIA)、运动引起的 ST 段下降 (ST)、ST 段峰值斜率 (SP)、透视镜看到的血管数 (NV)、地中海贫血 (Thal)。年龄分布在 20 岁到 80 岁之间; 性别包括男性 (0) 和女性 (1); CPT 分为 4 级, 分别是 0 (无疼痛), 1 (轻微疼痛), 2 (中级疼痛) 和 3 (重级疼痛); RER 有 3 个状态, 0 表示正常, 1 表示轻微异常, 2 表示严重异常; EIA 的类型为 0 (没有) 和 1 (有); 心电图 ST 段是指心电图上 QRS 波终点, 至 T 波开始前的一段水平线, ST 段改变往往提示有心肌缺血的可能。ST 段峰值斜率表示 ST 段抬高峰值处的斜率, 斜率为 0 时用 0 表示, 斜率位于 0 ~ 0.5 之间时用 1 表示, 斜率位于 0.5 ~ 1 之间时用 2 表示; Thal 的状态包括 1 (正常), 2 (固定缺陷) 和 3 (可逆缺陷)。本文将这两个数据集组合成了一个新的数据集, 它同时包含图像和数值数据, 命名为肺炎心脏病数据集 (PHD)。PHD 包含 606 个样本, 其中 424 个训练样本, 182 个测试样本。PHD 共包含 4 种类型: 同时患有肺炎和心脏病 (PH), 患有肺炎但未患心脏病 (PNH), 未患肺炎但患有心脏病 (NPH), 既没有患肺炎又没有患心脏病 (NPNH)。每一个样本包含一张图像和 13 个数值属性, 表 1 展示了数据集中的 4 种样本。

2.1.2 结果和讨论

如第 1 部分所述, 使用轻量化卷积神经网络 ShuffleNetv2 提取图像特征, 分类层使用两个全连接层和一个输出层, 两个全连接层分别包含 512 和 256 个神经元, 输出层包含 4 个神经元 (PHD 包含 4 种类型的疾病)。为了验证提出的融合方法的可行性, 进行了 3 组实验进行对比: 第一组实验 (ShuffleNetv2) 利用图 1 中的特征提取模块和分类层对肺部 CT 图像进行分类, 预测病人是否患有肺炎; 第二组实验 (DNN) 利用分类层对心脏病数值

表1 PHD 中四种类型的样本

Table 1 Four types of samples in PHD

Class	Image	Age	Sex	CPT	RBP/kPa	SC/(mg·dL ⁻¹)	FBS/(mg·dL ⁻¹)	RER	MHR/(times·min ⁻¹)	EIA	ST/mV	SP	NV	Thal
PH		66	0	3	20	226	0	1	114	0	2.6	0	0	2
PNH		54	1	0	14.7	239	0	1	126	1	2.8	1	1	3
NPH		65	0	2	20.7	269	0	1	148	0	0.8	2	0	2
NPNH		70	1	0	17.3	322	0	0	109	0	2.4	1	3	2

数据进行分类, 预测病人是否患有心脏病; 第三组实验(Fusion)利用提出的方法对肺部 CT 图像和心脏病数值数据进行分类, 预测病人是否患有肺炎和心脏病. 3 组实验的预测结果分别展示在表 2、表 3 和表 4, 各项评价指标展示在表 5. 在表 2~4 中, 横向表头表示样本的真实类别, 纵向表头表示样本的预测类别, 例表 2 中第 2 行第 4 列的 90 表示测试时有 90 个样本被预测为 No pneumonia; 第 4 行第 2 列的 91 表示, 测试集中共有 91 个样本的类型是 No pneumonia. 在表 5 中, TP 表示正样本被预测为正样本的数量, FP 表示负样本被预测为正样本的数量, FN 表示负样本被预测为负样本的数量, Precision 表示预测为正的样本中有多少是真正的正样本, Recall 表示样本中的正例有多少被预测正确了. F1-score 是对 Precision 和 Recall 的综合考虑, $F1\text{-score}=2 * Precision * Recall / (Precision+Recall)$. 图 3 展示了第三组实验结果(使用提出的方法)随着迭代次数的变动, 在训练集和验证集上的预测准确率和损失的变动. 从图中可以看出, 迭代 600 次之后, 模型开始趋于稳定.

表2 在 PHD 数据集上仅通过图像数据学习的预测结果

Table 2 Prediction results learned only from image data in PHD dataset

Prediction	Label		
	No pneumonia	Pneumonia	All
No pneumonia	79	11	90
Pneumonia	12	80	92
All	91	91	182

表3 在 PHD 数据集上仅通过结构化的数值数据学习的预测结果

Table 3 Prediction results learned only through structured numerical data

Prediction	Label		
	No pneumonia	Pneumonia	All
No pneumonia	72	16	88
Pneumonia	11	83	94
All	83	99	182

表4 本文方法在 PHD 数据集上的预测结果

Table 4 Predictive results of proposed method in PHD dataset

Prediction	Label				
	NPNH	NPH	PNH	PH	All
NPNH	33	12	5	2	52
NPH	8	29	0	0	39
PNH	2	2	24	10	38
PH	2	3	9	39	53
All	45	46	38	53	182

从表 5 可以看出, 仅使用肺炎图像的预测患者是否患有肺炎的平均准确率是 87.4%, 仅使用心脏病数值数据来预测患者是否患有心脏病的平均准确率是 85.2%. 将肺炎图像与心脏病数值数据相结合后, 可同时预测患者是否患有肺炎与心脏病. 根据实验结果可以看出, 组合后分类变成了一个四分类问题, 它获得了 68.7% 的较低的平均分类准确率. 根据表 4 的预测结果可知, 使用提出的方法, 预测肺炎的准确率 $(33+29+12+8+24+10+9+39)/182 \times 100\% = 90.1\%$, 预测心脏病的准确率等于 $(33+24+5+2+29+39+2+3)/182 \times 100\% = 75.3\%$. 比较结果表明, 提出的方法没有增加预测的准确率, 这与肺炎和心脏病数据之间的关系是我们人为构造的有关, 因为肺炎和心脏病这两种疾病没有内在联系, 即患有肺炎的患者并不会更容易患心脏病, 同理患有心脏病的患者也不会更容易患肺炎. 因此, 即使使用提出的方法, 也不会提高整体的分类准确率, 这与我们的设计初衷是一致的, 即只有图像数据和数值数据具有相辅相成、相互促进的关系时, 才有利于模型的分类. 虽然没有提高分类准确率, 但是提出的方法能够一次性处理多种类型的数据, 同时预测多种疾病, 在一定程度上减少了诊断时间. 因此, 实验 I 证明了, 如果将没有联系的数据进行融合用于分类时, 并不能增强模型分类的能力. 但是提出的方法具有同时预测多种疾病的能

表 5 在 PHD 数据集上三组实验的准确率和其他评价指标

Table 5 Accuracy and other evaluation indicators of three groups of experiments in PHD dataset

Model	Class	TP	FP	FN	Precision	Recall	F1-score	Accuracy
Fusion method	NHPH	33	19	12	0.635	0.733	0.680	0.687
	NPH	29	10	17	0.744	0.630	0.682	
	PNH	24	14	14	0.632	0.632	0.632	
	PH	39	14	14	0.736	0.736	0.736	
ShuffleNetv2(Only image data)	No pneumonia	79	11	12	0.878	0.868	0.873	0.874
	Pneumonia	80	12	11	0.870	0.879	0.874	
DNN(Only structured data)	No heart disease	72	16	11	0.818	0.867	0.842	0.852
	Heart disease	83	11	16	0.883	0.838	0.860	

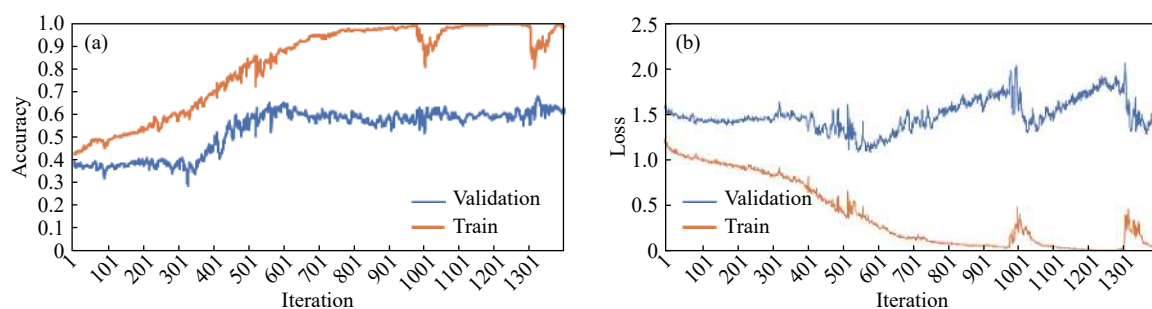


图 3 训练过程中的预测准确率和损失的变动。(a)准确率的变动;(b)损失的变动

Fig.3 Changes in predictive accuracy and loss during training: (a) changes in accuracy; (b) changes in the loss

力, 因此该方法可以用于具有相互联系的两种疾病的预测, 如肥胖和高血压。

2.2 实验 II

实验 II 用以判定一个人是否患有 COVID-19, 即本实验是一个二分类的任务。我们知道, 仅根据临床症状或者仅利用 CT 图像无法确定一个人是否患有 COVID-19。在没有进行核酸检测的情况下, 医生通常需要将病人的临床症状与 CT 图像结合起来, 以确定病人是否患有 COVID-19。因此, 利用本文提出的新的医疗辅助诊断方法与仅使用 CT 图像数据和仅使用结构化数值数据预测 COVID-19 的方法进行比较, 以证明提出的新方法的有效性。

2.2.1 数据集

实验 II 使用的数据集是从几家医院的公共网站^[30]收集的。该数据集(命名为 COVID)由 460 个样本组成, 其中训练集中包含 161 个没有患 COVID-19 的样本和 161 个患有 COVID-19 的样本, 测试集中包含 69 个没有患 COVID-19 的样本和 69 个患有 COVID-19 的样本。数据集中的每个样本包含患者的肺部 CT 图像和 10 个数值属性: 性别(Sex)、年龄(Age)、发热程度(F)、呼吸困难程度(D)、咳嗽程度(C)、血氧饱和度(OS)、既往病

史(PMH)、是否与阳性患者接触过(CP)、正常白细胞(WBC)和其他症状(Other)。性别包括男性(1)和女性(2); 年龄从 1 岁到 100 岁不等; F 由 3 个值表示, 1 代表不发热(<37.3 °C), 2 代表轻度发热(37.3 ~ 39 °C), 3 代表高度发热(>39 °C); D 由 4 个值表示, 1 表示正常, 2 表示轻度困难, 3 表示中度困难, 4 表示重度困难; C 由 3 个值表示, 1 表示没有咳嗽, 2 表示轻度咳嗽, 3 表示剧烈咳嗽; PMH 包含 2 种情况: 1 表示没有, 2 表示有; CP 有 3 种情况: 1 表示没有接触过, 2 表示接触过, 3 表示来自疫情区; WBC 有 3 种状态: 1 表示下降, 2 表示正常, 3 表示上升; 其他症状包括腹泻, 吞咽困难, 胸闷等, 用 1 表示没有其症状, 2 表示有其症状。由于数值数据的某些属性值是缺失的, 因此本文将缺失值用 0 代表。没有患 COVID-19 的患者的数值数据是在专业指导^[31]下构建的, 构建的数据集共包含两类: 没有患 COVID-19(NonCOVID)和患 COVID-19(COVID)。图 4 展示了数据集中的两种样本。

2.2.2 结果和讨论

实验 II 同实验 I 一样, 使用如图 1 所示的模型结构。与实验 I 相似, 设计了三组实验预测患者是否患有 COVID-19。第一组实验(ShuffleNetv2)利


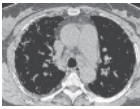
	Image	Sex	Age	F	D	C	OS	PMH	CP	WBC	Other
(a)		2	52	3	1	1	0.97	0	0	1	0
(b)		1	77	1	0	1	0.89	0	2	2	1

图4 COVID数据集中两种类型的样本。(a)未患COVID-19的样本;(b)患有COVID-19的样本

Fig.4 Two types of samples in COVID: (a) samples without COVID-19; (b) samples with COVID-19

用图1中的特征提取模块和分类层对肺部CT图像进行分类;第二组实验(DNN)利用分类层对数值数据进行分类;第三组实验(Fusion)利用提出的方法对肺部CT图像和数值数据进行分类。三组实验的预测结果分别展示在表6、表7和表8,各项评价指标展示在表9。图5展示了第三组实验随着迭代次数变动的训练集和验证集的预测准确率和损失的变动。在对138个样本进行分类时,还分别计算了单独使用图像进行分类和使用Fusion method的分类方法进行分类的时间。如表10所示,Fusion method的分类时间只比仅使用图像进行分类的时间增加了0.02 s。如表9所示,仅使用图像数据的预测准确率是75.4%,仅使用数值数据的预测准确率是88.4%,使用两种数据的预测准确率是94.2%,可见Fusion method能够提高患病的预测准确率。相比于仅使用图像分类,Fusion method的准确率提高了18.8%,但是花费的时间仅增加0.02 s。再观察指标F1-score,仅使用图像分类时,NonCOVID的F1-score是0.764,COVID的F1-score是0.742;仅使用数值数据分类时,NonCOVID的F1-score是0.869,COVID的F1-score是0.896;使用提出的融合方法分类时,NonCOVID和COVID的F1-score都是0.942。总体上来看,相比于前两组实验,提出的方法不仅显著提高了每一类的F1-score,而且显著提高了整体的分类准确率。之所以能够显著提高分类结果,是因为提出的方法将图像和数值数据融合,两类数据的特征相辅相成,互相增强,从而达到更好的分类结果。此外,为了证明本文提出的方法(Fusion method)对医疗辅助诊断的有效性,将本文提出的方法和现有的图像分类方法(ResNet50, VGG16, ShuffleNetv2, AlexNet)进行了对比。在实验环境相同的情况下,对测试集上138个样本进行了测试,实验结果如表11所示。可以从表中看出,对138个样本进行分类时,Fusion method达到了94.2%的准确率,超过了准确率为79.7%的ResNet50、准确率为77.5%的VGG16、准

确率为75.4%的ShuffleNetv2和准确率为73.2%的AlexNet。由实验结果可知,本文提出的将图像和数值数据进行融合的方法超越了单独用图像进行分类的方法(ResNet50、VGG16等),再一次证明了Fusion method更有益于医疗辅助诊断任务。

表6 在COVID数据集上仅通过图像数据学习的预测结果

Table 6 Prediction results learned only from image data in COVID dataset

Prediction	Label		
	NonCOVID	COVID	All
NonCOVID	55	20	75
COVID	14	49	63
All	69	69	138

表7 在COVID数据集上仅通过结构化的数值数据学习的预测结果

Table 7 Prediction results learned only by structured numerical data in COVID dataset

Prediction	Label		
	NonCOVID	COVID	All
NonCOVID	53	0	53
COVID	16	69	85
All	69	69	138

表8 本文方法在COVID数据集上的预测结果

Table 8 Predictive results of proposed method in COVID dataset

Prediction	Label		
	NonCOVID	COVID	All
NonCOVID	65	4	69
COVID	4	65	69
All	69	69	138

3 结论

本文为医疗领域提供了一种基于数据融合的医疗辅助诊断方法,在一定程度上提高了医疗诊断的准确率和效率。在两个数据集上进行了实验,实验I将肺炎CT图像和心脏病数值数据结合,使

表 9 在 COVID 数据集上三组实验的准确度和其他评价指标

Table 9 Accuracy and other evaluation indicators of three groups of experiments in COVID dataset

Model	Class	TP	FP	FN	Precision	Recall	F1-score	Accuracy
Fusion method	NonCOVID	65	4	4	0.942	0.942	0.942	0.942
	COVID	65	4	4	0.942	0.942	0.942	
ShuffleNetv2(Only image data)	NonCOVID	55	20	14	0.733	0.797	0.764	0.754
	COVID	49	14	20	0.778	0.710	0.742	
DNN(Only structured data)	NonCOVID	53	0	16	1.00	0.768	0.869	0.884
	COVID	69	16	0	0.812	1.00	0.896	

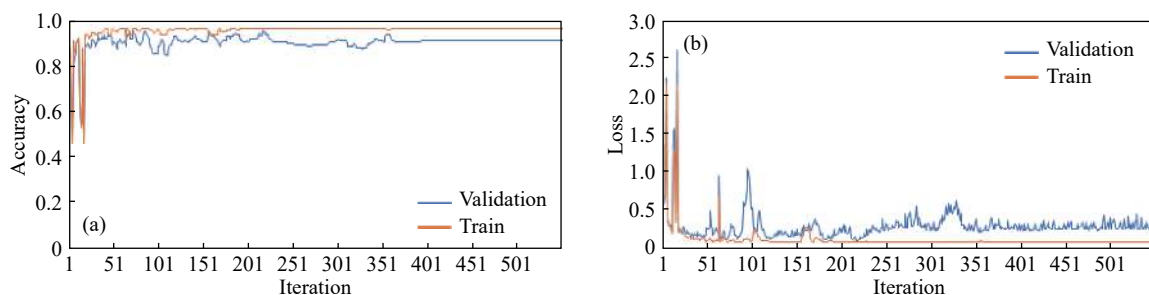


图 5 训练过程中的预测准确率和损失的变动. (a)准确率的变动; (b)损失的变动

Fig.5 Changes in predictive accuracy and loss during training: (a) changes in accuracy; (b) changes in the loss

表 10 本文方法和仅通过图像学习对 138 个样本进行分类的时间

Table 10 Time required to classify 138 samples using proposed method and using only image data

Model	Proposed method	Image only
Time	3.58	3.56

用结合的数据进行分类, 实现了同时预测多种疾病的效果, 节省了医学分析的时间. 实验 II 为了判断患者是否患有新冠肺炎, 以肺部 CT 图像和临床诊断信息作为诊断依据, 与仅使用肺部 CT 图像相比, 两种方法的预测时间相差极小, 但是提出的方

表 11 Fusion method、ResNet50、VGG16、ShuffleNetv2 和 AlexNet 的准确度和其他评价指标

Table 11 Accuracy and other evaluation indicators of Fusion method, ResNet50, VGG16, ShuffleNetv2 and AlexNet

Model	Class	TP	FP	FN	Precision	Recall	F1-score	Accuracy
Fusion method	NonCOVID	65	4	4	0.942	0.942	0.942	0.942
	COVID	65	4	4	0.942	0.942	0.942	
ResNet50	NonCOVID	56	15	13	0.789	0.812	0.800	0.797
	COVID	54	13	15	0.806	0.783	0.794	
VGG16	NonCOVID	54	16	15	0.771	0.783	0.777	0.775
	COVID	53	15	16	0.779	0.768	0.774	
ShuffleNetv2	NonCOVID	55	20	14	0.733	0.797	0.764	0.754
	COVID	49	14	20	0.778	0.710	0.742	
AlexNet	NonCOVID	50	18	19	0.735	0.725	0.730	0.732
	COVID	51	19	18	0.728	0.739	0.734	

法的预测结果准确地多. 综合来看, 本文提出的方法为提高医学诊断的准确性、节省医学诊断时间提供了一个很好的思路.

参 考 文 献

[1] Lecun Y, Bottou L, Bengio Y, et al. Gradient-based learning

applied to document recognition. *Proc IEEE*, 1998, 86(11): 2278

[2] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks. *Commun ACM*, 2017, 60(6): 84

[3] Deng J, Dong W, Socher R, et al. ImageNet: A large-scale hierarchical image database // 2009 *IEEE Conference on Computer*

- Vision and Pattern Recognition*. Miami, 2009: 248
- [4] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition [J/OL]. *ArXiv Preprint* (2014-09-04) [2021-01-12]. <https://arxiv.org/abs/1409.1556>
- [5] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition // 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Las Vegas, 2016: 770
- [6] Szegedy C, Liu W, Jia Y Q, et al. Going deeper with convolutions // 2015 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Boston, 2015: 1
- [7] Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift [J/OL]. *ArXiv Preprint* (2015-02-11) [2021-01-12]. <https://arxiv.org/abs/1502.03167>
- [8] Szegedy C, Vanhoucke V, Ioffe S, et al. Rethinking the inception architecture for computer vision // 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Las Vegas, 2016: 2818
- [9] Szegedy C, Ioffe S, Vanhoucke V, et al. Inception-v4, inception-resnet and the impact of residual connections on learning [J/OL]. *ArXiv Preprint* (2016-02-24) [2021-01-12]. <https://arxiv.org/abs/1602.07261>
- [10] Iandola F N, Han S, Moskewicz M W, et al. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5 MB model size [J/OL]. *ArXiv Preprint* (2016-02-24) [2021-01-12]. <https://arxiv.org/abs/1602.07360>
- [11] Howard A G, Zhu M, Chen B, et al. Mobilenets: Efficient convolutional neural networks for mobile vision applications [J/OL]. *ArXiv Preprint* (2016-02-24) [2021-01-12]. <https://arxiv.org/abs/1704.04861v1>
- [12] Zhang X Y, Zhou X Y, Lin M X, et al. ShuffleNet: an extremely efficient convolutional neural network for mobile devices // 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City, 2018: 6848
- [13] Sandler M, Howard A, Zhu M L, et al. MobileNetV2: inverted residuals and linear bottlenecks // 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City, 2018: 4510
- [14] Ma N N, Zhang X Y, Zheng H T, et al. ShuffleNet V2: practical guidelines for efficient CNN architecture design // 2018 *European Conference on Computer Vision (ECCV)*. Munich, 2018: 122
- [15] Howard A, Sandler M, Chen B, et al. Searching for MobileNetV3 // 2019 *IEEE/CVF International Conference on Computer Vision (ICCV)*. Seoul, 2019: 1314
- [16] Li L, Xu M, Liu H R, et al. A large-scale database and a CNN model for attention-based glaucoma detection. *IEEE Trans Med Imaging*, 2020, 39(2): 413
- [17] Yang H, Kim J Y, Kim H, et al. Guided soft attention network for classification of breast cancer histopathology images. *IEEE Trans Med Imaging*, 2020, 39(5): 1306
- [18] Xu X Y, Wang C D, Guo J X, et al. MSCS-DeepLN: Evaluating lung nodule malignancy using multi-scale cost-sensitive neural networks. *Med Image Anal*, 2020, 65: 101772
- [19] Mobiny A, Lu H Y, Nguyen H V, et al. Automated classification of apoptosis in phase contrast microscopy using capsule network. *IEEE Trans Med Imaging*, 2020, 39(1): 1
- [20] Zhou Y, Li G Q, Li H Q. Automatic cataract classification using deep neural network with discrete state transition. *IEEE Trans Med Imaging*, 2020, 39(2): 436
- [21] Wang Y, Wang N, Xu M, et al. Deeply-supervised networks with threshold loss for cancer detection in automated breast ultrasound. *IEEE Trans Med Imaging*, 2020, 39(4): 866
- [22] Liu T J, Guo Q Q, Lian C F, et al. Automated detection and classification of thyroid nodules in ultrasound images using clinical-knowledge-guided convolutional neural networks. *Med Image Anal*, 2019, 58: 101555
- [23] Yao C, Zhao J Z, Ma B Y, et al. Fast detection method for cervical cancer abnormal cells based on deep learning. *Chin J Eng*, <https://doi.org/10.13374/j.issn2095-9389.2021.01.12.001>
(姚超, 赵基准, 马博渊, 等. 基于深度学习的宫颈癌异常细胞快速检测方法. 工程科学学报, <https://doi.org/10.13374/j.issn2095-9389.2021.01.12.001>)
- [24] Zeng Y W, Liu X K, Xiao N, et al. Automatic diagnosis based on spatial information fusion feature for intracranial aneurysm. *IEEE Trans Med Imaging*, 2020, 39(5): 1448
- [25] Wang L T, Zhang L, Zhu M J, et al. Automatic diagnosis for thyroid nodules in ultrasound images by deep neural networks. *Med Image Anal*, 2020, 61: 101665
- [26] Kumar A, Fulham M, Feng D G, et al. Co-learning feature fusion maps from PET-CT images of lung cancer. *IEEE Trans Med Imaging*, 2020, 39(1): 204
- [27] Joyseeree R, Otálora S, Müller H, et al. Fusing learned representations from Riesz filters and deep CNN for lung tissue classification. *Med Image Anal*, 2019, 56: 172
- [28] Kingma D, Ba J. Adam: A method for stochastic optimization [J/OL]. *ArXiv Preprint* (2014-12-22) [2021-01-12]. <https://arxiv.org/abs/1412.6980>
- [29] Bio. Heart Disease UCI [J/OL]. *Kaggle* (2018-06-25) [2021-01-12]. <https://www.kaggle.com/ronitf/heart-disease-uci>
- [30] Società Italiana di Radiologia Medica e Interventistica. Covid-19 Database [J/OL]. *Database Online* (2020-03-18) [2021-01-12]. <https://www.sirm.org/category/senza-categoria/covid-19>
- [31] Zhao J, Zhang Y, He X, et al. COVID-CT-Dataset: A CT scan dataset about COVID-19 [J/OL]. *ArXiv Preprint* (2020-03-30) [2021-01-12]. <https://github.com/UCSD-AI4H/COVID-CT>