

# 基于深度学习的阿尔兹海默症多模态辅助诊断研究

Multi-Modal Data Computer-Aided Diagnosis of AD Based on Deep Learning Model

林雪峰 李 炜 (华中科技大学人工智能与自动化学院,湖北 武汉 430074)

**摘要:**阿尔兹海默症(Alzheimer’s Disease,AD)是一种在老年人群中常见的痴呆疾病,由于病程不可逆且无法治愈,常会对病人的生活质量产生极大影响,因此尽早诊断病情并对病程加以干预是唯一有效的手段。由于良好的实验效果,深度学习模型在医学图像领域受到了越来越多研究者的关注,但深度学习方法常需要较大的数据量作为支撑,而医学图像由于设备成本以及病例数量的限制,常存在着数据量不足的问题,因而在某些情况下会出现过拟合的问题。提出一种参数高效的深度学习模型,引入了可分离卷积、全局平均池化、残差结构,使得模型参数量成倍地减少,同时引入多模态数据,增大了输入样本的信息量,以求减少过拟合问题。最后,通过对照试验,验证了该文所提出模型的优越性。

**关键词:**阿尔兹海默症,深度学习,多模态数据,辅助诊断

**Abstract:**Alzheimer’s Disease (AD) is a common dementia disease in the geriatric population which is irreversible and cannot be cured,and it often affects the life quality of the patients. Early detection and early treatment may be the only effective way for helping the patients of AD.Due to excellent performance,deep learning models have attracted more and more researchers’ attention in the field of medical images.However,deep learning methods often require a large amount of data for training model,and because of the limitations of the patient number and the high equipment costs,the data of medical images are always insufficient,and overfitting problems may occur in some cases.To avoid the overfitting problems,this paper attempts to propose a parameter-efficient deep learning model.It introduces separable convolution, global average pooling, and residual structure into the model,and the number of model parameters are reduced by several times.At the same time, with using multimodal data, the information of the input is increased.Finally,the superiority of the model proposed in this paper is verified through controlled experiments.

**Keywords:**Alzheimer diseases,deep learning,multi-modal data,computer-aided diagnosis

在全球范围内,目前有超过3000万人患有阿尔茨海默症(Alzheimer’s disease,AD),由于预期寿命的延长和一般卫生和保健状况的改善,预计这一数字到2050年将增长为现在的两倍<sup>[1]</sup>。AD是痴呆的一种形式,其特征 $\beta$ -淀粉样蛋白肽沉积以及异常的tau积累和磷酸化,最终导致神经元死亡和神经突触丧失<sup>[2]</sup>。与AD有关的神经病变遵循特定的模式,初期从皮质下区域开始病变,并在疾病的晚期扩散至皮质层<sup>[3]</sup>。AD的最常见的经典临床表现形式是发作性记忆不足,其次是视觉空间损伤,出现时空定向问题,最后表现为显著的痴呆。

在针对AD的辅助诊断研究中,越来越多的研究人员对使用基于磁共振成像(MRI)的生物标记物来进行辅助诊断产生了兴趣。因为MRI相比电离辐射正电子发射断层扫描(PET),更加便宜,因而数据量更多。相比脑脊液(CSF)生物标记物,其侵入性更小。基于MRI的切片还可以提供有关同一扫描过程中大脑结构和功能的多模式信息,这在许多临床环境中通常是有用的。由于以上原因,基于MRI的图像来进行AD辅助诊断的研究近年来变得愈发热门。

深度学习在医学图像以及辅助诊断领域,因为其良好的效果,已经得到了广泛的应用<sup>[4]</sup>。其优秀的表征能力常依赖于大量的神经网络参数。在某些情况下,可能会因此产生过拟合现象。在计算机视觉领域,常通过增大训练数据量来解决过拟合问题。但医学图像领域相比计算机视觉领域,在可用的数据量上相差甚远。所以我们选择从模型入手,尝试构建一种参数高效的神经网络。在网络基础层上,我们引入了三维可分离卷积<sup>[5]</sup>、全局池化<sup>[6]</sup>,在网络结构上引入了残差结构<sup>[7]</sup>,使得所设计的神经网络

在保证实验效果的前提下,模型参数量成倍地减少。同时我们引入了临床诊断数据等多模态信息,在某种程度上相当于增大训练数据的信息量。从多方面减少了模型过拟合的可能性,大大提升了模型的诊断效果。

## 1 实验数据

本文所使用的所有数据均来自阿尔茨海默症影像学计划(Alzheimer’s Disease Neuroimaging Initiative,ADNI)。该计划由435名男性和350名女性组成,年龄在55至91岁之间。大多数受试者人种为白人(>94%)以及非西班牙裔(99.98%)。数据样本按照痴呆的程度被分为阿尔兹海默症(Alzheimer’s disease Group,AD)、阿尔兹海默症的前一阶段轻微认知损伤症(Mild Cognitive Impairment,MCI)、正常对照(Normal Control,NC)三组类型。表1总结了我们使用的所有数据。

表1 样本信息统计

	样本个数	年龄	男性/女性
AD	192	75.6±7	103/81
MCI	409	73.0±7	240/169
NC	184	74.6±6	92/100

对于所有参与者,我们使用了磁化快速梯度回波(MPRAGE)T1加权图像(结构项MRI)以及以下数据:人口统计数据(年龄、性别、种族和种族类别、教育程度),神经心理学认知评估测试如痴呆症评分量表(CDRSB),阿尔茨海默氏病评估量表(ADAS11、ADAS13),雷伊听觉语言学习测验(RAVLT)中的情景记忆评估以及APOe4基因分型。

在分类之前,所有T1加权(T1w)图像均已标准化处理到一

个公共空间(即 T1 模板)。具体的来说,我们将所有 T1w 图像(在 N4 偏置场校正之后)彼此非线性配准并进行迭代平均(即在每次迭代结束时重新创建组平均值)。该过程基于对称微分映射,总共进行了五次迭代。

2 实验方法

本文所设计的神经网络结构如图 1 所示。

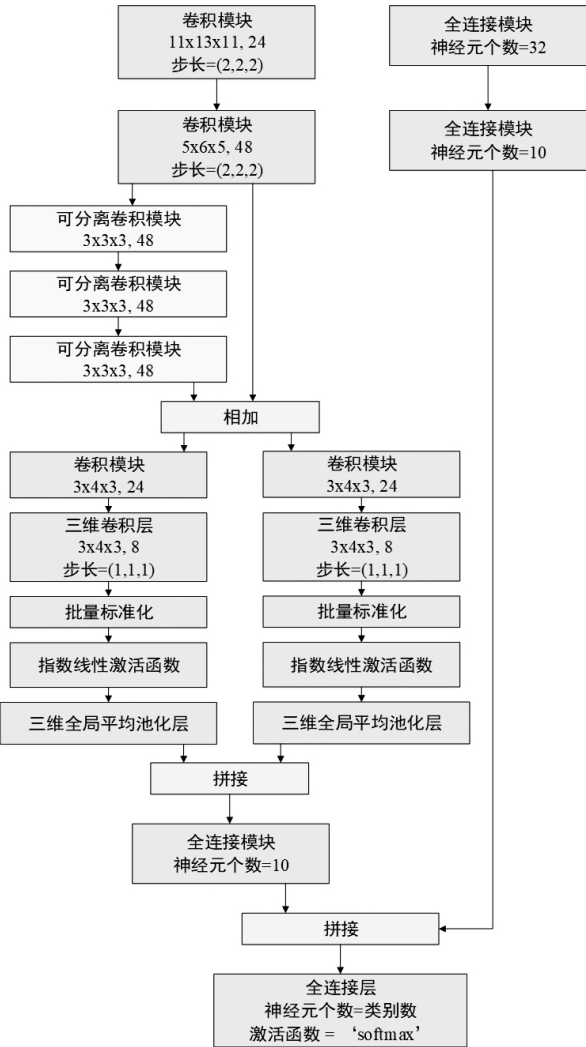


图 1 神经网络结构图

如图 1 和图 2 所示,我们所设计的分类网络可以划分成若干特征提取子模块,一些特殊结构以及部分基础网络层。包括的三维卷积模块、三维可分离卷积模块、三维全局平均池化层、残差结构、拼接层、全连接层。在设计本文所用的网络之前,我们大量阅读了前人所提出的一些经典的网络结构,比如微软亚洲研究院的

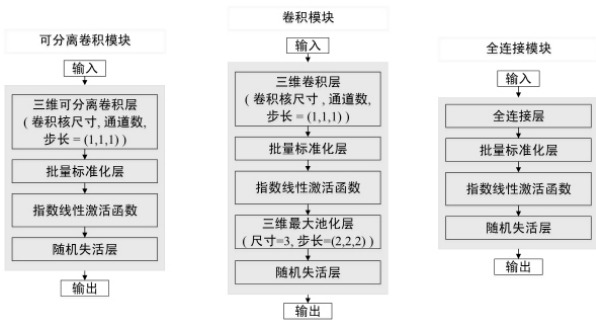


图 2 模块结构图

Kaiming He,Jian Sun 等人提出的 ResNet<sup>[7]</sup>,谷歌的 Chollet、Velickovic 等研究人员所提出的 Inception<sup>[8]</sup>和 Xception<sup>[5]</sup>以及由新加坡国立大学的 Lin、Chen 等人提出的 Network in Network<sup>[6]</sup>网络,这些都给了我们很大的启发。本文提出的网络中的三维可分离卷积模块、残差结构、三维全局平均池化层的设计灵感分别来自于以上文章,同时我们将二维结构的可分离卷积与全局平均池化改进为三维结构,使其更适合处理三维 MRI 图像。

分类网络的输入包括两个部分,一部分是经过预处理之后的 MRI 图像数据,标准维度为(121,145,121),另一部分是以人口统计数据(包含年龄、性别、种族和种族类别、教育程度等数据),神经心理学认知评估测试得分如痴呆症评分量表(CDRSB),简易智力状态检查量表(MMSE),阿尔茨海默氏病评估量表(ADAS11, ADAS13),雷伊听觉语言学习测验(RAVLT)中的情景记忆评估以及 APOe4 基因分型数据所组成的临床诊断数据。临床诊断数据在形式上为一维向量,一共包含 12 个特征值。由于阿尔兹海默症的致病原因非常复杂,且其早期病程(MCI)往往没有显著相关的病理学表现,所以越来越多的研究采用多模态数据进行 AD 的辅助诊断研究,吸纳多维度,不同观测方式得到的数据,以便获得更准确的诊断结果。

无论 MRI 图像数据还是临床诊断数据在输入神经网络之前都会先进行标准化操作,以消除量纲的影响,同时这样处理也更利于梯度下降寻优。经过标准化的数据,均值为 0,标准差为 1。其中临床数据的标准化,是针对每个单独特征因子进行的。

2.1 全连接层、全连接模块、全局平均池化

在分类网络中,全连接层常用在池化层或者卷积层之后,能起到整合局部信息的作用。全连接层每层包含若干个结点,每个结点都与上一层的所有结点相连。一般来说,每个结点还包含一个激活函数,对应的输入经过激活函数,完成输入输出的映射变换。这种变换可以是线性的也可以是非线性的,取决于激活函数的是线性函数还是非线性函数。激活函数的一般形式如表达式(1)所示。其中 x 是输入,y 是输出。 $\omega$  表示权重,b 表示偏差。

$$y=f(\omega * x+b)$$
(1)

全连接模块由四个基本层组成,包括使用线性激活函数的全连接层、批量标准化层(Batch normalization,BN)<sup>[9]</sup>、指数线性激活函数层(exponential linear unit,ELU)<sup>[11]</sup>、随机失活层(Dropout)层<sup>[10]</sup>。指数线性激活层的表达式如式(2)所示,当输入 x 大于等于 0 时, $f(x)=x$ ,当输入 x 小于 0 时, $f(x)=a(e^x-1)$ 。批量标准化层的作用主要是对全连接层的输出进行标准化处理,能够加快模型的收敛速度。而随机失活层(Dropout)层通过对输入进行随机失活处理,能够提升模型的泛化能力以及鲁棒性。

$$f(x)=\begin{cases} x, & x \geq 0 \\ a(e^x - 1), & x < 0 \end{cases}$$
(2)

全局平均池化的概念最早出自于 Network In Network,被公认为是一种可以替代全连接层的新方法。在设计模型时,可以将其放置在卷积模块之后,用来代替网络模型最后面几层中的全连接层。其实现原理是将卷积模块输出的特征图根据通道进行一个全局平均池化操作,一个通道对应的特征图在进行全局平均池化操作之后得到一个特征点。而这些特征点最后将组成特征向量输入下一层中。对比全连接层,全连接层放在卷积模块之后,主要是通过将全连接模块输出的特征图打散成一维向量,通过对这个向量进行乘法操作,得到降低维度后的特征向量。全连接层和全局平均池化的对比如图 3 所示。

全局平均池化这种操作是不需要数据参数的,相比全连接层中包含的大量输入输出映射变换,减小了参数量,无需大量训

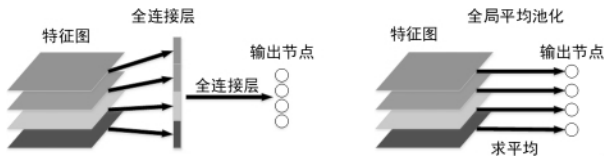


图3 全连接层与全局平均池化的实现方式对比

训练调优。可以说是从结构上进行了正则化,减小了参数空间,防止过拟合,增强了模型的鲁棒性。同时全局平均池化在特征图与特征向量(在某些情况下甚至对应最终的分类)的转换更加自然,剔除了全连接层的黑箱操作,也使得每个通道具有了实际的类别意义。同时,如果将全局平均池化用于取代网络最后的全连接层,还可以使得网络不用对输入图像尺寸的进行限制。

## 2.2 可分离卷积,三维卷积

三维卷积可以看做是二维卷积结构针对三维情况的扩展,在医学图像和视频处理领域,三维卷积已经得到了广泛的使用。本文所设计的网络模型中,三维卷积的卷积核结构为(3,4,3),padding方式为same。

在进行标准卷积操作时,对应输入的图像相应区域中的所有通道均同时进行了卷积操作。但这种操作实际上增大了卷积操作的参数量,同时令人疑惑的是,为什么我们在进行卷积操作时,要将图像的通道和图像区域进行捆绑,同时操作,这样的意义在哪里,效果真的好吗?为什么不能将图像的通道和空间区域分开进行处理?

基于以上考量,可分离卷积诞生了。可分离卷积分为depth-wise和pointwise两步。首先我们对输入图像的每个通道分别进行卷积操作,图像有多少个通道,就对应多少个卷积核,这样相当于提取了每个通道的空间特征,这就是depthwise操作。在每个通道完成卷积操作,得到其对应的特征图之后,再对这些特征图进行跨通道的标准卷积操作,不过使用的卷积都是(1,1,1)大小的卷积核,这样相当于得到了区域中每个点的特征值,这就是pointwise操作。可分离卷积不仅实现通道和空间区域的特征提取过程的分离,同时还大大减小了模型的参数计算量。以二维图像的卷积操作为例。假设输入图像高度为H,宽度为W,通道数为C。要求输出的通道数为k,使用的卷积核大小为最常见的(3,3)。那么普通卷积的计算量为 $H \times W \times C \times k \times 3 \times 3$ ,而使用可分离卷积的计算量为 $H \times W \times C \times 3 \times 3 + H \times W \times C \times k$ 。使用可分离卷积相比标准卷积参数压缩量如公式(3)所示,其中conv代表普通的卷积操作:

$$\frac{\text{depthwise+pointwise}}{\text{conv}} = \frac{H \times W \times C \times 3 \times 3 + H \times W \times C \times k}{H \times W \times C \times k \times 3 \times 3} = \frac{1}{k} + \frac{1}{3 \times 3} \quad (3)$$

可分离卷积模块则由可分离卷积层、批量标准化层(BN)、指数线性激活函数层(ELU)、随机失活层(Dropout)层四个基础层组成。而卷积模块与可分离卷积模块相似,不同之处在于使用的是标准卷积层,同时在指数线性激活函数层(ELU)与随机失活层(Dropout)层之间加入了三维最大池化层(Maxpooling)。

## 2.3 网络结构

网络可以看做由数个重复的模块组成,包含上文提到的卷积模块、可分离卷积模块、全连接模块。网络的输入包含两部分,临床数据诊断数据在经过两个全连接模块之后,在网络最后的分类层前与经过一系列处理得到的图像特征向量进行拼接,然后输入分类层,得到最终结果。图像数据从输入开始,分别经过两个卷积模块处理,然后经过一个残差结构。这个残差结构具体而言,是以网络第二个卷积模块的输出作为输入,这部分输入经过三个可分离卷积处理之后,与第二个卷积模块的输出直接进

行逐点相加。相加得到的特征向量维度为(None,7,8,7,48),其中None代表的是本批次的样本个数,48是特征向量的通道数,根据通道所在的维度将特征向量拆分成两个(None,7,8,7,24)的特征向量,对应输入下面的两个卷积模块,在经过全局平均池化之后,再将其沿通道所在维度拼接,最后输入神经元个数为10的全连接模块。而全连接模块的输出则与之前提到的经过处理的临床诊断数据特征向量拼接,输入分类层,得到最终标签值。

## 3 实验结果

网络模型主要使用Keras和TensorFlow框架构建,模型训练时使用的优化算法是Adam算法,Adam算法中参数 $\beta_1=0.9$ , $\beta_2=0.999$ , $\epsilon=1e-08$ 。学习率初始值为0.001,最大迭代周期为120,批量大小(batch size)大小为32。输入包含维度为(121,145,121)的经过预处理的MRI图像,以及特征维度为12的一维临床诊断数据。模型在AD/MCI/NC的三分类任务上,相比以往的模型取得了较大的进步。我们使用准确率(Accuracy)来对结果进行评价,本文所提出的模型在三分类任务上,经过五折交叉验证的准确率为97.10%,准确率的表达式如(4)所示。其中TP代表真正例(True Positive),TN代表真反例(True Negative),FP代表假正例(False Positive),FN代表假反例(False Negative)。

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \times 100\% \quad (4)$$

同时因为使用残差结构、可分离卷积、全局平均池化等结构,本文所设计的模型相对其它神经网络模型的,其所包含的参数得到了成倍地减小。模型总共的可训练参数为278571,相比之下,其他的针对三维图像的神经网络模型参数至少是百万级的。

为了验证模型的效果,我们使用一些经典方法对分类效果进行对比,包含VGG19<sup>[12]</sup>、Resnet50<sup>[7]</sup>、Densenet121<sup>[13]</sup>等针对二维图像的经典模型,以及由Payan<sup>[14]</sup>等人提出的在AD领域经典的三维网络模型。实验结果如表2所示。

表2 各个模型对应的三分类与二分类结果对比  
(准确率(%))以及标准差)

Model	AD/MCI	MCI/NC	AD/NC	AD/NC/MCI
VGG19	84.87 ± 0.64	80.42 ± 0.76	93.89 ± 0.56	79.61 ± 0.62
Resnet50	85.25 ± 0.86	76.55 ± 0.60	95.18 ± 0.68	78.20 ± 0.72
Densenet121	84.52 ± 0.83	79.86 ± 0.72	95.42 ± 0.58	82.38 ± 0.76
Payan et al.	86.84 ± 0.73	92.11 ± 0.74	95.39 ± 0.66	82.38 ± 0.68
We proposed	96.82 ± 0.43	98.86 ± 0.54	99.22 ± 0.42	97.10 ± 0.46

由表2可知,我们的模型无论在二分类还是三分类上都取得较大的进步,除了之前提到的模型结构因设计原因之外,引入临床诊断数据,使用多模态数据进行融合,在一定程度上对分类准确性的提高也起到了作用。

## 参考文献

- [1] Barnes D, Yaffe K. The Projected impact of risk factor reduction on Alzheimer's disease prevalence[J]. Alzheimer's & Dementia, 2011,7(4):511
- [2] Murphy M, Paul, Harry LeVine III. Alzheimer's disease and the amyloid- $\beta$  peptide[J]. Journal of Alzheimer's Disease, 2010,19(1):311-323
- [3] Braak H, Braak E. Development of Alzheimer-related neurofibrillary changes in the neocortex inversely recapitulates cortical myelogenesis[J]. Acta Neuropathologica, 1996,92(2):197-201
- [4] Sarraf S, Tofighi G. Deep learning-based pipeline to recognize Alzheimer's disease using fMRI data [C]//Future Technologies Conference. IEEE, 2017

(下转第109页)



学习层次的所有已制定完成的培养方案。学生可以在数据窗体内通过专业、院系、年级等条件筛选出所属的培养方案。选择方案后,系统自动将该方案的其余字段信息如“毕业要求总学分、方案最低学分”等写入培养计划主表,完成培养方案的选择。

### 3.2.2 课程选择

培养计划的课程选择与培养方案息息相关,学生所选课程总学分不仅需要达到毕业要求总学分,并且选择的课程需要在培养方案的规定课程内选择达到方案最低学分才能满足毕业要求。故利用脚本设计批量导入课程按钮,学生根据方案要求批量导入将要修的课程,导入课程后会自动累加学分,并且判断该门课程是否是培养方案中的课程,统计方案内课程总学分,实时更新到“已选所有课程学分”和“已选本方案总学分”两字段中。

由于研究生课程中存在免修课程如公共英语等,符合要求的学生可以申请免修该门课程,免修课程需要在培养计划中标识出来。当培养计划导入课程时,若该学生在该门课程的免修名单中,则需要将该门课程的状态改为免修。其对应关系如图 5 所示。

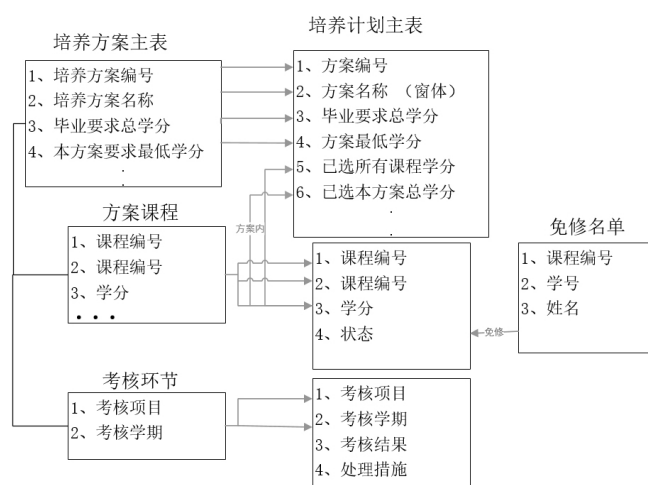


图 5 培养计划子主表对应关系

新系统将培养方案与培养计划关联后,不仅业务流程上更加规范,减少了教学秘书、辅导员的工作量,也减少了学生选课错误的发生。为学生制定培养计划、选课提供了便利,提升了效率,无论是可靠性、可维护性都有所提升。

## 4 结束语

本文根据研究生院对于信息管理平台的可靠性强、鲁棒性强、可改性强需求,选择基于云平台 PaaS 平台定制化涉及开发一款新的研究生院信息管理平台。可以简化学生的操作,减少老师的工作量,提高教务活动流程的规范性、数据的准确性以及管理工作的高效性。还可以增强领导监控能力、提高系统的稳定性和安全性。本文主要介绍了新系统的以下几个方面:

1)由于 PaaS 平台集成性强、开发性强并且无需忧心于开发所必需的服务器、数据库、网络和存储等基础结构的配置,且 B/S 架构与 C/S 架构相比,有负载低、容易操作、维护便捷等十分突出的优点,选择基于 PaaS 台作为新系统的开发平台,B/S 架构作为软件开发架构;

2)系统整体架构设计庞大且复杂,各模块之间各自独立又相互联系,一个功能块由多个场景组成。不同场景对不同的角色开放,可通过系统管理模块中角色、用户对管理工作流中的人员与权限进行分配。不同角色通过在各自场景进行操作,共同完成功能模块的具体流程,实现业务逻辑控制的严格把关;

3)教学管理模块是系统的核心交互模块,用户量大、使用频率高,对与并发性处理要求较高,其中培养方案与培养计划由研究生院、学院、学生共同完成。培养方案的制定直接影响培养计划制定,这两块功能联系紧密,却又相互独立,是教学管理模块中较为典型的模块交互方式。该种模式既能保证数据的关联性,又能保证数据的独立性,可有效提升系统的可靠性。

## 参考文献

- [1]叶冬.高校学籍管理系统设计与实现[J].计算机产品与流通,2019(10):197-198
- [2]雷唯.面向智慧城市大数据 PaaS 平台的研究[J].电子制作,2019(14):38-40,32
- [3]朱瑾.网络协作办公系统关键技术的研究及实现[D].西安:西安工业大学,2015
- [4]魏育华,陈允行.基于 B/S 模式的高校教师信息管理系统开发与探讨[J].网络安全技术与应用,2019(10):63-65
- [5]苗风华.SQL Server 2008 数据管理系统的优势研究[M].长春:长春师范大学,2014:2

[收稿日期:2020.1.3]

(上接第 60 页)

- [5]Francois Chollet.Xception:Deep Learning with Depthwise Separable Convolutions[C]// 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2017
- [6]Lin Min,Chen Qiang,Yan Shuicheng.Network In Network [J]. Computer Science, 2013
- [7]K He,X Zhang,S Ren,et al.Deep Residual Learning for Image Recognition [C]//IEEE Conference on Computer Vision and Pattern Recognition,Nevada,USA,2016
- [8]Szegedy Christian,Ioffe Sergey,Vanhoucke Vincent.Inception-v4,Inception-ResNet and the Impact of Residual Connections on Learning[J].Computer Vision and Pattern Recognition,2016
- [9]S Ioffe,C Szegedy.Batch Normalization:Accelerating Deep Network Training by Reducing Internal Covariate Shift[J].arXiv preprint arXiv:1502.03167, 2015
- [10]N Srivastava,G Hinton,A Krizhevsky,et al.Dropout:A Simple

Way to Prevent Neural Networks from Overfitting[J].J Mach Learn Res,2014(15):1929-1958

- [11]Clevert Djork-Arné, Unterthiner T,Hochreiter S.Fast and Accurate Deep Network Learning by Exponential Linear Units (ELUs)[J]. Computer Science, 2015
- [12]S Han,H Mao,WJ Dally,et al.Deep compression:Compressing deep neural networks with pruning, trained quantization and Huffman coding[J].arXiv:1510.00149, 2015
- [13]G Huang,Z Liu,L van der Maaten,et al.Densely Connected Convolutional Networks [C]//IEEE Conference on Computer Vision and Pattern Recognition, Hawaii, USA, 2017
- [14]Payan A,Montana G.Predicting Alzheimer's disease: a neuroimaging study with 3D convolutional neural networks [J]. Computer Science, 2015

[收稿日期:2020.1.2]