



## 4D卷积神经网络的自闭症功能磁共振图像分类

郭磊, 王骏, 丁维昌, 潘祥, 邓赵红, 施俊, 王士同

引用本文:

郭磊, 王骏, 丁维昌, 等. 4D卷积神经网络的自闭症功能磁共振图像分类[J]. *智能系统学报*, 2021, 16(6): 1021–1029.

GUO Lei, WANG Jun, DING Weichang, et al. Classification of the functional magnetic resonance image of autism based on 4D convolutional neural network[J]. *CAAI Transactions on Intelligent Systems*, 2021, 16(6): 1021–1029.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202009022>

## 您可能感兴趣的其他文章

### 融合整体与局部信息的武夷岩茶叶片分类方法

Classification of Wuyi rock tealeaves by integrating global and local information  
*智能系统学报*. 2020, 15(5): 919–924 <https://dx.doi.org/10.11992/tis.202003018>

### 基于小样本学习的LCD产品缺陷自动检测方法

An automatic small sample learning-based detection method for LCD product defects  
*智能系统学报*. 2020, 15(3): 560–567 <https://dx.doi.org/10.11992/tis.201904020>

### 深度学习的双人交互行为识别与预测算法研究

Human interaction recognition and prediction algorithm based on deep learning  
*智能系统学报*. 2020, 15(3): 484–490 <https://dx.doi.org/10.11992/tis.201812029>

### 面向自闭症辅助诊断的无监督模糊特征学习新方法

A novel unsupervised fuzzy feature learning method for computer-aided diagnosis of autism  
*智能系统学报*. 2019, 14(5): 882–888 <https://dx.doi.org/10.11992/tis.201808005>

### 基于改进卷积神经网络的多标记分类算法

A multi-label classification algorithm based on an improved convolutional neural network  
*智能系统学报*. 2019, 14(3): 566–574 <https://dx.doi.org/10.11992/tis.201804056>

### 基于卷积特征和贝叶斯分类器的人脸识别

Face recognition based on convolution feature and Bayes classifier  
*智能系统学报*. 2018, 13(5): 769–775 <https://dx.doi.org/10.11992/tis.201706052>

 微信公众平台



关注微信公众号, 获取更多资讯信息

DOI: 10.11992/tis.202009022

网络出版地址: <https://kns.cnki.net/kcms/detail/23.1538.tp.20210330.1640.004.html>

## 4D 卷积神经网络的自闭症功能磁共振图像分类

郭磊<sup>1</sup>, 王骏<sup>2</sup>, 丁维昌<sup>2</sup>, 潘祥<sup>1</sup>, 邓赵红<sup>1</sup>, 施俊<sup>2</sup>, 王士同<sup>1</sup>

(1. 江南大学人工智能与计算机学院, 江苏 无锡 214122; 2. 上海大学通信与信息工程学院, 上海 200444)

**摘要:** 静息态功能磁共振图像是随着时间变化的一系列三维图像。已有的 3D 卷积过程本质上是对三维图像数据或二维图像+时间维数据进行处理, 无法有效地融合静息态功能磁共振图像的时间轴信息。为此, 本文提出了新型的 4D 卷积神经网络识别模型。具体而言, 通过对输入的 fMRI 使用四维卷积核执行四维卷积, 在自闭症患者的功能磁共振图像中, 从空间和时间上提取特征, 从而捕获图像在时间序列上的变化信息。所开发的模型从输入图像中生成多个信息通道, 最终的特征表示结合了所有通道的信息。实验结果表明, 在保证模型泛化性能的前提下, 该方法融合了功能像的全局信息, 并且采集了功能像随时间变化的趋势信息, 进而解决了用卷积神经网络处理三维图像随时间变化的分类问题。

**关键词:** 深度学习; 卷积神经网络; 自闭症; 4D 卷积; 功能磁共振成像; 特征提取; 特征融合; 图像分类  
中图分类号: TP391 文献标志码: A 文章编号: 1673-4785(2021)06-1021-09

中文引用格式: 郭磊, 王骏, 丁维昌, 等. 4D 卷积神经网络的自闭症功能磁共振图像分类 [J]. 智能系统学报, 2021, 16(6): 1021-1029.

英文引用格式: GUO Lei, WANG Jun, DING Weichang, et al. Classification of the functional magnetic resonance image of autism based on 4D convolutional neural network[J]. CAAI transactions on intelligent systems, 2021, 16(6): 1021-1029.

### Classification of the functional magnetic resonance image of autism based on 4D convolutional neural network

GUO Lei<sup>1</sup>, WANG Jun<sup>2</sup>, DING Weichang<sup>2</sup>, PAN Xiang<sup>1</sup>,  
DENG Zhaohong<sup>1</sup>, SHI Jun<sup>2</sup>, WANG Shitong<sup>1</sup>

(1. School of Artificial Intelligence and Computer, Jiangnan University, Wuxi 214122, China; 2. School of Communication and Information Engineering, Shanghai University, Shanghai 200444, China)

**Abstract:** Resting-state functional magnetic resonance images are a series of three-dimensional (3D) images that change over time. The existing 3D convolution processes 3D image data or two-dimensional image and time-dimensional data, but it cannot effectively fuse the time axis information of a resting-state functional magnetic resonance image. To resolve this, a new four-dimensional (4D) convolutional neural network (CNN) recognition model is proposed in this paper. Specifically, by performing a 4D convolution using a 4D convolution kernel on the input functional magnetic resonance imaging, features are spatially and temporally extracted from the functional magnetic resonance image of a patient with autism, thereby capturing information about the changes in the image's time series. The developed model generates multiple information channels from the input image, and the final feature representation combines information from all channels. The experimental results show that to ensure the generalization performance of the model, the method fuses the global information of the functional image and collects its trend information over time, consequently solving the classification problem of 3D image changes with time using a CNN.

**Keywords:** deep learning; convolutional neural network; autism; 4D convolution; functional magnetic resonance imaging; feature extraction; feature fusion; image classification

自闭症 (autism spectrum disorder, ASD) 是一种广泛性发展障碍类疾病, 也是一种具有生物基础的发育障碍类疾病。该病症发病时间不等, 大部

分发病于儿童时期并伴随一生。其主要症状表现在社会交流障碍、语言交流障碍、情感缺陷等方面。这就导致了患者在日常的生活、交流和学习中有很大的障碍。因此, 自闭症的预测分类研究一直备受广大研究者的关注。研究者已经发现 ASD 患者与正常人 (typically developing individu-

收稿日期: 2020-09-16. 网络出版日期: 2021-03-31.  
基金项目: 江苏省自然科学基金项目 (BK20181339).  
通信作者: 王骏. E-mail: [wangjun\\_shu@shu.edu.cn](mailto:wangjun_shu@shu.edu.cn).

als, TD) 在大脑的功能连接上有明显的差异<sup>[1]</sup>。

静息态功能磁共振成像 (resting state functional magnetic resonance image, rs-fMRI) 通过在静息状态下采集血氧水平依赖信号, 对信号去噪<sup>[2]</sup>后, 使用多层同时扫描技术<sup>[3]</sup>对功能磁共振全脑影像快速采集, 以此来获得时间分辨率和空间分辨率较高的图像, 获得 rs-fMRI 后, 关键问题是如何从中有效提取特征。对于传统的特征提取方法, 以相关分析法为例, 即选取 ROI 种子点并计算它的平均时间序列, 然后与全脑所有体素的时间序列进行计算, 或者两个脑区之间的平均时间序列计算, 得到相关系数, 以此来反映不同脑区神经活动模式之间的时间相关性<sup>[4]</sup>。作为图像的预处理阶段, 该方法对 ROI 异常敏感, 种子点选取的位置、大小差异都可能导致较大的数据分析结果差异。

近年来, 卷积神经网络 (convolutional neural networks, CNN) 在医学图像分类任务中取得了非常好的效果, 随着结构的不断优化, 诞生了多种成熟的网络模型。2017年, Dvornik 等<sup>[5]</sup>使用具有 Long Short-Term Memory 的递归神经网络直接从静态 fMRI 的时间序列上对具有 ASD 的个体进行分类。但是, 从 fMRI 的感兴趣区域中提取平均时间序列作为输入数据无法兼顾全局信息, 易造成信息缺失; 2018年, Yu 等<sup>[6]</sup>开发了一种全自动的 3D 卷积神经网络框架, 用以识别和分类不同类型的功能性脑网络。但是, 脑网络在训练前需要花费大量的时间进行手动标记, 不可避免地会出现人工标签错误和标签间的差异; 2018年, Li 等<sup>[7]</sup>为了在 fMRI 中同时利用空间和时间信息, 研究了使用滑动窗口随时间测量时间统计量 (均值和标准差) 以及使用 3D 卷积神经网络 (CNN) 捕获空间特征的潜在好处。模型虽然在高维数据上获得了良好的空间识别性能, 但是在时间维度上的表现并不明显; 2019年, Khosla 等<sup>[8]</sup>提出了一种新颖的 3D 卷积神经网络 (CNN) 方法来实现集成学习策略, 提出的 CNN 方法利用了 rs-fMRI 数据的全分辨率 3D 空间结构, 并拟合了非线性预测模型。但是, 将 rs-fMRI 预处理后的多通道 3D 图像作为 CNN 的输入忽视了隐含在时间轴上的信息。对于以上提出的深度学习方法, 网络只能处理维数小于或等于 3 的灰度或 RGB 图像, 并且在图像进行预处理过程中, 很可能会丢失部分信息, 使得出的结果不具有全局性。对于功能磁共振图像而言, 如果将每个时间点上的 3D 图像视为一个通道进行 3D 卷积, 那么根据预处理后的数据集特性, 网络生成的通道数目会随着时间的推移而增多, 从而使得网络宽度过大, 可能会造成传播过程中信息丢失, 影响网络的表达能力<sup>[9]</sup>。另一方面, 可以将每个样本多个时间点上的 3D 图像视为一个图像立方体输入到 3D-CNN

中, 但是这种处理会增加大量的参数, 忽略静息态功能像随时间变化的趋势信息, 并且性能表现不佳。总而言之, 有必要寻找一种有效的方法从 rs-fMRI 的时间和空间信息中提取特征。

本文提出了 4D 卷积神经网络, 通过 rs-fMRI 对 ASD 进行分类, 所提出的 4D 卷积操作从功能磁共振图像中提取空间和时间上的特征, 可以得到功能磁共振图像随时间变化的信息。与几个流行的网络相比, 例如 AlexNet<sup>[10]</sup>、3D-CNN<sup>[11]</sup>、2CC3D<sup>[7]</sup>、T3D<sup>[12]</sup>, 所提出的网络能够有效地从功能磁共振图像中提取与 ASD 密切相关的时间空间特征, 较之以往的 3D 卷积神经网络, 4D-CNN 对含有时空特性的数据具有更好的适应性, 可以实现更好的分类效果。

## 1 卷积神经网络概述

深度学习是一类通过多层处理, 逐渐将初始的低层特征表示转化为高层特征表示的特征学习方法。这样的方法可以被训练用于监督或无监督学习, 由此产生的复杂模型已经被广泛用于视觉物体识别、人类行为识别、自然语言处理、音频分类、脑机交互、人体跟踪、图像恢复、去噪、分割等任务。卷积神经网络是一种深度模型, 在这种模型中, 输入层、隐藏层和输出层组成基本的网络结构, 其中隐藏层通常包括卷积层、池化层和全连接层。卷积层中卷积核相当于特征提取器, 每个卷积核对整幅图像数据进行有规律地扫描并依次计算各局部区域的特征值以匹配某种模式; 池化层对卷积层结果进行特征选择和信息过滤, 实现一定程度的平滑压缩; 全连接层通常搭建在隐藏层的最后, 对卷积层和池化层的结果进行分类。

1998年, 纽约大学的 Lecun 等<sup>[13]</sup>正式提出了卷积神经网络 LeNet-5, 确立了 CNN 的现代结构, 并在图像处理领域获得了极大成功。之后研究人员发现当神经网络饱和之后模型会产生过拟合现象, 直到 2012年, Krizhevsky 等<sup>[10, 14]</sup>在提出的 Alexnet 中采用 Dropout 和数据增强的方法来防止网络过拟合, 并使用非线性激活函数 ReLU (rectified linear unit) 提高网络收敛速度。2014年, Simonyan 等<sup>[15]</sup>采用小卷积堆叠的方式搭建 VGGNet, 证明了更深的网络能更好地提取特征, 值得一提的是, 虽然 VGGNet 比 AlexNet 的参数多, 但只需要更少的迭代次数就能收敛。此外, 在设计网络结构时, 除了考虑网络的深度, 还会考虑网络的宽度。同年, Szegedy 等<sup>[16]</sup>提出的 Inception v1 网络, 通过改变传统的卷积层设计思路, 在加大网络深度和宽度的同时减少了参数量, 并且增加了网络对尺度的适应性。随后, 通过对 Inception

v1 网络结构的不断改进, 诞生了 Inception v2<sup>[17]</sup>、Inception v3<sup>[18]</sup>、Inception v4<sup>[19]</sup>。在 VGG 网络中, 当网络深度达到一定界限的时候, 就会产生梯度消失现象, 2016 年, He 等<sup>[20]</sup>在提出的 ResNet 中引入了跳跃连接结构来解决梯度消失的问题, 进而可以进一步加大网络深度。2017 年, Huang 等<sup>[21]</sup>受 ResNet 和 Inception 的影响, 提出了 DenseNet, 不同于 ResNet 的是, DenseNet 模型建立的是前面所有层与后面层的密集连接, 通过特征重用和旁路设置, 既缓解了梯度消失问题, 又大幅度减少了网络的参数量, 实现了比 ResNet 更优的性能。

根据卷积核结构的不同, CNN 可分为一维卷积、二维卷积与三维卷积。一维卷积通常用于序列数据分析, 如从 rs-fMRI 的感兴趣区域中提取时间序列矩阵并输入到一维卷积神经网络中对自闭症进行分类<sup>[22]</sup>; 二维卷积被广泛用于计算机视觉的图像数据分析, 如具有代表性的网络 LeNet-5 识别手写数字图片<sup>[13]</sup>; 而三维卷积则可用于时空数据的分析, 如 3D-CNN 识别视频中的人类动作<sup>[11]</sup>。CNN 的卷积层通过卷积操作提取输入的特征, 低层卷积提取低级特征, 随着网络结构的加深, 更高层的卷积层提取更高级的特征。在卷

积层中每个卷积核在对输入图像进行扫描时, 仅处理其投影在图像上的感受野的数据, 且卷积核权重保持不变, 此即为卷积核的局部连接性与权值共享性。由于卷积核的局部连接特性, 模型可以计算区域特征, 充分利用局部区域内的空间信息; 由于卷积核的权值共享特性, 在模型训练时卷积层所需参数要远小于全连接层。因此 CNN 适用于特征维度较大、具有位置信息的图像数据的分析。

本文研究了 4D-CNN 在医学图像分类任务中对自闭症的功能核磁共振图像进行分类的应用, 通过用 4D 卷积核将多个连续的三维图像堆叠在一起, 可实现 4D 卷积。在降采样层中, 通过在上一层特征图上的局部领域上合并来降低特征图的分辨率, 从而去除冗余信息, 提高网络效率。最后通过全连接层将特征映射到样本标记空间实现分类任务。

## 2 基于 fMRI 数据的 4D 卷积神经网络

本文的技术路线如图 1 所示, 其中涉及原始数据的预处理、数据增强、模型训练和使用训练好的模型进行测试。

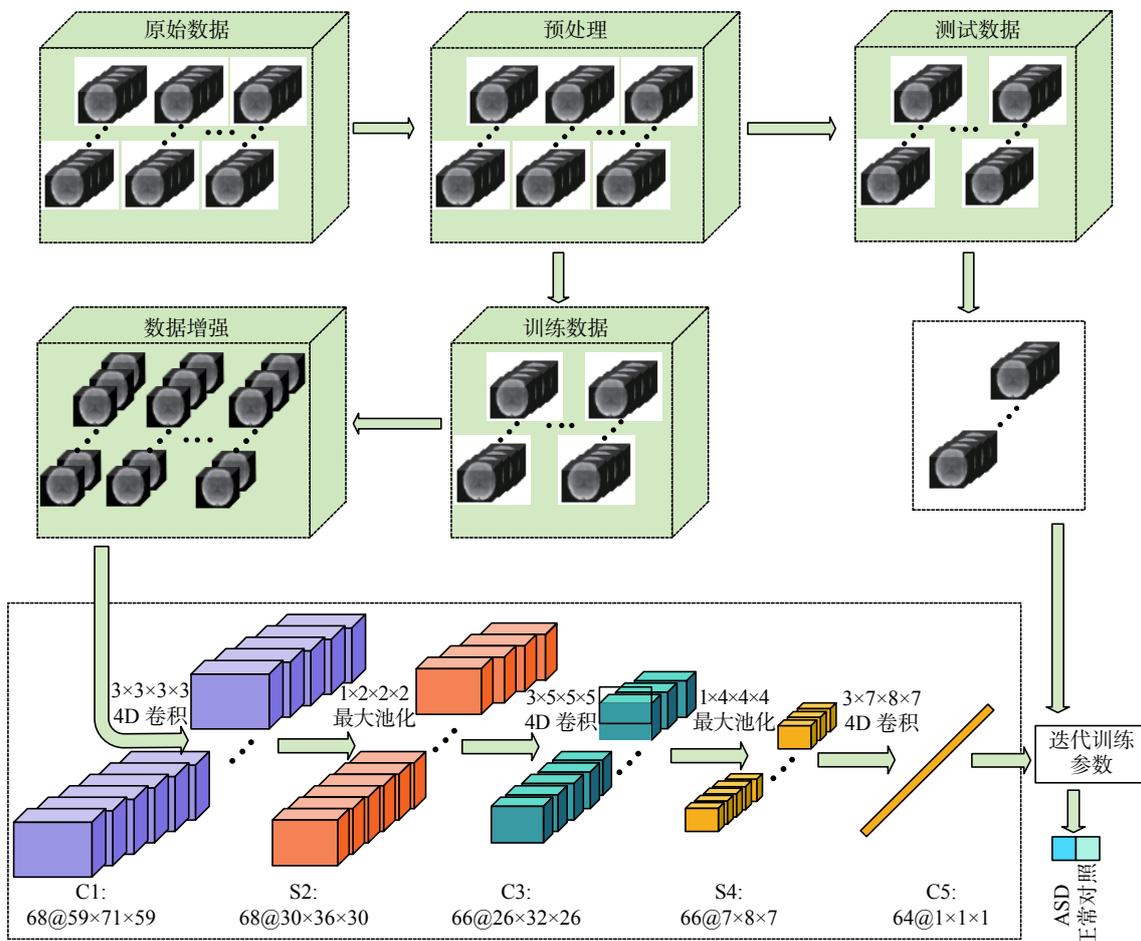


图 1 用于自闭症辅助诊断的 4D-CNN 流程

Fig. 1 4D-CNN procedure for the auxiliary diagnosis of autism

### 2.1 数据预处理

本文实验基于 ABIDE (autism brain imaging data exchange) 数据库, 其中包括来自多个数据中心的 1 112 名 ASD 患者和 TD 的 rs-fMRI 数据。在本文中, 考虑从 ABIDE 的 NYU、UM\_1 和 USM 这 3 个成像中心中获得 rs-fMRI, 其中各个数据中心的 ASD 组和 TD 组样本数如表 1 所示。使用 DPARSF<sup>[23]</sup> 对 rs-fMRI 数据进行预处理。具体而言, 将每个受试者的前 10 个获得的 rs-fMRI 丢弃。然后, 执行时间层校正 (slice timing) 和头部运动校正 (realign)。每个人的大脑尺寸大小、形状和位置差异较大, 无法将结果按照统一的标准进行分析, 因此需要把不同人的大脑影射在同一空间内, 本文使用 EPI (echo-planar image) 模板将数据统一到蒙特利尔神经学研究所 (montreal neurological institute, MNI) 制定的标准空间中, 从而完成对功能性进行空间标准化 (normalize) 处理。为了提高配准效果和增大图像的信噪比, 需要对图像进行空间平滑 (smooth)。经过预处理后, 本文得到了每个被试在单个时间点上尺寸大小为  $61 \times 73 \times 61$  的 rs-fMRI 数据。为了验证模型的泛化能力, 在数据预处理后将其分为训练集和测试集, 在不同成像中心下的具体数据分布如表 2 所示。

表 1 3 个成像中心的样本分布

Table 1 Sample distribution of the three imaging centers

数据集	ASD	正常对照	总数
NYU	79	105	184
UM_1	55	55	110
USM	43	58	101

表 2 3 个成像中心训练和测试的样本数量

Table 2 Sample size for training and testing at the three imaging centers

数据集	训练		测试		总数
	ASD	正常对照	ASD	正常对照	
NYU	53	70	26	35	184
UM_1	37	37	18	18	110
USM	29	39	14	19	101
总计	119	146	58	72	395

### 2.2 数据增强

由于每个被试采集的图像具有不同的时间长度, 同时, 也为了扩大数据集的规模, 避免因训练样本过少而引发的模型过拟合问题, 本文对训练

集应用数据增强策略。具体而言, 针对每个被试的 rs-fMRI, 在保证时间序列的前提下, 对时间点上的图像进行随机选取。为了保证数据的有效性和一致性, 选取 70 个时间点上的图像作为一个时间序列, 每个被试随机生成 5 组, 并标上相应的标签。通过把所有被试原始训练集中的 rs-fMRI 经过上述过程的处理, 得到了一个新的训练集。

### 2.3 4D 卷积神经网络分类模型

在三维卷积神经网络中, 当应用于时间序列上连续的二维特征图时, 三维卷积从空间和时间两个维度计算特征, 通过将 1 个三维核与多个相邻二维图叠加而成的立方体进行卷积来实现的, 在这种构造中, 卷积层中的特征图被连接到前一层中的多个相邻图像, 从而获得二维特征图的连续信息。在功能磁共振图像数据中, 时间序列上呈现的是三维空间结构, 本文应用四维卷积核与多个相邻的三维图叠加而成的四维特征进行卷积, 用于处理在时间序列上呈现三维空间结构的数据, 从而在空间和时间两个维度上计算特征。四维卷积过程如图 2 所示。其中, 卷积核在时间维度上的大小为 3, 连接采用颜色编码, 相同的颜色权重共享。将相同的 4D 内核应用于连续的 3D 立方块中, 通过移动步长提取前一层整个时间序列上的特征。

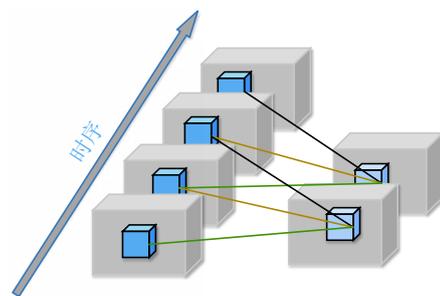


图 2 4D 卷积过程

Fig. 2 4D convolution process

图 1 给出了用于 rs-fMRI 的 4D-CNN 结构, 它由 1 个输入层、3 个卷积层、1 个全连接层和 1 个输出层组成。最后使用 softmax 函数将输出结果转换为概率表示, 以便更好地实现分类。每个卷积层由 1 个大小为  $P \times Q \times R \times T$  (空间维度为  $P \times Q \times R$ , 时间维度为  $T$ ) 的四维过滤器内核, 1 个激活单元和 1 个池化单元组成。卷积操作通过在输入图像或前一层卷积层的输出中对内核进行卷积来生成特征图, 并设置 ReLU 作为激励函数拟合网络的非线性映射, 防止网络梯度消失和加快训练速度。池化单元起到二次提取特征的作用, 经

过最大池化采样, 将高维特征转换成低维代表性特征, 有利于在连续多层的网络中进行更深层次的表示。最后, 通过反向传播算法<sup>[24]</sup>, 使来自损失函数的信息通过网络传递, 计算梯度来训练网络。

具体而言, 由于数据集中每个样本采集图像的时间长短有所差异, 为了获得相同的输入规模, 本文将统一选取每个被试的前  $M$  个时间点, 每一个时间点上特征图大小为  $X \times Y \times Z$  的 fMRI 作为特征输入到 4D-CNN 中。该网络由 6 层结构组成, 每层都包括可训练的参数。输入层包含  $M$  个通道, 每个通道大小为  $X \times Y \times Z$  的四维 fMRI 数据。对输入层应用大小为  $3 \times 3 \times 3 \times 3$  (空间维度为  $3 \times 3 \times 3$ , 时间维度为 3) 的核进行 4D 卷积, 卷积完成后得到  $M_1$  个特征图, 每个特征图大小为  $X_1 \times Y_1 \times Z_1$ 。S2 层是由  $M_1$  个大小为  $X_2 \times Y_2 \times Z_2$  的特征图组成的降采样层, 由 C1 层经过  $2 \times 2 \times 2 \times 1$  的最大池化得到。卷积层 C3 是用大小为  $5 \times 5 \times 5 \times 3$  (空间维度为  $5 \times 5 \times 5$ , 时间维度为 3) 的 4D 卷积核对 S2 层进行卷积得到, 规模为  $X_2 \times Y_2 \times Z_2 \times M_2$  (空间维度为  $X_2 \times Y_2 \times Z_2$ , 时间维度为  $M_2$ )。S4 层是由  $M_2$  个大小为  $X_3 \times Y_3 \times Z_3$  的特征图组成的降采样层, 由 C3 层经过  $4 \times 4 \times 4 \times 1$  的最大池化得到。根据 S4 层数据规模, 采用  $7 \times 8 \times 7 \times 3$  (空间维度为  $7 \times 8 \times 7$ , 时间维度为 3) 的卷积核对 S4 层进行卷积, 得到了 C5 层, 规模为  $X_4 \times Y_4 \times Z_4 \times M_3$  (空间维度为  $X_4 \times Y_4 \times Z_4$ , 时间维度为  $M_3$ )。输出层由与分类任务类别数量相同的特征组成, 每个特征都与 C5 层中的  $X_4 \times Y_4 \times Z_4 \times M_3$  个特征完全连接。通过多层卷积和池化, 将一个被试在时间序列上连续的功能磁共振图像转换为一个  $X_4 \times Y_4 \times Z_4 \times M_3$  维的特征向量。相比较以往处理功能磁共振图像的方法, 在卷积的过程中, 蕴含着连续的时间信息。在本设计中, 本文最终在  $X_4 \times Y_4 \times Z_4 \times M_3$  维特征向量上应用一个线性分类器进行分类。本文设计并评估了其他在不同阶段结合多个信息通道的 4D-CNN 架构, 结果表明该架构的性能最佳。

### 3 实验与分析

在本节中, 在 ASD 数据集上验证本文提出的 4D-CNN 模型, 通过对影响模型性能的超参数进行寻优, 来获得最佳的网络结构。为了验证 4D-CNN 在 ASD 分类上的优越性, 进一步在 ASD 的多个成像中心上比较了多种对比方法。

为了体现测试数据集上分类模型的有效性, 选取精确度 (ACC)、AUC 值、敏感度 (SEN) 和特

异性 (SPE) 作为实验的性能指标。表 3 和式 (1)~(3) 用于计算相应的性能指标:

$$ACC = \frac{TP + TN}{TP + FP + TN + FN} \quad (1)$$

$$SEN = \frac{TP}{TP + FN} \quad (2)$$

$$SPE = \frac{TN}{TN + FP} \quad (3)$$

表 3 混淆矩阵中的元素  
Table 3 Elements in the confusion matrix

真实情况	预测结果	
	ASD	正常对照
ASD	TP(真正例)	FN(假反例)
正常对照	FP(假正例)	TN(真反例)

#### 3.1 超参数寻优

在本节中, 通过 4D-CNN 来执行 ASD 诊断, 将预处理后的 rs-fMRI 作为网络的输入数据。本文将 ASD 组标记为 1, 将健康对照组标记为 0, 将 ASD 诊断视为二分类问题。为了找到适合 ASD 诊断的 4D-CNN 的最佳架构, 本文进行了架构搜索<sup>[25]</sup>策略, 即探索基于 4D-CNN 架构的最佳超参数配置。

对于卷积神经网络, 输入帧序列的时间深度在活动识别任务中起着关键作用<sup>[12]</sup>, 为了实现新架构的最佳配置, 本文对网络输入数据的时间深度进行了一系列测试, 表 4 列出了在不同时间深度下模型的测试性能。为寻求 4D-CNN 的最佳配置, 基于上述研究的结果进行了实验, 并采用 4D 卷积核 (多个连续的三维权值矩阵) 在卷积层之间建立更有效的时空连接, 便于最大限度地从 fMRI 中传递与自闭症相关的时间、空间特征, 以此改善网络提取特征的能力。其中, 为了选取使 4D-CNN 具有最佳性能的卷积核, 本文对 4D 卷积核在时间维度上的深度进行了验证。根据前一层特征图的规模, 选取不同尺寸的参数矩阵作为单个时间点的权值矩阵, 并默认卷积过程中卷积核的平移步长为 1, 表 5 列出了在不同卷积核规模下 4D-CNN 的测试性能。

表 4 不同时间深度下 4D-CNN 模型的评估结果  
Table 4 Evaluation results of the 4D-CNN models at different time depths

时间深度	ACC
35	0.5832
70	0.6201

**表 5 不同卷积核规模下 4D-CNN 模型的评估结果**  
**Table 5 Evaluation results of the 4D-CNN models under different convolution kernel scales**

卷积核	ACC
1	0.5774
2	0.6369
3	0.6820

在训练阶段,使用随机梯度下降 (SGD) 进行误差反向传播更新参数,经过验证,将学习率设置为 0.01,可避免在训练过程中收敛速度过慢或陷入局部最优。此外,将最大迭代次数设置为 150。

结果显示,卷积核规模过小,只能提取单个时间点或单个时间段上的特征;随着卷积核时间维度加深,网络需要训练的参数量呈指数上升,迭代更新参数速度减缓,并出现过拟合现象,带来的后果就是网络的泛化能力大大降低。所以,本文选取时间维度上深度为 3 的卷积核作为 4D-CNN 架构的特征提取器。

### 3.2 方法比较

将提出的网络与基于 ROI(regions of interest) 构建脑功能网络 (ROI-net) 方法进行 ASD 分类性能比较,此外,几个经典的网络对本文中用到的数据集进行诊断,包括 T3D、AlexNet、2CC3D、3D-CNN 以验证 4D-CNN 的有效性。表 6 将对比方法进行了介绍。

**表 6 用于比较的方法**  
**Table 6 Methods used for comparison**

方法	概述
ROI-net <sup>[26]</sup>	用标准模板将预处理后的 fMRI 划分为若干个脑区,提取出各个脑区的时间序列,计算任意两个脑区时间序列的相关性,形成相关矩阵
T3D <sup>[12]</sup>	在 DenseNet <sup>[21]</sup> 的基础上增加时间过渡层,提高特征提取能力,并将在二维网络中预训练的权重迁移到三维网络中,以此减小训练难度
AlexNet <sup>[10]</sup>	AlexNet 使用层叠的卷积层来提取图像的特征,并使用 Dropout 和数据增强策略抑制深层网络带来的过拟合问题
2CC3D <sup>[7]</sup>	对输入的 fMRI 数据使用滑动窗口创建了 2 通道图像,并将这些图像作为 3D 卷积神经网络的输入,最终输出与 ASD 相关的 fMRI 空间特征
3D-CNN <sup>[11]</sup>	从输入视频的连续帧中提取每帧画面的二维图像特征和相邻帧之间的时间轴特征
4D-CNN	提出的 4D-CNN 方法直接对功能磁共振图像进行特征提取,充分利用时间空间特征,更好地保留了全局信息

本文所用的数据集具有四维特性,无法直接将数据集用于比较的三维卷积神经网络方法中,因此采用系统抽样的方式对数据集进行处理,以便对比方法能更好地适用于本文中的数据集。具体而言,对每个被试的功能磁共振图像在时间轴上进行等距划分,即从第一个时间点开始,每连续 10 个时间点上的图像为一组,故前 50 个时间点上的图像随之分为 5 组,然后每组随机选出一个时间点上的图像作为卷积网络的输入,以此,每个被试选出 5 个图像分别进行 3D 卷积处理,输出 5 个分类结果,并将结果中占比高的类别作为本次输出结果。

### 3.3 分类结果及分析

表 7、8、9 分别展示了提出的 4D-CNN 与其他方法在 NYU、UM\_1 和 USM 数据集上的分类结果。可以看出,4D-CNN 使用四维卷积核对自闭症的功能磁共振图像进行时空特征学习,可以取得比 T3D、AlexNet、2CC3D 及 3D-CNN 等经典的三维卷积神经网络更好的性能。表 7 的实验结果表明提出的 4D-CNN 在 NYU 数据集上的表现明显优于其他方法,分类精度达到了 70.49%,并且呈现出良好的敏感性和特异性。归结原因在于,本文所提出的 4D 卷积神经网络结构直接从预处理后的功能磁共振图像中提取时空特征,既充分保留了图像的全局信息,又很好地避免了图像经传统方法预处理后带来的误差和冗余信息。进一步验证了从 fMRI 中提取四维特征对自闭症分类任务的有效性。在多个数据集上的实验结果说明了本文所提的网络对含有时间和空间特性的数据具有适用性。

**表 7 不同方法在自闭症分类中的性能比较 (NYU)**  
**Table 7 Comparison of the performances of different methods for autism classification (NYU)**

方法	ACC	AUC	SEN	SPE
ROI-net	0.5738	0.6052	0.5385	0.6000
T3D	0.5902	0.6429	0.6154	0.5714
AlexNet	0.6557	0.7071	0.6923	0.6286
3D-CNN	0.6066	0.6571	0.5769	0.6286
2CC3D	0.6230	0.6258	0.6538	0.6000
4D-CNN	<b>0.7049</b>	<b>0.7846</b>	<b>0.7307</b>	<b>0.6857</b>

此外,用 ROC 曲线来更加直观地鉴别模型优劣,ROC 曲线能很容易地查出任意限值时模型对疾病的识别能力,图 3、4、5 分别显示了在

NYU、UM\_1 和 USM 这 3 个数据集上不同方法所获得的 ROC 曲线。可以看出, 4D-CNN 的 ROC 曲线在整体上优于其他方法, 说明本文的方法对自闭症有更好的识别能力。ROC 曲线下的面积即为 AUC 值, AUC 值能很好地描述模型整体性能的高低, 从图 3 可以看出, 在 NYU 数据集上 4D-CNN 的 AUC 值为 0.7846, 明显高于其他方法。在图 4、5 中, 4D-CNN 的 AUC 值也高于大多数方法。

表 8 不同方法在自闭症分类中的性能比较 (UM\_1)

Table 8 Comparison of the performances of different methods for autism classification (UM\_1)

方法	ACC	AUC	SEN	SPE
ROI-net	0.5833	0.5826	0.7222	0.4444
T3D	0.6389	0.6204	0.5000	0.7778
AlexNet	0.6389	0.6358	0.5556	0.7222
3D-CNN	0.5556	0.5710	0.6111	0.5000
2CC3D	0.5833	0.5617	0.6111	0.5556
4D-CNN	<b>0.6944</b>	<b>0.7160</b>	<b>0.6667</b>	<b>0.7222</b>

表 9 不同方法在自闭症分类中的性能比较 (USM)

Table 9 Comparison of the performances of different methods for autism classification (USM)

方法	ACC	AUC	SEN	SPE
ROI-net	0.5758	0.5935	0.5000	0.6316
T3D	0.5152	0.5376	0.4286	0.5789
AlexNet	0.5758	0.6053	0.6429	0.5263
3D-CNN	0.6061	0.6203	0.5714	0.6316
2CC3D	0.6364	0.6729	0.7857	0.5263
4D-CNN	<b>0.6970</b>	<b>0.6692</b>	<b>0.7143</b>	<b>0.6842</b>

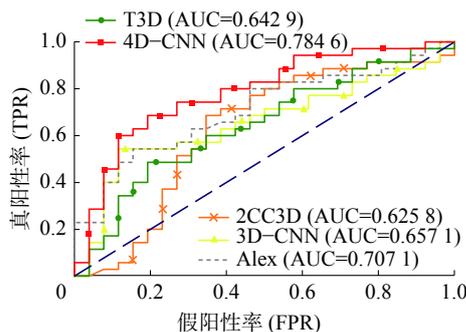


图 3 不同方法的 ROC 曲线 (NYU)

Fig. 3 ROC curve of different methods (NYU)

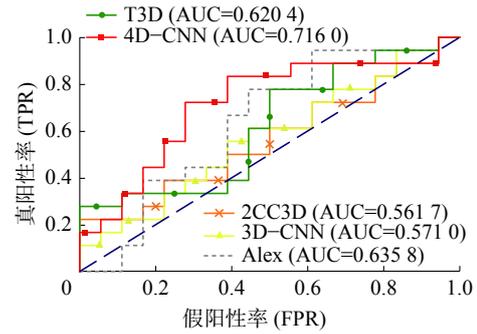


图 4 不同方法的 ROC 曲线 (UM\_1)

Fig. 4 ROC curve of different methods(UM\_1)

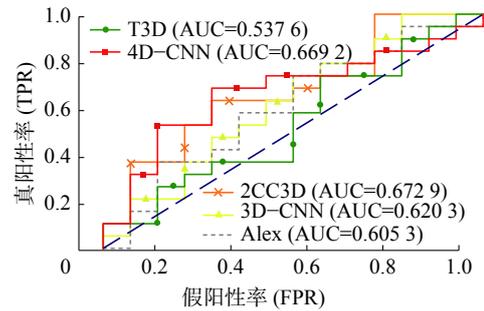


图 5 不同方法的 ROC 曲线 (USM)

Fig. 5 ROC curve of different methods (USM)

## 4 结束语

本文提出了一个四维的卷积神经网络对自闭症的功能磁共振图像进行分类。具体来说, 首先对所有的 fMRI 进行预处理, 然后通过提出的模型直接对图像进行训练。根据不断的迭代更新参数, 得到了一个具有显著识别能力的 4D-CNN。在提出的网络中, 训练过程完全保留了受试者的全局特征和连续的时间特征, 从而获得了较好的实验结果。对来自不同数据集的受试者实验表明, 所提出的 4D-CNN 方法可以使用 fMRI 数据有效地对脑疾病进行诊断。

## 参考文献:

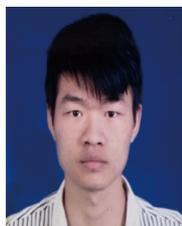
- [1] 张芬, 王穗苹, 杨娟华, 等. 自闭症谱系障碍者异常的大脑功能连接 [J]. *心理科学进展*, 2015, 23(7): 1196–1204. ZHANG Fen, WANG Suiping, YANG Juanhua, et al. Atypical brain functional connectivity in autism spectrum disorders [J]. *Advances in psychological science*, 2015, 23(7): 1196–1204.
- [2] CABALLERO-GAUDES C, REYNOLDS R C. Methods for cleaning the BOLD fMRI signal [J]. *Neuroimage*, 2017, 154: 128–149.
- [3] MOELLER S, YACCOUB E, OLMAN C A, et al. Multiband multislice GE-EPI at 7 tesla, with 16-fold acceleration using partial parallel imaging with application to high

- spatial and temporal whole-brain fMRI[J]. *Magnetic resonance in medicine*, 2010, 63(5): 1144–1153.
- [4] 常春云. 基于静息态功能磁共振成像的自闭症预测研究[D]. 北京: 北京交通大学, 2017.  
CHANG Chunyun. Research on prediction of autism spectrum disorder based on resting-state fMRI[D]. Beijing: Beijing Jiaotong University, 2017.
- [5] DVORNEK N C, VENTOLA P, PELPHREY K A, et al. Identifying autism from resting-state fMRI using long short-term memory networks[M]//WANG Qian, SHI Yinghuan, SUK H I, et al. *Machine Learning in Medical Imaging*. Cham: Springer, 2017: 362–370.
- [6] ZHAO Yu, DONG Qinglin, ZHANG Shu, et al. Automatic recognition of fMRI-derived functional networks using 3-D convolutional neural networks[J]. *IEEE transactions on biomedical engineering*, 2018, 65(9): 1975–1984.
- [7] LI Xiaoxiao, DVORNEK N C, PAPADEMETRIS X, et al. 2-Channel convolutional 3D deep neural network (2CC3D) for fMRI analysis: ASD classification and feature learning[C]//2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018). Washington, USA, 2018: 1252–1255.
- [8] KHOSLA M, JAMISON K, KUCEYESKI A, et al. Ensemble learning with 3D convolutional neural networks for functional connectome-based prediction[J]. *NeuroImage*, 2019, 199: 651–662.
- [9] LU Zhou, PU Hongming, WANG Feicheng, et al. The expressive power of neural networks: a view from the width[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach, USA, 2017: 6232–6240.
- [10] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[C]//Proceedings of the 25th International Conference on Neural Information Processing Systems. Lake Tahoe, Nevada, 2012: 1097–1105.
- [11] JI Shuiwang, XU Wei, YANG Ming, et al. 3D Convolutional neural networks for human action recognition[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2013, 35(1): 221–231.
- [12] DIBA A, FAYYAZ M, SHARMA V, et al. Temporal 3D convnets: new architecture and transfer learning for video classification [EB/OL]. (2017-11-22)[2020-9-20] <https://arxiv.org/abs/1711.08200>.
- [13] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition[J]. *Proceedings of the IEEE*, 1998, 86(11): 2278–2324.
- [14] HINTON G E, SRIVASTAVA N, KRIZHEVSKY A, et al. Improving neural networks by preventing co-adaptation of feature detectors[J]. *Computer science*, 2012, 3(4): 212–223.
- [15] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [EB/OL]// (2018-8-16)[2019-4-26]<https://arxiv.org/abs/1808.05377v1>.
- [16] SZEGEDY C, LIU Wei, JIA Yangqing, et al. Going deeper with convolutions[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston, USA, 2014: 1–9.
- [17] IOFFE S, SZEGEDY C. Batch normalization: accelerating deep network training by reducing internal covariate shift[C]//Proceedings of the 32nd International Conference on International Conference on Machine Learning. Lille, France, 2015: 448–456.
- [18] SZEGEDY C, VANHOUCKE V, IOFFE S, et al. Rethinking the inception architecture for computer vision[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, USA, 2016: 2818–2826.
- [19] SZEGEDY C, IOFFE S, VANHOUCKE V, et al. Inception-v4, inception-ResNet and the impact of residual connections on learning[C]//Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence. San Francisco, USA, 2017: 4278–4284.
- [20] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, USA, 2016: 770–778.
- [21] HUANG Gao, LIU Zhuang, VAN DER MAATEN L, et al. Densely connected convolutional networks[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, USA, 2017: 2261–2269.
- [22] EL GAZZAR A, CERLIANI L, VAN WINGEN G, et al. Simple 1-D convolutional networks for resting-state fMRI based classification in autism[C]//2019 International Joint Conference on Neural Networks (IJCNN). Budapest, Hungary, 2019: 1–6.
- [23] YAN Chaogan, WANG Xindi, ZUO Xinian, et al. DP-ABI: data processing & analysis for (resting-state) brain imaging[J]. *Neuroinformatics*, 2016, 14(3): 339–351.
- [24] HECHT-NIELSEN. Theory of the backpropagation neural network[C]//International 1989 Joint Conference on Neural Networks. Washington, USA, 1989: 593–605.
- [25] ELSKEN T, METZEN J H, HUTTER F. Neural architecture search: A survey[EB/OL]. (2018-8-16)[2020-9-

20] <https://arxiv.org/abs/1808.05377v1>.

- [26] SALVADOR R, SUCKLING J, COLEMAN M R, et al. Neurophysiological architecture of functional magnetic resonance images of human brain[J]. *Cerebral cortex*, 2005, 15(9): 1332–1342.

#### 作者简介:



郭磊, 硕士研究生, 主要研究方向为深度学习与医学图像处理。



王骏, 副教授, 博士, 主要研究方向为机器学习、模糊系统、医学影像分析。主持国家自然科学基金项目 1 项, 江苏省自然科学基金项目 1 项。2016 年获江苏省高校科研成果自然科学一等奖。获得国家发明专利 5 项, 发表学术论文 50 余篇。



王士同, 教授, 博士, 主要研究方向为模式识别、人工智能。曾获教育部、中船总公司、湖南省等省部级科技进步奖 10 项。获国务院政府特殊津贴, 省部级有突出贡献的中青年专家。发表学术论文百余篇。

## 第七届 IEEE 云计算与智能系统国际会议

2021 年 11 月 7 日至 8 日, 由中国人工智能学会与 IEEE 北京分会联合主办, 西安电子科技大学与西安国家民用航天基地管理委员会共同承办, 复旦大学、北京邮电大学、华中科技大学协办的 2021 年第七届 IEEE 云计算与智能系统国际会议将在线举办。

### 院士专家坐镇, 打造国际化权威学术平台

不论是学术研究者、企业决策者, 或是技术执行者和专业学习者, 这场会议绝对值得关注。

一方面, 大会由国家一级学会和国际组织 IEEE 共同主办, 奠定了大会的规格和量级; 另一方面, 大会邀请 CAAI 名誉理事长李德毅院士、中国科学院黄维院士、英国皇家工程院郭毅可院士、西安电子科技大学副校长石光明教授四位顶尖人工智能学者担任主席, 确保了大会的专业度和权威性。

主论坛部分, 中国科学院徐宗本院士、西安电子科技大学副校长石光明教授、CAAI 副秘书长白翔教授、微博 COO 王巍先生将分别做主题报告。分组讨论部分, 专家学者们将围绕大数据模型与数据挖掘、计算智能、计算机视觉、深度学习、机器学习、混合智能与复杂系统、神经网络应用, 进行分组讨论分享最新研究成果。

### 精选优秀论文, 活跃学术氛围促进成果交流

优秀论文征集和评选收录是大会的重要板块, 是大会学术成果沉淀的关键部分, 对活跃学术创新范围, 促进相关学科研究成果交流与传播具有极大的作用。

大会面向全球征集学术论文, 投稿文章经由程序委员会严格审核, 所有录用论文将收录至会议论文集, 并提交 IEEE Xplore 在线数据库检索, 挑选出的优秀论文扩充内容后可推荐至检索期刊。

时隔两年, 云计算与人工智能领域极具权威性的学术会议在金秋时节如约而至, 关于云计算、人工智能的核心技术和课题将在这里集中探讨。了解前沿趋势, 把握学术动态, 第七届 IEEE 云计算与智能系统国际会议不容错过。