

人工神经网络模型发展及应用综述

张 驰, 郭 媛, 黎 明

齐齐哈尔大学 计算机与控制工程学院, 黑龙江 齐齐哈尔 161000

摘 要: 人工神经网络与其他学科领域联系日益紧密, 人们通过对人工神经网络层结构的探索和改进来解决各个领域的问题。根据人工神经网络相关文献进行分析, 综述了人工神经网络算法以及网络模型结构的发展史, 根据神经网络的发展介绍了人工神经网络相关概念, 其中主要涉及到多层感知器、反向传播神经网络、卷积神经网络以及递归神经网络, 描述了卷积神经网络发展当中出现的部分卷积神经网络模型和递归神经网络中常用的相关网络结构, 分别综述了各个神经网络算法在相关领域的应用情况, 总结了人工神经网络的未来发展方向。

关键词: 人工神经网络; 多层感知器; 递归神经网络

文献标志码: A **中图分类号:** TP183 **doi:** 10.3778/j.issn.1002-8331.2102-0256

Review of Development and Application of Artificial Neural Network Models

ZHANG Chi, GUO Yuan, LI Ming

School of Computer and Control Engineering, Qiqihar University, Qiqihar, Heilongjiang 161000, China

Abstract: Artificial neural networks are increasingly closely related to other subject areas. People solve problems in various fields by exploring and improving the layer structure of artificial neural networks. Based on the analysis of artificial neural networks related literature, this paper summarizes the history of artificial neural network growth and presents relevant principles of artificial neural networks based on the development of neural networks, including multilayer perceptron, back-propagation algorithm, convolutional neural network and recurrent neural network, explains the classic convolutional neural network model in the development of the convolutional neural network and the widely used variant network structure in the recurrent neural network, reviews the application of each artificial neural network algorithm in related fields, summarizes the possible direction of development of the artificial neural network.

Key words: artificial neural network; multilayer perceptron; recurrent neural network

人工神经网络(Artificial Neural Networks, ANN)是一种可用于处理具有多个节点和多个输出点的实际问题的网络结构。虽然人类的大脑和人工神经网络的运用都具有极其强大的信息处理能力, 但是两者还是有许多不同之处。谷歌Deepmind最初被Demis Hassabis、Mustafa Suleyman以及Shane Legg创立出来, 在2016年创造出AlphaGo打败世界围棋冠军李世石后逐渐被人认可, 也说明人工神经网络具有巨大的潜力。与人脑处理信息方式有所不同, 运用人工神经网络开发出的机器人采用线性的思维方式处理获取到的信息, 计算机通过快速、精确的顺序数值运算, 在串行算术类型的任务处理上超过人类。但人脑的“并行处理体系”相对于人工神经网络领域具有绝对领先的能力。

McCulloch心理学家和Pitts数学家于1943年考虑寻找神经元背后的基本原理, 将阈值函数作为计算神经元的主要特性, 把逻辑演算表述为神经计算架构, 提出“神经网络”概念和M-P模型, 标志着人工神经网络ANN萌芽^[1]。Hebb假设突触权重的变化会如何控制神经元相互激励的方式, 在1949年出版的《行为的组织》中提出了Hebb突触以及Hebb学习规则, 为人工神经网络算法的发展构建了理论知识基础^[2]。20世纪60年代末, Rosenblatt开创了感知器, 感知器是建立在M-P模型基础上, 第一个物理构建并形成了具有学习能力的人工神经网络^[3]。Minsky和Papert在1969年出版*Perceptrons: an introduction to computational geometry*, 提出Rosenblatt的单层感知器只能学习线性可分模式, 无法处理

基金项目: 国家自然科学基金(61872204); 黑龙江省属高等学校基本科研业务费专项(135309462)。

作者简介: 张驰(1996—), 男, 硕士研究生, 研究方向为计算生物学与生物信息学; 郭媛(1974—), 女, 博士, 教授, 研究方向为计算机算法在光学检测及信息处理中的应用; 黎明(1982—), 通信作者, 女, 硕士, 副教授, 研究方向为计算生物学与生物信息学, E-mail: fionalee629@163.com。

收稿日期: 2021-02-23 **修回日期:** 2021-04-01 **文章编号:** 1002-8331(2021)11-0057-13

xor 等线性不可分问题^[4]。1984 年 Hopfield 神经网络 (Hopfield Neural Network, HNN) 首次引入, 从此基于 Hopfield 神经网络的动力学行为的理解应用于信息处理和工程研究起到了至关重要的作用^[5]。反向传播网络 (Backpropagation Neural Network, BPNN) 之后被提出用于解决多层神经网络所反应出来的问题^[6], 但是 BP 网络仍存在一部分缺点, 比如: 收敛速度慢以及大样本数据难以收敛, 容易出现局部最小化。1998 年 Lecun 等基于福岛邦彦提出的卷积和池化网络结构, 将 BP 算法运用到该结构的训练中, 形成了卷积神经网络 (Convolutional Neural Network, CNN) 的雏形 LeNet-5^[7]。2006 年由 Hinton、Osindero 和 Teh 提出深度信念网络 (Deep Belief Network, DBN)^[8]。近几年, 人工神经网络在各个领域都是非常热门的话题, 并且在图像处理、医学、生物学等领域均取得了非常多的成就。

本文针对人工神经网络领域中的几个模型 (多层感知器 (Multilayer Perceptron, MLP)、反向传播神经网络、卷积神经网络、递归神经网络 (Recursive Neural Network, RNN)) 基本结构进行介绍, 并对其相对热门的应用进行简单的概述。

1 多层感知器

多层感知器, 又称为多层前馈神经网络, 如图 1, 具有出色的非线性匹配和泛化能力。训练 MLP 使用反向传播算法, 可以减少 MLP 输出数据与实际所需数据之间的全局误差。

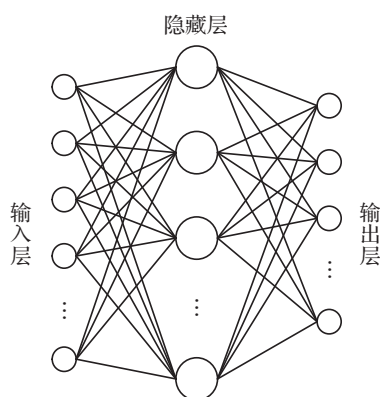


图1 MLP结构示意图

由于 MLP 具有非常好的非线性映射能力、较高的并行性以及全局优化的特点, 现今在图像处理、预测系统、模式识别等方面取得了不错的成就。

尽管 MLP 架构具有很多优点, 但在高维空间下的效率相对低下, 可能导致模型训练中过拟合的情况。并且由于隐藏层的存在加大了超参数的数量, 使得训练过程中在收敛缓慢的情况下需要处理很高的计算量。传统的 MLP 实值模型中单个神经元能接收的数据输入为单个实数, 在其进行多维信号输入时, 通常达不到令人

满意的效果。宫金良等人使用传统 MLP 对路面进行识别时发现高楼区域作为道路障碍并不能被模型识别出来, 处理结果的准确度和完整性上存在缺陷, 在经过 MLP 处理之前加入感兴趣区域 (Region of Interest, ROI), 通过预先规定的面积阈值对符合特征的区域进行提取, 能够排除图片数据中的杂质^[9]。此外, MLP 的神经元个数尚没有标准的方法来确定, 目前常用的交叉验证复杂度较高, 并且受限于数据量。陈小威等人提出 MLP 的泛化能力不足, 运用于大量气象数据时, 经遗传算法特征处理后的观测点 1 数据训练得到 MLP, 在气象观测点 2 相较于观测点 1 的精度大幅度下降, 所以 MLP 的泛化能力仍待提高^[10]。

针对 MLP 存在的问题, 提出了不同的改进算法。García-Pedrajas 等人提出一种广义多层感知器 (Generalized Multilayer Perceptron, GMLP) 的协同进化^[11]。模型基于模块的不同子群体进行协作, 每个子群体都是广义的多层感知器。与标准的多层感知器相比, 基于 GMLP 的网络结构具有相对较少的节点和连接数, 可以使用更少的节点定义非常复杂的表面。同时, 较小的网络进行演进能够提高网络的可解释性。受到大脑中神经胶质特征的启发, Ikuta 等人提出一种具有神经胶质网络的多层感知器, 其中神经胶质网络仅与 MLP 的第二层隐藏层进行连接, 通过计算机仿真结果证实具有神经胶质网络的 MLP 相对于标准的 MLP 具有更好的性能, 赋予了 MLP 中神经元的位置依赖性^[12]。Li 等提出一种基于简化几何代数 (Reduced Geometric Algebra, RGA) 的多层感知器扩展模型 RGA-MLP, 传统 MLP 模型将每个维度的信号视为一个实数进行单独处理, 基于 RGA 的模型中输入、输出、激活函数以及运算符都使用可交换乘法规则扩展到 RGA 域, 并且使用 RGA 版本的反向传播训练神经网络, 用于多维信号处理, 将多个通道视为一个单元而不是一个单独的组件, 可以实现更高的分类精度、更快的收敛速度以及更低的计算复杂度^[13]。Masulli 和 Penna 将基于主成分分析的增量输入维度 (IID) 算法应用于 MLP 中, 提高了多层感知器的学习速率^[14]。

Martinez-Morales 等人提出通过多目标蚁群优化算法对 MLP 参数进行优化的 MLP-MOACO 模型, 对发动机污染物相关系数进行计算以及估算发动机的废气排放^[15]。Mosavi 等人提出 MLP-GWO 模型, 该模型将 Gray Wolf 算法与标准 MLP 模型结合在一起并应用于土壤电导率预测, 实验结果证明混合 MLP-GWO 模型相对于标准 MLP 模型可以在隐藏层获取更加准确的连接权重, 从而提高预测精度^[16]。Liu 等基于 Adaboost (自适应 Boosting) 算法和 MLP (多层感知器) 神经网络, 提出了四种不同的混合方法用于高精度多步风速预测, 证明了 Adaboost 算法能有效提高 MLP 神经网络的性能^[17]。

2 BP神经网络

BP神经网络模型(反向传播算法)的网络体系结构是多层的,本质上是一种梯度下降局部优化技术,与网络权重的向后误差校正相关。

BP神经网络的多层结构使得模型的输出更加准确,但BP神经网络仍然存在一定的缺陷。针对XOR之类的非线性可分问题时,使用BP神经网络可能出现局部最小值导致无法找到全局最优解,并且在面对大样本数据时均方误差MSE过大导致难以收敛。王丽红等人将传统BP组合起来构成AdaBoost-BP模型,如图2,AdaBoost算法训练计算首个BP模型的错误率及权重,并将该权重作为下个BP网络的权重参数,依此类推进行迭代计算,其中单个传统BP网络隐藏层采用2层结构,该方法运用于短期销量预测时平均预测误差为18.89%,相较于传统BP网络的53.23%准确率有明显的改善,然而该模型在面对样本数据时间跨度较大的情况时误差偏大,仅在5天的样本数据下能有效地预测近期的销量变化^[18]。

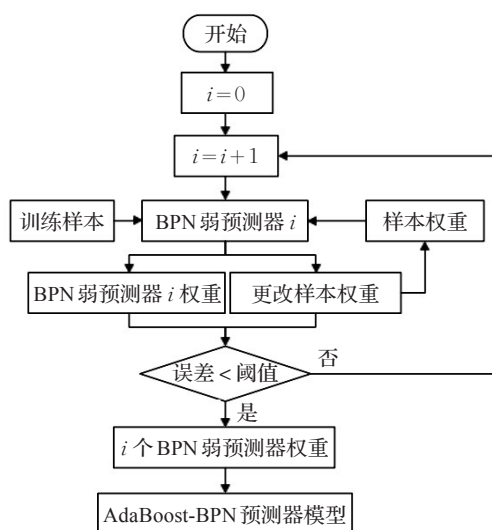


图2 AdaBoost-BP网络模型流程图

针对BP网络使用梯度下降容易使模型陷入局部最优的缺陷,黄宝洲等人改变传统BP调整自身阈值和权重参数的方式,使用粒子群优化算法获取BP网络的权重和阈值参数,选择种群大小40,进行150次迭代,在预测波浪波高以及波峰波谷实验结果中准确率虽然会随着预测步数的增加而下降,但相对于单一BP网络接近于真实数据^[19]。Ng等人提出一种改进的反向传播算法GBP,对激活函数的偏导进行修改从而增加反向传播的误差信号,对算法的学习率进行归一化,以改善和加快收敛速度,计算机仿真结果证实提出的广义BP网络(Generalized Backpropagation, GBP)相较于原始BP网络的全局收敛速度快得多,在回归问题中使用0.5的学习率和0.7的动量,传统模型需要进行6 000次以上的迭

代,但使用6个隐藏单元的GBP只需要平均1 471次迭代^[20]。Silaban等人将BFGS准牛顿应用于BP神经网络,使用经典反向传播和BFGS反向传播在玻璃数据集上测试了7种架构,提出具有BFGS的BP神经网络改善了学习过程的收敛性,平均改善收敛率为98.34%,并且当BFGS与BP一起使用时,准确率会有所提高^[21]。

Zhu等人结合Rumelhart的添加惯性脉冲动态调整学习速率,可以将学习率调整为较大值从而提升模型学习速度,并通过学习率进而影响节点的连接阈值和权重参数,消除BP网络学习过程中的无效迭代,将两者结合提出用以优化BP网络的改进算法,改善了传统BP网络局部最小值以及收敛速度慢的情况,用以确定隶属度函数并应用于蒸汽凝结设备的模糊诊断^[22]。Sang将BP神经网络与遗传算法优化的支持向量机对比分析,应用于供应链金融的信用风险评估,结果表明BP神经网络可以更好地识别和分类不同级别的企业,分类精度较高^[23]。Miao等人提出将反向传播算法和RBP算法结合用于模糊建模,可以良好地避免模糊化过程中存在的死区以及决策面突变等潜在问题^[24]。Jin等基于四川省某电网企业输变电投资项目的财务效益分析提出一种改进的模糊最优BP神经网络模型,使用IGSA算法(Improved Gravitational Search Algorithm, IGSA)优化模糊最优BP神经网络,实验结果表示BP神经网络的成本效益分析可以预测项目在一定范围内的变化中达到收支平衡的条件和财务收益^[25]。

3 卷积神经网络

卷积神经网络(CNN)是机器学习中常见的模型结构,在图像分类识别、语义分割、机器翻译等方面取得了良好的效果。传统CNN结构包括四种层结构:卷积层、池化层、全连接层、输出层。

卷积神经网络在图像领域的应用非常广泛并且取得了很大的成就。特别是在图像识别方面,经过卷积神经网络一系列运算,机器可以非常准确地识别图像特征信息。Lou等人将VGG16结合卷积神经网络CNN应用于人脸识别,同时收集丢弃的图像信息并将其应用于原始CNN,得到的改进模型相对于ICA算法以及传统的卷积神经网络等明显提升了性能以及图像的识别率^[26]。Zhang等人将CNN用于故障检测,传统的故障检测第一步处理信号,然后将特征放入分类器进行分类,使用基于卷积神经网络CNN的智能诊断检测算法将原始的输入信号转化得到二维的图像数据,并消除经验对特征提取的干扰,然后通过轴承数据验证了该算法的有效性,实验证实了该算法能很好地适应工作负载的变化^[27]。Deng等人提出了一种新型的带有可变形模块的基于区域的CNN(R-CNN)裂纹检测器,通过对比Faster

R-CNN、R-FCN 以及基于 FPN 的 Faster R-CNN, 得出由于 FPN 多尺度特征提取使得基于 FPN 的 Faster R-CNN 表现出最高的分类精度, 该模型可用于提高表面裂纹的检测精度^[28]。Cao 等人设计了三种深层 CNN 结构, 用于从原始音频片段和梅尔谱图中学习可分辨的情感特征, 利用合并的深度网络将两个异构的神经网络结合在一起, 利用了每个网络的优势, 最终提高了整体性能^[29]。

Mask R-CNN 在 Faster R-CNN 的基础上进行了改进, 是最著名的图像分割模型之一, 如图 3^[30], 其中 RPN 提取特征图中大小不一致的 RoI, 再对 RoI 做规范化处理, 使用 RoIAlign 替换 Faster R-CNN 中的 RoIPool 保证输入输出的对应关系。宋绍剑等人将 Mask R-CNN 应用于水下生物识别, 使用 mini-mask 替换原始的 mask 以节省内存, 但 100×100 的 mini-mask 相对于 $1\,024 \times 1\,024$ 的 mask 损失了像素精度, 导致水下轮廓模糊, 但最终测试结果准确率为 97.30%, 在与原始 mask 的结果无明显结果差异的情况下, 很大程度节约了内存^[30]。Hu 等人将 CNN 用于图像诊断, CT 图像由专家进行手动标记后, 使用提出的 Mask R-CNN 对 CT 图像中的肺部进行自动切割。实验中将 Mask R-CNN 与 K-means 结合得到了相较于其他方法最佳的分割效果, 分割精度达 $97.68\% \pm 3.42\%$, 平均运行时间 11.2 s^[31]。Hyojin 等人将 Mask R-CNN 用于 BDD 数据集提取可驾驶区域的模型, 以支持自动驾驶系统, 在 BDD 数据集上的训练证明了准确性高于现有的数据集 MS COCO^[32]。

传统 CNN 模型在图像领域虽然有广泛的应用, 但其劣势也较为明显, 较深层的卷积网络提取的更偏重局部信息, 并且深层卷积的计算量需求大, 输入图片大小固定, 在嵌入式设备应用方面具有局限性。王亚朝等人将基于卷积神经网络的方法应用于天气信息识别时, 发现对比支持向量机、随机森林等传统方法准确率提升, 但计算量的巨大导致模型退化, 并且由于数据源图片尺寸大小不一, 使用 CNN 进行训练时需对数据进行压缩裁剪, 导致损失部分重要信息, 影响模型识别率^[33]。CNN 的特征提取效果适合分类, 但传统 CNN 的最后一层分类效果相较于其他分类器如支持向量机、极限学习机等并不优秀。余丹等人使用 6 层卷积神经网络 (2 层

卷积、2 层子采样), 将原始 CNN 的最后一层转化成极限学习机, 在人脸识别中识别率达到约 97%, 并且在大数据下较与其他方法占优^[34]。

下面以几个经典的卷积神经网络模型进行展开。

3.1 LeNet-5

LeNet-5 网络结构并不是全连接网络, LeCun 等人使用多个卷积核, 采用卷积核权值共享的方法减少卷积神经网络中的连接数, 模型更加简洁易于计算。其网络体系由七层结构组成。

传统 LeNet-5 的效率较低, 单通道的网络结构进行特征提取时不完整并且模型收敛速率慢, 安源等人采用四通道网络, 对四个通道的卷积核和偏置参数进行设置, 采用 ReLU 激活函数, 在 MNIST 数据集上四通道模型准确率为 96.56%, 比传统 LeNet-5 高出 4.52%, 但这种多通道结构目前不能对数据规模进行动态调整^[35]。Hou 等人提出使用 FPGA 加速 LeNet-5 来改进原始的 LeNet-5 模型, 最后通过训练手写数字识别模型实验, 证实了改进后的模型在效率和准确度上都有很大的提升^[36]。针对滚动轴承故障诊断中传统 LeNet-5 网络识别准确率较低, 模型收敛速率慢, 泛化能力弱的问题, Wan 等人提出了一种采用改进的二维 LeNet-5 网络的滚动轴承故障诊断方法, 调整卷积核数量以及大小并执行批归一化, 全连接层除最后一层外均进行删除操作, 增强网络泛化能力, 实验表明改进方法具有更高的故障诊断精度和更少的训练时长^[37]。为了保护 Lenet-5-like 结构的卷积神经网络, Sun 等人提出在不同的层中添加保护锁模块, 通过在 C3 层、C1 & C3 层、C1 & C5 层和 C1 & C3 & C5 层中添加保护锁定模块, 得到了最好的预测结果为 90.26%^[38]。

3.2 AlexNet

AlexNet 在 LetNet 的基础上更进一步加深了网络结构, 是一个 5+3 的卷积神经网络, 包括 5 个卷积层、3 个全连接层, 如图 4。

AlexNet 虽然相比于传统 CNN 在图像领域具有更高的识别率、图像质量提高, 但由于其对特征进行提取时使用的卷积核不具有多样性, 在进行图像识别时仍然存在一定的误差。郭书杰等人在使用 AlexNet 对手势识别时提出 AlexNet 的非线性激活函数会导致训练过程中出现神经元死亡, 因此设计了包含三个批归一化的

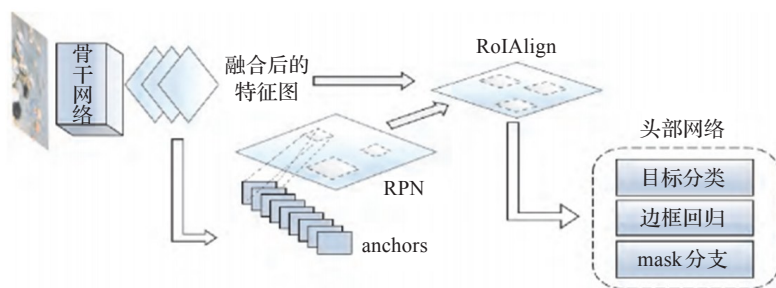


图3 Mask R-CNN 模型图

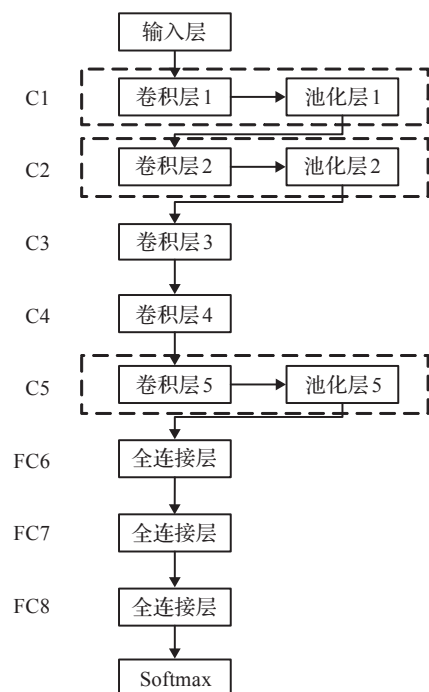


图4 AlexNet示意图

AlexNet结构(针对3、4、5层做批归一化)并且优化了模型超参数,结构优化后的AlexNet准确率提高了约4%,但该模型和原始模型同样受限于输入图像的复杂性和手势在图像面积中的占比^[39]。黄方亮等人提出了AlexNet_En模型,该模型在原始AlexNet的第四层后添加了一层与第四层相同的卷积操作,采用384个3×3卷积核确实增加了模型的准确率,在ImageNet数据集上达到94.00%,但该结构同样使模型复杂度变高,计算需求增加带来了一定的硬件负担^[40]。

Tao等人提出了一种改进的AlexNet,使用空间金字塔池化模式的网络结构解决了原网络模型适应性差的问题,对网络深度和整个连接层进行了调整,优化了整个网络的性能。实验结果表明,与公共和专用手指静脉数据集中的AlexNet模型相比,改进的网络模型在识别准确性和训练持续时间上都有显著改善^[41]。

AlexNet在场景分类应用中,传统AlexNet卷积核跨度大导致特征图的分辨率下降过快,Xiao等人提出一种改进的AlexNet模型,将大卷积核分解为两个步幅较小的小卷积核级联结构,实验证明改进模型在23种场景分类中的分类精度高于原始的AlexNet模型^[42]。Han等人提出一种改进的预训练AlexNet体系结构AlexNet-SPP-SS,结合了比例池-空间金字塔池(SPP)和边监督(SS)来改善原始AlexNet不收敛以及过拟合的问题,并证明了经过预训练的AlexNet-SPP-SS模型优于原始的AlexNet体系结构以及传统的场景分类方法^[43]。

3.3 VGG-16

VGG-16也是一个经典的卷积神经网络模型,其块由小型堆叠式卷积滤波器制成,如图5。已经显示出它

们具有与较大的卷积滤波器相同的效果,但是它们使用较少的参数。

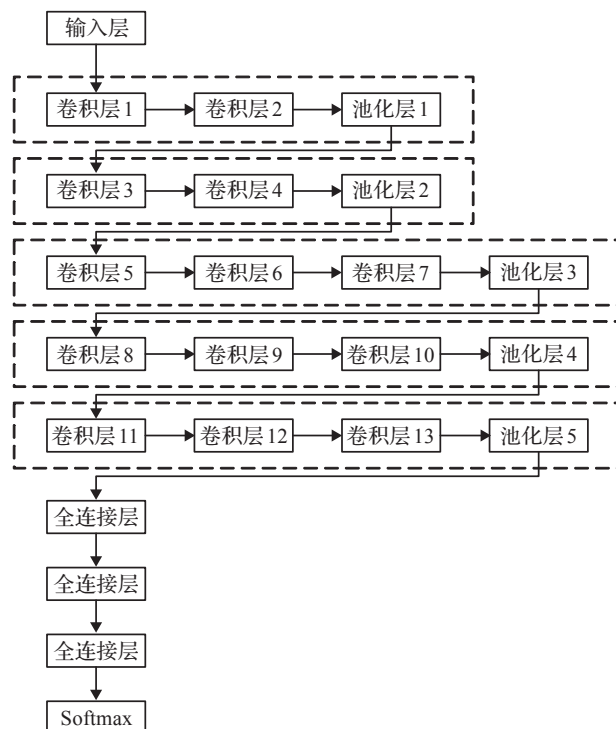


图5 VGG-16示意图

VGG-16在图像分类识别领域取得许多成就,为了提高蔬菜自动分类的准确性,Li等人利用开源的Gaffe深度学习框架,采用改进的VGG网络模型训练蔬菜图像,在批归一化层添加了VGG-M网络,提高了网络的收敛速度和准确度,增加批归一化成熟提升接近三倍的收敛速度,增大数据集训练规模提升模型的泛化能力,训练不同数量的蔬菜图像实验结果表明相对于传统VGG以及AlexNet准确率有了很大提升^[44]。徐志京等人将VGG16的网络全连接层改进得到简化的模型Reduce-VGGNet(R-VGGNet),用于青光眼眼底图像数据的识别,结果表明改进后的R-VGGNet准确率达到91.7%,相对于LeNet等提高了准确率^[45]。

Zhao等人根据无人驾驶汽车交通标志识别问题的实时性和准确性要求,对VGG神经网络结构进行了改进,提出VGG-8模型并改进了VGG-16模型,通过SGD和Nesterov Momentum优化网络,应用于视频中十个交通标志时VGG-8具有更高的准确率和运行速度^[46]。

VGG-16这种网络结构的深层次使得模型训练容易出现过拟合的问题,数据输入量大导致模型参数过多,传统VGG的参数占用空间约500 MB,这使得内存资源的占用过多可能会加大计算性能上的负担。鉴于VGG网络模型的结构参数主要分布在全连接层,在改进的时候将全连接层做优化成为一种方式,吉鑫等人将VGG-16模型的末端全连接层使用稀疏编码器代替,稀疏化隐藏层节点,在2 950幅图像数据中进行5次迭代,得到准

确率为92%,损失函数值相对较小^[47]。张旭欣等人在对卷积神经网络运用FPGA加速平台时,由于FPGA的内存容量远小于VGG网络模型的资源占用,故提出了二值VGG网络模型,其中激活函数和权重参数均使用符号函数取值,占用内存511.3 MB降低到了3.66 MB,在基于FPGA平台下的二值VGG网络相较于其他VGG加速器提速7倍,识别率为81%,该优化方式验证有效解决了资源负担但识别率可能仍有上升空间^[48]。

3.4 GoogLeNet

GoogLeNet相对于VGG、AlexNet等网络小、参数较少,性能相对优越,在ImageNet挑战赛(ILSVRC14)中取得第一名。GoogLeNet使用Inception网络结构,保持神经网络的稀疏性并且提高了性能。

GoogLeNet复杂性较高,针对小规模数据时可能无法达到大规模数据集所能达到的性能,Zhu等人提出一种新颖的双重微调策略来训练GoogLeNet模型,通过截断操作优化GoogLeNet的结构减小网络大小,用于极端天气识别,在天气数据集上进一步微调得到最后的模型,优化后的模型大小为原始GoogLeNet的31.23%,但识别准确率从94.74%提升至95.46%,识别速度也有所提高^[49]。

Bi等人将GoogLeNet应用于手写汉字识别,提出一种改进的GoogLeNet模型,在原始模型中添加了批归一化层,极大地提高了网络识别能力并且减少了训练时间^[50]。Xie等人为了减少稀疏CT图像中的伪影并且提升图像质量,将残差学习应用于GoogLeNet得到了一种新颖的改进GoogLeNet模型,结果表明该模型对于减少伪影和保持重构图像的质量是有效的^[51]。

Tuan使用预训练的三个神经网络AlexNet、GoogLeNet和SqueezeNet,并对神经网络进行了微调,用于COVID-19、病毒性肺炎和正常胸部X射线图像的分类,从不同性能指标的训练和测试数据中证实模型的有效性^[52]。

使用传统的GoogLeNet深层次网络结构做特征提取时可能会因为感受野扩大而导致特征消失,对准确率产生影响。要提升传统模型的性能,加大深层网络的深度和宽度会很大程度增加参数量,产生更大的计算负担,一般对传统模型进行结构优化。传统GoogLeNet包含9个Inception模块,张泽中等人在对胃癌病理图像提取特征时以Inception模块为单位进行实验,发现在第7个模块模型取得最优性能,最终保留前7个Inception, GPU占用由传统的65%降为43%,训练时间少了约4小时,第7个模块后衔接全卷积网络对特征分类输出,30次迭代后模型准确率为99.28%,但模型在提高灵敏度的前提下损失了部分特异性^[53]。孙中杰等人在张泽中的基础上对模型做了进一步改进,分块测试后删除了Inception_4e后的模型结构,之后直接衔接平均池化和

全连接层输出分类结果,见图6,在病理数据集上精简模型相较于传统GoogLeNet准确率上升约4%^[54]。

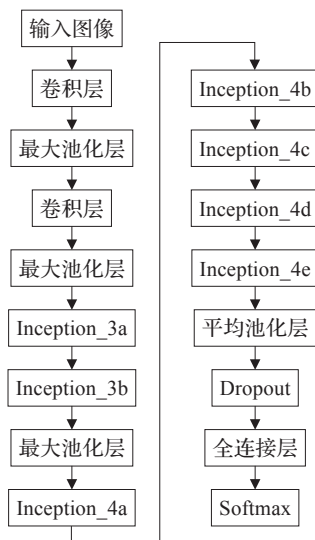


图6 精简的GoogLeNet示意图

3.5 ResNet

残差神经网络(ResNet)在多图像处理问题中能获取高精度的输出结果。其主要功能称为跳跃连接,有助于梯度流动,如图7。残差神经网络中He等人利用多层的神经网络结构来拟合残差映射的效果,从而解决加深神经网络深度导致的梯度消失以及精度下降等问题^[55]。

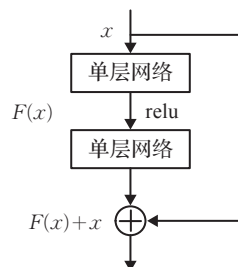


图7 ResNet跳跃连接

经典ResNet还存在很大的改进空间,残差单元中通过最终梯度所包含的梯度信息无法直接计算出其他梯度信息,导致残差单元增多时出现更多的卷积层无法获得梯度信息传递。李国强等人提出FCM-Resnet,提出跨层连接将所有卷积层与平均池化层相连,使每个残差单元都能传递梯度信息,在对比FCM-Resnet-56、FCM-Resnet-110和传统Resnet实验结果后提出的改进模型准确率为99.57%和99.63%,上升了约0.03%和0.02%,改进模型的稳定性和优化还存在改进空间^[56]。使用 1×1 卷积核来解决ResNet50输入输出数据维度不匹配时,在细颗粒图像分类领域会丢失信息并且影响计算结果,李晓双等人把跳跃连接中步长为2的卷积核替换为步长1,并在卷积操作前加入了平均池化,一定程度上保留了梯度信息,仅在小样本下证明了模型优化有效^[57]。

Deng 等人提出一种改进的 ResNet 模型,用于改进智能机械故障诊断算法的诊断准确性以及网络深度增加带来的操作速度下降等问题,设计多尺度特征融合模块提取多尺度故障特征信息,利用改进残差块提升了模型的计算速度^[58]。

Xia 等人以残差网络(ResNet)为主干,提取不同级别的语义信息,采取一种改进的多孔空间金字塔池化方法来提取多尺度深度语义信息提高网络处理边界信息的能力,然后通过不同规模的全球注意力上采样机制将深层语义信息与浅层空间信息融合在一起,从而提高了网络利用全局和局部特征的能力。在 Sentinel-2 卫星和陆地遥感卫星(Landsat)图像上的实验结果表明,该方法的分割精度和速度均优于现有方法,对实现实际的云影分割具有重要意义^[59]。

Wu 等人利用残差网络(ResNet)、双向门控单元(BiGRU)和注意力机制提出一种基于神经网络和主动学习(DABot)的新浪微博社交机器人检测框架,经过性能评估后,DABot的精度为0.988 7,说明该模型更加有效^[60]。

将五种经典卷积网络的优缺点以及适用场景的对比,总结为表1。

3.6 全卷积神经网络

全卷积网络(Fully Convolutional Networks,FCN)最初在语义分割的应用中取得了较为满意的成果,Long 等人在2015年提出该模型并对当前热门的 AlexNet、VGG net 以及 GoogLeNet 与全卷积网络相结合进行微调^[61]。全卷积神经网络将传统卷积神经网络的全连接操作全部用卷积操作来替代,FCN 同样具备传统卷积网络的特点,可以接受任意大小的输入数据,更加高效。但 FCN 经过上采样获得的结果中无法获取图像的详细信息,像素之间的相关性无法得到利用并且 FCN 缺乏先验知识约束,为了解决这些问题,He 等人将边缘信息作为先验知识引入 FCN,利用整体嵌套的边缘检测(HED)检测到的边缘信息校正 FCN 结果,提出一种 Edge-FCN,在 ESAR 和 GID 数据集上的实验结果证明该模型是有效的^[62]。Zhao 等人将基于边界项(BSLIC)

的简单线性迭代聚类(SLIC)与全卷积网络 FCN 结合在一起,使用 FCN 语义分割结果注释 BSLIC 获取的超像素区域,可以准确地识别图像中目标语义信息,并且在小边缘定位上也具有非常高的精度^[63]。最终在 PASCAL VOC 2012 数据集上的实验结果表明相对于传统的 FCN,改进算法明显提高了分割精度。

FCN 的应用领域不仅仅局限于图像处理,在视频领域也取得了一定的成果。Wang 等人将 FCN 应用于视频中进行动作估计,提出了一种混合全卷积网络(H-FCN),该网络由两部分组成,分别为外观 FCN 和运动 FCN,利用外观和运动线索来检测动作^[64]。最终,使用 Stanford40 Actions、UCF Sports 以及 JHMDB 三个数据集进行实验,结果证实了 H-FCN 的有效性。Fang 等人提出了一种新颖多模块全卷积网络(MM-FCN),用于立体视频的视觉注意力预测,该网络由三个模块组成,分别为 S-FCN、T-FCN 和 D-FCN,通过 S-FCN 和 T-FCN 学习到的显著性信息可以很大程度提高 D-FCN 的预测结果^[65]。实验结果证实 MM-FCN 在大规模立体视频人眼注视数据库上,用于预测立体视频的视觉注意力方面具有良好的前景。

FCN 存在一定局限性。传统模型由多个卷积层堆叠,卷积层堆叠过多会引起计算和优化困难,当需要大尺寸卷积核时,FCN 模型的复杂度会明显增大,计算量也相应加大。在与其他技术(如:空洞卷积、编解码、ASPP 等)融合时,会使参数量及计算复杂度明显上升^[66]。陈纪铭等人在使用 FCN 对视频数据做异常检测时发现分块速率过慢,无法训练出大样本等问题^[67]。杨朔等人在使用 FCN 对绿潮浓度图像进行分类时发现传统 FCN 的高倍上采样会使图像分类结果轮廓细节缺失,并且容易产生大面积缺口^[68]。未来全卷积网络应该更加注重于模块化设计而非仅仅改变网络宽度深度来提升性能,如何设计出更加轻量级并且高效的网络仍是重要方向。

4 递归神经网络

递归神经网络 RNN 是一种基于序列建模的人工神

表1 五种模型对比

名称	机制	特点	缺陷	适用场景
LeNet	卷积核权值共享	结构相对简单	效率较低,识别率低,收敛慢,泛化能力弱	最初用做图像规律识别任务,不建议用于复杂问题处理
AlexNet	ReLU 激活函数、Dropout 方法	复杂度低,参数少,表征能力强,节省时间	卷积核尺寸单一,模型优化进程缓慢,容易梯度消失	处理图像数据识别分类预测等任务
VGGNet	缩小卷积核尺寸,加深网络层次	泛化性能好,拓展性好	容易过拟合,参数量大,占用资源多	处理图像数据,适用于对精度要求较高的任务,不建议用于资源瓶颈的设备
GoogLeNet	Inception 单元模块	稀疏性,模块化,更好地提取上下文信息	占用资源大	针对微观复杂图像处理效果较好
ResNet	跳跃连接	准确率较高,解决梯度消失,训练时间短	反向传播中梯度减小	适用于深层次网络解决模型退化

神经网络,可以在神经元之间横向传输数据信息,并且部分表达数据之间的相关性,如图8。RNN的特点在于隐藏层中神经元相互连接,从而可以顺序传递滑动窗口中与时间有关的输入信息,并且可以考虑时间维度上距离较远的事件之间的时间相关性。

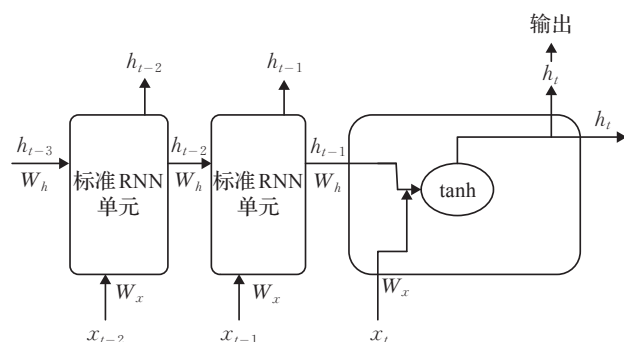


图8 RNN结构示意图

递归神经网络RNN的隐藏层结构使其在时间序列预测方面具有非常广泛的应用,Wei等人将MLP、RNN、LSTM、GRU分别应用于孔隙水压力(PWP),得出具有RNN结构的模型在针对时间序列数据时相较于MLP更为准确,尤其是LSTM和GRU可以描述输入与响应之间的时滞效应,相对于标准RNN更加精确可靠^[69]。Ling等人将RNN应用于核动力机械的故障预测,提出一种智能故障预测方法,将主成分分析PCA降维后的数据传递给完整的RNN模型,根据转速和振动信号分别提前60 h和44 h生成警报。实验结果表明,RNN模型可以有效地识别蠕变期间的故障^[70]。

Stender等人将CNN与RNN结合使用,用于刹车噪声检测和预测,发现结合模型可以克服传统方法的局限性,第一部分采用CNN显示出了优越的检测质量和特征提取性能,第二部分采用的RNN依赖于噪声的瞬时频谱特性,使用该模型预测刹车噪声的精度和准确度都非常高,该模型在声音检测方面展现出巨大的潜力^[71]。Bai等人将RNN用于智能图书馆,通过读者借阅记录以及借阅行为来预测读者在不同阶段的感知需求,为读者提供购书和个性化服务,模型中使用CRFID与RNN结合,通过图书馆记录的实际数据证实该模型在不同阶段感知读者需求是可行的^[72]。

传统RNN会产生梯度消失,并且在处理数据长期依赖时精度会大幅度下降,输入输出数据序列不匹配,模型的参数共享引起的缺失信息可能对时序特征产生影响,RNN将每个节点的先前隐藏状态进行编码作为整个模型的历史信息,但是忽略了每个节点之间的独立关系^[73]。针对传统模型的局限性,提出了LSTM和GRU等改进方案。

4.1 LSTM

1997年Hochreiter和Schmidhuber在标准RNN中引入门控单元概念,解决了标准RNN存在的梯度消失问

题^[74]。LSTM单元由遗忘门、输入门和输出门组成,如图9。通过这三个门可以使LSTM单元具有维持内存状态并且有选择地记住或遗忘信息的能力,传播过程中的无关信息将被遗忘丢弃。

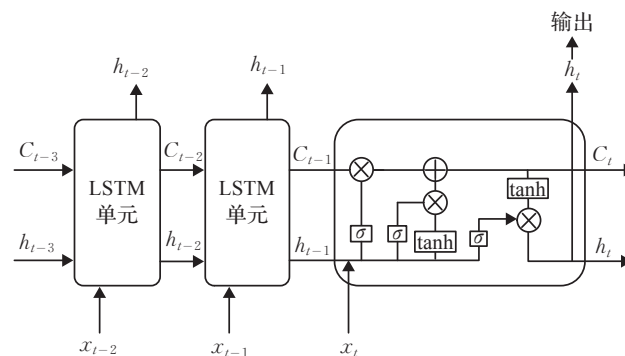


图9 LSTM结构示意图

对于时频信息的获取,LSTM相对于之前在该领域广泛应用的卷积神经网络更加合适。Wang等人将LSTM应用于语音增强,提出一种LSTM-卷积-BLSTM编解码器网络(LCLED),包含了转置卷积和跳跃连接,使用两个LSTM单元对上下文信息进行捕获,使用卷积层对频域特征进行提取,在多种噪音的情况下该网络模型仍具有良好的降噪功能,在语音增强方面具有更高的鲁棒性^[75]。Ma等人传统LSTM中引入卷积运算,提出一种CLSTM学习算法提取时频信息,并且通过卷积获取特征,通过堆叠CLSTM可以构建用于RUL预测的深层框架以提取更深层的信息,将其应用于滚珠轴承RUL预测,与现有模型相比,由于卷积运算导致训练时间略微增加但总体模型性能大大提高^[76]。Petmezas等人将LSTM与CNN结合提出CNN-LSTM模型应用于手动心电图(ECG)中,通过CNN将提取到的ECG信号特征传递给LSTM以实现时间动态记忆,从而更为准确地分类四种ECG类型^[77]。最终使用该模型在MIT-BIH心房颤动数据上进行训练,采用十折交叉验证了该模型能准确验证ECG类型(灵敏度为97.87%,特异性为99.29%),可以帮助临床医生实时检测常见类型的房颤。赵红蕊等人将LSTM与CNN结合用于股票价格预测并引入注意力机制(Convolutional Block Attention Module, CBAM),提出一种LSTM-CNN-CBAM混合模型,对比实验结果验证了在LSTM-CNN结合模型中加入CBAM模块的可行性^[78]。

Yu等人将LSTM应用于非线性系统建模,通常非线性建模使用的训练方法是时间反向传播BPTT,但是BPTT的速度较慢,所以提出一种改进的深度LSTM,结合了LSTM和多层感知器的优势,使用Lyapunov函数验证了该训练方法的稳定性,同时实验证明该模型针对非线性系统建模优于现有其他模型^[79]。

LSTM相对于RNN改进了梯度消失问题,但是在使

用小规模数据集时效果不够理想,模型的复杂度提升了,训练时间变长,效率相对低下。孙陈影等人在使用LSTM对语音数据集进行分类时,耗时60.62 min,RNN的2倍^[80]。余雅文等人对比了传统克里金方法和LSTM分别对自由空气重力异常的估计结果,虽然LSTM在准确度和稳定性上更好,但传统方法耗时小于1 s而LSTM则超过120 s,在效率方面LSTM还存在很大的提升空间^[81]。如果将LSTM应用于时间跨度比较大的数据时,会产生非常巨大的计算量和时长。黄婷婷等人将LSTM应用于金融时间序列预测,实验发现LSTM预测结果具有滞后性^[82]。

4.2 GRU

Cho等人提出的门控循环单元神经网络是LSTM的简化形态^[83]。GRU将输出门和遗忘门耦合为更新门,重置门对应LSTM的输入门,与LSTM相似的是,GRU也保留现有信息并在现有信息内容的基础上添加经过过滤的信息,模型具有存储功能。不同的是,GRU将LSTM中的内存控制剔除,简化了LSTM的计算量,如图10。

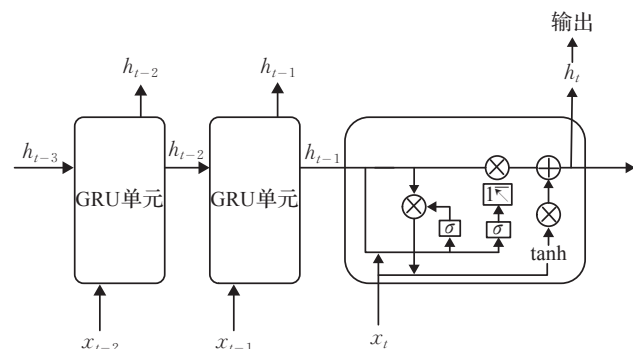


图10 GRU结构示意图

GRU简化LSTM的结构使得参数量减少,大幅缩短了训练时间。Liu等人使用门控循环单元GRU代替神经编程器解释器NPI中的LSTM从而改变NPI的核心结构,在确保精度相同的情况下,使用GRU结合的NPI相对于原始的基于LSTM的NPI性能提升了大约33%^[84]。Elsayed等人基于LSTM与全卷积网络LSTM-FCN应用于时间序列分类的分类结果,使用门控循环单元与全卷积网络相结合得到GRU-FCN,提出使用GRU在具有较小体系结构以及较小计算量的时间序列分类问题中,具有更高的分类准确性,更简单的硬件实现^[85]。

GRU在时序数据中的应用较为广泛,同时与处理高维数据的网络结构相结合,如:CNN等,能够处理更加复杂的问题。Wu等人将GRU与CNN结合,提出一种GRU-CNN混合神经网络模型,其中GRU部分负责提取时序数据的特征向量,CNN提取高维数据的特征向量,该模型应用于改善电力系统的短期负荷预测(STLF)中,对比BP神经网络、标准GRU以及标准CNN等预测方法,该模型能更好地处理时序数据并且提取数据集特征^[86]。Pan等利用GRU-CNN结合模型应用于长

江的水位预测问题,通过30年的长江水位数据,同时对比分析证实该模型优于小波神经网络(WANN)、LSTM以及统计学整合移动平均自回归模型ARIMA^[87]。桂智明等人通过CNN和GRU提取交通流的时空特征,提出卷积门控循环单元预测模型(ACGRU)对交通流数据的时空特征进行预测,在真实交通流数据集上的实验证实了该模型较于其他模型误差降低了约9%,具有更高的预测精度^[88]。

Tjandra等人提出一种基于张量分解方法的改进GRU模型TT-GRU,模型中使用CP分解和Tucker分解来表示权重矩阵,相比于标准GRU模型以及单纯使用CP分解的GRU模型和仅使用Tucker分解的模型,TT-GRU具有更好的性能,使用TT格式的低秩张量表示权重矩阵能够显著压缩参数量,同时保持模型准确性和性能表现^[89]。

GRU模型虽然在结构上相比LSTM较为简单,同样也解决了RNN梯度消失问题,但是单向GRU结构只能获取到某一时刻之前的历史信息,而无法获取到前后关系信息,在某些领域并不能满足实际需求。双向门控循环单元(Bidirection Gated Recurrent Unit, BiGRU)是将前向和后向两个GRU层输出结果拼接得到BiGRU的输出结果,见图11,前向由上文向下文学习数据间信息,后向相反。骆楠等人将BiGRU应用于树脂质量预测,实验最终模型拟合优度为0.982,预测精度优于GRU^[90]。BiGRU通常与注意力机制技术结合可以提高模型准确率,万子云等人将模型应用于MOOC平台检测作弊行为,CNN提取局部特征、BiGRU提取时序特征,结合注意力机制取得了98.51%的准确率,AUC为91.07%,但是BiGRU的模型收敛速度相对较慢,均大于1 ms^[91]。程琪苓等人同样将CNN与BiGRU结合构成编码器,应用于跨站脚本检测,提升了检测性能和分类效果^[92]。BiGRU模型可以满足训练过程中单向GRU需要获取数据间完整性关系信息的缺陷,但多个GRU进行拼接的结构同样容易导致过拟合,使得整体性能下降,并且训练后的网络泛化能力较低,不能适应测试数据。

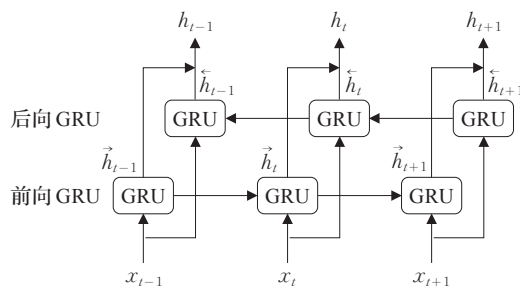


图11 BiGRU结构示意图

5 结语

近年来,人工智能领域较为火热,神经网络也进入到各大领域的视野,并且取得了不同程度的结合和

表2 四种人工神经网络模型对比

名称	机制	特点	缺陷	适用场景
MLP	各层神经元全连接	非线性映射、并行性高、全局优化	泛化能力不足、处理多维数据效果差	可以处理回归任务,应用于模式识别、优化计算,不建议做复杂问题处理
BP	梯度下降局部优化	非线性映射、自适应能力	易出现局部最优、预测精度较低、收敛较慢	结合粒子群、退火算法等优化算法应用于模式识别控制、数据建模等
CNN	卷积核特征提取,降采样保留信息	稀疏连接、权值共享	计算量大,输入图片尺寸固定	适合用于空间数据,应用于图像处理、机器翻译、语义分割等
RNN	隐藏层节点输出取决于当前节点输入和上个节点值	提取时序特征能力强、泛化能力相对好	输入输出序列不同、处理长期依赖精度下降	用于处理时序数据,应用于语音识别、自然语言理解等

发展。本文对人工神经网络发展中的部分经典网络算法进行了简单的梳理和概述,并对相关的研究应用做以总结,包含了理论基础以及人工神经网络一些算法的发展、改进研究以及算法应用,主要对多层感知器 MLP、BP 神经网络算法、卷积神经网络 CNN 以及递归神经网络四部分做以阐述,分别介绍了各个模型的优缺点,见表 2。

虽然人工神经网络目前应用非常广泛,但目前还面临许多问题仍需解决。

(1)目前人工神经网络中多通道结构模型(如:多通道卷积网络、多通道 LSTM 等)的通道数量设置比较固定,不能根据数据集进行动态调整。

(2)虽然传统模型在实验分析应用中都具有不错的准确率和能效,但是通过人工向模型添加特定噪声后,轻微的干扰会使精度严重下降,所以进一步提升人工神经网络模型的泛化能力,仍是人工神经网络模型发展的方向。

(3)面向深层次网络结构进行改进时,拓展模型的深度和宽度可能导致相反的效果,未来的工作可以围绕设计具有可区分特征的模块化结构,建立更加适应实际需求的模型。

(4)在使用卷积神经网络进行图像识别时,复杂图像或者识别物体占有较小面积时,识别精度低下,未来可以围绕发展使用弱监督或无监督学习对图像数据进行标注,归一化后再进行识别。

(5)目前面向递归神经网络的超参数调整还没有标准的算法,手动调参意味着模型精度受制于研究者经验或者费时费力的调试,超参数优化仍是一项需要解决的问题。

参考文献:

- [1] MCCULLOCH W S, PITTS W A. A logical calculus of the ideas immanent in nervous activity[J]. Bulletin of Mathematical Biophysics, 1943, 5: 115-133.
- [2] HEBB D O. The organization of behavior: a neuropsychological theory[M]. New Jersey: Lawrence Erlbaum Associates, 1949.
- [3] ROSENBLATT F. The perceptron: probabilistic model for

information storage and organization in the brain[J]. Psychological Review, 1958, 65(6): 386-408.

- [4] MINSKY L, PAPERT S A. Perceptrons: an introduction to computational geometry[M]. Cambridge: MIT Press, 1969.
- [5] HOPFIELD J J. Neurons with graded response have collective computational properties like those of two-state neurons[J]. Proc Natl Acad Sci, 1984, 81(10): 3088-3092.
- [6] RUMELHART D E, HINTON G, WILLIAMS R J. Learning representations by back-propagating errors[J]. Nature, 1986, 323(6088): 533-536.
- [7] 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.
- [8] HINTON G E, OSINDERO S, TEH Y A. A fast learning algorithm for deep belief nets[J]. Neural Computation, 2006, 18: 1527-1554.
- [9] 宫金良, 孙晓峰, 张彦斐. 多层感知器自监督在线修正的道路识别算法[J]. 交通运输系统工程与信息, 2019, 19(4): 101-107.
- [10] 陈小威, 朱文越, 钱仙妹, 等. 基于人工神经网络的近地面光学湍流估算[J]. 光学学报, 2020, 40(24): 15-21.
- [11] GARCÍA-PEDRAJAS N, ORTIZ-BOYER D, HERVAS-MARTINEZ C. Cooperative coevolution of generalized multi-layer perceptrons[J]. Neurocomputing, 2004, 56: 257-283.
- [12] IKUTA C, UWATE Y, NISHIO Y. Multi-layer perceptron with glial network for solving two-spiral problem[J]. IEICE Transactions on Fundamentals of Electronics Communications and Computer Sciences, 2011, E94A(9): 1864-1867.
- [13] LI Y P, CAO W M. An extended multilayer perceptron model using reduced geometric algebra[J]. IEEE Access, 2019, 7: 129815-129823.
- [14] MASULLI F, PENNA M. Improving learning speed in multilayer perceptrons through principal component analysis[J]. Proceedings of the SPIE, 1996, 2760: 85-95.
- [15] MARTINEZ-MORALES J, QUEJ-COSGAYA H, LAGUNAS-JIMENEZ J, et al. Design optimization of multilayer perceptron neural network by ant colony optimization applied to engine emissions data[J]. Science China-

- Technological Sciences, 2019, 62(6): 1055-1064.
- [16] MOSAVI A, SAMADIANFARD S. Predicting soil electrical conductivity using multi-layer perceptron integrated with grey wolf optimizer[J]. Journal of Geochemical Exploration, 2021, 220: 1-10.
- [17] LIU H, TIAN H Q, LI Y F, et al. Comparison of four Adaboost algorithm based artificial neural networks in wind speed predictions[J]. Energy Conversion and Management, 2015, 92: 67-81.
- [18] 王丽红. 基于BP-AdaBoost的电商短期销量预测模型[J]. 计算机系统应用, 2021, 30(2): 260-264.
- [19] 黄宝洲, 杨俊华, 卢思灵, 等. 基于改进粒子群优化神经网络算法的波浪捕获功率预测[J]. 太阳能学报, 2021, 42(2): 302-308.
- [20] NG S C, LEUNG S H, LUK A. Fast convergent generalized back-propagation algorithm with constant learning rate[J]. Neural Processing Letters, 1999, 9: 13-23.
- [21] SILABAN H, ZARLIS M, SAWALUDDIN. Analysis of accuracy and epoch on back-propagation BFGS quasi-Newton[C]//Proceedings of International Conference on Information and Communication Technology, Medan, 2017.
- [22] ZHU Q, LI J. Applied technology on artificial neural network in fault diagnosis system[C]//Proceedings of the 2nd International Conference on Civil Engineering and Material Engineering, Wuhan. Switzerland: Trans Tech Publications, 2014: 448-452.
- [23] SANG B. Application of genetic algorithm and BP neural network in supply chain finance under information sharing[J]. Journal of Computational and Applied Mathematics, 2021, 384: 1-11.
- [24] MIAO Z N, ZHENG H J. An improved back-propagation algorithm for fuzzy modeling[C]//Proceedings of the 3rd International Conference on Measuring Technology and Mechatronics Automation, Shanghai. Switzerland: Trans Tech Publications, 2011: 198-202.
- [25] JIN X, LIU Q, LONG H Z. Impact of cost-benefit analysis on financial benefit evaluation of investment projects under back propagation neural network[J]. Journal of Computational and Applied Mathematics, 2021, 384: 1-15.
- [26] LOU G X, SHI H Z. Face image recognition based on convolutional neural network[J]. China Communications, 2020, 17(2): 117-124.
- [27] ZHANG J Q, SUN Y, GUO L, et al. A new bearing fault diagnosis method based on modified convolutional neural networks[J]. Chinese Journal of Aeronautics, 2020, 33(2): 439-447.
- [28] DENG L, CHU H H, SHI P, et al. Region-based CNN method with deformable modules for visually classifying concrete cracks[J]. Applied Sciences-Basel, 2020, 10(7): 1-18.
- [29] CAO R, GUO K H, MA J H, et al. A deep convolutional neural network-based label completion and correction strategy for supervised medical image learning[C]//Proceedings of IEEE SmartWorld, Guangzhou. New York: IEEE, 2018: 1725-1730.
- [30] 宋绍剑, 朱靖旭. 基于Mask R-CNN和迁移学习的水下生物目标识别研究[J]. 计算机应用研究, 2020, 37(S2): 386-388.
- [31] HU Q H, SOUZA L F D, HOLANDA G B, et al. An effective approach for CT lung segmentation using mask region-based convolutional neural networks[J]. Artificial Intelligence in Medicine, 2020, 103: 1-36.
- [32] HYOJIN J, SOOSUN C. Brief paper: drivable area detection with region-based CNN models to support autonomous driving[J]. Journal of Multimedia Information System, 2020, 7(1): 41-44.
- [33] 王亚朝, 赵伟, 徐海洋, 等. 基于轻量型卷积神经网络的非固定场景天气识别算法[J]. 电子测量技术, 2019, 42(17): 152-156.
- [34] 余丹, 吴小俊. 一种卷积神经网络和极限学习机相结合的人脸识别方法[J]. 数据采集与处理, 2016, 31(5): 996-1003.
- [35] 安源, 刘春, 蔡朝晖, 等. 基于改进的LeNet-5网络在单通道图像分类中的研究[J]. 信息技术, 2020, 44(12): 8-10.
- [36] HOU Y, CHEN Z B. LeNet-5 improvement based on FPGA acceleration[C]//Proceedings of the 3rd Asian Conference on Artificial Intelligence Technology, Chongqing. London: Journal of Engineering, 2020: 526-528.
- [37] WAN L J, CHEN Y W, LI H Y, et al. Rolling-element bearing fault diagnosis using improved LeNet-5 network[J]. Sensors, 2020, 20(6): 1693-1716.
- [38] SUN L, WANG Y H, DAI L Y. Convolutional neural network protection method of lenet-5-like structure[C]//Proceedings of the 2nd International Conference on Computer Science and Artificial Intelligence, Shenzhen. New York: ACM, 2018: 77-80.
- [39] 郭书杰. 基于改进AlexNet的复杂背景手势识别方法研究[J]. 大连交通大学学报, 2021, 41(6): 95-99.
- [40] 黄方亮, 沈同平, 金力. 改进的AlexNet卷积神经网络用于中草药叶片分类[J]. 安庆师范大学学报(自然科学版), 2020, 26(3): 70-75.
- [41] TAO Z Y, HU Y L, LIN S. Finger vein recognition based on improved AlexNet[J]. Laser & Optoelectronics Progress, 2020, 57(8): 1-9.
- [42] XIAO L S, YAN Q, DENG S Y. Scene classification with improved AlexNet model[C]//Proceedings of the 12th International Conference on Intelligent Systems and Knowledge Engineering, Nanjing. New York: IEEE ISKE, 2017: 1-6.
- [43] HAN X B, ZHONG Y F, CAO L Q. Pre-trained AlexNet

- architecture with pyramid pooling and supervision for high spatial resolution remote sensing image scene classification[J]. Remote Sensing, 2017, 9(8): 1-22.
- [44] LI Z B, LI F, ZHU L, et al. Vegetable recognition and classification based on improved VGG deep learning network model[J]. International Journal of Computational Intelligence Systems, 2020, 13(1): 559-564.
- [45] 徐志京, 汪毅. 青光眼眼底图像的迁移学习分类方法[J]. 计算机工程与应用, 2021, 57(3): 144-149.
- [46] ZHAO Q Y, SHEN Y L, ZHANG Y. Video-based traffic sign detection and recognition[C]//Proceedings of the 2nd International Conference on Image and Video Processing, Shanghai. Bellingham: SPIE-Int Soc Optical Engineering, 2019: 1-8.
- [47] 吉鑫, 陈剑云, 完颜幸幸. 用于接触网可视化接地识别的改进 VGG-16 模型[J]. 传感器与微系统, 2021, 40(2): 58-72.
- [48] 张旭欣, 张嘉, 李新增, 等. 二值 VGG 卷积神经网络加速器优化设计[J]. 微电子技术, 2021, 47(2): 20-23.
- [49] ZHU Z Q, LI J F, ZHUO L, et al. Extreme weather recognition using a novel fine-tuning strategy and optimized GoogLeNet[C]//Proceedings of International Conference on Digital Image Computing-Techniques and Applications, Sydney. New York: IEEE, 2017: 839-845.
- [50] BI N, CHEN J H, TAN J. The handwritten Chinese character recognition uses convolutional neural networks with the GoogLeNet[C]//Proceedings of International Conference on Pattern Recognition and Artificial Intelligence, Montreal, 2019.
- [51] XIE S P, ZHANG P C, LUO L M, et al. Sparse-view CT reconstruction with improved GoogLeNet[C]//Proceedings of SPIE Conference on Medical Imaging-Biomedical Applications in Molecular, Houston. Bellingham: SPIE-Int Soc Optical Engineering, 2018.
- [52] TUAN D P. Classification of COVID-19 chest X-rays with deep learning: new models or fine tuning?[J]. Health Information Science and Systems, 2021, 9(2): 1-11.
- [53] 张泽中, 高敬阳, 吕纲, 等. 基于深度学习的胃癌病理图像分类方法[J]. 计算机科学, 2018, 45(11A): 263-268.
- [54] 孙中杰, 万涛, 陈东, 等. 深度学习在主动脉中膜变性病理图像分类中的应用[J]. 计算机应用, 2021, 41(1): 280-285.
- [55] HE K, ZHANG X, REN S, SUN J. Deep residual learning for image recognition[C]//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas. New York: IEEE, 2016: 770-778.
- [56] 李国强, 张露. 全卷积多并联残差神经网络[J]. 小型微型计算机系统, 2020, 41(1): 30-34.
- [57] 李晓双, 韩立新, 李景仙, 等. 基于优化残差网络的多模态音乐情感分类[J]. 计算机与现代化, 2020(12): 83-89.
- [58] DENG F Y, DING H, YANG S P, et al. An improved deep residual network with multiscale feature fusion for rotating machinery fault diagnosis[J]. Measurement Science and Technology, 2021, 32(2): 1-13.
- [59] XIA M, WANG T, ZHANG Y H, et al. Cloud/shadow segmentation based on global attention feature fusion residual network for remote sensing imagery[J]. International Journal of Remote Sensing, 2021, 42(6): 2022-2045.
- [60] WU Y H, FANG Y Z, SHANG S K, et al. A novel framework for detecting social bots with deep neural networks and active learning[J]. Knowledge-Based Systems, 2021, 211: 1-16.
- [61] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Boston. New York: IEEE, 2015: 3431-3440.
- [62] HE C, LI S L, XIONG D H J, et al. Remote sensing image semantic segmentation based on edge information guidance[J]. Remote Sensing, 2020, 12(9): 1-20.
- [63] ZHAO W, ZHANG H D, YAN Y J, et al. A semantic segmentation algorithm using FCN with combination of BSLIC[J]. Applied Sciences-Basel, 2018, 8(4): 1-15.
- [64] WANG L M, QIAO Y, TANG X O, et al. Actionness estimation using hybrid fully convolutional networks[C]//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition, Seattle. New York: IEEE, 2016: 2708-2717.
- [65] FANG Y M, ZHANG C, HUANG H Q, et al. Visual attention prediction for stereoscopic video by multi-module fully convolutional network[J]. IEEE Transactions on Image Processing, 2019, 28(11): 5253-5265.
- [66] 章琳, 袁非牛, 张文睿, 等. 全卷积神经网络研究综述[J]. 计算机工程与应用, 2020, 56(1): 25-37.
- [67] 陈纪铭, 陈利平. 一种优化 FCN 的视频异常行为检测定位方法[J]. 重庆邮电大学学报(自然科学版), 2021, 33(1): 126-134.
- [68] 杨朔, 陈丽芳, 石瑀, 等. 基于深度生成式对抗网络的蓝藻语音分割[J]. 计算机应用, 2018, 38(6): 1554-1561.
- [69] WEI X, ZHANG L L, YANG H Q, et al. Machine learning for pore-water pressure time-series prediction: application of recurrent neural networks[J]. Geoscience Frontiers, 2021, 12(1): 453-467.
- [70] LING J, LIU G J, LI J L, et al. Fault prediction method for nuclear power machinery based on Bayesian PPCA recurrent neural network model[J]. Nuclear Science and Techniques, 2020, 31(8): 1-11.
- [71] STENDER M, TIEDEMANN M, SPIELER D, et al. Deep learning for brake squeal: brake noise detection, characterization and prediction[J]. Mechanical Systems and Signal Processing, 2021, 149: 1-27.
- [72] BAI R Q, ZHAO J M, LI D G, et al. RNN-based demand

- awareness in smart library using CRFID[J].China Communications,2020,17(5):284-294.
- [73] 李大舟,于广宝,高巍,等.基于时序卷积网络的词级语言模型研究与应用[J].计算机工程与设计,2021,42(2):449-454.
- [74] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation, 1997, 9(8), 1735-1780.
- [75] WANG Z Y, ZHANG T, SHAO Y Y, et al. LSTM-convolutional-BLSTM encoder-decoder network for minimum mean-square error approach to speech enhancement[J]. Applied Acoustics, 2021, 172: 1-7.
- [76] MA M, MAO Z. Deep-convolution-based LSTM network for remaining useful life prediction[J]. IEEE Transactions on Industrial Informatics, 2021, 17(3): 1658-1667.
- [77] PETMEZAS G, HARIS K, STEFANOPOULOS L, et al. Automated atrial fibrillation detection using a hybrid CNN-LSTM network on imbalanced ECG datasets[J]. Biomedical Signal Processing and Control, 2021, 63: 1-9.
- [78] 赵红蕊,薛雷.基于LSTM-CNN-CBAM模型的股票预测研究[J].计算机工程与应用,2021,57(3):203-207.
- [79] YU W, GONZALEZ J, LI X O. Fast training of deep LSTM networks with guaranteed stability for nonlinear system modeling[J]. Neurocomputing, 2021, 422: 85-94.
- [80] 孙陈影,沈希忠. LSTM和GRU在城市声音分类中的应用[J].应用技术学报,2021,20(2):158-164.
- [81] 余雅文,付广裕.基于循环神经网络的重力异常数据推估研究[J].大地测量与地球动力学,2021,41(3):234-237.
- [82] 黄婷婷,余磊. SDAE-LSTM模型在金融时间序列预测中的应用[J].计算机工程与应用,2019,55(1):142-148.
- [83] CHO K, VAN MERRIENBOER B, GULCEHRE C, et al. Learning phrase representations using RNN encoder decoder for statistical machine translation[J]. arXiv: 1406. 1078, 2014.
- [84] LIU W, WANG Q, ZHU Y L, et al. GRU: optimization of NPI performance[J]. Journal of Supercomputing, 2020, 76(5): 3542-3554.
- [85] ELSAYED N, MAIDA A S, BAYOUMI M. Gated recurrent neural networks empirical utilization for time series classification[C]//Proceedings of IEEE Int Congress on Cybermat, Atlanta. New York: IEEE, 2019: 1207-1210.
- [86] WU L Z, KONG C, HAO X H, et al. A short-term load forecasting method based on GRU-CNN hybrid neural network model[J]. Mathematical Problems in Engineering, 2020, 2020: 1-10.
- [87] PAN M Y, ZHOU H N, CAO J Y, et al. Water level prediction model based on GRU and CNN[J]. IEEE Access, 2020, 8: 60090-60100.
- [88] 桂智明,李壮壮,郭黎敏.基于ACGRU模型的短时交通流预测[J].计算机工程与应用,2020,56(21):260-265.
- [89] TJANDRA A, SAKTI S, NAKAMURA S. Recurrent neural network compression based on low-rank tensor representation[J]. IEICE Transactions on Information and Systems, 2020, E103D(2): 435-449.
- [90] 骆楠,祁佳康,罗娜.基于双向门控循环单元神经网络的间歇过程最终产品质量预测[J].华东理工大学学报(自然科学版),2020,46(6):807-814.
- [91] 万子云,陈世伟,秦斌,等.基于深度学习的MOOC作弊行为检测研究[J].信息安全学报,2021,6(1):32-39.
- [92] 程琪苓,万良.改进编码-解码框架下的跨站脚本检测[J].计算机工程与设计,2021,42(1):44-50.