

---

智能科学与工程学院

## 论文写作指导

(卷积神经网络研究综述)

学 号: S321040137

专 业: 控制科学与工程

学生姓名: 姚宗汛

任课教师: 李娟 教授

2021 年 11 月

文章题目	卷积神经网络研究综述
创 新 点 自 述	<p>本文整理和总结了国内外关于深度学习的发展历程和最新的研究成果,对人工神经网络及经典的卷积神经网络所涉及到的概念和算法进行了简要介绍,将卷积神经网络算法进行了改进并应用。</p> <p>1、在 LeNet-5 网络模型的基础上进行改进,构造了若干各层具有不同神经元个数和层间连接方式的特征抽取滤波器层的卷积神经网络模型,将各个模型应用到光学数字识别问题上,通过这些不同的卷积神经网络模型在实验中学习过程表现出的特性和识别性能分析比较各种模型的优劣。</p> <p>2、通过借鉴自适应增强 (Adaboost) 的思想,构建了一个多列卷积神经网络模型,并将其应用在交通标示识别实际应用问题中,将数据进行预处理,训练卷积神经网络,实现卷积神经网络对交通标示的高性能识</p>

# 卷积神经网络研究综述

姚宗汛<sup>1</sup>

(1 哈尔滨工程大学智能科学与工程学院, 黑龙江 哈尔滨 150000)

**摘要:** 作为一个十余年来快速发展的崭新领域, 深度学习受到高度关注, 它在特征提取和建模上都有着相较于浅层模型显然的优势。深度学习善于从原始输入数据中挖掘越来越抽象的特征表示, 而这些表示具有良好的泛化能力。深层神经网络是目前的主要形式, 其神经元间的连接模式受启发于动物视觉皮层组织, 而卷积神经网络则是其中一种经典而广泛应用的架构。卷积神经网络的局部连接、权值共享及池化操作等特性使之可以有效地降低网络的复杂度, 减少训练参数的数目, 且也易于训练和优化。基于这些优越的特性, 它在各种信号和信息处理任务中的性能优于标准的全连接神经网络。该文首先概述了卷积神经网络的发展历史, 然后分别描述了神经元模型、多层感知器的结构。还分别介绍了卷积神经网络的监督学习、无监督学习训练方法。

**关键词:** 卷积神经网络, 深度学习, 人工智能

Doi:

中图分类号: TP24

文献标识码: A

文章编号:

## Summary of Convolutional Neural Network Research

YAO ZongXun<sup>1</sup>

(College of Intelligent Systems Science and Engineering, Harbin Engineering University, Harbin 150000, China)

**Abstract:** As a brand-new field that has been developing rapidly for more than ten years, deep learning has received high attention. It has obvious advantages over shallow models in feature extraction and modeling. Deep learning is good at mining more and more abstract feature representations from raw input data, and these representations have good generalization ability. Deep neural network is the main form at present, and the connection mode between neurons is inspired by the organization of animal visual cortex. And convolutional neural network is one of the classic and widely used structures. The characteristics of local connection, weight sharing and pooling operation of convolutional neural network can effectively reduce the complexity of the network, reduce the number of training parameters, and it is also easy to train and optimize. Based on these superior characteristics, its performance in various signal and information processing tasks is better than standard fully connected neural networks. This article first summarizes the development history of convolutional neural networks, and then describes the structures of neuron models and multilayer perceptrons. It also introduces the supervised learning and unsupervised learning training methods of convolutional neural networks.

**Keywords:** Convolutional neural network, deep learning, artificial intelligence

<sup>1</sup>人工神经网络<sup>[1]</sup> (Artificial Neural Network) 是对生物神经网络的一种模拟和近似, 是由大量神经元通过相互连接而构成的自适应非线性动态网络系统。1943 年, 心理学家和数理逻辑学 Pitts 提出了神经元的第 1 个数学模型——MP 模型<sup>[2]</sup>。MP 模型具有开创意义, 为后来的研究工作提供了依据。到了 20 世纪 50 年代末至 60 年代初, Rosenblatt 在 MP 模型的基础之上增加了学习功能, 提出了单层感知器模型, 第一次把神经网络的研究付诸实践。但是

单层感知器网络模型不能处理线性不可分问题。直至 1986 年, Rumelhart 等人提出了一种按误差逆传播算法训练的多层前馈网络——反向传播网络 (Backpropagation network), 解决了原来一些单层感知器所不能解决的问题。由于在 20 世纪 90 年代<sup>[3]</sup>, 各种浅层机器学习模型相继被提出, 较经典的如支持向量机, 而且当增加神经网络的层数时传统的 BP 网络会遇到局部最优、过拟合及梯度扩散等问题, 这些使得深度模型的研究被搁置。

收稿日期: 2021-11-26.

基金项目:

作者简介: 姚宗汛(1998-), 男, 硕士研究生.

通信作者: 姚宗汛, yaozx@hrbeu.edu.cn

# 1 CNN 概述

## 1.1 神经元

- 2 神经元<sup>[4]</sup>是人工神经网络的基本处理单元，一般是多输入单输出的单元，其结构模型如图 1 所示其中： $x_i$  表示输入信号； $n$  个输入信号同时输入神经元  $j$ ， $w_{ij}$  表示输入信号  $x_i$  与神经元  $j$  连接的权重值， $b_j$  表示神经元的内部状态即偏置值， $y_j$  为神经元的输出

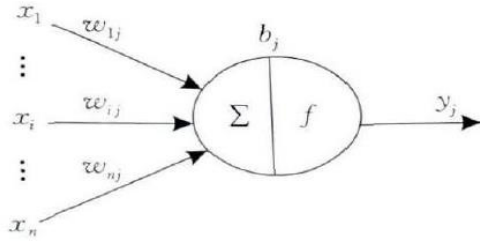


图 1 神经元模型  
Fig.1 Neuron model

## 1.2 多层感知机

多层感知器 (Multilayer perceptron) 是由输入层、隐含层 (一层或者多层) 及输出层构成的神经网络模型，它可以解决单层感知器不能解决的线性不可分问题输入层神经元接收输入信号，隐含层和输出层的每一个神经元与之相邻层的所有神经元连接，即全连接，同一层的神经元间不相连，且每个连接都有一个连接权值。隐含层和输出层中每一个神经元的输入为前一层所有神经元输出值的加权和。

当多层感知器用于分类时，其输入神经元个数为输入信号的维数，输出神经元个数为类别数，隐含层个数神经元个数视具体情况而定<sup>[6]</sup>但在实际应用中，由于受到参数学习效率影响，一般使用不超过 3 层的浅层模型。BP 算法可分为两个阶段：前向传播和后向传播，其后向传播始于 MLP 的输出层。

## 1.3 CNN 算法

1962 年，生物学家 Hubel 和 Wiesel 通过对猫脑视觉皮层的研究<sup>[7]</sup>，发现在视觉皮层中存在一系列复杂构造的细胞，这些细胞对视觉输入空间的局部区域很敏感，它们被称为“感受野”。感受野以某种方式覆盖整个视觉域，它在输入空间中起局部作用，因而能够更好地挖掘出存在于自然图像中强烈的局部空间相关性。将这些被称为感受野的细胞分为简单细胞和复杂细胞两种类型。在视觉皮层<sup>[8]</sup>中的神经网络有一个层级结构：外侧膝状体→简单细胞→复杂细胞→低阶超复杂细胞→高阶超复杂细胞。低阶超复杂细胞与高阶超复杂细胞之间的神经网络结构类似于简单细胞和复杂细胞间的神经网络结构。在该层级结构中，处于较高阶段的细胞通常会有这样一个倾向：选择性地响应刺激模式<sup>[9]</sup>更复杂的特征；同时还具有一个更大的感受野，对刺激模式位置的变化更加不敏感<sup>[10]</sup>。1980 年，Fukushima 根据前者的层级模型提出了结构与之类似的神经认知机。神经认知机采用简单细胞层和复杂细胞层

交替组成，其中 S 层级模型中的简单细胞层或者低阶超复杂细胞层相对应，层对应于复杂细胞层或者高阶超复杂细胞 C 层。S 层能够最大程度地响应感受野内的特定边缘刺激，提取其输入层的局部特征，C 层对来自确切位置的刺激具有局部敏感性。尽管在神经认知机中没有像 BP 算法那样的全局监督学习过程可利用，但它仍可认为是 CNN 的第一个工程实现网络，卷积和池化 (也称作下采样) 分别受启发于前者概念的简单细胞和复杂细胞，它能够准确识别具有位移和轻微形变的输入模式。随后,LeCun 等人基于 Fukushima 的研究工作使用 BP 算法设计并训练了 CNN, Lerner-5 是经典的 CNN 结构，后续有许多工作基于此进行改进，它在一些模式识别领域中取得了良好的分类效果

CNN 的基本结构由输入层、卷积层 (Convolutional Layer)、池化层 (Pooling Layer, 也称为取样层)、全连接层及输出层构成。卷积层和池化层一般会取若干个，采用卷积层和池化层交替设置，即一个卷积层连接一个池化层，池化层后再连接一个卷积层，依此类推。由于卷积层中输出特征面的每个神经元与其输入进行局部连接，并通过对应的连接权值与局部输入进行加权求和再加上偏置值，得到该神经元输入值，该过程等同于卷积过程，CNN 也由此而得名。

## 2 CNN 的一些改进算法

### 2.1 网中网结构

CNN 中的卷积滤波器是一种广义线性模型 (Generalized linear model), GLM 的抽象水平比较低，但通过抽象却可以得到对同一概念的不同变体保持不变的特征。Lin 等人提出了一种网中网 (Net in net) 模型，该模型使用微型神经网络 (Miniature neural network) 代替传统 CNN 的卷积过程，同时还采用全局平均池化来替换传统 CNN<sup>[11]</sup>的全连接层，它可以增强神经网络的表示能力。微神经网络主要是采用 MPL 模型用 MLP 来取代原来的通过在输入中滑动微型神经网络得到卷积层的特征面。与卷积的权值共享类似，对同一个特征面的所有局部感受野也共享，即对于同一个特征面相同。文献之所以选择 MLP，考虑到 MLP 采用 BP 算法进行训练，能与 CNN 结构融合，同时 MLP 也是一种深度模型，具有特征重用的思想<sup>[12]</sup>。

$$\delta t^j(t) = a_0 + a_1(t - t_{oc}) + a_2(t - t_{oc})^2 \quad (1)$$

MLP 卷积层能够处理更复杂的非线性问题，提取更加抽象的特征。在传统的结构中全连接层的参数过多，易于过拟合，因此它严重依赖于正则化技术。模型采用全局平均池化代替原来的全连接层，使模型的参数大大减少。它通过全局平均池化方法对最后一个 MLP 卷积层的每个特征面求取均值，再将这些数值连接成量，最后输到 Soft 分类层中。全局平均池化可看成是一个结构性的正<sup>[13]</sup>则

化算子，它可以增强特征面与类别的一致性。在全局平均池化层中没有需要优化的参数，因此能够避免过拟合。

$$\delta\rho = \frac{2.4224}{0.026 + \sin E} \exp(-0.13345H) \quad (2)$$

此外，全局平均池化层对空间信息进行求和，因此对输入的空间变化具有更强的鲁棒性。LIN 等人将该算法应用于数据集中，验证了该算法的有效性。XU 等人结合 NIN<sup>[14]</sup>模型提出了 ML-DNN 模型，使用与文献相同的数据库，将其与稀疏<sup>[15]</sup>疏编码等方法比较，表明了该模型的优越性

## 2.2 反卷积

由 Zeliu 等人提出的反卷积网络 (Deconvolution network) 模型与 CNN 的思想类似，只是在运算上有所不同。CNN 是一种自底而上的方法，其输入信号经过多层的卷积、非线性变换和下采样处理。而反卷积网络中的每层信息是自顶而下的，它对由已学习的滤波器组<sup>[16]</sup>与特征面进行卷积后得到的特征求和就能重构输入信号。随后，采用反卷积网络可视化中各网络层学习得到的特征，从而有利于分析并改进它的网络结构。反卷积网络也可看成是一个卷积模型，它同样需要进行卷积和池化过程，不同之处在于与 CNN<sup>[17]</sup>是一个逆过程。模型中的每一个卷积层都加上一个反卷积层。在卷积、非线性函数、最大池化后<sup>[18]</sup>，不仅将输出的特征作为下一层的输入，也将它送给对应的反卷积层。反卷积层需要依次进行一种近似的方法求最大池化的逆过程、矫正（使用非线性函数来保证所有输出均为非负数）及反卷积操作<sup>[19]</sup>（利用卷积过程中卷积核的转置作为核，与矫正后的特征作卷积运算），然后形成重构特征，模型由卷积结构<sup>[20]</sup>及反卷积结构组成，采用卷积结构对输入进行编码，而反卷积结构用来进行重构。SWWAE 的每一个阶段是一个“内容一位置”自动编码器<sup>[21]</sup>编码器由一个卷积层及紧随其后的一个最大池化层组成，通过最大池化层产生两个变量集：最大池化的输出记为 What 变量，它作为下一层的输入；将最大池化的位置信息记为变量、变量要横向传递到反卷积结构中。SWWAE 的损失包含 3 个部分<sup>[22]</sup>（判别损失、重构损失及中间重构损失）在各种半监督和监督任务中取得了很高的准确率<sup>[23]</sup>，它特别适用于具有大量无标注类别而有标注类别相对少的数据集的情况

$$\Phi_{k/k-1} = \exp(F(t_k)\Delta t) \approx I + F(t_k) \quad (3)$$

$$Q_k = \Phi_k G(t_k) Q_k G(t_k)^T \Phi_k^T \Delta t$$

## 3 无监督学习的应用

机器学习算法<sup>[24]</sup>可以分为有监督学习和无监督学习两种，有监督学习是指训练的样本带有标签，而无监督学习在训练过程中样本没有标签。在现实世界中，大部分样本是不带标签的，所以无监督学

习要比监督学习应用更广泛。常用的无监督学习算法主要有主成分分析方法等，等距映射方法、局部线性嵌入方法、拉普拉斯特征映射<sup>[25]</sup>方法、黑塞局部线性嵌入方法<sup>[26]</sup>和局部切空间排列方法等。深度学习是由多层神经网络组成，需要一层一层地抽取主要特征，忽略次要细节，所以深度学习中采用的无监督学习<sup>[27]</sup>方法需要满足三个条件：

- ① 可以从多维空间中抽取主要特征映射至低维空间；
- ② 具有递归性<sup>[28]</sup>；
- ③ 算法不能太过复杂，否则深层架构的计算量太大。

从原理上来说 PCA 等数据降维算法同样适用于深度学习，但是这些数据降维方法复杂度较高，并且其算法的目标太明确<sup>[29]</sup>，使得抽象后的低维数据中没有次要信息，而这些次要信息可能在更高层看来是区分数据的主要因素。所以现在深度学习中采用的无监督学习方法通常采用较为简单的算法和直观的评价标准。目前深度学习中的无监督学习主要分为两类，一类是确定型的自编码方法及其改进算法，其目标主要是能够从抽象后的数据中尽量无损地恢复原有数据<sup>[30]</sup>，一类是概率型的受限波尔兹曼机及其改进算法，其目标主要是使受限波尔兹曼机达到稳定状态时原数据出现的概率最大。

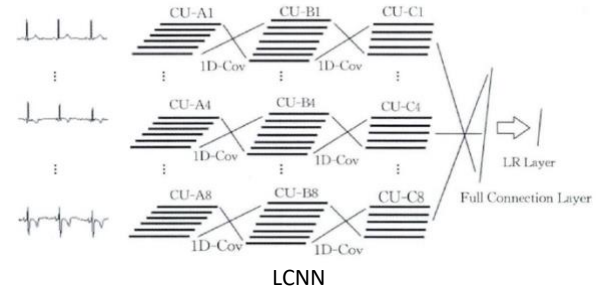


Fig.2 LCNN model

## 4 结论

近年来，CNN 的局部连接、权值共享、池化操作及多层结构等优良特性使其受到了许多研究者的关注。CNN 通过权值共享减少了需要训练的权个数、降低了网络的计算复杂度，同时通过池化操作使得网络对输入的局部变换具有一定的不变性如平移不变性、缩放不变性等，提升了网络的泛化能力。CNN 将原始数据直接输入到网络中，然后隐性地从训练数据中进行网络学习，避免了手工提取特征、从而导致误差累积的缺点，其整个分类过程是自动的。虽然 CNN 所具有的这些特点使其已被广泛应用于各种领域中，但其优势并不意味着可以解决或 CNN 仍有许多工作需要进一步去做。

## 参考文献

- [1] R. Mu and X. Zeng, "A review of deep learning research," *KSII Trans. Internet Inf. Syst.*, vol. 13, no. 4, pp. 1738–1764, 2019, doi: 10.3837/tiis.2019.04.001.
- [2] A. Mignan and M. Broccardo, "One neuron versus deep learning in aftershock prediction," *Nature*, vol. 574, no. 7776, pp. E1–E3, 2019, doi: 10.1038/s41586-019-1582-8. **B类**
- [3] "田大纲. 对偶性和线性规划问题的神经网络解法. 自动化学报, 2003, 29(2): 219-226." **B类**
- [4] J. Frazer *et al.*, "Disease variant prediction with deep generative models of evolutionary data.," *Nature*, vol. 599, no. December 2020, 2021, doi: 10.1038/s41586-021-04043-8. **A类**
- [5] J. Zhang *et al.*, "Fast and accurate automated recognition of the dominant cells from fecal images based on Faster R-CNN," *Sci. Rep.*, vol. 11, no. 1, pp. 1–8, 2021, doi: 10.1038/s41598-021-89863-4. **A类**
- [6] Y. E. Wu *et al.*, "Neural control of affiliative touch in prosocial interaction," *Nature*, vol. 599, no. 7884, pp. 262–267, 2021, doi: 10.1038/s41586-021-03962-w.
- [7] W. Shi *et al.*, "Research on Multi-aircraft Cooperative Air Combat Method Based on Deep Reinforcement Learning," *Zidonghua Xuebao/Acta Autom. Sin.*, vol. 47, no. 7, pp. 1610–1623, 2021, doi: 10.16383/j.aas.c201059. **B类**
- [8] Z.-H. Zhou, "Can Deep Learning Only Be Neural Networks?," pp. 6–6, 2020, doi: 10.1145/3336191.3372190. **A类**
- [9] D. Xiaoxiao and Z. Dongdai, "Study on the Influence of Learning engagement on Deep Learning in E-Learning Environment," *PervasiveHealth Pervasive Comput. Technol. Healthc.*, pp. 63–67, 2020, doi: 10.1145/3395245.3396428. **B类**
- [10] "ac ac 10 ac 10," vol. 359, no. October, p. 1992, 1992.
- [11] W. Lei, Q. Zhiwen, and Q. Ruonan, "Passenger capacity prediction based on genetic neural network," *Proc. - 2009 Int. Symp. Inf. Eng. Electron. Commer. IEEC 2009*, pp. 696–700, 2009, doi: 10.1109/IEEC.2009.152. **A类**
- [12] L. Shi, B. Li, C. Kim, P. Kellnhofer, and W. Matusik, "Towards real-time photorealistic 3D holography with deep neural networks," *Nature*, vol. 591, no. 7849, pp. 234–239, 2021, doi: 10.1038/s41586-020-03152-0.
- [13] H. A. Elmarakeby *et al.*, "Biologically informed deep neural network for prostate cancer discovery," *Nature*, vol. 598, no. 7880, pp. 348–352, 2021, doi: 10.1038/s41586-021-03922-4. **A类**
- [14] L. Shi, B. Li, C. Kim, P. Kellnhofer, and W. Matusik, "Author Correction: Towards real-time photorealistic 3D holography with deep neural networks (Nature, (2021), 591, 7849, (234-239)," **A类**  
10.1038/s41586-020-03152-0)," *Nature*, vol. 593, no. 7858, p. E13, 2021, doi: 10.1038/s41586-021-03476-5.
- [15] F. Yang-he and D. Silver, "Deep Multi-Agent Reinforcement Learning: A Survey," 2020.
- [16] 陈伟宏, 安吉尧, 李仁发, and 李万里, "深度学习认知计算综述," 自动化学报, vol. 43, no. 11, pp. 1886–1897, 2017. **B类**
- [17] 宋轶民, 马文贵, 张策. 基于神经网络的弹性连杆机构振动主动控制. 自动化学报, 2000, 26(5): 660-665. **B类**
- [18] A. R. Durmaz *et al.*, "A deep learning approach for complex microstructure inference," *Nat. Commun.*, vol. 12, no. 1, pp. 1–15, 2021, doi: 10.1038/s41467-021-26565-5. **A类**
- [19] N. Eter, "Deep learning in ophthalmology," *Ophthalmologie*, vol. 115, no. 9, pp. 712–713, 2018, doi: 10.1007/s00347-018-0713-1. **A类**
- [20] X. Yang, Y. Ye, X. Li, R. Y. K. Lau, X. AZhang, and X. Huang, "Hyperspectral image classification with deep learning models," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 9, pp. 5408–5423, 2018, doi: 10.1109/TGRS.2018.2815613. **A类**
- [21] W. J. Si and D. S. Wang, "Adaptive neural control for uncertain high-order stochastic nonlinear systems with input and output constraints," *Kongzhi Lilun Yu Yingyong/Control Theory Appl.*, vol. 36, no. 8, pp. 1250–1258, 2019, doi: 10.7641/CTA.2018.80427. **A类**
- [22] F. Stelzer, A. Röhm, R. Vicente, I. Fischer, and S. AYanchuk, "Deep neural networks using a single neuron: folded-in-time architecture using feedback-modulated delay loops," *Nat. Commun.*, vol. 12, no. 1, pp. 1–10, 2021, doi: 10.1038/s41467-021-25427-4. **A类**
- [23] Q. Wang and L. Zhang, "Inverse design of glass structure with deep graph neural networks," *Nat. Commun.*, vol. 12, no. 1, pp. 1–11, 2021, doi: 10.1038/s41467-021-25490-x.
- [24] T. Maekawa *et al.*, "Cross-species behavior analysis with attention-based domain-adversarial deep neural networks," *Nat. Commun.*, vol. 12, no. 1, pp. 1–11, 2021, doi: 10.1038/s41467-021-25636-x.
- [25] E. Tseng *et al.*, "imaging," no. 2021, pp. 1–7, doi: 10.1038/s41467-021-26443-0.
- [26] 段艳杰, 吕宜生, 张杰, 赵学亮, 王飞跃. 深度学习在控制领域的研究现状与展望. 自动化学报, 2016, 42(5): 643-654. doi: 10.16383/j.aas.2016.c160019 **B类**
- [27] 李翔龙, 殷国富, 罗红波. 进化神经网络在机床工具损耗预测中的应用. 自动化学报, 2004, 30(1): 114-119. **B类**
- [28] D. B. Zhang, Y. N. Wang, and H. X. Huang, "Fuzzy rough model based rough neural network modeling," *Zidonghua Xuebao/Acta Autom. Sin.*, vol. 34, no. 8, pp. 1016–1023, 2008, doi: 10.3724/SP.J.1004.2008.01016.
- [29] 杨若黎, 吴沧浦. 一种新的非线性规划神经网络模型. 自动化学报, 1996, 22(3): 293-300. **B类**
- [30] P. Javorka, A. Alam, M. Marso, and M. Heuken, "S t d s t s t," vol. 49, no. 8, pp. 1496–1498, 2002.

(引用、参考文献均采用 Mendeley 编辑)