

7大类卷积神经网络(CNN)创新综述!



尤而小屋 2023-01-18 00:00 发表于湖北

收录于合集
#神经网络 10 #卷积神经网络 4 #CNN 2 #深度学习 42

编者荐语

本综述将最近的 CNN 架构创新分为七个不同的类别，分别基于空间利用、深度、多路径、宽度、特征图利用、通道提升和注意力。

深度卷积神经网络 (CNN) 是一种特殊类型的神经网络，在各种竞赛基准上表现出了当前最优结果。深度 CNN 架构在挑战性基准任务比赛中实现的高性能表明，创新的架构理念以及参数优化可以提高 CNN 在各种视觉相关任务上的性能。



引言



通过 1989 年 LeCun 处理网格状拓扑数据 (图像和时间系列数据) 的研究，CNN 首次受到关注。CNN 被视为理解图像内容的最好技术之一，并且在图像识别、分割、检测和检索相关任务上表现出了当前最佳性能。

CNN 的成功引起了学界外的注意。在产业界，如谷歌、微软、AT&T、NEC 和 Facebook 这样的公司都设立了研究团队来探索 CNN 的新架构。目前，图像处理竞赛中的大多数领跑者都会采用基于深度 CNN 的模型。

自 2012 年以来，关于 CNN 架构的不同创新被提出来。这些创新可分为参数优化、正则化、结构重组等。但是据观察，CNN 网络的性能提升应主要归功于处理单元的重构和新模块的设计。自 AlexNet 在 ImageNet 数据集上展现出了非凡的性能后，基于 CNN 的应用变得越来越普及。



类似地，Zeiler 和 Fergus 介绍了特征分层可视化的概念，这改变了用深度架构 (如 VGG) 在简单的低空间分辨率中提取特征的趋势。如今，大多数新架构都是基于 VGG 引入的简单原则和同质化拓扑构建的。

另一方面，谷歌团队引入了一个非常著名的关于拆分、转换和合并的概念，称为 Inception 模块。初始块第一次使用了层内分支的概念，允许在不同空间尺度上提取特征。2015 年，为了训练深度 CNN，Resnet 引入的残差连接概念变得很有名，并且，后来的大多数网络像 Inception-ResNet, WideResNet, ResNext 等都在使用它。



与此类似，一些像 WideResnet, Pyramidal Nets、Xception 这样的架构都引入了多层转换的概念，通过额外的基数和增加的宽度来实现。因此，研究的重点从参数优化和连接再次调整，转向了网络架构设计 (层结构)。这引发了许多像通道提升、空间和通道利用、基于注意力的信息处理等新的架构概念。



本文结构如下：

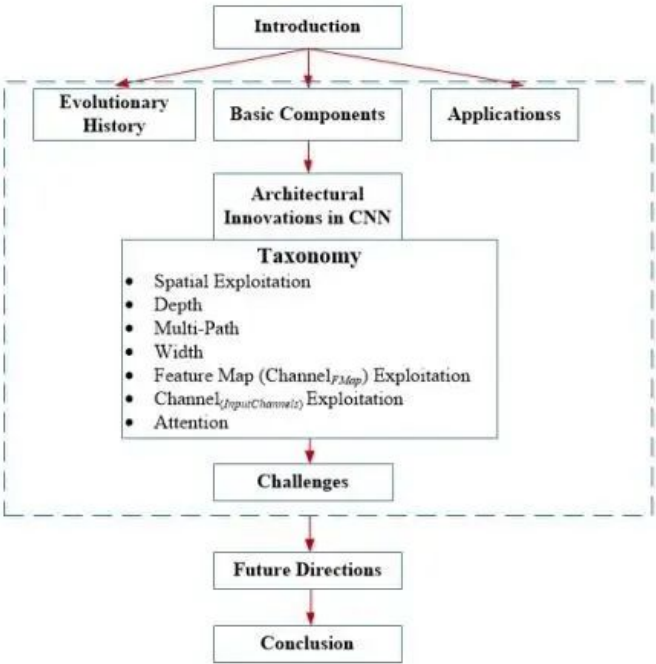


图 1: 文章结构

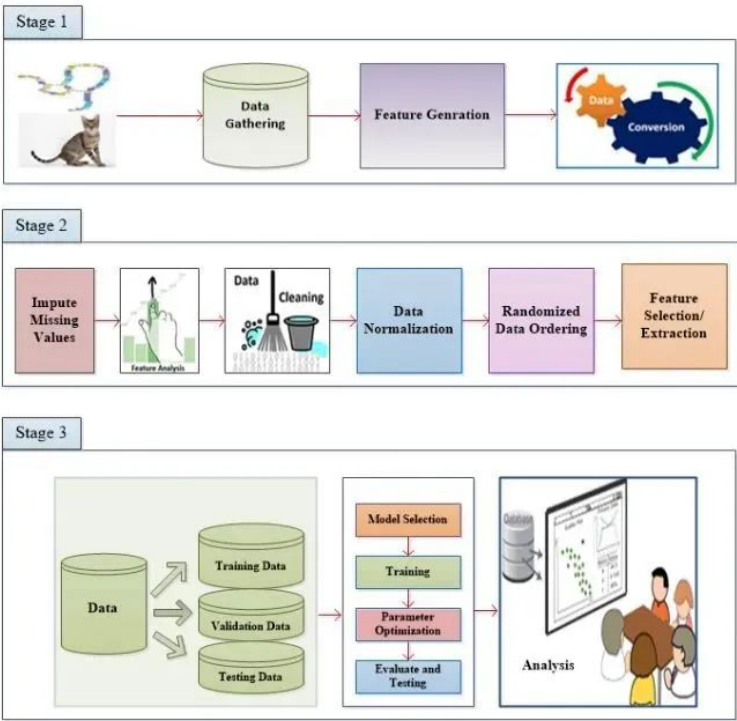


图 2: 典型模式识别 (OR) 系统的基本布局。PR 系统分为三个阶段: 阶段 1 和数据挖掘相关, 阶段 2 执行预处理和特征选择, 而阶段 3 基于模型选择、调参和分析。CNN 有良好的特征提取能力和强大的鉴别能力, 因此在一个 PR 系统中, 它可以用于特征提取/生成和模型选择阶段。

CNN 中的架构创新

自 1989 年至今, CNN 架构已经有了很多不同的改进。CNN 中的所有创新都是通过深度和空间相结合实现的。根据架构修改的类型, CNN 可以大致分为 7 类: 基于空间利用、深度、多路径、宽度、通道提升、特征图利用和注意力的 CNN。深度 CNN 架构的分类如图 3 所示。

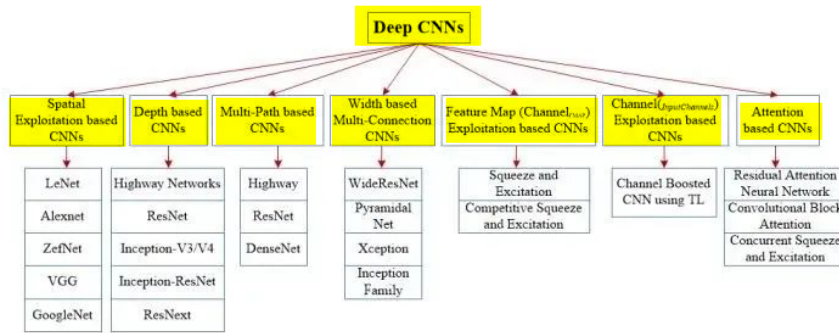


图 3: 深度 CNN 架构分类

基于空间利用的 CNN



CNN 有大量参数, 如处理单元数量 (神经元)、层数、滤波器大小、步幅、学习率和激活函数等。由于 CNN 考虑输入像素的邻域 (局部性), 可以使用不同大小的滤波器来探索不同级别的相关性。



因此, 在 2000 年初, 研究人员利用空间变换来提升性能, 此外, 还评估了不同大小的滤波器对网络学习率的影响。不同大小的滤波器封装不同级别的粒度; 通常, 较小的滤波器提取细粒度信息, 而较大的滤波器提取粗粒度信息。这样, 通过调整滤波器大小, CNN 可以在粗粒度和细粒度的细节上都表现很好。

基于深度的 CNN



深度 CNN 架构基于这样一种假设: 随着深度的增加, 网络可以通过大量非线性映射和改进的特征表示更好地逼近目标函数。网络深度在监督学习的成功中起了重要作用。理论研究已表明, 深度网络能够以指数方式比浅层网络更有效地表示特定的 20 个函数类型。

2001 年, Csáji 表示了通用近似定理, 指出单个隐藏层足够逼近任何函数, 但这需要指数级的神经元, 因而通常导致计算上行不通。在这方面, Bengio 和 elalleau 认为更深的网络有潜力在更少的成本下保持网络的表现能力。

2013 年, Bengio 等人通过实证表明, 对于复杂的任务, 深度网络在计算和统计上都更有效。在 2014-ILSVR 竞赛中表现最佳的 Inception 和 VGG 则进一步说明, 深度是调节网络学习能力的重要维度。

一旦特征被提取, 只要其相对于其他位置的近似位置被保留, 其提取位置就变得没那么重要了。池化或下采样 (如卷积) 是一种有趣的局部操作。它总结了感受野附近的类似信息, 并输出了该局部区域内的主要反应。作为卷积运算的输出结果, 特征图案可能会出现在图像中的不同位置。

基于多路径的 CNN



深度网络的训练颇具挑战性, 这也是近来很多深度网络研究的主题。深度 CNN 为复杂任务提供了高效的计算和统计。但是, 更深的网络可能会遭遇性能下降或梯度消失/爆炸的问题, 而这通常是由增加深度而非过拟合造成的。

梯度消失问题不仅会导致更高的测试误差, 还会导致更高的训练误差。为了训练更深的网络, 多路径或跨层连接的概念被提出。多路径或捷径连接可以通过跳过一些中间层, 系统地将一层连接到另一层, 以使特定的信息流跨过层。跨层连接将网络划分为几块。

这些路径也尝试通过使较**低层访问梯度**来解决**梯度消失问题**。为此，使用了不同类型的捷径连接，如零填充、基于投影、**dropout** 和 **1x1** 连接等。

激活函数是一种决策函数，有助于学习复杂的模式。选择适当的激活函数可以加速学习过程。卷积特征图的激活函数定义为等式 (3)。

$$T_l^k = f_A(F_l^k) \quad (3)$$

4

基于宽度的多连接 CNN



2012 至 2015 年，**网络架构的重点是深度的力量**，以及**多通道监督连接**在**网络正则化**中的重要性。然而，**网络的宽度和深度一样重要**。通过在一层之内并行使用多处理单元，多层感知机获得了在感知机上映射复杂函数的优势。

这表明**宽度和深度一样**是定义学习原则的一个重要参数。Lu 等人和 Hanin & Sellke 最近表明，带有**线性整流激活函数的神经网络要足够宽才能随着深度增加保持通用的近似特性**。并且，如果网络的最大宽度不大于输入维度，紧致集上的连续函数类无法被任意深度的网络很好地近似。

因此，**多层堆叠（增加层）可能不会增加神经网络的表征能力**。与深度架构相关的一个重要问题是，有些**层或处理单元可能无法学习有用的特征**。为了解决这一问题，研究的重点从深度和较窄的架构转移到了**较浅和较宽的架构上**。

5

基于特征图（通道特征图）开发的 CNN



CNN 因其分层学习和自动特征提取能力而闻名于 MV 任务中。**特征选择在决定分类、分割和检测模块的性能上起着重要作用**。**传统特征提取技术**中分类模块的**性能要受限于特征的单一性**。相较于传统技术，CNN 使用**多阶段特征提取**，根据**分配的输入来提取不同类型的特征**（CNN 中称之为**特征图**）。

但是，**一些特征图有很少或者几乎没有目标鉴别作用**。巨大的特征集有噪声效应，会导致**网络过拟合**。这表明，除了网络工程外，**特定类别特征图的选取对改进网络的泛化性能至关重要**。在这一部分，**特征图和通道会交替使用**，因为很多研究者已经用**通道**这个词代替了**特征图**。

6

基于通道（输入通道）利用的 CNN



图像表征在决定图像处理算法的性能方面起着重要作用。**图像的良好表征**可以定义来自紧凑代码的图像的**突出特征**。在不同的研究中，**不同类型的传统滤波器**被用来提取单一类型图像的不同级别信息。

这些不同的表征被用作模型的输入，以提高性能。CNN 是一个很好的特征学习器，它能根据问题**自动提取鉴别特征**。但是，CNN 的学习依赖于输入表征。

如果输入中缺乏多样性和类别定义信息，CNN 作为鉴别器的性能就会受到影响。为此，**辅助学习器的概念**被引入到 CNN 中来提升网络的输入表征。

7

基于注意力的 CNN



不同的抽象级别在定义神经网络的鉴别能力方面有着重要的作用。除此之外，选择与上下文相关的特征对于图像定位和识别也很重要。在人类的视觉系统中，这种现象叫做注意力。人类在一次又一次的匆匆一瞥中观察场景并注意与上下文相关的部分。在这个过程中，人类不仅注意选择的区域，而且推理出关于那个位置的物体的不同解释。

因此，它有助于人类以更好的方式来抓取视觉结构。类似的解释能力被添加到像 RNN 和 LSTM 这样的神经网络中。上述网络利用注意力模块来生成序列数据，并且根据新样本在先前迭代中的出现来对其加权。

不同的研究者把注意力概念加入到 CNN 中来改进表征和克服数据的计算限制问题。注意力概念有助于让 CNN 变得更加智能，使其在杂乱的背景和复杂的场景中也能识别物体。

论文: A Survey of the Recent Architectures of Deep Convolutional Neural Networks

A Survey of the Recent Architectures of Deep Convolutional Neural Networks

Asifullah Khan^{1,2*}, Anabia Sohail¹, Umme Zahoor¹ and Aqsa Saeed Qureshi¹

¹ Pattern Recognition Lab, DCIS, PIEAS, Nilore, Islamabad 45650, Pakistan

² Deep Learning Lab, Center for Mathematical Sciences, PIEAS, Nilore, Islamabad 45650, Pakistan

论文地址: <https://arxiv.org/abs/1901.06032>

摘要: 深度卷积神经网络 (CNN) 是一种特殊类型的神经网络，在各种竞赛基准上表现出了当前最优结果。深度 CNN 的超强学习能力主要是通过使用多个非线性特征提取阶段实现的，这些阶段能够从数据中自动学习分层表征。大量数据的可用性和硬件处理单元的改进加速了 CNN 的研究，最近也报道了非常有趣的深度 CNN 架构。

近来，深度 CNN 架构在挑战性基准任务比赛中实现的高性能表明，创新的架构理念以及参数优化可以提高 CNN 在各种视觉相关任务上的性能。鉴于此，关于 CNN 设计不同想法被探索出来，如使用不同的激活函数和损失函数、参数优化、正则化以及处理单元的重构。

然而，在表征能力方面的主要改进是通过重构处理单元来实现的。尤其是，使用块而不是层来作为结构单元的想法获得了极大的赞赏。

本综述将最近的 CNN 架构创新分为七个不同的类别。这七个类别分别基于空间利用、深度、多路径、宽度、特征图利用、通道提升和注意力。此外，本文还涵盖了对 CNN 组成部分的基本理解，并揭示了 CNN 目前面临的挑战及其应用。

转载自 | 深度学习这件小事



kaggle比赛: 肾脏疾病预测!

周志华亲授机器学习视频课上线了

【Pandas技巧】groupby+agg/transform

Pandas实用技巧: SAC机制!

机器学习: 图解PCA

Kaggle可视化：黑色星期五画像分析

尤而小屋，一个温馨的小屋。小屋主人，一手代码谋求生存，一手掌勺享受生活，欢迎你的光临



尤而小屋
尤而小屋，一个温馨且有爱的小屋🏠 小屋主人，一手代码谋求生存，一手掌勺享受生...
303篇原创内容

公众号

收录于合集 #神经网络 10

上一篇 · python深度学习18-生成式深度学习之DeepDream

喜欢此内容的人还喜欢

计算机是如何理解图片的？——基于深度学习的计算机视觉与卷积神经网络
福建省人工智能协会



一文看尽深度学习中的各种数据增强
CVHub



分析GRE隧道配置keepalive后隧道不up
口味儿

