

CNN 概述

CNN (Convolutional Neural Network), 中文名称“卷积神经网络”, 是一种专门处理具有类似网格结构数据的前馈神经网络。主要用于计算机视觉方向, 对于图像处理有出色表现。本文主要从 CNN 解决什么问题, CNN 的发展, CNN 原理以及 github 上 CNN 项目四个方面来介绍 CNN。

一 CNN 解决的主要问题

CNN 目前主要应用于计算机视觉领域, 例如图片的分类识别等。对于人眼来说看到的是色彩斑斓的照片, 但是对于 CNN 来说一张图片就是 $Height * Width * RGB$ 的像素矩阵。矩阵中每个像素点的值都在 0-255 之间, 计算机将输入的多维数组通过概率的方式对图片进行分类, 这就是 CNN 要解决的问题。

二 CNN 的发展

1958 年, David Hunter Hubel 和 Torsten Wiesel 通过对小猫进行实验, 观察猫的脑内神经元是如何响应投影在猫前方屏幕上的图像。实验表明, 小猫脑内的每个单个神经元只对特定的图像特征有反应, 比如有些神经元会被图像中横向的边缘所激活, 另一些神经元则被纵向的或者对角方向的边缘所激活, 而还有些神经元对亮光带和暗光带的反应模式也不相同。这说明, 单个神经元的作用域并不是整个图像, 而只是图像的一个子集, 这个子集就叫做**感受野**

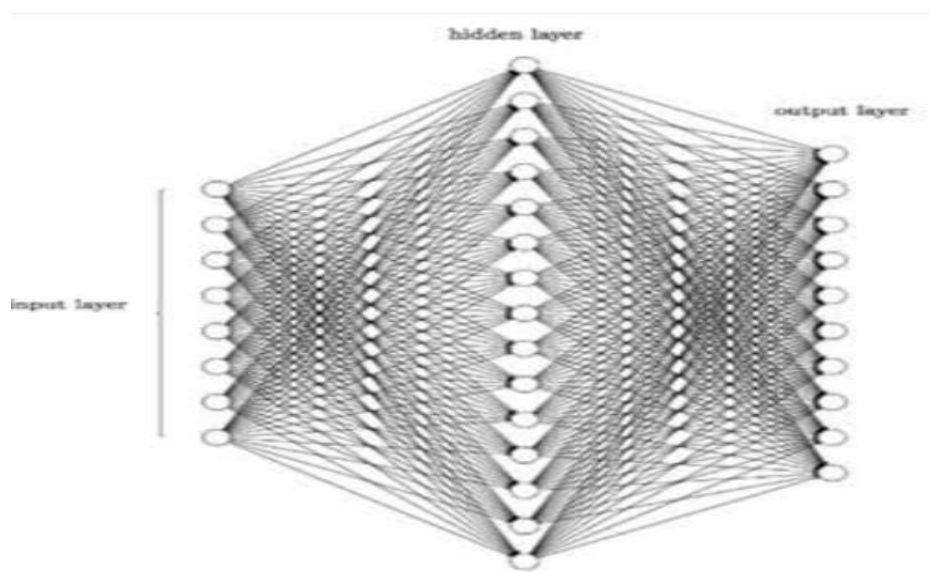
(Receptive Field)。相邻的神经元有着相似的, 相互重叠的感受野, 所有神经元有序的组织起来共同形成对图像的认知。整个实验

的过程及结论被记录在了他们 1959 年发表的论文中，这篇论文奠定了此后二十年间他们对大脑皮层研究取得优秀成果的基础。他们还在研究中发现，**视觉系统的信息处理是分层的。**

受感受野及视觉系统的分级信息处理的启发，第一个卷积神经网络——时延神经网络（Time Delay Neural Networks, TDNN）于 1989 年被 Hinton 等人提出，初衷是为了解决语音识别中传统方法隐马尔可夫模型（HiddenMarkov Model, HMM）无法适应语音信号中的动态时域变化的问题。该结构参数较少，进行语音识别不需要预先将音标与音频在时间线上进行对齐，实验证明 TDNN 相比 HMM 表现更好。随后，LeCun 于同年在识别手写邮政编码的系统中引入了卷积运算，并且通过反向传播的方法从图像中自动学习到了卷积核系数，使得整个训练过程完全自动化，实现了第一个真正意义上的卷积神经网络。真正使卷积神经网络大放异彩，在计算机视觉领域应用的如火如荼的事件，还要源于 2012 年举办的年度计算机视觉领域“奥林匹克”比赛——ImageNet 竞赛。Alex Krizhevsky 使用卷积神经网络将图片分类的错误率从 26%降低至了 15%，这绝对是一个惊人的进步。从那时起，众多的大佬级公司开始重视并应用卷积神经网络于他们的产品中。

三 CNN 原理

神经网络是一种运算模型，由大量的神经元相互连接而成，每个节点代表一种特定的输出函数，成为激励函数，每两个神经元之间的连接线表示该连接的权重。一个典型的全连接的神经网络模型如下：



卷积神经网络是一种深度神经网络结构，必然就有很多层。一般分为卷积层，池化层以及全连接层。卷积层负责特征提取，池化层负责将数据简化，全连接层负责决策判断。全连接层最后输出的一般是各个类别的概率。

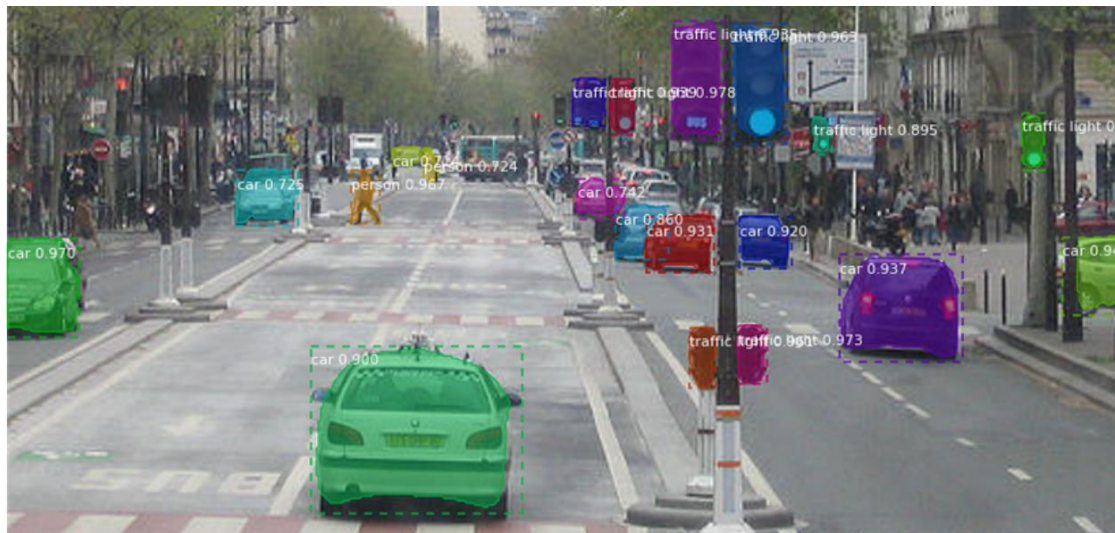
四 github 上的 CNN 项目

1. facebookresearch/Detectron

Detectron 是 Facebook AI 研究院（FAIR）于 2018 年初公开是我目标检测平台，包含了大量业内最具代表性的目标检测、图像分割、关键点检测算法。该框架主要基于 python 和 caffe2 实现，部分底层代码由 C++ 实现，另外部分算法如 group normal (GN) 则是基于 pytorch 框架。研究人员主要是将 Mask RCNN, RetinaNet, Faster RCNN, RPN 等算法集中到一个框架下，方便使用。截止 2019 年 12 月 13 日，本项目在 github 平台上 Watch 986 次，Star 22.5k 次，Fork 5k 次，共有 Issues 268 个，Pull request 22 个

2. matterport/Mask-RCNN

matterport/Mask-RCNN 是一个基于 python3, keras 和 tensorflow 的 mask rcnn 模型。这个模型对图像中的每一个目标产生候选框和分割掩膜。这个模型基于特征金字塔网络 (Feature Pyramid Network, FPN) 和一个 ResNet101 骨架。效果图如下:



截止 2019 年 12 月 13 日,本项目共 Watch 561 次,Star 15.1k 次, Fork 7k 次, 共有 Issues 1148 个, Pull request 73 个。

3. lengstorm/fast-style-transfer

这是基于 TensorFlow 的风格迁移项目,所谓风格迁移是指利用算法学习著名画作的风格,然后再把这种风格应用到另外一张图片上的技术。著名的图像处理应用 Prisma 是利用风格迁移技术,将普通用户的照片自动变换为具有艺术家的风格的图片。该项目可以快速地转换任何图像的风格。根据项目自己的测试,在 2015Titan X 显卡上将一幅 MIT Stata Center 的照片(1024*680)转换成弗朗西斯·毕卡比亚(Francis Picabia)的油画《Udine》的风格只要 100 毫秒。其速度之快甚至可以逐帧转换视屏并合成一个可以播放的视频。截止到 2019 年 12 月 13 日,本项目在

github 上用获得 Watch 328 次, Star 8.6 次, Fork 2.1k, 共有
Issues 60 个, Pull request 2 个。