

# 基于卷积神经网络的图像分类技术

童浩然 楚 军 沈静静

(安徽理工大学电气与信息工程学院, 安徽 淮南 232001)

**【摘要】**随着大量带标记的数据库的开源使用和带有高性能 GPU 的计算机的发展推广, 深度学习已然从理论走向实践, 开始广泛地活跃于图像分类的舞台之上, 其中表现最为突出的是卷积神经网络 (Convolutional neural networks, CNNs), 目前已经在大规模的识别与分类任务获得了瞩目的成果, 突破了传统分类方法的极限并且已经首次达到优于人眼识别的地步。本次实验的重点是将 CNN 应用在实际的图像分类操作中, 并对 CNN 进行优化与改进, 采用 Inception 架构高效并行处理数据, 并采取新型激活函数来提高识别率, 通过迁移学习在谷歌的深度学习框架 tensorflow 上进行试验, 结果表明, 通过对 CNN 架构的改进, 识别率得到了明显的提升, 证明改进的 CNN 具有相当的实用性。

**【关键词】**深度学习; 卷积神经网络; Inception 架构; 迁移学习

中图分类号: TP18

文献标识码: A

文章编号: 2095-2457(2017)32-0036-002

Image Classification Technology Based on Convolutional Neural Network

TONG Hao-ran CHU Jing SHEN Jing-jing

(School of Electrical and Information Engineering, Anhui University of Science and Technology, Huainan 232001, China)

**【Abstract】**With the open-source use of a large number of labeled databases and the development and promotion of computers with high-performance GPUs, deep learning has moved from theory to practice and has been widely active on the stage of image classification. The most prominent of these is Convolutional neural networks (CNNs) have achieved remarkable results in large-scale identification and classification tasks, breaking the limits of traditional classification methods and for the first time being better than human eye recognition. The focus of this experiment is to apply CNN in the actual image classification operation and to optimize and improve CNN. The Inception architecture is used to efficiently process data in parallel, and a new activation function is used to improve the recognition rate. Through the study of migration depth in Google The results show that the recognition rate has been significantly improved through the improvement of CNN architecture, which proves that the improved CNN is quite practical.

**【Key words】**Deep learning; Convolutional neural network; Inception architecture; Migration learning

## 0 引言

传统的图像分类技术最近几年来发展十分缓慢, 那些利用手动提取特征并进行分类的方法获得的关注越来越少。传统图像分类技术概括如下, 例如使用 SIFT<sup>[1]</sup>和 HOG<sup>[2]</sup>等方法进行特征提取后再利用无监督学习算法来生成特征描述符号, 然后利用监督学习算法对训练数据集中带标签的特征集进行学习并获得分类器, 从而实现对图像的分类。

随着大数据时代的到来和深度学习的火热兴起, 图像分类技术以日新月异的速度高速发展, 无数的成果接踵而来。尤其是 Alex<sup>[3]</sup>网络在著名的 ILSVRC<sup>[4]</sup>比赛上赢得冠军, 深度学习的实用性深得人心, 人们正式开始投身于深度学习网络在图像识别应用的汪洋大海中, 在物品定位, 姿态估计, 目标检测, 视频目标索引, 场景分类, 场景解析等应用中取得一个又一个瞩目的成就, 彻底取代了传统浅层网络和分类方法的地位。

在本文中, 将实验重心放在利用 Tensorflow 实际搭建 CNN 框架, 并对 Imagenet 数据库里面下载训练数据完成分类, 并主要进行两方面的创新, 首先是对 CNN 的 Alexnet 网络架构进行改进, 并通过迁移学习的方

法, 使之对数据库进行训练, 在图像分类的实验中取得了良好的效果。

## 1 深度网络

深度学习是传统神经网络的延伸和发展, 经过几十年的起起落落后, 已逐渐淡出研究界的视野, 但随着大数据时代的到来和硬件设备的提升再次活跃在机器学习的舞台之上, 其中在图像分类方面取得良好效果的当属 CNN。一个 CNN 主要有以下五种结构构成: 输入层, 卷积层 (Convolution), 池化层 (Pooling), 全连接层和 Softmax 层组成。输入层连接的是整个 CNN 的输入, 在图像识别的过程中, 表示的是一张图片的三维像素矩阵。CNN 从输入层开始, 将接收到的三维矩阵通过神经网络一层层传递下去直至完成分类。

### 1.1 卷积层

卷积层是 CNN 的核心部分和取得优于传统方法实验效果的关键, 主要是通过局部连接和权值共享来改善网络的架构, 从而实现减少参数, 降低计算的复杂度, 减轻实验过拟合的目的。卷积层的现源自偶然, 这里需要需要引入感受野的概念, 科学家通过对猫的视

觉皮层细胞研究发现，每一个视觉神经元都只会处理一小块视觉图像，根据这种一一对应的联系得到的启示，将过滤器引入神经网络中。过滤器是人工设定尺寸的三维矩阵，卷积层通过过滤器在节点间传送矩阵，将当前网络上的节点矩阵转换为下一层网络上的一个单位节点矩阵。

1.2 池化层

通常卷积层提取的特征结果往往拥有极大的数据量，为了解决计算量庞大的问题，池化层便应运而生。池化层往往出现在两个卷积层之间用来有效的缩小矩阵的尺寸来实现降采样，减少输入至最后全连接层的参数，从而达到减少计算量的目的。池化层的作用过程也是类似卷积层那样，通过移动过滤器来实现前向传播，不同点是不进行卷积计算加权和，只是单纯的采取最大值，这样的池化层称之为最大池化层。

2 改进

在 CNN 中卷积层的设计中，最重要的就是过滤器的选择，因为过滤器大小的选择直接影响特征提取的效果，选择较小的过滤器，优点是可以提取到更多更详尽的细节，然而如果选取的过滤器过小可能造成提取空白数据，导致特征大量丢失，从而影响分类效果，而且太小的过滤器选择会导致计算量增加，降低神经网络的效率。为了解决卷积神经网络的过滤器选择问题，本文采用 inception-v3 结构，和一般 CNN 中利用几个大小不同的过滤器串联起来的结构不同，它采用的是高效的并行处理架构，通过找出最优局部稀疏结构并用相似的用稠密组件代替。从而实现高效的降维，也可以在减少需要训练的参数的同时抑制过拟合，达到提高模型的推广能力的目的。

3 实验

3.1 ImageNet 数据集

ImageNet 是一个基于 WordNet 的大型图像数据库，在 ImageNet 中有大概 1500 万左右的不同类别图片，分别被关联到 WordNet 的两万个名词上，每个名词即为图像分类结果中的一个类别，ImageNet 中的图片都是从互联网上爬取的，并通过亚马逊的人工标注服务将其分类到 WordNet 上，ImageNet 里面的图像包括一个实体，或可能出现多个实体，本次试验主要研究单实体

图像的分类。

3.2 实验部分

为了验证改进的有效性，本次设计进行对比实验的方法，将实验分为三部分，第一部分是利用 CNN 基本的 LeNet-5 架构进行图像分类，第二部分是利用改进的 Inception-v3 架构进行分类，这部分激活函数选取 ReLU。第三部分是将 PReLU 函数加到 Inception 架构中，分别进行实验后记录结果，表 1 是对三种分类方法取得的识别率的比较。结果表明，改进的并行架构使得识别正确率得到了 5.9% 的提升的同时识别速率也得到了改善，同时激活函数的改变也使得在相同分类图像的识别效果得到了改善，改进的有效性得到了证明。

表 1

实验方法	识别率
LeNet-5	88.5%
ReLU-Inception	94.4%
PReLU-Inception	96.2%

4 结束语

和传统的分类方法相比，深度学习不仅在识别率方面具有明显的优越性，而且利用 CNN 的卷积层池化层等基本结构就可以提取图像特征并加以分类，这极大的简化了传统分类方法繁琐的步骤，是当今研究界最热门的课题之一，新的研究理论与成果与日俱增，并且已经初步投入工业使用中，基于深度学习的图像分类的产品也已推广使用。例如在谷歌，图像分类已经应用于无人驾驶汽车、YouTube、谷歌地图，图像搜索等。因此完全有理由相信深度学习必将取得更加瞩目的成果并且改变我们的生活。

【参考文献】

[1] J.: A large-scale hierarchical image database [C]. 2009.  
[2] A. classification with deep convolutional neural networks [J]. 2012.  
[3] Y.: Closing the gap to human-level performance in face verification [C]. 2012.  
[4] D. deep neural networks for image classification [J]. 2013.