

毕业设计（论文）



题 目 残差卷积神经网络在图像识别中的应用

学生姓名 赵昕

学 院 计算机科学与工程专 业 软件工程(嵌入式)

班 级 114053A 学 号 12014053007

指导教师 彭雅琴、钱敬冬

摘 要

图像识别，是指利用计算机对图像进行处理、分析和理解，以识别各种不同模式的目标和对像的技术。一般工业使用中，采用工业相机拍摄图片，然后再利用软件根据图片灰阶差做进一步识别处理。

课题的研究对象残差卷积神经网络（Residual Neural Network，ResNet）是2015年提出的深度卷积网络，一经出世，便在ImageNet中斩获图像分类、检测、定位三项的冠军。

本课题研究的内容是如何用残差卷积神经网络实现图像识别，并通过参数调整和优化提高算法的正确率。课题核心介绍了本系统的开发环境和残差卷积神经网络的算法结构，并通过对比实验证实了本系统的优越性。

关键词：图像识别；残差卷积神经网络；TensorFlow框架

**ABSTRACT**

Image recognition refers to the technique of using a computer to process, analyze, and understand images to identify various patterns of objects and objects. In general industrial use, an industrial camera is used to take pictures, and then software is used to further identify and process the image based on the difference in gray levels. The subject of the research project Residual Neural Network (ResNet) was a deep convolution network proposed in 2015. Once it was born, it captured the image classification, detection, and positioning champions in ImageNet.  
 The content of this study is how to use residual convolutional neural network to achieve image recognition, and improve the accuracy of the algorithm through parameter adjustment and optimization. The core of the project introduced the development environment of the system and the algorithm structure of the residual convolutional neural network, and confirmed the superiority of the system through comparative experiments.

**Keywords**: image recognition; residual convolutional neural network;

目录

1. 绪论
   1. 课题研究的背景和意义

近年来，随着计算机性能的提升、以及网络多媒体数据爆炸式的增长，海量多媒体数据的处理受到广泛关注。在计算机视觉领域，由于深度学习方法的发展，图像分类和物体检测的性能得到了巨大的提升。以卷积神经网络为代表的深度网络具有高度非线性的特点，它通过不断地叠加非线性函数来增强整个网络的表达能力，使之可以表达复杂的流形空间。

随着人们对于神经网络技术的不断研究和尝试，每年都会诞生很多新的网络结构或模型。这些模型大都有着经典神经网络的特点，但是又会有所变化。你说它们是杂交也好，是变种也罢，总之对于神经网络创新的各种办法那真叫大开脑洞。而这些变化通常影响的都是使得这些网络在某些分支领域或者场景下的表现更为出色（虽然我们期望网络的泛化性能够在所有的领域都有好的表现吧）。深度残差网络（Deep Residual Network）就是众多变种中的一个代表，而且在某些领域确实效果不错，

深度网络一个很重要的超参数就是深度，网络深度的提升往往伴随着网络性能的提升，一个比较好的例子就是VGG，该网络就是在AlexNex的基础上通过增加网络深度大幅度提高了网络性能。然而，随着网络深度的提升，训练深度网络也会变得越来越困难，极大的影响深度学习的效果。其中一种情况就是当网络层数达到一定的数目以后，网络的性能就会饱和，再增加网络的深度性能反而会开始退化，但是这种退化并不是由过拟合引起的，因为训练精度和测试精度都在下降，这说明当网络变得很深以后，深度网络就变得难以训练了。

ResNet的出现其实就是为了解决网络深度变深以后的性能退化问题。可以假设，当一个比较浅的网络达到了饱和的准确率时，如果后面再加上几个全等映射层，误差是不会增加的，即更深的网络不应该带来训练集上误差上升，ResNet正是应用了全等映射直接将前一层输出传到后面的思想。

ResNet shortcut没有权值，传递x后每个模块只学习残差F（x），且网络稳定易于学习，同时证明了随着网络深度的增加，性能将逐渐变好。可以推测，当网络层数够深时，优化Residual Function：F（x）=H（x）-x，易于优化一个复杂的非线性映射H（x）。

在ResNet推出之后，Google就借鉴了该算法思想，提出了Inception V4 和Inception-ResNet-V2，并通过融合这两个模型，在ILSVRC数据集上取得了惊人的3.08%的错误率。可见，ResNet及其思想对卷积神经网络的研究有着里程碑式的影响，具有很强的推广性，对其展开研究非常有意义。

* 1. 课题研究的现状
     1. 图像识别的研究现状

1.2.2残差卷积神经网络的研究现状

1.3内容安排

本课题研究的是残差卷积神经网络在图像识别中的应用，从算法原理到应用实例，从开发环境到运行结果，论文共分为四个部分，安排如下：

（1）第一章：该章节从课题的背景、意义和研究现状出发，对本课题研究的问题进行简单的介绍与阐述。

（2）第二章：该章节从开发系统环境、运行支持库和开发语言三个方面着手介绍本系统的开发环境。

（3）第三章：首先，介绍残差卷积神经网络的算法原理。其次，对代码分析所得的数据进行内容上的展示。接着，对系统的设计思路作出剖析并介绍功能。最后，通过对实验结果的归纳总结展示系统的优势。

（4）结束语：对本次课题研究中所采用的思路、方向和手段作总结说明，并陈述感想和收获。

1. 开发环境

2.1 Anaconda

Anaconda是一个强大的Python 库，它专门用于定义、优化、求值数学表达式，效率高，适用于多维数组。神经网络模型本身就是一群相互关联的数学表达式的集合，它的训练和测试的过程中就是数学表达式集的运算。Anaconda库中封装好的函数可以给模型的建立提供方便。

Anaconda的大部分代码都是使用Spyder编写的。Spyder是一种可编译为本地可执行代码的Python方言，与仅仅使用解释性Python语言相比，它能够提速代码运行过程。Anaconda同时支持GPU和CPU，在CPU上运行一个脚本，然后切换到GPU时代码本身并不需要做任何变化。Anaconda在处理具有大量数据的问题时，可以与C的速度媲美，通过利用GPU，甚至可以超过CPU上的C好几个数量级。不仅如此，Anaconda库还集成了很多优化程序，它能够减少计算量从而保证程序的运行时间尽可能最低。

Anaconda的应用范围十分广泛，包括基本的逻辑回归及其模型构建、使用LSTM对电影收视率进行分类等等。Anaconda可以仅仅使用Python语言来创建大部分先进的网络。

2.2 TensorFlow

TensorFlow的名字中已经说明了他最重要的两个概念——Tensor和Flow。Tensor就是张量。张量这个概念在数学或者物理学中可以有不同的解释，但在此论文中并不强调它本身的含义。在TensorFlow中，张量可以被简单地理解为多维数组。如果说TensorFlow的第一个词Tensor表明了他的数据结构，那么Flow则体现了它的计算模型。Flow翻译成中文就是“流”，它直观地表达了张量之间通过计算相互转化的过程。TensorFlow是一个通过计算图的形式来表述计算的编程系统。TensorFlow中的每一个计算都是计算图上的一个节点，而节点之间的边描述了计算之间的依赖关系。

2.3 Python

Python是一种流行的计算机程序设计语言，它诞生于1989年。

Python是完全面向对象的，对于Python而言，数字、函数等万事万物都是对象。和Java一样，Python也完全支持继承、重载、派生和多继承，复用性强是面向对象语言的一般优势。Python的面向对象甚至比Java和C#.net更彻底。

Python是解释型的，用Python语言写的程序可以跳过编译这一环节，直接从源代码运行程序。这减少了程序调试的时间，进而大大提高了开发的速度，也使得Python程序更加易于移植。

Python是一种非常简洁美观的高层语言，阅读一个良好的Python程序就感觉像是在读英语一样。Python的封装性好，调用方便，Python语法风格能够引导开发者专注于解决问题而不是去搞明白语言本身。Python的设计哲学旨在“用一种方法，最好是只有一种方法来做一件事”。在Python的设计之初，开发者在面临多种选择时，一般不会采用花俏的语法，而是选择明确的没有或者很少有歧义的语法，从而达到某个问题解决方案的统一化。语法的明确与统一使得Python极易上手。

Python是一门开源的语言，这使得它有很强的可塑性和传播性。使用者可以阅读Python的源代码，也可以直接对它做修改，甚至可以把它的一部分用于新的自由软件中。由于它的开源本质，Python已经被移植在Linux、Windows等诸多平台上。Python能够把其他语言制作的各种模板轻易的连接在一起，因此Python常被称为“胶水语言”。

Python拥有非常庞大的标准库。它功能齐全，可以帮助处理各种工作，例如正则表达式、文档生成、单元测试、线程、数据库、网页浏览器、XML、GUI（图形用户界面）、Tk和其他与系统有关的操作等。除了标准库以外，Python还拥有许多其他高质量的库，如Twisted和Python图像库等等。这使得Python应用前景非常广阔，其领域涉及了图形处理、数学处理、文本处理、数据库编程、网络编程和多媒体应用等等。

Python是一门简单而又强大的计算机语言，它拥有大量的数学库和图形库的支持，这给神经网络创造了良好的研究环境。以此可见，本课题以Python作为开发语言是明智的。

Python的创始人为Guido van Rossum。作为一种面向对象、解释性的高层语言和脚本语言。面向对象使得它支持多态、运算符重载、多重继承等高级概念，并使其成为C++和Java等高级系统语言的理想脚本工具。

Python是ABC语言的一种继承。ABC语言是Guido参加设计的一种教学语言，是专门为非程序员设计的，非常优美和强大，但因为其非开放性的限制而没有得到广泛的应用。Python弥补了ABC语言的缺陷，与其他语言如C、C++和Java结合的非常好，同时，也借鉴了Modula-3语言和C语言的习惯，不断地完善使得Python语言形成了自己的优势和特点。

Python是一种开源的、易学易用的脚本编程语言，它特别强调开发速度以及代码的清晰程度。能够适用于从简单到复杂的脚本任务，以及面向对象的独立的大型应用。对于语法规则上，Python有严格的要求，其中一条是：每航的首字符在这行的位置来决定某个模块的极限，也就是强制缩进，使得Python程序员们养成了良好的编程习惯，并且使得程序更加清晰和美观。

2.4 Spyder