基于Tensorflow的android端相册分类App设计与实现

摘要：图像的识别和分类在计算机视觉领域是一个久经不衰的研究课题。随着分类算法的日趋复杂，对硬件的要求也越来越高，随之而来的是硬件设备的体积越来越庞大。这就催生了一个需求，利用性能相对羸弱的移动端设备，凭借比较小的算力，达到和大型设备类似的图像分类效果。而2015年谷歌开源的Tensorflow深度框架，让各种深度学习网络运行在移动设备上得以实现。而对Tensorflow框架的具体实践也逐渐火热了起来。本文主要介绍，如何在android设备上运行Tensorflow框架，并基于框架，利用经典的用于图像处理的卷积神经网络inceptionV3进行Android端的相册图像识别。具体包括：

1. Tensorflow框架在移动端Jni接口的使用
2. InceptionV3模型的训练。
3. 训练数据的网络爬虫抓取。
4. 训练后的在移动端的移植。
5. Android端对相册和分类Jni库的调用。

本文的App是对 Tensorflow框架在移动端上的具体应用的一次实践，为成熟项目的开发提供了借鉴和经验。

近年来，基于深度学习的各类预训练模型，对许多传统领域，诸如计算机视觉，语音识别技术产生了极大的促进作用。其中首当其冲的便是计算机视觉。而计算机视觉领域，图像识别和分类作为基础技术得到了最广泛的关注，发展也日新月异。

从识别的对象上看：发展经历了从文字识别，数字图像识别，发展到如今的自然物体识别的过程。从发展的指标上看：识别的准确率不断上升，受限的条件逐渐被突破，图像的复杂度也不断加强。从识别的技术上看：传统的SVM已经非常成熟，而基于数据驱动的机器学习更是打开了人工智能的大门。

深度学习的训练方法，相对于传统的前向反馈网络，需要大量的训练层数以及迭代次数，对于算力的要求可以说是空前的。基于此，通常需要应用深度学习方法的模型会部署在大型服务器上，借助其强大的算力迅速得到结果。而作为大前端，包括web页和移动端，只扮演了一个数据传递和结果展示的角色，并没有真正参与到运算当中。但是，移动端对于实时性要求非常高，需要通过网络才能完成其功能的方式，让应用出现了限制和使用上的缺陷。所以，将深度学习框架移植到移动端，并且借助其比较小的算力，依然能够较快较好地完成相应运算的需求就应运而生了。

1、Tensorflow介绍

Tensorflow是谷歌大脑团队在第一代深度学习系统DistBelief的基础上，改进而来的通用计算框架。DistBelief本身在谷歌内部已经获得了巨大的成功，基于该框架的图像分类系统ImageNet，用Inception模型获得了ImageNet2014图像识别比赛的第一名。本文将使用Inception模型的改进版本InceptionV3，其介绍将在下一节中阐述。Tensorflow在此基础上，进一步提升了系统的通用性以及计算速度，他可以在更多的平台上更快更稳定地进行模型计算。

Tensorflow是基于数据流图（graph）的计算系统。数据流图是一个抽象概念。在节点（session）上进行数据运算，用入边和出边来表示数据的输入和输出。整个Tensorflow的计算过程，是数据从一段流到另一段的过程。

Tensorflow可以很方便地支持大部分神经网络，包括CNN，Rnn等。

2、InceptionV3模型简介

2012年，AlexNet作出历史性突破之后，主流的网络结构都是在让神经网络的深度更深，神经元数更多来让宽度更大，从而提升神经网络的能力。这样单纯地增大网络，带来了一系列的缺点。

1. 参数太多，在训练集中数据量不够多的情况下，容易过拟合。
2. 网络越大，计算复杂度越高，对算力要求过高，从而实际应用非常困难。
3. 网络深度过深，容易出现梯度弥散，难以优化模型。

而Inception结构的出现之后，为解决模型复杂度提供了一种思路。Inception模型的精髓在于“降维”，用多个低维度的卷积块，来替换高维度的卷积块。

大尺寸的滤波器的卷积（如5\*5，7\*7）引入的计算量非常大。一个5\*5卷积滤波器比一个3\*3卷积滤波器的计算量多25/9=2.78倍。大维度的卷积滤波器可以获得更多的信息，但可以通过多层的低纬度滤波器来代替这个大维度滤波器。5\*5的卷积，可以看作用一个3\*3的卷积和他的全连接层（相当于两个3\*3的卷积）来代替。这种方式可以理解成在卷积核大小上的因式分解。

Inception另一个主要的特性就是对卷积的堆叠。传统的卷积神经网络需

传统的卷积神经网络需要在1、3或者5这样的卷积边长当中选取一个合适的大小。而Inception给出的解决方案是，同时使用1、3、5边长的卷积滤波器，并把结果堆叠在一起，一方面增加了网络的宽度，另一方面增加了网络对尺度的适应性。

Inception在ImageNet挑战赛的成绩是，错误率低达5.1%。而本文当中将使用的便是InceptionV3模型，利用它来对照片的主体进行分类。

3、对已有模型的迁移训练

ImageNet图像分类的数据集有120万被标注的图片。如此巨大的数据大小，才能将神经网络训练到理想的准确率。而对于一般的研究人员来说，收集如此庞大的数据集并且标注其内容，其工作量非常庞大，几乎是不可能的。另外，用如此庞大的数据集重新训练一个模型会非常费时间，可能需要几天甚至几周。为了解决数据标注和训练时间过长的问题，迁移学习便应运而生。

迁移学习就是将已经在某个问题上训练好的模型，通过相对简单的调整，使其能够适应

4、Tensorflow的Android端移植

5、Android端图像预处理

6、Android对TensorFlow以及分类模型的调用