

一种移动式TensorFlow平台的卷积神经网络设计方法

李河伟

(北京物资学院, 北京 101149)

摘要: 近年来,随着卷积神经网络理论的快速发展和移动设备的普及,基于移动式TensorFlow平台构建定制化的卷积神经网络分类模型,成为深度学习爱好者的主要研究方向之一。将个人图片库作为训练集输入模型,通过调整阈值和权值,最终获得定制化的分类模型,满足个性化的分类需求。对移动式TensorFlow平台、卷积神经网络及其移植到移动平台上的相关步骤进行了研究。通过此开发流程,为进一步使用移动式TensorFlow平台解决现实图片分类问题提供了参考。

关键词: 移动式;TensorFlow;卷积神经网络;人工智能

中图分类号: TP18 **文献标识码:** A **文章编号:** 1009-3044(2017)22-0179-04

DOI: 10.14004/j.cnki.ckt.2017.2552

A Convolutional Neural Network Design Method for Mobile TensorFlow Platform

LI He-wei

(Beijing Wuzi University, Beijing 101149, China)

Abstract: In recent years, with the popularity of the rapid development of the theory of convolutional neural network and mobile devices, convolutional neural network classification model to build customized mobile platform Based on TensorFlow, has become one of the main research directions of deep learning enthusiasts. The personal image library is used as the input model of training set, and the customized classification model is finally obtained by adjusting the threshold and weight, which meets the personalized classification requirements. The related steps of mobile TensorFlow platform, convolutional neural network and its porting to mobile platform are studied. Through the development process, it provides a reference for the further use of mobile TensorFlow platform to solve the problem of realistic picture classification.

Key words: Mobile; TensorFlow; convolutional neural network; artificial intelligence

1 概述

即便是普通人也可以在有摄像头的手机上运行TensorFlow的Inception分类器,甚至是自定义的分类器。然后只要把摄像头对准希望做分类的东西,TensorFlow就会给出结果——它认为这是什么东西。2017年2月16日Pete Warden在TensorFlow Dev Summit 2017上的一篇以Mobile and Embedded TensorFlow为题的报告,引发了深度学习爱好者对于移动式TensorFlow平台的热情。

一个皮肤癌图像分类的应用案例:一位小哥拿到了一批皮肤癌的数据,使用一个pretrained的inception-v3这些数据做了一个inference,并把它搭建在手机上,完成一个app用来做皮肤癌的早期检测,并实现了很高的准确率。这是一个利用廉价的移动设备,结合计算机视觉和深度学习技术,有效检测是否有皮肤癌,大大节省医疗检测成本的成功案例。

TensorFlow作为一个开源的深度学习平台,实现了对卷积神经网络的良好支持。卷积神经网络提供了一种端到端的学习模型,模型中的参数可以通过传统的梯度下降方法进行训练,经过训练的卷积神经网络能够学习到图像中的特征,并且

完成对图像特征的提取和分类^[1]。它在图像分类、语音识别、目标检测等领域取得了一系列重大的研究成果。而移动式的TensorFlow平台更是为卷积神经网络的应用插上了飞翔的翅膀。下面,分别介绍移动式TensorFlow平台和卷积神经网络。

2 移动式TensorFlow平台

2.1 TensorFlow简介

TensorFlow是谷歌基于DistBelief进行研发的第二代人工智能开源学习系统^[2]。如今Google对待TensorFlow系统,有点类似于该公司对待旗下移动操作系统Android。它是一个采用数据流图(data flow graphs),用于数值计算的开源软件库。其命名来源于本身的原理,Tensor(张量)意味着N维数组,Flow(流)意味着基于数据流图的计算。Tensorflow运行过程就是张量从图的一端流动到另一端的计算过程。张量从图中流过的直观图像是这个工具取名为“TensorFlow”的原因。

Tensorflow的特性:

1)高度的灵活性:TensorFlow不是一个严格的“神经网络”库。只要你可以将你的计算表示为一个数据流图,你就可以使

收稿日期:2017-07-05

作者简介:李河伟(1988—),男,河南长葛人,硕士,主要研究方向为人工智能方法与应用。

本栏目责任编辑:唐一东

人工智能及识别技术 179

用TensorFlow。

2)可移植性(Portability):Tensorflow 可以运行在台式机、服务器、手机移动等等设备上。而且它可以充分使用计算资源,在多CPU和多GPU上运行。

3)多语言支持:Tensorflow 提供了一套易用的Python使用接口来构建和执行graphs,也同样提供了一套易于C++使用的接口(目前训练神经网络只支持python,C++接口只能使用已经训练好的模型)。未来还会支持Go、Java、Lua、JavaScript、R等等。

4)性能最优化:TensorFlow 给予了线程、队列、异步操作等最佳的支持,TensorFlow 可以把你手边硬件的计算潜能全部发挥出来,它可以充分利用多CPU和多GPU。

本文主要借助此平台搭建个人专属的自学习图片分类器模型,并实现在移动设备上实时地进行图片的识别分类工作。

2.2 移动式TensorFlow平台的搭建

1)TensorFlow 的安装可参考官方提供的教程。^[3]

下面仅演示在Ubuntu 16.4 64位python2.7环境下,通过pip安装TensorFlow的相关代码

Ubuntu/Linux 64-bit

```
bruce@bruce-desktop:~$ sudo apt-get install python-pip python-dev
```

Ubuntu/Linux 64-bit, CPU only, Python 2.7:

```
bruce@bruce-desktop:~$ sudo pip install --upgrade https://storage.googleapis.com/tensorflow/tflinux/cpu/tensorflow-0.8.0-cp27-none-linux_x86_64.whl
```

2) 然后下载安卓的SDK和NDK,并将其解压缩到你的TensorFlow目录里。

```
bruce@bruce-desktop:~$ wget https://dl.google.com/android/android-sdk_r24.4.1-linux.tgz
```

```
bruce@bruce-desktop:~$ tar xvfz android-sdk_r24.4.1-linux.tgz -C ~/tensorflow
```

下载额外的SDK编译工具

```
bruce@bruce-desktop:~$ cd ~/tensorflow/android-sdk-linux
```

```
bruce@bruce-desktop:~$ tools/android update sdk --no-ui
```

下载安卓NDK,并解压缩:

```
bruce@bruce-desktop:~$ wget https://dl.google.com/android/repository/android-ndk-r12b-linux-x86_64.zip
```

```
bruce@bruce-desktop:~$ unzip android-ndk-r12b-linux-x86_64.zip -d ~/tensorflow
```

下载分类器Inception

```
bruce@bruce-desktop:~$ cd ~/tensorflow
```

```
bruce@bruce-desktop:~$ wget https://storage.googleapis.com/download.tensorflow.org/models/inception5h.zip -O /tmp/inception5h.zip
```

```
bruce@bruce-desktop:~$ unzip /tmp/inception5h.zip -d tensorflow/examples/android/assets
```

为了使用安卓工具编译应用,需修改WORKSPACE文件

```
bruce@bruce-desktop:~$ gedit ~/tensorflow/WORKSPACE
```

拷贝下面的代码,并替换掉WORKSPACE文件里的相应行。

```
android_sdk_repository(
  name = "androidsdk",
  api_level = 24,
  build_tools_version = "24.0.3",
  path = "android-sdk-linux")
android_ndk_repository(
  name = "androidndk",
  path = "android-ndk-r12b",
  api_level = 21)
```

3) 开启USB调试和adb工具

为了测试调试设置是否成功,可以把手机连上电脑,并用下面的命令安装adb,并测试设备:

```
bruce@bruce-desktop:~$ sudo apt-get install android-tools-adb
```

```
bruce@bruce-desktop:~$ adb devices
```

编译APK

```
bruce@bruce-desktop:~$ cd ~/tensorflow
```

```
bruce@bruce-desktop:~$ bazel build //tensorflow/examples/android:tensorflow_demo
```

安装APK

至此,就可以在安卓设备上使用TensorFlow和Inception分类器了。需要注意的是,这个分类器会有偶尔出错的时候。原版的Inception分类器只能识别ImageNet挑战里出现的1000种图片分类。

2.3 TensorFlow的数据流图

TensorFlow 包含图、张量、操作、会话、变量等基本概念。其程序通常被组织成一个构建阶段和一个执行阶段。构建阶段,op的执行步骤会被描述成一个图;执行阶段,创建一个Session对象,使用会话执行图中的op,使用完成后关闭以释放资源。用数据流图可表示如图1:

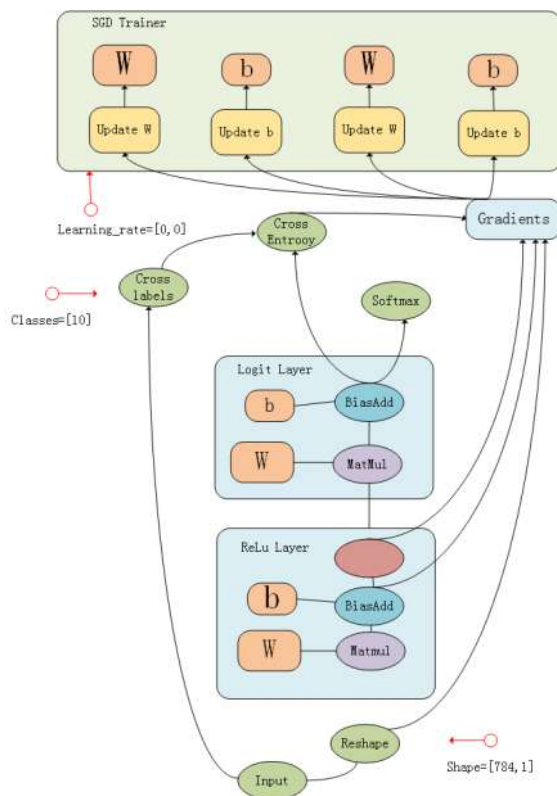


图1 TensorFlow的数据流图

3 卷积神经网络设计

LeCun等基于Fukushima的研究训练的LeNet-5模型是经典的卷积神经网络结构,后续的很多改进都是基于此模型的^[4]。李宏毅^[5]教授对此模型给出了以下3条解释。

- 1) Some patterns are much smaller than the whole image;
- 2) The same patterns appear in different regions;
- 3) Subsampling the pixels will not change the object.

3.1 卷积神经网络的基本结构

最早Yann LeCun教授提出的卷积神经网络,用于图像识别,充当分类器的功能。卷积神经网络LeNet5的结构,如图2 LeNet5结构图所示。

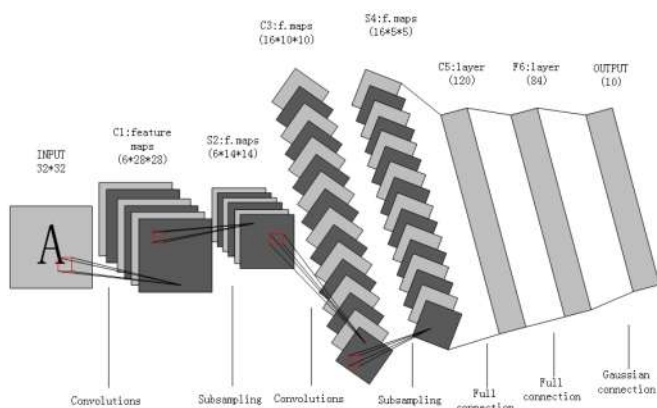


图2 LeNet5结构图

如图2 LeNet5结构图所示,卷积神经网络主要由5部分组成,分别是输入层、卷积层、池化层、全连接层和输出层^[6]。

输入层大小取输入图像的大小,若是彩色图像,则R、G、B三色分别用 x_1 、 x_2 、 x_3 表示。可以不对图像做预处理。

卷积层可通过卷积操作从小到大提取图像的特征,层数越多,特征的表达力越强。其具体操作,如图3卷积层操作示意图所示。

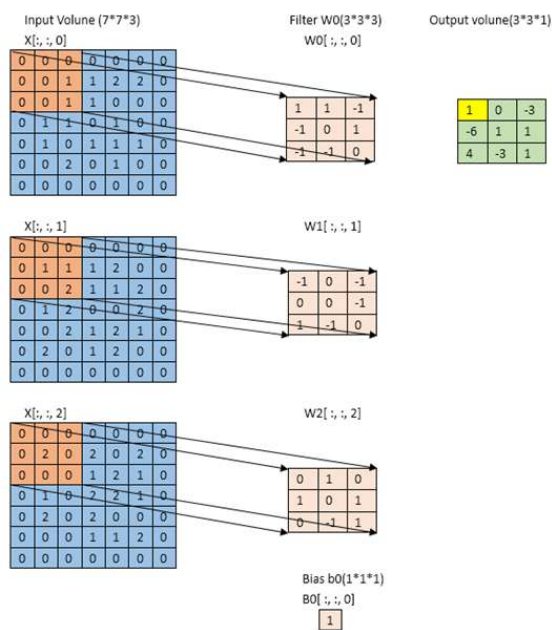


图3 卷积层操作示意图

池化层通过均值池化或者最大值池化,可以计算出卷积层输入的局部区域的重要特征值,实现降维和保持空间不变性,避免过拟合。池化层操作如图4池化层操作示意图。

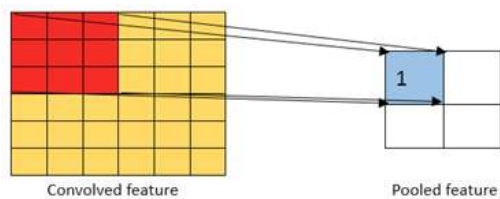


图4 池化层操作示意图

如图4池化层操作示意图,左侧黄色矩阵大小是6*6,要进行大小为3*3的池化(红色矩阵表示),对应到右侧矩阵中每个元素的值,是左侧红色矩阵每个元素的值求和后再除以红色矩阵的元素个数,即为平均值形式的池化。

全连接层会根据不同的应用情况,对前一层输出的特征进行映射。可选用sigmoid或tanh函数作为激活函数。

输出层的形式面向具体的应用。采用softmax回归就可以把卷积神经网络变成一个分类器,产生一个图像类别的向量或者属于哪一类的概率大小。

3.2 卷积神经网络的应用

卷积神经网络由于其强大的特征表达能力,被应用与图像识别、人脸识别、语音识别等方面。

判断给定图像属于哪个类别属于图像分类。He^[7]等人通过把空间金字塔池化层加入最后一个卷积层和第一个全连接层中,打破了传统CNN模型图像输入大小固定的限制,并在LSVRC-14的比赛中获得了很好的成绩。2016年,He^[8]等人采用快捷连接技术实现跨层连接的残差网络解决了退化问题,并在LSVRC-15的图像分类比赛中取得第一名。

在人脸识别中,DeepFace^[9]先对图像进行了3D人脸对齐的预处理,然后在把它输入的卷积神经网络中。FaceNet^[10]通过引入三元组损失函数进行端对端学习,在LFW和YouTube的人脸数据库中测试的准确率分别为99.63%和95.12%。DeepID3把人脸辨识-人脸确认的监督信号加入网络中间层,在最后的特征提取层中不共享权值,在LFW的人脸识别准确率为99.53%。

Hamid^[11]在CNN卷积层采用受限权值共享技术,并结合隐含马尔科夫链建立的语音识别模型比常规模型的识别错误率降低了10%。

3.3 移动设备下的卷积神经网络

对个人图片做分类,就必须用移动设备的图片文件数据重新训练,比如在安卓设备上使用定制化的分类器,让它学会通过识别不健康的叶子来发现营养不良或者有病害的植物。移植过程如下。

1) 编译图片训练优化器

```
bruce@bruce-desktop:~$ cd ~/tensorflow
bruce@bruce-desktop:~$ bazel build tensorflow/python/tools/optimize_for_inference
```

2) 对个人图片数据做训练

```
bruce@bruce-desktop:~$ bazel-bin/tensorflow/python/tools/optimize_for_inference
~input=tf_files/retrained_graph.pb \
~output=tensorflow/examples/android/assets/retrained_graph.pb
~input_names=Mul \
~output_names=final_result
```

3) 往 assets 目录下放训练后的标签

```
bruce@bruce-desktop:~$ cp ~/tensorflow/tf_files/retrained_labels.txt ~/tensorflow/examples/android/assets/
```

4) 编辑 java 文件,并替换相应变量

```
bruce@bruce-desktop:~$ gedit ~/tensorflow/tensorflow/examples/android/src/org/tensorflow/demo/TensorFlowImageListener.java
```

替换如下变量

```
private static final int INPUT_SIZE = 224;
private static final int IMAGE_MEAN = 128;
private static final float IMAGE_STD = 128;
private static final String INPUT_NAME = "Mul:0";
private static final String OUTPUT_NAME = "final_result:0";
private static final String MODEL_FILE = "file:///android_asset/retrained_graph.pb";
private static final String LABEL_FILE = "file:///android_asset/retrained_labels.txt";
```

5) 重新编译 APK

```
bruce@bruce-desktop:~$ cd ~/tensorflow
bruce@bruce-desktop:~$ bazel build //tensorflow/examples/android:tensorflow_demo.apk
```

重新安装 APK

```
bruce@bruce-desktop:~$ adb install -r -g bazel-bin/tensorflow/examples/android/tensorflow_demo.apk
```

至此,就可以使用个人定制的分类器在移动设备上进行图片分类的工作。

4 总结

TensorFlow自2015年开源以来,就被广泛应用与学术界,用于解决深度学习中神经网络的构建、参数优化等方面的困难。TensorFlow工作流程相对容易,API稳定,稳定性好,能与Numpy完美结合,并且能在各种类型的机器上运行,从超级计算机到嵌入式系统。谷歌对移动式TensorFlow平台的支持和研发,为极客提供了更大的开发平台和应用场所。卷积神经网络作为深度学习的一个热门分支,被应用于图像处理、语音识别、人脸识别等各个领域。以移动式TensorFlow为平台,构建并训练定制化的卷积神经网络成为了当前的一个研究热点。如何提高卷积神经网络模型的训练效率和优化构建方式,使其更好的应用于移动设备,是今后的进一步研究方向。

参考文献:

- [1] 李彦冬,郝宗波,雷航. 卷积神经网络研究综述[J]. 计算机应用, 2016, 36(9):2508-2515.
- [2] <https://github.com/tensorflow/tensorflow/tree/master/tensorflow/examples/android/>.
- [3] https://github.com/jikexueyuanwiki/tensorflow-zh/blob/master/SOURCE/get_started/os_setup.md.
- [4] Yann LeCun, Leon Bottou, Yoshua Bengio, et al. Gradient-based learning applied to document recognition. Proceedings of the IEEE, 1998, 86(11):2278-2324.
- [5] <https://www.youtube.com/watch?v=FrKWiv254g>.
- [6] 李旭冬,叶茂,李涛. 基于卷积神经网络的目标检测研究综述[J/OL]. [2017-01-13].
- [7] He Kai-Ming, Zhang Xiang-Yu, Ren Shao-Qing, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1904-1915.
- [8] He Kai-Ming, Zhang Xiang-Yu, Ren Shao-Qing, et al. Deep residual learning for image recognition// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, USA, 2016: 770-778.
- [9] Yaniv Taigman, Ming Yang, Marc' Aurelio Ranzato, et al. DeepFace: closing the gap to human-level performance in face verification//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Columbus, USA, 2014: 1701-1708.
- [10] Florian Schroff, Dmitry Kalenichenko, James Philbin. FaceNet: a unified embedding for face recognition and clustering//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, USA, 2015: 815-823.
- [11] Ossama Abdel-Hamid, Abdel-rahman Mohamed, Hui Jiang, et al. Applying convolutional neural networks concepts to hybrid NN-HMM model for speech recognition//Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Kyoto, Japan, 2012: 4277-4280.

(上接第173页)

从交互设计来说,需要考虑LED显示屏和智能手机APP应用的系统、操作、功能、界面尺寸与用户体验等,特别是智能手机APP应用需要对常见的两款智能手机系统进行研究,iOS和Android两种系统存在着明显的差异性,如显示屏的尺寸大小、操作方式、功能键等,以满足不同载体的各种要求。同时以用户体验为中心进行人性化设计也非常必要,可以采用对信息进行分类,对色彩进行和谐搭配,合理编排图形元素,创建清晰的导航和层级关系,和明确的交互方式,譬如界面信息分类按板块区域划分来显示,公交车辆实时情况为固定区域,而新闻娱乐及其他广告信息等可以分为不同时间段来进行滚动播放,同时设置搜索的功能,以此来优化用户体验。研究智能公交站牌的界面设计,需要根据各个公交线路、公交管理站点、区域人群、客流量、商户等来进行全方位的考量,让智能公交站牌无论是站点的LED显示屏的应用,还是随时随地可查阅信息的APP应用,都能够和用户建立起一种有机关系,使用户更加便捷智慧地乘车出行,全面刷新用户的体验与感受。

5 建设智慧城市的相关思考

公交站牌采用智能控制系统以及先进的科技信息显示,实时报告公交车运行状况,以满足和方便乘客需要,解决乘客在车站焦急等车的状况,这让城市居民的生活更加智慧和便捷。车载GPS全球定位系统,LED液晶显示屏,智能手机和APP应

用等技术和载体,让智能化公交站牌成为可行,同时也将凸显视觉应用设计的重要性和便捷性。

将LED显示屏的应用和智能手机APP应用给予结合,且在视觉设计中考虑区域特色,把西安城市印象、文化理念融入研究设计当中,拓展公众使用方式和适用范围,同时也推动数字信息时代经济的发展。这项研究以人为本服务于人,符合现阶段科学技术的发展,艺术设计视觉应用的研究,以及处于数字信息时代下我国经济建设的国情需求,建设智慧可持续发展的城市是必然和必要的。

参考文献:

- [1] (英)David Wood. 国际经典交互设计教程:界面设计[M]. 电子工业出版社,2015.
- [2] (英)维克托·迈尔-舍恩伯格,(英)肯尼斯·库克耶,盛阳燕,周涛. 大数据时代:生活、工作与思维的大变革[M]. 浙江人民出版社,2013.
- [3] 丁蕾. 智慧城市——旅游城市住宿预警视觉应用研究 [J]. 美与时代,2016(09).
- [4] “智慧城市”词条[EB/OL]. <http://baike.baidu.com/>
- [5] 上海正先智能公交电子站牌 走进“江城”芜湖[EB/OL]. <http://www.shuyang.tv/news/qiye/2017-07-12/2066.html>
- [6] “LED显示屏”词条[EB/OL]. <http://baike.baidu.com/>