学号：2020104238

姓名：黄人煌

课程报告——基于深度学习的商品检索系统

1.简介

近年来,随着互联网技术的发展,以及各种拍摄硬件的普及,使得各行各业中的图像数据都呈现爆炸式增长,从海量数据中找出真正符合需求的图像也变得越来越困难。

基于内容的图像检索系统CBIR(Content-based image retrieval)是被广泛关注和研究的计算机视觉领域的一个分支。其根据大规模的数字图像内容，在已有的图像数据集中查找到具有相同或者相似内容的图片。目前业界比较具有代表性的图像搜索应用场景主要包括:

1. 以电商为代表的商品购物搜索;
2. 以谷歌为代表的海量互联网图片搜索;
3. 基于移动设备的物体或者商品搜索。

鉴于深度学习在图像识别、检测、分类等领域的突出成就,本文设计了基于深度学习的商品检索系统,使得用户可以快速获取类似商品信息。

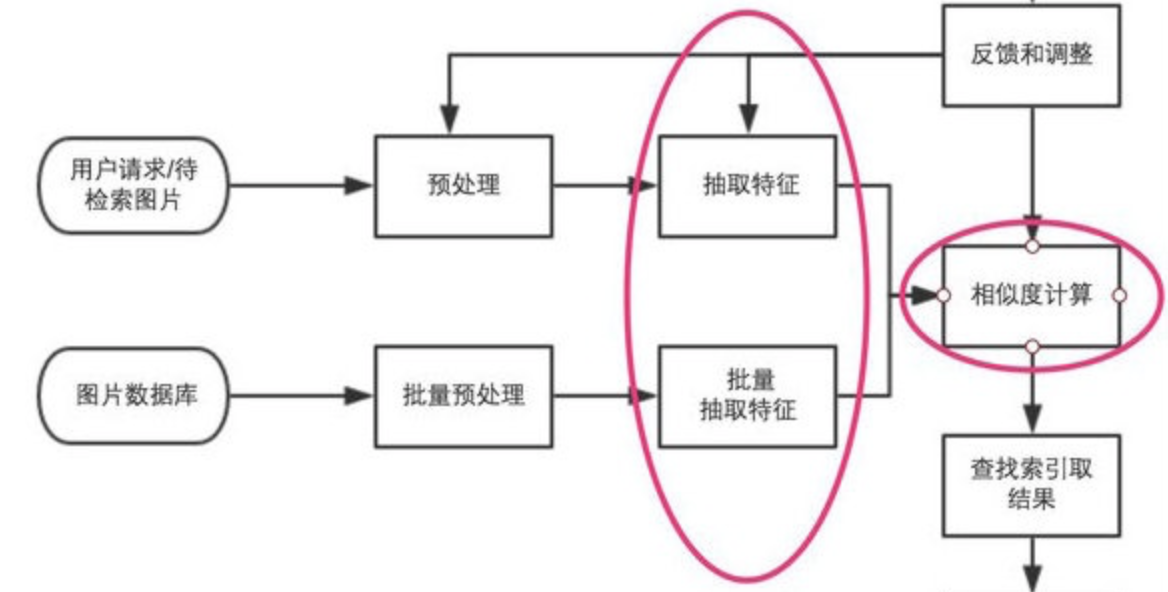


图1 整体架构

2.实现细节

2.1开发工具以及运行环境

IDE：Pycharm

Python：3.7

Packages：Keras + TensorFlow + Pillow + Numpy

运行环境：Ubuntu

2.2选取图像特征

传统的检索通常采用颜色、形状、纹理、能量等，我们使用深度学习模型（vgg16, resnet50,densenet121）进行特征提取，并使用Keras的预训练模型。

Keras是一个高层神经网络API，Keras由纯Python编写而成并基于Tensorflow、Theano以及CNTK后端。

2.3索引化数据集

对图像库中的每张图片抽取特征，存储在数据库中，并保存为H5文件。

h5文件是层次数据格式第5代的版本（Hierarchical Data Format，HDF5），用以存储和组织大规模数据。

H5将文件结构简化成两个主要的对象类型:

1. 数据集dataset,就是同一类型数据的多维数组。

2. 组：group,是一种容器结构,可以包含数据集和其他组,若一个文件中存放了不同种类的数据集,这些数据集的管理就用到了group。

2.4定义相似度

常用的相似性度量有：欧氏距离、曼哈顿距离、闵可夫斯基距离、切比雪夫距离、夹角余弦等。这里采用的夹角余弦。

夹角余弦(Cosine)：几何中夹角余弦可用来衡量两个向量方向的差异，机器学习中借用这一概念来衡量样本向量之间的差异。

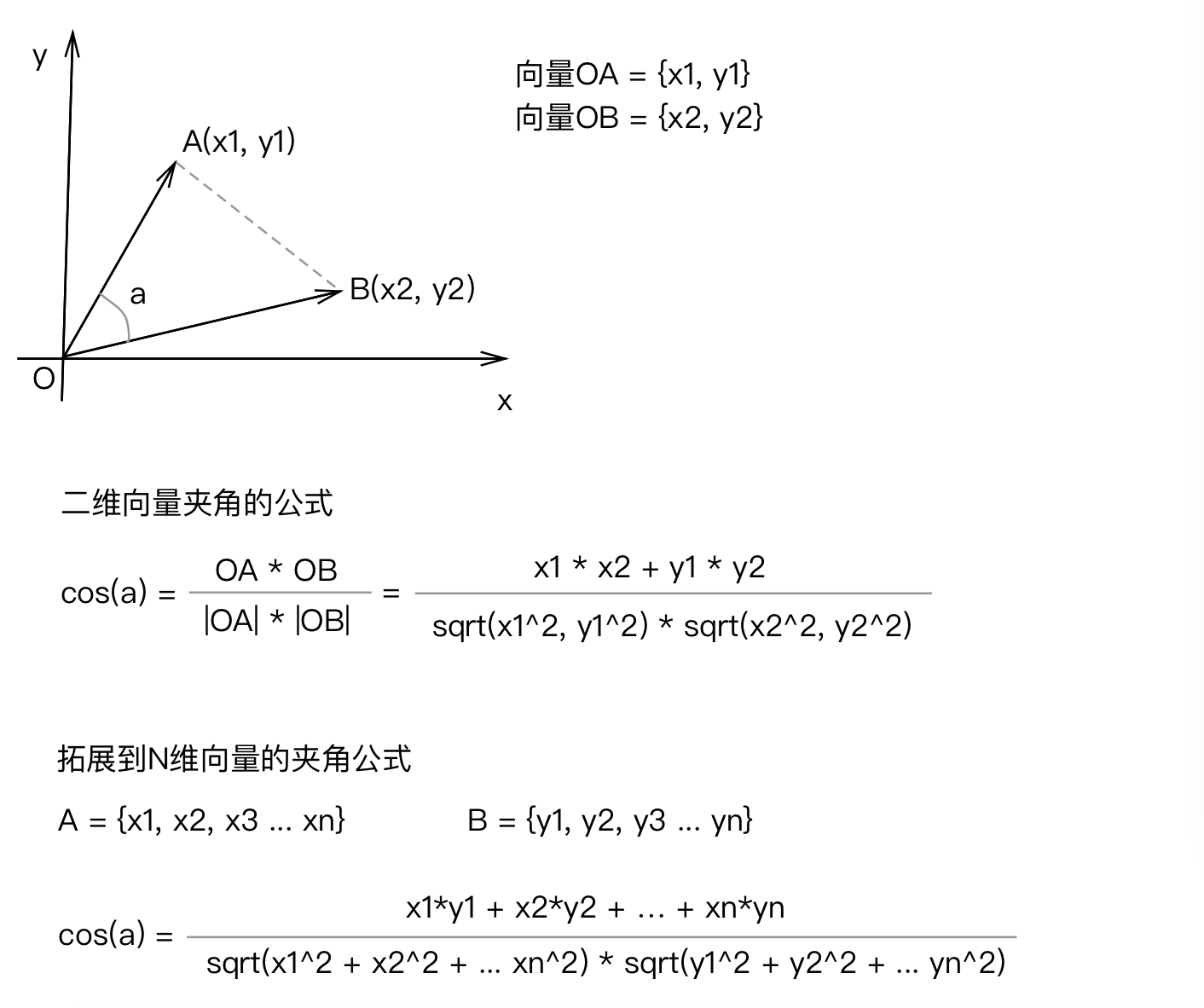


图 2 余弦夹角公式

2.5检索

首先提取测试图片的特征，再与数据库中的向量逐一计算距离，然后排序，取相似度Top3。

具体过程：

1. 选择图片

图3测试图片

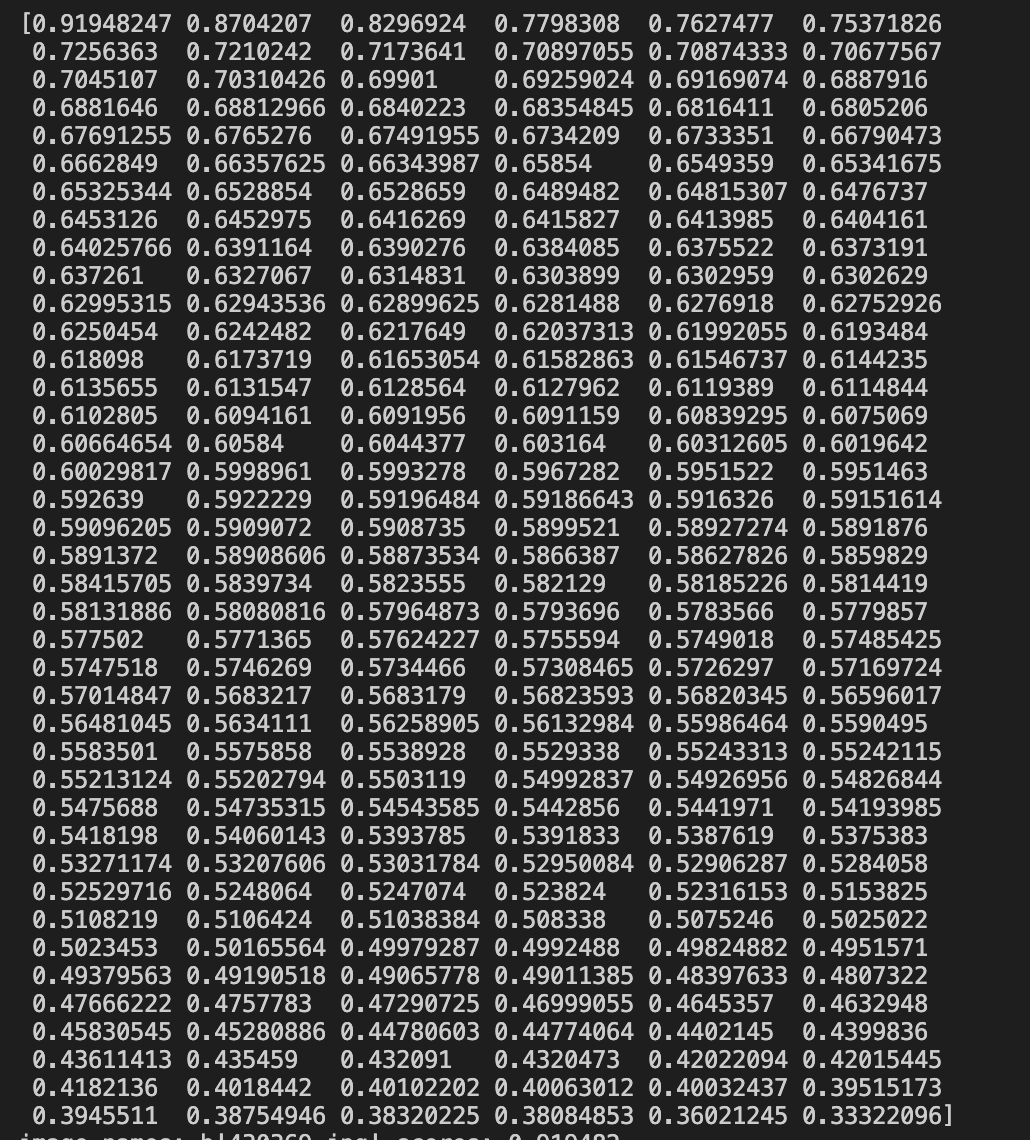
1. 数据库计算结果

图4计算结果

1. 排序



图5相似度最高的3个商品

3.课程总结

中间件这门课程我学到很多网站开发技术，我之前对C++语言较为熟悉，对Python语言和人工智能相关技术不太了解，所以实现的基于深度学习的商品检索系统还有待完善，尽管如此，通过动手实践，我对深度学习有了初步的了解，对我今后的使用和继续学习都有很大的帮助，相信在以后工作或学习中还会用到相关知识。

最后感谢老师的辛苦的付出。

4.参考文献

[1] https://zhuanlan.zhihu.com/p/80886534

[2] https://blog.csdn.net/qq\_19707521/article/details/78479532

[3] https://www.zhihu.com/question/25833727