高级数据结构课程实验一报告

洪方舟

2016013259

hongfz16@163.com

李帅

2016013270

lishuai16THU@163.com

周展平

2016013253

zhouzp16@163.com

**摘要**

本文对五篇图像检索相关论文进行综述，并且提出了我组最终项目的设想及架构。

**关键词**

图像检索；深度学习；二进制哈希；

# 论文综述

本文作者选取了三篇图像检索系统论文，一篇AlexNet深度学习网络论文，一篇模糊近邻查询综述论文。

## Deep learning of Binary Hash Codes for the Fast Image Retrieval

### Motivation

在图像检索问题中，基于哈希的算法常常需要通过一个相似度矩阵来计算两幅图像之间的相似程度。然而，相似的矩阵的构建与计算会耗费大量的时间与空间资源。受到深度学习算法的启发，论文作者尝试利用深度学习提取哈希特征。

### Framework

整个算法由以下几部分组成：首先，训练ImageNet网络；然后在CNN的最后添加一层与产生哈希特征表示相关的神经网络层；最后，可以利用训练好的网络进行检索，检索的步骤为：（1）将图像输入到训练好的网络中，得到一个输出的向量 （2）将向量的每一维量化为0或1 (3)用Hamming距离衡量图像之间的相似程度，设置合适的阈值，得到candidate set （4）计算candidate set中的图像与查询图像的特征之间的距离，这里的特征是指神经网络输出的未经量化的向量。根据欧氏距离对图像进行排序输出。

## Class-Weighted Convolutional Features for Visual Instance Search

### Motivation

之前许多利用CNN进行图像检索的方法都需要针对给定的数据集进行fine-tune,一些研究也发现了网络中蕴含的有关空间信息的知识，然而对于用于fine-tune的数据集的处理花费大量的精力。论文作者希望仅仅使用网络中的知识完成fine-tune步骤。

### Framework

整个算法有以下几部分组成：首先，训练CNN网络；然后输入用于fine-tune的图像集，针对每一个类别利用CAM加权计算出新的特征，并通过求和池化增强效果；之后，对于特征向量中的维度加权计算出新的向量，以此减少通道冗余；最后，依次进行I2 normalization、PCA-whitening、I2 normalization，将各个类别得到的特征描述合并成一个向量。

## ImageNet Classification with Deep Convolutional Neural Networks

该论文详细讲解了2012年ImageNet竞赛第一名的网络AlexNet。相较于之前较为成功的深度学习网络LeNet-5，该模型中提出了很多新的思想，使得深度学习重新受到人们的关注。

### Network Structure

该论文提出了如下创新结构，提高了学习的效率与效果：1. ReLUs线性整流函数，提高了梯度下降的效率，一定程度上解决了梯度爆炸与梯度消失的问题；2. 局部归一化，避免了对输入数据归一化的需求，使得模型能够更好的泛化；3. 重叠池化，使得模型更加难过拟合。

### Training Methods

同时该论文还提出了多种提高训练效率的模型训练方法：1. 数据集增强，减少过拟合，更好的捕捉原始图像的重要特征；2. Dropout方法，减少参数个数，防止过拟合。最终该网络应用于LSVRC-2010数据集，将Top-1错误和Top-5错误分别降到了37.5%和17.0%。

# 图像检索架构的提出

综合阅读的论文中的方法，我们提出如下图像检索系统的架构

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

# REFERENCES

1. Lin, K., Yang, H. F., Hsiao, J. H., & Chen, C. S. (2015). Deep learning of binary hash codes for fast image retrieval. 27-35.
2. Jimenez, A., Alvarez, J. M., & Giro-I-Nieto, X. (2017). Class-weighted convolutional features for visual instance search.
3. Noh, H., Araujo, A., Sim, J., Weyand, T., & Han, B. (2017). Large-Scale Image Retrieval with Attentive Deep Local Features. IEEE International Conference on Computer Vision (pp.3476-3485). IEEE.
4. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems (pp. 1097-1105).
5. Li, W., Zhang, Y., Sun, Y., Wang, W., Zhang, W., & Lin, X. (2016). Approximate Nearest Neighbor Search on High Dimensional Data---Experiments, Analyses, and Improvement (v1. 0). arXiv preprint arXiv:1610.02455.