基于支持向量机的图像分类算法

赵思源（919101960141 机械工程学院）

**摘要**：

**关键字**：图像分类、主成分分析、支持向量机、神经网络、最近邻算法

1.引言

伴随着计算机与模式识别技术的高速发展，图像分类在愈来愈多的领域都有着广泛的应用，如：安防领域的人脸识别和智能视频分析，交通领域的交通场景识别，互联网领域基于内容的图像检索和相册自动归类，医学领域的图像识别等。．可以说，图像分类技术已经应用于人们日常生活的方方面面，计算机自动分类与检测技术也在一定程度了减轻了人的负担，改变了人类的生活方式。

图像分类作为计算机视觉重要的基础问题之一，是物体检测、图像分割、物体跟踪、行为分析、人脸识别等其他高层视觉任务的基础：欲对实际复杂场景进行自动分析与理解，首先就需要确定图像中存在什么物体，因此研究准确且高效的图像分类算法有着重要的理论意义和实际意义。

2.研究现状

目前应用较为广泛的图像分类技术主要分为传统的诸如K-最近邻算法、支持向量机机器学习方法和目前高速发展的卷积神经网络模型，自LeNet5[1]网络提出以来，在 AlexNet[2]引起了大量关注后，应用于图像分类的卷积神经网络技术高速发展，逐渐诞生了GoogLeNet[3]模拟神经网络对猫的关键特征进行自行学习并分类、ResNet[4]解决了增加深度带来的退化问题、SENet[5]方便集成并提升性能等一系列优秀的模型。

但是，由于神经网络有大量的参数，容易在训练数据集规模较小时发生在训练集上准确率很高，在测试集上效果不尽人意的过拟合问题。且计算资源有限，训练网络需要花费很长的时间。支持向量机的优势由此体现：作为1992年和1995年发表的技术[6]，支持向量机在当时并没有立刻引起很大的效应，但其与传统模式识别以及人工神经网络相比，能够在样本数相对较少，特征维数高的情况下仍然取得很好的推广能力。结合模糊支持向量机、基于决策树的支持向量机、分级聚类的支持向量机等理论改进[7]，支持向量机在图像领域很常见，因为该领域很容易体现支持向量机的优势。

3.图像分类方法

本图像分类算法的流程如下：

3.1图像特征提取

在计算机中，图像以有序的多维数组形式储存，即使一张很小的图片，也需要大量的存储空间保存图像的全部信息。直接以图像的存储矩阵数据进行各类计算会造成大量的资源浪费与计算困难。因此，对图像特征进行提取可以用较低的维数概括图像的信息。使用图像特征而非原始图像数据为数据的存储和使用带来了极大的方便。

3.1.1色彩空间(HSV)转换

HSV是一种比较直观的颜色模型，不同于传统的RGB颜色通道，HSV颜色的参数为色调(H)，饱和度(s)，明度(v)，此形式下的图像更接近人们对彩色的感知经验，可以非常直观的表达颜色的色调、鲜艳程度与明亮程度；同时，相比RGB空间，在此空间下的图像颜色特征更加突出，方便根据图像的颜色特征进行识别、比对与分类。

说明图：hsv

为方便后续图像处理，用R、G、B代指原始图片像素点的各通道强度，并定义：



使用以下公式对原始数据集进行从RGB空间到HSV空间的转换：



3.1.2梯度直方图(HOG)特征

HOG特征是一种在计算机视觉和图像处理中用来描述图像梯度的特征，能够很好地描述图像中边缘部分的方向密度。

首先对原始的彩色图像进行灰度化处理并对颜色空间进行标准化处理，由此调节图像的对比度，起降低图像光照变化影响、消除噪声的作用。然后使用sobel算子依次计算划分后子图像在水平和竖直两方向的梯度，并由此计算图像梯度的大小与方向：



由于局部光照的变化以及前后景对比度的变化使得梯度强度的范围很大，因此接下来对图像的梯度进行归一化处理，通过将图像分为若干单元格并统计单元格中梯度之和，再将各个单元格组成大的、空间上连通的区间，将该区间中所有单元格的梯度串联起来，由此获得图像的HOG特征。

在原始图像上作出hog特征

3.1.3合成特征向量

在通过上述流程计算获得图像的HSV特征用以代表原始图像的颜色特征和HOG特征代表原始图像的边缘信息后，将每张图片的HOG和HSV特征向量连接起来并归一化，构成一个总的数据集的向量矩阵，该矩阵的每一行是单个图像的HSV+HOG特征。使用此向量矩阵代替原始的图像RGB特征，根据以往实验经验，HSV+HOG特征相比RGB，更好的反映了图像本身的特征信息，能够有效提高后续图像分类的正确率。

对比图：有无进行处理正确率区别

3.2 主成分分析法(PCA)降维

主成分分析法可以在保留原始数据的大量特征信息的前提下降低图像的维度，方便后续处理。假定原始矩阵为M，根据以下公式：



将原始矩阵进行奇异值分解，中间的非负实数对角矩阵对角线上的矩阵即为原始矩阵M的奇异值，通过奇异值的大小可以确定各维度包含原始数据信息量的大小。

图像：pca维数与包含信息量

图像：人脸各维度包含信息

维数越多，信息利用率也越高，但处理起来也更加困难，因此需要选择合适的维数。

完成PCA主成分分析法降维维数确定后，利用奇异值分解获得的酋矩阵各列代表的主成分方向，可将原始数据进行坐标变换，投影之PCA空间，完成主成分分析。

3.3 支持向量机构建

3.3.1核函数选择

3.3.2误差惩罚参数选择

4.实验结果

4.1与k-最近邻分类对比

4.2与人工神经网络对比

5.结束语

参考文献：

* [1]Y. LeCun,L. Bottou,Y. Bengio,P. Haffner. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE,1998,86(11).
* [2]Christian Szegedy,Wei Liu 0015,Yangqing Jia,Pierre Sermanet,Scott E. Reed,Dragomir Anguelov,Dumitru Erhan,Vincent Vanhoucke,Andrew Rabinovich. Going Deeper with Convolutions.[J]. CoRR,2014,abs/1409.4842.
* [3]Kaiming He,Xiangyu Zhang,Shaoqing Ren,Jian Sun 0001. Deep Residual Learning for Image Recognition.[J]. CoRR,2015,abs/1512.03385.
* [4]Hu Jie,Shen Li,Albanie Samuel,Sun Gang,Wu Enhua. Squeeze-and-Excitation Networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2019,42(8).
* [5]Corinna Cortes,Vladimir Vapnik. Support-Vector Networks.[J]. Machine Learning,1995,20(3).
* [6] V. N. Vapnik, An overview of statistical learning theory, IEEE Trans. Neural Netw. 1999, 10, 988-999.
* [7]黄勇,郑春颖,宋忠虎.多类支持向量机算法综述[J].计算技术与自动化,2005(04):61-63.