

谨以此论文献给我的导师和亲人！

————— 卢婧宇

细粒度图像分类

学位论文答辩日期： 2017 年 05 月 28 日

指导教师签字：

答辩委员会成员签字：

独 创 声 明

本人声明所呈交的学位论文是本人在导师指导下进行的研究工作及取得的科研成果。据我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含未获得 _____（注：如没有其他需要特别声明的，本栏可空）或其他教育机构的学位或证书使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示谢意。

学位论文作者签名： 签字日期： 年 月 日

学位论文版权使用授权书

本学位论文作者完全了解学校有关保留、使用学位论文的规定，并同意以下事项：

1. 学校有权保留并向国家有关部门或机构送交论文的复印件和磁盘，允许论文被查阅和借阅。
2. 学校可以将学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文。同时授权清华大学“中国学术期刊(光盘版)电子杂志社”用于出版和编入 CNKI《中国知识资源总库》，授权中国科学技术信息研究所将本学位论文收录到《中国学位论文全文数据库》。

(保密的学位论文在解密后适用本授权书)

学位论文作者签名：

签字日期： 年 月 日

导师签字：

签字日期： 年 月 日

细粒度图像分类

摘 要

细粒度图像分类是近几年新出现的一个研究方向，也是计算机视觉领域中的一个极其热门的研究领域。细粒度图像分类是对属于同一层次（例如不同的鸟种），并且具有相似的形状或视觉外观的分类对象进行分类的任务。虽然它与一般对象分类有关，但细粒度分类要求算法在通常仅由细微差别区分的高度相似的对象之间进行区分。伴随着近年来研究的越来越深入，目前已经有很多方法可以实现。本文的主要内容如下：

1. 介绍细粒度图像分类的定义以及研究目的，它与传统图像分类的区别以及在实际生活中的研究价值。目前网络中图像信息越来越多，针对这些信息的处理也变得更加重要，而细粒度图像分类是在传统图像分类基础上的深入研究。再总结它的发展现状、传统经典方法以及目前的研究进展。传统的细粒度分类方法依赖大量的人工标注，而如今的方法仅仅依靠类别标签就可以到达很高的准确率。以及介绍了目前发展所遇到的难题和解决方式。
2. 用传统方法在常用数据集上实现细粒度图像分类，本文采用 CUB-200-2011 鸟类数据集进行实验，并且选用三种不同方式进行特征提取。先介绍这三种方法的理论知识，再进行实验。输入的图像分为两种形式，一种是未经过处理的原始图像，另一种是添加了 Bounding Box 标注框信息的图像。提取特征部分采用 SIFT、HOG 和 LBP 三种方法，之后都经过词袋模型处理，最后通过支持向量机进行分类。并且分类器部分又选用了两种核函数进行实验，分别为径向基核函数和线性核函数。
3. 由实验得到的结果进行分析对比。当输入两种形式的图像时，观察每种方法得到的准确率的变化。当只用不同的核函数进行实验时，对比这两种核函数作用下的效果。当只有提取特征方式不同时，对比结果的准确率，找到较好的方法。并且对输出的混淆矩阵进行分析，将结果可视化，通过对矩阵的观察可以发现不同的方法对于各个类别的分类情况。

经过实验分析，虽然本文所采用的传统方法对于 CUB-200-2011 数据集的分类结果准确度偏低，但是可以从实验中进一步体会到传统方法在细粒度图像分类问题上的贡献。

关键词： 细粒度图像分类; 特征提取; 词袋模型; 支持向量机

Fine-Grained Image Classification

Abstract

Fine-grained image classification is a new research direction in recent years, and is also a very active research field in the field of computer vision. Fine-grained image classification is a task to classify objects that belong to the same level (such as different bird species) and have similar shapes or visual appearance. Although it is related to general object classification, the fine-grained classification requires that the algorithms distinguish highly similar objects that are usually only distinguished by subtle differences. With the deepening of research in recent years, there are many ways to achieve. The main work of this paper is as follows:

1. First, this paper introduces the definition of fine-grained image classification, its research significance, the difference between traditional image classification and the research value in real life. At present, there are more and more image information in the network, and the processing of these information becomes more important, and the fine-grained image classification is based on the traditional image classification. After that, this paper summarizes its development status, traditional classical methods and current research progress. Traditional fine-grained classification methods rely on a large number of manual annotation. But now the methods which only rely on the category labels can reach very high accuracy. And it introduces the problems and solutions to the present development.
2. Using the traditional method to achieve fine grain image classification on the common data set, this paper uses the CUB-200-2011 Birds bird data to set experiment, and selects three different ways to carry on the feature extraction. First, the theoretical knowledge of the three methods are introduced, and then experiments are carried out. The input images are divided into two forms, one is unprocessed raw image, the other is to add the image of bounding box labeled box information. The extracted feature part uses SIFT, HOG and LBP methods, after the word bag model processing, finally uses support vector machine. In addition, the classifier uses two kernel functions, radial basis function and linear kernel function respectively.
3. The results obtained by the experiments are analyzed and compared. When entering two forms of image, we observe the changes of the accuracy rate obtained by each method. When experiments are performed with different kernel functions, the effect of these two kernel functions is compared. When only the extraction of the

characteristics of the method is different, we compare the accuracy of the results and find a good method. This paper also analyzes the confusion matrix, which can make the results visible. Through the observation of the matrix, we can find different results for each category of classification.

Through the experimental analysis, although the traditional methods used in this paper for the CUB-200-2011 data set classification results accuracy is low, but we can understand the contribution of traditional methods in the fine-grained image classification problems.

Keywords: Fine-Grained Image Classification; Feature Extraction; Bag-of-Words Model; Support Vector Machines

目 录

1 绪论	1
1.1 细粒度图像分类的研究背景与意义	1
1.2 细粒度图像分类的研究现状	1
1.3 本文的研究内容及组织结构	3
2 细粒度图像分类基础	4
2.1 概念	4
2.2 流程框架	4
2.3 数据集	5
3 细粒度图像分类方法	7
3.1 图像特征提取	7
3.1.1 尺度不变特征变换	7
3.1.2 方向梯度直方图	9
3.1.3 局部二值模式	10
3.2 词袋模型	11
3.3 支持向量机	12
3.3.1 支持向量机的原理	12
3.3.2 核函数	14
4 实验及分析	16
4.1 实验设置	16
4.2 三种方法的实验分析	17
4.2.1 SIFT+BoW+SVM	17
4.2.2 HOG+BoW+SVM	20
4.2.3 LBP+BoW+SVM	22
4.3 实验总结	24
5 总结与展望	26
5.1 论文总结	26
5.2 展望	26
参考文献	27
致 谢	29

1 绪论

1.1 细粒度图像分类的研究背景与意义

本世纪是一个信息大爆炸的时代，伴随着计算机和网络技术的快速发展，全球互联网进入中速发展期，其用户规模实现平稳增长，现实生活中所接收到的信息也越来越多，尤其是多媒体信息。并且根据 2017 年互联网中心发布的一份关于互联网发展的权威报告^[1] 截至 2016 年 12 月，中国网站数量为 482 万个，年增长 14.1%。而作为目前媒体信息中出现的最多的图像信息，由于其包含大量信息，更让人们关注。图片信息之庞大，面对这种复杂的局面，仅仅依靠人工进行图像信息筛选十分困难，所以要借用计算机技术来帮助人们解决问题是非常必要的。人类可以使计算机了解图像，这种方法就是图像分类，这就说明对图像分类这个领域的探索具有巨大的实际价值。并且近几年计算机视觉、人工智能等领域逐渐进入人们的视野，这使图像分类技术有了新的进展，所获得的成果也体现在生活中的各个方面。

图像分类的对象是从确定的类别标签的集合中预测给出图像所属于的种类。在计算机视觉的研究过程里，图像分类主要分为粗糙分类和细分类两种形式。所谓的粗糙分类，就是传统图像分类，用来区分物体属于哪一个物种。细分类就是细粒度图像分类，用来区分对象属于物种中的哪一类。经过神经科学的逐渐发展和进步，研究人员认识到人类识别物体是借助物体具有区分度的局部特征，并且把它们组合起来实现的。这说明，学习一个高层概念需要建立结构性的模型。但是，传统的图片分类结果包括了大部分的无用的信息，这种情况下抑制了计算机算法经过一定次数的训练数据学习得到结构性的模型。此时图像分类中的一个更加深入的领域——细粒度图像分类 (Fine-Grained Image Classification) 走进我们的视线。在细粒度图像分类领域，待分类图像的所属类别的大体特征很相近，因此研究者可以从全部图像数据集的共同拥有的特征信息中学习到更好的层次结构模型。

1.2 细粒度图像分类的研究现状

传统图像分类的方法依赖于大量的人工标注，结果不理想而且过程还很繁琐，准确度大部分情况下依赖标注者的主观意识。并且，人工标注信息耗费的人力时间都很多，所以目前新兴的方法大部分偏向弱监督，无需人工标注只需要标签就可以自动训练，并且有一些算法的准确度可以达到百分之八十多。

比较经典的算法是, 2011 年 Wah^[2] 等人在发布 CUB-200-2011 数据库的技术报告中所做的基准测试, 其结果仅为 10.3%。算法过程是, 使用一幅原始的、未经过任何标注的测试图片, 利用训练过程得到的模型进行局部区域位置的定位; 之后提取特征, 经过词袋 (Bag-of-Words, BoW)^[3] 模型进行特征编码后, 输入到分类器完成分类。但如果在测试时给定了标注框和局部区域位置这些标注信息的话, 利用同样的方法, 得到的基准测试结果为 17.3%。之后, Berg T^[4] 等人提出一种方法基于部位的一对一特征, POOF 可以自主地从一组特定区域的具有固定位置和类别标注信息的图片集中学习不同的具有高度区别的中等特征。这种方法精确度最高可以达到 73.3%。

但是, 传统方法的准确度仍然不理想, 近年来兴起的深度学习为我们在图像分类领域指明了新的方向。总的来说, 深度学习对数据特征的无监督学习, 使它具有深层复杂的网络结构布局。现在深度学习技术在图像分类领域的应用已经展现出非常明显的优势。

特征表示是细粒度识别的一个关键问题, 卷积神经网络 (CNN)^[5] 被广泛应用于特征提取。然而, 细粒度表示存在两个挑战。首先, 传统的 CNN 表示需要固定大小的矩形作为输入, 这不可避免地包括背景信息。然而, 背景不太可能对细粒度的识别起任何重要作用, 因为所有的子类都有相似的背景。针对这个问题 Xiaopeng Zhang^[6] 等人提出了一种基于深度滤波的细粒度图像分类方法, 将 CNN 的深层过滤反应作为局部描述符, 并通过 SWFV-CNN 对其进行编码。第一步的目标是挑选深度过滤器, 以显著和一致地响应特定的模式。第二步是通过 fisher 向量的空间加权组合来选择 CNN 滤波器。实验结果表明, SWFV 的性能优于传统的 CNN, 与传统的 CNN 相辅相成, 进一步提高了性能。

现有的基于 CNN 的细粒度分类方法并不侧重于对局部语义的检测和利用。所以对于鸟类幼崽的分类来说识别头部的效果总是比身体差, 因为头部的尺寸很小。传统的 part-based CNN 方法对于部分网络之间的卷积过滤器的共享有困难。来自俄罗斯大学的 Han Zhang^[7] 等人关注到富有语义的局部, 认识到大多数卷积神经网络缺乏模型化对象语义部分的中层。故提出了一种新的 CNN 架构, 构建了一个 End-to-End 的网络, 将语义部分检测与抽象相结合进行细粒度分类。该网络有两个子网络, 一个用于检测一个用于识别, 该检测子网络具有一种新颖的自顶向下的生成小语义部分候选检测方法。分类子网络引入了一种新的部件层, 从检测子网络检测到的部分提取特征, 并结合它们进行识别。这种部分特定的学习为更深入地了解细粒度类别打开了门, 不仅仅是识别类标签。

目前, 大部分方法都侧重于深度学习, 这使细粒度图像分类的准确度有了很

大的提高。实现的方法有很多这里就不一一赘述了。

1.3 本文的研究内容及组织结构

本文共分五章，具体内容介绍如下：

第一章为绪论，对细粒度图像分类领域的研究背景与研究意义进行综述，简单阐述了细粒度图像分类方法的国内外研究现状，为后文各个章节内容的介绍作铺垫。

第二章主要介绍细粒度图像分类技术的基础知识，基本概念，细粒度图像分类方法的概述及时下常用的几种用于细粒度图像分类的数据库。

第三章主要介绍细粒度图像分类问题的传统解决方法，介绍三种提取特征的方法，词袋模型和分类器。

第四章主要介绍用第三章所提到的三种方法进行实验和实验分析。

第五章主要对论文进行总结和未来的展望。

2 细粒度图像分类基础

2.1 概念

细粒度图像分类作为一个近几年热门的研究方向，越来越受到各方面的关注。细粒度分类属于目标识别的一个子领域，它的目标是区分属于同一基本层次范畴的数百个子类。虽然它与一般对象分类有关，但细粒度分类要求算法在通常仅由细微差别区分的高度相似的对象之间进行区分。这部分问题的一个相同点是不同类别之间的相似度较高，相同类别内的差异性较大，所以细粒度图像分类与传统的图像分类相比难度更大。因此，对象和对应局部的定位和描述成为细粒度识别的关键。目前，精细图像分类的研究工作主要集中在花卉、鸟类、狗类等，而且也得到了很多的研究成果。细粒度图像分类研究的是需要掌握一定的专业常识才能进行对象分类的问题研究。由于近些年计算机发展的很快，又有很多新的方法出现解决了以往存在的种种方面的缺陷。

2.2 流程框架

与普通的图像分类不同，细粒度图像的信噪比较小，只有在细小的局部区域中可以找到包含足够区分度的信息，因此如何高效的找到并且利用这些信息是关键。如今，大多数的分类算法依据的流程如下：首先找到前景对象（鸟）及其局部区域（嘴、翅膀等），然后分别对这些有用区域进行特征提取，再将所提取的特征进行一定的处理，最后用分类器进行训练和预测。具体流程框架如图 2-1所示。

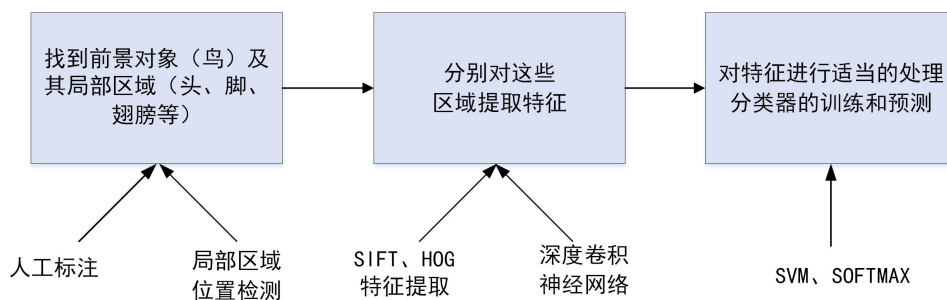


图 2-1 流程框架图

其中，第一步找到需要的前景对象，可以用到人工标注和局部区域位置检测的方式。提取特征时，有很多可以选择的特征算子比如 SIFT^[8]、HOG^[9] 特征并且

现在的大部分方法都把重点放在深度卷积神经网络上。第三步对特征进行适当处理和分类需要用到 SVM^[10] 等分类器。但具体运用哪种需要根据程序来决定。

2.3 数据集

目前大部分的数据集都是针对普通的图像分类设计的，细粒度图像分类数据库的获取相比之就难度非常大了，因为需要很丰富的专业知识才能完成数据的采集和标注。但随着细粒度图像分类领域的研究的越发深入，也有越来越多的细粒度图像数据库出现，这体现了细粒度图像分类在近年来发展之迅速。

目前比较常用的细粒度图像数据库，本文简要说明如下：

CUB-200-2011^[2]：来自加利福尼亚理工学院，数据集的内容是鸟类图像，包含 200 种不同类别，共 11788 张图像数据。关于标注情况，数据集提供了丰富的人工标注数据，每张图像包含 15 个局部区域位置，312 个二值属性，1 个标注框以及语义分割。

Stanford Dogs^[11]：来自斯坦福大学，数据记得内容是狗类图像，包含 120 种不同类别，共 20580 张图像数据，但只提供标注框这一个人工标注信息。

Oxford Flowers^[12]：来自牛津大学，数据集内容是花类图像，包含两个库分别是 17 类别和 102 类别。每个类包含 40 至 258 张，共 8189 张图像。只提供语义分割图像，不包含其他额外的标注信息。

下图 2-2 为这三种图像数据库的部分展示。但是 CUB-200-2011 数据库是目前细粒度图像分类领域最经典的，也是最常用的数据库。所以，本文将使用此数据集进行实验。在表 2-1 中列举了部分数据集。



图 2-2 三种数据集样例图片

表 2-1 数据集总表。其中 BBox 指标注框信息 (Bounding Box), Parts 指局部区域信息, Attributes 指属性标注信息。

名称	类别与图像数量	标注信息
CUB-200-2011	200 类 11788 张	BBox Parts Attributes 语义分割图像
Stanford Dogs	120 类 20580 张	BBox
Oxford Flowers	17/102 类 8189 张	语义分割图像
Cars	196 类 16185 张	BBox
FGVC-Aircraft	120 类 10200 张	BBox

3 细粒度图像分类方法

3.1 图像特征提取

细粒度图像分类中决定图像分类准确性的一个关键因素是要找到一个区分度较大的特征，这个特征一直是研究者所追寻的目标。当然，随着深度学习尤其是深度卷积神经网络的不断发展，从深度卷积神经网络中得到的特征比传统方法效果要好很多。但本文重点在于介绍传统方法，从基础出发对细粒度图像分类进行研究。接下来将介绍三种常用的特征提取方法。

3.1.1 尺度不变特征变换

在现实世界中进行物体识别会受到很多其他外在因素的干扰，这就要求局部图像特征要避免噪声影响。这些特征至少要对光照、三维投影变换以及普通物体变化具有一定的不变性。故尺度不变特征变换（Scale-Invariant Feature Transform, SIFT）^[8] 就可以得到这样的特征。它是由 British Columbia 大学大卫·劳伊教授总结了现有的基于不变量技术的特征检测方法，于 1999 年提出的，这种算法在 2004 年被进一步完善。

SIFT 提取特征的方法是在空间尺度中寻找局部极值点，也被称为关键点，它的特征是尺度方向和大小。把图像转换成一个包含很多局部特征向量的集合，这些特征向量具有位置、尺度、旋转不变性，并且对光照变化、三维投影或仿射也具有一定的不变性。

整个算法分五个部分，首先构建尺度空间。高斯卷积核^[13] 是可以实现尺度变换的唯一线性核，定义一副二维图像的尺度空间的方法见下面的公式：

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (3-1)$$

$$G(x_i, y_i, \sigma) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{(x - x_i)^2 + (y - y_i)^2}{2\sigma^2}\right) \quad (3-2)$$

其中 $L(x,y,\sigma)$ 为图像的尺度空间， $I(x,y)$ 为原始图像， $G(x,y,\sigma)$ 可变尺度的 2 维高斯函数。

利用高斯差分尺度空间，进行图像金字塔建立，构建过程分为两步，一个是对图像做高斯平滑，第二步是对图像进行降采样。得到其在不同尺度的图像，如

图 3-1所示。

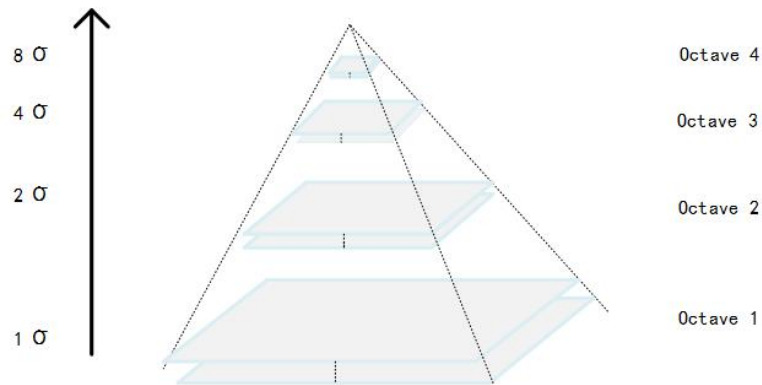


图 3-1 高斯金字塔

之后，检测尺度空间极值点。尺度规范化的 LoG 具有尺度不变性，是由高斯函数梯度算子 GoG 构建的尺度规范化的 GoG 算子。LoG 算子与高斯核函数的差有一定的直接联系，故提出了一种新的算子 DoG(Difference of Gaussians), 叫做高斯差分算子。关键点就是由 DoG 空间的局部极值点组成，它可以有效的在尺度空间检测到稳定的关键点。探寻尺度空间的极值点，每个采样点要和它相邻点进行对比, 找到最小和最大的点。如图 3-2所示，位于中间的点与它所处同一平面的 8 个点进行比较，并且也与上下两个平面的 26 个点进行比较。去除不好的特征点，要排除边缘响应。

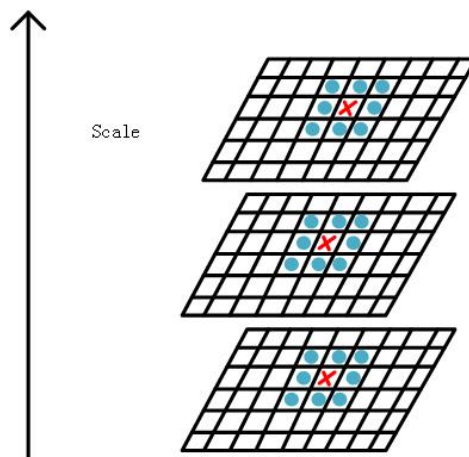


图 3-2 寻找极值点

第四步，利用关键点相邻像素的梯度方向分布特征为每个关键点确定方向参

数，使算子具有旋转不变性。像素点的梯度表示：

$$\text{grad}I(x, y) = \left(\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y} \right) \quad (3-3)$$

梯度幅值：

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \quad (3-4)$$

梯度方向：

$$\theta(x, y) = \tan^{-1} \left[\frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)} \right] \quad (3-5)$$

最后，关键点描述子的生成。将坐标轴旋转为关键点方向，以保证旋转不变性，对特征点产生 128 个数据，也就是 128 维的 SIFT 特征向量，再进行归一化。

3.1.2 方向梯度直方图

方向梯度直方图（Histograms of Oriented Gradients, HOG）特征^[9]是应用在计算机视觉和图像处理领域，用于目标检测的特征描述器。HOG 结合支持向量机（Support Vector Machine, SVM）^[10]进行行人检测的方法是法国研究人员 Dalal 在 2005 的 CVPR 上提出的，现在虽然有很多其他的算法被提出，但基本都是从 HOG 结合 SVM 的思想出发。

这种方法同上面介绍的 SIFT 特征有很多相似之处，此方法使用了图像的本身的梯度方向特征。不同点是这种方法是先将图像分成小的连通区域，被称为细胞单元，并采集计算各像素点的梯度或边缘的方向直方图，而且为了提高性能，还采用了重叠的局部对比度归一化^[14]技术。如图 3-3 所示为 HOG 特征的流程图。

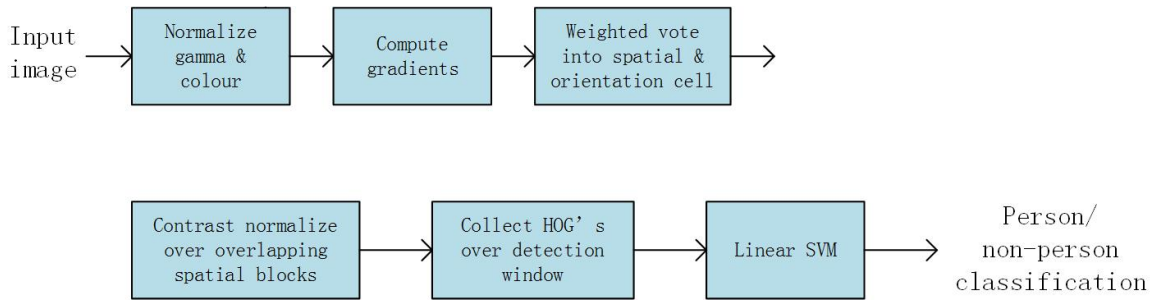


图 3-3 HOG 流程框架

首先将图像转化为灰度图，之后将图像进行归一化，图像归一化主要目的是提高检测器对光照的鲁棒性，由于现实情况的目标可能出现在各种不同的场合，所以为了得到好的效果检测器必须对光照不太敏感才可以。Gamma 压缩公式：

$$I(x, y) = I(x, y)^\gamma \quad (3-6)$$

其中 γ 可取 $1/2$ 。

之后，计算图像梯度是利用一阶微分进行处理。之后采用很多一阶微分模板来求梯度的近似值，但在结果表明模板 $[-1, 0, 1]$ 效果是最出众的。然后把整个图像分割为一个个的 Cell 单元，再对每个单元格构建梯度方向直方图。把相邻近的 Cell 单元组合成一个块，每一个单元格的特征会以不同的结果多次出现在最后的特征向量中，将归一化之后的块描述符就称之为 HOG 描述。最后将检测窗口中所有重叠的块进行 HOG 特征的收集，并将它们结合成最终的特征向量供分类使用。

HOG 的特点是所分的单元较小，在局部细胞单元上进行操作，所以可以保留一定的空间分辨率。并且这种归一化的操作可以使该特征对比度变化不敏感。

3.1.3 局部二值模式

局部二值模式 (Local Binary Pattern, LBP)^[15] 也是一种可以用来图像提取特征的方法，它主要是用来描述图像局的纹理特征。同样这种方法也具有旋转不变性和灰度不变性等明显的优点。最早提出此方法的是 T. Ojala, M. Pietikäinen 和 D. Harwood。

最早的 LBP 算子被定义在一个 3 乘 3 的区域中^[16]，将区域中心点像素的灰度值设置为阈值，与周围其他的 8 个点的灰度值比较，如果周围的值大于中心点的值，将次像素点的位置标记成 1，否则记为 0。如此便得到了一个 8 位二进制数，这也就是中心点像素的 LBP 值。这个 LBP 值就可以用来反映此窗口区域的纹理信息。如下图 3-4 所示

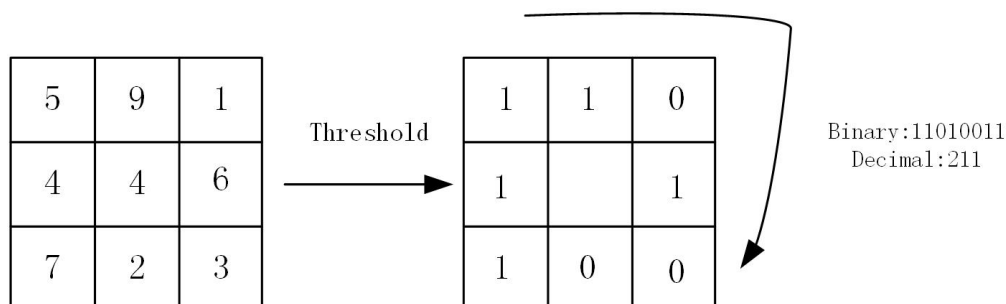


图 3-4 LBP 被定义在 3×3 窗口

故 LBP 操作的定义形式为:

$$LBP(x_c, y_c) = \sum_{p=0}^{p-1} 2^p s(i_p - i_c) \quad (3-7)$$

其中, (x_c, y_c) 是中心像素, i_c 代表亮度, i_p 代表相邻像素的亮度。s 是一个符号函数:

$$s(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{else} \end{cases} \quad (3-8)$$

按照这样的方法就可以获得很好的图像细节。但是随着技术的不断发展, 仅仅以 3×3 为窗口无法满足不同尺寸和频率纹理的需要, 所以为了适应不同的情况达到更好的效果, 研究人员不断对其提出了各种改进和优化。其中有圆形 LBP 算子、LBP 旋转不变模式和 LBP 等价模式等, 具体就不一一介绍了。

得到了每一点的 LBP 值后, 原始 LBP 特征依然是“一幅图片”。然后计算每个小区域的直方图, 就是每个数字出现的频率, 然后对该直方图进行归一化处理。最后将获得的每个小区域的统计直方图进行连接成为一个特征向量, 就得到了整幅图的 LBP 纹理特征向量。之后可以利用支持向量机 (Support Vector Machine, SVM) 或者其他机器学习算法进行分类。

3.2 词袋模型

最初词袋 (Bag-of-Words, BoW) 模型^[3] 是用来进行语义文本检索。此算法忽略掉文本的语法和语序等要素, 只是将文本视为一个个单独的词汇集合。近年来, BoW 模型被广泛运用到计算机视觉领域。所以, 充分应用文本检索算法的优势, 将图像视为一系列视觉单词的集合, 这些单词被定义为一个图像块的特征向量, 而 BoW 模型也就相当于图片中所有图像块的特征向量获得的直方图。所以用这种视觉单词构成的直方图来描述图像的特征, 使图像分类更加有效。BoW 算法步骤是, 首先提取一幅图片中的特征区域或关键点, 提取方法可以用到上面介绍的几种特征提取方法。然后一幅图像由若干个视觉单词组成, 构建视觉词典。如图 3-5 所示。

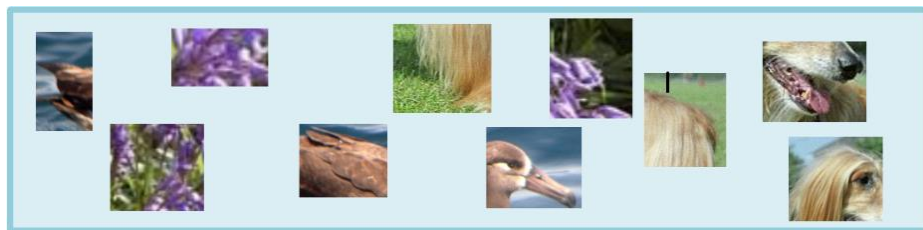


图 3-5 从图像中提取视觉词汇

将提取到的特征进行聚类，通常选用 **K-Means** 方法实现聚类^[17]。视觉词汇的聚类过程是指把很多描述符转化为一个个视觉词汇的过程。这种算法以 **K** 为参数，把 **N** 个对象分为 **K** 个簇，使簇内相似度比较高，而簇间相似度较低。对于一幅图像中所提取到的特征点，想要统计各个特征点属于哪些视觉单词，需要对图像中的特征点根据视觉词汇表的信息，再依据描述符与视觉词汇的距离远近进行归类，最终得出描述这幅图像的视觉单词并确定出物体的类别。

3.3 支持向量机

在机器学习中，支持向量机 (Support Vector Machine, SVM)^[10] 是有关联学习算法的监督学习模型，并利用关联学习算法分析数据进行分类和回归分析。给定一组训练的样本，每个都标注属于一个或两个类别。SVM 训练算法建立了一个模型，这个模型为一个类别或其他类别分配新的样本，故 SVM 成为非概率的二元线性分类器。除了执行线性分类外，支持向量机可以应用所谓的核函数有效地进行非线性分类，隐式将其输入映射到高维特征空间。SVM 分类器的学习需要构建一个从简单到复杂的模型：线性可分支持向量机、线性支持向量机和非线性支持向量机。

3.3.1 支持向量机的原理

首先在 n 维空间中找到一个超平面，可以将空间中的点进行分类^[18]。通常情况下，某个点离超平面的距离远或者近可以视为分类预测的确信或准确程度。SVM 分类器的目的就是要将这个间隔距离最大化。在虚线上的点被称做支持向量。如图 3-6 所示。

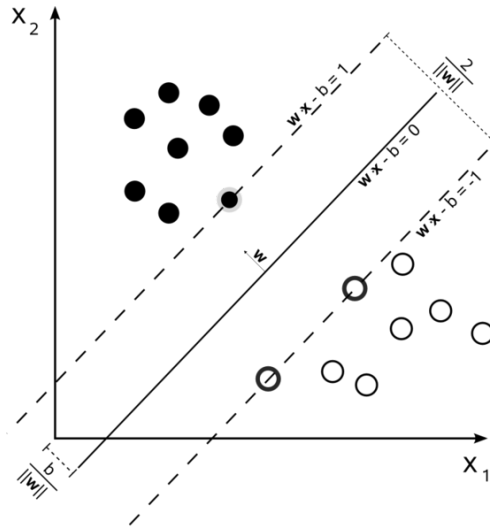


图 3-6 最优超平面^[10]

但在实际研究中，在二维平面中会常常遇到线性不可分的情况，面对这种问题通常采取的做法是把样例的特征映射到高维空间中，在高维空间就可以找到一个平面使其分类。如图 3-7所示。

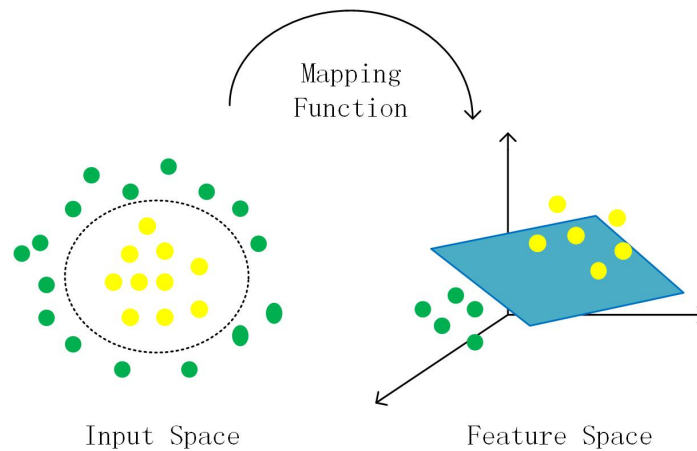


图 3-7 高维最优超平面

然而线性不可分的情况映射到高维空间时，可能会导致维度达到极高的程度，这种情况会使计算复杂。这时就需要用到核函数，它的作用尽管是可以将特征进行从低维到高维的转换，但是核函数的优势是它先在低维空间进行计算，再转换到高维空间进行分类，这就回避了直接在高维空间中进行冗长的计算的问题。具体的核函数内容下文将介绍。最后，应用松弛变量进行数据噪音的处理。

SVM 分类器的优势是它可以解释为凸优化问题，所以可以使用已知的高效算法发现目标函数的全局最小值。而其他分类方法都采用一种基于贪心学习的策略

来搜索假设空间，这种方法一般只能获得局部最优解。

3.3.2 核函数

当面对线性不可分的情况，SVM 处理的方法就用到核函数^[19]。它不用将处于原空间中的样本投射到新的空间中而就可以在原空间中直接计算出内积。所以计算两个向量在隐式映射之后的空间中的内积的函数被叫做核函数。核函数的定义表达式为，原始特征内积是 $\langle x, z \rangle$ ，映射后为 $\langle \phi(x), \phi(y) \rangle$ ，那么定义核函数 (Kernel) 为：

$$K(x, z) = (x^T z)^2 \quad (3-9)$$

这种方法避开了直接在高维空间中进行计算，但是得到的结果却是相同的。以下介绍几种核函数，主要介绍本文将要使用的两种核函数：

线性核函数 (Linear)：主要用于线性可分的情形。参数较少，速度较快，针对普通数据，分类结果比较理想。核的主要目的是解决“映射后空间中的问题”和“映射前空间中的问题”。公式如下：

$$K(x, y) = \langle X, Y \rangle \quad (3-10)$$

径向基核函数 (RBF)：主要用于线性不可分的情形。参数较多，分类效果与参数有很大关系。可以通过训练数据的反复检验来寻找适合的参数，但这是个比较耗费时间的过程。公式如下：

$$K(x, y) = \exp(-\gamma \|x - y\|^2) \quad (3-11)$$

多项式核函数：

$$K(x, y) = ((X, Y) + R)^d \quad (3-12)$$

高斯核函数：

$$K(x, y) = \exp\left\{-\frac{\|X - Y\|^2}{2\sigma}\right\} \quad (3-13)$$

对于应该选择哪种类型的核函数，需要按照具体的实际问题去分析，由于有的数据是线性可分的，有的是线性不可分，所以应该测试不同核和不同参数。

4 实验及分析

4.1 实验设置

通过前部分对基础概念的介绍，本章将开展对实验的分析和讨论。将围绕 SIFT+BoW+SVM、HOG+BoW+SVM、LBP+BoW+SVM 这三种方法展开实验。程序的运行是在 Ubuntu16.04 环境下，需要的数据库及其版本为 OpenCV3.0.1。数据集采用 CUB-200-2011 鸟类数据集。

对于 OpenCV 的安装配置就不在这里介绍了。但对于 CUB-200-2011 数据集这里简单说明一下。由于此数据集比较复杂，其中的局部区域（Part）标注信息以及属性（Attributes）标注信息本实验没有用到，只用到了标注框（Bounding Box, BBox）信息。所以本实验是进行三种方法之间的比较以及有无标注信息的比较。如表 4-3 为 BBox 文件的信息。

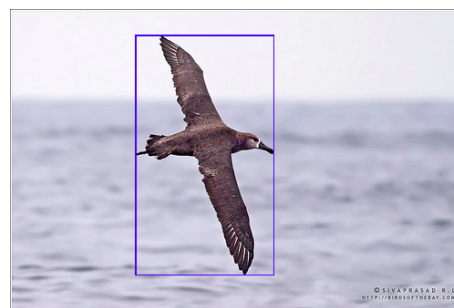
表 4-1 BBox 信息

图片编号	起始点横坐标	起始点纵坐标	标注框宽度	标注框长度
1	60.0	27.0	325.0	304.0
2	139.0	30.0	153.0	264.0
3	14.0	112.0	388.0	186.0

由于数据集没有直接给出加了标注框的图片，所以需要根据自己的实验需要按照文件中已给的位置和框大小的信息给图片进行加框标注。准备两种形式的图片，一种是不加任何标注信息，另一种加标注框信息。如下图 4-1 所示。



(a) 原始图像



(b) 带有标注框的图像

图 4-1 两种图像示例

4.2 三种方法的实验分析

4.2.1 SIFT+BoW+SVM

先进行 SIFT 特征提取，在进行编译时需要注意由于使用了 OpenCV 函数库，在 Ubuntu 系统终端直接进行编译，而不用 Cmake 需要命令行需要加上“pkg-config -cflags -libs opencv”否则无法找到所要调用的函数库。

特征点的提取和计算步骤是先加载输入图像，之后调整输入图像的尺寸，如果输入图像是彩色图片，需要将其转化成灰度图，然后调用 SIFT 库函数 `sift_feature` 来得到目标图片的特征点向量集合和特征点数目。提取到的特征点显示在图像中如下图 4-2 所示。



图 4-2 SIFT 特征提取结果

从图中信息可以看出，经过 SIFT 方法后，图像上所圈出的点就是被提取出的具有旋转不变性的特征点。并且从终端可以输出所提取特征的信息，如表 4-2 所示，表中只列出 4 个方向的描述符。

表 4-2 终端输出的部分结果

关键点数目	每个方向的描述符			
	0	1	2	3
139	0.184178	0.152615	0.000000	0.000000

将提取到的特征存放到 `descriptor` 参数中便于后面的使用。提取到特征后进行聚类过程，经过 Bag-of-Words 模型后得到的结果存在 `Train-Features-SIFT.txt` 中。如表 4-3 所示，可以看到所得到的处理都特征点归一化的数据，从结果中还可以看出每张图片提取的特征点数是不同的。

表 4-3 处理后部分结果

图片编号	特征点数目
1	36
2	43
3	35

接下来就是进行分类器的训练与识别了，用 **SVM** 分类器进行分类。一般利用 **BoW** 的时候，都选择支持向量机，有实验证明 **BoW** 结合 **SVM** 效果要好于其他的分类器，并且我先选择了用 **RBF** 核，得到的结果我是用两种形式体现。一种是直接将每幅图片测试后得到的所属类别用标号与图片一一对应起来如表 4-4所示。另一种是用混淆矩阵体现。

表 4-4 测试的部分结果

图片编号	所属类别	图片编号	所属类别	图片编号	所属类别
1	36	5	133	9	104
2	43	6	48	10	7
3	35	7	7	11	192
4	34	8	104	12	7

这里介绍一下混淆矩阵的概念。在机器学习领域，混淆矩阵 (**Confusion Matrix**) 是一种特定的矩阵，它可以使算法的性能以矩阵的形式直观的表现出来，通常进行监督学习时使用。矩阵的每一列代表预测出的值，在本实验中每一列中的数字代表测试出每类分别有多少张图片，每一行代表的是实际的类别，本实验中每行所有数字之和为每类所含图片的总数，所以矩阵的对角线上的数字就是真正预测准确的值。故在计算准确率时用对角线上的值与总和相比再取平均值。我将得到的矩阵结果用 **Matlab** 进行准确率分析，得到的准确率为 0.4%。由此可知将原始未经过处理的图像直接进行特征提取，得到的准确度结果并不理想非常低。

之后使用 **BBox** 标注框信息，虽然不能只将鸟框选出来，但是也去掉了一些无用的背景信息。还是运用 **Matlab** 将原始图像被框中部分截取出来。结果如图 4-3所示。

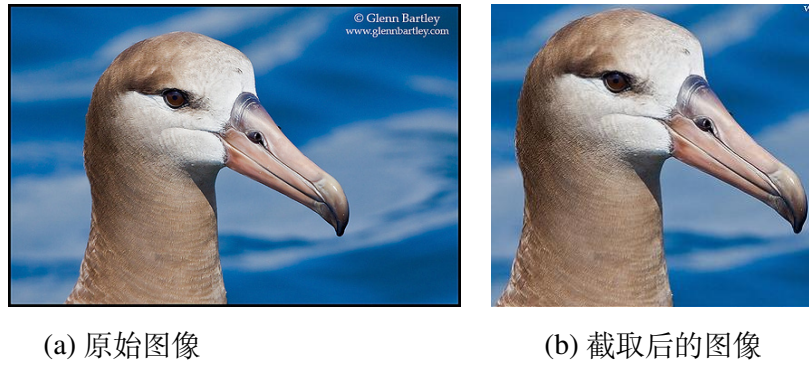


图 4-3 两种图像示例

将截取后的图像输入处理后，得到结果再求准确率，如图所示为 0.56%。结果并没有太大的变化。对于用 SIFT 方法来提取图像特征时，输入原始图像与只加入标注框信息的图像得到的结果相差不大。这说明如果想要用 SIFT 进行特征提取来达到一个更好的效果，需要在图片中加入大量的人工标注信息，因此本实验仅仅加入标注框信息得到的结果非常不理想。当输入原始图像时混淆矩阵中也可以直接观察到测试的结果，图下图 4-4所示。

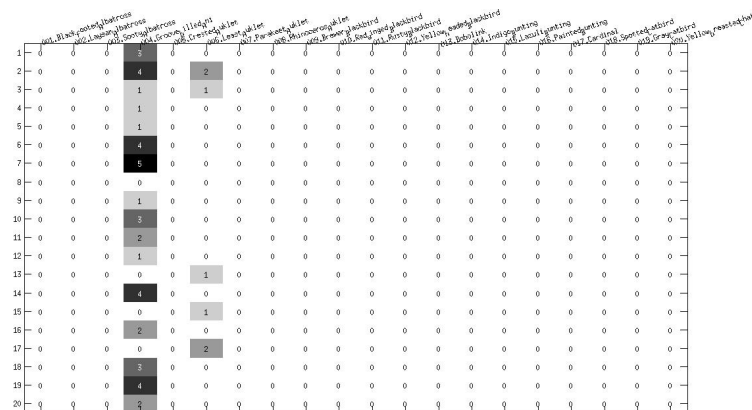


图 4-4 输入原始图像混淆矩阵部分示例

由混淆矩阵可以观察到，测试的结果中所有的类别中大部分图像被预测为第四种鸟类。这种情况说明类对于 SIFT 而言当结合 RBF 核进行分类时，对于第四种鸟类的分类效果更好。但同时也说明，分类结果误差比较大。

之后使用 SVM 的 Linear 核再进行一次分类，当输入未经过处理的图像时得到的准确率为 0.35%，当输入截取后的图片得到的准确率为 0.48%。两种方式得到的准确率如下表 4-5所示。

从结果可以看出对于 SIFT 而言与 RBF 核函数结合进行分类得到的结果要比与 Linear 结合使用少好一些，可能是由于所得到的特征数目少于数据集的图像数

表 4-5 SIFT+BoW+SVM 测试结果

核函数	输入原始图像	输入加标注框的图像
RBF	0.4%	0.56%
Linear	0.35%	0.48%

量。下图 4-5 为输入原始图像用 Linear 后得到的混淆矩阵部分示意图。

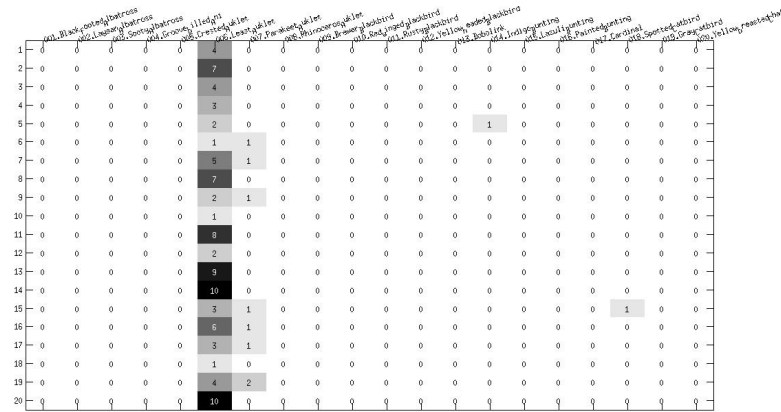


图 4-5 输入原始图像混淆矩阵部分示例

从混淆矩阵中还可以看出，分类结果中将大部分的图像都映射为第 6 类鸟，其中第 14 和 20 类中有 10 张图像的测试结果为第 6 类鸟，误差非常大，测试说明 Linear 更倾向对于第 6 类鸟的分类。但是通过观察对角线可以发现，对角线上的值几乎为 0，这也可以说明当直接输入原始图片进行分类时用 SIFT+BoW+SVM 这种传统方法的分类效果不理想。

4.2.2 HOG+BoW+SVM

用 HOG 特征进行特征提取，其实 HOG 是一个基于梯度的直方图提取算法，大部分用于人体检测十分有效。这里我们用其来提取鸟的特征也是可以的。因为这些参数是用很多图片训练而来的。先使用描述子 `setSVMDescriptor` 给用来对 HOG 特征进行分类的 SVM 分类器的系数赋值，此处的参数为 `HOGDescriptor::getDefaultPeopleDetector()` 表示采用系统默认的值，因为这些值是来自用大量图像训练得到的。并且输入的图像尺度不同，因此要使用多尺度检测，这里是用 HOG 类的方法 `detectMultiScale`。最后对检测出目标矩形框，得到的结果如图 4-6 所示。从图中可以观察到经过 HOG 特征提取，可以将图像的关键部分描述出来。

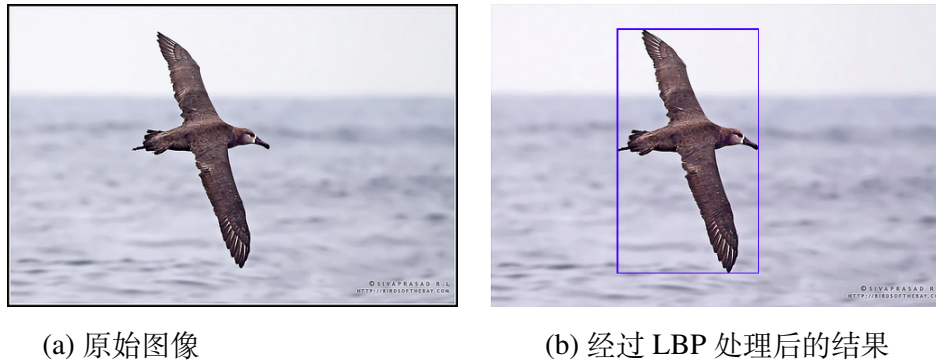


图 4-6 HOG 提取特征样例

经过词袋模型后，进入 SVM 分类器进行分类。未经处理的原始图片用 RBF 核得到的准确率可以达到 3.48%。当输入被标注框框出的图像后，得到的准确率确实有了提高，可以达到 9.13%。对于 HOG 算子来说当输入带有标注信息的图像，使关键的局部区域定位更加准确，提取到的特征点也就更有效。再观察一下当输入原始图像后混淆矩阵的输出情况，如图 4-7 所示。

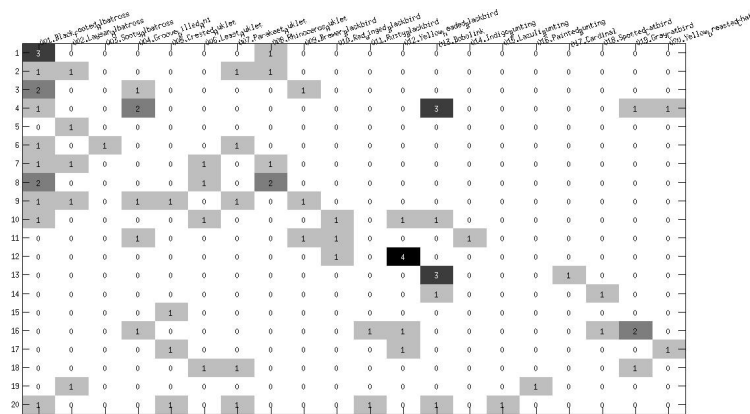


图 4-7 输入原始图像混淆矩阵部分示例

从图中可以发现用这种方式所得到的分类结果分布比较均匀，没有出现像图 4-5 中那样有大部分的值集中在某一个类别上，并且对角线上的值明显增多，也直观的表现出用 HOG 算子来提取特征效果好于 SIFT。并且计算出当输入带有标注框的图像后其准确率有了明显的提高，从混淆矩阵中可以直观的看出区别，如图 4-8 所示。

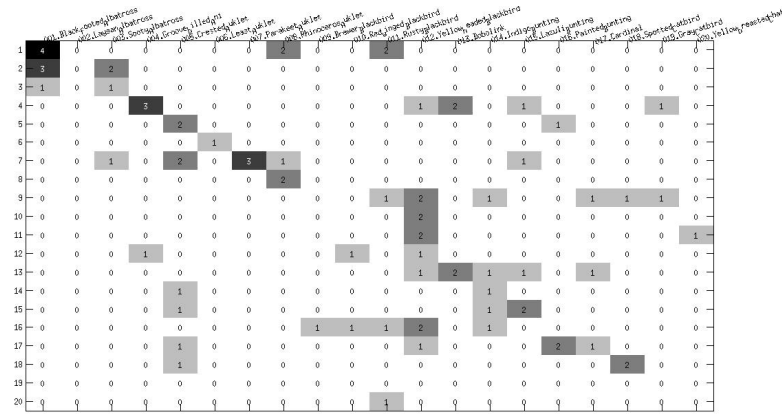


图 4-8 输入带标注图像混淆矩阵部分示例

与图 4-7相比分类准确的图像数量更多了，大部分的值分布在对角线附近，准确率明显提高。对于 HOG 算子来说当输入图像带有标注信息时，它所提取的特征更加准确，也更加可以体现 HOG 算子的优势。

用 Linear 核后，当输入未经过处理的图像和截取后的图像得到准确率分别为 2.05% 和 6.24%。两种方法的结果如下表 4-6所示。

表 4-6 HOG+BoW+SVM 测试结果

核函数	输入原始图像	输入加标注框的图像
RBF	3.48%	9.13%
Linear	2.05%	6.24%

由表中数据对比可以发现，对于 HOG 算子来说与 RBF 核结合使用效果更好。当输入原始图像时用这两种核函数得到的结果并没有太大差别，从混淆矩阵中也看不出很大变化，但是当输入带有标注信息的图像时可以发现，用 RBF 核的准确度更高。与 SIFT 相比无哪种情况用 HOG 提取特征得到的准确率更高，并且实验过程中发现运行速度也更快一些。另一个优点是 HOG 算子对细微动作的变化有很好的包容性，这样保证了由姿态不同所造成的误差。

4.2.3 LBP+BoW+SVM

用圆形的 LBP 算子进行特征提取，此方式是用圆形邻域替代了正方形邻域，改善后的 LBP 算子可以使半径为 R 的圆形邻域内有无数多个像素点。进而得到了半径为 R 的圆形区域包含有 P 个采样点的 LBP 算子。从而得到了半径为 R 的圆形

区域内含有 P 个采样点的 LBP 算子。以一张图片为例经过 LBP 提取得到物体局部纹理特征，如图 4-9 所示。

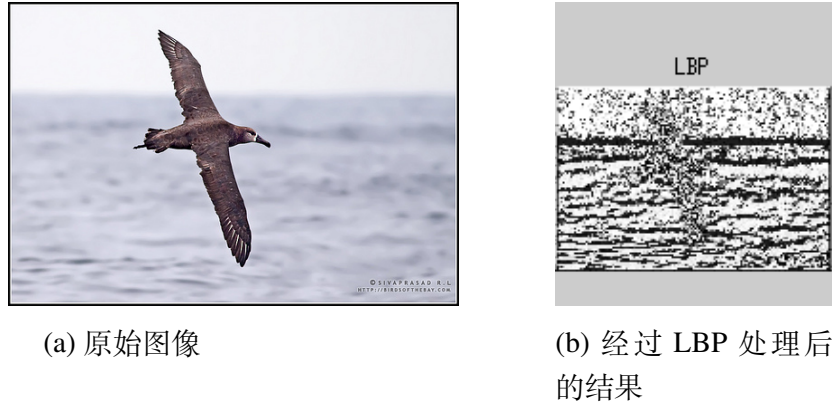


图 4-9 LBP 特征提取样例

由得到的纹理图像信息可以看出，所提取到的特征与其所处的位置有关系。经过 BoW 模型，得到提取的特征，之后输入到 SVM 分类器进行分类。得到用 RBF 核处理后的结果，准确率为 3.6%，用 Linear 核分别进行有无标注信息的实验，结果准确率都为 0.5%。两种种方式得到的结果如下表 4-7 所示。

表 4-7 LBP+BoW+SVM 测试结果

核函数	输入原始图像	输入加标注框的图像
RBF	3.6%	6.9%
Linear	0.5%	0.5%

由表中数据可以发现，当输入原始图像并用 RBF 核时得到的准确率与用 HOG 进行特征提取得到的结果接近。这意味着对于原始图像的处理在使用 RBF 核函数进行分类的情况下，用 HOG 或者 LBP 算子进行特征提取，二者的效果区别不大，但是当输入带有标注信息的图下时，HOG 的优势就表现出来了。由于 LBP 描述的纹理信息有局限性，并且对图像的边缘和方向信息无法无法准确的描述出来，所以目前大多是采用 LBP 与 HOG 算子相结合的方式。当用 Linear 核进行分类时，得到的准确率突然降低，说明对于 LPB 算子不适合用其进行分类。通过观察混淆矩阵可以发现，没有出现预测值集中到某一类别的情况。

在通过混淆矩阵进行观察，当用 RBF 核进行分类输入原始不经过处理的图像，得到的混淆矩阵如图 4-10 所示。

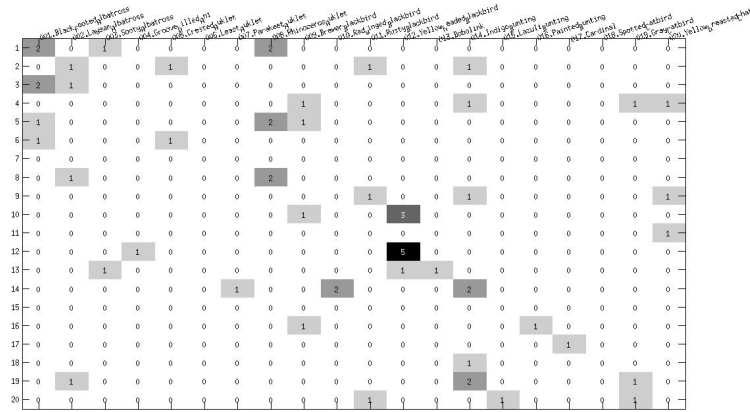


图 4-10 输入原始图像混淆矩阵部分示例

从图中可以观察到，用这种方法得到的结果与用 HOG 算子的结果近似，测试时得到的值比较分散，其中第 12 类预测的准确率最大，但还有很多类别没有预测对，对角线上的值为 0。总的来说当输入原始图像进行分类时，对于传统方法得到的效果是非常糟糕的。当使用 Linear 核进行分类时，准确率急剧下降，与用 SIFT 算子得到的结果相近，同样从混淆矩阵可以看出结果聚集在某一类别上，误差较大。

4.3 实验总结

当输入未经处理的图片时，只用 RBF 核三种方法中准确率最高的方法是 LBP+BoW+SVM 其次是 HOG+BoW+SVM，SIFT+ BoW+SVM 为最低。通过分析和查阅资料，运用 LBP 方法提取特征已经可以达到十分高效的结果，经过改善与发展已经应用于多个领域之中，尤其在是人脸识别、表情识别、行人检测领域已经获得了很大成功。LBP 特征可以一定程度上解决了复杂场景下特征描述问题。在一定程度上对未经过处理的图像提取特征后结合 SVM 进行分类，LBP 占有一定的优势。并且 SIFT 算法是最早提出的，其他两中都是后来提出的，肯定在提取特征的方法上都做了调整和优化。

当输入带有标注框的图像，并且用 RBF 核时三种方法中准确率最高的是 HOG+BoW+SVM，并且也是准确率提升幅度最大的。用 HOG 算子来提取特征在行人检测中得到的结果突出，结合 SVM 分类技术已经复变应用在图像识别与分析领域。由于给出了标注框并且 HOG 所分单元很小，这样提取的特征更加准确。

当用 Linear 核进行分类时，不同算法所体现出的效果不同。当输入的是原始图像时，仍然是用 SIFT 进行特征提取所得到的准确率最低，但此时用 HOG 算子得到的准确率最高。用 LBP 算子得到的准确率与用 SIFT 得到的相近，并且当输入

截图后的图像时准确率并没有提高，说明对于 LBP 算子来说进行分类时更加适合用 RBF 核。总的来说在此问题上使用 RBF 核的效果要高于使用 Linear 核的效果。三种方法所有结果如下表 4-8所示。

表 4-8 三种方法的实验结果

特征提取方式	原始图像		带标注框信息图像	
	Linear 核	RBF 核	Linear 核	RBF 核
SIFT	0.35%	0.4%	0.48%	0.56%
HOG	2.05%	3.48%	6.24%	9.13%
LBP	0.5%	3.6%	0.5%	6.9%

5 总结与展望

5.1 论文总结

近几年，细粒度图像分类问题越来越收到各方面的关注，普通的图像分类问题已经研究到比较成熟的地步。关于细粒度图像分类的研究价值可想而知是非常巨大的，并且还有一定的难度。目前存在的问题，一种情况是针对细粒度图像数据集比较难以采集获取，另一种是现有的很多方法严重依赖人工标注信息，但是这种方法的成本非常巨大，本惬意定程度上制约了算法的实用性。所以最近几年，各种方法都倾向于，仅仅只用类别标签进行分类，而且得到的分类效果也是非常不错的。并且，能否找到具有区分度的局部特征也是限制分类效果的一个重要因素。深度学习的兴起，以及深度卷积神经网络所获得的成功，使人们发现它的优势并用于细粒度图像分类。

本文介绍了细粒度图像分类的一些背景和发展现状，以及基本的理论和基础知识。对细粒度图像的分类流程框架和传统方法基本理论进行说明。对于所选用的三种方法进行试验，得到结果后进行数据分析和对比

5.2 展望

本文只是针对细粒度图像分类的传统方法进行介绍和实验，虽然是比较经典比较基础的方法，但是科学是总在进步的，不断有新的更加高效的算法被提出。本文的实验方法采用了类别标注信息和标注框信息，准确率非常低。近几年的方法基于深度神经网络的细粒度图像分类准确率可以达到 80% 以上，并且不需要人工标注信息，所以未来的研究重点必定是这种方式，并且通过对各种先进算法的比较，未来会找到效果更加的方式。

参考文献

- [1] 中国互联网络中心. 《第 39 次中国互联网络发展状况统计报告》. 传媒, 2017, (3):30–30.
- [2] Welinder P, Branson S, Mita T, et al. Caltech-UCSD Birds 200. Technical Report CNS-TR-2010-001, California Institute of Technology, 2010.
- [3] Harris Z S. Distributional structure. Word, 1954, 10(2-3):146–162.
- [4] Berg T, Belhumeur P N. Poof: Part-based one-vs.-one features for fine-grained categorization, face verification, and attribute estimation. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2013. 955–962.
- [5] Sharif Razavian A, Azizpour H, Sullivan J, et al. Cnn features off-the-shelf: an astounding baseline for recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2014. 806–813.
- [6] Zhang X, Xiong H, Zhou W, et al. Picking deep filter responses for fine-grained image recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016. 1134–1142.
- [7] Zhang H, Xu T, Elhoseiny M, et al. Spda-cnn: Unifying semantic part detection and abstraction for fine-grained recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016. 1143–1152.
- [8] Lowe D G. Object recognition from local scale-invariant features. Proceedings of the IEEE Conference on Computer Vision, 1999. 1150–1157.
- [9] Dalal N, Triggs B. Histograms of oriented gradients for human detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2005. 886–893.
- [10] Cortes C, Vapnik V. Support-vector networks. Machine learning, 1995, 20(3):273–297.
- [11] Khosla A, Jayadevaprakash N, Yao B, et al. Novel dataset for fine-grained image categorization. First Workshop on Fine-Grained Visual Categorization, Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Colorado Springs, CO, 2011.
- [12] Nilsback M E, Zisserman A. Automated flower classification over a large number of classes. Processings of IEEE Conference on Computer Vision, Graphics & Image Proceeding, 2008. 722–729.
- [13] Lindeberg T. Scale-space theory: A basic tool for analyzing structures at different scales. Journal of applied statistics, 1994, 21(1-2):225–270.
- [14] Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems, 2012. 1097–1105.
- [15] Ojala T, Pietikainen M, Harwood D. Performance evaluation of texture measures with classification based on kullback discrimination of distributions. Proceedings of the IEEE Conference on Computer Vision & Image Processing, 1994. 582–585.
- [16] Wang L, He D C. Texture classification using texture spectrum. Pattern Recognition, 1990, 23(8):905–910.

- [17] Wagstaff K, Cardie C, Rogers S, et al. Constrained k-means clustering with background knowledge. International Conference on Machine Learning, 2001. 577–584.
- [18] Boser B E, Guyon I M, Vapnik V N. A training algorithm for optimal margin classifiers. Proceedings of the fifth Annual Workshop on Computational Learning Theory, 1992. 144–152.
- [19] Bergman S. The kernel function and conformal mapping. Number 5, American Mathematical Soc, 1970.

致 谢

本论文是在郑海永老师的细心指导下完成，他那治学严谨，认真负责的态度深深的影响着我。从最初的课题的选择到最后论文的结束，郑老师都一直在耐心的指导我，并且给予我巨大的鼓励和支持，使我在整个课题研究过程中收获颇丰，我会将他所教给我的一切运用到今后的学习生活中。以此谨向郑老师致以我诚挚的谢意和崇高的敬意。

再次我还要感谢王如晨和常琳师姐对我的帮助，当遇到一些困难和阻碍时，师姐们给予我鼓励，让我又重拾信心。当我遇到不懂得地方师姐们常常提出建议并且共同与我讨论，找出解决方案，这个过程使我学到了如何解决问题的思路和方法。

本科四年的学习与生活是我人生中无妨忘却的经历，是我的宝藏。我要感谢大学中与我相处的老师和同学，让我的这四年充实且美好。并且，我要尤其感谢我的父母，他们在我身后默默的支持着我保护着我，是我的后盾。最后，感谢在百忙之中抽出时间审阅本论文的各位老师。