深圳大学考试答题纸

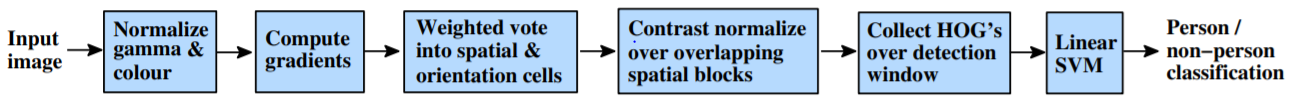
(以论文、报告等形式考核专用)  
二○ 二一 ～二○ 二二 学年度第 一 学期

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 课程编号 | 2801000033 | | 课程名称 | | Python程序设计 | | 主讲教师 | | 何志权 | 评分 |  |
| 学 号 | 2020281024 | | 姓名 | 陈志鸿 | | 专业年级 | | 电子信息工程 2020级 | | | |
|  | | | | | | | | | | | |
| 教师评语： | | | | | | | | | | | |
| 题目： | | **基于HoG 的图像特征提取及其分类研究** | | | | | | | |  | |

**1 介绍**

特征提取是图像处理中的一大领域,著名的提取算法有HoG(Histogram of Oriented Gradient)[1]、LBP(Local Binary Pattern)[2]和Haar-like[3]等等。近些年来,随着GPU算力的急速发展,深度学习也得到了迅速发展,使得图像特征提取的效率大大提升,各种分类任务的正确率不断的刷新提升。而深度学习存在着较差的可解释性和海量数据需求的问题,这是对机器视觉任务来说是伪命题,与之相反的是,传统特征提取方法可视性非常强,且现有的卷积神经网络可能与这些特征提取方法有一定类似性,因为每个滤波权重实际上是一个线性的识别模式,与这些特征提取过程的边界与梯度检测类似。因此对我们的学习来说,传统特征提取方法的学习是不可少的,本次实验是基于HoG 的图像特征提取及其分类研究。

HOG特征是一种在计算机视觉和图像处理中用来进行物体检测的特征描述子。它通过计算和统计图像局部区域的梯度方向直方图来构成特征。最早出现在2005年CVPR上，法国的研究人员Navneet Dalal 和Bill Triggs利用HOG特征+SVM进行行人检测,在当时得到了较好的检测效果。主要流程如下:



本次的实验也是基于上述流程进行一步一步实现,最终实现分类。

**2 方法**

**2.1 数据集准备**

本次分类任务是对猫、狗、人脸、蛇四种类别的分类。收集每种类别各4000张图片,其中87.5%作为训练集,12.5%作为预测集,且每个类别的占比相同。灰度图的形式读取图片,将图片重塑大小至256×256,最后给数据集的车、狗、人脸、蛇分别打上1,2,3,4的标签,以便训练预测以及后续的性能评估。

**2.2 特征提取**

1)采用不同的Gamma值(默认2.2)校正法对输入图像进行颜色空间的标准化(归一化),目的是调节图像的对比度,降低图像局部的阴影和光照变化所造成的影响,同时可以抑制噪音的干扰。2)计算图像每个像素的梯度(包括大小和方向),主要是为了捕获轮廓信息,同时进一步弱化光照的干扰。本实验使用Sobel算子进行计算。3)将图像划分成小cells,因考虑到设备的性能及维度过大问题,而本实验一个cell的大小为32×32个像素。3)统计每个cell的梯度直方图(不同梯度的个数),形成每个cell的描述子。4)对每个cell的梯度方向进行投票,本实验将cell的梯度方向180度分成9个方向块即9个bins,每个bin的范围为20,cell梯度的大小作为权值加到bins里面。5)本实验将每四个cell组成一个block(2×2cell/block),一个block内所有cell的特征描述子串联起来便得到该block的HOG特征描述子。6)将图像image内的所有block的bins串联起来就可以得到该图片的HoG特征向量。这一步中由于是串联求和过程,使得梯度强度的变化范围非常大。这就需要对block的梯度强度做归一化。归一化能够进一步地对光照、阴影和边缘进行压缩。

**2.3 降维**

在提取特征的步骤中,我们可以算出每一张图片提取出的特征向量的维度是2×2×9×7×7等于1764维,这是较大的维数,里面包含了许多冗余信息,所以我们应考虑对特征向量进行降维。本实验实现的是PCA(主成分分析)算法的降维,其主要过程如下:1)对所有样本进行中心化:。2)计算样本的协方差矩阵。3)对协方差矩阵做特征值分解。4)取最大的m个特征值所对应的单位特征向量:。5)得到投影矩阵。

**2.4 分类**

本实验的分类器利用了sklearn库中的svm模型,设置不同的惩罚参数(默认1.0)和核函数(默认’rbf’)的类型进行多次训练和测试。

**3 实验结果及分析**

本实验采用不同的惩罚程度、核函数、gamma校正值、特征维数的分类性能进行对比，采用了精确率、召回率和micro\_F1分数对实验结果进行性能评估。混淆矩阵如图1到图12所示。

**3.1 模型对分类性能的影响**

不同惩罚程度的模型性能对比如表1所示。随着惩罚程度的减少,也即松弛变量减小,训练集的正确率从0.866降低至0.753,这是因为对误分类的惩罚减小,允许容错。按照理论来讲此时模型的泛化能力应增强,但是测试集精确率、召回率和micro\_F1分数均从0.557减小至0.539,原因未能找到。此外,本实验还统计了此特征提取方法对应的每一类的分类情况,从表1中我们可以看出在车和人脸两类上召回率分别达到0.730和0.830左右,而狗和蛇两类分类的效果很差,分别只有0.334和0.324左右。

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 惩罚  程度 | 训练集 | 测试集(召回率) | | | | 测试集 | | | |
| 车 | 狗 | 人脸 | 蛇 | | 精确率 | 召回率 | micro\_F1 |
| 1.0 | 0.866 | **0.730** | 0.334 | **0.830** | 0.324 | | 0.557 | 0.557 | 0.557 |
| 0.8 | 0.833 | **0.722** | 0.334 | **0.830** | 0.304 | | 0.548 | 0.548 | 0.548 |
| 0.5 | 0.753 | **0.718** | 0.328 | **0.828** | 0.282 | | 0.539 | 0.539 | 0.539 |

表1 不同惩罚程度对分类性能的影响

不同和函数对模型性能的影响表2所示。在rbf、linear和poly三种核函数中,高斯核函数rbf的性能是最好的有0.557,这是因为训练的维度是14000×404,样本数n远远大于特征维数m,这也导致了linear性能最差;而多项式核函数poly的训练集的精确率有0.977,原因是多项式的阶数很高,网络变得更加复杂,但这也导致了过拟合。对于四种类别之间的分类召回率差异仍未得到很好的改善。

表2 不同核函数对分类性能的影响

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 核函数 | 训练集 | 测试集(召回率) | | | | 测试集 | | | |
| 车 | 狗 | 人脸 | 蛇 | | 精确率 | 召回率 | micro\_F1 |
| rbf | 0.866 | **0.730** | 0.334 | **0.830** | 0.324 | | 0.557 | 0.557 | 0.557 |
| linear | 0.639 | **0.660** | 0.316 | **0.782** | 0.296 | | 0.514 | 0.514 | 0.514 |
| poly | 0.977 | **0.646** | 0.292 | **0.750** | 0.467 | | 0.539 | 0.539 | 0.539 |

**3.2 特征对分类性能的影响**

不同gamma值对分类性能的影响如表3所示。随着gamma值从2.2减小至0.5,模型性能由0.557减小至0.530,车、人脸两类分类的召回率由0.730、0.830减小至0.694、0.722,蛇的召回率由0.324升至0.460,狗的召回率先有小幅度上升后减小至0.242。而训练集的精确率由0.866升至0.880,说明模型存在较为明显的过拟合问题。此外,车和人脸的分类召回率依然大幅度超过狗和蛇的分类精确率。

表3 不同gamma值对分类性能的影响

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| gamma | 训练集 | 测试集(召回率) | | | | 测试集 | | | |
| 车 | 狗 | 人脸 | 蛇 | | 精确率 | 召回率 | micro\_F1 |
| 2.2 | 0.866 | **0.730** | 0.334 | **0.830** | 0.324 | | 0.557 | 0.557 | 0.557 |
| 1.0 | 0.878 | **0.726** | 0.364 | **0.766** | 0.366 | | 0.555 | 0.555 | 0.555 |
| 0.5 | 0.880 | **0.694** | 0.242 | **0.722** | 0.460 | | 0.530 | 0.530 | 0.530 |

不同特征维数对分类性能的影响如表4所示。特征维数由100增加至704,训练集精确率由0.779增加至0.879,性能由0.589减小至0.555,四个类别的分类召回率出了狗均减小。原因是随着特征维数的增大,保留了更多的冗余信息,导致模型过拟合。

表4 不同特征维数对分类性能的影响

采用不同的惩罚程度、核函数、gamma校正值、特征维数分类后的混淆矩阵

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 特征  维数 | 训练集 | 测试集 | | | | | | |
| 车 | 狗 | 人脸 | 蛇 | 精确率 | 召回率 | micro\_F1 |
| 100 | 0.799 | **0.758** | 0.352 | **0.848** | 0.392 | 0.589 | 0.589 | 0.589 |
| 404 | 0.866 | **0.730** | 0.334 | **0.830** | 0.324 | 0.557 | 0.557 | 0.557 |
| 708 | 0.879 | **0.730** | 0.354 | **0.830** | 0.306 | 0.555 | 0.555 | 0.555 |

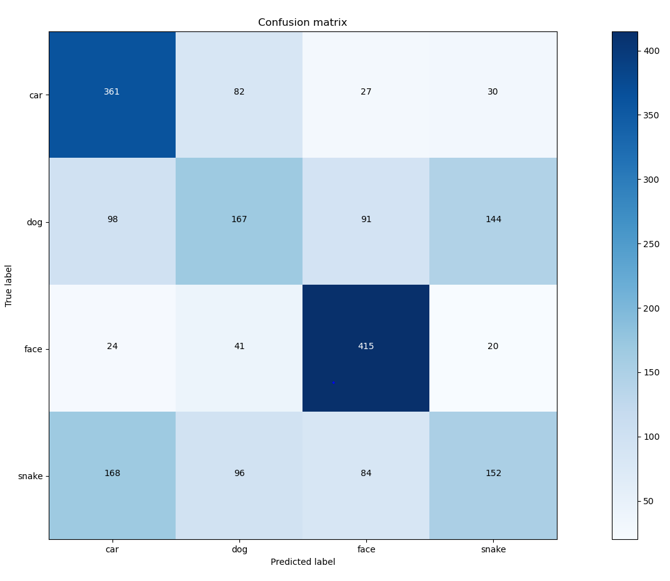
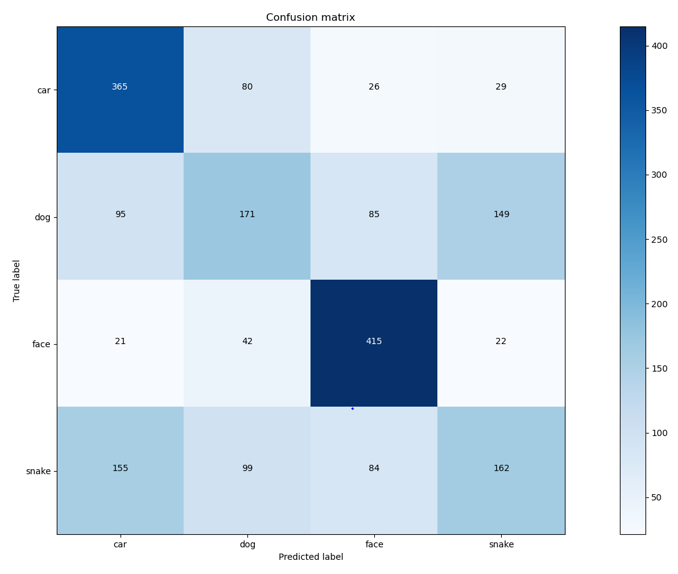


图1 惩罚程度为1.0的混淆矩阵 图2 惩罚程度为1.0的混淆矩阵

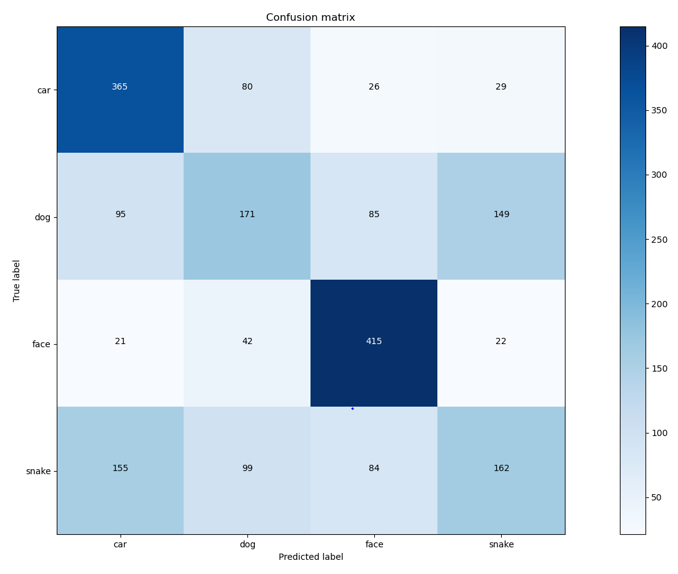
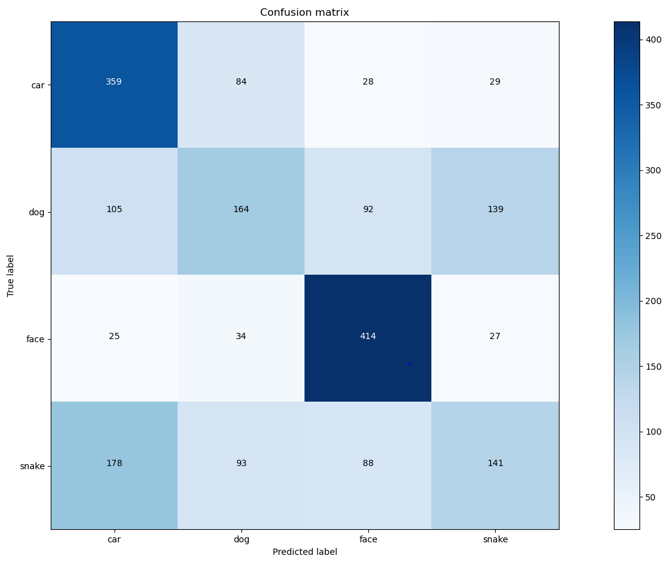


图3 惩罚程度为0.5的混淆矩阵 图4 核函数为rbf的混淆矩阵

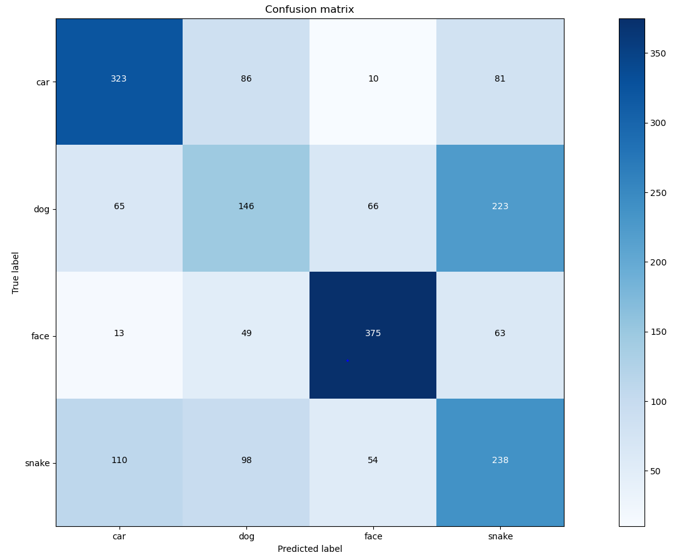
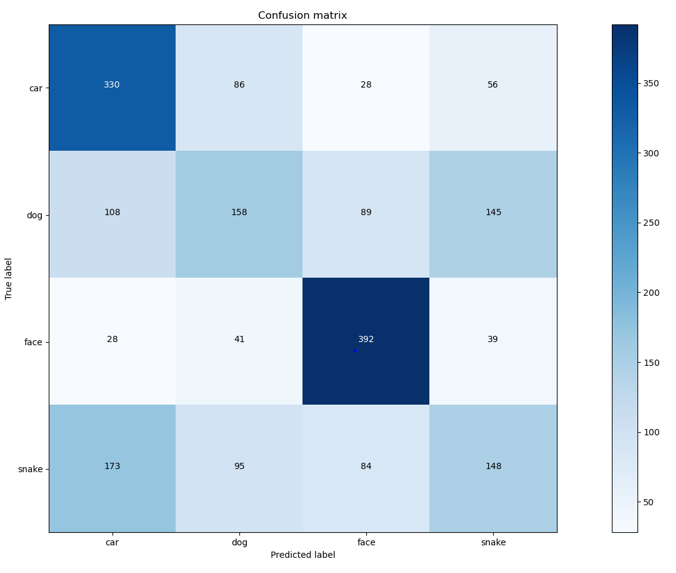


图5 核函数为linear的混淆矩阵 图6 核函数为poly的混淆矩阵

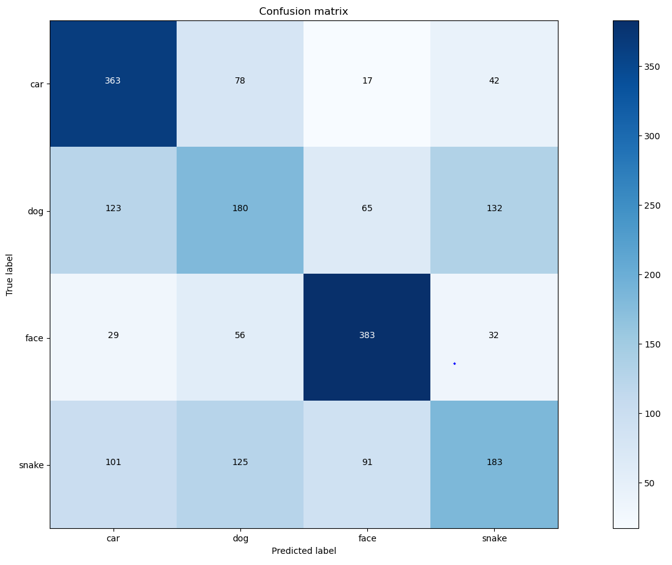
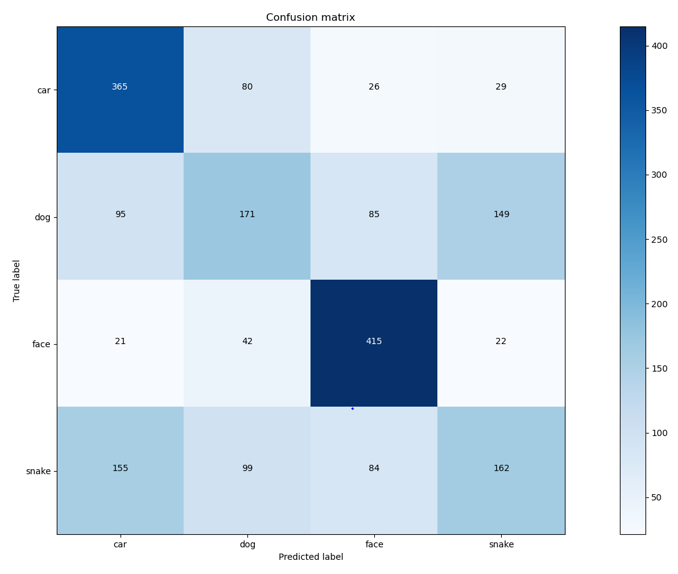


图7 gamma值为2.2的混淆矩阵 图8 gamma值为1.0的混淆矩阵

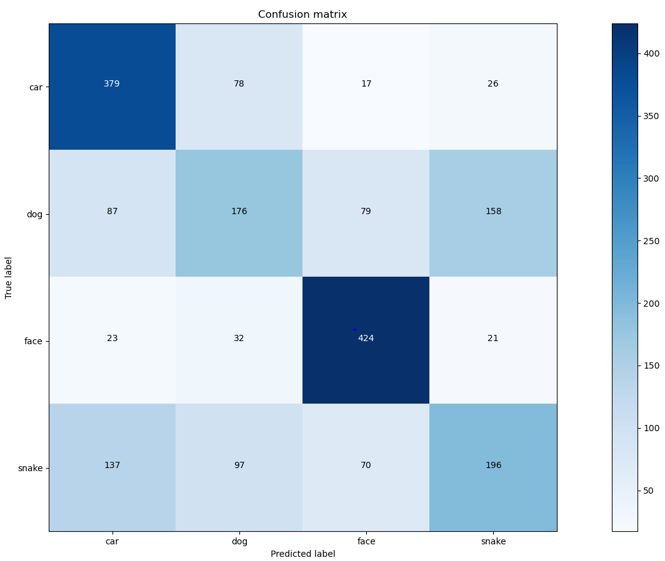
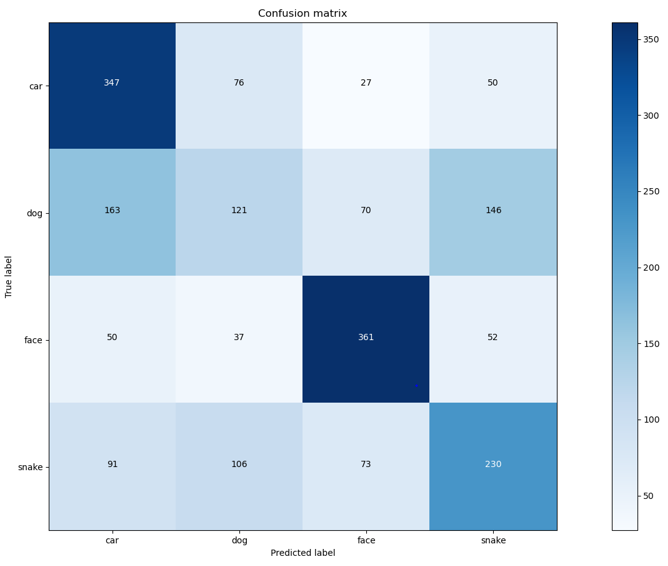


图9 gamma值为0.5的混淆矩阵 图10 特征数量为100的混淆矩阵

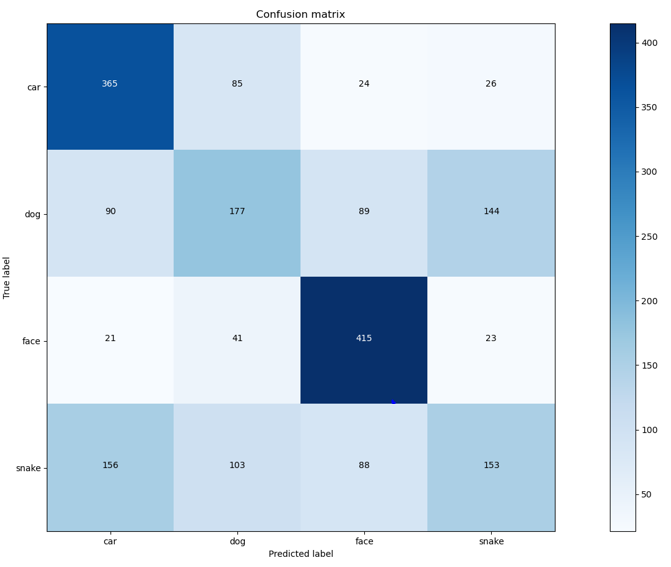
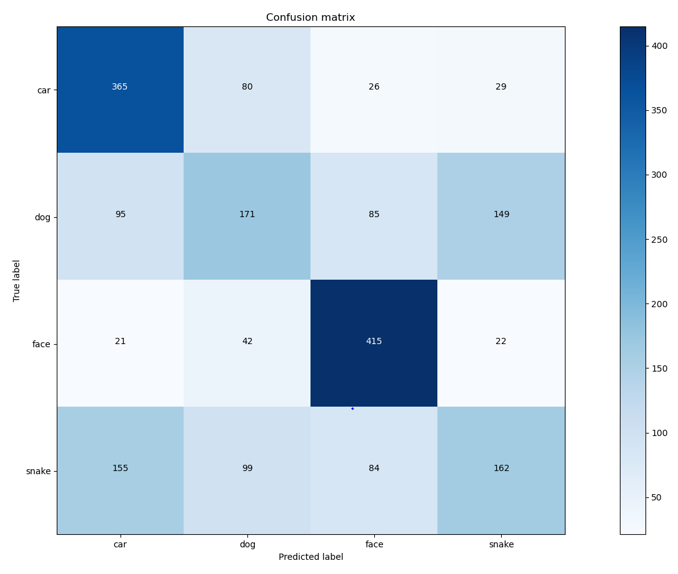


图11 特征数量为404的混淆矩阵 图12 特征数量为704的混淆矩阵

**4 讨论**

本实验通过修改模型参数和特征提取参数来对比分类性能,从表1到表4,无论怎么修改参数,分类性能最高为0.589,始终未超过0.600。因测试集分类精确率较低,因此可能存在着预测失衡的问题,因此使用召回率去衡量每个类别的分类效果会更加合适。将四个类别的分类召回率进行统计,发现车和人脸的分类效果十分好,最高可以分别达到0.758和0.848,HoG特征与SVM的结合之所以对这两类数据分类有较好的结果是因为HoG特征有对光照的不敏感,即使存在部分遮挡也可检测出来、能够反映人体的轮廓，并且它对图像中的人体的亮度和颜色变化不敏感的优点,因此HoG特征适合行人检测和车俩检测等领域。而本次实验对狗和蛇的分类性能很差,最高分别只有0.354和0.467,原因可能是HoG无法从含有狗和蛇的图片中有效提取它们的梯度和梯度方向,是否可以通过对图片的预处理比如裁剪操作来提高精度,还需要后续的实验。

**参考文献**

[1] Dalal N , Triggs B . Histograms of Oriented Gradients for Human Detection[C]// IEEE Computer Society Conference on Computer Vision & Pattern Recognition. IEEE, 2005.

[2] Ojala T , Pietikainen M , Maenpaa T . Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns[C]// IEEE Transactions on Pattern Analysis and Machine Intelligence. IEEE, 2002:971-987.

[3] Papageorgiou C P , Oren M , Poggio T . General framework for object detection[C]// Computer Vision, 1998. Sixth International Conference on. IEEE, 1998.