学号：21122873

**《机器学习基础》**

**课程论文**

**基于梯度直方图的特征提取在MNIST数据集训练与测试数据分布存在差异下的学习优化实践**

|  |  |
| --- | --- |
| **姓 名** | 郑力铖 |
| **学 号** | 21122873 |
| **论文评分** |  |

**2024年 1月13日**

目 录

1、引言............................................................1

2、小样本检测优化算法设计..........................................2

2.1支持向量机 ..................................................2

2.2卷积神经网络.................................................2

2.3梯度直方图...................................................2

3、数据实验 .......................................................3

3.1实验环境与数据...............................................3

3.2支持向量机分类准线结果.......................................3

3.3卷积神经网络分类准线结果.....................................4

3.4 HOG对比实验.................................................4

4、结论与展望......................................................6

基于梯度直方图的特征提取在MNIST数据集训练与测试数据分布存在差异下的学习优化实践

[摘要] 在训练数据相对测试数据较少的情况下，直接使用SVM和CNN方法的训练效果相对较差。因此，本文提出了一种基于梯度直方图的特征提取方法，旨在通过对梯度信息的敏感度分析，增强模型对训练数据的适应性。通过实验验证，我们展示了该方法在MNIST数据集中相较于普通方法的优越性，为解决实际应用中针对训练数据较少的分类问题提供了一种有效的解决途径。本研究对于提升模型泛化性能和小样本学习的实际应用具有一定意义。

[关键字]支持向量机 卷积神经网络 梯度直方图 小样本学习

**Experiment of Differences in Training and Testing Data Distribution with Feature Extraction Based on Gradient Histograms in MNIST Dataset**

**Abstract:** When the training data is relatively small compared to the test data, the training effect of directly using SVM and CNN methods is relatively poor. Therefore, this article proposes a feature extraction method based on gradient histograms, aiming to enhance the model's adaptability to training data by analyzing the sensitivity of gradient information. Through experimental verification, we have demonstrated the superiority of this method over ordinary methods in the MNIST dataset, providing an effective solution for the classification problem with limited training data in practical applications. This study has certain significance for improving model generalization performance and practical applications of small sample learning.

**Key words:** Support Vector Machine; Convolutional Neural Network; Gradient Histogram; Small Sample LearningS

**1、引言**

随着深度学习在图像分类领域取得显著成就，MNIST数据集一直是评估算法性能和模型有效性的重要基准之一。然而，在实际应用中，可能会面临训练数据相对测试数据较少的挑战，这会导致直接使用支持向量机（SVM）和卷积神经网络（CNN）等方法，在训练效果上相对较差的问题。为了解决这一问题，本文在训练前使用梯度直方图的特征提取方法，旨在通过对梯度信息的敏感度分析，增强模型对训练数据的适应性。

本文的核心思想是从图像的梯度信息中提取有意义的特征，以增强模型对训练数据的理解和泛化能力。相较于传统方法，梯度直方图提供了更丰富的信息，能够更好地捕捉图像中的细微变化和结构特征。通过对MNIST数据集的实验验证，我们展示了该方法相对于传统方法的优越性，尤其在训练数据较少的情况下，取得了更好的分类效果。

本研究的贡献不仅在于提出了一种新颖的特征提取方法，更在于通过实验证明了该方法在解决小样本学习问题上的有效性。相信这种基于梯度直方图的特征提取方法将为解决实际应用中针对训练数据较少的分类问题提供一种有效的解决途径。

**2、小样本检测优化算法设计**

**2.1 支持向量机（Support Vector Machine）**

支持向量机（英语：Support Vector Machine，常简称为SVM，又名支持向量网络）是在分类与回归分析中分析数据的监督式学习模型与相关的学习算法。给定一组训练实例，每个训练实例被标记为属于两个类别中的一个或另一个，SVM训练算法创建一个将新的实例分配给两个类别之一的模型，使其成为非概率二元线性分类器。SVM模型是将实例表示为空间中的点，这样映射就使得单独类别的实例被尽可能宽的明显的间隔分开。然后，将新的实例映射到同一空间，并基于它们落在间隔的哪一侧来预测所属类别。

**2.2 卷积神经网络（Convolutional Neural Network）**

卷积神经网络（英语：convolutional neural network，缩写：CNN）是一种前馈神经网络，它的人工神经元可以响应一部分覆盖范围内的周围单元，对于大型图像处理有出色表现。卷积神经网络由一个或多个卷积层和顶端的全连通层（对应经典的神经网络）组成，同时也包括关联权重和池化层（pooling layer）。这一结构使得卷积神经网络能够利用输入数据的二维结构。与其他深度学习结构相比，卷积神经网络在图像和语音识别方面能够给出更好的结果。

**2.3 梯度直方图（Histogram of Oriented Gradient）**

方向梯度直方图（英语：Histogram of oriented gradient，简称HOG）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。这项技术是用来计算局部图像梯度的方向讯息的统计值。在HOG特征描述符中，梯度方向的分布，也就是梯度方向的直方图被视作特征。图像的梯度(x和y导数)非常有用，因为边缘和拐角(强度突变的区域)周围的梯度幅度很大，并且边缘和拐角比平坦区域包含更多关于物体形状的信息。方向梯度直方图(HOG)特征描述符常和线性支持向量机(SVM)配合使用，用于训练高精度的目标分类器。

**3、数据实验**

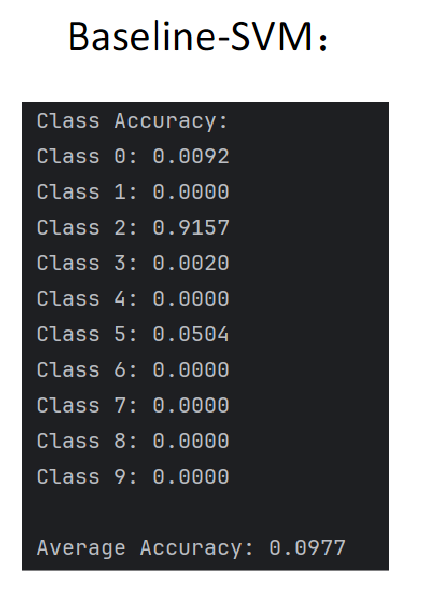
**3.1实验环境与数据**

环境：Python3.10, torch==1.13.1+cu117, numpy, opencv-python, scikit-learn.

训练集使用删减过后的MNIST 手写数字数据集，标注文件使用less\_train\_labs.txt。训练集中包含 0~9 十个类别各 200 张图片。测试集使 用color mnist的测试集，标注文件使用testset\_label.txt

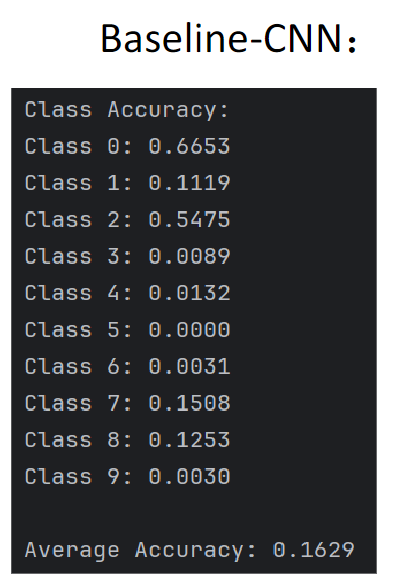
**3.2 支持向量机分类准线结果**

左图为作业要求中的Baseline-SVM，右图为自实现的未优化的SVM。

**3.3 卷积神经网络分类准线结果**

左图为作业要求中的Baseline-CNN，右图为自实现的未优化的CNN。



**3.4 对比实验**

**3.4.1 HOG参数调优**

直接对原数据计算HOG特征，实例化一个cv2.HOGDescriptor类，该类有四个常用参数：

winSize: 窗口大小，一般即设置成图片大小（本文中为(28,28)）

blockSize: 块大小，即HOG提取特征的分块。考虑到本文数据的特殊性，数字不需要特别细致的分块，且图片较小。即将blockSize固定在(14,14)。这样固定将图片分成四块，计算“田”字的四个块内的HOG特征。这也符合数字的特征，即最主要的信息集中在图片中心。

blockStride: 块步长。必须是blockSize的因数。需按照实际情况测试。

cellSize: 单元大小。必须是blockSize的因数。需按照实际情况测试。

因此下表即为针对blockStride和cellSize进行了对比实验。表中第一行为cellSize，第一列是blockStride。均为（2,2）处因为特征维数过高，导致svm复杂度较高而死机。

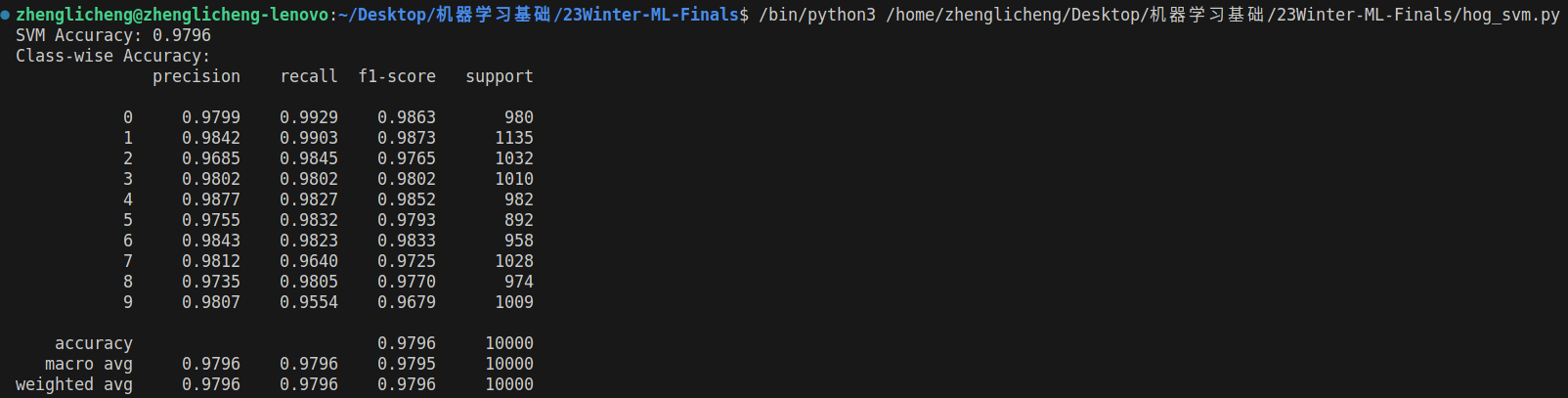
|  |  |  |  |
| --- | --- | --- | --- |
|  | （2,2） | （7,7） | （14,14） |
| （2,2） | **/** | **0.9777** | 0.9634 |
| （7,7） | 0.9748 | 0.9772 | 0.959 |
| （14,14） | 0.9595 | 0.9587 | 0.8942 |

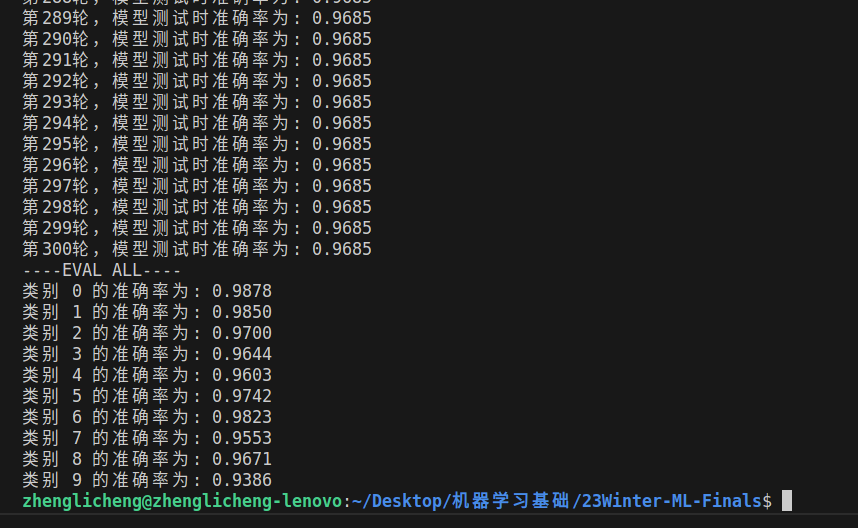
**3.4.2 HOG参数的进一步思考**

因为blockStride和cellSize都需要是blockSize的因数。而MNIST数据集的图片都是（28,28）的。而28的因数只有2，7和14。可变性较小。因此想实验一下，如果把图片通过cv2.resize到（32,32）之后，这样可以取2,4,8,16。最后测试后的选取的最佳参数是blockStride和cellSize皆为（4,4），此时的准确率是0.9793。

同样的，在28周围的24也是拥有丰富因数的一个数。原图cv2.resize到（24,24）之后因数便可取2,3,4,6,12。下表即为针对blockStride和cellSize进行了对比实验。表中第一行为cellSize，第一列是blockStride。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | （2,2） | （3,3） | （4,4） | （6,6） |
| （2,2） | **/** | 0.9781 | **0.9796** | 0.9787 |
| （3,3） | 0.9764 | 0.9780 | 0.9795 | 0.9789 |
| （4,4） | 0.9761 | 0.9790 | 0.9784 | 0.9786 |
| （6,6） | 0.9753 | 0.9768 | 0.9765 | 0.9777 |

最终实验结果是当blockStride=（2,2）且cellSize=（4,4）时，测试准确率最高，为97.96%。较baseline提升6%。

将该参数带入CNN，准确率为96.85%。较baseline提升2.1%

**4、结论与展望**

本文通过提出一种基于梯度直方图的特征提取方法，旨在弥补在训练数据相对较少的情况下，传统支持向量机（SVM）和卷积神经网络（CNN）等方法训练效果相对较差的问题。通过实验证明，该方法在MNIST数据集上相较于传统方法表现出更好的性能，HOG+SVM达到了接近98%的准确率，HOG+CNN达到了96.85%的准确率。不过在其他数据集上的泛化性能尚需进一步考察。未来的研究也可以进一步深入挖掘HOG的理论基础。期待将该方法推广到更多的领域，为解决实际应用中的小样本分类问题提供更全面和有效的解决方案。

**参考文献**

[1] Cortes, C., & Vapnik, V. (1995). "Support-vector networks. " Machine Learning, 20(3), 273-297.

[2] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, CA, USA, 2005, pp. 886-893 vol. 1, doi: 10.1109/CVPR.2005.177.

**附录**

代码仓库：<https://github.com/SHUzhekiNg/23Winter-ML-Finals>