

基于空间金字塔优化间距准则算法的小样本图像分类

张仕霞, 杨晓欢, 蒋新

(西南财经大学 天府学院, 四川 成都 610000)

摘要: 直接采用原始图像进行降维提取特征具有一定的局限性, 图像由于拍照远近、取景等原因会导致目标区域与背景区域大小差异, 同时目标区域的重要性远高于背景区域, 为此, 提出金字塔优化算法, 即通过图像分割后再进行卷积池化的方法提取特征, 然后采用优化的间距准则算法进行图像分类。在 ORL 和 Yale 以及加噪数据集上进行实验, 实验结果表明: 所提算法的鲁棒性和稳定性更好。

关键词: 金字塔; 特征提取; 图像分类; 间距准则

中图分类号: TP391.41 文献标识码: A 文章编号: 1003-7241(2025)03-0156-05

Few-Shot Image Classification Based on Spatial Pyramid Optimization Spacing Criterion Algorithm

ZHANG Shi-xia, YANG Xiao-huan, JIANG Xin

(Tianfu College of SWUFE, Chengdu 610000 China)

Abstract: It has certain limitations to directly use the original image for dimensionality reduction and feature extraction. The image leads to the size difference between the target area and the background area due to the distance of photographing, viewfinder and other reasons. At the same time, the importance of the target area is much higher than the background area. In order to solve this problem, this paper proposes a pyramid optimization algorithm, which extracts image features from multiple scales, and then uses the optimized spacing criterion algorithm for image classification. Experiments on ORL, Yale and noisy data sets show that the improved algorithm has better robustness and stability.

Keywords: Pyramid; feature extraction; image classification; spacing criterion

0 引言

近年来, 随着图像分类领域的飞速发展, 深度学习算法进行图像分类的研究越来越广泛^[1], 比如卷积神经网络在图像识别中的应用, 但是深度学习需要依赖大量数据, 数据量较少或标签缺失的问题抑制了深度学习在图像分类中的应用, 针对样本量较少或标签缺失的数据研究就显得尤为重要^[2], 目前有许多学者在对此进行研究, 如动态子空间算法^[3]以及空间金字塔降维等^[4]算法。

间距准则算法在小样本分类中已经有着较为广泛的应用^[5], 其核心是使得同类样本有着较高的相似度, 而不同类样本尽可能相似度越低越好, 比如最大间距准则, 为了解决经典算法的泛化能力提出了 L1 范数, 为了寻找最优参数利用粒子群算法进行相应的改进^[6]等, 针对原有的间距准则算法提出了许多改进算法^[7]。

本文主要针对图像的目标和背景区域大小的不一致导致的特征局限性问题, 提出了空间金字塔算法更好地

获取图像的特征, 并且和优化的间距准则算法结合使用。空间金字塔算法首先使用原始图像进行分割, 然后使用分割后的图像进行卷积和池化的方法降维到同一维度, 最后将得到的特征作为优化间距准则算法的数据输入进行分类训练。

1 相关理论

1.1 空间金字塔采样



(a) 原始图 (b) 2×2 分割图 (c) 4×4 分割图

图 1 空间金字塔采样

词袋算法早期主要应用于文本分类, 后来应用到了图像分类领域取得了较好的效果, 比如降维处理等, SPM 即把金字塔匹配的思想应用到词袋模型, 近年来较多学者对其进行了研究^[8], 本文具体实验使用如图 1(a)–图 1(c),

*基金项目: 全国高等院校计算机基础教育研究会基金项目 (2024-AFCEC-643)

收稿日期: 2023-09-06

其中图(a)是原始图像为 112×92 的灰度人脸图像,图1(b)是经过 2×2 的分割,分割后的图像大小为 56×46 ,其中 B_{ij} 表示第 i 行第 j 列的分割图,图1(c)是在图1(b)的基础上进行 2×2 的分割,分割后的图像大小为 28×23 。

本文获取数据信息的方法为首先对图1中的图(a)进行卷积和池化,采用值全为1的二维卷积和两次 2×2 的均值池化采样方法得到 28×23 的特征矩阵。对图1中的图 B_{ij} 采用卷积和 2×2 的均值池化采样方法得到四个 28×23 的特征,将图(a)采样特征矩阵和图(b)采样取得的 B_{ij} 特征矩阵及图(c)的特征矩阵进行线性组合得到 $[A, B_{11}, B_{12}, B_{21}, B_{22}, C_{11}, C_{12}, C_{13}, C_{14}, C_{21}, C_{22}, C_{23}, C_{24}, \dots, C_{44}]$ 的特征矩阵。

1.2 优化间距准则算法

间距准则算法的目的是使得原始图像降维后满足同类样本相似性尽可能高且异类样本相似性尽可能低,该算法在小样本图像分类中有着较为广泛的应用,目前有着许多的改进算法和实际应用。

经典的2DMMC算法的目标函数如下:

$$\begin{aligned} \max J(x) &= \sum_{i=1}^c n_i \left\| w^T (\bar{A}_i - \bar{A})^T \right\|_2 \\ &\quad - \sum_{j=1}^c \tau_j \sum_{i=1}^{n_i} \left\| w^T (A_j^i - \bar{A}_i)^T \right\|_2 \end{aligned} \quad (1)$$

s.t. $w^T \cdot w = 1$

式中, C 表示总共有多少个类别,对应的 N_i 表示该类别的总数, \bar{A}_i 表示对应类别的均值, \bar{A} 即为所选样本总均值,使得目标函数最大化的目的即求解投影矩阵 W 。

针对L2范数存在异常值问题,提出了L1-2DMMC算法,目标函数如下:

$$\begin{aligned} \max J(x) &= \sum_{i=1}^c n_i \left\| w^T (\bar{A}_i - \bar{A})^T \right\|_1 \\ &\quad - \sum_{j=1}^c \tau_j \sum_{i=1}^{n_i} \left\| w^T (A_j^i - \bar{A}_i)^T \right\|_1 \end{aligned} \quad (2)$$

s.t. $w^T \cdot w = 1$

针对L1范数求解复杂且不一定能得到全局最优的问题,提出了粒子群优化算法与L1-2DMMC结合的方法,此方法不仅简化了L1-2DMMC算法的计算复杂度,同时还更容易得到全局最优解,在文献[6]中得到了较好的验证,具体结合实验步骤如图2所示。

2 空间金字塔优化间距准则算法

本文采用空间金字塔算法,先将原始图像进行金字塔分割,利用卷积和池化的方法得到相应的数据特征,最后使用优化的间距准则算法进行图像分类。

输入:图像

输出:图像特征,分类结果

Step1:采用空间金字塔算法得到 $[A, B_{11}, B_{12}, B_{21}, B_{22}, C_{11}, C_{12}, C_{13}, C_{14}, C_{21}, \dots, C_{44}]$;

Step2:采用优化的间距准则算法进行分类。

为了验证本文算法的有效性,下面将在原始数据集和加噪数据集上实验进行验证。

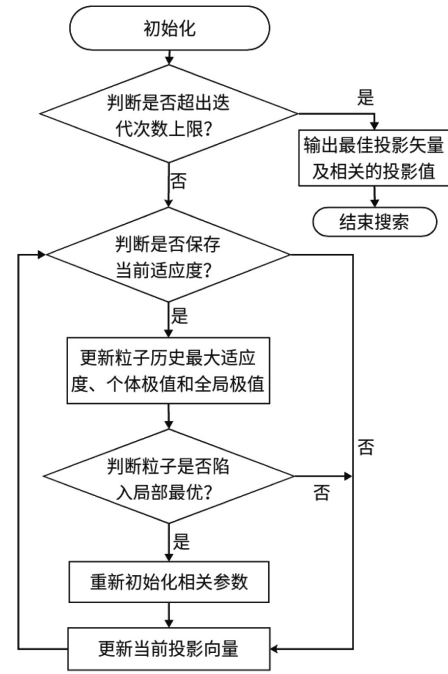


图2 优化间距准则算法

3 实验结果及分析

3.1 数据集介绍

ORL人脸数据库是由40个人,每人10张人脸图像构成的人脸数据库,图像大小宽度为92,高度为112。这些人脸图像之间有着较大差异,比如时间、光照、表情和面部细节等。如图3所示。



图3 ORL数据库图

Yale人脸数据库是由15个人,每人11张人脸图像构成的人脸数据库。每个人的图像包含有不同表情、姿态和光照,共165张图片,图像大小宽度为100,高度为100,如图4所示。



图4 Yale数据库图

3.2 原始数据集实验

下面实验中随机选取不同数量的 K 张图像用于训练,其中 $k=2, 3 \dots 7$, 剩余的图像用于测试算法的有效性,判断算法的好坏。实验结果见表 1。

取不同的样本数 $k=2, 3 \dots 7$, 包含了 2DMMC、L1-2DMMC 以及 PSO-L12DMMC 和本文算法, 实验结果如图 5、图 6 所示。

表 1 数据集上的分类精度

算法 精度/% 训练样本	ORL 数据集						Yale 数据集					
	K=2	K=3	K=4	K=5	K=6	K=7	K=2	K=3	K=4	K=5	K=6	K=7
本文	88.86	88.65	90.98	92.5	97.6	97.98	77.85	78.26	81.52	82.98	93.56	98.88
PSO_MMC[6]	87.81	87.86	90.42	91.5	96.25	97.5	76.67	77.5	80	81.11	92	98.33
L1_2DMMC	85.31	87.5	86.67	90.5	93.75	95.83	75.58	75	78.1	77.78	88	96.67
2DMMC	80.63	80	82.08	85.5	91.88	93.33	70.37	71.67	77.14	77.78	85.33	93.33

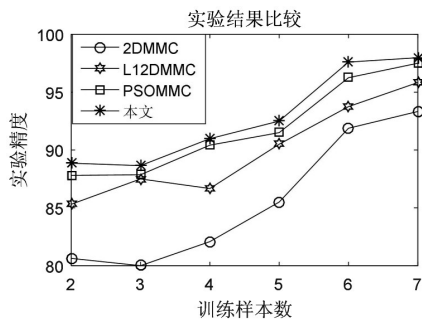


图 5 ORL 数据集上的准确率

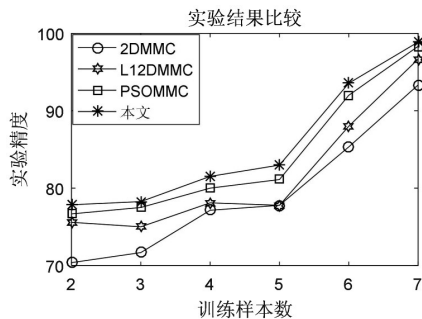


图 6 Yale 数据集上的准确率

通过表 1 和图 5 可以看出本文的算法在实验精度上取得了一定的提高,从 ORL 原始数据集上的实验可以看出,与经典的算法相比分别提高了 8.23、8.65、8.9、7、6.72 个百分点,和粒子群优化算法相比在训练样本数取 6 的时候改进最为明显优化了 1.35 个百分点,通过上面实

验数据分析可以看出本文算法稳定性更好。

通过表 1 和图 6 可以看出在 Yale 数据集上本文算法精度分别提高了 7.48、6.59、4.38、8.23、5.55 个百分点,从实验结果可以看出当训练样本取 7 时本文算法的精度为 98.88%,说明了本文算法有着较好的稳定性。

3.3 噪声数据集实验

下面将进行加噪声实验,在 ORL 数据集上加不同密度的椒盐噪声,如图 7 所示,在该数据集上加不同方差的高斯白噪声,如图 8 所示。



图 7 ORL 椒盐噪声图



图 8 ORL 高斯白噪声图

3.3.1 在 ORL 数据集上的加噪实验

实验中随机选取 4 张图像用于训练样本,剩余的图像用于测试算法的有效性,表 2 是 ORL 数据集加上不同噪声的实验结果。

表 2 ORL 加噪数据集准确率

算法 精度/% 噪声	椒盐噪声/密度							高斯白噪声/方差						
	0.01	0.05	0.1	0.2	0.3	0.4	0.5	0.01	0.05	0.1	0.2	0.3	0.4	0.5
本文	89.89	91.06	90.56	88.65	77.21	62.26	45.63	89.12	92.35	90.15	87.86	88.94	88.56	89.85
PSO_MMC[6]	89.17	90.83	90.42	87.92	75.83	52.5	35.83	88.75	91.25	89.58	86.67	87.92	87.08	89.17
L1_2DMMC	86.25	87.08	86.25	79.58	72.92	47.5	28.75	87.08	89.17	84.17	85.42	84.17	85.42	87.5
2DMMC	83.75	84.17	70	50.83	38.33	16.25	8.75	82.92	83.75	82.08	81.25	76.67	71.25	62.08

从表 2 数据可以看到实验的具体精度,但是不能很好地得到数据趋势和对比效果,为了更好地展示,绘制出图 9、图 10, ORL 数据在不同噪声上的实验精度对比效果图。

从实验结果可以看出图像加上噪声后识别精度都有减小,当噪声越大时本文算法的稳定性更高,从图 9 可以看出当椒盐噪声越来越大时虽然算法精度降低,但是本

文算法的实验精度有明显提高,当椒盐噪声为0.05时本文算法的精度比0.01提高了1.25个百分点,当椒盐噪声为0.1时比0.01提高了0.67个百分点;从高斯白噪声上的实验数据可以看出,噪声为0.05时的准确率比0.01时有明显提高,整体上看本文算法在高斯白噪声上的准确率基本没有明显变化,表明本文算法具有较好的稳定性。

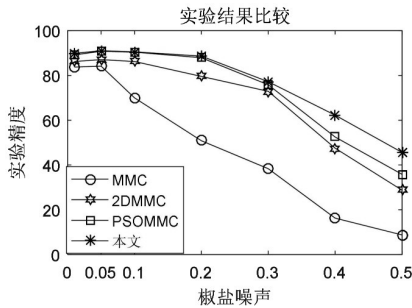


图9 椒盐噪声准确率

3.3.2 在Yale数据集上的加噪实验

同样的方式应用于Yale数据集中,其中Yale数据集上加椒盐噪声,参考图11所示;加高斯白噪声参考图12所示。

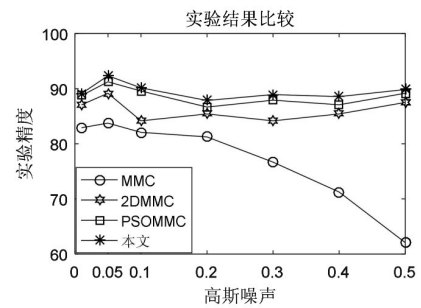


图10 高斯噪声准确率



图11 Yale椒盐噪声图



图12 Yale高斯白噪声图

表3 Yale加噪数据集准确率

算法 精度/% 噪声	椒盐噪声/密度							高斯白噪声/方差						
	0.01	0.05	0.1	0.2	0.3	0.4	0.5	0.01	0.05	0.1	0.2	0.3	0.4	0.5
本文	78.32	81.56	80.56	82.21	81.52	78.56	75.68	82.16	81.25	83.34	81.28	82.56	83.35	82.12
PSO_MMC[6]	77.78	81.11	80	81.11	77.78	74.44	73.33	81.11	78.89	82.22	80	81.88	82.22	81.11
L1_2DMMC	75.56	80	76.67	74.44	73.33	70	56.67	78.89	75.56	81.11	76.67	75.56	75.56	80
2DMMC	73.33	77.78	71.11	44.44	31.11	23.33	14.44	73.33	74.44	75.56	71.11	68.89	65.56	54.44

下面在Yale数据集上进行加噪实验,为了更好验证此处随机选取5张人脸图像用于训练样本,其他用于验证算法的有效性,表3是Yale数据集上加上具体不同噪声后的实验结果。

图13、图14是Yale数据在不同噪声上的实验精度对比效果图。

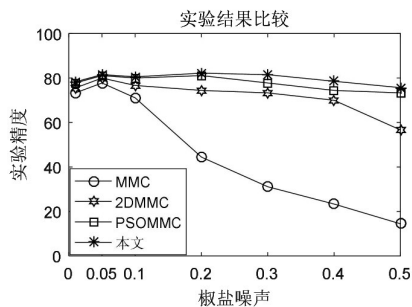


图13 椒盐噪声准确率

从Yale数据集上加噪声的实验结果可以看出,在经典算法上随着噪声的增加可以明显看到识别精度的下降,在本文算法上随着噪声的增加实验结果基本没有明

显变化,比如在椒盐噪声上,并不是噪声越大精度越低,和经典算法相比分别提高了4.99、3.78、9.45、37.77、50.41、55.23、61.24个百分点,有着较为明显的提高;和优化的算法相比改进较小,分别为0.54、0.45、0.56、1.1、4.12、2.35个百分点,整体上看本文算法的准确率都有提高;从高斯白噪声上的实验数据可以看出,高斯白噪声对本文算法实验精度影响较小,但是在经典算法上的识别精度有明显下降趋势。

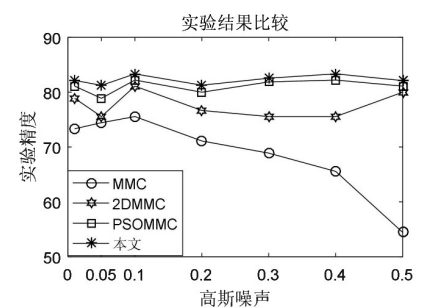


图14 高斯噪声准确率

(下转第188页)