基于神经网络的图像来源识别方法比较研究

孟 旭,孟 坤
(北京信息科技大学 计算机学院,北京 100101)

摘 要:随着手机等便携式智能电子设备的普及,图像已成为最重要的信息载体之一,在新闻、社交及司法等领域发挥着重要作用。在享用电子图像带来便捷性的同时,图像处理工具给不法分子通过篡改电子图像实施诈骗等犯罪活动提供了可能,识别图像来源、辨别图像真伪已成为遏制和惩罚此类犯罪活动的重要技术手段。该文讨论了神经网络在图像源识别中的应用方法,分别将原始图像和图像噪声作为模型输入数据,比较分析了神经网络的分类效果。从依赖数据属性、数据预处理方法以及应用模式等方面进行了实验。通过对实验结果进行分析,发现提取有代表性的图像块以及使用平滑的图像进行实验更有利于图像来源的识别。分别采用笔者建立的数据集(10 个相机)和 vision 数据集(35 个相机)作为分析数据集,图像来源分类的实验结果表明相对于简单估计相机传感器模式噪声的方法准确率提升了 35%,图像来源判断的实验结果准确率达到了 95%。

关键词:图像来源识别;噪声提取;神经网络;特征提取;传感器模式噪声
中图分类号:TP301 **文献标识码**:A **文章编号**:1673-629X(2022)01-0111-06
doi:10.3969/j.issn.1673-629X.2022.01.019

Comparison and Analysis of Image Source Identification Method Based on Neural Network

MENG Xu, MENG Kun
(School of Computer Science, Beijing Information Science and Technology University, Beijing 100101, China)

Abstract: With the popularization of portable smart electronic devices such as mobile phones, images have become one of the most important information carriers, playing an important role in news, social and judicial fields. While enjoying the convenience of electronic images, image processing tools make it possible for criminals to commit fraud and other criminal activities by tampering with electronic images. Identifying the source of the image and distinguishing the authenticity of the image has become an important technical means to deter and punish such criminal activities. We discuss the application method of neural network in image source identification, and compare and analyze the classification effect of neural network for the original image and image noise as model input data. Experiments are carried out in terms of dependent data attributes, data preprocessing methods, and application modes. According to the analysis of the experimental results, the extraction of overlapping image blocks and the use of smooth images for experiments are more conducive to the identification of image sources. By using our own dataset (10 cameras) and vision dataset (35 cameras) as the analysis data sets, the experimental results of image source classification show that the accuracy of the method of simply estimating the camera sensor pattern noise is improved by 35%. The accuracy of the experimental results of image source judgment reached 95%.

Key words: image source identification; noise extraction; neural network; feature extraction; sensor pattern noise

0 引 言

随着手机、平板电脑、数码相机等便携式设备的普及,人们每天都能产出大量的图像。这些图像为人们的生活带来了便利,同时也存在着一定的安全隐患。人们可以方便地对图像进行修改,这导致不真实的图像的数量大大增加。这些不真实的图像给新闻、司法等领域带来了极大的挑战。因此,为了保证图像的真实性和原始性,对图像进行来源识别是十分重要的。

图像来源识别是一种从图像本身识别图像来源的技术,可以通过提取图像本身特有的特征信息进行图像来源的识别。现有的图像来源识别方法有很多,基于传感器模式噪声(sensor pattern noise, SPN)的方法

收稿日期:2021-01-26 **修回日期**:2021-05-31
基金项目:国家自然科学基金项目(61502039)
作者简介:孟 旭(1994-),女,硕士研究生,研究方向为电子物证取证;孟 坤,副教授,研究方向为随机模型、网络性能评价、电子物证取证、计算智能。

是其中之一。在此方法中,人们通过提取图像的噪声残差来估计相机的 SPN,然后通过计算测试图像与相机的 SPN 的相似度来确定图像的来源。但是这种方法容易受到图像中其他模式噪声的影响,不够纯净的噪声信息容易影响图像来源识别的准确性。近几年随着神经网络的广泛应用,人们开始将神经网络用于图像来源识别中,通过神经网络学习图像的特征,以达到图像来源识别的目的。

该文针对图像来源识别的工作进行了一系列实验,探索了神经网络在不同的实验场景下的应用,分析在两种不同的应用模式下神经网络的分类效果。第一种应用场景是针对来自多个成像设备的图片进行分类,分析这些图片分别来自哪个成像设备。第二种应用场景是对某一张图像进行来源的判断,判断该图片是否来自于某个成像设备。使用两个不同的数据集分别进行实验,第一个数据集是笔者自己拍摄的,该数据集包括来自 4 个品牌的 10 种不同的设备拍摄的 1 069 张图像。第二个是 vision 数据集[1],共有来自 11 个品牌的 35 种不同设备拍摄的 4 167 张平滑的图像。分别对这两个数据集中的数据进行了不同的预处理,获得了原始图像、噪声图像、分块原始图像和分块噪声图像四种不同类型的数据集。将获得的数据集分别作为神经网络的输入进行训练,探索不同的图像数据分类的效果。

1 相关工作

传感器模式噪声是图像成像过程中生成的相机特有的噪声信息,可以根据图像的这部分噪声信息进行图像来源的识别。近几年人们提出了很多噪声信息的图像来源识别方法。例如,张明旺等人[2]对图像噪声提取方法进行了总结。Gupta 等人[3]基于光响应不均匀性(photo response non-uniformity,PRNU)噪声是非常弱的噪声信号,分别在图像的高频和低频分量中应用 PRNU 提取方法提取 PRNU。该方法能够提高大多数 PRNU 提取方法的效率,并且获得了 Mihcak 滤波器的最佳结果。Zeng 等人[4]提出了一种基于双树复数小波变换(dual-tree complex wavelet transform,DTCWT)[5-6]的方法从给定图像中提取 SPN,从而在强边缘周围的区域中实现了更好的性能。Zhao 等人[7]提出了一种使用权重函数的特征维数降低算法,该算法将不重要的像素方向的参考图案排除在外,保留了更重要的特征。Roy 等人[8]基于离散余弦变换残差特征[9]的提取以及随后的基于 AdaBoost 的基于随机森林的集成分类,提出了一种相机源识别方法。通过结合主成分分析的降维方法,提高了分类精度。

神经网络已被广泛应用于图像分析,并证明了其明显的优势,近几年越来越多的人将神经网络用于图像来源的识别中。Mandelli 等人[10]提出了一种基于 2 通道的卷积神经网络(convolutional neural network,CNN),该 CNN 学习了一种在色标级比较相机指纹和图像噪声的方法。所提出的解决方案比常规方法要快得多,并且可以确保较高的准确性。Kang 等人[11]提出的方法利用边缘分量干扰学习深度特征,对要输入深度网络的图像进行预处理,以提高摄像机模型识别精度。Tuama 等人[12]提出了一种可以自动并同时提取特征的方法,在学习过程中学习分类。之后 Rafi 等人[13]提出了 RemNet,该网络由两个主要的构建块组成:数据驱动的预处理块和分类块。Wang 等人[14]修改了类似于 AlexNet 的卷积神经网络结构,并为其配备了简单的局部二进制模式预处理操作。Huang 等人[15]提出将感兴趣的图像裁剪成较小的图像块,并采用局部到全局策略,根据图像块的多数表决权,就源摄像机做出最终决定。

从上述技术来看,基于 SPN 的方法已经广泛应用于图像来源识别领域,但是该方法对于图像噪声信息的提取的纯度要求较高,不纯净的噪声信息会影响图像来源识别的准确性。近几年人们越来越多地将神经网络用于图像来源识别领域,通过神经网络学习图像的特征来进行图像来源识别。

2 比较分析方法与数据集构造

该文使用神经网络进行图像来源识别的实验,将不同的实验数据作为输入,探究神经网络的分类效果。在两个数据集(笔者建立的数据集和 vision 数据集)上进行了实验,对两个数据集中的数据进行不同的预处理,获得了原始图像、噪声图像、分块原始图像和分块噪声图像四种不同类型的数据。实验设置两种不同的应用模式:针对来自多个成像设备的图片进行分类和判断某张图片是否来自于某个特定的成像设备。使用预处理后的数据在两种不同的应用模式下进行实验,对比分析方法与实验设置如表 1 所示。

表 1 对比分析方法与实验设置

应用模式	实验数据	
	笔者建立的数据集	vision 数据集
图像来源分类	原始图像	原始图像
	噪声图像	噪声图像
	原始图像块	原始图像块
	噪声图像块	噪声图像块
图像来源判断	原始图像	原始图像
	噪声图像	噪声图像
	原始图像块	原始图像块
	噪声图像块	噪声图像块

2.1　神经网络分类器设计

该文对 AlexNet 卷积神经网络[16]进行了修改,以适应模型的需求。提出的网络结构如表 2 所示,由三个卷积层、一个最大池化层和两个全连接层组成。

表 2　神经网络结构

结构	参数
COV1	kernel_size:5 * 5 stride=2
COV2	kernel_size:5 * 5 stride=2
COV3	kernel_size:3 * 3 stride=1
Max pooling	kernel_size:3 * 3 stride=2 dropout:0.2
FC1	In:30 752 Out:256 Dropout:0.5
FC2	In:256 Out:4 096 Dropout:0.5

神经网络结构具体如下:输入图像是大小为 255 * 255 的图像块。第一个卷积层(COV1)使用 64 个大小为 5 * 5 的卷积核处理输入的图像块;第二个卷积层(COV2)将第一层的输出作为输入,使用大小为 5 * 5 的卷积核进行卷积;第三个卷积层(COV3)应用 32 个大小为 3 * 3 的卷积核进行卷积。线性整流函数(ReLU)是非线性激活函数,应用于每个卷积层的输出。第三个卷积层之后是大小为 3 * 3 的最大池化层。完全连接层(FC1)和(FC2)分别具有 256 和 4 096 个神经元。ReLU 激活函数应用于完全连接层的输出。在最大池化层和两个完全连接层上分别应用 0.2 和 0.5 的丢失率,以防止出现过拟合问题。最后一个完全连接层的输出被馈送到 softmax 函数。

2.2　图像噪声提取

相机成像过程中会生成各种各样的噪声,这些噪声中会包含可以标识成像设备的信息,因此提取图像的噪声信息作为神经网络的输入进行训练。图像噪声的提取方法是使用去噪滤波器对原始图像进行降噪,然后由原始图像减去降噪后的图像就可以获得噪声图像了,如式(1)所示:

$$R = I - F(I) \tag{1}$$

其中,I 是原始图像,$F(\bullet)$ 是降噪的过程,该文采用如下 5 * 5 的高斯核进行图像噪声的计算。

$$\frac{1}{273}\begin{bmatrix} 1 & 4 & 7 & 4 & 1 \\ 4 & 16 & 26 & 16 & 4 \\ 7 & 26 & 41 & 26 & 7 \\ 4 & 16 & 26 & 16 & 4 \\ 1 & 4 & 7 & 4 & 1 \end{bmatrix}$$

2.3　图像块提取方法

提取图像块有两个作用:第一是扩充数据集;第二是提取更具代表性的图像块进行实验。由于完整的图像中包含各种各样的图像信息,其中有很多信息不能作为图像来源识别的特征,因此提取部分的图像块进行实验,从而减少其他信息对于图像来源识别的干扰。

该文主要提取图片边缘和中心区域的信息,即提取了图片的四个角、各边中心以及图像中心 9 个区域的图像块。提取的图像块如图 1 所示。

(a)原图　　(b)图像块

图 1　图像块提取

3　实验结果与分析

该文使用上述数据预处理获得的四类数据集分别进行了基于两种应用模式的实验。首先进行了简单的基于 SPN 的实验作为基准,然后基于两种不同的应用模式进行了基于神经网络的实验,最后探索分析了不同实验数据对于图像来源识别的分类效果。

3.1　原始数据集

实验中使用了两个不同的数据集。第一个是笔者建立的数据集,该数据集共有来自 4 个品牌的 10 种不同的设备拍摄的 1 069 张图像。第二个是由 vision 数据集提供的,来自 11 个品牌的 35 种不同设备拍摄的 4 167张平滑图像。数据集的具体情况如表 3 和表 4 所示。

表 3　自己建立的数据集

型号	数量
Apple 11	128
Apple xr	61
Apple 6plus	100
Apple iPad2018	100
HUAWEI Honor7X	116
HUAWEI JKM-TL00	120
HUAWEI TNY-AL00	92
HUAWEI nova3	182
Meizu 16spro	92
Xiaomi redmik20pro	77

表 4 vision 数据集

型号	数量
Samsung GalaxyS3Mini	78
Samsung GalaxyS3Mini	60
Samsung GalaxyS3	102
Samsung GalaxyS4Mini	112
Samsung GalaxyS5	100
Samsung GalaxyTrendPlus	151
Samsung GalaxyTab3	61
Samsung GalaxyTabA	126
Apple 4	109
Apple 4S	103
Apple 4S	133
Apple 5	100
Apple 5	106
Apple 5C	130
Apple 5C	113
Apple 5C	101
Apple 6	149
Apple 6	110
Apple 6Plus	169
Apple ipad2	159
Apple iPadMini	119
HUAWEI P8	126
HUAWEI P9	118
HUAWEI P9Lite	115
HUAWEI Ascend	84
HUAWEI Honor5c	80
Oneplus A3000	176
Oneplus A3003	150
Xiaomi RedmiNote3	174
LG D290	141
Sony XperiaZ1Compact	100
Wiko Ridge4G	140
Lenovo P70A	158
Microsoft Lumia640LTE	97
ASUS Zenfone2Laser	117

3.2 实验分析

该文分别使用上述数据集进行了三类实验:基于 SPN 的实验、使用神经网络实现图像来源的分类和使用神经网络实现图像来源的判断。

3.2.1 基于 SPN 的实验

基于 SPN 的实验是通过估计相机的 SPN 信息,计算图像噪声残差与相机 SPN 的相似性,进而识别图像的来源。该文使用 2.2 中的噪声提取方法获得图像的噪声残差,使用来自同一相机的 70% 的图像,采用均值计算的方法进行相机 SPN 的估计,然后使用剩余的 30% 的图像来验证识别的准确性。同时对两个数据集的数据进行了不同的预处理,获得了三类不同的数据集:第一是完整的图像;第二是使用 2.3 中提出的方法提取部分图像块;第三是使用文献[17-18]中提出的图像块提取方法计算权重对图像块进行加权。实验结果如图 2 所示。

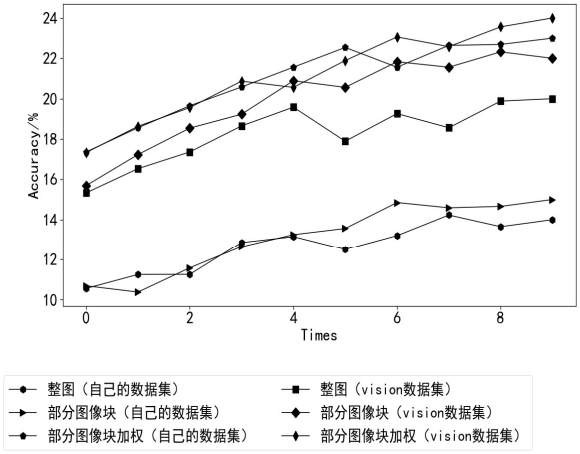

图 2 基于 SPN 的实验结果对比

通过以上实验可以看出,仅仅使用简单的噪声提取方法估计相机的 SPN 对于图像来源识别的效果不太理想,这是因为图像的噪声信息包含很多除 SPN 以外的其他模式噪声,这些噪声会干扰图像来源识别的准确率。其中使用部分图像块加权提取噪声信息和使用平滑的图像进行实验,实验结果准确率获得了一定的提升。因此,笔者认为选择具有代表性的部分图像块和使用平滑的图像(例如,天空图像)能够提取更加纯净的噪声信息,在一定程度上能够减少其他噪声的干扰,更有利于图像来源的识别。

3.2.2 使用神经网络实现图像来源的分类

使用神经网络实现图像来源分类的实验是使用 2.1 节中提出的神经网络来学习图像的特征信息,然后进行图像来源识别的。在此实验中,将图像切分为 255 * 255 的图像块(将边缘大小不够 255 * 255 的图像块删去),输入神经网络学习图像的特征。其中训练集占全部图像数量的 70%,测试集占 30%。分别使用经过不同预处理的数据集进行实验,探索神经网络在图像来源分类中的应用。

基于完整图像的实验。在这个实验中,从图像的左上角开始将图像分为多个 255 * 255 的图像块(将边缘大小不够 255 * 255 的图像块删去)。因此,笔者建立的数据集共有 29 967 个图像块,vision 数据集共有 361 932 个图像块。使用这些图像块进行两部分实验:

第一部分将获得的图像块直接输入神经网络进行训练;第二部分使用 2.2 中提出的方法提取图像的噪声,然后输入神经网络进行训练。实验结果表明,使用噪声图像进行实验,结果准确率比原始图像提升了 10%。

基于部分图像块的实验。在这个实验中,使用 2.3 中的方法提取大小为 255 * 255 的图像块。因此,笔者建立的数据集共有 9 612 个图像块,vision 数据集共有 37 503 个图像块。使用这些图像块分别进行两部分实验:第一部分将图像块直接输入到神经网络进行训练;第二部分使用 2.2 中提出的方法提取图像块的噪声,然后输入神经网络进行学习。实验结果表明,对于笔者建立的数据集,由于数据集的数据量相对较少,实验的结果有所下降。对于 vision 数据集,使用部分图像的实验的准确率比使用整个图像实验的准确率提升了 1% ~4%,同时,使用部分图像进行实验,耗费的时间更短。

使用神经网络实现图像来源分类的实验结果如图 3 所示。通过以上实验可以看出,使用部分图像块学习图像的特征信息,准确率提升了 26%,使用滤波后的图像进行实验,准确率提升了 35%。因此,笔者认为选择有代表性的图像信息进行实验比使用完整图像信息进行实验更有利于图像特征信息的提取。

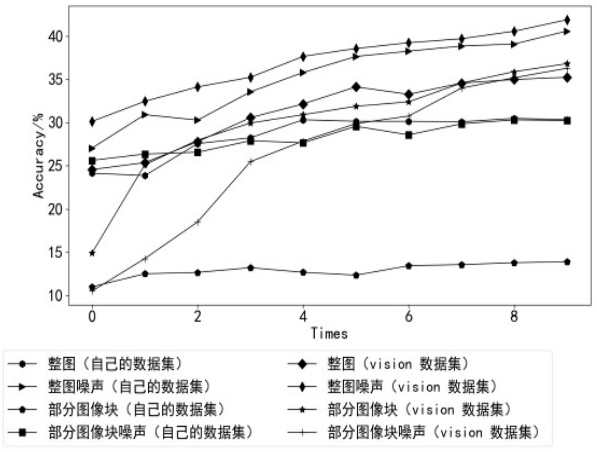

图 3　使用神经网络实现图像来源分类的实验结果对比

3.2.3　使用神经网络进行图像来源的判断

使用神经网络实现图像来源判断的实验是判断一个图像是否来自于某个特定的相机。在此实验中,使用神经网络训练图像的特征,然后将待测图像与标记好的图像进行比较,判断测试图像是否来自该相机。分别使用经过不同预处理的数据集进行实验,探索神经网络在图像来源判断中的应用。使用与 3.2.2 节中相同的数据集进行四个实验:整图、噪声图、图像块和噪声图像块。将图片输入神经网络进行训练获得来自同一相机的图像特征,提取测试特征的信息进行对比,判断该测试图像是否来自特定相机。实验结果如图 4 所示。

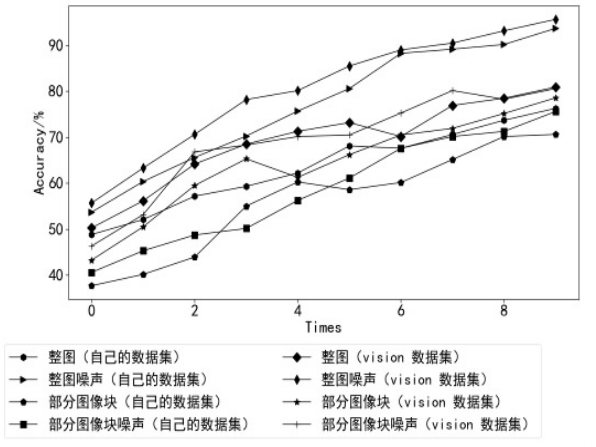

图 4　使用神经网络实现图像来源判断的实验结果对比

通过以上实验可以看出,使用部分图像块学习图像的特征信息进行图像来源判断的准确率相对较低。笔者认为原因是提取的图像块包含的图像信息较少,不利于图像特征的学习。使用滤波后的图像进行实验,准确率达到了 95%。因此,选择更具代表性的图像信息进行实验比使用完整的图像信息进行实验更有利于图像来源的判断。

4　结束语

该文探索了神经网络在图像来源分类和图像来源判断两个场景下的应用,分别在多个不同类型的数据集上进行了多次实验,通过对实验结果的分析,发现选择有代表性的图像信息和使用平滑的图像更加有利于图像来源的识别。在未来的工作中,将探索更好的图像处理方法,选择更具代表性的图像区域并提取更加纯净的图像特征信息,以进一步提高图像来源识别的准确性。

参考文献:

[1] SHULLANI D,FONTANI M,IULIANI M,et al. VISION:a video and image dataset for source identification[J]. EURASIP Journal on Information Security,2017,2017(1):1-16.

[2] 张明旺,肖延辉,田华伟,等. 图像中的设备指纹提取技术研究综述[J]. 激光与光电子学进展,2020,57(22):220003.

[3] GUPTA B,TIWARI M. Improving source camera identification performance using DCT based image frequency components dependent sensor pattern noise extraction method[J]. Digital Investigation,2018,24:121-127.

[4] ZENG H,WAN Y,DENG K,et al. Source camera identification with dual-tree complex wavelet transform[J]. IEEE Ac-

cess,2020,8:18874-18883.
[5] SELESNICK I W, BARANIUK R G, KINGSBURY N C. The dual-tree complex wavelet transform[J]. IEEE Signal Processing Magazine,2005,22(6):123-151.
[6] NICK K. The dual-tree complex wavelet transform:a new efficient tool for image restoration and enhancement[C]// 9th European signal processing conference. Rhodes,Greece: IEEE,1998.
[7] ZHAO Y,ZHENG N,QIAO Tong,et al. Source camera identification via low dimensional PRNU features[J]. Multimedia Tools and Applications,2019,78(7):8247-8269.
[8] ROY A,CHAKRABORTY R S,SAMEER U,et al. Camera source identification using discrete cosine transform residue features and ensemble classifier[C]//IEEE computer society conference on computer vision and pattern recognition workshops. Honolulu,HI,USA:IEEE,2017:1848-1854.
[9] HOLUB V,FRIDRICH J. Low-complexity features for jpeg steganalysis using undecimated dct[J]. IEEE Transactions on Information Forensics and Security,2015,10(2):219-228.
[10] MANDELLI S,COZZOLINO D,BESTAGINI P,et al. CNN-based fast source device identification[J]. IEEE Signal Processing Letters,2020,27:1285-1289.
[11] KANG C,KANG S. Camera model identification using a deep network and a reduced edge dataset[J]. Neural Computing and Applications,2020,32(17):13139-13146.
[12] TUAMA A,COMBY F. Camera model identification with the use of deep convolutional neural networks[C]//8th IEEE international workshop on information forensics and security. USA:Institute of Electrical and Electronics Engineers Inc. ,2017.
[13] RAFI A M,TONMOY T I,KAMAL U,et al. RemNet:remnant convolutional neural network for camera model identification[J]. arXiv:1902.00694,2020.
[14] WANG B,YIN J,TAN S,et al. Source camera model identification based on convolutional neural networks with local binary patterns coding[J]. Signal Processing:Image Communication,2018,68:162-168.
[15] HUANG N,HE J,ZHU N,et al. Identification of the source camera of images based on convolutional neural network [J]. Digital Investigation,2018,26:72-80.
[16] KRIZHEVSKY A, SUTSKEVER I, HINTON G. Imagenet classification with deep convolutional neural networks[J]. Communications of the ACM,2017,60(6):84-90.
[17] BONDI L, BAROFFIO L, et al. First Steps Toward Camera Model Identification With Convolutional Neural Networks [J]. IEEE Signal Processing Letters, 2017, 24(3): 259-263.
[18] FERREIRA A, CHEN H, et al. An Inception-Based Data-Driven Ensemble Approach to Camera Model Identification [C]. 10th IEEE International Workshop on Information Forensics and Security, December 10-13, 2018. USA: Institute of Electrical and Electronics Engineers Inc, 2018.

(上接第 110 页)

[19] HUANG K K,DAI D Q,REN C X,et al. Learning kernel extended dictionary for face recognition[J]. IEEE Transactions on Neural Networks and Learning Systems,2016,28(5):1082-1094.
[20] AHONEN T,HADID A,PIETIKAINEN M. Face description with local binary patterns:application to face recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2006,28(12):2037-2041.
[21] 马 振,刘凤连,汪日伟. 基于子模式下 LBP-HOG 特征融合的单样本人脸识别方法[J]. 光电子 · 激光,2019,30(12):1309-1316.
[22] WANG X,ZHANG B,YANG M,et al. Robust joint representation with triple local feature for face recognition with single sample per person[J]. Knowledge-Based Systems,2019,181:104790.
[23] 童 莹,沈越泓,魏以民. 基于旋转主方向梯度直方图特征的判别稀疏图映射算法[J]. 物理学报,2019,68(19):194202.
[24] TONG Y,ZHANG J,CHEN R. Discriminative sparsity graph embedding for unconstrained face recognition[J]. Electronics,2019,8(5):1-21.
[25] YANG M,ZHANG L,ZHANG D,et al. Relaxed collaborative representation for pattern classification[C]//2012 IEEE conference on computer vision and pattern recognition. Providence,RI,USA:IEEE,2012:2224-2231.