

基于 PCA-LDA 和 RIC-CNN 的旋转 无关手写识别

张墨逸, 袁小芳, 张 彬, 陈海燕

(兰州理工大学 计算机与通信学院, 甘肃 兰州 730050)

摘 要: 旋转无关手写字符识别对提高手写识别系统的实际应用准确性和鲁棒性至关重要。针对旋转字符识别准确率低、旋转不变坐标卷积神经网络(RIC-CNN)模型耗时长的问题,提出一种新的旋转无关手写字符识别方法。首先,在原RIC-CNN模型中引入非分离旋转矩不变量方法,并创建特征提取层,提取旋转不变数据特征,解决数据提取不充分的问题。接着,应用基于正交投影的PCA-LDA算法,该算法通过引入随机矩阵对特征进行正交投影,并将主成分分析(PCA)和线性判别分析(LDA)进行结合构建协方差矩阵和散射矩阵,从特征中筛选出有效的旋转不变特征,提高识别精度。最后,采用全连接神经网络(FCNN)对提取的特征进行分类识别。实验在空中手写数据集、MNIST数据集、中科院数据集与哈工大数据集上进行了验证。结果表明,该方法显著提升了旋转手写字符的识别精度,其中数字旋转数据的准确率达到97.72%。以MNIST数据集为例,该方法识别字符的时间仅为14.16秒,相比单独使用RIC-CNN模型,时间成本减少了19.66秒,充分证明了该方法的有效性。

关键词: 手写字符识别;非分离旋转矩不变量;旋转不变坐标卷积神经网络;正交投影;主成分分析;线性判别分析;全连接神经网络

中图分类号: TP391

文献标识码: A

文章编号: 1673-629X(2025)04-0093-07

doi: 10.20165/j.cnki.ISSN1673-629X.2024.0384

Rotation-independent Handwriting Recognition Based on PCA-LDA and RIC-CNN

ZHANG Mo-yi, YUAN Xiao-fang, ZHANG Bin, CHEN Hai-yan

(School of Computer and Communication, Lanzhou University of Technology, Lanzhou 730050, China)

Abstract: Rotation-independent handwriting character recognition is essential to improve the accuracy and robustness of handwriting recognition systems in practical applications. Aiming at the problems of low accuracy of rotated character recognition and long time consumption of the Rotation Invariant Coordinate Convolutional Neural Network (RIC-CNN) model, a new rotation-independent handwritten character recognition method was proposed. Firstly, the non-detached rotational moment invariant method was introduced into the original RIC-CNN model, and a feature extraction layer was created to extract the rotational invariant data features to solve the problem of insufficient data extraction. Then, the PCA-LDA based on orthogonal projection was applied, which introduced random matrices to project the features orthogonally, and combined Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) to construct covariance matrix and scattering matrix, so as to screen out effective rotational invariant features from the features and improve the recognition accuracy. Finally, the Fully-Connected Neural Network (FCNN) was used to classify and identify the extracted features. The experiments were verified on the aerial handwriting dataset, the MNIST dataset, the Chinese Academy of Sciences dataset and the Harbin Institute of Technology dataset. The results show that the proposed method significantly improves the recognition accuracy of rotated handwritten characters, and the accuracy of rotation data reaches 97.72%. Taking the MNIST dataset as an example, the time for the proposed method to recognize characters is only 14.16 seconds, which is 19.66 seconds less than the RIC-CNN model alone, which fully proves the effectiveness of the method.

Key words: handwriting character recognition; non-detached moment of rotation invariant; rotationally invariant coordinate convolutional neural network; orthographic projection; principal component analysis; linear discriminant analysis; fully connected neural networks

收稿日期: 2024-08-14

修回日期: 2024-12-17

基金项目: 国家自然科学基金项目(62161019)

作者简介: 张墨逸(1985-),女,副教授,博士,研究方向为机器视觉、机器学习、模式识别;通信作者: 袁小芳(1998-),女,硕士研究生,研究方向为模式识别与人工智能。

0 引言

手写字符识别一直是计算机视觉和模式识别的重要研究方向。随着智能技术的发展,旋转手写字符识别作为其重要分支,具有重要意义。然而,现实生活中旋转角度的多样性和形变的复杂性,使得旋转手写字符识别变得非常困难。

为解决以上问题,学者们提出许多创新方法和技术,主要有旋转不变特征提取算法、神经网络模型等。然而,常用的特征提取算法如 Zernike 不变矩^[1]、Hu 矩^[2]、尺度不变特征变换 (SIFT)^[3]、方向梯度直方图 (HOG)^[4]、旋转不变局部二进制模式 (ri-LBP)^[5]、旋转不变特征变换 (RIFT)^[6]、非分离旋转矩不变量^[7]等不能充分提取旋转字符特征,处理复杂任务时效果不佳。

在神经网络模型方面,如旋转等变向量场网络 (RotEqNet)^[8]、谐波网络 (H-Net)^[9]、群等变卷积网络 (G-CNN)^[10]等,虽能实现数据的旋转不变性,但都存在网络结构复杂、数据提取不充分、计算复杂度高等问题。目前,研究人员提出了一种新的旋转不变坐标卷积神经网络 (RIC-CNN)^[11]。在该网络中,设计了一

个旋转不变坐标卷积 (RIC-C),利用其将 CNN 中所有的标准卷积进行替换,构建了一个旋转不变坐标卷积神经网络。但该网络在大规模数据集上的训练时间成本非常高。

因此,该文在原 RIC-CNN 模型基础上对其进行改进,使其在充分提取数据旋转不变特征的同时极大地缩减 RIC-CNN 的时间成本。此外,考虑到提取的特征需要输入分类器进行分类识别,直接影响分类结果的有效性,对特征进行筛选显得尤为重要。而在文献[12]中,提出基于特征权重的决策树特征选择算法,对所得到的特征进行筛选过滤,其结果验证了特征筛选的必要性。故该文提出了基于正交投影的 PCA-LDA 算法,将提取的特征进行分析与筛选,从而保证最终特征的有效性。

1 方法流程

该文提出了一种基于 PCA-LDA 和非分离旋转矩不变量 RIC-CNN 的旋转无关手写字符识别方法。该方法主要分为三个阶段:特征提取、特征选择和分类识别,如图 1 所示。

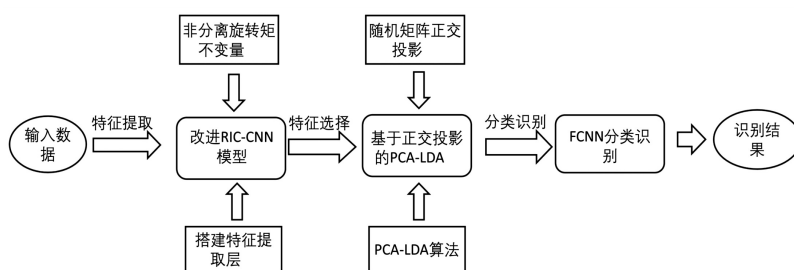


图 1 旋转无关手写字符识别流程

在特征提取阶段,引入非分离旋转矩不变量到 RIC-CNN 模型中,对输入目标进行分析并提取旋转不变特征。随后,在 RIC-CNN 中构建特征提取层,用于获取旋转不变特征向量。在特征选择阶段,采用基于正交投影的 PCA-LDA 算法对这些特征向量进行优化处理。首先,引用随机矩阵与特征向量进行正交投影,增加特征的随机性;之后,通过 PCA-LDA 算法对特征进行筛选,去除冗余信息,提升特征的针对性,从而优化特征向量;最后,在分类识别阶段,通过全连接神经网络对优化后的特征向量进行分类,从而获得相应的识别准确率。

2 非分离旋转矩不变量 RIC-CNN

2.1 非分离旋转矩不变量构造

非分离旋转矩不变量^[7],由不可分离的 Appell 矩组成。在文献[7]中证明了 Appell 多项式在旋转下表现为单项式,其中 Appell 矩是不可分离的。因此,利用 Appell 矩构造非分离旋转矩不变量。其计算公式如下:

首先,由生成函数定义的两个 Appell 多项式 u 与 v 如公式 1 和公式 2 所示。

$$v_{m,n}(x,y) = m! n! 2^{m+n} \sum_{i=0}^{\lfloor \frac{m}{2} \rfloor} (-1)^{i+j} \frac{(m+n-i-j)!}{i! j! (m-2i)! (n-2j)! 2^{2(i+j)}} x^{m-2i} y^{n-2j} \quad (1)$$

$$U_{m,n}(x,y) = \frac{m! n!}{2^{m+n}} \sum_{k=0}^{m+\lfloor \frac{n}{2} \rfloor} \sum_{j=\lceil \frac{m}{2} \rceil}^k (-1)^{j-k} \binom{m+n}{k} \binom{k}{j} \binom{2j}{m} \binom{2m+2n-2K}{n} x^{2j-m} y^{2(m-k)=n} \quad (2)$$

在文献[13]中提出了一个理论,该理论允许构造一组完整且独立的任意顺序旋转不变量,将该理论应

用于 Appell 多项式,如公式 3 所示。将从几何矩中获得的不变量公式 $x^p y^q$ 替换到 Appell 多项式中,得到

Appell 矩的旋转不变量。

$$\varphi_{pq} = \left(\sum_{k=0}^{q_0} \sum_{j=0}^{p_0} \binom{p}{k} \binom{q}{j} (-1)^{p_0-j} i^{p_0+q_0-k-j} A_{k+j, p_0+q_0-k-j} \right)^{p-q} \cdot \sum_{k=0}^p \sum_{j=0}^q \binom{p}{k} \binom{q}{j} (-1)^{q-j} i^{p+q-k-j} A_{k+j, p+q-k-j} \quad (3)$$

其中, $p > q$, p_0 与 q_0 都是用户自定义指数。

与可分离多项式如高斯-厄米特 (GH)^[14]、Zernike 矩^[1]相比,可分离多项式的零曲线始终遵循栅格,而非分离多项式的零曲线可以采取任意方向,并产生一种随机模式,更均匀地划分图像域,不受栅格的限制。这种特性对于非圆形大幅度多角度旋转的目标来说,更能提取到有效的数据特征。而原 RIC-CNN 模型会存在提取特征信息不充分、不全面等问题。因此,将该方法引入 RIC-CNN 模型中,使得模型提取的特征更加完整有效。

2.2 非分离旋转矩不变量引入 RIC-CNN

RIC-CNN,用于图像识别与分类的神经网络模

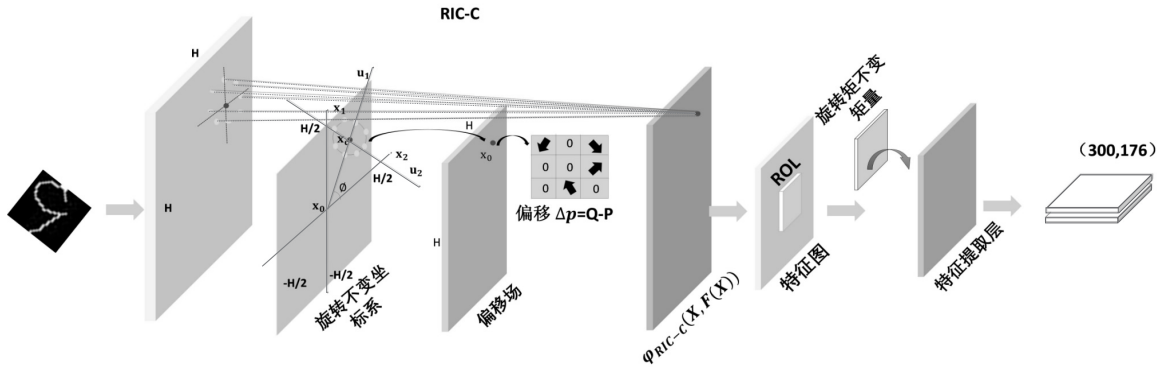


图2 非分离旋转矩不变量 RIC-CNN 模型

首先,该模型使用 RIC-C 卷积对输入数据进行卷积操作。通过 RIC-C 卷积,可以确保即使输入数据发生旋转,提取到的特征也能保持不变。接着,在原 RIC-CNN 模型中自定义一个卷积层,然后将非分离旋转矩不变量嵌入自定义的卷积层中。在模型的训练过程中,该方法与模型协同工作,一起提取数据的旋转不变特征,确保所得到的特征更加充分与完整。最后,在该模型中构建一个特征提取层,综合 RIC-C 卷积与非分离旋转矩不变量所得到旋转不变特征,实现有效的特征提取,便于后续将这些特征进行额外的分析与处理,从而获得更具有针对性的特征。

非分离旋转矩不变量在特征提取的过程中考虑整个图像的全局信息,从而能够得到更丰富的特征表达,增加数据的多样性,使得改进的 RIC-CNN 模型能够提供信息更加完整的旋转不变特征表示,有效提升模型整体识别性能。此外,该模型还无需对特征进行额外的训练,仅需对数据进行特征提取,从而节约了时

间。该网络通过在 CNN 中引入旋转不变坐标卷积 (RIC-C),使得 CNN 具有了旋转不变性。由文献 [11] 可知, RIC-C 定义如下:

$$\varphi_{\text{RIC-C}}(X_0, F(X)) = \sum_{Q \in T_{X_0}} W(P) * F(x_0 + Q) \quad (4)$$

其中, $P \in R$ 与 Q 具有相同的阶数且 $\varphi_{\text{RIC-C}}$ 对围绕原点 X_c 的任意图像旋转是不变的。用相应的 RIC-C 替换 CNN 模型中所有的传统卷积,得到了具有旋转不变性的 RIC-CNN。

通过其提取数据的特征,可以确保所得到的数据具有旋转不变性。然而,原 RIC-CNN 模型存在计算复杂度高、训练时间长等问题,需要强大的硬件资源来加速训练和推理过程,或者使用高性能的 GPU 或 TPU 来提升模型的计算效率。因此,该文对原 RIC-CNN 模型进行了改进,引入了非分离旋转矩不变量,并在模型中构建了一个特征提取层,用来提取数据的旋转不变特征,如图 2 所示。

间,缩短了训练过程。

3 基于正交投影的 PCA-LDA 算法

针对 RIC-CNN 模型所得到的旋转不变特征复杂度太高、冗余信息繁多的问题,该文采用了一种基于正交投影的 PCA-LDA^[15-16] 算法。首先将得到的特征与生成的随机矩阵进行正交相乘,增加数据随机性。随后采用 PCA 投影对数据进行降维处理,减少数据的复杂性和冗余性。最后采用 LDA 投影进行线性判别,使得数据更具有针对性与准确性。其流程如图 3 所示。

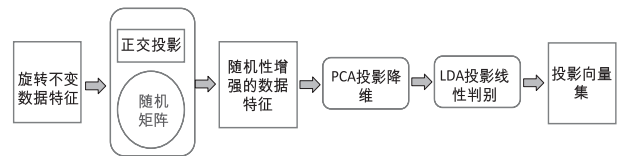


图3 基于正交投影的特征选取算法流程

该算法不仅增加了数据的随机性,提升了模型的泛化能力,还筛选出了关键的特征,成功降低了特征空

间的维度,从而减少了计算复杂度,有效降低了过拟合的风险。同时,这种简洁有效的数据表征显著提升了模型的训练效率和预测准确率。

3.1 正交投影

投影技术是一种常用的数据处理技术,主要目的是将原始特征空间中的数据投影到一个新的特征空间,以获取更具有判别性和泛化能力的特征表示。将所得到的旋转不变特征与随机矩阵进行正交相乘,得到正交投影矩阵,其原理如公式 5 所示。

$$X_{\text{proj}} = X * R \quad (5)$$

其中, X 为旋转不变特征矩阵, R 为随机正交矩阵,它们的形状分别为 $m * n$ 和 $n * n$ 。正交投影后的特征矩阵保持着原始特征 X 的维度,但特征表示已经发生变化,从而获得了更加随机性的特征表示,提高了模型的泛化能力。

3.2 PCA-LDA 算法

PCA 主要依靠样本在空间中的位置信息,将特征投影到这些方向所在的直线上,在投影的过程中消除了样本之间的相关性和噪声。计算步骤如下:

令投影矩阵 X_{proj} 为 X , 其均值向量为:

$$\mu = \frac{1}{m} \sum_{i=1}^m X_i \quad (6)$$

接着计算总协方差矩阵 C , 令 $\mu = 0$ 将数据的各个维度都中心化 0。

$$c = \frac{1}{m} \sum_{i=1}^m (x_i - \mu) (x_i - \mu)^T = \frac{1}{m} X X^T \quad (7)$$

对协方差矩阵进行特征值分解,得到特征值 $\lambda_1, \lambda_2, \dots, \lambda_n$ 和对应的特征向量 $\nu_1, \nu_2, \dots, \nu_n$ 。取前 k 个最大特征值所对应的特征向量组成主成分方向 U , 通过计算矩阵和特征向量,将数据变换到新的特征空间 Y 中,得到新的特征。

$$Y = U^T X \quad (8)$$

在使用 PCA 去除数据特征的冗余信息之后,采用

线性判别分析 (LDA) 优化算法进一步投影变化降低数据维度,使投影后的样本类间散度尽可能的大,类内离散度尽可能的小。设总均值向量为 d , 各类样本的均值为 d_i , 新的特征为 Y_{ij} (表示第 i 类的第 j 个样本), 共有 m 类, 每类有 g 个样本。因此, 样本的类内离散矩阵 S_w 与类间离散矩阵 S_b 分别为:

$$S_w = \sum_{i=1}^m \sum_{j=1}^k (Y_{ij} - d_i) (Y_{ij} - d_i)^T \quad (9)$$

$$S_b = \sum_{i=1}^m g(d - d_i) (d - d_i)^T \quad (10)$$

则特征值 λ 与对应的特征向量 ν 最佳投影方向之间的关系为:

$$\lambda = L(v) = \frac{\nu^T S_b \nu}{\nu^T S_w \nu} \quad (11)$$

选 k 个最大特征值对应的特征向量 $\nu = [w_1, w_2, \dots, w_k]$ 投影方向, 该投影矩阵更能准确地表示样本特征。

经过基于正交投影的 PCA-LDA 算法进行特征选择后, 使用全连接神经网络分类器^[17]对所选特征进行分类识别。全连接神经网络在面对新的、未见过的数据样本时展现出良好的泛化能力。其结构和参数可以根据具体任务的需求进行调整和优化, 从而适应不同的数据和问题。因此, 通过采用全连接神经网络分类器, 可以获得更优越的分类效果。

4 实验结果分析

4.1 实验数据

实验采用了四个不同的数据集进行验证, 分别是空中手写数据集、MNIST 数据集、哈工大数据集与中科院数据集。各数据集中包含训练集和测试集, 其中主要的信息如表 1 所示, 训练集部分样本如图 4 所示。然后, 对每张测试图像进行 0 到 360 度的随机旋转, 生成旋转测试图像, 构成旋转测试集, 如图 5 所示。

表 1 实验数据集

数据集	类别	总样本/张	训练数据/张	测试数据/张	旋转测试集/张
空中手写字符	数字(10 类)	10 000	6 000	4 000	4 000
MNIST	数字(10 类)	70 000	60 000	10 000	10 000
哈工大数据集	字母(26 类)	2 600	1 560	1 040	1 040
中科院数据集	汉字(50 类)	5 000	3 000	2 000	2 000

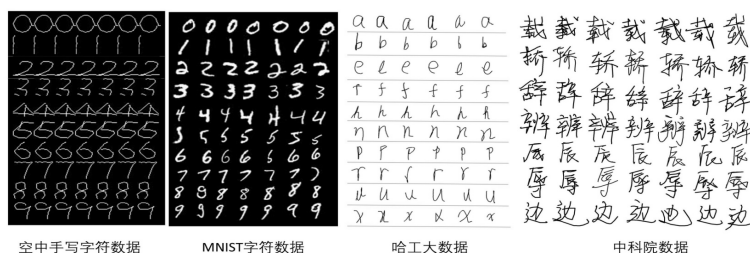


图 4 训练数据集中的一些样本

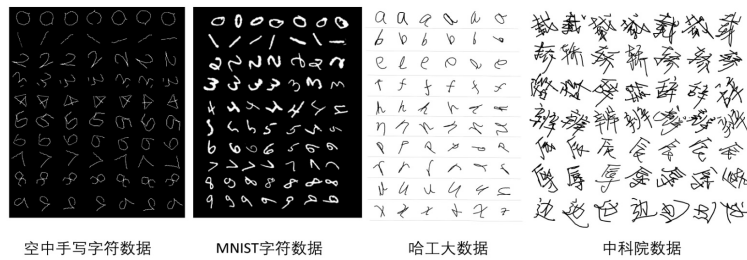


图5 旋转测试集中的一些样本

4.2 实验评估性能指标

该文主要以准确率和时间作为评估指标,准确率用于衡量模型的识别性能和分类准确性,而时间则反映了模型的计算效率和处理速度。通过这两个指标,不仅评估了该方法在实际应用中的精确度,还考察了其在处理任务时的实时性,从而全面反映了该方法的有效性和实用性。

准确率 Acc 的计算公式为:

$$\text{Acc} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (12)$$

其中,TP 是真正例的数量,TN 是真负例的数量,FP 是假证例的数量,FN 是假负例的数量。

时间 Time 的计算公式为:

$$\text{Time} = \text{Start} - \text{End} \quad (13)$$

其中,Start 指的是任务开始的时间点,End 指的是任务结束的时间点。

4.3 实验结果

4.3.1 不同数据集上的数据分析

本实验中,分别在四个不同的数据集上进行方法验证。将每个数据集中的原始测试集和旋转测试集进行 500 次迭代训练,对比原始测试集与旋转测试集的分类精度,其结果如表 2 所示。

表2 文中方法在不同数据集上的准确率 %

数据集	文中方法	
	原始数据准确率	旋转数据准确率
哈工大数据	99.12	97.03
中科院数据	95.33	78.99
MNIST 数据	99.34	96.97
空中手写字符	99.72	97.72

由表 1 的数据可以看出,文中方法无论是对于非旋转数据,还是旋转数据,其准确率均较高,尤其是数字,旋转数据的准确率达到 97.72%,验证了该方法的有效性。

4.3.2 各类方法在数据集上的性能

为了进一步验证文中方法,实验中使用不同的模型在 MNIST 数据集上进行对比实验。该数据集包含 60 000 个训练图像和 10 000 个测试图像,将每个测试图像随机旋转 0°到 360°,得到旋转测试集。分别对原

始测试集与旋转测试集进行 500 次迭代训练,对比不同方法的识别准确率。具体实验结果如表 3 所示。

表3 MNIST 数据集上各种方法的准确率

方法	MNIST 数据集		
	输入大小	原始数据/%	旋转数据/%
Zernike ^[1]	28 * 28	85.07	73.54
GA-CNN ^[18]	28 * 28	96.66	94.18
B-CNN ^[19]	32 * 32	96.95	80.02
H-Net ^[8]	28 * 28	98.79	91.21
G-CNN ^[10]	32 * 32	99.18	46.14
RIC-CNN ^[11]	32 * 32	98.68	94.87
CNN ^[20]	28 * 28	99.49	42.45
ORN ^[21]	28 * 28	99.37	79.31
RotEqNet ^[8]	32 * 32	99.12	73.59
文中方法	28 * 28	99.34	96.97

根据表中的数据,文中方法在旋转测试集上取得了最佳的准确率,达到了 96.97%,相较于原 RIC-CNN 模型 94.87% 的准确率提升了 2.1 百分点,显示了显著的优越性。为了进一步验证文中方法中每个模块的性能,进行了消融实验。消融实验的结果如表 4 所示。

表4 文中方法的消融实验

方法模块	准确率		时间
	非可分离算法	准确率 /%	
RIC-CNN	✓	94.87	33.82
PCA-LDA	✓	95.62	37.23
非可分离算法	✓	95.22	16.13
文中方法	✓	96.97	14.16

实验中使用了 MNIST 数据集中的旋转测试集进行了 500 次迭代,旨在评估 RIC-CNN 模型及其不同模块对准确率和时间性能的影响。结果显示,RIC-CNN 模型在旋转数据上的识别准确率为 94.87%。当引入非可分离旋转不变矩算法后,识别准确率提升至 95.62%,较 RIC-CNN 模型提高了 0.75 百分点。进一步结合基于正交投影的 PCA-LDA 算法后,识别准确率达到 96.97%,比单独使用 RIC-CNN 模型提升了 2.1 百分点。这些结果表明,每个模块的引入都显著提升了模型的性能,验证了该方法的有效性。

此外, RIC-CNN 模型单独运行的时间成本为 33.82 s。而在引入非分离旋转不变矩算法与基于正交投影的 PCA-LDA 算法后, 时间成本显著减少至 14.16 s, 在性能优化方面取得显著效果。

4.3.3 方法在各类数据上的性能

为了进一步验证文中方法的优越性, 将其与 Zernike 和 RIC-CNN 模型在不同类别(如数字、字母和汉字)旋转手写字符上的分类精度进行了对比, 结果如图 6~8 所示。

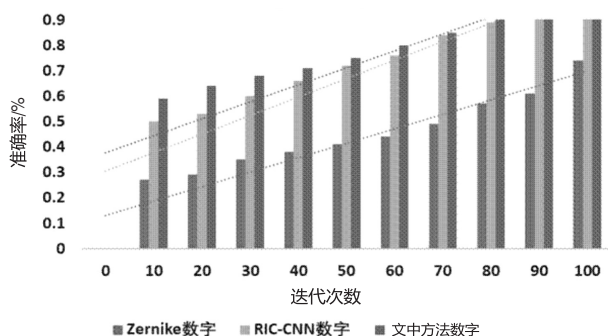


图 6 不同方法对数字的识别准确率

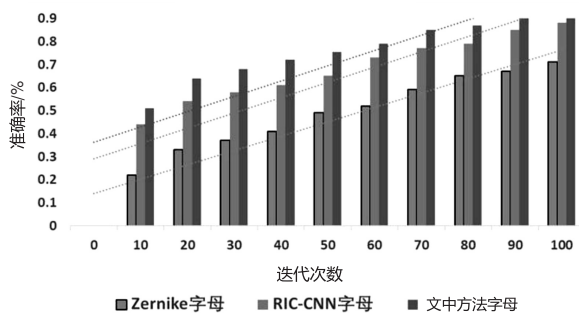


图 7 不同方法对字母的识别准确率

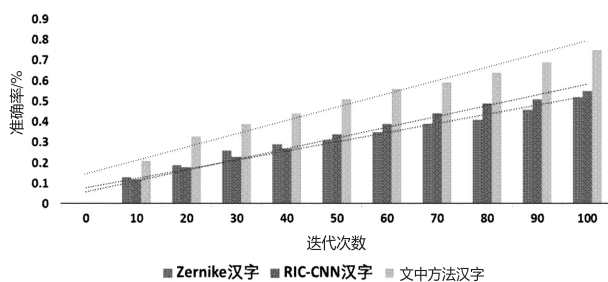


图 8 不同方法对汉字的识别准确率

从图 6~8 可以看出, 无论是数字、字母还是汉字, 相较于 Zernike 矩和 RIC-CNN 方法, 随着迭代次数的增加, 文中方法都具有较高的准确率。以复杂度更高的汉字为例, 这些汉字包含了多种字体和不同的书写风格。当迭代次数增加到 100 次时, 文中方法的准确率为 75%, 而 Zernike 矩和 RIC-CNN 方法的准确率分别为 46% 和 55%。由此可见, 文中方法在针对复杂度更高的手写字符具有更优的效果。为了进一步验证文中方法的识别效果, 分别使用其对汉字、数字和字母进行测试, 结果如图 9 所示。

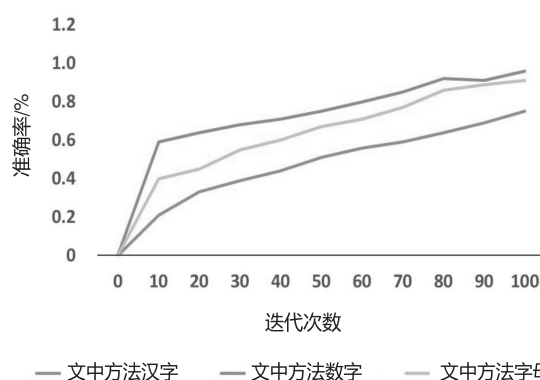


图 9 文中方法对字母、数字、汉字的识别准确率

在实验中, 分别从 MNIST 数据集、中科院数据集与哈工大数据集中选取了 1 560 张训练样本和 1 040 张测试样本, 并将测试样本进行 0° 到 360° 的随机旋转, 组成数字、字母与汉字三种类别的旋转测试集, 以验证文中方法。数字、字母与汉字的复杂度依次递增。从图 9 的结果来看, 当训练的迭代次数增加到 100 次时, 文中方法对数字、字母与汉字的准确率分别为 96.26%、95.55% 和 75%。准确率随着数据类型复杂度的增加而降低。由此可见, 文中方法在识别更复杂的数据时, 效果还有待提升。

4.3.4 方法的时效性

从表 1 中的数据可以看出, 文中方法在识别准确率方面表现优异。为了验证该方法的时效性, 从 MNIST 数据集中选取了 10 000 张训练样本和 5 000 张测试样本, 并对每张测试样本进行随机旋转, 生成旋转测试图像。然后, 使用这些图像对不同方法进行了时效性测试。在测试过程中, 按照 0 到 100 的迭代次数对各方法进行验证, 并将结果以折线图的形式展示, 如图 10 所示。

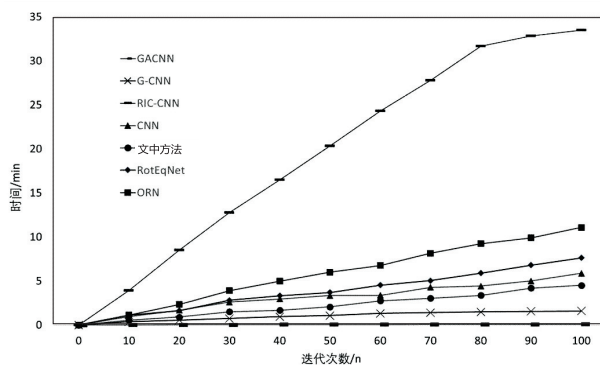


图 10 方法时间对比

从图中的数据可以看出, 相比于大型模型如 RIC-CNN 和 G-CNN, 该方法在时间成本上表现出显著优势, 其时间成本几乎减少了一半。此外, 随着迭代次数的增加, 该方法的时间波动非常小, 逐渐趋于稳定。因此, 以上数据清楚地证明了该方法在效率和稳定性方面的显著优势。

5 结束语

针对旋转字符准确率低、RIC-CNN 计算复杂度高、耗时长等问题,提出了一种基于PCA-LDA和非分离旋转矩不变量RIC-CNN的旋转无关手写字识别方法。利用非分离旋转矩不变量RIC-CNN提取旋转不变特征,针对所得到的特征具有冗余信息、复杂度高的问题,应用了基于正交投影的PCA-LDA算法,对特征进行分析与筛选,使其更加具有针对性,从而提升模型的训练效率和预测准确率。该方法在MNIST数据集、空中手写字、哈工大与中科院数据集上进行实验验证,根据实验结果得出以下结论:(1)该方法在旋转数据的识别性能上优于其他对比方法,显示出明显的优势。(2)特征选取算法引入随机矩阵,增加了数据的多样性,提升了模型的泛化能力。(3)该方法有效降低了时间成本,整体性能得到显著提升。然而,该方法在识别旋转样本中更加复杂的手写字(如不同类别的文字)时,效果欠佳。因此,未来需要探索其他解决方案以应对这一挑战。

参考文献:

- [1] 杨棉绒,牛丽平. 基于LGBM的Zernike特征选取及红外图像目标识别方法[J]. 红外与激光工程,2022,51(4): 20210309-1-20210309-6.
- [2] WU Z, JIANG S, ZHOU X, et al. Application of image retrieval based on convolutional neural networks and Hu invariant moment algorithm in computer telecommunications[J]. Computer Communications, 2020, 150: 729-738.
- [3] BRUNO A, ARDIZZONE E, VITABILE S, et al. A novel solution based on scale invariant feature transform descriptors and deep learning for the detection of suspicious regions in mammogram images[J]. Journal of Medical Signals & Sensors, 2020, 10(3): 158-173.
- [4] GHAFARI S, SOLEIMANI P, LI K F, et al. Analysis and comparison of FPGA-based histogram of oriented gradients implementations[J]. IEEE Access, 2020, 8: 79920-79934.
- [5] RASOOL M, KAUR A. A novel rotation invariant descriptor for texture classification with local binary patterns[C]//Soft computing and signal processing: proceedings of 3rd ICSCSP 2020, volume 1. Singapore: Springer, 2021: 385-396.
- [6] LI Y, LI B, ZHANG G, et al. sRIFD: a shift rotation invariant feature descriptor for multi-sensor image matching[J]. Infrared Physics & Technology, 2023, 135: 104970.
- [7] BEDRATYUK L, FLUSSER J, SUK T, et al. Non-separable rotation moment invariants[J]. Pattern Recognition, 2022, 127: 108607.
- [8] MARCOS D, VOLPI M, KOMODAKIS N, et al. Rotation equivariant vector field networks[C]//Proceedings of the IEEE international conference on computer vision. Las Vegas: IEEE, 2017: 5048-5057.
- [9] WORRALL D E, GARBIN S J, TURMUKHAMBETOV D, et al. Harmonic networks: deep translation and rotation equivariance[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. Honolulu: IEEE, 2017: 5028-5037.
- [10] COHEN T, WELLING M. Group equivariant convolutional networks[C]//International conference on machine learning. New York: PMLR, 2016: 2990-2999.
- [11] MO H, ZHAO G. RIC-CNN: rotation-invariant coordinate convolutional neural network[J]. Pattern Recognition, 2024, 146: 109994.
- [12] ZHOU H F, ZHANG J W, ZHOU Y Q, et al. A feature selection algorithm of decision tree based on feature weight[J]. Expert Systems with Applications, 2021, 164(4): 113842.
- [13] FLUSSER J. On the independence of rotation moment invariants[J]. Pattern Recognition, 2000, 33(9): 1405-1410.
- [14] YANG B, SHI X, CHEN X. Image analysis by fractional-order Gaussian-Hermite moments[J]. IEEE Transactions on Image Processing, 2022, 31: 2488-2502.
- [15] YE H, ZHU W, WANG C, et al. Faster voxelpose: real-time 3d human pose estimation by orthographic projection[C]//European conference on computer vision. Cham: Springer, 2022: 142-159.
- [16] LASALVIA M, CAPOZZI V, PERNA G. A comparison of PCA-LDA and PLS-DA techniques for classification of vibrational spectra[J]. Applied Sciences, 2022, 12(11): 5345.
- [17] 李伟, 霍雪松, 张明, 等. 基于残差全连接神经网络的电力监控系统异常行为检测方法[J]. 东南大学学报: 自然科学版, 2020, 50(6): 1062-1068.
- [18] HAO Y, HU P, LI S, et al. Gradient-aligned convolution neural network[J]. Pattern Recognition, 2022, 122: 108354.
- [19] DELCHEVALERIE V, BIBAL A, FRENAY B, et al. Achieving rotational invariance with Bessel-convolutional neural networks[J]. Advances in Neural Information Processing Systems, 2021, 34: 28772-28783.
- [20] ALZUBAIDI L, ZHANG J, HUMAIDI A J, et al. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions[J]. Journal of Big Data, 2021, 8(1): 40537-021-00444-8.
- [21] ZHOU Y, YE Q, QIU Q, et al. Oriented response networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. Honolulu: IEEE, 2017: 519-528.