

# 基于 Gabor 小波-卷积神经网络的 在线考试场景模糊人脸识别方法

何剑萍<sup>1</sup>, 蒋大锐<sup>2†</sup>, 杨波<sup>2</sup>

(1. 广州华商学院马克思主义学院, 广东 广州 511300; 2. 广州华商学院数据科学学院, 广东 广州 511300)

**摘要:**针对目前在线考试智能监考时因图像运动模糊、噪声等导致人脸识别准确率低的问题,提出了一种基于 Gabor 小波-卷积神经网络-支持向量机的人脸识别模型。将原始图像经 GW 滤波分解为包含了幅度和角度特征的协方差矩阵,从而提高模糊环境下人脸识别性能。还提出了一种改进的 CNN 网络学习 GW 生成的协方差矩阵,从而提取出人脸特征。应用 SVM 对人脸特征表示进行分类,最终输出人脸识别结果。通过实验验证,与 PCANet、VGGFace、ResNet50 模型相比,所提 GW-CNN-SVM 模型在低分辨率、运动模糊和噪声环境下识别性能更优。实验结果验证了所提 GW-CNN-SVM 模型对在线考试智能监考时低分辨率、运动模糊和噪声环境下的人脸识别具有更高的鲁棒性。

**关键词:**智慧教育;在线监考;人脸识别;特征提取;卷积神经网络;协方差矩阵

**中图分类号:**TP393

**文献标识码:**A

## Fuzzy Face Recognition Method in Online Examination Scene Based on Gabor Wavelet Convolutional Neural Network

HE Jianping<sup>1</sup>, JIANG Darui<sup>2†</sup>, YANG Bo<sup>2</sup>

(1. School of Marxism, Guangzhou Huashang College, Guangzhou, Guangdong 511300, China;

2. School of Data Science, Guangzhou Huashang University, Guangzhou, Guangdong 511300, China)

**Abstract:** In order to solve the problem of low accuracy of face recognition due to image motion blur and noise in intelligent invigilation of electronic assessment, a face recognition model based on Gabor wavelet convolutional neural network support vector machine is proposed. The original image is decomposed into a covariance matrix containing amplitude and angle features by GW filtering, thus improving the performance of face recognition in fuzzy environments. An improved CNN network learning the covariance matrix generated by GW is proposed to extract the face features. Apply SVM to classify facial feature representations and ultimately output facial recognition results. Through experimental verification, compared with PCANet, VGGFace, and ResNet50 models, the proposed GW-CNN-SVM model has better recognition performance in low resolution, motion blur, and noisy environments. The experimental results verify that the proposed GW-CNN-SVM model has higher robustness in low resolution, motion blur, and noisy environments during intelligent invigilation of school exams.

**Key words:** smart education; online invigilation; face recognition; feature extraction; convolutional neural network; covariance matrix

随着远程会议软件<sup>[1]</sup>的使用越来越广泛,在线教学模式和平台<sup>[2,3]</sup>已成院校教学不可或缺的一部分。在众多大学中,远程考试作为评估学生成绩的重要组成部分,继在线教学和远程学习之后,正逐渐成为一种主导趋势。且随着计算机视觉<sup>[4-6]</sup>技术不断发展,可将计算机视觉的分支技术人脸识别<sup>[7]</sup>引入在线考试监控领域,从而替代人力执行监考任务,不仅可以节省人力成本,还能大大增加考试的公平性。因此,准确的人脸识别将成为智能在线监考的前提和基础。目前,许多学者针对人脸识别进行了研究,提出了许多成果。文献[8]提出了一种基于 MTCNN 的高光谱人脸识别自动控制系统。该系统利用多任务卷积神经网络,提取最优谱带特征,由 Softmax 函数完成高光谱人脸分类识别。高光谱较可见光相比在信息维度方面有了极大的提高,可有效应对光照、姿态等因素产生的干扰。然而高光谱设备成本较高,尤其是应用于大规模在线考试领域,将产生巨大的设备购买及维护成本。文献[9]设计了一种轻量级遮挡人脸识别方法。该方法使用遮挡分割网络提取遮挡信息,掩码生成器根据遮挡信息生成掩码以掩蔽被遮挡损坏的特征,同时嵌入注意力机制提升对空间位置的敏感性,从而缓解遮挡条件下常用人脸识别模型性能急剧下降的问题。然而由于不同的姿态、不均匀的光照和网络通信时延的问题,Zoom 软件图像中存在低分辨率、运动模糊、噪声的问题,现有深度学习模型在复杂环境下的鲁棒性较低。

为改善上述问题,提出了一种基于 Gabor 小波-卷积神经网络-支持向量机(Gabor Wavelet-Convolutional Neural Network-Support Vector machine, GW-CNN-SVM)的人脸识别模型。

## 1 基于 GW-CNN-SVM 人脸识别模型

### 1.1 Gabor 小波及其协方差矩阵

GW<sup>[10]</sup>是一种复小波,其输出由实部和虚部组成。Gabor 小波是一种复小波,可以产生不同方向和尺度的判别信息。这意味着 GW 能够捕捉到图像中的多种特征,包括边缘、纹理等,而这些特征在图像识别中至关重要。通过对图像进行 GW 滤波,可以获得多个子带,每个子带代表一种特定尺度和方向的信息,这使得后续的分类任务可以更加准确和细致。如果图像经 GW 滤波,则将存在  $L \times D$  个子带( $L$  表示尺度的数量,  $D$  表示方向的数量)。具体而言, GW 子带通过将 Gabor 核与图像

$I(x, y)$  卷积如下:

$$g_{l,d}(z) = I(x, y) \cdot \Psi_{l,d}(x, y) \quad (1)$$

$$z = (x, y) \quad (2)$$

式中:  $l = 1, 2, \dots, L$  和  $d = 1, 2, \dots, D$ ;  $g_{l,d}(z)$  的元素为复数。因此,振幅  $M_{l,d}(z)$  和角度  $A_{l,d}(z)$  可通过  $g_{l,d}(z)$  的实部和虚部计算如下:

$$M_{l,d}(z) = \sqrt{\text{Re}_{l,d}^2(z) + \text{Im}_{l,d}^2(z)} \quad (3)$$

$$A_{l,d}(z) = \arctan \left[ \frac{\text{Re}_{l,d}(z)}{\text{Im}_{l,d}(z)} \right] \quad (4)$$

为提高模糊环境下 GW 识别性能,本研究基于协方差矩阵捕捉 GW 子带的相关性。假设子带中的系数为随机变量  $X_k$  的观测值,其中  $k = 1, 2, \dots, K$ , 则可以使用协方差矩阵  $C_{i,j} = \text{Cov}(X_i, X_j)$  表示人脸图像,  $\text{Cov}(\cdot)$  为两个变量  $X_i$  和  $X_j$  的协方差算子。更具体地说,给定 GW 的  $L$  尺度和  $D$  方向分解,可将幅度和角度子带分别定义为以下观测矩阵:

$$\mathbf{Z}_m = [M_{00}, M_{01}, M_{02}, \dots, M_{L \times D}] \quad (5)$$

$$\mathbf{Z}_a = [A_{00}, A_{01}, A_{02}, \dots, A_{L \times D}] \quad (6)$$

式中:  $M_{i,j}$  和  $A_{i,j}$  ( $i = 1, 2, \dots, L, j = 1, 2, \dots, D$ ) 分别为第  $i$  个尺度、第  $j$  个方向上的矢量化幅度/角度子带;  $\mathbf{Z}_m$  和  $\mathbf{Z}_a$  为的幅度和角度子带的观测矩阵,大小均为  $N \times K$ ,  $N$  为子带中系数的数量,  $K = L \times D$  为所有 GW 子带的数量;  $\mathbf{Z}_m$  和  $\mathbf{Z}_a$  的每列为 GW 的矢量化子带,表示为变量的观测值,每行为  $K$  维样本向量,表示为  $\{z_i\}, i = 1, 2, \dots, N$ 。根据 COV 算子的定义, GW 域中的协方差矩阵由下式给出:

$$\mathbf{C} = \frac{1}{N-1} \sum_{i=1}^N (z_i - \mu)^\top (z_i - \mu) \quad (7)$$

$$\mu = \frac{1}{N} \sum_{i=1}^N z_i \quad (8)$$

式中:  $\mu$  为采样点的平均值;  $\mathbf{C}$  为协方差矩阵,大小为  $K \times K$ 。

GW 分解通过多尺度和多方向的信息提取以及协方差矩阵对子带相关性的捕捉,为图像分类提供了丰富的特征信息。而 CNN 则能够进一步挖掘这些特征信息中的内在规律和模式,从而实现图像更加准确和鲁棒的分类。这种结合 GW 分解和 CNN 的方法在模糊环境下的人脸识别等应用中具有显著的优势。

### 1.2 基于 CNN 的融合协方差矩阵特征提取

为了提高模糊环境下人脸识别性能,将 GW 的幅度和角度子带的协方差矩阵(即融合协方差矩阵)带入 CNN 模型进行特征提取。

GW 域中构建融合协方差矩阵过程的框架图如图 1 所示。首先由 GW 分解人脸图像块,然后将这些子带定义为观测矩阵  $Z_m$  和  $Z_a$ 。通过使用协方差算子,生成两个协方差矩阵  $C_m$  和  $C_a$ 。接着,幅度子带的协方差矩阵的上三角部分与角度子

带的下三角部分进行级联,具体计算如下:

$$C_F = t_{riu}(C_m) + t_{riu}(C_a, -1) \tag{9}$$

式中:  $t_{riu}(C_m)$  为主对角线的上三角矩阵  $C_m$  的级联操作;  $t_{riu}(C_a, -1)$  为主对角线下方的下三角矩阵  $C_a$  的级联操作。

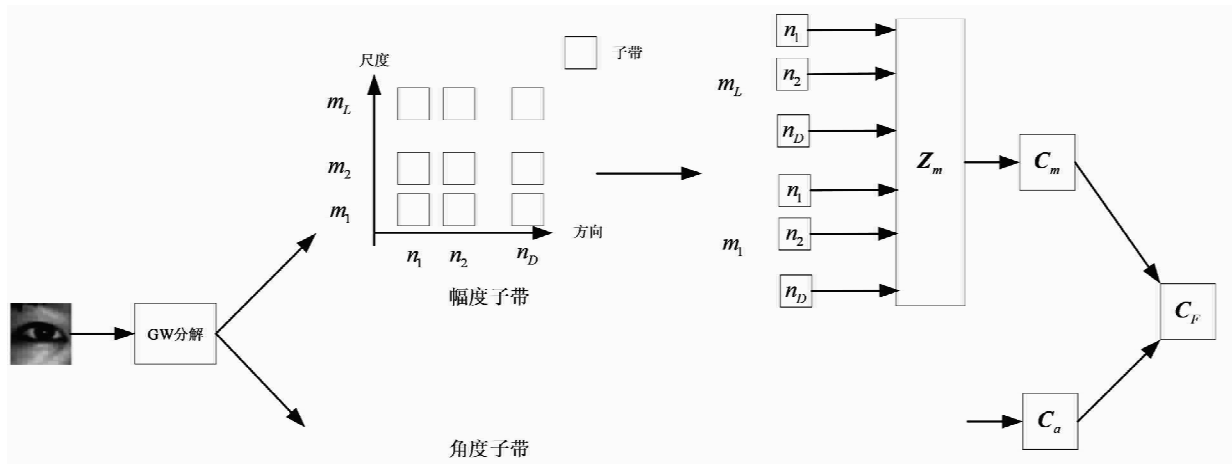


图 1 GW 域中构建融合协方差矩阵过程的框架图

由于融合的协方差矩阵包含了幅度和角度特征,因此可包含更多的面部特征,从而可在一定程度提升 CNN 模型训练性能。基于 CNN 及融合协方差矩阵的特征提取框架图如图 2 所示。在特征提取阶段,使用经过训练的 CNN 将融合的协方差矩阵转换为向量。人脸图像首先被划分为不重叠的  $M \times M$  个块,每个块将映射到融合的协方差矩阵中,具体定义为:

$$C_{F_i} = F_{COV} | G_W(B_i) \tag{10}$$

式中:  $B_i$  为第  $i$  个局部块;  $F_{COV}$  为协方差算子;  $G_W(\cdot)$  为局部块的 GW 分解。通过逐个映射,局部块将成为图像中相应位置的局部矩阵。最后,人脸图像将被转换为一个较大的矩阵(称为特征矩阵),该矩阵由许多较小的局部矩阵  $C_{F_i}$  组成。特征矩阵的大小取决于块的数量和 GW 分解大小(即 GW 子带的数量)。

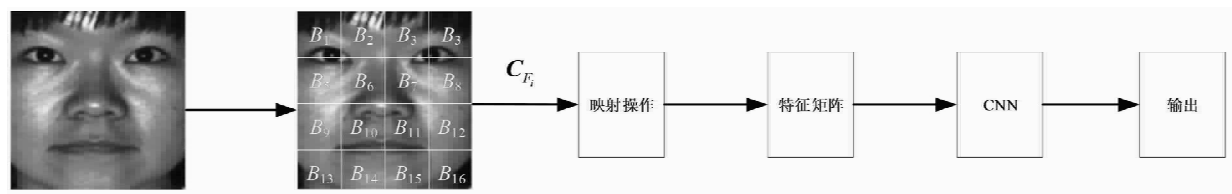


图 2 基于 CNN 及融合协方差矩阵的特征提取框架图

使用式(10)将数据集中的所有人脸图像转换为 GW 特征矩阵。在学习阶段,利用这些特征矩阵训练 CNN 模型,并实现面部图像识别。接下来,关键任务是设计能够有效地执行特征矩阵转换的 CNN 模型。本文设计了一个 8 层的映射 CNN 网络,其结构如图 3 所示。映射 CNN 网络包括输入层、2 个卷积层和 2 个最大池化层,以及 3 个全连接层。训练时使用交叉熵损失作为网络的损失函数。最终的输出矢量为一个 200 维的面部矢量特征。

全连接层在神经网络模型中可以进行分类,负责将学到的“分布式特征表示”映射到样本标记空间,从而进行分类。然而,在某些情况下,特别是在需要更高精度或更复杂的分类任务时,仅仅使用全连接层可能不是最优的。SVM 是一种经过广泛验证和测试的分类算法,它在处理高维数据和复杂分类任务时表现出色。通过结合 CNN 提取的特征和 SVM 的分类能力,可以获得比仅使用全连接层更高的分类准确性。

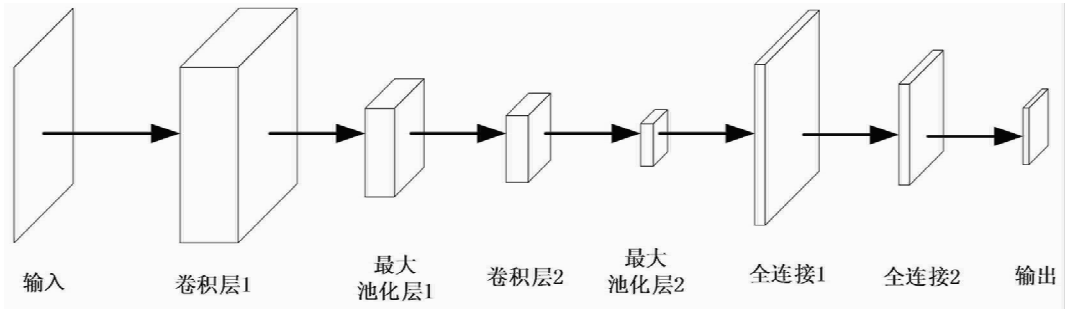


图 3 映射 CNN 网络结构图

### 1.3 SVM 面部分类器

支持向量机<sup>[11]</sup> (Support Vector Machines, SVM)是一种多类分类和回归的强大工具。SVM 估计不同类之间的最优分离超平面,同时最大化超平面和类的最近点之间的阈值从而实现区分不同的类。本研究中基于 SVM 实现面部矢量特征的分类。

将上述映射 CNN 网络获取的面部矢量特征作为支持向量机的输入,输入集合为  $X = \{x_1, x_2, \dots, x_n\}$ , 其中  $x_i$  表示样本特征,  $n$  代表的是样本特征的数量。

建立支持向量机模型如下:

$$f(x) = \sum_{i=1}^n \mathbf{W}\varphi_i + b \quad (11)$$

式中:  $b$  为估计偏差;向量  $\mathbf{W}$  由输入样本对应的权值构成;  $\varphi_i$  为非线性映射。

为了找到最优的分类面,需要使得分类面最小化。因此,这个问题被转化为求解最优化超平面的问题,也就是要求解式(12):

$$\min_{\vec{\omega}} = \frac{1}{2} \vec{\omega}^T \vec{\omega} + C \sum_{i=1}^N \xi(\vec{\omega}; \vec{\alpha}_i, \vec{\beta}_i) \quad (12)$$

式中:  $\vec{\alpha}_i$  和  $\vec{\beta}_i$  为包含一组面部矢量特征及其标签的训练对;参数  $\vec{\omega}$  和  $\varepsilon$  可确定线性超平面;  $C$  为正则化参数,用来平衡误差和最小化权重范数;  $N$  为待分类的个数。通过构造最优化超平面,可以得到一个满足条件的超平面  $f(x) = \mathbf{W} \cdot x + b$ 。

Hinge 损失函数是 SVM 中用来衡量分类错误的一种方式, Hinge 损失函数  $\xi(\vec{\omega}; \vec{\alpha}_i, \vec{\beta}_i)$  表示为:

$$\xi(\vec{\omega}; \vec{\alpha}_i, \vec{\beta}_i) = \max \{1 - \vec{\beta}_i [\vec{\omega}^T \vec{\varphi}(\vec{\alpha}_i) + \varepsilon], 0\}^2 \quad (13)$$

式中:参数  $\vec{\omega}$  和  $\varepsilon$  可确定线性超平面;映射函

数  $\vec{\varphi}(\cdot)$  将面部矢量特征  $\vec{\alpha}_i$  映射到高维空间。

## 2 实验与分析

### 2.1 数据集

本研究所用数据集来自 Zoom 软件提供的在线考试实时记录视频。Zoom 软件示意图如图 4 所示,在启用图库视图模式时,默认情况下每次屏幕显示 25 名(5 行×5 列)考生。通过设置,Zoom 软件还可以支持每次屏幕 49 个(7 行×7 列)参与者。每个录制视频都能够将考生的所有面部特征存储在 .mp4 格式的本地视频文件里。此外,Zoom 为每个考生提供了一个个人标签,以标记不同参与者。最终本研究收集了共 200 人 1361 幅原始图像。经过预处理,去除严重失真图像,最终数据集共包含 1303 幅图像。接着,按照 7:3 将数据集分为训练集和测试集。本文选用 Python 的 scikit-learn 库作为训练平台,其中参数设定如表 1 所示。

表 1 参数设定

参数	值
惩罚参数	1.0
核函数	RBF
核参数	0.1
损失函数	Hinge Loss
正则化类型	L2
叉验证折数	5
迭代次数	100

### 2.2 结果与分析

#### 2.2.1 GW 分解分析

为验证 GW 分解对分类效果的提升作用,进行分类实验,结果如图 5 所示。

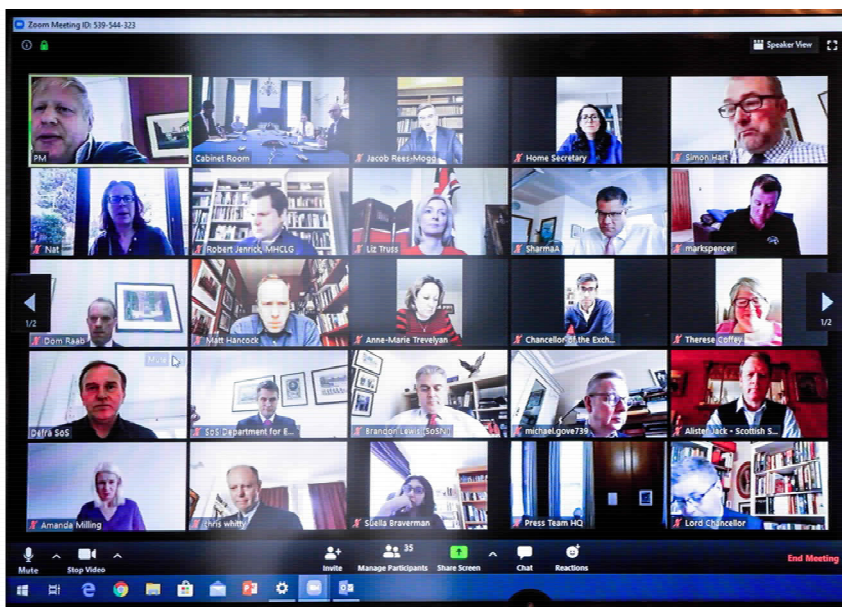


图4 Zoom软件示意图

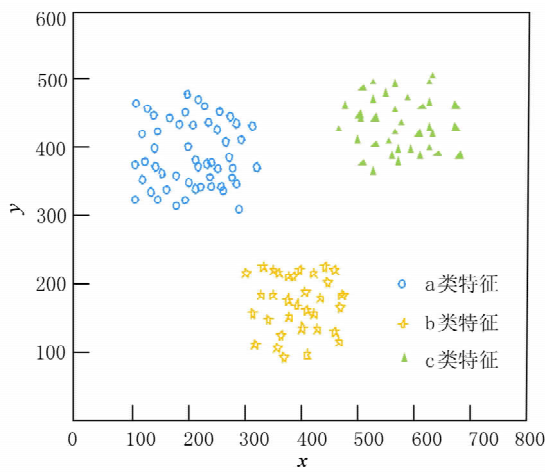


图5 GW分解结果

根据图5可以看出,不同类型的数据被有效区分,表明GW分解可能有效地提取了与分类任务相关的重要特征或信息,使得分类器能够基于这些信息进行准确的分类。

### 2.2.2 收敛曲线分析

为验证所提模型的应用性能,使用训练集对模型进行训练,并记录训练误差。与主流PCANet、VGGFace、ResNet50等模型进行对比。得到误差收敛曲线如图6所示。

根据图6可以看出,所提模型能够快速降低训练误差,当迭代次数达到6次时,误差达到了最小值0.01,而PCANet模型在迭代次数达到20次时,误差才达到了最小值0.05,VGGFace模型在迭代

次数达到17次时,误差才达到了最小值0.03,ResNet50模型在迭代次数达到27次时,误差才达到了最小值0.04。由此表明所提模型在达到误差最小值时所需的迭代次数远少于其他三个模型。这表明所提模型具有更快的收敛速度,能够在较少的迭代内收敛到更优的解。

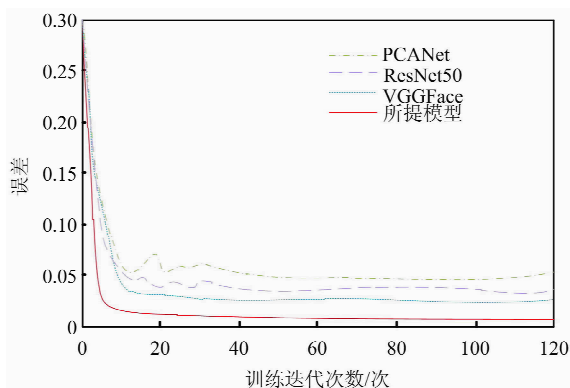


图6 收敛性能对比

### 2.2.3 不同光照准确率分析测试

选择13:00、14:30、16:30与18:00等4个时间段,收集人脸数据图像。进行光照分析测试,实验时对比指标选取准确率。准确率计算如下:

$$f_{acc} = \frac{n_{cor}}{n_{tol}} \quad (14)$$

式中: $f_{acc}$ 为准确率指标; $n_{cor}$ 为正确分类的样本个数; $n_{tol}$ 为总样本个数。不同光照准确率结果如表2所示。

表 2 不同光照准确率结果

时间 /h	PCANet /%	VGGFace /%	ResNet50 /%	所提模型 /%
13:00	82	81	88	98
14:30	88	84	89	98
16:30	84	85	90	96
18:00	85	86	92	97

根据上述结果可以看出,所提模型的准确率始终高于 95%,而对比方法的准确率在 81%~92% 之间,由此表明所提模型采用了更有效的特征提取方法,能够捕捉到人脸图像中的关键特征,从而提高了识别的准确率。这些特征可能对于光照变化具有更强的鲁棒性。

#### 2.2.4 不同分辨率准确率分析测试

为了评估所提 GW-CNN-SVM 模型对低分辨率下人脸图像的鲁棒性,试验时将图像调整为不同的尺寸:128×128,64×64,32×32,16×16,8×8,并与主流 PCANet、VGGFace、ResNet50 等模型进行对比。不同模型不同分辨率下对比结果如图 7 所示。

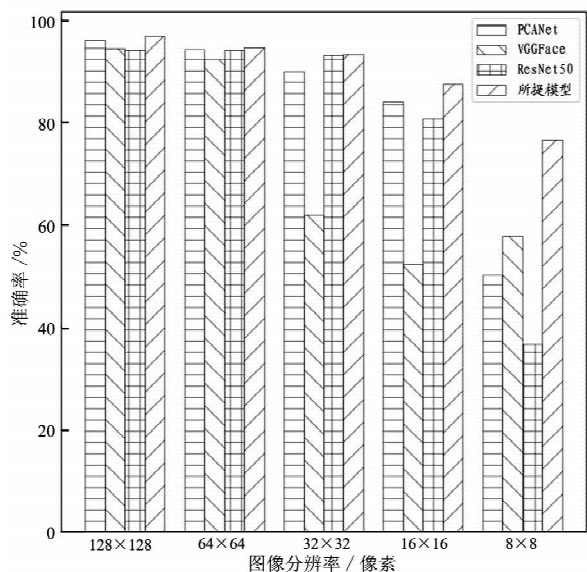


图 7 不同模型不同分辨率下对比结果

可以看出,所提 GW-CNN-SVM 模型的识别准确率在 128×128 像素下为 96.67%,明显高于 PCANet、VGGFace、ResNet50 等模型。同时,PCANet、VGGFace、ResNet50 模型在高分辨率下表现出良好的性能;然而,这些模型在低分辨率下识别准确率下降非常严重,尤其是 8×8 低分辨率时,所有模型准确率均低于 60%。与之相反,所提

GW-CNN-SVM 模型准确率约为 76.5%。试验结果验证了所提模型在低分辨率下具有较好的识别性能。

#### 2.2.5 运动模糊和噪声分析测试

实验还评估了运动模糊和噪声环境下人脸图像识别性能。为了进行运动模糊,在所有测试集面部图像中分别添加了以 10°角跨越 10 个像素的线性运动和以 30°角跨越 30 个像素的直线运动。同时,对测试集中原始图像分别添加了具有零均值和 0.01 方差以及具有零均值噪声和 0.03 方差的高斯白噪声。运动模糊和噪声测试数据部分展示效果如图 8 所示。

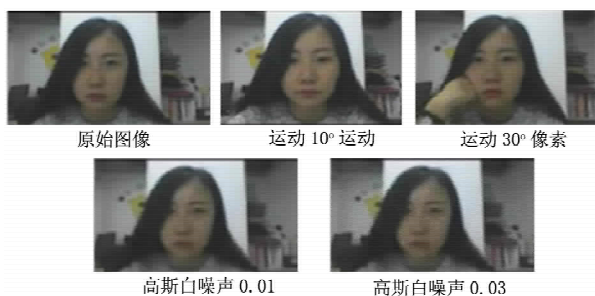


图 8 运动模糊和噪声测试数据部分展示效果

不同模型运动模糊和噪声下对比结果如图 9 所示。

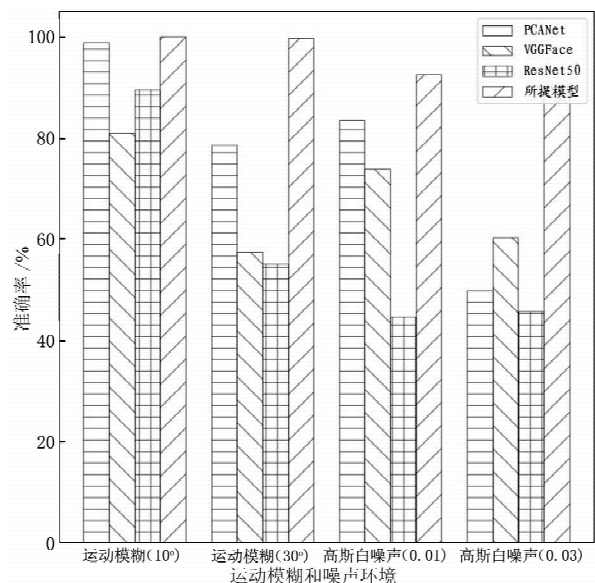


图 9 不同模型运动模糊和噪声下对比结果

可以看出,与 PCANet、VGGFace、ResNet50 模型相比,所提 GW-CNN-SVM 模型在运动模糊和噪声环境下识别性能更优。例如对于运动模糊

(30°)以及高斯白噪声(0.03),PCANet、VGG-Face、ResNet50准确率基本下降至60%以下,然而所提模型识别准确率均在85%以上。试验结果验证了所提GW-CNN-SVM模型在运动模糊和噪声环境下具有更高的鲁棒性。

### 3 结 论

对在线考试场景下具有模糊、噪声的人脸识别方法进行研究,提出了基于GW-CNN-SVM的人脸识别模型。首先,基于GW将人脸图像分解为幅度和角度子带的协方差矩阵。其次,为了提高GW特征提取的质量,采用改进的CNN模型学习GW协方差矩阵特征,从而提取面部矢量特征,并生成特征向量表示。最后,将生成特征向量表示作为支持向量机的输入输出识别结果。通过试验仿真,验证了所提模型对低分辨率、运动模糊、噪声环境具备较强的鲁棒性,可为在线考试时人脸监控与识别领域的发展提供一定借鉴作用。

### 参考文献

[1] 涂海亮. 基于远程会议的线上交互式教学系统设计[J]. 信息

与电脑(理论版),2022,34(5):82-84.

- [2] 刘信鑫,韩光,杨海根,等. 在线教学视觉隐私保护学生专注度监测研究[J]. 智能计算机与应用,2022,12(11):1-8+17.
- [3] 许声宏. 在线教学中的人工智能应用初探[J]. 西南师范大学学报(自然科学版),2022,47(9):111-115.
- [4] 郭隳. 国外高校在线考试作弊防范策略研究[J]. 海外英语,2023(7):92-95.
- [5] 颜慧. 人工智能视觉分析技术在课堂教学评价中的应用探析[J]. 电脑知识与技术,2023,19(4):29-32.
- [6] 李腊全,郑申海. 计算机视觉在大学生科研训练计划中的作用及实践探索[J]. 中国多媒体与网络教学学报(上旬刊),2022(2):57-60.
- [7] 姚静姝. 计算机视觉在人脸识别领域中的应用[J]. 电子技术与软件工程,2020(17):128-129.
- [8] 梁莹. 基于MTCNN的高光谱人脸识别自动控制系统[J]. 自动化与仪表,2023,38(6):10-13+28.
- [9] 田爽,黄威,代毅. 基于特征掩码的轻量级遮挡人脸识别方法[J]. 内蒙古大学学报(自然科学版),2023,54(3):272-280.
- [10] 王康毅,邵苏杰. 基于Gabor变换的多角度人脸表情识别方法[J]. 计算机仿真,2024,41(4):233-236+526.
- [11] 李昆仑,彭小华,熊婷. 基于支持向量机的动态人脸图像微表情识别[J]. 计算机仿真,2023,40(12):331-336.