

基于深度卷积神经网络的蝇类面部识别

陈彦彤, 陈伟楠, 张献中, 李雨阳, 王俊生

引用本文:

陈彦彤, 陈伟楠, 张献中, 等. 基于深度卷积神经网络的蝇类面部识别[J]. *光学精密工程*, 2020, 28(7): 1558–1567.

CHEN Yan-tong, CHEN Wei-nan, ZHANG Xian-zhong, et al. Fly facial recognition based on deep convolutional neural network[J]. *Optics and Precision Engineering*, 2020, 28(7): 1558–1567.

在线阅读 View online: <https://doi.org/10.37188/OPE.20202807.1558>

您可能感兴趣的其他文章

Articles you may be interested in

多模深度卷积神经网络应用于视频表情识别

Video-based facial expression recognition using multimodal deep convolutional neural networks

光学精密工程. 2019, 27(4): 963–970 <https://doi.org/10.3788/OPE.20192704.0963>

构建多尺度深度卷积神经网络行为识别模型

Action recognition model construction based on multi-scale deep convolution neural network

光学精密工程. 2017, 25(3): 799–805 <https://doi.org/10.3788/OPE.20172503.0799>

采用深度级联卷积神经网络的三维点云识别与分割

Recognition and segmentation of three-dimensional point cloud based on deep cascade convolutional neural network

光学精密工程. 2020, 28(5): 1187–1199 <https://doi.org/10.3788/OPE.20202805.1187>

基于卷积神经网络的光学遥感图像检索

Optical remote sensing image retrieval based on convolutional neural networks

光学精密工程. 2018, 26(1): 200–207 <https://doi.org/10.3788/OPE.20182601.0200>

基于改进深度卷积网络的铁路入侵行人分类算法

Pedestrian intruding railway clearance classification algorithm based on improved deep convolutional network

光学精密工程. 2018, 26(12): 3040–3050 <https://doi.org/10.3788/OPE.20182612.3040>

文章编号 1004-924X(2020)07-1558-10

基于深度卷积神经网络的蝇类面部识别

陈彦彤, 陈伟楠, 张献中, 李雨阳, 王俊生*

(大连海事大学 信息科学技术学院, 辽宁 大连 116026)

摘要:针对蝇类昆虫物种繁多、特征复杂等因素,导致蝇类识别准确率低、耗时较长等问题。本文借鉴深度学习方法中的人脸识别算法,提出一种基于深度卷积神经网络的蝇类面部识别方法。首先,在图像对齐过程中,使用多任务卷积神经网络并进行优化即应用深度可分离卷积减少计算参数,缩短图像预处理时间。其次,应用轮廓特征粗提取和具体部位特征细提取相结合的方式提取更加丰富的特征信息:即使用卷积池化粗提取出图像的轮廓特征值;同时,使用 Inception-ResNet 网络、Reduction 网络细提取出具体部位特征值。最终在网络训练时,结合上述方法使得提取到的特征信息更加精确全面。实验表明,所提方法的准确率达到 94.03%,相较于其他网络训练方法,该方法在保证较高准确率的情况下提升计算效率。

关键词:蝇类面部识别;深度卷积神经网络;多任务卷积神经网络;Inception-ResNet 网络;Reduction 网络

中图分类号:TP306 **文献标识码:**A **doi:**10.37188/OPE.20202807.1558

Fly facial recognition based on deep convolutional neural network

CHEN Yan-tong, CHEN Wei-nan, ZHANG Xian-zhong, LI Yu-yang, WANG Jun-sheng*

(College of information science and technology,

Dalian Maritime University, Dalian 116026, China)

* Corresponding author, E-mail: wangjsh@dlmu.edu.cn

Abstract: Given the large number of species of flies and their individual complex characteristics, recognizing a particular type of fly has proved to be time consuming and, for the most part, inaccurate. In this paper, a method for the facial recognition of a fly using deep learning technologies was proposed, specifically a Convolutional Neural Network (CNN), and its related face recognition algorithms. Initially, a multi-task convolutional neural network was utilized and optimized for the image alignment process. Thus, depth-wise separable convolutions were applied to reduce the number of calculation parameters as well as the image preprocessing time. Next, we combined the rough extraction of contour features and fine extraction of specific parts to derive more abundant feature information. The convolution and pooling layers were harnessed to elicit contour eigenvalues of the image, while Inception-ResNet and Reduction networks were administered simultaneously to obtain eigenvalues of specific parts. Finally, the above methods were coalesced to enhance the accuracy and comprehensibility of the resultant feature information for network training. Experimental results show that the mean average precision of the proposed method is 94.03%. When compared with other network training methods, this method not only improves the computational efficiency but also

收稿日期:2020-01-10;**修订日期:**2020-03-08.

基金项目:国家自然科学基金资助项目(No. 61901081);中央高校基本科研业务费专项资助(No. 3132020199)

ensures high accuracy.

Key words: face recognition of fly; deep convolutional neural network; multi-task convolutional neural network; Inception-ResNet; Reduction network

1 引言

随着世界各国之间的贸易交流日益频繁,旅客携带的外来昆虫传入我国境内几率增大,甚至会引发生态环境被破坏等问题。例如 2016 年 1 月,一名入境的澳门旅客携带的水果中存在地中海实蝇,该昆虫繁殖能力非常强,一只地中海实蝇雌虫经三代繁殖即可达 215 亿只,将会对农业生产造成极大经济损失;2018 年 10 月,越南芽庄入境航班旅客携带的芒果果肉中含有的番石榴果实蝇幼虫,导致我国当地果园大面积失收。因此,若无法对外来蝇类昆虫作出及时监测和识别,一旦发生某种昆虫过度繁衍,将造成不可估量的损失。由此可见,对蝇类昆虫物种的高效识别研究刻不容缓。

基于蝇类昆虫识别的传统方法有:基于颜色特征^[1]的蝇类昆虫识别,采集图像颜色直方图信息并将所有颜色直方图合并成特征向量作为蝇类昆虫的颜色特征。颜色特征不受图像旋转和平移变化的影响,但没有表达出颜色空间分布的信息,影响识别的准确性。基于纹理特征^[2]的蝇类昆虫识别,将整个图像形成一个纹理特征值矩阵,然后将这个矩阵转换成影像。基于空间特征^[3]的蝇类昆虫识别,将蝇类昆虫图像投影到特征子空间,利用测度计算未识别图像系数与训练样本系数之间的相似度,应用 K 最近邻(K-Nearest Neighbor, KNN)算法进行蝇类昆虫识别。虽然传统方法能够实现蝇类昆虫识别,但是传统方法提取的各类特征有局限性,识别精度较低。

随着人工智能的高速发展,深度学习在计算机视觉方向成为最流行的技术方法,尤其是在人脸识别方面取得了显著成果。人脸识别是将静态图像中检测出的人脸图像与数据库中的人脸图像进行对比,从中找出与之匹配人脸的过程,以达到身份识别与鉴定的目的。蝇类昆虫相似度较高,传统方法提取的特征不易区分导致识别困难。但

是通过分析发现,从不同蝇类昆虫面部提取的特征向量区别较大,因此可以借鉴深度学习中人脸识别算法对蝇类面部进行识别。

在基于深度学习的人脸识别方法中,Szegedy 等提出的 GoogLeNet^[4]采用多种尺寸的卷积核构成 Inception 结构,在保持网络稀疏性的前提下提高计算效率,却存在着网络层数较浅的问题,无法提取精确的特征向量。Bong-Nam Kang 等于 2017 年提出了深度卷积神经网络^[5](Deep Convolutional Neural Networks, DCNNs),由堆叠的多尺度卷积层块(Multi-scale Convolution Layer Blocks, MCLBs)构成,而 MCLBs 对 Inception 结构进行改进,加深卷积层数以呈现多尺度抽象,但是当 Inception 网络趋向于非常深时,会容易出现梯度消失的问题。He 等提出的残差网络(Residual Network, ResNet)采用快捷连接的方式实现特征跨层传递^[6],解决深度网络梯度消失这一问题,但在较深的网络层下计算效率仍然较低,同时还存在着特征提取不够精细的问题。

基于上述问题,本文在 DCNNs 网络的基础上,结合 Inception-ResNet 网络、Reduction 网络、轮廓和具体部位特征细提取相融合这三个部分,从而提出基于深度卷积神经网络的蝇类面部识别研究方法,并命名为蝇类面部卷积神经网络(Fly Facial Convolutional Neural Network, FFCNN)。通过改进卷积神经网络结构来提高蝇类面部识别的准确率,提高计算效率。

2 DCNNs 网络

DCNNs 网络是由堆叠的多尺度卷积层块构成,以呈现多尺度抽象。为了训练 DCNNs 网络,使用由标准脸图像、具有相同身份的人脸图像和不同身份的人脸图像组成的训练集,并应用由三重、成对、softmax 求和构成的损失函数。

MCLBs 是 DCNNs 网络主要的组成部分,分

别由 $1 \times 1, 3 \times 3, 5 \times 5$ 卷积和 3×3 最大池化层组成。由于计算时间在视觉任务中很重要,因此使用 1×1 卷积作为降维,这使得 DCNNs 网络能够足够快地训练。将一系列卷积输出串接为一个输出向量,与简单卷积层序列的输出相比,该输出向量中的抽象级别随着层级别的增加而增加。MCLBs 中的所有卷积层使用批处理规范化 (Batch Normalization, BN) 和非线性激活函数 (Rectified Linear Units, ReLU)。

在新型训练集上即三重人脸方面,采用 $T = (I_R, I_P, I_N)$ 即 (标准脸, 相同身份人脸图像, 不同身份人脸图像) 作为训练数据集,通过损失函数使具有相同身份的面部之间距离最小化,并使具有不同身份的面部之间距离最大化:

$$L_{\text{total}} = L_{\text{triplet}} + L_{\text{pairs}} + L_{\text{softmax}}, \quad (1)$$

其中总的损失函数如式(1)所示,即总的损失函数 L_{total} 由三重损失函数 L_{triplet} 和成对损失函数 L_{pairs} 和 softmax 损失函数^[7] L_{softmax} 相加求得:

$$L_{\text{triplet}} = \sum_{T} \max \left(0, 1 - \frac{\|F(I_R) - F(I_N)\|_2}{\|F(I_R) - F(I_P)\|_2 + m} \right). \quad (2)$$

三重损失函数如式(2)所示,即 $F(I_R)$ 是 I_R 经过 DCNNs 映射之后的输出,同理 $F(I_P)$ 是 I_P 经过 DCNNs 映射之后的输出, $F(I_N)$ 是 I_N 经过 DCNNs 映射之后的输出。 $\|F(I_R) - F(I_N)\|_2$ 是标准脸与不同身份的人脸图像的距离, $\|F(I_R) - F(I_P)\|_2$ 是标准脸与相同身份的人脸图像的距离。 m 是一个边缘限制比例值,是一个常数:

$$L_{\text{pairs}} = \sum_{(I_R, I_P) \in T} \|F(I_R) - F(I_P)\|_2^2. \quad (3)$$

成对损失函数由式(3)所示, I_R 和 I_P 均表示三重人脸数据集 T 中的标准脸和相同身份的人脸图像,成对损失函数 L_{pairs} 即为标准脸与相同身份的人脸图像距离之和,使得标准脸图像与相同身份的人脸图像之间距离最小。

而将 DCNNs 网络应用到蝇类面部发现,经过 MCLBs 等卷积层图像信息流失较大,提取不到较为精确的特征向量,无法达到较高的识别精度,并且耗时较长。为了提升面部识别的准确率以及降低计算复杂度,可通过改进卷积神经网络^[7]的方式来提升网络的性能。

3 基于深度卷积神经网络的蝇类面部识别

3.1 蝇类面部对齐

由于深度学习方法是将对齐后的面部图像通过深层网络提取面部特征,而现实问题是并不能全部捕捉到对齐的面部照片,因此识别精度不够高。针对这个问题采用多任务卷积神经网络 (Multi-task Convolutional Neural Network, MTCNN) 网络,而 MTCNN 网络主要由 P-Net、R-Net 和 O-Net 三个部分组成^[9],使用全卷积的 P-Net 在多尺度的待检图像上生成候选框,接着通过 R-Net 和 O-Net 来过滤,其总的损失函数如式(4)所示:

$$\min \sum_{i=1}^N \sum_{j \in \{\text{det}, \text{box}, \text{landmark}\}} \alpha_j \beta_i^j L_i^j, \quad (4)$$

其中: N 是训练样本的总数, α_j 表示各个损失所占的权重,在 P-Net 和 R-Net 中,设置 $\alpha_{\text{det}} = 1$, $\alpha_{\text{box}} = 0.5$, $\alpha_{\text{landmark}} = 0.5$,在 O-Net 中,设置 $\alpha_{\text{det}} = 1$, $\alpha_{\text{box}} = 0.5$, $\alpha_{\text{landmark}} = 1$ 。 L_i^{det} , L_i^{box} , L_i^{landmark} 分别代表面部检测损失函数,框回归损失函数和关键点损失函数, $\beta_i \in \{0, 1\}$ 表示样本类型指示符。因此,通过调试至 MTCNN 网络损失函数值最小从而检测面部在图像中的准确位置。但是, MTCNN 网络也存在很多问题,例如当图片越大 P-net 耗时越长;面部数量越多 O-net 和 R-net 耗时越长。针对这些问题,需要优化网络结构,在保证精度的前提下,尽可能减少计算量。本文提出应用深度可分离卷积 (Depthwise Separable Convolutions),把标准卷积分解成深度卷积 (Depthwise Convolution) 和逐点卷积 (Pointwise Convolution) 的方式降低计算量^[10]。

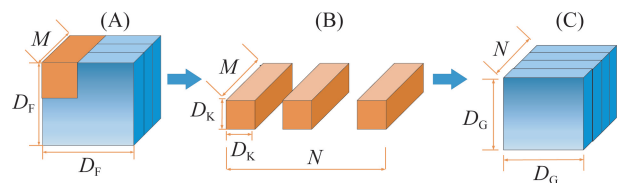


图 1 标准卷积

Fig. 1 Standard convolution

标准卷积如图 1 所示,输入通道数为 M ,输出通道数为 N ,其中(A)表示输入特征映射 F 尺寸为 (D_F, D_F, M) , (B)表示标准卷积 K 尺寸为 (D_K, D_K, M, N) , (C)表示输出特征映射 G 尺寸为 (D_G, D_G, N) ,则标准卷积的计算量为 $D_K \times D_K \times M \times N \times D_F \times D_F$ 。

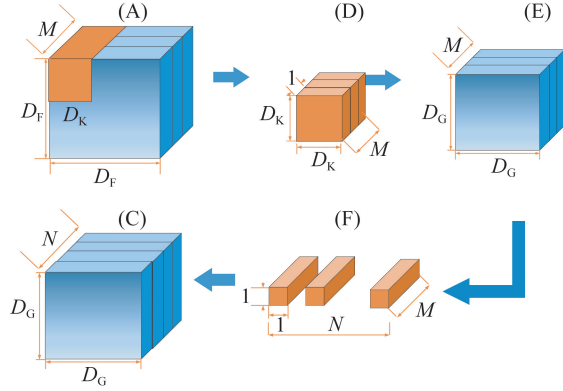


图 2 深度卷积与逐点卷积

Fig. 2 Depthwise convolution and pointwise convolution

深度可分离卷积由深度卷积和逐点卷积构成,如图 2 所示,由标准卷积拆分后可得。深度卷积对输入图像的每个通道进行卷积,从而达到滤波的功能。(D)表示深度卷积尺寸 $(D_K, D_K, 1, M)$,其输出特征用(E)表示为 (D_G, D_G, M) ,深度卷积计算量为 $D_K \times D_K \times M \times D_F \times D_F$ 。逐点卷积主要通过转换通道,使用 1×1 卷积将这些特征合并起来,(F)表示逐点卷积尺寸为 $(1, 1, M, N)$,得到最终输出特征仍用(C)表示为 (D_G, D_G, N) 。逐点卷积计算量为 $M \times N \times D_F \times D_F$ 。

深度可分离卷积计算量为深度卷积和逐点卷积之和,深度可分离卷积与标准卷积的计算量之比如式(5)所示:

$$\frac{D_K \times D_K \times M \times D_F \times D_F + M \times N \times D_F \times D_F}{D_K \times D_K \times M \times N \times D_F \times D_F} = \frac{1}{N} + \frac{1}{D_K^2}, \quad (5)$$

其中 N 和 D_K 的数值均大于 1,由此可得,深度可分离卷积的计算量比标准卷积的计算量要小。并且在 I/O 效率和性能不变的情况下,计算量有着明显下降。

在 MTCNN 网络中,通过框选定位以及识别出 5 个特征点位置后,再经过仿射变换,即二维坐标到二维坐标之间的线性变换,保持二维图形的“平直性”和“平行性”,如公式(6)所示:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & t_x \\ a_3 & a_4 & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}, \quad (6)$$

其中: (t_x, t_y) 表示平移量,而参数 a_i 则反映了图像旋转,缩放等变化。将参数 $t_x, t_y, a_i (i=1 \sim 4)$ 计算出,即可得到两幅图像的坐标变换关系,并实现了面面对齐。裁剪缩放到分辨率为 95×95 的标准面部图像,如图 3 所示。

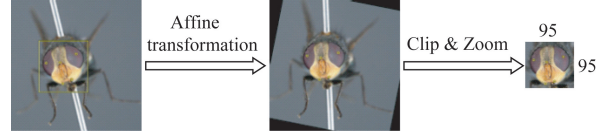


图 3 蝇类面面对齐

Fig. 3 Fly species facial alignment

3.2 蝇类面部的 FFCNN 网络

本文在 DCNNs 网络的基础上提出一种新的网络结构 FFCNN 网络,即通过 Inception-Resnet 网络、Reduction 网络,轮廓与具体部位粗细提取等组合来构建如图 4 所示。所有卷积层和完全连接层使用 ReLU 非线性激活函数和 BN 批处理规范化。平均池化取每个特征图的平均矢量,求出空间信息。由于空间信息由求平均值得出,所以平均池化具有平移不变性。FFCNN 网络的输入是蝇类昆虫 3 种面部之一,即 I_R (标准面部), I_P (相同昆虫的其他面部图像)或 I_N (不同昆虫的面部图像)。将提取到的具体部位特征放入损失函数中,使得同种类别昆虫距离更加接近,不同种类别昆虫距离更加疏远。并将轮廓特征与具体部位特征向量相融合为一个特征向量,通过验证其准确率判断识别效果。

首先输入预处理后蝇类面部图像,其次进入轮廓特征粗提取部分如框图 I 所示。该部分通过大量卷积池化能够粗提取出图像的轮廓特征值,并放置于全连接层中,使用辅助分类器判断类别,从而达到减少图像信息损失,使得识别更加全面的目的。同时,网络提取到蝇类面部的具体部位特征如框图 II 所示。此部分主要经过 Inception-ResNet 网络,在减少图像梯度散失的同时减少计算参数;以及 Reduction 网络在减少图像信息损失的同时减小图像尺寸,使得提取特征值变得更加精确化。并且在具体部位特征细提取后,通过

Bong-Nam Kang 等提出的损失函数算法,可以让同种类别的昆虫距离更加接近,不同种类昆虫距离更加疏远。

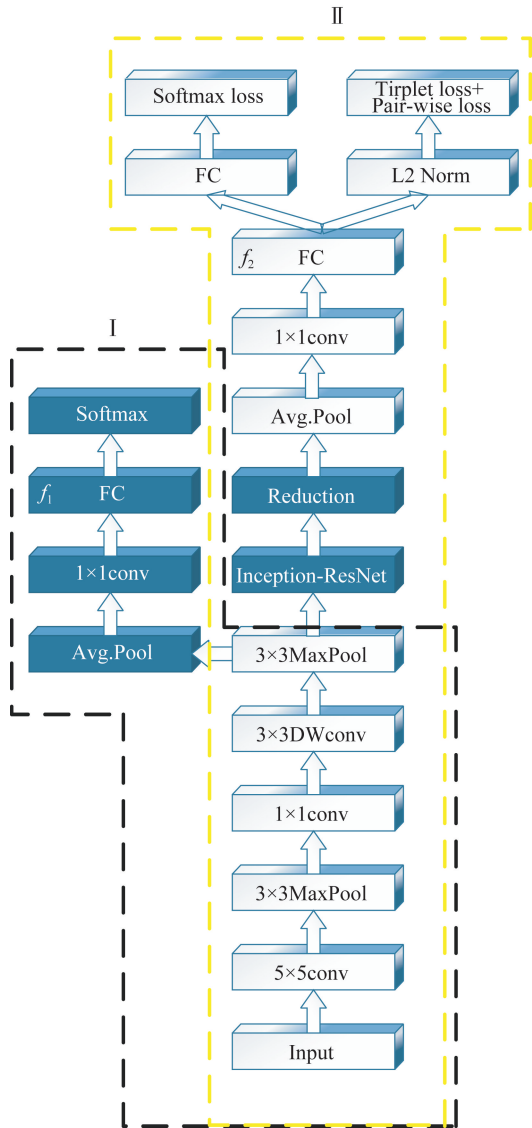


图 4 网络主框图

Fig. 4 Main network diagram

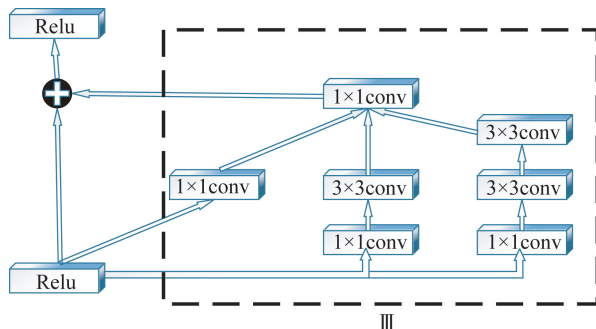


图 5 Inception-Resnet 框图

Fig. 5 Inception-Resnet diagram

为防止训练过程中的梯度消失问题和提高计算效率,加入 Inception-ResNet 网络如图 5 所示。将 ResNet 残差网络和 Inception 网络相结合,其中框图 III 表示 Inception 网络,并用残差连接替代滤波器的连结,将图 III 作为一个整体和剩余部分连接构成 ResNet 残差网络^[11]。首先用 1×1 卷积升高维度,然后用 3×3 卷积以最小卷积核最大程度的提取图像信息,最后再用 1×1 卷积进行降维。可以简化学习目标和难度,在防止梯度散失的同时,更加快速提取特征向量。

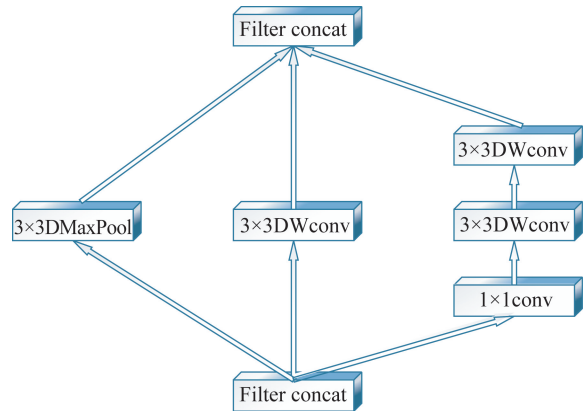


图 6 Reduction 框图

Fig. 6 Reduction diagram

图 6 的 Reduction 网络可以减小图像块尺寸,即每经过一个 Reduction 网络,特征图的尺寸就会显著减少。由于池化层在下采样的过程会损失过多信息,所以文中采用 Reduction 网络替代部分池化层来减小特征图的尺寸。其中 filter concat 层是一种累积和的叠加,把多个特征直接加在一起; 1×1 卷积用来减少特征维度;并且此网络加入了 3×3 深度可分离卷积可以在损失精度不多的情况下大幅度降低参数量。

在图 4 的网络主框图中,将给定面部图像的特征分别用 f_1, f_2 表示,框图 I 中的全连接层提取特征向量为 f_1 ;框图 II 中倒数第 2 个全连接层提取特征向量为 f_2 。将这些特征向量串接成一个特征向量,并利用 PCA^[12] 给它降维至 1 024 维,通过对数似然比的数值比较来实现面部验证。如公式(7)所示,其中, H_1 表示类内变化的熵和 H_E 表示类间变化的熵:

$$\log \frac{p(f_1, f_2 | H_1)}{p(f_1, f_2 | H_E)} \quad (7)$$

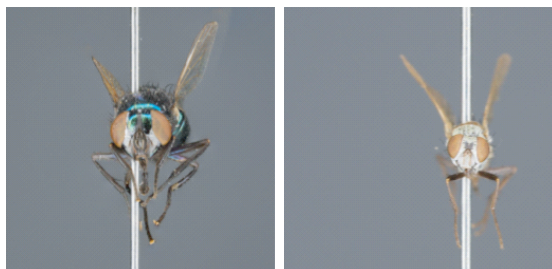
通过联合贝叶斯算法得到数据^[13],从而验证输入的两张面部是否为同一个昆虫。因此将面部的轮廓和具体部位特征提取相结合可以更精确地进行面部验证,从而可以了解特征提取程度,进而判断识别的效果。

4 实 验

本实验电脑配置为 Intel i7 处理器,显卡为 NVIDIA RTX2080Ti,16 G 内存,并基于 Caffe 深度学习框架的 GPU 版本实现了所提方法^[14]。利用前向网络上的标准反向传播,在动量为 0.9、学习率为 0.23,批量大小设置为 16 的情况下,通过随机梯度下降优化来训练所提出的 FFCNN 网络。从标准差为 0.01 的零均值高斯分布中提取权重来随机初始化所有的网络层。在数据集上对所提出的 FFCNN 进行了大约 24 000 次的迭代训练,训练的 epoch 为 100。本实验共包括数据集、识别过程、准确率及识别时间、ROC(Receiver Operating Characteristic)曲线及 LOSS 曲线共 4 个部分。

4.1 数据集

本实验选用由大连海关实验室提供的蝇类昆虫标本作为数据集样本,共选取 11 种蝇类标本包括叉叶绿蝇、大头金蝇、大洋翠蝇、横带花蝇、沈阳绿蝇、家蝇、厩腐蝇、巨尾阿丽蝇、亮绿蝇、丝光绿蝇、棕尾别麻蝇。选取蝇类标本其颜色形态不易变化且易于拍摄。图像采集设备为尼康(Nikon) COOLPIX A1000 数码相机,图像的分辨率为 4 608×3 456 pixel,格式为 JPG。为区分蝇类相似部分形态,将蝇类的 11 种样本分别以蝇类面部为中心随机角度拍摄图像,每种蝇类拍摄 60 张,共拍摄 660 幅图像。实验室拍摄的部分蝇类原始图像如图 7 所示。



(a)大洋翠蝇 (a)*Orthellia pacifica zimini* (b)横带花蝇 (b)*Anthomyia illocata*



(c)巨尾阿丽蝇 (c)*Aldrichina grahama* (d)丝光绿蝇 (d)*Lucilia sericata*

图 7 部分蝇类原始图像

Fig. 7 Original image of some fly species

在数据集制作过程中,为了增加训练的数据量,提高模型的泛化能力,避免数据量过少出现拟合现象,需要进行数据增强处理,保证每种蝇类的数目相同。通过图像翻转和平移、缩放原始图像的 20%、局部模糊等处理,增强图像后的数据集共得到 6 420 幅样本图像。并随机选取 60% 的图像,即 3 852 幅图像用作训练,选择 20% 的图像,即 1 284 幅图像作为验证集,剩下 20% 即 1 284 幅图像用来测试。

由于 MTCNN 检测方法对自然环境中光线、图像角度变化更具有鲁棒性,因此本实验使用 MTCNN 网络对原始图像进行面部定位和特征点检测。通过仿射变换和剪裁缩放得到对齐后统一尺寸为 95×95 的蝇类面部图像,同时记录对应蝇类的名字作为蝇类面部识别的标签。并展示部分数据集图像如图 8 所示。



(a)沈阳绿蝇 (a)*Lucilia Shen Yangensis* (b)棕尾别麻蝇 (b)*Boettcherisca peregrina* (c)叉叶绿蝇 (c)*Lucilia caesar*



(d)大洋翠蝇 (d)*Orthellia pacifica zimini* (e)大头金蝇 (e)*Chrysomya megacephala* (f)亮绿蝇 (f)*Lucilia illustris*

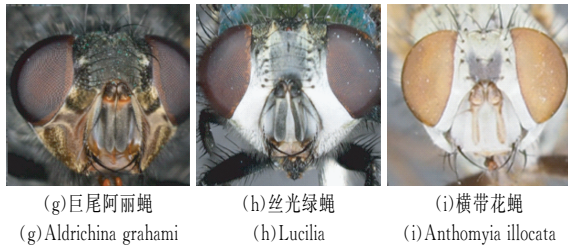


图 8 部分数据集图像

Fig. 8 Partial dataset images

此实验分别验证标准卷积和深度可分离卷积在 MTCNN 网络中的运行时间是不同的。并记录图 7 中列举部分蝇类在此网络中平均运行时间,如表 1 所示。从表 1 可以看出,应用深度可分离卷积确实可以减少参数量。通过不同卷积运行时间的比例可以看出,在 I/O 效率和性能不变的情况下,深度可分离卷积比标准卷积计算量降低近 8 倍。

表 1 不同卷积运行时间对比

Tab. 1 Comparison of different convolution time

(ms)

	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	平均值
标准卷积	64.5	66.8	67.4	65.8	66.1	65.4	68.2	65.9	66.3
深度可分离卷积	8.1	8.5	8.4	8.4	8.2	8.1	8.5	8.3	8.3

4.2 识别过程

本文使用 FFCNN 网络进行蝇类面部识别,应用 OpenCV 调用 Caffe 模型进行分类,具体过程如图 9 所示。为测试此网络模型训练效果的优劣,输入测试集非正面的蝇类图像。使用卷积池化粗提取出标准蝇类面部图像的轮廓特征值;同时,使用 Inception-ResNet, Reduction 网络细提

取具体部位特征值,并将两者的特征值串接为一组特征向量。最终识别结果为叉叶绿蝇的可能性达 98.995%,即识别正确并把正确的标签标注在输出图像中。因此,采用轮廓特征粗提取和具体部位特征细提取相结合的方式提取更加丰富的特征信息。与此同时,也验证了 FFCNN 网络模型的训练效果较好。

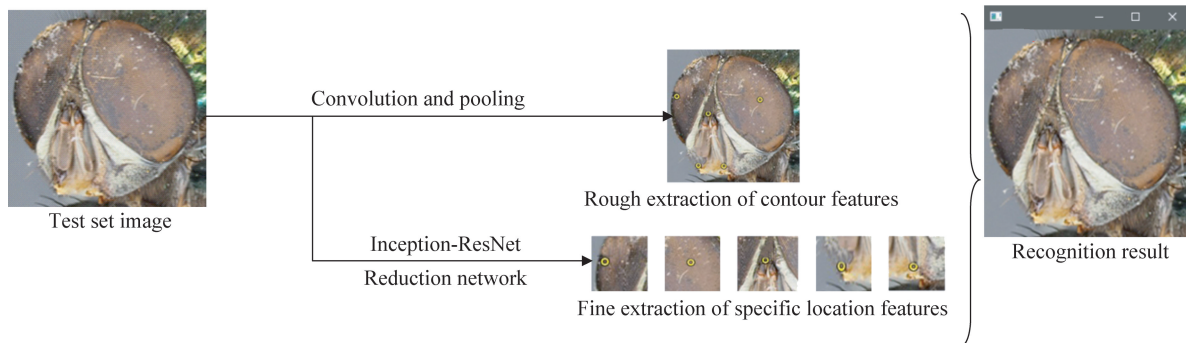


图 9 蝇类面部图像识别过程

Fig. 9 Face image recognition of fly

4.3 准确率及识别时间

通过多次随机输入两张蝇类面部图像,进行验证,从而得到各个网络下的准确率如表 2 所示。因为加入了轮廓与具体部位相结合特征提取和 Inception-ResNet 结构,不仅更加全面地进行了昆虫特征提取,还减少梯度散失更大程度地保留了特征信息。同时,为了图像尺寸缩减而引入的 Reduction 网络也在减少特征信息的流失。

FFCNN 网络的准确率高达到 94.03%,准确率最高。而其他几种方法没有应用此轮廓与具体部位相结合特征提取和 Inception-ResNet 网络以及 Reduction 网络,所以并没有取得较高的准确率。而验证准确率越高则说明其特征提取的信息越精准,从而证实了此网络的昆虫种类识别效果十分显著。

表 2 不同网络的准确率对比

Tab.2 Accuracy comparison of different networks (%)

方法	Inception-ResNet	Reduction	准确率
DeepID ^[15]	×	×	87.72
DCNNs ^[5]	×	×	90.01
ArcFace ^[16]	×	×	92.85
FFCNN	✓	✓	94.03

通过对数据集中面部图像测试,并将各种方法进行对比,并将其他用于人脸识别的方法如:将 DeepID,DCNNs,ArcFace 网络应用在蝇类面部。记录从不同网络输入时,平均每个蝇类面部图像识别所需时间,如图 10 所示。而且应用 Inception-ResNet 网络不仅减小梯度散失还同时在 Inception 的作用下提高计算效率,Reduction 网络在减小图像块尺寸时应用深度可分离卷积,减少大量计算参数,从对比可以看出 FFCNN 网络平均识别一张蝇类面部图像所用时间为 27 ms,虽然比近期发表的网络 ArcFace 网络识别时间要长,但是相对 DeepID,DCNNs 网络而言,耗时较少,计算效率较高。

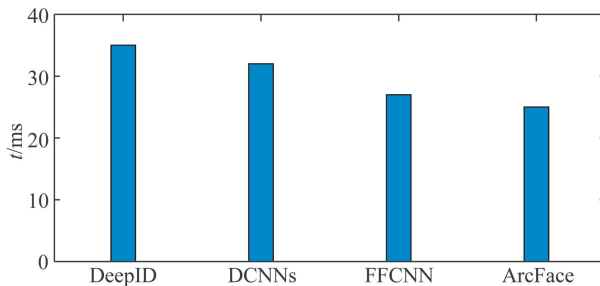


图 10 不同网络的识别时间对比
Fig. 10 Identification time comparison of different networks

4.4 ROC 曲线及 LOSS 曲线

若要测试此网络结构的性能,需绘制 ROC 曲线^[17]即感受性曲线。该方法简单直观,通过图示可观察分析方法准确性,并可用肉眼作出判断。其中横坐标为假正类率(False Positive Rate, FPR),纵坐标为真正类率(True Positive Rate, TPR)。其计算公式如式(8)和式(9)所示:

$$TPR = \frac{TP}{TP + FN}, \quad (8)$$

$$FPR = \frac{FP}{TN + FP}, \quad (9)$$

其中:FN 被判定为负样本,但实际为正样本;FP 被判定为正样本,但实际为负样本;TN 被判定为负样本,实际也是负样本;TP 被判定为正样本,

实际也是正样本。ROC 曲线将 FPR 和 TPR 以图示方法结合在一起,可准确反映学习器性能,是检测准确性的综合代表。并且靠近左上角的 ROC 曲线所代表的学习器准确性最高。实验结果如图 11 所示。经过对网络进行改进例如:轮廓与具体部位特征结合提取,加入 Inception-ResNet 网络以及 Reduction 网络,可以看出 FFCNN 网络曲线最靠近左上角,说明 FFCNN 网络的性能较好,准确率要高于其他网络。

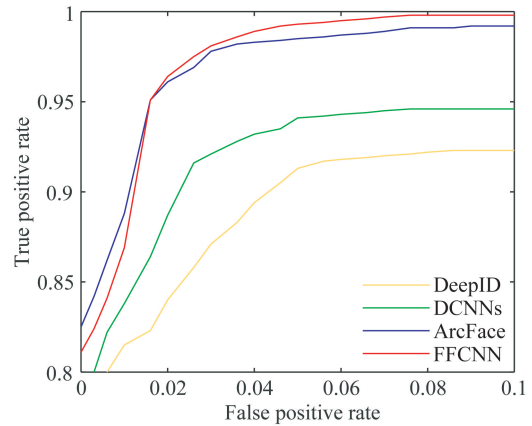


图 11 不同网络的 ROC 曲线
Fig. 11 ROC curves of different networks

通过对 DCNNs 网络损失函数的学习并应用在数据集上,可以得出 FFCNN 网络相应的损失函数的训练曲线如图 12 所示。可以看出在 epoch 为 0~40 时,曲线下落速度比较快,收敛较快;而随着横坐标的增大,epoch 为 40~100 时,曲线越发平缓,收敛较慢,损失函数的值也越发趋近于 0。迭代训练次数越多,损失函数值越小,最终趋近于 0。

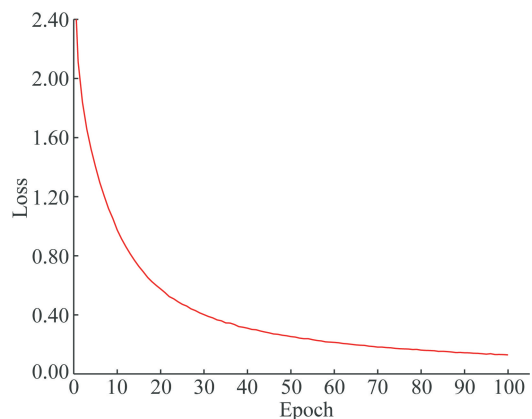


图 12 损失函数的训练曲线
Fig. 12 Training curve of loss function

5 结 论

本文借鉴深度学习在人脸识别中的方法,重点研究基于深度卷积神经网络在蝇类面部识别的应用。同时为解决蝇类面部识别精度较低,计算效率较慢等问题,对面部进行轮廓与具体部位特征细提取相结合,全面的提取特征向量,提升了准确率;将

Inception 和 ResNet 网络合理的融合,在防止梯度散失的同时提高了计算效率;应用 Reduction 网络通过深度可分离卷积使得计算量减小同时大幅度降低参数量。最终使 FFCNN 网络的准确率达到 94.03%,平均识别一张蝇类面部图像所用时间为 27 ms。同时,希望对未来的蝇类面部识别研究可以在非监督学习上取得显著进展,使蝇类昆虫识别问题在理论和实践上都得到更好地解决。

参考文献:

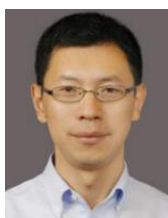
- [1] 刘颖,张倩楠,王富平,等. 基于颜色通道融合特征的现勘图像分类算法 [J]. 现代电子技术, 2020, 43(4): 67-72.
LIU Y, ZHANG Q N, WANG F P, *et al.*. Classification algorithm of prospecting image based on color channel fusion feature [J]. *Modern Electronics Technique*, 2020, 43(4): 67-72. (in Chinese)
- [2] 黄鸿,李政英,石光耀,等. 面向高光光谱影像分类的多特征流形鉴别嵌入 [J]. 光学 精密工程, 2019, 27(3): 726-738.
HUANG H, LI ZH Y, SHI G Y, *et al.*. Multi-feature manifold discriminant embedding for hyperspectral image classification [J]. *Opt. Precision Eng.*, 2019, 27(3): 726-738. (in Chinese)
- [3] 吴潇颖,李锐,吴胜昔. 基于 CNN 与双向 LSTM 的行为识别算法 [J]. 计算机工程与设计, 2020, 41(2): 361-366.
WU X Y, LI R, WU SH X. Behavior recognition algorithm based on CNN and bi-directional LSTM [J]. *Computer Engineering & Design*, 2020, 41(2): 361-366. (in Chinese)
- [4] HUANG, LAN, YANG, *et al.*. Deep learning versus Object-based Image Analysis (OBIA) in weed mapping of UAV imagery [J]. *International Journal of Remote Sensing*, 2020, 41(9): 221-227.
- [5] KANG B N, KIM Y, KIM D. Deep convolutional neural network using triplets of faces, deep ensemble, and score-level fusion for face recognition[C] 2017 *IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. *IEEE*, 2017: 76-84.
- [6] HE K, ZHANG X, REN S, *et al.*. Deep residual learning for image recognition[C] 2016 *IEEE Conference on Computer Vision and Pattern Recognition*. *IEEE*, 2016: 770-778.
- [7] SHI W W, GONG Y H, TAO X Y, *et al.*. Fine-grained image classification using modified DCNNs Trained by cascaded softmax and generalized large-margin losses [J]. *IEEE transactions on neural networks and learning systems*, 2019, 30(3): 683-694.
- [8] 王春哲,安军社,姜秀杰,等. 基于卷积神经网络的候选区域优化算法 [J]. 中国光学, 2019, 12(6): 1348-1361.
WANG CH ZH, AN J SH, JIANG X J, *et al.*. Candidate region optimization algorithm based on convolutional neural network [J]. *Chinese Optics*, 2019, 12(6): 1348-1361. (in Chinese)
- [9] ZHANG K, ZHANG Z, LI Z, *et al.*. Joint face detection and alignment using multitask cascaded convolutional networks [J]. *IEEE Signal Processing Letters*, 2016, 23(10): 1499-1503.
- [10] 陈希彤,卢涛. 基于全局深度分离卷积残差网络的高效人脸识别算法 [J]. 武汉工程大学学报, 2019, 41(3): 276-282.
CHEN X T, LU T. An efficient face recognition algorithm based on global depth separation convolution residual network [J]. *Journal of Wuhan Institute of Technology*, 2019, 41(3): 276-282. (in Chinese)
- [11] 杜振龙,沈海洋,宋国美,等. 基于改进 CycleGAN 的图像风格迁移 [J]. 光学 精密工程, 2019, 27(8): 1836-1844.
DU ZH L, SHEN H Y, SONG G M, *et al.*. Image style migration based on improved CycleGAN [J]. *Opt. Precision Eng.*, 2019, 27(8): 1836-1844. (in Chinese)
- [12] CHARLES B, PIERRE L, PIERRE A M. Exact dimensionality selection for Bayesian PCA [J]. *Scandinavian Journal of Statistics*, 2020, 47(1): 196-211.
- [13] MARCO S, CATHARINA E G, JOSE M G. Who

- learns better Bayesian network structures: Accuracy and speed of structure learning algorithms [J]. *International Journal of Approximate Reasoning*, 2019, 115(1): 235-253.
- [14] X CAO, D WIPF, F WEN, *et al.*. A practical transfer learning algorithm for face verification. [C] 2013 *IEEE International Conference on Computer Vision*, 2013: 3208-3215.
- [15] IGOR O R, IBAN T X, JONAS T W, UULKE A, *et al.*. A Deep Learning-based correction to EPID dosimetry for attenuation and scatter in the Unity MR-Linac system [J]. *Physica Medica*, 2020, 71(2): 124-131.
- [16] DENG J, GUO J, XUE N, *et al.*. Arcface: Additive angular margin loss for deep face recognition. [C] 2019 *IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2019: 482-493.
- [17] 张小荣, 胡炳樑, 潘志斌, 等. 基于张量表示的高光谱图像目标检测 [J]. *光学精密工程*, 2019, 27(2): 488-498.
- ZHANG X R, HU B Y, PAN ZH B, *et al.*. Target detection in hyperspectral images based on tensor representation [J]. *Opt. Precision Eng.*, 2019, 27(2): 488-498. (in Chinese)

作者简介:



陈彦彤(1989—),男,辽宁沈阳人,讲师,硕士生导师,2012年于吉林大学获得学士学位,2017年于中国科学院长春光学精密机械与物理研究所获得博士学位,主要从事图像处理及目标识别方面的研究。E-mail: chenyantong1@yeah.net



王俊生(1979—),男,黑龙江集贤人,教授,博士生导师,2002年、2007年于哈尔滨工业大学分别获得学士、博士学位,主要从事信号处理技术及应用、图像处理等方面的研究。E-mail: wangjish@dlnu.edu.cn