

doi:10.11835/j.issn.1000-582X.2022.09.013

基于随机权重分配策略的面目表情识别

张洋铭¹, 吴 凯², 王艺凡², 利 节²

(1. 复杂系统仿真总体重点实验室, 北京 100020; 2. 重庆科技学院 智能技术与工程学院, 重庆 400000)

摘要: 为了提高面部表情识别的精确度, 提出了一种基于数据增强策略面部表情识别, 区别于普通的在线随机数据增强, 将实验用到的训练数据集采用附加不同的权重分配策略进行增强数据, 并随机生成每次训练时的权重, 保证其训练数据的多样性并通过比较实验结果得出哪种权重的分布策略适用于面部表情识别数据集的增强, 同时解决了面部表情识别因数据集缺乏多样性识别精度不高等问题, 提升了人脸表情识别的准确性和鲁棒性, 此外还利用 VGG19 特征提取网络, 通过从数据中学习鲁棒性和区分性特征, 来实现高精度的面部表情识别。实验结果表明, 使用该方式增强后的数据进行训练的模型在 Fer2013 和扩展 Cohn-Kanade(CK+) 数据库上对 7 种表情的识别率相比其原始数据集均有提升。

关键词: 数据增强策略; 面部表情识别; VGG19

中图分类号: TP391

文献标志码: A

文章编号: 1000-582X(2022)09-135-06

Facial expression recognition based on random weight assignment strategy

ZHANG Yangming¹, WU Kai², WANG Yifan², LI Jie²

(1. General key Laboratory of complex system Simulation, Beijing 100020; 2. School of Intelligent Technology and Engineering, Chongqing University of Science and Technology, Chongqing 400000)

Abstract: In order to improve the accuracy of facial expression recognition, a facial expression recognition based on data enhancement strategy is proposed. The training data set used in the experiment is enhanced by adding different weight allocation strategies. The weights of each training are randomly generated to ensure the diversity of the training data, and which weight distribution strategy is suitable for the enhancement of facial expression recognition data by comparing the experimental results. At the same time, it solves the problem of facial expression recognition due to the lack of diversity and low accuracy of facial expression recognition, and improves the accuracy and robustness of facial expression recognition. In addition, it also uses VGG19 feature extraction network to achieve high-precision facial expression recognition by learning robustness and distinguishing features from the data. The experimental results

收稿日期: 2020-11-05

基金项目: 重庆市自然科学基金资助项目(cstc2020jcyj-msxmX0683); 重庆科技学院硕士研究生创新计划资助项目(YKJCX2020825); 重庆市研究生科技创新研究项目资助(CYS21505)。

Supported by the Chongqing Natural Science Foundation (cstc2020jcyj-msxmX0683); Innovation Program for Master Degree Students of Chongqing University of Science and Technology (YKJCX2020825); Chongqing Graduate Science and Technology Innovation Research Project (CYS21505).

作者简介: 张洋铭(1988—), 男, 博士, 主要从事复杂系统建模仿真, 决策评估理论与应用, 需求与战略方向研究, (E-mail)18992896079@163.com。

通信作者: 利节, 女, 重庆科技学院副教授, 硕士生导师, (E-mail)28605619@qq.com。

show that the model trained by the enhanced data in this way can improve the recognition rate of seven kinds of facial expressions on Fer2013 and extended Cohn-Kanade (CK+) database compared with the original data set.

Keywords: data enhancement strategy; facial expression recognition; VGG19

面部表情是人际交往过程中表达情感和含意的重要手段之一^[1]。近年来,面部表情识别在虚拟现实、智能教育系统、医疗卫生和数据驱动动画等领域都取得了较大进展^[2-4]。面部表情识别主要是根据给定的人脸图像来识别人脸的情绪和情感态度(如中性、愤怒、厌恶、恐惧、开心、伤心和惊叹)^[5]。一方面因在训练过程中表情标注主观性强,歧义较大,训练数据集标注难度大,导致可训练样少,甚至公开数据集均是缺乏多样性的小数据样本,导致训练困难和识别准确度不高。目前众多数据增强的预处理方法(比如随机裁剪,随机翻转等)可以解决数据样本缺失导致准确度不能提高的问题^[6],但采用具体模型适用于面部表情数据集的,才可以最大程度上提高识别的准确率。

笔者主要研究通过随机数据增强策略并结合基于特征的方法以提高人脸面部表情识别准确度,通过 VGG19 网络提取面部表情特征并进行准确识别。其中数据增强是采用 5 种数据增强方式(图像旋转、图像平移、图像缩放、图像翻转、图像投射)随机权重结合,通过实验结果得出了哪种数据增强的分配策略更适用于面部表情识别,即在保证可训练的数据集具有多样性的同时,得到更加准确的识别模型并掌握其权重的分布情况。

1 用于面目表情特征提取的网络 VGG19

VGG 在 2014 年由牛津大学著名研究组 VGG(visua geometry group)提出,斩获该年 Imagenet 竞赛中定位任务(localization task)的第一名和分类任务(classification task)的第二名^[7-8]。VGG19 的结构如下图 1 它是由 16 个 stride 为 1, padding 为 1 的 3×3 卷积核与 5 个 size 为 2, stride 为 2 的 maxpool 层加上 3 个全连接层最后添加一个 soft-max 层组成的。VGG 19 连续使用多个 3×3 卷积堆叠以便于优化并取代大体积的卷积核。多层非线性层在增加网络深度时,可以保证更复杂的模型学习而且代价更小(以至于达到参数更少的目标)。简单地来说,在 VGG 中,用 3 个 3×3 卷积核代替 7×7 的卷核,用 2 个 3×3 卷积核代替 5×5 的卷积核。其主要目的在于是在同一感受野的条件下,提高卷积网络的深度,在一定程度上提高神经网络提取特征的效果^[9]。

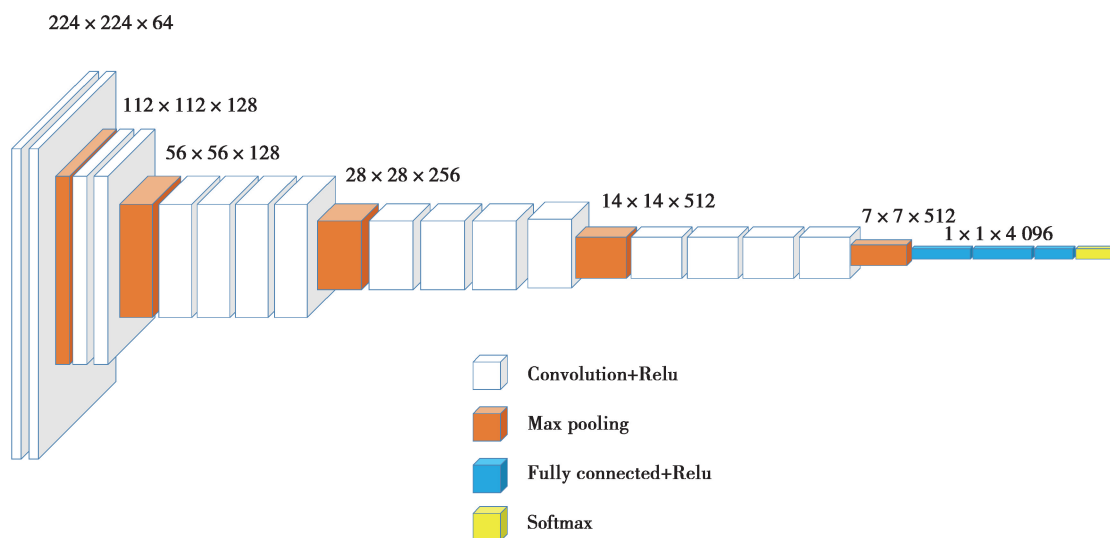


图 1 VGG19 结构图

Fig. 1 The structure diagram of VGG19

2 随机权重分布的数据增强

将输入图像并行通过穷举法策略,进行随机分配权重的5种数据增强方式,再将其送入到面目表情检测识别网络中通过采用数据增强策略以达到可提高面目表情识别准确率的目的。整体结构框图如图2所示。将原始数据 X 通过随机分配权重 $W_1 \sim W_5$ 得到分配后的图像数据 $X_1 \sim X_5$, 将其并行通过5种数据增强的图像变换并结合原始数据组成新的数据集 Y 。将新的数据集 Y 送入到图像特征提取 VGG19 网络模型中,进行后续的面目表情识别处理。

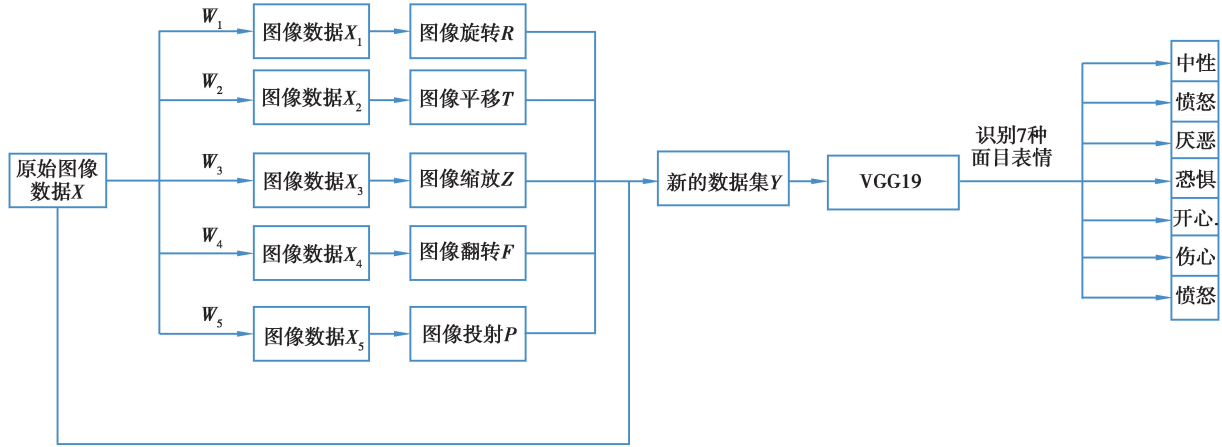


图2 随机权重分布的数据增强结构图

Fig. 2 Data enhancement structure diagram of random weight distribution

2.1 图像旋转(R)

图像旋转是指将图像以其中某一个点为旋转中心旋转一定角度,对图像对齐起着极其重要作用。以矩阵变换来表示设点 $P_0(x_0, y_0)$ 逆时针旋转 θ 角后对应点为那么,旋转前后点 $P(x_0, y_0)$ 的坐标分别为 $P(x, y)$

$$R = \begin{cases} x_0 = \gamma \cos \alpha, \\ y_0 = \gamma \sin \alpha, \end{cases} \quad (1)$$

$$R = \begin{cases} x = \gamma \cos(\alpha + \theta) = \gamma \cos \alpha \cos \theta - \gamma \sin \alpha \sin \theta = x_0 \cos \theta - y_0 \sin \theta, \\ y = \gamma \sin(\alpha + \theta) = \gamma \sin \alpha \cos \theta + \gamma \cos \alpha \sin \theta = x_0 \sin \theta + y_0 \cos \theta. \end{cases} \quad (2)$$

2.2 图像平移(T)

图像的平移操作就是将图像所有的像素点坐标分别加上水平与垂直操作上的偏移量。对于平移变换假设水平偏移量为 d_x , 垂直偏移量为 d_y , 则平移变换的坐标映射为

$$T = \begin{cases} x = x_0 + d_x, \\ y = y_0 + d_y. \end{cases} \quad (3)$$

其逆变换为

$$T = \begin{cases} x_0 = x - d_x, \\ y_0 = y - d_y. \end{cases} \quad (4)$$

x_1 与 y_1 的矩阵变换可表示为式(5)

$$T = \begin{bmatrix} x_1 & y_1 & 1 \end{bmatrix} = \begin{bmatrix} x_0 & y_0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ d_x & d_y & 1 \end{bmatrix}. \quad (5)$$

2.3 图像缩放(Z)

图像缩放是指对图像的大小进行调整的过程,是一种非平凡的过程,需要高效率,平滑和清晰,缩小图像

称为下采样,放大图像称为上采样。

$$\mathbf{Z} = \begin{bmatrix} k_x & 0 \\ 0 & k_y \end{bmatrix}^{-1} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}. \quad (6)$$

2.4 图像翻转(F)

图像的翻转变换是从 $A(x, y)$ (二维坐标) 到 $B(u, v)$ (二维坐标) 的一种线性变换,其数学表达式为

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} -1 & 0 & w \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & h \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & w \\ 0 & -1 & h \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}, \quad (7)$$

式(7)分别为水平翻转,垂直翻转,对角线翻转。且通过图像的翻转变换,可以使图像达到 180° 的翻转效果这样就可以加大其样本的多样性。

2.5 图像投射(P)

将一张图像投影到一个新的平面为图像的透射变换,其使用的变换公式为

$$\mathbf{P} = [x' \ y' \ w'] = [u \ v \ w] \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, \quad (8)$$

u, v 是原始图片左边,对应得到变换后的图片坐标 x, y , 其中: $x = x'/w', y = y'/w'$, 将变换矩阵

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \text{ 分为 4 个部分,用于透射变换。}$$

将 5 种数据增强方法进行基于随机权重分配策略结合其结果 Y 可表示为

$$Y = W_1 X R + W_2 X T + W_3 X Z + W_4 X F + W_5 X P, \quad (9)$$

式中 $W_1 \sim W_5$ 分别表示 5 种数据增强方式的随机权重分布。 $W_1 \sim W_5 X$ 分别表示进行并行处理图像数据。 RX_1 为进行随机旋转的图像样本; TX_2 为随机图像平移的图像样本; ZX_3 为随机图像缩放的样本; FX_4 为随机图像仿射变换的图像样本; PX_5 代表随机透射变换的图像样本。

3 实验结果分析

3.1 实验数据集

Fer2013 人脸面部表情数据集是人脸面部表情研究中最常用的数据集之一并且在研究当中占据了很高的地位,Fer2013 数据集已经为使用者划分好了训练集,验证集和测试集。Fer2013 数据集包含 35 887 张人脸图片,其中训练集 28 709 张,验证集 3 589 张,测试集 3 589 张。并该数据集中的图像均是灰度图像并且大小为 48×48 。样本被划分为 0=anger(生气)、1=disgust(厌恶)、2=fear(恐惧)、3=happy(开心)、4=sad(伤心)、5=surprised(惊讶)、6=normal(中性)7 类。

在 Cohn-Kanade 基础上扩充数据集产生了 CK+ 数据集,并在 2010 年发布。该数据集有 20% 的图像数据当作测试集并用于测试模型,80% 的图像数据用于训练模型。图 3,4 分别展示了 Fer2013 和 CK+ 人脸表情数据集对应的 7 种表情。



图 3 Fer2013 数据集的 7 种表情图像

Fig. 3 Seven facial expressions of Fer2013 dataset

3.2 训练数据并验证识别精准度

训练与测试数据集按 8 : 2 划分此整体数据集,并将训练集图像基于遗传算法的随机分配策进行权重的划分,并每次均采用相同的训练方式和识别网络架构进行实验。为验证其结果的通用性,实验中将使用的模



图 4 CK+数据集的 7 种表情图像

Fig. 4 Seven facial expressions of ck+ dataset

型为 VGG19,Resnet,Googlenet,为了提高模型的整体识别能力,通过微调将模型调整到最佳结构。训练的初始权重参数设置为 Epoch=30 时,batchsize=128,学习率为 0.001,优化算法采用随机梯度下降(SGD),并在每训练 1 个 Epoch 后使用此权重测试一次测试集,并记录每次实验的精准度。

通过随机算法生成不同的数据预处理子策略权重配比,比较实验结果准确度得出哪种比例数据预处理子策略配比方式更适用于面目表情识别网络并可以提升识别准确度。通过大量的等同条件重复实验,选取识别准确率最高的 6 种子策略权重配比方式,结果如表 1 所示。通过对比实验可看出图像旋转和图像平移 2 种数据增强子策略的权重比例增加,可一定程度上提高其模型的面目表情识别的准确率。

表 1 在 Fer2013 和 CK+ 数据集上数据预处理子策略配比最优权重

Table 1 Optimal weights of data preprocessing sub strategies on fer2013 and CK + datasets

5 种数据预处理子策略配比方式($W_1:W_2:W_3:W_4:W_5$)	Fer2013/%	CK+/%
2:3:1:1:1	79.229	78.198
2:2:1:1:0	80.001	77..213
3:4:1:2:1	79.712	78.888
3:6:0:1:1	80.009	81.387
7:4:1:1:1	82.701	79.780
4:3:0:1:0	77.000	81.785

为了验证实验的正确性,手动将数据预处理子策略配比方式调成 1:1:1:1:1 的形式,通过将数据预处理子策略的任意一种子策略手动调成 2,实验结果如表 2 所示。

表 2 在 Fer2013 和 CK+ 数据集上不同权重分配策略,准确度测试比较

Table 2 Accuracy test comparison with different weight allocation strategies on Fer2013 and CK + datasets

5 种数据预处理子策略配比方式($W_1:W_2:W_3:W_4:W_5$)	Fer2013/%	CK+/%
1:1:1:1:1	70.019	69.289
1:1:1:1:2	74.898	72.343
1:1:1:2:1	73.712	72.781
1:1:2:1:1	72.142	73.512
1:2:1:1:1	76.981	75.332
2:1:1:1:1	77.767	76.126

继续采用相同的数据预处理子策略权重配比方式,运用不同的网络结构在 Fer2013 训练集上训练表情识别模型,其性能在测试集上的结果比较如表 3,验证了其随机权重分配策略的通用性。但是,表情识别领域可能存在更好的权重分配策略,在今后的工作中将进一步深入探究。

表 3 采用不同网络结构与相同权重分配策略,平均准确度测试比较

Table 3 Comparison of average accuracy test with different network structures and the same weight allocation strategy

5 种数据预处理子策略配比方式($W_1:W_2:W_3:W_4:W_5$)	VGG19/%	Resnet/%	Googlenet
1:1:1:1:1	69.962	71.776	71.652
1:1:1:1:2	73.778	71.355	70.355
1:1:1:2:1	72.322	73.701	73.701
1:1:2:1:1	73.942	73.512	72.500
1:2:1:1:1	76.881	75.332	73.934
2:1:1:1:1	77.067	76.126	76.521

4 结束语

笔者提出了一种在现有的表情识别算法训练和识别的网络架构基础上增加了一种随机权重分配进行数据增强的预处理方式,解决了由于在训练过程中表情标注主观性强,歧义较大,导致可训练样本缺少,识别准确度不高等问题。实验结果表明,文中提出的方法可提高图像的数据库质量,提高分类器的性能和面目表情识别的准确率。

参考文献:

- [1] Anderson K, Mcowan P W. A real-time automated system for the recognition of human facial expressions[J]. IEEE Transactions on Cybernetics, 2006, 36(1):96-105.
- [2] Bengio Y, Lamblin P, Popovici D, et al. Greedy layer-wise training of deep networks[C]// Advances in Neural Information Processing Systems 19, Proceedings of the Twentieth Annual Conference on Neural Information Processing Systems. Vancouver, British Columbia, Canada; Morgan Kaufmann, 2006, 19: 4-7.
- [3] Chen D, Cao X, Wen F, et al. Blessing of Dimensionality: High-Dimensional feature and its efficient compression for face verification[C]//Computer Vision & Pattern Recognition, Piscataway: IEEE, 2013:3025-3032.
- [4] Chen H, Liang J, Wang Z, et al. Pinning controllability of autonomous Boolean control networks[J]. Science China (Information Sciences), 2016, 59(7):1-14.
- [5] Coates A, Ng A, Lee H. An analysis of single-layer networks in unsupervised feature learning[C]//Journal of Machine Learning Research - Proceedings Track. U.S.A.: MIT press, 2011, 15: 215-223.
- [6] 陈超, 齐峰. 卷积神经网络的发展及其在计算机视觉中的应用综述[J]. 计算机科学, 2019, 46(3): 63-73.
Chen C, Qi F. Review on development of convolutional eural network and its application in computer vision [J]. Computer Science, 2019, 46(3): 63-73.(in Chinese)
- [7] 高慧明. 基于卷积神经网络的肺结节检测及良恶性分类方法研究[D]. 太原: 太原理工大学, 2019.
Gao H M. Research on detection and benign and malignant classification method of pulmonary nodules based on convolutional neural network[D]. Taiyuan: Taiyuan University of Technology, 2019.(in Chinese)
- [8] 张璟. 基于改进卷积神经网络的交通标志牌识别研究[J]. 电脑知识与技术, 2019, 15(21): 195-197.
Zhang J. Research on traffic signs recognition based on improved convolutional neural network[J]. Computer Knowledge and Technology, 2019, 15(21): 195-197.(in Chinese)
- [9] 贺思艳, 任利娟, 田新诚. 基于卷积神经网络的滚动轴承故障诊断[J]. 兵工自动化, 2019, 38(3): 39-41.
He S Y, Ren L J, Tian X C. Rolling bearing fault diagnosis based on convolution neural network [J]. Ordnance automation, 2019, 38(3): 39-41.(in Chinese)