

基于深度学习的表情识别

2022 春季媒体与认知大作业汇报

程凯越 无 93 2018011113

清华大学电子工程系

2022 年 6 月 9 日



大纲

① 简介

② 方法

③ 结果

① 简介

② 方法

③ 结果

任务定义

基本

- 输入 一张人脸图像
- 输出 图像中人的表情

进一步

- 输入 一张包含人脸的图像
- 输出 图像中的人的表情

主要工作

- **人脸特征** 引入 dlib 的人脸关键点特征
- **数据处理** 引入超分辨率重建，重写 Dataset 和 DataLoader
- **模型** 尝试现有模型，自己搭建了一个深浅特征残差的网络
- **训练** tensorboardX 可视化，尝试 Adam+SGD 的训练方式
- **花里胡哨** 实时摄像头人脸检测 + 表情识别

实时摄像头人脸检测 + 表情识别

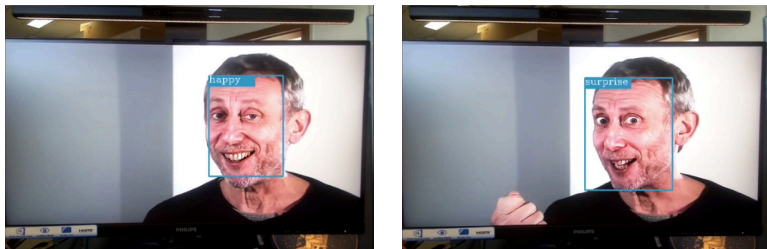


图 1: “nice 爷爷” 人脸检测与表情识别 (happy, surprise)

实时摄像头人脸检测 + 表情识别

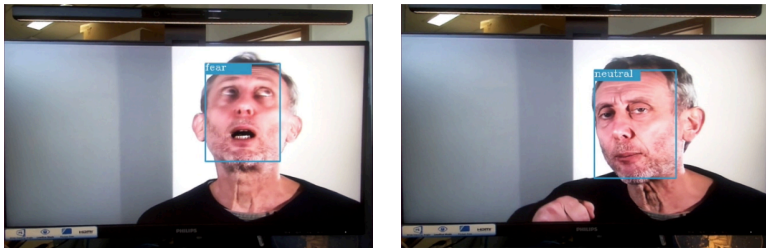


图 2: “nice 爷爷” 人脸检测与表情识别 (fear, neutral)

① 简介

② 方法

③ 结果

dlib 人脸特征提取与分类

dlib

- 输入 一张人脸图像
- 输出 128 维特征向量

分类方式

- SVM
- 随机森林
- 三层 MLP

dlib 人脸特征提取与分类

dlib 实现

- dlib 的 shape_predictor 和 face_recognition_model_v1
 - Code Link: dlib.net/face_landmark_detection.py.html
 - Code Link: dlib.net/face_recognition.py.html

分类实现

- sklearn 的 SVM 和随机森林
- 三层 MLP:128→1024→7

数据增强

数据增强方式

- 随机水平翻转
- 随机顺时针或逆时针旋转 $0 \sim 30$ 度

实现

- pytorch 的 transforms
- transforms.RandomHorizontalFlip()
- transforms.RandomRotation(30)

超分辨率重建

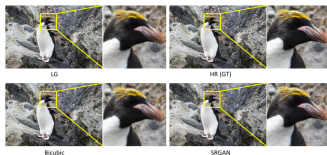


图 3: SRGAN 超分辨率重建 [LTH⁺17]

Link: blog.csdn.net/qianbin3200896/article/details/104181552

在我们的数据上:

- 输入 48×48
- 输出 192×192

网络结构

- 自己设计了一个网络
- 还尝试了经典网络：AlexNet, Densenet121, GoogLeNet, MobileNetV3, ResNet18, ResNet50, VGG11, VGG16
- 除了自己设计的网络，均采用 PyTorch 提供的预训练模型，仅更新分类模块和特征提取模块的最后几层参数

网络结构

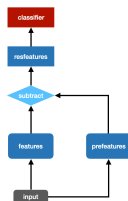


图 4: 网络结构

- 假定表情识别任务中，噪声来源于图像的浅层特征（人脸位置、背景物体等）
- 两个提取特征模块，一个提取深层特征，一个提取浅层特征
- 深层特征 - 浅层特征 = 残差特征
- 残差特征再过几个网络层，分类

训练及模型可视化

- tensorboardX 可视化模型
- tensorboardX 监测 Acc 和 Loss

优化训练方式

- Step1 超分辨率重建数据集, Adam, $1e-3$
- Step2 原始数据集, Adam, $1e-3$
- Step3 原始数据集, SGD, $1e-2$
- Step4 原始数据集, SGD, $1e-2$
- Step1,2 的 epoch 为 70; Step3,4 的 epoch 为 60
- batch size 均为 64
- Step1,2 的 milestone 为 [20,40,60]; Step3,4 的 milestone 为 [40]
- 在 milestone 更新学习率的方式均为 $\times 0.1$

测试方法

三种测试方法

- 整个数据集的测试
- 单张图像的测试
- Real Time 摄像头读取 + 检测 + 表情分类

人脸检测

- 直接使用 `mxnet_mtcnn_face_detection` 的工作 [ZZLQ16]
- GitHub Link:
github.com/YYuanAnyVision/mxnet_mtcnn_face_detection

① 简介

② 方法

③ 结果

dlib 人脸特征提取与分类

表 1: dlib 特征分类算法及测试集准确率

算法	超分辨率重建 Test Acc	原始 Test Acc
SVM	42.83	42.76%
RF	46.61	47.02%
MLP	45.84	46.84%

数据增强

表 2: 模型及测试集准确率 (超分辨率重建数据)

算法	数据增强 Test Acc	原始 Test Acc
AlexNet	61.61%	58.82%
Densenet121	54.47%	50.32%
GoogLeNet	55.14%	52.78%
MobileNetV3	57.64%	55.55%
ResNet18	58.20%	56.82%
ResNet50	58.14%	54.68%
VGG11	61.07%	58.20%
VGG16	61.49%	57.29%

超分辨率重建



图 5: 超分辨率重建前后的图像 (左侧为原图, 右侧为重建图)

超分辨率重建

表 3: 超分辨率重建前后模型准确率

算法	超分辨率重建 Test Acc	原始 Test Acc
AlexNet	61.61%	62.28%
MobileNetV3	57.64%	61.92%
ResNet50	58.14%	61.75%
VGG16	61.49%	63.59%

超分辨率重建算法可能引入了噪声对训练造成负面影响。

网络结构



图 6: TensorBoardX 绘制模型图

网络结构

表 4: 网络及测试集准确率 (超分辨率重建数据)

Model	Test Acc
Densenet121	54.47%
GoogLeNet	55.14%
MobileNetV3	57.64%
ResNet50	58.14%
ResNet18	58.20%
Ours	60.83%
VGG11	61.07%
VGG16	61.49%
AlexNet	61.61%

优化训练方式

表 5: 训练参数

Step	Dataset	lr	milestones	epoch	Acc
1	srdata	1e-3	[20,40,60]	70	60.51%
2	rawdata	1e-3	[20,40,60]	70	61.61%
3	rawdata	1e-2	[40]	60	61.57%
4	rawdata	1e-3	[40]	60	61.74%

备注:

- batch size 均为 64
- srdata 为超分辨率重建数据, rawdata 为原始数据

训练过程



图 7: TensorBoardX 监测训练 Loss 和 Acc

整活

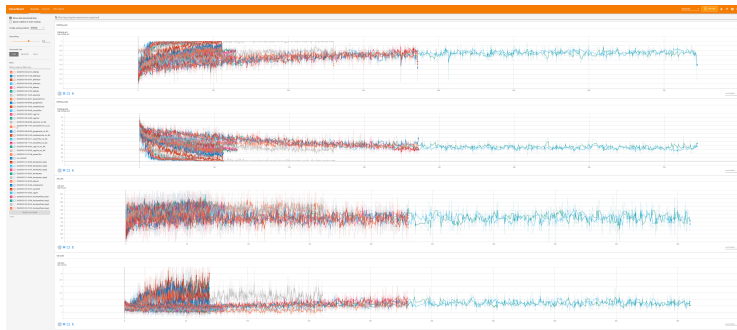


图 8: 把所有监测的数据都放在一起

测试

- Video Demo (Bilibili ID: BV1MA4y1R7F3)
- Video Demo (清华云盘)

Thank You!

L^AT_EX 模版参考: [Wen20]

- [LTH⁺17] Christian Ledig, Lucas Theis, Ferenc Huszár, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, et al.
Photo-realistic single image super-resolution using a generative adversarial network.
In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 4681–4690, 2017.
- [Wen20] Jiayi Weng.
Thu beamer theme.
Overleaf Templates, LaTeX beamer theme for Tsinghua University students. Chinese support, 2020.
- [ZZLQ16] Kaipeng Zhang, Zhanpeng Zhang, Zhifeng Li, and Yu Qiao.
Joint face detection and alignment using multitask cascaded convolutional networks.
IEEE signal processing letters, 23(10):1499–1503, 2016.