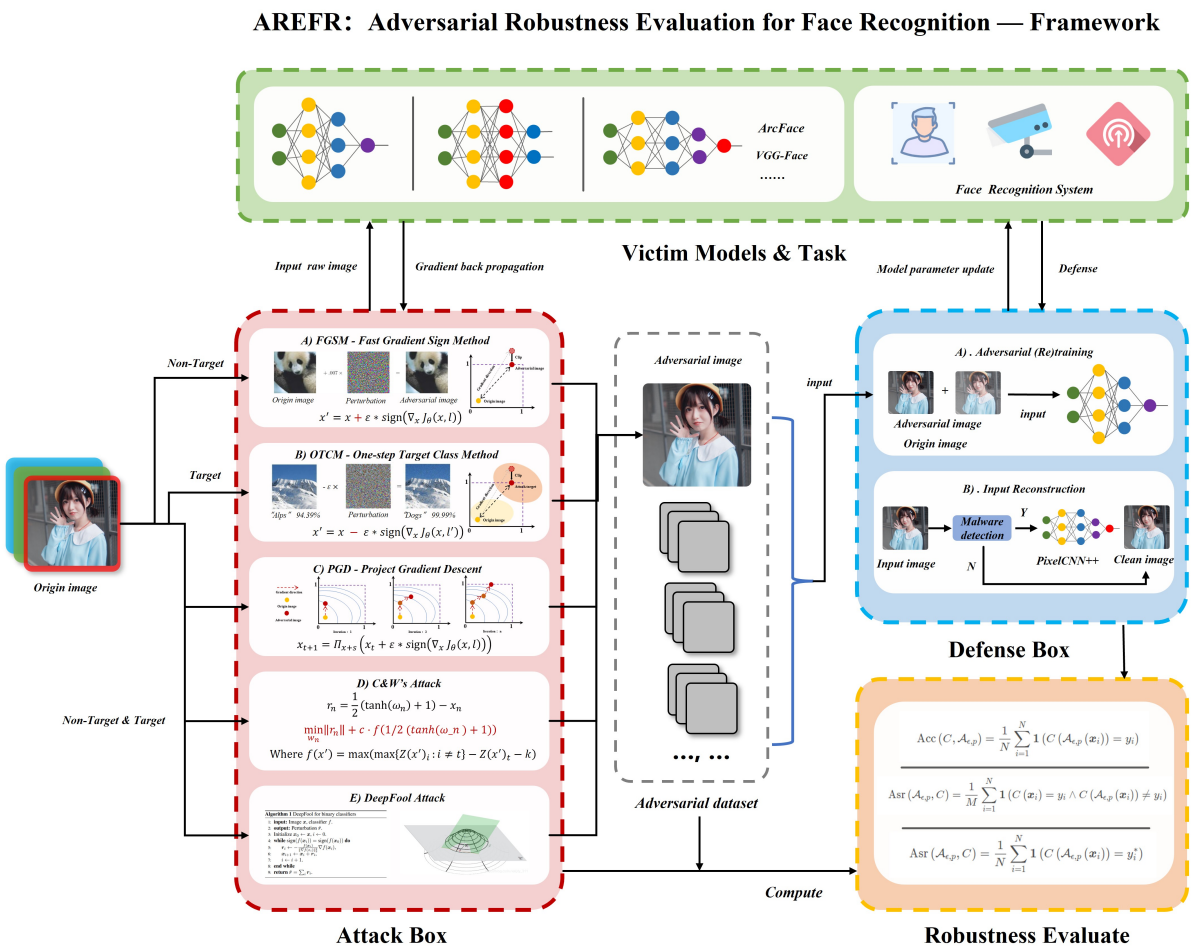
AREFR 使用手册



该文档为 人脸识别对抗鲁棒性评估平台(AREFR) 的中期报告暨解释性文档(中文版)，我们旨在向使用者说明该工程的所有功能，并且方便之后的扩展开发

> 开发者: 西安电子科技大学人工智能学院 *Silvester_Ruan*、*LYC*

一、工程文件结构

该项目主要关注人脸识别模型下的对抗鲁棒性测试， 并且着重关注物理场景下的模型鲁棒性。主要由人脸识别模型、攻击算法、评估代码、其他工具性的脚本组成，以下是工程文件内各个脚本的说明:

```
· Attack                                  : 内含实现的攻击算法源代码
· checkopints                             : 人脸识别backbone模型的训练权重
· data                                    : 人脸识别模型的测试数据
    - facebank                            : 用于攻击时进行测试的基准数据集里面包含大型
数据集中采集的人脸和现实生活中的人脸
    - lfw-align-128                       : LFW基准测试集，用于训练后的模型精度评估
    - cleaned list.txt                    : LFW测试集的干净数据索引
    - lfw_test_pair_target_attack         : 用于有目标攻击测试的数据对
    - lfw_test_pair.txt                   : 用于无目标攻击测试的数据对
```

```
· face_detect_feature                    : haar和lbp人脸检测的特征参数
· fig                                    : 进化黑盒攻击的迭代过程图
· model                                  : 人脸识别模型部分
    - loss                               : 自定义focalloss
    - metric                             : ArcFace/CosFace 度量函数
    - mobilenet.py                       : backbone, mobilenet
    - resnet.py                          : backbone, resnet

· sample                                 : 存放一些杂七杂八的图像，比如眼镜口罩的图像，以及一些测试图像
· attack_evaluate.py                     : 攻击效果评估代码
· attack_example.py                      : 生成单个对抗样本的代码
· config.py                              : 工程的所有超参数设置
· dataset.py                             : 训练数据的批量读取

· demo.py                                : 人脸检测与识别系统的运行代码
· face_alignment.py                      : 工具脚本 – 人脸对齐
· feature_dict.pkl                       : facebank内所有人物的特征向量文件，在运行demo时默认重新计算一遍

· put_glass.py                           : 工具脚本 – 添加眼镜
· put_mask.py                            : 工具脚本 – 添加口罩
· shape_predictor_68_face_landmarks.dat  : 用于面部检测和人脸对齐的68点特征文件
· take_picture.py                        : 工具脚本 – 向facebank内添加任务，允许用户进行拍照和上传照片

· test.py                                : 人脸识别模型训练后的测试代码
· train.py                               : 人脸识别模型训练代码
· utils.py                               : 一些工具函数
```

二、网络训练和测试

2.1 受害模型

在人脸识别的受害模型上, 项目目前支持四种模型的训练:

```
1、ArcFace + mobilenet
2、ArcFace + resnet
3、CosFace + mobilenet
4、CosFace + resnet
```

在 *conf.py* 内可以修改

```
""" 模型 超参数"""
backbone = 'fmobile'  # ['resnet', 'fmobile']
metric = 'arcface'  # ['cosface', 'arcface']
```

2.2 训练

在终端内输入

```
python train.py
```

即可使用默认超参数进行模型训练，若需要修改参数，我们提供两种方式，可以直接在*conf.py* 内修改，或者使用命令行，输入以下命令查看可修改的超参数及其说明：

```
python train.py -h
```

目前支持的可修改的超参数为：

超参数名称	含义	可选值或数据类型	默认值
--backbone	设置主干网络采用的模型	'resnet' 'fmobile'	'fmobile'
--metric	网络训练时计算损失采用的距离度量函数	'cosface' 'arcface'	'arcface'
--loss	网络训练时的损失函数	'focal_loss' 'cross_entropy'	'focal_loss'
--optimizer	梯度下降使用的优化器	'sgd' 'adam'	'sgd'
--batch_size	一个批次的样本数量	**int**	8
--epoch	训练轮数	**int**	60
--lr	梯度下降学习率	**float**	0.1

例如：

```
python train.py --backbone='fmobile' --epoch=2 --lr=0.001
```

3.9 ~ 3.18 Nerf 及仿真平台调研

3.18 跑通Nerf

3.25 收集沙盘数据，跑通Nerf