

大作业：人脸口罩检测

Lecturer: Changshui Zhang zcs@mail.tsinghua.edu.cn

Student:

1 作业主题

众志成城，抗击疫情。在全国人民尤其是医护人员的努力下，中国的疫情暂时趋于稳定，但是世界范围内的疫情仍不容乐观。在抗击疫情的过程中，人脸口罩检测（Face Mask Detection）是其中一项必要的工作，多家科技公司已经研发出了带有 AI 人脸检测算法的口罩佩戴检测设备。

目标检测（Object Detection）是重要的计算机视觉任务，而人脸检测（Face Detection）则是其中最久远的应用之一，也是本大作业的基础。文献 [8] 详尽地综述了目标检测任务的相关发展进程，请大家在阅读该综述的基础上，广泛阅读相关文献，深入思考，更好地完成本次训练。

在本次课程项目中，我们聚焦于使用机器学习方法检测**人脸是否佩戴口罩**。作业要求完整地完成数据处理、建模、实验、分析报告整个过程。

本次作业建模为一个两类检测问题：戴口罩的人脸和不戴口罩的人脸。对于每一个类别，我们采用 mAP (mean Average Precision) 作为检测效果的评价指标。关于 mAP 的定义和计算方法，可以参考[这里](#)和[这里](#)。在计算 mAP 时，请采用**PASCAL VOC 风格的 11 点插值法**。

本题目对计算机算力要求较高，建议在有 GPU 的个人计算机或服务器上完成实验。

2 报告内容

要求以小论文的格式提交实验报告，报告需要包括但不限于以下内容（基础分 80 分），在基本要求以外的实验、分析、讨论等均可以加分（满分 100 分，加满为止）。

评分时将更看重报告的完整度，在评价算法结果时将更看重改进的模型相对于基础模型的提高，而非最终结果的绝对值。需要注意的是，为了使得相对提高明显而故意把基础模型性能做差的行为是严格禁止的，这一点在评分时会有体现。

一、题目，摘要，关键词

二、简介 (10')

- 介绍目标检测任务及其主流解决方法，定义人脸口罩检测任务。
- 介绍深度学习方法在人脸口罩检测任务中的应用。
- 说明本文的主要贡献。

三、数据整理 (10')

- 介绍数据来源、内容和处理步骤。
- 数据清洗（如有必要）。
- 可视化地观察数据分布。

四、模型设计 (30')

- 构建至少 3 种不同的模型。可选择但不限于以下方法：
 1. Faster RCNN [6]
 2. SSD [3]
 3. YOLO [4]
 4. YOLOv2 [5]
 5. RetinaNet [2]
- (加分项) 尝试对所选模型中性能最好的一个做出自己的改进，改进可选择但不限于以下方面：
 1. 不损害模型性能的前提下，加快模型推理速度。
 2. 提高部分遮挡情况下的模型性能（例如手把嘴挡住不应该认为是戴了口罩）。
 3. 不采用预训练模型，从随机初始化开始训练检测器。
 4. 将模型部署上线，让用户可以在网页上通过上传图片或者打开摄像头来输入数据，并且在网页上实时展示检测结果。

五、实验设计及结果 (15')

- 数据集如何划分成训练、验证、测试集。
- 在报告每种模型性能时，须在表格中展示每类的 $mAP@.5$, $mAP@.7$, $mAP@.9$, $mAP@[.5:.95]$ 这四个指标，并且在图中画出 IoU 阈值取 0.5, 0.7, 0.9 时每类的 Precision-Recall 曲线。要注意合理绘制图表，保证每个图表的可读性的同时使得图表总数不要过多。

- 每种模型单独的最好结果对比（列出图表并进行讨论）。
- 每种模型在不同超参数下的表现（列出图表并进行讨论）。
- （加分项）其他实验设计及结果。

六、实验结果分析（15'）

实验结果分析的方式可以包括但不限于：

- 讨论数据和模型中每一部分的贡献（如果删去/更换部分数据输入/实验设定/模型结构，结果会发生什么变化）。
- 特征的重要性分析。
- 错误分析（模型在哪些数据下预测准确率高，哪些数据下预测准确率低）。可以采用 Derek 提供的分析工具 [1] 来辅助进行分析。
- 案例分析（在具体的案例上，不同模型表现的区别在哪里）。
- 模型和结果可视化分析。
- （加分项）其他方式的实验结果分析。

七、小组成员贡献

八、参考代码

九、结论

3 数据采集

人脸检测相关数据非常丰富，主要存在两种形式的口罩数据。(1) 真实口罩人脸识别数据集，从网络爬取样本，经过整理、清洗和标注后口罩数据。(2) 模拟口罩人脸识别数据集，通过给人脸检测公开数据集中的人脸戴上口罩，得到合成的口罩数据。下面推荐三个数据获取网站，其中，前两个为口罩检测相关数据集（已转载至[清华云盘链接](#)），最后一个为人脸检测相关数据集。利用已整理的公开数据集是一种解决问题的方式，同样鼓励自行检索和采集其他数据，可以根据算力和方法自由选择。

- AIZOO 的 FaceMaskDetection：
<https://github.com/AIZOOTech/FaceMaskDetection>

该网站开源了人脸口罩检测的主流框架的相应模型，并提供了相应的推理代码。该作者开源了如表格所示的 7,959 张人脸标注图片，数据集来自于 WIDER Face 和 MAFA 数据集，并重新修改了标注和校验。

数据集	WIDER	MAFA	合计
训练集/张	3,114	3,006	6,120
验证集/张	780	1,059	1,839

表 1: AIZOO 口罩检测数据集

- Real-World Masked Face Dataset (RMFD):

<https://github.com/X-zhangyang/Real-World-Masked-Face-Dataset>

武汉大学国家多媒体软件工程技术研究中心收集和标注的口罩数据集 [7]。包含从网络爬取样本，经过整理、清洗和标注后，含 525 人的 5 千张口罩人脸、9 万正常人脸。以及通过公开数据集中的人脸戴上口罩，得到 1 万人、50 万张人脸的模拟口罩人脸数据集。该博客中公布了合成数据的工具，但合成方法比较单一，同学们可以结合下一节提供的人脸检测数据集，提出适合的合成数据方法。

数据集	人物数量	人脸数量
真实数据	525	90,000
合成数据	10,000	50,000

表 2: RMFD 人脸口罩识别数据集

- 人脸检测、识别、跟踪、对准等丰富的工作列表和数据链接：

https://github.com/ChanChiChoi/awesome-Face_Recognition#datasets

该博客提供了丰富的人脸相关文章和数据。该博客主要给出了四个人脸检测的数据集，如表格所示。熟读该博客中的相关文章，可为本次大作业提供不少思路。

数据集	图像数量	人脸数量	发布年份	下载地址
Fddb	2,845	5,171	2010	Link
Wider-face	32,203	393,703	2015	Link
AFW	205	473	2013	Link
MALF	5,250	11,931	2015	Link

表 3: 人脸检测数据集

4 作业提交

- 大作业提交截止时间为 5 月 29 日 23:59。
- 编程语言不限，建议使用 python，并采用 pytorch 工具包，配置环境较为方便。

- 可以个人或 2 人小组的形式完成作业，每小组提交一份即可。报告中应当用独立段落说明成员贡献。注意，对小组完成的作业要求更高。
- 完成作业过程中可以参考已有的代码，但要在报告中用独立段落给出详尽说明，具体到自己提交的代码中哪一个文件哪些行，并提供参考来源的链接。若参考的代码过多，将影响评分；若在未标明参考他人的部分发现了雷同现象，本次作业将判定零分。
- 自己完成的全部代码均需要进行提交。库函数的代码无需提交，除非自己修改了某个已有的库。
- 需要手工从测试集中挑选五张戴口罩的测试图片 `mask1.jpg ~ mask5.jpg` 和五张不戴口罩的测试图片 `unmask1.jpg ~ unmask5.jpg`，作为示例放入最终提交的压缩包中。除了这十张图以外的数据无需提交。
- 需要从所做的模型中选择一个性能最优的进行提交。
- 提交的代码必须是自包含的 (self-contained)。即，从下载原始数据开始，不依赖于任何中间结果，必须要能够完整地复现训练、测试过程。
- 需要写一个 **Markdown** 格式的 `README.md`，内容至少应包括：
 - 按模块描述压缩包中每部分文件的作用。
 - 运行代码所需的软件环境和软件版本，并提供从裸的操作系统开始配置所需环境的命令。
 - 下载原始训练测试数据和整理数据的命令。
 - 训练所提交的最优模型的命令。按此命令训练出的模型，不应当与提交的最优模型性能差异过大。
 - 用上一步训练出来的模型在十张图片上测试的命令。
 - 用提交的最优模型在十张图片上测试的命令，以及预期的结果。

`README.md` 必须包含关于环境、数据、训练、测试的所有命令。`README.md` 中的命令依次复制粘贴到命令行里执行，应当能够复现整个流程。提交前建议反复检查并确认这一点，复现失败将极大影响评分。

- 最终作业以 zip 压缩包形式提交。除压缩包命名外，目录名和文件名采用英文，防止因为操作系统不同造成乱码。提交的压缩包展开后目录结构应当如下：

```
- 张三-李四-人脸口罩检测.zip
- codes/
  - README.md
  - test-images/
    - mask1.jpg - mask5.jpg
    - unmask1.jpg - unmask5.jpg
  - ...
- report.pdf
```

其中，`report.pdf` 为报告，... 部分为代码，可以包含多个文件或目录。

- 最终提交的压缩包体积**不得超过 20MB**，如模型或者图片太大导致超出这一限制，需要将较大的文件移出压缩包直至满足这一限制为止，并将这些文件放在互联网上（例如清华云盘）且在 README.md 中提供下载命令。

参考文献

- [1] Derek Hoiem, Yodsawalai Chodpathumwan, and Qieyun Dai. Diagnosing error in object detectors. In *European conference on computer vision*, pages 340–353. Springer, 2012.
- [2] Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollár. Focal loss for dense object detection. In *Proceedings of the IEEE international conference on computer vision*, pages 2980–2988, 2017.
- [3] Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, and Alexander C Berg. Ssd: Single shot multibox detector. In *European conference on computer vision*, pages 21–37. Springer, 2016.
- [4] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 779–788, 2016.
- [5] Joseph Redmon and Ali Farhadi. Yolo9000: better, faster, stronger. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 7263–7271, 2017.
- [6] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster r-cnn: Towards real-time object detection with region proposal networks. In *Advances in neural information processing systems*, pages 91–99, 2015.
- [7] Zhongyuan Wang, Guangcheng Wang, Baojin Huang, Zhangyang Xiong, Qi Hong, Hao Wu, Peng Yi, Kui Jiang, Nanxi Wang, Yingjiao Pei, et al. Masked face recognition dataset and application. *arXiv preprint arXiv:2003.09093*, 2020.
- [8] Zhengxia Zou, Zhenwei Shi, Yuhong Guo, and Jieping Ye. Object detection in 20 years: A survey. *arXiv preprint arXiv:1905.05055*, 2019.