

林培森

+49-163-622-8917|linpeishen@163.com | 1995/07/20 | 中共党员
GitHub: <https://github.com/linpeishen/i2dl/tree/master/projects>



教育背景

慕尼黑工业大学 研究型硕士 机器认知与人工智能 2019.04-2021.04

- 核心课程：深度学习、计算机视觉、机器学习

同济大学 工学学士 中德工程学院 机械电子工程 2014.09-2018.09

- 核心课程：设计语言 (Java)、信息技术 (C++)、网络通讯系统

GPA 86.5/100 所获荣誉：校级二等奖学金两次（学年排名前 15%），德国学术交流奖学金 (DAAD)

实习/项目经历

德国，慕尼黑 全卷积网络实现图片语义分割(AlexNet) 2019.05-2019.06

- **数据集预处理**：对原始图片和目标图片数据集进行中心裁剪到 240*240。目标图片与预设置的分割标签列表对比，得到整张图片各像素的对应标签。根据像素标签将其转换为 RGB 实现目标图片可视化。
- **构建训练全卷积网络**：以 AlexNet 结构为原型，构建一个分割网络。其参数在 pretrained 基础上，用上述数据集继续训练。此处调用 PyTorch 相关函数 (loss.backward, optim.step 等) 规范化训练网络。在设置 optim 参数时，选择 model.parameters 对全模型参数进行训练以提高准确率（对比 model.classifier.parameters），保存 validation 准确率最高模型。
- **测试**：对最优模型用 test 数据集进行测试，test 准确率超过 70%。
- **可视化**：使用 matplotlib.pyplot 库函数展示预测结果和 filter 权重。

德国，慕尼黑 深度残差网络(ResNet)实现实时人脸识别 2019.11-2020.01

- **搭建 ResNet**：根据 ResNet 相关论文对其结构进行总结，并搭建更加结构化的 ResNet。
- **数据集预处理**：以用户十秒短视频和 lfw (Labeled Faces in the Wild) 作为数据集，使用 dlib 库中的方向梯度直方图方法对视频图片中的脸部进行检测，然后将其处理成图片集。将图片集分割为训练验证两个数据集并对其进行转换（比如归一化，张量化，随机水平翻转等），提高模型的泛化性。
- **训练验证测试网络**：使用 GPU 对 ResNet 进行训练和验证，设置提前结束训练条件，保存最优模型，使用手机录像，截取 YouTube 上视频（高清）对模型进行测试。使用电脑自带摄像头（标清）实现实时人脸识别。测试时每一帧使用直方图均衡法平衡不同光照（比如室内灯光，自然光）影响，增加面部检测率。
- **结果**：高清上实现平均 74.6% 准确。标清和灯光条件下面部检测率从 1% 提升到 22%，自然光条件下面部检测率从 4% 到 36%，可以通过实时脸部识别判断是否为目标用户。

其他实习/项目经历

德国，慕尼黑 中国货币网城投公司信息支持性抓取 2020.01-2020.03

- **爬取网页所有内容**：因为网页异步加载，无法使用常规抓取库函数。使用 requests_html 中的相关函数解决此问题。
- **中国货币网不同页面相同 url 的区分**：因为该网页使用 ajax 动态加载技术，出现此问题。使用 google chrome 检查功能进行分析，找到相关信息的真实储存网址，并对其进行解析，解决此问题。
- **程序运行稳定**：因为网络不稳定，可能出现丢包情况。设置重确认，重连接条件解决此问题。
- **解决 PDF 文件损坏问题**：未知原因导致的 PDF 下载不完全。此问题的解决方法是，对下载好的 PDF 使用 PyPDF2 库中的 PdfFileReader 进行检查，如果出现不完整情况 (EOFError)，该文件将重新下载，直到完好或下载 5 次。损坏文件相关信息将被记录下来。
- **结果**：此爬虫可以稳定地批量爬取指定公司在该网站发布的相关财务及非财务报告，并可通过关键字对信息进行筛选，已爬取多个省份千余家城投公司目标数据，无一损坏。

技能与优势

- **编程能力**：熟练使用 Python (PyTorch)。对 C++，MATLAB，Linux，R 有一定了解。
- **机器学习相关知识**：掌握深度学习（如 CNN，RNN 等）和其他机器学习相关知识和应用（如聚类算法，EM 算法等）。
- **语言表达能力**：雅思：7.0（阅读 8.5）；CET-6：510；德福 (TestDaf)：16/20

Peisen Lin

+49-163-622-8917 | linpeishen@163.com | 1995/07/20 | male

GitHub: <https://github.com/linpeishen/i2dl/tree/master/projects>

Education

Technical University of Munich	Master of Science	Robotics, Cognition, Intelligence	2019.04-2021.04
--------------------------------	-------------------	-----------------------------------	-----------------

- Main courses: machine learning, deep learning, computer vision

Tongji University	Bachelor of Engineering	Mechatronics	2014.09-2018.09
-------------------	-------------------------	--------------	-----------------

- Main courses: Programming Language (Java)、Information Technology (C++)、Computer Networks and Communications

GPA 86.5/100

Honors/Awards: scholarship of Tongji University (Top 15%), DAAD scholarship

Project Experience

Germany, Munich	Fully Convolutional Network (FCN) for Instance Segmentation	2019.05-2019.06
-----------------	---	-----------------

- **Dataset Preprocessing:** The dataset is center cropped to 240*240. Each pixel is compared with presetting labels list then labeled. According the label of each pixel the pixel is converted to RGB then the converted picture is shown.
- **FCN Building and Training:** Based on the structure of AlexNet (The parameters are pretrained) a new Segmentation Neural Networks (NN) is built, all parameters of this NN are trained with the dataset. The functions of PyTorch are used here. The best model is saved.
- **Testing:** The best model is tested with the test dataset and the test accuracy is more than 70%.
- **Visualization:** With the functions of matplotlib.pyplot package the results of prediction and filter weights are shown.

Germany, Munich	Deep Residual Networks (ResNet) for Classification (Recurrence)	2019.11-2020.01
-----------------	---	-----------------

- **ResNet Building:** After learning the paper about the ResNet, a ResNet is built.
- **Dataset Preprocessing:** The original datasets are a 10-second video of users and lfw (Labeled Faces in the Wild). With function of dlib the faces are detected, and they are resized. Then they are spited as training, validation sets. All these sets are transformed (such as normalization, random horizon flip). With this step the network is more generalized.
- **Training, Validation, Testing:** With this GPU the ResNet is trained and validated. This process will early stop if the validation accuracy has not improved. The best model is saved and then tested with the videos captured by handy or from YouTube (HD). With the camera of Computer, the face can be real-time recognized (SD) and in this part each frame of the video is processed by equalizeHist to improve the possibility of face detection.
- **Results:** In HD the accuracy is 74.6% in average. The possibility of face detection in SD is improved from 1% to 22% with lights and from 4% to 36% with natural lighting. With the final model the face can be real-time recognized.

Other Project Experience

Germany, Munich	Information Crawling of Urban Investment Companies	2020.01-2020.03
-----------------	--	-----------------

- **All Web Content Crawling:** Because of the asynchronous loading of the web pages of chinamoney the normal function cannot be used. This problem is solved by using the functions of requests_html package.
- **Different Web Pages with the same URL Distinguish:** With the inspection of Chrome the loading process is analyzed, and the real information pages are found. This Problem is caused by ajax dynamic loading in this web.
- **Program Operation Stabling:** Rechecking and reconnecting conditions are set.
- **PDF Corruption Problem Solving:** With the function PdfFileReader the downloading PDFs are checked. If they are corrupted, they are redownloaded until they are intact or redownloaded 5 times. The corrupted PDFs are recorded.
- **Results:** This crawler can crawl the relevant reports published by the target companies on this web. In addition, it can filter the information by keywords. This crawler has crawled the target data of more than a thousand urban investment companies without corrupted PDFs.

Skills

-
- **Programming:** Skilled: Python (PyTorch); Understood: C++, MATLAB, Linux, R
 - **Knowledge in Machine Learning:** Master the knowledge of deep learning and other machine learning method (such as clustering and EM algorithms)
 - **Languages:** IELTS: 7.0 (8.5 for reading); CET-6: 510; TestDaf: 16/20