

FaceScore 项目

商业计划书

负责人: 李智敏

联系方式: 15868159657

邮箱: 707364882@qq.com

FaceScore 团队

二〇一八制

项目背景:

近年来,互联网和计算技术飞速发展,利用机器学习挖掘大数据的潜在价值具有非常广阔的应用前景。人脸识别作为机器学习技术中一个重要的分支,是模式识别研究中的热点。

由于人脸图像受环境、表情等多种变化因素的影响,导致人脸识别研究复杂而艰巨,是一项极富挑战性的研究课题。目前,还有许多的问题和关键技术有待进一步解决和完善,主要包括:

- 1.人脸特征提取阶段的完备性研究,即如何充分考虑局部特征和全局特征;
- 2.图像的平移/伸缩/旋转不变性等;
- 3.分类识别阶段的识别性能研究;
- 4.设计具有高精度识别率和快速分类的算法等。

社会需求:

随着社会的不断进步以及各方面对于快速有效的自动身份验证的迫切要求,生物特征识别技术在近几十年中得到了飞速的发展。当前的人脸识别技术主要被应用到了以下几个方面:刑侦破案,证件验证,视频监控,入口控制,表情分析。人脸识别也具有有一些有趣有益的应用,比如能够识别主人身份的智能玩具、家政机器人、具有真实面像的虚拟游戏玩家等。

具体目标:

我们以一个具体的例子说明。小明是一位事故中的受害者,现在警方请他描述犯罪嫌疑人长相。传统情况下,需要受害者或目击证人提供精确的描述,而一

般人的记忆时间非常短暂，这就使得还原犯罪嫌疑人的精确相貌非常困难。除此之外，需要画师根据描述画出最后的相貌，也对画师的技术要求提出了非常大的考验，不利于后期的筛选侦查。可见这是一个费时费力，效果也并不好的方法。如果我们使用机器学习技术中的 GAN 生成对抗网络技术，就可以预先将整个数据集进行特征化排序，提供一张标准人脸模型，只要让小明提供数个模糊特征(如鼻梁高，眼睛小，缺乏细致的程度描述)，根据贝叶斯公式，聚类检测技术就可以快速从数据库中筛选符合要求的犯罪嫌疑人，为警方办案提供极大的便利。

此外，我们还会使用随机森林，加强决策树，支持向量机等高级算法完善我们的检测和特征排序能力，同时把该技术应用在其他方面。如，根据调查网民在电商购买服装时有明星效应，都会购买时髦的，自己喜欢的服装，但无形中买家的心理会被模特所影响。我们可以通过技术修改对应的图片，让不同人看见自己最喜欢的那一部分，实现电商利润最大化。

总的来说，我们使用机器学习技术对人脸进行多特征的打分，排序，同时能根据模糊需求在库中快速匹配，进行自适应搜索同时动态调整图片。

项目投资:

因本项目之前申请了某大型上市企业内部立项，如签署此协议，将具有若干排他性条款，故略去相关细节。

来自成员的投资合计约叁万元，主要包含 GPU 计算设备的折价，相关资料的购买与必要性开支。

盈利方式与收益分配:

创业初期与试运营阶段不产生收益。

运营后我们的产品将封装为 API，在公网上对外以加密按量计费服务。具体运营方式已略去。

姓名	性别	专业	持股/收益比例
李智敏	男	16 计算机科学与技术	75%
王秀仪	女	16 计算机科学与技术	10%
吴联磊	男	15 计算机科学与技术	6%
杨思誉	男	16 计算机科学与技术	6%
朱育绮	女	16 会计学	3%

备注: 详细投资金额保密暂不公开。

团队简介:

我们团队主要由一群志同道合，有理想有追求的同学组成，他们对事认真负责，对人坦率真诚。

团队组成:

姓名	性别	专业	团队身份	负责内容
李智敏	男	16 计算机科学与技术	团队负责人	特征提取算法设计
王秀仪	女	16 计算机科学与技术	团队副负责人	数据清洗特征提取
吴联磊	男	15 计算机科学与技术	Linux 运维	进度控制 日志分析
杨思誉	男	16 计算机科学与技术	算法工程师	算法设计
周江策	男	14 广告设计	图像设计师	前端 UI 设计，图形学分析

朱育绮	女	16 会计学	财务管理	开发进度控制 财务管理
吕崇炜	男	15 计算机科学与技术	前端工程师	Web 开发后端 API 开发
郭献杰	男	15 计算机科学与技术	数据库管理	项目数据集的维护
黄诺琳	女	16 计算机科学与技术	市场顾问	需求分析人力资源

项目进度与困难:

目前我们的项目已经实现人脸区域的快速检测, 在阿里云的 GPU 服务器上, 根据已有论文复现了一个人脸区域权值计算的 Demo。能够使用卷积网络抽取低维特征, 但我们需要 GPU 计算服务器来加速我们的计算。但当前 GPU 的价格偏高, 在阿里云上一块 Tesla P4 显卡每小时的租用价格高达 15 元/19 元, 单卡训练时长达数千小时。如果我们使用私有服务器就能够大大降低我们的成本。

常见问题 QA:

Q:社会上已经有很多团队在分析和研究人脸特征了, 我们还有必要继续研究吗? 我们还能做什么?

A:有。人脸特征识别是计算机领域一个巨大无比的分支, 即便社会上现有的团队规模大, 能力强, 依然有很多空白需要填补。像物理,化学,生物等经典学科, 随着时间的推移, 不但没有消失在历史的潮流中, 反而越发达的国家越努力发展高新科技。前人的研究为我们打下了坚实的基础, 站在巨人的肩膀上, 我们可以取得进一步的突破。

Q:你们有哪些竞争优势和劣势?

A:我们所研究的项目属于互联网产业的一小部分, 正如大家印象中那样, 互联网产业附加值高, 对环境产生的附加伤害少, 作为朝阳产业, 在相对于传统产业, 如化工, 能源, 服饰, 伞业等, 我们能够在创业潮中获得更多的竞争优势。

我们的团队由敢于拼搏，有坚持有梦想的同学组成，在合作中我们不但能够学到更多知识，也能为社会注入新鲜力量。

我们无法取得社会团队一样的大量资金支持，用于购买服务器对我们的网络进行训练。同时，因为机器学习的门槛较高，给我们带来较高的实践难度。

发展目标：

我们会将训练好的神经网络封装成多个 API，在公网上实现加密按量计费，客户按需购买我们的服务。因为项目性质的特殊性，人脸头像图片数据集涉及了各方隐私，所以我们需要谨慎处理。在积累一定经验之后，我们会申请知识产权保护和工商注册，寻找更多企业客户和我们合作。

开发进度安排：

- ✧ 2018.4 因为机器学习的门槛非常高，所以我们目前以及未来一段时间内需要复现已有论文伪代码的 demo。学习 TensorFlow, MXNet, Caffe 等神经网络设计，原理等知识。
- ✧ 2018.4-2018.7 前我们会完成整个数据集的清洗，特征化。
- ✧ 2018.7-2018.9 前我们会实现整个 CNN 网络的大部分设计和调试。
- ✧ 2018.07-2018.12：相关网站的开发，API 封装，研究报告的撰写。申请知识产权保护并在工商注册登记，寻找企业客户合作和投资。

2019.01-2019.03：平台正式进入公测，同时不断修正 bug，产生正面社会效益。

备注：如您还有更多问题，请致电团队负责人。