

附件 2

中山大学 大学生创新训练计划项目申请书

项目名称: 基于计算机视觉的智能云相册应用

相关学科: 计算机科学与技术

申 请 人: 黎建忠、徐通、韦政源

所在院系: 移动信息工程学院

指导教师: 王昌栋

预期成果类别:

☐ 社科类社会调查报告及学术论文

☒ 自然科学类学术论文

教 务 处 制

填表时间 二〇一五年十月八日

项目名称	基于计算机视觉的智能云相册应用										
申请人或申请团队	姓名 (负责人请加*号)	学号		性别	身份证号		年级	专业	联系电话	手机	E-mail
	*黎建忠	13354146		男	441881199508261937		2013	软件工程	13750022107	13750022107	lijzh9@mail2.sysu.edu.cn
	徐通	13354374		男	13112719941173878		2013	软件工程	13149309741	13149309741	tt_leader@hotmail.com
	韦政源	13354347		男	445322199410080413		2013	软件工程	13631216505	13631216505	weizhy3@hotmail.com
指导教师	姓名	性别	年龄	职称	职务		所属院系	联系电话	手机	E-mail	
	王昌栋	男	31	讲师	讲师		移动信息工程学院	0756-3668567	13450460216	wchangd3@mail.sysu.edu.cn	
研究起止时间		2015 年 10 月至 2016 年 10 月									
一、申请理由 (包括自身具备的知识条件、自己的特长、兴趣及开展本项目的基础等)											
<p>团队由移动信息工程学院三名 2013 级学生组成。在前两年的软件设计、数据结构、算法设计等课程中，我们打下了扎实的理论与实践基础，对软件工程领域充满了热情。团队成员富有想象力与创造力，在不同的细分领域各有所长。其中：</p> <ol style="list-style-type: none"> 1. 黎建忠同学热衷钻研开源技术，具备 Android 开发的经验，熟悉移动应用的设计规范，对于面向对象程序设计和 Web 开发也有着不错的理解和认识，能根据需求快速阅读相关资料并解决问题。 2. 徐通同学是在任的微软校园精英，对于微软的 Azure 云服务平台有一年的使用经验，熟悉云应用后端平台的搭建，同时还是移动信息工程学院本科生科研项目（URP）的成员，跟随老师在机器学习和计算机视觉领域进行学习，具备一定的研究经历，对于人脸识别和图像分类领域也有一定的经验和基础。 3. 韦政源同学熟悉网络编程与常见网络通信协议，曾搭建过校内公用 FTP 服务器和具备内容管理系统的博客平台，可以熟练地实现云服务与手机终端之间的网络通信，确保手机应用在云服务器上有良好的后端支持。 <p>综上所述，团队成员的专业知识丰富，实践能力强，对移动互联网和云计算的未来持乐观态度。此外，通过学院设置的课程实验与工程实训等项目的训练，我们能适应高强度的学习与工作，拥有良好的英文资料研读能力，可以确保本项目的圆满完成。</p>											
二、项目创新特色概述(50 字以内)											

1. 云端与移动终端结合，对照片进行智能识别分类
2. 提供了更人性化、用户友好程度更高的照片浏览和分享方式

三、计划项目实施思路（3000 字以内）

1. 研究意义与目的，同类研究工作国内外研究现状与存在的问题等

① 研究意义与目的：

智能手机在现代生活中扮演着十分的重要角色。在出游、聚会、演出等许多场合，人们常常使用智能手机拍照，将这些美好的记忆永久珍藏。然而随着照片的增多，这些照片的整理、检索和分享也变得越来越麻烦。

考虑到屏幕尺寸以及在手机上进行文件操作的不便，为了整理这些照片，人们不得不将照片导入到计算机，再对它们进行人工分类、排序和标记。这一过程往往是枯燥而繁琐的，也需要耗费大量的时间和精力。日后如果想要查找某张特定的照片，无论是桌面端还是移动端的照片管理应用，它们中的绝大多数都没有提供以照片内容为基础的语义化搜索功能。这也就意味着，在最坏的情况下，用户必须手动翻遍所有相册才能找到想要的照片。这是一个移动互联网的时代。“分享”是移动互联网的重要主题。如何精简照片从拍下到被分享出去的这一系列操作的步骤，提供更人性化，用户友好程度更高的照片分享模式，也是业界关注的热点之一。

因此，我们拟开发一款基于人脸识别与计算机视觉的Android云相册应用，帮助智能手机理解图片内容，并以此为基础，对图片进行自动分类整理，如分成肖像、风景、建筑等，方便用户检索和浏览。对于肖像类图片，可以进一步识别出人脸，关联用户的通讯录与社交账号，方便分享。图片的智能识别需要大量的计算资源，适合在云端服务器上完成。云端与移动终端的结合可以高效地完成识别任务，将极大地提升智能手机上该类需求的用户体验。

② 国内外研究现状与存在的问题：

通过对各大智能手机应用市场的调研与观察，我们发现，面对海量的照片，现有的绝大多数相册应用却只提供了基本的展示、编辑、以及按日期简单排序的功能，并未解决上文中提到的诸多痛点。手机的性能不断提升，照片的管理方式却没有得到应有的优化，有悖于科技以人为本的理念。

近年来人工智能、计算机视觉与云计算等领域发展迅猛，由著名学者Yann LeCun, Andrew Ng推崇的深度学习在最近三年更是受到了学术界和工业界的热捧，该细分领域中也先后涌现了RCNN区域特征卷积神经网络（文献[1]），SPP空间金字塔聚合（文献[2]）等优秀的算法。除此之外，由加州伯克利大学贾杨清博士领导开发的Caffe开源深度学习框架与微软公司的Project Oxford所提供的API与SDK等一系列开发工具则大大地降低了这些领域的开发门槛。

国内外的不少研究都致力于将人工智能与计算机视觉等领域的相关理论和技术应用于图片的智能识别与分类上。文献[3]提出了一种基于双层支持向量机（two-layered support vector

machine, SVM) 分类器的语义化图片分类算法。其中第一层根据图片的元数据 (metadata) 和局部的底层视觉特征来估计图片语义似然 (likelihood), 第二层则以此进一步得出全局的图片语义。文献[4]则在文献[3]的基础上, 进一步构建了一个手机相册系统, 使得用户可以根据照片所描述的事件、人物以及所摄内容等语义上有意义的类别来对照片进行浏览。但是该系统的局限性在于分类模型的训练必须预先在离线的PC机上完成, 然后才人工将训练数据迁移至移动终端, 用于照片分类。囿于当时 (2008年) 的硬件成本和移动互联网的普及程度, 该系统也未能考虑用户的社交与分享需求。

在本项目中, 我们拟开发的智能云相册应用可以帮助用户自动备份照片到云存储空间, 同时在云端实时进行智能识别, 得到的识别结果则通过网络返回移动终端, 伴随友好的界面呈现给用户。除此之外, 通过关联用户的通讯录与社交账号, 配合人脸识别结果, 一种更人性化、用户友好程度更高的照片分享模式也就应运而生。

③ 参考文献:

- [1] Ross Girshick, Jeff Donahue, Trevor Darrell, Jitendra Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation", CVPR 2014
- [2] He, K., Zhang, X., Ren, S., & Sun, J. (n.d.). "Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition," Computer Vision - ECCV 2014 Lecture Notes in Computer Science, 346-361.
- [3] Yang, S., Kim, S., & Ro, Y. (n.d.). "Semantic Home Photo Categorization," IEEE Trans. Circuits Syst. Video Technol. IEEE Transactions on Circuits and Systems for Video Technology, 324-335.
- [4] Sang-Kyun Kim, Jae-Won Lee, Ryong Lee, Eui-Hyeon Hwang, and Min Gyo Chung, "User-Friendly Personal Photo Browsing for Mobile Devices," ETRI Journal, vol. 30, no. 3, June. 2008, pp. 432-440.

2. 研究内容及工作方案

研究内容:

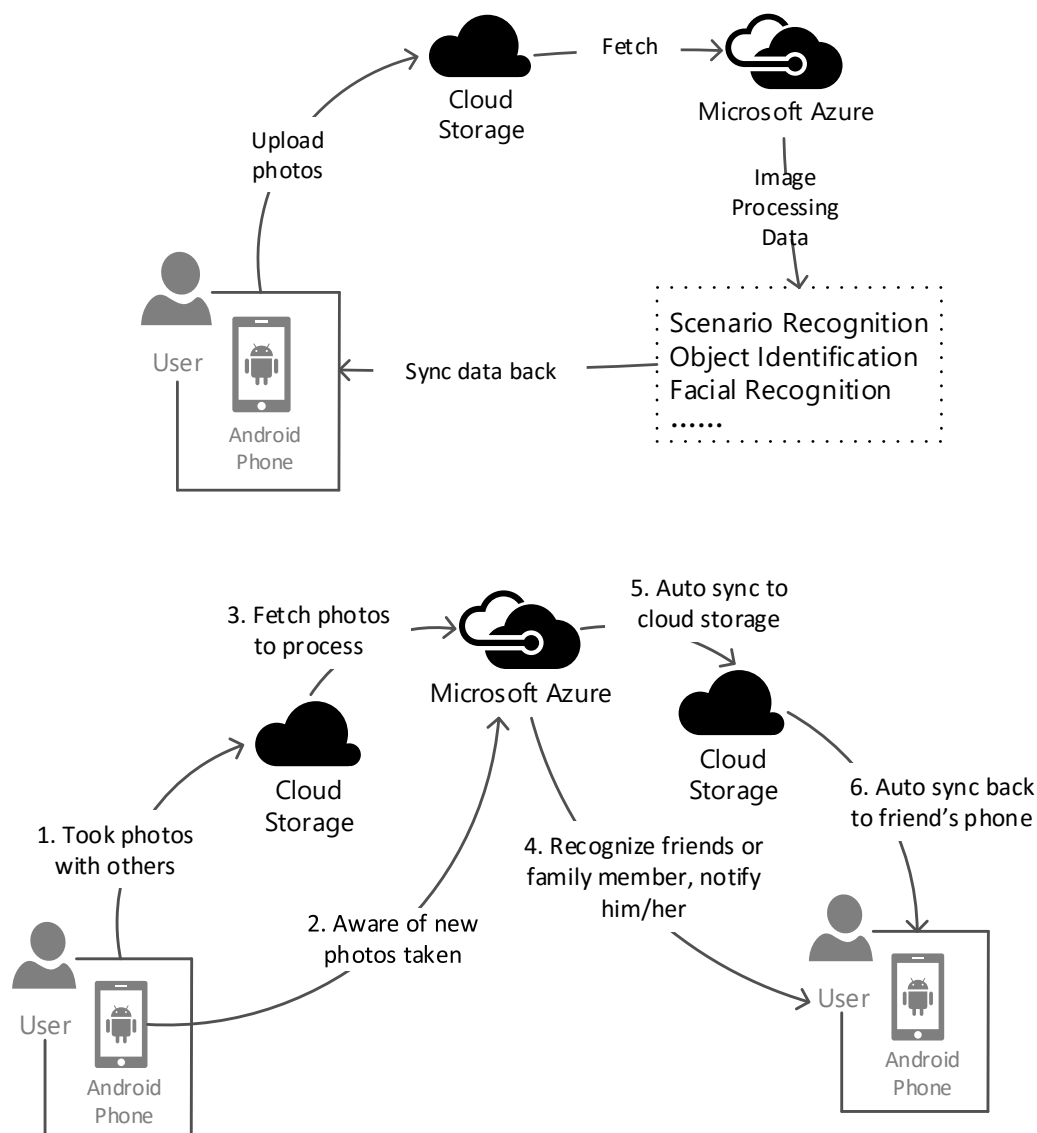
- 1) 熟悉Google Material Design的设计准则, 打磨出友好的用户界面 (User Interface)。
- 2) 深入理解Android平台的文件管理机制, 提供简单、直观的照片浏览交互方式。
- 3) 深入理解云计算的基本模式, 将云端与移动终端结合, 充分发挥云计算的优势。
- 4) 了解人脸识别、计算机视觉等领域的常用算法, 熟悉相关框架与代码库的使用。
- 5) 熟悉云平台的开发环境, 以及掌握常用的 Web 框架和各种开源代码库的使用。
- 6) 理解敏捷软件开发 (Agile Software Development) 的准则, 确保团队高效合作。

工作方案：

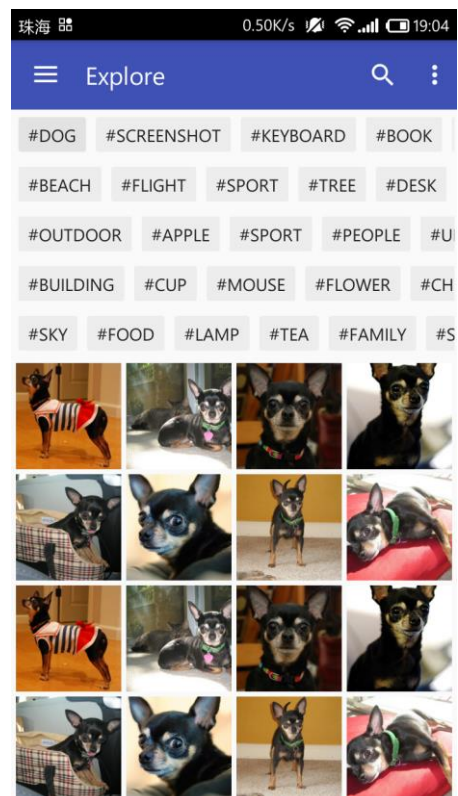
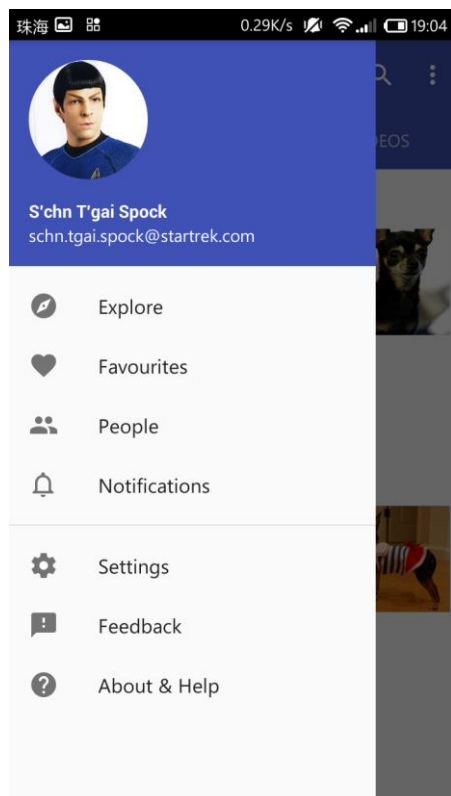
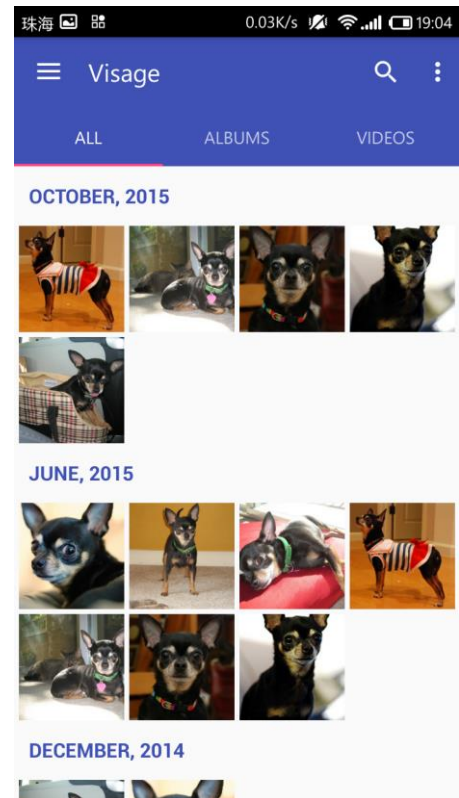
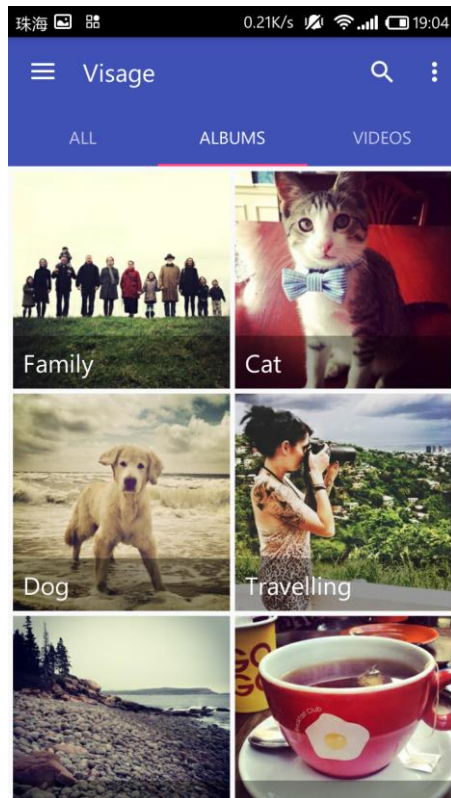
我们拟开发一款基于人脸识别与计算机视觉的智能云相册应用，具备包括但不限于以下所列举的特性（Features）：

- 1) 在网络条件允许的前提下，将手机上的图片自动备份至云端存储空间。
- 2) 在云端对海量图片进行计算机视觉处理，自动分类整理出肖像、风景、建筑等类别，打上与照片内容相关的关键词标签，方便进行搜索。
- 3) 对于肖像类图片，进一步识别出人脸，关联用户的通讯录与社交账号，配合社交账号的分组信息对图片进行自动分组。
- 4) 以上的识别结果将同步至移动终端，伴随友好的界面呈现给用户，并提供简单易用的照片整理和分享方式。

应用架构如下：



交互界面示意:



3. 拟解决主要问题

- 1) 人脸识别与计算机视觉处理结果的准确率和效率问题
- 2) 移动终端与云端的隐私保护与数据安全问题
- 3) 不同网络环境下，识别数据与照片的同步问题

4. 研究支持条件（使用仪器或设备）

本项目是在指导老师王昌栋的指导下，以云计算应用开发课程大项目为依托开展。团队成员的专业知识丰富，实践能力强，对移动互联网和云计算的未来持乐观态度。

项目中图片的智能识别需要大量的计算资源，适合在云端服务器上完成。联网的移动应用也需要性能良好的云端服务器作为后端支持。团队拟租用微软公司的Azure Cloud云平台的服务，以支持云端的开发，费用将出自项目预算资金。

除此之外，应用的设计、开发、测试等过程中需要用到部分商业软件，需要购买付费授权。这部分费用也以教育优惠价格计入项目预算。

5. 预期研究成果（研究论文、设计、专利、产品、鉴定、推广应用等）

- 1) 一款基于人脸识别和计算机视觉的智能云相册应用，在各大应用市场上架；
- 2) 本项目的开发文档与研究报告，并投稿至相关领域的专业期刊；
- 3) 对项目中的创新性成果申请国家专利。

6. 项目创新之处（原始创新：重大科学发现、技术发明；集成创新：融合多种相关技术，形成新产品、新产业；引进消化吸收再创新：在引进国内外先进技术的基础上，学习、分析、借鉴，形成具有自主知识产权的新技术）

智能手机在现代生活中扮演着十分重要的角色。而手机上海量照片的整理与分享问题，是长期以来都没有得到很好解决的痛点。本项目的特色与创新点在于：

- 1) 采用人脸识别与计算机视觉的相关技术，可以对海量图片进行处理，自动根据图片内容进行分类，如分成肖像、风景、静物等。对于肖像类图片，还可以进一步识别出人脸，方便分享与整理。与仅能根据日期进行分类的其他应用相比，本项目的分类机制和交互模式显然更符合用户的直觉与习惯。
- 2) 针对移动终端计算能力、电量和存储空间都有限制的特点，我们将图片的识别处理过程放在了云端，充分利用了云计算的优势，以达到提高识别效率、节省手机存储空间和降低手机功耗的目的。
- 3) 遵循Google官方的Material Design设计准则。精心设计的用户界面是影响使用体验的直接决定因素。好的设计风格可以让用户在使用的全过程都充满了愉悦与惊喜。

四、计划项目管理（800 字以内）

1、计划项目人员分工

黎建忠：项目管理，撰写文档，用户界面设计，实现 Android 前端的功能需求。

徐 通：搭建云端开发环境，计算机视觉处理与人脸识别，优化识别效率。

韦政源：进行云平台后端开发，解决移动终端与云端的通信与同步问题。

2、计划项目研究时间安排（查阅资料 、选题、项目研究方案、开题报告、实验研究、数据统计、处理与分析、研制开发、结题、撰写研究论文和总结报告、参加结题答辩和成果推广等）

2015 年 10 月-12 月：应用原型设计，搭建本地与云端的开发环境，细化团队成员分工。

2016 年 1 月-6 月：团队成员根据分工协作开发，逐步实现应用预期的基础功能与特性。

2016 年 7 月-9 月：优化用户界面，功能特性测试，视项目进度调整团队成员分工。

2016 年 9 月-10 月：后期测试，撰写开发文档与研究报告，应用在各大应用市场上架。

五、经费预算及使用计划（材料费、资料费、版面费、鉴定费、专利费等） 总计：10000 元人民币

具体预算：

支出明细科目	金额(元)	计算依据和明细
印刷费	300	打印论文与参考资料
邮费	200	日常邮费开支
差旅费	1000	项目前期调研，与合作成员交流所需
市内交通费	300	日常交通费开支
材料费	700	开发设备一年的维护费用
办公用品	6500	购买微软 Azure Cloud 云计算服务:3000 元 购买用于真机测试的开发用机：2000 元 购买其他必需专业软件，GtiHub 私有项目托管服务，注册 Google Play 开发者账号等软件服务：1500 元 (以上均以教育优惠价格计算)
资料费	1000	购买相关专业图书一批
合计	10000	

指导教师意见	<div>签名:</div> <div>年 月 日</div>	院系意见	<div>签名(公章):</div> <div>年 月 日</div>
学校审核意见	<div>签名(公章):</div> <div>年 月 日</div>		
备注			