**软件需求分析与开发计划**

**题目：面向深度学习的图像识别和目标跟随系统设计**

**项目提出者： 赵振刚**

**项目成员： 程欣 石强 齐昱博 万璐敏**

**撰 写 人： 石强**

# 项目背景

随着安全入口控制和金融贸易方面应用需要的快速增长，生物统计识别技术得到了新的重视。目前，微电子和视觉系统方面取得的新进展，使该领域中高性能自动识别技术的实现代价降低到了可以接受的程度。而人脸识别是所有的生物识别方法中应用最广泛的技术之一，在国外，人脸识别技术早已被大量使用在国家重要部门以及军警等安防部门。在国内，对于人脸识别技术的研究始于上世纪90年代，目前主要应用在公安、金融、网络安全、物业管理以及考勤等领域。

# 人脸识别技术的研究意义

人脸识别是机器视觉和模式识别领域最富有挑战性的课题之一，同时也具有较为广泛的应用意义。人脸识别技术是一个非常活跃的研究领域，它覆盖了数字图像处理、模式识别、计算机视觉、神经网络、心理学、生理学、数学等诸多学科的内容。

同时，人脸识别是生物特征识别技术的一个重要方向，与其他生物特征相比，人脸识别具有主动性、非接触性和用户友好等许多优点，多年来一直受到许多研究者的关注。但是目前的人脸识别方法主要集中在二维图像方面，由于受到光照、姿势、表情变化的影响，识别的准确度受到很大限制，脸像随年龄变化，而且容易被伪装。因此，人脸识别技术目前只能用于某些识别准确率要求不是很高的场合。

基于视觉通道信息的面部感知系统，包括人脸检测和跟踪、面部特征定位、面部识别、人脸归类（年龄、种族、性别等的判别）、表情识别、唇读等分系统，如图1所式，可以看出，继人脸检测和跟追之后，面部特征定位通常是面部感知的一个必备环节，是后续工作的基础，具有重要的意义。尽管人脸识别不能说是其他面部感知模块的必备功能，但是，可以肯定的是，利用已知的身份信息，结合特定人的先验知识，可以提高表情分析、唇读和语音识别、手势识别乃至手写体识别的可靠性。而计算机对使用者身份确认的最直接的应用就是基于特定使用者的环境设置：如使用者的个性化工作环境，信息的共享和隐私保护等等。

# 人脸识别的国内外发展概况

作为一种最直接、最自然、最容易被人接受的生物特征识别技术，人脸识别致力于探索如何使机器能够自动地根据用户的人脸图像来鉴别用户的身份。人脸识别发展到现在，已有30多年的历史，是模式识别和计算机视觉的一个非常活跃的研究热点。人脸识别的研究涉及模式识别、计算机视觉、人工智能、图像处理、心理学、生理学和认知科学等，与计算机人机交互领域和基于其它生物特征的身份识别方法都有密切联系。虽然人们可以毫不费力地通过脸部图像来鉴别互相的身份，然而由于成像过程中各种影响因素的变化常常导致同一个人的人脸图像发生非常大的变化，因此建立自动系统完成识别任务是非常具有挑战性的。虽然目前国内外已经有许多实用系统问世，但是只有在非常苛刻的成像条件下，才能得到比较令人满意的识别效果。因此，人脸识别研究仍然远远没有到达完善的境地，还有非常大的发展空间。

## 国外的发展概况

1893年，Bertillon采用语句描述方法对人脸分类。

20世纪60年代，Bledsoe提出了人脸识别的半自动系统模型与特征提取方法。

70年代，美、英等发达国家开始重视人脸识别的研究工作并取得进展。1972年，Harmon用交互式人脸识别方法在理论上与实践上进行了详细的论述。就在这一年，Sakai设计了人脸图像自动识别系统。

80年代初T. Minami研究出了优于Sakai的人脸图像自动识别系统。

90年代，由于计算机技术、数字图像处理、模式识别技术的发展，加上人们对人脸图像自动识别的迫切需求，各国军、警方及有关部门高度重视，大公司鼎力相助，对它的研究变得非常热门。

1996年美国军方组织了人脸自动识别系统大赛，勒克菲勒大学的Facelt系统获得冠军。最近，美国的LAU公司研制的人脸图像自动识别系统，是以人眼辨别人的原理，基于生物测量学、人像复原技术开发的装置。用人脸12--42个特征点，对人群中寻找的人进行定量定性识别，已经用在机场、火车站、公共场所、重点控制地区。

得益于深度学习的出现，人脸识别在过去的三四年取得了很快的发展。在计算机视觉领域，深度学习中应用得最好、最成功的就是卷积神经网络（CNN）。神经网络的最大价值在于对特征的自动提取和抽象，它免去了人工提取特征的繁琐，可以自动找出复杂且有效的高阶特征。这一点类似于人的学习过程，先理解简单概念，再逐渐递进到复杂概念，神经网络每加深一层，可以提取的特征就更抽象。

LFW数据集（Labeled Faces in the Wild）是目前用得最多的人脸图像数据库。该数据库共13，233幅图像，其中5749个人，其中1680人有两幅及以上的图像，4069人只有一幅图像。图像为250\*250大小的JPEG格式。绝大多数为彩色图，少数为灰度图。该数据库采集的是自然条件下人脸图片，目的是提高自然条件下人脸识别的精度。目前，人工在该数据集上的准确率在0.9427~0.9920。在该数据集的第六种评价标准下（无限制，可以使用外部标注的数据），许多深度学习方法已经赶上（超过）人工识别精度，比如face++，DeepID3，FaceNet等。

## 国内的发展概况

人脸识别系统如今在大多数领域中起到举足轻重的作用，尤其是用在机关单位的安全和考勤、网络安全、银行、海关边检、物业管理、军队安全、智能身份证、智能门禁、司机驾照验证、计算机登录系统。

我国在这方面也取得了较好的成就，国家863项目“面像检测与识别核心技术”通过成果鉴定并初步应用，就标志着我国在人脸识别这一当今热点科研领域掌握了一定的核心技术。北京科瑞奇技术开发股份有限公司在2002年开发了一种人脸鉴别系统，对人脸图像进行处理，消除了照相机的影响，再对图像进行特征提取和识别。这对于人脸鉴别特别有价值，因为人脸鉴别通常使用正面照，要鉴别的人脸图像是不同时期拍摄的，使用的照相机不一样。系统可以接受时间间隔较长的照片，并能达到较高的识别率，在计算机中库藏2300人的正面照片，每人一张照片，使用相距1--7年、差别比较大的照片去查询，首选率可以达到50%，前20张输出照片中包含有与输入照片为同一人的照片的概率可达70% 。

2005年1月18日，由清华大学电子系人脸识别课题组负责人苏光大教授主持承担的国家“十五”攻关项目《人脸识别系统》通过了由公安部主持的专家鉴定。鉴定委员会认为，该项技术处于国内领先水平和国际先进水平。

# 系统的需求分析及方案选择

人脸识别系统现在应用于许多领域中，但是人脸识别技术也是一项近年来兴起的，且不大为人所知的新技术。在我国以及其他国家都有大量的学者正在研究之中，不断的更新人脸识别技术，以便系统的识别准确率达到新的高度。

## 需求分析

人脸识别系统主要包括四个组成部分，分别为：人脸图像采集及检测、人脸图像预处理、人脸图像特征提取以及匹配与识别。

人脸图像采集及检测：不同的人脸图像都能通过摄像镜头采集下来，比如静态图像、动态图像、不同的位置、不同表情等方面都可以得到很好的采集。当用户在采集设备的拍摄范围内时，采集设备会自动搜索并拍摄用户的人脸图像。

人脸图像预处理：对于人脸的图像预处理是基于人脸检测结果，对图像进行处理并最终服务于特征提取的过程。系统获取的原始图像由于受到各种条件的限制和随机 干扰，往往不能直接使用，必须在图像处理的早期阶段对它进行灰度校正、噪声过滤等图像预处理。

人脸图像特征提取：人脸识别系统可使用的特征通常分为视觉特征、像素统计特征、人脸图像变换系数特征、人脸图像代数 特征等。人脸特征提取就是针对人脸的某些特征进行的。人脸特征提取，也称人脸表征，它是对人脸进行特征建模的过程。

人脸图像匹配与识别：提取的人脸图像的特征数据与数据库中存储的特征模板进行搜索匹配，通过设定一个阈值，当相似度超过这一阈值，则把匹配得到的结果输 出。人脸识别就是将待识别的人脸特征与已得到的人脸特征模板进行比较，根据相似程度对人脸的身份信息进行判断。这一过程又分为两类：一类是确认，是一对一 进行图像比较的过程，另一类是辨认，是一对多进行图像匹配对比的过程。

## 开发环境需求分析

### 硬件平台

联想E420 Intel Core i3-2330M CPU 2.20GHz×4

### 开发环境

Ubuntu 14.04 64-bit

### 开发工具

Google深度学习框架Tensorflow 1.0，计算机视觉库OpenCV，Python包管理工具Anaconda 3，代码编辑器Sublime 3

具体到人脸检测的问题，在深度学习技术出现之前，几乎所有人脸检测都是采用滑动窗口式的方法。通过设计分类器，检测窗口区域是否为人脸。此外还需要对图片进行缩放再进行检测。这是2000-2012年以来主流的方法。

神经网络的最大价值在于对特征的自动提取和抽象，它免去了人工提取特征的繁琐，可以自动找出复杂且有效的高阶特征。这一点类似于人的学习过程，先理解简单概念，再逐渐递进到复杂概念，神经网络每加深一层，可以提取的特征就更抽象。在本课题中正是使用FaceNet模型来解决人脸识别的问题，其中FaceNet是由训练集：MS-Celeb-1M （10575个人共计453453张人脸图像）、网络模型（Inception-ResNet-v1）、Loss function（triplets loss）、预处理（人脸对齐Face Alignment：MTCNN）测试：（LFW）。

选择深度学习方法的原因：

第一，最重要的因素是深度学习可以做到传统方法无法企及的精度。这是关键中的关键。许多人脸识别的深度学习方法，如face++，DeepID3，FaceNet等，在LFW数据集上的识别精度已经赶上甚至超过人工识别精度。

第二，深度学习算法的通用性很强。在传统算法里面，针对不同的物体的目标检测需要定制化不同的算法。相比来看，基于深度学习的算法更加通用，比如faster RCNN在人脸、行人、一般物体检测任务上都可以取得非常好的效果。

第三，深度学习获得的特征（feature）有很强的迁移能力。所谓特征迁移能力，指的是在A任务上学习到一些特征，在B任务上使用也可以获得非常好的效果。例如在ImageNet（物体为主）上学习到的特征在场景分类任务上也能取得非常好的效果。

第四，工程开发、优化、维护成本低。深度学习计算主要是卷积和矩阵乘，针对这种计算优化，所有深度学习算法都可以提升性能。另外，通过组合现有的层（layer），可以实现大量复杂网络结构和一些算法，开发维护的成本低。