



**二 〇 一 九 届 毕 业 设 计**

**开题报告**

**学 院：信息工程学院**

**专 业：计算机科学与技术**

**姓 名：宋启迪**

**学 号：201524020401**

**指导教师：南春丽**

**二〇一九年三月**

长安大学毕业设计开题报告表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 课题名称 | 基于TensorFlow框架的手写数字识别系统 | | | | |
| 课题来源 | 自选课题 | 课题类型 | 专题研究 | 指导教师 | 南春丽 |
| 学生姓名 | 宋启迪 | 学 号 | 201524020401 | 专 业 | 计算机科学与技术（卓越工程师） |
| 1. **课题背景及意义：**   自从AlphaZero击败人类顶尖围棋高手后，人们关于人工智能的讨论就喋喋不休，从而出现了人工智能有益论和人工智能有害论两派。有许多人人认为人工智能会使一大批人失去他们所从事的行业，无情的被机器人取代。而更多人则认为，人工智能可以帮助我们创造更多的财富，创造更美好的明天。而目前我所能接触到的人工智能技术，都为人类的发展做出了极大的贡献，人类不能因噎废食。例如在图像识别领域，使用Google的TensorFlow人工智能框架可以很方便的搭建一个卷积神经网络，并且立即投入使用。  手写数字识别重大研究价值意义在于人机交互方向，提高人机交互的流畅性、方便性以及似人性。在于数字信息自动处理领域，提高处理效率，节省物力财力，加快信息传输，创造出无数的财富。手写体是人类在交流中使用的最常见的标准和常规媒介。即使引入了诸如键盘，声音命令等新技术，它也是一种有效且高效的信息记录方式。通常，手写识别系统是一种用于识别人类手写的任何语言的机制，无论是扫描的手写图像或者在电子设备上使用手写笔进行实时手写，也可以分别称为离线和在线手写[1]。此外，该系统的应用可分为数字，字符和英文字母三种。它广泛用于许多应用，如语言翻译，银行支票和关键字定位。手写数字识别系统的理论价值如下:  1)阿拉伯数字官方符号，被世界各国通用，且应用于生活的方方面面。因此研究手写数字识别具有很高的应用价值。  2)相较于汉字、英文单词等常用字符，数字识别的类型数较小，有助于做深入分析以及验证一些新的理论。  3)虽然人们已经对模式识别的手写数字领域进行了相当长时间的研究，但是手写数字识别系统仍然有很多不足，在某些情况下，还无法与人类的识别能力相提并论，人工智能还有很长的路要走。  4)手写数字识别与许多领域都有关联性，尤其是中英文拼音的识别。事实上，很多研究人员就是同时研究这两者的，由于相似性较强，结合在一起研究更易取得研究成果。  在大规模的数据统计中(如:人口普查等)需要输入极大量的数据，之前则完全要靠手工输入。而近几年来在这类任务中采用手写识别的技术已成为一种趋势。因为数据经常会大规模集中录入系统，所以可以在数字录入时加一些书写限制条件，或者在用户手写数字输入时给出限制提示，从而避免手写体数字的千差万别[2]，得到良好的手写数字源，用于机器学习和自动识别。当前绝大部分操作系统，都会在用户录入数据时增加限制条件，让用户按照规范进行手写数字，另外还采用合适的图形用户界面对识别结果做全面的检查，最终保证正确无误。  银行、表格、证券是手写数字识别可以大放异彩的又一个领域。随着我国经济的迅速发展，无时无刻不产生大量的账单、支票、税务报表等等手写数据[3]。与上面提到的统计报表[4]处理相比，在这个领域的应用难度更大，因为对识别的精度要求更高，处理的表格种类更多等，这样对识别及预处理的核心算法要求也大大提高了。如果在这个方向上都能做到机器自动识别的话，就能够取得巨大的经济效益。 | | | | | |
| 1. **国内外发展状况：**   随着科学技术与发展, 计算机视觉已经广泛运用于图像理解、地图导航、医疗制药、无人机和无人驾驶汽车等领域。而这些应用的核心技术就是图像处理[5]、图像定位[6]和图像分类[7]等视觉识别任务。  模式识别是六十年代初迅速发展起来的一门学科。由于它研究的是如何用机器来实现人(及某些动物)对事物的学习、识别和判断能力，因而受到了很多科技领域研究人员的注意，成为人工智能研究的一个重要方面。一个模式识别系统的基本职能是对系统所要处理的模式归属于哪一类做出判别，从该系统的模式输入到系统做出判别之间，主要包括信息检测、预处理、特征提取和分类几大环节。  字符识别是模式识别领域中的一个非常活跃的分支。一方面是由于问题本身的难度使之成为一个极具挑战性的课题;另一方面，是因为字符识别不是一项孤立的应用技术，其中包含的模式识别领域中其他分支都会遇到的一些基本和共性的问题。从50年代开始，许多的研究者就在这一研究领域开展了广泛的探索并为模式识别的发展产生了积极的影响。   字符识别，从采用的输入设备来分，可分为脱机识别(又称为光学字符识别 Optical Character Recognition )OCR和联机识别，脱机字符又分为印刷体和手写字符识别，从对书写者要求来分，手写字符又分为限制性和非限制性的手写字符识别。在联机手写字符识别中，计算机能够通过与计算机相连的手写输入设备获得输入字符笔划的顺序.笔划的方向以及字符的形状，所以相对0CR来说它更容易识别一些。但联机字符识别有一个重要的不足就是要求输入者必须在指定的设备上书写，然而人们在生活中大部分的书写情況是不满足这一要 求的，比如人们填写各种表格资料，开具支票等。如果需要计算机去认识这些已经成为文字的东西，就需要OCR技术。比起联机字符识别来，OCR 不要求书写者在特定输入设备上书写，它可以与平常一样书写，所以OCR的应用更为广泛。OCR所使用的输入设备可以是任何一种图像采集设备，如扫描仪、数字相机等。通过使用这类采集设备，OCR系统将书写者已经写好的文字作为图像输入到计算机中，然后由计算机去识别。由于OCR的输入只是简单的一副图像，它就不能像联机输入那样比较容易的从物理上获得字符笔划的顺序信息，因此OCR是一个更具挑战性的问题。  脱机字符识别(OCR) 分为印刷体OCR和手写OCR。印刷体字符比手写体字符少了随机性，它的识别相对容易些，难点已经不在识别环节，而在于字符的分割上。印刷体识别的错误绝大多数都是错误的分割引起的。对于手写体0CR, 无论是联机还是脱机识别，手写体的识别都要经历由限制性手写体识别到非限制性手写识别体两个阶段。本文将已手写体数字为代表，讨论非限制性手写体字符的识别。脱机字符是别的研究最早始于上个世纪六十年代，是为了应付汉英翻译的需要。八十年代后的研究中心转移到脱机手写字符的识别上。对于小类别数的字符集如数字、字母的识别，已经可以做到对书写不加任何的限制。非限制性手写OCR的研究始终以阿拉伯数字为主导。这事因为，第一，十个阿拉伯数字是全世界的一套通用字符。第二，在数字的许多应用场合，如报表、账单、支票等，手写体还难以被印刷体所替代，而且对识别的可靠性要求极高。三，由于类别数少，所以模式识别中的许多方法研究均可以以数字识别作为实验背景。对脱机手写体字符的研究，人们由简单集成笔画密度、笔画方向和背景特征方法过渡到特征匹配方法，进而过渡到结合神经网络方法，随着对识别可靠性要求的提高，九十年代以后，多分类器集成方法成为了一个研究重点。 | | | | | |
| 1. **课题研究的主要内容：**   深度学习是当下的研究热点之一，目的在于建立、模拟人脑进行分析学习的神经网络，模仿人脑的机制来解释图像、声音和文本等数据，而且许多研究成果，都离不开对大脑认知原理的研究，尤其是视觉原理的研究。美国神经生物学家发现了视觉系统的分级信息处理方式，即可视皮层是分级的。人类的视觉从原始信号摄入开始，瞳孔摄入像素，接着大脑皮层某些细胞发现边缘和方向，然后抽象判定眼前的物体的形状，然后进一步抽象来判断眼前的物体是何物。  因此根据大脑分层逐级认知的这个特点，法国科学家Yann LeCun提出Convolutional Neural Networks卷积神经网络（CNN）并应用在手写字体识别上（MNIST）。  CNN是一种多层神经网络，用滤波器进行特征抽取，使用卷积核作为特征抽取器，并自动训练特征抽取器，就是说，卷积核以及阈值等参数都需要由网络去学习优化，通过大量的训练数据来不断优化提升，而且图像可以直接作为网络的输入，避免了传统的识别算法中复杂的特征提取和数据重建过程。  CNN的基本架构包含三个主要的层——卷积层、池化层和全连接层。卷积层用来学习输入数据的特征，由很多的卷积核组成；池化层降低卷积层输出的特征向量，典型应用有平均池化和最大池化；全连接层将卷积层和池化层堆叠起来以后，就能够形成一层或多层全连接层，实现高阶的推理能力。   1. **课题的研究方法和手段：**   多年的研究实践表明，对于完全没有限制的手写数字，几乎可以肯定:没有一种简单的方案能达到很高的识别率和识别精度，因此，最近这方面的努力向着更为成熟、复杂、综合的方向发展。研究工作者努力把新的知识运用到预处理，特征提取，分类当中。近年来，人工智能中专家系统方法、人工神经网络方法已应用于手写数字识别。在手写数字识别的研究中，神经网络技术和多种方法的综合是值得重视的方向。针对模式特征的不同选择及其判别决策方法的不同，可将模式识别方法大致分为5大类[8],这5种识别方法均可实现手写数字识别，但它们特点不同，必须根据条件进行选择。   (1)统计模式法  这是以同类模式具有相同属性为基础的识别方法。用来描述事物属性的参量叫做待征，它可以通过模式的多个样本的测量值统计分析后按一定准则来提取。例如:在手写数字识别系统中，我们可以把每个数字的图形分为若干个小方块  (2)句法结构方法  在形式语言和自动机的基础上产生了句法结构这一方法。其基本原理是:对每一个模式都用一个句法来表示，而对一个待识别的未知样本，通过抽取该样本的基元来构造该样本的句子，然后分析此句子满足什么样的句法，从而推断出他该属于哪个模式类。这种方法的优点是它能反映模式的结构特征  (3)逻辑特征法  就是其特征的选择对一.类模式识别问题来说是独一无二的,即在一类问题中只有1个模式具有某1种(或某1组合的)逻辑特征，此方法律立了关于知识表示及组织，目标搜索及匹配的完整体系;对需通过众多规则的推理达到识别目标的问题，有很好的效果，但当样品有缺损，背景不清晰，规则不明确甚至有歧义时，效果不好。  (4)模糊模式方法  就是在模式识别过程中引入了模糊集的概念,由于隶属度函数作为样品与模板相似程度的量度，故能反映整体的、主要的特性，模糊模式有相当不匀称的抗干扰与畸变，从而允许样品有相当程度的干扰与畸变,但准确合理的隶属度函数往往难以建立。目前有学者在研究，并将其引入神经网络方法形成模糊神经网络识别系统。  (5)神经网络方法  就是使用人工神经网络方法实现模式识别。可处理某些环境信息十分复杂，背景知识不清楚，推理规则不明确的问题，允许样品有较大的缺损、畸变。神经网络方法的缺点是其模型在不断丰富完善中，目前能识别的模式类不够多，神经网络方法允许样品有较大的缺损和畸变，其运行速度快，自适应性能好，具有较高的分辨率。  上述几种识别方法各有特点。神经网络方法由于处理的并行性，可以快速同时处理大容量的数据，工作时具有高速度和潜在超高速，并且，网络的最终输出是由所有神经元共同作用的结果，一个神经元的错误对整体的影响很小，所以其容错性也非常的好。基于以上的考虑，本文的手写数字识别采用了神经网络的方法。  **五．预期成果：**  实现一个完整的数字识别系统，并且手写数字成功率达到99%以上  **六．开发计划**：  **1** 调研准备 2月25日-3月1日（第1周）  **2** 熟悉语言及相关技术 3月4日-3月15日（第2-3周）  **3** 软件设计及其开发 3月18日-4月26日（第4-9周）  **4** 调试 4月29日-5月10日（第10-11周）  **5** 撰写毕业设计论文 5月13日-5月31（第12-14周）  **6** 提交论文评审 6月3日-6月7日 （第15周）  **7** 毕业答辩 6月10日-6月14日（第16周）  **七．完成任务所具备的条件因素:**  1.首先是指导老师的指导和帮助，同学之间的探讨；  2.深入理解Tensor Flow框架和CNN原理；  3.熟练掌握 Python 语言和 Pycharm编译工具。  4.熟练查阅相关的文献、资料  **八．参考文献**   1. [手写数字识别的前景与难点](http://kns.cnki.net/kcms/detail/detail.aspx?filename=SJSM201601048&dbcode=CJFQ&dbname=CJFD2016&v=" \t "kcmstarget)[J]. 张晓.  数码世界. 2016(01) 2. [脱机手写数字识别技术研究](http://kns.cnki.net/kcms/detail/detail.aspx?filename=DNZS201629081&dbcode=CJFQ&dbname=CJFD2016&v=" \t "kcmstarget)[J]. 张玉叶,王尚强,王淑娟,王春歆.  电脑知识与技术. 2016(29) 3. 楼怡航, 基于数字图像处理的车牌识别技术[J].电子制作:2019 4. 郭佳. 基于图像的表格识别算法与自动录入系统[D].北京邮电大学,2018. 5. Waris M A，Iosifidis A，Gabbouj M． Object proposals using CNNbased edge filtering［C］/ / International Conference on Pattern Recognition. New York: IEEE，2017:627-632 6. Zhang P，Zhuo T，Huang W，et al． Online object tracking based on CNN with spatial-temporal saliency guided sampling [J]. Neurocomputing，2017，257:115-127 7. Smirnov E A，Timoshenko D M，Andrianov S N． Comparison of regularization methods for image net classification with deep convolutional neural networks [J]. Aasri Procedia，2014，6( 1) : 89-94 8. [基于深度神经网络的手写数字识别模拟研究](http://kns.cnki.net/kcms/detail/detail.aspx?filename=KXJS201905029&dbcode=CJFQ&dbname=CJFDTEMP&v=" \t "kcmstarget)[J]. 宋晓茹,吴雪,高嵩,陈超波.  科学技术与工程. 2019(05)   **指导教师意见及建议：**  指导教师签名：  年 月 日 | | | | | |