动智能与大数据相结合开发的教学管理系统，将能够很好地解决目前学校 PC 端的教学管理系统存在的弊端，推动教学移动智能化的发展，提升教学效率。同时通过该系统学生可以根据自己的空闲时间随时随地学习，也可以将学生的动态及时更新发送给学生。出勤记录和报告生成变得容易。并且系统的灵活性增高，故障率下降，可以更好服务学校的老师和学生。将来，该系统可以实现大多数教育系统的自动化，并且可以用于设计跨平台的教务系统。

参考文献：

[1]杨鹏，梁维意. 基于数字化教学管理系统模式的教学质量管理研究—民办高校传统教学管理模式改革及探究[J].资源建设，2011.

[2]罗轶玮，姚建文. 基于微服务架构的临床医学教学管理系统[J].中国医学教育技术，2019.

[3]彭银，赵秀丽. 基于 Hadoop 虚拟仿真教学管理系统研究[J].中国设备工程，2020.

[4]王华君. 基于大数据的网络教学管理研究[J].中国多媒体与网络教学学报，2019.

[5]白杨. 大数据背景下的教育管理模式革新[J].课程教育研究，2018.

[6]蒋东兴，付小龙，袁芳，蒋磊宏. 高校智慧校园技术参考模型设计[J]. 中国电化教育，2016.

[7]贺占魁，黄涛. 高校智慧教室的建设理念、模式与应用展望—以华中师范大学为例[J]. 现代教育技术，2018.

[8]段修俊.基于移动智能终端的教学管理系统的设计实现[J].课程教育研究，2019.

基于机器学习的截图识别翻译应用研究

◆张龙坤　何舟桥　万武南

（成都信息工程大学网络空间安全学院　四川　610225）

摘要：随着互联网和移动互联网技术的发展，人工智能、大数据、云计算、机器学习的循序渐进的研究和具体应用，也催生出了文字识别和文字翻译的新应用。本文着重分析图像截取、文字识别和文字翻译应用结合研究的现状以及应用前景，并对现有的文字识别算法做简要概述和 Tesseract ORC 算法做具体分析。图像截取、文字识别和文字翻译应用结合的功能将成为新一代的办公软件和社交软件的基本功能，新的用户体验也必将推动着办公软件和社交软件的革新。本文提出，图像截取、文字识别和文字翻译三种应用结合的基本模型结构，并对该模型结构结合现有的机器学习算法做示例实现，过示例演示对模型结构做具体解释。

关键词：图像截图；文字识别；文字翻译；网络办公；机器学习；模型应用

基金项目：成都信息工程大学 2019-2021 年第一阶段本科教学工程项目（BKJX2019129、BKJX2019033）；成都信息工程大学 2018 年大学生创新创业项目（201810621146）；四川省科技厅重点研发（2017GZ0314）；成都市科技惠民项目（2016-HM01-00217-SF）；四川省教育厅重点项目（16ZA0212）

随着 5G 时代的到来，信息大爆炸、人工智能、大数据、机器学习等成为这个时代的新标签。人们对这些新技术的应用也获得了新的进步和发展。尤其是在当代，网络办公和网络课堂正如雨后春笋般出现，图像交流和图像传播已经成了新的趋势，图像截取、文字识别和文字翻译三种应用结合成为人们网络办公和网络社交的基本需求。但是市面上与需求匹配的应用却发展滞后。突出的问题是文字识别和文字翻译单个功能在发展，但是没有将图像识别、文字识别和文字翻译三种应用进行密切结合。而这个结合点正是人们现在的特殊需要和新需求。

出现新的需求，就会有新的应用为之诞生，也会推动着相应的技术和算法取得新的突破。本文将对文字识别算法 Tesseract OCR 分析和介绍。图像截取、文字识别和文字翻译应用结合的功能将成为新一代的办公软件和社交软件的基本功能。新的用户体验也必将推动着办公软件和社交软件的革新。本文给出，图像截取、文字识别和文字翻译三种应用结合的基本模型，并对模型做出简单的示例演示。

1　应用结合研究现状和应用前景

1.1 目前的发展和存在的问题

随着互联网技术的高速发展，人工智能、机器学习、文字识别，智能翻译等新的高科技领域研究和发展方兴未艾。对这些领域新的应用创新也层出不穷。其中图像截取，文字识别和文字翻译的研究和应用如火如荼，但是这三种应用的结合使用，却没有跟上信息社会的发展脚步，略显滞后。市面上对这三种应用的结合也有许多不尽如人意的地方。目前的大多数办公软件、社交软件和翻译软件都没有将这三种应用得到很好的结合。有些软件要么只有图像截取、要么只有文字翻译功能，却很少有将三种应用全面结合的。支撑这三种应用的有关技术的发展和相关算法的研究也存在一定的瓶颈。文字识别存在不准确和模糊现象和文字翻译结果略显机械呆板的情况，这些问题的解决都需要技术的进一步发展和算法的新突破。

1.2 应用需求分析

信息大爆炸时代已经到来，传统经济向数字经济转型发展，人们的需求日益变化。产品设计也从以技术为核心、以功能为卖点到以用户体验为中心进行了转变。人们对办公软件、社交软件和翻译软件的办公效率和使用体验也提出了新的要求。当下，网络办公、网络课堂发展异军突起，人们对图像处理、文字识别、文字翻译等功能的需求显著增加。在网上办公、网上课堂和网上社交等必然难以离开文字，文字是人们使用这些应用的进行交流的主要途径。技术的发展使得我们对图像的传输和广泛使用成了可能。新的需求推动着对图像处理的研究，例如图像中的文字识别。随着经济全球化的发展，国与国之间的交流日益密切，人们对除了母语之外的语言文字的接触也日益频繁。文字处理有了新的广阔的应用场景，即对图像文字的识别提取和对文字翻译。

1.3 应用前景

随着科技的发展，人们对图像截取、文字识别和文字翻译应用结合的需求已经成了新的发展趋势。图像交流和图像传播具有高效的特点，而人们对各种软件和应用的体验也正在向简洁、高效和智能等方面聚集。对图像中的文字识别和再理解，也将成为人们的基本需求。有了新的需求推动，必然会产生新的用户体验设计。图像截取、文字识别和文字翻译三种应用的结合必然会成为下一代办公软件和社交软件的必备的基础功能，也必然成为有关软件对用户体验新的研究点，进而成为新的价值增长点。

2　基于机器学习文字识别算法

2.1 机器学习文字识别算法简述

新的需求已经产生，新的用户体验不仅依赖于应用结合的创新，还需要从技术这个根本点出发。图像截图、文字识别和文字翻译三种应用的结合的实现和新发展，需要文字识别和文字翻译有关算法和技术的新突破。

文字识别和文字翻译当前的主流机器学习算法的研究已经取得

了丰硕的成果。

文字识别算法：基于深度学习的 OCR（Optical Character Recognition，光学字符识别）技术，例如 Tesseract OCR 和 CRNN OCR 和 attention OCR。OCR 电子设备（例如扫描仪或数码相机）检查纸上打印的字符，通过检测暗、亮的模式确定其形状，然后用字符识别方法将形状翻译成计算机文字的过程。

支持“向量机算法”，是一种监督式学习的方法，可广泛地应用于统计分类以及回归分析。它是将向量映射到一个更高维的空间里，在这个空间里建立有一个最大间隔超平面。在分开数据的超平面的两边建有两个互相平行的超平面，分隔超平面使两个平行超平面的距离最大化。假定平行超平面间的距离越大，分类器的总误差越小。

文字翻译算法：循环神经网络和双向 RNN 是比较成熟的文字翻译算法，这两种算法的核心是通过分析大量的文档从而得出相应的模型以实现对文字的翻译。

这些主流的机器学习算法，都有本身的优点和缺陷，需要进一步的研究和发展。

2.2 Tesseract OCR 算法

笔者在实现图像截取，文字识别和文字翻译应用基本模型时，通过使用 Tesseract OCR 算法来实现文字识别功能。在此，简单地分析和介绍一下，Tesseract OCR 算法。

Tesseract 使用了一种通过文本行累加的简单而有效的偏斜检测算法。该算法中将 Blobs 组织成文本行，并且分析这些行和区域以用来获取固定间距或成比例的文本。然后会经过两遍识别，在第一遍中，尝试依次识别每个单词，将每个令人满意的单词都将作为训练数据传递给自适应分类器。然后自适应分类器将会更准确地识别页面下方的文本行，找到文本行后，使用二次样条曲线更精确地拟合基线。通过将 blobs 划分为几组来拟合基线，并为原始笔直基线合理地连续位移。

二次样条曲线以最小二乘拟合到人口最多的分区（假定为基线）。当从一个完整的单词识别出来的结果不太满意时，tesseract 通过字符级别的分割 blob 来改善结果。多边形轮廓的一些凹的顶点是作为候选的分割点，以及相反方向的凹点或者线段。这三部分可以成功的分割连接的字符。当潜在的分割点已用完，但是还不能满足要求，文字识别不好时，就用到了“联合器”。“联合器”会尝试最优先搜索，把分割的 blob 最大可能联合成候选字符。最后来解决含糊不清的空格，检查 x-height，定位 small-cap 的文本，和对大小写处理。

字库训练：由于 tesseract 利用多边形近似法而不是字符粗略的轮廓这种不同寻常的处理方式，也就带来了识别率不高的问题。因此为了提高不同场景下对文字的识别率，利用 jTessBoxEditor 软件，需要对字库进行训练：具体操作步骤如下：

（1） 图片样本采集
（2） 图片样本标注
（3） 图片样本降噪
（4） 将图片转换为 TIF 格式
（5） 将 TIF 格式的图片样本合并为一个 TIF 文件
（6） 生成 TIF 文件的 box 盒子文件
（7） 补充和修正 box 盒子文件
（8） 生成 lstmf 文件
（9） 提取语言的 LSTMF 文件
（10）训练

3　截屏翻译识别软件模型结构及实现

3.1 模型示例简述

简易 STR（截屏翻译识别）软件是一款基于 Python、运用了文字识别、文字翻译、截屏等技术的文字工具，致力于解决文字转换给人们带来的困扰。它包含了文字识别、文字提取、截屏等三大模块，三者既可以各自独立工作，也可结合起来。文字翻译模块可指定英语、中文、韩语、法语等国际通用语言之间相互转换，可从翻译文本框中选择性的转换从图片中识别出的文字，翻译后的文本可全文复制以备他用；图片截取模块可截取当前屏幕的选择区域，也可全屏截取，截取图片后可以使用文字提取模块将截图中的文字提取到翻译文本框中，也可将图片复制到剪切板，以便发送给他人，或做其他的编辑。

3.2 模型示例系统框架图和设计总流程

本软件采用 Python 语言编程算法设计，使用最多的是逻辑结构的判断语句，还有顺序结构，还有循环结构。本系统的大多数算法简单，软件方便操作，容易上手。软件设计时简单分为可视化窗口设计，截图功能设计、文字识别功能设计和翻译功能设计四大部分进行编程，使用 python 语言中的 tkinter 进行 GUI 设计，简单朴素，功能齐全，使用 tkinter 编辑可视化窗口，操作简单。系统框架如图 1，软件流程如图 2。

3.3 模型示例的各模块分析

3.3.1 图像截取的实现

使用 tkinter 的 canvas 画布并监听鼠标的单击位置和释放位置来实现截图，可以对屏幕进行截图，使用者可以使用本功能截取屏幕上想要的部位，也可以对图片利用本软件的文字提取模块对截取图中的文字进行提取，从未获得图片中文字的文本。截取的图片会以一个单独的窗口呈现，对窗口里的截取的图片可以进行复制，以作其他用途。大致步骤如下所示：

（1） 产生截屏需求
（2） 点击图片截图按钮
（3） 调用 tkinter 模块中的 canvas 画布
（4） 截取图片磁盘存储
（5） 截取图片可视化呈现
（6） 图片利用，文字提取，图片拷贝

3.3.2 图像文字识别的实现

本软件使用由 Google 公司维护的一款功能强大的图像识别软件 tesseract，其可识别超过 100 种语言，本软件根据例如各种软件无法识别的提示框，各种语言的外刊等几十种不同应用场景来训练了字库，将其对文字的识别率提高到了 98%。提取图片中的文字，将图片中的文字转化为可以进行编辑的文本形式。方便使用者对图片中的文字进行编辑，并且结合了图片截取功能，使得办公更加方便。大致步骤如下所示：

（1） 对获取图片，提取文字
（2） 点击文字提取按钮
（3） 程序通过 pytesseract 库调用 tesseract
（4） 发出请求，返回提取结果
（5） 对提取结果进行简单处理
（6） 在 InText 文本框中呈现，翻译编辑等

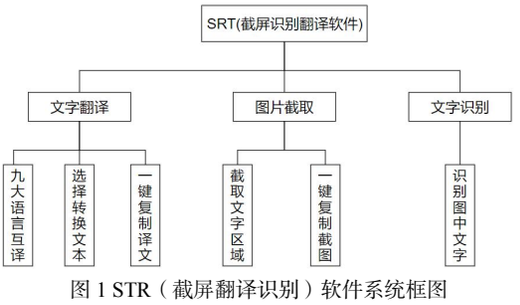

图 1 STR（截屏翻译识别）软件系统框图

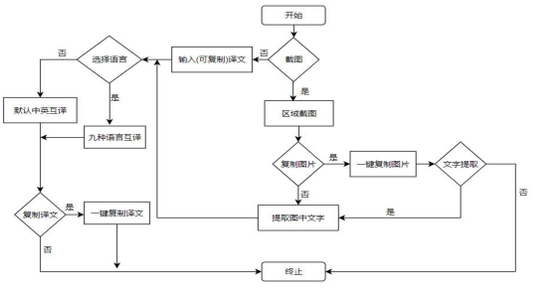

图 2 简易 STR（截图翻译识别）软件流程图

3.3.3 文字翻译的实现

软件基于强大的百度翻译平台来实现对不同种语言的转换，支持世界上主流语言的互相翻译：目前支持的语言有中文、英语、日语、

韩语、法语、阿拉伯语、俄罗斯语、西班牙语、葡萄牙语等国际主流语言的相互翻译。只需将想要翻译的文字放到 InText 文本框里，选择对应的语言和要翻译成的语言即可完成翻译。译文会自动呈现在 OutText 文本框中。其步骤如下所示：

（1） 将原文放到 InText 文本框中

（2） 选定原文的语言，选定译文语言

（3） 点击文字翻译按钮

（4） OutText 文本框呈现翻译的译文

3.3.4 实现结果展示

通过对图像截取、文字识别、文字翻译这三大原本各自独立功能的结合，并依附强大的百度翻译，Google Tesseract-OCR 图像识别软件，最终形成了 STR（截屏翻译识别）软件的雏形。打开主界面，如需截取相应的文字区域，则点击图片截取按钮来截图，然后通过文字提取按钮来识别图片中的文字，再指定想要转换的源语言和目的语言来转换全文或者选择其中的一部分文字来进行转换，点击文字翻译按钮后会将全文或者选中的文字翻译成指定的语言，输出到下方的文本框中。

运行主界面如图 3，文字翻译结果展示如图 4。

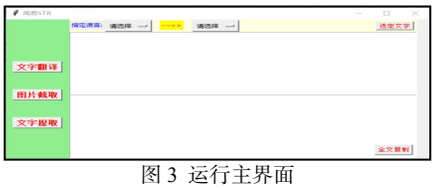

图 3 运行主界面

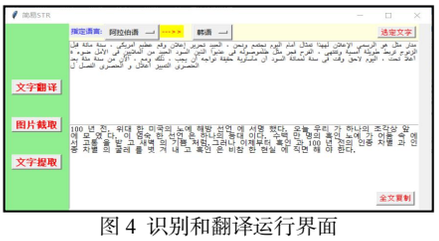

图 4 识别和翻译运行界面

4 结束语

网络空间已经成为人们主要的精神生活虚拟空间，网络空间的信息主要的形式载体就包含图片，而图像交流和图像传播已经成为一个新的趋势。对图像种文字地提取和翻译已经成为新的需求。图像截取、文字识别和文字翻译的应用结合迫切需要得到实现。文字识别和文字翻译的新技术和新算法也需要不断地突破，为新的应用结合提供技术支撑。本文据此提出，图像截取、文字识别和文字翻译应用结合的基本模型，并通过示例对模型做了基本演示，更加具体和可行的模型和具有更加丰富功能的示例，需要进一步的丰富和完善。

参考文献：

[1]刘昌伟.自主可控环境下办公软件发展新趋势分析[J].网络空间安全，2018.

[2]洪锦魁.Python GUI 设计 tkinter 菜鸟编程[M].清华大学出版社，2019.

[3]陈文鹏.计算机智能图像识别算法研究[J].无线互联科技，2019.

[4]刘水丽，吴恋，吴文宇.基于深度学习文字识别技术发展现状及展望[J].电脑知识与技术，2019.

[5]官文天，刘春晓，林朗，等.基于微信平台的学生管理系统的研究设计[J].软件，2015.

[6]张晓轩.计算机智能图像识别的算法与技术研究[J].软件，2018.

[7]杨梦铎，李凡长，张莉，等.机器学习十年研究进展[J]，2015.

[8]盛宦，刘伟.计算机文字识别的发展及应用[J].科技信息，2008.

基于某研究院 OA 系统在线自动归档的研究

◆李烨　张良旭　陈哲

（核工业理化工程研究院　天津　300160）

摘要：随着 OA（综合业务）系统的公文规模日益增多给传统归档方式带来挑战。本文通过分析档案系统和综合业务系统的业务特征，运用 webservice、XML、URL 技术，建立一套规范的接口模型，实现综合业务系统公文自动归档的功能，为其他应用系统的自动归档提供有价值的参考。

关键词：公文自动归档；流程设计；webservice 接口

近年来核工业理化工程研究院(以下简称核理化院)科研任务重，归档压力也在逐年增加。目前，核理化院现有的档案业务系统支持在线归档，由部门档案员完成录入，通过录入案卷文件信息生成该系列档案。当录入大量资料时，在此种一文一录的模式下，重复工作现象严重。尤其是综合业务平台中形成的公文，每年生成近百份待归档资料。而综合业务系统只提供简单的归档业务，导致这些资料，一方面通过网络办公形成的数据散乱，另一方面归档部门还在通过手工著录方式进行档案归档，造成资源人力、物力的浪费[1]。为此，网络信息中心通过开发档案系统与综合业务系统的接口，实现公文自动归档。

1 公文自动归档可行性分析

业务层面，核理化院现有公文归档制度依据国家、集团法规制定，文书档案的收集按照 2016 年制度第 27 号《核工业理化工程研究院管理类文件材料归档范围和档案保管期限规定》执行。文书档案的整理依据 DA/T22-2015 中华人民共和国档案行业标准《归档文件整理规则》执行，制度中对公文归档著录项、归档模板样式作出规范，数据标准统一。

技术层面，档案管理系统与综合业务系统虽然表结构、名称并不相同，但存储的内容实质相同，如在档案系统数据库中存在题名字段而综合管理系统中对应文件名字段。同时，接口自动推送本质上是数据的提取、交换和处理。在核理化院的信息化部门统筹规划下均采用 Oracle 11g 版本数据库，作为同质数据库，为数据提取、交换提供便利，不需通过接口工具对数据库字段再整述，运用数据库、中间件技术实现跨系统的自动归档合。最后，Oracle 数据库作为一个关系型数据库，具有可扩展性强、数据安全性强的特点，保障数据库传输的可靠性。

综上所述，运用数据库、中间件技术实现跨系统的自动归档。

2 总体方案设计

档案具有准确性、原始性的特点。因此，该系统对于综合业务系统电子文件的归档将采取包含了元数据描述信息的电子文件主动推送至档案管理系统的模式。档案管理系统对业务系统推送过来的归档