OCR泛化课题的研究方案

# （一）立项依据与研究内容

1．项目的立项依据

1.1 研究背景和意义

随着互联网及移动设备的发展，除了传统的柜台业务外，用户使用银行app办理业务，业务员使用手机或平板等终端设备为用户办理业务成为常态。而在业务办理中，常常涉及到大量身份信息、凭证信息的输入，手动输入的方式费时且费力，因此OCR（Optical Character Recognition，光学字符识别）成为了一种切实可行的解决方案。

在银行现有的场景中，OCR处理的图像来源包括扫描仪或高拍仪扫描获得的、手机拍摄证件或纸张和手机拍摄屏幕等多种不同途径，图像中需要识别的目标内容包括卡证、票据、表格等多种不同形式文字载体，获得的图像也具有不同格式、不同质量、不同分辨率。如何使用统一的框架或模型在凭证识别场景中对OCR进行泛化增强成为了一个非常值得研究的问题。另一方面，扫描的卡、票、证、表等材料都具有一定的格式，电子化过程除了文字识别外，材料的版面复原也是现在一个非常常见的需求，而在当前国内外研究中，这仍是一个难点。

1.2 国内外研究现状

OCR（Optical Character Recognition）的概念是最早由德国科学家Tausheck提出的。早期的OCR研究主要针对扫描的印刷文本，通常采用光学的方式将纸质文档中的文字转换成为黑白点阵的图像文件，并通过识别软件将图像中的文字转换成文本格式，供文字处理软件进一步编辑加工。而现在的光学文字识别研究，应用范围更加广泛，包含自然场景下的文字检测、手写文字识别、文档版式还原等。如何利用已知信息来提高识别正确率是当前最重要的课题，ICR（Intelligent Character Recognition）的名词也因此而产生。

传统的OCR识别技术主要包含文档图像输入、图像预处理、版面分析、字符切割、字符特征提取、字符识别、版面恢复和后处理校正这几大步骤。

目前，在版面识别方面，国内外学者都提出了一些有效的方法。Akmal Jahan等人提出利用行高和字间距来定位表格等非文本信息，提取页面中的不同元素，在实验中对于非文档元素的识别准确率达到75%。Ying等人提出两个算法来提高表格边界识别的准确率，消弱表格内空行等对于版面分析带来的影响。Kboubi等人则提出将多个OCR系统结合来提高表格识别的准确率。

在表格识别与复原方面，主要的技术难点在于表格线的提取和表格的描述重绘。目前，主流的表格线提取方法是通过Hough变换提取表格线，大多数研究都选择基于此技术进行表格提取。而对于表格的描述，则有各种不同的策略。张慧提出了基于模型驱动的表格识别方法，采用Hough变换提取表格线，然后使用定义好的描述模板将表格信息描述出来。刘昱在印刷体表格识别的研究中，提取表格线坐标，识别文本信息，利用连通区域的中心坐标与表格线坐标进行表格重绘。

在文本检测方面，传统的OCR研究中，大量使用基于连通分量的方法，B. Epshtein等人使用了SWT(stroke width transform)算法，J. Matas等人提出了MSER(maximally stable extremal region)算法来从背景中提取字符候选，并将这些候选区域合并来获得文字区域对象。随着深度学习的发展，越来越多学者也开始使用基于对象检测和语义分割的方式来进行文本检测。Tian Zhi等人在16年提出了基于Faster RCNN的CTPN模型（基于连接预选框网络的文本检测），它对于自然场景中的文本检测在ICDAR13数据集上获得了82%的f1 score评分。17年R2CNN模型被提出，它对于旋转文本的检测进行了优化，在ICDAR13数据集上获得了87%的f1 score。19年最新的CVPR会议上，Baek, Youngmin等人在论文“Character Region Awareness for Text Detection”论文中提出了通过挖掘每个字符和字符之间的联系来有效地检测文本区域的方法，并在ICDAR13数据集上将文本检测评分提高到了95%。

在文本识别方面，现在基于深度学习的方法也是主流。Su和Lu首先将文本识别看作序列识别问题，使用RNN网络模型来解决。P. He等人进一步将CNN和RNN进行结合。Shi Baoguang等人提出了ASTER模型，引入了注意力机制并分别构建了文本校正模型和识别模型，结合seq2seq解码器，在多个数据集上都取得了非常好的结果。

除了将文本检测与文本识别问题分别解决之外，现在也有许多学者研究端到端的学习，利用一个神经网络解决这两个问题。Tao Wang等人在12年率先将端到端神经网络训练应用到文本识别领域，在ICDAR03数据集上获得67%的f1 score。18年的CVPR会议上，Liu等人提出了FOTS网络，在ICDAR13和ICDAR15的数据集上都取得了端到端识别领域当前最好成绩，分别是84.7%和65.3%。

相比于两阶段的文本检测+文本识别方案，端到端的神经网络训练解决方案具有更快的速度，但是在现行研究中，对于文本检测和文本识别的精度相对于分阶段的解决方案而言比较低。

当前业内成熟的OCR软件和开源工具包有Tesseract，Abbyy FineReader，国内的百度、腾讯及有道也有效果较好的在线API。

1.3 当前研究的局限性分析

当前研究提出的算法和模型大多适用于云端进行运算，对于算力要求高，响应时间较长，对于线下场景的用户体验差。当前研究对于复杂格式的文档版面还原能力较弱，当包含复杂版面格式的文档中缺少明显的边线或者样式时，版面还原难度大。在文字识别方面，大多针对单一语言或者仅适应印刷体或手写字，对于混合场景的泛化支持能力有待提高。

知识图谱作为一种具有语义信息的知识网络，具有强大的信息理解和推理能力。将知识图谱技术与OCR研究相结合，在银行OCR场景中，构建领域知识图谱，利用领域知识对识别的文字进行验证，根据文本在图像中的上下位关系进行推理，将对文字识别准确率的提升和材料格式还原都有极大的帮助。对于本项目的识别可以提供内容意义层面的辅助，提高识别结果的可理解性和智能对照优化。

1.4研究思路和目标

主要研究路线分为三阶段：

第一阶段，对终端设备输入的图像质量进行评估，对于图像清晰度等质量标准进行判断，对拍摄材料尺寸进行判断，结果及时反馈给用户（终端设备使用者）。这一部分功能需要部署到终端支持离线运行，且反馈速度快。

第二阶段，构建特定场景下的通用文字识别模型，对于发票、身份证、表格等不同内容分别构建模型，使得模型在银行业务场景中针对特定类型版式还原和文字识别（支持多种字体、多种语言，适应印刷字和手写字）都能取得较好的效果。

第三阶段，对于表格OCR场景的单元格检测进行优化，对于边线不清晰、图像噪声较多以及表格无线框样式的情况做进一步优化，实现单元格的有效定位。

参考文献

1. AKMAL JAHAN M A C, RAGEL R G. Locating tables in scanned documents for reconstructing and republishing; proceedings of the Information and Automation for Sustainability (ICIAfS), 2014 7th International Conference on, F 22-24 Dec. 2014, 2014 [C].
2. YING L, BAI K, MITRA P, et al. Improving the Table Boundary Detection in PDFs by Fixing the Sequence Error of the Sparse Lines; proceedings of the Document Analysis and Recognition, 2009 ICDAR '09 10th International Conference on, F 26-29 July 2009, 2009 [C].
3. B. Epshtein, E. Ofek, and Y. Wexler, “Detecting text in natural scenes with stroke width transform,” in Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on. IEEE, 2010, pp. 2963–2970
4. J. Matas, O. Chum, M. Urban, and T. Pajdla, “Robust wide-baseline stereo from maximally stable extremal regions,” Image & Vision Computing, vol. 22, no. 10, pp. 761–767, 2004.
5. Tian Z , Huang W , He T , et al. Detecting Text in Natural Image with Connectionist Text Proposal Network[J]. 2016.
6. Jiang Y , Zhu X , Wang X , et al. R2CNN: Rotational Region CNN for Orientation Robust Scene Text Detection[J]. 2017.
7. Baek Y , Lee B , Han D , et al. Character Region Awareness for Text Detection[J]. 2019.
8. B. Su, S. Lu, "Accurate scene text recognition based on recurrent neural network", Proc. Asian Conf. Comput. Vis., pp. 35-48, 2014.
9. P. He, W. Huang, Y. Qiao, C. C. Loy, X. Tang, "Reading scene text in deep convolutional sequences", Proc. AAAI Conf. Artif. Intell., pp. 3501-3508, 2016.
10. Baoguang S , Mingkun Y , Xinggang W , et al. ASTER: An Attentional Scene Text Recognizer with Flexible Rectification[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018:1-1.
11. Babenko B, Belongie S. End-to-end scene text recognition[C]// IEEE International Conference on Computer Vision. 2012.
12. Liu X , Liang D , Yan S , et al. FOTS: Fast Oriented Text Spotting with a Unified Network[J]. 2018.

2．项目的研究内容、研究目标，以及拟解决的关键科学问题

2.1 研究目标

1. 实现可通用部署到多种终端设备的影像质量评估模型

2. 实现具有泛化能力可支持多语言、多模式、多场景的OCR模型

3. 实现针对各种表格类型（有框线/无框线）文档图像优化的版面复原模型

2.2 研究内容

1. 构建输入影像质量评估模型

建立模型对输入的影像质量进行评估与检测。图像来源方面包括扫描仪（分辨率从70DPI～300DPI）、高拍仪（高度不固定），及手机拍摄实物纸张、手机拍摄屏幕。不同来源、不同影像质量，导致后续的识别难度较高。为此，首先要对影像质量进行基本的判断，需实现一个对影像材料质量分析和合规性判断的模型，可用于对提交材料进行初步判断，若质量低于某个阈值，则返回重新提交。影像质量包括对于影像材料尺寸合规性的判断、图像倾斜度的判断及对反光、清晰度、曝光程度等图像质量评价指标的计算。

除保证准确率之外，模型要具有通用可部署性，支持部署到各种图像采集设备并能在可接受时间内得出结果给予用户反馈。

2. 构建并优化具有泛化能力的OCR模型

对于输入图片，消除不同文字背景带来的差异，解决不同图像格式不同压缩比及图像尺寸带来的数据源不一致，对书写方向倾斜、打印颜色深浅不一等问题做校正。

训练基于深度学习模型，有较好的多字体（支持MS Office中常用印刷字体）、多语言（同时支持简体中文、繁体中文、英文和数字）适应能力的印刷体识别模型和泛化能力较强的手写体中文识别模型，且能支持印刷体、手写体的混合识别。泛化能力体现在对于未经过训练的，且来自不同场景的图像也能准确识别其中的手写体文字。

通过版式分析和还原，结合识别到的材料尺寸，按比例复原电子材料。

通过后处理纠偏对于识别结果进行分析，对于不合理的识别结果，提供置信度较高的参考替换结果。

3. 优化表格OCR场景单元格检测

针对表格识别场景，优化表格线/无线条单元格检测。在报表打印墨迹不清晰或拍摄角度导致断线、手机拍摄实物纸张存在翻折或卷曲、手机拍摄屏幕存在屏闪线等噪声干扰，以及无线条样式表格的情况，有效定位出表单元格。

2.3 拟解决的关键科学问题

1. 设计具有高准确率的影像质量评估模型

2. 针对多种终端通用部署优化压缩影像质量评估模型

3. 构建支持手写体、印刷体混合识别的高精度OCR模型

4. 针对银行业务场景下的多字体、多语言的通用识别进行优化

5. 实现多种不同材料（卡、证、券）的高精度版式识别及复原

6. 针对表格类型，对于有框线/无框线单元格检测进行优化，提高定位准确率。

3．拟采取的研究方案及可行性分析

3.1 研究方案

对于影像质量评估模型构建，针对尺寸估计、失焦、反光、曝光过高或过低等问题判断，采用多种不同算法，对于算法计算效率和准确率进行权衡，选择适合部署到终端的方案。

对于OCR模型通用泛化能力研究，设计并实验多种神经网络模型，引入知识图谱技术辅助后处理与版面还原。

对于表格场景下的单元格检测优化，在通过图像增强手段优化图像之后，结合知识图谱与神经网络方案对于单元格位置进行推理。

3.2 技术路线及可行性分析

3.2.1 输入影像质量评估

3.2.1.1 图像曝光分析

将图像转为灰度图，计算图片在灰度图上的均值和方差，当存在亮度异常时，均值会偏离均值点（可以假设为128），方差也会偏小；通过计算灰度图的均值和方差，可以评估图像是否存在过曝光或曝光不足。

3.2.1.2 图像倾斜度判断与校正

针对前后景差别较为明显，边缘清晰的图片，利用边缘检测和霍夫直线变换提取图像边缘，根据矩形特征提取图像中扫描的票证区域并获取四个角点。使用四个角点和图像边缘计算夹角就可以得到倾斜角。获取倾斜角之后可以使用仿射变换来修正图像倾斜和透视。

针对目标边缘不清晰的图像（如扫描仪扫描比较薄的纸张获得的图像），用MSER算法提取文字行区域，获得参考直线，和图像边缘计算夹角，获得倾斜角，通过旋转可以获得校正后的图像。

3.1.1.3 图像清晰度分析

结合采用 tenegrad法和brenner法对倾斜校正后的图像清晰度进行评估，任一方法检测得到的清晰度得分低于阈值，则判断为不清晰，给予重新拍摄的反馈。阈值通过对于各类不同设备拍摄到的不同格式图像进行综合分析计算得到。

3.1.1.4 图像反光判断

使用文字检测相关算法（MSER，SWT等）对包含文字的区域进行检测，对这些区域进行曝光分析，如果存在某个文字区域的曝光度相较于其他更高，则认为存在反光区域，需要提醒用户重新采集图像。

3.1.1.5 影像材料尺寸判断

在图像采集时要求放置身份证或者工作证等大小已知的物品作为参照物，根据参照物和识别的材料在倾斜校正后在图像中的大小计算材料实际大小。结合知识库中各类材料尺寸和比例，判断扫描材料是否合规。

3.1.1.6 模型构建及验证

a. 对数据集随机划分测试集和训练集，在训练集上迭代对获取各项图像问题采用不同类型的传统图像处理算法进行判断，不断调整阈值直到收敛。

b. 对数据集随机划分测试集和训练集，采用神经网络模型，对各类图像问题分别训练二分类器来判别图像是否符合要求。

c.使用C++进行编程并对上述两种方案进行评估，在保证准确率和召回率的前提下，综合考虑部署到终端的难易程度，选择最终方案。

3.2.2 OCR模型通用泛化能力提升研究

3.2.2.1 图像预处理

因为图像来源于不同采集设备，因此存在不同尺寸和格式，在对图像进行处理之前，需要先做一步预处理将它们统一成同一种存储格式，并resize到同样的尺寸。

在识别文字或者做图像分割之前，还需要对图像进行增强操作，其中的步骤主要包括图像二值化、降噪、倾斜校正。

其中，二值化将图像中的颜色分为黑白两色，其中一种即为文字内容。可以考虑使用wellner自适应阈值二值化算法,对于光照不均匀的纸张图像可以得到较好的二值化结果。

图像降噪则是将图像输入时产生的噪点等无用的干扰信息去除，让文字的特征更加明显。图像降噪常用的滤波算法包括均值滤波、高斯滤波、双边滤波和中值滤波，除此之外， stroke filter对于文本分割有提升效果。

倾斜校正是将扫描或拍摄得到的图像中产生的倾斜和透视情况进行修正，常用的方法在3.2.1.2中有描述。

3.2.2.2 文字区域定位

银行业务中涉及到的文字检测场景相对较为简单，因此可以考虑使用MSER算法+NMS算法进行提取。对于银行卡、发票等格式相对固定的检测对象，可以根据模板规则等进行推理校验，优化文字区域定位。另外实现一种基于深度学习的文字检测模型，并对两种算法进行对比。

3.2.2.3 文字识别

本研究中的文字识别目标包括多种印刷字体及手写字体，覆盖多种语言（简体中文、繁体中文、英文、数字）且能支持混合模式识别。

在获得文本区域边界之后，传统方法是分割字符，对单字符进行识别，然而一来字符分割本身是一个难点，尤其是对于手写文字，二来分割后的字符失去了语义信息。因此，本研究将分别在现在较为流行且有较好效果的CRNN模型和attention OCR模型基础上进行研究，在保留词汇、语句之间的上下文关系条件下识别文字。本研究将比较这两种模型效果并针对银行业务场景对模型进行优化。

3.2.2.4 版面还原

根据文字区域定位和文字识别结果进行版面还原，使用影像质量评估模型计算得到的材料原始尺寸进行等比例复原。

3.2.2.5 纠偏及后处理

根据现有业务数据构建领域知识图谱，并利用构建的领域知识图谱对于文字的定位进行合理化验证和纠正，优化版面还原结果。结合领域知识图谱和工商信息库等信息源对于文字识别结果进行查询和推理，如果发现不合理的文字，进行标记并提示可能的正确结果。

3.2.2.6 模型评估

对于文字区域定位，对每一张测试图片统计定位到文字区域的总数和正确定位的数量，求定位正确率，取均值作为模型最终的定位正确率。

对于文字识别的评估，将纠偏提示的正确结果作为识别最终结果，统计纠偏后正确识别文字的数量与总文字数，取二者的商作为识别正确率。

3.2.3 表格OCR场景单元格检测优化研究

3.2.3.1 带框线/样式表格单元格定位

对于线框较为清晰的图片，可以先使用边缘检测及霍夫直线等方法获取表格边框，裁剪图片获得表格区域。对于直线进行提取，获取直线交点对图片进行切分从而获取各个单元格所在位置。

3.2.3.2 无框线单元格定位

进一步优化的目标，是对于边线不清晰或者受到污染的图片甚至是无框线或样式对单元格进行分界的含表格图片进行识别。

首先考虑使用传统图像处理方法对噪点进行过滤再尝试提取边线。

对于无明显图像特征分界的情况，利用构建的领域知识图谱，对识别到的文字信息及位置信息进行语义推理，推测各文字块之间的语义关系，从而实现表格边线界定和单元格分割和定位。

3.2.3.3模型评估

通过上述研究对于不同文本块所在表格中定位可以用是在第几行第几列，在测试集上对于正确定位的数量和定位总数进行统计。

4．本项目的特色与创新之处

1. 除了对于卡、票、券、证等常见材料的版面还原之外，对于各类表格（有框线、边线模糊及无框线）进行文字识别及版面还原。

2. 构建领域知识图谱，利用领域知识辅助OCR识别结果的校正和后处理，利用语义关系推理辅助各类材料的版面复原。

3. 针对银行业务领域，对于多字体、多语言、多来源图像的文字识别效果进行提升。

5．年度研究计划及预期研究结果

5.1 年度研究计划

|  |  |  |  |
| --- | --- | --- | --- |
| 阶段 | 时间进度 | 任务 | 任务描述 |
| 影像评估模型构建与优化  （第一期） | 2019.11.1-2019.12.31 | 收集数据，调研图像问题，调研终端设备类型。 | 工作：   1. 收集银行各类业务场景下获取的包含需识别文字的影像材料，对于图像存在或者可能存在的问题（如反光、模糊等）进行调研和分析。 2. 调研采集的各类影像中包含材料的常见规范尺寸并建立知识库 3. 调研各类影像采集设备架构，初步确定后续模型开发部署方案   成果：   1. 影像质量评估模型开发具体技术方案 2. 影像材料尺寸知识库 |
| 2020.1.1-2020.1.31 | 构建影像质量评估模型 | 工作：   1. 采用多种不同算法对影像材料存在的各类问题进行分析判断。 2. 实现影像尺寸估计算法，估计输入影像尺寸并与知识库中规范尺寸进行比对给出尺寸合规性判断。 3. 统计各种算法的运行效率和判别准确率   成果：   1. 准确率达到预期95%的影像质量评估模型 |
| 2020.2.1-2020.3.31 | 优化封装影像质量评估模型 | 工作：   1. 根据各算法运行效率和准确率统计结果进行权衡，选择合适的算法。 2. 将算法和模型进行封装并尝试部署到各类终端设备。 3. 在各终端设备上对影像质量评估模型进行测试，并根据测试结果再进行优化。   成果：   1. 封装完成并可部署到手机、高拍仪等设备的影像质量评估应用demo |
| 通用OCR模型构建与泛化研究  （第一期） | 2020.4.1-2020.5.31 | 图像预处理及文本定位 | 工作：   1. 使用图像处理算法对不同来源的输入图像进行预处理优化和格式统一 2. 实现基于MSER和NMS算法的文本定位算法，并对结果进行评估 3. 实现一个基于深度学习模型的文本定位算法并在实验数据集上和传统方法进行比较   成果：   1. 文本定位准确率达到85%以上的文本定位算法v1.0 |
| 2020.6.1-2020.8.31 | 训练具有通用泛化能力的文字识别模型 | 工作：   1. 构建基于深度学习模型的中文识别模型 2. 不断优化模型，并根据训练结果重新优化图像预处理算法   成果：   1. 文字识别准确率达到80%以上的中文文本识别模型v1.0 |
| 2020.9.1-2020.9.30 | 构建领域知识图谱 | 工作：   1. 针对银行需要使用OCR的业务相关领域构建领域知识图谱，结合外部知识库进行扩充和优化   成果：   1. 银行业务领域知识图谱 |
| 2020.10.1-2020.11.31 | 版面复原 | 工作：   1. 对于卡、票、证、券等不同类型文档，根据文本定位结果、文字识别结果以及领域知识图谱进行版面复原 2. 根据版面复原结果优化文本定位算法   成果：   1. 基于领域知识图谱辅助的文本定位算法v2.0 |
| 2020.12.1-2020.12.30 | OCR后处理 | 工作：   1. 结合领域知识图谱、复原版面位置对于识别的文字内容进行检查，对于置信度较低的结果，搜索可能的正确替换值   成果：   1. 基于知识图谱辅助后处理，文字识别（修正后）准确率达到85%以上的中文文本识别模型v2.0 |
| 复杂表格OCR场景检测优化研究  （第二期） | 2021-2022 | 无线框表格的单元格定位检测优化研究 | 工作：  对于通用中文OCR模型进一步优化。对于有线框或者明显样式，但图片包含较多噪声或者污点的表格图像的单元格检测进行优化，结合领域知识图谱研究对于无线框表格的单元格定位算法并进行优化。  成果：   1. 优化的OCR模型v3.0 2. 针对表格OCR场景优化的单元格检测算法 |

5.2 预期研究结果

第一期

1.影像质量评估模型：能在测试集上满足召回率达到95%以上，准确率95%以上。

2.通用OCR模型：各类文档的版面复原定位正确率（=正确定位的个数/定位总数）达90%以上；混合手写与印刷体识别准确率（=正确识别的个数/识别综述）达85%以上。

**第二期**

1. 优化的通用OCR模型：各类文档的版面复原定位正确率（=正确定位的个数/定位总数）达95%以上；混合手写与印刷体识别准确率（=正确识别的个数/识别综述）达90%以上。

2. 表格OCR场景优化模型：有框线表单元格定位正确率（=正确定位的个数/定位总数）达95%以上；无框线表单元格定位准确率（=正确定位的个数/定位总数）达90%以上。

# （二）研究基础与工作条件

1．**研究基础**（与本项目相关的研究工作积累和已取得的研究工作成绩）；

课题团队的负责人蔡鸿明教授为ACM/IEEE Senior member，中国图学学会常务理事，计算机图学专委会主任，获中国高校科学技术奖科技进步奖，中国高校科学技术奖自然科学奖，上海市科技进步奖省部科技奖等科技奖励4项，在信息可视化、智慧应用相关领域具有较多积累并在业界具有一定的影响力。

目前课题团队已将OCR技术使用到航天等领域的非结构化历史数据的数字化中。在航天元器件数仓构建项目中，结合了领域术语库对OCR结果进行语义分析推理，提高了特定业务场景下的中文识别准确率。另外曾基于OCR完成纸质表单信息化研究，将纸质表单信息转化为RDF模型表示，并基于关联对于实例进行合并，相关论文（iRMP: From Printed Forms to Relational Data Model）录用于会议2016 IEEE 18th International Conference on High Performance Computing and Communications。团队曾构建了复杂模型的半自动化标注语义方法，建立了面向数据服务的语义关联发现方法，并实现基于相似度计算的本体匹配及扩展，也实现了基于本体的案例查询系统(发明专利1)，为知识图谱的建立和更新提供了较好的基础。

2．**工作条件**（包括已具备的实验条件，尚缺少的实验条件和拟解决的途径，包括利用国家实验室、国家重点实验室和部门重点实验室等研究基地的计划与落实情况）；

申请人所在的上海交通大学团队的教师及研究生长期从事计算机软件应用相关课题的研究，有长期的工作实践，以及大量的理论成果和技术积累，研究成果较丰富，并拥有多个自有版权的软件平台和专利，积累了智慧应用方面的技术及实践。依托教育部人工智能重点实验室和上海交通大学人工智能研究院，可以为项目研究提供支撑实现技术。