



南京大學

本科畢業論文

院 系 匡亞明學院

專 業 計算機科學與技術

題 目 基於超聲波和 IMU 的室內定位系統

年 級 2016 級 學 號 161240005

學生姓名 陳勇虎

指導老師 謝磊 職 稱 副教授

提交日期 2020 年 5 月 20 日

南 京 大 学

本科生毕业论文（设计）指导情况记录

开 题 简 况	论文题目：
	<div>1、选题质量（简述选题与专业培养目标、专业要求关系、题目难度、工作量、创新性、理论性、实用性）</div> <div>2、开题意见：</div> <div>指导教师签名：</div> <div>年 月 日</div>
中 期 检 查	<div>指导教师检查论文的进展情况：（指导和培养学生查阅文献资料、综合运用知识、研究方案设计、研究方法和手段运用和外文应用等能力简况）</div> <div>指导教师签名：</div> <div>年 月 日</div>

南 京 大 学

本科生毕业论文（设计）指导教师评阅意见

指导教师评语：

指导教师签名：

年 月 日

南 京 大 学

本科生毕业论文（设计）评阅教师评阅意见

评阅教师评语：

评阅教师签名：

年 月 日

南 京 大 学

本科生毕业论文（设计）答辩记录、成绩评定

答辩记录：

答辩记录人签名：

答辩小组评语：

答辩小组成员：_____

成绩_____

组长签名：

答辩时间： 年 月 日

南京大学本科生毕业论文(设计、作品)中文摘要

题目：基于超声波和 **IMU** 的室内定位系统

院系：匡亚明学院

专业：计算机科学与技术

本科生姓名：陈勇虎

指导老师（姓名、职称）：谢磊副教授

摘要：

手写中文文本的多样性、差异性让它的识别成为一个富有前景但又充满挑战的计算机视觉问题。

关键词： 手写中文；文本识别；深度学习

南京大学本科生毕业论文(设计、作品)英文摘要

THESIS: IMU

DEPARTMENT: School of Computer Science

SPECIALIZATION: Computer Science and Technology

UNDERGRADUATE: Yonghu Chen

MENTOR: Professor Lei Xie

ABSTRACT:

The diversity of handwritten Chinese text make it a promising but challenging computer vision problem.

KEY WORDS: Handwritten Chinese, Text recognition, Deep learning

目 录

目 录	V
1 绪论	1
1.1 研究背景及意义	1
1.2 国内外研究现状	2
1.2.1 dccc	2
1.3 研究内容与本文主要工作	2
1.4 论文的整体结构	2
2 室内声源定位相关研究工作	3
2.1	3
2.2	3
2.3	3
2.4	3
3 相关工作	5
3.1 图片	5
3.2 公式	6
3.3 表格	7
3.4 算法	7
4 系统实现与实验分析	9
4.1 实现细节	9
4.2 文本分行结果	9
4.3 识别结果	9
4.3.1 准确率	9
5 总结与展望	11
5.1 工作总结	11
5.2 前景展望	11

参考文献	13
致 谢	15

第一章 绪论

1.1 研究背景及意义

在室内环境 (办公室, 住所) 中想要寻找一个小物品, 诸如钥匙, 硬币, 可以说是一件令人费神的事情了^[1]。在游戏世界和虚幻的电影场景中, 我们都或多或少接触过类似于“寻宝罗盘”的神奇宝物, 它可以类似一个掌上的指南针, 时刻指示“宝物”的方向。随着各式各样移动智能设备的广泛涌现, 例如智能手机, 智能手表, 这些设备不仅自带了诸多传感器, 同时可以提供很好的 UI 界面。因此如果可以利用这些设备实现类似于寻宝罗盘的功能, 用于寻找我们室内环境中的小物品, 无疑是有趣且富有创造性的。

在过去的 40 年中^[2], 手写中文文本识别 (HCTR) 的研究获得了很大的进展, 效果得到了很大的提升^[3]。但是, 由于手写中文文本的多样性, 它依然是一个具有研究意义和挑战性的问题^[4]。不同的文本有不同的书写风格, 如图 1-1。

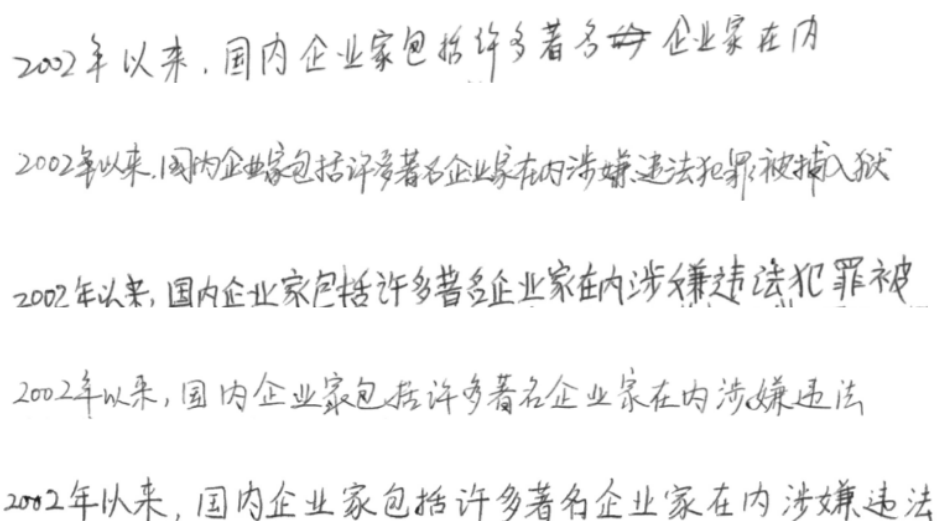


图 1-1: 不同的书写风格。对于同一句话, 有不同的书写风格: 倾斜, 写错字, 工整, 潦草等。

1.2 国内外研究现状

1.2.1 dccc

作为深度神经网络中处理序列的一个重要模型，循环神经网络，如图??，在训练和测试过程中不需要知道视觉序列对象中每个元素的位置。但是，对于循环神经网络，非常重要的一点是将输入图片通过图片预处理转化为一串图片特征^[5,6]。但是通常的基于循环卷积神经网络的网络，因为预处理不在系统训练流程之内，所以无法用从头到尾的方式进行训练，不是很方便。

1.3 研究内容与本文主要工作

本文旨在对图片级手写中文文本做出识别分类。主要工作如下：

1. 在标准公开数据集里获得了一定的识别准确度。
2. 在标准公开数据集上击败了一些相关工作的结果。
3. 建立分本分行规划表格，很好地处理了文本的分行，降低了训练开销。

1.4 论文的整体结构

本文的各章节组织结构如下：

第一章：绪论。简要说明了手写中文文本识别的研究来由及相关工作。并概括地描述了这篇文章的工作，总结了本文结构。

第二章：识别系统

第三章：实验。介绍了实验进行的配置环境，文中使用的测度，文本分行的结果，识别的结果并分析了得到这种结果背后的缘由。

第四章：总结与讨论。总结全文工作，讨论存在的问题和今后可以继续研究的方向。

第二章 室内声源定位相关研究工作

2.1

2.2

2.3

2.4

第三章 相关工作

如何使用这个模板

3.1 图片

一行一图，如图 3-1

2002年以来，国内企业信用评级者为企业家在因涉嫌违法犯罪被
司法机关不断调查，此方面的报道也是屡见不鲜，不是明升暗降
哪个被判了，或者是这个案子开庭了，那个案子判决了。总之，几乎 则都有
样的新闻。
企业家落马、判刑、入狱，甚至死无葬身之地，媒体关注的
往往不是法律问题，而是更多的企业家经营和管理上的问题，在媒体上
是各种各样的，不是经济学家、管理学家，而是与法律定论来对比，这
中不正常的现象。企业家不管在经营、管理上有什么问题，最终的
与后果是在监狱，最终的法律记录如果是经由法院判决有罪，那
重要而应该是在法律问题！

图 3-1: 待分行文本

一行两个图

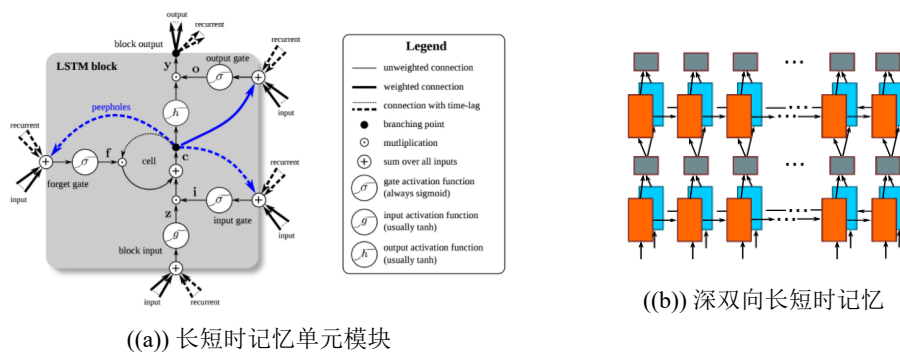


图 3-2: (a) 一个长短时记忆单元模块。(b) 深度双向长短时记忆的结构。

多行多图

2002年以来,国内企业包括许多著名企业家在内涉嫌违法犯罪被
 狱的人数不断增多,这方面的报道也是屡见不鲜,不过,明什么
 哪什么被判了,或者是个案子开庭了,那个案子判了。总之,几乎 明部有
 样的新闻。
 企业家落马,判刑,入狱,甚至犯罪被判处死刑,媒体关注的
 往往不是法律问题,而更多的企业家经营和管理上的问题。在媒体上
 表各种意见的,不乏经济学家、管理专家,但很少有法律专家参与讨论,这
 中不正带的现象。企业家不是在经营、管理上有什么大问题,最后判
 为入狱是在进监狱,最后判了,如果要是经由法院判处有罪,那
 重要的应该是法律问题!

((a)) 全局损失切割

2002年以来,国内企业包括许多著名企业家在内涉嫌违法犯罪被
 狱的人数不断增多,这方面的报道也是屡见不鲜,不过,明什么
 哪什么被判了,或者是个案子开庭了,那个案子判了。总之,几乎 明部有
 样的新闻。
 企业家落马,判刑,入狱,甚至犯罪被判处死刑,媒体关注的
 往往不是法律问题,而更多的企业家经营和管理上的问题。在媒体上
 表各种意见的,不乏经济学家、管理专家,但很少有法律专家参与讨论,这
 中不正带的现象。企业家不是在经营、管理上有什么大问题,最后判
 为入狱是在进监狱,最后判了,如果要是经由法院判处有罪,那
 重要的应该是法律问题!

((b)) 局部损失切割

图 3-3: 分行结果比较。(a) 全局损失切割; (b) 局部损失切割; (c) 局部水平投影切割; (d) 投影损失切割

3.2 公式

$$\frac{\partial L}{\partial a_k^t} = d(s)^2 (y_k^t - \frac{\sum_{lab(1,k)} \alpha_t(s) \beta_t(s)}{y_k^t}) \quad (3-1)$$

$$d_{0j} = \sum_{k=1}^j w_{ins}(a_k), \quad \text{for } 1 \leq j \leq n$$

$$d_{ij} = \begin{cases} d_{i-1,j-1} & \text{for } a_j = b_i \\ \min \begin{cases} d_{i-1,j} + w_{del}(b_i) \\ d_{i,j-1} + w_{ins}(a_j) \\ d_{i-1,j-1} + w_{sub}(a_j, b_i) \end{cases} & \text{for } a_j \neq b_i \end{cases} \quad \text{for } 1 \leq i \leq m, 1 \leq j \leq n. \quad (3-2)$$

$$\begin{aligned}\beta_T(|l'|) &= y_b^T \\ \beta_T(|l'| - 1) &= y_{l|l}^T \\ \beta_T(s) &= 0, \forall s < |l'| - 1\end{aligned}\tag{3-3}$$

递归公式

$$\beta_t(s) = \begin{cases} (\beta_{t+1}(s)d(s) + \beta_{t+1}(s+1)d(s+1))y'_s, & \text{if } l'_s = b \text{ or } l_{s+2}' = l'_s \\ (\beta_{t+1}(s)d(s) + \beta_{t+1}(s+1)d(s+1) + \beta_{t+1}(s+2)d(s+2))y'_s, & \text{otherwise} \end{cases}\tag{3-4}$$

3.3 表格

		国	内	企	业	包	括	许	多
	0	1	2	3	4	5	6	7	8
国	1	0	1	2	3	4	5	6	7
著	2	1	1	2	2	3	4	5	6

表 3-1: 编辑距离（乐文斯汀距离计算过程示例表格。字符串“国内企业包括许多”与“国著名括许多”乐文斯汀距离是 3。

3.4 算法

算法 3.1 Beam Search

- 1: 将初始节点插入到集束中。
 - 2: **while** 遍历未结束 **do**
 - 3: 遍历集束中所有节点的后续节点。
 - 4: **if** 该节点是目标节点 **then**
 - 5: 算法结束。
 - 6: **else**
 - 7: 扩展该节点，取集束宽度的节点入堆。
 - 8: **end if**
 - 9: **end while**
-

集束宽度可以在搜索过程中保持为一个定值，也可以根据搜索的进行而变化。搜索算法可以根据搜索的结果进行调整，比如，当以一个小的集束宽度搜索解却无法找到适合解的时候，可以增大集束宽度重新进行一次搜索。

第四章 系统实现与实验分析

4.1 实现细节

我们在 Tensorflow 框架上实现了我们的网络系统。实验在一个搭载 2.40GHz 英特尔志强 Xeon E5-2673 CPU, 32GB RAM 和一块英伟达 1080Ti 12GB 显存的服务器电脑上运行。网络系统使用 Adam 训练算法。

4.2 文本分行结果

尽管如此, 在局部损失切割和局部水平投影切割之后, 每一个竖直段的分行结果的对应关系却很难处理。在一些特殊情况下, 无法做到每一竖直段分行关系的对应。所以这两个方法不适用。

4.3 识别结果

4.3.1 准确率

我们根据数据集中人的笔迹将数据集分为了 HWDB1-HWDB3, 并实现了 Wang 等人^[7] 和 Mishra 等人^[8] 的方法, 通过调用百度的文字识别系统^[9], 进行对比实验得到以下结果。

方法	HWDB1	HWDB2	HWDB3
Wang 等人 ^[7]	74.0	60.0	68.0
Mishra 等人 ^[8]	80.8	63.6	73.5
百度通用文字识别 ^[9]	64.8	36.8	60.8
我们的方法（没有字典信息）	81.5	67.5	73.6
我们的方法	81.8	67.8	73.9

表 4-1: 识别准确率

第五章 总结与展望

5.1 工作总结

5.2 前景展望

参考文献

- [1] HONGZI Z, YUXIAO Z, ZIFAN L, et al. HyperEar: Indoor Remote Object Finding with a Single Phone[J]. IEEE 39th International Conference on Distributed Computing SysTems(ICDCS), 2019.
- [2] DAYU Y, WEI S, XUDAN W, et al. Review of development status of indoor location technology in China[J]. Navigation and Position, 2019, 7(4): 5–12.
- [3] FUJISAWA H. Forty years of research in character and document recognition: an industrial perspective[J]. Pattern Recognition, 2008, 41(8): 2435–2446.
- [4] XU L, YIN F, WANG Q-F, et al. A touching character database from Chinese handwriting for assessing segmentation algorithms[C] // Frontiers in Handwriting Recognition (ICFHR), 2012 International Conference on. 2012: 89–94.
- [5] GRAVES A, LIWICKI M, FERNÁNDEZ S, et al. A novel connectionist system for unconstrained handwriting recognition[J]. IEEE transactions on pattern analysis and machine intelligence, 2009, 31(5): 855–868.
- [6] SU B, LU S. Accurate scene text recognition based on recurrent neural network[C] // Asian Conference on Computer Vision. 2014: 35–48.
- [7] WANG T, WU D J, COATES A, et al. End-to-end text recognition with convolutional neural networks[C] // Pattern Recognition (ICPR), 2012 21st International Conference on. 2012: 3304–3308.
- [8] MISHRA A, ALAHARI K, JAWAHAR C. Scene text recognition using higher order language priors[C] // BMVC 2012-23rd British Machine Vision Conference. 2012.
- [9] BAIDU. BAIDU Text Recognition[EB/OL]. BAIDU, 2018 (2018/05/10) [2018/05/10].
<https://cloud.baidu.com/product/ocr.html>.

致 谢

非常感谢在 Dislab 实验室度过的两年时光，让我在实验室里拓宽眼界的同时，也不断纠正了自己的学习方法和学习态度。我的毕业论文撰写和系统实现中，也或多或少的遇到许多困难，也走过很多弯路。我很感谢我的导师谢磊副教授，从我毕设选题阶段，到我研究工作的开展，再到最后的论文撰写中，谢老师都给予了我细心的指导和点拨，给我纠正了很多错误的方向，避免我走了很多弯路。在实验室的两年，谢磊老师也给我的研究工作不断指引方向，在此，我向谢磊老师致以最真挚的谢意。

同样，我也要感谢在 Dislab 实验室的同一课题组的各位师兄师姐，鲁欣然师兄在我的成长道路上指引了很多，从细节方面给予我帮助，王楚豫师兄同样，给了我研究工作上的很多帮助，让我的研究工作变的富有创新，也是他们给我提出的宝贵建议，让我可以不断改进自己。

感谢大学四年来四种相伴的同学和朋友，因为你们的存在，我的大学才会在忙碌中透露出精彩。感谢每一位悉心教导我的老师，感谢钱柱中教授的指引，让我能够去香港游学增长知识和阅历，让我在问题求解的课程中感受到计算机和算法的魅力。