

基于区间二型模糊的手写文字识别不确定性研究

尚建华¹ 杨盛毅^{2*} 杨 雯¹ 吴 亮¹
SHANG Jianhua YANG Shengyi YANG Wen WU Liang

摘 要

随着人工智能技术的不断发展, 手写文字的检测与识别技术在各类实际场景中发挥着不可或缺的关键作用。传统的文本识别系统依赖固定的置信度阈值评判识别结果的可靠性, 但在处理模糊、不规范、相似的手写文字时存在识别错误的问题, 根本原因是无法精准量化预测结果的不确定性。针对这一问题, 文章提出了融合区间二型模糊系统 (IT2FS) 的手写文字不确定性的评估的校正框架。该框架基于 PaddleOCR 获取初始的识别结果和置信度。将置信度作为 IT2FS 的输入进行模糊推理, 量化每个识别结果的不确定性, 并构建人机协同校正机制, 引导用户优先处理不确定性较高的文字区域。实验表明, 该框架能有效识别 OCR 引擎原始结果中的潜在错误 (即便在总体置信度较高时), 并通过有针对性的人工校正干预, 将输入手写文字图像的平均准确率从 68.7% 提升至 90%, 能有效减少传统的识别结果的复核时间。

关键词

手写文字识别; 区间二型模糊系统; 不确定性量化; 人工校正

doi: 10.3969/j.issn.1672-9528.2026.01.008

0 引言

在数字时代快速发展背景下, 文本检测与识别技术作为连接物理世界与数字信息的关键桥梁, 已成为计算机视觉和模式识别领域的重要研究方向。该技术旨在从自然场景图像

或视频中精准定位文字区域并转换为可编辑文本, 不仅为后续处理提供基础, 更在自动驾驶、智能文档处理、工业自动化和多模态人机交互等实际场景中发挥出不可替代的作用。与通用目标检测任务不同, 文本检测面临独特挑战: 自然场景中的文本呈现出极大多样性, 包括字体、颜色、尺度、长宽比和语言的多变; 文本朝向的任意性 (水平、倾斜、弯曲); 以及复杂背景、光照变化、遮挡和模糊等因素的干扰。而手写文字识别更具特殊性, 需应对书写风格的多样性、字符形态变化和背景复杂性等挑战^[1]。

1. 贵州民族大学数据科学与信息工程学院 贵州贵阳 550000
 2. 贵州民族大学物理与机电工程学院 贵州贵阳 550000
- [基金项目] 贵州省数据科学与计算智能科技创新人才团队项目 (黔科合平台人才 - CXTD [2025] 038)

- [9] CAI L R, LUO L, LI Z J, et al. Game-based dynamic clustering routing strategy for mobile wireless sensor networks[J]. IEEE internet of things journal, 2024, 11(5): 26481-26491.
- [10] 李建忠, 李慧超, 高生虎. 基于无线传感器网络的远程监测与控制系统设计与实现 [J]. 集成电路应用, 2024, 41(2): 280-281.
- [11] 陈小健, 宋承继. 基于能耗优化算法的移动无线传感器网络问题研究 [J]. 工业仪表与自动化装置, 2019(1): 123-126.
- [12] HEINZELMAN W B, CHANDRAKASAN A P, BALAKRISHNAN H. An application-specific protocol architecture for wireless microsensor networks[J]. IEEE transactions on wireless communications, 2002, 1(4): 660-670.
- [13] SABET M, NAJI H R. A decentralized energy efficient hierarchical cluster-based routing algorithm for wireless sensor

- networks[J]. AEU-international journal of electronics and communications, 2015, 69(5): 790-799.
- [14] 张丹, 刘雅喆, 董雷刚, 等. 无线传感器网络定位算法在环境监测中的应用研究 [J]. 电脑知识与技术, 2017, 13(2): 253-255.
- [15] 梁修荣. 基于同态加密的无线传感器网络数据安全存储方法 [J]. 自动化与仪器仪表, 2023(12): 76-80.
- [16] KHASTEH S H, ROKHSATI H. On transmission range of sensors in sparse wireless sensor networks[J]. Results in engineering, 2023, 18: 101108.

【作者简介】

罗仑 (1994—), 男, 广东汕头人, 硕士, 研究方向: 网络与信息安全、计算机应用、大数据应用。

(收稿日期: 2025-05-27 修回日期: 2025-12-29)

文本检测技术主要分为两大路线：基于回归的方法，通过直接回归文本边界框坐标进行定位，如CTPN、TextBoxes专注于水平文本检测，TextBoxes++、EAST等通过改进锚点设计实现对任意朝向文本的检测，CTD、LOMO等则通过预测多边形顶点实现不规则形状文本的精准刻画^[2]。基于分割的方法则通过像素级分类生成文本区域概率图，再通过后处理得到文本轮廓，PixelLink、PSENet等通过链接像素或渐进式尺度扩张解决相邻文本粘连问题^[3]。DBNet提出可微分二值化模块，将二值化步骤嵌入网络实现端到端学习，显著提升性能；PAN通过轻量级设计在精度和速度间取得平衡^[4]。

文本识别技术可分为规则文本识别（处理印刷字体、扫描文档等整齐文本）和不规则文本识别（应对弯曲、旋转、透视变换等复杂情况）^[5]。主要的深度学习方法如下：基于CTC的编码-解码架构采用CNN提取视觉特征，结合RNN建模序列依赖，通过CTC解码实现序列对齐，CRNN作为经典代表在规则文本上效果优异^[6]；基于注意力机制的序列到序列方法通过注意力机制动态聚焦图像不同区域，提升对长文本和不规则文本的识别能力；基于分割的识别方法将文本识别转化为像素级字符分割问题^[7]；基于Transformer的方法通过自注意力机制建立全局依赖关系，SRN、ABINet等模型通过视觉-语义双模态融合在挑战性场景中表现出色^[8]。手写文字识别技术从传统依赖手工设计特征（HOG、LBP）和统计模型（SVM、HMM）的方法，发展为基于CNN和RNN的深度学习架构^[9]。CRNN模型有效解决变长文本序列识别难题，研究者通过引入ResNet、U-Net、特征金字塔和自注意力机制等模块，提升模型在复杂版面和多变书写风格下的性能^[10]。针对汉字识别的特殊性，细粒度技术如笔画分割和部首表示学习实现对罕见字乃至零样本汉字的识别^[11]。为适应移动和边缘设备需求，轻量化模型架构设计和基于笔画的数据增强技术成为重要研究方向^[12]。尽管深度学习取得了显著的进展，但在处理实际不确定性的手写文字时仍然存在识别错误。这些不确定性可能来源于图像质量不佳、书写模糊、字符间距不规则，以及模型自身对复杂模式的理解不足。传统的OCR系统通常依赖单一的置信度阈值来判断识别结果的可靠性，这种方法难以准确反映识别过程中的实际不确定性^[13]。可能导致识别的文字不准确，从而影响系统的可靠性和实用性。

针对这一问题，本文引入了区间二型模糊进行不确定性评估机制处理。利用PaddleOCR框架中的DBNet算法进行文本检测，CRNN模型进行文本识别^[14]。在此基础上构建一个IT2FS来评估识别结果的置信度和不确定性。通过对识别结果的置信度进行度量化，并结合人工校正反馈，从而提高系统在处理不确定性手写文字的鲁棒性和准确性^[15]。

1 问题描述与系统构建

1.1 问题描述

在手写文本的文本的检测与识别处理中，问题描述为：系统输入一张包含手写文本区域的图像，每个文本区域均包含其边界框坐标信息和待识别的文本内容；同时，系统配备有经过训练的手写文本识别模型。文本识别的核心目的是设计一种有效的识别方法，将检测到的文本区域合理地分配给识别模型，确保所有文本区域的内容能够被准确、高效地识别，并输出结构化的可编辑文本。

1.2 系统构建

本文提出的基于区间二型模糊系统的手写文字识别不确定性评估系统采用模块化设计，主要包括图像预处理模块、文本检测模块、文本识别模块、不确定性评估模块、人机协同模块。整个系统的工作流程如下：

首先，将输入的手写文字图像进行预处理模块进行灰度转换、对比度增强和锐度增强等操作；然后，预处理后的图像被送入文本检测模块，精确定位图像中的文本区域。将检测到的文本区域通过CRNN模型进行文字识别，输出识别结果和对应的置信度；随后，识别结果的置信度被输入到IT2FS不确定性评估模块，通过模糊推理得到量化的不确定性值；最后，基于不确定性评估结果去引导人机协同校正。

2 模型设计

2.1 区间二型模糊理论

区间二型模糊是传统一型模糊系统的扩展，旨在更好地处理不确定性。与一型模糊使用精确的隶属度函数不同，二型模糊使用模糊隶属度函数，即隶属度本身也是一个模糊集。当这个模糊隶属度函数是一个区间时，就形成了区间二型模糊系统。这种设计使其在处理高度不确定性信息时具有更强的鲁棒性，适用于手写文字识别这种存在大量模糊和变异数据的场景。区间二型模糊集可以用公式表示为：

$$\tilde{A} = \int_{x \in X} \int_{u \in J_x} \frac{1}{(x, u)} \quad (1)$$

式中： $J_x \subseteq [0, 1]$ 是主隶属度区间； x 是输入变量； u 是次隶属度变量。在实现中区间二型模糊集使用上下隶属度函数表示为：

$$\mu_{\tilde{A}}(x) = [\underline{\mu}(x), \overline{\mu}(x)] \quad (2)$$

式中： $\underline{\mu}(x)$ 和 $\overline{\mu}(x)$ 分别表示下隶属度函数和上隶属度函数。这两个函数共同定义了模糊隶属度的不确定性范围。与传统模糊系统的单一数值隶属度不同，IT2FS的隶属度是一个区间，意味着每个输入值不再映射到一个精确的隶属度，而是映射到一个隶属度区间。

2.2 区间二型模糊系统设计

2.2.1 模糊集构建

本文构建的置信度对应的不确定性规则如表 1 所示。图像处理中的各个步骤的流程图,如图 1 所示。

表 1 置信度对应表

规则	置信度水平	不确定性
1	极低	极高
2	低	高
3	中等	中等
4	高	低
5	极高	极低

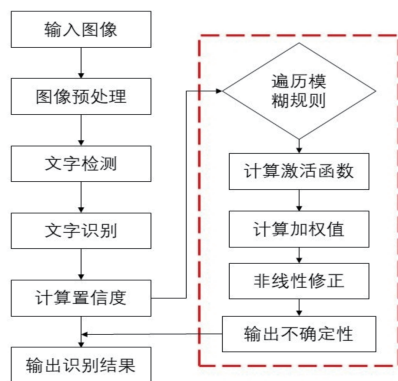


图 1 图像处理流程

2.2.2 不确定性的校正策略

根据 IT2FS 评估的不确定性值,将识别结果分为 3 个等级:当不确定性值超过 0.4 时,系统强制要求人工校正或再次输入识别;不确定性值在 0.2~0.4 之间时,系统建议人工校正;不确定性值低于 0.1 时,系统认为识别结果可靠。这种分级策略能够有效地将有限的人力资源集中到最需要关注的区域,提高校正效率。当用户对识别结果进行校正时,系统会记录校正前后的文本差异,并计算校正率。

3 实验与结果分析

3.1 实验环境与数据集

本实验基于 Python 编程语言实现和 Tkinter 库构建图形用户界面,并利用 PaddleOCR 作为核心的文字识别系统。选取 CASIA-HWDB 数据集和自建测试集进行试验。

3.2 评估指标

为全面评估模型的性能,本文使用了以下评价指标:识别准确率(Accuracy):这是衡量模型正确识别字符的能力的标准指标,计算公式为:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (3)$$

式中:TP 表示正确识别的字符数量;TN 表示正确识别的空白区域;FP 表示误识别的字符数量;FN 表示漏识别的字符数量。

准确率(Precision)表示模型识别正确字符的精确度或可靠性。计算公式为:

$$P = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (4)$$

召回率(Recall)表示模型发现正例的能力或覆盖率。计算公式为:

$$R = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (5)$$

F_1 分数是准确率和召回率的调和平均数。用于在准确率和召回率之间寻找一个平衡点。计算公式为:

$$F_1 = 2 \cdot \frac{P \cdot R}{P + R} \quad (6)$$

不确定性预测精度(Uprecision):该指标主要衡量模糊系统的不确定性评分与实际识别误差之间的匹配程度。计算公式为:

$$U = \frac{\sum \mu_p(x) \cdot A(x)}{N} \quad (7)$$

式中: $\mu_p(x)$ 为最终的置信度区间; $A(x)$ 为识别结果的准确性; N 为样本总数。

3.3 实验结果与分析

首先通过一个具体案例展示框架的工作流程与效果。输入一张原始图像如图 2 所示,对其中的手写文本进行检测如图 3 所示,将检测到的文本框进行识别,识别结果如图 4 所示。将识别的内容与真实的文本进行对比。发现识别的结果与真实文本存在一定的误差。这主要是由于相似字形的误判、文本检测区域不完整、手写字体的结构不规范导致的识别不准确。

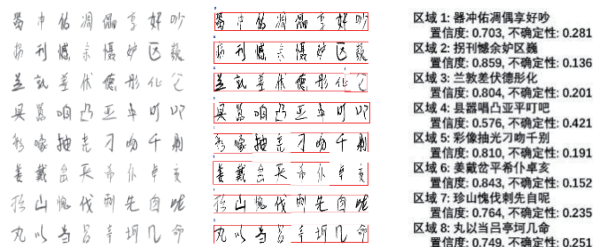


图 2 原始图像 图 3 检测图 图 4 识别图

表 2 每行文字指标得分

文本区域	准确度	精确率	召回率	F_1 分数
第 1 行	0.625	0.625	0.625	0.625
第 2 行	0.625	0.714	0.625	0.667
第 3 行	0.750	0.857	0.750	0.800
第 4 行	0.500	0.500	0.500	0.500
第 5 行	0.625	0.625	0.625	0.625
第 6 行	0.875	0.875	0.875	0.875
第 7 行	0.625	0.625	0.625	0.625
第 8 行	0.875	0.875	0.875	0.875

表 3 性能对比总结

性能	简单规则模式	模糊规则
平均字符准确度	0.687	1.00
平均置信度	0.763	0.746
平均不确定性	0.253	0.246
人工校正率	N/A	90.00%

由表 2 每行文字指标得分可以发现，第 4 行的原始置信度最低，系统给出了最高的不确定性评估值。经检查该行存在书写连笔严重且存在字形歧义，所以模型给出了错误识别。第 1 行和第 7 行的原始置信度中等，但系统也给出了较高的不确定性值。将其正确归入“建议复核”区间。第 2、3、5、6 行原始置信度高，且不确定性评估值低。这表明系统能有效识别出可靠结果，有效避免了不必要的复核。根据分级策略，系统会自动接受第 2、3、5、6 行的结果；标记第 1、7 行为建议校正；标记第 4 行为必须校正。由表 3 性能对比报告显示了两种模式的差异，模糊规则模式通过引入人工校正，将 OCR 原始字符 68.7% 的平均准确率提升至符合预期的水平。尽管两种模式基于完全相同的原始识别结果，但模糊规则系统能精准识别出原始结果中不可接受的错误，证明其在处理手写字体时，通过动态权重调整和多重规则推理，可以对识别结果可靠性进行一个智能判断，能有效降低人工复核成本。

4 总结

本文设计并实现了一个集成了区间二型模糊系统（IT2FS）的手写文字识别不确定性评估框架，有效解决了现有 OCR 系统在面对不确定性输入时可靠性不足的难点。在未来的工作中可以将其推广至其他存在不确定性问题的识别领域，如医疗影像报告识别、语音识别、工业产品缺陷检测等，为更广泛的人工智能应用提供更加可靠的方法。

参考文献:

- [1]LONG S B, HE X, YAO C. Scene text detection and recognition: the deep learning era [J]. International journal of computer vision, 2020, 129: 161–184.
- [2]吴湘平. 图像文本识别的关键技术研究 [D]. 哈尔滨: 哈尔滨工业大学, 2021.
- [3]陈太兵. 基于深度学习的手写中文体识别的研究 [D]. 太原: 太原理工大学, 2024.
- [4]LIAO M H, ZOU Z S, WAN Z Y, et al. Real-time scene text detection with differentiable binarization [J]. IEEE transactions on pattern analysis and machine intelligence 2022, 45(1): 919 - 931.
- [5]ATIENZA R. Vision transformer for fast and efficient scene text recognition [C]//Document Analysis and Recognition –

ICDAR 2021. Berlin:Springer, 2021: 319–334.

- [6]AGRAWAL V , JAGTAP J, KANTIPUDI M P. An overview of hand-drawn diagram recognition methods and applications [J]. IEEE access, 2024, 12: 19739–19761.
- [7]柳璨, 吴成茂. 基于分数阶信息的鲁棒图形模糊聚类分割算法 [J]. 计算机工程与设计, 2019, 40(3): 774–781.
- [8]崔立志, 李俊. 基于自注意力与多粒度卷积的文本识别方法 [J]. 兵器装备工程学报, 2023, 44(S1): 300-306.
- [9]DA SILVA B T, OLIVEIRA A L I, DE SOUZA R M C R , et al. Adaptive confidence calibration for handwritten text recognition using fuzzy logic[J]. Engineering applications of artificial intelligence, 2024, 129: 107-118.
- [10]ZHOU X Y, YAO C, WEN H, et al. EAST: an efficient and accurate scene text detector[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway:IEEE, 2017:5551–5560.
- [11]LIU Y L, SHEN C H, JIN L W, et al. ABCNet v2: adaptive Bezier-curve network for real-time end-to-end text spotting[J]. IEEE transactions on pattern analysis and machine intelligence, 2022, 44(11): 8048-8064.
- [12]LIAO M H, PANG G, HUANG J,et al. Mask textspotter v3: segmentation proposal network for robust scene text spotting[EB/OL].(2020-07-18)[2025-12-03].<https://doi.org/10.48550/arXiv.2007.09482>.
- [13]王健, 李静, 张鹏, 等. 基于可微分二值化的自然场景文本检测算法研究 [J]. 计算机学报, 2023, 46(5): 1021-1035.
- [14]WANG W H, XIE E Z, LI X, et al. PAN++: towards efficient and accurate end-to-end spotting of arbitrary-shaped text[J]. IEEE transactions on pattern analysis and machine intelligence, 2022, 44(9): 5349 - 5367.
- [15]HUANG M X, LIU Y L, PENG Z H, et al. SwinTextSpotter: better synergy between text detection and recognition via transformer[EB/OL].(2022-03-19)[2025-12-02]. <https://arxiv.org/abs/2203.10209>.

【作者简介】

尚建华（1998—），男，贵州毕节人，硕士研究生，研究方向：统计建模与模式识别。

杨盛毅（1986—），通信作者（email:yangshy@163.com），男，贵州黔南人，博士，教授，研究方向：无人飞行器总体设计、自主控制与决策、环境感知与目标识别。

（收稿日期：2025-10-01 修回日期：2026-01-06）