

贺先域, 刘伟, 陶家俊. HCR-RecNet: 基于深度学习的手写化学反应式识别研究 [J]. 智能计算机与应用, 2026, 16(4): 238-245. DOI: 10.20169/j.issn.2095-2163.26013102

HCR-RecNet: 基于深度学习的手写化学反应式识别研究

贺先域, 刘伟, 陶家俊

(湖南中医药大学 信息科学与工程学院, 长沙 410208)

摘要: 化学反应式是描述物质转化过程的核心表达形式, 广泛存在于教学、考试、实验记录及工业文档中。将手写化学反应式由图像形式转换为一维编码文本, 有助于降低数据冗余, 并为智能批改与化学信息处理提供结构化输入。然而, 手写化学反应式符号类型多样、书写风格差异显著, 且包含分子式、反应符号及反应条件等复杂结构, 传统的整体识别方法难以稳定识别。针对上述问题, 本文提出一种面向手写化学反应式的分治识别框架 HCR-RecNet。该框架由反应式检测模块与后处理模块构成, 其中基于 YOLOv13 的检测模块负责对反应式中的关键符号进行定位与分类, 后处理模块对检测结果进行结构化解析, 包括分子式区域的精细识别、检测结果的空间顺序建模以及在统一符号编码约束下的序列重构, 从而生成规范的一维化学反应式表达。在自建的 8 352 张手写化学反应式数据集上, HCR-RecNet 的整体识别准确率达到 86.24%, 显著优于整体识别方法。实验结果表明, 该方法在保持反应式整体结构一致性的同时, 提高了对关键子结构的识别稳定性, 为手写化学反应式的自动解析提供了一种有效解决方案。

关键词: 手写化学反应式识别; YOLOv13 模型; 后处理模块; 字符序列; 检测-识别框架

中图分类号: TP391.41 **文献标志码:** A **文章编号:** 2095-2163(2026)04-0238-08

HCR-RecNet: Deep learning-based handwritten chemical reaction expression recognition

HE Xianyu, LIU Wei, TAO Jiajun

(School of Informatics, Hunan University of Chinese Medicine, Changsha 410208, China)

Abstract: Handwritten chemical reaction expressions appear frequently in teaching materials, exams, laboratory notes, and industrial documents. Converting them from images to one-dimensional encoded text reduces data redundancy and supplies structured input for intelligent grading and chemical information processing. Reliable recognition remains difficult because symbol categories are diverse, handwriting styles vary widely, and reaction expressions contain coupled components such as molecular formulae, reaction operators, and reaction conditions. This work presents HCR-RecNet, a divide-and-conquer framework for handwritten chemical reaction expression recognition. The framework contains a detection stage and a post-processing stage. The detection stage is built on YOLOv13 to localize and classify key symbols. The post-processing stage performs structured parsing of detections. It refines molecular-formula regions, models the spatial order of symbols, and reconstructs sequences under a unified symbol-encoding constraint to produce standardized one-dimensional representations. Experiments on a self-constructed dataset of 8,352 handwritten reaction images report an overall recognition accuracy of 86.24%, which is markedly higher than that of holistic recognition baselines. The results indicate improved robustness on critical substructures while maintaining global structural consistency, supporting automatic digitization and downstream analysis of handwritten reaction content.

Key words: handwritten chemical reaction expression recognition; YOLOv13 model; post-processing module; character sequence; detection-recognition framework

0 引言

化学反应式是描述物质转化及反应条件的核心

表达形式, 广泛应用于教材、试卷、实验记录和工业文档等场景。在教学评价与化学信息管理等场景下, 将手写反应式准确转化为一维编码文本, 是实现

基金项目: 湖南省自然科学基金面上项目(2022JJ30438); 长沙市自然科学基金(kq2202260)。

作者简介: 贺先域(1999—), 男, 硕士研究生, 主要研究方向: 计算机视觉; 陶家俊(1994—), 男, 硕士, 助教, 主要研究方向: 人工智能, 计算机视觉。

通信作者: 刘伟(1982—), 男, 博士, 教授, 硕士生导师, 主要研究方向: 人工智能, 医药信息学, 化学信息学。Email: weiliu@hnu.cm.edu.cn。

收稿日期: 2026-01-31

智能批改与自动处理的基础。然而,当前手写反应式多以图像形式保存,数据冗余高且难以直接利用,亟需研究高效的自动识别方法以生成机器可读文本表示。

手写化学反应式识别具有较大挑战。其组成符号类型多样、空间结构复杂、且手写风格差异明显,字符形态与相对位置不稳定,使其同时具备文本序列与二维结构特征,导致传统光学字符识别(OCR)方法难以保持稳定性能。

早期研究多采用规则驱动或多阶段框架处理手写化学表达式。Chang等^[1]提出统一识别框架,通过符号分组、结构分析与语义验证实现二维表达式的解析,并利用图结构建模刻画空间关系;亦有研究通过分阶段建模有机结构与无机符号,实现结构式与字符混合表达的识别^[2]。然而,此类方法流程复杂、规则依赖性强,难以适应结构更复杂或书写形式多样的反应式场景。

随着笔式交互设备的发展,部分研究开始关注在线手写化学表达式识别,利用笔画顺序与时序信息辅助切分与解析。Yang等^[3]提出两级在线识别框架,将结构分析与符号识别结合,实现实时解析;姜映映等^[4]在切分与后处理阶段引入化学知识与语法规则。但此类方法通常依赖特定输入形式,难以直接迁移至离线图像场景。

近年来,深度学习方法推动了手写化学反应式识别的发展。部分研究将反应式视为序列识别任务,基于卷积神经网络(CNN)与循环神经网络(RNN)构建端到端模型,并通过编码规则将二维表达映射为一维序列。Wang等^[5]提出轻量级卷积循环神经网络(CRNN)框架,实现高效识别。然而,该类整体式方法在复杂空间布局、多符号类型或局部书写误差情况下,鲁棒性仍存在局限。

针对反应式中的关键子结构识别,已有研究在分子式与结构式识别方面开展探索。早期方法基于隐马尔可夫模型(HMM)及其与支持向量机(SVM)结合的两阶段框架,增强结构识别能力^[6-7];随后引入CNN与数据增强技术,提升特征表达能力与模型鲁棒性^[8]。在此基础上,部分工作结合Transformer与高效主干网络,实现从手绘结构到SMILES的自动转换^[9-10]。Tao等^[11]基于无锚点检测与图重建提升未知分子泛化能力;彭晓旺等^[12]通过域转换与双向生成模型改善识别性能。DECIMER与MolScribe分别构建图像到SMILES的映射,并引入几何约束以增强结构合法性^[13-14];李愿军等^[15]通过多模型

融合提升复杂结构识别稳定性。此外,Liu等^[16]基于CRNN进行手写化学式识别,但未显式引入化学规则约束。总体而言,上述方法验证了对子结构进行精细识别的可行性,但多聚焦于单一表达形式,尚未形成面向化学反应式整体解析的统一建模框架。

祝磊等^[17]在YOLOv5s框架下引入EfficientNet主干网络、VoV-GSCSP模块及NAM注意力机制,实现模型轻量化并提升尾灯灯语检测性能;冯奥等^[18]针对卷烟目标检测任务优化特征融合结构,提高复杂背景下的小目标识别能力;廖佳逸等^[19]在YOLOv8n中融合CA与CBAM注意力机制并引入WIoU损失函数,提升口罩佩戴检测精度与速度。此外,相关研究通过改进特征金字塔结构与检测头设计,增强模型在小目标、复杂背景及密集场景下的检测性能^[20-23]。

综合来看,现有方法在反应式整体结构建模与关键子结构精细识别之间仍存在不足:单纯序列建模难以充分利用空间结构信息,而缺乏对分子式等关键部分的专门处理,易放大局部误差。为此,本文提出面向手写化学反应式的分治识别框架。该方法首先通过目标检测对各组成部分进行定位与分类,并依据空间顺序构建初始结构序列;随后对分子式等关键区域进行精细识别,并结合统一符号编码与序列重构生成一维表达。该框架在保持整体结构一致性的同时,实现对子结构的专门建模,为手写化学反应式的自动解析提供了有效方案。

1 数据集

目前鲜有公开可用的手写化学反应式数据集能够满足复杂反应式结构识别的研究需求。为此,本文构建了一个手写化学反应式数据集,用于支持后续识别实验。

1.1 数据采集

在数据采集阶段,本文组织志愿者依据预设模板书写手写化学反应式样本,内容涵盖常见无机反应、离子反应及包含气体生成与沉淀标记的典型反应式,以模拟教学与考试等实际场景。原始数据以整页扫描图像形式获取。针对单页包含多行样本的情况,采用基于OpenCV的自动裁切方法批量提取单条反应式。最终从扫描页中获得1392张有效手写反应式图像,扫描示例如图1所示。

1.2 数据增强

考虑到手写化学反应式在实际应用中易受书写风格、扫描质量及噪声干扰等因素影响,原始样本在规模与多样性方面仍存在不足。为此,本文对数据

集进行增强处理,包括高斯模糊与随机噪声扰动等操作,以模拟不同成像与书写条件。经增强后,数据规模由1 392张扩展至8 352张手写反应式图像。

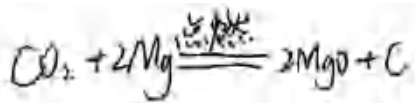


图1 手写化学反应式

Fig. 1 Handwritten chemical reaction equation

1.3 数据集划分

在完成数据增强后,本文按照8:2的比例对数据集进行划分,其中训练集包含6 696张样本,测试集包含1 656张样本。该划分方式在保证样本数量充足的同时,也有助于对模型在未见样本上的泛化性能进行评估。

1.4 标注策略

针对分治识别框架中不同任务的需求,本文根据符号编码表对数据集分别构建了两类标注形式,符号编码见表1。

表1 符号编码方案

Table 1 Symbol encoding scheme

符号	编码	符号	编码	符号	编码
=	=	⇌	~<>	→	~>>
↓	~-!	↑	~!~	Δ	~\
点燃	~!!	通电	~!@	高温	~\
光	~@@	放电	~!	高温高压	~##
催化剂	~%%	电炉	~&&	电解	~\$\$

目标检测标注:为支持反应式中各组成元素的定位与分类,本文对图像中的数字、化学分子式、反

应条件及相关符号进行标注,用于训练目标检测模型实现区域级结构检测。各目标均采用边界框形式标注,记录其位置与类别信息。

整体识别标注:针对整体识别任务,依据预定义符号编码表对整张手写反应式进行字符级标注,将符号顺序映射为一维编码序列,用于描述最终文本表达。

2 方法

2.1 总体工作流程

本文提出的手写化学反应式识别方法HCR-RecNet采用分治识别流程,整体框架如图2所示。该方法以完整手写反应式图像为输入,通过反应式检测与后处理两个阶段,实现结构解析与一维文本化表示。首先,基于YOLOv13的检测模块对反应式各组成元素进行定位与分类,包括数字、化学分子式、反应符号及反应条件等,输出其空间位置与类别信息。随后,后处理模块在检测结果基础上完成结构化解析,包括分子式识别与序列重构,实现反应式的统一编码与文本化表达。

2.2 YOLOv13

YOLOv13是近期提出的目标检测方法^[24],旨在保持实时性能的同时提升复杂场景中多目标结构与语义关系的建模能力。相较于以往主要依赖局部卷积或区域注意力机制的设计,YOLOv13引入超图(Hypergraph)建模思想,将高阶语义关联纳入检测框架,在复杂布局与密集目标场景中具有更强鲁棒性。

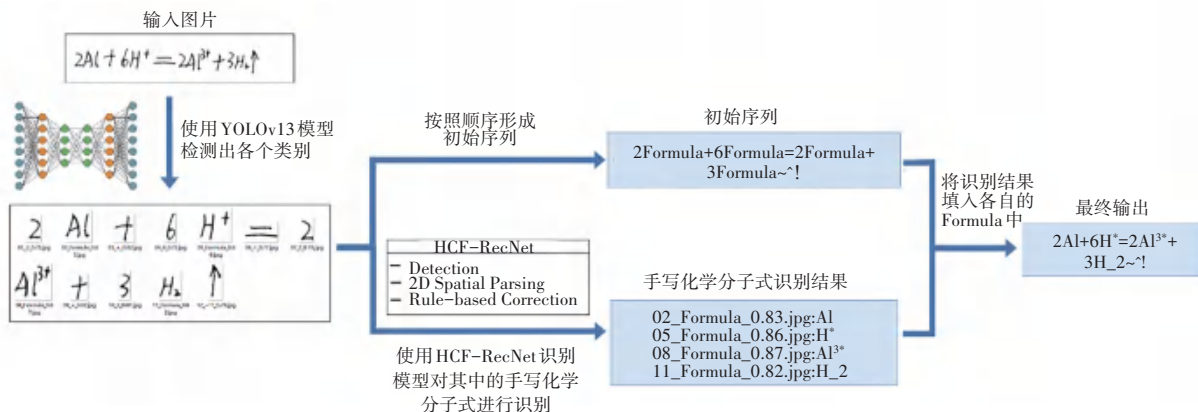


图2 HCR-RecNet整体工作框架

Fig. 2 Overall framework of HCR-RecNet

该模型延续Backbone-Neck-Head三阶段架构,并对核心模块进行了重构。基于轻量化卷积骨干网络,模型在特征提取与融合阶段引入超图驱动

的高阶关联模块(如HyperACE),强化了语义关系建模能力;同时提出全链路特征协同机制,通过双向信息交互,改善多尺度融合中小目标特征弱化问题。

在实现层面,采用深度可分离卷积等轻量化算子,在引入超图建模的同时保持较低参数量与计算开销,从而满足实时检测需求。

2.3 后处理模块

在完成反应式各组成元素的检测与定位后,本文引入后处理模块对检测结果进行结构化解析。该模块基于检测输出的类别与空间信息,对反应式整体结构及关键子结构进行重构,而不参与检测过程。

后处理模块包括手写化学分子式识别与序列重构两部分。前者对检测得到的分子式区域进行独立识别,获取字符级表达;后者结合检测框的空间顺序与统一符号编码,对各组成部分进行排序与整合,生

成结构一致的一维反应式序列。

2.3.1 手写化学分子式识别方法

化学分子式是反应式中的关键子结构,内部包含上下标与电荷等二维依赖字符,其识别精度直接影响整体解析结果。为此,本文在 HCR-RecNet 中将分子式识别作为独立子任务,引入前期提出的 HCF-RecNet 模型完成该任务。

在检测阶段,模型将判定为 Formula 类别的区域作为候选分子式,根据检测框位置从原图裁剪子图,并输入 HCF-RecNet 进行识别。HCF-RecNet 总体工作流程如图 3 所示。

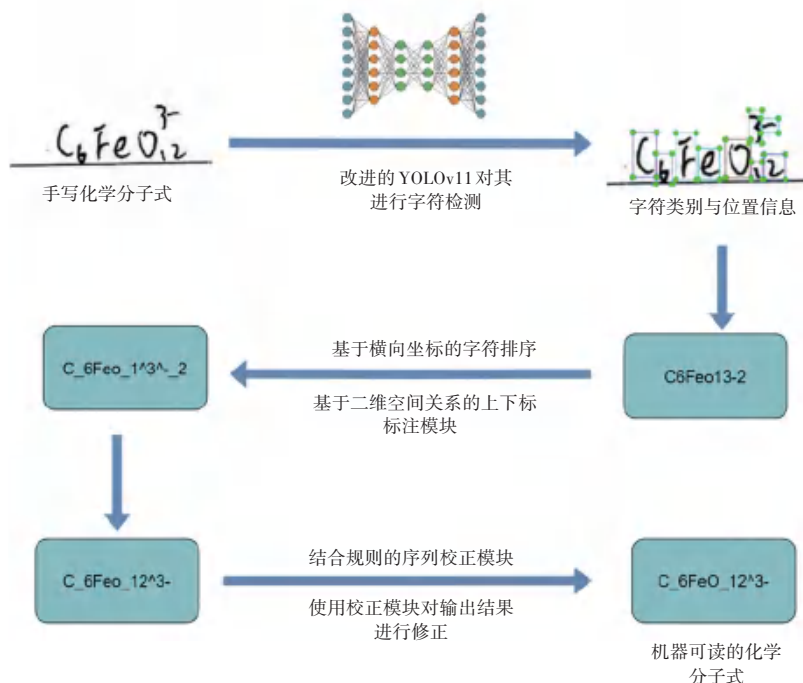


图3 HCF-RecNet 整体工作框架

Fig. 3 Overall framework of HCF-RecNet

HCF-RecNet 是面向手写化学反应式的识别框架,通过字符级检测与二维空间关系建模实现分子式内部结构解析。该方法利用改进的目标检测模型定位与分类字符,并在后处理阶段依据相对空间关系重建上下标结构,同时引入化学规则约束以增强语义一致性与稳定性,在多种手写场景下具有较好性能。

在 HCR-RecNet 中,HCF-RecNet 作为后处理模块的重要组成部分,用于分子式区域的精细识别。通过将分子式识别与反应式整体结构建模解耦,在保证结构一致性的同时提升复杂手写反应式的识别稳定性。分子式结果将在后续序列重构阶段与初始结构表示融合,生成规范的一维反应式表达。

2.3.2 序列重构方法

该方法以检测输出的类别标签与空间坐标为基础,通过空间排序与符号编码约束实现结构重建。首先依据检测框的水平位置对结果排序,确定反应式从左至右的书写顺序;针对反应符号及其关联条件,结合纵向位置关系进行内部调整:上方条件视为前置并排列于符号左侧,下方条件视为后置并排列于符号右侧,从而保持结构关系的准确表达。

基于上述排序与调整结果构建初始序列,以检测类别作为占位符表示各组成元素,分子式统一以 Formula 标记。随后依据预定义符号编码表对非分子式符号进行统一映射。最后,将 HCF-RecNet 输出的分子式结果替换序列中的 Formula 占位符,完

成整体整合。

通过上述重构过程, HCR-RecNet 融合检测与分子式识别结果, 生成结构一致的一维手写化学反应式表示, 为后续化学信息处理提供可靠输入。

2.4 实验环境与配置

本研究实验平台基于 Windows10 操作系统, 处理器为 Intel Core i5 - 13600KF (5.1 GHz), 内存 32 GB, 配备 NVIDIA GeForce RTX4080Super (16 GB 显存)。目标检测模块基于 Ultralytics 框架实现, 采用 YOLOv13 系列模型训练。输入尺寸统一为 640×640, 批大小设为 16, 最大训练轮数为 100, 并加载预训练权重进行迁移学习。为保证实验可复现性, 固定随机种子并启用确定性训练策略, 同时开启自动混合精度以提升计算效率。

3 实验结果与分析

3.1 评价指标

为系统评估 HCR-RecNet 在手写化学反应式识别任务中的性能, 本文从整体序列识别效果与检测模块性能两个层面设置评价指标, 对模型输出结果进行定量分析。

3.1.1 完全匹配率

本研究采用完全匹配率 (Exact Match Rate, EMR) 作为模型性能的主要评价指标。当模型输出的一维反应式序列在符号类别与排列顺序上均与人工标注结果完全一致时, 该样本被判定为正确识别, 该指标能够直观反映模型在整体层面上对化学反应式的识别可靠性。设完全正确识别的样本数量为 correct_num, 样本总数为 all, 则 EMR 定义如下式:

$$EMR = \frac{\text{correct_num}}{\text{all}} \quad (1)$$

3.1.2 字符级匹配率

为进一步刻画模型在存在局部识别误差时的表

现, 本文引入字符级匹配率 (Character Match Rate, CMR) 作为补充评价指标。该指标基于预测序列与真实序列之间的公共子序列, 用于衡量字符层面的匹配程度。字符级匹配率是通过计算两个字符串间相同子字符串的长度, 再乘以 2, 并除以这两个字符串长度之和得到的, 具体表达式如下:

$$CMR = 2 \times \frac{\sum_{i=1}^n \mathcal{L}(s_i)}{\mathcal{L}(a) + \mathcal{L}(b)} \quad (2)$$

3.1.3 检测模块评价指标

在反应式检测阶段, 本文采用多种常用指标对基于 YOLOv13 的检测模块进行综合评估。通过混淆矩阵统计真阳性 (TP)、假阳性 (FP)、真阴性 (TN) 与假阴性 (FN), 分析不同符号类别的误检与漏检情况。在此基础上, 准确率 (Accuracy) 衡量整体预测正确性, 精确率 (Precision) 与召回率 (Recall) 分别反映误检控制能力与目标覆盖能力, $F1 - Score$ 综合评价二者平衡性能。为刻画不同空间匹配条件下的检测效果, 采用平均精度均值 (mAP) 作为核心指标, 其中 mAP0.5 表示固定 IoU 阈值下的检测性能, mAP0.5:0.95 依据 COCO 评价协议在多个 IoU 阈值区间内取平均, 更全面表征模型的空间定位能力。此外, 同时报告推理时间与模型参数量, 以评估模型在实际应用中的效率与复杂度。

3.2 不同标检测模型对比实验

为评估检测模块在手写化学反应式场景下的性能, 本文选取多种主流目标检测模型进行对比实验。所有模型均在相同的数据集划分与训练配置下进行, 仅替换检测网络结构, 以保证结果的可比性。对比结果见表 2, 不同模型在复杂手写符号布局下的检测性能存在明显差异, 其中新一代 YOLO 系列模型在整体检测精度与稳定性方面表现更为突出。

表 2 目标检测模型的对比实验

Table 2 Comparison of Object Detection Models

模型	EMR/%	CMR/%	Inference Time/ms	Parameter Count/M	Precision/%	Recall/%	F1 - Score/%	mAP50/%	mAP50-95/%
YOLOv8x ^[25]	83.90	98.46	3.70	68.21	94.81	97.10	95.94	98.30	80.40
YOLOv9e ^[26]	84.32	98.56	6.50	69.50	94.93	97.74	96.31	98.22	80.65
YOLOv10x ^[27]	48.89	91.73	4.90	31.77	88.73	87.54	88.13	93.15	73.11
YOLOv11x ^[28]	85.28	99.02	5.00	58.93	93.54	98.57	95.99	98.01	80.92
YOLOv13x	86.24	98.60	4.30	64.00	95.29	99.01	97.11	96.86	80.37

3.3 各预训练模型对比实验

为进一步提升模型性能,在目标检测模型确定为YOLOv13后,本文进一步比较了YOLOv13不同预训练模型在手写化学反应式检测任务中的性能。

表3 YOLOv13 预训练模型的对比实验

Table 3 Comparison of YOLOv13 pretrained models

模型	EMR/%	CMR/%	Inference Time/ms	Parameter Count/M	Precision/%	Recall/%	F1-score/%	mAP50/%	mAP50-95/%
YOLOv13n	83.42	98.18	0.80	2.53	94.50	96.40	95.44	97.44	78.82
YOLOv13s	85.10	98.44	0.90	9.00	95.10	97.98	96.52	97.20	79.63
YOLOv13l	84.20	98.24	2.80	27.62	94.55	96.99	95.75	97.14	80.41
YOLOv13x	86.24	98.60	4.30	64.00	95.29	99.01	97.11	96.86	80.37

3.4 与 OCR 模型对比实验

为验证所提出HCR-RecNet在手写化学反应式识别任务中的优势,本文选取4种具有代表性的OCR模型作为对比方法,包括CRNN^[29]、GRCNN^[30]、FAN^[31]和TrOCR^[32]。其中,CRNN为经典的卷积循环文本识别框架;GRCNN在特征提取阶段引入门控机制以增

为保证公平性,各模型在相同数据集与训练配置下进行训练,仅预训练权重不同。对比结果见表3。综合考虑检测精度与模型规模,本文选取性能最优的预训练模型作为后续实验的默认配置。

强表达能力;FAN采用注意力机制强化关键区域建模;TrOCR基于Transformer架构构建端到端视觉文本识别模型。实验结果如图4所示,可以观察到,通用OCR方法在手写化学反应式场景下整体表现有限,而HCR-RecNet在EMR指标上显著优于所有对比模型,展现出明显优势。

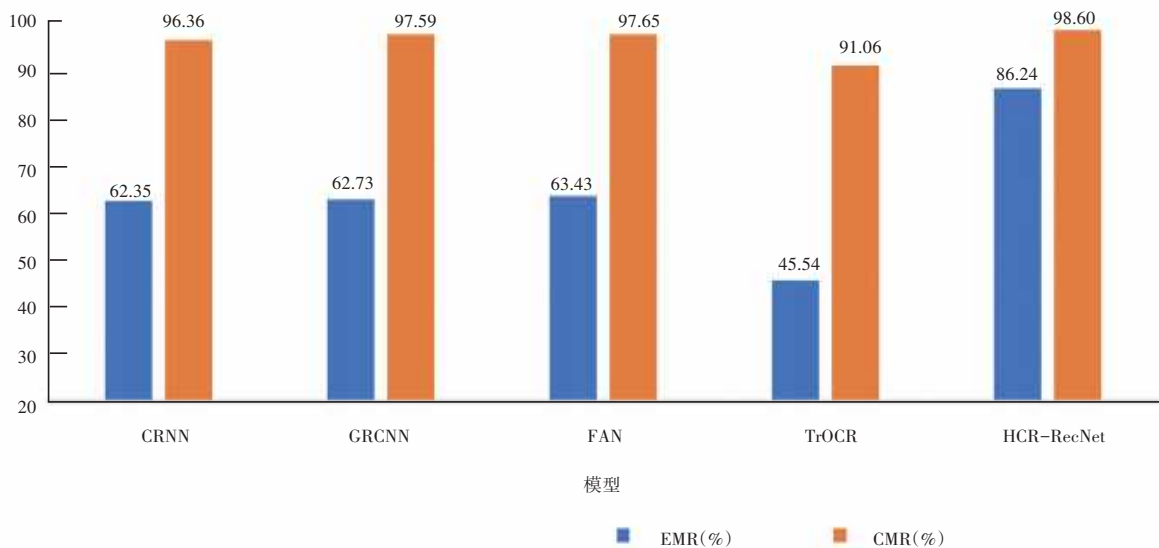


图4 与 OCR 模型的对比实验

Fig. 4 Comparison with OCR models

3.5 结果分析

实验结果表明,通用OCR模型在手写化学反应式识别任务中难以同时兼顾符号多样性与空间结构复杂性,整体识别效果受到明显限制,尤其在分子式、反应符号与反应条件混合出现的场景下,基于整体文本建模的识别方式容易产生结构性错误。相比之下,本文提出的HCR-RecNet采用分治识别策略,将反应式检测与分子式精细识别相结合,并通过后处理模块完成结构化序列重构,在复杂手写反应式

场景下表现出更高的稳定性与准确性。在自建手写化学反应式数据集上,该方法的整体识别准确率达到86.24%,显著优于对比的OCR方法,验证了本文所提出框架在手写化学反应式自动识别中的有效性。

4 结束语

本文围绕手写化学反应式自动识别问题,提出了基于分治思想的识别框HCR-RecNet。针对符号

类型复杂、空间布局多样及局部结构识别困难等特点,将反应式整体结构解析与分子式精细识别进行解耦,通过目标检测与后处理结合,实现手写化学反应式的结构化解析与一维序列表达。在方法上,框架由基于 YOLOv13 的反应式检测模块与后处理模块构成。检测模块完成关键符号定位与分类,后处理模块实现分子式识别与序列重构,并结合空间顺序建模与统一符号编码约束,提高复杂手写反应式的整体识别稳定性。实验方面,构建了包含 8 352 张手写化学反应式的自建数据集,并与多种 OCR 模型进行对比。结果表明, HCR-RecNet 的整体识别准确率达到 86.24%,相较于整体式 OCR 方法具有明显优势,验证了本文分治识别策略的有效性。

本文研究主要针对由分子式、反应符号及反应条件构成的手写反应式,尚未覆盖包含化学结构式的复杂场景。未来将进一步引入手写化学结构式识别与结构建模方法,实现更复杂手写反应式的统一解析。

参考文献

- [1] CHANG Ming, HAN Shi, ZHANG Dongmei. A unified framework for recognizing handwritten chemical expressions[C]//Proceedings of 2009 10th International Conference on Document Analysis and Recognition. Piscataway, NJ: IEEE, 2009: 1345-1349.
- [2] MAPARI S. A neural network-based framework to recognize handwritten chemical expression [C]//Proceedings of World Conference on Information Systems for Business Management. Cham; Springer, 2023: 405-413.
- [3] YANG Jufeng, SHI Guangshun, WANG Kai, et al. A study of on-line handwritten chemical expressions recognition [C]//Proceedings of the 19th International Conference on Pattern Recognition. Piscataway, NJ: IEEE, 2008: 1-4.
- [4] 姜映映,王绪刚,敖翔,等. 手写化学公式的在线切分识别[C]//第2届中国人机交互学术会议(CHCI06)论文集. 2006: 111-116.
- [5] WANG Xiaofeng, HE Zhihuang, WU Zhize, et al. Handwritten chemical equations recognition based on lightweight networks [C]// Proceedings of the International Conference on Intelligent Computing. Cham; Springer, 2022: 317-329.
- [6] ZHANG Yang, SHI Guangshun, YANG Jufeng. HMM-based online recognition of handwritten chemical symbols [C]// Proceedings of the 10th International Conference on Document Analysis and Recognition. Piscataway, NJ: IEEE, 2009: 1255-1259.
- [7] ZHANG Yang, SHI Guangshun, WANG Kai. A SVM-HMM based online classifier for handwritten chemical symbols [C]// Proceedings of the 20th International Conference on Pattern Recognition. Piscataway, NJ: IEEE, 2010: 1888-1891.
- [8] ZHENG L, ZHANG T, YU X. Recognition of handwritten chemical organic ring structure symbols using convolutional neural networks [C]// Proceedings of the 2019 International Conference on Document Analysis and Recognition Workshops. Piscataway, NJ: IEEE, 2019: 165-168.
- [9] RAJAN K, BRINKHAUS H O, ZIELESNY A, et al. Advancements in hand-drawn chemical structure recognition through an enhanced DECIMER architecture [J]. Journal of Cheminformatics, 2024, 16(1): 78. DOI: 10.1186/s13321-024-00872-7.
- [10] OUYANG Hengjie, LIU Wei, TAO Jiajun, et al. ChemReco: Automated recognition of hand-drawn carbon-hydrogen-oxygen structures using deep learning [J]. Scientific Reports, 2024, 14(1): 17126. DOI: 10.1038/s41598-024-67496-7.
- [11] TAO Jiajun, LIU Wei, PENG Xiaowang, et al. Recognition of hand-drawn hydrocarbon structure formulas using anchor-free detector [C]//Proceedings of the 21st Pacific Rim International Conference on Artificial Intelligence. PRICAI, 2024: 98-110.
- [12] 彭晓旺,刘伟,陶家俊,等. 基于Transformer的手绘与标准印刷化学分子结构图像双向循环生成方法[J]. 智能计算机与应用, 2025, 15(4): 77-85. DOI: 10.20169/j. issn. 2095-2163. 25030303.
- [13] RAJAN K, ZIELESNY A, STEINBECK C. DECIMER 1.0: Deep learning for chemical image recognition using transformers [J]. Journal of Cheminformatics, 2021, 13(1): 1-16.
- [14] QIAN Y, GUO J, TU Z, et al. MolScribe: Robust molecular structure recognition with image-to-graph generation [J]. Journal of Chemical Information and Modeling, 2023, 63(7): 1925-1934.
- [15] 李愿军,陈建虎,张亚标,等. 基于多模型融合的手写化学结构式识别方法[J]. 现代信息科技, 2026, 10(1): 58-62. DOI: 10.19850/j. cnki. 2096-4706. 2026. 01. 012.
- [16] LIU Xiaoxue, ZHANG Ting, YU Xinguo. An end-to-end trainable system for offline handwritten chemical formulae recognition [C]// Proceedings of the 2019 International Conference on Document Analysis and Recognition. Piscataway, NJ: IEEE, 2019: 577-582.
- [17] 祝磊,欧阳万棋,敖思铭. 基于YOLOv5s轻量化改进的车尾灯语检测[J]. 智能计算机与应用, 2025, 15(9): 56-63. DOI: 10.20169/j. issn. 2095-2163. 250909.
- [18] 覃宜霜,陶雯,贾建双,等. 基于改进YOLOv5的卷烟品规展示视频检测研究[J]. 智能计算机与应用, 2025, 15(6): 127-133. DOI: 10.20169/j. issn. 2095-2163. 25022201.
- [19] 冯奥,廖佳逸,蒋银莹,等. 基于改进YOLOv8n的口罩佩戴检测算法[J]. 智能计算机与应用, 2025, 15(10): 145-149. DOI: 10.20169/j. issn. 2095-2163. 251022.
- [20] 羊森海,陈丹. 基于改进YOLOv7的抓取图像小目标检测算法[J]. 智能计算机与应用, 2025, 15(7): 99-103. DOI: 10.20169/j. issn. 2095-2163. 250714.
- [21] 杨凡,张洁,蒋安祥. 改进YOLOv8n的迷彩伪装目标检测算法[J]. 智能计算机与应用, 2025, 15(6): 7-14. DOI: 10.20169/j. issn. 2095-2163. 24122503.
- [22] 彭小龙,王正勇,何小海,等. 基于改进YOLOv8的岩石纹层图像分割研究[J]. 智能计算机与应用, 2025, 15(6): 1-6. DOI: 10.20169/j. issn. 2095-2163. 24121204.
- [23] 姜乃祺,陈俊,陈芳,等. 基于改进YOLOv8n的轻量化密集行人检测模型[J]. 智能计算机与应用, 2025, 15(11): 137-141. DOI: 10.20169/j. issn. 2095-2163. 251122.
- [24] LEI Mengqi, LI Siqi, WU Yihong, et al. YOLOv13: Real-time object detection with hypergraph-enhanced adaptive visual perception [J]. arXiv preprint arXiv, 2506.17733, 2025. DOI: 10.48550/arXiv.2506.17733.

-
- [25] TERVEN J, CÓRDOVA-ESPARZA D M, ROMERO-GONZÁLEZ J A. A Comprehensive review of YOLO architectures in computer vision; From YOLOv1 to YOLOv8 and YOLO-NAS [J]. *Machine Learning and Knowledge Extraction*, 2023, 5(4): 1680–1716. DOI: 10.3390/make5040083.
- [26] WANG C Y, YEH I H, MARK LIAO H Y. YOLOv9: Learning what you want to learn using programmable gradient information [C]//*Proceedings of the European Conference on Computer Vision*. Cham: Springer, 2024; 1–21.
- [27] WANG A, CHEN H, LIU L, et al. YOLOv10: Real-time end-to-end object detection[C]//*Proceedings of Advances in Neural Information Processing Systems*. NeurIPS, 2024; 107984 – 108011. DOI: 10.52202/079017–3429.
- [28] KHANAM R, HUSSAIN M. YOLOv11: An overview of the key architectural enhancements [J]. *arXiv preprint arXiv*, 2410.17725, 2024. DOI: 10.48550/arXiv.2410.17725.
- [29] SHI Baoguang, BAI Xiang, YAO Cong. An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2016, 39 (11): 2298–2304. DOI: 10.1109/TPAMI.2016.2646371.
- [30] WANG Jianfeng, HU Xiaolin. Gated recurrent convolution neural network for OCR [C]//*Proceedings of Advances in Neural Information Processing Systems*. NeurIPS, 2017; 335 – 344. DOI: 10.5555/3294771.3294803.
- [31] CHENG Zhanzhan, BAI Fan, XU Yunlu, et al. Focusing attention: Towards accurate text recognition in natural images [C]//*Proceedings of the IEEE International Conference on Computer Vision*. Piscataway, NJ: IEEE, 2017; 5076–5084.
- [32] LI Minghao, LV T, CHEN Jingye, et al. TROCR: Transformer-based optical character recognition with pre-trained models [C]//*Proceedings of the AAAI Conference on Artificial Intelligence*. AAAI, 2023; 13094–13102.