

# 基于深度学习的土壤学文献图形数值抽取技术框架初步构建<sup>①</sup>

刘 杰<sup>1</sup>, 马海艺<sup>1,2</sup>, 郭志英<sup>1,2</sup>, 郑梦思<sup>1</sup>, 王昌昆<sup>1,2</sup>, 潘贤章<sup>1,2\*</sup>

(1 土壤与农业可持续发展全国重点实验室(中国科学院南京土壤研究所), 南京 211135; 2 中国科学院大学, 北京 100049)

**摘要:** 针对土壤学文献中图形数值抽取效率低的问题, 本文提出了基于深度学习的图形数值抽取技术框架。首先, 对土壤学文献图形中常见要素及其表示符号进行梳理, 并收集相关图形进行标注, 形成训练数据集; 然后, 利用基于全局图像信息的 YOLO v8 模型, 通过多轮迭代优化训练适合于土壤学文献图形要素的识别模型; 最后, 研发图像坐标到数值坐标的转换算法, 实现二维散点和柱状图数值的自动提取。经独立样本检验, 所训练的模型可有效识别文献中的图形要素, 且解算的数值与手工提取结果高度吻合(线性回归决定系数  $R^2$  大于 0.99)。据此, 本文构建的基于深度学习的图形数值抽取技术框架具有较强的可行性, 为土壤学文献中图形数据进一步利用提供了一种新途径。

**关键词:** 深度学习; 土壤学文献; 图形要素识别; 数值提取; YOLO v8

中图分类号: S15; TP399 文献标志码: A

## Preliminary Construction of Technical Framework for Numerical Value Extraction from Figures in Soil Literatures Based on Deep Learning

LIU Jie<sup>1</sup>, MA Haiyi<sup>1,2</sup>, GUO Zhiying<sup>1,2</sup>, JIA Mengsi<sup>1</sup>, WANG Changkun<sup>1,2</sup>, PAN Xianzhang<sup>1,2\*</sup>

(1 State Key Laboratory of Soil and Sustainable Agriculture, Institute of Soil Science, Chinese Academy of Sciences, Nanjing 211135, China; 2 University of Chinese Academy of Sciences, Beijing 100049, China)

**Abstract:** To address the issue of low efficiency in extracting numerical values from figures, based on deep learning, a technical framework to extract numerical values from figures in soil literatures was proposed. Firstly, the common figure elements and their symbols were sorted out, and some figures were collected and manually labelled to form a training dataset. Secondly, using YOLO v8 base model, which uses the global image to detect multiple targets through one-time process, an optimized model suitable for the detection of figure elements in soil literatures was trained through several rounds of training. Thirdly, to convert the identified figure elements to real values, an algorithm was designed to automatically calculate the numerical values in 2D scatter and histogram figures. Using figures that were not involved in the training, the results showed this technique could effectively extract the figure elements and the numerical values were in high agreement with the manually extracted values (the linear regression coefficient of determination  $R^2 > 0.99$ ). Therefore, the technical framework proposed in this study has strong feasibility, which provides a new approach for the efficient use of figure data in soil literatures.

**Key words:** Deep learning; Soil research literatures; Identification of figure elements; Numerical extraction; YOLO v8

土壤学文献是土壤研究成果表达、传播, 以及相关科学知识继承和交流的主要方式。其内容主要包括文字和非文字, 文字是成果内容的主要表达形式, 而非文字类则对理解所要表达的内容至关重要。其中, 图形是主要的非文字表达方式, 是直观传达土壤数据模式和趋势的重要方式。但与表格数据相比, 图形既

不利于读者了解单个数据点的数值, 更不方便数据的后续使用, 限制了其价值的进一步发挥。为了弥补这种缺陷, 一般期刊编辑部会要求作者在论文发表的同时提供原始数据, 以供下载使用, 但这并不能解决所有问题, 因为作者提供的数据可能仅仅是少部分, 大量的数据仍仅体现在图形中。因此, 如何从图形回溯

<sup>①</sup>基金项目: 国家重点研发计划项目(2020YFC1807401)、国家科技基础资源调查专项项目(2021FY100703)和中国科学院网络安全和信息化专项应用示范项目(CAS-WX2022SF-0201)资助。

\* 通信作者(panxz@issas.ac.cn)

作者简介: 刘杰(1986—), 男, 河南新乡人, 博士研究生, 主要从事土壤大数据挖掘及应用研究。E-mail: liujie@issas.ac.cn

土壤属性等原始数值,一直是困扰读者和研究人员的问题之一。

获取图形中数值的最原始方法,是手工使用尺子在屏幕或书本上直接量算,但该方法对从大量图形中获取数据较为困难。随着技术的发展,采用相关软件,将图像校准到对应的坐标系统,直接鼠标点选相关要素获取数值,明显提升了工作效率和数据精度,有力支持了土壤学领域诸多主题的 Meta 分析研究<sup>[1-5]</sup>。但这种针对单幅图片的处理方法,在应对大量图形数据时,提取效率仍然达不到要求,如基于大量文献的 Meta 分析研究。近年来,深度学习等人工智能技术,如卷积神经网络(Convolutional neural network, CNN)、循环神经网络(Recurrent neural network, RNN),在图形数据抽取中展现出较好的效果,能够准确检测图形中的目标信息<sup>[6-7]</sup>。另外,图文识别技术(Optical character recognition, OCR)同样发展迅速<sup>[8]</sup>,将 OCR 与图像目标检测技术相结合,可以更便捷智能地获取图形中的数字信息,这种综合方法有望进一步提高图形数值抽取的自动化程度和效率。

利用深度学习等技术,国内外学者对图形数值提取方法进行了较多的探索。Al-Zaidy 等<sup>[9]</sup>提出了一种自动读取图表数据的方法,且可提供图表文本摘要。Cliché 等<sup>[10]</sup>提出了 Scatteract 框架,利用深度学习和 OCR 技术检索散点图中点的图表坐标。Jung 等<sup>[11]</sup>提出了一种交互式图表数据提取系统 ChartSense,其使用基于深度学习的分类器确定图表类型,然后利用针对不同类型图表优化的半自动交互式提取算法从图像中提取数据。Liu 等<sup>[12]</sup>提出了一种单一深度神经网络框架,由对象检测、文本识别和对象匹配模块组成,可同时处理柱状图和饼图,在模拟柱状图和饼图上的检测成功率分别达到了 79.4% 和 88.0%,但训练域以外效果较差。Daggubati 等<sup>[13]</sup>使用张量投票中的二阶张量场,建立了针对条形图中数值提取的方法流程。Zhang 等<sup>[14]</sup>提出了一种端到端的图表识别方法,其使用不同的深度框架基于不同类型的图表提取关键点,以获得图形的原始数据。Luo 等<sup>[15]</sup>提出了一种 ChartOCR 方法,从条形图、折线图、饼图等图表中提取数据,通过将深度框架与基于规则的方法相结合,可以获得令人满意的泛化能力。

综上可知,利用图像目标检测和特征提取的方法,可以检测坐标轴刻度、坐标数字和数据点等关键信息,结合 OCR,可识别数字信息。然而,由于土壤学文献中图形类型和风格多样,在没有标准模板的

情况下,图形要素识别方法缺乏通用性,因此,从图形直接提取数值的研究仍有待进一步深入。在综合分析大量技术进展的基础上,本文提出了土壤学文献图形数值抽取的新思路:首先,基于深度学习技术优化目标检测方法,形成土壤学文献图形要素识别模型;其次,利用 OCR 技术获取坐标数字信息;最后,设计坐标转换算法获取图形要素真实坐标值。另外,在当前图像目标检测算法中,YOLO(You only look once)能够一次性识别图像中的多种目标及其位置,且具有速度快、精度高的特点<sup>[16]</sup>,但不能识别土壤学文献图形中的要素目标。因此,通过专项优化 YOLO 模型,本文初步构建了基于深度学习的土壤学文献图形数值抽取技术框架。

## 1 研究方法

### 1.1 基本思路

土壤学文献有多种常见图形,如柱状图、条形图、折线图、饼图、散点图、样点分布地图等,以及多个类型组合的复杂图形开展研究,自动化提取多种图形较难实现。作为一个初步技术框架,本文主要针对散点图、柱状图等二维图形开展研究。文献图形中包括图形符号、坐标轴标签、坐标轴刻度、说明性文字、公式、图例等多种要素,点、线拐点、柱状要素顶端等关键位置及其表征的数值是研究人员最感兴趣的。因此,识别图形中符号类型及其在图形坐标系中的位置信息,对土壤文献图形的数值提取最为重要。

将图形要素关键位置看作是一个个待识别的目标,就可以用图像目标检测的方法进行自动识别。获取目标的中心位置、高度、宽度等关键信息后,再结合 OCR 识别数字信息,就可以通过一定的公式解算出图形符号的原始数值。结合 YOLO 多目标一次性识别的特点,本文优化 YOLO 基础模型训练了土壤学文献图形要素识别模型,并利用 OCR 技术识别的坐标信息来解算图形的真实数值。

### 1.2 目标检测

图像目标检测(Object detection)是计算机视觉领域的核心问题之一,目的是从图像中精确高效地识别、定位大量预定义类别的目标实例,目前深度学习的广泛应用使目标检测的精度和效率都得到了较大提升<sup>[17]</sup>。利用卷积神经网络,有研究者提出了 R-CNN(Regions with CNN features)进行图像目标检测<sup>[18]</sup>,并发展出 Faster R-CNN<sup>[19]</sup>及 SSD(Single shot multibox detector)<sup>[20]</sup>等检测方法。2016 年,Redmon 等<sup>[16]</sup>提出了多尺度特征融合的 YOLO 模型,其以良

好的速度-精度平衡，近年来在目标检测领域中的应用处于领先地位<sup>[21]</sup>，在农业、交通、安防、电力等领域也得到了广泛应用<sup>[22-25]</sup>。

YOLO v8 的基本原理是：首先，使用一个预训练的骨干网络(如 CSPDarkNet53)提取图像中的特征；其次，使用多个不同尺度的特征图来捕捉不同大小和比例的物体，再将特征图输入到一组卷积和全连接层中，预测物体的位置和类别，并进行边界框筛选；最

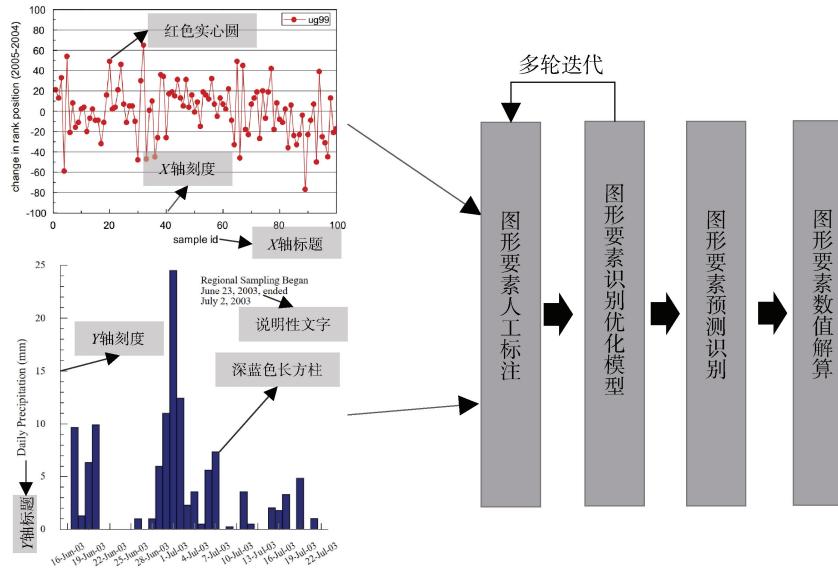


图 1 土壤学文献图形数值提取技术流程

Fig.1 Technological process of numerical value extraction from figures in soil literatures

### 1.3.1 图形要素标注

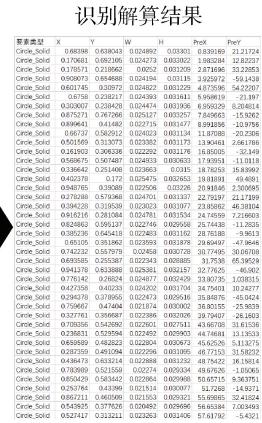
**1.3.1 图形要素标注** YOLO v8 是对 COCO (Common objects in context) 数据集进行训练的一个基础模型，可以识别人、交通工具、动物、日常用品等 80 种预训练目标，但缺乏针对文献图形要素的识别能力。因而，需要对土壤学文献散点和柱状图中的要素符号进行系统梳理，并人工标注，形成土壤学文献图形要素基础数据集。图像标注的基本目标是根据视觉内容和可获得的指导信息来确定对应的文本语义描述<sup>[26]</sup>，在综合考虑任务及标注系统便捷性等因素后，使用开源的 Label Studio 标注软件，对论文图形要素进行人工标注，详细步骤如下。

1) 制定图形要素标签体系。土壤学文献散点图、柱状图等图形中符号系统相当复杂，不易穷尽所有。作为一个初步阶段的技术验证性研究，本文梳理了土壤学文献中二维散点图和柱状图中的主要表示符号。在散点图中，符号类型主要包括圆形、三角形、正方形、菱形、长方形、十字形、五角星形、米字形及斜十字形等，前 7 项包括空心和实心类型，圆形包括半圆实心等，方形及菱形包括上下和斜对角等类型；在

后，将目标的边界框和识别概率输出，即可获得物体类别及其位置信息。因其网络模型结构较为复杂，本文不作详细阐述。

### 1.3 技术框架

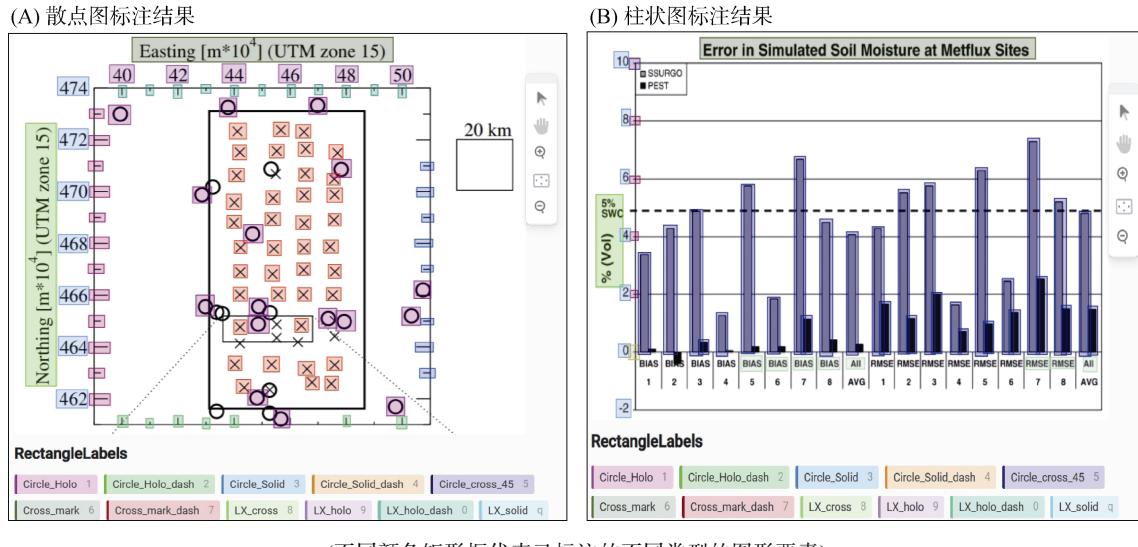
基于深度学习的土壤学文献图形数值提取技术流程大致分为 4 个部分：图形要素人工标注、YOLO 模型迭代训练、图形要素目标检测、图形要素数值解算。详细的技术流程如图 1 所示。



柱状图中,符号类型包含空心、实心、斜向上线填充、斜向下线填充、网状线填充,以及密网线和稀疏网线填充等类型。除此之外,还有一些辅助图形要素,包括X轴标签、Y轴标签,下X轴刻度线、上X轴刻度线,左Y轴刻度线、右Y轴刻度线,X和Y轴旁数字等,这些要素对于后续图形中真实坐标信息获取至关重要。目前,已经梳理出用于土壤文献图形要素的标签类型共47类。

2) 利用 Label Studio 标注软件, 自定义 47 类文献图形要素标签。基于采集的土壤学文献中的 194 幅二维图形, 将其导入 Label Studio 系统, 采用人工方式逐图进行各项要素标签的标注, 具体标注过程不在此详述, 部分图形要素标注结果如图 2 所示。

**1.3.2 模型优化训练** 不同于之前常用的基于滑动窗口的目标检测算法，YOLO v8 是基于整个图像进行的，能一次性输出所有待检测目标的信息。其对于输入的图像尺寸没有限制，但都将被调整为固定大小(比如宽度和高度都是 640 或 320)，并划分为固定大小的网格，每个网格单元负责检测落入其中的目



(不同颜色矩形框代表已标注的不同类型的图形要素)

图 2 散点图(A)和柱状图(B)人工标注结果

Fig. 2 Manually labelled scatterplots (A) and histograms (B)

标。YOLO v8 提供了一系列参数规模不等的模型，本文使用参数量最大的 YOLO v8x 作为基础模型，对其进行训练优化，详细步骤如下。

1) 划分数据集。将含有标签的 194 个图形文件分为训练集、测试集和验证集，其中，训练集图形文件 135 个，标签数 6 251 个；测试集图形文件 39 个，标签数 2 025 个；验证集图形文件 20 个，标签数 1 112 个。

2) 设置训练参数进行初始训练。训练迭代周期设为 400，批量大小设为 16，初始学习率设置为 0.01，动量因子为 0.937。

3) 模型优化。一方面通过优化图形要素标注，另一方面通过调整模型参数，最终经过多轮迭代训练形成适合于土壤学文献图形的要素识别模型。

**1.3.3 图形要素数值解算** 模型检测的点状、柱状要素信息，为其在对应图像空间的中心点坐标( $X^I$ ， $Y^I$ )、高度( $H$ )和宽度( $W$ )，下面将详述散点图和柱状图中关键要素真实坐标数值的解算过程。

1) 散点图。首先，提取识别结果中  $X$ 、 $Y$  轴旁数字信息，之后利用 OCR 技术识别对应区域的数字；其次，找出数字对应的刻度线要素在图像空间的坐标范围( $X_{\min}^I$ 、 $X_{\max}^I$  和  $Y_{\min}^I$ 、 $Y_{\max}^I$ )，及其真实坐标空间  $X$ 、 $Y$  区间  $X_{\min}$ 、 $X_{\max}$  和  $Y_{\min}$ 、 $Y_{\max}$ ；再次，建立真实坐标空间与图像坐标空间的换算公式(1)，进而获取图形中每个点的真实数值。

$$\begin{cases} X = X_{\min} + \frac{X^I - X_{\min}^I}{X_{\max}^I - X_{\min}^I} \times (X_{\max} - X_{\min}) \\ Y = Y_{\max} - \frac{Y^I - Y_{\min}^I}{Y_{\max}^I - Y_{\min}^I} \times (Y_{\max} - Y_{\min}) \end{cases} \quad (1)$$

2) 柱状图。首先，提取识别结果中  $Y$  轴旁数字信息，并利用 OCR 技术识别  $Y$  轴对应区域的数字；其次，找出数字对应的刻度线要素的图像空间坐标( $Y_{\min}^I$ 、 $Y_{\max}^I$ )，及其真实坐标空间  $Y$  区间( $Y_{\min}$ 、 $Y_{\max}$ )；再次，建立柱状要素图像空间高度与真实坐标数值的转换公式(2)，进而获取图形中每个柱状要素的真实数值。

$$Y = Y_{\min} + \frac{H}{Y_{\max}^I - Y_{\min}^I} \times (Y_{\max} - Y_{\min}) \quad (2)$$

#### 1.4 结果评价

为验证本文技术框架的数值提取结果，利用常用的图形数值提取工具 Web Plot Digitizer (<https://automeris.io/WebPlotDigitizer/>)，手工提取图形中数值，将自动解算的图形数值与手工提取结果进行线性拟合分析，并用决定系数( $R^2$ )评价两种方法获取数据的一致性。

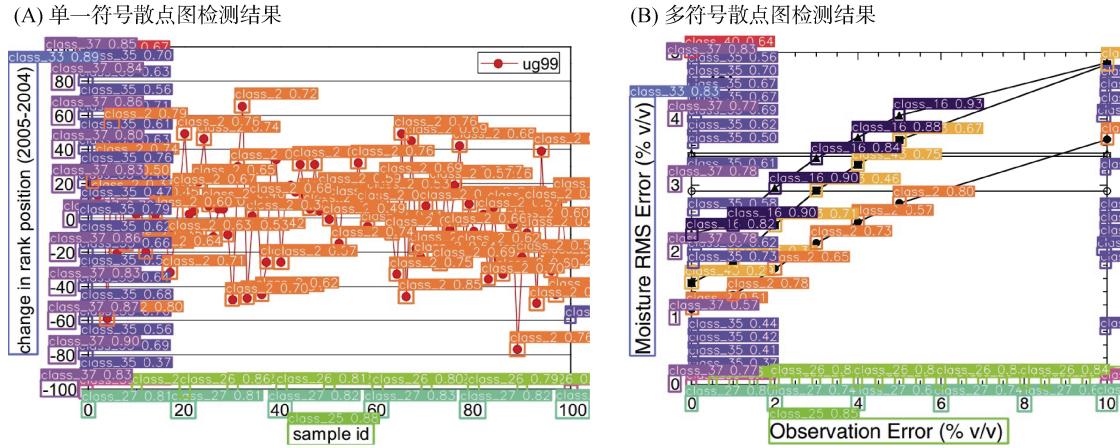
## 2 应用实例验证

### 2.1 散点图要素识别

为验证本研究所建模型对单一和多类型点状符号散点图的识别效果，分别选取了相应的图形进行预测。从图形要素检测结果(图 3)可以看出，大部分点状符号均能被本研究所建模型检出， $X/Y$  轴名称、 $X/Y$  轴数值及其对应刻度线也能被检出。在单一符号散点图中，共有 100 个点状符号，模型检出 76 个，检出率 76%；在多符号散点图中，分别有 7 个实心圆点、7 个实心方块点、6 个实心三角点，每类符号的点均被模型准确检出，检出率 100%。可见，在二

维散点图中，本研究训练的土壤学文献图形要素识别模型具有较好的点状符号识别效果。不管散点图中存在一种或多种点状符号，在点状符号清晰且分

散的图形中检测效果均较好，而当点状符号密集有遮挡时检测效果较差，类似情况在其他图像目标检测时也经常遇到<sup>[27-28]</sup>。



(图中不同颜色矩形框代表识别出的不同图形要素，框旁文字为要素类型及置信度；下同)

图 3 单一符号(A)和多类型符号(B)散点图中图形要素检测结果  
Fig. 3 Identification results of elements in scatterplots with single dot (A) and multiple dots (B)

## 2.2 柱状图要素识别

为验证本研究所建模型对单色、多色柱状图的识别效果，分别选取相应的图形进行预测。从图形要素检测结果(图 4)可以发现，大部分柱状要素均可以被本研究所建模型识别，X/Y 轴名称、Y 轴数值及其对应刻度也能被检出。在单一颜色柱状图中，共有 23 个柱状要素，模型检出 25 个，存在 2

个错误检测(错检率 8.70%)；在多颜色柱状图中，共有 12 个柱状要素，模型检出 12 个，完全检出，检测率 100%。这表明在二维柱状图中，训练的图形要素识别模型具有较好的柱状符号检测效果，在柱状要素边界清晰(如不同颜色、分散的柱状图)的图形中检测效果较好，而当柱状要素密集时检测效果相对较差。

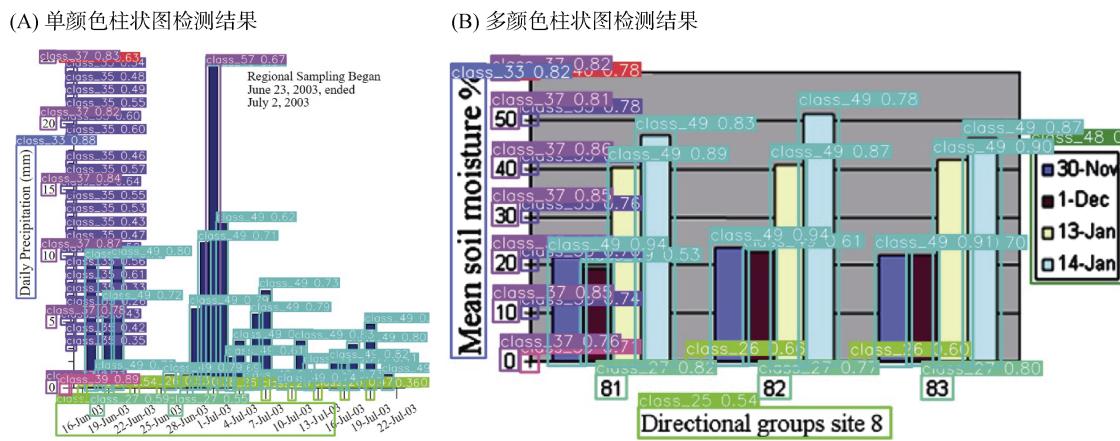


图 4 单色(A)和多色(B)柱状图中图形要素检测结果

Fig. 4 Identification of elements in histograms with single type column (A) and multiple type columns (B)

## 2.3 数值解算结果验证

为进一步验证图形要素数值解算结果的精度，利用 Web Plot Digitizer 手工提取了散点图和柱状图的数值，与本研究所建模型识别解算的结果进行了对比，线性回归分析结果(图 5)表明，不管是散点图还

是柱状图，本研究所建模型提取数值和人工提取数值间相关性极强，决定系数( $R^2$ )均大于 0.99。在散点图中，两种方法提取的数值拟合线与 1 : 1 线基本重合，而在柱状图中稍有偏离或相互平行。可见，通过对土壤学文献中检测出的图形要素进行解算，可以获取精

度较高的图形要素数值。然而,模型识别解算的柱状图要素值普遍在 1:1 线下方,说明模型识别解算的数值偏高,这可能是因为模型检测到的柱状要素高度包含外框,与真实的柱状要素相比偏高,导致解算值稍高。由此,在下一步研究中,需要针对点状要素密集、柱状图边界不清或密集等图形,通过改进要素标注策略、引入新技术、优化解算方法等途径,进一步提升土壤学文献二维图形要素检测准确率和数值解算精度。

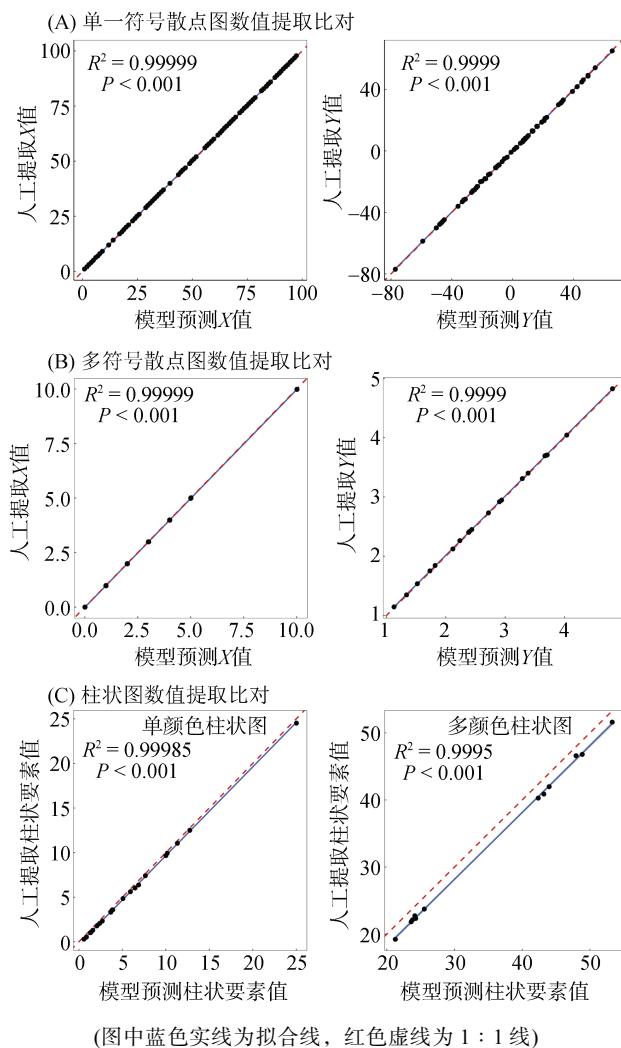


图 5 散点图(A、B)和柱状图(C)数值提取结果比对  
Fig. 5 Comparison of numerical values extracted by different methods in scatterplots (A and B) and histograms (C)

### 3 结论

本研究基于 YOLO v8 基础模型,通过手工标注常见图形要素构建训练数据集,经过多轮迭代训练形成了适合于土壤学文献的图形要素自动识别模型,研发了要素模型识别信息到图形真实数值的解算方法,初步构建了土壤学文献图形数值自动高精

度抽取的技术框架,实现了二维散点图、柱状图中数值的自动提取。对土壤学文献二维图形进行识别解算验证,结果表明,点状符号清晰且分散的散点图要素检测效果较好,且解算与手工提取的数值高度一致;柱状要素边界清晰的柱状图要素检测效果较好,虽然解算与手工提取的数值相关性极高,但存在数值高估的现象。

### 参考文献:

- [1] Beillouin D, Corbeels M, Demenois J, et al. A global meta-analysis of soil organic carbon in the Anthropocene[J]. *Nature Communications*, 2023, 14(1): 3700.
- [2] 施宇森, 王杉杉, 方伟, 等. 基于 Meta 分析研究毛竹入侵致土壤 pH 提升及养分和微生物群落结构的变化[J]. *土壤学报*, 2024, 61(3): 862–877.
- [3] 操一凡, 沈宗专, 刘珊珊, 等. Meta 分析评估中国木霉对枯萎病防控效果及其影响因素[J]. *土壤学报*, 2019, 56(3): 716–727.
- [4] 董丽, 史学正, 徐胜祥, 等. 基于 Meta 分析研究不同管理措施对中国农田土壤剖面有机碳的影响[J]. *土壤*, 2021, 53(6): 1290–1298.
- [5] 邓倩, 马娟娟, 孙西欢, 等. 生物质炭对农田土壤腐殖质的影响——Meta 分析[J]. *土壤*, 2024, 56(1): 42–48.
- [6] 卢宏涛, 张秦川. 深度卷积神经网络在计算机视觉中的应用研究综述[J]. *数据采集与处理*, 2016, 31(1): 1–17.
- [7] 郑远攀, 李广阳, 李晔. 深度学习在图像识别中的应用研究综述[J]. *计算机工程与应用*, 2019, 55(12): 20–36.
- [8] 王珂, 杨芳, 姜杉. 光学字符识别综述[J]. *计算机应用研究*, 2020, 37(S2): 22–24.
- [9] Al-Zaidy R A, Giles C L, Al-Zaidy R A, et al. A machine learning approach for semantic structuring of scientific charts in scholarly documents[C]//Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence. February 4–9, 2017, San Francisco, California, USA. ACM, 2017: 4644–4649.
- [10] Cliche M, Rosenberg D, Madeka D, et al. Scatteract: Automated extraction of data from scatter plots[M]. Machine Learning and Knowledge Discovery in Databases, Cham: Springer International Publishing, 2017: 135–150.
- [11] Jung D, Kim W, Song H, et al. ChartSense[C]//Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. May 6 – 11, 2017, Denver, Colorado, USA. ACM, 2017: 6706–6717.
- [12] Liu X, Klabjan D, NBless P. Data extraction from charts via single deep neural network[J]. 2019: p. arXiv: 1906.
- [13] Daggubati S C, Sreevalsan-Nair J, Dadhich K. BarChartAnalyzer: Data extraction and summarization of bar charts from images[J]. SN Computer Science, 2022, 3(6): 500.
- [14] Zhang L M, Wang G X, Chen L Y. ChartMaster: An end-to-end method to recognize, redraw, and redesign chart[J]. *Discrete Dynamics in Nature and Society*, 2021, 2021: 9559430.

- [15] Luo J Y, Li Z K, Wang J P, et al. ChartOCR: Data extraction from charts images via a deep hybrid framework[C]//2021 IEEE Winter Conference on Applications of Computer Vision (WACV). January 3-8, 2021, Waikoloa, HI, USA. IEEE, 2021: 1916–1924.
- [16] Redmon J, Divvala S, Girshick R, et al. You only look once: Unified, real-time object detection[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). June 27-30, 2016, Las Vegas, NV, USA. IEEE, 2016: 779–788.
- [17] 赵永强, 饶元, 董世鹏, 等. 深度学习目标检测方法综述[J]. 中国图象图形学报, 2020, 25(4): 629–654.
- [18] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//2014 IEEE Conference on Computer Vision and Pattern Recognition. June 23-28, 2014, Columbus, OH, USA. IEEE, 2014: 580–587.
- [19] Ren S Q, He K M, Girshick R B, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 39: 1137–1149.
- [20] Liu W, Anguelov D, Erhan D, et al. SSD: Single shot multibox detector[J]. Computer Vision – ECCV 2016, Cham: Springer International Publishing, 2016: 21–37.
- [21] 邵延华, 张铎, 楚红雨, 等. 基于深度学习的YOLO目标检测综述[J]. 电子与信息学报, 2022, 44(10): 3697–3708.
- [22] 高宗, 李少波, 陈济楠, 等. 基于YOLO网络的行人检测方法[J]. 计算机工程, 2018, 44(5): 215–219, 226.
- [23] 郭敬东, 陈彬, 王仁书, 等. 基于YOLO的无人机电力线路杆塔巡检图像实时检测[J]. 中国电力, 2019, 52(7): 17–23.
- [24] 王宇宁, 庞智恒, 袁德明. 基于YOLO算法的车辆实时检测[J]. 武汉理工大学学报, 2016, 38(10): 41–46.
- [25] 赵德安, 吴任迪, 刘晓洋, 等. 基于YOLO深度卷积神经网络的复杂背景下机器人采摘苹果定位[J]. 农业工程学报, 2019, 35(3): 164–173.
- [26] 卢汉清, 刘静. 基于图学习的自动图像标注[J]. 计算机学报, 2008, 31(9): 1629–1639.
- [27] 刘芳, 刘玉坤, 林森, 等. 基于改进型YOLO的复杂环境下番茄果实快速识别方法[J]. 农业机械学报, 2020, 51(6): 229–237.
- [28] 刘莫尘, 褚镇源, 崔明诗等. 基于改进yolo v8-pose的红熟期草莓识别和果柄关键点检测[J]. 农业机械学报, 2023, 54(s2 ): 244–251.