

基于深度学习的可视化仪表盘生成技术研究

陈亮, 王璐, 李生春, 刘昌宏

引用本文

陈亮, 王璐, 李生春, 刘昌宏. [基于深度学习的可视化仪表盘生成技术研究](#)[J]. 计算机科学, 2023, 50(3): 238-245.

CHEN Liang, WANG Lu, LI Shengchun, LIU Changhong. [Study on Visual Dashboard Generation Technology Based on Deep Learning](#) [J]. Computer Science, 2023, 50(3): 238-245.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[深度学习模型的后门攻击研究综述](#)

Backdoor Attack on Deep Learning Models:A Survey

计算机科学, 2023, 50(3): 333-350. <https://doi.org/10.11896/jsjcx.220600031>

[特征增强损失与前景注意力人群计数网络](#)

Crowd Counting Network Based on Feature Enhancement Loss and Foreground Attention

计算机科学, 2023, 50(3): 246-253. <https://doi.org/10.11896/jsjcx.220100219>

[跨层融合和感受野扩增的SSD目标检测算法](#)

SSD Object Detection Algorithm with Cross-layer Fusion and Receptive Field Amplification

计算机科学, 2023, 50(3): 231-237. <https://doi.org/10.11896/jsjcx.211100281>

[极化自注意力约束颜色溢出的图像自动上色](#)

Polarized Self-attention Constrains Color Overflow in Automatic Coloring of Image

计算机科学, 2023, 50(3): 208-215. <https://doi.org/10.11896/jsjcx.220100149>

[基于特征融合的边缘引导乳腺超声图像分割方法](#)

Segmentation Method of Edge-guided Breast Ultrasound Images Based on Feature Fusion

计算机科学, 2023, 50(3): 199-207. <https://doi.org/10.11896/jsjcx.211200294>

基于深度学习的可视化仪表板生成技术研究

陈亮¹ 王璐¹ 李生春² 刘昌宏²

1 西安工程大学计算机科学学院 西安 710048

2 重庆中烟工业有限责任公司黔江卷烟厂 重庆 409000

(chenliang@xpu.edu.cn)

摘要 仪表板是支持制造企业开展数据可视化分析和经营决策的重要手段。为解决可视化仪表板设计与实现过程中用户对专业知识依赖性强及流程迭代繁杂的问题,提出了一种基于深度学习技术 YOLOv5s 算法的可视化仪表板自动识别与生成方法。首先,基于 YOLOv5s 算法对仪表板图像及手绘草图中包含的可视化图表组件进行检测,并针对在检测过程中出现的手绘草图中不规则线条对识别图表的干扰及误检等问题,引入 CA 注意力机制来增强模型对重要特征的关注及目标精确定位能力,从而提高模型的识别精度;其次,将图表检测模型部署在 Web 中,服务器根据模型检测结果调用封装好的可视化图表组件代码,生成多组件组合的初始仪表板;最后,基于此 Web 设计,开发了一款数据可视化仪表板构建平台,为用户提供可修改配置仪表板样式及数据的详细选项,以方便用户快速构建完整的仪表板。通过收集 Tableau,Power BI 等可视化工具产生的仪表板图像及企业应用过程中手绘仪表板草图形成数据集,基于该数据集进行实验验证,改进的模型识别精度比原 YOLOv5s 模型提升了 2.1%,mAP 为 98.4%,并通过系统部署应用验证了图表识别方法及开发的平台可有效识别及生成相应图表组件,满足用户的基本需求。

关键词: 可视化;深度学习;目标检测;仪表板生成

中图法分类号 TP391.41

Study on Visual Dashboard Generation Technology Based on Deep Learning

CHEN Liang¹, WANG Lu¹, LI Shengchun² and LIU Changhong²

1 School of Computer Science, Xi'an Polytechnic University, Xi'an 710048, China

2 China Tobacco Chongqing Industrial Co. Ltd. Qianjiang Cigarette Factory, Chongqing 409000, China

Abstract Dashboard is an important tool to support manufacturing enterprises in data visualization analysis and business decision-making. In order to solve the problems of users' strong dependence on professional knowledge and complicated process iteration in the design and implementation of visual dashboards, a method for automatic recognition and generation of visual dashboards based on the YOLOv5s algorithm in deep learning technology is proposed. Firstly, based on the YOLOv5s algorithm, the visual chart components contained in the dashboard images and hand-drawn sketches are detected, and in order to address the problems of interference and false detection caused by irregular lines in hand-drawn sketches during the detection process, the CA attention mechanism is introduced to enhance the ability of the model to focus on important features and accurately locate the target, so as to improve the recognition accuracy of the model. Secondly, deploy the chart detection model in the web, and the server calls the encapsulated visual chart component code according to the model detection results to generate the initial dashboard of multi-component combination. Finally, a data visualization dashboard building platform is developed based on this web design, which provides users with detailed options to modify and configure the dashboard style and data, so that users can quickly build a complete dashboard. Through the collection of dashboard images generated by visualization tools such as Tableau and Power BI and hand-drawn dashboard sketches during the enterprise application process to form a dataset for experimental validation, the improved model increases the recognition accuracy by 2.1% compared to the original YOLOv5s model, and the mAP is 98.4%. The system deployment application verifies that the chart recognition method and the developed platform can effectively identify and generate the corresponding chart components to meet the basic needs of users.

Keywords Visualization, Deep learning, Object detection, Dashboard generation

到稿日期:2023-01-12 返修日期:2023-02-13

基金项目:国家自然科学基金(51675108);陕西省教育厅重点科学研究计划(22JS021)

This work was supported by the National Natural Science Foundation of China(51675108) and Key Scientific Research Program of the Education Department of Shaanxi Province, China(22JS021).

通信作者:王璐(1466756411@qq.com)

1 引言

近年来,工业大数据分析 & 数据可视化技术为处理数据分析与结果展示任务提供了丰富的手段,帮助企业管理人员从庞大而复杂的数据中发现和诊断业务问题,制定数据驱动型的企业经营决策方案。Cook 等^[1]指出可视化分析是以视觉为基本通道,通过可视化技术和交互界面,将人的知识或经验融入到数据分析和推理决策过程中,以迭代求精的方式将数据复杂度降低到用户和计算机可处理的范围,从而获取有效知识。而可视化仪表板^[2]是将多个数据视图结合起来,每个视图展示数据某一方面的属性,并允许用户通过交互手段探索数据中隐藏的规律,让企业管理人员能直观地看到数据的特征关系,为决策提供有效的信息支持。仪表板作为各项数据的汇总呈现形式,不仅在科学计算及商业智能领域中起到重要作用,同时对指导企业经营决策来说也是一种不可或缺的方法^[3]。

目前常用的数据可视化仪表板构建方法有两种:1)使用描述性设计语言表达仪表板结构,如通过调用基于 JavaScript 的开源数据可视化图表组件库(如 D3, Echarts 等),在自开发的 Web 应用中将数据转化成容易理解的图表方式;2)创建可视化应用模板,如使用商用的数据可视化分析工具(如 Tableau, Power BI, Datav 等)中提供的图表初始化模板,结合工具提供的方便灵活的自定义操作选项,定制面向最终用户的可视化仪表板。这两种方法从不同角度为仪表板的创建带来了更多的可访问性,包括减少用户所需专业知识和实现开销。但用户在设计实现过程中常遇到以下问题。一是普通用户在使用可视化图表库时,需具备一定的编程基础及开发能力,这通常需要付出较大的学习及开发成本,并且,由于缺乏可视化专业知识,普通用户在使用商用仪表板创建工具时会强烈依赖工具所提供的模板^[4],难以根据业务需求和数据特点灵活地选取合适的图表组件表达业务数据中蕴含的知识,极大地限制了数据可视化分析的应用场景的拓展。二是企业用户设计、实现仪表板的实际 workflow 是螺旋式向上循环迭代的过程,需付出巨大的时间与交流开销,开发人员需对数据分析人员给定的数据进行整理分析及探索理解,根据业务需求选择合适的图表设计进行数据展示,并需要界面设计人员提供整体布局及样式设计。但由于在数据整理与界面设计时使用的工具不同,导致在具体实现时易出现与设计意图或数据特点不符的偏差,需根据实际效果反馈进行反复修改调整至最佳方案,增加了工作量。

为了解决以上问题,Vázquez-Ingelmo 等^[5]提出利用领域通用工程和代码模板自动生成定制仪表板的方法,使用代码模板连接外部人工智能算法来推断仪表板功能,以减少对编程的依赖;Da 等^[6]提出利用机器学习算法指导用户提供反馈和解释来定制或推荐仪表板;Aksu 等^[7]采用一组关键性能指标及属性和值,通过决策模型推荐选择特定的图表作为可视化元素;Elias 等^[8]采用创建探索视图的方法,提出了一组针对新手用户的仪表板附加设计指南,使可视化新手用户能够轻松地构建商业智能仪表板。Bigelow 等^[9]提出了一款基于设计师使用的绘图工具与 D3 脚本的桥接模型,该模型允许设计师将设计工作导入到可视化生成工具中重新编辑,以解决

工作流迭代及实现偏差问题;Elshehaly 等^[10]提出了一个仪表板生成引擎 QualDash,该引擎接受一系列度量规范结构并能生成可进一步定制 GUI 元素的可视化模板,通过使用可定制的模板来简化仪表板的生成过程。Savva 等^[11]提出了一个可以自动重新设计可视化的系统 ReVision,将图表的位图图像作为输入,使用计算机视觉和机器学习技术来识别图表类型。这些方法涵盖了一些先进的仪表板设计原则、有限的代码模板及推荐算法,降低了对文本编程的依赖性。然而,用户在使用这些方法创建仪表板时仍需掌握详细的相关专业知识,并通过复杂的流程来完成,这对大多数最终用户而言难度较高。

近年来,深度学习技术的发展驱动着数据可视化流程朝着智能化方向演进^[12]。本文提出了一种基于深度学习的仪表板识别与生成方法:首先通过训练单阶段目标检测模型使其从仪表板示例图像和手绘草图中学习不同图表组件的特征,对仪表板图表组件进行类型识别及定位;其次,使用轻量级框架将仪表板识别模型部署到 Web 中,服务器根据模型识别结果调用封装好的图表组件代码,生成多图表组件组合的初始仪表板;最后,通过开发包含图表组件详细配置选项的创建可视化仪表板的系统,支持用户对生成的组件及初始仪表板进行样式及数据的自定义配置,从而构建完整的可视化仪表板。通过此方法可以简化用户创建仪表板的步骤,提高仪表板的设计与实现效率。

2 可视化仪表板生成总体方案设计

基于深度学习的仪表板识别与生成方法的总体思想是用户通过网络搜索下载或手绘符合其设计意图的仪表板图像,将其上传到平台中进行自动识别,生成图像中相应的图表组件并组合成初始仪表板;接着用户可以将生成的初始仪表板导入平台提供的配置模块中,进行图表组件的样式配置、交互设置等操作,并通过加入业务数据完成数据绑定。通过此方法,普通用户可以快捷地实现仪表板的创建及生成,减少对编程及专业知识的依赖。仪表板设计人员可高效地将设计理念快速转化成能演示的仪表板原型,并进行实际效果评估及反馈。因此,本文设计了仪表板识别与生成系统,其工作原理如图 1 所示。

图 1 中,基于深度学习的仪表板识别与生成系统主要包含仪表板图表组件识别及仪表板配置生成两部分。在仪表板识别模块中,本文通过收集计算机生成的仪表板示例图像及手绘仪表板草图制作成数据集,进行目标检测模型的训练,使其通过学习不同图表组件的特征,识别出柱状图、折线图等常见图表的类型、预测置信度分数及位置信息。接着,将训练好的仪表板识别模型部署到 Web 服务器中,并封装多类可视化图表代码形成图表组件代码库。服务器根据模型的识别结果,调用相应的组件代码生成多个图表组件,并根据其位置信息组合成可进一步交互配置的初始仪表板。

仪表板配置与生成模块中,通过开发具有详细配置选项的交互式界面,为用户提供仪表板及组件的样式配置、数据绑定等选项。用户可以将生成的初始仪表板导入到配置模块中,通过图形化点击、拖拽等操作进行仪表板组件的颜色、背景及布局等样式配置;此外,用户还可以加入业务数据,对图表组件进行数据绑定,完成可视化仪表板的数据展示及分析。



图1 系统工作原理图
Fig. 1 Diagram of system working principle

3 图表组件识别模型介绍

深度学习技术旨在对大量图像、文本等数据进行非线性变化,以学习高层次抽象特征并进行检测、推理以及识别等任务。Ma 等^[13]使用双阶段目标检测算法及机器学习验证的方法训练模型,以对仪表盘进行识别。但双阶段目标检测算法^[14]先采用区域生成机制得到候选框,再通过卷积神经网络进行特征提取及分类,需分别训练两个网络模型,结构复杂且检测速度较慢。而单阶段算法^[15]直接将整张图像送入网络中进行特征提取以预测物体的类别及位置,网络结构简单,检测速度快,泛化能力强,常被用于多种领域的目标检测。因此,本文基于单阶段算法进行仪表板的检测识别。

3.1 YOLOv5s 模型的整体结构

Redmon 等提出了单阶段目标检测模型 YOLO^[16],其核心思想是将网络输入的整张图像划分为网格单元,每个网格负责预测中心点落在该网格的目标、边界框及置信度分数。YOLOv5 是 YOLO 系列的改进,模型大小从 s, m, l, x 依次递增。其中 YOLOv5s 模型因检测速度快、网络轻、迁移能力强而被广泛运用于不同任务的实时检测,因此本文基于 YOLOv5s 模型进行可视化仪表盘图表组件的识别检测。

YOLOv5s 网络由输入端 Input、骨干网络 Backbone、核心模块 Neck、输出预测端 Head 这 4 部分组成。在输入端使用 Mosaic 数据增强将输入图像尺寸等比例缩放到网络的输入大小,通过随机缩放、裁剪以及拼接不同图像来丰富检测目标的背景,增强数据集。YOLOv5s 网络还通过自适应计算不同训练集中的最佳锚框值,提升了不同大小及不同类别的目标检测效果与网络的训练速度,降低了模型的内存需求。

骨干网络 Backbone 部分中的 Focus 结构对输入图片进行剪裁切片,以减少卷积带来的模型计算量,提升网络的速度。CSP 结构将输入的特征图经过两层卷积后与原始输入进行 add 操作,增加各层之间反向传播的梯度值,避免因网络加深而带来梯度消失的问题。SPP 空间金字塔池化结构通过

3 个不同大小的池化层进行下采样,并将池化结果与输入的特征图进行拼接,将多尺度特征转换成固定维数的特征输出,进行特征融合,扩大感受野。

Neck 模块通过特征金字塔 FPN 结构^[17]将高层特征上采样和低层特征进行自顶向下的连接,把深层的语义特征传到浅层,从而增强语义表达。路径聚合网络 PAN 结构^[18]引入自底向上的金字塔结构,将低层的定位特征信息传递到高层顶部,加强定位信息。此模块结合语义信息及定位信息,解决了多尺度特征融合的问题,极大地提升了模型的检测效果。

Head 模块将 Backbone 中提取的特征经过 Neck 模块的压缩和融合后,采用 3 个不同大小的输出进行多尺度预测。

3.2 模型改进

由于检测对象为多种可视化图表,其特征相似性较高,YOLOv5s 网络在识别仪表盘图表方面会出现将图表的一部分错误识别为单独图表的问题以及手绘草图中的不规则线条干扰识别目标的问题。因此,针对 YOLOv5s 网络对图表细粒度特征提取不充分的问题,本文引入基于空间和坐标的协同注意力机制模块 (Coordinate Attention, CA)^[19]。CA 模块通过捕获跨信道特征信息及方向感知和位置感知信息,进一步区分空间领域不同像素之间的权重关系,加强模型对重要特征的关注,弱化无关特征对识别目标造成的干扰,提高模型对目标的关注度及精确定位的能力,从而增强网络的性能。

CA 注意力机制对特征信息进行编码时分为坐标信息嵌入及注意力权重生成两个阶段。在坐标信息嵌入阶段,将 $H \times W \times C$ 大小的输入特征图,使用 $(H, 1)$ 和 $(1, W)$ 的池化核在水平方向和垂直方向分别进行全局平均池化,得到 $H \times 1 \times C$ 和 $1 \times W \times C$ 的特征图 z_h 与 z_w ,如式(1)、式(2)所示。这种方法将不同方向的感知特征图进行融合,使注意力模块获得了空间单方向的长距离关系,以及保存精确位置信息。

$$z_c^h(h) = \frac{1}{W} \sum_{0 \leq i \leq W} x_c(h, i) \tag{1}$$

$$z_c^w(w) = \frac{1}{H} \sum_{0 \leq i \leq H} x_c(i, w) \tag{2}$$

在注意力权重生成阶段,首先对经过水平和垂直方向处理生成的特征图进行级联操作,并送入 1×1 卷积模块,将其维度降为原来的 C/r ,然后使用批量归一化(BN)及非线性激活函数进行特征转化,生成 $1 \times (H+W) \times C/r$ 的中间特征图 $f = \delta(F_1([\mathbf{z}^h, \mathbf{z}^w]))$ 。接着,将生成的特征图 f 沿空间维度划分为两个独立的张量 $f^h \in R^{C/r \times H}$ 和 $f^w \in R^{C/r \times W}$ 。将两个张量送入两个 1×1 的卷积模块转换成与输入相同的通道数,

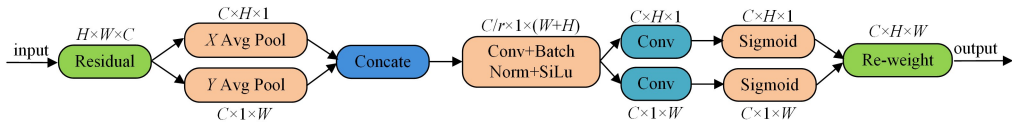


图2 CA注意力模块的结构

Fig. 2 CA attention module structure

本文将 CA 模块应用于 YOLOv5s 网络的骨干网络与 Neck 模块中间,并将其命名为 CA-YOLOv5s。将图像输入到模型后,首先通过骨干网络进行多维特征提取,将提取的特征图输入到 CA 注意力模块中,通过精确定位提取对目标区域更有判别力的特征表示,来增强网络对各类型图表特征的学习能力;然后将提取的特征输入到 Neck 模块中加强网络的语义特征和位置信息的提取,并进行多维特征融合;最后通过损失函数进行计算预测,输出结果。改进后的模型的整体结构如图 3 所示,各模块的结构如图 4 所示。

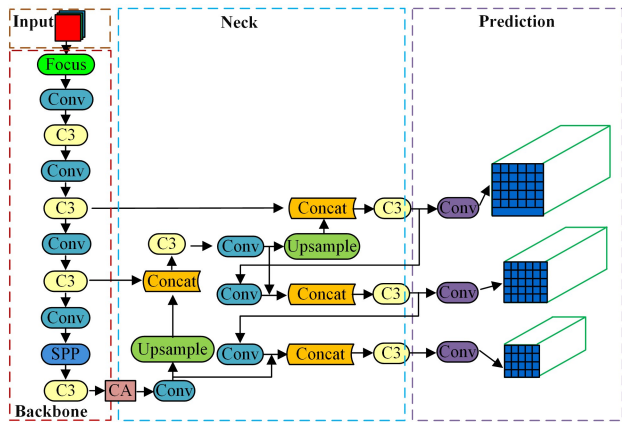


图3 CA-YOLOv5s模型的整体结构

Fig. 3 Overall structure of CA-YOLOv5s model

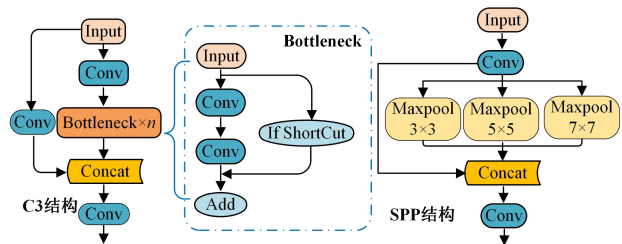


图4 CA-YOLOv5s各模块的结构

Fig. 4 Structure of each module of CA-YOLOv5s

3.3 损失函数

YOLOv5s 模型在 Head 部分采用了 3 种损失函数来计算分类损失、定位损失和置信度损失。训练阶段的目标定位损失采用 CIoU Loss 计算,CIoU 将预测框与真实框的重叠面积、中心点距离、纵横比同时加入计算,通过更多的维度来

并经过 Sigmoid 激活函数后,分别获得在垂直方向上的注意力权重 g^h 和水平方向上的注意力权重 g^w ,如式(3)、式(4)所示:

$$g^h = \sigma(F_h(f^h)) \quad (3)$$

$$g^w = \sigma(F_w(f^w)) \quad (4)$$

最后将两个权重通过乘法加权计算,输出最终的注意力权重特征图,如式(5)所示,CA注意力模块的结构如图2所示。

$$y_c(i, j) = x_c(i, j) \times g_c^h(i) \times g_c^w(j) \quad (5)$$

考虑它们之间的差异,进一步提升了训练的稳定性 and 收敛速度。计算式如式(6)~式(8)所示:

$$CIoU(B, B_{gt}) = IoU(B, B_{gt}) - \frac{(\rho^2(B, B_{gt}))}{c^2} - \alpha v \quad (6)$$

$$v = \frac{4}{\pi^2} \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2 \quad (7)$$

$$\alpha = \frac{v}{1 - IoU(B, B_{gt}) + v} \quad (8)$$

其中, ρ 为预测框 B 和真实框 B_{gt} 的中心点距离; c 为预测框和真实框的最小包围矩形的对角线长度; v 为其宽高比相似度,用于将预测框与真实框宽高比差值进行归一化; α 为 v 的权衡长宽比例的影响因子。目标定位损失可以表示为 $L_{loc} = L_{CIoU} = 1 - CIoU$ 。

本文的目标置信度损失及分类损失采用归一化交叉熵损失函数 BCEWithLogitsLoss^[20] 代替原本的交叉熵损失函数 BCELoss 进行多标签二分类计算。BCELoss 用于创建衡量目标和输出之间的二进制交叉熵的标准,在计算时需先对输入做 Sigmoid 处理,再进行归一化计算,增加了计算量与时间。而 BCEWithLogitsLoss 将两者合并到一个层中,函数内部先对输出的置信度进行 Sigmoid 操作,再进行 BCELoss 计算,省略了原损失函数的归一化步骤,在几乎不影响准确率的情况下提升了检测速度和数值稳定性。使用 BCEWithLogitsLoss 计算的目标置信度损失 L_{conf} 包括正样本目标置信度损失 L_{obj} 和负样本目标置信度损失 L_{noobj} , 可表示为 $L_{conf} = L_{obj} + L_{noobj}$, 其中 L_{obj} 与 L_{noobj} 的计算式如下:

$$C = \frac{1}{1 + e^{-x}} \quad (9)$$

$$L_{obj} = - \sum_{i=0}^S \sum_{j=0}^B I_{ij}^{obj} [C_i^j \log(C_j^j) + (1 - C_i^j) \log(1 - C_j^j)] \quad (10)$$

$$L_{noobj} = - \lambda_{noobj} \sum_{i=0}^S \sum_{j=0}^B I_{ij}^{noobj} [C_i^j \log(C_j^j) + (1 - C_i^j) \log(1 - C_j^j)] \quad (11)$$

式(9)为 sigmoid 激活函数,式(10)和式(11)中, S 代表 YOLOv5s 检测头部尺寸值的网格数, B 表示先验框的数量, I_{ij}^{obj} 和 I_{ij}^{noobj} 为判断第 i 个网格的第 j 个先验框是否有预测目标,若含有预测目标,则 $I_{ij}^{obj} = 1, I_{ij}^{noobj} = 0$; 否则, $I_{ij}^{obj} = 0, I_{ij}^{noobj} = 1$ 。 C_i^j 和 C_j^j 分别表示经过 Sigmoid 激活函数输出的置信度的预测值和真实值, λ_{noobj} 是用来平衡正负样本的常数系数。类别预测损失只计算正样本的分类损失,如式(12)所示:

$$L_{cls} = - \sum_{i=0}^{s^2} \sum_{j=0}^B I_{ij}^{obj} \sum_{c \in classes} [p_i^j(c) \log(p_j^i(c)) + (1 - p_i^j(c)) \log(1 - p_j^i(c))] \quad (12)$$

4 实验设计与结果分析

为了评估改进的 CA-YOLOv5s 算法的优势,本文通过收集计算机生成的仪表板图像及手绘仪表板草图构成数据集,训练 CA-YOLOv5s 模型及原 YOLOv5s 模型进行比较验证,并引入 CenterNet^[21],SSD^[22] 及 RetinaNet^[23] 模型在自构建数据集上进行对比实验分析。最后,将本文改进的 CA-YOLOv5s 方法的结果与文献[13]使用的双阶段 Faster R-CNN^[24] 方法的结果进行比较,以验证 CA-YOLOv5s 模型识别仪表板图表组件的有效性。

4.1 实验环境及数据集

本文实验基于 Windows10 操作系统,使用 Pytorch 1.12.0 版本的框架、NVIDIA GeForce RTX2070 显卡进行运算,GPU 加速计算方面采用 CUDA12.0 及 CUDNN8.3.2。

实验共采集了 3600 张仪表板图像样本,包括计算机生成的仪表板图像及手绘仪表板草图。计算机生成的图像来自 Tableau,DataV,Power BI 等平台的模板案例下载以及互联网抓取。手绘草图是在企业应用过程中收集的不同设计人员以及实验人员手绘的图像。因仪表板中常用图表组件的特征较为相似,本文实验将数据集划分为仪表板单图表组件图像以及仪表板混合图像两类。其中单图表组件包含七大类常用的可视化图表,分别为柱状图 Bar、条形图 HBar、折线图 Line、柱状折线图 BarLine、饼状图 Pie、环形图 Doughnut、散点图 Scatter。手绘草图的识别受不规则线条影响较大,收集的每类手绘草图的数量比计算机生成图像多 50 张,以提高手绘草图的识别精度。

数据集采用 PASCAL VOC 格式,使用 LabelImg 工具对图像进行手工标注,标注的文件为 xml 格式,然后使用脚本文件将数据集标签转化为 YOLO 的文本格式,记录图表的中心坐标、宽高以及类型,表示为 (x,y,w,h,t) ,并进行归一化操作。最后将数据集按 8:2 的比例划分为训练集和测试集,详细信息如表 1 所列。

表 1 仪表板数据集分类表

Table 1 Dashboard dataset classification

类型	描述	计算机生成 图像数目	手绘图像 数目
Bar	柱状图	200	250
HBar	条形图	200	250
Line	折线图	200	250
BarLine	柱状折线图	200	250
Doughnut	环形图	200	250
Pie	饼状图	200	250
Scatter	散点图	200	250
标准仪表板	混合图表	200	250

4.2 评价指标

为更好地说明本文模型的检测性能,本实验使用均值平均精度 mAP@0.5(IoU 阈值为 0.5 的平均精度值)、精确度 Precision、召回率 Recall 作为评价指标,计算式如下:

$$Precision = \frac{TP}{TP + FP} \quad (13)$$

$$Recall = \frac{TP}{TP + FN} \quad (14)$$

$$AP = \int_0^1 p(R) dR \quad (15)$$

$$mAP = \frac{1}{n} \sum_j^{class\ number} Class_AP_j \quad (16)$$

其中,TP,FP,FN 分别表示正确检测出的图表组件数目、错误检测出的图表组件数目,以及未被正确检测出的图表组件数目。精确度 Precision 为模型预测所有目标时,预测类别正确的比例。召回率 Recall 为模型预测真实目标正确的比例。AP 为单类别平均精度,是精确度与召回率的结合,计算方法为 PR 曲线下面积,这是判断该类别与分类器表现性能的标准。mAP 为各类别 AP 的平均值,用于对所有目标类别检测的效果取平均值,可以代表整个检测模型的表现性能。另外,本实验引入模型计算量 FLOPs,即浮点运算次数,以及每秒识别帧数 FPS 来评估改进模型的性能。

4.3 模型训练及结果分析

本文实验将模型初始学习率设为 0.01,批处理大小为 8,训练周期为 100 轮。将模型训练结果可视化,图 5 给出了 CA-YOLOv5s 模型训练 100 轮次后训练集与验证集中损失函数的收敛情况及各指标的变化情况。

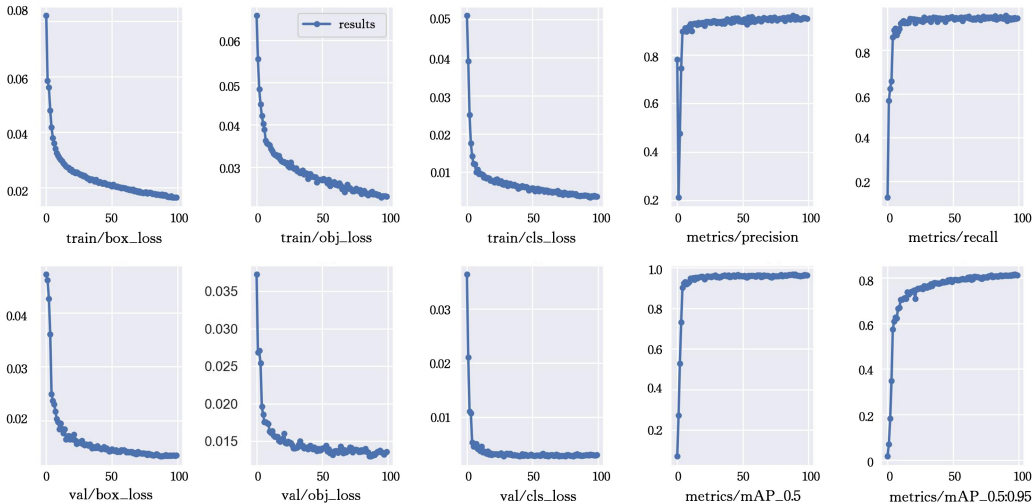


图 5 损失函数收敛情况及指标变化情况
Fig. 5 Loss function convergence and indicator changes

由图 5 可知,模型训练 50epoch 时,损失函数及各指标值都趋于稳定,训练 100epoch 后,都达到了较好的效果。其中边界框回归损失 box_loss、置信度损失 obj_loss 以及类别损失 cls_loss 的误差均低于 0.03,在对应验证集 val 损失中,3 个损失函数的误差均低于 0.02,模型稳定收敛。精确度 Precision 与召回率 Recall 在 50epoch 之后无较大波动,mAP@0.5 值最终趋于 1,说明模型训练效果较好。

为评估改进的 CA-YOLOv5s 模型在可视化仪表板识别方面的效果,本文在自构建数据集中训练原 YOLOv5s 模型,并采用 mAP、模型计算量 GFLOPs 与每秒处理帧数 FPS 这 3 个指标综合对比评估改进模型的整体性能,结果如表 2 所列。其次,将 CenterNet,RetinaNet,SSD,以及 YOLOv5s 与 CA-YOLOv5s 模型的各类图表的 AP 值进行详细对比,以评估本文算法的性能优势,结果如表 3 所列。

表 3 其他目标检测算法与本文方法的对比
Table 3 Comparison of other object detection algorithms with the proposed method

模型	(单位:%)						
	柱状图	折线图	条形图	饼状图	环形图	散点图	mAP
CenterNet	85.0	94.0	95.8	97.0	93.3	94.3	93.4
RetinaNet	86.7	93.2	96.5	97.5	92.6	94.0	93.6
SSD	84.5	93.7	96.2	97.2	92.7	93.4	93.2
YOLOv5s	85.5	98.0	99.2	98.2	98.0	97.8	96.3
CA-YOLOv5s	94.2	98.8	99.6	99.6	98.7	98.0	98.4

由表 3 可知,在当前实验环境下,CA-YOLOv5s 方法得到的 7 种图表组件的平均识别精度 mAP 从原 YOLOv5s 模型的 96.3% 提升到了 98.4%,提高了 2.1%。其中,柱状图的 AP 值有显著提升,从 85.5% 提高到 94.2%;条形图 HBar、柱状折线图 Barline 和饼图 Pie 的 AP 值均达到 99.6%,其他各类图表组件的识别精度 AP 值也有提高。上述结果证明了引入 CA 注意力模块能够增强模型对重要特征的关注,能更好地提取特征,有效提高检测精度。并且相比 CenterNet,RetinaNet,SSD 方法,CA-YOLOv5s 模型的检测效果更好,算法性能更优。

文献[13]使用 Faster R-CNN 算法训练两个模型,分别对计算机生成的图表组件及手绘图表组件进行候选检测。本文将 CA-YOLOv5s 结果与其对应的图表组件识别精度及仪表板的 mAP 值进行对比分析,对比结果如表 4 所列。

表 4 本文方法与文献[13]的 AP 值的对比
Table 4 AP comparison of the proposed method with literature [13]
(单位:%)

	柱状图	折线图	条形图	仪表板 mAP	手绘 mAP	总 mAP
本文	94.2	98.8	99.6	—	—	98.4
文献[13]	96.6	96.4	92.4	77.9	85.0	—

由表 4 可知,本文提出的 CA-YOLOv5s 方法对折线图和条形图的识别精度分别为 98.8% 和 99.6%,均高于文献[13]的 Faster R-CNN 方法;柱状图的识别精度比其低 2.4%。Faster R-CNN 对计算机生成的标准仪表板识别精度为 77.9%,手绘仪表板草图的识别精度为 85%。本文基于 YOLOv5s 的方法仅训练一个模型对计算机生成的仪表板及手绘仪表板进行识别,最终 mAP 值为 98.4%。

表 2 改进模型的性能比较
Table 2 Performance comparison of improved models

	mAP/%	FLOPs/G	FPS/(f/s)
YOLOv5s	96.3	15.8	43.7
YOLOv5s+CA	98.0	15.8	41.0
YOLOv5s+BCEWithLogitsLoss	96.5	15.7	46.0
YOLOv5s+CA+BCEWithLogitsLoss	98.4	15.8	43.3

由表 2 可知,与原 YOLOv5s 网络相比,加入 CA 模块的网络 mAP 值提升了 1.7%,但网络检测速度下降了 2.7 f/s。将 BCELoss 损失函数替换为 BCEWithLogitsLoss 时,检测速度有一定的改善。最后将 CA 模块与改进的损失函数结合,在 FLOPs 相同和 FPS 略有减少的情况下检测精度值提升了 2.1%,这说明引入 CA 与改进 BCELoss 损失函数的模型整体性能较优。

表 3 其他目标检测算法与本文方法的对比
Table 3 Comparison of other object detection algorithms with the proposed method
(单位:%)

模型	柱状图	折线图	条形图	饼状图	环形图	散点图	柱状折线图	mAP
CenterNet	85.0	94.0	95.8	97.0	93.3	94.3	94.6	93.4
RetinaNet	86.7	93.2	96.5	97.5	92.6	94.0	94.8	93.6
SSD	84.5	93.7	96.2	97.2	92.7	93.4	94.5	93.2
YOLOv5s	85.5	98.0	99.2	98.2	98.0	97.8	97.2	96.3
CA-YOLOv5s	94.2	98.8	99.6	99.6	98.7	98.0	99.6	98.4

5 系统应用与评估

5.1 系统部署

本文基于 B/S 架构和 RESTful 微服务程序设计风格,开发了一款可视化仪表板创建平台,首先选用轻量级 Flask 框架搭建 Web 服务,将训练好的仪表板图表组件识别模型部署到其中,获取用户上传的图像并进行存储与识别,基于 Echarts 及 AntV 封装多类可视化图表代码形成组件库。用户在仪表板识别界面上传仪表板模板图像或手绘草图后,Web 服务器通过 API 接口调用挂载的仪表板图表识别引擎来识别图像中的图表组件的类型及布局信息,并通过绑定及响应 post 消息预测函数来进行预测。最后将预测结果通过 Web 服务返回,服务器根据返回的结果,选择预测置信度分数大于 0.7 的图表类型,从图表组件代码库中调用封装好的组件代码,渲染生成对应的图表组件,并根据其位置信息组合成初始仪表板。

接着,平台采用组件化和拖拽式思想构建面向可视化的工程,通过封装文本、表格等功能类组件及包含详细配置选项的交互式配置界面,允许用户在缺乏编程知识的情况下,将通过仪表板识别模块识别生成的初始仪表板导入到组件配置模块中,对各组件进行颜色样式配置及数据绑定等。配置界面由 4 部分组成:功能面板、组件面板、画布和配置面板。其中左侧组件面板存放着封装好的柱状图、饼图等数图组件及文本等工具组件,用户可以通过简单地拖拽组件的操作来修改仪表板中的组件。点击组件时,右侧配置面板显示当前选中组件所支持的样式及数据配置的选项。用户可以将 CSV、数据库等多种类型的数据源绑定到各图表组件中来完成数据的动态展示及分析。

5.2 应用案例

在某制造企业生产加工的过程中,不同加工设备及不同阶段的工序会对产品质量产生影响,为方便管理人员高效地了解车间生产数据的规律和产品质量异常问题,企业引入数据可视化与分析技术,使用可视化仪表板对车间在制品的合格率及物料异常等性能进行监控,发现产品缺陷以及提升产品质量,实现数据可视化驱动质量管理。本文将部署开发的系统应用于制造企业的数据可视化过程中。图 6(a)一图 6(d)给出了一位设计师使用本文开发的工具对于质量监控的仪表板原型设计进行评估的整体流程。图 6(a)设计师手绘的仪表板草图,设计师将其上传到系统中,系统调用图表组件识别

引擎进行检测识别。图 6(b)给出了模型识别结果,包含了 7 个图表的置信度分数及图表类型,详细内容为:“Line 0.84”“HBar 0.74”“Bar 0.82”“Pie1 0.83”“Pie2 0.87”“Doughnut 0.87”“BarLine 0.90”。图 6(c)给出了服务器根据结果调用封装好的图表组件代码生成的相应图表组件。图 6(d)是将其导入到配置模块中,进行后续的样式定制和数据绑定,以生成最终仪表板的界面。结果表明本文方法在不规则线条影响下可以有效地对手绘仪表板草图进行识别与检测,并根据封装的图表组件代码库生成仪表板。因此,设计师可以通过此方法实现快速构建、配置及生成仪表板,高效地将概念化设计转换为可演示的仪表板,简化工作流程,减少实现开销。

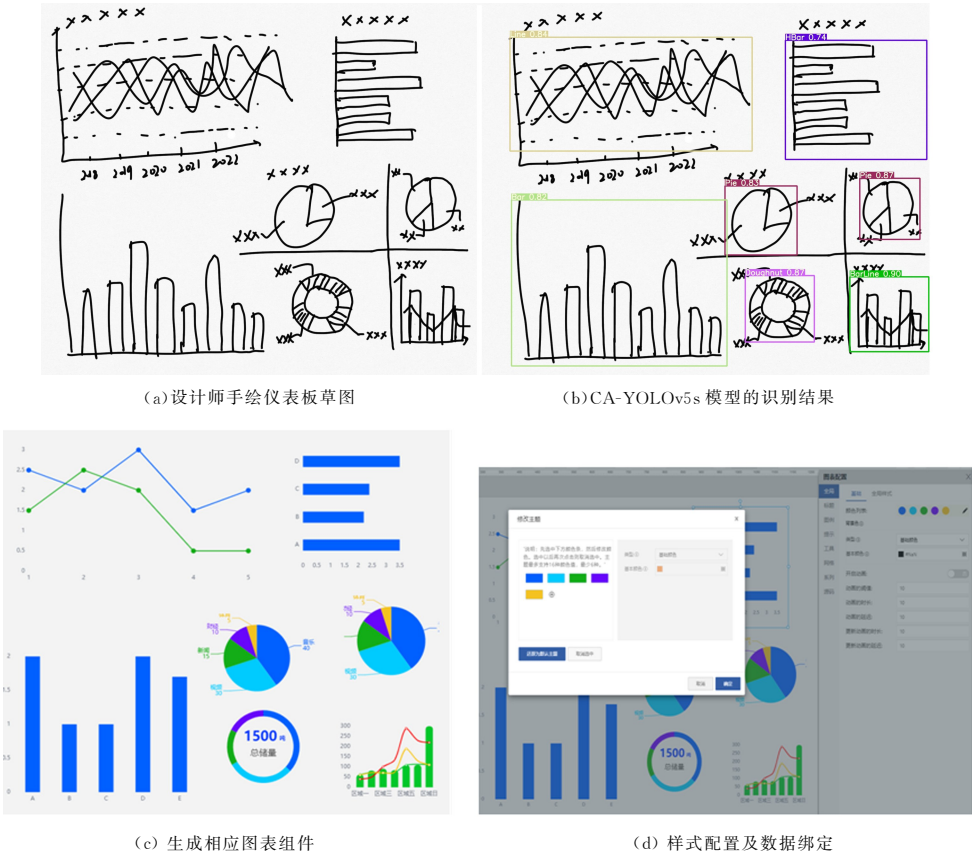


图 6 使用可视化工具识别创建仪表板的整体流程

Fig. 6 Overall process of identifying and creating dashboards using visualisation tools

将 Tableau,Power BI 等可视化平台中下载的标准仪表板模板图像及实验人员手绘的仪表板设计草图组成数据集,训练基于协同注意力机制的 CA-YOLOv5s 模型来识别仪表板图表,并通过实验及企业应用案例证明了本文方法可有效地检测识别仪表板。但在实际应用中出现了数据集来源及可视化场景单一、可检测识别的仪表板图表组件类型有限的问题。因此,本文开发的 Web 应用将用户上传的仪表板模板图像或手绘草图存储到服务器中,以扩充模型数据集。后期定时重新训练模型,使其学习到更多应用场景中不同类型的仪表板图表组件特征,以及不同的手绘与设计风格,提高识别模型的泛化能力及鲁棒性。

结束语 本文基于深度学习技术中的单阶段目标检测算法 YOLOv5s,识别用户上传的可视化仪表板图像中的图表组件的类型及位置信息,并引入协同注意力机制模块,解决了原

引擎进行检测识别。图 6(b)给出了模型识别结果,包含了 7 个图表的置信度分数及图表类型,详细内容为:“Line 0.84”“HBar 0.74”“Bar 0.82”“Pie1 0.83”“Pie2 0.87”“Doughnut 0.87”“BarLine 0.90”。图 6(c)给出了服务器根据结果调用封装好的图表组件代码生成的相应图表组件。图 6(d)是将其导入到配置模块中,进行后续的样式定制和数据绑定,以生成最终仪表板的界面。结果表明本文方法在不规则线条影响下可以有效地对手绘仪表板草图进行识别与检测,并根据封装的图表组件代码库生成仪表板。因此,设计师可以通过此方法实现快速构建、配置及生成仪表板,高效地将概念化设计转换为可演示的仪表板,简化工作流程,减少实现开销。

图表数据及其他文本内容进行识别。未来将通过模型改进提高识别范围,实现对仪表板更全面的识别及生成,进一步简化仪表板的创建流程。

参考文献

[1] COOK K A, THOMAS J J. Illuminating the path: The research and development agenda for visual analytics[R]. Richland, WA (United States): Pacific Northwest National Lab, 2005.

[2] SARIKAYA A, CORRELL M, BARTRAM L, et al. What do we talk about when we talk about dashboards? [J]. IEEE Transactions on Visualization and Computer Graphics, 2018, 25(1): 682-692.

[3] KEIM E D, KOHLHAMMER J, ELLIS G. Mastering the Information Age: Solving Problems with Visual Analytics[M]// Golsar: Eurographics Association, 2010.

[4] GRAMMEL L, TORY M, STOREY M A. How information visualization novices construct visualizations [J]. IEEE Transactions on Visualization and Computer Graphics, 2010, 16(6): 943-952.

[5] VÁZQUEZ-INGELMO A, GARCÍA-PEÑALVO F J, THERÓN R, et al. Connecting domain-specific features to source code: Towards the automatization of dashboard generation [J]. Cluster Computing, 2020, 23(3): 1803-1816.

[6] DA COL S, CIUCANU R, SOARE M, et al. DashBot: An ML-Guided Dashboard Generation System[C]// Proceedings of the 30th ACM International Conference on Information & Knowledge Management. New York: Association for Computing Machinery, 2021: 4696-4700.

[7] AKSU Ü, DEL-RÍO-ORTEGA A, RESINAS M, et al. An approach for the automated generation of engaging dashboards [C]// Proceedings of OTM Confederated International Conferences on the Move to Meaningful Internet Systems. Cham: Springer, 2019: 363-384.

[8] ELIAS M, BEZERIANOS A. Exploration views: understanding dashboard creation and customization for visualization novices [C]// Proceedings of IFIP Conference on Human-Computer Interaction. Berlin: Springer, 2011: 274-291.

[9] BIGELOW A, DRUCKER S, FISHER D, et al. Iterating between tools to create and edit visualizations [J]. IEEE Transactions on Visualization and Computer Graphics, 2016, 23(1): 481-490.

[10] ELSHEHALY M, RANDELL R, BREHMER M, et al. QualDash: Adaptable generation of visualisation dashboards for healthcare quality improvement [J]. IEEE Transactions on Visualization and Computer Graphics, 2020, 27(2): 689-699.

[11] SAVVA M, KONG N, CHHAJTA A, et al. Revision: Automated classification, analysis and redesign of chart images[C]// Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology. New York, United States: Association for Computing Machinery, 2011: 393-402.

[12] LIU C, LAI C F, JIANG R K, et al. Visualization driven by deep learning [J]. Journal of Computer Aided Design and Graphics, 2020, 32(10): 1537-1548.

[13] MA R, MEI H, GUAN H, et al. Ladv: Deep learning assisted authoring of dashboard visualizations from images and sketches [J]. IEEE Transactions on Visualization and Computer Gra-

phics, 2020, 27(9): 3717-3732.

[14] AZIZ L, SALAM M S B H, SHEIKH U U, et al. Exploring deep learning-based architecture, strategies, applications and current trends in generic object detection: A comprehensive review [J]. IEEE Access, 2020, 8: 170461-170495.

[15] JIAO L, ZHANG F, LIU F, et al. A survey of deep learning-based object detection [J]. IEEE Access, 2019, 7: 128837-128868.

[16] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2016: 779-788.

[17] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2017: 2117-2125.

[18] LIU S, QI L, QIN H, et al. Path aggregation network for instance segmentation[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2018: 8759-8768.

[19] HOU Q, ZHOU D, FENG J. Coordinate attention for efficient mobile network design[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2021: 13713-13722.

[20] REZATOFIGHI H, TSOI N, GWAK J Y, et al. Generalized intersection over union: A metric and a loss for bounding box regression[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2019: 658-666.

[21] ZHOU X, WANG D, KRÄHENBÜHL P. Objects as points [J]. arXiv:1904.07850, 2019.

[22] LIU W, ANGUELOV D, ERHAN D, et al. Ssd: Single shot multibox detector[C]// Proceedings of European Conference on Computer Vision. Cham: Springer, 2016: 21-37.

[23] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]// Proceedings of the IEEE International Conference on Computer Vision. Los Alamitos: IEEE Computer Society Press, 2017: 2980-2988.

[24] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 39(6): 1137-1149.

CHEN Liang, born in 1977, Ph.D, associate professor, master's supervisor, is a member of China Computer Federation. His main research interests include industrial big data, big data analysis and visualization.

WANG Lu, born in 1996, postgraduate. Her main research interests include data visualization and image identification.

(责任编辑:何杨)