

改进 YOLOX-m 的火焰检测方法研究

战乃岩, 张晓禾, 姜泽旭, 于傲芝

(吉林建筑大学 应急科学与工程学院, 长春 130119)

摘要: 如今火灾自动报警技术已逐步朝着智能化、网络化和自动化的方向发展, 然而, 目前的火焰实时检测技术存在火焰实时检测精度低和网络计算参数量大等问题; 针对以上问题, 对 YOLOX-m 目标检测模型进行研究, 提出改进 YOLOX-m 的轻量级火焰检测模型; 通过将主干网络 CSPDarknet-53 替换为 ShuffleNetV2, 在降低计算量的同时提高网络精度, 在 ShuffleNetV2 结构中插入 RFB 模块扩大感受野, 在保持分辨率和精确定位检测目标的同时提升检测大目标的能力, 将 Neck 部分的上采样替换为 Pixel Shuffle 以降低特征损失, 为使网络能够关注到更关键的信息, 增加注意力机制 CBAM, 从而提高模型整体性能; 经过算法优化和实验测试, 改进模型比 YOLOX-m 模型精度提高了 2.87 个百分点, 参数量减少 37.9%, 计算量降低 30.7%; 改进模型成功应用于森林火灾、城市火灾等实际场景, 通过对比可以相对更加精确地检测火焰。

关键词: 火焰检测; YOLOX; 通道混洗; 感受野增强; 注意力机制

Research on Flame Detection Method Based on Improved YOLOX-m

ZHAN Naiyan, ZHANG Xiaohe, JIANG Zexu, YU Jingzhi

(School of Emergency Science and Engineering, Jilin Jianzhu University, Changchun 130119, China)

Abstract: Nowadays, the automatic fire alarm technology is gradually developing towards intelligence, networking, and automation. However, current real-time flame detection techniques have the characteristics of low detection accuracy and high computational parameters. To address these issues, this paper focuses on the YOLOX-m object detection model, and proposes an improved lightweight fire detection model. The proposed model improves the YOLOX-m model by replacing the backbone network CSPDarknet-53 with the ShuffleNetV2, which reduces computational complexity while improving network accuracy. Additionally, the RFB module is inserted into the ShuffleNetV2 structure to increase the receptive field, enhancing the detection capability for large objects while maintaining the resolution and precise localization. The upsampling in the Neck is replaced with the Pixel Shuffle to minimize the feature loss. Furthermore, the attention mechanism CBAM is incorporated to enable the network to focus on the crucial information, thus improving the overall performance of the model. Through experimental testing of the optimized algorithm, the improved model increases the accuracy by 2.87% compared to the YOLOX-m model, with a reduction of 37.9% in parameters and a decrease of 30.7% in computational complexity. The improved model is successfully applied in real scenarios such as forest and urban fires, which can detect flames more accurately.

Keywords: flame detection; YOLOX; channel shuffle; enhanced receptive field; attention mechanism

0 引言

火灾是一种由气候变化或人类活动引发的具有极高危害性的自然灾害, 可能造成大量生态、人员和财产损失。“十四五”国家消防工作规划中提出了一系列措施, 重点强调了信息化手段运用的提升, 旨在加强火灾防控工作的效能, 包括远程监控、物联网监测、电气监控等方面, 加强消防监督检查和火灾调查装备的轻型化与集

成化。通过这些措施, 消防工作将更加高效和智能化^[1]。利用现代科技手段和智能设备, 对火灾进行实时监测、预警和管理的一种方式。通过将传感器、摄像头、智能化控制系统等设备与互联网、云计算等技术结合, 实现对火灾的及时发现、精确定位和迅速响应。由此可见, 政府逐渐重视火灾检测的智能化, 且随着火焰识别等技术的出现, 火灾检测的自动化水平有望进一步提高, 对于提高火灾安全性能、加速反应速度、提高可

收稿日期: 2023-12-13; 修回日期: 2024-02-02。

作者简介: 战乃岩(1975-), 女, 博士, 教授。

通讯作者: 张晓禾(1999-), 女, 硕士研究生。

引用格式: 战乃岩, 张晓禾, 姜泽旭, 等. 改进 YOLOX-m 的火焰检测方法研究[J]. 计算机测量与控制, 2025, 33(3): 20-29.

靠性、降低人力成本以及方便管理和维护都具有重要的意义,帮助人们更早地发现和及时处理火灾事件,减少火灾带来的伤害和损失。

20世纪90年代开始,随着计算机图像处理技术的发展,目标检测算法兴起,利用监控视频流进行图像分析以获取火情信息的视觉检测方法已经逐渐成为火灾探测的重要手段。主要基于计算机视觉和图像处理技术,依靠手工设计的特征提取器和分类器来实现目标检测。这一时期的火焰检测算法虽然实现了远程监控火灾情况的目标,但需要人工依据日常经验提取颜色、纹理等特征^[2]。因此,随着科技的进步,新兴的智能火灾检测技术如深度学习、计算机视觉等正逐渐应用于火灾检测领域,具有更高的准确性和灵敏度。

自2012年深层卷积网络 AlexNet 横空出世,打开了深度学习在目标检测领域的大门^[3]。深度学习不需要人工提取特征值,只需设计卷积网络,读取视频流,即可检测相应目标,降低了人工选取特征的工作量及其带来的主观性和不确定性。然而,这些方法仍然会消耗大量的计算资源,并且模型推理时间较长。2015年,文献[4]中首次提出的单阶段算法YOLO^[4]是目标检测技术的一个重大突破,其特点在于不需要分别进行预选框的生成和框内物体的识别,检测速度明显高于此前的双阶段检测算法。如今研究人员在单阶段检测算法的基础上进行改进和优化,实现更加高精度、高速率地检测火焰和烟雾。2020年,文献[5]的研究者基于双阶段网络Faster R-CNN(The Faster Region-based Convolution Network)进行森林火灾情形下的烟雾检测试验^[5]。同年,在文献[6]中,通过使用MobileNetV2替换SSD的VGG主干网络的方法,利用特征金字塔结构实现多尺度特征提取,提高火焰与烟雾检测精度^[6]。到2021年,文献[7]整合了Yolov5和Efficient Net完成火灾探测过程,采用Ensemble高效学习网络负责学习全局信息,以避免误报^[7]。

2021年,YOLOX^[8]算法问世,因其优秀的实时检测速度兼顾精准的检测性能,在YOLO系列算法中脱颖而出。2022年,文献[9]提出改进的T-YOLOX,在CSP层添加轻量级注意力模块,通过全连接层、ReLU和Sigmoid激活函数来削弱噪声,同时将主干网络最后一层替换为MobileViT Block,提升对特征信息的感知能力^[9]。2023年,文献[10]将主干和颈部网络进行多尺度融合,利用改进Ghost-CSP简化模型,提升检测速度^[10]。

针对以上场景中的不足,本文提出一种YOLOX的改进算法,通过替换主干网络为基于分组卷积的轻量化神经网络,并在主干网络中扩大感受野,改进上采样方式,增加注意力机制,实现在提高检测精度的同时,能

够大幅降低网络计算量。

1 YOLOX 目标检测原理

YOLOX是基于YOLOv3的改进算法,融合了YOLO系列网络的优点,保留YOLOv5的CSP-Darknet53主干网络,共3个部分,分别为主干网络、颈部网络和检测头。与YOLOv5相比,YOLOX将原本的回归与分类二合一的检测头替换为两支输出路径,一支负责物体预测的类别,另一支负责判断是否含有需要检测的类别,最后对特征信息进行融合。采用不同的分支来进行运算,有利于效果的提升。另外,YOLOX在Focus的基础上用Anchor-free代替传统的Anchor-base,舍弃锚框,只将物体中心点所在的位置作为正样本,由原来检测3组预选框的中心的坐标和宽高减少至只检测一组数值,缓解了正负样本不平衡的问题,降低时间成本。根据模型参数数量和计算量的不同,YOLOX具有YOLOX-s、YOLOX-m、YOLOX-l、YOLOX-nano和YOLOX-tiny等不同的版本,在实际模型的应用中,模型越大,检测效果越好,但这样会导致需要更高的算力设备来运行,模型越小检测速度越快,但检测精度会相对较低,这同样为火焰检测任务带来困难,综合考虑,选用计算参数适中的YOLOX-m模型是本文火焰检测任务的最优选择。YOLOX-m的整体结构如图1所示。

2 基于ShuffleNetV2的YOLOX改进算法

2.1 主干网络改进

由图1,YOLOX-m的主干网络采用CSPDarknet-53结构,主要由Focus、CBS模块、CSP模块和SPP模块组成。将图片(640, 640, 3)输入主干网络,依次输出三组特征图,分别为小目标(20, 20)、中目标(40, 40)和大目标(80, 80)^[11]。本文保留Focus模块,对CBS和CSP模块进行改进。首先将由卷积层、Batch Normalization层(下称BN层)和SiLU激活函数组成的CBS结构替换为ShuffleNetV2的基本组成单元Shuffle Unit(步长为2)(图c),Shuffle Unit将残差网络(图a)中的逐点卷积替换为卷积核为 1×1 的分组卷积,显著减少网络所需的参数数量和计算量,并插入通道重排(Channel Shuffle),将特征图重新排列组合,避免因分组后各组(各特征图)之间没有联系带来信息丢失,保证计算的准确性和可靠性^[12]。残差边增加平均池化层,将两个分支在通道数上堆叠,最终输出特征图尺寸减半,通道数翻倍^[13]。其次将CSP结构替换为Shuffle Unit(步长为1)(图b),此时输出通道数扩大到输入通道数的两倍。基于上述改进方法,本文将原主干网络中的D2-D5中的CBS和CSP模块均替换为Shuffle Unit,改进后的结构如图3所示。

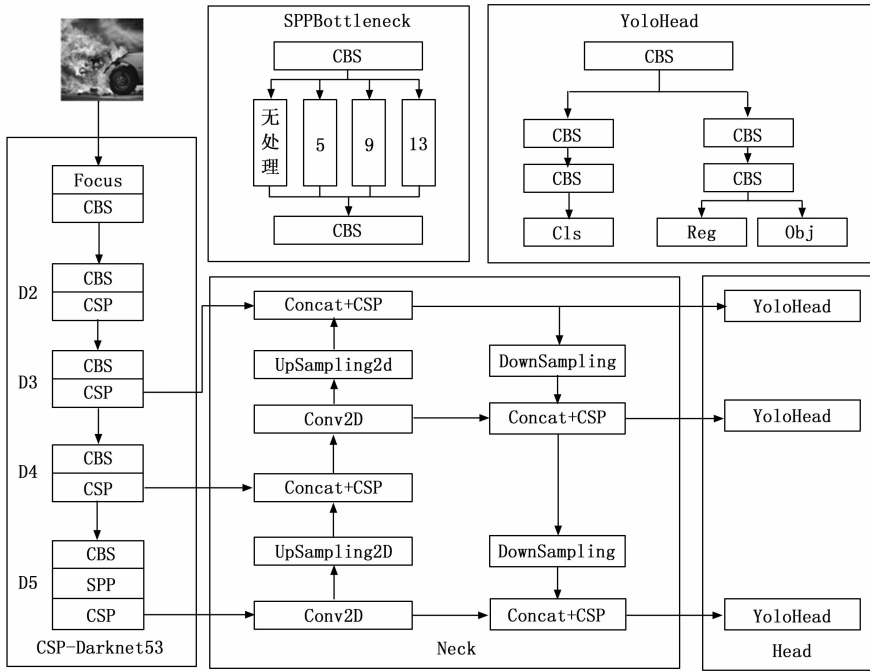


图 1 初始网络结构图

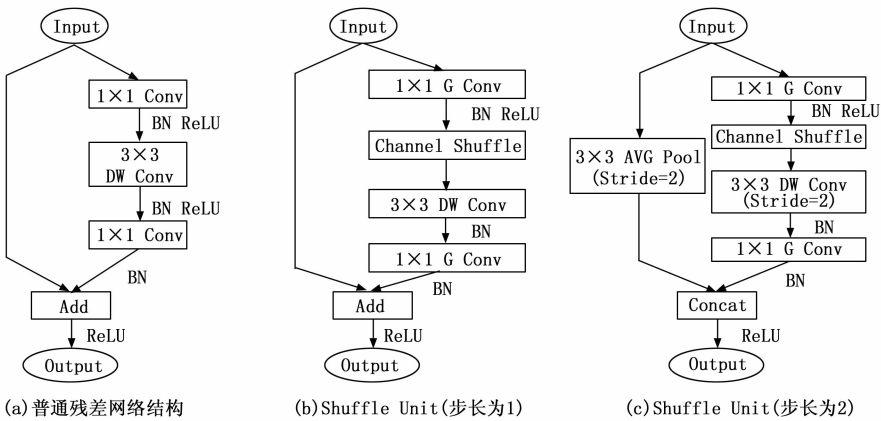


图 2 通道混洗结构图

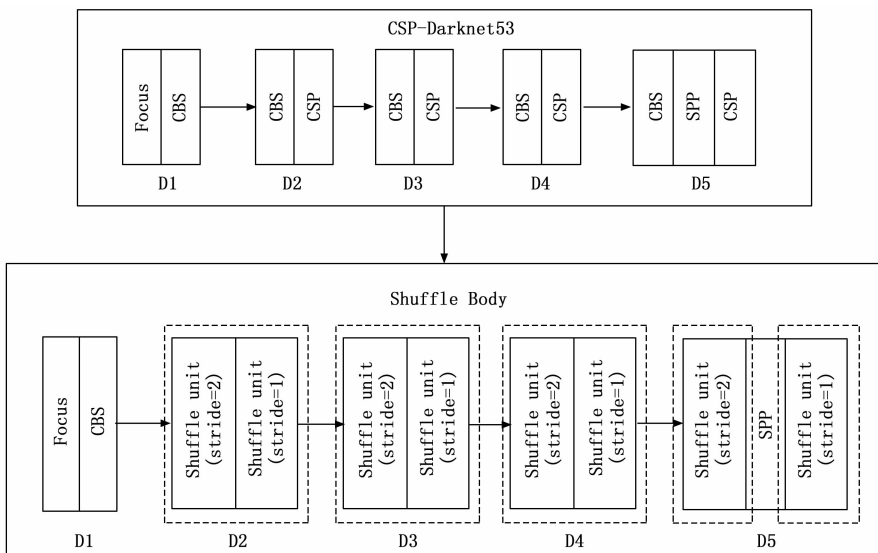


图 3 改进后的主干结构图

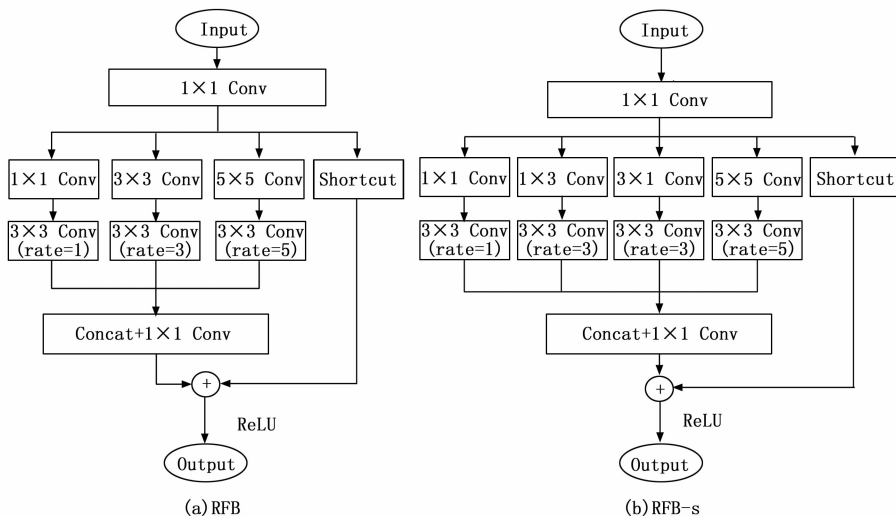


图 4 感受野模块结构图

2.2 感受野增强

本文在主干网络 D2-D4 中插入感受野模块 (RFB, receptive field block), 将该结构置于步长为 1 和步长为 2 的 Shuffle Unit 之间, 通过修改间隔长度, 获得不同尺度的感受野, 捕获多尺度上下文信息, 可在保持参数数量和相同卷积核尺寸的情况下获取更高分辨率的特征, 以扩大网络的感受野, 减少计算量, 具有高效性、多尺度适应性和准确性^[14]。D5 中 SPP 模块也有相同效果, 故 D5 不作改变。RFB 结构以空洞卷积为基础单元代替普通卷积, 首先通过 1×1 卷积调整通道数; 接着分别进行卷积核为 1×1 、 3×3 和 5×5 的普通卷积; 然后对其后两支进行膨胀率为 3 和 5 的空洞卷积; 最后将 3 个分支的特征进行堆叠^[15]。本文采用的 RFB-s 是 RFB 结

构的轻量版, 在原版基础上将中间分支分为两支, 其中中间步骤 (3×3 普通卷积) 改为一支使用 1×3 普通卷积和另一支使用 3×1 普通卷积; 第三支用卷积核 3×3 的普通卷积代替原来的 5×5 卷积。轻量版 RFB 在保证扩大网络感受野效果的同时, 很大程度上降低了该模块的计算量, 使用 RFB 旨在扩大感受野, 由于 SPP 模块同样可以实现此目的, 因此 RFB 模块只用于 D2-D4。本文将该结构置于步长为 1 和步长为 2 的 Shuffle Unit 之间, 大大提高了测量精度, 改进后的主干网络图如图 5 所示。

2.3 特征提取结构改进

为了解决特征提取过程中火焰图像与普通物体特征提取存在差异的问题, 本文对 YOLOX 的 Neck 部分进

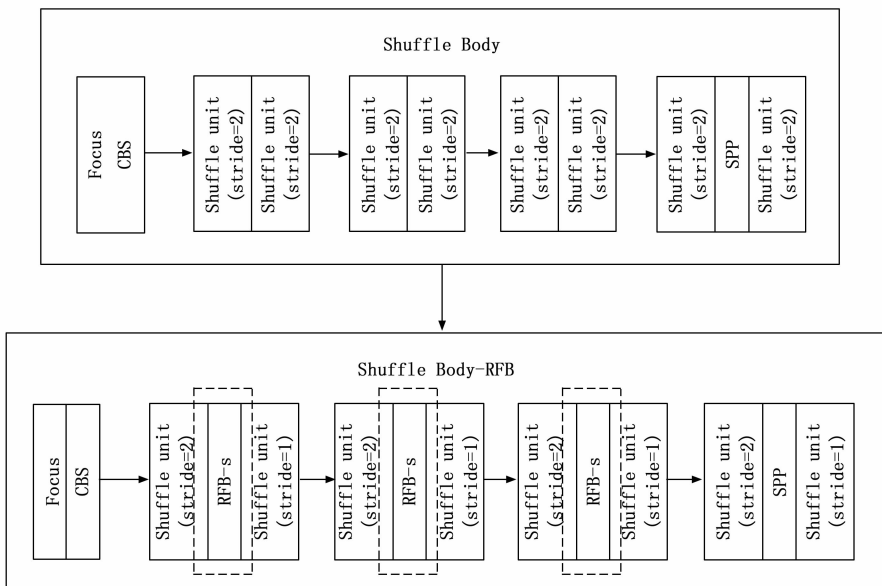


图 5 扩大感受野的改进主干网络图

行了改进。原 YOLOX 在主干网络提取的部分采用了 PANet 网络来加强特征提取，在 FPN 结构的基础上有所改进，增加了一条自下而上的融合路径，尽可能保留原始图像的浅层信息，增强了网络的表征能力。然而，由于火焰图像本身存在分辨率较低、边缘分界线模糊且纹理细节信息丢失严重的问题，所以即使使用了 PANet，仍不能有效地提取火焰图像的特征。对此，本文提出基于 PANet 改进的特征提取方式，用来提取火焰图像的深层特征，以便网络能够更好地检测火焰。YOLOX 原文采用的最邻近元插值法^[16]十分简单，不需要复杂的计算，方法是选择与所求像素距离最近的像素灰度值直接作为待求像素值。然而操作简便的同时也带来了问题，只选取最临近像素的灰度值作为上采样结果极有可能偏离实际图像特征，产生较大误差。本文将上采样替换为 Pixel Shuffle (像素重组)，原理是将一个 H (高) $\times W$ (宽) $\times C$ (通道数) 的分辨率的图像输入网络，通过上采样将其扩大为 $rH \times rW \times C$ 的高分辨率图像。再通过普通卷积得到分辨率为 $H \times W$ ，通道数为 $r^2 C$ 个的特征图，然后进行通道混洗筛选输出分辨率为 $rH \times rW$ ， $\frac{C}{r^2}$ 的特征图，即上采样后图像通道数变为原来的 $\frac{1}{r^2}$ ，其中 r 为上采样因子，即特征图宽高的扩大倍率。通过此方法降低上采样操作中损失的特征量，尽可能保留更多的信息，有利于特征提取，高效地进行上采样操作。由于普通上采样操作后通道数不变，使用 Pixel Shuffle 的上采样方法与普通上采样法后得到的通道数不同，因此替换上采样方法后需在网络加入逐点卷积调整通道数，改进后的 PANet 网络如图 6 所示。

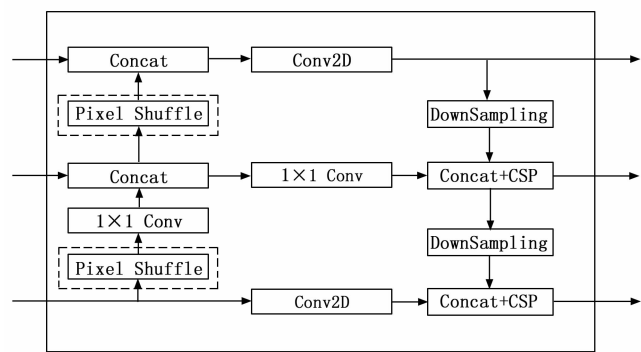


图 6 改进后的颈部网络结构图

2.4 添加注意力机制

对于一张图片，人眼会不由自主地关注更敏感的有效信息。同理，添加注意力机制能够赋予网络这种能力。以火焰检测为例，注意力可以使模型更多关注在火焰本身，适当忽略背景。注意力机制包括单路和多路两种，本文采用的 CBAM (卷积块注意力机制) 属于多路

注意力机制。该机制同时在通道和空间两个维度上提升网络注意力，由 CAM (通道注意力模块) 和 SAM (空间注意力模块) 两部分组成。首先 CAM 模块 (图 7) 将输入特征图进行最大池化和平均池化，可以学习到每个通道在特征表示中的重要性。使用全局平均池化操作来获取通道的全局信息，并使用全连接层来学习通道权重，进而将通道权重与原始特征图相乘以得到具有更强表征能力的通道注意力特征图。而 SAM 模块 (图 8) 负责处理图像分类，着眼于图像中重要的区域，进行通道拼接，通过对输入特征图的空间维度进行注意力调节，可以学习到不同位置的特征区域的重要性。该模块使用一个卷积操作来获取特征图的空间信息，并使用一个全连接层学习空间权重。然后，将空间权重与原始特征图相乘以得到同样具有更强表征能力的空间注意力特征图^[17]。使用 CBAM 注意力机制可以帮助网络获得目标的重要特征信息，提升模型的检测性能，在特征图的通道和空间维度上自适应地选择重要的信息，使模型更加关注重要的特征区域。CBAM 的作用原理如图 9 所示。具体实现方式是对输入进行加权再输出，给网络需要多加关注的地方较大权重，给不需要关注如背景等较小权重，以此方法更高效地利用网络检测火焰。由于 CBAM 模块不改变通道数，可以用于任意需要提升注意力的位置，本文将注意力机制模块添加于主干网络与颈部网络之间，对 (20, 20)、(40, 40) 和 (80, 80) 3 个不同宽高的特征图均进行注意力提升，通过学习到的注意力权重，自适应地调节不同空间维度的特征图的重要性，以提升模型的性能。

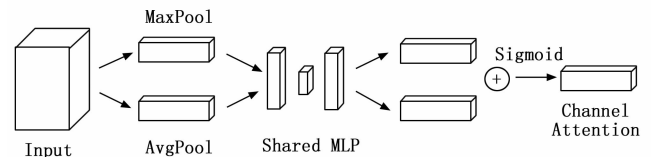


图 7 通道注意力模块

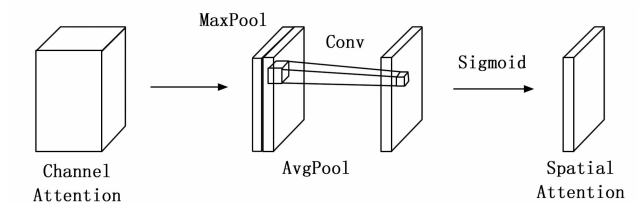


图 8 空间注意力模块

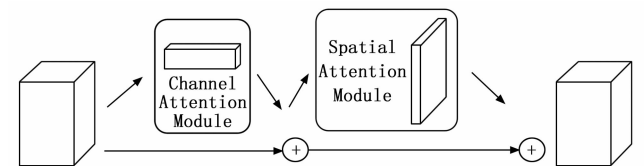


图 9 卷积注意力模块

2.5 损失函数

损失函数是一种用来估算预测值和标签值间差距大小的函数。损失越小,则模型的鲁棒性越高。YOLOX采用BCE计算Cls(类别预测)和Obj(置信度)损失,用IoU损失函数计算Reg(边界框回归)损失^[18]。其中IoU是检测框和预设框的交并比,交并比越大,说明检测框与火焰的实际框重合面积越大。IoU计算公式及最终损失函数为:

$$L_{IoU} = 1 - IoU \quad (1)$$

式中,

$$IoU = \frac{|B^{pr} \cap B^{gt}|}{|B^{pr} \cup B^{gt}|} \quad (2)$$

式中, B^{pr} 表示检测框的区域, B^{gt} 表示火焰真实框的区域。由公式可知当检测框和真实框并无交集时, $IoU = 0$,此时 $L_{IoU} = 1$,意为模型没有损失,显然与事实不符。因此文献^[19]提出优化损失函数GIoU,其计算公式及最终函数表示如下^[19]:

$$L_{GIoU} = 1 - GIoU \quad (3)$$

式中,

$$GIoU = IoU - \frac{C - (B^{pr} \cup B^{gt})}{C} \quad (4)$$

式中, C 表示包含检测框和真实框的最小矩形的面积,作为损失函数的惩罚项。可以在一定程度上抑制两框无交集情况对损失函数的影响。然而当一个框包含于另一框时, $GIoU$ 退化为 IoU ,针对该情况,在文献^[20]中提出 $DIoU$ 损失函数^[20],计算公式如下:

$$DIoU = IoU - \frac{d^2}{c^2} \quad (5)$$

式中,

$$d = \rho(B^{pr}, B^{gt}) \quad (6)$$

式中, d 指检测框和真实框中心坐标的距离, c 表示包含两框最小矩形的对角线长度,以两距离比值的平方作为惩罚项。然而当两中心点重合时, d 值为0。对于此情况,文献^[20]在DIOU的基础上,又提出CIoU损失函数,引入宽高比作为惩罚项。计算公式为:

$$CIoU = IoU - \frac{d^2}{c^2} - \alpha \quad (7)$$

式中,

$$\alpha = \frac{\nu}{(1 - IoU) + \nu} \quad (8)$$

$$\nu = 4/\pi \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w^{pr}}{h^{pr}} \right) \quad (9)$$

式中, α 是权重参数, ν 代表检测框和真实框宽比的相似程度。式中 C_w 、 C_h 分别是包含两框最小矩形的宽度和高度。在 $DIoU$ 的基础上增加了长和宽的 $loss$,这样预测框就会更加的符合真实框。损失函数表示为:

$$L_{DIoU} = 1 - DIoU \quad (10)$$

$$L_{CIoU} = 1 - CIoU \quad (11)$$

$CIoU$ 、 $DIoU$ 损失函数考虑了真实框与检测框的宽高比,但依然与真实的宽高损失有一定出入。本文采用的 $EIoU$ ^[21],在YOLOX采用的 IoU 的基础上将宽高比的影响因子拆开,分别计算检测框和真实框的长和宽,使得检测框的宽高分别与真实框的宽高之差最小,加快收敛。 $EIoU$ 的损失函数如下:

$$L_{EIoU} = 1 - EIoU \quad (12)$$

式中,

$$EIoU = IoU - \frac{\rho^2(B^{pr}, B^{gt})}{c^2} - \frac{\rho^2(w^{pr}, w^{gt})}{C_w^2} - \frac{\rho^2(h^{pr}, h^{gt})}{C_h^2} \quad (13)$$

2.6 改进后的网络结构

本文提出了一种基于ShuffleNetV2的改进YOLOX-m网络检测火焰。将主干网络中的CBS和CSPLayer分别替换为Shuffle Unit (stride = 2)和Shuffle Unit (stride = 1),并两者之间插入RFB结构,以扩大感受野。此外,将特征提取网络中的上采样替换为Pixel Shuffle,减少信息损失。最后,在主干网络和颈部网络中间插入CBAM注意力机制模块,提升模型检测性能。最终网络结构图如图10所示。

3 实验结果与分析

3.1 实验环境

改进YOLOX-m模型使用版本PyTorch深度学习框架搭建。在GPU为NVIDIA GeForce RTX 3090,内存为80 G,操作系统为Windows 10的PC机上训练模型。输入图片(Input_size)尺寸为(640, 640),训练批次大小(Batch_size)为8,即每次传递给模型用于训练的数据个数为8。优化器类型(Optimiser_type)使用Adam优化器,最大学习率(Init_lr)为 1×10^{-3} ,最小学习率(Min_lr)为 1×10^{-5} ,学习率的大小决定目标函数收敛的速率。总共训练300个Epoch训练,所用的参数如表1所示。

表1 训练参数

Name	Value
Input_size	640×640
Batch_size	8
Optimiser_type	Adam
Init_lr	1×10^{-3}
Min_lr	1×10^{-5}
Epoch	300

3.2 数据集

本文在训练过程中采用BoWFire Dataset(ht-

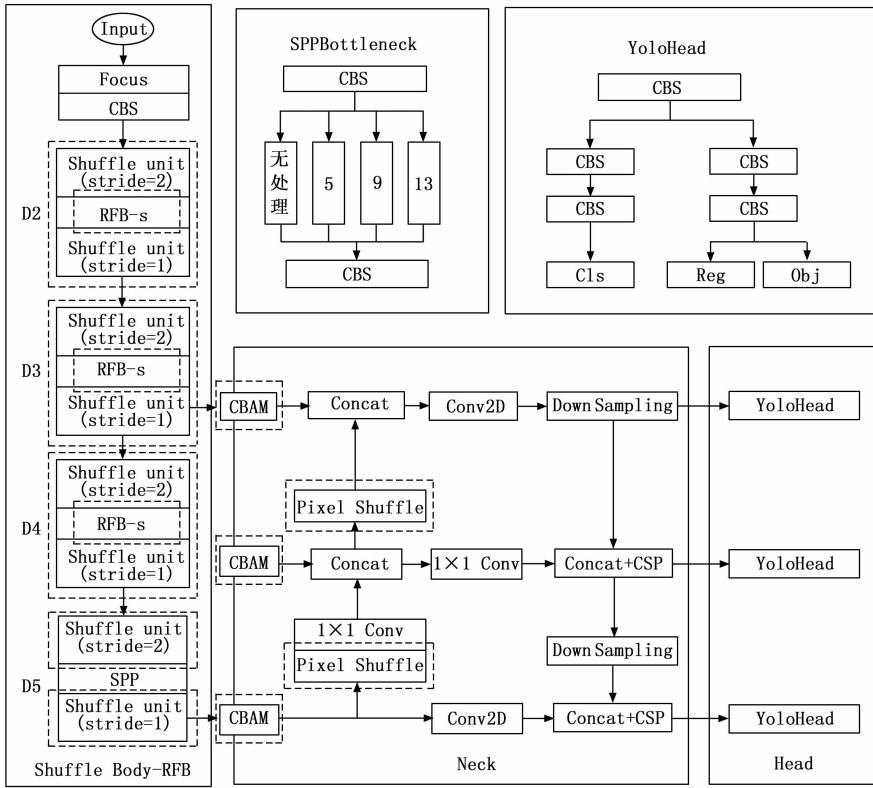


图 10 改进后的网络结构图

tps: //bitbucket.org/gbdi/bowfire-dataset/downloads/) 和部分互联网采集图片自制数据集进行试验, 包含 3 612张各种不同尺寸的火焰图像, 其中包含训练集 2 600张图片, 验证集 289 张图片和测试集 723 张图片。数据集中的图片经预处理得到 (640, 640, 3) 的图像输入网络中进行统一处理。火焰数据集虽有数据真实的优点, 但由于部分图像拍摄时间较早, 经过数年流传, 画质已然模糊。另外, 数据集中许多图片为城市监控摄像或无人机拍摄视频截图, 清晰程度与火焰可见度受到天气、天色明暗程度、拍摄距离、视线是否被阻挡等因素影响。为此, 将火焰数据集在翻转、平移、裁剪、缩放、对比度、噪声等随机数据增强方法的基础上, 采用马赛克 (Mosaic) 增强方法^[22]。马赛克数据增强将四幅图像混合拼接成一个图像, 即可丰富图像背景也可增加更多的小目标, 使得模型在做归一化操作时可一次性计算四张图片的特征, 增强了模型的鲁棒性。但由于四张图片在使用马赛克数据增强的方法拼接的过程中, 原图的像素特征会减少, 这也间接导致原本标定的小目标的图像特征会变得更模糊, 较小的目标会变得更小, 最终导致模型的泛化能力整体变差。为解决这一问题, 在原本的数据集中选取百分之五十进行 Mix up 数据增强, 以此扩大数据集, 增加训练数据, 并在训练后百分之三十代 (第

210epoch 起) 停止使用数据增强方法, 在不失原有的部分关键特征的基础上, 让其接近真实场景进行训练, 有效提升模型整体的检测能力。

3.3 检测指标

本文选用 Precision、AP 和 FLOPs 作为模型的评估指标, 其中:

1) 预测指标 Precision 是检测精度, 计算公式为:

$$Precision = \frac{TP}{TP + FP} \quad (14)$$

式中, TP 为 True Positive, 即预测结果是火焰, 标签值也是火焰, 预测正确的情况; FP 为 False Positive, 即预测结果为火焰, 但没有标签值, 预测错误的情况。当 Precision 越大时, FP 越小, 错误地将没有标签的物体预测成火焰的个数也越少, 可以理解为预测出的正例纯度越高, 误检越少。

2) AP 是以 Precision (精度) 和 Recall (召回率) 为坐标形成的曲线与坐标轴围成的面积。AP 是目标检测中最重要的指标, 体现模型的整体检测水平。其中 Recall 的计算公式为:

$$Recall = \frac{TP}{TP + FN} \quad (15)$$

式中, FN 为 False Negative, 即没有预测出火焰, 但标签值为火焰, 预测错误的情况。若一个模型的 AP 越

大, 说明 PR 曲线与坐标轴围成的面积越大, *Precision* 与 *Recall* 在整体上也相对较高。

3) FLOPs 指浮点计算次数, 用来衡量模型的复杂度, 数字越大, 则浮点计算次数越多, 模型越复杂。

3.4 消融实验

为了得到更客观公正的实验结果, 本文通过反复测试, 做对照组实验, 验证 ShuffleNet、RFB、Pixel Shuffle、CBAM 及其组合对检测性能提升的有效性, 获得数据如表 2 所示。本文算法 AP、Precision 分别提升了 2.87 个百分点, 计算量降低了 30.7%, 证明本文提出的改进型 YOLOX 算法是有效果的。

表 2 消融实验

	Shuffle NetV2	RFB	Pixel Shuffle	CBAM	EIoU	Precision	AP	FLOPs ($\times 10^9$)
1	—	—	—	—		75.91	62.03	73.726
2	✓					77.21	62.49	54.402
3	✓	✓				77.85	63.22	57.103
4	✓		✓			77.42	63.39	42.923
5	✓	✓	✓			77.36	63.90	51.090
6	✓	✓	✓	✓		78.74	64.16	51.096
7	✓	✓	✓	✓	✓	78.87	64.90	51.096

实验 1 采用原版 YOLOX-m, AP 为 62.03%。由实验 2 可得, 仅改变主干网络时, 网络在大幅度降低计算量的同时, 精度提高了 0.46 个百分点, ShuffleNetv2 将原来的普通卷积替换为分组卷积和逐深度卷积, 降低了网络的参数量; 由于前面的操作, 不可避免地降低了提取特征的能力, 因此加入通道重排, 加强各个特征图之间的联系, 减少信息丢失。实验 3 中, 在主干网络加入 RFB 模块后, AP 再次提升了 0.73 个百分点, 但计算量略有提升。由实验 4, 若不加入 RFB, 将上采样替换为 Pixel Shuffle, 此时 AP 提升 1.36 个百分点, 计算量再次降低 21.1%。实验 5 中, 若更改主干网络并加入 RFB 同时替换上采样, 则精度可以提升 1.87 个百分点。在此基础上, 实验 6 增加注意力机制模块, AP 提升至 64.16%。实验 7 中, 将 YOLOX 原损失函数 IoU 更改为 EIoU, 加快收敛, 最终精度提升至 64.90%, 计算量降低至 51.096 GFLOPs (1 GFLOPs = 10^9 FLOPs)。实验表明每个模块都是有意义的, 均对模型具有优化作用。

3.5 对比实验

为验证本文改进模型的火焰检测性能, 将改进的 YOLOX-m (Improved-YOLOX) 与主流算法对比。选取 SSD^[23]、Faster-RCNN^[24] 和 YOLOX-m 模型, 训练相同数据集并设置相同训练参数, 训练结果如表 3 所示。由表 3 分析可得, 本文模型比 YOLOX 检测精度提

升了 2.87%, 计算量降低了 22.63 FLOPs, 参数量 (Params) 减少了 9.7×10^6 个。改进 YOLOX 的 AP 值相较于其他主流算法精确度 (Precision)、AP、计算量、模型权重方面均有不同程度的优化, 证明本文算法在火焰检测方向具有先进性。

表 3 对比实验

Model	AP	FLOPs ($\times 10^9$)	Params ($\times 10^6$)
SSD	49.82	60.756	23.612
Faster-RCNN	47.38	369.719	136.689
YOLOX-m	62.03	73.726	25.281
Improved-YOLOX-m	64.90	51.096	15.594

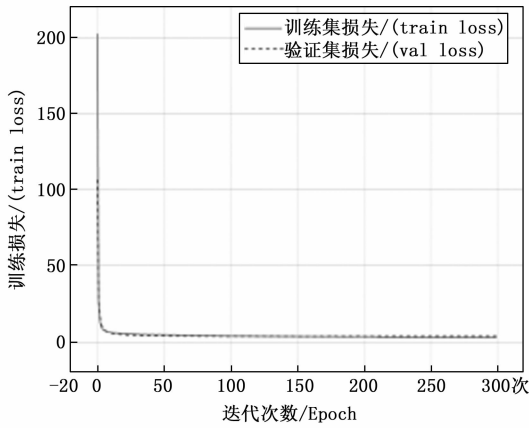
3.6 训练与检测结果

按照本文提出的优化算法进行训练的结果如图 11 中 (a) 和 (b) 所示。图 (a) 是损失值随训练逐渐降低的图像。由于训练中未使用预训练模型, 初始损失值高于 200, 并在前 5 代训练中迅速降低至 6, 然后在后面的训练中缓慢趋向于 0。图 (b) 是平均检测精度随训练的变化趋势, 前期迅速升高至 60% 开始趋于平缓, 缓慢上升并伴随有一定波动。将测试集输入训练完成的模型, 得到的结果如图 (c) 和 (d) 所示。图 (c) 和图 (d) 分别是由训练中最佳结果所得的 Precision 和 AP。

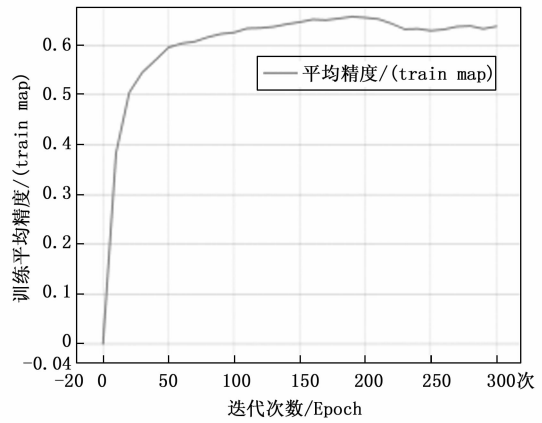
另外, 本文选用对比实验中四组模型训练好的权重预测火焰图所得结果如图 12 中 (a) 和 (b) 所示。图 (a) 是室外所摄电动车起火图片检测结果; 图 (b) 是室内圣诞树着火的图片检测结果, 红框为检测框, 数字是所框物体是火焰 (而非其他物体) 的置信度。图中从左向右依次为本文算法、SSD、Faster-RCNN 和 YOLOX 的检测结果, 4 种算法均可以有效检测火焰。然而 SSD 算法和 Faster-RCNN 算法置信度过高, 检测框过大, 出现误检状况 (如将火焰照在墙上的光检测成火焰)。YOLOX 算法表现尚可, 而本文算法可进一步确定火焰燃烧范围。由图中对比可知, 无论室内或是室外, 改进后的模型的检测性能更好, 能够更加精确地检测出图片中是否存在火焰, 并提供更准确的预测框。

4 结束语

针对火焰实时监测中识别速度较慢, 识别精度较低等问题, 本文对 YOLOX-m 网络进行改进, 精简主干网络, 扩大感受野, 提高网络提取特征的能力, 使模型的精度提高了 2.87 个百分点, 参数量减少了 9.7×10^6 个, 计算量减少了 30.7%, 在提高精度和降低网络复杂度都有一定的效果, 具有实用性。考虑到实际监控系

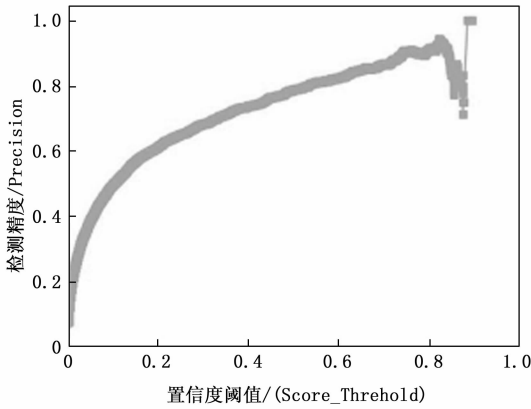


(a) 损失函数



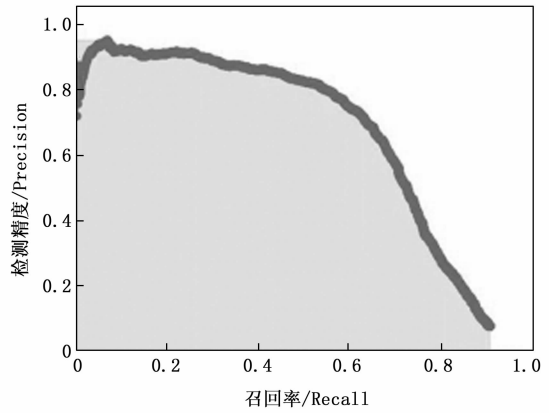
(b) 平均精度曲线

火焰类别检测精度: 78.87%
(置信度阈值=0.5)



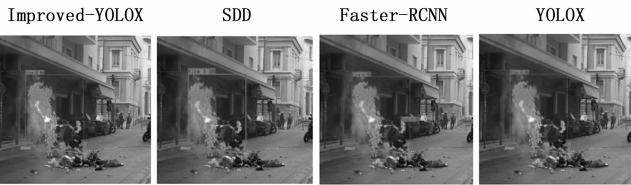
(c) 准确率变化曲线

火焰类别平均检测精度: 64.90%

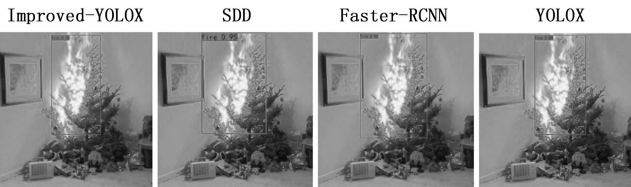


(c) 准确率变化曲线

图 11 改进后模型训练结果



(a) 室外场景



(b) 室内场景

图 12 检测结果

统的硬件设备算力较低，在后续研究中将尝试进一步减少网络计算量，同时保持检测精度，扩大适用范围，降低火焰监测成本，提高预防火灾的能力。

参考文献:

[1] 肖方. 消防救援局召开《“十四五”国家消防工作规划》

宣贯会 [J]. 中国消防, 2022 (4): 14.

[2] 严云洋, 唐岩岩, 郭志波, 等. 融合色彩和轮廓特征的火焰检测 [J]. 微电子学与计算机, 2011, 28 (10): 137 - 141.

[3] KRIZHEVSKY A, SUTKEVER I, HINTON G. ImageNet classification with deep convolutional neural networks [J]. Communications of the ACM, 2017, 60 (6): 84 - 90.

[4] 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, 2016: 779 - 788.

[5] WANG D C, CUI X, PARK E, et al. Adaptive flame detection using randomness testing and robust features [J]. Fire Safety Journal, 2013, 55: 116 - 125.

[6] 张勇. 基于改进型 SSD 的火焰及烟雾检测系统的设计与实现 [D]. 长沙: 湖南师范大学, 2021.

[7] XU R, LIN H, LU K, et al. A forest fire detection system based on ensemble learning [J]. Forests, 2021, 12 (2): 217.

- [8] GE Z, LIU S T, WANG F, et al. YOLOX: exceeding YOLO series in 2021 [J]. arXiv: 2107. 08430, 2021: 1 - 7.
- [9] 张剑飞, 柯 赛. 改进 YOLOX 火灾场景检测方法的研究 [J]. 计算机与数字工程, 2022, 50 (2): 318 - 322.
- [10] 张 龙. 基于 YOLOX-S 深度学习算法的多路视频火灾烟雾检测系统研究 [D]. 太原: 中北大学, 2023.
- [11] 汪子健, 高焕兵, 侯宇翔, 等. 改进 YOLOX-nano 的火灾火焰烟雾检测 [J]. 计算机系统应用, 2023, 32 (3): 265 - 274.
- [12] MA N, ZHANG X, ZHENG H T, et al. ShuffleNetV2: practical guidelines for efficient CNN architecture design [M/OL]. ArXiv, 2018 [2023-06-16]. <http://arxiv.org/abs/1807.11164>.
- [13] 王洪义, 孔梅梅, 徐荣青. 基于改进 YOLOV5 的火焰检测算法 [J]. 计算机与现代化, 2023 (1): 103 - 107.
- [14] LIU S, HUANG D, WANG Y. Receptive field block net for accurate and fast object detection [M/OL]. ArXiv, 2018 [2023-06-16]. <http://arxiv.org/abs/1711.07767>.
- [15] 乔 炎, 甄 彤, 李智慧, 等. 改进 YOLOv5 的安全帽佩戴检测算法 [J]. 计算机工程与应用, 2023, 59 (11): 203 - 211.
- [16] 王冠博, 赵一帆, 李 波, 等. 改进 YOLO v4-tiny 的火焰实时检测 [J]. 计算机工程与科学, 2022, 44 (12): 2196 - 2205.
- [17] 汪慧轩. 结合亚像素卷积的轻量级 YOLO 安全帽检测算法研究 [D]. 上海: 东华大学, 2022.
- [18] 邢志伟, 阚 犇, 刘子硕等. 基于改进 YOLOX-s 的机场跑道冰雪状态感知 [J]. 上海交通大学学报, 2023, 57 (10): 1292 - 1304.
- [19] 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/CVF Conference on Computer Vision and Pattern Recognition, 2019: 658 - 666.
- [20] ZHENG Z, WANG P, LIU W, et al. Distance-IoU loss: Faster and better learning for bounding box regression [C] // Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34 (7): 12993 - 13000.
- [21] ZHANG Y F, REN W Q, ZHANG Z, et al. Focal and efficient IOU loss for accurate bounding box regression [J]. Neurocomputing, 2022, 506: 146 - 157.
- [22] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. Yolov4: Optimal speed and accuracy of object detection [J]. Arxiv Preprint Arxiv: 2004.10934, 2020: 1 - 17.
- [23] LIU W, ANGUELOV D, ERHAN D, et al. Ssd: Single shot multibox detector [C] // Computer Vision-ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11 - 14, 2016, Proceedings, Part I 14. Springer International Publishing, 2016: 21 - 37.
- [24] GIRSHICK R. Fast r-cnn [C] // Proceedings of the IEEE International Conference on Computer Vision, 2015: 1440 - 1448.
- [25] LIN Y, LIU Z, SUN M. Knowledge graph embedding via dynamic mapping matrix [C] // Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics, Baltimore, Maryland, USA, 2022: 13209 - 13224.
- [26] HE S, LIU K, JI G, et al. Learning to represent knowledge graphs with gaussian embedding [C] // Proceedings of International Conference on Information and Knowledge Management, 2015: 623 - 632.
- [27] XIAO H, HUANG M, HAO Y, et al. TransG: a generative mixture model for knowledge graph embedding [C] // Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, Berlin, Germany, 2015, 1: 2316 - 2325.
- [28] SUN Z, DENG Z, NIE J, et al. RotatE: knowledge graph embedding by relational rotation in complex space [C] // International Conference for Learning Representations (2019), 2019: 1 - 19.
- [29] LIN Y, LIU Z, SUN M. Knowledge representation learning with entities, attributes and relations [C] // Proceedings of the 25th International Joint Conference on Artificial Intelligence, 2016: 2866 - 2872.
- [30] LIN Y, LIU Z, SUN M. Learning entity and relation embeddings for knowledge graph completion [C] // Proceedings of AAAI, 2015: 2181 - 2187.
- [31] XIE R, LIU Z, SUN M. Representation learning of knowledge graphs with hierarchical types [C] // International Joint Conference on Artificial Intelligence, 2016: 2965 - 2971.
- [32] JI G, HE S, XU L, et al. Knowledge graph embedding via dynamic mapping matrix [C] // Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics, 2015: 687 - 696.
- [33] YU J, CAI Y, SUN M, et al. SpaceE: knowledge graph embedding by relational linear transformation in the entity space [C] // Proceedings of the 33rd ACM Conference on Hypertext and Social Media, 2022: 64 - 72.
- [34] LI R, LI C, ZHAO J, et al. HousE: knowledge graph embedding with householder parameterization [C] // Proceedings of the 39th International Conference on Ma-

(上接第 11 页)