

复杂室内环境下轻量级手势识别算法^{*}

师红宇 刘蒙蒙 杜文 张哲于 李怡
(西安工程大学计算机科学学院 西安 710600)

摘要: 针对室内环境背景复杂、手部多样、识别角度多变等因素导致手势识别算法检测率低,算法复杂难以在移动端设备部署,提出了一种 SA-YOLOv8 手势识别算法。首先,利用改进后的 CB-ShuffleNet V2 轻量级网络作为主干网络提取手势特征,在保证准确率的同时降低模型参数与计算量,方便模型部署在智能家居设备,保证识别的实时性。其次,在 Neck 层引入渐进特征金字塔网络(AFPN)实现手势信息的多尺度特征融合,通过自适应空间融合操作避免复杂因素干扰,保留手部细节信息,提高模型鲁棒性。最后,在损失函数阶段引入 Shape-IoU 损失函数,增加模型对非规则手势与远距离小尺度手势识别的敏感力与准确性。实验结果表明,SA-YOLOv8 在 ASL-6 与完整 ASL 数据集上平均精度均值(mAP)mAP@0.5 分别达到 99.80%与 99.83%,相较于原始 YOLOv8 模型提高了 4.47%与 4.5%,模型参数量下降 80.18%,计算量减少 77.46%。改进后的算法在手势识别方面效果提升明显,且模型更加轻量,适合部署在移动端设备中。

关键词: 手势识别;YOLOv8;智能家居;轻量级网络;实时性

中图分类号: TP391.41;TN911.7 **文献标识码:** A **国家标准学科分类代码:** 520.2

Lightweight gesture recognition algorithm for complex indoor environments

Shi Hongyu Liu Mengmeng Du Wen Zhang Zheyu Li Yi
(School of Computer Science, Xi'an Polytechnic University, Xi'an 710600, China)

Abstract: To address the low detection rates in gesture recognition algorithms caused by complex indoor environments, diverse hand appearances, and variable recognition angles, and to facilitate deployment on mobile devices, we propose a novel SA-YOLOv8 gesture recognition algorithm. Initially, an improved CB-ShuffleNetV2 lightweight network is utilized as the backbone for extracting gesture features, ensuring accuracy while reducing model parameters and computational load, facilitating real-time recognition on smart home devices. Subsequently, an asymptotic feature pyramid network (AFPN) is integrated into the Neck layer for multi-scale feature fusion of gesture information, employing adaptive spatial fusion operations to mitigate interference from complex factors and preserve detailed hand information, thereby enhancing the model's robustness. Finally, the Shape-IoU loss function is introduced during the loss calculation phase, increasing the model's sensitivity and accuracy for irregular and small-scale gestures at a distance. The experiments demonstrate that SA-YOLOv8 achieves an average detection precision mAP@0.5 of 99.80% on the ASL-6 dataset and 99.83% on the full ASL dataset, marking a 4.47% and 4.5% improvement over the original YOLOv8 model, along with an 80.18% reduction in parameter volume and a 77.46% decrease in computational demand. The improved algorithm shows a significant enhancement in gesture recognition performance and is more lightweight, making it suitable for deployment on mobile devices.

Keywords: gesture recognition; YOLOv8; smart home; lightweight network; real-time

0 引言

随着互联网与科学技术的飞速发展,各种智能产品渐渐融入人们的工作、学习以及生活中,人们逐渐开始追求

舒适安全、智能便捷的生活方式,智能家居应运而生。传统家居设备多以手动或者无线方式进行控制^[1],依赖性较强,智能家居可以通过手势直接控制智能家居设备,方便用户操作,极大地提高了人机交互体验^[2]。目前市场上手

收稿日期:2024-10-21

^{*} 基金项目:陕西省重点研发计划(2022GY-074,2022GY-058)项目资助

势控制的智能家居设备存在识别率较低、操作复杂、实用性低、响应速度慢等问题,没有得到大范围推广。为了更好地提高智能家居设备的人机交互体验,手势识别算法亟需进一步的优化与改进。

近年来,手势识别领域吸引了众多学者进行深入研究。手势识别可分为静态与动态手势,动态手势可看为一组静态手势序列组成^[3],早期传统的手势识别方法分为基于传感器的识别方法和基于视觉的识别方法。基于传感器的识别方法通过可穿戴式物理设备^[4]采集手势运动信息,虽实时性高,但用户佩戴传感器设备繁琐不便,且成本较高。基于视觉的识别方法通过深度相机、背景差分法、模板匹配法、支持向量机等方法对手势进行识别,但易受环境因素影响,鲁棒性差,导致手势识别的准确率较低。

深度学习和神经网络技术的广泛应用为手势识别研究领域带来了创新的探索路径。吴晓凤等^[5]通过改进 Faster-RCNN 算法实现了同步检测和识别手势的目的,检测准确率达到 99.21%,但检测时间较长,实时性差。Ding 等^[6]提出一种融合神经网络与单次多框检测器(single shot multiBox detector, SSD)架构的方法,该方法在光线条件和肤色变化等干扰因素下能够提升手势识别的准确度,检测准确率为 95.36%,但其模型较为复杂,模型体积较大,导致检测存在延迟。牛群峰等^[7]提出一种基于改进 ResNet50 的表面肌电信号手势识别方法,实现了对相似度高且多种类手势的精准识别,最高达到 96.40%,但 ResNet50 模型复杂度较高,拥有较低实时性。梁彩虹等^[8]针对检测实时性与模型轻量化提出了一种基于 MobileNet 的轻量级卷积神经网络手势识别算法,模型检测准确率为 87.33%,模型性能有待提升。Zhou 等^[9]对 YOLO(you only look once)模型进行改进,提出了一种静态手势识别网络 PEA-YOLO,该算法提高了对于手势特征的定位和识别性能,优化了检测速度,模型检测准确率达到 97.87%,但在处理复杂背景下的小尺寸手势时检测性能表现不足。Peng 等^[10]提出了一种 HGDR-Net 的手势识别算法,该算法解决了原始 YOLO 对小尺寸手势检测效果差和定位准确度不高的问题,检测准确率达到 99.2%,但在光照变化、手势多变等问题干扰时,算法仍存在提升空间。

综合上述研究,可以看到手势识别模型在追求轻量化与实时性的同时,可能面临着低准确率的问题,另一方面,在复杂识别背景下,在保证高识别准确率时模型过于复杂也会导致识别延迟。并且在实际运用中室内环境因装修多样、家具摆放、光线不均匀等识别背景复杂多变,手部多样且复杂度高(纹理、肤色、遮挡),用户在不同距离与角度进行识别手势,导致手势识别难度进一步增大,可能会产生漏检、误检等问题。因此手势识别技术需要进行算法优化以提高其在实际应用中识别的准确率与实时性。经过对诸多文献和深度学习算法的综合对比后,本文选用改进

YOLOv8 算法来对智能家居背景下的手势进行识别研究。

1 智能家居手势识别流程

智能家居设备主要通过摄像头捕捉手势动作,将动态的手势动作分解成静态手势图像,对手势图像进行预处理之后输入手势识别模型,最后输出手势识别结果,将输出的手势结果映射成家居设备的开关指令,其工作流程如图 1 所示。

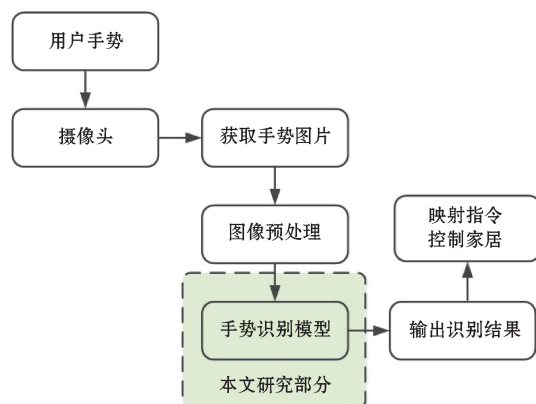


图1 智能家居设备识别手势流程

Fig. 1 Gesture recognition flowchart for smart home devices

2 SA-YOLOv8 网络结构

SA-YOLOv8 以 YOLOv8 模型为原型进行改进,由 Backbone 网络、Neck 网络与 Head 层网络 3 部分组成。如图 2 所示,首先,通过改进的轻量级网络 CB-ShuffleNet V2 作为模型的 Backbone 网络,使模型保持较高准确率的同时降低模型的参数与计算量,方便模型部署在智能家居设备并保证手势识别的实时性。其次,在 Neck 层引入渐进特征金字塔网络(asymptotic feature pyramid network, AFPN)实现手势局部与全局信息的多尺度特征融合,利用自适应空间融合操作动态调整融合权重,避免各种复杂因素的干扰,增强模型鲁棒性。同时,针对远距离识别与手势多样多变等特性在损失函数阶段引入 Shape-IoU 损失函数,通过考虑边框形状与尺度的度量,提高模型对非刚性物体手势和小尺度手势识别的敏感性与准确性。

2.1 CB-ShuffleNet V2 网络

1) 卷积块注意力模块(CBAM)

CBAM^[11]作为一种混合注意力机制,通过在卷积神经网络(CNN)中综合引入通道和空间注意力来增强特征表示,减少由于背景干扰、手势形状变化、用户间差异性问题导致的手势误检问题,有助于网络更准确地识别手势图像中的关键信息,提高网络泛化性与鲁棒性。同时, CBAM 模块主要依赖一维卷积和简单的全连接层,相对于传统的卷积层来说参数数量较少,其结构如图 3 所示。

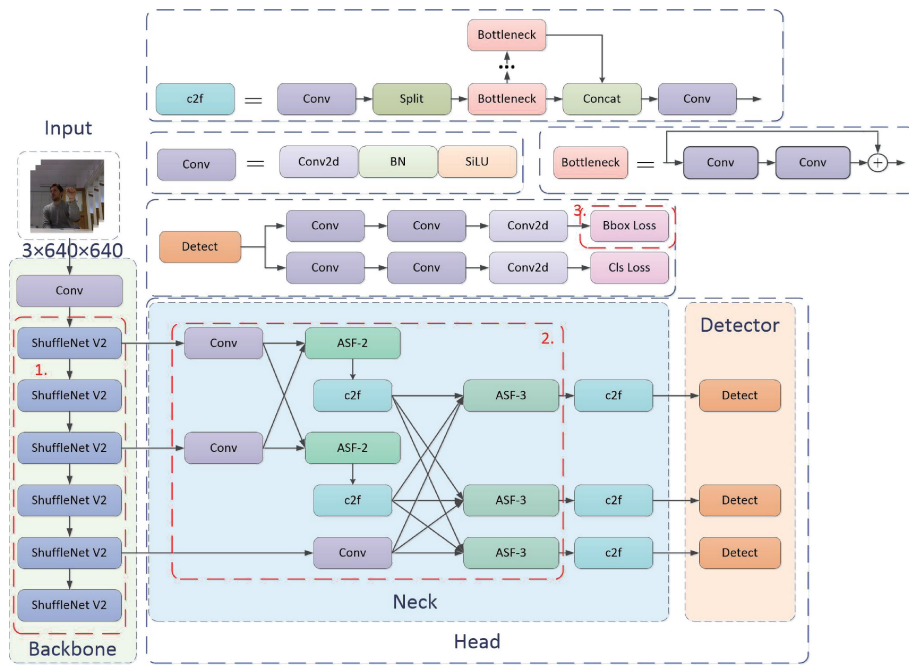


图2 SA-YOLOv8网络结构

Fig. 2 SA-YOLOv8 network structure

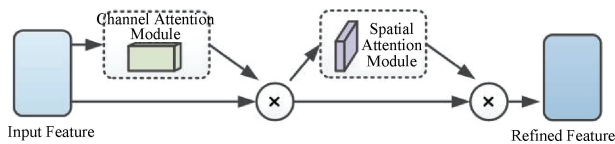


图3 CBAM模块结构

Fig. 3 CBAM network structure

2) B-ShuffleNet V2 网络

针对智能家居设备资源有限以及对准确率与实时性小模型的需求,SA-YOLOv8采用改进的CB-ShuffleNet

V2 轻量级网络作为其 Backbone 部分。ShuffleNet V2^[12]通过分组卷积、深度可分离卷积与通道重排操作显著减少模型的参数量,降低模型的计算成本,但同时可能会牺牲一定程度的特征表示能力。如图4所示,本文在ShuffleNet V2的基础上增加一组 3×3 的深度可分离卷积,通过增加网络的感受野来增强其特征表达能力,同时在网络最后的一个阶段加入CBAM模块,旨在提升ShuffleNet V2对全局信息的处理和关键手势特征的捕捉,同时保持了其轻量级的特性。

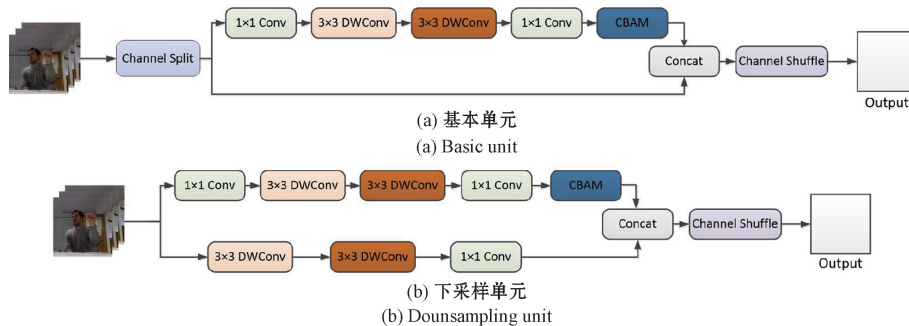


图4 CB-ShuffleNet V2网络结构

Fig. 4 CB-ShuffleNet V2 network structure

2.2 AFPN

针对复杂识别场景背景,以及手部特征具有个体差异性问题,SA-YOLOv8在Neck层引入AFPN^[13]。如图5所示,本文将AFPN原本结构中4个特征融合头减少至3个,更进一步的减轻模型复杂度。经调整后的AFPN在优化特征的渐进式融合过程中直接在非相邻层级P3、

P4、P5间进行特征交互,以增强模型对多尺度尤其小尺度手势的识别和定位能力,保留更多手部细节信息。并且,AFPN融合了自适应空间特征融合(ASF)模块,根据手势特征图中不同区域的重要性动态调整融合权重,避免在室内环境中因为复杂背景与光线变化等各种干扰因素降低手势识别的准确率。

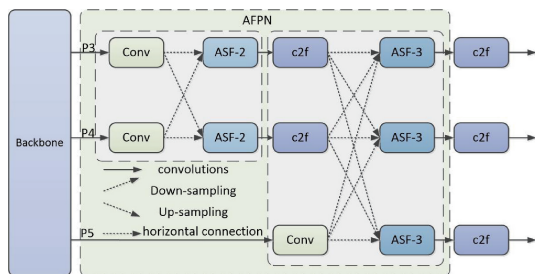


图5 渐进特征金字塔网络 AFPN 结构

Fig. 5 Asymptotic feature pyramid network structure

2.3 Shape-IoU 损失函数

YOLOv8 采用的损失函数 CIoU Loss 在考虑样本均衡性方面、长宽比差异较大与不规则手势判断能力和小手势检测方面具有局限性。

为针对 CIoU Loss 的不足,本研究引入考虑边框形状与尺度度量的 Shape-IoU 损失函数^[14],提高对多尺度尤其小型手势识别的准确性和鲁棒性,Shape-IoU 的计算公式如下。

$$IoU = \frac{|B \cap B^{gt}|}{|B \cup B^{gt}|} \quad (1)$$

$$ww = \frac{2 \times (w^{gt})^{scale}}{(w^{gt})^{scale} + (h^{gt})^{scale}} \quad (2)$$

$$hh = \frac{2 \times (h^{gt})^{scale}}{(w^{gt})^{scale} + (h^{gt})^{scale}} \quad (3)$$

$$d^{shape} = hh \times \frac{(x_c - x_c^{gt})^2}{c^2} + ww \times \frac{(y_c - y_c^{gt})^2}{c^2} \quad (4)$$

$$\Omega^{shape} = \sum_{t=w,h} (1 - e^{-w_t})^\theta, \theta = 4 \quad (5)$$

$$\begin{cases} w_w = hh \times \frac{|w - w^{gt}|}{\max(w, w^{gt})} \\ w_h = ww \times \frac{|h - h^{gt}|}{\max(h, h^{gt})} \end{cases} \quad (6)$$

式中:scale 表示缩放因子,与手势大小相关; ww 表示表示水平方向权重系数; hh 表示垂直方向上的权重系数,两者的值与 GT 框形状相关; w 与 h 为先验框的长和宽; x_c 与 y_c 为先验框中心点坐标; w^{gt} 与 h^{gt} 为 GT 框的长和宽; x_c^{gt} 与 y_c^{gt} 为 GT 框中心点坐标。

相应的边界框回归损失如下:

$$L_{Shape-IoU} = 1 - IoU + d^{shape} + 0.5 \times \Omega^{shape} \quad (7)$$

Shape-IoU 损失函数通过引入辅助边框的尺度因子来控制辅助边框的尺度大小,并结合形状因子来调整损失计算,以便更快更有效地获得回归结果。尤其在处理小尺寸和形状多变手势的情况下,更针对性地解决了 CIoU Loss 原本存在的问题。

3 实验与结果

3.1 实验设备配置

1)数据集

本文采用公开权威 ASL 数据集,为更针对智能家居

应用场景,选取 A、Y、I、J、L、W 等 6 种手势分别对应于窗帘、空调、门锁的开与关指令,如图 6 所示。为了获得更多的数据且提高模型泛化能力与鲁棒性,本文采用数据增强的方法对采集的图片进行翻转、剪切、噪声等操作,对数据集进行扩充。经过扩充后的数据集共 4 560 张,并以 8 : 1 : 1 的比例划分为训练集、验证集与测试集。

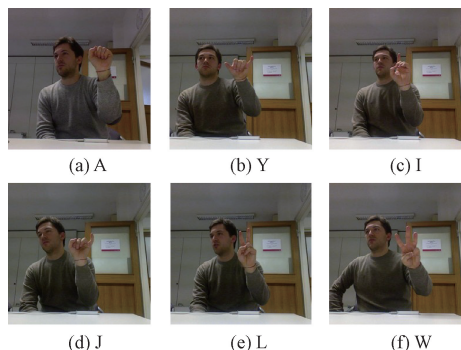


图6 数据集例图

Fig. 6 Dataset example graph

2)实验环境

实验所使用的配置环境如表 1 所示,实验未采用硬件加速技术。训练期间 GPU 利用率为 90%~100%,CPU 利用率为 30%~70%,GPU 显存、系统内存与数据 I/O 等方面未限制模型性能,BatchSize 最大为 16,设更大数值导致内存溢出迫使训练终止。

表1 实验环境配置

Table 1 Experimental environment configuration

配置	版本
操作系统	Ubuntu20.04
GPU	NVIDIA Tesla V100 32 G
GPU 卡数	1
CPU	8163CPU@2.50 GHz
CPU 核数	8
Python	3.9
CUDA	11.3
Pytorch	1.12
BatchSize	16

3.2 评价指标

为全面分析本文模型的各项性能,评估算法的有效性,选取精准率(precision, P)、召回率(recall, R)、平均精度均值(mAP)作为评价指标,其公式如下所示:

$$R = \frac{TP}{TP + FN} \quad (8)$$

$$P = \frac{TP}{TP + FP} \quad (9)$$

$$AP = \int_0^1 P(R) dR \quad (10)$$

$$mAP = \frac{1}{N} \sum_{n \in N} AP(n) \quad (11)$$

式中: TP 表示检测正确的数量; FP 表示检测错误的数量; FN 表示漏检的数量; AP 值表示 PR 曲线下面积、 AP 表示所有类别的 AP 值总和; N 表示类别; mAP 为全部类

别 AP 值的平均值。

3.3 消融实验

为了验证各改进方法对 YOLOv8 网络模型的影响, 对上述改进在相同数据集与参数环境下进行消融实验对比, 结果如表 2 所示, 其中“√”代表使用。

表 2 消融实验对比

Table 2 Comparison of ablation experiments

模型	CB-ShuffleNet V2	AFPN	Shape-IoU	$P/\%$	$R/\%$	$mAP@0.5/\%$	参数量/ $(\times 10^6)$	计算量/GFLOPs
1				97.10	93.61	95.33	11.1	28.4
2	√			98.54	93.90	97.92	3.4	8.3
3		√		98.02	94.30	96.80	3.1	9.3
4			√	98.71	94.70	98.34	11.1	28.4
5	√	√		98.92	94.24	98.46	2.2	6.4
6	√		√	99.01	95.32	99.50	2.6	6.8
7		√	√	97.62	93.11	97.31	3.1	9.7
8	√	√	√	99.90	96.63	99.80	2.2	6.4

由表 2 可知, 模型 8 即 SA-YOLOv8 将改进后的 CB-ShuffleNet V2 网络替换原模型骨干网络、引入 AFPN 模块和 Shape-IoU 损失函数共同改进 YOLOv8, 使得模型精准率与 $mAP@0.5$ 分别提升 2.80% 与 4.47%, 达到了 99.90% 与 99.80%, 参数量下降 80.18%, 计算量减少 77.46%, 实验表明本文方法在提升模型精度的同时还可以保持轻量化与手势识别的实时性。

3.4 对比实验

1) 主干网络对比

为验证本文对 ShuffleNet V2 网络的改进有效性, 以 YOLOv8 为基础进行验证, 实验结果如表 3 所示。

由表 3 可知, 相较于原始 DarkNet-53 网络, 使用 ShuffleNet V2 的模型在精准率和 $mAP@0.5$ 上分别提升了 0.81% 和 1.07%, 同时参数量减少了 72.9%, 计算量下降了 71.46%。进一步地, 使用 CBAM 改进的 ShuffleNet V2 网络虽因额外引入深度可分离卷积和 CBAM 模块增加了小部分参数量, 但实现了精确度和 $mAP@0.5$ 的显著提升, 分别提升了 1.44% 和 2.59%。综上所述, CB-ShuffleNet V2 在综合性能评估中表现出卓越的均衡性, 整体提高了网络性能, 证明了本文对骨干网络的改进措施有效提升了模型的检测精度和效率。

表 3 主干网络对比

Table 3 Comparison of backbone networks

主干网络	$P/\%$	$R/\%$	$mAP@0.5/\%$	参数量/ $(\times 10^6)$	计算量/GFLOPs
DarkNet-53	97.10	93.61	95.33	11.1	28.4
ShuffleNet V2	97.91	93.74	96.40	3.0	8.1
ECA-ShuffleNet V2	98.30	93.43	97.36	3.2	8.3
SimAM-ShuffleNet V2	98.17	92.77	96.80	3.0	8.1
SA-ShuffleNet V2	98.35	93.61	97.53	3.6	8.4
CB-ShuffleNet V2	98.54	93.90	97.92	3.4	8.3

2) AFPN 特征头对比

为验证本研究减少 AFPN 网络特征融合头的有效

性, 分别通过在 Neck 层引入 4 个和 3 个特征头的 AFPN 进行对比, 实验结果如表 4 所示。

表 4 AFPN 特征头对比

Table 4 Comparison of AFPN feature heads

模型	$P/\%$	$R/\%$	$mAP@0.5/\%$	参数量/ $(\times 10^6)$	计算量/GFLOPs
AFPN(4)	98.35	93.90	97.26	3.6	9.8
AFPN(3)	98.02	94.30	96.80	3.1	9.3

由表 4 可知,本文将 AFPN 的特征头减至 3 个,虽在准确率的提升上略逊于原 AFPN,但这种调整在不牺牲太多精度的前提下,提高模型的计算效率,实现了准确率和参数量之间更好的权衡,更适合于资源受限的手势识别任务。

3)模型性能对比
为验证本文所提改进方法的有效性与优势,将其与目标检测算法中较先进经典的算法在相同数据集上进行对比,选择精准率、召回率、mAP@0.5、模型大小与每秒传输帧数作为评价指标衡量算法复杂度,实验结果如表 5 所示。

表 5 与目标检测算法对比结果 (ASL-6)
Table 5 Comparison results with object detection algorithms (ASL-6)

模型	P / %	R / %	mAP@0.5 / %	模型大小 / MB	帧率 / fps
Faster R-CNN	90.81	90.31	89.20	107	39.1
SSD	89.70	87.82	87.01	85	43
YOLOv5	92.03	91.34	94.52	21.5	106.1
YOLOv7	91.73	90.42	90.62	71.4	90.8
YOLOv8	97.10	93.61	95.33	13.8	114.3
SA-YOLOv8	99.90	96.63	99.80	9.1	157.9

由表 5 可知,SA-YOLOv8 模型大小为 9.1 MB,占用内存较少,Faster R-CNN、SSD、YOLOv7 模型大小约为本文模型的 7~11 倍,难以部署在智能家居设备上,同时实时性与特征提取能力不足。YOLOv5 展现出了较好的综合性能,但 YOLOv8 作为基线模型在各方面均优于 YOLOv5 及其他模型,本研究的 SA-YOLOv8

在各项指标上都表现出较为明显的优势,模型更轻量化且推理速度快,每秒处理帧值达到 157.6 fps,具有更好的实时性,更适合在边缘侧和移动端智能家居设备上部署。
为更全面评估本文方法的泛化能力,与其他经典算法在完整 ASL 数据集上进行对比,实验结果如表 6 所示。

表 6 与目标检测算法对比结果 (ASL)
Table 6 Comparison results with object detection algorithms (ASL)

模型	P / %	R / %	mAP@0.5 / %	模型大小 / MB	帧率 / fps
YOLOv5s-ASFF-SE ^[15]	96.1	—	96.8	—	45
CBELU-YOLOv5x ^[16]	97.8	97.5	98.8	—	—
IMP-YOLOv5 ^[17]	93.8	91.8	94.7	—	196
Faster R-CNN	92.36	89.20	90.53	107	38.6
SSD	90.21	89.03	88.20	85	43
YOLOv5	93.41	90.70	95.0	22	106.3
YOLOv7	92.60	90.14	92.94	71.5	91.2
YOLOv8	97.57	94.0	95.74	13.8	112.0
SA-YOLOv8	99.92	96.10	99.83	9.1	158.2

由表 6 可知,在完整 ASL 数据集上,IMP-YOLOv5 模型每秒传输帧数最高,实时性能好,但其识别准确率性能不如本文模型。本文的 SA-YOLOv8 在各个评估指标上仍然展现出强大的竞争力,表明其在处理多样化手势识别任务时较强的泛化能力。

为了进一步验证本文所提出的手势识别方法的优越性,将本文方法与其他手势识别方法在同数据集上进行比较,结果如表 7 所示,可知与现有手势识别方法相比,本文提出的方法识别性能有明显提升。

3.5 检测结果可视化

为了更好地验证本文改进方法的有效性与泛化性,

表 7 与手势识别方法对比结果
Table 7 Comparison results with gesture recognition algorithms

方法	准确率 / %
改进的 3D 卷积神经网络 ^[18]	96.30
YOLO+LSTM ^[19]	98.07
对比学习 ^[20]	99.42
残差双注意力+跨级特征融合 ^[21]	99.68
SA-YOLOv8(ASL)	99.83
SA-YOLOv8(ASL-6)	99.80

在测试集中选取背景杂乱、小尺寸手势、光线分布不均、不规则手势等情况下较为代表性的手势图片以及同样情况下的自制图像进行测试,识别效果对比如图7所示。



图7 检测结果对比

Fig. 7 Comparison of detection results

图7A~C行依次代表模型YOLOv8、SA-YOLOv8以及SA-YOLOv8在自制图像上的表现,由图7可以看出,SA-YOLOv8的识别率较YOLOv8均有提升,并在自制图像中展现了良好性能。图8所示为模型改进前后的可视化实验结果曲线对比,横坐标为训练轮次,纵坐标为mAP@0.5指标,可以看出SA-YOLOv8比YOLOv8模型性能更好。

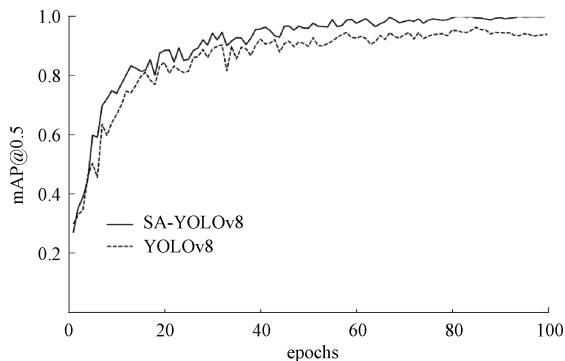


图8 改进前后曲线对比

Fig. 8 Comparison chart of curves before and after improvement

4 结论

本文提出一种SA-YOLOv8的手势识别方法,用于在复杂环境下智能家居设备识别手势。在针对室内环境背景复杂、手部多样、识别角度多变以及算法复杂难以在移动端设备部署等特点,首先采用改进后的CB-ShuffleNet V2网络作为模型骨干网络,在保证提取特征的同时大幅度减少模型参数量与计算量;其次在Neck层引入AFPN,增强模型对多尺度手势的特征融合能力,通过自适应空间融合操作动态调整融合权重,增强模型鲁棒性;最后,在损失函数部分引入Shape-IoU损失函数,提高模型对识别非

规则手势变化的敏感度。本文方法对于识别背景复杂、手势多变等问题,可以达到较好的识别效果与鲁棒性,并且极少的参数量与计算量为算法在资源有限的智能家居设备上部署提供非常大的便利。未来计划研究更先进的模型压缩技术,进一步落实高性能轻量化模型在移动端设备上的部署,设计直观、自然的智能家居控制系统,提供实用性的人机交互体验。

参考文献

- [1] 张家源,刘建华,傅周超,等. 基于手势识别的智能家居人机交互系统[J]. 电脑知识与技术, 2023, 19(13): 105-107.
ZHANG J Y, LIU J H, FU ZH CH, et al. A smart home human-computer interaction system based on gesture recognition [J]. Computer Knowledge and Technology, 2023, 19(13): 105-107.
- [2] ROBERGE A, BOUCHARD B, MAITRE J, et al. Hand gestures identification for fine-grained human activity recognition in smart homes [J]. Procedia Computer Science, 2022, 201: 32-39.
- [3] QI J, MA L, LUI Z C, et al. Computer vision-based hand gesture recognition for human-robot interaction: A review [J]. Complex Intelligent Systems, 2023, 10(1): 1581-1606.
- [4] HU B Y, WENG L, LIU K L, et al. Structural design of magnetostrictive sensing glove and its application for gesture recognition[J]. Sensor Review, 2024, 44(2): 113-121.
- [5] 吴晓凤,张江鑫,徐欣晨. 基于Faster R-CNN的手势识别算法[J]. 计算机辅助设计与图形学学报, 2018, 30(3): 468-476.
WU X F, ZHANG J X, XU X CH. Hand gesture recognition algorithm based on Faster R-CNN [J]. Journal of Computer-Aided Design & Computer Graphics, 2018, 30(3): 468-476.
- [6] DING C, LIN J, YOU J, et al. A gesture recognition method based on deep learning [J]. Control and Information Technology, 2018(6): 96-99.
- [7] 牛群峰,石磊,贾昆明,等. 基于改进ResNet50的表面肌电信号手势识别[J]. 国外电子测量技术, 2024, 43(4): 181-189.
NIU Q F, SHI L, JIA K M, et al. SEMG gesture recognition based on improved ResNet50 [J]. Foreign Electronic Measurement Technology, 2024, 43(4): 181-189.
- [8] 梁彩虹,钟艳涓,汤赵斌,等. 一种基于MobileNet的轻量级卷积神经网络手势识别算法[J]. 信息与电脑(理论版), 2023, 35(5): 169-172.
LIANG C H, ZHONG Y J, TANG ZH B, et al. A

- Lightweight convolutional neural network gesture recognition algorithm based on MobileNet [J]. Information & Computer (Theory Edition), 2023, 35(5): 169-172.
- [9] ZHOU W N, LI X L. PEA-YOLO: A lightweight network for static gesture recognition combining multiscale and attention mechanisms [J]. Signal, Image and Video Processing, 2023, 18(1): 597-605.
- [10] PENG Y Q, ZHAO X S, TAO H F, et al. Hand gesture recognition against complex background based on deep learning[J]. Robot, 2019, 41(4): 534-542.
- [11] 胡久松, 刘张驰, 余谦, 等. 融入 GhostNet 和 CBAM 的 YOLOv8 烟雾识别算法[J]. 电子测量与仪器学报, 2024, 38(8): 201-207.
- HU J S, LIU ZH CH, YU Q, et al. YOLOv8 smoke detection algorithm integrated with GhostNet and CBAM [J]. Journal of Electronic Measurement and Instrumentation, 2024, 38(8): 201-207.
- [12] MA N, ZHANG X, ZHENG H T, et al. ShuffleNet V2: Practical guidelines for efficient CNN architecture design [C]. European Conference on Computer Vision (ECVV). Springer, 2018: 116-131.
- [13] YANG G, LEI J, ZHU Z, et al. AFPN: Asymptotic feature pyramid network for object detection [C]. 2023 IEEE International Conference on Systems, Man, and Cybernetics (SMC). IEEE, 2023: 2184-2189.
- [14] ZHANG H, ZHANG S. Shape-IoU: More accurate metric considering bounding box shape and scale[J]. ArXiv preprint ArXiv: 2312.17663, 2023.
- [15] 陈帅, 袁宇浩. 改进 YOLOv5 的手语字母识别算法研究[J]. 小型微型计算机系统, 2023, 44(4): 838-844.
- CHEN SH, YUAN Y H. Research on improved hand gesture recognition algorithm based on YOLOv5 [J]. Mini and Micro Computer Systems, 2023, 44(4): 838-844.
- [16] ATTIA N F, AHMED M T F S, ALSHEWIMY M A M. Efficient deep learning models based on tension techniques for sign language recognition [J]. Intelligent Systems with Applications, 2023, 20: 200284.
- [17] 潘格, 许有熊, 刘晓峰. 基于改进 YOLO v5 的手语字母的识别方法[J]. 南京工程学院学报(自然科学版), 2023, 21(1): 27-32.
- PAN G, XU Y X, LIU X F. Recognition algorithm of sign language alphabet based on improved YOLO v5 [J]. Journal of Nanjing Institute of Technology (Natural Science Edition), 2023, 21(1): 27-32.
- [18] SHARMA S, KUMAR K. ASL-3DCNN: American sign language recognition technique using 3 - D convolutional neural networks[J]. Multimedia Tools and Applications, 2021, 80(17): 26319-26331.
- [19] RIVERA-ACOSTA M, RUIZ-VARELA J M, ORTEGA-CISNEROS S, et al. Spelling correction real-time american sign language alphabet translation system based on YOLO network and LSTM [J]. Electronics, 2021, 10(9): 1035.
- [20] 卞雨玮, 华立涛, 周媛. 基于对比学习的信息缺失手势识别新方法 [J]. 电子测量技术, 2023, 46(7): 180-186.
- BIAN Y W, HUA L T, ZHOU Y. Novel method for gesture recognition with missing information based on contrastive learning [J]. Electronic Measurement Technology, 2023, 46(7): 180-186.
- [21] 吴佳璐, 田秋红, 岳金鸿. 基于残差双注意力与跨级特征融合模块的静态手势识别[J]. 计算机系统应用, 2022, 31(11): 111-119.
- WU J L, TIAN Q H, YUE J H. Static gesture recognition based on residual double attention module and cross-level feature fusion [J]. Computer Systems & Applications, 2022, 31(11): 111-119.

作者简介

师红宇(通信作者), 硕士, 高级工程师, 主要研究方向为图像处理、深度学习、智能检测。

E-mail: shy510213@163.com

刘蒙蒙, 硕士研究生, 主要研究方向为手势识别、图像处理。

E-mail: Liumeng1252@163.com

杜文, 硕士研究生, 主要研究方向为密集人群估计、图像处理。

E-mail: duvin0512@163.com

张哲于, 硕士研究生, 主要研究方向为目标检测、图像处理。

E-mail: zzy1195264354@163.com

李怡, 博士, 副教授, 主要研究方向为人工智能、计算机检测。

E-mail: 119586420@163.com