

改进 FlowNetS 的光流估计算法研究

王雅妮¹, 翟正军^{1,2}, 代巍¹, 申思远¹

(1. 西北工业大学 计算机学院, 陕西 西安 710129;

2. 西北工业大学 计算机测控与仿真技术研究所, 陕西 西安 710072)

摘要:近年来,基于深度学习的光流估计方法在准确性和效率方面取得了显著进展。然而,其仍面临着对训练数据依赖性强、对特定场景敏感、计算资源消耗大、物理约束利用不足以及可解释性差等问题。该文提出了一种基于卷积神经网络的叠加融合方法,旨在提升光流特征的表达能力,提高预测精度并降低内存消耗。具体而言,设计了一种叠加融合模块即连续加法块(Continue-ADD-Block)。连续加法块通过添加一个 Conv7 层并采用连续下采样进行融合,有效地整合了不同尺度下的信息,增强了对复杂场景和多尺度运动的处理能力。在 Flying Chairs, KITTI Flow 2015 和 MPI-Sintel (clean 和 final) 数据集上的实验结果表明,连续加法块在降低资源消耗和内存占用的同时,在复杂场景下取得了更高的精度,表现为更低的最大的平均端点误差(End Point Error, EPE)即更高的精度,并且在 Flying Chairs 数据集中精度提升 2.86%,在 KITTI Flow 2015 数据集中精度提升 0.70%,在 MPI-Sintel (clean 和 final) 数据集中精度分别提升 7.40%、2.43%。这表明该方法在复杂场景下具有更强的鲁棒性,为光流估计领域提供了一种新的解决方案。

关键词:光流估计;分层策略;多尺度估计;叠加融合;卷积神经网络

中图分类号:TP391.41

文献标识码:A

文章编号:1673-629X(2025)07-0001-07

doi:10.20165/j.cnki.ISSN1673-629X.2025.0041

Advanced Research on Improving Optical Flow Estimation Based on FlowNetS

WANG Ya-ni¹, ZHAI Zheng-jun^{1,2}, DAI Wei¹, SHEN Si-yuan¹

(1. School of Computer Science, Northwestern Polytechnical University, Xi'an 710129, China;

2. Institute of Computer Measurement, Control and Simulation Technology, Northwestern Polytechnical University, Xi'an 710072, China)

Abstract: Deep learning-based methods for optical flow estimation have shown remarkable progress in recent years, achieving notable gains in both accuracy and efficiency. However, these methods continue to suffer from several shortcomings, including a strong dependency on extensive training datasets, sensitivity to specific environmental conditions, high computational costs, a lack of effective incorporation of physical constraints, and limited interpretability. To address these issues, we propose additive fusion methodology using convolutional neural networks (CNNs) to enhance the representational capacity of optical flow features, improve the accuracy of predictions, and simultaneously reduce memory requirements. We specifically introduce a novel stacked fusion module, termed the Continue-ADD-Block. This module effectively consolidates multi-scale information through the inclusion of a Conv7 layer and the application of consecutive downsampling for the fusion process. This integration process strengthens the model's capacity to cope with complex scenes characterized by multi-scale motion. Empirical evaluations conducted on the Flying Chairs, KITTI Flow 2015, and MPI-Sintel (both clean and final versions) datasets, demonstrate that the Continue-ADD-BLOCK achieves superior accuracy (expressed as a lower maximal average End-Point Error (EPE)) in complex scenarios. Critically, this performance gain is achieved while simultaneously reducing resource consumption and the memory footprint. Specifically, the approach demonstrated accuracy improvements of 2.86% on Flying Chairs, 0.70% on KITTI Flow 2015, and 7.40% and 2.43% on the MPI-Sintel (clean and final) datasets, respectively. These findings highlight the proposed method's enhanced robustness under challenging circumstances, thus offering a novel and effective approach for optical flow estimation.

Key words: optical flow estimation; hierarchical strategy; multi-scale estimation; additive fusion; convolutional neural network

收稿日期:2024-11-28

修回日期:2025-03-28

基金项目:陕西省教育科学研究计划项目(21JK0684);工信部基金优选项目(MJZ1-8N22)

作者简介:王雅妮(2002-),女,硕士研究生,研究方向为视觉导航;通信作者:翟正军(1965-),男,教授,研究方向为航空电子与机载计算机技术、嵌入式智能系统、视觉导航。

0 引言

光流法因反映物体或相机运动,在相机位姿估计中备受关注。作为二维矢量场,光流描述像素运动,应用于运动感知、估计、3D 重建等领域。深度学习前,研究集中于手工特征和数学模型,如 Horn 和 Schunck 的变分理论算法及 Lucas 和 Kanade 的局部平滑约束算法,但处理噪声和模糊时鲁棒性有限。特征点匹配、块匹配、变分方法和基于滤波器的光流估计方法各有优劣。块匹配计算简单鲁棒,但成本高、对速度敏感、易受光照影响。变分方法优化光流场,但需预定义信息且计算量大。滤波器方法快速鲁棒,但精度稍低。这些方法依赖图像亮度恒定和局部平滑,为深度学习光流估计奠定基础。

近年来,深度学习技术的兴起为光流估计带来了新的机遇。FlowNet 网络率先应用卷积神经网络,利用网络学习图像特征与光流之间的映射关系,但处理大尺度运动和遮挡有局限。随后,RAFT 网络引入 Transformer 架构和注意力机制,擅长处理大尺度运动和遮挡^[1]。此外,PWC-Net 网络利用金字塔结构和光流卷积层进一步提高光流估计精度,有效地处理复杂场景^[2]。为了满足移动设备和资源受限场景的需求,LiteFlowNet 网络应运而生。该网络采用轻量化设计,显著降低了计算成本,同时保持了一定的精度水平,使其更适合移动设备和资源有限的环境^[3]。

为增强模型泛化能力,研究者进一步渴望探索无监督光流估计方法。其中 DDFlow 网络^[4]和 ARFlow^[5]网络均不同程度提高了精度,但计算复杂度增加,而 SimFlow^[6]网络能够平衡模型的速度与精度,但对极端复杂运动模式性能受限。此外,光流估计与其他任务结合进行多任务学习,如与深度估计、分割等结合,进一步提升性能。

总之,基于深度学习的光流估计技术正不断发展,未来有望在效率、精度和鲁棒性方面取得更大突破,为光流估计领域注入新的活力。

2021 年后,深度学习光流估计研究较少,且存在场景敏感、资源需求高等问题。深入探索场景敏感性有重要价值,该文对此做出如下贡献:

(1)提出了一种基于卷积神经网络的新型光流估计算法。该算法采用端到端的训练方式,以两幅连续图像作为输入,直接输出光流场,无需手动设计特征或复杂计算。

(2)为了增强光流特征表达,避免预测特征的奇异性,提出了一种分层策略,采用连续加法块(Continue-ADD-Block)连续顺序地融合预测的光流特征。该模型可以显著提升预测的光流特征。

1 相关工作

光流估计是计算机视觉的关键技术,其性能依赖于光流特征质量。增强光流特征能提升运动估计的准确性和鲁棒性,利于视频分析、目标跟踪和自动驾驶等领域^[7]。该文采用多尺度融合技术和分层策略,增强网络架构,提高模型表示能力。

1.1 分层策略

分层策略将复杂问题逐层解决,广泛应用于各领域。该模型遵循从低级到高级、具体到抽象的信息处理流程,实现网络分层设计^[8]。网络从输入层起,经卷积层逐层提取图像多尺度特征,由精细至粗糙。随后,池化层减少特征冗余,下采样压缩特征图^[9]。为恢复空间分辨率,引入反卷积层上采样,扩大特征图尺寸。最终,预测层整合各层信息,输出精确光流预测。分层设计确保网络逐步深入理解图像,每层承担特定任务,共同构成高效准确的图像光流预测网络,实现从像素级特征提取到全局光流估计的紧密相连与相互依赖。

该文设计分层特征提取框架,利用卷积层和隐式池化操作捕获图像多尺度特征。卷积层参数精心设置,定义感受野,影响特征图分辨率和抽象层次。通过设置步长,间接实现池化效果,增强模型泛化能力。每层设预测层,独立运行,基于当前层特征进行光流预测^[10]。特征图分辨率随网络深度降低,特征抽象层次加深,形成特征提取金字塔结构。该分层表示使网络同时捕获图像细节和全局信息,为光流估计提供丰富多尺度特征基础。

1.2 多尺度光流估计

在分层结构的基础上引入叠加融合机制,将高层光流预测上采样与低层特征结合,辅助下层预测,保留分层优点并利用跨尺度信息互补^[11]。多尺度光流估计处理大尺度运动和细节丢失,提高准确性和鲁棒性,适应不同尺度运动^[12]。该文采用的叠加融合方法包括插值方法和加权平均方法。Continue-ADD-Block 的创新之处在于对新添加的卷积层的输出进行上采样,即最高分辨率的光流场 flow7,使用插值方法使其与 flow6、flow5、flow4、flow3 和 flow2 的尺寸匹配,但是不会单独将最高分辨率的光流估计结果与较低分辨率的光流结果叠加融合。相反,该文从最高分辨率开始,依次与相应较低分辨率的光流结果融合,以满足不同的分辨率。Continue-ADD-Block 的过程可以用公式 1 表示。

$$\begin{aligned} \text{flow}_i &= \text{flow}_i + \text{interpolate}(\text{flow}_{\{i+1\}}, \text{size} = \\ &(\text{flow}_i.\text{shape}[2], \text{flow}_i.\text{shape}[3]), \\ &\text{mode} = \text{'bilinear'}, \text{align_corners} = \text{False}) \end{aligned} \quad (1)$$

在进行上采样后,Continue-ADD-Block 使用加权

平均法叠加融合特征。通过结合逐层预测和适度融合,该文构建了一个高效且鲁棒的光流估计模型,它将不同尺度的光流信息视为同等重要,连续叠加融合它们以获得整合了不同尺度信息的最终结果。

2 网络结构

为了使预测得到的光流更加准确,采用端到端的训练方式,其输入为图像对和真实光流值,通过文中模型结构直接预测输出光流场。模型结构包含下采样、上采样、叠加融合三个部分^[13]。下采样主要由卷积层构成,上采样主要由反卷积层等其他层构成,最后在输出光流特征之前使用叠加融合模块,模型架构如图 1 所示。在 FlowNetS 架构基础上提出的叠加融合模块 Continue-ADD-Block,旨在提升对复杂敏感场景的鲁棒性,并提高光流估计的精度,其模型结构如图 2 所示。

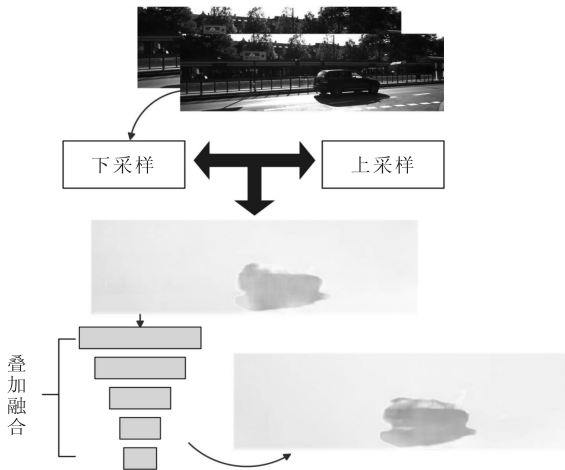


图 1 FlowNetS with Continue-ADD-Block 模型原理

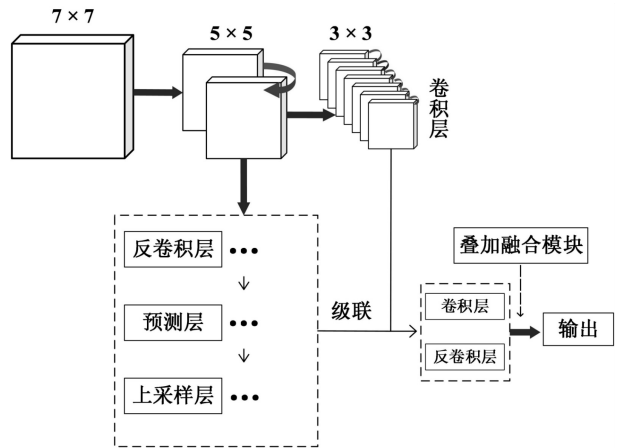


图 2 FlowNetS with Continue-ADD-Block 模型结构

针对传统 FlowNetS 光流估计网络在处理复杂场景和多尺度运动方面的局限性,在原模型的基础上提出了连续且具有层次化的叠加融合模块 Continue-ADD-Block。该模块在原有结构的基础上,新增了 Conv7 层,并在反卷积阶段采用了一种全新的多尺度融合策略。传统的 FlowNetS 网络在 Conv6 层之后不再进行下采样,直接进行光流预测,这使得它在处理大尺度运动和细节信息方面存在一定的局限性。该文提出的 Continue-ADD-Block 新增的 Conv7 层能够提取更深层的特征,从而更好地捕捉全局运动信息。此外,为了更加有效地融合不同尺度上的光流信息,创新性地引入了一种分层融合策略,旨在克服原方法可能导致的特征表达单一化问题,使得平均精度达不到预期效果。此模块不仅强调了特征在时间序列上的连续性捕捉,还通过逐步添加与集成不同层次的特征块实现了对复杂动态信息的多维度解析与融合,其连续性的过程原理见图 3。

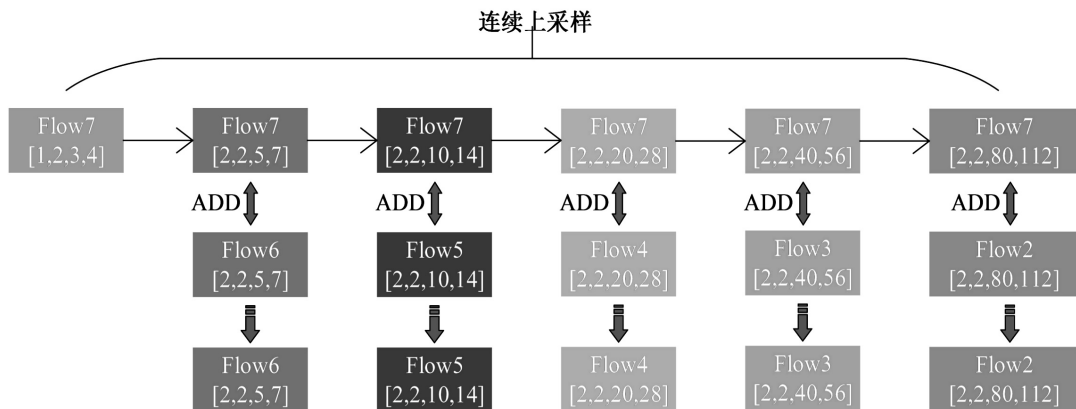


图 3 Continue-ADD-BLOCK 原理

Continue-ADD-Block 的 flow7 在每次上采样后,都会将上采样后的 flow7 赋值给 flow7 本身,也就是说,flow7 的值会被不断更新,它不再是原始的最高分辨率光流结果,而是经过多次上采样后的结果。在进行叠加融合的时候,同样是将上采样后的 flow7 叠加

到当前尺度的结果上,但此时 flow7 已经经过了多次上采样才使得其分辨率与当前尺度的结果一致,可以通过不断地叠加,flow7 能够累积来自不同尺度的信息,可能得到更精确的光流结果;同时由于每次上采样后都会更新 flow7,因此在内存中只需要保存一个

flow7,为提出的方法减轻了资源消耗和内存体积。

3 数据集

数据集的选择对光流估计模型的性能至关重要。该文选用了以下三种广泛认可的公开数据集: Flying Chairs、MPI-Sintel (clean 和 final) 和 KITTI Flow 2015。Flying Chairs 专注于合成图像中的运动物体,提供高质量的 ground truth 数据,用于研究算法的准确性^[14]。

表 1 数据集的比较

数据集	数据种类	数据来源	场景种类	复杂程度	训练集帧对数	测试集帧对数	分辨率
Flying Chairs	合成图像	人工合成	单一物体运动	相对简单	22 872	暂未公开	1 280×720
MPI-Sintel	渲染场景	计算机图形学渲染	复杂场景,多种运动	较为复杂	clean:1 064 final:1 048	clean:376 final:376	1 024×436 (训练集)
KITTI	真实世界图像	汽车采集	道路场景,多种运动	最为复杂	200	200 (无真值)	1 242×375 (训练集)

4 实验结果分析

4.1 实验设计

为了验证创新模块 Continue-ADD-Block 的性能,该文在光流估计领域开源的、权威的数据集 Flying Chairs、MPI-Sintel (clean 和 final) 和 KITTI Flow 2015 数据集上进行了实验,并使用端点误差 (End Point Error, EPE) 进行评估。

实验时,除 Flying Chairs 数据集暂无公开测试集,该文用到的其余数据集的数据分为训练集 (80%) 和验证集 (20%),测试集是单独的一部分,其数据量见表 1。KITTI Flow 2015 数据因稀疏性采用简单增强策略,进行随机裁剪 (320×448)、垂直翻转和水平翻转;其他数据集则全面增强,包括随机平移、旋转。

在光流估计任务中,为增强对噪声和异常值的抵抗力,EPE 作为主要评估指标,用来衡量估计的光流矢量与真实光流矢量之间的差异程度。简单来说,它衡量的是预测的光流矢量末端和真实光流矢量末端之间的欧几里得距离,且该指标同时适用于以上提到的三种数据集^[15]。

$$E_{PE} = \sqrt{(u_{est} - u_{gt})^2 + (v_{est} - v_{gt})^2} \quad (2)$$

其中, E_{PE} 代表端点误差, u_{est} 代表估计水平光流, u_{gt} 代表真实水平光流, v_{est} 代表估计垂直光流, v_{gt} 代表真实垂直光流。计算得到的 EPE 值越小,其估计得到的光流误差越小,即模型精度越高。

4.2 实验过程

基于实验设计,在实验过程中共分为四组实验,分别使用 Flying Chairs、KITTI Flow 2015、MPI-Sintel clean 和 MPI-Sintel final 数据集。每组均在原始模型和带有 Continue-ADD-Block 的方法上迭代 300 次进

MPI-Sintel 则使用高质量的渲染场景,提供更复杂的运动和照明条件,考验算法对真实世界场景的鲁棒性,其中 clean 部分在图像渲染时未添加任何后期处理效果,final 部分添加包括运动模糊、景深和纹理等后期处理效果。KITTI 则侧重于真实世界场景,提供从汽车上采集的图像和 ground truth 数据,用于评估算法在实际应用中的性能,为自动驾驶等领域提供重要参考。三个数据集的详情对比如表 1 所示。

行两次实验,记录在测试集上得到的 EPE 的最大值、最小值和平均值。对这四个组在带有该模块的新方法上训练得到的模型进行命名,加以区分,具体模型名称见表 2。

表 2 模型名称

名称	FlowNetS	with ADD-Block
Flying Chairs	F-FlowNetS	F-Continue-ADD-Block
Sintel clean	Sc-FlowNetS	Sc-Continue-ADD-Block
Sintel final	Sf-FlowNetS	Sf-Continue-ADD-Block
KITTI Flow2015	K-FlowNetS	K-Continue-ADD-Block

注意:所有这些模型都是在相同数量的训练迭代、参数条件和硬件设备下训练的。

为验证模块 Continue-ADD-Block 对光流估计在复杂敏感场景的识别能力提高的效果,进行消融实验并分析结果。将刚才四个组得到的结果,该文单独对比记录每个小组中数据集在原始方法和带有 Continue-ADD-Block 的模型方法上训练测试得到的 EPE 变化值,通过比较同一组 (同一个数据集) 上训练后测试得到的 EPE 最大值在原始模型方法中始终高于该文提出的方法,即验证模型方法添加 Continue-ADD-Block 后,其对复杂环境光流估计得到的 EPE 值降低,误差降低,即精度提高。记录并保存不同小组,数据集在测试后得到的光流估计效果图,比较其光流估计效果。

根据以上实验过程,总结得出结论,并分析实验结果。

4.3 实验结果

记录每组 2 次实验得到的测试集 EPE 的最大值、平均值和最小值的变化。通过以上四组的对比实验,带有 Continue-ADD-Block 的模块在最差的情况下估

计得到的光流结果误差下降,其精度得到提升,在 Flying Chairs 数据集、MPI-Sintel clean 数据集和 MPI-Sintel final 数据集上,最坏的情况下即测试得到的 EPE 最大值分别降低 5.90%、4.97%、1.90%,同时其平均端点误差在四组实验上也下降了,依次降低 30.34%、21.93%、4.80%、3.45%,其端点误差 EPE 最大值、平均值和最小值的对比如图 4 所示。

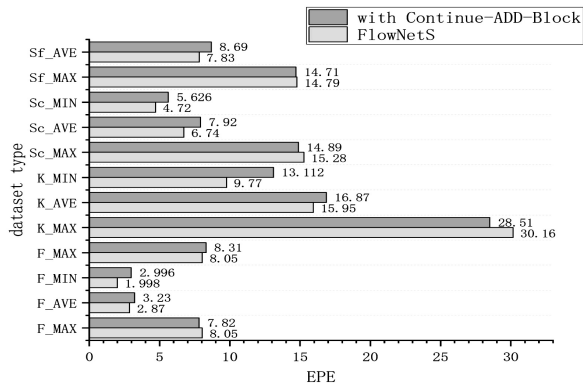


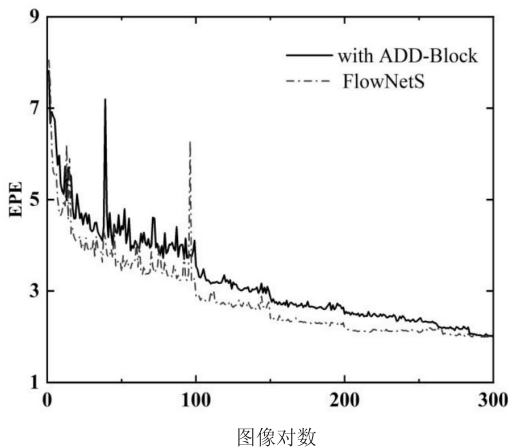
图 4 EPE 的最大值、平均值和最小值的变化对比
在图 4 中, F 代表 Flying Chairs 数据集、K 代表

KITTI Flow 2015 数据集、Sc 代表 MPI-Sintel clean 数据集、Sf 代表 MPI-Sintel final 数据集。柱状图可以很好地表示不同模块对相同数据评价标准的对比性。综合以上数据,认为 Continue-ADD-Block 在光流估计任务中在面对复杂环境的精度效果上好于 Straight-ADD-Block。

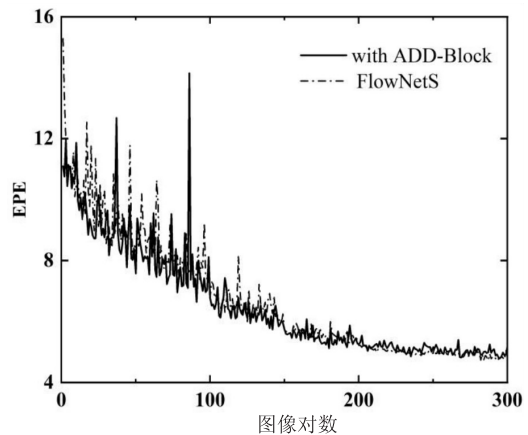
4.4 消融分析

接下来将对 Continue-ADD-Block 模块进行消融分析,图 5 展示了 Flying Chairs 数据集、KITTI Flow 2015 数据集和 MPI-Sintel(clean 和 final)数据集分别在 FlowNetS 模型和带有 Continue-ADD-Block 叠加融合模块的模型结构中训练 300 次得到的端点误差变化曲线。

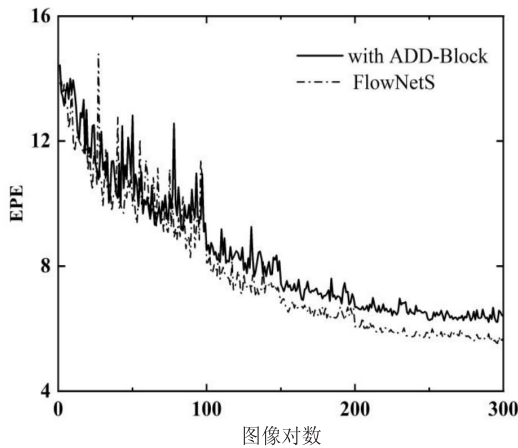
据图 5 可以观察到带有 Continue-ADD-Block 叠加融合模块的模型结构下降速度在部分图像中不如 FlowNetS 模型快,但是其最大值却低于 FlowNetS 模型的结果,端点误差最大值降低预示着该模型的鲁棒性有所提高,在后续进一步的改善中未来将取得更快的收敛速度。



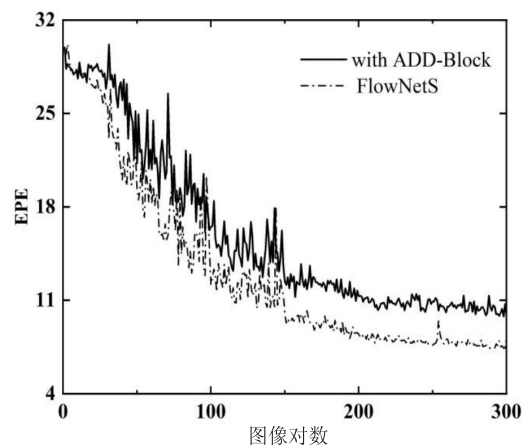
(a) F-FlowNetS 和 F-Continue-ADD-Block 对比



(b) Sc-FlowNetS 和 Sc-Continue-ADD-Block 对比



(c) Sf-FlowNetS 和 Sf-Continue-Add-Block 对比图



(d) K-FlowNetS 和 K-Continue-Add-Block 对比图

图 5 不同数据集上的消融实验 EPE 变化曲线

表 3 详细展示了两结构在不同数据集上训练得到的模型测试的端点误差。以下数据均在相同训练迭

代次数、参数条件、硬件设备下训练得到;其中 EPE max 代表最大的端点误差,EPE min 代表最小的端点

误差即最佳,EPE ave 代表平均端点误差。

表 3 消融实验结果

模型名称	EPE max	EPE ave	EPE min
F-FlowNetS	8.05	2.87	1.998
F-Continue-ADD-Block	7.82	3.25	1.992
Sc-FlowNetS	15.28	6.74	4.72
Sc-Continue-ADD-Block	14.15	7.54	5.38
Sf-FlowNetS	14.79	7.83	5.53
Sf-Continue-ADD-Block	14.43	8.39	6.21
K-FlowNetS	30.16	15.95	9.77
K-Continue-ADD-Block	29.95	13.17	7.35

以上数据中,在带有 Continue-ADD-Block 的模型结构上训练得到的模型测试的最大端点误差在 Flying Chairs 数据集、KITTI Flow 2015 数据集和 MPI-Sintel(clean 和 final)数据集中均小于原始结构模型得到的数据,并且在 Flying Chairs 数据集中精度提升 2.86%,在 KITTI Flow 2015 数据集中精度提升 0.70%,在 MPI-Sintel(clean 和 final)数据集中精度分别提升 7.40%、2.43%。由此可以得出,Continue-Add-Block 模块在该任务上取得较好的效果。

以下将展示 Flying Chairs 数据集、KITTI Flow 2015 数据集和 MPI-Sintel clean 数据集的部分图像对,并分别展示将其通过结果精度最高的两个模型结构,即 FlowNetS-F 和 ADD-Block-F,得到不同的光流估计结果图像。

(1)Flying Chairs 数据集部分样例在不同模型下的测试效果如图 6 所示,其中(a)是连续的图像对,(b)是 FlowNetS-F 模型测试得到的光流效果图,(c)是 ADD-Block-F 模型测试得到的光流效果图。可以直观地看到 ADD-Block-F 的模型效果更好。

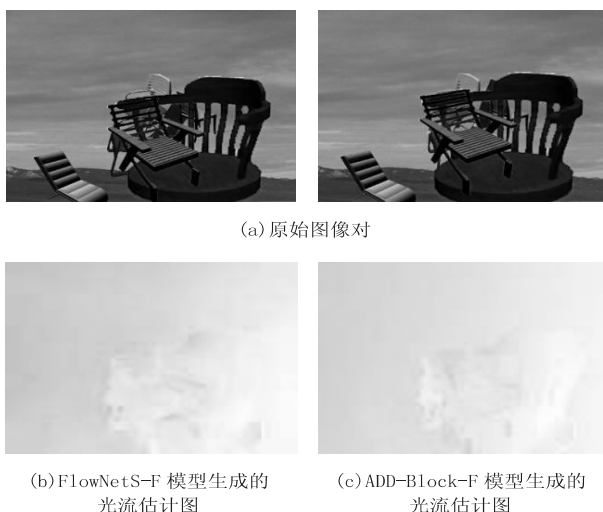


图 6 Flying Chairs 数据集样本测试效果

(2)KITTI Flow 2015 数据集在不同模型下的测试效果如图 7 所示。第一行从左至右的两张图是连续的图像对,第二行从左至右的第一张图是 FlowNetS-F 模型测试得到的光流效果图,第二张图是 ADD-Block-F 模型测试得到的光流效果图。可以直观地看到 ADD-Block-F 的模型效果更好。

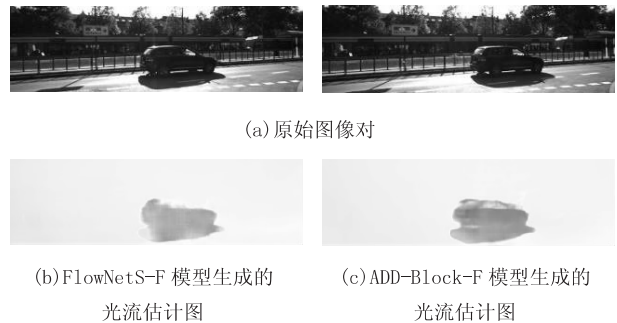


图 7 KITTI Flow 2015 数据集测试效果

(3)MPI-Sintel clean 数据集在不同模型下的测试效果如图 8 所示。第一行从左至右的两张图是连续的图像对,第二行从左至右的第一张图是 FlowNetS-F 模型测试得到的光流效果图,第二张图是 ADD-Block-F 模型测试得到的光流效果图。可以直观地看到 ADD-Block-F 的模型效果更好。

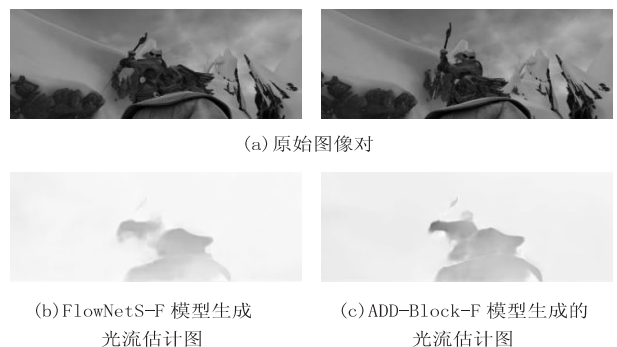


图 8 MPI-Sintel clean 数据集测试效果

5 结束语

该文提出一种基于深度学习的光流估计叠加融合法,用 Continue-ADD-Block 提取特征,精准预测且内存占用小。在四种不同的数据集中端点误差均下降,即精度都得到提升,在 Flying Chairs 数据集中精度提升 2.86%,在 KITTI Flow 2015 数据集中精度提升 0.70%,在 MPI-Sintel(clean 和 final)数据集中精度分别提升 7.40%、2.43%,对光流估计至关重要。但单目视觉局限,未来研究将探索双目视觉,以获取更丰富、精确的环境感知信息。

参考文献:

- [1] TEED Z, DENG J. Raft: recurrent all-pairs field transforms for optical flow[C]//Computer vision - ECCV 2020: 16th European conference. Glasgow: Springer, 2020: 402-419.

- [2] SUN D, YANG X, LIU M Y, et al. Pwc-net: CNNs for optical flow using pyramid, warping, and cost volume [C]//Proceedings of the IEEE conference on computer vision and pattern recognition. Salt Lake City: IEEE, 2018; 8934–8943.
- [3] HUI T W, LOY C C. Liteflownet3: resolving correspondence ambiguity for more accurate optical flow estimation [C]//Computer vision – ECCV 2020: 16th European conference. Glasgow: Springer, 2020; 169–184.
- [4] LIU P, KING I, LYU M R, et al. DdfLOW: learning optical flow with unlabeled data distillation [C]//Proceedings of the AAAI conference on artificial intelligence. [s. l.]: AAAI, 2019; 8770–8777.
- [5] LIU L, ZHANG J, HE R, et al. Learning by analogy: reliable supervision from transformations for unsupervised optical flow estimation [C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. Seattle: IEEE, 2020; 6489–6498.
- [6] IM W, KIM T K, YOON S E. Unsupervised learning of optical flow with deep feature similarity [C]//Computer vision – ECCV 2020: 16th European conference. Glasgow: Springer, 2020; 172–188.
- [7] ZHAI M, XIANG X, LV N, et al. Optical flow and scene flow estimation: a survey [J]. Pattern Recognition, 2021, 114; 107861.
- [8] SHIRI F M, PERUMAL T, MUSTAPHA N, et al. A comprehensive overview and comparative analysis on deep learning models; CNN, RNN, LSTM, GRU [J]. arXiv; 2305. 17473, 2023.
- [9] BIANCHI F M, LACHI V. The expressive power of pooling in graph neural networks [J]. Advances in Neural Information Processing Systems, 2023, 36; 71603–71618.
- [10] HUANG Z, LIU H, LV C. Gameformer: game – theoretic modeling and learning of transformer – based interactive prediction and planning for autonomous driving [C]//Proceedings of the IEEE/CVF international conference on computer vision. Paris: IEEE, 2023; 3903–3913.
- [11] XIANG X, ABDEIN R, LV N, et al. InvFlow: involution and multi – scale interaction for unsupervised learning of optical flow [J]. Pattern Recognition, 2024, 145; 109918.
- [12] WANG Z, CHEN Z, ZHANG C, et al. LCIF-Net: local criss – cross attention based optical flow method using multi – scale image features and feature pyramid [J]. Signal Processing: Image Communication, 2023, 112; 116921.
- [13] ZHAO R, XIONG R, DING Z, et al. MRDFlow: unsupervised optical flow estimation network with multi – scale recurrent decoder [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2021, 32(7); 4639–4652.
- [14] DOSOVITSKIY A, FISCHER P, ILG E, et al. FlowNet: learning optical flow with convolutional networks [C]//Proceedings of the IEEE international conference on computer vision. Santiago: IEEE, 2015; 2758–2766.
- [15] SAVIAN S, MORERIO P, DEL BUE A, et al. Towards equivariant optical flow estimation with deep learning [C]//Proceedings of the IEEE/CVF winter conference on applications of computer vision. Hawaii: IEEE, 2023; 5088–5097.