Real-World Evaluation of Optical Flow with Varying Lighting Conditions

Author: Zhuoyue Ge[1]
Supervisors: Sander Gielisse[1], Responsible Professor: Jan van Gemert[1]
[1]EEMCS, Delft University of Technology

Background

- Optical flow
  - movement of objects between consecutive frames in a video
  - wide applications, including autonomous driving and video surveillance

Figure 1. Optical flow visualization showing two frames of a sequence, ground-truth optical flow (color coded), and the color code to read the vector at each pixel.

Recent deep learning models have achieved strong benchmark results, however...

- Simulation to reality gap
  - trained on synthetic datasets (e.g., FlyingChairs[5])
  - lacks the complexity of real-world scenes. struggle in real-world scenarios
- Struggle with challenging lighting conditions such as:
  - Glare
  - Rapid lighting intensity change
  - Shadows

Reasearch Question

How well do optical flow models evaluated on synthetic datasets perform in real-world scenarios with varying lighting conditions?

Methodology

- Dataset Collection
  - Lighting conditions: Glare, moving shadows, light intensity, and outdoor shadows.
  - Static camera and objects in the scene, only lighting changes.
- Frame Selection
  - Manually select frame pairs using semi-automated tools.
  - Export in KITTI-compatible format.
- Model Evaluation
  - Models benchmarked: RAFT[1], GMFlow[2], SEA-RAFT[3], and FlowDiffuser[4].
  - Evaluation metrics:
    - End Point Error (EPE) – Euclidean distance between predicted and ground truth flow
    - F1-all score – percentage of outlier pixels (EPE > 3px and relative error > 5%)
- Performance Analysis

Results

Model	Glare		Moving Shadows		Lighting Intensity		Outdoor Shadows	
	EPE	F1-all	EPE	F1-all	EPE	F1-all	EPE	F1-all
RAFT	0.31	0.13	13.35	14.43	12.24	16.35	1.94	3.01
GMFlow	3.03	5.36	7.15	34.90	4.35	28.67	2.57	7.21
SEA-RAFT	0.26	0.58	4.09	14.94	6.84	13.16	1.53	3.34
FlowDiffuser	0.37	0.39	28.56	42.77	22.89	30.92	3.92	6.16

Table 1. Mean EPE and F1-all scores across four lighting conditions for each model.

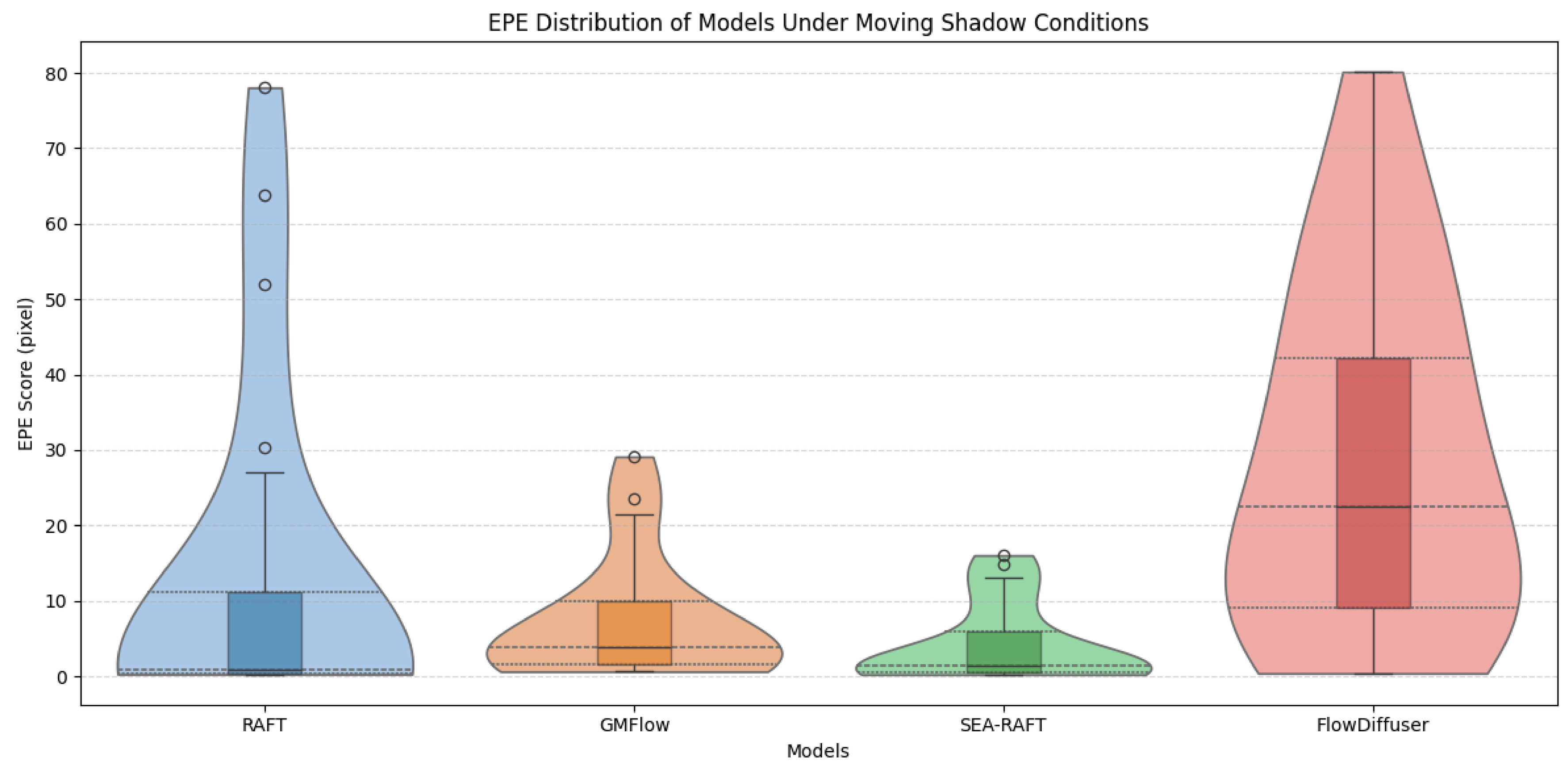

Figure 3. EPE distribution of all four models under the moving shadow condition.

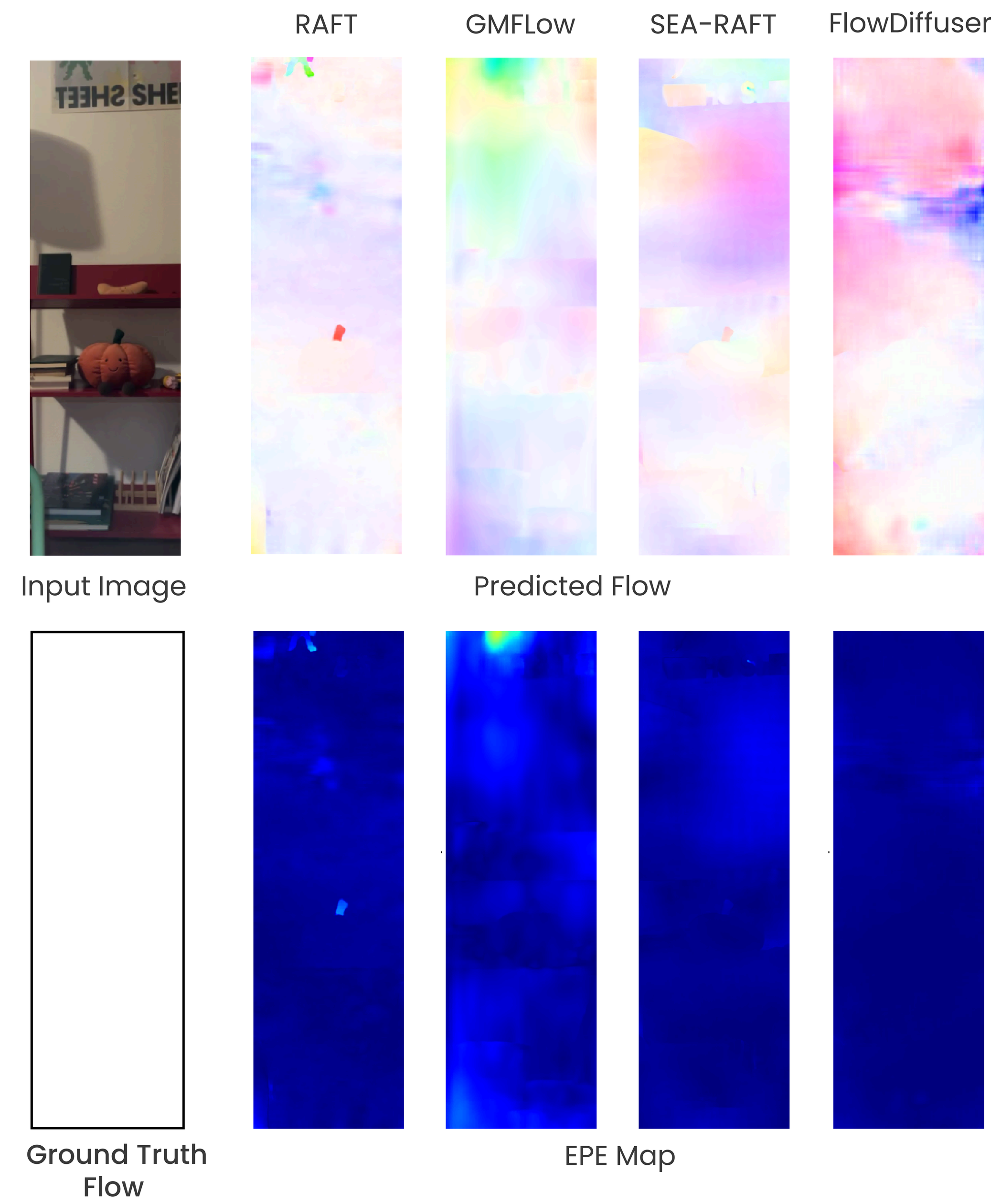

Figure 4. Example optical flow predictions from each model on a sample from the Glare dataset.

Conclusion

- Regional lighting changes can induce incorrect optical flow estimates across the entire image.
- SEA-RAFT is the most robust among the four models, likely due to its diverse training data, but it still struggles under complex lighting.
- Significant variation in EPE is observed within the same scene and using the same model.
- Results expose the limitations of current architectures in handling real-world lighting variability.

Future Work

- Dynamic Scene Simulation
  - Add camera and object movement to better align with real-world scenes.
- Data Expansion
  - Use a larger number of frames and more diverse lighting conditions.
- Model Adaptation
  - Improve current models under challenging lighting conditions (e.g., learn lighting invariant features).

Figure 2. Example collected frame pairs under different lighting variations.

Reference

- [1] Z. Teed and J. Deng, "RAFT: Recurrent all-pairs field transforms for optical flow," in Proc. Eur. Conf. Comput. Vis. (ECCV), 2020, pp. 402–419.
- [2] H. Xu, J. Zhang, J. Cai, H. Rezatofighi, and D. Tao, "GMFlow: Learning optical flow via global matching," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), 2022, pp. 8111–8120.
- [3] Y. Wang, L. Lipson, and J. Deng, "SEA-RAFT: Simple, efficient, accurate RAFT for optical flow," in Proc. Eur. Conf. Comput. Vis. (ECCV), 2024, pp. 36–54.
- [4] L. Luo, Y. Wang, Y. Wang, and J. Deng, "FlowDiffuser: Advancing optical flow estimation with diffusion models," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), 2024, pp. 12345–12354.
- [5] A. Dosovitskiy et al., "FlowNet: Learning optical flow with convolutional networks," in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), 2015, pp. 2758–2766.

z.ge-3@student.tudelft.nl