DistractFlow: Improving Optical Flow Estimation via Realistic Distractions and Pseudo-Labeling (Supplementary Material)

1. Implementation and Training Details

We use the official codes for RAFT [4],[1] GMA [3],[2] and FlowFormer [1].[3] We use 2 NVIDIA A100 GPUs for all training experiments. We follow RAFT and GMA learning parameters except for the batch size. Specifically, we use a batch size of 8 for RAFT and GMA, when reproducing the baseline numbers and the numbers with DistractFlow. For FlowFormer, we follow their original setting including the batch size to reproduce this baseline. When applying DistractFlow, we use a batch size of 2 for FlowFormer due to GPU memory constraint.

Table 1. EPE and coverage at varying confidence thresholds(τ). Test results obtained using RAFT on Sintel (final) dataset (trained on FlyingChair+FlyingThings3D).

Threshold (τ)	0.0	0.37	0.7	0.8	0.90	0.95	1.0
EPE	2.73	1.12	0.93	0.85	0.74	0.66	0.28
Coverage	1.0	0.89	0.87	0.85	0.81	0.74	0.18×10^{-4}

2. Calibration

Table 1 shows that when choosing pixels with a higher confidence threshold, the EPE of the selected pixels decreases while the coverage (i.e., percentage of selected pixels) also decreases. This verifies that our confidence measure of Eq. (5) in the paper properly captures the correctness of the prediction. Therefore, when applying a confidence threshold, we obtain a set of more accurately predicted pixels, rather than simply a smaller subset.

3. Sintel Clean Results

As can be seen in the table 2, our proposed DistractFlow also provides consistent improvements for state-of-the-art models on Sintel Clean (train), despite that this is an easier split with low-noise, synthetic images.

[1] https://github.com/princeton-vl/RAFT

[2] https://github.com/zacjiang/GMA

[3] https://github.com/drinkingcoder/FlowFormer-Official

Table 2. Optical flow estimation results on Sintel (train/clean) dataset. The training is under the same as semi-supervised setting of Table 1 in the main paper.

Method	Model	Sintel (Clean)
Supervised	RAFT [4]	1.42
FlowSupervisor [2]		1.30
DistractFlow (Our)		**1.25**
Supervised	GMA [3]	1.35
DistractFlow (Our)		**1.22**
Supervised	FlowFormer [1]	0.95
DistractFlow (Our)		**0.90**

References

[1] Zhaoyang Huang, Xiaoyu Shi, Chao Zhang, Qiang Wang, Ka Chun Cheung, Hongwei Qin, Jifeng Dai, and Hongsheng Li. Flowformer: A transformer architecture for optical flow. In *Proceedings of the European Conference on Computer Vision*, 2022. 1

[2] Woobin Im, Sebin Lee, and Sung-Eui Yoon. Semi-supervised learning of optical flow by flow supervisor. In *Proceedings of the European Conference on Computer Vision*, 2022. 1

[3] Shihao Jiang, Dylan Campbell, Yao Lu, Hongdong Li, and Richard Hartley. Learning to estimate hidden motions with global motion aggregation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 9772–9781, 2021. 1

[4] Zachary Teed and Jia Deng. Raft: Recurrent all-pairs field transforms for optical flow. In *Proceedings of the European Conference on Computer Vision*, pages 402–419. Springer, 2020. 1