University of Stuttgart Germany

imprs-is

SFB-TRR 161 Quantitative Methods for Visual Computing

DFG Deutsche Forschungsgemeinschaft

Reviving Unsupervised Optical Flow: Concept Reevaluation, Multi-Scale Advances and Full Open-Source Release

Azin Jahedi, Marc Rivinius, Noah Berenguel Senn and Andrés Bruhn

WACV 2026

University of Stuttgart
Computer Vision Group

Presenter: Azin Jahedi
March 2026

WACV 2026
Tucson, Arizona • Mar 6 - 10, 2026

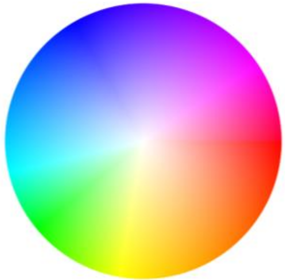
What is Optical Flow?

- Given: Two consecutive images of an image sequence

image at time t

image at time t+1

optical flow



Middlebury color coding
(Baker et al., IJCV 2011)

- Wanted: Displacement field between the two images → optical flow

Motivation

- Optical flow has many applications in different domains:
  - Autonomous driving
  - Robot navigation
  - Video compression
  - Action recognition
- Supervised methods typically lack generalization on unseen domain.
- Collecting ground truth optical flow is cumbersome.
- Unsupervised methods do not require ground truth optical flow.
- Most approaches are based on older architectural backbones (PWC-Net).
- The current SOTA requires extensive resources and has proven challenging to reproduce.
- Progress in the field has mainly plateaued.

An open-source SOTA unsupervised optical flow method would be useful.

Motivation

- In this paper
    - We reinvestigate unsupervised training strategies.
        - Yielding effective training settings
    - We present two SOTA unsupervised methods of
        - Sun-RAFT (simple unsupervised RAFT)
            - Based on RAFT (Teed and Deng, ECCV 2020)
        - Muun-RAFT (multi-scale unsupervised RAFT)
            - Based on MS-RAFT (Jahedi et al., ICIP 2022)
    - We fully open-source our PyTorch-based code.

Unsupervised Learning of Optical Flow

Preliminaries

Unsupervised optical flow

Preliminaries

- Photometric Loss

census-transformed
backward-warped
second image with the
current forward flow

occlusion mask

$$\mathcal{L}_{ph,fw} = \frac{1}{n} \sum O_{12} \odot \rho(I_1, \tilde{I}_2^{f_{12}})$$

sub-quadratic penalization of difference

  - ρ computes the soft Hamming distance of census transformed images [1].

  - Evaluates visual similarity of two potentially corresponding pixels, based on the current flow.

1. Meister et al. Unsupervised learning of optical flow with a bidirectional census loss, AAAI 2018

Unsupervised optical flow

Preliminaries

- Edge-aware image-driven l_{th}-order smoothness loss [1]

$$\mathcal{L}_{sm_{o_l},fw} = \frac{1}{n}\sum \exp(-\frac{\alpha}{3}\sum_{c=1}^{3}|\frac{\partial I_{1,c}}{\partial x}|) \odot |\frac{\partial^l f_{12}}{\partial x^l}| + \exp(-\frac{\alpha}{3}\sum_{c=1}^{3}|\frac{\partial I_{1,c}}{\partial y}|) \odot |\frac{\partial^l f_{12}}{\partial y^l}|$$

  - Used as regularization, reducing fluctuations in the flow field, where no edges in the reference image are detected.

1. Tomasi and Manduchi, Bilateral filtering for gray and color images, ICCV 1998

Unsupervised optical flow

Preliminaries

- Self-Supervision via augmentation [1,2]
  - Apply photometric, occlusion and diverse geometric augmentations.
    - Compute the flow from augmented images -> Student flow
  - Compute the flow from unaugmented (clean) images.
    - Transform the output flow accordingly -> Teacher flow

forward-backward
consistency mask
of the student flow

$$\mathcal{L}_{self,fw} = \frac{1}{n}\sum \|\mathcal{S}(f_{12}^T) - f_{12}^S\| \odot M_{12}^T \odot (1 - M_{12}^S)$$

gradient stopping

forward-backward
consistency mask
of the teacher flow

1. Liu et al., Learning by analogy: reliable supervision from transformations for unsupervised optical flow estimation CVPR 2020
2. Liu et al. DDFlow: learning optical flow with unlabeled data distillation, AAAI 2019

Unsupervised optical flow

Total Loss

- Weighted sum of all forward and backward losses

Weight of specific loss term

$$\mathcal{L} = \sum_{s=1}^{N_{\text{scales}}} \sum_{i=1}^{T_s} \sum_{t \in \{ph, sm, self\}} \omega_t \, \gamma_t^{s,i} \, \mathcal{L}_t^{s,i}$$

Exponential weight (increasing over iterations)

- N=1 for Sun-RAFT, N=3 for Muun-RAFT

Reinvestigation of Previously Advised Concepts

Self-Supervised Learning

Settings

- Self-Supervision settings
  - Photometric, occlusions and diverse geometric augmentation: random flip, scale & crop
    - Unlike DDFlow [1] and SMURF [2]: fixed zoom and crop

1. Liu et al. DDFlow: learning optical flow with unlabeled data distillation, AAAI 2019
2. Stone et al., Smurf: self-teaching multi-frame unsupervised RAFT with full-image warping, CVPR 2021

Self-Supervised Learning

Settings

- Self-Supervision settings
  - Photometric, occlusions and diverse geometric augmentation: random flip, scale & crop
    - Unlike DDFlow [1] and SMURF [2]: fixed zoom and crop
  - Forward-Backward masking
    - Apply self-supervision where teacher is confident but the student is not.
  - Much more weight to teach early estimates
    - (γ=0.95, ω=1.2) instead of (γ=0.8, ω=0.3) in SMURF. First iteration : 6.82 vs. 0.026!

1. Liu et al. DDFlow: learning optical flow with unlabeled data distillation, AAAI 2019
2. Stone et al., Smurf: self-teaching multi-frame unsupervised RAFT with full-image warping, CVPR 2021

Self-Supervised Learning

Settings

- Self-Supervision settings
  - Photometric, occlusions and diverse geometric augmentation: random flip, scale & crop
    - Unlike DDFlow [1] and SMURF [2]: fixed zoom and crop
  - Forward-Backward masking
    - Apply self-supervision where teacher is confident but the student is not.
  - Much more weight to teach early estimates
    - (γ=0.95, ω=1.2) instead of (γ=0.8, ω=0.3) in SMURF. First iteration : 6.82 vs. 0.026!
- We do not perform full-image warping [2] for computing the photometric loss
  - Not helpful for us.
- Training Schedule and weight initialization

1. Liu et al. DDFlow: learning optical flow with unlabeled data distillation, AAAI 2019
2. Stone et al., Smurf: self-teaching multi-frame unsupervised RAFT with full-image warping, CVPR 2021

Reviving Unsupervised Optical Flow

Outcome

- Sun-RAFT: Simple unsupervised RAFT
  - A two-frame alternative to SMURF, but
    - Simpler
    - Much less resource-intensive (2 GPUs vs. 8 GPUs)
    - New SOTA for two-frame unsupervised optical flow

- Muun-RAFT: Multi-scale unsupervised RAFT
  - Based on MS-RAFT (Jahedi et al., ICIP 2022; Jahedi et al., IJCV 2024)
  - Multi-scale version of Sun-RAFT
  - Exploits a novel context-based upsampling scheme

MS-RAFT (Jahedi et al., ICIP 2022)

Architecture

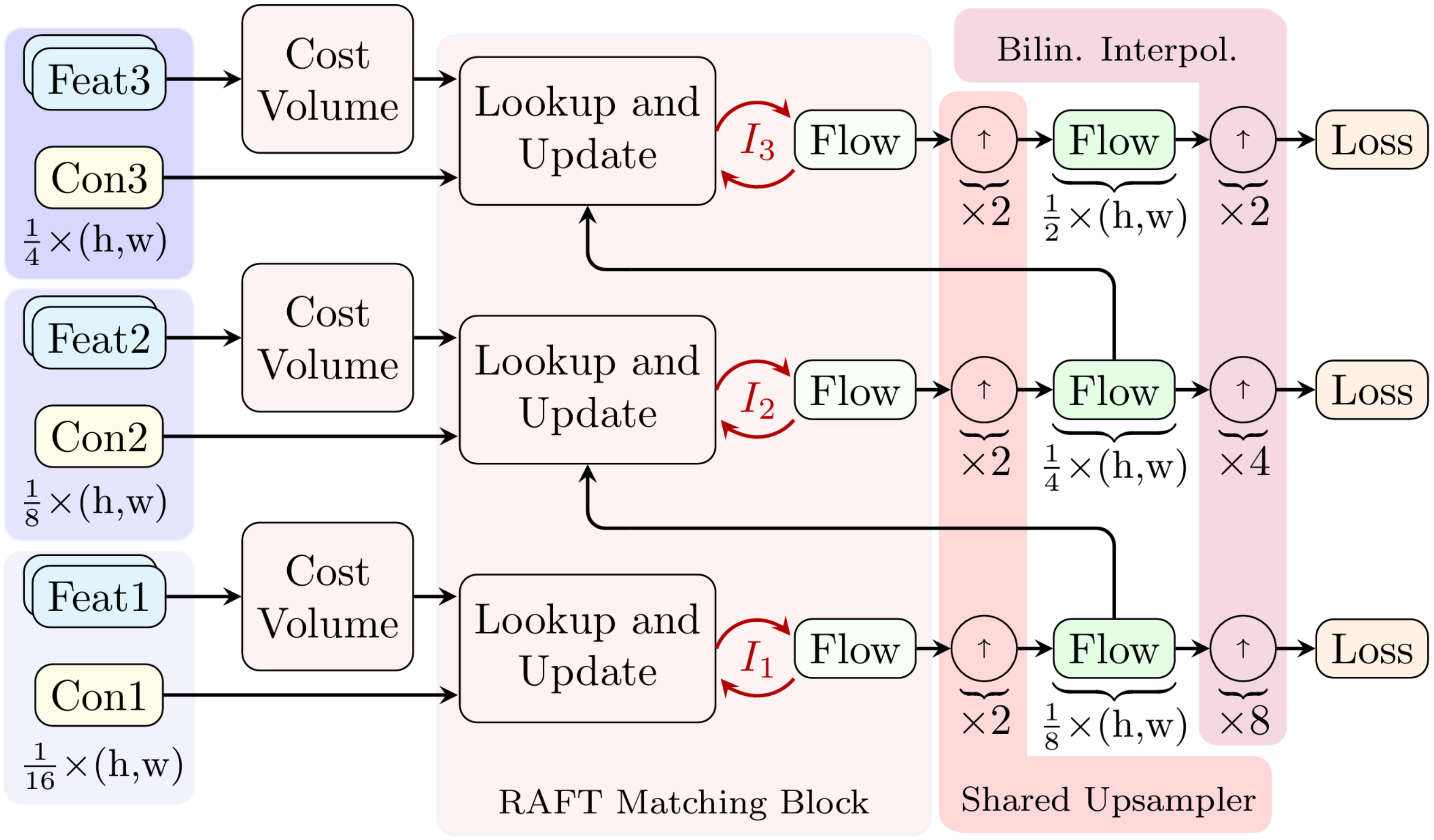


Figure from Jahedi et al., ICIP 2022

Reviving Unsupervised Optical Flow

Performance of the multi-scale vs. single-scale?

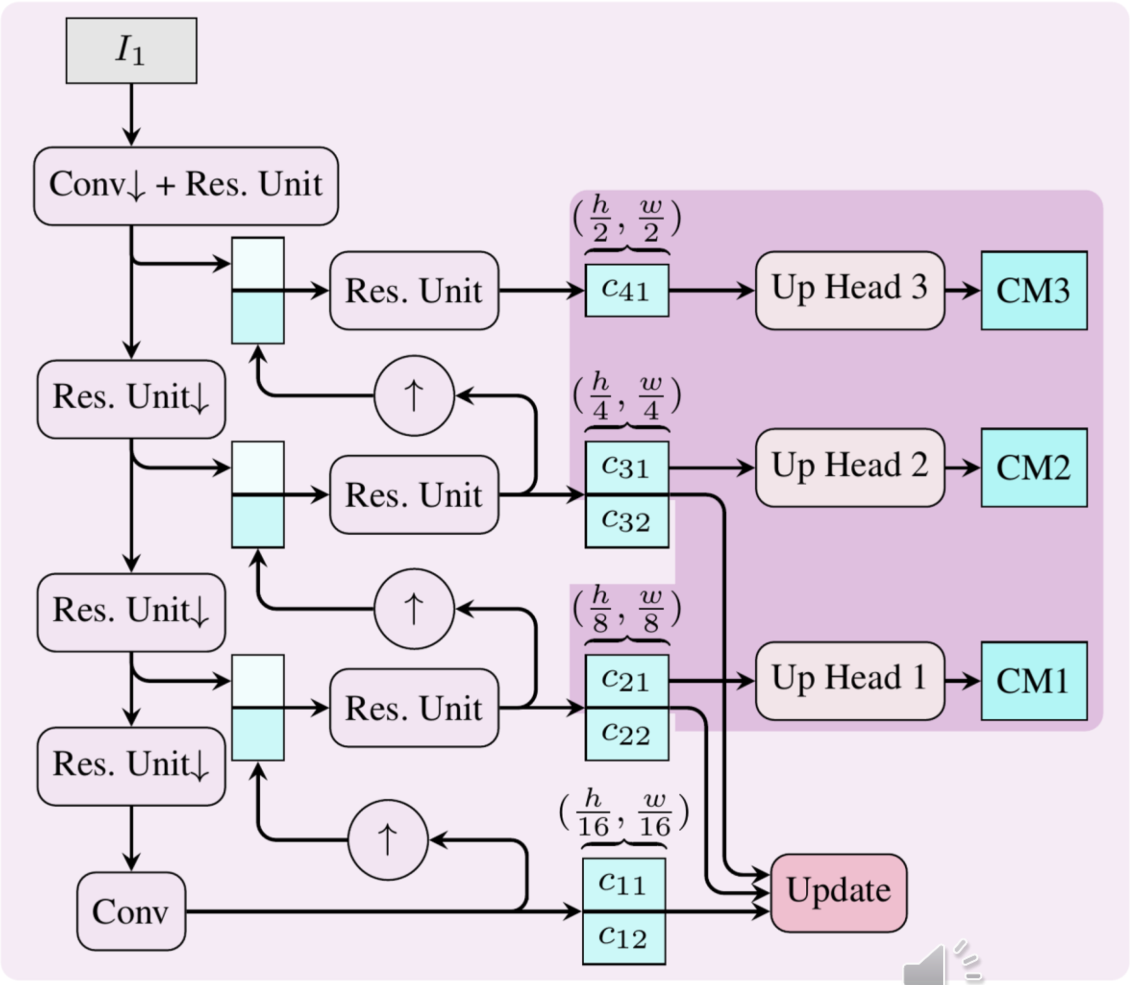
- Interestingly not more accurate than single-scale
  if MS-RAFT is directly integrated.
  - Unlike the supervised case.

- Remedy?
  - Compute the loss from high-quality intermediate and final outputs!
  - Novel architectural sub-network:
    - Gradual context-based upsampling!

Reviving Unsupervised Optical Flow

Gradual Context-based Upsampling

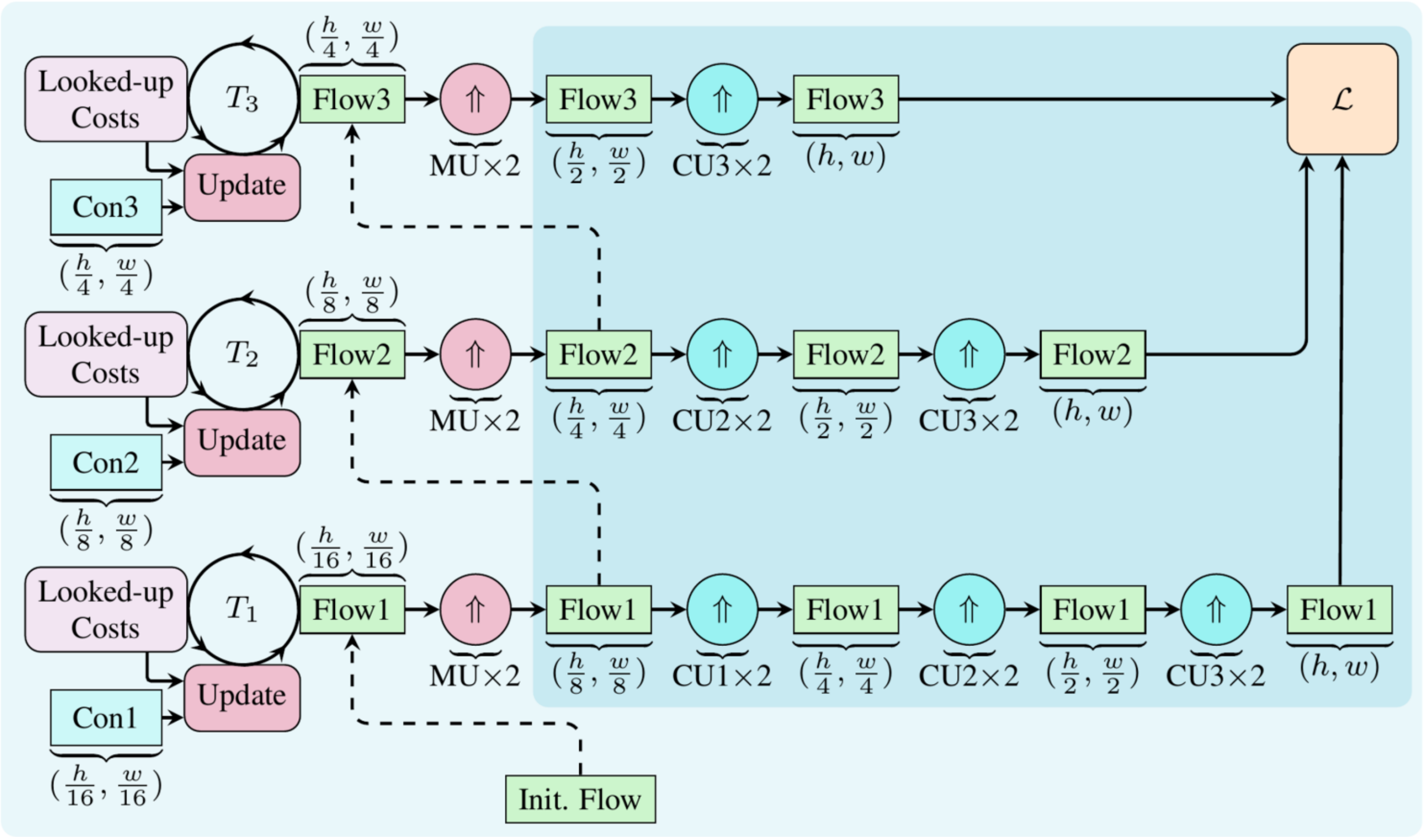
- Step 1:
  - Compute learned upsampling masks from context features of all scales.



Reviving Unsupervised Optical Flow

Gradual Context-based Upsampling

- Step 2:
  - Use gradual context-based upsampling, instead of the bilinear upsampler of MS-RAFT!



Reviving Unsupervised Optical Flow

Impact of the Gradual Context-based Upsampling (GCU)

Upsampled estimate at 1/16
using bilinear upsampling

Reviving Unsupervised Optical Flow

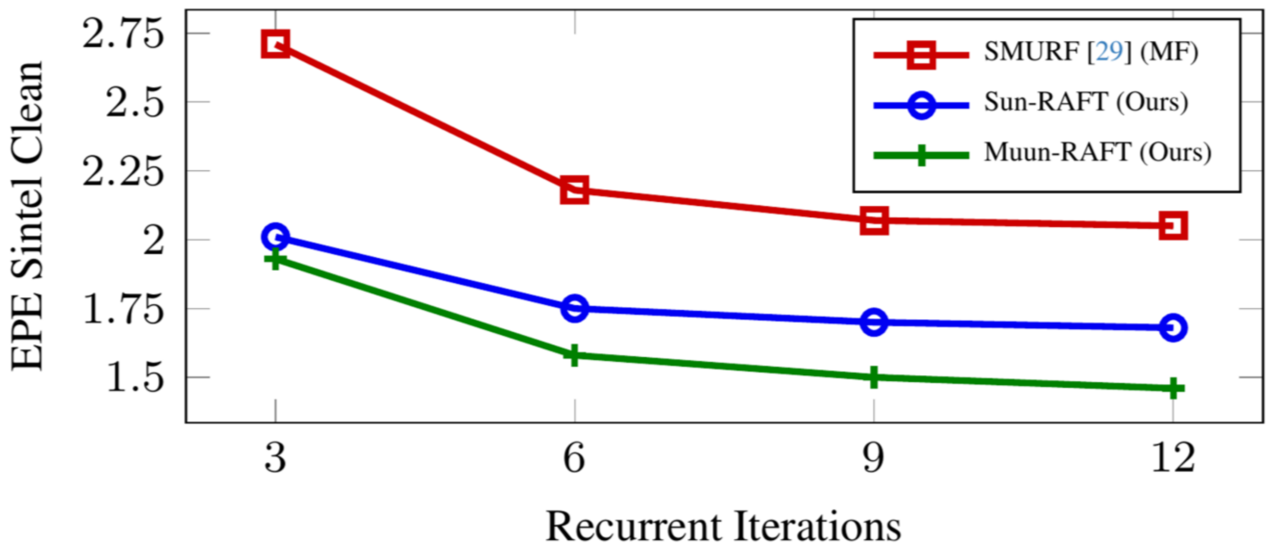
Impact of the Gradual Context-based Upsampling (GCU)

Upsampled estimate at 1/16
using our gradual context-based upsampling

Reviving Unsupervised Optical Flow

Outcome

- Sun-RAFT and Muun-RAFT
  - New SOTA on MPI-Sintel
  - Both methods outperform the multi-frame SMURF method on Sintel-clean with only 3 iterations.



Reviving Unsupervised Optical Flow

Results on Chairs, Sintel and KITTI

- Muun-RAFT's result on Sintel-clean after pre-training on Chairs is on-par with SMURF after fine-tuning on Sintel!
- Both Sun-RAFT and Muun-RAFT achieve a new SOTA on Sintel. (Sun-RAFT is also SOTA on KITTI)
- Muun-RAFT outperforms the SOTA multi-frame SMURF by 28% on Sintel-clean.

Method	Trained on	Chairs Val	Sintel train		KITTI train	
			Clean	Final	EPE	FL
DDFlow	Chairs Train	2.97	4.83	4.85	17.26	–
UFlow	Chairs Train	2.55	3.43	4.17	15.68	32.69
SMURF-EAS (2F)	Chairs Train	1.72	2.19	3.35	7.94	26.51
SMURF (2F)	Chairs Train	1.99	2.48	3.32	12.71	31.04
Sun-RAFT	Chairs Train	1.90	2.22	3.52	9.64	28.71
Muun-RAFT	Chairs Train	**1.71**	**2.02**	**3.08**	**8.61**	**22.06**
DDFlow	Sintel Test	3.46	{2.92}	{3.98}	12.69	–
UFlow	Sintel Test	3.25	3.01	4.09	7.67	17.41
SMURF-EAS (MF)	Sintel Test	1.99	1.99	2.80	4.47	12.55
SMURF-EAS (2F)	Sintel Test	–	2.15	2.99	–	–
Sun-RAFT	Sintel Test	2.23	1.69	2.60	4.76	12.63
Muun-RAFT	Sintel Test	**2.03**	**1.44**	**2.56**	**4.39**	**12.35**
DDFlow	KITTI Test	6.35	6.20	7.08	[5.72]	–
UFlow	KITTI Test	5.05	6.34	7.01	2.84	9.39
SMURF-EAS (MF)	KITTI Test	3.26	3.38	4.47	2.01	6.72
SMURF-EAS (2F)	KITTI Test	–	–	–	2.45	7.53
Sun-RAFT	KITTI Test	3.49	3.97	5.20	**2.17**	**7.66**
Muun-RAFT	KITTI Test	**3.32**	**3.35**	**4.89**	2.34	8.64

Summary

- Reviving unsupervised optical flow
  - Reinvestigated unsupervised training strategies
    - To obtain effective settings
  - Sun-RAFT
    - A simpler alternative to SMURF
    - New SOTA results on Sintel and KITTI for two-frame unsupervised optical flow
  - Muun-RAFT
    - Multi-scale version of Sun-RAFT
    - Introduces a novel gradual context-based upsampling scheme
    - New SOTA on Sintel

- Both methods
  - PyTorch-based and open-source
  - Code available at: https://cv-stuttgart.github.io/Reviving-Unsupervised-OptcialFlow