Optical Flow Training under Limited Label Budget via Active Learning

Shuai Yuan, Xian Sun, Hannah Kim, Shuzhi Yu, and Carlo Tomasi
Department of Computer Science, Duke University, USA

Introduction

"Accuracy-annotation cost" trade-off

- Optical flow: especially ***hard*** to label
- **Supervised**: better accuracy
  Unsupervised: lower annotation cost
  Semi-supervised: balancing both
- Draw **"Error-label ratio" curves** to visualize **label efficiency**
- **Even a few labels help significantly**

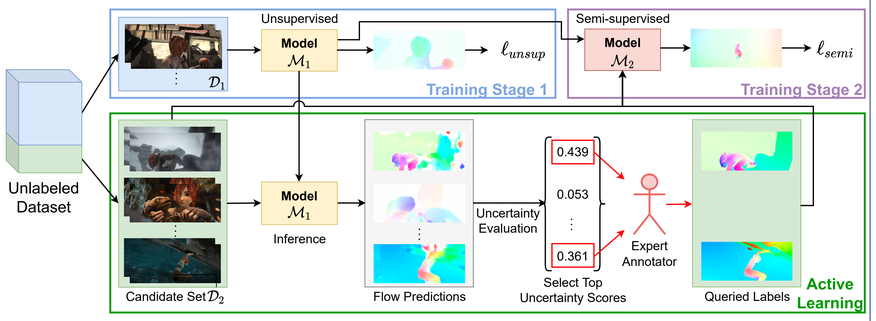
Active learning picks samples for labeling

- Given a label budget, ***which*** samples to label? Random Selection → **Active Learning**
- **New problem setup**: Semi-supervised optical flow estimation **under certain label budget**
- **Active learning picks most effective samples to label → fully use label budget**

Network Architecture

- **Network**: unsup SOTA: ARFlow[1]
- **Unsup loss**: multi-scale bidirectional photometric loss, smoothness loss, and augmentation loss as in ARFlow[1]
- **Sup loss**: multi-scale robust L_1 loss as in PWC-Net[2]
- **Semi-sup loss**: mixed supervised and unsupervised loss
- **Code**: https://github.com/duke-vision/optical-flow-active-learning-release

Active Learning Method



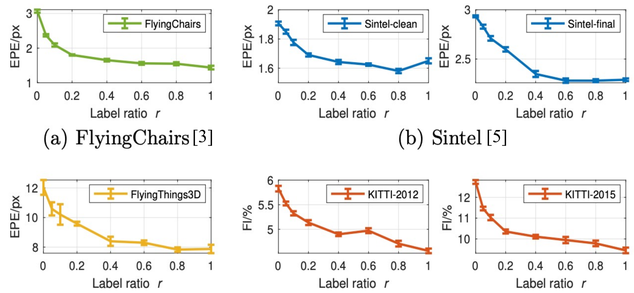
1. **Training Stage 1**: Train an unsup model $\mathcal{M}_1$ on $\mathcal{D}_1$
2. **Active Learning**: Use $\mathcal{M}_1$ to infer flow on $\mathcal{D}_2$ and evaluate uncertainty scores. Top samples query labels
3. **Training Stage 2**: Fine-tune $\mathcal{M}_1$ on the partially labeled $\mathcal{D}_2$ with semi-supervised training to get final model $\mathcal{M}_2$

Uncertainty scores

- *Photo loss*: photometric loss
- *Occ ratio*: ratio of occluded pixels from consistency check
- *Flow grad norm*: estimated flow's gradient magnitude

Semi-supervised Training Results

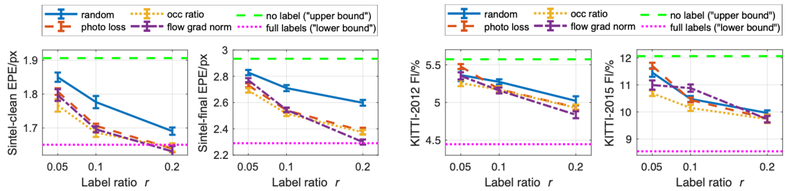
"Error-label ratio" curves: even a few labels help a lot



(a) FlyingChairs[3] (b) Sintel [5] (c) FlyingThings[4] (d) KITTI[6]

Active Learning Results

Validation: *occ ratio* works better than baseline



(a) Sintel (b) KITTI

Benchmark test: *occ ratio* works (even w/ 5-20% labels)

	Label ratio r	Method	Train Clean all	Train Final all	Test Clean all	Test Clean noc	Test Clean occ	Test Final all	Test Final noc	Test Final occ
unsup	$r = 0$	SelFlow [29]	(2.88)	(3.87)	6.56	2.67	38.30	6.57	3.12	34.72
		UFlow [19]	(2.50)	(3.39)	5.21	2.04	31.06	6.50	3.08	34.40
		ARFlow [28]	(2.79)	(3.73)	**4.78**	**1.91**	**28.26**	**5.89**	**2.73**	**31.60**
semi-sup	$r = 0.05$	Ours(rand)	(2.09)	(2.99)	4.04	1.52	24.65	5.49	2.62	28.86
		Ours(occ)	(1.95)	(2.38)	4.11	1.63	**24.39**	**5.28**	**2.49**	**28.03**
		Ours(occ-2x)	(1.94)	(2.55)	**3.96**	**1.45**	24.42	5.35	2.50	28.58
	$r = 0.1$	Ours(rand)	(2.36)	(3.18)	**3.91**	**1.47**	**23.82**	5.21	2.46	27.66
		Ours(occ)	(1.64)	(1.98)	4.28	1.68	25.49	5.31	**2.44**	28.68
		Ours(occ-2x)	(1.75)	(2.30)	4.06	1.63	23.94	**5.09**	2.49	**26.31**
	$r = 0.2$	Ours(rand)	(2.17)	(2.93)	3.89	1.56	**22.86**	5.20	2.50	27.19
		Ours(occ)	(1.35)	(1.63)	4.36	1.86	24.76	5.09	2.45	26.69
		Ours(occ-2x)	(1.57)	(2.05)	**3.79**	**1.44**	23.02	**4.62**	**2.07**	**25.38**
sup	$r = 1$	PWC-Net [47]	(2.02)	(2.08)	4.39	1.72	26.17	5.04	2.45	26.22
		IRR-PWC [15]	(1.92)	(2.51)	3.84	1.47	23.22	4.58	2.15	24.36
		RAFT [49]	(0.77)	(1.27)	**1.61**	**0.62**	**9.65**	**2.86**	**1.41**	**14.68**

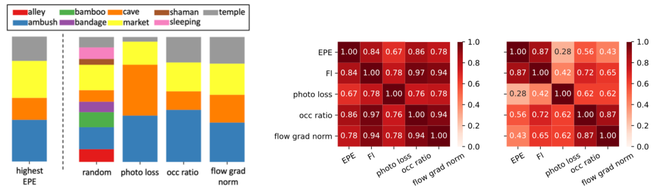
(a) Sintel

	Label ratio r	Method	Train 2012 EPE	Train 2015 EPE	Test 2012 Fl-noc	Test 2012 EPE	Test 2015 Fl-all	Test 2015 Fl-noc	Test 2015 Fl-bg	Test 2015 Fl-fg
unsup	$r = 0$	SelFlow [29]	(1.69)	(4.84)	4.31	2.2	14.19	9.65	12.68	21.74
		UFlow [19]	(1.68)	(2.71)	4.26	1.9	11.13	8.41	9.78	17.87
		ARFlow [28]	(1.44)	(2.85)	-	1.8	11.80	-	-	-
semi-sup	$r = 0.05$	Ours(rand)	(1.25)	(2.61)	3.90	1.6	9.77	6.99	8.33	17.02
		Ours(occ)	(1.22)	(2.29)	3.90	**1.5**	**9.65**	**6.94**	**8.20**	**16.91**
	$r = 0.1$	Ours(rand)	(1.21)	(2.56)	3.75	1.5	9.51	6.69	8.01	17.01
		Ours(occ)	(1.21)	(1.98)	**3.74**	1.5	**8.96**	**6.28**	**7.74**	**15.04**
	$r = 0.2$	Ours(rand)	(1.16)	(2.10)	3.50	1.5	8.38	**5.68**	7.37	**13.44**
		Ours(occ)	(1.13)	(1.73)	**3.49**	1.5	**8.30**	5.69	**7.25**	13.53
sup	$r = 1$	PWC-Net [47]	(1.45)	(2.16)	4.22	1.7	9.60	6.12	9.66	9.31
		IRR-PWC [15]	-	(1.63)	3.21	1.6	7.65	4.86	7.68	7.52
		RAFT [49]	-	(0.63)	-	-	**5.10**	**3.07**	**4.74**	**6.87**

(b) KITTI

Model Analysis

Active learning picks challenging samples to label



References

[1] Liu, L., *et al.*: Learning by analogy: Reliable supervision from transformations for unsupervised optical flow estimation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 6489–6498 (2020)
[2] Sun, D., *et al.*: PWC-Net: CNNs for optical flow using pyramid, warping, and cost volume. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 8934–8943 (2018)
[3] Dosovitskiy, A., *et al.*: Flownet: Learning optical flow with convolutional networks. In: Proceedings of the IEEE International Conference on Computer Vision. pp. 2758–2766 (2015)
[4] Mayer, N., *et al.*: A large dataset to train convolutional networks for disparity, optical flow, and scene flow estimation. In: Proceedings of the IEEE International Conference on Computer Vision (2016)
[5] Butler, D.J., *et al.*: A naturalistic open source movie for optical flow evaluation. In: Proceedings of the European Conference on Computer Vision. pp. 611–625. Part IV, LNCS 7577, Springer-Verlag (2012)
[6] Menze, M. and Geiger, A.: Object scene flow for autonomous vehicles. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2015)

Paper & Appendix

Code

Acknowledgments: This material is based upon work supported by the National Science Foundation under Grant No. 1909821 and by the Intelligence Advanced Research Projects Agency under contract number 2021-21040700001.