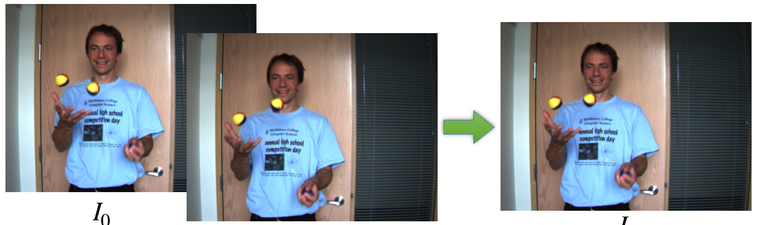
CDFI: Compression-Driven Network Design for Frame Interpolation

Tianyu Ding[1,2*], Luming Liang[2*], Zhihui Zhu[3], Ilya Zharkov[2]

[1]Johns Hopkins University [2]Microsoft [3]University of Denver

Problem of Frame Interpolation

- Synthesize the middle frame given two consecutive frames in a video



Contributions of CDFI

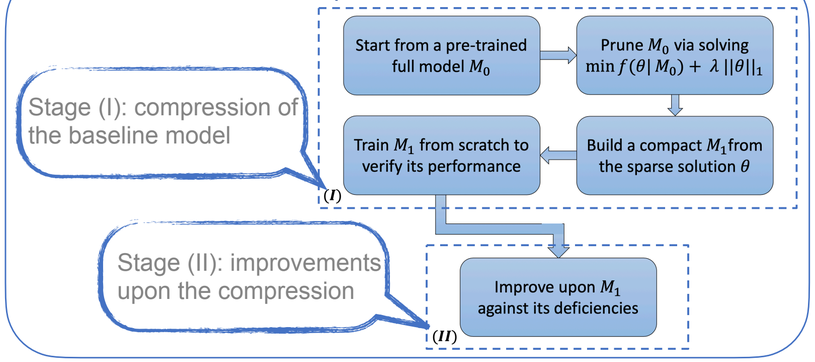
- It leverages model compression to significantly reduce the model size
- It makes room for further improvements
- The framework is generic and does not rely on particular architectures

Concretely, we first compress AdaCoF and show that a 10X compressed AdaCoF performs similarly as its original counterpart; then we improve upon this compressed model with simple modifications

- the final model significantly outperforms AdaCoF with only a quarter in size
- it also performs favorably against other methods

From top to bottom:
the overlaid I_0 and I_1;
the ground-truth $I_{0.5}$;
the result of AdaCoF;
the result of 0.1X AdaCoF;
the result of our model

Pipeline of CDFI



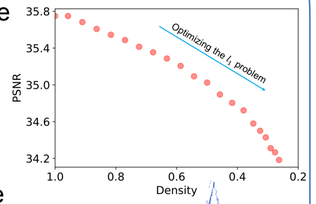
First Stage: Compression of the Baseline Model

- Given a pre-trained full model M_0, we prune its weights θ by solving

$$\min_{\theta} \; f(\theta \,|\, M_0) + \lambda \|\theta\|_1 \qquad (1)$$

with an orthant-based method (OB-ProxG) to effectively promote sparsity in M_0

- We build a small model M_1 from the sparse solution θ of (1): the density of each layer is used as the compression ratio of that layer
- We train the compressed model M_1 from scratch (much faster than that of M_0)

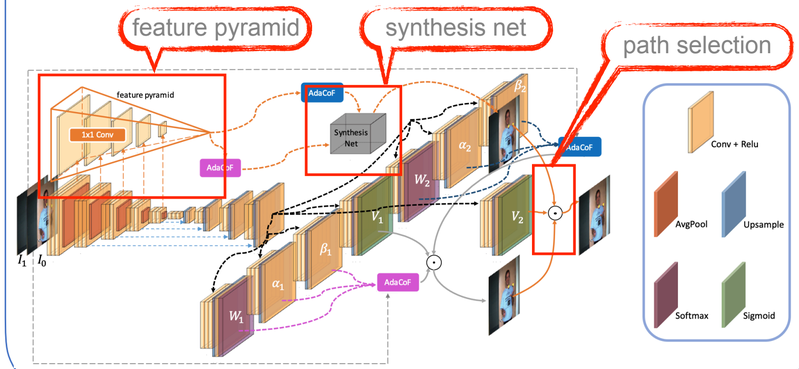


PSNR against the density of AdaCoF when optimizing (1)

	Vimeo-90K	Middlebury	UCF101-DVF	#of Params
	PSNR, SSIM, LPIPS	PSNR, SSIM, LPIPS	PSNR, SSIM, LPIPS	
AdaCoF	34.35, 0.956, 0.019	35.72, 0.959, 0.019	35.16, **0.950**, 0.019	21.84M
Compressed AdaCoF	34.10, 0.954, 0.020	35.43, 0.957, 0.018	35.10, **0.950**, 0.019	2.45M
AdaCoF+	34.56, 0.959, 0.018	36.09, 0.962, 0.017	35.16, **0.950**, 0.019	22.93M
Compressed AdaCoF+	34.44, 0.958, 0.019	35.73, 0.960, 0.018	35.13, **0.950**, 0.019	2.56M
Our Final Model	**35.17**, **0.964**, **0.010**	**37.14**, **0.966**, **0.007**	**35.21**, **0.950**, **0.015**	4.98M

Second Stage: Improve upon the Compression

- The compression makes room for further improvements
- Observing that AdaCoF is short of handling severe occlusion and capturing finer feature details, we are adding three specific components on top of the compressed AdaCoF



Qualitative and Quantitative Results

	Training dataset	Vimeo-90K [62]			Middlebury [1]			UCF101-DVF [36]			Parameters (million)
		PSNR	SSIM	LPIPS	PSNR	SSIM	LPIPS	PSNR	SSIM	LPIPS	
†SepConv - $\mathcal{L}_1$ [44]	proprietary	33.80	0.956	0.027	35.73	0.959	0.017	34.79	0.947	0.029	21.6
†SepConv - $\mathcal{L}_F$ [44]	proprietary	33.45	0.951	0.019	35.03	0.954	0.013	34.69	0.945	0.024	21.6
†CtxSyn - $\mathcal{L}_{Lap}$ [41]	proprietary	34.39	0.961	0.024	36.93	0.964	0.016	34.62	0.949	0.031	–
†CtxSyn - $\mathcal{L}_F$ [41]	proprietary	33.76	0.955	0.017	35.95	0.959	0.013	34.01	0.941	0.024	–
†SoftSplat - $\mathcal{L}_{Lap}$ [42]	Vimeo-90K	**36.10**	**0.970**	0.021	**38.42**	**0.971**	0.016	**35.39**	**0.952**	0.033	–
†SoftSplat - $\mathcal{L}_F$ [42]	Vimeo-90K	**35.48**	**0.964**	**0.013**	**37.55**	0.965	**0.008**	35.10	0.948	0.022	–
†DAIN [2]	Vimeo-90K	34.70	**0.964**	0.022	36.70	0.965	0.017	35.00	**0.950**	0.028	24.02
AdaCoF [32]	Vimeo-90K	34.35	0.956	0.019	35.72	0.959	0.019	35.16	**0.950**	**0.019**	21.84
AdaCoF+ [32]	Vimeo-90K	34.56	0.959	0.018	36.09	0.962	0.017	35.16	**0.950**	**0.019**	22.93
EDSC - $\mathcal{L}_C$ [11]	Vimeo-90K	34.86	0.962	0.016	36.76	**0.966**	0.014	35.17	**0.950**	**0.019**	8.9
EDSC - $\mathcal{L}_F$ [11]	Vimeo-90K	34.57	0.958	**0.010**	36.48	0.963	**0.007**	35.04	0.948	**0.015**	8.9
BMBC [45]	Vimeo-90K	35.06	**0.964**	0.015	36.79	0.965	0.015	35.16	**0.950**	**0.019**	11.0
CAIN [16]	Vimeo-90K	34.65	0.959	0.020	35.11	0.951	0.019	34.98	**0.950**	0.021	42.8
Ours	Vimeo-90K	35.17	**0.964**	**0.010**	37.14	**0.966**	**0.007**	**35.21**	**0.950**	**0.015**	**4.98**

Link to our code: https://github.com/tding1/CDFI