Federated Online Adaptation for Deep Stereo

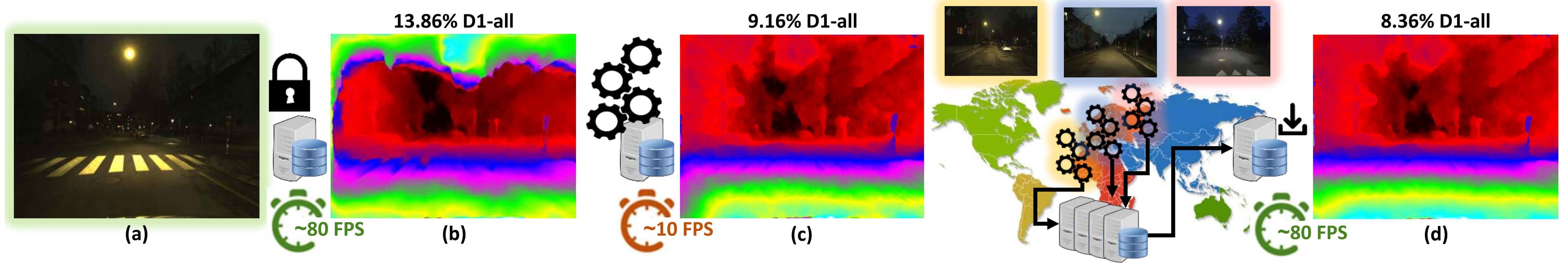
Matteo Poggi   Fabio Tosi

University of Bologna

Project Page:

CVPR SEATTLE, WA JUNE 17-21, 2024

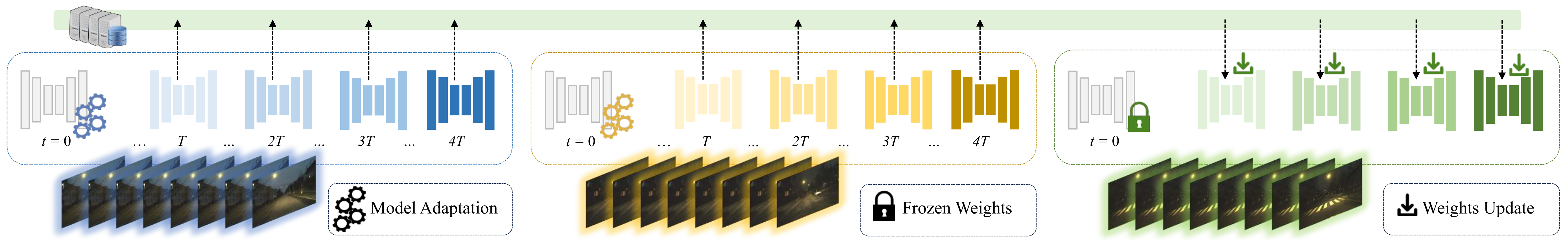
Introduction



Problem: Stereo networks struggle at generalizing to challenging domains. Online adaptation can soften the problem, yet is unfeasible on embedded hardware

Our Proposal: A distributed adaptation framework, where resource-constrained nodes can benefit from the adaptation carried out by other clients

Federated Adaptation Framework



Server executes:

set $t = 0$, load pre-trained $w_t = w_0$
register adapting clients A, listening clients C
initialize buffers $W = [\,][\,]$, $H = [\,][\,]$
while True **do**
  for each client $k \in A$ **in parallel do**
    $W[k] \leftarrow$ ClientUpdate(k, w_t^k, T)
  end for
  for each block i in w_t **do**
    $w_{t+1}[i] \leftarrow \frac{1}{||A||} \sum_{k \in A} W[k][i]$
    send w_{t+1} to C
  end for
  flush buffer $W = [\,]$
  $t \leftarrow t + 1$
end while

ClientUpdateFULL(k, w^k, T):

for each step τ from 0 to T **do**
  sample batch b_τ
  update $w^k \leftarrow w - \eta \nabla \sum_i \ell_i(w^k, b_\tau)$
end for
return w^k to server

Online Adaptation for Stereo: at time step t, the network processes a stereo pair b_t and updates its weights w_t (FULL, Eq. 1) or a subset (MAD, Eq. 2)

$$w_{t+1} \leftarrow w_t - \eta \nabla \sum_i \ell_i(w_t, b_t) \quad (1)$$

$$i = \text{sample}(\text{softmax}(H)) \qquad w_{t+1}[i] \leftarrow w_t[i] - \eta \nabla \ell_i(w_t, b_t)[i] \quad (2)$$

Federated Adaptation for Stereo: a central **server** receives updates by *active* clients – those performing adaptation – aggregates the weights, and sends them to *listening* clients.

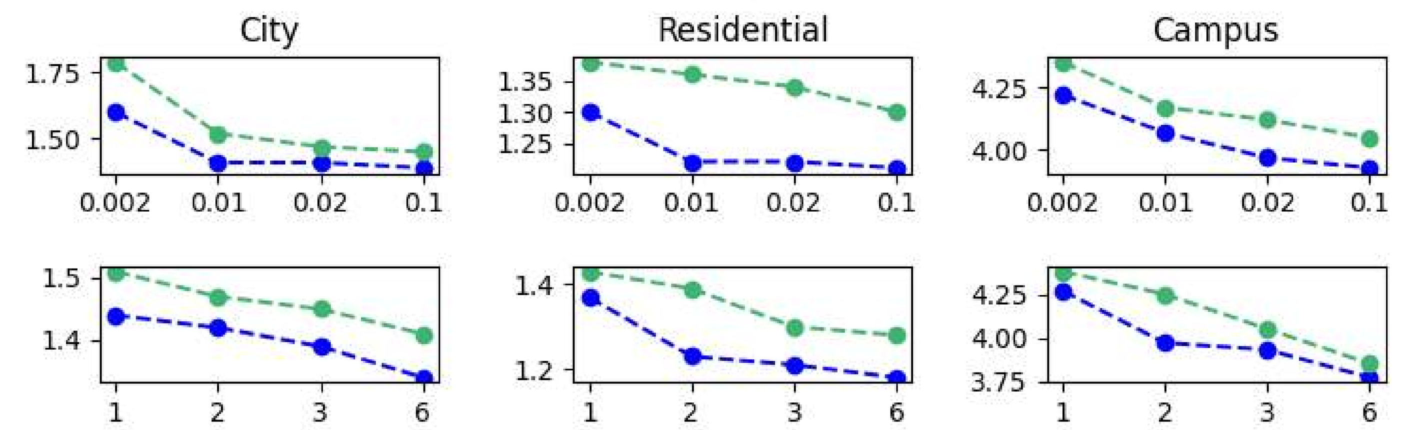
Active clients submit updated weights to the server every T steps, either sending all weights (**FedFULL**) or a subset (**FedMAD**).

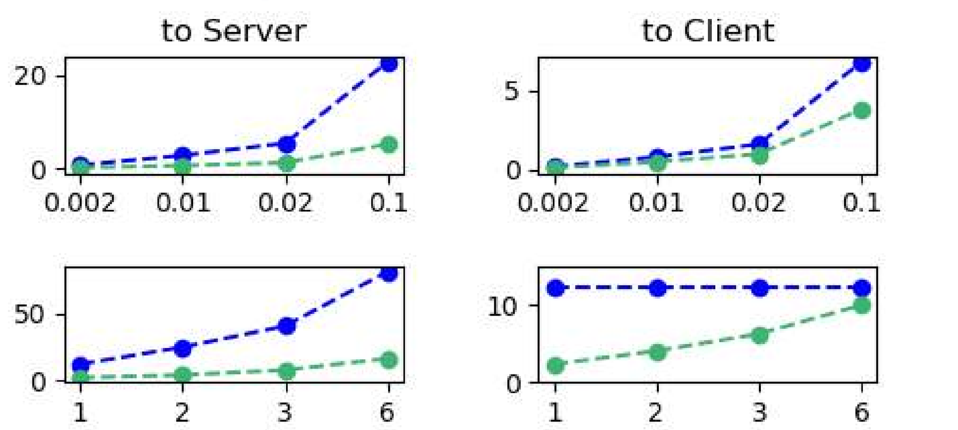
ClientUpdateMAD(k, w^k, T, H):

for each step τ from 0 to T **do**
  sample batch b_τ
  update $w^k \leftarrow w - \eta \nabla \sum_i \ell_i(w^k, b_\tau)$
  for each block i in H **do**
    $H[k][i]$ += 1 **if** i was updated
  end for
end for
$j \leftarrow$ sample(softmax($H[k]$))
$H[k][j] = 0.9 \cdot H[k][j]$
return $j, w^k[j]$ to server

Experimental Results

Ablation Study – update frequency and number of adapting clients





Experiments on DrivingStereo and DSEC

Model	Adapt. mode	Rainy D1-all (%)	Rainy EPE (px)	Dusky D1-all (%)	Dusky EPE (px)	Cloudy D1-all (%)	Cloudy EPE (px)	Data Traffic To Server (MB/s)	Data Traffic To Client (MB/s)	Runtime 3090 (ms)	Runtime AGX (ms)
RAFT-Stereo	No Adapt.	11.52	1.59	3.08	0.88	4.18	1.02	-	-	264	> 1000
CREStereo		17.43	3.61	7.08	1.23	4.08	1.07	-	-	415	
IGEV-Stereo		11.70	1.85	3.57	0.95	5.27	1.26	-	-	389	
UniMatch		14.84	2.69	7.51	1.27	5.78	1.25	-	-	85	
CoEX	No Adapt.	13.48	2.53	11.00	1.58	4.46	1.16	-	-	16	130
HITNet		14.08	2.74	8.88	1.37	4.17	1.14	-	-	29	311
TemporalStereo		18.53	3.94	13.61	1.80	6.02	1.31	-	-	33	✗
MADNet	No Adapt.	27.14	3.90	24.73	2.45	11.00	1.77	-	-	6	64
MADNet 2 **(ours)**		16.47	3.03	13.16	1.66	6.72	1.35	-	-	4	43

(a) No Adaptation – pre-trained on SceneFlow

Model	Adapt. mode	Rainy D1-all (%)	Rainy EPE (px)	Dusky D1-all (%)	Dusky EPE (px)	Cloudy D1-all (%)	Cloudy EPE (px)	Data Traffic To Server (MB/s)	Data Traffic To Client (MB/s)	Runtime 3090 (ms)	Runtime AGX (ms)
MADNet 2	FULL	10.19	1.70	11.36	1.54	5.76	1.27	-	-	30	492
	MAD	11.12	1.78	13.36	1.61	5.93	1.26	-	-	12	65
MADNet 2	FedFULL	11.57	2.00	10.65	1.44	5.45	1.20	20.6	6.8	4	43
	FedMAD	11.71	2.10	10.12	1.41	5.60	1.21	4.6	3.6	4	43

(b) Single-agent vs Federated Adaptation – photometric loss

Model	Adapt. mode	Rainy D1-all (%)	Rainy EPE (px)	Dusky D1-all (%)	Dusky EPE (px)	Cloudy D1-all (%)	Cloudy EPE (px)	Data Traffic To Server (MB/s)	Data Traffic To Client (MB/s)	Runtime 3090 (ms)	Runtime AGX (ms)
MADNet 2	FULL++	10.34	2.27	4.41	1.04	5.20	1.63	-	-	20	470
	MAD++	10.06	2.01	5.25	1.09	4.34	1.09	-	-	8	48
MADNet 2	FedFULL++	8.33	1.73	4.13	1.00	4.55	1.13	28.8	9.6	4	43
	FedMAD++	8.58	1.74	4.40	1.01	4.65	1.16	6.5	4.5	4	43

(c) Single-agent vs Federated Adaptation – proxy labels

Model	Adapt. mode	Night #1 D1-all (%)	Night #1 EPE (px)	Night #2 D1-all (%)	Night #2 EPE (px)	Night #3 D1-all (%)	Night #3 EPE (px)	Night #4 D1-all (%)	Night #4 EPE (px)	Data Traffic To Server (MB/s)	Data Traffic To Client (MB/s)	Runtime 3090 (ms)	Runtime AGX (ms)
RAFT-Stereo	No adapt.	13.04	3.41	21.64	4.26	10.91	1.91	10.07	1.68	-	-	1030	> 8000
CREStereo		11.34	2.38	23.48	3.19	15.37	2.39	12.42	1.75	-	-	1242	
IGEV-Stereo		9.14	1.85	11.97	1.96	12.65	2.01	10.01	1.66	-	-	1250	
UniMatch		34.29	5.43	39.80	5.32	26.75	3.29	26.29	3.28	-	-	480	
CoEX	No adapt.	6.26	1.72	10.81	1.87	8.60	1.64	8.31	1.53	-	-	53	539
HITNet		6.49	1.54	9.57	1.71	8.28	1.62	7.88	1.47	-	-	112	1400
TemporalStereo		7.17	1.68	10.22	1.92	8.66	1.62	8.40	1.49	-	-	118	✗
MADNet 2 **(ours)**	No Adapt.	8.94	1.97	13.86	2.32	10.63	1.83	10.55	1.69	-	-	12	111

(a) No adaptation – pre-trained on SceneFlow

Model	Adapt. mode	Night #1 D1-all (%)	Night #1 EPE (px)	Night #2 D1-all (%)	Night #2 EPE (px)	Night #3 D1-all (%)	Night #3 EPE (px)	Night #4 D1-all (%)	Night #4 EPE (px)	Data Traffic To Server (MB/s)	Data Traffic To Client (MB/s)	Runtime 3090 (ms)	Runtime AGX (ms)
MADNet 2	FULL	5.65	1.41	9.16	1.60	8.12	1.50	8.97	1.46	-	-	102	1238
	MAD	5.79	1.52	8.87	1.60	7.89	1.49	8.50	1.46	-	-	30	253
MADNet 2	FedFULL	5.50	1.43	8.36	1.52	7.63	1.48	7.57	1.37	13.8	4.6	12	111
	FedMAD	5.52	1.43	8.39	1.53	7.91	1.50	7.79	1.39	2.9	2.0	12	111

(b) Single-agent vs Federated Adaptation – photometric loss

Model	Adapt. mode	Night #1 D1-all (%)	Night #1 EPE (px)	Night #2 D1-all (%)	Night #2 EPE (px)	Night #3 D1-all (%)	Night #3 EPE (px)	Night #4 D1-all (%)	Night #4 EPE (px)	Data Traffic To Server (MB/s)	Data Traffic To Client (MB/s)	Runtime 3090 (ms)	Runtime AGX (ms)
MADNet 2	FULL++	4.69	1.28	7.13	1.43	6.20	1.35	6.06	1.27	-	-	45	808
	MAD++	5.66	1.43	7.76	1.49	6.57	1.39	6.47	1.30	-	-	16	172
MADNet 2	FedFULL++	4.99	1.33	7.03	1.41	6.43	1.37	6.18	1.28	21.7	7.1	12	111
	FedMAD++	4.99	1.34	7.13	1.42	6.48	1.38	6.23	1.28	7.3	5.8	12	111

(c) Single-agent vs Federated Adaptation – proxy labels

Qualitative Results