

Procjena optičkog toka dubokim povratnim modelima

Autor: Frano Rajić

Mentor: prof. dr. sc. Siniša Šegvić

Sadržaj

1. Optički tok

Što je optički tok i kako se vizualizira

2. Duboki model RAFT

Pregled arhitekture modela RAFT

3. Eksperimenti autora

Prikaz postignutih rezultata

4. Evaluacija

Primjeri procjena optičkog toka koje daje RAFT

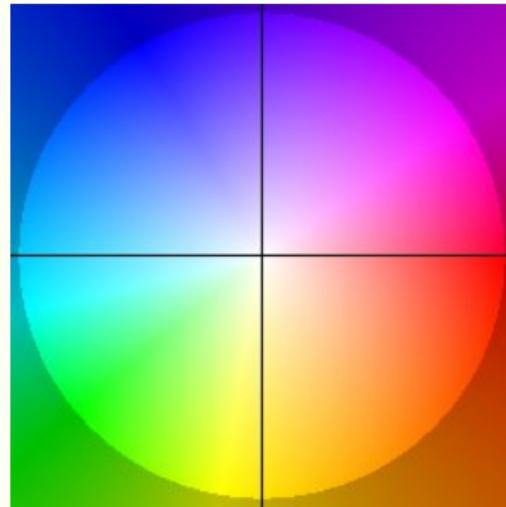
Optički tok

Što je optički tok?

- ▢ prividno kretanje u slijedu slika
- ▢ svakom pikselu $(x, y) \in \mathbb{N}^2$ pridijeljuje vektor $(u, v) \in \mathbb{R}^2$ koji govori gdje se taj piksel pomaknuo na idućoj slici:

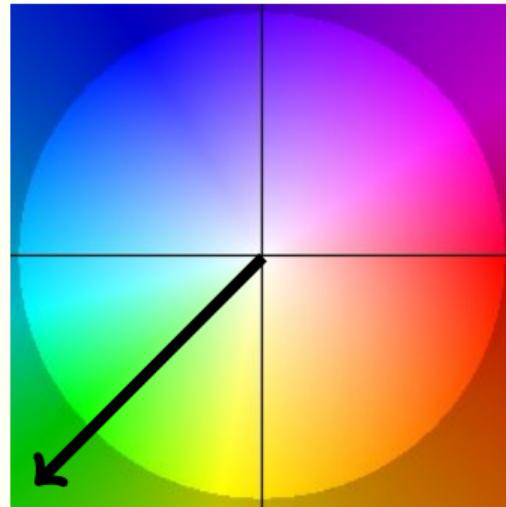
$$(x', y') = (x + u, y + v)$$

Ključ

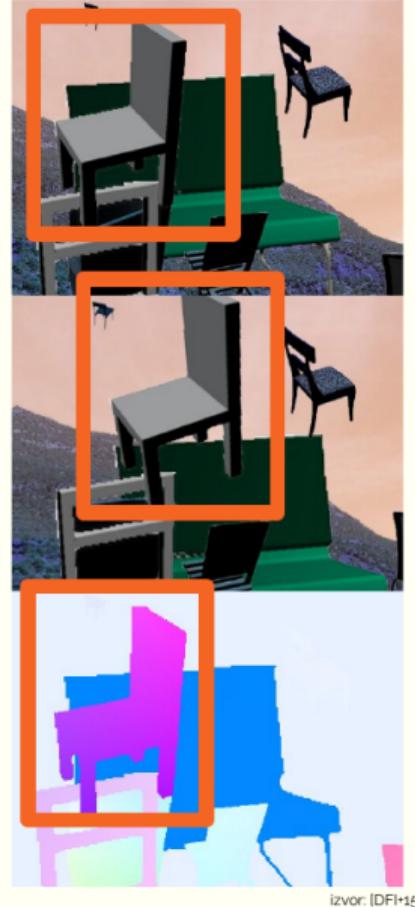
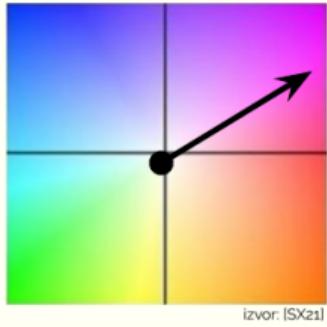


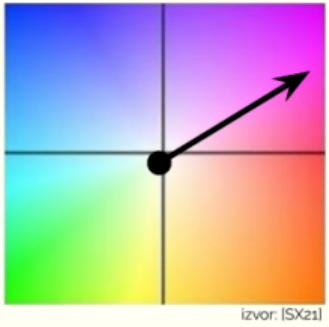
Izvor: [CGN14]

Ključ



Izvor: [CGN14]





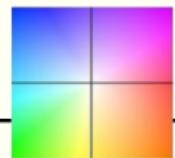
izvor: [TD20]

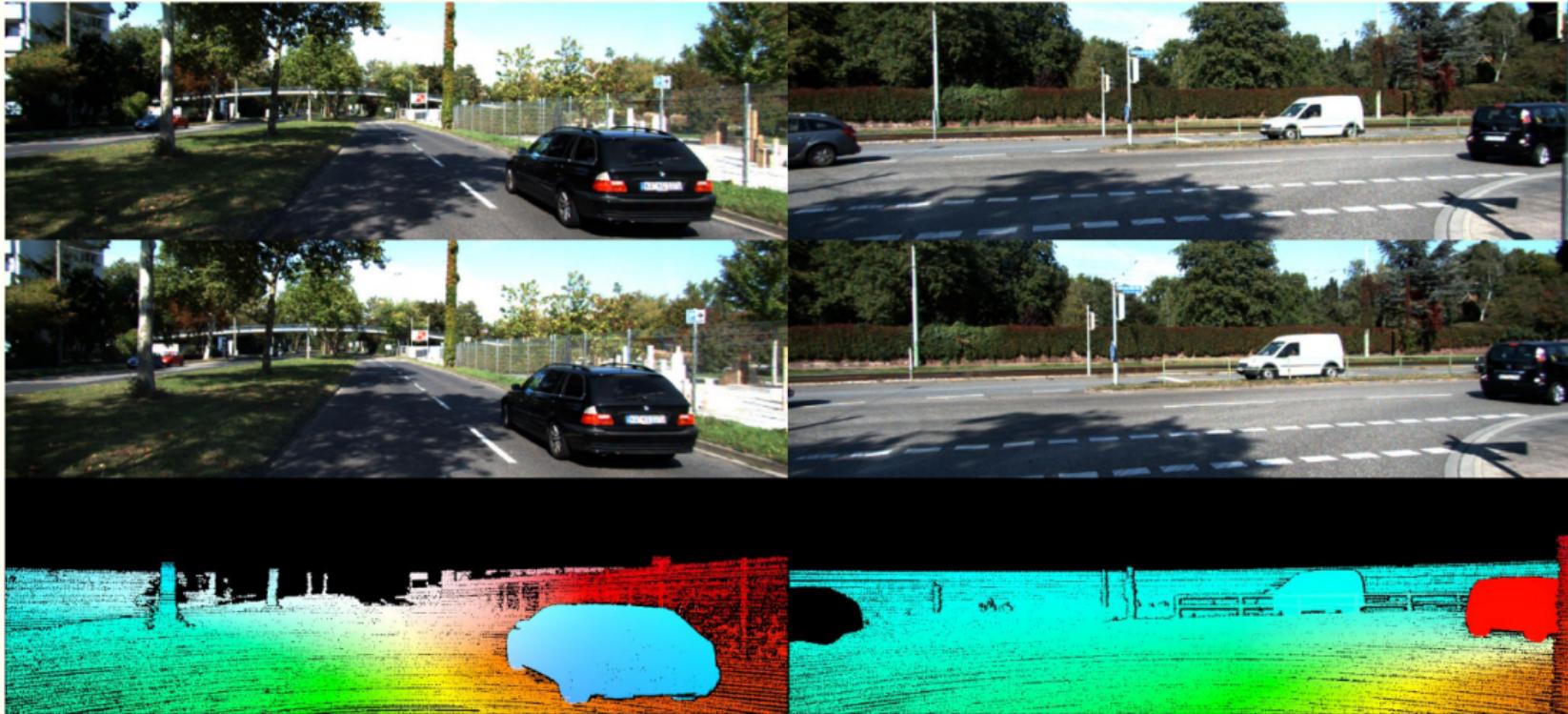
Skupovi podataka

- ▶ noviji skupovi su zahtjevni:
 - ▶ okluzija
 - ▶ veliki pomaci piksela
 - ▶ različiti uvjeti osvjetljenja
- ▶ primjerice:
 - ▶ MPI-Sintel
 - ▶ KITTI-2015

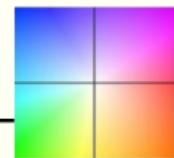


MPI-Sintel





KITTI-2015

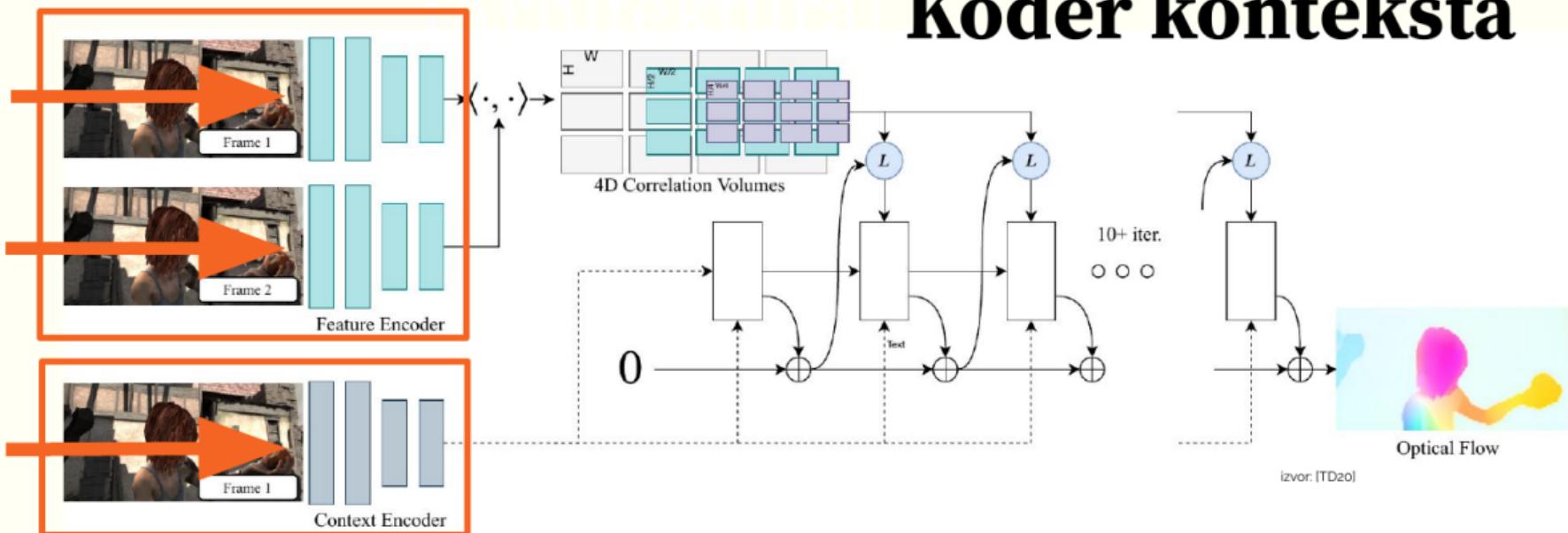


izvor: [SX21]

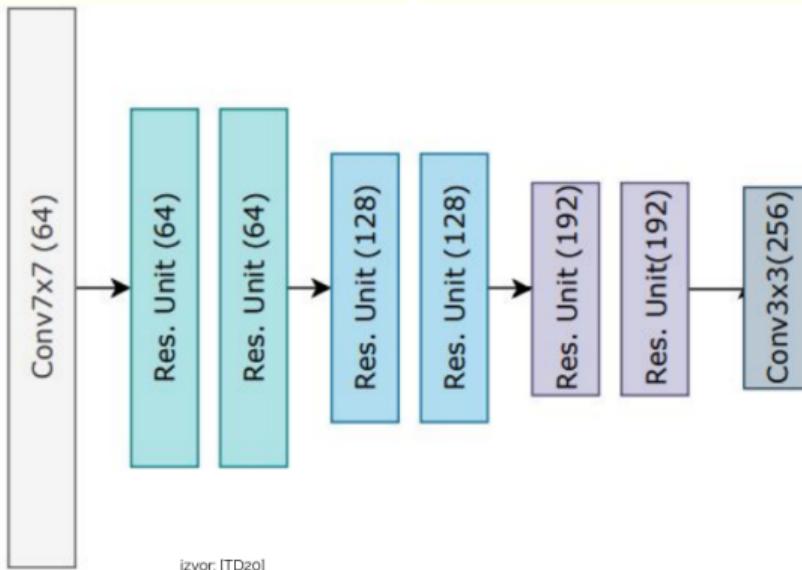
Duboki model RAFT

- Duboki povratni model
- Zachary Teed i Jia Deng: "RAFT: Recurrent all-pairs field transforms for optical flow"
- *ECCV2020 Best Paper Award*
- *state-of-the-art* procjena optičkog toka
 - na skupu MPI-Sintel
 - na skupu KITTI-2015

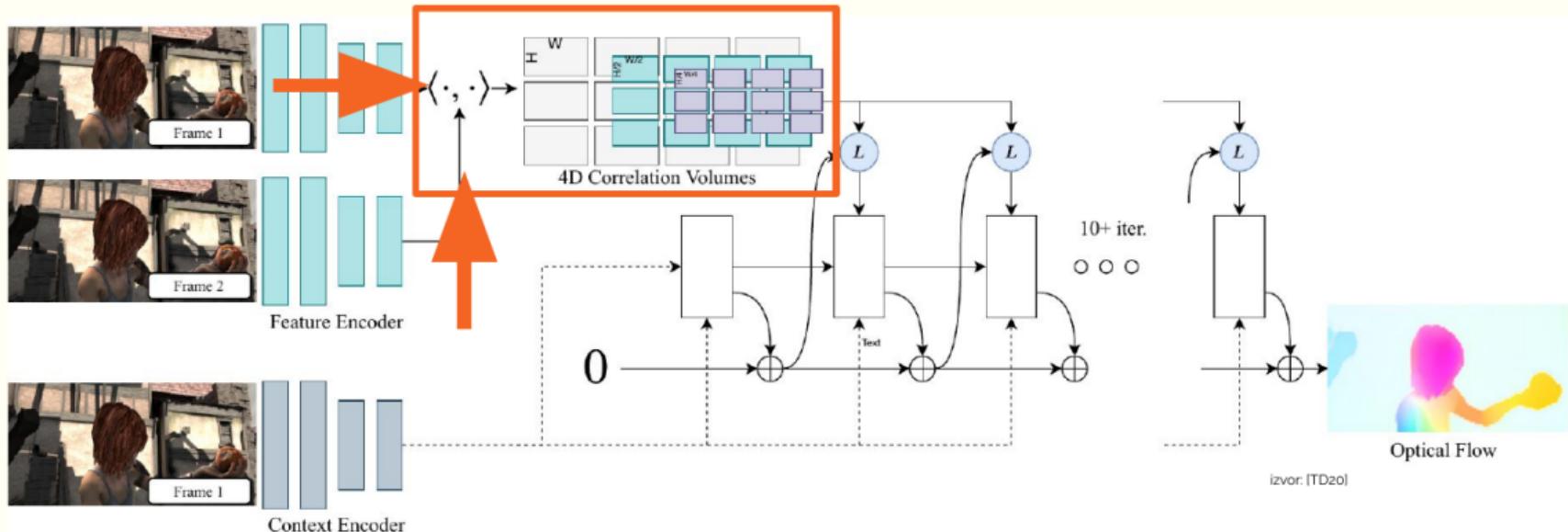
[1/3] [Arhitektura] Koder značajki Koder konteksta



[1/3] [Arhitektura] Koder značajki Koder konteksta



[2/3] [Arhitektura] Korelacijski sloj



[2/3] [Arhitektura] Korelacijski sloj



Image 1

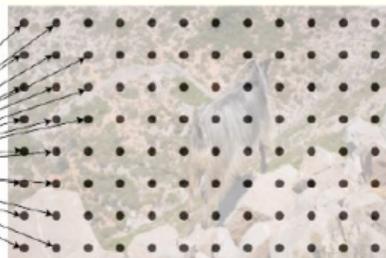
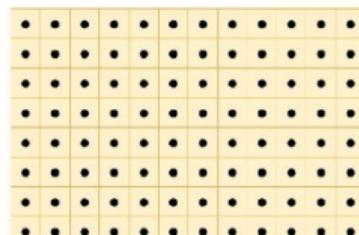
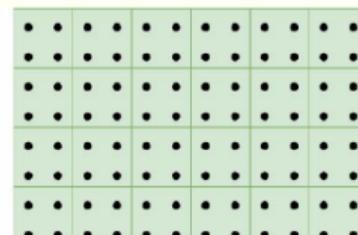


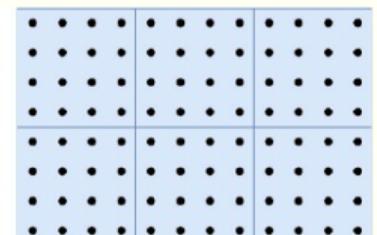
Image 2



$$\mathbf{C}^1 \in H \times W \times H \times W$$



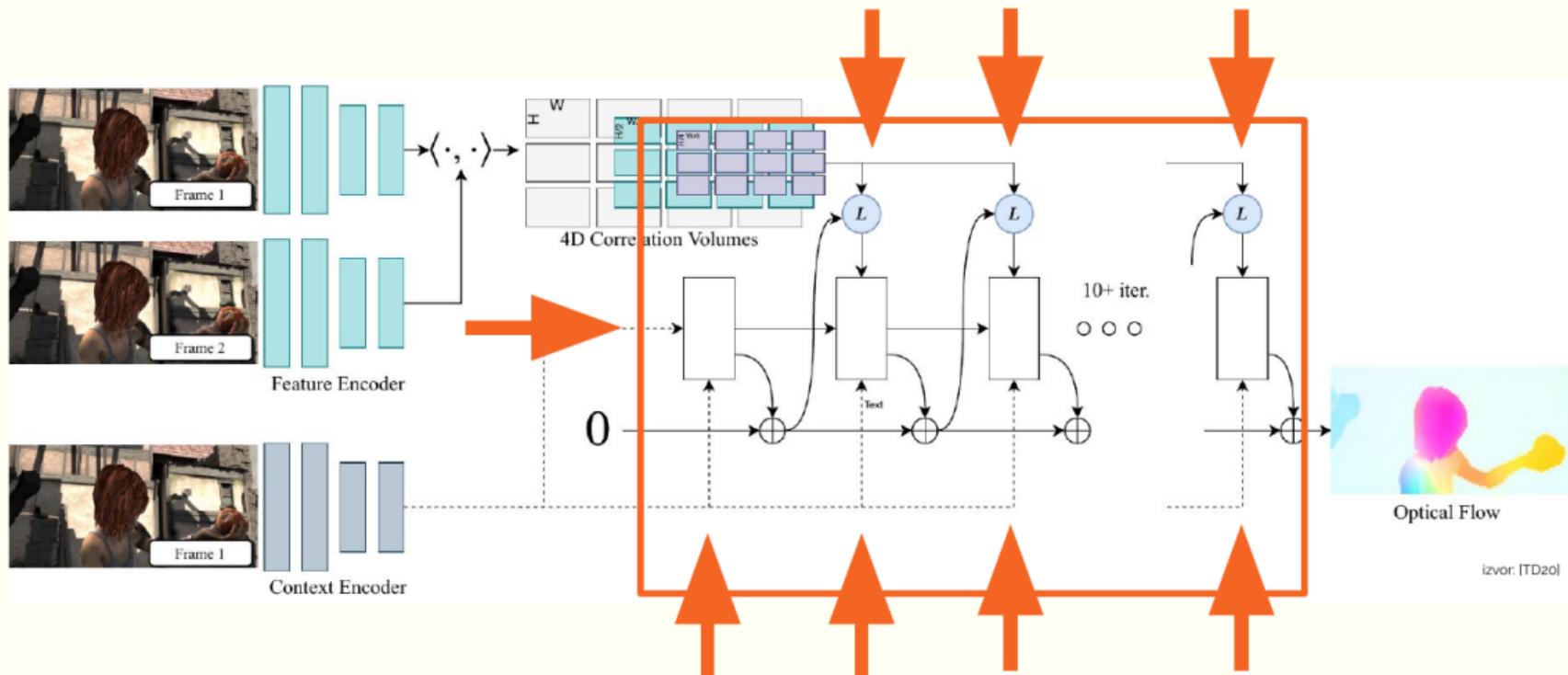
$$\mathbf{C}^2 \in H \times W \times H/2 \times W/2$$



$$\mathbf{C}^3 \in H \times W \times H/4 \times W/4$$

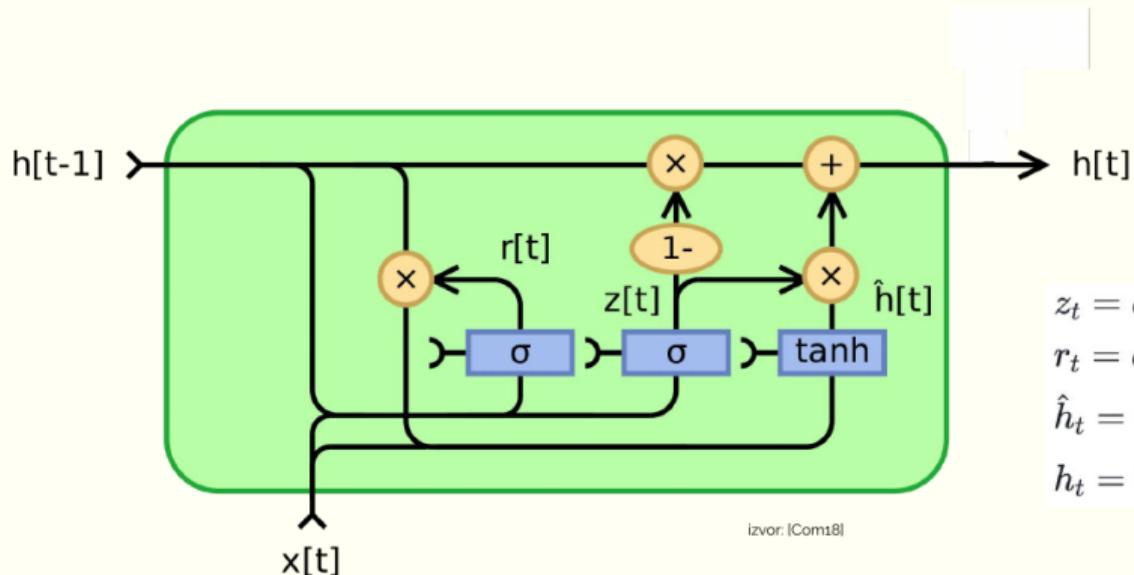
izvor: ITD201

[3/3] [Arhitektura] GRU element



izvor: ITD20

[3/3] [Arhitektura] GRU element



$$z_t = \sigma(\text{Conv}_{3 \times 3}([h_{t-1}, x_t], W_z))$$

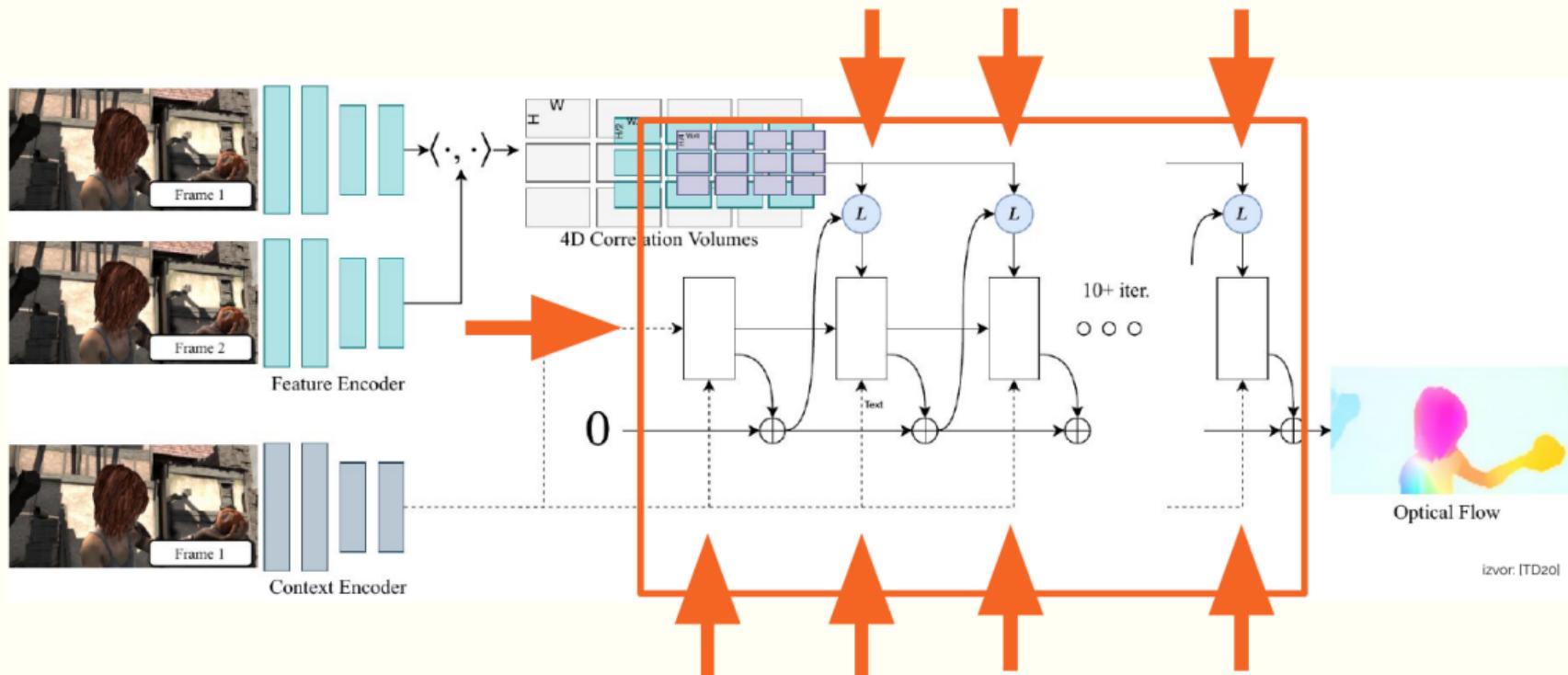
$$r_t = \sigma(\text{Conv}_{3 \times 3}([h_{t-1}, x_t], W_r))$$

$$\hat{h}_t = \tanh(\text{Conv}_{3 \times 3}([r_t \odot h_{t-1}, x_t], W_h))$$

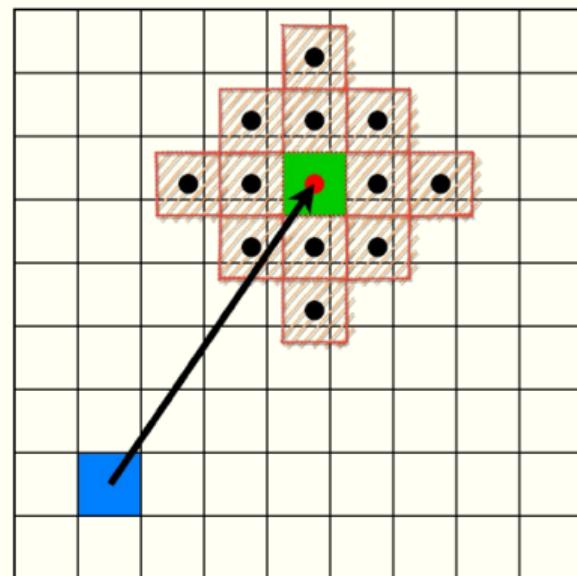
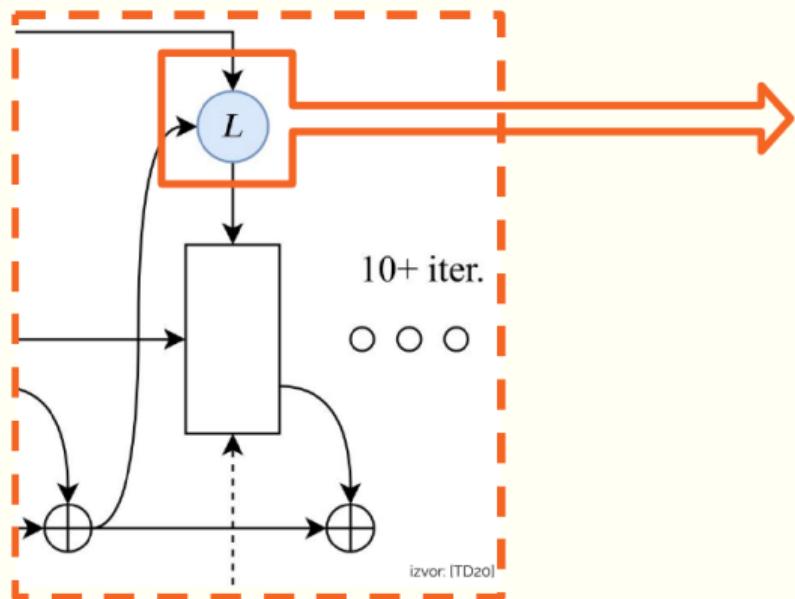
$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \hat{h}_t$$

izvor: [Com18]

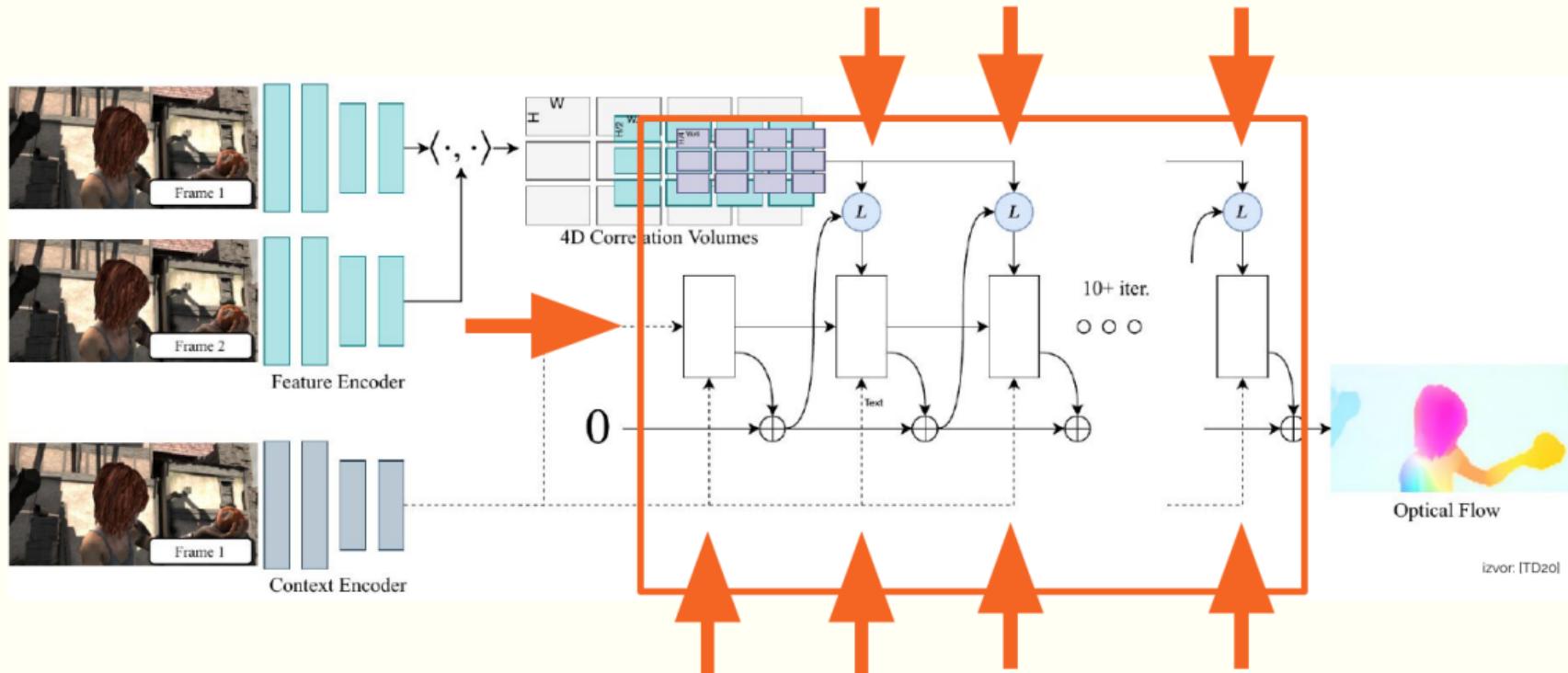
[3/3] [Arhitektura] GRU element



[3/3] [Arhitektura] GRU element



[3/3] [Arhitektura] GRU element



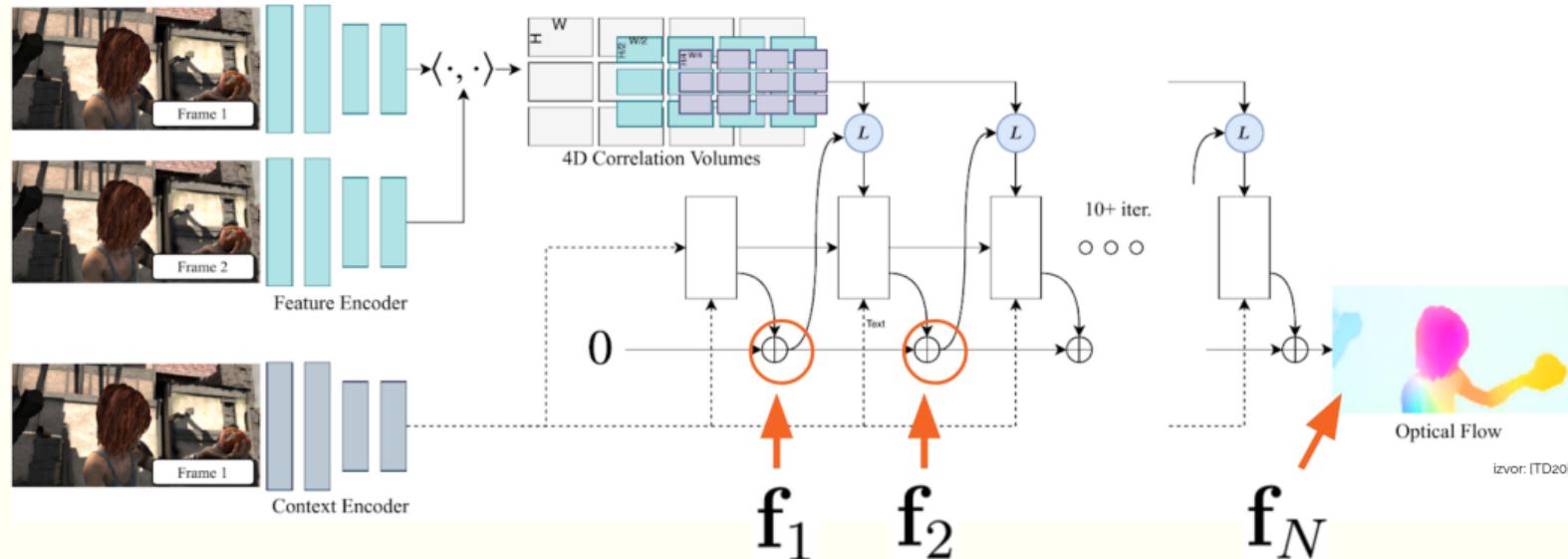
izvor: ITD20

Funkcija gubitka

- ▀ N iteracija daje N procjena optičkog toka: $\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_N\}$

Funkcija gubitka

- N iteracija daje N procjena optičkog toka: $\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_N\}$



izvor: [TD20]

Funkcija gubitka

- ▶ N iteracija daje N procjena optičkog toka: $\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_N\}$
- ▶ funkcija gubitka \mathbf{f}_g - oznaka točnog optičkog toka. $\gamma = 0.8$

$$\mathcal{L} = \sum_{i=1}^N \gamma^{N-i} \|\mathbf{f}_g - \mathbf{f}_i\|_1$$

Eksperimenti autora

Implementacija

- PyTorch
- optimizacijski algoritam AdamW
- *gradient clipping*

Training Data	Method	MPI-Sintel (train)		KITTI-2015 (train)		MPI-Sintel (test)		KITTI-2015 (test)	
		Clean	Final	EE	Fl-all	Clean	Final	Fl-all	
-	FlowFields [1]	-	-	-	-	3.75	5.81	15.31	
-	FlowFields++ [20]	-	-	-	-	2.94	5.49	14.82	
S	DCFFlow [27]	-	-	-	-	3.54	5.12	14.86	
S	MRFlow [26]	-	-	-	-	2.53	5.38	12.19	
C + T	HD3 [29]	3.84	8.77	13.17	24.0	-	-	-	
	LiteFlowNet [10]	2.48	4.04	10.39	28.5	-	-	-	
	PWC-Net [23]	2.55	3.93	10.35	33.7	-	-	-	
	LiteFlowNet2 [11]	2.24	3.78	8.97	25.9	-	-	-	
	VCN [28]	2.21	3.68	8.36	25.1	-	-	-	
	MaskFlowNet [30]	2.25	3.61	-	<u>23.1</u>	-	-	-	
	FlowNet2 [13]	<u>2.02</u>	3.54 ¹	10.08	30.0	3.96	6.02	-	
	RAFT-S	2.21	<u>3.35</u>	<u>7.51</u>	26.9	-	-	-	
	RAFT	1.43	2.71	5.04	17.4	-	-	-	
C+T+S/K	FlowNet2 [13]	(1.45)	(2.01)	(2.30)	(6.8)	4.16	5.74	11.48	
	HD3 [29]	(1.87)	(1.17)	(1.31)	(4.1)	4.79	4.67	6.55	
	IRR-PWC [12]	(1.92)	(2.51)	(1.63)	(5.3)	3.84	4.58	7.65	
	ScopeFlow [2]	-	-	-	-	<u>3.59</u>	<u>4.10</u>	<u>6.82</u>	
	RAFT	(0.77)	(1.20)	(0.64)	(1.5)	2.08	3.41	5.27	
C+T+S+K+H	LiteFlowNet2 ¹ [11]	(1.30)	(1.62)	(1.47)	(4.8)	3.48	4.69	7.74	
	PWC-Net+ [24]	(1.71)	(2.34)	(1.50)	(5.3)	3.45	4.60	7.72	
	VCN [28]	(1.66)	(2.24)	(1.16)	(4.1)	2.81	4.40	6.30	
	MaskFlowNet [30]	-	-	-	-	2.52	4.17	<u>6.10</u>	
	RAFT	(0.76)	(1.22)	(0.63)	(1.5)	<u>1.94</u>	<u>3.18</u>	5.10	
	RAFT (<i>warm-start</i>)	(0.77)	(1.27)	-	-	1.61	2.86	-	

Rezultati autora

Izvor: ITD20

Evaluacija

Evaluacija RAFT modela

1. Primjer 1 – [MPI-Sintel](#)
2. Primjer 2 – [samo optički tok](#)
3. Primjer 2 – [video + optički tok \[cro\]](#)

Literatura |

-  Giannis Chantas, Theodosios Gkamas, and Christophoros Nikou, *Variational-bayes optical flow*, Journal of Mathematical Imaging and Vision **50** (2014).
-  Wikimedia Commons, *Gradient recurrent unit, fully gated version*, 2018.
-  cro3x3, *Ulaz u crnu riku - maraton lada 2020*.
-  Alexey Dosovitskiy, Philipp Fischer, Eddy Ilg, Philip Hausser, Caner Hazirbas, Vladimir Golkov, Patrick Van Der Smagt, Daniel Cremers, and Thomas Brox, *Flownet: Learning optical flow with convolutional networks*, Proceedings of the IEEE international conference on computer vision, 2015, pp. 2758–2766.
-  Syed Tafseer Haider Shah and Xiang Xuezhi, *Traditional and modern strategies for optical flow: an investigation*, SN Applied Sciences **3** (2021).
-  Zachary Teed and Jia Deng, *Raft: Recurrent all-pairs field transforms for optical flow*, 2020.