

# Методы поиска областей открытия/закрытия в видео и способы их улучшения

Великанов Максим

*Video Group*

*CS MSU Graphics&Media Lab*

# Содержание

- **Введение**
- Методы поиска областей открытия/закрытия
- Методы построения оптического потока
- Заключение

# Введение

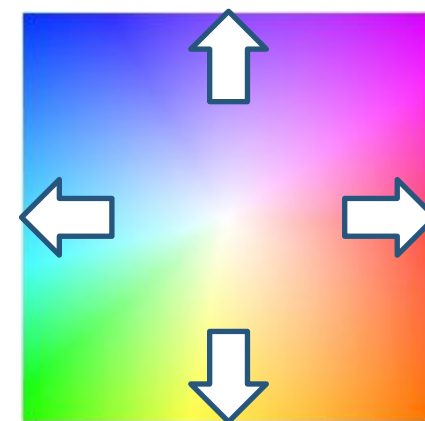
## Оптический поток



1 кадр

2 кадр

оптический поток



направление движения

# Введение

## Области открытия/закрытия



видео



оптический поток

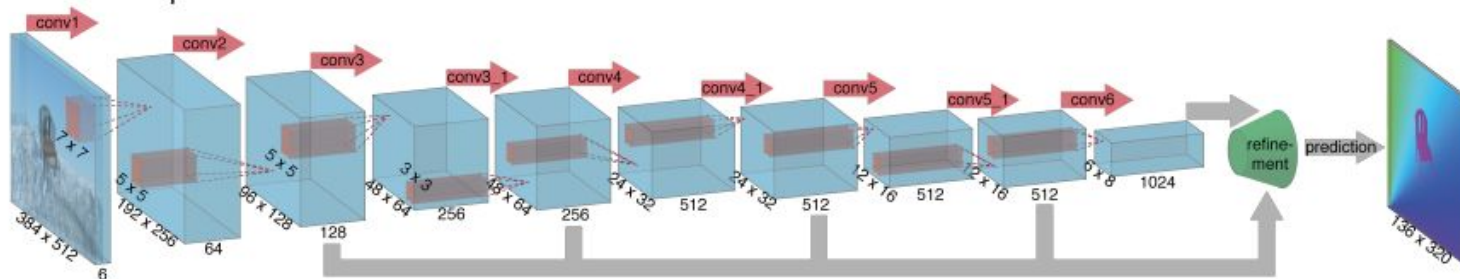


области открытия/закрытия

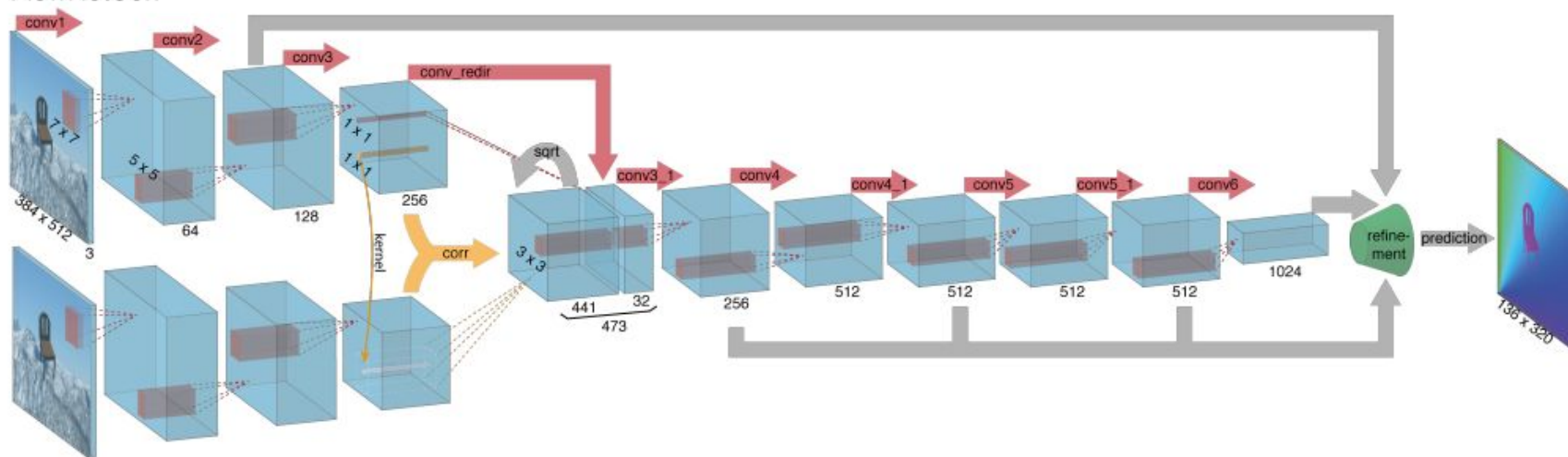
# Введение

## FlowNet — encoder

FlowNetSimple

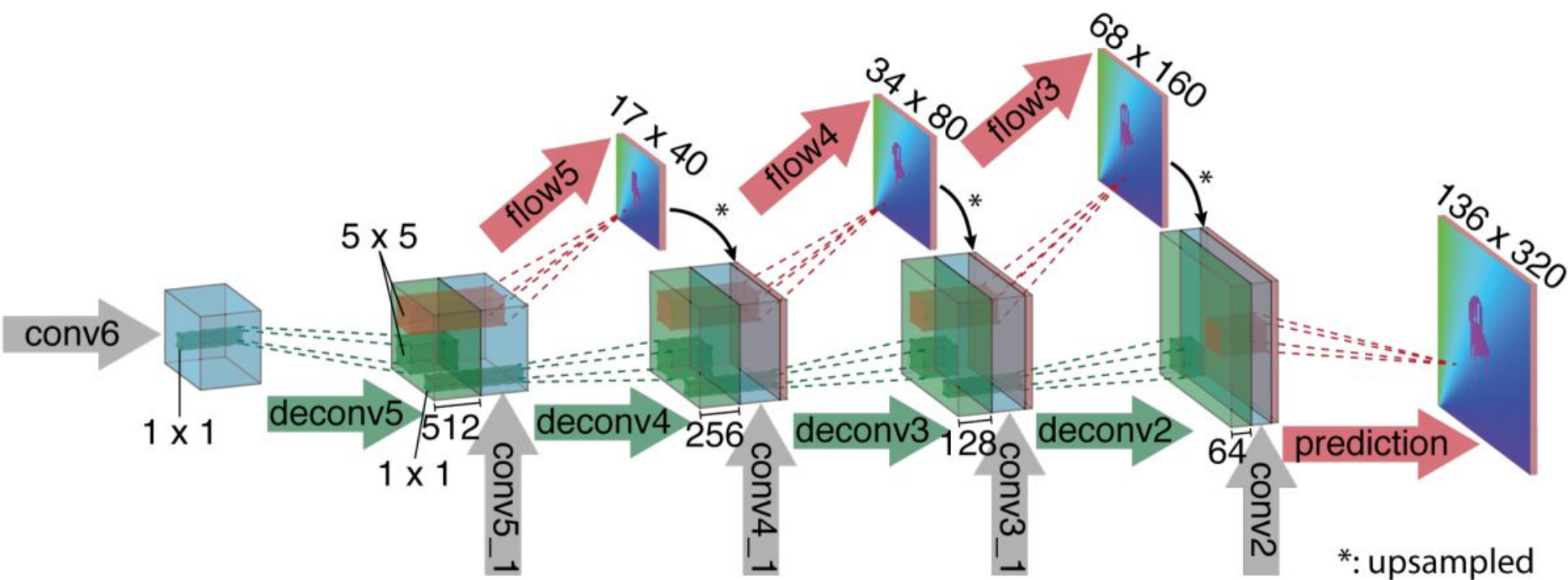


FlowNetCorr

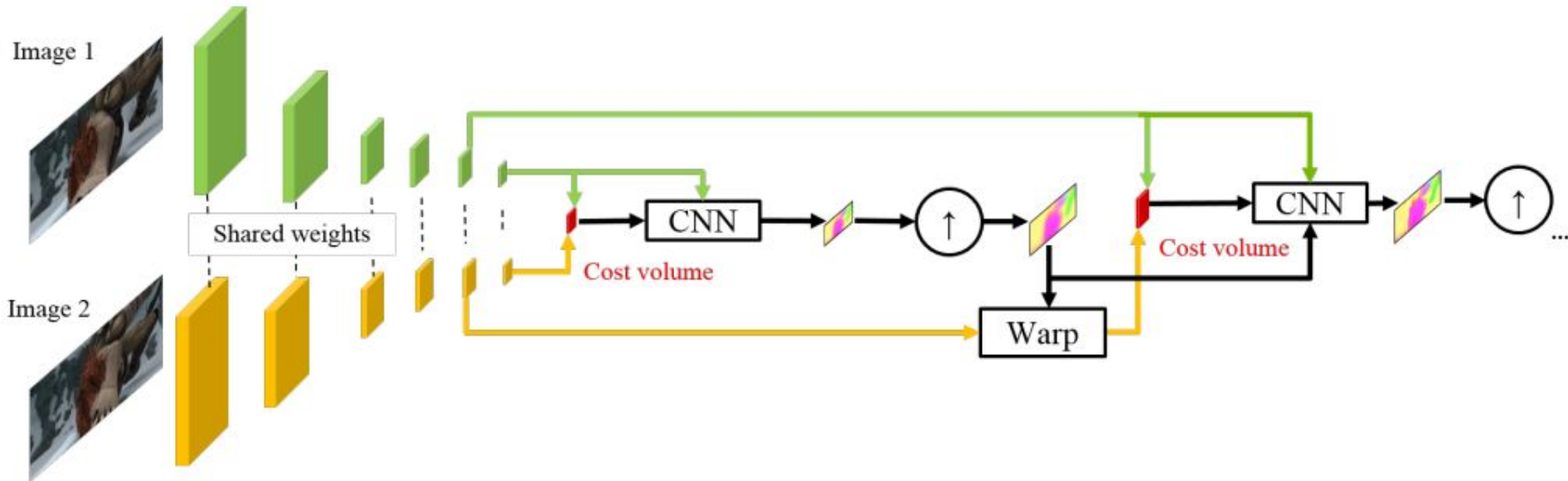


# Введение

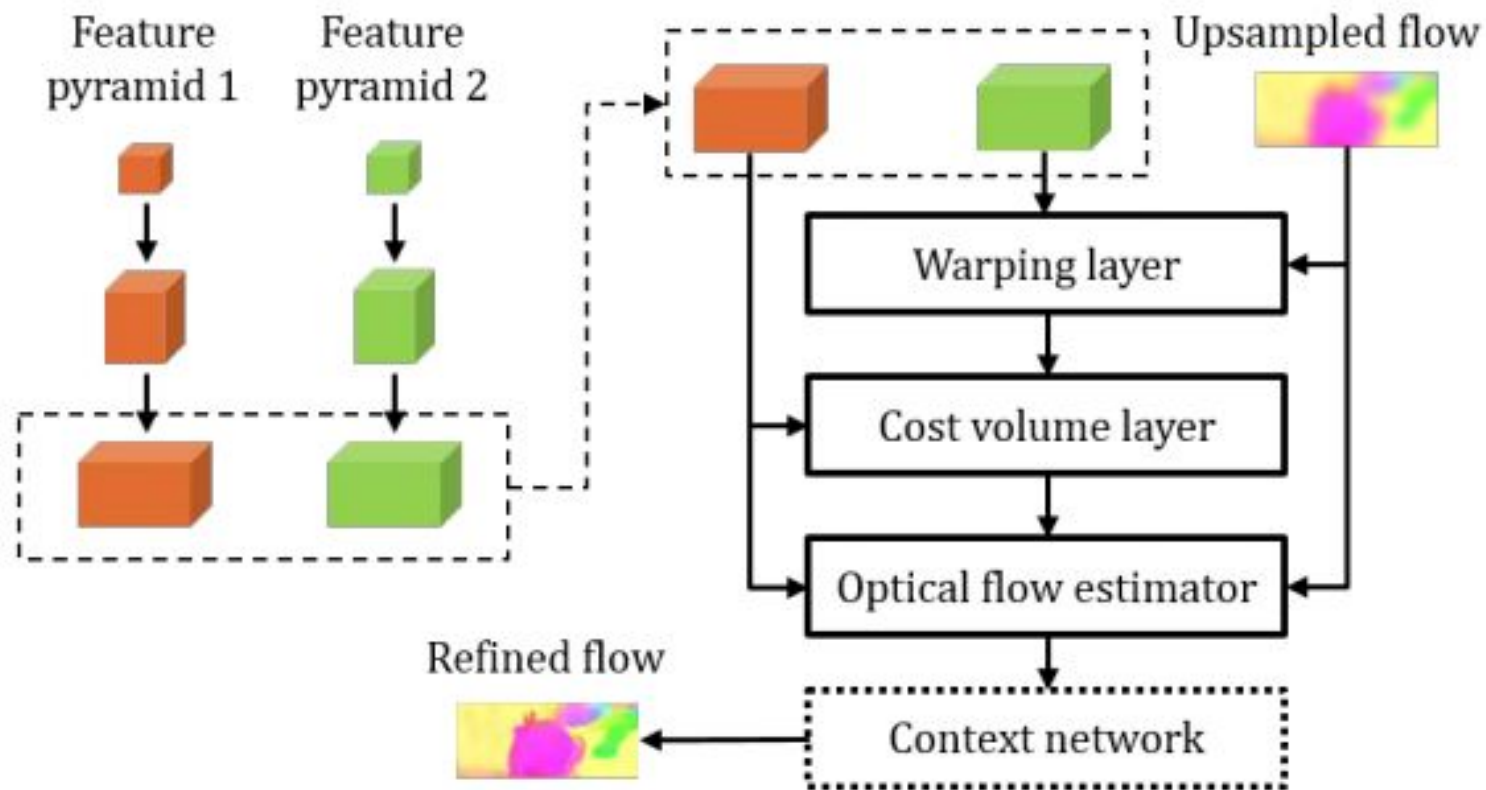
## FlowNet — decoder



# Введение PWC-Net (1)



# Введение PWC-Net (2)





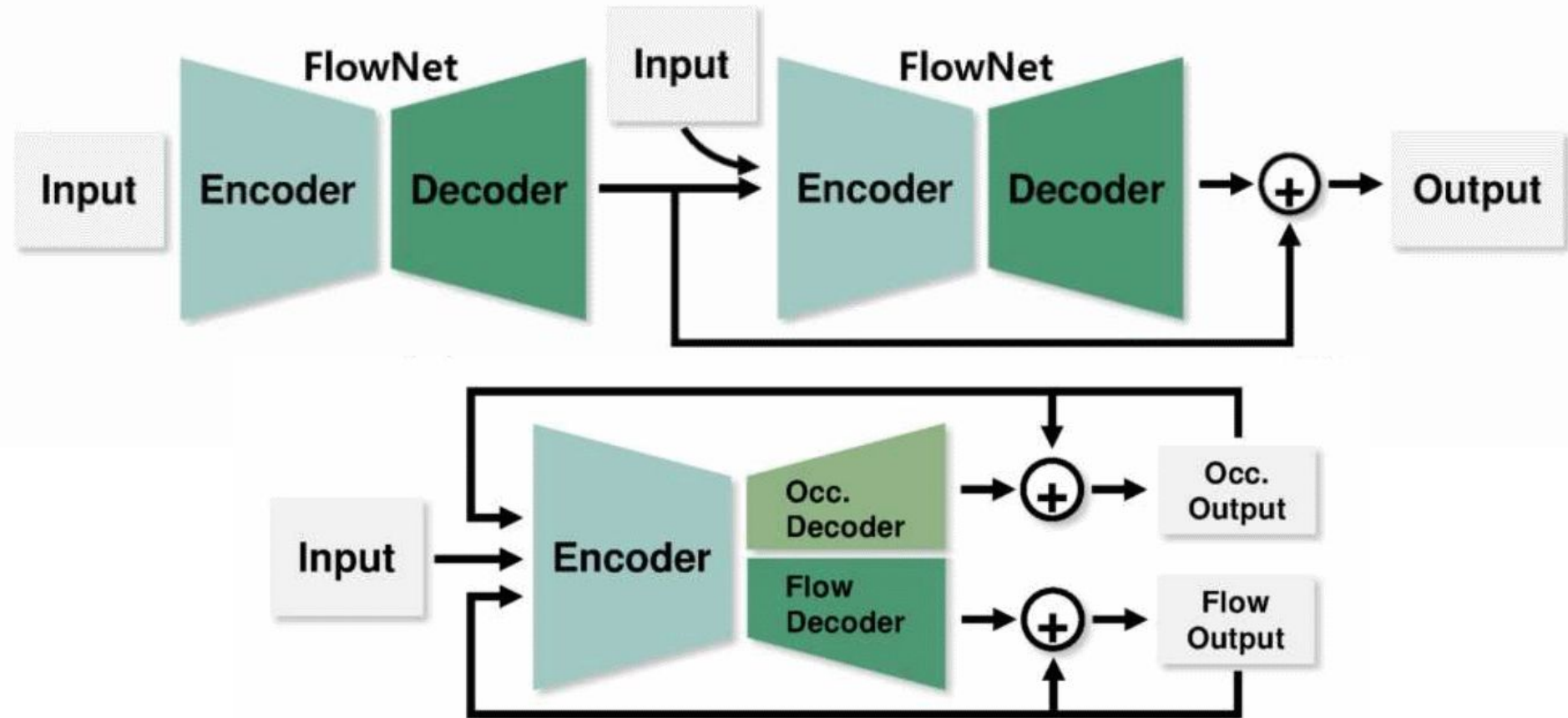
# Содержание

- Введение
- Методы поиска областей открытия/закрытия
  - **IRR-PWC**
  - FlowNet-CSSR
  - ContinualFlow
- Методы поиска оптического потока
- Заключение

- Прогоняют поток через одну и ту же сеть несколько раз, в основе — зарекомендовавшие себя архитектуры
- Одновременное вычисление потока и областей открытия/закрытия
- Прямое и обратное вычисление оптического потока
- Билатеральная фильтрация
- “Super-resolution” для областей о/з

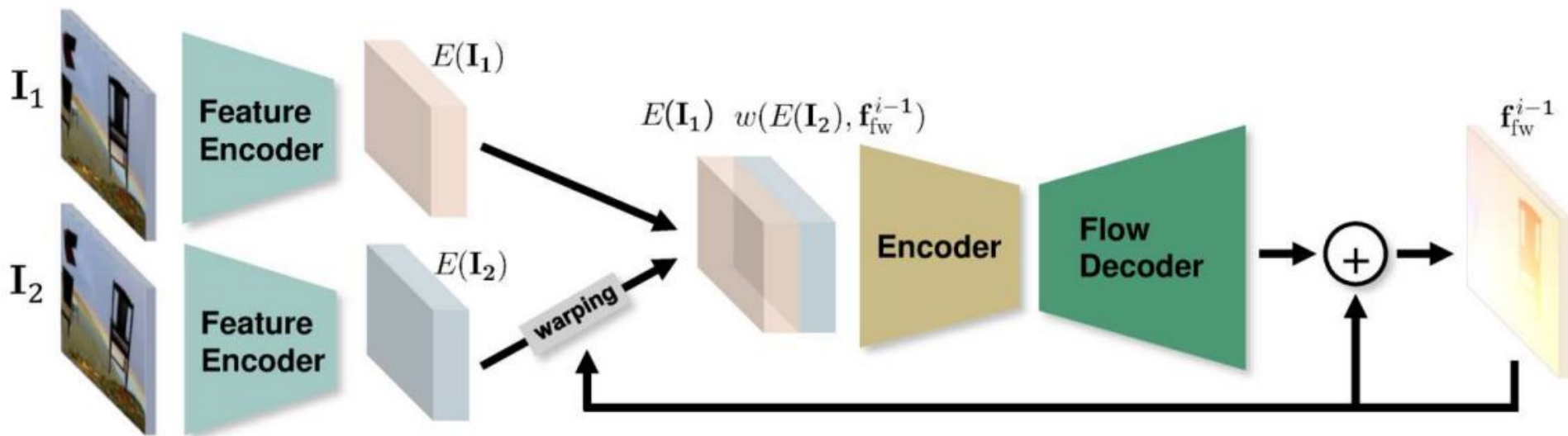
# IRR-PWC

## Общая архитектура



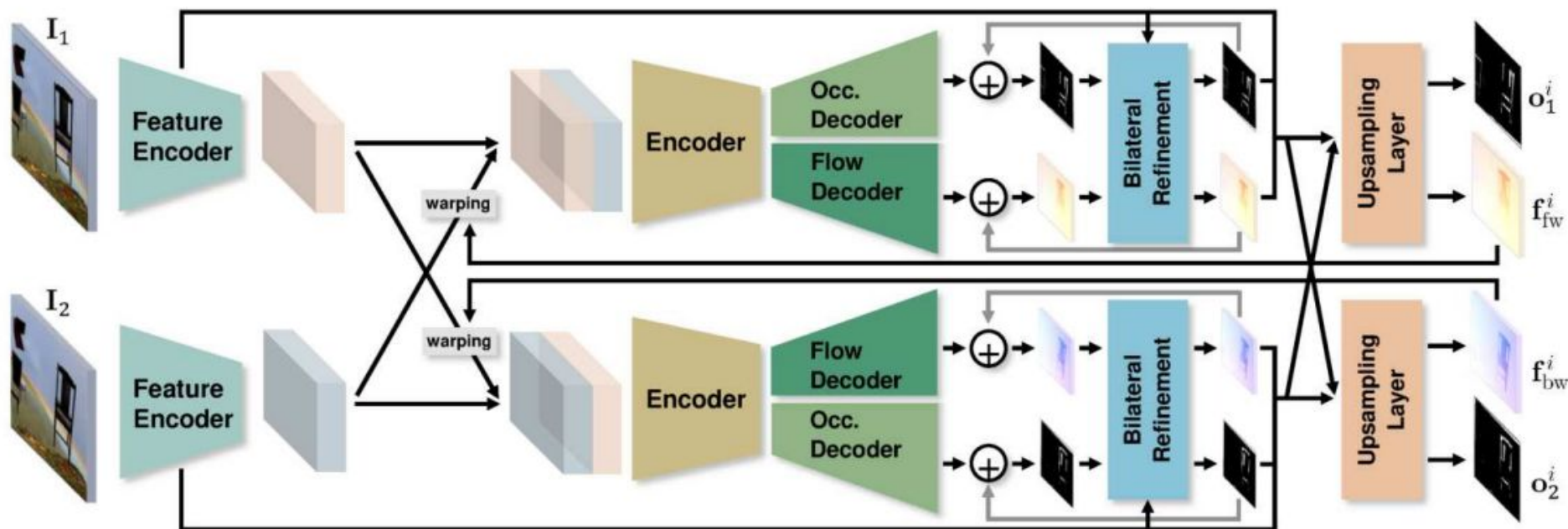
# IRR-PWC

## Применение IRR к FlowNetS



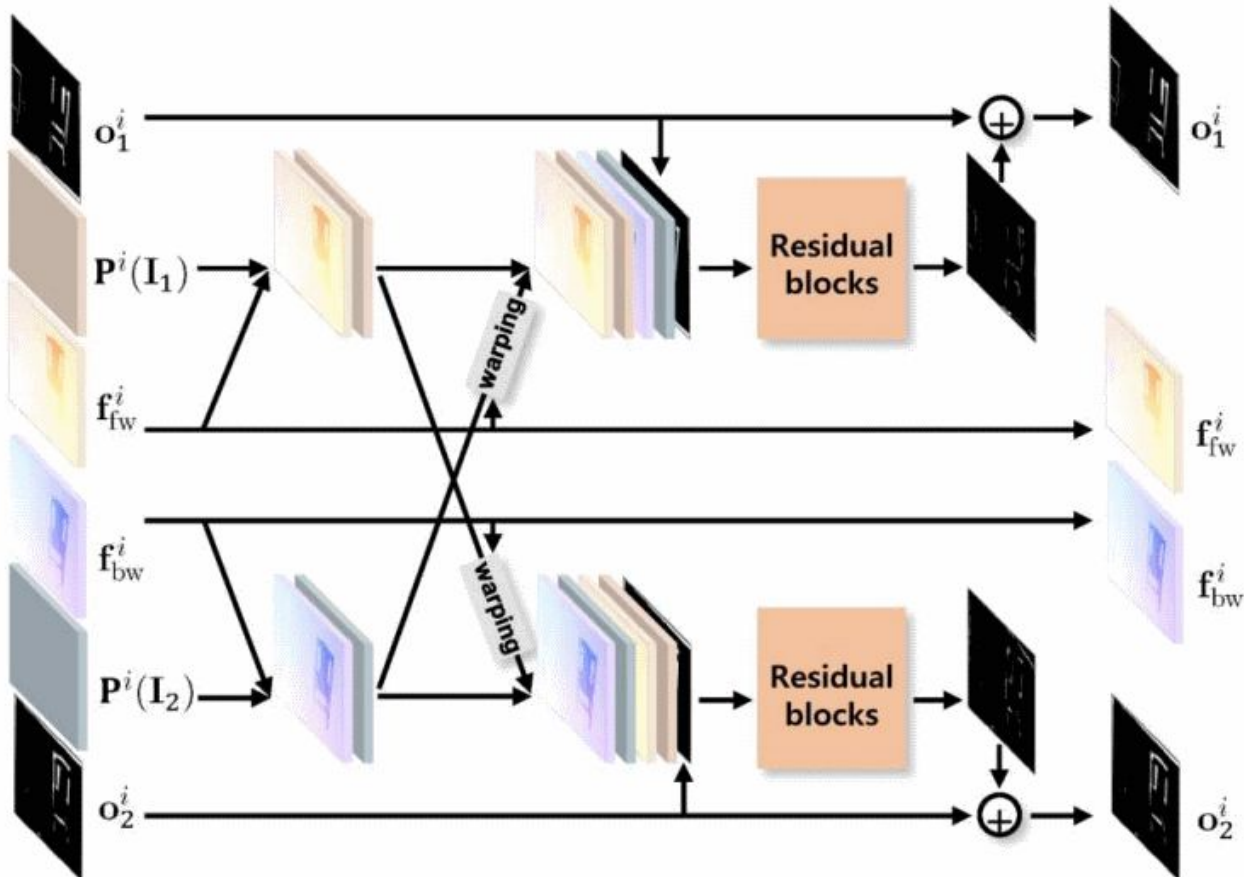
# IRR-PWC

## Дополнительные улучшения



# IRR-PWC

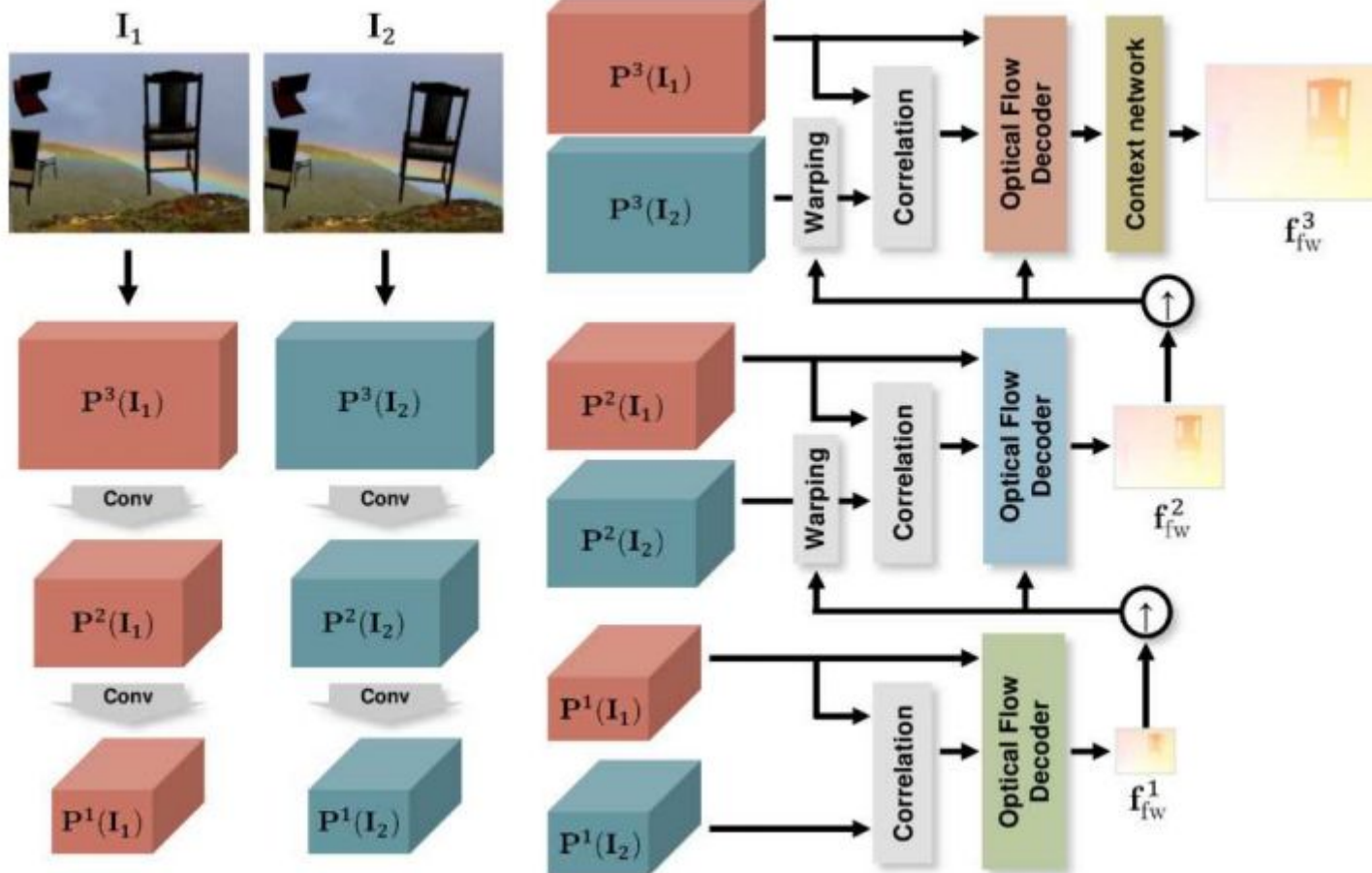
## Увеличение разрешения областей o/z



H. Junhwa and S. Roth, "Iterative residual refinement for joint optical flow and occlusion estimation," in *CVPR*, 2019

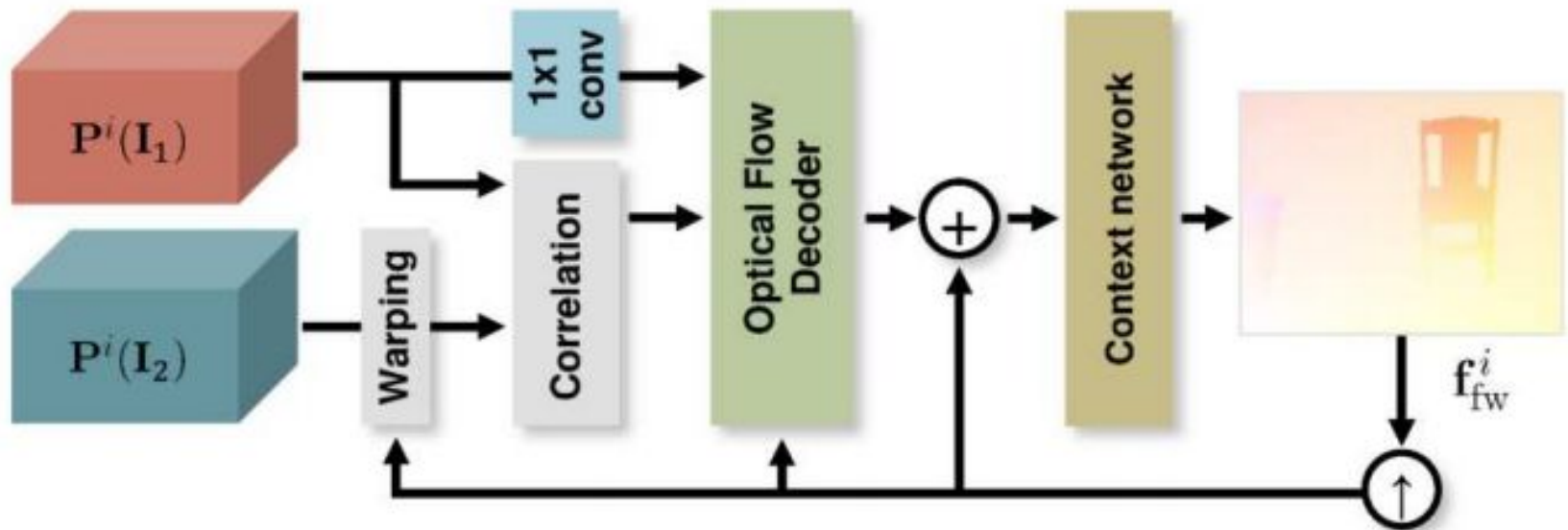
# IRR-PWC

## Применение к PWC-Net (1)



# IRR-PWC

## Применение к PWC-Net (2)





## Функция ошибки (1)

Для каждой итерации обработки  $i$ :

$$l_{\text{flow}}^i = \frac{1}{2} \sum (\| \mathbf{f}_{\text{fw}}^i - \mathbf{f}_{\text{fw,GT}} \|_2 + \| \mathbf{f}_{\text{bw}}^i - \mathbf{f}_{\text{bw,GT}} \|_2)$$

Взвешенная бинарная кросс-энтропия:

Области о/з  
в прямом и обратном  
направлении

$$l_{\text{occ}}^i = -\frac{1}{2} \sum (w_1^i o_1^i \log o_{1,\text{GT}} + \bar{w}_1^i (1 - o_1^i) \log(1 - o_{1,\text{GT}}) + w_2^i o_2^i \log o_{2,\text{GT}} + \bar{w}_2^i (1 - o_2^i) \log(1 - o_{2,\text{GT}}))$$

Нормализация  
по размеру кадра  
и площади  
областей

$$w_1^i = \frac{H \cdot W}{\sum o_1^i + \sum o_{1,\text{GT}}}$$

$$\bar{w}_1^i = \frac{H \cdot W}{\sum (1 - o_1^i) + \sum (1 - o_{1,\text{GT}})}$$

H. Junhwa and S. Roth, "Iterative residual refinement for joint optical flow and occlusion estimation," in *CVPR*, 2019

## Функция ошибки (2)

- $$l_{\text{FlowNet}} = \frac{1}{N} \sum_{i=1}^N \sum_{s=s_0}^S \alpha_s (l_{\text{flow}}^{i,s} + \lambda \cdot l_{\text{occ}}^{i,s})$$

$$l_{\text{PWC-Net}} = \frac{1}{N} \sum_{i=1}^N \alpha_i (l_{\text{flow}}^i + \lambda \cdot l_{\text{occ}}^i)$$

$S$  — количество upsampling-слоев FlowNetS

$\alpha_s$  — коэффициенты для потока в refinement-модуле

$\lambda$  — коэффициент, выравнивающий вклад потока и областей о/з на каждой итерации

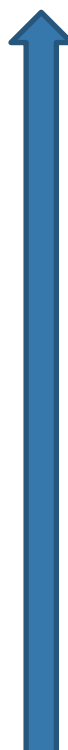
1. FlyingChairsOss: прямой и обратный поток, области открытия/закрытия, 600 000 итераций ( $S_{short}$  из FlowNet 2.0)
2. Дообучение на FlyingThings3D
3. Дообучение на Sintel или KITTI с улучшенным протоколом PWC-Net+

	Bi	Occ	IRR	Chairs Full	ChairsOcc Validation	Sintel Clean Training	Sintel Final Training	Rel. Param.
FlowNet [14]				2.39	2.27	4.35	5.44	0 %
	✓			2.43	2.30	4.40	5.53	0 %
		✓		2.29	2.18 (0.690)	4.26 (0.521)	5.51 (0.493)	+38.5%
			✓	2.36	2.22	3.77	5.00	0 %
	✓	✓		2.31	2.20 (0.691)	4.21 (0.515)	5.46 (0.488)	+38.5%
	✓		✓	2.14	2.00	3.45	4.96	0 %
		✓	✓	2.22	2.10 (0.689)	3.56 (0.507)	5.03 (0.486)	+38.5%
	✓	✓	✓	2.05	1.91 (0.699)	3.40 (0.528)	5.08 (0.502)	+38.5%
	✓	✓+	<b>1.92</b>	<b>1.77 (0.736)</b>	<b>3.32 (0.596)</b>	<b>4.92 (0.560)</b>	+40.7%	
PWC-Net [52]				2.03	1.89	3.13	4.41	0 %
	✓			2.06	1.87	2.98	4.14	0 %
		✓		1.94	1.79 (0.706)	3.16 (0.616)	4.35 (0.581)	+87.4%
			✓	2.01	1.83	2.79	4.10	-61.2%
	✓	✓		1.99	1.82 (0.696)	3.01 (0.618)	4.39 (0.581)	+87.4%
	✓		✓	2.08	1.90	2.80	4.13	-61.2%
		✓	✓	1.91	1.73 (0.700)	2.64 (0.630)	4.09 (0.593)	-34.7%
	✓	✓	✓	1.98	1.81 (0.698)	2.69 (0.633)	4.03 (0.598)	-34.7%
	✓	✓+	<b>1.67</b>	<b>1.48 (0.757)</b>	<b>2.34 (0.677)</b>	<b>3.95 (0.624)</b>	-26.4%	

H. Junhwa and S. Roth, "Iterative residual refinement for joint optical flow and occlusion estimation," in *CVPR*, 2019

# IRR-PWC

## MPI-Sintel (оптический поток)



Method	Training		Test		Parameters
	Clean	Final	Clean	Final	
ContinualFlow_ROB <sup>†§</sup> [42]	–	–	3.34	<b>4.53</b>	14.6 M
MFF <sup>§</sup> [46]	–	–	3.42	4.57	N/A
<b>IRR-PWC (Ours)</b>	(1.92)	(2.51)	3.84	4.58	6.36M
PWC-Net+ <sup>†</sup> [53]	(1.71)	(2.34)	3.45	4.60	8.75M
ProFlow <sup>§</sup> [37]	–	–	2.82	5.02	–
PWC-Net-ft-final [53]	(2.02)	(2.08)	4.39	5.04	8.75M
DCFlow [61]	–	–	3.54	5.12	–
FlowFieldsCNN [6]	–	–	3.78	5.36	5.00M
MR-Flow [59]	1.83	3.59	<b>2.53</b>	5.38	–
LiteFlowNet [24]	(1.35)	(1.78)	4.54	5.38	5.37M
S2F-IF [62]	–	–	3.50	5.42	–
SfM-PM [38]	–	–	2.91	5.47	–
FlowFields++ [49]	–	–	2.94	5.49	–
FlowNet2 [26]	(2.02)	(3.14)	3.96	6.02	162.5 M

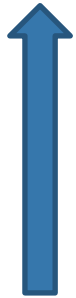
H. Junhwa and S. Roth, “Iterative residual refinement for joint optical flow and occlusion estimation,” in CVPR, 2019

# IRR-PWC

## MPI-Sintel (области открытия/закрытия)



Method	Type	Sintel Training	
		Clean	Final
<b>IRR-PWC (Ours)</b>	supervised	<b>0.712</b>	<b>0.669</b>
FlowNet-CSSR [27]	supervised	0.703	0.654
OccAwareFlow [58]	unsupervised	0.54	0.48
Back2FutureFlow [30]	unsupervised	0.49	0.44
MirrorFlow [25]	estimated	0.390	–



### Достоинства:

- Один из лучших по качеству оптического потока
- Лучший по качеству областей открытия/закрытия
- Размер сети небольшой из-за переиспользования модулей
- Выложены скрипты обучения и веса обученных моделей

### Недостаток:

- В 6 раз медленнее PWC-Net (5.5 fps vs 33.3 fps)

# Содержание

- Введение
- Методы поиска областей открытия/закрытия
  - IRR-PWC
  - **FlowNet-CSSR**
  - ContinualFlow
- Методы поиска оптического потока
- Заключение



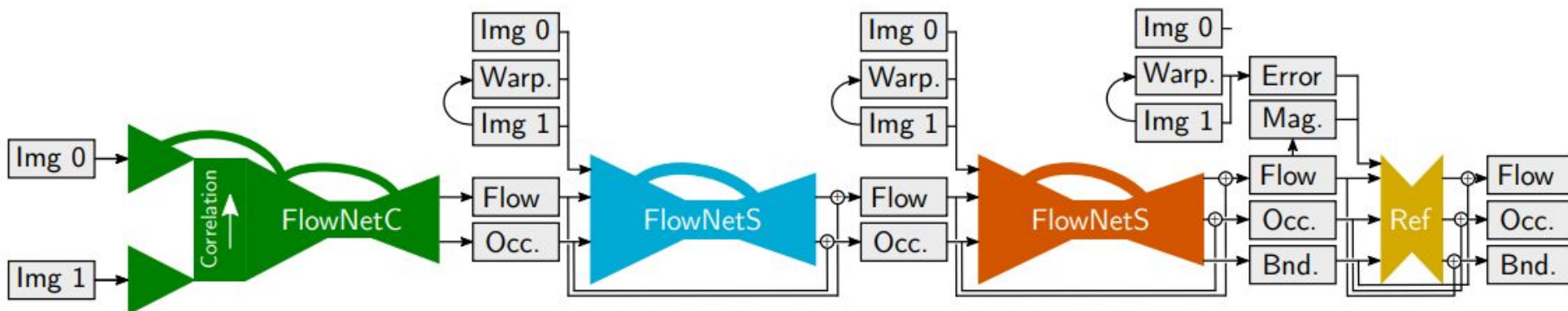
Построение сети аналогично FlowNet 2.0

Вычисляют:

- Оптический поток
- Области о/з
- Диспаратет
- Границы объектов по движению и глубине

# FlowNet-CSSR

## Архитектура (1)

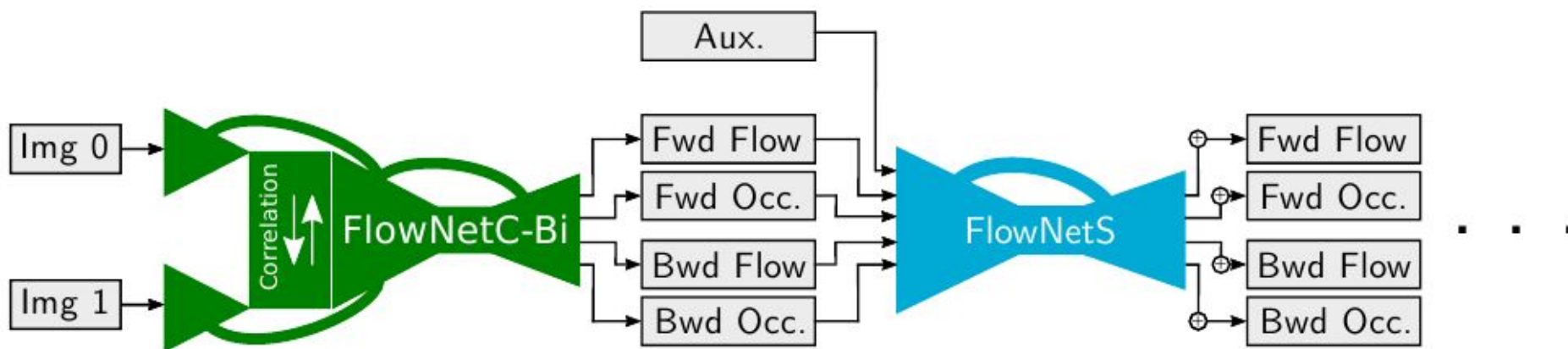


(a). Основная архитектура

E. Ilg et al., "Occlusions, motion and depth boundaries with a generic network for disparity, optical flow or scene flow estimation," in *ECCV*, 2018

# FlowNet-CSSR

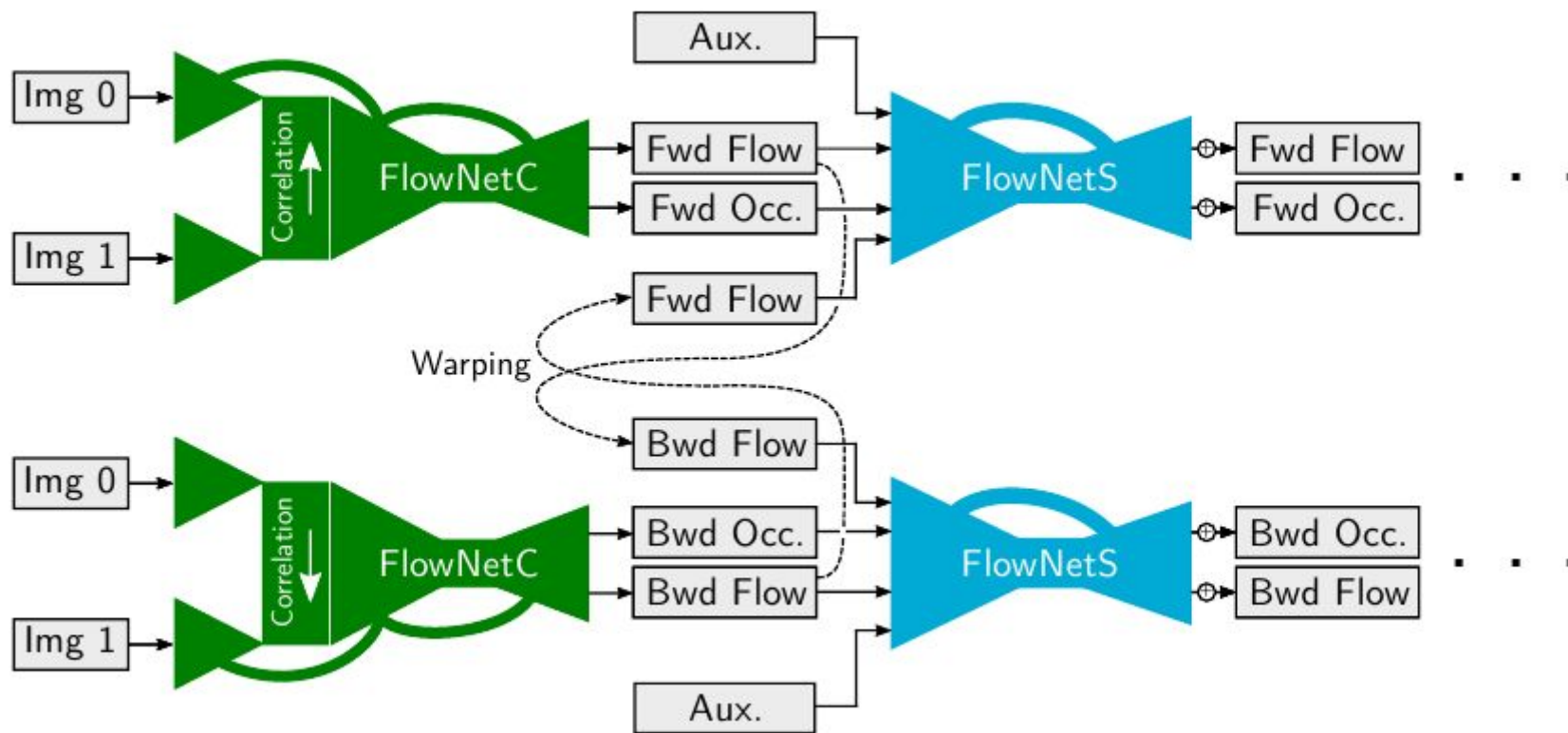
## Архитектура (2)



(b). Прямое и обратное вычисление оптического потока и областей открытия/закрытия

# FlowNet-CSSR

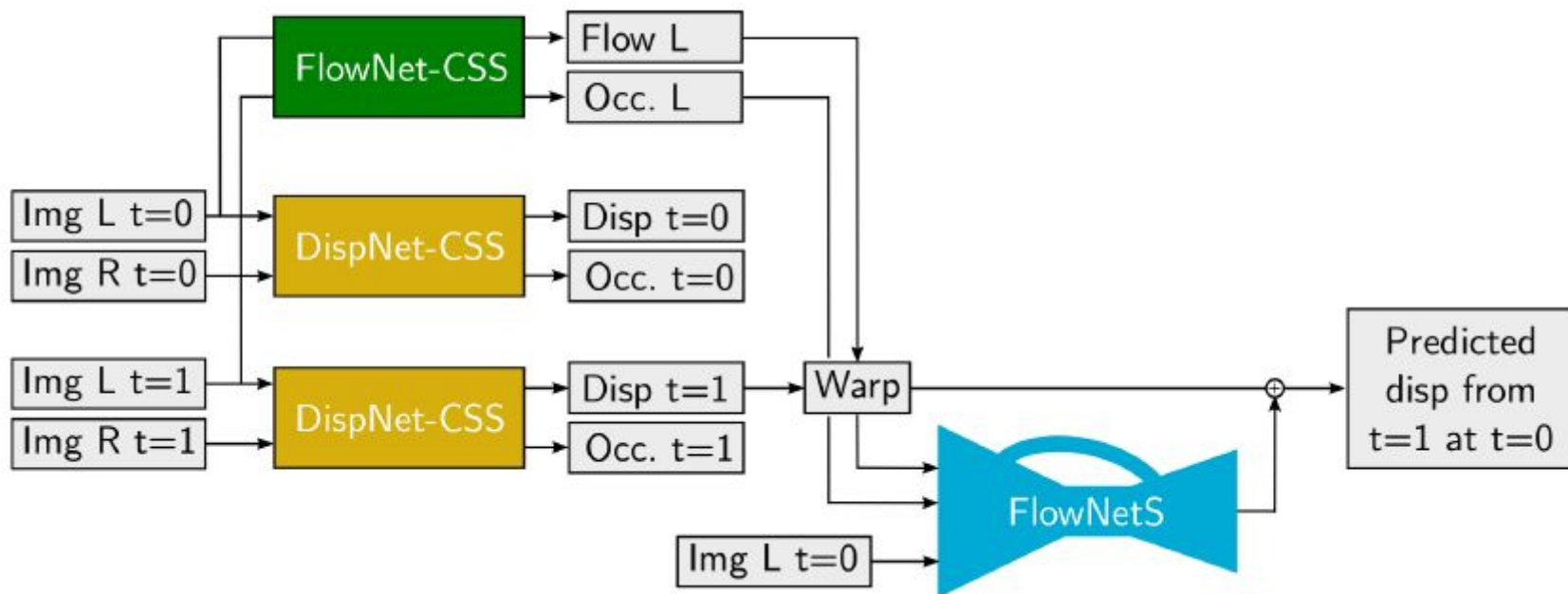
## Архитектура (3)



(с). Еще один вариант с прямым и обратным вычислением

# FlowNet-CSSR

## Архитектура (4)



(d). Использование диспаратета

Для каждой сети в стеке:

1. FlyingChairs — 600 000 итераций
2. FlyingThings3D — 250 000 итераций
3. FlyingChairsSDHom — (опционально) 250 000 итераций

Функция ошибки:

- Оптический поток — EPE
- Области о/з, границы движения — бинарная кросс-энтропия

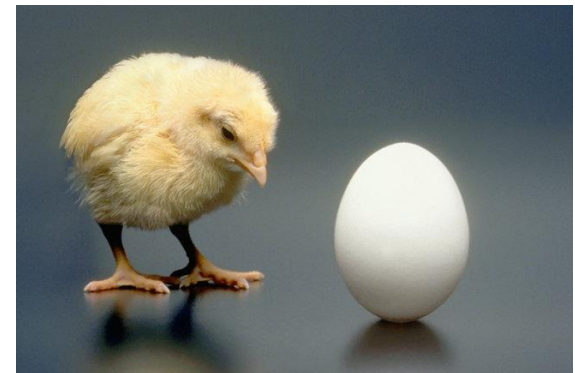
# FlowNet-CSSR

## Исследование разных конфигураций метода



Configuration	EPE	F-measure
FlowNetC estimating flow	3.21	-
FlowNetC estimating occlusions	-	<b>0.546</b>
FlowNetC estimating flow + occlusions	<b>3.20</b>	0.539
FlowNetC-Bi estimating fwd/bwd flow and fwd occlusions	3.26	0.542

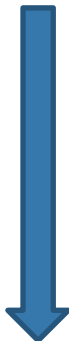
Configuration	EPE	F-measure
Only flow as in FlowNet2-CS [18]	2.28	-
+ occlusions (Figure 1(a))	<b>2.25</b>	<b>0.590</b>
+ bwd direction (Figure 1(b))	2.77	0.572
+ mutual warping (Figure 1(c))	2.25	0.589



E. Ilg et al., "Occlusions, motion and depth boundaries with a generic network for disparity, optical flow or scene flow estimation," in *ECCV*, 2018

# FlowNet-CSSR

Сравнение методов: области 0/3



Method	Type	F-Measure	
		clean	final
FlowNet2 [18]	consistency	0.377	0.348
MirrorFlow [17]	estimated	0.390	0.348
S2DFlow [27]	estimated	0.470	0.403
Ours	estimated	<b>0.703</b>	<b>0.654</b>



# FlowNet-CSSR

## Сравнение методов: ОПТИЧЕСКИЙ ПОТОК



Method	Sintel (clean)		Sintel (final)		KITTI (2012)		KITTI (2015)		Runtime (s)
	AEE		AEE		<i>AEE</i>	<i>OUT-noc</i>	<i>AEE</i>	<i>F1-all</i>	
	<i>train</i>	<i>test</i>	<i>train</i>	<i>test</i>	<i>train</i>	<i>test</i>	<i>train</i>	<i>test</i>	
<b>Standard</b>									
EpicFlow [35]	2.27	4.12	3.56	6.29	<b>3.09</b>	7.88%	9.27	26.29%	42
FlowfieldsCNN [2]	-	3.78	-	5.36	-	4.89%	-	18.68%	23
DCFlow [50]	-	<b>3.54</b>	-	5.12	-	-	-	14.86%	9
<b>CNN based</b>									
FlowNet2 [18]	<b>2.02</b>	3.96	<b>3.14</b>	6.02	4.09	-	10.06	-	0.123
FlowNet2-ft [18]	(1.45)	4.16	(2.01)	5.74	(1.28)	-	(2.30)	11.48%	0.123
SpyNet [34]	4.12	6.69	5.57	8.43	9.12	-	-	-	<b>0.016</b>
SpyNet-ft [34]	(3.17)	6.64	(4.32)	8.36	(4.13)	12.31%	-	35.07%	<b>0.016</b>
PWC-Net [42]	2.55	-	3.93	-	4.14	-	10.35	33.67%	0.030
PWC-Net-ft [42]	(2.02)	4.39	(2.08)	<b>5.04</b>	-	4.22%	(2.16)	9.80%	0.030
<b>With joint occ est.</b>									
MirrorFlow [17]	-	3.32	-	6.07	-	4.38%	-	10.29%	660
S2D flow [27]	-	18.48	-	6.82	-	-	-	-	2280
Our FlowNet-CSS	2.08	3.94	3.61	6.03	3.69	-	<b>9.33</b>	-	0.068
Our FlowNet-CSS-ft	(1.47)	4.35	(2.12)	5.67	(1.19)	<b>3.45%</b>	(1.79)	<b>8.60%</b>	0.068
Our FlowNet-css	2.65	-	4.05	-	5.05	-	11.74	-	0.033

# Occlusions, Motion and Depth Boundaries with a Generic Network for Disparity, Optical Flow or Scene Flow Estimation

---

Eddy Ilg\*, Tonmoy Saikia\*, Margret Keuper and  
Thomas Brox

\* equal contribution

### Достоинства:

- Универсальное решение многих задач
- Лучший результат по областям о/з на момент публикации (Sintel)
- Лучший результат по диспаритету (KITTI)
- Хорошая скорость: 14.7 fps

### Недостаток:

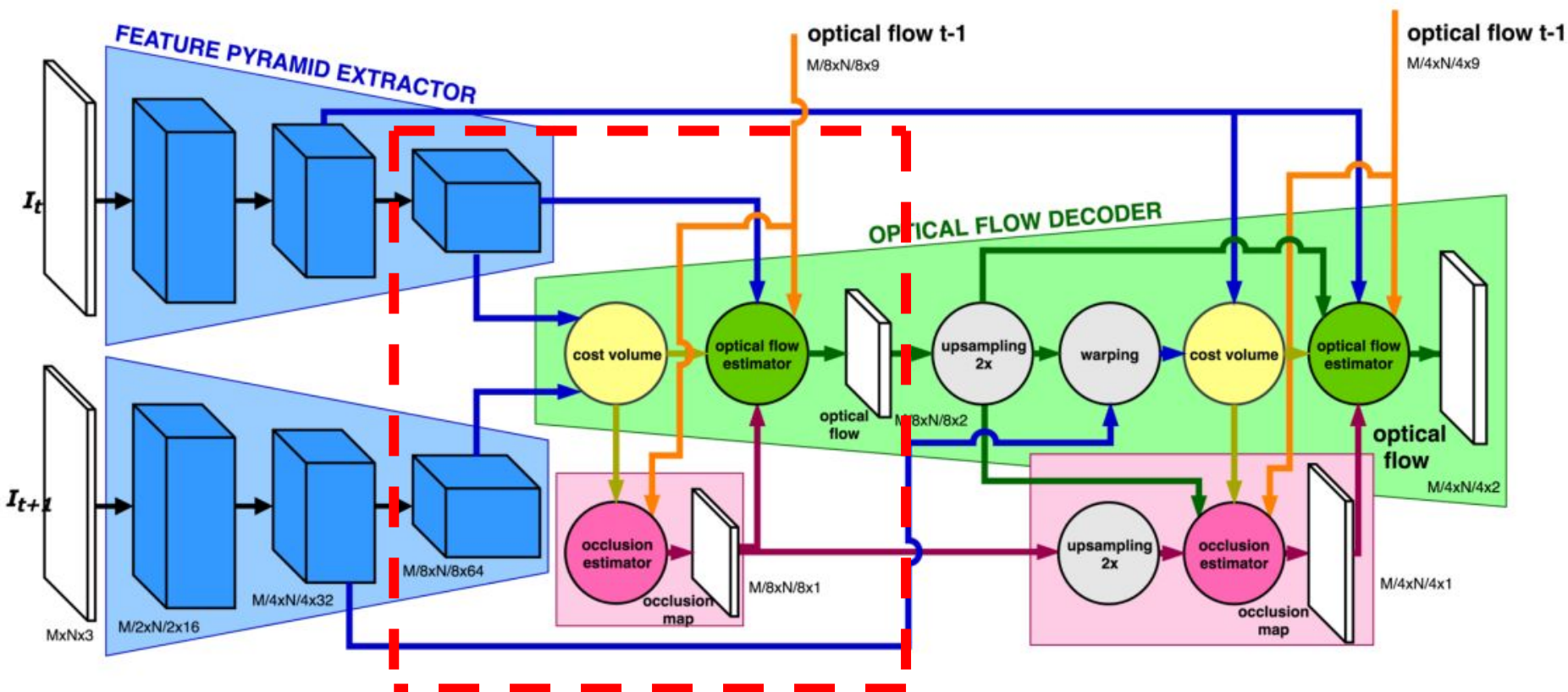
- Не выложены скрипты для обучения моделей

# Содержание

- Введение
- Методы поиска областей открытия/закрытия
  - IRR-PWC
  - FlowNet-CSSR
  - **ContinualFlow**
- Методы поиска оптического потока
- Заключение

# ContinualFlow

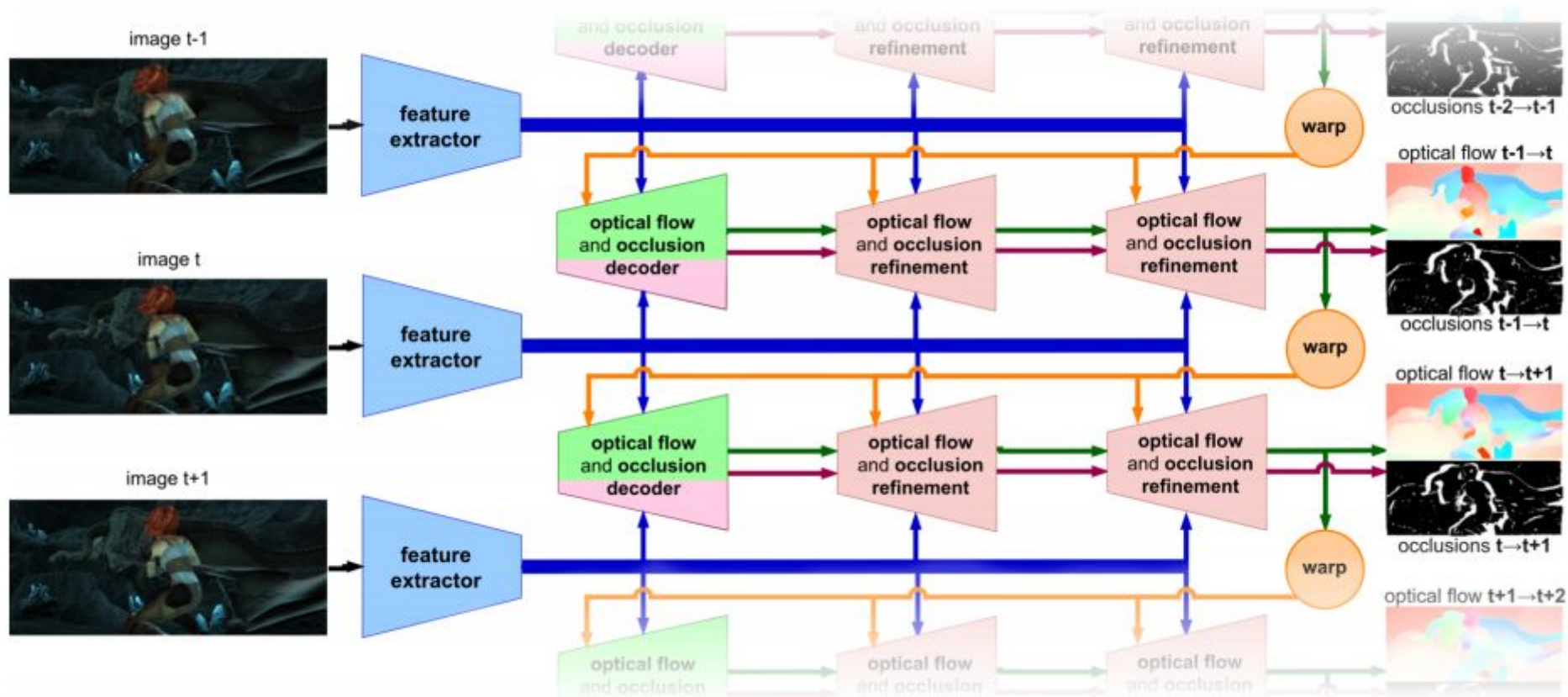
## PWC-Net + области открытия/закрытия



Один уровень пирамиды

# ContinualFlow

## Использование временного домена (1)



## Использование временного домена (2)

- Недифференцируемое прямое преобразование потока

$$\hat{F}_{t-1}(\mathbf{x} + \text{round}(F_{t-1}(\mathbf{x}))) = F_{t-1}(\mathbf{x})$$

- Дифференцируемое обратное преобразование  
(требуется дополнительно вычислить обратный поток)

# ContinualFlow

## Сравнение различных конфигураций



	FlyingThings					KITTI'15 noc		KITTI'15 occ		Sintel Clean			Sintel Final		
	all	occ-bg	occ-fg	noc-bg	noc-fg	Fl-all	all	Fl-all	all	all	occ	noc	all	occ	noc
common: baseline	<b>Occlusion map learning</b>														
+OC	22.79	25.31	53.88	10.64	26.78	37.73	7.82	43.56	14.16	3.45	9.29	2.38	5.36	12.03	4.17
	<b>18.01</b>	<b>18.27</b>	<b>47.53</b>	<b>7.10</b>	<b>20.13</b>	<b>23.98</b>	<b>5.22</b>	<b>31.12</b>	<b>10.60</b>	<b>2.45</b>	<b>7.46</b>	<b>1.53</b>	<b>4.02</b>	<b>9.99</b>	<b>2.91</b>
common: baseline+OC	<b>The specialised refinement block</b>														
+R	18.01	18.27	47.53	<b>7.10</b>	<b>20.13</b>	23.98	5.22	31.12	10.60	2.45	7.46	1.53	4.02	9.99	2.91
	<b>17.80</b>	<b>17.49</b>	<b>45.90</b>	7.31	21.46	<b>21.14</b>	<b>4.78</b>	<b>28.61</b>	<b>9.83</b>	<b>2.30</b>	<b>7.11</b>	<b>1.42</b>	<b>3.87</b>	<b>9.68</b>	<b>2.76</b>
common: baseline+OC+W <sub>bf</sub>	<b>Number of refinement blocks</b>														
+1xRW <sub>bf</sub>	14.28	14.24	35.58	5.82	<b>17.56</b>	<b>21.72</b>	<b>4.41</b>	<b>29.48</b>	<b>9.33</b>	<b>2.26</b>	<b>6.71</b>	<b>1.47</b>	<b>3.76</b>	<b>8.93</b>	2.80
+2xRW <sub>bf</sub>	<b>14.26</b>	<b>14.13</b>	<b>35.60</b>	5.78	17.62	21.77	4.45	29.62	9.35	<b>2.26</b>	6.72	<b>1.47</b>	<b>3.76</b>	8.96	<b>2.79</b>
+3xRW <sub>bf</sub>	14.30	<b>14.13</b>	35.71	<b>5.75</b>	17.77	21.98	4.50	29.86	9.40	<b>2.26</b>	6.74	<b>1.47</b>	3.77	8.99	2.80
+5xRW <sub>bf</sub>	14.43	14.24	36.16	<b>5.75</b>	17.93	22.48	4.58	30.35	9.49	2.28	6.80	1.48	3.80	9.03	2.83
common: baseline+OC+W <sub>bf</sub> +RW <sub>bf</sub>	<b>Multi-frame sequence initialisation</b>														
2 frames	-	-	-	-	-	25.08	5.50	32.59	11.56	2.48	7.72	1.48	3.84	9.64	2.75
2 frames+2pass	-	-	-	-	-	<b>23.06</b>	<b>5.03</b>	<b>30.92</b>	<b>11.00</b>	<b>2.41</b>	<b>7.60</b>	<b>1.41</b>	<b>3.74</b>	<b>9.48</b>	<b>2.66</b>
3 frames	-	-	-	-	-	21.72	4.41	29.48	9.33	2.26	6.71	<b>1.47</b>	3.76	8.93	2.80
3 frames+2pass	-	-	-	-	-	<b>21.65</b>	<b>4.36</b>	<b>29.42</b>	<b>9.23</b>	<b>2.26</b>	<b>6.71</b>	1.48	<b>3.73</b>	<b>8.92</b>	<b>2.76</b>
4 frames	-	-	-	-	-	<b>21.53</b>	4.30	<b>29.32</b>	9.05	<b>2.23</b>	6.59	1.46	3.75	8.83	2.82
4 frames+2pass	-	-	-	-	-	21.54	<b>4.30</b>	29.33	<b>9.02</b>	2.24	<b>6.59</b>	<b>1.46</b>	<b>3.73</b>	<b>8.80</b>	<b>2.80</b>
5 frames	-	-	-	-	-	<b>21.48</b>	<b>4.25</b>	<b>29.27</b>	<b>8.92</b>	<b>2.21</b>	<b>6.51</b>	<b>1.45</b>	3.80	8.85	2.87
5 frames+2pass	-	-	-	-	-	<b>21.48</b>	<b>4.25</b>	29.28	<b>8.92</b>	<b>2.21</b>	6.52	1.46	<b>3.79</b>	<b>8.83</b>	<b>2.86</b>



$$\mathcal{L}_F^s = \sum_{\mathbf{x}} \gamma(\mathbf{x}) \|F^s(\mathbf{x}) - F_{gt}^s(\mathbf{x})\|_2$$

$\gamma(x)$  — маска существования GT для потока

# ContinualFlow

## Функция ошибки: области о/з



Используется взвешенная бинарная кросс-энтропия:

$$\begin{aligned} \mathcal{L}_O^s = & -w_{noc} \sum_{\mathbf{x}: O_{gt}(\mathbf{x})=1} \rho(\mathbf{x}) \log \Pr(O(\mathbf{x}) = 1|X) \\ & -w_{occ} \sum_{\mathbf{x}: O_{gt}(\mathbf{x})=0} \rho(\mathbf{x}) \log \Pr(O(\mathbf{x}) = 0|X) \end{aligned}$$

$w_{noc}$ ,  $w_{occ}$  — доли пикселей, не принадлежащих и принадлежащих области открытия/закрытия  
 $\rho(x)$  — маска существования GT

- $$\mathcal{L} = \sum_{s=1}^S \alpha^s \mathcal{L}_F^s + \alpha_O \sum_{s=1}^S \alpha^s \mathcal{L}_O^s$$

$S$  — количество уровней в PWC-Net

$\alpha^s$  — коэффициент для каждого уровня пирамиды

$\alpha_O$  — вклад областей о/з в ошибку

# ContinualFlow

## Обучение



Предобучение на парах изображений:

1. FlyingChairs — 1 200 000 итераций
2. FlyingThings3D — 500 000 итераций

Дообучение с временными связями — 500 000 итераций:  
Driving, KITTI, virtualKITTI, Sintel, HD1K,  
FlyingChairsSDHom

# ContinualFlow

## Результаты: KITTI



F1 (%)	KITTI'15 occ (%)			KITTI'15 noc (%)		
	bg	fg	all	bg	fg	all
<b>ContinualFlow_ROB</b>	<b>8.54</b>	17.48	<b>10.03</b>	<b>5.90</b>	14.99	7.55
LFNet_ROB [anon]	11.18	10.20	11.01	6.14	6.87	<b>6.27</b>
PWC-Net_ROB [35]	11.22	13.69	11.63	7.12	10.29	7.69
ProFlow_ROB [23]	14.15	21.82	15.42	8.44	17.90	10.15
FF++_ROB [33]	15.32	19.27	15.97	7.82	15.33	9.18
ResPWCR_ROB [anon]	16.63	16.18	16.55	10.10	12.23	10.49
AugFNG_ROB [anon]	19.77	<b>9.95</b>	18.14	13.75	<b>6.71</b>	12.47
DMF_ROB [40]	30.74	30.07	30.63	19.32	25.60	20.46

# ContinualFlow

## Результаты: Sintel



	Sintel Final			Sintel Clean		
	all	noc	occ	all	noc	occ
<b>ContinualFlow_ROB</b>	<b>4.528</b>	2.723	<b>19.248</b>	3.341	1.752	<b>16.292</b>
PWC-Net_ROB [35]	4.903	<b>2.454</b>	24.878	3.897	1.726	21.637
ProFlow_ROB [23]	5.015	2.659	24.192	<b>2.709</b>	<b>1.013</b>	16.549
AugFNG_ROB [anon]	5.500	2.978	26.052	3.606	1.603	19.939
LFNet_ROB [anon]	5.966	3.278	27.893	4.815	2.333	25.065
FF++_ROB [33]	6.496	2.990	35.057	3.953	1.148	26.836
ResPWCR_ROB [anon]	6.530	3.849	28.371	5.674	3.138	26.380
DMF_ROB [40]	7.475	3.575	39.245	5.368	1.742	34.899

# ContinualFlow

## Общая таблица Sintel



	EPE all
GroundTruth <sup>[1]</sup>	0.000
SelfFlow <sup>[2]</sup>	4.262
VCN <sup>[3]</sup>	4.520
ContinualFlow_ROB <sup>[4]</sup>	4.528
MFF <sup>[5]</sup>	4.566
IRR-PWC <sup>[6]</sup>	4.579
PWC-Net+ <sup>[7]</sup>	4.596

### Достоинства:

- Первый метод обучения с учителем, успешно использовавший более 2 кадров
- Один из лучших результатов по потоку на Sintel

### Недостатки:

- В 5 раз дольше PWC-Net
- Невысокое качество областей открытия/закрытия



# Содержание

- Введение
- Методы поиска областей открытия/закрытия
- Методы поиска оптического потока
  - **MFF**
  - PWC-Net+
- Заключение

- Как учитывать предыдущие кадры?
  - Простое затухание влияния предыдущих
  - Эмпирический выбор наилучшего кандидата
  - Fusion-сеть
- Можно использовать любой двухкадровый метод
- Для каждого предыдущего кадра вычисляется прямой и обратный поток, объединение потоков fusion-сетью

Получение кандидата на поток с помощью предыдущего кадра

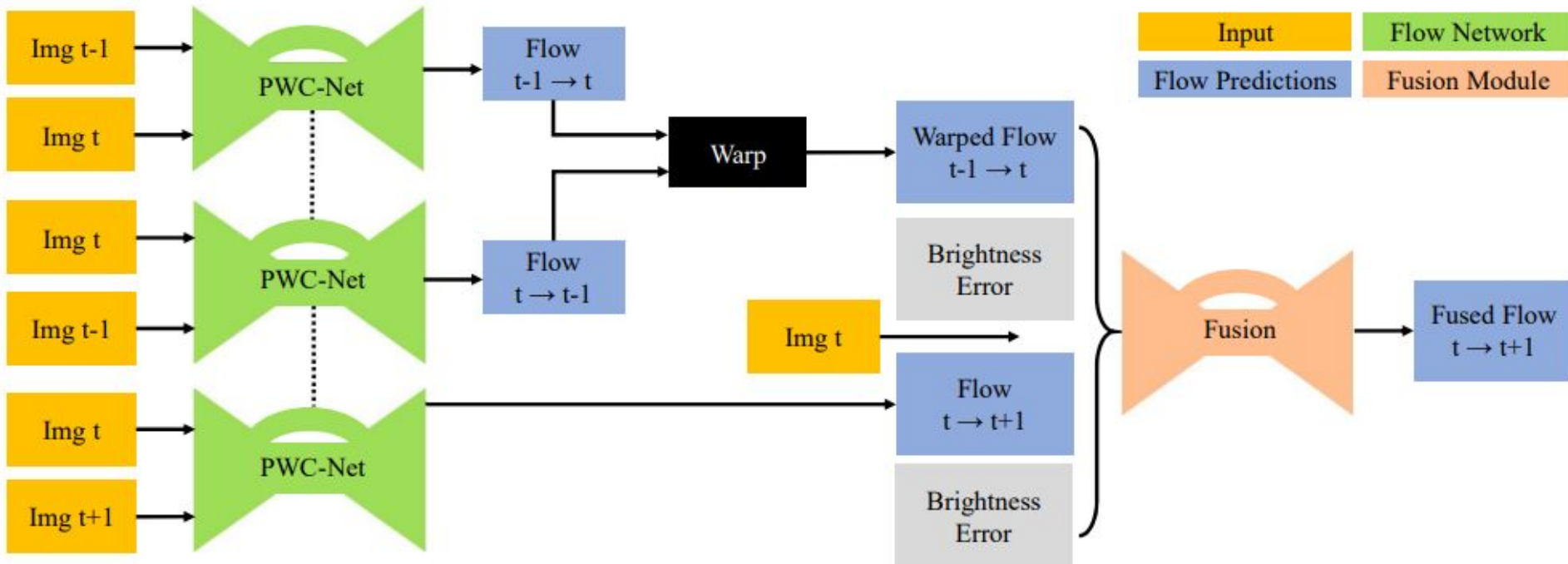
$$\hat{F}_{t \rightarrow t+1} = W^{-1}(F_{t-1 \rightarrow t}, F_{t \rightarrow t-1})$$

Oracle flow — исследование потенциального улучшения с помощью сравнения кандидатов с GT

	FlowNetS	Oracle	PWC-Net	Oracle
EPE All	6.12	4.35	2.34	1.80
EPE Inside	4.03	2.62	1.60	1.20
EPE Outside	28.97	21.83	10.43	7.90
EPE Occlusion	7.44	4.63	2.41	1.82
	FlowNetS	Oracle	PWC-Net	Oracle
EPE All	2.07	1.89	1.19	1.03
EPE Inside	1.91	1.75	1.19	1.03
EPE Outside	11.47	9.79	8.00	6.90
EPE Occlusion	8.16	7.00	5.83	4.88

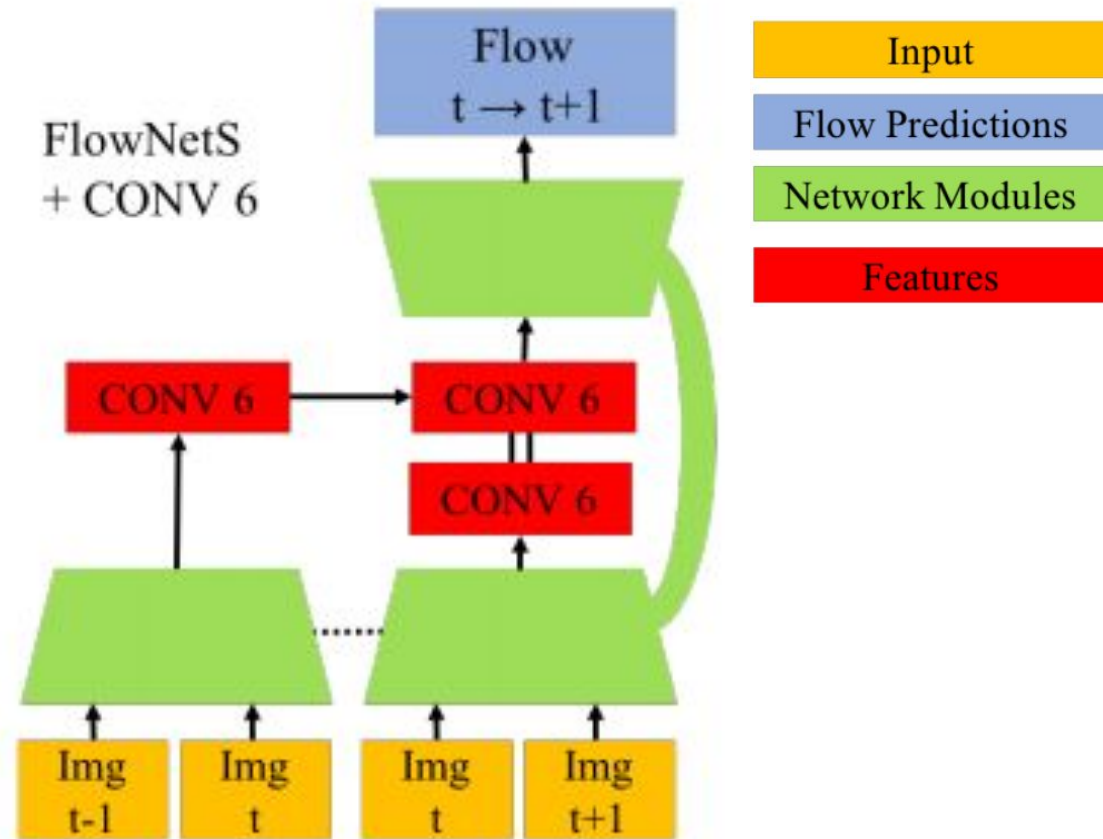
# MFF

## Архитектура



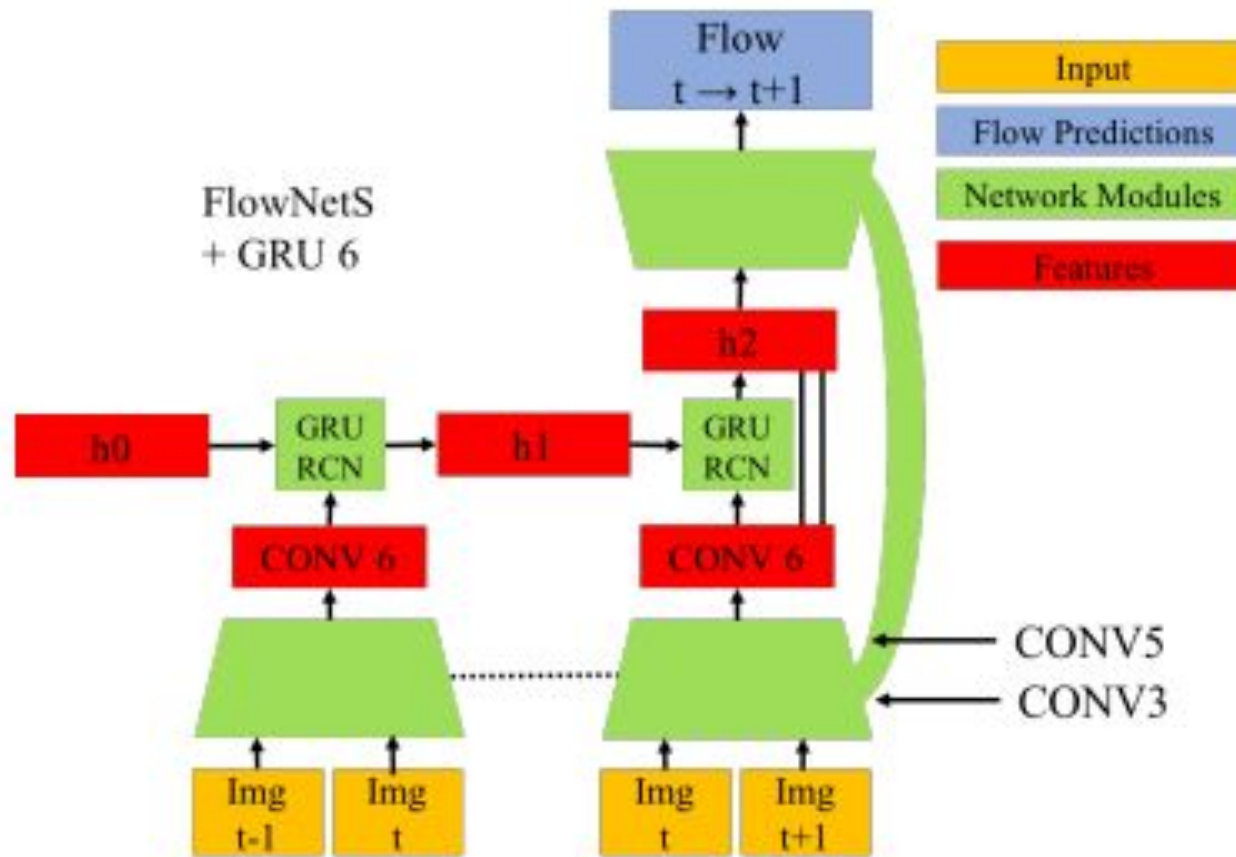
# MFF

## Baseline: FlowNetS++



# MFF

## Baseline: GRU-RCN



### Обучение:

- FlyingChairs
- FlyingThings3D

### Сравнение с бейзлайном:

- virtualKITTI (2016)
- Monkaa (2016)

### Сравнение с остальными алгоритмами:

- KITTI
- Sintel

•

$$L(\Theta) = \sum_{l=l_0}^L \alpha_l \sum_{\mathbf{x}} (|\mathbf{w}_{\Theta}^l(\mathbf{x}) - \mathbf{w}_{GT}^l(\mathbf{x})| + \epsilon)^q + \gamma \|\Theta\|_2^2$$

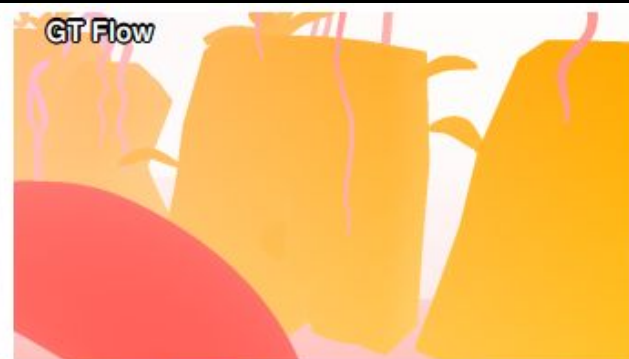
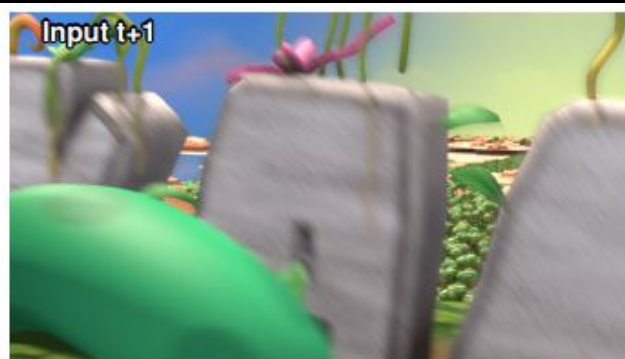
$L_2$ -регуляризация

$\alpha_l$  — коэффициент уровня пирамиды

$\epsilon$  — устойчивая ошибка Шарбонье

$q$  — параметр толерантности к ошибкам





	FlowNetS	FlowNetS++	GRU 3	GRU 4	GRU 5	GRU 6	Fusion	Oracle
EPE All	6.12	5.90	5.26	5.40	5.15	5.32	<b>5.00</b>	4.35
EPE Inside	4.03	3.87	3.61	3.64	3.58	3.59	<b>3.14</b>	2.62
EPE Outside	28.97	27.57	23.26	24.60	<b>22.28</b>	24.25	25.15	21.83
EPE Occlusion	7.44	7.11	<b>5.93</b>	6.27	5.82	6.18	6.14	4.63
	FlowNetS	FlowNetS++	GRU 3	GRU 4	GRU 5	GRU 6	Fusion	Oracle
EPE All	2.07	2.06	2.56	2.45	2.34	2.27	<b>1.97</b>	1.89
EPE Inside	1.91	1.89	2.37	2.27	2.16	2.09	<b>1.8</b>	1.75
EPE Outside	11.47	<b>11.12</b>	13.13	12.76	12.65	12.43	11.43	9.79
EPE Occlusion	8.16	8.02	9.17	9.07	8.84	8.71	<b>7.89</b>	7.00

	PWC-Net	GRU 3	GRU 4	GRU 5	GRU 6	Fusion	Oracle
EPE All	2.34	2.17	2.13	2.12	2.16	<b>2.07</b>	1.80
EPE Inside	1.60	1.44	1.41	1.40	1.42	<b>1.37</b>	1.20
EPE Outside	10.43	10.01	9.94	10.02	9.86	<b>9.71</b>	7.90
EPE Occlusion	2.41	2.29	<b>2.24</b>	<b>2.24</b>	2.26	2.27	1.82
	PWC-Net	GRU 3	GRU 4	GRU 5	GRU 6	Fusion	Oracle
EPE All	1.19	1.26	1.23	1.27	1.27	<b>1.18</b>	1.03
EPE Inside	1.19	1.26	1.23	1.27	1.27	<b>1.18</b>	1.03
EPE Outside	8.00	8.16	8.11	8.55	8.42	<b>7.94</b>	6.90
EPE Occlusion	5.83	5.97	5.90	5.80	5.79	<b>5.67</b>	4.88

### Достоинства:

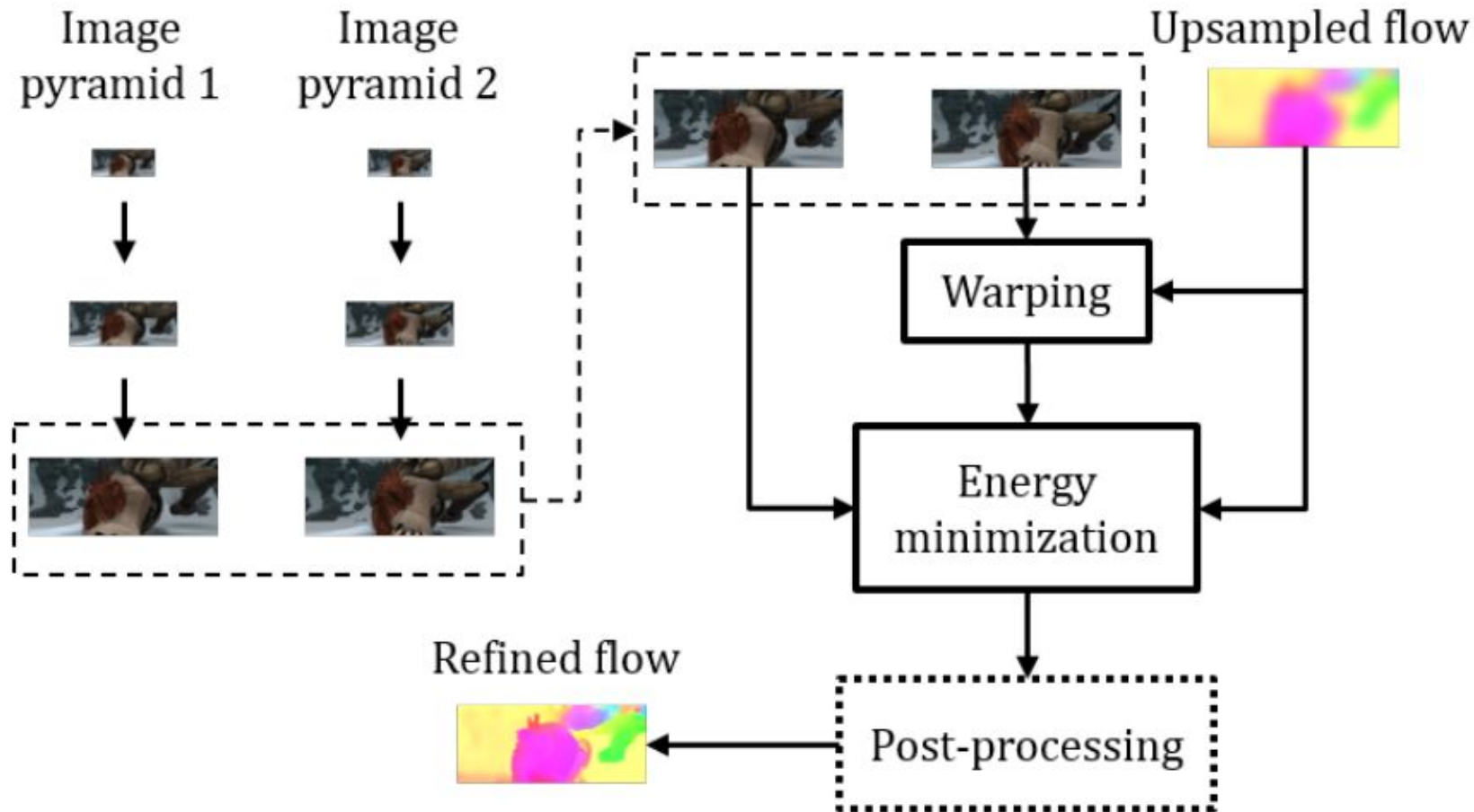
- Один из лучших методов по оптическому потоку
- Простая идея, общий метод
- Выложены скрипты обучения и обученные модели

### Недостаток:

- В 2 раза медленнее по сравнению с PWC-Net

# Содержание

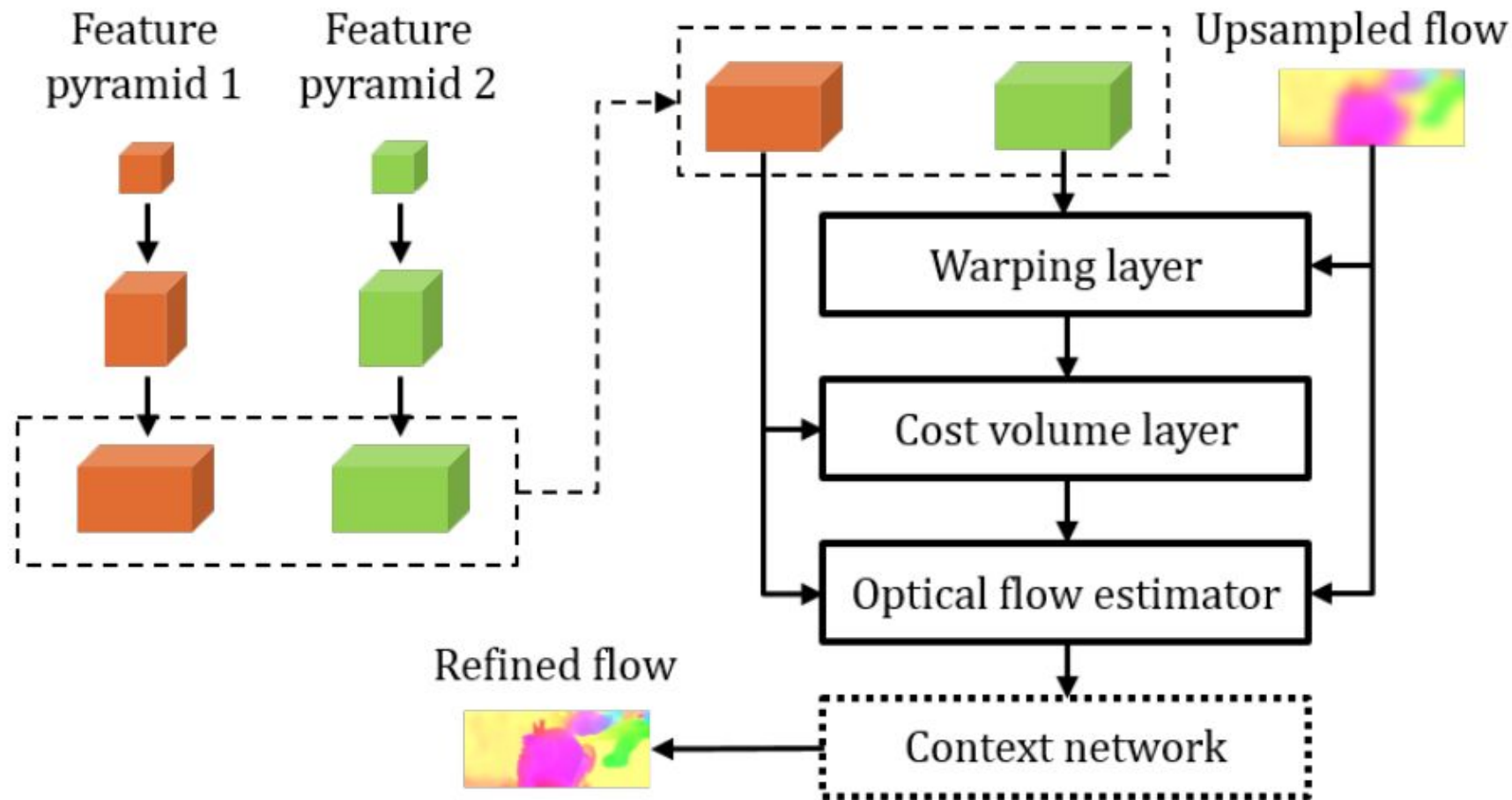
- Введение
- Методы поиска областей открытия/закрытия
- Методы поиска оптического потока
  - MFF
  - **PWC-Net+**
- Заключение



D. Sun et al., "Models matter, so does training: an empirical study of CNNs for optical flow estimation," in *IEEE TPAMI*, 2019

# PWC-Net+

## Архитектура PWC-Net



D. Sun et al., "Models matter, so does training: an empirical study of CNNs for optical flow estimation," in *IEEE TPAMI*, 2019

Ошибка при обучении:

$$\mathcal{L}(\Theta) = \sum_{l=l_0}^L \alpha_l \sum_{\mathbf{x}} |\mathbf{w}_{\Theta}^l(\mathbf{x}) - \mathbf{w}_{GT}^l(\mathbf{x})|_2 + \gamma |\Theta|_2^2$$

Ошибка при дообучении:

$$\mathcal{L}(\Theta) = \sum_{l=l_0}^L \alpha_l \sum_{\mathbf{x}} \left( |\mathbf{w}_{\Theta}^l(\mathbf{x}) - \mathbf{w}_{GT}^l(\mathbf{x})| + \epsilon \right)^q + \gamma |\Theta|_2^2$$

коэффициент  
уровня

пирамиды

устойчивая ошибка  
Шарбонье

$L_2$ -регуляризация

D. Sun et al., "Models matter, so does training:  
an empirical study of CNNs for optical flow  
estimation," in *IEEE TPAMI*, 2019



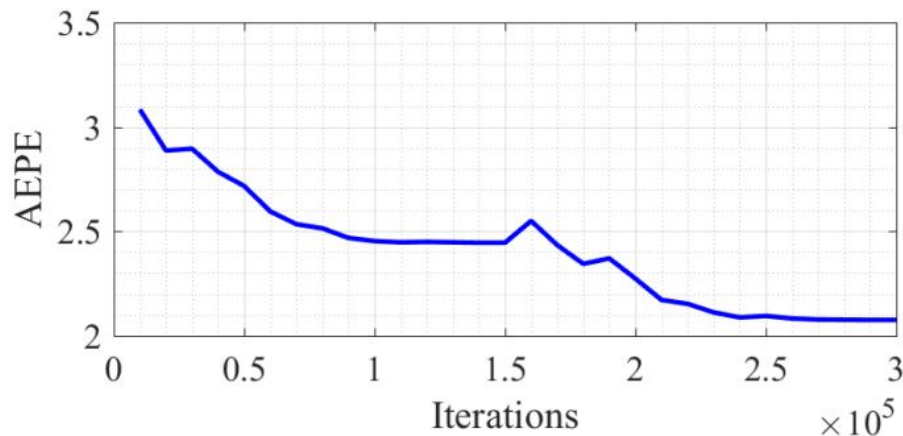
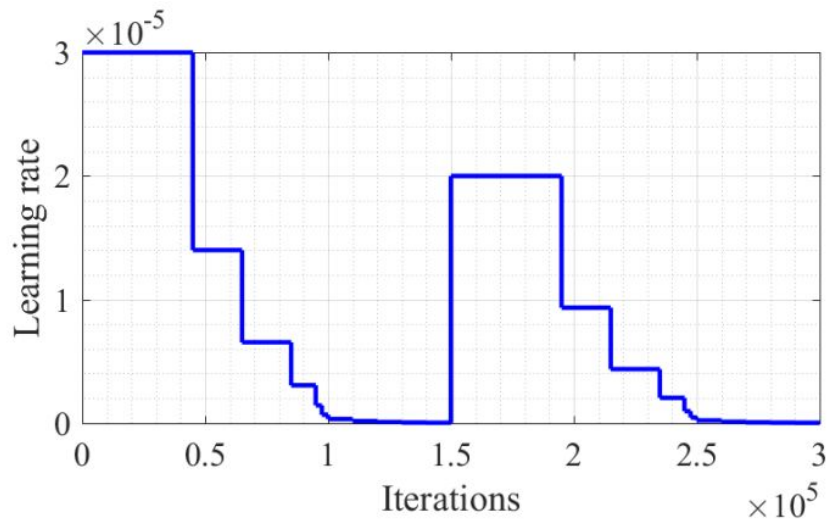
# PWC-Net+

## Обучение



1. FlyingChairs — 1 200 000 итераций
2. FlyingThings3D — 500 000 итераций
3. Дообучение на Sintel и KITTI

Первый протокол тренировки: дообучение на Sintel Final, 300 000 итераций



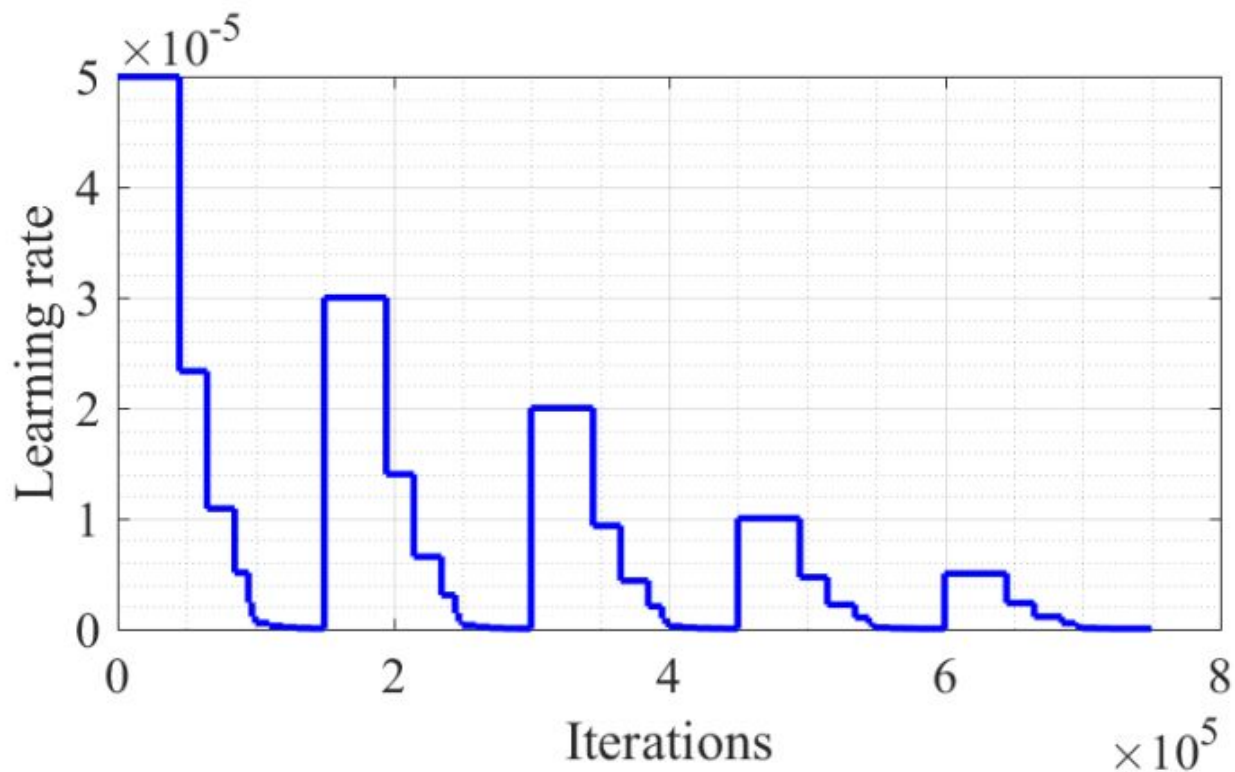
D. Sun et al., “Models matter, so does training: an empirical study of CNNs for optical flow estimation,” in *IEEE TPAMI*, 2019

# PWC-Net+

## Второй протокол обучения



Дообучение на Sintel, KITTI15 и HD1K, 800 000 итераций



D. Sun et al., "Models matter, so does training: an empirical study of CNNs for optical flow estimation," in *IEEE TPAMI*, 2019

Methods	Training		Test		Time (s)
	Clean	Final	Clean	Final	
PatchBatch [61]	-	-	5.79	6.78	50.0
EpicFlow [28]	-	-	4.12	6.29	15.0
MRFlow [35]	1.83	3.59	<b>2.53</b>	5.38	480
FlowFieldsCNN [30]	-	-	3.78	5.36	23.0
SpyNet-ft [11]	(3.17)	(4.32)	6.64	8.36	0.16
FlowNet2 [10]	2.02	3.14	3.96	6.02	0.12
FlowNet2-ft [10]	(1.45)	(2.01)	4.16	5.74	0.12
FlowNetS+	(2.80)	(2.76)	6.49	6.54	<b>0.01</b>
FlowNetC+	2.31	2.34	5.04	5.47	0.05
PWC-Net-small	2.83	4.08	-	-	0.02
PWC-Net-small-ft	(2.27)	(2.45)	5.05	5.32	0.02
PWC-Net	2.55	3.93	-	-	0.03
PWC-Net-ft	(1.70)	(2.21)	3.86	5.13	0.03
PWC-Net-ft-final	(2.02)	(2.08)	4.39	5.04	0.03
PWC-Net_ROB	(1.81)	(2.29)	3.90	4.90	0.03
PWC-Net+	(1.71)	(2.34)	3.45	<b>4.60</b>	0.03

D. Sun et al., “Models matter, so does training: an empirical study of CNNs for optical flow estimation,” in *IEEE TPAMI*, 2019

Methods	KITTI 2012			KITTI 2015		
	AEPE	AEPE	Fl-Noc	AEPE	Fl-all	Fl-all
	<i>train</i>	<i>test</i>	<i>test</i>	<i>train</i>	<i>train</i>	<i>test</i>
EpicFlow [28]	-	3.8	7.88%	-	-	26.29 %
FlowFields [62]	-	-	-	-	-	19.80%
MRFlow [35]	-	-	-	-	14.09 %	12.19 %
DCFlow [17]	-	-	-	-	15.09 %	14.83 %
SDF [29]	-	2.3	3.80%	-	-	11.01 %
MirrorFlow [63]	-	2.6	4.38%	-	9.93%	10.29%
SpyNet-ft [11]	(4.13)	4.7	12.31%	-	-	35.07%
FlowNet2 [10]	4.09	-	-	10.06	30.37%	-
FlowNet2-ft [10]	(1.28)	1.8	4.82%	(2.30)	(8.61%)	10.41 %
LiteFlowNet-CVPR	(1.26)	1.7	-	(2.16)	(8.16%)	10.24 %
LiteFlowNet-arXiv	<b>(1.05)</b>	1.6	<b>3.27%</b>	(1.62)	<b>(5.58%)</b>	9.38 %
PWC-Net	4.14	-	-	10.35	33.67%	-
PWC-Net-ft-CVPR	(1.45)	1.7	4.22%	(2.16)	(9.80%)	9.60%
PWC-Net-ft	(1.08)	<b>1.5</b>	<b>3.41%</b>	<b>(1.45)</b>	<b>(7.59%)</b>	<b>7.90%</b>

D. Sun et al., “Models matter, so does training: an empirical study of CNNs for optical flow estimation,” in *IEEE TPAMI*, 2019

Frame 46 of "Market\_5" (training, clean)

Ground truth

W/o context

W/o DenseNet

PWC-Net

PWC-Net-Sintel-fit

Frame 5 of "Ambush\_3" (test, final)

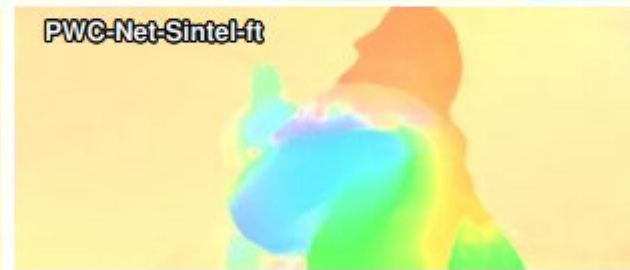
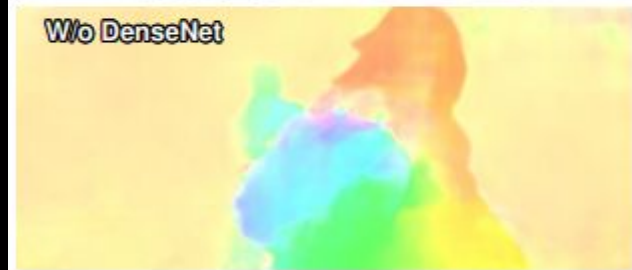
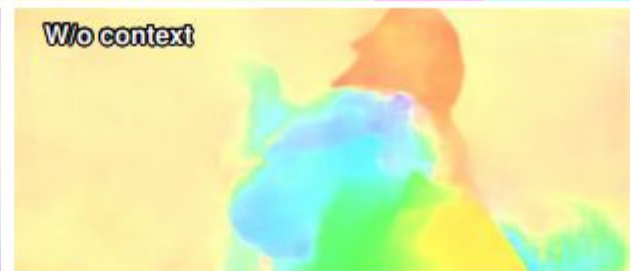
Frame 6

W/o context

W/o DenseNet

PWC-Net

PWC-Net-Sintel-fit



Frame 11 of Market\_1



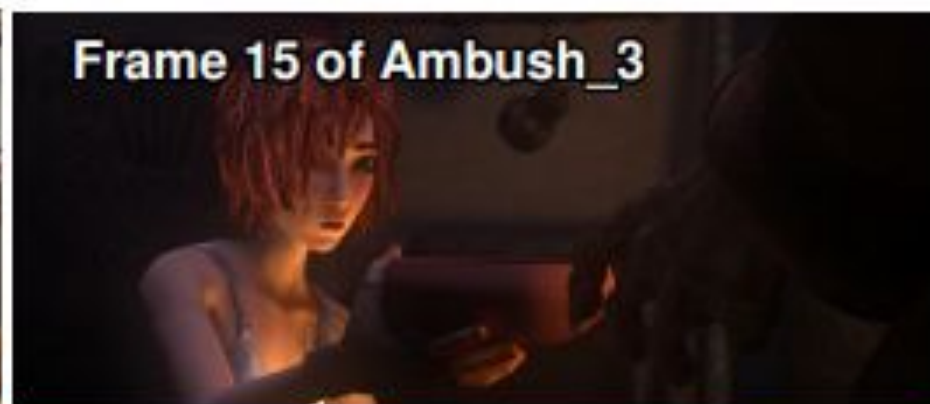
FlowNetC



FlowNetC+



Frame 15 of Ambush\_3



FlowNetC



FlowNetC+



Methods	FlowNetS	FlowNetC	FlowNet2	SpyNet	PWC-Net	PWC-Net-small
#parameters (M)	38.67	39.17	162.49	1.2	8.75	4.08
Parameter Ratio	23.80%	24.11%	100%	0.74%	5.38%	2.51%
Memory (MB)	154.5	156.4	638.5	9.7	41.1	22.9
Memory Ratio	24.20%	24.49%	100%	1.52%	6.44%	3.59%
Training (days)	4	6	>14	-	4.8	4.1
Forward (ms)	11.40	21.69	84.80	-	28.56	20.76
Backward (ms)	16.71	48.67	78.96	-	44.37	28.44

D. Sun et al., “Models matter, so does training: an empirical study of CNNs for optical flow estimation,” in *IEEE TPAMI*, 2019

### Достоинства:

- Авторы провели подробное исследование производительности PWC-Net для различных способов обучения, улучшили свой предыдущий результат
- Лучшие модели выложены вместе со скриптами обучения

### Недостаток:

- Сложность расписания обучения, применимость только для Sintel



# Содержание

- Введение
- Методы поиска областей открытия/закрытия
- Методы поиска оптического потока
- **Заключение**

# Заключение

## Методы поиска областей о/з



Название	Тип	Sintel Clean	Sintel Final
IRR-PWC	с учителем	0.71	0.67
FlowNet-CSSR	с учителем	0.70	0.65
Метод Александры	с учителем	0.56	0.53
OscAwareFlow	без учителя	0.54	0.48
ContinualFlow_ROB	с учителем	-	0.48
MODOF	не нейросеть	-	0.48
Back2FutureFlow	без учителя	0.49	0.44
S2DFlow	не нейросеть	0.38	0.35
MirrorFlow	не нейросеть	0.39	-

# Заключение

В докладе были рассмотрены:

- Методы одновременного вычисления областей открытия/закрытия и оптического потока
- Способы улучшения уже зарекомендовавших себя методов поиска оптического потока

# Планы

- Использование методов улучшения поиска оптического потока для поиска областей открытия/закрытия
- Разработать метод поиска областей открытия/закрытия в стереовидео, который бы превзошел текущий используемый в VQMT3D для построения карт доверия метод

# Литература

1. H. Junhwa and S. Roth, “Iterative Residual Refinement for Joint Optical Flow and Occlusion Estimation,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 5754-5763, 2019.
2. E. Ilg, T. Saikia, M. Keuper, and T. Brox, “Occlusions, motion and depth boundaries with a generic network for disparity, optical flow or scene flow estimation,” in *Proceedings of the European Conference on Computer Vision*, pp. 614-630, 2018.
3. M. Neoral, J. Šochman, and J. Matas, “Continual Occlusion and Optical Flow Estimation,” in *Asian Conference on Computer Vision*, pp. 159-174, Springer, Cham, 2018.
4. Z. Ren, O. Gallo, D. Sun, M.-H. Yang, E. Sudderth, and J. Kautz, “A fusion approach for multi-frame optical flow estimation,” in *IEEE Winter Conference on Applications of Computer Vision*, pp. 2077-2086, 2019.
5. D. Sun, X. Yang, M. Y. Liu, and J. Kautz, “Models Matter, So Does Training: An Empirical Study of CNNs for Optical Flow Estimation,” in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019.