



# SDPL: Shifting-Dense Partition Learning for UAV-View Geo-Localization

Quan Chen, Tingyu Wang, Zihao Yang, Haoran Li,  
Rongfeng Lu, Yaoqi Sun, Bolun Zheng, Chenggang Yan

March 2024

Домницкий Егор Алексеевич,  
студент М4130 ИПКН,  
инженер ФБИТ

## Задача Cross-View Matching

- Имеется набор данных: спутниковая или аэрофотосъемка географических или рукотворных объектов
- Беспилотное средство (воздушное или наземное) оснащено оптической моно- или стереосистемой камер и получает видеопоток
- В режиме реального времени или отсроченно необходимо распознать объекты в кадре (т.е. получить геотег/идентификатор объекта в поле зрения или ограниченной области)



SUES-200 Cross-view Image Matching Benchmark

# Принципы Cross-View representation learning

- Большая разница в визуальных признаках между платформами: зависимость от угла наблюдения, освещения, повороты, смещение, масштаб
- Handmade дескрипторы и статистические подходы не обеспечивают устойчивую дистинктивность/схожесть
- **Подход DL: выучиваем кросс-доменный инвариант, общее латентное пространство (Classification, Metric Learning)**



(a) Ground Image



(b) Matching Satellite Image (c) Unmatching Satellite Image

## Референсные модели (Drone<>Satellite)

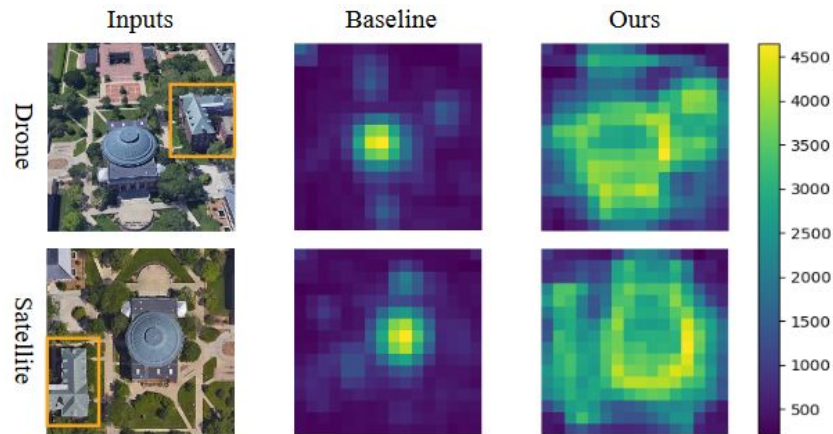
- CVM Net (Hu et al, IEEE Conference on CV and PR, 2018)  
CNN + NetVLAD + Soft margin ranking loss
- SAFA (Shi et al, Neural Information Processing Systems, 2019)  
CNN + извлечение Spatial Aware Embedding (макспул по каналам и генерация маски важности) + Triplet loss
- LCM (Ding et al, Remote Sensing 2020)  
ResNet + FC + Cls , cross entropy loss, cosine distance  
(все более ранние работы + бейзлайны бенчмарков похожи)
- CVFT (Shi et al, AAAI 2020)  
CNN + Cross domain feature alignment (regression, cross-view feature transport) + Triplet loss

## Референсные модели (Drone $\leftrightarrow$ Satellite)

- FSRA (Dai et al, IEEE Transactions 2021)  
ViT-s + Feature Segmentation and Region Alignment (FSRA)  
Сегментация областей при помощи хитмапа, сравнение фич из агрегированных областей, Triplet Loss
- LPN (Wang et al, IEEE Transactions 2022)  
CNN+Square Partition Strategy (далее), cross entropy loss
- MBF (Zhu et al, Sensors 2023)  
Фьюжн ResNet признаков с BERT эмбедами описания в ViT, cross entropy + triplet loss
- MCCG (Shen et al, IEEE Transactions 2023)  
ConvNeXt + кросс пространственные свертки + triplet loss между эмбедами и cross entropy loss для классификации

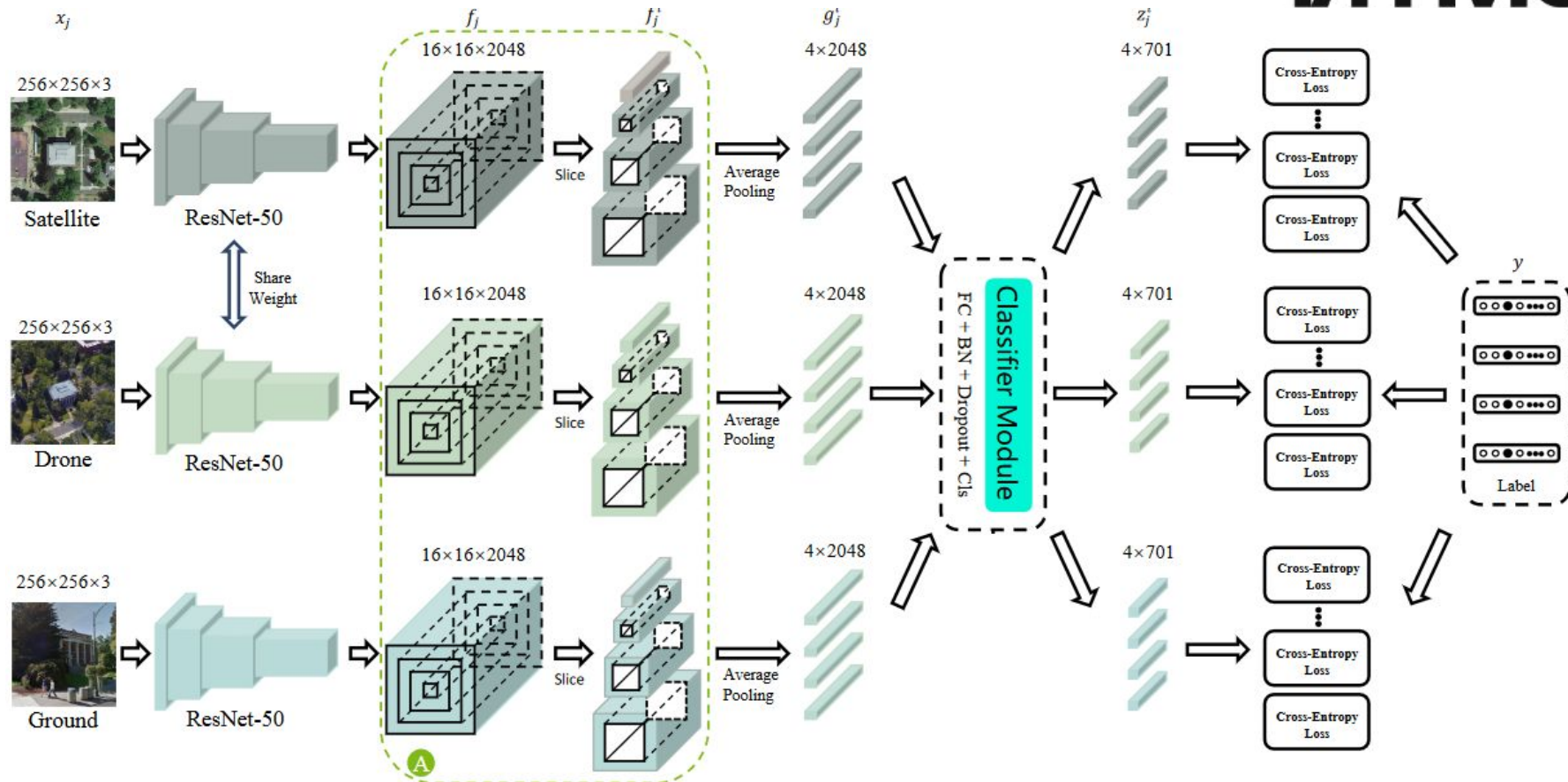
## Local Partition Network (Wang et al)

- Идея вдохновлена принципом фокусировки зрения: распознаваемый объект находится в “фокусе”, но окружение также задает необходимый контекст
- Хотим явно фокусироваться на окружении
- Как сайд-эффект: устойчивость к повороту





# Local Partition Network (Wang et al)

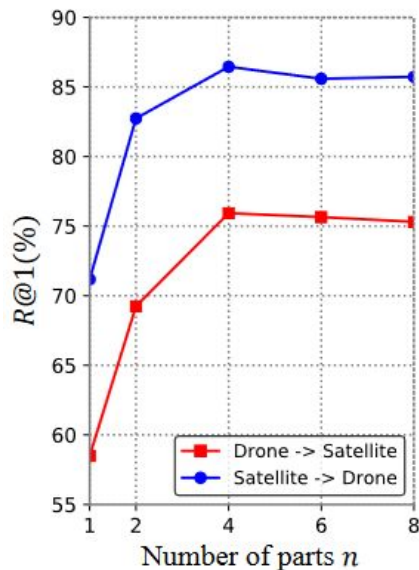


**Local Partition Network (Wang et al)**

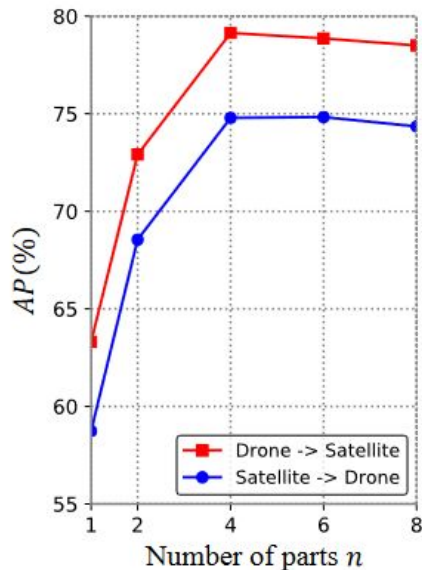
Method	University-1652			
	Drone $\rightarrow$ Satellite		Satellite $\rightarrow$ Drone	
	R@1	AP	R@1	AP
Instance Loss [2]	58.49	63.31	71.18	58.74
Contrastive Loss [23]	52.39	57.44	63.91	52.24
Triplet Loss ( $M = 0.3$ ) [57]	55.18	59.97	63.62	53.85
Triplet Loss ( $M = 0.5$ ) [57]	53.58	58.60	64.48	53.15
Soft Margin Triplet Loss [9]	53.21	58.03	65.62	54.47
Ours (w/o Google)	74.18	77.39	85.16	73.68
Ours	<b>75.93</b>	<b>79.14</b>	<b>86.45</b>	<b>74.79</b>
Ours (w/ CVUSA distractors)	70.61	73.53	-	-

Method	Publication	Backbone	CVUSA			
			R@1	R@5	R@10	R@Top1%
CVFT [7]	AAAI'20	VGG16	61.43	84.69	90.94	99.02
SAFA [1]	NIPS'19	VGG16	89.84	96.93	98.14	99.64
Ours	-	VGG16	79.69	91.70	94.55	98.50
Ours	-	ResNet-50	<b>85.79</b>	<b>95.38</b>	<b>96.98</b>	<b>99.41</b>
CVFT [7] + Ours	-	VGG16	68.20	88.00	92.69	99.30
SAFA [1] + Ours	-	VGG16	<b>92.83</b>	<b>98.00</b>	<b>98.85</b>	<b>99.78</b>





(a)



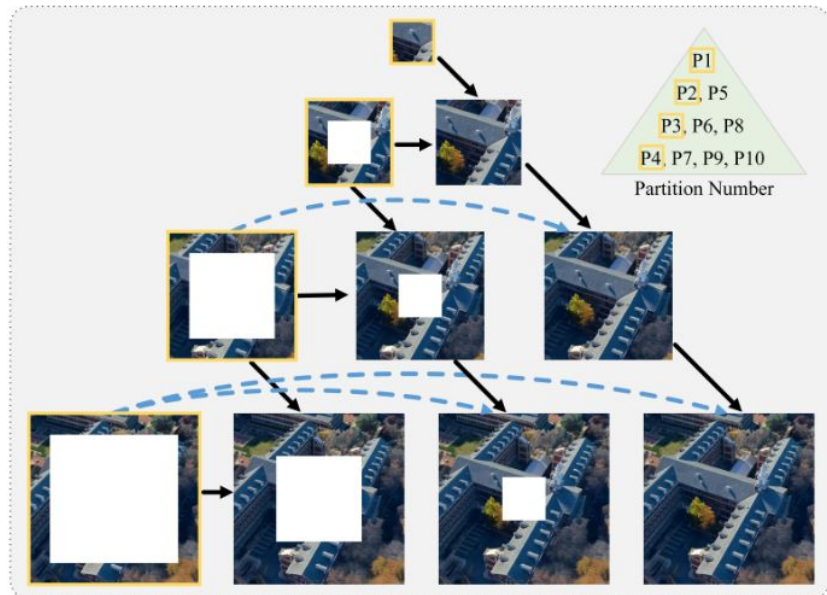
(b)

TABLE V  
ABLATION STUDY ON ROTATING IMAGES DURING INFERENCE ON  
UNIVERSITY-1652.

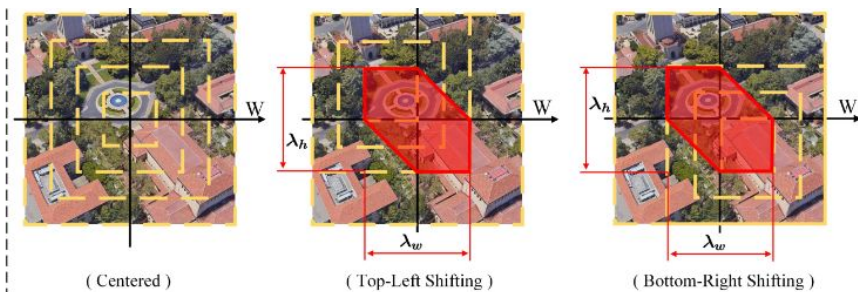
Rotation Query	Angle Gallery	Drone $\rightarrow$ Satellite		Satellite $\rightarrow$ Drone	
		R@1	AP	R@1	AP
0°	0°	75.93	79.14	86.45	74.79
16°	0°	75.64	78.86	85.16	72.78
45°	0°	72.04	75.62	85.16	72.27
67°	0°	70.39	74.09	85.73	73.06
90°	0°	68.80	72.67	86.31	75.31
180°	0°	70.76	74.47	85.45	74.03
204°	0°	69.92	73.68	84.45	72.22
270°	0°	69.06	72.49	86.73	75.12
317°	0°	72.29	75.87	84.17	71.85
32°	75°	73.19	76.69	83.17	66.41
216°	87°	69.54	73.27	83.45	65.29

Is LPN robust to the shifted query image?

# Shifting-Dense Partition (Chen et al)

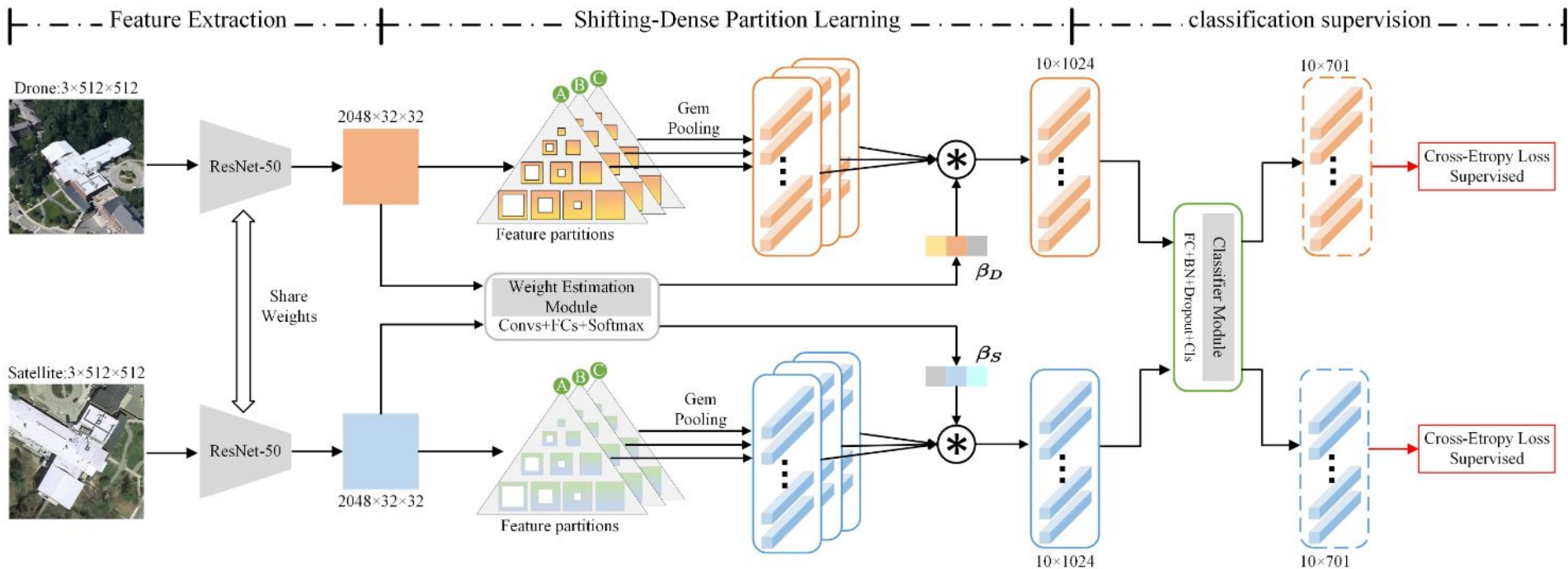


Dense partition strategy



- Увеличение числа разбиений, начиная с 4, не приводит к улучшению качества в LPN
- Можно постепенно увеличивать область “фокуса”, накладывая базовые разбиения
- Для того, чтобы сделать модель устойчивой к смещениям центра - можно на этапе обучения извлекать разбиения со сдвигом

# Shifting-Dense Partition (Chen et al)



## Анализ SDPL

Method	Publication	University-1652 [9]			
		Drone→Satellite		Satellite→Drone	
		Recall@1	AP	Recall@1	AP
Instance Loss [9]	ACMMM'2020	59.69	64.80	73.18	59.40
Soft Margin Triplet Loss [57]	CVPR'2018	54.77	59.93	66.87	55.19
LCM [15]	Remote Sens'2020	66.65	70.82	79.89	65.38
LPN [11]	TCSVT'2021	77.71	80.80	90.30	78.78
SDPL <sup>‡</sup> (w. ResNet-50)	—	82.57	85.08	88.73	80.79
SDPL (w. ResNet-50)	—	85.25	87.48	88.45	79.79
MCCG [34] (w. ConvNeXt)	TCSVT'2023	89.28	91.01	94.29	89.29
Swin-B [52]	CVPR'2021	84.15	86.62	90.30	83.55
FSRA [18]	TCSVT'2021	84.51	86.71	88.45	83.37
SwinV2-B [56]	CVPR'2022	86.99	89.02	91.16	85.77
Dai <i>et al.</i> [58]	TIP'2023	82.22	84.78	87.59	81.49
MBF <sup>‡</sup> [10]	Sensors'2023	89.05	90.61	93.15	88.17
SDPL (w. SwinV2-B)	-	90.16	91.64	93.58	89.45



## Анализ SDPL (SUES - 200)

Drone→Satellite								
Method	150m		200m		250m		300m	
	Recall@1	AP	Recall@1	AP	Recall@1	AP	Recall@1	AP
SUES-200 [21]	55.65	61.92	66.78	71.55	72.00	76.43	74.05	78.26
LCM [15]	43.42	49.65	49.42	55.91	54.47	60.31	60.43	65.78
LPN [11]	61.58	67.23	70.85	75.96	80.38	83.80	81.47	84.53
FSRA [18]	68.25	73.45	83.00	85.99	90.68	92.27	91.95	93.46
MBF [10]	85.62	88.21	87.43	90.02	90.65	92.53	92.12	93.63
MCCG [34]	82.22	85.47	89.38	91.41	93.82	95.04	95.07	96.20
<b>SDPL</b>	82.95	85.82	92.73	94.07	96.05	96.69	97.83	98.05
Satellite→Drone								
Method	150m		200m		250m		300m	
	Recall@1	AP	Recall@1	AP	Recall@1	AP	Recall@1	AP
SUES-200 [21]	75.00	55.46	85.00	66.05	86.25	69.94	88.75	74.46
LCM [15]	57.50	38.11	68.75	49.19	72.50	47.94	75.00	59.36
LPN [11]	83.75	66.78	88.75	75.01	92.50	81.34	92.50	85.72
FSRA [18]	83.75	76.67	90.00	85.34	93.75	90.17	95.00	92.03
MBF [10]	88.75	84.74	91.25	89.95	93.75	90.65	96.25	91.60
MCCG [34]	93.75	89.72	93.75	92.21	96.25	96.14	98.75	96.64
<b>SDPL</b>	93.75	83.75	96.25	92.42	97.50	95.65	96.25	96.17



## Анализ SDPL (смещения на University 1652)

Padding Pixel (P)	LPN (ResNet) [11]		FSRA (ViT) [18]		SDPL (ResNet)	
	Recall@1	AP	Recall@1	AP	Recall@1	AP
(0,0)	78.08 <sub>-0</sub>	80.94 <sub>-0</sub>	86.41 <sub>-0</sub>	88.34 <sub>-0</sub>	85.25 <sub>-0</sub>	87.48 <sub>-0</sub>
(+20,0)	76.72 <sub>-1.36</sub>	79.72 <sub>-1.22</sub>	85.51 <sub>-0.90</sub>	87.59 <sub>-0.75</sub>	84.44 <sub>-0.81</sub>	86.80 <sub>-0.68</sub>
(+40,0)	73.04 <sub>-5.04</sub>	76.47 <sub>-4.47</sub>	82.77 <sub>-3.64</sub>	85.30 <sub>-3.04</sub>	82.46 <sub>-2.79</sub>	85.15 <sub>-2.33</sub>
(+60,0)	69.54 <sub>-8.54</sub>	73.36 <sub>-7.58</sub>	77.95 <sub>-8.46</sub>	81.18 <sub>-7.16</sub>	78.68 <sub>-6.57</sub>	81.91 <sub>-5.57</sub>
(+80,0)	61.13 <sub>-16.95</sub>	65.80 <sub>-15.14</sub>	70.90 <sub>-15.51</sub>	74.99 <sub>-13.35</sub>	74.48 <sub>-10.77</sub>	78.33 <sub>-9.15</sub>
(+100,0)	47.87 <sub>-30.21</sub>	53.50 <sub>-27.44</sub>	62.10 <sub>-24.31</sub>	67.05 <sub>-21.29</sub>	68.19 <sub>-16.34</sub>	73.57 <sub>-13.91</sub>
(+20,+20)	74.75 <sub>-3.33</sub>	77.94 <sub>-3.00</sub>	83.59 <sub>-2.82</sub>	85.95 <sub>-2.39</sub>	83.24 <sub>-2.01</sub>	85.76 <sub>-1.72</sub>
(+40,+40)	73.04 <sub>-5.04</sub>	69.56 <sub>-11.38</sub>	74.15 <sub>-12.26</sub>	77.80 <sub>10.54</sub>	77.90 <sub>-7.35</sub>	81.20 <sub>-6.28</sub>
(+60,+60)	54.37 <sub>-23.71</sub>	59.38 <sub>-21.56</sub>	57.60 <sub>-28.81</sub>	62.88 <sub>-25.46</sub>	65.56 <sub>-19.69</sub>	70.45 <sub>-17.03</sub>
(+80,+80)	36.95 <sub>-41.13</sub>	42.79 <sub>-38.15</sub>	41.30 <sub>-45.11</sub>	47.15 <sub>-41.19</sub>	54.11 <sub>-31.14</sub>	60.12 <sub>-27.36</sub>
(+100,+100)	24.03 <sub>-54.05</sub>	29.76 <sub>-51.18</sub>	28.07 <sub>-58.34</sub>	33.73 <sub>-54.61</sub>	45.19 <sub>-40.06</sub>	51.84 <sub>-35.64</sub>
(-20,-20)	76.41 <sub>-1.67</sub>	79.44 <sub>-1.50</sub>	84.35 <sub>-2.06</sub>	86.62 <sub>-1.72</sub>	84.06 <sub>-1.19</sub>	86.47 <sub>-1.01</sub>
(-40,-40)	70.28 <sub>-7.80</sub>	74.04 <sub>-6.90</sub>	78.10 <sub>-8.31</sub>	81.24 <sub>-7.10</sub>	81.00 <sub>-4.25</sub>	83.92 <sub>-3.56</sub>
(-60,-60)	62.90 <sub>-15.18</sub>	67.32 <sub>-13.62</sub>	67.73 <sub>-18.68</sub>	71.97 <sub>-16.37</sub>	76.81 <sub>-8.44</sub>	80.27 <sub>-7.21</sub>
(-80,-80)	48.94 <sub>-29.14</sub>	54.43 <sub>-26.51</sub>	54.60 <sub>-31.81</sub>	59.80 <sub>-28.54</sub>	69.95 <sub>-15.30</sub>	74.34 <sub>-13.14</sub>
(-100,-100)	19.14 <sub>-58.94</sub>	24.79 <sub>-56.15</sub>	42.45 <sub>-43.96</sub>	48.10 <sub>-40.24</sub>	60.09 <sub>-25.16</sub>	65.62 <sub>-21.86</sub>
(+20,-20)	78.73 <sub>+0.65</sub>	81.48 <sub>+0.54</sub>	84.23 <sub>-2.18</sub>	86.55 <sub>-1.79</sub>	83.77 <sub>-1.48</sub>	86.21 <sub>-1.27</sub>
(+40,-40)	73.18 <sub>-4.90</sub>	76.60 <sub>-4.34</sub>	77.90 <sub>-8.51</sub>	81.09 <sub>-7.25</sub>	81.15 <sub>-4.10</sub>	84.03 <sub>-3.45</sub>
(+60,-60)	62.78 <sub>-15.30</sub>	67.28 <sub>-13.66</sub>	67.29 <sub>-19.12</sub>	71.62 <sub>-16.72</sub>	76.94 <sub>-8.31</sub>	80.41 <sub>-7.07</sub>
(+80,-80)	48.81 <sub>-29.27</sub>	54.31 <sub>-26.63</sub>	53.90 <sub>-32.51</sub>	59.21 <sub>-29.13</sub>	70.12 <sub>-15.13</sub>	74.54 <sub>-12.94</sub>
(+100,-100)	35.86 <sub>-42.22</sub>	41.78 <sub>-39.16</sub>	41.69 <sub>-44.72</sub>	47.41 <sub>-40.93</sub>	61.80 <sub>-23.45</sub>	67.14 <sub>-20.34</sub>
(-20,+20)	76.98 <sub>-1.10</sub>	79.87 <sub>-1.07</sub>	83.46 <sub>-2.95</sub>	85.85 <sub>-2.48</sub>	82.32 <sub>-2.93</sub>	84.94 <sub>-2.54</sub>
(-40,+40)	68.77 <sub>-9.31</sub>	72.51 <sub>-8.43</sub>	74.47 <sub>-11.94</sub>	78.05 <sub>-10.29</sub>	76.40 <sub>-8.85</sub>	79.85 <sub>-7.63</sub>
(-60,+60)	54.45 <sub>-23.63</sub>	59.43 <sub>-21.51</sub>	58.05 <sub>-28.36</sub>	63.27 <sub>-25.07</sub>	66.43 <sub>-18.82</sub>	71.17 <sub>-16.31</sub>
(-80,+80)	36.81 <sub>-41.27</sub>	42.69 <sub>-38.25</sub>	41.69 <sub>-44.72</sub>	47.52 <sub>-40.82</sub>	54.80 <sub>-30.45</sub>	60.78 <sub>-26.70</sub>
(-100,+100)	24.16 <sub>-53.92</sub>	29.90 <sub>-51.04</sub>	27.93 <sub>-58.48</sub>	33.67 <sub>-54.67</sub>	44.19 <sub>-41.06</sub>	50.89 <sub>-36.59</sub>

## Реализация



- SUES-200 (300m)
- Drone → Satellite
- Pytorch
- Kaggle (GPU P100)
- 1 epoch (15 min на 1 эпоху, валидация на трейне)

**Спасибо  
за внимание!**

**it**MO *re than a*  
**UNIVERSITY**

Домницкий Е.А.  
egor.dom0923@gmail.com  
+7 952 202 93 78