



ГОСУДАРСТВЕННЫЙ НАУЧНЫЙ ЦЕНТР РОССИИ
ЦЕНТРАЛЬНЫЙ НАУЧНО-ИССЛЕДОВАТЕЛЬСКИЙ И ОПЫТНО-КОНСТРУКТОРСКИЙ
ИНСТИТУТ РОБОТОТЕХНИКИ И ТЕХНИЧЕСКОЙ КИБЕРНЕТИКИ

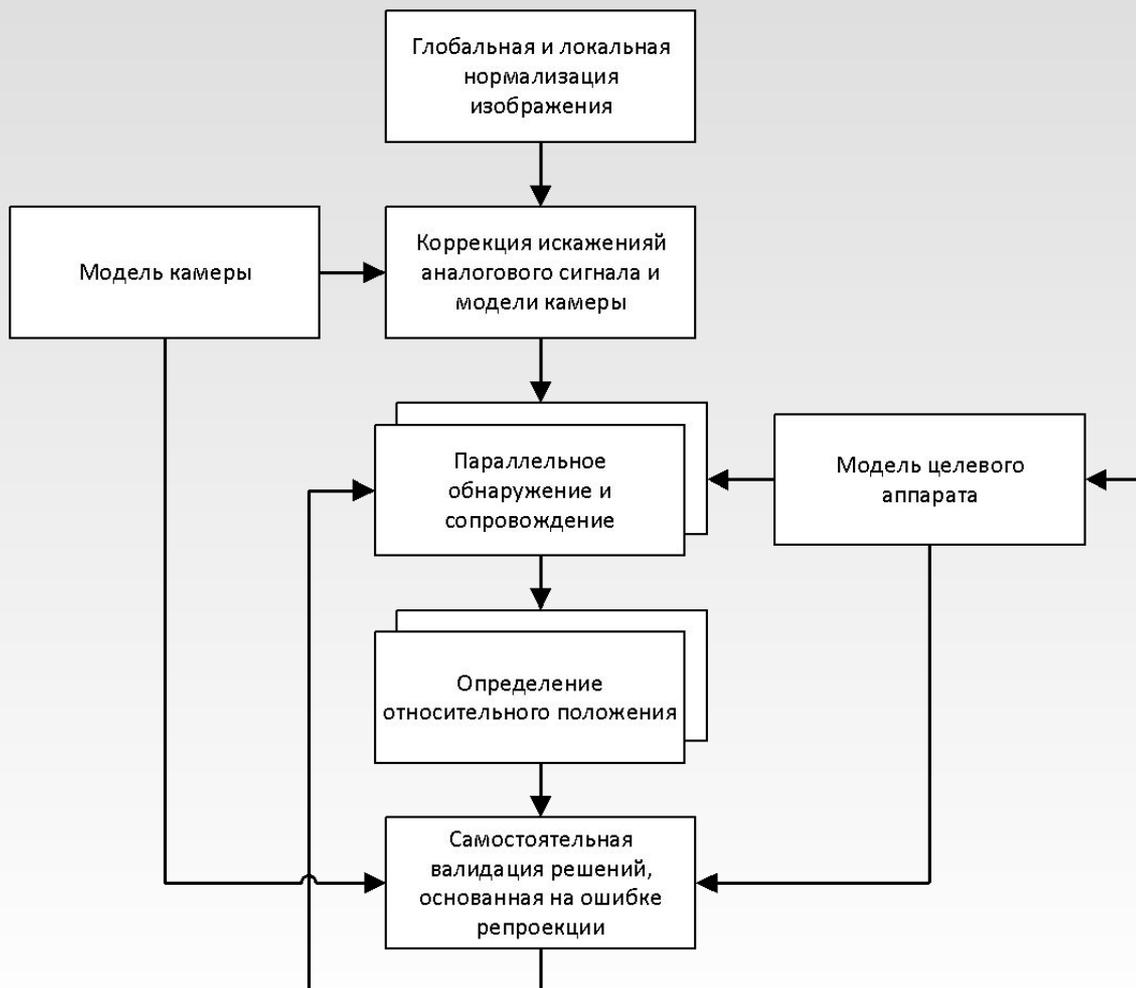
ИССЛЕДОВАНИЕ СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ НА ЗАДАЧЕ ОБНАРУЖЕНИЯ ОБЪЕКТОВ НА ИЗОБРАЖЕНИЯХ КОСМИЧЕСКОЙ СТЫКОВКИ

***Иван Фомин, Светлана Орлова
Дмитрий Громошинский, Александр Бахшиев***

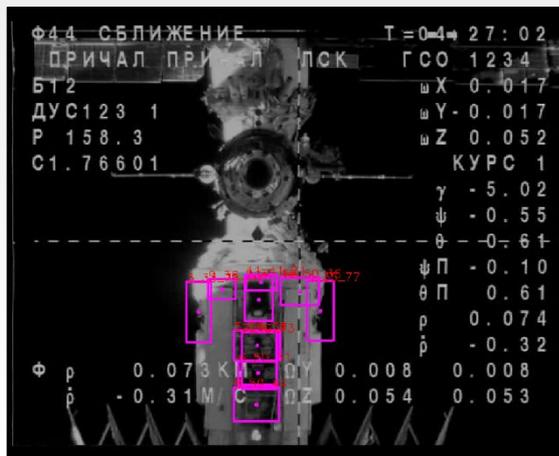
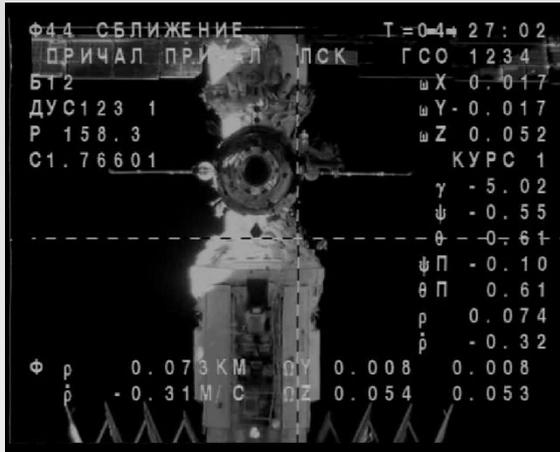
Россия, 194064, г. Санкт-Петербург, Тихорецкий пр., 21
тел.: (812) 552-0110 (812) 552-1325 факс: (812) 556-3692 <http://www.rtc.ru> e-mail: rtc@rtc.ru



Общая схема системы



Постановка задачи обнаружения объектов



- Есть последовательность изображений с камеры из космоса

$$I = \text{image}(\text{width}, \text{height})$$

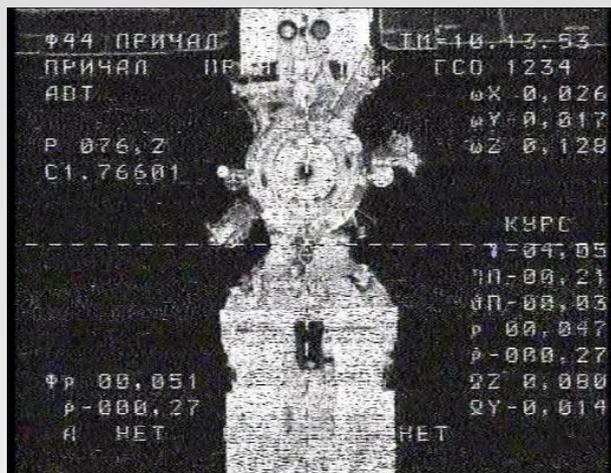
- Требуется определить положение всех видимых объектов
- Важна позиция центра каждого объекта, иногда высота и ширина

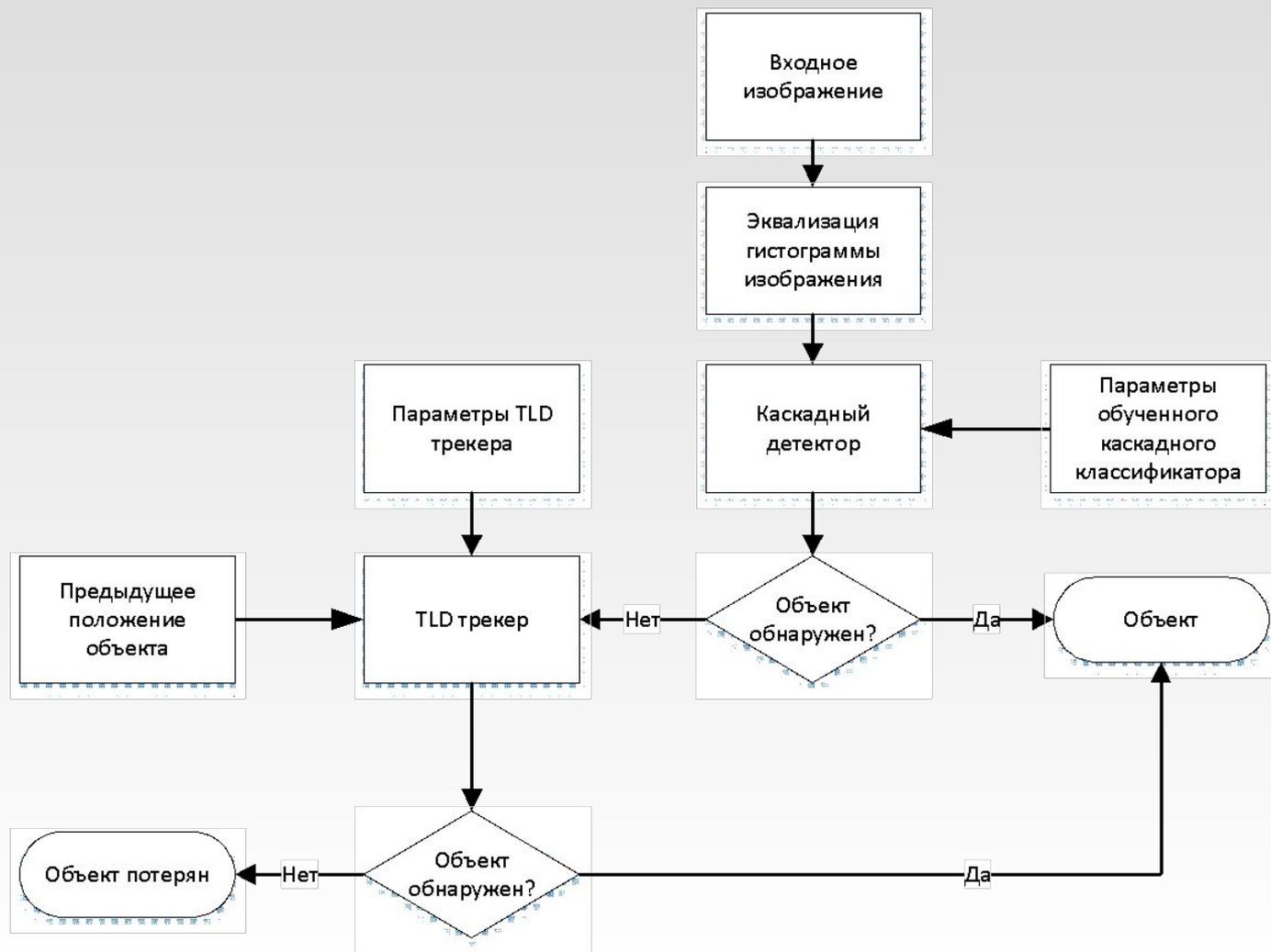
$$I \xrightarrow{\text{detector}} \{ \text{feat}(cx, cy, \text{width}, \text{height}) \}_{i=1}^n$$

- Любое изображение может быть искажено
- Детектор должен быть инвариантен к искажениям



Примеры различных искажений







Постановка задачи

- ❑ Подготовить обучающий и тестовый наборы данных
- ❑ Выбрать нейросетевые системы обнаружения объектов
- ❑ Преобразовать обучающий и тестовый наборы данных в соответствии с выбранными системами
- ❑ Обучить и протестировать нейронные сети и каскадный детектор
- ❑ Сравнить результаты работы детектора и нейронной сети
- ❑ Принять решение о допустимости использования выбранной системы в дальнейших работах



- ❑ **Faster R-CNN:** Предположения о положении объектов генерируются специальной сетью, вся система работает на GPU, свертка по всему изображению, регионы проецируются на результаты свертки, работает медленно (0.3 с на изображении размером 516x413), архитектура слоев от сети VGG-CNN-M-1024.
- ❑ **YOLO v2:** Предположения выбираются самой сетью: изображение разбивается по сетке и в каждой ячейке берется набор предположений, классификатор работает очень быстро, проверяя порядка 800 примеров за 0.02 с на изображении размером 516x413. Архитектура слоев от MobileNet.



Постановка задачи обучения нейронной сети

□ Имеется система обнаружения S , основанная на нейронной сети N . Сеть содержит слои L_i которые содержат веса $W = W_{i,j}$.

Нейронная сеть инициализируется перед обучением случайно или заранее заданными весами:

$$W = W^0 \text{ или } W = W^R$$

В процессе обучения в систему передаются изображения и положения объектов, и веса изменяются:

$$\begin{array}{c} I_1 \dots I_N \\ gt_{1,1} \dots gt_{1,M_1} \dots gt_{N,M_N} \end{array} \xrightarrow{\text{learn}} W = W^K$$

Обученная система определяет координаты объектов:

$$S(I, W^K) \xrightarrow{\text{detection}} \{det(cx, cy, w, h)\}_{i=1}^{N_D}$$



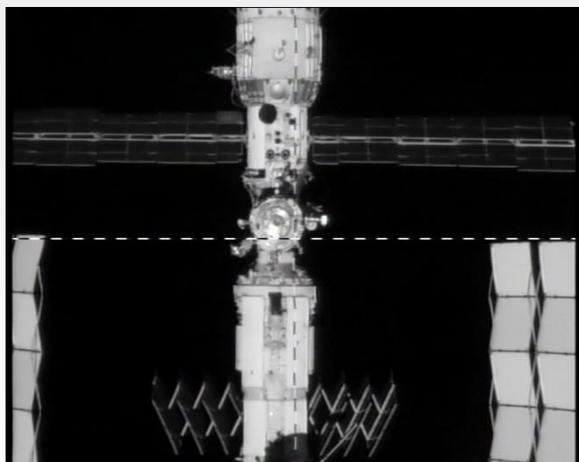
Описание набора данных: примеры вида различных узлов



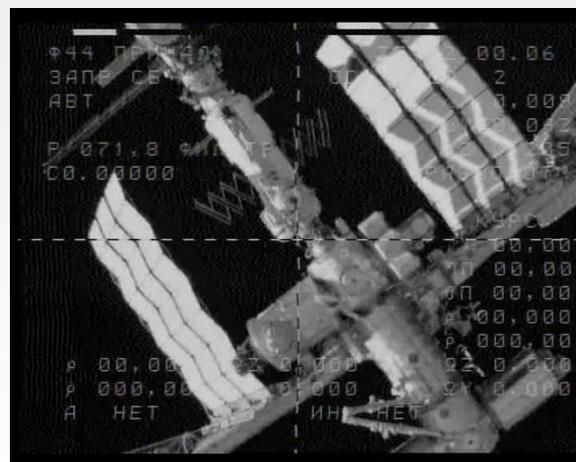
Узел 1



Узел 2



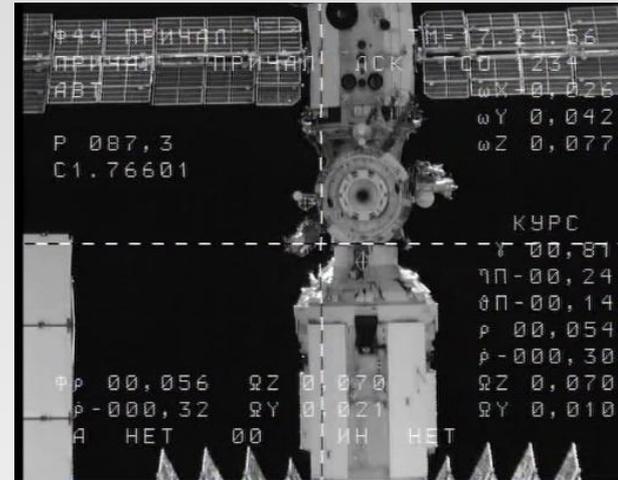
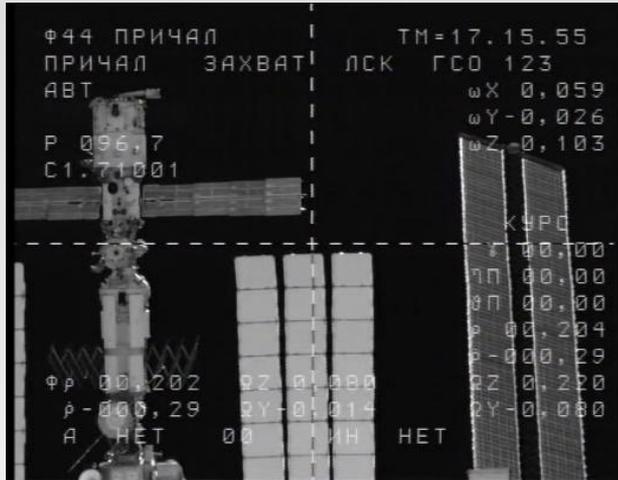
Узел 3



Узел 4



Описание набора данных: вид второго узла с разного расстояния





Описание набора данных: количество изображений примеров

Параметр	Узел 1	Узел 2	Узел 3	Узел 4
Число объектов	20	33	38	51
Обучающий набор	2307	3443	5070	5146
Тестовый набор	122	182	267	271
Число примеров	10084	35002	54445	56276

Имя объекта	Число примеров	Имя объекта	Число примеров
DUnit_Center_Small_1	591	Marker_B_6	1119
Comb_Double_Part_7	479	Marker_Circle_1	961
Black_Rect_2	1377	Ledge_3	1317
TopUnit_2	460	SP_Right_4	585
Marker_1	1288	DUnit_Center_6	2199
Right_Part_10:	1591	Right_Edge_Part_4	1171
Top_Double_Part_5	792	Left_Part_9	1691
Marker_C_4	1115	Marker_TC_2	1247
Left_Rect_4	736	DUnit_Top_5	1176
Marker_LC_3	1066	Right_Rect_5	711
Bottom_Double_Part_6	529	SP_Left_3	590
Left_Edge_Part_3	1105	DUnit_LeftCircleAnt_8	616
Cross_8	201	Black_Rect_Top_11	1059
Marker_Rhombus_2	1341	DUnit_LeftAnt_7	1673
DUnit_RightCircleAnt_10	843	DUnit_BottomAnt_11	1095
DUnit_1	2068	Marker_RC_5	1037
DUnit_RightAnt_9	1264		



Задача сравнения результатов (1/2)

Результаты работы нейронной сети это **прямоугольники** с **гладкой** функцией принадлежности объекту:

$$S(I, W^K) \xrightarrow{\text{detection}} \{rect(l, t, r, b, score)\}_{i=1}^{N_D}, score = [0, 1]$$

Обнаружение верно если метрика *Intersection over Union* с любым размеченным прямоугольником больше 0.5:

$$\frac{rect \cap gt_k}{rect \cup gt_k} > 0.5$$

Используя эти данные можно посчитать различные метрики:

True Positive Rate (TP), False Negative Rate (FN), False Positive Rate (FP), кривые точности-полноты, Mean Average Precision (mAP), и т.д.



Задача сравнения результатов(2/2)

- Результаты каскадного детектора это **координаты** положения объектов в кадре с **почти бинарной** функцией принадлежности:

$$C(I) \xrightarrow{\text{detection}} \{pos(x, y, score)\}_{i=1}^{N_D}, score = 0 \mid (0.8; 1]$$

Условие правильности обнаружения авторами **не определено**. Выбрано правило, что обнаружение верное, если двумерная Евклидова дистанция до любого известного объекта меньше 5 пикселей

Метрики, которые можно рассчитать:

True Positive Rate (TP), False Negative Rate (FN), False Positive Rate (FP) и другие, использующие их – точность и полнота, например.

Результаты зависят от выбранного условия.



Сравнение результатов

Дистанция	Нейронная сеть		Каскадный детектор		Нейронная сеть		Каскадный детектор	
	Точность	Полнота	Точность	Полнота	Ложн.	Проп.	Ложн.	Проп.
110-75 м.	0.832	0.464	0.518	0.422	0.168	0.536	0.482	0.578
75-62 м.	0.84	0.54	0.646	0.543	0.160	0.458	0.354	0.457
62-52 м.	0.69	0.458	0.514	0.555	0.310	0.542	0.486	0.445
52-45 м.	0.753	0.479	0.427	0.463	0.247	0.521	0.573	0.537
45-30 м.	0.664	0.591	0.422	0.427	0.336	0.409	0.578	0.573
30-23 м.	0.608	0.562	0.457	0.58	0.392	0.438	0.543	0.420
23-19 м.	0.644	0.762	0.243	0.57	0.356	0.238	0.757	0.429

$$\text{ТОЧНОСТЬ} = \frac{TP}{TP + FP}$$

$$\text{ПОЛНОТА} = \frac{TP}{TP + FN}$$

$$\text{ЧАСТ. ЛОЖН.} = \frac{FP}{FP + TP}$$

$$\text{ЧАСТ. ПРОП.} = \frac{FP}{FN + TP}$$

	Узел 1	Узел 2	Узел 3	Узел 4
VGG-1024, все объекты	0,790	0,363	0,309	0,394
VGG-1024, аугментация	0,765	0,345	0,297	0,508
VGG-1024, высокий контраст	0,779	0,377	0,250	0,454
ZF, все объекты	0,708	0,316	0,219	0,345
ZF, высокий контраст	0,727	0,365	0,217	0,360
YOLO v2, все объекты	0,851	0,716	0,455	0,406

mAP, Mean Average Precision – обобщенная метрика из соревнования Pascal VOC, упрощающая Precision-Recall кривую до 11 точек, в каждой из которой берется максимум из всех значений слева от точки



Влияние предобработки изображений

	Ф1	Ф2	Ф3	Ф4	Ф5
VGG-1024, все объекты	0,846	0,849	0,843	0,851	0,857
YOLO v2, все объекты	0,851	0,849	0,850	0,849	0,846

- Ф1 – нет фильтра
- Ф2 – повышение контрастности
- Ф3 – выравнивание гистограммы
- Ф4 – выравнивание гистограммы и повышение контрастности
- Ф5 – повышение контрастности и повышение резкости



Заключение

- Собраны обучающий и тестовый наборы данных
- Выполнено тестирование нескольких сетей и каскадного детектора на одном и том же наборе данных
- Выполнено сравнение результатов работы
- Система YOLO v2 превышает качество работы Faster R-CNN, и с учетом работы в реальном времени в дальнейшем будет использоваться в задачах обнаружения объектов

Дальнейшие исследования:

- Исследование влияния архитектуры и различных гиперпараметров сетей на качество обнаружения
- Внедрение в существующую систему поддержки стыковки



ГОСУДАРСТВЕННЫЙ НАУЧНЫЙ ЦЕНТР РОССИИ
ЦЕНТРАЛЬНЫЙ НАУЧНО-ИССЛЕДОВАТЕЛЬСКИЙ И ОПЫТНО-КОНСТРУКТОРСКИЙ
ИНСТИТУТ РОБОТОТЕХНИКИ И ТЕХНИЧЕСКОЙ КИБЕРНЕТИКИ

Спасибо за внимание!



Россия, 194064, г. Санкт-Петербург, Тихорецкий пр., 21
тел.: (812) 552-0110 (812) 552-1325 факс: (812) 556-3692 <http://www.rtc.ru> e-mail: rtc@rtc.ru