Разработка и исследование нейросетевых методов мультимодального трехмерного обнаружения объектов

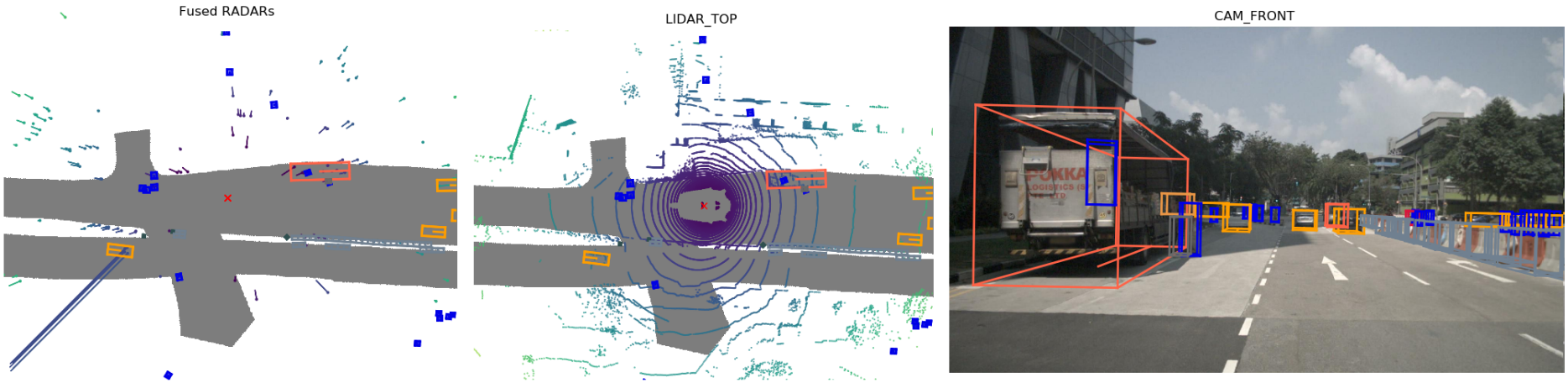
Матыкина О.В.
matykina.ov@phystech.edu

Магистерская программа
«Методы и технологии искусственного интеллекта»

Презентация статуса работы

Введение

Трехмерное обнаружение объектов, целью которого является предсказание местоположения, размеров и категорий объектов вблизи автономного транспортного средства, является важной частью системы восприятия



Мотивация

Table 1
The influence level of various weather conditions on sensors.

Modality	Light rain <4 mm/h	Heavy rain >25 mm/h	Dense smoke /Mist vis<0.1 km	Fog vis<0.5 km	Haze/Smog vis>2 km	Snow	Strong light
LiDAR (λ 850–950 nm and 1550 nm)	2	3	5	4	1	5	2
Radar (24, 77 and 122 GHz)	0	1	2	0	0	2	0
Ground-Penetrating Radar (100–400 MHz)	0	0	0	0	0	1	0
Camera	3	4	5	4	3	2 (dynamic) 3 (static)	5
Stereo camera	Almost same as regular camera						

1

[1] Yuxiao Zhang и др. «Perception and sensing for autonomous vehicles under adverse weather conditions: A survey». В: *ISPRS Journal of Photogrammetry and Remote*

Мотивация

	Modality	mAP (*val*)	NDS (*val*)	Latency (ms)
BEVFusion (Ours)	C+L	**68.5**	**71.4**	**119.2**
RCTrans(ours)	RC	**52.0**	**59.4**	60.5

Постановка задачи

Общая формула задачи трёхмерного обнаружения объектов:

$$B = f_{\text{det}}(I, R), \tag{1}$$

где $B = \{B_1, \ldots, B_N\}$ - это N трёхмерных предсказаний на сцене, f_{det} - нейросетевая модель 3д-детекции, I - изображения с нескольких камер, R - облако точек радара.
Трёхмерное предсказание представляет собой параллелепипед, который ограничивает объект, и записывается так:

$$B = [x_c, y_c, z_c, l, w, h, v_x, v_y, \theta, \text{class}], \tag{2}$$

где (x_c, y_c, z_c) это трёхмерные координаты центра, l, w, h - длина, ширина и высота параллелепипеда соответственно, v_x, v_y - скорости объекта вдоль осей x и y, θ - угол поворота вдоль вертикальной оси, class соответствует категории объекта, например, машина, грузовик, пешеход и т.д.

Постановка задачи

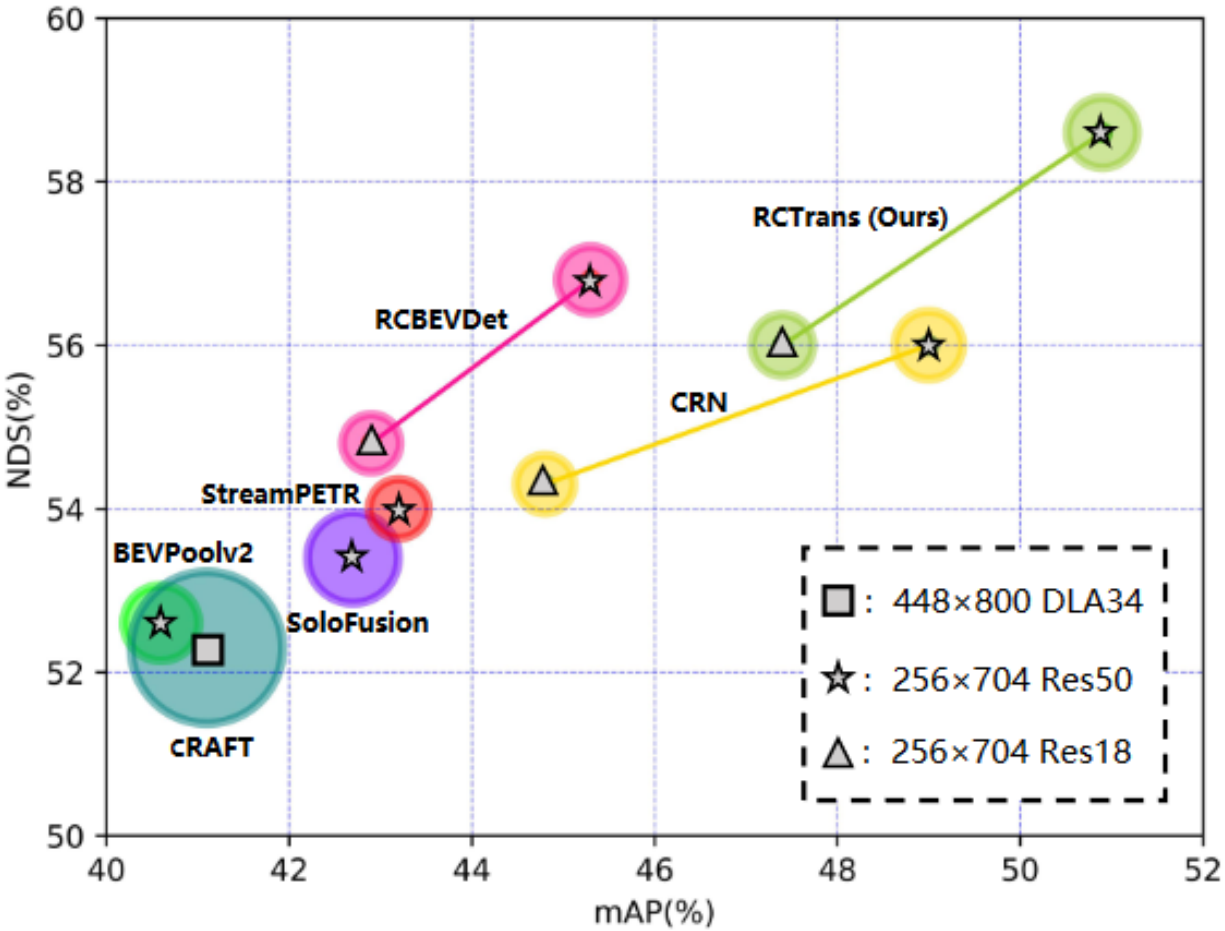
Основная метрика качества в задаче трёхмерного обнаружения объектов - mean average precision (mAP).
На датасете nuScenes используется широко известная метрика Average Precision (AP), но соответствие предсказаний и разметки (ground truth) определяется по расстоянию между центрами в горизонтальной плоскости (без учета высоты), а не по IoU.
Для каждого порога соответствия рассчитывается AP как площадь под кривой Precision-Recall в диапазоне Recall > 0.1 и Precision > 0.1.
В конце метрика усредняется порогам, а затем вычисляется среднее значение по всем классам:

$$\text{mAP} = \frac{1}{|C||D|} \sum_{c \in C} \sum_{d \in D} \text{AP}_{c,d}, \tag{3}$$

где C - множество классов, D - пороги расстояния между центрами $\{0.5, 1, 2, 4\}$ м.

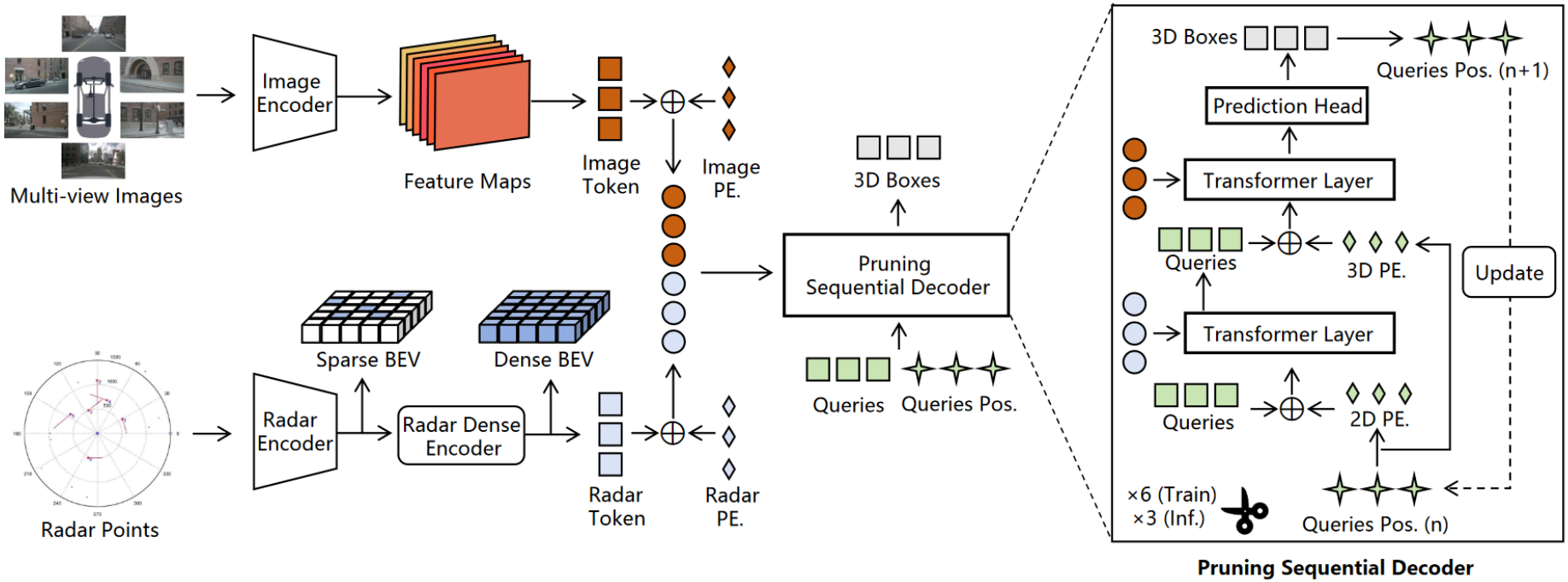
Базовый метод

RCTrans



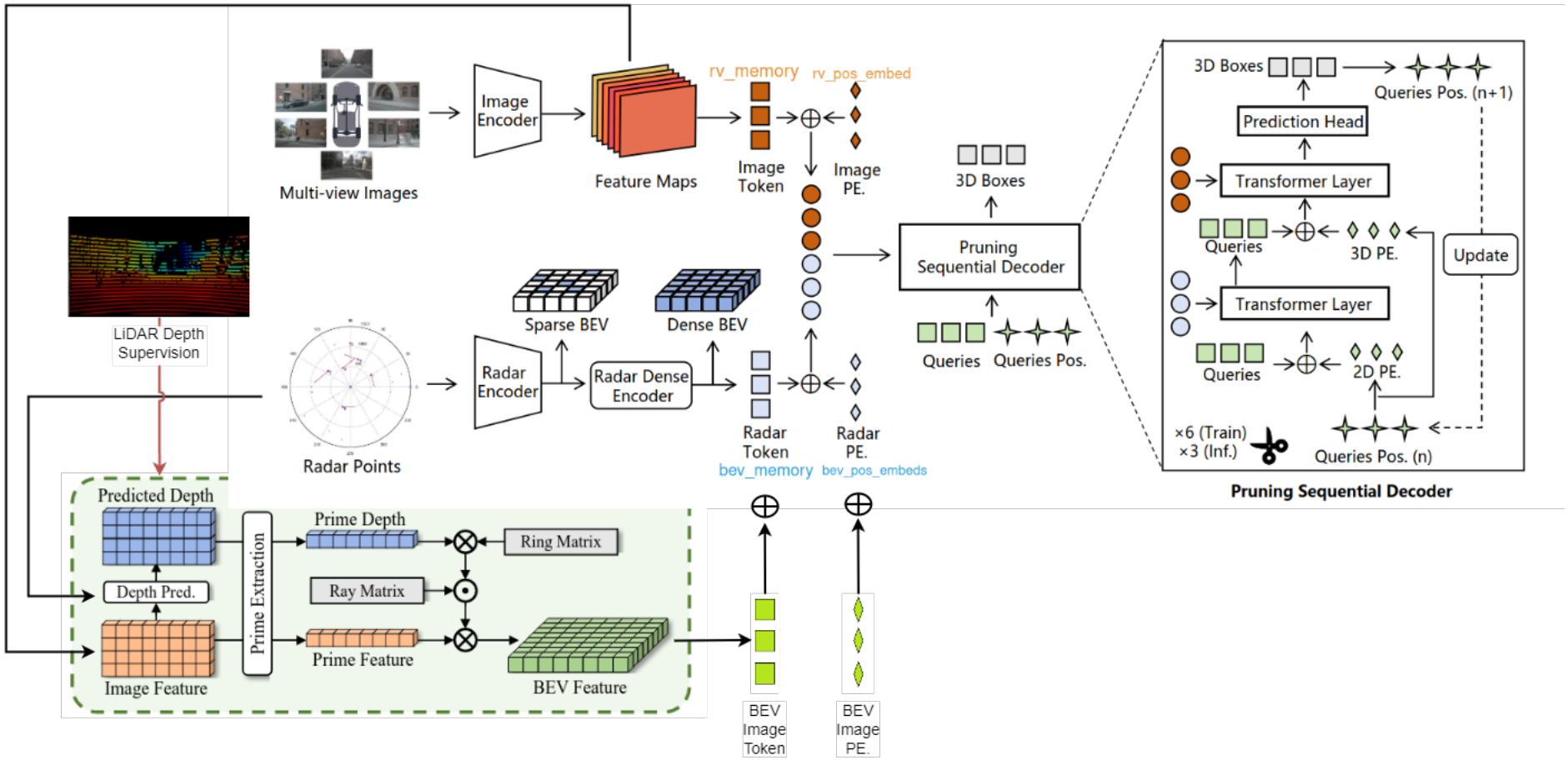
Сравнение RCTrans с другими камерными и камерно-радарными методами 3D-детекции. Размер круга иллюстрирует время инференса

Базовый метод



Общая архитектура Radar-Camera Transformer

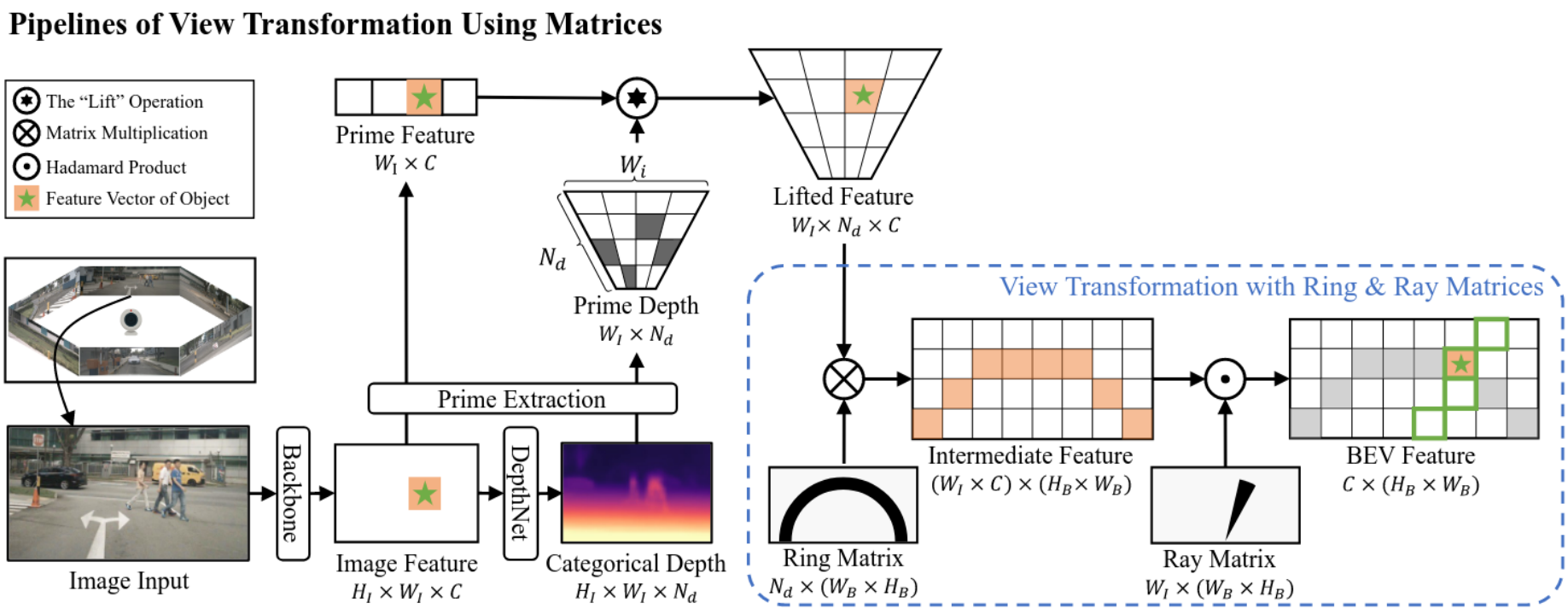
Предложенные изменения



- Дополнительное вычисление BEV-признаков по изображениям с помощью выученной глубины
- Глубина вычисляется по облаку точек радара и признакам изображения

Эксперименты

Проведен эксперимент по добавлению глубины к модели для вычисления BEV-признаков по изображениям, глубина вычислялась по признакам изображения



MatrixVT

Эксперименты

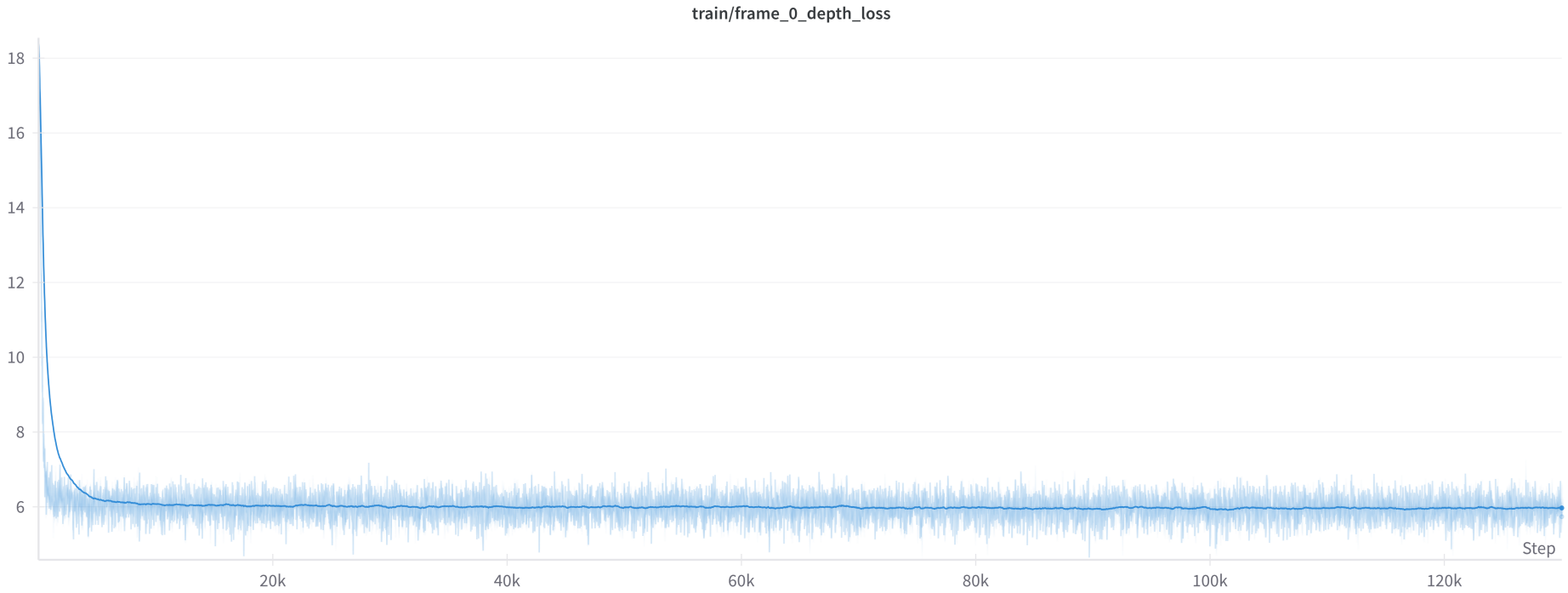
Метрика	RCTrans (90 epochs)	Our (18 epochs)
mAP	0.4741	0.1564
mATE	0.5399	0.9740
mASE	0.2735	0.3245
mAOE	0.5566	0.9801
mAVE	0.2081	0.7089
mAAE	0.1899	0.2501
NDS	0.5602	0.2545

Таблица: Сравнение метрик между оригинальным RCTrans и моделью с модификациями

Нужно продолжать обучение для формирования выводов по гипотезе

Эксперименты

Зачем добавлять радар для вычисления глубины?



Функция потерь между (gt) лидарной глубиной и глубиной, полученной по признакам изображения

- В условиях плохой видимости глубина по изображениям будет низкой точности
- Падение лосса останавливается

Заключение

- Предложены модификации SOTA-метода 3D-детекции
- Начаты эксперименты по проверке гипотезы о том, что добавление глубины позволит повысить метрики

Спасибо за внимание!

matykina.ov@phystech.edu