

Применение нейронных сетей в задаче одновременного картирования и локализации

Муравьев Кирилл, группа 599

Научный руководитель
Яковлев Константин, ФИЦ ИУ РАН

vSLAM: актуальность задачи

- Мобильные роботы
- Малые БПЛА
- Автономное функционирование
- Планирование траектории
- Ограничения по сенсорам



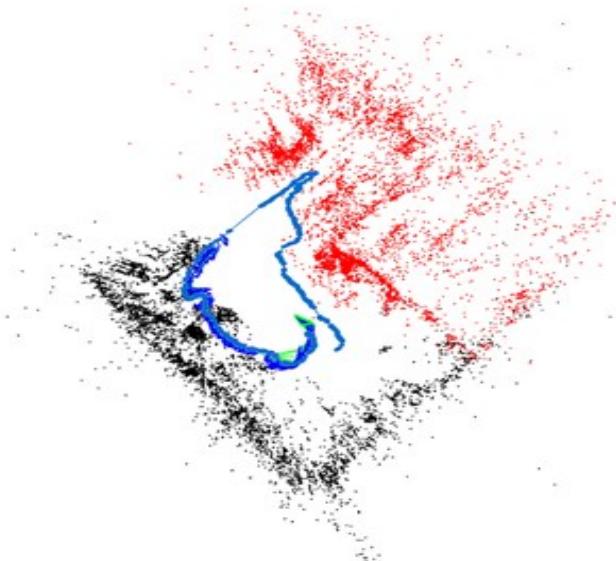
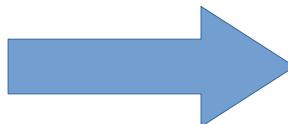
vSLAM: постановка задачи

Дано:

- Последовательность кадров $\{P_t\}$

Необходимо найти:

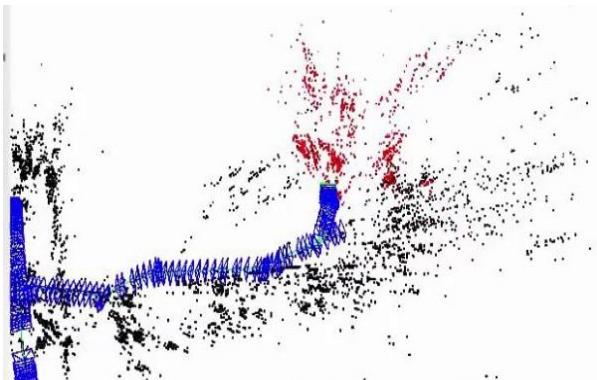
- Координаты точек сцены M_i
- Координаты всех положений камеры R_t



vSLAM: подходы к решению задачи

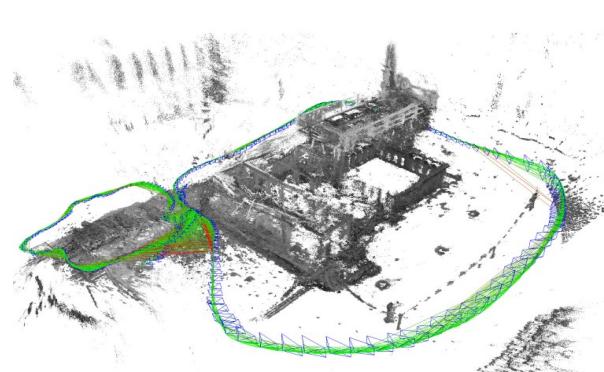
ORB-SLAM

- Данные: видео
- Карта: разреженная
- Скорость: высокая



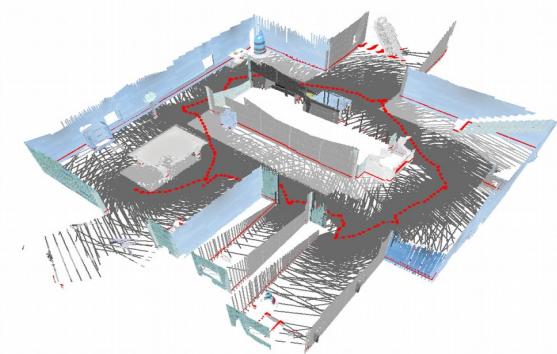
LSD-SLAM

- Данные: видео
- Карта: semi-dense
- Скорость: низкая



RTABMAP

- **Данные: RGB-D**
- Карта: плотная
- Скорость: высокая



RTABMAP + Depth Reconstruction:

данные — видео, карта — плотная, скорость — высокая

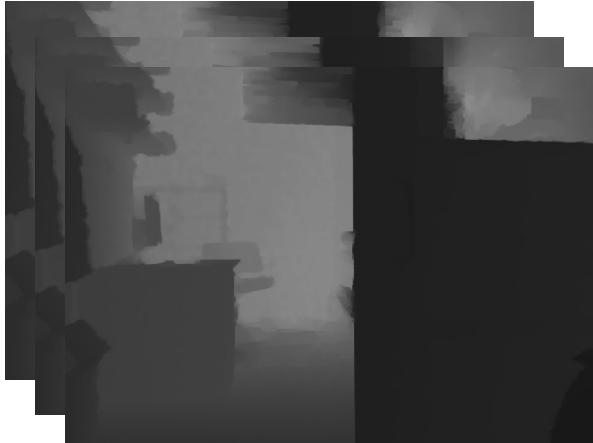
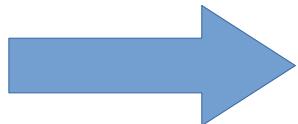
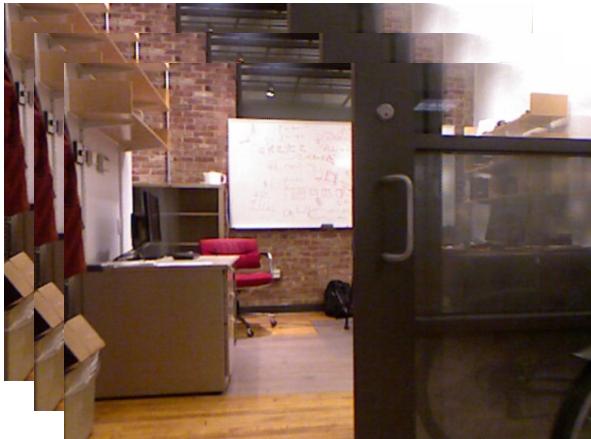
Восстановление глубины: постановка задачи

Дано:

- Последовательность кадров $\{P_t\}$

Необходимо найти:

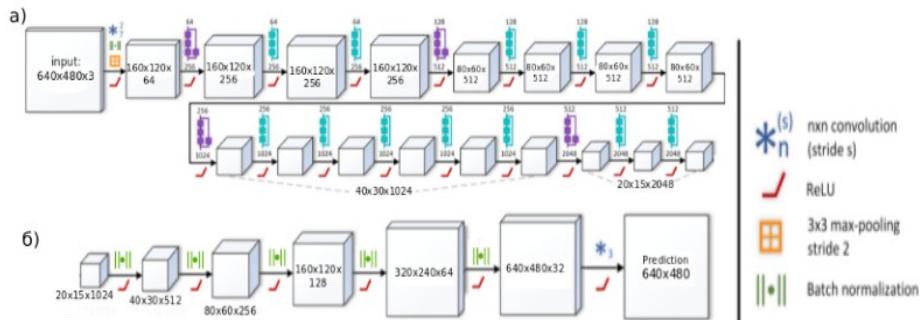
- $\{D_t\}$ — расстояния от камеры до точек сцены



Восстановление глубины: известные методы

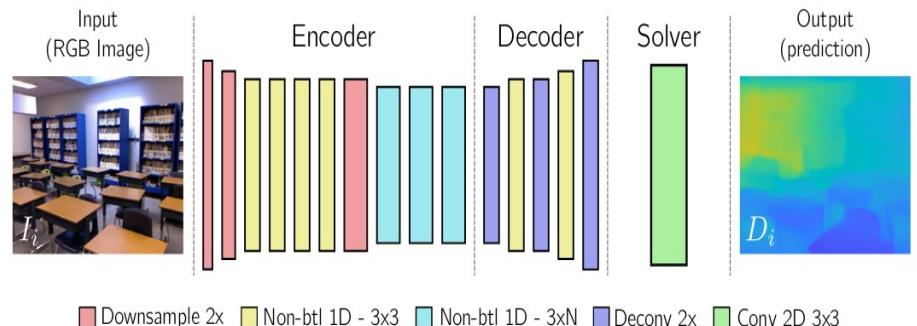
Laina et al, 2016

- Высокое качество восстановления
- Низкая скорость работы



Spek et al, 2018

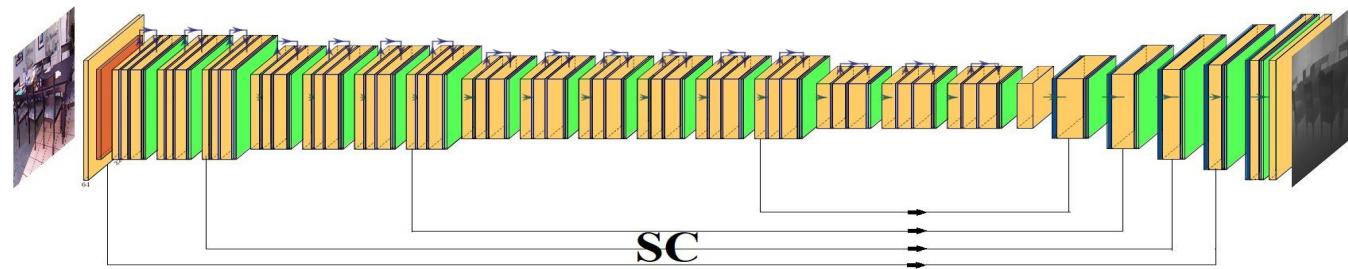
- Высокая скорость работы
- Низкое качество восстановления



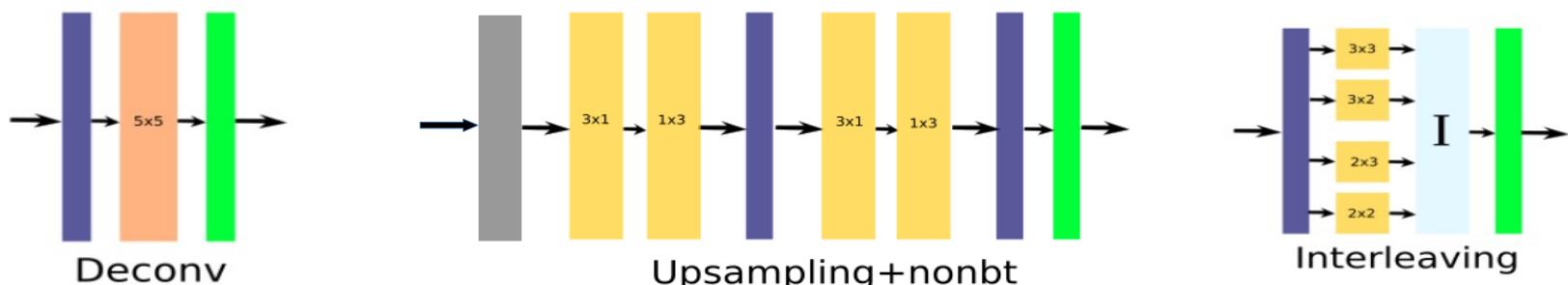
Восстановление глубины с помощью нейросетей: предложенные архитектуры

Энкодер:
ResNet-50

Декодер:
5 блоков развертки



Варианты блоков декодера:



Conv2D

BatchNorm

ReLU

Deconv

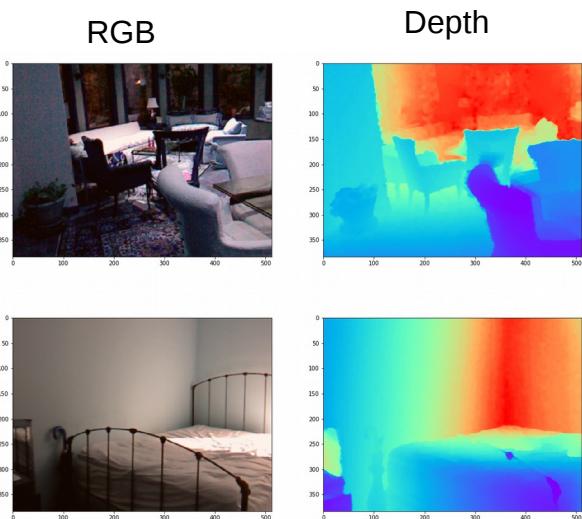
Unpooling

Interleaving

Восстановление глубины с помощью нейросетей: детали обучения

Обучающий датасет: NYU Depth v2

- 464 видеосцены
- Более 400000 картинок
- Глубины от 0.5м до 10м



Функции потерь

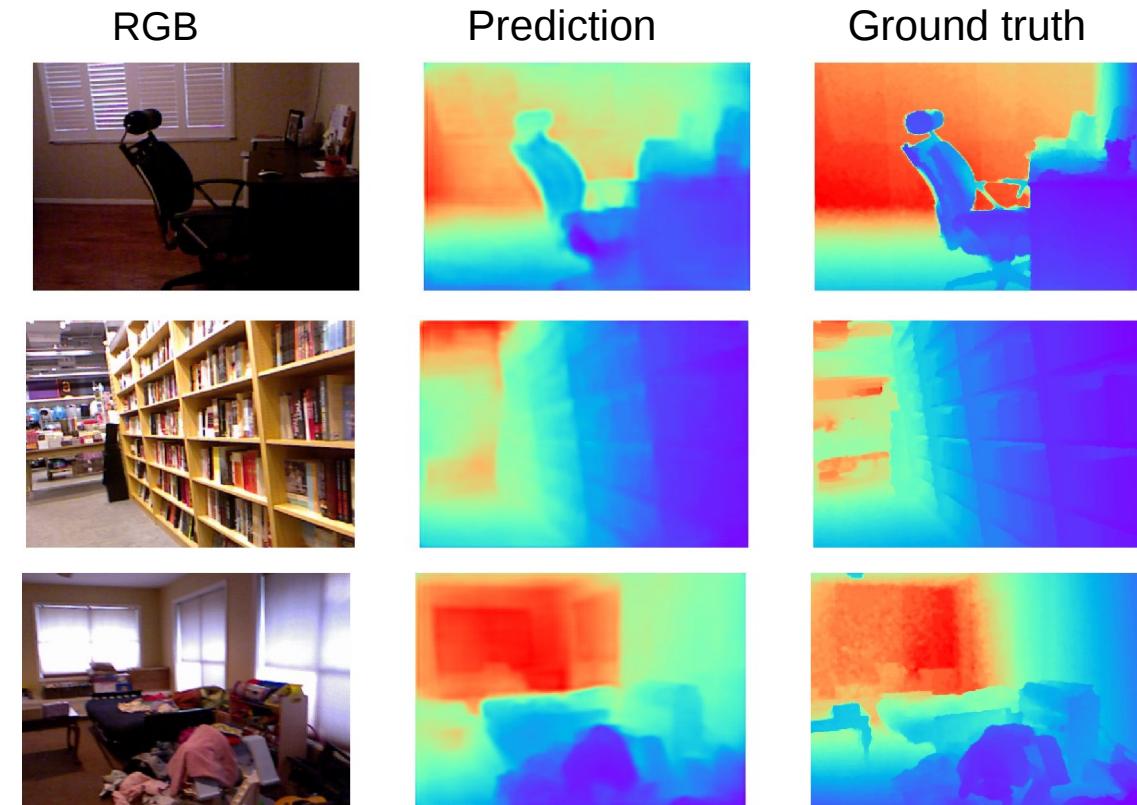
- MSE: $L(y, \hat{y}) = (y - \hat{y})^2$
- Relative squared error: $L(y, \hat{y}) = (\frac{y - \hat{y}}{y})^2$
- MSE + Rel: $L(y, \hat{y}) = (y - \hat{y})^2(1 + \frac{2}{y^2})$
- BerHu: $L(y, \hat{y}) = \begin{cases} |y - \hat{y}|, & |y - \hat{y}| < \delta \\ \frac{(y - \hat{y})^2 + \delta^2}{2\delta}, & |y - \hat{y}| \geq \delta \end{cases}$

Метрики качества

- RMSE: $L(y, \hat{y}) = \sqrt{(y - \hat{y})^2}$
- Relative error: $L(y, \hat{y}) = \frac{|y - \hat{y}|}{y}$

Восстановление глубины с помощью нейросетей: качество на NYU Depth v2

Архитектура	RMSE, м	Rel error	Скорость, FPS
Upsampling + nonbt	0.639	0.173	16
Interleaving	0.675	0.180	16
Deconv	0.683	0.186	13
FCRN (Laina et al, 2016)	0.573	0.127	12
Cream (Spek et al, 2018)	0.687	0.190	30



Восстановление глубины с помощью нейросетей: тестирование в реальном времени

Устройство: NVIDIA Jetson TX2

- 128 CUDA-ядер
- Процессорный модуль 50x87 мм
- Энергопотребление 15 Вт

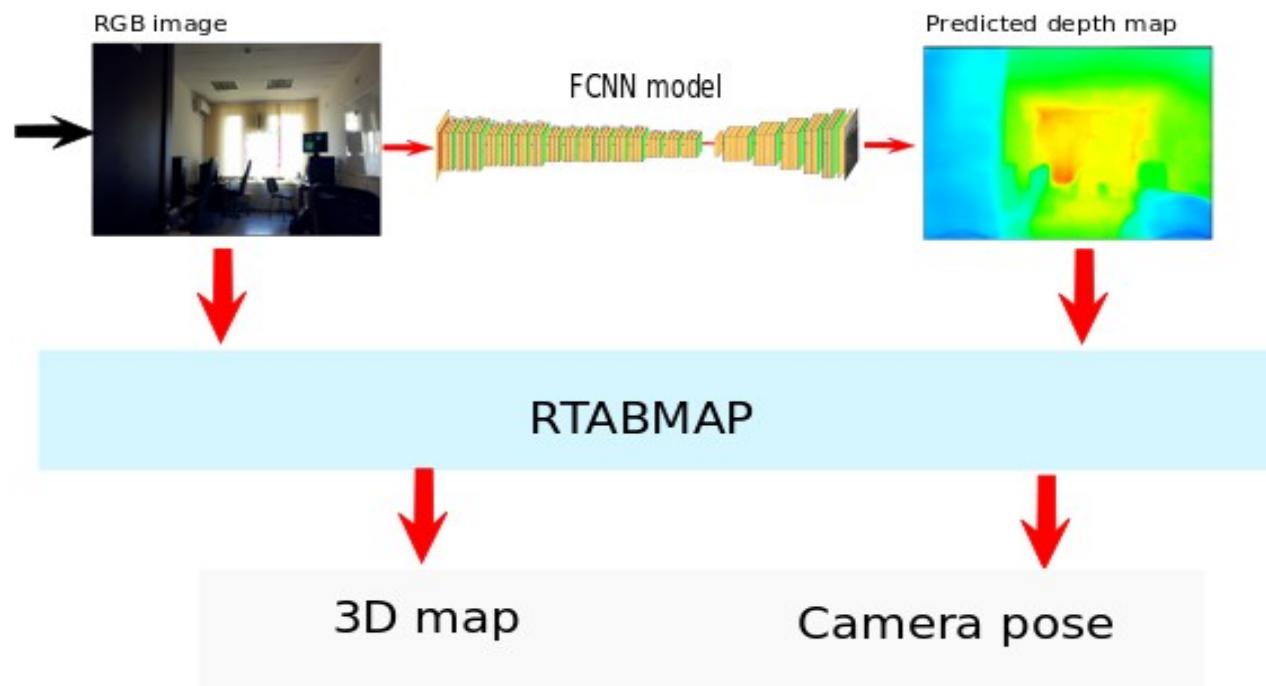
Параметры

- Декодер Upsampling + nonbt
- Разрешение 320x240
- TensorRT, float16



SLAM с применением нейросетей

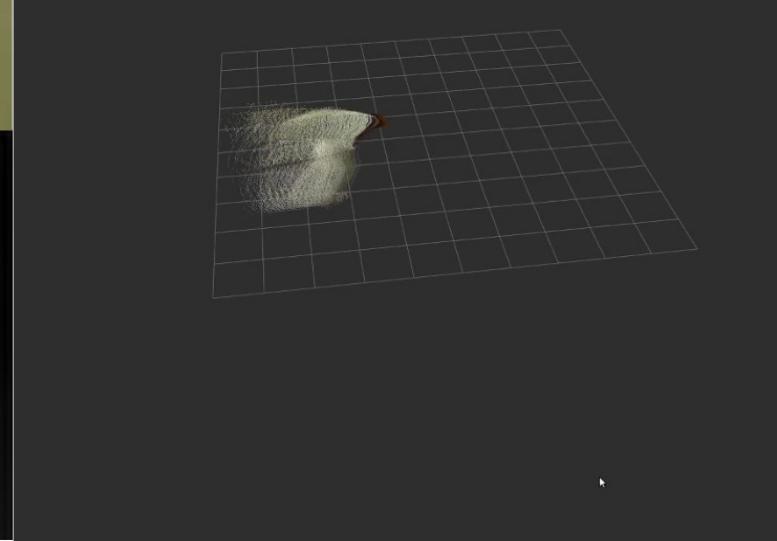
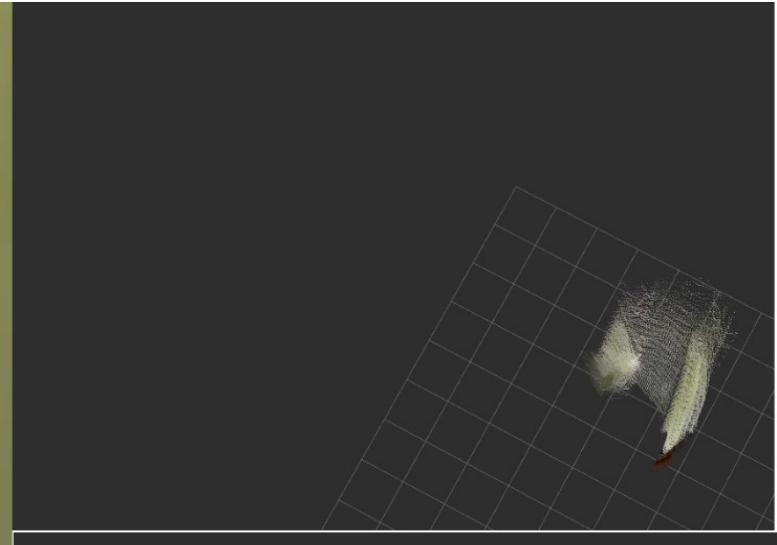
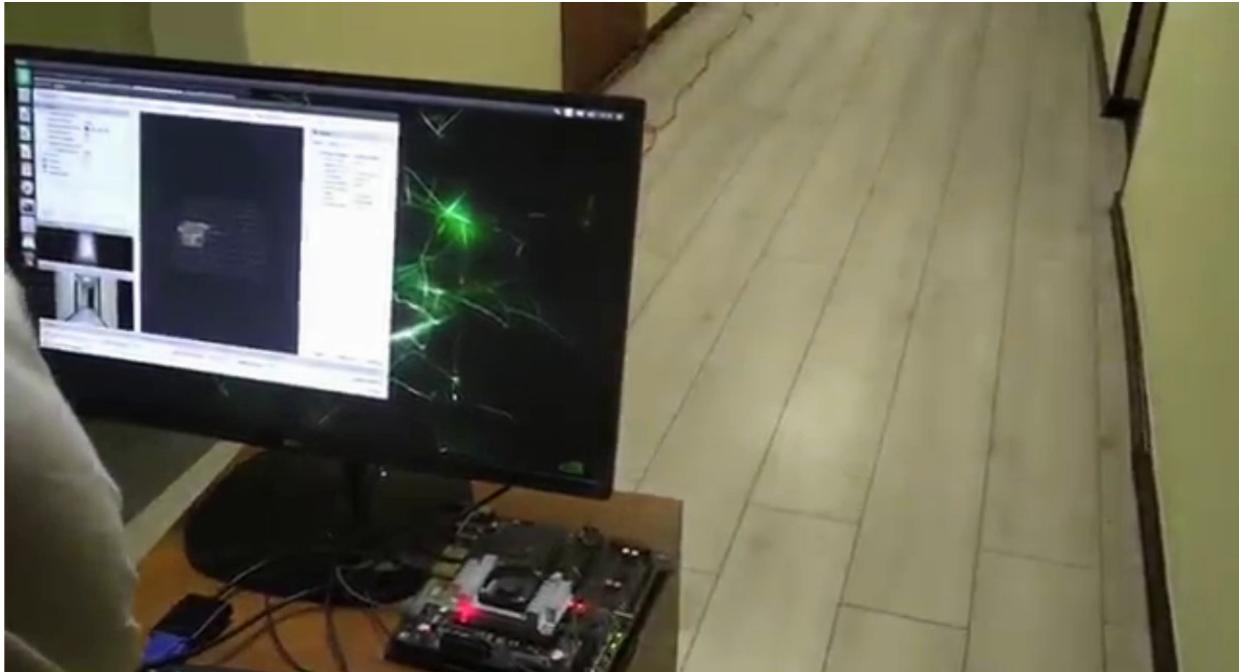
Схема алгоритма



Построенная карта



Код алгоритма под ROS: https://github.com/cnndepth/tx2_fcnn_node



Результаты

- Предложены архитектуры нейросетей для восстановления глубины
- Достигнуты real-time скорость и высокое качество восстановления
- Реализован метод vSLAM: FCNN Depth Reconstruction + RTABMAP
- Проведены эксперименты в реальном времени на NVIDIA Jetson
- Исходный код¹ интегрирован с ROS
- Сделан доклад на семинаре БТС-ИИ 2019
- Подана статья на конференцию ECMR-2019

¹https://github.com/cnndepth/tx2_fcnn_node