

Оценка транспортного потока по данным с камеры видеонаблюдения

Федоров А.Н., к. ф.-м. н. Иванов С.А.

Постановка задачи

- Разработать систему подсчета и классификации транспортных средств по направлениям их движения на перекрестке пр. Победы и ул. Молодогвардейцев г. Челябинска
- Обнаружение транспортного средства
- Отслеживание транспортного средства
- Оценка направления транспортного средства

Современное положение

- CCTV (система телевидения замкнутого контура)
- Современные системы детектирования объектов
- Faster R-CNN
- YOLO
- SSD

Сопутствующие работы

- *Обнаружение объектов*
- *Отслеживание нескольких объектов*

Обнаружение объектов

Одноступенчатые детекторы:

- YOLO, SSD
- Anisimov D, Khanova T. Towards lightweight convolutional neural networks for object detection. In: 2017 14th IEEE international conference on advanced video and signal based surveillance (AVSS). 2017, p. 1–8.
- Sang J, Wu Z, Guo P, Hu H, Xiang H, Zhang Q, Cai B. An improved YOLOv2 for vehicle detection. Sensors. 2018;18(12):4272.

Обнаружение объектов

Двухступенчатые детекторы:

- Семейство R-CNN
- Girshick R. Fast R-CNN. In: Proceedings of the IEEE international conference on computer vision 2015 Inter, 2015, p. 1440–8.
- Girshick R, Donahue J, Darrell T, Malik J. Rich feature hierarchies for accurate object detection and semantic segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2012, p. 580–7.
- He K, Gkioxari G, Dollar P, Girshick R. Mask R-CNN. In: 2017 IEEE international conference on computer vision (ICCV). vol. 2017 Oct, 2017, p. 2980–8. IEEE.

Обнаружение объектов

Двухступенчатые детекторы:

- Ren S, He K, Girshick R, Sun J. Faster R-CNN: towards real-time object detection with region proposal networks. IEEE Trans Pattern Anal Mach Intell. 2017; 39(6): 1137–49.
- Li S. 3D-DETRNet: a single stage video-based vehicle detector. 2018.
- Wang X, Cheng P, Liu X, Uzochukwu B. Focal loss dense detector for vehicle surveillance. In: 2018 international conference on intelligent systems and computer vision (ISCV). vol. 2018 May, 2018, p. 1–5. IEEE.
- Lin TY, Goyal P, Girshick R, He K, Dollar P. Focal loss for dense object detection. In: Proceedings of the IEEE international conference on computer vision. 2017 Oct, p. 2999–3007.
- Hu X, Xu X, Xiao Y, Chen H, He S, Qin J, Heng PA. SINet: A scale-insensitive convolutional neural network for fast vehicle detection. In: IEEE transactions on intelligent transportation systems. vol. 20, no. 3. 2019, p. 1010–9.

Отслеживание нескольких объектов

- Multiple Hypothesis Tracking

Kim C, Li F, Ciptadi A, Rehg JM. Multiple hypothesis tracking revisited. In: 2015 IEEE international conference on computer vision (ICCV). vol. 22. 2015, p. 4696–704. IEEE.

- Joint Probabilistic Data Association Filter

Rezatofighi SH, Milan A, Zhang Z, Shi Q, Dick A, Reid I. Joint probabilistic data association revisited. In: 2015 IEEE international conference on computer vision (ICCV). 2015, p. 3047–55. No. December, IEEE.

- SORT tracker

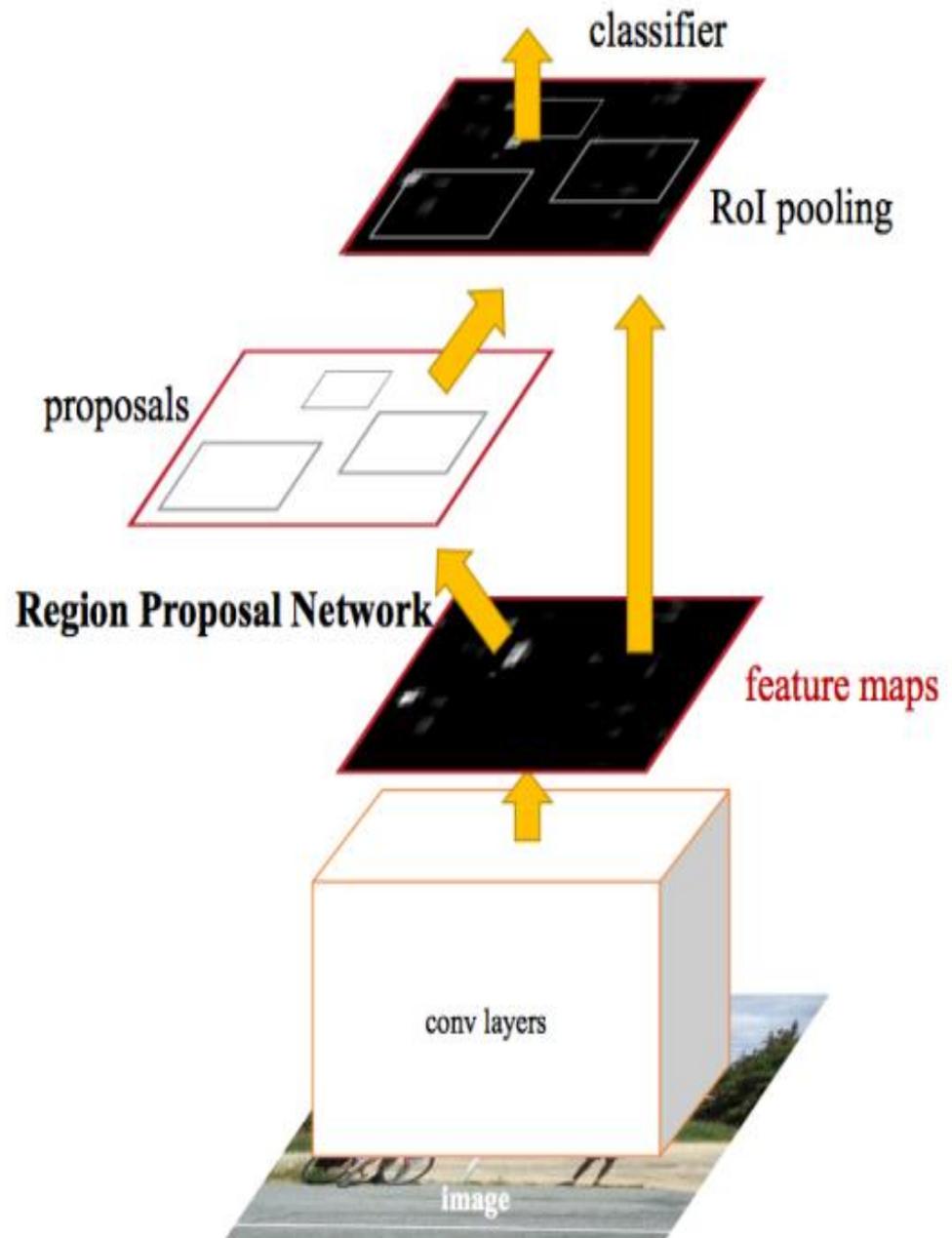
Bewley A, Ge Z, Ott L, Ramos F, Upcroft B. Simple online and realtime tracking. In: 2016 IEEE international conference on image processing (ICIP). vol. 2016-Aug, p. 3464–8. IEEE. 2016.

- CNN trackers

- Chu P, Ling H. FAMNet: Joint learning of feature, affinity and multi-dimensional assignment for online multiple object tracking. 2019.

- Wojke N, Bewley A, Paulus D. Simple online and realtime tracking with a deep association metric. In: 2017 IEEE international conference on image processing (ICIP). vol. 2017 Sept, 2017, p. 3645–9. IEEE.

Faster R-CNN



Dataset

COCO Annotator

60 000 объектов на 982 кадрах

Проведена аугментация данных



Dataset

- 786 изображений для обучения, 196 для проверки.
- Из-за ограниченного объема памяти графического процессора мы также масштабировали каждое изображение до разрешения 1280 × 512.

Тип машины	Всего экземпляров	Среднее количество экземпляров на изображение
Автомобиль	53083	54.06
Фургон	2783	2.83
Грузовая машина	2415	2.46
Трамвай	1298	1.32
Автобус	1234	1.26
Троллейбус	611	0.62

Оценка транспортных потоков

- Была собрана статистика прохождения транспортных средств с 8:00 до 12:00 и с 17:00 до 19:00 по 4 наиболее популярным направлениям движения.

Модуль обнаружения

Была улучшена базовая производительность Faster R-CNN с помощью:

- 1) магистрали пирамидальных сетей (FPN)
- 2) дополнительной ветви маски
- 3) оптимизации формы якорей
- 4) потери фокуса и адаптивного объединения функций

Ветвь маски

- He K, Gkioxari G, Dollar P, Girshick R. Mask R-CNN. In: 2017 IEEE international conference on computer vision (ICCV). vol. 2017 Oct, 2017, p. 2980–8. IEEE

Оптимизация формы якорей

Существует эффективный способ выбрать якоря с помощью кластеризации K-Means с помощью меры расстояния, основанной на пересечении по объединению (IOU).

Поиск по сетке.

Цель поиска состояла в том, чтобы найти набор аспектов и шкал, которые максимизируют IOU между объектами в обучающем наборе и якорями, которые сопоставляются с ними во время обучения.

Потеря фокуса

Фокальные потери (FL) служат заменой обычной перекрестной энтропии и могут быть определены как:

$$p_t = \begin{cases} p & \text{if } y = 1 \\ 1 - p & \text{otherwise,} \end{cases}$$

$$FL(p_t) = -\alpha_t(1 - p_t)^\gamma * \log(p_t)$$

Здесь $y \in \pm 1$ указывает целевой класс, а $p \in [0,1]$ является оценочной вероятностью для класса с меткой $y = 1$. Параметр γ используется для уменьшения веса легких негативов, $\alpha_t \in [0,1]$ является дополнительным уравнивающим фактором и определяется как α для класса 1 и $1 - \alpha$ для класса - 1.

$$\gamma = 2, \alpha = 0.25$$

Адаптивное объединение функций

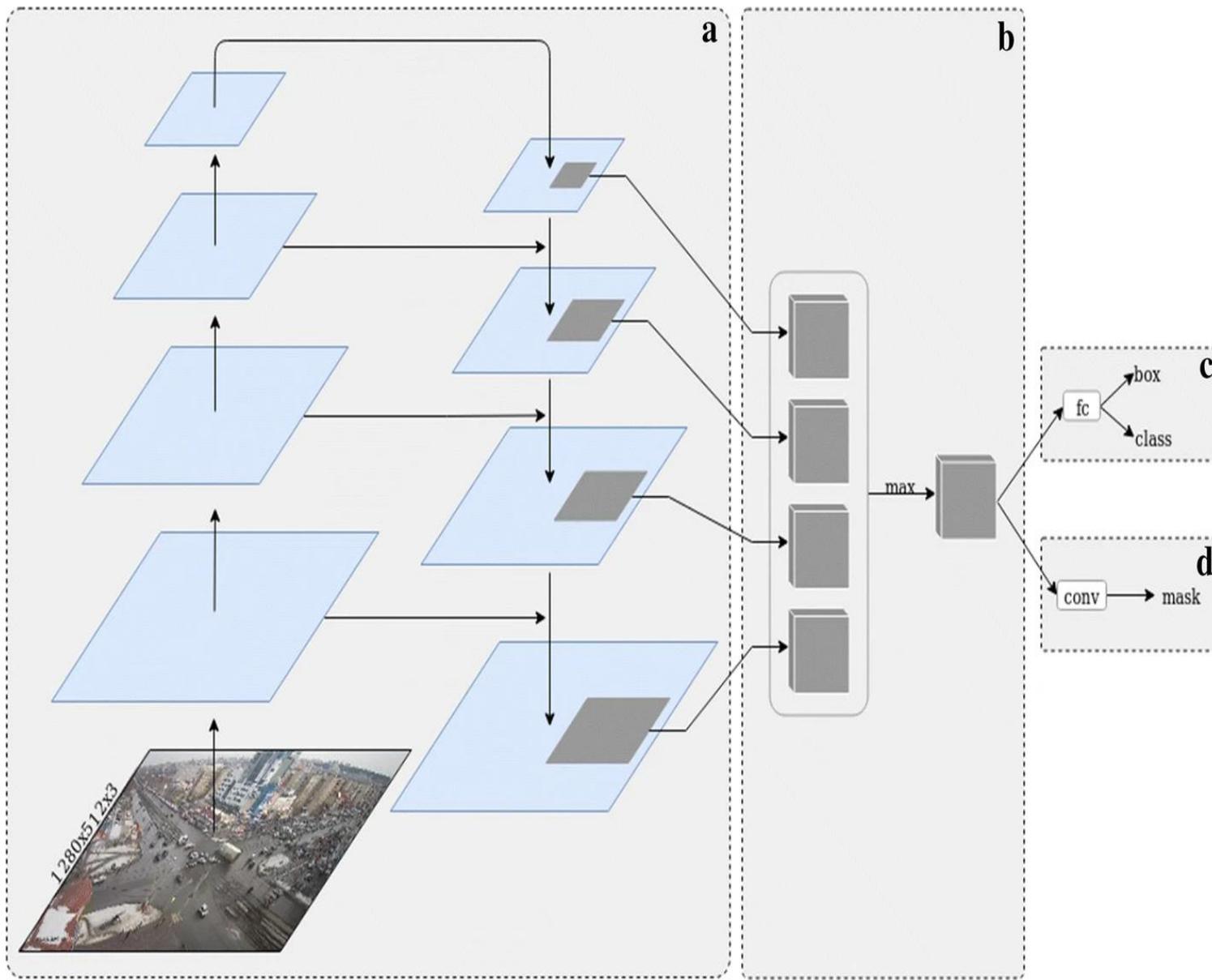
Производительность модели была повышена с помощью адаптивного объединения функции ROI для каждого уровня FPN:

объединение функции с нескольких уровней с помощью функции `max`.

Детали сети обнаружения

- Использована модель ResNet-50 в качестве основы вместе со стандартной конфигурацией FPN от Detection.
- Стохастический градиентный спуск с импульсом был использован для обучения сети. Начальная скорость обучения равна $1e-3$ и снижена в 10 раз после итераций 20k и 27k. В общей сложности обучение было прекращено после 30k итераций. Batch size, равный 4, для всех экспериментов.
- Для ускорения сходимости были использованы веса, предварительно обученные на наборе данных MS COCO. Перед обучением всей сети мы подстроили последние уровни RPN и ветви классификации/маски на 1k итераций с линейной скоростью снижения скорости обучения от $25e-3$ до $1e-3$.

Топология сети



Модуль отслеживания

Модуль отслеживания предлагаемой системы отвечает на два вопроса:

Откуда взялся каждый автомобиль?

Куда он направлялся?

Трекер SORT.

SORT обеспечивает необходимую функциональность для отслеживания объектов в нескольких кадрах.

Модуль отслеживания

- Сначала разделяем перекресток на несколько непересекающихся регионов.
- Пара исходных и целевых дорог однозначно описывает направление движения транспортного средства.
- Проблема незаконченных траекторий.

Эксперименты

- Все эксперименты проводились на Ubuntu 18.04 с графическим процессором Nvidia RTX 2080 Ti, 12 ядрами процессора (процессор Intel® Core i7-8700 @ 3,20 ГГц) и 32 ГБ оперативной памяти. Мы создали наше решение на основе официальной реализации Facebook Detectr.

Протокол оценки

Оценки:

- среднюю точность (AP),
- среднюю точность при IOU_{0,5} (AP_{0,5}),
- среднюю точность при IOU_{0,75} (AP_{0,75}),
- полнота AP для ограничивающих рамок (AP^{bb}),
- средняя точность для сегментации экземпляра (AP^{segm}).

Для задачи оценки потока трафика:

- средняя абсолютная ошибка,
- средняя абсолютная процентную ошибку.

Результаты

Производительность обнаружения

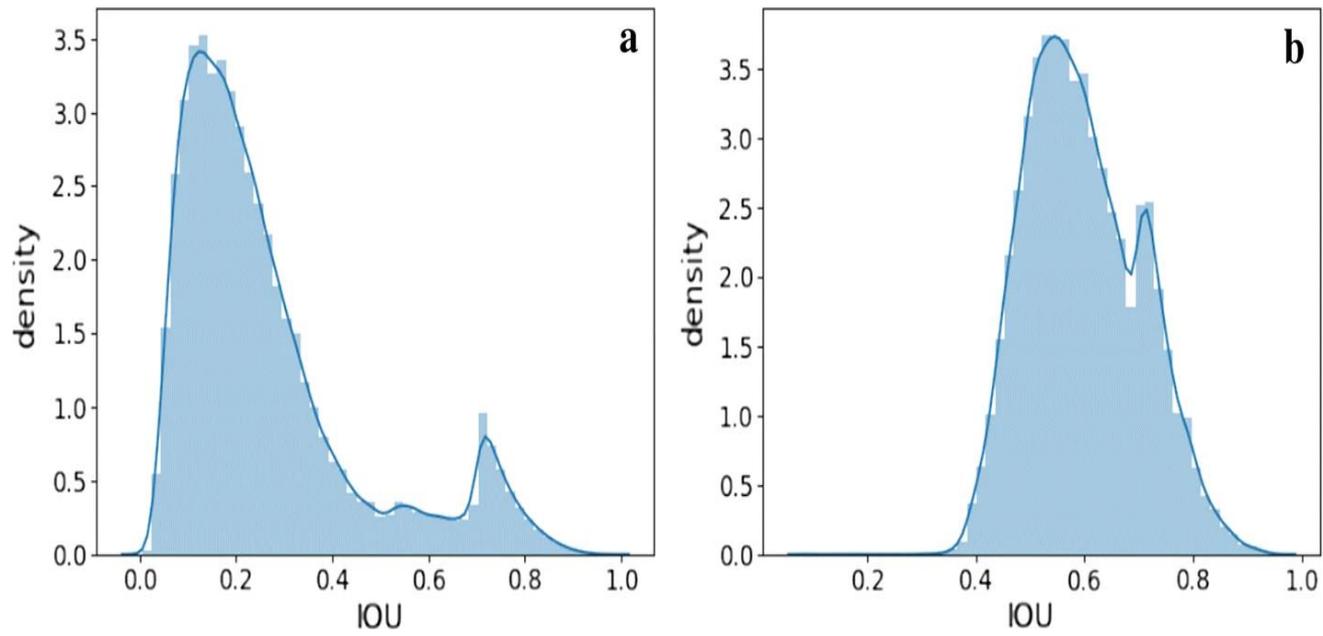
В таблице приведены результаты базовой модели Faster R-CNN наряду с влиянием каждого предложенного улучшения.

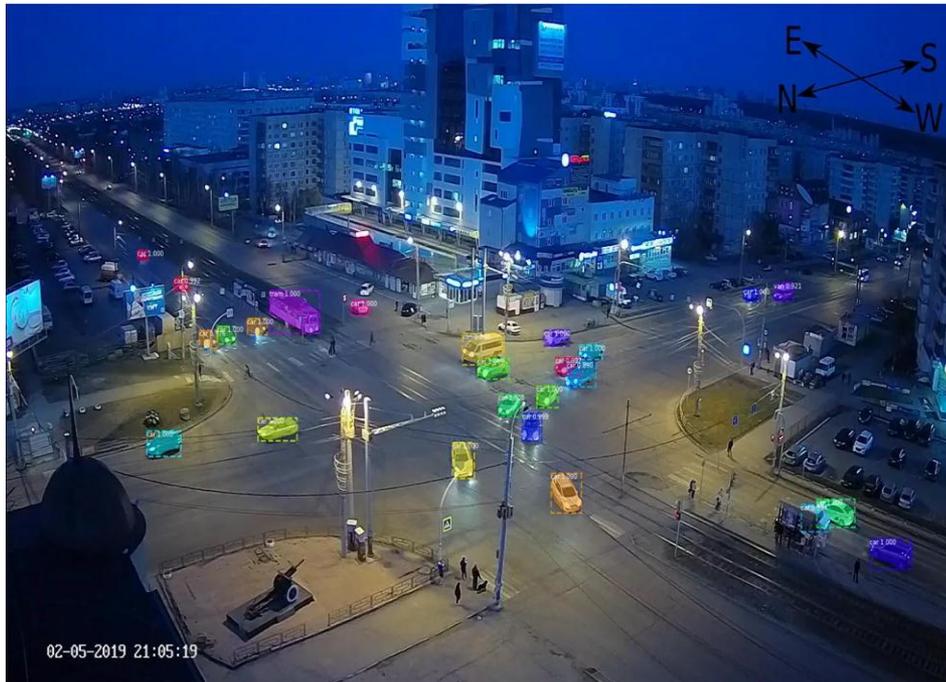
	Baseline					Overall gain
Mask branch	X	✓	✓	✓	✓	
Anchors optimization	X	X	✓	✓	✓	
Focal loss	X	X	X	✓	✓	
Adaptive feature pooling	X	X	X	X	✓	
A_{pbb}	57.3	58.2	58.4	58.9	59.2	+1.9
A_{50}^{pbb}	79.0	79.7	79.7	79.7	80.3	+1.3
A_{75}^{pbb}	67.5	68.8	68.9	69.4	70.0	+2.5
A_{psegm}	–	56.1	56.4	56.7	57.2	+1.1
A_{50}^{segm}	–	78.9	79.4	79.5	80.2	+1.3
A_{75}^{segm}	–	68.0	69.2	68.8	69.1	+1.1

Italic values indicate the performance gain relative to the Faster R-CNN baseline when applying all the proposed improvements

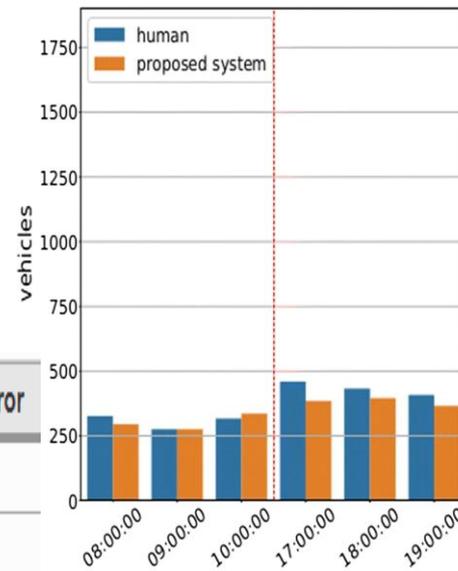
Производительность обнаружения

Мы ожидали большего выигрыша от оптимизации якорей, поскольку на рисунке четко указано, что якоря, выбранные с помощью процедуры поиска по сетке, обеспечивают лучший охват основных истинных блоков

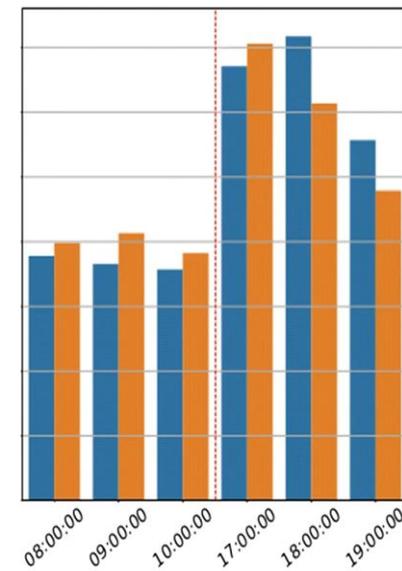




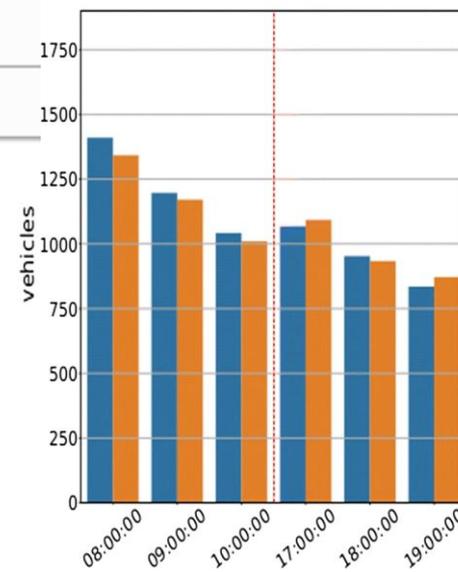
Direction	Mean absolute error	Mean absolute percentage error
North-East	34.17	8.49
West-East	129.0	9.85
East-West	34.5	3.14
South-West	22.67	7.53
Average	55.08	7.25



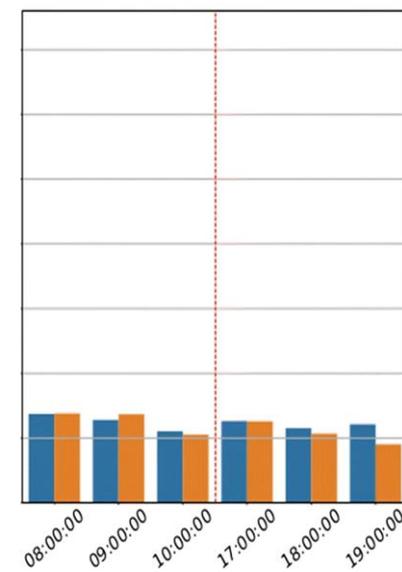
a North-East



b West-East



c East-West



d South-West

Исследование ошибок

- Длительные окклюзии между транспортными средствами в плотном транспортном потоке.
- Разработка эффективных алгоритмов для повторной идентификации транспортных средств.

Доступность данных и материалов

Весь код доступен онлайн,

<https://github.com/alnfedorov/traffic-analysis>.

Fedorov A., Nikolskaia K., Ivanov S., Shepelev V., Minbaleev A. Traffic flow estimation with data from a video surveillance camera // Journal of Big Data. 2019. Vol. 6. Article 73. DOI: [10.1186/s40537-019-0234-z](https://doi.org/10.1186/s40537-019-0234-z), indexing: [Scopus Top10%](#).