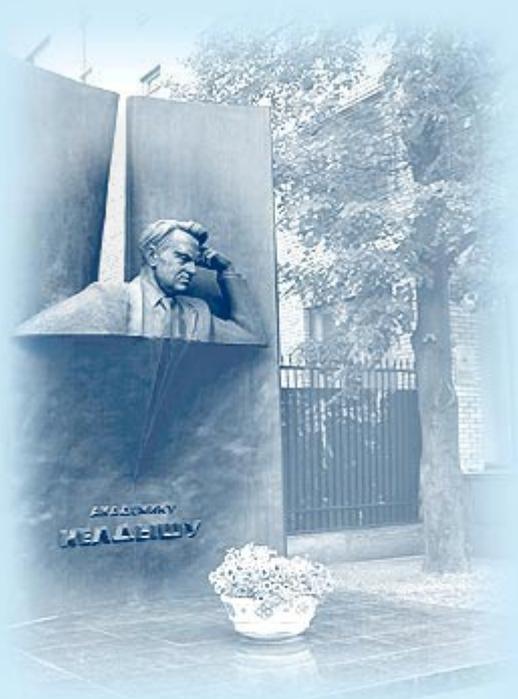




ИПМ им.М.В.Келдыша РАН • Электронная библиотека

Препринты ИПМ • Препринт № 104 за 2021 г.



ISSN 2071-2898 (Print)
ISSN 2071-2901 (Online)

А.В. Подопрсветов, Д.А. Анохин,
К.И. Кий, И.А. Орлов

Сравнительный анализ
метода геометризованных
гистограмм и нейросетевого
метода для распознавания
дорожной разметки

Рекомендуемая форма библиографической ссылки: Сравнительный анализ метода геометризованных гистограмм и нейросетевого метода для распознавания дорожной разметки / А.В. Подопрсветов [и др.] // Препринты ИПМ им. М.В.Келдыша. 2021. № 104. 22 с.
<https://doi.org/10.20948/prepr-2021-104>
<https://library.keldysh.ru/preprint.asp?id=2021-104>

**Ордена Ленина
ИНСТИТУТ ПРИКЛАДНОЙ МАТЕМАТИКИ
имени М.В.Келдыша
Российской академии наук**

А.В.Подопросветов, Д.А.Анохин, К.И.Кий, И.А.Орлов

**Сравнительный анализ метода
геометризованных гистограмм
и нейросетевого метода
для распознавания дорожной разметки**

Москва — 2021

Подопросветов А.В., Анохин Д.А., Кий К.И., Орлов И.А.

Сравнительный анализ метода геометризированных гистограмм и нейросетевого метода для распознавания дорожной разметки

В данной работе приведено сравнение двух подходов к определению дорожной разметки по видео последовательностям, а именно метода поиска разметки с помощью геометризированных гистограмм и метода, основанного на нейронных сетях. Для проведения сравнительного анализа работы алгоритмов используется независимый открытый датасет TuSimple. Так как исследуемые методы имеют разную архитектуру, их работа оценивается по следующим метрикам: точность (Accuracy), скорость (relative FPS), общая вычислительная сложность алгоритма (TFlops).

Ключевые слова: геометризированная гистограмма, нейронная сеть, распознавание образов, дорожная разметка

Podoprosvetov A.V., Anokhin D.A., Kiy K.I., Orlov I.A.

Comparative analysis of the geometrized histogram method and the neural network method for road markings recognition

This paper compares two approaches to determining road markings from video sequences, namely, the method of finding the markings using geometrized histograms and the method based on neural networks. An independent open dataset TuSimple is used to conduct a comparative analysis of the algorithms. Since the investigated methods have different architectures, their work is evaluated according to the following metrics: Accuracy, speed (relative FPS), general computational complexity of the algorithm (TFlops).

Key words: *geometrized histogram, neural network, road markings*

Работа выполнена при поддержке Российского фонда фундаментальных исследований, проекты 19-08-01159, 19-01-00123.

Оглавление

Введение	3
1. Постановка задачи сравнения	4
2. Критерии сравнения	4
4. Описание нейросетевого метода.....	16
5. Сравнительный анализ работы методов	19
Заключение.....	21
Список литературы.....	22

Введение

Технологии обработки изображений для обнаружения объектов или событий часто требуют обработки потока кадров. Вследствие этого существует множество подходов к обработке видеорядов, в том числе в реальном времени. Их можно разделить как по целям применения, например: автономное вождение, медицинские приложения, системы безопасности, так и по методам исполнения. Классические методы сегментации [1], методы многоклассовой сегментации, основанные на обучении [2, 3] и методы, основанные на нейросетях [3, 4], дают несколько подходов к решению задач сегментации объектов на изображениях в реальном времени. Данные методы также содержат процедуру кластеризации в пространстве признаков, что является существенным для решения указанной задачи.

Кроме того, в связи со взрывным ростом исследований и разработок в области технологий компьютерного зрения, объединяющих сенсорную информацию для локализации и планирование пути в задачах автовождения, появились развитые системы помощи водителю (ADAS) и высокоуровневые системы автономного вождения (SDS), такие как Waymo [8], Uber [9], Lyft [18], Mobileye [19], Google car [20] и Tesla [21]. В частности, наиболее современные исследования и проекты [19, 21, 22] по ADAS и SDS больше ориентированы на использование камер, чем на другие сенсоры, такие как LiDAR, в связи с уменьшением стоимости и значительным повышением точности современных камер. Для систем ADAS и SDS характерно использование методов по обнаружению и отслеживанию транспортных средств, обнаружению препятствий, пониманию сцены и семантической сегментации, обнаружению полосы движения для позиционирования транспортного средства, удержания или смены полосы движения. В работах [4, 5, 23, 24] описан ряд исследований по обнаружению разметки. Подробный обзор данных методов содержится в [6-7]. Однако есть существенные ограничения на область их применения, поскольку необходимые вычислительные мощности, обеспечивающие достаточную точность, часто превышают рамки ресурсов, выделенных под задачу. Большинство традиционных методов обнаружения разметки основаны на двухэтапных подходах семантической сегментации [26, 27, 28]. На первом этапе в этих подходах используется нейронная сеть, которая предназначена для выполнения классификации на уровне пикселей, назначающая каждому пикселю в изображении двоичную метку: наличие полосы или ее отсутствие. При этом при классификации пикселей специально не учитываются зависимости между пикселями и их структурой. На втором этапе выполняется дополнительная постобработка данных, позволяющая определить такие характеристики, как уникальность или прямолинейность обнаруженной разметки. Постобработка может быть реализована с помощью модели условного случайного поля, дополнительных нейронных сетей или сложных методов CV, таких как RANSAC, но ее вычислительная сложность значительна,

и ее следует аккуратно комбинировать с первым этапом классификации путем ручной настройки. Эти подходы трудно масштабировать для различных сред и наборов данных. Другими методами обнаружения дорожной разметки являются подходы на основе генеративной состязательной нейронной сети (GAN) [23, 24], которые автоматически определяют указанные характеристики.

1. Постановка задачи сравнения

В данной работе будет рассматриваться сравнение двух подходов к решению задачи определения дорожно-транспортной разметки на изображениях.

Первый подход заключается в использовании метода, основанного на принципе геометризированных гистограмм. Приведем его краткое описание, подробное изложено ниже. Из исходного изображения строится граф цветовых ступок, на его основе определяются ростки глобальных объектов. По анализу геометрии ростков строятся глобальные объекты, классификация которых производится по форме и яркостно-цветовым характеристикам. Назовем данный метод прямым.

Другой метод основан на нейросетевом подходе, он является одним из общепринятых. Задача обнаружения полос рассматривается как задача выделения каждого отдельного интересующего объекта. Архитектура нейронной сети состоит из двух ветвей. Ветвь сегментации создает двоичную маску полосы, ветвь кластеризации полосы дополнительно разлагает выделенные пиксели полосы на разные объекты. После исключения фоновых пикселей с помощью двоичной карты; выделенные ветви группируются вместе и нумеруются. Чтобы параметризовать выделенную разметку, предлагается применить преобразование перспективы и затем аппроксимацию полосы движения.

Сравнительный анализ был проведен по открытому датасету TuSimple. Набор данных TuSimple состоит из 6408 изображений дорог на автомагистралях США. Разрешение изображения 1280×720 px. Набор данных состоит из 3626 изображений для обучения, 358 изображений для проверки и 2782 изображений для тестирования, называемых тестовым набором TuSimple, изображения которого относятся к разным погодным условиям.

2. Критерии сравнения

Для сравнения точности в качестве критерия оценки берется метрика, описанная в работе [17]. Точность рассчитывается как среднее отношения количества верно определенных вершин на изображении ко всем вершинам.

$$\text{Точность} = N_{\text{correct}}/N_{\text{gt}},$$

где $N_{correct}$ – число правильно спрогнозированных вершин, а N_{gt} – число вершин в оригинальной разметке полосы движения. Точка считается правильной, если разница между точной и предсказанной точкой меньше определенного порога. Кроме того, считаются ложноположительные (FP) и ложно отрицательные (FN) характеристики.

$$FP = F_{pred} / N_{pred},$$

$$FN = M_{pred} / N_{gt},$$

где F_{pred} – число ошибочно спрогнозированных полос, N_{pred} – число спрогнозированных полос, M_{pred} – число пропущенных полос и N_{gt} – число всех полос.

Для сравнения вычислительной сложности оценивается число операций с плавающей точкой в двух алгоритмах. Стоит отметить, что первый алгоритм выполняется только на центральном процессоре, а второй алгоритм выполняется на графическом процессоре.

3. Основные принципы метода геометризованных гистограмм

В отличие от других методов сегментации изображений, метод геометризованных гистограмм построен так, что основная обработка видеоданных будет вестись параллельным образом. В основе метода геометризованных гистограмм лежит идея математической редукции сложных задач. Эта идея предполагает введение новых понятий и логических систем, таких, что вложение в эти системы исходных задач и переформулирование их в новых терминах делает их разрешимыми с помощью специальной математической техники или дает эффективные методы их численного решения. Так как пиксельная сегментация является лишь промежуточным продуктом в построении систем понимания изображений и плохо годится для решения задач с обратной связью по зрению, то в методе геометризованных гистограмм она не применяется.

3.1. Конструкция геометризованных гистограмм. С помощью техники, описанной в [10–12], каждому цветному изображению ставится в соответствие структурный граф цветовых сгустков STG. Чтобы построить STG, изображение разбивается на полосы одинаковой ширины St_n со сторонами, параллельными горизонтальной или вертикальной оси плоскости изображения O_s . Вводится понятие геометризованной гистограммы изображения [10], которая служит аналогом обычной гистограммы. Вместе с тем обычная гистограмма является общим инвариантом изображения, так как она остается инвариантной при любом взаимно однозначном преобразовании прямоугольника изображения. В ней совершенно не учитывается геометрия объектов изображения. В то время

как геометризованная гистограмма остается инвариантной только для преобразований внутри полос, при которых точки передвигаются перпендикулярно оси O_s . Так как мы имеем дело с узкими полосами, эти преобразования почти не меняют геометрии объектов, принадлежащих изображению.

Геометризованная гистограмма достаточно точно описывает распределение значений функции, задающей изображение в прямоугольнике кадра. Получается геометризованная гистограмма с помощью проектирования пикселей полосы на основание полосы. Мы будем обозначать L_z подмножества полосы St_n , в которых функция $f(x,y)$, задающая изображение, принимает фиксированное значение z , и называть их множествами уровня z . Для монохромного изображения [10], ввиду дискретного характера изображений, проекция любого L_z на нижнее основание полосы есть объединение интервалов $Pr(L_z) = \cup_k I_{kz}$ на нижней оси полосы.

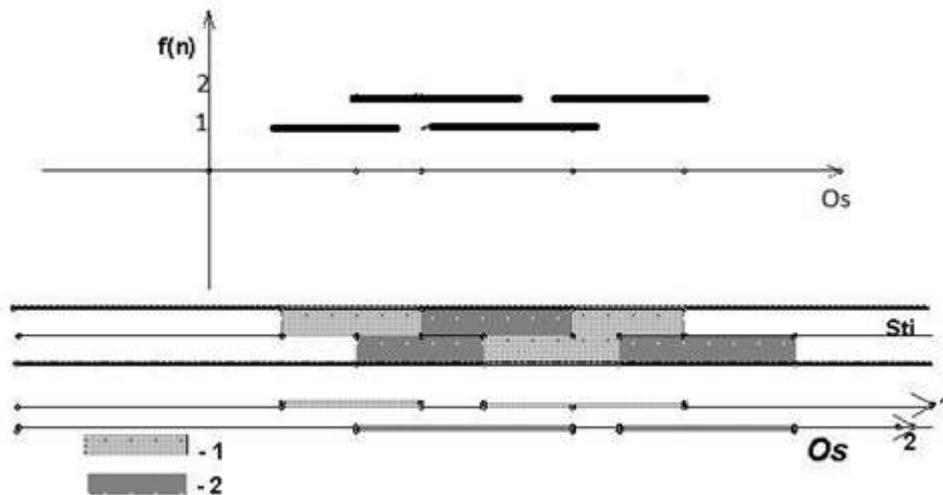


Рис. 1. Получение интервалов геометризованной гистограммы для модельного изображения

Заметим, что описание геометрии множеств L_z на плоскости изображения есть описание структуры распределения значений функции $f(x, y)$, задающей изображение. Описание этой геометрии на всей плоскости носит сложный характер. Использование описания распределения значений функции f в задачах сегментации выведет решение задачи о сегментации изображения на новый уровень по сравнению с изучением частотных закономерностей в области ее значений. Задача приближенного описания множеств L_z в узких полосах легко поддается решению. Это решение дается геометризованной гистограммой полосы. Чтобы сравнивать множество уровней в разных полосах, можно считать, что все интервалы лежат на оси O_s . Рисунок 1 иллюстрирует получение интервалов геометризованной гистограммы в полосе. В полосе

задана функция изображения с двумя значениями, которая постоянна в выделенных прямоугольниках. Под полосой расположены два экземпляра оси O_s . На них показаны системы интервалов, соответствующие областям одинакового значения функции изображения. На самом деле эти системы интервалов лежат на одной оси. Два экземпляра взяты для наглядности. При проектировании для каждого интервала S_g вычисляются его границы $begS_g$ и $endS_g$ на оси O_s , а также его мощность $CardS_g$ – число точек полосы, в которых функция изображения принимает заданное значение и которые проектируются на данный интервал. Объединение систем интервалов по всем полосам хорошо описывает распределение значений монохромной функции, задающей изображение. Эта конструкция обобщается на случай векторной функции, задающей цветное изображение [10]. При проектировании удастся разделить все множество точек полосы на подмножества, в которых насыщение, оттенок, и яркость варьируются в некоторых диапазонах. Эти подмножества проектируются на интервалы на оси O_s . В каждой полосе получаются системы интервалов, каждый из которых (S_g) характеризуется следующими параметрами:

- Положение интервала $[begS_g, endS_g]$ S_g на оси O_s ;
- Диапазон $\Delta HS_g = [H_{minS_g}, H_{maxS_g}]$ и среднее значение оттенка H_{meanS_g} ;
- Диапазон $\Delta SS_g = [S_{minS_g}, S_{maxS_g}]$ и среднее значение S_{meanS_g} насыщения;
- Диапазон $\Delta IS_g = [I_{minS_g}, I_{maxS_g}]$ и среднее значение яркости I_{meanS_g} .

Мощность интервала $CardS_g$ приблизительно равна числу точек полосы, лежащих в полосе над интервалом, с цветовыми характеристиками, заключенными в интервалах, указанными выше для S_g на оси O_s .

Обозначим $dens(S_g) = CardS_g / (endS_g - begS_g + 1)$ плотность интервала S_g . Естественно, если изображение в полосе состоит из однородно закрашенных цветовых пятен с фиксированными цветовыми характеристиками, то подобно рис. 1 каждому такому цветовому пятну при проектировании будет соответствовать интервал геометризованной гистограммы с теми же фиксированными цветовыми характеристиками. Однако изображения сцен реального мира с тенями, областями неоднородного освещения и влажными частями намного сложнее. Геометризованные гистограммы полос таких изображений содержат сложные системы пересекающихся интервалов с заданными диапазонами и средними значениями яркостно-цветовых характеристик. Для того, чтобы давать семантическое описание изображений и решать задачи их понимания на основе геометризованных гистограмм, требуется введение большого числа новых понятий и аксиоматизированных конструкций; также необходимо разработать эффективные программные средства для построения объектов, удовлетворяющих постулированным свойствам и их интерпретации.

Существенным свойством рассматриваемого в препринте алгоритма получения геометризованной гистограммы в полосе является то, что она может быть получена за один проход массива точек полосы. Это обеспечивает ее построение в реальном времени и является ключевым моментом параллельной реализации метода. Даже для монохромной функции изображения, построение геометризованной гистограммы за один проход изображения является занятной программистской задачей, которая может предлагаться на олимпиадах по программированию. Она реализуется кодом, объем которого более 300 строк. Код программы построения геометризованной гистограммы монохромной функции изображения выложен на [18], поэтому ввиду его размера код в препринте не приводится. Построение геометризованной гистограммы цветного изображения за один его проход является существенно более сложной программистской задачей. Код ее реализации содержит более 2000 строк. Существенное место в коде занимает сортировка пикселей согласно с их яркостно-цветовыми характеристиками. При сортировке мы используем цветовые координаты $(G/(G+B), G/(G+R), I)$, введенные первым автором в [10]. Здесь R, G, B – обычные цветовые координаты, а I – полутоновая интенсивность. Введем характеристическую функцию CF для данной цветовой сегмента. Заметим, что функции $G/(G+B), G/(G+R)$ задают цветовые координаты на цветовом треугольнике [10]. Если оттенок пикселя принадлежит желтой части цветового треугольника, тогда CF совпадает с $G/(G+B)$. Когда мы движемся в следующий цветовой диапазон (зеленый, голубой, красный), значение $G/(G+B)$ сдвигается на M , где M – число градаций функции $G/(G+B)$. Геометризованная гистограмма CF , дополненная для каждого ее интервала I_{kz} классической гистограммой другой цветовой координаты $G/(G+R)$ и диапазоном, и средним значением полутоновой компоненты спроектированных на интервал пикселей, называется геометризованной гистограммой цветного изображения в полосе St_i . На основе полученных данных для каждого интервала вычисляются диапазоны и средние значения координат H (оттенок), S (насыщение), I (полутоновая интенсивность). Далее характеристики интервалов применяются в этой цветовой системе координат.

Пример геометризованной гистограммы выделенной полосы приведен на рис. 2. Точка в правой части изображения, расположенного в верхней части рисунка, выделяет горизонтальную его полосу. В нижней части рисунка показаны все интервалы геометризованной гистограммы этой полосы. По горизонтальной оси откладываются значения характеристической функции геометризованной гистограммы CF . По вертикальной оси откладываются координаты концов интервалов (измеренные по горизонтальной оси изображения). Овалами (при движении слева направо) выделены интервалы геометризованной гистограммы, соответствующие: 1. частям белого

автомобиля между габаритами и зонами сигналов торможения; 2. частям зоны стоп-сигналов на белом автомобиле, которые имеют красный цвет; и 3. частям соответствующих зон растительности на обочинах (верхний и нижний овалы). Квадраты в верхней части визуализации геометризованной гистограммы показывают цвет того из интервалов, определяемых стоп-сигналами, к которому подведена стрелка мыши.

Так как мы имеем дело с реальными объектами, то им соответствуют наборы интервалов геометризованной гистограммы, яркостно-цветовые характеристики которых варьируются. На сайте [18] выложена программа с инструкциями, которая позволяет построить и визуализировать, как на рис. 2, геометризованную гистограмму любого изображения в формате BMP.

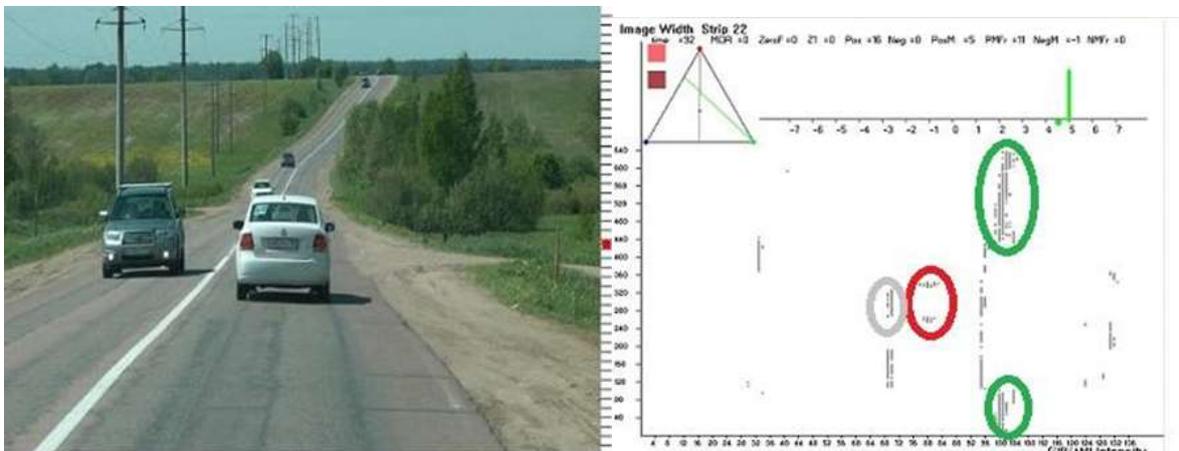


Рис. 2. Геометризованная гистограмма выделенной полосы

Программная реализация построения геометризованных гистограмм цветного изображения (кадра видеопоследовательности) и его полутоновой компоненты находится в Блоке 1 итоговой блок-схемы программной системы из раздела 3. Для реальных изображений применяется правило, что реальным объектам соответствуют несколько геометрически близких интервалов геометризованной гистограммы, у которых яркостно-цветовые характеристики имеют близкие значения. Объединение геометризованных гистограмм полос дает геометризованную гистограмму всего изображения.

3.2. Построение цветовых сгустков и графа STG. Для того чтобы выделить на геометризованной гистограмме полосы части реальных объектов, в ней необходимо применить некоторые процедуры для объединения ее интервалов, им соответствующим. С помощью оригинальной операции кластеризации [10] интервалы геометризованной гистограммы объединяются в кластеры – цветовые сгустки, которые характеризуются такими же, как и у интервалов, яркостно-цветовыми параметрами, мощностью и плотностью. На

этом шаге проводится объединение интервалов геометризованной гистограммы, интервалы локализации которых имеют сильное пересечение [10].

Программная реализация этой процедуры разделяется на следующие шаги:

1. Все интервалы геометризованной гистограммы отображаются на три отрезка прямой, соответствующей средней линии полосы. На первом отрезке остаются интервалы максимальной плотности, на втором – интервалы наибольшего локального насыщения, на третьем – интервалы максимального значения полутоновой компоненты;
2. Интервалы с номерами, “выжившими” на всех трех отрезках средней линии, являются зародышами, с которых стартует процедура кластеризации.
3. К каждому зародышу присоединяются интервалы гистограммы, которые имеют с ними сильное пересечение [10] и близкие яркостно-цветовые характеристики.

Далее работает процедура объединения цветовых сгустков. При этом объединяются цветовые сгустки, имеющие соседние интервалы локализации и сходные яркостно-цветовые параметры. Итоговые цветовые сгустки b (color bunches в английском варианте термина) характеризуются интегральными параметрами ΔN_b , N_{meanb} , ΔS_b , S_{meanb} , ΔI_b , I_{meanb} , $Cardb$, получаемыми суммированием и усреднением соответствующих интервалов геометризованной гистограммы, из которых они состоят. Введем так же, как и для интервалов геометризованной гистограммы, плотность цветового сгустка как $dens(b) = Cardb / L([begb, endb])$ (мощность, поделенная на длину интервала). Процедуры построения цветовых сгустков и их объединения принадлежат Блоку 2 программы.

Цветовые сгустки объединяются в граф. В полосе соединяются ребром соседние цветовые сгустки (с соседними интервалами локализации), а в соседних полосах – цветовые сгустки, интервалы локализации которых пересекаются. Неформально каждый сгусток дает описание некоторой части реального объекта в полосе, его проекцию на ось O_s и описание значений численных цветовых характеристик этой части объекта. STG можно интерпретировать геометрически с помощью наложения отрезков его сгустков ($[begb, endb]$) на центральную линию соответствующей полосы и окрашивания этих отрезков в цвет, определяемый N_{meanb} , S_{meanb} , I_{meanb} . Пример множества цветовых сгустков изображения на рис. 2, наложенных на его полутоновую компоненту, можно посмотреть на рис. 3. Примеры цветовых сгустков, наложенных на цветные изображения, можно также посмотреть в открытом доступе в [13, 14].

Цветовые сгустки являются аналогом суперпикселов, применяемых в классических методах сегментации. Также с помощью программы из [18] можно построить и визуализировать множество цветовых сгустков для любого

изображения в формате BMP. Приведенные изображения цветowych сгустков и получаемые с помощью программы из [18] показывают, что граф цветowych сгустков хорошо описывает яркостно-цветовые свойства изображений и геометрию реальных объектов в полосах и в изображении в целом.



Рис. 3. Цветовые сгустки изображения из рис.2. Черточками выделены границы зеленой обочины, полученные с помощью программной системы

3.3. Построение поисковой решетки на STG. На множестве цветowych сгустков строится “решетка поиска” $\text{SearchLat}(\text{STG})$ [12], которая позволяет производить глобальный анализ изображения. Если мы положим на среднюю линию каждой полосы разбиения изображения интервалы геометрической локализации $[\text{begb}, \text{endb}]$ всех цветowych сгустков полосы, то получим некоторое ее покрытие.

Определение 1.1. Цветовые сгустки, имеющие в некоторой точке средней линии максимальную плотность, называются доминирующими цветowymi сгустками. Остальные сгустки называются доминируемыми.

Ясно, что доминирующие цветовые сгустки образуют покрытие средней линии. Оказывается [12], что всегда можно выбрать линейно упорядоченную последовательность базисных цветowych сгустков, которые образуют покрытие средней линии.

Алгоритм для построения $\text{SearchLat}(\text{STG})$ состоит из следующих шагов:

1. Для каждого доминирующего сгустка находятся доминирующие сгустки слева и справа, интервалы локализации являются соседними для интервала локализации выбранного сгустка.
2. Начиная с левого конца строится упорядоченная система доминирующих цветковых сгустков, соседних друг другу.

Доминирующие цветковые сгустки, включенные в линейно упорядоченное покрытие, называются базисными цветковыми сгустками. Базисные цветковые сгустки всех полос образуют решетку поиска изображения SearchLat(STG) [12].

В дополнение к поисковой решетке строится структура BunchLoc(STG). В каждой полосе для любой точки x оси O_s данная структура указывает номера всех цветковых сгустков, которые проходят через данную точку. Программа, реализующая построение структуры BunchLoc(STG), реализуется в два шага:

1. Перебираются все цветковые сгустки, и для каждой точки $x \in O_s$ определяется число сгустков, проходящих через данную точку. Вычисляется N_b — максимальное число сгустков, которые могут проходить через произвольную точку оси O_s .
2. Строится массив, высота которого равна N_b и который содержит для каждой точки $x \in O_s$ список цветковых сгустков, проходящих через эту точку.

Структура BunchLoc(STG) позволяет для каждого цветкового сгустка b определить, какие цветковые сгустки в соседних полосах связаны с данным сгустком ребром в STG (имеют интервалы локализации, пересекающиеся с интервалом локализации b). Структуры SearchLat и BunchLoc позволяют находить все соседние цветковые сгустки для выбранного сгустка b . После выполнения операций сегментации цветковым сгусткам присваиваются метки объектов, которым они принадлежат. Эти структуры позволяют строить отношения соседства для выделенных на изображении объектов и выполнять глобальный анализ их взаимодействия [12]. Их конструирование выполняется в Блоке 3.

3.4. Построение глобальных объектов. Чтобы решать задачи поиска объектов с заданными геометрией и яркостно-цветовыми характеристиками на изображениях и задачи понимания изображений, необходимо строить образы частей предполагаемых реальных глобальных объектов в STG. Под глобальными объектами мы понимаем объекты, расположенные более чем в одной полосе изображения. Реальные глобальные объекты могут быть двух типов. К первому типу относятся областные объекты. Они в полосах, как правило, при проектировании на STG индуцируют доминирующие цветковые сгустки. Глобальные объекты второго типа – это малые и тонкие объекты, такие как дорожная разметка, строительные конусы для обозначения запретных зон на дороге, знаки аварийной остановки, дорожные знаки, стоп-сигналы автомобилей и т.д. Объекты такого типа могут при проектировании давать как

доминирующие (на близком расстоянии), так и доминируемые цветовые сгустки. Целый объект или его часть, лежащие в нескольких полосах, дают в некотором смысле непрерывные системы цветовых сгустков в STG. Для нахождения образов таких объектов, необходимо формализовать понятие непрерывной системы цветовых сгустков. Многие задачи по поиску ориентиров и объектов в кадре можно переформулировать строго как задачи поиска некоторых непрерывных абстрактных объектов на графе цветовых сгустков.

Для этих целей в [11] введены понятия левых и правых контрастных кривых или левых и правых ростков глобальных объектов на STG и определен двудольный граф левых и правых контрастных кривых LRG. Если изображение разбито на горизонтальные полосы, то неформально левая или правая контрастная кривая (росток глобального объекта) есть цепочка доминирующих цветовых сгустков в соседних полосах с подобными цветовыми характеристиками. При этом левые или правые концы цветовых сгустков меняются от полосы к полосе “непрерывно”, и соседние в той же полосе слева или справа цветовые сгустки имеют контрастные цветовые характеристики. Эта цепочка строится снизу вверх, переходя из полосы в полосу [11]. Для того чтобы избежать операций полного перебора, непрерывное продолжение цветового сгустка в соседние полосы (построение левых и правых ростков глобальных объектов) осуществляется с использованием структуры BunchLoc. Конструкция левых и правых непрерывных цепочек контрастных доминирующих цветовых сгустков (левых и правых ростков глобальных контрастных объектов) выполняется в Блоке 4 программной системы. Также в этом блоке строится двудольный граф связи между левыми и правыми ростками глобальных объектов. Двудольный граф LRG показывает, какие левые и правые ростки в полосах можно соединить цепочками цветовых сгустков без контраста. Следовательно, эти ростки могут быть частями одного объекта. Для решения различных задач понимания изображений необходимо находить прямолинейные участки в граничных элементах построенных левых и правых ростков. Методы, основанные на методе наименьших квадратов, являются неустойчивыми относительно ошибок сегментации. Устойчивый метод, основанный на анализе гистограмм углов наклона отрезков, соединяющих концы интервалов локализации сгустков, входящих в построенный росток глобального объекта, разработан в [12]. Часть программы, реализующая предложенный метод, находится в Блоке 5 диаграммы. Этот блок заканчивает часть программного комплекса, который дает аппарат для решения различных задач понимания изображений. Общий объем кода больше 40000 строк. Предполагается выложить части кода в интернет. Также будет выложена в [18] dll, которая реализует все перечисленные модули для общего анализа изображений. Программный комплекс, который позволяет строить ростки

левых и правых глобальных объектов для любого цветного изображения, выложен на сайт [18]. Пример левого ростка глобального множества для изображения рис. 2 приведен на рис. 4.

Подобные определения вводятся для поиска малых и тонких объектов. Единственная разница состоит в том, что в цепочках цветных сгустков участвуют контрастные доминируемые сгустки. На этом языке могут быть переформулированы многие классические задачи понимания изображений.

Например, так можно сформулировать задачу детектирования и понимания дорожной разметки [14]. При этом удастся отыскивать как белую, так и цветную разметку, производить нахождение временной цветной разметки при наличии старой белой и анализировать криволинейную разметку так же хорошо, как и прямолинейную.

Поисковая решетка SearchLat активно используется при сегментации и решении задачи понимания изображений. С помощью специальных массивов описывается, какие левые и правые ростки проходят через любой цветовой сгусток. Также отмечается тот факт, что таких ростков не имеется. С помощью SearchLat легко строятся отношения соседства для ростков глобальных множеств [12], что при обычных подходах представляет серьезную проблему.



Рис. 4. Непрерывная слева система цветных сгустков, соответствующая части зеленой обочины.

Таким образом, готовится процедура объединения ростков. Также с помощью SearchLat строится система рассуждений, которая позволяет выделять кандидатов на часть дорожной разметки в полосе. Так как кандидаты на части дороги всегда принадлежат SearchLat, то для любого цветового сгустка достаточно проанализировать его связи в яркостно-цветовой и в геометрической областях с соседними базисными сгустками SearchLat, чтобы отнести его к кандидатам на часть разметки в полосе.

Поисковая решетка SearchLat и структура BunchLoc также позволяют выделять мелкие контрастные объекты на изображениях. Используя структуру BunchLoc, для каждого доминируемого цветового сгустка b можно определить все базисные цветовые сгустки из SearchLat, интервалы локализации которых пересекают интервал локализации b .



Рис. 5. Блок-схема этапов работы (блоков программного комплекса) метода геометризованных гистограмм, серым фоном выделены параллельно исполняемые части метода

Определение 1.2. Доминируемый цветовой сгусток b является контрастным цветовым сгустком в STG, если он имеет контрастные цветовые характеристики со всеми базисными сгустками из SearchLat, с которыми он имеет пересекающиеся интервалы локализации.

Замечание. Для доминирующих сгустков определение контрастности было дано в [11]. Оно утверждает, что доминирующий сгусток контрастный, если его яркостно-цветовые характеристики контрастны к яркостно-цветовым характеристикам соседних доминирующих сгустков.

Используя это определение, с помощью формальных процедур, на STG находятся образы контрастных малых объектов из изображения, таких как фары автомобилей, сигнальные огни летательных объектов, сигналы светофоров, стоп-сигналы и сигналы поворотов автомобилей, и т.д. Для подтверждения истинности найденных малых объектов используются данные глобальной сегментации. Эти данные подтверждают, что малый контрастный объект находится рядом с большим контрастным объектом (сигнальные огни

вертолета) или внутри него (стоп-сигналы автомобиля). Взаимодействие локального и глобального анализа обеспечивает надежность идентификации.

Для большей наглядности на рис 5 приведена схема работы алгоритма, а также выделены те его части, которые теперь выполняются в многопоточном режиме. Существенной особенностью метода является возможность работы без информации о внутренних и внешних параметрах камеры, снимающей сцену. Данные о таких параметрах, как и информация о подстилающей поверхности существенны при отбрасывании фальшивых объектов. Недостаток этих данных компенсируется решением одновременно нескольких задач (поиск неба, областей дороги, обочины, дорожной разметки, других объектов на дороге) и сравнительного анализа полученных результатов. Для этого строятся системы рассуждения, которые делают выводы о надежности отдельных элементов, и на этой основе производится отбрасывание ложных объектов. Пока возможности этой системы реализованы приблизительно на 30%. Ее развитие позволит существенно улучшить результаты. Также реализуется одновременная обработка разбиений изображений на горизонтальные и вертикальные полосы. Это позволит более точно оценивать параметры формы автомобилей на дороге и точнее находить параметры других важных объектов, таких как мосты, переходы через дороги. Также имеются возможности определения цветов автомобилей, анализировать состояние сигнальных зон автомобилей (торможение, поворот), находить светофоры и определять их состояние на большом удалении от них.

4. Описание нейросетевого метода

4.1. Нейронная сеть LaneNet. В основе метода используется нейронная сеть LaneNet (см. рис. 6), которая сочетает в себе преимущества бинарного распознавания с функцией кластеризации потерь, разработанной для сегментации одного объекта. Подробно описание данного алгоритма приведено в работе [18]. В выходных данных LaneNet каждому пикселю дорожной полосы назначается идентификатор соответствующей полосы, при этом каждая дорожная полоса на кадре формирует свой собственный объект.

Задача определения объекта состоит из двух частей: сегментации и кластеризации. Для повышения производительности, с точки зрения как скорости, так и точности, две части совместно обучаются в многозадачной сети (см. рис. 6). Ветвь сегментации LaneNet (см. рис. 6, нижняя ветвь) обучена выводить карту двоичной сегментации, указывающую, какие пиксели принадлежат дорожной полосе, а какие нет. На данном этапе на кадре определяются полосы движения, несмотря на возможное отсутствие сегментов полосы на кадре, что обусловлено такими факторами, как загоразивание

автомобилем, выцветание полосы и пр. Сеть сегментации обучается с помощью стандартной функции потерь кросс-энтропии. Поскольку два класса (полоса/фон) сильно неуравновешены, применяется ограниченное обратное взвешивание классов (bounded inverse class weighting).

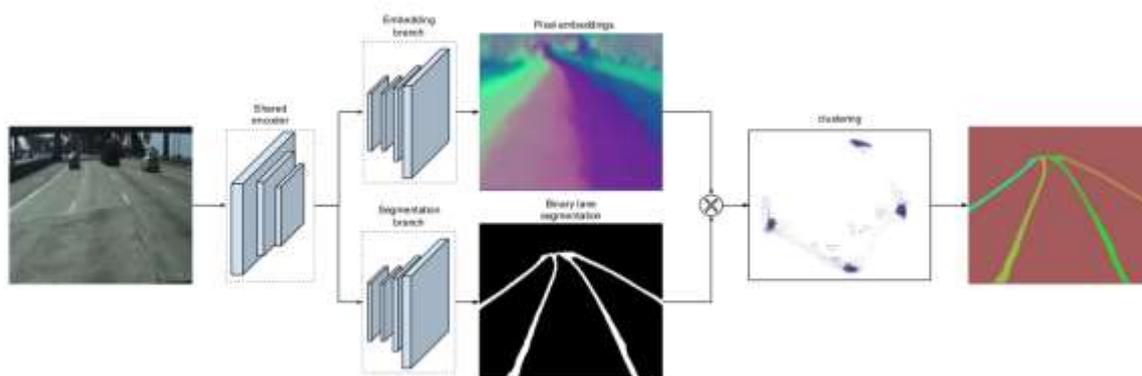


Рис. 6. Архитектура нейронной сети LaneNet.

Чтобы определить пиксели дорожной полосы, идентифицированные ветвью сегментации, обучается вторая ветвь LaneNet для кластеризации найденной области. (см. рис. 6, верхняя ветвь). В работе используется метод, основанный на изучении расстояний от данных, предложенный В. De Brabandere и др. [19], который можно легко встраивать в стандартные сети с прямой связью и который специально разработан для приложений реального времени. Используя функцию потерь кластеризации рис. 7, ветвь кластеризации объектов обучается выводить **положение** для каждого пикселя полосы так, чтобы расстояние между положением пикселей, принадлежащих одной полосе, было небольшим, в то время как расстояние между положением пикселей, принадлежащих разным полосам, было максимальным.

Таким образом, пиксели одной и той же полосы будут сгруппированы вместе, образуя уникальные кластеры для каждой полосы. Это достигается за счет введения двух терминов: члена дисперсии (L_{var}), который применяет тянущее усилие к каждому положению по направлению к среднему положению полосы движения, и члена расстояния (L_{dist}), который отталкивает центры кластеров друг от друга. Оба члена являются переменными: сила сближения активна только тогда, когда его положение находится дальше, чем δv от его центра кластера, а сила отталкивания между центрами активна только тогда, когда они ближе, чем δd друг к другу.

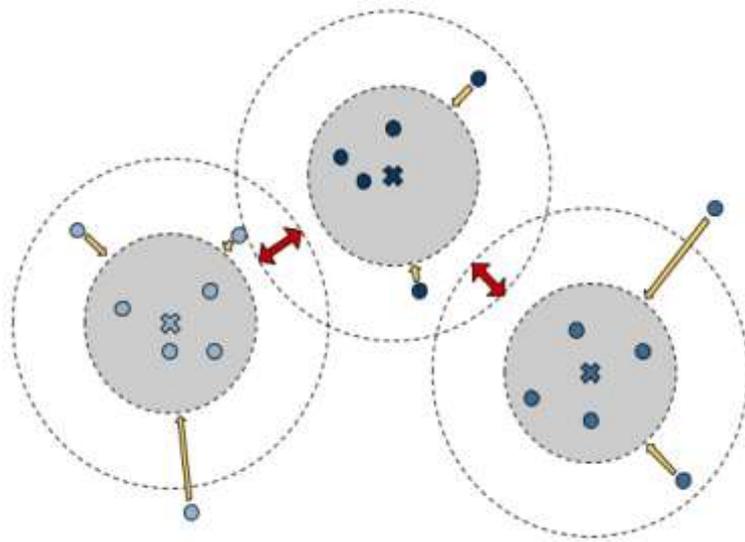


Рис 7. Иллюстрация функция контрастных потерь или потерь триплетов.

Как только сеть сойдется, положения пикселей дорожек будут сгруппированы вместе (см. рис. 6), так что каждый кластер будет располагаться дальше, чем δd друг от друга, а радиус каждого кластера будет меньше, чем δv . Кластеризация выполняется итеративной процедурой. Установив $\delta d > 6\delta v$ в приведенных выше потерях, можно выбрать случайное положение полосы и порог вокруг нее с радиусом $2\delta v$, чтобы выбрать все положения, принадлежащие той же полосе. Это повторяется до тех пор, пока все положения пикселей полос не будут присвоены какой-либо полосе. Чтобы избежать выбора выброса для определения порогового значения, сначала используется средний сдвиг для смещения ближе к центру кластера, а затем выполняется определение порога (см. рис. 6).

Сетевая архитектура LaneNet основана на сети кодировщика-декодера ENet [20], которая преобразована в сеть с двумя ветвями. Поскольку кодировщик ENet содержит больше параметров, чем декодер, полное совместное использование кодировщика между двумя задачами приведет к неудовлетворительным результатам. Таким образом, в то время как исходный кодировщик ENet состоит из трех этапов (этапы 1,2,3), LaneNet использует только первые два этапа (1 и 2) между двумя ветвями, оставляя этап 3 энкодера ENet и полный декодер ENet как основу каждой отдельной ветви. Последний слой ветви сегментации выводит одноканальное изображение (двоичная сегментация), тогда как последний слой ветви кластеризации выводит N-мерное изображение с N размерной композицией. Схематично изображено на рис. 6.

4.2. Подготовка кривой с использованием H-Net. Как было описано в предыдущем разделе, вывод LaneNet представляет собой набор пикселей на полосу. Аппроксимация пикселей происходит с помощью полинома третьей степени и матрицы преобразования перспективы H-Net, которая вычисляется из

решения задачи оптимизации методом наименьших квадратов для построения многочлена по выходным точкам. Сетевая архитектура H-Net намеренно сделана небольшой и построена из последовательных блоков сверток 3x3, пакетной обработки и ReLU. Размер уменьшается с использованием максимального количества слоев объединения, и в конце добавляются 2 полностью связанных слоя.

5. Сравнительный анализ работы методов

5.1. Описание эксперимента. Был осуществлен эксперимент по сравнению двух подходов в решении задачи определения дорожной разметки. Задачи решались на персональном компьютере (ПК) среднего сегмента. Основные характеристики Core i7-8700, DDR4 32gb, GPU GTX 1080Ti. Сравнительный анализ был проведен по открытому датасету TuSimple. Набор данных состоял из 2782 изображений для тестирования. Тестовые изображения относятся к разным погодным условиям.

Проведен сравнительный анализ работы нейросетевого метода и метода геометризованных гистограмм. Справедливо сказать, что сравнение по быстродействию или вычислительной сложности не является корректным из-за разницы архитектур представленных алгоритмов. Метод геометризованных гистограмм выполняется параллельно на всех ядрах CPU, а нейросетевой метод выполняется на GPU. Стоит отметить, что нейросетевое решение предполагает возможную реализацию для промышленного использования на программируемой логической интегральной схеме (ПЛИС), что приведет к существенному уменьшению времени работы. Несмотря на это, в работе приведены оценки по быстродействию на ПК, так как прямой метод не предполагает реализации на ПЛИС, но и не исключает ее.

Точность рассчитывается как среднее отношения количества верно определенных пикселей на изображении ко всем размеченным пикселям.

$$\text{Точность} = N_{\text{correct}}/N_{\text{gt}},$$

где N_{correct} - число правильно спрогнозированных пикселей, а N_{gt} - число пикселей в оригинальной разметке полосы движения. Точка считается правильной, если разница между точкой и предсказанной точкой меньше определенного порога. Кроме того, считаются ложноположительные (FP) и ложноотрицательные (FN) характеристики.

5.2. Результаты сравнения В табл. 1,2 можно увидеть, что прямой метод основанный на подходе геометризованных гистограмм, работает почти в два раза быстрее, но его точность требует доработки. Однако следует отметить, что построение кривой прямым методом уже содержит в себе классификатор

найденных кривых и выделяет одну или две сплошные полосы, прерывистую полосу и разметку желтого цвета.



Рис. 8. На рисунке два блока изображений. Первый блок (два столбца по три изображения в каждом) - работа нейронной сети. В верхнем ряду показаны точки оригинальной разметки, в среднем ряду - выход LaneNet, в нижнем ряду - прогноз нейронной сети после аппроксимации полиномом. Второй блок изображений - работа прямого метода с распознанной разметкой прерывистой линии.

Таблица 1

Быстродействие алгоритмов

	time (ms)	fps	Разрешение (px)	Вычислительная сложность
LaneNet + H-Net	19	52.6	512x256	201.4 TFlops
GeomGistogram	10,2	98	640x480	3.876 TFlops

Таблица 2

Точность работы алгоритмов

	Метрика	Значение	FP	FN
LaneNet + H-Net	Accuracy	96.4%	0.0780	0.0244
GeomGistogram	Accuracy	89.20%	0.2170	0.1785

Заключение

Планируется усовершенствовать работу прямого метода для повышения точности сегментации изображения. На представленных кадрах видно, что алгоритм, основанный на геометризированных гистограммах не всегда определяет все типы разметки, что обусловлено необходимостью индивидуальной настройки кластеризации объектов, поэтому далее планируется рассматривать больше кандидатов на разметку. Пока сравнительный анализ кандидатов на разметку проводится не для всех типов разметки. Доработка систем понимания разметки позволит существенно улучшить результаты. Также в системе заложена возможность детектирования радикальной смены типов разметки (появление временной разметки, меняющей направление движения). В качестве одной из перспективных доработок может быть использование комбинации подходов таким образом, чтобы кластеризация происходила на графе цветовых сгустков. Также будет рассмотрен подход, при котором часть блока обработки данных осуществляется с помощью нейронной сети. При указанных доработках прогнозируется повышение точности прямого метода до уровня точности методов, основанных исключительно на нейронных сетях.

Список литературы

1. *Felzenszwalb P. F., Huttenlocher D.P.* Efficient graph-based image segmentation // *International Journal of Computer Vision*. 2004. V. 59. № 2. P. 167–181.
2. *Shotton J. Winn J. Rother C. Criminisi C.A.* TextonBoost for image understanding multi-class objects recognition and segmentation by jointly modeling texture, layout, and context// *International Journal of Computer Vision*. 2004. V. 81. № 1. P. 2–23.
3. *Badrinarayanan V. Kendall A. Cipola R.* SegNet: a deep convolutional encoder-decoder architecture for image segmentation // *IEEE Trans. PAMI*. 2017. V. 39. № 12. P. 2481–2495.
4. *Tian Y., Gelernter J. and Wang X.* Lane marking detection via deep convolutional neural network // *Neurocomputing* 2018. V. 280 P. 46–55
5. Cao, J.; Song, C.; Song, S.; Xiao, F.; Peng, S. Lane Detection Algorithm for Intelligent Vehicles in Complex Road Conditions and Dynamic Environments. *Sensors* 2019, 19, 3166. <https://doi.org/10.3390/s19143166>
6. *Norote S.P., Bhujbal P.N., Norote A.S. and Dhane D.M.* A review of recent advances in lane detection and departure warning system // *Pattern Recognition*. 2018. V. 73 № 2. P. 216–234.
7. *Son J., Yoo H., Kim S., Sohn K.* Real-time invariant lane detection for lane departure warning system *Expert Systems with Applications*. 2015. V. 42 P. 1816–24.
8. *Rosenband D.* Inside waymo’s self-driving car: My favorite transistors. In 2017 Symposium on VLSI Circuits, pages C20–C22. IEEE, 2017.
9. <https://www.cnbc.com/2020/01/28/ubers-self-driving-carsare-a-key-to-its-path-to-profitability.html>.
10. *Kiy K. I.* A new real-time method for description and generalized segmentation of color images // *Pattern Recognition and Image Analysis*. 2010. V. 20. № 2. P. 398–401.
11. *Kiy K.I.* Segmentation and detection of contrast objects and their application in robot navigation // *Pattern Recognition and Image Analysis*. 2015. V. 25. № 2. P. 338–346.
12. *Kiy K.I.* A new method of global image analysis and its application in understanding road scenes // *Pattern Recognition and Image Analysis*. 2018. V. 28. № 3. P. 483–494.

13. *Kiy K.I.* An image understanding system based on the geometrized histograms method: finding the sky in road scenes // CEUR Workshop Proceedings. V. 2210. P. 291-299. <http://ceur-ws.org/Vol-2210/paper38.pdf> .
14. *Dosaev R.V., Kiy K.I.* A new real-time method for finding temporary and permanent road marking // CEUR Workshop Proceedings. V. 2391. P. 86-96. <http://ceur-ws.org/Vol-2391/paper12.pdf/>
15. *Подопросветов А.В., Павловский В.Е., Кий К.И., Анохин Д.А.* Робомобиль Нива: алгоритмы управления. // Сборник трудов ПУМСС XXI Международной конференции. 2019. т. II, с. 102-106.
16. *Neven D, De Brabandere B., Georgoulis S., Proesmans M., Gool L.* Towards end-to-end lane detection: an instance segmentation approach. // In 2018 IEEE intelligent vehicles symposium (IV), pages 286–291. IEEE, 2018
17. TuSimple. <https://github.com/TuSimple/tusimple-benchmark/issues/3>
18. *Neven D., Brabandere B., Georgoulis S., Proesmans M., Van Gool. I.*, Towards End-to-End Lane Detection: an Instance Segmentation Approach // 2018, arXiv:1802.05591 {HYPERLINK"<https://github.com/MaybeShewill-CV/lanenet-lane-detection>" \h}
19. *De Brabandere B., Neven D., Van Gool L.* CoRR abs/1708.02551, 2017.
20. *Paszke A., Chaurasia A., Kim S., Culurciello E.*, ENet: A deep neural network architecture for real-time semantic segmentation. // CoRR abs/1606.02147, 2016.
21. <https://www.cnbc.com/2019/11/05/lyft-is-developing-selfdriving-cars-at-its-level-5-lab-in-palo-alto.html>.
22. *Yoffie D.* Mobileye: The future of driverless cars. // 2014.
23. *Sharon L., Poczter and Luka M. Jankovic.* The google car: driving toward a better future // Journal of Business Case Studies (JBCS), 10(1):7–14, 2014.
24. *Dikmen M., Burns C.*, Autonomous driving in the real world: Experiences with tesla autopilot and summon. In Proceedings AutomotiveUI '16, p. 225–228, 2016.
25. *Ghafoorian M. Nugteren C., Baka N., Booij O., Hofmann M.* EL-GAN: Embedding loss driven generative adversarial networks for lane detection. // In ECCV Workshop, 2019.
26. *Liu T., Chen Z., Yang Y., Wu Z., Li H.* Lane detection in low-light conditions using an efficient data enhancement : Light conditions style transfer. arXiv:2002.01177, 2020.
27. *Luc P., Couprie C., Chintala S., Verbeek J.* Semantic segmentation using adversarial networks. arXiv preprint arXiv:1611.08408, 2016.