

**МИНИСТЕРСТВО ОБРАЗОВАНИЯ РЕСПУБЛИКИ БЕЛАРУСЬ**  
**БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ**  
**ФАКУЛЬТЕТ ПРИКЛАДНОЙ МАТЕМАТИКИ И ИНФОРМАТИКИ**  
**Кафедра компьютерных технологий и систем**

**МЕТОДЫ И АЛГОРИТМЫ СЕМАНТИЧЕСКОЙ СЕГМЕНТАЦИИ  
ИЗОБРАЖЕНИЙ**

Курсовой проект

Ничипорука Романа Олеговича  
студента 3-го курса  
специальности «Информатика»  
Научный руководитель:  
Старший преподаватель кафедры  
компьютерных технологий и сетей  
С. В. Шолтанюк

Минск, 2023

# ОГЛАВЛЕНИЕ

<b>ВВЕДЕНИЕ</b>	<b>4</b>
<b>1 Общие понятия. Обзор основных методов и алгоритмов семантической сегментации</b>	<b>5</b>
1.1 Понятие сегментации изображений . . . . .	5
1.2 Семантическая сегментация. . . . .	6
1.3 Методы и алгоритмы семантической сегментации изображений .	8
1.4 Граничные алгоритмы . . . . .	8
1.5 Пороговые алгоритмы . . . . .	11
1.5.1 Метод с глобальным порогом . . . . .	12
1.5.2 Метод с адаптивным порогом . . . . .	14
1.6 Алгоритмы наращивания областей . . . . .	17
1.7 Алгоритмы сегментации на основе кластеризации . . . . .	18
1.8 Нейронные сети . . . . .	19
<b>2 Методы оценки эффективности сегментации</b>	<b>23</b>
2.1 Методы оценки эффективности сегментации . . . . .	23
2.1.1 Устойчивость . . . . .	23
2.1.2 Дополнительные ограничения . . . . .	24
2.1.3 Оптимальные условия применения метода . . . . .	25
2.1.4 Точность средней области интереса . . . . .	25
2.1.5 Достоинства . . . . .	26
2.1.6 Ограничения . . . . .	26
<b>3 Анализ избранных методов и алгоритмов</b>	<b>27</b>
3.1 Анализ алгоритмов сегментации . . . . .	27
3.1.1 Анализ граничных алгоритмов . . . . .	27
3.1.2 Анализ пороговых алгоритмов . . . . .	27
3.1.3 Анализ алгоритмов наращивания областей . . . . .	28
3.1.4 Анализ алгоритмов кластеризации . . . . .	28
3.1.5 Анализ нейронных сетей в задаче сегментации изображений	29

<b>ЗАКЛЮЧЕНИЕ</b>	<b>31</b>
<b>СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ</b>	<b>32</b>

# ВВЕДЕНИЕ

Компьютерное зрение получило большое развитие в последние годы благодаря достижениям в области искусственного интеллекта и инновациям в области глубокого обучения нейронных сетей. Сегментация при этом, является важной задачей данной сферы, решения которой люди стараются найти разными путями, включая нейронные сети.

Семантическая сегментация изображений представляет собой задачу разделения изображения на группы пикселей, соответствующих различным объектам, а также их классификации. Она отличается от задач позиционирования и обнаружения объектов тем, что требует более детального определения границ и структуры объектов на изображении. Вместо простого определения наличия объекта на изображении, семантическая сегментация ставит целью классифицировать каждый пиксель и точно определить его принадлежность к определенному классу объектов. Это сложная задача, которая требует учета контекста и семантической информации, чтобы достичь точных и качественных результатов в выделении и классификации объектов на изображении.

В настоящее время существует ряд методов и алгоритмов для сегментации изображений, различающихся по точности и производительности, начиная с поиска классических признаков изображения, заканчивая моделями глубокого обучения на которых основано большое количество современных алгоритмов компьютерного зрения.

Цель курсового проекта состоит в том, чтобы осуществить обзор и сравнение различных методов и алгоритмов семантической сегментации.

Задачи курсового проекта:

- описать различные подходы к семантической сегментации;
- проанализировать принцип работы классических алгоритмов и методов;
- проанализировать принцип работы алгоритмов на основе нейронных сетей;
- определить оптимальные условия и задачи для выбранных алгоритмов и методов семантической сегментации.

# ГЛАВА 1

## Общие понятия. Обзор основных методов и алгоритмов семантической сегментации

### 1.1 Понятие сегментации изображений

Сегментация изображения – это метод, при котором цифровое изображение разбивается на различные подгруппы, называемые сегментами изображения, что помогает уменьшить сложность изображения и упростить дальнейшую обработку или анализ изображения. Сегментация – это присвоение меток пикселям. Все элементы изображения или пиксели, принадлежащие к одной и той же категории, имеют общую метку, назначенную им.

Сегментация изображений является расширением задачи классификации изображений, которая также включает в себя задачу локализации объектов. Таким образом, сегментация изображения представляет собой надмножество классификации изображений, при этом она позволяет модели не только классифицировать объекты, но и точно определить их границы, создавая сегменты, охватывающие каждый объект на изображении..

На рис. 1.1 представлены различные задачи компьютерного зрения.

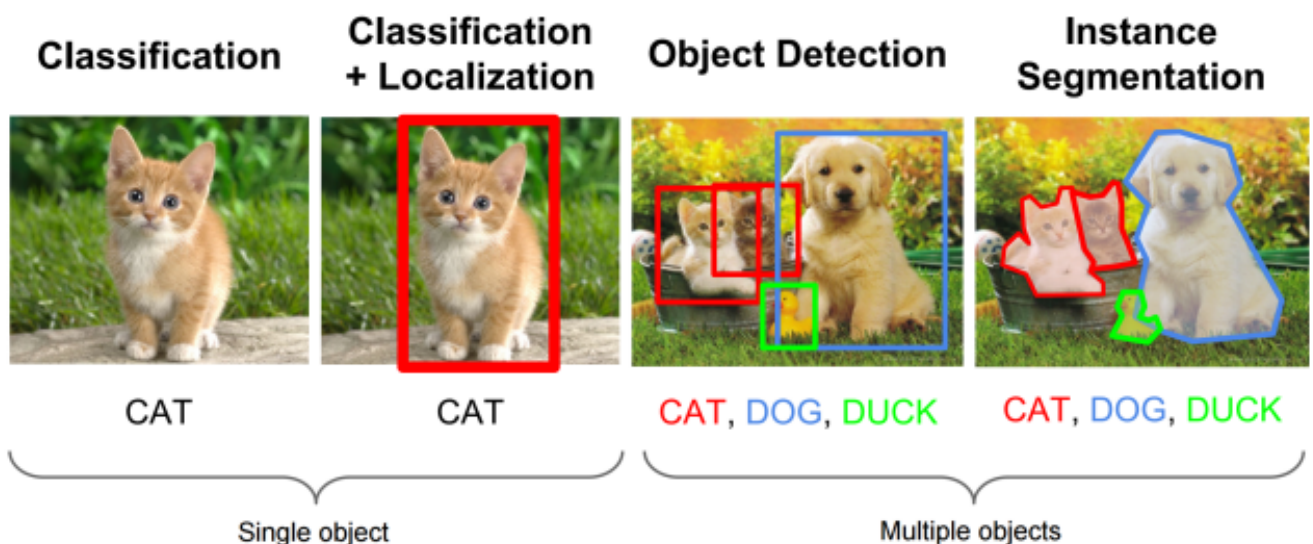


Рисунок 1.1 Различные задачи компьютерного зрения

На рис. 1.2 представлены различные виды сегментаций.



Рисунок 1.2 Виды сегментаций

## 1.2 Семантическая сегментация.

Семантическая сегментация – это задача поиска групп пикселей изображения, каждая из которых характеризует один смысловой объект. Пиксели, принадлежащие к определенной группе, просто относятся к этой группе без учета другой информации или контекста.

При семантической сегментации требуется не только обнаружить и классифицировать объекты, но и определить их границы.

Сегментация изображений находит применение во многих сферах, включая управление беспилотным транспортом, промышленную робототехнику, автомобильную промышленность и медицину. В этих областях семантическая сегментация позволяет проводить более точный анализ различных типов изображений, таких как спутниковые снимки и карты, рентгенографии, электронно-микроскопические снимки, а также изображения, полученные с камер автопилотируемых автомобилей и роботов.

На рис. 1.3 представлен пример семантической сегментации.



Рисунок 1.3 Семантическая сегментация

Если говорить точнее, то имеем следующее:

1. В управлении беспилотным транспортом сегментация изображений помогает распознавать дорожные знаки, пешеходов, препятствия и другие объекты на дороге, что позволяет принимать более точные решения и обеспечивать безопасность в автономном режиме.
2. В индустриальной робототехнике семантическая сегментация может использоваться для распознавания и классификации объектов в производственной среде, что помогает роботам взаимодействовать с окружающими предметами и выполнять задачи автоматизации с большей точностью.
3. В автомобильной промышленности сегментация изображений применяется для обнаружения и классификации объектов на дороге, таких как автомобили, пешеходы, велосипедисты, а также для анализа дорожной инфраструктуры, включая разметку и дорожные знаки. Это важная задача для развития автономных систем и повышения безопасности на дорогах.

4. В медицине семантическая сегментация используется для анализа медицинских изображений, таких как рентгенограммы и электронно-микроскопические снимки. Она помогает в точной локализации и классификации аномалий, опухолей и других патологических изменений, что важно для диагностики и планирования лечения.

## 1.3 Методы и алгоритмы семантической сегментации изображений

Для решения задач семантической сегментации используются различные подходы. Ниже приведены некоторые из них:

- граничные алгоритмы;
- пороговые алгоритмы;
- алгоритмы наращивания областей;
- алгоритмы сегментации на основе кластеризации;
- нейронные сети;
- и другие.

Дальше рассмотрим некоторые из них.

## 1.4 Граничные алгоритмы

Граничные методы являются одним из подходов в области обработки изображений и компьютерного зрения, направленных на выделение контуров объектов. Они применяются в различных приложениях, включая распознавание образов, сегментацию изображений, анализ медицинских снимков и другие.

Работа граничных методов обычно состоит из двух этапов. На первом этапе происходит поиск граничных пикселей, то есть точек на изображении, которые находятся на границе объектов. Для этого применяются различные операторы, такие как операторы Робертса, Превитта, Собеля и другие. [5] Эти операторы выявляют различные характеристики границ, такие как изменение яркости



или градиент интенсивности, и помогают определить местоположение границ объектов.

Пусть исходное изображение рис. 1.4, включая черно-белые и цветные изображения,  $A$  представляется в виде двумерной матрицы.

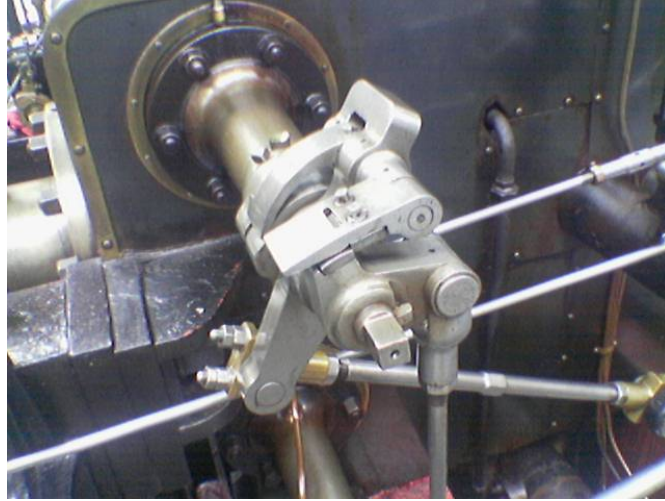


Рисунок 1.4 Исходное изображение

Оператор Робертса представляет собой один из простейших способов нахождения первых частных производных в точке  $x_*$ , который описывается следующими масками:

$$G_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} * A \quad \text{and} \quad G_2 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} * A \quad (1.1)$$

где точка  $x_*$  расположена в левом верхнем углу.

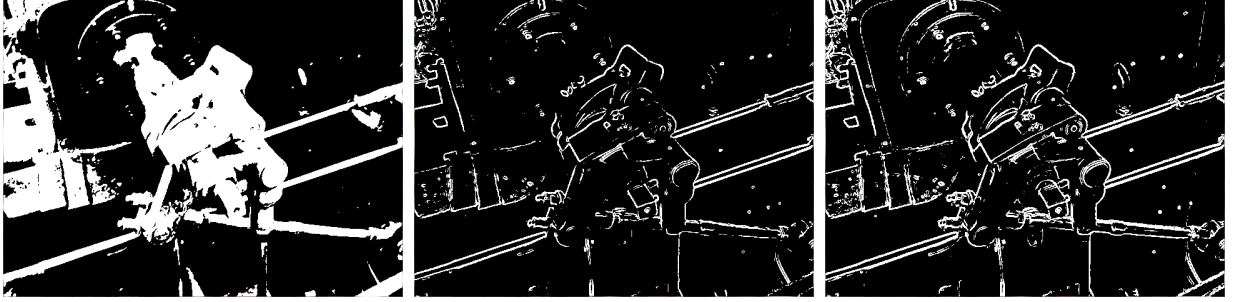
Оператор Превитта, так же как и оператор Робертса, вычисляет первые частные производных в точке, однако описывается другими масками:

$$G_1 = \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} * A \quad \text{and} \quad G_2 = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} * A \quad (1.2)$$

Оператор Собеля является модификацией оператора Превитта, а видоизменение заключается в использовании весового коэффициента 2 для средних элементов.

Маски, используемые оператором Собеля имеют следующий вид:

$$G_1 = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} * A \quad \text{and} \quad G_2 = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} * A \quad (1.3)$$



(a) Оператор Робертса

(b) Оператор Превитта

(c) Оператор Собеля

Рисунок 1.5 Результаты применения операторов

Величина перепада  $G$  получаемого изображения вычисляется из полученных значений  $G_1$ ,  $G_2$  по правилу:

$$G = \sqrt{G_1^2 + G_2^2}, \quad (1.4)$$

or

$$G = |G_1| + |G_2|, \quad (1.5)$$

где (1.4) - Евклидова метрика, (1.5) - Метрика городских кварталов.

На втором этапе происходит замыкание границ, то есть соединение граничных пикселей в непрерывные контуры. Это может быть достигнуто с помощью различных алгоритмов, таких как алгоритм замыкания границ поиском в глубину (Depth-First Search) или алгоритм связности (Connected Component Labeling). Замыкание границ позволяет получить более полные и непрерывные контуры объектов на изображении.

Однако основным недостатком граничных методов является их сильная зависимость от шума на изображении. Шум может привести к ложным границам или пропуску реальных границ объектов, что ведет к пересегментации. Пересегментация означает разделение одного объекта на несколько отдельных сегментов из-за некорректного определения границ. Это может быть проблемой при

дальнейшем анализе изображений или распознавании объектов.

Для снижения влияния шума и улучшения результатов сегментации объектов могут применяться различные методы фильтрации или сглаживания изображений перед применением граничных методов. Также существуют более сложные алгоритмы сегментации, которые комбинируют граничные методы с другими подходами, такими как методы на основе регионов или методы на основе графов, чтобы достичь более точных результатов сегментации объектов на изображении.

## 1.5 Пороговые алгоритмы

Пороговые алгоритмы являются простыми и широко используемыми методами обработки изображений. Они основаны на идее разделения гистограммы изображения на различные части по пороговым значениям. Пиксели, значения которых превышают пороговое значение, устанавливаются равными 1, а пиксели, значения которых меньше порогового значения, устанавливаются равными 0.

В результате чего, изображение преобразуется в двоичную карту, что приводит к процессу, который часто называют бинаризацией. Пороговое значение изображения очень полезно, если разница в значениях пикселей между двумя целевыми классами очень велика, и легко выбрать среднее значение в качестве порога [4].

Пороговое значение часто используется для бинаризации изображения, чтобы можно было использовать дополнительные алгоритмы, такие как обнаружение контура и идентификация, которые работают только с бинарными изображениями.

Существует два основных метода пороговых алгоритмов:

- Метод с глобальным порогом;
- Метод с адаптивным порогом.

### 1.5.1 Метод с глобальным порогом

В глобальных методах сегментации происходит работа со всем изображением сразу. В ходе работы находится порог сегментации  $T$ , с помощью которого происходит деление на черное и белое, причем величина порога  $T$  остается неизменной в течение всего процесса сегментации [4]. К пороговым методам сегментации относятся:

- сегментация с нижним порогом;
- сегментации с верхним порогом;
- сегментация с двойным ограничением;
- неполная пороговая обработка;
- многоуровневое пороговое преобразование.

Одним из наиболее простых методов преобразования изображения является сегментация с нижним порогом, в котором рассматривается лишь одно значение порога:

$$F'(m, n) = \begin{cases} 0, & F(m, n) \geq T, \\ 1, & F(m, n) < T, \end{cases} \quad (1.6)$$

где  $F(m, n)$  – исходное изображение;  $m, n$  – размерность матрицы изображения;  $T$  – порог значений яркости изображения.

Если в формуле (2.4) для точки изображения выполняется первое условие, то такая точка является точкой объекта, если же выполняется второе условие, то точка будет точкой фона.

В некоторых случаях можно использовать вариант метода сегментации с нижним порогом, в результате которого получается негатив исходного изображения. Такой метод называется сегментацией с верхним порогом и представляется формулой:

$$F'(m, n) = \begin{cases} 0, & F(m, n) < T, \\ 1, & F(m, n) \geq T. \end{cases} \quad (1.7)$$

Метод сегментации с двойным ограничением, используется для выделения определенных областей, в которых значения яркости пикселей находятся в заданном диапазоне. Этот метод основывается на установлении верхнего порога

для значений яркости пикселей и отбрасывании всех пикселей, чьи значения яркости превышают этот порог.

Формула для сегментации с двойным ограничением может быть записана следующим образом:

$$F'(m, n) = \begin{cases} 0, F(m, n) \leq T_1, \\ 1, T_1 < F(m, n) \leq T_2, \\ 0, F(m, n) > T_2, \end{cases} \quad (1.8)$$

где  $T_1, T_2$  – первый и второй порог яркости изображения.

Если требуется получить простое для дальнейшего анализа изображение, можно применить алгоритм неполной пороговой обработки. В результате этого алгоритма фон изображения удаляется вместе со всеми его деталями, которые присутствовали на исходном изображении. Формула для неполной пороговой сегментации может быть записана следующим образом:

$$F'(m, n) = \begin{cases} F(m, n), F(m, n) > T, \\ 0, F(m, n) \leq T. \end{cases} \quad (1.9)$$

В случае, когда требуется получить изображение, содержащее сегменты с различными уровнями яркости, можно применить метод многоуровневого порогового преобразования. Однако в результате этого преобразования полученное изображение уже не будет бинарным, то есть не будет состоять только из черного и белого цветов. Формула для этого преобразования может быть записана следующим образом:

$$F'(m, n) = \begin{cases} 1, F(m, n) \in D_1, \\ 2, F(m, n) \in D_2, \\ \dots \\ n, F(m, n) \in D_n, \\ 0, . \end{cases} \quad (1.10)$$

где  $D_i$  – некоторые области изображения,  $i = 1, 2, \dots, n$

### 1.5.2 Метод с адаптивным порогом

Адаптивные (локальные) методы сегментации производят разбивку изображения на несколько областей, для каждой из которых необходимо вычислить порог, основываясь на информации об интенсивности пикселей. Это позволяет более гибко адаптироваться к изменениям освещения или контрастности в разных областях изображения. Метод с адаптивным порогом особенно полезен для изображений с неравномерным освещением или изменяющимся фоном.

Алгоритмы данного класса предусматривают разделение изображения на блоки определенного размера. Размер блока должен быть минимальным, чтобы сохранить исходные особенности и детали изображения, но в то же время достаточно большим, чтобы минимизировать влияние шумов на результат [3].

При использовании алгоритмов разбиения на блоки применяется функция сглаживания для результирующего раstra при адаптивной сегментации. Эта функция позволяет получить удовлетворительный результат без использования дополнительных фильтров.

Наиболее популярные адаптивные методы сегментации изображений:

- метод Ниблэка;
- метод Бернсена;
- метод Савола.

Метод Ниблэка принадлежит к семейству алгоритмов с адаптивным порогом и определяет пороговое значение  $T$  для каждого пикселя путем перемещения по всему изображению прямоугольного окна определенного размера. Порог рассчитывается с помощью двух параметров – локального среднего значения  $m$  и стандартного отклонения  $S$ . Данные параметры в свою очередь рассчитываются индивидуально для каждого пикселя в пределах его окна. Для расчета локального порога используется следующая формула:

$$T_i = m + k * S, i = 1, 2, ..., n, \quad (1.11)$$

где  $m$  – среднее значение яркости пикселей в пределах локального окна;  $k$  – заранее подобранный коэффициент, определяющий значимость стандартного отклонения;  $n$  – количество локальных порогов на изображении;  $S$  – стандартное

отклонение, которое может быть выражено по формуле:

$$S = \sqrt{\frac{1}{N} \sum_{i=1}^N (p_i - m)^2}, \quad (1.12)$$

где  $p_i$  – значение яркости  $i$ -го пикселя в локальном окне;  $N$  – количество пикселей внутри окна.

Метод Ниблэка выбирает оптимальный порог  $T_*$  с помощью вычисления минимизации между  $T_i$ :

$$T_* = \arg\{\min_{i=1 \rightarrow n} \{T_i\}\}, \quad (1.13)$$

Огромным достоинством данного метода является его способность корректно обрабатывать изображения, на которых присутствует плавное изменение освещения, что достигается благодаря использованию адаптивного порога.

На рис. 1.6 представлен результат применения метода Ниблэка:



(a) Оригинальное изображение

(b) Изображение после применения метода Ниблэка с  $k = 8$

Рисунок 1.6 Демонстрация применения метода Ниблэка

Метод Бернсена предполагает деление всего изображения на квадраты размером  $r \times r$ , имеющие центр в точке  $(m, n)$ . Для каждого пикселя изображения в пределах квадрата используется порог, имеющий значение по следующей формуле:

$$T(m, n) = \frac{j_{max} + j_{min}}{2}, \quad (1.14)$$

где  $j_{max}$ ,  $j_{min}$  – являются наибольшим и наименьшим уровнем яркости в квадрате:

$$j_{max} - j_{min} \leq \epsilon, \quad (1.15)$$

где  $\epsilon$  – заданная пороговая величина, то исследуемый квадрат содержит объекты только одного класса: объектов или фона.

К недостаткам данного метода можно отнести большую зашумленность фоновых участков, а также невозможность применения интегральных изображений, поскольку при расчете порогового значения  $T$  не используется среднее значение пикселей в пределах окна, что в свою очередь приводит к повышению вычислительной сложности.

На рис. 1.7 представлен результат применения метода Бернсена:

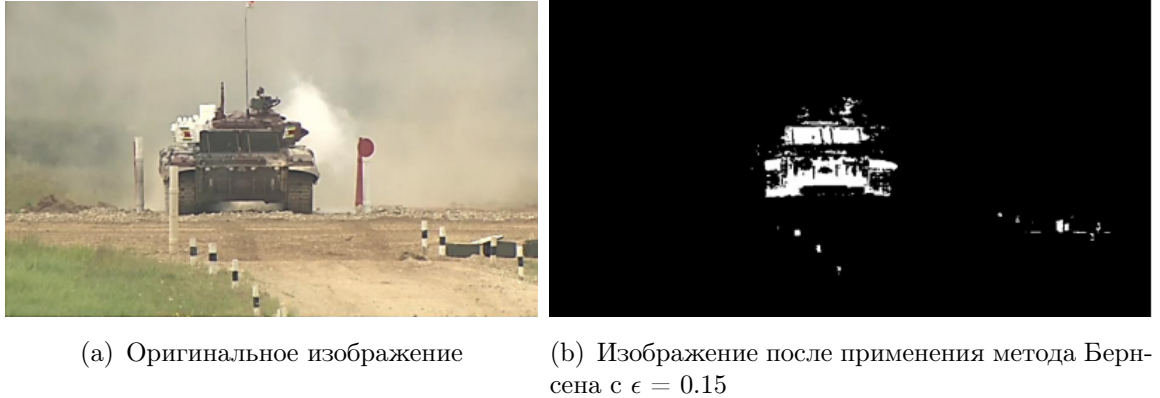


Рисунок 1.7 Демонстрация применения метода Бернсена

Метод Савола является модификацией метода Ниблэка, который использует локальное пороговое значение. Это значение вычисляется с использованием среднего значения  $m(x, y)$  и стандартного отклонения  $S(x, y)$  интенсивностей пикселей в пределах окна размером  $w \times w$ , центрированного вокруг пикселя  $(x, y)$ . Пороговое значение  $T(x, y)$  рассчитывается по следующей формуле:

$$T_i = m(x, y) \left( 1 + k \left( \frac{S(x, y)}{R} - 1 \right) \right), \quad (1.16)$$

где  $R$  – максимально возможное значение стандартного отклонения ( $R = 128$  для черно-белого изображения), а  $k$  – параметр, который принимает положительное значение из диапазона  $[0; 0.5]$ .

На рис. 1.8 представлен результат применения метода Савола:





(а) Оригинальное изображение



(б) Изображение после применения метода Савола с  $k = 0.15$

Рисунок 1.8 Демонстрация применения метода Савола

## 1.6 Алгоритмы наращивания областей

Алгоритмы наращивания областей в простейшем варианте основаны на выборе пикселя и анализе его соседей для проверки их яркостей [5]. Если значения яркости близки, то эти пиксели объединяются в одну группу, формируя область.

На рис. 1.9 представлена идея работы алгоритма «Region growth»:

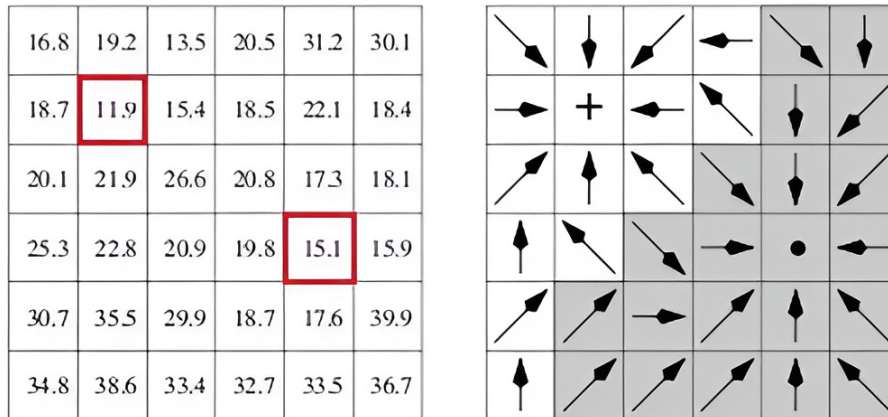


Рисунок 1.9 Исходное изображение

Более эффективные варианты алгоритмов используют не отдельные пиксели в качестве начальной точки, а небольшие области изображения. Каждая область проверяется на однородность, и если результат неудовлетворительный, область разбивается на более мелкие. Этот процесс продолжается до тех пор, пока все области не пройдут проверку на однородность. Затем формируются окончательные области путем наращивания этих более мелких областей. Наибо-

лее распространенные алгоритмы нахождения областей включают в себя такие алгоритмы, как алгоритм водоразделов и алгоритм центроидного связывания.

Однако недостатком таких алгоритмов является необходимость вручную выбирать начальную точку для сегментируемой области, что ограничивает возможность полной автоматизации метода.

## 1.7 Алгоритмы сегментации на основе кластеризации

Кластеризация - это задача группировки объектов в наборе данных на основе их сходства, так чтобы объекты внутри одного кластера были более похожи друг на друга, чем на объекты из других кластеров. Входными данными для задачи кластеризации являются признаковые описания объектов – наборы характеристик, свойственные каждому объекту. Выходными данными служат кластеры (непересекающиеся множества). Чтобы привести задачу сегментации к задаче кластеризации, нужно отобразить точки изображения в пространство признаков, а также ввести в этом пространстве признаков метрику.

По способу выделения кластеров все алгоритмы автоматической классификации можно разделить на две большие группы – иерархические и неиерархические. Иерархические алгоритмы позволяют обнаружить вложенные кластеры. Для этого строится либо дерево кластеров, называемое дендрограммой, либо так называемая диаграмма достижимости (по которой можно построить дендрограмму). Неиерархические алгоритмы вычисляют кластеры, исходя из оптимизации некоторого заранее заданного (явно или неявно) критерия качества [7].

Задача сегментации на основе кластеризации находит широкое применение в анализе спутниковых данных и обладает следующими особенностями:

- большой объём обрабатываемых данных (порядка  $10^6$  –  $10^7$  пикселей);
- отсутствие априорной информации о количестве и вероятностных характеристиках классов (при исследовании природных объектов получение априорной информации зачастую связано со значительными и не всегда оправданными затратами ресурсов);
- наличие «шума» и выбросов в данных.

## 1.8 Нейронные сети

Нейронные сети в упрощенном виде можно представлять как способ моделирования в технических системах принципов организации и механизмов функционирования головного мозга человека. В настоящее время наиболее популярным решением проблемы семантической сегментации является использование нейронных сетей. Благодаря им точность семантической сегментации значительно увеличилась. Среди основных характеристик ИНС выделяют обучение на примерах и обобщение полученных знаний, подавление шумов, отказоустойчивость.

Нейронные сети, выполняющие сегментацию, обычно используют структуру кодер-декодер, в которой за кодером следует узкое место, а декодер или уровни повышения дискретизации следуют непосредственно после узкого места.

Архитектуры кодер-декодер для семантической сегментации стали популярными с появлением такой архитектуры, как SegNet в 2015 г..

SegNet был в первую очередь ориентирован на приложения для понимания сцены. Следовательно, он спроектирован так, чтобы быть эффективным как с точки зрения памяти, так и с точки зрения времени вычислений во время вывода. Он также значительно меньше по количеству обучаемых параметров, чем другие конкурирующие архитектуры, из-за чего он достаточно хорош для дорожных сцен [6].

SegNet предлагает использовать комбинацию блоков свертки и понижающей дискретизации, чтобы сформировать представление входных данных. Затем декодер реконструирует входную информацию, чтобы сформировать карту сегментов, выделяющую регионы на входе и группирующие их по классам. Наконец, декодер имеет сигмовидную функцию активации в конце, которая сжимает выходные значения в диапазоне  $[0,1]$ . На рис. 1.10 показана архитектура кодер-декодера.

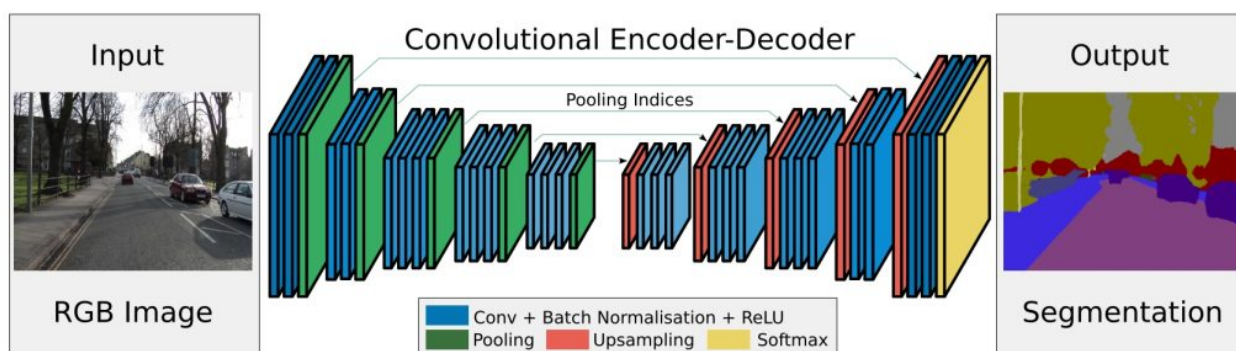


Рисунок 1.10 Архитектура кодер-декодера

В то же время, когда была представлена архитектура SegNet, было также предложено другое независимое решение для задачи сегментации - U-Net.

U-Net изначально создавалась для сегментации биомедицинских изображений. Она отлично зарекомендовала себя в этой области, отлично справившись с сегментацией опухолей головного мозга на материалах снимков магнитно-резонансной томографии.

Архитектура U-Net использовала пропускные соединения в глубоком обучении как способ решения проблемы потери информации, которая часто возникает в слоях понижения дискретизации типичных энкодер-декодер сетей. Пропускные соединения - это соединения, которые устанавливаются прямо от энкодера к декодеру, обходя узкое место. Это позволяет передавать информацию о контексте и деталях объектов между различными уровнями архитектуры и улучшает точность сегментации.

На рис. 1.11 показана архитектура нейронной сети U-net:

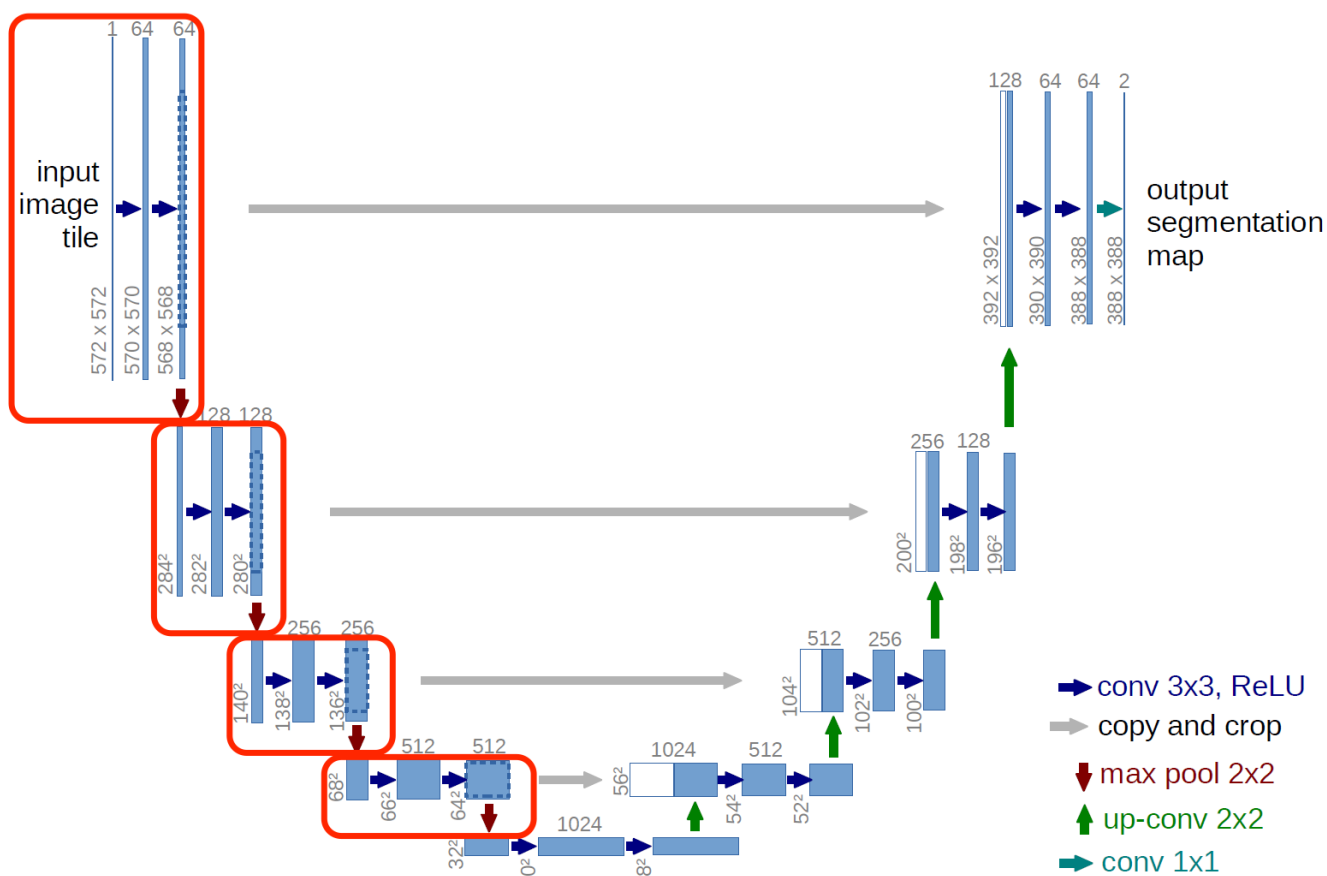


Рисунок 1.11 Архитектура нейронной сети U-net

Кроме этих двух нейросетей стоит также отметить ENet, которая была разработана с учетом требований эффективности и скорости работы, что делает ее особенно полезной для задач, где требуется обработка в реальном времени, например, в автономных транспортных средствах, системах видеонаблюдения, робототехнике и других приложениях, где требуется быстрая и точная сегментация изображений.

ENet, в отличие от вышеизложенных моделей, состоит из блоков с нетривиальной структурой, называемой bottleneck.

Архитектура bottleneck представлена на рис. 1.12:

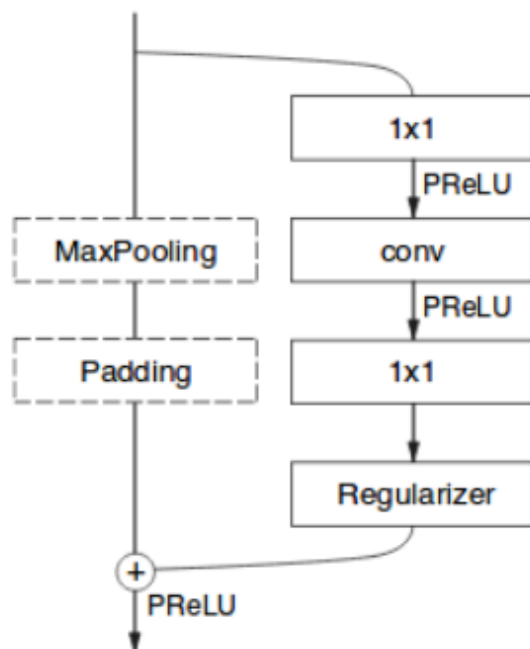


Рисунок 1.12 Архитектура bottleneck

Bottleneck состоит из основной ветви с макс-пулингом и паддингом и вспомогательной (со свертками), затем результаты ветвей сливаются и проходят через активацию PReLU.

Несмотря на большую глубину ENet, он содержит в 10-тки раз меньше параметров, чем U-net и SegNet, что хорошо сказывается на скорости работы и делает эту модель применимой для сегментации на мобильных устройствах в реальном времени.

## ГЛАВА 2

### Методы оценки эффективности сегментации

#### 2.1 Методы оценки эффективности сегментации

Для оценки эффективности методов и алгоритмов семантической сегментации зачастую рассматриваются следующие характеристики:

- Устойчивость;
- Дополнительные ограничения;
- Оптимальные условия применения метода.

Отдельно для методов сегментации с использованием нейронных сетей применим другие методы:

- Точность средней области интереса;
- Достоинства;
- Ограничения.

##### 2.1.1 Устойчивость

Устойчивость относится к способности алгоритма или модели поддерживать стабильные и последовательные результаты при изменении условий или входных данных. Устойчивость является важным свойством, поскольку она гарантирует, что сегментация будет достаточно надежной и нечувствительной к случайным или незначительным изменениям.

Если алгоритм или модель сегментации является устойчивым, то он демонстрирует следующие характеристики:

- Независимость от начального состояния: Устойчивый алгоритм не должен сильно менять результаты сегментации при незначительных изменениях начальных условий или параметров. Это означает, что малые изменения

не должны приводить к радикальным изменениям в границах или форме сегментов;

- **Консистентность:** Устойчивый алгоритм должен давать согласованные результаты при повторном запуске на одних и тех же данных или на разных подвыборках данных из одного и того же набора. Это позволяет получать предсказуемые и стабильные результаты;
- **Устойчивость к шуму:** Устойчивый алгоритм должен быть способен обрабатывать шум во входных данных без значительного искажения сегментации. Это означает, что случайные или незначительные изменения в данных не должны приводить к ошибкам или неправильным границам сегментов;
- **Устойчивость к вариациям:** Устойчивый алгоритм должен быть способен поддерживать схожие результаты при вариациях в условиях съемки, освещении, точности датчиков и других факторах, которые могут влиять на качество данных. Это гарантирует, что сегментация будет надежной и применима в различных сценариях.

Устойчивость является желательным свойством для алгоритмов и моделей сегментации, поскольку она обеспечивает надежность и последовательность результатов. Однако, в зависимости от конкретной задачи и особенностей данных, некоторые компромиссы между устойчивостью и точностью могут быть необходимы.

### 2.1.2 Дополнительные ограничения

Дополнительные ограничения могут включать в себя различные факторы, которые оказывают влияние на процесс сегментации и требуют дополнительных учета или ограничений.

Ниже приведены некоторые примеры дополнительных ограничений, которые могут быть применены в контексте сегментации:

- **Дополнительная информация:** Если доступна дополнительная информация о сегментируемых объектах или сцене, такая как цветовая информация, текстуры или контекстные данные, они могут быть использованы в



качестве дополнительных ограничений для улучшения процесса сегментации;

- Зависимость от параметров: Различные алгоритмы сегментации имеют свои собственные параметры, которые определяют их поведение и результаты. Например, в методах пороговой сегментации параметр порога определяет, какие пиксели будут отнесены к объекту, а какие - к фону. Выбор оптимальных параметров может быть сложной задачей и требует экспериментов или оптимизации;
- И другие.

Дополнительные ограничения могут быть специфичными для конкретной задачи или сценария сегментации. Их применение зависит от целей и требований сегментации, а также от доступных данных и контекста задачи.

### **2.1.3 Оптимальные условия применения метода**

Оптимальные условия применения метода относятся к идеальным или наилучшим условиям, при которых метод достигает наивысшей эффективности, точности и результативности в решении конкретной задачи. Это означает, что метод применяется в оптимальном контексте, с учетом оптимальных факторов и параметров, чтобы достичь наилучших возможных результатов.

### **2.1.4 Точность средней области интереса**

Точность средней интересующей области (mean Intersection over Union, mIoU) является метрикой, используемой для оценки эффективности сегментации изображений. Она измеряет сходство между предсказанной сегментационной маской и эталонной маской, учитывая как пиксели, принадлежащие области сегментации, так и пиксели, не принадлежащие этой области.

Для вычисления mIoU сначала вычисляется Intersection over Union (IoU) для каждого класса. IoU для каждого класса определяется как отношение площади пересечения между предсказанной маской и эталонной маской этого класса к площади их объединения. Затем, для получения mIoU, вычисляется среднее значение IoU по всем классам.

Формула для вычисления mIoU выглядит следующим образом:

$$mIoU = \frac{\sum_{i=1}^n (IoU_i)}{n} \quad (2.1)$$

Где  $n$  - количество классов,  $IoU_i$  - Intersection over Union для каждого класса. mIoU является популярной метрикой для оценки качества сегментации изображений, особенно при работе с несбалансированными классами. Свыше 0.5 обычно считается хорошим значением mIoU, при этом 1 соответствует идеальной сегментации.

### 2.1.5 Достоинства

Достоинство алгоритма означает его положительные характеристики или преимущества. Оно определяет, насколько хорошо алгоритм выполняет свою задачу и какие преимущества он предлагает по сравнению с другими алгоритмами или подходами.

### 2.1.6 Ограничения

Ограничения алгоритма означает недостатки или условия, с которыми алгоритм может столкнуться или которые могут оказывать влияние на его эффективность или точность. Например, ограничением может быть то, что ваш алгоритм плохо обрабатывает определенный тип данных или например необходимость предобработки входных данных.

## ГЛАВА 3

### Анализ избранных методов и алгоритмов

#### 3.1 Анализ алгоритмов сегментации

На основании предыдущих глав мы можем оценить каждый из вышеуказанных алгоритмов, описать их преимущества и недостатки сравнительно друг с другом.

##### 3.1.1 Анализ граничных алгоритмов

Граничные алгоритмы сегментации зачастую не рассматриваются как отдельное решение проблемы, а используются в паре с другими алгоритмами, для уточнения результата сегментации.

Для них свойственны следующие характеристики:

- Устойчивость: результат сегментации может быть затронут любым видом шума, а низкий контраст изображения может негативно сказаться на точности обнаружения границ;
- Дополнительные ограничения: обнаружение всех границ может затруднить поиск связи между интересующей областью и обнаруженными границами;
- Оптимальные условия применения метода: изображения с высоким контрастом между различными областями.

##### 3.1.2 Анализ пороговых алгоритмов

Пороговые алгоритмы сегментации чаще используются для выбора интересующей области изображения, игнорируя ненужные части. В медицинской сфере пороговые алгоритмы используются для обработки и анализа медицинских изображений, таких как рентгеновские снимки, снимки МРТ, УЗИ и другие. Представляя собой один из этапов в процедуре обработки медицинских данных.

Для них свойственны следующие характеристики:

- Устойчивость: успешность применения данного метода зависит от условий освещения и правильно подобранного порогового значения;
- Дополнительные ограничения: результаты сегментации зависят от выбранного порогового значения. Повышенная чувствительность к шуму и неоднородностям в интенсивности также может влиять на процесс сегментации:
- Оптимальные условия применения метода: простые изображения с высоким контрастом между различными областями.

### 3.1.3 Анализ алгоритмов наращивания областей

Алгоритмы наращивания областей применяются в робототехнике для обнаружения и сегментации объектов или препятствий в окружающей среде робота. Они позволяют роботу распознавать и взаимодействовать с различными объектами или структурами в окружающей среде.

Для них свойственны следующие характеристики:

- Устойчивость: наращивание области может быть чувствительно к шуму, что приведет к неверному определению областей;
- Дополнительные ограничения: точность и качество сегментации зависят от выбранных начальных параметров. Для каждой сегментированной области требуется ручной выбор начальной точки, что означает, что метод не может быть полностью автоматизирован;
- Оптимальные условия применения метода: простые изображения с высоким контрастом между различными областями.

### 3.1.4 Анализ алгоритмов кластеризации

Алгоритмы кластеризации являются наиболее распространенным подходом к сегментации спутниковых изображений. Например, они применяются для дистанционного зондирования Земли из космоса.

Для них свойственны следующие характеристики:

- Устойчивость: могут быть чувствительны к шуму и неоднородностям в интенсивности;
- Дополнительные ограничения: зависит от конкретного алгоритма, например, серьёзным недостатком алгоритма k-средних является необходимость задания числа классов, которое на практике чаще всего неизвестно и не существует эффективного алгоритма нахождения их числа, а также не способность выделять кластеры разной структуры;
- Оптимальные условия применения метода: зависит от конкретного алгоритма, например, для алгоритма k-средних: выборка осмысленных начальных центроид.

### 3.1.5 Анализ нейронных сетей в задаче сегментации изображений

Использованием нейронных сетей находит свое применение в различных задачах семантической сегментации, так как они обладают гибкостью и могут изучать сложные зависимости между пикселями изображения и их классами. Все это достигается благодаря тому, что их архитектуры включают слои, которые сохраняют пространственную информацию и обеспечивают более точную сегментацию, а также возможности эффективно использовать большие наборы данных для обучения.

Анализ нейросетевых алгоритмов:

- SegNet:

Точность средней области интереса(mIoU) составляет около 56.1%.

В достоинства данной нейронной сети можно записать ее хорошую производительность и быстрое действие.

Ограничения представляют собой необходимость адаптации для сегментации, а также невысокую точность сегментации.

- U-Net:

Точность средней области интереса(mIoU) составляет около 77.5%.

В достоинства данной нейронной сети можно записать ее хорошие результаты сегментации даже на небольших датасетах, а также высокую точность.

В ограничения данной сети можно записать невысокую точность сегментации для некоторых задач.

- ENet:

Точность средней области интереса(mIoU) составляет около 58.3%.

В достоинства данной нейронной сети можно записать то что она работает в реальном времени.

Ограничением является потребность в мощном процессоре.

Приведенные оценки описывались в [5].

## ЗАКЛЮЧЕНИЕ

Методы и алгоритмы семантической сегментации играют важную роль в различных областях компьютерного зрения и имеют широкий спектр применений. В курсовом проекте мы рассмотрели различные подходы к семантической сегментации, начиная от классических методов до современных глубоких нейронных сетей.

Основным результатом проделанной работы является обзор и сравнение различных подходов к семантической сегментации. В ходе работы были рассмотрены классические подходы к решению задачи, а также современные глубокие методы, использующие сверточные нейронные сети.

Современные глубокие методы, такие как U-Net, SegNet и ENet, демонстрируют высокую точность и способность выделять объекты семантически однородными регионами. Они позволяют получить более точные и семантически богатые результаты, нежели классические подходы. Однако, эти методы требуют большого объема размеченных данных для обучения и вычислительных ресурсов для эффективной работы.

В заключение, курсовой проект осветил методы и алгоритмы семантической сегментации и их применимости в различных областях. Современные методы семантической сегментации предоставляют мощные инструменты для анализа и понимания содержания изображений. Однако, они также требуют дальнейших исследований и инноваций для решения вызовов, связанных с ограничениями данных и вычислительной эффективности.

## СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

- [1] Bradski, G. Learning OpenCV / G. Bradski, A. Kaehler. – Sebastopol, CA : O'Reilly Media, Inc., 2008. – 281-331 p.
- [2] Solem, J.E. Programming Computer Vision with Python. / J.E. Solem. - San Francisco, CA :: O'Reilly Media, Inc., 2016. – 312 p.
- [3] Годунов, А. И. Сегментация изображений и распознавание объектов на основе технологии сверточных нейронных сетей / А. И. Годунов, С. Т. Баланян, П. С. Егоров // Надежность и качество сложных систем. – 2021. № 3. С. 62–73.
- [4] Лукашик, Д.В. Анализ современных методов сегментации изображений / Д.В. Лукашик // Экономика и качество систем связи. – 2022. – №2. – С. 57-65.
- [5] Шабалина, Д.Е. Семантическая сегментация изображений в проекте «Duckietown» / Д.Е. Шабалина, К.С. Ланчуковская, Т.В. Лях, К.В. Чайка // Вестник НГУ. Серия: Информационные технологии. – 2021. – Т. 19, №3. – С. 26-29.
- [6] Piramanayagam, S. Supervised Classification of Multisensor Remotely Sensed Images Using a Deep Learning Framework / S. Piramanayagam, E. Saber, W. Schwartzkopf, F. W. Koehler. - New York, CA : Rochester Institute of Technology, 2018. - 25 p.
- [7] Пестунов, И. А. Алгоритмы Кластеризации в Задачах Сегментации Спутниковых Изображений / И. А. Пестунов, Ю. Н. Синявский // Сибирское отделение РАН, 2012. - 16 с.