

УДК: 004.582

EDN: [CCORPV](https://ccorpv.com)

DOI: <https://doi.org/10.47813/2782-2818-2022-2-3-0101-0116>



## Обзор методов сегментации и обнаружения объектов на изображении в реальном времени для предотвращения аварийных ситуаций РЖД

А.Т. Тисецкий<sup>1</sup>, Д.И. Ковалев<sup>2,3</sup>, Т.П. Мансурова<sup>3</sup>

<sup>1</sup>Сибирский федеральный университет, Красноярск, Российская Федерация

<sup>2</sup>Красноярский государственный аграрный университет, Красноярск, Российская Федерация

<sup>3</sup>Красноярский краевой Дом науки и техники РосСНИО, Красноярск, Российская Федерация

**Аннотация.** С развитием железнодорожной индустрии, информатизации общества и автоматизации многих технологических процессов, появляется возможность создания аппаратно-программных комплексов автоматического управления, диагностики и безопасности движения локомотивов. Одной из важнейших систем данного комплекса является система обнаружения объектов на железнодорожных путях, разрывов железнодорожного полотна и его поворотов. Подобная система может быть разработана в виде камеры, установленной на локомотиве, и систем обработки информации, находящихся на борту каждого подвижного состава, или в виде глобальной системы, осуществляющей удаленную обработку информации с нескольких локомотивов. Независимо от реализации системы существует необходимость создания блока обнаружения объектов на изображениях, приходящих с камер. Для реализации данного блока необходимо выделять железнодорожную полосу на изображении и детектировать объекты в режиме реального времени. Для выделения полосы используются методы сегментации. В статье приведены алгоритмы нескольких из них и выбран наиболее предпочтительный вариант. Задача обнаружения объектов в видеопотоке в режиме реального времени решается при помощи сверточных нейронных сетей. В статье приведены краткие описания нескольких сетей, проанализированы результаты работы описанных нейронных сетей и выбрана сеть, наиболее подходящая для решения поставленной задачи.

**Ключевые слова:** железная дорога, обнаружение объектов, компьютерное зрение, сегментация изображения, разработка приложения, математическая модель, автоматизация процесса.

**Для цитирования:** Тисецкий, А., Ковалев, Д., & Мансурова, Т. (2022). Обзор методов сегментации и обнаружения объектов на изображении в реальном времени для предотвращения аварийных ситуаций РЖД. *Современные инновации, системы и технологии - Modern Innovations, Systems and Technologies*, 2(3), 0101–0116. <https://doi.org/10.47813/2782-2818-2022-2-3-0101-0116>

# Study of methods of segmentation and detection of objects in the image in real time to prevent accidents of Russian Railways

A.T. Tisetsky<sup>1</sup>, D.I. Kovalev<sup>2,3</sup>, T.P. Mansurova<sup>3</sup>

<sup>1</sup>*Siberian Federal University, Krasnoyarsk, Russian Federation*

<sup>2</sup>*Krasnoyarsk State Agrarian University, Krasnoyarsk, Russian Federation*

<sup>3</sup>*Krasnoyarsk Science and Technology City Hall, Krasnoyarsk, Russian Federation*

**Abstract.** With the development of the railway industry, the informatization of society and the automation of many technological processes, it becomes possible to create a complex of automatic control, diagnostics and safety of locomotive traffic. One of the most important systems of this complex is the system for detecting objects on the railway tracks, breaks in the railway track and its turns. Such a system can be designed as a locomotive-mounted camera and information processing systems on board each rolling stock, or as a global system that performs remote processing of information from several locomotives. Regardless of the implementation of the system, there is a need to create a block for detecting objects in images coming from cameras. To implement this block, it is necessary to select a railway lane in the image and detect objects in real time. Segmentation methods are used to select a band. The article presents the algorithms of several of them and chooses the most preferred option. The task of detecting objects in a video stream in real time is solved using convolutional neural networks. The article provides brief descriptions of several networks, analyzes the results of the described neural networks, and selects the network that is most suitable for solving the problem.

**Keywords:** railroad, object detection, computer vision, image segmentation, application development, mathematical model, process automation.

**For citation:** Tisetsky, A., Kovalev, D., & Mansurova, T. (2022). Study of methods of segmentation and detection of objects in the image in real time to prevent accidents of Russian Railways. *Modern Innovations, Systems and Technologies*, 2(3), 0101–0116. <https://doi.org/10.47813/2782-2818-2022-2-3-0101-0116>

---

## ВВЕДЕНИЕ

На сегодняшний день индустрия железнодорожных перевозок все еще остается довольно важной и связанной со многими сферами современной жизни. Помимо удобства, предоставляемого этой индустрией, имеет место большое количество инцидентов, связанных с ней. Один из таких – это аварийные ситуации, связанные с нахождением посторонних объектов на пути движущегося состава. Как правило, такие инциденты происходят там, где железнодорожная полоса проходит в местах, находящихся рядом с другими возможными потоками транспорта, а также в местах, где полосу может пересекать человек. То есть, самые распространенные места – это пересечения с дорожным полотном, железнодорожные станции, иные часто посещаемые людьми маршруты, находящиеся вблизи от железной дороги. Инциденты на

железнодорожных путях имеют множество негативных последствий, начиная от изменений расписания движения транспорта, задержек, проблем с логистикой, авариями транспорта, заканчивая реальными человеческими жертвами.

Правительства и железнодорожные компании по всему миру уделяют большое внимание безопасности на железных дорогах. По статистике – это приводит к уменьшению инцидентов, однако до полного их устранения еще очень далеко.

Долгосрочная программа развития (ДПР) ОАО «Российские железные дороги» разработана и синхронизирована с положениями Государственных программ Российской Федерации. К основным задачам данной программы относятся: обновление парка подвижного состава, в том числе тягового, с учётом заключения с его производителями контрактов жизненного цикла; развитие инфраструктуры для обеспечения перспективных объёмов перевозок и повышение производственной эффективности; обеспечение необходимого уровня безопасности движения и экологической безопасности; переход на «цифровую железную дорогу». Данные планы позволяют задуматься о возможности создания системы автоматического управления, диагностики и безопасности движения локомотива [1].

Для создания подобной системы необходима подсистема безопасности, которая позволит обнаруживать посторонние объекты на пути следования железнодорожного состава, а также разрывы самого железнодорожного полотна не только для быстрой остановки состава во избежание аварий, но также для обеспечения оперативного реагирования для устранения источников аварийных ситуаций.

Можно выделить две основных задачи при реализации подсистемы безопасности:

- 1) необходимо четко выделять границы железнодорожного полотна;
- 2) определять объекты, находящиеся на изображении.

Для решения первой задачи подходят методы сегментации изображения, а системы обнаружения объектов на видео в реальном времени позволяют решить вторую задачу. Сегментация изображений и построение систем понимания изображений в реальном времени являются ключевыми элементами в информатизации и автоматизации многих отраслей, в том числе и для отрасли железнодорожных перевозок. Таким образом, актуальной задачей является анализ и выбор соответствующих моделей и систем.

В данной работе проводится обзор нескольких методов сегментации изображения, а также рассматривается несколько нейронных сетей, реализующих

модели обнаружения объектов в реальном времени. В результате анализа осуществляется выбор наиболее подходящих сетей для решения поставленных задач.

## **МЕТОДЫ СЕГМЕНТАЦИИ ИЗОБРАЖЕНИЙ**

Прежде всего определим, что сегментация изображения — это процесс разделения цифрового изображения на несколько сегментов. Цель сегментации заключается в упрощении и/или изменении представления изображения, чтобы его было проще и легче анализировать [2]. Сегментация используется для выделения отдельных объектов на изображении.

### **Сегментация с использованием кластеризации**

Одним из методов сегментации изображения является метод кластеризации – *k*-средних. Данный метод используется, чтобы разделить изображение на определенное количество – *k*-кластеров.

Алгоритм метода *k*-средних:

1. Случайным образом или по каким-либо правилам выбрать *k*-центров кластеров;
2. Оценить «расстояние» от каждого пикселя до центров кластеров и отнести каждый пиксель к ближайшему из кластеров;
3. Далее пересчитать центры кластеров таким образом, чтобы они оказались в центре масс кластера;
4. Повторять пункты 2 и 3, пока на 3 пункте центры кластеров не окажутся в новых центрах масс.

Поскольку кластеризация происходит на изображении, то центры кластеров обладают показателями цвета, расположения по осям, иногда яркости и т.д. Следовательно, в качестве «расстояния» берется сумма квадратов разностей показателей центра и пикселя. Поскольку цвет, расстояние и яркость могут иметь разную размерность, перед расчётом следует нормализовать данные показатели.

Недостатком данного метода является сильная зависимость от начального расположения центров и их количества, вследствие чего сегментация может пройти неудачно или выделить лишние сегменты в рамках одного объекта. Также отсутствует возможность сегментации разноцветных объектов.

К преимуществам можно отнести простоту реализации алгоритма и скорость его выполнения.

### Метод выделения границ

Методы выделения границ позволяют определять границы объектов за счет перепада в яркости цвета на изображении. Одним из таких методов является оператор Кэнни – это многоступенчатый алгоритм, предназначенный для выявления границ объектов.

Прежде всего для данного метода необходимо преобразовать изображения в чёрно-белый вид. Это можно осуществить путем использования различных фильтров, таких как YUV, HSL, HSV и др.

Следующим шагом является сглаживание изображения для устранения шумов. Данный процесс возможно осуществить путем применения фильтра Гаусса [3]. Чаще всего этот процесс реализуют за счет маски Гауссианы с размером 5 и  $\sigma = 1.4$  (рисунок 1).

	2	4	5	4	2
	4	9	12	9	4
$\frac{1}{115}$	5	12	15	12	5
$\sigma=1.4$	4	9	12	9	4
	2	4	5	4	2

Рисунок 1. Маска Гауссианы.

Следующим шагом является поиск градиентов на изображении. Данный шаг очень часто реализуют при помощи оператора Собеля, который основывается на применении фильтров к изображению (рисунок 2).

-1	0	+1
-2	0	+2
-1	0	+1

G<sub>x</sub>

+1	+2	+1
0	0	0
-1	-2	-1

G<sub>y</sub>

**Рисунок 2.** Фильтры Собеля.

При помощи данных фильтров мы находим величину и направление градиента в каждом пикселе. При этом направление градиентов округляется таким образом, чтобы оно было равно 0°, 45°, 90°, 135°.

Затем объявляются пиксели границ, данными пикселями становятся пиксели с локальным максимумом. Они определяются в зависимости от направления и величины градиента таким образом, чтобы в заданном направлении градиента его величина у данного пикселя была больше, чем у соседних.

Основным недостатком данного метода является его сильная зависимость от шума, как правило, необходима правильная предобработка начального изображения. Также данный метод может оставлять рваные края при выделении границ.

## **ПРИМЕНЕНИЕ НЕЙРОННЫХ СЕТЕЙ ДЛЯ СЕГМЕНТАЦИИ**

На сегодняшний день лучших результатов по сегментации изображений достигают при помощи нейронных сетей. При решении данной задачи чаще всего используют полносверточные нейронные сети.

Полносверточные нейронные сети – это такие сверточные нейронные сети, в которых используются только слои свертки. На выходе данной сети возможно получение изображения, что позволяет данные сети использовать для сегментации изображения.

Сверточная нейронная сеть – специальная архитектура искусственных нейронных сетей, предложенная Яном Лекуном в 1988 году [4] и нацеленная на эффективное распознавание образов. Сверточная нейронная сеть состоит из чередующихся сверточных слоёв и субдискретизирующих слоёв. Структура сети – однонаправленная (без обратных связей), принципиально многослойная. Для обучения используются

стандартные методы, чаще всего метод обратного распространения ошибки. Функция активации нейронов – любая, по выбору исследователя.

Суть сверточных слоев заключается в процессе свертки. Свертка – это операция, в ходе которой ядро свертки (матрица – фильтр) перемещается по входной матрице, после чего результат матричного умножения участков входной матрицы и ядра свертки записывается в новую матрицу (карту признаков) (рисунок 3).

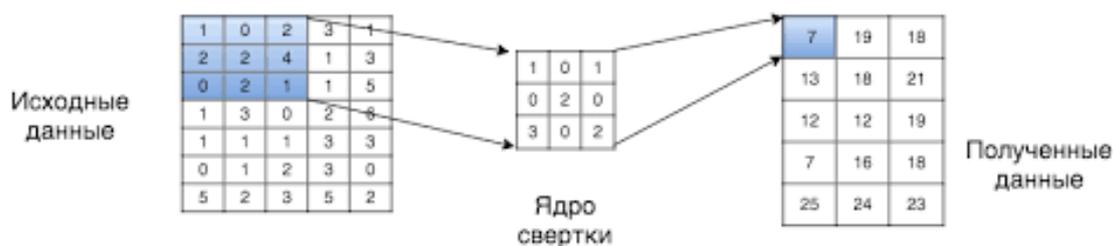


Рисунок 3. Процесс свертки.

Количество ядер совпадает с количеством каналов входного изображения. Так, если на вход сверточному слою подаётся RGB изображение, то сверточный слой будет содержать в себе 1 фильтр глубиной 3 (рисунок 4).

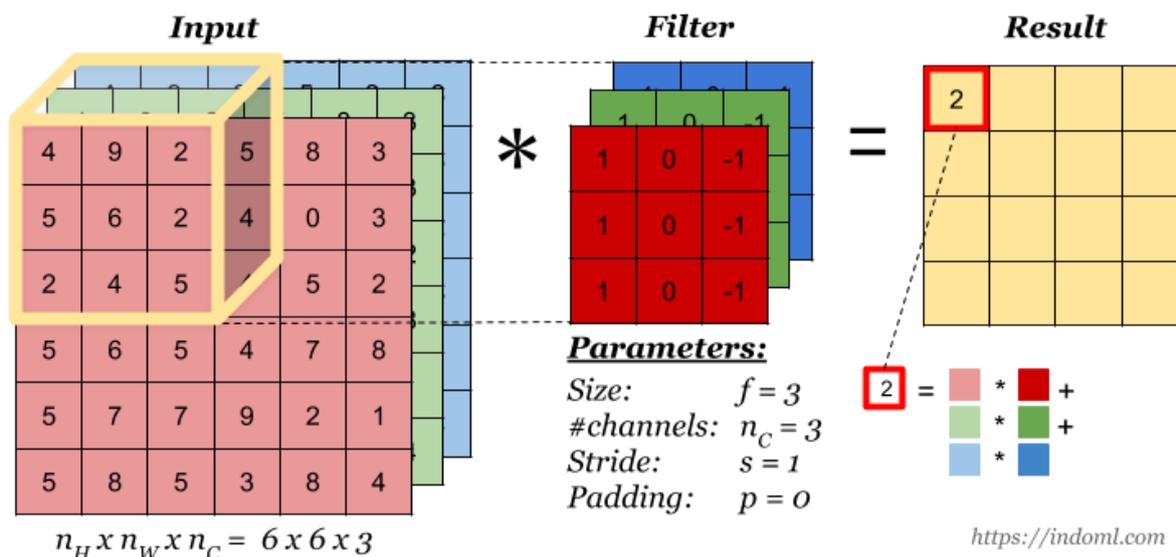


Рисунок 4. Свертка трехканального изображения.

На каждом сверточном слое может быть несколько фильтров, глубинна которых должна быть равна количеству входных каналов. В итоге, после сверточного слоя

получаем матрицы на нескольких каналах, количество которых будет равно количеству фильтров.

Помимо слоев свертки в полносверточной сети также присутствуют слои транспонированной свертки. На данных слоях происходит процесс, когда каждый элемент входных данных перемножается на фильтр, в итоге получается несколько матриц размерностью равные ядру, они помещаются в результирующую матрицу так, что при пересечении значений складываются (рисунок 5).

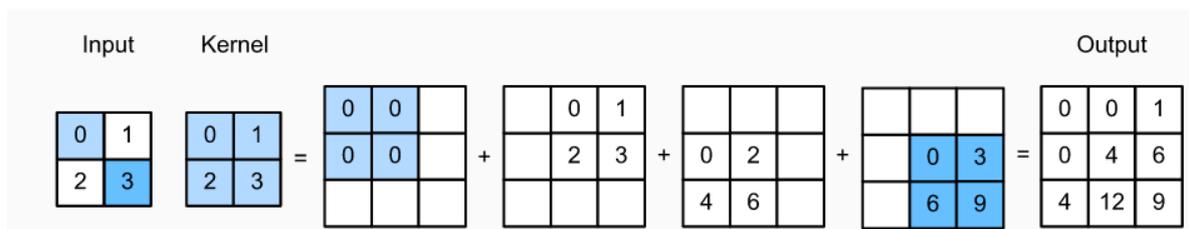


Рисунок 5. Транспонированная свертка.

Помимо слоя свертки, в сети участвует слой подвыборки (пулинга), который представляет собой нелинейное уплотнение карты признаков, при этом группа пикселей (обычно размера  $2 \times 2$ ) уплотняется до одного пикселя, проходя нелинейное преобразование. Наиболее употребительна при этом функция максимума [5] (рисунок 6).

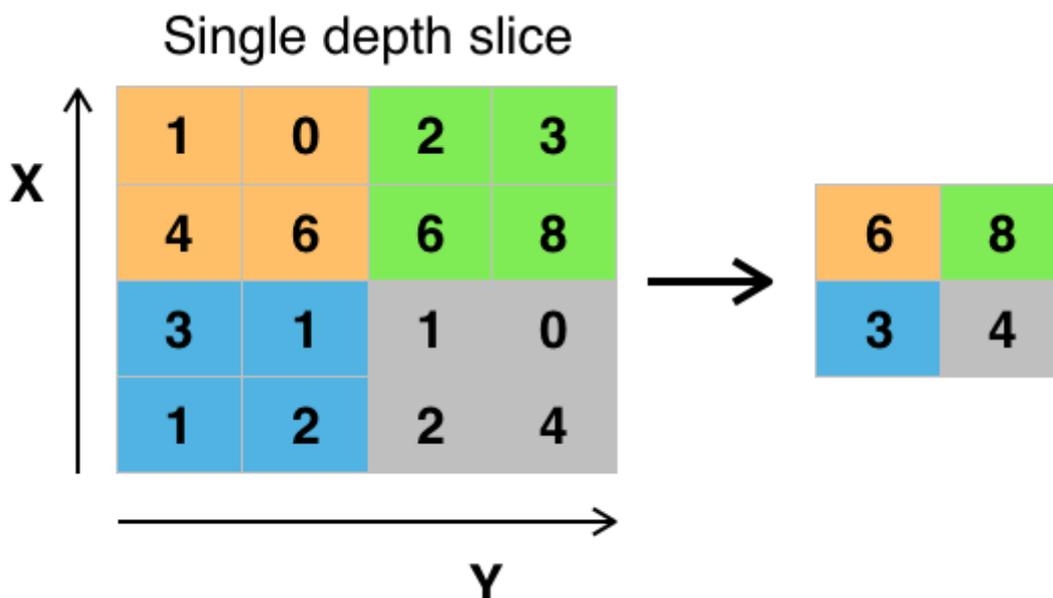


Рисунок 6. Процесс пулинга.

На сегодняшний день одной из лучших сегментирующих нейронных сетей признана сеть U-Net. Данная нейронная сеть в своей структуре помимо всех перечисленных выше слоев, чтобы получить более точные результаты, на каждом шаге после транспонированного слоя использует пропускные соединения, объединяя выходные данные транспонированных сверточных слоев с картами функций из сверточных слоев на тех же уровнях. После каждой конкатенации мы снова применяем две последовательные регулярные свертки, чтобы модель могла научиться собирать более точный результат (рисунок 7) [6].

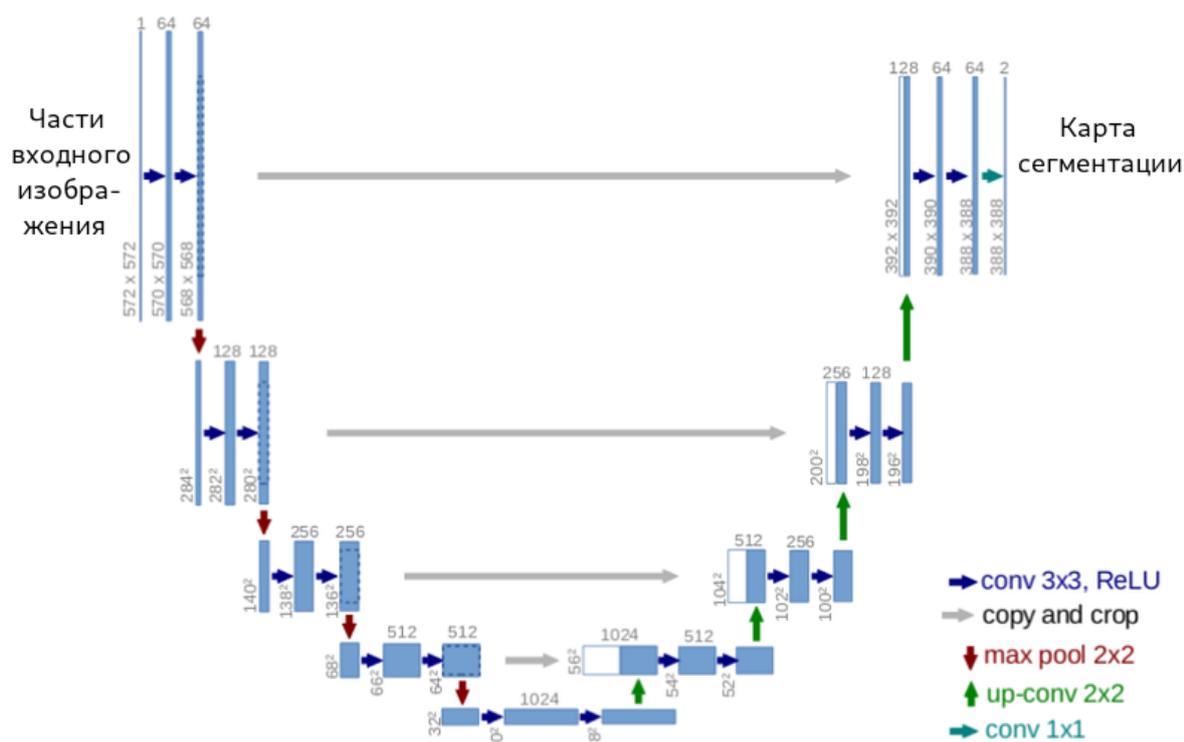


Рисунок 7. U-Net.

### Выбор метода

Проанализировав различные методы сегментации, был сделан вывод, что наиболее подходящим для решения поставленной задачи методом является метод, основанный на применении нейронных сетей для сегментации. Поскольку основным объектом, который необходимо исследовать, является железнодорожное полотно. Соответственно, чтобы выделять границы и определять разрывы на полотне, необходимо

постоянно четко определять границы железнодорожного полотна и следить за разрывами на них, остальные же объекты на изображении в анализе не нуждаются, если они не пересекают железнодорожное полотно.

Достичь данных результатов возможно, используя для обучения размеченные данные, на которых будут четко выделены границы железной дороги на изображении, а также изображения, на которых железная дорога будет пересекаться с объектом.

## **НЕЙРОННЫЕ СЕТИ ДЛЯ ОБНАРУЖЕНИЯ ОБЪЕКТОВ НА ВИДЕО В РЕЖИМЕ РЕАЛЬНОГО ВРЕМЕНИ**

Существующие архитектуры свёрточных нейронных сетей для обнаружения объектов на изображениях можно разделить на две категории: одноэтапные и двухэтапные. Двухэтапные нейросетевые алгоритмы обнаружения объектов на изображении включают в себя следующие два этапа.

Первый этап заключается в поиске подозрительных зон на изображении, т. е. зон, возможно содержащих интересующие нас объекты. Данный этап может выполняться без использования нейронных сетей с помощью информации о контрасте, по ключевым точкам или перебирая все возможные положения объекта с помощью процедуры селективного поиска. Однако зоны, полученные вышеперечисленными методами, зачастую имеют очень низкое качество в том смысле, что много областей либо содержат слишком большое количество фона, либо содержат только небольшую часть объекта, либо содержат более одного объекта. Поэтому более предпочтительно на данном этапе использовать свёрточные нейронные сети, не содержащие полносвязных слоев [7].

Второй этап заключается в классификации зон, найденных на первом этапе. На данном этапе для классификации всегда используются обычные свёрточные искусственные нейронные сети.

Одноэтапные нейросетевые алгоритмы обнаружения объектов на изображении не включают в себя стадию поиска подозрительных зон на изображении, а сразу нацелены на обнаружение объектов. Преимуществом данного вида алгоритмов является их простота и относительно высокая скорость работы, а среди их недостатков можно отметить более низкую точность обнаружения объектов по сравнению с двухэтапными алгоритмами, а также меньшую гибкость алгоритма [7].

В данной работе будут рассмотрены некоторые популярные нейронные сети.

## CenterNet

CenterNet – нейронная сеть, использующая подход, основанный на ключевых точках, для обнаружения объектов. Она рассматривает центр прямоугольника выделения объекта, как сам объект, а также, как ключевую точку, а затем использует этот предсказанный центр для нахождения координат и смещений ограничивающего объект прямоугольника.

Изначально на вход подается изображение, оно проходит предобработку для выделения признаков с помощью одной из 4 сетей на выбор: ResNet18, ResNet101, Deep Layer Aggregation Networks (DLA), Stacked Hourglass Networks. После субдискретизации с заданным шагом, при помощи слоев сети, генерируется тепловая карта для всех классов, на ней отображены ключевые точки входного изображения. Следующие слои определяют размеры и расположение прямоугольных областей для обнаруженных точек. И чтобы избежать ошибок дискретизации после субдискретизации, определяются локальные смещения для полученных точек [8].

## Single Shot Multibox Detector

Метод Single Shot Detector (SSD) используется для описания архитектур, в которых используется одна сверточная нейронная сеть (feedforward convolutional network) для непосредственного предсказания расположения областей и их классов, без применения второго этапа классификации. В этом методе на выходе нейронной сети формируются несколько тысяч прогнозов для возможных регионов расположения объектов разной формы на разных масштабах, затем с помощью подавления *немаксимумов* (Non-Maximum Suppression) происходит выбор нескольких наиболее вероятных областей. Такая единая структура, одновременно с учетом различных масштабов изображения обеспечивает методу SSD наиболее высокие показатели по скорости и качеству обнаружения объектов по сравнению с остальными современными подходами [9].

## RetinaNet

Архитектура свёрточной нейронной сети RetinaNet состоит из 4 основных частей, каждая из которых имеет своё назначение:

- 1) Backbone – основная (базовая) сеть, служащая для извлечения признаков из поступающего на вход изображения. Данная часть сети является вариативной и в её основу могут входить классификационные нейросети, такие как ResNet, VGG, EfficientNet и другие;
- 2) Feature Pyramid Net (FPN) – свёрточная нейронная сеть, построенная в виде пирамиды, служащая для объединения достоинств карт признаков нижних и верхних уровней сети, первые имеют высокое разрешение, но низкую семантическую, обобщающую способность; вторые — наоборот;
- 3) Classification Subnet – подсеть, извлекающая из FPN информацию о классах объектов, решая задачу классификации;
- 4) Regression Subnet – подсеть, извлекающая из FPN информацию о координатах объектов на изображении, решая задачу регрессии [10].

### YOLOv4

Архитектура свёрточной нейронной сети YOLOv4 состоит из 3 основных частей, каждая из которых имеет своё назначение:

а) Backbone, с помощью которого можно повысить точность, спроектировать более глубокую сверточную нейронную сеть и увеличить сложность модели. В архитектуре YOLOv4 используется CSPDarknet53 в качестве блока Backbone.

б) Neck, основной целью данного блока является добавление дополнительных слоев между блоком Backbone и блоком Head, чтобы получить «более богатую» пространственную и семантическую информацию. В архитектуре YOLOv4 используются PAN и SPP в качестве блока Neck.

в) Head реализует процесс, такой же, как и в архитектуре YOLOv3 – сверточная нейронная сеть определяет координаты bounding-boxes (x, y, w, h) вместе с оценкой достоверности для класса. Цель состоит в том, чтобы разделить изображение на сетку, состоящую из нескольких ячеек, а затем для каждой ячейки предсказать вероятность наличия объекта с помощью anchor-boxes. На выходе получается вектор с координатами bounding-boxes и классами вероятностей [11].

### СРАВНЕНИЕ ЭФФЕКТИВНОСТИ РАБОТЫ НЕЙРОННЫХ СЕТЕЙ

Для сравнения эффективности работы данных нейронных сетей было принято решение использовать данные о прохождении MS COCO набора данных для каждой из

сетей, проведенные на GPU с архитектурой Pascal: Titan X (Pascal), Titan Xp, GTX 1080 Ti, Tesla P100 GPU. В качестве метрик были выбраны: скорость работы нейронной сети в кадрах в секунду (fps) и AP – это среднее значение для нескольких IoU (минимальное IoU для рассмотрения положительного совпадения). AP @ [.5: .95] соответствует среднему AP для IoU от 0,5 до 0,95 с размером шага 0,05. Для соревнования COCO AP – это среднее значение по 10 уровням IoU по 80 категориям (AP @ [. 50: .05: .95]: от 0,5 до 0,95 с размером шага 0,05). Здесь IoU измеряет перекрытие между двумя областями. Это используется для определения процента перекрытия предсказанной областью нахождения объекта его реальной области нахождения. Для некоторых наборов данных для IoU предопределяется некий порог (например, 0,5), и это используется для классификации предсказания. Данные представлены в таблице 1.

**Таблица 1.** Параметры работы нейронных сетей.

Метод	Backbone	Size	FPS	AP	AP <sub>50</sub>	AP <sub>75</sub>	AP <sub>s</sub>	AP <sub>m</sub>	AP <sub>l</sub>
CenterNet	Hourglass-104	512x512	7.8	45.1	63.9	49.3	26.6	47.1	57.7
CenterNet	DLA-34	512x512	28	41.6	60.3	45.1	21.5	43.9	56.0
SSD	VGG-16	512x512	22	28.8	48.5	30.3	10.9	31.8	43.5
RetinaNet	ResNet-50	-	10.8	37.1	56.9	40.0	20.1	40.1	48.0
YOLOv4	CSPDarknet-	512x512	43	43.0	64.9	46.5	24.3	46.1	55.2

53

На основе полученных данных можно сделать вывод о том, что самой быстрой и второй по точности является нейронная сеть YOLOv4. Следовательно, это наиболее подходящая для наших целей нейронная сеть. Однако не стоит забывать об одном ее недостатке, она обладает наибольшим количеством признаков для обучения, это приводит к необходимости использовать наборы данных больших размеров при обучении, при этом данная проблема слегка упрощается, если использовать различные методы аугментации данных, такие как повороты, размытие изображения, мозаика и др.

## ЗАКЛЮЧЕНИЕ

В рамках данной работы были исследованы и изучены методы сегментации изображений и выбран наилучший метод, соответствующий задаче – выделение железнодорожной полосы на изображении путем его сегментирования. Лучшим методом

было признано использование нейронных сетей для сегментации, поскольку они позволяют выделять заранее обученные объекты на всем изображении.

Также проведен обзор сверточных нейронных сетей, предназначенных для обнаружения объектов в видео потоке в реальном времени. Были изучены алгоритмы работы нескольких из них. После изучения алгоритмов были проанализированы результаты тестов на наборе данных MS COCO. На основе этих тестов была составлена таблица, в которой приведены результаты теста, соответствующие запуску на GPU с архитектурой Pascal для каждой нейронной сети. Изучив полученную таблицу, была выбрана архитектура сверточной нейронной сети, как наиболее подходящая для обнаружения объектов на изображении в реальном времени.

Целью дальнейших исследований является программная реализация приведенных в работе алгоритмов для решения задачи обнаружения объектов на пути следования подвижного состава.

#### СПИСОК ЛИТЕРАТУРЫ

- [1] Распоряжение Правительства РФ № 466-р от 19.03.2019 г. Долгосрочная программа развития ОАО «РЖД» до 2025 года.
- [2] Koheri Arai, Ali Ridho Barakbah. *Heirarchical K-means: An algorithm for Centroid initialization for K-means*. Saga University, 2007.
- [3] Удалов Т.А. *Программный комплекс для выделения треков частиц на растровых изображениях видеопотока по набору геометрических примитивов: магистерская диссертация по направлению подготовки: 09.04.01 - Микропроцессорные системы*. Барнаул, 2016.
- [4] LeCun Y. et al. Backpropagation applied to handwritten zip code recognition, *Neural computation*, 1989, 1(4-q), 541-551.
- [5] *Банковская система: устойчивость и перспективы развития: сборник научных статей десятой международной научно-практической конференции по вопросам банковской экономики*. УО «Полесский государственный университет», г. Пинск 25 октября 2019 г. Министерство образования Республики Беларусь [и др.]; редкол.: К.К. Шебеко [и др.]. Пинск: ПолесГУ, 2019, 321, 326 с.
- [6] Ronneberger O., Fischer P., Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation, *Computer Science Department and BIOSS Centre for Biological Signalling Studies*, University of Freiburg, Germany.

- [7] Бондаренко В.А., Павлова В.А., Тупиков В.А., Холод Н.Г. Алгоритм нейросетевого распознавания надводных объектов в реальном времени, *Известия Тульского государственного университета. Технические науки*, 2021, 1, 19-33. EDN LBWTUH.
- [8] Zhou X., Wang D., Krähenbühl P. "Objects as points," 2019. arXiv:1904.07850. [URL]. Доступ: <http://arxiv.org/abs/1904.07850>
- [9] Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, and Alexander C Berg. SSD: Single shot multibox detector. *In Proceedings of the European Conference on Computer Vision (ECCV)*, 2016, 21-37.
- [10] Jiaqi Wang, Kai Chen, Shuo Yang, Chen Change Loy, and Dahua Lin. Region proposal by guided anchoring. *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, 2965-2974.
- [11] Bochkovskiy Alexey, Wang Chien-Yao, Liao HongYuan Mark. *YOLOv4: Optimal speed and accuracy of object detection*. arXiv preprint arXiv:2004.10934, 2020.

## REFERENCES

- [1] Decree of the Government of the Russian Federation No. 466-r dated March 19, 2019. Long-term development program of Russian Railways until 2025.
- [2] Koheri Arai, Ali Ridho Barakbah. *Heirarchical K-means: An algorithm for Centroid initialization for K-means*. Saga University, 2007.
- [3] Udalov T.A. *A software package for extracting particle tracks on bitmap images of a video stream by a set of geometric primitives: master's thesis in the direction of preparation: 09.04.01 - Microprocessor systems*. Barnaul, 2016.
- [4] LeCun Y. et al. Backpropagation applied to handwritten zip code recognition, *Neural computation*, 1989, 1(4.-q), 541-551.
- [5] *Banking system: sustainability and development prospects: a collection of scientific articles of the tenth international scientific and practical conference on banking economics*. EE "Polesye State University", Pinsk October 25, 2019 Ministry of Education of the Republic of Belarus [and others]; editorial board: K.K. Shebeko [i dr.]. Pinsk: PolesGU, 2019, 321, 326.
- [6] Ronneberger O., Fischer P., Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation, *Computer Science Department and BIOSS Centre for Biological Signalling Studies*. University of Freiburg, Germany.
- [7] Bondarenko V.A., Pavlova V.A., Tupikov V.A., Kholod N.G. Algorithm for neural network recognition of surface objects in real time, *Proceedings of the Tula State University*.

Technical Sciences, 2021, 1, 19-33. EDN LBWTUH.

[8] Zhou X., Wang D., Krähenbühl P. “*Objects as points*,” 2019, arXiv:1904.07850. [URL].  
Доступ: <http://arxiv.org/abs/1904.07850>.

[9] Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, and Alexander C Berg. SSD: Single shot multibox detector. *In Proceedings of the European Conference on Computer Vision (ECCV)*, 2016, 21-37.

[10] Jiaqi Wang, Kai Chen, Shuo Yang, Chen Change Loy, and Dahua Lin. Region proposal by guided anchoring. *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, 2965-2974.

[11] Bochkovskiy Alexey, Wang Chien-Yao, Liao HongYuan Mark. YOLOv4: Optimal speed and accuracy of object detection. arXiv preprint arXiv:2004.10934, 2020.

#### ИНФОРМАЦИЯ ОБ АВТОРАХ / INFORMATION ABOUT THE AUTHORS

**Тисецкий Артем Тимофеевич**, магистрант кафедры информатики, Институт космических и информационных технологий Сибирского федерального университета, Красноярск, Российская Федерация  
**E-mail:** tat500@mail.com

**Tisetsky Artem T.**, Master student of the Department of Informatics, Institute of Space and Information Technologies, Siberian Federal University, Krasnoyarsk, Russian Federation  
**E-mail:** tat500@mail.com

**Ковалев Дмитрий Игоревич**, аспирант кафедры информационных технологий и математического обеспечения информационных систем, Красноярский государственный аграрный университет, Российская Федерация  
**E-mail:** grimm7jow@gmail.com

**Kovalev Dmitry I.**, Postgraduate Student, Department of Information Technologies and Software for Information Systems, Krasnoyarsk State Agrarian University, Russian Federation  
**E-mail:** grimm7jow@gmail.com

**Мансурова Тамара Павловна**, научный сотрудник Красноярского краевого Дома науки и техники Российского Союза научных и инженерных общественных объединений, Российская Федерация  
**E-mail:** mansurovatp@mail.ru

**Mansurova Tamara P.**, Researcher of the Krasnoyarsk Science and Technology City Hall of the Russian Union of Scientific and Engineering Public Associations, Russian Federation  
**E-mail:** mansurovatp@mail.ru

*Статья поступила в редакцию 07.06.2022; одобрена после рецензирования 07.07.2022; принята к публикации 12.07.2022.*

*The article was submitted 07.06.2022; approved after reviewing 07.07.2022; accepted for publication 12.07.2022.*