



Министерство науки и высшего образования Российской Федерации  
Федеральное государственное бюджетное образовательное учреждение  
высшего образования  
«Московский государственный технический университет  
имени Н. Э. Баумана  
(национальный исследовательский университет)»  
(МГТУ им. Н. Э. Баумана)

---

ФАКУЛЬТЕТ «Информатика и системы управления»

КАФЕДРА «Программное обеспечение ЭВМ и информационные технологии»

---

## ОТЧЕТ

*ПО НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЙ РАБОТЕ*

НА ТЕМУ:

«Классификация методов распознавания надводных  
объектов с аэрофотоснимков»

Студент ИУ7-73Б  
(Группа)

\_\_\_\_\_  
(Подпись, дата)

Г. А. Миронов  
(И. О. Фамилия)

Руководитель НИР

\_\_\_\_\_  
(Подпись, дата)

К. Л. Тассов  
(И. О. Фамилия)

2022 г.

## РЕФЕРАТ

Расчетно-пояснительная записка 27 с., 13 рис., 2 табл., 28 источн., 1 прил.

**Ключевые слова:** РАСПОЗНАВАНИЕ, НАДВОДНЫЕ ОБЪЕКТЫ, СКРЫТЫЕ МАРКОВСКИЕ СЕТИ, СВЕРТОЧНЫЕ НЕЙРОННЫЕ СЕТИ

Объектом классификации являются методы распознавания надводных объектов с аэрофотоснимков.

Цель работы – классификация методов распознавания надводных объектов с аэрофотоснимков. Для достижения данной цели требуется решить следующие задачи.

Для достижения данной цели необходимо решить следующие задачи:

- описать процесс распознавания объекта с изображения;
- изучить современные подходы к решению данного класса задач;
- описать формальную постановку задачи;
- описать существующие методы.

# СОДЕРЖАНИЕ

<b>РЕФЕРАТ</b>	<b>3</b>
<b>ОБОЗНАЧЕНИЯ И СОКРАЩЕНИЯ</b>	<b>5</b>
<b>ВВЕДЕНИЕ</b>	<b>6</b>
<b>1 Общие сведения</b>	<b>7</b>
1.1 Распознавание . . . . .	7
1.2 Марковские цепи с дискретным временем . . . . .	8
1.3 Нейронные сети . . . . .	9
1.4 Сверточные нейронные сети . . . . .	10
1.4.1 Свертка . . . . .	11
1.4.2 Двухэтапные алгоритмы . . . . .	12
1.4.3 Одноэтапные алгоритмы . . . . .	13
<b>2 Постановка задачи</b>	<b>14</b>
<b>3 Существующие методы</b>	<b>15</b>
3.1 Скрытые Марковские модели . . . . .	15
3.2 Метод k-ближайших соседей . . . . .	16
3.3 Нейросетевые алгоритмы . . . . .	16
3.3.1 Двухэтапные CNN . . . . .	16
3.3.2 Одноэтапные CNN . . . . .	17
3.3.3 Параметры для сравнения . . . . .	18
3.3.4 Сравнение . . . . .	19
<b>4 Классификация</b>	<b>22</b>
<b>ЗАКЛЮЧЕНИЕ</b>	<b>23</b>
<b>СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ</b>	<b>26</b>
<b>ПРИЛОЖЕНИЕ А Презентация</b>	<b>27</b>

## ОБОЗНАЧЕНИЯ И СОКРАЩЕНИЯ

В настоящей расчетно-пояснительной записке применяют следующие сокращения и обозначения.

ANN — (англ. Artificial Neural Network) Искусственная нейронная сеть

CNN — (англ. Convolutional Neural Network) Свёрточная нейронная сеть

RoI — (англ. Region Of Interest) Область интереса

HMM — (англ. Hidden Markov Model) Скрытая Марковская модель

kNN — (англ. k-Nearest Neighbors) k-Ближайших соседей

## ВВЕДЕНИЕ

В настоящее время все больше внимания привлекает тема автоматизации судоходства, а так же общее повышение уровня безопасности в процессе эксплуатации судов.

Методы распознавания различных надводных объектов имеют особое значение в данном контексте, так как позволяют решить множество проблем: избежание столкновений судов, осуществление автономного плавания и пр. [1].

Не менее важной является задача распознавания малых надводных объектов, поскольку традиционные методы обнаружения, основанные на использовании радара, не подходят для задачи обнаружения близко расположенных и малых объектов [2].

Кроме того, обнаружение активности рыболовецких судов по-прежнему является сложной задачей для многих стран, расположенных на архипелагах, например – Индонезии. В настоящее время для мониторинга огромной морской акватории используется технология, использующая датчики SAR для обнаружения кораблей, разрабатываемая с 1985 года. Однако стоимость использования данной технологии является одним из препятствий для дальнейшего развития технологии [3].

Целью настоящей работы является классификация методов распознавания надводных объектов с аэрофотоснимков. Для достижения данной цели требуется решить следующие задачи:

- описать процесс распознавания объекта с изображения;
- изучить современные подходы к решению данного класса задач;
- описать формальную постановку задачи;
- описать существующие методы.

# 1 Общие сведения

## 1.1 Распознавание

Стоит отдельно отметить, что в русском языке термин «распознавание» является достаточно объемным и включает в себя, непосредственно, обнаружение, классификацию и кластеризацию распознанных объектов.

В данной работе речь идет о рассматриваемой задаче распознавания объектов на аэрофотоснимке, т. е. изображении. Изображения можно получить в трех основных диапазонах: видимом, инфракрасном и радиоволновом спектрах. При этом, способ представления данных на изображении зависит от конкретной задачи – это может быть трехканальный RGB, одноканальная градация серого и т.д.

Так же, существует несколько подходов к распознаванию объектов на изображении в зависимости от имеющихся данных.

Например, последовательность изображений содержит больше данных, чем единичное изображение. Благодаря этому можно повысить точность распознавания, применяя методы распознавания объектов в последовательности сигналов, учтя, тем самым, контекст происходящего на изображениях последовательности. При этом, последовательность может представлять собой как снимки одного и того же объекта с разных ракурсов, так и содержать различные состояния объекта в результате многократной съемки со стационарной позиции.

К тому же, при наличии карт глубины изображения или же изображения в альтернативном спектре можно комбинировать различные методы, повышая тем самым точность распознавания за счет использования всех источников данных [4].

В любом случае, обнаружение объекта на изображении сводится к выделению области изображения, в которой находится объект. Обычно такие области задаются парами координат  $(x_1, y_1)$  и  $(x_2, y_2)$ , задающие левый нижний и правый верхний углы области соответственно.

Результатом классификации, равно как и кластеризации, является разбиение распознанных объектов на группы. Такие группы принято называть классами. Основным отличием кластеризации является отсутствие необходимости в явном задании классов – они формируются в процессе работы на

основании признаков распознаваемых объектов, в то время как классификация подразумевает разбиение объектов на заранее заданные классы.

В настоящее время основным методом распознавания каких-либо объектов является применение нейронных сетей.

В первую очередь, это связано с основным преимуществом нейронных сетей перед альтернативными методами – отсутствие необходимости в явном задании признаков распознаваемых объектов.

## 1.2 Марковские цепи с дискретным временем

С помощью Марковских цепей можно представлять реальные процессы, являющиеся или приводимые к Марковским процессам. Процесс называется Марковским, если выполняется следующее равенство:

$$\begin{aligned} P\{\xi(t_{n+1}) < x_{n+1} | \xi(t_n) = \\ &= x_n, \dots, \xi(t_1) = x_1\} = \\ &= P\{\xi(t_{n+1}) < x_{n+1} | \xi(t_n) = x_n\}, \end{aligned} \quad (1.1)$$

То есть условная функция распределения вероятностей значений  $\xi(t_{n+1})$  для данного процесса в момент времени  $t_{n+1}$  не зависит от значений процесса в моменты времени  $t_1, \dots, t_{n-1}$ , а определяется лишь значением  $\xi(t_n) = x_n$  в момент времени  $t_n$ .

Для описания цепей Маркова можно использовать граф вероятностей переходов. Вершины данного графа обозначают возможные состояния рассматриваемой системы, а стрелки от одной вершины к другой указывают возможные переходы между состояниями. Число над стрелкой обозначает вероятность такого перехода. Например, пусть множество состояний  $X = \{1, 2, 3\}$ , а матрица вероятностей переходов имеет вид:

$$P = \begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 0 & 1/2 \\ 2/3 & 0 & 1/3 \end{bmatrix}. \quad (1.2)$$

Тогда граф вероятностей переходов следующим образом:

Цепь Маркова полностью определяется матрицей вероятностей переходов и начальным распределением.

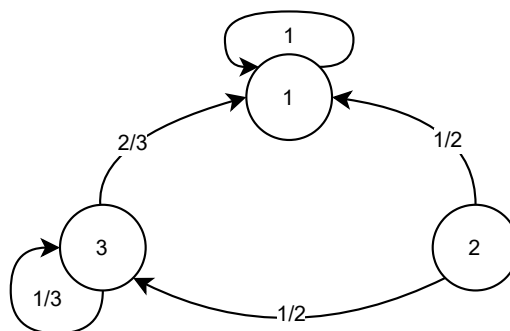


Рисунок 1.1 – Пример графа вероятностей перехода

### 1.3 Нейронные сети

ANN определяется как массово параллельный распределенный процессор, состоящий из простых процессорных блоков, который обладает естественной склонностью накапливать эмпирические знания [5].

Их название и структура вдохновлены человеческим мозгом, а алгоритм работы основывается на способе, которым биологические нейроны передают сигналы друг другу.

Простейшая ANN включает в себя входной слой, выходной (или целевой) слой и, между ними, скрытый слой. Слои соединены через узлы (искусственные нейроны). Эти соединения образуют «сеть» – нейронную сеть – из взаимосвязанных узлов. Пример приведен на рисунке 1.2.

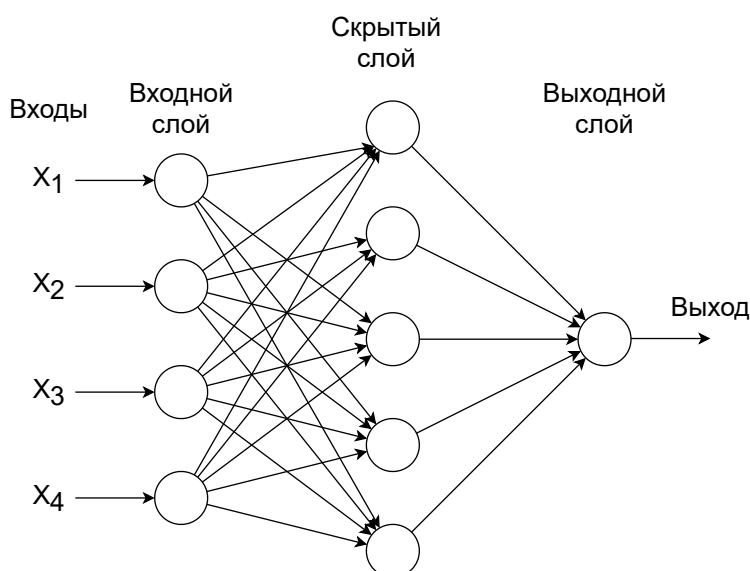


Рисунок 1.2 – Пример схемы ANN

В общем случае искусственный нейрон можно образом, приведенном на рисунке 1.3.



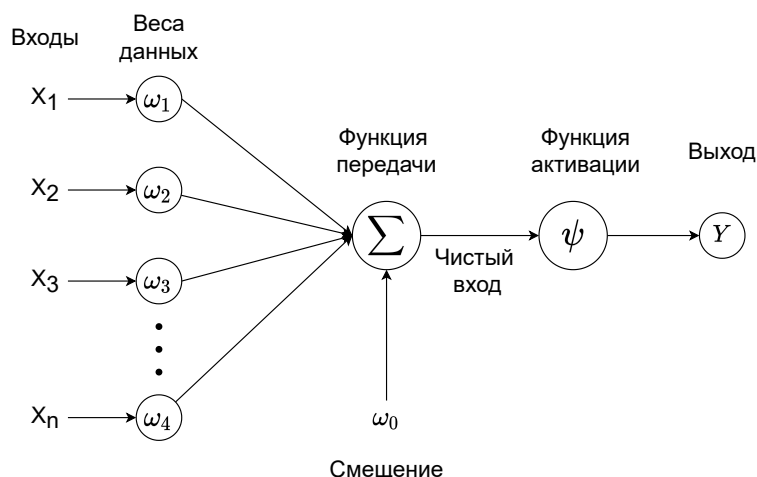


Рисунок 1.3 – Общая схема искусственного нейрона

Данное представление применимо к любому виду нейронных сетей – вне зависимости от типа, нейронные сети реализуются путем упорядочивания нейронов в слои и последующим связыванием соответствующих слоев между собой. [5]

В процессе обучения нейронной сети используется так называемая обучающая выборка – заранее подготовленный набор данных, отражающий суть рассматриваемой предметной области [5]. В зависимости от содержимого обучающей выборки результирующие весовые конфигурации нейронной сети (т. е. веса связей между нейронами, а так же смещения отдельных взятых нейронов) могут отличаться [5]. В связи с этим одна и та же структура нейронной сети может переиспользована для работы в различных предметных областях.

## 1.4 Сверточные нейронные сети

Одним из основных видов нейронных сетей, применяемых для распознавания является CNN [6].

CNN представляет собой тип ANN, которая имеет архитектуру с глубокой обратной связью и выделяется на фоне остальных ANN с полносвязными слоями своей способностью к обобщению. CNN работает с сильно абстрагированными характеристиками объектов, особенно это касается пространственных данных, что позволяет добиться более эффективно идентифицировать объекты в сравнении с другими типами ANN [6]. Одним из отличительных свойств CNN является способность к фильтрации посторонних шумов во входных данных.

Модель CNN состоит из конечного набора уровней обработки, которые могут изучать различные характеристики входных данных (например, изображения) с несколькими уровнями абстракции. Начальные уровни изучают и извлекают высокоуровневые свойства, а более глубокие уровни изучают и извлекают более низкоуровневые свойства. Базовая концептуальная модель CNN была показана на рисунке 1.4.

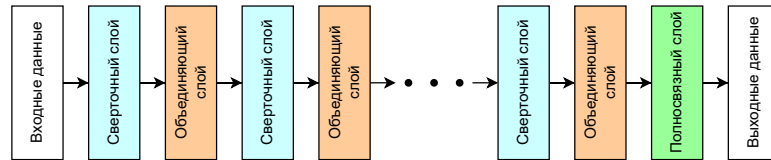


Рисунок 1.4 – Концептуальная модель CNN

Существующие архитектуры CNN для обнаружения объектов на изображениях можно разделить на две категории: одноэтапные (one-stage) и двухэтапные (two-stage) [7].

### 1.4.1 Свертка

#### Ядро

Прежде, чем рассматривать процесс свертки, необходимо определить понятие «ядро», используемое при свертке. Ядро представляет собой сетка из дискретных значений или чисел, где каждое значение известно как вес этого ядра. Пример двумерного ядра приведен на рисунке 1.5.

0	1
-1	2

Рисунок 1.5 – Пример двумерного ядра с размерностью  $2 \times 2$

Ядро инициализируется случайными значениями, которые изменяются в ходе обучения CNN.

#### Пример свертки

Разберем пример свертки для изображения в градациях серого, т.к. такое изображение содержит лишь один канал, передаваемый на вход CNN.

Пусть дано изображение в градациях серого, представленное на рисунке 1.6.

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

Рисунок 1.6 – Пример изображения в градациях серого с размерностью  $4 \times 4$

Далее рассмотрим первые два шага процесса свертки, представленные на рисунках 1.7 и 1.8, соответственно.

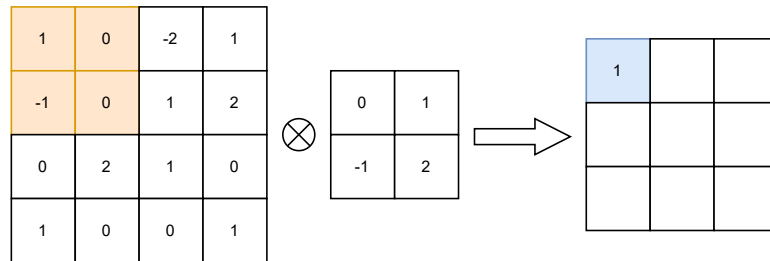


Рисунок 1.7 – Пример изображения в градациях серого с размерностью  $4 \times 4$

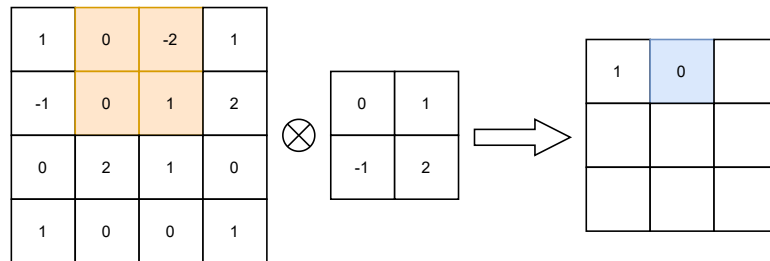


Рисунок 1.8 – Пример изображения в градациях серого с размерностью  $4 \times 4$

Аналогичным образом свертка продолжается до полного заполнения результирующей матрицы. Стоит отметить, что, в зависимости от размера окна и размера перекрытия окон, будет меняться размер результирующей матрицы.

### 1.4.2 Двухэтапные алгоритмы

В таких нейросетевых алгоритмах выделяют два этапа: поиска RoI (англ. Candidate Region Extraction) на изображении и последующей классификации RoI, найденных на первом этапе. При этом под RoI на изображении подразумеваются зоны, потенциально содержащие искомые объекты [8].

Стоит отметить, что первый этап может происходить без использования нейронных сетей. Для этого можно использовать информацию о контрасте,

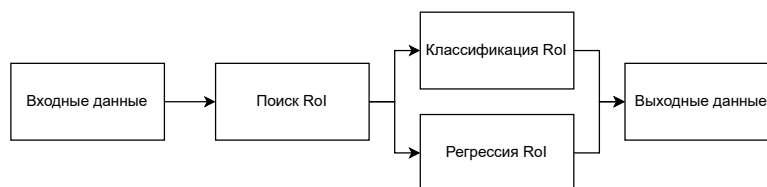


Рисунок 1.9 – Схема работы двухэтапного алгоритма

ключевые точки или перебор всех возможных положений объекта с помощью процедуры **selective search** [9].

RoI, полученные вышеперечисленными методами, могут обладать серьезными недостатками, например:

- содержать слишком большое количество фона;
- содержать лишь небольшую часть объекта;
- содержать более одного объекта.

В связи с этим на первом этапе более предпочтительным методом является применение CNN, не содержащих полносвязных слоев [9].

На втором этапе CNN применяются к обнаруженным RoI.

Преимуществом данных алгоритмов является высокая точность распознавания объектов, однако, платой за это является время, необходимое для выделения «подозрительных» зон на изображении [8].

### 1.4.3 Одноэтапные алгоритмы

Данные нейросетевые алгоритмы не включают в себя этап поиска RoI на изображении.

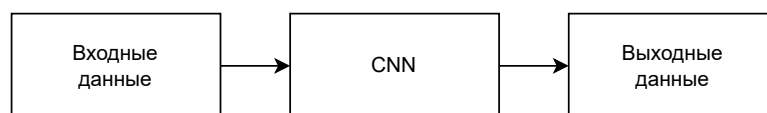


Рисунок 1.10 – Схема работы одноэтапного алгоритма

Преимуществами одноэтапных алгоритмов являются их простота и относительно высокая скорость работы. К недостаткам же можно отнести более низкую точность детектирования объектов по сравнению с двухэтапными алгоритмами, а также меньшую гибкость алгоритма с точки зрения рассматриваемых изображений [7].

## 2 Постановка задачи

Задача заключается в обнаружении надводного объекта на изображении по заранее заданному множеству критериев и его классификацию согласно предварительно выделенным классам.

Метод распознавания надводных объектов с аэрофотоснимков соотносит входные данные – единичное изображение или же последовательность изображений, с классом объектов из заранее определенного множества.

В задаче можно выделить два этапа: предварительная обработка изображения и классификация.

## 3 Существующие методы

### 3.1 Скрытые Марковские модели

Одним из статических методов распознавания объектов с изображений являются скрытые Марковские модели с дискретным временем. НММ – это вероятностная модель, построенная на цепях Маркова, в частности, цепях Маркова с дискретным временем [10].

НММ используют статистические свойства сигналов и учитывают непосредственно их пространственные характеристики, поэтому они широко применяются в задачах распознавания. Входным сигналом для НММ служит последовательность  $O = o_1, o_2, \dots, o_T$ , где  $o_i$  - сигнал, наблюдаемый в момент времени  $i = \overline{1, T}$  [10].

На рисунке 3.1 приведены сам процесс (модель)  $x(\tau)$  и наблюдаемые неизвестные параметры  $y(\tau)$ .

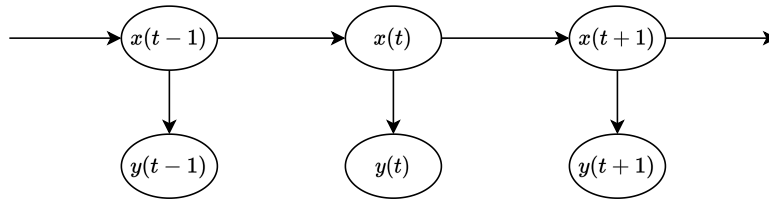


Рисунок 3.1 – Скрытая Марковская модель

Для задания НММ требуется обеспечить процесс перехода между состояниями, связи между процессом  $x(\tau)$  и наблюдаемыми неизвестными параметрами  $y(\tau)$ , а также знать исходное распределение состояний  $p_0(x)$  [11].

Последовательность наблюдений  $O$  может извлекаться из изображения различными способами. В силу этого описательные способности полученных моделей могут различаться. Обычно прибегают к методу сканирования изображения прямоугольным окном, причем размерность перекрытия окон при сканировании подбирается экспериментально. Затем, для уменьшения пространства признаков каждый извлеченный блок проходит некоторую обработку. Для задачи распознавания с изображений чаще всего используются преобразование Карунена-Лоэва (англ. KLT) или дискретно-косинусное преобразование (англ. DCT) [12].

Несмотря на то, что НММ может применяться в качестве самостоятельного классификатора, сложность задания моделируемого процесса все

еще остается существенной проблемой. Для ее решения, можно использовать CNN, предварительно пропустив входное изображение через нее. Это позволит выделять векторы признаков с меньшим пространством [13].

## 3.2 Метод $k$ -ближайших соседей

$k$ NN – это один из самых простых, но эффективных методов машинного обучения, чаще всего использующийся в задачах классификация, несмотря на то, что данный метод подходит и для задач регрессии. В  $k$ NN объект классифицируется на основе оценки  $k$ -ближайших соседей в  $n$ -мерном пространстве известных образцов [14].

Пример работы данного алгоритма для двухмерного случая приведен на рисунке 3.2.

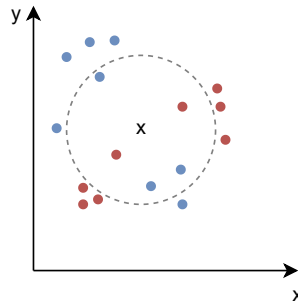


Рисунок 3.2 – Пример использования  $k$ NN для двухмерного случая с  $k = 5$

Существует множество оптимизаций, призванных улучшить данный метод. Однако у  $k$ NN остаются его наиболее сложно решаемые проблемы – определение значения  $K$ , размер памяти, необходимый при работе с большими объемами данных, а так же выбор подходящего способа определения расстояния между объектами в конкретной задаче [15].

## 3.3 Нейросетевые алгоритмы

### 3.3.1 Двухэтапные CNN

#### R-CNN

Первой CNN, разработанной для обнаружения объектов, является модель Region-based CNN, которая использует подходы на основе скользящего окна (sliding window) для обнаружения объектов [16].

Здесь авторы разделяют всю задачу на три модуля. В первом модуле из каждого входного изображения извлекаются RoI, которые могут содержать какой-либо объект (с помощью процедуры selective search), затем во втором модуле авторы используют аффинное искажение изображения, чтобы сделать все извлеченные RoI фиксированного размера (или фиксированного соотношения сторон), а затем пропускают эти искаженные RoI через AlexNet CNN для извлечения конечных признаков (векторов признаков фиксированного размера). Наконец, третий этап представляет собой набор линейных SVM (support vector machine), которые причисляют каждый вектор какому-либо классу и отдельный регрессор обрамляющих окон [16].

Последующие версии алгоритма – Fast R-CNN [17] и Faster R-CNN [18] – призваны оптимизировать время работы алгоритма, а так же повысить точность распознавания. Так, в Faster R-CNN процедура selective search была заменена на сеть предложений регионов (RPN). RPN – это CNN, используемая для создания высококачественных RoI.

### 3.3.2 Одноэтапные CNN

#### YOLO

Алгоритм YOLO (You Only Look Once) является одноэтапным и может непосредственно распознавать объекты, а также их местоположение с помощью сквозной обученной модели CNN [19].

В алгоритме YOLO входное изображение разбивается на фиксированное число сеток, а затем из каждой сетки предсказывается фиксированное число местоположений обрамляющих окон (bounding boxes) и вероятностей. Затем используется пороговое значение для выбора и определения местоположения объекта на изображении [19].

Основной проблемой YOLO является более низкая точность при распознавании больших и малых объектов, а так же присущая всем одноэтапным алгоритмам потеря точности распознавания в сравнении с двухэтапными алгоритмами [19].

Данное семейство алгоритмов включает в себя множество оптимизаций оригинального YOLO, среди которых:

- YOLOv2 (YOLO9000) [20];



- YOLOv3 [21];
- YOLOv3 tiny [22].

Последующие варианты алгоритма имеют малозначительные изменения, призванные улучшить его в конкретных задачах [23]. Последней опубликованной версией является YOLOv7 [24].

## SSD

Алгоритм SSD (Single Shot MultiBox Detector) является альтернативным взглядом на YOLO – от последнего взяты наиболее удачные, с точки зрения авторов, решения, а так же привнесены свои идеи для повышения mAP и снижения времени работы алгоритма [25].

Отличительной особенностью является распознавание объектов за один прогон с помощью заданной сетки окон (default box) на пирамиде изображений, закодированной в сверточных тензорах. Таким образом алгоритм распознает как большие, так и маленькие объекты за один прогон сети [25].

Второй версией алгоритма является MobileSSD, который является комбинацией двух нейросетевых алгоритмов – MobileSSD и SSD [26].

### 3.3.3 Параметры для сравнения

Для оценки работы нейронной сети используется величина AP (Average Precision), вычисляемую по следующей формуле:

$$AP = \frac{\text{кол-во верно распознанных объектов}}{\text{общее кол-во распознанных объектов}}. \quad (3.1)$$

Однако, AP вычисляется для объектов одного класса, в связи с чем для классификаторов с несколькими возможными классами объектов используется величина mAP (англ. mean AP) – среднее значение AP.

Кроме того, данную величину можно оценивать в зависимости от IoU (Intersection over Union), вычисляемую по формуле:

$$IoU = \frac{\text{площадь пересечения областей}}{\text{площадь объединения областей}}. \quad (3.2)$$

IoU описывает то, насколько предсказанные CNN обрамляющие окна близки к «истинным». Данная величина принимает значения в диапазоне

$[0; 1]$ , соответственно.

### 3.3.4 Сравнение

В таблицах 3.1 и 3.2 приведены результаты сравнения работы рассмотренных выше нейросетевых алгоритмов [21].

Результаты, приведенные в таблице 3.1, получены посредством бенчмарка COCO [27] в [21].

Таблица 3.1 – Сравнение существующих методов распознавания объектов

	$AP_{IoU}$			$AP_{object\ size}$		
	$AP$	$AP_{50}$	$AP_{75}$	$AP_S$	$AP_M$	$AP_L$
<i>Двухэтапные</i>						
Faster R-CNN+++	34.9	55.7	37.4	15.6	38.7	50.9
Faster R-CNN w FPN	36.2	59.1	39.0	18.2	39.0	48.2
Faster R-CNN by G-RMI	34.7	55.5	36.7	13.5	38.1	52.0
Faster R-CNN w TDM	36.8	57.7	39.2	16.2	39.8	52.1
<i>Одноэтапные</i>						
YOLOv2	21.6	44.0	19.2	5.0	22.4	35.5
SSD513	31.2	50.4	33.3	10.2	34.5	49.8
DSSD513	33.2	53.3	35.2	13.0	35.4	51.1
YOLOv3 608 × 608	33.0	57.9	34.4	18.3	35.4	41.9

В таблице 3.2 приведен результат сравнение проводившегося на основе набора данных PASCAL VOC 2007 и 2012 [28], проученный в [21] и [25].

Таблица 3.2 – Сравнение существующих методов

Метод	Критерии сравнения				
	mAP	Кол-во слоев	Кол-во параметров	FPS	Разрешение
<i>Двухэтапные алгоритмы</i>					
Fast R-CNN	70.0	—	—	0.5	—
Faster R-CNN VGG-16	73.2	—	—	7	—
Faster R-CNN ResNet	76.4	—	—	5	—
<i>Одноэтапные алгоритмы</i>					
YOLO	63.4	—	—	45	448 × 448
SSD300	74.3	—	—	46	300 × 300
SSD512	76.8	—	—	19	512 × 512
YOLOv2 416 × 416	76.8	65	6.70452e+07	67	416 × 416
YOLOv3	79.2	222	6.15507e+07	70	416 × 416
YOLOv3 tiny	61.4	37	8.67824e+06	188	416 × 416

## Вывод

### Скрытые Марковские модели

Данный метод позволяет классифицировать объекты на изображении, основываясь на заранее заданном моделируемом процессе, используя для этого последовательности изображений.

Необходимость в явном задании процесса и предобработке изображений с целью выделения признаков – основные проблемы данного метода. При этом, размерность пространства признаков можно уменьшить используя определенные преобразования.

Один из способов упрощения моделируемого процесса – использование CNN в качестве блока предобработки изображений, что позволит избавиться от необходимости явной разработки процесса.

### Метод $k$ -ближайших соседей

Данный метод является простейшим классификатором, позволяющим классифицировать объект на изображении согласно заранее заданному множеству классов.

Основные проблемы метода:

- необходимость в явном задании  $n$ -мерного пространства признаков;
- необходимость в явном задании функции расстояния в  $n$ -мерном пространстве признаков;
- эмпирический подход к нахождению значения  $K$ ;
- ошибки классификации «выбывающих» из класса объектов;
- линейная зависимость времени работы от количества объектов в худшем случае;
- линейная зависимость объема оперативной памяти, необходимой в процессе работы, от количества объектов.

Тем не менее, для относительно малого числа признаков и объектов, при условии изолированных классов, данный метод может быть использован в виду относительной простоты реализации.

## Нейросетевые алгоритмы

Нейросетевые алгоритмы выделяются на фоне альтернативных методов своей способностью к самостоятельному выделению признаков на основе обучающей выборки и самостоятельному выделению кластеров объектов на основе полученных признаков.

Кроме того, при распознавании объектов на изображении нейросетевые алгоритмы нуждаются в минимальной предобработке, заключающейся в аффинных преобразованиях входного изображения.

К преимуществам двух- и одноэтапных CNN относятся:

- способность к фильтрации шумов на входном изображении;
- малое число настраиваемых весов в сравнении с полносвязными ANN;
- константное время работы и объем оперативной памяти для конкретной CNN.

При этом современные одноэтапные алгоритмы практически не уступают двухэтапным с точки зрения точности распознавания (для YOLOv3 mAP=79.2, для Faster R-CNN – 76.4), существенно превосходя их с точки зрения времени работы (YOLOv3 tiny FPS = 188, Faster R-CNN – 5). Это позволяет использовать одноэтапные CNN в задачах распознавания в реальном времени.

Тем не менее, у CNN слишком большое число варьируемых параметров: количество слоёв, размерность ядра свёртки для каждого из слоёв, количество ядер для каждого из слоёв, шаг сдвига ядра при обработке слоя и т.д.

Параметры сети существенно влияют на результаты работы, однако в настоящее время выбираются эмпирически. Существует несколько выверенных и отлаженных CNN, но отсутствуют рекомендации, согласно которым нужно проектировать нейросетевой алгоритм для новой задачи.

Так же, недостатком CNN, как и остальных нейросетевых алгоритмов является варьирование результатов обучения, связанная с самой структурой алгоритмов.

## 4 Классификация

Рассмотренные выше методы распознавания надводных объектов можно классифицировать по ряду признаков, приведенных на рисунке 4.1.

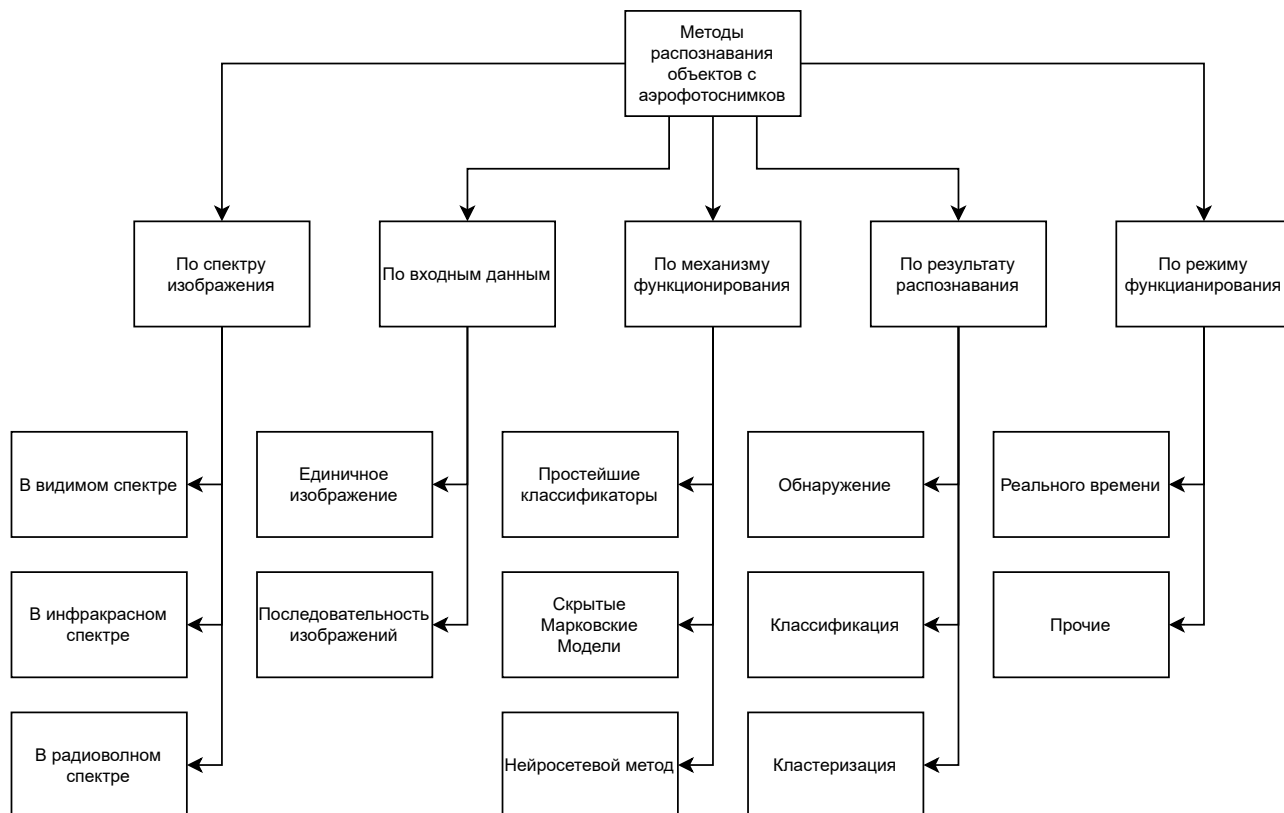


Рисунок 4.1 – Классификация методов распознавания объектов с аэрофотоснимков

## ЗАКЛЮЧЕНИЕ

В данной работе была проведена классификация существующих методов распознавания надводных объектов с аэрофотоснимков. Результат классификации приведен на рисунке 4.1.

Точность распознавания объекта зависит от множества факторов помимо выбранного метода распознавания, в первую очередь – от зашумленности входных данных.

Нейросетевые алгоритмы выделяются на фоне альтернативных методов более высокой точностью распознавания, а так же отсутствием необходимости в явном задании критериев распознавания. Наиболее популярным видом ANN для распознавания являются CNN, т.к. данные нейросетевые алгоритмы обладают свойством фильтрации шумов, что особенно актуально в данной задаче.

## СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. *Lee S.-J., Roh M.-I., Oh M.-j.* Image-based ship detection using deep learning // Ocean Systems Engineering. — 2020. — Т. 10. — С. 415–434. — DOI: 10.12989/ose.2020.10.4.415.
2. *Z. Chen* Deep learning for autonomous ship-oriented small ship detection / Z. Chen [и др.] // Safety Science. — 2020. — Т. 130. — С. 104812.
3. *M. Marzuki* Fishing boat detection using Sentinel-1 validated with VIIRS Data / M. Marzuki [и др.] // IOP Conference Series: Earth and Environmental Science. — 2021. — Т. 925. — С. 012058. — DOI: 10.1088/1755-1315/925/1/012058.
4. *Rokach, Lior.* Ensemble Methods for Classifiers // The Data Mining and Knowledge Discovery Handbook. — 2005. — С. 957–980. — DOI: 10.1007/0-387-25465-X\_45.
5. *Sharkawy A.-N.* Principle of Neural Network and Its Main Types: Review // Journal of Advances in Applied & Computational Mathematics. — 2020. — Т. 7. — С. 8–19. — DOI: 10.15377/2409-5761.2020.07.2.
6. *A. Ghosh* Fundamental Concepts of Convolutional Neural Network / A. Ghosh [и др.]. — 2020. — DOI: 10.1007/978-3-030-32644-9\_36.
7. *Zhang H., Cloutier R.* Review on One-Stage Object Detection Based on Deep Learning // EAI Endorsed Transactions on e-Learning. — 2022. — Т. 7. — С. 174181. — DOI: 10.4108/eai.9-6-2022.174181.
8. *Du L., Zhang R., Wang X.* Overview of two-stage object detection algorithms // Journal of Physics: Conference Series. — 2020. — Т. 1544. — С. 012033. — DOI: 10.1088/1742-6596/1544/1/012033.
9. *А.В. Бондаренко* АЛГОРИТМ НЕЙРОСЕТЕВОГО РАСПОЗНАВАНИЯ НАДВОДНЫХ ОБЪЕКТОВ В РЕАЛЬНОМ ВРЕМЕНИ / А. В. Бондаренко [и др.] // Известия Тульского государственного университета. Технические науки. — 2021. — Т. 1. — С. 19–33.
10. *L. Carlson* Early Classification of Intent for Maritime Domains Using Multinomial Hidden Markov Models / L. Carlson [и др.] // Frontiers in Artificial Intelligence. — 2021. — Т. 4. — С. 702153. — DOI: 10.3389/frai.2021.702153.

11. *S. Zouaoui-Elloumi* Modeling Ship Behavior based on Hidden Markov Models / S. Zouaoui-Elloumi [и др.] // Proceedings of Signal Processing, Pattern Recognition, and Applications (SPPRA 2011). — 2011. — DOI: 10.2316/P.2011.721-107.
12. *Гультяева Т., Попов А.* ПРИМЕНЕНИЕ СКРЫТЫХ МАРКОВСКИХ МОДЕЛЕЙ С ОДНОМЕРНОЙ ТОПОЛОГИЕЙ К ЗАДАЧЕ РАСПОЗНАВАНИЯ ЛИЦ // Информатика и проблема телекоммуникаций. — 2006. — Т. 1. — С. 150—154.
13. *Tur A., Keles H.* Evaluation of hidden Markov models using deep CNN features in isolated sign recognition // Multimedia Tools and Applications. — 2021. — Т. 80. — DOI: 10.1007/s11042-021-10593-w.
14. *Zhang Z.* Introduction to machine learning: K-nearest neighbors // Annals of Translational Medicine. — 2016. — Т. 4. — С. 218—218. — DOI: 10.21037/atm.2016.03.37.
15. *K. Taunk* A Brief Review of Nearest Neighbor Algorithm for Learning and Classification / K. Taunk [и др.] // 2019 International Conference on Intelligent Computing and Control Systems (ICCS). — 2019. — С. 1255—1260. — DOI: 10.1109/ICCS45141.2019.9065747.
16. *R. Girshick* Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation / R. Girshick [и др.] // Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. — 2013. — DOI: 10.1109/CVPR.2014.81.
17. *Girshick R.* Fast R-CNN. — 2015. — DOI: 10.1109/ICCV.2015.169.
18. *S. Ren* Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks / S. Ren [и др.] // IEEE Transactions on Pattern Analysis and Machine Intelligence. — 2015. — Т. 39. — DOI: 10.1109/TPAMI.2016.2577031.
19. *P. Jiang* A Review of Yolo Algorithm Developments / P. Jiang [и др.] // Procedia Computer Science. — 2022. — Т. 199. — С. 1066—1073. — ISSN 1877-0509. — DOI: <https://doi.org/10.1016/j.procs.2022.01.135>.



20. *Redmon J., Farhadi A.* YOLO9000: Better, Faster, Stronger // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). — 2017. — С. 6517—6525. — DOI: 10.1109/CVPR.2017.690.
21. *Redmon J., Farhadi A.* YOLOv3: An Incremental Improvement. — 2018.
22. *D. Li* Yolo-tiny-MSD: A tiny neural network for object detection / D. Li [и др.] // Journal of Physics: Conference Series. — 2021. — Т. 1873. — С. 012073. — DOI: 10.1088/1742-6596/1873/1/012073.
23. *L. Hao* Enhanced YOLO v3 Tiny Network for Real-time Ship Detection from Visual Image / L. Hao [и др.] // IEEE Access. — 2021. — Т. PP. — С. 1—1. — DOI: 10.1109/ACCESS.2021.3053956.
24. *Wang C.-Y., Bochkovskiy A., Liao H.-y.* YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. — 2022. — DOI: 10.48550/arXiv.2207.02696.
25. *W. Liu* SSD: Single Shot MultiBox Detector / W. Liu [и др.]. — 2016. — DOI: 10.1007/978-3-319-46448-0\_2.
26. *A. Younis* Real-Time Object Detection Using Pre-Trained Deep Learning Models MobileNet-SSD / A. Younis [и др.] // International Journal for Research in Applied Science & Engineering Technology (IJRASET). — 2020. — С. 44—48. — DOI: 10.1145/3379247.3379264.
27. COCO - Common Objects in Context [Электронный ресурс]. — Режим доступа: <https://cocodataset.org/> (Дата обращения: 20.11.2022).
28. PASCAL VOC Dataset | Papers With Code [Электронный ресурс]. — Режим доступа: <http://host.robots.ox.ac.uk/pascal/VOC/> (Дата обращения: 20.11.2022).

# ПРИЛОЖЕНИЕ А

## Презентация