

Министерство образования и науки Российской Федерации
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
Институт прикладной математики и компьютерных наук
Кафедра теоретических основ информатики

ДОПУСТИТЬ К ЗАЩИТЕ В ГЭК

Руководитель ООП

канд. техн. наук, доцент

 А. Л. Фукс

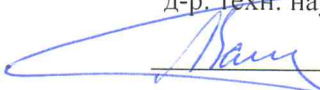
« 11 » 06 2018 г.

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА БАКАЛАВРА
РЕКОНФИГУРИРУЕМАЯ ПРОГРАММНАЯ СИСТЕМА ДЛЯ РАСПОЗНАВАНИЯ ПАТОЛОГИЙ
ПО ВИДЕОДАНЫМ КОЛОНОСКОПИИ
по основной образовательной программе подготовки бакалавров
направление подготовки 02.03.02 Фундаментальная информатика и информационные технологии

Никитенко Александр Андреевич

Руководитель ВКР

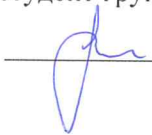
д-р. техн. наук, доцент

 А. В. Замятин

« _____ » _____ 2018 г.

Автор работы

студент группы № 1441

 А. А. Никитенко

Томск-2018

Реферат

Выпускная работа 47 с., 30 рис., 11 табл.

IMAGE RECOGNITION, POLYP DETECTION, DEEP LEARNING, CONVOLUTIONAL NEURAL NETWORK, SEGMENTATION, PYTHON, OPEN CV, РАСПОЗНОВАНИЕ

Объект исследования – кадры видеозаписей колоноскопии.

Цель работы – разработка реконфигурируемой системы для распознавания патологий на видео колоноскопии.

Методы работы – теоретические, практические.

Результаты работы – построена система автоматического распознавания полипов.

Оглавление

Введение	4
Глава 1. Обзор предметной области	5
1.1 Методы решения задачи распознавания изображений.....	5
1.2 Алгоритмы машинного обучения	7
1.2.1 Предобработка	7
1.2.2 Выделение областей интереса	8
1.2.3 Проектирование признаков.....	8
1.2.3.1 Гистограмма ориентированных градиентов	9
1.2.3.2 Локальные бинарные шаблоны.....	11
1.2.4 Классификация.....	13
1.2.5 Сверточные нейронные сети.....	14
1.2.5.1 Слои сверточной нейронной сети.....	16
1.2.5.2 Функции активации и обучение	17
1.2.5.3 Построение архитектуры сверточной нейронной сети	19
1.2.6 Ансамбли сверточных нейронных сетей	20
1.2.7 Сверточные нейронные сети основанные на регионах (RCNN)	21
1.2.7.1 Селективный поиск	22
1.3 Существующие подходы к распознаванию полипов.....	23
1.4 Метрики оценки качества	24
Глава 2. Разработка алгоритма распознавания.	26
2.1 Описание исходных данных	26
2.2 Анализ исходных данных	26
2.3 Разработка алгоритмов распознавания	27
2.3.1 Генерация данных для классификации.....	30
2.3.1.1 Несбалансированность выборки	31
2.3.1.2 Извлечение сложных отрицательных примеров	32
2.3.1.3 Аугментация	33
2.3.2 Построение классификатора	34
Глава 3. Разработка системы распознавания и проведение экспериментов.....	35
3.1 Проектирование.....	35
3.2 Реализация	38
3.2.1 Выбор технологий	38
3.2.2 Интрфейс пользователя.....	42
3.3 Экспериментальные исследования.....	41
Заключение.....	45
Литература	46

Введение

Полипы – это избыточное разрастание слизистой оболочки внутри толстой кишки. От 10 до 30% из них могут перерождаться в злокачественные опухоли. По причине поздней диагностики летальность достигает 40% в течение года с момента выявления болезни. Почти в 50% случаев колоректального рака, у больных обнаруживаются метастазы в печени.

Самым информативным методом обследования является колоноскопия. Этот метод позволяет осмотреть толстую кишку на всем ее протяжении.

Представленная статистика и возможность сокращения медицинских ошибок с помощью запоминания особенностей полипов, позволяющих их детектировать, показывает важность разработки и необходимость использования новых программных методов и инструментов, классифицирующих аномальные разрастания в толстом кишечнике.

Целью работы является разработка реконфигурируемой системы для распознавания патологий по видеоданным колоноскопии.

Для достижения данной цели были поставлены следующие задачи:

1. Изучение методов распознавания изображений
2. Изучение существующих подходов к распознаванию полипов
3. Разработка системы для тестирования, сравнения и использования методов распознавания на реальных данных

1. Обзор предметной области

Цель распознавания – нахождение предметов реального мира на изображениях и видеоданных. Результатом распознавания, в зависимости от задачи, могут являться метки, определяющие наличие объекта на изображении - при решении задачи классификации; координаты прямоугольника, охватывающего объект - при решении задачи обнаружения, либо пиксели, принадлежащие объекту - при решении задачи семантической сегментации.

1.1 Методы решения задачи распознавания изображений

Задачи распознавания изображений и видеозаписей характеризуются рядом сложностей. К ним относятся: внешние факторы, связанные с тем, что изображения одних и тех же объектов могут сильно отличаться в разных условиях и зависеть от точки наблюдения, освещения, масштаба, перекрытия, сложности фона и, внутренние – внутриклассовая изменчивость и деформация (различные формы объекта).

В зависимости от вышеперечисленных факторов, для решения задач распознавания объектов на изображении применяются различные группы методов.

Метод цветowych/яркостных фильтров можно применять в случаях, когда объект существенно отличается от фона, а освещение равномерно и не изменяется. Отделив такие объекты от фона, можно воспользоваться анализом геометрических признаков для принятия решений о классификации.

Если объект на фоне существенным образом не выделяется или имеет сложную окраску, то применение метода цветowych/яркостных фильтров не даст хороших результатов. В данном случае более эффективен метод выделения и анализа контуров. Для этого мы выделяем на изображении границы, то есть места резкого изменения градиента яркости, например, с помощью метода Canny. Затем проверяем выделенные границы на соответствие геометрическим контурам объекта, это можно сделать, применив

метод Хафа (Hough Transform), для поиска в границах окружностей, эллипсов, линий или применить методы контурного анализа.

Если изображение имеет множество мелких деталей, то анализ контуров может быть затруднён. В этом случае можно применить метод сопоставления с шаблоном. Данный метод применяется для поиска участков изображений, которые наиболее схожи с некоторым заданным шаблоном – конкретным объектом одного класса.

Метод сопоставления с шаблоном находит точные совпадения точек шаблона с точками изображения. Если же изображение повёрнуто или масштабировано относительно параметров шаблона, то данный метод работает плохо. Для преодоления этих ограничений применяют методы, основанные на особых точках. Особая точка (key point) это небольшая область, которая существенным образом выделяется на изображении. Существует несколько методов определения таких точек, это могут быть углы (Harris corner detector) или блобы (blob), то есть небольшие области одинаковой яркости, с достаточно чёткой границей выделяющиеся на общем фоне. Для особой точки вычисляют дескриптор - характеристику особой точки. Дескриптор вычисляют по заданной окрестности особой точки, как направления градиентов яркости разных частей этой окрестности. Существует несколько методов расчёта дескрипторов для особых точек: SIFT, SURF, ORB и др. Особые точки можно применять для поиска объекта на изображении. Для этого нам необходимо взять изображение искомого объекта и выполнить следующие действия:

- 1) На картинке с объектом найти особые точки объекта и вычислить их дескрипторы
- 2) На анализируемом изображении найти особые точки и вычислить их дескрипторы

- 3) Сравнить дескрипторы особых точек объекта и дескрипторы особых точек, найденных на изображении
- 4) Если найдено достаточное количество соответствий, то пометить область с соответствующими точками.

Методы детектирования движения объектов – выделение движущихся объектов на основании нескольких изображений или кадров видеозаписи одной и той же сцены.

Область применения перечисленных выше методов ограничена по причине того, что все они имеют плохую обобщающую способность. Это значит, что модель описывает один конкретный объект или очень узкий класс объектов. Для того, чтобы успешно распознавать не один объект, а большие и разнообразные классы объектов, можно применять методы машинного обучения.

Глубокое обучение - это совокупность методов машинного обучения, основанных на обучении представлением (feature/representation learning). В контексте распознавания изображений основным алгоритмом глубокого обучения является сверточная нейронная сеть.

1.2 Алгоритмы машинного обучения

В общем случае методы распознавания на основе машинного обучения состоят из этапов предобработки, выделения областей интереса, извлечения признаков и классификации.

1.2.1 Предобработка

На данном этапе могут использоваться методы для создания условий, повышающих эффективность и качество выделения и распознавания объектов.

1.2.2 Выделение областей интереса

Самый простой способ решить задачу детектирования – свести её к задаче классификации. Чтобы обнаружить объект, необходимо взять специфический для него классификатор и применить его к участкам изображения.

Одним из способов извлечения таких участков является метод скользящего окна, представляющий собой прямоугольную область с фиксированной шириной и высотой, которая «скользит» по изображению. Для каждого из участков применяется классификатор, чтобы определить, имеет ли окно искомый объект.

«Пирамида изображения» представляет собой многомасштабное представление изображения. В сочетании с скользящим окном мы можем находить объекты разных размеров в различных местах. В нижней части пирамиды мы имеем исходное изображение с первоначальным размером. На каждом последующем слое изображение изменяется и сглаживается (обычно с помощью гауссовского размытия). Изображение уменьшается до тех пор, пока не будет выполнен некоторый критерий остановки.

1.2.3 Проектирование признаков

На входном изображении слишком много несущественной информации, которая не требуется для классификации. Поэтому первым шагом в классификации изображений является упрощение изображения путем извлечения важной информации, содержащейся в изображении, и исключения остального.

Каждый пиксел представляется числом, в случае чёрно-белого изображения, или тройкой чисел, в случае цветного. В данном формате удобно воспринимать изображение, но для решения задачи распознавания, такого количества информации оказывается слишком много.

Способ описания изображения первичными признаками (то есть интенсивностью яркости) имеет максимальную размерность и неинвариантен к повороту и масштабу, поэтому модели, сопоставляющие эти описания, наименее эффективны в вычислительном плане. Более перспективным считается подход к пониманию изображений путем анализа их вторичного описания.

В связи с многообразием классов зрительных образов и целей распознавания, число способов вторичного описания изображений практически неограниченно. Это приводит к тому, что на пути создания надежных и быстродействующих алгоритмов возникает проблема выбора вторичного описания изображений наименьшей размерности, с возможно большей информативностью в условиях действия мешающих факторов. В общей теории формализованного определения таких описаний на сегодняшний день не существует, поэтому разработчик обычно базируется на опыте и интуиции.

Генеральным критерием при выборе вторичных признаков является существование этих признаков у всех классов различаемых объектов.

При обработке изображений наиболее универсальными характерными признаками являются: форма, яркость (или цвет), текстура и др. Менее различимые для восприятия человеком - это спектральные, гистограммные, корреляционные и другие характеристики.

1.2.3.1 Гистограмма ориентированных градиентов

Одним из популярных признаков является гистограмма ориентированных градиентов [4], которая изначально применялась для распознавания пешеходов, однако впоследствии получила большое распространение в других областях, показывая высокий показатель распознаваемости.

Основной идеей алгоритма является допущение, что внешний вид и форма объекта на участке изображения могут быть описаны распределением градиентов интенсивности.

Алгоритм применяется для одноканального изображения, которое обычно получают путем усреднения трехканального. Во избежание потери значимых признаков, данный метод будет применяться для различных каналов.

Алгоритм работы метода следующий:

1) Вычисление градиента. На первом шаге рассчитываются значения градиентов. Самым распространенным методом является применение одномерной дифференцирующей маски в горизонтальном и/или вертикальном направлении. Этот метод требует фильтрации цветовой или яркостной составляющей при помощи фильтрующих ядер:

-1	0	1
-1	0	1

Рисунок 1 - Фильтрующие ядра

Первое ядро находит резкие изменения интенсивности по вертикальным линиям, а второе - по горизонтальным. Ни один из них не срабатывает, если область не имеет сильных изменений в яркости.

Затем мы можем найти величину и направление градиента, используя следующую формулу:

$$g = \sqrt{g_x^2 + g_y^2}$$
$$\theta = \arctan \frac{g_y}{g_x}$$

В каждом пикселе градиент имеет величину и направление.

2) Разделение изображения на ячейки определенной размерности

3) Вычисление гистограмм в ячейках. Каждый пиксел в ячейке участвует во взвешенном голосовании для каналов гистограммы направлений, основанном на значении градиентов. Ячейки могут быть прямоугольной или круглой формы, каналы гистограммы равномерно распределяются от 0 до 180

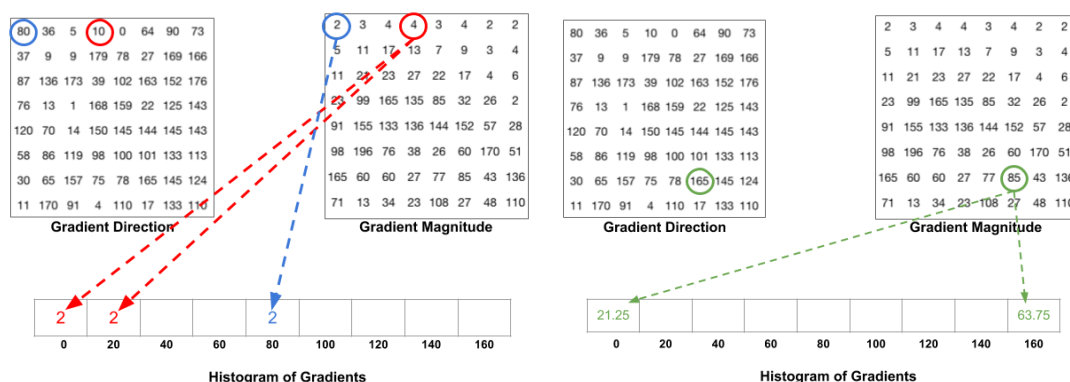


Рисунок 2- Пример вычисления гистограммы

4) Нормализация блоков. Для инвариантности к изменению яркости гистограммы можно нормализовать. Обычно нормализуются связанные блоки гистограмм определенного размера.

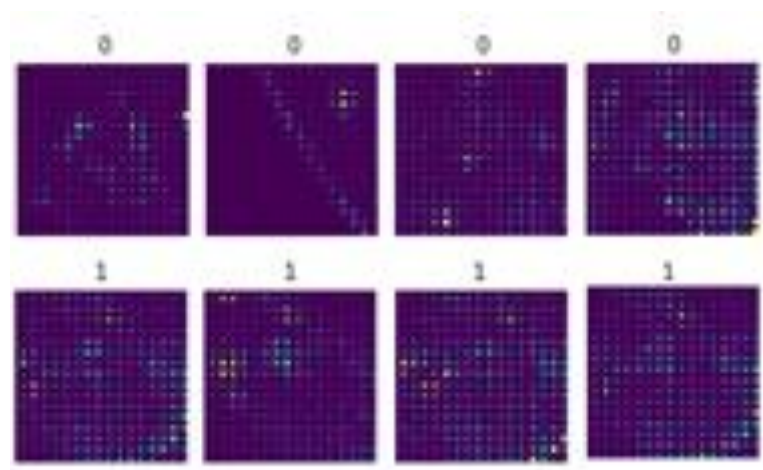


Рисунок 3 - Визуализация гистограммы ориентированных градиентов

5) Объединение блоков в вектор признаков. Таким образом, дескриптор HOG, является вектором нормированных гистограмм ячеек из всех областей блока.

1.2.3.2 Локальные бинарные шаблоны

Другим признаком для описания участков изображений был взят локальный бинарный шаблон, используемый для классификации текстур. Он является одним из часто применяемых, и используется не только для классификации текстур, но и в других задачах – в распознавании биомедицинских изображениях (в т.ч. маммографии) аэрофотосъемке, распознавании лиц.

Данный метод сильно уменьшает размерность вектора признаков и имеет высокую скорость расчета, а также инвариантен к освещению.

Он принимает центральный пиксель в качестве порога и сравнивает значение яркости в каждом пикселе окрестности с ним. Если это значение больше порога (или равное значение), то пиксель принимает значение 1. Если же меньше — 0. Полученное восьмибитное число характеризует окрестность пикселя. Всего вариантов таких чисел 256. Таким образом мы присваиваем каждому пикселю изображения одну из 256 меток, характеризующих его. Далее из этих данных не составляет труда построить гистограмму и сравнивать фрагменты изображения по гистограммам LBP. Однако хотелось бы использовать окрестности произвольного радиуса с произвольным числом значащих пикселей. В 2002 году был предложен метод расширенных бинарных паттернов (extended LBP), который удовлетворяет этим условиям. Теперь окрестность выглядит следующим образом:

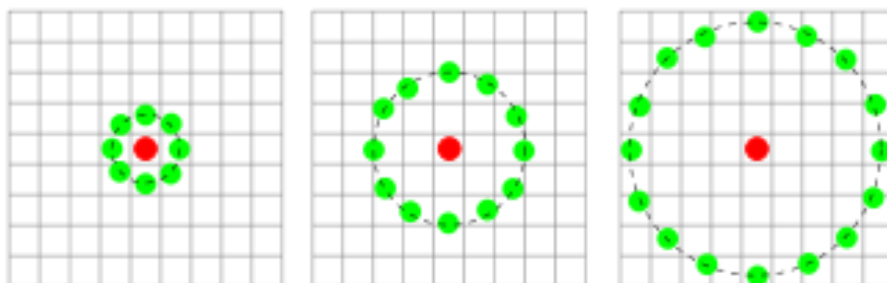


Рисунок 4 -Примеры построения точек для локальных бинарных шаблонов

1.2.4 Классификация

После того, как мы подготовили вектора признаков, мы можем перейти к построению алгоритма классификации для построения правила, отделяющего полипы и не полипы, и распределить новые данные в одну из этих категорий.

Для решения задачи классификации в машинном обучении существует несколько групп методов.

Методы статистической классификации основаны на оценивании плотностей распределения классов по обучающей выборке. Ярким представителем является наивный байесовский классификатор. Метрические алгоритмы классификации применяются в тех задачах, где удаётся естественным образом задавать объекты не их признаковыми описаниями, а матрицей попарных расстояний между объектами. При построении нейронных сетей соединяется большое количество относительно простых элементов, а обучение сводится к построению оптимальной структуры связей и настройке параметров связей. Логические алгоритмы классификации - деревья решений, представляют собой композиции простых, легко интерпретируемых правил.

Классификация на основе разделимости - большая группа методов классификации основана на явном построении разделяющей поверхности в пространстве объектов. Ярким представителем данной группы является метод опорных векторов, являющийся одним из самых эффективных алгоритмов в машинном обучении.

Исходные вектора представляется как вектор x в n -мерном пространстве, нормализованный значениями $[0,1]$. Каждая из этих точек принадлежит одному из двух классов. Далее происходит построение гиперплоскости, разделяющей эти классы. Метод опорных векторов заключается в поиске разделяющей гиперплоскости с максимальным зазором в этом пространстве. Как только он будет найден, мы можем

классифицировать новые вектора в зависимости от того, по какую сторону гиперплоскости он попадает.

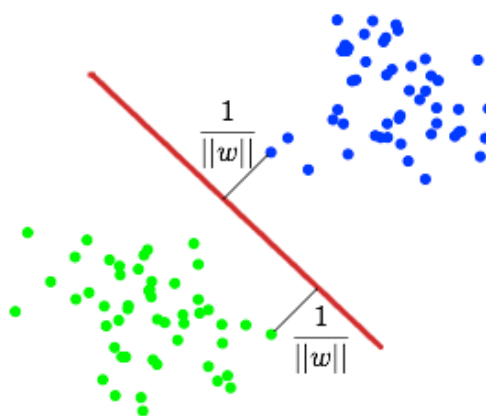


Рисунок 5 - Визуализация построения разделяющей прямой

1.2.5 Сверточные нейронные сети

В предыдущих методах для классификации фрагментов изображений извлекались признаки, подобранные вручную. Выбор был основан на предположении, что определенный признак способен в общем случае разграничить классы объектов. Данный подход является эвристическим, поэтому не самым оптимальным, так как не исключает вероятность выбора признака, который может привести к снижению значения функции ошибки.

Сверточная нейронная сеть (СНС) принимает само изображение как входной сигнал, в процессе обучения настраивает оптимальные признаки и принимает решение о классификации.

Сверточная нейронная сеть является особой архитектурой искусственной нейронной сети, имитирующая особенности зрительной коры головного мозга, состоящая из нескольких многомерных слоев и предназначенная для эффективного распознавания сложных изображений. Впервые такая модель была предложена Яном Лекуном (Yann Lecun) в 1998 году и предназначалась для распознавания рукописных символов. Особенность предложенной

Лекуном модели нейронной сети заключалась во введении в архитектуру многослойного персептрона сверточных слоев, в которых каждый нейрон был связан только небольшой областью нейронов предыдущего слоя. Такая особенность позволяла выделять на исходном изображении примитивные особенности, а на последующих слоях сети, объединяя выделенные признаки, получать все более сложные элементы таким образом детектируя сложные объекты на изображениях.

Во время создания первой СНС вычислительные мощности компьютера были сравнительно малыми, а накопленных к тому времени визуальных данных было недостаточно для того, чтобы такие математические модели получили широкое распространение. В настоящее время СНС и машинное обучение являются одними из главных направлений развития большинства сфер, связанных с информационными технологиями. Однако базовые принципы построения архитектуры таких сетей мало изменились с 1998 года.

Сверточные нейронные сети могут иметь слои следующих типов:

- Входной слой: входное изображение, включая несколько цветовых каналов.
- Сверточный слой (*Convolution*): все нейроны слоя, в отличие от персептрона, связаны только с частью нейронов предыдущего слоя.
- Слой субдискретизации (*Pooling, Subsampling*): выделение наиболее значимых признаков предыдущего слоя и значительное сокращение размерности последующих слоев сети.
- Полносвязный слой (*Fully-connected*): представляет собой скрытый слой искусственной нейронной сети типа персептрон.

Также, после сверточного слоя и полносвязного слоя может применяться слой Исклучения (*Dropout*), который предотвращает переобучение сети.

1.2.5.1 Слой сверточной нейронной сети

Сверточный слой нейронной сети является одним из основных и предназначен для выделения признаков изображения и их преобразования, которые, в свою очередь, в дальнейшем на более глубоких слоях, используются для получения более сложных признаков и, в конечном итоге, определяют класс распознаваемого объекта. Основной характеристикой данного слоя являются так называемые фильтры – многомерные матрицы весов связи нейронов предыдущего слоя с нейронами сверточного слоя.

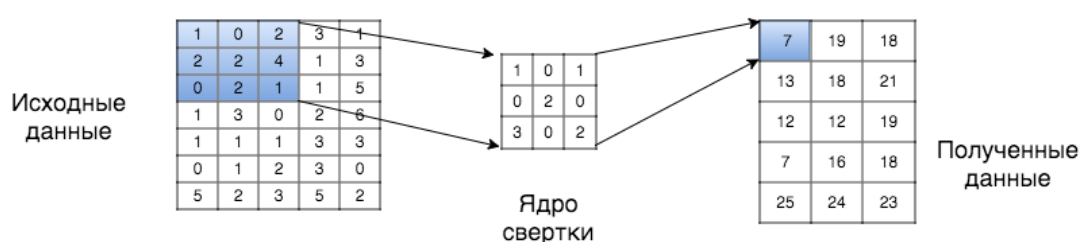


Рисунок 6 - Слой свертки

Слой субдискретизации используется для сокращения размерности данных с целью уменьшения вероятности быстрого переобучения, а также для сокращения вычислительных затрат и расхода памяти. Обычно слой субдискретизации используется после проведения операции свертки и преобразует сигналы сверточного слоя, выделяя наиболее значимые, по определенным критериям. В данном слое используется окно определенного размера для выделения областей нейронов предыдущего слоя, которые будут связаны с нейроном слоя субдискретизации. Затем, по определенным правилам выбирается и рассчитывается один сигнал нейрона слоя субдискретизации.

Наиболее типичные способы формирования сигнала данного слоя – это нахождение максимального значения среди сигналов предыдущего слоя, находящихся в пределах окна субдискретизации, либо расчёт среднего значения.

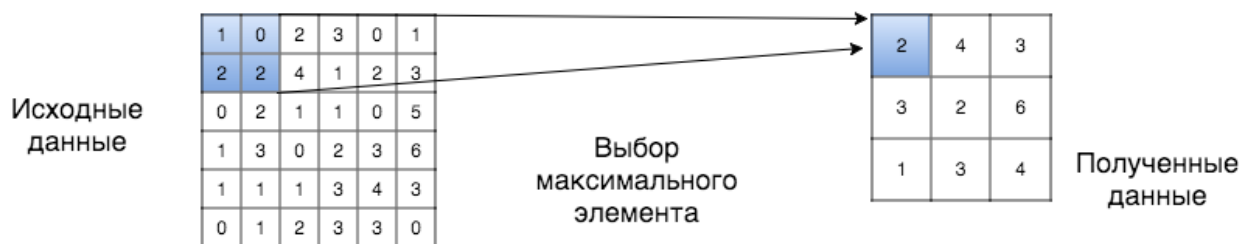


Рисунок 7 - Слой субдискретизации

Полносвязный слой является одномерным и в нем каждый нейрон связан с каждым нейроном предыдущего слоя на всех уровнях, если у предыдущего слоя присутствует параметр глубины. Основное назначение данного слоя – преобразование сигналов, полученных на сверточных уровнях сети к одномерному виду и выделение признаков на одномерном уровне.

Данный слой также может использоваться в качестве последнего (выходного) слоя СНС, результатом которого является вероятность принадлежности входного изображения определенному классу (функция softmax).

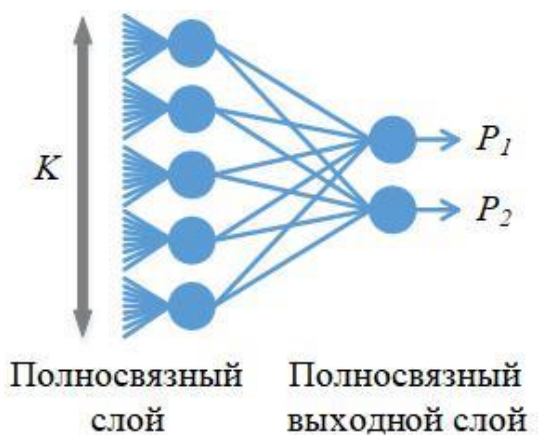


Рисунок 8 - Полносвязный слой

1.2.5.2 Функции активации и обучение

Одним из этапов разработки нейронной сети является выбор функции активации нейронов, которая используется для формирования выходного сигнала нейронов.

Функция активации сигмоиды (sigmoid) - монотонно возрастающая, всюду дифференцируемая S-образная нелинейная функция с насыщением. Сигмоид позволяет усиливать слабые сигналы и не насыщаться от сильных сигналов. Сигмоида выражается формулой:

$$f(x) = \frac{1}{(1 + e^{-x})}$$

Благодаря такой функции можно добиться непрерывного выходного значения в диапазоне $[0,1]$. Кроме того, она легко дифференцируема, что позволяет упростить обучение сети по методу обратного распространения ошибки. Это самая распространенная функция активации во многих нейронных сетях.

Гиперболический тангенс (hyperbolic tangent, tanh) принимает на входе произвольное вещественное число, а на выходе дает вещественное число в интервале от -1 до 1 . Подобно сигмоиде, гиперболический тангенс может насыщаться. Однако, в отличие от сигмоиды, выход данной функции центрирован относительно нуля. Следовательно, на практике всегда предпочтительнее использовать гиперболический тангенс, а не сигмоиду.

$$f(x) = \frac{(e^{2x} - 1)}{(e^{2x} + 1)}$$

В последние годы большую популярность приобрела функция активации под названием «выпрямитель» (ReLU). Данная функция вычисляется по формуле:

$$f(x) = \max(0, x)$$

и реализует простой пороговый переход в нуле.

Вычисление сигмоиды и гиперболического тангенса требует выполнения ресурсоемких операций, таких как возведение в степень, в то время как ReLU может быть реализован с помощью простого порогового преобразования матрицы активаций в нуле. Кроме того, ReLU не подвержен

насыщению. Так же применение ReLU существенно повышает скорость сходимости стохастического градиентного спуска по сравнению с сигмной и гиперболическим тангенсом. Считается, что это обусловлено линейным характером и отсутствием насыщения данной функции. Этот вид функции активации чаще всего используется в сверточных слоях.

Функция активации softmax — это обобщение логистической функции для многомерного случая. Функция преобразует вектор x размерности K в вектор $f(x)$ той же размерности, где каждая координата $f(x)_i$ полученного вектора представлена вещественным числом в интервале $[0,1]$ и сумма координат равна 1.

$$f(x)_i = \frac{e^{x_i}}{\sum_{k=1}^K e^{x_k}}$$

Координаты полученного вектора при этом трактуются как вероятности того, что объект принадлежит к классу i , и обычно данной функцией завершается нейронная сеть.

В СНС используется обучение с учителем, подразумевающее использование обучающей выборки для сравнения выходного сигнала сети с эталонным значением обучающей выборки, расчет ошибки и подстройку весов связей сети с целью уменьшения значения этой ошибки. В СНС в качестве алгоритма обучения используется алгоритм градиентного спуска или обратного распространения ошибки и его модификации.

1.2.5.3 Построение архитектуры сверточной нейронной сети.

Обычную архитектуру CNN для распределения изображений по k классам можно разделить на две части: цепочка чередующихся слоев свертки/подвыборки $Conv \rightarrow Pool$ (иногда с несколькими слоями свертки подряд, что позволяет выделить более глубокие признаки) и несколько полносвязных слоев (принимающих каждый пиксель как независимое значение) с слоем softmax в качестве завершающего.

Простейшая архитектура, реализующая данный подход – LeNet, имеющая следующую архитектуру:

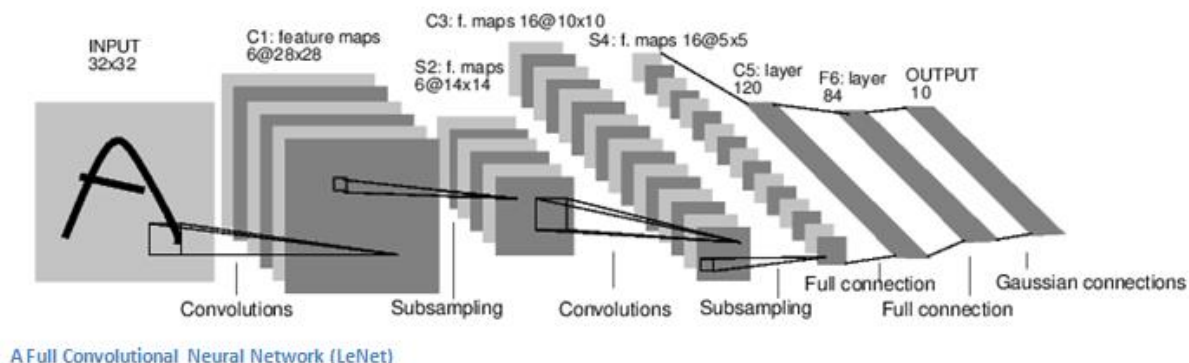


Рисунок 9 - LeNet - простейшая сверточная нейронная сеть

1.2.6 Ансамбли сверточных нейронных сетей

Ансамбль нейронных сетей – это набор нейросетевых моделей, принимающий решение путем усреднения результатов работы отдельных моделей. Обычно лучшие результаты распознавания достигаются с помощью ансамбля моделей, а не одной модели.

Существует несколько различных типов ансамблей, которые объединяют несколько методов машинного обучения в одну прогностическую модель, чтобы:

- Уменьшить дисперсию путем создания дополнительных данных для обучения (variance) — bagging; Дисперсия - это ошибка от чувствительности к небольшим колебаниям в тренировочном наборе. Высокая дисперсия может привести к тому, что алгоритм моделирует случайный шум в данных обучения, а не предполагаемые выходы
- Уменьшить смещение путем объединения сильных классификаторов (bias) — boosting; Смещение - ошибка от неверных допущений в алгоритме обучения. Высокое смещение может привести к тому, что

алгоритм пропустит соответствующие отношения между функциями и целевыми выходами (недообучение).

- Улучшить прогнозы — stacking.

Укладка (Stacking) один из наиболее общих типов ансамбля, который включает в себя обучение обучающего алгоритма, чтобы объединить прогнозы нескольких других алгоритмов обучения. В данной работе используется одна из форм укладки, которая предполагает получение среднего количества выходов моделей в ансамбле. Поскольку усреднение не принимает никаких параметров, нет необходимости обучать этот ансамбль (только его модели).

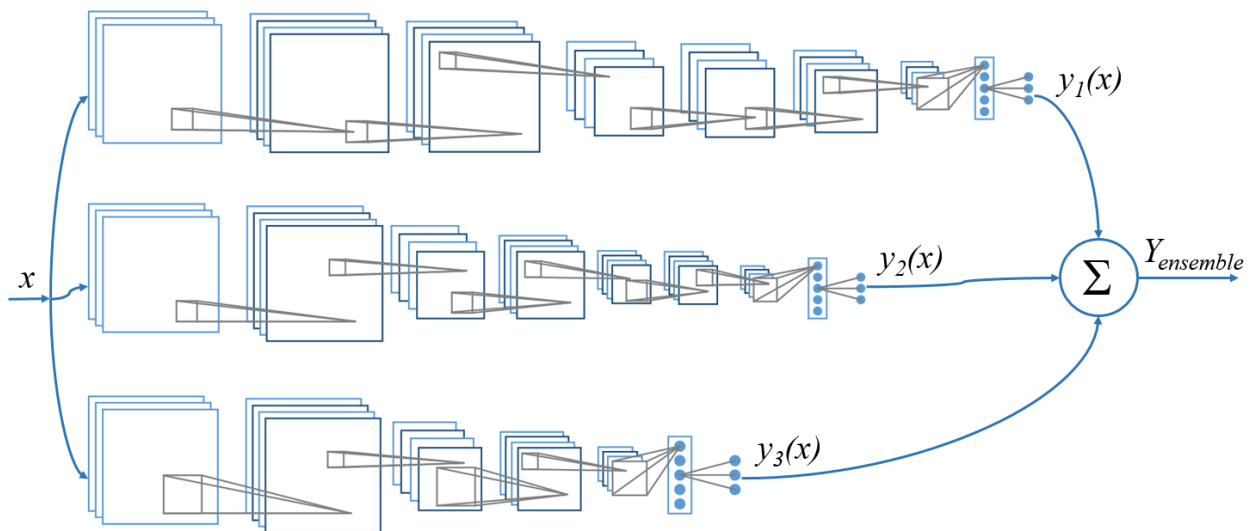


Рисунок 10 – Пример ансамбля сверточных нейронных сетей

1.2.7 Сверточные нейронные сети основанные на регионах (RCNN)

Скользящее окно имеет один существенный недостаток – мы ищем все возможные местоположения в разных масштабах, что требует больших временных затрат.

R-CNN подход был предложен Ross Girshick в 2012 году. RCNN аналогичны CNN, но не проверяют все области на изображении, а генерируют множество окон-кандидатов (Region Proposal) - прямоугольных областей изображения, которые, предположительно, ограничивают какой-то объект, а

затем запускают классификатор только на этих окнах, предварительно отмасштабировав их к единому размеру.

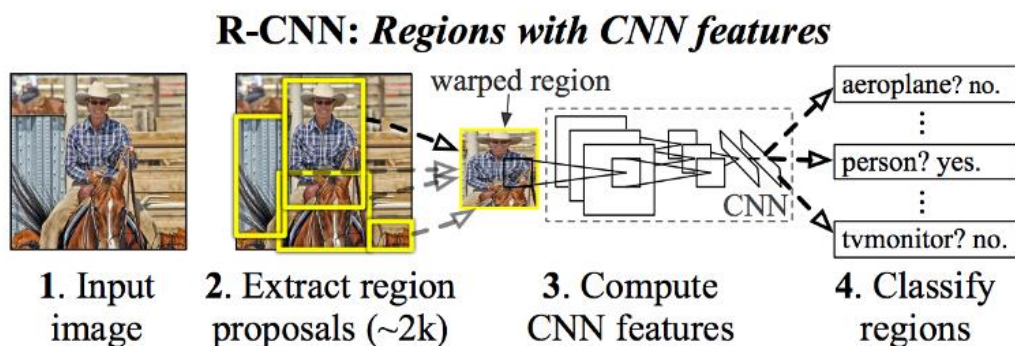


Рисунок 11 -Общая схема работы RCNN

Наиболее распространённые и эффективные подходы к генерации окон-кандидатов:

- Селективный поиск (Selective search). Основан на том наблюдении, что объекты обычно состоят из одной или нескольких частей, каждая из которых имеет почти монотонный окрас. Кандидатами являются области с примерно одинаковым цветом, а также всевозможные объединения таких областей, смежных друг с другом.

- Коробка контуров (Edge box). Контуров часто очерчивают границы объекта. В роли кандидатов выступают области, содержащие замкнутые или почти замкнутые контуры.

1.2.7.1 Селективный поиск

Селективный поиск - это алгоритм предложения региона, используемый при обнаружении объекта, основанный на вычислении иерархической группировки похожих регионов на основе соответствия цвета, текстуры, размера и формы.

Селективный поиск начинается с детальной сегментации изображения в зависимости от интенсивности пикселей, с использованием метода сегментации на основе графов и выборочного поиска. Алгоритм выборочного поиска принимает эти сегменты в качестве начального ввода и выполняет следующие шаги:

- 1) Добавляет все ограничивающие поля, соответствующие сегментированным частям, в список региональных предложений
- 2) Объединяет смежные сегменты группы на основе сходства
- 3) Переходит к шагу 1

На каждой итерации формируются и добавляются более крупные сегменты в список предложений региона. Поэтому мы создаем предложения регионов из меньших сегментов в более крупные, в восходящем подходе. Выборочный поиск использует 4 метода подобия, основанных на цвете, текстуре, размере и совместимости с формами.

1.3 Существующие подходы к распознаванию полипов

Одним из основных методов детектирования патологий на видеоданных колоноскопии являются выделение областей интереса, предположительно содержащих полипы, путем использования детектора границ и генерации текстурных и цветовых признаков с помощью дискретного вейвлет-преобразования [1]. Однако такое решение из-за большой вариации цветов и текстур полипов не является эффективным, поэтому для улучшения показателей распознавания в [2] было предложено использовать информацию о форме полипа (полученное значение TPR равно 83 % при значении FPR равном 0,15 на одном кадре видео). Дополнение существующих методов стадией двойной классификации выделенных границ областей интереса с помощью деревьев решений и дальнейшей классификацией областей интереса с помощью схемы голосования позволило сократить долю неверно классифицированных полипов на одном кадре видео до 0,1 при сохранении значения TPR на том же уровне.

В одной из наиболее современных работ [3], посвященных данной процедуре диагностики, использование сверточной нейронной сети для классификации полипов на видеозаписи колоноскопии позволило получить значение TPR классификации 65 % при FPR 0,01 на кадре видео. Полученный результат является одним из самых лучших из существующих на сегодняшний день компьютерных систем медицинской диагностики, предназначенных для детектирования полипов на видеоданных колоноскопии.

1.4 Метрики оценки качества

Основой проверки качества является тестовая выборка, в которой представлено соответствие между образцами и их классами. Для того чтобы решить, хорошо или плохо справляется с работой новая версия алгоритма, нам необходима численная метрика его качества.

В простейшем случае такой метрикой может быть доля образцов по которым классификатор принял правильное решение.

$$Accuracy = \frac{P}{N}$$

где, P – количество образцов по которым классификатор принял правильное решение, а N – размер обучающей выборки.

Учитывая, что классы в тестовых выборках для задач локализации и классификации разбалансированы, такой широко используемый показатель качества как *Accuracy* не подходит.

В подобных случаях уместно использовать метрики, именуемые точностью (precision) и полнотой (recall).

Их значения можно рассчитать по матрице неточностей (confusion matrix). Матрица неточностей – это матрица размера N на N , где N — это количество классов. Столбцы этой матрицы резервируются за экспертными решениями, а строки за решениями классификатора.

Таблица 1 Матрица неточностей для бинарной классификации

Категория		Экспертная оценка	
		Положительная	Отрицательная
Оценка системы	Положительная	TP	FP
	Отрицательная	FN	TN

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

Чтобы определить, насколько точно на изображении выделен объект, считают похожесть его предсказанного и истинного ограничивающих прямоугольников. Распространённый показатель похожести – Intersection over Union (пересечение над объединением).

$$IoU = \frac{AoO}{AoU},$$

где AoO (Area of Overlap) – пересечение истинного ограничивающего прямоугольника и предсказанного, а AoU (Area of Union) – объединение

Принято считать, что объект локализован верно, если $IoU > 0.5$. По этому критерию предсказания разделяются на верные и неверные, а дальше считаются точность, полнота или другая выбранная метрика.

2. Разработка алгоритма распознавания

2.1 Описание исходных данных

В качестве исходных данных для проведения экспериментов были взяты данные медицинской диагностики прямого кишечника CVC-ClinicDB [13], включающие в себя 612 кадров 29 различных записей колоноскопии, содержащие непосредственно кадры видеозаписи, а также соответствующие им бинарные изображения, представляющие собой маску, на которой белым цветом помечены участки изображения, соответствующие полипу.

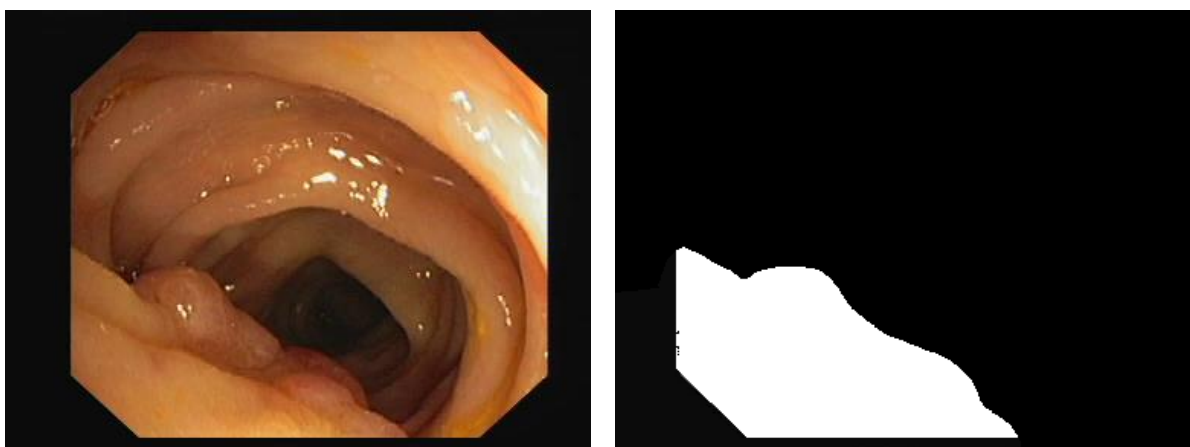


Рисунок 12 - Формат исходных данных

2.2 Анализ исходных данных

Анализ представленных кадров позволил сделать вывод о следующих сложностях поставленной задачи:

Во-первых, это внутриклассовая изменчивость - полипы имеют разнообразную сложную форму, текстуру и размеры, которые тяжело описать определенными правилами.

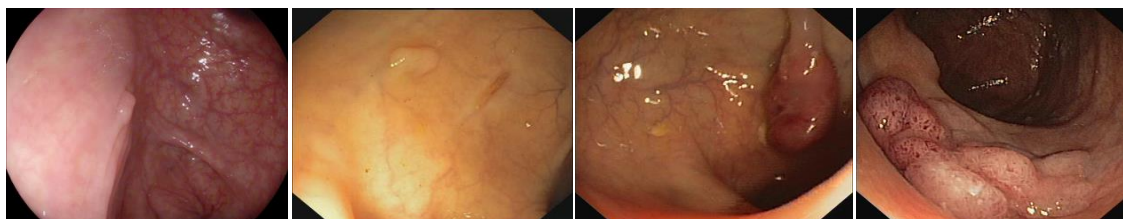


Рисунок 13 - Разнообразие форм полипов

Во-вторых – это сложный фон, с одной стороны по текстуре схожий с некоторыми полипами, а с другой - содержащий в себе сторонние объекты (водяные и воздушные структуры, кровеносные сосуды, массивные стенки кишечника), затрудняющие распознавание.

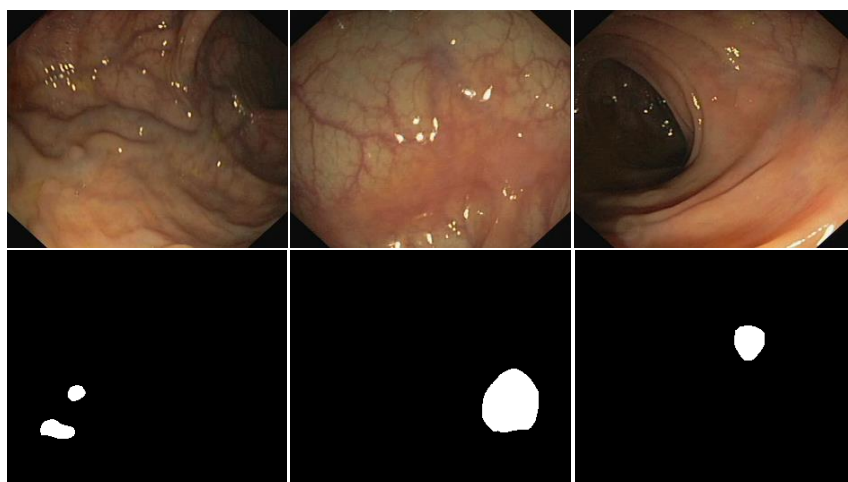


Рисунок 14 - Сложный фон

Дополнительно, свойства изображения на кадрах разных записей отличаются по яркости, преобладающему цвету, контрасту и качеству, и содержат в себе помехи, вызванные быстрым движением камеры.

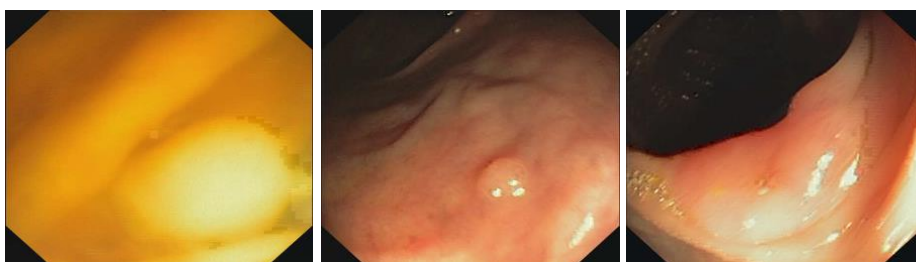


Рисунок 15 – Разные свойства камеры и помехи

2.3 Разработка алгоритмов распознавания

Учитывая сложности присущие задаче обнаружения полипов, а также все плюсы и минусы методов распознавания, для проведения экспериментов было решено использовать методы машинного и глубокого обучения.

Для того, чтобы оценить работу алгоритмов, необходимо разделить весь набор данных на обучающий и тестовый (контрольный), в соотношении 9:1. При разработке выполняется настройка алгоритма по обучающей подвыборке, затем оценивается его ошибка на объектах контрольной.

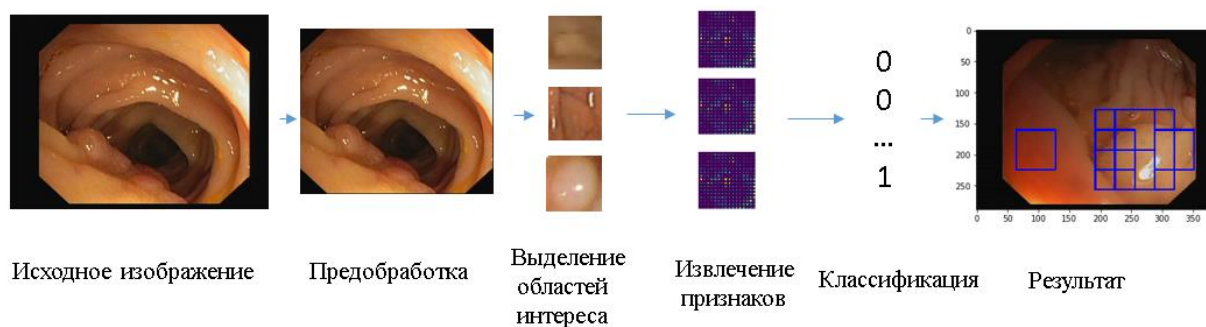


Рисунок 16 - Этапы обнаружения объектов

Первым этапом распознавания является предобработка. На изображениях присутствуют неинформативные черные рамки, которые вырезаются. После этого изображения приводятся к единому размеру 224x180.

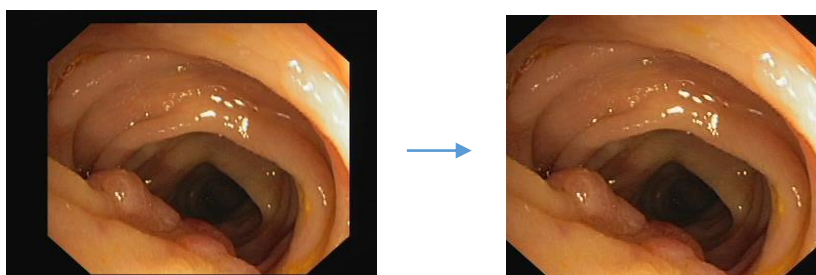


Рисунок 17 - Обрезание изображения

Для выделения областей интереса было выбрано два способа - скользящее окно при котором выбираются все возможные местоположения в разных масштабах и метод Селективного поиска, используемый в RCNN, который генерируют множество окон-кандидатов областей с примерно одинаковым цветом, а также всевозможные объединения таких областей, смежных друг с другом.

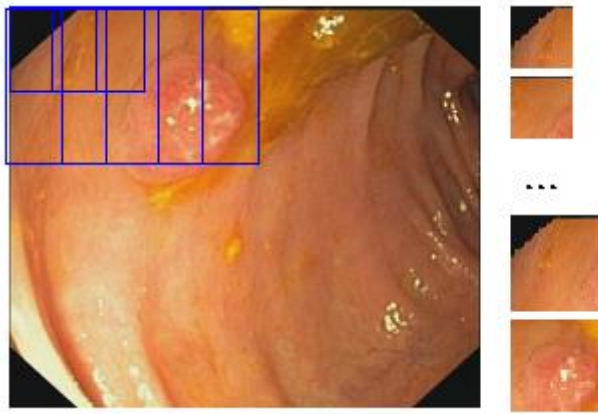


Рисунок 18 - Визуализация работы скользящего окна

Размер скользящего окна зависит от возможных размеров полипов. Проведя исследования встречающихся размеров полипов на приведенных к единому размеру изображениях было установлено, что минимальное значение меньшей стороны ограничивающего прямоугольника - 32 пикселя, а максимальное значение – 192 пикселя.

Шаг в этом диапазоне определяет качество распознавания и вычислительную сложность алгоритма. Чем меньше шаг – тем точнее и медленней будет работать алгоритм. Все окна были приведены к размеру 32x32, под размер минимального полипа.

Другие методы, используемые в RCNN – селективный поиск и коробка контуров. Поверхность кишечника имеет сложную структуру, а значит и множество контуров, полипы не всегда имеют замкнутые контуры, поэтому метод основанный контурах не подходит. Более приемлем метод основанный на сегментации по цвету.

При селективном поиске, поскольку на вход классификатора должен подаваться вектор фиксированного размера, а предложенные кандидаты имеют разные размеры, их нужно преобразовать к тому размеру, который сеть принимает на вход. В RCNN предлагается масштабировать каждый из кандидатов, но тогда изображение неестественно искажается, образуя лишние или теряя имеющиеся особенности.

В качестве улучшения метода было предложено, как и при генерации данных для классификатора, пускать скользящее окно в одном масштабе размером с минимальную сторону прямоугольника. Количество окон-кандидатов увеличится, но остается в разы меньше, чем при методе скользящего окна.

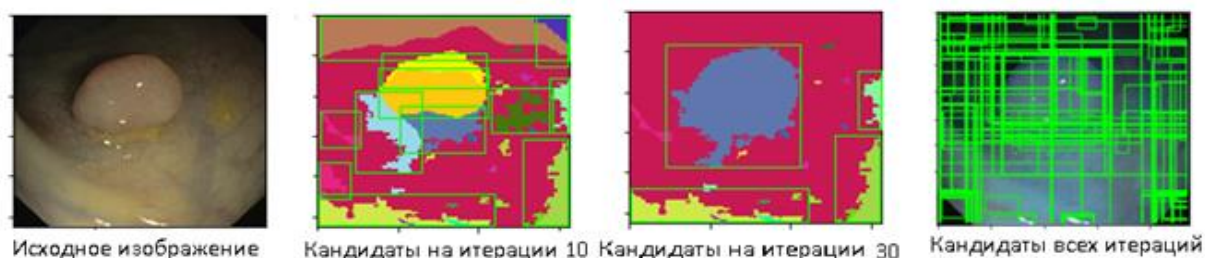


Рисунок 19 - Пример работы сегментации на основе графов

2.3.1 Генерация данных для классификации

На вход модели классификатора будет подаваться фрагменты изображения, полученные с помощью скользящего окна. Для того, чтобы классификатор выдавал хорошие результаты, необходимо его обучить на “правильных” данных.

Первый вариант – обучить модель на скользящих окнах. Но это приведет к тому, что в обучающую выборку с одной стороны, попадут окна, содержащие небольшой фрагмент полипа, который не будет полностью отражать признаки полипов, что приведет к зашумлению модели классификатора, а с другой стороны, содержать внутреннюю область полипа, которая может быть неотличима от стенок кишечника, что так же будет приводить к ухудшению модели.

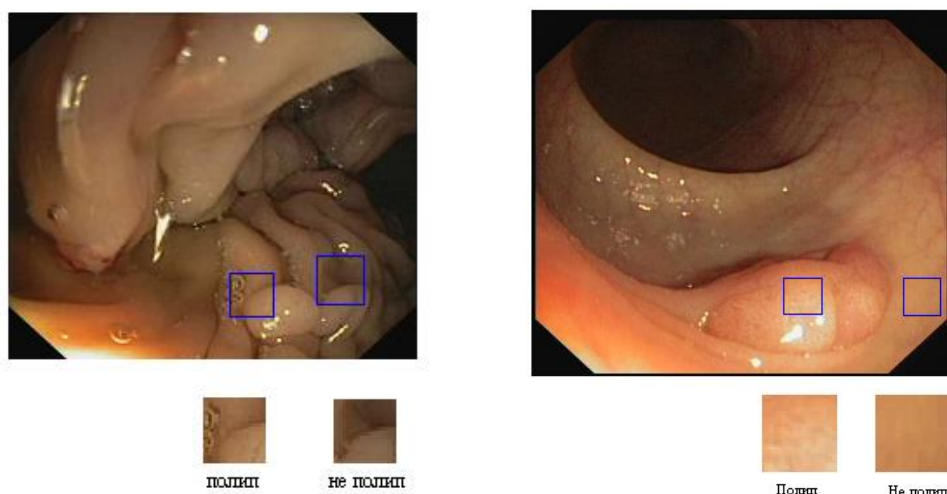


Рисунок 20 - Проблемы скользящего окна для генерации данных

Для генерации положительных образцов было принято решение вырезать ограничивающий прямоугольник вокруг полипа, и пустить скользящее окно в одном масштабе, размером с минимальную сторону прямоугольника. Для генерации отрицательных данных можно было бы вырезать случайные поверхности кишечника, но тогда, мы бы упустили большую часть информации, ведь поверхность кишечника имеет структуру, схожую с полипами.

Один из вариантов – применить скользящее окно по всему изображению для извлечения отрицательных образцов. Информация не будет упущена, но возникнет несбалансированность выборки.

2.3.1.1 Несбалансированность выборки

Выборка является несбалансированной если объектов одного класса существенно меньше, чем объектов другого класса (менее 10% от большего).

Классификаторы пытаются минимизировать число ошибок на обучающей выборке, не учитывая цены ошибок, что может привести к тому, что при высокой точности все объекты будут отнесены к наибольшему классу, а значит понизится точность и полнота.

Существует несколько подходов к работе с такими выборками:

- 1) Уменьшение большего класса (undersampling) – удаление части из объектов большего класса
- 2) Увеличение меньшего класса (oversampling) – дублирование объектов меньшего класса, что аналогично замене веса классов

Большинство вероятностных моделей слабо зависят от баланса классов. Но на такие модели, как SVM, дисбаланс классов может серьёзно повлиять. SVM строит обучающую гиперплоскость так, что примерно одно и то же число положительных и отрицательных примеров находится на разделяющей полосе или на неправильной её стороне. Изменение баланса классов может повлиять на это число, а значит, и на положение границы.

2.3.1.2 Извлечение сложных отрицательных примеров

Проблему несбалансированности выборки может решить метод извлечения сложных отрицательных примеров (hard negative mining) [14].

Обучив классификатор на равном количестве положительных и отрицательных образцов, полученных путем выделения случайных областей, не содержащих полип, применим классификатор к фрагментам, полученным с помощью метода скользящего окна. На ложно-положительных образцах дообучим классификатор и будем повторять данный процесс до тех пор, пока не извлечём все сложные отрицательные примеры.

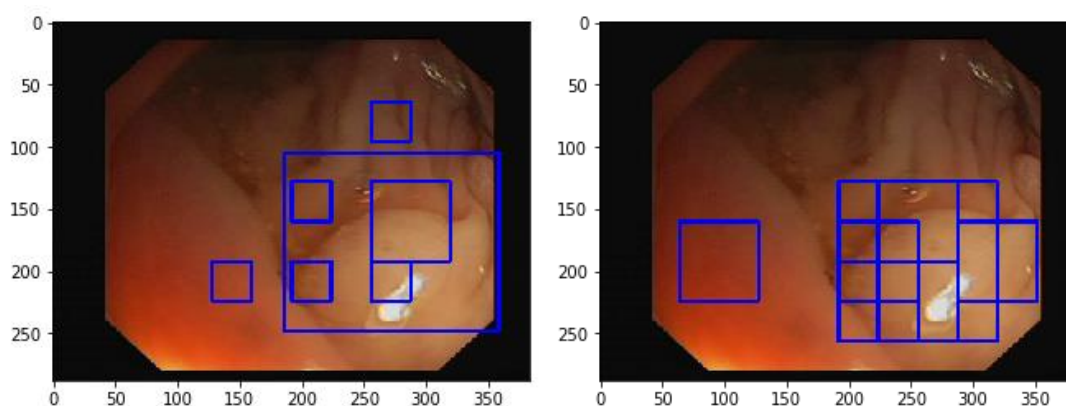


Рисунок 21 - Улучшение распознаваемости с помощью метода извлечения сложных отрицательных примеров

2.3.1.3 Аугментация

Еще один из способов улучшить модель - улучшить данные при помощи аугментации [15]. Аугментация данных (data augmentation) – это методика создания дополнительных обучающих данных из уже имеющихся. Для достижения хороших результатов классификаторы должны обучаться на большом объеме данных. Следовательно, если исходный обучающий набор содержит ограниченное количество изображений, необходимо выполнить аугментацию, чтобы улучшить результаты модели.

Улучшения результатов можно достигнуть с помощью таких преобразований, как сдвиг, повороты и зеркальные отражения, масштабирование.

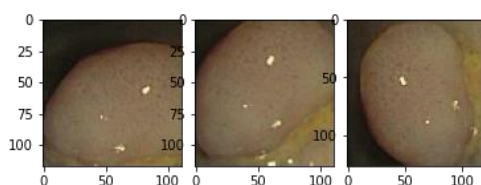


Рисунок 22 - Повороты

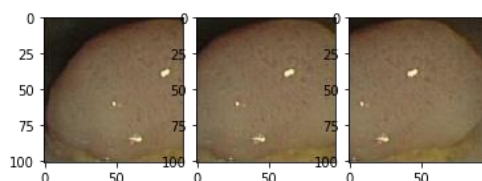


Рисунок 23 - Смещение

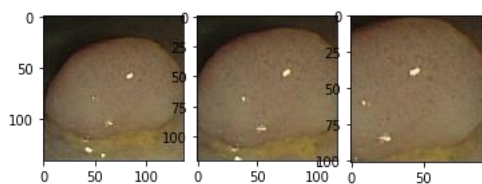


Рисунок 24- Масштабирование

Дополнительно, этот метод решает проблему несбалансированности данных.

2.3.2 Построение классификатора

При моделировании на основе гистограммы ориентированных градиентов и локальных бинарных шаблонов, обычно используется метод опорных векторов, так как является одним из самых современных и наиболее точных классификаторов.

В сложных моделях сверточных нейронных сетей нет необходимости, поскольку размер изображений небольшой - 32x32. Поскольку в каждой конкретной задаче показатели точности моделей могут отличаться, потребуется выбирать архитектуру экспериментально, стоя ее по правилам, описанным в главе 1.

2.3.3 Объединение результатов

Поскольку при сегментации извлекаются перекрывающиеся области, в результате классификации может возникнуть ситуация обнаружения одного полипа несколько раз, что затрудняет тестирование. Сократить перекрывающиеся группы обнаруженных полипов можно с помощью метода подавления немаксимумов [19]. Все обнаруженные области проверяются на пересечение. Если пересечений больше некоторого порога, стандартно – 50%, то из них выбирается большая область.

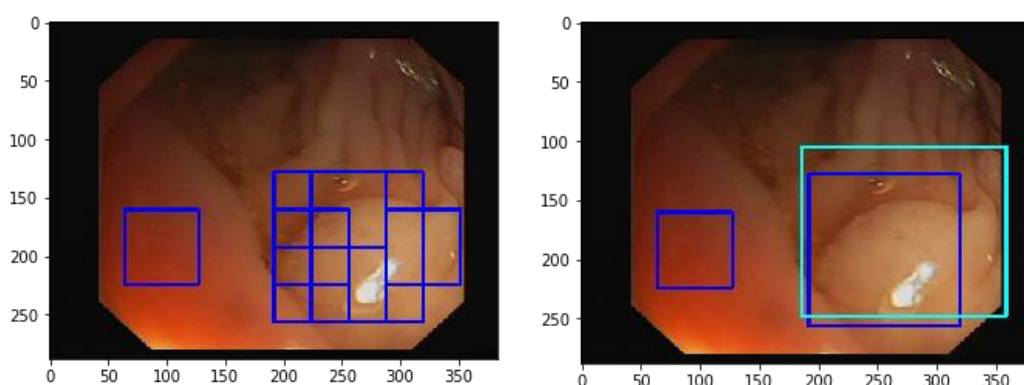


Рисунок 25 - Результат работы алгоритма подавления немаксимумов

3. Разработка системы распознавания и проведение экспериментов

Для проведения экспериментов по оценке методов распознавания и предоставления сервиса с целью применения их на видеоданных колоноскопии, необходимо разработать веб-приложение, позволяющее хранить, визуализировать и сравнивать методы, а также применять на практике методы с требуемыми параметрами.

Основные критериями к разрабатываемой системе являются:

1. Возможность оценки сравнения методов.
2. Возможность сохранения параметров, моделей и результатов оценки методов.
3. Возможность использования сохранённых моделей и методов для распознавания видеоматериала.
4. Возможность проверки видеоданных на наличие полипов.

3.1 Проектирование

В качестве абстракций действий методы были разбиты на следующий классы:

Классы Preprocess, Features, Segmentation реализуют соответствующие их именам методы. Класс Detection отвечает за обнаружение объектов, создание, сохранение и загрузку параметров методов. Класс Classify реализует классификатор или их ансамбль. С помощью встроенных методов можно управлять классификаторами.

Класс Data отвечает за выдачу данных, а DataGenerator – за кросс-валидацию, создание данных для классификации, локализации и аугментацию.

Параметры алгоритмов, модели и результаты хранятся в БД. Эти данные не являются структурированными, а значит реляционная база данных в данном случае не подходит. Поэтому была выбрана NoSql документно-ориентированная БД.

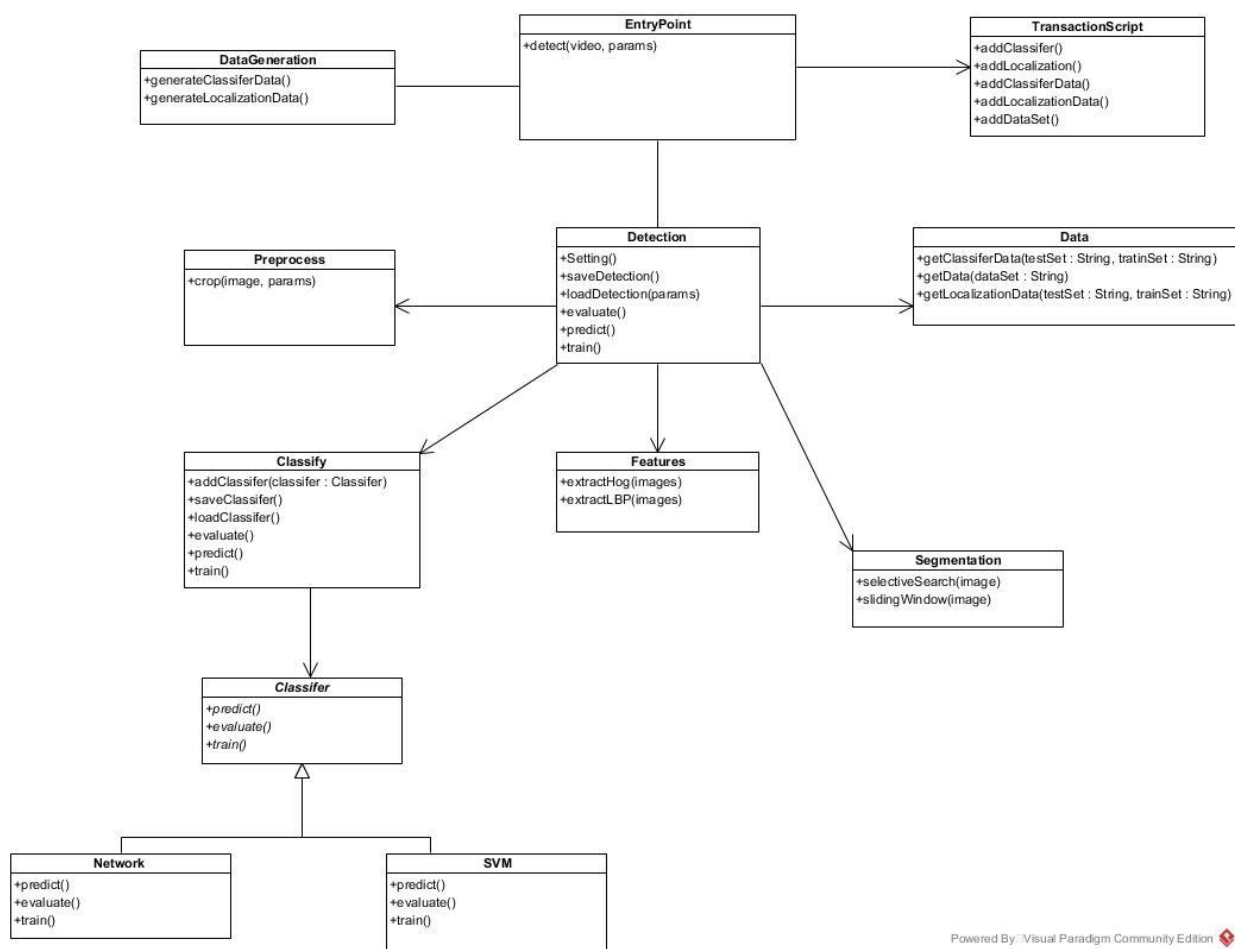


Рисунок 26 - Диаграмма классов методов распознавания

Документно-ориентированная СУБД — СУБД, специально предназначена для хранения иерархических структур данных (документов) и обычно реализуемая с помощью подхода NoSQL. Документы могут быть организованы (сгруппированы) в коллекции. Их можно считать отдалённым аналогом таблиц реляционных СУБД, но коллекции могут содержать другие коллекции.

В документно-ориентированной БД существует несколько стратегий моделирования данных.

Стратегия полного встраивания подразумевает, что все данные встроены в один документ. Смысл встраивать все связанные данные в один документ заключается в том, что некоторые запросы к данным становится делать намного проще. Но это приводит к дублированию данных, что вызывает трудности их редактирования. Стратегия “без встраивания” легка в разработке

и проста в обслуживании, но запросы получаются более трудоемкими. Стратегия «частичного встраивания» подразумевает сочетание предыдущих методов и является оптимальным вариантом. Полученная база данных представлена ниже.

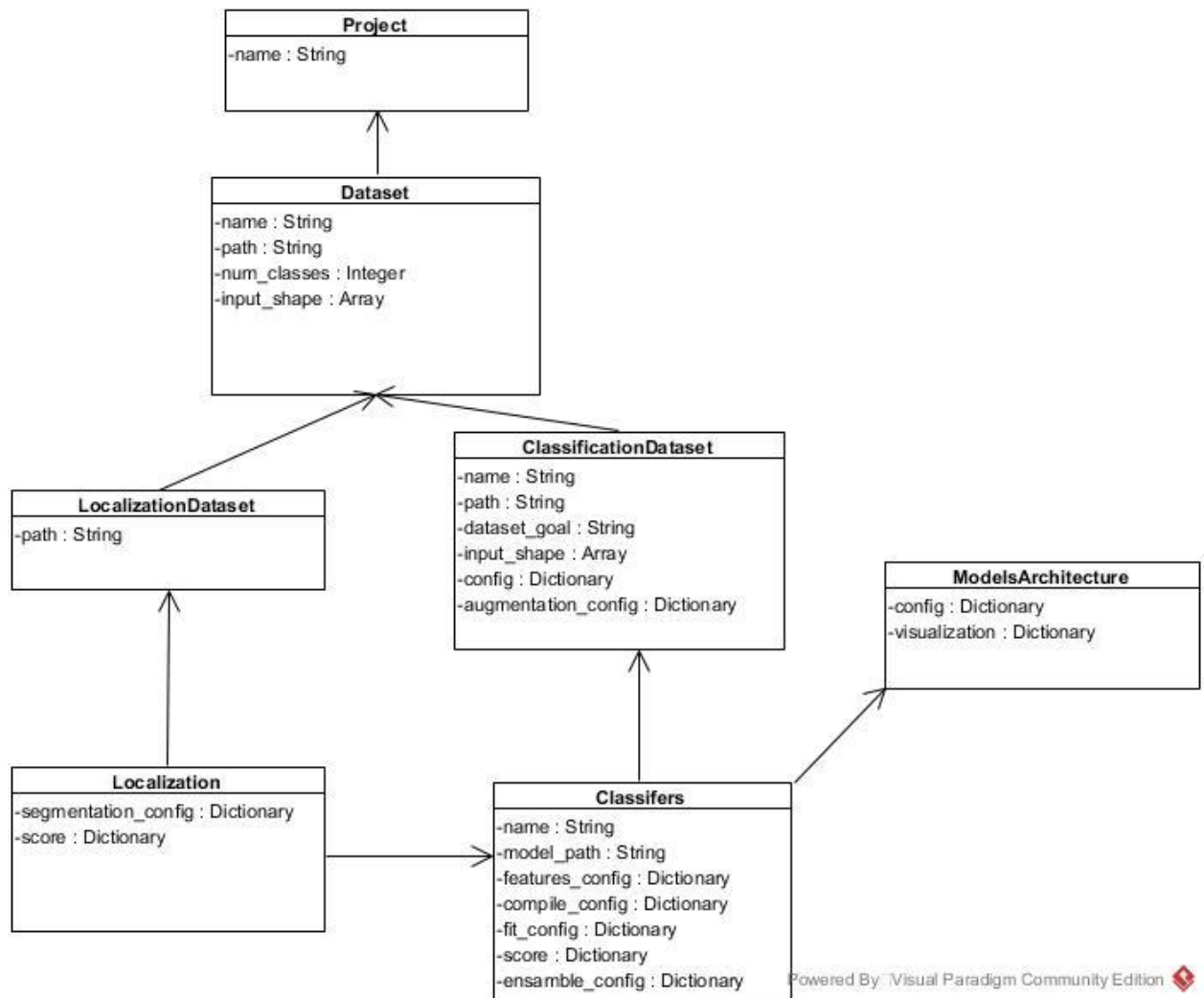


Рисунок 27 Концептуальная модель базы данных

Для работы с БД был реализован класс TSMongo, реализующий паттерн Transaction Script. Методы данного класса добавляют/выдают наборы данных, методы и модели.

```
{'_id': ObjectId('5ae0623a89349f3a961025d0'),
  'ansamble': None,
  'architecture_id': ObjectId('5ae05adf89349f3a961025c4'),
  'compile_config': {'loss': 'binary_crossentropy',
    'metrics': ['accuracy'],
    'optimizer': 'adam'},
  'dataset': {'test': ObjectId('5adb13bc89349f12277479b2'),
    'train': ObjectId('5adb001c89349f1909d5aa16')},
  'fit_config': {'fit_params': {'batch_size': 32, 'epochs': 12},
    'type': 'fit'},
  'model_path': 'media/dataset/models/simple_cnn.h5',
  'score': {'conf_matrix_config': {'fn': 309,
    'fp': 742,
    'tn': 3312,
    'tp': 300},
    }
}
```

Пример записи в документе Classifiers

3.2 Реализация

3.2.1 Выбор технологий

При выборе технологий ключевую роль играло наличие мощных библиотек работы с изображениями и глубоким обучением, а также возможность создания сервиса. Идеальным вариантом оказался язык программирования python, имеющий в своем арсенале такие библиотеки как opencv и scikit-learn – библиотеки компьютерного зрения и машинного обучения, реализующие большую часть классических методов анализа изображений, keras – мощный фреймворк работы с нейронными сетями, являющийся оберткой над Tenser-flow и Theano, библиотеку Matplotlib для визуализации и Jupyter Notebook - интерактивную среду разработки, предоставляющий "блокнот" для браузера. Так же на нем реализованы фреймворки для созданий веб-приложений Flask и Djano. Первый представляет собой гибкий микрофреймворк, а второй – фреймворк реализующий шаблон MVC и по умолчанию включающий в себя большое количество функционала, такого как ORM, админ-панель и др.

В связи с небольшим количеством планируемого веб-сервиса, включающего по большей части методы распознавания и использующего

документно-ориентированную базу данных для разработки системы, был выбран Flask.

Самая популярная на сегодняшний день NoSQL СУБД – MongoDB. Для работы с ней в языке python реализована библиотека PyMongo, поэтому в качестве СУБД была выбрана именно она.

3.2.2 Интерфейс пользователя

Зайдя на страницу распознавания, пользователю предлагается загрузить видео и выбрать один из методов распознавания, находящийся в БД.

Automated object detection

[Main](#) [Projects](#) [Architectures](#) [Classifiers](#) [Localization](#)

Upload new File

ShortVD_wp_66.wmv

Time/recall



Select model

- Ensamble 88.12 ▾
- Ensamble 88.12
- Ensamble 79.94
- Ensamble 79.40
- Ensamble 76.12
- CNN 75.64
- CNN 73.11
- CNN 71.01
- test2 69.98
- HOG 62.34
- CNN 61.78
- LBP 59.14
- GMTP 39.98
- GMTP 21.44
- test2 11.05

Рисунок 28 - Страница загрузки видео

Для обработки изображений исходное видео необходимо разбить на отдельные кадры. Для увеличения производительности захватывается из видео каждый десятый кадр, и, если полип найден – промежуточные значения вычисляются как среднее между двумя кадрами.

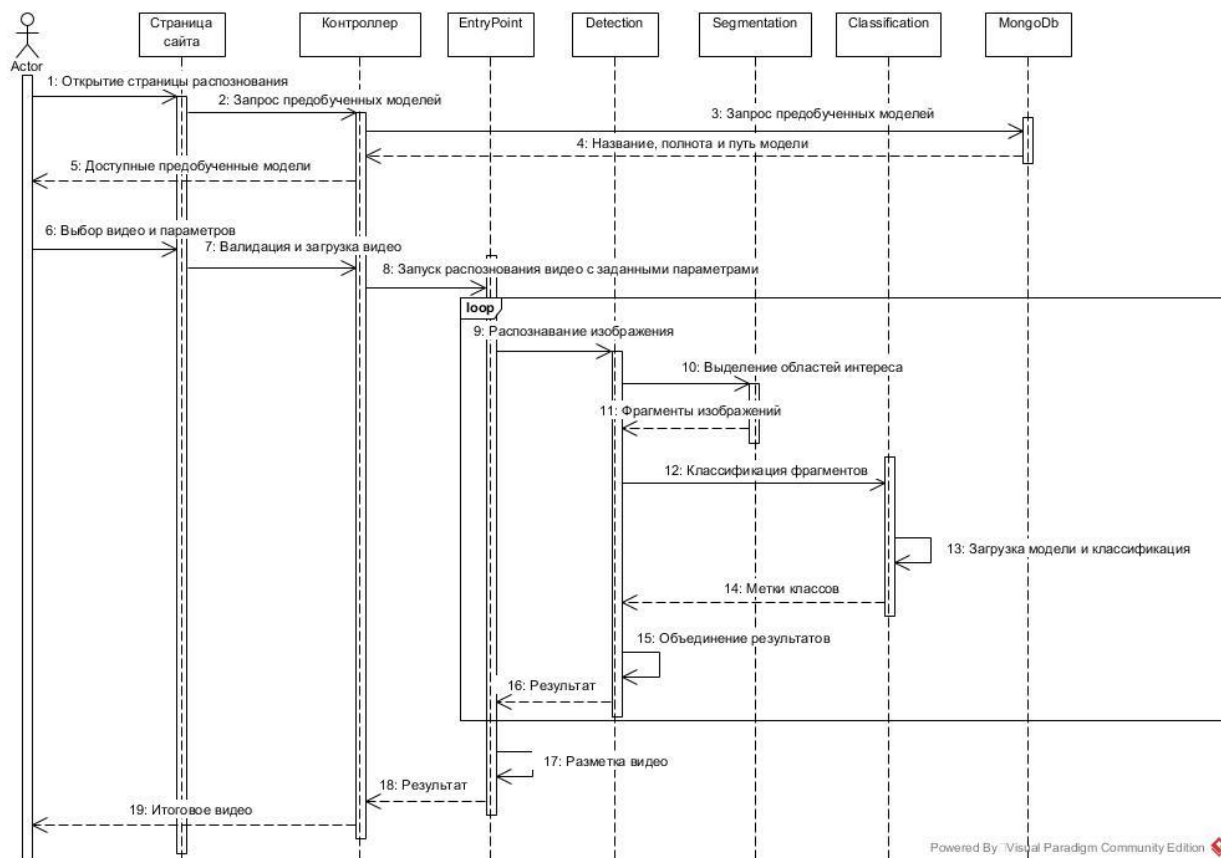


Рисунок 29 – Диаграмма последовательности сценария распознавания

После окончания обработки сайт перешлет пользователя на скачивание видео, где распознанные полипы отмечены ограничивающим прямоугольником.

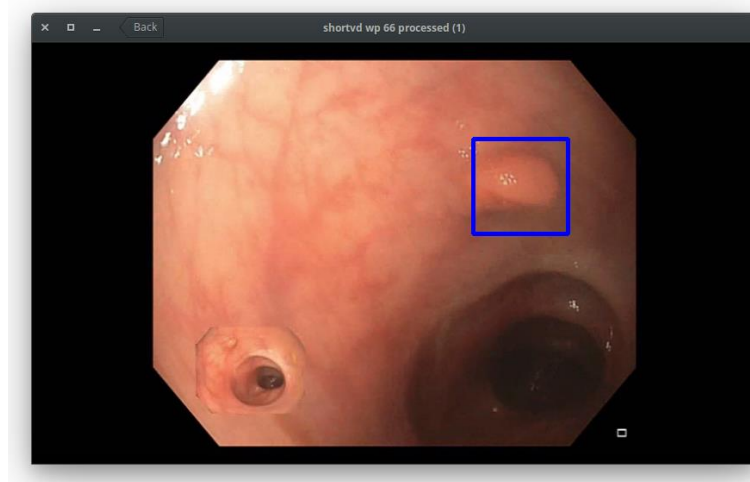


Рисунок 30 - Распознанное видео

3.3 Проведение экспериментов

Сначала на примере НОГ была проверена зависимость применения метода на разных цветовых каналах и усредненного значения (в градациях серого).

Таблица 2 Зависимость качества от цветового канала

Канал	Точность	Полнота
R	73	57
G	68	62
B	57	70
Gray-scale	72	63

На канале красного цвета точность относительно черно-белого незначительно повысилась, но снизилась полнота. На синем же канале полнота повысилась, но сильно снизилась точность. Следовательно, все-таки лучше использовать усредненное черно-белое изображение.

Далее был исследован тот же метод в сравнении с локальными бинарными шаблонами. Метод гистограммы ориентированных градиентов показал большую точность по сравнению с ним. Были проверены все возможные параметры методов (для НОГ – размер сетки и блоков нормализации, для LBP – радиус и количество точек).

Таблица 3 Сравнение НОГ и LBP

Метод	Параметры	Точность	Полнота
НОГ	(2,2), (2,2)	72	63
LBP	$p = 5, r = 11$	64	59

Проверены методы улучшения данных и обучения. Получены следующие результаты:

Таблица 4 Улучшение данных и обучения

Данные	Размер обучающей выборки	Метод	Точность %	Полнота %
ОП	1981	LBP	64	59
ОП	1981	HOG	72	63
ОП + Hard Negative	2333	LBP	76	68
ОП + Hard Negative	2333	HOG	82	68
ОП + HN + смещения + повороты + масштабирование	10482	LBP	78	71
ОП + HN + смещения + повороты + масштабирование	10482	HOG	84	76

Максимальная точность и полнота обнаружения при IoU>50% с применением этих методов составили 78 и 71 соответственно, что говорит о том, что точность распознавания данных методов не достаточна для практического использования.

Таблица 5 Точность обнаружения полипов HOG и LBP

Метод	Параметры	Время работы	Точность	Полнота
HOG	(2,2), (2,2)	3.6 сек	78	71
LBP	p = 5, r = 11	2.9 сек	62	68

Следующими методами для исследования стали нейронные сети. В соответствии со схемой построения архитектуры сверточной нейронной сети были реализованы следующие модели:

Таблица 6 Архитектура первой сверточной сети

Тип слоя	Функция активации	Размерность
Input		32x32x3
Conv2d	ReLu	32x32x20
MaxPooling		16x16x20
Conv2d	ReLu	16x16x40
MaxPooling		8x8x20
Dence	ReLu	128
Output	softmax	2

Таблица 7 Архитектура второй сверточной сети

Тип слоя	Функция активации	Размерность
Input		32x32x3
Conv2d	ReLu	32x32x20
Conv2d	ReLu	32x32x40
MaxPooling		16x16x40
Dence	ReLu	256
DropOut		
Dence	ReLu	256
Output	softmax	2

Таблица 8 Архитектура третьей сверточной сети

Тип слоя	Функция активации	Размерность
Input		32x32x3
Conv2d	ReLu	32x32x20
Conv2d	ReLu	32x32x40
MaxPooling		16x16x40
Conv2d	ReLu	16x16x60
MaxPooling		8x8x60
Dence	ReLu	256
Output	softmax	2

Одиночные модели и ансамбли показали следующие результаты:

Таблица 9 Результаты сверточных нейронных сетей и ансамблей

Модель	Точность	Полнота
1	84	79
2	79	78
3	88	86
1+2	85	87
1+3	91	89
2+3	89	88
1+2+3	92	91

Лучшая одиночная и лучшая ансамблевая модель была сравнена с RCNN той же архитектуры в локализации.

Таблица 10 Сравнений скорости и качества работы

Модель	Время работы	Точность	Полнота
3	3.3	85	82
HOG	3.6	78	71
1+2+3	3.8	90	88
RCNN	1.6	81	79

Методы глубокого обучения обошли классический метод машинного обучения. Работать алгоритм RCNN стал чуть быстрее, но при этом начал упускать некоторые регионы с полипом. В итоге время работы, точность и полнота локализации разных методов составила:

Таблица 11 Ключевые результаты

Модель	Время работы	Точность	Полнота
HOG	3.6	78	71
RCNN	1.6	84	80
1+2+3	3.8	90	88

Заключение

В ходе работы была построена система для распознавания патологий на видео колоноскопии, были разработаны и протестированы несколько основных методов распознавания.

Ансамбли сверточных нейронных сетей показали самую высокую эффективность для решения задачи классификации полипов. При обнаружении полипов оба рассмотренных метода сегментации показали хороший результат, и могут меняться в зависимости от временных ограничений и требуемой точности. Однако обучающая выборка является недостаточной для построения качественной модели.

В дальнейшем планируется добавить методы предобработки для усиления признаков, увеличить объем обучающих данных, улучшить архитектуру ансамблей сверточных нейронных сетей, а также усовершенствовать веб-приложения для более эффективной разработки и использования.

Список использованной литературы

1. Alexandre, L.A., Nobre, N., Casteleiro, J.: Color and position versus texture features for endoscopic polyp detection. In: International Conference on BioMedical Engineering and Informatics, BMEI 2008, vol. 2, pp. 38–42. IEEE.
2. Tajbakhsh, N., Chi, C., Gurudu, S.R., Liang, J.: Automatic polyp detection from learned boundaries. In: 2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI), 2014. pp. 97–100. IEEE.
3. Tajbakhsh N., Gurudu S.R., Liang J. A Comprehensive Computer-Aided Polyp Detection System for Colonoscopy Videos. In: Ourselin S., Alexander D., Westin CF., Cardoso M. (eds) Information Processing in Medical Imaging. IPMI 2015. Lecture Notes in Computer Science, 2015. Vol. 9123. Springer, Cham.
4. N. Dalal, B. Triggs. Histograms of oriented gradients for human detection. In Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, volume 1, pages 886–893. IEEE, 2005.
5. R. Girshick, J. Donahue, T. Darrell, J. Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. arXiv:1311.2524
6. J. R. Uijlings, K. E. van de Sande, T. Gevers, A. W. Smeulders. Selective search for object recognition. International journal of computer vision, 104(2):154–171, 2013.
7. C. L. Zitnick and P. Dollár. Edge boxes: Locating object proposals from edges. In Computer Vision–ECCV 2014, pages 391–405. Springer, 2014.
8. Y. Tang. Deep Learning using Linear Support Vector Machines. arXiv:1306.0239
9. A. Nøklund. Improving Back-Propagation by Adding an Adversarial Gradient. arXiv:1510.04189v2
10. K. He, X. Zhang, S. Ren, J. Sun. Deep Residual Learning for Image Recognition. arXiv:1512.03385
11. Li B. and Meng M.Q.H., Automatic Polyp Detection for Wireless Capsule Endoscopy Images, Expert Systems with Applications, 39(12), p.10952-10958, 2012.

12. Tamaki T., Yoshimuta J., Kawakami M., Raytchev B., Kaneda K., Yoshida S., Takemura Y., Onji K., Miyaki R. and Tanaka S., Computer-Aided Colorectal Tumor Classification in NBI Endoscopy using Local
13. CVC-ClinicDB // MICCAI Sub-Challenge on Automatic Polyp Detection Challenge in Colonoscopy Videos. 2015. URL: <https://polyp.grand-challenge.org/site/Polyp/CVCClinicDB/> (дата обращения: 12.04.2017).
14. Shaohua Wan Zhijun Chen Tao Zhang Bo Zhang Kong-kat Wong, Bootstrapping Face Detection with Hard Negative Examples. arXiv:1608.02236
15. Sebastien C. Wong, Adam Gatt, Victor Stamatescu, Understanding data augmentation for classification: when to warp. arXiv:1609.08764
16. R. Girshick, J. Donahue, T. Darrell, J. Malik. Region-based convolutional networks for accurate object detection and segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2016. Vol. 38, No 1. P. 142–158.
17. Описание Image Processing Toolbox [Электронный ресурс] – Электрон. дан. – URL: http://matlab.exponenta.ru/imageprocess/book5/5_6.php (дата обращения: 25.03.2017)
18. Ensembling ConvNets using Keras [Электронный ресурс] – Электрон. дан. – URL: <https://towardsdatascience.com/ensembling-convnets-using-keras-237d429157eb> (дата обращения: 20.04.2017)
19. Jan Hosang, Rodrigo Benenson, Bernt Schiele. Learning non-maximum suppression. arXiv: 1705.02950

Отчет о проверке на заимствования №1

Автор: Aleksander Nikitenko berteiusgedoer@gmail.com / ID: 5737787

Проверяющий: Aleksander Nikitenko (berteiusgedoer@gmail.com) / ID: 5737787)

Отчет предоставлен сервисом «Антиплагиат»- <http://www.antiplagiat.ru>

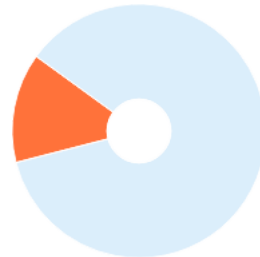
ИНФОРМАЦИЯ О ДОКУМЕНТЕ

№ документа: 1
Начало загрузки: 29.05.2018 13:32:20
Длительность загрузки: 00:00:03
Имя исходного файла: ВКРН
Размер текста: 2125 кБ
Символов в тексте: 50105
Слов в тексте: 5938
Число предложений: 326

ИНФОРМАЦИЯ ОБ ОТЧЕТЕ

Последний готовый отчет (ред.)
Начало проверки: 29.05.2018 13:32:24
Длительность проверки: 00:00:02
Комментарии: не указано
Модули поиска:

ЗАИМСТВОВАНИЯ	ЦИТИРОВАНИЯ	ОРИГИНАЛЬНОСТЬ
14,46% 	0% 	85,54% 



Заимствования — доля всех найденных текстовых пересечений, за исключением тех, которые система отнесла к цитированиям, по отношению к общему объему документа.

Цитирования — доля текстовых пересечений, которые не являются авторскими, но система посчитала их использование корректным, по отношению к общему объему документа. Сюда относятся оформленные по ГОСТу цитаты; общеупотребительные выражения; фрагменты текста, найденные в источниках из коллекций нормативно-правовой документации.

Текстовое пересечение — фрагмент текста проверяемого документа, совпадающий или почти совпадающий с фрагментом текста источника.

Источник — документ, проиндексированный в системе и содержащийся в модуле поиска, по которому проводится проверка.

Оригинальность — доля фрагментов текста проверяемого документа, не обнаруженных ни в одном источнике, по которым шла проверка, по отношению к общему объему документа.

Заимствования, цитирования и оригинальность являются отдельными показателями и в сумме дают 100%, что соответствует всему тексту проверяемого документа.

Обращаем Ваше внимание, что система находит текстовые пересечения проверяемого документа с проиндексированными в системе текстовыми источниками. При этом система является вспомогательным инструментом, определение корректности и правомерности заимствований или цитирований, а также авторства текстовых фрагментов проверяемого документа остается в компетенции проверяющего.

№	Доля в отчете	Доля в тексте	Источник	Ссылка	Актуален на	Модуль поиска	Блоков в отчете	Блоков в тексте
[01]	4,55%	4,55%	Скачать - 0 байт	http://nauchkor.ru	13 Апр 2017	Модуль поиска Интернет	15	15
[02]	2,33%	2,33%	Лабораторные работы. ТВ - ...	http://studfiles.ru	30 Июл 2016	Модуль поиска Интернет	5	5
[03]	1,7%	1,85%	http://diss.vlsu.ru/uploads/me...	http://diss.vlsu.ru	06 Ноя 2016	Модуль поиска Интернет	4	5

Еще источников: 13

Еще заимствований: 5,88%