**Исследование методов машинного обучения для поиска объектов на изображении**

**Актуальность:** машинное (компьютерное) зрение используется для решения многих прикладных задач, возникающих в современном мире, таких как автономное управление беспилотным транспортом, автоматизированные системы безопасности, поддержка работы медицинских сотрудников. Технологии машинного зрения основаны на методах машинного обучения. Классическая задача в компьютерном зрении – это определение наличия в видеоданных некоторого характерного объекта, особенности или активности. Эта задача может быть достоверно и легко решена человеком, существуют также алгоритмы машинного зрения, позволяющие распознавать один или несколько предварительно заданных объектов или классов объектов. Однако до сих пор данная задача не решена удовлетворительно в общем случае: распознавание случайных объектов в случайных ситуациях.

Нейронные сети – это подобласть машинного обучения.

**Объектом** исследования являются методы машинного обучения, используемые для поиска объектов на изображении.

**Предмет** исследования – это характеристики, характерные свойства, особенности и возможности методов машинного обучения, используемых для поиска объектов на изображении. (… методов компьютерного зрения)

**Цель**: исследование методов машинного зрения, обнаружение их сильных и слабых сторон в задаче обнаружения и распознавания объектов, для увеличения эффективности их практического применения путём комбинации методов. Комбинация простых методов, учитывающая их сильные и слабые стороны...

**Практическая значимость**: полученный в ходе исследования алгоритм комбинации методов машинного обучения может быть использован взамен отдельных методов. Он позволяет ускорить процесс распознавания объектов, снизить процент ошибочных обнаружений и, в целом, повысить эффективность поиска объектов на изображении.

**Анализ существующих методов**:

Машинное зрение — это научное направление в области искусственного интеллекта и связанные с ним технологии получения изображений объектов реального мира, их обработки и использования полученных данных для решения разного рода прикладных задач без участия (полного или частичного) человека.

На сегодняшний день задача поиска объектов на изображении может быть решена при помощи следующих подходов: цветовые фильтры, выделение и анализ контуров, сопоставление с шаблоном, работа с особыми точками, методы машинного обучения.

Цветовые фильтры используются в том случае, если объект существенно выделяется на фоне по цвету, в таком случае для его обнаружения достаточно правильно подобрать соответствующий фильтр.

Если заранее известна форма объекта, то можно применять метод выделения и анализа контуров. Таким образом, в случае поиска круглого объекта анализируются окружности на изображении.

Сопоставление с шаблоном может быть использовано, если имеется изображение искомого объекта. Данный метод заключается в поиске на изображении областей, совпадающих с шаблонным изображением объекта.

При работе с особыми точками на изображении объекта выделяются особые точки, например углы, и сопоставляются с такими же особенностями на анализируемом изображении.

Методы машинного обучения основываются на получении классификатора путём обучения на тренировочной выборке объектов и дальнейшем его применении на всех частях некоторым образом разделенного изображения. Главным преимуществом этих методов является высокая обобщающая способность: в то время, как подходы сопоставления шаблонов или особых точек распознают только одно конкретное изображение, методы машинного обучения способны определять объект, класс или группу объектов.

Задачу поиска объектов на видео можно разделить на три подзадачи или этапа: обнаружение, распознавание и сопровождение. На этапе обнаружения необходимо на изображении выделить области поиска интересующих нас объектов. Затем, на этапе распознавания, для каждого из найденных объектов уточняется его тип. Сопровождение заключается в локализации распознанных объектов на следующих кадрах.

Машинное обучение можно применять, по крайней мере, для решения задач обнаружения и распознавания, сводимых к построению классификатора изображений. Классификатор изображений состоит из двух частей: метод извлечения признаков и, собственно, сам классификатор. Среди методов извлечения признаков, используемых для построения классификаторов изображений, наиболее популярны методы HOG, BoW, Виолы-Джонса.

Гистограммы направленных градиентов (HOG) – это метод извлечения признаков из изображений, который напоминает метод вычисления дескрипторов SIFT для особых точек, но в этом методе вычисление происходит не для окрестности, а для всего изображения. Общий алгоритм вычисления HOG состоит в следующем: изображение разделяется на части – ячейки, затем для каждой ячейки осуществляется построение гистограммы направлений яркости градиента, после чего гистограммы ячеек нормируются по контрасту и объединяются в классы.

Мешок слов (BoW)– это метод извлечения признаков из изображений, который является адаптацией метода частотного анализа текстов. Применение методов машинного обучения для распознавания объектов на изображении предполагает наличие учебного набора, состоящего как минимум из двух частей: набор изображений, содержащих объект, и набор изображений, его не содержащих. Аналогично методу частотного анализа текстов, для использования мешка слов необходимо для начала собрать словарь. Алгоритм построения словаря состоит в следующем. Сначала выбирается метод выявления особых точек и метод вычисления их дескрипторов. Затем для каждого изображения, входящего в учебный набор, определяются особые точки и вычисляются дескрипторы для них. После этого похожие дескрипторы объединяются в группы, то есть выполняется кластеризация множества полученных дескрипторов. Полученные кластеры будут играть роль визуальных слов и составлять словарь для частотного анализа изображения. Признаки BoW извлекаются из изображения следующим образом: методом, который был выбран на этапе составления словаря, определяются особые точки на изображении и вычисляются для них дескрипторы, найденные дескрипторы разбираются по кластерам словаря и затем для каждого кластера посчитывается количество найденных дескрипторов. Таким образом, вектор BoW-признаков – это количество найденных дескрипторов в каждом кластере словаря.

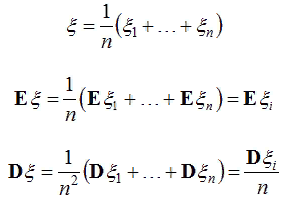
Метод Виолы-Джонса, который часто используется для поиска лиц на изображении, построен на основе признаков Хаара. Признак Хаара вычисляется следующим образом: выбирается прямоугольная область на изображении, эта область разбивается на несколько смежных прямоугольных частей, в каждой из них суммируется яркость точек, после чего вычисляется разность между полученными суммами. Итоговые разность и являются значением признака Хаара. Для формирования характеристики изображения используют несколько разных признаков Хаара, каждый из которых имеет свои параметры: размер области для вычисления признака, количество частей, позиция на изображении.

Существует также модель классификатора изображений, обладающая встроенным механизмом извлечения признаков из изображения – свёрточная нейронная сеть. Механизм извлечения признаков является частью сети и «самонастраивается» в процессе обучения.

Важной задачей поиска объектов на изображении является локализация объектов. Данная задача состоит в определении наличия объекта на изображении и вычислении занимаемой им области изображения. Задача локализации решается при помощи скользящего окна: на каждом шаге рассматривается одна из частей изображения, и те из них, в которых обнаружен объект, сужаются пока не достигнут размера объекта. Результатом работы алгоритма является список, состоящий из позиции окна, соответствующих ему значения масштаба изображения и значения выхода классификатора. Далее окна приводятся к единому масштабу и удаляются дубликаты окон.

Метод машинного обучения, где несколько моделей обучаются для решения одной и той же проблемы и объединяются для получения лучших результатов называется ансамблевым методом. Основная предпосылка заключается в том, что результат работы нескольких моделей будет более точен, чем результат только одной модели.

Если рассматривать значения базовых алгоритмов на объекте как независимые случайные величины с одинаковым матожиданием и одинаковой конечной дисперсией, то понятно, что случайная величина имеет такое же матожидание, но меньшую дисперсию:



Наиболее популярными ансамблевыми методами являются: стекинг, бэггинг, бустинг. При стекинге используется несколько разнородных методов. Их обучают и объединяют для построения прогноза, основанного на результатах различных слабых моделей. В случае бэггинга однородные модели обучают на разных наборах данных и объединяют, получая прогноз путём усреднения. При использовании бустинга несколько однородных моделей последовательно обучаются, исправляя ошибки друг друга.

В настоящей работе предлагается алгоритм ансамблирования рассмотренных методов поиска объекта на изображении, учитывающий их сильные и слабые стороны и характерные особенности. К таким особенностям можно отнести, например, применимость к обнаружению и распознаванию лиц, к динамическим или статическим изображениям, чувствительность к шуму, цвету и другим характеристикам изображения.

**Формально** идея комбинирования заключается в следующем. Пусть даны *k* простых методов машинного обучения. Тогда *Mi* (*i=1,...,k) –*рассматриваемые методы, комбинируются с весами *wi*. Коэффициенты *wi*. определяются исходя из условий задачи и выявленных в настоящей работе характеристик методов:

,

где *ci* – коэффициенты, определяемые из особенностей метода, *N* – количество различных коэффициентов.

Полученный мета-алгоритм будет иметь вид

,

где оператор ◦ обозначает некоторую связь между методами. В качестве связи может быть использован метод голосования по всем или нескольким методам, вычисление среднего или последовательное применение одних методов на основе ответов других.

Эффективность результирующей модели вычисляется как отношение качества работы комбинации методов к усреднённому качеству моделей по отдельности:

,

где *E* – эффективность комбинации моделей, *SMi* – качество работы *i*-ого метода.