



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 «ФАКУЛЬТЕТ»

КАФЕДРА «КАФЕДРА»

РАСЧЕТНО-ПОЯСНИТЕЛЬНАЯ ЗАПИСКА
К ВЫПУСКНОЙ КВАЛИФИКАЦИОННОЙ РАБОТЕ
НА ТЕМУ:

«Крутая тема»

Студент группы <ГРУППА>

(Подпись, дата)

<Студент>

(И.О. Фамилия)

Руководитель ВКР

(Подпись, дата)

<Руководитель>

(И.О. Фамилия)

Нормоконтролер

(Подпись, дата)

<Нормоконтролер>

(И.О. Фамилия)

РЕФЕРАТ

Расчетно-пояснительная записка 20 с., 1 рис., 1 табл., 8 ист., 0 прил.

КЛЮЧЕВЫЕ СЛОВА

отвлеченный водитель, распознавание по фото, классификация изображений.

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	5
1 Задача классификации	6
2 Обзор методов классификации	6
2.1 Методы на основе глубокого обучения	6
2.2 Методы основанные на экземплярах	8
2.3 Методы на основе деревьев принятия решений	12
3 Сравнение методов классификации	16
3.1 Формулировка критериев сравнения методов классификации	16
3.2 Сравнительный анализ методов классификации	16
ЗАКЛЮЧЕНИЕ	19
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	20

ВВЕДЕНИЕ

В современном мире автомобили стали неотъемлемой частью повседневной жизни людей, и безопасность на дорогах становится все более актуальной проблемой.

Согласно статистике[1] ежегодно около 3000 человек погибают в автокатастрофах из-за невнимательного вождения.

С помощью компьютерных технологий можно распознавать действия водителя по фотографии. Такие технологии могут быть полезны как для уведомления водителя, так и для выявления причины ДТП.

Цель работы — проанализировать существующие методы распознавания действий водителя по фото.

Для достижения поставленной цели требуется выполнить следующий задачи:

- провести обзор существующих методов распознавания действий человека за рулем автомобиля по фото;
- классифицировать существующие методы распознавания действий водителя по фото;
- сформулировать критерии сравнения методов распознавания действий водителя по фото;
- провести сравнительный анализ рассмотренных методов.

1 Задача классификации

Задачи классификации — это задачи, в которых объект должен быть отнесен к одному из n классов на основе индекса сходства его характеристик с каждым классом. Под классами понимается набор подобных объектов. Объекты считаются похожими на основе совпадающих характеристик, таких как цвет, форма, размер и т.д. Классы определяются на основе их уникальных меток. В частности, задача «распознавания действий человека за рулем автомобиля по фото» является задачей классификации.

Для решения задачи классификации используются такие методы как:

- методы на основе глубокого обучения;
- методы основанные на экземплярах;
- методы на основе деревьев принятия решений.

2 Обзор методов классификации

2.1 Методы на основе глубокого обучения

Методы на основе глубокого обучения представляют собой один из подходов для решения задачи классификации изображений. Они основаны на использовании глубоких нейронных сетей.

С точки зрения классификации изображений наиболее эффективным методом являются сверточные нейронные сети (CNN)[2]. Их способность работать непосредственно с пикселями изображения позволяет извлекать сложные признаки на разных уровнях абстракции, что делает их хорошо подходящими для большинства задач классификации изображений. Кроме того, многократное применение сверток и слоев пулинга позволяет CNN эффективно обрабатывать изображения различных размеров.

Общий алгоритм сверточных нейронных сетей включает в себя несколько ключевых этапов.

- 1) Сверточный слой. Входные данные подвергаются свертке с несколькими фильтрами для извлечения различных признаков. Каждый фильтр обнаруживает определенные характеристики, такие как края, углы и текстуры. Это позволяет сети автоматически изучать различные уровни абстракции в изображениях.
- 2) *Pooling* слой. После каждого сверточного слоя обычно следует *pooling* слой, который уменьшает размер признаков карт, сохраняя наиболее важные признаки. Это способствует снижению количества параметров и вычислительную сложность модели, а также увеличивает инвариантность к масштабированию и смещению объектов на изображении.
- 3) Полносвязанный слой. После нескольких сверточных слоев и слоев пулинга, выходы подаются на полносвязные слои, которые выполняют классификацию на основе полученных признаков.
- 4) Функция активации. Между слоями нейронов используется функция активации, например ReLU (Rectified Linear Unit)[3], которая добавляет нелинейность к модели.
- 5) Обучение. Все параметры (веса и смещения) в сети обучаются с использованием методов градиентного спуска[4] и обратного распространения ошибки[5] на больших объемах размеченных данных.

Преимущества методов на основе глубокого обучения.

- 1) Автоматическое извлечение признаков без необходимости ручной настройки алгоритмов.
- 2) Глубокие нейронные сети обладают способностью моделировать сложные зависимости между пикселями изображения и их классификацией.
- 3) Высокая точность классификации: методы на основе глубокого обучения достигают высоких показателей точности в классификации изображений, особенно при использовании больших наборов данных и правильной настройке гиперпараметров модели.

Однако методы на основе глубокого обучения также имеют некоторые ограничения:

- требовательность к вычислительным ресурсам;
- необходимость большого количества размеченных данных;
- общность модели, некорректные экземпляры в наборе обучающих данных могут привести к переобучению.

2.2 Методы основанные на экземплярах

Методы основанные на экземплярах представляют собой один из подходов решения задачи классификации. Данный подход предполагает, что похожие входные данные приводят к похожим результатам, что является важным предположением для классификации. Наиболее распространенные примеры методов основанных на экземплярах – это k-ближайших соседей (k-Nearest Neighbors, KNN) и метод опорных векторов (Support Vector Machines, SVM).

Метод опорных векторов

Метод опорных векторов является одним из наиболее распространенных методов классификации и регрессии. Он основан на идее нахождения гиперплоскости максимальной ширины, разделяющей два класса данных. SVM может быть применен для решения задач как бинарной, так и многоклассовой классификации. На рисунке 1 представлена визуализация алгоритма SVM.

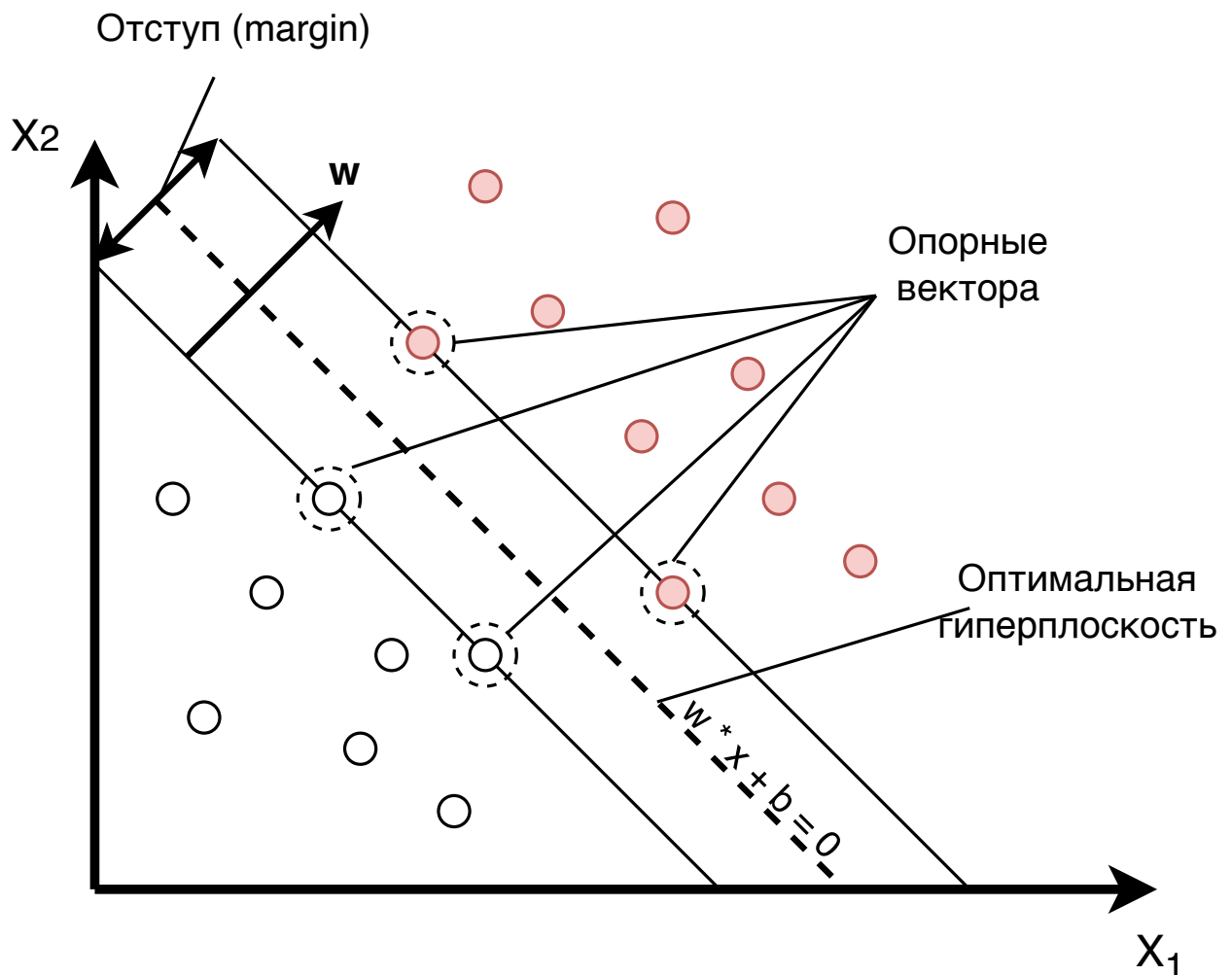


Рисунок 1 – Визуализация алгоритма SVM.

SVM оперирует в многомерном пространстве данных, где каждое изображение представлено в виде вектора признаков. Основным шагом в методе SVM – это построение оптимальной разделяющей гиперплоскости. Этот процесс сводится к решению оптимизационной задачи с целью максимизации отступа (margin) между обучающими объектами и разделяющей гиперплоскостью.

Уравнение гиперплоскости в пространстве можно представить формулой

$$w^T x + b = 0. \quad (1)$$

где w — вектор нормали к гиперплоскости;
 x — точка;
 b — смещение.

Расстояние от точки x до гиперплоскости рассчитывается формулой

$$d = \frac{|w^T x + b|}{|w|}. \quad (2)$$

Отступ(margin) равен расстоянию между двумя параллельными плоскостями, проходящими через опорные вектора, его можно рассчитать как разницу расстояний двух опорных векторов до гиперплоскости.

Цель состоит в том, чтобы найти оптимальные значения w и b , которые определяют гиперплоскость, удовлетворяя при этом определенным ограничениям. Запас должен быть максимальным, что означает минимизацию величины w .

В силу того, что весь процесс нахождения гиперплоскости сводится к нахождению скалярных произведений векторов, переход к размерности большей чем 2 осуществляется путем замены скалярного произведения на так называемое ядро – функцию, которая является скалярным произведением в пространстве текущей размерности.

Вектор признаков для SVM обычно получается путем извлечения характеристик изображений. Эти характеристики могут включать в себя различные аспекты изображения, такие как цветовые гистограммы, текстурные дескрипторы, угловые точки или другие характеристики, которые могут быть выделены в процессе предобработки изображений.

Например, можно использовать методы извлечения признаков, такие как гистограммы направленных градиентов (HOG), которые описывают распределение градиентов яркости в изображении.

Рассмотренный метод имеет ряд преимуществ и недостатков при решении задачи классификации изображений.

Преимущества:

- высокая способность обобщения;
- устойчивость к переобучению;
- возможность работы с небольшими наборами данных.

Недостатки:

- требует сбора признаков;
- высокая вычислительная сложность в случае большого количества признаков;
- чувствительность к масштабированию признаков.

Метод k -ближайших соседей

Метод k -ближайших соседей(KNN) применяется для решения задач классификации и регрессии, при этом, он основан на предпосылке о том, что близкие объекты более вероятно принадлежат к одному классу.

Последовательность действий алгоритма KNN.

1. Определение параметра k , который представляет собой число ближайших соседей, которые будут учитываться при определении класса для нового объекта.
2. Расчет расстояния от нового объекта до всех остальных в обучающем наборе. Это расстояние может быть рассчитано разными способами, зачастую используется евклидово расстояние[6] или манхэттенское расстояние[7].
3. Сортировка обучающего набора по возрастанию расстояния до нового объекта.
4. Выбор первых k объектов из отсортированного набора. Класс объекта определяется на основе большинства классов этих k соседей.

При очень маленьком k модель может стать чрезмерно чувствительной к шумам в данных, при слишком большом k – классификация может «смазываться» .

Важным шагом также является правильный выбор метрики расстояния, которая может значительно влиять на результаты классификации.

Для классификации изображений этим методом нужно преобразовать изображение в числовой вектор. Обработка большого количества изображений может потребовать большое количество вычислительных ресурсов и времени.

Рассмотренный метод имеет ряд преимуществ и недостатков при решении задачи классификации изображений.

Преимущества:

- не требуется обучение, так как хранит набор данных;
- адаптация к изменяющимся данным, поскольку алгоритм не требует повторного обучения;
- устойчивость к переобучению.

Недостатки:

- высокая вычислительная сложность;
- зависимость от выбора метрики и количества соседей;
- требует сбора признаков.

2.3 Методы на основе деревьев принятия решений

Методы на основе деревьев принятия решений представляют собой алгоритмы классификации, которые основываются на принципе создания дерева, где каждый узел представляет собой решающее правило, а каждая ветвь – возможный исход этого правила. В рамках задачи классификации изображений можно применить следующие методы.

Дерево принятия решений

Дерево принятия решений (Decision Tree) – это базовый метод, где каждый узел дерева представляет собой один из признаков изображения, а каждая ветвь – пороговое значение этого признака. Классификация производится путем прохождения по дереву, начиная с корневого узла и переходя на следующий узел в зависимости от значения признака, пока не будет достигнут листовой узел, который представляет конечный класс.

Алгоритм простого дерева принятия решений.

1. Выбор признака. Из множества признаков выбирается тот, который наилучшим образом разделяет данные на разные классы. Обычно используется мера неоднородности, такая как критерий Джини[8] или энтропия, для определения приоритетности выбора признака.

2. Разделение узлов. Данные разделяются на две (или более) части на основе выбранного признака. Каждый раздел создает дочерний узел в дереве.
3. Повторение процесса. Процесс выбора признака и разделения данных повторяется для каждого дочернего узла, пока не будет выполнен критерий останова (например, достигнута определенная глубина дерева или узлы содержат данные одного класса).
4. Листовые узлы. Когда выполнен критерий останова, узлы становятся листьями и представляют собой окончательные классификации.
5. Прогноз. При поступлении нового наблюдения оно проходит по дереву, используя значения его признаков, и попадает в соответствующий лист, предсказывая тем самым класс данного наблюдения.

Простое дерево принятия решений часто служит строительным блоком для более продвинутых ансамблевых методов, таких как случайные леса или градиентный бустинг. В ансамблевых методах несколько слабых обучаемых моделей, таких как простые деревья решений, обучаются на разном подмножестве данных и затем комбинируются для улучшенной предсказательной способности.

Случайный лес

Алгоритм случайный лес (Random Forest) – это ансамблевый метод машинного обучения, который используется для классификации, регрессии и других задач. Он строится на основе концепции, где множество деревьев решений объединяются для достижения более точных прогнозов.

Алгоритм случайного леса.

1. Выбор случайной подвыборки. Из общего набора данных случайным образом выбирается подмножество данных (с повторениями).
2. Построение деревьев решений. Для каждой подвыборки строится отдельное дерево решений. Процесс построения дерева аналогичен алгоритму дерева решений.
3. Голосование. Когда поступает новое наблюдение, оно проходит через все построенные деревья, и каждое дерево дает свой прогноз. Затем происходит голосование, и классификация определяется большинством голосов.
4. Использование случайности. Random Forest внедряет случайность в процесс построения деревьев, что позволяет каждому дереву быть уникальным. Это помогает уменьшить переобучение и повысить обобщающую способность модели.
5. Усреднение прогнозов. Поскольку Random Forest состоит из множества деревьев, его прогнозы усредняются, что обычно приводит к улучшению точности по сравнению с одним деревом решений.

Важно отметить, что Random Forest обладает рядом дополнительных преимуществ, таких как возможность оценки важности признаков и автоматическое управление переобучением. Кроме того, он позволяет эффективно обрабатывать большие объемы данных с высокой размерностью признаков.

Однако, существует несколько недостатков, связанных с Random Forest, таких как более высокое время обучения по сравнению с одним деревом решений, потребность в тщательной настройке гиперпараметров и сложность интерпретации по сравнению с одиночным деревом решений.

Градиентный бустинг

Градиентный бустинг (Gradient Boosting) – это метод машинного обучения, который используется для построения прогностических моделей, таких как регрессия и классификация. Он основан на комбинировании слабых algo-

ритмов обучения (например, деревьев решений) в композицию, которая обладает более высокой предсказательной способностью.

Алгоритм градиентного бустинга.

1. Инициализация модели. Строится начальная модель, которая может быть, например, классом с наибольшим количеством экземпляров для задачи классификации.
2. Вычисление остатков. Вычисление остатков или ошибок предсказания начальной модели для каждого обучающего примера. Эти остатки становятся новой целевой переменной для следующей модели.
3. Обучение слабой модели. Эта модель пытается предсказать остатки, полученные на предыдущем шаге, обычно для этого используются деревья решений небольшой глубины.
4. Вычисление коэффициента, на который мы будем умножать предсказания новой модели, чтобы получить более точное предсказание композиции. Этот коэффициент вычисляется с использованием градиентного спуска или другого оптимизационного метода.
5. Обновление композиции моделей, путем добавления новой модели, умноженной на вычисленный коэффициент.
6. Повторение шагов 2-5 пока не достигнем определенного критерия останова, такого как количество итераций или улучшение метрики качества.

Градиентный бустинг обладает высокой предсказательной способностью и способен автоматически отбирать признаки, однако он чувствителен к переобучению и требует тщательной настройки гиперпараметров.

3 Сравнение методов классификации

3.1 Формулировка критериев сравнения методов классификации

Для сравнения рассмотренных алгоритмов можно выделить следующие критерии.

1. Автоматическое создание признаков: этот критерий обозначает, способен ли выбранный метод автоматически генерировать признаки из исходных данных, или же требуется вручную задавать признаки. Генерация означает процесс преобразования исходных данных в более усовершенствованный формат, который может улучшить эффективность анализа и классификации.
2. Автоматическое извлечение признаков: относится к способности метода автоматически определять и использовать наиболее важные или релевантные аспекты данных для классификации.
3. Устойчивость к шуму: означает способность метода справляться с ошибками, неточностями или нежелательной информацией в данных.
4. Точность в сложных задачах классификации: измеряет, насколько верно метод определяет классы в трудных или многоклассовых задачах.
5. Обработка сложных признаков: оценивает, насколько эффективно метод может обрабатывать сложные признаки, такие как непрерывные, категориальные или временные.

3.2 Сравнительный анализ методов классификации

На основе глубоких нейронных сетей

1. Автоматическое создание/извлечение признаков: данные методы работают напрямую с изображениями, без необходимости создавать и извлекать признаки, это происходит автоматически.
2. Устойчивость к шуму: средняя, в сравнении с другими рассмотренными методами, так как шум может приводить к изменениям в распределении данных и переобучению.

3. Точность в сложных задачах классификации: высокая, в сравнении с другими рассмотренными методами, это связано с их способностью автоматически выявлять и анализировать признаки на всех уровнях абстракций.
4. Обработка сложных признаков: глубокие нейронные сети способны обрабатывать сложные признаки, так как они автоматически извлекают признаки на разных уровнях сложности.

На основе экземпляров

1. Автоматическое создание/извлечение признаков: зависят от разработанных и извлеченных человеком признаков для обучения.
2. Устойчивость к шуму: сравнительно низкая, поскольку данные методы зависят от близости измерений между точками данных.
3. Точность в сложных задачах классификации: средняя, в силу того, что могут возникнуть трудности в случаях, когда классы не могут быть легко разделены в пространстве признаков.
4. Обработка сложных признаков: данный метод имеет малую точность на больших объемами данных и низким уровнем сигнала, которые обычно связаны со сложными признаками.

На основе деревьев принятия решений

1. Автоматическое создание/извлечение признаков: зависят от разработанных и извлеченных человеком признаков для обучения.
2. Устойчивость к шуму: сравнительно высокая, поскольку данные методы делают предположения о структуре данных и не полагаются на расстояние между точками данных для классификации.
3. Точность в сложных задачах классификации: средняя, методы на основе деревьев принятия решений имеют сравнимую точность с методами на основе экземпляров, но уступают методам на основе глубоких нейронных сетей.
4. Обработка сложных признаков: обычные деревья принятия решений,

как правило, не очень хорошо справляются с сложными признаками, потому что их стратегия разделения может привести к переобучению. Однако более современные алгоритмы, такие как случайные леса и градиентный бустинг, могут выполнять иерархическое или несколько уровневое разбиение, что позволяет им лучше справляться со сложной структурой данных.

Сравнительная характеристика рассмотренных методов представлена в таблице 1.

Таблица 1 – Таблица сравнение методов.

Метод	На основе глубоких нейронных сетей	На основе экземпляров	На основе деревьев принятия решений
Обработка сложных признаков	+	—	±
Точность в сложных задачах классификации	высокая	средняя	средняя
Устойчивость к шуму	средняя	низкая	высокая
Автоматическое извлечение признаков	+	—	—
Автоматическое создание признаков	+	—	—

ЗАКЛЮЧЕНИЕ

В ходе научно-исследовательской работы были изучены существующие методы решения задачи классификации, частным случаем которой является задача распознавания действий водителя по фото. Можно сделать вывод, что метод на основе глубоких нейронных сетей является наиболее эффективным для решения поставленной задачи.

В ходе выполнения данной работы были выполнены следующие задачи:

- проведен обзор существующих методов распознавания действий человека за рулем автомобиля по фото;
- классифицированы существующие методы распознавания действий водителя по фото;
- сформулированы критерии сравнения методов распознавания действий водителя по фото;
- проведет сравнительный анализ рассмотренных методов.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Distracted Driving Statistics [Электронный ресурс]. – Режим доступа: <https://www.forbes.com/advisor/legal/auto-accident/distracted-driving-statistics>, свободный – (15.10.2023)
2. Convolutional neural networks for image classification. [Электронный ресурс]. – Режим доступа: https://www.researchgate.net/publication/325770590_Convolutional_neural_networks_for_image_classification, свободный – (10.12.2023)
3. Deep Learning using Rectified Linear Units (ReLU). [Электронный ресурс]. – Режим доступа: https://www.researchgate.net/publication/323956667_Deep_Learning_using_Rectified_Linear_Units_ReLU, свободный – (10.12.2023)
4. What is gradient descent? [Электронный ресурс]. – Режим доступа: <https://www.ibm.com/topics/gradient-descent>, свободный – (10.12.2023)
5. Backpropagation neural networks. [Электронный ресурс]. – Режим доступа: <https://www.sciencedirect.com/science/article/abs/pii/016974399380052J>, свободный – (10.12.2023)
6. Euclidean Distance. [Электронный ресурс]. – Режим доступа: <https://www.sciencedirect.com/topics/engineering/euclidean-distance>, свободный – (10.12.2023)
7. Manhattan distance. [Электронный ресурс]. – Режим доступа: <https://xlinux.nist.gov/dads/HTML/manhattanDistance.html>, свободный – (10.12.2023)
8. Analysis of Depth of Entropy and GINI Index Based Decision Trees. [Электронный ресурс]. – Режим доступа: https://www.researchgate.net/publication/357777690_Analysis_of_Depth_of_Entropy_and_GINI_Index_Based_Decision_Trees_for_Predicting_Diabetes, свободный – (10.12.2023)