

1. ОПИСАНИЕ ОБЪЕКТА УПРАВЛЕНИЯ

1.1 Системы и методы распознавания лиц

Распознавание лиц – практическое применение теории распознавания образов, задача которого состоит в автоматической локализации лица на изображении, а также в идентификации персоны по лицу [1]. Поэтому для начала необходимо понимание теории распознавания образов.

Сама теория распознавания образов — это раздел информатики и смежных дисциплин, развивающий основы и методы классификации и идентификации предметов, явлений, процессов, сигналов, ситуаций и т. п. объектов, которые характеризуются конечным набором некоторых свойств и признаков. Необходимость в таком распознавании возникает в самых разных областях - от военного дела и систем безопасности до оцифровки аналоговых сигналов. Проблема распознавания образов приобрела выдающееся значение в условиях информационных перегрузок, когда человек не справляется с линейно-последовательным пониманием поступающих к нему сообщений, в результате чего его мозг переключается на режим одновременности восприятия и мышления, которому свойственно такое распознавание[2].

Алгоритмы распознавания образов зависят от типа вывода метки, от того, является ли обучение контролируемым или неконтролируемым, а также от того, является ли алгоритм статистическим или нестатистическим по своей природе. Статистические алгоритмы можно далее разделить на генеративные и дискриминационные.

Алгоритмы можно разделить на следующие категории по основе их общих подходов и применений в распознавании образов:

1. Статическое распознавание образов, к ней относят:

- Байесовские классификаторы: используют теорему Байеса для расчета вероятности определенного класса с учетом наблюдаемых данных.
- Скрытые марковские модели (HMM): моделируют последовательности наблюдений и часто используются для распознавания речи и жестов.

2. Нейронные сети, к ним относят:

- Нейронные сети прямого распространения: состоят из слоев взаимосвязанных узлов и широко используются для распознавания изображений и речи.

- Сверточные нейронные сети (CNN): особенно эффективны для задач распознавания изображений за счет использования сверточных слоев для обнаружения пространственных закономерностей.

- Рекуррентные нейронные сети (RNN): подходят для данных последовательностей, таких как временные ряды или обработка естественного языка.

3. Деревья решений:

- C4.5 и CART (деревья классификации и регрессии): методы рекурсивного разделения, которые разбивают данные на подмножества на основе значений признаков.

4. Обучение на основе экземпляров:

- k-ближайшие соседи (k-NN): классифицирует точки данных на основе класса большинства их k ближайших соседей в пространстве признаков.

5. Кластеризации:

- k-средние: разделяет данные на k кластеров на основе сходства.
- Иерархическая кластеризация: строит иерархию кластеров, часто представленную в виде дендрограммы.

6. Машины опорных векторов

- Kernel SVM: расширяет базовую SVM за счет использования функции ядра для отображения входного пространства в многомерное пространство признаков, где данные могут быть более разделимыми.

- SVM с мягкими границами: в отличие от стандартной SVM с жесткими границами, которая направлена на идеальное разделение классов, SVM с мягкими границами допускает некоторые неправильные классификации для нахождения более обобщенной границы решения.

7. Ансамблевое обучение:

- Случайные леса: строит несколько деревьев решений и объединяет их прогнозы для повышения точности и уменьшения переобучения.

- Машины повышения градиента: последовательно создает серию слабых учеников, каждый из которых исправляет ошибки своего предшественника.

8. Генетические алгоритмы:

- Эволюционные алгоритмы: используйте принципы естественного отбора для разработки решений проблемы на протяжении нескольких поколений.

- Оптимизация роя частиц (PSO) — PSO — это алгоритм популяционной оптимизации, вдохновленный социальным поведением стад птиц или стайной рыбы.

9. Снижение размерности:

- Анализ главных компонент (PCA): уменьшает размерность данных, сохраняя при этом их дисперсию.

- t-распределенное стохастическое внедрение соседей (t-SNE): визуализирует многомерные данные, представляя аналогичные экземпляры с соседними точками.

10. Марковские модели:

- Модели цепей Маркова: представляют системы с последовательностью состояний, где переходы между состояниями являются вероятностными.

- Скрытые полумарковские модели (HSMM). HSMM расширяют традиционные скрытые марковские модели, допуская состояния переменной продолжительности.

Примеры систем:

- Amazon Rekognition: это сервис обработки изображений и видео от Amazon Web Services. Он предоставляет API для распознавания лиц, а также для анализа эмоций, определения пола и возраста, идентификации знаменитостей и других функций.

- Face++: это платформа и API для распознавания лиц, разработанная китайской компанией Megvii. Она обладает высокой точностью и может использоваться для идентификации лиц, анализа эмоций, определения возраста и пола и других задач.

- Microsoft Azure Face API: это облачный сервис от Microsoft, предоставляющий API для распознавания лиц. Он позволяет идентифицировать лица на фотографиях и в видео, а также проводить анализ эмоций, определять возраст и пол и выполнять другие задачи.

- Google Cloud Vision API: это сервис от Google, который включает в себя функции распознавания лиц. Он может определять лица на изображениях и проводить анализ эмоций, определять возраст и пол и выполнять другие задачи с использованием машинного обучения.

- OpenFace: это открытое программное обеспечение для распознавания лиц, разработанное компанией Carnegie Mellon University. Оно предоставляет набор инструментов и библиотек для обнаружения и идентификации лиц на изображениях и в видео.

- Kairos: это платформа для распознавания лиц, которая предоставляет API для идентификации лиц, анализа эмоций, определения возраста и пола и других функций. Она может использоваться в различных отраслях, включая безопасность, маркетинг и развлечения.
- IBM Watson Visual Recognition: это сервис от IBM, который включает в себя функции распознавания лиц. Он может определять лица на изображениях и в видео, а также проводить анализ эмоций, определять возраст и пол и выполнять другие задачи с использованием искусственного интеллекта.
- FaceID: это система распознавания лиц, разработанная компанией Apple. Она используется для разблокировки устройств, авторизации платежей и других задач, связанных с идентификацией лиц.

1.2 Нейронные сети и их виды

Нейронные сети это ветвь моделей машинного обучения , которые построены с использованием принципов нейронной организации, открытых коннекционизмом в биологических нейронных сетях, составляющих мозг животных .

Нейронный сети основаны на наборе связанных единиц или узлов, называемых искусственными нейронами , которые в общих чертах моделируют нейроны биологического мозга. Каждое соединение, подобно синапсам в биологическом мозге, может передавать сигнал другим нейронам. Искусственный нейрон получает сигналы, затем обрабатывает их и может передавать сигналы подключенным к нему нейронам. «Сигналом» соединения является действительное число , а выходной сигнал каждого нейрона вычисляется с помощью некоторой нелинейной функции суммы его входных сигналов. Соединения называются ребрами . Нейроны и ребра обычно имеют вес , который корректируется по мере обучения. Вес увеличивает или уменьшает силу сигнала при соединении. Нейроны могут иметь порог, при котором сигнал отправляется только в том случае, если совокупный сигнал пересекает этот порог.

Искусственный нейрон являются упрощённой моделью естественного нейрона, а математически обычно представляют как некоторую нелинейную функцию от единственного аргумента — линейной комбинации всех входных сигналов. Данную функцию называют функцией активации или функцией срабатывания, передаточной функцией. Полученный результат посылается на единственный выход.

Математически нейрон представляет собой взвешенный сумматор, единственный выход которого определяется через его входы и матрицу весов следующим образом:

$$y = f(u), \text{ где } u = \sum_{i=1}^n \omega_i x_i + \omega_0 x_0$$

Здесь x_i и ω_i — соответственно сигналы на входах нейрона и веса входов, функция называется индуцированным локальным полем, а $f(u)$ — передаточной функцией. Возможные значения сигналов на входах нейрона считают заданными в интервале $[0,1]$. Они могут быть либо дискретными (0 или 1), либо аналоговыми. Дополнительный вход x_0 и соответствующий ему вес ω_0 используются для инициализации нейрона.

Передаточная функция определяет зависимость сигнала на выходе нейрона от взвешенной суммы сигналов на его входах.

Связи, по которым выходные сигналы одних нейронов поступают на входы других, часто называют синапсами по аналогии со связями между биологическими нейронами. Каждая связь характеризуется своим весом. Связи с положительным весом называются возбуждающими, а с отрицательным — тормозящими.

На рисунке 1 представлена схема искусственного нейрона, состоящая из следующих частей: 1 — нейроны, выходные сигналы которых поступают на вход; 2 — сумматор входных сигналов; 3 — вычислитель передаточной функции; 4 — нейроны, на входы которых подаётся выходной сигнал; ω_i — веса входных сигналов.

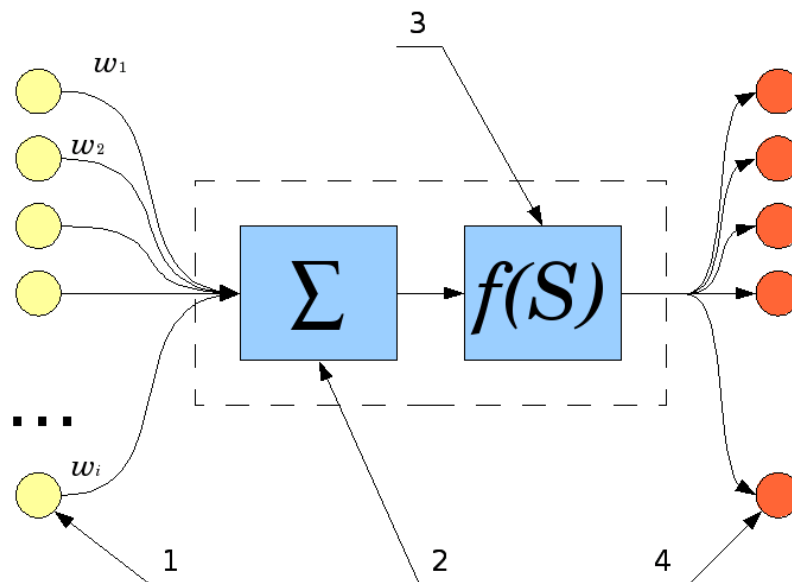


Рисунок 1 – Схема искусственного нейрона

Обычно нейроны объединяются в слои. Разные слои могут выполнять разные преобразования на своих входах. Сигналы передаются от первого слоя (входной слой) к последнему слою (выходной слой), возможно, после многократного прохождения слоев.

Нейронные сети обучаются (или обучаются) путем обработки примеров, каждый из которых содержит известные «входные данные» и «результат», образуя между ними взвешенные по вероятности ассоциации, которые хранятся в структуре данных самой сети. Обучение нейронной сети по заданному примеру обычно проводится путем определения разницы между обработанным выходным сигналом сети (часто прогнозом) и целевым выходным сигналом. Эта разница и есть ошибка. Затем сеть корректирует свои взвешенные ассоциации в соответствии с правилом обучения и использованием этого значения ошибки. Последовательные корректировки заставят нейронную сеть выдавать выходные данные, которые все больше похожи на целевые выходные данные. После достаточного количества таких корректировок обучение может быть прекращено по определенным критериям. Это форма контролируемого обучения.

Такие системы «учатся» выполнять задачи, рассматривая примеры, как правило, без программирования правил для конкретных задач. Например, при распознавании изображений они могут научиться распознавать изображения, на которых есть кошки, анализируя примеры изображений, которые были вручную помечены как «кот» или «нет кота», и используя результаты для идентификации кошек на других изображениях. Они делают это, не зная заранее о кошках, например, о том, что у них есть шерсть, хвосты,

усы и кошачьи морды. Вместо этого они автоматически генерируют идентифицирующие характеристики на основе обрабатываемых примеров.

Классификация нейронных сетей по характеру связей:

- Прямое распространения — характеризуются направлением потока информации между ее слоями. Его поток является однонаправленным, что означает, что информация в модели течет только в одном направлении — вперед — от входных узлов через скрытые узлы (если таковые имеются) и к выходным узлам без каких-либо циклов или петель. В отличие от рекуррентных нейронных сетей, которые имеют двунаправленный поток.
- Радиально-базисных функций — использует радиальные базисные функции в качестве функций активации. Выход сети представляет собой линейную комбинацию радиальных базисных функций входов и параметров нейрона.
- Глубокая сеть доверия - один из типов глубинных нейронных сетей, состоящая из нескольких скрытых слоев, в которых нейроны внутри одного слоя не связаны друг с другом, но связаны с нейронами соседнего слоя.
- Рекуррентные — характеризующиеся направлением потока информации между ее слоями. В отличие от однонаправленной нейронной сети прямого распространения, это двунаправленная искусственная нейронная сеть, что означает, что она позволяет выходным данным некоторых узлов влиять на последующий ввод в те же узлы.

1.3 Свёрточные нейронные сети в распознавании лиц

Свёрточная нейронная сеть специальная архитектура искусственных нейронных сетей, предложенная Яном Лекуном в 1988 году. Структура сети — однонаправленная (без обратных связей), принципиально многослойная.

Работа свёрточной нейронной сети интерпретируется как переход от конкретных особенностей изображения к более абстрактным деталям, и далее к ещё более абстрактным деталям вплоть до выделения понятий высокого уровня. При этом сеть самонастраивается и вырабатывает сама необходимую иерархию абстрактных признаков (последовательности карт признаков), фильтруя маловажные детали и выделяя существенное.

Сверточные сети состоят слоёв, которые разделены на 3 типа:

- слой свёртка;

- слой пулинга.

Слой свёртки - это основной блок свёрточной нейронной сети. Слой свёртки включает в себя для каждого канала свой фильтр, ядро свёртки которого обрабатывает предыдущий слой по фрагментам (суммируя результаты поэлементного произведения для каждого фрагмента). Весовые коэффициенты ядра свёртки (небольшой матрицы) неизвестны и устанавливаются в процессе обучения. Скалярный результат свёртки попадает в функцию активации, которая представляет собой некую нелинейную функцию.

Слой пулинга(иначе подвыборки, субдискретизации) - представляет собой нелинейное уплотнение карты признаков, при этом группа пикселей (обычно размера 2×2) уплотняется до одного пикселя, проходя нелинейное преобразование. Используются функции максимума, минимума и среднего значения. Преобразования затрагивают непересекающиеся прямоугольники или квадраты, каждый из которых ужимается в один пиксель, при этом выбирается пиксель, имеющий максимальное значение. Операция пулинга позволяет существенно уменьшить пространственный объём изображения. Пулинг интерпретируется так: если на предыдущей операции свёртки уже были выявлены некоторые признаки, то для дальнейшей обработки настолько подробное изображение уже не нужно, и оно уплотняется до менее подробного.

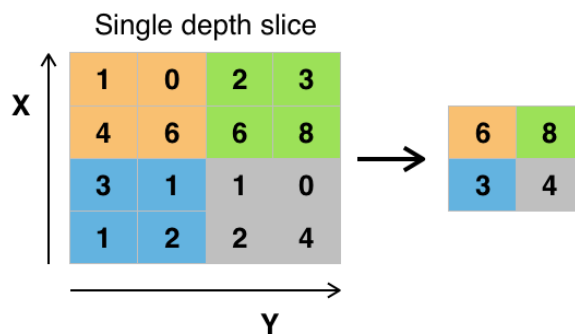


Рисунок - Пулинг с функцией максимума и фильтром 2×2 с шагом 2

Вот несколько конкретных технологий, которые используют сверточные нейронные сети для распознавания лиц:

- FaceNet - это технология, разработанная Google, которая использует сверточные нейронные сети для создания уникальных векторных представлений лиц. Она позволяет сравнивать и идентифицировать лица на основе этих векторных представлений.
- DeepFace - это технология, разработанная Facebook, которая также использует сверточные нейронные сети для распознавания лиц. Она

способна определять и идентифицировать лица на фотографиях с высокой точностью.

- OpenFace - это открытая библиотека, которая использует сверточные нейронные сети для распознавания лиц. Она предоставляет возможность извлекать признаки лиц и сравнивать их для идентификации.
- Dlib - это библиотека машинного обучения, которая включает в себя реализацию сверточных нейронных сетей для распознавания лиц. Она предоставляет инструменты для обнаружения лиц, извлечения признаков и идентификации.