

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 несколько множеств, характеризующиеся различными методами:

- задача классификации, который в свою очередь можно разделить на параметрический(пример: линейный дискриминантный анализ) и не параметрический(пример: свёрточные нейронные сети);
- кластерный анализ(пример: корреляционная кластеризация);
- ансамблевое обучение(пример: усреднение по ансамблю);
- регрессионный анализ(пример: анализ главных компонент).

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.

Исходя из требований технического задания был выбран способ распознавания лиц с помощью нейронных сетей, а именно свёрточных нейронных сетей.

1.2 Нейронные сети и их виды

Нейронные сети – это ветвь моделей машинного обучения, которые построены с использованием принципов нейронной организации, открытых соединениях в биологических нейронных сетях, составляющих мозг животных [3].

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. 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.

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 $f(u)$ является передаточной функцией, определяющая зависимость сигнала на выходе нейрона от взвешенной суммы сигналов на его входах. 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.

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.

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.

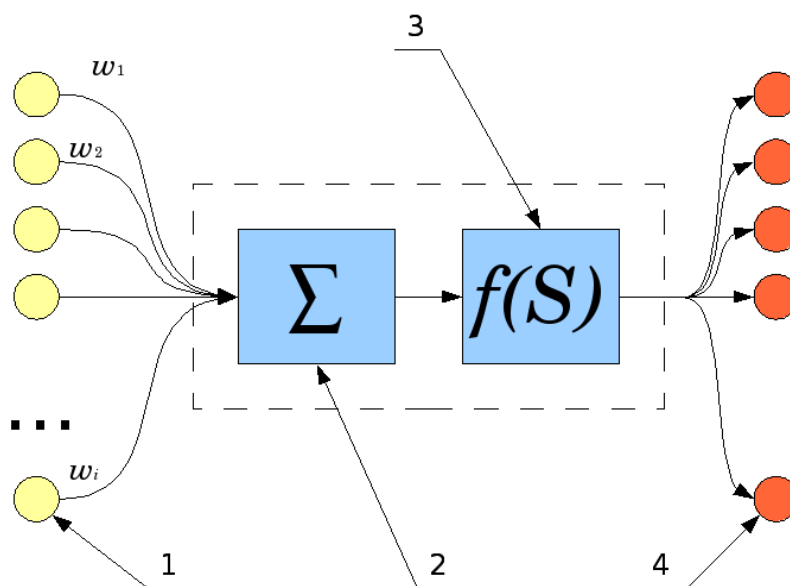


Рисунок 1 – Схема искусственного нейрона

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.

Классификация нейронных сетей по характеру связей:

- прямого 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. Примером сети прямого распространения являются свёрточные нейронные сети.

- 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 это однонаправленная (без обратных связей), принципиально многослойная [4].

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 2 типа:

- слой свёртки;
- 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, а 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. Слой пулинга всегда следует за слоем свёртки.

В свёрточной нейронной сети в операции свёртки используется лишь ограниченная матрица весов небольшого размера, которую «двигают» по всему обрабатываемому слою (в самом начале – непосредственно по входному изображению), формируя после каждого сдвига сигнал активации для нейрона следующего слоя с аналогичной позицией. То есть для различных нейронов выходного слоя используется одна и та же матрица весов, которую также называют ядром свёртки. Её интерпретируют как графическое кодирование какого-либо признака, например, наличие наклонной линии под определённым углом. Тогда следующий слой, получившийся в результате операции свёртки такой матрицей весов, показывает наличие данного признака в обрабатываемом слое и её координаты, формируя так называемую карту признаков. Естественно, в свёрточной нейронной сети набор весов не один, а набор, кодирующий элементы изображения (например линии и дуги под разными углами). При этом такие ядра свёртки не закладываются исследователем заранее, а формируются самостоятельно путём обучения сети классическим методом обратного распространения ошибки. Проход каждым набором весов формирует свой собственный экземпляр карты признаков, делая нейронную сеть многоканальной (много независимых карт признаков на одном слое). При переборе слоя матрицей весов её передвигают обычно не на полный шаг (размер этой матрицы), а на небольшое расстояние. Так, например, при размерности матрицы весов 5×5 её сдвигают на один или два нейрона (пикселя) вместо пяти, чтобы не «перешагнуть» искомый признак.

Типовая сеть состоит из большого количества слоёв. После начального слоя (входного изображения) сигнал проходит серию свёрточных слоёв, в которых чередуется свёртка и субдискретизация (пулинг). Чередование слоёв позволяет составлять «карты признаков», на каждом следующем слое карта уменьшается в размере, но увеличивается количество каналов. На практике это означает способность распознавания сложных иерархий признаков. Обычно после прохождения нескольких слоёв карта признаков вырождается в вектор или даже скаляр, но таких карт признаков возникают сотни. На выходе свёрточных слоёв сети дополнительно устанавливают несколько слоёв полносвязной нейронной сети (перцептрон), на вход которых подаются окончательные карты признаков.

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.

Вывод: распознавание лиц является практическим применением теории распознавания образов; существует большой спектр алгоритмов для распознавания образов, которые можно систематизировать по основным методам, лежащим в основе; нейронные сети математическая интерпретация организации и функционирования сетей нервных клеток живого организма; основной составной частью нейронных сетей является персептрон, принимающий числовые значения умножение на веса каждого из входов, сумма результатов которых передаётся в передаточную функцию; свёрточные нейронные сети состоят из 2 чередующихся слоёв: свёртки и пулинга; свёрточные нейронные сети широко используются для идентификации различных объектов на изображениях.