

АЛГОРИТМЫ ОБРАБОТКИ ГРАФИЧЕСКОЙ ИНФОРМАЦИИ НА МОБИЛЬНЫХ УСТРОЙСТВАХ

В данной статье автором рассматриваются алгоритмы распознавания и обработки графической информации на мобильных устройствах. Тема является актуальной, так как сейчас наибольшее распространение получили мобильные приложения, и нужны инструменты, позволяющие работать в таких приложениях максимально комфортно.

Введение

Распознавание образов является одной из самых изученных задач в таких областях, как цифровая обработка изображений, компьютерное зрение, биометрия, создание интеллектуальных систем безопасности и контроля доступа. Тем не менее, в области распознавания образов продолжают представлять большой научный и практический интерес такие задачи как распознавание лиц, жестов, текстов, отпечатков пальцев, моделей помещений и другой графической информации. Распознавание изображения – очень сложная задача с теоретической и практической точек зрения [1].

I. Распознавание изображений

Распознавание изображений ежедневно используется тысячами компаний и миллионами потребителей. Распознавание изображений обеспечивается методами глубокого обучения, в частности Convolutional Neural Networks (CNN), архитектурой нейронных сетей, которая имитирует работу зрительной части мозга и анализ данных изображения. Распознавание образов CNN и нейронных сетей является ключевым компонентом глубокого обучения компьютерному зрению, которое имеет множество приложений, включая электронную коммерцию, игры, автомобилестроение, производство и образование.

Человеческий глаз видит изображение как набор сигналов, интерпретируемых зрительной корой головного мозга. Распознавание изображений имитирует этот процесс. Компьютеры «видят» изображение как набор векторов (многоугольники с цветовой аннотацией) или растр (холст пикселей с дискретными числовыми значениями для цветов).

В процессе распознавания нейронной сети векторное или растровое кодирование изображения превращается в конструкции, которые отображают физические объекты и особенности. Системы компьютерного зрения могут логически

анализировать эти конструкции, сначала упрощая изображения и извлекая наиболее важную информацию, а затем организуя данные посредством выделения и классификации признаков [2].

Одним из типов алгоритма распознавания изображений является классификатор изобра-

жений. Он принимает изображение (или часть изображения) в качестве входных данных и предсказывает, что изображение содержит. Входными данными является метка класса, такая как собака, кошка или таблица. Алгоритм должен быть обучен, чтобы учиться и различать классы.

В простом случае, чтобы создать алгоритм классификации, который может идентифицировать изображения с собаками, мы тренируем нейронную сеть с тысячами изображений собак и фонов без собак. Алгоритм научится извлекать признаки, которые идентифицируют объект «собака», и правильно классифицировать изображения, содержащие собак. Хотя большинство алгоритмов распознавания изображений являются классификаторами, другие алгоритмы могут использоваться для выполнения более сложных действий. Например, рекуррентная нейронная сеть может использоваться для автоматической записи подписей, описывающих содержание изображения.

После подготовки обучающих изображений нам понадобится система, которая сможет их обрабатывать и использовать для прогнозирования новых неизвестных изображений. Эта система представляет собой искусственную нейронную сеть. Нейронные сети представляют собой взаимосвязанную совокупность узлов, называемых нейронами или перцептронами. Каждый нейрон берет одну часть входных данных, обычно один пиксель изображения, и применяет простое вычисление, называемое функцией активации, для получения результата. Каждый нейрон имеет числовой вес, который влияет на его результат. Этот результат подается на дополнительные нейронные слои, пока в конце процесса нейронная сеть не сгенерирует прогноз для каждого входа или пикселя (см. рис. 1).

Этот процесс повторяется для большого количества изображений, и сеть изучает наиболее подходящие веса для каждого нейрона, которые обеспечивают точные прогнозы, в процессе, называемом обратным распространением. После того, как модель обучена, она применяется к новому набору изображений, которые не участвовали в обучении (набор для проверки или проверки), для проверки ее точности.

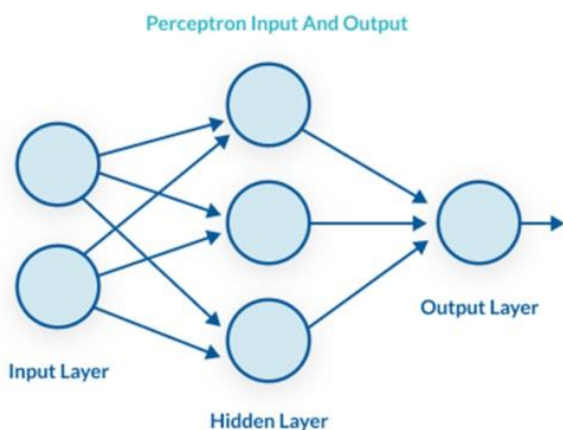


Рис. 1 – Слои нейронной сети

II. Контролируемая классификация

Контролируемая классификация входных данных в методе распознавания образов использует алгоритмы контролируемого обучения, которые создают классификаторы на основе данных обучения из различных классов объектов. Классификатор затем принимает входные данные и назначает соответствующий объект или метку класса. В компьютерном зрении используются методы распознавания образов под наблюдением для оптического распознавания символов (OCR), обнаружения лиц, распознавания лиц, обнаружения объектов и классификации объектов. Метод неконтролируемой классификации работает путем поиска скрытых структур в немаркированных данных с использованием методов сегментации или кластеризации. Обычные неконтролируемые методы классификации включают в себя: кластеризацию, гауссовские модели смеси, скрытые марковские модели.

В обработке изображений и компьютерном зрении для распознавания объектов и сегментации изображений используются неконтролируемые методы распознавания образов. В машинном обучении распознавание образов - это присвоение некоторого вида выходного значения (или метки) заданному входному значению (или экземпляру), согласно некоторому определенному алгоритму распознавания паттерна (PR) как области исследований, значительно развитой в 1960-х годах. Это была междисциплинарная тема, охватывающая, в частности, достижения в области статистики, инженерии, искусственного интеллекта, информатики, психологии и физиологии. Распознавание образов обычно классифицируется в соответствии с типом процедуры обучения, используемой для генерации выходного значения. Контролируемое обучение предполагает, что был предоставлен набор обучающих дан-

ных (обучающий набор), состоящий из набора экземпляров, которые были должным образом помечены вручную с правильным выводом [3]. Затем процедура обучения генерирует модель, которая пытается достичь двух иногда противоречащих друг другу целей: максимально эффективно работать с данными обучения и максимально обобщать новые данные (обычно это означает максимально простую задачу для некоторого технического определения).

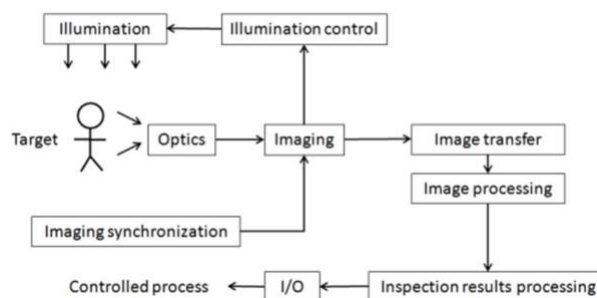


Рис. 2 – Базовая структура системы технического зрения

III. Заключение

Распознавание изображений, также известное как компьютерное зрение, позволяет приложениям, использующим определенные алгоритмы глубокого обучения, понимать изображения или видео. В этих сценариях изображения являются данными в том смысле, что они вводятся в алгоритм, алгоритм выполняет запрошенную задачу, и алгоритм выводит решение, представленное изображением. Одно из распространенных решений для приложений компьютерного зрения включает в себя распознавание лиц - будь то для маркировки друзей в Facebook или полиции, устанавливающей личность потенциального подозреваемого - исключительно на основе изображения. Другое использование для распознавания изображений в области медицины, где искусственный интеллект, используя распознавание изображений, может наблюдать рентгеновское излучение и расшифровывать диагнозы исключительно на основе изображения.

1. Абраменко А. Принципы распознавания / А. Абраменко – К.: Компьютер-пресс, 2017. – 123 с.
2. Авербух В.Л., Бахтерев М.О., Васёв П.А., Кузнецов Я.Д. Развитие программных средств научной визуализации // Параллельные вычислительные технологии (ПаВТ'2014): труды международной научной конференции (1-4 апреля 2014 г., г. Ростов-на-Дону). – Челябинск: Издательский центр ЮУрГУ, 2014. – С. 359.
3. Гаврилов Г.П. Логический подход к искусственному интеллекту / Г.П. Гаврилов – М.: Мир, 2018. – 256 с.

Посивец А. А., магистрант кафедры информационных технологий автоматизированных систем БГУИР