

УДК 004.8

**МЕТОДЫ И СИСТЕМЫ РАСПОЗНАВАНИЯ ОБРАЗОВ,
ИЗОБРАЖЕНИЙ И СИГНАЛОВ**



**METHODS AND SYSTEMS OF PATTERN RECOGNITION,
IMAGES AND SIGNALS**

Янаева Марина Викторовна

кандидат технических наук, доцент
кафедры информационные системы и программирование,
Кубанский государственный технологический институт
yanaevam@mail.ru

Осипова Анна Дмитриевна

студент
Кубанский государственный технологический институт
miss5x@yandex.ru

Богацкий Никита Валерьевич

студент
Кубанский государственный технологический институт
miss5x@yandex.ru

Аннотация. Распознавание образов актуальная тема, находящая применение в статистическом анализе данных, обработке сигналов, анализе изображений, поиске информации, биоинформатике, сжатию данных, компьютерной графике и машинном обучении. Одной из наиболее перспективных областей использования распознавания образов является распознавание лиц.

Ключевые слова: машинное обучение, распознавание, образы, паттерны, наборы данных, нейронные сети, лица.

Yanaeva Marina Viktorovna

Candidate of Technical Sciences,
Associate Professor
of the Department of Information
Systems and Programming,
Kuban State Technological Institute
yanaevam@mail.ru

Osipova Anna Dmitrievna

Student,
Kuban State Technological Institute
miss5x@yandex.ru

Bogatsky Nikita Valeryevich

Student,
Kuban State Technological Institute
miss5x@yandex.ru

Annotation. Pattern recognition is a topical topic that finds application in statistical data analysis, signal processing, image analysis, information retrieval, bioinformatics, data compression, computer graphics and machine learning. One of the most promising areas of using pattern recognition is facial recognition.

Keywords: machine learning, recognition, images, patterns, datasets, neural networks, faces.

Распознавание образов – это процесс, в ходе которого машина находит и распознает закономерности (паттерны) в данных, применяя алгоритмы машинного обучения. Его можно рассматривать как своего рода высоко-развитую классификацию данных [1].

Это сложный процесс анализа входных данных, извлечения шаблонов, сравнения их с определенными стандартами и использования результатов для руководства будущими действиями системы.

Следовательно, распознаватель образов должен обладать разнообразным набором функций. Автоматическое распознавание знакомых паттернов является одним из основных. Однако во многих случаях для правильного функционирования система должна уметь распознавать незнакомые объекты и классифицировать их, а также распознавать объекты даже тогда, когда данных о них может показаться недостаточно.

Существует три основных подхода, которые используют алгоритмы распознавания образов:

1. Статистический. Этот подход основан на статистической теории принятия решений. Распознаватель образов извлекает количественные признаки из данных вместе с несколькими выборками и сравнивает эти признаки. Однако в нем не затрагивается вопрос о том, как эти функции связаны друг с другом.

2. Структурный (он же синтаксический). Этот подход ближе к тому, как работает человеческое восприятие. Он извлекает морфологические признаки из одной выборки данных и проверяет, как они связаны и взаимосвязаны.

3. Нервный. В этом подходе используются искусственные нейронные сети. По сравнению с упомянутыми выше, он обеспечивает большую гибкость в обучении и наиболее близок к естественному интеллекту [2].

Каждый алгоритм распознавания образов, основанный на машинном обучении, включает в себя следующие шаги.

- ввод данных. Большие объемы данных поступают в систему через различные датчики;
- предварительная обработка или сегментация. На этом этапе система группирует входные данные, чтобы подготовить наборы для будущего анализа;
- выбор объекта (извлечение). Система выполняет поиск и определяет отличительные признаки подготовленных наборов данных;
- классификация. На основе признаков, обнаруженных на предыдущем шаге, данным присваивается класс (или кластер), или вычисляются прогнозируемые значения (в случае алгоритмов регрессии);
- постобработка. В соответствии с результатом распознавания система выполняет будущие действия.

Примеры распознавания образов можно легко найти в природе, например, люди распознают лица или домашние животные реагируют на свои имена. В технологиях алгоритмы распознавания образов, обученные с помощью машинного обучения, применяются в различных областях, начиная от повседневных задач и заканчивая узкоспециализированными областями [1].

Одной из наиболее перспективных областей использования распознавания образов является распознавание лиц. Google, Microsoft, Apple и многие другие компании встроили данную технологию в приложения для составления альбомов людей. Она проверяет, кто находится в аэропортах и других людных местах, и является новейшим биометрическим средством разблокировки вашего мобильного телефона, где приложений для распознавания лиц предостаточно. Необходимо подтвердить свою личность для банковского перевода? Просто посмотрите в камеру. Хочется узнать кто за дверью? Видеодомофон с функцией распознавания лиц сообщит вам об этом при условии, что вы загрузили фотографию лица этого человека. Другие системы используются для обнаружения пропавших людей или поимки сотрудников, которые обманывают о часах, проведенных в офисе.

Как же это работает? Прежде всего, компьютер должен узнать, что такое лицо. Это можно сделать, обучив алгоритм (обычно глубокую нейронную сеть) огромному количеству фотографий, на которых лица расположены в известных положениях. Каждый раз, когда алгоритму предоставляется изображение, он оценивает, где находится лицо. Сначала сеть будет мусором, но если повторять эти действия несколько раз, алгоритм совершенствуется и в конечном итоге овладевает искусством распознавания лиц. Это этап распознавания лиц [3].

Далее идет часть распознавания искомого лица. Это делается различными способами, но обычно используется вторая нейронная сеть. Он получает серию изображений лиц и учится – в течение многих раундов – как лучше отличить одно от другого. Некоторые алгоритмы явно отображают лицо, измеряя расстояния между глазами, носом и ртом и так далее. Другие отображают лицо, используя более абстрактные черты. В любом случае сеть выводит вектор для каждого лица – строку чисел, которая однозначно идентифицирует человека среди всех остальных в обучающем наборе.

При развертывании в реальном времени программное обеспечение работает с видеоматериалами в режиме реального времени. Компьютер сканирует кадры видео, обычно снятые в местах скопления людей, таких как входы на футбольные стадионы, вокзалы, аэропорты и т.д.

Сначала он распознает лица в кадре, а затем выдает векторы для каждого. Затем векторы лиц сверяются с векторами людей из списка наблюдения. Любые совпадения, превышающие заданный порог, затем ранжируются и отображаются. Типичный порог совпадения составлял 60 %, но планку можно установить и выше, чтобы уменьшить количество ложных срабатываний.

Независимые тесты, проведенные Национальным институтом стандартов и технологий США, показали, что в период с 2014 по 2018 год системы распознавания лиц стали в 20 раз лучше находить совпадения в базе данных из 12 миллионов портретных фотографий. За этот период частота сбоев снизилась с 4 % до 0,2 %, что является значительным увеличением точности, которое во многом обусловлено глубокими нейронными сетями.

Но такая высокая производительность зависит от идеальных условий: когда четкий снимок неизвестного человека сверяется с базой данных других высококачественных фотографий. В реальном мире изображения могут быть размытыми или плохо освещенными, люди могут отводить взгляд от камеры, или у них может быть шарф на лице, или они намного старше, чем на их эталонной фотографии. Все это склонно снижать точность. Также тесты Национальным институтом стандартов и технологий США показали, что технология испытывает ужасные трудности с близнецами, и даже лучшие алгоритмы путают их.

Предвзятость уже давно преследует алгоритмы распознавания лиц. Проблема возникает, когда нейронные сети обучаются на разном количестве лиц у разных групп людей. Итак, если система обучена на миллионе белых мужских лиц, но меньше женщин и цветных людей, она будет менее точной для последних групп. Меньшая точность означает больше ошибочных идентификаций и, возможно, больше людей будут ошибочно остановлены и допрошены.

Так, например, программное обеспечение Amazon для распознавания лиц ошибочно идентифицировало 28 человек как людей, которые ранее были арестованы. Оно непропорционально неправильно идентифицировало афроамериканцев и латиноамериканцев. Amazon заявила, что ACLU использовал неправильные настройки.

Литература

1. URL : <https://www.pravmir.ru/kamery-sledyat-za-vami-vsyudu-kak-rabotaet-sistema-raspoznaniya-licz-i-nuzhno-li-ee-boyatsya/> (дата обращения 26.07.22)
2. Аркадьев А.Г., Браверман Э.М. Обучение машины распознаванию образов. – М. : Книга по Требованию. – 2012. – 110 с. – ISBN 978-5-458-25154-9
3. Александров В.В., Горский Н.Д. ЭВМ видит мир. – Л. : Машиностроение : Ленингр. отд-ние, 1990. – 136 с. – ISBN 5-217-00611-0 : 80 к.

References

1. URL : <https://www.pravmir.ru/kamery-sledyat-za-vami-vsyudu-kak-rabotaet-sistema-raspoznaniya-licz-i-nuzhno-li-ee-boyatsya/> (date of the application: 26.07.22)
2. Arkadiev A.G., Braverman E.M. Teaching a machine to recognize patterns. – M. : Book on Demand. – 2012. – 110 p. – ISBN 978-5-458-25154-9
3. Alexandrov V.V., Gorsky N.D. The computer sees the world. – L. : Mechanical engineering: Leningrad. department, 1990. – 136 p. – ISBN 5-217-00611-0 : 80 k.