Научная статья
Original article
УДК 621

ИССЛЕДОВАНИЕ НЕЙРОННЫХ СИСТЕМ ДЛЯ РАСПОЗНАВАНИЯ ТЕКСТОВЫХ ИЗОБРАЖЕНИЙ

RESEARCH OF NEURAL SYSTEMS FOR RECOGNITION OF TEXT IMAGES

Морозов Артемий Александрович, Студент магистратуры 2 курс, факультет «Машиностроения», ФГАОУ ВО Московский политехнический университет, Россия, г. Москва

Morozov Artemiy Alexandrovich, 2nd year master's student, Faculty of Mechanical Engineering, Moscow Polytechnic University, Russia, Moscow

Аннотация: В статье проводится исследование нейронных систем для распознавания текстовых изображений. Рассматриваются понятие, структура и принцип работы нейронных сетей. Приводятся классификация нейронных сетей по способу обучения и количеству слоёв и этапы процесса их обучения. Выявляются характерные черты, отличия и специфика работы свёрточных нейронных сетей, чаще всего использующихся для распознавания текстовых изображений. Рассматриваются особенности обучения нейросетей и алгоритм работы систем распознавания текстов.

Annotation: The article deals with the study of neural systems for recognition of text images. The concept, structure and principle of operation of neural networks are considered. The classification of neural networks according to the method of training and the number of layers and the stages of the process of their training are

given. The characteristic features, differences and specifics of the work of convolutional neural networks, most often used for recognition of text images, are revealed. The features of learning neural networks and the algorithm of text recognition systems are considered.

Ключевые слова: нейронные сети, свёрточные нейронные сети, текстовые изображения, символы, распознавание текста.

Keywords: neural networks, convolutional neural networks, text images, symbols, text recognition.

В настоящее время распознавание образов, включая символы текстовых изображений, является одной из актуальных задач различных областей человеческой деятельности [1]. Масштабы практического применения алгоритмов распознавания образов увеличиваются по мере роста производительности компьютеров, делая востребованным поиск новых подходов и методов распознавания. При анализе больших объёмов данных активно применяются интеллектуальные методы, дающие возможность автоматически осуществлять обработку информации, извлекая необходимые сведения, что позволяет существенно упростить и ускорить весь процесс обработки [2]. Наиболее перспективным направлением решения задач распознавания символов и анализа данных являются глубокие нейронные сети, позволяющие добиваться высоких результатов в интеллектуальном анализе текстовых изображений. Это делает актуальным изучение специфики использования нейронных сетей для распознавания символов текстовых изображений.

Целью работы является исследование нейронных систем для распознавания текстовых изображений. Для её достижения были использованы методы анализа и синтеза научных публикаций и литературных источников по рассматриваемой теме.

Нейронная сеть представляет собой обучаемую математическую модель, тип и структура которой выбираются и синтезируются для решения задач

определённого типа [3]. Обучение нейронной сети осуществляется на основе данных о решении задачи экспертом либо информации о решении подобных задач в прошлом, поэтому для её обучения требуется подбор представительных данных и запуск алгоритма обучения, который автоматически воспринимает структуру введённых сведений. Нейронные сети способны распознавать примитивы достаточно сложной структуры, что обуславливает их широкое применение в задачах распознавания отдельных классов изображений, таких как текстовые изображения.

Нейронная сеть сформирована из нейронов, имеющих определённое ограниченное число входов, каждому из которых соответствует некоторый вес [4]. Задачей нейрона является определение взвешенной суммации своих входов. Для её решения, заключающегося в адекватном формировании выходного сигнала, используется целевая функция, значение которой является выходом нейрона. Нейронные сети могут быть классифицированы по следующим основным признакам:

1. По способу обучения:
   - с учителем, подающим на вход нейросети образы, производящим сравнение результирующих выходов с необходимыми и настраивающим веса сети так, чтобы сократить различия;
   - без учителя, когда после обработки входных образов нейросеть самоорганизуется, настраивая веса в соответствии с определённым алгоритмом.
2. По числу слоёв:
   - однослойная, состоящая из группы нейронов, формирующих слой, причём сигналы от входного слоя подаются непосредственно на выходной слой, преобразующий сигнал и выводящий результат;
   - многослойная, состоящая из входного и выходного слоёв и расположенных между ними скрытых слоёв, что придаёт сети большие вычислительные способности.

Процесс обучения нейросети новому классу задач включает несколько стадий [5]:

- постановка задачи и определение основных характеризующих предмет исследования параметров;
- выбор модели нейронной сети, в наибольшей степени подходящей для решения задач подобного типа;
- подготовка набора обучающих примеров, содержащего входные и соответствующие им выходные данные;
- проведение подготовленных данных через нейросеть, выдающую определённый результат, сравниваемый с известным эталонным;
- подстройка весовых коэффициентов межнейронных соединений с учётом процента ошибок;
- продолжение обучение до момента снижения процента ошибок до желаемого уровня либо до приведения нейронной сети в стационарное состояние.

Для решения задачи распознавания текстового изображения обычно используются свёрточные нейронные сети [6]. Как и обычные нейросети, они сформированы из нейронов с обучаемыми весами и смещениями, каждый из которых обладает несколькими входами, генерирует числовое значение и посылает его другим нейронам. Свёрточные сети представляют собой функции, которые принимают на вход вектор со значениями в неограниченном интервале и отдают на выход вектор со значениями, расположенными в ограниченном интервале. На их последнем слое располагается функция потерь, необходимая для обучения. Отличие свёрточной нейросети от обычной заключается в том, что первая напрямую учитывает подачу изображения на вход, что позволяет провести оптимизацию с целью облегчения процесса обучения и распознавания.

Обучение свёрточной нейросети распознаванию текстовых изображений осуществляется с учётом следующих ограничений [7]:

1. Извлечение признаков. Все нейроны получают входной сигнал от локального рецептивного поля предшествующего слоя, извлекая его локальные признаки. После извлечения признаков их местоположение не имеет значения, поскольку их расположение относительно других признаков приблизительно установлено.
2. Отображение признаков. Все вычислительные слои сети сформированы из множества карт признаков, каждая из которых имеет вид плоскости, на которой всем нейронам необходимо совместно использовать единое множество синаптических весов.
3. Инвариантность к смещению. Она реализуется посредством карт признаков с применением свёртки с ядром малого размера, выполняющей функцию «сплющивания».
4. Подвыборка. За каждым из слоёв свёртки располагается вычислительный слой, который осуществляет локальное усреднение и подвыборку, посредством которых достигается сокращение разрешения для карт признаков. Данная операция приводит к снижению чувствительности выходного сигнала оператора, отображающего признаки, к смещению и иным формам преобразований.

Обучение нейросети осуществляется на эталонных данных, соответствующих идеально написанным буквенным символам [8]. Последующее обучение производится на других данных, которые соответствуют различно написанным символам, при этом важно, чтобы символы не были схожи, поскольку в противном случае нейронная сеть настроится на распознавание символов только с похожим написанием, что будет приводить к ошибочным результатам. Заключительное обучение проводится вновь на эталонных образах, что обусловлено разучиванием сети определять хорошо написанные символы после обучения на неидеальных образцах.

Общий алгоритм работы систем распознавания текстов состоит из следующих этапов [9]:

1. Нахождение содержащей текст области и её локализация.
2. Предварительное улучшение качества и бинаризация локализованной области.
3. Выявление структуры обнаруженного блока текста и определение порядка чтения.
4. Сегментация текстового массива на слова, строки и символы.
5. Создание признакового описания всех символов.
6. Распознавание каждого отдельного символа.
7. Проведение словарной проверки.

Предварительная обработка изображений используется для устранения шумов, размытых границ символов и бликов, значительно снижающих эффективность распознавания текста [10]. Для предварительной обработки могут использоваться алгоритмы преобразования изображения в оттенки серого и ранговой обработки сигнала. Для рукописного текста дополнительно используется выпрямление символов, позволяющее устранить часто присутствующий наклон [11].

Важным этапом обработки является сегментация текстового изображения, представляющая собой выделение полезных сведений из изображения с их последующей обработкой. Сегментация в области текстового распознавания осуществляется последовательно: сперва на изображении выделяются фрагменты слов, затем фрагменты изображений, на которых присутствуют слова, а после распознанное изображение слова разделяется на символы, которые подвергаются распознаванию нейросетью.

Таким образом, одной из основных задач машинного обучения является задача распознавания текстовых изображений и выделения графических примитивов. Нейронные сети, несмотря на длительный период обучения, сложность реализации и высокие требования производительности, дают

возможность получать верный ответ с вероятностью, стремящейся к 99 %, обладают способностью к распознаванию нестандартных образов и обучению, позволяющему строить нечувствительные к воздействию искажений и шумов модели. Именно за счёт использования наборов зашумлённых векторов сеть может быть обучена работать с искажёнными шумами изображениями, которые чаще всего встречаются в практической деятельности.

Список литературы

1. Вальке А.А., Лобов Д.Г. Алгоритмы распознавания символов // Динамика систем, механизмов и машин. – 2018. – Т. 6, № 4. – С. 164-168.
2. Николаев П.Л. Классификация книг по жанрам на основе текстовых описаний посредством глубокого обучения // International Journal of Open Information Technologies. – 2022. – Т. 10, № 1. – С. 36-40.
3. Красновидов А.В., Алексеев А.С. Распознавание зашумлённых текстовых символов с помощью обучаемой нейронной сети // Интеллектуальные технологии на транспорте. – 2018. – № 2 (14). – С. 28-33.
4. Кулакович А.Ю. Программная реализация однослойной нейронной сети для распознавания цифровых символов // Инженерный вестник Дона. – 2018. – № 3 (50). – С. 58-72.
5. Денискин А.В. Применение нейронных сетей для разработки автоматизированной системы распознавания шрифтов // E-Scio. – 2018. – № 12 (27). – С. 150-155.
6. Саргин Д.В. Распознавание областей выделения текста на изображении методами машинного обучения: маг. дис.: 03.04.01 / Даниил Владимирович Саргин; Моск. физ.-техн. инст (гос. ун-т). – М., 2018. – 50 с.
7. Прокопеня А.С., Азаров И.С. Сверточные нейронные сети для распознавания изображений // Big Data and Advanced Analytics. – 2020. – № 6-1. – С. 271-280.
8. Пшеничкин Е.В., Цуканов М.В., Рыженков Д.В. Распознавание рукописных символов с помощью нейронных сетей методом с обратным

распространением ошибки // Инновационная наука. – 2018. – № 2. – С. 14-16.

9. Болотова Ю.А., Спицын В.Г., Осина П.М. Обзор алгоритмов детектирования текстовых областей на изображениях и видеозаписях // Компьютерная оптика. – 2017. – Т. 41, № 3. – С. 441-452.
10. Казнин А.А. Распознавание рукописного и печатного текста при разработке требований к программному обеспечению // Бюллетень науки и практики. – 2019. – Т. 5, № 12. – С. 246-256.
11. Бобров К.А., Шульман В.Д., Власов К.П. Анализ технологий распознавания текста из изображения // Международный журнал гуманитарных и естественных наук. – 2022. – № 3-2 (66). – С. 124-128.

Bibliography

1. Valke A.A., Lobov D.G. Algorithms for character recognition // Dynamics of systems, mechanisms and machines. - 2018. - V. 6, No. 4. - S. 164-168.
2. Nikolaev P.L. Genre classification of books based on text descriptions through deep learning // International Journal of Open Information Technologies. - 2022. - T. 10, No. 1. - S. 36-40.
3. Krasnovidov A.V., Alekseev A.S. Recognition of noisy text characters using a trained neural network // Intelligent Technologies in Transport. - 2018. - No. 2 (14). - S. 28-33.
4. Kulakovich A.Yu. Software implementation of a single-layer neural network for digital character recognition // Engineering Bulletin of the Don. - 2018. - No. 3 (50). - S. 58-72.
5. Deniskin A.V. Application of neural networks for the development of an automated font recognition system // E-Scio. - 2018. - No. 12 (27). – S. 150-155.
6. Sargin D.V. Recognition of text selection areas in an image using machine learning methods: Mag. dis.: 03.04.01 / Daniil Vladimirovich Sargin; Moscow physics and technology inst (state university). - M., 2018. - 50 p.

7. Prokopenya A.S., Azarov I.S. Convolutional Neural Networks for Image Recognition // Big Data and Advanced Analytics. - 2020. - No. 6-1. – S. 271-280.
8. Pshenichkin E.V., Tsukanov M.V., Ryzhenkov D.V. Recognition of handwritten characters using neural networks using the backpropagation method. Innovatsionnaya nauka. - 2018. - No. 2. - S. 14-16.
9. Bolotova Yu.A., Spitsyn V.G., Osina P.M. Overview of algorithms for detecting text areas in images and videos // Computer Optics. - 2017. - T. 41, No. 3. - S. 441-452.
10. Kaznin A.A. Recognition of handwritten and printed text in the development of software requirements // Bulletin of science and practice. - 2019. - V. 5, No. 12. - S. 246-256.
11. Bobrov K.A., Shulman V.D., Vlasov K.P. Analysis of text-from-image recognition technologies // International Journal of the Humanities and Natural Sciences. - 2022. - No. 3-2 (66). - S. 124-128.

© Морозов А.А., 2022 Научно-образовательный журнал для студентов и преподавателей «StudNet» №6/2022.

Для цитирования: Морозов А.А. ИССЛЕДОВАНИЕ НЕЙРОННЫХ СИСТЕМ ДЛЯ РАСПОЗНАВАНИЯ ТЕКСТОВЫХ ИЗОБРАЖЕНИЙ// Научно-образовательный журнал для студентов и преподавателей «StudNet» №6/2022.