

ОБНАРУЖЕНИЕ ОБЪЕКТОВ НА ИЗОБРАЖЕНИИ С ИСПОЛЬЗОВАНИЕМ СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ

Верхов К.А.

*Белорусский государственный университет информатики и радиоэлектроники,
г. Минск, Республика Беларусь*

Научный руководитель: Тонкович И.Н. – канд.хим.наук, доцент, доцент кафедры ПИКС

Аннотация. Рассмотрены строение и принципы работы сверточных нейронных сетей. Описаны слои, из которых состоит сверточная нейронная сеть, операции, выполняемые на каждом слое и назначение слоев. Выявлены категории, на которые делятся современные алгоритмы на основе сверточных нейронных сетей, указаны основные различия этих категорий и приведены примеры существующих алгоритмов.

Ключевые слова: компьютерное зрение, обнаружение объектов, машинное обучение, сверточные нейронные сети

Введение. Обнаружение объектов на изображении является актуальной задачей в компьютерном зрении. Учитывая современный тренд на автоматизацию, требуется наличие быстрых и эффективных систем для распознавания объектов [1]. А благодаря росту производительности аппаратного обеспечения и появлению наборов данных большого объема с различными изображениями, алгоритмы компьютерного зрения эффективно решают задачи классификации изображений, обнаружения объектов и сегментирования изображений.

Алгоритмы обнаружения объектов делятся на два вида: основанные на обработке изображения и на сверточных нейронных сетях. Алгоритмы, основанные на сверточных нейронных сетях, практически полностью заменили алгоритмы на основе обработки изображения, по причине того, что они намного точнее и быстрее.

Алгоритмы обнаружения объектов с использованием сверточных нейронных сетей не только являются объектами большого числа научных исследований, но и применяются в различных областях повседневной жизни: определение дыма на видео, системы беспилотного управления автомобилем, системы охранного мониторинга, анализ снимков с беспилотных летательных аппаратов.

В данной статье описаны строение и принцип работы сверточных нейронных сетей, а также рассмотрены алгоритмы, использующие эти сети.

Основная часть. Сверточные нейронные сети в основном состоят из следующих слоев: входного слоя, слоев свертки, активации, субдискретизации, полносвязной нейронной сети [2]. На практике слой свертки и слой активации не разделяют, подразумевая, что после каждой операции свертки результаты обязательно передаются в функцию активации. Слои свертки и полносвязной нейронной сети для своей работы используют не только функцию активации, но и такие параметры как вес и отклонение нейрона. В свою очередь, слой активации и слой субдискретизации используют конкретно заданные функции. Для того, чтобы сверточная нейронная сеть правильно обнаруживала и классифицировала объекты, она обучается на тренировочном наборе данных. Процесс обучения заключается в изменении параметров и весов сверточного слоя и слоя полносвязной нейронной сети.

Слой свертки является основой сверточной нейронной сети – в нем происходит большинство вычислений сети. Свертка значительно снижает сложность обучения модели и уменьшает количество связей и веса параметров, что упрощает ее обучение, по сравнению с обучением полносвязной нейронной сети того же размера. Виды сверток, которые обычно используют в данном слое, следующие: обычная свертка, транспонированная свертка и свертка с разделением по глубине.

В процессе обычной свертки на матрицу свертки поэлементно умножается каждый

пиксель изображения, суммируется и записывается в соответствующую позицию в выходном изображении [3]. В итоге создается изображение меньшего размера, на котором сильнее выделены признаки, присутствующие на исходном изображении. Транспонированная свертка, в отличие от обычной свертки, не уменьшает выходное изображение, а наоборот увеличивает его. Эта операция часто используется для семантического анализа и распознавания изображений. Свертка с дробным шагом создает карту признаков путем уменьшения шага транспонированной свертки, тем самым увеличивая разрешение выходного изображения. Свертка с разделением по глубине, вместо применения свертки для всего изображения, применяет ее для каждого цветового канала отдельно. Использование этой операции помогает снизить количество необходимых расчетов и уменьшить размер итоговой модели.

Слой активации подразумевает под собой передачу результатов свертки в функцию активации. Функция активации помогает нейронной сети узнавать и изучать шаблоны и повторяющиеся признаки в массивах данных, а также определяет, какая информация будет передана на следующий слой сети.

В качестве функций активации обычно используются следующие функции: линейный выпрямитель (*Rectified Linear Unit, ReLU*), линейный выпрямитель с «утечкой» (*Leaky rectified linear unit, Leaky ReLU*), рандомизированный линейный выпрямитель с «утечкой» (*Randomized LeakyReLU, RReLU*), экспоненциальная линейная функция (*Exponential Linear Units, ELU*). В подавляющем большинстве случаев используется функция линейный выпрямитель. Главная задача этой функции – отсечение отрицательной части скалярной величины. Данная функция представляется следующим образом (формула 1) [2]:

$$f(x) = \max(0, x). \quad (1)$$

Слой субдискретизации очень распространен в современных сверточных нейронных сетях. Он располагается между двумя последовательными сверточными слоями и используется для сжатия количества передаваемой на следующий слой информации и параметров. Это помогает избежать переобучения сети. Например, если на вход сверточной нейронной сети подается изображение, то задача слоя субдискретизации в сжатии этого изображения.

Слой субдискретизации значительно уменьшает размер матрицы, тем самым снижая чувствительность сверточного слоя к положению изображения, уменьшая количество параметров на слое полносвязной нейронной сети и ускоряя расчеты.

Алгоритмы обнаружения объектов, как и машинное обучение в целом, в последние годы очень сильно развивались. Популярные на текущий момент алгоритмы обнаружения объектов можно разделить на две категории [2]:

- двухэтапные;
- одноэтапные.

К двухэтапным относятся алгоритмы основанные на разбиении изображения на регионы. Например *R-CNN*, *Fast R-CNN*, *Faster R-CNN* и другие их разновидности [4]. Двухэтапность этих алгоритмов заключается в том, что перед этапом обнаружения и классификации объекта нейросеть должна провести этап анализа изображения и выделения на нем участков с возможным присутствием искомым объектов.

Одноэтапные алгоритмы используют сверточную нейронную сеть для того, чтобы сразу определить местоположение объекта и классифицировать его без дополнительной обработки изображения. К ним, например, относятся алгоритмы *YOLO (You Only Look Once)*, *SSD (Single Shot Multibox Detector)* и их разновидности.

Двухэтапные алгоритмы для выделения регионов изображения, на которых может находится объект, используют дополнительную сверточную нейронную сеть, которая создает карту признаков и на ее основе выделяет подходящие регионы. И только после этого, используя вторую сверточную нейронную сеть. Алгоритм классифицирует и определяет искомые объекты, проходя скользящим окном по выделенным ранее регионам изображения.

Одноэтапные алгоритмы обнаруживают и классифицируют объекты без предварительного выделения возможных регионов. Они сразу создают карту признаков, определяют искомые объекты и уточняют их границы. По сравнению с двухэтапными точность одноэтапных алгоритмов немного ниже, но при этом есть значительные улучшения в скорости работы.

Заключение. Благодаря наличию слоев свертки и субдискретизации, сверточные нейронные сети отлично подходят для обнаружения объектов на изображении. Недостатками являются длительность обучения таких сетей и неидеальная эффективность алгоритмов, использующих эти сети. Для решения этих проблем можно выделить несколько направлений развития алгоритмов с использованием сверточных нейронных сетей:

1 Автоматизированное обнаружение объектов. Для обучения нейросетям требуется тренировочный набор данных, в котором для каждого изображения указано, какой именно объект на нем изображен. Создание технологии, которая бы позволяла алгоритмам обнаружения объектов автоматически определять объект на изображении, значительно упростит создание обучающих наборов данных и уменьшит время, необходимое для обучения модели.

2 Создание алгоритма, который включал бы в себя преимущества как одноэтапных алгоритмов, так и двухэтапных.

3 Разработка эффективной нейронной сети для выделения признаков.

4 Применение созданных генеративно-состязательной сетью изображений для обучения нейронных сетей, используемых в алгоритмах обнаружения объектов. Использование в обучающей выборке естественных и синтезированных изображений поможет сделать алгоритм более надежным и точным.

Список литературы

1. Верхов, К.А. Обнаружение объектов на изображении с использованием машинного обучения / К.А. Верхов // Новые информационные технологии в научных исследованиях: материалы XXV Юбилейной Всероссийской научно-технической конференции студентов, молодых ученых и специалистов; Рязань: ИП Коняхин А.В. (Book Jet), 2020 – С. 226-227.

2. Overview of Object Detection Algorithms Using Convolutional Neural Networks / J.Ren, Y.Wang // Journal of Computer and Communications. – 2022. – Vol. 10, N 1. – Pp. 115–132.

3. Верхов, К.А. Методы обнаружения объектов на изображении / К.А. Верхов // Электронные системы и технологии: сборник материалов 57-й научной конференции аспирантов, магистрантов и студентов учреждения образования «Белорусский государственный университет информатики и радио-электроники», Минск, 19–23 апреля 2021 г. / редкол.: Д.В. Лихачевский [и др.]. – Минск: БГУИР, 2021. – С. 538-540.

4. Верхов, К.А. Обнаружение объектов на изображении с помощью метода Regional Proposal Network / К.А. Верхов // Новые информационные технологии в научных исследованиях: материалы XXVI Всероссийской научно-технической конференции студентов, молодых ученых и специалистов; Рязань: ИП Коняхин А.В. (Book Jet), 2021 – С. 213-214.

UDC 621.3.049.77–048.24:537.2

DETECTION OF OBJECTS IN THE IMAGE USING CONVOLUTIONAL NEURAL NETWORKS

Verkhov K.A.

Belarusian State University of Informatics and Radioelectronics, Minsk, Republic of Belarus

Tonkovich I.N. – PhD, assistant professor, associate professor of the department of ICSD

Annotation. The structure and principles of operation of convolutional neural networks are considered. The layers that make up the convolutional neural network, the operations performed on each layer and the purpose of the layers are described. The categories into which modern algorithms based on convolutional neural networks are divided are identified, the main differences between these categories are indicated, and examples of existing algorithms are given.

Keywords. computer vision, object detection, machine learning, convolutional neural networks