

1. Как решается проблема, которая заключается в том, что на изображении объект может быть в разном масштабе?

Проблема, которая заключается в том, что на изображении объект может быть в разном масштабе, обусловлена различными расстояниями от камеры до объекта в процессе съёмки.

Есть несколько способов решения данной проблемы, первое – предварительно обработать изображения, а если всё-таки данные поступают необработанные, то имеет смысл использовать свёрточную нейронную сеть.

Существует три основных категории решения данной проблемы с использованием CNN:

- Пирамида изображений – через сеть пропускаются несколько вариантов одного и того же изображения в разных масштабах и далее совмещаются выходы.
- Кодер-декодер архитектура – различные масштабы изображения преобразуются в кодере, а разрешение восстанавливается до исходного с помощью декодера.
- Пространственное объединение в пирамиду – входящий сигнал преобразуется фильтрами с разными степенями расширения и тем самым разными полями зрения.

2. Какой вид обучения используется для решения задачи кластеризации?

Кластеризация – задача разбиения заданной выборки объектов (ситуаций) на непересекающиеся подмножества, называемые кластерами, так чтобы каждый кластер состоял из схожих объектов, а объекты разных кластеров существенно отличались.

Задача кластеризации относится к широкому классу задач обучения без учителя. (Задачи, в которых известны только описания множеств объектов

(обучающей выборки), и требуется обнаружить внутренние взаимосвязи, зависимости, закономерности, существующие между объектами).

3. Для чего может быть использована свертка с ядром 1×1 ?

Свёртка 1×1 просто отображает входной пиксель со всеми его каналами на выходной пиксель, несмотря ни на что вокруг себя. Она часто используется для уменьшения количества каналов глубины, так как часто очень медленно умножать объемы с чрезвычайно большой глубиной. Фактически это скалярное произведение тензора на вектор, размерность которого равна глубине вектора.

Например, если на входе имеем $[32 \times 32 \times 6]$, то свёртка 1×1 будут эффективно выполнять 6-мерное скалярное произведение (поскольку входная глубина состоит из 6 каналов). В результате получим одну карту признаков $[32 \times 32 \times 1]$.

Если хотим уменьшить глубину до N каналов, то надо сделать N сверток с различными векторами (фильтрами).

4. Для чего нужен слой Flatten?

Flatten – слой выравнивания. Преобразует формат изображения из двумерного массива (28×28) в одномерный с размерностью 784 пикселя. Слой извлекает строки пикселей из входного черно-белого изображения и выстраивает их в один ряд. Этот слой не имеет параметров для обучения, он только переформатирует данные. Это необходимо нам потому что многослойный перцептрон работает с одномерными векторами.

5. Почему используется кросс-энтропия?

Потому что в задачах с вероятностной классификацией, в качестве функции потерь наиболее эффективно использовать кросс-энтропию. Для

определенного выходного вероятностного вектора \vec{y} , сравниваемого с фактическим вектором $\vec{\hat{y}}$, потеря (для k -го класса) будет определяться как

$$L(\vec{y}, \vec{\hat{y}}) = -\sum_{i=1}^k \vec{\hat{y}}_i \log y_i$$

Потери будут меньше для вероятностных задач, в основном из-за того, что данная функция предназначена для максимизации уверенности модели в правильном определении класса, и её не заботит распределение вероятностей попадания образца в другие классы.