**Тема 3.4 Задача обнаружение объектов. Основы подхода**

**План**

1. Задача обнаружение объектов
2. Архитектура Overfeat
3. Алгоритм Non-Maximum Suppression
4. Архитектура R-CNN

**Задача обнаружение объектов**

На предыдущих занятиях мы уже рассмотрели примеры некоторых задач компьютерного зрения. В некоторых задачах компьютерного зрения требуется обнаружить экземпляры объектов каждого класса и установить их позиции и геометрические размеры. Такая задача называется обнаружение объектов.

C задачей обнаружение объектов связаны сразу несколько типа задач компьютерного зрения. К таким задачам относятся как:

* Задача обнаружение объектов, в неизвестном числе
* Задача классификация + локализация, характерна для одного объектов в кадре. Задачи объектной сегментации – поиск объектов, и их сегментация.
* Задача обнаружения ключевых точек объекта,
* Задача паноптической сегментации
* и некоторые другие.

**Архитектура Overfeat**

Одной из первых успешных архитектур, реализующих обнаружение объектов при помощи только нейронных сетей была архитектура Overfeat. Принцип работы архитектуры Overfeat основан на продвижение некоторого окна по изображению. У окна меняется размер и положение. В каждом состоянии окне извлекается часть изображения. Для извлечённой части происходит классификация объекта и предсказание его размеров. Таким образом оказывается, обнаружение объектов — это одновременно и задача классификации, и задача регрессии. Отметим, что в сегодняшней терминологии Overfeat – решает задачу локализации в каждом окне.

Состояние окна изображения, для которого найден объект называется регион кандидат. Координаты центра и габариты такого окна называются **ограничивающей рамкой или просто рамкой (bounding box, bbox)**.

**Алгоритм** **Non-Maximum Suppression**

При решении задачи обнаружение объектов возникает следующая проблема: для одного объекта может быть предложено несколько регионов кандидатов. Разные регионы могут иметь разный уровень уверенности классификации – то есть разный выход вероятности нахождения объекта в регионе – так называемый Score. Для решения этой проблемы используется алгоритм **Non-Maximum Suppression** – алгоритма не максимального сжатия.

Алгоритм Non-Maximum Suppression состоит из следующих шагов.

* + Выбор из исходного списка регионов кандидатов со значениями score выше порога.
  + Вычисление для каждой пары оставшихся регионов значения пересечения площадей.
  + Из всех рамок с пересечением площадей выше порога остается одна, для которой score максимален
  + Процедура может повторяться итеративно пока не прекратится удаление рамок для всех найденных объектов на изображении.

**Архитектура R-CNN**

Архитектура overfeat имеет очень большое время работы из за необходимости перебора большого числа положений окна на изображении. Частично эта проблема решна в архитектуре R-CNN. Сеть R-CNN осуществляет предсказание области нахождения объекта и его класса в несколько этапов поэтому такой подход называется много-этапное обнаружение. Основное отличие R-CNN от overfeat это подготовка перед началом обучения. На этом этапе Производится выбор порядка 2000 регионов кандидатов, в которых предположительно есть объекты (**Regeons of Interest, ROI**). В оригинальной статье выбор регионов кандатов было предложено проводить при помощи метода селективный поиск (selective search) – при помощи сегментации так называемыми суперпикселями , так как это показано на слайде. Также отметим, что изначально все реигоны кандидаты могу иметь разные размеры – поэтому они все растягиваются до одного размера, который и необходим энкодеру сверточной сети.

**Резюме**

Архитектура R-CNN стала прообразом для всех последующих подходов к решению задачи обнаружения объектов. По крайней мере для так называемых **многоэтапных подходов.** Подробней о состоянии многоэтапных подходов мы поговорим на следующей лекции.