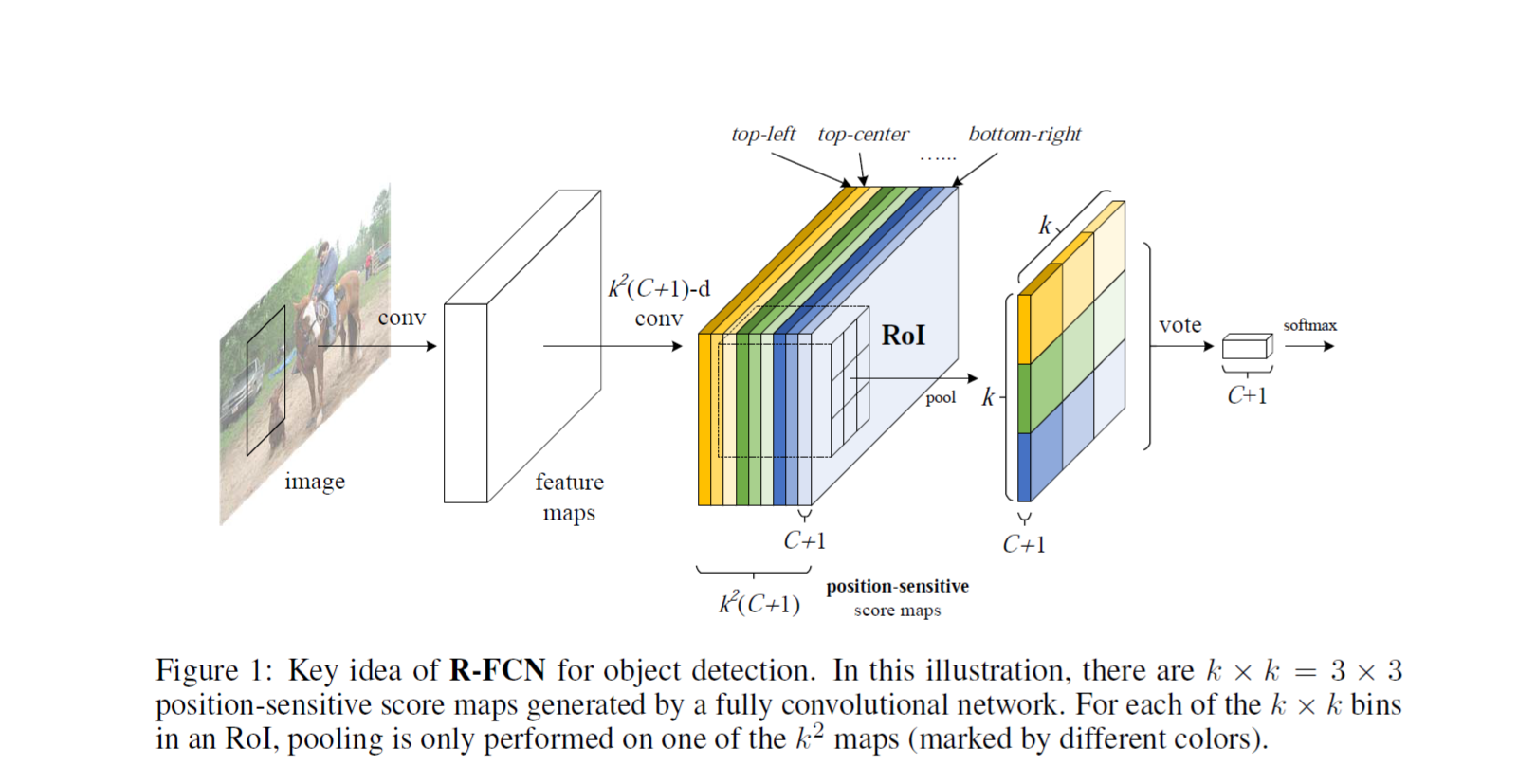
ЗАДАНИЕ 1.

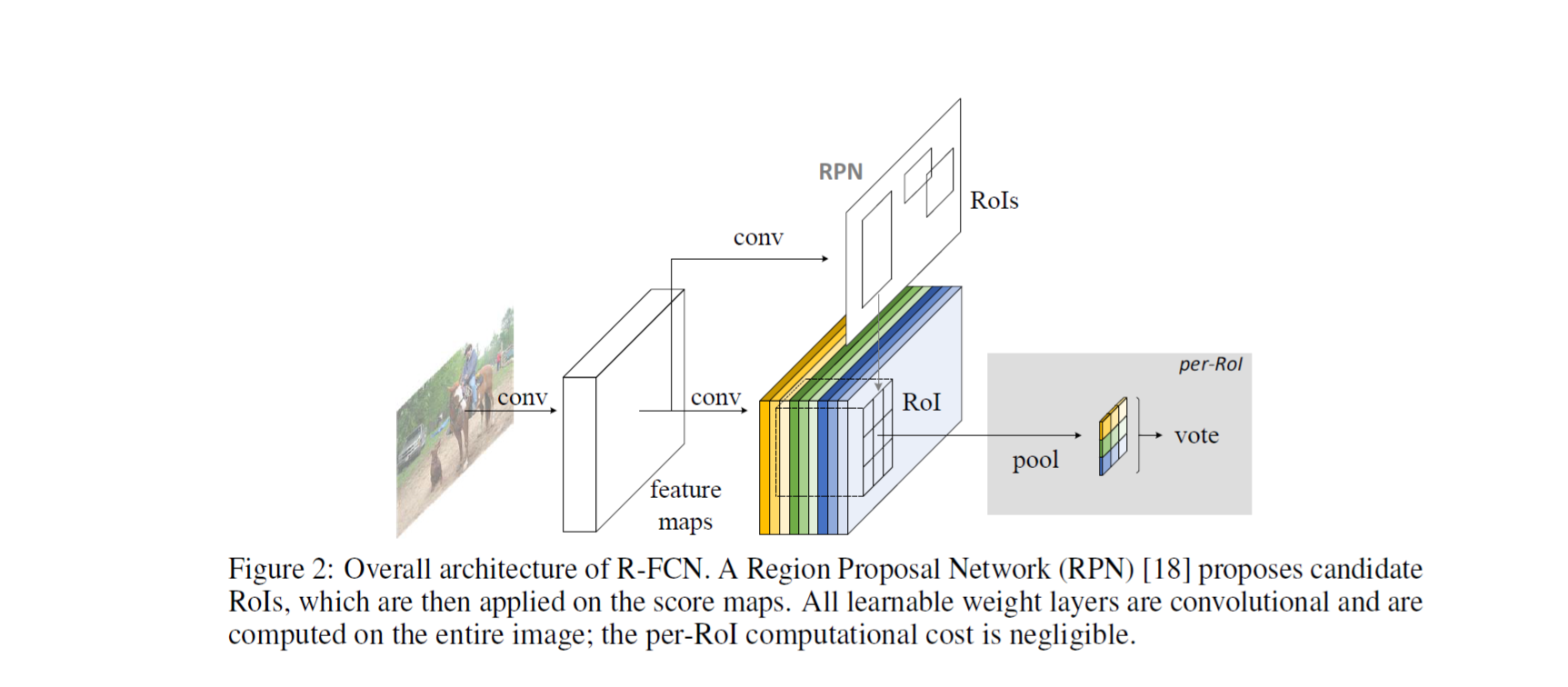
Сделайте краткий обзор какой-нибудь научной работы посвященной тому или иному алгоритму для object detection, который не рассматривался на уроке. Проведите анализ: Чем отличается выбранная вами на рассмотрение архитектура нейронной сети от других архитектур? В чем плюсы и минусы данной архитектуры? Какие могут возникнуть трудности при применении данной архитектуры на практике?

Ссылка на статью: <https://arxiv.org/pdf/1605.06409.pdf>

R-FCN: Object Detection via

Region-based Fully Convolutional Networks





Сеть R-FCN для детектирования объектов является конкурентом сети Fast/Faster

R-CNN и устраняет главный недостаток последней - применение дорогостоящей подсети для каждого региона сотни раз. Для достижения этой цели используется карты оценок с учетом позиции (position-sensitive score maps) для решения противоречия между трансляционной инвариантностью в классификации изображений и трансляционной зависимостью от локализации – региона при обнаружении объекта.

**Описание архитектуры.**

Учитывая регионы предложения (RoI), архитектура R-FCN предназначена для классификации регионов RoI в объекты определенных категорий и фона. Вначале слева направо идет ResNet-101, которая имеет только 100 сверточных слоев и порождает карты характеристик (feature maps). Она предварительно обучена на ImageNet. Далее идут сверточные слои, собственно, R-FCN — это обучаемые слои с весами и вычисляются веса на всем изображении (по регионам стоимость вычислений незначительна). Последний сверточный слой создает набор масштаба чувствительных к положению карт оценок (position-sensitive score maps) для каждой категории и, таким образом, имеет выходной слой с \* (C + 1) -каналами с категориями объектов C и 1 категория для фона. Содержимое масштаба карты оценок (position-sensitive score maps) соответствует пространственной сетке k Х k, описывающей относительные позиции. Например, при k Х k = 3 Х 3= 9 карт оценок кодируют случаи положения {вверху слева, вверху в центре, ..., внизу справа} объекта определенной категории. R-FCN заканчивается чувствительным к положению слоем pooling RoI. Этот слой объединяет выходные данные последнего сверточного слоя и генерирует оценки для каждого RoI. Сеть региональных предложений (RPN) предлагает кандидатов RoIs, которые затем накладываются на карты оценок. Позиционно-чувствительный RoI слой проводит выборочное объединение (pooling), и каждый из k Х k бинов агрегирует ответы только из одной карты оценки (score map) из набора k Х k карт оценок. Благодаря сквозному обучению этот слой RoI обрабатывает последние сверточный слой для обучения специализированных чувствительных к положению карт оценок.

**Плюсы и минусы данной архитектуры**

Плюсы данной архитектуры – простота архитектуры, быстрота обучения за счет не применения дорогостоящей подсети для расчета для каждого региона, а расчета по всему изображению, возможность подключить в качестве генератора карты признаков стандартную сверточную сеть, использование только сверточных слоев – полностью сверточная сеть.

Минусы – в некоторых случаях результаты mAP меньше чем для Faster R-CNN.

**Трудности при применении данной архитектуры на практике**

Настройкачувствительного к положению слоя - pooling RoI – какой взять масштаб. В статье использовался 3 Х 3.