

РАСПОЗНАВАНИЕ И СЕГМЕНТАЦИЯ ИЗОБРАЖЕНИЙ ИГРЫ НА МУЗЫКАЛЬНЫХ ИНСТРУМЕНТАХ

Еспаев Галымжан Берикулы

магистрант, Казахский национальный исследовательский технический университет имени К. И. Сатпаева, РК, г. Алматы

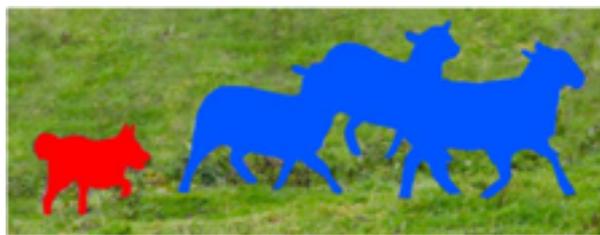
Кеншимов Чингиз Амирханович

научный руководитель, доктор PhD, Казахский национальный исследовательский технический университет имени К. И. Сатпаева, РК, г. Алматы

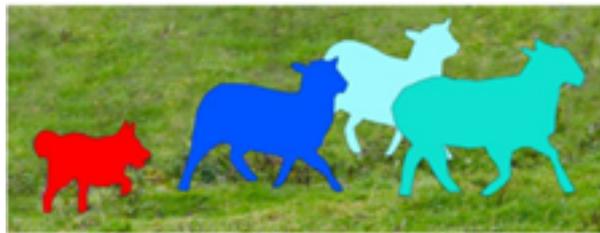
Сегментация в компьютерном зрении (компьютерном зрении) - это процесс разделения изображений на несколько сегментов (также называемый набором супер пикселей). Цель сегментации состоит в том, чтобы упростить и изменить внешний вид изображения, чтобы его можно было легко анализировать. Сегментация изображений обычно используется для выделения объектов и границ (линий, кривых и т. д.) На изображениях. Точнее, сегментация изображений-это процесс присвоения таких символов каждому пикселю изображения с общими визуальными характеристиками пикселей с одинаковыми символами. Результатом сегментации изображений является множество сегментов. Вместе они покрывают все изображения или множество контуров, отделенных от изображений. Все пиксели в сегменте схожи по некоторым характеристикам или вычисленным свойствам, таким как цвет, яркость или текстура. Соседние сегменты существенно различаются в этой характеристике [1].

Существует два основных типа сегментации изображений-семантическая сегментация (Semantic Segmentation) и экземплярная сегментация (Instance Segmentation). При семантической сегментации все объекты одного типа обозначаются знаком одного класса, а при экземплярной сегментации объекты-аналоги приобретают свои индивидуальные признаки.

Обнаружение или локализация объектов-это серьезный шаг от грубого до точного воспроизведения цифровых изображений. Он не только дает классы объектов изображения, но и обеспечивает расположение классифицированных объектов изображения. Местоположение указывается в виде центроидов или ограничительных рамок. Семантическая сегментация предсказывает признаки для каждого пикселя входного изображения и дает точный результат. Каждый пиксель обозначается в соответствии с классом объекта, на котором он построен. Продолжая эту эволюцию, сегментация экземпляра дает разные признаки отдельным экземплярам объектов, относящихся к одному классу. Поэтому сегментацию экземпляра можно рассматривать как способ одновременного решения задачи определения объектов, а также задачи семантической сегментации (показано на рис.1).



Semantic Segmentation



Instance Segmentation

Рисунок 1. Виды сегментации

Detectron2 - это программная система нового поколения Facebook AI Research, реализующая самые современные алгоритмы обнаружения объектов. Это переработка с нуля предыдущей версии detectron, созданной на базе Maskrcnn-benchmark [2].

Благодаря модульной реализации Detectron2 появилась возможность подключения пользовательских реализаций модулей к любой части системы обнаружения объектов. Detectron2 имеет другие модули, такие как Cascade R-CNN, включая модули Faster R-CNN, Mask R-CNN, RetinaNet и densepose, которые доступны в первом Detectron. Благодаря добавлению аппаратного ускорения к GPU можно значительно увеличить скорость работы и обучение моделей. Кроме того, теперь легко распространять обучение между несколькими серверами GPU, что облегчает масштабирование и чтение для очень больших наборов данных [3].

Поскольку в сегментации изображений людей, играющих на музыкальных инструментах, используется экземплярная сегментация, в Detectron2 используется модуль Mask R-CNN.

В этой архитектуре объекты классифицируются и локализуются с использованием рамок ограничений и семантической сегментации, которые классифицируют каждый пиксель по набору категорий. Каждая зона интереса получает маску сегментации. Знак класса и ограничительная рамка выводятся в качестве конечного результата. Его архитектура-это расширение Faster R-CNN. Faster-CNN состоит из глубокой конвульсивной сети, которая представляет области и детектор, использующие области.

Сегментация изображений. Для сегментации картинок необходимо собрать картинки (картинки людей, играющих на музыкальных инструментах). Создается аннотация к собранному изображению, и этот формат аннотации сохраняется как файл COCO json. Аннотация составлена для 4 класса (cellist, clarinetist, flutist, guitarist). Код можно написать в Google Colab. Перед обучением (обучением) модели необходимо проверить правильность или неправильность аннотирования изображений. Если картинки будут соответствовать аннотации, то можно приступать к обучению модели. По окончании обучения можно увидеть результат (показан на рис.2).



Рисунок 2. Результат сегментации изображения

Список литературы:

1. Сегментация (обработка изображений) [Электронный ресурс]. – Режим доступа: [https://ru.wikipedia.org/wiki/Сегментация_\(обработка_изображений\)](https://ru.wikipedia.org/wiki/Сегментация_(обработка_изображений)) (дата обращения: 27.04.21)
2. Detectron2 [Электронный ресурс]. – Режим доступа: <https://github.com/facebookresearch/detectron2>. (дата обращения: 27.04.21)
3. Распознавание объектов в Detectron2 [Электронный ресурс]. – Режим доступа: <https://deeplearningrussia.wordpress.com/2019/10/30/распознавание-объектов-в-detectron2/>. (дата обращения: 27.04.21)