

РАЗРАБОТКА АЛГОРИТМОВ СЕМАНТИЧЕСКОЙ СЕГМЕНТАЦИИ ИЗОБРАЖЕНИЙ

Ильин В.И.

НИ ТПУ, ИШИиЭС, группа 6В41ПИШ, email: wasilij1711@tpu.ru

Введение

Развитие систем машинного зрения является очень перспективным направлением. Одной из важных частей программного обеспечения машинного зрения является аналитика видеоданных и обнаружение на них объектов окружения. Для качественного решения этой задачи необходимо реализовать алгоритм семантической сегментации изображения, применяемый при обработке полученных материалов.

Целью работы стало создание алгоритма семантической сегментации изображения, который можно было бы применить при решении практических задач.

Задачи, поставленные для достижения цели выглядят следующим образом.

1. Выбрать готовый вариант и доработать, либо же разработать архитектуру нейросети.
2. Найти и подготовить базу данных для обучения и тестов нейросети.
3. Провести эксперимент по обучению и тестам нейросети.

Семантическая сегментация

Сначала необходимо разобраться в двух вопросах: «Что такое семантическая сегментация?» и «Как можно применить семантическую сегментацию?». Семантическая сегментация – это процесс разбиения получаемого изображения на отдельные элементы, которые присваиваются какому-либо классу или категории. Пример сегментации представлен на рис. 1.

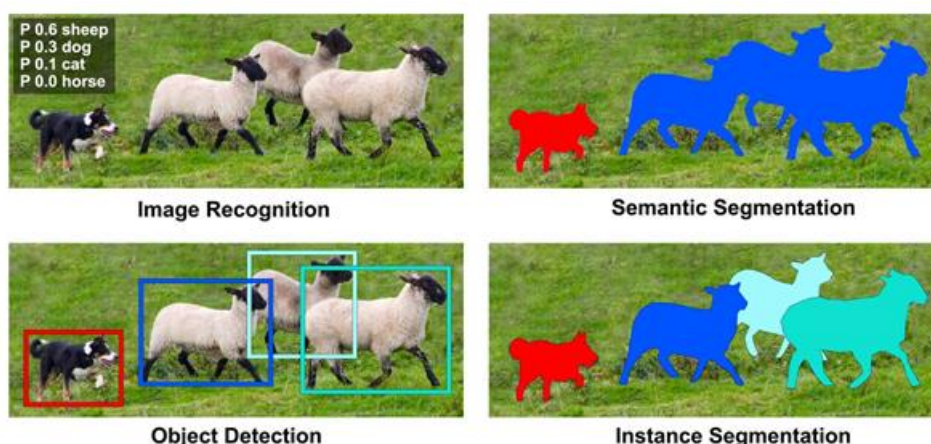


Рис. 1. Семантической сегментации изображения [1]:
Осуществлялось разбиение по 4 классам: собаки, кошки, овцы и лошади

Применяются подходы семантической сегментации в областях, где необходимо четкое деление на объекты при анализе изображений. Примерами таких областей служат машинное зрение, медицинское изображение и автономная навигация.

Модель нейросети

После анализа возможных моделей нейросети была выбрана модель сверточной нейросети. Выбрана была именно эта модель из-за ряда преимуществ, присущих ей. Во-первых, сверточные нейросети позволяют использовать удобные параллельные вычисления, что позволяет использовать и обучать нейросеть на графическом процессоре, что и будет использоваться в дальнейшем. Во-вторых, сверточные нейросети имеют меньшее количество настраиваемых весов, в отличие от полноразмерных, что позволяет снизить сложность выполнения, а при обучении прийти к обобщению демонстрируемой информации, а не попиксельному запоминанию [2].

Архитектуру было решено взять у сверточной нейросети U-Net. U-Net – это сверточная нейросеть, архитектура которой представляет полносвязную сеть, модифицированную так, чтобы она могла работать с меньшим количеством примеров [3]. Названа нейросеть так из-за того, что её архитектура представляет из себя конвертацию слоёв в форме буквы U, что представлено на рис. 2.

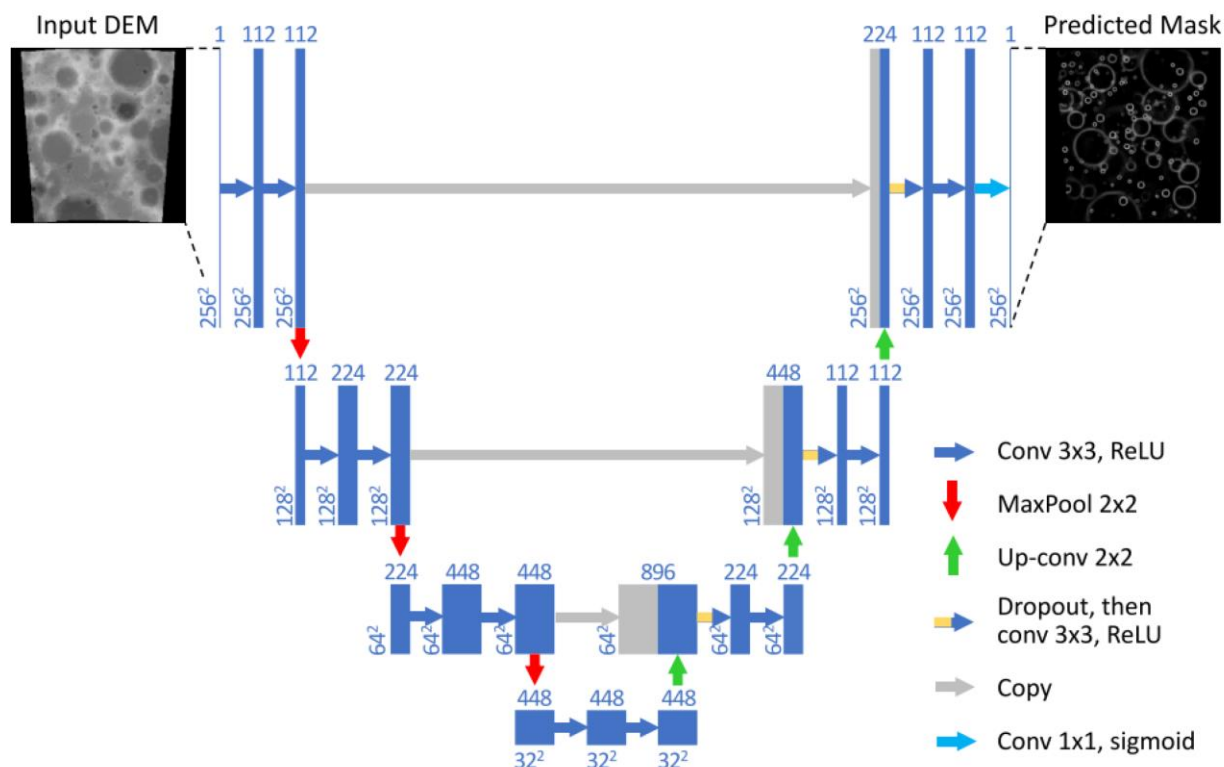


Рис. 2. Архитектура нейросети U-Net

Реализация

Реализация нейронной сети осуществляется на языке программирования Python. Связан такой выбор с тем, что язык программирования Python хорошо подходит для работы с данными в той или иной форме, что необходимо при обработке информации с фотографий. Также Python поддерживает большое количество пользовательских библиотек и расширений, некоторые из которых активно используются при создании нейросети, например, библиотеки NumPy, Keras, Matplotlib и другие. Сами тесты и обучения осуществляются на платформе Google Colaboratory, так как эта платформа позволяет подключать графическую память с серверов при запуске кода, что необходимо для корректной работы сверточной нейросети.

Заключение

Нейросети, позволяющие осуществлять семантическую сегментацию изображений очень сильно помогают при решении некоторых задач современности, поэтому стоит понимать их логику и принципы работы, чтобы создавать не только уже существующие модели, но и перерабатывать их, создавая новые поколения моделей, работающие в разы быстрее и корректнее.

Список литературы

1. Сегментация экземпляров с помощью Mask R-CNN // Habr.ru: Машинное обучение : сайт. – 2022. – URL: <https://habr.com/ru/articles/665300/> (дата обращения: 19.09.2024)
2. Бородин Г.Д. Краткий обзор и классификация искусственных нейронных сетей // Известия Тульского государственного университета. Технические науки. – 2021. – URL: <https://elibrary.ru/item.asp?edn=blhmxx&ysclid=lkwnnvbluk797943160> (дата обращения: 17.09.2024)
3. Olaf Ronneberger Convolutional Networks for Biomedical Image Segmentation / Olaf Ronneberger, Philipp Fischer, Thomas Brox. – Германия : Computer Science Departmen, 2015. – 8 с.