

## СИСТЕМА РАСПОЗНАВАНИЯ ОБЪЕКТОВ НА СНИМКАХ ЗЕМНОЙ ПОВЕРХНОСТИ НА ОСНОВЕ СВЕРТОЧНОЙ НЕЙРОННОЙ СЕТИ U-NET

Куприянова Д.В.

Некревич Ю.И.

Белорусский государственный университет информатики и радиоэлектроники  
г. Минск, Республика Беларусь

Никульшин Б.В. – канд. техн. наук, доцент

Представлены результаты эксперимента с применимостью сверточной нейронной сети U-Net для сегментации множества объектов на снимках земной поверхности.

Развитие авиакосмической промышленности привело не просто к активному освоению космического пространства, но и появлению множества потоков информации, получаемых от спутников, находящихся на орбите Земли, и, как результат, постоянно возрастающим требованиям к качеству и скорости их автоматизированной обработки. Получаемые снимки имеют большую ценность во многих отраслях: сельское хозяйство, метеорология, георазведка и т.д. Однако существенной проблемой является качественный анализ снимков земной поверхности, поскольку они содержат большой объем информации при относительно невысокой детализации самих объектов.

Целью работы является разработка системы сегментации и классификации объектов на аэрофотоснимках различного пространственного разрешения. В качестве поддерживаемых классов объектов определены строения, поле, вода, дорога, зеленые насаждения (например, лес).

Схематично разработанная система представлена на рисунке 1.

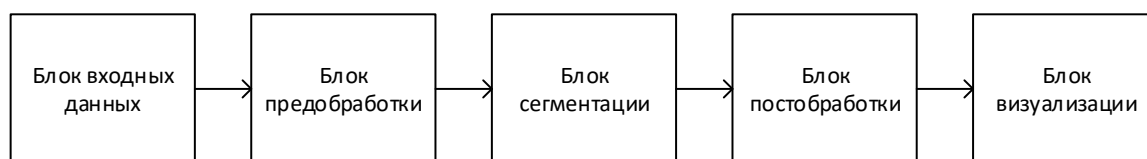


Рисунок 2 – Структурная схема предлагаемой системы

Блок входных данных загружает изображение в одном из поддерживаемых форматов (tiff, jpg, ppm), после чего изображение подвергается предварительной обработке, включающей конвертацию в изображение в градациях серого и масштабирование к заданному размеру.

Затем полученное изображение передается на блок сегментации. В данном случае используется сверточная нейронная сеть U-Net, созданная в 2015 году для сегментации биомедицинских изображений в отделении Computer Science Фрайбургского университета [1]. Архитектура сети представляет собой полносвязную сверточную сеть, модифицированную так, чтобы она могла работать с меньшим количеством обучающих образов при увеличении точности сегментации.

Архитектура сети показана на рисунке 2, состоит из двух частей (сужающейся и расширяющейся) и соответствует схеме кодер-декодер:

- кодировщик уменьшает пространственное измерение с помощью объединения слоев;
- декодер восстанавливает детали объекта и пространственное измерение.

Также существуют быстрые соединения от кодера к декодеру, чтобы «помочь» декодеру лучше восстановить детали объекта.

Блок постобработки в перспективе будет выполнять анализ полученных результатов, выявление явно ошибочных участков (например, разрывы в близко расположенных отрезках дорожного полотна).

Блок визуализации сохраняет полученный результат и выводит его на экран.

Тестовый набор включает 25 изображений (пример оригинального снимка земной поверхности показан на рисунке 3.а, результат распознавания – на рисунке 3.б), заранее размеченных вручную. Обучение для поддерживаемых классов выполнялось отдельно. Для оценки качества распознавания всех поддерживаемых классов объектов использовался коэффициент Жаккара:

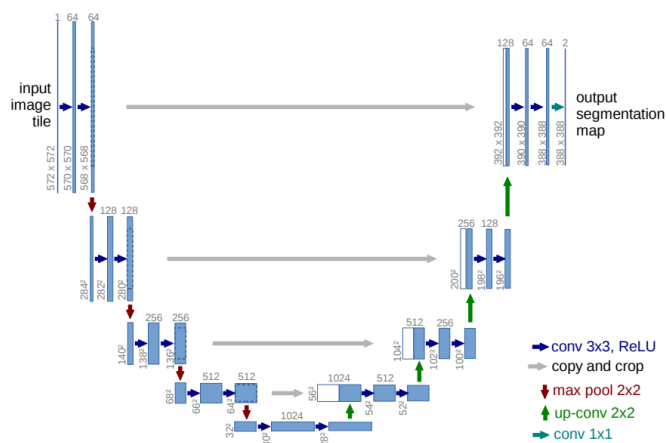


Рисунок 2 – Базовая архитектура U-Net

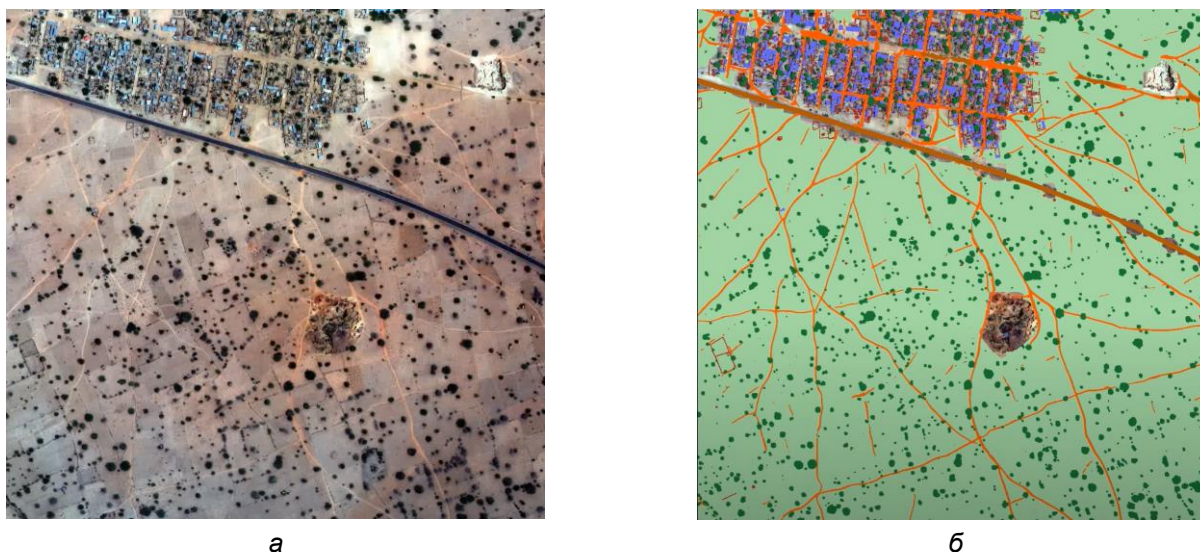


Рисунок 3 – Пример фотографии земной поверхности: а – оригинальный снимок, б – после сегментации

$$J = \frac{1}{n} \sum_{i=1}^n \frac{A_i \cap B_i}{A_i \cup B_i}, \quad (1)$$

где  $n$  – количество классов, которые поддерживаются разработанным алгоритмом;

$A_i$  и  $B_i$  – координаты областей исходной и предсказанной масок соответственно.

Тестирование проводилось с применением центрального процессора Intel Core i7-9700 и видеокарты NVIDIA RTX 2060 RTX с 8 Гб видеопамати.

Представлен алгоритм сегментации объектов на снимках земной поверхности с помощью 5 заранее построенных моделей на основе сверточной нейронной сети U-Net (соответственно для строений, полей, воды, дороги, зеленых насаждений).

Полученный усредненный коэффициент Жаккара для всех поддерживаемых классов объектов составляет 0,64, что является довольно низким коэффициентом, что связано с низким качеством обучающей выборки и высокими требованиями к необходимому аппаратному обеспечению, проявляющимися в длительном процессе обучения.

В дальнейшем планируется усовершенствовать тестовый набор данных для обучения, исследовать алгоритмы постобработки для выявления явных ошибок (например, разрывы в близко расположенных отрезках дорожного полотна), а также исследовать альтернативные архитектуры нейронных сетей.

**Список использованных источников:**

1. U-Net: Convolutional Networks for Biomedical Image Segmentation [Electronic Resource] / ArXiv. – Mode of access: <https://arxiv.org/abs/1505.04597>. – Date of access: 06.04.2022.