

СЕГМЕНТАЦИЯ СПУТНИКОВЫХ ИЗОБРАЖЕНИЙ НА ОСНОВЕ СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ

Ерофеев В.С., студент гр.250502

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

Перцев Д.Ю. – канд. техн. наук, доцент

Аннотация. В работе проводится сравнительный анализ архитектур сверточных нейронных сетей U-Net и Feature Pyramid Network применительно к задаче сегментации спутниковых изображений. Полученные результаты показывают, что обе архитектуры демонстрируют высокую точность сегментации, при этом U-Net обеспечивает лучшую детализацию мелких объектов, а Feature Pyramid Network – более устойчивую сегментацию объектов различного масштаба.

Ключевые слова. Спутниковые изображения, U-Net, Feature Pyramid Network, сверточные нейронные сети, обнаружение объектов.

Применение традиционных методов обработки изображений при сегментации объектов сложной формы и различного масштаба оказываются малоэффективными. В то же время современные подходы, основанные на сверточных нейронных сетях (CNN), демонстрируют высокую эффективность в решении задач подобного рода. Среди них следует выделить такие модели, как U-Net и Feature Pyramid Network (FPN), которые изначально успешно применялись в медицинской диагностике и затем были адаптированы к обработке спутниковых данных.

Модель U-Net изначально была предложена для сегментации биомедицинских изображений [1]. Она состоит из симметричных энкодера и декодера с пропусками между соответствующими уровнями, что позволяет эффективно восстанавливать пространственную информацию и обеспечивать точную сегментацию объектов. Данная сеть показала высокую эффективность при ограниченном объеме обучающих данных и широко используется в задачах сегментации различных типов изображений.

Модель FPN была разработана для улучшения обнаружения объектов различного масштаба [2]. Она использует иерархическую структуру признаков, в которых информация с разных уровней объединяется с помощью нисходящего пути и боковых соединений. Это позволяет эффективно сегментировать объекты различного размера и формы. FPN применяется в задачах обнаружения и сегментации объектов на изображениях высокой сложности.

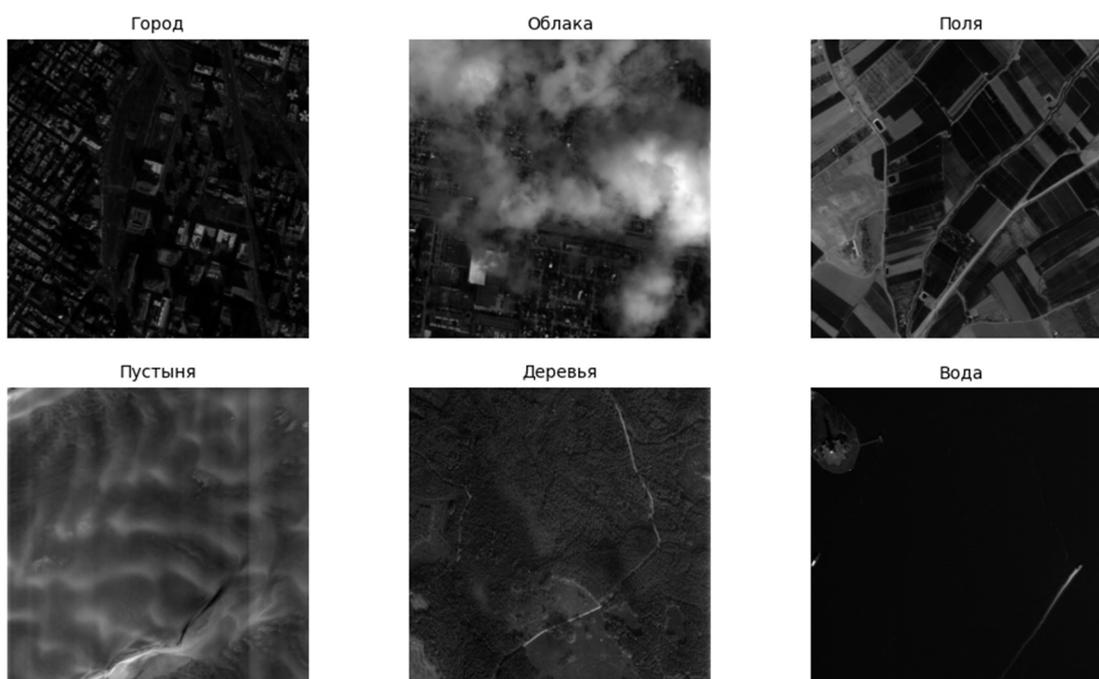


Рисунок 1 – Примеры классов

Для проведения эксперимента использовался набор спутниковых снимки, охватывающий различные географические регионы. Изображения были размечены по шести классам: город, облака, деревья, поля, пустыни, деревья и вода. Примеры классов представлены на рисунке 1. Соотношение классов одинаковое. Все изображения были разделены на три группы: для обучения (80% от общей выборки), для проверки (10%) и тестирования (10%). При обучении к каждому изображению применялась аугментация данных (повороты, размытие, повышение четкости) и нормализация. Разметка выполнялась вручную с использованием специализированного программного обеспечения CVAT, обеспечивая высокое качество обучающей выборки. Всего было обработано 250 изображений. Для извлечения признаков в моделях U-Net и FPN выбран энкодер ResNet-18 [3]. Данный выбор обусловлен балансом между вычислительной эффективностью и способностью модели извлекать детальные признаки, что особенно важно для сегментации сложных географических объектов. Также использовались предобученные на ImageNet веса, чтобы модель начинала обучение с уже «наученного» представления визуальных признаков. Для многоклассовой семантической сегментации используется функция активации SoftMax2D, которая нормализует выходные значения по каждому пикселю, распределяя вероятность принадлежности к одному из шести классов. Модель обучалась на протяжении 100 эпох. Все это позволило достичь стабилизации метрик обучения и достаточной сходимости для точного выделения объектов.

Для оценки эффективности обучения моделей U-Net и FPN были построены графики изменения метрик Dice Loss и IoU в зависимости от количества эпох (см. рисунки 2 и 3 для моделей U-Net и FPN соответственно). График метрики IoU отражает, как изменяется точность сегментации по мере обучения модели. При этом рост значения свидетельствует о том, что модель лучше справляется с задачей сегментации по мере увеличения эпох, правильно определяя границы объектов на изображениях. График функции Dice Loss показывает, как уменьшается ошибка модели при обучении. Снижение данного значения указывает на то, что модель с каждой эпохой точнее предсказывает маски объектов, уменьшая расхождение между предсказанными и истинными сегментами. В таблице 1 представлены итоговые значения по выбранным метрикам.

Как видно из полученных результатов, модель U-Net продемонстрировала более качественные результаты. На рисунках 4 и 5 представлены примеры масок для спутниковых снимков, полученных при съемке города и леса, с использованием указанных моделей.

Визуальный анализ масок, полученный в результате сегментации, показал, что U-Net точнее определяет границы объектов, особенно в случаях с мелкими деталями, такими как узкие полосы воды или небольшие участки деревьев.

Таблица 1 – Значения метрик Dice Loss и IoU, которые были получены при обучении в течении 100 эпох.

| Метрика | U-Net | FPN |
|-----------|--------|--------|
| IoU | 0.9187 | 0.9054 |
| Dice Loss | 0.0525 | 0.0583 |

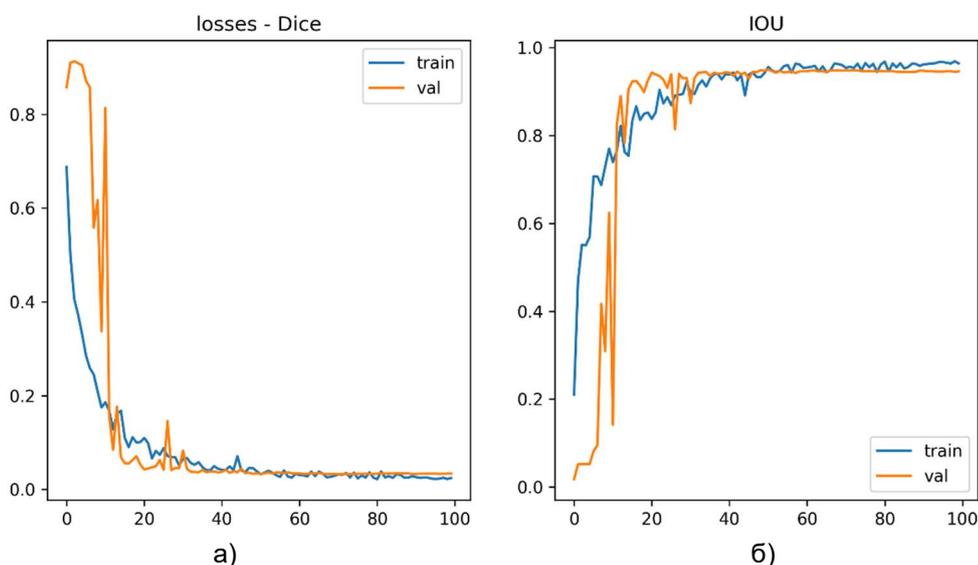


Рисунок 2 – График обучение U-Net: а – значение Dice Loss; б – значение IoU

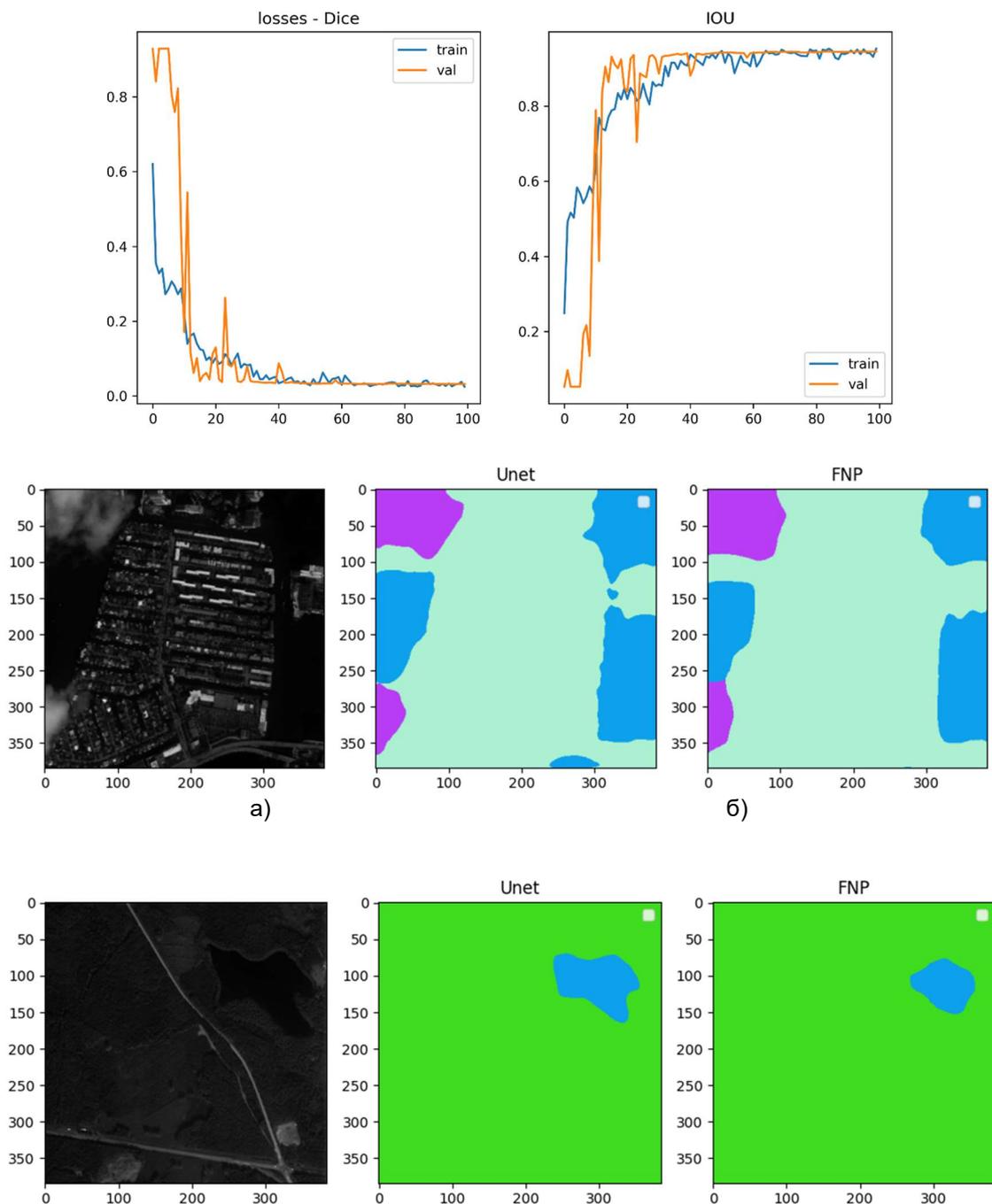


Рисунок 3 – График обучение FPN: а – значение Dice Loss; б – значение IoU

Рисунок 4 – Результат сегментации спутникового снимка, полученного при съемке города

Рисунок 5 – Результат сегментации спутникового снимка, полученного при съемке леса

Результаты проведенного исследования показали, что модель U-Net превосходит FPN по ключевым метрикам: IoU и Dice Loss (0.9187 и 0.0525 против 0.9054 и 0.0583 соответственно). Визуальный анализ подтвердил, что U-Net более точно сегментирует мелкие объекты и определяет границы между различными классами.

Таким образом, архитектура U-Net является более предпочтительной для задач, требующих высокой точности сегментации и детального распознавания объектов на спутниковых изображениях.

Список использованных источников:

1. Ronneberger O., Fischer P., Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation: [электронный ресурс] // arXiv. – 2015. – URL: <https://arxiv.org/pdf/1505.04597.pdf> (дата обращения: 14.04.2025).
2. Lin T.-Y., Dollár P., Girshick R. [и др.]. Feature Pyramid Networks for Object Detection: [электронный ресурс] // arXiv. – 2016. – URL: <https://arxiv.org/pdf/1612.03144.pdf> (дата обращения: 14.04.2025).
3. He K., Zhang X., Ren S., Sun J. Deep Residual Learning for Image Recognition: [электронный ресурс] // arXiv. – 2015. – URL: <https://arxiv.org/pdf/1512.03385.pdf> (дата обращения: 14.04.2025).

SEGMENTATION OF SATELLITE IMAGES BASED ON CONVOLUTIONAL NEURAL NETWORKS

Yerofeyev V.S.¹

Belarusian State University of Informatics and Radioelectronics¹, Minsk, Republic of Belarus

Pertseu D.Y.– PhD in Technical Sciences

Annotation. In the paper we compare the U-Net and Feature Pyramid Network convolutional neural network architectures for the satellite image segmentation. The results show that both architectures demonstrate high segmentation accuracy, with U-Net providing better detail of small objects and Feature Pyramid Network providing more stable segmentation of objects of different scales.

Keywords. Satellite images, U-Net, Feature Pyramid Network, Convolutional neural networks, object detection.