

**МОДИФИКАЦИЯ НЕЙРОСЕТЕВОЙ МОДЕЛИ U-NET ДЛЯ ПОВЫШЕНИЯ ЭФФЕКТИВНОСТИ
СЕГМЕНТАЦИИ ИЗОБРАЖЕНИЙ**

К.А. Костин, С.А. Семёнов

Научный руководитель: доцент, к.т.н. С.В. Аксёнов

Национальный исследовательский Томский политехнический университет,

Россия, г. Томск, пр. Ленина, 30, 634050

E-mail: kak@tpu.ru

**MODIFICATION OF NEURAL NETWORK MODEL U-NET TO IMPROVE THE EFFICIENCY OF
IMAGE SEGMENTATION**

K.A. Kostin, S.A. Semenov

Supervisor: Associate Professor, PhD. S.V. Axyonov

Tomsk Polytechnic University, Russia, Tomsk, Lenin str., 30, 634050

E-mail: kak@tpu.ru

***Abstract.** In the present study we developed two modifications of the U Net model, the effective segmentation algorithm and architecture used for bioimage analysis. The research uses CDNET database to evaluate the performance of the suggested models and basic architecture. The obtained results demonstrate the effectiveness of the modifications. The first model can detect all the segment locations correctly and the second one has few errors.*

Введение. Задачи сегментации изображений являются крайне актуальными в медицинской сфере, геоинформационных системах и технологиях для умного транспорта. Первые методы решения этой задачи включали подходы, основанные на цветовой фильтрации изображений [1], метрических алгоритмах кластеризации [2] или морфологических алгоритмах компьютерного зрения [3]. Такие методы, зачастую, не позволяли с достаточной точностью решать задачи на реальных данных окружающего мира. В настоящее время вышеописанным подходам пришли на смену методы машинного и, в частности, глубинного обучения, позволяющие решать сложные задачи различных предметных областей с использованием больших наборов накопленных данных. Одними из наиболее известных и успешных решений в сфере глубинного обучения на сегодняшний день являются: SegNet [4], Mask R-CNN [5], U-net [6], FgSegNet [7], другие архитектуры сверточных нейронных сетей, использующие методологию автокодирования изображений [8; 9; 10].

Целью нашего исследования является поиск методов универсализации архитектур нейронных сетей для сегментации изображений путём их модификации.

Экспериментальная часть. В работе был выбран алгоритм U-net, применявшийся ранее для сегментации биомедицинских изображений [11]. Архитектура свёрточной нейронной сети U-net использует принципы построения автоэнкодеров с применением пирамиды признаков [12], позволяющей учитывать различные масштабы сегментируемых объектов.

В исследовании проводился эксперимент по обучению и тестированию этой сети на данных естественной среды, полученных из набора CDNET 2014 [13]. Для проведения исследования была

выбрана категория baseline и серия изображений highway из CDNET, предполагающая задачу сегментации объектов-автомобилей на видео-записи транспортного потока на загородной трассе. Общее количество изображений используемых для эксперимента составляет 1700 кадров видео.

В ходе проведения исследований было осуществлено несколько экспериментов с различными версиями алгоритма U-net: исходной версией, модификацией 1 и модификацией 2. Каждая последующая модификация алгоритма была произведена на основе результатов тестирования предыдущей версии.

Первой модификацией алгоритма являлась реализация функции ошибки сети, учитывающей несбалансированность соотношения пикселей изображения, относящихся к сегментам, и остальных пикселей изображения. Перед началом обучения рассчитывается общее соотношение количества пикселей, относящихся к маске сегмента, и остальных пикселей – это используется для расчёта повышающего коэффициента балансировки. Расчёт функции ошибки осуществляется по двум компонентам: первая компонента – это значение ошибки, рассчитанное только по пикселям, соответствующим маске, умноженное на коэффициент балансировки; вторая компонента – это ошибка, рассчитанная по прочим пикселям изображения. Итоговое значение ошибки рассчитывалось как сумма двух компонент.

Второй модификацией алгоритма являлось изменение архитектуры сети путём добавления слоёв нормализации батчей. Данный тип слоёв для нейронных сетей впервые был предложен в [14] и предназначен для стандартизации распределения значений выходов свёрточных слоёв сети перед применением нелинейной функции активации, такой, например, как ReLU или LeakyReLU. Основным эффектом применения данных слоёв является более плавное изменение функции ошибки, снижение вероятности переобучения, осуществление более аккуратного градиентного спуска во время обучения сети.

Результаты. Эксперименты с исходной сетью U-net и двумя предложенными модификациями для задачи сегментации изменяющихся объектов на кадрах видео естественной среды осуществлялись на вышеописанных данных набора CDNET 2014. Для эксперимента из исходного набора вручную было отобрано 60 изображений для осуществления тестирования обученной сети, оставшиеся примеры в наборе были разделены на обучающую и валидационную подвыборки в отношении 5:1. Все модификации алгоритма обучались в равных условиях: 100 эпох обучения, 4 – размер батча обучающих примеров.

Закключение. По результатам тестирования получено, что архитектура сети U-net без изменений не позволяет выделить сегменты на изображениях, и такой алгоритм в исходном виде не применим для решения выбранной задачи. Первая модификация позволяет детектировать местоположения сегментов, однако точность такой архитектуры недостаточно высока, наблюдается высокое количество ложных срабатываний сети. Вторая модификация практически полностью решает проблему ложных срабатываний, и, при этом, имеет малое количество ошибок – местоположения сегментов достаточно близки к их истинному местоположению.

СПИСОК ЛИТЕРАТУРЫ

1. Jamil A.M. Saif, Ali Abdo Mohammed Al-Kubati, Abdultawab Saif Hazaa, Mohammed Al-Moraish. Image Segmentation Using Edge Detection and Thresholding // Proceedings of the 13th International Arab Conference on Information Technology ACIT'2012, Dec. 10-13. – 2012. – pp. 473-476.

2. Noel O'Connor¹, Sorin Sav, Tomasz Adamek, Vasileios Mezaris, Ioannis Kompatsiaris, Tsz Ying Lui, Ebroul Izquierdo, Christian Ferran Bennström, Josep R Casas. Region and Object Segmentation Algorithms in the Qimera Segmentation Platform // Proceedings of the 3rd International Workshop on Content-Based Multimedia Indexing. – 2003. – pp. 95-103.
3. Akshay P.Vartak, Dr. Vijay Mankar. Morphological Image Segmentation Analysis // International Journal of Computer Science and Applications. – 2013. – Vol. 6(2). – pp. 161-165.
4. Vijay Badrinarayanan, Alex Kendall, Robert Cipolla. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation [Электронный ресурс]. – режим доступа: <https://arxiv.org/pdf/1511.00561.pdf> (дата обращения: 01.03.2019).
5. Kaiming He, Georgia Gkioxari, Piotr Dollar, Ross Girshick Mask R-CNN [Электронный ресурс]. – режим доступа: <https://arxiv.org/pdf/1703.06870.pdf> (дата обращения: 01.03.2019).
6. Olaf Ronneberger, Philipp Fischer, Thomas Brox U-Net: Convolutional Networks for Biomedical Image Segmentation [Электронный ресурс]. – режим доступа: <https://arxiv.org/pdf/1505.04597.pdf> (дата обращения: 01.03.2019).
7. Long Ang Lim, Hacer Yalim Keles. Foreground segmentation using convolutional neural network for multiscale feature encoding // Pattern Recognition Letters. – 2018. – Vol. 112. – pp. 256-266.
8. Nikolas Lessmann, Bram van Ginneken, Pim A. de Jong, Ivana Isgum Iterative fully convolutional neural networks for automatic vertebra segmentation and identification [Электронный ресурс]. – режим доступа: <https://arxiv.org/pdf/1804.04383.pdf> (дата обращения: 01.03.2019).
9. Leena Silvester M., Govindan V.K. Convolutional Neural Network Based Segmentation // Proceedings of the 5th International Conference on Information Processing, ICIP 2011, Bangalore, India, August 5-7, 2011, CCIS, Vol.157. – 2011. – pp. 190-197.
10. Stiaan Wiehman, Hendrik de Villiers Semantic segmentation of bioimages using convolutional neural networks // Proceedings of the 2016 International Joint Conference on Neural Networks (IJCNN), July 24-29. – 2016. – DOI: 10.1109/IJCNN.2016.7727258.
11. ISBI Challenge: Segmentation of neuronal structures in EM stacks [Электронный ресурс]. – режим доступа: http://brainiac2.mit.edu/isbi_challenge/ (дата обращения: 01.03.2019).
12. Tsung-Yi Lin, Piotr Dollar, Ross Girshick, Kaiming He, Bharath Hariharan, Serge Belongie Feature Pyramid Networks for Object Detection [Электронный ресурс]. – режим доступа: <https://arxiv.org/pdf/1612.03144.pdf> (дата обращения: 01.03.2019).
13. ChangeDetection.NET (CDNET) A video database for testing change detection algorithms [Электронный ресурс]. – режим доступа: <http://jacarini.dinf.usherbrooke.ca/dataset2014/> (дата обращения: 01.03.2019).
14. Sergey Ioffe, Christian Szegedy Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift [Электронный ресурс]. – режим доступа: <https://arxiv.org/pdf/1502.03167.pdf> (дата обращения: 01.03.2019).