

УДК 004.932.1

СРАВНИТЕЛЬНЫЙ АНАЛИЗ КАЧЕСТВА СЕГМЕНТАЦИИ И ВЫЧИСЛИТЕЛЬНОЙ ПРОИЗВОДИТЕЛЬНОСТИ АРХИТЕКТУРЫ U-NET ПРИ ОБРАБОТКЕ ИЗОБРАЖЕНИЙ РАЗЛИЧНОЙ КАНАЛЬНОСТИ

Ерофеев В.С., студент

Белорусский государственный университет информатики и радиоэлектроники¹
г. Минск, Республика Беларусь

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

Аннотация. В работе проводится исследование влияния количества каналов входных изображений на результат сегментации и вычислительную эффективность нейронной сети U-Net. Экспериментальные результаты показывают, что использование RGB-изображений обеспечивает улучшение качества сегментации и более высокую вычислительную эффективность, несмотря на увеличение объема входных данных.

Ключевые слова. Семантическая сегментация, U-Net, нейронные сети, обнаружение объектов, нагрузочное тестирование, GPU.

Семантическая сегментация – важная задача компьютерного зрения, для решения которой широко применяется модель U-Net [1]. Одним из факторов, влияющих на качество сегментации и вычислительную эффективность, является число каналов входного изображения. Увеличение числа каналов потенциально позволяет модели извлекать более информативные признаки, однако одновременно повышает вычислительную сложность обработки данных.

В экспериментах использовалась архитектура U-Net с одинаковыми гиперпараметрами обучения. Изменялся только размер входного тензора: в первой конфигурации применялись изображения в градациях серого (1 канал), во второй – RGB-изображения (3 канала). Обучение выполнялось на одном и том же наборе данных.

Эксперименты проводились с использованием графического ускорителя NVIDIA Tesla T4 с объемом видеопамяти 16 ГБ. В качестве центрального процессора использовался Intel Xeon с тактовой частотой 2,00 ГГц. Объем оперативной памяти составлял 12 ГБ. Все эксперименты проводились в одинаковых вычислительных условиях без изменения аппаратной конфигурации, что обеспечивало корректность и сопоставимость результатов сравнительного анализа моделей.

Для оценки качества сегментации использовались метрики Dice Loss, характеризующая степень перекрытия масок, F1-score – гармоническое среднее precision и recall, а также Intersection over Union (IoU), определяющая отношение площади пересечения предсказанной и истинной областей к площади их объединения.

Использование нескольких метрик позволяет комплексно оценить качество сегментации, учитывая как полноту обнаружения объектов, так и точность локализации их границ.

Тестирование моделей проводилось на данных, которые ранее не использовались в обучении, для объективности оценки качества сегментации. Результаты тестирования представлены в таблице 1.

Таблица 1 – Результаты тестирования моделей

Метрика	Одноканальная модель	Трехканальная модель
Dice Loss	0,055	0,032
F1-score	0,949	0,970
IoU	0,911	0,946

Сравнительный анализ показывает, что переход к RGB-данным привёл к снижению Dice Loss на 41,8%, увеличению F1-score на 2,1% и росту IoU на 3,5%. Это подтверждает, что цветовые признаки улучшают качество сегментации и точность выделения объектов.

Нагрузочное тестирование проводилось с batch size от 1 до 64. Измерялись общее время выполнения, время прямого прохода (forward time), накладные расходы (overhead), пропускная способность (throughput), загрузка CPU и GPU, а также использование памяти.

Результаты нагрузочного тестирования для одноканальных и трехканальных изображений показаны в таблицах 2 и 3, а также на рисунках 1 и 2.

Результаты демонстрируют, что увеличение batch size улучшает пропускную способность и эффективность использования ресурсов, при этом RGB-данные позволяют достичь более высокой производительности по сравнению с одноканальными изображениями. Несмотря на то, что число входных каналов увеличилось в три раза, модель с RGB-данными показала большую пропускную способность. Это объясняется особенностями работы современных графических процессоров.

Таблица 2 – Результаты нагрузочного тестирования одноканальной модели

Batch Size	1	2	4	8	16	32	64
Iterations	1024	512	256	128	64	32	16
Total Time (s)	9,38	8,22	7,96	7,85	7,10	7,14	7,23
Forward Time (s)	5,83	2,80	1,43	0,81	0,40	0,51	0,20
Overhead (s)	3,55	5,42	6,54	7,04	6,70	6,62	7,03
Throughput (image/s)	175,68	365,72	717,34	1270,99	2560,75	1989,77	5173,35
Throughput (TB/day)	4,1280	8,5935	16,8556	29,8648	60,1709	46,7545	121,5603
CPU Load (%)	49,504	49,435	49,316	49,250	49,197	48,445	46,463
CPU Mem Used (MB)	1429,3	6739,2	8365,4	8060,3	7431,5	7168,5	7106,0
GPU Load (%)	89,91	95,73	96,79	96,74	95,91	99,38	93,75
GPU Mem Used (MB)	844,19	1046,19	1440,19	2216,19	3848,19	7100,19	13592,19

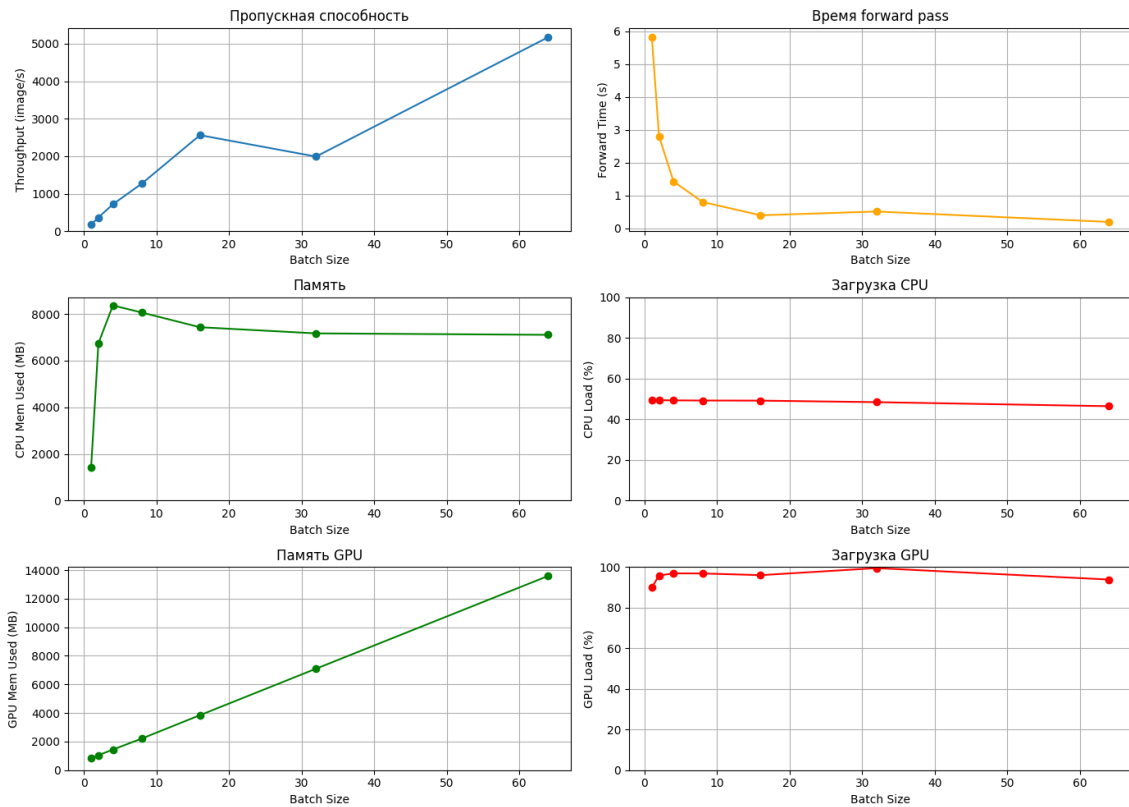


Рисунок 1 – Графики результатов нагрузочного тестирования одноканальной модели

Таблица 3 – Результаты нагрузочного тестирования трехканальной модели

Batch Size	1	2	4	8	16	32	64
Iterations	1024	512	256	128	64	32	16
Total Time (s)	8,92	5,75	5,60	5,43	4,82	4,71	4,69
Forward Time (s)	7,12	2,43	1,25	0,64	0,39	0,41	0,14
Overhead (s)	1,79	3,31	4,35	4,79	4,42	4,30	4,55
Throughput (image/s)	143,70	419,96	818,77	1593,97	2560,40	2478,75	7032,88
Throughput (TB/day)	2,3628	6,9051	13,462	26,208	42,099	40,756	115,63
CPU Load (%)	48,543	50,412	50,002	49,710	48,53	48,128	47,015
CPU Mem Used (MB)	1406,10	6759,85	8218,80	7546,90	7365,30	7059,00	6865,70
GPU Load (%)	71,97	94,26	95,83	95,88	95,00	95,25	92,75
GPU Mem Used (MB)	601,88	729,88	997,88	1501,88	2593,88	4765,88	9097,88

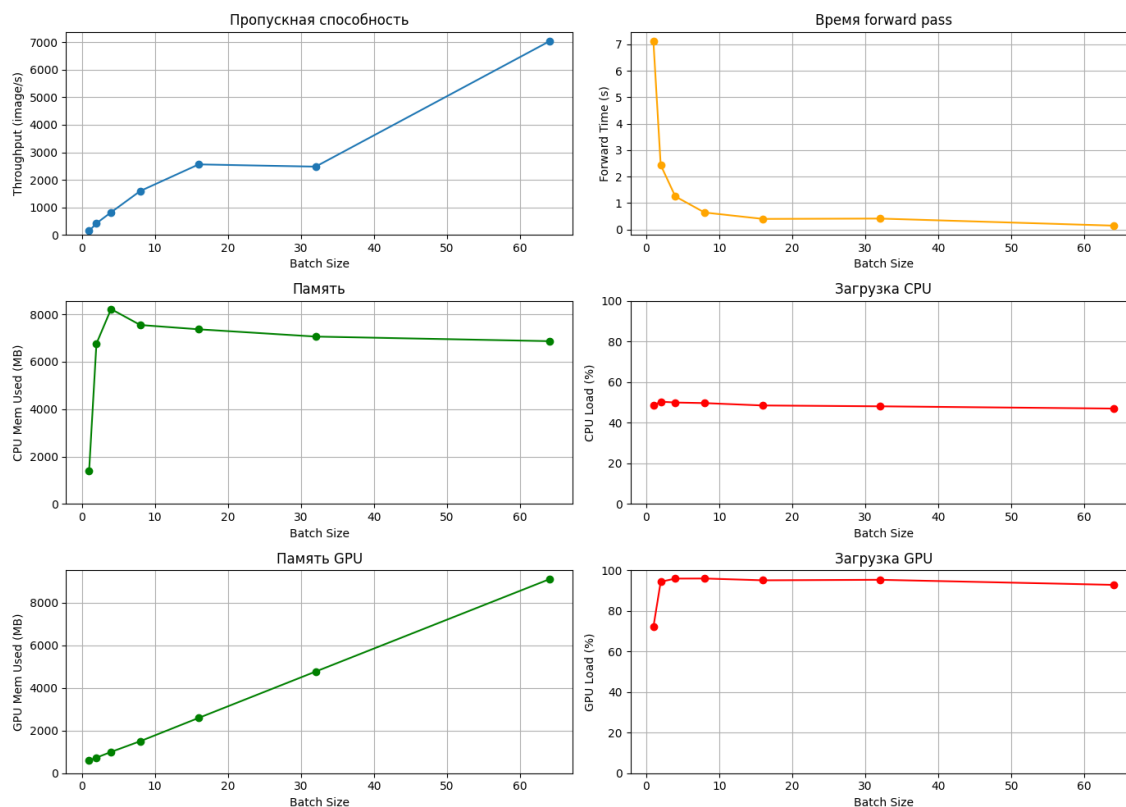


Рисунок 2 – Графики результатов нагрузочного тестирования трехканальной модели

Вычислительная сложность свёрточных операций растёт пропорционально числу каналов, однако эффективность GPU определяется не только количеством операций, но и степенью параллелизма вычислений. Одноканальные изображения формируют относительно небольшие вычислительные задачи, что приводит к неполной загрузке вычислительных блоков видеокарты. В результате значительная часть времени тратится на служебные операции, включая запуск вычислительных ядер и синхронизацию потоков. Для одноканальных данных накладные расходы составили 3,55 с. при batch size равном 1, тогда как для RGB-модели – лишь 1,79 с., что указывает на недостаточную вычислительную плотность задач в первом случае.

Увеличение числа каналов повышает арифметическую интенсивность вычислений и позволяет более эффективно использовать параллельную архитектуру GPU. При этом относительная доля накладных расходов уменьшается, а вычисления переходят в режим, ограниченный вычислительной мощностью устройства, а не задержками памяти или служебными операциями. Дополнительно, оптимизация библиотек cuDNN [2] под большие тензоры позволяет выбирать более эффективные алгоритмы свёртки, улучшая доступ к памяти и загрузку вычислительных блоков. Таким образом, RGB-вход обеспечивает более высокий уровень утилизации GPU, что и приводит к росту производительности.

Особенно заметный рост пропускной способности наблюдается на batch size равном 64, когда cuDNN перераспределяет вычислительные ресурсы в пользу более эффективных GEMM-ядер (SGEMM/volta_sgemm). При небольших размерах batch свёртки в основном обрабатывались через комбинацию Winograd и FFT, что создаёт ощутимые накладные расходы.

Полученные результаты показывают, что увеличение объёма входных данных не обязательно приводит к снижению скорости обработки. Напротив, при использовании современных GPU увеличение вычислительной нагрузки может улучшать эффективность вычислений. RGB-данные одновременно повышают качество сегментации и обеспечивают более стабильное масштабирование производительности при увеличении batch size.

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: 27.03.2026).
2. NVIDIA. cuDNN: CUDA Deep Neural Network Library. NVIDIA Developer. URL: <https://developer.nvidia.com/cudnn> (дата обращения: 27.03.2026).

UDC 004.932.1

COMPARATIVE ANALYSIS OF SEGMENTATION QUALITY AND COMPUTING PERFORMANCE OF ARCHITECTURE U-NET FOR IMAGE PROCESSING OF VARIOUS CHANNELS

Yerofeyev V., student

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

Pertseu D. – PhD in Technical Sciences

Annotation. The paper investigates the effect of the number of input image channels on the segmentation result and computational efficiency of the U-Net neural network. Experimental results show that the use of RGB images provides improved segmentation quality and higher computational efficiency, despite an increase in the amount of input data.

Keywords. Semantic segmentation, U-Net, neural networks, object detection, load testing, GPU.