

УДК 004.891.2

ОБНАРУЖЕНИЕ ПОЛИПОВ НА КОЛОНОСКОПИЧЕСКИХ ИЗОБРАЖЕНИЯХ ПРИ ПОМОЩИ АЛГОРИТМОВ НА БАЗЕ НЕЙРОСЕТЕВОЙ АРХИТЕКТУРЫ YOLO

Хрящев В.В., к.т.н., доцент кафедры цифровых технологий и машинного обучения Ярославского государственного университета им. П.Г. Демидова, e-mail: v.khryashchev@uniyar.ac.ru

Приоров А.Л., д.т.н., профессор кафедры цифровых технологий и машинного обучения Ярославского государственного университета им. П.Г. Демидова, e-mail: pri@uniyar.ac.ru

Лебедев А.А., к.т.н., ведущий инженер центра искусственного интеллекта и цифровой экономики Ярославского государственного университета им. П.Г. Демидова, e-mail: lebedevdes@gmail.com

Котов Н.А., аспирант кафедры цифровых технологий и машинного обучения Ярославского государственного университета им. П.Г. Демидова, e-mail: n.kotov2@uniyar.ac.ru

DETECTION OF POLYPS ON COLONOSCOPIC IMAGES USING ALGORITHMS BASED ON THE YOLO NEURAL NETWORK ARCHITECTURE

Khryashchev V.V., Priorov A.L., Lebedev A.A., Kotov N.A.

Neural network algorithms for detecting colon polyps in images obtained during colonoscopic examination is investigated. The Kvasir-SEG image database was used to train and test deep machine learning algorithms. YOLOv6, YOLOR, YOLOv7, YOLOv7X, and YOLOv8 networks previously trained on the basis of MS COCO images were used as neural network architectures. Due to the small volume of images in the Kvasir-SEG database, data augmentation was actively used. As a result of the application of detection algorithms to the test set of endoscopic images, the highest values of the metrics $AP@[0.25..0.75]$ – equal to 98.4, and $AP@0.50$ – equal to 98.6, were obtained for a neural network detector based on the YOLOv8 network. The obtained results can be used in the development of a video stream analysis system in an endoscopic system operating in real time, during colonoscopic examinations.

Key words: endoscopy, colonoscopic images, intestinal polyps, neural network detection algorithms, deep learning, YOLO neural network architecture.

Ключевые слова: эндоскопия, колоноскопические изображения, полипы кишечника, алгоритмы нейросетевого детектирования, глубокое обучение, нейросетевая архитектура YOLO.

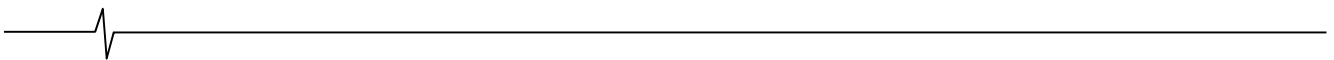
Введение

Алгоритмы и методы анализа видеоизображений в медицине находят применение в такой перспективной области диагностики, как эндоскопические исследования [1-3]. Одной из эндоскопических процедур является колоноскопия, представляющая собой дорогостоящую и достаточно трудоемкую инвазивную процедуру, требующую высокой квалификации врача, проводящего исследование.

Одним из маркеров патологий являются полипы толстой кишки. Фактически это аномальные разрастания ткани, выступающие над слизистой оболочкой. Они часто являются предвестниками достаточно распространенного колоректального рака. Полипы могут возникать в любом месте желудочно-кишечного тракта (ЖКТ), однако чаще всего их обнаруживают в колоректальной области. В толстом кишечнике выделяют несколько видов классификации полипов: по количеству, по патоморфологическим данным, по особенностям прикрепления к стенке кишки [1, 2]. Размер колоректальных полипов может существенно различаться: от совсем небольших (≤ 5 мм) до весьма значительных (≥ 10 мм).

Исследовано использование нейросетевых алгоритмов детектирования полипов толстой кишки на изображениях, полученных при колоноскопическом обследовании. Для обучения и тестирования алгоритмов глубокого машинного обучения использовалась база изображений Kvasir-SEG. В качестве нейросетевых архитектур использовались сети YOLOv6, YOLOR, YOLOv7, YOLOv7X, YOLOv8, предварительно обученные на базе изображений MS COCO. В связи с малым объемом изображений в базе Kvasir-SEG активно использовалась аугментация данных. В результате применения алгоритмов детектирования к тестовому набору эндоскопических изображений получены наивысшие значения метрик $AP@[0.25..0.75]$ – равное 98,4; и $AP@0.50$ – равное 98,6; для нейросетевого детектора на базе сети YOLOv8. Полученные результаты могут использоваться при разработке системы анализа видеопотока в эндоскопической системе, работающей в режиме реального времени, при проведении колоноскопических исследований.

Статистика эндоскопических исследований показывает, что уровень обнаружения полипов врачом высокой квалификации составляет примерно 80 % [3-5]. Новые эндоскопические устройства и диагностические инструменты улучшили точность обнаружения полипов, но проблема их пропуска остается по-прежнему актуальной [6-9]. Таким образом, задача разработки автоматических методов обнаружения полипов на видеопотоке с эндоскопа, работающих в режиме реального времени и дающих «второе мнение» для врачей, проводящих диагностику, является актуальной на данном этапе развития медицинских приборов.



Для решения этой задачи использовались классические методы машинного обучения, основанные на анализе цвета и текстуры полипов [10]. Позднее их заменили методы глубокого машинного обучения, основанные на сверточных нейронных сетях (СНС) [11-14], с автоматическим формированием признаков во время процедуры обучения. Их использование стало возможным при появлении общедоступных наборов размеченных эндоскопических изображений [15-17].

Обобщенная структурная схема разработки нейросетевого алгоритма для решения задачи обнаружения полипов на эндоскопических изображениях кишечника приведена на рис. 1. Кроме самой эндоскопической системы она включает размеченную медицинскими специалистами базу изображений, полученных с эндоскопической системы в виде отдельных кадров соответствующих видеопоследовательностей. Также в схеме используются стандартные или модифицированные алгоритмы аугментации, значительно расширяющие исходную базу изображений. Важной составляющей является также выбор нейросетевой архитектуры и процедура обучения соответствующего нейросетевого алгоритма. После выполнения указанных этапов на выходе алгоритма получаем прямоугольную область, соответствующую области нахождения полипа на исследуемом эндоскопическом изображении.



Рис. 1. Обобщенная структурная схема разработки нейросетевого алгоритма обнаружения полипов кишечника

Целью настоящей работы является исследование эффективности применения алгоритмов обнаружения

полипов при проведении колоноскопических исследований на базе современной нейросетевой архитектуры YOLO [18-20].

Обзор баз колоноскопических видеоизображений

Рассмотрим существующие открытые базы колоноскопических изображений и видеопоследовательностей с разметкой данных:

- Kvasir-SEG [21] – содержит 1000 изображений толстой кишки с экспертной разметкой полипов для задач обнаружения и сегментации;
- Kvasir [22] – содержит 8000 изображений всего ЖКТ с разметкой нескольких классов объектов и патологий;
- Nerthus [23] – содержит 21 видеопоследовательность исследования толстой кишки с разметкой классов: очищенный кишечник/не очищенный кишечник;
- HyperKvasir [24] – содержит 110079 изображений и 373 видеопоследовательности всего ЖКТ с разметкой 40 различных классов;
- CVC-Clinic [25] – содержит 612 изображений толстой кишки с разметкой полипов для задачи сегментации;
- Kvasir-Instrument [26] – содержит 590 изображений толстой кишки с разметкой медицинских инструментов для задач обнаружения и сегментации изображений.

Для проведения настоящего исследования выбрана база Kvasir-SEG, содержащая 1072 полипа и предлагающая одновременно ограничивающие прямоугольные рамки и бинарные маски для сегментации. В рассматриваемом наборе данных разрешение изображений варьируется от 332×487 до 1920×1072 пикселей. Такое число обучающих данных недостаточно для современных нейросетевых методов обнаружения объектов, поэтому здесь активно используется процедура аугментации (размножения, обогащения) данных, которая производится с учетом специфики получения эндоскопических изображений [8-9].

Архитектура сверточной нейронной сети YOLO

В рамках исследования выбраны последние версии детекторов семейства YOLO, а именно: YOLOv6, YOLOR, YOLOv7, YOLOv7X, YOLOv8. Важно отметить, что каждая последующая архитектура содержит различные модификации и самой архитектуры, и методов процедуры обучения. Рассмотрим общую архитектуру такой нейронной сети на примере YOLOv7. Процесс обработки эндоскопического изображения на основе алгоритма, использующего эту архитектуру, представлен на рис. 2.

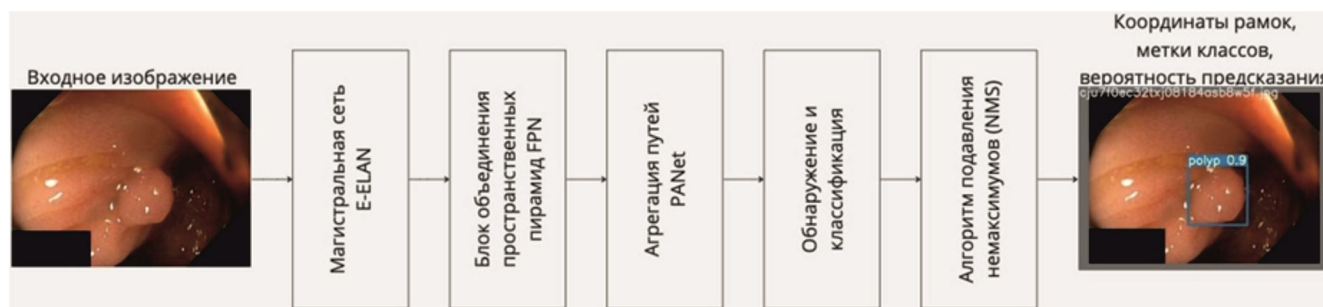


Рис. 2. Архитектура сверточной нейронной сети YOLO

Нейронная сеть семейства YOLO состоит из трех основных частей – магистральной, объединения пространственных пирамид и предсказания [19].

В качестве магистральной части сети применяется нейронная сеть с использованием блоков E-ELAN, служащая для извлечения признаков из входного изображения. Этот блок контролирует кратчайший и самый длинный градиентный пути для лучшей сходимости модели, используя элементы расширения, перетасовки и слияния. В нем входные данные подаются непосредственно через блок свертки 1×1 , в то время как два других подключения к блоку получены путем свертки входных данных двумя блоками свертки 3×3 с тем же канальным множителем. Далее все функции объединяются, и для изучения обширной информации вновь применяется свертка 1×1 .

К магистральной части добавляется часть объединения пространственных пирамид, выделяющая наиболее важные характерные особенности с разных уровней магистралей через сверточные слои CBS и SPPFCSPC. В конечных блоках пирамиды признаков свертки заменены блоками RepConv, необходимыми для обновления параметров модели.

Далее следуют слои обнаружения и классификации, дающие предсказание модели. Также используется механизм назначения меток, использующий главную и вспомогательные ветки предсказания, а также программный определитель меток. Тем самым на этапе тренировки модели используются признаки в промежуточных слоях, улучшающие итоговую предсказательную способность модели.

Сравнительный анализ нейросетевых алгоритмов обнаружения полипов

Для алгоритмов на основе архитектур YOLO использовались следующие параметры обучения: размер батча (последовательности изображений) составлял 16 изображений, в качестве алгоритма оптимизации использовался алгоритм Adam (Adaptive Moment Estimation), сочетающий

в себе идеи накопления движения и более слабого обновления весов для типичных признаков. Начальная скорость обучения -10^{-3} , общее число эпох обучения – 250. Другие параметры данного алгоритма выбирались следующим образом: $\beta_1 = 0,9$; $\beta_2 = 0,999$; $\varepsilon = 10^{-8}$.

Во время обучения использовались функции потерь с бинарной кросс-энтропией для сетей YOLOv6, YOLOv7 и YOLOR, а также функция потерь, основанная на индексе Жаккара (IoU), – для сети YOLOv8.

Для проведения сравнения с алгоритмами, приведенными в [9, 27], база Kvasir-SEG разделялась на тестовый и обучающий наборы (120 – тестовый и 880 – обучающий). Для обучения и тестирования нейронных сетей, координаты ограничивающих рамок в файлах разметки базы пересчитывались из формата Pascal-VOC в формат YOLO.

Для оценивания качества работы алгоритмов в [27] предложено использовать метрики средней точности, рассчитанные при значениях порога IoU 0,25, 0,50 и 0,75. Далее данные метрики обозначены как $AP@0,25$, $AP@0,50$ и $AP@0,75$ соответственно. Кроме того, для комплексной оценки по множеству порогов IoU предложена метрика $AP@[0,25..0,75]$, которая рассчитывается как среднее 11-ти значений AP , рассчитанных при 11 различных порогах IoU в пределах от 0,25 до 0,75 с шагом 0,05.

В табл. 1 приведены результаты тестирования алгоритмов для всех указанных нейросетевых архитектур на базе эндоскопических изображений Kvasir-SEG.

На рис. 3 представлены графики зависимости точности (P) от полноты (R) на тестовых изображениях для исследованных в работе нейросетевых алгоритмов детектирования. Видно, что лучшие результаты по обнаружению полипов показал алгоритм на основе сети YOLOv8. Значение $AP@0,50$ для него достигает 98,6; а усредненная по диапазону порогов метрика $AP@[0,25..0,75]$ – 98,4; что говорит о высокой способности детектирования полипов в широком диапазоне порогов детектирования.

Таблица 1. Тестирование нейросетевых алгоритмов обнаружения полипов

Метрика	$AP@[0,25..0,75]$	$AP@0,25$	$AP@0,5$	$AP@0,75$
Известные аналоги из [27]				
EfficientDet-D0	47,6	68,5	50,5	22,8
Faster R-CNN	78,7	89,5	84,2	56,6
RetinaNet50	87,0	93,9	90,9	69,7
RetinaNet101	87,4	94,8	90,9	71,3
YOLOv3+spp	81,0	88,6	85,3	75,9
YOLOv4	85,1	91,2	82,3	75,9
ColonSegNet	80,0	90,0	81,7	67,1
Предложенные алгоритмы на базе YOLOv5 из [9]				
HAOOSV512	90,2	96,6	93,8	74,9
HAOOSYS	88,6	94,0	90,8	73,8
HAOOSYM	91,8	95,9	93,3	82,8
HAOOSYL	92,5	96,7	93,9	85,1
Алгоритмы, исследованные в настоящей работе				
YOLOv6	89,2	93	90	84,1
YOLOR	87,8	93,2	88,2	79,3
YOLOv7	80,6	89,5	80,6	66,4
YOLOv7X	80,9	88,7	80,9	70
YOLOv8	98,4	98,7	98,6	97,6

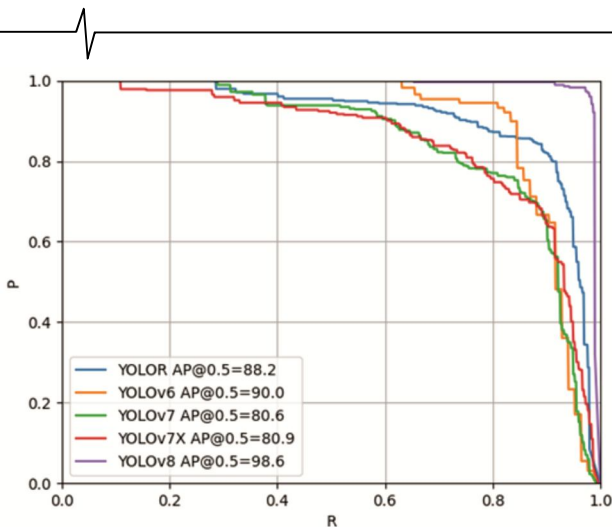


Рис. 3. Кривые на плоскости точность-полнота для исследованных нейросетевых алгоритмов

Можно отметить, что тестирование сетей YOLOv7, YOLOv7X и YOLOR показало значение метрик $AP@0,50$ меньше, чем для YOLOv6, несмотря на обратные показатели по данной метрике на больших эталонных базах. Это связано с большим числом параметров сети относительно количества обучающих примеров в базе Kvasir-SEG, несмотря на применение методов аугментации во время процесса обучения.

Таким образом, установлено, что алгоритм на базе нейронной сети YOLOv8 достигает наилучшего результата по метрикам средней точности среди всех рассматриваемых алгоритмов обнаружения полипов, причем это преимущество особенно существенно по метрике $AP@0,75$.

Заключение

Результаты проведенных исследований нейросетевых алгоритмов детектирования полипов кишечника позволяют сформулировать следующие выводы и закономерности:

- в настоящее время наблюдается существенное отставание размера имеющихся баз эндоскопических изображений от требований современных нейросетевых алгоритмов и методов глубокого машинного обучения;

- в результате тестирования нейросетевых алгоритмов обнаружения полипов на изображениях из базы Kvasir-SEG наилучший результат получен для нейросетевого алгоритма YOLOv8, для которого значение метрик $AP@[0,25..0,75]$ равно 98,4 и $AP@0,50$ – 98,6;

- при сравнении результатов работы нейросетевых алгоритмов YOLOv8 и YOLOv5 оказывается, что для первых они лучше по метрикам $AP@[0,25..0,75]$ – на 5,9 – 9,8 %; $AP@0,25$ – на 2 – 4,7 %; $AP@0,50$ – на 4,7 – 7,8 % и $AP@0,75$ – на 12,5 – 23,8 %;

- исследования, проведенные в работе, будут использованы для построения на основе нейросетевого алгоритма обнаружения полипов модуля анализа видеопотока в реальной эндоскопической системе.

Литература

1. Палевская С.А., Короткевич А.Г. Эндоскопия желудочно-кишечного тракта. М.: ГЭОТАР-Медиа, 2020. 752 с.

2. Никонов Е.Л., Аксенов В.А., Кашин С.В., Нехайкова Н.В. Международный опыт скрининга колоректального рака. Доказательная гастроэнтерология. 2017. Т. 6. № 3. С. 30-35.

3. Lee J. Resection of diminutive and small colorectal polyps: what is the optimal technique? Clinical endoscopy. 2016, vol. 49 (4). 355 p.

4. Lebedev A., Khryashchev V., Stefanidi A., Stepanova O., Kashin S., Kuvaev R. Convolutional neural network for early detection of gastric cancer by endoscopic video analysis. Proc. SPIE 11433. Twelfth International Conference on Machine Vision (ICMV 2019).

5. Батухтин Д.М., Пеганова Е.В., Митракова Н.Н. Анализ узкоспектральных эндоскопических изображений на внутренней поверхности пищевода. Вестник Поволжского государственного технологического университета. Серия: радиотехнические и инфокоммуникационные системы. 2014. № 4 (23). С. 45-57.

6. Коваленко Д.А., Гнатюк В.С. Ассоциация сцен в эндоскопических видео. GraphiCon 2017: Обработка и анализ биомедицинских изображений. Пермь. 2017. 269-274 с.

7. Matyja M., Pasternak A., Szura M., Wysocki M., Pędzwiatr M., Rembiasz K. How to improve the adenoma detection rate in colorectal cancer screening? Clinical factors and technological advancements. Archives of medical science: AMS. 2019, vol. 15 (2). 424 p.

8. Лебедев А.А., Хрящев В.В., Среднякова А.С., Казина Е.М. Разработка алгоритма детектирования полипов на эндоскопических изображениях с использованием сверточных нейронных сетей. Цифровая обработка сигналов. 2021. № 2. 55-60 С.

9. Лебедев А. Исследование нейросетевых алгоритмов обнаружения объектов на видеоизображениях в медицинских системах прикладного телевидения: диссертация на соискание ученой степени кандидата технических наук. Ярославль. 2022.

10. Karkanis S.A., Iakovidis D.K., Maroulis D.E., Karras D.A., Tzivras M. Computer-aided tumor detection in endoscopic video using color wavelet features. IEEE Trans. on Information Technology in Biomedicine. 2003, vol. 7 (3), pp. 141-152.

11. Николенко С.И., Кадуринов А.А., Архангельская Е.О. Глубокое обучение. Питер: СПб. 2018. 480 с.

12. Tajbakhsh N., Shin J.Y., Gurudu S., Hurst R., Kendall C. et al. Convolutional neural networks for medical image analysis: Full training or fine tuning? IEEE transactions on medical imaging. 2016, vol. 35(5), pp. 1299-1312.

13. Shin H.-C., Roth H., Gao M., Lu L., Xu Z. et al. Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning. IEEE Trans. on Medical Imaging. 2016, vol. 35 (5), pp. 1285-1298.

14. Wang Y., Tavanapong W., Wong J., Oh J.H., De Groen P.C. Polyp-alert: Near real-time feedback during colonoscopy. Computer Methods and Programs in Biomedicine. 2015, vol. 120 (3), pp. 164-179.

15. Shin Y., Qadir H.A., Aabakken L., Bergsland J., Balasingham I. Automatic Colon Polyp Detection Using Region Based Deep CNN and Post Learning Approaches. IEEE Access. 2018, vol. 6, pp. 40950-40962.

16. Lee J., Jeong J., Song E., Ha C., Lee H. et al. Real-time detection of colon polyps during colonoscopy using deep learning: systematic validation with four independent datasets. *Scientific Reports*. 2020, vol. 10 (1), pp. 1-9.
17. Yamada M., Saito Y., Imaoka H., Saiko M., Yamada S. et al. Development of a real-time endoscopic image diagnosis support system using deep learning technology in colonoscopy. *Scientific reports*. 2019, vol. 9 (1), pp. 1-9.
18. Li C., Li L., Jiang H., Weng K., Geng Y. et al. YOLOv6: A single-stage object detection framework for industrial applications. *arXiv preprint arXiv:2209.02976*, 2022.
19. Wang C.-Y., Bochkovskiy A., Liao H.-Y. M. Yolov7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. *arXiv preprint arXiv:2207.02696*, 2022.
20. Jocher G., Chaurasia A., Qiu J. YOLO by Ultralytics. <https://github.com/ultralytics/ultralytics>, 2023. Accessed: February 30, 2023.
21. Jha D., Smedsrud P., Riegler M., Halvorsen P., Lange T. et al. Kvasir-seg: A Segmented Polyp Dataset. *Proc. of International Conference on Multimedia Modeling*. 2020, pp. 451-462.
22. Pogorelov K., Randel K.R., Griwodz C., Eskeland S.L., Lange T. et al. Kvasir: A multi-class image dataset for computer aided gastrointestinal disease detection. *Proc. of the 8th ACM on Multimedia Systems Conference*. 2017, pp. 164-169.
23. Pogorelov K., Randel K.R., Lange T., Eskeland S.L., Griwodz C. et al. Nerthus: A bowel preparation quality video dataset. *Proc. of the ACM on Multimedia Systems Conference*. 2017, pp. 170-174.
24. Borgli H., Thambawita V., Smedsrud P.H., Hicks S., Jha D. et al. Hyperkvasir, a comprehensive multi-class image and video dataset for gastrointestinal endoscopy. *Scientific Data*. 2020, vol. 7 (1), pp. 1-14.
25. Bernal J., Sánchez F.J., Fernández-Esparrach G., Gil D., Rodríguez C., Vilarriño F. Wm-dova maps for accurate polyp highlighting in colonoscopy: Validation vs. saliency maps from physicians. *Computerized Medical Imaging and Graphics*. 2015, vol. 43, pp. 99-111.
26. Jha D., Ali S., Emanuelsen K., Hicks S.A., Thambawita V. et al. Kvasir-Instrument: Diagnostic and therapeutic tool segmentation dataset in gastrointestinal endoscopy. *International Conference on Multimedia Modeling*. 2021, pp. 218-229.
27. Jha D., Ali S., Tomar N.K., Johansen H.D., Johansen D. et al. Real-Time Polyp Detection, Localization and Segmentation in Colonoscopy Using Deep Learning. *Computer Vision and Pattern Recognition*. 2021. pp. 40496-40510.

Уважаемые коллеги!

Приглашаем Вас принять участие в формировании тематических выпусков журнала «Цифровая обработка сигналов» и размещению рекламы продукции (услуг) Вашей организации на его страницах. В случае положительного решения просим представить в редакцию журнала Ваши предложения по плановому размещению информационных материалов и макет рекламы продукции (услуг) с указанием желаемого её месторасположения: обложка (2-я, 3-я или 4-я стр.), цветная внутренняя полоса (объем полосы).

Журнал «Цифровая обработка сигналов» издается с 1999 года. Выходит ежеквартально, тиражом 200 экз.

Научно-технический журнал «Цифровая обработка сигналов» включен в Перечень изданий, рекомендуемый ВАК РФ для публикации результатов научных исследований соискателями ученой степени доктора и кандидата технических наук в области радиотехники, связи, вычислительной техники, электроники, приборостроения, информационных технологий, информационно-измерительных и управляющих систем. Журнал «Цифровая обработка сигналов» включен в базу данных Web of Science - Russian Science Citation Index.

Планируемые сроки издания отдельных номеров журнала:

- № 3 сентябрь 2023 г. Тематический выпуск: «Цифровая обработка изображений».
- № 4 декабрь 2023 г. Тематический выпуск: «ЦОС в радиотехнике и системах телекоммуникаций».
- № 1 март 2024 г. Тематический выпуск: «ЦОС в инфокоммуникационных системах».
- № 2 июнь 2024 г. Тематический выпуск по материалам 26-й Международной научно-технической конференции «Цифровая обработка сигналов и ее применение-DSPA».

Ориентировочная стоимость рекламных услуг:

- 4-я (внешняя) страница цветной обложки – 25 тысяч рублей.
- 2-я и 3-я (внутренние) страницы цветной обложки – 15 тысяч рублей.
- 1\2 цветной внутренней полосы – 8 тысяч рублей.

Ждем Ваших предложений.

С наилучшими пожеланиями, зам. главного редактора

д.т.н., профессор Витязев Владимир Викторович, телефон 8-903-834-81-81.

Предложения прошу направлять по адресу: E-mail: vityazev.v.v@rsreu.ru или info@dspa.ru