

22. Junayed Pasha, Zeinab Elmi, Sunit Purkayastha, Amir M. Fathollahi-Fard, Ying-En Ge, Yui-Yip Lau. The Drone Scheduling Problem: A Systematic State-of-the-Art Review, *IEEE Transactions on Intelligent Transportation Systems*, 2022, Vol. 23, Issue 9, pp. 14224-14247. DOI: 10.1109/TITS.2022.3155072.
23. Qi X., Chong J., Zhang Q., Yang Z. Collaborative computation offloading in the multi-UAV fledged mobile edge computing network via connected dominating set, *IEEE Trans. Veh. Technol.*, 2022, 71 (10), pp. 10832-10848. Available at: <https://doi.org/10.1109/tvt.2022.3188554>.
24. Vaswani A. et al. Attention is all you need, *31st Conference on Neural Information Processing Systems (NIPS 2017)*, 2017, pp. 11.
25. Wu X., Gao Z., Yuan S., Hu Q., Dang Z. A dynamic task allocation algorithm for heterogeneous UAV swarms, *Sensors*, 2022, 22 (6), 2122, pp. 18. DOI: 10.3390/s22062122.
26. Ahmad S., Zhang J., Khan A., Khan U.A., Hayat B. JO-TADP: learning-based cooperative dynamic resource allocation for MEC-UAV-enabled wireless network, *Drones*, 2023, 7 (5), 303, pp. 35. DOI: 10.3390/drones7050303.
27. Bai Y., Zhao H., Zhang X., Chang Z., Jäntti R., Yang K. Toward autonomous multi-UAV wireless network: a survey of reinforcement learning-based approaches, *IEEE Commun. Surv. Tutorials.*, 2023, 25 (4), pp. 3038-3067. DOI: 10.1109/COMST.2023.3323344.
28. Zhu R., Huang M., Sun K., Hou Y., Wan Y., He H. Deep Reinforcement Learning Based Task Offloading for UAV-Assisted Edge Computing, *IEEE International Conference on Unmanned Systems (ICUS)*, 2023: IEEE, pp. 1104-1111.
29. Ding Y., Yang Z., Pham Q.V., Hu Y., Zhang Z., Shikh-Bahaei M. Distributed machine learning for UAV swarms: computing, sensing, and semantics, *IEEE Internet Things J.*, 2023. DOI: 10.1109/IJOT.2023.3341307.
30. Sutton R.S., Barto A.G. Reinforcement Learning: An Introduction. A Bradford Book, 2018, 526 p.

**Абросимов Вячеслав Константинович** – Главный научно-исследовательский испытательный межвидовой центр перспективного вооружения Минобороны России; e-mail: avk787@yandex.ru; г. Москва, Россия; тел.: 89168153512; д.т.н.; член-корреспондент РАН; с.н.с.

**Долгов Глеб Александрович** – Главный научно-исследовательский испытательный межвидовой центр перспективного вооружения Минобороны России; e-mail: gleb\_dolgov@mail.ru; г. Москва, Россия; тел.: +79133043935; с.н.с.

**Михайлова Екатерина Сергеевна** – Главный научно-исследовательский испытательный межвидовой центр перспективного вооружения Минобороны России; e-mail: ekaterinaolimp99@mail.ru; г. Москва, Россия; тел.: 89175192916; м.н.с.

**Abrosimov Vyacheslav Konstantinovich** – Main Research and Testing Interspecific Center for Advanced Weapons of the Russian Ministry of Defense; e-mail: avk787@yandex.ru; Moscow, Russia; phone: +79168153512; dr. of eng. sc.; Corresponding Member of RARAS; senior researcher.

**Dolgov Gleb Alexandrovich** – Main Research and Testing Interspecific Center for Advanced Weapons of the Russian Ministry of Defense; e-mail: gleb\_dolgov@mail.ru; Moscow, Russia; phone: +79133043935; senior researcher.

**Mikhailova Ekaterina Sergeevna** – Main Research and Testing Interspecific Center for Advanced Weapons of the Russian Ministry of Defense; e-mail: ekaterinaolimp99@mail.ru; Moscow, Russia; phone: +79175192916; junior researcher.

УДК 004.932:004.89

DOI 10.18522/2311-3103-2025-2-19-30

**М.А. Астапова**

### **ИНТЕЛЛЕКТУАЛЬНОЕ УПРАВЛЕНИЕ РОБОТОТЕХНИЧЕСКИМИ СРЕДСТВАМИ В ЗАДАЧЕ СЕГМЕНТАЦИИ НОР ГРЫЗУНОВ С ИСПОЛЬЗОВАНИЕМ ГЛУБОКИХ СВЕРТОЧНЫХ АРХИТЕКТУР**

*Исследуется применение нейросетевых архитектур для семантической сегментации нор грызунов с целью мониторинга их популяции на сельскохозяйственных полях. В частности, рассматриваются три модели для семантической сегментации: сверточный автокодировщик (CAK), SegNet и U-Net. Эти модели применяются для анализа изображений, полученных с беспилотных*

летательных аппаратов (БпЛА) и наземных роботизированных средств (РТС), что позволяет автоматически выявлять норы, минимизируя необходимость в трудозатратах при обработке больших объемов данных. Для обучения и тестирования моделей была подготовлена выборка, включающая 247 RGB-изображений, содержащих 1098 размеченных нор. Оценка показателей качества семантической сегментации проводилась с использованием метрики Джаккара (IoU), в результате чего были получены следующие значения: 0,511 для САК, 0,548 для SegNet и 0,529 для U-Net. Была проведена оценка вычислительных ресурсов, необходимых для внедрения этих моделей в бортовые вычислительные устройства (БВУ) мобильных РТС. Рассмотрены два критерия: количество операций с плавающей точкой (GFLOPS) и количество параметров моделей. Результаты показали, что SegNet требует 2,23 GFLOPS и имеет 0,76 миллиона параметров, что в 2,58 и 2,33 раза меньше по сравнению с САК и U-Net соответственно. Количество операций с плавающей точкой для SegNet также оказалось на 2,43 и 1,88 раза ниже, чем у САК и U-Net соответственно. В результате, SegNet превзошла САК и U-Net как в эффективности сегментации, так и в требуемых вычислительных ресурсах. Данная работа выполнялась в рамках реализации системы компьютерного зрения сельскохозяйственной РТС.

*Сельскохозяйственная робототехника; интеллектуальное управление роботами; компьютерное зрение; семантическая сегментация; глубокое обучение.*

**М.А. Астапова**

### **INTELLIGENT CONTROL OF ROBOTICS AT RODENT BURROW SEGMENTATION USING DEEP CONVOLUTIONAL ARCHITECTURES**

*In this paper, we investigate the application of neural network architectures for semantic segmentation of rodent burrows for monitoring their population in agricultural fields. In particular, three models for semantic segmentation are considered: convolutional autoencoder (CAE), SegNet, and U-Net. These models are applied to analyze images obtained from unmanned aerial vehicles (UAVs) and ground robotic means, which allows for automatic burrow detection, minimizing the need for labor costs in processing large amounts of data. A sample of 247 RGB images containing 1098 labeled burrows was prepared for training and testing the models. The quality indicators of semantic segmentation were assessed using the Jaccard metric (IoU), which resulted in the following values: 0.511 for CAE, 0.548 for SegNet, and 0.529 for U-Net. An assessment of the computational resources required to implement these models in on-board computing units (OCUs) of mobile robotic means was conducted. Two criteria were considered: the number of floating-point operations (GFLOPS) and the number of model parameters. The results showed that SegNet requires 2.23 GFLOPS and has 0.76 million parameters, which is 2.58 and 2.33 times less than SAE and U-Net, respectively. The number of floating-point operations for SegNet was also 2.43 and 1.88 times lower than that of SAE and U-Net, respectively. As a result, SegNet outperformed SAE and U-Net in both segmentation efficiency and required computational resources. This work was carried out as part of the implementation of a computer vision system for an agricultural robotic means.*

*Agricultural robotics; intelligent robot control; computer vision; semantic segmentation; deep learning.*

**Введение.** Одним из важнейших этапов контроля численности грызунов является выявление нор на сельскохозяйственных территориях [1, 2]. Традиционно этот процесс осуществляется вручную – работники обходят поля и подсчитывают норы, что требует значительных временных и трудовых затрат. В связи с этим актуальной задачей становится автоматизация данного процесса. Перспективным направлением является использование роботизированных систем, включая беспилотные летательные аппараты (БпЛА) и наземные робототехнические средства, оснащённые системами компьютерного зрения. Эти технологии обеспечивают автономное перемещение по полю, а также сбор и анализ изображений или видеоданных поверхности сельскохозяйственных угодий. Автоматизированная обработка этих данных направлена на локализацию нор, что предполагает точное определение их пространственного расположения.

Задача локализации объектов может быть решена различными методами: через детекцию, которая включает определение координат прямоугольной рамки, охватывающей локализуемые объекты, или с помощью сегментации, заключающейся в формировании маски, повторяющей форму объекта определённого класса на изображении [3]. В отличие от детекции, основным отличием методов сегментации является именно определение

формы объектов. Задача сегментации делится на несколько типов: семантическую [4], объектную [5] и паноптическую [6]. Дополнительные процедуры в объектной и паноптической сегментации увеличивают сложность обработки обучающих выборок и снижают устойчивость процесса обучения. Также эти процедуры значительно повышают вычислительную сложность моделей и замедляют обработку данных, что ограничивает их применение на бортовых вычислительных устройствах сельскохозяйственных роботов. На мобильных робототехнических средствах необходимо выбирать модели семантической сегментации, обладающие минимальной вычислительной сложностью при достижении ими приемлемых результатов сегментации.

На данный момент в открытом доступе отсутствуют выборки изображений сельскохозяйственных полей, на которых расположены норы грызунов, что делает актуальной задачу сбора и разметки таких данных.

В работе предлагается сравнительное исследование глубоких сверточных нейросетевых архитектур, выполняющих семантическую сегментацию: CAK [7–9], SegNet [10], U-Net [11].

Описанное в данной работе исследование выполнялось в рамках реализации системы компьютерного зрения сельскохозяйственного робототехнического средства [12] представленного на рис. 1.



Рис. 1. Сельскохозяйственная робототехническая платформа

Одной из функций мобильного робототехнического средства является автоматическая локализация нор грызунов, которая выполняется для сброса туда отравленного зерна. Система травления состоит из RGB камеры, которая выполняет съемку поверхности поля, на котором располагаются норы грызунов; движущейся каретки, соединенной с трубкой, через которую выполняется сбрасывание отравляющего вещества против грызунов; каретка двигается по направляющим полозьям; движение осуществляется шаговыми двигателями. Сброс отравляющих веществ выполняется специализированной системой, состоящей из: воздушного компрессора, который через трубку из контейнера выдувает в обнаруженную нору отравляющие вещества. Бортовое вычислительное устройство (БВУ) мобильной платформы получает RGB-кадры поля, которые затем подвергаются семантической сегментации. В результате сегментации определяются координаты найденных нор, в которые система сбрасывает отравляющие вещества.

**Обзор существующих подходов.** В исследовании [13] изучается обнаружение и классификация кластеров нор полёвки Брандта с использованием спутниковых снимков GF-2 и модели Faster R-CNN. Рассмотрены три подхода: OBIC (объектно-ориентированный анализ), BTC (текстурная классификация) и BVIC (по вегетационному индексу). OBIC показал наилучшие результаты (AP 63,80%, F1 = 0,722), BTC – (AP 55,95%, F1 = 0,666), BVIC – наименьшую точность (AP 29,45%, F1 = 0,437).

В работе [14] представлен алгоритм для обнаружения местоположения нор, созданных грызунами, с использованием аэрофотосъемки с помощью БПЛА. Для классификации изображений использовалась модель опорных векторов (SVM) с радиальной базисной функцией. Точность классификации зависела от ширины ядра SVM: максимальная точность была достигнута при ширине ядра 0,1, с точностью 99,07% на обучающих данных и 96,5% на тестовых данных. Общая точность предложенного метода по обнаружению местоположения нор составила 96,15%.

Авторы работы [15] использовали машинное обучение и аэрофотосъемку с БпЛА для обнаружения нор полёвки в сельском хозяйстве. Модель YOLOv3, обученная на мультимодальных данных, показала высокую точность в обнаружении нор полёвки, достигая средней точности (mAP) 0,86 и F1-оценки около 93,39%.

В исследовании [16] предложена усовершенствованная модель CGT-YOLOv5n для обнаружения нор грызунов в сложных луговых условиях с БпЛА. Она включает: SAM (повышает точность детекции мелких объектов и снижает влияние теней), ODCov (адаптирует модель к разным изображениям), TSCODE (оптимизирует контекст для классификации и регрессии). При IoU = 0,5 средняя точность (mAP<sub>0,5</sub>) достигает 92,8% (+3,3% к базовой модели), а mAP<sub>0,5:0,95</sub> — 46,2% (+4,3%).

Авторами [17] была разработана методика автоматического, точного и эффективного обнаружения нор полевых мышей в степных экосистемах. Для этого использовались БпЛА для сбора изображений и методы глубокого обучения для их анализа. Были протестированы шесть моделей глубокого обучения для детекции объектов: Faster R-CNN, R-FCN, Cascade R-CNN, SSD, RetinaNet и YOLOv4. Наилучшие результаты показала модель YOLOv4, достигшая точности обнаружения (Precision) 94,3% и полноты (Recall) 93,7%.

В статье [18] представлена усовершенствованная модель детекции объектов BSM-YOLO, основанная на архитектуре YOLOv5, для точного и эффективного обнаружения нор грызунов в степных экосистемах. Модель BSM-YOLO была обучена на наборе данных, содержащем 2397 изображений нор грызунов. Результаты показали, что BSM-YOLO достигла средней точности обнаружения (mAP) в 94,5%, что на 5,4% выше по сравнению с базовой моделью YOLOv5s. Кроме того, скорость детекции увеличилась на 8,7 кадров в секунду.

В работе [19] предложен метод обнаружения нор полевых мышей на пастбищах с использованием глубокого обучения. Разработана CNN-модель для анализа изображений с БпЛА. Обучение проводилось на 5181 изображении, тестирование на 518 изображениях. Faster R-CNN и улучшенный SSD показали схожую точность, однако SSD оказался быстрее в 5 раз. Лучшие метрики: Precision – 98,8% (SSD), Recall – 97,5% (Faster R-CNN), F1-Score – 97,9% (Faster R-CNN).

Большинство современных подходов требуют ручного анализа и ограничены задачами классификации без точной локализации нор. Таким образом, в рамках исследования предполагается проведение сбора и разметки выборки изображений с норами грызунов в контексте задачи сегментации, сравнительного исследования эффективности различных архитектур глубоких нейронных сетей, предназначенных для выполнения семантической сегментации изображений.

**Материалы и методы.** В рамках работы была собрана и размечена выборка изображений нор различных грызунов. Изображения были получены из открытых источников в сети Интернет посредством поисковых запросов в системе Google Image [20] «rodent burrows», «rodent holes», «mice burrows», «mice holes», «hamster burrows», «hamster holes» и т.д. В результате была составлена выборка из 247 изображений. При этом, на каждом изображении находится хотя бы одна нора. Каждое из 247 изображений было размечено вручную в соответствии с задачей семантической сегментации. Разметка выполнялась посредством программного средства [21]. В табл. 1 содержится наиболее значимые параметры размеченной выборки изображений.

Таблица 1

Параметры выборки

Параметр выборки	Значение параметра
Общее количество изображений	247
Общее количество нор на изображениях	1098
Среднее количество нор на одном изображении	4,4
Общее количество пикселей	$3 \times 10^9$
Количество пикселей, принадлежащих классу «норы»	$7,2 \times 10^6$
Количество пикселей, принадлежащих классу «фон»	$2,95 \times 10^9$

Примеры изображений из собранной и размеченной выборки приведены на рис. 2.

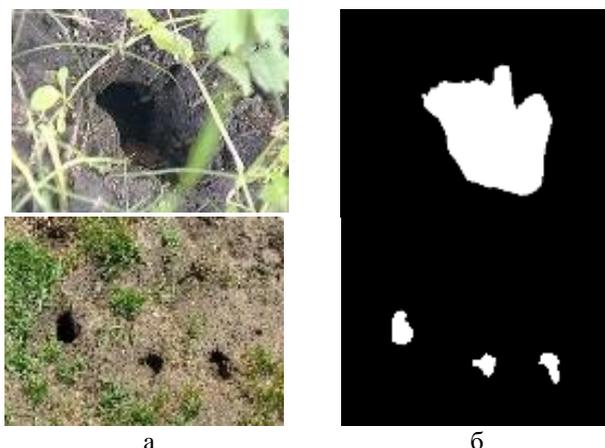


Рис. 2. (а) Изображения и (б) Размеченные сегментационные карты выборки данных

**Описание используемых нейронных сетей.** В работе было выполнено сравнительное исследование трех распространенных нейросетевых архитектур: САК, U-Net, SegNet.

**САК.** Архитектура нейронной сети САК приведена на рис. 3.

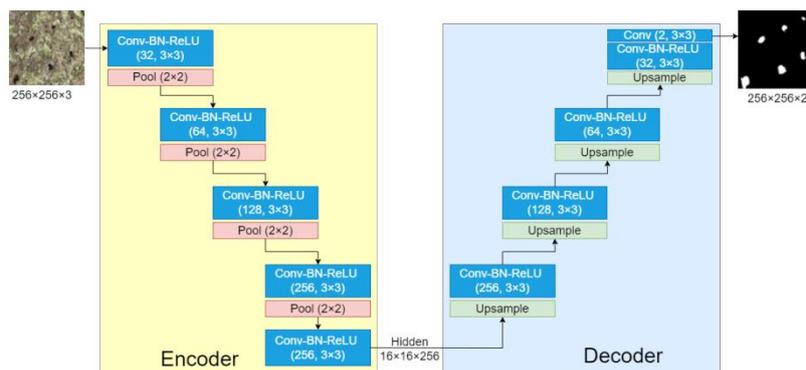


Рис. 3. Архитектура САК

Нейронная сеть данной архитектуры состоит из двух симметричных частей – энкодера и декодера (блоки Encoder и Decoder на рис. 3). Энкодер выполняет обработку поступающего на его вход RGB изображения размером 256×256 пикселей и выделяет в нем наиболее значимую информацию. Энкодер состоит из блоков обработки информации, куда последовательно включены слои свертки, пакетной нормализации (Batch Normalization – BN) [22], активации ReLU [23]. На рис. 3 это изображено как единый блок Conv-BN-ReLU, при этом, в скобках для каждого блока указано количество ядер свертки и размер каждого ядра свертки (например, 16, 3×3). Блоки Conv-BN-ReLU перемежаются блоками субдискретизации – отбора значимой информации и снижения размерности (Pool). В данной операции исходная карта признаков разбивается на равные участки размером 2×2, и на каждом таком участке производится отбор максимального значения. По мере углубления обработки информации в автокодировщике, количество карт признаков увеличивается при снижении размера самих карт признаков, в результате чего формируется промежуточное представление, которое на рис. 3 обозначено как Hidden. Декодер выполняет обратную операцию: обрабатывая промежуточное представление, декодер выполняет построение результирующей сегментационной маски, поэтапно повышая размер промежуточных карт признаков и уменьшая их количество. Само увеличение размера карт признаков осуществляется посредством операции билинейной интерпо-

ляции карт признаков (блок Upsample на рис. 3). Обработка карт признаков, полученных в результате операции Upsample, также выполняется блоками Conv-BN-ReLU, описанных выше. На выходе архитектуры формируется множество результирующих карт признаков, количество которых равняется количеству классов сегментации, а размер выходных карт равен размеру входного изображения. При этом, каждая карта выходных признаков содержит в себе сегментационные маски одного класса. На выходе нейронной сети применяется активационная функция softmax:

$$\text{softmax}(z_i) = \frac{e^{z_i}}{\sum_{j=1}^N e^{z_j}},$$

где  $z_i$  – это значение  $i$ -го нейрона выходного слоя до применения активационной функции,  $N$  – количество классов.

**SegNet.** Особенности архитектуры SegNet отражены на рис. 4.

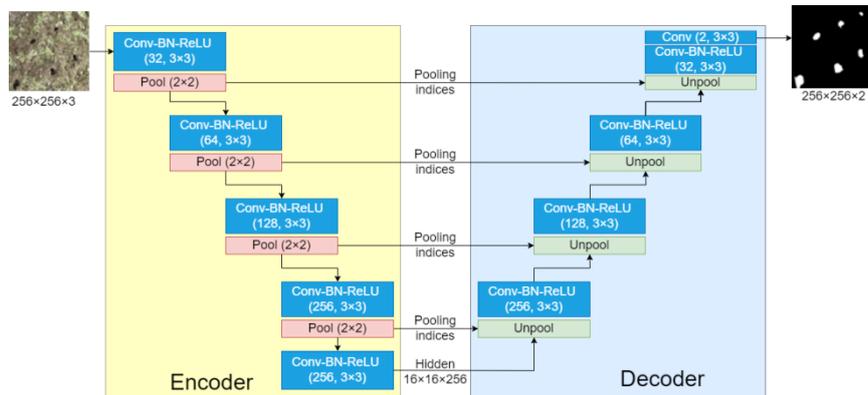


Рис. 4. Архитектура SegNet

SegNet также является глубокой сверточной архитектурой, также реализующей принципы автокодировщика. Отличительной особенностью этой нейронной сети является применение специализированных средств для повышения размерностей карт признаков в декодере: на этапе обработки информации кодировщиком при снижении размерностей карт признаков посредством операции субдискретизации (блок Pool на рис. 4) для каждого отобранного максимального элемента сохраняется также и его индекс. Эти индексы учитываются уже в декодере на этапе повышения размерности карт признаков (блок Unpool на рис. 4) следующим образом: в новую карту признаков большей размерности копируются элементы из карт признаков, полученных на предыдущем слое. Самое важное, что копирование выполняется в элементы, индексы которых совпадают с индексами максимальных элементов на соответствующем этапе субдискретизации в энкодере. Данная операция имеет два существенных преимущества: во-первых, в декодере выполняется учет информации, получаемой в энкодере, а во-вторых, операция повышения размерности представляет собой простое копирование и не требует дополнительных математических вычислений. Остальные элементы данной нейронной сети полностью воспроизводят САК.

**U-Net.** Архитектура U-Net приведена на рис. 5.

Нейронная сеть архитектуры U-Net также реализует принципы обработки информации автокодировщиком. В этой нейронной сети, также как и в SegNet в декодере происходит учет информации, получаемой в энкодере. Однако, реализовано это при помощи другой операции: выполняется конкатенация (блок Concatenate на рисунке 6 увеличиваемой карты признаков с картой признаков того же размера, получаемой из симметричного слоя энкодера. Сформированная таким образом карта признаков обрабатывается далее следующим сверточным слоем. Повышение размерности карт признаков, также как и в САК, выполняется посредством операции билинейной интерполяции (блок Upsample на рис. 5). Остальные элементы данной нейронной сети полностью воспроизводят САК.

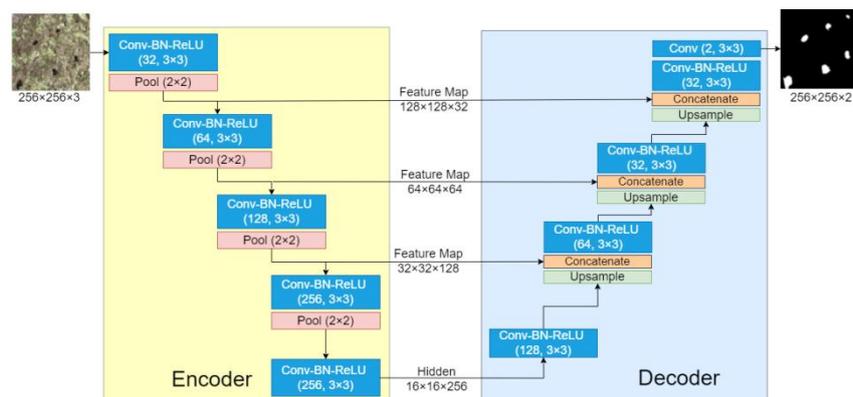


Рис. 5. Архитектура U-Net

Все приведенные нейронные сети реализуют архитектуру автокодировщика. Другой общей особенностью всех приведенных архитектур является одинаковое устройство сверточных слоев и в энкодере, и в декодере. Различаются эти слои только лишь в особенностях передачи информации между энкодером и декодером.

**Оценка вычислительной сложности рассматриваемых моделей.** В современной робототехнике в качестве БВУ используются маломощные с вычислительной точки зрения ЭВМ, обладающие низким уровнем энергопотребления. Таким образом, для реализации глубоких нейронных сетей необходима также оценка требований к вычислительным ресурсам, необходимым для их функционирования.

Для определения вычислительных требований к рассматриваемым нейросетевым моделям была выполнена оценка необходимого для каждой модели количества операций с плавающей точкой (floating point operations – FLOPS), количества параметров моделей и размера необходимой для их хранения памяти. В случае маломощных БВУ наиболее предпочтительными являются те, которые требуют меньшее количество операций с плавающей точкой и обладают меньшим количеством параметров. Результаты сведены в табл. 2.

Таблица 2

**Вычислительная сложность и требования к памяти для рассматриваемых нейронных сетей**

Модель	Количество операций с плавающей точкой, GFLOPs	Количество параметров	Размер, МБ
САК	5,42	1 958 882	7,84
SegNet	2,23	760 130	2,93
U-Net	4,21	1 773 890	7,1

Из таблицы видно, что наиболее оптимальным с вычислительной точки зрения является применение архитектуры SegNet, которая имеет в 2,58 раз меньше параметров по сравнению с САК и в 2,33 раза меньше параметров по сравнению с U-Net. Также SegNet требует в 2,43 раза меньше операций с плавающей точкой по сравнению с САК и в 1,88 раз меньше операций с плавающей точкой по сравнению с U-Net.

**Описание экспериментов.** Каждая из описанных в предыдущем разделе нейронных сетей была обучена на выборке данных, описанной в разделе «Описание набора данных». Обучение выполнялось следующим образом: сначала общая выборка была разделена на обучающую и тестовую части: 50% попало в обучающую часть, 50% тестовую. Такое разбиение объясняется, прежде всего, ограниченным размером выборки, что обуславливает необходимость увеличения объема тестовой выборки для повышения надежности результатов метрик. Разбиение выполнялось посредством генератора случайных чисел из равномерного распределения, взятого из библиотеки random языка программирования python. Параметр зерна генератора был взят равным нулю (seed = 0). Обучение выполнялось при помощи библиотеки глубокого обучения pytorch [24]. Каждая модель

обучалась 2000 эпох при помощи модифицированного алгоритма стохастического пакетного (mini-batch) градиентного спуска Adam [25] со стандартными параметрами из библиотеки pytorch. Для предотвращения переобучения были применены следующие методы аугментации изображений обучающей выборки: случайное зеркальное отображение изображений по горизонтали и по вертикали, поворот изображений на случайный угол, взятый из равномерного распределения с диапазоном значений  $[-15, 15]$  угловых градусов. В качестве функции потерь была взята функция перекрестной энтропии:

$$E(\hat{y}, y) = -y \log \hat{y} - (1 - y) \log (1 - \hat{y}),$$

где  $y$  – это значение истинной метки класса  $\hat{y}$  – значение метки класса, сгенерированное нейронной сетью.

Для общей оценки эффективности семантической сегментации используется индекс Джаккара. Данный индекс вычисляется как отношение пересечения истинной площади маски объекта и площади маски объекта, сгенерированной нейронной сетью, к площади их объединения (Intersection Over Union – IoU):

$$IoU = \frac{|S_p \cap S_t|}{|S_p \cup S_t|} = \frac{TP}{TP + FP + FN},$$

где  $S_p$  – площадь маски, сгенерированной нейронной сетью,  $S_t$  – это площадь истинной маски объекта, TP (True Positive) – это количество истинноположительно распознанных пикселей, FP (False Positive) – это количество ложноположительно распознанных пикселей, FN (False Negative) – это количество ложноотрицательно распознанных пикселей. Данная метрика вычисляется для каждого из классов, представленных в выборке. Для того, чтобы оценить общую эффективность семантической сегментации, используется среднее значение метрики Джаккара:

$$mIoU = \frac{1}{N} \sum_{i=1}^N IoU_i,$$

где  $N$  – это количество классов,  $IoU_i$  – значение метрики Джаккара, вычисленная для  $i$ -го класса.

**Результаты и обсуждение.** Результаты описанных моделей сведены в табл. 3.

Таблица 3

**Результаты сегментации IoU на тестовом наборе**

Модель	$mIoU$	IoU для отдельных классов	
		Фон	Норы
CAK	0,746	0,983	0,511
<b>SegNet</b>	<b>0,766</b>	<b>0,983</b>	<b>0,548</b>
U-Net	0,756	0,983	0,529

Архитектура SegNet показала лучшие результаты по производительности на описанной выборке данных. Кроме того, как показано в табл. 2, данная нейронная сеть также имеет меньшее количество параметров (в 2,58 и в 2,33 раз меньше по сравнению с CAK и U-Net соответственно), а также меньшее количество необходимых операций с плавающей точкой по сравнению с другими рассматриваемыми нейронными сетями (в 2,43 и в 1,88 раз меньше по сравнению с CAK и U-Net соответственно). Лучшие результаты, полученные с использованием наиболее простой модели (с точки зрения количества параметров и требуемого количества операций с плавающей точкой), могут быть объяснены следующим образом: в данном случае нейронная сеть обучалась и тестировалась на выборке данных невысокой сложности. Это обусловлено тем, что выборка содержала всего два класса объектов, сами объекты (норы грызунов) имеют гомогенный характер. В таких условиях использование сложных моделей оказывается избыточным, что подтверждено результатами экспериментальных испытаний [26]. Продемонстрированные SegNet результаты показывают наибольшую перспективность данной нейронной сети для реализации на БВУ мобильных роботов не только с точки зрения эффективности сегментации нор грызунов, но и с точки зрения требований к вычислительным возможностям БВУ.

Другим важным выводом из результатов является то, что при сегментации нор грызунов наилучшие показатели демонстрируют модели, в декодерах которых используются механизмы учета карт признаков, сформированных на промежуточных слоях энкодера. Это подтверждается более высокими значениями IoU, достигаемыми нейросетевыми архитектурами SegNet и U-Net, каждая из которых реализует этот механизм по-своему (рис. 3–5). Примеры сегментационных масок представлены на рис. 6.

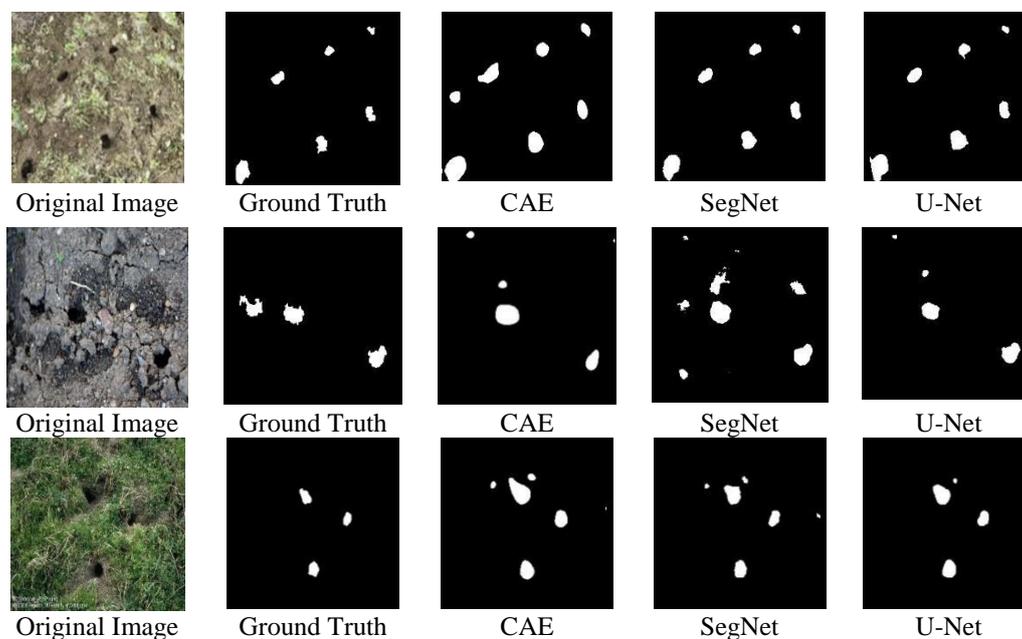


Рис. 6. Визуализация сегментационных масок нор грызунов

**Заключение.** В рамках данной работы в задаче локализации нор грызунов на изображениях сельскохозяйственных полей были применены нейросетевые архитектуры семантической сегментации (CAE, SegNet, U-Net), которые были обучены протестированы на новой выборке изображений, собранной в рамках этой работы. Также, в целях оценки реализуемости указанных архитектур на БВУ мобильных роботов была выполнена оценка необходимых для каждой архитектуры вычислительных ресурсов. В результате архитектура SegNet показала лучшие результаты как с точки зрения эффективности выполнения семантической сегментации ( $mIoU=0,766$ , значение  $IoU$  для мышинных нор равняется 0,548), так и с точки зрения необходимых вычислительных ресурсов (SegNet имеет в 2,58 и в 2,33 раз меньше параметров по сравнению с CAE и U-Net соответственно, а также требует в 2,43 и в 1,88 раз меньше операций с плавающей точкой по сравнению с CAE и U-Net соответственно). Такие результаты объясняются, в первую очередь, особенностями обучающей выборки: всего два класса разметки, единообразие или относительная гомогенность данных в выборке. Полученные результаты позволяют выбрать SegNet в качестве архитектуры, использующейся в системе компьютерного зрения сельскохозяйственной робототехнической платформы. В дальнейшем планируется выполнить исследование работоспособности SegNet непосредственно на мобильной сельскохозяйственной робототехнической платформе в ходе автономного сбора данных с реальных сельскохозяйственных полей. Также планируется исследовать семантическую сегментацию в смежной задаче локализации нор грызунов на аэрофотоснимках сельскохозяйственных полей, полученных посредством БПЛА.

*Работа выполнена в рамках бюджетной темы № FFZF-2025-0012.*

## БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Witmer G.* Rodents in Agriculture: A Broad Perspective // *Agronomy*. – 2022. – Vol. 12, No. 6. – P. 1458.
2. *Old J. M., Lin S.H., Franklin M.J.M.* Mapping out bare-nosed wombat (*Vombatus ursinus*) burrows with the use of a drone // *BMC Ecology*. – 2019. – Vol. 19, No. 1. – P. 1-10.
3. *Фаворская М.Н., Нишчхал Н.* Верификация разливов нефти на водных поверхностях по аэрофотоснимкам на основе методов глубокого обучения // *Информатика и автоматизация*. – 2022. – Т. 21, № 5. – С. 937-962.
4. *Hao S., Zhou Y., Guo Y.* A brief survey on semantic segmentation with deep learning // *Neurocomputing*. – 2020. – Vol. 406. – P. 302-321.
5. *Hafiz A. M., Bhat G. M.* A survey on instance segmentation: state of the art // *International Journal of Multimedia Information Retrieval*. – 2020. – Vol. 9, No. 3. – P. 171-189.
6. *Kirillov A., He, K., Girshick R., Rother C., Dollár P.* Panoptic segmentation // *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. – 2019. – P. 9404-9413.
7. *Geddes A.M.W.* The relative importance of pre-harvest crop pests in Indonesia // *Bulletin-Natural Resources Institute*. – 1992. – No. 47.
8. *Leirs H.* Management of rodents in crops: the Pied Piper and his orchestra // *ACIAR Monograph Series*. – 2003. – Vol. 96. – P. 183-190.
9. *Singleton G., Belmain S., Brown P. R., Aplin K., Htwe N. M.* Impacts of rodent outbreaks on food security in Asia // *Wildlife Research*. – 2010. – Vol. 37, No. 5. – P. 355-359.
10. *John A.* Rodent outbreaks and rice pre-harvest losses in Southeast Asia // *Food Security*. – 2014. – Vol. 6, No. 2. – P. 249-260.
11. *Chow J., Su Z., Wu J., Tan P. S., Mao X., Wang Y. H.* Anomaly detection of defects on concrete structures with the convolutional autoencoder // *Advanced Engineering Informatics*. – 2020. – Vol. 45. – P. 101105.
12. *Krestovnikov K.D., Erashov A.A.* Development of architecture and generalized structure of modules for a distributed control system for general-purpose robotic complexes // *Robotics and Cybernetics*. – 2022. – Vol. 10, No. 3. – P. 201-212. (In Russ.).
13. *Sun C., Bao Y., Bao Y., Vandansambuu B., Bayarsaikhan S., Gantumur B., Wu K.* Detection and classification of Brandt's vole burrow clusters utilizing GF-2 satellite imagery and faster R-CNN model // *Frontiers in Ecology and Evolution*. – 2024. – Vol. 12. – P. 1310046.
14. *Heydari M., Mohamadzamani D., Parashkouhi M. G., Ebrahimi E., Soheili A.* An Algorithm for Detecting the Location of Rodent-Made Holes through Aerial Filming by Drones // *Archives of Pharmacy Practice*. – 2020. – Vol. 1. – P. 55.
15. *Ezzy H., Charter M., Bonfante A., Brook A.* How the small object detection via machine learning and UAS-based remote-sensing imagery can support the achievement of SDG2: A case study of vole burrows // *Remote Sensing*. – 2021. – Vol. 13. – No. 16. – P. 3191.
16. *Li C., Luo X., Pan X.* CGT-YOLOv5n: A Precision Model for Detecting Mouse Holes Amid Complex Grassland Terrains // *Applied Sciences*. – 2023. – Vol. 14, No. 1. – P. 291.
17. *Du M., Wang D., Liu S., Lv C., Zhu Y.* Rodent hole detection in a typical steppe ecosystem using UAS and deep learning // *Frontiers in Plant Science*. – 2022. – Vol. 13. – P. 992789.
18. *Xie T., Luo X., Pan X.* BSM-YOLO: A Dynamic Sparse Attention-Based Approach for Mousehole Detection // *IEEE Access*. – 2024.
19. *Wan J., Jian D., Yu D.* Research on the method of grass mouse hole target detection based on deep learning // *Journal of Physics: Conference Series*. – IOP Publishing, 2021. – Vol. 1952. – No. 2. – P. 022061.
20. *Google Pictures*. – URL: <https://www.google.ru/imghp?hl=ru&ogbl> (дата обращения: 11.01.2025).
21. *Labelme. Image Polygonal Annotation with Python*. – URL: <https://github.com/wkentaro/labelme> (дата обращения: 11.01.2025).
22. *Ioffe S., Szegedy C.* Batch normalization: Accelerating deep network training by reducing internal covariate shift // *International Conference on Machine Learning*. – PMLR, 2015. – P. 448-456.
23. *Glorot X., Bordes A., Bengio Y.* Deep sparse rectifier neural networks // *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*. – *JMLR Workshop and Conference Proceedings*, 2011. – P. 315-323.
24. *PyTorch*. – URL: [pytorch.org](https://pytorch.org) (дата обращения: 11.01.2025).
25. *Bergmann P., Löwe S., Fauser M., Sattlegger D., Steger C.* Improving unsupervised defect segmentation by applying structural similarity to autoencoders // *arXiv preprint arXiv:1807.02011*. – 2018.
26. *Астапова М.А., Уздяев М.Ю., Кондратьев В.М.* Прогнозирование урожайности зеленых культур на основе мониторинга морфометрических параметров посредством машинного зрения и нейронных сетей // *Известия Кабардино-Балкарского научного центра РАН*. – 2024. – Т. 26, № 3. – С. 11-20.

## REFERENCES

1. Witmer G. Rodents in Agriculture: A Broad Perspective, *Agronomy*, 2022, Vol. 12, No. 6, pp. 1458.
2. Old J.M., Lin S.H., Franklin M.J. M. Mapping out bare-nosed wombat (*Vombatus ursinus*) burrows with the use of a drone, *BMC Ecology*, 2019, Vol. 19, No. 1, pp. 1-10.
3. Favorskaya M.N., Nishchkal N. Verifikatsiya razlivov nefti na vodnykh poverkhnostyakh po aerofotosnimkam na osnove metodov glubokogo obucheniya [Verification of oil spills on water surfaces based on deep learning methods using aerial photographs], *Informatika i Avtomatizatsiya* [Informatics and Automation], 2022, Vol. 21, No. 5, pp. 937-962.
4. Hao S., Zhou Y., Guo Y. A brief survey on semantic segmentation with deep learning, *Neurocomputing*, 2020, Vol. 406, pp. 302-321.
5. Hafiz A.M., Bhat G.M. A survey on instance segmentation: state of the art, *International Journal of Multimedia Information Retrieval*, 2020, Vol. 9, No. 3, pp. 171-189.
6. Kirillov A., He, K., Girshick R., Rother C., Dollár P. Panoptic segmentation, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 9404-9413.
7. Geddes A. M. W. The relative importance of pre-harvest crop pests in Indonesia. *Bulletin-Natural Resources Institute*, 1992, No. 47.
8. Leirs H. Management of rodents in crops: the Pied Piper and his orchestra. *ACIAR Monograph Series*, 2003, Vol. 96, pp. 183-190.
9. Singleton G., Belmain S., Brown P.R., Aplin K., Htwe N.M. Impacts of rodent outbreaks on food security in Asia. *Wildlife Research*, 2010, Vol. 37, No. 5, pp. 355-359.
10. John A. Rodent outbreaks and rice pre-harvest losses in Southeast Asia, *Food Security*, 2014, Vol. 6, No. 2, pp. 249-260.
11. Chow J., Su Z., Wu J., Tan P.S., Mao X., Wang Y.H. Anomaly detection of defects on concrete structures with the convolutional autoencoder, *Advanced Engineering Informatics*, 2020, Vol. 45, pp. 101105.
12. Krestovnikov K.D., Erashov A.A. Development of architecture and generalized structure of modules for a distributed control system for general-purpose robotic complexes, *Robotics and Cybernetics*, 2022, Vol. 10, No. 3, pp. 201-212. (In Russ.).
13. Sun C., Bao Y., Bao Y., Vandansambuu B., Bayarsaikhan S., Gantumur B., Wu K. Detection and classification of Brandt's vole burrow clusters utilizing GF-2 satellite imagery and faster R-CNN model, *Frontiers in Ecology and Evolution*, 2024, Vol. 12, pp. 1310046.
14. Heydari M., Mohamadzamani D., Parashkouhi M. G., Ebrahimi E., Soheili A. An Algorithm for Detecting the Location of Rodent-Made Holes through Aerial Filming by Drones, *Archives of Pharmacy Practice*, 2020, Vol. 1, pp. 55.
15. Ezy H., Charter M., Bonfante A., Brook A. How the small object detection via machine learning and UAS-based remote-sensing imagery can support the achievement of SDG2: A case study of vole burrows, *Remote Sensing*, 2021, Vol. 13, No. 16, pp. 3191.
16. Li C., Luo X., Pan X. CGT-YOLOv5n: A Precision Model for Detecting Mouse Holes Amid Complex Grassland Terrains, *Applied Sciences*, 2023, Vol. 14, No. 1, pp. 291.
17. Du M., Wang D., Liu S., Lv C., Zhu Y. Rodent hole detection in a typical steppe ecosystem using UAS and deep learning, *Frontiers in Plant Science*, 2022, Vol. 13, pp. 992789.
18. Xie T., Luo X., Pan X. BSM-YOLO: A Dynamic Sparse Attention-Based Approach for Mousehole Detection, *IEEE Access*, 2024.
19. Wan J., Jian D., Yu D. Research on the method of grass mouse hole target detection based on deep learning, *Journal of Physics: Conference Series*, IOP Publishing, 2021, Vol. 1952, No. 2, pp. 022061.
20. Google Pictures. Available at: <https://www.google.ru/imghp?hl=ru&ogbl> (accessed: 11 January 2025).
21. Labelme. Image Polygonal Annotation with Python. Available at: <https://github.com/wkentaro/labelme> (accessed: 11 January 2025).
22. Ioffe S., Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift, *International Conference on Machine Learning*, PMLR, 2015, pp. 448-456.
23. Glorot X., Bordes A., Bengio Y. Deep sparse rectifier neural networks, *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, JMLR Workshop and Conference Proceedings, 2011, pp. 315-323.
24. PyTorch. Available at: [pytorch.org](https://pytorch.org) (accessed: 11 January 2025).
25. Bergmann P., Löwe S., Fauser M., Sattlegger D., Steger C. Improving unsupervised defect segmentation by applying structural similarity to autoencoders, *arXiv preprint arXiv:1807.02011*, 2018.
26. Astapova M.A., Uzdyayev M.Yu., Kondrat'ev V.M. Prognozirovanie urozhaynosti zelenykh kul'tur na osnove monitoringa morfometricheskikh parametrov posredstvom mashinnogo zreniya i neyronnykh setey [Prediction the yield of green crops based on monitoring morphometric parameters using machine vision and neural networks], *Izvestiya Kabardino-Balkarskogo nauchnogo tsentra RAN* [News of the Kabardino-Balkarian Scientific Center of RAS], 2024, Vol. 26, No. 3, pp. 11-20.

**Астапова Марина Алексеевна** – Санкт-Петербургский Федеральный исследовательский центр Российской академии наук; e-mail: astapova.m@iias.spb.su; г. Санкт-Петербург, Россия; тел.: +79675329363; младший научный сотрудник лаборатории автономных робототехнических систем.

**Astapova Marina Alekseevna** – St. Petersburg Federal Research Center of the Russian Academy of Sciences; e-mail: astapova.m@iias.spb.su; St. Petersburg, Russia; phone: +79675329363; junior researcher, Laboratory of autonomous robotic systems.

УДК 621.385.21

DOI 10.18522/2311-3103-2025-2-30-40

**И.Л. Житяев, М.С. Картель, Ю.Ю. Житяева, А.А. Авакян, В.А. Смирнов**

**ИССЛЕДОВАНИЕ МЕМРИСТИВНЫХ НАНОРАЗМЕРНЫХ СТРУКТУР  
С ПРОФИЛИРОВАННЫМ ЭЛЕКТРОДОМ ДЛЯ НЕЙРОМОРФНОЙ  
ЭЛЕКТРОНИКИ**

*Представлены результаты разработки наноразмерных мемристорных структур, применение которых является перспективным для аппаратной реализации систем искусственного интеллекта. Предложена конструкция мемристорной ячейки на основе пленки оксида титана толщиной от 3 до 50 нм, верхний электрод которой представлял профилированную структуру в виде двух высокоаспектных острижных наноразмерных структур (ВОНС), у которой одно из острий было с радиусом 10 нм, а радиус острия второй варьировался в диапазоне от 10 до 50 нм. В качестве материала для верхнего электрода была выбрана платина из-за ее уникальных физико-химических свойств, в том числе высокой химической инертности в широком диапазоне температур и агрессивных сред, низкого удельного электрического сопротивления и устойчивости к окислению. Эти характеристики делают платину оптимальным материалом для использования в электронных устройствах и сенсорных системах, где требуется долговременная стабильность и минимальные потери энергии при передаче сигнала. Представлены результаты моделирования распределения напряженности электрического поля в межэлектродном промежутке мемристорной ячейки. Моделирование проводилось с использованием программного обеспечения COMSOL Multiphysics, в которой решаются системы нелинейных дифференциальных уравнений в частных производных методом конечных элементов, при разности потенциалов между электродами 5 В. На основе результатов моделирования получены и проанализированы зависимости величины напряженности электрического поля от геометрических параметров мемристорной ячейки. Выявлено локальное усиление напряженности электрического поля по периметру интерфейса оксида с ВОНС. Усиление неоднородности напряженности электрического поля возрастает с увеличением толщины оксидной пленки и может достигать 13,4%. Полученные результаты могут быть использованы при разработке нейроморфной электронной компонентной базы робототехнических систем и систем искусственного интеллекта на основе мемристоров.*

*Мемристор; нейроэлектроника; нейроморфные системы; резистивное переключение; наноразмерные оксидные пленки; напряженность электрического поля.*

**I.L. Jityaev, M.S. Kartel, Yu.Yu. Jityaeva, A.A. Avakyan, V.A. Smirnov**

**INVESTIGATION OF MEMRISTIVE NANOSCALE STRUCTURES WITH  
PROFILED INTERFACE FOR NEUROMORPHIC ELECTRONICS**

*The article presents the results of the development of nanoscale memristive structures, the application of which is promising for the hardware implementation of artificial intelligence systems. A design of a memristive cell based on a titanium oxide film with a thickness ranging from 3 to 50 nm is proposed. The upper electrode of the cell features a profiled structure in the form of two high-aspect-ratio nanoscale tip structures (HANTS), where one tip has a radius of 10 nm, and the radius of the second tip varies in the range of 10 to 50 nm. Platinum was chosen as the material for the upper electrode due to its unique physicochemical properties, including high chemical inertness across a wide range of temperatures and aggressive environments, low electrical resistivity, and resistance to oxidation. These characteristics make platinum an optimal material for use in electronic devices and sensor systems where long-term stability and minimal energy losses during signal transmission are required. The results of modeling the electric field strength distribution in the interelectrode gap of the memristive cell are presented. The modeling was*