

ИЗГОТОВЛЕНИЕ И ТЕСТИРОВАНИЕ АППАРАТНОЙ ИМПУЛЬСНОЙ НЕЙРОСЕТИ С МЕМРИСТОРНЫМИ СИНАПСАМИ ДЛЯ БИОМОРФНОГО НЕЙРОПРОЦЕССОРА

© 2021 г. А. Н. Бобылев¹, А. Н. Бусыгин¹, А. А. Губин¹, А. Д. Писарев¹, С. Ю. Удовиченко^{1,*}

¹ Тюменский государственный университет, Тюмень, Россия

*E-mail: udotgu@mail.ru

Поступила в редакцию 14.05.2021 г.

После доработки 11.06.2021 г.

Принята к публикации 11.06.2021 г.

Методами магнетронного распыления и электронной литографии изготовлен комбинированный мемристорно-диодный кроссбар, являющийся основой запоминающей матрицы биоморфного нейропроцессора. Проведено аппаратное тестирование работы запоминающей матрицы в режиме синапсов однослойного персептрона, который может рассматриваться в качестве первого слоя биоморфной нейросети и выполнять первичную обработку поступающей информации в биоморфном нейропроцессоре. Продемонстрирована генерация новой ассоциации при переобучении, вызванном поступлением новой входной информации. Показано влияние соседних нейронов на синаптический ток, возникающее вследствие паразитных токов между соседними ячейками в мемристорно-диодном кроссбаре. С помощью построенной аппаратной импульсной нейросети поступление новой неизвестной информации можно отождествлять с генерацией новых ассоциаций в биоморфном нейропроцессоре, при совершенствовании нейросети научиться осмысливать эту информацию и, следовательно, совершить переход от слабого искусственного интеллекта к сильному.

DOI: 10.1134/S1992722321060054

ВВЕДЕНИЕ

В импульсных нейросетях реализуется биологически подобный механизм ассоциативного самообучения, который сложно реализовать в традиционных сетях с точечными нейронами [1]. Импульсные нейронные сети превосходят нейросети на точечных нейронах в точности и вычислительной мощности, кроме того, они лучше приспособлены для аппаратной реализации из-за работы по принципу “интегрировать-и-сработать” (“integrate-and-fire”) [2]. Исследования по созданию аппаратной нейросети на основе мемристорных синапсов были начаты в [3], где продемонстрирована экспериментальная неимпульсная нейросеть из однослойного персептрона с интегрированным кроссбаром металл-оксидных мемристоров.

Принципы ассоциативного самообучения и формирование новой ассоциации (нового знания) в импульсной нейросети с мемристорными синапсами по правилу Хебба впервые представлены в [4]. Аппаратные реализации этой идеи в виде нейросети из трех нейронов с синапсами на дискретных мемристорах предложены в [5–10]. При экспериментальной демонстрации ассоциативного самообучения в нейросети с мемристора-

ми на основе органического полимера парилена явно показана генерация новых импульсов, связанных с возникновением новой ассоциации [9]. В [10] проводилось численное моделирование нейросети с мемристорными синапсами, характеристики которых были предварительно получены экспериментально [9].

Однако предложенные электрические цепи аппаратной реализации ассоциативной памяти не могут быть использованы для построения сверхбольшой аппаратной нейросети с высокой интеграцией элементов и энергоэффективностью, что связано с отсутствием интеграции мемристоров в кроссбары, а также наличием в схемах нейронов и синапсов большого количества активных электронных элементов с высоким энергопотреблением.

Аппаратные неимпульсные нейронные сети на основе мемристорного кроссбара с мемристорно-транзисторными ячейками 1T1M реализованы исследовательской группой Hewlett-Packard в виде персептрона [11] и сети Хопфилда [12]. Размер селективного полевого транзистора в использованных ячейках на порядок превышает минимальный размер мемристора. Увеличение степени интеграции возможно при замене транзистора

на диод Зенера. Для сверхбольших кроссбаров диод Зенера более предпочтителен, поскольку имеет размер, соизмеримый с размером мемристора.

В [13] предложена разработка биоморфного нейропроцессора на основе нового компонента наноэлектроники – мемристорно-диодного кроссбара с биполярными мемристорами, подходящего для традиционных задач обработки информации, а также для воспроизведения работы кортикальной колонки мозга или ее фрагмента. Нейропроцессор реализует импульсную аппаратную нейросеть с помощью оригинальных электрической и программной биоморфных моделей нейрона. Сформулированы принципы построения биоморфной нейросети [14].

Создание нейропроцессоров на основе мемристоров и мемристорных кроссбаров затруднено, так как разрабатываемые твердотельные мемристоры на оксидах переходных металлов характеризуются невысокой стабильностью и воспроизводимостью электрических характеристик. Однако предложенный биоморфный нейропроцессор лишен перечисленных выше недостатков. Распределенный характер биоморфной нейросети снижает требования к воспроизводимости и стабильности характеристик мемристоров, а также повышает отказоустойчивость схемы нейропроцессора. Кроме этого, в аппаратную реализацию нейросети можно добавить электрическую схему, которая будет воспроизводить работу астроцита (одного из видов глиальных клеток мозга) путем увеличения проводимости оставшихся мемристорных синапсов при обнаружении поврежденного, что дополнительно повысит отказоустойчивость схемы нейропроцессора и увеличит биоморфность нейросети [15].

В [16] представлена магнетронная технология изготовления полупроводниковых слоев диода Зенера и мемристорного слоя, обеспечивающая требуемые характеристики диода и мемристоров. Проведены исследования электрофизических свойств изготовленного мемристорно-диодного кроссбара с числом ячеек $1D2M\ 2 \times 2$. Выполнена обработка сигналов с помощью умножения матрицы напряжений на вектор проводимостей ячеек и ассоциативного самообучения. Приведено сравнение экспериментальных результатов и SPICE-моделирования процессов обработки сигналов в изготовленном мемристорно-диодном кроссбаре и определен разброс параметров мемристоров, при котором наблюдается устойчивая работа кроссбара. Среднеквадратичное отклонение сопротивления ячеек может достигать от 60 до 100% и более.

Впервые продемонстрирована генерация новой ассоциации (нового знания) в нейросети из двух виртуальных и одного аппаратного нейрона

с синапсами в виде изготовленного мемристорно-диодного кроссбара в отличие от ассоциативного самообучения в существующих аппаратных нейросетях с синапсами на базе дискретных мемристоров [16].

В настоящем сообщении приведены результаты тестирования работы аппаратной импульсной нейросети с мемристорными синапсами, представляющей собой запоминающую матрицу в режиме расчета синапсов однослойного перцептрона. Перцептрон может рассматриваться в качестве первого слоя 3D-нейросети процессора, выполняющего первичную обработку поступающей информации [14].

ОБЪЕКТЫ И МЕТОДЫ ИССЛЕДОВАНИЯ

С помощью магнетронной технологии [16] изготовлен мемристорно-диодный кроссбар запоминающей матрицы с числом ячеек 4×2 , являющийся интегральным массивом синапсов аппаратного перцептрона. Ячейки матрицы $1D2M$ представляют собой двухслойное соединение комплементарных биполярных мемристоров и одного разделяющего диода Зенера. Изготовление мемристорно-диодного кроссбара $TiN/Ti_{0.93}Al_{0.07}O_x/p-Si/n-Si/W$ (рис. 1а) проводили в магнетронном модуле, входящем в нанотехнологический комплекс “НаноФаб-100” (NT-MDT, Россия). Слой смешанного оксида металлов осаждали при одновременном распылении двух катодов из титана и алюминия, а полупроводниковые слои диода – при последовательном распылении катодов из легированного бором и фосфором кремния ($p-Si$ и $n-Si$ соответственно). Мемристоры изготавливали по технологии кроссбар путем последовательного напыления пленок через маски электронного резиста (полиметилметакрилат), выполненные на электронном микроскопе JSM-6510LV-EDS (Jeol, Япония). Вольт-амперная характеристика (ВАХ) ячейки кроссбара приведена на рис. 1б.

Максимальное отношение сопротивлений ячейки кроссбара в низко- и высокопроводящем состояниях достигается при напряжении $U = 0.3$ В и составляет 54.2. Среднее сопротивление ячеек в низкопроводящем состоянии равно 1.56 МОм со среднеквадратичным отклонением 0.43 МОм при $U = 2$ В. Характеристики изготовленных мемристоров в составе мемристорно-диодного кроссбара исследовали ранее [15, 17]. Удельные сопротивления предельных состояний при циклических переключениях ρ_{off} и ρ_{on} составляют 1.14 ± 0.18 кОм м и 1.61 ± 0.05 Ом м соответственно. Напряжение переключения ячейки оказалось больше, чем у отдельного мемристора [16], поскольку часть напряжения падает на диоде. Большое сопротивление закрытого диода приво-

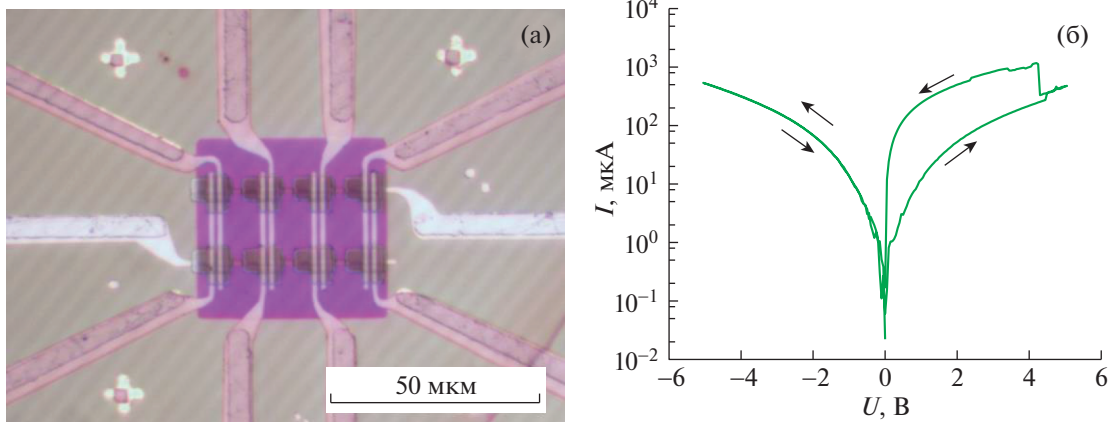


Рис. 1. Микрофотография мемристорно-диодного кроссбара $\text{TiN}/\text{Ti}_{0.93}\text{Al}_{0.07}\text{O}_x/p\text{-Si}/n\text{-Si}/\text{W}$ (а) и зависимость тока ячейки I от приложенного напряжения U ячейки изготовленного кроссбара (б).

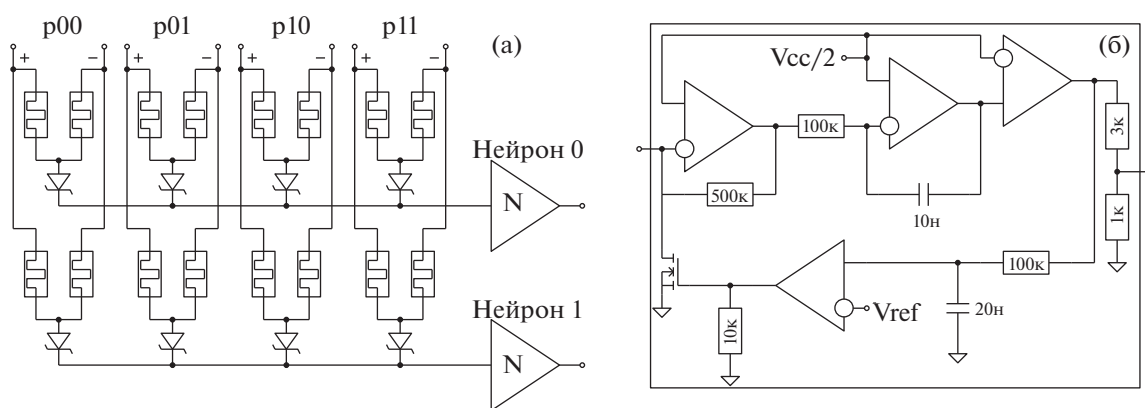


Рис. 2. Электрическая схема аппаратного персептрона: а – мемристорно-диодный кроссбар в качестве массива синапсов, б – электрическая схема нейрона.

дит к стягиванию гистерезиса в обратной ветви ВАХ ячейки, так как вклад сопротивления диода преобладает над вкладом малого сопротивления мемристора в их общую ВАХ. Снижение напряжения переключения ячейки возможно при уменьшении толщины активного слоя мемристора и увеличении нелинейности селективного элемента. Распределенный характер нейросети снижает требования к воспроизводимости и стабильности характеристик мемристоров, и на первый план выходит отношение предельных значений сопротивлений для существования множества синаптических состояний. Это подтверждается в [16], где работоспособность кроссбара сохранялась при разбросе предельных сопротивлений ячеек до 100%.

Аппаратный персептрон (рис. 2) построен на основе мемристорно-диодного кроссбара с четырьмя парами входных проводников и двумя выходными шинами. Соответственно, кроссбар

содержит восемь ячеек, которые являются синапсами нейросети. Слой персептрона образован двумя нейронами, построенными на основе операционных усилителей. Электрическая схема нейрона состоит из преобразователя ток–напряжение, аналогового интегратора, компаратора, схемы задержки в виде интегрирующей RC-цепи и полевого транзистора.

Реализованный в аппаратной нейросети механизм самообучения пластичности, зависимой от времени импульса (spike-timing dependent plasticity (STDP)), учитывает причинность. Если пресинаптический импульс пришел раньше и перекрывается во времени с постсинаптическим импульсом, то он может быть причиной активации постсинаптического нейрона, и вес синапса возрастает. Если постсинаптический импульс возник раньше пресинаптического импульса, то он не может быть причиной активации постсинаптического нейрона, и вес синапса уменьшается.

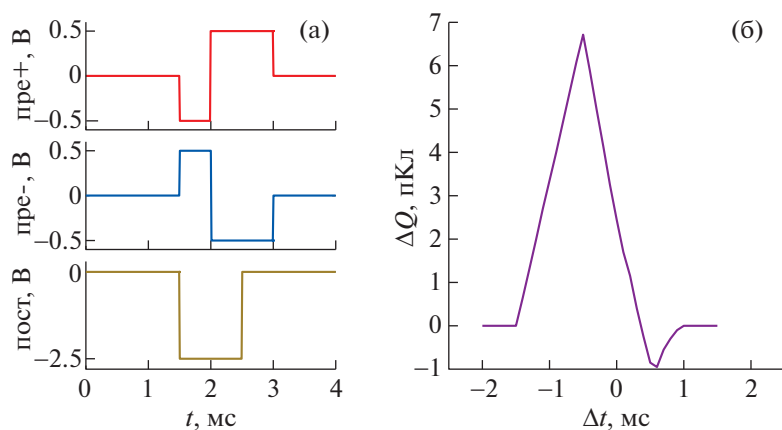


Рис. 3. Пластичность, зависящая от времени импульса: а – пре- и постсинаптические напряжения ячейки запоминающей матрицы при $\Delta t = 0$, б – функция пластичности $\Delta Q(\Delta t)$ – изменение заряда конденсатора в интеграторе нейрона при прохождении одного информационного импульса после изменения массы синапса.

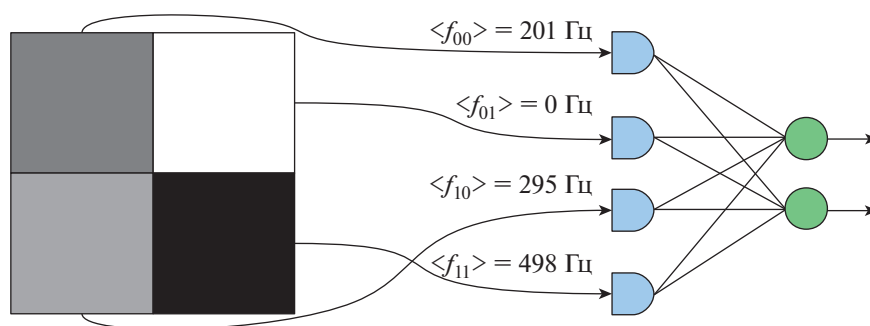


Рис. 4. Входное изображение и архитектура аппаратной нейросети: скругленные справа квадраты – входные виртуальные нейроны, круги – выходные нейроны.

Специальная форма входных импульсов позволяет реализовать ассоциативное обучение нейросети по механизму STDP. Изменение веса синапса зависит от разницы между временем срабатывания пресинаптического и постсинаптического нейронов $\Delta t = t_{\text{пре}} - t_{\text{пост}}$.

Форма пре- и постсинаптических импульсов для обеспечения механизма STDP и соответствующая функция пластичности представлены на рис. 3. Величина ΔQ показывает изменение заряда конденсатора в интеграторе нейрона при прохождении одного информационного импульса после изменения веса синапса.

Нейросеть, состоящая из четырех виртуальных входных нейронов и двух аппаратных выходных нейронов (рис. 4), обучалась для выполнения задачи распознавания набора входных изображений с разрешением 2×2 пикселя.

Значения яркостей пикселей преобразовывались в среднюю частоту последовательности входных импульсов с помощью микроконтроллера. Выходные импульсы активировавшихся нейронов фиксировались этим же микроконтролле-

ром. Результат распознавания изображения определял выходной нейрон сети с большим числом выходных импульсов, зарегистрированных во время присутствия последовательности входных импульсов.

РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

Кривая обучения (рис. 5а) показывает долю правильно классифицированных входных изображений от числа циклов (эпох) обучения. Каждый цикл обучения состоит из последовательной подачи на вход нейросети импульсов, соответствующих 128 изображениям, образованных путем добавления шума к исходным двум эталонным картинкам. В каждой эпохе набор изображений одинаковый.

Расположение прямоугольников на рис. 5б соответствует положению мемристоров на рис. 2а. Изначально проводимость мемристоров была низкой, за исключением одного из них. В процессе ассоциативного самообучения проводимость

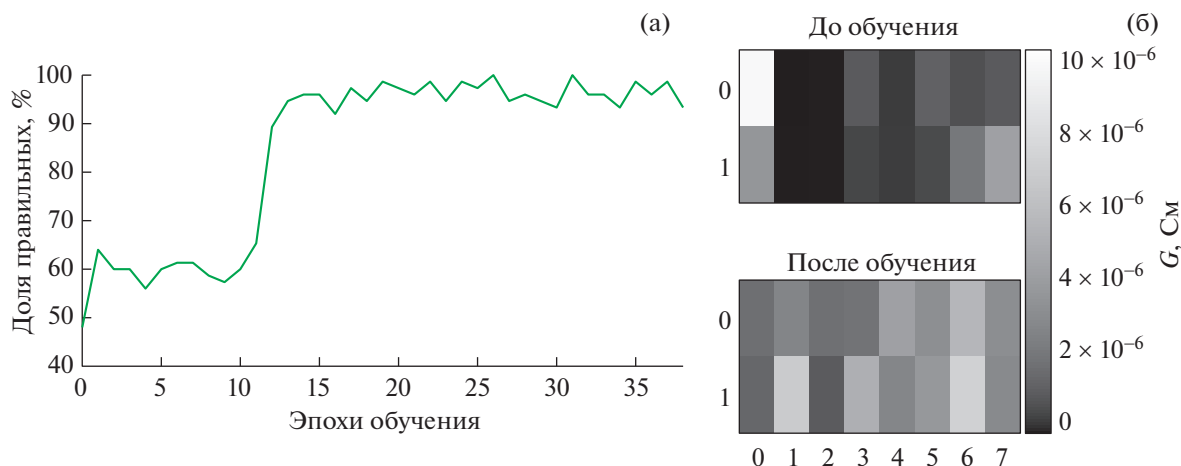


Рис. 5. Результат ассоциативного самообучения аппаратной нейросети: а – изменение доли правильных классификаций, б – изменение проводимости мемристоров кроссбара.

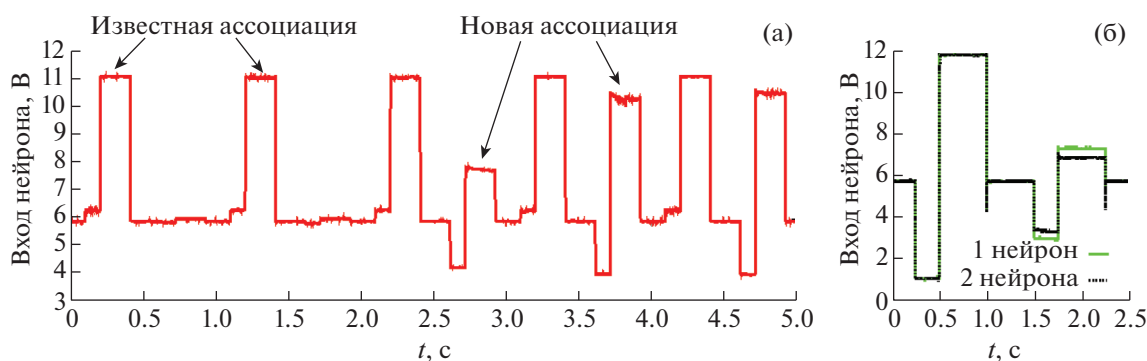


Рис. 6. Генерация новой ассоциации на фоне известной: а – появление и нарастание сигнала от изначально слабого синапса, б – изменение амплитуды импульсов при подключении второго нейрона в схему.

этого мемристора уменьшилась, а остальных выросла.

Исследуемая аппаратная импульсная нейросеть на основе мемристорно-диодного кроссбара отличается от нейросети большим числом нейронов и синапсов [16]. Кроме этого, при последующем переобучении нейросеть опирается на ассоциации, сформированные в процессе обработки входных сигналов, а не заложенные искусственно, как в [16].

Процесс генерации новой ассоциации при переобучении, связанном с поступлением новой информации, показан на рис. 6а.

Для демонстрации процесса быстрого формирования новой ассоциации были выбраны импульсы большой длительности, чтобы переобучение произошло за небольшое количество импульсов. Этот же процесс описан в [9], где представлены начальный и конечный результаты формирования новой ассоциации после большого числа импульсов.

Наращение входного напряжения нейрона, пропорциональное синаптическому току, вызвано усилением синапса при формировании новой ассоциации. Генерация новой ассоциации при переобучении происходит в изготовленном мемристорно-диодном кроссбаре в отличие от существующих нейросетей с синапсами на базе дискретных мемристоров [5–10].

Подключение второго нейрона вызывает уменьшение синаптического тока первого нейрона на 9.5% (рис. 6б) из-за существования небольших паразитных токов через соседние ячейки в кроссбаре. Часть суммарного тока синапсов первого нейрона втекает во второй нейрон, поэтому наблюдаемая амплитуда импульсов уменьшается. Численное моделирование работы запоминающей матрицы, образованной с помощью идеальных моделей мемристоров и диодов, показало, что деградация выходного напряжения из-за паразитных токов составляет 70% при числе ячеек 10^6 [18]. Это достигается за счет высокой нели-

нейности ВАХ в модели диода Зенера, которая проявляется в отношении его сопротивлений в закрытом и открытом состояниях, равном 10^9 . Решить проблему паразитных токов в изготовленной матрице большого размера можно путем оптимизации параметров ячейки в первую очередь выбором материалов $p-n$ -перехода и понижением напряжения пробоя диода, а также выбором материалов проводников мемристора с целью повышения барьера Шоттки.

ЗАКЛЮЧЕНИЕ

Изготовлена и протестирована импульсная нейросеть с мемристорными синапсами на основе мемристорно-диодного кроссбара и аппаратных нейронов, представляющая собой однослойный персептрон. Персептрон может рассматриваться в качестве первого слоя 3D-нейросети [14], выполняющего первичную обработку поступающей информации в биоморфном нейропроцессоре [13]. Получена экспериментальная кривая обучения, показывающая ожидаемое увеличение доли правильных классификаций с ростом числа эпох обучения. Продемонстрирована генерация новой ассоциации при переобучении, вызванном поступлением новой входной информации.

Показано влияние соседних нейронов на синоптический ток, возникающее вследствие паразитных токов между соседними ячейками в мемристорно-диодном кроссбаре, которое необходимо учитывать при разработке аппаратных нейросетей с большими кроссбарами. Решить проблему паразитных токов в матрицах большого размера можно путем подбора материалов ячейки с целью увеличения нелинейности ее вольт-амперной характеристики.

Таким образом, с помощью построенной аппаратной импульсной нейросети поступление новой неизвестной информации можно отождествлять с генерацией новых ассоциаций в биоморфном нейропроцессоре, при совершенствовании нейросети [14] научиться осмысливать эту информацию и, следовательно, совершить переход от слабого к сильному искусственному интеллекту.

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (гранты № 19-07-00272 “Электрофизические свойства комбинированного мемристорного-диодного кроссбара — нового компонента нанoeлектроники, предназначенного для изготовления запоминающей и логической матриц нейропроцессора” и № 19-37-90030 “Генерация нового знания в нейросети на основе массива мемристорных синапсов в запоминающей матрице биоморфного нейропроцессора и принципы увеличения быстродействия и энергоэффективности

обработки информации на специализированном устройстве по сравнению с существующими вычислительными средствами”).

СПИСОК ЛИТЕРАТУРЫ

1. *Khacef L., Abderrahmane N., Miramond B.* // Int. Jt. Conf. Neural Netw. (IJCNN 2018). 2018. P. 1. <https://doi.org/10.1109/ijcnn.2018.8489241>
2. *Lobo J.L., Ser J.D., Bifet A., Kasabov N.* // Neural Netw. 2020. V. 121. P. 88. <https://doi.org/10.1016/j.neunet.2019.09.004>
3. *Prezioso M., Merrikkh-Bayat F., Strukov D.B. et al.* // Nature. 2015. V. 521. P. 61. <https://doi.org/10.1038/nature14441>
4. *Pershin Y.V., Di Ventra M.* // Neural Netw. 2010. V. 23. № 7. P. 881. <https://doi.org/10.1016/j.neunet.2010.05.001>
5. *Wang Z., Wang X., Novel A.* // IEEE Trans. Circuits Syst. I: Regul. Pap. 2018. V. 65. № 7. P. 2210. <https://doi.org/10.1109/TCSI.2017.2780826>
6. *Yang L., Zeng Z., Huang Y., Wen S.* // IEEE Trans. Cogn. Dev. Syst. 2018. V. 10. № 4. P. 1133. <https://doi.org/10.1109/TCDS.2018.2859303>
7. *Wang Z., Rao M., Han J.-W. et al.* // Nat. Commun. 2018. V. 9. P. 3208. <https://doi.org/10.1038/s41467-018-05677-5>
8. *Zhang X., Long K.* // IEEE Access. 2019. V. 7. P. 15262. <https://doi.org/10.1109/ACCESS.2019.2894634>
9. *Minnekhanov A.A., Emelyanov A.V., Lapkin D.A. et al.* // Sci. Rep. 2019. V. 9. P. 10800. <https://doi.org/10.1038/s41598-019-47263-9>
10. *Demin V.A., Nekhaev D.V., Surazhevsky I.A. et al.* // Neural Netw. 2021. V. 134. P. 64. <https://doi.org/10.1016/j.neunet.2020.11.005>
11. *Hu M., Graves C.E., Li C. et al.* // Adv. Mater. 2018. V. 30. № 9. P. 1705914. <https://doi.org/10.1002/adma.201705914>
12. *Cai F., Kumar S., Vaerenbergh T.V.* // Nat. Electron. 2020. V. 3. P. 409. <https://doi.org/10.1038/s41928-020-0436-6>
13. *Pisarev A.D., Busygin A.N., Udovichenko S.Yu., Maevsky O.V.* // Microelectron. J. 2020. V. 102. P. 104827. <https://doi.org/10.1016/j.mejo.2020.104827>
14. *Filippov V.A., Bobylev A.N., Busygin A.N. et al.* // Neural Comput. Appl. 2020. V. 32. P. 2471. <https://doi.org/10.1007/s00521-019-04383-7>
15. *Liu J., Mcdaid L.J., Harkin J. et al.* // IEEE Trans. Neural Netw. Learn. Syst. 2018. V. 30. № 3. P. 865. <https://doi.org/10.1109/TNNLS.2018.2854291>
16. *Pisarev A., Busygin A., Bobylev A. et al.* // Microelectron. Eng. 2021. V. 236. P. 111471. <https://doi.org/10.1016/j.mee.2020.111471>
17. *Bobylev A.N., Udovichenko S.Yu.* // Russ. Microelectron. 2016. V. 45. № 6. P. 396. <https://doi.org/10.1134/S1063739716060020>
18. *Pisarev A.D., Busygin A.N., Udovichenko S.Yu., Maevsky O.V.* // Microelectron. Eng. 2018. V. 198. P. 1. <https://doi.org/10.1016/j.mee.2018.06.008>