

**А.М. Буряков
Е.И. Жемеров
Н.А. Ильин**

2024

**От биологических нейронов
к нейроморфным чипам:
введение в новые горизонты
искусственного интеллекта**

Монография



УДК 53.06
ББК 32.813
Б 919

Рецензент: Абдуллаев Даниил Анатольевич, к.т.н., научный сотрудник ИНМЭ РАН.

Буряков, Арсений Михайлович
Жемеров, Евгений Игоревич
Ильин, Никита Александрович

Б 919 От биологических нейронов к нейроморфным чипам: введение в новые горизонты искусственного интеллекта. Монография. – М.: Мир науки, 2024. – Сетевое издание. Режим доступа: <https://izd-mn.com/PDF/53MNNPM24.pdf> – Загл. с экрана.

ISBN 978-5-907891-33-3
DOI: 10.15862/53MNNPM24

В данной научно-популярной книге авторы приглашают читателя в увлекательное путешествие по миру нейронных сетей, нейроморфных чипов и других прорывных технологий, вдохновлённых человеческим мозгом. Через доступное изложение научных исследований и примеры новейших технических разработок, читатель узнает о том, как биологические нейроны послужили основой для создания энергоэффективных вычислительных систем, способных адаптироваться и обучаться, подобно живым организмам. Эта книга станет проводником в будущее искусственного интеллекта, покажет перспективы медицинских приложений, интерфейсов «мозг–компьютер» и способов решения сложнейших задач за счёт инновационных подходов, где наука, биология и инженерия сливаются воедино. Научно-популярная книга подготовлена при поддержке Министерства науки и высшего образования Российской Федерации в рамках гранта №075-15-2022-1131.

Научно-популярная книга издается в авторской редакции.

Авторский коллектив: Буряков Арсений Михайлович, Жемеров Евгений Игоревич, Ильин Никита Александрович

ISBN 978-5-907891-33-3

© Буряков Арсений Михайлович
© Жемеров Евгений Игоревич
© Ильин Никита Александрович
© ООО Издательство «Мир науки», 2024

Оглавление

Введение	4
Миниатюрные нейроинтерфейсы: современные разработки и перспективы	7
Проблемы существующих нейроморфных архитектур	9
Новаторские подходы к созданию надежных синапсов.....	10
Примеры существующих нейрочипов и их направления развития	11
Структура и функционирование искусственных нейронов.....	12
Архитектура искусственных нейронных сетей.....	14
Методы обучения нейронных сетей.....	14
Переход к нейроморфным чипам	16
Архитектура нейрочипов в сравнении с традиционными процессорами	16
Методы обработки данных и энергоэффективность	17
Сферы применения нейрочипов	17
Использование электронной микроскопии в разработке нейроморфных чипов.....	18
Заключение.....	28
Список литературы.....	30

Введение

Мозг человека — один из самых загадочных и удивительных органов, обладающий невероятной способностью выполнять разнообразные функции. Он способен не только хранить и воспроизводить тысячи нот, звуков и ритмов, например, «К Элизе» Людвига ван Бетховена, но и координировать движение десяти пальцев и двух стоп, чтобы исполнить это произведение. Мозг мгновенно узнает знакомые лица — родителей, братьев, сестёр, любимого кота по имени Мурзик или даже известных личностей, таких как Бейонсе. Он способен осмысливать как гигантские масштабы галактик, так и крошечные молекулы РНК, служит источником изобретательности, благодаря которой человечество создало всё — от электромобилей до передовых поисковых систем. Такая многофункциональность мозга во многом объясняется его невероятной сложностью. В человеческом мозге насчитывается свыше ста миллиардов нейронов, каждый из которых образует в среднем около тысячи синаптических связей с другими клетками. Это приводит к созданию порядка ста триллионов соединений — поистине впечатляющая «электропроводка», обеспечивающая как стабильность, так и гибкость нейронных сетей. Мозг обладает удивительной способностью к самоорганизации: новые связи формируются, изменяются и исчезают в ответ на внешние стимулы и накопленный опыт, что лежит в основе обучения и памяти. Эта способность к изменению и адаптации особенно интересна нейробиологам, стремящимся понять механизмы формирования синапсов, их сохранения или удаления, а также эволюции целых нейронных сетей со временем.

Для изучения столь сложных процессов исследователи часто обращаются к более простым биологическим системам: от нервных узлов пиявок и плодовых мушек до культур нейронов позвоночных, выращенных в лабораторных условиях. Такие клеточные культуры, содержащие десятки тысяч нейронов, значительно упрощают экспериментальные задачи по сравнению с человеческим

мозгом. Для изучения активности нейронов в таких системах применяются современные методы, включая электрофизиологические подходы, которые позволяют фиксировать электрические сигналы клеток. Среди них особое значение имеют патч-кламп, используемый для детального изучения активности отдельных клеток, и мультиэлектродные матрицы (МЭА), позволяющие регистрировать сигналы от больших сетей нейронов. Несмотря на свою эффективность, каждый из этих методов имеет ограничения: патч-кламп охватывает лишь небольшое число клеток и ограничивает длительность экспериментов, тогда как МЭА, обеспечивая длительное наблюдение за активностью, затрудняет точное определение взаимосвязи между электродами и конкретными клетками.

Чтобы преодолеть барьеры, возникающие при изучении нейронных систем, учёные разрабатывают новые инструменты, сочетающие передовые достижения микроэлектроники и нейробиологии. Одним из таких прорывных решений стал «нейрочип» — инновационная технология, способная одновременно стимулировать и регистрировать активность множества нейронов в культуре на протяжении всего периода их развития. Подобные достижения, находящиеся на пересечении биологии и технологий, становятся частью более широкого контекста эволюции искусственного интеллекта (ИИ), концепция которого была предложена ещё в 1956 году.

Идея ИИ заключается в имитации человеческого интеллекта с помощью технологий. Сегодня искусственный интеллект достиг стадии, на которой он предлагает эффективные решения для сложных задач в таких ключевых областях, как наука, финансы, безопасность и транспорт. Использование «нейрочипов» и других подобных технологий может стать ключевым элементом в создании более совершенных моделей ИИ, вдохновлённых нейронными системами [1].

Одним из наиболее перспективных направлений в области ИИ является создание нейроморфных чипов, которые имитируют работу человеческого

мозга. Мозг состоит из нейронов и синапсов и при этом его энергозатраты составляют всего 20 ватт. Это стало вдохновением для создания чипов, которые могут эффективно решать задачи ИИ с высокой вычислительной мощностью и низким потреблением энергии [2]. В отличие от традиционных чипов, таких как процессоры и графические карты, нейроморфные чипы, имитируя нейронные сети мозга, позволяют значительно ускорить обработку данных и снизить энергозатраты, что является важным шагом для создания эффективных ИИ-систем [3]. Также, нейроморфные чипы могут быть использованы в таких областях, как обработка данных, робототехника, искусственный интеллект и системы с интерфейсами "мозг-компьютер". Они играют важную роль в решении таких сложных задач, как распознавание образов, управление движением, принятие решений и анализ больших данных.

Ключевым элементом в нейроморфных чипах являются импульсные нейронные сети (SNN), использующие импульсы (спайки) для передачи информации. Такой подход обеспечивает высокую энергоэффективность и скорость работы, что делает эти чипы перспективными для создания мощных ИИ-систем.

Нейронные сети уже доказали свою эффективность в решении ряда задач, таких как компьютерное зрение [4], распознавание речи [5], интерфейсы «мозг-компьютер» [6], визуализация данных [7]. Однако, несмотря на впечатляющие достижения в этих областях, текущие нейроморфные архитектуры остаются ограниченными: они реализованы в программном виде, а аппаратное обеспечение часто не обладает пластичностью, то есть способностью к обучению, и требует значительных энергозатрат.

Миниатюрные нейроинтерфейсы: современные разработки и перспективы

Недавно швейцарские исследователи из Федеральной политехнической школы Лозанны (EPFL) представили миниатюрный чип, способный преобразовывать нейронную активность в текст с точностью свыше 90%. В отличие от более крупных и энергоёмких аналогов, таких как имплант Neuralink, этот нейрочип минимально инвазивен и чрезвычайно компактен, что делает его особенно перспективным для широкого спектра применений.

Чип, названный MiBMI (miniature brain-computer interface), состоит всего из двух тонких элементов размером 8 мм² и потребляет минимальное количество энергии. Он использует уникальный метод идентификации специфических «нейромаркеров» для каждой буквы, что позволяет резко сократить объём обрабатываемой информации с тысяч байт до сотен байт для каждого символа. Такой подход не только уменьшает размер устройства и снижает энергопотребление, но и ускоряет процесс обучения пользователей. В настоящее время MiBMI способен распознавать 31 символ, и учёные планируют увеличить этот набор до 100 символов.

Кроме того, NeuroChip, обновлённая версия первоначальной платформы NeuroX, представляет собой инструмент, специально разработанный для исследования генетических аспектов нейродегенеративных заболеваний. Эта платформа сочетает высокую производительность генотипизации с возможностью выявления редких генетических вариантов, ассоциированных с такими заболеваниями, как болезнь Альцгеймера, Паркинсона, боковой амиотрофический склероз и другие. Генетический анализ с использованием NeuroChip позволяет исследователям не только подтверждать клинические диагнозы, выявляя причинные мутации, но и определять генетические факторы, которые могут повышать риск развития заболевания или влиять на его течение. Благодаря высокой пропускной способности и доступной стоимости около 40

долларов за образец, NeuroChip становится важным инструментом для массового скрининга, ускоряя отбор образцов для более глубоких геномных исследований и помогая учёным лучше понять механизмы этих сложных заболеваний.

Объединяя возможности MiBMI и NeuroChip, современные нейротехнологии открывают новые горизонты в понимании и лечении нейродегенеративных заболеваний. Например, разработанный EPFL нейрочип может не только преобразовывать мысли в текст, но и служить основой для интерфейсов, помогающих людям с двигательными и речевыми нарушениями общаться с окружающим миром. В то же время, NeuroChip предоставляет мощный инструмент для изучения генетических основ нейродегенерации, способствуя поиску новых терапевтических мишеней и улучшению стратегий клинических испытаний.

Таким образом, дальнейшие успехи в разработке нейрочипов не только помогают учёным глубже понять принципы функционирования нейронных сетей и генетические механизмы нейродегенеративных заболеваний, но и значительно расширяют возможности человека, восстанавливая утраченные навыки общения и взаимодействия с окружающим миром. Эти достижения открывают перспективы для создания универсальных мозг-компьютерных интерфейсов, способных адаптироваться к различным неврологическим состояниям и расстройствам, что обещает революцию в медицине и нейронауках.

Проблемы существующих нейроморфных архитектур

Термин "нейроморфные" относится к вычислительным системам и архитектурам, которые имитируют структуру и принципы работы биологических нейронных сетей. Эти системы разрабатываются для эффективной обработки информации, используя аналоговые или цифровые элементы, напоминающие нейроны и синапсы, с целью повысить производительность и энергоэффективность по сравнению с традиционными вычислительными подходами. Современные нейроморфные архитектуры, вдохновлённые принципами работы биологических нейронных сетей, представляют собой перспективное направление в области вычислительных технологий. Они обещают значительно повысить эффективность и производительность в задачах обработки информации, особенно в системах искусственного интеллекта. Важнейшими элементами таких систем являются синапсы, которые хранят веса связей в нейронной сети и имеют возможность адаптировать этот вес в процессе обучения. В качестве синапсоподобных устройств наиболее часто предлагаются мемристоры [8]. Однако современные мемристорные материалы в основном подвержены влиянию ионных смещений, приводящих к невоспроизводимости и эффектам старения. Таким образом ещё одной важной веткой исследования является разработка методов и механизмов для создания надёжных синапсов.

Новаторские подходы к созданию надежных синапсов

Одной из ключевых задач в области нейроморфных вычислений является разработка методов и механизмов для создания надежных и эффективных синапсов. Прорывным решением в этом направлении стала демонстрация оптически управляемого синаптического поведения в магнитной пленке [9]. Эта технология устраняет проблемы невоспроизводимости и старения материалов, объединяя высокую скорость и эффективность оптических процессов с энергонезависимостью и стабильностью магнитных свойств. Эти достижения открывают новые возможности для разработки нейроморфных чипов — устройств, вдохновленных структурой и принципами работы биологической нервной системы. Такие чипы имитируют работу мозга, что позволяет эффективно решать задачи, требующие параллельной обработки информации, адаптации и обучения.

Примеры существующих нейрочипов и их направления развития

Ключевым преимуществом нейроморфных чипов является их способность к обучению, что делает их аналогичными человеческому мозгу, который адаптируется и изменяет свои реакции на основе накопленного опыта. Это качество позволяет нейроморфным устройствам эффективно решать задачи, требующие динамической адаптации и самообучения, таких как обработка сложных данных, распознавание образов и управление в реальном времени.

Однако, несмотря на значительные успехи, современные нейроморфные чипы сталкиваются с рядом технических ограничений. Среди них — высокие энергозатраты, которые препятствуют широкому применению этих устройств в масштабных системах, а также недостаточная гибкость, связанная с ограниченной возможностью адаптации существующих решений к изменяющимся условиям. Эти проблемы существенно сдерживают потенциал нейроморфных технологий, особенно в задачах, где требуется высокая производительность при низком энергопотреблении.

Для преодоления этих ограничений активно исследуются и разрабатываются новые подходы, основанные на создании материалов с уникальными функциональными свойствами. Такие материалы должны не только имитировать работу нейронов и синапсов, но и обладать способностью к обучению, адаптации и долговременной стабильности. Например, использование ферроидных сред или магнитных пленок, демонстрирующих синаптическое поведение, может позволить создать более эффективные и долговечные нейроморфные устройства.

Intel создала крупнейшую в мире нейроморфную систему — Hala Point, которая использует 1,15 миллиарда нейронов и 1152 процессора Loihi 2. Эта система разработана для более эффективного и устойчивого ИИ, помогая решать проблемы в области науки и инженерии с помощью нейроморфных вычислений, которые работают по принципу человеческого мозга, минимизируя

энергозатраты. Hala Point может выполнять до 20 петаопераций в секунду и поддерживает глубокие нейронные сети с эффективностью до 15 TOPS/W, что намного выше, чем у традиционных процессоров (ЦПУ и ГПУ). Система может быть использована для исследований в таких областях, как научное моделирование, управление инфраструктурой и даже обучение ИИ в реальном времени. Это первый шаг к созданию более энергоэффективных и масштабируемых ИИ-систем, которые могут работать в реальном времени, решая задачи, такие как обработка видео или речи.

Нейронные сети, работающие на основе нанофотоники, представляют собой еще одно перспективное направление. Создание умных машин, имитирующих биологические нейронные сети (BNNs), стало целью ученых сразу после появления современных компьютеров. Большинство исследований нейронных сетей (ANNs) сегодня основано на программных моделях, которые используют традиционные компьютеры с архитектурой фон Неймана. Идея имитации BNN с помощью электронных или фотонных устройств, также называемая нейроморфными вычислениями, появилась в конце 1980-х годов. В отличие от нейронных сетей, основанных непосредственно на биологических нейронах, такие искусственные сети используют электронные или фотонные системы с искусственными нейронами, чтобы косвенно имитировать нейробиологические структуры, встречающиеся в BNN.

Структура и функционирование искусственных нейронов

Биологическими аналогами искусственных нейронов являются реальные нейроны, которые являются строительными блоками биологических нейронных сетей (BNNs). В BNN содержится миллиарды нейронов различных типов и размеров. На рисунке 1 изображена схема упрощенного биологического нейрона, который состоит из четырех основных частей: дендритов, аксона, сомы и синапсов.

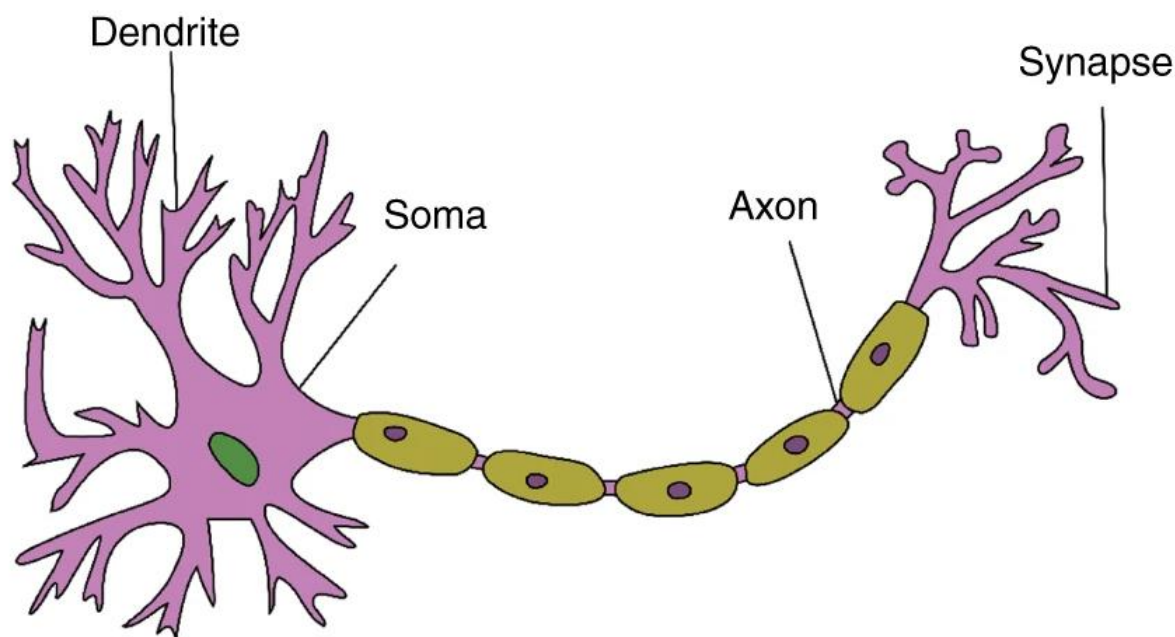


Рисунок 1 Схематическое изображение биологических нейронов [10].

Нейронные сети (ANNs) следуют упрощенной модели, вдохновленной BNN. Основным элементом этой модели является искусственный нейрон, предложенный МакКаллохом и Питтсом, который выполняет вычислительные операции в ANN. Для выполнения задачи происходит параллельная активация множества искусственных нейронов. Как показано на рисунке 2, входы (x) от других нейронов поступают в искусственный нейрон, где они суммируются, а затем вырабатывается выход (y), который передается другим нейронам через аксоны. Синапсы между нейронами имеют веса (w), а сама функция активации (или передаточная функция f) является нелинейной. Эти функции могут быть, например, ступенчатой или сигмоидной. Математическое представление искусственного нейрона выглядит так:

$$y_i = f\left(\sum_{i=0}^n w_{ij} x_i\right)$$

Аналогия между искусственными и биологическими нейронами заключается в том, что межнейронные соединения соответствуют аксонам и дендритам, суммирование и пороговое значение — активации в соме, а веса соединений и память — синапсам.

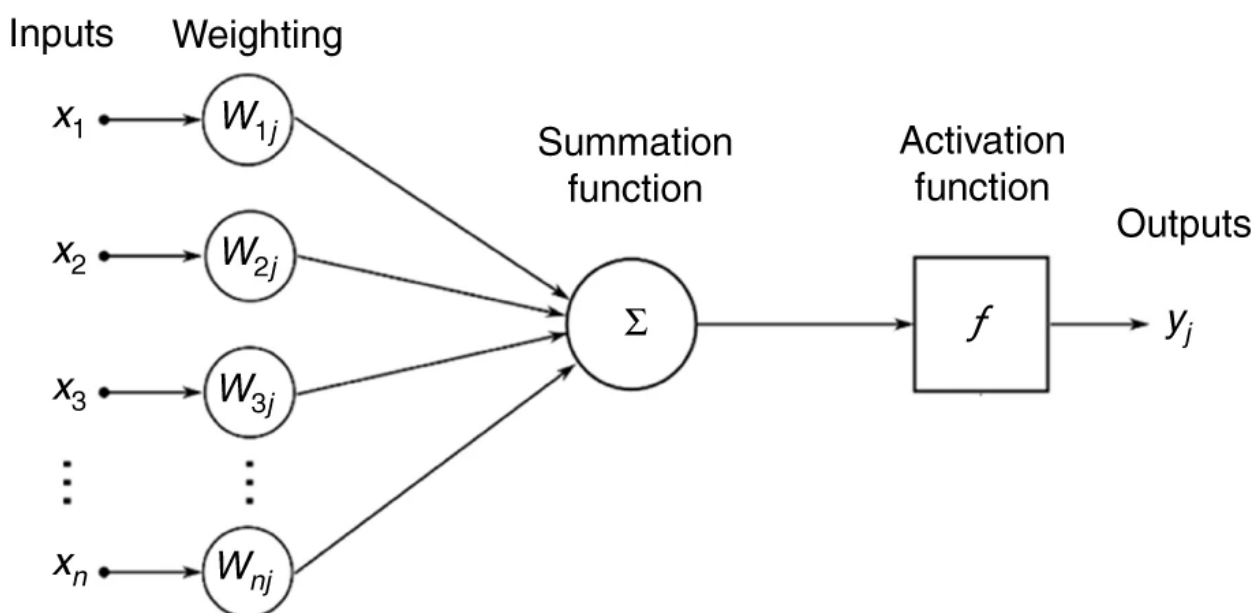


Рисунок 2. Схематическое изображение искусственных нейронов [11].

Архитектура искусственных нейронных сетей

Искусственные нейроны могут быть организованы в различных топологических структурах для формирования нейронных сетей (ANN). Одной из распространенных архитектур является feedforward сеть, изображенная на рисунке 3. В этой сети информация передается только в одном направлении. Несколько искусственных нейронов объединяются в слой ввода, после чего сигнал проходит через один или несколько скрытых слоев. В конце структура выходит на выходной слой, который предоставляет результат. В отличие от традиционного компьютера, который выполняет задачи согласно заранее подготовленной программе, нейронная сеть может научиться выполнять задачу, обучаясь на примерах.

Методы обучения нейронных сетей

Одним из типичных методов обучения является обучение с учителем, когда происходит обновление весов на основе ошибки, полученной при сравнении фактических значений с целевыми. Также существует метод обучения без учителя — спайковая пластичность, зависящая от времени (STDP),

вдохновленный особенностями работы биологических нейронов. Исследования показали, что биологические нейроны кодируют информацию в тайминге одиночных всплесков активности. STDP является механизмом для реализации правила Хебба, что позволяет обучать нейронные сети без внешнего контроля.

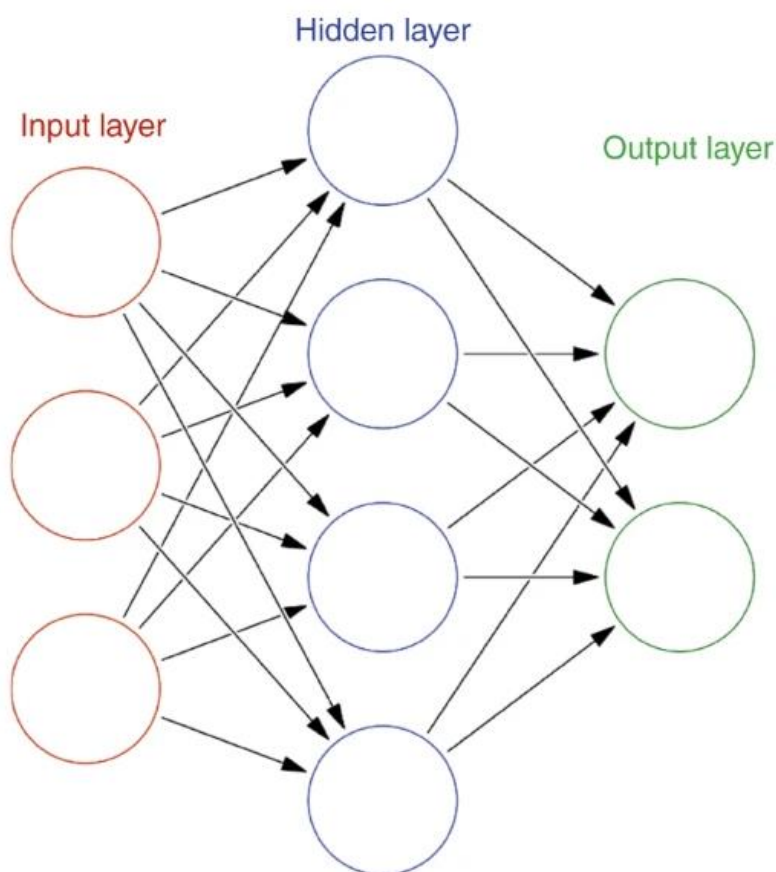


Рисунок 3. Структура искусственных нейронных сетей прямого распространения [11]

Переход к нейроморфным чипам

С ростом интереса к искусственному интеллекту (ИИ) и глубокому обучению (Deep Learning) всё больше внимания уделяется разработке специализированных процессоров, способных эффективно справляться с возрастающими вычислительными нагрузками. Среди таких решений особое место занимают нейроморфные чипы, или нейрочипы. Эти устройства, вдохновлённые структурой и принципами работы биологических нейронных сетей, предоставляют новые возможности для выполнения сложных задач с минимальными энергозатратами и высокой скоростью обработки данных. Для понимания их потенциала и отличий от традиционных процессоров необходимо рассмотреть ключевые аспекты: архитектуру, методы обработки данных, энергоэффективность и сферы применения.

Архитектура нейрочипов в сравнении с традиционными процессорами

Традиционные процессоры функционируют на основе централизованной архитектуры. Центральные процессоры (CPU) выполняют последовательные инструкции, управляя задачами через временное разделение ресурсов, тогда как GPU предлагают высокую степень параллелизма, что особенно полезно для обработки больших массивов данных. Однако они опираются на стандартизированные вычислительные блоки, ограниченные фиксированными алгоритмами и маршрутами передачи данных, что снижает их гибкость при работе с задачами ИИ.

Нейрочипы отличаются распределенной архитектурой, состоящей из множества искусственных нейронов и синапсов, взаимодействующих локально и параллельно, что напоминает работу биологических систем. Они используют спайковую обработку информации, где данные передаются через электрические импульсы (спайки), моделируя временную динамику сигналов. Ключевым

преимуществом является адаптивность: синаптические веса могут изменяться в процессе обучения, что позволяет системе адаптироваться к новым условиям.

Методы обработки данных и энергоэффективность

Традиционные чипы обрабатывают данные по принципу последовательных или параллельных вычислений с фиксированными маршрутами и блоками. Это делает их универсальными, но менее эффективными для задач, связанных с обработкой больших объемов данных в реальном времени.

Нейрочипы предлагают параллельную и событийно-ориентированную обработку, где информация передаётся только при возникновении определённых событий. Локальная обработка данных каждым нейроном позволяет минимизировать задержки и снизить энергозатраты. Такие подходы делают нейрочипы особенно эффективными для задач ИИ, требующих обработки в реальном времени, как, например, распознавание образов или управление автономными системами.

Энергоэффективность является одним из важнейших преимуществ нейроморфных чипов. Высокое энергопотребление традиционных процессоров связано с постоянной работой всех вычислительных блоков и значительным тепловыделением, что требует сложных систем охлаждения и ограничивает их использование в масштабируемых системах. Событийно-ориентированная обработка и локальная обработка данных позволяют нейрочипам работать только тогда, когда это необходимо, минимизируя избыточные вычисления. Это приводит к значительному снижению энергопотребления и тепловыделения, упрощая охлаждение и повышая надежность системы.

Сферы применения нейрочипов

Нейроморфные чипы находят применение в широком спектре областей, требующих высокой скорости обработки данных, адаптации и

энергоэффективности. В настоящее время они активно используются в системах распознавания образов, где их архитектура позволяет эффективно обрабатывать изображения и видео, что делает их незаменимыми в задачах компьютерного зрения, таких как автономные транспортные средства, видеонаблюдение и медицинская диагностика. Например, чип Intel Loihi демонстрирует высокую производительность в анализе сенсорных данных и управления роботами.

В робототехнике нейрочипы применяются для создания автономных систем, способных адаптироваться к изменениям окружающей среды. IBM TrueNorth используется в задачах, связанных с управлением движением и взаимодействием роботов с людьми. В области медицинских технологий нейрочипы играют важную роль в разработке имплантатов и других устройств, обрабатывающих нейронные сигналы в реальном времени. Например, они применяются для управления протезами или мониторинга мозговой активности. Кроме того, в исследованиях искусственного интеллекта (ИИ) нейроморфные процессоры, такие как BrainScaleS, используются для моделирования сложных нейронных сетей, что открывает новые возможности в разработке обучаемых систем для анализа больших данных.

Использование электронной микроскопии в разработке нейроморфных чипов

Одним из ключевых методов исследования нейроморфных чипов является использование электронного микроскопа, который позволяет детально изучать микроструктуру материалов, из которых они состоят, а также наблюдать их поведение на атомарном уровне. Например, в рамках разработки новых синапсopодобных элементов, таких как реконфигурируемые проводящие доменные стенки, можно использовать электронную микроскопию для изучения их морфологии и взаимодействия с другими материалами, что позволяет точно контролировать и оптимизировать их характеристики.

Далее, рассмотрим практическое применение метода электронной микроскопии для нейроморфного чипа, интегрированного в крупномасштабную интегральную схему (LSI). На рисунке 1 показан элемент нейрона, изготовленный на LSI, а также реальная фотография нейронной сети. Рисунок 4 представляет схему цепи нейрона, который состоит из четырех транзисторов и двух инверторов, соединенных последовательно. Этот элемент нейрона генерирует бинарное состояние, которое изменяется в зависимости от входного сигнала. Входной сигнал, а также положительный и отрицательный выходы, имеют однонаправленную связь. Когда входной сигнал превышает пороговое напряжение инвертора, генерируется стабильное состояние, при котором положительный выходной сигнал становится V_{ss} , а отрицательный — V_{dd} . Если же входной сигнал ниже порогового напряжения, генерируется активное состояние, при котором положительный выходной сигнал становится V_{dd} , а отрицательный — V_{ss} . Теоретически этот элемент нейрона представляет собой простой буферный блок, который выполняет функцию двух инверторов.

Для изготовления таких нейронных элементов используется технология 180-нм CMOS Si. На LSI создаются 25×25 элементов нейронов, и 12×12 из них расположены через одну строку и колонку, функционируя как нейроны ввода/вывода (I/O), что позволяет проверить работу нейроморфной системы. Скрытые нейронные элементы расположены между I/O-нейронами. Поскольку LSI состоит из обычных цифровых цепей, такие элементы могут быть легко изготовлены с использованием традиционного метода Si CMOS FET. До интеграции синапсных устройств из тонкой пленки AOS, нейронные элементы изолированы и не соединены между собой.

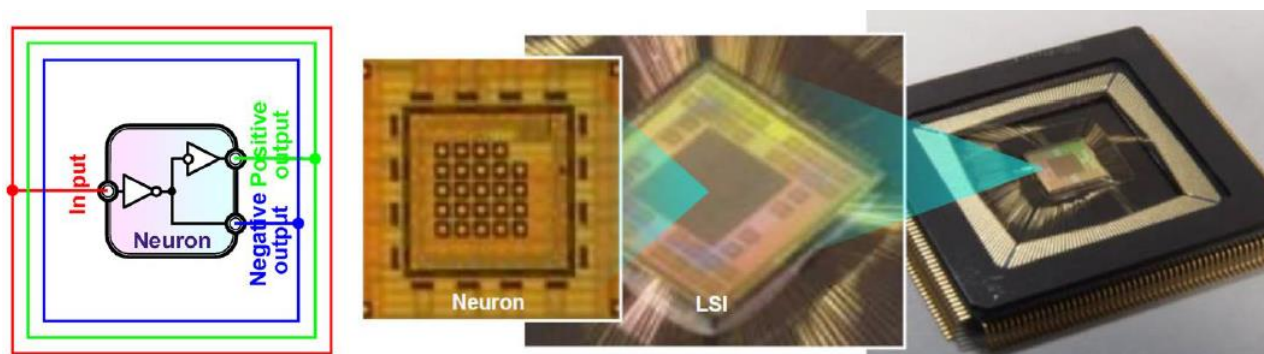


Рисунок 4. Схема нейронного элемента, изготовленного на LSI. Фотография реального нейронного элемента [12].

Рисунок 5 иллюстрирует нейроморфный чип, интегрированный с LSI и синапсными устройствами из аморфно-металлооксидного полупроводника (AOS). Этот тип полупроводников является аморфным и состоит из окислов металлов, что делает его удобным для использования в тонкопленочных устройствах. Аморфно-металлооксидные полупроводники, такие как α -IGZO и α -GTO, обладают хорошими электрическими свойствами и не требуют тепловой обработки при производстве, что позволяет интегрировать их с другими компонентами на одном подложке. Рисунок 2а показывает структуру чипа, а рисунок 5b — срез, полученный с помощью сканирующего электронного микроскопа (СЭМ), который позволяет увидеть микроструктуру и взаимодействия компонентов чипа на атомарном уровне.

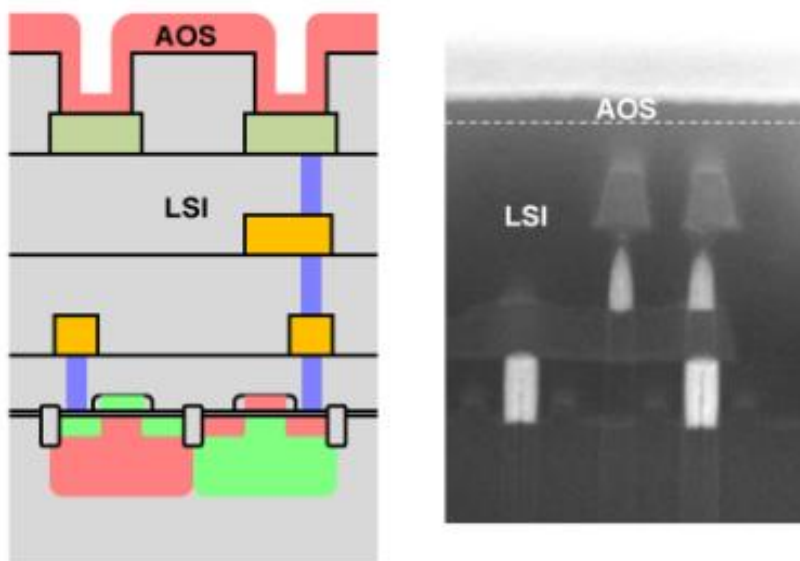


Рисунок 5. (а) Структура нейроморфного чипа, интегрированного с LSI и синапсными устройствами из аморфно-металлооксидного полупроводника (AOS). (б) Фотография среза нейроморфного чипа, полученная с помощью сканирующего электронного микроскопа (СЭМ) [12].

В дальнейших исследованиях можно увидеть, как использовались мемристоры на основе оксидов, например, NbOx для создания нейронов, как это показано на Рисунке 6. Это исследование демонстрирует создание нейросетей, полностью построенных на мемристорах, что открывает возможности для реализации нейроморфных систем с высокой плотностью и функциональностью. [13]

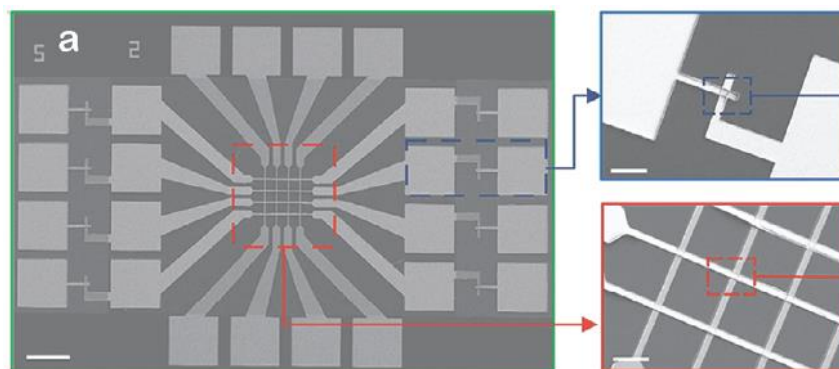


Рисунок 6. СЭМ монолитно интегрированной мемристивной нейронной сети. Масштабная линейка: 100 мкм [13].

Рисунок 7 иллюстрирует дизайн устройства Резистивная память с произвольным доступом (RRAM) для нейроморфных вычислений, использующий в своем дизайне многослойные структуры. Структура устройства RRAM, состоящего из различных слоёв, включая Al_2O_3 в качестве туннельного барьера, TiO_2 и TiO_x с разной оксигеновой стехиометрией. Эти слои были изготовлены с помощью атомно-слоевого осаждения (ALD) и распыления (sputtering).

На рисунке 7а показан SEM-снимок массива перекрестных проводников 16×16 , использующих структуру с отверстиями для подключения (via-hole), что позволяет добиться более равномерного и надежного переключения устройства. Эта структура была выбрана для минимизации эффекта на краях и в угловых областях, которые могут вызывать нестабильное поведение устройства при высоких полях. На рисунках 7б и 7с показаны фотографии одиночного трёхслойного устройства RRAM и среза устройства, что позволяет оценить его микроструктуру и взаимодействие различных слоёв.

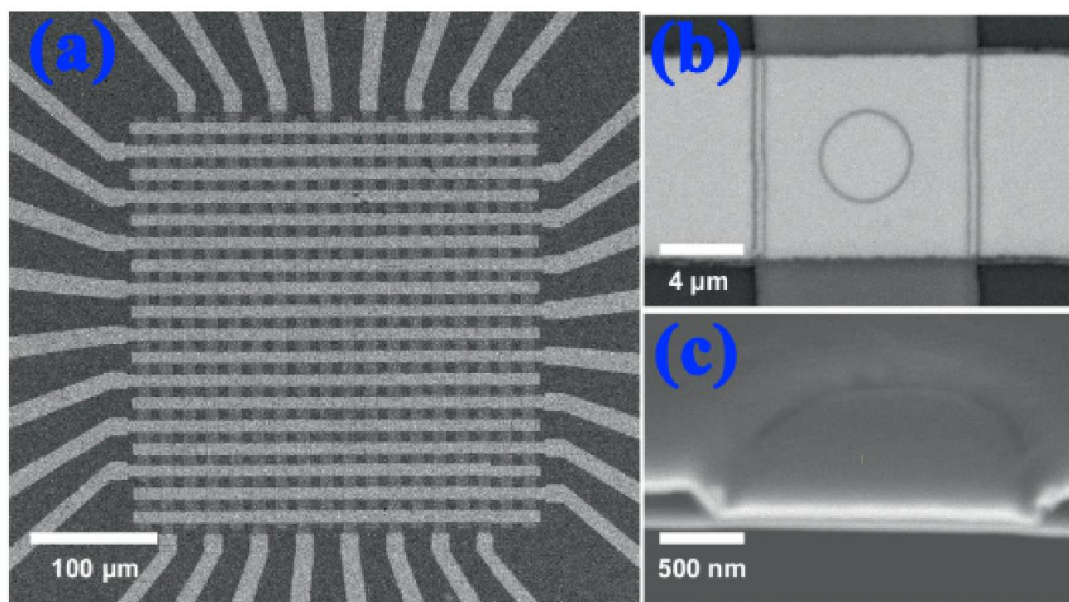


Рисунок. 7: Устройство RRAM. Изображения, полученные с помощью СЭМ, показывают: (а) массив 16×16 элементов в виде перекрестной решетки, (б) одиночный элемент RRAM с трехслойной структурой, (с) срез разрезанного устройства RRAM. [14]

Рисунок 8 демонстрирует фотографию, полученную с помощью СЭМ, на которой изображён массив мемристоров 16×8 , изготовленных в HRL. Вдоль края массива можно увидеть отверстия для подключения (vias) от электродов к металлу 8 в схеме CMOS. Это позволяет интегрировать мемристоры с схемой CMOS, обеспечивая стабильную работу устройства. Следует отметить, что расстояние между мемристорами в данном массиве составляет 10 мкм, что существенно менее агрессивно по сравнению с необходимыми плотностями синапсов для нейроморфных систем, но это решение было выбрано для минимизации рисков при интеграции различных сложных элементов в одном чипе. В [15] показано, что для достижения более высоких плотностей синапсов можно использовать расстояние порядка 100 нм. Эти результаты будут использованы для масштабирования чипа в будущих поколениях.

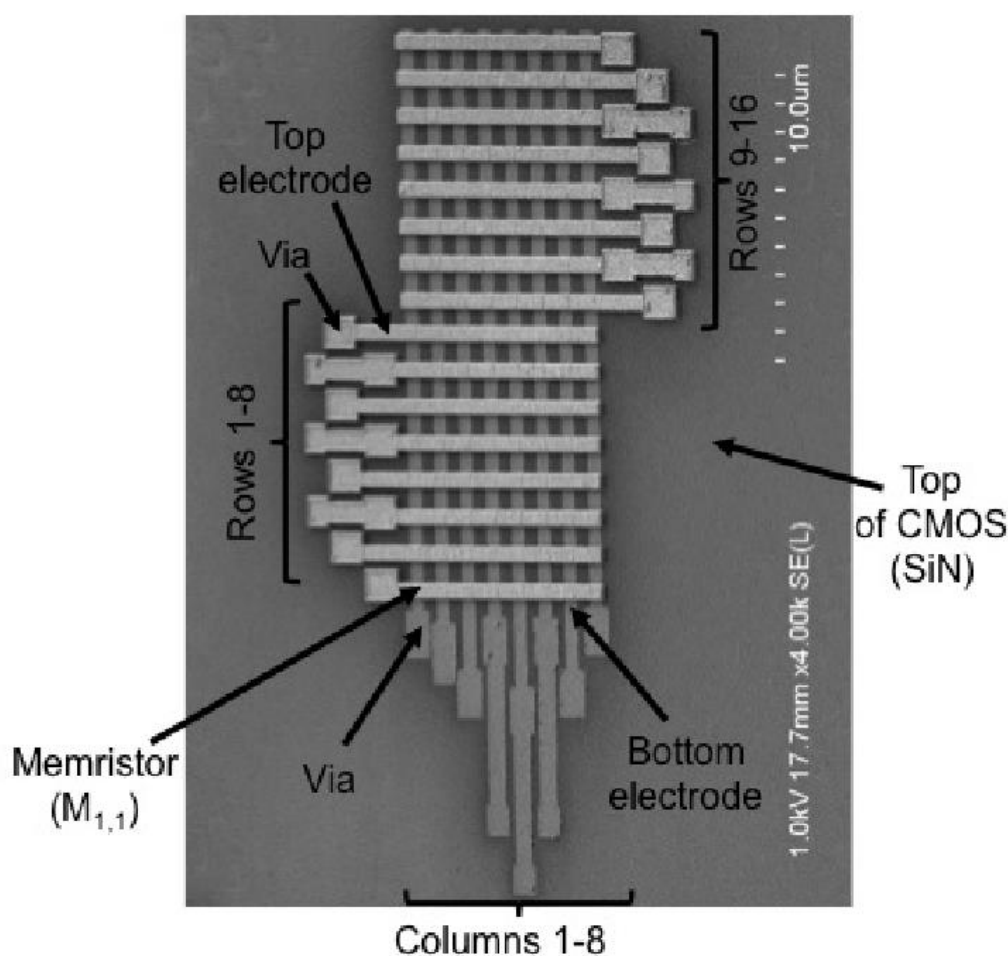


Рисунок 8. Фотография с использованием СЭМ массива мемристоров 16×8 , изготовленных в HRL. [16].

Рисунок 9 представляет собой схему предлагаемого нейроморфного решения для прямой стимуляции нервов (neuromorphic ES) с целью подавления острого воспаления в поврежденных сухожилиях. В данной схеме гибкий электрод оборачивается вокруг симпатической цепи, иннервирующей сегмент сухожилия, обеспечивая прямую стимуляцию. Также используется 2D FGM IDC нейроморфный стимулятор для генерации бионических импульсов.

2D FGM IDC (Flexible Graphene Monolayer Interdigital Capacitor) на основе атомно тонких монослоев MoS_2 действует как нейроморфный стимулятор, генерирующий бионические импульсы. Эти импульсы стимулируют симпатические нервы, что способствует высвобождению норадреналина (NE), который ингибирует воспалительные цитокины, связанные с повреждением сухожилия, минимизируя при этом повреждение нервов.

На рисунке 9: Верхняя панель показывает обёртывание гибких стимуляционных портов вокруг симпатической цепи для прямой стимуляции нервов. Нижняя левая часть иллюстрирует, как норадреналин действует на рецепторы миелоидных клеток, таких как моноциты/макрофаги и гранулоциты, окружающие поврежденное сухожилие, тем самым подавляя цитокины, ответственные за острое воспаление. Нижняя правая часть демонстрирует работу 2D FGM IDC, подключенного к гибкой печатной плате (FPC) и имплантируемым гибким электродам через анизотропную проводящую пленку, вставленной в адаптерную печатную плату (PCB) для соединения с анализатором/монитором.

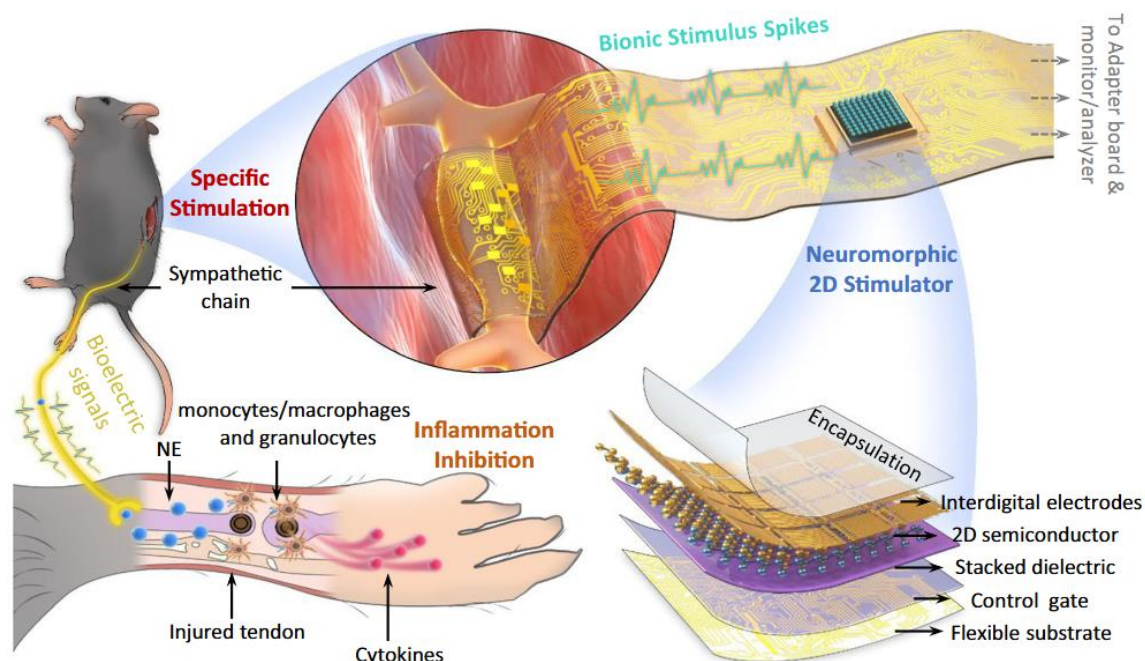


Рисунок 9. Нейроморфная прямая стимуляция нервов для подавления воспаления при повреждении сухожилия.

Схема предлагаемого нейроморфного решения для прямой стимуляции нервов (neuromorphic ES), где гибкий электрод оборачивается вокруг симпатической цепи, иннервирующей сегмент сухожилия, для обеспечения прямой стимуляции, а также используется 2D FGM IDC нейроморфный стимулятор для генерации бионических импульсов. Сочетание прямой стимуляции симпатической нервной системы и бионических импульсов способствует высвобождению норадреналина (NE), который ингибирует воспалительные цитокины, связанные с повреждением сухожилия, при минимальном повреждении нервов. NE: норадреналин [17].

Рисунок 10 показывает СЭМ изображение центральной области 2D FGM IDC. На изображении: желтые пунктирные квадраты обозначают межпечатные (source/drain) электроды. Синие пунктирные квадраты в центре межпечатных электродов представляют каналы MoS_2 . Красные пунктирные квадраты указывают на управляющие затворы (control gates), расположенные под каждым каналом.

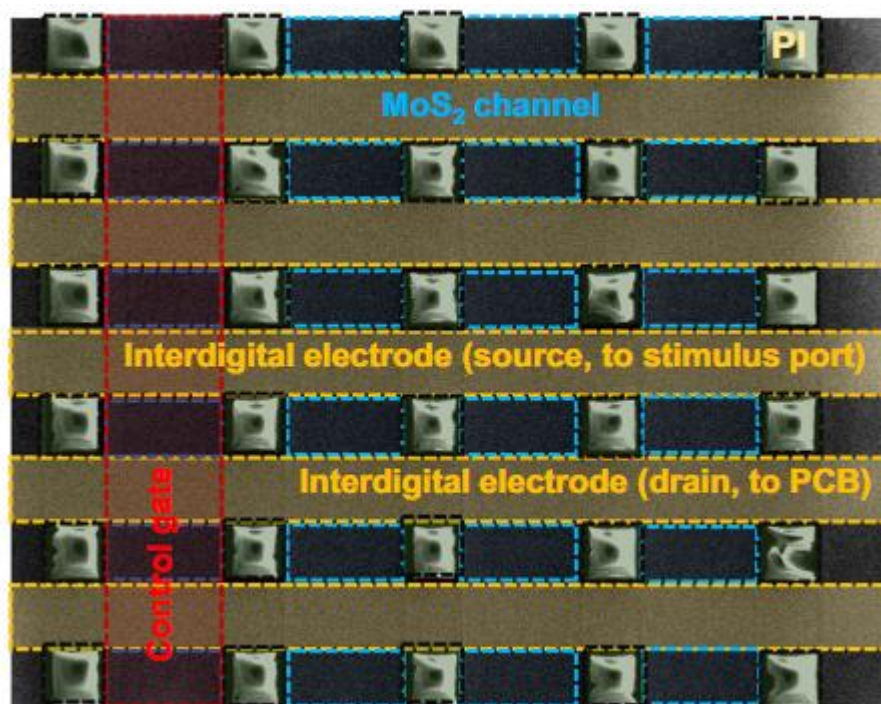


Рисунок 10. СЭМ-изображение 2D FGM IDC. Жёлтые, синие и красные пунктирные квадраты обозначают межпечатные (source/drain) электроды, канал MoS_2 и управляющие затворы соответственно. Масштабная линейка: 50 мкм [17].

Рисунок 11 иллюстрирует реализацию оптической нейронной сети (Optic Artificial Neural Network, OANN) на основе массива мемристоров 1PT1R для задачи распознавания образов. С учетом вышеупомянутых факторов, мы изготовили массив мемристоров 1PT1R и использовали его для имитации человеческой зрительной системы в оптической перцептронной сети. Для реализации массива мемристоров 1PT1R в качестве классификатора оптических изображений мы разработали и изготовили печатную плату в сочетании с оптической измерительной установкой. Рисунок 11(а) показывает схему оптической измерительной установки, которая включает матрицу лазерных источников и модули регулировки света. Как показано на рисунке 11, массив 1PT1R размером 16×3 , соединенный проводами на чип-носителе, использовался для считывания оптических изображений и обучения OANN для задачи распознавания изображений.

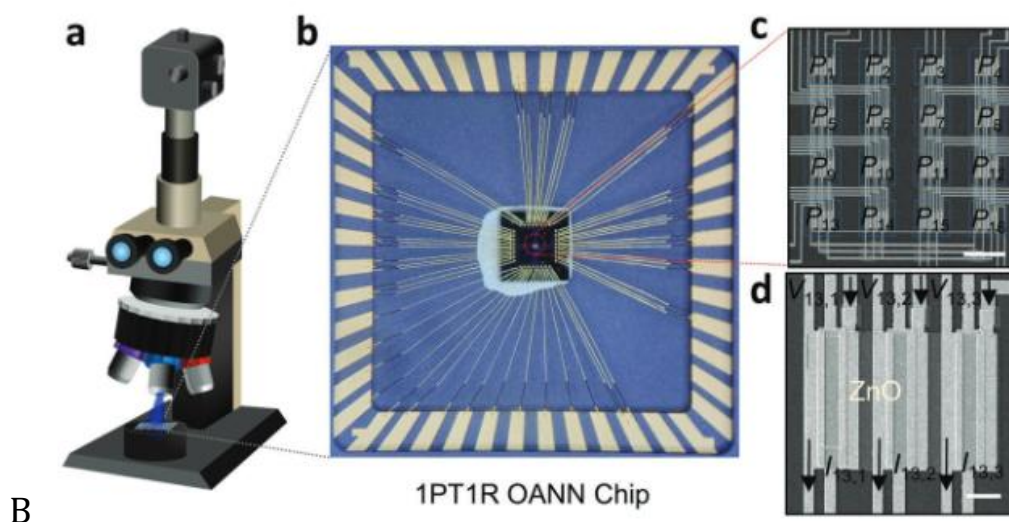


Рисунок 11: Реализация Оптической Нейронной Сети на Основе Мемристоров 1PT1R. 11(a): Схема оптической измерительной установки. 11(b): Макроскопическое изображение массива мемристоров 16×3 на чип-носителе. 11(c): СЭМ-изображение массива мемристоров 16×3. 11(d): Увеличенное СЭМ-изображение устройств в каждом пикселе [18].

Рисунок 11 показывает электронное микроскопическое изображение изготовленного массива 1PT1R. Чип OANN состоял из сенсорного массива 4×4 пикселей (P_i , $i = 1, 2, \dots, 16$), и каждый пиксель содержал три устройства 1PT1R (см. Рисунок 11), что определяется количеством классификационных нейронов [19]. Массив мемристоров 1PT1R (One Transistor One Resistor) используется для моделирования синаптических связей в нейронной сети. Каждый мемристор может изменять своё сопротивление, что позволяет имитировать синаптическую пластичность — способность синапсов изменять силу связи на основе опыта. Оптическая измерительная установка включает матрицу лазерных источников и модули регулировки света, что позволяет подать оптические сигналы на мемристорный массив для обучения нейронной сети. Wire-bonded соединения обеспечивают надёжный контакт между мемристорным массивом и чип-носителем, что необходимо для передачи сигналов и питания устройств. OANN состоит из сенсорного массива 4×4 пикселей, где каждый пиксель содержит три устройства 1PT1R, отвечающие за классификацию образов.

Заключение

Нейрочипы представляют собой значительный шаг вперед в области аппаратного обеспечения для искусственного интеллекта. Их архитектура, основанная на принципах биологических нейронных сетей, обеспечивает высокую параллельность, энергоэффективность и адаптивность, что делает их превосходными для выполнения задач глубокого обучения. В отличие от традиционных процессоров, нейрочипы оптимизированы специально для работы с нейронными сетями, что позволяет значительно ускорить процессы обучения и инференса, а также снизить энергопотребление.

Также, интеграция мемристоров и нейроморфных архитектур открывает новые горизонты для разработки высокоэффективных ИИ-систем. Использование передовых материалов и технологий, таких как аморфно-металлооксидные полупроводники и нанофотоника, позволяет создавать более плотные и функциональные нейроморфные чипы. Это, в свою очередь, способствует развитию приложений, требующих высокой скорости обработки данных и низкого энергопотребления, таких как медицинские устройства, автономные системы и интерфейсы "мозг-компьютер".

Однако, несмотря на значительные достижения, остаются вызовы, связанные с повышением плотности синапсов, улучшением стабильности мемристормых элементов и дальнейшей оптимизацией архитектур нейроморфных чипов. Будущие исследования будут направлены на преодоление этих ограничений, что позволит создать ещё более мощные и адаптивные ИИ-системы, способные решать задачи, аналогичные тем, что выполняются биологическими мозгами.

В заключение, нейроморфные чипы представляют собой перспективное направление в развитии искусственного интеллекта, объединяя достижения в области материаловедения, электроники и нейробиологии. Их способность к эффективной обработке информации, адаптации и низкому энергопотреблению

делает их незаменимыми для создания будущих ИИ-систем, способных работать в реальном времени и решать сложные задачи в различных областях человеческой деятельности.

Список литературы

1. McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (2006). A *proposal for the dartmouth summer research project on artificial intelligence*, august 31, 1955. AI magazine, 27(4), 12-12.
2. Wenqiang Zhang, Bin Gao, Jianshi Tang, Peng Yao, et al., *Neuro-inspired computing chips*//Nat Electron,3,7,371-382,2020-7-21.
3. A. Krikelis, C.C. Weems, *Associative processing and processors*//Computer, 27, 11, 12-17,1994-11.
4. A. Kumar, *Computer-Vision-Based Fabric Defect Detection: A Survey*//IEEE Trans. Ind. Electron.,55,1,348-363,2008-1.
5. Harpreet Kaur, Deepika Koundal, Virender Kadyan, *Image Fusion Techniques: A Survey*//Arch Computat Methods Eng,28,7,4425-4447,2021-12.
6. Neil Savage, *The mind-reading devices that can free paralysed muscles*//Nature,555,7695,S12-S14,2018-3-8.
7. Peng Wu, Junheng Cheng, Feng Chu, *Large-scale energy-conscious bi-objective single-machine batch scheduling under time-of-use electricity tariffs via effective iterative heuristics*//Ann Oper Res,296,1-2,471-494,2021-1.
8. Jafar Shamsi, Amirali Amirsoleimani, Sattar Mirzakuchaki, Majid Ahmadi, *Modular neuron comprises of memristor-based synapse*//Neural Comput & Applic,28,1,1-11,2017-1.
9. A. Chakravarty, J. H. Mentink, S. Semin, Th. Rasing, *Training and pattern recognition by an opto-magnetic neural network*//Applied Physics Letters,120,2,2022-1-10.

- 10.R.E. Uhrig, *Introduction to artificial neural networks*//Proceedings of IECON '95 - 21st Annual Conference on IEEE Industrial Electronics,1,33-37.
- 11.Hinton, G. (1990). *Mental simulation*.
- 12.Mutsumi Kimura, Yuki Shibayama, Yasuhiko Nakashima, *Neuromorphic chip integrated with a large-scale integration circuit and amorphous-metal-oxide semiconductor thin-film synapse devices*//Sci Rep,12,1,2022-3-30.
- 13.Xuegang Duan, Zelin Cao, Kaikai Gao, Wentao Yan, et al., *Memristor-Based Neuromorphic Chips*//Advanced Materials,36,14,2024-4.
- 14.Jaeseoung Park, Ashwani Kumar, Yucheng Zhou, Sangheon Oh, et al., *Multi-level, forming and filament free, bulk switching trilayer RRAM for neuromorphic computing at the edge*//Nat Commun,15,1,2024-4-25.
- 15.Kuk-Hwan Kim, Siddharth Gaba, Dana Wheeler, Jose M. Cruz-Albrecht, et al., *A Functional Hybrid Memristor Crossbar-Array/CMOS System for Data Storage and Neuromorphic Applications*//Nano Lett.,12,1,389-395,2012-1-11.
- 16.Jose M Cruz-Albrecht, Timothy Derosier, Narayan Srinivasa, *A scalable neural chip with synaptic electronics using CMOS integrated memristors*//Nanotechnology,24,38,384011,2013-9-27.
- 17.Rong Bao, Shuiyuan Wang, Xiaoxian Liu, Kejun Tu, et al., *Neuromorphic electro-stimulation based on atomically thin semiconductor for damage-free inflammation inhibition*//Nat Commun,15,1,2024-2-13.
- 18.Bingjie Dang, Keqin Liu, Xulei Wu, Zhen Yang, et al., *One-Phototransistor–One-Memristor Array with High-Linearity Light-Tunable Weight for Optic Neuromorphic Computing*//Advanced Materials,35,37,2023-9.
- 19.M. Prezioso, F. Merrih-Bayat, B. D. Hoskins, G. C. Adam, et al., *Training and operation of an integrated neuromorphic network based on metal-oxide memristors*//Nature,521,7550,61-64,2015-5.

Буряков Арсений Михайлович

Жемеров Евгений Игоревич

Ильин Никита Александрович

**От биологических нейронов к нейроморфным чипам: введение в новые горизонты
искусственного интеллекта**

Монография издана в авторской редакции

Сетевое издание

Ответственный за выпуск – Алимова Н.К.

Научное издание

Системные требования:

операционная система Windows XP или новее, macOS 10.12 или новее, Linux.

Программное обеспечение для чтения файлов PDF.

Объем данных 2 Мб

Принято к публикации «16» декабря 2024 года

Режим доступа: <https://izd-mn.com/PDF/53MNNPM24.pdf> свободный. – Загл. с экрана. – Яз.
рус., англ.

ООО «Издательство «Мир науки»

«Publishing company «World of science», LLC

Адрес:

Юридический адрес – 127055, г. Москва, пер. Порядковый, д. 21, офис 401.

Почтовый адрес – 127055, г. Москва, пер. Порядковый, д. 21, офис 401.

<https://izd-mn.com/>

**ДАННОЕ ИЗДАНИЕ ПРЕДНАЗНАЧЕНО ИСКЛЮЧИТЕЛЬНО ДЛЯ ПУБЛИКАЦИИ НА
ЭЛЕКТРОННЫХ НОСИТЕЛЯХ**