

Моделирование импульсных нейронных сетей с обучающимися синапсами RRAM, вдохновленное строением мозга

Modeling-based design of brain-inspired spiking neural networks with RRAM learning synapses

G. Pedretti, S. Bianchi, V. Milo, A. Calderoni, N. Ramaswamy, and D. Ielmini

Dipartimento di Elettronica, Informazione e Bioingegneria (DEIB), Politecnico di Milano and IU.NET 2017

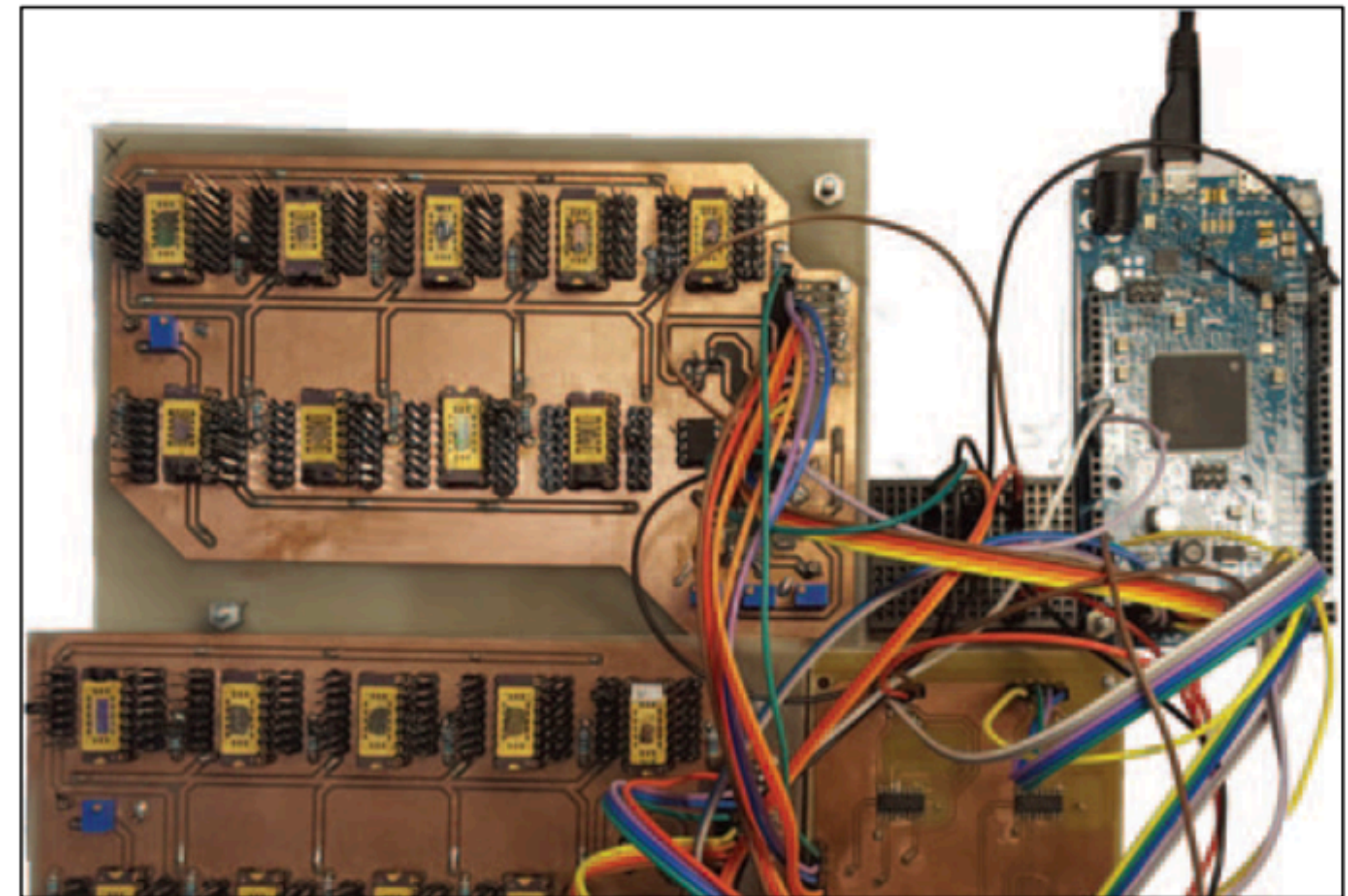
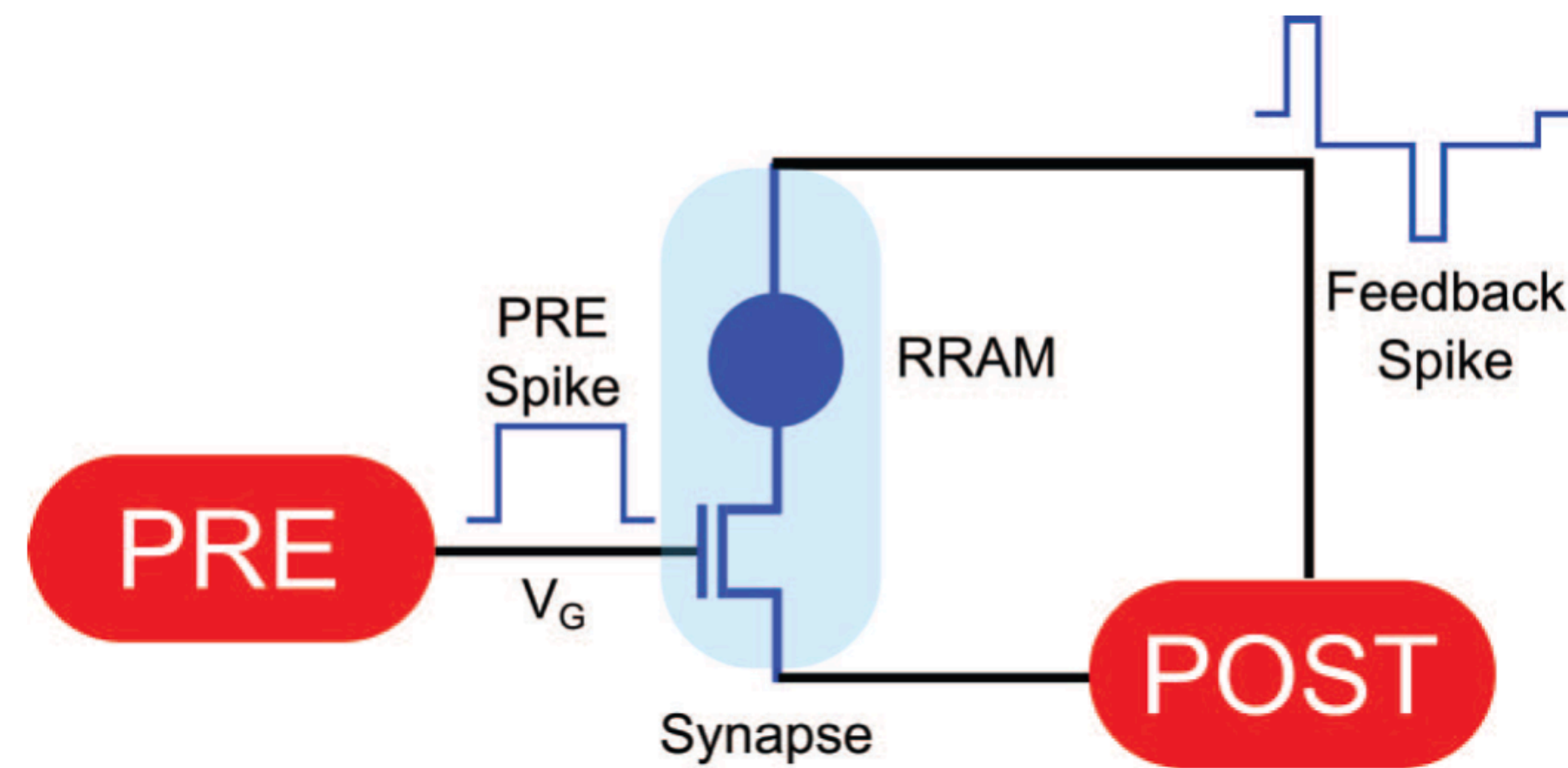
Publisher: IEEE

Мирпулатов Исломбек Пулат-угли

Введение

- Нейроморфные вычисления с RRAM или памятью с фазовым переходом (PCM) позволяют распознавать рукописные символы [1] и лица [2], тем самым обеспечивая основу для компьютерного зрения и навигации дронов/роботов/автомобилей.
- Нейронные сети с импульсной зависимостью от времени (STDP) способны воспроизводить биореалистичное обучение без учителя онлайн [3, 4]
- Нейроморфная схема на основе RRAM выполняет обучение и распознавание статических и динамических паттернов путем реализации правила обучения STDP на аппаратном уровне [5]

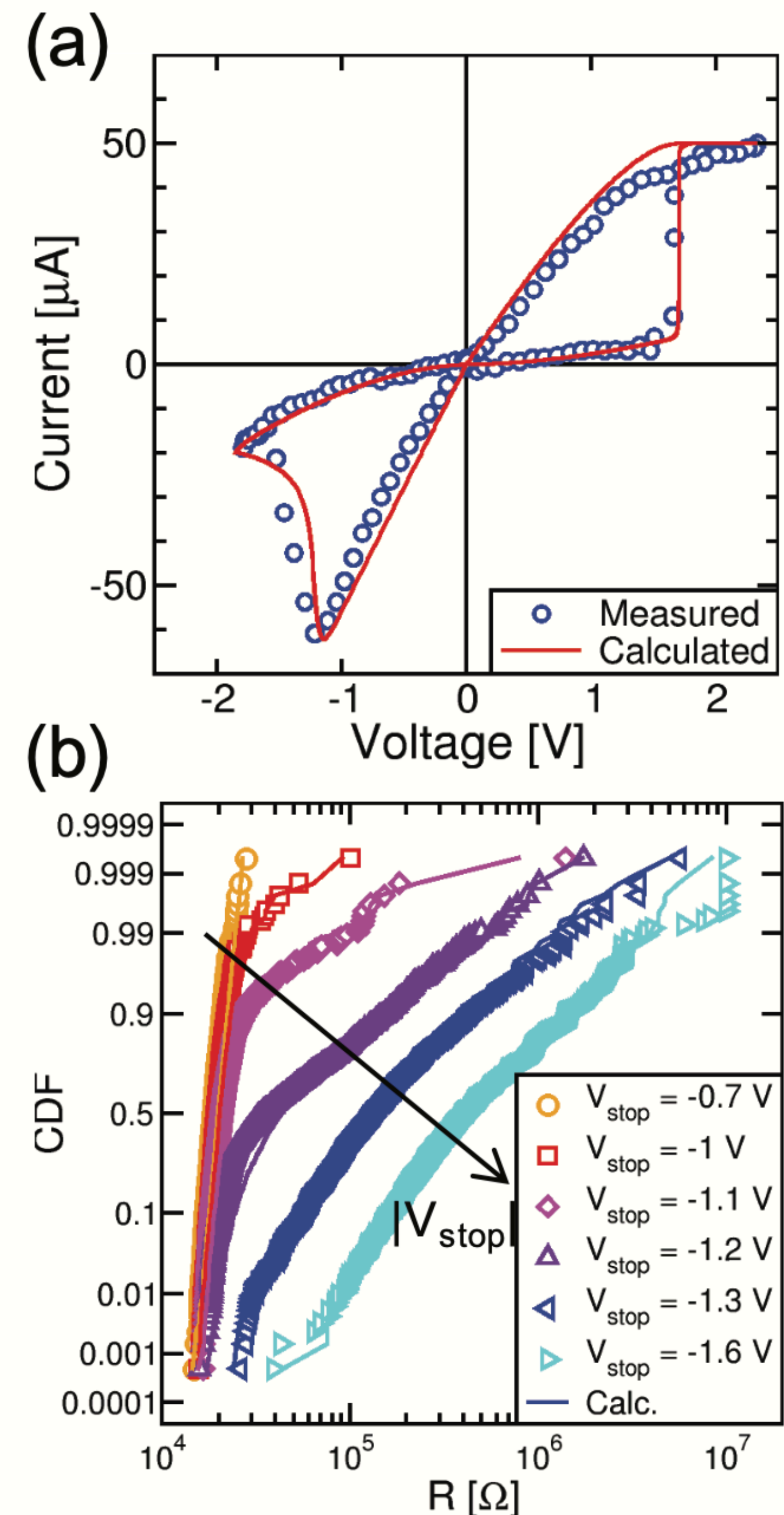
Нейроморфное аппаратное обеспечение



В качестве синаптического элемента использовались устройства RRAM на основе HfO с током установки/сброса $< 50 \mu\text{A}$ и средним окном сопротивления около 10X

Модель Монте-Карло

Модель МС была разработана для описания процессов установки/сброса, происходящих в каждую эпоху (то есть, событие представления шаблона/шума) в процессе обучения



Аналитическая компактная модель

- Для прогнозирования среднего обучающего поведения сети с помощью модели МС требуются длительные симуляции для усреднения эффектов изменчивости. Чтобы ускорить оценку обучения STDP, разработали высокоуровневую аналитическую компактную модель, принимающую уравнения скорости для синаптического веса в паттерне и на заднем плане. Уравнения скорости включают все возможные последовательности входных импульсов (например, шаблон/шум, шум/шаблон и т.д.) в качестве движущей силы для потенцирования/депрессии

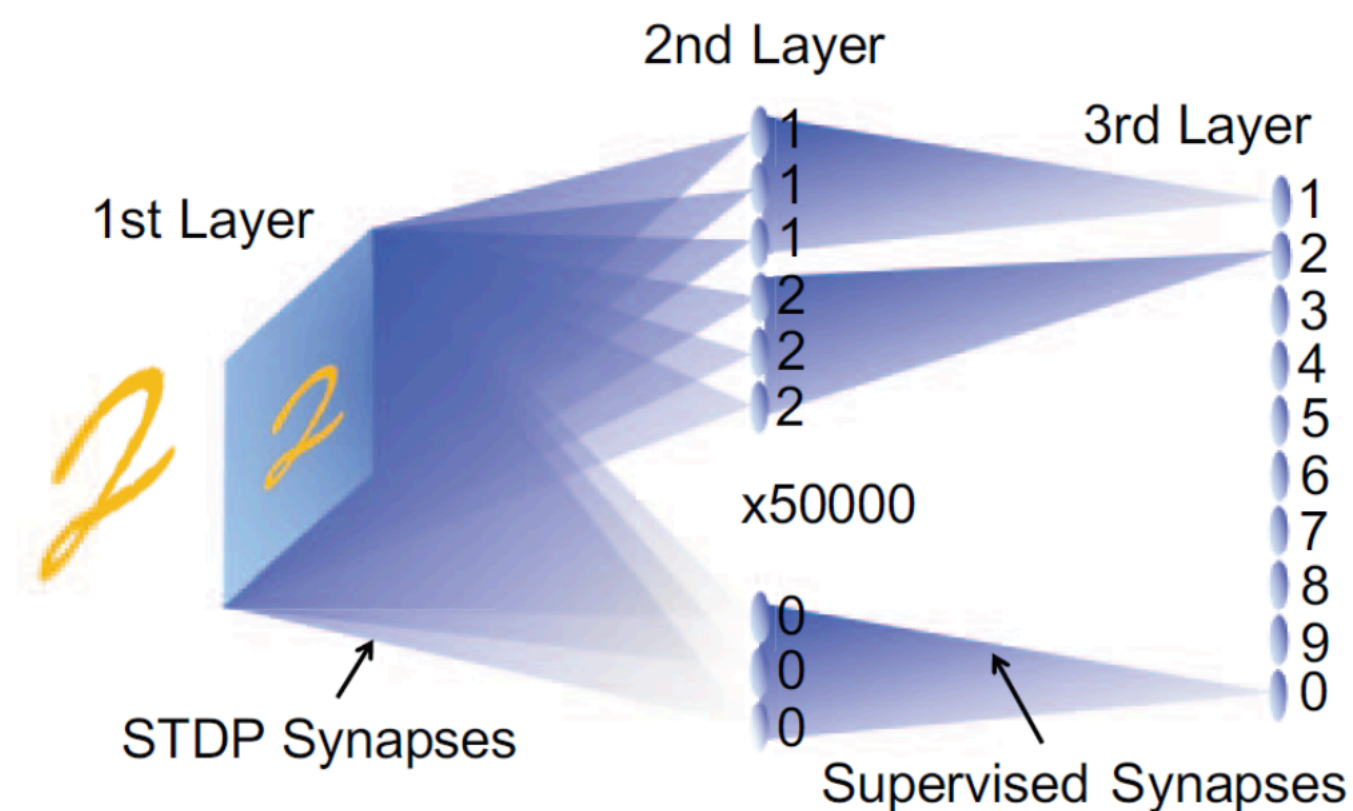
$$\frac{dG_P}{dt} = ANR_N(G_{\max} + G_{\min} - 2G_P) + C(G_{\max} - G_P)(G_P - \alpha NG_{\min})(P - N) * R_P \quad (1)$$

$$\frac{dG_B}{dt} = A^*NR_N(G_{\max} + G_{\min} - 2G_B) + D(\beta G_{\max} - G_B)(G_B - G_{\min})(N - P) * R_N R_P * N(1 - P) \quad (2)$$

- **Влияние плотности шума N**
- По мере увеличения N депрессия ускоряется, однако обучение также становится менее стабильным, о чем свидетельствуют колебания веса со временем.
- **Влияние плотности паттерна P**
- Потенцирование паттернов синапсов наиболее эффективно, если плотность паттернов P намного больше, чем N .
- **Влияние вероятности паттерна/шума**
- Оптимальная вероятность $R_P = R_N = 50\%$, что может быть достигнуто путем систематического чередования между входными каналами шаблона и шума.

Оптимизированное распознавание цифр

- Далее показана принятая схема, состоящая из трехслойного перцептрона с 28x28 PRE в 1-м слое, 50 000 POST во 2-м слое и 10 классификационных нейронов в 3-м слое.
- Синапсы между 1-м и 2-м уровнями обучаются (STDP) обучением без учителя, а синапсы между 2-м и 3-м уровнями обучаются обучением с учителем для классификации.
- Обучение без учителя проводилось либо с фиксированным N ($N = 5\%$ и $N = 10\%$), либо с оптимизированным N.
- Оптимизированный N дает наилучшую эффективность около 92%



Выводы

- Была представлена новая методология разработки и оптимизации импульсной нейронной сети для обучения без учителя, вдохновленного строением мозга, с использованием STDP в синапсах RRAM. Мы показываем, что модели МК цепей RRAM и аналитические компактные модели динамики STDP точно предсказывают поведение обучения в современной импульсной сети с синапсами RRAM.
- Далее демонстрировали повышение эффективности обучения до 92% за счет использования оптимизированного шума во время обучения без учителя рукописным цифрам из базы данных MNIST, тем самым открывая путь для предиктивного проектирования и контроля обучающих схем STDP.

Литература

- 1. G. W. Burr, et al., IEEE Trans. Electron Devices 62, 3498 (2015).
- 2. P. Yao, et al., Nat. Commun. 8, 15199 (2017).
- 3. P. A. Merolla, et al., Science 345, 668–673 (2014).
- 4. N. Qiao, et al., Front. Neurosci. 9:141. (2015).
- 5. G. Pedretti, et al., Sci. Rep. 7:5288 (2017).
- 6. P. U. Diehl and M. Cook, Front. Comput. Neurosci. 9, 99 (2015).
- 7. S. Ambrogio, et al., IEEE Trans. Electron Devices 63, 1508 (2016).
- 8. S. Ambrogio, et al., IEEE Trans. Electron Devices 61, 2378 (2014).
- 9. S. Ambrogio, et al., Symp. VLSI Tech. Dig., 196 (2016).