

## МЕТОДЫ ПРЯМОГО ОБУЧЕНИЯ ГЛУБОКИХ ИМПУЛЬСНЫХ НЕЙРОННЫХ СЕТЕЙ

Ю.М. Вувуникян, Чэнь Ваньли

*Гродненский государственный университет имени Янки Купалы,  
ул. Ожешко, 22, 220023, г. Гродно, Беларусь, [yuv64@mail.ru](mailto:yuv64@mail.ru)*

Импульсные или спайковые нейронные сети (SNN) способны моделировать пространственно-временную информацию и имеют широкие перспективы применения для выявления поведения, подобного мозгу. Недавние схемы, такие как предварительное обучение искусственных нейронных сетей (ИНС) или прямое обучение на основе обратного распространения, обеспечивают высокоэффективное контролируемое обучение SNN. Но другая потенциальная проблема заключается в том, что в импульсных нейронных сетях пиковая активность не дифференцируема, что усложняет контролируемое обучение SNN. Поэтому нам нужны новые алгоритмы для прямого обучения глубоких SNN. В настоящей работе предложен эффективный метод пакетной нормализации, зависящий от порога, для алгоритма пространственно-временного обратного распространения.

**Ключевые слова:** Импульсная (спайковая) нейронная сеть; метод обратного распространения; пакетная регуляризация; метод градиентного спуска; итеративная модель LIF.

## DIRECT LEARNING METHODS FOR DEEP PULSED NEURAL NETWORKS

Yu. M. Vuvunikian, Chen WanLi

*Grodno State University, Ozheshko, 22, Grodno 220023, Belarus  
Corresponding author: [yuv64@mail.ru](mailto:yuv64@mail.ru)*

Pulsed or spiking neural networks (SNNs) are capable of modeling spatiotemporal information and have broad application prospects for detecting brain-like behavior. Recent schemes such as artificial neural network (ANN) pretraining or backpropagation (BP) forward learning provide highly efficient supervised learning for SNNs. But another potential problem is that in bursty neural networks, peak activity is not differentiable, which makes it difficult to supervise SNN training. Therefore, we need new algorithms for direct learning of deep SNNs. In this paper, we propose an efficient threshold-dependent batch normalization method for the spatiotemporal backpropagation algorithm.

**Keywords:** Pulse (spike) neural network; backpropagation method; batch regularization; gradient descent method; iterative model LIF.

## **Введение**

В настоящее время существует два основных подхода к обучению высокопроизводительных SNN.

Первый подход заключается в преобразовании предварительно обученной ИНС в SNN, что обычно занимает сотни временных шагов. Следовательно, хотя эти SNN имеют сравнимую точность со структурно подобными ANN, большое количество временных шагов может вызвать серьезные задержки сигнала и увеличить объем вычислений.

Другой подход заключается в обучении SNN напрямую на основе градиентного спуска, не полагаясь на предварительно обученные нейронные сети и уменьшая временной шаг [2,3].

Текущие SNN, обученные алгоритму STBP, ограничены неглубокими архитектурами и не могут обеспечить удовлетворительную производительность на крупномасштабных наборах данных. Поэтому в этой статье предлагается использовать метод пакетной нормализации, зависящий от порога (tdBN), основанный на пространственно-временном обратном распространении (STBP – Spatio-Temporal BackPropagation), называемый «STBP-tdBN». И строить сеть с глубокими всплесками, используя tdBN, расширяя непосредственно обученную сеть SNN, используя небольшой временной шаг, достигает эффекта прямого обучения нейронных сетей с глубокими пиками.

### **1. Полученные результаты**

1.1. Глубокие нейронные сети (DNN) могут избежать исчезновения или взрыва градиентов, когда они динамически равноудалены (то есть каждое сингулярное значение входной и выходной матриц Якоби поддерживается близким к 1).

В работе [2] Chen и др. предложили «динамическую изометрию блоков» в качестве общего статистического инструмента для всех сложных последовательно-параллельных ГНС, изучая моменты первого и второго порядка каждого блока в нейронной сети и анализируя их влияние на распределение градиента. Теоретически объяснена роль инициализации веса, пакетной нормализации и быстрого соединения в DNN.

#### **1.2. Итерационная модель LIF:**

Итеративная модель LIF была предложена [1] Wu и др. в 2019 году на основе оригинальной модели Leaky Intergrate and Fired.

Найдены импульсные и спектральные характеристики нелинейных эволюционных операторов [5], определяемых моделью LIF.

### 1.3. Пороговая нормализация партии (TDBP).

Пакетная нормализация (BN), как компонент регуляризации DNN, является широко используемым методом нейронной сети, который может обеспечить стабильную конвергенцию и более глубокие нейронные сети.

Однако непосредственно обученные SNN требуют специально разработанных методов нормализации из-за дополнительного временного измерения и специального механизма активации SNN. Поэтому мы используем метод нормализации партии, зависящий от порогового значения (threshold-dependent batch normalization, tdBN [3]).

### 1.4. Общий алгоритм обучения.

Мы вводим общий алгоритм обучения STBP-tdBN для обучения глубокого SNN с нуля с помощью нашего tdBN. При обратном распространении ошибки мы используем последний уровень в качестве уровня декодирования, а конечный результат  $Q$  выражается следующим равенством:

$$Q = \frac{1}{T} \sum_{t=1}^T M o^{n,t}$$

где  $o^{n,t}$  – шип, выпущенный последним выходным слоем,  $M$  — матрица уровня декодирования, а  $T$  обозначает количество временных шагов. Затем мы пропускаем выходные данные через слой softmax. Функция потерь определяется как кросс-энтропия. С учетом выхода  $Q$  и вектора меток  $Y$  функция потерь  $L$  определяется выражением

$$L = - \sum_{i=1}^n y_i \log(p_i), \quad p_i = \frac{e^{q_i}}{\sum_{j=1}^n e^{q_j}}.$$

В итеративной модели LIF метод STBP-tdBN распространяет обратно градиент потерь  $L$  как в пространственной, так и во временной областях. Применяя формулу дифференцирования сложной функции, производные

$\frac{\partial L}{\partial o_i^{t,n}}$  и  $\frac{\partial L}{\partial u_i^{t,n}}$  можно вычислить с помощью равенств

$$\begin{aligned} \frac{\partial L}{\partial o_i^{t,n}} &= \sum_{j=1}^{l(n+1)} \frac{\partial L}{\partial u_j^{t,n+1}} \frac{\partial u_j^{t,n+1}}{\partial o_i^{t,n}} + \frac{\partial L}{\partial u_i^{t+1,n}} \frac{\partial u_i^{t+1,n}}{\partial o_i^{t,n}}, \\ \frac{\partial L}{\partial u_i^{t,n}} &= \frac{\partial L}{\partial o_i^{t,n}} \frac{\partial o_i^{t,n}}{\partial u_i^{t,n}} + \frac{\partial L}{\partial u_i^{t+1,n}} \frac{\partial u_i^{t+1,n}}{\partial u_i^{t,n}}, \end{aligned}$$

где  $o^{t,n}$  и  $u^{t,n}$  представляют спайковый и мембранный потенциал нейрона в слое  $n$  в момент времени  $t$ .

### 1.5. Теория градиентной нормы

Теория нормы градиента — это разработанная в последние годы теория, целью которой является преодоление проблемы исчезновения или взрыва градиента в различных структурах нейронных сетей.

В этой статье мы используем «блочную динамическую изометрию», предложенную в работе [2] в 2020 году, для анализа влияния tdBN на непосредственно обученную SNN, которая рассматривает сеть как систему блоков.

Экспериментальные результаты на статистическом наборе данных показаны в приведенной таблице.

*Таблица:* Полученные экспериментальные данные.

<b>Набор данных</b>	<b>Модель</b>	<b>Методы</b>	<b>Шаг времени</b>	<b>Точность</b>
CIFAR-10	(Sengupta et al. 2019)	ANN-SNN	2500	90.53%
	(Hu et al. 2018)	ANN-SNN	350	91.32%
	(Rathi et al. 2020)	Hybird Training	200	91.01%
	(Lee et al. 2020)	Spike-based BP	100	89.93%
	(Wu et al. 2019)	STBP	12	89.49%
	наша модель	STBP-tdBN	4	91.89%
ImageNet	(Sengupta et al. 2019)	ANN-SNN	2500	68.86%
	(Sengupta et al. 2019)	ANN-SNN	2500	64.47%
	(Han, Srinivasan, and Roy 2020)	ANN-SNN	1024	65.63%
	(Hu et al.2018)	ANN-SNN	768	71.53%
	(Rathi et al. 2020)	Hybird	250	65.38%

		Training		
	(Rathi et al. 2020)	Hybird Training	250	61.35%
	наша модель	Hybird Training	6	65.72%

## Заключение

В этой статье мы используем метод нормализации, который напрямую обучает высокопроизводительные глубокие SNN. Включая теорию градиентной нормы, мы демонстрируем, что метод может эффективно сбалансировать входные стимулы и пороги нейронов во время обучения, тем самым способствуя конвергенции обучения. Мы достигаем приемлемой точности на наборах данных статических изображений со значительно меньшей задержкой вывода по сравнению с другими моделями SNN.

Поэтому в этой работе предлагается возможная схема прямого обучения для глубоких SNN. Он поддерживает высокую эффективность механизма связи на основе спайков, позволяет SNN решать более сложные задачи классификации, облегчает реализацию на соответствующем оборудовании и практическое применение SNN.

## Библиографические ссылки

1. Wu Y, Deng L, Li G, et al. Direct training for spiking neural networks: Faster, larger, better[C] // Proceedings of the AAAI Conference on Artificial Intelligence. 2019, № 33(01). P. 1311–1318.
2. Chen Z, Deng L, Wang B, et al. A comprehensive and modularized statistical framework for gradient norm equality in deep neural networks[J] // IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020. № 44(1). P. 13–31.
3. Wu Y, Deng L, Li G, et al. Spatio-temporal backpropagation for training high-performance spiking neural networks[J] // Frontiers in neuroscience, 2018. № 12. P. 331.
4. Yao M, Gao H, Zhao G, et al. Temporal-wise Attention Spiking Neural Networks for Event Streams Classification[C] // Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021.P. 10221–10230.
5. Вувуникян Ю.М. Эволюционные операторы с обобщенными импульсными и спектральными характеристиками: монография. Гродно : ГрГУ, 2007. 224 с.