

Обучение импульсных нейронных сетей с помощью комбинированных хеббовских правил

Д.И. Антонов
Ульяновский государственный
технический университет
Ульяновск, Россия
d.antonov@ulireran.ru

Б.Б. Батуев
ИРЭ им. В.А. Котельникова РАН
Москва, Россия
buligarmouth@gmail.com

С.В. Сухов
УФИРЭ им. В.А. Котельникова РАН
Ульяновск, Россия
ssukhov@ulireran.ru

Аннотация — В статье представлен новый метод обучения импульсных (спайковых) нейронных сетей (ИмНС). Данный метод основан на локальных правилах изменения силы синаптической связи. Разработанный метод обучения с учителем использует комбинацию хеббовских правил - так называемой пластичности, зависящей от времени импульса (STDP), и постоянной долгосрочной депрессии (All-LTD). ИмНС были реализованы в пакете Brian 2.0. Численные эксперименты проводились на общедоступных наборах данных.

Ключевые слова — импульсная нейронная сеть, зависящая от времени пластичность, обучение с учителем

I. ВВЕДЕНИЕ

Несмотря на существенный прогресс в развитии методов обучения искусственных нейронных сетей, продолжается поиск альтернативных методов обучения на основе нашего понимания принципов работы мозга. В русле данных исследований появились модели биологически правдоподобных импульсных нейронных сетей (ИмНС) [1]. Интерес к исследованию ИмНС обусловлен их энергоэффективностью и алгоритмом обработки сигналов, изменяющихся во времени, внутренне присущем данному виду нейронных сетей.

Предыдущие поколения искусственных нейронных сетей обучались с помощью метода обратного распространения ошибки. Этот метод с трудом обобщается на случай ИмНС. Для обучения ИмНС более подходящими являются правила, основанные на локальном изменении синапсов (весов). Соответствующее правило изменения синапсов называется пластичностью, зависящей от времени спайка (импульса) (spike-time-dependent plasticity, STDP) [2, 3]. Согласно этому правилу, сила синаптической связи w изменяется в зависимости от временной разницы возбуждения спайков в пре- и постсинаптическом нейронах $\Delta t = t_{\text{post}} - t_{\text{pre}}$ (рис. 1). STDP обычно используется при обучении без учителя [4]. В нашем подходе мы модифицировали STDP для использования в задачах обучения с учителем.

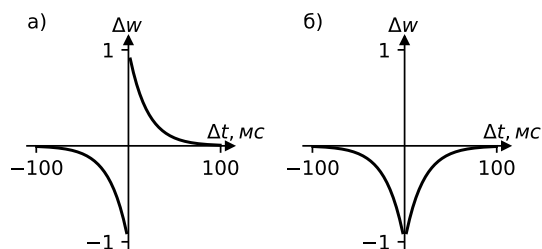


Рис. 1. Правила изменения синаптической связи: обычное STDP (а), All-LTD вариант (б)

II. МЕТОД

Логика работы обычного механизма STDP состоит в следующем. Если пресинаптический спайк предшествует постсинаптическому ($\Delta t > 0$), то можно предполагать, что пресинаптический спайк вызвал постсинаптический. В этом случае усиление соответствующей синаптической связи позволяет нейрону выделять во входном сигнале повторяющиеся последовательности (выделять характерные черты сигнала). Комбинация STDP и латерального торможения [4] приводит к тому, что каждый из нейронов реагирует на отдельный признак.

Для реализации обучения с учителем мы внесли в STDP некоторые изменения. Мы поделили нейроны в выходном слое сети на подгруппы. Каждая из подгрупп нейронов обучается реагировать на признаки объектов определенного класса (рис. 2). В отличие от обычного STDP нейроны должны обучиться реагировать только на уникальные признаки класса и игнорировать признаки, общие для нескольких классов. Игнорирование общих признаков достигается путем декоррелирующего (антихеббовского, депрессирующего) STDP – All-LTD (All-time long-time depression) (рис. 1б). Если пресинаптический нейрон вызвал спайк в постсинаптическом нейроне, All-LTD подавляет такое поведение путем уменьшения синаптической связи.

Процесс обучения происходит следующим образом. Для каждого класса выделяется подгруппа нейронов. При подаче на вход сети сигнала определенного класса нейроны, предназначенные для распознавания этого класса, обучаются с помощью STDP. Нейроны, ответственные за распознавание других классов, при

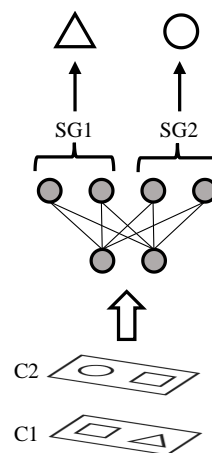


Рис. 2. Принцип обучения ИмНС, основанный на выделении признаков классов. Для объектов двух классов C1 и C2 соответствующие нейронные группы SG1 и SG2 выделяют уникальные признаки (круги и треугольники) и игнорируют общие признаки (квадраты)

этом обучаются с помощью All-LTD. Процесс последовательно повторяется для всех образцов из тренировочного набора.

Таким образом, мы представляем новый принцип обучения с учителем, который основан не на следовании какому-то обучающему сигналу и не на максимизации некоторой целевой функции, а на выделении признаков объектов.

III. ЧИСЛЕННЫЙ ЭКСПЕРИМЕНТ

Архитектура ИмНС, использованная в наших экспериментах, аналогична архитектуре сети в работе [4]. Сеть была построена на интегрирующих нейронах с утечкой. Сеть обучалась классифицировать изображения из набора MNIST [5]. ИмНС смоделирована в пакете Brian 2 [6]. Сеть состоит из двух слоёв, первый слой содержит 784 пуассоновских нейрона (по количеству пикселей в изображениях), второй слой состоит из 30 возбуждающих и 30 тормозящих нейронов.

Работа сети, изображённой на рис. 3, заключается в циклическом выполнении следующих шагов:

- на вход сети подаётся изображение размера 28x28 пикселей, каждый пиксель которого попадает на один из 784 нейронов пуассоновского слоя и вызывает генерацию последовательности импульсов с частотой, равной интенсивности данного пикселя;
- далее сигнал по синапсам группы S_1 попадает на возбуждающие нейроны по принципу «один-ко-всем»; группа S_1 ответственна за обучение сети;
- импульсы, сгенерированные возбуждающими нейронами, передаются по синапсам группы S_2 в тормозящие нейроны второго слоя по принципу «один-к-одному»;
- от тормозящих нейронов сигнал для обеспечения латерального торможения идёт по группе синапсов S_3 назад к возбуждающим нейронам по принципу «один-ко-всем-кроме-иницирующего».

Возбуждающие нейроны используются для предсказания класса. Мы использовали три класса изображений из набора MNIST. Для каждого класса было выделена подгруппа, содержащая десять возбуждающих нейронов второго слоя. Каждый из нейронов подгруппы обучался одному из признаков изображений соответствующего класса. Обучение предполагало изменение весов синапсов, идущих от пуассоновских нейронов к возбуждающим. Эти синапсы обучались с помощью STDP – All-LTD локальных правил.

При обучении изображения показывались сети по одному в течение 350 мс. Перед показом следующего изображения в течение 150 мс на вход сети сигналы не подавались, чтобы параметры сети релаксировали к значениям, соответствующим покою. Обучение проводилось на 5000 изображениях MNIST. После обучения точность классификации проверялась на 1000 тестовых изображениях. Для каждого показанного сети

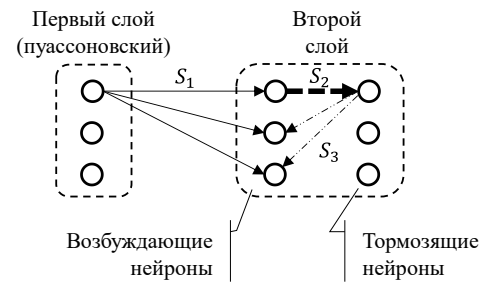


Рис. 3. Архитектура нейронной сети, количество нейронов в слоях показано условно

изображения класс определялся по наибольшей активности популяций нейронов, отведенных для каждого класса. Полученная при обучении в течение одной эпохи точность классификации составила $74,9 \pm 19,8\%$. Для сравнения мы также обучили сеть в соответствии с методом [7], а именно: ИмНС была обучена с помощью STDP; классификация при этом осуществлялась с помощью внешнего классификатора, основанного на дереве решений. При использовании внешнего классификатора точность составила $74,6 \pm 6,1\%$.

IV. ЗАКЛЮЧЕНИЕ

Нами предложен и реализован метод обучения импульсной сети-классификатора, основанный на локальных методах обучения. Метод базируется на выделении уникальных признаков объектов. Обученная сеть показала возможность классифицировать набор MNIST с той же точностью, что и при использовании внешнего классификатора. В данной работе метод был опробован для классификации стационарных пространственных данных. В будущем мы планируем протестировать наш метод для классификации сигналов, изменяющихся во времени.

Работа выполнена в рамках государственного задания ИРЭ им. В.А. Котельникова РАН.

ЛИТЕРАТУРА

- [1] Maass, W. Networks of spiking neurons: The third generation of neural network models / W. Maass // Neural Networks. – 1997. – Vol. 10(9). – P. 1659-1671.
- [2] Caporale, N. Spike Timing-Dependent Plasticity: A Hebbian Learning Rule / N. Caporale, Y. Dan // Annual Review of Neuroscience. – 2008. – Vol. 31(1). – P. 25-46.
- [3] Tavanaei, A. Deep learning in spiking neural networks / A. Tavanaei, M. Ghodrati, S.R. Kheradpisheh et al. // Neural Networks. – 2019. – Vol. 111. – P. 47-63.
- [4] Diehl, P.U. Unsupervised learning of digit recognition using spike-timing-dependent plasticity / P.U. Diehl, M. Cook // Front. Comput. Neurosci. – 2015. – Vol. 9(99). DOI: 10.3389/fncom.2015.00099.
- [5] Deng, L. The MNIST database of handwritten digit images for machine learning research / L. Deng // IEEE Signal Processing Magazine. – 2012. – Vol. 29(6). – P. 141-142.
- [6] Goodman, D. Brian: a simulator for spiking neural networks in Python / D. Goodman, R. Brette // Front. Neuroinform. – 2008. – Vol. 2(5). – P. 350. DOI:10.3389/neuro.11.005.2008.
- [7] MNIST spiking neural network [Electronic resource]. – Access mode: <https://www.kaggle.com/code/dlaronov/mnist-spiking-neural-network/notebook> (10.10.2023)