



Институт радиотехники
и электроники РАН
им. В.А.Котельникова



Д.И. АНТОНОВ

Алгоритмы обучения импульсных нейронных сетей для нейроморфных систем обработки информации

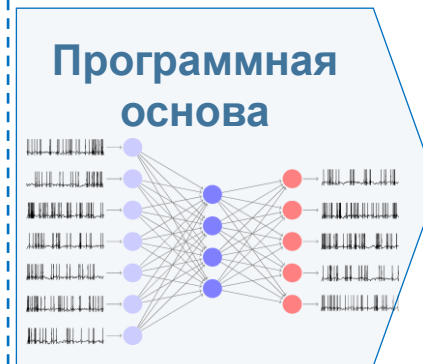
Москва
2025

Связь алгоритмов обучения импульсных нейронных сетей и нейроморфных процессоров (1)

Импульсные нейронные сети (ИмНС)



Нейроморфные системы обработки информации



Связь алгоритмов обучения импульсных нейронных сетей и нейроморфных процессоров (2)

Нейроморфные системы обработки информации прямо сопряжены с физической реализацией импульсных нейронных сетей (ИмНС)

Базовые свойства ИмНС:

- кодировка информации с помощью импульсов
 - локальные методы обучения*
- лежат в основе высокой энергоэффективности нейроморфных процессоров на ИмНС

*локальные методы обучения имманентно присущи ИмНС, возможно обучение ИмНС также и нелокальными методами

Актуальность темы

Магистральное направление развития искусственного интеллекта — развитие нейроморфных технологий на базе биоподобных импульсных нейронных сетей (ИмНС)

Принципы работы ИмНС и искусственных нейронных сетей (НС) 2-го поколения различны: в ИмНС кодирование информации и передача сигналов имеют импульсную природу, что имитирует свойства биологических нейронов

Для нейронных сетей 2-го поколения разработаны комплексы методов:

- **методы снижения потерь от «катастрофического забывания» при непрерывном обучении, когда дообучение сети на новых категориях данных вызывает забывание «старых» знаний**
- **методы эффективного обучения нейронных сетей**
- **методы регуляризации для преодоления переобучения сетей**

Для искусственных нейронных сетей 3-го поколения (импульсных нейронных сетей) требуется развитие методов решения аналогичных проблем

Структура доклада

1. Метод преодоления катастрофического забывания при непрерывном обучении импульсных нейронных сетей
 2. Метод локального обучения импульсных нейронных сетей на основе правил STDP и All-LTD
 3. Метод регуляризации импульсных нейронных сетей
 4. Научные и практические результаты работы
-

Структура доклада

1. Метод преодоления катастрофического забывания при непрерывном обучении импульсных нейронных сетей
2. Метод локального обучения импульсных нейронных сетей на основе правил STDP и All-LTD
3. Метод регуляризации импульсных нейронных сетей
4. Научные и практические результаты работы

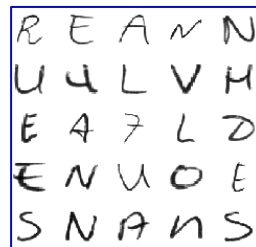
Непрерывное обучение и катастрофическое забывание

Непрерывное обучение

Задача 1



Задача 2



Базы данных MNIST и EMNIST — базы данных рукописных изображений цифр и букв

Процесс непрерывного обучения

Задача 1: обучение на изображениях цифр

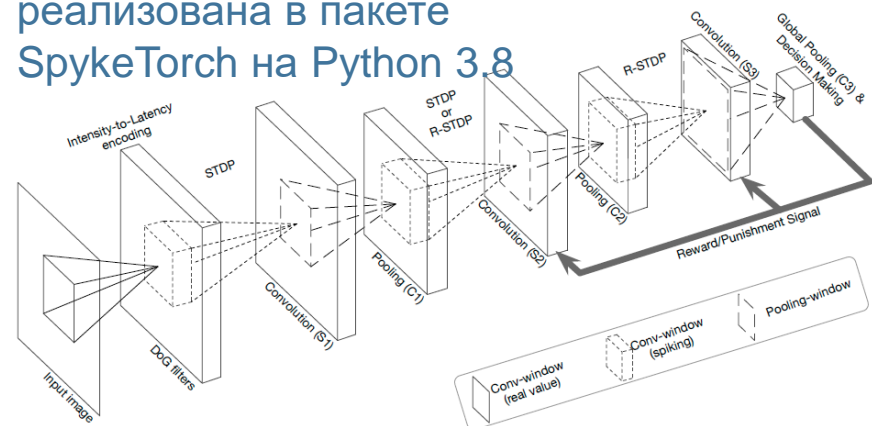
Задача 2: переобучение на изображениях букв



Исследованные способы предотвращения катастрофического забывания

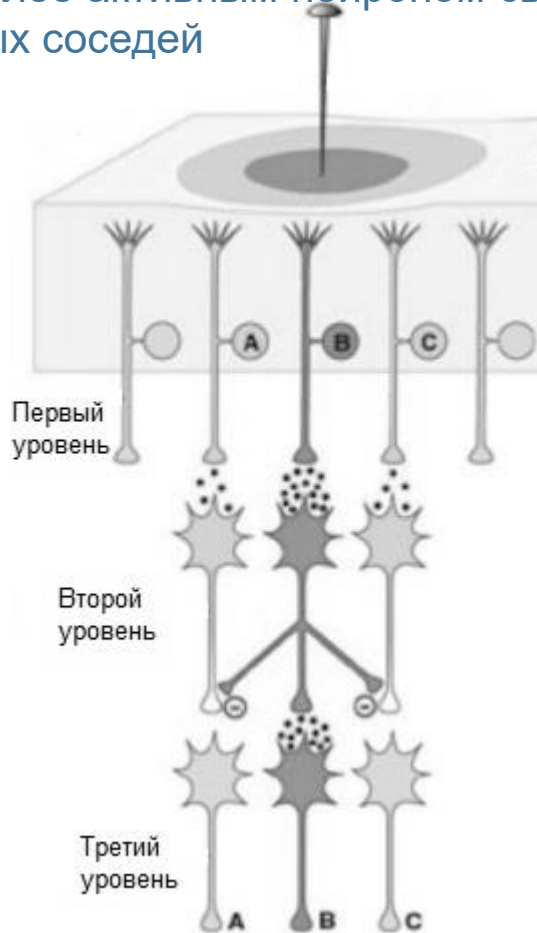
1. Разграничение областей знания с помощью латерального торможения
2. Сохранение «старых» знаний с помощью манипуляций с весами

Импульсная нейронная сеть реализована в пакете SpykeTorch на Python 3.8

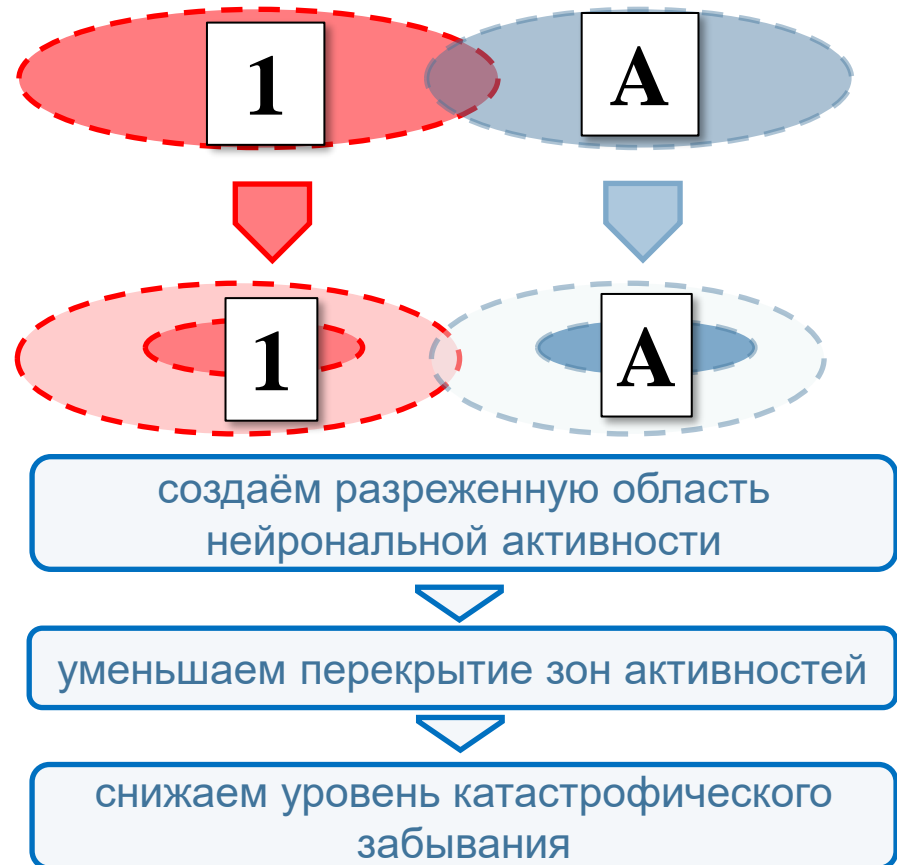


1-ый способ: латеральное торможение

В нейробиологии **латеральное торможение** - процесс подавления наиболее активным нейроном своих слабых соседей



Перекрывание картин активностей нейронов для прежних (цифры) и новых (буквы) данных - одна из причин катастрофического забывания



1-ый способ: исследование латерального торможения

Задача 1 и **Задача 2** выполнялись послойным обучением при различных значениях радиуса латерального торможения в каждом слое

Таблица

№	Радиусы латерального торможения в слоях		Точность распознавания, %		
	R1 (слой S1)	R2 (слой S2)	Задача 1 (цифры)	Задача 2	
				(буквы)	(цифры)
1	0	0	90,8	78,4	48,1
2	3	1	93,6	87,3	73,3
3	5	1	94,8	86,5	72,2
4	5	2	94,8	88,9	78,6

Вывод по исследованию влияния латерального торможения

радиусы латерального торможения

точность классификации набора цифр в Задаче 2

уровень катастрофического забывания

= 0 (в Таблице - №1)		менее 50%			выше
> 0 (в Таблице - № 2, 3, 4)		72% - 79%			ниже

Антонов Д.И., Сухов С.В. Механизмы предотвращения катастрофического забывания в импульсных нейронных сетях / Сборник научных трудов XXIV Международной конференции Нейроинформатика-2022. – 2022. – МФТИ. – С. 55-64.

2-ой способ: как сохранить «старые» знания

Предположение:

найти способ предотвращения катастрофического забывания



найти способ изменения весов, позволяющий сети сочетать одновременно стабильность и пластичность*

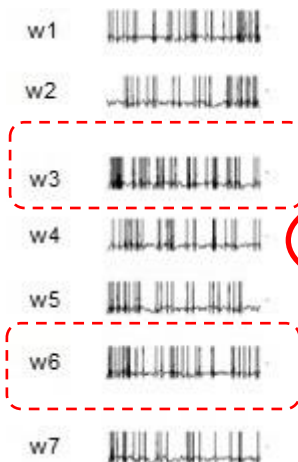
* пластичность \Leftrightarrow изменчивость

2.1. Фиксировать номера (индексы) весов

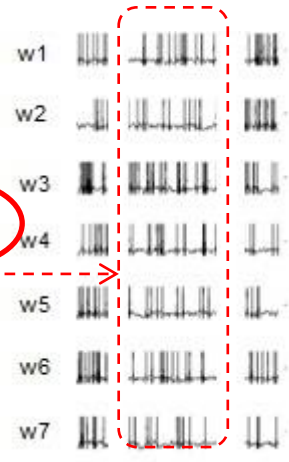
Выбор принципа:
Каким образом будем менять веса при дальнейшем обучении?

2.2. Фиксировать диапазон в каждом весе

Выбор важных весов (индексов) на основе значений: «чем больше, тем важнее»



Варианты выбора весов для изменения при дальнейшем обучении



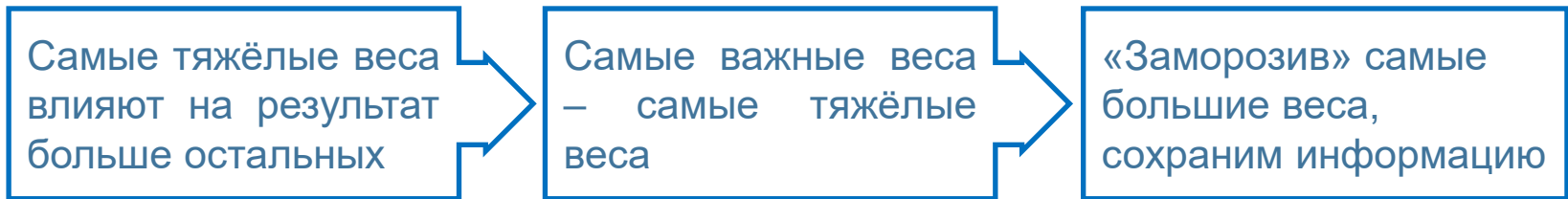
Выбор важных диапазонов весов на основе динамики Ланжевена

Antonov D., Sviatov K., Sukhov S. Continuous learning of spiking networks trained with local rules / Neural Networks. – 2022. – Vol. 155. – P. 512-522.

2-ой способ:

2.1. алгоритм фиксации весов для обучения

Предположение:



Алгоритм:

Этап I. обучение на начальном наборе данных обычным образом

Этап II. отбор подмножества $W_{\max} = \{\text{наибольшие веса после обучения на начальном наборе}\}$

Этап III. «заморозка» подмножества весов W_{\max} , исключая любую их модификацию

Этап IV. обучение сети на новом наборе данных с использованием только «незамороженных» весов

2-ой способ:

2.2. алгоритм выделения диапазона весов для обучения

Алгоритм:

Этап I. обучение на начальном наборе данных обычным образом

Этап II. продолжение обучения на начальном наборе с добавлением гауссова шума:

$$\Delta w_{ij} = F(w_{ij}) + \sigma N(0,1) ,$$

Δw_{ij} - изменение веса (индексы i и j - пост- и пре-синаптические нейроны)

$F(w_{ij})$ определяет величину изменения связи между i и j нейронами

$N(0,1)$ - стандартное нормальное распределение

σ - амплитуда шума

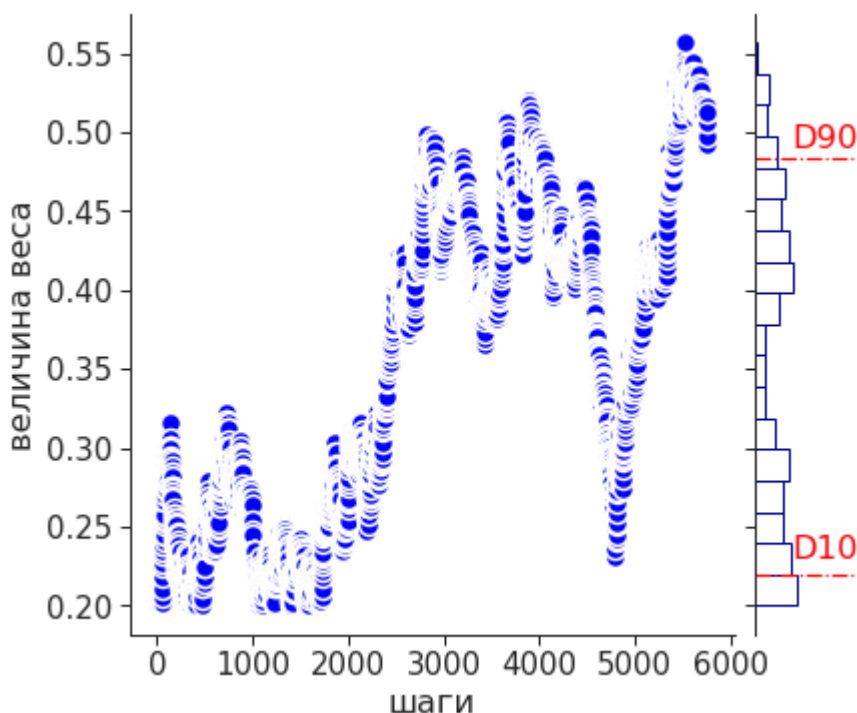
Этап III. определение приемлемых диапазонов изменения значений весов:

для каждого веса строится вероятностное распределение $p(w_{ij})$ и определяются нижняя и верхняя границы для последующего обучения

Этап IV. обучение на новом наборе данных с ограничением диапазона изменения значений каждого веса

2-ой способ:

2.2. исследование влияния диапазона весов



Типичная траектория блуждания веса
(броуновское блуждание)

Гистограмма на правой стороне
графика - распределение вероятности

Результат серии опытов показывает наличие **оптимального диапазона значений весов**, который не позволяет забываться предыдущему знанию и, в то же время, в меньшей степени препятствуют приобретению сетью нового знания

Нижняя и верхняя границы интервала	Точность распознавания Задача 2, %	
	EMNIST	MNIST
Q1 (25%) и Q3 (75%)	66,3	78,7
D10 (10%) и D90 (90%)	78,3	82,2

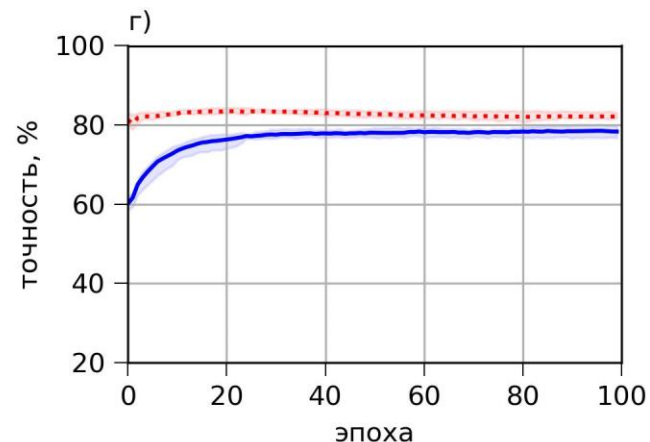
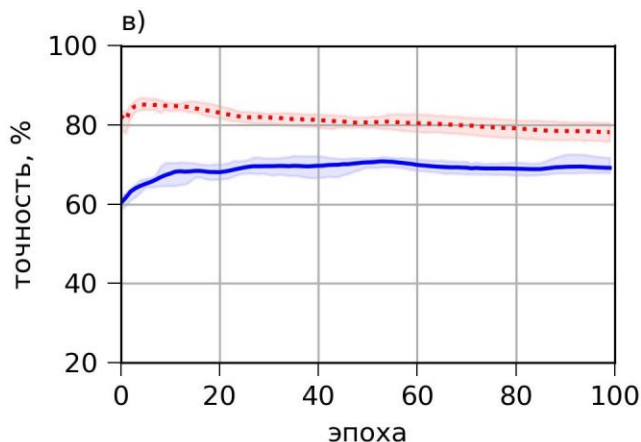
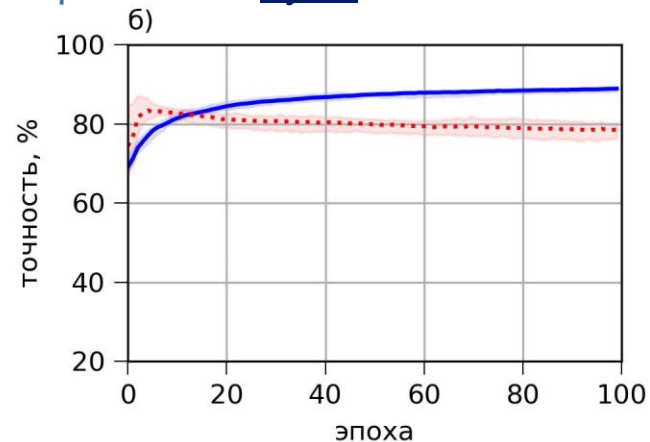
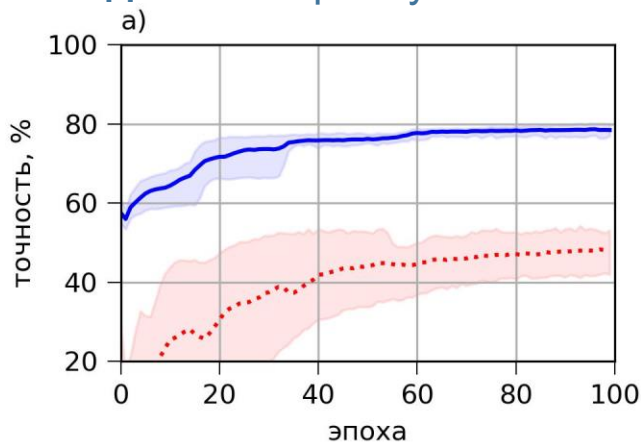
Antonov, D.I. Noise data impact on catastrophic interference of spiking neural networks / AIP Conference Proceedings. - 2022. - Vol. 2466. – P. 070003.

Сравнительные результаты по п.1

Задача 1: обучение на изображениях цифр из набора **MNIST**

Задача 2: переобучение на изображениях букв из **EMNIST**

Точность классификации цифр (красная пунктирная линия) и букв (синяя сплошная линия) в процессе обучения Задаче 2 в случае



(а) катастрофическое забывание

(б) латеральное торможение (**способ 1**)

(в) фиксация самых больших весов (**способ 2.1**)

(г) ограничение диапазона весов (динамика Ланжевена) (**способ 2.2**)

Положение №1. Метод непрерывного обучения ИМНС на основе динамики Ланжевена позволяет снизить потери от катастрофического забывания ИМНС; при обучении ИМНС на данных из баз MNIST и EMNIST в 1,7 раз

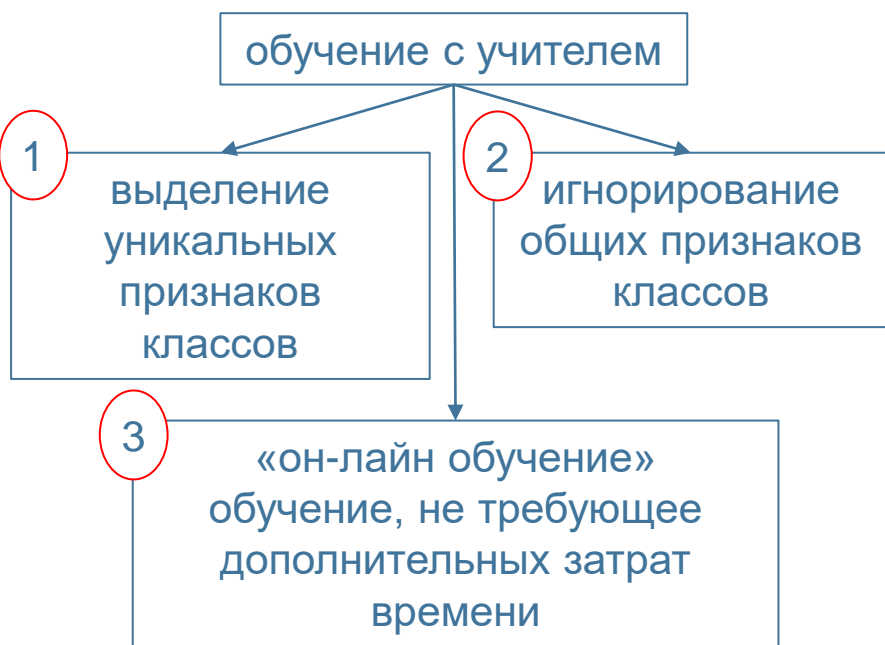
Структура доклада

1. Метод преодоления катастрофического забывания при непрерывном обучении импульсных нейронных сетей
2. Метод локального обучения импульсных нейронных сетей на основе правил STDP и All-LTD
3. Метод регуляризации импульсных нейронных сетей
4. Научные и практические результаты работы

Ограничения для правила обучения с учителем

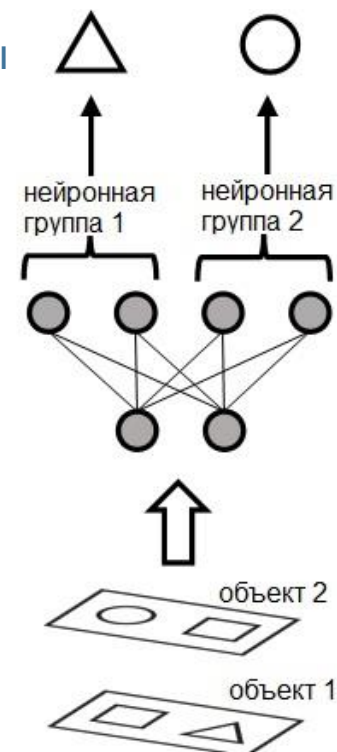
Ограничения:

- **био-подобный метод**
 - прямо применим для обучения импульсной нейронной сети
 - экономный по времени
- **метод для задач классификации изображений**



Изображены объекты двух классов 1 и 2, для которых нейронные группы №1 и №2:

- **выделяют уникальные признаки** (круги и треугольники)
- **игнорируют общие признаки** (квадраты)

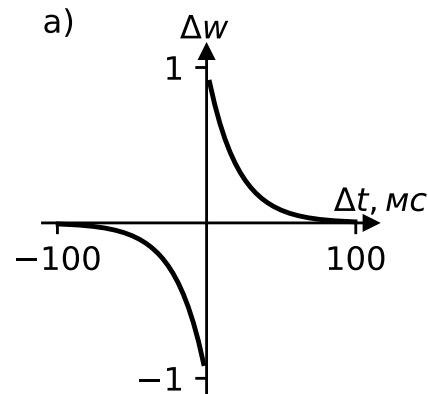


Локальные правила STDP и All-LTD

STDP (spike-timing-dependent plasticity)

правило STDP приводит к кластеризации данных

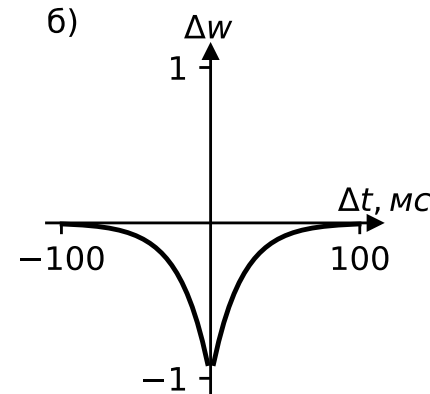
$$\Delta w(\Delta t) = \begin{cases} A_{pre} \cdot \exp(-\Delta t/\tau_{pre}), & \Delta t > 0 \\ A_{post} \cdot \exp(\Delta t/\tau_{post}), & \Delta t < 0 \end{cases}$$



Правило STDP учит нейрон отвечать на определённый признак

All-LTD (all-long-term depression)

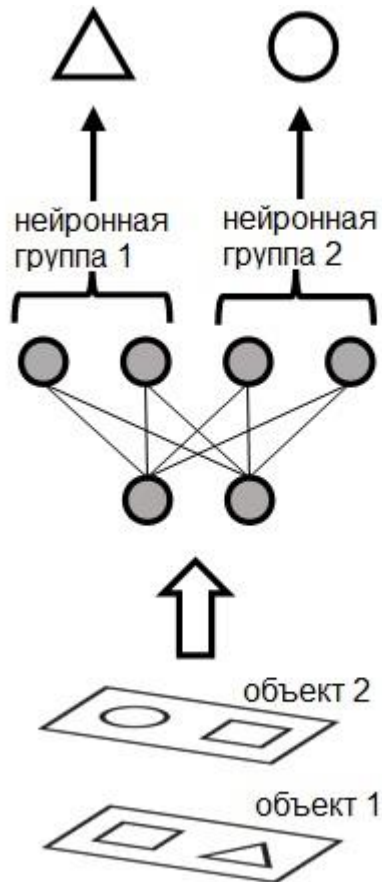
$$\Delta w(\Delta t) = \begin{cases} -A_{pre} \cdot \exp(-\Delta t/\tau_{pre}), & \Delta t > 0 \\ A_{post} \cdot \exp(\Delta t/\tau_{post}), & \Delta t < 0 \end{cases}$$



Правило All-LTD учит нейрон «молчать» в ответ на определённые признаки

Antonov D., Batuev B., Sukhov S. Spiking neural networks training with combined Hebbian rules / X International conference on information technology and nanotechnology (ITNT). — Samara. Russian Federation. — 2024. — P.1-5.

Особенности реализации правила обучения нейронной сети



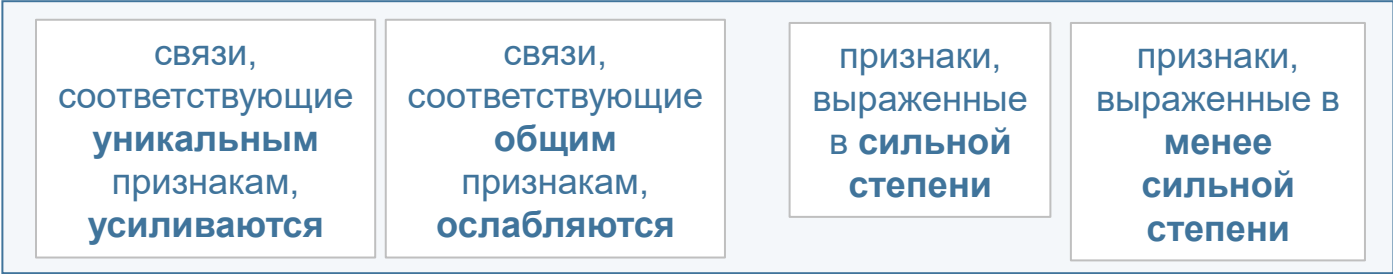
латеральное торможение при
посредстве тормозных
нейронов

адаптивный порог
позволяет выровнять
активность возбуждающих
нейронов на признаки

механизм действия

выделяет разнообразные
признаки классов

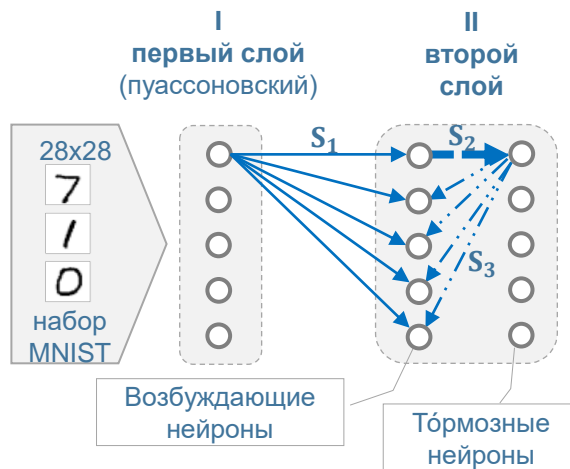
выделяются важные
признаки



Антонов Д.И., Сухов С.В. Обучение импульсных нейронных сетей-классификаторов с помощью локальных правил / XXV Международная конференция Нейроинформатика-2023. – 2023. – Москва, МИФИ.

Архитектура импульсной нейронной сети

ИмНС реализована в пакете Brian 2.0 на Python 3.8



(1) генерация импульсов с частотой = интенсивность пикселей

(2) пуассоновские нейроны \Rightarrow синапсы $S_1 \Rightarrow$ возбуждающие нейроны [связь «один-ко-всем»]

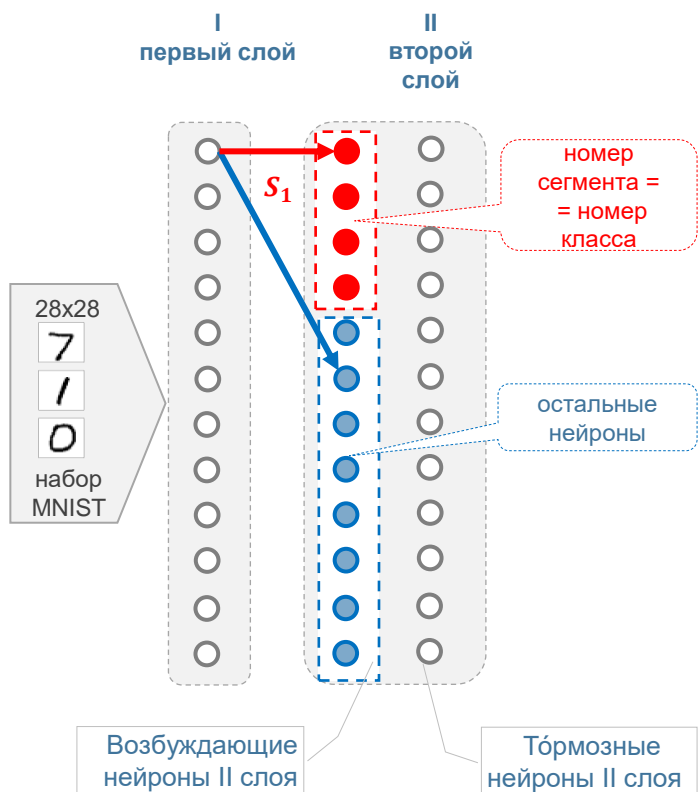
(3) возбуждающие нейроны при достижении порога генерируют импульсы

(4) возбуждающие нейроны \Rightarrow синапсы $S_2 \Rightarrow$ тормозные нейроны [связь «один-к-одному»]

(5) тормозные нейроны \Rightarrow синапсы $S_3 \Rightarrow$ возбуждающие нейроны [связь «один-ко-всем-кроме-иницирующего»]

Программа для ЭВМ. Программа обучения импульсной нейронной сети по правилу STDP + all-LTD / Д.И. Антонов, С.В. Сухов; правообладатель УлГТУ. - заявка 2025616742 от 31.03.2025. - свидетельство № 2025619058. - дата регистрации 14.04.2025.

Схема локального правила обучения 'STDP+AII-LTD'



Класс данных определяется по максимальной суммарной активации сегмента нейронов

При обучении возбуждающие нейроны делятся на сегменты:

- количество сегментов = количество классов
- номер сегмента \leftrightarrow номер класса

пуассоновский нейрон \Rightarrow
 \Rightarrow синапс группы $S_1 \Rightarrow$
 \Rightarrow возбуждающий нейрон

TRUE

возбуждающий нейрон в сегменте с номером = номер класса

FALSE

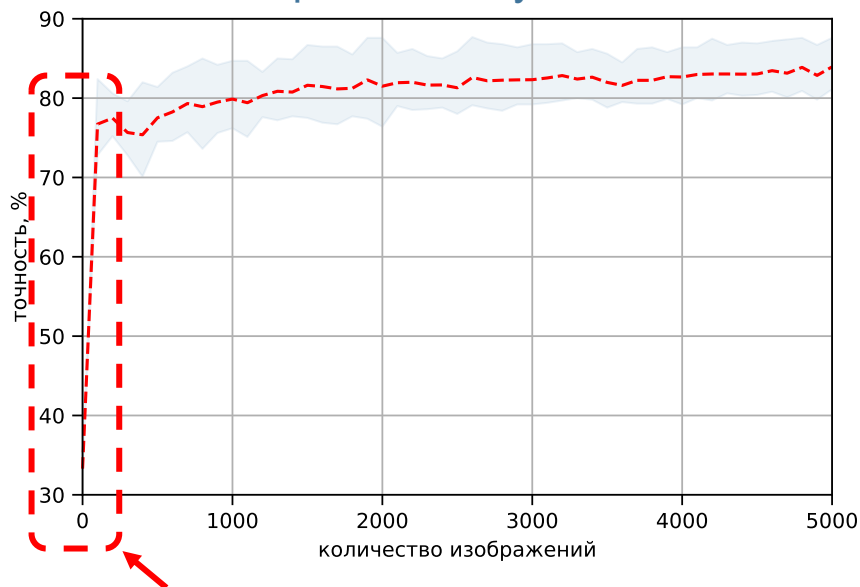
возбуждающий нейрон в сегменте с номером \neq номер класса

STDP правило обучения синапса S_1

AII-LTD правило обучения синапса S_1

Результаты экспериментов

Изменение точности классификации в процессе обучения



В рамках серии экспериментов в процессе обучения ИмНС после обучения на каждых 100 изображениях точность тестировалась на 1000 изображений

В опытах использовались данные из набора **MNIST**:

- для обучения 5000 изображений
- для теста 1000 изображений

Было использовано 3 класса изображений из 10 возможных

Точность $\approx 80\%$ достигается сетью, обученной лишь на 100 изображениях базы MNIST

Точность классификации по итогам серии экспериментов:

- STDP + All-LTD = $74.9 \pm 19.8\%$
- STDP + All-LTD с адаптивным порогом = $80.0 \pm 13.0\%$

Положение №2. Алгоритм обучения ИмНС на основе правил STDP и All-LTD позволяет реализовывать архитектуру ИмНС для обучения с учителем без внешних классификаторов и позволяет сократить размер набора данных для обучения (до 100 изображений при использовании базы данных MNIST) в случае задачи классификации изображений при обучении с учителем

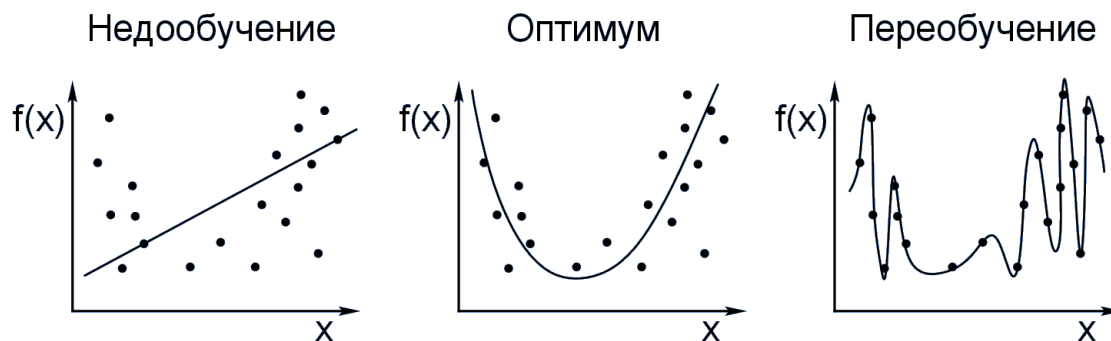
Структура доклада

1. Метод преодоления катастрофического забывания при непрерывном обучении импульсных нейронных сетей
2. Метод локального обучения импульсных нейронных сетей на основе правил STDP и All-LTD
3. Метод регуляризации импульсных нейронных сетей
4. Научные и практические результаты работы

Регуляризация искусственных нейронных сетей (ИНС)

Желаемым результатом любого обучения является **умение обобщать знания (generalize)** для решения новых задач

В процессе обучения ИНС вычисляется и минимизируется функция потерь (Loss function). Но минимизация функции потерь может привести к переобучению модели.



Переобучение (overfitting) — это результат обучения, при котором ИНС начинает учитывать как существенные, так и несущественные признаки, шум

Пути преодоления переобучения ИНС

Наиболее часто используются три метода регуляризации ИНС:

- **L1-регуляризация** (Lasso regression) использует модуль весов в качестве штрафного члена к функции потерь

$$F_{loss} = E(Y - \hat{Y}) + \lambda \sum_{i=0}^n |w_i|$$

- **L2-регуляризация** (Ridge regression) использует квадрат весов в качестве штрафного члена функции потерь

$$F_{loss} = E(Y - \hat{Y}) + \lambda \sum_{i=0}^n w_i^2$$

где

- λ – параметр
- E – функция ошибки, измеряющая разницу между истинной Y и предсказанной \hat{Y} величинами

L2-регуляризация минимизирует веса

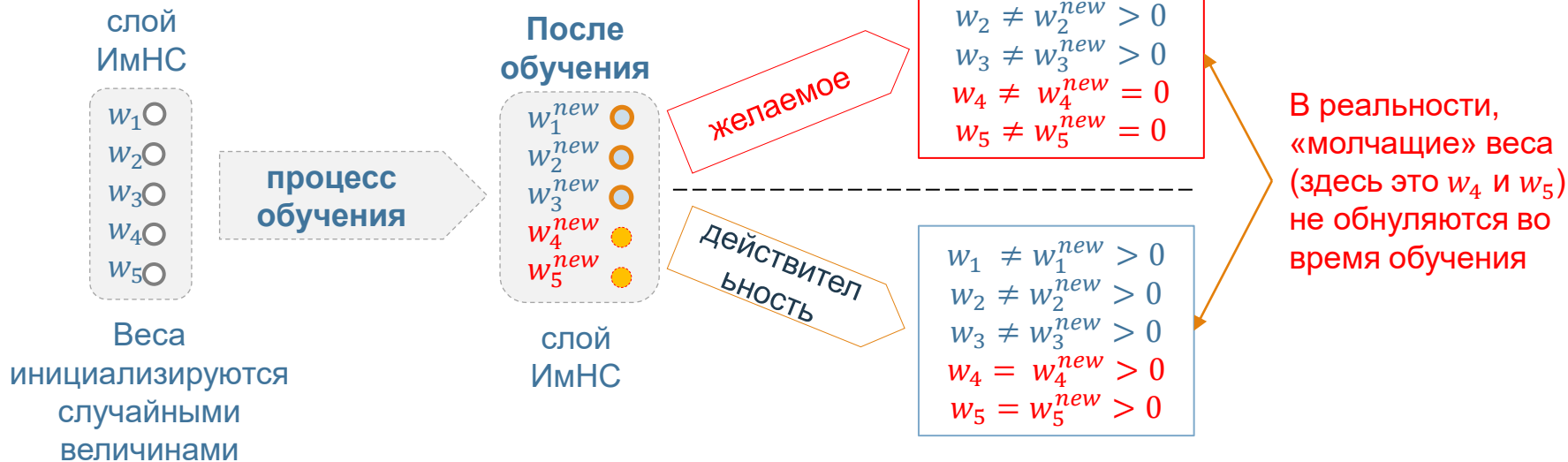
L1-регуляризация делает избыточные веса равными нулю, создавая разреженную модель

- **Дропаут** (dropout) подразумевает случайный выбор весов или целых нейронов для удаления из процесса обучения
Эффект дропаута заключается в том, что сеть становится менее чувствительной к определенным весам нейронов

Что происходит при обучении ИмНС

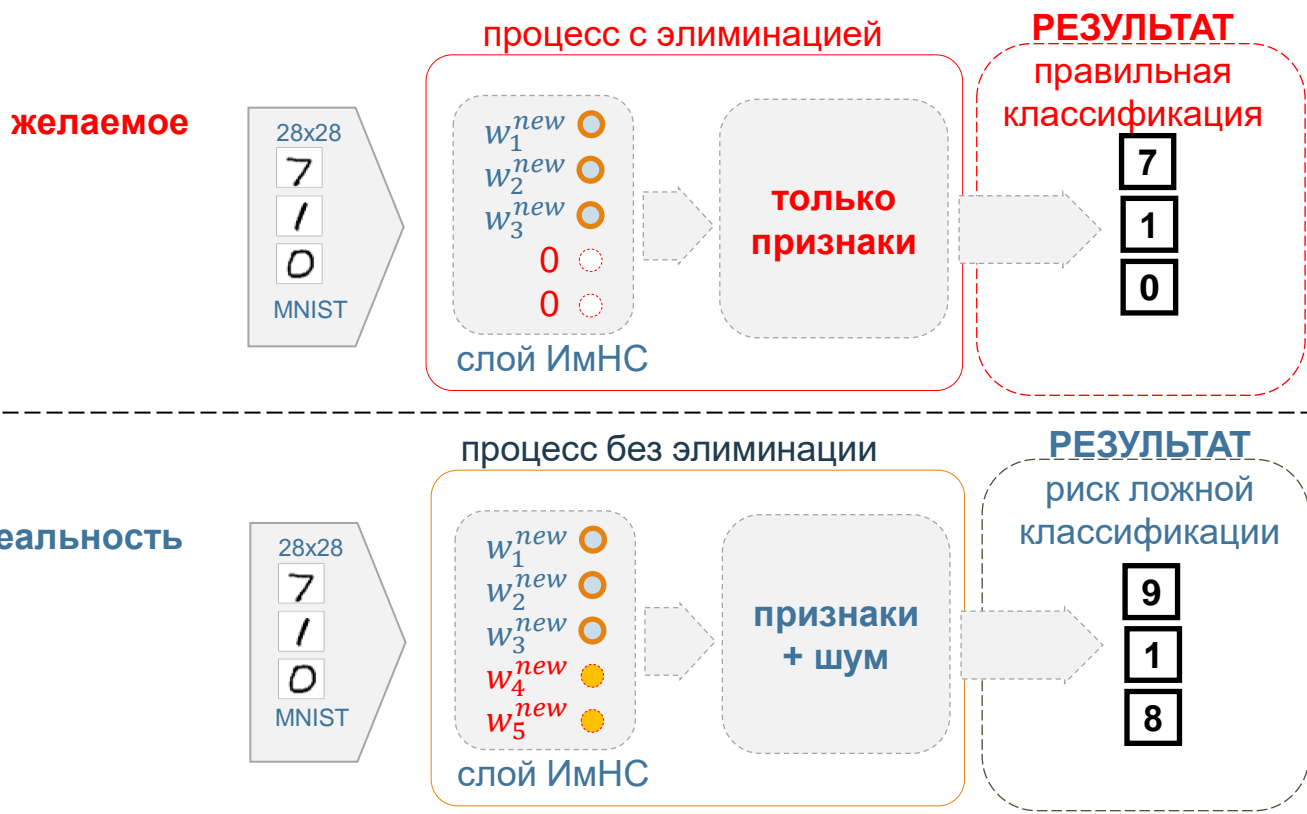
Методы регуляризации ИНС 2-го поколения (дропаут, L1-регуляризация, L2-регуляризация и т. д.) не подходят для ИмНС

Пусть w_1, w_2, w_3 используются при обучении ИмНС, а w_4 и w_5 не используются, тогда:



Antonov D., Sukhov S. Weight regularizations in spiking neural networks / Материалы XXXII Всероссийского семинара «Нейроинформатика, её приложения и анализ данных». - 2024. – Красноярск. – С. 10-17.

Содержание процесса регуляризации ИмНС



Избыточные веса (здесь это $w_4 \neq 0$ и $w_5 \neq 0$) могут быть впоследствии использованы не по назначению: для распространения шума или ненужной информации

Антонов Д.И., Сухов С.В. Регуляризация весов в импульсных нейронных сетях / Информационные и математические технологии в науке и управлении. – 2025. - №1(37). – С. 16-24.

Метод регуляризации для ИмНС «используй или потеряешь»

Веса изменяются только в случае активации пресинаптических нейронов

Веса «молчащих» пресинаптических нейронов остаются такими же, как и при инициализации



Биологическая концепция «используй или потеряешь» ('use it or lose it'):

если синаптическая связь в мозге не используется, она обрезается

Придание весам зависимости от времени позволяет организовать удаление (прунинг, pruning) неиспользуемых весов

$$w(t) = w(t_s) \exp \left[-k_p \left(\frac{t - t_s}{\tau_{prun}} \right)^2 \right]$$

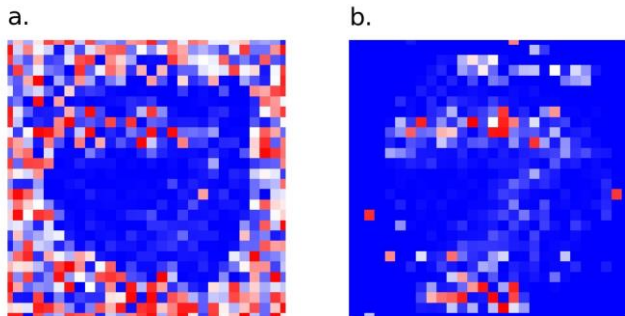
где

- t_s - время последнего импульса
- τ_{prun} - характерное время, в течение которого происходит прунинг
- k_p - коэффициент обрезки (принимает значение 0 или 1)

Оптимальное значение $\tau_{prun} = 7500$ мс, пропорционально времени прохождения одного изображения через сеть и количеству классов данных

Программа для ЭВМ. Программа регуляризации импульсной нейронной сети при обучении / Д.И. Антонов, С.В. Сухов - свидетельство № 2025682750. - дата 27.08.2025.

Тестирование метода регуляризации «используй или потеряешь»



Обученные сети тестировались на 4 различных наборах изображений, размером 28x28 пикселей :

- с дельта-коррелированным шумом
- с шумом, обработанные гауссовым фильтром
- цифры (из **MNIST**), на которых ИмНС обучалась
- цифры (из **MNIST**), на которых ИмНС не обучалась

Изображение 784 весов синаптической группы S_1 , связанной с одним из возбуждающих нейронов:

- a) модель без регуляризации
- b) те же веса в модели с регуляризацией

Активность нейронов (количество импульсов)

№	Набор для тестирования	Модель без регуляризации	Модель с регуляризацией
1	Дельта-коррелированный шум	3162	1
2	Гауссово размытые изображения	3265	1
3	«Известные» цифры	416	344
4	«Неизвестные» цифры	346	240

Положение №3. Метод регуляризации позволяет избегать переобучения ИмНС, снижая влияние случайного шума и несущественных признаков, выявляя важные признаки; реакция ИмНС с регуляризацией на шумовые данные практически отсутствует, реакция ИмНС с регуляризацией на данные неизвестных классов по сравнению с ИмНС без регуляризации снижена в 1,8 раза при использовании данных из базы MNIST

Структура доклада

1. Метод преодоления катастрофического забывания при непрерывном обучении импульсных нейронных сетей
2. Метод локального обучения импульсных нейронных сетей на основе правил STDP и All-LTD
3. Метод регуляризации импульсных нейронных сетей
4. Научные и практические результаты работы

Научные и практические результаты работы

1. Разработан метод непрерывного обучения ИмНС, снижающий потери от катастрофического забывания

Доказано, что **метод непрерывного обучения ИмНС на основе применения динамики Ланжевена позволяет снизить потери от катастрофического забывания ИмНС при обучении сети на новых данных в 1,7 раз**

2. Разработан метод локального обучения ИмНС на основе правил STDP и All-LTD

Доказано, что метод локального обучения 'STDP+All-LTD' не требует дополнительных внешних классификаторов, **для обучения ИмНС до уровня точности 80% на данных базы MNIST достаточно обучающей выборки из 100 изображений**

3. Разработан метод регуляризации для ИмНС

Доказано, что практически отсутствует реакция ИмНС с применением регуляризации на шумовые данные, **реакция ИмНС с применением регуляризации на данные неизвестных классов по сравнению с ИмНС без регуляризации снижена в 1,8 раза**, что повышает точность ИмНС и улучшает её способность к обобщению