

ОПТИМИЗАЦИЯ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ С ПОМОЩЬЮ ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЙ

© 2022 г. Н. А. Вершков^{a,*} (ORCID: 0000-0001-5756-7612),
М. Г. Бабенко^{a,b,**} (ORCID: 0000-0001-7066-0061),
А. Н. Черных^{b,c,d,***} (ORCID: 0000-0001-5029-5212),
В. А. Кучуков^{a,****} (ORCID: 0000-0002-1839-2765), Н. Н. Кучеров^{a,*****} (ORCID: 0000-0003-0337-0093),
Н. Н. Кучукова^{a,*****} (ORCID: 0000-0002-8070-0829)

^a Северо-Кавказский центр математических исследований,
Северо-Кавказский федеральный университет,
355017, г. Ставрополь, ул. Пушкина, 1, Россия

^b Институт системного программирования РАН,
109004, г. Москва, ул. А. Солженицына, д. 25, Россия

^c Центр научных исследований и высшего образования,
22860, Нижняя Калифорния, Энсенада, ш. Тихуана-Энсенада, 3918, Мексика

^d Южно-Уральский государственный университет,
454080, Челябинск, проспект Ленина, 76, Россия

*E-mail: nvershkov@ncfu.ru

**E-mail: mgbabenko@ncfu.ru

***E-mail: chernykh@cicese.mx

****E-mail: vkuchukov@ncfu.ru

*****E-mail: nkuchеров@ncfu.ru

*****E-mail: nkuchukova@ncfu.ru

Поступила в редакцию 01.07.2022 г.

После доработки 16.08.2022 г.

Принята к публикации 22.09.2022 г.

В статье рассматривается вопрос оптимизации искусственных нейронных сетей с точки зрения быстродействия с помощью вейвлет-преобразования. Существующие подходы внедрения вейвлет-преобразования в нейронные сети подразумевают или преобразование до нейронной сети, или с использованием архитектуры “вейвнета”, требующей новых подходов к обучению и применению нейронной сети. Предлагаемый подход опирается на представление модели нейрона в виде неркурсивного адаптивного фильтра и применение вейвлет-фильтра для получения низкочастотной части изображения, снижая тем самым размер изображения и отфильтровывая помехи, имеющие, как правило, высокочастотную природу. Предлагаемая модель вейвлет-преобразования опирается на классическое представление слоя нейронной сети прямого распространения или сверточного слоя. Предлагаемый подход позволяет проектировать нейронные сети с вейвлет-преобразованием на базе существующих библиотек и не требует вносить изменения в алгоритм обучения нейронных сетей. Для подтверждения теоретических положений предлагаемый алгоритм был проверен на трех MNIST-подобных сетях. Для программистов-практиков предлагаемый алгоритм был проверен на реальных изображениях для различения животных и показал аналогичные результаты, что и на MNIST-подобных тестах.

DOI: 10.31857/S0132347422060073

1. ВВЕДЕНИЕ

Одной из самых широко распространенных задач, решаемых искусственными нейронными сетями (ИНС), является распознавание визуальных образов. ИНС способны распознавать рукописные и печатные символы на бумаге, подписи на официальных документах, водяные знаки и пр.

подобные объекты. Применение ИНС позволяет существенно облегчить труд человека, повысить надежность и точность бизнес-процессов за счет снижения человеческого фактора. В общем виде распознавание образов — это отнесение визуального объекта к одному из классов (задача классификации), когда классы заранее определены, или выявление классов по некоторым общим призна-

кам и отнесение объекта к одному из них (задача кластеризации).

Вопросам решения задачи распознавания образов уделяется много внимания и посвящено большое количество литературы [1–4]. Одной из основополагающих работ, посвященных этой теме, считается книга Ахмеда Н. и Рао К.Р. [5] о роли ортогональных преобразований в цифровой обработке сигналов, в т.ч. при распознавании образов. Книга стала обобщением целого ряда работ авторов, посвященных обобщенным ортогональным преобразованиям, функциям Уолша, преобразованиям Хаара, а также методам вычисления спектра в различных базисах. Несмотря на значительный срок, работы Ахмеда Н. и Рао К.Р. до сих пор актуальны и находят применение в современных исследованиях.

Использование ИНС для распознавания образов, несмотря на значительное количество исследований, достаточно ограничено. Большинство авторов рассматривают ИНС в виде обучаемого классификатора, предпочитая формировать вектор признаков и осуществлять отбор признаков стандартными методами [6–9]. Действительно, классическая схема распознавания образов, предложенная в [5], включает в себя ортогональное преобразование, формирующее вектор признаков, алгоритм оптимизации вектора признаков с целью уменьшения размерности и обучаемый классификатор, чаще всего представляющий собой ИНС прямого распространения. Дело в том, что теория ИНС с самого зарождения опирается на представление многомерной нелинейной функции в виде суперпозиции одномерных функций, аргументами которых являются аргументы исходной функции или их комбинация [10–12]. Кроме того, расширение теоремы Колмогорова–Арнольда работами Хехт–Нильсена [13] ограничивает размер скрытого слоя ИНС снизу. В связи с этим постулатом формирования вектора признаков и его оптимизация в ИНС невозможны.

В то же время целый ряд работ [14–16] позволяют исследовать ИНС как информационную систему. Использование математической модели нейрона [17] как адаптивного сумматора (или рекурсивного адаптивного фильтра) позволяют применить основные положения теории информации к теории ИНС [18]. В частности, в работе [19] авторы рассмотрели возможность применения пространственно-темпоральной модели нейрона для реализации возможности обработки нестационарных процессов на ИНС. Рассматриваются также возможности использования преобразования Фурье для фильтрации периодических помех, возникающих на космических снимках [9].

Попытки построения ортогональных нейронных сетей предпринимаются и в настоящее время,

однако, как правило, носят узкоспециализированный характер, например, в медицине [20].

Основным недостатком преобразования Фурье является его низкая информативность во временной области. Обработка нестационарных сигналов требует различные степени разрешения как во временной, так и в частотной областях. Использование вейвлетов позволяет преодолеть ограниченность преобразования Фурье во временной области. Попытки создать гибридную нейронную сеть с использованием вейвлет-преобразования предпринимались многократно для различных направлений. Так в работе [21] вейвлет-преобразование применялось для предварительной обработки временных рядов и были получены результаты, значительно улучшающие качество работы ИНС по сравнению с классической моделью. Однако, в работе не описан метод выполнения преобразования. В работе [22] авторы исследовали возможности по использованию вейвлет-преобразования для распознавания символов. Опираясь на опубликованные результаты, авторы утверждают, что использование вейвлет-преобразования позволяет сократить время обучения ИНС минимум в 3 раза, что, несомненно, является значимым результатом. Причины успеха, которые называют авторы, связаны с физической природой преобразования Хаара и со снижением размерности вектора признаков. К сожалению, в работе не приведена архитектура ИНС, поэтому сложно судить о глубине применения ИНС для решения задачи преобразования Хаара.

Подводя итог вышеприведенному обзору необходимо отметить, что процесс использования ИНС для классификации и кластеризации изображений требует решения 2-х основных задач: максимальная унификация ИНС, т.е. интеграция процессов нахождения вектора признаков в слой ИНС и оптимизация ИНС, т.е. снижения ее размерности, путем оптимизации вектора признаков. При этом решение задачи оптимизации должно происходить с учетом первой задачи, т.е. в рамках архитектуры ИНС.

Появление сверточных ИНС [23, 25, 26] в корне изменило подход к решению задачи распознавания. Резкое снижение количества обучаемых параметров ИНС позволило значительно повысить качество обучения: так для датасета MNIST [32] за менее, чем 20 эпох достигается качество 98% для сверточного слоя 1d. Для слоя 2d скорость и качество обучения еще выше. Сверточная архитектура ИНС включает в себя 3 основных парадигмы: локальное восприятие, разделяемые веса и субдискретизация. За счет этого происходит снижение размерности ИНС, однако, как правило, последние слои ИНС, предназначенные для классификации признаков, остаются слоями прямого распространения и количество связей в

них значительно. Поэтому количество признаков остается критическим параметром.

В работе [18] авторы предложили для формирования вектора признаков дискретное преобразование Фурье (ДПФ) в слое ИНС прямого пространства. Предлагаемый подход опирается на представление ДПФ в виде фильтра. Такое представление позволило использовать модель нейрона МакКаллока–Питтса [17] для формирования частотной характеристики в каждом нейроне первого слоя. При этом слой выступает как линейный частотный преобразователь. Использование вместо ДПФ таких ортогональных преобразований как, например, дискретное косинусное преобразование (ДКП) позволяет снизить количество нейронов в последующих слоях ИНС. Однако отбор значимых признаков осуществлялся с помощью дисперсионного критерия и требовал отдельного исследования в процессе обучения.

Использование Фурье-подобных преобразований осложняется еще тем, что частотное представление сигнала не всегда позволяет выделить лучший вектор признаков по сравнению с временным представлением. Например, когда входной сигнал ИНС не является стационарным (в этом случае шум при представлении одного кластера не является гауссовым), тогда частотный спектр также не будет стационарным в силу влияния каждого временного отсчета на каждую частотную составляющую.

Если же вместо Фурье-подобных преобразований использовать вейвлет-преобразование, то можно получить выигрыш в количестве значимых признаков без проведения дополнительных исследований. Вейвлет-преобразование представляет на выходе аппроксимацию, т.е. низкочастотную часть сигнала и отдельно – высокочастотную. Поскольку шум обычно имеет высокочастотную природу, то данный подход может быть использован для очищения сигналов от шума путем отбрасывания высокочастотной составляющей входного сигнала, выделенной путем вейвлет-преобразования. Применение вейвлетов в ИНС на сегодняшний день не является новостью [24, 25], появился даже термин “вейвнет”, означающий нейроподобную среду с использованием вейвлетов в качестве активационных функций. Однако подобный подход имеет 2 момента, усложняющих обучение ИНС: во-первых, в процессе обучения изменяются не только веса нейронов, но и параметры вейвлетов, а во-вторых, ИНС начинает принимать несвойственную ей пирамидальную архитектуру, которая также скажется на алгоритме обучения. И, наконец, подобный подход к использованию вейвлетов в ИНС потребует создания собственных библиотек под-

держки архитектуры “вейвнетов”, что ставит под сомнение целесообразность данного подхода.

Поэтому авторы в рамках данной статьи поставили себе задачу унификации и оптимизации ИНС при использовании вейвлетов, а также необходимость оставаться в рамках популярных библиотек для построения вейвлет-ориентированных ИНС.

Статья построена следующим образом: во втором разделе рассматривается вейвлет-преобразование как способ снижения размера изображения и формирования вектора признаков оптимального размера и максимальной информативности. В третьем разделе проводится сравнительная характеристика ИНС, не использующей вейвлет-преобразования, с ИНС, которая реализует в первом слое вейвлет-преобразование. При этом анализируется не только быстродействие, но и качество обучения ИНС. В заключение сформулированы результаты применения вейвлет-преобразования, а также определены направления дальнейших исследований.

2. ИСПОЛЬЗОВАНИЕ ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЙ ДЛЯ СЖАТИЯ ВХОДНОГО СИГНАЛА ИНС

Как отмечалось ранее во введении, преобразование Фурье, несмотря на его очень широкое применение, обладает целым рядом недостатков при анализе нестационарных сигналов. К этим недостаткам относятся, в первую очередь, низкая пространственная локализация, т.к. спектр сигнала формируется с учетом всех значений сигнала:

$$S(\omega) = \sum_n s_n(t) u_n(t) \quad (1)$$

Ортогональность базисных функций имеет большое значение [26–28]: однозначность значений коэффициентов разложения по базису существует только в случае, если

$$\int u_m(t) u_n(t) dt = \begin{cases} c, & \text{если } m = n \\ 0, & \text{если } m \neq n \end{cases}$$

В случае, когда $c = 1$ базисные функции ортонормированы. Для улучшения временной локализации используют дополнительную функцию $g(t - t_0)$, которая позволит преобразовать (1) к виду:

$$S(\omega) = \sum_n s_n(t) g_n(t - t_0) u_n(t)$$

Такое преобразование называют оконным преобразованием Фурье. Оконное преобразование частично решает вопрос временной локализации, однако обладает собственными недостатками, к важнейшему из которых относят принцип неопределенности Гейзенберга. В этом

случае говорят не о конкретной частоте, а о диапазоне частот, присутствующих в сигнале в определенном диапазоне времени.

Вейвлет-анализ имеет много общего с классическим ортогональным преобразованием [24, 25], фактически является разновидностью ортогональных преобразований. Как, например, преобразование Фурье в качестве базовой функции использует комплексную экспоненту $e^{i\omega_0 t}$, где $i = \sqrt{-1}$, n – номер гармоники, ω_0 – основная (центральная) частота, t – время, так и вейвлет-преобразование имеет базовую функцию. Например, широко известный на практике вейвлет Хаара описывается функцией вида

$$\psi(t) = \begin{cases} +1, & 0 \leq t < 0.5 \\ -1, & 0.5 \leq t < 1 \\ 0, & t < 0, \quad t \geq 1 \end{cases} \quad (2)$$

Рассматривая преобразование Хаара, необходимо отметить, что половина коэффициентов преобразования Хаара соответствуют корреляции соседних точек в пространстве входных последовательностей, четверть – связям четырех соседних точек и т.д. В отличие от ДПФ и преобразования Уолша–Адамара преобразование Хаара обладает свойством как локальной, так и глобальной чувствительности, благодаря чему именно преобразование Хаара стало первым вейвлет-преобразованием, применяемым на практике.

Базис пространства $L^2(R)$ конструируют из порождающей функции $\psi(t)$, например (2), с помощью двух преобразований независимой переменной: $\psi(t) \Rightarrow \psi(t+k)$ и $\psi(t) \Rightarrow \psi(a^m t)$. Первое преобразование реализует сдвиг порождающей функции по оси времени для покрытия ограниченной функцией $\psi(t)$ (например, функцией Хаара) всей области преобразования, а второе (масштабирование) – для формирования покрытия спектра сигнала. Объединив оба преобразования в одно выражение, получим

$$\psi(t) \Rightarrow \psi(a^m t + k) \quad (3)$$

Подобный подход в теории связи называют частотно-временной матрицей, применяют при синтезе сложных сигналов, он хорошо описан в [30]. Таким образом, произвольная функция в гильбертовом пространстве может быть разложена по базису $\psi_{m,k}(t)$ и представлена в виде ряда

$$s(t) = \sum_{m,k} S_{mk} \psi_{mk}(t) \quad (4)$$

Спектр вейвлет-преобразования является двумерным в пространстве переменных m и k в соответствии с (3). Однако, и одномерный спектр, построенный для определенного значения m , дает информацию, которая может быть использована

для анализа. Так, коэффициент масштабирования a^m влияет на ширину функции по оси времени, а значит, и на полосу частот, занимаемую этой функцией. Небольшие значения коэффициента масштабирования соответствуют узким базисам во временной области и, соответственно, более высоким частотам и, наоборот, рост коэффициента масштабирования приводит к более “длинным” базисам и низким частотам. Таким образом, каждая строка спектра, проведенная параллельно оси времени, дает представление об активности исследуемого процесса в определенной полосе частот на определенном промежутке времени и, следовательно, обладает чувствительностью как во временной, так и в частотной области. Рассмотрим теперь подробнее реализацию вейвлет-преобразования в ИНС как прямого пространства, так и сверточных.

Поскольку вейвлет-преобразование является разновидностью ортогональных преобразований и определяется выражением (4), то по аналогии с [18] для ИНС прямого распространения можно воспользоваться моделью МакКаллока–Питтса [17]:

$$y_{k,l} = f \left(\sum_{i=1}^n w_i^{k,l} x_i^{k,l} \right), \quad (5)$$

где l, k – номер слоя и номер нейрона в слое, $y_{k,l}$ – выход нейрона, $x_i^{k,l}$ – входы нейрона, $w_i^{k,l}$ – веса (синапсы) входных сигналов, f – выходная функция нейрона. Сравнив выражения (1) и (5) можно убедиться в том, что значение под функцией полностью идентично одной строке спектра вейвлета, а, следовательно, может быть получен набор коэффициентов вейвлет-преобразования в первом приближении. Для этого необходимо использовать слой нейронов, в котором в качестве весов занесены значения вейвлет-функции $\psi(a^m t + kl_\psi)$, где $a, m = \text{const}$, k – номер нейрона в слое, а l_ψ – длина вейвлета. Все остальные значения весов должны быть обнулены, а также должен быть введен запрет на изменение весов в этом слое. Этот способ апробирован при вычислении ДПФ, ДКП и т.п. [18].

Предложенный выше способ дает стабильный результат, но обладает значительным недостатком: для получения одной строки вейвлет-спектра используется полноценный слой с размерностью равной размерности сигнала (более подробно это будет показано в следующем разделе). При этом большая часть весов будет обнулена, т.к. размерность вейвлета по сравнению с размерностью входного сигнала обычно невелика. Можно, конечно, отказаться от нейронов, у которых веса обнулены, но в этом случае придется “расплетать” входной сигнал, т.е. перенаправлять соот-

ветствующие разряды входного сигнала на соответствующие входы нейронов.

Для снижения количества используемых связей можно использовать сверточный слой. Процедура вейвлет-преобразования может быть описана как операция пропускания входного сигнала через полуполосный цифровой фильтр с частотной характеристикой $h(n)$ (высокочастотный фильтр) или $g(n)$ (низкочастотный фильтр):

$$\begin{cases} x(n) * h(n) = \sum_k x(k) h(n-k) \\ x(n) * g(n) = \sum_k x(k) g(n-k) \end{cases} \quad (6)$$

Если сигнал на входе ИНС представляет собой одномерную последовательность длиной n , то, используя одномерный сверточный слой с ядром $h(n)$ или $g(n)$, на выходе получим коэффициенты вейвлет-преобразования. Чтобы сократить количество слоев ИНС, можно использовать один слой с двумя (или более) различными ядрами. Для этого создается сверточный слой с одним входом, двумя выходами и с шагом, равным размерности вейвлета. В этом случае будет создан сверточный слой с двумя ядрами, в которые заносятся значения $h(n)$ и $g(n)$.

Для реализации многомерных вейвлет-преобразований придется задействовать количество слоев, равное количеству измерений. На примере двумерного вейвлет-преобразования это понять несложно. Пусть входное изображение представлено матрицей пикселей $l \times l$. На первом этапе осуществляется построчное вейвлет-преобразование, например, вида (6). Затем осуществляется транспонирование полученного результата и повторное вейвлет-преобразование. К сожалению, объединить эти этапы не представляется возможным, т.к. ко второму этапу преобразования можно перейти только после полного выполнения вейвлет-преобразования всех строк. При увеличении размерности входного сигнала пропорционально растет и количество слоев вейвлет-преобразования.

Таким образом, используя слои прямого пространства или сверточные слои, предоставляемые стандартными библиотеками, например, Keras, PyTorch и т.п., можно осуществлять выполнение операции вейвлет-преобразования. 1d вейвлет-преобразование фактически разделяет спектр сигнала на 2 части: низкочастотную и высокочастотную. Для целей классификации (кластеризации) будем использовать низкочастотную часть, т.к. основная часть шума имеет высокочастотную природу и, как показывает практика, не оказывает значительного влияния на качество классификации. Хотя, при необходимости, могут быть задействованы и обе составляющие [21] или высо-

кочастотная часть, например, для определения границ.

3. ПРАКТИЧЕСКИЕ АСПЕКТЫ ПРИМЕНЕНИЯ ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЯ В ИНС

Для практической оценки возможности оптимизации ИНС с помощью вейвлет-преобразования была использована несложная сверточная сеть с двумя сверточными слоями и двумя слоями прямого распространения для реализации обучающего классификатора [31]. ИНС реализована с помощью стандартной библиотеки PyTorch [36]. В качестве данных для обучения ИНС и контроля обучения использовались 3 базы данных: MNIST [32], KMNIST [33] и Fashion-MNIST [34]. MNIST – база образцов рукописного написания арабских цифр. MNIST является стандартом, предложенным Национальным институтом стандартов и технологий США с целью калибровки и сопоставления методов распознавания изображений с помощью машинного обучения в первую очередь на основе ИНС. KMNIST (Kuzushiji-MNIST) является заменой набора данных MNIST, представленного в оригинальном формате MNIST. Поскольку MNIST ограничена 10 классами, то выбран один символ для представления каждого из 10 рядов Хираганы при создании Kuzushiji-MNIST. KMNIST создан Центром открытых данных в гуманитарных науках ROIS-DS (CODH). Fashion-MNIST – это также альтернатива MNIST. Он создан немецкой компанией Zalando и охватывает в общей сложности 70 000 различных изображений из 10 категорий. Подразделение Fashion-MNIST по размеру, формату, комплекту обучения и тестирования точно такое же, как и исходный MNIST. Использование трех различных баз, использующих одинаковый размер и организацию данных, сформулировано несколько пренебрежительным отношением многих программистов-практиков к базе MNIST и желанием получить более надежные гарантии работоспособности предлагаемого метода. Применение же MNIST-подобных баз данных позволило получить сравнительные данные без изменения структуры ИНС, т.е. сохранены условия для качественного сравнения результатов тестирования.

Для оценки результатов обучения использовались результаты контрольных замеров на тестовых данных. Максимальное количество эпох, установленное для обучения, равно 20. В испытаниях принимали участие реализации вейвлет-преобразования с помощью линейных слоев ИНС и с помощью сверточных слоев.

Контрольные замеры были выполнены на всех 3 базах данных без использования вейвлет-преобразования, результаты представлены на рис. 1.

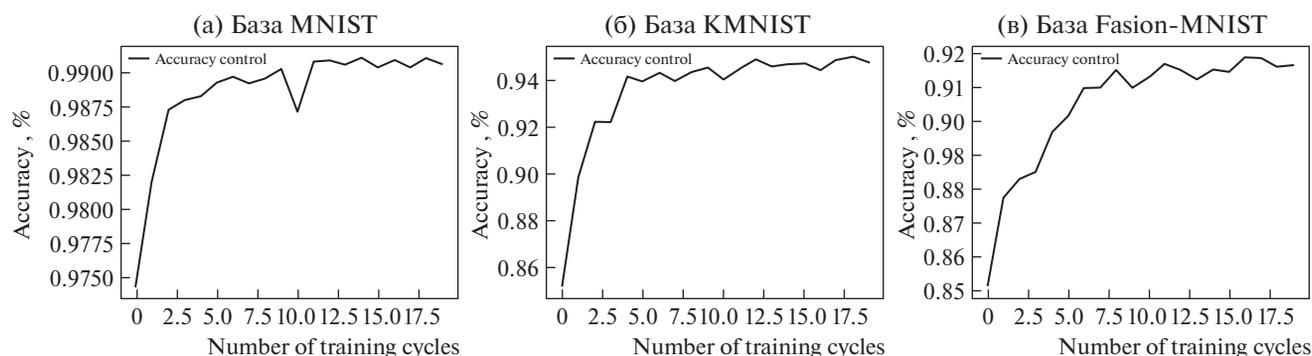


Рис. 1. Результаты обучения ИНС без вейвлет-преобразования.

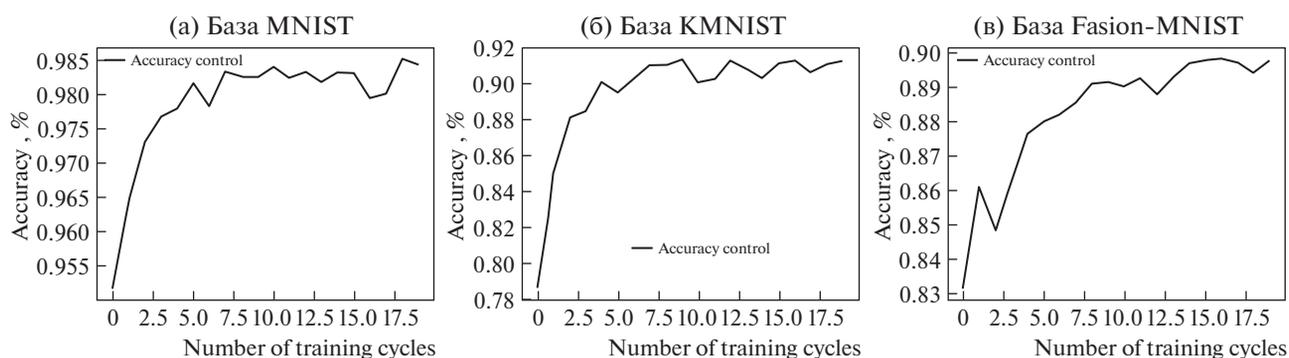


Рис. 2. Результаты обучения ИНС с 1 линейным слоем вейвлет-преобразования.

В результате обучения были получены следующие значения: обучение на базе MNIST дало максимальное значение качества обучения 99.1%, на базе KMNIST – 95.01%, на базе Fashion-MNIST – 91.87%. При этом производились замеры времени, которое было затрачено на 1 цикл обучения, данные об этих измерениях представлены в табл. 1.

Далее на вход ИНС был добавлен один линейный слой размером 784 нейрона (28×28 – размеры входного изображения). В качестве весов для каждого нейрона были установлены значения низкочастотного фильтра вейвлета Хаара (он же вейвлет Добеши 1 порядка). Причем в каждом следующем нейроне значения сдвигались на величину $s \times l$, где s – номер нейрона, а l – длина низкочастотного фильтра (для вейвлета Хаара равна 2). Результаты обучения с использованием линейного слоя вейвлет-преобразования пред-

ставлены на рис. 2, а временные затраты – в таблице 2.

Получены следующие результаты обучения: для базы MNIST максимальное значение качества обучения составило 98.51%, на 0.59% ниже, чем без вейвлет-преобразования, для базы KMNIST – 91.37%, потеря качества обучения составила 3.64%, а для базы Fashion-MNIST – 89.81, потеря качества – 2.06%.

Если рассчитать в процентах, то выигрыш от использования вейвлет-преобразования по среднему времени обучения для базы MNIST составил 46.6% при потере качества обучения на 0.59%, для базы KMNIST – 54.4% при потере качества на 3.64%, для базы Fashion-MNIST – 50.6% при потере качества 2.06%. Таким образом, применение даже одного линейного низкочастотного фильтра вейвлет-преобразования позволяет получить зна-

Таблица 1. Затраты времени на 1 цикл обучения, мс

База	Макс. время	Мин. время	Среднее время
MNIST	0.2884	0.0758	0.1178
KMNIST	0.2015	0.0896	0.1286
Fashion-MNIST	0.3241	0.0948	0.1550

Таблица 2. Затраты времени на 1 цикл обучения с 1 линейным слоем вейвлет-преобразования, мс

База	Макс. время	Мин. время	Среднее время
MNIST	0.2067	0.0398	0.0629
KMNIST	0.1494	0.0369	0.0586
Fashion-MNIST	0.1254	0.0518	0.0766

Таблица 3. Затраты времени на 1 цикл обучения с 1 сверточным слоем вейвлет-преобразования, мс

База	Макс. время	Мин. время	Среднее время
MNIST	0.1498	0.0409	0.0617
KMNIST	0.1199	0.0389	0.0571
Fashion-MNIST	0.1167	0.0389	0.0597

чительный выигрыш ($\approx 50\%$) в затратах времени при незначительной потере качества обучения.

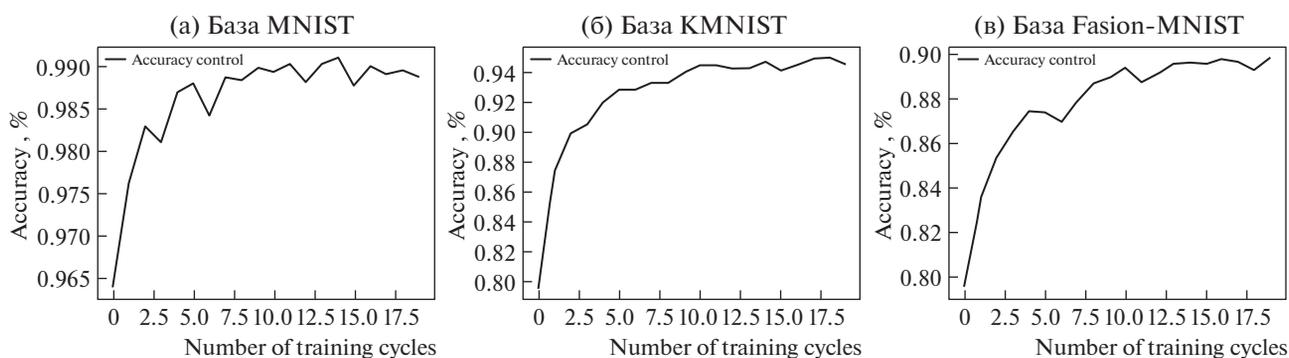
Проведем сравнение характеристик исходной ИНС с нейронной сетью с вейвлет-преобразованием Хаара в первом сверточном слое. Для этого создадим сверточный слой $nn.Conv1d(1, 2, kernel_size = dec_len, stride = dec_len, bias = False)$, где 1 – количество входов, 2 – количество выходов, $kernel_size$ – размер ядра (для вейвлета Хаара 2), $stride$ – шаг свертки, dec_len – длина фильтра вейвлета, полученная с помощью библиотеки PyWavelets, сверточный слой построен на базе библиотеки PyTorch [36] на основе модуля nn . В ядро преобразования занесем значение вейвлет-фильтра и также, как и для линейного слоя ранее, запретим изменение весов слоя. Результаты обучения приведены на рис. 3, а временные показатели – в таблице 3.

Сравнивая таблицы 2 и 3 нетрудно заметить, что результаты по затратам времени очень близки, небольшой выигрыш заметен только для базы Fashion-MNIST. Тем не менее, предпочтительнее использовать сверточный слой. Выигрыш заключается в том, что линейный слой выполняет только один вид вейвлет-преобразования – высоко-

частотное или низкочастотное. Для получения одновременно обоих результатов потребуется 2 линейных слоя ИНС. Сверточный слой формирует ядра таким образом, что преобразование от 1 входа к 2 выходам осуществляется обособленно двумя ядрами, куда и можно занести значения вейвлет-фильтра.

И, наконец, учитывая недоверчивое отношение программистов-практиков к экспериментам, выполненным с использованием баз данных MNIST, был проведен эксперимент по распознаванию вида животного по фотографии. Для этого был использован набор изображений собак и кошек с сайта Microsoft [35]. Для эксперимента была создана сверточная сеть на базе библиотеки PyTorch [36], содержащая 3 сверточных слоя и 2 слоя прямого распространения с двумя выходами. Результаты ее обучения с использованием вейвлета и без него на протяжении 10 эпох представлены на рис. 4, а затраты времени – в таблице 4.

Из рис. 4 видно, что модель ИНС с использованием вейвлета проигрывает в точности распознавания на 2.9% (76.1% против 79%), но выигрывает более, чем на 50% по среднему времени цикла обучения (табл. 4). Т.е. статистика, полученная

**Рис. 3.** Результаты обучения ИНС с 1 сверточным слоем вейвлет-преобразования.

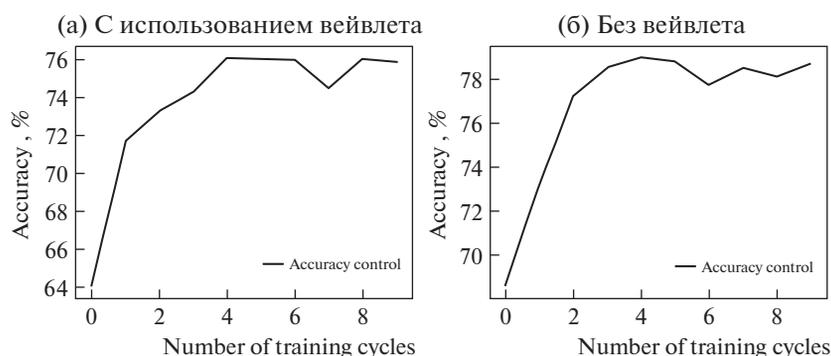


Рис. 4. Результаты обучения ИНС с 1 сверточным слоем вейвлет-преобразования.

на базах MNIST, подтверждается и на базах с реальными изображениями в тех же пропорциях: потеря 2-3% в точности распознавания при приблизительно двукратном выигрыше в скорости обучения.

Таким образом, вейвлет-преобразование в сверточном слое осуществляется одновременно и независимо, не требует формирования “вейвнетов” пирамидальной структуры. Кроме того, в качестве ядра преобразования может быть использован фильтр любого масштаба преобразования, а не только последовательно (например, в первом слое может быть использован фильтр Добеши сразу 2, 3 и более высокого порядка без использования первого). Следовательно, вейвлет-преобразование может быть интегрировано в архитектуру ИНС без дополнительных библиотек и доработок. Использование вейвлет-преобразования позволяет формировать удачный вектор признаков, чувствительный как к временным, так и частотным параметрам, а использование одной его части (низкочастотной) позволяет осуществлять оптимизацию объема вектора признаков (снижение в 2 раза) при незначительной потере качества распознавания и без необходимости проведения дополнительной оптимизации, как, например, с использованием дисперсионного критерия.

4. ЗАКЛЮЧЕНИЕ

Проведенные исследования показали эффективность использования вейвлет-преобразования для оптимизации работы ИНС. Вейвлет-преобразование гораздо более эффективно для формирова-

ния и снижения размерности вектора признаков, чем Фурье-подобные преобразования, т.к. обладает чувствительностью как во временной, так и в частотной области. Использование слоев ИНС, созданных стандартными библиотеками, для вейвлет-преобразования позволяет встроить процедуру формирования и оптимизации вектора признаков в архитектуру ИНС без создания дополнительных библиотек и проведения дополнительных исследований по отбору признаков. Результаты исследования, описанные в 3 разделе статьи показывают, что использование вейвлет-преобразования Хаара позволяет снизить временные затраты вдвое (пропорционально уменьшению размерности вектора признаков) при потере качества обучения 0.5–3%, что является хорошим показателем. Подобный подход, несомненно, будет востребован в тех случаях, когда необходимо осуществлять обработку больших объемов изображений (например, видео), а также при обработке изображений, имеющих априори высокую избыточность (высокое разрешение – HD, Full HD).

Для полноценного использования вейвлет-преобразований в ИНС необходимо провести еще ряд исследований. В первую очередь, это – построение полноценного двумерного спектра вейвлет-преобразования и его обработка средствами стандартных библиотек для проектирования ИНС. Это позволит полностью встроить процедуру вейвлет-преобразования в архитектуру ИНС и получить более широкий спектр возможностей по обработке и распознаванию изображений. Кроме того, значительный интерес представляет вейвлет-обработка цветных изображений, т.е. параллельная обработка 3-х слоев одновременно. Возможно, что здесь заложены значительные возможности по снижению временных затрат ИНС.

5. ФИНАНСИРОВАНИЕ

Работа выполнена при поддержке Российского научного фонда, проект № 19-71-10033.

Таблица 4. Затраты времени на 1 цикл обучения со слоем вейвлет-преобразования и без него, мс

	Макс. время	Мин. время	Среднее время
С вейвлетом	0.3351	0.0558	0.1066
Без вейвлета	0.343	0.1047	0.2259

СПИСОК ЛИТЕРАТУРЫ

1. Методы компьютерной обработки изображений / Под ред. В.А. Соифера. 2-е изд., испр. М.: ФИЗМАТЛИТ, 2003 г.
2. Компьютерное зрение. Современный подход: Пер. с англ. М.: Издательский дом “Вильямс”, 2004 г.
3. Шапиро Л., Стокман Дж. Компьютерное зрение: Пер. с англ. М.: Бином. Лаборатория знаний, 2006 г.
4. Hastie T., Tibshirani R., Friedman J. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. New York: Springer, 2009. Т. 2. С. 1–758.
5. Ахмед Н., Пао К.Р. Ортогональные преобразования при обработке цифровых сигналов: Пер. с англ. / Под ред. И.Б. Фоменко. М.: Связь, 1980 г.
6. Ranzato M.A., Poultney C., Chopra S., Sun Y. Efficient learning of sparse representations with an energy-based model // Advances in neural information processing systems. 2006. Т. 19.
7. Галушкин А.И., Томашевич Д.С., Томашевич Н.С. Методы реализации инвариантности к аффинным преобразованиям двумерных изображений // Приложение к журналу “Информационные технологии”. 2001. № 1. С. 1–19.
8. Филлипс П.Дж., Мартин Э., Уилсон С.Л., Пржибоски М. Введение в оценку биометрических систем // Открытые системы. 2000. № 3. С. 21–27.
9. Гусев В.Ю., Крапивенко А.В. Методика фильтрации периодических помех цифровых изображений // Труды МАИ. Радиотехника. Электроника. Телекоммуникационные системы. 2012. № 50.
10. Колмогоров А.Н. О представлении непрерывных функций нескольких переменных в виде суперпозиций непрерывных функций одного переменного и сложения. Докл. АН СССР. 1957. Т. 114. № 5. С. 953–956.
11. Арнольд В.И. О представлении функций нескольких переменных в виде суперпозиции функций меньшего числа переменных // Мат. просвещение. 1958. Вып. 3. С. 41–61.
12. Горбань А.Н. Обобщенная аппроксимационная теорема и вычислительные возможности нейронных сетей. Сибирский журнал вычислительной математики. 1998. Т. 1. № 1. С. 12–24.
13. Hecht-Nielsen R. Neurocomputing. Addison-Wesley, 1989. 433 p.
14. Widrow B. Adaptive sampled-data systems, a statistical theory of adaptation. IRE WESCON Convention Record, 1959. V. 4. P. 74–85.
15. Цыпкин Я.З. Информационная теория идентификации. М., Наука. Физматлит, 1995 г.
16. Харкевич А.А. Теория информации. Оpozнание образов. Избранные труды в трех томах. Т. III. М., “Наука”, 1973 г.
17. Мак-Каллок У., Пяттс У. Логическое исчисление идей, относящихся к нервной активности // Автоматы. М.: Изд. иностр. лит., 1956. С. 362–384.
18. Vershkov N.A., Kuchukov V.A., Kuchukova N.N., Babenko M. The Wave Model of Artificial Neural Network. In Proc. of the 2020 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering, EIConRus. 2020. P. 542–547.
19. Мальхина Г.Ф., Меркушева А.В. Метод контроля состояния подсистемы (объекта) при неполной измерительной информации о совокупности параметров, определяющих ее динамику. Научное приборостроение. 2004. Т. 14. № 1. С. 72–84.
20. Kim J.S., Cho Y., Lim T.H. Prediction of locations in medical images using orthogonal neural networks // European Journal of Radiology Open. 2021. Т. 8. С. 100388.
21. Jamal A., Ashour M.A.H., Helmi R.A.A., Fong S.L. A Wavelet-Neural Networks Model for Time Series // 2021 IEEE 11th IEEE Symposium on Computer Applications & Industrial Electronics (ISCAIE). IEEE, 2021. С. 325–330. <https://doi.org/10.1109/ISCAIE51753.2021.9431777>
22. Хаустов П.А., Григорьев Д.С., Спицын В.Г. Разработка системы оптического распознавания символов на основе совместного применения вероятностной нейронной сети и вейвлет-преобразования // Известия Томского политехнического университета. Инжиниринг георесурсов. 2013. Т. 323. № 5. С. 101–105.
23. LeCun Y., Bengio Y. Convolutional networks for images, speech, and time series // The handbook of brain theory and neural networks. 1995. Т. 3361. № 10. С. 1995.
24. Fujieda S., Takayama K., Hachisuka T. Wavelet convolutional neural networks // arXiv preprint arXiv:1805.08620. 2018.
25. Veitch D. Wavelet Neural Networks and their application in the study of dynamical system. Department of Mathematics University of York. 2005
26. Нагорнов О.В., Никитаев В.Г., Простокишин В.М., Тюфлин С.А., Проничев А.Н., Бухарова С.А., Чистов К.С., Кашафутдинов Р.З., Хоркин В.А. Вейвлет анализ в примерах. М.: НИЯУ МИФИ, 2010.
27. Kerenidis I., Landman J., Mathur N. Classical and quantum algorithms for orthogonal neural networks // arXiv preprint arXiv:2106.07198. 2021.
28. Wang J., Chen Y., Chakraborty R., Yu S.X. Orthogonal convolutional neural networks // Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2020. С. 11505–11515. <https://doi.org/10.1109/CVPR42600.2020.01152>
29. Смоленцев Н.К. Основы теории вейвлетов. Вейвлеты в МАТЛАБ. М.: ДМК Пресс, 2019.
30. Сикарев А.А., Лебедев О.Н. Микроэлектронные устройства формирования и обработки сложных сигналов. М.: Изд-во “Радио и связь”, 1983
31. Koehler G. MNIST Handwritten Digit Recognition in PyTorch. Nextjournal, 2020, <https://nextjournal.com/gkoehler/pytorch-mnist> Accessed 28.02.2022.
32. Qiao, Yu THE MNIST DATABASE of handwritten digits (2007). Accessed 04.03.2022.
33. GitHub – rois-codh/kmnist: Repository for Kuzushiji-MNIST, Kuzushiji-49, and Kuzushiji-Kanji Accessed 04.03.2022.
34. A MNIST-like fashion product database. Benchmark <https://github.com/zalandoresearch/fashion-mnist>. Дата обращения 04 марта 2022.
35. Cats and Dogs Dataset https://download.microsoft.com/download/3/E/1/3E1C3F21-ECDB-4869-8368-6DEBA77B919F/kagglecatsanddogs_3367a.zip Дата обращения 04 марта 2022.
36. PyTorch (англ.) <https://pytorch.org/get-started/previous-versions/> Дата обращения 04 марта 2022.