Быстрые пирамидальные нейронные сети глубокого обучения для адаптивной цифровой обработки сигналов

А. Ю. Дорогов

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина)

vaksa2006@yandex.ru

***Аннотация.* В работе рассматривается класс быстрых перестраиваемых преобразований, предназначенных для адаптивной цифровой обработки сигналов. Перестраиваемые преобразования интерпретируются как нейронные сети с линейными функциями активации. Показано, что предложенная структура пирамидальной нейронной сети является вариантом структуры быстрого преобразования Фурье. Предложен абсолютно-сходящийся алгоритм быстрого обучения пирамидальной сети к произвольному набору функций. Доказано, что быстрая пирамидальная сеть является нейронной сетью глубокого обучения.**

***Ключевые слова:* быстрое перестраиваемое преобразование; нейронная сеть; пирамидальная структура; глубокое обучение; адаптивная цифровая обработка сигналов**

I. Введение

В работе [1] была предложена архитектура пирамидальной нейронной сети (PyraNet) предназначенной для визуального распознавания образов. Пирамидальная сеть PyraNet имеет иерархическую многослойную структуру, обрабатывающую 2D образы. Выходами сети является 1D слой, представляющий категории входных образов. Пирамидальная нейронная сеть для своего обучения использует тот же арсенал методов, что и классические многослойные нейронные сети. Эти методы обучения основаны на алгоритмах обратного распространения ошибок. Общий недостаток этих методов – не гарантируемая сходимость и большие временные затраты особенно при обучении глубоких нейронных сетей [2].

В данной работе будут рассмотрены пирамидальные с сети с регулярной самоподобной структурой. Для сетей данного типа можно предложить быстрые алгоритмы обучения, сходящиеся за конечное число шагов. Сети данного класса имеют глубокую связь с алгоритмом быстрого преобразования Фурье и обобщенного быстрого преобразования. Первые предложения по построению обобщенного ортогонального преобразования были высказаны Эндрюсом и Каспари [3] в 70-х годах прошлого века. А первые подходы к обучению подобных преобразований были развиты в работах А.И. Солодовникова и его научной группы в середине 80-х годов [4]. В то время подобный класс преобразований называли приспобленными быстрыми преобразованиями.

Возможность перестройки значений весовых коэффициентов и многослойная структура алгоритма роднит быстрые перестраиваемые преобразования с многослойными нейронными сетями прямого распространения. В рамках данной парадигмы быстрые линейные перестраиваемые преобразования являются частным случаем многослойных нейронных сетей и отличаются от последних линейными функциями активации и нулевыми смещениями в нейронах. Для обозначения нового класса сетей используется термин быстрые нейронные сети (БНС) [5]. Теоретическое развитие направления в основном касалось одномерных преобразований, но полученные результаты легко переносятся на двумерный и многомерные случаи.

Благодаря своей структуре БНС обладают специфичными алгоритмами обучения принципиально отличающимся от классического ErrorBackPropagation отсутствием механизма обратного распространения ошибки, что обеспечивает высокую скорость обучения и абсолютную сходимость. В основе алгоритмов обучения БНС лежит доказанное свойство структурной фрактальности, которое можно выразить системным инвариантом морфологического уровня [6]. Идея метода обучения БНС к одной или нескольким функциям основана на представлении каждой функции заданного набора в виде предфрактального произведения, отвечающего мультипликативной форме представления элементов матрицы быстрого преобразования.

В данной работе показано использование пирамидальной БНС для построения многоканального дискриминатора сигналов. Дано описание математической одномерной модели БНС, представлены методы построения топологии и обучения БНС для задачи корреляционных измерений в дискретном пространстве.

Одним из основных способов распознавания сигналов при корреляционных измерениях состоит в сопоставлении сигнала с эталоном. Если сходство между неизвестным сигналом и эталоном велико, то сигнал помечается как соответствующий эталонному. Простейшей мерой сходства является линейный дискриминант Фишера [7]. В частном случае этот дискриминант представляет собой меру взаимной энергии сигналов и выражается через нормированное скалярное произведение

$$d = (x,e) \Big/ \sqrt{(x,x)(e,e)} > \alpha \ ,$$

где x – неизвестный сигнал представленный вектором, e – вектор эталонного сигнала, скобки $(,)$ – обозначают скалярное произведение векторов, скаляр α определяет пороговый уровень различения сигналов. В многоканальном дискриминаторе используются несколько эталонных сигналов.

II. Математическая модель одномерных БНС

В работе Гуда [8] впервые было показано, что элементы матрицы быстрого преобразования Фурье можно представить в виде произведения элементов матриц базовых операций. На рис. 1 представлен граф быстрого преобразования в топологии Гуда с явным выделением базовых операций. Особенностью топологии Гуда является тождественность топологических связей для всех межслойных переходов. Базовую операцию определим матрицей $W^m_{i^m}$, где m – номер слоя, и i^m – номер базовой операции в пределах слоя. Элементы матриц базовых операций обозначим через $w^m_{i^m}(u_m, v_m)$, где u_m, v_m – определяют номер строки и номер столбца матрицы.

В алгоритмах быстрых преобразований удобно использовать нумерацию индексов, начиная с нулевого значения. Гуд показал, что элементы матрицы быстрого преобразования Фурье размерности $N = 2^n$ могут быть представлены в виде произведения элементов базовых операций:

$$h(U,V) = w^0_{i^0}(u_0, v_0)\, w^1_{i^1}(u_1, v_1) \ldots w^{n-2}_{i^{n-2}}(u_{n-2}, v_{n-2})\, w^{n-1}_{i^{n-1}}(u_{n-1}, v_{n-1}) \quad (1)$$

где топология преобразования определяется кортежами:

$$\begin{aligned} U^m &= \langle v_{m-1} v_{m-2} \ldots v_0 u_{n-1} u_{n-2} \ldots u_m \rangle, \\ V^m &= \langle v_m v_{m-1} \ldots v_0 u_{n-1} u_{n-2} \ldots u_{m+1} \rangle, \\ i^m &= \langle v_{m-1} \ldots v_0 u_{n-1} u_{n-2} \ldots u_{m+1} \rangle. \end{aligned} \qquad (2)$$

Здесь кортежи используются для поразрядного представления чисел в позиционной системе счисления. Например, для системы счисления с основанием 2 имеем:

$$u = \langle u_{n-1} u_{n-2} \ldots u_0 u_1 \rangle = u_{n-1} 2^{n-1} + u_{n-2} 2^{n-2} + \ldots + u_1 2 + u_0 \,.$$

В кортежах (2) индекс m определяет номер нейронного слоя, нумерация слоёв сети начинается с нуля, и заканчивается номером $n-1$, общее число слоёв равно n. Значение U^m определяет номер рецептора, значение V^m – номер нейрона в слое m, i^m – номер базовой операции в пределах слоя. В контексте нейронных сетей базовую операцию уместно называть нейронным ядром.

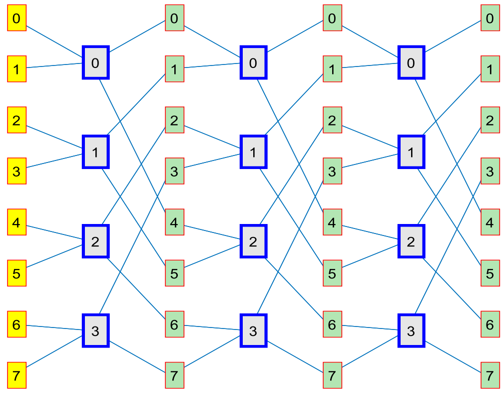

Рис. 1. Топологическая схема Гуда для размерности быстрого преобразования 8 с явным выделением базовых операций

Обобщённая теорема мультипликативной факторизации элементов матриц быстрых преобразований и правило построения топологического графа представлены в работе автора [9].

III. Пирамидальная быстрая нейронная сеть

Из выражения (2) следует, что для нулевого слоя номер базовой операции (нейронного ядра) определяется выражением

$$i^0 = \langle u_{n-1} u_{n-2} \ldots u_1 \rangle \,,$$

а глобальный номер рецептора нулевого нейронного слоя выражением

$$U^0 \ = \langle u_{n-1} u_{n-2} \ldots u_0 \rangle \,. \qquad (3)$$

Локальные переменные u_0, v_0 определяют позиционный номер рецептора и позиционный номер аксона в пределах каждого ядра нулевого слоя (нумерация рецепторов и аксонов начинается с нуля). Будем полагать, что число p_0 определяет число рецепторов ядра нулевого слоя, а число g_0 – число аксонов. Аналогичные обозначения p_m, g_m, где $m = 0, 1, \ldots, n-1$, будем использовать для размерностей рецепторных и аксоновых полей остальных слоёв сети. Выберем числа $g_m = 1$ для всех $m = 1, 2, \ldots, n-1$. В этом случае из (2) следует, что аксоны последнего слоя получат глобальные позиционные номера

$$V^{n-1} = \langle 0_{n-1} 0_{n-2} \ldots 0_1 v_0 \rangle \,.$$

Таким образом, нейронная сеть будет иметь ровно g_0 выходов, а из (3) следует, что число входов будет равно $N = p_0 \cdot p_1 \cdot \ldots \cdot p_{n-1}$. Построенная сеть является пирамидальной, с размерностью N по входу и g_0 по выходу. На рис. 2 приведен пример трехслойной пирамидальной сети для структуры $(p_0, p_1, p_2) = (4,2,2)$, $(g_0, g_1, g_2) = (3,1,1)$.

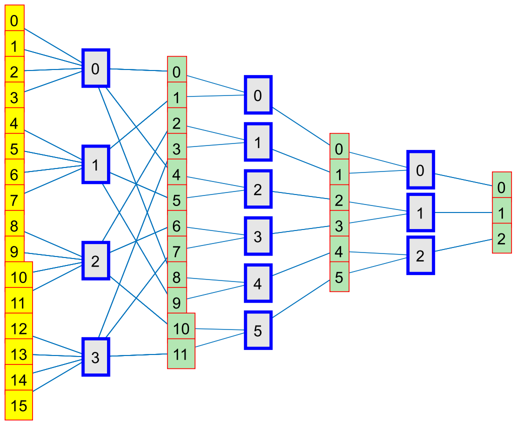

Рис. 2. Треслойная пирамидальная БНС с регулярной топологией

Детали построения топологического графа БНС изложены в работе [9].

IV. Алгоритм обучения сети к сигнальным функциям

Рассмотрим обучение быстрой пирамидальной сети для задачи многоканальной корреляции сигналов. Будем полагать, что эталонный сигнал, задан функцией $f(u)$ на дискретном интервале длиной $N = p_0 p_1 \ldots p_{n-1}$, где p_m произвольные целые числа. Представим аргумент функции в позиционной многоосновной системе счисления с основаниями $p_0, p_1, \ldots, p_{\kappa-1}$. Формула перехода, как известно, имеет вид:

$$u = \langle u_{n-1} u_{n-2} \ldots u_0 \rangle = u_{n-1} p_{n-2} p_{n-3} \ldots p_0 + \cdots \\ \cdots + u_{n-2} p_{n-3} p_{n-4} \ldots p_0 + \ldots + u_1 p_0 + u_0.,$$

где $u_i \in [0,1,\ldots p_i - 1]$ – разрядные переменные. В результате данного преобразования сигнал представляется как многомерная функция $f\langle u_{n-1} u_{n-2} \ldots u_0 \rangle$. Каждый аргумент функции определяет некоторый масштабный срез сигнала. Зафиксируем все аргументы функции кроме u_m. Варьируя свободный аргумент u_m, получим выборку S_m (с числом элементов p_m). Фрактальным фильтром [9] частотной локализации m называется произвольный функционал $F(S_m)$, определённый на выборке S_m. Операцию фрактальной фильтрации можно записать в виде:

$$f_{out} \langle u_{n-1} u_{n-2} \ldots u_{m+1} u_{m-1} \ldots u_0 \rangle = \underset{u_m}{F} \left(f_{inp} \langle u_{n-1} u_{n-2} \ldots u_0 \rangle \right).$$

В простейшем варианте фрактальный фильтр выполняет суммирование значений функции по аргументу u_m. Если $m = 0$, то такой фильтр генерализует сигнал, сглаживая мелкие детали. Фрактальная фильтрация по аргументу u_m приводит к сокращению интервала определения сигнала в p_m раз.

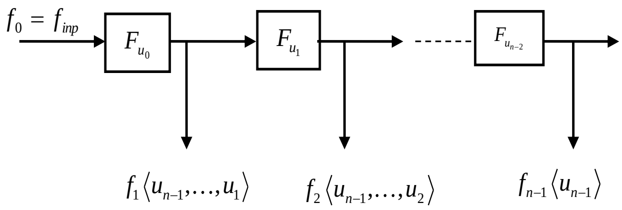

Рис. 3. Цепочка фрактальных фильтров

Рассмотрим цепочку фрактальных фильтров, показанную на рис. 3. Выходные сигналы для фильтров цепочки определяются рекуррентным соотношением:

$$f_m \langle u_{n-1} \ldots u_m \rangle = \underset{u_{m-1}}{F} \left(f_{m-1} \langle u_{n-1} \ldots u_m u_{m-1} \rangle \right).$$

Введем функции:

$$\varphi_{i^m}(u_m) = \frac{f_m \langle u_{n-1} u_{n-2} \ldots u_m \rangle}{f_{m+1} \langle u_{n-1} u_{n-2} \ldots u_{m+1} \rangle}, \qquad m = 0,1,\ldots n-2,$$

где $i^m = \langle u_{n-1} u_{n-2} \ldots u_{m+1} \rangle$.

Используя определения функций, можно записать

$$f_m \langle u_{n-1} u_{n-2} \ldots u_m \rangle = \varphi_{i^m}(u_m) \, f_{m+1} \langle u_{n-1} u_{n-2} \ldots u_{m+1} \rangle, \\ m = 0,1,\ldots n-2. \tag{4}$$

Из рекуррентных соотношений (4) непосредственно следует:

$$f(u) = f_0 \langle u_{n-1} u_{n-2} \ldots u_0 \rangle = \phi_{i^0}(u_0) \phi_{i^1}(u_1) \ldots \phi_{i^{n-2}}(u_{n-2}) \phi_{i^{n-1}}(u_{n-1}). \tag{5}$$

Нетрудно заметить, что выражения (1) и (5) подобны, учитывая, что для построенной пирамидальной сети

$$i^m = \langle 0_{m-1} \ldots 0_1 v_0 u_{n-1} u_{n-2} \ldots u_{m+1} \rangle,$$

можно записать следующее правило обучения нейронных ядер:

$$w_{i^0}^0(u_0, v_0) = \phi_{i^0}^k(u_0), \quad m = 0, \\ w_{i^m}^m(u_m, 0) = \phi_{i^m}^k(u_m), \quad m = 1,2,\ldots,n-1, \\ i^m = \langle v_0 u_{n-1} u_{n-2} \ldots u_{m+1} \rangle,$$

где k – номер эталонной сигнальной функции. Взаимно-однозначное соответствие $k \leftrightarrow v_0$ определяет упорядочение сигнальных функций в выходном слое.

V. Задача многоканальной корреляции

Будем полагать, что входной и эталонные сигналы нормированы к энергии, так что $(x,x)=1$ и $(e,e)=1$. В этом случае дискриминант определяется скалярным произведением

$$d=(x,e).$$

Скалярное произведение можно рассматривать как произведение вектора-строки x на одно-столбцовую матрицу $H=e'$. Или в покоординатном выражении

$$d=\sum_U x(U)h(U,0).$$

А для многоканального дискриминатора матрица H будет состоять из g_0 столбцов, и выполнять операцию

$$d(k)=\sum_U x(U)h(U,V).$$

Будем полагать, что матрица H является матрицей быстрого пирамидального преобразования с топологией Гуда. Поскольку сеть имеет g_0 выходов, то

$$V^{n-1}=\langle 0_{n-1}0_{n-2}\ldots 0_1 v_0\rangle.$$

На каждом выходе будет формироваться значение скалярного произведения, отвечающего эталонному сигналу. Пирамидальная сеть, показанная на рис. 2 может быть обучена к трем сигнальным функциям, заданным на интервале длинной 16.

Замечание. Поскольку для предложенной архитектуры пирамидальной сети выполняются условия: $g_m=1$ для всех $m=1,2,\ldots,n-1$, то из (2) следует

$$\begin{aligned}
U^m &=\langle 0_{m-1}0_{m-2}\ldots 0_1 v_0 u_{n-1}u_{n-2}\ldots u_m\rangle,\\
V^m &=\langle 0_m 0_{m-1}\ldots 0_1 v_0 u_{n-1}u_{n-2}\ldots u_{m+1}\rangle,\\
i^m &=\langle 0_{m-1}\ldots 0_1 v_0 u_{n-1}u_{n-2}\ldots u_{m+1}\rangle.
\end{aligned}$$

В частности, для нулевого слоя имеем:

$$\begin{aligned}
U^0 &=\langle u_{n-1}u_{n-2}\ldots u_0\rangle,\\
V^0 &=\langle v_0 u_{n-1}u_{n-2}\ldots u_1\rangle,\\
i^0 &=\langle u_{n-1}u_{n-2}\ldots u_1\rangle.
\end{aligned}$$

При варьировании в кортеже V^m переменных v_0 и u_i кортеж определяет множество номеров аксонов в пределах слоя. Нетрудно заметить, что это множество распадаются на непересекающиеся подмножества отличающиеся значением разряда v_0. Подчеркнем, что в БНС ветвления аксонов не допускается, и каждый нейрон слоя имеет точно один аксон [10].

Кортеж U^m определяют множества номеров рецепторов в каждом слое. С каждым аксоном нейронного ядра связано подмножество рецепторного поля, для которого разряды номеров рецепторов и аксона совпадают по значениям с одноименными разрядами номера нейронного ядра. Рецепторные поля нейронных ядер слоя не пересекаются.

Из сказанного следует вывод, что пирамидальная нейронная сеть предложенной архитектуры, распадается на независимые группы нейронов индексированных разрядом v_0. Каждая группа нейронов способна обучаться независимо от других групп. Таким образом, пирамидальная нейронная сеть обладает уникальной возможностью последовательного дообучения к эталонам без потери ранее накопленных знаний.

VI. Пластичность пирамидальной нейронной сети

В биологии термин пластичность используется как качественная характеристика способности нейронной сети обучаться при воздействии внешних факторов. Для искусственных нейронных сетей адекватной количественной оценкой может служить число независимых настроек, называемое также числом степеней свободы. Это значение как правило меньше полного количества настраиваемых синаптических весов нейронной сети. Для БНС получены точные формулы расчета числа степеней свободы [9]. Исходными данными для расчета является структурная модель нейронной сети. Структурная модель это взвешенный граф вершинами которого являются нейронные ядра. Вес каждого ядра определяется парой чисел, представляющих размерность по входу и выходу, а вес дуг определяется рангами операторов межъядерной связи. На рис. 4 представлена структурная модель для сети, показанной на рис.2.

Для БНС ранги операторов межъядерных связей равны 1 и на графе структурной модели не показаны. Формула расчёта числа степеней свободы для БНС имеет вид:

$$S(H)=\sum_{m=0}^{n-1} p_m g_m q_m-\sum_{m=0}^{n-2} D_m,$$

где p_m, g_m размерности рецепторного и аксонового поля ядра в слое m, q_m – число ядер в слое, D_m – количество одноранговых связей в межслойном переходе с номером m. Для структурной модели, представленной на рис. 4 непосредственно из графа получим $S(H)=4\cdot 3\cdot 4+2\cdot 1\cdot 6+2\cdot 1\cdot 3-12-6=48$. Напомним, что данная сеть может обучиться к трём эталонным функциям, заданным на интервале длиной 16. При произвольном выборе трёх функций необходимо задать 48 значений, что совпадает с числом степеней свободы пирамидальной сети.

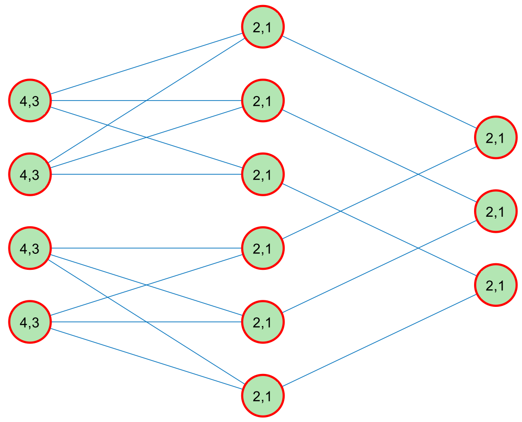

Рис. 4. Структурная модель пирамидальной нейронной сети

Таким образом, рассматриваемая сеть является глубокой в том смысле, что ее потенциал обучения используется полностью, и покрывает все допустимое многообразие эталонных сигналов.

VII. Заключение

Пирамидальные регулярные нейронные сети можно рассматривать как частный случай быстрых нейронных сетей, которые в свою очередь относятся к широкому классу самоподобных многослойных нейронных сетей [10]. Самодобные многослойные сети с регулярной структурой обладают уникальным свойством аналитического представления графа топологической модели, что позволяет для пирамидальных сетей разработать быстрые абсолютно сходящиеся алгоритмы обучения с конечным числом вычислительных операций. Кроме того, процесс обучения для каждого эталона может выполняться независимо от других, т. е. пирамидальная сеть данного типа обладает важным качествам накапливать знания, без необходимости переобучения сети при добавлении новых данных. В данной статье вывод основных соотношений основан на алгоритме быстрого преобразования с топологией Гуда, подобные выражения могут быть получены для топологии Кули−Тьюки или любой иной топологии быстрых преобразований.

Представленные в работе результаты показывают, что пирамидальные сети могут быть эффективно использованы для построения многоканальных корреляторов сигналов и изображений. Подобным образом могут быть построены дискриминаторы для локализации изображений и объектов в одномерных и многомерных пространствах. Более того, регулярный лес корреляционных сетей позволяет реализовать быстрые нейронные сети глубокого обучения [11], способные дообщаться без потери ранее накопленных знаний. Поскольку подобные сети имеют гарантированное время обучения при абсолютной сходимости алгоритма, то они могут использоваться в самообучающихся робототехнических системах реального времени.

Список литературы

[1] Phung S.L. and A. Bouzerdoum, “A pyramidal neural network for visual pattern recognition,” IEEE Transactions on Neural Networks, vol. 18, no. 2, pp. 329–343. 2007.

[2] Ihsan Ullahy and Alfredo Petrosinoy About Pyramid Structure in Convolutional Neural Networks arXiv:1608.04064v1 [cs.CV] 14 Aug 2016.

[3] Andrews H.C., Caspari K.L. A General Techniques for Spectral Analysis // IEEE. Tr. Computer. 1970. Vol C-19.-Jan, No 1. P.16–25.

[4] Солодовников А.И., Спиваковский А.М. Основы теории и методы спектральной обработки информации. Л., 1986. 272 с.

[5] Дорогов А.Ю. Быстрые нейронные сети: Проектирование, настройка, приложения. // Лекции по нейроинформатике Ч.1. В трудах школы-семинара «Современные проблемы нейроинформатики», науч.-техн. конф. “Нейроинформатика-2004” 28-30 января 2004г. Москва. Изд. МИФИ, 2004. С. 69–135.

[6] Дорогов А.Ю., Шестопалов М.Ю. Нейросетевое моделирование регулярных фракталов // Нейрокомпьютеры разработка и применение. №6, 2007. С. 3–15.

[7] Фишер Р.А. Статистические методы для исследователей. Москва, Госиздат, 1958. 267 с.

[8] Good I.J. The Interaction Algorithm and Practical Fourier Analysis // Journal of Royal Statistical Soseity. Ser.B. 1958. Vol.20. No.2. P. 361–372.

[9] Дорогов А.Ю. Теория и проектирование быстрых перестраиваемых преобразований и слабосвязанных нейронных сетей. СПб.: «Политехника», 2014. 328 с.

[10] Дорогов А.Ю. Самоподобные структуры многослойных нейронных сетей. // XXII Международная научно-техническая конференция «Нейроинформатика-2020» (10-16 октября 2020 г.): Сборник научных трудов. М. НИЯУ МИФИ, 2020. С. 214-224.

[11] Дорогов А.Ю. Быстрые нейронные сети глубокого обучения. // Сборник докладов III Международной научной конференции по проблемам управления в технических системах (CTS'2019).. Санкт-Петербург. 30 октября – 1 ноября 2019 г. СПб.: СПбГЭТУ «ЛЭТИ». С. 275-280.