

Корреляционный дискриминатор изображений в классе быстрых нейронных сетей

А. Ю. Дорогов

Санкт-Петербургский государственный электротехнический университет
«ЛЭТИ» им. В.И. Ульянова (Ленина)
vaks2006@yandex.ru

Аннотация. В работе представлены методы построения топологии и параметрического обучения быстрых нейронных сетей (БНС) для задач корреляционной обработки изображений. Описана математическая модель двумерных БНС. Отмечена связь быстрых алгоритмов с самоподобными структурами. Показан способ мультипликативного представления произвольных изображений в дискретном пространстве. Предложены алгоритмы построения топологии БНС и обучения для реализации корреляционного дискриминатора изображений.

Ключевые слова: изображение; линейный дискриминант; быстрая нейронная сеть; фрактальные произведения; самоподобные структуры

I. ВВЕДЕНИЕ

Одним из основных способов распознавания изображений состоит в сопоставлении изображения с эталоном. Если сходство между неизвестным изображением и эталоном велико, то изображение помечается как соответствующее эталонному. Простейшей мерой сходства является линейный дискриминант Фишера [1]. В частном случае этот дискриминант представляет собой меру взаимной энергии изображений и выражается через нормированное скалярное произведение

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

где x – неизвестное изображение, e – эталонное изображение, скобки $(,)$ – обозначают скалярное произведение изображений, скаляр α определяет пороговый уровень различения сигналов.

Разработку алгоритма быстрого преобразования Фурье (БПФ) безусловно, можно считать одним из выдающихся достижений второй половины 20 века. Дальнейшим развитием этого направления является класс перестраиваемых быстрых алгоритмов. Сохраняя структуру графа и изменяя только параметры базовых операций в перестраиваемых алгоритмах можно переходить от одного быстрого ортогонального преобразования к другому. Первые предложения по построению обобщенного ортогонального преобразования были высказаны Эндрюсом и Каспари [2] 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 [3]. В то время подобный класс преобразований называли приспособленными быстрыми преобразованиями.

Возможность перестройки значений весовых коэффициентов и многослойная структура алгоритма роднит быстрые перестраиваемые преобразования с многослойными нейронными сетями прямого распространения. Иногда используют термин – ортогональные нейронные сети. 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. Теоретическое развитие направления в основном касалось одномерных преобразований, но полученные результаты легко переносятся на двумерный и многомерные случаи. Для обозначения нового класса сетей используется термин быстрые нейронные сети (БНС) [4].

Благодаря своей структуре БНС обладают специфичными алгоритмами обучения принципиально отличающимся от классического ErrorBackPropagation отсутствием 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 [5]. 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 БНС для построения дискриминатора изображений. Дано описание математической модели двумерной БНС, представлены методы построения топологии и обучения БНС для задачи корреляционных измерений.

II. МАТЕМАТИЧЕСКАЯ МОДЕЛЬ ОДНОМЕРНЫХ БНС

В работе Гуда [6] 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 $W_{i_m}^m$, где m – номер слоя, и i_m – номер базовой операции в пределах слоя. Элементы матриц базовых операций обозначим через $w_{i_m}^m(u_m, v_m)$, 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:

$$h(U, V) = w_{p_0}^0(u_0, v_0) w_{p_1}^1(u_1, v_1) \dots w_{p_{n-2}}^{n-2}(u_{n-2}, v_{n-2}) w_{p_{n-1}}^{n-1}(u_{n-1}, v_{n-1}) \quad (1)$$

где топология преобразования определяется кортежами:

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

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

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

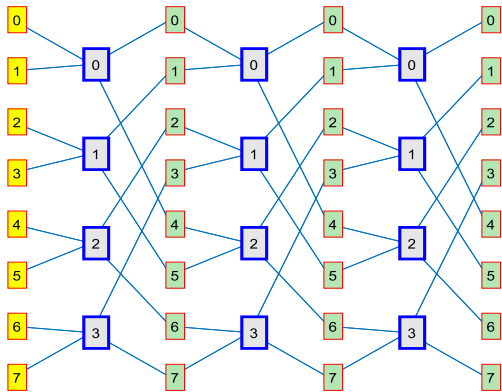


Рис. 1. Топологическая схема Гуда для размерности быстрого преобразования 8 с явным выделением базовых операций

В кортежах (2) значение U^m определяет номер рецептора, а значение V^m – номер нейрона в слое m . Нумерация слоёв сети начинается с нуля, полное число слоёв равно n . Обобщённая теорема мультипликативной факторизации элементов матриц быстрых преобразований и правило построения топологического графа представлены в работе [7].

III. МАТЕМАТИЧЕСКАЯ МОДЕЛЬ ДВУМЕРНЫХ БНС

Обозначим через $F(U_y, U_x)$ матрицу изображения размерностью $N_y \times N_x$. При воздействии на изображение линейного преобразования $h(U_y, U_x; V_y, V_x)$ получается массив из $M_y \times M_x$ коэффициентов. Двумерное преобразование выполняется по правилу:

$$S(V_y, V_x) = \sum_{U_y=0}^{N_y-1} \sum_{U_x=0}^{N_x-1} F(U_y, U_x) h(U_y, U_x; V_y, V_x) \quad (3)$$

Необходимым условием существования быстрого алгоритма является возможность мультипликативной декомпозиции значений обеих размерностей изображения в равное число сомножителей:

$$N_y = p_0^y p_1^y \dots p_{n-1}^y, \quad N_x = p_0^x p_1^x \dots p_{n-1}^x.$$

Индексы x, y здесь означают принадлежность к осям координат исходного изображения. Данное условие не является жестким ограничением, поскольку некоторые сомножители могут быть единичными. Тем не менее, чем больше число неединичных сомножителей в декомпозициях тем выше вычислительная эффективность быстрого алгоритма. Используя сомножители декомпозиций, координаты точек изображения представим в позиционной системе счисления со смешанными основаниями:

$$U_y = \langle u_{n-1}^y u_{n-2}^y \dots u_1^y u_0^y \rangle, \quad U_x = \langle u_{n-1}^x u_{n-2}^x \dots u_1^x u_0^x \rangle, \quad (4)$$

где вес m -го разряда определяется выражением $p_{m-1}^* p_{m-2}^* \dots p_1^* p_0^*$, а u_m^* – является разрядной переменной, принимающей значения $[0, p_m^* - 1]$ (звездочка здесь и далее заменяет индексы x, y). Аналогично можно представить координаты спектральных коэффициентов в плоскости $[V_y, V_x]$ следующим образом:

$$V_y = \langle v_{n-1}^y v_{n-2}^y \dots v_1^y v_0^y \rangle, \quad V_x = \langle v_{n-1}^x v_{n-2}^x \dots v_1^x v_0^x \rangle.$$

Для математического описания топологического графа быстрого алгоритма удобно использовать поразрядную форму представления чисел. Например, для топологии Гуда граф может быть описан набором кортежей

$$\begin{aligned} U_*^m &= \langle v_{m-1}^* v_{m-2}^* \dots v_0^* u_{n-1}^* u_{n-2}^* \dots u_m^* \rangle, \\ V_*^m &= \langle v_m^* v_{m-1}^* \dots v_0^* u_{n-1}^* u_{n-2}^* \dots u_{m+1}^* \rangle, \\ i_*^m &= \langle v_{m-1}^* v_{m-2}^* \dots v_0^* u_{n-1}^* u_{n-2}^* \dots u_{m+1}^* \rangle, \end{aligned}$$

где позиционное представление координат U_y^m, U_x^m и V_y^m, V_x^m определяют топологию во внутренних слоях быстрого алгоритма. На основе данных кортежей может быть построен топологический граф алгоритма [7]. На рис.2 приведен пример построения подобного графа для размерностей $N_y \times N_x = 8 \times 8$, $M_y \times M_x = 8 \times 8$.

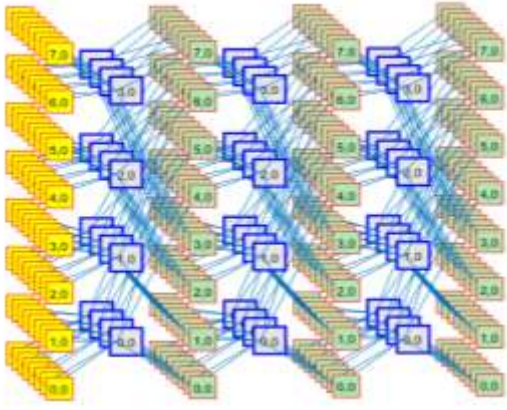


Рис. 2. Топологический граф быстрого алгоритма двумерного преобразования

В общем случае топологии направлений x и y могут различаться. На графе топологии показаны базовые операции $w_{i_x^m, i_y^m}^m(u_{i_x^m}^y, u_{i_y^m}^x; v_{i_x^m}^y, v_{i_y^m}^x)$ (выделены синим цветом) представляющие собой четырехмерные матрицы размерности $[p_m^y, p_m^x; p_m^y, p_m^x]$. Также как и в одномерном случае, каждый элемент четырехмерной матрицы преобразования H выражается через элементы базовых операций в виде произведения:

$$h(U_y, U_x; V_y, V_x) = w_{i_x^{n-1}, i_y^{n-1}}^{n-1}(u_{i_x^{n-1}}^y, u_{i_y^{n-1}}^x; v_{i_x^{n-1}}^y, v_{i_y^{n-1}}^x) w_{i_x^{n-2}, i_y^{n-2}}^{n-2}(u_{i_x^{n-2}}^y, u_{i_y^{n-2}}^x; v_{i_x^{n-2}}^y, v_{i_y^{n-2}}^x) \dots w_{i_x^0, i_y^0}^0(u_{i_x^0}^y, u_{i_y^0}^x; v_{i_x^0}^y, v_{i_y^0}^x), \quad (5)$$

где поразрядные выражения индексов ядер слоя m для выбранной топологии имеют вид:

$$i_*^m = \langle v_{m-1}^* v_{m-2}^* \dots v_0^* u_{n-1}^* u_{n-2}^* \dots u_{m+1}^* \rangle.$$

IV. МУЛЬТИПЛИКАТИВНОЕ ПРЕДСТАВЛЕНИЕ ДИСКРЕТНЫХ ИЗОБРАЖЕНИЙ

Алгоритм мультипликативного представления основан на идеях последовательной фрактальной фильтрации и для двумерного случая представляет собой кратно-масштабную обработку изображений, последовательно сжимающую его размеры вплоть до единственной точки [7]. Схему фрактальной фильтрации можно представить в виде пирамиды показанной на рис. 3. Основанием пирамиды является исходное изображение, $F(U_y, U_x)$ для которого аргументы U_y и U_x представлены в позиционной системе счисления (см. (4)). В этом позиционном представлении зафиксируем все разряды кроме двух младших u_0^y и u_0^x . Если варьировать эти разряды по всем возможным значениям, то получим выборку размером $p_0^y \times p_0^x$.

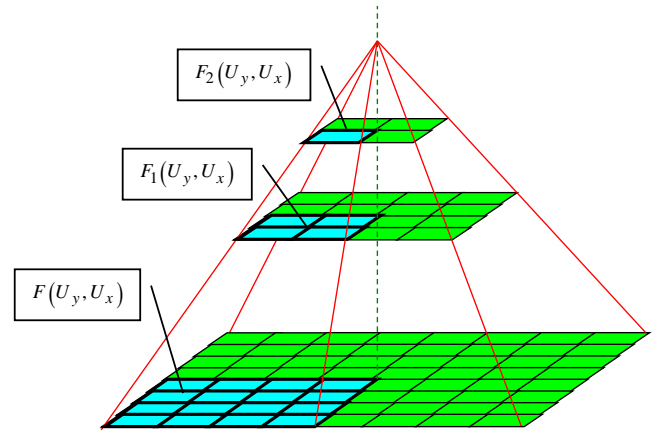


Рис. 3. Алгоритм мультипликативной декомпозиции изображений

Под фрактальным фильтром будем понимать произвольный функционал Φ , заданный на этой выборке. Формально фильтр можно записать в виде выражения:

$$F_1(\langle u_{n-1}^y, u_{n-2}^y, \dots, u_1^y \rangle, \langle u_{n-1}^x, u_{n-2}^x, \dots, u_1^x \rangle) = \Phi_{(u_0^y, u_0^x)} [F(\langle u_{n-1}^y, u_{n-2}^y, \dots, u_1^y, u_0^y \rangle, \langle u_{n-1}^x, u_{n-2}^x, \dots, u_1^x, u_0^x \rangle)]$$

Очевидно, что изображение F_1 будет кратно уменьшенным по размерам по отношению к исходному изображению. Функционалом, например, может быть правило вычисления среднего значения выборки или ее медианы. Исходное изображение теперь можно формально представить в виде произведения:

$$F(\langle u_{n-1}^y, u_{n-2}^y, \dots, u_1^y, u_0^y \rangle, \langle u_{n-1}^x, u_{n-2}^x, \dots, u_1^x, u_0^x \rangle) = F_1(\langle u_{n-1}^y, u_{n-2}^y, \dots, u_1^y \rangle, \langle u_{n-1}^x, u_{n-2}^x, \dots, u_1^x \rangle) f_{j_0^y, j_0^x}(u_0^y, u_0^x),$$

где $f_{j_0^y, j_0^x}(u_0^y, u_0^x)$ – набор двумерных функций-множителей, зависящих от разрядных переменных u_0^y и u_0^x , а индексы j_0^y, j_0^x выделяют функцию из набора. Для получения функций множителей достаточно выполнить скалярное деление изображения F на изображение F_1 при варьировании всех разрядных переменных. Изображение F_1 в свою очередь также может быть представлено как произведение изображения F_2 на множители из набора $f_{j_1^y, j_1^x}(u_1^y, u_1^x)$. Повторяя операцию фрактальной фильтрации и разложения многократно, достигнем вершины пирамиды изображений и получим в результате мультипликативную декомпозицию исходного изображения:

$$F(\langle u_{n-1}^y, u_{n-2}^y, \dots, u_1^y, u_0^y \rangle, \langle u_{n-1}^x, u_{n-2}^x, \dots, u_1^x, u_0^x \rangle) = f_{j_{n-1}^y, j_{n-1}^x}(u_{n-1}^y, u_{n-1}^x) \cdot f_{j_{n-2}^y, j_{n-2}^x}(u_{n-2}^y, u_{n-2}^x) \dots f_{j_1^y, j_1^x}(u_1^y, u_1^x) f_{j_0^y, j_0^x}(u_0^y, u_0^x), \quad (6)$$

где $j_m^y = \langle u_{n-1}^y, u_{n-2}^y, \dots, u_{m+1}^y \rangle$ и $j_m^x = \langle u_{n-1}^x, u_{n-2}^x, \dots, u_{m+1}^x \rangle$.

V. ПРОЕКТИРОВАНИЕ ТОПОЛОГИИ И ОБУЧЕНИЕ КОРРЕЛЯЦИОННОЙ НЕЙРОННОЙ СЕТИ

Будем полагать, что входное и эталонное изображения 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 изображения x на одно-плоскостную матрицу $H = e'$. Или в покоординатном выражении

$$d = \sum_{U_x, U_y} x(U_x, U_y) h(U_x, U_y; V_x, V_y).$$

Будем полагать, что матрица H является матрицей быстрого преобразования с топологией Гуда. Поскольку сеть имеет только один выход, то

$$V_y = \langle 0_{n-1}^y, 0_{n-2}^y \dots 0_1^y, 0_0^y \rangle, \quad V_x = \langle 0_{n-1}^x, 0_{n-2}^x \dots 0_1^x, 0_0^x \rangle.$$

Таким образом, все разрядные переменные $v_m^* = 0$ для $m = 0, 1, \dots, n-1$. Т.е. каждая базовая операция также будет иметь только один выход с локальным номером (0,0). Выражения (2) для топологии сети в этом случае примут вид:

$$\begin{aligned} U_*^m &= \langle 0_{m-1}^* 0_{m-2}^* \dots 0_0^* u_{n-1}^* u_{n-2}^* \dots u_m^* \rangle, \\ V_*^m &= \langle 0_m^* 0_{m-1}^* \dots 0_0^* u_{n-1}^* u_{n-2}^* \dots u_{m+1}^* \rangle, \\ i_*^m &= \langle 0_{m-1}^* 0_{m-2}^* \dots 0_0^* u_{n-1}^* u_{n-2}^* \dots u_{m+1}^* \rangle, \end{aligned} \quad (7)$$

На рис. 4 представлен топологический граф коррелирующей сети для изображения $N \times N = 2^3 \times 2^3$. Каждая базовая операция данной нейронной сети представляет собой один нейрон, имеющий четыре входа.

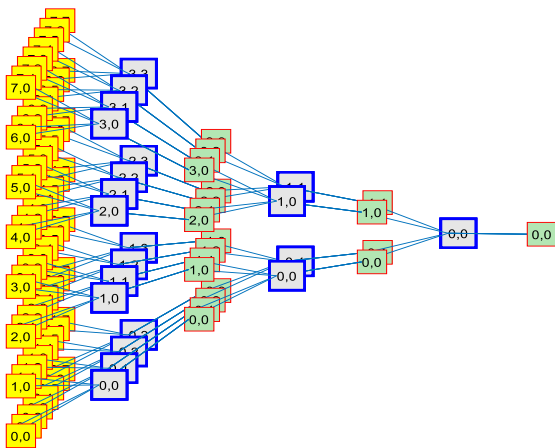


Рис. 4. Топология коррелирующей сети

Сравним между собой выражения (5) и (6). Нетрудно заметить, что при топологии (7) они совпадают. Отсюда сразу следует правило обучения нейронной сети:

$$w_{i_*^m, i_*^m}^m (u_m^y u_m^x; 0_m^y 0_m^x) = f_{i_*^m, i_*^m}^m (u_m^y, u_m^x).$$

VI. ЗАКЛЮЧЕНИЕ

Быстрые нейронные сети относятся к широкому классу самоподобных многослойных нейронных сетей [8]. Самоподобные многослойные сети с регулярной структурой обладают уникальными свойствами аналитического представления графа топологической модели, что позволяет разработать быстрые абсолютно сходящиеся алгоритмы обучения с конечным числом вычислительных операций. В данной статье алгоритм быстрого преобразования с топологией Гуда использован как пример, подобный алгоритм может быть реализован для топологии Кули-Тьюки или любой иной топологии быстрых преобразований.

Представленные в работе результаты показывают, что сети могут быть эффективно использованы для корреляционных измерений 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 [9].

СПИСОК ЛИТЕРАТУРЫ

- [1] Фишер Р.А. Статистические методы для исследователей. Москва, Госиздат, 1958. 267 с.
- [2] 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.
- [3] Солодовников А.И., Спиваковский А.М. Основы теории и методы спектральной обработки информации. Л., 1986. 272 с.
- [4] Дорогов А.Ю. Быстрые нейронные сети: Проектирование, настройка, приложения. // Лекции по нейроинформатике Ч.1. В тр. школы-семинара «Современные проблемы нейроинформатики», науч.-техн. конф. «Нейроинформатика-2004» 28-30 января 2004г. Москва. Изд. М.: МИФИ, 2004, с. 69-135.
- [5] Дорогов А.Ю., Шестопалов М.Ю. Нейросетевое моделирование регулярных фракталов. // Нейрокомпьютеры: разработка и применение. №6, 2007. с. 3-15.
- [6] 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.
- [7] Дорогов А.Ю. Теория и проектирование быстрых перестраиваемых преобразований и слабосвязанных нейронных сетей. СПб.: «Политехника», 2014. 328 с.
- [8] Дорогов А.Ю. Самоподобные структуры многослойных нейронных сетей. // XXII Международная научно-техническая конференция «Нейроинформатика-2020» (10-16 октября 2020г.): Сборник научных трудов. М. НИЯУ МИФИ, 2020. с. 214-224.
- [9] 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.