

ВОПРОСЫ ИСПОЛЬЗОВАНИЯ РЕПЛИКАТИВНЫХ НЕЙРОННЫХ СЕТЕЙ В ЗАДАЧАХ СЖАТИЯ ИЗОБРАЖЕНИЙ

Abstract: The experimental analysis of influence of some parameters of replicative neural net and training set for the speed and quality of training is executed; the formula for calculation of factor of the compression is offered; recommendations on formation of training set are given.

Key words: replicative neural net, method of decreasing dimensionality, neural net compression of images.

Анотація: Виконано експериментальний аналіз впливу деяких параметрів реплікативної нейронної мережі та навчальної множини на швидкість і якість навчання; запропонована формула для обчислення коефіцієнту стиснення, що реалізується на реплікативній нейронній мережі; подані рекомендації щодо формування навчальної множини.

Ключові слова: реплікативна нейронна мережа, метод зниження розмірності, нейромережеве стиснення зображень.

Аннотация: Выполнен экспериментальный анализ влияния некоторых параметров репликативной нейронной сети и обучающего множества на скорость и качество обучения; предложена формула для вычисления коэффициента сжатия, реализуемого на репликативной нейронной сети; даны рекомендации по формированию обучающего множества.

Ключевые слова: репликативная нейронная сеть, метод понижения размерности, нейросетевое сжатие изображений.

1. Введение

В качестве возможного практического применения искусственных нейронных сетей (ИНС) авторами ряда работ [2–5] рассматривается сжатие изображений, для осуществления которого предлагается, в частности, метод *понижения размерности*, основанный на следующих предположениях:

- 1) источник сжимаемых данных рассматривается как генератор векторов в многомерном пространстве;
- 2) генерируемые векторы занимают подпространство с размерностью меньшей, чем та, которая используется генератором, что открывает возможности для их более компактного представления.

Для осуществления отображения исходных данных в пространство низкой размерности в [2] было предложено использовать ИНС, обобщенная структурная схема которой представлена на рис. 1.

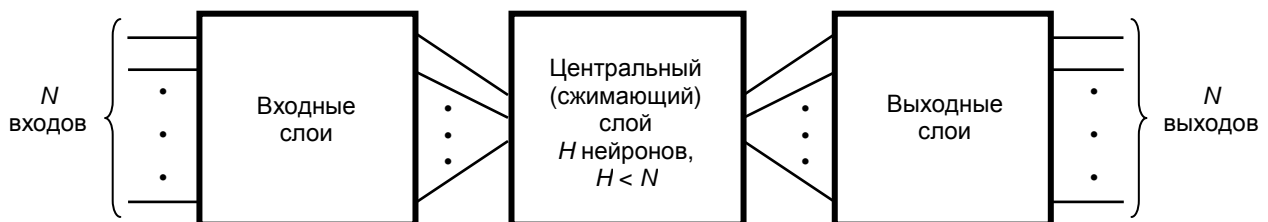


Рис. 1. Обобщенная структурная схема репликативной ИНС

В зависимости от необходимости выделить архитектурную или функциональную особенность, ИНС подобного вида называют также *репликативными* [2] или ИНС с “узким горлом”. Термин “репликация” отражает следующую особенность функционирования ИНС: воспроизведение входного вектора на выходе сети с минимальным уровнем искажений. Термин “узкое горло” предполагает наличие в сети скрытого слоя, количество нейронов, предназначенных для получения компактного (сжатого) представления, данных в котором, меньше, чем в остальных

слоях.

Сжатие изображений на репликативных ИНС выполняется по следующему алгоритму [3]:

- 1) разбиение сжимаемого изображения на непересекающиеся прямоугольные блоки;
- 2) построчное разложение блоков: формирование векторов-строк (рис. 2);
- 3) формирование обучающего множества сети из полученных на втором шаге векторов, при котором входной и целевой выходной векторы совпадают;
- 4) обучение ИНС на множестве, полученном на шаге 3 по алгоритму обратного распространения ошибки;
- 5) результат сжатия – массив выходных значений нейронов сжимающего скрытого слоя.

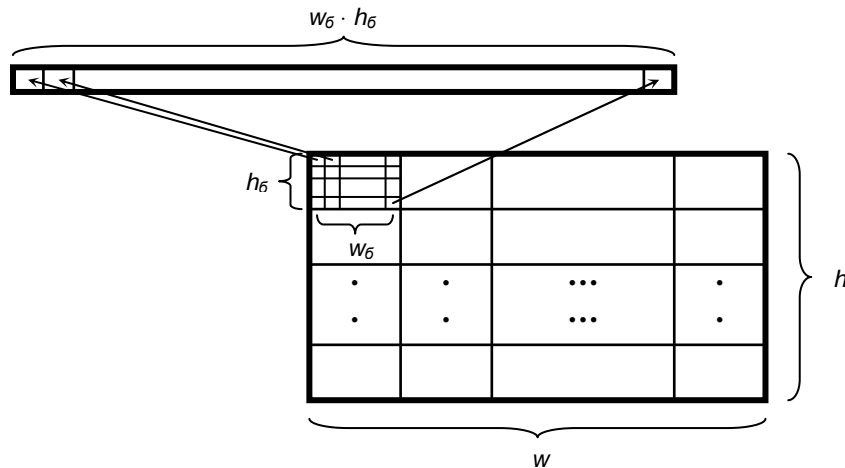


Рис. 2. Схема формирования векторов обучающего множества для репликативной ИНС

Для восстановления сжатого изображения необходимы лишь те слои ИНС, которые следуют после сжимающего (рис. 3). Входными данными в этом случае становятся сжатые векторы, размерность которых обученной ИНС восстанавливается до исходной.



Рис. 3. Обобщенная структурная схема восстанавливающей части репликативной ИНС

Простейшая оценка коэффициента k сжатия, реализуемого на репликативной ИНС, имеет вид

$$k = \frac{N}{H}, \quad (1)$$

где N – размерность входных векторов сети (исходная размерность);

H – размерность векторов выходных значений нейронов сжимающего слоя.

2. Постановка задачи

Несмотря на наличие сообщений [3, 4] об успешных экспериментах по сжатию изображений на репликативных ИНС, остаются недостаточно изученными влияние обучающего множества, а также отдельных параметров ИНС на степень сжатия и уровень искажений в восстановленном

изображении. Кроме того, отсутствует обоснование величины максимально достижимого коэффициента сжатия, реализуемого на репликативной ИНС. В результате возникают трудности при разработке методов построения ИНС с требуемыми техническими и эксплуатационными характеристиками.

В настоящей работе предлагается подход к построению и формированию обучающего множества репликативной ИНС, позволяющий уменьшить уровень потерь в восстановленном изображении с учетом требуемого коэффициента сжатия и повысить скорость обучения.

3. Основные результаты

Определение величины коэффициента сжатия

В работах, посвященных сжатию изображений на репликативных ИНС (например, [3, 4]) величину коэффициента сжатия предлагается вычислять в соответствии с (1). В этом выражении учитывается только отношение объемов исходного и сжатого изображений, но не учитывается та информация, которая необходима для восстановления сжатого изображения. Такую информацию можно не учитывать, когда обрабатывается некоторое множество однотипных изображений (например, аэрофотоснимков), поскольку в этом случае обучение ИНС может быть выполнено не для каждого отдельно взятого изображения, а для некоторого их подмножества. При использовании ИНС для сжатия изображений с уникальными характеристиками обучение ИНС должно выполняться применительно к каждому из обрабатываемых изображений. Вычисление коэффициента сжатия в соответствии с (1) в этом случае может приводить к завышенным значениям коэффициента и, вследствие этого, получению неверных результатов.

Вычисление коэффициента сжатия с учетом информации, необходимой для восстановления сжатого изображения, может быть выполнено с использованием следующих соотношений.

Пусть V_u – размер исходного изображения в битах, V_a – размер сжатого изображения (архива) с учетом информации, необходимой для восстановления. Коэффициент сжатия естественно определить как отношение этих размеров:

$$k = \frac{V_u}{V_a}. \quad (2)$$

Предположим, что для осуществления сжатия исходное изображение было разбито на блоки-векторы размерностью N . Количество векторов K при этом будет равно $\frac{V_u}{N}$.

При использовании трехслойной ИНС с H нейронами в единственном скрытом слое размер архива составит

$$V_a = KH + NH. \quad (3)$$

Здесь первое слагаемое представляет собой размер сжатого изображения (K векторов размерностью H), а второе – информацию, необходимую для восстановления – массив синаптических коэффициентов восстанавливающей части ИНС.

Таким образом, на основе (2) и (3) получаем

$$k = \frac{V_u}{KH + NH} = \frac{V_u}{\frac{V_u}{N}H + NH}. \quad (4)$$

При фиксированном количестве нейронов “узкого горла” (например, минимально возможном, когда $H = 2$), увеличение коэффициента сжатия возможно лишь за счет увеличения размерности исходных векторов – N . Однако из (4) видно, что при этом растет и объем массива синаптических коэффициентов восстанавливающей части ИНС [1].

Продифференцировав (4) по N , получим

$$\frac{dk}{dN} = \frac{V_u}{H} \cdot \frac{1 - \frac{V_u}{N^2}}{\left(\frac{V_u}{N} + N\right)^2}. \quad (5)$$

Выражение (5) обращается в нуль при $\frac{V_u}{N^2} = 1$, что соответствует максимуму функции (4) и, следовательно, наибольший коэффициент сжатия будет тогда, когда размерность исходных блоков-векторов изображения

$$N = \sqrt{V_u}. \quad (6)$$

Таким образом, несмотря на то, что по мере увеличения размера блоков, составляющих исходное изображение, отношение размерностей $\frac{N}{H}$ возрастает, коэффициент сжатия, начиная с момента $N = \sqrt{V_u}$, уменьшается за счет увеличения “размеров” восстанавливающей сети.

Пренебрежение данной закономерностью приводит к существенным ошибкам вычисления коэффициента сжатия. Так, например, в [3] приведены данные, касающиеся результатов сжатия нескольких изображений различного размера трехслойной ИНС с $H = 20$ и разбиением на блоки размерностью $N = 64$.

Для одного из этих изображений $V_{u1} = 256 \times 256 = 65536$ байт и $K_1 = \frac{V_{u1}}{N} = \frac{65536}{64} = 1024$,

для другого $V_{u2} = 128 \times 144 = 18432$ байт и $K_2 = \frac{V_{u2}}{N} = \frac{18432}{64} = 288$. При этом предполагается,

что в обоих случаях коэффициент сжатия один и тот же: $k = \frac{N}{H} = \frac{64}{20} = 3,2$.

Вместе с тем вычисление коэффициента сжатия по формуле (4), т.е. с учетом информации, необходимой для восстановления, дает следующие результаты:

для первого изображения: $k_1 = \frac{V_{u1}}{K_1 H + NH} = \frac{65536}{1024 \cdot 20 + 64 \cdot 20} \approx 3,02$;

для второго изображения: $k_2 = \frac{V_{u2}}{K_2 H + NH} = \frac{18432}{288 \cdot 20 + 64 \cdot 20} \approx 2,62$.

Как видно, во втором случае предполагаемый коэффициент сжатия был завышен не менее чем на 20 %.

Более существенная ошибка при вычислении коэффициента сжатия допускается в [4]. Для компрессии здесь предложено использовать пятислойную ИНС с 4096 входами и выходами, 40 нейронами в сжимающем слое и 410 нейронами в каждом промежуточном слое. При этом изображение размером 192×256 пикселей разбивается на блоки размерностью $N = 64 \times 64 = 4096$

пикселей общим количеством $K = \frac{192 \times 256}{4096} = 12$. Предполагаемый коэффициент сжатия, по

мнению авторов, составляет $k = \frac{4096}{40} \approx 102,4$, однако, поскольку

$$k = \frac{V_u}{KH_1 + H_1H_2 + H_2N}, \quad (7)$$

где H_1 – количество нейронов в центральном скрытом слое;

H_2 – количество нейронов во втором промежуточном слое;

в действительности $k = \frac{192 \cdot 256}{12 \cdot 40 + 40 \cdot 410 + 410 \cdot 4096} \approx 0,03 (!)$, т.е. при указанных параметрах ИНС

и обучающего множества вместо компрессии данных имеет место увеличение их объема.

Еще одним фактором, влияющим на способ вычисления коэффициента сжатия, являются разрядности представления исходных и сжатых данных, а также синаптических коэффициентов ИНС. Полагая, что в общем случае эти разрядности могут иметь разное значение, получим следующее выражение для коэффициента сжатия:

$$k = \frac{V_u R_u}{KHR_a + NHR_{НС}}, \quad (8)$$

где R_u – разрядность элементов исходного изображения;

R_a – разрядность выходов нейронов сжимающего слоя (разрядность архива);

$R_{НС}$ – разрядность синаптических коэффициентов нейронной сети.

Как правило, все параметры ИНС имеют одну и ту же разрядность (например, 32 бита: вещественное число формата IEEE – 754). В случае сжатия полутоновых или монохромных изображений (8 и 1 бит/пиксель соответственно) уменьшение размерности исходных данных может частично или полностью компенсироваться увеличением разрядности архива, что приведет к

снижению коэффициента сжатия в худшем случае в $\frac{R_a}{R_u}$ раз. При этом, несмотря на то, что

искусственное понижение разрядности архивных данных для полутоновых изображений является приемлемым, разрядность синаптических коэффициентов должна оставаться неизменной. Поэтому окончательное выражение для вычисления коэффициента сжатия принимает следующий вид:

$$k = \frac{V_u}{KH + NHR_{HC}}. \quad (9)$$

4. Использование перекрывающихся блоков

Для проведения эксперимента по сжатию полутонового изображения на репликативной ИНС нами была разработана и реализована в среде Borland CBuilder 6.0 программа, моделирующая работу многослойной ИНС со следующими регулируемыми параметрами сети:

- количество слоев сети (2 – 5);
- вид активационной функции нейронов каждого слоя (линейная или сигмоидная);
- коэффициент скорости обучения;
- максимальное количество итераций обучения;
- максимальное количество эпох обучения;
- разрядность представления выходных данных (32, 16, 8 или 4 бита)

и обучающего множества:

- ширина и высота блока изображения;
- наличие перекрытия блоков.

Было использовано полутоновое (Grayscale) изображение со следующими характеристиками:

- ширина – 100 пикселей;
- высота – 114 пикселей;
- разрядность цвета – 8 бит/пиксель;
- размер блока – $4 \times 3 = 12$ пикселей.

Сжатие выполнялось на ИНС со следующими фиксированными параметрами:

- количество слоев сети – 3;
- вид активационной функции: сигмоидная (thx) для всех слоев;
- коэффициент скорости обучения – 0,1;
- разрядность выходов нейронов сжимающего слоя – 8 бит.

Для обучения ИНС использовался алгоритм обратного распространения ошибки.

Результаты обучения на множестве неперекрывающихся блоков представлены в табл. 1.

Таблица 1. Результаты сжатия полутонового изображения с использованием неперекрывающихся блоков

Количество нейронов "узкого горла"	Количество векторов	Суммарное количество предъявлений векторов (23 эпохи)	Достигнутая минимальная ошибка обучения, %	Реализуемый коэффициент сжатия
6	950	21850	9,8	1,98
3	–	–	15,8	3,95
2	–	–	21,5	5,93

Для каждой моделируемой ИНС было проведено не менее пяти сжатий. Исходное изображение и типичные образцы восстановленных битовых карт приведены на рис. 4.



Рис. 4. Результаты использования неперекрывающегося разбиения:

а) исходное изображение; б), в), г) – изображения, восстановленные после сжатия с коэффициентом 1,98, 3,95 и 5,93 соответственно

Анализ приведенных образцов изображений показывает, что при увеличении коэффициента сжатия происходит общее снижение качества восстановленного изображения. Кроме того, из рассмотрения рис. 4 (б, в, г) следует: помимо общего его снижения, возникают многочисленные дефекты, расположенные, в основном, на границах контрастных переходов, т.е. там, где аппроксимируемая ИНС функция изменяет свое значение весьма быстро. Для обеспечения более гладкого вида аппроксимируемой функции и, таким образом, устранения данного негативного эффекта, в настоящей работе предлагается обучающее множество сети формировать из перекрывающихся блоков (рис. 5).

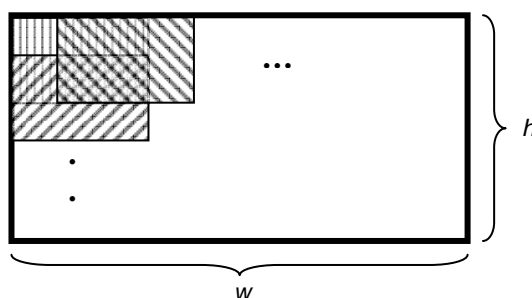


Рис. 5. Схема разбиения изображения на перекрывающиеся блоки

При разбиении с перекрытием количество блоков-векторов значительно возрастает, в связи с чем для обеспечения равноценных условий обучения ИНС количество обучающих эпох для блоков с перекрытием было уменьшено до двух. Количество предъявлений векторов сети при этом осталось приблизительно одинаковым.

Результаты обучения ИНС на множестве перекрывающихся блоков представлены в табл. 2.

Таблица 2. Результаты сжатия полутонового изображения с использованием перекрывающихся блоков

Количество нейронов "узкого горла"	Количество векторов	Суммарное количество предъявлений векторов (2 эпохи)	Достигнутая минимальная ошибка обучения, %	Реализуемый коэффициент сжатия
6	10864	21728	12,5	1,98
3	–	–	18,5	3,95
2	–	–	19,6	5,93

Из этой таблицы видно, что, несмотря на небольшое количество эпох обучения, полученная ошибка обучения незначительно отличается от ошибки, которая имела место при использовании неперекрывающегося разбиения. Кроме того, восстановленные изображения практически не

содержат дефектов, описанных ранее и свойственных сети, обученной на множестве неперекрывающихся блоков (рис. 6).



Рис. 6. Результаты использования перекрывающегося разбиения:

а) исходное изображение; б), в), г) – изображения, восстановленные после сжатия с коэффициентом 1,98, 3,95 и 5,93 соответственно

Таким образом, расширение обучающего множества “переходными” блоками-векторами улучшает общий результат сжатия за счет устранения дефектов, возникающих при использовании неперекрывающегося разбиения.

5. Разрядность сжатых данных

В процессе эксперимента были использованы следующие разрядности представления сжатых данных: 32, 16, 8 и 4 бита. Выбор требуемой разрядности осуществлялся в результате применения ступенчатой передаточной функции нейронов. При этом выходные значения нейронов округлялись применительно к соответствующей разрядной сетке. Визуальный анализ сжатых и восстановленных изображений показал, что по мере снижения разрядности до используемой для представления исходных данных (8 бит/пиксель) качество восстановленных изображений остается практически неизменным. Дальнейшее снижение разрядности, однако, приводит к появлению значительной текстуризации (рис. 7).



Рис. 7. Образцы изображений, восстановленных после сжатия с коэффициентом 1,98 и разрядностью выходных значений: а) 32 бита; б) 16 бит; в) 8 бит; г) 4 бита

6. Выводы

Коэффициент сжатия, реализуемый на трехслойной репликативной ИНС, предлагается вычислять по формуле, учитывающей размер описания восстанавливаемой части ИНС.

Формирование обучающего множества из перекрывающихся блоков изображения приводит к ускорению процесса обучения ИНС и улучшению качества восстановленного изображения.

Использование одинаковой разрядности для представления входных и выходных векторов позволяет получить наилучшее качество восстановленного изображения при минимальном размере архива.

СПИСОК ЛИТЕРАТУРЫ

1. Осовский С. Нейронные сети для обработки информации: Пер. с польск. Н.Д. Рудинского. – М.: Финансы и статистика, 2004. – 344 с.
2. Роберт Хехт-Нильсен. Нейрокомпьютинг: история, состояние, перспективы // Открытые системы. – 1998. – № 4. – С. 10–14.
3. Verma B., Blumenstein M., Kulkarni S. A New Compression Technique Using an Artificial Neural Network // Journal of Intelligent Systems. – 1999. – Vol. 9. – P. 39–53.
4. Cottrell G.W., Munro P., Zipser D. Image compression by back propagation: An example of extentional programming // Proc. 9th Annual Conf., Cognitive Soc. – 1987. – P. 461–473.
5. Jiang J. Image compression with neural networks – A survey. Signal Processing: Image Commun. – 1999. – N 14. – P. 737–760.