

УДК 004.032.2

DOI: 10.15827/0236-235X.123.430-434

Дата подачи статьи: 25.04.18

2018. Т. 31. № 3. С. 430–434

## **Исследование сжатия растровых изображений с использованием искусственных нейронных сетей**

А.А. Генов<sup>1</sup>, д.т.н., профессор, ведущий научный сотрудник, kt-mati@mail.ru

К.Д. Русаков<sup>2</sup>, младший научный сотрудник, rusakov.msk@yandex.ru

А.А. Моисеев<sup>3</sup>, студент, moisandry@gmail.com

В.В. Осипов<sup>1</sup>, к.т.н., доцент, старший научный сотрудник, ovv\_42@mail.ru

<sup>1</sup> Центр визуализации и спутниковых информационных технологий ФНЦ НИИСИ РАН, г. Москва, 117218, Россия

<sup>2</sup> Институт проблем управления им. В.А. Трапезникова РАН, г. Москва, 117997, Россия

<sup>3</sup> Московский государственный технический университет им. Н.Э. Баумана, г. Москва, 105005, Россия

Современные темпы роста объемов информации, хранящихся на жестких дисках, передаваемых по сети Интернет и локальным сетям предприятий, обусловили актуальность задачи сжатия, передачи и хранения данных. Большая часть передаваемых по сети данных представлена в виде мультимедийного контента.

Сегодня все более популярными становятся алгоритмы сжатия визуальной информации, основанные на нейросетевом аппарате. В отличие от классических алгоритмов, основанных на устранении избыточности, данные алгоритмы базируются на искусственных нейронных сетях. Направление актуально в связи с развитием математических алгоритмов обучения сети, что в дальнейшем позволит улучшить существующие методы сжатия.

Проведенный анализ публикаций показал, что в настоящее время конкретная информация о влиянии архитектуры искусственной нейронной сети на процесс обучения и качество их работы на реальном мультимедийном контенте отсутствует. Актуальна задача выбора топологии сети, наиболее подходящей для сжатия визуальной информации.

Целью авторов статьи является описание возможностей одного из типов искусственных нейронных сетей – многослойного персептрона – в области сжатия и восстановления изображений произвольного вида. Рассматриваются топологии искусственных нейронных сетей и алгоритмы их обучения, анализируется эффективность работы этих сетей. Описывается архитектура бутылочного горлышка, наиболее часто используемая при решении задачи сжатия и восстановления изображений. Приводится один из способов кодирования и декодирования данных, полученных при работе сетей. В статье описывается проведенный вычислительный эксперимент, приведены полученные результаты.

Результаты показали, что использование многослойного персептрона с входным вектором свыше восьми значений менее эффективно. В итоге предложена наиболее подходящая архитектура сети, которую можно использовать на практике.

**Ключевые слова:** нейронная сеть, алгоритм сжатия, изображение, машинное обучение.

Сжатие данных широко используется в самых разнообразных областях информационных технологий. Учитывая современные темпы роста объемов информации, хранящихся на жестких дисках, передаваемых по сети Интернет и локальным сетям предприятий, стоит задача сжатия, передачи и хранения данных. Большая часть видеоматериалов и других мультимедийных проектов размещается в Интернете, на различных видеосервисах, в социальных сетях, поэтому передаваемые по сети данные представлены в виде мультимедийного контента. В силу этого возникает задача по созданию более эффективных методов сжатия.

На сегодняшний день набирают популярность алгоритмы сжатия визуальной информации, основанные на нейросетевом аппарате. Данные алгоритмы базируются на искусственных нейронных сетях (ИНС, нейросети) в отличие от классических алгоритмов, основанных на устранении избыточности. Данное направление актуально в связи с развитием математических алгоритмов обучения сети, что в дальнейшем позволит улучшить существующие методы сжатия. Ускорить процесс сжатия и восстановления информации можно также за счет

растущей с каждым годом производительности современной вычислительной техники.

Однако существует проблема подбора топологии сети, наиболее подходящей для задачи сжатия. На данный момент нет конкретной информации о влиянии архитектуры сети на процесс обучения и качество работы на реальном мультимедийном контенте. Цель научной работы состоит в анализе возможностей одного из типов ИНС – многослойного персептрона – в области сжатия и восстановления растровых изображений.

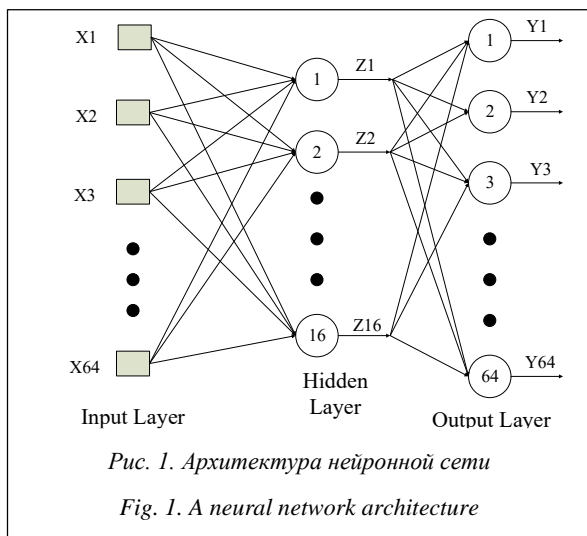
Во многих публикациях рассматривается аппарат ИНС. Так, в работах [1–3] описаны механизм функционирования нейросетей, взаимодействие искусственных нейронов, правила смены их состояний, расчет весовых коэффициентов.

В работе [4] рассматриваются методы сжатия геофизических данных с использованием многослойного персептрона в сети с архитектурой бутылочного горлышка. Предлагается комплексный алгоритм по сжатию геофизических данных при помощи ИНС.

В работах [5–8] приведены основные подходы к сжатию изображений на многослойном персеп-

троне, показан один из методов квантования полученных данных, описаны алгоритмы обучения сетей, такие, как алгоритм градиентного спуска и алгоритм Rprop (Resilient Propagation – упругое распространение). Однако в рассмотренных выше публикациях не была найдена информация по воздействию топологии сети на эффективность их работы.

Основная топология сети приведена на рисунке 1. Архитектура бутылочного горлышка предполагает обязательное наличие входного и выходного слоев одинаковой размерности, а также как минимум одного скрытого слоя, размерность которого много меньше размерности входного и выходного слоев.



Изображение представлено в виде вектора пикселей. Каждый вектор состоит из  $n$  пикселей, где  $n$  – размерность входного слоя сети. Числовые значения каждого пикселя подаются на вход ИНС. Полученные данные на выходах скрытого слоя составляют вектор сжатых данных размерности  $m$ . Вектор содержит вещественные числа, которые подвергаются кодированию и передаче получателю [1, 7]. Пример одного из способов кодирования представлен на рисунке 2.

Сжатые данные на скрытом слое подаются на следующий слой, где происходит их декодирование. В результате приведенных выше действий должно получиться исходное изображение.

Особенности описанных нейронных сетей:

- постоянный коэффициент сжатия  $K_c = \frac{m}{n}$  ;
- возможность разделения сети на две части: кодирующую и декодирующую.

На рисунке 1 представлена классическая архитектура сети ИНС,  $n = 64$ ,  $m = 16$ . Такая сеть имеет обозначение 64-16-64, где 64 – размерность входного и выходного слоев, 16 – размерность скрытого слоя. Приведенная архитектура обеспечивает коэффициент сжатия  $K_c = 4$ . При рассмотрении задачи

сжатия изображений данная топология используется в качестве базовой.

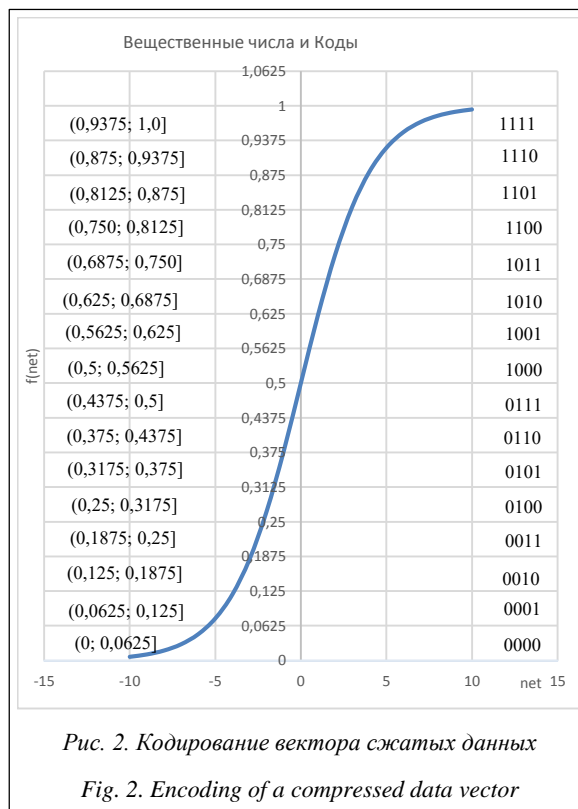
В качестве функции активации используется сигмоидальная [9, 10], которая обеспечивает диапазон значений  $[-1; 1]$  и позволяет применить градиентные методы по причине непрерывной дифференцируемости функции.

Так как для изображений числовые значения пикселей могут находиться в широком диапазоне, необходимо выполнить нормализацию вектора входных значений перед подачей на вход нейронной сети.

В реализации используется генератор случайных чисел с диапазоном значений  $[-1; 1]$  для инициализации весов ИНС. Возможно применение других методов, способствующих уменьшению количества локальных минимумов функции ошибки обучения, что позволяет уменьшить количество циклов обучения.

Для обучения ИНС применяются алгоритм наискорейшего спуска и метод обратного распространения ошибки [9, 11, 12], имеющие большое количество недостатков, основной из которых – длительность поиска минимизации направления весов. Однако другие алгоритмы (алгоритм наискорейшего спуска с адаптируемым шагом, алгоритм сопряженных градиентов) либо потребляют большое количество памяти, либо требуют больших вычислительных ресурсов, что недопустимо при работе с изображениями.

Обучение ИНС проходило на пятидесяти растровых изображениях с разрешением 512×512 пик-



селей в соответствии с моделью Train/Test/Validation в соотношении 75:15:10 %, где стадия Train отвечала за обучение, Test – за проверку качества обучения, Validation – за оценку качества работы ИНС.

Результаты обучения сети с архитектурой 64-16-64 показали, что количество циклов для каждого изображения сильно зависит от характеристик изображения. Наибольшее количество циклов обучения было на растровых изображениях с резкими перепадами интенсивности и большим диапазоном ее значений. Количество циклов на остальных изображениях распределилось в диапазоне примерно от 4 до 450 тыс., в среднем составив 100 тыс. циклов.

На реальных данных архитектура сети 64-16-64 показала себя не лучшим образом, восстановив изображение с достаточно низкой точностью. На рисунке 3 приведен пример исходного и восстановленного изображений.



Рис. 3. Конфигурация сети 64-16-64

Fig. 3. Network configuration 64-16-64

Быстрее обучение проходило на сети с конфигурацией 64-4-64, коэффициент сжатия  $K_c = 16$ . Количество циклов – 84 тысячи. Однако стоит отметить, что при уменьшении числа нейронов скрытого слоя уменьшилось число связей, поэтому приведенная архитектура работает хуже, чем 64-16-64. Такая ИНС подвергается эффекту переобучения.

Для примера рассмотрим конфигурацию вида 64-64-64 с коэффициентом сжатия  $K_c = 16$ . Среднее количество циклов обучения увеличилось незначительно и составило 120 тысяч. По сравнению с базовой архитектурой данная сеть показала лучшие результаты, изображение стало более контрастным.

Результат работы сети 64-64-64 показан на рисунке 4.

Проведенные исследования позволяют сделать вывод о том, что ИНС с архитектурой бутылочного горлышка, имеющие вектор входных значений размерностью 64 и один скрытый слой, не позволяют восстановить изображение с высокой точностью. Это можно объяснить тем, что при таком размере входного вектора количество возможных значений велико, приблизительно  $256^{64}$ , и однослойный персептрон просто не в состоянии запомнить всю обу-



Рис. 4. Конфигурация сети 64-64-64

Fig. 4. Network Configuration 64-64-64

чающую выборку. Происходит процесс переобучения, что ведет к ошибкам при восстановлении данных.

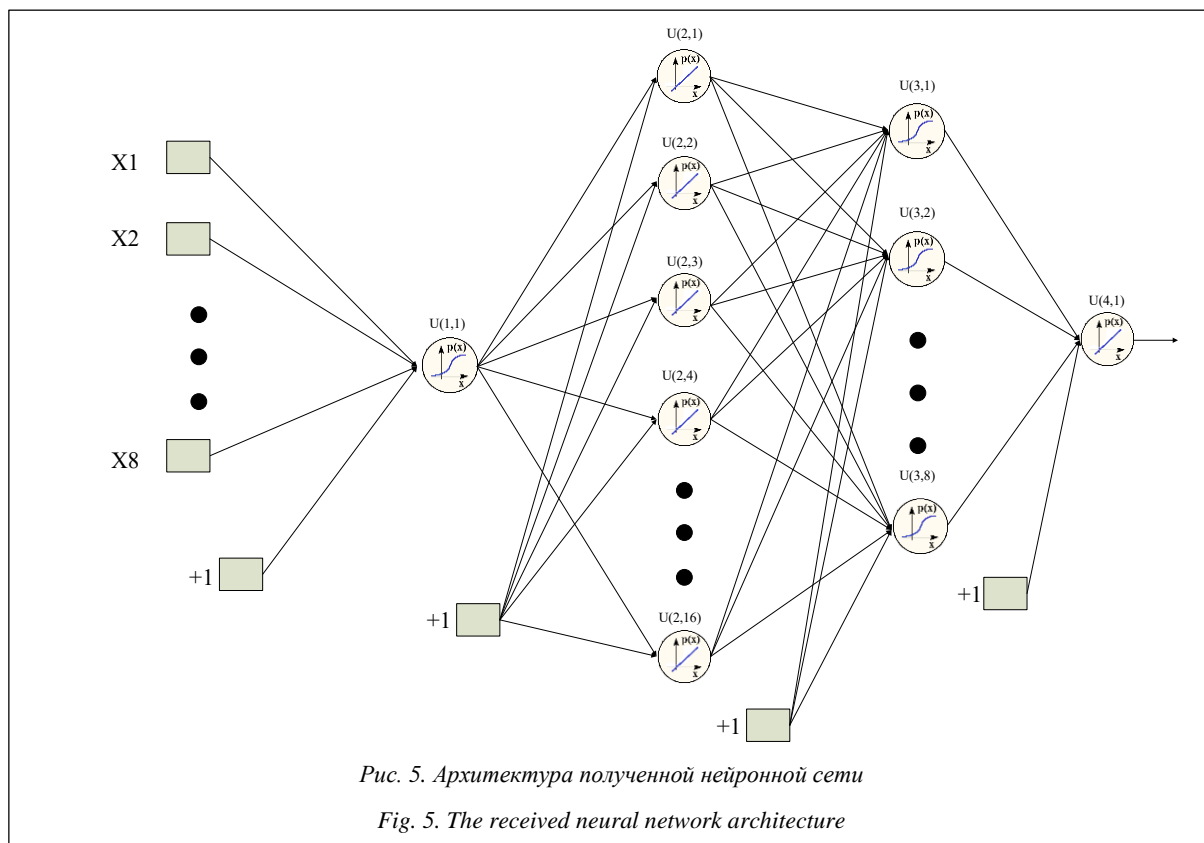
Для устранения выявленной проблемы было принято решение увеличить количество внутренних слоев и уменьшить количество нейронов на входе и выходе сети.

При выборе архитектуры были приняты во внимание скорость обучения и коэффициент ошибочных ответов.

В результате подобрана архитектура сети, представленная на рисунке 5.

Полученная архитектура ИНС имеет пять слоев, три из которых скрытые. Первый слой – вектор входных значений  $\{X_1, X_2, \dots, X_n\}$ , на который подается битовое представление сжимаемого изображения. Слой состоит из одного нейрона с сигмоидной функцией активации. Значение данного слоя  $Z$  подвергается квантованию, в результате чего получается битовая маска  $\{1001\}$ , которая является сжатым фрагментом изображения. Полученный фрагмент можно сохранить в файл или передать по сети. Для восстановления фрагмента необходимо выполнить обратное преобразование квантования, из битовой маски получить вещественное число, которое подается на третий скрытый слой. Так как пятый слой имеет линейную функцию активации,  $Y$  является целочисленным значением в диапазоне  $[0, 255]$ , что соответствует восстановленному байту изображения. Благодаря такой внутренней архитектуре данная ИНС способна запомнить все возможные значения и достаточно точно транслировать входной вектор бит в соответствующее байт-значение. При таком подходе сеть обучается не на наборе изображений, а на векторе бит-значений.

Достаточно сгенерировать набор случайных значений, который охватывал бы весь рассматриваемый диапазон. Благодаря этому была изменена стандартная обучающая выборка, представляющая последовательность байтов изображения, на случайную последовательность значений. Используя данный подход, уменьшаются потери качества изображения на уровне нейронной сети и обеспечивается коэффициент сжатия  $K_c = 2$ . Результат работы сети 8-1-16-8-1 показан на рисунке 6.



Восстановленное изображение не имеет серьезных искажений по сравнению с полученными предыдущими способами. На изображении присутствует плавность цветовых градиентов, отсутствуют точки повышенной резкости.

### Заключение

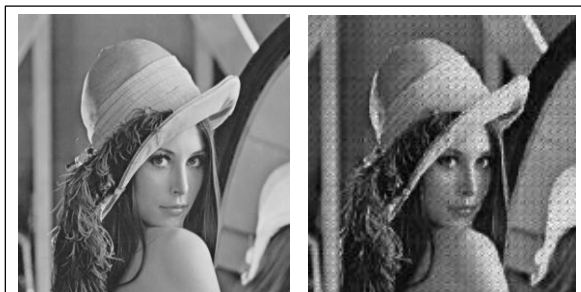
Таким образом, в работе были рассмотрены и проанализированы несколько базовых архитектур многослойного персептрона. Для оценки качества сжимаемых изображений использовалась метрика

$$K_k = \frac{Y_{\Gamma}}{Y} \cdot 100\%, \text{ где } Y_{\Gamma} - \text{количество верно угаданных битовых векторов; } Y - \text{полное множество битовых векторов.}$$

Для сети 64-16-64  $K_k = 64\%$ , для 64-64-64  $K_k = 70\%$  и для 8-1-16-8-1  $K_k = 95\%$ . Полученные результаты свидетельствуют о том, что предложенная структура сети более подходит для решения задач сжатия и восстановления растровых изображений. Для данного класса изображений наиболее приемлемый результат показала ИНС вида 8-1-16-8-1.

### Литература

1. Хайкин С. Нейронные сети: полный курс. М.: Вильямс, 2006. 1104 с.
2. Watta P., Desaie B., Dannug N., Hassoun M. Image compression using backprop. Detroit, Wayne State Univ. Publ., 1998, 3 p.
3. Круглов В.В. Нечеткая логика и искусственные нейронные сети. М.: Физматлит, 2001. 224 с.
4. Булаев В.И., Мунасыпов Р.А. Сжатие геофизических данных с применением вейвлет-преобразования и нейронных сетей // Вестн. УГАТУ. 2008. № 1. С. 200–207.
5. Masalmah Y.M. Image compression using neural networks. Puerto Rico, Univ. of Puerto Rico Publ., 2008. 4 с.
6. Осовский С. Нейронные сети для обработки информации; [пер. с польск. И.Д. Рудинского]. М.: Библио-Глобус, 2016. 448 с.
7. Куликов А.И., Михальченко Н.В. Сжатие растровых изображений нейронными сетями Цао Ена // GraphiCon. 2001. С. 231–236.
8. Брюхомицкий Ю.А. Нейросетевые модели для систем информационной безопасности. Таганрог: Изд-во ТРТУ, 2005. 160 с.
9. Воронцов К.В. Лекции по искусственным нейронным сетям. Т.: Изд-во ТРТУ, 2007. 29 с.



*Рис. 6. Улучшение качества изображения с сетью 8-1-16-8-1*

*Fig. 6. Improving image quality with the network 8-1-16-8-1*

10. Севастьянов А.А. Нейросетевая регуляризация решения обратных некорректных задач прикладной спектроскопии. М.: Финансы и статистика, 2003. 195 с.  
11. Вороновский Г.К. Генетические алгоритмы, нейронные

сети и проблемы виртуальной реальности. Х.: Основа, 1997. 112 с.

12. Короткий С. Нейронные сети: алгоритм обратного распространения. Х.: Основа, 2011. 155 с.

Software & Systems

DOI: 10.15827/0236-235X.123.430-434

Received 25.04.18

2018, vol. 31, no. 3, pp. 430–434

### Research on compression of raster images using artificial neural networks

A.A. Genov<sup>1</sup>, Dr.Sc. (Engineering), Professor, Leading Researcher, kt-mati@mail.ru

K.D. Rusakov<sup>2</sup>, Junior Researcher, rusakov.msk@yandex.ru

A.A. Moiseev<sup>3</sup>, Student, moisandry@gmail.com

V.V. Osipov<sup>1</sup>, Ph.D. (Engineering), Associate Professor, Senior Researcher, ovv\_42@mail.ru

<sup>1</sup> Center of Visualization and Satellite Information Technologies SRISA, Moscow, 117218, Russian Federation

<sup>2</sup> A. Trapeznikov Institute of Control Sciences of RAS, 117997, Russian Federation

<sup>3</sup> Bauman Moscow State Technical University, Moscow, 105005, Russian Federation

**Abstract.** Modern rates of information growth stored on hard disks transferred over the Internet and local enterprise networks has made it necessary to solve the problem of compressing, transferring and storing data. Most of the transferred data is in the form of multimedia content.

Nowadays, the algorithms for compressing visual information based on the neural network apparatus are becoming more popular. Unlike classical algorithms, which are based on the elimination of redundancy, these algorithms are based on artificial neural networks. The field is relevant due to the development of mathematical algorithms for network learning, which will improve existing compression methods in the future.

The analysis of publications showed that nowadays there is no particular information about the influence of the artificial neural network architecture on a learning process and the quality of their work in real multimedia content. The important task is to select a network topology, which is most suitable for compressing visual information.

The purpose of the article is to describe the capabilities of one of the types of artificial neural networks called a multi-layer perceptron in the area of compression and recovery of images of an arbitrary type. The paper analyzes topologies of artificial neural networks, algorithms for their learning, and the efficiency of their work. It also describes the architecture of a “bottle-neck”, which is most often used in solving the problem of image compression and recovery. The authors give one of the ways of encoding and decoding data obtained during network operation. The paper describes a computational experiment and gives its results.

The experiment showed that using a multilayer perceptron with an input vector of more than eight values turned out to be less effective. As a result, the authors propose the most suitable network architecture to use in practice.

**Keywords:** neural network, compression algorithm, image, machine learning.

### References

1. Haykin S. *Neural Networks: Full Course*. 2nd ed. Moscow, Vilyams Publ., 2006, 1104 p.
2. Watta P., Desai B., Dannug N., Hassoun M. *Image Compression Using Backprop*. Detroit, Wayne State Univ., 1998, 3 p.
3. Kruglov V.V. *Fuzzy Logic and Artificial Neural Networks*. Moscow, Fizmatlit Publ., 2001, 224 p.
4. Bulaev V.I., Munasypov R.A. Compression of geophysical data using wavelet transform and neural networks. *Bulletin of USATU*. 2008, no. 1, pp. 200–207 (in Russ.).
5. Masalmah Y.M. *Image Compression Using Neural Networks*. Puerto Rico, Univ. of Puerto Rico, 2008, 4 p.
6. Osovsky S. *Neural Networks for Information Processing*. Moscow, Biblio-Globus Publ., 2016.
7. Kulikov A.I., Mikhachenko N.V. *Compression of Raster Images by Cao Yen Neural Networks*. GraphiCon Publ., 2001, pp. 231–236 (in Russ.).
8. Bryukhomitsky Yu.A. *Neural Network Models for Information Security Systems*. Taganrog, TRTU, 2005, 160 p.
9. Vorontsov K.V. *Lectures on Artificial Neural Networks*. Taganrog, TRTU Publ., 2007, 29 p.
10. Sevastyanov A.A. *Neural Network Regularization of the Inverse Ill-Posed Problems of Applied Spectroscopy*. Moscow, Finansy i statistika Publ., 2003, 195 p.
11. Voronovsky G.K. *Genetic Algorithms, Neural Networks and Virtual Reality Problems*. Kharkiv, OSNOVA Publ., 1997, 112 p.
12. Korotky S. *Neural Networks: Back Propagation Algorithm*. Kharkiv, OSNOVA Publ., 2011, 155 p.