

АЛГОРИТМЫ СЖАТИЯ ИЗОБРАЖЕНИЙ НА ОСНОВЕ НЕЙРОСЕТЕВЫХ ТЕХНОЛОГИЙ

К. А. Щасный

*Белорусский государственный университет, г. Минск,
e-mail: kiril3467@yandex.by,
науч. рук.: Е. А. Головатая, старший преподаватель*

В данной работе ставится задача рассмотреть алгоритмы сжатия изображений с помощью нейронных сетей, а также реализовать один из них. В работе были рассмотрены несколько видов нейронных сетей, способных сжимать изображения. С помощью библиотеки ConvNetJS, а также с использованием образцов рукописного написания цифр MNIST, было реализовано сжатие изображений на основе архитектуры нейронных сети автоэнкодер.

Ключевые слова: сжатие изображений; нейронные сети; автоэнкодер; JavaScript; MNIST.

Выделяют 3 основных типа нейронных сетей, способных к сжатию изображений: сеть Кохонена и ее вариации, нейронные сети с ассоциативной памятью, например, сеть Хопфилда и автоэнкодеры.

Сеть Кохонена относится к методам обучения без учителя. Сама сеть и ее вариации часто используется для сжатия изображений с потерей качества. Она позволяет выделять похожие фрагменты данных в классы. Номер класса обычно занимает гораздо меньше места в памяти, чем ядро класса. Если передать получателю все ядра классов и номера классов, кодирующие каждый фрагмент данных, то данные могут быть восстановлены. При этом неизбежны потери, если число классов меньше числа различных фрагментов данных.

«Сжатие данных» для сетей Хопфилда является побочным явлением, так как главным их предназначением является восстановление исходного образа из зашумленных или поврежденных входных данных.

Автоэнкодеры – это нейронные сети прямого распространения, которые обучаются восстанавливать входной сигнал на выходе сети. Они характеризуются наличием скрытого слоя, как правило, меньшей размерности. В общем случае он представляет из себя код, описывающий данные, поступающие на вход сети, и использующийся для их восстановления в декодере.

Завершающим этапом работы стала реализация алгоритма сжатия. Для этого, с использованием библиотеки ConvNetJS, был реализован автоэнкодер, архитектура которого показана на Рис. 1. В качестве оптимизирующего алгоритма был использован адаптивный градиентный спуск с усреднением и отсечением (adadelat).

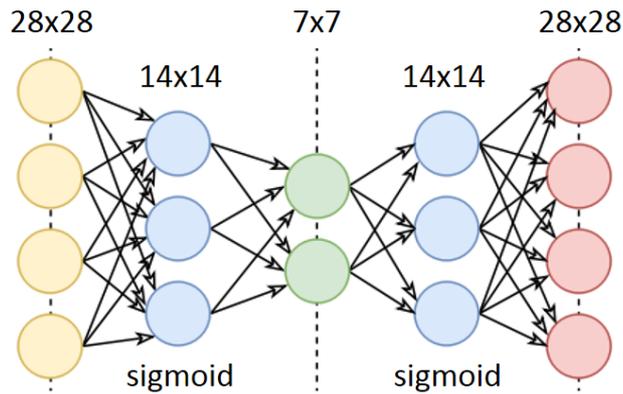


Рис. 1. Архитектура автоэнкодера

Обучение сети велось на изображениях чисел из набора MNIST, каждое размером 28x28 пикселей. Соответственно, входной и выходной слои сети имели по 28*28 нейронов. Центральный слой имел в своем составе 7*7 нейронов, что, как нетрудно подсчитать, подразумевает сжатие в 16 раз.

В процессе исследования работы алгоритма были отмечены следующие особенности: активизация задачи классификации сети и возможность переобучения на одном изображении. В первом случае сеть иногда выдавала не то изображение, на котором обучалась, но с той же цифрой, то есть сеть выполняла задачу классификации, одновременно сжимая данные.

На Рис. 2а представлен результат работы сети, обученной на наборе MNIST, состоящей из 10000 изображений. В этом случае, учитывая, что размер входного слоя 28*28 нейронов, а скрытого – 7*7, коэффициент сжатия составил $28*28/7/7 = 16$ раз.

Результат работы сети, переобученной на том изображении, которое впоследствии и сжималось представлен на Рис. 2б. В качестве обучающей выборки использовался набор с 500-800 копиями этого изображения. Как результат – с одной стороны, отказ работать с другими картинками, с другой – неплохие результаты сжатия. Даже с 5*5 нейронов в скрытом слое, что означает сжатие более чем в 31 раз, и 800 (в предыдущем случае было 10000) объектов в обучающей выборке, визуально вход и выход сети практически неразличимы. Однако, в данном случае для разжатия изображения на стороне получателя возникает необходимость передачи или хранения параметров декодера сети, например, в формате JSON. То есть, фактический коэффициент сжатия становится меньше за счет хранения этих параметров. Но, с другой стороны, если сжать его с помощью алгоритма сжатия, эффективно работающего с текстом или большим количеством повторяющихся элементов (например, RLE или

алгоритмы группы LZ), то результат может оказаться вполне приемлемым.

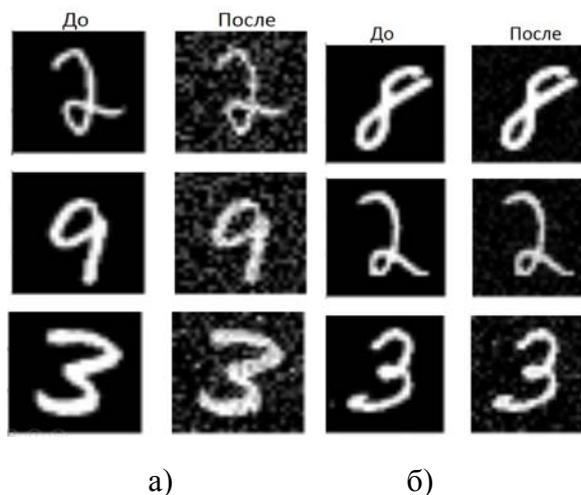


Рис. 2. Результаты обучения сети: а) на всей выборке, б) на одном изображении

Особенности нейронных сетей, заключающиеся в способности выявлять некоторые зависимости по входным данным, способствуют применению их к решению задачи сжатия данных, в том числе, изображений. В процессе работы удалось достичь главной цели исследования – сжать изображение. Основными недостатками данного подхода являются: побочное влияние результатов классификации на выдачу, долгое время обучения сети, особенно при усложнении архитектуры и увеличении размера изображения, необходимость передачи служебной информации. В перспективе планируется обучить нейронную сеть для работы с реальными полноразмерными изображениями.

Библиографические ссылки

1. Тропченко А. Ю., Тропченко А. А. Методы сжатия изображений, аудиосигналов и видео: Учебное пособие – СПб: СПбГУ ИТМО, 2009. С. 108.
2. Хайкин, Саймон. Нейронные сети: полный курс 2-е издание. : Пер. с англ. – М. : Издательский дом «Вильямс», 2006. С. 1104.
3. Ian Goldfellow, Yoshua Bengio, Aaron Courville. Deep learning – Boston, The MIT Press, 2016. С. 800
4. Eric Roberts. Neural networks and image compression[Electronic resource]. – Mode of access: <https://cs.stanford.edu/people/eroberts/courses/soco/projects/neural-networks/Applications/imagecompression.html>. Date of access: 04.2018.