



О.П. Солдатова, Ю.М. Шепелев

СЖАТИЕ ВИЗУАЛЬНЫХ ДАННЫХ С ИСПОЛЬЗОВАНИЕМ МНОГОСЛОЙНОГО ПЕРСЕПТРОНА

(Самарский государственный аэрокосмический университет имени академика С.П. Королева (национальный исследовательский университет))

Цель данной научной работы состоит в анализе возможностей одного из типов нейронных сетей – многослойного перцептрона – в области сжатия и восстановления изображений произвольного вида. Тема сжатия изображений с помощью многослойного перцептрона достаточно часто освещается в научных работах. Тем не менее, на данный момент нет конкретных данных о влиянии архитектурных параметров сети на процесс обучения и качество работы сети на реальных данных.

В рамках данной работы была разработана система, позволяющая конструировать многослойные перцептроны с архитектурой «бутылочного горлышка» [1, 2], наиболее часто используемой при решении задачи сжатия и восстановления изображений. В системе имеется возможность задать произвольное количество скрытых слоев, а также настроить количество нейронов в каждом слое. Созданная система хорошо подходит для исследования различных архитектур с последующим использованием на реальных данных. Программа разработана на языке программирования C++ с применением фреймворка Qt в интегрированной среде разработки Visual Studio 2013.

Базовая структура сети приведена на рисунке 1. Архитектура «бутылочного горлышка» предполагает обязательное наличие входного и выходного слоя одинаковой размерности, а также как минимум одного скрытого слоя, размерность которого выбирается много меньшей размерности входного и выходного слоев.

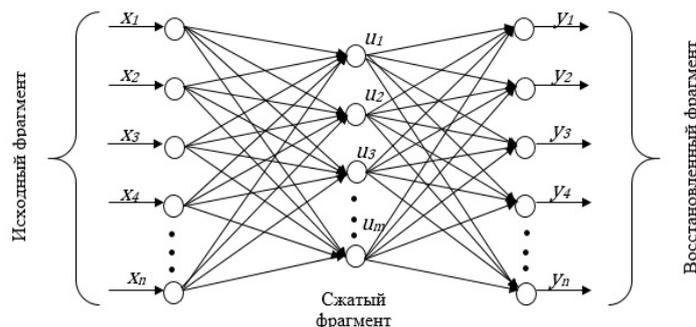


Рис. 1. Нейронная сеть с архитектурой «бутылочного горлышка»

Изображение разбивается на квадратные фрагменты. Каждый фрагмент состоит из n пикселей, где n – размерность входного слоя сети. Значения интенсивности каждого пикселя фрагмента подаются на вход сети. Данные, полученные на выходах скрытого слоя, образуют вектор сжатых данных размерности m . Данные представляют собой вещественные числа, которые могут быть закоди-



рованы и затем переданы по сети. Пример одного из способов кодирования представлен на рисунке 2 [1].

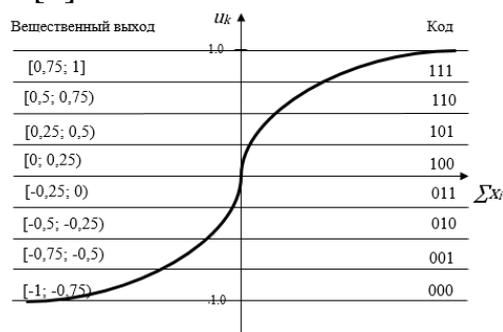


Рис. 2. Кодирование вектора сжатых данных

Сжатые данные, полученные на выходах скрытого слоя, подаются на выходной слой, где происходит восстановление данных. В результате восстановления должно получиться изображение, которое с достаточной точностью повторяет исходное.

Подобные нейронные сети имеют несколько особенностей:

- 1) коэффициент сжатия является постоянной величиной и равен $\frac{m}{n}$;
- 2) сеть может быть разделена на две части: одна производит сжатие изображения, другая – восстановление. Это позволяет использовать сеть в распределенных приложениях.

«Классической» версией сети является вариант с одним скрытым слоем, показанный на рисунке 1, причем выбирают $n = 64$, $m = 16$. Этот вид имеет обозначение 64-16-64, где 64 – размерность входного и выходного слоев, 16 – размерность скрытого слоя. Он обеспечивает сжатие изображения в четыре раза. Обычно при рассмотрении задачи сжатия изображений данная версия используется в качестве базовой.

Традиционно для подобных сетей в качестве функции активации используется сигмоидальная функция [2], обеспечивающая диапазон выходных значений $[-1; 1]$ и позволяющая использовать градиентные методы в силу непрерывной дифференцируемости.

Поскольку для различных изображений диапазон интенсивностей пикселей, вообще говоря, различен, необходимо перед подачей на входы нейронной сети выполнить нормализацию значений входного вектора относительно его длины.

Для инициализации весов сети в системе используется генератор случайных чисел. Существуют и другие методы, которые в некоторых случаях позволяют уменьшить количество локальных минимумов функции ошибки обучения, что приводит к уменьшению циклов обучения (например, метод имитации отжига).

Для обучения нейронной сети используется алгоритм наискорейшего спуска и метод обратного распространения ошибки [2, 3]. Несмотря на большое количество недостатков, среди которых основной – высокая длительность поиска направления минимизации весов, он достаточно хорошо подходит для ре-



шения поставленной задачи. Прочие методы обладают либо большой вычислительной сложностью (алгоритм наискорейшего спуска с адаптируемым шагом, алгоритм сопряженных градиентов), либо большим расходом памяти (алгоритм переменной метрики), что недопустимо для системы работы с изображениями в силу высокой длительности операций доступа к отдельным пикселям изображения и необходимости хранения его отдельных фрагментов.

Так как выходные значения распределены в диапазоне $[0; 1]$, для получения результирующих значений интенсивности необходимо умножить на длину входного вектора.

Сети, созданные в системе, обучались на 50 изображениях с разрешением 256×256 в оттенках серого. Для каждого варианта архитектуры было создано несколько сетей для получения достаточно достоверной оценки среднего количества циклов обучения. Процесс обучения осуществлялся в соответствии с моделью Train/Test/Validation в соотношении 70:20:10, где на стадии Train осуществляется непосредственно процесс обучения, Test – проверка качества обучения, Validation – финальная оценка.

При анализе обучения сети с архитектурой 64-16-64 выяснилось, что количество циклов обучения для каждого изображения сильно зависит от самого изображения. Дольше всего сеть обучалась на изображениях с резкими перепадами интенсивности и большим диапазоном ее значений. На небольшом количестве изображений (трех из общего числа в выборке) нейронные сети не смогли пройти обучение. Для остальных изображений количество циклов обучения распределилось в диапазоне примерно от 3 до 400 тысяч циклов, в среднем составив 110 тысяч циклов.

При работе на реальных данных сети с архитектурой 64-16-64 показали себя не лучшим образом, восстановив изображение с достаточно низкой точностью. На рисунке 3 приведен пример исходного и восстановленного изображений. Восстановленное изображение имеет менее резкие контуры.

Значительно быстрее обучается сеть с конфигурацией 64-4-64, обеспечивающая сжатие в 16 раз. Среднее количество циклов обучения для изображения составило 86 тысяч циклов. Однако вместе с сокращением числа нейронов скрытого слоя сократилось и число связей, поэтому нейронная сеть на реальных данных работает ожидаемо хуже, чем 64-16-64. Такая нейронная сеть слишком быстро «забывает» данные, поступающие в процессе обучения.



Рис. 3. Результат работы сети 64-16-64



Также для примера рассмотрим сеть вида 64-64-64 с коэффициентом сжатия, равным 1. Количество циклов обучения в среднем увеличилось по сравнению с сетью 64-16-64 незначительно и составило около 124 тысяч циклов, несмотря на сильное увеличение количества связей. Такая сеть показывает лучшие результаты, чем базовый вариант: изображение стало более контрастным. Однако точность восстановления все равно остается средней. Результат работы сети 64-64-64 показан на рисунке 4.



Рис. 4. Результат работы сети 64-64-64

Полученные результаты позволяют говорить о том, что нейронные сети с архитектурой «бутылочного горлышка», имеющие один скрытый слой, не позволяют с высокой точностью восстановить изображение. Кроме того, процесс обучения не для всех изображений проходит корректно. Увеличение размера сети приводит к улучшению результата, но при этом увеличиваются временные затраты на обучение, уменьшается степень сжатия, и результат, тем не менее, остается недостаточно близким к исходному.

Из полученных сетей лучшие результаты показала сеть 64-64-64. Однако она не выполняет фактического сжатия и потому ее не следует применять на практике. В случаях, когда не требуется высокой точности восстановления, можно использовать сеть 64-16-64.

Полученные данные говорят о том, что требуются дальнейшие исследования как в области повышения качества процесса обучения, так и в области усложнения архитектуры сети.

Литература

- 1 Watta P. Image Compression Using Backprop [Текст] / P. Watta, B. Desai, N. Dannug, M. Hassoun. – Detroit: Computation and Neural Networks Laboratory, Department of Electrical and Computer Engineering, Wayne State University, 1998. – 3 с.: ил.
- 2 Masalmah Y.M. – Image Compression Using Neural Networks [Текст] / Y.M. Masalmah. – Puerto Rico: Electrical and Computer Engineering Department, University of Puerto Rico, 2008. – 4 с.: ил.
- 3 Siva Nagi Reddy K. Image Compression and Reconstruction Using a New Approach By Artificial Neural Network [Текст] / K. Siva Nagi Reddy, Dr. B.R. Vikram, L. Koteswara Rao, B. Sudheer Reddy // International Journal of Image Processing (IJIP) – Taiwan, 2012. – Volume 6.