

Шумоподавляющий автоэнкодер на основе архитектуры резервуарной нейронной сети LogNNet

П. П. Борисков, К. Е. Рудковский, А. А. Величко

Петрозаводский государственный университет

boriskov@petsu.ru, rudkovskii.k333@gmail.com, velichkogf@gmail.com

Аннотация. В исследовании представлен новый автоэнкодер для сжатия и шумоподавления данных (сигналов и изображений) на основе архитектуры резервуарной нейронной сети LogNNet. LogNNet — сеть прямого распространения, в которой первая (входная) весовая матрица не обучается, а генерируется хаотическими отображениями. LogNNet имеет перспективы для внедрения в мобильных устройствах искусственного интеллекта, поскольку эффективно работает, используя меньшие объемы оперативной памяти. Симметричный резервуарный автоэнкодер по типу LogNNet хорошо восстанавливает зашумленные регулярные сигналы, а также 2D данные в виде зашумленных копий изображения. Он обучается два раза быстрее и требует два раза меньше оперативной памяти по сравнению со стандартной схемой.

Ключевые слова: резервуарные вычисления; автоэнкодер, сжатие данных; шумоподавление

I. ВВЕДЕНИЕ

В настоящее время мы переживаем революцию искусственного интеллекта и нейротехнологий. Нейронные сети производят обработку больших объемов данных [1], выявляют угрозы в сетевой инфраструктуре [2], применяются в мобильных устройствах [3] и системах управления [3]. Искусственный интеллект основанный на нейронных сетях находит применение в концепции «Internet of Things» (IoT) [5] не только на верхних уровнях экосистем (серверы, облачные сервисы), но и на локальном уровне IoT для вычислений на edge устройствах (микроконтроллеры, датчики). Ограниченные ресурсы edge devices, их низкая вычислительная мощность создают трудности для внедрения искусственного интеллекта на низкоразовых платформах и требуют развития специальных архитектур и алгоритмов нейронных сетей.

В работах [6, 7] была разработана архитектура нейронной сети LogNNet и проведено апробирование ее функционирование на плате Arduino UNO с низким объемом ~2kB RAM. LogNNet продемонстрировала хорошие результаты по распознаванию рукописных цифр MNIST с точностью ~82%. Отличительной особенностью архитектуры LogNNet является наличие резервуара, где входная информация преобразуется с помощью хаотического отображения. Поскольку в первом исполнении использовалась логистическое отображение, данная резервуарная архитектура была названа Logistic Neural Network. Хаотическое преобразование трансформирует информацию из одного многомерного пространства в другое, после чего возможна эффективная классификация результата выходным линейным классификатором. Нейронная сеть

LogNNet является сетью прямого распространения, где сигналы распространяются исключительно от входа к выходу, а входную весовую матрицу можно рассматривать как хаотический нелинейный фильтр.

Традиционные методы шумоподавления, такие как цифровые фильтры, пороговое обработка вейвлетов, сингулярное разложение, эффективны для сохранения полезной информации при удалении шума из сигналов. Однако эти методы часто требуют существенного опыта в предметной области, поскольку имеют множество модификаций и параметров. Например, выбор материнского вейвлета и уровня разложения значительно влияет на его эффективность [8].

Автоэнкодеры (АК) являются неотъемлемой частью современных нейросетевых технологий глубокого обучения [9]. Основная функция автоэнкодера в обработке данных есть выделения наиболее значимых признаков, что особенно важно в таких задачах как распознавания классификация, кластеризация и т.п. Малозначимые характеристики данных, как правило, удаляются и далее не участвуют в обработке. Кроме того, сжатие информации АК приводит к удалению в них шумового фона (с нулевым средним), что означает еще одну функцию автоэнкодера – шумоподавление. Шумоподавляющие автоэнкодеры являются простыми и перспективными модулями для повышения производительности IoT приложений и устройств путем шумоподавления наблюдаемых данных, которые состоят из исходных данных и шума [10].

В данном исследовании представлена концепция резервуарного АК в стандартной (трехслойной) архитектуре сети прямого распространения для выделения шума из данных (сигналов и изображений). Данный АК вместо обучаемой первой (кодирующей) матрицы имеет резервуарную матрицу. Вначале проведен сравнительный анализ восстановления узкополосного (амплитудной-модулированного) сигнала из шума с помощью стандартного и резервуарного АК. Далее, рассмотрены результаты шумоподавления с помощью этих АК изображения в градации серого. Во всех вариантах шум был белым и аддитивным. Расчеты выполнены в Python.

II. РЕЗЕРВАУАРНАЯ СЕТЬ LOGNNET

Архитектура нейронной сети LogNNet для распознавания рукописных цифр из базы MNIST показана на рис. 1, где $S=784$ – размер входного массива данных Y , значение P определяет число нейронов в скрытом слое S_h , а $M=10$ – количество вариантов цифр и нейронов выходного слоя S_{out} .

На вход нейронной сети LogNNet подается рукописный образ размером 28×28 пикселей из базы MNIST [24]. Изображение представлено в градации серого (8 бит), черному цвету соответствует значение пикселя 255, а белому 0. Изображение шаблона (цифры) преобразуется в линейный массив Y используя T-pattern преобразованием по столбцам. Нулевой элемент массива $Y[0]=1$ является нейроном смещения. Вектор Y подается на резервуар и преобразует информацию согласно (1) и далее линейный классификатор классифицирует результат согласно (2):

$$S_h = fh(W_1 \cdot Y), \quad (1)$$

$$S_{out} = fout(W_2 \cdot S_h). \quad (2)$$

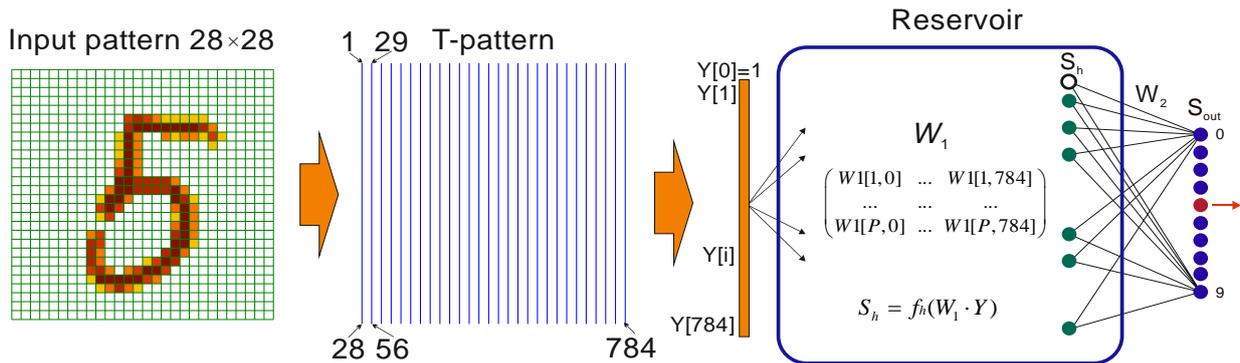


Рис. 1. Архитектура нейронной сети LogNNet для распознавания рукописных цифр из базы MNIST

Весовые коэффициенты W_1 задаются с помощью хаотического отображения, нумерация элементов матрицы $W_1[j,i]$ имеет пределы $j=1 \dots P, i=0 \dots 784$:

$$W_1 = \begin{pmatrix} x_1 & x_2 & \dots & x_n \\ x_{n+1} & \rightarrow & \dots & \cdot \\ \dots & \rightarrow & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \end{pmatrix}. \quad (3)$$

Весовые коэффициенты W_2 обучаются методом обратного распространения ошибки с использованием обучающего множества изображений MNIST. Тестирование точности распознавания LogNNet производится прогоном тестового множества MNIST через нейронную сеть.

III. АВТОЭНКODER НА ОСНОВЕ LOGNNET: ШУМОПОДАВЛЕНИЕ РЕГУЛЯРНОГО СИГНАЛА

На рис. 2 представлен простейший (трехслойный) автоэнкодер как сеть прямого распространения, в котором выход обучается повторить вход, через сжимающий скрытый слой («бутылочное горлышко»). С помощью первой весовой матрицы W_1 входной сигнал кодируется в выход скрытого слоя, а с помощью второй весовой матрицы W_2 и нейронов выходного слоя жатый сигнал восстанавливается (декодируется).

В случае фиксированной случайной матрицы W_1 АК функционирует как резервуарная сеть LogNNet, где настраиваются только веса выходной матрицы W_2 . В начале в качестве тестовых примеров использовались сигналы в виде суммы зашумленных синусоид с разными частотами. Эффективность шумоподавления оценивалась отношением MSE (mean square error) входного сигнала к выходному (MSE_R):

В формулах (1) и (2) S_h – вектор нейронов скрытого слоя, находящиеся в резервуаре, S_{out} – вектор нейронов выходного слоя, Y – вектор входных данных, W_1, W_2 – матрицы весовых коэффициентов, fh – тождественная функция активации, с последующим нормированием значений в диапазон от -0.5 до 0.5, а f_{out} – логистическая функция активации. Максимальные, минимальные и средневзвешенные значения S_h являются нормировочными значениями функции fh и вычисляются по первым 1000 элементам обучающего множества.

$$MSE_R = \frac{MSE^{(in)}}{MSE^{(out)}} = \frac{\sum_{i=1}^n (s^{(in)}[i] - s_o[i])^2}{\sum_{i=1}^n (s^{(out)}[i] - s_o[i])^2}. \quad (2)$$

Примеры шумоподавления с помощью стандартного АК входного сигнала при низком и высоком уровне шума показаны на рис. 3. Зависимости параметра шумоподавления MSE_R от числа эпох для стандартного и резервуарного АК представлены на рис. 4.

Эти результаты показывают, что для коротких серий стандартный и резервуарный АК дают схожую эффективность (рис. 4а). Для длинных серий (рис. 4б) стандартный АК превосходит резервуарный примерно в 4 раза.

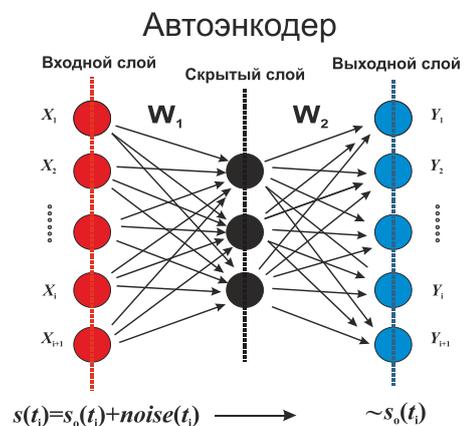


Рис. 2. Общая схема АК. В архитектуре автоэнкодера по типу резервуарной сети LogNNet матрица весов W_1 случайна и фиксирована.

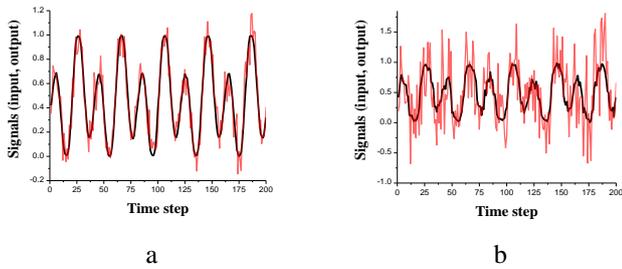


Рис. 3. Осциллограммы входного $s(t)$ (зашумленного, красная кривая) и выходного (шумоподавленного, черная кривая) аддитивных синусоидальных сигналов с частотами $f_0=250$ и $f_1=500$ низкого и высокого шум. а — Низкий шум с дисперсией 0.01, $\mu=0.01$ и $MSE_R=34.6$. б — Высокий шум с дисперсией 0.16, $\mu=0.001$ и $MSE_R=28.94$. Число нейронов во входном и выходном слоях $N=400$, число нейронов в скрытом слое $N_h=200$, общая длина серий $N_{learn} = 2000$, epoch=1000.

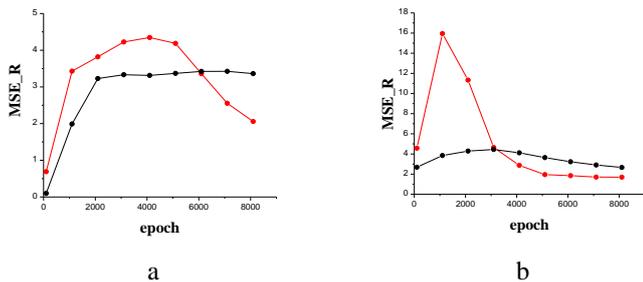


Рис. 4. Зависимости параметра шумоподавления MSE_R от числа эпох обучения стандартного АК (красная кривая) и АК по типу LogNet (черная кривая) для двух длин обучающих серий N_{learn} . Число нейронов во входном и выходном слоях $N=400$, число нейронов в скрытом слое $N_h=200$, $\mu=0.01$. а — $N_{learn}=500$, б — $N_{learn}=2000$.

IV. АВТОЭНКODER НА ОСНОВЕ LOGNET: ШУМОПОДАВЛЕНИЕ ИЗОБРАЖЕНИЙ

Для шумоподавления изображений рассмотрим аналогичную схему АК (рис. 2а), где пиксели входного и выходного изображения развернуты (flatten) в вектора.

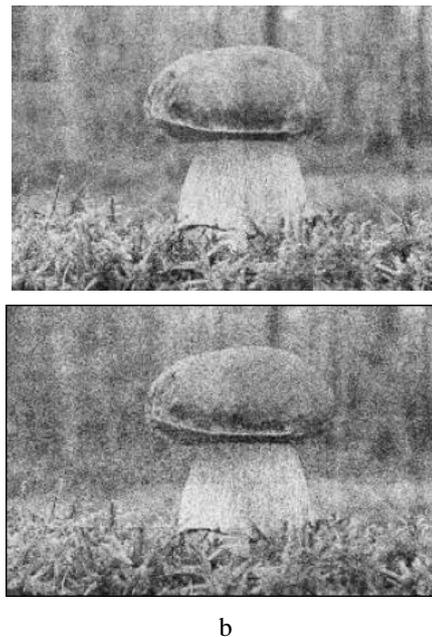
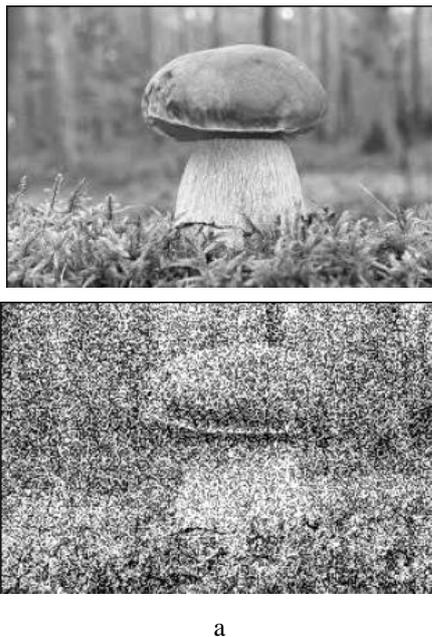


Рис. 5. Эталонное (верх) и зашумленное (низ) изображение гриба (а) и результат его восстановления (б) с помощью стандартного (верх) и резервуарного (низ) АК

На рис. 5 представлены варианты шумоподавления изображения гриба с помощью АК. Размер изображения (183×275), то есть примерно 50 000 пикселей, равное числу входных и выходных нейронов. Коэффициент сжатия скрытого слоя варьировался в пределах 0.1–0.5. Для создания обучающей выборки генерировалось различное зашумление эталона (рис. 5а, верх), которое достигало 100 шаблонов (пример шаблона показан на рис. 5а, низ).

Поскольку в данном случае число нейронов АК огромно, желательно выполнять вычисления поблочно, то есть делить общее изображение на неперекрывающиеся блоки и для каждого из них выполнять операцию шумоподавления. Также следует контролировать количество эпох обучения (не более 50) для learn gate 0.001, чтобы не было переобучения. Кроме того, для снижения степени зернистости восстановленного изображения необходимо в резервуарном АК уменьшить (до 0.1) показатель экспоненты функций активаций (сигмоид) нейронов скрытого слоя.

Как видно, резервуарный (трехслойный) АК хорошо справляется с восстановлением изображения, параметр шумоподавления MSE_R достигает почти 10. Результат стандартного АК чуть лучше со значением $MSE_R = 12.95$. В табл. 1 даны параметры обучения стандартного и резервуарного АК.

ТАБЛИЦА I. ПАРАМЕТРЫ ОБУЧЕНИЯ ПРИ ШУМОПОДАВЛЕНИИ ИЗОБРАЖЕНИЙ С ПОМОЩЬЮ АК

Параметры обучения	Стандартный АК	Резервуарный АК
Число эпох/ Learning rate	20/0.001	20/0.001
Число обучающих шаблонов	60	60
Размер блока	50×50	70×70
Размер входного/скрытого слоя	2500/1000	4900/3920

V. ЗАКЛЮЧЕНИЕ

Как показывают результаты шумоподавления сигналов и изображений, резервуарный АК может конкурировать со стандартным АК для коротких сигналов и небольшого числа шаблонов, обеспечивая более быстрое обучение и экономию оперативной памяти (в 2 раза). Это делает его перспективным для устройств с ограниченными ресурсами, например, мобильных приложений IoT.

СПИСОК ЛИТЕРАТУРЫ

- [1] S. K. Baliarsingh, S. Vipsita, A. H. Gandomi, A. Panda, S. Bakshi and S. Ramasubbareddy, "Analysis of high-dimensional genomic data using MapReduce based probabilistic neural network." *Comput. Methods Programs Biomed.*, vol.195, p. 105625, 2020.
- [2] I. H. Sarker, "Deep Cybersecurity: A Comprehensive Overview from Neural Network and Deep Learning Perspective," *SN Comput. Sci.*, vol. 2, p. 154, 2021.
- [3] Z. Obermeyer and E. J. Emanuel, "Predicting the Future — Big Data, Machine Learning, and Clinical Medicine," *N. Engl. J. Med.*, vol.375, p. 1216–9, 2016.
- [4] J. G. Erdmann, V. Hummela, K. Leipzig von and J. Schuhmacher, "Development and implementation of an autonomous control system for target-optimised use of intralogistics transport systems in the Learning Factory Werk 150 at Reutlingen University," *Procedia Manuf.*, vol.45, 405–10, 2020.
- [5] S. Hamdan, M. Ayyash and S. Almajali, "Edge-Computing Architectures for Internet of Things Applications: A Survey," *Sensors*, vol. 20, 2020.
- [6] A. Velichko, "Neural Network for Low-Memory IoT Devices and MNIST Image Recognition Using Kernels Based on Logistic Map," *Electronics*, vol. 9, p. 1432, 2020.
- [7] Y. A. Izotov, A. A. Velichko and P. P. Boriskov, "Method for fast classification of MNIST digits on Arduino UNO board using LogNNNet and linear congruential generator," *J. Phys. Conf. Ser.*, vol.2094, p. 32055, 2021.
- [8] S. Solbo and T. Eltoft, "Homomorphic wavelet-based statistical despeckling of sar images," *IEEE Trans Geosci Remote Sens*, vol. 42(4), p. 711-721, 2004.
- [9] K. Bajaj, D. K. Singh and M. A. Ansari, Autoencoders Based Deep Learner for Image Denoising. *Procedia Computer Science*, vol. 171, p. 1535–1541, 2020.
- [10] Y. Sun, M. Peng, Y. Zhou, Y. Huang and S. Mao, "Application of machine learning in wireless networks: Key techniques and open issues," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 4, pp. 3072–3108, 2019.