

А. В. Шостак

Національний аерокосмічний університет імені М.Є. Жуковського, Харків, Україна

ПРО ОСОБЛИВОСТІ ФОРМУВАННЯ ВХІДНИХ ДАНИХ У СІАМСЬКІЙ НЕЙРОННІЙ МЕРЕЖІ

Анотація. Проаналізовано різні способи формування вхідних даних та оцінок дескрипторів сіамської нейронної мережі (СНМ) для порівняння зображень рукописних цифр. Запропоновано спосіб використання квазівипадкової N -вимірної послідовності векторів, сформованих відповідно до методу Соболя, для формування дескрипторів, які, спільно із зображеннями, є вхідними даними для навчання СНМ та її подальшого використання. Виконано тестування сіамської нейронної мережі з використанням отриманих оцінок дескрипторів зображень рукописних цифр. Під час тестування СНМ використовувався набір MNIST. Результат тестування моделі СНМ дав значення показника асигасу, що дорівнює 0.9706. Результати тестування засвідчили, що розглянуті оцінки дескрипторів h_1 і h_2 зменшують кількість помилок під час тестування порівняно з використанням дескриптора h_0 на основі квазівипадкової N -вимірної послідовності векторів.

Ключові слова: сіамська нейронна мережа, дескриптор, тестування нейронної мережі.

Вступ

Сіамська нейронна мережа (СНМ) зазвичай містить три підмережі [1, 2]. Перші дві підмережі однакові, їхні входи є входами всієї СНМ і вони обчислюють представлення входів мережі. Остання підмережа виконує порівняння представлень і видає результат роботи

СНМ у вигляді числа в діапазоні від 0 до 1. Сферою використання СНМ є системи, у яких виконується порівняння вхідних даних на ідентичність [1–3].

СНМ формує значення показника, яке оцінює ступінь близькості вхідних даних [1–4].

У відомих виданнях не повною мірою описано особливості формування вхідних даних і використання представлень для навчання і тестування СНМ.

Основна частина

Узагальнену структуру СНМ представлено й описано в [1, 2].

Вхідні дані x_1 і x_2 за допомогою перших двох підмереж перетворюються на N -вимірні представлення (або дескриптори) $h(x_1)$ та $h(x_2)$. Перші дві підмережі ідентичні за своїми параметрами W . Третя підмережа СНМ обчислює евклідову відстань між представленнями вхідних даних і на підставі неї формує значення оцінки ступеня подібності в діапазоні від 0 до 1.

Пропонується для навчання СНМ та її тестування використовувати архітектуру тільки з однією вхідною підмережею (рис. 1). На один вхід СНМ надходить N -вимірний дескриптор j -ї цифри h_0j , на

другий вхід – відображення цифри x з бібліотеки MNIST [5].

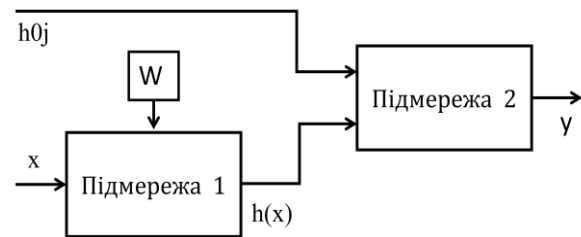


Рис. 1. Архітектура СНМ під час навчання і тестування

Підмережі 1 і 2 на рисунку 1 аналогічні відповідним підмережам узагальненої структури СНМ [2]. Дескриптор $h(x)$ є поданням вхідного зображення x . Дескриптори для десяти цифр h_0 є елементами квазівипадкової N -вимірної послідовності з десяти векторів, сформованої відповідно до методу Соболя [6, 7].

Було проаналізовано квазівипадкові N -вимірні послідовності векторів, що генеруються за методами Холтона [8], Соболя [6, 7] і R-послідовність [7]. N -вимірні послідовності векторів за методом Соболя показала найбільше значення за показником середня відстань між векторами. Тому послідовність векторів за методом Соболя і була використана для формування десяти дескрипторів цифр, які подаються на перший вхід СНМ (рис. 1) і використовуються для навчання СНМ. Перші десять 10-вимірних векторів квазівипадкової послідовності Соболя, згенеровані з використанням бібліотеки SciPy, мають такий вигляд:

- 0) [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
- 1) [0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5]
- 2) [0.75 0.25 0.25 0.25 0.75 0.75 0.25 0.75 0.75 0.75]
- 3) [0.25 0.75 0.75 0.75 0.25 0.25 0.75 0.25 0.25 0.25]
- 4) [0.375 0.375 0.625 0.875 0.375 0.125 0.375 0.875 0.875 0.625]
- 5) [0.875 0.875 0.125 0.375 0.875 0.625 0.875 0.375 0.375 0.125]
- 6) [0.625 0.125 0.875 0.625 0.625 0.875 0.125 0.125 0.125 0.375]
- 7) [0.125 0.625 0.375 0.125 0.125 0.375 0.625 0.625 0.625 0.875]
- 8) [0.1875 0.3125 0.9375 0.4375 0.5625 0.3125 0.4375 0.9375 0.9375 0.3125]
- 9) [0.6875 0.8125 0.4375 0.9375 0.0625 0.8125 0.9375 0.4375 0.4375 0.8125]

Навчання СНМ та її тестування проводилося за допомогою Google Colaboratory. Основні складові підмереж 1 і 2 (рис. 1) докладно описані в [2, 4].

При компіляції моделі СНМ використовувалася функція контрастних втрат [3, 4], основні параметри функції fit навчання СНМ - розмір пакета рівний 16 і число епох рівне 10.

Результат тестування моделі СНМ за допомогою методу evaluate такий - значення показника assiguas дорівнює 0.9706, а loss - 0.0226.

Після навчання СНМ з тренувального набору зображень рукописних цифр на виході підмережі 1 (рисунок 1) було отримано кластери 10-вимірних дескрипторів цифр, а також на основі цих кластерів обчислено оцінки дескрипторів цифр.

Оцінки дескрипторів цифр виконувались в такий спосіб [2]:

1. h_1 – дескриптор, що забезпечує мінімум суми евклідових відстаней на відповідному кластері,

2. h_2 – дескриптор, який є рішенням методу Нелдера-Міда на відповідному кластері.

Наприклад, дескриптор h_{10} має такий вигляд:

$h_{10} = [-0.9991518 \quad -0.99878484 \quad -0.9936094 \quad -0.9974742 \quad 0.99989784 \quad -0.9944306 \quad -0.9998869 \quad -0.97155184 \quad -0.9997868 \quad -0.97745085]$.

Для пошуку дескрипторів h_2 було використано метод Нелдера-Міда з Python-бібліотеки SciPy. Дескриптор h_{20} для цифри 0 має такий вигляд:

$h_{20} = [-0.99668455 \quad -0.9942764 \quad -0.99217147 \quad -0.9915273 \quad 0.9961964 \quad -0.9922386 \quad -0.9978741 \quad -0.9734324 \quad -0.9987575 \quad -0.9803164]$.

У табл. 1 наведено розраховані з використанням дескриптора h_2 характеристики кластерів дескрипторів цифр - мінімальна відстань (r), максимальна відстань (радіус кластера) між дескрипторами

кластера і дескриптором h_2 (R), середня всерединікластерна відстань (ssR) і середнє квадратичне відхилення (sD). Кількість даних у табл. 1 для цифр від 0 до 9 відповідно така - 2961, 3423, 2948, 3073, 2926, 2709, 2975, 3107, 2875, 3003.

Кластер дескрипторів зображень цифри 9 найбільш некомпактний за показником R (за радіусом кластера). Кластер дескрипторів зображень цифри 3 найнекомпактніший за показником ssR (за середньою всерединікластерною відстанню). Найкомпактніший за показником ssR кластер дескрипторів цифри 0. Найкомпактніший за показником R кластер дескрипторів зображень цифри 1.

Тестування роботи дескрипторів h_0 , h_1 і h_2 виконувалося в архітектурі СНМ з рисунка 1 за методикою, наведеною в [2].

Таблиця 1 – Характеристики кластерів дескрипторів цифр

Цифра	r	R	ssR	sD
0	0.00996	5.17015	0.30645	0.58298
1	0.23036	4.00553	0.97232	0.55612
2	0.03950	4.60219	0.89012	0.97916
3	0.09604	5.32569	1.16170	0.87623
4	0.07548	4.04438	0.92428	0.70884
5	0.08129	4.81103	1.05936	0.92620
6	0.04481	4.22272	0.69917	0.74386
7	0.03546	4.87599	1.00177	0.84371
8	0.15787	4.46806	1.10556	0.82220
9	0.03347	5.87126	0.73154	0.86083

У табл. 2 представлені підсумки роботи СНМ при подачі на другий вхід даних для цифри 0 з тестового набору з використанням дескрипторів h_0 , h_1 і h_2 для цифр від 0 до 9.

Таблиця 2 – Результати тестування на тестовому наборі зображень із цифрою 0 з використанням дескрипторів h_0 , h_1 і h_2 для цифр від 0 до 9

Цифра	0	1	2	3	4	5	6	7	8	9
Помилки для h_0	22	18	5	8	3	0	1	5	0	0
y	3.86e-05	0.96	1.00	0.99	0.94	0.42	0.91	0.99	0.43	0.38
Помилки для h_1	16	32	1	0	0	1	0	1	0	0
y	5.43e-06	1.00	0.92	0.05	0.07	0.74	0.49	0.94	0.02	0.00
Помилки для h_2	15	35	1	0	0	1	0	1	0	0
y	5.58e-06	0.99	0.93	0.07	0.06	0.76	0.48	0.95	0.04	0.00

У стовпчику з цифрою 0 показано результат тестування, коли на перший вхід СНМ (рис. 1) подають значення дескриптора h_{00} для цифри 0 (рядки 2, 3), а на вхід x – дані для цифри 0. Цифра 22 у другому рядку стовпчика 0 означає, що для h_{00} вихід СНМ для 22 зображень цифри 0 з 980 був меншим за 0.5, причому 3.86e-05є мінімальне значення виходу СНМ. Тобто для 22 зображень цифри 0 винесено рішення, що вони не відповідають дескриптору h_{00} цифри 0. На рис. 2 показана цифра з міткою 0, що відповідає значенню $y=0.00020055$, коли на першому вході дескриптор h_{00} .

Цифра 18 у другому рядку стовпчика з цифрою 1 показує, що для h_{01} вихід СНМ для 18 зображень цифри 0 з 980 був більшим за 0.5, максимальне значення y дорівнює 0.96. Тобто для 18 зображень ви-

несено рішення, що вони відповідають дескриптору h_{01} цифри 1.

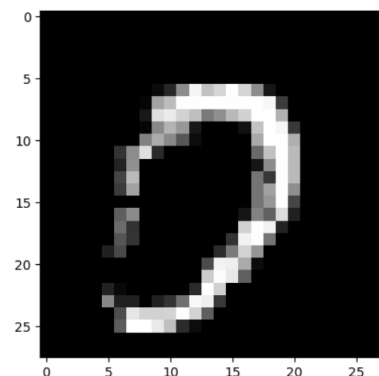


Рис. 2. Цифра з міткою 0 (для $h_{00} y = 0.00020055$)

На рис. 3 показана цифра з міткою 0 для значення $y=0.9560281$, коли на першому вході дескриптор h_{01} .

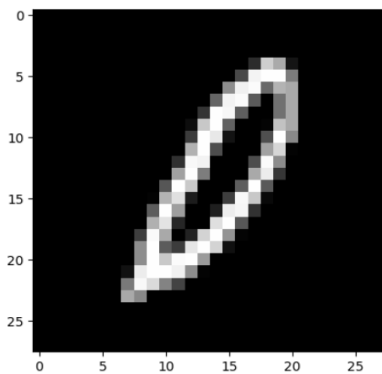


Рис. 3. Цифра з міткою 0 (для $h_{01} y = 0.9560281$)

Таким чином, за використання дескриптора h_0 було 62 помилки на 980 зображень цифри 0, що становить 6.3%, за використання дескриптора h_1 - 51 помилка (5.2%) і за використання дескриптора h_2 - 53 помилки (5.4%).

Наведені в табл. 2 результати тестування зображень із цифрою 0 показують, що розглянуті оцінки дескрипторів h_1 і h_2 значно зменшують кількість помилок під час тестування.

Висновки

У роботі запропоновано спосіб використання квазिवипадкової N -вимірної послідовності векторів, сформованих відповідно до методу Соболя, для

формування дескрипторів, які, спільно із зображеннями, є вхідними даними для навчання СНМ та її подальшого використання.

У роботі також проаналізовано способи побудови оцінок дескрипторів для порівняння і розпізнавання зображень для СНМ.

Використана для навчання і тестування архітектура СНМ складається з двох підмереж. Перша підмережа формує уявлення вхідного зображення, що аналізується. На перший вхід другої підмережі подають дескриптор, який використовують під час навчання або отримані оцінки дескрипторів, а на другий вхід - дескриптор зображення, що аналізується.

Під час тестування СНМ використовувався набір MNIST. Результат тестування моделі СНМ дав значення показника accuracy, що дорівнює 0.9706.

Результати тестування засвідчили, що розглянуті оцінки дескрипторів h_1 і h_2 зменшують кількість помилок під час тестування порівняно з використанням дескриптора h_0 на основі квазिवипадкової N -вимірної послідовності векторів.

Під час випробування 980 зображень із цифрою 0 у разі використання дескриптора h_0 кількість помилок становила 62, за використання дескриптора h_1 - 51 помилка і за використання дескриптора h_2 - 53 помилки.

У подальших дослідженнях слід приділити увагу вибору та аналізу впливу показників квазिवипадкової N -вимірної послідовності дескрипторів на якість навчання та на точність порівняння зображень у сіамській нейронній мережі.

СПИСОК ЛІТЕРАТУРИ

1. Chicco D. Siamese Neural Networks: An Overview. *Artificial Neural Networks*. 2021. MIMB, vol. 2190. P. 73–94. URL: https://link.springer.com/protocol/10.1007/978-1-0716-0826-5_3
2. Шостак А. В. Про особливості формування дескрипторів у сіамській нейронній мережі. *Системи управління, навігації та зв'язку*. Полтава: НУ ПП, 2021. Вип. 4(66). С. 91–96. DOI: <https://doi.org/10.26906/SUNZ.2021.4.079>
3. Contrastive loss for Siamese networks with Keras and TensorFlow. URL: <https://www.pyimagesearch.com/2021/01/18/contrastive-loss-for-siamese-networks-with-keras-and-tensorflow/>
4. Image similarity estimation using a Siamese Network with a contrastive loss. URL: https://keras.io/examples/vision/siamese_contrastive/
5. The Mnist database of handwritten digits. URL: <http://yann.lecun.com/exdb/mnist/>
6. Owen, A.B. On Dropping the First Sobol' Point. In: Keller, A. (eds) Monte Carlo and Quasi-Monte Carlo Methods. MCQMC 2020. *Springer Proc. in Mathematics & Statistics*, vol 387. Springer, Cham. DOI: https://doi.org/10.1007/978-3-030-98319-2_4
7. Roberts M. The Unreasonable Effectiveness of Quasirandom Sequences. 2020. URL: <https://extremelarning.com.au/unreasonable-effectiveness-of-quasirandom-sequences/>
8. Halton J. H. On the efficiency of certain quasi-random sequences of points in evaluating multidimensional integrals. *Numer. Math.* 1960. Vol. 2. P. 84–90. DOI: <https://doi.org/10.1007/BF01386213>

Received (Надійшла) 29.05.2024

Accepted for publication (Прийнята до друку) 21.08.2024

On the features of the formation of input data in the Siamese neural network

A. Shostak

Abstract. Various methods of forming input data and descriptor estimates of the Siamese neural network (SNM) for comparing images of handwritten digits are analyzed. A method of using a quasi-random N -dimensional sequence of vectors, formed according to Sobol's method, is proposed for the formation of descriptors, which, together with images, are input data for training SNM and its further use. Siamese neural network testing was performed using the received evaluations of image descriptors of handwritten digits. The MNIST set was used during SNM testing. The result of testing the SNM model gave a value of the accuracy indicator equal to 0.9706. The test results showed that the considered estimates of the descriptors h_1 and h_2 reduce the number of errors during testing compared to the use of the descriptor h_0 based on a quasi-random N -dimensional sequence of vectors.

Keywords: Siamese neural network, descriptor, neural network testing.