

Н.Л. Дорош, Т.М. Фененко

## **ДОСЛІДЖЕННЯ ДЕСКРИПТОРІВ ЩОДО РОЗПІЗНАВАННЯ ЦИФР НАБОРУ MNIST**

*Анотація. Кращі результати розпізнавання цифр отримані на основі нейронних мереж і мають помилку менше 1%. Успішні алгоритми розпізнавання, в тому числі і глибокого навчання, приховані від користувача і складні в описі, тому не втратили свою актуальність алгоритми на основі дескрипторів. Метою роботи є вибір та дослідження дескрипторів для розпізнавання набору MNIST. Виконано розпізнавання цифр на основі 12 дескрипторів із застосуванням моделей з бібліотеки Scikit-Learn Python. За результатами розпізнавання методом k-середніх з'ясовано, що доцільно обрати 8 дескрипторів.*

*Ключові слова: розпізнавання, цифри рукопису MNIST, дескриптори, Ху-моменти, гістограми, Python, Scikit-Learn, метод k-середніх*

**Постановка проблеми.** Робота присвячена вирішенню задачі розпізнавання цифр. Актуальність дослідження пов'язана з практичною проблемою і представляє інтерес в багатьох сферах діяльності людини. Наприклад, існує проблема розпізнавання номерів автомобілів, вагонів та інших об'єктів, які отримані з використанням відеозйомки. Але безлічі зображень, що представляють інтерес в тому, або іншому додатку, не належать до класу множин, ретельно вивчених у процесі багатовікових математичних досліджень. Істотні позитивні результати з розпізнавання зображень не можуть бути отримані на підставі лише загальних рекомендацій теорії розпізнавання. Необхідно вивчати і враховувати виняткову специфікацію зображення, як об'єкта формального аналізу [1,2]. В роботі проведено аналіз дескрипторів, які були використані в вирішенні задачі розпізнавання цифр рукопису.

**Аналіз останніх досліджень і публікацій.** На підставі результатів двадцятирічних досліджень можна переконатися, що проблема розпізнавання цифр хоча і добре вивчена, але до теперішнього часу представ-

ляє чималий інтерес. Є велика кількість робіт, присвячених розпізнаванню рукописних цифр, які представлені зображеннями [3-7]. При цьому використовуються різні алгоритми класифікації, застосування яких складається з наступних етапів [8]:

1-отримати дані (вивчити, проаналізувати і візуалізувати для розуміння їх сутності);

2 - зробити попередню обробку даних: масштабування, перетворення даних;

3 - вибрати модель розпізнавання;

4 - навчити модель на навчальних даних;

5 - перевірити модель на тестовому наборі даних;

6 - оцінити точність роботи моделі.

Таким чином, перш, ніж перейти до застосування алгоритму, необхідно провести збір і підготовку даних. Цей процес досить трудомісткий. Використання наборів даних NIST і MNIST [9] дозволило спростити роботу з підготовки даних і одразу перейти до алгоритмів розпізнавання. У набори MNIST кожна цифра є зображенням, яке являє собою квадрат 28 на 28 пікселів. Усього набір містить 70 000 зображень, з яких 60 000 використовується для навчання моделі, а 10 000 - для її перевірки. Є десять цифр або десять класів. В результаті розпізнавання видається помилка, яка є оцінкою точності.

В [6] показано зведену таблицю результатів класифікації різними алгоритмами на наборі MNIST. Кращі результати розпізнавання мають помилку менше 1%. Вони отримані із застосуванням великих згортальних нейронних мереж [7]. Успішні алгоритми розпізнавання, в тому числі і глибокого навчання, приховані від користувача і складні в описі, тому не втратили свою актуальність алгоритми розпізнавання на основі дескрипторів.

**Мета дослідження.** Метою роботи є дослідження дескрипторів і зменшення їх кількості для розпізнавання цифр MNIST з застосуванням бібліотек системи Python.

**Викладення основного матеріалу дослідження.** При виборі найбільш інформативних ознак необхідно враховувати як властивості

самих об'єктів, так і можливості роздільної здатності первинних формувачів сигналу зображення. Важливою особливістю більшості геометричних ознак є інваріантність щодо розвороту зображення об'єкта, а шляхом нормування геометричних ознак досягається інваріантність щодо масштабу зображення об'єкту. Для розпізнавання цифр рукопису MNIST було обрано набір з 12 дескрипторів, які запропоновані [10] стосовно розпізнавання зображень, а саме: сім перших моментів  $H_1, H_2, H_3, H_4, H_5, H_6, H_7$  [11] і ще п'ять -  $E_n$  (число Ейлера),  $E_x$  (коефіцієнт заповнення),  $E_c$  (ексцентриситет),  $Y_n, X_n$  (координати центру ваги), які обрані на підставі аналізу морфологічних ознак об'єкту [12].

Для розширення інформації про властивості дескрипторів проведено аналіз їх гістограм, які наведені на рисунку 1 (для навчального набору) і рисунку 2 (для тестового набору).

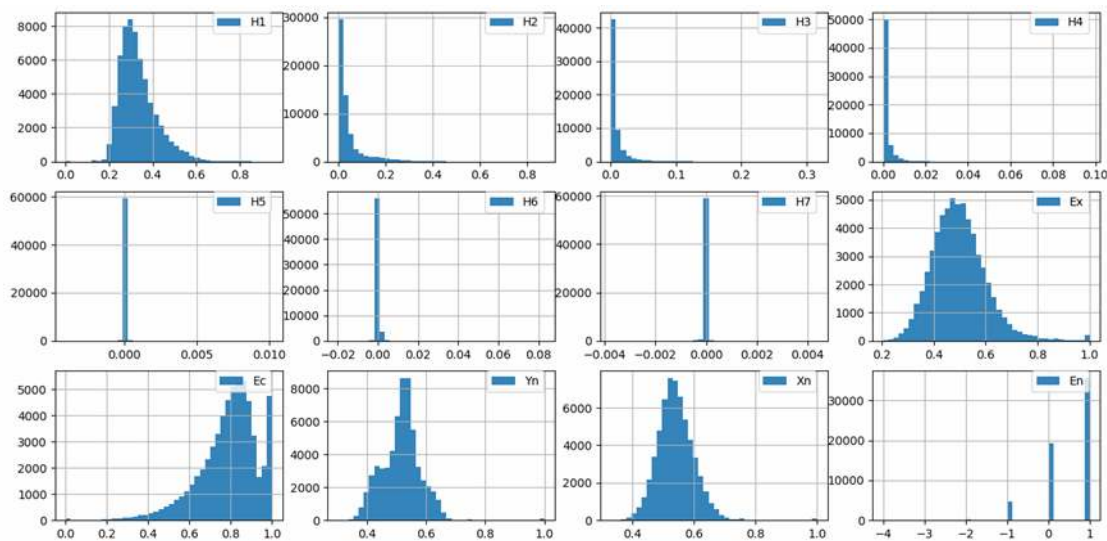


Рисунок 1 – Гістограми дескрипторів для навчального набору (60000)

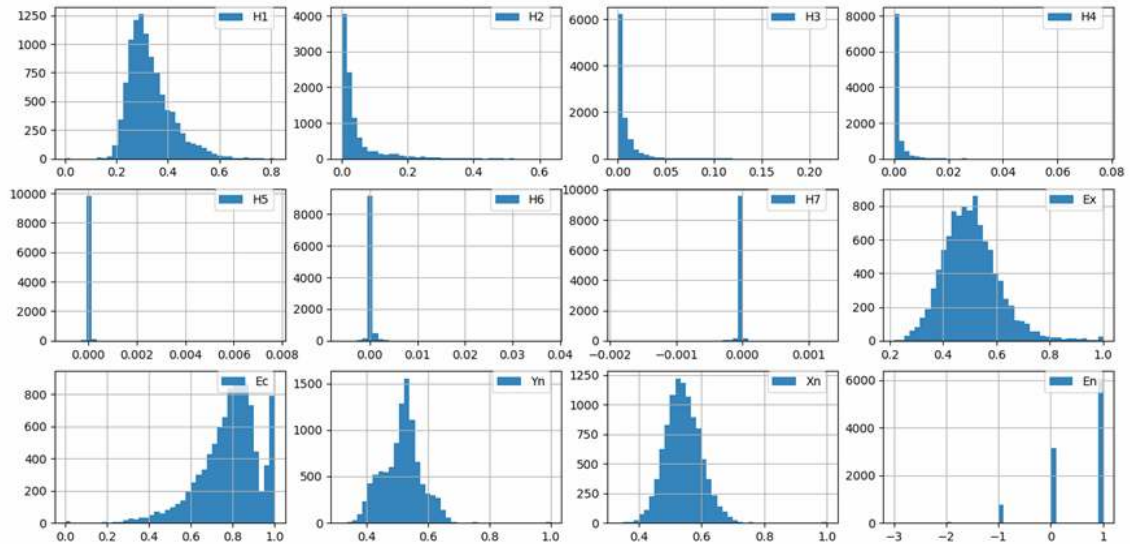


Рисунок 2 – Гістограми дескрипторів для тестового набору (10000)

На основі аналізу гістограм дескрипторів, який можна вважати попереднім, з'ясовано:

- гістограми відповідних дескрипторів в навчальному та тестовому наборах даних мають схожу конфігурацію;
- розподіл даних на гістограмах відрізняється від нормального, отже, для розпізнавання не підходить алгоритм «Наївний Байес»;
- значення дескриптора H5 практично для всіх точок нульове;
- небагато значень дескрипторів H6 і H7 відрізняються від нуля.

З'ясуємо, скільки зображень з навчального та тестового набору цифр відповідають таким значенням:  $H5 > 0,005$ ;  $H6 > 0,02$ ;  $\text{abs}(H7) > 0,0037$  (цифрові значення обрані інтуїтивно по гістограмі навчального набору даних).

Таким чином, у наборі рукописних цифр для навчання і тестування (70000) є мала кількість зображень (7), для яких значення дескрипторів H5, H6, H7 з певною точністю відрізняється від нуля. Можна припустити, що ці сім зображень є шумом (брак).

Проведено розпізнавання цифр із застосуванням класифікатора на основі методу k-середніх з  $n\_neighbors = 10$  бібліотеки Scikit-Learn системи Python.

Точність моделі з набором 12 дескрипторів на тестових даних становить 78,15%.

Для того, щоб дізнатися, де модель помиляється можна використувати матрицю відмінностей (рисунок 3), обчислити яку можна за допомогою бібліотеки Scikit-Learn, а візуалізувати за допомогою Seaborn. В роботі точність розпізнавання виконана з використанням перехресної перевірки.



Рисунок 3 – Матриця відмінностей

Проведемо додаткові дослідження з використанням методу k-середніх для перевірки розумності вибору всіх 12 дескрипторів. Результати розпізнавання схематично показані на рисунку 4. З рисунку 4 видно, що точність розпізнавання на основі 12 дескрипторів становить 78,14% і є найвищою серед наявних варіантів.

Попередній аналіз дескрипторів, давав підстави припустити, що дескриптори H5, H6 і H7 не вносять вклад в якість розпізнавання заданого тестового набору цифр методом k-середніх. Це припущення виправдано. Після додавання цих дескрипторів точність залишалася 46,37%. Також необхідно виключити з набору Eс. Після додавання Eс точність зменшилась до 47,44% (була 49,18%).

На рисунку 5 наведені результати розпізнавання методом k-середніх для всіх 12 дескрипторів, з набором з 9 дескрипторів (без H5, H6, H7) і з набором з 8 дескрипторів (без H5, H6, H7 і Eс).

<u>KNeighborsClassifier</u>
'H1 Точность - 18.2%'
'H1 H2 Точность - 29.02 %'
'H1 H2 H3 Точность - 44.1%'
'H1 H2 H3 H4 Точность - 46.37%'
'H1 H2 H3 H4 H5 Точность - 46.37%'
'H1 H2 H3 H4 H5 H6 Точность - 46.37%'
'H1 H2 H3 H4 H5 H6 H7 Точность - 46.37%'
'H1 H2 H3 H4 H5 H6 H7 Ex Точность - 49.18%'
'H1 H2 H3 H4 H5 H6 H7 Ex Ec Точность - 47.44%'
'H1 H2 H3 H4 H5 H6 H7 Ex Ec Yn Точность - 64.14
'H1 H2 H3 H4 H5 H6 H7 Ex Ec Yn Xn Точность - 70.11%'
'H1 H2 H3 H4 H5 H6 H7 Ex Ec Yn Xn En Точность - 78.14%'

Рисунок 4 – Результати розпізнавання з різним набором дескрипторів

Як видно з рисунку 5 точність розпізнавання тестового набору цифр на підставі набору H1, H2, H3, H4, Ex, Yn, Xn, En (виключили H5, H6, H7, Ec) навіть поліпшилась і становить 78,58%. Таким чином в набір дескрипторів увійшло 8 елементів замість 12.

'H1 H2 H3 H4 H5 H6 H7 Ex Ec Yn Xn En Точность - 78.14%'
'H1 H2 H3 H4 Ex Ec Yn Xn En Точность - 78.14%'
'H1 H2 H3 H4 Ex Yn Xn En Точность - 78.58%'

Рисунок 5 – Результати розпізнавання з вибіркоvim набором дескрипторів

**Висновки.** Проведено дослідження дескрипторів, визначені статистичні характеристики та гістограми щодо наборів цифр набору MNIST (навчання та тестування). На основі досліджень зроблено ряд припущень.

Проведено розпізнавання рукописних цифр набору MNIST з використанням класифікатора на основі методу k-середніх з  $n\_neighbors = 10$ . Використано набори дескрипторів з 12, які було обрано. Точність розпі-

знавання –78,14%, оцінка точності проведена з використанням перехресної перевірки.

Допущення про виключення з набору дескрипторів Н5, Н6, Н7 підтвердилися, було також з'ясовано, що необхідно виключити з набору Ес.

Таким чином, для розпізнавання набору рукописних цифр методом k-середніх з  $n\_neighbors = 10$  доцільно взяти 8 дескрипторів замість 12. Виключити п'ятий, шостий і сьомий Ху-моменти і ексцентриситет. Точність розпізнавання склала 78,58% проти 78,14%.

#### **ЛИТЕРАТУРА / ЛІТЕРАТУРА**

1. Шлезингер М., Главач В. Десять лекций по статистическому и структурному распознаванию. - К.: Наукова думка, 2004. — 545 с.
2. Барсегян А. А., Куприянов М. С., Степаненко В.В., Холод И.И. Методы и модели анализа данных: OLAP и Data Mining, - СПб.: БХВ-Петербург, 2004. – 336с.
3. Плас Дж. Вандер. Python для сложных задач: наука о данных и машинное обучение. — СПб.: Питер, 2018. — 576 с.
4. Нейронные сети с самоорганизацией в задачах классификации и обработки изображений / Г.А. Ососков, С.Г. Дмитриевский, А.В. Стадник // Искусств. интеллект. - 2004. - № 3. - С. 574-586. - Библиогр.: 6 назв. - рус.
5. Дорош Н. Л., Храпач Ю. А. Программное средство для распознавания цифр на изображениях// Материалы международной научной конференции «Интеллектуальные системы принятия решений и проблемы вычислительного интеллекта» (ISDMCI'2012). – Херсон: ХНТУ, 2012. – С. 353-355.
6. MNIST. Who is the best in MNIST? [Електронний ресурс] – Режим доступу. — URL:  
[https://rodrigob.github.io/are\\_we\\_there\\_yet/build/classification\\_datasets\\_results.html](https://rodrigob.github.io/are_we_there_yet/build/classification_datasets_results.html) / (дата звернення 18.10.2019)
7. Распознавание рукописных цифр с использованием сверточных нейронных сетей в Python с Keras [Електронний ресурс] – Режим доступу. — URL:

<https://www.machinelearningmastery.ru/handwritten-digit-recognition-using-convolutional-neural-networks-python-keras> // (дата звернення 18.10.2019).

8. Жерон, Орельен. Прикладное машинное обучение с помощью Scikit-Learn и TensorFlow: концепции, инструменты и техники для создания интеллектуальных систем./Пер. с англ. — СПб.: ООО "Альфа-книга, 2018. — 688 с.

9. The MNIST database of handwritten digits. [Электронный ресурс] – Режим доступа. — URL: <http://yann.lecun.com/exdb/mnist/> /(дата звернення 18.10.2019).

10. Гонсалес Р., Вудс Р., Эддинс С. Цифровая обработка изображений в среде MatLab. М: Техносфера, 2006. – 616 с.

11. Hu\_moments\_in\_Python. [ Электронный ресурс] – Режим доступа. — URL:[https://github.com/adailtonjn68/hu\\_moments\\_in\\_python/blob/master/hu\\_moments.py](https://github.com/adailtonjn68/hu_moments_in_python/blob/master/hu_moments.py)/ (дата звернення 24.11.2019).

12. Яне Б. Цифровая обработка изображений. М: Техносфера, 2007.–584 с.

#### **REFERENCES**

1. Shlezinger M., Glavach V. Desyat leksiyy po statisticheskomu i strukturnomu raspoznavaniyu. - K.: Naukova dumka, 2004. — 545 s.

2. Barsegyan A. A., Kupriyanov M. S., Stepanenko V.V., Holod I.I. Metodyi i modeli analiza dannyih: OLAP i Data Mining, - SPb.: BHV-Peterburg, 2004. – 336s.

3. Plas Dzh. Vander. Python dlya slozhnyih zadach: nauka o dannyih i mashinnoe obuchenie. — SPb.: Piter, 2018. — 576 s.

4. Neyronnyie seti s samoorganizatsiey v zadachah klassifikatsii i obrabotki izobrazheniy / G. A. Ososkov, S. G. Dmitrievskiy, A. V. Stadnik // Iskusstv. intellekt. - 2004. - # 3. - S. 574-586. - Bibliogr.: 6 nazv. - rus.

5. Dorosh N. L., Hrapach Yu. A. Programmnoe sredstvo dlya raspoznavaniya tsifr na izobrazheniyah// Materialyi mezhdunarodnoy nauchnoy konferentsii «Intellektualnyie sistemyi prinyatiya resheniy i problemyi vyichislitel'nogo intellekta» (ISDMCI'2012). – Herson: HNTU, 2012. – S. 353-355.



6. MNIST. Who is the best in MNIST? [Electronic resource] - Access mode.— URL:[https://rodrigob.github.io/are\\_we\\_there\\_yet/build/classification\\_datasets\\_results.html](https://rodrigob.github.io/are_we_there_yet/build/classification_datasets_results.html) / (date of appeal 18.10.2019)
7. Raspoznavanie rukopisnyih tsifr s ispolzovaniem svertochnyih neyronnyih setey v Python s Keras [Electronic resource] - Access mode.— URL: <https://www.machinelearningmastery.ru/handwritten-digit-recognition-using-convolutional-neural-networks-python-keras/> / (date of appeal 18.10.2019)
8. Zheron, Orelen. Prikladnoe mashinnoe obuchenie s pomoschyu Scikit-Learn i TensorFlow: kontseptsii, instrumentyi i tehniki dlya sozdaniya intellektualnyih sistem./Per. s angl. — SPb.: OOO "Alfa-kniga, 2018. — 688 s.
9. The MNIST database of handwritten digits. [Electronic resource] - Access mode.— URL: <http://yann.lecun.com/exdb/mnist/> / (date of appeal 18.10.2019).
10. Gonsales R., Vuds R., Eddins S. Tsifrovaya obrabotka izobrazheniy v srede MatLab. M: Tehnosfera, 2006. — 616 s.
11. Hu\_moments\_in\_Python. [Electronic resource] - Access mode. — URL:[https://github.com/adailtonjn68/hu\\_moments\\_in\\_python/blob/master/hu\\_moments.py](https://github.com/adailtonjn68/hu_moments_in_python/blob/master/hu_moments.py) / (date of appeal 24.11.2019).
12. Yane B. Tsifrovaya obrabotka izobrazheniy. M: Tehnosfera, 2007. — 584 s.

Received 07.02.2020.

Accepted 11.02.2020.

### ***Исследование дескрипторов для распознавания цифр набора MNIST***

*Проведены выбор и исследование дескрипторов для набора MNIST. Выполнено распознавание цифр на основе 12 дескрипторов с применением моделей из библиотеки Scikit-Learn Python. Получены результаты сравнительного анализа распознавания с разным количеством дескрипторов методом k-средних, на основании которого установлено, что целесообразно перейти к группе дескрипторов из 8 элементов.*

### ***Research of descriptors for digit recognition of MNIST dataset***

*The work is dedicated to solving the task of digit recognition.*

*Based on the results of twenty years of research, we can verify that the problem of digit recognition, even being well studied, is still of considerable interest.*

*The MNIST set has become a test set one and is used by many authors for testing of recognition algorithms. In the work "MNIST. Who is the best at MNIST" there is a spreadsheet with result of handwritten numbers recognition made by different algorithms, which are combined into groups.*

*The best result of recognition have an error of less than 1%. They are obtained using neural networks. Successful algorithms of recognition, including deep learning, are hidden from user and they are difficult in description. That is why the descriptor-based recognition algorithm is still relevant.*

*The goal of the work is the study of influence of descriptors and reduction of their quantity for recognition of MNIST set handwritten numbers*

*For recognition of the MNIST digits, a set of 12 descriptors was chosen, namely: seven  $X$ -moments,  $E_n$  (Euler's number),  $E_x$  (filling coefficient),  $E_c$  (eccentricity),  $Y_n$ ,  $X_n$  (coordinates of the center of gravity). Statistical characteristics and histograms in relation to the training and test sets MNIST were determined. Based on their research, a number of assumptions were made.*

*Digit recognition with usage of classifier based on  $k$ -means method with  $n\_neighbors = 10$  of Scikit-Learn Python system library was done. Preliminary analysis of descriptors gave the reason to assume, that the fifth, sixth and seventh Hu-moments doesn't contribute into result of digit recognition of test set using  $k$ -means method. This assumption is justified with researches, which showed, that there is a need to exclude eccentricity from the set of descriptors*

*Thus, for recognition of a set of handwritten digits by the  $k$ -means method with  $n\_neighbors = 10$ , it is advisable to take 8 descriptors instead of 12, excluding the fifth, sixth and seventh Hu-moments and eccentricity. Recognition accuracy was 78.58% compared to 78.14%.*

**Дорош Наталья Леонидовна** - к.т.н., доцент, кафедра інформаційних технологій і систем, Національна металургічна академія України.

**Фененко Татьяна Михайловна** - старший преподаватель, кафедра інформаційних технологій і систем, Національна металургічна академія України.

**Дорош Наталія Леонідівна** - к.т.н., доцент, кафедра інформаційних технологій і систем, Національна металургічна академія України.

**Фененко Тетяна Михайлівна** - старший викладач, кафедра інформаційних технологій і систем, Національна металургічна академія України.

**Dorosh Nataliia** - associate professor, department of information technologies and systems, National Metallurgical Academy of Ukraine.

**Fenenko Tatyana** - senior teacher, department of information technologies and systems, National Metallurgical Academy of Ukraine.