

# ДОСЛІДЖЕННЯ МЕТОДІВ РОЗПІЗНАВАННЯ РУКОПИСНИХ СИМВОЛІВ НА ОСНОВІ НЕЙРОННИХ МЕРЕЖ

Вінницький національний технічний університет

## *Анотація*

*Доведено актуальність проблеми розпізнавання рукописних символів, проведено огляд основних методів розпізнавання рукописних символів, вказано основні переваги і недоліки кожного з методів, здійснено порівняння та вибір нейронної мережі для розв'язання даної задачі.*

**Ключові слова:** розпізнавання символів, рукописний символ, нейронна мережа, багат шаровий перцептрон, алгоритм зворотного розповсюдження помилки.

## *Abstract*

*The relevance of the problem of handwriting recognition was proved, an overview of the main methods of recognition of handwritten characters was given, the main advantages and disadvantages of each of the methods were given, the comparison and selection of the neural network for the solution of this problem was made.*

**Keywords:** character recognition, handwriting symbol, neural network, multilayer perceptron, algorithm for reverse error propagation.

## **Вступ**

На сучасному етапі розвитку людства все частіше виникає потреба у розв'язання так званих важко формалізованих задач (для розв'язання важко формалізованих задач часто використовують методи штучного інтелекту [1]). Однією з таких задач є розпізнавання рукописних текстів, і, зокрема, рукописних символів. Розпізнавання рукописних символів може бути корисним у багатьох сферах життєдіяльності людини, наприклад, криміналістиці, музейній справі, власних потребах і т. д. Це буває потрібно, наприклад, для оцифрування різноманітних текстових матеріалів, написаних рукою людини (у тому числі, як сучасних, так і старовинних документів, рукописів, конспектів, нотаток, книг тощо); для введення у комп'ютер швидкописаних рукописних слів на сенсорному екрані; для створення людино-машинного інтерфейсу і т. п. Тому розв'язання задачі розпізнавання рукописних символів є досить важливою та актуальною.

Метою роботи є порівняння існуючих підходів щодо розпізнавання рукописних символів у поєднанні з аналізом нейронних мереж, з метою обґрунтування вибору останньої для розв'язання задачі розпізнавання рукописних символів.

## **Основна частина**

Основні підходи розпізнавання тексту умовно можна поділити на дві групи [2]:

- підходи, засновані на бінаризації вихідного зображення. Вихідне зображення конвертується у бінарну матрицю та використовується як шаблон для розпізнавання [2, 3];

- підходи щодо аналізу вихідного зображення шляхом розтину його на геометричні примітиви. При цьому фрагменти графічного зображення порівнюються з шаблонами, що зберігаються у базі застосування.

Підходи, засновані на бінаризації. Вихідне зображення символу конвертується у бінарну матрицю фіксованих розмірів, у якій одиницями позначаються точки, що відповідають введеному символу. Перевагами таких підходів є висока точність та швидкість розпізнавання зображення при використанні значних обчислювальних ресурсів та пам'яті. Недоліками даних підходів є:

- залежність аналізатора від інтерфейсу. Розмір бінарної матриці залежить від інтерфейсу програми, оскільки кожний елемент матриці відповідає окремому пікселю (групі пікселів) введеного символу [4];

- залежність шаблону від розміру та позиції зображення. Оскільки кожен елемент бінарної матриці відповідає окремому пікселю (чи групі пікселів), то введення користувачем однакового символу різних розмірів у різних місцях екрану, приведе до того, що нейронна мережа отримає різні шаблони для розпізнавання.

Як наслідок, для коректного розпізнавання символів база шаблонів повинна бути досить великою і містити усі можливі варіанти введення одного й того самого символу.

Підходи щодо аналізу вихідного зображення шляхом розтину його на геометричні примітиви. Найкращі методи розпізнавання символів використовують оптичні характеристики цих символів [3, 5]. Тобто, існує декілька характерних для символу властивостей і система перевіряє наявність таких у зображенні. Даний пошук виконується досить швидко і має високі показники точності розпізнавання.

Основою реалізації такої системи розпізнавання є виділення контурів зображення. Пропонований підхід передбачає знаходження об'єкта при деформації його контурів.

Перевагами даного підходу є:

- незалежність аналізатора від інтерфейсу. Алгоритм аналізує зображення однаково, не залежно від розміру екрану чи інтерфейсу програми [4];

- незалежність шаблону від розмірів чи позиції символу. Оскільки один і той самий символ буде містити однаковий набір примітивів, його шаблон не залежатиме від розміру символу чи його позиції на екрані [4].

Недоліком даного підходу є залежність якості розпізнавання від почерку користувача (кожен користувач має свій почерк). Поганий, нечіткий почерк ускладнить роботу системи і зробить неможливим якісне розпізнавання геометричних примітивів.

Використання багатошарової нейронної мережі дозволить значно прискорити та покращити реалізацію розпізнавання за рахунок можливості її навчання. Потрібний продукт повинен мати змогу самоналаштуватися під потреби замовника, будучи при цьому повністю універсальним програмним забезпеченням (тобто не вимагати втручання програміста). Однак, для ефективного розв'язання поставленої задачі потрібно обрати деякий вид нейромережі, проаналізувавши існуючі варіанти реалізації останньої.

На сьогоднішній день існує багато варіантів реалізації нейронних мереж, які мають різні характеристики для вирішення задач у різних галузях людської діяльності [6 – 9]. Для розв'язання задачі розпізнавання рукописних символів найчастіше використовують чотири популярні варіанти реалізації нейромережі: нейронна мережа зі зворотним поширенням помилки, нейронна мережа Хеммінга, нейронна мережа Хопфілда та нейронна мережа Хеммінга [6]. Нижче, коротко проаналізуємо дані варіанти, приймаючи до уваги найзначущі для нас критерії.

Перший критерій – швидкодія. Оскільки, система повинна працювати максимально ефективно (швидко та точно), потрібна мережа мінімального розміру, оскільки навчання мережі є процесом складним з точки зору обчислень. Навчання мережі повинно вимагати якомога менше операцій. З метою збільшення швидкодії можна було б залучити додаткову вимогу: можливість проведення навчання на цілочисельній арифметиці. Однак це зменшить точність рішення, що негативно вплине на другий, також дуже важливий критерій: точність розпізнавання. Отже, зосередимось лише на ефективності з точки зору мінімальної кількості операцій ділення, множення та додавання (оскільки операції додавання та віднімання з точки зору комп'ютера еквівалентні за швидкодією). Використання лише одношарової мережі веде до зменшення якості розпізнавання і навчання. Оскільки однією з вимог системи є постійна адаптація та швидка реакція на зміни, потрібно використовувати багатошарову мережу [3]. Отже, для першого критерію найкращим типом мережі буде звичайна багатошарова нейромережа зі зворотнім розповсюдженням помилки. Це обґрунтовується тим, що навчання можна оптимізувати обравши функцію активації з нескладною похідною. Таким чином, ми зменшуємо кількість арифметичних операцій процесора. Інші, вищезгадані типи нейромереж мають занадто складну структуру, що тягне за собою додаткові витрати процесорного часу на реалізацію процесу розпізнавання символів.

Алгоритм навчання нейронної мережі за допомогою процедури зворотного поширення передбачає наявність деякої зовнішньої ланки, що надає мережі, крім вхідних, також і цільові вихідні образи. Тому в якості методу навчання обрано саме алгоритм зворотного розповсюдження помилки, а в якості типу мережі – багатошаровий перцептрон.

Проте, слід врахувати те, що застосування багатошарового перцептрона з традиційною структурою при вирішенні реальних завдань розпізнавання і класифікації зображень викликає певні труд-

нощі. По-перше, зображення, як правило, мають велику розмірність, внаслідок чого зростає число нейронів і синаптичних зв'язків у мережі. У свою чергу, це вимагає збільшення навчальної вибірки, внаслідок чого збільшується час і обчислювальна складність процесу навчання. По-друге, ігнорується топологія вхідних даних. Компоненти вхідного шару можуть бути представлені в будь-якому порядку, без урахування мети навчання. Однак, зображення мають строгу двовимірну структуру, в якій існує залежність між просторово сусідніми пікселями. Від цих недоліків вільні так звані згорткові нейронні мережі, які є особливим класом багатшарових перцептронів, спеціально створених для розпізнавання двовимірних поверхонь з високим ступенем інваріантності до масштабування, зміщення, повороту, зміни ракурсу та іншим просторовим спотворенням [6]. Проте слід врахувати, що рішення застосувати згорткову нейронну мережу негативно відобразиться на швидкості розпізнавання, порівняно з багатшаровою нейронною мережею.

Отже, при отриманні зображення рукописного символу великої роздільної здатності, або зашумленого зображення, написаного зі специфічним почерком, у системі розпізнавання рукописних символів можна скомбінувати декілька варіантів нейронних мереж, і, окрім багатшарового перцептронну зі зворотнім розповсюдженням помилки, застосувати згорткову нейронну мережу.

### Висновки

У ході проведеного дослідження проведено аналіз основних підходів щодо розпізнавання рукописних символів, визначено їх переваги та недоліки. Обґрунтовано доцільність використання нейронної мережі для розв'язання задачі розпізнавання рукописних символів. Проаналізовано кілька основних варіантів реалізації нейронних мереж. Розглянуто випадки розпізнавання символів при їхньому зашумленому зображенні, або у випадку зображенні великої роздільної здатності. У результаті дослідження для найбільш ефективного розпізнавання рукописних символів було вирішено спробувати скомбінувати використання багатшарової нейронної мережі зі зворотнім поширенням помилки та згорткової нейронної мережі.

### СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Месюра В. І. Основи проектування систем штучного інтелекту. Навчальний посібник / В. І. Месюра, Л. М. Ваховська. – В.: ВДТУ, 2000. – 96 с.
2. Катасёв А. С. Распознавание рукописных символов на базе искусственной нейронной сети / А. С. Катасёв, Д. В. Катасёва, А. П. Кирпичников. – Вестник технологического университета. 2015. Т.18, №11, С. 173 – 176.
3. Тарик Рашид. Создаём нейронную сеть / Тарик Рашид. – СПб.: ООО “Альфа-книга”, 2017. – 320 с.
4. Гайдуков Н. П. Обзор методов распознавания рукописного текста / Н. П. Гайдуков, Е. О. Савкова. – Донецкий национальный технический университет, 2013. – 6 с.
5. Чебанюк О. В. Методика розпізнавання рукописного тексту на основі аналізу векторів руху за допомогою сенсорних пристроїв / О. В. Чебанюк, Д. А. Долотов – Вісник східноукраїнського національного університету імені Володимира Даля 2015, С. 230 – 236.
6. Солдатова О. П. Применение сверточной нейронной сети для распознавания рукописных цифр / О. П. Солдатова, А. А. Гаршин / Самарский государственный аэрокосмический университет имени академика С. П. Королева. – 2011. – 6 с.
7. Kolesnytskyj, O. K., Bokotsey, I. V., & Yaremchuk, S. S. (2010). Optoelectronic implementation of pulsed neurons and neural networks using bispin-devices. *Optical Memory and Neural Networks (Information Optics)*, 19(2), 154-165. doi:10.3103/S1060992X10020062
8. Kozemiako, V. P., Kolesnytskyj, O. K., Lischenko, T. S., Wojcik, W., & Sulemenov, A. (2013). Optoelectronic spiking neural network. Paper presented at the Proceedings of SPIE - the International Society for Optical Engineering, , 8698 doi:10.1117/12.2019340
9. Гудфеллоу Я. Глубокое обучение / Я. Гудфеллоу, И. Бенджио, А. Курвилль пер. с англ. А. А. Слинкина – 2-е изд., испр. – М.: ДМК Пресс, 2018. – 652 с.

**Зінов'єв Євгеній Вікторович** – студент групи ІКН-17МС, факультет інформаційних технологій та комп'ютерної інженерії, Вінницький національний технічний університет, Вінниця, e-mail: zinovieff55@gmail.com

**Арсенюк Ігор Ростиславович** – к. т. н., доцент, доцент кафедри комп'ютерних наук, Вінницький національний технічний університет, м. Вінниця

**Yevgeny V. Zinoviev** – student of group ІКН-17МС, Faculty of Information Technologies and Computer Engineering, Vinnytsia National Technical University, Vinnytsia, e-mail: zinovieff55@gmail.com

**Igor R. Arsenyuk** – Cand. Sc., Assistant Professor of the Chair of Computer Science, Vinnytsia National Technical University, Vinnytsia