

122 КОМП'ЮТЕРНІ НАУКИ ТА ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ

УДК 004.896

doi: 10.32782/2225-6733.43.2021.1

© Проніна О.І.¹, Остапенко Р.В.², Альошин С.В.³

РОЗПІЗНАВАННЯ ЖЕСТІВ ЗА ДОПОМОГОЮ ЗГОРТКОВОЇ НЕЙРОННОЇ МЕРЕЖІ

В статті наведено матеріал стосовно розпізнавання жестів за допомогою згорткової нейронної мережі. Нейронна мережа була натренована, щоб у режимі реального часу виявляти жести мови жестів та ідентифікувати їх. В рамках роботи було проведено аналіз наукових публікацій, існуючих методів розпізнавання та вилучення об'єктів на зображенні. Для розпізнавання жестів була побудована математична модель згорнутої нейронної мережі. Після чого вона була натренована на обраному датасеті. Для тренування було створено набір фотографій, що складається з 4000 фотографій для 50 жестів, це для кожного жесту було створено 80 фотографій. Програмне забезпечення було написано на мові Python, використовувались ряд бібліотек, а сам YOLOv5, NumPy, PyTorch, OpenCV. В експериментальних дослідженнях було проведено порівняння результатів розпізнавання на навчальній вибірці жестів та на тестовій вибірці. Для визначення адекватності роботи розробленої згорнутої нейронної мережі було проведено тестування на визначення помилок при розпізнаванні жестів.

Ключові слова: мова жестів, згорткова нейронна мережа, розпізнавання об'єктів, YOLOv5, програмне забезпечення.

O.I. Pronina, R.V. Ostapenko, S.V. Aloschin. Gesture recognition using a rolled neural network. In the modern world, more and more attention is paid to the needs of a person with disabilities. Developing software that allows such people to communicate is an urgent scientific and practical task. Its implementation with the help of artificial intelligence, namely neural networks, allows to increase the level of gesture recognition. In addition, the use of neural networks can improve not only the recognition accuracy rate, but also the recognition speed and efficiency. Which in turn enhances the process of communication of people with disabilities. Today, there are many algorithms for object recognition, but the complexity of the issue under consideration lies in the fact that gestures in sign language are very similar to each other. There are a number of problems that are associated with identifying the hand, the one who speaks using sign language, and recognizing the gesture itself in real time. Software development with the help of actual libraries and improvement of the recognition speed is the main task of the research. The article provides a review of modern literature, where the authors consider the problems of gesture recognition. It was decided to use a convolutional neural network as the main tool for gesture recognition. To develop a mathematical model of a convolutional neural network, the main settings and their indicators were chosen. The convolutional neural network was trained to determine the learning error and minimize it. After training, the developed mathematical model was put into the basis of the software. The article is devoted

¹ канд. техн. наук, доцент, ДВНЗ «Приазовський державний технічний університет», м. Маріуполь, ORCID: 0000-0001-7085-8027, pronina.lelka@gmail.com

² магістр, ДВНЗ «Приазовський державний технічний університет», м. Маріуполь

³ ст. викладач, ДВНЗ «Приазовський державний технічний університет», м. Маріуполь, alioshin.serg@gmail.com

to the description of software development and its testing on the created data set, which includes 4000 photos of various gestures. Namely, 80 photos for each of the 50 selected gestures that were included in the database. To test the developed software and the mathematical model that underlies it, experimental studies were carried out in real time on the recognition of gestures of the sign language.

Key words: *sign language, convolutional neural network, object recognition, YOLOv5, software.*

Постановка проблеми. Жести виконують безліч функцій в спілкуванні, навчанні і розумінні як для тих, хто їх бачить, так і для тих, хто їх створює. Жести особливо ефективні, коли вони схожі на думки, які вони представляють, і мають перевагу перед словами. Жести можуть відображати багато значення більш безпосередньо, ніж мова, конгруентно представляючи багато концепції. Розробка і використання жестів, відповідних значенням, може поліпшити розуміння і навчання.

Глухонімі люди – це люди на найглибшому психологічному рівні. Багато з цих людей навіть не знайомі з мовами жестів, і було помічено, що це дає велике полегшення на психологічному рівні, коли вони дізнаються про жести, щоб з'єднатися з іншими, висловлюючи свою любов або емоції.

Близько 5% населення світу страждає втратою слуху. Глухі та німі люди використовують мову жестів як основний засіб вираження своїх думок і ідей оточуючим за допомогою різних жестів рук і тіла [1]. Основною проблемою спілкування на мові жестів є те, що зазвичай людина не знає мову жестів, тому не може допомогти людині з обмеженнями. Крім того, люди під час розмови виражають свої думки дуже швидко. Люди з обмеженнями також дуже швидко будують речення за допомогою мови жестів. І основною проблемою є швидкість розпізнавання цих жестів та їх точність.

Метою даної роботи є побудова математичної моделі згорткової нейронної мережі та її реалізації у вигляді програмного забезпечення для розпізнавання мови жестів. Для розпізнавання жестів буде використовуватися бібліотека YOLOv5 з налаштуванням основних параметрів для підвищення швидкості розпізнавання та виявлення об'єктів в режимі реального часу.

Аналіз останніх досліджень і публікацій. Розпізнавання об'єктів є достатньо популярною темою наукових досліджень, оскільки багато де може бути використана. Розпізнавання жестів руками – одне з найважливіших завдань у галузі дослідження мови жестів. Оскільки це важка сфера, яка страждає від складних проблем, пов'язаних з контекстом дослідження, вона потребує використання надійних дескрипторів. В області розпізнавання жестів на сьогодні є безліч робіт.

Наприклад, в роботі [2] наведено систему розпізнавання жестів руками з використанням датчика Kinect. Автори, окрім розробленої системи, також запропонували нову метрику відстані для вимірювання несхожості рук, звану Відстань рухаючого пальця-Землі (FEMD). Тим самим вони підіймають питання того, що для тренування нейронної мережі одним з чинників є те, що не всі руки схожі між собою.

Новий підхід до сегментації рук на основі виявлення країв описаний у роботі [3]. Кроки підходу до сегментації можна сформулювати таким чином: усунення шуму шляхом застосування двостороннього фільтра на зображенні, виявлення країв за допомогою методу «Canny», закриття країв краю, заповнення області руки, видалення всіх небажаних областей краю. Даний підхід дозволяє локалізувати руки для подальшої ідентифікації жесту, що вони зображають.

У роботі [4] автори запропонували новий метод розпізнавання алфавіту американської мови жестів (ASL) за допомогою недорогої камери та датчиків. Цей новий метод використовує функцію глибокого контрасту на основі алгоритму класифікації на піксель для сегментації конфігурації руки. Для забезпечення етапу сегментації рук автори використовували спеціальні рукавички. Потім розробляється та впроваджується ієрархічний метод пошуку режимів для локалізації позицій суглоба кисті за кінематичних обмежень. Нарешті, побудований класифікатор випадкових лісів (RF) для розпізнавання знаків ASL за допомогою спільних кутів.

У роботі [5] автори запропонували високоточний метод розпізнавання статичних жестів за даними глибини. Після чого вони використовують багатосаровий випадковий ліс (MLRF), за

допомогою якого вдається класифікувати вектори ознак, що призводить до розпізнавання ручних знаків.

Система, розроблена у роботі [6], використовує комбінований дескриптор RGB та глибини для класифікації жестів руками. Використовуються два взаємопов'язані модулі: перший виявляє руку в області взаємодії та виконує класифікацію користувача, а другий – розпізнавання жестів.

В роботі [7] автори запропонували нову схему розпізнавання мови жестів для ідентифікації алфавітів у мові жестів. Для цього вони вилучили 10 різних просторових об'єктів, що були далі використані для завдання розпізнавання. Для розпізнавання жестів була обрана нейронна мережа зворотного поширення, яка використовується для класифікації ознак. Також була використана функція 2D (орієнтація руки та номер Ейлера).

У роботі [8] пропонується нова система розпізнавання жестів на основі суперпікселів. Ця система базується на новій суперпіксельній метриці відстані землерубки разом із камерою глибини Kinect для вимірювання несхожості між жестами рук. Це вимірювання є не тільки стійким до спотворень та артикуляцій, але також є інваріантним до масштабування, трансляції та обертання за допомогою належної попередньої обробки.

У роботі [9] автори пропонують метод розпізнавання 14 жестів на мові жестів з використанням контурних функцій, де аналіз кривизни часу використовується для опису силуету форми руки, а потім класифікатор SVM вивчається за допомогою виділених ознак.

У роботі [10] автори розробляють систему на основі CNN для аналізу жестів руками на платформах iPhone та рук, де двійкові маски для рук класифікуються на кілька жестів.

Виклад основного матеріалу. У класичних підходах до розпізнавання та класифікації об'єктів спочатку зображення представляється числовими дескрипторами, що характеризують структури пікселів найбільш унікальним способом для певного набору зображень, що належать до того ж класу. У випадку зображень, що представляють різні класи, значення конкретних дескрипторів повинні бути максимально різними. На практиці застосовуються різні методи попередньої обробки, що ведуть до різних визначень дескрипторів. Усі вони зменшують розмір вхідного зображення до відносно невеликого розміру описового вектору зображення [2].

Усі ці методи попередньої обробки призводять до представлення зображення обмеженою кількістю діагностичних ознак. Об'єкти повинні представляти вихідні вектори (зображення жестів) таким чином, щоб забезпечити найвищу однорідність у межах одного класу та найвищі відмінності для зображень, що представляють різні класи. Обробка зображення складається з трьох кроків:

- імпортування зображення за допомогою оптичного сканера або цифрової фотографії;
- аналіз та маніпулювання зображенням, що включає стиснення даних і покращення зображення, а також визначення шаблонів, які не є для людського ока, як супутникові фотографії;
- вихід – це останній етап, на якому результатом може бути змінене зображення або звіт, заснований на аналізі зображення.

Нейронні мережі зараз широко використовуються для обробки та розпізнавання зображень, у системах розпізнавання мовлення, відеоаналітики, інтелектуальної безпеки, тощо. В роботі було вирішено використовувати згорткову нейронну мережу, тому що її архітектура дозволяє максимально якісно працювати з розпізнаванням об'єктів в реальному часі.

Загальний процес використання згорткової нейронної мережі для розпізнавання мови жестів для програмного забезпечення, що було розроблено, включає в себе два етапи.

На першому етапі відбувається створення штучної нейронної мережі. Цей етап містить наступні кроки:

- збір набору даних (4000 фотографій, для кожного з 50 жестів було створено по 80 фотографій у форматі .jpg);
- попередня обробка набору даних (розмітка жестів за допомогою програми LabelImg – інструменту для створення графічних анотацій);
- завантаження набору даних, а саме імпорт вхідних зображень, що будуть використані для навчання штучної нейронної мережі;
- тренування моделі, що включає в себе процес вибору математичної моделі, функцій активації, витрат, вибір кількості епох, оптимізації тощо;

– тестування, включає в себе оцінку точності результатів навчання моделі (необхідною і достатньою було встановлено точність в 0.7 та вище);

– збереження штучної нейронної мережі, а саме збереження отриманої після навчання моделі з метою подальшого використання у програмному забезпеченні.

Другий етап описує роботу з додатком та містить наступні кроки:

– імпорт штучної нейронної мережі для подальшої роботи з нею, а саме використання як окремого модуля при роботі додатку;

– використання штучної нейронної мережі під час роботи програмного забезпечення, процес відбувається наступним чином: при увімкненні додатку включається камера пристрою (ПК/ноутбуку) та відбувається процес розпізнавання та класифікації жестів у реальному часі;

– відображення результатів розпізнавання, а саме жест людини виділяється рамкою з позначенням точності розпізнавання цього жесту та його розшифрування. На точність розпізнавання також впливає освітлення: чим світліше, тим вище точність розпізнавання;

– оцінка точності результатів навчання: якщо точність вище 0,7, то програма записує цей жест, інакше не записує його;

– вивід результатів, вивід у консоль всіх жестів, що пройшли перевірку на попередньому кроці. Один із 50 наявних у програмі жестів виконує роль розділювального символу, що робить зручним спілкування за допомогою програми.

Розроблену математичну модель, що включає в себе модель згорткової нейронної мережі необхідно натренувати, для тренування моделі використовуємо наступну команду:

```
python train.py --img 320 --batch 16 --epochs 200 --data dataset.yml --weights yolov5s.pt --workers 2,
```

де були передані основні параметри тренування. Для того щоб обрати параметри тренування проводилось додаткове дослідження для виявлення оптимальних значень.

Розмір зображення автоматично задається як 320px * 320px. В моделі використано пакетну нормалізацію (batch normalization), щоб нормалізувати результати прискорення тренування. Кількість епох при навчанні було встановлено 200 з метою покращення навчання моделі. Значення для вагів (weights) згорткової нейронної мережі беруться з файлу yolov5s.pt. Параметр workers визначає кількість робочих потоків для обробки запитів. Інші параметри та значення налаштування наведено у таблиці 1.

Таблиця 1

Параметри та значення налаштування згорткової нейронної мережі

| Гіперпараметр | Значення |
|-------------------|-----------------------|
| Learning Rate | 0.001 |
| Epochs | 200 |
| Number of Classes | 50 |
| Algorithm | YOLOv5 |
| Optimizer | Adam |
| Activation | Linear, Leaky |
| Filter Size | [64,128,256,512,1024] |
| Mask | 0–8 |
| Decay | 0.0005 |

Результат тренування моделі по завершенні 200 епох наведено на рис. 1. Навчання зайняло 43.653 години для обраних 50 жестів мови жестів.

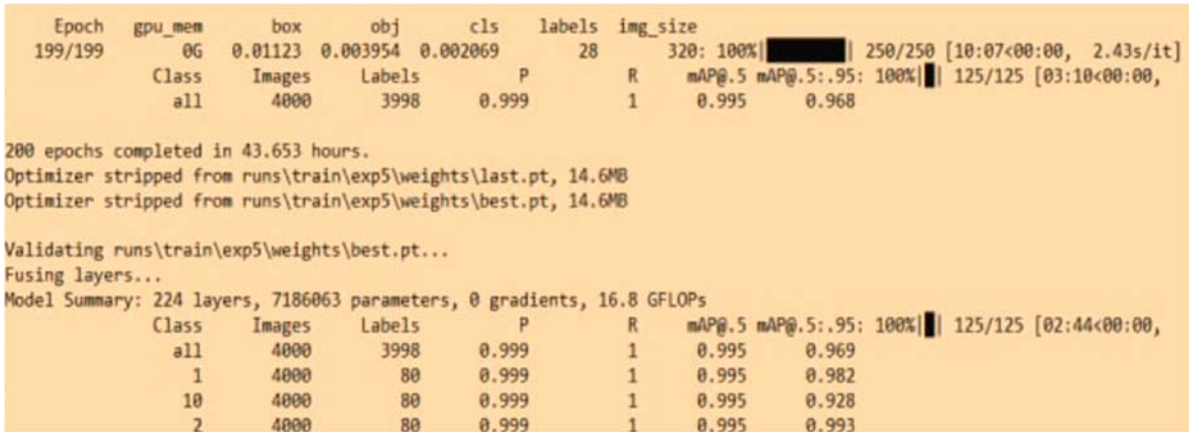


Рис. 1 – Результат тренування моделі

Результат тренування згорткової нейронної мережі наведено на рисунку 2. У лівій частині рис. 2 показано рівні помилок у тренувальних даних (верхні) та валідаційних (нижні). Висока валідаційна точність при низьких валідаційних втратах означає, що модель навчена коректно.

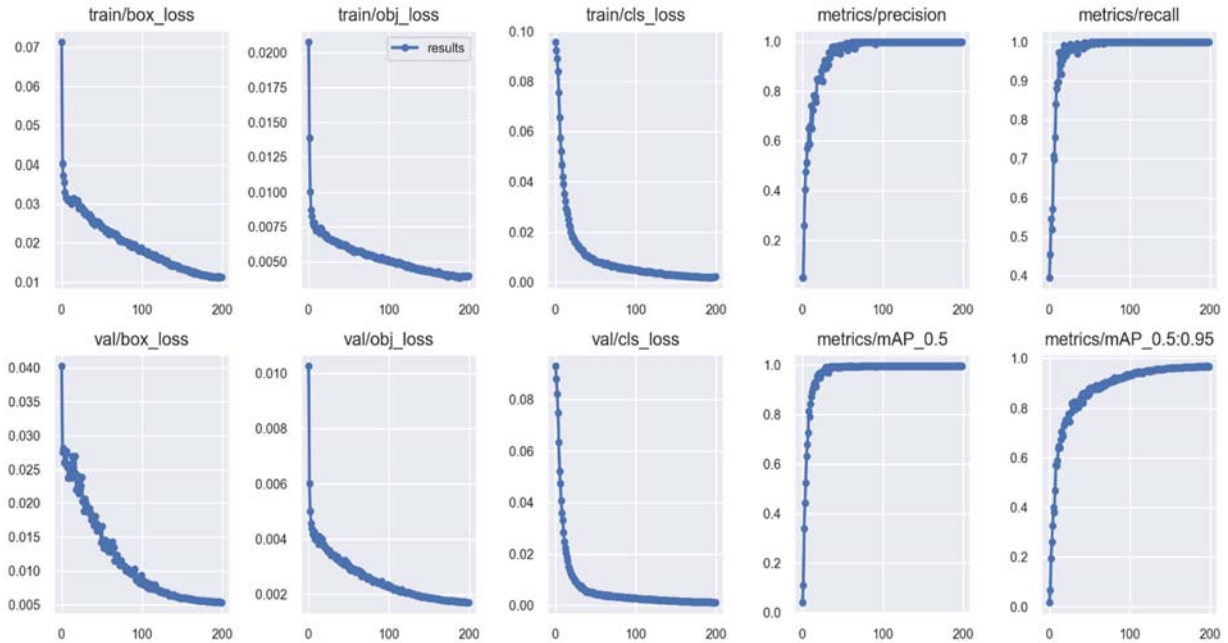


Рис. 2 – Результат тренування моделі

Таким чином, за допомогою мови програмування Python, бібліотеки YOLOv5 було створено модель для розпізнавання жестів рук у режимі реально часу, а також за допомогою бібліотек NumPy, PyTorch, OpenCV було розроблено додаток, що дозволяє розпізнавати жести та полегшати спілкування за допомогою використання розробленого програмного забезпечення.

Приклад розпізнавання жесту за допомогою розробленого програмного забезпечення, в основі якого покладено розроблену математичну модель згорткової нейронної мережі, наведено на рис. 3. Жести, що мають точність більше 70%, записуються в консоль, таким чином інтерпретація жестів для розмови стає можливою.

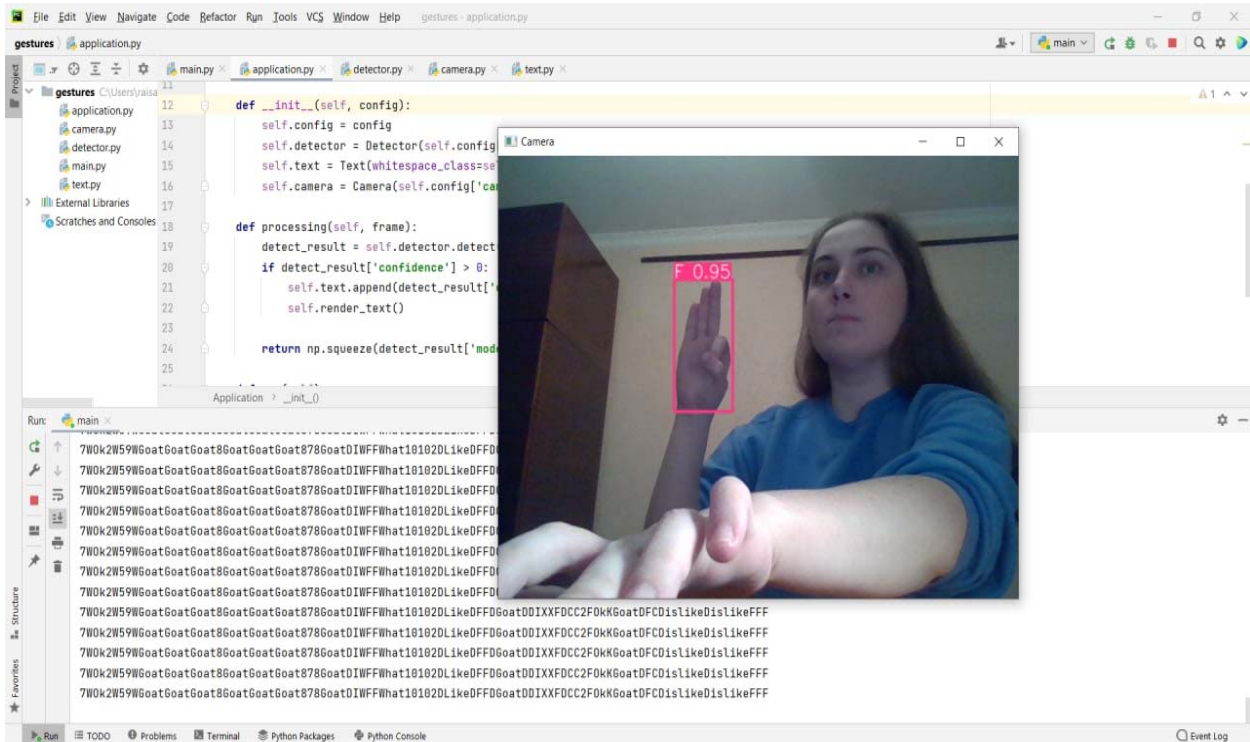


Рис. 3 – Результат розпізнавання жесту

З метою перевірки адекватності навчання розробленої математичної моделі наведено фрагмент розпізнавання жестів з навчальної та тестової вибірки за допомогою створеної згорткової нейронної мережі (фрагмент наведено у таблиці 2). Таким чином, з таблиці 2 можна зробити висновок, що модель розпізнає як навчальні, так і тестові дані, це означає, що модель не перенавчилася, що є одним з основних проблем навчання моделей нейронних мереж.

Таблиця 2

Фрагмент порівняння точності навчальної та тестової вибірки

| Назва жесту | Значення точності | |
|-------------|-------------------|---------|
| | Навчальна | Тестова |
| 1 | 0.95 | 0.82 |
| 2 | 0.95 | 0.89 |
| 3 | 1 | 0.92 |
| 4 | 1 | 0.89 |
| 5 | 1 | 0.91 |
| 6 | 0.90 | 0.85 |
| 7 | 1 | 0.94 |

Для перевірки розробленого програмного забезпечення проводилося дослідження. Для цього виконане тестування точності розпізнавання жестів. 12 жестів було розпізнано з точністю 92%, 7 жестів розпізнані з точністю 94%. Результати розпізнавання ряду жестів зображені на рис. 4.

Аналізуючи отримані результати можна зробити висновок, що більшість значень знаходяться в одних діапазонах (від 0,83 до 0,95). Розроблене програмне забезпечення розпізнавання жестів руки виявляє як об'єкти у реальному часі, так і жести з відеокадрів з точністю 0,8986%.



Рис. 4 – Діаграма результатів розпізнавання жестів

У розробленій математичній моделі все ще є можливості для покращення, оскільки в даний час запропонована модель виявляє статичні жести. Швидкість розпізнавання самого жесту є достатньою для розмови, але необхідно доробити модель, щоб вона розпізнавала жести, які змінюються з високою швидкістю. Крім того математична модель може бути покращена для виявлення кількох жестів за раз. Запропонований метод може бути використаний для покращення системи життєзабезпечення, які використовуються для взаємодії людей з обмеженнями за допомогою комп'ютера.

Висновки

У ході виконання роботи було створено математичну модель згорткової нейронної мережі для розпізнавання мови жестів і реалізовано її у вигляді програмного забезпечення. Було створено набір даних зображень жестів руки, що було необхідно для навчання математичної моделі, що розробляється. Також було проведено навчання розробленої математичної моделі для 50 класів за допомогою бібліотеки YOLOv5. Для цього було експериментально виявлено значення основних параметрів нейронної мережі. Створено пакетну нормалізацію (batch normalization). Для реалізації програмного забезпечення було використано технології Python, NumPy, PyTorch, OpenCV. Було проведено тренування програмного забезпечення та тестування моделі згорткової нейронної мережі для вирішення задачі розпізнавання зображень жестів руки (перевірка на навчальних даних показала точність 97%, а на тестових даних – 90%), що є достатнім результатом.

Також було виконано оцінку якості розпізнавання жестів руки програмним забезпеченням при різному освітленні. Перевірка на навчальних даних показала точність 97%, а на тестових даних – 89%. Розроблена система може розпізнавати жести рук (виявляти їх як об'єкти) у режимі реального часу, з точністю $0,8986 \div 0,9\%$.

Запропоноване програмне забезпечення для розпізнавання жестів рук можна застосовувати для вдосконалення допоміжних живих системи, які використовуються для взаємодії людини та комп'ютера. Крім цього розроблене програмне забезпечення дозволить комунікувати з людьми з обмеженими можливостями.

Перелік використаних джерел:

1. Boppana L. Assistive Sign Language Converter for Deaf and Dumb / L. Boppana, R. Ahamed // 2019 International Conference on Internet of Things and IEEE Green Computing and Communications and IEEE Cyber, Physical and Social Computing and IEEE Smart Data. – 2019. – Pp. 2-4. – Mode of access: <https://doi.org/10.1109/iThings/GreenCom/CPSCCom/SmartData.2019.00071>.
2. RGB-D datasets using microsoft Kinect or similar sensors: a survey / Z. Cai, J. Han, Li Liu,

- L. Shao // *Multimedia Tools and Applications*. – 2016. – Vol. 75. – Pp. 1-43. – Mode of access: <https://doi.org/10.1007/s11042-016-3374-6>.
3. Chaudhury K.N. Acceleration of the shifttable $o(1)$ algorithm for bilateral filtering and non-local means / K.N. Chaudhury // *IEEE Transactions on Image Processing*. – 2013. – Vol. 22. – Pp. 1291-1300. – Mode of access: <https://doi.org/10.1109/TIP.2012.2222903>.
 4. Dong C. American sign language alphabet recognition using microsoft Kinect / C. Dong, M. Leu, Z. Yin // *IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. – 2015. – Pp. 44-52. – Mode of access: <https://doi.org/10.1109/CVPRW.2015.7301347>.
 5. Kuznetsova A. Real-time sign language recognition using a consumer depth camera / A. Kuznetsova, L. Leal-Taixe, B. Rosenhahn // *2013 IEEE International Conference on Computer Vision Workshops*. – 2013. – Pp. 83-90. – Mode of access: <https://doi.org/10.1109/ICCVW.2013.18>.
 6. Ohn-Bar E. Hand gesture recognition in real time for automotive interfaces: A multimodal vision-based approach and evaluations / E. Ohn-Bar, M. Manubhai Trivedi // *IEEE Transactions on Intelligent Transportation Systems*. – 2014. – Vol. 15, iss. 6. – Pp. 2368-2377. – Mode of access: <https://doi.org/10.1109/TITS.2014.2337331>.
 7. Pandey P. An efficient algorithm for sign language recognition / P. Pandey, V. Jain // *International Journal of Computer Science and Information Technologies (IJCSIT)*. – 2015. – Vol. 6. – Pp. 5565-5571.
 8. Wang C. Superpixel-based hand gesture recognition with Kinect depth camera / C. Wang, Z. Liu, S. Chan // *IEEE Transactions on Multimedia*. – 2015. – Vol. 17, iss. 1. – Pp. 29-39. – Mode of access: <https://doi.org/10.1109/TMM.2014.2374357>.
 9. Hand gesture recognition with multiscale weighted histogram of contour direction normalization for wearable applications / Y. Ren, X. Xie, G. Li, Z. Wang // *IEEE Transactions on Circuits and Systems for Video Technology*. – 2018. – Vol. 28, iss. 2. – Pp. 364-377. – Mode of access: <https://doi.org/10.1109/TCSVT.2016.2608837>.
 10. Towards robust ego-centric hand gesture analysis for robot control / H. Song, W. Feng, N. Guan, X. Huang, Z. Luo // *IEEE International Conference on Signal and Image Processing (ICSIP)*. – 2016. – Pp. 661-666. – Mode of access: <https://doi.org/10.1109/SIPROCESS.2016.7888345>.

References:

1. Boppana L., Ahamed R. Assistive Sign Language Converter for Deaf and Dumb / L. Boppana, R. Ahamed. *2019 International Conference on Internet of Things and IEEE Green Computing and Communications and IEEE Cyber, Physical and Social Computing and IEEE Smart Data*, 2014, pp. 2-4. doi: 10.1109/iThings/GreenCom/CPSSCom/SmartData.2019.00071.
2. Cai Z., Han J., Liu Li, Shao L. RGB-D datasets using microsoft Kinect or similar sensors: a survey. *Multimedia Tools and Applications*, 2016, vol. 75, pp. 1-43. doi: 10.1007/s11042-016-3374-6.
3. Chaudhury K.N. Acceleration of the shifttable $o(1)$ algorithm for bilateral filtering and non-local means. *IEEE Transactions on Image Processing*, 2013, vol. 22, pp. 1291-1300. doi: 10.1109/TIP.2012.2222903.
4. Dong C., Leu M., Yin Z. American sign language alphabet recognition using microsoft Kinect. *2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2015, pp. 44-52. doi: 10.1109/CVPRW.2015.7301347.
5. Kuznetsova A., Leal-Taixe L., Rosenhahn B. Real-time sign language recognition using a consumer depth camera. *2013 IEEE International Conference on Computer Vision Workshops*, 2013, pp. 83-90. doi: 10.1109/ICCVW.2013.18.
6. Ohn-Bar E., Manubhai Trivedi M. Hand gesture recognition in real time for automotive interfaces: A multimodal vision-based approach and evaluations. *IEEE Transactions on Intelligent Transportation Systems*, 2014, vol. 15, iss. 6, pp. 2368-2377. doi: 10.1109/TITS.2014.2337331.
7. Pandey P., Jain V. An efficient algorithm for sign language recognition. *International Journal of Computer Science and Information Technologies (IJCSIT)*, 2015, vol. 6, pp. 5565-5571.
8. Wang C., Liu Z., Chan S. Superpixel-based hand gesture recognition with Kinect depth camera. *IEEE Transactions on Multimedia*, 2015, vol. 17, iss. 1, pp. 29-39. doi: 10.1109/TMM.2014.2374357.

9. Ren Y., Xie X., Li G., Wang Z. Hand gesture recognition with multiscale weighted histogram of contour direction normalization for wearable applications. *IEEE Transactions on Circuits and Systems for Video Technology*, 2018, vol. 28, iss. 2, pp. 364-377. doi: 10.1109/TCSVT.2016.2608837.
10. Song H., Feng W., Guan N., Huang X., Luo Z. Towards robust ego-centric hand gesture analysis for robot control. *2016 IEEE International Conference on Signal and Image Processing (ICSIP)*, 2016, pp. 661-666. doi: 10.1109/SIPROCESS.2016.7888345.

Рецензент: О.Є. П'ятикоп
канд. техн. наук, доц., ДВНЗ «ПДТУ»

Стаття надійшла 23.08.2021

УДК 004.032.26:629.11

doi: 10.32782/2225-6733.43.2021.2

© Федосова І.В.¹, Подольніков І.В.², Гальчук Д.С.³

РОЗПІЗНАВАННЯ ДЕРЖАВНИХ НОМЕРІВ ТРАНСПОРТНИХ ЗАСОБІВ ЗА ДОПОМОГОЮ ЗГОРТКОВОЇ НЕЙРОННОЇ МЕРЕЖІ

Стаття розкриває тему розробки системи розпізнавання державних номерів різноманітних транспортних засобів, використовуючи сучасні підходи і алгоритми згорткових нейронних мереж. Система зосереджена на необмежених сценаріях захоплення, де державні знаки можуть бути значно спотворені через похилі види та інші недоліки цифрового зображення. Під час аналізу зовнішніх джерел було виявлено, що аналогічні системи мають безліч недоліків, які виражаються у: нестійкості розпізнавання низькоякісних зображень, погану ідентифікацію об'єктів (розпізнавання об'єктів, які не відносяться до транспортного засобу), а також жорсткому обмеженні обчислювальних потужностей та виробників відеокарт. Як наслідок було виявлено проблему, поставлено мету роботи, проаналізовано наукові дослідження та публікації, а також побудовано математичну модель, проведено навчання моделі нейронної мережі та тестування якості розпізнавання. Після розробки системи були виявлені й усунуті її недоліки та підвищена продуктивність. У підсумку отримано систему, що може використовуватись на відеоадаптерах Nvidia, але при необхідності (відеочіп Radeon або чіп, інтегрований у процесор) нейронна мережа може використовувати процесор. Також однією з основних переваг стала трьошарова робота нейронної мережі, завдяки котрій спочатку відбувається пошук транспортного засобу, після чого на вихідному зображенні проводиться пошук державного номеру і наприкінці починається розпізнавання номерного знаку. До більш непомітних, але важливих переваг, можна віднести: швидкість роботи, низькі системні вимоги, а також розпізнавання номерних знаків будь-якого формату та країни. Результати тестування показують, що запропонований метод без будь-якої адаптації параметрів та налаштувань для конкретного сценарію працює аналогічно найсучаснішим комерційним системам у традиційних сценаріях та навіть перевершує деякі із них.

Ключові слова: розпізнавання номерних знаків, нейронні мережі, глибоке навчання.

¹ д-р пед. наук, професор, ДВНЗ «Приазовський державний технічний університет», м. Маріуполь, ORCID: 0000-0003-3923-8270, irivasilevna1964@gmail.com

² магістр, ДВНЗ «Приазовський державний технічний університет», м. Маріуполь

³ студент, ДВНЗ «Приазовський державний технічний університет», м. Маріуполь