

УДК 004.032.26:004.931

РОЗРОБКА НЕЙРОННОЇ МЕРЕЖІ ДЛЯ РОЗПІЗНАВАННЯ ЛЮДЕЙ НА АПАРАТНІЙ ПЛАТФОРМІ RASPBERRY PI

Сербський Г.О.

email: hlib.serbskyi@nure.ua

Науковий керівник – к.т.н., проф. Зубков О.В.

Харківський національний університет радіоелектроніки, каф. МТС
м. Харків, Україна

The necessity of introducing object recognition algorithms in modern video surveillance systems is shown. Variants of standard neural network architectures and simple solutions for embedded systems are analyzed. A neural network has been developed that allows people to be recognized on the Raspberry PI hardware platform. The network allows us to process the video stream in near real time. The accuracy of object classification using the developed neural network is 5% lower than that of the well-known YOLO algorithm.

Одним із головних елементів сучасних систем безпеки є відеокамери, що забезпечують: локальне або дистанційне спостереження, детектування рухомих об'єктів. Але ефективність їх автоматичної роботи без людини низька, бо в зоні огляду можуть з'являтися такі рухомі об'єкти, як тварини, птахи і т.д., рух яких буде помилково ідентифікуватись як рух зловмисників. Розвиток штучного інтелекту, а саме нейронних мереж, дозволив розпізнавати та класифікувати рухомі об'єкти [1,2]. Існують різні архітектури нейронних мереж, такі як AlexNet, VGGNet, ResNet, YOLO та інші [3,4], які вже пройшли безліч тестів та довели свою ефективність у різних завданнях комп'ютерного зору. Використання таких архітектур дозволяє значно скоротити час та зусилля, необхідні для розробки та впровадження систем безпеки, але потребує значних апаратних ресурсів центрального процесора. Їх можна оптимізувати для роботи на вбудовуваних платформах з обмеженими апаратними ресурсами, таких як Raspberry Pi, Nvidia Jetson та інші. Існують інструменти та методи, такі як TensorRT, quantization та pruning [5], які дозволяють знизити вимоги до обчислювальних ресурсів та підвищити продуктивність моделей на малопотужних пристроях. Також можна синтезувати архітектури нейронних мереж, оптимізовані для розпізнавання потрібних класів об'єктів. Тому метою досліджень стало створення простої архітектури нейронної мережі, яка не поступається популярним архітектурам нейронних мереж, але має суттєво меншу кількість параметрів і дозволяє працювати на апаратній платформі Raspberry Pi зі швидкістю, наближеною до швидкості відеопотоку з камер спостереження і адаптованої до умов експлуатації системи безпеки.

В ході наукових досліджень були проаналізовані можливості 3 популярних архітектур для вбудування: TFLite, MobileNet-SSDLite, YOLOv8n [6,7]. Перша ефективно працює на Raspberry PI або ESP32, але містить всього 2 згорткових шари і дозволяє обробляти зображення з роздільною здатністю 92 пікселя, що забезпечує високу швидкість, але значно зменшує точність розпізнавання. Дві останніх архітектури мають від 60 до 162 шарів та працюють із розмірами зображень 320-640 пікселів. Вони забезпечують значно вищу точність розпізнавання, але потребують використання додаткових апаратних прискорювачів GPU чи потужних процесорів для обробки відеопотоків близько до реального часу.

В проведеному дослідженні було вирішено створити варіант нейронної мережі, що дозволить обробляти зображення із розміром 320 пікселів, але буде встигати обробляти відопотік близько до реального часу.

Спроектowana нейронна мережа складається з 6 згорткових шарів для підвищення якості виділення характерних ознак на зображеннях у порівнянні з TFLite. У першому та другому згорткових шарах виконувалася згортка з 32 ядрами розміром 3×3 та функцією активації 'relu'. Ця функція активації не поступається при навчанні функції 'сигмоїд', але вимагає значно менше обчислень при обробці зображень. Для зменшення проблем чутливості до початкової ініціалізації ваг та покращення збіжності навчання після кожного згорткового шару додатково використовувалася техніка BatchNormalization. Далі виконується операція MaxPool, що зменшує розмір зображення по вертикалі та горизонталі вдвічі. У 3-му та 4-му шарах структура аналогічна, але використовується згортка з 64 ядрами, а в 5-му та 6-му шарах – зі 128 ядрами. В результаті після згорткових шарів виходить 128 карт ознак розміром 40×40, що дозволяє виявляти об'єкти умовно середніх і великих розмірів. Для кінцевої класифікації об'єктів наприкінці архітектури мережі додані 2 повнозв'язні шари з 512 та 6 нейронами за кількістю класів: людина, птах, собака, кішка, автомобіль, вантажівка. Модель синтезована з використанням API Keras. При компіляції моделі використаний SGD.

Навчання нейронної мережі виконувалось з використанням створеного датасету із 400 зображень впродовж 300 епох.

Тестування ефективності роботи нейронної мережі було проведено на Raspberry PI 4B з об'ємом оперативної пам'яті 2 Гбайт в умовах різної інтенсивності освітлення, що вимірювалась за допомогою датчика BH1750FVI. За результатами розпізнавання людини в 800 кадрах з камери при однаковій яскравості освітлення обчислювалась імовірність вірного розпізнавання людини. Результати досліджень представлені у даних таблиці 1.

Таблиця 1 – Результати вимірювань вірогідності правильного розпізнавання людей

Яскравість освітлення, lux	Імовірність вірного розпізнавання	
	YOLOv8n	Наша нейронна мережа
2100	0.93	0.88
600	0.86	0.82
100	0.67	0.64
18	0.28	0.29

Аналіз даних таблиці 1 показує, що ефективність вірного розпізнавання створеної нейронної мережі нижче ніж у відомої архітектури YOLOV8n на 5%. Це пояснюється більш простою архітектурою та відсутністю об'єднання карт ознак із різних згорткових шарів. Однак відмінність ефективності зменшується зі зменшенням яскравості освітлення. Це пояснюється використанням під час навчання нашої мережі зображень, зроблених при низькому освітленні. Основною перевагою створеної мережі є висока швидкість – до 12 кадрів/с, що у 3 рази вище, ніж для алгоритму YOLOV8n.

Список використаних джерел:

1. Sisavatha C., Yu L.. Design and implementation of security system for smart home based on IOT technology. *Procedia Computer Science*. 2021. Vol. 183. pp. 4-13
2. Taiwo O., Ezugwu A.E., Oyelade O.N., Almutairi M.S. Enhanced Intelligent Smart Home Control and Security System Based on Deep Learning Model. *Wireless Communications and Mobile Computing*. 2022. Vol. 2022. pp.1-22
3. Nugraha G.S., Darmawan M.I., Dwiyanaputra R.. Comparison of CNN's architecture GoogleNet, AlexNet, VGG-16, Lenet-5, Resnet-50 in Arabic handwriting pattern recognition. *KINETIK*. 2023. Vol. 8. No. 2. pp. 545-554
4. Han L., Ma C., Liu Y., Jia J. SC-YOLOv8: A Security Check Model for the Inspection of Prohibited Items in X-ray Images. *Electronics* 2023. Vol.12. pp.1-14
5. Lianga T. and others. Pruning and Quantization for Deep Neural Network Acceleration: A Survey. Elsevier. 2021. Vol.461. pp. 370-403
6. Hafsa F., Shayesta N., Maryam B. Tensorflow-Based Automatic Personality Recognition Used In Asynchronous Video Interviews. *Journal of Engineering Science*. 2022. Vol. 13, Iss. 5. pp.262-268
7. Cao J. and oth. Front Vehicle Detection Algorithm for Smart Car Based on Improved SSD Model. *Sensors*. 2020. Vol. 20. pp.1-21