

СИСТЕМА ІНВАРІАНТНОГО РОЗПІЗНАВАННЯ ОБ'ЄКТІВ НА ЗАЛІЗНИЧНОМУ ТРАНСПОРТІ НА ОСНОВІ НЕЙРОННИХ МЕРЕЖ

Тіганов О. С., аспірант Чорноморського національного університету ім. Петра Могили (м. Миколаїв), e-mail: oleg.tiganov@gmail.com, ORCID: 0000-0003-1297-6230

Кравець І. О., к.т.н., доцент кафедри інтелектуальних інформаційних систем Чорноморського національного університету ім. Петра Могили (м. Миколаїв), e-mail: agniloga@chmni.edu.ua, ORCID: 0000-0002-1361-1017

Фісун М. Т., д.т.н., проф., завідувач кафедри інженерії програмного забезпечення Чорноморського національного університету ім. Петра Могили (м. Миколаїв), e-mail: mykola.fisun@gmail.com, ORCID: 0000-0003-1297-6230

Запропоновано систему розпізнавання об'єкта на зображенні з камери системи відеоконтролю залізничного переїзду на основі згорткових нейронних мереж. Розроблено систему розпізнавання об'єкта на зображенні з використанням архітектур YOLO та SSD. Використано фреймворк Caffe, та бібліотеку OpenCV.

Ключові слова: нейронні мережі, згорткові нейронні мережі, системи комп'ютерного зору.

Вступ. Один з найважливіших напрямків розвитку автоматизації – розвиток систем, що забезпечують безпеку використання залізничного транспорту. Сучасне обладнання більшості залізничних переїздів дає змогу тільки попередити водіїв світловими та звуковими сигналами, перекрити дорогу за допомогою шлагбаумів. Але якщо небезпечна ситуація все ж сталася, співробітники залізниці не мають інформації про стан переїзду. Деякі переїзди обладнані аналоговими системами для контролю стану небезпечної зони, але така система має ряд недоліків. Рішенням проблеми може стати встановлення систем відеоконтролю. У свою чергу, відеонагляд має здійснюватися черговим, це вносить в систему ненадійний людський фактор. Отже, актуальним є створення програмного модуля, який без участі людини на основі даних з системи відеонагляду безпомилково та своєчасно визначить наявність небезпеки та дозволить вжити відповідних заходів для її усунення.

Основний матеріал. Загальна схема процесу пошуку об'єкта на зображенні має наступний вигляд (рис. 1).

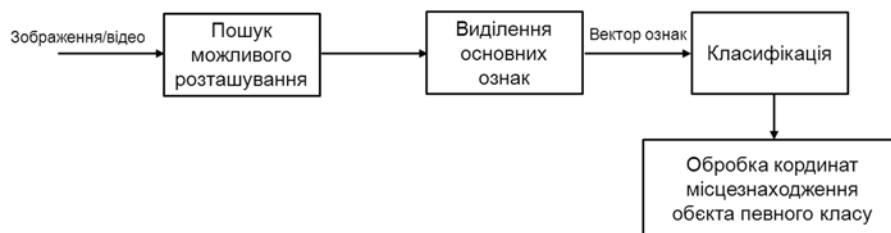


Рисунок 1 – Загальна схема процесу пошуку об'єкта на зображенні

На основі аналізу науково-технічної літератури можна виділити два основних напрями для вирішення проблеми інваріантного розпізнавання образів: часткове (інваріантність щодо повороту, зсуву, масштабу) і загальне (інваріантність щодо форми, ракурсу, текстури). Для задач часткового інваріантного розпізнавання вже зараз існує ряд математичних методів, таких як нейронні мережі, неокогнітрон, згорткові нейронні мережі вищих порядків, статистичні моделі. Основна проблема полягає в практичних складнощах реалізації таких методів. Задача загального інваріантного розпізнавання не вирішена досі. Є окремі успіхи на якість розпізнавання для обмежених класів об'єктів, як правило, обличчя осіб, дорожніх знаків, банківських чеків, які демонструють точність розпізнавання навіть вище, ніж у людини.

Серед найсучасніших методів інваріантного розпізнавання об'єктів можна виділити наступні: метод Віюлі-Джонса та ознаки Хаара; метод головних компонент; неокогнітрон та згорткові нейронні мережі [1].

Перші два методи найкраще підходять для розпізнавання одного класу об'єктів, що має невелику внутрішньокласову мінливість, наприклад, розпізнавання облич.

Метод Віюлі-Джонса [2] використовує зображення в інтегральному уявленні, що дозволяє швидко обчислювати необхідні об'єкти; використовує ознаки Хаара, за допомогою яких відбувається пошук потрібного об'єкта (в даному контексті, обличчя і його рис); використовує бустінг для вибору найбільш оптимальних ознак для шуканого об'єкта на даній частині зображення; всі ознаки надходять на вхід класифікатора, який дає результат «так» або «ні»; використовуються каскади ознак для швидкого відкидання вікон, де не знайдено обличчя.

Метод головних компонент (principal component analysis) – застосовується для витягу з потрібної інформації з даних великої розмірності. При розпізнаванні об'єктів метод PCA дозволяє відобразити великий одновимірний вектор пікселів, побудований з двовимірного зображення одного об'єкта у різних варіантах, в компактні основні компоненти простору ознак. Результат перетворення можна назвати підпростором проєкції. Підпростір розраховується шляхом визначення власних векторів коваріаційної матриці, отриманої з набору зображень об'єктів.

Неокогнітрон та згорткові нейронні мережі, які є його удосконаленням, використовуються для загального інваріантного розпізнавання об'єктів.

Проведений аналітичний огляд інтелектуальних методів розпізнавання фотографічних зображень та виділення об'єкта з відео потоку показує, що найкращі результати на сьогодні дають згорткові нейронні мережі. Згорткові нейронні мережі (ЗНМ) враховують просторову топологію зображення; не чутливі до спотворень зображення; мають високу швидкість навчання і роботи (особливо при використанні графічних процесорів).

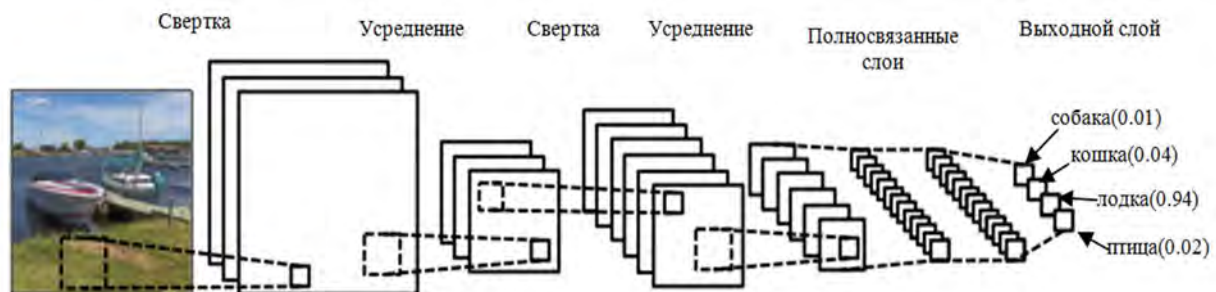


Рисунок 2 – Загальний вигляд загорткової нейронної мережі

Архітектура ЗНМ складається з багатьох шарів. Є два основні типи шарів: згорткові (Convolutional) і підвибіркові (Subsampling), Згорткові та підвибіркові шари чергуються один з одним. Але останнім часом є тенденція до скорочення кількості підвибіркових. Тому підвибіркових шарів часто менше, ніж згорткових або немає взагалі.

У кожному загортковому шарі є набір з декількох фільтрів (ядер згортки), які мають невелике рецептивне поле, але займають усю глибину вхідного об'єму, причому нейрони одного фільтра мають однакові ваги, що застосовуються до всіх локальних ділянок попереднього шару (тим самим здійснюючи згортку). Зображення попереднього шару ніби сканується невеликим вікном і помножується на набір ваг, а результат зберігається на відповідному нейроні поточного шару. Таким чином набір площин є картою ознак (feature maps), де кожна площина знаходить «свої» ознаки зображення в будь-якому місці попереднього шару. На рис. 3 представлена схема нейронів загорткових шарів

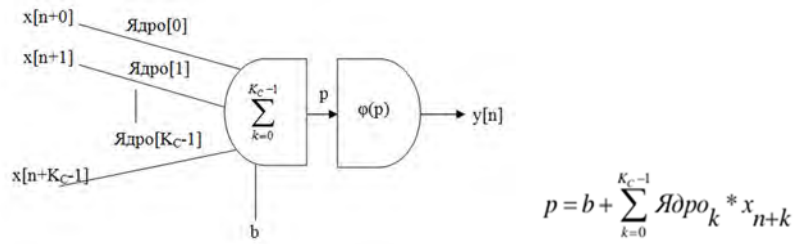


Рисунок 3 – Загальна схема нейрона згорткового шару

На рис. 3 прийнято наступні умовні позначення: КС – загальна кількість нейронів, що входять до рецептивного поля n-го нейрона згорткового шару, ядро [k] – ваги нейрона, що налаштовуюються, b – зсув n-го нейрона, причому b і ядро [k] – одне й те ж для всієї карти згорткового шару і виконує роль фільтру, який шукає певний фрагмент зображення. x[n + k] – вхідні дані для n-го нейрона згорткового шару, k = 0...КС – 1. Вхідні дані x[n + k] дають зважену суму з параметрами, що налаштовуюються, тобто з ядром [k], тобто відбувається операція згортки:

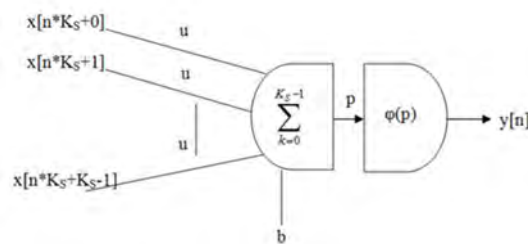


Рисунок 4 – Загальна схема нейрона підвборочного шару

Наступний за згортчим шаром, підвборочний шар зменшує масштаб площин даних, для полегшення виділення основних ознак в наступних шарах. Найефективнішою є «максимальна підвбірка». Остання частина згорткової мережі – це класифікатор ознак, як правило, представлений кількашаровим перцептроном або опорно-векторною машиною. Кількість нейронів у шарі-класифікаторі визначається кількістю класів, до яких слід віднести вхідне зображення.

Було застосовано та порівняно два найсучасніших алгоритму згорткових нейронних мереж SSD [3, 4] та YOLO [4], що мають найвищі швидкість та точність детекції.

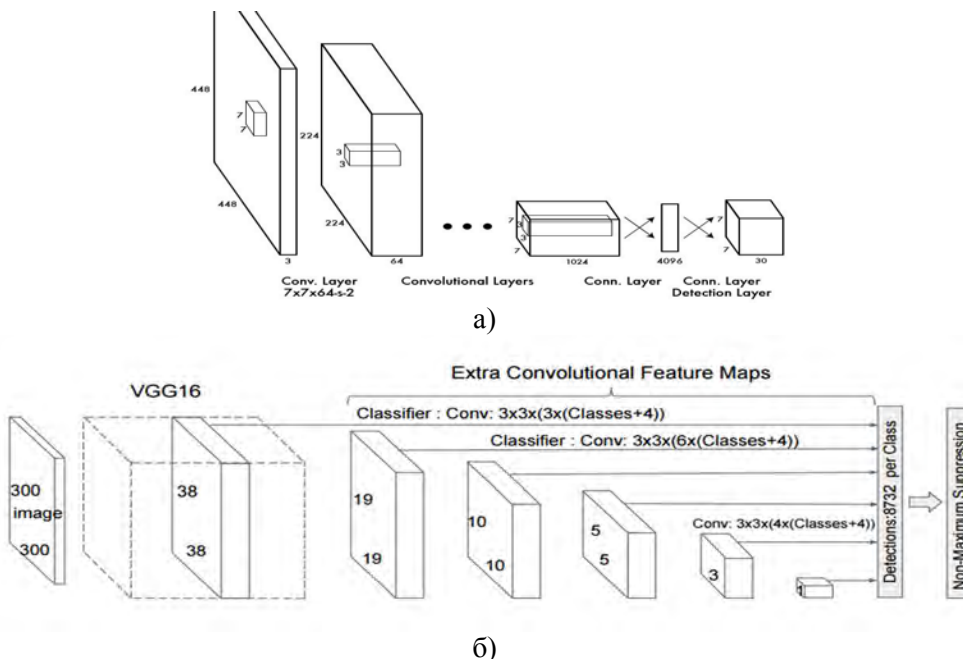


Рисунок 5 – Архітектури YOLO (a) та SSD (б) [2, 3]

YOLO дуже швидка. Під час тестування на одному з найпотужніших графічних процесорів (Nvidia Titan X GPU) базова мережа обробляє 45 кадрів за секунду. YOLO оцінює усе зображення, коли робить прогноз, на відміну від методів, що базуються на ковзаючому вікні або пропозиціях регіонів. Система поділяє зображення за допомогою сітки та для кожної клітинки передбачає декілька обмежувальних рамок і ймовірність наявності в них об'єкта певного класу.

SSD – є наступним поколінням ЗНМ для розпізнавання об'єктів, вона також розглядає зображення цілком, але розглядає сітки різних масштабів. Додавши ряд поліпшень, вдалося значно підвищити точність, порівняно з попереднім варіантом, не втративши при цьому у швидкості розпізнавання. Застосовуючи ці фільтри до різних карт ознак, отриманих з різних шарів на пізніших етапах, мережі дозволяють виконувати детекцію у різних масштабах. Це вирішує значну проблему попередніх мережі і дозволяє досягти високої точності виявлення об'єктів різних форм та розмірів на зображеннях з порівняно низькою роздільністю.

Для реалізації було обрано архітектуру SSD, яка дає найбільшу точність при прийнятній швидкості роботи навіть на відносно «слабких» конфігураціях обладнання.



Рисунок 6 – Результат обробки зображення переїзду мережею SSD

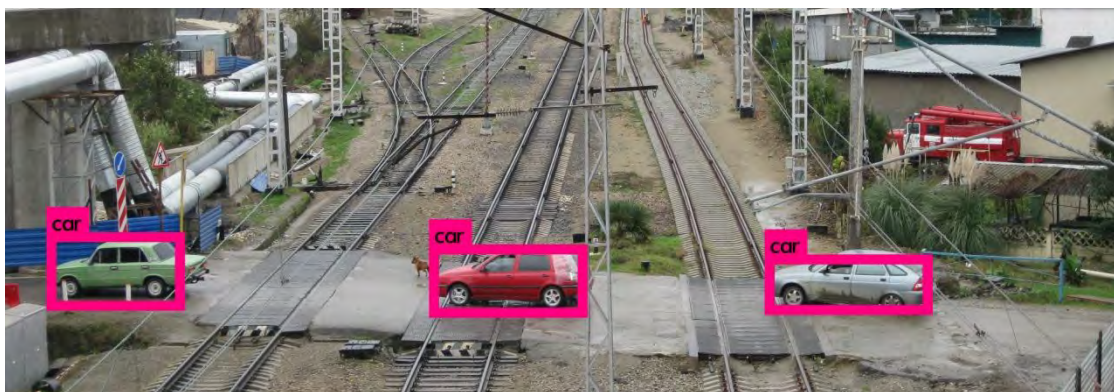


Рисунок 7 – Результат обробки зображення переїзду мережею YOLO

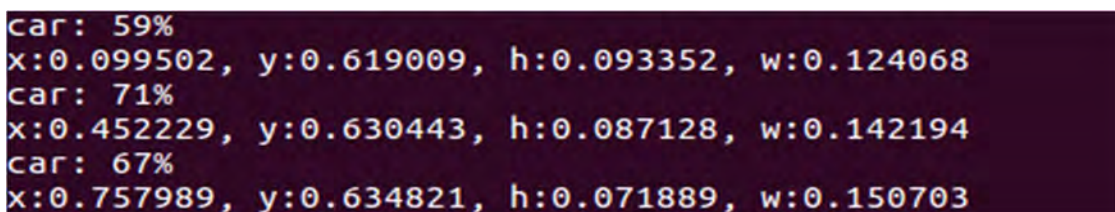


Рисунок 8 – Консольний вигляд результатів обробки зображення переїзду мережею YOLO

На основі архітектури SSD було розроблено програмний додаток для розпізнавання об'єкта на зображенні з камери системи відеоконтролю залізничного переїзду. Додаток створений на основі фреймворку Caffe [5–8], що використовує бібліотеку OpenCV для роботи з відеопотоком.

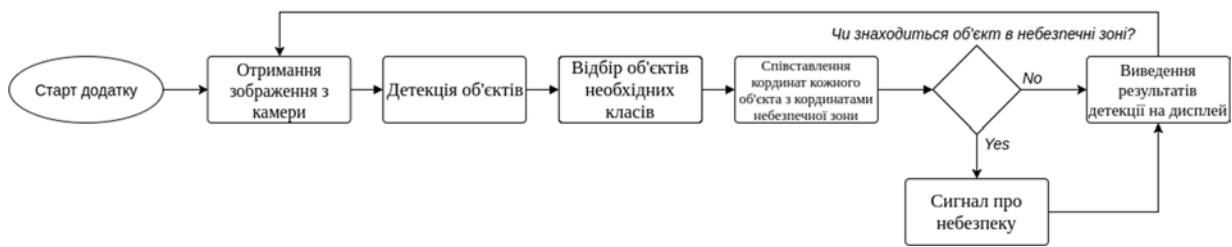


Рисунок 9 – Блок-схема розробленої системи інваріантного розпізнавання об'єктів

Процес роботи додатку можна розділити на декілька етапів. У першу чергу, система отримує на вхід зображення, що, у випадку з системами реального часу, витягується з відеопотоку. В ролі інструмента для цього, а також для подальшої обробки зображення та виведення результатів на екран, використано бібліотеку OpenCV [9, 10]. Наступним кроком, після того як зображення отримано, є його обробка за допомогою певної архітектури згорткової нейронної мережі, тобто локалізація об'єктів. Для визначення місцезнаходження об'єкта та класу, якому він належить, було використано фреймворк глибокого машинного навчання Caffe, що надає можливості для навчання та розгортання моделі нейронної мережі за підготовленою заздалегідь архітектурою.

Для навчання нейронної мережі для локалізації об'єктів використовується метод навчання з учителем, тому що він відповідає потребам алгоритму (використовується невелика кількість класів і він легший у використанні). Для методу навчання з учителем потрібен величезний набір розмічених даних (в нашому випадку – це набір зображень з вручну поміченим місцезнаходженням певних об'єктів). На щастя в наш час у вільному доступі є величезна кількість вільних даних, серед яких є набори, які підійдуть для навчання. Але такі набори зазвичай включають у себе декілька десятків класів об'єктів, тоді як для додатку, що розробляється, достатньо лише декількох (людина, автомобіль, автобус, мотоцикл). Це робить обґрунтованим подальше видозмінення набору і параметрів мережі, та навчання мережі на зміненому наборі даних. Проте це потребує розробки окремого алгоритму обробки даних набору.

Також навчання настільки глибокої нейронної мережі потребує великих об'ємів обчислюваних ресурсів, використання яких економічно недоцільне при розробці прототипу додатку. Тому при розробці та тестуванні додатків використовувалася заздалегідь навчена модель. При подальшому використанні системи можливо здійснити навчання нової моделі на наборі, спеціально підготовленому для задачі, що вирішується за її допомогою. Доцільно використати для навчання сервіси, що надають послуги хмарних обчислень.

Додаток потребує Unix-подібної операційної системи з встановленими пакетами libprotobuf-dev, libleveldb-dev, libsnappy-dev, libopencv-dev, libhdf5-serial-dev, protobuf-compiler. Запускається у консольному режимі за допомогою Python-скрипта. Робоча інформація і повідомлення про небезпеку виводяться в консоль. Зображення з камери з накладеними на нього обмежувальними рамками знайдених об'єктів а також полігон небезпечної зони виводяться на екран у окремому вікні «railroad crossing» (рис. 10).



Рисунок 10 – Приклад зображення, що виводиться на екран при роботі додатку

Висновки. У результаті аналізу літератури було виявлено, що найбільш ефективним з інтелектуальних методів інваріантного розпізнавання об'єктів з відео потоку є згорткові нейронні мережі. Серед них найбільш удосконалені архітектури є YOLO та SSD. На основі архітектури SSD розроблено систему розпізнавання об'єкта на зображенні з камери системи відеоконтролю залізничного переїзду.

REFERENCES

1. Krizhevsky, I. Sutskever, and G. Hinton. ImageNet classification with deep convolutional neural networks. In NIPS, 2012.
2. Ole, H.J. Implementing the Viola-Jones face detection algorithm: Master's thesis / H.J. Ole. – DTU, 2008. <http://old.kv.by/index 2001513402>.
3. Redmon, J., Divvala, S., Girshick, R., Farhadi, A. You only look once: Unified, real-time object detection. In: CVPR. (2016)
4. Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng Yang Fu, Alexander C. Berg. «SSD: SingleShotMultiBoxDetector» in European Conference on Computer Vision (ECCV), 2016.
5. Caffe; Deep learning framework by the BVLC [Електронний ресурс]. – Режим доступу : <http://caffe.berkeleyvision.org>. (Дата звернення: 15.11.2017.)
6. Caffe: A fast open framework for deep learning [Електронний ресурс]. – Режим доступу : <https://github.com/BVLC/caffe> (Дата звернення: 15.11.2017.)
7. LeCun, Y. Gradient-based Learning Applied to Document Recognition [Text] / Y. LeCun, L. Bottou, Y. Bengio, P. Haffner // Proceedings of the IEEE. v. 86, 1998, pp. 2278–2324.
8. Fukushima, K. Neocognitron: A neural network model for a mechanism of visual pattern recognition [Text] / K. Fukushima, S. Miyake, T. Ito // IEEE Trans. Syst., Man, Cyber., V. SMC-13. – 1983. – P. 826–834.
9. LeCun, Y. Backpropagation applied to handwritten zip code recognition [Text] / Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, L. D. Jackel // Neural Computation, vol. 1, no. 4, 1989, pp. 541–551.
10. LeCun, Y. Backpropagation applied to handwritten zip code recognition [Text] / Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, L. D. Jackel // Neural Computation, vol. 1, no. 4, 1989, pp. 541–551.

Тиганов О. С., Кравец І. А., Фісун Н. Т. СИСТЕМА ИНВАРИАНТНОГО РАСПОЗНАВАНИЯ ОБЪЕКТОВ НА ЖЕЛЕЗНОДОРОЖНОМ ТРАНСПОРТЕ НА ОСНОВЕ НЕЙРОННЫХ СЕТЕЙ

Предложена система распознавания объекта на изображении с камеры системы видеоконтроля железнодорожного переезда на основе сверточных нейронных сетей. Разработана система распознавания объекта на изображении с использованием архитектур YOLO и SSD. Использован фреймворк Caffe, и библиотека OpenCV/

Ключевые слова: *нейронные сети, сверточные нейронные сети, системы компьютерного зрения.*

Tihanov O., Kravets I., Fisun M. THE SYSTEM OF OBJECT'S INVARIANT RECOGNITION FOR RAILWAY TRANSPORT BASED ON NEURAL NETWORKS

The system of object' invariant recognition on the camera image at a railway crossing, based on convolutional neural networks, is proposed. The developed system detects an objects on the image, using YOLO and SSD architectures, the Caffe framework and the OpenCV library.

Keywords: *neural networks, convolution neural networks, object' invariant recognition, computer vision.*

© Тіганов О. С., Кравец І. О., Фісун М. Т.

Статтю прийнято
до редакції 2.06.17