



Б. П. Борківський, В. М. Теслюк

Національний університет "Львівська політехніка", м. Львів, Україна

ВИКОРИСТАННЯ НЕЙРОМЕРЕЖЕВИХ ЗАСОБІВ ДЛЯ РОЗПІЗНАВАННЯ ОБ'ЄКТІВ У МОБІЛЬНИХ СИСТЕМАХ З ОБХОДОМ ПЕРЕШКОД

Запропоновано вдосконалення нейромережевого методу розпізнавання об'єктів для пошуку доречних об'єктів в межах заданого контексту. Проаналізовано останні дослідження та підходи до вирішення проблеми розпізнавання об'єктів і виявлено ключові особливості декількох архітектур нейронних мереж. Наведено метод, що базується на поєднанні нейронної мережі, пристосованої до пошуку образів, та застосуванні додаткових алгоритмів для вилучення зайвих об'єктів, а саме алгоритму видалення дублікатів об'єктів та алгоритму фільтрації об'єктів відповідно до контексту. Застосовано готову архітектуру Single Shot Detector як основу нейронної мережі для розпізнавання об'єктів, треновану на наборі даних COCO. Розроблено алгоритм фільтрації об'єктів, який містить попереднє створення конфігурації, для визначення множин доречних класів об'єктів для різних контекстів. Розроблено гнучку конфігурацію, яка може бути модифікована та розширена новими умовами роботи, за потреби, без додаткових змін у логіці програмної частини. Проведено низку експериментів для визначення точності роботи системи у нових середовищах, під час яких систему протестовано з використанням набору даних ODSR-IHS, який містить зображення, отримані під час роботи роботів-порохотягів, і досягнуто точності розпізнавання 5 класів об'єктів на рівні 83 %. Виявлено під час проведення експериментів здатність системи знаходити більшу кількість значущих об'єктів, ніж передбачалось авторами набору даних, зокрема людей у кадрі, що свідчить про перспективність використання системи у різному оточенні. Встановлено, що завдяки використанню диверсифікованого набору даних із 80 класами об'єктів та відбором тільки значущих предметів розроблена система є ефективною під час використання у різних середовищах. Запропонований метод вирішує проблему перевикористання єдиної системи розпізнавання об'єктів у різних середовищах, що спрощує її використання у мобільних системах, які можуть працювати у різному оточенні.

Ключові слова: оброблення зображень; згортоква мережа; виявлення об'єктів; контекстне видалення.

Вступ / Introduction

Здатність сучасних програмно-апаратних систем аналізувати стан навколишнього середовища та відповідно реагувати на нього є невід'ємною складовою частиною сьогодення, що дає змогу автоматизувати рутинні завдання чи такі, які вимагають постійної концентрації, якої людина не може досягти. У випадку з завданнями теорії комп'ютерного зору – його здатність розпізнавати складні образи і предмети з візуальних даних (зображення, відео) досягається з використанням нейромережевих підходів.

Сучасні системи розпізнавання технічних об'єктів корисні в різних сферах науки та техніки, таких як системи безпеки, системи виявлення порушень, соціальні та робототехнічні системи тощо. Для того, щоб підвищити якість розпізнавання об'єктів, системи розробляються для виконання конкретного завдання, яке передбачає використання невеликого обсягу даних, а саме обмежену кількість типів об'єктів для розпізнавання. У такому разі для розв'язання іншої задачі необхідно проєктувати окрему систему. Тобто такий підхід не є уні-

версальним і не може бути використаний у разі великого обсягу вхідних даних, чи на пристрої з обмеженим обчислювальними ресурсами.

Прикладом задачі, яка потребує узагальненої дії, може бути роботизована система, яка допомагає людям з вадами зору під час пошуку речей та ін. Якщо така система може знайти тільки кілька об'єктів, її практична цінність буде низькою. Але, якщо така система може знаходити багато речей, які часто використовуються, і може працювати як у приміщенні, так і на вулиці, тоді її спроможність допомагати людям буде значно вищою.

Також існує проблема у використанні систем, які вміють працювати з великою кількістю різних об'єктів. За певних обставин нам потрібно враховувати тільки певний набір об'єктів. Наприклад, під час оброблення відео з дорожньої камери дані можуть бути використані в системі для відстеження людей на перехрестях, і потрібно враховувати пішоходів і велосипедистів, або ж вона може бути використана для пошуку транспортних засобів, і потрібно шукати тільки автомобілі та автобуси. Саме тому розроблення єдиного алгоритму, який може обробляти одну й ту саму інформацію по-різному,

Інформація про авторів:

Борківський Богдан Петрович, аспірант, кафедра автоматизованих систем управління. Email: bohdan.p.borkivskyi@lpnu.ua;
<https://orcid.org/0000-0003-3301-476X>

Теслюк Василь Миколайович, д-р техн. наук, професор, завідувач кафедри автоматизованих систем управління.
Email: vasyi.m.teslyuk@lpnu.ua; <https://orcid.org/0000-0002-5974-9310>

Цитування за ДСТУ: Борківський Б. П., Теслюк В. М. Використання нейромережевих засобів для розпізнавання об'єктів у мобільних системах з обходом перешкод. Науковий вісник НЛТУ України. 2023, т. 33, № 4. С. 84–89.

Citation APA: Borkivskyi, B. P., & Teslyuk, V. M. (2023). Application of neural network tools for object recognition in mobile systems with obstacle avoidance. *Scientific Bulletin of UNFU*, 33(4), 84–89. <https://doi.org/10.36930/40330412>

тобто одночасно працювати у різних середовищах, є актуальним завданням.

Об'єкт дослідження – процес розпізнавання значущих об'єктів у різних середовищах.

Предмет дослідження – методи та засоби розпізнавання об'єктів з виокремленням релевантних об'єктів у різних середовищах.

Мета роботи – підвищити ефективність розпізнавання об'єктів у мобільних системах з обходом перешкод у спосіб розроблення узагальненого алгоритму пошуку об'єктів у різних середовищах.

Для досягнення зазначеної мети визначено такі *основні завдання дослідження*:

1. Здійснити літературний аналіз методів і засобів розпізнавання об'єктів.
2. Обрати оптимальну архітектуру нейронної мережі для пошуку об'єктів.
3. Розробити узагальнений алгоритм пошуку об'єктів у різних середовищах з використанням рішення з пункту 2.
4. Оцінити точність алгоритму.

Аналіз останніх досліджень та публікацій. Під час розроблення систем комп'ютерного зору велику увагу потрібно приділити вибору моделі, яка буде покладена в основу її алгоритму, адже від характеристик обраної моделі залежить якість та швидкість роботи системи.

Розробники в галузі штучного інтелекту та комп'ютерного зору постійно вдосконалюють наявні та намагаються винайти нові підходи [2], і є декілька, які набули широкого застосування в усьому світі, оскільки довели свою ефективність у прикладних задачах. Однак кожен з алгоритмів має як сильні, так і слабкі позиції, які необхідно дослідити, щоб визначити, який з алгоритмів найкраще підходить для розв'язання поставленої задачі.

Дослідники Університету Берклі запропонували підхід R-CNN [5], спрямований на вирішення проблеми з виділення величезної кількості регіонів. Науковці запропонували метод, за допомогою якого використовується вибіркового пошуку для отримання тільки 2000 регіонів із зображення, які назвали пропозиціями регіонів (regions proposals). Тому замість того, щоб намагатися класифікувати величезну кількість регіонів, можна просто працювати з 2000 регіонами.

Цей підхід використовує вибіркового пошуку для пошуку пропозицій регіону. Вибірковий пошук – це повільний і трудомісткий процес, що впливає на продуктивність мережі. Тому дослідники з Microsoft Research розробили алгоритм Faster R-CNN [6] для виявлення об'єктів, який усуває алгоритм вибіркового пошуку та дає змогу мережі самостійно вивчати пропозиції регіонів. Сильною позицією моделі Faster R-CNN є висока точність розпізнавання об'єктів, оскільки архітектурною особливістю цієї моделі є наявність одразу двох підмереж [8, 10]. Одна виконує звичайну задачу використання згорткової мережі для отримання карти ознак із вхідних даних. Роль другої підмережі виконує мережа пропозицій регіонів (Region Proposal Network), яка замість заздалегідь визначених регіонів шукає відповідні регіони самостійно, що в поєднанні з великою навчальною вибіркою дає високий результат на тестових зображеннях.

YOLO – ще одна мережа, яку часто використовують для вирішення завдань пошуку об'єктів. Перевага цієї моделі [1] відображена в назві – "ви дивитесь тільки один раз", тобто вона не має недоліків Fast R-CNN. Мо-

дель побудована переважно з використанням згорткових шарів, за винятком кількох FC шарів.

SSD (англ. *Single Shot Detector*) [4] має дещо подібну до YOLO методологію, що також відображено в його назві – "single frame detector" – "однокадровий детектор". Мережа складається зі згорткових шарів, розділених на дві частини. Перша частина, яка базується на згортковій нейронній мережі VGG-16 [3], є типовою моделлю класифікації, і використовується для отримання карти ознак, після чого, використовуючи додаткові згорткові шари різного розміру, отримуємо додаткові карти ознак меншого розміру.

Результати дослідження та їх обговорення / Research results and their discussion

Вибір архітектури моделі розпізнавання образів.

Описані вище архітектурні особливості одночасно є слабким місцем мережі R-CNN. Оскільки в роботі моделі беруть участь дві підмоделі, це зменшує швидкість роботи всієї моделі і, відповідно, час оброблення кожного зразка.

Оскільки зображення проходить через шари мережі YOLO тільки один раз, це зменшує час роботи моделі та збільшує продуктивність. Але ця модель не позбавлена недоліків – через особливості роботи моделі та організації її шарів, зокрема останнього шару розміром 7x7 та можливості побудови тільки 2 регіонів для кожної локації – загальна кількість запропонованих регіонів становить тільки 98. Через ці просторові обмеження моделі важко знайти невеликі об'єкти, або об'єкти, які з'являються групами.

Через особливості організації шарів у мережі SSD, на виході отримуємо 8732 регіонів (проти 98 в YOLO), що значно збільшує кількість потенційних кандидатів регіонів.

Для порівняння ефективності роботи моделей розпізнавання образів використовується метрика mAP (англ. mean average precision). mAP поєднує в собі компроміс між влучністю (precision) та повнотою (recall) і враховує як хибнопозитивні, так і хибнонегативні спрацьовування. Ця властивість робить mAP придатною метрикою для більшості задач розпізнавання. mAP обчислюють шляхом знаходження середньої точності (AP) для кожного класу, після чого усереднюється для кількох класів:

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i.$$

Для кращого розуміння ефективності кожної моделі ознайомимося з результатами [7] тестування моделей (табл. 1) на однакових даних.

Табл. 1. Порівняння роботи моделей /
Performance comparison of models

Модель	Швидкість (кадри за секунду)	Точність (mAP)
Faster R-CNN	7	73,2
YOLO	45	63,4
SSD	59	74,3

Враховуючи розглянуті особливості моделей, їх переваги та недоліки, які підтверджуються результатами порівняльного тестування, найкращим варіантом моделі є SSD.

Вибір набору даних. Реалізації системи з використанням методів штучного інтелекту типу "навчання з

учителем" важливо зібрати якісний масив даних, який згодом буде використаний для навчання моделі. Основним критерієм у виборі масиву даних є його різноманітність, як за типом об'єктів, так і за кількістю унікальних екземплярів. Оскільки практичного застосування в системах розпізнавання декількох класів об'єктів не буде, кількість класів повинна починатися від декількох десятків одиниць.

MS COCO (англ. *Microsoft Common Objects in Context*) – це масштабний набір даних для розв'язання задач розпізнавання та сегментації об'єктів, створення підписів до зображень, який створено за підтримки таких ІТ-компаній, як Microsoft та Facebook [9].

Оновлена версія цього набору, випущена у 2017 р., містить 120 тис. зображень, які розділені на 80 класів. На відміну від попередніх наборів, у цьому наборі даних зберігається баланс між різними типами класів, а для маркування побутових об'єктів доступно близько 30 класів, що дає змогу створити добре узагальнену систему розпізнавання.

Розроблення узагальненого алгоритму розпізнавання об'єктів. Типовий конвеєр виявлення об'єктів має один компонент для створення пропозицій для класифікації. Пропозиції – це не що інше, як регіони-кандидати для об'єкта, який нас цікавить. Більшість підходів використовують ковзне вікно над картою ознак і призначають оцінку достовірності залежно від ознак, обчислених у цьому вікні. Сусідні вікна мають дещо схожі оцінки і вважаються регіонами-кандидатами. Це призводить до появи сотень пропозицій. Оскільки метод створення пропозицій повинен мати високу повноту, на цьому етапі ми залишаємо вільні обмеження. Однак оброблення цих множинних пропозицій за допомогою класифікаційної мережі є громіздкою. Це призводить до використання методу, який фільтрує пропозиції на підставі критерію IoU (англ. *Intersection over Union*), що називають Non-maximum Suppression [11].

Значення IoU обчислюють за формулою

$$IoU = \frac{A \cap B}{A \cup B},$$

де: A – зона покриття першого об'єкта; B – зона покриття другого об'єкта.

Для забезпечення узагальненої роботи системи використовуються модель, здатна розпізнавати різні типи об'єктів. Але в деяких ситуаціях недоцільно шукати всі можливі об'єкти: автомобіль, який видно далеко у вікні кімнати, або побутову техніку, коли автоматизована система пересувається кімнатою.

Для забезпечення пошуку тільки релевантних об'єктів алгоритм розпізнавання об'єктів розширено кроком фільтрації за режимом роботи, який порівнює інформацію про кожен окремо взятий класифікований об'єкт і фільтрує її відповідно до інформації про класи вказаного режиму. Відповідно кількість релевантних об'єктів для кожного режиму визначається за формулою

$$k_i = \sum_{j=1}^n N_j \in K_i,$$

де: k_i – кількість релевантних об'єктів; n – кількість попередньо знайдених об'єктів; N_j – клас j -го об'єкта; K_i – множина класів i -го режиму роботи.

В основу цього алгоритму фільтрації покладено створення відповідної конфігурації, виходячи з того, який тип пошуку використовується. Наприклад, для

відстеження пішоходів відповідними класами будуть людина, велосипедист і т.ін., для відстеження транспорту відповідними класами – автомобіль, автобус, фургон, скутер і т.ін. Вхідними даними для алгоритму є список раніше знайдених об'єктів, причому, виходячи з режиму роботи системи, враховуються тільки відповідні об'єкти з конфігурації.

Основні кроки алгоритму:

Крок 1. На вхід алгоритм отримує масив B зі списком пропозицій, відповідними значеннями впевненості S та певним пороговим значенням N .

Крок 2. Вибираємо пропозицію з найбільшим значенням впевненості і переносимо її із вхідного масиву у вихідний масив D .

Крок 3. Порівнюємо цей запис з рештою записів вхідного масиву, обчислюючи значення метрики IoU. Якщо отримане значення більше від заданого порогового значення (елементи дуже схожі) – видаляємо цю пропозицію.

Крок 4. Повторюємо кроки 2-3 доти, поки вхідний масив не стане порожнім.

Крок 5. Повертаємо вихідний масив D з унікальними пропозиціями регіонів.

Крок 6. Кожен клас знайдених об'єктів з масиву D порівнюється з множиною класів режиму K згідно з конфігурацією, і видаляється з масиву, якщо цей клас не є елементом множини.

Вихідний масив D містить тільки унікальні та релевантні розпізнані об'єкти.

Визначення точності алгоритму. Для перевірки працездатності узагальненої системи розпізнавання образів вирішено випробувати її для розпізнавання об'єктів у полі зору мобільного робота-порохотяга. У цьому дослідженні використано масштабний загальнодоступний еталонний набір даних під назвою "Object Detection for Sweeping Robots In the Home Scene" (ODSR-IHS). Набір даних містить 6 тис. зображень і 16 тис. екземплярів, поділених на 14 категорій. Дані в основному відображають складне середовище у приміщенні, з різними формами цілей і значною часткою малих цілей.

Цей набір даних містить 14 класів, серед яких присутні класи меблів, таких як ліжко, стіл, стілець, і більш специфічні для роботів-порохотягів класи – дроти і шкарпетки. Попри те, що за основу алгоритму обрано модель, треновану на наборі даних COCO, який містить 90 класів об'єктів для розпізнавання, перетин класів наборів даних COCO та ODSRIHS є невеликим. Тому для визначення якості розпізнавання об'єктів алгоритмом відбудеться валідація пошуку усіх класів, та тих, які є спільними для усіх наборів даних.

У межах цього дослідження проведено декілька експериментів, з використанням різної кількості класів запропонованого набору даних, результати яких наведено у табл. 2. Під час першого проведеного експерименту отримано точність розпізнавання на рівні 48,7 %. Для розуміння такого низького показника проаналізовано коректність роботи реалізованих алгоритмів, і помилок не виявлено. Під час аналізу набору даних встановлено численні помилки анотації даних – замість об'єктів 14 класів для розпізнавання у наборі даних містилось 33 мітки класу, що більш ніж удвічі перевищує очікуваний показник. Причиною такої невідповідності виявились помилки назв класів – зайві символи у назві класи, хибне кодування назв класів, використання китайської мови для анотацій. Для отримання достовірних результатів усі помилки усунено.

Після виправлення помилок з анотаціями проведено новий експеримент (рис. 1) для встановлення точності розпізнавання, і за результатами обчислення розпізнавання усіх 14 класів об'єктів показник середньої точності розпізнавання зріс до 62,0 %.



Рис. 1. Приклад незнайденого предмета – дроту / Example of missed object – wire

Для розуміння точності розпізнавання відомих класів об'єктів, тобто тих, які присутні в обох наборах даних, на абсолютно нових даних іншої природи, аніж тих, які використовувались при навчанні, проведено експеримент із залученням виключно спільних класів об'єктів, а саме ліжко, диван, стіл, стілець, книга. За результатами проведеного експерименту (рис. 2) показник середньої точності розпізнавання об'єктів досягнув позначки 82,9 %.



Рис. 2. Успішний пошук усіх меблів / Successful detection of all furniture pieces

Табл. 2. Порівняння результатів експериментів / Comparison of experiment results

	Опис класів	Точність (mAP)
Експеримент 1	33 класи набору даних без змін (з помилками)	48,7
Експеримент 2	14 класів – усі класи набору даних з виправленнями	62,0
Експеримент 3	5 класів, які є спільними для обох наборів даних	82,9

Оскільки під час навчання моделі використовувався різноманітніший набір даних, ніж під час тестування, додатково проведено експерименти з виявлення нових типів об'єктів, не анотованих авторами набору даних (рис. 3 і 4). Враховуючи відсутність міток класів для нових типів об'єктів, неможливо поррахувати точність розпізнавання під час таких експериментів, тільки зроби́ти суб'єктивну оцінку якості розпізнавання.

На відміну від наявних підходів, спрямованих на використання вузькоспеціалізованих моделей розпізнавання об'єктів, здатних розрізняти тільки декілька класів, запропонований алгоритм знаходження об'єктів дає

змогу працювати одночасно із сотнею різних класів. Під час перевірки справності роботи узагальненого алгоритму розпізнавання об'єктів використано спеціалізований набір даних ODSR-IHS, який містить анотації до 14 класів об'єктів.



Рис. 3. Розпізнавання предметів на кухні. Зелений колір – правильно знайдений предмет, жовтий колір – додаткові знахідки розробленого алгоритму, не анотовані в наборі даних / Objects recognition in the kitchen. Green colour – item detected correctly, yellow colour – extra findings of developed algorithm not annotated in dataset



Рис. 4. Розпізнана людина / Detected person

Оскільки більшість класів не є відомими для розробленого алгоритму, йому не вдалось ідентифікувати такі об'єкти, що негативно вплинуло на точність розпізнавання. Порівнюючи еталонне для розробленого алгоритму значення точності 74,3 % при розпізнаванні 90 відомих класів об'єктів, з результатом експерименту розпізнавання 14 класів, з яких деякі є невідомими, отримано нижче значення точності 62,0 %, що свідчить про вплив кількості невідомих класів об'єктів на точність визначення предметів. Підтвердженням цього є результати іншого експерименту, де розпізнавалась обмежена кількість типів об'єктів, а саме 5 класів, спільних для обох наборів даних. За такого обмеження, алгоритм продемонстрував точність розпізнавання на рівні 82,9 %, що на 10 % більше, ніж оригінальна точність на відомих даних, хоча кількість класів, яка бралась до уваги, значно менша. Це означає, що цей узагальнений алгоритм розпізнавання нездатен замінити вузькоспеціалізовану модель знаходження специфічних типів об'єктів (див. рис. 1), але в задачах неспеціалізованого пошуку показує високі результати (рис. 2).

Така невелика кількість класів використаного набору даних не може покрити більшості об'єктів, які зус-

трічались на зображеннях. Водночас розроблений алгоритм знаходив більше об'єктів, ніж ановано авторами набору даних, що дає змогу отримати більше інформації зі зображення, ніж моделлю, тренованою виключно на даному наборі даних (див. рис. 3).

Додаткова інформація, знайдена розробленим алгоритмом, окрім меблів, містить предмети побуту, спортивне знаряддя, кухонне приладдя (наприклад духову шафу, мікрохвильову піч, холодильник), що яскраво демонструє успішність роботи розробленого узагальненого алгоритму пошуку об'єктів.

Вагомим є також здатність моделі знаходити на зображенні людей (див. рис. 4), що істотно спрощує використання роботів-асистентів, які розуміють що перед ними не перешкода, а безпосередньо об'єкт взаємодії.

Обговорення результатів дослідження. Порівнюючи отримані результати дослідження з працями [1, 4, 5, 6, 10], можна дійти висновку, що запропонований метод розпізнавання об'єктів додає гнучкості у використанні та покращує ефективність пошуку предметів.

Попри високу точність розпізнавання розмитих зображень 82,8 %, завдяки використанню підмережі пропозицій регіонів RetinaRPN у роботі [5], оновлена підмережа не здатна визначати доцільність запропонованих регіонів, а отже запропоноване рішення не є ефективним у системах, що працюють в різних середовищах.

У роботі [6] автори змінили розмір вхідних і вихідних шарів обох підмереж моделі, що дало змогу одночасно опрацьовувати декілька зображень. Попри удосконалення точності пошуку об'єктів, порівняно з базовою моделлю, автори досягли точності тільки 42,6 %.

Хоча у роботі [10] використовується воксельна підмережа пошуку регіонів-кандидатів, яка істотно покращує якість розпізнавання об'єктів – до 84,28 %, а також дещо оптимізує швидкість роботи, а саме до 16 кадрів на секунду, проте такий підхід все ще істотно повільніший за розроблений метод, який базується на моделі SSD, а тому загальна ефективність нашого рішення є значно вищою.

У праці [1] автори додатково використовують мережу агрегування шляхів і досягли точності розпізнавання 74,5 %, але результати є контекстно незалежними, що не дає змоги використовувати його у системах зі змінним контекстом. Використання блоку спостереження за ознаками, наведене у роботі [4], дає змогу досягнути такої ж точності розпізнавання, що і метод фільтрування релевантних об'єктів, але попри високу точність розпізнавання самих об'єктів не здатен визначати доцільність передбачень.

Отже, за результатами виконаної роботи можна сформулювати такі наукову новизну та практичну значущість результатів дослідження.

Наукова новизна отриманих результатів дослідження – вдосконалено метод розпізнавання об'єктів з використанням нейронних мереж шляхом створення та впровадження алгоритму фільтрації результатів роботи моделі відповідно до заданого контексту.

Практична значущість результатів дослідження – вдосконалий метод пошуку релевантних об'єктів можна використовувати у роботизованих системах військового спрямування за різних умов, наприклад під час розмінувальних робіт для виявлення небезпечних об'єктів чи під час рятувально-пошукових робіт, а та-

кож у роботизованих системах соціального спрямування, наприклад у роботах-асистентах.

Висновки / Conclusions

Унаслідок виконаної роботи досліджено наявні програмні рішення для вирішення задачі пошуку об'єкта на зображенні, проаналізовано наявні підходи до проектування та побудови нейронних мереж для розпізнавання об'єктів. Проведено порівняльну характеристику відомих моделей розпізнавання об'єктів, а саме SSD (англ. *Single Shot Detector*), YOLO та Faster R-CNN, та проведено експерименти для визначення якісних характеристик запропонованих моделей. Як наслідок – вибір зроблено на моделі SSD.

Для забезпечення якісної роботи моделі, для її навчання обрано набір даних COCO (англ. *Common Objects in Context*), який містить 80 класів різноманітних даних. Удосконалено метод розпізнавання об'єктів з використанням нейронних мереж шляхом створення та впровадження алгоритму фільтрації результатів роботи моделі відповідно до обраного режиму роботи.

Поєднання швидкої моделі SSD, великого набору даних COCO з 80 різними класами об'єктів та алгоритму фільтрації дало змогу створити ефективну систему виявлення об'єктів. Особливістю розробленої системи є вміння знаходити велику кількість об'єктів і одночасно ігнорувати нерелевантні об'єкти в заданому контексті роботи, що дає змогу використовувати єдину систему без жодних змін у різних середовищах, досягаючи однакової високої точності розпізнавання тільки релевантних об'єктів.

Для визначення ефективності роботи системи проведено експерименти на наборі даних ODSR-IHS, і досягнуто точності розпізнавання предметів 83 %. Таку точність досягнуто під час розпізнавання анованих даних, але за результатами експериментів встановлено, що система здатна розпізнавати більшу кількість об'єктів, ніж передбачав набір даних, що свідчить про перспективність використання системи у різних середовищах.

Приклад ситуації, коли рішення може бути корисним – це допоміжні роботи. Якщо вони працюють на вулиці, то можуть допомогти людині зорієнтуватися в місті. Якщо вони працюють у приміщенні, та ж система може перемішати людину будівлю і допомагати знаходити необхідні речі.

References

1. Chuyi Li, Lulu Li, Yifei Geng, Hongliang Jiang, Meng Cheng, Bo Zhang, Zaidan Ke, Xiaoming Xu, & Xiangxiang Chu. (2023). YOLOv6 v3.0: A Full-Scale Reloading. <https://doi.org/10.48550/arXiv.2301.05586>
2. Huan Wang, Can Qin, Yue Bai, Yulun Zhang, & Yun Fu. (2022). Recent Advances on Neural Network Pruning at Initialization. <https://doi.org/10.48550/arXiv.2103.06460>
3. Hussam, Qassim, David, Feinzimer, & Abhishek, Verma. (2017). Residual Squeeze VGG16. <https://doi.org/10.48550/arXiv.1705.03004>
4. Jingru Yi, Pengxiang Wu, & Dimitris N. (2019). Metaxas. ASSD: Attentive Single Shot Multibox Detector. <https://doi.org/10.48550/arXiv.1909.12456>
5. Pinhao, Song, Pengteng, Li, Linhui, Dai, Tao, Wang, & Zhan, Chen. (2022). Boosting R-CNN: Reweighting R-CNN Samples by RPNs Error for Underwater Object Detection. <https://doi.org/10.48550/arXiv.2206.13728>
6. Sebastian, Cygert, & Andrzej, Czyzowski. (2021). Robust Object Detection with Multi-input Multi-output Faster R-CNN. <https://doi.org/10.48550/arXiv.2111.13065>

7. Sik-Ho Tsang. Review: SSD – Single Shot Detector (Object Detection). (2018). URL: <https://towardsdatascience.com/review-ssd-single-shot-detector-object-detection-851a94607d11>
8. Teslyuk, V., Kazarian, A., Kryvinska, N., & Tsmots, I. (2021). Optimal artificial neural network type selection method for usage in smart house systems. *Sensors*, 21(1), 1–14, 47.
9. Tsung-Yi, Lin, Michael, Maire, Serge, Belongie, Lubomir, Bourdev, Ross, Girshick, James, Hays, Pietro, Perona, Deva, Ramanan, C. Lawrence, Zitnick, & Piotr Dollár. (2015). Microsoft CO-
CO: Common Objects in Context. <https://doi.org/10.48550/arXiv.1405.0312>
10. Yilun, Chen, Shu, Liu, Xiaoyong, Shen, & Jiaya, Jia. (2019). Fast Point R-CNN. <https://doi.org/10.48550/arXiv.1908.02990>
11. Zekun Luo, Zheng Fang, Sixiao Zheng, Yabiao Wang, & Yanwei Fu. (2021). NMS-Loss: Learning with Non-Maximum Suppression for Crowded Pedestrian Detection. <https://doi.org/10.1145/3460426.3463588>

B. P. Borkivskyi, V. M. Teslyuk

Lviv Polytechnic National University, Lviv, Ukraine

APPLICATION OF NEURAL NETWORK TOOLS FOR OBJECT RECOGNITION IN MOBILE SYSTEMS WITH OBSTACLE AVOIDANCE

An improved neural network-based method of object recognition for searching relevant objects in scope of given context is presented. The study includes analysis of existing approaches aiming to solve the problem of object recognition and discovers key features of several neural network architectures. The improved method is a combination of neural network-based approach for pattern recognition and algorithms aiming to remove redundant objects, namely a duplicated objects removal algorithm and a context-aware objects filtering algorithm. The base for the object detection neural network is Single Shot Detector model architecture, trained on COCO dataset with extensive data. The objects filtering algorithm anticipates creating configuration to determine sets of relevant object classes for different contexts. If necessary, this configuration can be modified and extended with new types of contexts without any additional changes in program code. The paper provides results of several conducted experiments aiming to estimate the performance of the developed system in new environments by finding accuracy using ODSR-IHS dataset, that contains images of sweeping robots working in various situations. The experiments determined that the accuracy of detecting 5 classes of objects is 83 %. After the analysis of experiments the ability to find more relevant objects than was intended by dataset authors was discovered, including a person, it testifies that applying such a system in different environments is promising. Using diverse dataset for training that consists of 80 objects classes and selecting only significant objects makes the developed system an effective tool for working in different situations. The benefits of this study are that the developed method enables solving the problem of reusing the same object detection system in varying situations, simplifying system usage in mobile systems that can work in multiple environments.

Keywords: image processing; convolutional network; object detection; contextual deletion.