

DOI: <https://doi.org/10.36910/6775-2524-0560-2024-54-18>

УДК 004.934

Недашківський Сергій Миколайович, заступник голови з питань цифрового розвитку, цифрових трансформацій і цифровізації

<https://orcid.org/0000-0002-8222-6905>

Державне агентство водних ресурсів України, м. Київ, Україна

МЕТОДОЛОГІЯ ТА ПРИНЦИПИ ВИЯВЛЕННЯ ОБ'ЄКТІВ ЗАДОПОМОГОЮ ДЕФОРМОВАНИХ ЗГОРТКОВИХ МЕРЕЖ

Недашківський С.М. Методологія та принципи виявлення об'єктів за допомогою деформованих згорткових мереж. У динамічній сфері комп'ютерного зору впровадження штучного інтелекту започаткувало трансформаційну еру, пропонуючи неперевершену точність і ефективність ідентифікації об'єктів на зображеннях і відео. Це дослідження заглиблюється в сферу виявлення об'єктів, керованих штучним інтелектом, з особливим акцентом на ключовій ролі метаданих у покращенні розуміння та корисності при роботі з розпізнаними об'єктами. Співпраця між штучним інтелектом і метаданими не тільки підвищує точність виявлення об'єктів, але й відкриває інноваційні шляхи для вилучення й аналізу інформації. Метадані охоплюють ключові деталі, такі як клас об'єкта, місце виявлення, час виникнення та взаємозв'язки між об'єктами, надаючи інформацію для наступних програм, таких як автономні транспортні засоби, спостереження та доповнена реальність. Дослідницька стаття служить демонстрацією інтеграції вилучення метаданих і керування ними з системами виявлення об'єктів на основі штучного інтелекту, що підвищує точність ідентифікації та відстеження об'єктів. Це дослідження висвітлює взаємодію між штучним інтелектом і комп'ютерним зором, формуючи ландшафт, де точність і адаптивність заново визначають межі можливостей виявлення об'єктів. Співпраця між штучним інтелектом і метаданими стає ключовим рушієм у підвищенні загальної ефективності та результативності систем розпізнавання об'єктів, пропонуючи зазірнути в майбутнє інтелектуального аналізу зображень і відео.

Ключові слова: штучний інтелект, згорткові мережі, наука про дані, аналіз даних, обробка даних, презентації даних.

Nedashkivskiy S. Methodology And Principles of Object Detection in Deformable Convolutional Networks. In the dynamic field of computer imaging, the introduction of artificial intelligence began a transformation, contributing to the unsurpassed accuracy and efficiency of object detection in images and videos. This study delves into the field of AI-driven objects detection, with a particular focus on the key role of metadata in improving the understanding and utility of dealing with detected objects. The collaboration between artificial intelligence and metadata not only completes the accuracy of object detection, but also opens up innovative ways to extract and analyze information. Metadata captures key details such as object class, location of detection, time of occurrence, and relationships between objects, providing information for downstream applications such as autonomous vehicles, observables, and augmented reality. The research paper serves to demonstrate the integration of metadata extraction and management with artificial intelligence-based object detection systems, resulting in accurate object identification and tracking. This research illuminates the interaction between artificial intelligence and computer vision, shaping a landscape where accuracy and adaptability redefine the limits of object detection capabilities. The collaboration between artificial intelligence and metadata is becoming a key driver in increasing the overall efficiency and effectiveness of the object recognition system, offering a glimpse into further intelligent analysis of images and videos.

Key words: artificial intelligence, convolutional networks, data science, data analysis, data Processing, data Presentations.

Вступ та постановка проблеми. Виявлення об'єктів являє собою фундаментальну концепцію комп'ютерного зору, яка відіграє ключову роль у різних програмах різних галузей. Воно передбачає ідентифікацію та локалізацію об'єктів у зображенні чи відеопотоці. Основною метою є не лише розпізнавання присутності об'єктів, але й визначення їх конкретного розташування у візуальних даних.

Однією із відомих галузей, де виявлення об'єктів може бути застосованим є автономні транспортні засоби, це має вирішальне значення для забезпечення безпеки та ефективності навігації. Виявляючи та відстежуючи такі об'єкти, як пішоходи, транспортні засоби та перешкоди, автономні транспортні засоби можуть приймати обґрунтовані рішення, щоб уникнути зіткнень і пересуватися в складних умовах. Виявлення об'єктів також широко використовується в системах відеоспостереження, що дозволяє контролювати громадські місця, аеропорти та критичну інфраструктуру шляхом ідентифікації та відстеження підозрілих об'єктів.

Методи виявлення об'єктів зазвичай належать до підходів на основі нейронних мереж або до підходів, не заснованих на використанні нейронних мереж, для яких стає необхідним спочатку визначити ознаки об'єкту, а потім використовувати такий метод, як машина опорних векторів, щоб виконати його класифікацію. З іншого боку, нейронні методи здатні здійснювати наскрізне виявлення об'єктів без спеціального визначення функцій і зазвичай базуються на згорткових нейронних мережах.

Аналіз останніх досліджень і публікацій. Наукові діячі сьогодення внесли значний вклад у розробку методологій по виявленню об'єктів у деформованих згорткових мережах. Був проведений ряд досліджень для вирішення даної проблеми.

У [1] було запропоновано мережу регіональних пропозицій (RPN) для створення запропонованих областей і використано іншу мережу для використання цих запропонованих областей для виявлення об'єктів. Бокси, які, швидше за все, містять об'єкти навчальних зображень, були запропоновані за допомогою ранжирувальних блоків, які називаються прив'язками. Faster R-CNN показала такий результат: 73,2% mAP на VOC2007, ресурсі, що є набором даних для розпізнавання зображень. Однак, Faster R-CNN не повною мірою реалізувала виявлення об'єктів у режимі реального часу, оскільки функції обчислення вимагали багато часу. У роботі [2] продемонстрована техніка розпізнавання об'єктів за допомогою нейронної мережі YOLO (You Only Look Once), з використанням регресії – метрики для вимірювання того, наскільки добре передбачені обмежувальні рамки фіксують об'єкти. Це робиться шляхом обчислення середньоквадратичної похибки (MSE) між координатами передбачуваних та існуючих обмежувальних рамок. Вхідні зображення поділяються на кілька клітинок, і ймовірність класифікації та блоків прогнозується з кожної клітинки. Прогнози були зроблені на основі глобальної репрезентації із зображення та показали відмінний результат. Використовуючи концепцію регресії YOLO, а також спеціальний метод прив'язки Faster R-CNN, розроблено спосіб виявлення об'єкта всередині зображення за допомогою єдиної глибокої нейронної мережі [3], що створює прогнози ознак з карт різного масштабу. У кожному місці на карті об'єктів передбачений простір обмежувальних рамок було розбито на групу рамок за замовчуванням із різними пропорціями та масштабами.

Крім того, варто зазначити праці наступних науковців: Лоханде Рідам, Накхале Яш, Петкар Неха, Шейх Танзіла, Джанве Др., Лічоде Проф [4], Дуан Кайвен, Бай Сонг, Се Лінгсі, Ці Хунган, Тянь Ці [5], Абдул Бассім, Амін Самман, Мнаті Духа [6], Ядав Атул, Чатурведі Пратюш, Рані Шаллу [7], Кусума Патрісія, Соевіто Бенфано [8], Голлапаллі Гіта, Тотакура Ясвант, Касім Шалом, Доппалапуді Каліян [9], Сонг Ігуо, Лю Женю, Тан Руїнінг, Дуань Гуйфан, Тан Цзяньронг [10], Ден Зіхан, Лі Енг [11], Бхат Пуджа, Маквана Дхрув, Джоші Акшит, Раол Крішнарадж [12], Романо Бруно, Фаготто Ерік [13], Джалал Ахмад [14], Александріно Леандро, Оляей Хаді, Альбукерке Андре, Георгієва Петія, Драммонд Мігель [15], Мумтаз Фаріха, Ашраф Х., Джанджі Нур [16], Цю Хецянь, Лі Хунлянь, Ву Цінбо, Ши Хенкань, ВанЛаньсяо, Мен Фанман, Сюй, Ліньфен [17] та інші.

Проте, беручи до уваги вище зазначену наукову документацію, питання, пов'язане з методологіями по виявленню об'єктів у деформованих згорткових мережах, все ще залишається недостатньо дослідженим та потребує подальшого опрацювання.

Постановка завдання. Метою роботи є дослідження принципів виявлення об'єктів у деформованих згорткових мережах.

Викладення основного матеріалу дослідження. Розуміння важливості застосування такої концепції як виявлення об'єктів та користі поєднання її з такими винаходами штучного інтелекту, як нейронні мережі, а найчастіше саме зі згортковими нейронними мережами вимагає розуміння особливостей згорткових нейронних мереж.

Директор зі стратегії планування ресурсів підприємства в Cabernet Networks Group, Техас, США Чалла Нараяна [18] сформулював б обґрунтувань важливості виявлення об'єктів з використанням можливостей нейронних мереж:

1. Виявлення об'єктів допомагає інтерпретувати вміст візуальної сцени, що дозволяє комп'ютерам ідентифікувати та знаходити об'єкти, надаючи контекст для більш складної обробки. Виявлення об'єктів має вирішальне значення в промислових умовах для моніторингу руху продукту, автоматизації процесів і забезпечення контролю якості. Крім того, воно сприяє більшій незалежності людей з вадами зору, розпізнаючи та характеризуючи об'єкти та їх розташування.

2. Уникнення перешкод і навігація: виявлення об'єктів має важливе значення для виявлення та уникнення перешкод, орієнтування в оточенні та прийняття рішень на основі наявності об'єктів. Це особливо важливо для автономних транспортних засобів, дронів і робототехніки. Для систем безпеки спостереження виявлення об'єктів дозволяє ідентифікувати та контролювати об'єкти та людей у реальному часі, покращуючи виявлення загроз та реагування на них. У роздрібній торгівлі воно автоматизує процедури оформлення замовлення, відстежує продукти на полицях і керує запасами для підвищення продуктивності та задоволення споживачів.

3. Медична візуалізація та діагностика: у медичній візуалізації виявлення об'єктів

допомагає знаходити та ідентифікувати конкретні структури, сприяючи діагностиці та лікуванню. Це забезпечує точність ідентифікації об'єктів на зображеннях, сприяючи кращим результатам лікування.

4. Інтерактивні можливості та розваги: виявлення об'єктів слідує за рухами та жестами користувача в інтерактивних програмах та іграх, створюючи ефект занурення. Воно автоматично додає теги та класифікує фільми та зображення, спрощуючи пошук і організацію великих колекцій медіа.

5. Автоматизація та прискорення процесів: виявлення об'єктів автоматизує завдання в різних галузях, зменшуючи потребу в ручному втручанні та прискорюючи процедури. Програми доповненої реальності розпізнають об'єкти реального світу за допомогою виявлення об'єктів, покращуючи взаємодію з користувачем шляхом накладання цифрових даних або графіки.

6. Моніторинг навколишнього середовища та дослідження дикої природи: дослідники та природоохоронці отримують важливі дані для відстеження змін навколишнього середовища, моніторингу дикої природи та вивчення поведінки тварин. Моніторинг відіграє ключову роль у розумінні та збереженні екосистем.

Оскільки у цифрову епоху, кількість візуальних даних, від відеоспостереження до медичних зображень і безпілотних автомобілів, є занадто великою, вона вимагає точного виявлення об'єктів. Штучний інтелект, особливо з такими моделями глибокого навчання, як згорткові нейронні мережі, став надійним рішенням, яке з більшою точністю долає такі проблеми, як оклюзія та варіації масштабу.

Виявлення об'єктів зазвичай включає такі фундаментальні елементи:

– ідентифікація положення та розмірів об'єктів на зображенні. Це часто можна візуалізувати, окреслюючи об'єкти обмежувальними рамками, вказуючи їх просторові координати (x, y) і розміри (ширина та висота);

– присвоєння мітки або категорії кожному виявленому об'єкту із зазначенням його типу. Наприклад, у сцені з різними об'єктами функція виявлення об'єктів може визначити, чи наявні автомобілі, пішоходи, тварини чи інші об'єкти, і призначити відповідні мітки;

– моделі глибокого навчання, такі як згорткові нейронні мережі, часто використовуються для виявлення об'єктів. Ці моделі призначені для одночасного виконання задач локалізації та класифікації. Завдяки навчанню на великих наборах даних із позначеними зображеннями вони можуть розпізнавати об'єкти та точно визначати їх положення;

– великі набори даних із анотованими зображеннями. Вони зазвичай потрібні для ефективного навчання нейронних моделей виявлення об'єктів. Ці набори даних зазвичай включають зображення, де цікаві об'єкти позначені обмежувальними рамками та пов'язаними мітками класів. Під час навчання моделі вчать на цих прикладах точно передбачати положення та класи об'єктів на нових, небачених зображеннях.

Згорткова нейронна мережа являє собою регуляризований тип прямої нейронної мережі, яка сама вивчає розробку функцій за допомогою фільтрів. Зникаючі градієнти, які спостерігаються під час зворотного поширення в нейронних мережах, є можливість мінімізувати за допомогою регуляризованих вагових коефіцієнтів, однак кількість з'єднань залишиться без змін. Наприклад, для кожного нейрона в повністю пов'язаному шарі буде потрібно 10 000 вагових коефіцієнтів для обробки зображення розміром 100×100 пікселів. Однак, застосовуючи ядра каскадної згортки (або крос-кореляції), для обробки клітин розміром 5×5 потрібно лише 25 нейронів. Функції вищого рівня формуються на основі клітин максимального набору порівняно з функціями нижчого рівня.

Перевагою використання згорткової нейронної мережі є її можливість мати десятки або сотні шарів, кожен з яких навчається виявляти різні характеристики зображення. Фільтри застосовуються до кожного тренувального зображення з різною роздільною здатністю, а вихідні дані кожного згорнутого зображення використовуються як вхідні дані для наступного шару. Фільтри можуть починатися з дуже простих функцій, таких як яскравість та границі, і збільшуватися в складності до функцій, які однозначно визначають об'єкт.

Згорткова нейронна мережа стає здатною до деформування, тобто деформованою, коли вона включає в себе згорткові шари, що деформуються. Згорткові шари, що деформуються, є типом згорткового шару, який дозволяє динамічно коригувати місця просторової вибірки в згортковому ядрі під час прямого проходу на основі вивчених зсувів. Ця гнучкість дає змогу мережі краще адаптуватися до деформацій об'єктів і варіацій у вхідних даних.

У стандартній або недеформованій мережі згорткові ядра мають фіксовані регулярні сітки для вибірки вхідних характеристик. Незважаючи на неадекватність, у багатьох випадках, вони можуть мати проблеми при роботі з об'єктами, які зазнають деформацій.

Деформовані згорткові нейронні мережі вирішують це обмеження шляхом введення деформованих згорткових шарів. Ці рівні дозволяють мережі вивчати просторові перетворення для кожної позиції в згортковому ядрі, дозволяючи мережі зосереджуватися на відповідних областях вхідних даних і краще обробляти деформовані структури в зображеннях.

Одним з ефективних типів згорткових нейронних мереж у використанні для виявлення об'єктів є R-CNN, яка являє собою мережу, що функціонує на локальному рівні та яку можна застосовувати для виявлення цілей. Її робота полягає в тому, щоб створити приблизно 2000 рекомендацій щодо певної області із зображення, потім стиснути його в попередньо навчену модель згорткової нейронної мережі, а потім застосувати метод опорних векторів для сортування функцій. Нарешті, границя запропонованої області змінюється за допомогою лінійної регресії, а область накладання усувається. Перевага R-CNN полягає в тому, що вона використовує переваги глибокого навчання, що підвищує точність і швидкість виявлення об'єктів. Процес виявлення об'єкту за регіональним типом з використанням нейронної мережі R-CNN наведено на рисунку 1.

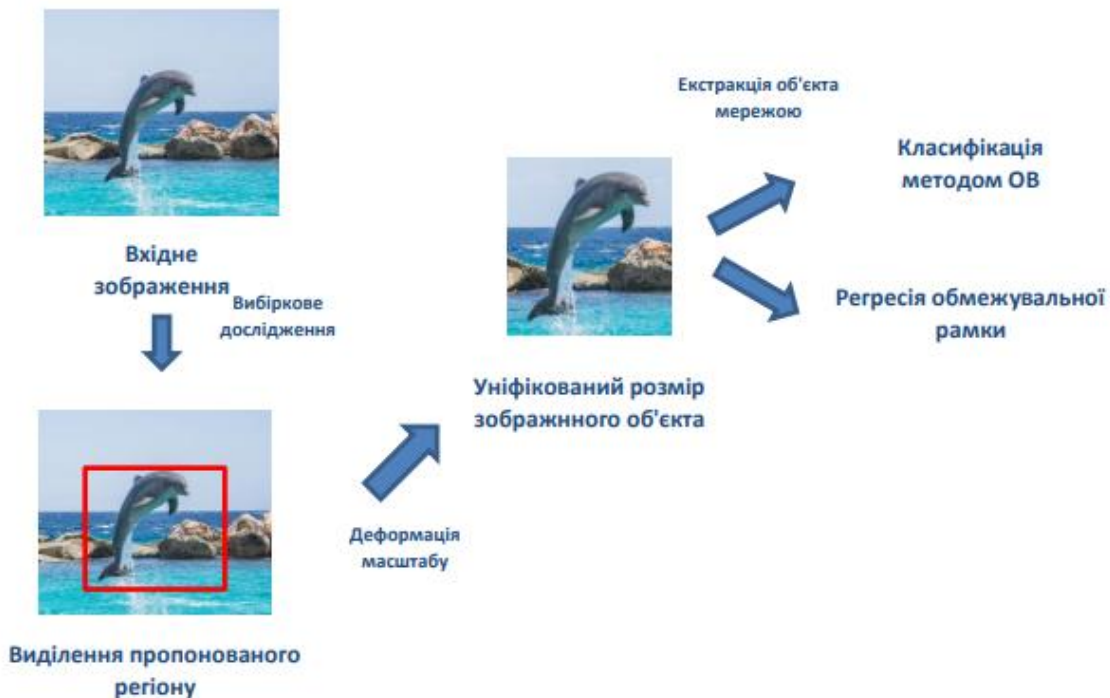


Рис. 1 – Процес виявлення об'єкту за регіональним типом з використанням нейронної мережі R-CNN

Однак, R-CNN також має деякі недоліки:

- низька ефективність виявлення. Через те, що кожен кадр-кандидат виділяється окремо, це призводить до великої кількості повторюваних обчислень, тому для виявлення потрібно багато часу.

- порушена точність виявлення. Операція масштабування поля-кандидата призводить до зміни форми та пропорції вихідного зображення, що впливає на розпізнавання ознак.

- складний навчальний процес. Генерація блоків-кандидатів, виділення ознак, класифікація та регресія виконуються окремо, що вимагає багаторазового ручного втручання, а зберігання проміжних даних також займає додатковий простір.

Використовується схема представлення орієнтованих об'єктів, яка застосовує зовнішні прямокутні рамки зі зміщенням для кодування та декодування орієнтованих обмежувальних рамок. Як показано на рисунку 2(a), неорієнтована обмежувальна рамка кодується як $(x, y, w, h, \Delta\alpha, \Delta\beta)$. Де (x, y) – центроїдна координата, ширина та висота зовнішнього горизонтального прямокутника рамки; $\Delta\alpha$ – горизонтальна координата найвищої вершини v_2 , зміщена від x ; $\Delta\beta$ – вертикальна

координата крайньої правої вершини v_3 , зміщена від u . Таким чином, можна кодувати і декодувати орієнтовані об'єкти, що дає змогу отримати теоретично орієнтовані обмежувальні прямокутники шляхом прямої регресії зміщення.

Приймаються п'ять рівнів функцій $\{P_2, P_3, P_4, P_5, P_6\}$ у якості вхідних даних та встановлюється згортковий шар 3×3 і два однотипні згорткові шари 1×1 до кожного рівня функцій. Необхідно призначити три горизонтальні прив'язки з трьома співвідношеннями сторін $\{1:2, 1:1, 2:1\}$ для кожного просторового розташування на всіх рівнях об'єктів. Основа має площу пікселів $\{32^2, 64^2, 128^2, 256^2, 512^2\}$ на $\{P_2, P_3, P_4, P_5, P_6\}$ відповідно. Кожна основа позначається 4-вимірним вектором $a = (a_x, a_y, a_w, a_h)$, де (a_x, a_y) є координатою центру основи, a_w і a_h представляють ширину та висоту основи. Один із двох однорідних згорткових шарів 1×1 є гілкою регресії: виведення зсуву $\delta = (\delta_x, \delta_y, \delta_w, \delta_h, \delta_\alpha, \delta_\beta)$ пропозицій відносно прив'язок. У кожному розташуванні карти об'єктів генеруються A пропозиції (A – це кількість прив'язок у кожному розташуванні, і в цій роботі воно дорівнює 3), таким чином, гілка регресії має $6A$ виходів. Шляхом декодування результатів регресії можна отримати орієнтовану пропозицію. Процес декодування описується так:

$$\begin{aligned} \Delta\alpha &= \delta_\alpha \cdot w, \Delta\beta = \delta_\beta \cdot h \\ w &= a_w \cdot e^{\delta_w}, h = a_h \cdot e^{\delta_h} \\ x &= \delta_x \cdot a_w + a_x, y = \delta_y \cdot a_h + a_y \end{aligned} \quad (1)$$

де (x, y) – координата центру прогнозованої пропозиції;

w і h – ширина та висота зовнішнього прямокутника прогнозованої орієнтованої пропозиції;

$\Delta\alpha$ і $\Delta\beta$ – зміщення відносно середин верхньої та правої сторін зовнішнього прямокутника.

Нарешті, створюються орієнтовані пропозиції відповідно $(x, y, w, h, \Delta\alpha, \Delta\beta)$.

На рисунку 2 зображена схему представлення орієнтованих об'єктів, яка називається представленням зміщення середньої точки. Чорні крапки є серединами кожної сторони горизонтальної рамки, яка є зовнішнім прямокутником орієнтованої обмежувальної рамки O . Помаранчеві крапки означають вершини орієнтованої обмежувальної рамки O . Зокрема, використовується орієнтована обмежувальна рамка O з шістьма параметрами $O = (x, y, w, h, \Delta\alpha, \Delta\beta)$, щоб представити об'єкт, обчислений за рівнянням (1).

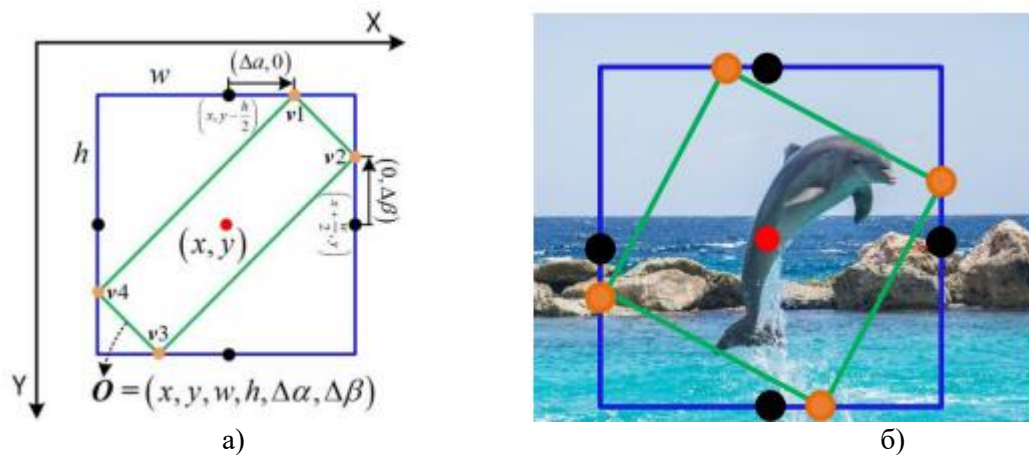


Рис. 2 – Ілюстрація представлення зміщення середньої точки: а) принципова діаграма зображення зміщення середньої точки; б) приклад представлення зміщення середньої точки

За допомогою шести параметрів можна отримати набір координат $v = (v_1, v_2, v_3, v_4)$ чотирьох вершин для кожної пропозиції. Тут $\Delta\alpha$ – зсув v_1 відносно середини $(x, y - h/2)$ верхньої сторони горизонтального прямокутника. Відповідно до симетрії, $\Delta\alpha$ являє собою зсув v_3 відносно нижньої середньої точки $(x, y + h/2)$. $\Delta\beta$ означає зсув v_2 відносно правої середньої точки $(x + w/2, y)$, а $-\Delta\beta$ є зміщенням v_4 відносно лівої середньої точки $(x - w/2, y)$. Таким чином, координати чотирьох вершин можна виразити так:

$$v_1 = \left(x, y - \frac{h}{2}\right) + (\Delta\alpha, 0)$$

$$\begin{aligned}v_2 &= \left(x + \frac{w}{2}, y\right) + (0, \Delta\beta) \\v_3 &= \left(x, y + \frac{h}{2}\right) + (-\Delta\alpha, 0) \\v_4 &= \left(x - \frac{w}{2}, y\right) + (0, -\Delta\beta)\end{aligned}\quad (2)$$

За допомогою способу представлення необхідно реалізувати регресію для кожної орієнтованої пропозиції шляхом прогнозування параметрів (x, y, w, h) для її зовнішнього прямокутника та визначення параметрів $(\Delta\alpha, \Delta\beta)$ для її середнього зміщення.

У процесі навчання нейронна мережа вчиться мінімізувати розбіжність між прогнозованими параметрами та базовими значеннями істинності для заданого набору навчальних зразків. Функція втрат призначена для мінімізації відхилень як у положенні (x, y) , так і в орієнтації $(\Delta\alpha, \Delta\beta)$ орієнтованих обмежувальних рамок. Це дає змогу мережі ефективно вловлювати орієнтири об'єктів та їх середніх точок.

У міру навчання згортована нейронна мережа вдосконалює свою здатність передбачати представлення зміщення середньої точки для орієнтованих об'єктів у заданому наборі даних. Вивчене представлення дозволяє мережі добре узагальнювати невидимі приклади, покращуючи її продуктивність у задачах виявлення об'єктів, особливо коли йдеться про об'єкти неправильної форми або орієнтації.

Переваги використання запропонованого представлення зміщення середньої точки при виявленні об'єктів включають

- підвищену надійність: модель більш стійка до змін орієнтації об'єкта, що робить її придатною для сценаріїв, коли об'єкти можуть бути нахилені або повернуті;
- покращену локалізацію: явне представлення середніх точок і зсувів допомагає точно локалізувати об'єкт, сприяючи кращому передбаченню обмежувальної рамки;
- покращене узагальнення: навчене представлення дає змогу мережі добре узагальнювати різні форми та орієнтації об'єктів, що призводить до покращення продуктивності реальних наборів даних;
- зменшену неоднозначність: представлення зсуву середньої точки зменшує неоднозначність у локалізації об'єкта, забезпечуючи більш чітке розуміння меж об'єкта.

Висновки. У даному документі був зроблений огляд методів виявлення об'єктів на відео та зображеннях використовуючи згортовані нейронні мережі, зокрема локальний метод виявлення на основі нейронної мережі R-CNN. Загалом, існує процедура скринінгу для виявлення цілі на основі запропонованої області. Тому, з розвитком технологій точність стає все більшою, а процес з роками стає все простішим. Але загалом швидкість виявлення є досить повільною і не може задовольнити вимоги програми. Проте, з еволюцією та розвитком системи, за останні роки, її точність також постійно вдосконалюється і вона має переваги в практичному застосуванні. В останні роки завдяки науковим досягненням у галузі комп'ютерного зору широко використовуються глибокі нейронні мережі, алгоритм виявлення об'єктів, заснований на згортованих нейронних мережах, перевершив традиційний алгоритм. Однак залишаються труднощі, пов'язані, наприклад з питанням того, як навчити мережу виявляти об'єкти за умови обмеженого набору даних, чи як досягти достатніх вимог до точності в реальній області застосування.

На завершення, запропоноване представлення зсуву середньої точки, інтегроване в нейронну мережу для виявлення об'єктів, пропонує більш ефективний і надійний підхід до обробки орієнтованих об'єктів. Підвищена точність локалізації та узагальнення для різноманітних сценаріїв робить цю схему представлення багатобічним досягненням у сфері комп'ютерного зору та виявлення об'єктів.

Перспективами подальших досліджень є розробка більш ефективної методології по виявленню об'єктів за допомогою деформованих згортованих мереж.

Список бібліографічного опису

1. Ren S., He K., Girshick R., Sun J., Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39. 2016. № 6. P. 1137 – 1149.
2. Redmon J., Divvala S., Girshick R., Farhadi A. You Only Look Once: Unified, real-time object detection. *Computer Vision and Pattern Recognition (CVPR), International Conference on IEEE*. 2016. P. 779 – 788.
3. Liu W., Anguelov D., Erhan D., Szegedy C., Reed S., Fu C., Alexander C., Berg SSD: Single Shot MultiBox

Detector. *CoRR*. 2015 DOI:abs/1512.02325

4. Lokhande R., Nakhale Y., Petkar N., Sheikh T., Janwe Dr., Lichode P. Object Detection. *International Journal of Advanced Research in Science, Communication and Technology*. 2023. P. 316-320.

DOI:10.48175/IJARSCT-14042.

5. Duan K., Bai S., Xie L., Qi H., Tian Q. CenterNet++ for Object Detection. *IEEE transactions on pattern analysis and machine intelligence*. 2023. DOI:10.1109/TPAMI.2023.3342120.

6. Abdul B., Ameen S., Mnaty D. Object Detection Algorithm Using Deep Learning. 2024. № 9.

7. Yadav A., Chaturvedi P., Rani S. Object Detection and Tracking using YOLOv8 and DeepSORT. 2024. DOI:10.56155/978-81-955020-7-3-7.

References

8. Kusuma P., Soewito B. Multi-Object Detection Using YOLOv7 Object Detection Algorithm on Mobile Device. *Journal of Applied Engineering and Technological Science (JAETS)*. 2023. № 5. P. 305-320. DOI:10.37385/jaets.v5i1.3207.

9. Gollapalli G., Thotakura Y., Kasim S., Doppalapudi K. Multi Object Detection. *International Journal for Research in Applied Science and Engineering Technology*. 2023. № 11. P. 1222-1229. DOI:10.22214/ijraset.2023.55817.

10. Song Y., Liu Z., Tang R., Duan G., Tan J. Cross-domain object detection by local to global object-aware feature alignment. *Neural Computing and Applications*. 2023. P. 1-14. DOI:10.1007/s00521-023-09248-8.

11. Deng Z., Li A. Object Detection Algorithms Based on Convolutional Neural Networks. *Highlights in Science, Engineering and Technology*. 2024. № 81. P. 243-251. DOI:10.54097/vyfg4e34.

12. Bhatt P., Makwana D., Joshi A., Raol K. Object Extraction and Detection Using U2-Net and YOLOv7. 2023. DOI:10.21203/rs.3.rs-3808815/v1.

13. Romão B., Fagotto E. Convolutional Neural Networks for Object Detection. 2024. DOI:10.4271/2023-36-0097.

14. Jalal A. Integrating Semantic Segmentation and Object detection for Multi-object labeling in Aerial Images. 2024.

15. Alexandrino L., Olyaei Hadi., Albuquerque André., Georgieva Petia., Drummond Miguel. 3D Object Detection for Self-Driving Vehicles Enhanced by Object Velocity. *IEEE Access*. 2024. DOI:10.1109/ACCESS.2024.3353051.

16. Mumtaz F., Ashraf H., Jhanjhi N. Moving Object Detection Using Deep Learning Algorithms. 2023. DOI:10.20944/preprints202312.1589.v1.

17. Qiu H., Li H., Wu Q., Shi H., Wang L., Meng F., Xu Linfeng. Learning Offset Probability Distribution for Accurate Object Detection. *ACM Transactions on Multimedia Computing, Communications, and Applications*. 2023. № 20. DOI:10.1145/3637214.

18. Challa N. Artificial Intelligence for Object Detection and its Metadata. 2023. № 2. P. 121-133. DOI:10.17605/OSF.IO/FG3SQ.