Вісник Національного технічного університету "ХПІ", 2024, № 1–2 (11–12)
ISSN 2079-0031 (Print) ISSN 2411-0558 (Online)

УДК 004.93 **DOI: 10.20998/2411-0558.2024.01.05**

В. В. ЛЮБЧЕНКО, д-р техн. наук, проф., Національний університет "Одеська політехніка", м. Одеса**,**
Д. В. МОРОЗ, асп., Національний університет "Одеська політехніка", м. Одеса

МОДЕЛІ ТА МЕТОДИ ВИЯВЛЕННЯ ОБ'ЄКТІВ В ЦИФРОВІЙ ОБРОБЦІ ЗОБРАЖЕНЬ

Виявлення об'єктів в цифрових зображеннях може стикатися з такими проблемами, що виникають в процесі реєстрації зображень, як наявність шуму, низька якість/роздільна здатність зображень, неоднорідність освітлення, перекриття об'єктів та інші. Ці проблеми ускладнюють процес виявлення об'єктів та призводять до помилок у роботі алгоритмів обробки зображень. Для розв'язання цих проблем досліджуються переваги, недоліки та технічні особливості моделей і методів виявлення об'єктів у цифрових зображеннях для їх обґрунтованого вибору в практичних застосуваннях. Іл.: 2. Бібліогр.: 11 назв.

Ключові слова: комп'ютерний зір; розпізнавання об'єктів; обробка зображень; технології штучного інтелекту.

Вступ і постановка проблеми. Виявлення об'єктів у цифровій обробці зображень вирішує фундаментальну проблему, яка полягає в тому, щоб дати можливість машинам розпізнавати і знаходити об'єкти у візуальних даних. Ця проблема виникає через необхідність інтерпретації та взаємодії з візуальним світом і охоплює велику кількість застосувань у різних галузях.

Основний виклик полягає в тому, щоб отримати з зображення чи відео релевантну інформацію, необхідну для прийняття рішень, автоматизації процесів діяльності людини та розуміння навколишнього середовища. Комп'ютерний зір, як галузь знань, досяг великих успіхів, але йому все ще важко зрівнятися з точністю людського сприйняття. Людина може легко виявити та ідентифікувати об'єкти, присутні на зображеннях, зорова система людини працює швидко і точно і може виконувати складні завдання, наприклад, миттєве і точне виявлення кількох об'єктів або виявлення перешкод на шляху, навіть не замислюючись про це.

© В. В. Любченко, Д. В. Мороз, 2024

Завдяки доступності великих обсягів даних, швидким графічним процесорам і кращим алгоритмам, комп'ютери можна легко навчити виявляти і класифікувати кілька об'єктів на зображенні з високою точністю. Більше того, їх можна використовувати у вирішенні проблем, які не може вирішити людина через обмежені можливості її психофізіологічної системи зору. Іншими перевагами застосування комп'ютерів є щвидка ідентифікація та локалізація виявлення об'єктів, що дає змогу підраховувати об'єкти в сцені, точно знаходити і відстежувати їх, а також точно маркувати їх. Тому можна зробити висновок, що бачення є символом інтелекту, а комп'ютер є можливістю втілення логіки, яка притаманна людському баченню. Коли ми дивимося на фотографію, нам потрібно менше кількох мілісекунд, щоб зрозуміти, хто на ній зображений. У людському мозку формується потік інформації, що дає змогу миттєво розпізнавати об'єкти.

Проблема виявлення об'єктів є різнобічною. Потрібна стійкість до змін масштабу об'єкта, його орієнтації, варіації точки огляду тощо. Один і той же предмет може мати абсолютно різні форми і розміри. Комп'ютерний зір потребує багато досліджень, щоб прочитати об'єкт і зрозуміти, що він означає. Об'єкт, який розглядається з різних кутів, може виглядати абсолютно по-різному [1].

Тому прикладні дослідження в напрямку виявлення об'єктів, ідентифікації зображень/текстур залишається актуальною проблемою, а досягнення в галузі штучного інтелекту і їх застосування в моделюванні системи бачення людини лише підтверджує і мотивує до пошуку оптимальних розв'язків цієї проблеми.

Мета дослідження. Метою дослідження є узагальнення переваг, недоліків та технічних особливостей моделей і методів виявлення об'єктів у цифрових зображеннях для їх обґрунтованого вибору в практичних застосуваннях.

Аналіз сучасних досліджень. Методи виявлення об'єктів комп'ютерного зору охоплюють як традиційні підходи, так і підходи на основі штучного інтелекту (ШІ). Класичні (традиційні) методи, які передують методам ШІ, грунтуються на:

1) порівнянні шаблонів, коли порівнюються області зображення з певним шаблоном або зразком;

2) використанні кольору та текстури, які застосовують аналіз кольору та текстури об'єктів в зображенні;

3) використанні контурів і застосуванні алгоритмів виявлення країв або дескрипторів форми для визначення об'єктів на зображенні.

Методи на основі ШІ широко використовуються в комп'ютерному зорі вже багато років і мають переваги з точки зору обчислювальної ефективності та інтерпретації.

Метод виявлення об'єктів – це метод комп'ютерного зору, який використовується для ідентифікації та позиціонування об'єктів на зображенні або відео. Зокрема, виявлення об'єктів може визначити положення об'єктів (або те, як вони рухаються) у певній сцені, завдяки обмежувальній рамці навколо виявлених об'єктів.

Задача виявлення об'єктів складається з 3-х завдань:

1) визначити, чи є на зображенні об'єкт;
2) виявити, де знаходиться цей об'єкт в зображенні;
3) зрозуміти, що це за об'єкт.

Що відрізняє виявлення об'єктів від інших завдань, так це унікальна здатність знаходити об'єкти на зображеннях і відео. Це дає змогу підраховувати та відстежувати об'єкти.

З метою відтворення візуальної системи людини в комп'ютері, були розроблені відповідні моделі і методи. Однією з перших фундаментальних робіт цього напрямку є дослідження Девіда Мара [2]. Такі поширені методи комп'ютерного зору, як розпізнавання об'єктів, розпізнавання образів і оброблення зображень, дають змогу передбачати і вирішувати критичні завдання виявлення об'єктів без особливого втручання людини.

Розпізнавання об'єктів – це техніка комп'ютерного зору, що використовується для ідентифікації, визначення місця розташування і класифікації об'єктів на цифрових зображеннях і, відвовідно, в реальних сценах. Це прикладний підхід до ШІ, який перепрофілює комп'ютери на детектори об'єктів, даючи їм змогу сканувати зображення і відео реального світу. Вони розуміють особливості об'єктів та інтерпретують їхнє призначення так само, як це робить людина.

Успіх алгоритмів розпізнавання об’єктів залежить від якості даних, необхідних для навчання. Що більше даних, то швидше модель зможе класифікувати об’єкти на основі відомих характеристик. Імовірність точної ідентифікації об’єкта залежить від атрибутів зображення. У ШІ система розраховує бал довіри і передбачає мітки та класи об’єктів. Однак алгоритмічні розрахунки під час розпізнавання об’єктів дещо складні та потребують глибокого розуміння для отримання результатів.

Розпізнавання об’єктів поділяють на чотири етапи: розпізнавання зображень, локалізацію об’єктів, виявлення об’єктів і сегментацію зображень. Вони застосовують контрольовані методи машинного навчання, такі як метод опорних векторів (Support Vector Machine), adaboosting, boosting і дерева рішень. Розглянемо більш детально етапи розпізнавання об’єктів.

1. Розпізнавання зображень – це важливий етап усього процесу, який використовується для визначення категорії аналізованого зображення. Наприклад, якщо на зображенні присутній об’єкт заданого типу, система розпізнавання об’єктів аналізує основні характеристики об’єкта (наприклад, форму, відносні розміри, структурні елементи), порівнює їх із тисячами навчених зображень і видає на виході клас об’єкта.

2. Локалізація об’єктів визначає точне положення кожного об’єкта на зображенні. При отриманні вхідного зображення створюється обмежувальна рамка, в якій містяться координати положення, висота і ширина кожного з об’єктів, а також визначається клас об’єкта для кожного положення.

3. Виявлення об’єктів має на меті ідентифікувати та класифікувати кожен конкретний об’єкт або набір об’єктів на зображенні. Під час виявлення об’єктів система автоматично визначає наявність об’єкта і передбачає його клас.

4. Сегментація зображень полягає в навчанні нейронної мережі або алгоритму машинного навчання ідентифікувати окремі об’єкти на основі пікселів зображення. Замість того, щоб створювати границі, пікселі об’єкта аналізуються окремо, і їхнє положення виділяється для підтвердження присутності об’єкта. У разі частково прихованих або

ISSN 2079-0031 (Print) ISSN 2411-0558 (Online)

захованих об’єктів система не може знайти затінені частини зображення і тому не повертає значення.

Алгоритми розпізнавання об’єктів. На успішність алгоритмів розпізнавання об’єктів впливають два фактори: ефективність алгоритму та кількість об’єктів або ознак на зображенні. Процес полягає в зіставленні зображеннь в алгоритмі машинного навчання для виявлення відповідних ознак, ідентифікації та визначення місця розташування об’єктів, присутніх на зображенні. Ознаки можуть бути функціональними або геометричними. Результатом завжди є лінійне або бінарне передбачення класу. Розглянемо етапи розпізнавання.

1. *Виділення ознак* – це процес пошуку характеристик або властивостей об’єктів, які можуть бути використані для їх ідентифікації або аналізу. Вони можуть бути отримані за допомогою алгоритмів машинного навчання при застосуванні навчених моделей згорткових нейронних мереж (CNN), таких як Alexnet та inception. Алгоритми створюють карту ознак зображення, полегшуючи ідентифікацію об’єктів

2. *Обмежувальні рамки*. Кожна частина зображення оточена обмежувальною або якірною рамкою. Обмежувальна рамка – це прямокутнник, який обмежує розташування об’єктів в зображенні і полегшує класифікацію. Обмежувальні рамки допомагають отримати інформацію про об’єкт із більш ніж 25 елементами даних, такими як графічні координати, бали ймовірності, висота і ширина.

3. *Формування гіпотез*. Кількість видобутих ознак зображення і якість навчальних даних, що надходять в алгоритм, є ключовими елементами під час формування гіпотези. Після видобування ознак система генерує оцінку ймовірності та присвоює її об’єктам, присутнім на зображенні. Це робиться головним чином для того, щоб знизити навантаження на класифікатор машинного навчання. Підсумковий результат розраховується на основі оцінки ймовірності та передбачення класу для кожного об’єкта на зображенні.

4. *Перевірка гіпотез*. На цьому етапі відбувається перевірка попередньої гіпотези, що призводить до отримання середнього бала класифікації – метрики, яку використовує алгоритм для розрахунку ефективності передбачень класів для різних об’єктів на зображенні.

Розгорнута модель ШІ перевіряє передбачення класу за відповідними характеристиками об'єкта (форма, розмір, колір тощо) і обмежувальною рамкою, що оточує об'єкт. Після перевірки обох параметрів система виставляє підсумкову композитну оцінку.

5. *Розпізнавання і зіставлення*. Після того, як алгоритм класифікував ознаки, він зіставляє обмежувальну рамку з координатами об'єкта. Ця інформація надходить у support vector machines (SVM), який використовує інструменти зростання FP (Feature Pyramid – піраміда ознак) для передбачення класу об'єкта в режимі реального часу.

6. *Лінійна регресія*. Етап лінійної регресії в процесі машинного навчання використовується для прогнозування числових значень на основі вхідних ознак. У контексті розпізнавання об'єктів на зображеннях, лінійна регресія використовується для знаходження зв'язку між ознаками об'єктів на зображенні та їхніми числовими характеристиками.

Результат роботи алгоритму розпізнавання об'єктів складається з набору класів об'єктів, а також деталей їх локації на зображенні.

Наразі одним з поширених класичних методів виявлення об'єктів на зображеннях є метод Віоли-Джонса та його модифікації [3, 4]. Проте ані він, ані інші математико-статистичні методи не можуть впоратися із такими проблемами, як, наприклад, зміна освітлення або деформація об'єкта, оскільки об'єкт може бути не тільки твердим тілом, а й таким, яке може деформуватися та змінювати свою форму, що ускладнює розпізнавання. В останні роки поява методів виявлення об'єктів на основі ШІ, таких як моделі глибокого навчання, значно підвищила точність і надійність систем виявлення об'єктів. Такі методи ШІ стають дедалі популярнішими завдяки їхній здатності автоматично навчатися та адаптуватися до різноманітних і складних сценаріїв.

Більшість досліджень присвячені CNN, оскільки задача виявлення об'єктів схожа на задачу класифікації зображень. Тоді як задача класифікації зображення вимагає лише визначення, що знаходиться на зображенні, задача виявлення вимагає визначення та локалізації одного або декількох об'єктів, отже задача виявлення об'єктів вимагала зробитит крок вперед від класифікації зображень [5]. Отже, з'явилось сімейство методів регіональних згорткових нейронних мереж (R-CNN), початкова

мета яких полягала в тому, щоб взяти вхідне зображення та створити набір обмежувальних рамок як вихід, де кожна обмежувальна рамка містить об'єкт, а також категорію (наприклад, автомобіль чи пішохід) об'єкта [6].

Ще одним методом виявлення об'єктів на зображеннях є SSD (Single Shot MultiBox Detector). Цей підхід дискретизує вихідний простір обмежувальних рамок на набір стандартних рамок з різним співвідношенням сторін і масштабом для кожного місця на карті об'єктів. Крім того, мережа комбінує прогнози з декількох карт об'єктів з різною роздільною здатністю, щоб природно обробляти об'єкти різного розміру [7].

Традиційні алгоритми виявлення об'єктів, такі як R-CNN та швидкий R-CNN, базуються на Region Proposal Network (RPN), які генерують велику кількість регіонів-кандидатів, які потім класифікуються та узгоджуються. Цей процес займає багато часу і робить ці алгоритми непрактичними для застосування в реальному часі.

SSD базується на архітектурі VGG16 CNN, що є потужним екстрактором ознак, найбільшим недоліком якого є низька продуктивність. SSD вирішує цю проблему із допомогою feature pyramid network (FPN). FPN поєднує ознаки з різних шарів CNN для формування піраміди ознак, що містить ознаки різного масштабу. Це дозволяє SSD виявляти об'єкти різного розміру [8].

SSD також використовує згладжену функцією втрат L1 (Smooth L1 loss function), яка є більш стійкою до викидів, ніж традиційна функція втрат L1, і краще підходить для виявлення об'єктів [9].

Проте згадані вище методи використовую дворівневу архітектуру нейронних мереж, через що виникають такі проблеми:

– ці методи не оглядають зображення повністю, а розглядають лише окремі регіони на зображенні, що впливає на точність виявлення об'єктів;

– через те, що цей підхід основується на великій кількості ітерацій огляду зображення, швидкість роботи алгоритму значно зменшується.

Для вирішення цих проблем було розроблено методи ШІ з однорівневою архітектурою (рис. 1). Найяскравішим прикладом таких методів є YOLO (You Only Look Once). YOLO розглядає виявлення об'єктів як задачу регресії до просторово розділених обмежувальних

ISSN 2079-0031 (Print) ISSN 2411-0558 (Online)

рамок і пов'язаних з ними ймовірностей класів. Одна нейронна мережа прогнозує обмежувальні рамки та ймовірності класів об'єктів безпосередньо з повних зображень за одну оцінку. Оскільки весь конвеєр виявлення є єдиною мережею, його можна оптимізувати наскрізь, безпосередньо впливаючи на точність та швидкість розпізнавання [10].

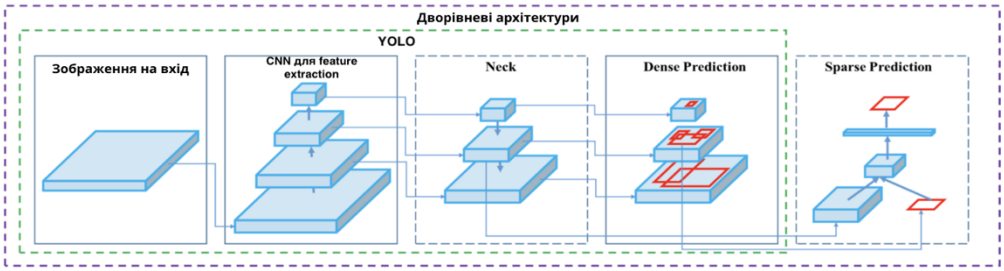

Рис. 1. Архітектура мережі YOLO та її порівняння з дворівневими нейронними мережами [10]

YOLO потрібен тільки один "погляд" на зображення, тому метод ковзного вікна не підходить у цій ситуації. Замість цього, зображення перетворюється на сітку з комірками розміром $SvoS$. Кожна комірка може містити кілька різних об'єктів для розпізнавання.

Кожна комірка відповідає за прогнозування кількості bounding boxes. Також кожна комірка прогнозує довірче значення (confidence value) для кожної області, обмеженої bounding box. Тобто, це значення визначає ймовірність знаходження того чи іншого об'єкта в даній області. В разі, якщо якась комірка сітки не має певного об'єкта, важливо, щоб довірче значення для цієї області було низьким.

Водночас кожна комірка відповідає за передбачення ймовірностей класів об'єктів. Це не говорить про те, що певна комірка містить певний об'єкт, тільки ймовірність знаходження об'єкта. Припустимо, якщо комірка передбачає автомобіль, це не гарантує, що автомобіль насправді присутній у ній. Це говорить лише про те, що якщо присутній об'єкт, то цей об'єкт найімовірніше автомобіль (рис. 2).

ISSN 2079-0031 (Print) ISSN 2411-0558 (Online)

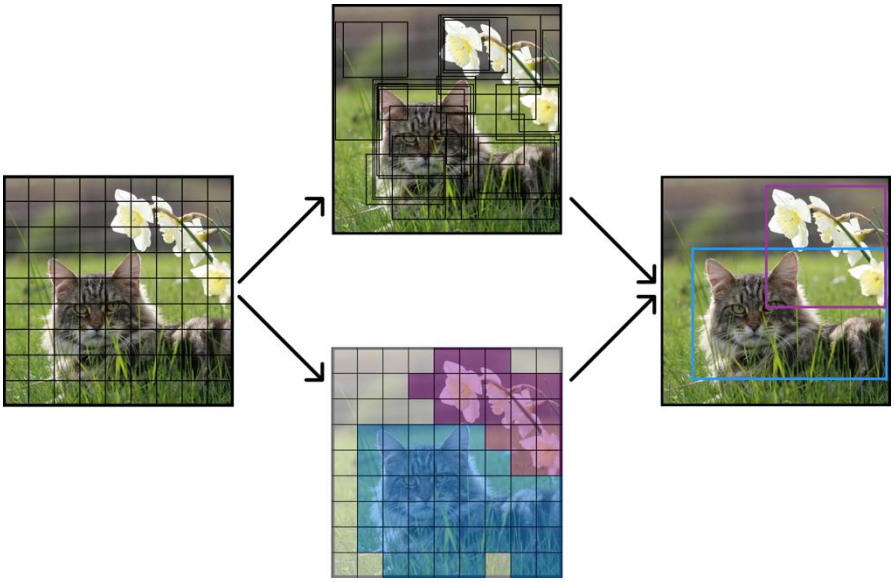

Рис. 2. Приклад роботи YOLO

YOLO використовує anchor boxes для прогнозування bounding boxes. Ідея полягає в тому, щоб заздалегідь визначити дві різні фігури. Таким чином, ми можемо об’єднати два передбачення з двома anchor boxes (загалом, можна використовувати більшу кількість anchor boxes). Ці anchor boxes розраховують за допомогою набору даних COCO (Common Objects in Context) та кластеризації за методом k-середніх [10].

Підхід отримав величезне поширення серед спільноти, тому науковці знаходили нові можливості для його покращення. Наприклад, YOLOv3 запровадив такі вдосконалення, як багатомасштабне навчання та нову функцію втрат для покращення стабільності та точності навчання. Багатомасштабний підхід YOLOv3 дозволяє охоплювати ширший діапазон розмірів об’єктів за один прохід, що робить його більш універсальним [13].

Експериментальне дослідження. Для проведення експерименту було обрано моделі YOLOv3, SSD та Faster R-CNN через різні архітектури цих моделей. Порівняння цих моделей також забезпечує порівняння підходів, до яких відносяться ці методи.

В якості тестових даних було взято послідовність зі 120 кольорових зображень розміру 854×480 пікселів. На послідовності зображено кота та собаку, що стрибають по траві.

Для порівняння результатів було обрано метрику mAP (Mean Average Precision), яка розраховується за формулою $mAP = \frac{1}{N}\sum_{i=1}^{N} AP_i$. Також, для порівняння було використано показник FPS (Frames per second) – показник кількості оброблених зображень за секунду часу.

Таблиця 1

Результати експериментального дослідження

Алгоритм	**mAP**	**FPS**
YOLO v3	81,15%	65
SSD	82,23%	38
Faster R-CNN	86,10%	5

Експеримент показав, що Faster R-CNN виявився кращим за точністю виявлення об’єктів (mAP – 86,10%), але значно гірший за швидкістю – усього 5 кадрів в секунду. YOLO v3 виявився найкращим з порівняних алгоритмів за швидкістю виявлення – 65 кадрів/секунду, mAP – 81,15%. SSD виявився швидшим за Faster R-CNN – він здатний виявляти об’єкти на послідовності зображень зі швидкістю 38 кадрів/секунду, а також показав точніший результат за YOLO v3 – 82,23% проти 81,15% у YOLO v3. Отже, якщо для певної задачі швидкість має найбільший пріоритет, варто обрати YOLO v3. В іншому випадку найкращий результат покаже Faster R-CNN.

Висновки. Технологія розпізнавання об’єктів є актуальною темою. Класичні методи комп’ютерного зору знаходять застосування завдяки їх обчислювальній ефективності та легкості інтерпретації. Однак, ці методи зазвичай не можуть впоратися з такими проблемами, як зміна освітленості або деформація об’єктів, особливо нежорстких об’єктів, які змінюють форму та розмір. А найголовніше, вони не дають можливості визначити або описати характерні ознаки об’єктів, які не завжди здатна сформулювати людина через складність їх формалізації.

ISSN 2079-0031 (Print) ISSN 2411-0558 (Online)

Дослідження виявлення об’єктів при цифровій обробці зображень показують, що вибір моделі та методу залежить від конкретних вимог програми. Традиційні методи, такі як порівняння шаблонів, аналіз кольору і текстури та використання контурів, ефективні в простих сценаріях, але обмежені при зміні умов освітлення або деформації об’єктів. Зі збільшенням обсягу даних і розвитком обчислювальних можливостей методи, засновані на штучному інтелекті, особливо глибокі нейронні мережі, стали домінувати у виявленні об’єктів завдяки їхній високій точності та адаптивності.

Експериментальне порівняння моделей YOLOv3, SSD та Faster R-CNN показало їхні переваги та недоліки. Faster R-CNN забезпечує найвищу точність виявлення об’єктів (86,10% точність виявлення), але має низьку швидкість обробки (5 кадрів на секунду) і не підходить для роботи в реальному часі. Модель YOLOv3 показала найшвидший результат (65 кадрів на секунду), але з дещо нижчою точністю (точність виявлення – 81,15%), що робить її ідеальною для застосувань, де швидкість є пріоритетом. Завдяки балансу точності (виявлення з точністю 82,23%) і швидкості (38 кадрів в секунду), модель SSD ідеально підходить для додатків, які вимагають і того, і іншого.

Методи глибокого навчання, такі як YOLO і його вдосконалені версії, значно підвищили швидкість і точність, завдяки архітектурним інноваціям. Наприклад, YOLOv3 реалізує багатомасштабне навчання, що дозволяє більш ефективно обробляти об’єкти різного розміру.

Як правило, вибір конкретної моделі або методу виявлення об’єктів повинен ґрунтуватися на вимогах точності, швидкості та складності зображення. Подальші дослідження в цій галузі спрямовані на підвищення стійкості до змін умов освітлення, деформації об’єктів та інших факторів, що ускладнюють процес виявлення. Технології штучного інтелекту та глибокого навчання продовжують розвиватися, відкриваючи нові можливості для вдосконалення системи виявлення об’єктів на цифрових зображеннях.

Перспективи подальшого дослідження виявлення об’єктів в цифровій обробці зображень включають розробку моделей з більш високою точністю і швидкістю, адаптованих до різних умов освітлення і

змін об’єктів. Важливо дослідити роль методів пошуку ключових точок в задачі виявлення об’єктів. Такий підхід може покращити швидкість, а також допоможе відслідковувати рух об’єктів на зображеннях та порівнювати об’єкти для пітдвердження унікальності об’єкту.

Список літератури

1. Exposit (2021), "Computer Vision Object Detection: challenges faced", available at: www.exposit.com/blog/computer-vision-object-detection-challenges-faced/
2. Marr D. (2010), "Vision: A Computational Investigation into the Human Representation and Processing of Visual Information", The MIT Press, available at: www.direct.mit.edu/books/book/3299/VisionA-Computational-Investigation-into-the-Human
3. Viola, P., Jones, M.J. (2004), “Robust Real-Time Face Detection”, International Journal of Computer Vision 57, pp. 137–154.
4. Лисенко, Г. Л., Тарновський, М. Г., Кузьменко, Л. В. (2017), Сучасні тенденції у вирішенні задач виявлення та розпізнавання об′ єктів на зображеннях. Оптико-електронні інформаційно-енергетичні технології, 33(1), 18-23.
5. Tatan V. (2021),” Understanding Object Detection”, available at: towardsdatascience.com/understanding-object-detection-1d810164b8a6
6. Gandhi R. (2018), “R-CNN, Fast R-CNN, Faster R-CNN, YOLO — Object Detection Algorithms”, available at: https://towardsdatascience.com/r-cnn-fast-r-cnn-faster-r-cnn-yolo-object-detection-algorithms-36d53571365e
7. Liu W., Anguelov D., Erhan D., Szegedy C., Reed S., Cheng-Yang Fu, Alexander C. Berg, (2015), “SSD: Single Shot MultiBox Detector”, available at: https://arxiv.org/abs/1512.02325
8. Bhavya Sree, B., Yashwanth Bharadwaj, V., & Neelima, N. (2021). An Inter-Comparative Survey on State-of-the-Art Detectors—R-CNN, YOLO, and SSD. In Intelligent Manufacturing and Energy Sustainability: Proceedings of ICIMES 2020 (pp. 475-483). Springer Singapore.
9. Zhang S., Wen L., Shi H., (2019), “Single-Shot Scale-Aware Network for Real-Time Face Detection”, Int J Comput Vis 127, pp. 537–559, available at: https://doi.org/10.1007/s11263-019-01159-3
10. Redmon J., Divvala S., Girshick R., Farhadi A., (2015), “You Only Look Once: Unified, Real-Time Object Detection”, available at: https://arxiv.org/abs/1506.02640
11. Redmon J., Farhadi A., (2018), “YOLOv3: An Incremental Improvement”, available at: https://arxiv.org/abs/1804.02767
12. Nepal U., Eslamiat H., (2022), “Comparing YOLOv3, YOLOv4 and YOLOv5 for Autonomous Landing Spot Detection in Faulty UAVs. Sensors”, 22(2):464., available at: https://doi.org/10.3390/s22020464

Статтю представив д-р. техн. наук, проф. НТУ "ХПІ" О. А. Серков.

Надійшла (Received) 27.05.2024

Liubchenko Vira, Doctor of Science, Professor
National University “Odesa Polytechnic”
Shevtchenko Ave., 1, Odesa, Ukraine, 65044
Tel.: +380-50-392-21-50, email: lvv@op.edu.ua
ORCID ID: 0000-0002-4611-7832

ISSN 2079-0031 (Print) ISSN 2411-0558 (Online)

Moroz Dmytro, PhD student
National University “Odesa Polytechnic”
Shevtchenko Ave., 1, Odesa, Ukraine, 65044
Tel.: +380-93-234-78-27, email: mordmytro@stud.op.edu.ua
ORCID ID: 0009-0002-1082-5033

УДК 004.93

Моделі та методи виявлення об’єктів в цифровій обробці зображень / Любченко В.В., Мороз Д.В. // Вісник НТУ "ХПІ". Серія: Інформатика та моделювання. – Харків: НТУ "ХПІ". – 2024. – № 1 – 2 (11 – 12). – С. 61 – 74.

Виявлення об'єктів в цифрових зображеннях може стикатися з такими проблемами, що виникають в процесі реєстрації зображень, як наявність шуму, низька якість/роздільна здатність зображень, неоднорідність освітлення, перекриття об'єктів та інші. Ці проблеми ускладнюють процес виявлення об'єктів та призводять до помилок у роботі алгоритмів обробки зображень. Для розв’язання цих проблем досліджуються переваги, недоліки та технічні особливості моделей і методів виявлення об'єктів у цифрових зображеннях для їх обґрунтованого вибору в практичних застосуваннях. Іл.: 2. Бібліогр.: 11 назв

Ключові слова: комп’ютерний зір, розпізнавання об’єктів, обробка зображень, технології штучного інтелекту.

UDK 004.93

Models and methods of object detection in digital image processing / Lyubchenko V.V., Moroz D.V. // Herald of the National Technical University "KhPI". Series of "Informatics and Modeling". – Kharkov: NTU "KhPI". – 2024. – № 1 – 2 (11 – 12). – P. 61 – 74.

Object detection in digital images can face such problems that arise in the process of image registration as the presence of noise, low image quality/resolution, illumination heterogeneity, overlapping objects, and others. These problems complicate the process of object detection and lead to errors in image processing algorithms. To solve these problems, we study the advantages, disadvantages, and technical features of models and methods for detecting objects in digital images for their reasonable selection in practical applications. Figs. 2. Refs. 11 titles.

Keywords: computer vision, object recognition, image processing, artificial intelligence technologies.