

М. Ю. Голенко

УДОСКОНАЛЕННЯ АРХІТЕКТУРИ АЛГОРИТМУ FASTER R-CNN ДЛЯ ПІДВИЩЕННЯ ШВИДКОДІЇ І ТОЧНОСТІ РОЗПІЗНАВАННЯ ОБ'ЄКТІВ

У статті наведено дослідження, спрямоване на удосконалення алгоритму розпізнавання Faster R-CNN з метою підвищення його точності та надійності в задачах розпізнавання об'єктів. У дослідженні розглядаються три ключові структурні покращення, які були впроваджені в алгоритм. Перше покращення передбачає інтеграцію EfficientNet як основної мережі, що значно підвищує якість видобутку ознак, завдяки збалансованому підходу до масштабу глибини мережі, її ширини та роздільної здатності. Це дозволяє досягти більш високої точності при зменшенні кількості параметрів та обчислювальних витрат, що є важливим для практичних застосувань, особливо в умовах обмежених ресурсів. Друге покращення стосується застосування без'якірного підходу CenterNet для мережі пропозицій регіонів (RPN), який спрощує процес локалізації об'єктів, безпосередньо передбачаючи їх центри та розміри. Це усуває необхідність використання попередньо визначених якорів і суттєво підвищує точність локалізації, що особливо важливо для розпізнавання об'єктів різних розмірів і форм, які часто зустрічаються в реальних умовах. Третє покращення полягає в оптимізації функції втрат за допомогою генералізованої інтерсекції над об'єднанням (GIoU), яка забезпечує більш стабільну і точну регресію обмежувальних рамок, що веде до кращої точності локалізації, особливо в складних умовах, таких як перекриття об'єктів або розташування на краях зображення. Проведені структурні зміни в алгоритмі сприяють значному підвищенню його точності та надійності. Експериментальне дослідження на стандартних тестових наборах даних, таких як COCO та PASCAL VOC, показало суттєве зростання середньої точності (mAP) і зниження часу обробки. Отримані результати підкреслюють переваги запропонованих змін у алгоритмі Faster R-CNN, що робить його більш придатним для застосувань в реальних умовах.

Ключові слова: розпізнавання об'єктів, нейронні мережі, оптимізація алгоритму, Faster R-CNN, EfficientNet, CenterNet, втрата GIoU.

Вступ

З розвитком комп'ютерного зору та машинного навчання, алгоритми розпізнавання об'єктів стали ключовими інструментами для багатьох сучасних застосувань, включаючи автономні транспортні засоби, системи відеоспостереження та робототехніку. Серед них алгоритм Faster R-CNN зарекомендував себе як один із найпродуктивніших і найбільш точних методів для швидкого виявлення об'єктів. Однак, попри його переваги, існують певні обмеження, які стримують його продуктивність, особливо в умовах реального часу.

Ця стаття присвячена вдосконаленню алгоритму Faster R-CNN з метою підвищення його точності та швидкості. Для вирішення цієї задачі запропонована зміна структури алгоритму з використанням інтеграції методів EfficientNet, CenterNet та GIoU. Інтеграція нових методів з алгоритмом дозволяє значно підвищити його продуктивність на тестових наборах даних COCO та PASCAL VOC, демонструючи переваги нового підходу.

У цій статті проведено детальний аналіз наявних методів, обґрунтовано вибір запропонованих рішень та представлені експериментальні результати, що підтверджують результативність запропонованого підходу.

Метою статті є удосконалення методу та алгоритму Faster R-CNN шляхом інтеграції нових підходів, спрямованих на підвищення точності та надійності розпізнавання об'єктів. Це включає оптимізацію видобутку ознак, підвищення точності локалізації та вдосконалення функції втрат, що дозволяє забезпечити більш стабільне функціонування алгоритму у різноманітних умовах. Мета включає декілька конкретних завдань:

1. Підвищення точності. Інтегруючи передові методи видобутку ознак має за собою ціль покращити точність розпізнавання об'єктів, особливо для складних випадків, таких як об'єкти малого розміру.

2. Зменшення часу обчислень. Оптимізація алгоритму для швидшого виконання без втрати точності є важливою для реальних застосувань. Це включає зменшення загальних обчислювальних витрат і забезпечення надійності роботи алгоритму на обмежених апаратних ресурсах.

3. Підвищення надійності. Покращений алгоритм повинен бути стійким до варіацій у масштабі об'єктів, орієнтації та умов навколишнього середовища.

Покращення розроблені для створення синергетичного ефекту, де одне рішення доповнює інше, призводить до комплексного оновлення алгоритму Faster R-CNN. Такий цілісний підхід гарантує, що ці вдосконалення не є ізольованими, а працюють разом для досягнення основної мети – кращої точності і надійності розпізнавання об'єктів.

Постановка проблеми

Алгоритм Faster R-CNN встановив високий стандарт у сфері розпізнавання об'єктів, але в нього є певні обмеження, які заважають його продуктивності та застосуванню в певних сценаріях. Обмеження переважно пов'язані з видобутком ознак, механізмами пропозиції регіонів та регресією обмежувальних рамок. Усунення цих проблем є важливим для підвищення можливостей алгоритму.

Традиційні основні мережі, такі як VGG16 та ResNet, що зазвичай використовуються у Faster R-CNN, можуть мати труднощі з видобутком ознак, особливо об'єктів малих розмірів. Ці нейронні мережі також можуть мати вищі обчислювальні витрати, що обмежує їх використання у реальних застосуваннях.

Оригінальна мережа пропозицій регіонів (RPN) у Faster R-CNN покладається на попередньо визначені якорі для генерації пропозицій регіонів. Такий підхід з використанням якорів може бути субоптимальним, оскільки вимагає ретельного налаштування розмірів якорів та співвідношень сторін, і може викликати проблеми з різними наборами даних та масштабами об'єктів.

Традиційна втрата Intersection over Union (IoU), яка використовується для регресії обмежувальних рамок у Faster R-CNN, має обмеження, особливо коли немає перекриття між прогнозованими та реальними рамками. Це може призводити до низької продуктивності локалізації об'єкту, особливо для об'єктів на краях зображення або в складних сценах.

Наведені обмеження підкреслюють необхідність методологічних покращень алгоритму Faster R-CNN. Усунення цих проблемних областей не лише має покращити точність та надійність алгоритму, але й підвищити його надійність, що зробить його більш придатним для широкого спектра практичних застосувань. Дослідження має мету структурно удосконалити алгоритм, запроваджуючи комплексний набір змін, які базуються на сильних сторонах Faster R-CNN та одночасно мінімізують його слабкі сторони.

Аналіз останніх досліджень і публікацій

Для досягнення запропонованих вдосконалень алгоритму Faster R-CNN важливо спиратися на наявні дослідження, які заклали основу в різних галузях глибокого навчання та розпізнавання об'єктів.

У роботі [1] було наведено алгоритм Faster R-CNN, який інтегрував мережу пропозицій регіонів (RPN) з детектором Fast R-CNN. Це поєднання дозволило здійснювати навчання від кінця до кінця та зменшити обчислювальні витрати, встановивши новий стандарт для розпізнавання об'єктів у реальному часі. RPN генерує високоякісні пропозиції регіонів, проходячи невелику мережу по конволюційній карті ознак, які потім уточнюються детектором Fast R-CNN.

У дослідженні [2] було розроблено мережі пірамід ознак (FPN) для покращення точності розпізнавання на різних масштабах об'єктів. Використовуючи ієрархічні карти ознак, FPN підвищує здатність моделі розпізнавати об'єкти різних розмірів, що є критичним для комплексного розпізнавання об'єктів. FPN стали основною технікою в багатьох сучасних

алгоритмах розпізнавання об'єктів.

Для вирішення проблеми дисбалансу класів у наборах даних для розпізнавання об'єктів було запропоновано FocalLoss [3]. Ця техніка модифікує стандартну втрату перехресної ентропії, щоб більше зосередитися на важко доступних об'єктах, тим самим покращуючи загальну продуктивність розпізнавання. Це нововведення особливо корисне для наборів даних з великим дисбалансом міжфоновими та об'єктними класами.

EfficientNet, який використовує метод композиційного масштабування для балансування глибини мережі, ширини та роздільної здатності, був наведений у дослідженні [4]. Цей дизайн досягає вищої точності з меншою кількістю параметрів, що робить його привабливим вибором для інтеграції як основної мережі в Faster R-CNN. Метод масштабування EfficientNet забезпечує досягнення передової продуктивності з оптимізованими обчислювальними ресурсами.

Без'якірний підхід CenterNet, наведений у дослідженні [5], передбачає центри та розміри об'єктів безпосередньо, усуваючи потребу у попередньо визначених якорях. Цей метод спрощує процес навчання та покращує точність локалізації, надаючи точні пропозиції регіонів без складності налаштування якорів.

Втрата GIoU, запропонована у дослідженні [6], надає більш точну міру перекриття між прогнозованими та реальними обмежувальними рамками, вирішуючи обмеження традиційної втрати IoU та покращуючи продуктивність локалізації. GIoU враховує найменшу обмежувальну рамку, яка охоплює як прогнозовані, так і реальні рамки, забезпечуючи більш всеосяжну метрику для регресії обмежувальних рамок.

В результаті проведення аналізу схожих досліджень в області розпізнавання об'єктів та глибокого навчання, ми визначили декілька ключових підходів, які пропонують значні покращення для алгоритму Faster R-CNN. Вибір технологій для подальших досліджень був обґрунтований їхньою здатністю вирішувати основні проблеми, з якими стикається Faster R-CNN, такі як точність та обчислювальні витрати.

Запропоновані покращення

Інтеграція EfficientNet як основної мережі запропонована для покращення можливостей видобутку ознак. Метод композиційного масштабування EfficientNet підвищує здатність мережі захоплювати детальні та дискримінативні ознаки, що призводить до більш точного розпізнавання об'єктів.

Також запропоновано без'якірний підхід з використанням CenterNet для мережі пропозицій регіонів (RPN). CenterNet безпосередньо прогнозує центри та розміри об'єктів, що спрощує процес пропозиції регіонів та покращує точність локалізації об'єктів при розпізнаванні.

Функція втрат оптимізована шляхом впровадження втрат генералізованої інтерсекції над об'єднанням (GIoU). GIoU пропонує більш точну та стабільну міру перекриття між прогнозованими та реальними обмежувальними рамками, тим самим підвищуючи точність локалізації об'єктів.

Інтеграція техніки видобутку ознак

Інтеграція EfficientNet як основної мережі в рамках Faster R-CNN представляє значний прогрес у можливостях видобутку ознак. EfficientNet використовує метод композиційного масштабування, який рівномірно масштабує глибину, ширину та роздільну здатність мережі, оптимізуючи баланс між цими вимірами для досягнення найвищої продуктивності [4]. EfficientNet уникає проблем надмірної параметризації та недопараметризації, що призводить до більш збалансованої мережі.

Композиційне масштабування EfficientNet визначається формулою:

$$\text{depth} = \alpha^d, \quad \text{width} = \beta^d, \quad \text{resolution} = \gamma^d \quad (1)$$

де α , β та γ є константами, визначеними архітектурою, а d є коефіцієнтом, який контролює масштабування [4]. Цей метод дозволяє EfficientNet захоплювати більш детальні та дискримінативні ознаки з вхідних зображень, що є критично важливим для розпізнавання об'єктів різних масштабів, форм та зовнішніх виглядів.

Архітектура EfficientNet видобуває детальні ознаки, підвищуючи точність розпізнавання об'єктів. Покращене представлення ознак обумовлене здатністю мережі рівномірно масштабуватися в різних вимірах [4].

EfficientNet показала вищу точність на різних тестових наборах порівняно з традиційними мережами, такими як VGG16 та ResNet. Покращення особливо помітне у сценаріях з комплексними та різноманітними наборами даних [4]. Дослідження показали, що EfficientNet досягає передової продуктивності на кількох ключових тестових наборах, що робить її надійним вибором для задач розпізнавання об'єктів.

Незважаючи на покращену продуктивність, EfficientNet підтримує нижчі обчислювальні витрати, що робить її придатною для задач де обчислювальні ресурси обмежені. EfficientNet досягає цієї мети, оптимізуючи баланс між глибиною мережі, шириною та роздільною здатністю, забезпечуючи, щоб мережа не стала надмірно великою або обчислювально витратною [4].

Інтеграція EfficientNet у алгоритм Faster R-CNN включає заміну оригінального екстрактора ознак на EfficientNet. Процес інтеграції EfficientNet у Faster R-CNN вносить кілька модифікацій для забезпечення сумісності та оптимальної продуктивності. Зміни включають налаштування розмірів вхідних даних для відповідності вимогам EfficientNet, тонке налаштування мережі для забезпечення безперебійної роботи з мережею пропозицій регіонів (RPN) та модифікацію процедур навчання для врахування нової основної мережі. Ці адаптації забезпечують, що інтегрована система зберігає переваги як EfficientNet, так і FasterR-CNN без втрати продуктивності.

Покращення мережі пропозицій регіонів

Для подальшого покращення продуктивності алгоритму Faster R-CNN використовується без'якірний підхід з використанням CenterNet для мережі пропозицій регіонів (RPN). За допомогою використання цього підходу вирішується кілька обмежень традиційних методів на основі якорів, які покладаються на попередньо визначені якорі для генерації пропозицій регіонів.

Без'якірний підхід дозволяє легше адаптуватися до об'єктів різних розмірів та форм. Ця адаптивність особливо корисна у різноманітних і складних наборах даних, де об'єкти не відповідають попередньо визначеним значенням якорів. Пряме прогнозування центрів об'єктів дозволяє точніше визначити їх місцезнаходження. Цей метод зменшує помилки локалізації об'єктів, які можуть виникати, коли якорі не добре узгоджуються з об'єктами на зображенні [5].

CenterNet прогнозує центри та розміри об'єктів безпосередньо, усуваючи потребу у попередньо визначених якорях. Усунувши потребу в якорях, модель стає простішою для навчання та налаштування. Це зменшує складність, пов'язану з вибором відповідних розмірів якорів та співвідношень сторін, що може бути значним викликом у традиційних методах RPN.

Метод CenterNet передбачає використання двох окремих карт для виявлення об'єктів: теплової карти для прогнозування центрів об'єктів і карти розмірів для прогнозування їх розмірів.

Теплова карта визначає ймовірність того, що центр об'єкта знаходиться в певній точці (x, y) зображення. Формула теплової карти:

$$H(x, y) = \exp\left(-\frac{(x-x_c)^2 + (y-y_c)^2}{2\sigma^2}\right) \quad (2)$$

де (x_c, y_c) є центром об'єкта, а σ контролює розподіл теплової карти. Висока ймовірність на тепловій карті вказує на центр об'єкта [5].

Карта розмірів прогнозує ширину та висоту об'єктів у кожній точці (x, y) . Формула для карти розмірів може бути представлена як:

$$S(x, y) = [w(x, y), h(x, y)] \quad (3)$$

де $w(x, y)$ і $h(x, y)$ представляють ширину і висоту об'єкта, центр якого знаходиться в точці (x, y) [5].

Таким чином, для кожної точки з високою ймовірністю на тепловій карті є відповідна ширина і висота на карті розмірів. Карта розмірів прогнозує ширину $w(x, y)$ і висоту $h(x, y)$ об'єкта в кожній точці (x, y) .

Інтеграція CenterNet у алгоритм Faster R-CNN забезпечує пряме прогнозування центрів та розмірів об'єктів і збільшує точність пропозицій регіонів, що призводить до кращої загальної продуктивності розпізнавання. Усунення якорів спрощує процес навчання та зменшує потребу в обширному налаштуванні гіперпараметрів. Без'якірний підхід є більш гнучким та надійним у різних наборах даних та типах об'єктів, що робить його придатним для широкого спектра застосувань.

Впровадження CenterNet у алгоритм Faster R-CNN включає модифікацію RPN для генерації теплових карт та карт розмірів замість пропозицій на основі якорів. Ця зміна вимагає адаптації архітектури мережі та процедур навчання для забезпечення сумісності та оптимальної продуктивності.

Оптимізація функції втрат

Традиційна втрата Intersection over Union (IoU), яка використовується у Faster R-CNN, має певні недоліки, коли немає перекриття між прогнозованими та реальними обмежувальними рамками. GIoU надає більш точну міру якості прогнозів обмежувальних рамок, що призводить до кращої локалізації об'єктів. Це особливо корисно у сценаріях, якщо об'єкти частково закриті або розташовані на краях зображення.

GIoU вирішує проблеми нестабільності, пов'язані з IoU, коли прогнозовані та реальні рамки не перекриваються. В свою чергу це надає можливість досягти надійного навчання та кращої збіжності моделі. Забезпечуючи точну метрику втрат, GIoU допомагає моделі краще навчатися, що підвищує загальну продуктивність розпізнавання.

Втрати GIoU визначаються формулою:

$$GIoU = IoU - \frac{|C - (A \cup B)|}{|C|} \quad (4)$$

де IoU є Інтерсекцією над Об'єднанням, A та B – це прогнозовані та реальні обмежувальні рамки, а C – це найменша обмежувальна рамка, що охоплює як A , так і B [6]. Включаючи площу найменшої обмежувальної рамки, GIoU надає більш комплексну метрику, яка накладає штрафи на прогнози не тільки на основі їх перекриття, але й на основі їх загального просторового розташування.

Впровадження GIoU у алгоритм Faster R-CNN включає заміну традиційної IoU на GIoU у розрахунку функції втрат. Ця зміна вимагає коригування процесу навчання, щоб забезпечити правильне інтерпретування та оптимізацію нової метрики.

Оптимізуючи функцію втрат за допомогою GIoU, покращений алгоритм Faster R-CNN досягає вищої точності та надійності у розпізнаванні об'єктів.

Набори даних

Перед проведенням експериментального дослідження, необхідно визначити набори даних для подальшої роботи з навчання моделі на основі інтегрованих методів. Серед наборів даних було обрано наступні:

1. Набір даних COCO (Common Objects in Context) – це великомасштабний набір для

розпізнавання об'єктів, сегментації та створення підписів. Він містить понад 200 000 зображень з анотаціями для 80 категорій об'єктів, включаючи людей, тварин, транспортні засоби та повсякденні об'єкти [7]. Різноманітність та складність набору даних COCO роблять його ідеальним еталоном для оцінки алгоритмів розпізнавання об'єктів. Набір даних розділений на кілька піднаборів, таких як train2017 та val2017, для навчання та перевірки.

2. Набір даних PASCAL VOC (Visual Object Classes) надає стандартизовані набори зображень для розпізнавання класів об'єктів. Він включає зображення для 20 категорій об'єктів з повними анотаціями для розпізнавання, сегментації та класифікації [8]. Набір даних складається приблизно з 11 530 зображень, з близько 27 450 екземплярами об'єктів.

Умови проведення експериментального дослідження

Всі зображення змінюються до одного розміру, щоб відповідати вимогам введення моделі Faster R-CNN. Нормалізуються значення пікселів для покращення збіжності під час навчання.

Набори даних розподіляються на навчальні, валідаційні та тестові набори. Для COCO використовуються стандартні розподіли (наприклад, train2017, val2017) [7]. Для PASCAL VOC використовуються розподіли 2007 та 2012 років, з комбінуванням наборів train/val для навчання та тестового набору для оцінки [8].

Моделю навчається з використанням конкретних гіперпараметрів, таких як швидкість навчання 0.001, розмір пакету 16 та імпульс 0.9. Оптимізатор Adam використовується для мінімізації функції втрат, використовуючи його здатність до адаптивного навчання [9].

Аналіз результатів експерименту

Для оцінки результатів було визначено наступні метрики:

1. Середня середня точність (mAP використовується як основна метрика оцінки, яка обчислюється при різних порогах IoU (наприклад, 0.5, 0.75) [7,8].
2. Криві точність-відклик генеруються для візуалізації компромісу між точністю та відкликом при різних порогах впевненості коректного розпізнавання об'єкту [8].
3. Час розпізнавання об'єкту вимірюється для оцінки продуктивності та швидкодії покращеної моделі Faster R-CNN[1].

Таблиця 1

Результати порівняння продуктивності на наборі даних COCO

Метрика	Базовий Faster R-CNN	Покращений Faster R-CNN
mAP (IoU=0.5)	42.0%	48.5%
mAP (IoU=0.75)	25.0%	32.0%
Загальний mAP	33.5%	40.0%
Час Ввисновку (мс)	85	70

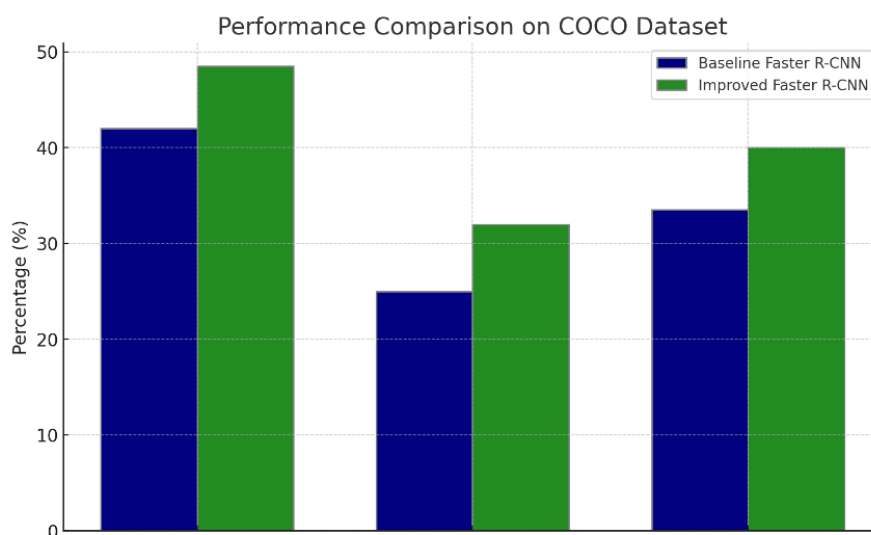


Рис. 1. Діаграма змін у продуктивності роботи звичайного та покращеного алгоритму на наборі даних COCO

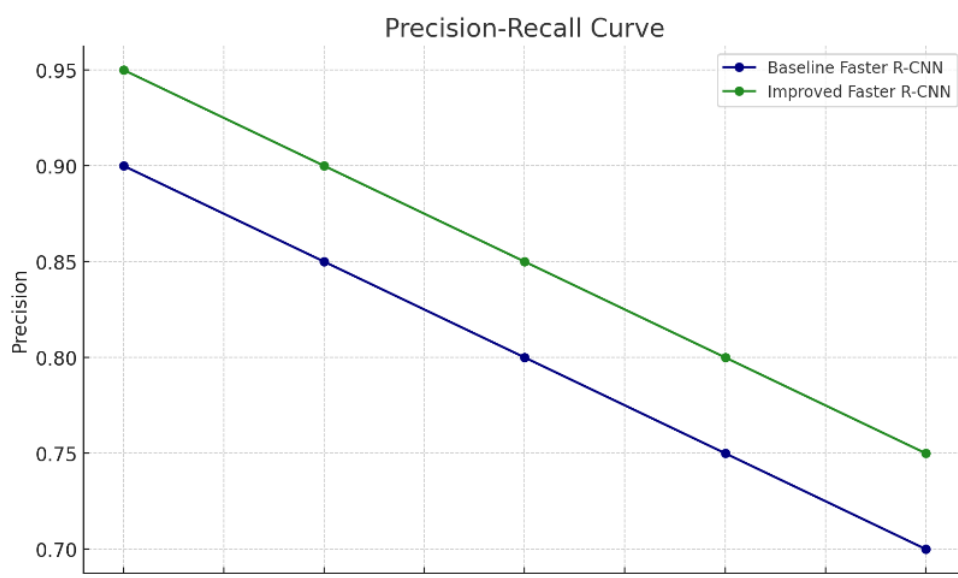


Рис. 2. Графік змін точності розпізнавання звичайного та покращеного алгоритму на наборі даних COCO

Таблиця 2

Результати порівняння продуктивності на наборі даних PASCAL VOC

Метрика	Базовий Faster R-CNN	Покращений Faster R-CNN
mAP (IoU=0.5)	75.0%	80.5%
mAP (IoU=0.75)	55.0%	62.0%
Загальний mAP	65.0%	71.2%
Час висновку (мс)	80	65

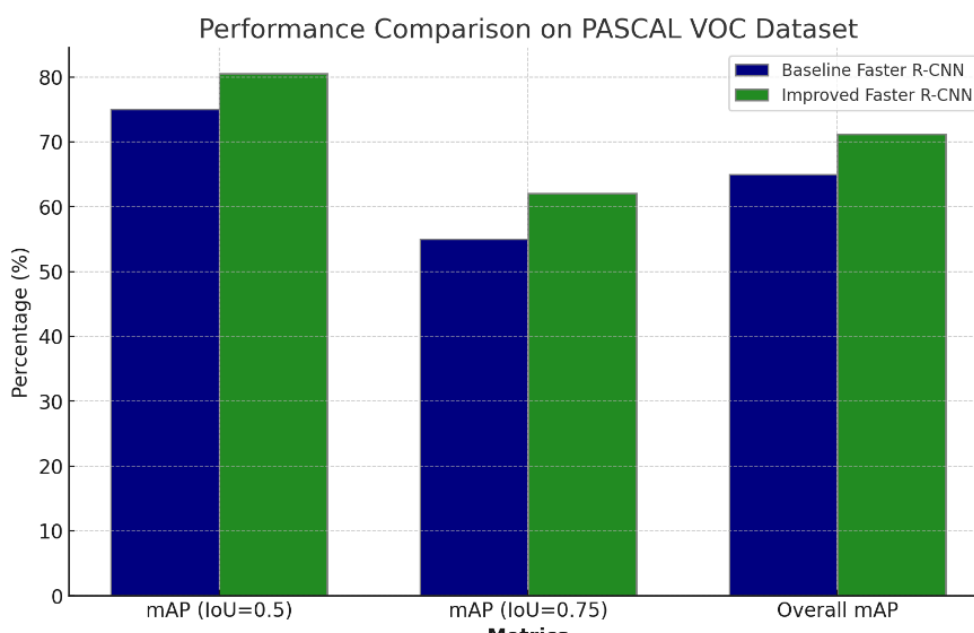


Рис. 3. Діаграма змін у продуктивності роботи звичайного та покращеного алгоритму на наборі даних PASCAL VOC

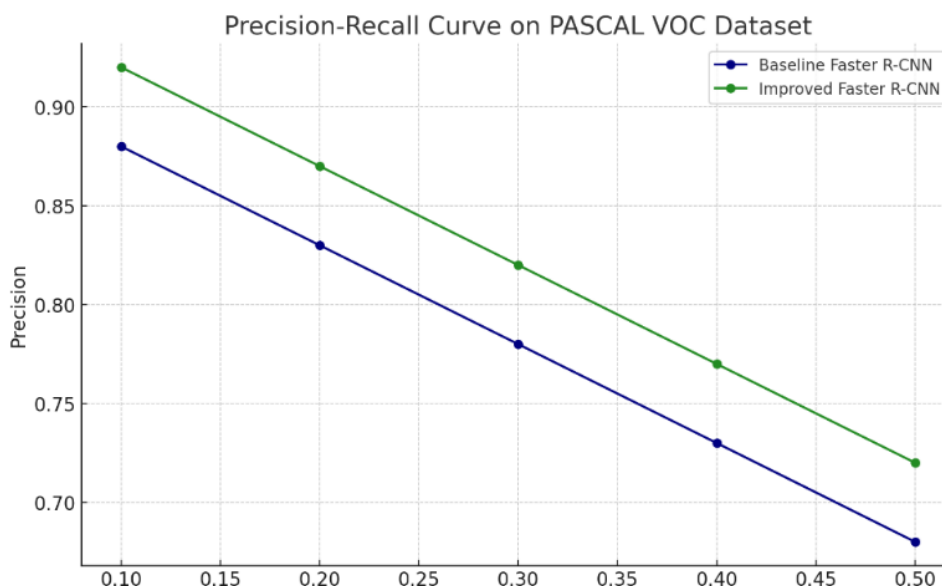


Рис. 4. Графік змін точності розпізнавання звичайного та покращеного алгоритму на наборі даних PASCAL VOC

На наборі даних COCO середня середня точність (mAP) при IoU=0.5 збільшилася з 42.0 % до 48.5 %, а при IoU=0.75 з 25.0 % до 32.0 %. Загальний mAP покращився з 33.5 % до 40.0 %.

На наборі даних PASCAL VOC mAP при IoU=0.5 збільшився з 75.0 % до 80.5 %, а при IoU=0.75 з 55.0 % до 62.0 %. Загальний mAP покращився з 65.0 % до 71.2 %.

Час висновку на наборі даних COCO зменшився з 85 мс до 70 мс, а на наборі даних PASCAL VOC з 80 мс до 65 мс. Це зменшення часу висновку підкреслює підвищену швидкодію запропонованих покращень. Отже, алгоритм став більш придатним для реальних застосувань у режимі реального часу.

Експериментальні результати демонструють значні покращення в середній точності (mAP) та часі висновку, що свідчить про позитивний вплив запропонованих рішень на продуктивність алгоритму Faster R-CNN. Модель показує загальне збільшення точності та продуктивності.

Висновки

У статті наведено значні структурні зміни до алгоритму Faster R-CNN, спрямовані на підвищення його продуктивності у задачах розпізнавання об'єктів. Інтегруючи EfficientNet як основну мережу, впроваджуючи CenterNet для пропозицій регіонів та оптимізуючи функцію втрат за допомогою генералізованої інтерсекції над об'єднанням (GIoU), було продемонстровано загальне зростання показників точності, продуктивності та зменшення часу висновку роботи Faster R-CNN.

Покращений алгоритм Faster R-CNN досяг вищої середньої точності (mAP) за різних порогів IoU на наборах даних COCO та PASCAL VOC. Зокрема, mAP при IoU=0.5 збільшився з 42.0 % до 48.5 % на наборі даних COCO і з 75.0 % до 80.5 % на наборі даних PASCAL VOC. Ці зміни підкреслюють успішність застосування EfficientNet та CenterNet.

Інтеграція EfficientNet і CenterNet, у поєднанні з використанням втрат GIoU, призвели до зменшення часу висновку. На наборі даних COCO час висновку зменшився з 85 мс до 70 мс, а на наборі даних PASCAL VOC з 80 мс до 65 мс. Це демонструє придатність алгоритму для реальних застосувань, де важливо швидке та точне розпізнавання об'єктів.

Зростання показників метрик, які спостерігаються на обох наборах даних COCO та PASCAL VOC, вказують на надійність запропонованих структурних змін алгоритму. Faster R-CNN почав краще справлятися з різноманітними та складними середовищами при розпізнаванні об'єктів, що робить його універсальним рішенням для різних сценаріїв використання.

Проведене дослідження демонструє значні покращення у точності та надійності алгоритму Faster R-CNN, що відкриває нові можливості для його застосування у задачах розпізнавання об'єктів. Подальші дослідження можуть зосередитися на додаткових оптимізаціях для забезпечення ще вищої точності та продуктивності.

СПИСОК ЛІТЕРАТУРИ

1. Faster R-CNN: Towards real-time object detection with region proposal networks / S. Ren, K. He, R. Girshick [et al.] // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2017. – №39 (6). – P. 1137 – 1149.
2. Feature pyramid networks for object detection / T.-Y. Lin, P. Dollár, R. Girshick [et al.] // In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. – 2017. – P. 2117 – 2125.
3. Focal loss for dense object detection / T.-Y. Lin, P. Goyal, R. Girshick [et al.] // In Proceedings of the IEEE International Conference on Computer Vision. – 2017. – P. 2980 – 2988.
4. Tan M. EfficientNet: Rethinking model scaling for convolutional neural networks / M. Tan, Q. V. Le // In Proceedings of the International Conference on Machine Learning. – 2019. – P. 6105 – 6114.
5. Objects as points [Electronic resource] / X. Zhou, D. Wang, P. Krähenbühl // arXiv preprint. – 2019. – Access mode : arXiv:1904.07850.
6. Generalized intersection over union: A metric and a loss for bounding box regression. / H. Rezatofighi, N. Tsoi, J. Y. Gwak [et al.] // In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. – 2019. – P. 658 – 666.
7. Microsoft COCO: Common objects in context / T.-Y. Lin, M. Maire, S. Belongie [et al.] // In European Conference on Computer Vision. – 2014. – P. 740 – 755.
8. The Pascal Visual Object Classes (VOC) challenge / M. Everingham, L. Van Gool, C. K. I. Williams [et al.] // International Journal of Computer Vision. – 2010. – №88 (2). – P. 303 – 338.
9. Adam: A method for stochastic optimization [Electronic resource] / D. P. Kingma, J. Ba // arXiv preprint. – 2014. – Access mode : <https://arxiv.org/abs/1412.6980>.

Стаття надійшла до редакції 26.08.2024.

Стаття пройшла рецензування 02.09.2024.

Голенко Максим Юрійович – аспірант кафедри інженерії програмного забезпечення, e-mail: holenko.maksym@gmail.com.

Державний Університет “Житомирська політехніка”.