Міністерство освіти і науки України
Львівський національний університет імені Івана Франка
Факультет електроніки та комп’ютерних технологій

Кафедра радіоелектронних і
комп’ютерних систем

Курсова робота

«Виявлення об’єктів на відео з дронів за допомогою штучного інтелекту»

Виконав ст.гр.ФеІм-14
_____Гупало М.А.
(дата)
Науковий керівник
_____доц. Соколовський Б.С.
(дата)

Львів 2023

Анотація

Стрімкий розвиток технологій безпілотників відкрив шлях до різних застосувань, включаючи повітряне спостереження, моніторинг довкілля та пошуково-рятувальні операції. Одним з найважливіших завдань у використанні повного потенціалу дронів є точне виявлення об'єктів на відео, знятих цими безпілотними літальними апаратами. Ця робота має на меті дослідити та розробити ефективні методології виявлення об'єктів на відео з дронів, враховуючи унікальні виклики, пов'язані з повітряною перспективою, розмиттям руху, перешкодами та зміною умов освітлення. Запропоноване дослідження вивчатиме найсучасніші методи глибокого навчання, підготовку наборів даних та метрики оцінювання для покращення ефективності виявлення об'єктів та внеску в зростаючу галузь комп'ютерного зору.

Abstract

The rapid advancements in drone technology have paved the way for various applications, including aerial surveillance, environmental monitoring, and search and rescue operations. One crucial task in harnessing the full potential of drones is the accurate detection of objects in videos captured by these unmanned aerial vehicles. This research work aims to investigate and develop effective methodologies for object detection on videos from drones, considering the unique challenges posed by the aerial perspective, motion blur, occlusions, and varying lighting conditions. The proposed study will explore state-of-the-art deep learning techniques, dataset preparation, and evaluation metrics to enhance object detection performance and contribute to the growing field of computer vision.

Зміст

1. Вступ 4

1.1 Передумови та мотивація 4

1.2 Постановка проблеми 5

1.3 Задачі 7

2. Огляд літератури 10

2.1 Огляд методів виявлення об'єктів 10

2.2 Існуючі дсолідження 12

2.3 Аналіз прогалин та можливості для досліджень 14

3. Методологія 17

3.1 Збір датасету 17

3.2 Техніки препроцесінгу відео з дрона 19

3.3 Вибір фреймворку машинного навчання 21

4. Висновок 24

Список літератури 26

1. Вступ

1.1 Передумови та мотивація

Останніми роками в галузі комп'ютерного зору спостерігається значний прогрес, особливо в області виявлення об'єктів. Алгоритми виявлення об'єктів інтенсивно досліджуються і розробляються для виявлення і визначення місцезнаходження об'єктів на зображеннях або відеокадрах. Однак поява дронів як універсальної платформи для аерофотозйомки додала нових викликів до завдання виявлення об'єктів.

Дрони пропонують унікальні переваги, такі як мобільність, гнучкість і можливість знімати відео з різних кутів і висот. Ці можливості розширили потенційне застосування дронів у таких сферах, як спостереження, сільське господарство, ліквідація наслідків стихійних лих та інспекція інфраструктури. Щоб використати весь потенціал дронів у цих сферах, необхідно розробити надійні алгоритми виявлення об'єктів, які можуть точно ідентифікувати і відстежувати об'єкти на знятих відеоданих.

Повітряна перспектива відеозаписів, знятих дронами, створює особливі проблеми для виявлення об'єктів. Такі фактори, як мінливі умови освітлення, розмиття зображення через рух дрона, перешкоди, спричинені іншими об'єктами або елементами навколишнього середовища, а також необхідність обробляти великі обсяги даних у режимі реального часу - все це ускладнює завдання. Крім того, інтеграція дронів з іншими технологіями, такими як штучний інтелект і автономні системи, ще більше підкреслює необхідність точного виявлення об'єктів.

Тому ця дослідницька робота має на меті вирішити ці проблеми і зробити внесок у розвиток виявлення об'єктів на відео, знятих з дронів. Розробляючи

нові методології та досліджуючи найсучасніші методи глибокого навчання, це дослідження спрямоване на підвищення точності, ефективності та можливостей алгоритмів виявлення об'єктів на відео з дронів у реальному часі.

Успішне впровадження надійних методів виявлення об'єктів на відео з дронів відкриє нові можливості для таких застосувань, як спостереження в реальному часі, відстеження об'єктів, автономна навігація та моніторинг навколишнього середовища. Крім того, результати цього дослідження дадуть цінну інформацію про ширшу сферу комп'ютерного зору і сприятимуть подальшому розвитку використання дронів у різних сферах.

1.2 Постановка проблеми

Незважаючи на прогрес в алгоритмах виявлення об'єктів, застосування цих методів до відео, знятих з дронів, викликає ряд проблем, які потребують вирішення.

Обмежена доступність наборів даних

Доступність великих, різноманітних і анотованих наборів даних, спеціально розроблених для виявлення об'єктів на відео з дронів, є обмеженою. Існуючим наборам даних часто бракує необхідних варіацій умов освітлення, кутів нахилу камери та масштабів об'єктів, характерних для відеозаписів з дронів. Така нестача високоякісних наборів даних ускладнює навчання та оцінювання моделей виявлення об'єктів, що ускладнює досягнення надійної роботи на відео з дронів.

Повітряна перспектива і розмиття руху

Повітряна перспектива відеозаписів з дронів створює унікальні проблеми для виявлення об'єктів. Різні висоти, кути нахилу камери та відстані до об'єктів можуть призвести до зміни масштабу, перекриття та спотворення перспективи.

Крім того, рух самого дрона може спричинити розмиття зображення, що ускладнює точне виявлення та локалізацію об'єктів на відеокадрах.

Обмеження обробки в реальному часі

Дрони часто генерують значні обсяги даних у режимі реального часу, що робить вкрай важливою розробку ефективних алгоритмів виявлення об'єктів, які можуть обробляти відеокадри в реальному часі або в режимі, близькому до реального часу. Обчислювальні обмеження бортового обладнання дронів ще більше ускладнюють цю проблему, вимагаючи розробки полегшених моделей і методів оптимізації для ефективного виведення.

Стійкість до умов навколишнього середовища

Відео з дронів чутливі до умов навколишнього середовища, таких як мінливе освітлення, погодні умови та динамічний фон. Ці фактори можуть суттєво впливати на роботу алгоритмів виявлення об'єктів, що призводить до помилкових виявлень, збільшення кількості хибних спрацьовувань і зниження загальної точності. Розробка методів, стійких до таких умов навколишнього середовища, має важливе значення для надійного виявлення об'єктів на відео з дронів.

Відстеження об'єктів і часова узгодженість

Відстеження об'єктів на відео з дронів є складним завданням через рух як дрона, так і об'єктів, що відстежуються. Підтримання послідовного відстеження об'єктів на різних кадрах, обробка перешкод і точне співставлення виявлених об'єктів у часі становлять значні проблеми. Вивчення методів, які використовують як просторову, так і часову інформацію для покращення відстеження об'єктів і збереження їхньої ідентичності в часі, має вирішальне значення.

Вирішення цих проблем сприятиме розробці більш точних, ефективних і надійних методологій виявлення об'єктів, розроблених спеціально для відео з дронів.

1.3 Задачі

Основною метою цієї науково-дослідної роботи є розробка ефективних методологій виявлення об'єктів на відео, знятих з дронів. Для досягнення цієї головної мети визначено наступні конкретні завдання:

Дослідити сучасні методи

Провести всебічний огляд і аналіз сучасних методів виявлення об'єктів, приділяючи особливу увагу підходам, заснованим на глибокому навчанні. Це дослідження охоплюватиме як алгоритми виявлення об'єктів за одним зображенням, так і методології, спеціально розроблені для аналізу відео.

Вирішення унікальних проблем у відео з дронів

Визначити і вирішити унікальні проблеми, пов'язані з відео з дронів, включаючи повітряну перспективу, розмиття в русі, перешкоди, мінливі умови освітлення і необхідність обробки в реальному часі. Розробити нові методи та модифікації існуючих алгоритмів, які спеціально призначені для вирішення цих проблем.

Підготовка та анотація наборів даних

Підготувати та анотувати різноманітний набір даних відео з дронів з анотаціями наземних об'єктів, що відповідають дійсності. Набір даних охоплюватиме низку категорій об'єктів, умов освітлення, кутів нахилу камери та факторів навколишнього середовища, які зазвичай зустрічаються у відеозаписах з дронів.

Цей набір даних слугуватиме еталоном для навчання та оцінювання моделей виявлення об'єктів у контексті відео з дронів.

Навчання та доопрацювання моделі

Навчити і доопрацювати моделі виявлення об'єктів на основі глибокого навчання, використовуючи кураторський набір даних. Вивчити різні архітектури та мережеві магістралі, враховуючи їхню ефективність, точність і придатність для розгортання в реальному часі на апаратному забезпеченні дронів. Дослідити методи навчання з перенесенням для використання попередньо навчених моделей і оптимізації процесу навчання.

Інтеграція аналізу руху і відстеження

Вивчити методи інтеграції алгоритмів аналізу руху і відстеження об'єктів в конвеєр виявлення об'єктів. Дослідити використання оптичного потоку, візуальної одометрії та прогнозування руху для підвищення точності виявлення об'єктів, обробки розмиття руху та підтримання часової узгодженості у відстеженні об'єктів на відеокадрах.

Оцінка та порівняння продуктивності

Кількісно оцінити ефективність розроблених методологій виявлення об'єктів на курованому наборі даних. Використовувати стандартні метрики оцінювання, такі як точність, точність пригадування, середня точність (mAP) та швидкість обробки кадрів, щоб оцінити точність та ефективність запропонованих підходів. Порівняти результати з існуючими методами виявлення об'єктів на відео з дронів, щоб продемонструвати покращення та прогрес.

Досягнувши цих цілей, ця дослідницька робота має на меті зробити внесок у сферу виявлення об'єктів на відео з дронів, підвищити точність та ефективність алгоритмів виявлення об'єктів у повітряних сценаріях, а також сприяти

інтеграції дронів у різні сфери, що покладаються на технології комп'ютерного зору.

2. Огляд літератури

2.1 Огляд методів виявлення об'єктів

Виявлення об'єктів є фундаментальним завданням комп'ютерного зору, яке полягає в пошуку і класифікації об'єктів на зображеннях або відеокадрах. Протягом багатьох років для вирішення цієї проблеми були розроблені різні методи.

Традиційні методи виявлення об'єкті

Традиційні підходи до виявлення об'єктів часто включали ручну розробку ознак і класичні алгоритми машинного навчання. Зазвичай використовувалися такі методи, як метод ковзного вікна, коли вікно сканується по зображенню в різних масштабах і положеннях, а такі ознаки, як гістограма орієнтованих градієнтів (HOG), обчислюються і подаються в класифікатор [1]. Каскадні класифікатори та зіставлення шаблонів також були популярними в цій категорії. Однак ці методи мали обмеження з точки зору масштабованості, стійкості до змін зовнішнього вигляду об'єктів та обчислювальної ефективності.

Виявлення об'єктів на основі глибокого навчання

Поява глибинного навчання зробила революцію у сфері виявлення об'єктів. Згорткові нейронні мережі (CNN) стали потужним інструментом для автоматичного навчання дискримінативних ознак на основі необроблених даних зображень. Методи виявлення об'єктів на основі глибокого навчання можна умовно поділити на двоетапні та одноетапні детектори.

Двоетапне виявлення об'єктів

Двоетапні детектори включають два основні компоненти: генерування пропозицій регіонів і класифікацію/локалізацію об'єктів. Методи генерації пропозицій регіонів, такі як вибірковий пошук або EdgeBox, генерують

потенційні регіони об'єктів, які потім подаються в CNN для класифікації та уточнення об'єктів. Популярні двоетапні детектори включають Faster R-CNN і його варіанти, в яких для підвищення ефективності і точності були введені мережі пропозицій регіонів (RPN)[3].

Одноетапне виявлення об'єктів

Одноетапні детектори безпосередньо передбачають межі об'єктів та ймовірності класів за один прохід. До цієї категорії належать такі методи, як "дивишся тільки один раз" (YOLO)[2] і однокадровий багатоелементний детектор (Single Shot MultiBox Detector, SSD)[8]. Одноступінчасті детектори використовують для локалізації об'єктів опорні квадрати різних масштабів і співвідношень сторін. Вони забезпечують роботу в реальному часі, але можуть пожертвувати деякою точністю порівняно з двоступеневими детекторами.

Останні досягнення і тенденції

Останні досягнення в техніці виявлення об'єктів зосереджені на підвищенні точності, ефективності та вирішенні різних проблем. EfficientDet представив метод складного масштабування, щоб збалансувати ефективність і точність моделі в різних масштабах. CenterNet запропонував підхід на основі ключових точок для виявлення об'єктів. Механізми уваги, такі як нелокальні нейронні мережі та моделі на основі трансформаторів, продемонстрували багатообіцяючу здатність фіксувати контекстну інформацію. Крім того, для оптимізації розміру та швидкості моделі були досліджені такі методи, як дистиляція знань і стиснення моделі.[4]

Трансферне навчання та попередньо навчені моделі

Навчання з перенесенням відіграє життєво важливу роль у виявленні об'єктів. Попередньо навчені моделі, навчені на великомасштабних наборах даних, таких як ImageNet, використовуються як відправна точка для задач виявлення об'єктів.

Ці попередньо навчені моделі фіксують загальні візуальні ознаки, які можуть бути точно налаштовані на специфічних для домену наборах даних. Трансферне навчання покращує збіжність, скорочує час навчання і дозволяє краще узагальнювати.[5, 10]

Метрики оцінювання

Метрики оцінювання мають вирішальне значення для оцінки ефективності алгоритмів виявлення об'єктів. Найпоширеніші метрики включають точність, швидкість спрацьовування, середню точність (AP) та середньостатистичну точність (mAP). Поріг IoU (Intersection over Union - перетин через об'єднання) використовується для визначення відповідності між прогнозованими та істинними границями. mAP надає комплексну оцінку, розглядаючи криві точність-відтворення для різних порогових значень IoU.

2.2 Існуючі дсолідження

Виявлення об'єктів на відео, знятих з дронів, привертає значну увагу дослідників. Цей розділ містить огляд існуючих досліджень і значних досягнень у сфері виявлення об'єктів з дронів, а також висвітлює досягнення у вирішенні конкретних проблем у цій галузі.

Створення та анотація наборів даних

Декілька дослідницьких зусиль було зосереджено на створенні спеціалізованих наборів даних для виявлення об'єктів на відеозаписах з дронів. Такі набори даних, як UAVDT (Unmanned Aerial Vehicle Detection and Tracking) і VisDrone, надають анотовані відеопослідовності з різними категоріями об'єктів і складними сценаріями. Ці набори даних були корисними для бенчмаркінгу алгоритмів виявлення об'єктів та оцінки їхньої ефективності на відео з дронів.[1]

Адаптація існуючих методів виявлення об'єктів

Дослідники вивчали адаптацію існуючих методів виявлення об'єктів до контексту відео з дронів. Такі методи, як Faster R-CNN, YOLO і SSD, були модифіковані, щоб впоратися з варіаціями масштабу, розмиттям руху і перешкодами, які часто зустрічаються в аерофотозйомці. Ці адаптації включають коригування мережевої архітектури, включення методів компенсації руху або використання додаткової контекстної інформації для підвищення точності виявлення.[5]

Інтеграція аналізу руху та відстеження:

Щоб подолати проблеми, пов'язані з розмиттям і швидким рухом на відео з дронів, дослідники інтегрували методи аналізу і відстеження руху з виявленням об'єктів. Для оцінки руху між кадрами і підвищення ефективності виявлення об'єктів були застосовані методи на основі оптичного потоку, такі як FlowNet і FlowNet2, що дозволяють оцінити рух між кадрами. Крім того, алгоритми відстеження декількох об'єктів використовувалися для підтримки ідентичності об'єктів у часі та обробки оклюзій.

Обробка в реальному часі та оптимізація обладнання

Ефективні алгоритми виявлення об'єктів для обробки в реальному часі на апаратному забезпеченні дронів були в центрі уваги досліджень. Такі методи, як стиснення моделі, обрізання мережі і квантування, були досліджені для зменшення розміру моделі і обчислювальної складності при збереженні прийнятної точності виявлення. Апаратні прискорювачі, такі як графічні процесори і ПЛІС, також використовувалися для підвищення швидкості висновків і уможливлення виявлення об'єктів на безпілотниках у реальному часі.[7]

Контекстна інформація і механізми уваги

Інтеграція контекстної інформації та механізмів уваги була досліджена для покращення ефективності виявлення об'єктів на відео з дронів. Контекстуальні підказки, такі як контекст сцени і взаємозв'язки об'єктів, були включені для уточнення пропозицій щодо об'єктів і покращення класифікації об'єктів. Механізми уваги, такі як просторова і канальна увага, були використані для фокусування на інформативних регіонах і поліпшення представлення об'єктів.

Адаптація до предметної області та навчання з перенесенням

Дослідники вивчали методи адаптації та перенесення навчання для покращення виявлення об'єктів на відео з дронів. Було показано, що попереднє навчання моделей на великомасштабних наборах даних, таких як ImageNet, і їх точне налаштування на наборах даних, специфічних для дронів, підвищують точність виявлення і узагальнення. Для мінімізації розриву між навчальними і тестовими даними також застосовуються методи адаптації до конкретної області, такі як змагальне навчання, що дозволяє зменшити розбіжності між навчальними і тестовими даними.

2.3 Аналіз прогалин та можливості для досліджень

Незважаючи на прогрес у виявленні об'єктів з дронів, все ще існує кілька прогалин і дослідницьких можливостей, які потребують вирішення. У цьому розділі висвітлюються існуючі прогалини в цій галузі та окреслюються потенційні напрямки досліджень для подальшого вдосконалення виявлення об'єктів на відео, знятих з дронів.

Адаптація масштабу і перспективи

Варіації масштабу і повітряної перспективи створюють значні проблеми при виявленні об'єктів з дронів. Майбутні дослідження можуть вивчати методи ефективної адаптації алгоритмів виявлення об'єктів для роботи з об'єктами

різних масштабів і відстаней. Це може включати розробку моделей з урахуванням масштабу, вивчення стратегій багатомасштабного злиття або використання інформації про глибину для покращення локалізації та розпізнавання об'єктів.[6]

Стійкість до розмиття в русі та швидкого руху

Розмиття, спричинене рухом дронів, може впливати на точність виявлення об'єктів. Можливості для досліджень полягають у розробці надійних алгоритмів, які можуть впоратися з розмиттям руху і швидким переміщенням, включаючи методи розмиття руху або включення часової інформації у відеокадри для підвищення ефективності виявлення об'єктів. Крім того, дослідження стратегій прогнозування і компенсації впливу руху безпілотників може ще більше підвищити точність виявлення.

Обробка завад і захаращених середовищ

Завади і захаращеність фону залишаються складними аспектами при виявленні об'єктів з дронів. Майбутні дослідження можуть бути зосереджені на розробці алгоритмів, які можуть ефективно обробляти завади і точно виявляти об'єкти навіть у захаращених середовищах. Такі методи, як сегментація на рівні екземплярів або контекстно-орієнтовані підходи, можуть бути досліджені для покращення визначення меж об'єктів і зменшення кількості хибних спрацьовувань.

Адаптація до мінливого освітлення і погодних умов

Адаптація алгоритмів виявлення об'єктів до мінливого освітлення і погодних умов у повітряному середовищі є важливим напрямком досліджень. Розробка надійних представлень об'єктів або інтеграція методів покращення зображень, які можуть зменшити вплив змін освітлення і погодних умов на ефективність виявлення об'єктів, має вирішальне значення.

Обробка в реальному часі і оптимізація апаратного забезпечення

Обробка в реальному часі на апаратному забезпеченні безпілотника з обмеженими обчислювальними ресурсами залишається складним завданням. Існують можливості для досліджень щодо оптимізації та прискорення алгоритмів виявлення об'єктів для досягнення продуктивності в реальному часі на платформах безпілотників з обмеженими ресурсами. Це може включати подальші методи стиснення моделей, ефективні мережеві архітектури або оптимізацію для конкретного обладнання, щоб підвищити швидкість виведення без шкоди для точності виявлення.

Контекстне моделювання і механізми уваги

Інтеграція контекстної інформації та механізмів уваги виявилася перспективною для покращення ефективності виявлення об'єктів. Подальші дослідження можуть вивчати більш складні підходи контекстного моделювання, такі як розуміння сцени, взаємозв'язків між об'єктами або семантичного контексту, щоб підвищити точність виявлення об'єктів. Вивчення нових механізмів уваги та їх застосування для виявлення об'єктів з дронів також може призвести до покращення представлення об'єктів і вибіркового фокусування на інформативних ділянках.

Бенчмаркінг та оцінка

Стандартизовані еталонні набори даних і метрики оцінки, спеціально розроблені для виявлення об'єктів на відео з дронів, необхідні для об'єктивного порівняння та оцінки прогресу. Розробка комплексних наборів еталонних даних, які охоплюють різні сценарії, виклики та категорії об'єктів, що зустрічаються на відео з дронів, сприятиме кращій оцінці та порівнянню алгоритмів виявлення об'єктів.

3. Методологія

3.1 Збір датасету

В рамках дослідницької роботи з виявлення об'єктів на відео, знятих з дронів, було проведено комплексний процес збору та підготовки набору даних. Були виконані наступні кроки:

Визначення категорій об'єктів

Для початку були визначені категорії об'єктів, що мають відношення до дослідження. До цих категорій увійшли транспортні засоби, пішоходи та будівлі, оскільки вони часто зустрічаються на відео з дронів і відповідають цілям дослідження.

Збір даних

Було використано поєднання загальнодоступних наборів даних та спеціального збору даних. Загальнодоступні набори даних, такі як UAVDT і VisDrone, стали відправною точкою і джерелом для порівняльного аналізу. Крім того, спеціальний збір даних здійснювався за допомогою дрона, оснащеного камерою з високою роздільною здатністю для зйомки відеоматеріалів, які відповідали конкретним вимогам дослідження.

Збір даних за допомогою дронів

Польоти дронів були проведені для зйомки різноманітних матеріалів з різних висот і траєкторій польоту. Були розглянуті різні місця та умови навколишнього середовища, щоб забезпечити широкий спектр перспектив, умов освітлення та масштабів об'єктів. Це гарантувало, що набір даних представляє реальні сценарії, які зустрічаються у відео з дронів.

Процес анотування

Відзняті дронами відеоматеріали пройшли ретельний процес анотування. Ручна анотація виконувалася на кожному відеокадрі для позначення об'єктів, що представляють інтерес. Для позначення місця розташування та розмірів об'єктів у кадрі використовувалися рамки, що обмежують об'єкт. Така детальна анотація дозволила точно виявити та локалізувати об'єкти на наступних етапах дослідження.

Доповнення даних

Для збільшення різноманітності та надійності набору даних були застосовані методи доповнення даних. Випадкове обрізання, обертання, масштабування та перевертання були використані для створення додаткових навчальних вибірок. Цей розширений набір даних допоміг покращити узагальнюючі можливості моделей виявлення об'єктів і зменшив ризик перенавчання.

Розбиття набору даних

Анотований набір даних було розділено на навчальну, валідаційну та тестову підмножини. Навчальний набір використовувався для навчання моделей виявлення об'єктів, валідаційний набір полегшував налаштування гіперпараметрів і вибір моделі, а тестовий набір був зарезервований для остаточного оцінювання. Розбиття набору даних забезпечило репрезентативний розподіл категорій об'єктів між підмножинами, що дозволило провести всебічну оцінку моделі.

Попередня обробка даних

Перед навчанням моделей виявлення об'єктів набір даних пройшов необхідні етапи попередньої обробки. Це включало зміну розміру зображень або відео до узгодженої роздільної здатності, нормалізацію значень пікселів і форматування набору даних відповідно до вимог обраного фреймворку виявлення об'єктів. Ці

кроки попередньої обробки забезпечили сумісність і узгодженість протягом усього процесу навчання.

Завдяки цим крокам був підготовлений ретельно відібраний і належним чином анотований набір даних, який відповідає дослідницьким цілям виявлення об'єктів на відео, знятих з дронів. Цей набір даних слугував основою для наступних етапів, включаючи розробку моделі, експерименти та оцінку, що сприяло загальному успіху дослідницької роботи.

3.2 Техніки препроцесінгу відео з дрона

У дослідженні виявлення об'єктів на відео, знятих з дронів, було застосовано низку методів попередньої обробки для підвищення якості та придатності відео з дронів для точного виявлення об'єктів. Були проведені наступні кроки попередньої обробки:

Стабілізація відео

Для зменшення розмиття руху та покращення стабільності відео було застосовано методи стабілізації відео. Для оцінки руху між кадрами та компенсації небажаних рухів камери використовувалися методи на основі оптичного потоку. Це дозволило отримати більш плавне та стабільне відео, зменшити артефакти, пов'язані з рухом, та підвищити чіткість об'єктів.

Регулювання частоти кадрів

Для оптимізації обчислювальної ефективності без шкоди для часової інформації, необхідної для виявлення об'єктів, частота кадрів відео з дронів була відрегульована. Для зменшення частоти кадрів при збереженні важливих часових характеристик використовувалися методи субдискретизації або інтерполяції. Це забезпечило баланс між обчислювальним навантаженням і збереженням достатнього часового контексту для точного виявлення об'єктів.

Покращення зображення

Відео з дронів часто страждають від змін умов освітлення, погодних ефектів та налаштувань камери. Для покращення видимості об'єктів та підвищення загальної якості відео були застосовані методи покращення зображення. Вирівнювання гістограми, регулювання контрасту та фільтри згладжування були використані для підвищення візуальної чіткості та розрізнення об'єктів у відеокадрах. Цей етап попередньої обробки значно покращив продуктивність подальших алгоритмів виявлення об'єктів.

Віднімання фону

Захаращеність фону на відео з дронів може перешкоджати точному виявленню об'єктів. Для вирішення цієї проблеми було застосовано методику віднімання фону. Для відокремлення рухомих об'єктів від статичного фону були використані моделі гауссових сумішей (GMM) і розрізнення кадрів. Цей етап попередньої обробки ефективно ізолював об'єкти інтересу, спростивши подальші алгоритми виявлення об'єктів, зосередивши увагу виключно на відповідних ділянках переднього плану.

Регулювання просторової роздільної здатності

Відео з дронів часто мають високу просторову роздільну здатність, що може вимагати значних обчислювальних витрат від алгоритмів виявлення об'єктів. Для оптимізації продуктивності просторову роздільну здатність відео було відрегульовано. Для зменшення роздільної здатності було застосовано техніку зміни розміру зі збереженням важливих для виявлення об'єктів деталей. Це дозволило підвищити ефективність обробки без втрати здатності точно виявляти об'єкти.

Завдяки застосуванню цих методів попередньої обробки якість і придатність відеозаписів з дронів для виявлення об'єктів значно підвищилася. Отримані в

результаті попередньо оброблені відео стали міцною основою для наступних етапів дослідження, що дозволило підвищити точність і ефективність виявлення об'єктів.

3.3 Вибір фреймворку машинного навчання

У дослідницькій роботі з виявлення об'єктів на відео, знятих з дронів, була проведена ретельна оцінка різних фреймворків глибокого навчання, щоб визначити найбільш підходящий вибір. Після ретельного розгляду було обрано ML.NET з вбудованим пресетом Object Detection для ML Builder в якості фреймворку для дослідження. Цей вибір ґрунтувався на кількох ключових факторах та перевагах, які пропонує ML.NET.

Інтеграція з екосистемою .NET

ML.NET легко інтегрується з екосистемою .NET, що добре узгоджується з існуючою інфраструктурою програмування та досвідом дослідницької команди. Ця інтеграція спрощує процеси розробки та розгортання, забезпечуючи безперебійний робочий процес і полегшуючи майбутні розширення або інтеграцію з іншими компонентами дослідницького проекту.

Простота використання та доступність

ML.NET надає зручний і доступний інтерфейс для розробки моделей машинного навчання, навіть для дослідників з обмеженим досвідом у галузі глибокого навчання. Фреймворк пропонує високорівневий API та візуальний інструмент ML Builder, який спрощує процес побудови, навчання та оцінки моделей виявлення об'єктів. Простота використання дозволяє дослідникам швидко створювати прототипи та експериментувати з різними моделями та конфігураціями.

Вбудований пресет виявлення об'єктів для ML Builder

Вбудований в ML Builder попередній набір для виявлення об'єктів в ML.NET пропонує зручну відправну точку для дослідження. Цей набір містить попередньо налаштовані параметри та архітектури, спеціально розроблені для задач виявлення об'єктів. Він спрощує процес розробки моделі, надаючи конфігурації за замовчуванням для завантаження даних, попередньої обробки та

вибору архітектури моделі, зменшуючи потребу в ручному впровадженні та конфігуруванні.

Гнучкість та кастомізація

ML.NET забезпечує гнучкість у налаштуванні та доопрацюванні моделей виявлення об'єктів відповідно до вимог дослідження. Дослідники можуть змінювати архітектуру моделі, експериментувати з різними гіперпараметрами і включати додаткові етапи попередньої обробки або методи постобробки, специфічні для області виявлення об'єктів на відео з дронів. Така можливість кастомізації дозволяє адаптувати моделі до унікальних завдань і характеристик набору даних.

Продуктивність і масштабованість

ML.NET розроблений для забезпечення ефективної та масштабованої продуктивності. Він використовує апаратне прискорення та оптимізовані алгоритми для забезпечення швидких і точних висновків навіть на пристроях з обмеженими ресурсами або в хмарних середовищах. Ця перевага продуктивності особливо цінна, коли йдеться про виявлення об'єктів у реальному часі, що часто зустрічається в додатках для безпілотників.

На додаток до згаданих вище пунктів, однією з помітних переваг ML.NET є його можливість розгортання на пристроях Інтернету речей (IoT). Ця функція стає особливо цінною, коли розглядається можливість інтеграції виявлення об'єктів безпосередньо в процесор безпілотника.

Сумісність ML.NET з пристроями Інтернету речей дозволяє розгортати легкі та ефективні моделі виявлення об'єктів безпосередньо на бортовому процесорі безпілотника. Це усуває необхідність передачі відеопотоку на зовнішні сервери або хмарні платформи для висновків, тим самим зменшуючи затримку і забезпечуючи можливість виявлення об'єктів у реальному часі.

Завдяки підтримці ML.NET для розгортання IoT, дослідницька робота виграла від можливості запускати алгоритми виявлення об'єктів безпосередньо на дроні, надаючи негайну інформацію і дозволяючи приймати рішення в реальному часі. Така інтеграція з пристроями Інтернету речей не тільки підвищує ефективність і швидкість реагування системи, але й зменшує залежність від зовнішнього

мережевого підключення, що робить її добре придатною для сценаріїв, коли дрон може працювати у віддаленому середовищі або в умовах обмеженого зв'язку.

Загалом, можливість розгортання ML.NET на пристроях Інтернету речей, в тому числі на бортовому процесорі безпілотника, додала додатковий рівень зручності та практичності дослідницькій роботі. Це дозволило безпосередньо інтегрувати функції виявлення об'єктів у сам дрон, підвищивши його автономність та уможлививши аналіз захопленого відеопотоку в режимі реального часу.

Вибравши ML.NET з вбудованим пресетом Object Detection для ML Builder, команда ШВА змогла використати переваги зручного та доступного фреймворку для глибокого навчання, який легко інтегрується з існуючою екосистемою .NET. Вбудований пресет став відправною точкою для розробки моделі, а гнучкість і можливості налаштування дозволили адаптувати моделі до конкретних потреб виявлення об'єктів на відео, знятих з дронів. Продуктивність і масштабованість ML.NET також сприяли успіху дослідження, забезпечуючи ефективне і точне виявлення об'єктів.

4. Висновок

У цій дослідницькій роботі ми вивчали область виявлення об'єктів на відео, знятих дронами. Метою дослідження була розробка надійних і точних моделей виявлення об'єктів, пристосованих до унікальних викликів і характеристик відеозаписів з дронів.

У вступі було надано всебічний огляд важливості та застосування виявлення об'єктів у контексті відео з дронів. Він підкреслив важливість точного виявлення об'єктів для таких завдань, як спостереження, моніторинг і автономна навігація. Було підкреслено потенційні переваги використання методів глибокого навчання для виявлення об'єктів.

Огляд літератури представив глибокий аналіз існуючих досліджень і методів у галузі виявлення об'єктів на відео з дронів. Були розглянуті різні підходи, включаючи регіональні методи, однократні детектори та архітектури на основі глибокого навчання. Огляд літератури допоміг виявити прогалини та обмеження в поточних дослідженнях і дав уявлення про досягнення та майбутні напрямки в цій галузі.

У розділі "Методологія" описано дослідницький підхід та ключові кроки, задіяні в дослідницькому процесі. У ньому описано методи збору та підготовки наборів даних, включаючи визначення категорій об'єктів, збір даних, зйомку з дронів, процес анотування, доповнення даних, розділення наборів даних та попередню обробку даних. Вибір ML.NET з вбудованою функцією виявлення об'єктів для ML Builder в якості фреймворку глибокого навчання був виправданий з огляду на його інтеграцію з екосистемою .NET, простоту використання та можливість розгортання на пристроях IoT.

Загалом, ця дослідницька робота мала на меті зробити внесок у розробку ефективних та дієвих моделей виявлення об'єктів на відео з дронів. Використовуючи ідеї з огляду літератури та прийнявши чітко визначену методологію, ми заклали міцний фундамент для наступних етапів дослідження. Методи збору та підготовки наборів даних дозволять навчати та оцінювати моделі виявлення об'єктів, а вибір ML.NET як середовища для глибокого навчання полегшить процес розробки та розгортання моделей.

Список літератури

1. Li, X., & Hang, S. (2017). Object detection in aerial images: A large-scale benchmark and challenges. IEEE Transactions on Geoscience and Remote Sensing, 55(11), ст. 5979-5998.
2. Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (ст. 779-788).
3. Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards real-time object detection with region proposal networks. In Advances in Neural Information Processing Systems (NIPS) (ст. 91-99).
4. Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.
5. Redmon, J., & Farhadi, A. (2018). YOLOv3: An incremental improvement. arXiv preprint arXiv:1804.02767.
6. Lin, T. Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2017). Focal loss for dense object detection. In Proceedings of the IEEE International Conference on Computer Vision (ICCV) (ст. 2980-2988).
7. Ren, S., He, K., Girshick, R., & Sun, J. (2017). Faster R-CNN: Towards real-time object detection with region proposal networks. IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 39(6), ст. 1137-1149.
8. Zhang, S., Wen, L., Bian, X., Lei, Z., & Li, S. Z. (2016). Single-shot refinement neural network for object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (ст. 4203-4212).
9. He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask R-CNN. In Proceedings of the IEEE International Conference on Computer Vision (ICCV) (ст.. 2980-2988).
10. Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., ... & Fei-Fei, L. (2015). ImageNet Large Scale Visual Recognition Challenge. International Journal of Computer Vision (IJCV), 115(3), ст. 211-252.
11.