

О. М. КОНДРАТОВ, аспірант, старший викладач кафедри інформаційних систем та технологій Національного технічного університету «Харківський політехнічний інститут», Харків, Україна; e-mail: kondratovolexiy@gmail.com; ORCID: <https://orcid.org/0000-0001-6367-9944>

О. М. НИКУЛІНА, д-р техн. наук, професор, завідувачка кафедри інформаційних систем та технологій Національного технічного університету «Харківський політехнічний інститут», Харків, Україна; професор кафедри цифрових технологій та проектно-аналітичних рішень ТОВ «ТУ «МЕТІНВЕСТ ПОЛІТЕХНІКА»», Запоріжжя, Україна; e-mail: elniknik02@gmail.com; ORCID: <https://orcid.org/0000-0003-2938-4215>

ПРОГРАМНА РЕАЛІЗАЦІЯ ІЗ ВИКОРИСТАННЯМ ТРАНСФОРМЕРА З ОПТИЧНИМ ПОТОКОМ ТА GEONET ДЛЯ ІДЕНТИФІКАЦІЇ ПАРАМЕТРІВ ДИНАМІЧНИХ ОБ'ЄКТІВ

Сьогодні міждисциплінарні дослідження в галузі комп'ютерних наук та інженерії стають все більш актуальними через зростаючу потребу обробляти дані в реальному часі для виявлення та відстеження об'єктів. Зокрема, ідентифікація параметрів динамічних об'єктів є критично важливою у таких сферах, як автономні транспортні системи, робототехніка, системи спостереження і моніторингу. Ефективне автоматичне отримання та обробка інформації з відеоджерел є перспективною галуззю для науковців і практиків, що працюють у суміжних областях. Метою цього дослідження є вдосконалення процесів виявлення та відстеження динамічних об'єктів шляхом розробки та впровадження інформаційної технології, заснованої на використанні сучасних методів машинного навчання, таких як DETR (Detection Transformer), оптичний потік та GeoNet. Методи дослідження включають розробку програмного забезпечення на основі мови програмування Python із застосуванням сучасних бібліотек та фреймворків для обробки зображень і відео. Для виявлення об'єктів було використано метод DETR, що дозволяє з високою точністю визначати положення об'єктів у кадрі. Оптичний потік застосовувався для визначення напрямку та швидкості їх переміщення, а GeoNet — для аналізу глибини сцени та геометричних параметрів. Запропонована технологія була протестована на різних відеозаписах, які відображають складні сценарії з динамічними умовами, включаючи зміну освітлення, перекриття об'єктів та швидкі зміни руху. Результати дослідження показали високу точність і надійність запропонованого підходу для ідентифікації параметрів динамічних об'єктів у різних умовах. Поєднання методів дозволило суттєво підвищити точність і стійкість системи виявлення та відстеження, особливо в умовах зміни середовища або низької якості відеозаписів. Висновки дослідження вказують на ефективність використання запропонованої інформаційної технології для практичного застосування у сфері автономних систем, робототехніки та відеоспостереження.

Ключові слова: Дистанційна ідентифікація динамічних об'єктів, виявлення об'єктів, оптичний потік, ідентифікація швидкості, глибоке навчання, згорткові нейронні мережі.

Вступ. Сьогодні швидка еволюція комп'ютерних технологій змінює наукові підходи до сучасних викликів і відкриває нові шляхи для створення нових методів або вдосконалення існуючих. Це особливо важливо для прикладних наук, де обчислювальні методи прискорюють розв'язання складних проблем, що вимагають обробки великих обсягів даних [1].

У галузі комп'ютерного зору досягнення в машинному навчанні та виявленні об'єктів революціонізували спосіб відстеження та аналізу динамічних об'єктів [2].

Одним з ключових досягнень є використання трансформера виявлення (DETR) для виявлення об'єктів та оптичного потоку (Optical Flow) для відстеження руху у відео послідовностях. Комбінація цих методів дозволяє точно ідентифікувати параметри динамічних об'єктів у різних середовищах. Спочатку моделі виявлення об'єктів мали обмежений спектр можливостей. Однак завдяки зростанню доступності відеоданих високої роздільної здатності та досягненням у глибокому навчанні моделі ідентифікації параметрів динамічних об'єктів значно еволюціонували, що дозволяє точно відстежувати і аналізувати об'єкти в реальному часі. Це розширило використання таких моделей у різних додатках, від автономних систем до спостереження, демонструючи критичну роль обчислювальних технологій у вирішенні сучасних проблем. Подальший розвиток комп'ютерних технологій, зокрема збільшення

обчислювальної потужності та зростання машинного навчання, відкрив нові напрями для науковців та інженерів у розширенні своїх методологій. Досягнення в області комп'ютерного зору призвели до створення моделей виявлення об'єктів, таких як DETR, яка дозволяє здійснювати виявлення об'єктів у реальному часі в відеопослідовностях, та Оптичний потік, який використовується для відстеження руху між кадрами. У динамічних середовищах точно ідентифікувати параметри об'єктів, такі як позиція, швидкість і траєкторія, є критично важливим для додатків, починаючи від автономних транспортних засобів і до спостереження. Характеристики реальних даних, такі як великий обсяг, швидкість і різноманіття, створюють труднощі для моделей ідентифікації параметрів динамічних об'єктів, оскільки їм необхідно швидко та точно обробляти великі обсяги відеоданих [3–4].

Сьогодні важливо використовувати моделі виявлення та відстеження об'єктів, які можуть адаптуватися до змінних умов середовища та працювати в реальному часі. Для ідентифікації параметрів динамічних об'єктів цей підхід є важливим для визначення таких характеристик, як розмір об'єкта, швидкість і напрямок. Дослідники в області комп'ютерного зору погоджуються, що інтеграція DETR та оптичного потоку відкриває нові можливості для міждисциплінарної співпраці між комп'ютерними науковцями та інженерами.

© Кондратов О. М., Нікуліна О. М., 2024



Дослідницька стаття: Цю статтю опубліковано видавництвом НТУ «ХПИ» у збірнику «Вісник Національного технічного університету "ХПИ" Серія: Системний аналіз, управління та інформаційні технології». Ця стаття поширюється за міжнародною ліцензією [Creative Common Attribution \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/). Конфлікт інтересів: Автор/и заявив/или про відсутність конфлікту.



Ключовим моментом для моделей виявлення об'єктів є їх здатність залишатися гнучкими та адаптованими до постійно змінюваних умов середовища. Багато галузей прагнуть впроваджувати такі технології для автоматизації та передової аналітики. Однак слід зазначити, що різні додатки можуть мати різні вимоги та цілі. В таких умовах цікаво дослідити, як різноманітні динамічні середовища впливають на ефективність моделей виявлення об'єктів.

Це дослідження має на меті відповісти на такі наукові питання.

1. Як можна покращити точність виявлення динамічних об'єктів у відео з використанням сучасних моделей машинного навчання? Це питання стосується ефективності різних моделей, таких як DETR для виявлення об'єктів, та аналізу їх застосування в реальних умовах, зокрема при обробці відео-потoku.

2. Яка роль оптичного потоку у відстеженні об'єктів і як він може допомогти в розпізнаванні руху в реальному часі? Вивчення того, як оптичний потік допомагає в аналізі руху, і як його можна використовувати для визначення напрямку та швидкості переміщення об'єктів в реальному часі, є важливим аспектом.

3. Як ефективно інтегрувати методи глибокого навчання, такі як DETR і GeoNet, для покращення точності та надійності систем виявлення і відстеження об'єктів у динамічних сценах?

4. Важливо розглянути, як поєднання DETR для виявлення об'єктів та GeoNet для глибокого навчання може підвищити точність виявлення та відстеження, зокрема в складних умовах, таких як зміни освітлення або перекриття об'єктів.

5. Які фактори впливають на ефективність моделей в реальних умовах, зокрема на точність виявлення та відстеження в умовах змінюваного оточення та швидкості? Вивчення того, як різні фактори, такі як зміна освітлення, туман, або швидкість руху об'єктів, можуть впливати на продуктивність моделей виявлення та відстеження, є важливою частиною цього дослідження.

Ця стаття представляє комплексний підхід до аналізу динамічних об'єктів у відеопотоках, поєднуючи DETR, оптичний потік та GeoNet для глибокої оцінки параметрів об'єктів, таких як положення, швидкість, напрямок руху та глибина. Розроблений конвеєр включає завантаження відео, обробку кадрів, виявлення об'єктів, аналіз руху та оцінку глибини. Система продемонструвала високу ефективність у різних сценаріях, включаючи складні фони та динамічні середовища [4–8].

Мета та задачі дослідження. Мета статті полягає в використанні методів ідентифікації параметрів динамічного об'єкту з використанням DETR та Optical Flow для побудови моделі в технології DETR.

Для досягнення мети поставлені задачі дослідження.

1. Описати аналітичну та реалізовану модель.
2. Описати реалізовану обробку відео-потoku.
3. Описати використання моделі DETR.
4. Описати використання оптичного потоку.

5. Описати розроблену модель з використанням GeoNet.

6. Аналіз результатів.

Аналітична модель. Схема ідентифікації параметрів динамічного об'єкту з використанням DETR та Optical Flow показана на рис. 1 і включає елементи. Обробка кадрів відеопотоку. Кожен кадр (1 кадр, 2 кадр і т. д.) аналізується з використанням Відео вхід. Детектор (DETR), що дозволяє визначити положення та межі об'єктів у кадрі. Застосування алгоритму оптичного потоку для визначення векторів руху об'єктів між послідовними кадрами. Аналіз параметрів об'єктів. На основі отриманих векторів руху проводиться аналіз параметрів об'єктів, включаючи швидкість, напрямок руху, траєкторію та інші характеристики [3, 6, 8, 9].

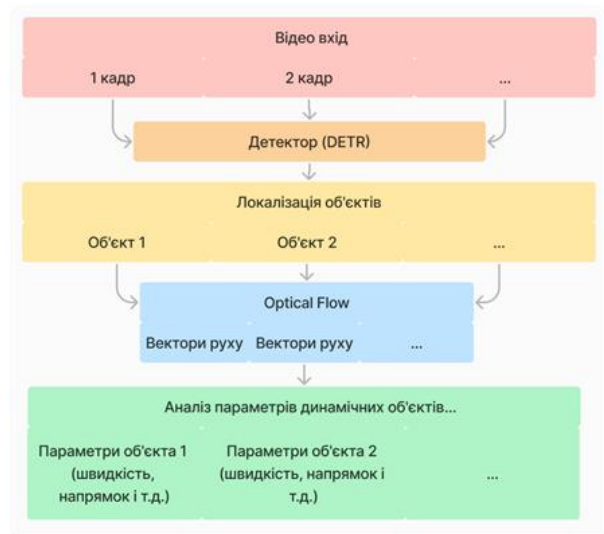


Рис. 1. Аналітична модель

Реалізована модель. На рис. 2, представлено реалізовану модель, яка включає обробку відео потоку, використання моделі DETR, використання оптичного потоку, формування і відображення результатів. Колорами відмічено різні елементи реалізованої моделі.



Рис. 2. Реалізована модель

Розширена реалізована модель. На рис. 3 представлено розширений опис реалізованої моделі. Яка включає наступні елементи у вигляді конвеєру:



Рис. 3. Розширений опис реалізованої моделі

1. Обробку відео потоку. Збереження відео файлу з відеопотоку. Екстракція кадрів з відеофайлу. Отримання кадрів з відео файлу для подальшої їх обробки.

2. Модель DETR. Кожен кадр обробляється моделлю для виявлення об'єктів, тобто виділення обмежувальних рамок об'єктів знайдених на кадрах. Отримання кадрів з рамками.

3. Обчислення оптичного потоку. Визначення векторів руху. Аналіз швидкості та напрямку. Отримання візуалізації руху.

Обробка відео-потоку. У розробленому застосунку здійснюється обробка відео-потоку з використанням моделі DETR для виявлення об'єктів та обчислення оптичного потоку між послідовними кадрами. Відео спочатку завантажується з YouTube (якщо файл не існує на диску), після чого аналізується кадр за кадром. Для кожного кадру проводиться виявлення об'єктів, відображення коробок навколо об'єктів та обчислення оптичного потоку для оцінки руху об'єктів. Застосунок використовує моделі на базі PyTorch і TensorFlow для максимізації продуктивності на GPU або TPU (Google Colab).

Опис використання моделі DETR. Модель DETR (Detection Transformer) використовується для виявлення об'єктів у кожному кадрі відео. Вона отримує зображення, попередньо оброблене через `DetrImageProcessor`, і проводить детекцію за допомогою трансформера. Виходи моделі включають логіти та координати предиктованих коробок для кожного виявленого об'єкта. Детекція здійснюється з використанням `softmax` для прогнозування класів об'єктів і їх точних координат. Виявлені об'єкти накладаються на оригінальне зображення через функцію `draw_boxes`, що дозволяє візуалізувати результати.

Використовується декодер, який має схожу структуру до кодера з деякими додатковими компонентами. Список шарів включає багатоголову самоувагу, суму та нормування, багатоголову увагу, суму та нормування, нейронну мережу прямого поширення, як і в кодері використовується багатозаровий перцептрон [3, 10].

Опис використання оптичного потоку. Оптичний потік обчислюється між двома зображеннями,

`prev_frame_gray` та `next_frame_gray`, за допомогою методу `Farneback`, який використовує параметри для покращення точності виявлення руху. Параметри включають використання пірамідного масштабу (0.5), трьох рівнів піраміди, вікно 15x15 пікселів для розрахунку потоку, три ітерації уточнення, поліноміальне розширення з розміром сусідства 5 пікселів, а також Гауссовий фільтр з відхиленням 1.2. Ці налаштування дозволяють отримати більш точні результати для аналізу руху об'єктів у відео, що важливо для відстеження динамічних об'єктів.

На рис. 4, представлено результат роботи використання оптичного потоку.

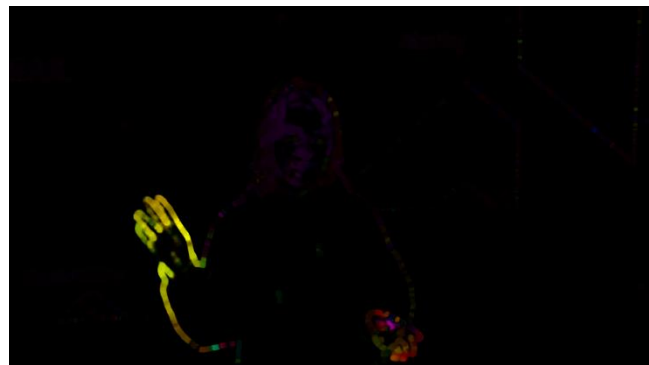


Рис. 4. Результат роботи використання оптичного потоку

Опис додаткової розробленої моделі з використанням GeoNet. Реалізована інтеграція GeoNet для обробки відео шляхом оцінки глибини зображень (`monocular depth estimation`). Що ідентифікує, ще один із найважливіших параметрів об'єкта, а саме глибину.

GeoNet працює із відеофраємами, які попередньо витягуються з відео, зберігаються у вигляді зображень та формуються для подальшої обробки. Входи обробляються через GeoNet, що передбачає отримання параметрів.

Етап підготовки даних. GeoNet здійснює попередню обробку кадрів відео з використанням скрипта. Цей етап включає зміну розмірів зображень, створення необхідних послідовностей кадрів (наприклад, для аналізу оптичного потоку або оцінки глибини) і

видалення статичних об'єктів для покращення результатів.

Інференс. Основний обчислювальний етап виконується за допомогою `geonet_main.py`, який аналізує дані та генерує результати, такі як глибинні карти, оптичний потік або оцінки позиції. Для цього використовуються попередньо натреновані моделі, збережені у вигляді контрольних точок (checkpoints).

Вивід результатів. Глибинні карти (depth maps) зберігаються у визначеному каталозі та візуалізуються для оцінки результатів роботи моделі.

На рис. 5, представлено реалізована модель з використанням GeoNet.

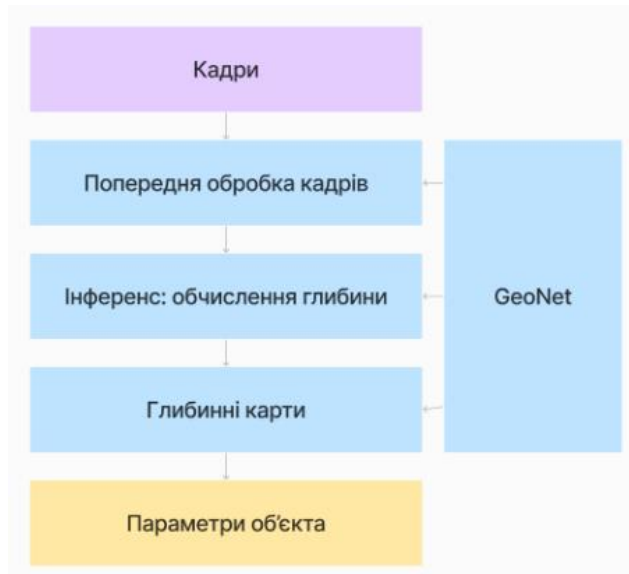


Рис. 5. Реалізована модель з використанням GeoNet

На рис. 6 представлено загальну реалізовану модель з використанням GeoNet. Яка реалізована для паралельного обчислення з DETR та оптичним потоком.

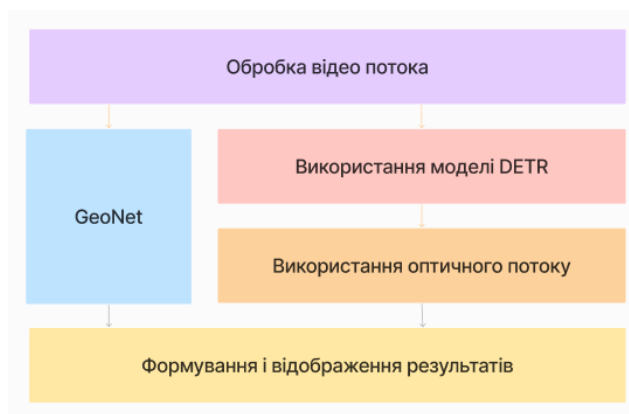


Рис. 6. Загальна реалізована модель з використанням GeoNet

Оцінка глибини (Depth Estimation). Результати роботи GeoNet продемонстрували можливість точного визначення глибини на основі кадрів відео. Згенеровані глибинні карти надають детальну інформацію про структуру сцени, що корисно для відстеження об'єктів у просторі та побудови 3D-моделей [8, 11–12].

Підготовка та обробка даних. Процес підготовки даних для GeoNet включає збереження кадрів із відео, їх форматування та видалення статичних об'єктів. Це дозволяє зменшити вплив шуму та покращити якість оцінки глибини.

Аналіз результатів. У результаті виконання обробки відео з використанням DETR, модель успішно виявляє об'єкти на кожному кадрі, і середній час обробки для одного кадру може бути вимірний, що дозволяє оцінити ефективність моделі.

В обчисленні оптичного потоку за допомогою методу Farneback на основі двох послідовних кадрів отримано точні дані про напрямок та швидкість руху об'єктів, що дозволяє покращити відстеження візуальних змін у кадрі.

Для кожного кадру відео вимірюється час обробки для детекції об'єктів та для обчислення оптичного потоку, що дозволяє оцінити ефективність кожного етапу обробки та оптимізувати продуктивність системи.

Інтеграція з іншими методами. У поєднанні з моделлю DETR результати глибинних карт використані для покращення розпізнавання об'єктів, а також для визначення їх положення та руху в просторі, що робить систему потужним інструментом для аналізу відео-потоків.

Для покращення реалізованої моделі та майбутнього представлення у вигляді цілісної системи доцільним є використання методів бустінгу та бегінгу. Бустінг дозволяє послідовно комбінувати кілька моделей, кожна з яких коригує помилки попередньої, що забезпечує високу точність виявлення та відстеження об'єктів. Застосування цього підходу може значно покращити продуктивність системи в умовах складних середовищ, таких як зміна освітлення або перекриття об'єктів. У поєднанні з DETR та оптичним потоком бустінг здатний підвищити ефективність аналізу динамічних об'єктів у відеопотоках.

Бегінг, у свою чергу, забезпечує зниження варіативності моделей шляхом паралельного навчання декількох базових моделей на різних підмножинах даних. Це сприяє підвищенню стійкості системи до шумів у вхідних даних та підвищує загальну надійність. Інтеграція цих підходів із застосуванням мета-алгоритму дозволяє створити ансамбль, який об'єднує результати всіх компонентів у єдину систему. Така архітектура забезпечить оптимальну взаємодію між DETR, GeoNet, оптичним потоком та іншими компонентами, що значно покращить ефективність і адаптивність системи в реальних сценаріях [13].

Висновки. Динамічні об'єкти є важливим елементом в багатьох сферах, таких як автономні автомобілі, робототехніка, моніторинг безпеки та медична діагностика. Ідентифікація їх параметрів, таких як швидкість, напрямок руху та інші відіграє ключову роль у розумінні та прогнозуванні їх поведінки.

Використання згорткових мереж для виявлення ознак з відеопослідовностей, інтеграція технік трансформерів для покращення точності виявлення та ідентифікації об'єктів, оптимізація методів навчання для обробки великих обсягів даних та динамічних змін у відеопослідовностях – і елементи допомагають у ство-

ренні ефективної системи для ідентифікації параметрів динамічних об'єктів з використанням переваг як згорткових нейронних мереж, так і трансформерів та оптичного потоку. Використання моделі, яка поєднує у собі DETR та метод Optical Flow для ефективної ідентифікації параметрів динамічного об'єкту. DETR є потужним інструментом для точного виявлення об'єктів на зображеннях.

Комп'ютерні технології суттєво змінюють підходи до вирішення сучасних проблем, зокрема в галузі комп'ютерного зору. Інтеграція методів, таких як DETR для виявлення об'єктів, оптичний потік для аналізу руху та GeoNet для оцінки глибини, відкриває нові можливості для точного відстеження динамічних об'єктів у реальному часі. Це дозволяє визначити такі параметри, як швидкість, траєкторія та глибина, що є критично важливим для застосувань у автономних транспортних системах, відеоспостереженні та інших галузях. Завдяки обчислювальним потужностям GPU і TPU реалізовані методи демонструють високу ефективність у складних середовищах.

Особливу увагу заслуговує аналітична та реалізована модель обробки відео, яка включає етапи екстракції кадрів, виявлення об'єктів за допомогою DETR і аналіз руху через оптичний потік. Використання попередньо натренованих моделей дозволяє отримати точні результати навіть у складних умовах, таких як зміни освітлення чи перекриття об'єктів.

Інтеграція глибинних карт GeoNet значно покращує здатність ідентифікації об'єктів у просторі, що є важливим для 3D-реконструкції сцен.

Усі ці методи були оптимізовані для роботи в режимі реального часу, що дозволяє використовувати їх на мобільних пристроях і в системах із обмеженими ресурсами.

Оцінка ефективності реалізованої системи базується на метриках, таких як час обробки кадрів, точність виявлення та стабільність результатів. Інтеграція DETR і GeoNet забезпечує високу адаптивність до змінних умов середовища, що робить ці методи універсальними для різних застосувань.

Подальший розвиток цього підходу сприятиме підвищенню точності та швидкості обробки, що відкриває перспективи для ширшого впровадження в міждисциплінарні проекти. Загалом, дослідження демонструє важливість інноваційних обчислювальних методів у вирішенні складних задач аналізу відеопотоків.

Реалізована модель, яка об'єднує сучасні алгоритми DETR, оптичний потік та GeoNet, продемонструвала високу ефективність для ідентифікації параметрів динамічних об'єктів у складних сценаріях. Інтеграція цих методів дозволила суттєво покращити точність, швидкість і стійкість обробки відеоданих, зокрема в умовах змінного освітлення, шуму та перекриття об'єктів. Геометрична інформація, отримана за допомогою GeoNet, забезпечила додатковий рівень деталізації, зокрема аналіз глибини сцени, що значно розширило можливості системи.

Результати дослідження підкреслюють перспективність міждисциплінарного підходу до розробки сис-

тем ідентифікації динамічних об'єктів. Застосування поєднаних методик має великий потенціал для практичного використання у таких сферах, як автономні транспортні системи, відеоспостереження та робототехніка. Подальший розвиток і оптимізація цих підходів сприятиме створенню більш досконалих рішень для аналізу відеоданих у реальному часі.

Список використаної літератури

1. Wang Z., Turko R., Shaikh O., Park H., Das N., Hohman F., Kahng M., Chau D. *CNN Explainer: Learning Convolutional Neural Networks with Interactive Visualization*. URL: <https://arxiv.org/abs/2004.15004> (дата звернення: 06.11.2024).
2. Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A., Kaiser L., Polosukhin I. *Attention Is All You Need*. URL: <https://arxiv.org/abs/1706.03762> (дата звернення: 06.11.2024).
3. Carion N., Massa F., Synnaeve G., Usunier N., Kirillov A., Zagoruyko S. *End-to-End Object Detection with Transformers*. URL: <https://arxiv.org/abs/2005.12872v3> (дата звернення: 06.11.2024).
4. Zou Z., Chen K., Shi Z., Shi Z., Guo Y., Ye J. *Object Detection in 20 Years: A Survey*. URL: https://arxiv.org/pdf/1905.05055.pdf?fbclid=IwAR0ILGAWTtWU-9-iH6lZyPFXyXA5JRWarM_XoSJ78QEhmm-txvr_iGEzCio (дата звернення: 06.11.2024).
5. Ammar A., Chebbah A., Fredj H., Souani C. *Comparative Study of latest CNN based Optical Flow Estimation*. URL: <https://ieeexplore.ieee.org/document/9806070/references#references>. (дата звернення: 06.11.2024).
6. Yin Z., Shi J. *GeoNet: Unsupervised Learning of Dense Depth, Optical Flow and Camera Pose*. URL: <https://arxiv.org/abs/1803.02276v2> (дата звернення: 06.11.2024).
7. Girshick R., Donahue J., Darrell T., and Malik J. Region-based convolutional networks for accurate object detection and segmentation. *IEEE transactions on pattern analysis and machine intelligence*. 2016. Vol. 38, no. 1. P. 142–158.
8. Нікуліна О. М., Северин В. П., Кондратов О. М., Ольховий О. М. Моделі дистанційної ідентифікації параметрів динамічних об'єктів з використанням трансформерів виявлення та оптичного потоку. *Вісник НТУ «ХПІ»*. Серія: Системний аналіз, управління та інформаційні технології. Харків. НТУ «ХПІ», 2024. № 1 (11). С. 52–57.
9. Нікуліна О. М., Кондратов О. М. Модель ідентифікації параметрів динамічного об'єкту з використанням DETection TRansformer та Optical Flow. *Інформаційні технології: наука, техніка, технологія, освіта, здоров'я: Тези доповідей XXXII міжнародної науково-практичної конференції MicroCAD-2024, 22-24 травня 2024 р.* Харків, НТУ «ХПІ». 2024. С. 1047.
10. Нікуліна О. М., Кондратов О. М. Методи дистанційної ідентифікації динамічних параметрів об'єкта. *Інформаційні технології: наука, техніка, технологія, освіта, здоров'я: Тези доповідей XXXI міжнародної науково-практичної конференції MicroCAD-2023, 17-20 травня 2023 р.* Харків, НТУ «ХПІ». 2023. С. 1047.
11. Gracyk A., Chen X. *GeONet: a neural operator for learning the Wasserstein geodesic*. URL: <https://arxiv.org/abs/2209.14440> (дата звернення: 06.11.2024).
12. Inomata T., Kimura K., Hagiwara M. *Object Tracking and Classification System Using Agent Search*. URL: https://www.jstage.jst.go.jp/article/ieejagents/129/11/129_11_2065/_pdf/-char/ja (дата звернення: 06.11.2024).
13. Gavrylenko S., Chelak V., Hornostal O. Construction Method Of Fuzzy Decision Trees For Identification The Computer System State. *2022 XXXII International Scientific Symposium Metrology and Metrology Assurance (MMA)*. 2022. P. 1–5.

References (transliterated)

1. Wang Z., Turko R., Shaikh O., Park H., Das N., Hohman F., Kahng M., Chau D. *CNN Explainer: Learning Convolutional Neural Networks with Interactive Visualization*. URL: <https://arxiv.org/abs/2004.15004> (accessed: 06.11.2024).
2. Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A., Kaiser L., Polosukhin I. *Attention Is All You Need*. URL: <https://arxiv.org/abs/1706.03762> (accessed: 06.11.2024).

3. Carion N., Massa F., Synnaeve G., Usunier N., Kirillov A., Zagoruyko S. *End-to-End Object Detection with Transformers*. URL: <https://arxiv.org/abs/2005.12872v3> (accessed: 06.11.2024).
4. Zou Z., Chen K., Shi Z., Shi Z., Guo Y., Ye J. *Object Detection in 20 Years: A Survey*. URL: https://arxiv.org/pdf/1905.05055.pdf?fbclid=IwAR0ILGAWTtWU-9-iH6lZyPFXYXA5JRWarM_XoSJ78QEhmnn-txvr_iGEzCio (accessed: 06.11.2024).
5. Ammar A., Chebbah A., Fredj H., Souani C. *Comparative Study of latest CNN based Optical Flow Estimation*. URL: <https://ieeexplore.ieee.org/document/9806070/references#references>. (accessed: 06.11.2024).
6. Yin Z., Shi J. *GeoNet: Unsupervised Learning of Dense Depth, Optical Flow and Camera Pose*. URL: <https://arxiv.org/abs/1803.02276v2> (accessed: 06.11.2024).
7. Girshick R., Donahue J., Darrell T., and Malik J. Region-based convolutional networks for accurate object detection and segmentation. *IEEE transactions on pattern analysis and machine intelligence*. 2016, vol. 38, no. 1, pp. 142–158.
8. Nikulina O. M., Severin V. P., Kondratov O. M., Olhovoy O. M. Models of remote identification of parameters of dynamic objects using detection transformers and optical flow. *Vestnik Nats. tekhn. un-ta "KhPI": sb. nauch. tr. Temat. vyp.: Sistemnyy analiz, upravlenie i informatsionnye tekhnologii [Bulletin of the National Technical University "KhPI": a collection of scientific papers. Thematic issue: System analysis, management and information technology]*. Kharkiv, NTU "KhPI" Publ., no. 1 (11), pp. 52–57.
9. Nikulina O. M., Kondratov O. M. Dynamic object parameter identification model using DETection TRansformer and Optical Flow. *Information technologies: science, technology, technology, education, health: Abstracts of reports of XXII international scientific and practical conference MicroCAD-2024, May 22–24, 2024, Kharkiv, NTU "KhPI"*, 2024, p. 1047.
10. Nikulina O. M., Kondratov O. M. Methods of remote identification of dynamic object parameters. *Information technologies: science, technology, technology, education, health: Abstracts of reports of the XXI international scientific and practical conference MicroCAD-2023, May 17–20, 2023, Kharkiv, NTU "KhPI"*, 2023, p. 1047.
11. Gracyk A., Chen X. *GeONet: a neural operator for learning the Wasserstein geodesic*. URL: <https://arxiv.org/abs/2209.14440> (accessed: 06.11.2024).
12. Inomata T., Kimura K., Hagiwara M. *Object Tracking and Classification System Using Agent Search*. URL: https://www.jstage.jst.go.jp/article/ieejieiss/129/11/129_11_2065/_pdf/-char/ja (accessed: 06.11.2024).
13. Gavrylenko S., Chelak V., Hornostal O. Construction Method Of Fuzzy Decision Trees For Identification The Computer System State. *2022 XXXII International Scientific Symposium Metrology and Metrology Assurance (MMA)*. 2022, pp. 1–5.

Надійшла (received) 08.11.2024

UDC 004.8+004.9

O. M. KONDRATOV, Postgraduate, senior lecturer of Department Information Systems and Technologies National Technical University «Kharkiv Polytechnic Institute», Kharkiv, Ukraine; e-mail: kondratovolexiy@gmail.com; ORCID: <https://orcid.org/0000-0001-6367-9944>

O. M. NIKULINA, Doctor of Technical Sciences, Professor, Head of Department Information Systems and Technologies National Technical University «Kharkiv Polytechnic Institute», Kharkiv, Ukraine; Professor of Department IT, Analysis and Project Decisions, Technical University "METINVEST POLYTECHNICS", LLC, Zaporizhzhia, Ukraine; e-mail: elniknik02@gmail.com; ORCID: <https://orcid.org/0000-0003-2938-4215>

SOFTWARE IMPLEMENTATION USING TRANSFORMER WITH OPTICAL FLOW AND GEONET FOR IDENTIFYING PARAMETERS OF DYNAMIC OBJECTS

Today, interdisciplinary research in computer science and engineering has become increasingly relevant due to the growing demand for real-time data processing in object detection and tracking applications. The identification of dynamic object parameters plays a crucial role in various domains such as autonomous transportation systems, robotics, and surveillance. Effective automated acquisition and processing of video data represent a promising field for scientists and practitioners working in these interconnected disciplines. This research aims to enhance object detection and tracking processes by developing and implementing an information technology solution based on modern machine learning methods, including DETR (Detection Transformer), Optical Flow, and GeoNet. The research methodology involves designing software using Python programming language and modern libraries and frameworks for image and video processing. The DETR method was employed for precise object detection within video frames. Optical Flow was used to determine the direction and velocity of object movement, and GeoNet provided depth and geometric scene analysis. The proposed technology was tested on diverse video recordings depicting complex scenarios with dynamic conditions, such as varying lighting, object occlusions, and rapid motion changes. The results demonstrate the high accuracy and reliability of the proposed approach for identifying dynamic object parameters under various conditions. The integration of these methods significantly improved the precision and robustness of the detection and tracking system, particularly in challenging environments or low-quality video scenarios. The study concludes that the proposed information technology is effective and can be applied in practical fields such as autonomous systems, robotics, and video surveillance.

Keywords: Remote identification of dynamic objects, object detection, optical flow, velocity identification, deep learning, convolutional neural networks.

Повні імена авторів / Author's full names

Автор 1 / Author 1: Кондратов Олексій Михайлович, Kondratov Oleksii Mikhailovich

Автор 2 / Author 2: Нікуліна Олена Миколаївна, Nikulina Olena Mykolaivna