DOI:

УДК 519.2

© Нікуліна О. М., Северин В. П., Кондратов О. М., Ольховий О. М., 2024

Дослідницька стаття: Цю статтю опубліковано видавництвом НТУ «ХПІ» у збірнику «Вісник Національного технічного університету "ХПІ" 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.



О. М. Нікуліна, д-р техн. наук, професор, завідувачка кафедри інформаційних систем та технологій Національного технічного університету «Харківський політехнічний інститут», Харків, Україна; ORCID: https://orcid.org/0000-0003-2938-4215; e-mail: elniknik02@gmail.com

В. П. Северин, д-р техн. наук, професор, професор кафедри системного аналізу та інформаційно-аналітичних технологій Національного технічного університету «Харківський політехнічний інститут», Харків, Україна; ORCID: https://orcid.org/0000-0002-2969-6780; e-mail: valerii.severyn@khpi.edu.ua

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

О. М. Ольховий, старший викладач кафедри інформаційних систем та технологій Національного технічного університету «Харківський політехнічний інститут», Харків, Україна; ORCID: https://orcid.org/0009-0004-6409-2706, e-mail: aleksey.olhovoy@khpi.edu.ua

Модель та методи ідентифікації параметрів динамічних об’єктів з використанням трансформера виявлення та оптичного потоку

Завдання ідентифікації параметрів динамічних об’єктів є важливим для різноманітних галузей, включаючи робототехніку, автономні транспортні засоби, системи відеоспостереження та багато інших. Традиційні методи часто стикаються з проблемами недостатньої точності та ефективності в умовах швидко змінюваних середовищ та складних динамічних сценаріїв. Розглядаються сучасні методи ідентифікації параметрів динамічних об’єктів, зокрема з використанням технологій трансформера виявлення (ТРВИ) або Detection Transformer (DETR) та оптичного потоку (Optical Flow). Трансформер виявлення є одним з новітніх підходів у галузі комп’ютерного зору, що використовує архітектуру трансформера для завдань об’єктного детектування. Цей трансформер інтегрує процеси виявлення об’єктів і визначення їхніх меж у єдину модель енд-ту-енд, що значно покращує точність та швидкість обробки. Використання трансформерів дозволяє моделі ефективно обробляти інформацію з усього зображення одночасно, що сприяє кращому розпізнаванню об’єктів навіть у складних умовах. Оптичний потік є методом аналізу руху, що визначає швидкість та напрямок руху пікселів між послідовними кадрами відео. Цей метод дозволяє отримати детальну інформацію про динаміку сцени, що є критично важливим для точного відстеження та ідентифікації параметрів рухомих об’єктів. Пропонується інтеграція ТРВИ та оптичного потоку для підвищення точності ідентифікації параметрів динамічних об’єктів. Комбінація цих двох методів дозволяє використовувати переваги обох підходів: високу точність детектування об’єктів та детальну інформацію про їхній рух. Проведені експерименти показують, що запропонована модель значно перевершує традиційні методи як у точності визначення параметрів об’єктів, так і у швидкості обробки даних. Ключові результати дослідження свідчать про те, що інтеграція ТРВИ та оптичного потоку забезпечує надійне і швидке визначення параметрів рухомих об’єктів у реальному часі, що може бути застосовано у різних практичних сценаріях. Дослідження також показало потенціал для подальшого вдосконалення методів обробки даних та їхнього застосування у складних динамічних середовищах. Отримані результати відкривають нові перспективи для розробки інтелектуальних систем моніторингу та керування, що здатні адаптуватися до швидкозмінних умов навколишнього середовища, підвищуючи ефективність та безпеку їхньої роботи.

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

Вступ. Одними з ключових задач штучного інтелекту є виявлення та ідентифікація динамічних об’єктів у великих обсягах цифрових даних, що зазнають постійних змін. Ці задачі стають особливо актуальними у таких галузях як комп’ютерний зір, відеоспостереження, автономне керування транспортними засобами та багато інших.

Для розв’язання цих задач використовуються передові методи глибокого навчання, серед яких особливе місце займають засоби детектування та ідентифікації, такі як згорткові нейронні мережі (ЗНМ, CNN), кодувальні та декодувальні архітектури, а також нейронні мережі прямого поширення (НМПП, FFN). Використання цих методів є ключовим для ефективного виявлення та ідентифікації об’єктів у реальному часі, зокрема динамічних об’єктів, що зазнають змін у формі, розмірі, положенні та інших параметрах [1].

У статті розглянуті підходи до використання методів детектування на основі CNN, а також інших складових глибокого навчання, таких як кодувальні та декодувальні архітектури у контексті дистанційної ідентифікації динамічних об’єктів (ДІДО) у зображеннях. Також досліджені переваги та обмеження цих методів та надані висновки щодо їхнього потенціалу для розв’язання актуальних задач в різних областях застосування [2].

З вибором технології трансформера виявлення (ТРВИ) або Detection Transformer (DETR) є кілька причин, чому цей підхід привабливий для ДІДО у зображеннях [3].

Трансформерна архітектура: DETR базується на трансформерних архітектурах, які виявилися дуже ефективними в обробці послідовностей даних, таких як мовний текст. Використання трансформерів у відображенні зображень дозволяє моделі здійснювати увагу до глобальних та локальних залежностей у зображенні, що може бути корисним при ідентифікації об’єктів.

Підхід End-to-End: DETR пропонує архітектуру, яка дозволяє здійснювати виявлення об’єктів та їхню класифікацію у вигляді однієї моделі без необхідності в окремих компонентах, таких як anchor boxes або non-maximum suppression. Це спрощує процес навчання та інференсу, зменшуючи кількість гіперпараметрів та оптимізуючи швидкість роботи моделі [3, 4].

Технологія DETR результативно використана в різних областях, включаючи відеоаналітику, медичні зображення та інші. Це свідчить про універсальність використаної моделі та її потенціал у різних завданнях, пов’язаних з ідентифікацією об’єктів [5].

Отже, обираючи технологію DETR для ДІДО у зображеннях, можна розраховувати на високу точність та ефективність роботи моделі, яка здатна адаптуватися до різних сценаріїв та умов [5–7].

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. Проаналізувати модель технології DETR.

2. Розглянути технологію DETR.

3. Розглянути структуру кодера та декодера самоуваги в технології DETR.

4. Розглянути механізми уваги та самоуваги в технології DETR.

Модель в технології ТРВИ. Схема ідентифікації параметрів динамічного об’єкту з використанням DETR та Optical Flow показана на рис. 1 і включає такі елементи.

1) Обробка кадрів відеопотоку. Кожен кадр (1 кадр, 2 кадр і т.п.) аналізується з використанням Відео Вхід.

2) Детектор (DETR), що дозволяє точно визначати положення та межі об’єктів у кадрі.

3) Використання Оптичного Потоку. Застосування алгоритмів оптичного потоку для визначення векторів руху об’єктів між послідовними кадрами. Це дає можливість оцінити швидкість та напрямок руху об’єктів.

4) Аналіз параметрів об’єктів. На основі отриманих векторів руху проводиться аналіз параметрів об’єктів, включаючи швидкість, напрямок руху, траєкторію та інші динамічні характеристики.

На рис. 2, в, представлено розроблену модель [8]

 (1)

де  –  кадр;

 – індекс кадру;

 – набір обмежуючих рамок для кожного кадру.

Кожний кадр аналізується з використанням відео входу, детектору DETR, що дозволяє точно визначати положення та межі об’єктів у кадрі. Модель DETR приймає вхідне зображення і генерує набір обмежуючих рамок для кожного кадру

, (2)

де  та  – послідовні кадри;

 – оптичний потік.

Застосування алгоритмів оптичного потоку для визначення векторів руху об’єктів між послідовними кадрами. Оптичний потік визначається, як поле векторів, що описує рух кожної точки кадру.

, (3)

де  – кількість точок, що належать об’єкту;

 – вектори руху;

 – швидкість;

 – напрямок руху.

На основі отриманих векторів руху проводиться аналіз параметрів об’єктів, включаючи швидкість та напрямок руху. Швидкість об’єкта розраховується, як величина середнього вектора руху.

, (4)

де  – напрямок руху об’єкта.

Напрямок руху об’єкта визначається, як середнє значення напрямків векторів руху.

Завдяки застосуванню технології DETR, вдалося значно підвищити точність виявлення об’єктів у відеопотоці, особливо в умовах складного фону та динамічного середовища. Використання оптичного потоку дозволило точно визначати параметри руху об’єктів, що є критично важливим для задач реального часу, таких як системи безпеки, автономні транспортні засоби та відеоспостереження [3, 6, 9].

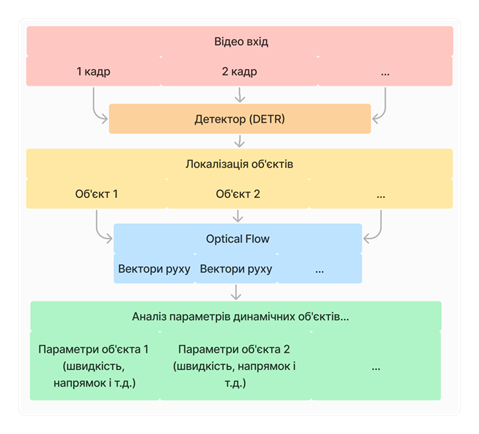


Рис. 1. Ідентифікації параметрів динамічного об’єкту з використанням DETR та Optical Flow.

Технологія ТРВИ. На рис. 2, а, представлено вхідний шар для зображень, який застосовується для виявлення просторових ознак. На рис. 2, б, представлено кодер, який складається з декількох шарів самоуваги та нормалізації, що перетворюють вхідні ознаки у високорівневі представлення. На рис. 2, в, представлено декодер використовує багатоголову увагу для передбачення виходів на основі представлень, отриманих від кодера.

Класична нейронна мережа з прямим поширенням, яка обробляє вихід декодера для отримання кінцевих результатів.

Структура кодера та декодера. На рис. 3, а, представлено кодер, який складається з шарів. Кожен шар має два основні підблоки – багатоголову самоувагу (Multi-Head Self-Attention) та НМПП, де НМПП відповідає нормалізації та прямому поширення.

На рис. 3, б, представлено декодер, який має схожу структуру до кодера з деякими додатковими компонентами. Список шарів включає багатоголову самоувагу, суму та нормування, багатоголову увагу, суму та нормування, НМПП, як і в кодері використовується багатошаровий перцептрон [3, 10].

|  |
| --- |
|  |
| Рис. 2. ТРВИ: а – ЗНМ (CNN); б – Кодер (Encoder); в – Декодер (Decoder); г – НМПП (FFN) |

Згорткова нейрона мережа використовується для екстракції ознак з вхідного зображення.

Етап попередньої обробки зображення, де ЗНМ витягує відповідні просторові ознаки [1, 5].

Кодер – це частина трансформера, яка приймає на вхід ознаки, що отримані ЗНМ, та обробляє їх за допомогою механізмів самоуваги. Завдання кодера полягає в кодуванні вхідних ознак у контекстні представлення, які зберігають інформацію про всі об’єкти в сцені. Декодер приймає на вхід закодовані ознаки з кодера та використовує механізми уваги для передбачення набору фіксованих об’єктів. Виходом декодера є кінцеві координати та категорії виявлених об’єктів.

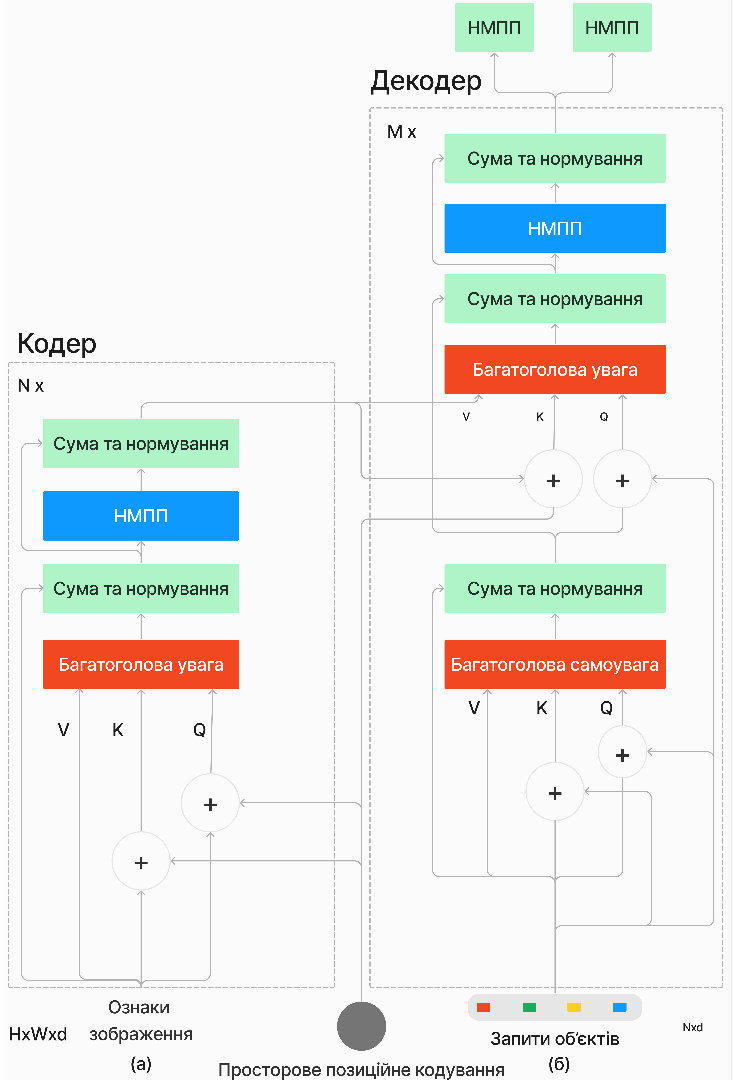
Багатоголова увага та самоувага у трансформерів. Багатоголова увага у трансформерів (Multi-Head attention). Трансформери мають мережеву архітектуру, що базується на механізмах уваги для машинного перекладу. Маючи елемент запиту (наприклад, цільове слово у вихідному реченні) і набір ключових елементів (наприклад, вихідні слова у вхідному речення), багатоголовий модуль уваги адаптивно агрегує ключовий вміст відповідно до вагами уваги, які вимірюють сумісність пар запит-ключ. Щоб дозволити моделі фокусуватися на вмісті з різних підпросторів представлення і різних позицій, виходи різних голов (heads) уваги лінійно агрегуються з перенавчальними важелями. Цей механізм забезпечує трансформерам здатність вловлювати складні залежності між елементами вхідних даних, що робить їх ефективними для широкого спектру задач обробки природної мови та інших областей.

Рис. 3. Модель Кодер – Декодер: а) Кодер; б) Декодер

Функція багатоголової уваги обчислюється за формулою [11]

(5)

де  – індексує елемент запиту;

 – ознака представлення запиту;

 – індексує елемент ключа;

 – ознака представлення ключа;

 – розмірність ознаки;

 та  – множини елементів запиту та ключа відповідно;

 – індекс голови уваги;

 –  вхідна проекція матриці;

 –  вихідна проекція матриці.

Вхідна та вихідна проекція матриці мають перенавчальні важелі

, (6)

де  –  розмірність уваги для кожної голови уваги;

 –  кількість голів уваги;

, (7)

де – важелі уваги;

,та – індекси голови, запиту та ключа відповідно.

Важелі уваги нормалізуються за формулою

, (8)

де  – перенавчальні важелі, вхідні запитів та ключів проекції матриці на ‑ту голову;

 та  – ознаки представлення, зазвичай є конкатенацією/сумою вмісту елементів та позиційних вкладок для розрізнення різних просторових позицій.

, (9)

де  –  вхідний вектор або матриця вхідних векторів;

 – матриця ваг для перетворення в запити;

 –  матриця ваг для перетворення в ключі;

 – матриця ваг для перетворення в значення;

 –  вектор запиту, отриманий з вхідного вектору або матриці вхідних векторів;

 – вектор ключа, отриманий з вхідного вектору або матриці вхідних векторів;

 –  вектор значення, отриманий з вхідного вектору або матриці вхідних векторів;

Створені вектори запитів, ключів і значень використовуються в механізмі уваги, щоб обчислити ваги уваги і зважити відповідні значення

, (10)

де  –  розмірність векторів ключів.

 – матриця ключів;

 – вектор-стовпець матриці ключів;

 – матриця значень;

 – вектор-стовпець матриці значень;

Механізм уваги призначає ваги ключам на основі їх схожості з запитом. Ці ваги використовуються для агрегування векторів значень у зважену суму, яка потім використовується в моделі трансформера. Оцінки уваги дозволяють моделі фокусуватися на різних частинах вхідної послідовності для кожної вихідної позиції, що дозволяє захоплювати залежності незалежно від їхньої відстані у послідовності. Увага цього типу застосовується для взаємодії між різними частинами послідовностей в трансформерних моделях. Вона розбиває вхідні дані на кілька голів, кожна з яких використовується для відповідних обчислень. Кожна голова вивчає взаємозв’язки між словами у різних контекстах та відповідності між різними частинами послідовності.

Результати обчислень кожної голови конкатенуються та проходять через додатковий проектний шар перед об’єднанням для подальшої обробки.

Багатоголова самоувага у трансформерів (Multi-Head Self-Attention) – це спеціалізований тип багатоголової уваги, де вхідні дані подаються одночасно на всі голови для обробки.

Кожна голова вивчає відносини між різними словами в межах цієї послідовності, тобто самоподібності.

Цей підхід дозволяє моделі вивчати взаємозв’язки між різними частинами послідовності без потреби у внутрішніх чи зовнішніх взаємодіях [2, 11].

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 [10, 12].

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 створенні ефективної системи для ідентифікації параметрів динамічних об’єктів з використанням переваг як згорткових нейронних мереж, так і трансформерів та оптичного потоку [13].

Використання моделі, яка поєднує у собі DETR та метод Optical Flow для ефективної ідентифікації параметрів динамічного об’єкту [2, 3, 5]. DETR є потужним інструментом для точного виявлення об’єктів на зображеннях. Використано цю модель для локалізації параметрів динамічних об’єктів на кадрах відео.

Запропоновані модель та методи будуть використані для ідентифікації параметрів динамічних об’єктів, таких як автономні автомобілі, роботи, безпілотні літальні апарати, а також для моніторингу безпеки та медичної діагностики.

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: 06.06.2024)..

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: 06.06.2024).

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: 06.06.2024).

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=IwAR0ILGAWTwU-9-iH6lZyPFXYXA5JRWarM\_XoSJ78QEhmnn-txvr\_iGEzCio (дата звернення: 06.06.2024).

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: 06.06.2024).

6. Girshick R., Donahue J., Darrell T., Malik J. Rich feature hierarchies for accurate object detection and semantic segmentation. URL: https://arxiv.org/abs/1311.2524 (дата звернення: 06.06.2024).

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. , Рекова Н. Ю. Аналіз інформаційних технологій для дистанційної ідентифікації динамічних об’єктів. Вісник НТУ «ХПІ». Серія: Системний аналіз, управління та інформаційні технології. Харків. НТУ «ХПІ», 2023. № 1 (9). С. 110–115.

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 ХXХІІ 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 ХXХІ міжнародної науково-практичної конференції MicroCAD-2023, 17-20 травня 2023 р. Харків, НТУ «ХПІ». 2023. С. 1047.

11. Zhu X., Hu H., Lin S., Dai J. Deformable ConvNets v2: More Deformable, Better Results. URL: https://arxiv.org/abs/1811.11168 (дата звернення: 06.06.2024).

12. Inomata T., Kimura K., Hagiwara M. Object Tracking and Classification System Using Agent Search. URL: https://www.jstage.jst.go.jp/article/ieejeiss/129/11/129\_11\_2065/\_pdf/-char/ja (дата звернення: 06.06.2024).

13. Krizhevsky A., Sutskever I., Hinton G. E. Imagenet classiﬁcation with deep convolutional neural networks. Advances in neural information processing systems. 2012. P. 1097‑1105.

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: 06.06.2024).

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: 06.06.2024).

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: 06.06.2024).

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=IwAR0ILGAWTwU-9-iH6lZyPFXYXA5JRWarM\_XoSJ78QEhmnn-txvr\_iGEzCio (accessed: 06.06.2024).

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: 06.06.2024).

6. Girshick R., Donahue J., Darrell T., Malik J. Rich feature hierarchies for accurate object detection and semantic segmentation. URL: https://arxiv.org/abs/1311.2524 (accessed: 06.06.2024).

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. М, Rekova N. Y. Analysis of information technologies for remote identification of dynamic objects*.* 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 (9), pp. 110–115.

9. Nikulina O. M., Kondratov O. М. 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. М. 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. Zhu X., Hu H., Lin S., Dai J. Deformable ConvNets v2: More Deformable, Better Results. URL: https://arxiv.org/abs/1811.11168 (accessed: 06.06.2024).

12. Inomata T., Kimura K., Hagiwara M. Object Tracking and Classification System Using Agent Search. URL: https://www.jstage.jst.go.jp/article/ieejeiss/129/11/129\_11\_2065/\_pdf/-char/ja (accessed: 06.06.2024).

13. Krizhevsky A., Sutskever I., Hinton G. E. Imagenet classiﬁcation with deep convolutional neural networks. Advances in neural information processing systems. 2012, pp. 1097‑1105.

Надійшла (received) 06.06.2024

UDC 519.2

О. М. Nikulina, Doctor of Technical Sciences, Associate Professor, Head of Department Information Systems and Technologies National Technical University «Kharkiv Polytechnic Institute», Kharkiv, Ukraine; ORCID: https://orcid.org/0000-0003-2938-4215; e-mail: elniknik02@gmail.com

V. P. Severyn, Doctor of Technical Sciences, Professor, Professor of Department System Analysis and Information-Analytical Technologies National Technical University «Kharkiv Polytechnic Institute», Kharkiv, Ukraine; ORCID: https://orcid.org/0000-0002-2969-6780; e-mail: valerii.severyn@khpi.edu.ua

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

O. M. Olhovoy, Senior lecturer of Department Information Systems and Technologies National Technical University «Kharkiv Polytechnic Institute», Kharkiv, Ukraine; ORCID: https://orcid.org/0009-0004-6409-2706; e-mail: aleksey.olhovoy@khpi.edu.ua

Model and methods for identifying dynamic object parameters using detection transformer and optical flow

This paper examines modern methods for identifying the parameters of dynamic objects, particularly using Detection Transformer (DETR) and Optical Flow technologies. The task of identifying the parameters of dynamic objects is crucial for various fields, including robotics, autonomous vehicles, surveillance systems, and many others. Traditional methods often face challenges of insufficient accuracy and efficiency in rapidly changing environments and complex dynamic scenarios. Detection Transformer (DETR) is one of the latest approaches in computer vision that uses a transformer architecture for object detection tasks. DETR integrates the processes of object detection and boundary determination into a single end-to-end model, significantly improving accuracy and processing speed. The use of transformers allows the model to effectively process information from the entire image simultaneously, facilitating better object recognition even in complex conditions. Optical Flow is a motion analysis method that determines the speed and direction of pixel movement between consecutive video frames. This method provides detailed information about the scene dynamics, which is critical for accurately tracking and identifying the parameters of moving objects. In our study, we propose integrating Detection Transformer and Optical Flow to enhance the accuracy of dynamic object parameter identification. The combination of these two methods leverages the advantages of both approaches: high precision in object detection and detailed information about their movement. Experiments conducted show that the proposed model significantly outperforms traditional methods in both the accuracy of parameter determination and data processing speed. Key research findings indicate that integrating DETR and Optical Flow provides reliable and fast parameter identification of moving objects in real time, applicable in various practical scenarios. The study also revealed the potential for further refinement of data processing methods and their application in complex dynamic environments. The obtained results open new perspectives for developing intelligent monitoring and control systems capable of adapting to rapidly changing environmental conditions, enhancing the efficiency and safety of their operation.

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: Нікуліна Олена Миколаївна, Nikulina Olena Mykolaivna

Автор 2 / Author 2: Северин Валерій Петрович, Severyn Valerii Petrovych

Автор 3 / Author 3: Кондратов Олексій Михайлович, Kondratov Oleksii Mikhailovich

Автор 4 / Author 4: Ольховий Олексій Михайлович, Olhovoy Oleksii Mikhailovich