

**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 проєктно-аналітичних рішень ТОВ «ТУ «МЕТІНВЕСТ ПОЛІТЕХНІКА»», Запоріжжя, Україна; ORCID: <https://orcid.org/0000-0003-2938-4215>; e-mail: [elniknik02@gmail.com](mailto: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](mailto: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: [kondratovolexy@gmail.com](mailto:kondratovolexy@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](mailto:aleksey.olhovoy@khpi.edu.ua)

### МОДЕЛІ ДИСТАНЦІЙНОЇ ІДЕНТИФІКАЦІЇ ПАРАМЕТРІВ ДИНАМІЧНИХ ОБ'ЄКТІВ З ВИКОРИСТАННЯМ ТРАНСФОРМЕРІВ ВИЯВЛЕННЯ ТА ОПТИЧНОГО ПОТОКУ

Задачі дистанційної ідентифікації параметрів динамічних об'єктів є важливими для різноманітних галузей, включаючи комп'ютерний зір, робототехніку, автономні транспортні засоби, системи відеоспостереження та багато інших. Традиційні методи розв'язання цих задач стикаються з проблемами недостатньої точності та ефективності визначення динамічних параметрів в умовах швидко змінюваних середовищ та складних динамічних сценаріїв. Розглядаються сучасні методи ідентифікації параметрів динамічних об'єктів з використанням технологій трансформерів виявлення та оптичного потоку. Трансформер виявлення є одним з новітніх підходів у галузі комп'ютерного зору, що використовує архітектуру трансформера для завдань детектування об'єктів. Цей трансформер інтегрує процеси виявлення об'єктів і визначення їхніх меж у єдину модель end-to-end, 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 трансформерів виявлення 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 трансформерів виявлення та оптичного потоку забезпечує надійне і швидке визначення параметрів рухомих об'єктів у реальному часі, що може бути застосовано у різних практичних сценаріях. Проведені дослідження також показали 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, трансформер виявлення, 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 відео даних, що зазнають постійних змін. Ці задачі стають особливо актуальними у таких галузях як комп'ютерний зір, системи відеоспостереження, робототехніка, 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), кодувальні та декодувальні архітектури, а також нейронні мережі прямого поширення (НМПП). 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 ЗНМ, 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 і надані висновки щодо їхнього потенціалу для розв'язання актуальних задач в різних областях застосування [2].

З вибором технології трансформерів виявлення (ТРВИ) є кілька причин, чому цей підхід привабливий для ДІДО у зображеннях [3].

Трансформери виявлення базуються на трансформерних архітектурах, які виявилися дуже ефективними в обробці послідовностей даних, таких як мовний текст. Використання трансформерів у відображенні зображень дозволяє моделі визначати глобальні та локальні залежності у зображенні, що може бути корисним при ідентифікації об'єктів.

Трансформери виявлення використовують підхід end-to-end, що пропонує архітектуру, яка дозволяє здійснювати виявлення об'єктів та їхню класифікацію

© 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.



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, і оптимізує швидкість роботи моделі [3, 4].

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

Отже, обираючи технологію ТРВИ 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 аналізі моделей дистанційної ідентифікації параметрів динамічних об'єктів з використанням трансформерів виявлення та оптичного потоку.

Для досягнення мети поставлені задачі дослідження.

1. Розглянути схему дистанційної ідентифікації з використанням ТРВИ та оптичного потоку.
2. Розглянути моделі технології ТРВИ.
3. Розглянути структуру кодера та декодера в технології ТРВИ.

4. Розглянути моделі багатоголової уваги та самоуваги в технології трансформерів виявлення.

**Технологія дистанційної ідентифікації.** Схема дистанційної ідентифікації параметрів динамічного об'єкта з використанням ТРВИ та оптичного потоку показана на рис. 1 і включає такі елементи.

1. Блок «Відео вхід». Кожен кадр відео даних (1 кадр, 2 кадр і т. п.) аналізується з використанням блоку «Відео вхід». Результати аналізу подаються на детектор виявлення.

2. Детектор (DETR) у вигляді трансформера виявлення. Детектор дозволяє визначати положення та межі об'єктів у кадрі.

3. Локалізація об'єктів. Дані положення та меж об'єктів у кадрі, визначені детектором виявлення, використовуються для локалізації об'єктів.

4. Оптичний потік (Optical Flow). Застосування алгоритмів оптичного потоку для визначення векторів руху об'єктів між послідовними кадрами дає можливість оцінити швидкість та напрямок руху об'єктів.

5. Аналіз параметрів динамічних об'єктів. На основі отриманих векторів руху проводиться аналіз параметрів об'єктів, включаючи швидкість руху, напрямок руху, траєкторію та інші динамічні характеристики.

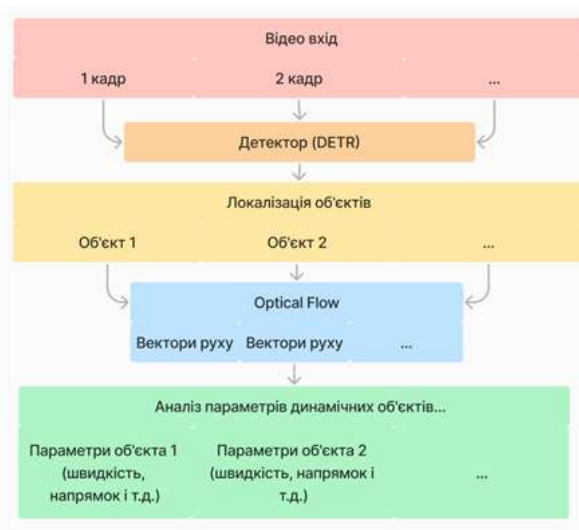


Рис. 1. Схема дистанційної ідентифікації параметрів динамічних об'єктів

На рис. 2 показана схема технології трансформерів виявлення, що включає ЗНМ (а), кодер (б), декодер (в), НМПП (г).

Для виявлення просторових ознак у ЗНМ застосовується вхідний шар для зображень. У кодері який складається з декількох шарів, вхідні ознаки перетворюються у високорівневі представлення. Декодер формує набір обмежуючих рамок для передбачення виходів на основі представлень, отриманих від кодера.

Розроблена модель декодера має вигляд [8]

$$B_t = \text{DETR}(I_t) \quad (1)$$

де  $I_t$  – кадр;

$t$  – індекс кадру;

$B_t$  – набір обмежуючих рамок для кожного кадру.

Кожний кадр аналізується з використанням відео входу та детектора ТРВИ, що дозволяє точно визначати

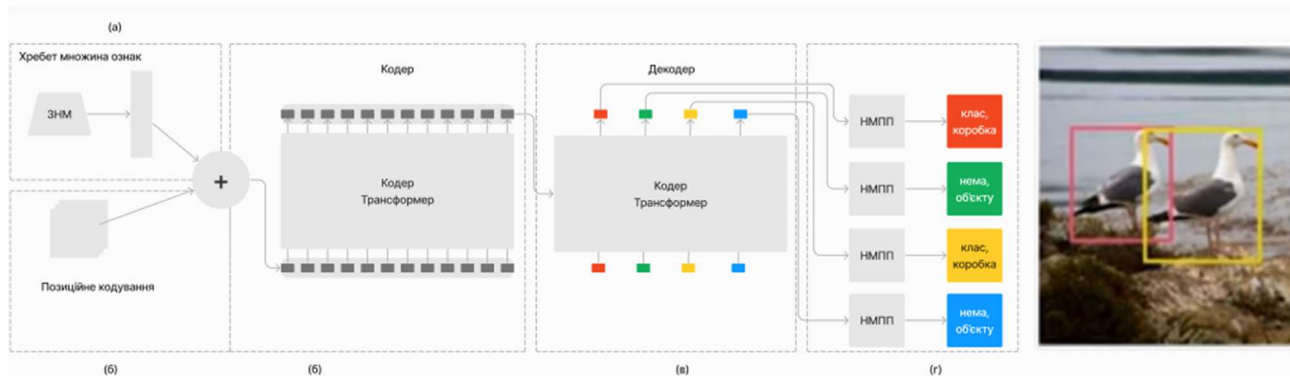


Рис. 2. Схема технології трансформерів виявлення: а – ЗНМ; б – кодер; в – декодер; г – НМПП

положення та межі об'єктів у кадрі. Блок оптичного потоку приймає вхідне зображення і генерує набір обмежувачів рамок для кожного кадру

$$F_{t \rightarrow t+1} = \text{OpticalFlow}(I_t, I_{t+1}), \quad (2)$$

де  $I_t$  та  $I_{t+1}$  – послідовні кадри;

$F_{t \rightarrow t+1}$  – оптичний потік.

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

$$v = \frac{1}{N} \sum_{i=1}^N \|F_{t \rightarrow t+1}(i)\|, \quad (3)$$

де  $N$  – кількість точок, що належать об'єкту;

$F_{t \rightarrow t+1}$  – вектори руху;

$v$  – швидкість;

$\theta$  – кут напрямку руху.

На основі отриманих векторів руху проводиться аналіз параметрів об'єктів, включаючи швидкість та напрямок руху. Напрямок руху об'єкта визначається, як середнє значення напрямків векторів руху. Швидкість об'єкта розраховується як величина середнього вектору руху. Кут напрямку руху об'єкта обчислюється за формулою

$$\theta = \frac{1}{N} \sum_{i=1}^N \text{atan2}(F_{t \rightarrow t+1}^y(i), F_{t \rightarrow t+1}^x(i)). \quad (4)$$

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

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

**Модель кодер – декодер.** Схема моделі з'єднання елементів кодер – декодер показана на рис. 3. На рис. 3а надано схему кодера, який складається з шарів. Кожен шар має два основні підблоки – багатоголову самоувагу (Multi-Head Self-Attention) та НМПП, де використовується багатощаровий перцептрон.

На рис. 3б показано схему декодера, який має схожу структуру до кодера з додатковими компонентами. Список шарів включає багатоголову самоувагу, суму та нормування, багатоголову увагу (Multi-Head attention), суму та нормування, НМПП, де як і в кодері використовується багатощаровий перцептрон [3, 10].

Кодер – це частина трансформера, яка приймає на вхід ознаки, що отримані ЗНМ, та обробляє їх за допомогою механізмів багатоголової 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 багатоголової уваги для визначення набору фіксованих об'єктів. Виходом декодера є кінцеві координати та категорії виявлених об'єктів.

На етапі попередньої обробки зображення ЗНМ витягує відповідні просторові ознаки [1, 5].

Трансформери мають мережеву архітектуру, що базується на механізмах уваги для машинного перекладу. Маючи елемент запиту і набір ключових елементів, багатоголовий модуль уваги адаптивно агрегує ключовий зміст з відповідними вагами уваги, які вимірюють сумісність пар запит-ключ. Елементом запиту може бути цільове слово у вихідному реченні, а набір ключових елементів можуть складати вихідні слова у вхідному реченні.

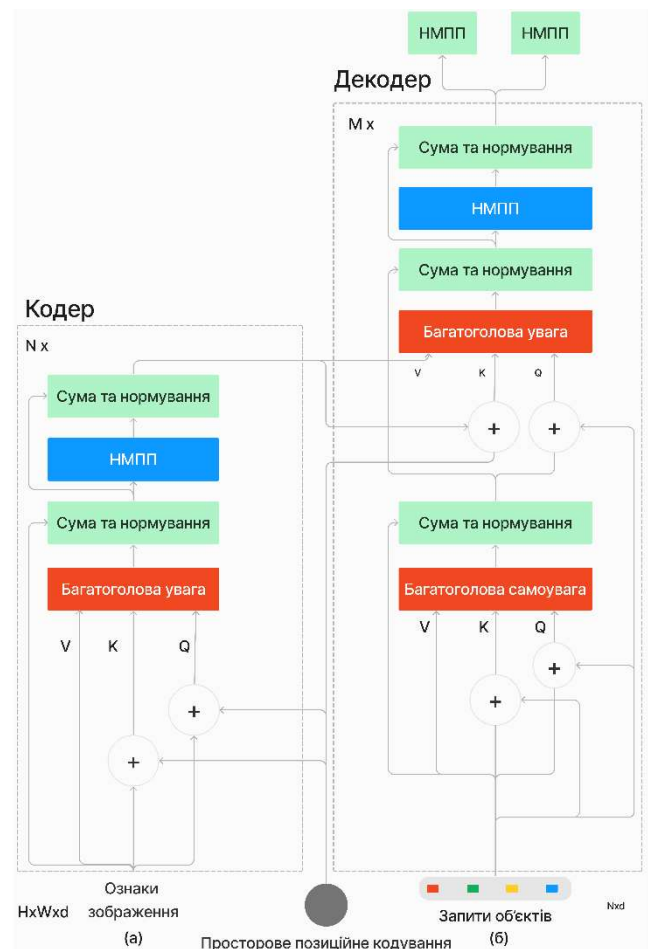


Рис. 3. Модель кодер – декодер: а) кодер; б) декодер

Щоб дозволити моделі фокусуватися на вмісті з різних підпросторів представлення і з різних позицій, виходи різних голів уваги лінійно агрегуються з перенавчальними важелями. Цей механізм забезпечує трансформерам здатність визначити складні залежності між елементами вхідних даних, що робить їх ефективними для широкого спектру задач обробки природної мови та інших областей.

Функція багатоголової уваги обчислюється за формулою [11]

$$\text{MultiHeadAttn}(z_q, x) = \sum_{m=1}^M W_m \left[ \sum_{k \in \Omega_k} A_{mqk} * W'_m x_k \right], \quad (5)$$

де  $q \in \Omega q$  – індекс елемента запиту;  
 $z_q \in \mathbb{R}^C$  – ознака представлення запиту;  
 $m$  – індекс голови уваги;  
 $M$  – кількість голів уваги;  
 $W_m \in \mathbb{R}^{C \times C_v}$  – вихідна проєкція матриці.  
 $k \in \Omega k$  – індекс елемента ключа;  
 $A_{mqk}$  – важелі уваги;  
 $W'_m \in \mathbb{R}^{C_v \times C}$  – вхідна проєкція матриці;  
 $x_k \in \mathbb{R}^C$  – ознака представлення ключа;  
 $\Omega q$  та  $\Omega k$  – множини елементів запиту та ключа відповідно.

Вхідна та вихідна проєкції матриці мають перенавчальні важелі, що визначаються розмірністю уваги для кожної голови уваги

$$C_v = C / M, \quad (6)$$

де  $C$  – розмірність ознаки.

Важелі уваги обчислюються за формулою

$$A_{mqk} \propto e^{\frac{z_q^T U_m^T V_m x_k}{\sqrt{C_v}}}, \quad (7)$$

де  $U_m, V_m \in \mathbb{R}^{C_v \times C}$  – перенавчальні важелі вхідних запитів та ключів для проєкції матриці на  $m$ -ту голову;

$m, q$  та  $k$  – індекси голови, запиту та ключа відповідно.

Важелі уваги нормалізуються за формулою

$$\sum_{k \in \Omega k} A_{mqk} = 1. \quad (8)$$

Ознаки представлення  $z_q$  та  $x_k$  зазвичай є конкатенацією/сумою вмісту елементів та позиційних вкладок для розрізнення різних просторових позицій.

Вектори запитів, ключів і значень обчислюються за формулами:

$$q_i = X_i W^Q, \quad k_i = X_i W^K, \quad v_i = X_i W^V, \quad (9)$$

де  $X_i$  – вхідний вектор або матриця вхідних векторів;

$W^Q$  – матриця ваг для перетворення в запити;

$W^K$  – матриця ваг для перетворення в ключі;

$W^V$  – матриця ваг для перетворення в значення;

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

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

$v_i$  – вектор значення, отриманий з вхідного вектору або матриці вхідних векторів.

Щоб обчислити ваги уваги і зважити відповідні значення створені вектори запитів, ключів і значень використовуються в механізмі уваги з функцією

$$\text{Attention}(q_i, k_j, v_j) = \sum_{j=1}^n \text{softmax} \left( \frac{q_i \cdot k_j}{\sqrt{d_k}} \right) v_j, \quad (10)$$

де  $d_k$  – розмірність векторів ключів.

$K$  – матриця ключів;

$k_j$  – вектор-стовпець матриці ключів;

$V$  – матриця значень;

$v_j$  – 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 проєкційний шар перед об'єднанням для подальшої обробки.

Багатоголова самоувага у трансформерів – 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 за принципом 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 – ці засоби 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 ТРВИ та метод оптичного потоку для ефективної ідентифікації параметрів динамічних об'єктів [2, 3, 5]. Трансформери виявлення є потужним інструментом для точного виявлення динамічних об'єктів на зображеннях. Цю модель доцільно використовувати для локалізації параметрів динамічних об'єктів на кадрах відео даних.

Запропоновані моделі та методи будуть використані для дистанційної 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: 06.05.2024).
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: 06.05.2024).
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: 06.05.2024).
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=IwAR0ILGAWTuwU-9-iH6lZyPFXyXA5JRWarM\\_XoSJ78QEhmnn-txvr\\_iGEzCio](https://arxiv.org/pdf/1905.05055.pdf?fbclid=IwAR0ILGAWTuwU-9-iH6lZyPFXyXA5JRWarM_XoSJ78QEhmnn-txvr_iGEzCio) (дата звернення: 06.05.2024).
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: 06.05.2024).
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: 06.05.2024).
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 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. Zhu X., Hu H., Lin S., Dai J. *Deformable ConvNets v2: More Deformable, Better Results*. URL: <https://arxiv.org/abs/1811.11168> (дата звернення: 06.05.2024).
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](https://www.jstage.jst.go.jp/article/ieejieiss/129/11/129_11_2065/_pdf-char/ja) (дата звернення: 06.05.2024).
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.
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 06.05.2024).
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 06.05.2024).
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=IwAR0ILGAWTuwU-9-iH6lZyPFXyXA5JRWarM\\_XoSJ78QEhmnn-txvr\\_iGEzCio](https://arxiv.org/pdf/1905.05055.pdf?fbclid=IwAR0ILGAWTuwU-9-iH6lZyPFXyXA5JRWarM_XoSJ78QEhmnn-txvr_iGEzCio) (accessed 06.05.2024).
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 06.05.2024).
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 06.05.2024).
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., Reкова N. Y. Analiz informatsiinykh tekhnolohii dlia dystantsiinoi identyfikatsii dinamichnykh ob'ektiv [Analysis of information technologies for remote identification of dynamic objects]. *Visnyk NTU «KhPI». Seriya: Systemnyi analiz, upravlinnia ta informatsiini tekhnolohii* [Bulletin of NTU "KhPI". Series: System analysis, management and information technologies]. Kharkiv, NTU "KhPI" Publ., 2023, no. 1 (9), pp. 110–115.
9. Nikulina O. M., Kondratov O. M. Model identyfikatsii parametriv dynamichnoho ob'ektu z vykorystanniam DEtection TRansformer ta Optical Flow [Dynamic object parameter identification model using DEtection TRansformer and Optical Flow]. *Informatsiini tekhnolohii: nauka, tekhnika, tekhnolohiia, osvita, zdorovia: Tezy dopovidei XXXII mizhnarodnoi naukovo-praktychnoi konferentsii MicroCAD-2024, 22–24 travnia 2024 r.* [Information technologies: science, technology, technology, education, health: Abstracts of reports of XXXII international scientific and practical conference MicroCAD-2024, May 22–24, 2024]. Kharkiv, NTU "KhPI" Publ., 2024, p. 1047.
10. Nikulina O. M., Kondratov O. M. Metody dystantsiinoi identyfikatsii dinamichnykh parametriv ob'ekta [Methods of remote identification of dynamic object parameters]. *Informatsiini tekhnolohii: nauka, tekhnika, tekhnolohiia, osvita, zdorovia: Tezy dopovidei XXXI mizhnarodnoi naukovo-praktychnoi konferentsii MicroCAD-2023, 17–20 travnia 2023 r.* [Information technologies: science, technology, technology, education, health: Abstracts of reports of the XXXI international scientific and practical conference MicroCAD-2023, May 17–20, 2023]. Kharkiv, NTU "KhPI" Publ., 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 06.05.2024).
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](https://www.jstage.jst.go.jp/article/ieejieiss/129/11/129_11_2065/_pdf-char/ja) (accessed 06.05.2024).
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.

## 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 06.05.2024).

Надійшла (received) 09.05.2024

UDC 004.8+004.9

**O. M. NIKULINA**, Doctor of Technical Sciences, Associate 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; ORCID: <https://orcid.org/0000-0003-2938-4215>; e-mail: [elniknik02@gmail.com](mailto: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](mailto: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](mailto: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@khp.edu.ua](mailto:aleksey.olhovoy@khp.edu.ua)

### **MODELS OF REMOTE IDENTIFICATION OF PARAMETERS OF DYNAMIC OBJECTS USING DETECTION TRANSFORMERS AND OPTICAL FLOW**

The tasks of remote identification of parameters of dynamic objects are important for various fields, including computer vision, robotics, autonomous vehicles, video surveillance systems, and many others. Traditional methods of solving these problems face the problems of insufficient accuracy and efficiency of determining dynamic parameters in conditions of rapidly changing environments and complex dynamic scenarios. Modern methods of identifying parameters of dynamic objects using technologies of detection transformers and optical flow are considered. Transformer detection is one of the newest approaches in computer vision that uses transformer architecture for object detection tasks. This transformer integrates the object detection and boundary detection processes into a single end-to-end model, which greatly improves the accuracy and speed of processing. The use of transformers allows the model to effectively process information from the entire image at the same time, which contributes to better recognition of objects even in difficult conditions. Optical flow is a motion analysis method that determines the speed and direction of pixel movement between successive video frames. This method allows obtaining detailed information about the dynamics of the scene, which is critical for accurate tracking and identification of parameters of moving objects. The integration of detection transformers and optical flow is proposed to increase the accuracy of identification of parameters of dynamic objects. The combination of these two methods allows you to use the advantages of both approaches: high accuracy of object detection and detailed information about their movement. The conducted experiments show that the proposed model significantly outperforms traditional methods both in the accuracy of determining the parameters of objects and in the speed of data processing. The key results of the study indicate that the integration of detection transformers and optical flow provides reliable and fast determination of parameters of moving objects in real time, which can be applied in various practical scenarios. The conducted research also showed the potential for further improvement of data processing methods and their application in complex dynamic environments. The obtained results open new perspectives for the development of intelligent monitoring and control systems capable of adapting to rapidly changing environmental conditions, increasing the efficiency and safety of their work.

**Keywords:** remote dynamic object identification, object detection, detection transformer, 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, Sevryn Valerii Petrovych

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

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