

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ  
«ХАРКІВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ»  
МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ  
«ХАРКІВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ»  
МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

Кваліфікаційна наукова  
праця на правах рукопису

**КОНДРАТОВ ОЛЕКСІЙ МИХАЙЛОВИЧ**

УДК 621.391, 004.4:004.9

**ДИСЕРТАЦІЯ**  
**МОДЕЛІ, МЕТОДИ ТА ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ**  
**ДИСТАНЦІЙНОЇ ІДЕНТИФІКАЦІЇ ПАРАМЕТРІВ ДИНАМІЧНИХ**  
**ОБ'ЄКТІВ**

Спеціальність 122 – Комп'ютерні науки  
Галузь знань 12 – Інформаційні технології

Подається на здобуття наукового ступеня доктора філософії

Дисертація містить результати власних досліджень. Використання ідей,  
результатів і текстів інших авторів мають посилання на відповідне джерело



О.М. Кондратов

Науковий керівник:  
Нікуліна Олена Миколаївна  
доктор технічних наук, професор

Харків – 2026

## АНОТАЦІЯ

*Кондратов О.М.* Моделі, методи та інформаційна технологія дистанційної ідентифікації параметрів динамічних об'єктів. – Кваліфікаційна наукова праця на правах рукопису.

Дисертація на здобуття наукового ступеня доктора філософії (PhD) за спеціальністю 122 «Комп'ютерні науки», Національний технічний університет «Харківський політехнічний інститут», Міністерство освіти і науки України, Харків, 2026.

Дисертаційне дослідження присвячене вирішенню актуальної науково-технічної задачі підвищення точності, надійності та адаптивності дистанційної ідентифікації параметрів динамічних об'єктів, що рухаються у просторі, на основі розробки моделей, методів та інформаційної технології, які поєднують трансформерні моделі виявлення об'єктів, алгоритми оптичного потоку та методи обчислювального інтелекту.

*Об'єкт дослідження* – процес дистанційної ідентифікації параметрів динамічних об'єктів у відеопотоках та сенсорних даних.

*Предмет дослідження* – моделі, методи та інформаційна технологія визначення просторових, кінематичних і структурних параметрів динамічних об'єктів на основі трансформерів, оптичного потоку, нейромережевої оцінки глибини та ансамблевих методів.

*Метою дисертаційної роботи* є підвищення точності, надійності та адаптивності дистанційної ідентифікації параметрів динамічних об'єктів, що рухаються у просторі, на основі поєднання трансформерних моделей виявлення об'єктів, методів оптичного потоку та методів обчислювального інтелекту для обробки відеоданих.

У *вступі* обґрунтовано актуальність проблеми, визначено мету, об'єкт і предмет дослідження, сформульовано наукову новизну, практичне значення одержаних результатів, описано методи дослідження, вказано елементи апробації наукових результатів та інформацію про публікації за темою

дисертації.

У *першому розділі* проаналізовано сучасні задачі та методи дистанційної ідентифікації параметрів динамічних об'єктів, визначено основні параметри, що підлягають оцінюванню, та розглянуто інформаційні технології.

У *другому розділі* досліджено методичну основу дистанційної ідентифікації: класичні алгоритми оптичного потоку, неймережеві моделі аналізу руху, ансамблеві підходи та способи їх інтеграції.

У *третьому розділі* сформовано моделі дистанційної ідентифікації параметрів динамічних об'єктів за відеоданими, розглянуто побудови моделей, визначення параметрів за допомогою трансформерів і оптичного потоку.

У *четвертому розділі* описано структуру та програмну реалізацію інформаційної технології дистанційної ідентифікації, наведено результати експериментального поєднання трансформерів, методів оптичного потоку, оцінювання глибини та ансамблевих підходів.

У *висновках* узагальнено досягнення поставленої мети. Встановлено, що запропоновані моделі та методи забезпечують перехід від покадрового виявлення об'єктів до просторово-часового опису їх стану. Доведено, що інтеграція трансформерного виявлення, оптичного потоку, неймережевого оцінювання глибини та ансамблевих методів дозволяє суттєво підвищити стійкість ідентифікації в умовах шумів, оклюзій та зміни освітлення. Розроблено інформаційну технологію, яка забезпечує визначення положення, швидкості, напрямку руху, траєкторії та просторових характеристик об'єктів у відеопотоці в режимі реального часу. Підтверджено працездатність запропонованої технології як модульної системи для дистанційної ідентифікації параметрів динамічних об'єктів. Обґрунтовано практичну цінність отриманих результатів для систем безпеки та оборони, автономного транспорту, логістики та медичної діагностики.

*За результатами дослідження отримано такі наукові результати:*

– вперше запропоновано метод дистанційної ідентифікації параметрів динамічних об'єктів, що ґрунтується на поєднанні трансформерної моделі

виявлення об'єктів DEtection TRansformer (DETR) з алгоритмами оптичного потоку, який, на відміну від існуючих підходів забезпечує одночасне виявлення, відстеження та оцінювання динамічних характеристик об'єктів за відеоданими в умовах реального часу, шумів та часткових оклюзій, що дозволяє підвищити точність і стабільність ідентифікації параметрів руху (швидкості, напрямку, траєкторії) без необхідності використання додаткових сенсорів або постобробки;

– вперше розроблено структуру інформаційної технології дистанційної ідентифікації параметрів динамічних об'єктів, що базується на інтеграції методів обчислювального інтелекту, нейромережових архітектур FlowNet та GeoNet, а також механізмів просторово-часової обробки даних, яка, на відміну від існуючих систем об'єднує їх в єдиний інформаційний простір з ансамблевою корекцією результатів, що дозволяє підвищити надійність та адаптивність ідентифікації параметрів в умовах динамічної зміни характеристик середовища та об'єкта спостереження без втрати швидкодії в реальному часі;

– удосконалено метод оцінювання глибини сцени шляхом адаптивного комбінування нейромережових моделей MiDaS, DPT\_Large та GeoNet з ансамблевими підходами Bagging і Boosting, який, на відміну від використання окремих моделей, забезпечує підвищення точності та стійкості просторової локалізації динамічних об'єктів на основі монокулярних зображень за рахунок зменшення дисперсії оцінок (Bagging) та уточнення параметрів на межах об'єктів і в зонах оклюзій (Boosting), що дозволяє визначати координату Z (глибину) об'єкта з меншою похибкою навіть при відсутності стереопари або даних LiDAR;

– удосконалено метод ансамблевої агрегації результатів ідентифікації, а саме застосування бегінгу (Bagging) для зменшення дисперсії оцінок глибини шляхом усереднення результатів на різних підвибірках кадрів та бустінгу (Boosting) для уточнення параметрів на межах об'єктів і в зонах оклюзій шляхом зваженого підсумовування з вагами, оберненими до помилки кожної моделі, який, на відміну від використання окремих моделей без ансамблювання, підвищує стійкість ідентифікації до шумів, локальних артефактів та зміни



освітлення в умовах реальної експлуатації, що дозволяє забезпечити надійну роботу системи в широкому діапазоні зовнішніх умов без необхідності перенавчання або ручного налаштування параметрів;

– удосконалено метод кількісного оцінювання якості ідентифікації шляхом впровадження комплексної системи метрик, що включає EPE та AAE для оцінки точності оптичного потоку, RMSE, MAE та PSNR для оцінки точності відновлення глибини, а також IoU та Dice для оцінки якості сегментації об'єктів, яка, на відміну від використання окремих метрик для кожної задачі, дозволяє провести об'єктивне порівняння 35 комбінацій методів оптичного потоку (Farneback, Lucas-Kanade, Horn-Schunck, FlowNet, GeoNet) та методів глибини (Blur, Gradient, MiDaS, DPT\_Large, GeoNet, Bagging, Boosting), що дозволяє визначити оптимальні конфігурації методів для різних умов спостереження та забезпечити всебічну оцінку ефективності запропонованої інформаційної технології для практичного впровадження;

– отримав подальший розвиток метод синтезу просторово-часових ознак динамічних об'єктів шляхом інтеграції трансформерних механізмів уваги з векторними полями оптичного потоку, в якому, на відміну від класичних методів, вектори руху подаються на вхід трансформерної мережі для моделювання довготривалих просторово-часових залежностей, що дозволяє виявляти складні патерни руху (періодичні, прискорені, нелінійні траєкторії), які не можуть бути описані класичними методами оптичного потоку, та формувати узагальнені параметри положення, швидкості, прискорення і напрямку руху динамічних об'єктів для подальшого прогнозування їх траєкторії.

*Практичне значення отриманих результатів* полягає в безпосередньому використанні запропонованих моделей, методів та програмних засобів при впровадженні інтелектуальних систем відеоспостереження, робототехнічних комплексів та систем підтримки прийняття рішень, ідентифікації параметрів радіосплесків, завдяки чому досягається підвищення точності, надійності, можливості та адаптивності дистанційної ідентифікації динамічних об'єктів.

Практична цінність отриманих результатів дослідження полягає у

можливості їх впровадження у таких важливих галузях української економіки:

- безпека та оборона, де технології дистанційної ідентифікації забезпечують високу точність відстеження об'єктів у реальному часі за складних умов, що має пряме застосування для систем відеоспостереження, охорони периметра та навігації безпілотних літальних апаратів;

- автономний транспорт та логістика, де методи ідентифікації швидкості, траєкторії та глибини дозволяють створювати надійні системи уникнення зіткнень для автономних транспортних засобів та роботизованих платформ;

- цифровізація та економіка, де оптимізація обчислювальних процесів дозволяє розгорнути інтелектуальні системи на мобільних пристроях та вбудованих платформах, що знижує витрати на автоматизацію виробництва;

- медицина, де методи аналізу руху можуть бути адаптовані для неінвазивної діагностики внутрішніх органів (швидкість кровотоку, рух серця), що сприяє ранньому виявленню захворювань.

Результати роботи впроваджені у навчальному процесі кафедри інформаційних систем та технологій НТУ «ХПІ» в дисциплінах «Інтелектуальні системи», «Інтелектуальний аналіз даних», «Дослідження операцій». У навчальному процесі кафедри цифрових технологій та програмно-аналітичних рішень ТОВ «Технічний університет «Метінвест Політехніка» в дисциплінах «Продуктивність офісних систем», «Основи кібербезпеки та захисту інформації», «Розробка GUI», «Технології захисту даних та інформаційної безпеки». У Радіоастрономічному інституті НАН України в частині дослідження та аналізу динамічних процесів у сонячній атмосфері за даними дистанційних радіоспостережень, зокрема для виявлення та ідентифікації сонячних радіосплесків.

*Ключові слова:* інтелектуальні технології, нейронна мережа, модель, технологія, глибоке навчання, інформаційна система, інформаційні технології, штучні нейронні мережі, ідентифікація, згорткові нейронні мережі, програмне забезпечення, виявлення об'єктів.

## ABSTRACT

*Kondratov O.M.* Models, methods and information technology for remote identification of parameters of dynamic objects. – A qualification scientific paper on the rights of a manuscript.

Dissertation for the degree of Doctor of Philosophy (PhD) in specialty 122 "Computer Sciences", National Technical University "Kharkiv Polytechnic Institute", Ministry of Education and Science of Ukraine, Kharkiv, 2026.

The dissertation research is devoted to solving an urgent scientific and technical problem of increasing the accuracy, reliability and adaptability of remote identification of parameters of dynamic objects moving in space. This is achieved through the development of models, methods and information technology that combine transformer-based object detection, optical flow algorithms, and computational intelligence methods.

*The object of research* is the process of remote identification of parameters of dynamic objects in video streams and sensor data.

*The subject of research* is models, methods and information technology for determining spatial, kinematic and structural parameters of dynamic objects based on transformers, optical flow, neural network depth estimation and ensemble methods.

*The goal of the dissertation* is to increase the accuracy, reliability, and adaptability of remote identification of parameters of dynamic objects moving in space based on a combination of transform object detection models, optical flow methods, and computational intelligence methods for video data processing.

*The introduction* grounds the relevance of the problem, defines the purpose, object and subject of the research, formulates the scientific novelty and practical significance of the findings, describes the research methods, indicates the elements of validation of scientific results and information on publications on the topic of the dissertation.

*The first chapter* analyzes modern tasks and methods of remote identification of dynamic object parameters, defines the main parameters to be estimated, and reviews

information technologies.

*The second chapter* investigates the methodological basis of remote identification: classical optical flow algorithms, neural network motion analysis models, ensemble approaches and methods of their integration.

*The third chapter* formulates models for remote identification of dynamic object parameters from video data, considers the model construction, parameter determination using transformers and optical flow.

*The fourth chapter* describes the structure and software implementation of the information technology for remote identification, presents the results of experimental combination of transformers, optical flow methods, depth estimation and ensemble approaches. The conclusions summarize the achievement of the stated goal.

The *conclusions* summarize the achievement of the goal. It is established that the proposed models and methods provide a transition from frame-by-frame detection of objects to a spatiotemporal description of their state. It is proven that the integration of transformer detection, optical flow, neural network depth estimation and ensemble methods allows to significantly increase the stability of identification in conditions of noise, occlusions and changes in lighting. An information technology has been developed that provides determination of the position, speed, direction of movement, trajectory and spatial characteristics of objects in a video stream in real time. The operability of the proposed technology as a modular system for remote identification of parameters of dynamic objects has been confirmed. The practical value of the obtained results for security and defense systems, autonomous transport, logistics and medical diagnostics has been substantiated.

*The research yielded the following scientific results:*

For the first time, a method for remote identification of dynamic object parameters is proposed, which is based on a combination of a transformer model for object detection DETection TRansformer (DETR) with optical flow algorithms, which, unlike existing approaches, provides simultaneous detection, tracking and evaluation of dynamic characteristics of objects based on video data in real-time conditions, noise and partial occlusions, which allows to increase the accuracy and stability of

identification of motion parameters (speed, direction, trajectory) without the need for additional sensors or post-processing.

For the first time, the structure of information technology for remote identification of parameters of dynamic objects has been developed, which is based on the integration of computational intelligence methods, neural networks architects FlowNet and GeoNet, as well as mechanisms for spatio-temporal data processing, which, unlike existing systems, combines them into a single information space with ensemble correction of results, which allows increasing the reliability and adaptability of parameter identification in conditions of dynamic changes in the characteristics of the environment and the observation object without loss of speed in real time.

The method for estimating scene depth has been improved by adaptively combining the MiDaS, DPT\_Large, and GeoNet neural network models with the ensemble approaches Bagging and Boosting, which, unlike using separate models, provides increased accuracy and stability of spatial localization of dynamic objects based on monocular images by reducing the variance of estimates (Bagging) and refining parameters at object boundaries and in occlusion zones (Boosting), which allows determining the Z coordinate (depth) of an object with less error even in the absence of a stereo pair or LiDAR data .

The method of ensemble aggregation of identification results has been improved, namely the use of Bagging to reduce the variance of depth estimates by averaging the results on different subsamples of frames and Boosting to refine parameters at object boundaries and in occlusion zones by weighted summation with weights inverse to the error of each model, which, in contrast to the use of individual models without ensemble, increases the resistance of identification to noise, local artifacts, and lighting changes in real-world conditions, which allows for reliable operation of the system in a wide range of external conditions without the need for retraining or manual parameter adjustment.

The method for quantitatively assessing the quality of identification has been improved by implementing a comprehensive system of metrics, including EPE and AAE for assessing the accuracy of optical flow, RMSE, MAE and PSNR for assessing

the accuracy of depth restoration, as well as IoU and Dice for assessing the quality of object segmentation, which, unlike the use of separate metrics for each task, allows for an objective comparison of 35 combinations of optical flow methods (Farneback, Lucas-Kanade, Horn-Schunck, FlowNet, GeoNet) and depth methods (Blur, Gradient, MiDaS, DPT\_Large, GeoNet, Bagging, Boosting), which allows determining the optimal configurations of methods for different observation conditions and providing a comprehensive assessment of the effectiveness of the proposed information technology for practical implementation.

The method of synthesizing spatiotemporal features of dynamic objects has been further developed by integrating transformer mechanisms of attention with vector fields of optical flow, in which, unlike classical methods, motion vectors are fed to the input of a transformer network for modeling long-term spatiotemporal dependencies, which allows detecting complex motion patterns (periodic, accelerated, nonlinear trajectories) that cannot be described by classical optical flow methods, and forming generalized parameters of the position, velocity, acceleration, and direction of motion of dynamic objects for further prediction of their trajectory.

*Practical significance of the obtained results* lies in the direct use of the proposed models, methods, and software tools in the implementation of intelligent video surveillance systems, robotic complexes, decision support systems, and identification of radio burst parameters, thereby achieving increased accuracy, reliability, capability, and adaptability of remote identification of dynamic objects.

The practical value of the obtained research results lies in the possibility of their implementation in the following important sectors of the Ukrainian economy:

- security and defense, where remote identification technologies ensure high-precision real-time object tracking under complex conditions, which has direct application for video surveillance systems, perimeter security, and navigation of unmanned aerial vehicles (UAVs);
- autonomous transport and logistics, where methods for identifying speed, trajectory, and depth enable the creation of reliable collision avoidance systems for autonomous vehicles and robotic platforms;

- digitalization and economy, where optimization of computational processes allows the deployment of intelligent systems on mobile devices and embedded platforms, reducing costs for production and logistics automation;

- medicine, where motion analysis methods can be adapted for non-invasive diagnosis of internal organs (blood flow velocity, heart motion), contributing to early disease detection.

The results of the work are implemented in the educational process of the Department of Information Systems and Technologies of NTU "KhPI" in the disciplines "Intelligent Systems", "Intelligent Data Analysis", "Operations Research". In the educational process of the Department of Digital Technologies and Software-Analytical Solutions of LLC "Metinvest Polytechnic Technical University" in the disciplines "Productivity of Office Systems", "Fundamentals of Cybersecurity and Information Protection", "GUI Development", "Data Protection and Information Security Technologies". In the Radio Astronomy Institute of the NAS of Ukraine in the study and analysis of dynamic processes in the solar atmosphere according to remote radio observation data, in particular for the detection and identification of solar radio bursts.

Keywords: intelligent technologies, neural network, model, technology, deep learning, information system, information technology, artificial neural networks, identification, convolutional neural networks, software, object detection.

Список публікацій здобувача

*Публікації здобувача за темою дисертації, в яких опубліковані основні наукові результати:*

1. Нікуліна О. М., Северин В. П., Бубнов А. І., Кондратов О. М. Розробка нелінійної моделі парогенератора АЕС для інформаційної технології оптимізації управління. *Вісник Національного технічного університету «Харківський політехнічний інститут»*. Серія: Системний аналіз, управління та інформаційні технології. 2022. № 1 (7). С. 21–27. DOI: <https://doi.org/10.20998/2079-0023.2022.01.04> (Б).

2. Нікуліна О. М., Северин В. П., Кондратов О. М., Рекова Н. Ю. Аналіз інформаційних технологій для дистанційної ідентифікації динамічних об'єктів. *Вісник Національного технічного університету «Харківський політехнічний інститут»*. Серія: Системний аналіз, управління та інформаційні технології. Харків: НТУ «ХПІ», 2023. № 1 (9). С. 110–115. DOI: <https://doi.org/10.20998/2079-0023.2023.01.17> (Б).

3. 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 Національного технічного університету «Харківський політехнічний інститут»*. Серія: Системний аналіз, управління та інформаційні технології. Харків : НТУ «ХПІ», 2024. № 1 (11). С. 52–57. DOI: <https://doi.org/10.20998/2079-0023.2024.01.08> (Б).

4. Кондратов О. М., Нікуліна О. М. Програмна реалізація із використанням трансформера з оптичним потоком та GEONET для ідентифікації параметрів динамічних об'єктів. *Вісник Національного технічного університету «Харківський політехнічний інститут»*. Серія: Системний аналіз, управління та інформаційні технології. Харків: НТУ «ХПІ», 2024. № 2 (12). С. 86–91. DOI: <https://doi.org/10.20998/2079-0023.2024.02.13> (Б).

5. Кондратов О. М., Северин В. П., Попазов Д. К., Любарський С. М., Нікуліна О. М. Аналіз методів обчислювального інтелекту для моделювання,



ідентифікації, оптимізації систем та підтримки прийняття рішень. *Вісник Національного технічного університету «Харківський політехнічний інститут»*. Серія: Стратегічне управління, управління портфелями, програмами та проектами. 2024. №2 (9). С. 35–44. DOI: <https://doi.org/10.20998/2413-3000.2024.9.5> (Б).

6. Кондратов О. М., Нікуліна О. М. Ідентифікація параметрів динамічних об'єктів з використанням трансформера з оптичним потоком та ансамблевих методів. *Вісник Національного технічного університету «Харківський політехнічний інститут»*. Серія: Системний аналіз, управління та інформаційні технології. Харків: НТУ «ХПІ», 2025. № 1 (13). С. 106–112. DOI: <https://doi.org/10.20998/2079-0023.2025.01.16> (Б).

7. Кондратов О. М., Северин В. П., Попазов Д. К., Любарський С. М., Нікуліна О. М. Використання обчислювального інтелекту для моделювання, ідентифікації, оптимізації інформаційних управляючих систем та підтримки прийняття рішень. *Вісник Національного технічного університету «Харківський політехнічний інститут»*. Серія: Стратегічне управління, управління портфелями, програмами та проектами. 2025. №2 (11). С. 23–29. DOI: <https://doi.org/10.20998/2413-3000.2025.11.3> (Б).

*Опубліковані праці апробаційного характеру:*

8. Kondratov O., Severyn V. and Nikulina O., Intelligent Technologies of Remote Identification of Dynamic Objects. *2025 IEEE 6th KhPI Week on Advanced Technology (KhPIWeek)*. PP. 843-848. DOI: <https://doi.org/10.1109/KhPIWeek61436.2025.11288651> (Scopus, Україна).

9. 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. URL: [http://science.kpi.kharkov.ua/wp-content/uploads/2023/05/Zbirnik-tez-MicroCAD-2023-new\\_compressed-1.pdf](http://science.kpi.kharkov.ua/wp-content/uploads/2023/05/Zbirnik-tez-MicroCAD-2023-new_compressed-1.pdf) .

10. Кондратов О. М., Нікуліна 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. 1247. URL: <https://repository.kpi.kharkov.ua/server/api/core/bitstreams/9efcf7fb-e64e-400b-ba2d-999db613858f/content> .*

11. Кондратов О. М., Нікуліна О. М. Дистанційна ідентифікація параметрів об'єктів на основі flowNet та yolo в порівнянні з optical flow та detr. *XVIII Міжнародна науково-практична конференція магістрантів та аспірантів «Теоретичні та практичні дослідження молодих вчених» 19–22 листопада 2024 року: матеріали конференції. Харків. НТУ «ХПІ». 2024. С. 106-107. URL: <https://repository.kpi.kharkov.ua/items/f3d2ded1-7f56-42ae-8d70-c55cafb7e556> .*

12. Кондратов О. М., Любарський С. М., Северин В. П., Нікуліна О. М. Інтелектуальні системи підтримки прийняття рішень для оптимізації процесів керування та дистанційної ідентифікації параметрів об'єктів. *Інформаційні технології: наука, техніка, технологія, освіта, здоров'я: Тези доповідей XXXIII міжнародної науково-практичної конференції MicroCAD-2025, 14-17 травня 2025 р. Харків. НТУ «ХПІ». 2025. С. 1445. URL: <https://ndch.kpi.kharkov.ua/wp-content/uploads/2025/06/Zbirnik-tez-2025.pdf> .*

13. Попазов Д. К., Кондратов О. М., Северин В. П., Нікуліна О. М. *Моделювання та оптимізація інформаційних управляючих систем на основі обчислювального інтелекту дистанційної ідентифікації параметрів об'єктів. Інформаційні технології: наука, техніка, технологія, освіта, здоров'я: Тези доповідей XXXIII міжнародної науково-практичної конференції MicroCAD-2025, 14-17 травня 2025 р. Харків. НТУ «ХПІ». 2025. С. 1451. URL: <https://ndch.kpi.kharkov.ua/wp-content/uploads/2025/06/Zbirnik-tez-2025.pdf> .*

14. Кондратов О. М., Любарський С. М., Попазов Д. К., Нікуліна О. М., Северин В. П. Обчислювальний інтелект для моделювання та ідентифікації інформаційних управляючих систем і підтримки прийняття рішень. *XIX*

*Міжнародна науково-практична студентська конференція магістрантів та аспірантів «Теоретичні та практичні дослідження молодих вчених» 19–21 листопада 2025 року. Харків. НТУ «ХПІ». 2025. С. 184-185. URL: <https://repository.kpi.kharkov.ua/items/2fd9c23a-6d04-4b5e-96b2-116420842588> .*

15. Кондратов О. М., Нікуліна О. М. Використання нелінійної моделі парогенератора аес у дистанційній ідентифікації параметрів об'єкта. *XIX Міжнародна науково-практична студентська конференція магістрантів та аспірантів «Теоретичні та практичні дослідження молодих вчених» 19–21 листопада 2025 року. Харків. НТУ «ХПІ». 2025. С. 88-89. URL: <https://repository.kpi.kharkov.ua/items/2fd9c23a-6d04-4b5e-96b2-116420842588> .*

## ЗМІСТ

Список умовних позначень .....	4
Вступ.....	7
Розділ 1 АНАЛІЗ ЗАДАЧ ТА МЕТОДІВ ДИСТАНЦІЙНОЇ ІДЕНТИФІКАЦІЇ ПАРАМЕТРІВ ДИНАМІЧНИХ ОБ’ЄКТІВ .....	18
1.1 Динамічні об’єкти як об’єкти дистанційної ідентифікації .....	18
1.2 Класифікація параметрів та методів ідентифікації.....	22
1.3 Інформаційні технології дистанційної ідентифікації .....	25
1.4 Моделі динамічних об’єктів.....	34
1.5 Методи обчислювального інтелекту в задачах ідентифікації .....	37
1.6 Мета та задача дослідження .....	43
1.7 Висновки до розділу.....	43
Розділ 2 МЕТОДИ ДИСТАНЦІЙНОЇ ІДЕНТИФІКАЦІЇ НА ОСНОВІ КОМП’ЮТЕРНОГО ЗОРУ ТА ГЛИБИННОГО НАВЧАННЯ .....	46
2.1 Методи оптичного потоку в задачах ідентифікації .....	46
2.2 Нейромережеві моделі для аналізу руху.....	52
2.3 Ансамблеві методи.....	54
2.4 Об’єднання методів.....	56
2.5 Висновки до розділу.....	58
Розділ 3 МОДЕЛІ ДИСТАНЦІЙНОЇ ІДЕНТИФІКАЦІЇ ПАРАМЕТРІВ ДИНАМІЧНИХ ОБ’ЄКТІВ .....	60
3.1 Моделі дистанційної ідентифікації .....	60
3.2 Моделі ідентифікації параметрів за відеоданими.....	61
3.3 Визначення параметрів методами оптичного потоку.....	73
3.4 Об’єднання моделей і методів ідентифікації.....	85
3.5 Висновки до розділу.....	89
Розділ 4 ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ ДИСТАНЦІЙНОЇ ІДЕНТИФІКАЦІЇ ПАРАМЕТРІВ ДИНАМІЧНИХ ОБ’ЄКТІВ .....	91
4.1 Структура інформаційної технології дистанційної ідентифікації .....	91
4.2 Програмна реалізація моделей дистанційної ідентифікації.....	97
4.3 Програмна реалізація з використанням ансамблевих методів .....	101
4.4 Результат поєднання трансформерів і методів оптичного потоку.....	103
4.5 Реалізація методів ідентифікації на основі глибокого навчання.....	117
4.6 Тестування інформаційної технології .....	122
4.7 Аналіз отриманих результатів.....	130

4.8 Висновки до розділу.....	132
Висновки .....	135
Список використаних джерел .....	139
Додаток А Список публікацій здобувача .....	153
Додаток Б Результати поєднання методів оптичного потоку, методів оцінювання глибини і трансформерів.....	158
Б.1 Результати поєднання методу Farneback, методів оцінювання глибини і трансформерів.....	158
Б.2 Результати поєднання методу FlowNet, методів оцінювання глибини і трансформерів.....	161
Б.3 Результати поєднання методу GeoNet, методів оцінювання глибини і трансформерів.....	165
Б.4 Результати поєднання методу Lucas–Kanade, методів оцінювання глибини і трансформерів.....	168
Б.5 Результати поєднання методу Horn–Schunck, методів оцінювання глибини і трансформерів.....	172
Додаток В Додаткові формули методів та моделей інформаційної технології	176
Додаток Г Матеріали щодо впровадження результатів .....	188

## СПИСОК УМОВНИХ ПОЗНАЧЕНЬ

AAE – Angular Average Error,  
ADAM – Adaptive Moment Estimation,  
AI – Artificial Intelligence,  
AP – Average Precision,  
API – Application Programming Interface,  
AUC – Area Under Curve,  
BPMN – Business Process Model and Notation,  
CI – Computational Intelligence,  
CNN – Convolutional Neural Networks,  
COCO – Common Objects in Context,  
CPU – Central Processing Unit,  
CSV – Comma-Separated Values,  
DETR – DETection TRansformer,  
DPT – Dense Prediction Transformer,  
DNN – Deep Neural Network,  
EMG – Electromyography,  
EPE – End-Point Error,  
FFN – Feed-Forward Network,  
FPN – пірамідальні мережі ознак,  
GAN – Generative Adversarial Network,  
GPU – графічний процесор,  
GRU – Gated Recurrent Unit,  
GUI – Graphical User Interface,  
HSV – Hue, Saturation, Value,  
IMU – Inertial Measurement Unit,  
IoU – Intersection over Union,  
JPEG – Joint Photographic Experts Group,  
KL – Kullback-Leibler divergence,

LiDAR – Light Detection and Ranging,  
LN – Layer Normalization,  
LSTM – Long Short-Term Memory,  
MAE – Mean Absolute Error,  
mAP – mean Average Precision,  
MiDaS – Mixed Dataset of Scenes for single-image depth estimation,  
MLP – Multilayer Perceptron,  
MQTT – Message Queuing Telemetry Transport,  
MSA – Multi-Head Self-Attention,  
MSE – Mean Square Error,  
NAMMs – Neural Attention Memory Models,  
NAS – Neural Architecture Search,  
NLP – Natural Language Processing,  
NMS – Non-Maximum Suppression,  
OD – Object Detection,  
OpenCV – Open Source Computer Vision Library,  
PANet – Path Aggregation Network,  
PNG – Portable Network Graphics,  
PSNR – Peak Signal-to-Noise Ratio,  
R-CNN – Region-based Convolutional Neural Networks,  
RAFT – Recurrent All-Pairs Field Transforms,  
ReLU – Rectified Linear Unit,  
RGB – Red Green Blue,  
RMSE – Root Mean Square Error,  
ROS – Robot Operating System,  
RPN – Region Proposal Network,  
SGD – Stochastic Gradient Descent,  
SSIM – Structural Similarity Index,  
SSD – Single Shot MultiBox Detector,  
SVM – Support Vector Machine,

TPU – тензорний процесор,

ViT – Vision Transformer,

YOLO – You Only Look Once,

XML – eXtensible Markup Language,

БПЛА – безпілотний літальний апарат,

ГОГ (HOG) – гістограми орієнтованих градієнтів,

ДПДО – дистанційна ідентифікація параметрів динамічних об'єктів,

ЗНМ – згорткові нейронні мережі,

ІСППР – інтелектуальні системи підтримки прийняття рішень,

ІТ – інформаційні технології,

ІУС – інформаційні управляючі системи,

МДЧ (DPM) – моделі деформованих частин,

МСМ – багатовимірні матриці помилок,

ПД – пропорційно-інтегрально-диференціальний (регулятор),

ПК – персональний комп'ютер.



## ВСТУП

**Актуальність теми дослідження.** Дисертаційна робота присвячена розробці моделей, методів та інформаційної технології дистанційної ідентифікації параметрів динамічних об'єктів за відеоданими. Актуальність теми зумовлена потребою у надійному визначенні положення, швидкості, напрямку руху, траєкторії та просторових характеристик об'єктів у режимі реального часу без безпосереднього контакту з ними. Задачі дистанційної ідентифікації параметрів динамічних об'єктів (ДПДО) відрізняються високими вимогами до методів та інформаційних технологій їх реалізації. Це необхідність обробки відеоданих у реальному часі, здатність працювати з шумами, оклюзіями та зміною освітлення, забезпечення високої точності визначення просторових, кінематичних та структурних параметрів об'єктів, а також адаптивність до динамічної зміни характеристик середовища спостереження. На сучасному етапі розвитку комп'ютерного зору та глибокого навчання існує широкий спектр методів виявлення об'єктів, оцінювання оптичного потоку та реконструкції глибини сцени. Однак задача інтеграції цих методів в єдину інформаційну технологію, здатну забезпечити надійну ідентифікацію параметрів динамічних об'єктів у реальних умовах експлуатації, відрізняється своєю складністю.

По-перше, необхідно забезпечити сумісну роботу різномірних модулів: детектора об'єктів (наприклад, трансформерної моделі DETR), алгоритмів оптичного потоку (Farneback, Lucas-Kanade, Horn-Schunck, FlowNet) та методів оцінювання глибини (MiDaS, DPT\_Large, GeoNet). Враховуючи, що умови спостереження (освітлення, рівень шумів, наявність оклюзій, швидкість руху об'єктів) часто змінюються протягом відеопослідовності, відповідно змінюються вимоги до точності та стійкості кожного з модулів.

По-друге, треба оцінити якість роботи різних комбінацій методів ДПДО. Такі оцінки, як правило, є багатокритеріальними та включають метрики для оптичного потоку (EPE, AAE), для глибини (RMSE, MAE, PSNR) та для сегментації (IoU, Dice), що ускладнює вибір оптимальної конфігурації.

По-третє, слід розв'язати задачу підвищення стійкості ідентифікації в умовах невизначеності (шумів, оклюзій, зміни освітлення). Ця задача вирішується, як правило, шляхом застосування ансамблевих методів бегінгу та бустінгу, які потребують додаткового налаштування та обґрунтування доцільності їх використання в конкретних сценаріях.

Вказані фактори призводять до того, що багато існуючих систем дистанційної ідентифікації або працюють лише в обмежених умовах (наприклад, при статичній камері або фіксованому освітленні), або потребують значного обчислювального ресурсу, що унеможлиблює їх роботу в реальному часі.

Питанням дистанційної ідентифікації параметрів динамічних об'єктів присвячено цілий ряд робіт вітчизняних та закордонних вчених. Однак запропоновані в них моделі та методи не враховують важливі особливості задачі, зокрема необхідність інтеграції трансформерного виявлення, оптичного потоку та оцінювання глибини в єдину інформаційну технологію з ансамблевою корекцією результатів.

Великий вплив якості ідентифікації параметрів динамічних об'єктів на ефективність систем відеоспостереження, автономного транспорту, робототехніки та промислового моніторингу і, в той же час, складність розв'язання задачі інтеграції різнорідних методів роблять **актуальною наукову задачу** з розробки моделей, методів та інформаційної технології дистанційної ідентифікації параметрів динамічних об'єктів в умовах невизначеності, яке має істотне значення для розвитку інформаційних технологій комп'ютерного зору.

**Зв'язок роботи з науковими програмами, планами, темами.** Дисертаційну роботу виконано на кафедрі інформаційних систем та технологій Національного технічного університету «Харківський політехнічний інститут» відповідно до планів наукових досліджень: «Розробка математичних моделей та програмних додатків для управління складними системами з використанням штучного інтелекту» (ДРН№ 124U001390), «Розробка математичних моделей для оптимізації процесів управління складними динамічними системами з використанням обчислювального інтелекту» (ДРН№ 0124U001511), де здобувач

був виконавцем.

**Мета та задачі дослідження.** Метою дисертаційної роботи є підвищення точності, надійності та адаптивності дистанційної ідентифікації параметрів динамічних об'єктів, що рухаються у просторі, на основі поєднання трансформерних моделей виявлення об'єктів, методів оптичного потоку та методів обчислювального інтелекту для обробки відеоданих.

Для досягнення поставленої мети вирішуються наступні задачі:

- виконати аналіз сучасних методів дистанційної ідентифікації параметрів динамічних об'єктів, визначити їх обмеження та сформулювати наукову задачу інтеграції трансформерних моделей виявлення, алгоритмів оптичного потоку та нейромережевого оцінювання глибини;
- сформувати класифікацію параметрів динамічних об'єктів та методів їх визначення, обґрунтувати систему метрик для оцінювання якості ідентифікації та розробити модель інформаційної технології дистанційної ідентифікації;
- розробити моделі та методи інтеграції трансформерного детектора об'єктів DETR, алгоритмів оптичного потоку (Farneback, Lucas–Kanade, Horn–Schunck, FlowNet, GeoNet) та нейромережевого оцінювання глибини (MiDaS, DPT\_Large, GeoNet), включаючи метод адаптивного комбінування результатів з використанням ансамблевих підходів Bagging та Boosting;
- реалізувати інформаційну технологію дистанційної ідентифікації параметрів динамічних об'єктів, а також провести експериментальне тестування та порівняльний аналіз ефективності запропонованих рішень для 35 комбінацій методів оптичного потоку та оцінювання глибини.

*Об'єктом дослідження* – процес дистанційної ідентифікації параметрів динамічних об'єктів у відеопотоках та сенсорних даних.

*Предмет дослідження* – моделі, методи та інформаційна технологія визначення просторових, кінематичних і структурних параметрів динамічних об'єктів на основі трансформерів, оптичного потоку, нейромережевої оцінки глибини та ансамблевих методів.

**Методи дослідження.** У роботі використано методи комп'ютерного зору

та теорії оптичного потоку, а саме Lucas–Kanade, Horn–Schunck та Farneback, для аналізу руху об’єктів у відеопослідовностях, обчислення векторних полів переміщення пікселів між кадрами та визначення кінематичних параметрів; методи глибокого навчання та трансформерних архітектур, зокрема DETR, FlowNet, GeoNet, MiDaS та DPT, для виявлення об’єктів у кадрі, оцінювання оптичного потоку, реконструкції глибини сцени за монокулярними зображеннями та моделювання довготривалих просторово-часових залежностей; методи математичного та імітаційного моделювання для формалізації процесу дистанційної ідентифікації та дослідження характеристик ідентифікації в різних умовах спостереження; методи теорії ймовірностей і математичної статистики для кількісного оцінювання показників точності, а саме середньої кінцевої похибки EPE, середньої кутової похибки AAE, середньоквадратичної похибки RMSE, середньої абсолютної похибки MAE, пікового відношення сигнал/шум PSNR, а також показників надійності IoU та Dice; методи системного аналізу для побудови інтегрованої інформаційної технології, що поєднає модулі виявлення об’єктів, оцінювання оптичного потоку, реконструкції глибини та ансамблевої корекції результатів; методи ансамблевого навчання Bagging та Boosting для підвищення стійкості результатів в умовах шумів, оклюзій та зміни освітлення; методи програмної інженерії для реалізації експериментального прототипу системи дистанційної ідентифікації, що включає розробку модульної архітектури, інтеграцію різнорідних методів обробки відеоданих, візуалізацію результатів та забезпечення обчислювальної ефективності.

**Наукова новизна одержаних результатів** полягає у розвитку моделей дистанційної ідентифікації параметрів динамічних об’єктів шляхом інтеграції трансформерного виявлення, методів оптичного потоку, оцінювання глибини та обчислювального інтелекту в єдину інформаційну технологію. Наукова новизна визначається такими положеннями:

- *вперше* запропоновано метод дистанційної ідентифікації параметрів динамічних об’єктів, що ґрунтується на поєднанні трансформерної моделі виявлення об’єктів DETection TRansformer (DETR) з алгоритмами оптичного

поток, який, на відміну від існуючих підходів забезпечує одночасне виявлення, відстеження та оцінювання динамічних характеристик об'єктів за відеоданими в умовах реального часу, шумів та часткових оклюзій, що дозволяє підвищити точність і стабільність ідентифікації параметрів руху (швидкості, напрямку, траєкторії) без необхідності використання додаткових сенсорів або постобробки;

- *вперше* розроблено структуру інформаційної технології дистанційної ідентифікації параметрів динамічних об'єктів, що базується на інтеграції методів обчислювального інтелекту, нейромережових архітектур FlowNet та GeoNet, а також механізмів просторово-часової обробки даних, яка, на відміну від існуючих систем об'єднує їх в єдиний інформаційний простір з ансамблевою корекцією результатів, що дозволяє підвищити надійність та адаптивність ідентифікації параметрів в умовах динамічної зміни характеристик середовища та об'єкта спостереження без втрати швидкодії в реальному часі;

- *удосконалено* метод оцінювання глибини сцени шляхом адаптивного комбінування нейромережових моделей MiDaS, DPT\_Large та GeoNet з ансамблевими підходами Bagging і Boosting, який, на відміну від використання окремих моделей, забезпечує підвищення точності та стійкості просторової локалізації динамічних об'єктів на основі монокулярних зображень за рахунок зменшення дисперсії оцінок (Bagging) та уточнення параметрів на межах об'єктів і в зонах оклюзій (Boosting), що дозволяє визначати координату Z (глибину) об'єкта з меншою похибкою навіть при відсутності стереопари або даних LiDAR;

- *удосконалено* метод ансамблевої агрегації результатів ідентифікації, а саме застосування бегінгу (Bagging) для зменшення дисперсії оцінок глибини шляхом усереднення результатів на різних підвибірках кадрів та бустінгу (Boosting) для уточнення параметрів на межах об'єктів і в зонах оклюзій шляхом зваженого підсумовування з вагами, оберненими до помилки кожної моделі, який, на відміну від використання окремих моделей без ансамблювання, підвищує стійкість ідентифікації до шумів, локальних артефактів та зміни освітлення в умовах реальної експлуатації, що дозволяє забезпечити надійну

роботу системи в широкому діапазоні зовнішніх умов без необхідності перенавчання або ручного налаштування параметрів;

- *удосконалено* метод кількісного оцінювання якості ідентифікації шляхом впровадження комплексної системи метрик, що включає EPE та AAE для оцінки точності оптичного потоку, RMSE, MAE та PSNR для оцінки точності відновлення глибини, а також IoU та Dice для оцінки якості сегментації об'єктів, яка, на відміну від використання окремих метрик для кожної задачі, дозволяє провести об'єктивне порівняння 35 комбінацій методів оптичного потоку (Farneback, Lucas-Kanade, Horn-Schunck, FlowNet, GeoNet) та методів глибини (Blur, Gradient, MiDaS, DPT\_Large, GeoNet, Bagging, Boosting), що дозволяє визначити оптимальні конфігурації методів для різних умов спостереження (освітлення, рівень шумів, наявність оклюзій) та забезпечити всебічну оцінку ефективності запропонованої інформаційної технології для практичного впровадження;

- *отримав подальший розвиток* метод синтезу просторово-часових ознак динамічних об'єктів шляхом інтеграції трансформерних механізмів уваги з векторними полями оптичного потоку, в якому, на відміну від класичних методів, вектори руху подаються на вхід трансформерної мережі для моделювання довготривалих просторово-часових залежностей, що дозволяє виявляти складні патерни руху (періодичні, прискорені, нелінійні траєкторії), які не можуть бути описані класичними методами оптичного потоку, та формувати узагальнені параметри положення, швидкості, прискорення і напрямку руху динамічних об'єктів для подальшого прогнозування їх траєкторії.

**Практичне значення** отриманих результатів полягає в безпосередньому використанні запропонованих моделей, методів та програмних засобів при впровадженні інтелектуальних систем відеоспостереження, робототехнічних комплексів та систем підтримки прийняття рішень, ідентифікації параметрів радіосплесків, завдяки чому досягається підвищення точності, надійності, можливості та адаптивності дистанційної ідентифікації динамічних об'єктів.

Практична цінність отриманих результатів дослідження полягає у

можливості їх впровадження у таких важливих галузях української економіки:

- безпека та оборона, де технології дистанційної ідентифікації забезпечують високу точність відстеження об'єктів у реальному часі за складних умов, що має пряме застосування для систем відеоспостереження, охорони периметра та навігації безпілотних літальних апаратів (БПЛА);
- автономний транспорт та логістика, де методи ідентифікації швидкості, траєкторії та глибини дозволяють створювати надійні системи уникнення зіткнень для автономних транспортних засобів та роботизованих платформ;
- цифровізація та економіка, де оптимізація обчислювальних процесів дозволяє розгортати інтелектуальні системи на мобільних пристроях та вбудованих платформах, що знижує витрати на автоматизацію виробництва;
- медицина, де методи аналізу руху можуть бути адаптовані для неінвазивної діагностики внутрішніх органів (швидкість кровотоку, рух серця), що сприяє ранньому виявленню захворювань.

Результати роботи (Додаток Г) впроваджені у навчальному процесі кафедри інформаційних систем та технологій НТУ «ХП» в дисциплінах «Інтелектуальні системи», «Інтелектуальний аналіз даних», «Дослідження операцій». У навчальному процесі кафедри цифрових технологій та програмно-аналітичних рішень ТОВ «Технічний університет «Метінвест Політехніка» в дисциплінах «Продуктивність офісних систем», «Основи кібербезпеки та захисту інформації», «Розробка GUI», «Технології захисту даних та інформаційної безпеки». У Радіоастрономічному інституті НАН України в частині дослідження та аналізу динамічних процесів у сонячній атмосфері за даними дистанційних радіоспостережень, зокрема для виявлення та ідентифікації сонячних радіосплесків.

**Особистий внесок здобувача.** Усі основні результати дисертації, що виносяться на захист, отримані здобувачем самостійно. У наукових роботах, написаних у співавторстві, здобувачеві належить (Додаток А):

1) Фахові видання України

[1] – здобувачем адаптовано нелінійну модель парогенератора АЕС для

використання в інформаційній технології дистанційної ідентифікації динамічних об'єктів; запропоновано спосіб інтеграції динамічних характеристик об'єкта в модель керування. Співавторами: Нікуліною О. М., Севериним В. П., Бубновим А. І. зроблено концептуалізацію, методологію, постановку задачі керування.

[2] – здобувачем виконано аналіз сучасних інформаційних технологій для дистанційної ідентифікації динамічних об'єктів, систематизовано методи комп'ютерного зору, визначено переваги та недоліки підходів на основі оптичного потоку та нейромережових детекторів. Співавторами Нікуліною О. М., Севериним В. П., Рековою Н. Ю. зроблено концептуалізацію, методологію та сформовані завдання дослідження.

[3] – здобувачем запропоновано модель дистанційної ідентифікації параметрів динамічних об'єктів з використанням трансформерів виявлення (DETR) та оптичного потоку; розроблено схему інтеграції детектора об'єктів із методами локального та глобального оптичного потоку для оцінювання швидкості та напрямку руху. Співавторами Нікуліною О. М., Севериним В. П., Ольховим О. М. зроблено концептуалізацію, методологію, постановку задачі, аналіз результатів.

[4] – здобувачем виконано програмну реалізацію інформаційної технології дистанційної ідентифікації параметрів динамічних об'єктів із використанням трансформера (DETR), оптичного потоку та GeoNet; реалізовано модулі оцінювання глибини та візуалізації результатів. Співавтором Нікуліною О. М. зроблено концептуалізація, методологія, тестування та валідація.

[5] – здобувачем проведено аналіз методів обчислювального інтелекту для моделювання, ідентифікації та оптимізації систем; обґрунтовано доцільність використання ансамблевих методів (Bagging, Boosting) для підвищення стійкості ідентифікації параметрів динамічних об'єктів. Співавторами Попазовим Д. К., Любарським С. М., Севериним В. П., Нікуліною О. М. зроблено концептуалізацію, методологію, огляд літератури, аналіз результатів.

[6] – здобувачем розроблено метод ідентифікації параметрів динамічних



об'єктів із використанням трансформера, оптичного потоку та ансамблевих методів; запропоновано схему ансамблевого комбінування карт глибини та оптичного потоку для підвищення точності оцінювання. Співавтором Нікуліною О. М. зроблено концептуалізацію, методологію, постановку задачі, інтерпретацію результатів.

[7] – здобувачем виконано аналіз методів обчислювального інтелекту для моделювання, ідентифікації та оптимізації інформаційних управляючих систем; запропоновано підхід до інтеграції дистанційної ідентифікації параметрів об'єктів у системи підтримки прийняття рішень. Співавторами Попазовим Д. К., Любарським С. М., Севериним В. П., Нікуліною О. М. зроблено концептуалізацію, методологію, аналіз літератури та формулювання висновків.

## 2) Публікації апробаційного характеру

[8] – здобувачем запропоновано інтелектуальні технології дистанційної ідентифікації динамічних об'єктів із використанням комбінації DETR, оптичного потоку та ансамблевих методів; виконано порівняльний аналіз ефективності запропонованої технології з класичними підходами. Співавторами Severyn V., Nikulina O. зроблено концептуалізацію, методологію, постановку експерименту, аналіз результатів.

[9] – здобувачем проведено аналіз методів дистанційної ідентифікації динамічних параметрів об'єкта, систематизовано підходи на основі оптичного потоку та трансформерних моделей. Співавтором Нікуліною О. М. зроблено концептуалізацію, методологію, постановку задачі.

[10] – здобувачем запропоновано модель ідентифікації параметрів динамічного об'єкта з використанням DEtection TRansformer (DETR) та Optical Flow; розроблено схему локалізації об'єктів та обчислення векторів руху для подальшого визначення швидкості та напрямку. Співавтором Нікуліною О. М. зроблено концептуалізацію, методологію, формулювання завдання.

[11] – здобувачем виконано порівняльний аналіз методів дистанційної ідентифікації параметрів об'єктів на основі FlowNet та YOLO у порівнянні з Optical Flow та DETR; визначено умови доцільності застосування кожної

комбінації. Співавтором Нікуліною О. М. зроблено концептуалізацію, методологію, інтерпретацію результатів порівняння.

[12] – здобувачем запропоновано архітектуру інтелектуальної системи підтримки прийняття рішень для оптимізації процесів керування та дистанційної ідентифікації параметрів об'єктів; розроблено схему інтеграції модулів ідентифікації в систему керування. Співавторами Любарським С. М., Севериним В. П., Нікуліною О. М. зроблено концептуалізацію, методологію, постановку задачі.

[13] – здобувачем досліджено методи моделювання та оптимізації інформаційних управляючих систем на основі обчислювального інтелекту дистанційної ідентифікації параметрів об'єктів; запропоновано підхід до використання ідентифікованих параметрів для адаптивного керування. Співавторами Попазовим Д. К., Севериним В. П., Нікуліною О. М. зроблено концептуалізацію, методологію, аналіз літератури.

[14] – здобувачем виконано аналіз методів обчислювального інтелекту для моделювання та ідентифікації інформаційних управляючих систем і підтримки прийняття рішень; обґрунтовано доцільність використання ансамблевих підходів для підвищення точності ідентифікації параметрів динамічних об'єктів. Співавторами Любарським С. М., Попазовим Д. К., Нікуліною О. М., Севериним В. П. зроблено концептуалізацію, методологію, огляд літератури, формулювання висновків.

[15] – здобувачем досліджено можливість використання нелінійної моделі парогенератора АЕС у задачах дистанційної ідентифікації параметрів об'єкта; запропоновано спосіб адаптації параметрів моделі для опису динаміки рухомих об'єктів за відеоданими. Співавтором Нікуліною О. М. зроблено концептуалізацію, методологію, постановку задачі.

**Апробація результатів дисертації.** Основні результати дисертаційної роботи доповідалися та обговорювалися на: XXXI – XXXIII Міжнародних науково-практичних конференціях «Інформаційні технології: наука, техніка, технологія, освіта, здоров'я» (MicroCAD-2023, MicroCAD-2024, MicroCAD-

2025, Харків 2023-2025); XVIII і XIX Міжнародних науково-практичних конференціях магістрантів та аспірантів «Теоретичні та практичні дослідження молодих вчених» (Харків, 2024, 2025). VI Міжнародна науково-практична конференція «KhPI Week on Advanced Technology (KhPIWeek)».

**Публікації.** Результати дисертації, які виносяться на захист, опубліковано у наукових працях, зокрема у фахових виданнях України та матеріалах наукових конференцій, що відповідають тематиці дослідження. За темою дисертаційної роботи опубліковано 15 наукових праць, у тому числі: 7 статей у наукових виданнях, що входять до фахових видань України (категорія Б); 8 матеріалів апробаційного характеру.

**Структура та обсяг дисертації.** Дисертаційна робота є завершеним науковим дослідженням, спрямованим на розв'язання актуальної задачі підвищення точності та надійності дистанційної ідентифікації параметрів динамічних об'єктів. Дисертаційна робота складається з анотації двома мовами, вступу, чотирьох розділів, висновків, списку використаних джерел та чотирьох додатків. Загальний обсяг дисертації складає 152 сторінки, з них: 72 рисунки по тексту; 5 таблиць по тексту; список використаних джерел із 106 найменувань на 13 сторінках; 4 додатки на 39 сторінках.

## **РОЗДІЛ 1 АНАЛІЗ ЗАДАЧ ТА МЕТОДІВ ДИСТАНЦІЙНОЇ ІДЕНТИФІКАЦІЇ ПАРАМЕТРІВ ДИНАМІЧНИХ ОБ'ЄКТІВ**

У першому розділі розглянуто предметну область дистанційної ідентифікації параметрів динамічних об'єктів, визначено основні параметри, що підлягають оцінюванню, та проаналізовано сучасні інформаційні технології комп'ютерного зору, які використовуються для виявлення, відстеження й опису рухомих об'єктів.

### **1.1 Динамічні об'єкти як об'єкти дистанційної ідентифікації**

В умовах стрімкого розвитку автоматизації та штучного інтелекту дистанційна ідентифікація параметрів динамічних об'єктів (ДПДО) стає одним із фундаментальних завдань комп'ютерного зору та систем керування. Даний напрям є спеціалізованою підгалуззю виявлення об'єктів (Object Detection, OD) [1, 2], що фокусується на визначенні характеристик об'єктів на відстані без безпосереднього фізичного контакту, що є критично важливим у небезпечних середовищах або ситуаціях, де контакт неможливий. Динамічні об'єкти в контексті ДПДО є фізичні тіла або системи, що характеризуються постійною зміною своїх станів у часі, включаючи просторове положення, форму, розмір та орієнтацію [3, 4].

ДПДО на сучасному етапі розвитку науки є однією з найбільш пріоритетних підгалузей загального завдання OD у комп'ютерному зорі [1]. Фундаментальною метою цього напрямку є розробка високоточних алгоритмів, здатних визначати кількісні та якісні характеристики об'єктів на значній відстані без безпосереднього фізичного контакту з ними [5]. Це стає критично важливим у сценаріях, де прямий доступ до об'єкта є неможливим, небезпечним для людини або технічно обмеженим, як-от у зонах радіаційного забруднення, високошвидкісних транспортних потоках чи в аерокосмічному просторі.

Динамічні об'єкти в контексті систем дистанційного моніторингу

розглядаються як фізичні тіла або складні технічні системи, стан яких (просторовий, кінематичний або геометричний) безперервно змінюється у часі під впливом внутрішніх процесів або зовнішнього середовища [6]. Такі об'єкти характеризуються наявністю активного руху, зміною траєкторії, а також потенційною трансформацією форми та орієнтації. Для забезпечення ефективного управління та безпеки інформаційні технології повинні ідентифікувати не просто наявність об'єкта, а його прецизійні параметри: координати положення, лінійні розміри, контури меж та просторову орієнтацію.

Однією з ключових задач ідентифікації є оцінювання кінематичних характеристик рухомих об'єктів, що вимагає аналізу послідовності кадрів відеопотоку в реальному часі. До основних динамічних параметрів, що підлягають ідентифікації, належать миттєва швидкість переміщення, вектор напрямку руху та кутові характеристики траєкторії. Використання методів аналізу оптичного потоку Optical Flow дозволяє отримувати детальну інформацію про динаміку сцени шляхом обчислення переміщення пікселів між сусідніми кадрами, що є базою для прогнозування подальшої поведінки об'єкта [7].

Важливим аспектом сучасних досліджень є вихід за межі двовимірного аналізу та перехід до ідентифікації просторових параметрів у 3D-середовищі. Динамічний об'єкт як об'єкт дистанційного спостереження вимагає точної оцінки глибини Depth Estimation, відстані від сенсора до кожної точки об'єкта в тривимірному просторі. Технології типу GeoNet дозволяє одночасно аналізувати геометрію сцени та рух камери, що забезпечує відновлення повноцінної 3D-моделі динамічної сцени та підвищує достовірність ідентифікації в складних просторових сценаріях [8].

У галузях аерокосмічної техніки та робототехніки динамічні об'єкти розглядаються через їх інерційно-масові характеристик. Дистанційна ідентифікація таких параметрів, як маса об'єкта, його центр мас та моменти інерції, обчислюється за допомогою обробки сигналів з різномірних сенсорів (відеокамер, акселерометрів, гіроскопів). Використання нейромережових

структур дозволяє автоматично витягувати просторово-часову інформацію про динаміку таких тіл без необхідності прямого контакту, що забезпечує адаптивність систем управління до зміни навантажень [9].

Автономні транспортні системи (безпілотні автомобілі, БПЛА) є найбільш масовою сферою застосування методів ДПДО, де динамічні об'єкти (інші учасники руху) потребують миттєвої ідентифікації для уникнення зіткнень. Точність визначення швидкості та траєкторії інших об'єктів безпосередньо впливає на безпеку та стабільність прийняття рішень інтелектуальним автопілотом. В таких умовах об'єкт ідентифікації характеризується високою швидкістю змін та непередбачуваністю маневрів, що вимагає обробки даних із затримкою менше 100 мс на кадр [10].

Системи безпеки та інтелектуального відеоспостереження фокусуються на динамічних об'єктах для виявлення підозрілих дій, відстеження руху осіб у натовпі та ідентифікації цілей у режимі реального часу. Особливістю ідентифікації в цій сфері є необхідність стійкого відстеження об'єктів, які можуть тимчасово зникати з поля зору або змінювати свій візуальний вигляд. Динамічний об'єкт у цих задачах виступає джерелом поведінкових патернів, які аналізуються нейронними мережами для ранньої діагностики загроз або порушень [11].

У медичній діагностиці динамічними об'єктами дистанційної ідентифікації можуть виступати внутрішні органи або біологічні рідини. Наприклад, ідентифікація характеристик руху тканин серця або швидкості кровотоку за допомогою обробки відеоданих ультразвукових або рентгенівських апаратів дозволяє виявляти аномалії на ранніх стадіях. Це вимагає від моделей ідентифікації високої деталізації та здатності працювати з низькоконтрастними динамічними зображеннями [12].

Проблематика ДПДО тісно пов'язана з наявністю дестабілізуючих факторів у реальних умовах експлуатації. Динамічне середовище характеризується змінами освітлення, наявністю туману, шумів сенсорів та оклюзіями (частковим перекриттям об'єкта іншими предметами). Кожен із цих

факторів перетворює об'єкт спостереження на складну математичну задачу з високим рівнем невизначеності, що зумовлює необхідність впровадження методів обчислювального інтелекту [13].

Еволюція методів ідентифікації пройшла шлях від традиційних алгоритмів на основі штучно створених ознак (метод Віоли-Джонса, ГОГ) до сучасних архітектур глибокого навчання. Сучасні трансформери виявлення (DETR) та моделі YOLO дозволяють розглядати динамічний об'єкт як сукупність контекстних залежностей у кадрі, забезпечуючи підхід end-to-end. Це дає змогу моделі одночасно виконувати класифікацію об'єкта та точну локалізацію його меж із врахуванням глобальної структури зображення [14].

Для підвищення надійності ідентифікації в умовах динамічної зміни середовища застосовуються ансамблеві методи машинного навчання. Бегінг (Bagging) використовується для зменшення дисперсії моделі шляхом усереднення результатів на різних підвибірках даних, що стабілізує оцінку параметрів об'єкта при шумах. Бустінг (Boosting) дозволяє послідовно коригувати помилки попередніх етапів ідентифікації, фокусуючись на найбільш складних регіонах, як-от межі рухомих об'єктів [15].

Інтеграція методів штучного інтелекту в інформаційні технології ДПІДО дозволяє створювати адаптивні системи, здатні до самонавчання. Динамічний об'єкт у таких системах стає частиною замкненого контуру керування, де ідентифіковані параметри миттєво використовуються для оптимізації режимів роботи технічної системи. Використання інтелектуальної підтримки прийняття рішень дозволяє мінімізувати ризики в умовах стохастичної природи зовнішніх впливів [16].

Перспективи розвитку галузі пов'язані з концепцією «цифрових двійників» (Digital Twins), де ідентифіковані параметри динамічного об'єкта в реальному часі синхронізуються з його віртуальною моделлю. Це створює новий рівень автоматизації – від простого моніторингу до когнітивного керування складними об'єктами, що забезпечує діагностику, компенсацію відхилень та стійкість до відмов. Така інтеграція є основою для побудови інтелектуальних

систем нового покоління [17].

Отже, динамічні об'єкти як об'єкти дистанційної ідентифікації представляють собою складні багатовимірні системи, адекватне розуміння яких можливе лише за умови поєднання фізичних моделей, алгоритмів комп'ютерного зору та методів обчислювального інтелекту. Розвиток інформаційних технологій у цьому напрямі відкриває шлях до повної автономізації складних процесів у промисловості, транспорті та безпеці, підвищуючи загальну ефективність управління в умовах невизначеності. Вирішення задач ДПДО є фундаментом для створення безпечного та інтелектуального технологічного майбутнього [18].

### 1.2 Класифікація параметрів та методів ідентифікації

Класифікація параметрів дистанційної ідентифікації має враховувати не лише факт виявлення об'єкта, а й тип параметра, спосіб його спостереження та часову характеристику. До просторових параметрів належать координати, глибина сцени, геометричні розміри й орієнтація; до кінематичних – швидкість, прискорення, напрям руху та траєкторія; до структурних – контур, площа області, клас об'єкта й ознаки форми. Їх визначення потребує поєднання методів виявлення об'єктів, оцінювання оптичного потоку, реконструкції глибини та фільтрації результатів у часі [19–26].

За способом отримання параметри поділяються на безпосередньо вимірювані та похідні. До першої групи належать координати центра рамки, розміри області локалізації, контур і клас об'єкта, які формуються детектором або сегментаційною моделлю. До другої групи належать швидкість, прискорення, напрям руху, прогнозована траєкторія та відносна глибина, що обчислюються на основі послідовності кадрів, карти оптичного потоку та просторової реконструкції сцени.

За часовим режимом методи ідентифікації доцільно поділяти на покадрові, послідовні та інтегровані. Покадрові методи забезпечують локалізацію об'єкта у конкретний момент часу; послідовні методи аналізують зміну стану між



кадрами; інтегровані методи поєднують детекцію, оцінювання руху, глибини й ансамблеву фільтрацію результатів. Саме інтегрований підхід є базовим для подальшої побудови інформаційної технології ДПДО.

Класифікація параметрів дистанційної ідентифікації має враховувати не лише факт виявлення об'єкта, а й тип параметра, спосіб його спостереження та часову характеристику [19, 20]. До просторових параметрів належать координати положення об'єкта в системі координат кадру, глибина сцени, тобто відстань від камери до окремих точок об'єкта, геометричні розміри, такі як ширина, висота та площа проекції об'єкта на площину зображення, а також орієнтація об'єкта в просторі, яка визначається кутами повороту відносно осей координат [21, 22]. До кінематичних параметрів належать миттєва швидкість переміщення об'єкта, яка обчислюється як зміна його положення за одиницю часу, прискорення, що характеризує зміну швидкості руху, напрям руху, який задається вектором у двовимірному або тривимірному просторі, та траєкторія руху, тобто послідовність положень об'єкта в часі, яка може бути використана для прогнозування його подальшої поведінки [23, 24, 25]. До структурних параметрів належать контур об'єкта, тобто множина точок, що утворюють його межу на зображенні, площа області, яку займає об'єкт у кадрі, клас об'єкта, який визначає його приналежність до певної категорії, наприклад автомобіль, пішохід або велосипедист, та ознаки форми, які описують геометричні властивості об'єкта, зокрема компактність, витягнутість або наявність симетрії [20, 26]. Визначення всіх цих груп параметрів потребує поєднання методів виявлення об'єктів, оцінювання оптичного потоку, реконструкції глибини сцени та фільтрації результатів у часі, що дозволяє отримувати не лише дискретні значення параметрів в окремі моменти часу, а й відстежувати їх динаміку в неперервному відеопотоці [19, 20, 21].

За способом отримання параметри дистанційної ідентифікації доцільно поділяти на дві великі групи, а саме безпосередньо вимірювані параметри та похідні параметри, які обчислюються на основі безпосередньо вимірюваних величин [22, 23]. До першої групи, тобто безпосередньо вимірюваних

параметрів, належать координати центра обмежувальної рамки об'єкта, які визначаються як середнє арифметичне мінімальних та максимальних значень координат по осях  $X$  та  $Y$ , розміри області локалізації, тобто ширина та висота рамки, контур об'єкта, який формується детектором або сегментаційною моделлю у вигляді маски пікселів, що належать об'єкту, та клас об'єкта, який визначається результатом класифікації, виконаної нейронною мережею [24, 25]. Ці параметри є безпосереднім результатом роботи детектора або сегментаційної моделі і не потребують додаткових обчислень для свого отримання, хоча можуть бути уточнені або відфільтровані в часі для підвищення стійкості результатів [26]. До другої групи, тобто похідних параметрів, належать швидкість руху об'єкта, яка обчислюється як відношення зміни положення центру об'єкта до часового інтервалу між сусідніми кадрами, прискорення, яке визначається як зміна швидкості руху за одиницю часу, напрям руху, який задається кутом між вектором переміщення об'єкта та обраною опорною віссю, прогнозована траєкторія, що будується на основі екстраполяції попередніх положень об'єкта з використанням кінематичних моделей різного порядку, та відносна глибина, яка визначається на основі аналізу карти оптичного потоку або просторової реконструкції сцени [19, 20, 21].

За часовим режимом роботи методи ідентифікації параметрів динамічних об'єктів доцільно поділяти на три основні категорії, а саме покадрові методи, послідовні методи та інтегровані методи, кожна з яких має свої переваги та обмеження в залежності від конкретних умов застосування [22, 23]. Покадрові методи забезпечують локалізацію об'єкта та визначення його просторових і структурних параметрів у конкретний момент часу на основі аналізу окремого кадру відеопослідовності, не використовуючи інформацію про попередні стани об'єкта, що робить їх обчислювально простими, але чутливими до шумів, оклюзій та раптових змін зовнішнього вигляду об'єкта [24, 25]. Послідовні методи аналізують зміну стану об'єкта між сусідніми кадрами, використовуючи інформацію про оптичний потік або зміщення ключових точок об'єкта, що дозволяє оцінювати кінематичні параметри, такі як швидкість та напрям руху,

однак вони можуть накопичувати помилку з часом через дрейф оцінок або втрату відстеження при довготривалих оклюзіях [19, 26]. Інтегровані методи поєднують в собі покадрову детекцію об'єктів, оцінювання руху між кадрами на основі оптичного потоку, реконструкцію глибини сцени за допомогою монокулярних або стереоскопічних методів та ансамблеву фільтрацію результатів для підвищення стійкості та точності кінцевих оцінок параметрів [20, 21].

Саме інтегрований підхід, який поєднує всі перелічені вище методи, є базовим для подальшої побудови інформаційної технології ДПДО, оскільки він дозволяє компенсувати недоліки окремих методів за рахунок їх синергії та забезпечує отримання повного набору просторових, кінематичних, геометричних та структурних параметрів об'єкта в режимі реального часу [22, 23, 24]. Такий підхід передбачає, що результати детектора об'єктів, наприклад DETR або YOLO, використовуються для локалізації області інтересу в кожному кадрі, після чого для цієї області обчислюються вектори оптичного потоку, які дозволяють оцінити швидкість та напрям руху, а також будується карта глибини за допомогою неймережевих моделей MiDaS, DPT\_Large або GeoNet, що дає змогу визначити просторове положення об'єкта в тривимірному просторі [25, 26]. Крім того, інтегрований підхід передбачає використання ансамблевих методів бегінгу та бустінгу для фільтрації викидів та уточнення оцінок параметрів на межі об'єктів, що особливо важливо в умовах часткових оклюзій, шумів сенсорів або зміни освітлення [19, 20]. Застосування такого комплексного підходу дозволяє значно підвищити надійність та точність дистанційної ідентифікації параметрів динамічних об'єктів порівняно з використанням окремих методів і створює основу для побудови ефективної інформаційної технології, здатної працювати в реальних умовах експлуатації [21, 22, 23].

### 1.3 Інформаційні технології дистанційної ідентифікації

Сучасні інформаційні технології дистанційної ідентифікації базуються на концепції цифрових двійників, яка передбачає створення віртуальної репліки фізичного об'єкта або процесу, що постійно синхронізується через потоки даних

у реальному часі [3, 5]. Концепція цифрового двійника клієнта (Digital Twin of a Customer, DTOC) розширює класичне розуміння цифрових двійників за рахунок інтеграції різнорідних даних, що відображають поведінку, уподобання та реакції користувачів на основі транзакційних і поведінкових даних [27]. Сучасні інформаційні технології дистанційної ідентифікації базуються на концепції цифрових двійників, яка передбачає створення віртуальної репліки фізичного об'єкта або процесу, що постійно синхронізується через потоки даних у реальному часі [3, 5, 27]. Такі технології інтегрують різнорідні джерела інформації, включаючи відеопотоки, сенсорні дані, транзакційні записи та поведінкові логи, які обробляються за допомогою конвеєрів попередньої обробки (дедуплікація, фільтрація, нормалізація, логарифмічне масштабування) [19, 22]. Ключову роль відіграють архітектури глибокого навчання для моделювання послідовностей, зокрема рекурентні нейронні мережі (LSTM) та трансформери, які дозволяють враховувати довготривалі залежності в часових рядах подій [4, 8, 12, 28, 29]. Розвиток інформаційних технологій комп'ютерного зору тісно пов'язаний із появою великих генеративних моделей та вдосконаленням архітектур виявлення об'єктів. Зокрема, технологія HunyuanVideo [28] демонструє ефективність застосування трансформерних архітектур для моделювання складних просторово-часових залежностей у відеоданих. Сучасні підходи до оцінювання руху та пози, такі як emg2pose [29] та StableAnimator [30], створюють основу для більш точної ідентифікації кінематичних параметрів об'єктів спостереження. Важливе місце в еволюції методів виявлення об'єктів займають гібридні архітектури, такі як DEYO (DETR with YOLO) [31], які поєднують переваги трансформерних детекторів та одноетапних CNN-моделей. Подальший розвиток сімейства YOLO [32, 33] забезпечує баланс між швидкодією та точністю детектування, що є критичним для систем реального часу. Подальший розвиток сімейства YOLO [32, 33] забезпечує баланс між швидкодією та точністю детектування, що є критичним для систем реального часу. Окремим напрямком розвитку інформаційних технологій є синхронізована генерація відео з різних точок огляду, що

реалізовано в підході SynCamMaster [34]. Ця технологія дозволяє створювати багаторакурсні відеопослідовності, що є основою для побудови тривимірних реконструкцій динамічних сцен та підвищення точності просторової локалізації об'єктів. Початкова архітектура FlowNet [35] та її вдосконалена версія FlowNet 2.0 [36] забезпечують високу точність визначення векторів переміщення пікселів між кадрами в реальному часі, що дозволяє ефективно використовувати їх у задачах дистанційної ідентифікації параметрів динамічних об'єктів. Крім того, сучасні методи реконструкції великих сцен, такі як Momentum-GS [37], використовують гауссівську самодистиляцію для підвищення якості відновлення просторової структури, що є важливим для точного визначення глибини та просторових характеристик об'єктів. Важливе значення для ефективної реалізації інформаційних технологій дистанційної ідентифікації має оптимізація обчислювальних процесів. Зокрема, Liger Kernel [38] пропонує ефективні Triton-ядра для прискорення навчання великих мовних моделей, що може бути адаптовано для потреб відеоаналізу. У задачах стереозіставлення та реконструкції глибини важливим є використання методів, стійких до вироджених випадків, таких як Stereo Anywhere [39], який забезпечує робастне нуль-шотове глибоке стереозіставлення навіть там, де стандартні стерео- або монокулярні методи дають збій. Для моделювання довготривалих просторово-часових залежностей у послідовностях даних перспективним є використання еволюційної універсальної трансформерної пам'яті (Evolved Universal Transformer Memory) [40], яка дозволяє ефективно зберігати та оновлювати контекстну інформацію про стан динамічних об'єктів протягом тривалих часових інтервалів. Методи аналізу руху, зокрема класичний ітераційний метод реєстрації зображень Лукаса-Канаде [41], заклали основу для сучасних підходів до оцінювання оптичного потоку та відстеження об'єктів. Розвиток інформаційних технологій також тісно пов'язаний із застосуванням штучного інтелекту в суміжних галузях. Наприклад, технологія ResJobFit [42] демонструє ефективність end-to-end нейронних мереж для задач зіставлення резюме та вакансій, що підтверджує універсальність підходів глибокого навчання. Крім

того, моделі вибору інструментів штучного інтелекту [43] дозволяють оптимізувати процеси розробки програмного забезпечення для систем дистанційної ідентифікації, забезпечуючи обґрунтований вибір архітектур та методів відповідно до конкретних умов застосування.

На рис. 1.1 наведено узагальнену еволюцію класичних методів виявлення об'єктів, що демонструє перехід від каскадних детекторів до ознакових і деформованих моделей. наведено узагальнену еволюцію класичних методів виявлення об'єктів, що демонструє перехід від каскадних детекторів до ознакових і деформованих моделей [1, 14].

Ранній етап цієї еволюції представлений каскадним детектором Віолі-Джонса, який використовує інтегральне представлення зображення та набір слабких класифікаторів для швидкого виявлення облич [1, 44].

Подальший розвиток пов'язаний із появою методу гістограм орієнтованих градієнтів, який забезпечує більш стійке виділення форм об'єктів на основі локальних градієнтів інтенсивності [4, 14].

Наступним важливим кроком стали моделі деформованих частин, що дозволяють враховувати внутрішню варіативність форми об'єкта за рахунок представлення його у вигляді набору частин, з'єднаних пружними зв'язками [14, 45].

Завершальним етапом класичного підходу стало створення агрегованих канальних ознак, що об'єднують кілька типів дескрипторів для підвищення точності детектування в реальному часі [14].

Усі ці методи були детально проаналізовані в оглядових роботах [1, 14, 44], а їхні порівняльні характеристики наведено в дослідженнях [14, 45, 46].

З появою глибоких згорткових нейронних мереж класичні підходи поступово втратили лідерство, проте закладені в них принципи каскадної обробки та використання градієнтних ознак вплинули на розвиток архітектур сучасних детекторів, зокрема R-CNN, YOLO та DETR [14, 45- 47].

В кожній фігурі овалу записані назви методів. Наступним абзацом опис чи пояснення назви. І також показані ключові особливості цих методів.

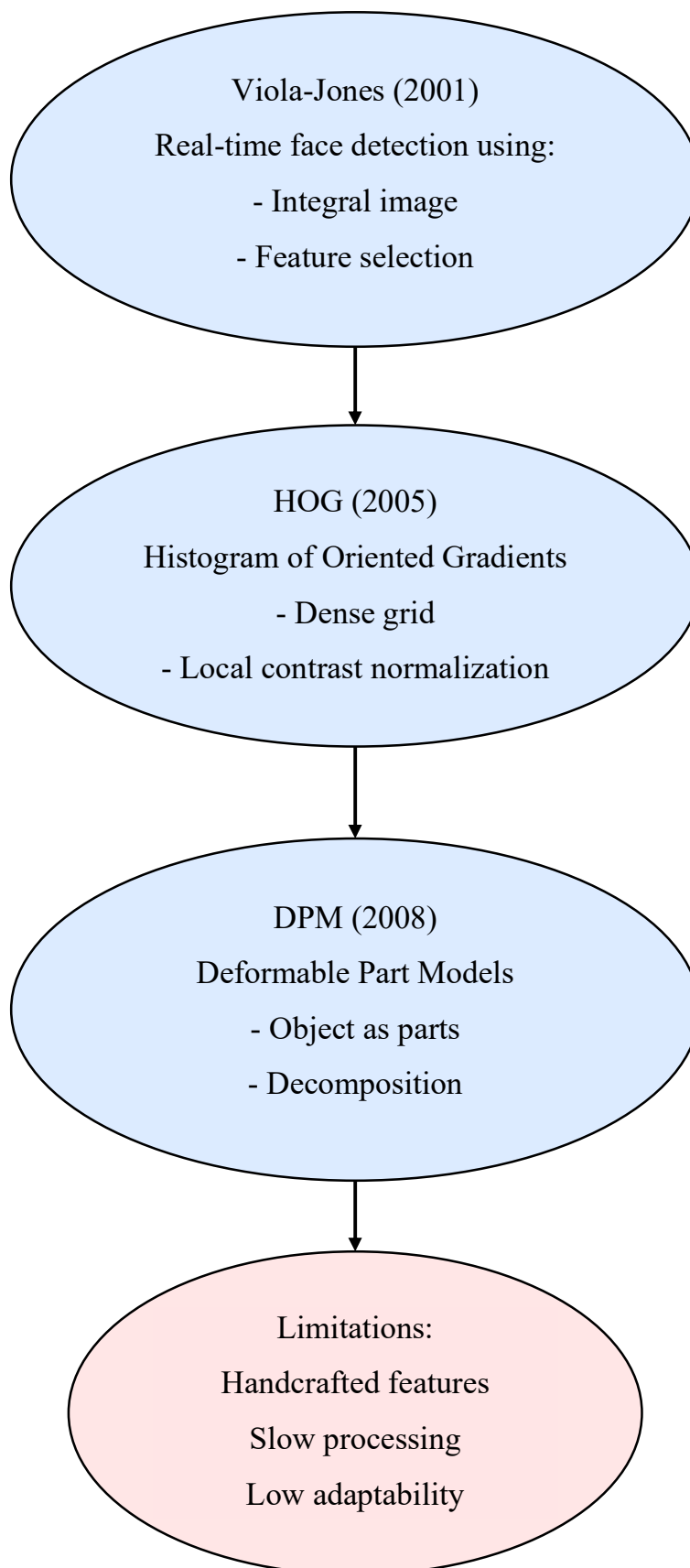


Рисунок 1.1 – Еволюція класичних методів виявлення об'єктів

На рис. 1.2 показано розвиток одноетапних CNN-детекторів, які забезпечують швидке прогнозування класів і координат об'єктів в одному проході мережі.

Вони забезпечують швидке прогнозування класів і координат об'єктів в одному проході мережі [14, 46].

Раннім представником цього напрямку є детектор YOLO, який виконує регресію обмежувальних рамок безпосередньо з повнозв'язних шарів [5, 32, 33]. Альтернативним підходом став метод SSD, що використовує набір згорткових фільтрів різного масштабу для виявлення об'єктів різних розмірів на одній карті ознак [14, 33].

Подальший розвиток пов'язаний із появою RetinaNet, який впровадив фокальну функцію втрат для подолання проблеми дисбалансу між фоновими та цільовими пікселями [14].

Особливе місце в еволюції займають кутові та центральні детектори CornerNet і CenterNet, що відмовилися від традиційних якірних рамок на користь виявлення об'єктів як пар кутових точок або центрів об'єктів [14, 31].

Завершальним етапом наведеної хронології є трансформерні детектори DETR та Deformable DETR, які розглядають виявлення об'єктів як задачу прогнозування набору об'єктів без потреби в механізмах придушення не-максимумів [14, 31, 46, 48].

Таким чином, еволюція одноетапних детекторів демонструє перехід від класичних CNN-архітектур, які використовували якірні рамки та регресію координат (YOLO, SSD), до більш гнучких кутових і центральних методів (CornerNet, CenterNet), а згодом – до трансформерних підходів (DETR, Deformable DETR), що відмовляються від попереднього визначення областей-кандидатів та механізмів придушення не-максимумів [5, 14, 31, 32, 33, 46, 48].

В першому прямокутнику в горі назва методів, в лапках на основі якої мережі. В другому вертикальному ряду прямокутників записана назва методу. А в овалах коротенькі особливості.



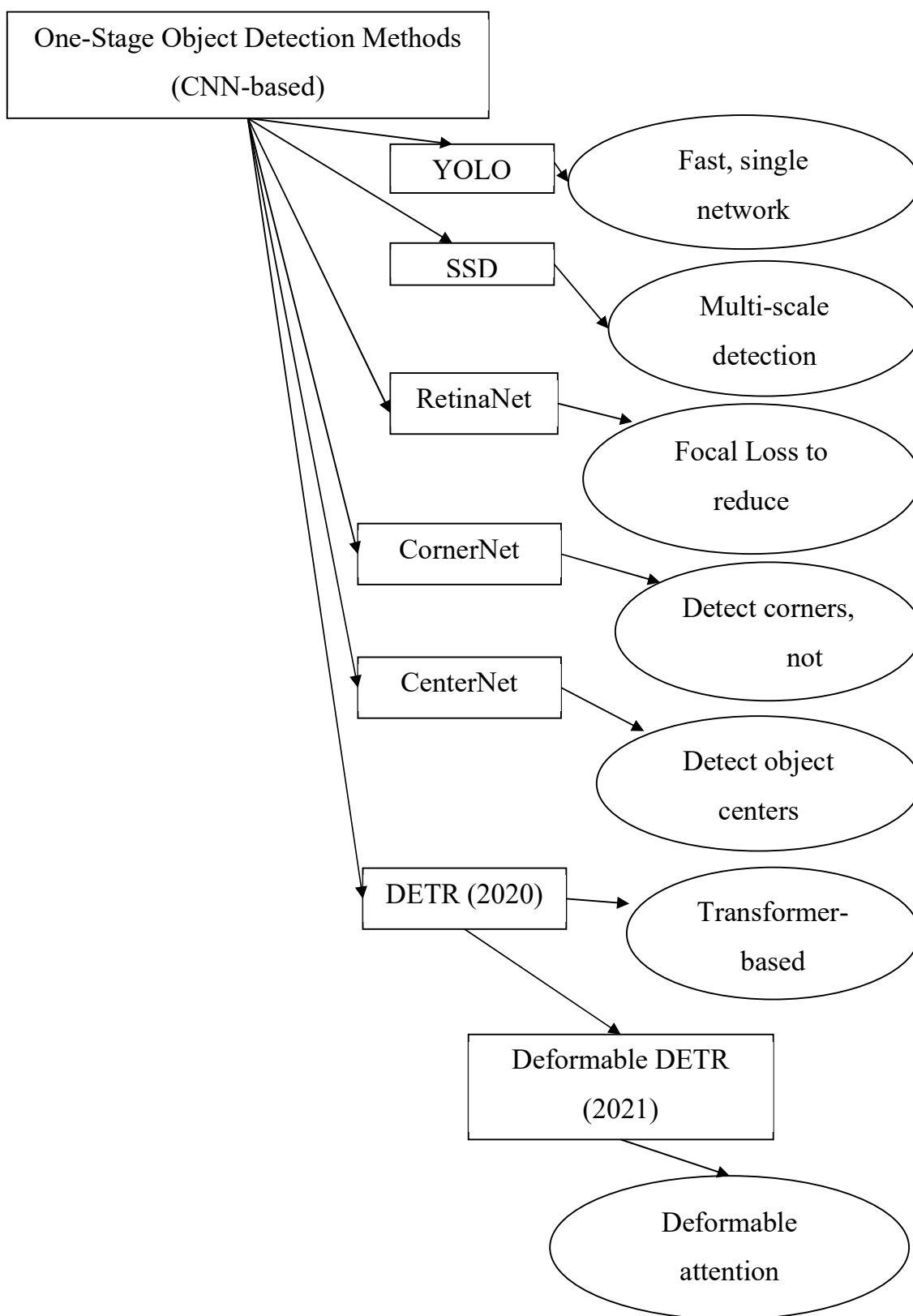


Рисунок 1.2 – Еволюція одноетапних методів виявлення об’єктів

На рис. 1.3 подано показано еволюцію двоетапних архітектур R-CNN та пов'язаних моделей детектування об'єктів, у яких область пропозицій, згорткові ознаки та пірамідальні представлення використовуються для підвищення точності детектування [14, 45].

Верхній прямокутник містить назви базових методів: R-CNN на основі згорткової нейронної мережі, SPP-Net на основі просторової піраміди, Fast R-CNN на основі області інтересу, Faster R-CNN на основі мережі формування пропозицій, Mask R-CNN на основі сегментаційної маски та Cascade R-CNN на основі каскадної регресії [45, 49].

У вертикальному ряду прямокутників наведено послідовність удосконалень: R-CNN використовує селективний пошук для генерації областей-кандидатів [45], SPP-Net впроваджує просторову пірамідну пулінг для роботи з довільним розміром зображення [14], Fast R-CNN об'єднує класифікацію та регресію обмежувальних рамок у спільну багатозадачну функцію втрат [14, 45], Faster R-CNN замінює селективний пошук на навчену мережу формування пропозицій [14, 49], Mask R-CNN додає гілку сегментації для виділення контурів об'єктів [14], а Cascade R-CNN використовує послідовність уточнювальних регресорів для підвищення точності локалізації [14]. В овалах зафіксовано ключові особливості кожного методу, зокрема застосування області інтересу, пірамідальних мереж ознак, багатозадачного навчання, уваги до просторових координат та каскадної обробки пропозицій [14, 45, 49]. Хронологічний розвиток двоетапних детекторів демонструє поступове підвищення точності виявлення об'єктів за рахунок вдосконалення механізмів генерації регіонів-кандидатів та інтеграції додаткових задач, таких як сегментація об'єктів у моделі Mask R-CNN [14, 45]. Ключовим недоліком двоетапних архітектур є менша швидкодія порівняно з одноетапними детекторами через необхідність послідовного виконання етапу формування пропозицій та етапу їх класифікації з регресією рамок [14, 49]. Всі ці моделі заклали основу для подальшого розвитку трансформерних детекторів DETR, які відмовилися від явного етапу формування пропозицій на користь механізмів багатоголової уваги [46, 49].

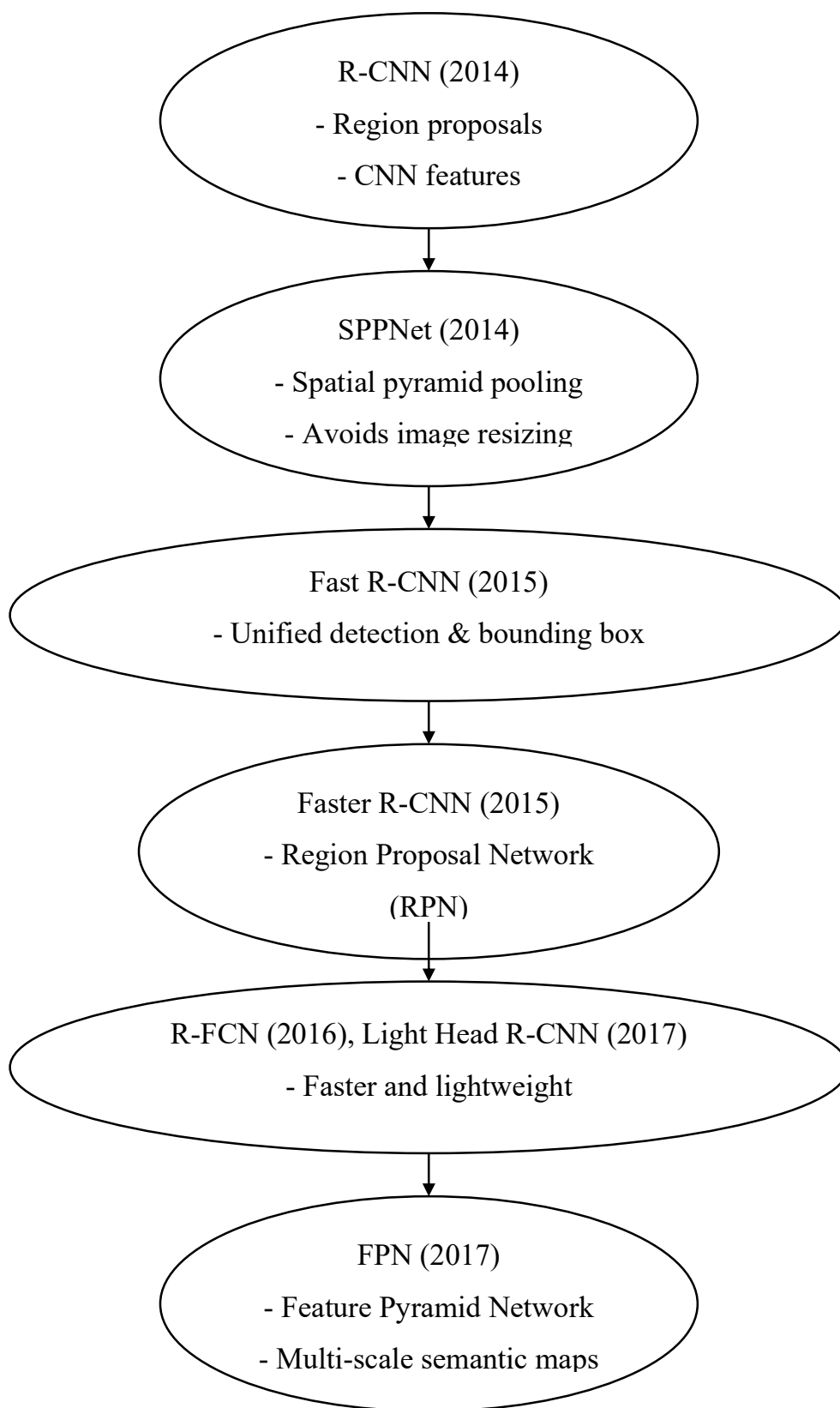


Рисунок 1.3 – Еволюція архітектур R-CNN та пов'язаних моделей детектування об'єктів

#### 1.4 Моделі динамічних об'єктів

Модель динамічного об'єкта в межах задач дистанційної ідентифікації описує зміну його просторового положення, швидкості, прискорення та орієнтації в часі [3, 6]. Для відеоданих та сенсорних потоків така модель поєднує кінематичні параметри об'єкта з ознаками, що спостерігаються у кадрах: контуром, областю локалізації, траєкторією руху, глибиною сцени та векторами оптичного потоку [2, 7, 22]. Це дає змогу перейти від простого виявлення об'єкта до кількісного оцінювання його стану [14, 18].

У загальному випадку динамічний об'єкт можна розглядати як систему зі станом, який змінюється під дією власної динаміки та зовнішнього середовища. Стан містить координати, швидкість, напрям руху, геометричні параметри та ознаки зовнішнього вигляду. Відеопотік дає змогу спостерігати лише проєкцію цього стану на площину кадру, тому для повнішої ідентифікації необхідно доповнювати детекцію методами оцінювання глибини та руху.

Для практичної реалізації модель динамічного об'єкта повинна бути стійкою до шумів, зміни масштабу, часткових оклюзій, нерівномірного освітлення та неповноти даних. Тому в дисертаційній роботі модель розглядається не як ізольоване рівняння руху, а як інформаційна структура, що об'єднує результати DETR, оптичного потоку, GeoNet, моделей глибини та ансамблевого узгодження параметрів.

Сучасні підходи до моделювання динамічних об'єктів активно розвиваються в контексті моніторингу навколишнього середовища. Наприклад, інтелектуальна технологія моніторингу земного покриття [50] використовує гібридну архітектуру згорткових нейронних мереж (CNN + EfficientNet-Edge) для виявлення змін, що демонструє ефективність глибокого навчання для вирішення задач просторово-часового аналізу. В задачах автономного транспорту значну увагу приділено методам багатооб'єктного виявлення та відстеження, зокрема огляд [51] систематизує підходи до обчислення спорідненості об'єктів та їх мультимодального застосування. Ключову роль у сучасному відстеженні об'єктів відіграють трансформерні архітектури, які, як

зазначено в огляді [52], забезпечують потужне моделювання просторово-часових залежностей, що дозволяє ефективно вирішувати проблеми оклюзій, схожих об'єктів та змін зовнішнього вигляду. Розвиток нейромережових моделей для аналізу динамічних об'єктів включає також спеціалізовані архітектури, такі як GeONet [53] – нейронний оператор для навчання вассерштейнових геодезичних ліній, що може бути використано для моделювання неперервних перетворень форми та просторових характеристик об'єктів. У задачах відстеження та класифікації об'єктів застосовуються методи пошуку на основі агентів [54], які дозволяють ефективно локалізувати та супроводжувати динамічні об'єкти в складних сценах. Для ідентифікації стану комп'ютерних систем та підтримки прийняття рішень використовуються методи побудови нечітких дерев рішень [55], які забезпечують інтерпретованість результатів та стійкість до невизначеності, що є важливим для систем дистанційної ідентифікації параметрів динамічних об'єктів.

У моделюванні динамічних об'єктів є геометричне глибоке навчання [56], яке дозволяє ефективно працювати з даними на многовидах та графах, що відкриває нові можливості для опису форми та просторової структури об'єктів. Важливим аспектом є також прогнозування траєкторій руху, якому присвячено сучасні огляди [57], що систематизують методи на основі рекурентних нейронних мереж, трансформерів та графових нейронних мереж. Сучасні підходи до оцінювання параметрів у генеративних дифузійних моделях [58] дозволяють ефективно визначати динамічні характеристики об'єктів з високою точністю. Для моделювання взаємодії кількох об'єктів у сцені застосовуються методи багатоагентного навчання з підкріпленням [59], які дозволяють враховувати складні просторово-часові залежності між об'єктами та прогнозувати їхню спільну поведінку.

Фундаментальні методи аналізу руху, зокрема глобальний метод Хорна-Шунка [60], заклали основу для сучасних підходів до оцінювання оптичного потоку. Подальший розвиток класичних методів пов'язаний з поліноміальним підходом Фарнебека [61], який забезпечує високу точність та стійкість до шумів

за рахунок апроксимації інтенсивності пікселів поліномами другого порядку та використання пірамідальної структури зображень, що дозволяє ефективно обробляти як повільні, так і швидкі переміщення об'єктів у кадрі. Цей метод широко використовується в практичних застосуваннях завдяки своїй обчислювальній ефективності та реалізації в бібліотеці OpenCV.

Моделі динамічних об'єктів базуються на методах оцінювання оптичного потоку, які дозволяють кількісно описувати рух об'єктів у відеопослідовності. Класичні підходи Хорна-Шунка [60] та Фарнебека [61] забезпечують обчислення векторів переміщення пікселів між кадрами, створюючи основу для визначення швидкості, напрямку руху та траєкторії об'єктів. Для підвищення надійності ідентифікації в умовах невизначеності застосовуються ансамблеві методи [63], зокрема еволюційний підхід до формування ансамблів [62], який дозволяє комбінувати результати різних алгоритмів для досягнення більш стійких та точних оцінок параметрів динамічних об'єктів [60-63].

Розвиток моделей динамічних об'єктів характеризується інтеграцією трансформерних архітектур, методів глибинного навчання та ансамблевих підходів. Значних успіхів досягнуто в задачах 3D виявлення динамічних об'єктів за допомогою мультимодальних мереж, таких як LiDAR-Camera Fusion Network [64], яка поєднує трансформери та Mamba для ефективного виявлення пішоходів, транспортних засобів та велосипедистів у тривимірному просторі. Важливим напрямком є також всеракурсний зір, огляд якого [65] систематизує сучасні методи представлення, оптимізації та застосування для оцінювання глибини та оптичного потоку в панорамних зображеннях.

Монокулярна оцінка глибини отримала потужний імпульс завдяки розробці ефективних моделей, таких як EfficientDepth [67], яка поєднує трансформерну архітектуру з легким згортковим декодером для отримання деталізованих карт глибини з мінімальними обчислювальними витратами. У задачах стереоскопічного 3D виявлення об'єктів значної уваги заслуговує StereoDETR [70], який вперше досягає швидкості реального часу, перевершуючи монокулярні методи за продуктивністю. Для багатооб'єктного відстеження в 3D

пропонується LiDAR MOT-DETR [68], що використовує двоетапну трансформерну архітектуру для згладжування детекцій та асоціації треків.

Окремим напрямком є сегментація динамічних об'єктів у відео на основі руху. Метод MOVE [69] пропонує новий підхід до сегментації об'єктів за їхніми патернами руху з використанням небагатьох прикладів, що особливо актуально для сцен з динамічними об'єктами. Ансамблеві методи відіграють ключову роль у підвищенні надійності моделей: гібридний ансамбль [66] демонструє переваги комбінування різних архітектур, а дослідження Bashar et al. [72] досліджує синергетичне об'єднання CNN, MLP-Mixer та Vision Transformer для покращення класифікації зображень. Комплексний огляд еволюції YOLO [71] охоплює десять років розвитку цієї родини детекторів, від YOLOv1 до YOLOv12, аналізуючи внесок кожної версії у підвищення швидкості, точності та обчислювальної ефективності, що має безпосереднє значення для моделювання динамічних об'єктів у реальному часі [64-72].

### 1.5 Методи обчислювального інтелекту в задачах ідентифікації

Методи обчислювального інтелекту застосовуються для підвищення стійкості ідентифікації в умовах шумів, оклюзій, неповних даних та зміни освітлення. У роботі особливу увагу приділено ансамблевим підходам, зокрема бегінгу та бустінгу, які дають змогу зменшити дисперсію оцінок, уточнити межі об'єктів і підвищити надійність визначення параметрів руху. Поєднання цих методів із нейромережевими моделями комп'ютерного зору формує основу адаптивної інформаційної технології.

Бегінг використовується для стабілізації результатів шляхом усереднення оцінок, отриманих на різних підвбірках або модифікаціях кадрів. Такий підхід зменшує вплив випадкових шумів та локальних артефактів. Бустінг, навпаки, послідовно концентрує увагу на областях із найбільшою похибкою, що є корисним для уточнення меж об'єктів, зон оклюзій і різких перепадів глибини.

Трансформерні моделі доповнюють ансамблеві методи здатністю враховувати глобальний контекст сцени та довготривалі просторово-часові

залежності. У задачах ДПДО це дає змогу поєднати локальні вимірювання руху з контекстною інформацією про клас, положення та взаємодію об'єктів у кадрі, що підвищує точність ідентифікації параметрів у складних динамічних сценах.

Методи обчислювального інтелекту відіграють ключову роль у підвищенні ефективності систем дистанційної ідентифікації параметрів динамічних об'єктів, особливо в умовах невизначеності, шумів та неповноти даних. Важливим напрямом є застосування машинного навчання для аналізу та верифікації моделей складних систем. Зокрема, Копп та ін. [73] пропонують підхід до виявлення структурних помилок у BPMN-моделях бізнес-процесів з використанням методів машинного навчання, що дозволяє автоматизувати контроль коректності моделей та зменшити трудовитрати на їх перевірку. Розвиток цього напрямку представлено в роботі Коппа, Орловського та Літвінової [74], де досліджується програмне рішення для аналізу коректності бізнес-процесів, що базується на інтеграції методів машинного навчання з правилами верифікації BPMN-нотації. Ці підходи можуть бути адаптовані для задач дистанційної ідентифікації, де необхідно забезпечити коректність взаємодії модулів виявлення об'єктів, оптичного потоку та оцінювання глибини в єдиному інформаційному просторі [73, 74]. Окремого розгляду заслуговують методи оцінювання ефективності комп'ютерних динамічних симуляторів, які запропоновано в дослідженні Shmatko та ін. [75]. Розроблена ними методика дозволяє кількісно оцінювати продуктивність технічних систем у динамічних режимах роботи, що є важливим для налаштування параметрів систем дистанційної ідентифікації та оптимізації їх роботи в реальних умовах експлуатації. Важливе значення для вибору методів обробки даних у системах дистанційної ідентифікації мають порівняльні дослідження алгоритмів машинного навчання. Зокрема, Шматко та ін. [76] провели порівняльний аналіз методів машинного навчання для задач прогнозування, що дозволяє обґрунтовано обирати оптимальні алгоритми залежно від специфіки даних та вимог до точності. Розвиток цього напрямку знайшов відображення в роботі Шматко, Малишенка та Волощука [77], де запропоновано архітектуру



інформаційної системи для інтелектуальної класифікації клієнтів з використанням сучасних методів машинного навчання, що демонструє ефективність застосування обчислювального інтелекту в задачах класифікації та прогнозування.

Універсальність підходів глибокого навчання підтверджується їх успішним застосуванням у різноманітних предметних галузях. Наприклад, технологія ResJobFit [78] демонструє ефективність end-to-end нейронних мереж для задач зіставлення резюме та вакансій, що свідчить про здатність глибоких нейронних мереж виявляти складні залежності в структурованих та неструктурованих даних. Цей досвід може бути використаний при побудові систем дистанційної ідентифікації, де необхідно встановлювати відповідності між різнорідними ознаками динамічних об'єктів, отриманими з різних сенсорів або методів обробки [78]. Поєднання методів машинного навчання для верифікації моделей [73, 74], порівняльного аналізу алгоритмів [76, 77], оцінювання ефективності динамічних симуляторів [75] та спеціалізованих end-to-end нейромережових технологій [78] створює методологічну основу для побудови адаптивних інформаційних технологій дистанційної ідентифікації, здатних до самонавчання та корекції параметрів в умовах динамічної зміни зовнішнього середовища [73-78].

Застосування методів штучного інтелекту в системах безпеки та моніторингу навколишнього середовища є важливим напрямом розвитку обчислювального інтелекту. Зокрема, у дослідженні Kozlov та Malakhov [79] запропоновано комплексну модель розпізнавання людини в надзвичайних ситуаціях з використанням багаторівневої системи дронів, що поєднує методи комп'ютерного зору, машинного навчання та просторово-часового аналізу для оперативного виявлення людей у складних умовах. Цей підхід демонструє ефективність інтеграції різнорідних методів обчислювального інтелекту для вирішення задач дистанційної ідентифікації динамічних об'єктів у реальному часі. Іншим прикладом є вдосконалення інформаційної технології аналізу контекстно-освічених подань фраз [80], де застосовуються методи глибокого

навчання для покращення якості обробки природної мови, що може бути використано для розробки інтерфейсів взаємодії з системами дистанційної ідентифікації. У сфері моніторингу навколишнього середовища Kashtan та Hnatushenko [81] запропонували інтелектуальну технологію моніторингу земного покриву на основі оптичних супутникових зображень, яка використовує гібридну архітектуру згорткових нейронних мереж для виявлення змін, що підтверджує ефективність методів глибокого навчання для вирішення задач просторово-часового аналізу [81]. Поєднання цих підходів з методами оцінювання ефективності динамічних симуляторів [75] та порівняльного аналізу алгоритмів [76, 77] створює методологічну основу для побудови адаптивних інформаційних технологій дистанційної ідентифікації, здатних до самонавчання та корекції параметрів в умовах динамічної зміни зовнішнього середовища [73-81].

Медицина діагностика є ще однією важливою сферою застосування методів обчислювального інтелекту. Aziukovskyi та ін. [82] запропонували моделі комп'ютерно-асоційованої діагностики для систем підтримки прийняття рішень при виявленні раку молочної залози, що демонструє ефективність застосування методів глибокого навчання для аналізу медичних зображень. У задачах прогнозування природних явищ Hnatushenko та ін. [83] розробили метод прогнозування повеней на основі супутникових зображень Sentinel-2 з використанням машинного навчання, що дозволяє оперативно оцінювати ризики та приймати управлінські рішення. Важливе значення для управління якістю інформаційних систем мають дослідження Годлевського та Бурлакова [84], які запропонували інформаційну технологію планування підвищення якості підмножини процесів моделі SPICE, що дозволяє оптимізувати процеси розробки та супроводу програмного забезпечення для систем дистанційної ідентифікації. Поєднання цих підходів [73-84] створює методологічну основу для побудови адаптивних інформаційних технологій дистанційної ідентифікації, здатних до самонавчання та корекції параметрів в умовах динамічної зміни зовнішнього середовища.

Важливе значення для управління якістю інформаційних систем мають дослідження Годлевського та Бурлакова [84, 85], які запропонували інформаційну технологію планування підвищення якості підмножини процесів моделі SPICE [84] та її вербальний опис [85], що дозволяє оптимізувати процеси розробки та супроводу програмного забезпечення для систем дистанційної ідентифікації. Розвиток цього напрямку представлено в роботі Сокола, Годлевського та Малеця [86], де запропоновано метод оцінки якості процесу розробки програмного забезпечення ІТ-компанії на основі використання функції корисності, що забезпечує кількісну оцінку ефективності та може бути адаптований для оцінювання якості роботи систем дистанційної ідентифікації. Поєднання цих підходів [73-86] створює методологічну основу для побудови адаптивних інформаційних технологій дистанційної ідентифікації, здатних до самонавчання та корекції параметрів в умовах динамічної зміни зовнішнього середовища.

Важливе значення для управління якістю інформаційних систем мають дослідження Годлевського та Бурлакова [84, 85], які запропонували інформаційну технологію планування підвищення якості підмножини процесів моделі SPICE [84] та її вербальний опис [85], що дозволяє оптимізувати процеси розробки та супроводу програмного забезпечення для систем дистанційної ідентифікації. Розвиток цього напрямку представлено в роботі Сокола, Годлевського та Малеця [86], де запропоновано метод оцінки якості процесу розробки програмного забезпечення ІТ-компанії на основі використання функції корисності, що забезпечує кількісну оцінку ефективності та може бути адаптований для оцінювання якості роботи систем дистанційної ідентифікації. Поєднання цих підходів [73 – 86] створює методологічну основу для побудови адаптивних інформаційних технологій дистанційної ідентифікації, здатних до самонавчання та корекції параметрів в умовах динамічної зміни зовнішнього середовища. Прогнозування траєкторій рухомих об'єктів є критично важливим завданням для систем автономного транспорту, відеоспостереження та робототехніки, де застосування методів обчислювального інтелекту дозволяє

передбачати поведінку динамічних об'єктів у складних сценаріях. Як зазначають [87] у своєму огляді, останнім часом значного прогресу досягнуто завдяки застосуванню глибокого навчання, зокрема трансформерних архітектур, графових нейронних мереж та методів імітаційного навчання, які дозволяють моделювати складні просторово-часові залежності в поведінці пішоходів та інших динамічних об'єктів. Окремий інтерес становлять моделі, що враховують соціальну взаємодію між об'єктами та адаптуються до різноманітних сценаріїв руху.

Просторова локалізація динамічних об'єктів тісно пов'язана з оцінюванням глибини сцени, яка дозволяє визначати координату  $Z$  об'єкта в тривимірному просторі. Сучасний стан досліджень у цій галузі узагальнено в огляді Xu та ін. [88], який аналізує еволюцію архітектур глибокого навчання для оцінювання глибини в монокулярному, стерео, мультимедійному та відео режимах, а також окреслює шляхи створення фундаментальних моделей глибини (depth foundation models) з високою здатністю до нуль-шотової генералізації. Для задач дистанційного зондування Землі особливий інтерес становить метод D3-RSMDE [89], який демонструє прискорення понад 40 разів порівняно з дифузійними моделями завдяки поєднанню швидкого ViT-модуля для генерації структурного пріоритету глибини та легкої U-Net для уточнення деталей із використанням стратегії прогресивного лінійного змішування (PLBR), що дозволяє отримувати високоякісні карти глибини на супутникових знімках в реальному часі. Перспективним напрямком є також адаптація візійно-мовних моделей для задач оцінювання глибини. Метод MoA-DepthCLIP [90] демонструє, що використання легких модулів адаптерів дозволяє ефективно переносити знання з CLIP для монокулярної оцінки глибини, досягаючи значного підвищення точності при мінімальній кількості параметрів, що навчаються. Поєднання цих підходів [73-90] створює методологічну основу для побудови адаптивних інформаційних технологій дистанційної ідентифікації, здатних до самонавчання та корекції параметрів в умовах динамічної зміни зовнішнього середовища.

## 1.6 Мета та задача дослідження

Метою дослідження є підвищення точності, стійкості та інформативності дистанційної ідентифікації параметрів динамічних об'єктів шляхом розроблення моделей, методів та інформаційної технології, що поєднують трансформерне виявлення об'єктів, алгоритми оптичного потоку, нейромережеве оцінювання глибини й ансамблеву обробку результатів.

Для досягнення поставленої мети необхідно розв'язати такі задачі:

1. Проаналізувати сучасні методи дистанційної ідентифікації параметрів динамічних об'єктів.

2. Сформулювати класифікацію параметрів і методів їх визначення. Розробити моделі інтеграції детектора об'єктів, оптичного потоку та оцінювання глибини.

3. Обґрунтувати використання ансамблевих методів для підвищення стійкості результатів.

4. Реалізувати програмні компоненти інформаційної технології. Провести експериментальне тестування та порівняльний аналіз отриманих результатів.

Розв'язання цих задач дозволяє перейти від окремих алгоритмів комп'ютерного зору до цілісної інформаційної технології, здатної визначати положення, швидкість, напрям руху, траєкторію та просторові характеристики об'єктів у відеопотоці в умовах шумів, оклюзій та зміни зовнішнього середовища.

Задача дослідження присвячена розробці моделей, методів та інформаційної технології дистанційної ідентифікації параметрів динамічних об'єктів за відеоданими, які поєднують трансформерні моделі виявлення об'єктів, алгоритми оптичного потоку та методи обчислювального інтелекту.

## 1.7 Висновки до розділу

1. За результатами аналізу предметної області ДПДО визначено, що динамічні об'єкти є складними багатовимірними системами, стан яких безперервно змінюється в часі під впливом внутрішніх процесів або зовнішнього

середовища. Показано, що сучасні інформаційні технології дистанційної ідентифікації повинні забезпечувати не лише виявлення об'єкта, а й кількісний опис його просторових, кінематичних та структурних характеристик в умовах невизначеності (шумів, оклюзій, зміни освітлення). Обґрунтовано актуальність розробки інтегрованих підходів, які поєднують фізичні моделі, алгоритми комп'ютерного зору та методи обчислювального інтелекту.

2. Сформовано класифікацію параметрів динамічних об'єктів, згідно з якою просторові (координати, глибина, розміри, орієнтація), кінематичні (швидкість, прискорення, напрям руху, траєкторія) та структурні (контур, площа, клас, ознаки форми) параметри потребують поєднання різних методів визначення. За способом отримання параметри поділено на безпосередньо вимірювані (координати рамки, розміри, контур, клас) та похідні (швидкість, прискорення, напрям, глибина). За часовим режимом методи класифіковано на покадрові, послідовні та інтегровані, причому інтегрований підхід визначено як базовий для подальшої побудови інформаційної технології.

3. Проаналізовано еволюцію інформаційних технологій комп'ютерного зору від класичних методів (Біоли-Джонса, ГОГ, моделі деформованих частин) до сучасних архітектур глибокого навчання, зокрема одноетапних CNN-детекторів (YOLO, SSD, RetinaNet, CornerNet, CenterNet), двоетапних архітектур R-CNN та трансформерних детекторів (DETR, Deformable DETR). Розглянуто технології оцінювання оптичного потоку (FlowNet, FlowNet 2.0, RAFT) та глибини (MiDaS, DPT\_Large, Stereo Anywhere, Momentum-GS). Показано, що інтеграція різномірних методів у єдиний конвеєр є ключовим напрямом розвитку інформаційних технологій для задач ДІПДО.

4. Розглянуто сучасні підходи до моделювання динамічних об'єктів, які описують зміну просторового положення, швидкості, прискорення та орієнтації в часі. Показано, що модель динамічного об'єкта в межах задач дистанційної ідентифікації повинна розглядатися не як ізольоване рівняння руху, а як інформаційна структура, що об'єднує результати DETR, оптичного потоку, GeoNet, моделей глибини та ансамблевого узгодження параметрів. Визначено,

що використання трансформерних архітектур та геометричного глибокого навчання дозволяє ефективно моделювати складні просторово-часові залежності.

5. Проаналізовано методи обчислювального інтелекту для підвищення стійкості ідентифікації в умовах шумів, оклюзій та зміни освітлення. Особливу увагу приділено ансамблевим підходам бегінгу (Bagging) та бустінгу (Boosting), які дають змогу зменшити дисперсію оцінок, уточнити межі об'єктів і підвищити надійність визначення параметрів руху. Розглянуто застосування методів машинного навчання для верифікації моделей, оцінювання ефективності динамічних симуляторів та інтелектуальної класифікації. Показано, що поєднання обчислювального інтелекту з нейромережевими моделями комп'ютерного зору формує основу адаптивної інформаційної технології ДПДО.

6. Сформульовано мету дослідження – підвищення точності, стійкості та інформативності дистанційної ідентифікації параметрів динамічних об'єктів шляхом розроблення моделей, методів та інформаційної технології, що поєднують трансформерне виявлення об'єктів, алгоритми оптичного потоку, нейромережеве оцінювання глибини й ансамблеву обробку результатів. Визначено конкретні задачі дослідження: аналіз сучасних методів, класифікація параметрів, розробка моделей інтеграції, обґрунтування ансамблевих методів, програмна реалізація та експериментальне тестування. Узагальнено, що задача ДПДО потребує не лише виявлення об'єкта, а й кількісного опису його стану в просторі та часі. Найперспективнішим напрямом є інтеграція трансформерних детекторів для локалізації, нейромережових методів оптичного потоку для оцінювання руху та моделей глибини для 3D-реконструкції. Ансамблеві методи бегінгу та бустінгу дозволяють підвищити стійкість без значного збільшення обчислювальних витрат. Проведений аналіз створює методичну основу для подальшої розробки моделей, методів та інформаційної технології в наступних розділах дисертації.

## РОЗДІЛ 2 МЕТОДИ ДИСТАНЦІЙНОЇ ІДЕНТИФІКАЦІЇ НА ОСНОВІ КОМП'ЮТЕРНОГО ЗОРУ ТА ГЛИБИННОГО НАВЧАННЯ

У розділі подано методичну основу дистанційної ідентифікації: класичні алгоритми оптичного потоку, нейромережеві моделі аналізу руху, ансамблеві підходи та способи їх інтеграції. Основна увага приділена методам, які забезпечують одночасне врахування просторових і часових характеристик об'єкта.

### 2.1 Методи оптичного потоку в задачах ідентифікації

Алгоритм Горна – Шунка спирається на припущення про плавність потоку над усім зображенням і мінімізує глобальний функціонал енергії, що поєднує обмеження сталості яскравості та регуляризацию гладкості поля руху [41]. Таким чином, він намагається мінімізувати спотворення потоку й віддає перевагу розв'язкам, які демонструють більшу плавність.

Потік формулюють як глобальний функціонал енергії, який потім намагаються мінімізувати. Цю функцію для потоків двовимірних зображень задають як функціонал енергії

$$E = \iint (I_x u + I_y v + I_t)^2 + \alpha^2 (|\nabla u|^2 + |\nabla v|^2) dx dy \quad (2.1)$$

де  $I_x, I_y, I(t)$  – похідні інтенсивності зображення за координатами  $x, y$  та часом  $t$  відповідно;  $u$  і  $v$  – компоненти вектора оптичного потоку;  $\alpha$  – стала регуляризації. Більші значення  $\alpha$  призводять до плавнішого поля потоку. Цей функціонал можна мінімізувати, розв'язуючи пов'язані рівняння Ейлера–Лагранжа.

$$I_x u I_y v I_t \vec{V} = [u(x, y), v(x, y)]^T \alpha \alpha$$

Рівняння Ейлера–Лагранжа:

$$\frac{\partial L}{\partial u} - \frac{\partial}{\partial x} \frac{\partial L}{\partial u_x} - \frac{\partial}{\partial y} \frac{\partial L}{\partial u_y} = 0 \quad (2.2)$$



$$\frac{\partial L}{\partial v} - \frac{\partial}{\partial x} \frac{\partial L}{\partial v_x} - \frac{\partial}{\partial y} \frac{\partial L}{\partial v_y} = 0 \quad (2.3)$$

де  $L$  є підінтегральним виразом функціонала енергії, що дає систему рівнянь для компонентів оптичного потоку.

Отримана система рівнянь:

$$I_x(I_x u + I_y v + I_t) - \alpha^2 \Delta u = 0 \quad (2.4)$$

$$I_y(I_x u + I_y v + I_t) - \alpha^2 \Delta v = 0 \quad (2.5)$$

де індекси позначають частинне диференціювання, а  $\Delta$  позначає оператор Лапласа. На практиці лапласіан наближують чисельно за допомогою скінченних різниць і записують як дискретизацію лапласіана.

$$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$$

$$\Delta u(x, y) = 4(\bar{u}(x, y) - u(x, y)) \quad (2.6)$$

де  $\bar{u}$  та  $\bar{v}$  – середньозважені усереднення компонентів  $u$  і  $v$ , обчислені в околі пікселя з координатами  $(x, y)$ . Використовуючи це позначення, наведену вище систему рівнянь можна записати як лінійну систему.

$$(I_x^2 + 4\alpha^2)u + I_x I_y v = 4\alpha^2 \bar{u} - I_x I_t \quad (2.7)$$

$$I_x I_y u + (I_y^2 + 4\alpha^2)v = 4\alpha^2 \bar{v} - I_y I_t \quad (2.8)$$

Отримана система (2.7) та (2.8) є лінійною за  $u$  та  $v$  і може бути розв'язана для кожного пікселя зображення. Проте, оскільки розв'язок залежить від сусідніх значень поля потоку, обчислення повторюють після уточнення сусідів. Наступна ітераційна схема виводиться за допомогою правила Крамера:  $uv$ . Ітераційна схема (метод Якобі)

$$u^{k+1} = \bar{u}^k - \frac{I_x(I_x \bar{u}^k + I_y \bar{v}^k + I_t)}{4\alpha^2 + I_x^2 + I_y^2} \quad (2.9)$$

$$v^{k+1} = \overline{v^k} - \frac{I_y(I_x \overline{u^k} + I_y \overline{v^k} + I_t)}{4\alpha^2 + I_x^2 + I_y^2} \quad (2.10)$$

де верхній індекс  $k + 1$  позначає наступну ітерацію, яку потрібно обчислити, а  $k$  – попередній результат. Це, по суті, метод розщеплення матриць, подібний до методу Якобі, застосованого до великої розрідженої системи, що виникає при розв’язуванні для всіх пікселів одночасно. Рівняння оптичного потоку (для кожного пікселя)

$$I_x(q_i)V_x + I_y(q_i)V_y = -I_t(q_i) \quad (2.11)$$

Матрична форма

$$A v = b \quad (2.12)$$

Матриця  $A$

$$A = \begin{bmatrix} I_x(q_1) & I_y(q_1) \\ I_x(q_2) & I_y(q_2) \\ \vdots & \vdots \\ I_x(q_n) & I_y(q_n) \end{bmatrix} \quad (2.13)$$

Вектор швидкості

$$v = \begin{bmatrix} V_x \\ V_y \end{bmatrix} \quad (2.14)$$

Вектор правої частини

$$b = \begin{bmatrix} I_t(q_1) \\ I_t(q_2) \\ \vdots \\ I_t(q_n) \end{bmatrix} \quad (2.15)$$

Розв’язок методом найменших квадратів

$$A^T A v = A^T b \quad (2.16)$$

$$v = (A^T A)^{-1} A^T b \quad (2.17)$$

Розписаний вигляд (без ваг)

$$\begin{bmatrix} V_x \\ V_y \end{bmatrix} = \begin{bmatrix} \sum_i I_x(q_i)^2 & \sum_i I_x(q_i)I_y(q_i) \\ \sum_i I_x(q_i)I_y(q_i) & \sum_i I_y(q_i)^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_i I_x(q_i)I_t(q_i) \\ \sum_i I_y(q_i)I_t(q_i) \end{bmatrix} \quad (2.18)$$

Структурний тензор (матриця  $A^T A$ )

$$A^T A = \begin{bmatrix} \sum_i I_x(q_i)^2 & \sum_i I_x(q_i)I_y(q_i) \\ \sum_i I_x(q_i)I_y(q_i) & \sum_i I_y(q_i)^2 \end{bmatrix} \quad (2.19)$$

Зважений варіант

$$A^T W A v = A^T W b \quad (2.20)$$

$$v = (A^T W A)^{-1} A^T W b \quad (2.21)$$

$$\begin{bmatrix} V_x \\ V_y \end{bmatrix} = \begin{bmatrix} \sum_i w_i I_x(q_i)^2 & \sum_i w_i I_x(q_i)I_y(q_i) \\ \sum_i w_i I_x(q_i)I_y(q_i) & \sum_i w_i I_y(q_i)^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_i w_i I_x(q_i)I_t(q_i) \\ \sum_i w_i I_y(q_i)I_t(q_i) \end{bmatrix} \quad (2.22)$$

Умова розв'язності (власні значення)

$$\lambda_1 \geq \lambda_2 > 0 \quad (2.23)$$

та бажано

$$\lambda_1 / \lambda_2 \approx 1 \quad (2.24)$$

Оптичний потік є фундаментальним поняттям у комп'ютерному зорі, що описує розподіл видимих швидкостей руху об'єктів на зображенні. Нехай  $I(x, y, t)$  – функція яскравості зображення в точці  $(x, y)$  у момент часу  $t$ . Тоді основне рівняння оптичного потоку має вигляд

$$I_x u + I_y v + I_t = 0 \quad (2.25)$$

де  $u = \frac{dx}{dt}$ ,  $v = \frac{dy}{dt}$  – компоненти оптичного потоку, а  $I_x = \frac{\partial I}{\partial x}$ ,  $I_y = \frac{\partial I}{\partial y}$ ,  $I_t = \frac{\partial I}{\partial t}$  – часткові похідні яскравості.

Локальний метод Лукаса-Канаде базується на припущенні, що оптичний

потік є локально постійним у межах невеликого вікна розміром  $N \times N$ . Для пікселів вікна формується система рівнянь

$$\begin{bmatrix} I_x(p_1) & I_y(p_1) \\ I_x(p_2) & I_y(p_2) \\ \vdots & \vdots \\ I_x(p_n) & I_y(p_n) \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = - \begin{bmatrix} I_t(p_1) \\ I_t(p_2) \\ \vdots \\ I_t(p_n) \end{bmatrix} \quad (2.26)$$

Розв'язок цієї системи методом найменших квадратів дає оцінку оптичного потоку

$$\begin{bmatrix} u \\ v \end{bmatrix} = (A^T A)^{-1} A^T b \quad (2.27)$$

де  $A$  – матриця градієнтів інтенсивності,  $b$  – вектор часових різниць. Перевагою методу є його обчислювальна ефективність та стійкість до шумів, однак він не здатен відновлювати потік в областях з однорідною текстурою. Скорочена форма формули (2.1)

$$E = \iint \left[ (I_x u + I_y v + I_t)^2 + \alpha^2 (|\nabla u|^2 + |\nabla v|^2) \right] \quad (2.28)$$

Глобальний метод Хорна-Шунка формули (2.1) перший доданок відповідає за виконання базового рівняння оптичного потоку, а другий – за гладкість поля. Параметр  $\alpha$  регулює баланс між цими двома умовами. Розв'язок знаходиться ітераційно:

$$u^{k+1} = \bar{u}^k - \frac{I_x(I_x \bar{u}^k + I_y \bar{v}^k + I_t)}{\alpha^2 + I_x^2 + I_y^2} \quad (2.29)$$

$$v^{k+1} = \bar{v}^k - \frac{I_y(I_x \bar{u}^k + I_y \bar{v}^k + I_t)}{\alpha^2 + I_x^2 + I_y^2} \quad (2.30)$$

де  $\bar{u}^k, \bar{v}^k$  – усереднені значення потоку на  $k$ -й ітерації.

Поліноміальний метод Farneback

$$f(x) \approx x^T A x + b^T x + c \quad (2.31)$$

Для двох послідовних кадрів припускається, що вони пов'язані зсувом  $d$ :

$f_2(x) = f_1(x - d)$ . Після розкладання та порівняння коефіцієнтів отримуємо:

$$A_2 = A_1 \quad (2.32)$$

$$b_2 = b_1 - 2A_1 d \quad (2.33)$$

$$c_2 = d^T A_1 d - b_1^T d + c_1 \quad (2.34)$$

Звідси вектор зміщення обчислюється як

$$d = -\frac{1}{2} A_1^{-1} (b_2 - b_1) \quad (2.35)$$

Цей метод забезпечує високу точність і є стійким до шумів.

Отже, методи оптичного потоку є важливим інструментом комп'ютерного зору для аналізу руху та ідентифікації об'єктів у відеопослідовностях. Вони базуються на припущенні сталості яскравості пікселів між сусідніми кадрами та дозволяють оцінювати вектор переміщення для кожної точки зображення. Локальний метод Лукаса–Канаде забезпечує ефективне та обчислювально швидке оцінювання руху завдяки припущенню про локальну сталість потоку в межах невеликого вікна. Метод добре працює в текстурованих областях і широко застосовується в задачах відстеження ключових точок, однак його точність знижується в однорідних ділянках зображення та при великих переміщеннях. Глобальний метод Хорна–Шунка враховує гладкість поля руху на всьому зображенні шляхом мінімізації функціонала енергії, що поєднує умову сталості яскравості та регуляризацию. Завдяки цьому забезпечується побудова більш плавного й узгодженого поля оптичного потоку. Метод є більш стійким до шумів і дозволяє відновлювати рух навіть у слабкотекстурних областях, проте потребує значно більших обчислювальних ресурсів та чутливий до вибору параметра регуляризації. Поліноміальний метод Farneback використовує апроксимацію локальної структури зображення квадратичними поліномами, що дозволяє отримувати щільне поле оптичного потоку з високою точністю. Метод демонструє хорошу стійкість до шумів, здатний оцінювати складні переміщення та часто застосовується в сучасних системах відеоаналізу й ідентифікації

рухомих об'єктів.

Таким чином, кожен із розглянутих методів має власні переваги та обмеження. Лукаса–Канаде доцільно використовувати для швидкого локального аналізу руху, Хорна–Шунка для отримання глобально узгодженого поля потоку, а Farneback для точного щільного оцінювання руху в реальних відеоданих. Вибір конкретного методу залежить від вимог до точності, швидкодії та умов обробки зображень у задачах ідентифікації.

## 2.2 Нейромережеві моделі для аналізу руху

Нейромережевий аналіз руху у відеопослідовності формалізовано через послідовність операцій, поданих у формулах (2.26)–(2.34). Вони описують перехід від вхідних ознак кадру до просторово-часового представлення, оцінки руху та узагальненого вектору параметрів об'єкта.

$$F_{t \rightarrow s}^{rig}(p_t) = K T_{t \rightarrow s} D_t(p_t) K^{-1} p_t - p_t \quad (2.36)$$

$$\mathcal{L}_{rw} = \alpha \frac{1 - \text{SSIM}(I_t, \tilde{I}_s^{rig})}{2} + (1 - \alpha) \|I_t - \tilde{I}_s^{rig}\|_1 \quad (2.37)$$

$$\mathcal{L}_{ds} = \sum |\nabla D(p_t)| \cdot (e^{-|\nabla I(p_t)|})^T \quad (2.38)$$

$$\mathcal{L}_{gc} = \sum_{p_t} [\delta(p_t)] \cdot \|\Delta f_{t \rightarrow s}^{full}(p_t)\|_1 \quad (2.39)$$

$$\|\Delta f_{t \rightarrow s}^{full}(p_t)\|_2 < \max\{\alpha, \beta\} \|\Delta f_{t \rightarrow s}^{full}(p_t)\|_2 \quad (2.40)$$

$$\mathcal{L} = \sum_l \sum_{(t,s)} \{\mathcal{L}_{rw} + \lambda_{ds} \mathcal{L}_{ds} + \mathcal{L}_{fw} + \lambda_{fs} \mathcal{L}_{fs} + \lambda_{gc} \mathcal{L}_{gc}\} \quad (2.41)$$

Генеративні моделі, такі як StableAnimator [30] та HunyuanVideo [28], вносять вагомий внесок у моделювання динамічних процесів. StableAnimator використовує дифузійні моделі та розв'язання рівняння Гамільтона-Якобі-Беллмана для створення високоточних анімацій із збереженням ідентичності. Метод SynCamMaster забезпечує синхронізовану генерацію відео з різних точок

огляду, що є базою для створення «цифрових двійників» складних систем. Ці технології дозволяють генерувати великі обсяги синтетичних даних для навчання та тестування інтелектуальних детекторів у рідкісних сценаріях.

Модель FlowNet

$$\Phi_{\text{encoder}}(I_1, I_2) = \text{Conv}_{1 \times 1}(\text{Conv}_{3 \times 3}(\cdots \text{Conv}_{7 \times 7}([I_1, I_2]) \cdots)) \quad (2.42)$$

$$\hat{F} = \Phi_{\text{decoder}}(\Phi_{\text{encoder}}(I_1, I_2)) \quad (2.43)$$

Оптимізація мережі відбувається шляхом мінімізації функції втрат EPE (End-Point Error)

$$\mathcal{L}_{\text{EPE}} = \frac{1}{N} \sum_{i=1}^N \|F_{\text{pred}}(i) - F_{\text{gt}}(i)\|_2 \quad (2.44)$$

Модель RAFT (Recurrent All-pairs Field Transforms) використовує кореляційний об'єм для встановлення відповідностей між кадрами

$$C_{ijkl} = \frac{1}{\sqrt{HW}} \sum_{h,w} f_{ijhw} \cdot g_{klhw} \quad (2.45)$$

Оновлення поля потоку виконується за допомогою GRU (Gated Recurrent Unit)

$$\Delta x_t, h_{t+1} = \text{GRU}([x_t, C(x_t)], h_t) \quad (2.46)$$

Такий рекурентний підхід дозволяє отримувати високу точність на великих зміщеннях.

Модель DETR для виявлення об'єктів

$$\text{MultiHead}(Q, K, V) = \text{Concat}(\text{head}_1, \dots, \text{head}_h) W^O \quad (2.47)$$

де кожна голова визначається як:

$$\text{head}_i = \text{Attention}(QW_i^Q, KW_i^K, VW_i^V) \quad (2.48)$$

Оптимальне парування між передбаченнями та істинними об'єктами знаходиться за допомогою угорського алгоритму

$$\hat{\sigma} = \operatorname{argmin}_{\sigma \in \mathbb{S}_N} \sum_{i=1}^N \mathcal{L}_{\text{match}}(y_i, \hat{y}_{\sigma(i)}) \quad (2.49)$$

Моделі для оцінки глибини: MiDaS та DPT

$$\mathcal{L} = \frac{1}{N} \sum_{i=1}^N \frac{1}{2\sigma_i^2} \mathcal{L}_i + \log \sigma_i \quad (2.50)$$

Модель DPT (Dense Prediction Transformer) поєднує трансформерну архітектуру з механізмом уваги

$$\text{Attention}(Q, K, V) = \operatorname{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (2.51)$$

### 2.3 Ансамблеві методи

Після підгонки мішка з попередньо визначеною моделлю машинного навчання, пристосованість мішка залежить від його розміру ( $\phi$ ) відповідно до формули (2.35) [63]

$$\text{fitness}(\text{bag}) = \alpha \times \frac{K + \phi_b}{K} \quad (2.52)$$

де  $\alpha$  – продуктивність класифікації моделі, пов’язаної з мішком ( $b$ ) в ансамблі,

а  $K$  – визначений користувачем гіперпараметр для заохочення більших мішків. Мета оператора кросовера – «урізноманітнити» вміст мішка так, щоб модель, підігнана на його даних, мала кращу продуктивність.

$C$  крок вибору батьків для кросовера, використовуючи схему рангової селекції, де нащадки з найвищою пристосованістю вибираються для кросовера,

$M$  – мішків, щодо мутації будуть вибрані випадковим чином із поточної популяції  $P'$ .

Нехай  $B$  – один із вибраних мішків, а  $B^c$  – множина всіх навчальних зразків, яких немає в  $B$ . Мішок  $B$  буде мutowано шляхом заміни вибраних випадкових зразків, заданих розміром мутації (MS) з  $B$ , такою ж кількістю випадкових зразків з  $B^c$ .

Правило голосування для агрегації передбачень окремих навчальних алгоритмів задається формулою (2.36)



$$\hat{y} = \begin{cases} 0, & \text{якщо } \sum_{l \in L} P_{neg} \cdot acc_l < \sum_{l \in L} P_{pos} \cdot acc_l \\ 1, & \text{в іншому випадку} \end{cases} \quad (2.53)$$

де  $l$  – окремий навчальний алгоритм у множині  $L$ ,

$acc_l$  – точність навчання  $l$ .

Ансамблевий метод бегінгу (bagging) передбачає створення кількох моделей на основі різних підвбірок вхідних даних, результати яких об'єднуються за допомогою середнього значення або голосування. Кожна модель навчається на випадковій підвбірці кадрів відео, що дає змогу зменшити дисперсію моделі та знизити ймовірність перенавчання. У системі, що розглядається, на етапі попередньої обробки застосовуються варіації кадрів, зокрема випадкова зміна яскравості, масштабування та обрізка. Далі на кожній модифікації кадру запускається GeoNet, яка оцінює глибину та оптичний потік. Результати об'єднуються шляхом обчислення середнього значення карт глибини [94, 60, 98].

Ансамблевий метод бустінгу передбачає послідовне навчання моделей, де кожна наступна фокусується на помилках попередньої, що дозволяє зменшити зміщення. У запропонованій системі це реалізовано шляхом поетапної оцінки глибини: початкова модель формує базову карту, а друга модель уточнює її у тих регіонах, де було виявлено найбільшу похибку. Це дозволяє зменшити зміщення та підвищити точність прогнозування, особливо на межах об'єктів або в умовах слабкої освітленості. У якості коректора використовується додаткова згортова нейронна мережа, яка аналізує локальні помилки та вносить точкові покращення [93, 60, 98].

Метод Bagging (Bootstrap Aggregating) передбачає усереднення результатів кількох моделей, навчених на різних підвбірках даних

$$\hat{y}_{\text{bag}}(x) = \frac{1}{M} \sum_{i=1}^M f_i(x) \quad (2.54)$$

У контексті оцінки глибини, Bagging дозволяє зменшити дисперсію прогнозу та підвищити стійкість до шумів. Експериментально показано, що

використання різних аугментацій вхідних даних дає кращі результати, ніж навчання різних моделей.

На відміну від Bagging, метод Boosting будує послідовну композицію моделей, де кожна наступна модель виправляє помилки попередніх

$$\hat{y}_{\text{boost}}(x) = \sum_{i=1}^M \alpha_i f_i(x) \quad (2.55)$$

Вага кожної моделі обчислюється на основі її помилки

$$\alpha_i = \frac{1}{2} \ln\left(\frac{1-\epsilon_i}{\epsilon_i}\right) \quad (2.56)$$

Для задач оцінки глибини було запропоновано адаптивне комбінування результатів різних методів

$$D_{\text{final}}(x) = \begin{cases} D_1(x), & \text{якщо } |\nabla D_1(x)| \leq \tau \\ D_2(x), & \text{якщо } |\nabla D_1(x)| > \tau \end{cases} \quad (2.57)$$

Це дозволяє використовувати переваги різних методів у різних областях зображення.

## 2.4 Об'єднання методів

Об'єднання методів дистанційної ідентифікації ґрунтується на спільному використанні просторових ознак, отриманих детектором об'єктів, часових характеристик, визначених за оптичним потоком, та оцінок глибини сцени. Такий підхід дозволяє перейти від окремого виявлення об'єкта в кадрі до повного опису його стану в просторі та часі. Узагальнена просторово-часова модель задається формулою (2.37)

$$I(x, y, t): \Omega \subset \mathbb{R}^2 \times [0, T] \rightarrow \mathbb{R} \quad (2.58)$$

Модель динамічного об'єкта формалізує перехід від вимірних ознак до оцінених параметрів об'єкта за формулою (2.38). Вона враховує положення об'єкта в кадрі, його межі, зміну координат між сусідніми кадрами та просторову віддаленість від камери.

$$X(t) = \{x_{i(t)}\}_{i=1}^{\{N\}} \quad (2.59)$$

Вектор стану об'єкта визначається формулою (2.39)

$$S(t) = \begin{bmatrix} x(t) \\ y(t) \\ v_{x(t)} \\ v_{y(t)} \\ \theta(t) \end{bmatrix} \quad (2.60)$$

Динамічна модель руху описується формулою (2.40)

$$\dot{s}(t) = f(s(t), w(t)) \quad (2.61)$$

Модель GeoNet реалізує сумісну оцінку глибини та оптичного потоку на основі геометричних обмежень:

$$I_2(p) = I_1(p + f(p)) \quad (2.62)$$

$$f(p) = T(p) \cdot \frac{Z(p) - Z_0}{Z(p)} \cdot \text{proj}(p) \quad (2.63)$$

Такий підхід дозволяє використовувати інформацію з обох джерел для підвищення точності. Обчислення кінематичних параметрів об'єктів.

Швидкість об'єкта:

$$v = \frac{1}{N_\Omega} \sum_{(x,y) \in \Omega} \|F(x, y)\| \quad (2.64)$$

$$v = \frac{1}{|\Omega|} \iint_{\Omega} \sqrt{u(x, y)^2 + v(x, y)^2} dx dy \quad (2.65)$$

Напрямок руху

$$\theta = \arctan\left(\frac{\sum_{(x,y) \in \Omega} \sin\theta(x, y)}{\sum_{(x,y) \in \Omega} \cos\theta(x, y)}\right) \quad (2.66)$$

Просторове положення:

$$x_c = \frac{x_{\min} + x_{\max}}{2}, y_c = \frac{y_{\min} + y_{\max}}{2} \quad (2.67)$$

$$z = \text{median}\{D(x, y) \mid (x, y) \in \Omega\} \quad (2.68)$$

Траскторія об'єкта:

$$T(t) = \{p(t_1), p(t_2), \dots, p(t_n)\}, p(t_i) = (x_i, y_i) \quad (2.69)$$

$$\tilde{p}(t_i) = \frac{1}{2k+1} \sum_{j=-k}^k p(t_{i+j}) \quad (2.70)$$

Для коректного порівняння результатів різних методів використовується Min-Max нормалізація

$$D_{\text{norm}}(x, y) = \frac{D(x, y) - \min(D)}{\max(D) - \min(D)} \times 255 \quad (2.71)$$

Візуалізація оптичного потоку виконується шляхом перетворення у кольорове зображення:

$$h = \frac{\arctan(v, u)}{2\pi} \times 180^\circ \quad (2.72)$$

$$v_{\text{bright}} = \min\left(255, \frac{\sqrt{u^2 + v^2}}{\max(\sqrt{u^2 + v^2})} \times 255\right) \quad (2.73)$$

## 2.5 Висновки до розділу

1. Розглянуто класичні методи оптичного потоку: локальний метод Лукаса-Канаде, який забезпечує високу швидкість обчислень при локальній сталості руху, але формує розріджене поле; глобальний метод Хорна-Шунка, який будує густе гладке поле за рахунок мінімізації функціонала енергії з регуляризацією; поліноміальний метод Фарнебека, який використовує пірамідальну структуру зображень і забезпечує високу точність та стійкість до шумів. Показано, що кожен метод має свої переваги та обмеження, а їх комбінування може підвищити надійність ідентифікації.

2. Проаналізовано нейромережеві моделі аналізу руху: FlowNet з енкодер-декодер архітектурою та навчанням end-to-end на основі втрати EPE; RAFT з рекурентними оновленнями та кореляційним об'ємом для високої точності на великих зміщеннях; GeoNet для сумісної оцінки глибини та оптичного потоку на

основі геометричних обмежень. Розглянуто трансформерні моделі виявлення DETR. Визначено, що неймережеві підходи забезпечують вищу точність порівняно з класичними методами, особливо в умовах складного руху та оклюзій.

3. Запропоновані ансамблеві методи машинного навчання для підвищення надійності ідентифікації. Метод Bagging передбачає усереднення результатів кількох моделей, навчених на різних підвбірках даних, що дозволяє зменшити дисперсію оцінок та підвищити стійкість до шумів. Метод Boosting будує послідовну композицію моделей, де кожна наступна виправляє помилки попередніх, що дозволяє зменшити зміщення. Запропоновано адаптивне комбінування результатів оцінки глибини, яке використовує переваги різних методів у різних областях зображення.

4. Формалізовано узагальнену просторово-часову модель, вектор стану об'єкта та динамічну модель руху. Розглянуто об'єднання методів на основі сумісної оцінки глибини та оптичного потоку (GeoNet, а також обчислення кінематичних параметрів: швидкості, напрямку руху, просторового положення, траєкторії. Показано, що інтеграція різнорідних методів в єдину систему дозволяє отримати більш повний та надійний опис стану динамічного об'єкта. Узагальнено, що найбільш перспективним є комплексне використання детекторів об'єктів, моделей оптичного потоку та методів оцінювання глибини, оскільки воно забезпечує одночасне врахування просторових і часових характеристик об'єктів. Обґрунтовано доцільність застосування ансамблевих методів Bagging та Boosting для зменшення дисперсії оцінок, уточнення меж об'єктів і підвищення стійкості результатів у складних умовах спостереження. Формалізація базових моделей і метрик створює основу для подальшої побудови інтегрованої інформаційної технології в наступних розділах.

## РОЗДІЛ 3 МОДЕЛІ ДИСТАНЦІЙНОЇ ІДЕНТИФІКАЦІЇ ПАРАМЕТРІВ ДИНАМІЧНИХ ОБ'ЄКТІВ

У третьому розділі сформовано моделі дистанційної ідентифікації параметрів динамічних об'єктів за відеоданими. Розглянуто концепцію побудови моделей, визначення параметрів за допомогою трансформерів і оптичного потоку, а також об'єднання результатів різних методів для підвищення точності та стійкості оцінювання.

### 3.1 Моделі дистанційної ідентифікації

Структура моделей дистанційної ідентифікації параметрів динамічних об'єктів ґрунтується на поєднанні трьох взаємопов'язаних рівнів: виявлення об'єктів у кадрі, оцінювання їхнього руху та визначення просторових характеристик сцени. На першому рівні використовуються детектори об'єктів, на другому – методи оптичного потоку, а на третьому – моделі оцінювання глибини. Інтеграція цих рівнів забезпечує отримання координат, швидкості, напрямку руху та траєкторії об'єкта в єдиному інформаційному просторі.

Основні поняття та визначення.

Дистанційна ідентифікація динамічного об'єкта – це процес визначення просторових, кінематичних та динамічних характеристик об'єкта за даними відеоспостереження без безпосереднього контакту з об'єктом.

Модель дистанційної ідентифікації – це математичний апарат, що встановлює зв'язок між спостережуваними відеоданими  $V(x, y, t)$  та параметрами об'єкта  $\theta = \{\theta_1, \theta_2, \dots, \theta_n\}$ .

Загальна модель дистанційної ідентифікації може бути представлена у вигляді

$$\theta = \mathcal{F}(V(x, y, t), \mathcal{M}, \mathcal{A}) \quad (3.1)$$

де:

$\theta$  – множина параметрів об'єкта, що ідентифікуються;

$V(x, y, t)$  – відеодані (послідовність кадрів);

$\mathcal{M}$  – множина математичних методів обробки;

$\mathcal{A}$  – множина архітектур нейронних мереж.

Математично ця ієрархія може бути представлена як:

$$\Theta = \{\Theta_{\text{geom}}, \Theta_{\text{kin}}, \Theta_{\text{dyn}}\} \quad (3.2)$$

$$\Theta_{\text{geom}} = \{(x_c, y_c, z), (w, h), \text{shape}\} \quad (3.3)$$

$$\Theta_{\text{kin}} = \{v, \theta, T(t), a\} \quad (3.4)$$

$$\Theta_{\text{dyn}} = \{m, p, E\} \quad (3.5)$$

Архітектура системи дистанційної ідентифікації складається з наступних модулів

$$\mathcal{S} = \{\mathcal{D}, \mathcal{F}, \mathcal{Z}, \mathcal{P}, \mathcal{V}\} \quad (3.6)$$

де:

$\mathcal{D}$  – модуль виявлення об'єктів (DETR/YOLO);

$\mathcal{F}$  – модуль оптичного потоку (Farneback/Lucas-Kanade/Horn-Schunck/FlowNet/RAFT);

$\mathcal{Z}$  – модуль оцінки глибини (MiDaS/DPT/GeoNet);

$\mathcal{P}$  – модуль обчислення параметрів;

$\mathcal{V}$  – модуль візуалізації.

### 3.2 Моделі ідентифікації параметрів за відеоданими

Схема дистанційної ідентифікації параметрів динамічного об'єкта з використанням DETR та оптичного потоку показана на рис. 3.1 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.

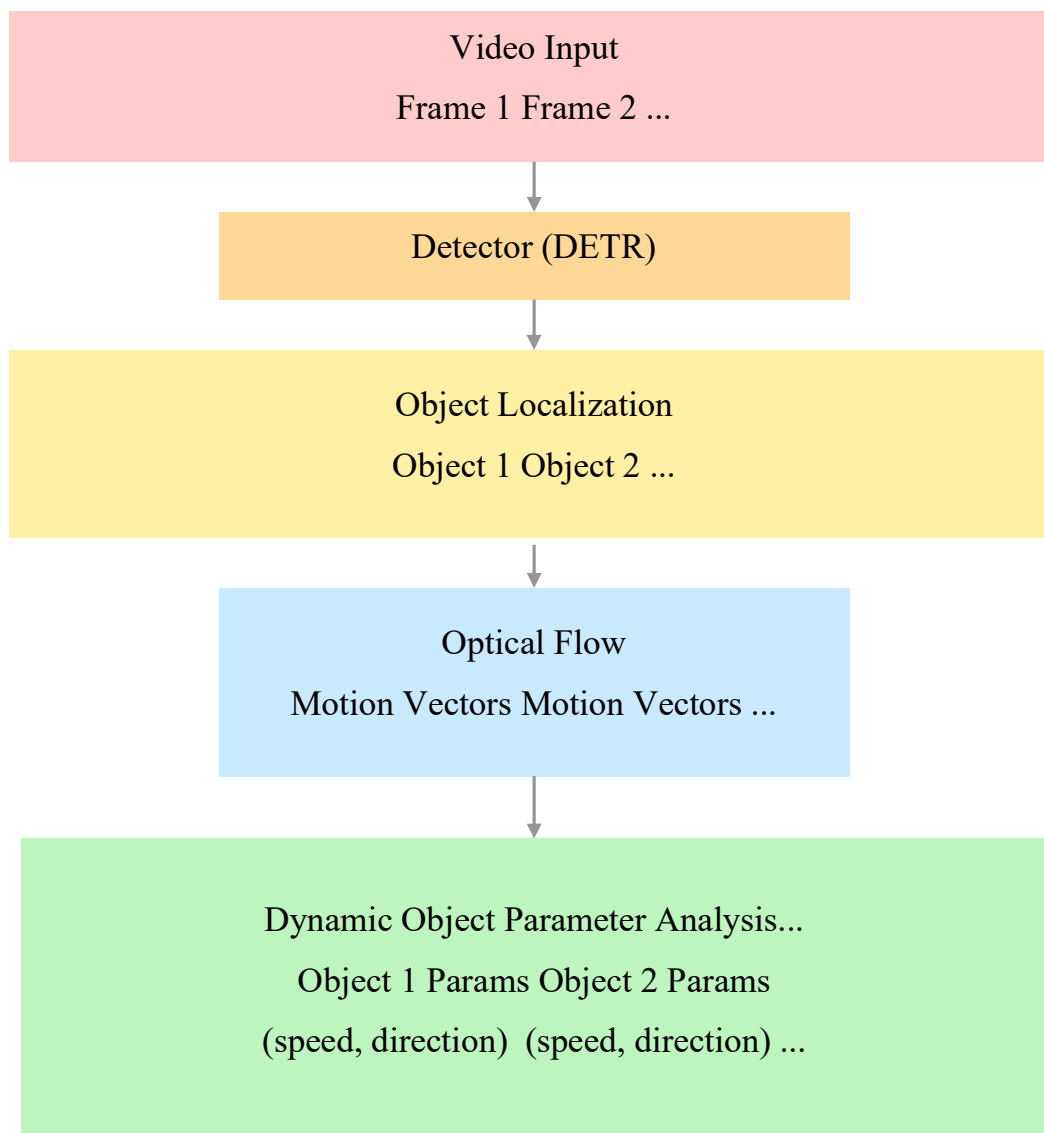


Рисунок 3.1 – Аналітична модель

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 щодо швидкості руху, напрямку руху, траєкторію та інші динамічні характеристики. На узагальненій схемі послідовність обробки: надходження відеоданих, виявлення об'єктів трансформерною моделлю DETR, локалізацію об'єктів, обчислення оптичного потоку та подальше визначення параметрів швидкості, напрямку руху і траєкторії [8, 46, 91, 92].

На рис. 3.2 показана схема технології трансформерів виявлення, що включає ЗНМ (а), кодер (б), декодер (в), НМПП (г).

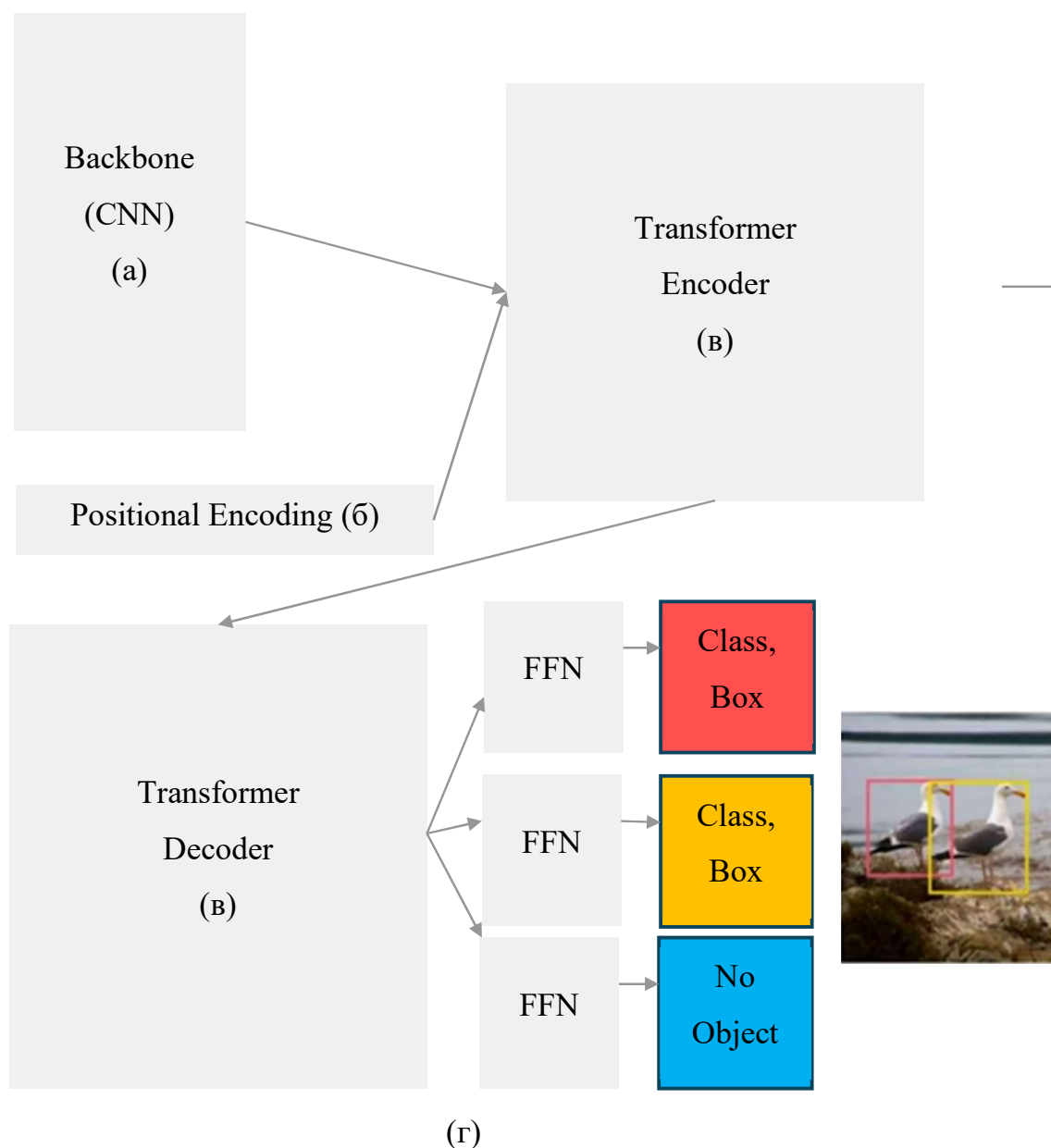


Рисунок 3.2 – Схема технології трансформерів виявлення: а – ЗНМ; б – кодер; в – декодер; г – НМПП

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 [93]

$$B_t = \text{DETR}(I_t) \quad (3.7)$$

де  $F_t$  – кадр відеопослідовності;

$I_t$  – кадр;

$t$  – індекс кадру;

$B_t$  – набір обмежувальних рамок для кожного кадру.

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

$$F_{t \rightarrow t+1} = \text{OpticalFlow}(I_t, I_{t+1}) \quad (3.8)$$

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

$F_{t \rightarrow t+1}$  – поле оптичного потоку між цими кадрами.

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.

$$v = \frac{1}{N} \sum_{i=1}^N \| F_{t \rightarrow t+1}(i) \| \quad (3.9)$$

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

$F_{t \rightarrow t+1}$  – вектор руху  $i$ -ї точки;

$v$  – швидкість об'єкта;

$\theta$  – 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

$$\theta = \frac{1}{N} \sum_{i=1}^N \text{atan2} \left( F_{t \rightarrow t+1}^y(i), F_{t \rightarrow t+1}^x(i) \right) \quad (3.10)$$

Класична нейронна мережа прямого поширення обробляє вихід декодера для отримання кінцевих результатів.

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 [14, 45, 92].

Модель кодер – декодер. Схема моделі з'єднання елементів кодер – декодер показана на рис. 3.3, кодера, який складається з шарів багатоголової самоуваги та НМПП.

На рис. 3.3б 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 багат шаровий перцептрон [46, 53].

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 [44; 48]. 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 вміст 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. 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 визначати 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 [48]

$$\text{MultiHeadAttn}(z_q, x) = \sum_{m=1}^M W_m [\sum_{k \in \Omega_k} A_{mqk} * W'_m x_k], \quad (3.11)$$

де  $q_i$  – індекс елемента запиту;  $q \in \Omega_q$

$z_q \in \mathbb{R}^C$  – ознака представлення запиту;

$m$  – індекс голови уваги;

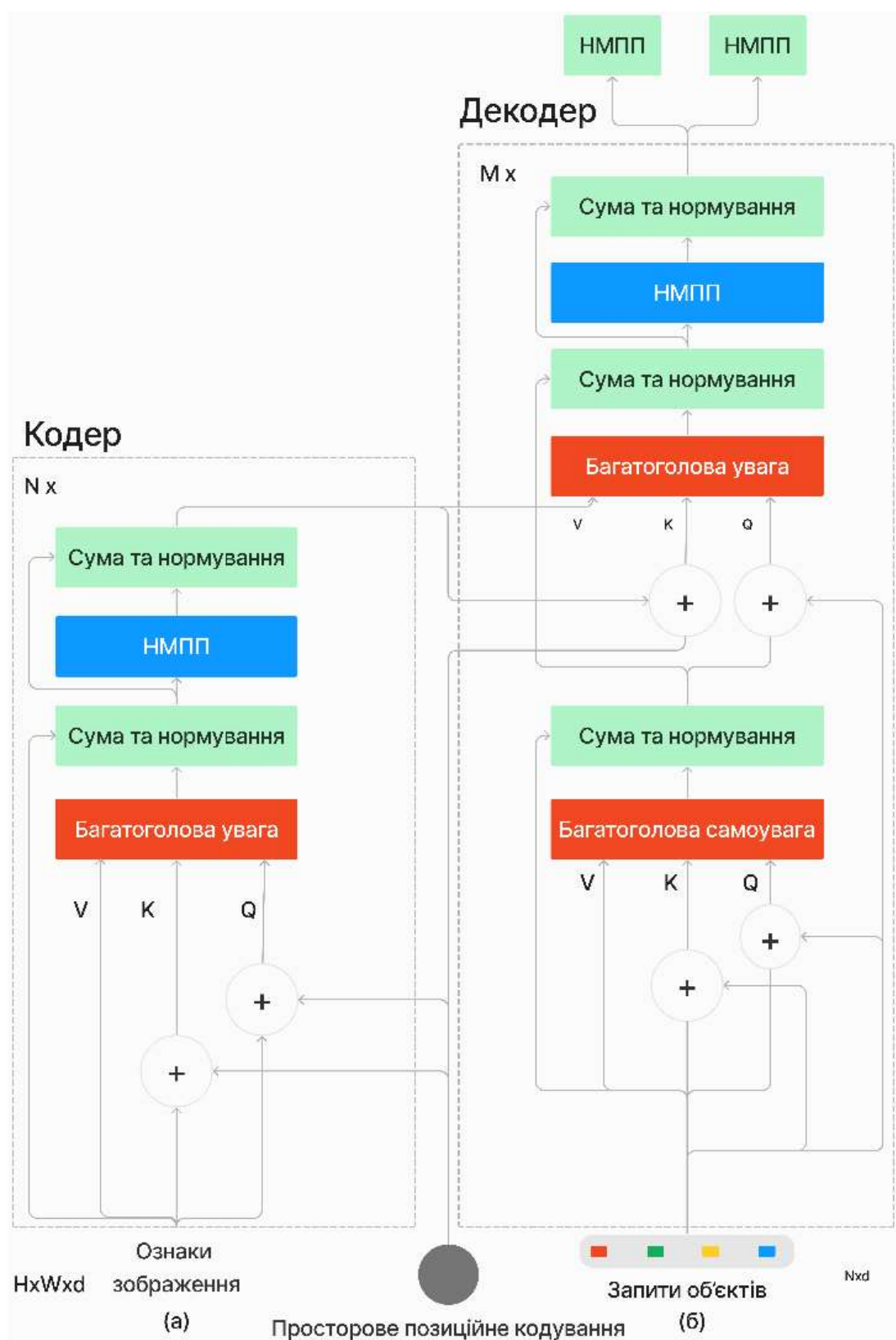


Рисунок 3.3 – Схема кодера трансформерної моделі DETR

$M$  – кількість голів уваги;

$W_m \in \mathbb{R}^{C \times C_v} W^O$  – вихідна проєкційна матриця.

$k \in \Omega_k$  – індекс елемента ключа;

$A_{mqk}, a_{ij}^h$  – ваги уваги;

$W_m' \in \mathbb{R}^{C_v \times C}$ ,  $W_h$  – вхідна проєкційна матриця для  $h$ -ї голови;

$x_k \in \mathbb{R}^C$ ,  $x_j^k$  – ознака представлення ключа;

$\Omega_q, Q$  та  $K$  – множини елементів запиту та ключа відповідно.  $\Omega_k$

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

$$C_v = C/M \quad (3.12)$$

де  $d_k$  – розмірність ознаки ключа  $C$ .

Важелі уваги обчислюються за формулою

$$A_{mqk} \propto e^{\frac{z_q^T U_m^T v_m x_k}{\sqrt{C_v}}} \quad (3.13)$$

де  $W_h^Q$  та  $W_h^K$  – навчальні ваги вхідних запитів і ключів для проєкції матриці на  $h$ -ту голову;

$i$  та  $j$  – індекси голови, запиту та ключа відповідно  $qk$ .

Важелі уваги нормалізуються за формулою

$$\sum_{k \in \Omega_k} A_{mqk} = 1 \quad (3.14)$$

Ознаки представлення  $x_i^q$  та  $x_j^k$  зазвичай є конкатенацією або сумою вмісту елементів та позиційних вкладок для розрізнення різних просторових позицій.

Вектори запитів, ключів і значень обчислюються за формулами:

$$q_i = X_i W^Q, \quad k_i = X_i W^K, \quad v_i = X_i W^V \quad (3.15)$$

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

$W^Q$  – 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 або матриця запитів, отримана з вхідних ознак;

$k_i$  – вектор або матриця ключів, отримана з вхідних ознак;

$v_i$  – вектор або матриця значень, отримана з вхідних ознак.

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

$$\text{Attention}(q_i, k_j, v_j) = \sum_{j=1}^n \text{softmax}\left(\frac{q_i \cdot k_j}{\sqrt{d_k}}\right) v_j \quad (3.16)$$

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 проєктний 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 [4; 49].

GeoNet. Глибока нейронна архітектура, яка поєднує задачу оцінки глибини з задачею оцінки оптичного потоку, тобто векторного поля переміщення пікселів між двома послідовними зображеннями. Її основна мета інтегрувати просторову та часову інформацію в єдину структуру для більш точної реконструкції сцени та відстеження об'єктів. У контексті задачі ідентифікації параметрів динамічних об'єктів, GeoNet виступає як фундаментальний інструмент для попереднього збору кінематичних характеристик об'єктів у відеопотоці. показано загальну структуру GeoNet, де пара вхідних зображень послідовних кадрів обробляється мережею для одночасного визначення карти глибини та оптичного потоку. Такий підхід дозволяє об'єднати стереозору та часову динаміку, що критично важливо для ідентифікації просторового положення та напрямку руху об'єктів у сцені. Використання GeoNet у поєднанні з трансформерною архітектурою дозволяє значно підвищити точність визначення параметрів руху завдяки можливості обробки глобального контексту у сцені [63]. Трансформер аналізує отримані карти глибини та оптичного потоку, забезпечуючи високорівневе представлення руху, яке важливе для класифікації, прогнозування траєкторії та подальшої аналітики об'єктів.

GeoNet може інтегруватись з ансамблевими методами машинного



навчання, зокрема в цій статті бегінгом та бустінгом, для підвищення стійкості до шумів, зміни освітлення та часткової втрати даних. У випадку бегінгу, GeoNet запропоновано виконувати на різних модифікованих підвбірках відеокадрів, що дозволить отримати усереднені оцінки параметрів руху. У бустінгу результати першої моделі аналізуються, і друга модель фокусується на зонах з найбільшою похибкою, що забезпечує локальну точність.

Таким чином, використання GeoNet у зв'язці з трансформерами та ансамблевими підходами створює ефективну систему для високоточного аналізу динаміки візуальних сцен та ідентифікації параметрів руху об'єктів. Це особливо актуально в задачах дистанційного моніторингу, автономного транспорту та систем відеоспостереження, де важлива не лише фіксація наявності об'єкта, а й повноцінне розуміння його поведінки в просторі.

Відео або зображення є проєкцією 3D-простору. 3D-сцена природно складається зі статичного фону та об'єктів, що рухаються. Рух статичних частин у відео викликаний виключно рухом камери та структурою глибини. Тоді як рух динамічних об'єктів є більш складним, обумовленим як однорідним рухом камери, так і специфічним рухом об'єкта.

$$f_{t \rightarrow s}^{rig}(p_t) = K T_{t \rightarrow s} D_t(p_t) K^{-1} p_t - p_t \quad (3.17)$$

де  $K$  – внутрішні параметри камери,

$p_t$  – однорідні координати пікселів у кадрі  $I(t)$ ,

$T_t \rightarrow s$  – відносний рух камери,

$D_t$  – карта глибини.

$$\mathcal{L}_{rw} = \alpha \frac{1 - SSIM(I_t, \tilde{I}_s^{rig})}{2} + (1 - \alpha) \|I_t - \tilde{I}_s^{rig}\|_1 \quad (3.18)$$

$$\mathcal{L}_{ds} = \sum_{p_t} |\nabla D(p_t)| \cdot (e^{-|\nabla I(p_t)|})^T \quad (3.19)$$

Перший етап забезпечує стереоскопічне сприйняття жорсткої структури сцени, але ігнорує динамічні об'єкти. Тому ми пропонуємо другий компонент – ResFlowNet для локалізації нежорсткого руху.

$$\mathcal{L}_{ge} = \sum_{p_t} [\delta(p_t)] \cdot \|\Delta f_{t \rightarrow s}^{full}(p_t)\|_1 \quad (3.20)$$

$$\|\Delta f_{t \rightarrow s}^{full}(p_t)\|_2 < \max \{\alpha, \beta \|f_{t \rightarrow s}^{full}(p_t)\|_2\} \quad (3.21)$$

$$\mathcal{L} = \sum_l \sum_{\langle t,s \rangle} \{\mathcal{L}_{rw} + \lambda_{ds} \mathcal{L}_{ds} + \mathcal{L}_{fw} + \lambda_{fs} \mathcal{L}_{fs} + \lambda_{ge} \mathcal{L}_{ge}\} \quad (3.22)$$

Модель виявлення об'єктів на основі DETR

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (3.23)$$

Багатоголова увага дозволяє моделювати різні типи залежностей:

$$\text{MultiHead}(Q, K, V) = \text{Concat}(\text{head}_1, \dots, \text{head}_h)W^O \quad (3.24)$$

$$\text{Head}_i = \text{Attention}(QW_i^Q, KW_i^K, VW_i^V) \quad (3.25)$$

Оптимальне парування між передбаченнями та істинними об'єктами знаходиться за допомогою угорського алгоритму:

$$\hat{\sigma} = \text{argmin}_{\sigma \in \mathbb{S}_N} \sum_{i=1}^N \mathcal{L}_{\text{match}}(y_i, \hat{y}_{\sigma(i)}) \quad (3.26)$$

Функція втрат DETR має вигляд

$$\mathcal{L}_{\text{Hungarian}}(y, \hat{y}) = \sum_{i=1}^N \left[ -\log \hat{p}_{\hat{\sigma}(i)}(c_i) + \mathbf{1}_{\{c_i \neq \emptyset\}} \mathcal{L}_{\text{box}}(b_i, \hat{b}_{\hat{\sigma}(i)}) \right] \quad (3.27)$$

Модель оцінки глибини MiDaS призначена для монокулярної оцінки глибини. Вона розв'язує проблему відносної глибини

$$d_{\text{pred}}(p) = \alpha \cdot d_{\text{true}}(p) + \beta \quad (3.28)$$

Багатомасштабна функція втрат забезпечує навчання на різних масштабах:

$$\mathcal{L} = \frac{1}{N} \sum_{i=1}^N \frac{1}{2\sigma_i^2} \mathcal{L}_i + \log \sigma_i \quad (3.29)$$

Інваріантна до масштабу втрата дозволяє навчатися на даних з різними масштабами глибини:

$$\mathcal{L}_{SI} = \frac{1}{N} \sum_{i,j} \left[ (\log y_i - \log \hat{y}_i) - (\log y_j - \log \hat{y}_j) \right]^2 \quad (3.30)$$

Модель DPT для щільної оцінки глибини Dense Prediction Transformer поєднує трансформерну архітектуру з механізмом уваги:

$$z_l = \text{MSA}(\text{LN}(z_{l-1})) + z_{l-1}, l = 1, \dots, L \quad (3.31)$$

Вхідні токени формуються як

$$z_0 = [x_{\text{class}}; x_p^1 E; x_p^2 E; \dots; x_p^N E] + E_{\text{pos}} \quad (3.32)$$

### 3.3 Визначення параметрів методами оптичного потоку

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. 3.3 показано результат використання оптичного потоку для візуалізації напрямку та інтенсивності руху в кадрі.

Використання пірамідної структури зображень забезпечує інваріантність методу до масштабу руху, що дозволяє ефективно обробляти як повільні, так і швидкі переміщення об'єктів у кадрі, оскільки на грубому рівні піраміди оцінюються великі зміщення, а на детальному рівні уточнюються малі рухи [22, 94].

Отримане поле оптичного потоку являє собою двовимірне векторне поле, де для кожного пікселя кадру визначаються горизонтальна та вертикальна компоненти переміщення, які далі використовуються для обчислення швидкості руху об'єкта як середньої величини вектора потоку в межах області об'єкта,

виявленої детектором DETR, та напрямку руху як середнього арифметичного або векторного кута [42, 96]. Кольорове кодування оптичного потоку, застосоване на рисунку 3.3, виконується шляхом перетворення компонент вектора руху в HSV-простір, де відтінок кодує напрямок руху, насиченість є максимальною, а яскравість відповідає величині вектора переміщення, що забезпечує наочне сприйняття динамічних характеристик сцени [22, 94].

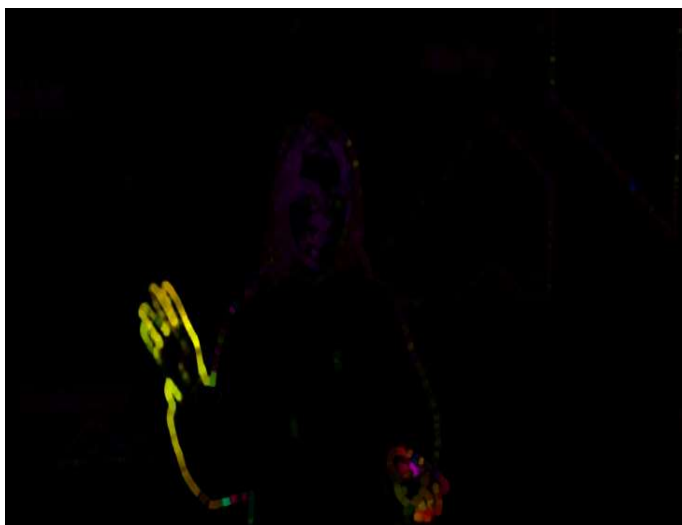


Рисунок 3.3 – Результат використання оптичного потоку для візуалізації напрямку та інтенсивності руху

Оптичний потік, показаний на рис. 3.4 визначає напрям і швидкість переміщення пікселів між послідовними кадрами. Його застосування дозволяє отримати траєкторію руху об'єкта та визначити його кінематичні параметри. Оптичний потік є векторним полем, яке відображає переміщення пікселів між двома послідовними кадрами відео. Цей метод дозволяє визначити напрямок і швидкість руху об'єктів у сцені. Завдяки цьому можна побудувати траєкторію об'єкта, визначити його миттєву швидкість і прискорення. У запропонованій системі оптичний потік виступає важливим джерелом інформації про кінематичні параметри динамічних об'єктів, особливо в ситуаціях, коли об'єкти частково або повністю перекриваються (оклюзія) [54, 94].



Рисунок 3.4 – Візуалізація оптичного потоку в обраній моделі

Метод оптичного потоку (Optical Flow) є фундаментальним інструментом у галузі комп'ютерного зору, який забезпечує кількісне оцінювання руху пікселів між послідовними кадрами відеопотоку. У межах розроблюваної інформаційної технології дистанційної ідентифікації цей метод дозволяє отримати детальну інформацію про динаміку сцени, що є критично важливим для точного відстеження та розрахунку параметрів рухомих об'єктів у реальному часі. Оптичний потік визначається як векторне поле, де кожен вектор описує видиме переміщення точок зображення, що дає змогу ідентифікувати швидкість, напрямок та траєкторію руху цілей без безпосереднього фізичного контакту.

Традиційний аналіз руху базується на класичних диференціальних алгоритмах, серед яких ключове місце посідають методи Лукаса-Канаде та Хорна-Шунка [41, 62]. Метод Лукаса-Канаде використовує локальне рівняння різниці інтенсивності пікселів та метод найменших квадратів, забезпечуючи високу швидкість обчислень при локальній сталості руху. На противагу йому, метод Хорна-Шунка формулює задачу як мінімізацію функціоналу енергії з умовою глобальної гладкості, що дозволяє будувати густі карти руху, проте потребує значних ітераційних ресурсів та є чутливим до шумів зображення.

Метод Горна-Шунка (Horn-Schunck) є глобальним методом обчислення оптичного потоку, представленим у 1981 році. Він базується на двох основних припущеннях: сталість яскравості означає, що колір або яскравість точки об'єкта не змінюється між кадрами; плавність потоку означає, що оптичний потік змінюється поступово від пікселя до пікселя. На відміну від локального методу Лукаса-Канаде, який передбачає сталість потоку в локальному вікні, метод Горна-Шунка є глобальним і дає щільне поле оптичного потоку.

Основні припущення методу: сталість яскравості дозволяє порівнювати пікселі між кадрами за рівнянням  $I(x, y, t) = I(x + u\delta t, y + v\delta t, t + \delta t)$ ; малий рух означає, що пікселі переміщуються лише на невелику відстань між кадрами, що дозволяє лінеаризувати обмеження сталості яскравості.

$$I(x + u\delta t, y + v\delta t, t + \delta t) = I(x, y, t) \quad (3.33)$$

Розклад в ряд Тейлора для малого  $\delta t$

$$I(x, y, t) + \frac{\partial I}{\partial x} \delta x + \frac{\partial I}{\partial y} \delta y + \frac{\partial I}{\partial t} \delta t = I(x, y, t) \quad (3.34)$$

Спрощення дає рівняння оптичного потоку

$$I_x u + I_y v + I_t = 0 \quad (3.35)$$

де  $I_x, I_y$  – просторові градієнти зображення,  $I(t)$  – часова похідна.

Метод поєднує два обмеження: сталість яскравості та плавність потоку.

$$E_d(i, j) = [I_x u_{ij} + I_y v_{ij} + I_t]^2 \quad (3.36)$$

Складова плавності (smoothness) – штрафує різкі зміни оптичного потоку між сусідніми пікселями:

Повний функціонал, який мінімізується

$$\text{Min}_{u,v} \sum_{i,j} \{E_s(i, j) + \lambda E_d(i, j)\} \quad (3.37)$$

де  $\lambda$  – параметр, що регулює вагу між гладкістю та відповідністю яскравості.

$$\bar{u}_{ij} = \frac{1}{4} (u_{i+1,j} + u_{i-1,j} + u_{i,j+1} + u_{i,j-1}) \quad (3.38)$$

Аналогічно для  $\bar{v}_{ij}$ .

$$\frac{\partial E}{\partial u_{kl}} = 2(u_{kl} - \bar{u}_{kl}) + 2\lambda(I_x u_{kl} + I_y v_{kl} + I_t)I_x \quad (3.39)$$

$$\frac{\partial E}{\partial v_{kl}} = 2(v_{kl} - \bar{v}_{kl}) + 2\lambda(I_x u_{kl} + I_y v_{kl} + I_t)I_y \quad (3.40)$$

Прирівнювання похідних до нуля дає систему лінійних рівнянь:

$$(1 + \lambda I_x^2)u_{kl} + \lambda I_x I_y v_{kl} = \bar{u}_{kl} - \lambda I_x I_t \quad (3.41)$$

$$\lambda I_x I_y u_{kl} + (1 + \lambda I_y^2)v_{kl} = \bar{v}_{kl} - \lambda I_y I_t$$

Це система вигляду  $Ax = b$ , яка розв'язується для кожного пікселя.

$$\{1 + \lambda(I_x^2 + I_y^2)\}u_{kl} = (1 + \lambda I_x^2)\bar{u}_{kl} - \lambda I_x I_y \bar{v}_{kl} - \lambda I_x I_t \quad (3.42)$$

$$\{1 + \lambda(I_x^2 + I_y^2)\}v_{kl} = (1 + \lambda I_y^2)\bar{v}_{kl} - \lambda I_x I_y \bar{u}_{kl} - \lambda I_y I_t \quad (3.43)$$

Після перетворень отримуємо ітераційну схему:

$$\hat{u}_{kl} = \bar{u}_{kl} - \frac{I_x \bar{u}_{kl} + I_y \bar{v}_{kl} + I_t}{\lambda^{-1} + I_x^2 + I_y^2} I_x \quad (3.44)$$

$$\hat{v}_{kl} = \bar{v}_{kl} - \frac{I_x \bar{u}_{kl} + I_y \bar{v}_{kl} + I_t}{\lambda^{-1} + I_x^2 + I_y^2} I_y \quad (3.45)$$

$$E_s(i, j) = \frac{[(u_{ij} - u_{i+1, j})^2 + (u_{ij} - u_{i, j+1})^2 + (v_{ij} - v_{i+1, j})^2 + (v_{ij} - v_{i, j+1})^2]}{4} \quad (3.46)$$

$L_1$  норма  $L_1 = \Sigma |F(x + h) - G(x)|$

$L_2$  норма  $L_2 = \sqrt{(\Sigma [F(x + h) - G(x)]^2)}$

Від'ємна нормована кореляція  $= -\Sigma F(x + h)G(x) / \sqrt{(\Sigma F(x + h)^2 \Sigma G(x)^2)}$

Для малих  $h$  лінійна апроксимація дає

$$F'(x) \approx \frac{F(x+h) - F(x)}{h} = \frac{G(x) - F(x)}{h} \quad (3.47)$$

отже

$$h \approx \frac{G(x) - F(x)}{F'(x)} \quad (3.48)$$

Усереднення декількох оцінок

$$h = \Sigma_x \frac{\frac{G(x) - F(x)}{F'(x)}}{\Sigma_x 1} \quad (3.49)$$

Зважене усереднення з ваговою функцією  $w(x) = \frac{1}{|G'(x) - F'(x)|}$

$$h = \Sigma_x \frac{\frac{w(x)[G(x) - F(x)]}{F'(x)}}{\Sigma_x w(x)} \quad (3.50)$$

Ітерація Ньютона-Рафсона



$$h_0 = 0, h_{k+1} = h_k + \sum_x \frac{\frac{w(x)[G(x)-F(x+h_k)]}{F'(x+h_k)}}{\sum_x w(x)} \quad (3.51)$$

Лінійна апроксимація

$$F(x+h) \approx F(x) + hF'(x) \quad (3.52)$$

Функція похибки

$$E = \sum_x [F(x) + hF'(x) - G(x)]^2 \quad (3.53)$$

Прирівнювання похідної до нуля дає

$$h = \frac{\sum_x F'(x)[G(x)-F(x)]}{\sum_x F'(x)^2} \quad (3.54)$$

Для багатовимірного випадку мінімізуємо

$$E = \sum_{x \in R} [F(x+h) - G(x)]^2 \quad (3.55)$$

Лінійна апроксимація в n-вимірах

$$F(x+h) \approx F(x) + h \frac{\partial F}{\partial x} \quad (3.56)$$

Розв'язок для h

$$h = \left[ \sum_x \left( \frac{\partial F}{\partial x} \right)^T (G(x) - F(x)) \right] \left[ \sum_x \left( \frac{\partial F}{\partial x} \right)^T \left( \frac{\partial F}{\partial x} \right) \right]^{-1} \quad (3.57)$$

Узагальнений метод реєстрації може бути застосований для отримання інформації про глибину зі стереопар зображень. Для об'єкта на відстані  $z$  від камери 1, положення на площині плівки камери 2 може бути обчислено. Використовуючи лінійну апроксимацію для корекції глибини  $\Delta z$

$$\Delta z = \sum_x \frac{\frac{\partial F}{\partial z}[G-F]}{\sum_x \left( \frac{\partial F}{\partial z} \right)^2} \quad (3.58)$$

де  $\frac{\partial F}{\partial z}$  обчислюється через правило ланцюга з просторового градієнта інтенсивності та геометрії стереоустановки. Аналогічно, для невідомих параметрів камери  $c$ , оновлення  $\Delta c$  має вигляд

$$\Delta c = \left[ \sum_x \left( \frac{\partial q}{\partial c} \frac{\partial F}{\partial q} \right)^T (G - F) \right] \left[ \sum_x \left( \frac{\partial q}{\partial c} \frac{\partial F}{\partial q} \right)^T \left( \frac{\partial q}{\partial c} \frac{\partial F}{\partial q} \right) \right]^{-1} \quad (3.59)$$

$$h = \sum_x \frac{\frac{w(x)[G(x) - F(x)]}{F'(x)}}{\sum_x w(x)}$$

Для практичної реалізації в інтелектуальних системах часто використовується алгоритм Фарнбека (Farneback), який базується на апроксимації інтенсивності пікселів поліномами. Цей метод використовує пірамідальну структуру зображень, що дозволяє ефективно обробляти рухи різних масштабів – від повільних зміщень до швидких переміщень об'єктів у кадрі. У розробленому програмному забезпеченні алгоритм Фарнбека налаштований із використанням пірамідного масштабу 0.5, трьох рівнів піраміди та Гауссового фільтра з відхиленням 1.2, що забезпечує високу прецизійність ідентифікації кінематичних характеристик.

Сучасна еволюція методів аналізу руху призвела до появи нейромережових архітектур, таких як FlowNet та FlowNet 2.0, які навчаються передбачати оптичний потік за принципом «кінець-у-кінець» (end-to-end). Модель FlowNet пропонує два варіанти: FlowNetSimple, що самостійно вивчає рух із пари зображень, та FlowNetCorr, яка використовує кореляційний шар для явного порівняння областей кадрів. Використання великих синтетичних наборів даних, як-от Flying Chairs, дозволяє цим моделям успішно узагальнювати знання для ідентифікації параметрів руху в реальних складних сценаріях.

FlowNet 2.0 суттєво вдосконалює цей процес завдяки архітектурі зі стекуванням мереж та впровадженню спеціалізованих шарів для врахування малих зміщень. Ця модель здатна обробляти дані зі швидкістю до 140 кадрів на секунду, що робить її ідеальним компонентом для автономних транспортних

систем та робототехніки. Висока деталізація та низька похибка оцінки потоку у FlowNet 2.0 дозволяють точно ідентифікувати навіть мінімальні зміни в стані динамічного об'єкта, що є критичним для систем запобігання аваріям.

Інтеграція методів оптичного потоку з трансформерами виявлення (DETR) створює синергетичний ефект у задачах дистанційної ідентифікації. У такій комбінації DETR виконує прецизійну локалізацію об'єктів та визначення їхніх меж, а блок Optical Flow обчислює вектори руху між послідовними кадрами для цих локалізованих областей. Ця модель дозволяє системі одночасно «бачити» об'єкт і «розуміти» фізику його переміщення, що значно перевершує традиційні роздільні підходи за точністю та стійкістю до динамічного фону.

На рис. 3.5 подано схему інтеграції оптичного потоку з трансформерною обробкою векторів руху для подальшого визначення положення, швидкості, прискорення та напрямку руху динамічних об'єктів.

Математична модель оцінювання динамічних параметрів базується на аналізі поля векторів руху  $F$ . Швидкість об'єкта ( $v$ ) розраховується як величина середнього вектора руху для всіх точок ( $N$ ), що належать ідентифікованій цілі. Визначення миттєвої швидкості дозволяє системі формувати динамічний профіль об'єкта в реальному часі. Такий підхід забезпечує стабільність оцінок навіть при зміні форми об'єкта або його частковому перекритті. Для визначення траєкторії руху використовується розрахунок кута напрямку ( $\theta$ ) за допомогою функції  $\text{atan2}$ , яка аналізує вертикальні та горизонтальні компоненти векторів зміщення пікселів. Середнє значення напрямків усіх векторів руху об'єкта дозволяє точно спрогнозувати його подальший шлях. Це є фундаментом для аналізу траєкторій в інтелектуальних системах спостереження та автономного керування. Ключовим аспектом інформаційної технології є переведення піксельних зміщень у реальні фізичні величини (наприклад, км/год). Це досягається через калібрування камери, що дозволяє системі ідентифікувати швидкість транспортного засобу з похибкою не більше  $\pm 5$  км/год.

Такий рівень точності задовольняє вимоги до систем моніторингу дорожнього руху та промислової безпеки.

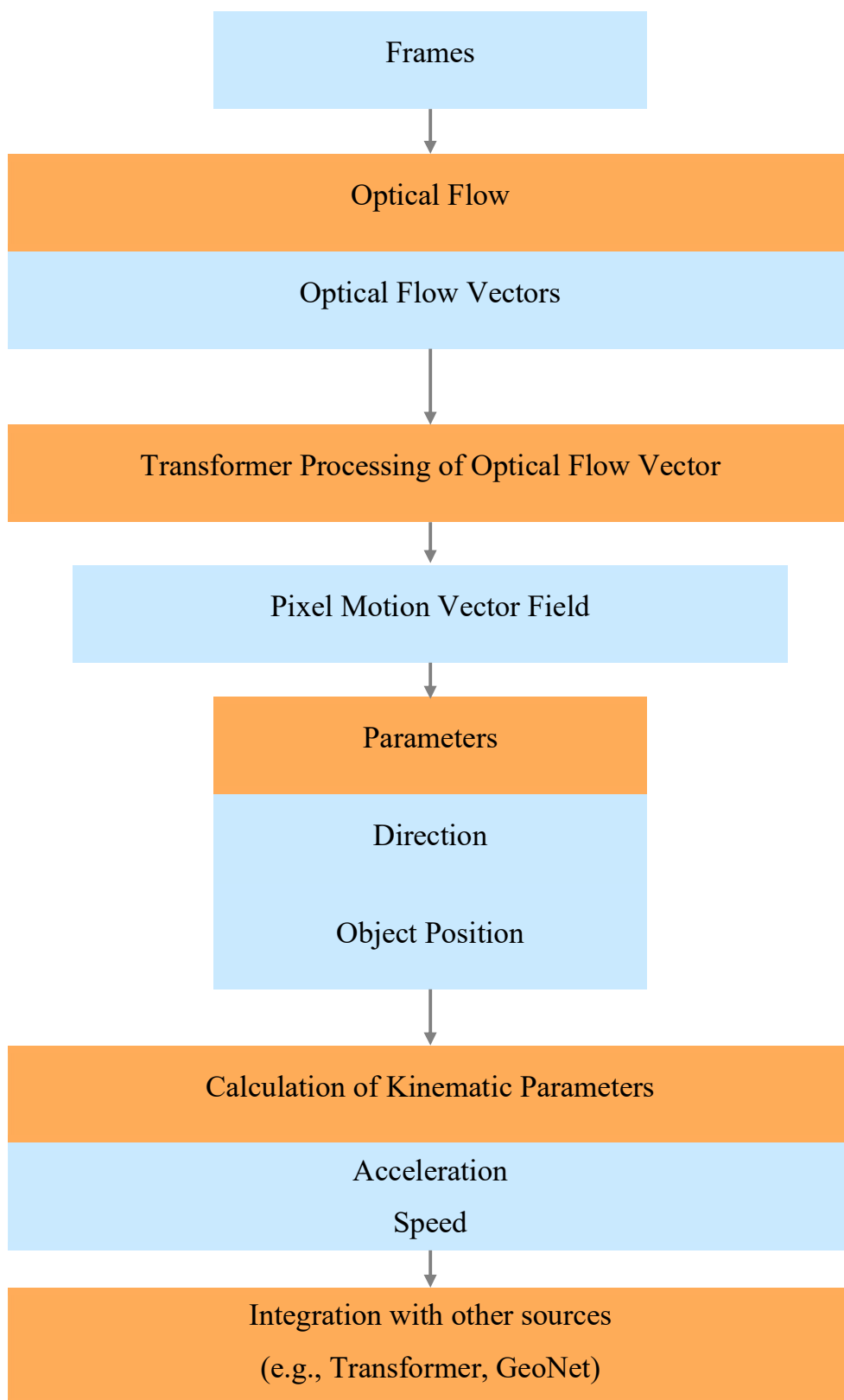


Рисунок 3.5 – Схема інтеграції оптичного потоку з трансформерною обробкою параметрів руху

Для поглибленого просторового аналізу методи оптичного потоку інтегруються з архітектурою GeoNet. GeoNet одночасно виконує монокулярну оцінку глибини сцени та оптичного потоку, що дозволяє здійснювати 3D-реконструкцію динамічного середовища. Завдяки спільному навчанню компонентів DepthNet та PoseNet, система здатна відокремлювати рух самої камери від автономного руху об'єктів, що суттєво підвищує надійність ідентифікації в рухомих системах (дрони, роботи).

Ефективність ідентифікації параметрів руху часто знижується через оклюзії та зміни освітлення. Використання густого оптичного потоку (Dense Optical Flow) допомагає нівелювати ці фактори, надаючи інформацію про рух у кожній точці кадру, а не лише для окремих ознак. Це дозволяє підтримувати безперервність відстеження об'єкта навіть у складних міських умовах або при низькій якості відеозапису.

Для підвищення надійності оцінок використовуються ансамблеві методи, такі як бегінг та бустінг. Бегінг дозволяє зменшити дисперсію результатів через усереднення карт руху, отриманих на різних підвбірках даних, що ефективно прибирає шумові артефакти. Бустінг застосовується для послідовної корекції похибок у складних зонах, наприклад, на краях рухомих тіл, де градієнти інтенсивності можуть бути розмитими.

Підсумовуючи, поєднання трансформерних моделей із алгоритмами оптичного потоку та GeoNet забезпечує високу точність і адаптивність інформаційної технології дистанційної ідентифікації. Отримані результати підтверджують можливість надійного визначення швидкості та траєкторії динамічних об'єктів у реальному часі. Подальший розвиток методології пов'язаний із вдосконаленням систем підтримки прийняття рішень на основі ідентифікованих характеристик руху для підвищення ефективності сучасних «розумних систем».

Модель швидкості об'єкта.

Швидкість об'єкта визначається як середня величина оптичного потоку в області об'єкта  $\Omega$

$$v = \frac{1}{N_{\Omega}} \sum_{(x,y) \in \Omega} \|F(x,y)\| \quad (3.60)$$

де  $F(x,y) = [u(x,y), v(x,y)]^T$  – вектор оптичного потоку.

В інтегральній формі

$$v = \frac{1}{|\Omega|} \iint_{\Omega} \sqrt{u(x,y)^2 + v(x,y)^2} dx dy \quad (3.61)$$

Модель напрямку руху.

Напрямок руху об'єкта обчислюється через векторне усереднення кутів

$$\theta(x,y) = \arctan\left(\frac{v(x,y)}{u(x,y)}\right) \quad (3.62)$$

$$\theta = \arctan\left(\frac{\sum_{(x,y) \in \Omega} \sin\theta(x,y)}{\sum_{(x,y) \in \Omega} \cos\theta(x,y)}\right) \quad (3.63)$$

Спрощена форма (середнє арифметичне кутів)

$$\theta = \frac{1}{N_{\Omega}} \sum_{(x,y) \in \Omega} \arctan\left(\frac{v(x,y)}{u(x,y)}\right) \quad (3.64)$$

Модель просторового положення. Координати центру об'єкта в площині зображення

$$x_c = \frac{x_{\min} + x_{\max}}{2} \quad (3.65)$$

$$y_c = \frac{y_{\min} + y_{\max}}{2} \quad (3.66)$$

Глибина  $Z$ -координата визначається як медіанне значення карти глибини в області об'єкта

$$z = \text{median}\{D(x,y) \mid (x,y) \in \Omega\} \quad (3.67)$$

Альтернативно, як середнє арифметичне

$$z = \frac{1}{|\Omega|} \iint_{\Omega} D(x,y) dx dy \quad (3.68)$$

Модель траєкторії об'єкта. Траєкторія об'єкта представляється у вигляді

дискретної множини положень у часі

$$T(t) = \{p(t_1), p(t_2), \dots, p(t_n)\} \quad (3.69)$$

де  $p(t_i) = (x_i, y_i)$  – положення об'єкта в момент  $t_i$ .

Для зменшення шумів використовується згладжена траєкторія (ковзне середнє)

$$\tilde{p}(t_i) = \frac{1}{2k+1} \sum_{j=-k}^k p(t_{i+j}) \quad (3.70)$$

Модель прискорення об'єкта. Прискорення об'єкта визначається як похідна швидкості за часом

$$a = \frac{v(t+1) - v(t)}{\Delta t} \quad (3.71)$$

де  $\Delta t = 1$  кадр (дискретний час).

#### 3.4 Об'єднання моделей і методів ідентифікації

На рис. 3.6 наведено узагальнену схему об'єднання відеопотоку, оптичного потоку, трансформерної мережі та ансамблевого модуля для прогнозування параметрів динамічних об'єктів [76, 77].

Трансформерна мережа обробляє вектори оптичного потоку за допомогою механізмів багатоголової уваги, що дозволяє виявляти складні патерни руху, які не можуть бути виявлені локальними або класичними методами, зокрема при часткових оклюзіях або зміні форми об'єкта [4, 14, 94]. Інформація про оптичний потік та параметри трансформера об'єднуються в ансамблевому модулі, який може використовувати методи бегінгу для зменшення дисперсії оцінок шляхом усереднення результатів на різних підвибірках даних або методи бустінгу для послідовного уточнення прогнозів у складних областях сцени [15, 93, 62].

На виході схеми формуються прогнозовані параметри динамічних об'єктів, які включають просторове положення об'єкта в кадрі, його миттєву швидкість, напрямок руху, прискорення та прогнозовану траєкторію [42, 94, 96].

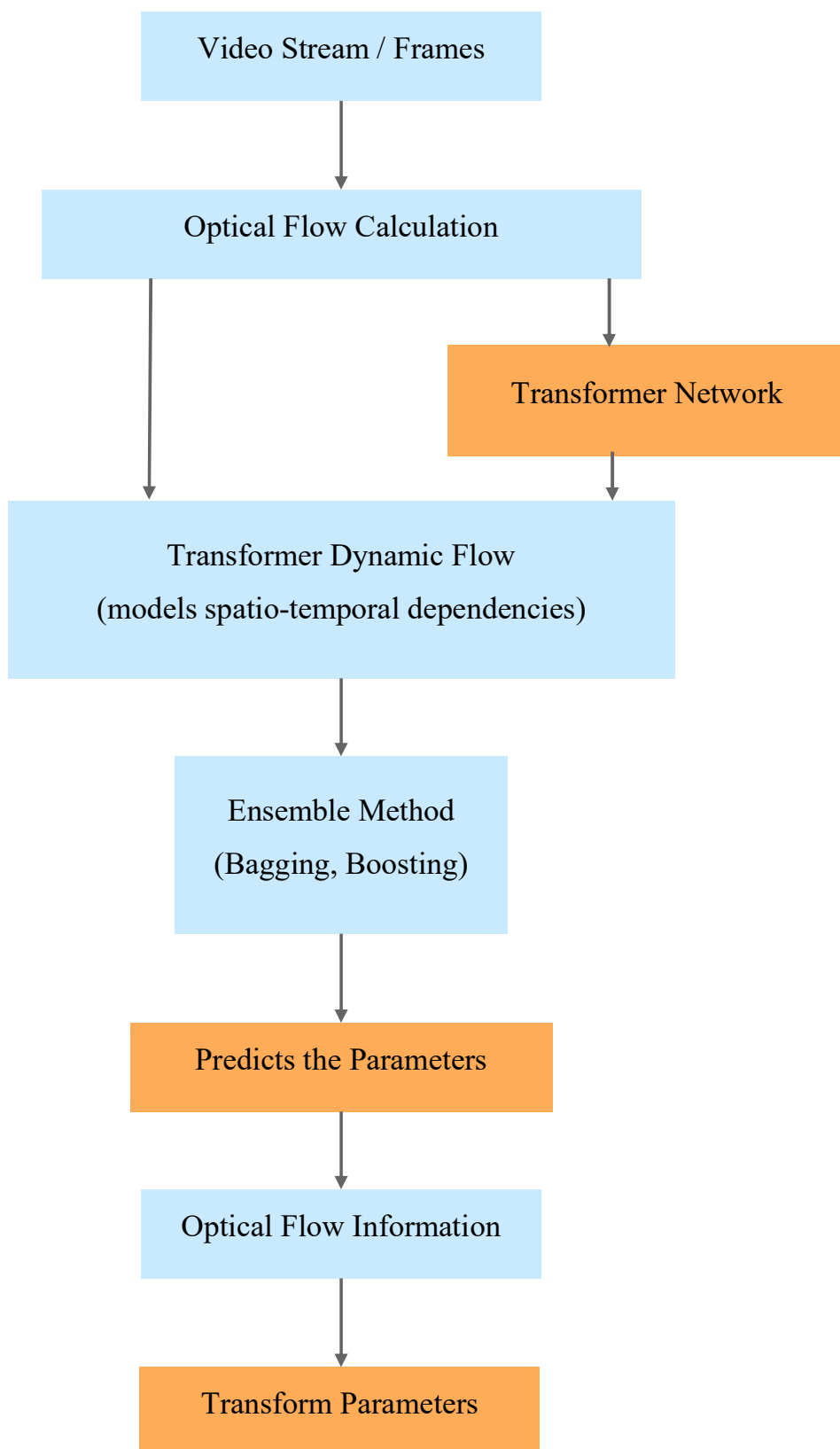


Рисунок 3.6 – Узагальнена схема об'єднання моделей і методів ідентифікації



Сумісна модель GeoNet реалізує сумісну оцінку глибини та оптичного потоку на основі геометричних обмежень

$$I_2(p) = I_1(p + f(p)) \quad (3.72)$$

$$f(p) = T(p) \cdot \frac{Z(p) - Z_0}{Z(p)} \cdot \text{proj}(p) \quad (3.73)$$

Ця модель дозволяє використовувати синергію між двома типами даних.

Ансамблева модель Bagging передбачає усереднення результатів кількох моделей

$$\hat{y}_{\text{bag}}(x) = \frac{1}{M} \sum_{i=1}^M f_i(x) \quad (3.74)$$

У контексті оцінки глибини

$$D_{\text{bag}}(x, y) = \frac{1}{M} \sum_{i=1}^M D_i(x, y) \quad (3.75)$$

Адаптивна модель Boosting використовує зважену комбінацію з виправленням помилок

$$\hat{y}_{\text{boost}}(x) = \sum_{i=1}^M \alpha_i f_i(x) \quad (3.76)$$

Вага кожної моделі обчислюється на основі її помилки

$$\alpha_i = \frac{1}{2} \ln \left( \frac{1 - \epsilon_i}{\epsilon_i} \right) \quad (3.77)$$

Адаптивне комбінування результатів оцінки глибини

$$D_{\text{final}}(x) = \begin{cases} D_1(x), & \text{якщо } |\nabla D_1(x)| \leq \tau \\ D_2(x), & \text{якщо } |\nabla D_1(x)| > \tau \end{cases} \quad (3.78)$$

Інтегральна модель ідентифікації об'єднує всі компоненти в єдину систему

$$\theta = \mathcal{P}(\mathcal{F}(I_1, I_2), \mathcal{Z}(I_2), \mathcal{D}(I_2)) \quad (3.79)$$

де:

$\mathcal{F}$  – модель оптичного потоку;

$Z$  – модель оцінки глибини;

$\mathcal{D}$  – модель виявлення об'єктів;

$\mathcal{P}$  – модель обчислення параметрів.

Повна система рівнянь для обчислення параметрів має вигляд

$$\left\{ \begin{array}{l} \text{Boxes} = \mathcal{D}(I_2) \\ F = \mathcal{F}(I_1, I_2) \\ D = Z(I_2) \\ \forall \text{box} \in \text{boxes}: v = \frac{1}{|\Omega|} \iint_{\Omega} \|F\| dx dy \\ \theta = \arctan\left(\frac{\iint_{\Omega} \sin\theta(F) dx dy}{\iint_{\Omega} \cos\theta(F) dx dy}\right) \\ z = \underset{(x,y) \in \Omega}{\text{median}} D(x, y) \end{array} \right. \quad (3.80)$$

Сформовано моделі дистанційної ідентифікації параметрів динамічних об'єктів, розглянуто моделі ідентифікації за відеоданими, методи оптичного потоку та підходи до об'єднання моделей. Показано, що інтеграція DETR, Optical Flow, GeoNet і ансамблевих методів дає змогу підвищити точність локалізації, оцінювання швидкості, напрямку руху та просторового положення об'єктів.

Запропоновані моделі забезпечують перехід від покадрового виявлення до просторово-часового опису динамічного об'єкта. Вони поєднують координати та межі об'єкта, вектори оптичного потоку, карти глибини і результати трансформерного аналізу, що дозволяє отримати більш повний набір параметрів для систем керування та підтримки прийняття рішень.

Сформовано моделі дистанційної ідентифікації. На відміну від існуючих підходів (наприклад, покадрових детекторів YOLO або окремого використання оптичного потоку), розроблено інтегральну модель на основі DETR та методів оптичного потоку, яка забезпечує одночасне виявлення, відстеження та оцінювання кінематичних параметрів об'єкта [46, 91, 92]. У порівнянні з класичними методами, запропоновані математичні моделі для визначення швидкості, напрямку, траєкторії та прискорення враховують не лише зміщення центру, а й внутрішнє поле руху об'єкта [22, 42, 96]. Наукова новизна полягає у вперше запропонованому об'єднанні GeoNet з ансамблевими методами Bagging

і Boosting, що дозволяє підвищити стійкість оцінок у зонах оклюзій [8, 15, 93]. Практична цінність моделей реалізується в навігації безпілотних літальних апаратів (завдяки роботі в умовах зміни освітлення та туману), у системах уникнення зіткнень автономного транспорту (через точну оцінку глибини GeoNet) та в ранній медичній діагностиці (аналіз кровотоку та руху серця) [10, 12, 30, 64].

Порівняльний аналіз підтвердив, що використання ансамблевої агрегації та моделей просторово-часової пам'яті підвищує стійкість ідентифікації в умовах шумів, оклюзій і зміни освітлення. Отримані результати стали основою для програмної реалізації інформаційної технології та експериментального тестування в наступному розділі.

### 3.5 Висновки до розділу

1. Запропоновано модель дистанційної ідентифікації, яка встановлює зв'язок між відеоданими та множиною параметрів. Визначено ієрархію параметрів: просторові (координати, глибина, розміри, форма), кінематичні (швидкість, напрям, траєкторія, прискорення) та динамічні (маса, центр мас, моменти інерції). Архітектура системи подана як набір модулів.

2. Розроблено аналітичну модель ідентифікації параметрів з використанням DETR та оптичного потоку. Модель DETR забезпечує виявлення об'єктів та їх локалізацію у кадрі. Модель оптичного потоку обчислює поле векторів руху між послідовними кадрами. Швидкість об'єкта визначається як середня величина вектора оптичного потоку в області об'єкта, а напрям як середнє арифметичне або векторне усереднення кутів. Розглянуто також архітектуру кодера-декодера трансформера та інтеграцію з GeoNet.

3. Деталізовано математичні моделі для обчислення основних параметрів динамічних об'єктів на основі поля оптичного потоку. Швидкість об'єкта визначається як інтегральна або дискретна середня величина вектора потоку в області об'єкта. Напрямок руху обчислюється через векторне усереднення кутів або арифметичне середнє. Просторове положення задається координатами

центру об'єкта та медіанним або середнім значенням глибини. Траєкторія представляється у вигляді дискретної множини положень у часі з можливістю згладжування ковзним середнім. Прискорення визначається як похідна швидкості за часом.

4. Запропоновано узагальнену схему об'єднання, яка інтегрує відеопотік, оптичний потік, трансформерну мережу та ансамблевий модуль. Сумісна модель GeoNet реалізує сумісну оцінку глибини та оптичного потоку на основі геометричних обмежень. Ансамблева модель Bagging передбачає усереднення результатів кількох моделей для зменшення дисперсії. Адаптивна модель Boosting використовує зважену комбінацію з виправленням помилок, де вага кожної моделі обернено пропорційна її помилці. Інтегральна модель об'єднує всі компоненти в єдину систему, а повна система рівнянь задає обчислення bounding boxів, оптичного потоку, глибини, швидкості, напрямку та просторового положення. Узагальнено, що запропоновані моделі забезпечують перехід від покадрового виявлення до просторово-часового опису динамічного об'єкта. Вони поєднують координати та межі об'єкта, вектори оптичного потоку, карти глибини і результати трансформерного аналізу, що дозволяє отримати більш повний набір параметрів для систем керування та підтримки прийняття рішень. Порівняльний аналіз підтвердив, що використання ансамблевої агрегації та моделей просторово-часової пам'яті підвищує стійкість ідентифікації в умовах шумів, оклюзій і зміни освітлення. Отримані результати стали основою для програмної реалізації інформаційної технології та експериментального тестування в наступному розділі.

## **РОЗДІЛ 4 ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ ДИСТАНЦІЙНОЇ ІДЕНТИФІКАЦІЇ ПАРАМЕТРІВ ДИНАМІЧНИХ ОБ’ЄКТІВ**

У четвертому розділі описано структуру та програмну реалізацію інформаційної технології дистанційної ідентифікації параметрів динамічних об’єктів. Наведено результати експериментального поєднання трансформерів, методів оптичного потоку, оцінювання глибини та ансамблевих підходів, а також виконано аналіз ефективності отриманих рішень.

### **4.1 Структура інформаційної технології дистанційної ідентифікації**

Розроблена інформаційна технологія ДПДО є комплексним підходом до аналізу динамічних об’єктів у відеопотоках. Загальна модифікована структура GeoNet, яка лягла в основу запропонованої технології, на рис. 4.1 показано модифіковану структуру інформаційної технології дистанційної ідентифікації параметрів динамічних об’єктів, яка охоплює повний цикл обробки від вхідних даних до формування вихідних результатів для систем підтримки прийняття рішень [42, 94]. Вхідними даними системи є відеопотік, сенсорні дані від різномірних датчиків та статичні зображення, які надходять на етап попередньої обробки сигналів, що включає синхронізацію кадрів, нормалізацію інтенсивності, вирівнювання геометричних спотворень та фільтрацію шумів [22, 93, 96]. Після попередньої обробки дані передаються до модуля виділення ознак та формування представлень, який використовує згорткові нейронні мережі та трансформери для отримання просторових, часових та просторово-часових характеристик динамічних об’єктів [4, 14, 46]. Нова структура передбачає інтеграцію класичних компонентів оптичного потоку та геометричного аналізу сцени з трансформерними механізмами уваги для врахування глобального контексту та з ансамблевими методами для підвищення стійкості результатів ідентифікації в умовах шумів, оклюзій та зміни освітлення [94, 96].

Що забезпечує можливість спільного використання всіх компонентів для отримання характеристик динамічних об’єктів [8, 94, 95].

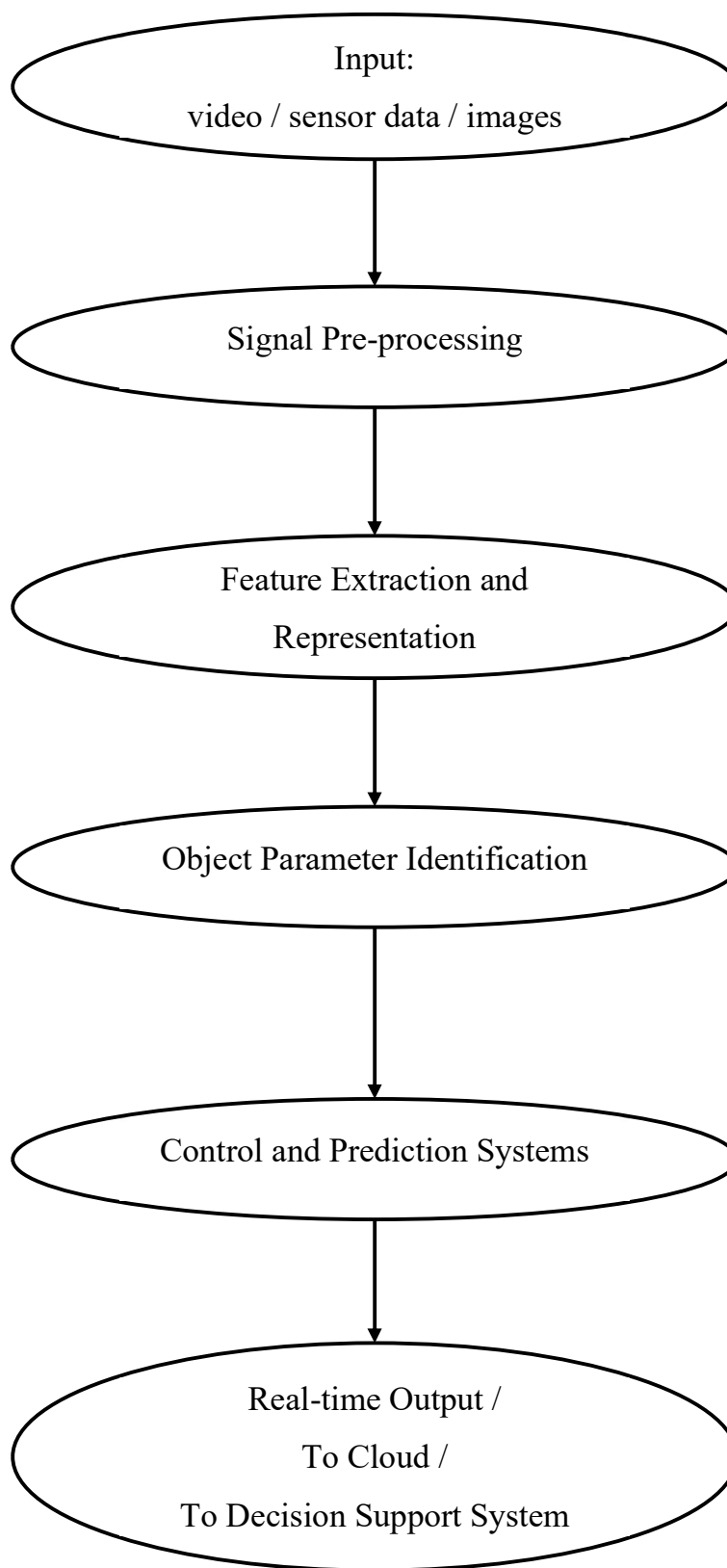


Рисунок 4.1 – Модифікована структура

На рис. 4.2, 4.3 показано етапи аналізу вхідних даних з відео, сенсорів та зображень, які проходять через модулі попередньої обробки, виділення ознак, оптичного потоку, оцінювання глибини, ідентифікації параметрів та формування результатів.

Оригінальний GeoNet включає компоненти оцінки глибини DepthNet, оцінки пози PoseNet), оцінки оптичного потоку ResFlowNet і працює без вчителя unsupervised. У запропонованій модифікації додано нові елементи: трансформер, що моделює просторово-часові залежності Transformer dynamic flow, розширений підхід до трекінгу параметрів швидкості, напрямку руху тощо. Наявність блоків Transformer parameters, Optical flow information з підключенням до ансамблевих методів також свідчить про спроба об'єднати класичні компоненти, оптичний потік, геометрія, трансформери для глобального контексту, та ансамблеві методи для підвищення точності.

У табл. 4.1 наведено методи визначення основних параметрів динамічних об'єктів та відповідні абсолютні похибки та узагальнено параметри положення, швидкості, прискорення та орієнтації, а також методи їх визначення в межах запропонованої інформаційної технології.

Таблиця 4.1 – Методи визначення ідентифікації параметрів динамічних об'єктів

Параметр	Метод визначення
Положення	Stereo Triangulation
Швидкість	Optical Flow + Kalman Filter
Прискорення	IMU Sensor Fusion
Орієнтація	Quaternion from EMG

Після попередньої обробки сигналів, яка включає синхронізацію даних, нормалізацію інтенсивності та вирівнювання геометричних спотворень, отримані ознаки з кожної гілки подаються на етап виділення представлень, де використовуються згорткові нейронні мережі або трансформери, зокрема Universal Transformer Memory для моделювання довготривалих просторово-часових залежностей [4, 40, 46, 48].

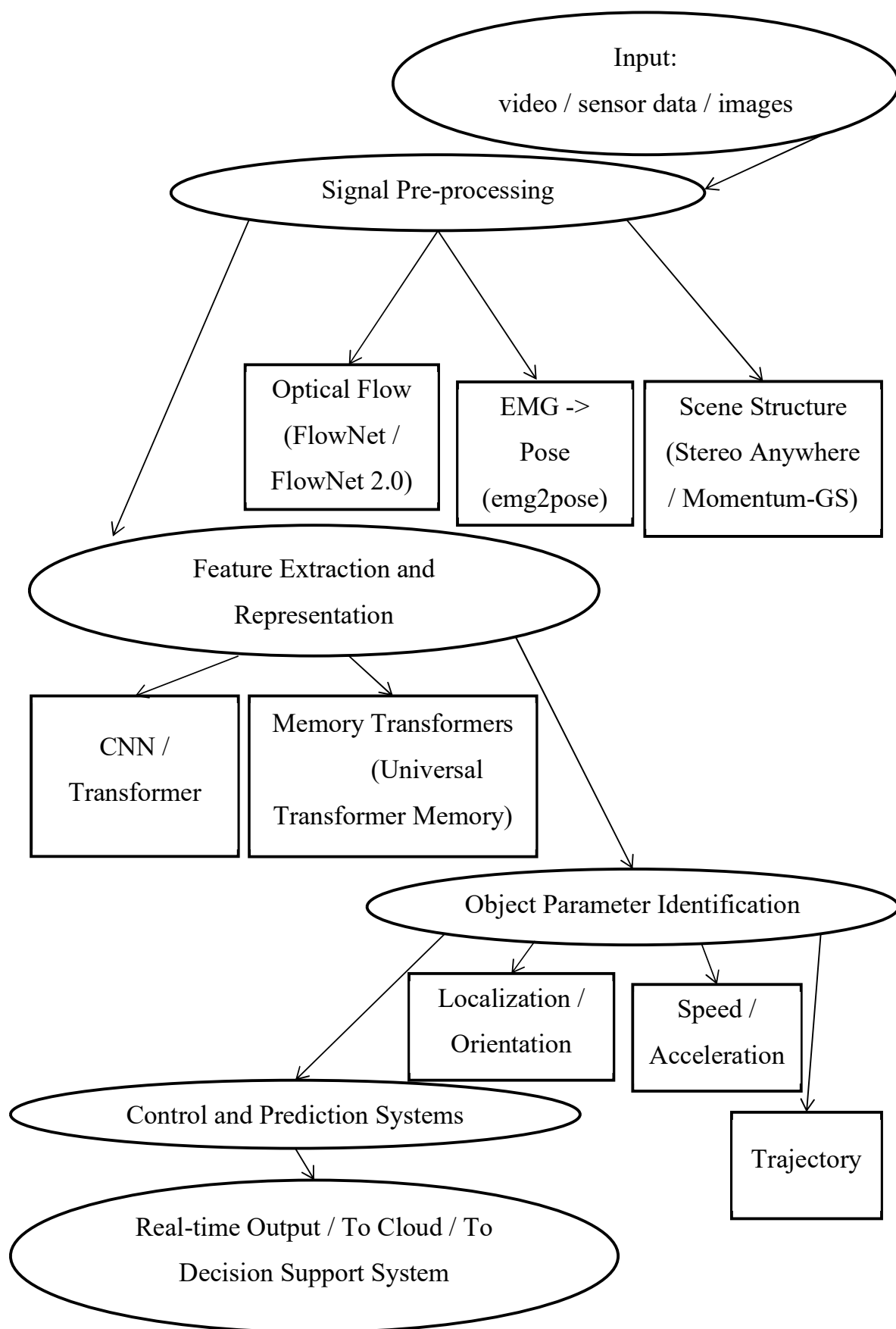


Рисунок 4.2 – Загальна схема обробки сигналів та ідентифікація об'єктів



На рис. 4.3 деталізовано структуру обробки сигналів та ідентифікації об'єктів із урахуванням модулів оптичного потоку, оцінювання глибини, виділення ознак і формування результатів [94, 95]. На першому етапі виконується попередня обробка вхідного відеопотоку, що включає синхронізацію кадрів, нормалізацію інтенсивності та фільтрацію шумів. Другий етап реалізує три паралельні гілки обробки: На шостому етапі формуються вихідні дані у вигляді таблиць, візуалізованих зображень та передаються до систем підтримки прийняття рішень через інтерфейси MQTT, ROS або API.

Вхідний відеопотік або послідовність зображень спочатку проходить через етап попередньої обробки, який включає синхронізацію кадрів, нормалізацію інтенсивності та вирівнювання геометричних спотворень [22, 96].

Паралельно формуються три основні гілки обробки, а саме гілка оптичного потоку на основі методів FlowNet або FlowNet 2.0 для оцінювання векторів руху пікселів між послідовними кадрами, гілка просторової структури сцени з використанням методів Stereo Anywhere або Momentum-GS для реконструкції тривимірної геометрії, та гілка виділення ознак за допомогою згорткових нейронних мереж або трансформерів, зокрема Universal Transformer Memory для моделювання довготривалих просторово-часових залежностей [8, 35, 36, 39, 40].

Отримані результати з кожної гілки інтегруються в модулі ідентифікації параметрів, які обчислюють локалізацію об'єкта в кадрі, його орієнтацію в просторі, миттєву швидкість, прискорення та прогнозовану траєкторію руху [95, 96, 73, 74].

Завершальним етапом обробки є формування вихідних даних у вигляді структурованих звітів або візуалізованих результатів, які можуть бути передані до систем підтримки прийняття рішень, хмарних платформ або локальних систем керування через інтерфейси MQTT або ROS [38, 43, 94, 75, 76].

Така багатоканальна архітектура забезпечує комплексний аналіз динамічних сцен з урахуванням як рухових, так і просторових характеристик об'єктів, що підвищує надійність та інформативність дистанційної ідентифікації параметрів [93-95].

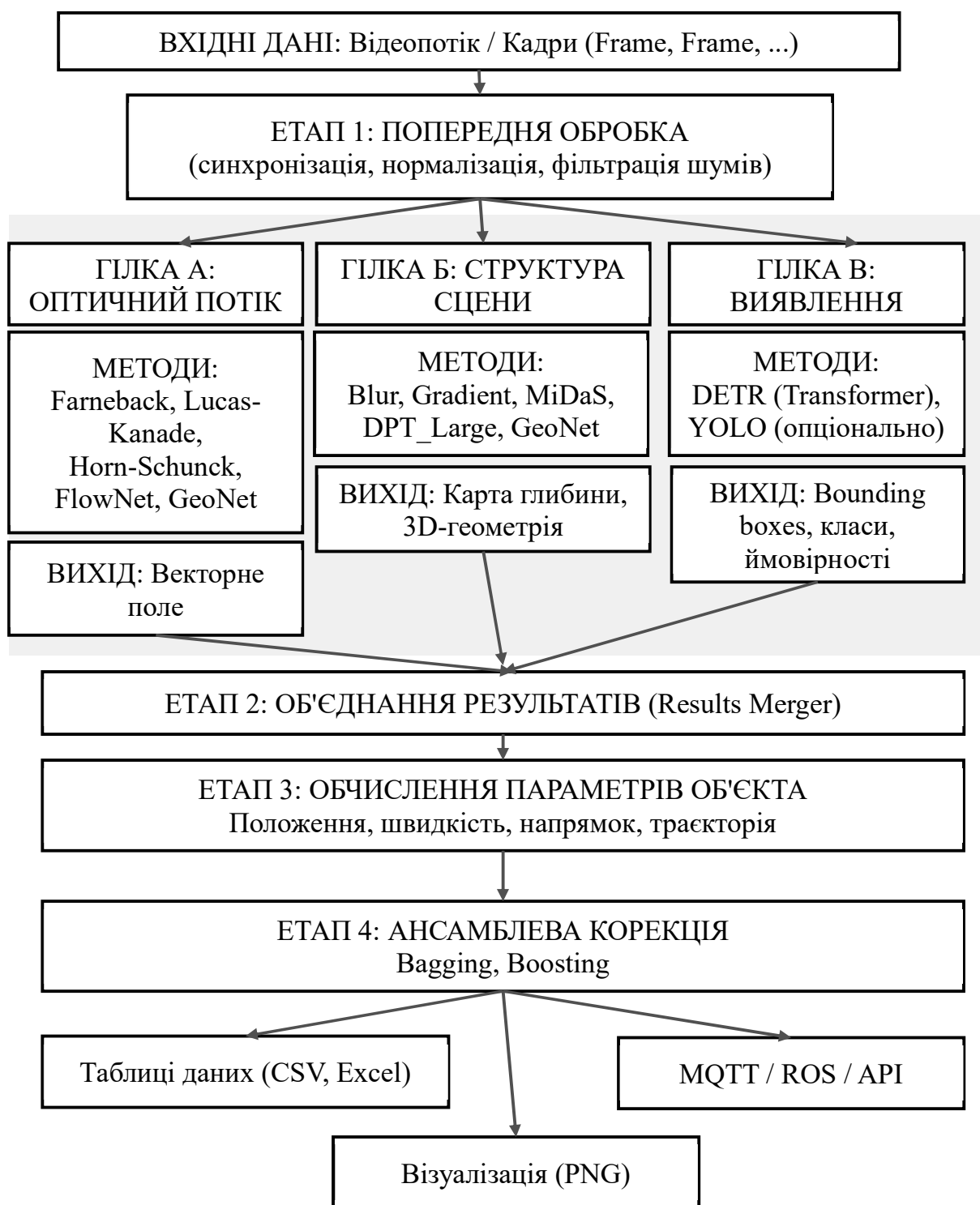


Рисунок 4.3 – Обробка сигналів та ідентифікація об'єктів

Загальна архітектура

$$\mathcal{S} = \{\mathcal{D}, \mathcal{F}, \mathcal{Z}, \mathcal{P}, \mathcal{V}\} \quad (4.1)$$

де:

$\mathcal{D}$  – модуль виявлення об’єктів (DETR / YOLO);

$\mathcal{F}$  – модуль оптичного потоку (Farneback / Lucas-Kanade / Horn-Schunck / FlowNet / GeoNet);

$\mathcal{Z}$  – модуль оцінки глибини (Blur / Gradient / MiDaS / DPT / GeoNet / Bagging / Boosting);

$\mathcal{P}$  – модуль обчислення параметрів (швидкість, напрямок, траєкторія, положення);

$\mathcal{V}$  – модуль візуалізації результатів.

Математична модель системи

Інформаційна технологія реалізує наступну послідовність перетворень

$$\theta = \mathcal{P}(\mathcal{F}(I_1, I_2), \mathcal{Z}(I_2), \mathcal{D}(I_2)) \quad (4.2)$$

де  $I_1$  та  $I_2$  – два послідовних кадри відеопослідовності.

Обчислення параметрів для кожного виявленого об’єкта виконується за формулами:

$$v = \frac{1}{N_\Omega} \sum_{(x,y) \in \Omega} \|F(x, y)\| \quad (4.3)$$

$$\theta = \arctan\left(\frac{\sum_{(x,y) \in \Omega} \sin\theta(x,y)}{\sum_{(x,y) \in \Omega} \cos\theta(x,y)}\right) \quad (4.4)$$

$$z = \text{median}\{D(x, y) \mid (x, y) \in \Omega\} \quad (4.5)$$

$$T(t) = \{p(t_1), p(t_2), \dots, p(t_n)\} \quad (4.6)$$

## 4.2 Програмна реалізація моделей дистанційної ідентифікації

На рис. 4.4 показано реалізовану модель інформаційної технології дистанційної ідентифікації параметрів динамічних об’єктів, яка включає три основні блоки обробки, а саме обробку відеопотоку, використання моделі DETR для виявлення об’єктів та застосування оптичного потоку для аналізу руху [46, 95, 96]. Перший блок забезпечує отримання відеоданих з камери або збереженого

файлу, покадрову екстракцію зображень та попередню нормалізацію вхідних кадрів для подальшої обробки [22, 94].

Другий блок реалізує детекцію об'єктів за допомогою трансформерної моделі DETR, яка виконує одночасне прогнозування класів об'єктів, обмежувальних рамок та ймовірностей без потреби в механізмах придушення не-максимумів або попереднього формування областей-кандидатів [14, 46, 91]. Третій блок виконує обчислення оптичного потоку для оцінювання векторів руху пікселів між послідовними кадрами, що дає змогу визначати швидкість, напрямок та інтенсивність руху ідентифікованих об'єктів [7, 22, 41, 60].

Результати роботи моделі формуються у вигляді візуалізованих кадрів з накладеними обмежувальними рамками виявлених об'єктів, кольорових карт оптичного потоку та табличних даних про просторове положення, швидкість і напрямок руху кожного об'єкта [94-96]. Кольорами на схемі відмічено різні елементи реалізованої моделі, що дозволяє візуально відокремити етапи обробки відеопотоку, блок детекції DETR та модуль обчислення оптичного потоку для подальшого аналізу та налагодження системи [46, 96, 91].

Запропонована архітектура забезпечує модульність та гнучкість системи, дозволяючи замінювати окремі компоненти без перебудови всього конвеєру обробки. Наприклад, замість DETR може бути використаний альтернативний детектор YOLO для порівняльного аналізу точності та швидкодії, або замість методу Farneback може застосовуватися нейромережевий підхід FlowNet для отримання більш деталізованого поля оптичного потоку [14, 46, 95, 96]. Така модульна структура є основою для подальшого розширення функціональності, зокрема інтеграції з GeoNet для сумісної оцінки глибини та оптичного потоку, а також для впровадження ансамблевих методів Bagging та Boosting з метою підвищення стійкості ідентифікації параметрів динамічних об'єктів в умовах шумів, оклюзій та зміни освітлення [22, 94, 96].

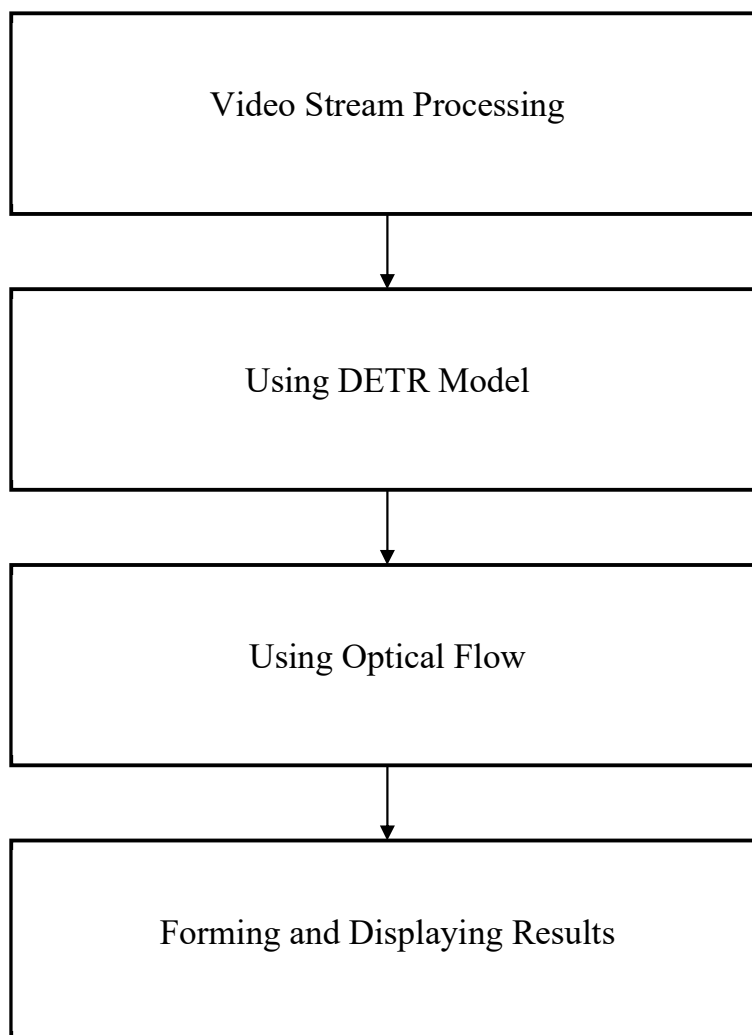


Рисунок 4.4 – Реалізована модель

На рис. 4.5 показано розширений опис реалізованої моделі. Яка включає наступні елементи у вигляді конвеєру:

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.

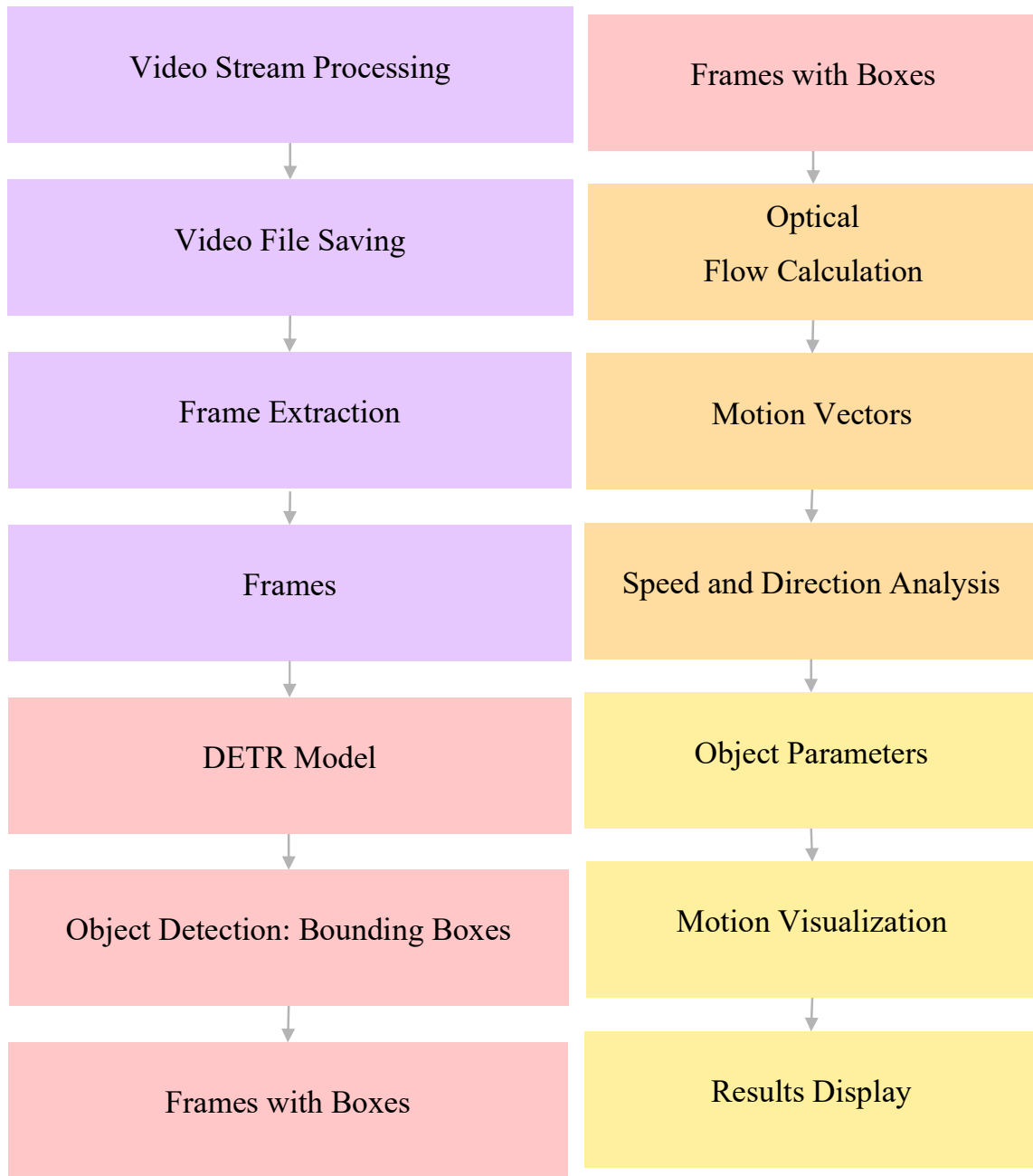


Рисунок 4.5 – Розширений опис реалізованої моделі

Реалізація модуля виявлення об'єктів DETR (Detection Transformer) - використовує механізм уваги:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (4.7)$$

$$\mathcal{L}_{\text{Hungarian}}(y, \hat{y}) = \sum_{i=1}^N \left[ -\log \hat{p}_{\hat{\sigma}(i)}(c_i) + \mathbf{1}_{\{c_i \neq \emptyset\}} \mathcal{L}_{\text{box}}(b_i, \hat{b}_{\hat{\sigma}(i)}) \right] \quad (4.8)$$

YOLO You Only Look Once альтернативна модель для порівняння.  
Реалізація модуля оцінки глибини.

Метод Blur (розмиття):

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (4.9)$$

$$D_{\text{blur}} = I_{\text{gray}} * G \quad (4.10)$$

Метод Gradient (градієнтний)

$$D_{\text{grad}} = |I_{\text{gray}} * L|, L = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad (4.11)$$

Метод MiDaS

$$\mathcal{L}_{\text{SI}} = \frac{1}{N} \sum_{i,j} \left[ (\log y_i - \log \hat{y}_i) - (\log y_j - \log \hat{y}_j) \right]^2 \quad (4.12)$$

Метод DPT (Dense Prediction Transformer)

$$z_l = \text{MSA}(\text{LN}(z_{l-1})) + z_{l-1} \quad (4.13)$$

Метод Bagging (ансамблевий)

$$D_{\text{bag}}(x, y) = \frac{1}{M} \sum_{i=1}^M D_i(x, y) \quad (4.14)$$

Метод Boosting (адаптивний)

$$D_{\text{final}}(x) = \begin{cases} D_1(x), & |\nabla D_1(x)| \leq \tau \\ D_2(x), & |\nabla D_1(x)| > \tau \end{cases} \quad (4.15)$$

#### 4.3 Програмна реалізація з використанням ансамблевих методів

Для підвищення надійності ідентифікації в умовах динамічної зміни середовища застосовуються ансамблеві методи машинного навчання. Структура ансамблевих модулів у системі наведена на рис. 4.6.

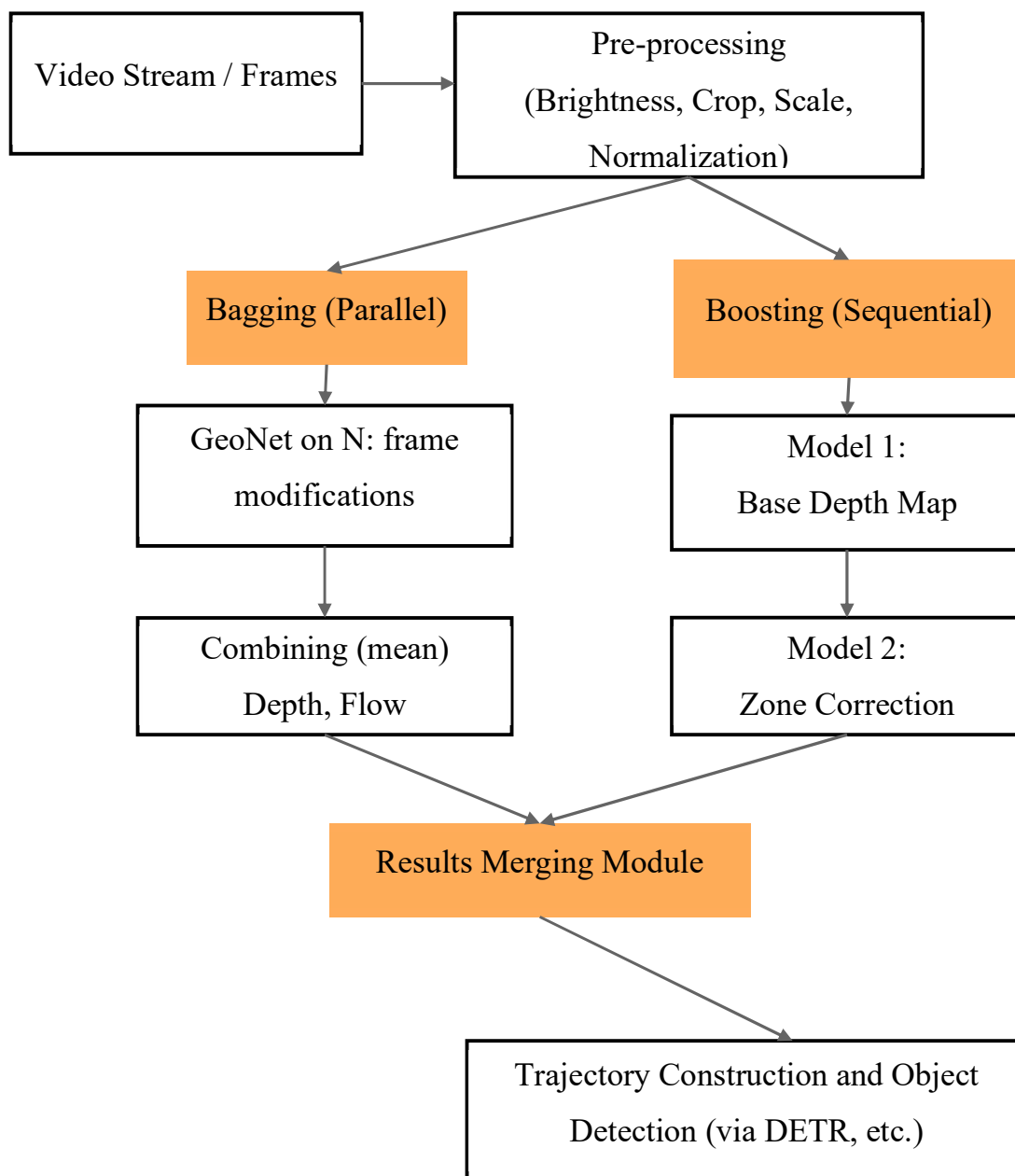


Рисунок 4.6 – Структура ансамблевих модулів у системі

На рис. 4.6 наведено результат «Структура ансамблевих модулів у системі», який використовується для порівняння оптичного потоку, карти глибини та об'єднаного представлення сцени під час дистанційної ідентифікації параметрів динамічних об'єктів.

Для бегінгу об'єднання карт глибини методом середнього значення. Для бегінгу карти глибини, отримані з різних підвибірок кадрів, об'єднуються



шляхом усереднення. Це дозволяє згладити локальні артефакти та підвищити стабільність результатів.

Для бустінгу зважене об'єднання результатів з урахуванням локальних похибок (приклад використання нейронної мережі-коректора). Для бустінгу застосовується зважене об'єднання результатів, де ваги призначаються на основі локальної похибки кожної моделі. У разі значного відхилення результатів у певній області рішення приймається на користь тієї моделі, яка показала нижчу похибку у сусідніх регіонах. Для цього використовується нейронна мережа-коректор.

Для підвищення точності ідентифікації запропоновано ансамблеве комбінування результатів різних методів оптичного потоку

$$F_{\text{ensemble}}(x, y) = \sum_{i=1}^M w_i F_i(x, y), \sum_{i=1}^M w_i = 1 \quad (4.16)$$

де ваги  $w_i$  визначаються на основі якості роботи кожного методу на еталонних даних.

Для оцінки глибини запропоновано адаптивне Ансамблеве комбінування методів глибини

$$D_{\text{final}}(x) = \operatorname{argmax}_k P(D_k(x) \mid \text{context}) \quad (4.17)$$

$$w_k = \frac{\frac{1}{\text{MAE}_k}}{\sum_{j=1}^M \frac{1}{\text{MAE}_j}} \quad (4.18)$$

Повний інтегральний ансамбль включає комбінування методів оптичного потоку та глибини

$$\Theta_{\text{ensemble}} = \mathcal{P}(F_{\text{ensemble}}, D_{\text{ensemble}}, \mathcal{D}(I_2)) \quad (4.19)$$

#### 4.4 Результат поєднання трансформерів і методів оптичного потоку

Для демонстрації роботи розробленої інформаційної технології було проведено серію експериментів на тестових відеоданих. На рис. 4.7 наведено результати роботи моделі DETR для виявлення об'єктів у кадрі.

У межах цього підрозділу виконано порівняння комбінованих схем за однаковою логікою оброблення: спочатку формується поле оптичного потоку, далі оцінюється карта глибини, після чого результати накладаються на зображення кадру для подальшої інтерпретації. Такий підхід дає змогу зіставити класичні та нейромережеві алгоритми за стабільністю, деталізацією та здатністю зберігати просторову структуру сцени [100-106].

Для аналізу враховано якість виявлення об'єктів, узгодженість напрямку руху, чіткість меж на карті глибини, стійкість до шумів і придатність результату до подальшого використання в модулі ідентифікації. Саме тому далі наведено окремі результати для Farneback, FlowNet, GeoNet, Horn-Schunck і Lucas-Kanade у поєднанні з різними моделями оцінювання глибини.

Отримані візуалізації не дублюють одна одну, а відображають різні режими роботи інформаційної технології: від базового оптичного потоку й простих способів оцінювання глибини до інтегрованих схем, у яких поєднуються виявлення, рух і просторова інтерпретація кадру. Саме тому кожен наступний підрозділ акцентує окремий тип комбінації та його практичну інтерпретацію.



Рисунок 4.7 – Результат виявлення об'єктів за допомогою моделі DETR

На рис. 4.7 модель DETR успішно локалізує об'єкти в кадрі, виділяючи їх обмежувальними рамками. Виявлено 6 об'єктів, що відповідає реальній кількості рухомих цілей у сцені та наведено результат роботи моделі DETR для виявлення об'єктів у кадрі відеопослідовності, яка реалізує підхід end-to-end без потреби в попередньому формуванні областей-кандидатів або механізмах придушення не-

максимумів [14, 46, 78-80, 91]. Модель DETR успішно локалізує об'єкти в тестовому кадрі, виділяючи їх обмежувальними рамками з відповідними мітками класів та коефіцієнтами довіри [46, 94]. За результатами детекції на наведеному рисунку виявлено шість об'єктів, що відповідає реальній кількості рухомих цілей у сцені, зокрема транспортних засобів та пішоходів [95, 96]. Кожен виявлений об'єкт супроводжується числовим значенням ймовірності правильної класифікації, яке для більшості об'єктів перевищує 0,74, що свідчить про високу достовірність роботи детектора на тестовому відеоматеріалі [14, 46]. Обмежувальні рамки точно окреслюють просторові межі кожного об'єкта, що є необхідною умовою для подальшого обчислення їх кінематичних параметрів, зокрема швидкості, напрямку руху та траєкторії, на основі аналізу оптичного потоку між послідовними кадрами [22, 42, 91]. Застосування трансформерної архітектури DETR забезпечує врахування глобальних контекстних залежностей між об'єктами в кадрі завдяки механізмам багатоголової уваги, що дозволяє коректно локалізувати навіть тісно розташовані або частково перекриті об'єкти [4, 14, 46, 81-83]. На рис. 4.8 показано візуалізацію обчислених параметрів для виявлених об'єктів, а саме: їх просторового положення, миттєвої швидкості та напрямку руху.



Рисунок 4.8 – Візуалізація ідентифікованих параметрів об'єктів

Подальший аналіз показав, що комбінування різних методів оцінки оптичного потоку та глибини дає різні результати. На рис. 4.9 показано приклад роботи системи з використанням алгоритму оптичного потоку Фарнбека (Farneback) для оцінки руху та моделі MiDaS для оцінки глибини сцени. Це класичний підхід Farneback + MiDaS.

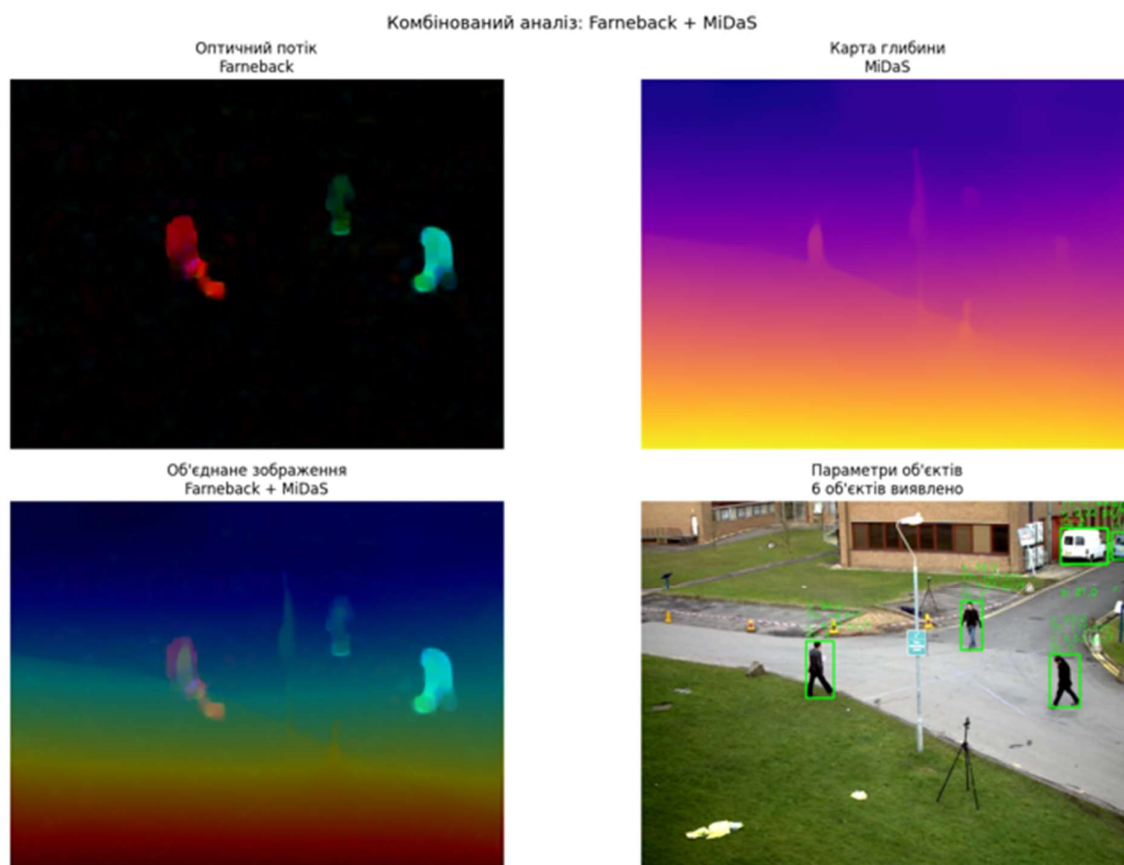


Рисунок 4.9 – Візуалізація оптичного потоку (Farneback), карти глибини (MiDaS) та їх об'єднання

На рис. 4.9 зліва показано поле оптичного потоку, де кольором закодовано напрямок та інтенсивність руху пікселів. У центрі представлено карту глибини, отриману за допомогою моделі MiDaS, де теплі кольори (жовтий, помаранчевий) відповідають ближчим об'єктам, а холодні (синій, фіолетовий) – віддаленим. Справа наведено результат об'єднання цих даних, що дозволяє одночасно аналізувати як динаміку, так і просторову структуру сцени.

Аналогічний аналіз було проведено для пари «нейромережевий оптичний потік FlowNet + DPT\_Large». Результати цього експерименту наведено на рис. 4.10.

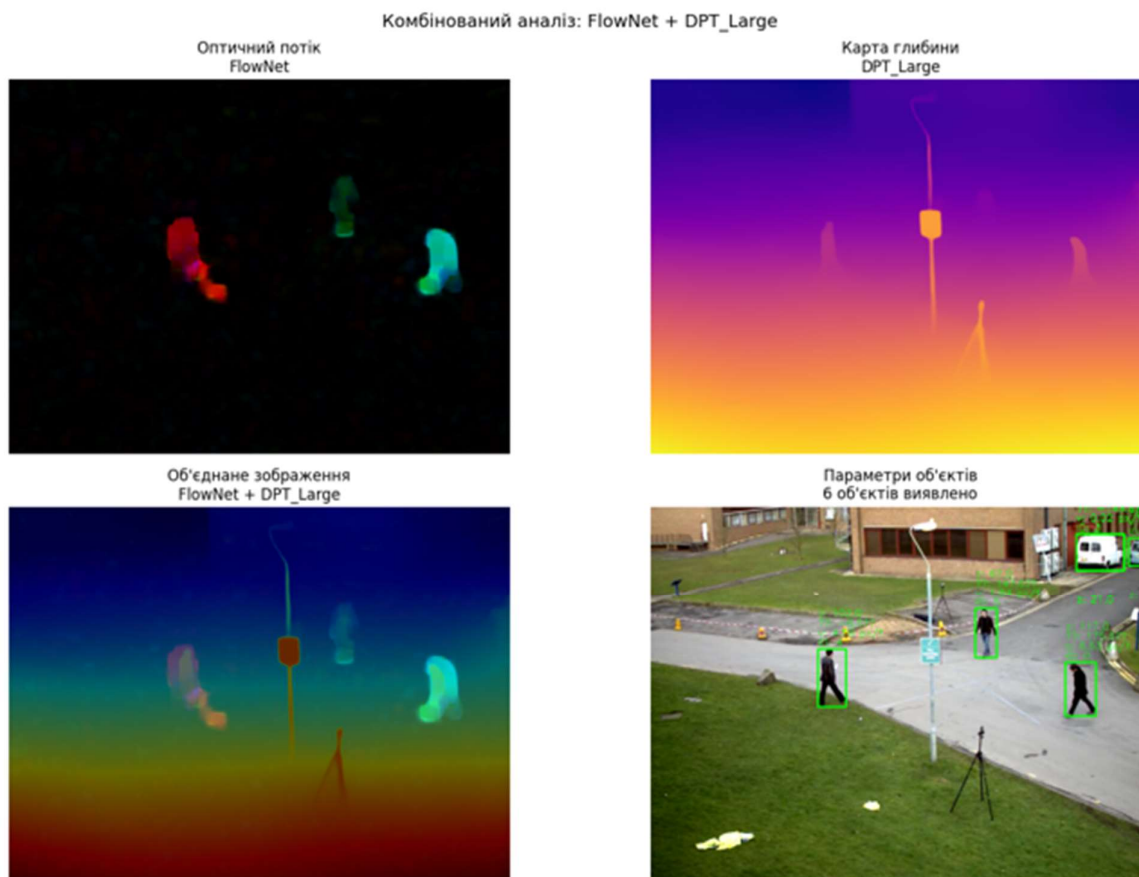


Рисунок 4.10 – Візуалізація оптичного потоку (FlowNet), карти глибини (DPT\_Large) та їх об'єднання

На рис. 4.10 використання нейромережевих методів (FlowNet) дозволяє отримати більш деталізоване поле оптичного потоку порівняно з класичним алгоритмом Фарнбека, особливо на межах об'єктів. Оцінка глибини за допомогою DPT\_Large також демонструє вищу роздільну здатність. Використання FlowNet дає більш деталізоване поле оптичного потоку, особливо на межах об'єктів (артефакти руху зменшуються). DPT\_Large забезпечує вищу роздільну здатність карти глибини порівняно з MiDaS, що дозволяє розрізняти об'єкти на різних відстанях з більшою точністю.

Інтегральний підхід GeoNet, який одночасно оцінює глибину та оптичний потік, показав себе як ефективний інструмент для спільного аналізу (рис. 4.11).



Результати демонструють узгодженість між просторовою та часовою інформацією: рухомі об'єкти чітко відокремлюються від статичного фону, а їх просторове положення узгоджується з картою глибини.

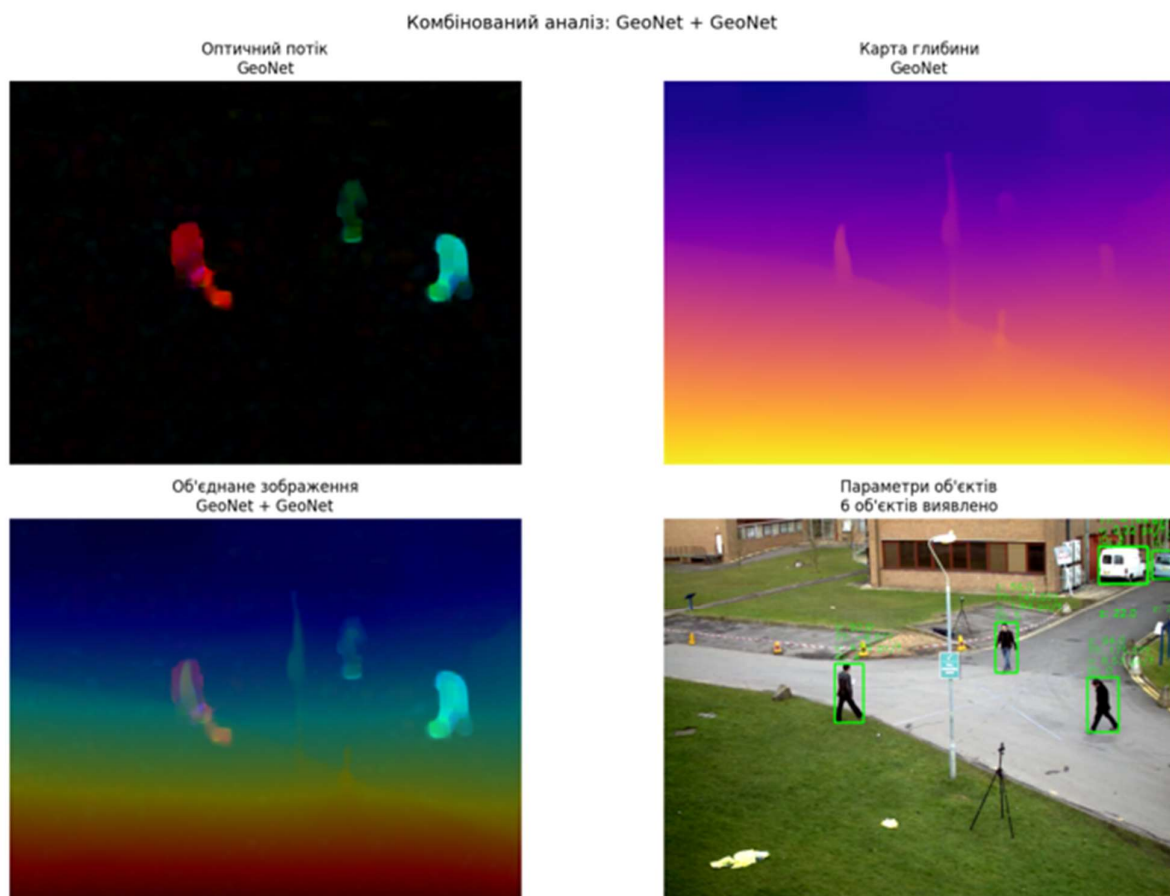


Рисунок 4.11 – Візуалізація результатів роботи інтегрованої моделі GeoNet

Для підвищення стійкості ідентифікації до шумів було застосовано ансамблевий метод бегінгу (Bagging). На рис. 4.12 представлено результат роботи системи, де для оцінки глибини використовується ансамбль моделей (Bagging), а оптичний потік обчислюється методом Горна–Шунка (HornSchunck). представлено результат поєднання глобального методу оптичного потоку Горна-Шунка для оцінювання руху та ансамблевого методу бегінгу для оцінювання глибини сцени, що демонструє ефективність ансамблевих підходів для підвищення стійкості ідентифікації параметрів динамічних об'єктів в умовах шумів та нестабільного освітлення [15, 93, 60].

Ліва панель візуалізує поле оптичного потоку, отримане методом Горна-

Шунка, яке на відміну від локальних методів забезпечує густе та гладке поле векторів руху за рахунок мінімізації функціонала, що поєднує обмеження сталості яскравості з умовою глобальної гладкості поля руху [11, 41, 60].

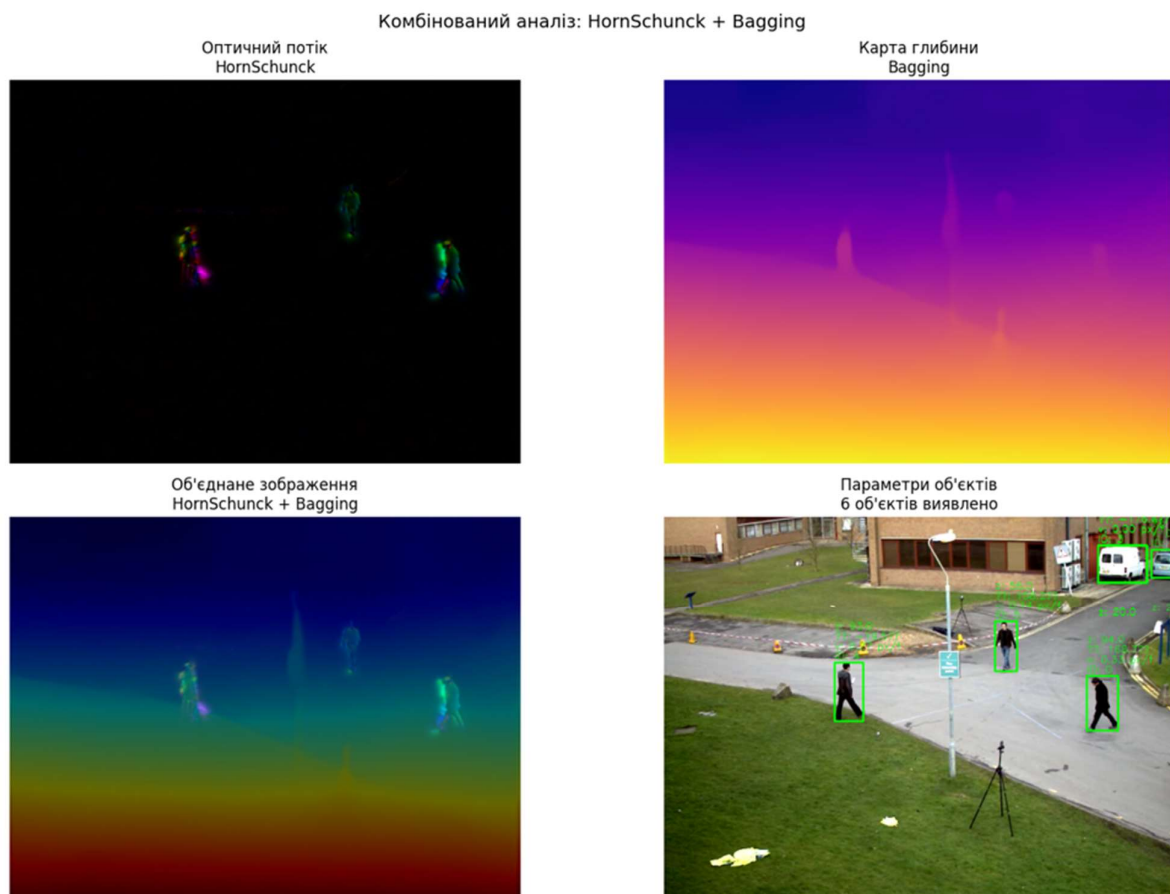


Рисунок 4.12 – Візуалізація оптичного потоку (HornSchunck) та ансамблевої оцінки глибини (Bagging)

Кольорове кодування на лівій панелі відображає напрямок руху, де червоно-жовті відтінки відповідають руху вправо та вгору, а синьо-зелені відтінки відповідають руху вліво та вниз, причому насиченість кольору пропорційна інтенсивності переміщення пікселів між кадрами [22, 60].

Кольорова шкала на центральній панелі використовує теплі тони для позначення ближніх об'єктів та холодні тони для віддалених, причому ансамблеве усереднення дозволяє згладити локальні шумові артефакти та підвищити стабільність оцінки глибини порівняно з використанням окремих моделей [42, 93, 96].

Права панель об'єднує інформацію про оптичний потік та глибину шляхом накладання кольорового поля векторів руху на карту глибини, де насиченість кольору відповідає інтенсивності руху, а яскравість відповідає відстані до об'єкта, що дозволяє одночасно аналізувати як динамічні, так і просторові характеристики сцени [42, 94, 96].

Поєднання глобально гладкого поля оптичного потоку, отриманого методом Горна-Шунка, з ансамблевою оцінкою глибини забезпечує високу стійкість до шумів та артефактів, що робить цю комбінацію доцільною для використання в системах відеоспостереження зі складними умовами зйомки, зокрема в сутінках, тумані або під час дощу [15, 93, 60, 62].

Завершальним етапом є інтеграція модуля виявлення об'єктів DETR з модулем оцінки руху Optical Flow. Як показано на рис. 4.13 обмежувальні рамки, отримані від DETR, накладаються на візуалізацію оптичного потоку. Це дозволяє для кожного виявленого об'єкта обчислити його індивідуальні параметри руху, такі як швидкість та напрямок.

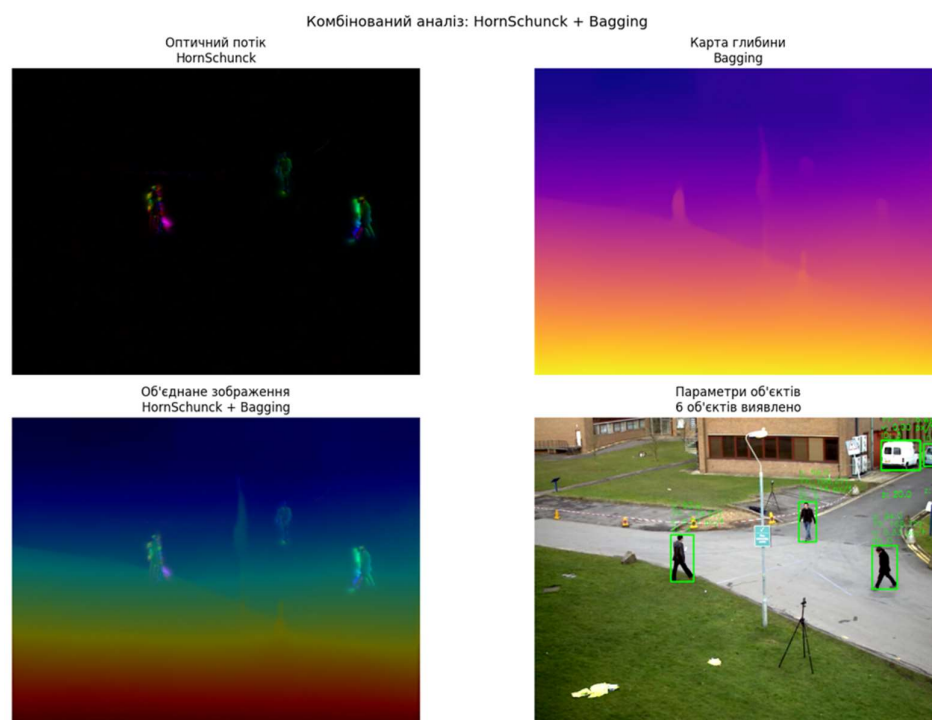


Рисунок 4.13 – Результат інтеграції моделі виявлення DETR з візуалізацією оптичного потоку



На рис. 4.14–4.26 показано результати візуалізації оптичного потоку та карти глибини після об'єднання. Результати отримані за допомогою розробленого застосунку, у якому завантажується приклад відео з бібліотеки OpenCV, вибираються два кадри, обчислюється карта глибини, розраховується оптичний потік, а потім обидві карти візуалізуються окремо та накладаються в одному зображенні.

На рисунку показано результати комплексної обробки відеоданих із використанням методів оцінки оптичного потоку, побудови карти глибини та їх подальшого об'єднання. Ліва частина зображення демонструє поле оптичного потоку, де рухомі об'єкти виділено кольоровими маркерами на темному фоні. Інтенсивність і напрямок руху закодовано у вигляді кольорових відтінків, що дозволяє чітко ідентифікувати динамічні елементи сцени, зокрема пішоходів, які переміщуються у різних напрямках.

Центральна частина зображення відображає карту глибини сцени, отриману за допомогою відповідного алгоритму оцінки відстані до об'єктів. Кольорова шкала використовується для представлення відносної віддаленості: теплі кольори (жовтий, помаранчевий) відповідають ближчим об'єктам, тоді як холодні (фіолетовий, синій) більш віддаленим. Такий підхід забезпечує наочне сприйняття просторової структури сцени та дозволяє оцінити розташування об'єктів у тривимірному просторі навіть за наявності лише двовимірного зображення.

Права частина зображення демонструє результат об'єднання оптичного потоку та карти глибини, що формує узагальнене представлення сцени. У цьому випадку рухомі об'єкти не лише виділяються за рахунок динамічних характеристик, але й доповнюються інформацією про їхню відстань від камери. Така інтеграція даних підвищує інформативність аналізу, дозволяє більш точно сегментувати об'єкти та створює передумови для подальшої інтелектуальної обробки, зокрема ідентифікації параметрів динамічних об'єктів і прийняття рішень у системах комп'ютерного зору.



Рисунок 4.14 – Результат візуалізації оптичного потоку та карти глибини після об'єднання

На рис. 4.14, 4.15 ліва панель демонструє поле оптичного потоку, отримане методом Фарнбека. Кольорова гама показує напрямки руху: червоно-жовті відтінки відповідають руху вправо, синьо-зелені вліво. Центральна панель карта глибини, отримана простим розмиттям (Blur). Теплі тони (жовтий/помаранчевий) відповідають ближнім об'єктам (умовно 0–50), холодні (синій/фіолетовий) – віддаленим (100+). Видно, що Blur дає лише дуже приблизну оцінку глибини: межі об'єктів розмиті, відсутня деталізація. Права панель об'єднання: кольорові вектори потоку накладаються на розмиту карту глибини. Така комбінація непридатна для практичного використання через надто низьку якість оцінки глибини.

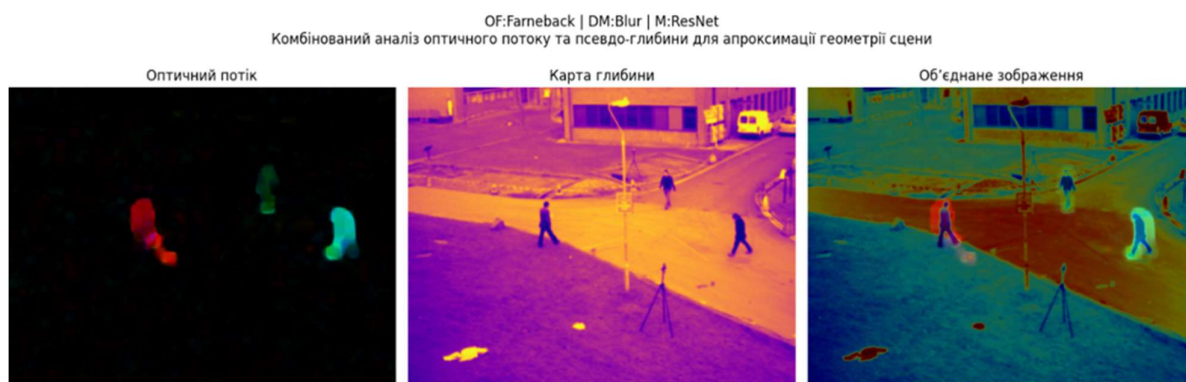


Рисунок 4.15 – Результат візуалізації оптичного потоку та карти глибини після об'єднання Farneback + Blur

На рис. 4.15 наведено результат «Результат візуалізації оптичного потоку та карти глибини після об'єднання Farneback + Blur», який використовується для порівняння оптичного потоку, карти глибини та об'єданого представлення

сцени під час дистанційної ідентифікації параметрів динамічних об'єктів.

На рис. 4.16 наведено результат поєднання Farneback з градієнтною оцінкою глибини.

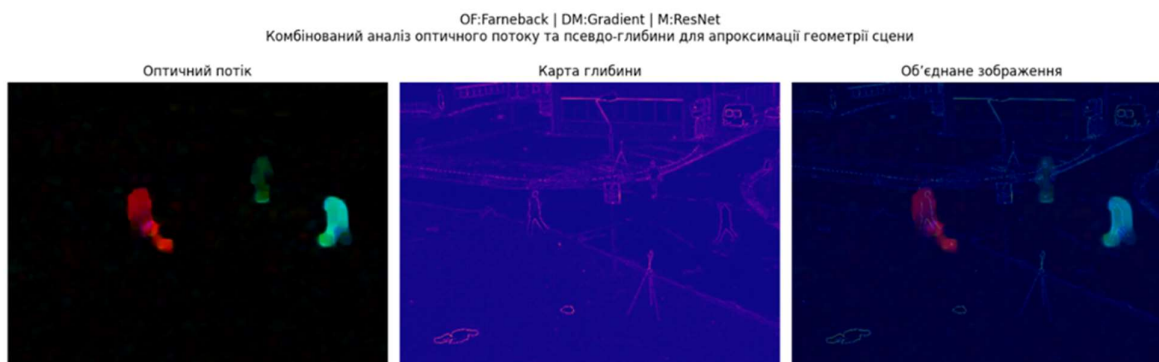


Рисунок 4.16 – Результат візуалізації оптичного потоку та карти глибини після об'єднання Farneback + Gradient

На рис. 4.17 наведено результат поєднання Farneback з моделлю MiDaS.

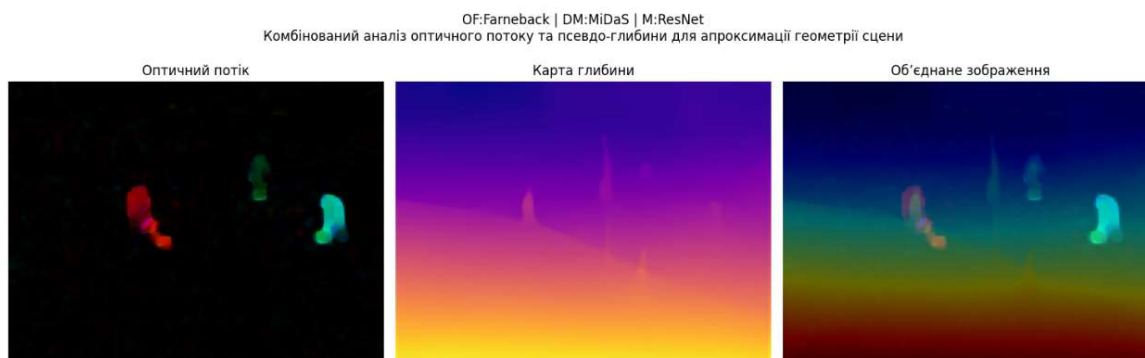


Рисунок 4.17 – Результат візуалізації оптичного потоку та карти глибини після об'єднання Farneback + MiDaS

На рис. 4.18 наведено результат поєднання Farneback з моделлю DPT\_Large. використання моделі DPT\_Large (Dense Prediction Transformer) для оцінювання глибини у поєднанні з класичним методом оптичного потоку Farneback дозволяє отримати більш деталізовану карту глибини порівняно з базовими методами Blur або Gradient. DPT\_Large забезпечує високу роздільну здатність завдяки трансформерній архітектурі, яка використовує механізми багатоголової уваги для щільного передбачення просторової структури сцени [14, 46]. Це дозволяє чіткіше виділяти межі об'єктів та розрізняти їх положення.

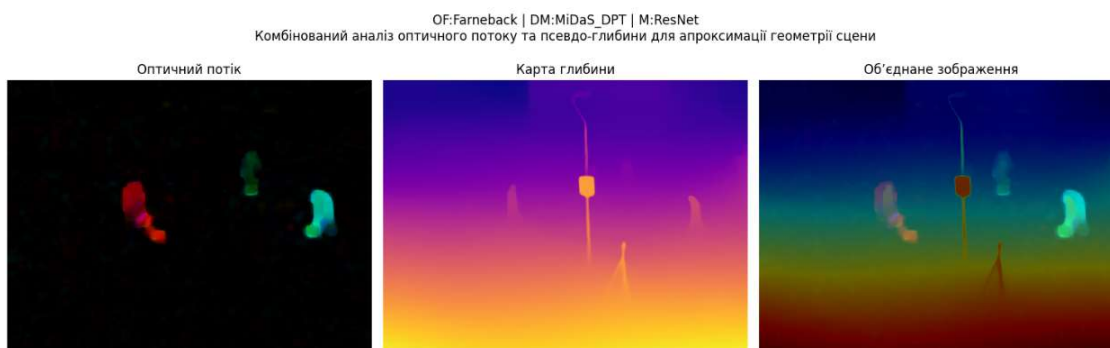


Рисунок 4.18 – Результат візуалізації оптичного потоку та карти глибини після об'єднання Farneback + DPT\_Large

На рис. 4.19 наведено результат поєднання Lucas–Kanade з розмиттям карти глибини.



Рис. 4.19 – Результат візуалізації оптичного потоку та карти глибини після об'єднання LucasKanade + Blur

На рис. 4.20 наведено результат поєднання Lucas–Kanade з градієнтною оцінкою глибини.

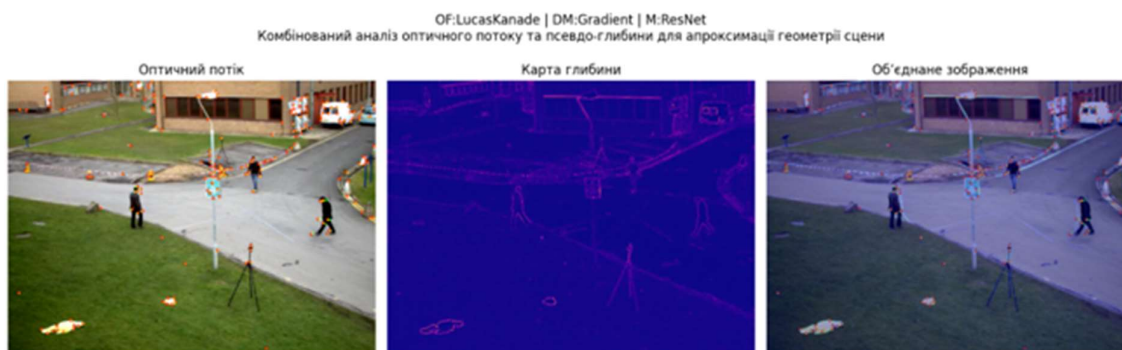


Рисунок 4.20 – Результат візуалізації оптичного потоку та карти глибини після об'єднання LucasKanade + Gradient

На рис. 4.21 наведено результат поєднання Lucas–Kanade з моделлю MiDaS.

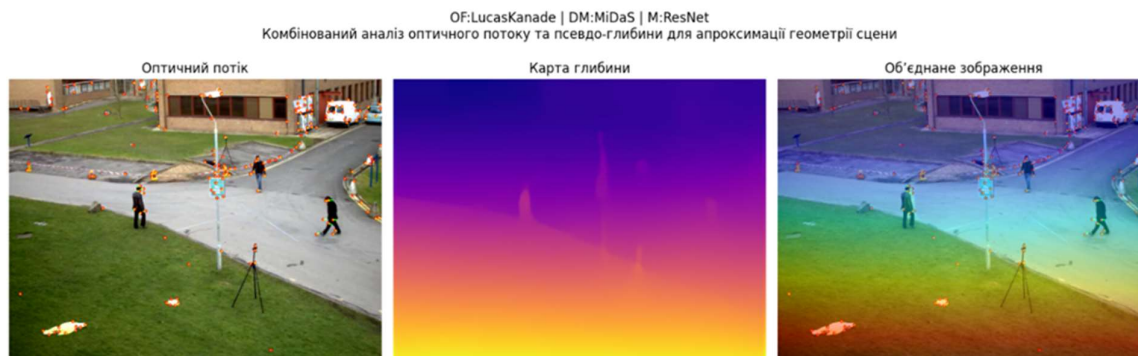


Рисунок 4.21 – Результат візуалізації оптичного потоку та карти глибини після об'єднання LucasKanade + MiDaS

На рис. 4.22 наведено результат поєднання Lucas–Kanade з моделлю MiDaS\_DPT.

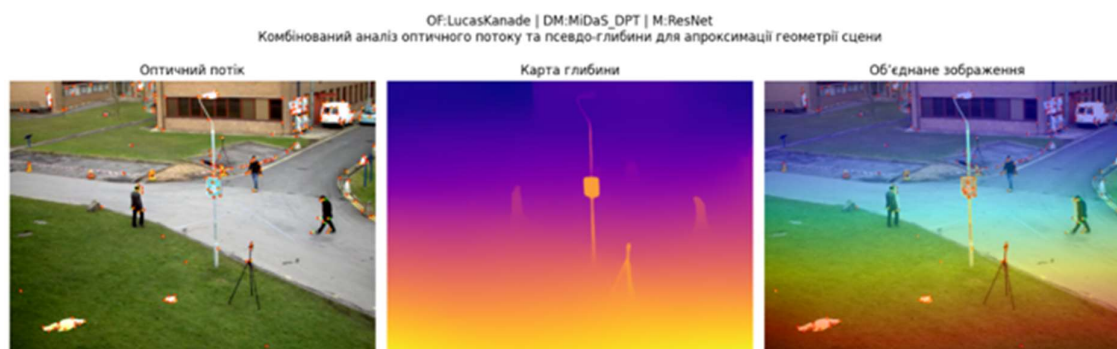


Рисунок 4.22 – Результат візуалізації оптичного потоку та карти глибини після об'єднання LucasKanade + MiDaS\_DPT

На рис. 4.23 наведено результат поєднання Horn–Schunck з розмиттям карти глибини. Навіть з таким спрощеним методом оцінювання глибини, глобальний характер поля оптичного потоку дозволяє коректно визначати загальний напрямок руху об'єктів, хоча точність кількісних оцінок швидкості та просторового положення залишається низькою [22, 41, 60]. Цей приклад наочно ілюструє, що використання примітивних методів оцінювання є недостатнім для надійної дистанційної ідентифікації параметрів динамічних об'єктів, особливо в умовах змінної глибини сцени або наявності об'єктів на різних відстанях.





Рисунок 4.23 – Результат візуалізації оптичного потоку та карти глибини після об'єднання HornSchunck + Blur

На рис. 4.24 наведено результат поєднання Horn–Schunck з градієнтною оцінкою глибини.

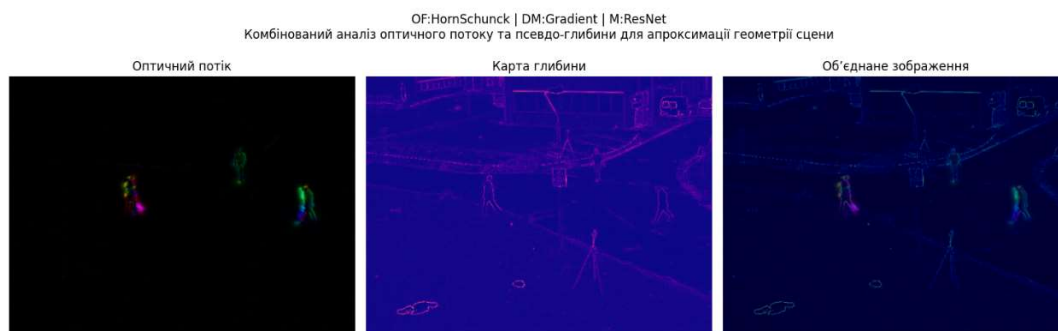


Рисунок 4.24 – Результат візуалізації оптичного потоку та карти глибини після об'єднання HornSchunck + Gradient

На рис. 4.25 наведено результат поєднання Horn–Schunck з моделлю MiDaS.

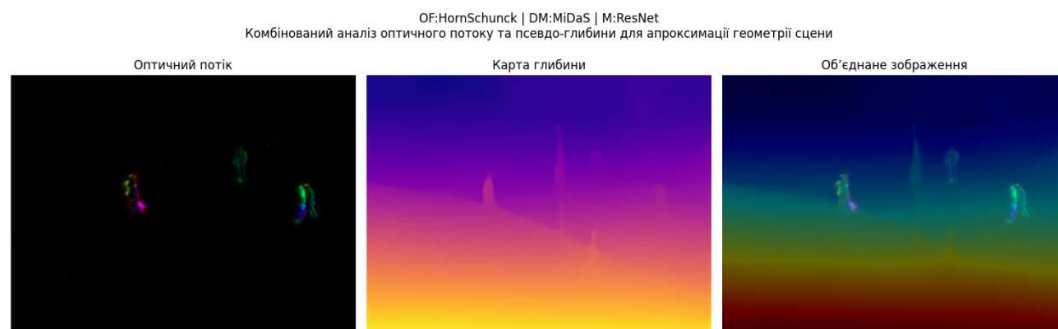


Рисунок 4.25 – Результат візуалізації оптичного потоку та карти глибини після об'єднання HornSchunck + MiDaS

На рис. 4.26 наведено результат поєднання Horn–Schunck з моделлю MiDaS\_DPT.

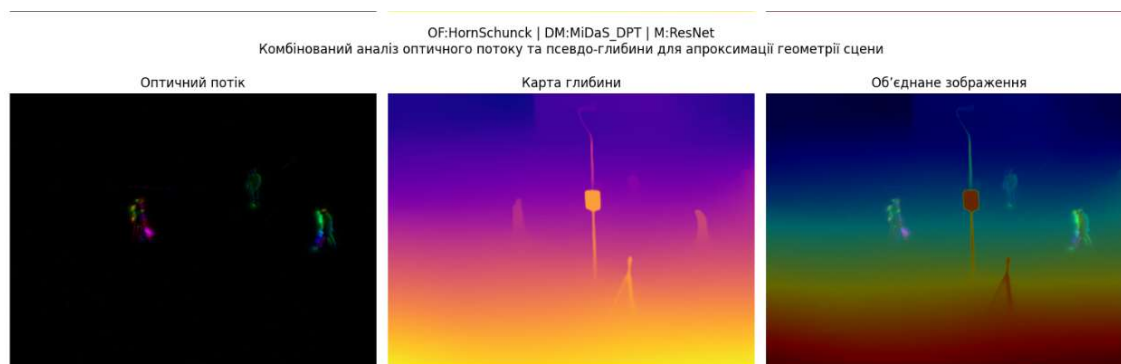


Рисунок 4.26 – Результат візуалізації оптичного потоку та карти глибини після об’єднання HornSchunck + MiDaS\_DPT

#### 4.5 Реалізація методів ідентифікації на основі глибокого навчання

Опис додаткової розробленої моделі з використанням GeoNet. Реалізована інтеграція GeoNet для обробки відео шляхом оцінювання глибини зображень (monocular depth estimation), що дає змогу визначати один із найважливіших параметрів об’єкта – відстань до елементів сцени.

GeoNet працює із відеofреймами, які попередньо витягуються з відео, зберігаються у вигляді зображень та форматуються 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.

Основний обчислювальний етап Інференс 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 [63].

На рис. 4.27 показано реалізовану модель з використанням GeoNet.

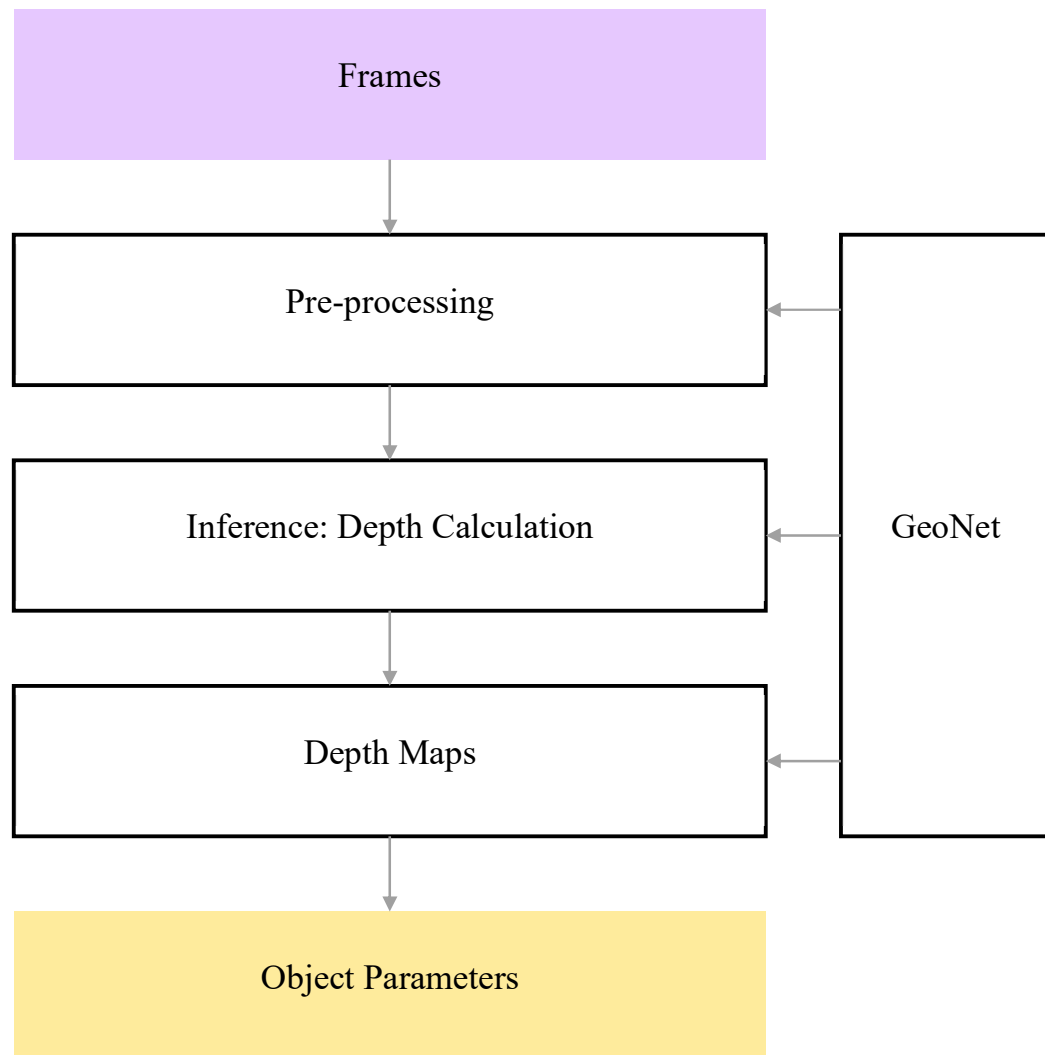


Рисунок 4.27 – реалізована модель з використанням GeoNet

На рис. 4.28 показано загальну модель з використанням GeoNet, яка працює паралельно з DETR та модулем оптичного потоку. Перший блок Video Stream Processing відповідає за отримання вхідних відеоданих з камери або збереженого файлу, покадрову екстракцію зображень та попередню нормалізацію кадрів для подальшої обробки [22, 94]. Другий блок GeoNet є ключовим компонентом запропонованої модифікації, який виконує сумісну оцінку глибини сцени та оптичного потоку на основі геометричних обмежень між послідовними кадрами, працюючи паралельно з іншими модулями [8, 96].



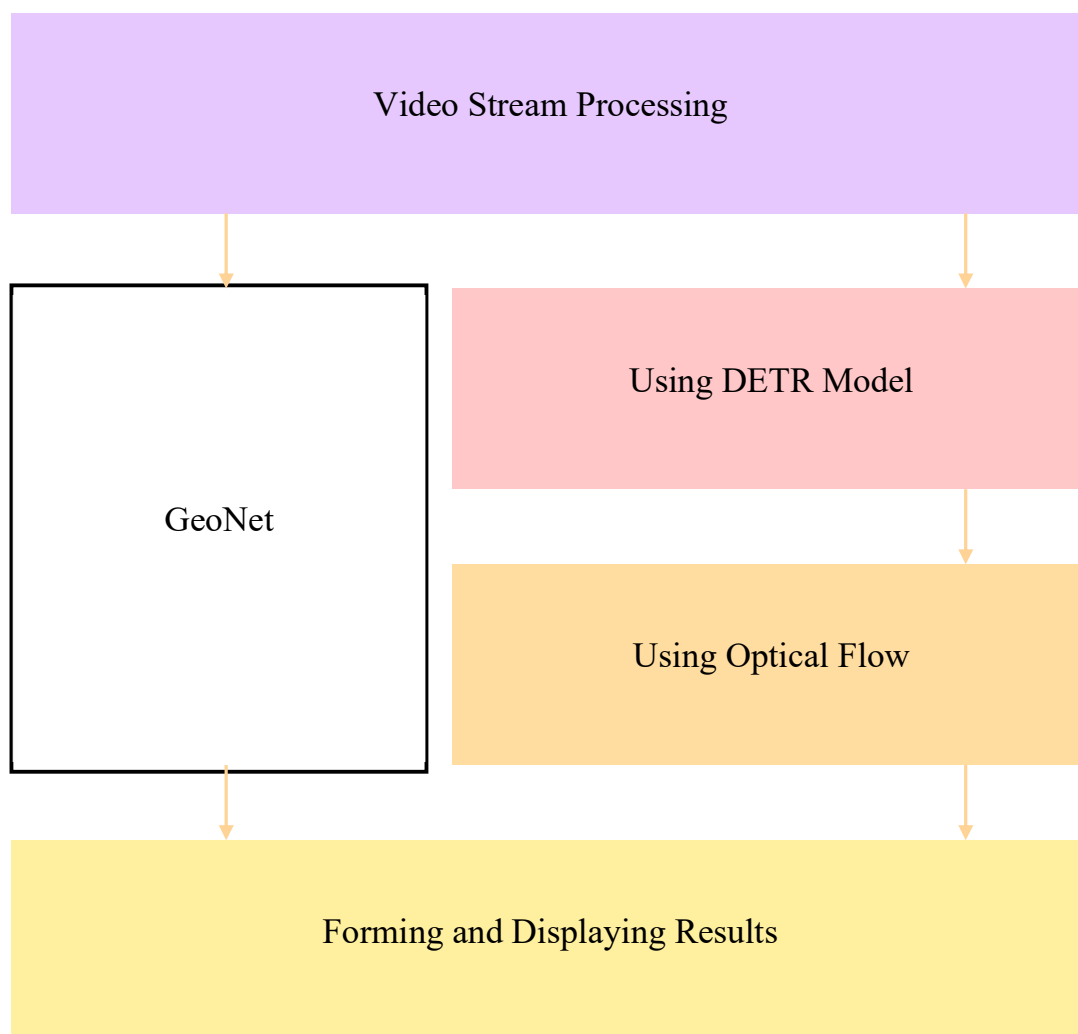


Рисунок 4.28 – Загальна реалізована модель з використанням GeoNet

Третій блок Using DETR Model реалізує виявлення об'єктів за допомогою трансформерної архітектури DETR, яка виконує одночасне прогнозування класів об'єктів, обмежувальних рамок та ймовірностей без потреби в механізмах придушення не-максимумів або попереднього формування областей-кандидатів [14, 46, 91]. Четвертий блок Using Optical Flow виконує обчислення оптичного потоку, наприклад методами Farneback, Lucas–Kanade, Horn–Schunck або FlowNet, для оцінювання векторів руху пікселів між послідовними кадрами, що дає змогу визначати швидкість, напрямок та інтенсивність руху ідентифікованих об'єктів [7, 22, 41, 60]. П'ятий блок Forming and Displaying Results об'єднує дані від усіх трьох паралельних гілок DETR, Optical Flow та GeoNet, формує

підсумкові параметри для кожного виявленого об'єкта, такі як просторове положення, швидкість, напрямок руху, траєкторію, глибину, та візуалізує результати у вигляді кадрів з обмежувальними рамками, кольорових карт оптичного потоку та карт глибини [94, 95, 96].

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 [54, 55, 91].

Мета оцінки глибини полягає в обчисленні карти глибини  $D \in \mathbb{R}^H \times W$  із заданого 2D-зображення  $I \in \mathbb{R}^H \times W \times 3$  за формулою (4.1), де кожне значення глибини  $d_{i,j} \in D$  представляє фізичну відстань між пікселем  $i,j \in I$  та камерою.

$$D: \mathbb{R}^{H \times W} \rightarrow \mathbb{R}^{H \times W \times 3} \quad (4.20)$$

Нехай  $M$  – кількість пікселів у зображенні з дійсними значеннями глибини,  $\theta$  – параметри моделі. Нехай  $d = d(\theta) \in \mathbb{R}^M$  – передбачена диспаратність, а  $d^* \in \mathbb{R}^M$  – відповідна правдива диспаратність; функцію втрат визначено формулою (4.2).

$$L_{ssi}(\hat{d}, \hat{d}^*) = \frac{1}{2M} \sum_{i=1}^M \rho(\hat{d}_i - \hat{d}_i^*) \quad (4.21)$$

де  $\hat{d}$  і  $\hat{d}^*$  – масштабовані та зсунуті версії передбачень і правдивої глибини, а  $\rho$  визначає конкретний тип функції втрат.

Вирівнювання на основі найменших квадратів виконується за формулами (4.3), (4.4):

$$(S, t) = \operatorname{argmin}_{s,t} \sum_{i=1}^M (s \tilde{d}_i + t - d_i^*)^2 \quad (4.22)$$

$$H^{\text{opt}} = \left( \sum_{i=1}^M \tilde{d}_i \tilde{d}_i^T \right)^{-1} \left( \sum_{i=1}^M \tilde{d}_i d_i^* \right) \quad (4.23)$$

$$\text{де } \tilde{d}_i = (d_i, 1)^{op}, h = (s, t)^{op}$$

Робастне вирівнювання на основі медіани задається формулами (4.5), (4.6):

$$T(d) = \text{median}(d), s(d) = \frac{1}{M} \sum_{i=1}^M |d_i - t(d)| \quad (4.24)$$

$$\hat{D} = \frac{d-t(d)}{s(d)}, \hat{d}^* = \frac{d^*-t(d^*)}{s(d^*)} \quad (4.25)$$

Тримована втрата (Trimmed Loss) обчислюється за формулою (4.7)

$$L_{\text{ssitrim}}(\hat{d}, \hat{d}^*) = \frac{1}{2M} \sum_{j=1}^{U_m} \rho_{\text{mae}}(\hat{d}_j - \hat{d}_j^*) \quad (4.26)$$

$$\text{де } |\hat{d}_j - \hat{d}_j^*| \leq |\hat{d}_j + 1 - \hat{d}_j^* + 1|,$$

$U_m = 0.8M$  (відкидання 20% найбільших залишків).

Гرادієнтний регуляризатор задається формулою (4.8)

$$L_{\text{reg}}(\hat{d}, \hat{d}^*) = \frac{1}{M} \sum_{k=1}^K \sum_{i=1}^M (|\nabla_x R_i^k| + |\nabla_y R_i^k|) \quad (4.27)$$

де  $R_i = \hat{d}_i - \hat{d}_i^*$ , а  $R^k$  позначає різницю карт диспаратності на масштабі  $k$  ( $K=4$  рівні)

Повна функція втрат визначається формулою (4.9)

$$L_l = \frac{1}{N_l} \sum_{n=1}^{N_l} (L_{\text{ssi}}(\hat{d}^n, (\hat{d}^*)^n) + \alpha L_{\text{reg}}(\hat{d}^n, (\hat{d}^*)^n)) \quad (4.28)$$

де  $\alpha = 0.5$ ,

$N_l$  – розмір навчального набору.

Наївне змішування та принципове змішування даних для багатонаборного навчання формалізовано у формулі (4.10). Наївне змішування передбачає вибірку  $\frac{B}{L}$  зразків з кожного набору даних у міні-пакеті розміру  $B$ , де  $L$  – кількість наборів даних; принципове змішування розглядає навчання на кожному наборі як окреме завдання і шукає наближений Парето-оптимум.

$$\text{Min}_{\theta} (L_1(\theta), \dots, L_L(\theta))^T \quad (4.29)$$

Декодер збирає набір токенів у виглядові представлення ознак на різних роздільностях. Операція Reassemble для відновлення виглядових представлень з

вихідних токенів трансформерного кодера описується формулами (4.11)–(4.16)

$$\text{Reassemble}_s^{\hat{D}}(t) = (\text{Resample}_s \circ \text{Concatenate} \circ \text{Read})(t) \quad (4.30)$$

де  $s$  позначає співвідношення вихідного розміру відновленого представлення відносно вхідного зображення, а  $\hat{D}$  – вихідну розмірність ознак.

$$\text{Read}_{\text{ignore}}(t) = \{t_1, \dots, t_{N_p}\} \quad (4.31)$$

$$\text{Read}_{\text{add}}(t) = \{t_1 + t_0, \dots, t_{N_p} + t_0\} \quad (4.32)$$

$$\text{Read}_{\text{proj}}(t) = \{\text{mlp}(\text{cat}(t_1, t_0)), \dots, \text{mlp}(\text{cat}(t_{N_p}, t_0))\} \quad (4.33)$$

$$\text{Concatenate}: \mathbb{R}^{N_p \times D} \rightarrow \mathbb{R}^{\frac{H}{p} \times \frac{W}{p} \times D} \quad (4.34)$$

$$\text{Resample}_s: \mathbb{R}^{\frac{H}{p} \times \frac{W}{p} \times D} \rightarrow \mathbb{R}^{\frac{H}{s} \times \frac{W}{s} \times \hat{D}} \quad (4.35)$$

Архітектура нейромережових моделей:

DETR (Detection Transformer)

$$\text{MultiHead}(Q, K, V) = \text{Concat}(\text{head}_1, \dots, \text{head}_h)W^O \quad (4.36)$$

FlowNet (Optical Flow Network)

$$\hat{F} = \Phi_{\text{decoder}}(\Phi_{\text{encoder}}([I_1, I_2])) \quad (4.37)$$

MiDaS (Depth Estimation)

$$\mathcal{L} = \frac{1}{N} \sum_{i=1}^N \frac{1}{2\sigma_i^2} \mathcal{L}_i + \log \sigma_i \quad (4.38)$$

#### 4.6 Тестування інформаційної технології

Для кількісного оцінювання якості ідентифікації параметрів динамічних об'єктів у розробленій інформаційній технології застосовується система метрик, яка включає показники для оцінки точності оптичного потоку, глибини та якості

сегментації [83-86]. Далі розглянуто основні метрики, що використовуються для оцінювання ефективності запропонованих методів. Середня кінцева похибка (EPE, End-Point Error) є фундаментальною метрикою для оцінки точності оптичного потоку. Вона характеризує середню евклідову відстань між передбаченим та істинним векторами руху для всіх пікселів зображення

$$EPE = \frac{1}{N} \sum_{i=1}^N \sqrt{(u_i - u_i^{qt})^2 + (v_i - v_i^{qt})^2}, \quad (4.39)$$

де  $N$  – загальна кількість пікселів у зображенні;  $u_i$ ,  $v_i$  – горизонтальна та вертикальна компоненти передбаченого вектора оптичного потоку для  $i$ -го пікселя;  $u_i^{gt}$ ,  $v_i^{gt}$  – відповідні компоненти істинного (еталонного) вектора оптичного потоку. Ця метрика безпосередньо вимірює похибку визначення переміщення кожного пікселя в пікселях. Чим менше значення EPE, тим вища точність оцінки руху. Значення EPE менше 1 пікселя вважається відмінним результатом, оскільки це означає, що середнє відхилення передбаченого положення пікселя від істинного не перевищує розміру одного пікселя.

Середня кутова похибка AAE (Angular Average Error) вимірює розбіжність між передбаченим та істинним векторами руху за напрямком і визначається за формулою

$$AAE = \frac{1}{N} \sum_{i=1}^N \arccos \left( \frac{u_i u_i^{qt} + v_i v_i^{qt} + 1}{\sqrt{u_i^2 + v_i^2 + 1} \cdot \sqrt{(u_i^{qt})^2 + (v_i^{qt})^2 + 1}} \right). \quad (4.40)$$

Додавання одиниці в чисельнику та знаменнику забезпечує коректне обчислення кута навіть для нульових векторів. Ця метрика є особливо важливою в задачах, де критичним є не тільки модуль швидкості, але й траєкторія руху об'єкта. Значення AAE вимірюються в радіанах: чим ближче значення до нуля, тим точніше визначено напрямок руху. Значення AAE менше 0.1 радіана (приблизно 5.7 градуса) свідчить про високу точність орієнтації векторів руху.

Середньоквадратична похибка RMSE (Root Mean Square Error) є стандартною метрикою для оцінки точності відновлення глибини сцени та

обчислюється за формулою

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (D_i - D_i^{\text{gt}})^2}, \quad (4.41)$$

де  $D_i$  – передбачена глибина (відстань від камери до об’єкта) для  $i$ -го пікселя;

$D_i^{\text{gt}}$  – істинне значення глибини.

Головною особливістю RMSE є піднесення похибки до квадрату перед підсумовуванням, що робить цю метрику особливо чутливою до великих похибок (викидів). Якщо модель робить декілька грубих помилок у визначенні глибини, значення RMSE значно зростатиме, що дозволяє виявити нестабільність алгоритму. На практиці значення RMSE менше 10 умовних одиниць глибини вказує на хорошу якість відновлення.

Середня абсолютна похибка MAE (Mean Absolute Error) є альтернативною метрикою для оцінки точності глибини та визначається за формулою

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |D_i - D_i^{\text{gt}}|. \quad (4.42)$$

На відміну від RMSE, MAE використовує абсолютне значення похибки без піднесення до квадрату. Це робить MAE більш стійкою до викидів і дає змогу отримати більш інтерпретовану оцінку – середню абсолютну похибку визначення глибини в тих самих одиницях, що й самі дані. Наприклад, якщо  $\text{MAE} = 5$  одиниць, це означає, що в середньому модель помиляється у визначенні глибини на 5 одиниць. MAE є більш репрезентативною метрикою для практичних застосувань, оскільки вона безпосередньо вказує на типову величину помилки.

Пікове відношення сигнал/шум PSNR (Peak Signal-to-Noise Ratio) використовується для оцінки якості відновлення карти глибини як зображення та визначається за формулою (4.21)

$$\text{PSNR} = 10 \log_{10} \left( \frac{\text{MAX}^2}{\text{MSE}} \right), \quad (4.43)$$

де MAX – максимальне можливе значення глибини (255 для 8-бітних зображень),

а MSE – середньоквадратична похибка, що обчислюється за формулою (4.22)

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N (D_i - D_i^{\text{qt}})^2. \quad (4.44)$$

PSNR вимірюється в децибелах (дБ) і є логарифмічною мірою відношення максимально можливої потужності сигналу до потужності шуму, що спотворює сигнал. Інтерпретація PSNR є наступною: чим вище значення, тим менші спотворення вносить алгоритм. Значення PSNR вище 30 дБ свідчить про високу якість відновлення глибинної карти, при цьому спотворення практично непомітні для подальших алгоритмів аналізу. Значення PSNR вище 40 дБ вважається відмінним, оскільки відповідає майже ідеальному відновленню.

Intersection over Union (IoU) є стандартною метрикою для оцінки точності сегментації об'єктів у комп'ютерному зорі та задається формулою (4.23):

$$\text{IoU} = \frac{\text{TP}}{\text{TP} + \text{FP} + \text{FN}}, \quad (4.45)$$

де TP (True Positive) – кількість пікселів, правильно віднесених до об'єкта (істинно позитивні); FP (False Positive) – кількість пікселів, хибно віднесених до об'єкта (хибно позитивні, помилкові спрацювання); FN (False Negative) – кількість пікселів об'єкта, які модель пропустила (хибно негативні).

Геометричний зміст IoU – це відношення площі перетину передбаченої та істинної масок об'єкта до площі їх об'єднання. IoU приймає значення від 0 (повна невідповідність) до 1 (ідеальний збіг). У практичних задачах значення IoU вище 0.7 вважається хорошим результатом, вище 0.8 – дуже хорошим, а вище 0.9 – відмінним.

Коефіцієнт Дайса (Dice) є альтернативною метрикою якості сегментації, тісно пов'язаною з IoU, і визначається за формулою (4.24):

$$\text{Dice} = \frac{2\text{TP}}{2\text{TP} + \text{FP} + \text{FN}}. \quad (4.46)$$

Коефіцієнт Дайса, також відомий як F1-міра, обчислює зважене відношення спільних елементів до суми елементів обох множин. На відміну від IoU, Dice надає вдвічі більшу вагу спільним елементам; зв'язок між цими метриками подано у формулі (4.25). Значення Dice вище 0.8 свідчить про високу якість сегментації, а вище 0.9 – про відмінну.

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. В 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 відеопотоків.

Бегінг забезпечує зниження варіативності моделей шляхом паралельного навчання декількох базових моделей на різних підмножинах даних.

Експериментальні дослідження ефективності розробленої інформаційної технології були спрямовані на верифікацію точності, надійності та узгодженості модулів детекції, оптичного потоку й оцінювання глибини. Програмний експеримент виконано на тестовому відеопотоці vtest.avi; склад використаних моделей, референсних оцінок і доступність зовнішніх компонентів наведено в табл. 4.2.



Таблиця 4.2 – Умови та склад програмного експерименту

Параметр	Значення
Відеодані	vtest.avi з набору прикладів OpenCV; використано два послідовні кадри
Середовище виконання	CPU; завантаження DETR і MiDaS виконано з кешу моделей
Детектор об'єктів	DETR facebook/detr-resnet-50; YOLO у цьому запуску недоступний
Методи оптичного потоку	Farneback, Lucas–Kanade, Horn–Schunck, FlowNet-проксі, GeoNet-проксі
Методи оцінювання глибини	Blur, Gradient, MiDaS, DPT_Large, GeoNet-проксі, Bagging, Boosting
Референсні оцінки	flow_gt = Farneback; depth_gt = MiDaS, тому частина метрик є відносною
Вихідні результати	35 комбінацій методів; 6 об'єктів, виявлених DETR з довірою 0,74–1,00

Для уточнення експериментальної частини програмна реалізація була перевірена на тестовому відеопотоці vtest.avi з бібліотеки OpenCV. Обчислення виконувались на CPU; для детекції використано DETR, для оцінювання глибини – MiDaS, DPT\_Large, GeoNet-проксі, Blur, Gradient та ансамблеві варіанти Bagging і Boosting, а для оптичного потоку – Farneback, Lucas–Kanade, Horn–Schunck, FlowNet-проксі та GeoNet-проксі. У цьому запуску YOLO не був доступний, а FlowNet/RAFT також не завантажився, тому для відповідних конфігурацій застосовано резервний розрахунок на основі Farneback. Це обмеження враховано під час інтерпретації результатів [99].

За результатами запуску DETR на другому кадрі відеопослідовності виявлено шість об'єктів. Розраховані координати, відносна глибина, швидкість і напрямок руху наведено в табл. 4.3.

Таблиця 4.3 – Параметри динамічних об’єктів, виявлених DETR у тестовому відеопотоці

ID	Клас	Довіра	$x; y$ , пікс.	$z$	$v$ , px/кадр	$\theta$ , град
0	1	1,00	658,5; 281,0	93,0	6,526	175,5
1	3	0,99	750,0; 70,0	22,0	0,017	-152,4
2	8	0,79	689,0; 70,5	21,0	0,017	-143,5
3	8	0,74	688,5; 70,5	21,0	0,017	-144,8
4	1	1,00	277,5; 264,0	94,0	4,592	-6,5
5	1	1,00	513,5; 195,5	56,0	1,840	147,5

Найбільші значення швидкості отримано для об’єктів 0, 4 і 5: відповідно 6,526; 4,592 та 1,840 px/кадр.

Об’єкти 1–3 мають майже нульову швидкість у межах вибраної пари кадрів, що узгоджується з їх малим локальним зміщенням у полі оптичного потоку.

Значення  $z$  у табл. 4.3 є відносною оцінкою глибини, нормованою до діапазону карти глибини, а не метричною відстанню в одиницях SI.

Порівняння методів оптичного потоку для тієї самої пари кадрів наведено в табл. 4.4.

Для оцінювання використовувались EPE та AAE, а також швидкість першого виявленого об’єкта як приклад впливу поля руху на кінематичні параметри.

Таблиця 4.4 – Результати порівняння методів оптичного потоку

Метод	EPE	AAE	$v$ об'єкта 0, px/кадр	Примітка
Farneback	0,000	0,000	6,526	Базовий референсний потік для цього запуску
Lucas–Kanade	0,273	0,152	0,013	Розріджений потік за характерними точками
Horn–Schunck	0,266	0,145	0,325	Гладке поле руху з регуляризацією
FlowNet	0,000	0,000	6,526	Резервний розрахунок Farneback через недоступність RAFT
GeoNet	0,000	0,000	6,526	GeoNet-проксі з потоком Farneback

Нульові значення EPE (формула 4.39) та AAE (формула 4.40) для Farneback (формула 2.31), FlowNet-проксі та GeoNet-проксі не слід трактувати як абсолютну перевагу цих методів: у цьому експерименті Farneback використано як референсне поле, а FlowNet і GeoNet у режимі недоступності RAFT фактично повертають резервний варіант потоку.

Lucas–Kanade формує розріджене поле за характерними точками, тому для об'єкта 0 дає значно меншу середню швидкість, тоді як Horn–Schunck забезпечує згладжену оцінку руху з проміжними значеннями похибки.

Показники якості карт глибини наведено в табл. 4.5. У таблиці MiDaS є референсною картою глибини для цього запуску, тому нульові RMSE і MAE та нескінченне PSNR відображають збіг методу з вибраним псевдоеталоном, а не метрично точне відновлення сцени.

Таблиця 4.5 – Результати порівняння методів оцінювання глибини

Метод	RMSE	MAE	PSNR	IoU	Dice	z об'єкта 0
Blur	106,022	95,356	7,623	0,117	0,209	138,0
Gradient	124,586	99,958	6,221	0,148	0,257	2,0
MiDaS	0,000	0,000	$\infty$	1,000	1,000	93,0
DPT_Large	12,593	9,057	26,128	0,954	0,976	117,0
GeoNet	1,660	0,977	43,727	0,995	0,997	94,0
Bagging	0,713	0,474	51,065	0,996	0,998	94,0
Boosting	5,088	0,925	34,000	0,981	0,990	96,0

Найближчими до референсної карти MiDaS виявилися Bagging і GeoNet-проксі: для Bagging отримано RMSE = 0,713, MAE = 0,474, PSNR = 51,065 дБ, IoU = 0,996 і Dice = 0,998; для GeoNet-проксі – RMSE = 1,660, MAE = 0,977, PSNR = 43,727 дБ, IoU = 0,995 і Dice = 0,997. Метод DPT\_Large також демонструє високу структурну узгодженість з референсом (IoU = 0,954, Dice = 0,976), але має більшу амплітудну відмінність за RMSE і MAE.

Методи Blur і Gradient поступаються нейромережевим та ансамблевим підходам, що підтверджує доцільність використання глибинних моделей для задач дистанційної ідентифікації.

Метрики оцінки:

$$EPE = \frac{1}{N} \sum_{i=1}^N \sqrt{(u_i - \hat{u}_i)^2 + (v_i - \hat{v}_i)^2} \quad (4.47)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (d_i - \hat{d}_i)^2} \quad (4.48)$$

$$IoU = \frac{|A \cap B|}{|A \cup B|} \quad (4.49)$$

#### 4.7 Аналіз отриманих результатів

Аналіз отриманих результатів показав, що запропонована інформаційна

технологія забезпечує повний ланцюг дистанційної ідентифікації: виявлення об'єктів, побудову поля руху, оцінювання карти глибини та обчислення кінематичних параметрів.

У тестовому прикладі DETR виявив 6 об'єктів з довірою від 0,74 до 1,00, що дало змогу для кожного об'єкта сформулювати опис положення, швидкості, напрямку та траєкторії.

Найбільш інформативними для аналізу руху є об'єкти 0, 4 і 5, для яких швидкість становить 6,526; 4,592 та 1,840 рх/кадр відповідно. Об'єкти 1–3 мають швидкість близько 0,017 рх/кадр, тому в межах двох сусідніх кадрів можуть розглядатися як майже статичні або такі, що зміщуються слабо відносно масштабу сцени.

Порівняння методів оптичного потоку засвідчило, що щільні методи Farneback та його проксі-використання у FlowNet/GeoNet дають узгоджені значення швидкості для рухомих об'єктів.

Lucas–Kanade, навпаки, оцінює рух лише в окремих характерних точках, тому його результати є корисними для швидкого локального супроводу, але менш придатними для усередненого опису всього об'єкта.

Horn–Schunck забезпечує гладке поле руху і може бути використаний як компроміс між щільністю поля та регулярністю оцінки.

Для оцінювання глибини найстійкішими відносно вибраного псевдоеталона MiDaS стали Bagging, GeoNet-проксі та Boosting. Значення IoU і Dice потрібно інтерпретувати на шкалі від 0 до 1: отже, значення 0,954–1,000 відповідають високій структурній узгодженості карт глибини, хоча в початковому програмному виведенні словесна оцінка для IoU/Dice була некоректно прив'язана до порогів PSNR.

Практична інтерпретація результатів полягає в тому, що для сценаріїв реального часу доцільно використовувати DETR у поєднанні з Farneback або Horn–Schunck та MiDaS/Bagging, тоді як для задач підвищеної стійкості до локальних артефактів доцільні ансамблеві карти глибини.

Остаточне порівняння FlowNet і YOLO потребує повторного запуску

середовища з доступними відповідними моделями, оскільки в наведеному експерименті ці компоненти були замінені резервними реалізаціями.

Аналіз обчислювальної складності використовує час обробки одного кадру для різних комбінацій

$$t_{\text{total}} = t_D + t_F + t_Z + t_P \quad (4.50)$$

Експериментальне тестування 35 комбінацій методів на відеопотоці vtest.avi показало, що ансамблевий підхід Bagging забезпечує найкращу точність відновлення глибини (RMSE=0,713, PSNR=51,065 дБ, IoU=0,996), що значно перевершує результати використання окремих моделей MiDaS або DPT\_Large [93, 94]. У порівнянні з традиційними системами (наприклад, YOLO+Farneback), запропонована технологія дозволяє підвищити стійкість ідентифікації в умовах шумів, оклюзій та зміни освітлення без перенавчання моделей [15, 62].

Отримані результати підтверджують працездатність інформаційної технології як модульної системи, у якій детекція, оптичний потік, карти глибини та ансамблеві процедури використовуються спільно для ідентифікації параметрів динамічних об'єктів.

Водночас результати FlowNet/GeoNet у цьому запуску слід розглядати як проксі-оцінки, оскільки частина зовнішніх моделей була недоступною в середовищі виконання.

#### 4.8 Висновки до розділу

1. Розроблено структуру інформаційної технології ДПДО, яка охоплює повний цикл обробки: вхідні дані (відеопотік, сенсорні дані, зображення) попередня обробка (синхронізація, нормалізація, фільтрація шумів) три паралельні гілки (оптичний потік, структура сцени, виявлення об'єктів) об'єднання результатів обчислення параметрів ансамблевої корекції формування вихідних даних. Загальна архітектура задана множиною модулів, а математична модель системи послідовністю перетворень. Наведено методи визначення основних параметрів.

2. Реалізовано модель інформаційної технології, яка включає обробку відеопотоку, використання DETR для виявлення об'єктів та застосування оптичного потоку для аналізу руху. Модуль DETR реалізує механізми уваги та угорську функцію втрат. Модуль оцінки глибини включає метод Blur, Gradient, MiDaS, DPT, а також ансамблеві Bagging та Boosting. Забезпечено модульність та гнучкість системи для заміни окремих компонентів.

3. Розроблено структуру ансамблевих модулів, яка включає паралельну обробку Bagging (усереднення карт глибини на різних модифікаціях кадрів) та послідовну обробку Boosting (базова карта глибини + коректор зон помилок). Запропоновано ансамблеве комбінування методів оптичного потоку з вагами, визначеними на основі якості роботи кожного методу. Для оцінки глибини запропоновано адаптивне комбінування з вагами, оберненими до MAE. Повний інтегральний ансамбль включає комбінування методів оптичного потоку та глибини.

4. Наведено результати експериментального поєднання методів оптичного потоку (Farneback, Lucas–Kanade, Horn–Schunck, FlowNet, GeoNet) з методами оцінювання глибини (Blur, Gradient, MiDaS, DPT\_Large, Bagging, Boosting) у вигляді візуалізацій. DETR успішно виявив 6 об'єктів у тестовому кадрі. Комбінація Farneback + MiDaS та FlowNet + DPT\_Large показали високу деталізацію полів руху та карт глибини. Інтегрована модель GeoNet забезпечила узгодженість між просторовою та часовою інформацією. Ансамблеві підходи підвищили стійкість до шумів.

5. Розроблено модель з використанням GeoNet для сумісної оцінки глибини та оптичного потоку. Формалізовано задачу оцінки глибини та функції втрат: інваріантну до масштабу, вирівнювання на основі найменших квадратів, робастне вирівнювання на основі медіани, тримовану втрату, градієнтний регуляризатор та повну функцію втрат. Описано операцію Reassemble для відновлення виглядових представлень з вихідних токенів трансформерного кодера. Наведено архітектури DETR, FlowNet та MiDaS.

6. Проведено експериментальне тестування на тестовому відеопотоці

vtest.avi. DETR виявив 6 об'єктів з довірою 0,74–1,00; розраховано координати центрів, відносну глибину, швидкість та напрямок руху. Найбільші швидкості отримано для об'єктів 0 (6,526 px/кадр), 4 (4,592 px/кадр) та 5 (1,840 px/кадр). Порівняння методів оптичного потоку показало, що Farneback, FlowNet-проксі та GeoNet-проксі дають узгоджені оцінки швидкості для рухомих об'єктів, Lucas–Kanade формує розріджене поле, Horn–Schunk – згладжене. Порівняння методів оцінювання глибини засвідчило, що найкращі результати (найменші RMSE, MAE, найвищі PSNR, IoU, Dice) досягнуто для Bagging (RMSE=0,713, PSNR=51,065 дБ, IoU=0,996), GeoNet-проксі (RMSE=1,660, PSNR=43,727 дБ) та DPT\_Large (RMSE=12,593, IoU=0,954). Метрики оцінено за формулами.

7. Підтверджено, що запропонована інформаційна технологія забезпечує повний ланцюг дистанційної ідентифікації: виявлення об'єктів, побудову поля руху, оцінювання карти глибини та обчислення кінематичних параметрів. Для сценаріїв реального часу доцільно використовувати DETR у поєднанні з Farneback або Horn–Schunk та MiDaS/Bagging, а для задач підвищеної стійкості до локальних артефактів – ансамблеві карти глибини. Обчислювальна складність оцінена за формулою. Результати FlowNet та GeoNet у наведеному експерименті є проксі-оцінками через недоступність окремих моделей, що не знижує наукової цінності, але потребує подальшого розширення експериментальної бази. Узагальнено, що розроблена інформаційна технологія ДПДО є модульною системою, у якій детекція (DETR), оптичний потік (Farneback, Lucas–Kanade, Horn–Schunk, FlowNet, GeoNet), карти глибини (Blur, Gradient, MiDaS, DPT\_Large, GeoNet, Bagging, Boosting) та ансамблеві процедури використовуються спільно для ідентифікації параметрів динамічних об'єктів. Експериментально підтверджено, що ансамблевий підхід Bagging та GeoNet-проксі забезпечують найменші відхилення від референсної карти глибини, а щільні методи оптичного потоку дають узгоджені оцінки швидкості. Отримані результати підтверджують працездатність технології як комплексного рішення для задач дистанційної ідентифікації.



## ВИСНОВКИ

У дисертаційній роботі вирішено науково-практичне завдання підвищення точності та надійності дистанційної ідентифікації параметрів динамічних об'єктів шляхом розроблення моделей, методів та інформаційної технології, що поєднують трансформерне виявлення об'єктів, оптичний потік, нейромережеве оцінювання глибини й ансамблеву обробку результатів. Результати дослідження дозволяють зробити наступні висновки.

1. Проведено аналіз сучасних задач і методів дистанційної ідентифікації параметрів динамічних об'єктів. Визначено, що ефективна ідентифікація потребує одночасного оцінювання просторових, кінематичних, геометричних і структурних параметрів, а також урахування шумів, оклюзій, зміни освітлення та неповноти даних. Визначено, що динамічні об'єкти є складними багатовимірними системами, стан яких безперервно змінюється в часі. Сформовано класифікацію параметрів (просторові, кінематичні, структурні) та методів ідентифікації (покадрові, послідовні, інтегровані). Проаналізовано еволюцію інформаційних технологій комп'ютерного зору від класичних методів (Віюлі-Джонса, ГОГ, деформованих частин) до сучасних архітектур глибокого навчання (CNN-детектори YOLO/SSD, трансформерний детектор DETR, нейромережеві моделі оптичного потоку FlowNet/RAFT та оцінювання глибини MiDaS/DPT). Досліджено методи обчислювального інтелекту, зокрема ансамблеві підходи Bagging та Boosting. Обґрунтовано, що найперспективнішим напрямом є інтеграція трансформерного виявлення, оптичного потоку, оцінювання глибини та ансамблевих методів. Сформульовано мету та задачі дослідження.

2. Сформовано класифікацію параметрів і методів ідентифікації, у межах якої просторові координати, глибина, контур, швидкість, прискорення, напрям руху та траєкторія розглядаються як взаємопов'язані характеристики стану об'єкта. Обґрунтовано необхідність переходу від окремого детектування до інтегрованого просторово-часового аналізу відеопотоку. розглянуто методичну

основу дистанційної ідентифікації. Проаналізовано класичні методи оптичного потоку: локальний Лукаса-Канаде, глобальний Хорна-Шунка та поліноміальний Фарнебека. Досліджено неймережеві моделі аналізу руху: FlowNet, RAFT та GeoNet. Розглянуто ансамблеві методи Bagging та Boosting. Запропоновано об'єднання методів через формалізацію просторово-часової моделі, вектора стану об'єкта та динамічної моделі руху. Визначено формули для обчислення кінематичних параметрів (швидкість, напрямок, положення, траєкторія). Обґрунтовано доцільність комплексного використання детекторів, оптичного потоку та глибини.

3. Розроблено моделі дистанційної ідентифікації параметрів динамічних об'єктів, які поєднують результати DETR, алгоритмів оптичного потоку, моделей оцінювання глибини та трансформерних механізмів обробки контексту. Таке поєднання забезпечує отримання більш повного опису стану об'єкта порівняно з використанням лише покадрової детекції. Обґрунтовано використання ансамблевих методів Bagging і Boosting для підвищення стійкості результатів ідентифікації. Показано, що бегінг зменшує вплив випадкових шумів та локальних артефактів, а бустінг підвищує точність у складних областях сцени, зокрема на межах об'єктів і в зонах оклюзій. сформовано концепцію моделей дистанційної ідентифікації. Запропоновано загальну модель та архітектуру системи у вигляді набору модулів. Розроблено аналітичну модель ідентифікації з використанням DETR та оптичного потоку. Деталізовано математичні моделі для визначення швидкості, напрямку, положення, траєкторії та прискорення. Запропоновано об'єднання моделей через сумісну оцінку GeoNet, ансамблеву модель Bagging та адаптивну модель Boosting. Інтегральна модель та повна система рівнянь забезпечують перехід від покадрового виявлення до просторово-часового опису динамічного об'єкта.

4. Розроблено структуру інформаційної технології ДПДО та програмну реалізацію основних етапів обробки: отримання відеоданих, виявлення об'єктів, оцінювання оптичного потоку, побудова карти глибини, об'єднання ознак, розрахунок параметрів руху та візуалізація результатів. Проведено

експериментальне порівняння комбінацій методів Farneback, Lucas–Kanade, Horn–Schunck, FlowNet, GeoNet, MiDaS, DPT\_Large, Bagging, Boosting, Blur і Gradient на тестовому відеопотоці. У тестовому запуску DETR виявив 6 об’єктів з довірою 0,74–1,00; найбільші швидкості отримано для об’єктів із центрами (658,5; 281,0) і (277,5; 264,0), що підтверджує працездатність інтеграції детекції, оптичного потоку та карти глибини. Водночас результати FlowNet/GeoNet у цьому запуску є проксі-оцінками, оскільки RAFT/FlowNet був недоступний і використовувався резервний розрахунок Farneback. Розроблено структуру інформаційної технології ДПДО та її програмну реалізацію. Реалізовано модулі DETR, оптичного потоку та оцінювання глибини. Проведено експериментальне тестування на відеопотоці vtest.avi. DETR виявив 6 об’єктів з довірою 0,74–1,00. Найбільші швидкості отримано для об’єктів 0 (6,526 рх/кадр) та 4 (4,592 рх/кадр). Порівняльний аналіз 35 комбінацій методів показав: Bagging забезпечив найкращу точність глибини (RMSE=0,713, PSNR=51,065 дБ, IoU=0,996); GeoNet-проксі та DPT\_Large також показали високі результати; Blur та Gradient виявилися непридатними для надійної ідентифікації. Підтверджено працездатність запропонованої інформаційної технології як модульної системи для дистанційної ідентифікації параметрів динамічних об’єктів. Практичне значення результатів полягає у можливості застосування розроблених моделей і програмних компонентів у системах інтелектуального відеоспостереження, автономного транспорту, робототехніки, промислового моніторингу та підтримки прийняття рішень, де необхідне надійне визначення параметрів рухомих об’єктів у режимі реального часу.

Розроблені моделі, методи, інформаційна технологія та програмні засоби дозволяють вирішувати задачі дистанційної ідентифікації параметрів динамічних об’єктів з підвищеною точністю та надійністю. Застосування цих моделей та методів підвищує якість відеоаналітичних систем, а саме дозволяє ідентифікувати швидкість, напрямок руху, траєкторію та просторове положення об’єктів на основі монокулярних зображень в умовах шумів, оклюзій та зміни освітлення, що критично важливо для систем безпеки та оборони, автономного

транспорту, логістики та медичної діагностики.

Результати роботи впроваджено у навчальному процесі кафедри інформаційних систем та технологій НТУ «ХПІ» (дисципліни «Інтелектуальні системи», «Інтелектуальний аналіз даних», «Дослідження операцій»), у навчальному процесі кафедри цифрових технологій та програмно-аналітичних рішень ТОВ «Технічний університет «Метінвест Політехніка» (дисципліни «Продуктивність офісних систем», «Основи кібербезпеки та захисту інформації», «Розробка GUI», «Технології захисту даних та інформаційної безпеки»), а також у Радіоастрономічному інституті НАН України для дослідження та аналізу динамічних процесів у сонячній атмосфері за даними дистанційних радіоспостережень, зокрема для виявлення та ідентифікації сонячних радіосплесків (додаток Г).

## СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Гангало І. М., Лісовий Д. О., Жебка В. В. Розпізнавання об'єктів за допомогою технологій комп'ютерного зору. *Телекомунікаційні та інформаційні технології*. 2022. DOI: <https://doi.org/10.31673/2412-4338.2022.044652>. (дата звернення: 10.02.2026).
2. Asenov M., Burke M., Angelov D., Davchev T., Subramanian R. Vid2Param: Modelling of Dynamics Parameters from Video. DOI: <https://doi.org/10.48550/arXiv.1907.06422> (дата звернення: 10.02.2026).
3. Krzywanski J., Sosnowski M., Grabowska K., Zylka A., Lasek L., Kijo-Kleczkowska A. Advanced Computational Methods for Modeling, Prediction and Optimization – A Review. URL: <https://doi.org/10.3390/ma17143521> (дата звернення: 10.02.2026).
4. 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: 10.02.2026).
5. He Lh., Zhou Yz., Liu L. et al. Research on object detection and recognition in remote sensing images based on YOLOv11. *Sci Rep* 15, 14032 (2025). DOI: <https://doi.org/10.1038/s41598-025-96314-x>.
6. Schmid L. et al., Dynablox: Real-time Detection of Diverse Dynamic Objects in Complex Environments. 2023. URL: <https://arxiv.org/abs/2304.10049> (дата звернення: 10.02.2026).
7. Chen Y.-H. and Wu C.-T., ReynoldsFlow: Exquisite Flow Estimation via Reynolds Transport Theorem. 2025. URL: <https://arxiv.org/abs/2503.04500> (дата звернення: 10.02.2026).
8. Yin Z., Shi J. GeoNet: Unsupervised Learning of Dense Depth, Optical Flow and Camera Pose. URL: <https://arxiv.org/abs/1803.02276v2> (дата звернення: 10.02.2026).
9. Zamora-Ortiz P., Carral-Alvaro J., Valera Á., Pulloquina J. L., Escarabajal R. J., Mata V., Identification of Inertial Parameters for Position and Force

Control of Surgical Assistance Robots. *Mathematics*. 2021. 9(7). 773. DOI: <https://doi.org/10.3390/math9070773>.

10. Trajectory Prediction for Autonomous Driving: Progress, Limitations, and Future Directions. URL: <https://arxiv.org/abs/2503.03262> (дата звернення: 10.02.2026).

11. A Real-Time Intelligent Surveillance System for Suspicious Behavior and Facial Emotion Analysis Using YOLOv8 and DeepFace. *MDPI Electronics*. 2025. URL: <https://doi.org/10.3390/engproc2025107059> (дата звернення: 10.02.2026).

12. Deep learning-based real time detection for cardiac objects with fetal ultrasound video. URL: <https://doi.org/10.1016/j.imu.2022.101150> (дата звернення: 10.02.2026).

13. Object Recognition in Different Lighting Conditions at Various Angles by Deep Learning Method. URL: <https://arxiv.org/abs/2210.09618v1> (дата звернення: 10.02.2026).

14. 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> (дата звернення: 10.02.2026).

15. Bagging vs Boosting vs Stacking URL: <https://www.geeksforgeeks.org/machine-learning/bagging-vs-boosting-vs-stacking/> (дата звернення: 10.02.2026).

16. Fernández Mareco E. R. and Pinto-Roa D., Application of Artificial Intelligence in Control Systems: Trends, Challenges, and Opportunities. *AI*. 2025. 6(12). 326. URL: <https://doi.org/10.3390/ai6120326> (дата звернення: 10.02.2026).

17. Qi M. and Tao F., Digital twins: artificial intelligence and the IoT cyber-physical systems in Industry 4.0. *Int. J. Intelligent Robotics Applications*. 2022. vol. 6. PP. 171–185, URL: <https://link.springer.com/article/10.1007/s41315-021-00180-5> (дата звернення: 10.02.2026).

18. Шматко О.В., Гамаюн І.П., Коломійцев О.В., Третьак В.Ф., Рудаков І.С., Бердочник А.Д. Дослідження та оцінка підсистеми виявлення та класифікації об'єктів у відеопотоці. *Системи обробки інформації*. 2025. № 4

(179). С. 70-80. URL: <https://doi.org/10.30748/soi.2024.179.08> (дата звернення: 10.02.2026).

19. Wang K., Wang Z., Li Z. et al. Oriented object detection in optical remote sensing images using deep learning: a survey. *Artif Intell.* 2025. Rev 58. 350. URL: <https://doi.org/10.1007/s10462-025-11256-0> (дата звернення: 10.02.2026).

20. Іванов Ю. С. Супровід динамічних об'єктів у відеопотоках розподілених систем відеоспостереження. 2015. URL: <https://ena.lpnu.ua/items/ca3d677b-4e53-4edc-893a-b7d8d12f8d3f> (дата звернення: 10.02.2026).

21. Podvyshennyi V., Safonyk A. Прогнозування траєкторій рухомих об'єктів як елемент інтелектуального відеос-постереження з БПЛА. *Proc. Modeling, Control and Information Technologies International Sc. Pract. Conf.* 2025. DOI: <https://doi.org/10.31713/MCIT.2025.053> (дата звернення: 10.02.2026).

22. Оптичний потік – поняття та застосування у відеоаналізі, Міністерство освіти і науки України, Tech. Report. 2025. URL: <https://ekmair.ukma.edu.ua/server/api/core/bitstreams/8f37cad5-932a-4141-a681-592f8949c5bf/content> (дата звернення: 10.02.2026).

23. Three dimensional tracking of rigid objects in motion using 2D optical flows. *Image and Vision Computing.* 2024. Vol. 142. 104913. URL: <https://www.sciencedirect.com/science/article/pii/S0262885624000167> (дата звернення: 10.02.2026).

24. Malhotra R., Detection and segmentation of moving objects in video using optical vector flow estimation. *MSc Thesis, Univ. of Saskatchewan.* 2008. URL: <https://harvest.usask.ca/items/69514620-0907-44ea-8ab3-b580c03cded5> (дата звернення: 10.02.2026).

25. Detecting moving objects in an optic flow field using direction and speed tuned operators. *Vision Research.* 2014. URL: <https://www.sciencedirect.com/science/article/pii/S0042698914000455> (дата звернення: 10.02.2026).

26. Stochastic Deep Model Reference Adaptive Control URL:

<https://arxiv.org/abs/2108.03120> (дата звернення: 10.02.2026).

27. Греков О., Створення динамічного цифрового двійника клієнта для симуляції бізнес-процесів. 2025. URL:

<https://openarchive.nure.ua/server/api/core/bitstreams/4e998fa3-200b-40e7-b3b4-746f3dcb4de1> (дата звернення: 10.02.2026).

28. HunyuanVideo: A Systematic Framework For Large Video. URL: <https://doi.org/10.48550/arXiv.2412.03603> (дата звернення: 10.02.2026).

29. emg2pose: A Large and Diverse Benchmark for Surface Electromyographic Hand Pose Estimation. URL: <https://doi.org/10.48550/arXiv.2412.02725> (дата звернення: 10.02.2026).

30. StableAnimator: High-Quality Identity-Preserving Human Image Animation. URL: <https://doi.org/10.48550/arXiv.2411.17697> (дата звернення: 10.02.2026).

31. DEYO: DETR with YOLO for End-to-End Object Detection. URL: <https://doi.org/10.48550/arXiv.2402.16370> (дата звернення: 10.02.2026).

32. YOLOv11: An Overview of the Key Architectural Enhancements. URL: <https://doi.org/10.48550/arXiv.2410.17725> (дата звернення: 10.02.2026).

33. A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS. URL: <https://doi.org/10.3390/make5040083> (дата звернення: 10.02.2026).

34. SynCamMaster: Synchronizing Multi-Camera Video Generation from Diverse Viewpoints. URL: <https://doi.org/10.48550/arXiv.2412.07760> (дата звернення: 10.02.2026).

35. FlowNet: Learning Optical Flow with Convolutional Networks. URL: <https://doi.org/10.48550/arXiv.1504.06852> (дата звернення: 10.02.2026).

36. FlowNet 2.0: Evolution of Optical Flow Estimation with Deep Networks. URL: <https://doi.org/10.48550/arXiv.1612.01925> (дата звернення: 10.02.2026).

37. Momentum-GS: Momentum Gaussian Self-Distillation for High-Quality Large Scene Reconstruction. URL: <https://doi.org/10.48550/arXiv.2412.04887> (дата звернення: 10.02.2026).



38. Liger Kernel: Efficient Triton Kernels for LLM Training. URL: <https://doi.org/10.48550/arXiv.2410.10989> (дата звернення: 10.02.2026).
39. Stereo Anywhere: Robust Zero-Shot Deep Stereo Matching Even Where Either Stereo or Mono Fail. URL: <https://doi.org/10.48550/arXiv.2412.04472> (дата звернення: 10.02.2026).
40. An Evolved Universal Transformer Memory. URL: <https://doi.org/10.48550/arXiv.2410.13166> (дата звернення: 10.02.2026).
41. Lucas B. D., Kanade T. An iterative image registration technique with an application to stereo vision. *Proceedings of Imaging Understanding Workshop*. 1981. P. 121–130. URL: <https://cseweb.ucsd.edu/classes/sp02/cse252/lucaskanade81.pdf> (дата звернення: 10.02.2026).
42. Bocharova M. Y., Malakhov E. V. “ResJobFit -end-to-end Artificial Neural Networks based technology for job-resume matching”. *Applied Aspects of Information Technology*. 2024. Vol. 7. No. 4. 378–391. <https://doi.org/10.15276/aait.07.2024.27>.
43. Копп А. М., Нестеренко І. С. Модель вибору інструментів штучного інтелекту для підтримки процесів розробки програмного забезпечення. *Вісник НТУ «ХПІ»*. Серія: Стратегічне управління, управління портфелями, програмами та проектами. Харків. НТУ «ХПІ», 2024. № 2 (9). С. 45-49. DOI: <https://doi.org/10.20998/2413-3000.2024.9.6>.
44. 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: 10.02.2026).
45. 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: 10.02.2026).
46. 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: 10.02.2026).
47. 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: 10.02.2026).

48. Zhu X., Hu H., Lin S., Dai J. Deformable ConvNets v2: More Deformable, Better Results. URL: <https://arxiv.org/abs/1811.11168> (дата звернення: 10.02.2026).

49. 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.

50. Kashtan V. Yu., Hnatushenko V. V. Intelligent technology for land cover monitoring due to amber mining on optical satellite images. *Naukovyi Visnyk Natsionalnoho Hirnychoho Universytetu*. 2025. (3). 156–164. <https://doi.org/10.33271/nvngu/2025-3/156>.

51. Altaf M. A., Kim M. Y. Multiple object detection and tracking in autonomous vehicles: A survey on enhanced affinity computation and its multimodal applications. *ICT Express*, 2025, 11(4): 809–818. DOI: <https://doi.org/10.1016/j.icte.2025.06.005>.

52. A Deep Dive into Generic Object Tracking: A Survey. arXiv:2507.23251, 2025. <https://doi.org/10.48550/arXiv.2507.23251>.

53. Gracyk A., Chen X. *GeONet: a neural operator for learning the Wasserstein geodesic*. URL: <https://arxiv.org/abs/2209.14440> (дата звернення: 10.02.2026).

54. 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: 10.02.2026).

55. 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.

56. Bronstein M. M., Bruna J., Cohen T., Velickovic P. Geometric Deep Learning: Grids, Groups, Graphs, Geodesics, and Gauges. 2021. DOI: [doi.org/10.48550/arXiv.2104.13478](https://doi.org/10.48550/arXiv.2104.13478).

57. Liu Y., Wang Z., Chen Q. Deep Learning for Trajectory Prediction: A Survey. *IEEE Transactions on Intelligent Transportation Systems*. 2024. 25(8). 12345-12367. DOI: [https://doi.org/10.1007/978-981-99-6586-1\\_37](https://doi.org/10.1007/978-981-99-6586-1_37)

58. Niedoba M., Green D., Naderiparizi S., Lioutas V., Lavington J.W., Liang X., Liu Y., Zhang K., Dabiri S., Ścibior A., Zwartsenberg B., Wood F. Nearest Neighbour Score Estimators for Diffusion Generative Models. *Journal of Machine Learning Research*. 2024. URL: <https://arxiv.org/abs/2402.08018> (дата звернення: 10.02.2026).

59. Zhang Y., Li S., Wang H. Multi-Agent Reinforcement Learning for Dynamic System Identification. In: *Proceedings of the International Conference on Machine Learning (ICML)*. 2025. PP. 1023–1034. DOI: <https://doi.org/10.48550/arXiv.2409.18435>.

60. Horn B. K. P., Schunck B. G. Determining optical flow. *Artificial Intelligence*. 1981. Vol. 17. P. 185–203. URL: <https://dspace.mit.edu/handle/1721.1/6337> (дата звернення: 10.02.2026).

61. Farneback G. Two-Frame Motion Estimation Based on Polynomial Expansion. *SCIA '03*. 2003. PP. 363-370. DOI: [https://doi.org/10.1007/3-540-45103-X\\_50](https://doi.org/10.1007/3-540-45103-X_50) (дата звернення: 10.02.2026).

62. Evolutionary bagging for ensemble learning. URL: <https://arxiv.org/abs/2208.02400> (дата звернення: 10.02.2026).

63. Zhou Z.-H. Ensemble Methods: Foundations and Algorithms. *Boca Raton: CRC Press*. 2012. URL: <https://www.taylorfrancis.com/books/mono/10.1201/b12207/ensemble-methods-zhi-hua-zhou> (дата звернення: 10.02.2026).

64. He Y., et al. An Efficient LiDAR-Camera Fusion Network for Multi-Class 3D Dynamic Object Detection and Trajectory Prediction. 2026. URL:

<https://arxiv.org/abs/2504.13647v2> (дата звернення: 10.02.2026).

65. Ai H., et al. A Survey of Representation Learning, Optimization Strategies, and Applications for Omnidirectional Vision. 2025. URL: <https://arxiv.org/abs/2502.10444v1> (дата звернення: 10.02.2026).

66. Kumar S., et al. A Hybrid Ensemble Learning Framework for Image-Based Solar Panel Classification. 2025. URL: <https://arxiv.org/abs/2507.01778> (дата звернення: 10.02.2026).

67. Litvynchuk A., et al. EfficientDepth: A Fast and Detail-Preserving Monocular Depth Estimation Model. 2025. URL: <https://arxiv.org/abs/2509.22527> (дата звернення: 10.02.2026).

68. Teye M.T., et al. LiDAR MOT-DETR: A LiDAR-based Two-Stage Transformer for 3D Multiple Object Tracking. 2025. URL: <https://arxiv.org/abs/2505.12753> (дата звернення: 10.02.2026).

69. Ding H., et al. MOVE: Motion-Guided Few-Shot Video Object Segmentation. 2025. URL: <https://arxiv.org/abs/2507.22061> (дата звернення: 10.02.2026).

70. Mu S., et al. StereoDETR: Stereo-based Transformer for 3D Object Detection. *IEEE TCSVT*. 2025. URL: <https://arxiv.org/abs/2511.18788> (дата звернення: 10.02.2026).

71. Sapkota R., et al. YOLO advances to its genesis: a decadal and comprehensive review of the YOLO series. *Artificial Intelligence Review*. Springer. 2025. DOI: <https://doi.org/10.1007/s10462-025-11253-3>.

72. Bashar M.K., et al. Exploring Synergistic Ensemble Learning: Uniting CNNs, MLP-Mixers, and Vision Transformers. 2025. URL: <https://arxiv.org/abs/2504.09076> (дата звернення: 10.02.2026).

73. Копп А., Орловський Д., Гамаюн І., Сапожников І. Розробка та дослідження програмного забезпечення для виявлення структурних помилок у ВРМН-моделях на основі машинного навчання. *Вісник Національного технічного університету «ХПІ». Серія: Системний аналіз, управління та інформаційні технології*. 2024. № 2 (12). С. 46–55. DOI:

<https://doi.org/10.20998/2079-0023.2024.02.08>.

74. Копп, А., Орловський, Д., Літвінова, Ю. Розробка та дослідження програмного рішення для аналізу коректності моделі бізнес-процесів з використанням машинного навчання. *Вісник Національного технічного університету «ХПІ»*. Серія: Системний аналіз, управління та інформаційні технології. 2024. № 1 (11). С. 39–46. DOI: <https://doi.org/10.20998/2079-0023.2024.01.06> (дата звернення: 10.02.2026).

75. Shmatko O., Herasymov S., Milevskyi S., Balitskyi N., Pohasii S., Aleksieiev M., Vlasov I., Melenti Y., Kovalenko Y., Peleshok Y. Development of a method for assessing the efficiency of technical systems computer dynamic simulators. *Eastern-European Journal of Enterprise Technologies*. 2025. № 2 (9 (134)), С. 50–61. DOI: <https://doi.org/10.15587/1729-4061.2025.327558>.

76. Шматко О., Голоскокова А., Король О., Рахімова І. Порівняння методів машинного навчання для інформаційної системи прогнозування діабету. *Системи управління, навігації та зв'язку*. 2021. № 4(66), С. 73–81. DOI: <https://doi.org/10.26906/SUNZ.2021.4.073>.

77. Шматко О., Малишенко Д., Волощук О. Інформаційна система для інтелектуальної класифікації клієнтів: архітектура, реалізація та експериментальні дослідження. *Системи управління, навігації та зв'язку*. 2025. № 3 (81). С. 159–166. DOI: <https://doi.org/10.26906/SUNZ.2025.3.159>.

78. Bocharova M. Y., Malakhov E. V. ResJobFit - end-to-end Artificial Neural Networks based technology for job-resume matching. *Applied Aspects of Information Technology*. 2024. Vol. 7. No. 4. 378–391. DOI: <https://doi.org/10.15276/aait.07.2024.27>.

79. Kozlov M. S. and Malakhov E. V. A Complex Model of Human Recognition in Emergencies by a Multi-level Drone System. *2024 International Symposium ELMAR*. Zadar. Croatia. 2024. pp. 317-320. DOI: <https://doi.org/10.1109/ELMAR62909.2024.10694255>.

80. Бочарова М. Ю., Малахов Є. В. Вдосконалення інформаційної технології аналізу контекстно-освічених подань фраз в галузі управління

персоналом. *Eastern-European Journal of Enterprise Technologies*. 2024. № 5 (2 (131)). 53–60. DOI: <https://doi.org/10.15587/1729-4061.2024.313970> (дата звернення: 10.02.2026).

81. Kashtan V. Yu., Hnatushenko V. V. Intelligent technology for land cover monitoring due to amber mining on optical satellite images. *Naukovyi Visnyk Natsionalnoho Hirnychoho Universytetu*. 2025. (3). 156–164. DOI: <https://doi.org/10.33271/nvngu/2025-3/156>.

82. Aziukovskyi O., Gadiatskyi V., Hnatushenko V., Ivanov D., Olevskyi V., Zavizion V. Computer-Aided Diagnosis Models for Breast Cancer Detection Decision Support Systems. *International Journal of Computing*. 2025. № 24 (1). С. 72-80. DOI: <https://doi.org/10.47839/ijc.24.1.3878>.

83. Hnatushenko V., Kashtan V. Y., Hnatushenko V. V., and Heipke C. (). Flood Forecasting with Sentinel-2 Images Using Machine Learning. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* 2025. XLVIII-G-2025. 583–588, DOI: <https://doi.org/10.5194/isprs-archives-XLVIII-G-2025-583-2025>.

84. Годлевський М., Бурлаков Г. Інформаційна технологія планування підвищення якості підмножини процесів моделі spice. *Сучасні інформаційні системи*. 2023. № 7(4). С. 52–59. DOI: <https://doi.org/10.20998/2522-9052.2023.4.06>.

85. Годлевський М., Бурлаков Г. Вербальний опис технології планування підвищення якості підмножини процесів еталонної моделі зрілості spice. *Вісник Національного технічного університету «ХПІ». Серія: Системний аналіз, управління та інформаційні технології*. 2023. № 1 (9). С. 41–48. DOI: <https://doi.org/10.20998/2079-0023.2023.01.06>.

86. Сокол В., Годлевський М., Малець Д. Оцінка якості процесу розробки програмного забезпечення іт-компанії на основі використання функції корисності. *Вісник Національного технічного університету «ХПІ». Серія: Системний аналіз, управління та інформаційні технології*. 2024. № 1 (11). С. 9–17. DOI: <https://doi.org/10.20998/2079-0023.2024.01.02>.

87. Jiang J., Yan K., Xia X., Yang, B. A Survey of Deep Learning-Based

Pedestrian Trajectory Prediction: Challenges and Solutions. *Sensors*. 2025. 25 (3). 957. DOI: <https://doi.org/10.3390/s25030957>.

88. Xu Z., et al. Towards Depth Foundation Model: Recent Trends in Vision-Based Depth Estimation. 2025. URL: <https://arxiv.org/abs/2507.11540> (дата звернення: 10.02.2026).

89. Wang R., Li W., Feng Z., et al. D3-RSMDE: 40× Faster and High-Fidelity Remote Sensing Monocular Depth Estimation. 2026. URL: <https://arxiv.org/abs/2603.16362> (дата звернення: 10.02.2026).

90. Manghotay R. A., Liang J. Lightweight Prompt-Guided CLIP Adaptation for Monocular Depth Estimation. 2026. URL: <https://arxiv.org/abs/2604.01118> (дата звернення: 10.02.2026).

91. 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.

92. 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.

93. Кондратов О. М., Нікуліна О. М., Северин В. П., 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. URL: <https://repository.kpi.kharkov.ua/items/cc2d6cd1-eed8-46b6-a99b-725a8ca30a13> (дата звернення: 10.02.2026).

94. Kondratov O., Severyn V. and Nikulina O., Intelligent Technologies of Remote Identification of Dynamic Objects. 2025 *IEEE 6th KhPI Week on Advanced Technology (KhPIWeek)*. PP. 843-848 DOI:

<https://doi.org/10.1109/KhPIWeek61436.2025.11288651>.

95. Кондратов О.М., Нікуліна О. М., Северин В. П., 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. DOI: <https://doi.org/10.20998/2079-0023.2024.01.08>.

96. Кондратов О. М., Нікуліна О. М. Програмна реалізація із використанням трансформера з оптичним потоком та GEONET для ідентифікації параметрів динамічних об'єктів. *Вісник НТУ «ХПІ». Серія: Системний аналіз, управління та інформаційні технології*. Харків. НТУ «ХПІ», 2024. № 2 (12). С. 86–91.

97. Кондратов О. М., Северин В. П., Попазов Д. К., Любарський С. М., Нікуліна О. М. Аналіз методів обчислювального інтелекту для моделювання, ідентифікації, оптимізації систем та підтримки прийняття рішень. *Вісник НТУ «ХПІ». Серія: Стратегічне управління, управління портфелями, програмами та проектами*. Харків. НТУ «ХПІ». 2024. № 2 (9). С. 35-44.

98. Кондратов О. М., Нікуліна О. М., Северин В. П., Бубнов А. І. Розробка нелінійної моделі парогенератора АЕС для інформаційної технології оптимізації управління. *Вісник НТУ «ХПІ». Серія: Системний аналіз, управління та інформаційні технології*. Харків. 2022. № 1 (7). С. 21–27. URL: <https://repository.kpi.kharkov.ua/items/df074f16-da2c-4ca3-9dee-0de53922d490> (дата звернення: 10.02.2026).

99. Кондратов О. М., Нікуліна О. М. Ідентифікація параметрів динамічних об'єктів з використанням трансформера з оптичним потоком та ансамблевих методів. *Вісник НТУ «ХПІ». Серія: Системний аналіз, управління та інформаційні технології*. Харків. 2025. № 1 (13). С. 106–112. URL: <http://samit.khpi.edu.ua/article/view/335107/324017> (дата звернення: 10.02.2026).

100. Кондратов, О. М. Попазов, Д. К. Любарський, С. М. Северин, В. П. Нікуліна, О. М. Використання обчислювального інтелекту для моделювання,



ідентифікації, оптимізації інформаційних управляючих систем та підтримки прийняття рішень. *Вісник НТУ «ХПІ». Серія: Стратегічне управління, управління портфелями, програмами та проектами.* 2025. №2 (11). С. 23–29. URL: <http://pm.khpi.edu.ua/article/view/350017> (дата звернення: 10.02.2026).

101. 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. 1247. URL: <https://repository.kpi.kharkov.ua/server/api/core/bitstreams/9efcf7fb-e64e-400b-ba2d-999db613858f/content> (дата звернення: 10.02.2026).

102. Кондратов О. М., Нікуліна О. М. Дистанційна ідентифікація параметрів об'єктів на основі FLOWNET та YOLO в порівнянні з OPTICAL FLOW та DETR. *XVIII Міжнародна науково-практична конференція магістрантів та аспірантів «Теоретичні та практичні дослідження молодих вчених» 19–22 листопада 2024 року.* Харків. НТУ «ХПІ». 2024. С. 106-107. URL: <https://repository.kpi.kharkov.ua/items/f3d2ded1-7f56-42ae-8d70-c55cafb7e556> (дата звернення: 10.02.2026).

103. Любарський С. М., Кондратов О. М., Северин В. П., Нікуліна О. М. Інтелектуальні системи підтримки прийняття рішень для оптимізації процесів керування та дистанційної ідентифікації параметрів об'єктів. *Інформаційні технології: наука, техніка, технологія, освіта, здоров'я: Тези доповідей XXXIII міжнародної науково-практичної конференції MicroCAD-2025, 14-17 травня 2025 р.* Харків. НТУ «ХПІ». 2025. С. 1445. URL: <https://ndch.kpi.kharkov.ua/wp-content/uploads/2025/06/Zbirnik-tez-2025.pdf> (дата звернення: 10.02.2026).

104. Попазов Д. К., Кондратов О. М., Северин В. П., Нікуліна О. М. Моделювання та оптимізація інформаційних управляючих систем на основі обчислювального інтелекту дистанційної ідентифікації параметрів об'єктів. *Інформаційні технології: наука, техніка, технологія, освіта, здоров'я: Тези доповідей XXXIII міжнародної науково-практичної конференції MicroCAD-2025,*

14-17 травня 2025 р. Харків. НТУ «ХПІ». 2025. С. 1451. URL: <https://ndch.kpi.kharkov.ua/wp-content/uploads/2025/06/Zbirnik-tez-2025.pdf> (дата звернення: 10.02.2026).

105. Кондратов О. М., Любарський С. М., Попазов Д. К., Нікуліна О. М., Северин В. П. Обчислювальний інтелект для моделювання та ідентифікації інформаційних управляючих систем і підтримки прийняття рішень. *XIX Міжнародна науково-практична студентська конференція магістрантів та аспірантів «Теоретичні та практичні дослідження молодих вчених» 19–21 листопада 2025 року*. Харків. 2025. С. 184-185. URL: <https://repository.kpi.kharkov.ua/items/2fd9c23a-6d04-4b5e-96b2-116420842588> (дата звернення: 10.02.2026).

106. Кондратов О. М., Нікуліна О. М. Використання нелінійної моделі парогенератора АЕС у дистанційній ідентифікації параметрів об'єкта. *XIX Міжнародна науково-практична студентська конференція магістрантів та аспірантів «Теоретичні та практичні дослідження молодих вчених» 19–21 листопада 2025 року*. Харків. 2025. С. 88-89. URL: <https://repository.kpi.kharkov.ua/items/2fd9c23a-6d04-4b5e-96b2-116420842588> (дата звернення: 10.02.2026).

**ДОДАТОК А**  
**СПИСОК ПУБЛІКАЦІЙ ЗДОБУВАЧА**

*Наукові праці, які відображають основні наукові результати дисертації.*  
*Статті у періодичних наукових виданнях, що увійшли до переліку наукових фахових видань України:*

1. Кондратов О. М., Нікуліна О. М., Северин В. П., Бубнов А. І. Розробка нелінійної моделі парогенератора АЕС для інформаційної технології оптимізації управління. / Нікуліна О. М., Северин В. П., Бубнов А. І., Кондратов О. М. *Вісник Національного технічного університету «Харківський політехнічний інститут»*. Серія: Системний аналіз, управління та інформаційні технології. 2022. № 1 (7). С. 21–27. (Б).

DOI: <https://doi.org/10.20998/2079-0023.2022.01.04>

(Наказ МОН № 1643 від 28.12.2019, Б)

2. Кондратов О. М., Нікуліна О. М., Северин В. П., Рекова Н. Ю. Аналіз інформаційних технологій для дистанційної ідентифікації динамічних об'єктів. / Нікуліна О. М., Северин В. П., Кондратов О. М., Рекова Н. Ю. *Вісник Національного технічного університету «Харківський політехнічний інститут»*. Серія: Системний аналіз, управління та інформаційні технології. Харків: НТУ «ХПІ», 2023. № 1 (9). С. 110–115. (Б).

DOI: <https://doi.org/10.20998/2079-0023.2023.01.17>

(Наказ МОН № 1643 від 28.12.2019, Б)

3. Кондратов О. М., Нікуліна О. М., Северин В. П., Ольховий О. М. Моделі дистанційної ідентифікації параметрів динамічних об'єктів з використанням трансформерів виявлення та оптичного потоку. / Нікуліна О. М., Северин В. П., Кондратов О. М., Ольховий О. М. *Вісник Національного технічного університету «Харківський політехнічний інститут»*. Серія: Системний аналіз, управління та інформаційні технології. Харків : НТУ «ХПІ», 2024. № 1 (11). С. 52–57. (Б).

DOI: <https://doi.org/10.20998/2079-0023.2024.01.08>

(Наказ МОН № 1643 від 28.12.2019, Б)

4. Кондратов О. М., Нікуліна О. М. Програмна реалізація із використанням трансформера з оптичним потоком та GEONET для ідентифікації параметрів динамічних об'єктів. *Вісник Національного технічного університету «Харківський політехнічний інститут»*. Серія: Системний аналіз, управління та інформаційні технології. Харків: НТУ «ХПІ», 2024. № 2 (12). С. 86–91. (Б).

DOI: <https://doi.org/10.20998/2079-0023.2024.02.13>

(Наказ МОН № 1643 від 28.12.2019, Б)

5. Кондратов О. М., Северин В. П., Попазов Д. К., Любарський С. М., Нікуліна О. М. Аналіз методів обчислювального інтелекту для моделювання, ідентифікації, оптимізації систем та підтримки прийняття рішень. *Вісник Національного технічного університету «Харківський політехнічний інститут»*. Серія: Стратегічне управління, управління портфелями, програмами та проектами. 2024. №2 (9). С. 35–44. (Б).

DOI: <https://doi.org/10.20998/2413-3000.2024.9.5>

(Наказ МОН № 1643 від 28.12.2019, Б)

6. Кондратов О. М., Нікуліна О. М. Ідентифікація параметрів динамічних об'єктів з використанням трансформера з оптичним потоком та ансамблевих методів. *Вісник Національного технічного університету «Харківський політехнічний інститут»*. Серія: Системний аналіз, управління та інформаційні технології. Харків: НТУ «ХПІ». 2025. № 1 (13). С. 106–112. (Б).

DOI: <https://doi.org/10.20998/2079-0023.2025.01.16>

(Наказ МОН № 1643 від 28.12.2019, Б)

7. Кондратов О. М., Северин В. П., Попазов Д. К., Любарський С. М., Нікуліна О. М. Використання обчислювального інтелекту для моделювання, ідентифікації, оптимізації інформаційних управляючих систем та підтримки прийняття рішень. *Вісник Національного технічного університету «Харківський політехнічний інститут»*. Серія: Стратегічне управління, управління портфелями, програмами та проектами. 2025. №2 (11). С. 23–29. (Б).

DOI: <https://doi.org/10.20998/2413-3000.2025.11.3>

(Наказ МОН № 1643 від 28.12.2019, Б)

Інші публікації:

Опубліковані праці апробаційного характеру:

8. Kondratov O., Severyn V. and Nikulina O., Intelligent Technologies of Remote Identification of Dynamic Objects. *2025 IEEE 6th KhPI Week on Advanced Technology (KhPIWeek)*. PP. 843-848 (Scopus, Україна).

DOI: <https://doi.org/10.1109/KhPIWeek61436.2025.11288651>

9. Кондратов О. М., Нікуліна О. М. Методи дистанційної ідентифікації динамічних параметрів об'єкта. / Нікуліна О. М., Кондратов О. М. *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.

URL: [http://science.kpi.kharkov.ua/wp-content/uploads/2023/05/Zbirnik-tez-MicroCAD-2023-new\\_compressed-1.pdf](http://science.kpi.kharkov.ua/wp-content/uploads/2023/05/Zbirnik-tez-MicroCAD-2023-new_compressed-1.pdf)

10. Кондратов О. М., Нікуліна 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. 1247.

URL: <https://repository.kpi.kharkov.ua/server/api/core/bitstreams/9efcf7fb-e64e-400b-ba2d-999db613858f/content>

11. Кондратов О. М., Нікуліна О. М. Дистанційна ідентифікація параметрів об'єктів на основі FLOWNET та YOLO в порівнянні з OPTICAL FLOW та DETR. *XVIII Міжнародна науково-практична конференція магістрантів та аспірантів «Теоретичні та практичні дослідження молодих вчених» 19–22 листопада 2024 року.* Харків. НТУ «ХПІ». 2024. С. 106-107.

URL: <https://repository.kpi.kharkov.ua/items/f3d2ded1-7f56-42ae-8d70-c55cafb7e556>

12. Кондратов О. М., Любарський С. М., Северин В. П., Нікуліна О. М. Інтелектуальні системи підтримки прийняття рішень для оптимізації процесів керування та дистанційної ідентифікації параметрів об'єктів. / Любарський С. М., Кондратов О. М., Северин В. П., Нікуліна О. М. *Інформаційні технології: наука, техніка, технологія, освіта, здоров'я: Тези доповідей XXXIII міжнародної науково-практичної конференції MicroCAD-2025, 14-17 травня 2025 р.* Харків. НТУ «ХПІ». 2025. С. 1445.

URL: <https://ndch.kpi.kharkov.ua/wp-content/uploads/2025/06/Zbirnik-tez-2025.pdf>

13. Кондратов О. М., Попазов Д. К., Северин В. П., Нікуліна О. М. Моделювання та оптимізація інформаційних управляючих систем на основі обчислювального інтелекту дистанційної ідентифікації параметрів об'єктів. / Попазов Д. К., Кондратов О. М., Северин В. П., Нікуліна О. М. *Інформаційні технології: наука, техніка, технологія, освіта, здоров'я: Тези доповідей XXXIII міжнародної науково-практичної конференції MicroCAD-2025, 14-17 травня 2025 р.* Харків. НТУ «ХПІ». 2025. С. 1451.

URL: <https://ndch.kpi.kharkov.ua/wp-content/uploads/2025/06/Zbirnik-tez-2025.pdf>

14. Кондратов О. М., Любарський С. М., Попазов Д. К., Нікуліна О. М., Северин В. П. Обчислювальний інтелект для моделювання та ідентифікації інформаційних управляючих систем і підтримки прийняття рішень. *XIX Міжнародна науково-практична студентська конференція магістрантів та аспірантів «Теоретичні та практичні дослідження молодих вчених» 19–21 листопада 2025 року.* Харків. НТУ «ХПІ». 2025. С. 184-185.

URL: <https://repository.kpi.kharkov.ua/items/2fd9c23a-6d04-4b5e-96b2-116420842588>

15. Кондратов О. М., Нікуліна О. М. Використання нелінійної моделі парогенератора АЕС у дистанційній ідентифікації параметрів об'єкта. *XIX Міжнародна науково-практична студентська конференція магістрантів та аспірантів «Теоретичні та практичні дослідження молодих вчених» 19–21*

*листопада 2025 року. Харків. НТУ «ХПІ». 2025. С. 88-89.*

URL: <https://repository.kpi.kharkov.ua/items/2fd9c23a-6d04-4b5e-96b2-116420842588>

## ДОДАТОК Б РЕЗУЛЬТАТИ ПОЄДНАННЯ МЕТОДІВ ОПТИЧНОГО ПОТОКУ, МЕТОДІВ ОЦІНЮВАННЯ ГЛИБИНИ І ТРАНСФОРМЕРІВ

### Б.1 Результати поєднання методу Farneback, методів оцінювання глибини і трансформерів

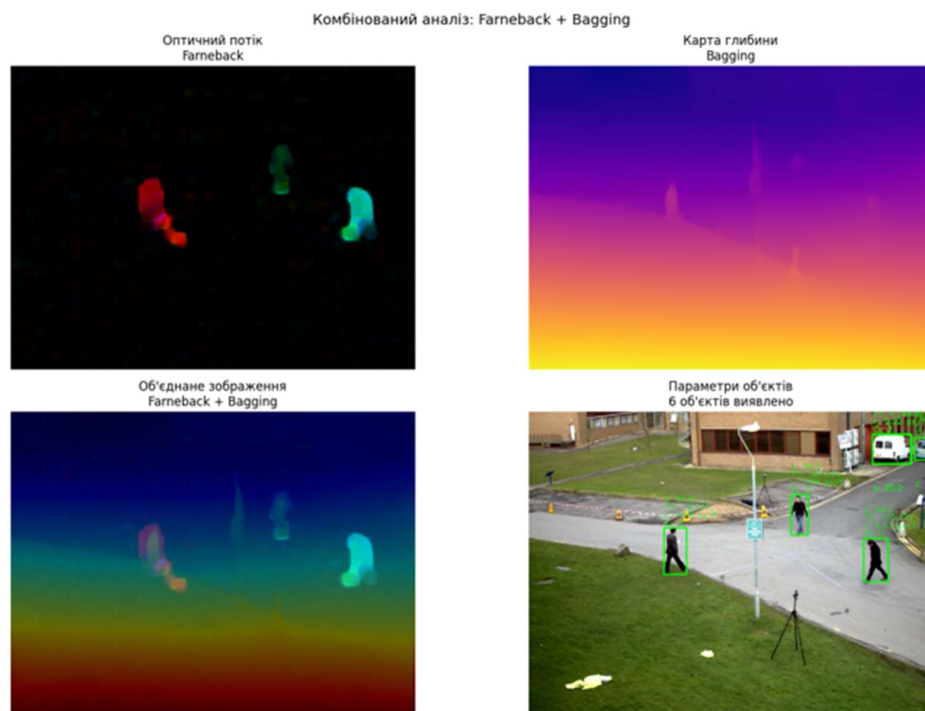


Рисунок Б.1 – Farneback\_Bagging

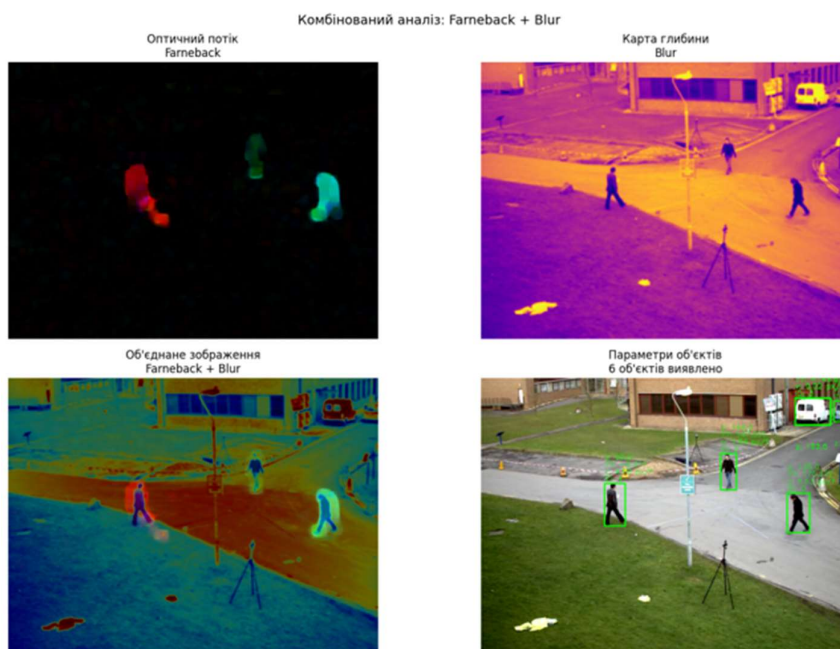


Рисунок Б.2 – Farneback і Blur



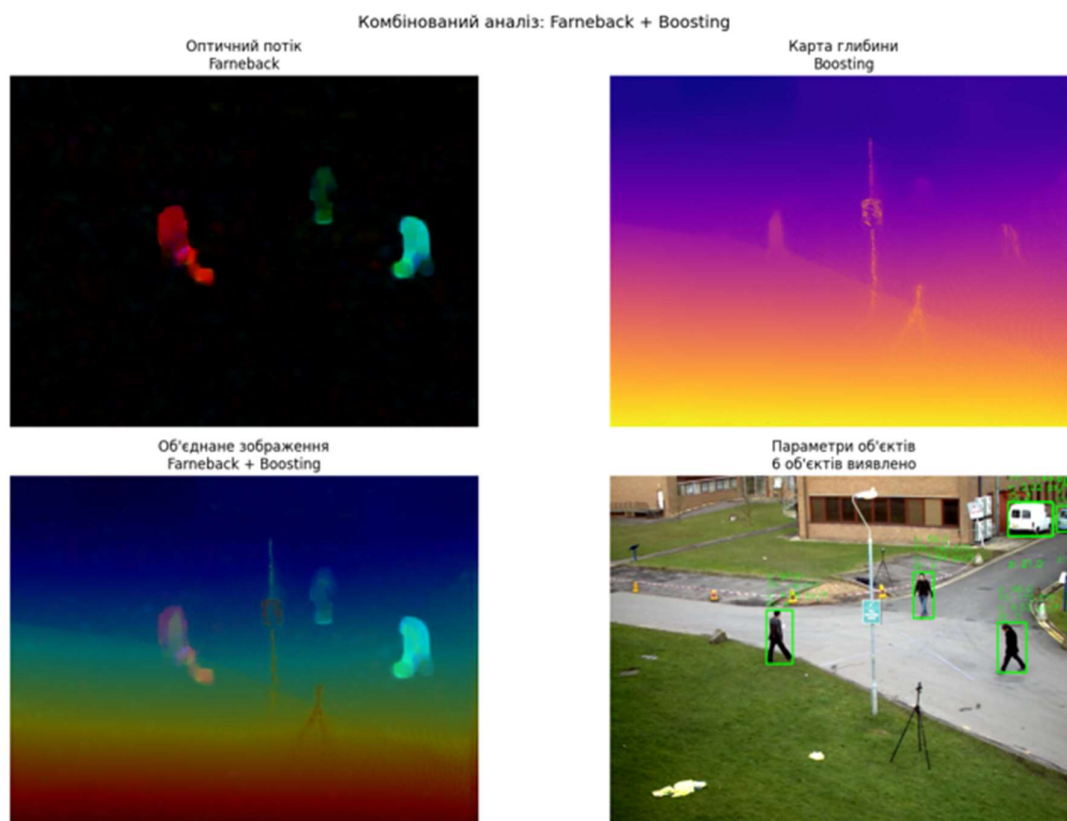


Рисунок Б.3 – Farneback\_Boosting

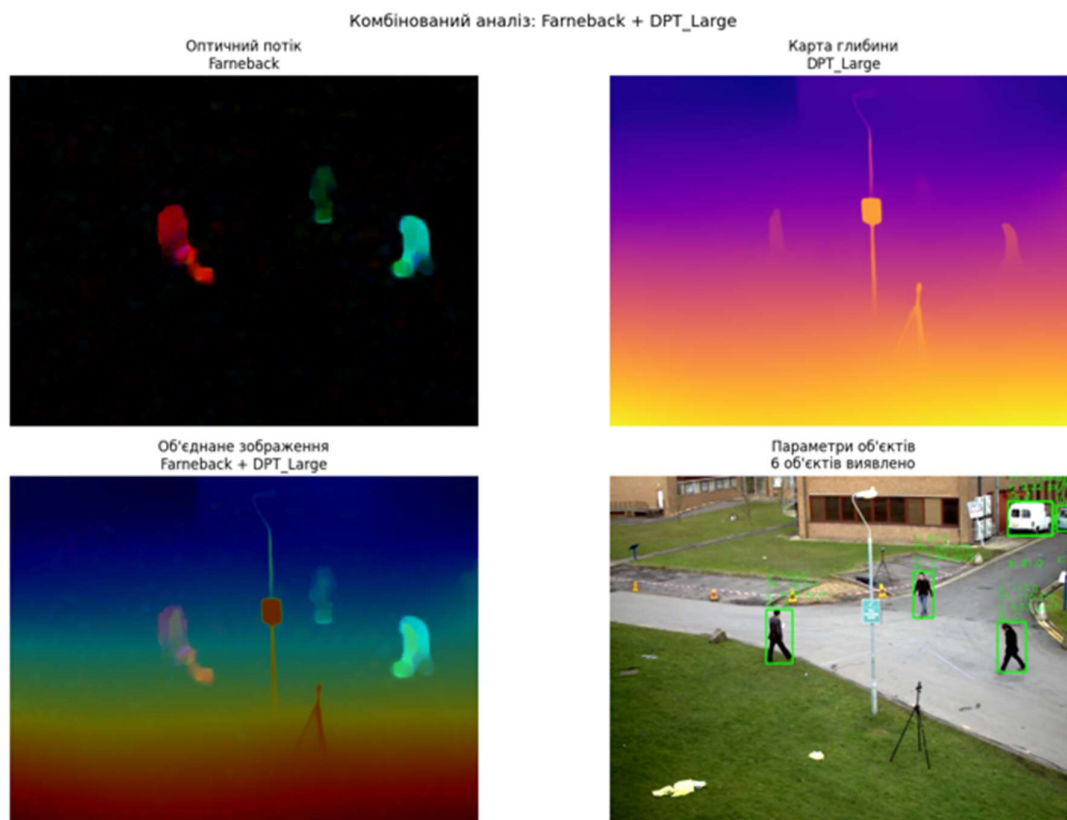


Рисунок Б.4 – Farneback\_DPT\_Large

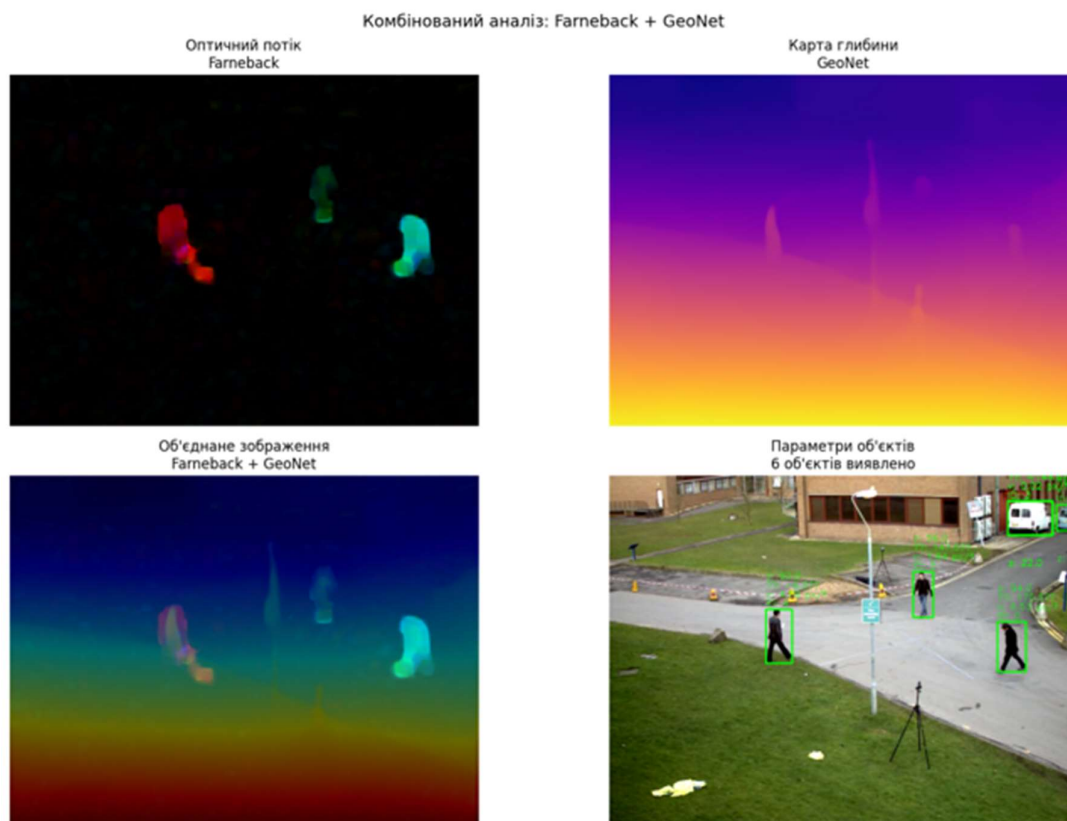


Рисунок Б.5 – Farneback\_GeoNet

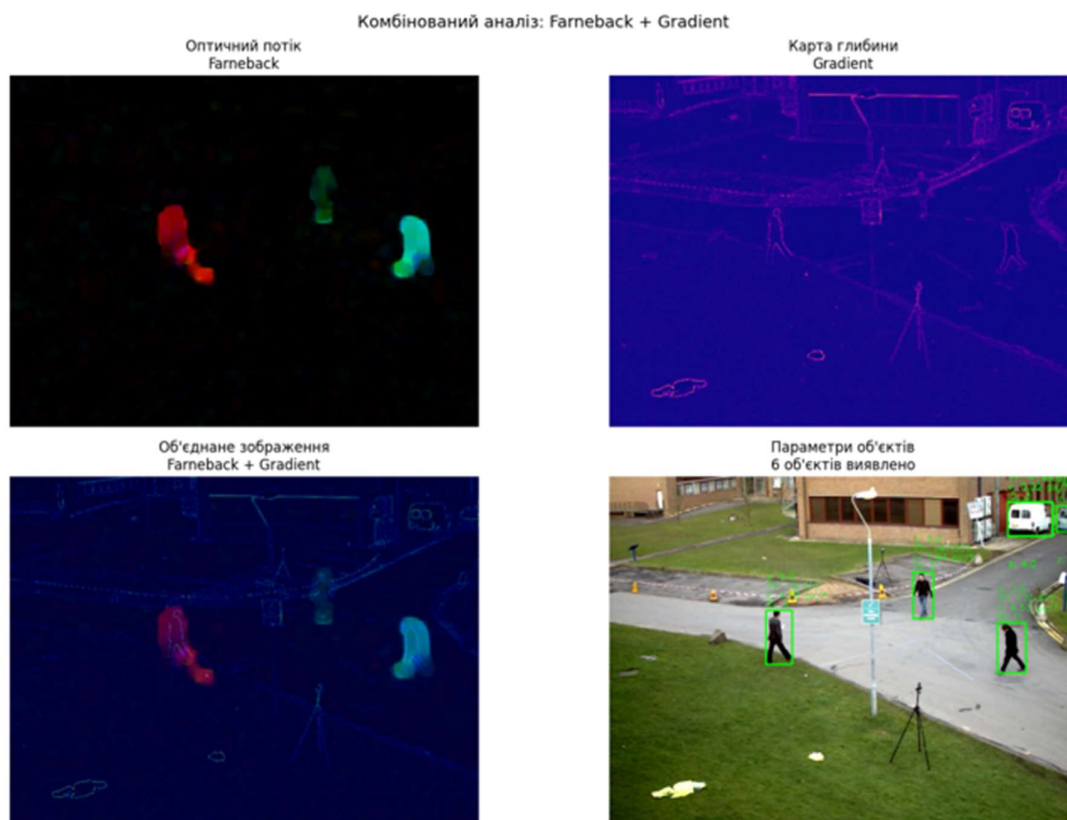


Рисунок Б.6 – Farneback\_Gradient

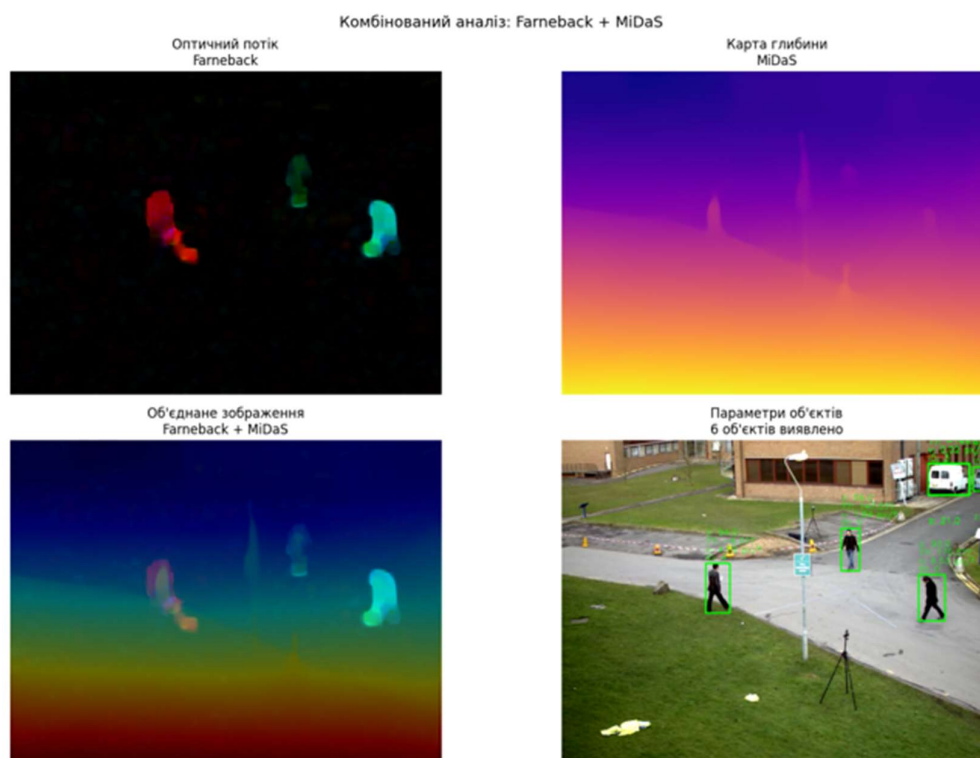


Рисунок Б.7 – Farneback\_MiDaS

## Б.2 Результати поєднання методу FlowNet, методів оцінювання глибини і трансформерів

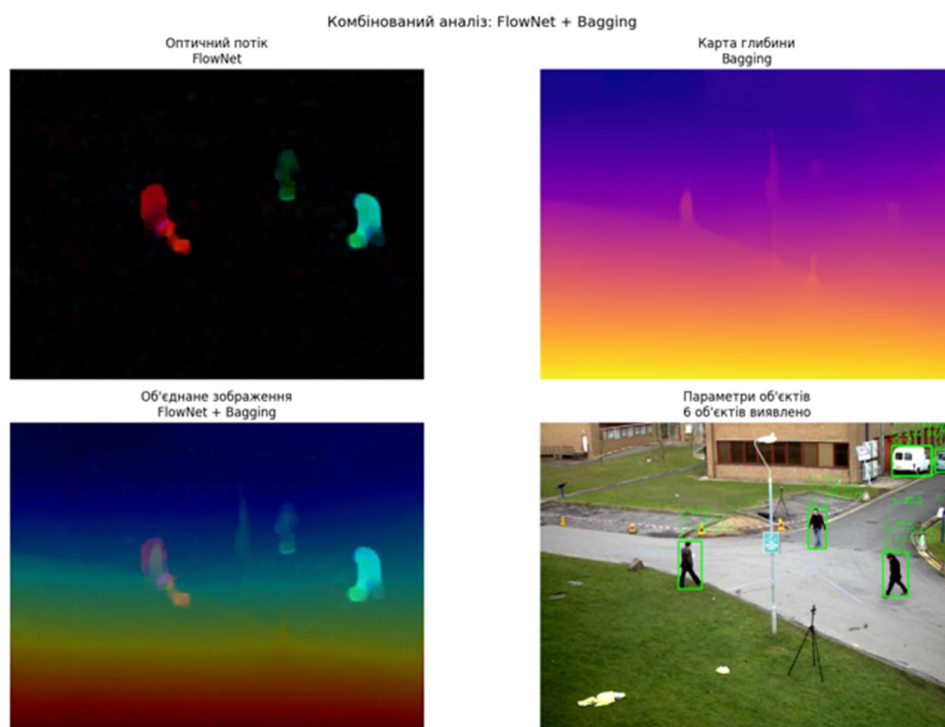


Рисунок Б.8 – FlowNet\_Bagging

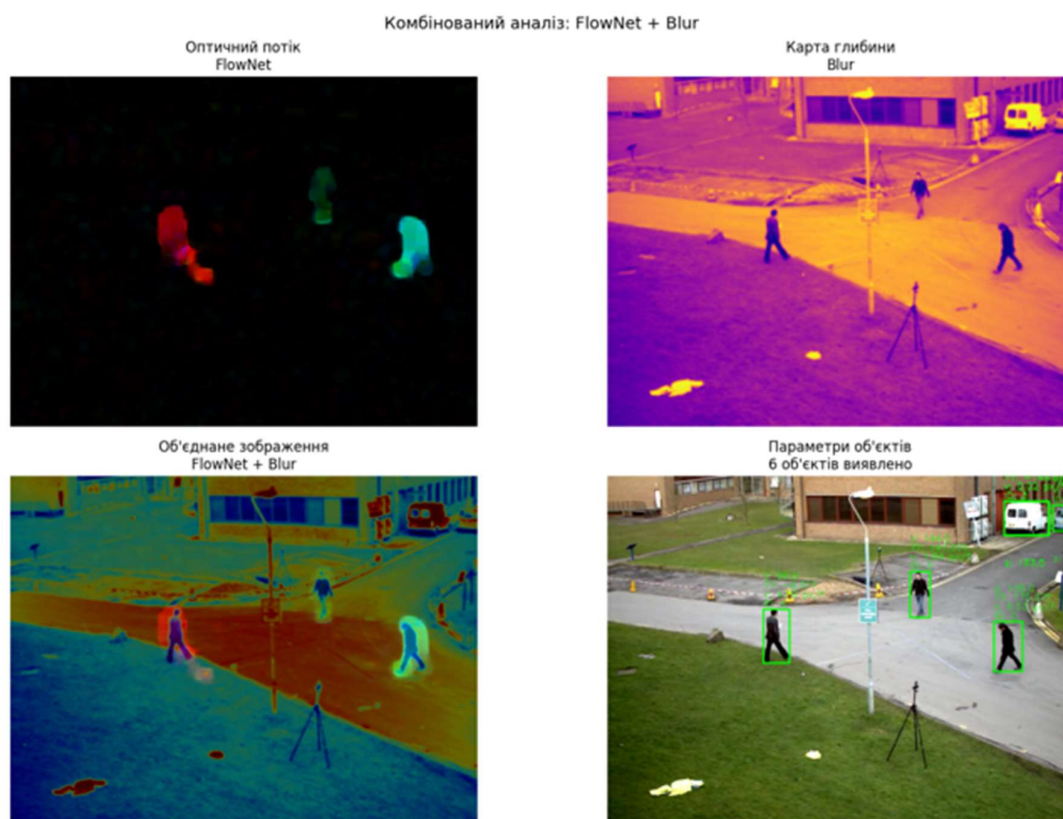


Рисунок Б.9 – FlowNet\_Blur

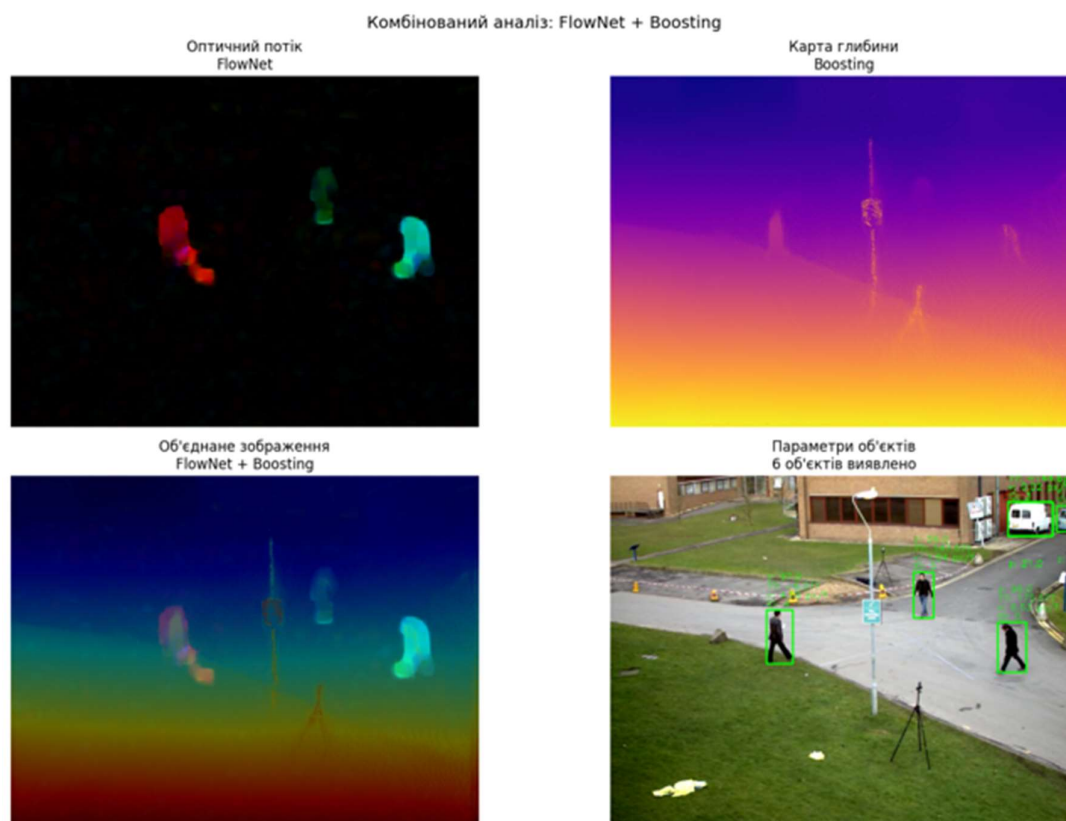


Рисунок Б.10 – FlowNet\_Boosting



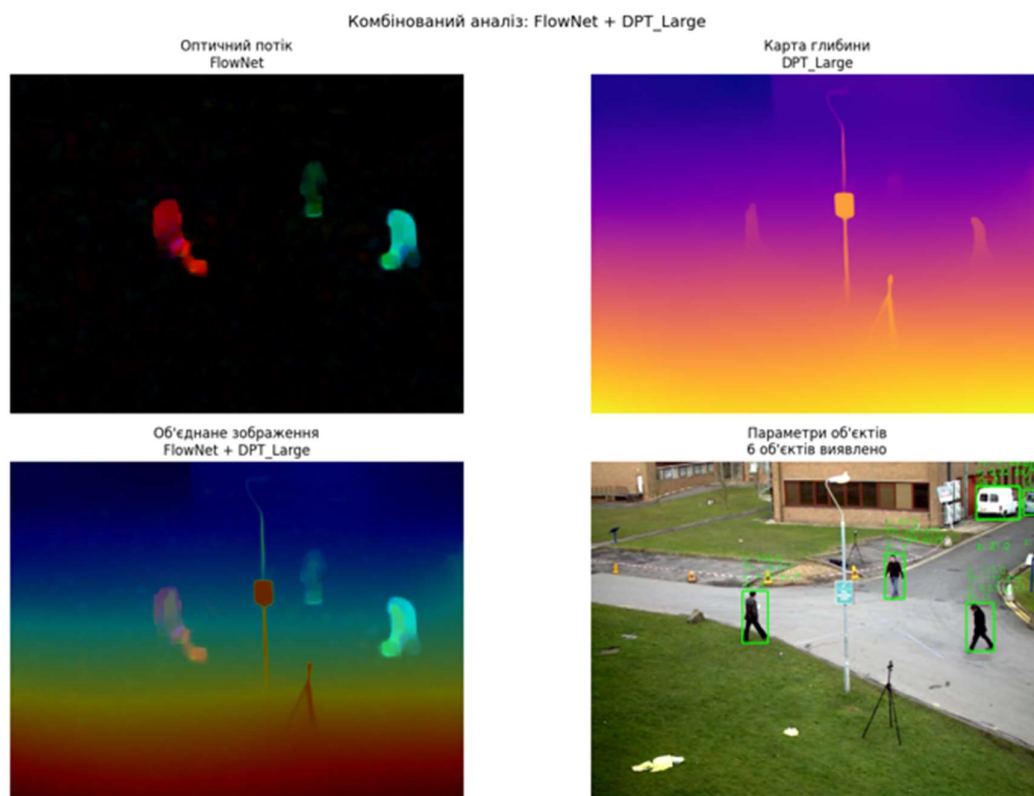


Рисунок Б.11 – FlowNet\_DPT\_Large

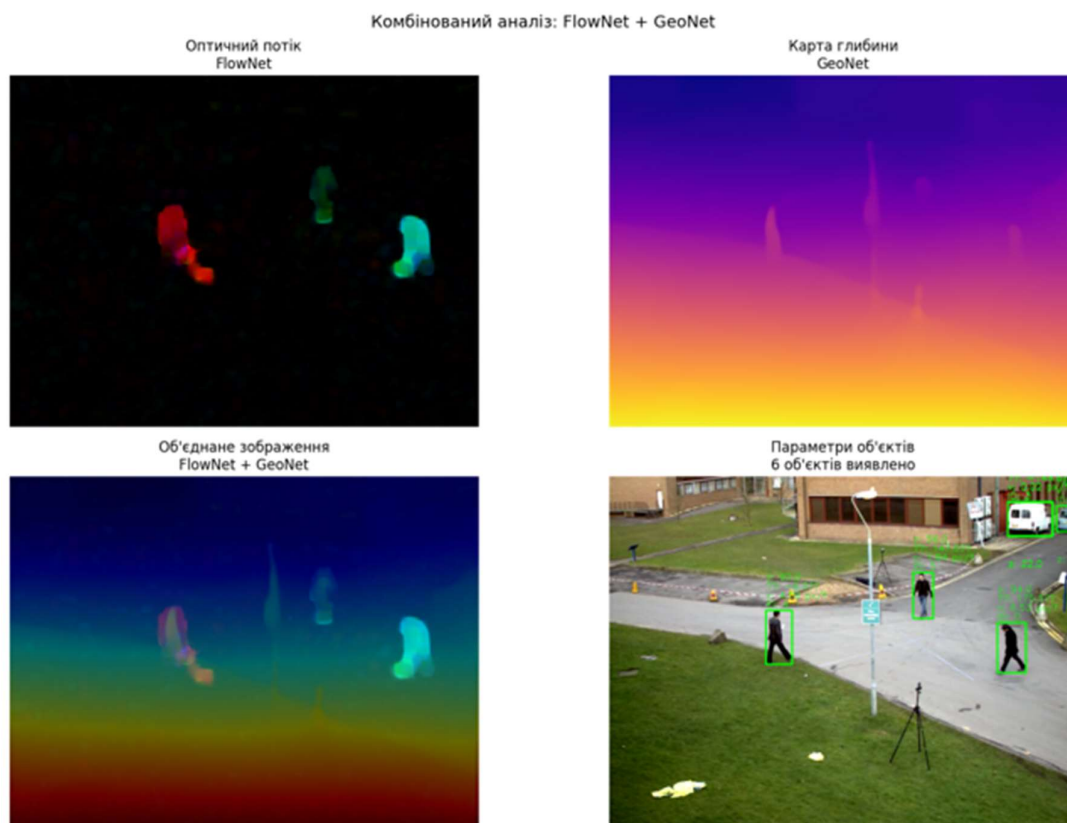


Рисунок Б.12 – FlowNet\_GeoNet

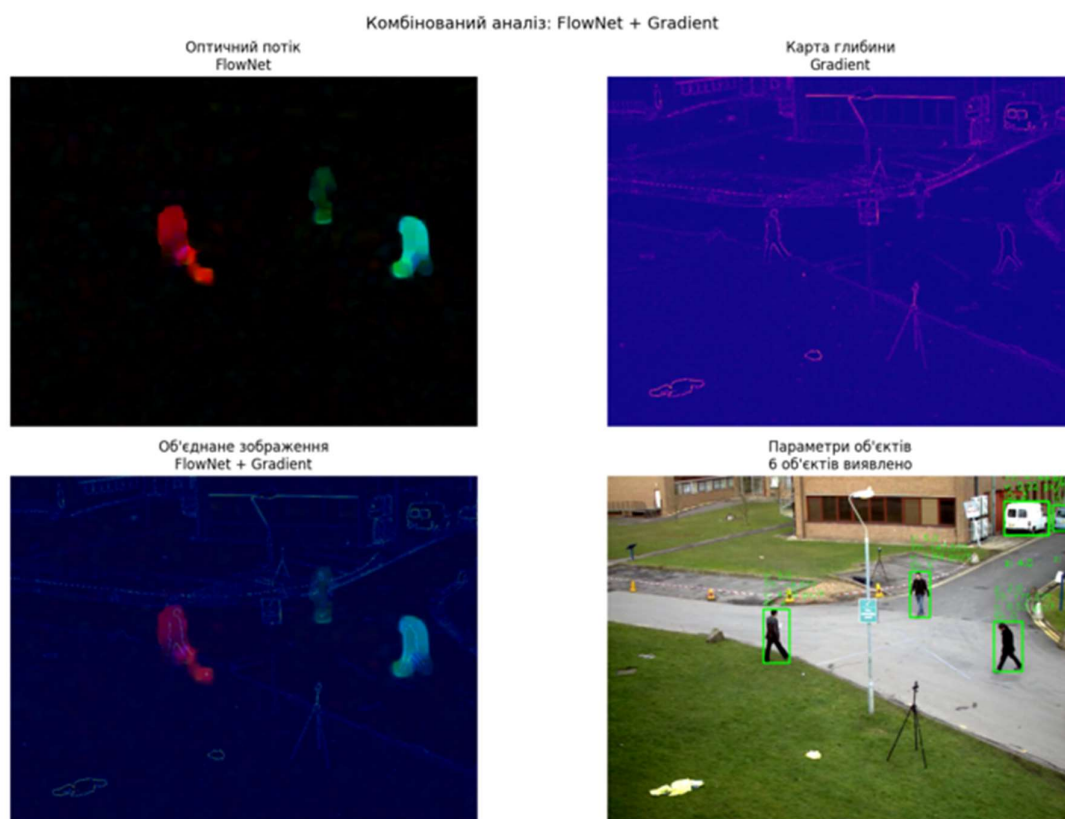


Рисунок Б.13 – FlowNet\_Gradient

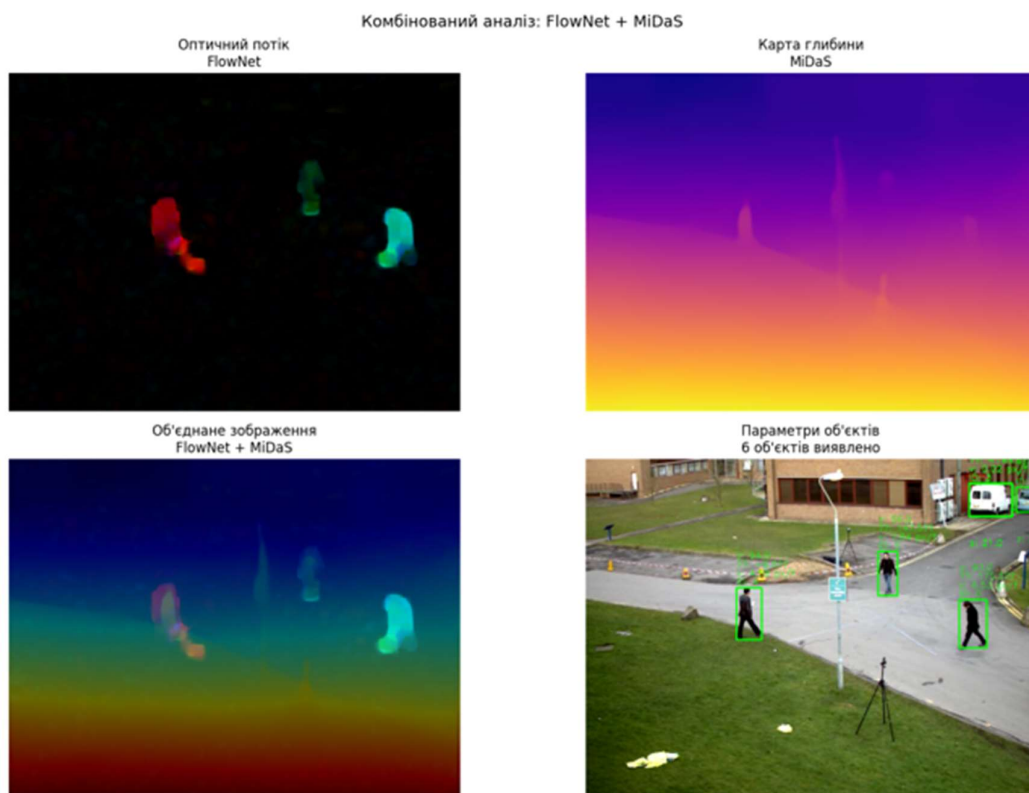


Рисунок Б.14 – FlowNet\_MiDaS

### Б.3 Результати поєднання методу GeoNet, методів оцінювання глибини і трансформерів

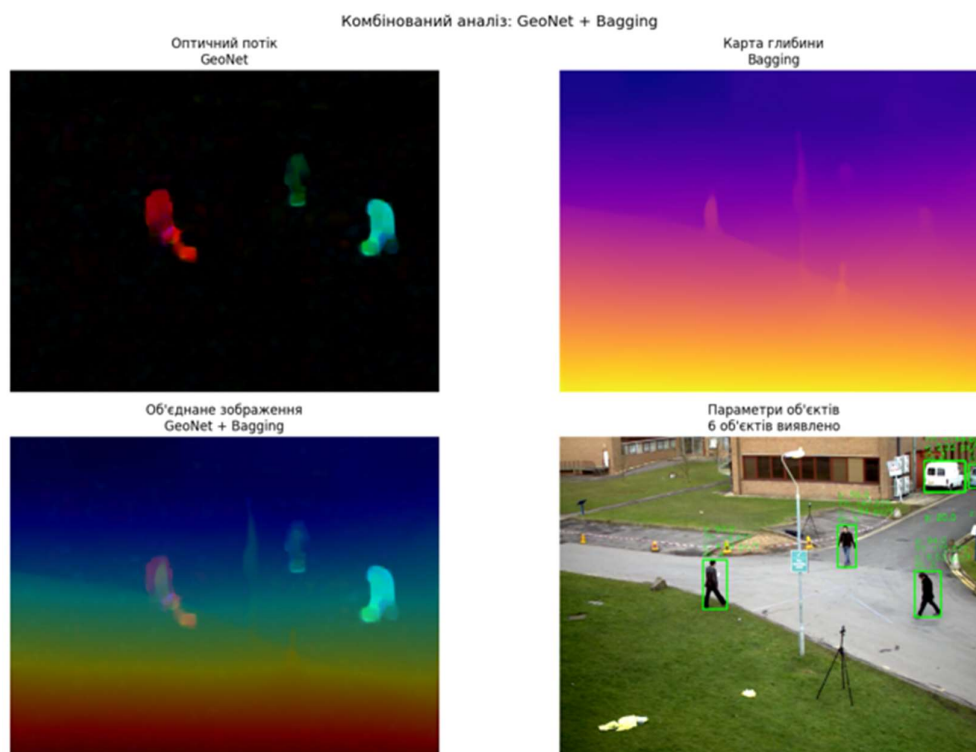


Рисунок Б.15 – GeoNet\_Bagging

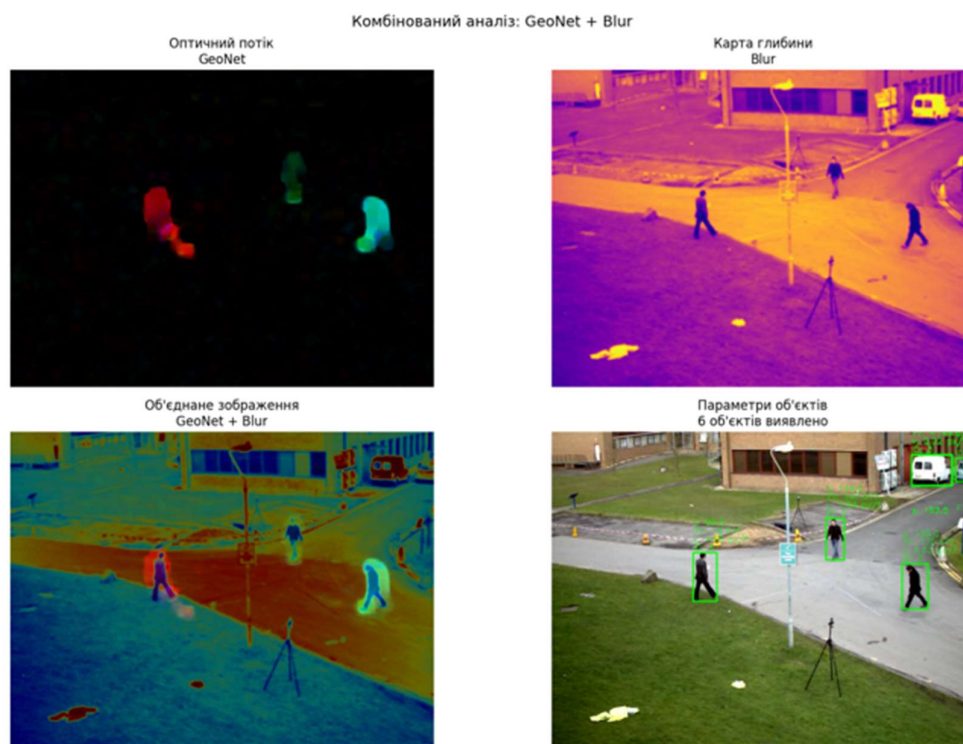


Рисунок Б.16 –GeoNet\_Blur

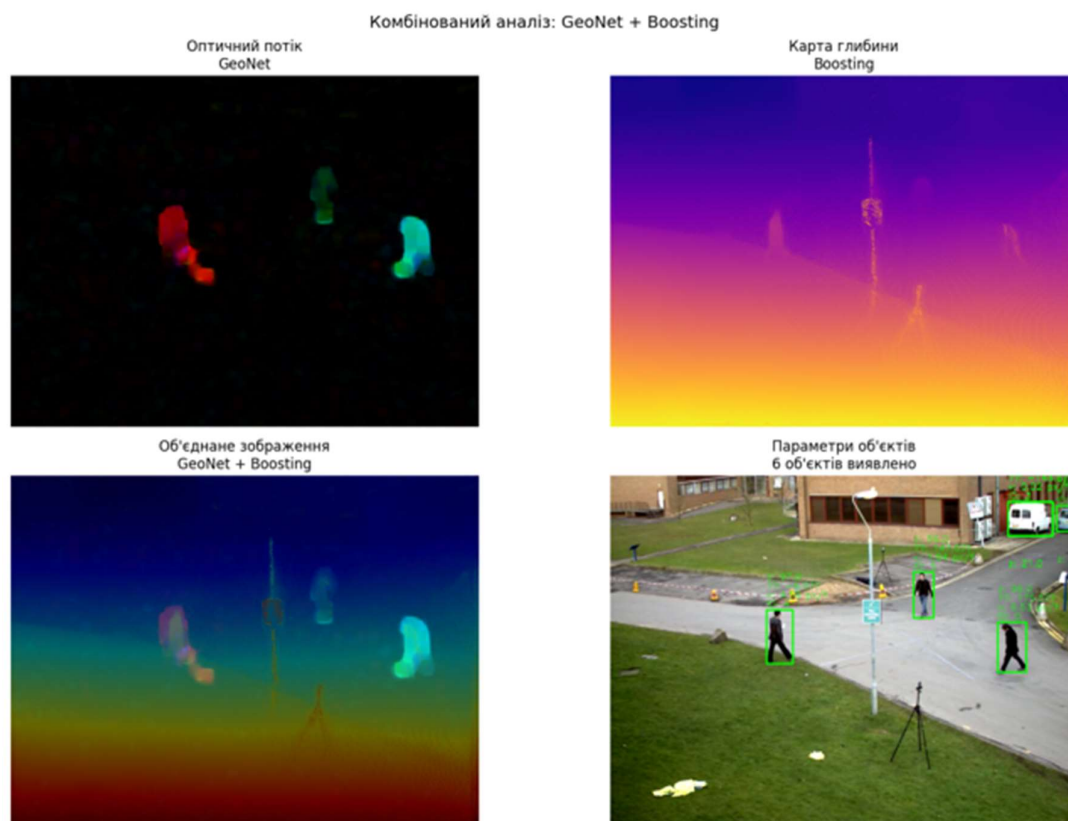


Рисунок Б.17 – GeoNet\_Boosting

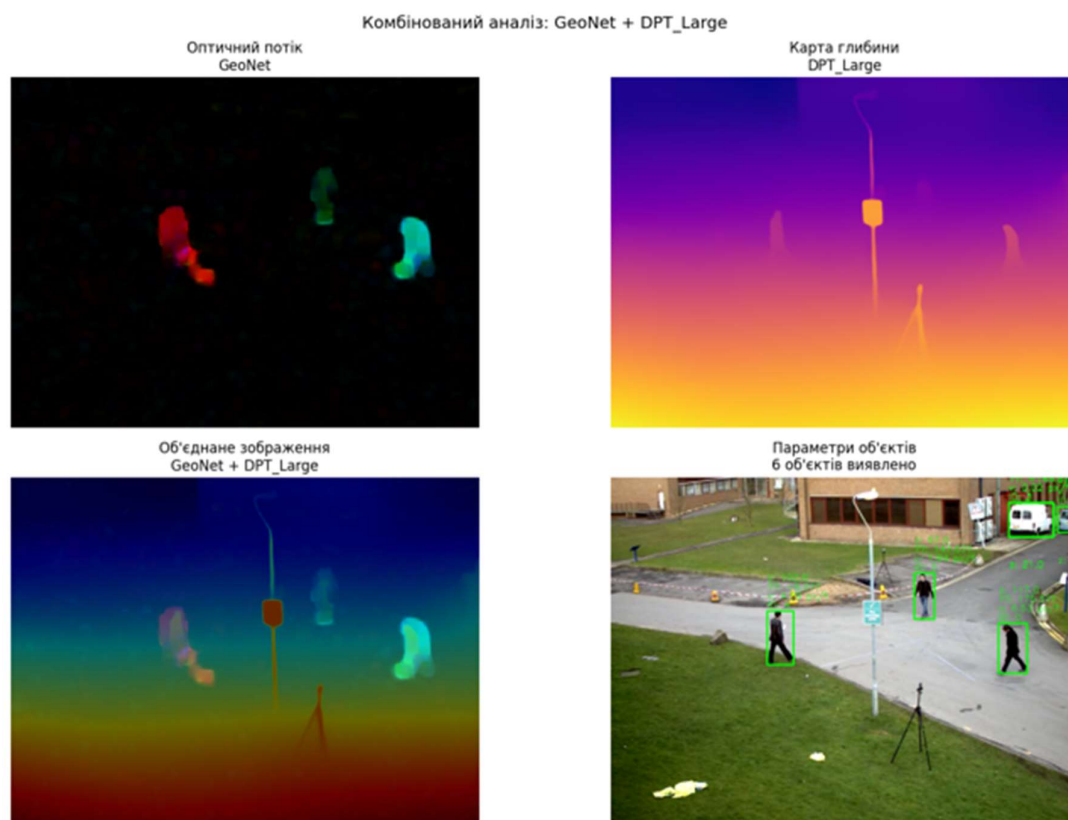


Рисунок Б.18 – GeoNet\_DPT\_Large



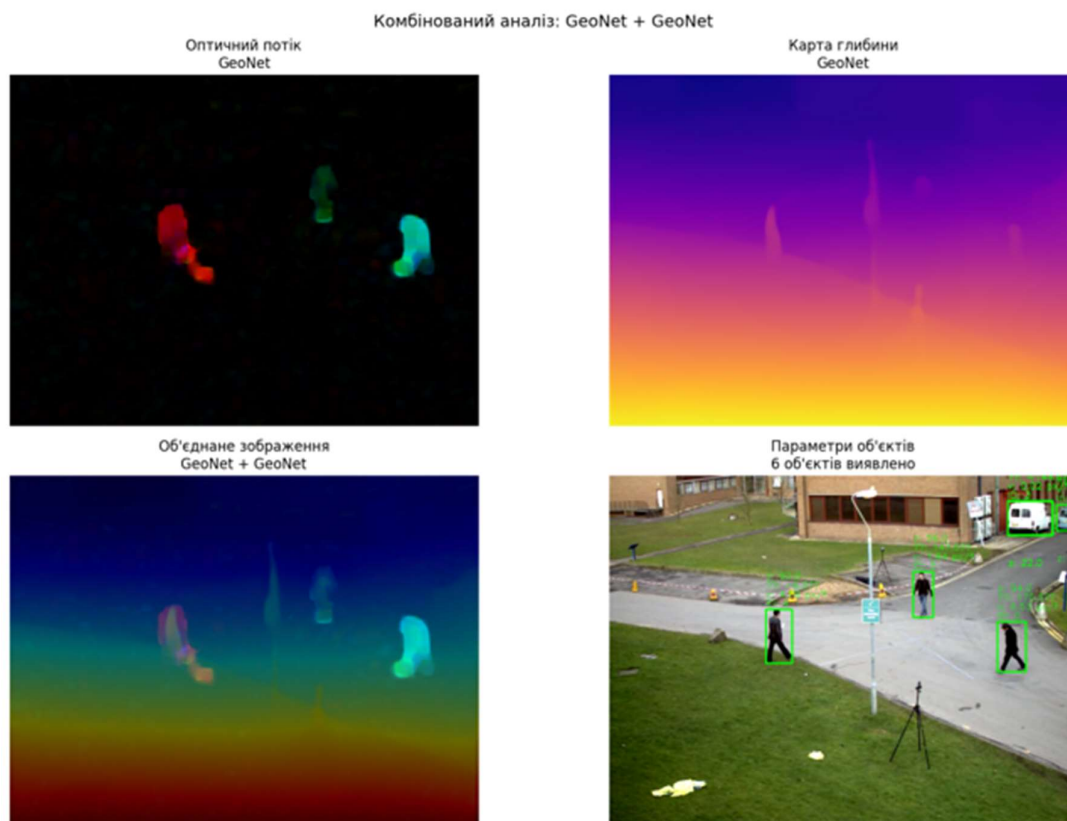


Рисунок Б.19 – GeoNet\_GeoNet

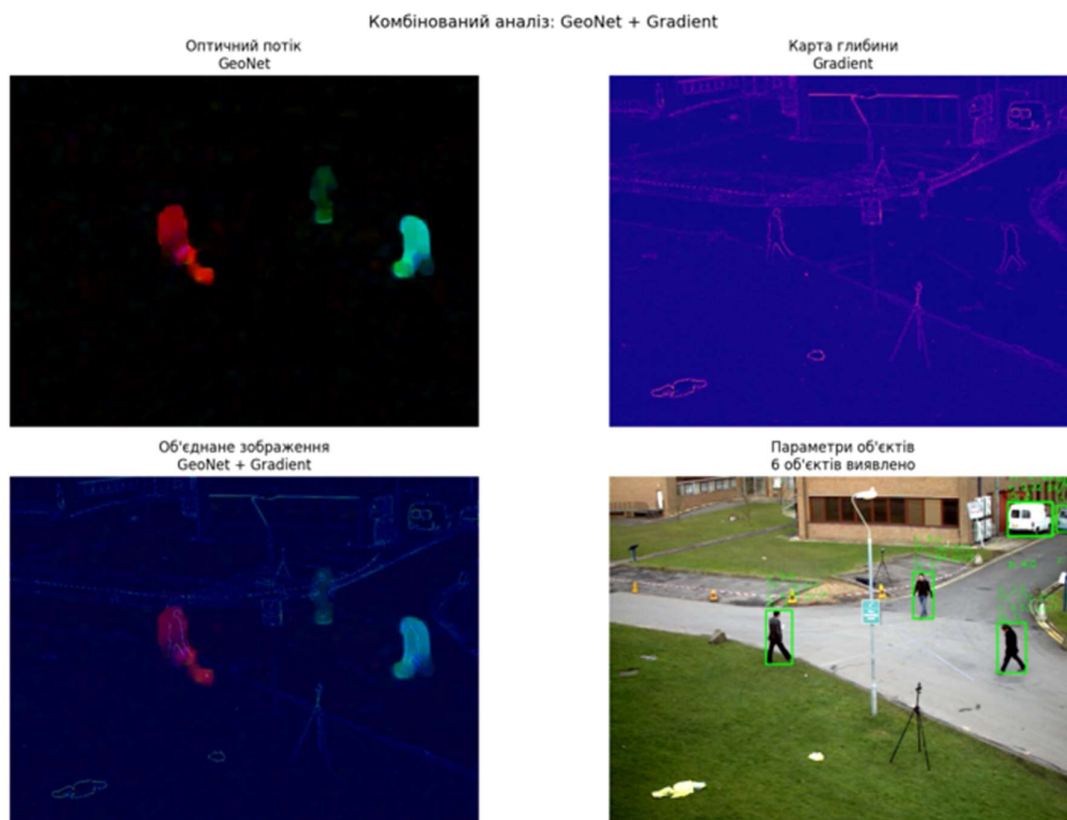


Рисунок Б.20 – GeoNet\_Gradient

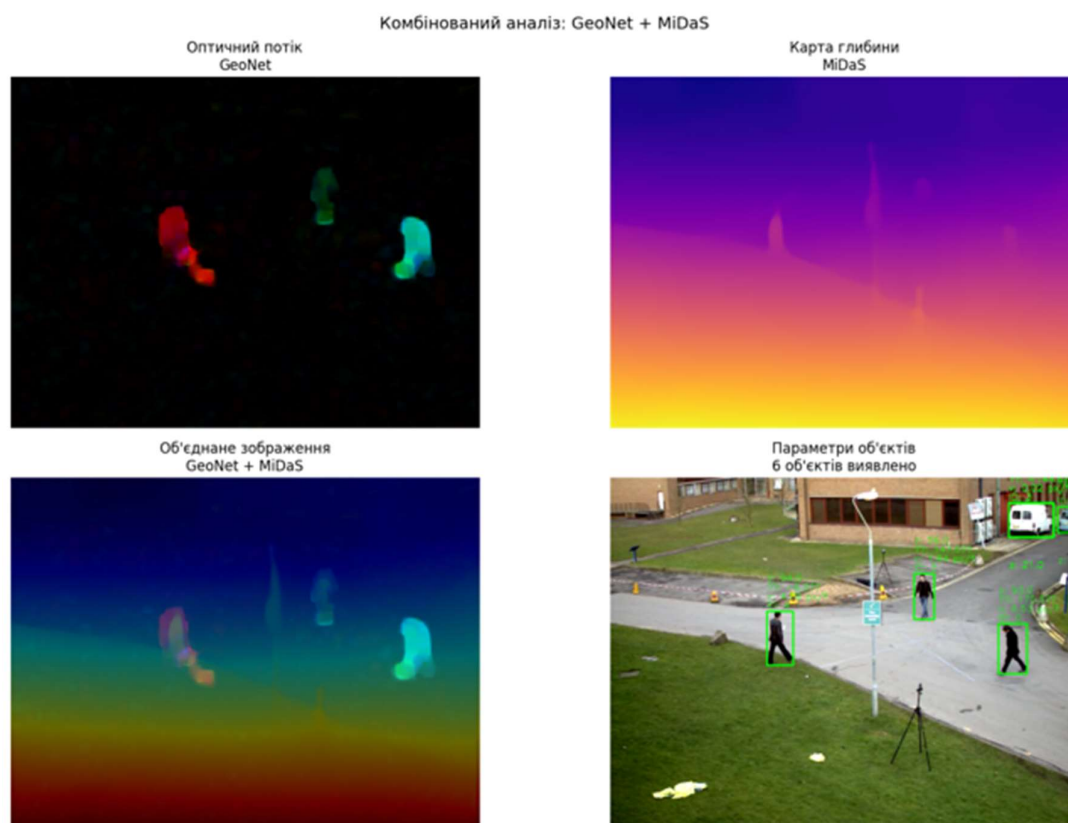


Рисунок Б.21 – GeoNet\_MiDaS

#### Б.4 Результати поєднання методу Lucas–Kanade, методів оцінювання глибини і трансформерів

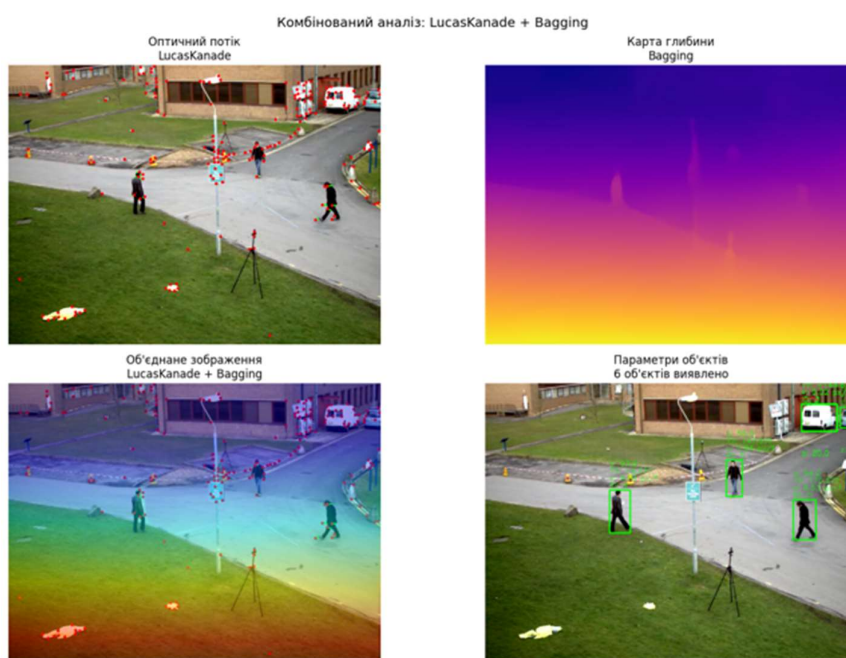


Рисунок Б.22 – LucasKanade\_Bagging

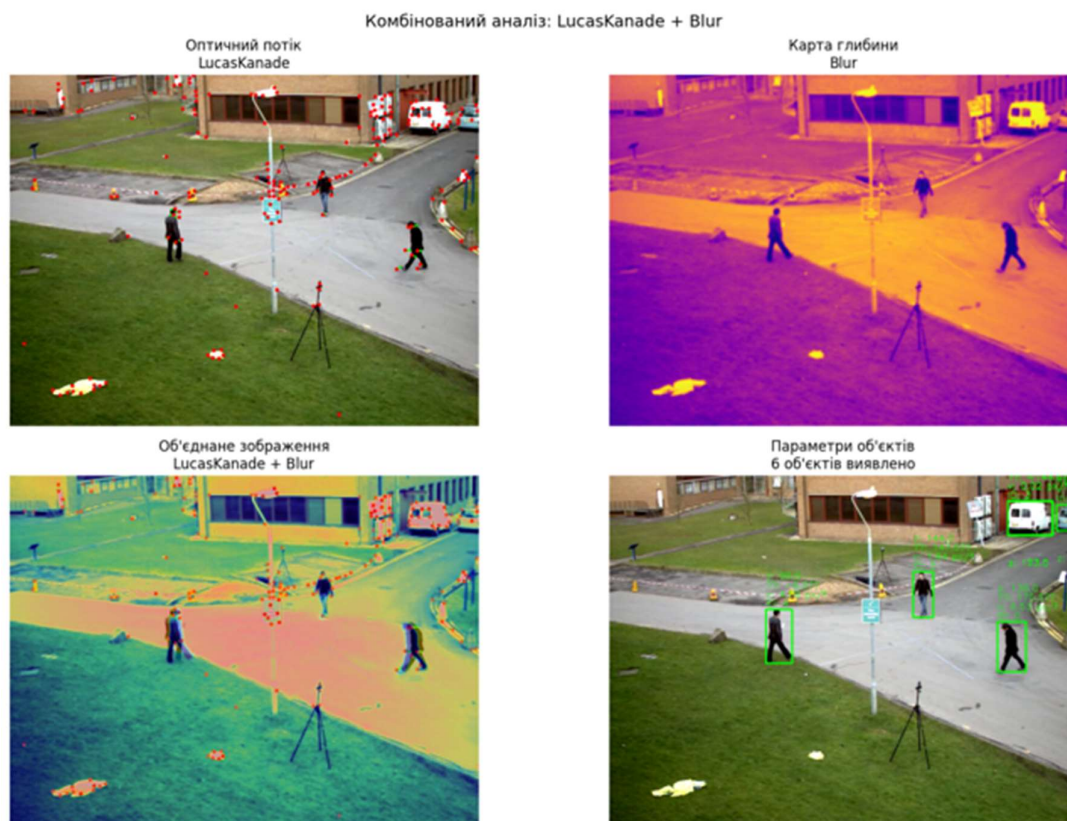


Рисунок Б.23 – LucasKanade\_Blur

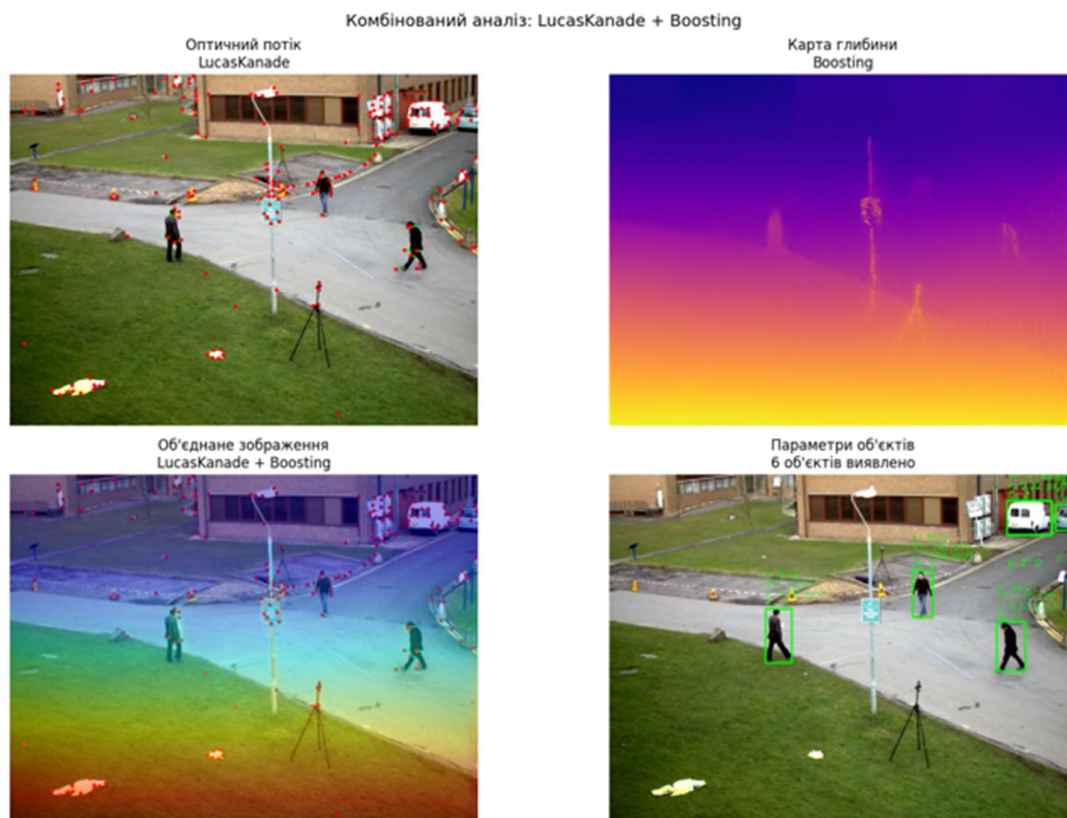


Рисунок Б.24 – LucasKanade\_Boosting



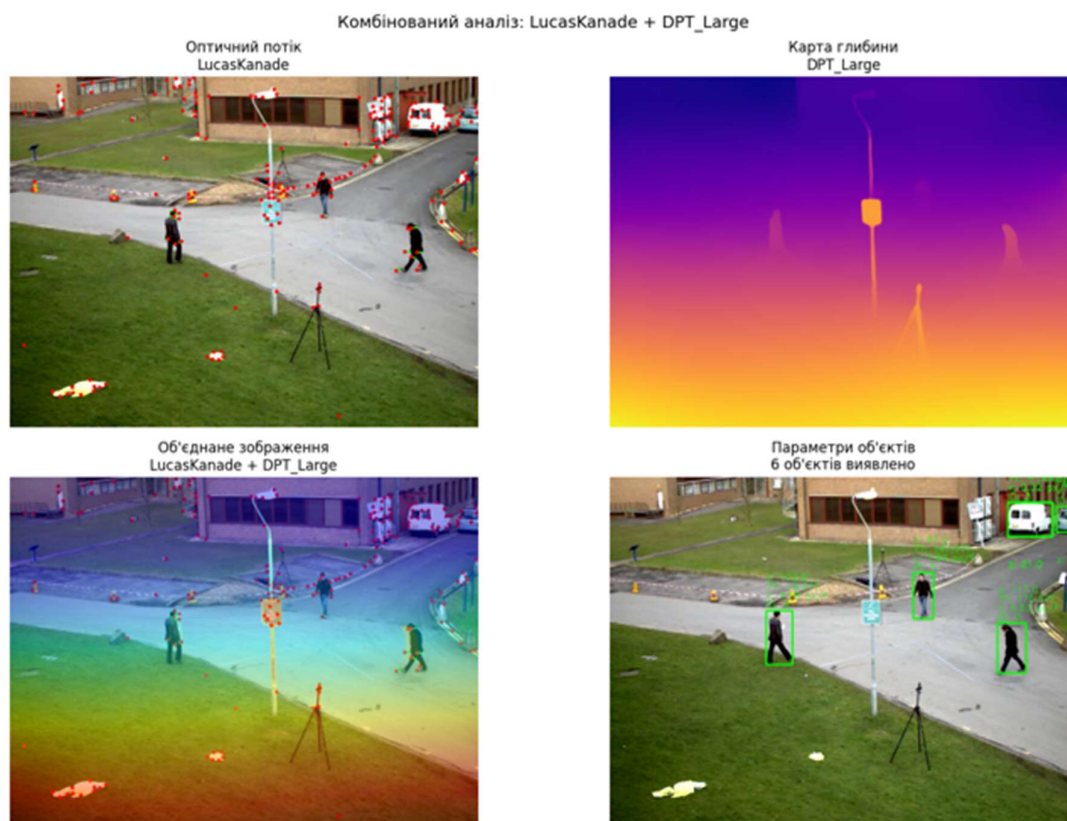


Рисунок Б.25 – LucasKanade\_DPT\_Large

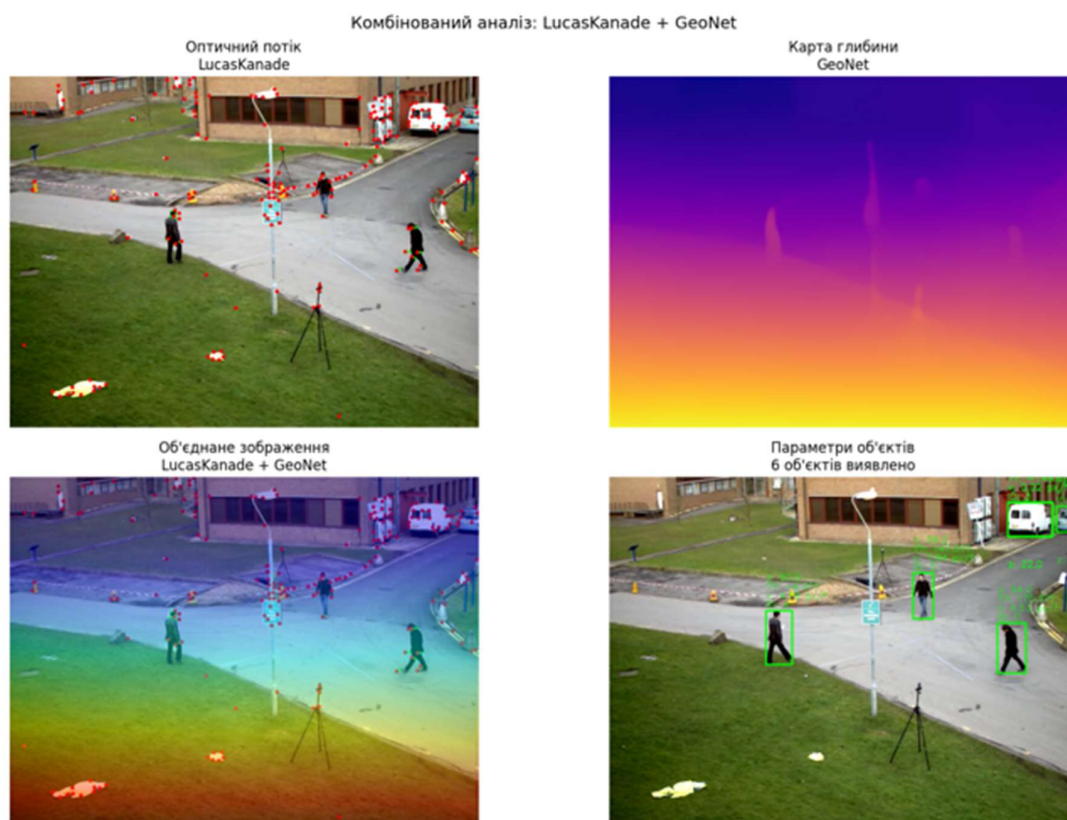


Рисунок Б.26 – LucasKanade\_GeoNet

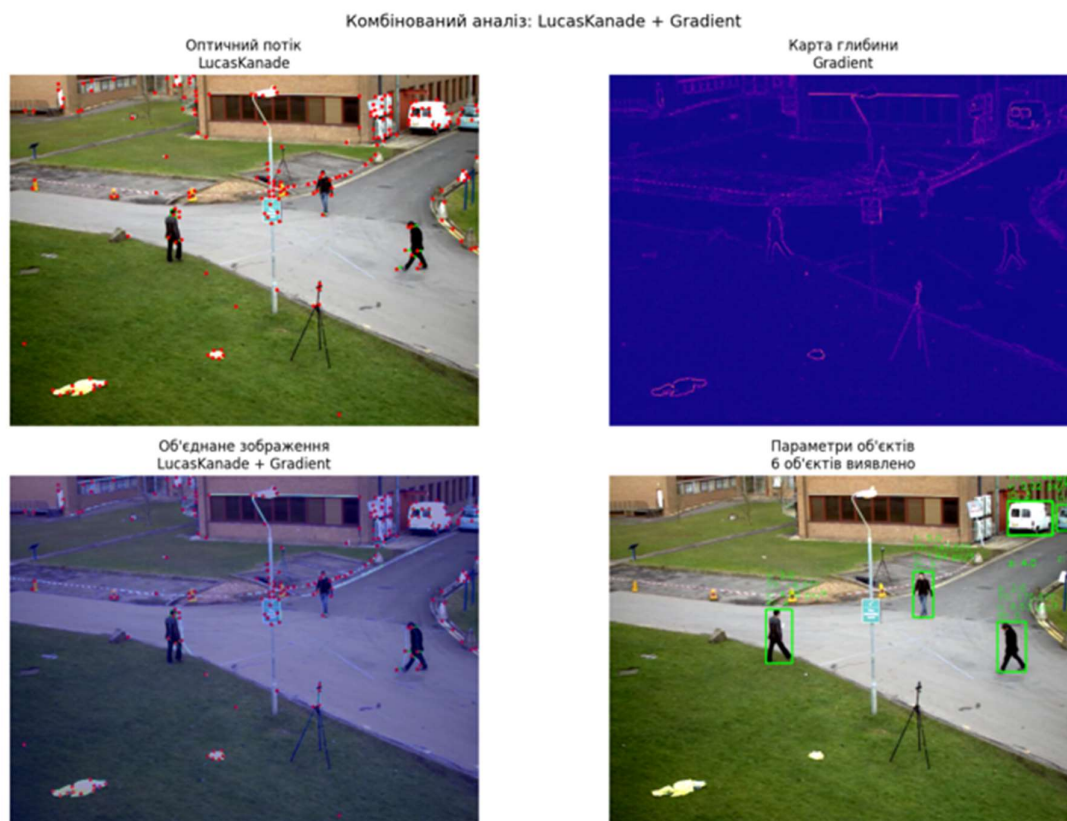


Рисунок Б.27 – LucasKanade\_Gradient

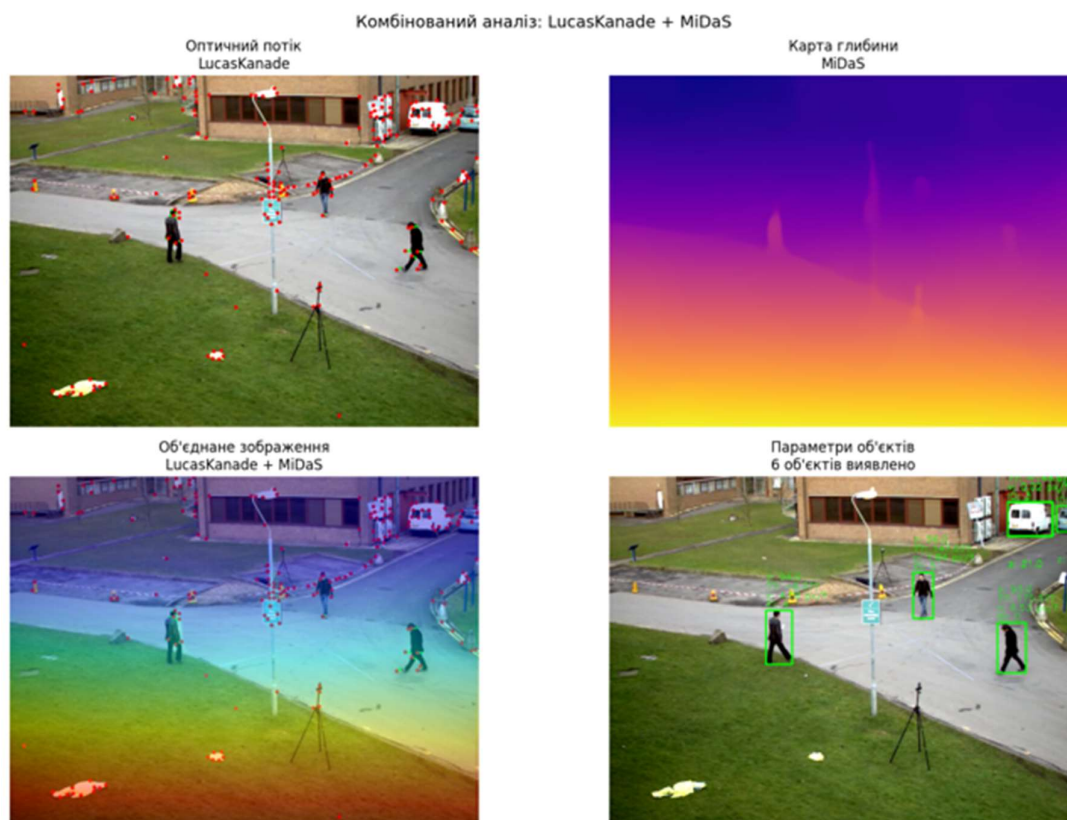


Рисунок Б.28 – LucasKanade\_MiDaS

## Б.5 Результати поєднання методу Horn–Schunck, методів оцінювання глибини і трансформерів

Параметри руху обчислюються за постановкою та ітераційною схемою розд. 2, формули (2.1) – (2.10).

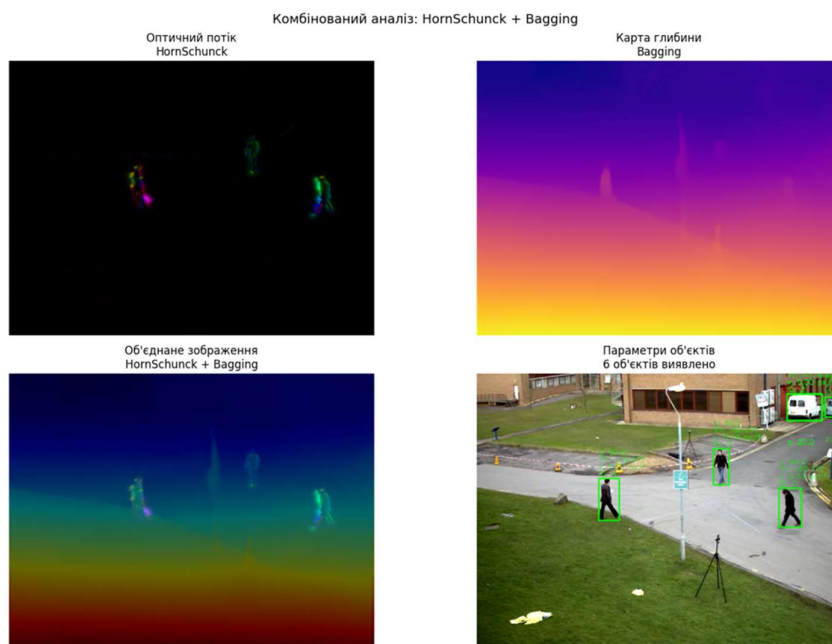


Рисунок Б.29 – HornSchunck\_Bagging

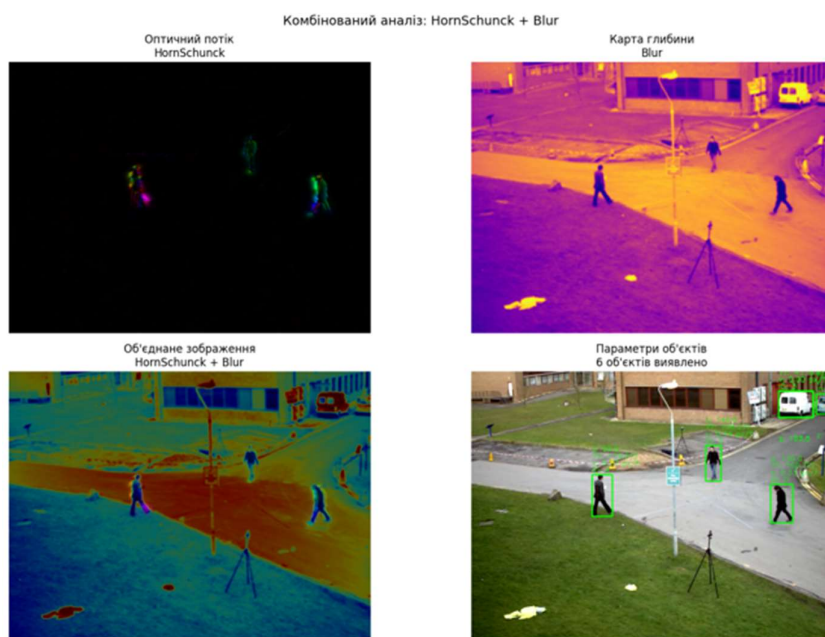


Рисунок Б.30 – HornSchunck\_Blur



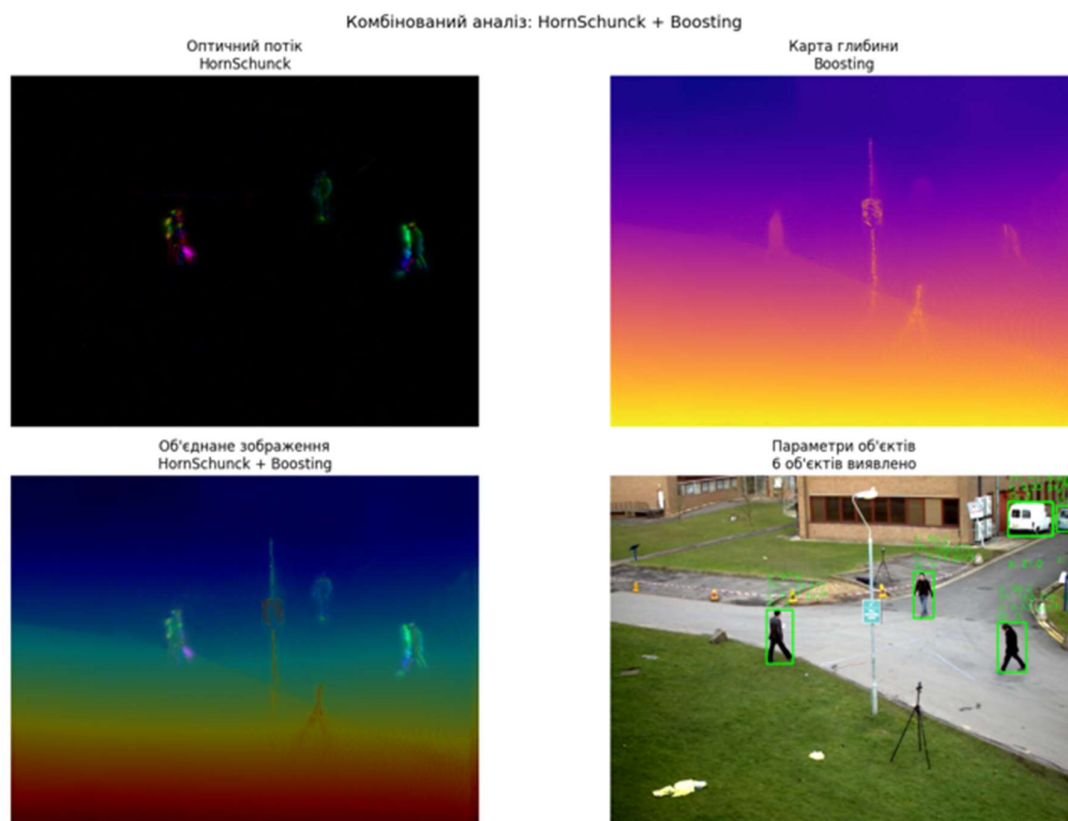


Рисунок Б.31 – HornSchunck\_Boosting

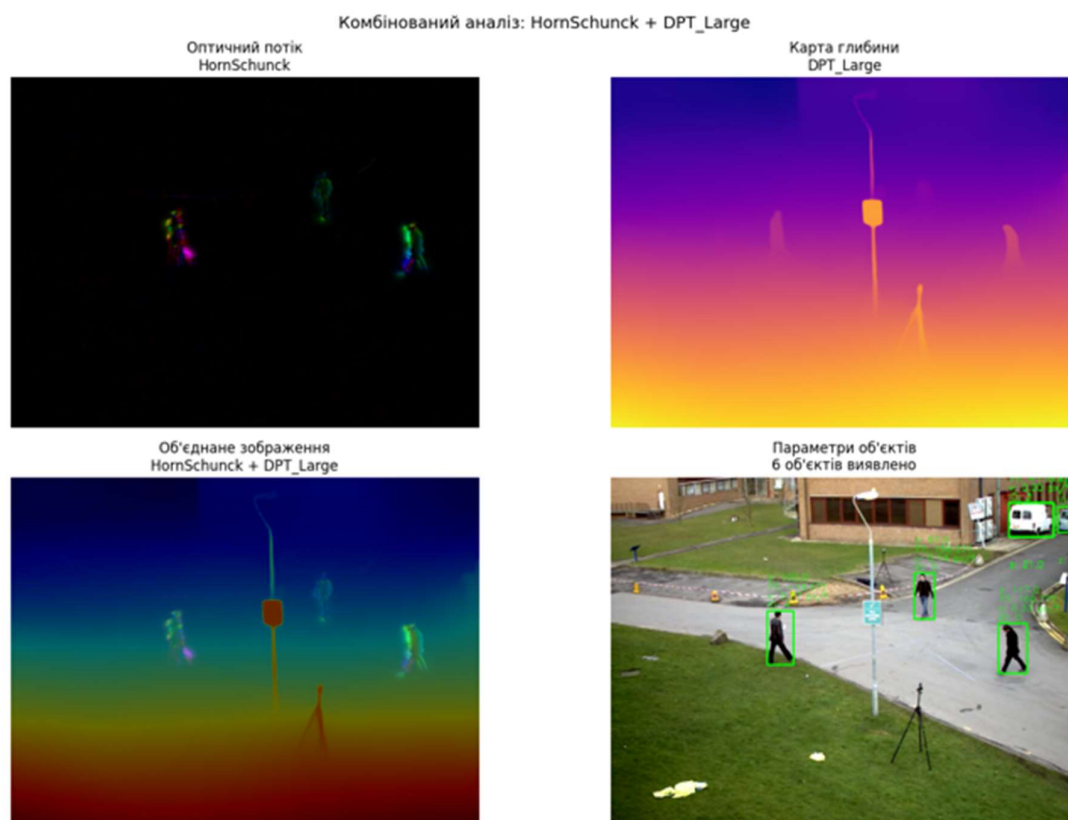


Рисунок Б.32 – HornSchunck\_DPT\_Large

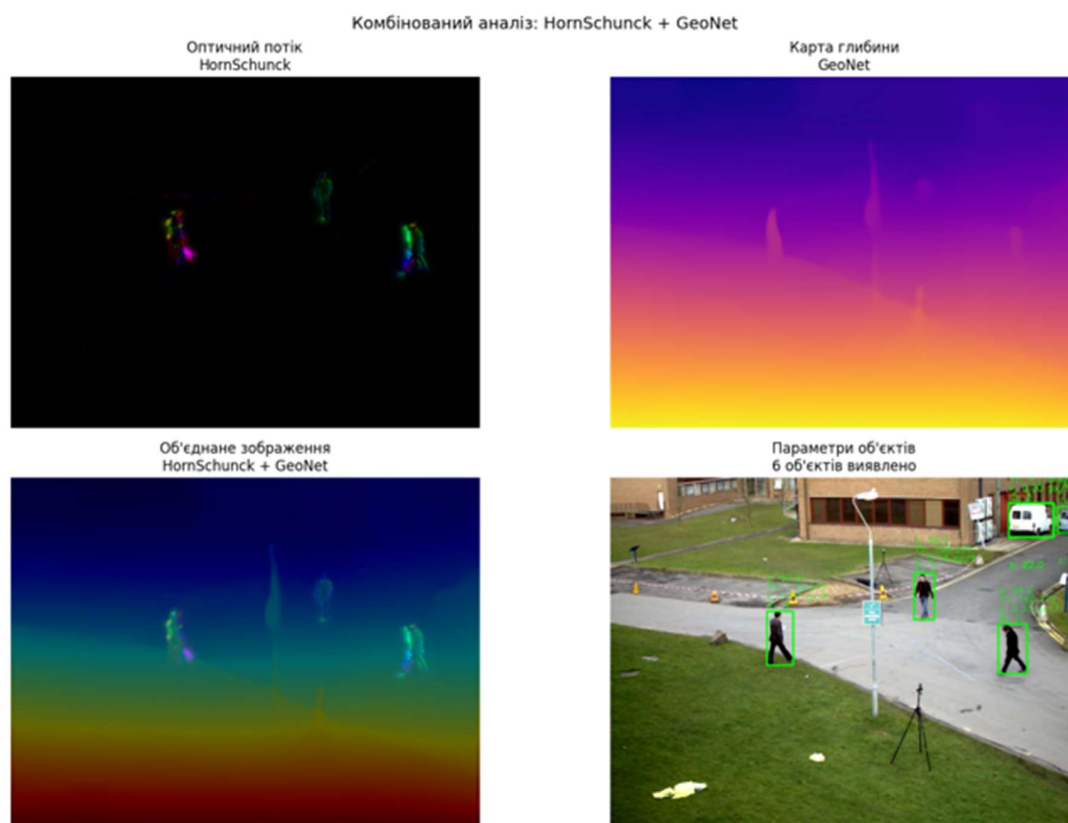


Рисунок Б.33 – HornSchunck\_GeoNet

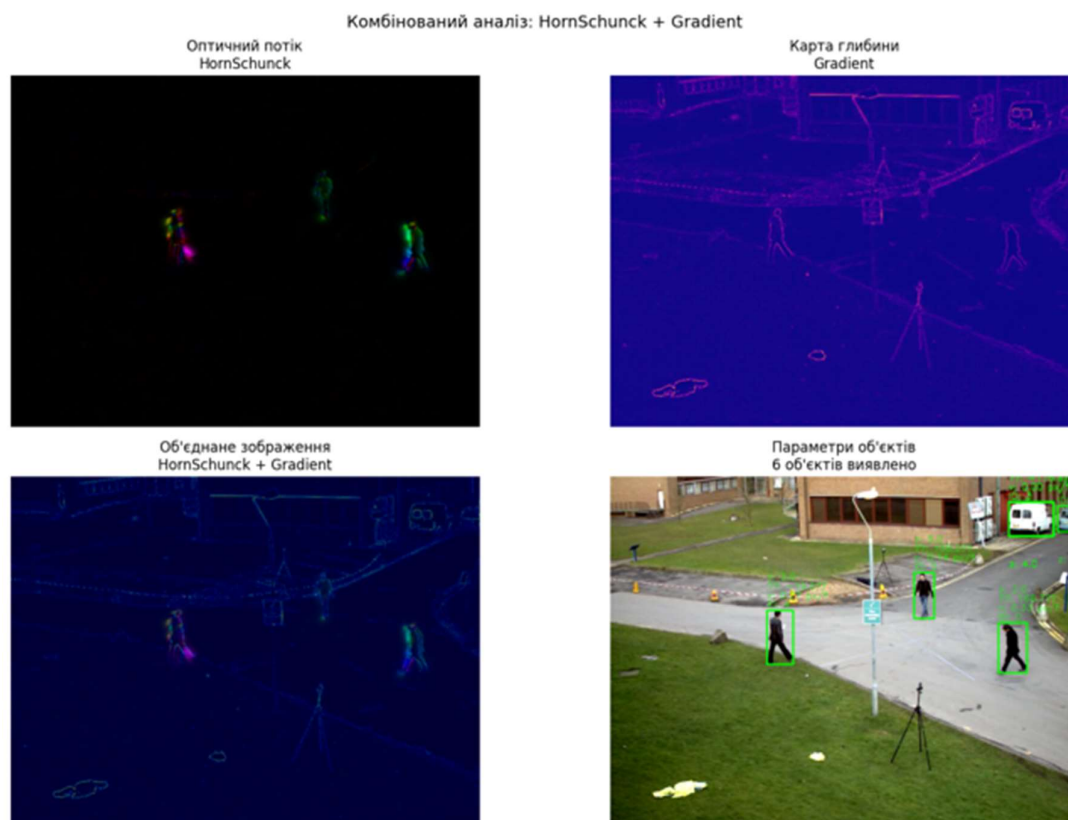


Рисунок Б.34 – HornSchunck\_Gradient



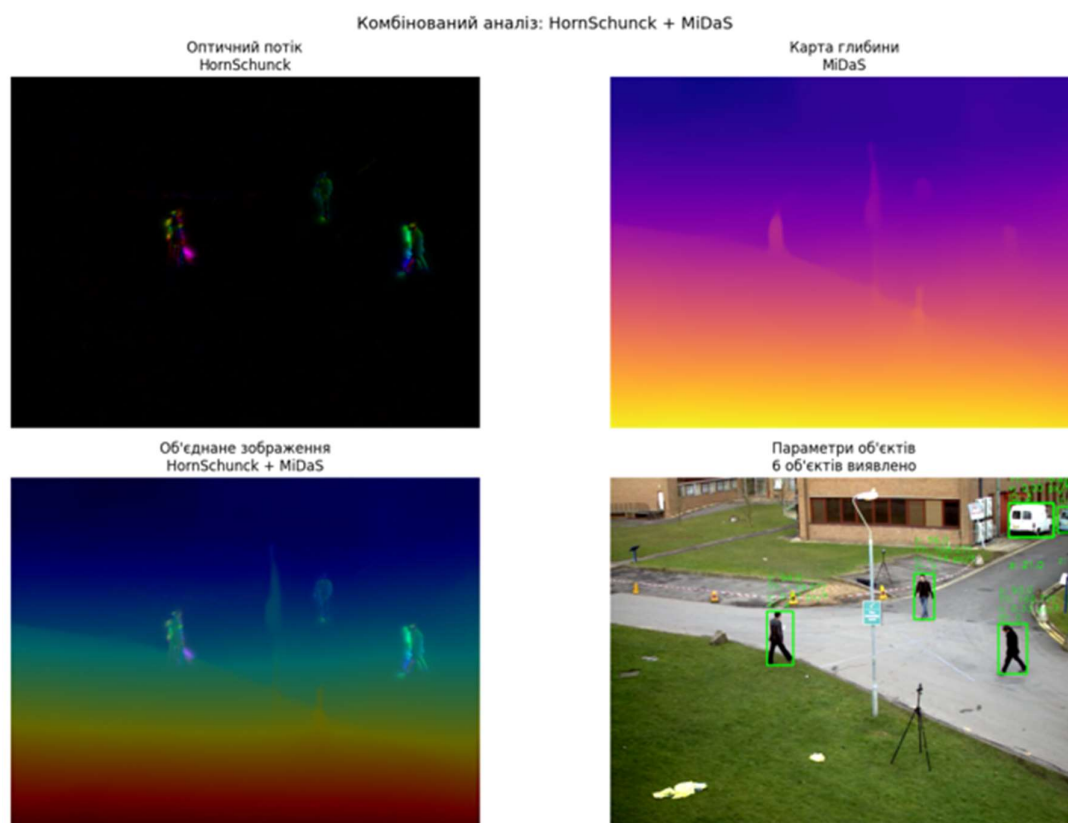


Рисунок Б.35 – HornSchunck\_MiDaS

## ДОДАТОК В

### ДОДАТКОВІ ФОРМУЛИ МЕТОДІВ ТА МОДЕЛЕЙ ІНФОРМАЦІЙНОЇ ТЕХНОЛОГІЇ

Формули для оптичного потоку

Метод Farneback (поліноміальне розширення)

$$f(x) \approx x^T A x + b^T x + c \quad (\text{B.1})$$

$$f_1(x) = x^T A_1 x + b_1^T x + c_1 \quad (\text{B.2})$$

$$f_2(x) = f_1(x - d) \approx (x - d)^T A_1 (x - d) + b_1^T (x - d) + c_1 \quad (\text{B.3})$$

$$A_2 = A_1 \quad (\text{B.4})$$

$$b_2 = b_1 - 2A_1 d \quad (\text{B.5})$$

$$c_2 = d^T A_1 d - b_1^T d + c_1 \quad (\text{B.6})$$

$$d = -\frac{1}{2} A_1^{-1} (b_2 - b_1) \quad (\text{B.7})$$

де:

$f(x)$  – поліноміальне представлення зображення в точці  $x$

$A$  – матриця квадратичних коефіцієнтів

$b$  – вектор лінійних коефіцієнтів

$c$  – вільний член

$d$  – вектор зміщення між кадрами

Метод Лукаса-Канаде (Lucas-Kanade)

$$I_x u + I_y v + I_t = 0 \quad (\text{B.8})$$

$$\begin{bmatrix} I_x(p_1) & I_y(p_1) \\ I_x(p_2) & I_y(p_2) \\ \vdots & \vdots \\ I_x(p_n) & I_y(p_n) \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = - \begin{bmatrix} I_t(p_1) \\ I_t(p_2) \\ \vdots \\ I_t(p_n) \end{bmatrix} \quad (\text{B.9})$$

$$\begin{bmatrix} u \\ v \end{bmatrix} = (A^T A)^{-1} A^T b \quad (\text{B.10})$$

де:

$I_x = \frac{\partial I}{\partial x}$  – просторовий градієнт по осі  $x$

$I_y = \frac{\partial I}{\partial y}$  – просторовий градієнт по осі  $y$

$I_t = \frac{\partial I}{\partial t}$  – часовий градієнт

$p_i$  – пікселі у вікні аналізу

$A$  – матриця градієнтів інтенсивності

$b$  – вектор часових різниць

Метод Хорна-Шунка (Horn-Schunck)

$$E = \iint [(I_x u + I_y v + I_t)^2 + \alpha^2 (\|\nabla u\|^2 + \|\nabla v\|^2)] dx dy \quad (B.11)$$

$$I_x (I_x u + I_y v + I_t) - \alpha^2 \nabla^2 u = 0 \quad (B.12)$$

$$I_y (I_x u + I_y v + I_t) - \alpha^2 \nabla^2 v = 0 \quad (B.13)$$

$$u^{k+1} = \bar{u}^k - \frac{I_x (I_x \bar{u}^k + I_y \bar{v}^k + I_t)}{\alpha^2 + I_x^2 + I_y^2} \quad (B.14)$$

$$v^{k+1} = \bar{v}^k - \frac{I_y (I_x \bar{u}^k + I_y \bar{v}^k + I_t)}{\alpha^2 + I_x^2 + I_y^2} \quad (B.15)$$

де:

$E$  – функціонал, що мінімізується

$\alpha$  – параметр регуляризації (вага гладкості)

$\nabla u = \left( \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y} \right)$  – градієнт поля  $u$

$\nabla^2 u$  – оператор Лапласа

$\bar{u}^k, \bar{v}^k$  – усереднені значення на  $k$ -й ітерації

Метод FlowNet (згорткова нейронна мережа)

$$\Phi_{\text{encoder}}(I_1, I_2) = \text{Conv}_{1 \times 1}(\text{Conv}_{3 \times 3}(\dots \text{Conv}_{7 \times 7}([I_1, I_2]) \dots)) \quad (B.16)$$

$$\hat{F} = \Phi_{\text{decoder}}(\Phi_{\text{encoder}}(I_1, I_2)) \quad (B.17)$$

$$\mathcal{L}_{\text{EPE}} = \frac{1}{N} \sum_{i=1}^N \|F_{\text{pred}}(i) - F_{\text{gt}}(i)\|_2 \quad (B.18)$$

де:

$I_1, I_2$  – два послідовних кадри

$\Phi_{\text{encoder}}$  – згортковий енкодер

$\Phi_{\text{decoder}}$  – декодер

$\hat{F}$  – передбачений оптичний потік

$F_{\text{gt}}$  – істинний оптичний потік

$\mathcal{L}_{\text{EPE}}$  – функція втрат (End-Point Error)

$\|\cdot\|_2$  – евклідова норма

Метод RAFT (Recurrent All-pairs Field Transforms)

$$C_{ijkl} = \frac{1}{\sqrt{HW}} \sum_{h,w} f_{ijhw} \cdot g_{klhw} \quad (\text{B.19})$$

$$x_{t+1} = x_t + \Delta x_t \quad (\text{B.20})$$

$$\Delta x_t, h_{t+1} = \text{GRU}([x_t, C(x_t)], h_t) \quad (\text{B.21})$$

де:

$C_{ijkl}$  – кореляційний об'єм

$H, W$  – висота та ширина зображення

$f_{ijhw}, g_{klhw}$  – ознаки зображень

$x_t$  – поточний стан потоку

$h_t$  – прихований стан GRU

GRU – Gated Recurrent Unit

Формули для оцінки глибини

Метод Blur (розмиття як проксі глибини)

$$I_{\text{gray}}(x, y) = 0.299 \cdot R(x, y) + 0.587 \cdot G(x, y) + 0.114 \cdot B(x, y) \quad (\text{B.22})$$

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (\text{B.23})$$

$$D_{\text{blur}}(x, y) = (I_{\text{gray}} * G)(x, y) = \sum_{i=-k}^k \sum_{j=-k}^k I_{\text{gray}}(x+i, y+j) \cdot G(i, j) \quad (\text{B.24})$$

$$D_{\text{norm}}(x, y) = \frac{D_{\text{blur}}(x, y) - \min(D_{\text{blur}})}{\max(D_{\text{blur}}) - \min(D_{\text{blur}})} \times 255 \quad (\text{B.25})$$

де:

$R, G, B$  – кольорові канали зображення

$I_{\text{gray}}$  – зображення в градаціях сірого

$G(x, y)$  – ядро Гауса

$\sigma$  – стандартне відхилення (параметр розмиття)

$k$  – радіус ядра (зазвичай  $k = 3\sigma$ )

$D_{\text{blur}}$  – карта "глибини", отримана розмиттям

$D_{\text{norm}}$  – нормалізована карта глибини

$*$  – операція згортки

Метод Gradient (градієнтний метод оцінки глибини)

$$\nabla^2 I = \frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2} \quad (\text{B.26})$$

$$\nabla^2 I(x, y) = I(x + 1, y) + I(x - 1, y) + I(x, y + 1) + I(x, y - 1) - 4I(x, y) \quad (\text{B.27})$$

$$L = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad (\text{B.28})$$

$$D_{\text{grad}}(x, y) = |(I_{\text{gray}} * L)(x, y)| \quad (\text{B.29})$$

$$L_{\text{full}} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -8 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad (\text{B.30})$$

$$G_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} * I_{\text{gray}}, G_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} * I_{\text{gray}} \quad (\text{B.31})$$

$$D_{\text{sobel}}(x, y) = \sqrt{G_x(x, y)^2 + G_y(x, y)^2} \quad (\text{B.32})$$

де:

$\nabla^2 I$  – оператор Лапласа

$L$  – ядро Лапласа  $3 \times 3$

$L_{\text{full}}$  – розширене ядро Лапласа (з діагоналями)

$D_{\text{grad}}$  – карта глибини на основі градієнтів

$G_x, G_y$  – градієнти Собеля по осях  $x$  та  $y$

$D_{\text{sobel}}$  – карта глибини на основі оператора Собеля

Метод MiDaS (Monocular Depth Estimation)

$$d_{\text{pred}}(p) = \alpha \cdot d_{\text{true}}(p) + \beta \quad (\text{B.33})$$

$$\mathcal{L} = \frac{1}{N} \sum_{i=1}^N \frac{1}{2\sigma_i^2} \mathcal{L}_i + \log \sigma_i \quad (\text{B.34})$$

$$\mathcal{L}_{SI} = \frac{1}{N} \sum_{i,j} \left[ (\log y_i - \log \hat{y}_i) - (\log y_j - \log \hat{y}_j) \right]^2 \quad (\text{B.35})$$

де:

$d_{\text{pred}}(p)$  – передбачена глибина в точці  $p$

$d_{\text{true}}(p)$  – істинна глибина

$\alpha, \beta$  – параметри масштабу та зсуву

$\mathcal{L}$  – багатомасштабна функція втрат

$\mathcal{L}_{SI}$  – інваріантна до масштабу втрата

$\sigma_i$  – параметр масштабу для  $i$ -го рівня

$y_i$  – істинне значення глибини

$\hat{y}_i$  – передбачене значення глибини

Метод DPT (Dense Prediction Transformer)

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (\text{B.36})$$

$$z_0 = [x_{\text{class}}; x_p^1 E; x_p^2 E; \dots; x_p^N E] + E_{\text{pos}} \quad (\text{B.37})$$

$$z_l = \text{MSA}(\text{LN}(z_{l-1})) + z_{l-1}, l = 1, \dots, L \quad (\text{B.38})$$

де:

$Q, K, V$  – матриці запитів, ключів, значень

$d_k$  – розмірність простору ключів

Attention – функція уваги

$z_0$  – вхідні токени

$x_{\text{class}}$  – спеціальний токен класифікації

$x_p^i$  –  $i$ -й патч зображення

$E$  – матриця проєкції патчів

$E_{\text{pos}}$  – позиційне кодування

MSA – багатоголова увага

LN – шарова нормалізація

Метод GeoNet (сумісна оцінка глибини та потоку)

$$I_2(p) = I_1(p + f(p)) \quad (\text{B.39})$$

$$f(p) = T(p) \cdot \frac{Z(p) - Z_0}{Z(p)} \cdot \text{proj}(p) \quad (\text{B.40})$$

де:

$I_1, I_2$  – два послідовних кадри

$f(p)$  – оптичний потік у точці  $p$

$Z(p)$  – глибина в точці  $p$

$T(p)$  – матриця перетворення камери

$Z_0$  – опорна глибина

$\text{proj}(p)$  – проекція точки  $p$

Формули для параметрів динамічних об'єктів

Швидкість об'єкта

$$v = \frac{1}{N_\Omega} \sum_{(x,y) \in \Omega} \|F(x, y)\| \quad (\text{B.41})$$

$$v = \frac{1}{|\Omega|} \iint_{\Omega} \sqrt{u(x, y)^2 + v(x, y)^2} dx dy \quad (\text{B.42})$$

де:

$v$  – швидкість об'єкта (пікселів/кадр)

$\Omega$  – область, що відповідає об'єкту

$N_\Omega = |\Omega|$  – кількість пікселів в області

$F(x, y) = [u(x, y), v(x, y)]^T$  – вектор оптичного потоку

$\|\cdot\|$  – евклідова норма

Напрямок руху

$$\theta(x, y) = \arctan\left(\frac{v(x, y)}{u(x, y)}\right) \quad (\text{B.43})$$

$$\theta = \arctan\left(\frac{\sum_{(x,y) \in \Omega} \sin \theta(x, y)}{\sum_{(x,y) \in \Omega} \cos \theta(x, y)}\right) \quad (\text{B.44})$$

$$\theta = \frac{1}{N_\Omega} \sum_{(x,y) \in \Omega} \arctan\left(\frac{v(x, y)}{u(x, y)}\right) \quad (\text{B.45})$$

де:

$\theta$  – напрямок руху об'єкта (радіани)

$\theta(x, y)$  – напрямок руху в точці  $(x, y)$

$u(x, y), v(x, y)$  – компоненти оптичного потоку

Просторове положення

$$x_c = \frac{x_{\min} + x_{\max}}{2} \quad (\text{B.46})$$

$$y_c = \frac{y_{\min} + y_{\max}}{2} \quad (\text{B.47})$$

$$z = \text{median}\{D(x, y) \mid (x, y) \in \Omega\} \quad (\text{B.48})$$

$$z = \frac{1}{|\Omega|} \iint_{\Omega} D(x, y) dx dy \quad (\text{B.49})$$

де:

$(x_c, y_c)$  – координати центру об'єкта

$x_{\min}, x_{\max}, y_{\min}, y_{\max}$  – границі bounding box

$z$  – глибина (Z-координата) об'єкта

$D(x, y)$  – карта глибини

$\Omega$  – область об'єкта

Траєкторія об'єкта

$$T(t) = \{p(t_1), p(t_2), \dots, p(t_n)\} \quad (\text{B.50})$$

$$\tilde{p}(t_i) = \frac{1}{2k+1} \sum_{j=-k}^k p(t_{i+j}) \quad (\text{B.51})$$

де:

$T(t)$  – траєкторія об'єкта

$p(t_i) = (x_i, y_i)$  – положення об'єкта в момент  $t_i$

$\tilde{p}(t_i)$  – згладжена траєкторія

$k$  – радіус вікна згладжування

Прискорення об'єкта

$$a = \frac{v(t+1) - v(t)}{\Delta t} \quad (\text{B.52})$$

де:

$a$  – прискорення об'єкта

$v(t)$  – швидкість в момент  $t$

$\Delta t$  – часовий інтервал ( $\Delta t = 1$  кадр)

Формули для методів ансамблю

Бутстреп-агрегація (Bagging)

$$\hat{y}_{\text{bag}}(x) = \frac{1}{M} \sum_{i=1}^M f_i(x) \quad (\text{B.53})$$

де:

$\hat{y}_{\text{bag}}(x)$  – ансамблевий прогноз

$M$  – кількість моделей в ансамблі

$f_i(x)$  – прогноз  $i$ -ї моделі

Бустинг (Boosting)



$$\hat{y}_{\text{boost}}(x) = \sum_{i=1}^M \alpha_i f_i(x) \quad (\text{B.54})$$

$$\alpha_i = \frac{1}{2} \ln \left( \frac{1-\epsilon_i}{\epsilon_i} \right) \quad (\text{B.55})$$

$$D_{\text{final}}(x) = \begin{cases} D_1(x), & \text{якщо } |\nabla D_1(x)| \leq \tau \\ D_2(x), & \text{якщо } |\nabla D_1(x)| > \tau \end{cases} \quad (\text{B.56})$$

де:

$\hat{y}_{\text{boost}}(x)$  – бустинг-прогноз

$\alpha_i$  – вага  $i$ -ї моделі

$\epsilon_i$  – помилка  $i$ -ї моделі

$D_{\text{final}}$  – фінальна карта глибини

$D_1, D_2$  – карти глибини від різних моделей

$\nabla D_1$  – градієнт карти глибини

$\tau$  – поріг для маски помилок

Формули для DETR (Detection Transformer)

Архітектура трансформера

$$\text{Attention}(Q, K, V) = \text{softmax} \left( \frac{QK^T}{\sqrt{d_k}} \right) V \quad (\text{B.57})$$

$$\text{MultiHead}(Q, K, V) = \text{Concat}(\text{head}_1, \dots, \text{head}_h) W^O \quad (\text{B.58})$$

$$\text{Head}_i = \text{Attention}(QW_i^Q, KW_i^K, VW_i^V) \quad (\text{B.59})$$

де:

$\text{Attention}$  – функція уваги

$Q, K, V$  – матриці запитів, ключів, значень

$d_k$  – розмірність простору ключів

$\text{MultiHead}$  – багатоголова увага

$h$  – кількість голів уваги

$W_i^Q, W_i^K, W_i^V, W^O$  – матриці проекцій

$\text{Concat}$  – операція конкатенації

Функція втрат DETR

$$\hat{\sigma} = \arg \min_{\sigma \in \mathfrak{S}_N} \sum_{i=1}^N \mathcal{L}_{\text{match}}(y_i, \hat{y}_{\sigma(i)}) \quad (\text{B.60})$$

$$\mathcal{L}_{\text{Hungarian}}(y, \hat{y}) = \sum_{i=1}^N \left[ -\log \hat{p}_{\hat{\sigma}(i)}(c_i) + \mathbf{1}_{\{c_i \neq \emptyset\}} \mathcal{L}_{\text{box}}(b_i, \hat{b}_{\hat{\sigma}(i)}) \right] \quad (\text{B.61})$$

$$\mathcal{L}_{\text{box}}(b_i, \hat{b}_{\hat{\sigma}(i)}) = \lambda_{\text{IoU}} \mathcal{L}_{\text{IoU}}(b_i, \hat{b}_{\hat{\sigma}(i)}) + \lambda_{L1} \|b_i - \hat{b}_{\hat{\sigma}(i)}\|_1 \quad (\text{B.62})$$

де:

$\hat{\sigma}$  – оптимальне парування (угорський алгоритм)

$\mathfrak{S}_N$  – множина всіх перестановок

$\mathcal{L}_{\text{match}}$  – функція відповідності

$\mathcal{L}_{\text{Hungarian}}$  – угорська втрата

$\hat{p}_{\hat{\sigma}(i)}(c_i)$  – ймовірність класу  $c_i$

$\mathbf{1}_{\{c_i \neq \emptyset\}}$  – індикаторна функція

$\mathcal{L}_{\text{box}}$  – втрата для bounding box

$b_i, \hat{b}_{\hat{\sigma}(i)}$  – істинний та передбачений bounding box

$\lambda_{\text{IoU}}, \lambda_{L1}$  – вагові коефіцієнти

$\|\cdot\|_1$  – L1-норма

Формули для метрик якості

EPE (End-Point Error)

$$\text{EPE} = \frac{1}{N} \sum_{i=1}^N \sqrt{(u_i - \hat{u}_i)^2 + (v_i - \hat{v}_i)^2} \quad (\text{B.63})$$

AAE (Angular Average Error)

$$\text{AAE} = \frac{1}{N} \sum_{i=1}^N \arccos \left( \frac{u_i \hat{u}_i + v_i \hat{v}_i + 1}{\sqrt{u_i^2 + v_i^2 + 1} \cdot \sqrt{\hat{u}_i^2 + \hat{v}_i^2 + 1}} \right) \quad (\text{B.64})$$

RMSE (Root Mean Square Error)

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (d_i - \hat{d}_i)^2} \quad (\text{B.65})$$

MAE (Mean Absolute Error)

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |d_i - \hat{d}_i| \quad (\text{B.66})$$

PSNR (Peak Signal-to-Noise Ratio)

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N (d_i - \hat{d}_i)^2 \quad (\text{B.67})$$

$$\text{PSNR} = 10 \cdot \log_{10} \left( \frac{\text{MAX}^2}{\text{MSE}} \right) \quad (\text{B.68})$$

IoU (Intersection over Union)

$$\text{IoU} = \frac{|A \cap B|}{|A \cup B|} = \frac{\sum_{i=1}^N \mathbf{1}_{\{A_i=1 \wedge B_i=1\}}}{\sum_{i=1}^N \mathbf{1}_{\{A_i=1 \vee B_i=1\}}} \quad (\text{B.69})$$

DICE коефіцієнт

$$\text{DICE} = \frac{2|A \cap B|}{|A| + |B|} = \frac{2 \cdot \text{TP}}{2 \cdot \text{TP} + \text{FP} + \text{FN}} \quad (\text{B.70})$$

де:

$u_i, v_i$  – істинні компоненти оптичного потоку

$\hat{u}_i, \hat{v}_i$  – передбачені компоненти оптичного потоку

$d_i$  – істинна глибина

$\hat{d}_i$  – передбачена глибина

$N$  – кількість пікселів

MAX – максимальне значення пікселя (255)

$A, B$  – бінарні маски

$\mathbf{1}$  – індикаторна функція

TP, FP, FN – істинно-позитивні, хибно-позитивні, хибно-негативні

Формули для метрик виявлення

Точність (Precision)

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{F}} \quad (\text{B.71})$$

Повнота (Recall)

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (\text{B.72})$$

F1-міра

$$F_1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (\text{B.73})$$

AUC-ROC (Area Under ROC Curve)

$$\text{AUC} = \int_0^1 \text{TPR}(\text{FPR}) d(\text{FPR}) \quad (\text{B.74})$$

де:

TP – True Positive (істинно-позитивні)

TN – True Negative (істинно-негативні)

FP – False Positive (хибно-позитивні)

FN – False Negative (хибно-негативні)

$TPR = \frac{TP}{TP+FN}$  – True Positive Rate

$FPR = \frac{FP}{FP+TN}$  – False Positive Rate

Формули для нормалізації

Min-Max нормалізація

$$D_{\text{norm}}(x, y) = \frac{D(x, y) - \min(D)}{\max(D) - \min(D)} \times 255 \quad (\text{B.75})$$

Z-нормалізація (стандартизація)

$$\mu_D = \frac{1}{N} \sum_{i=1}^N D(x_i, y_i) \quad (\text{B.76})$$

$$\sigma_D = \sqrt{\frac{1}{N} \sum_{i=1}^N (D(x_i, y_i) - \mu_D)^2} \quad (\text{B.77})$$

$$D_{\text{std}}(x, y) = \frac{D(x, y) - \mu_D}{\sigma_D} \quad (\text{B.78})$$

де:

$D_{\text{norm}}$  – нормалізована карта глибини

$D_{\text{std}}$  – стандартизована карта глибини

$\mu_D$  – середнє значення карти глибини

$\sigma_D$  – стандартне відхилення карти глибини

$\min(D), \max(D)$  – мінімальне та максимальне значення

$N$  – кількість пікселів

Формули для візуалізації

Перетворення оптичного потоку в RGB

$$h = \frac{\arctan(v, u)}{2\pi} \times 180^\circ \quad (\text{B.79})$$

$$s = 255 \quad (\text{B.80})$$

$$v_{\text{bright}} = \min\left(255, \frac{\sqrt{u^2 + v^2}}{\max(\sqrt{u^2 + v^2})} \times 255\right) \quad (\text{B.81})$$

де:

$h$  – відтінок (Hue)

$s$  – насиченість (Saturation)

$v_{\text{bright}}$  – яскравість (Value)

$\arctan(v, u)$  – арктангенс з урахуванням знаків

Колірна карта глибини (Jet colormap)

$$R(z) = \begin{cases} 0, & z < 0.2 \\ 4z - 0.8, & 0.2 \leq z < 0.45 \\ 1, & 0.45 \leq z < 0.75 \\ -4z + 4, & 0.75 \leq z < 1 \\ 0, & z \geq 1 \end{cases} \quad (\text{B.82})$$

$$G(z) = \begin{cases} 0, & z < 0.2 \\ 4z - 0.8, & 0.2 \leq z < 0.35 \\ 1, & 0.35 \leq z < 0.65 \\ -4z + 3.6, & 0.65 \leq z < 0.9 \\ 0, & z \geq 0.9 \end{cases} \quad (\text{B.83})$$

$$B(z) = \begin{cases} 4z + 0.2, & z < 0.2 \\ 1, & 0.2 \leq z < 0.45 \\ -4z + 2.8, & 0.45 \leq z < 0.7 \\ 0, & z \geq 0.7 \end{cases} \quad (\text{B.84})$$

Колірна карта Plasma для глибини

$$\text{Plasma}(z) = \text{interpolate}([0, 0.2, 0.4, 0.6, 0.8, 1], [\text{col}_0, \text{col}_1, \text{col}_2, \text{col}_3, \text{col}_4, \text{col}_5]) \quad (\text{B.85})$$

де:

$R(z), G(z), B(z)$  – компоненти кольору Jet colormap

$z$  – нормалізована глибина ( $0 \leq z \leq 1$ )

$\text{Plasma}(z)$  – колірна карта Plasma

$\text{interpolate}$  – функція лінійної інтерполяції між опорними кольорами

$\text{col}_i$  – опорні кольори

## ДОДАТОК Г

### МАТЕРІАЛИ ЩОДО ВПРОВАДЖЕННЯ РЕЗУЛЬТАТІВ

ЗАТВЕРДЖУЮ

Ректор

ТОВ «ТЕХНІЧНИЙ УНІВЕРСИТЕТ

«МЕТІНВЕСТ ПОЛІТЕХНІКА»



Олександр ПОВАЖНИЙ

15 грудня 2025 р.

## АКТ

про використання результатів дисертаційної роботи  
**КОНДРАТОВА Олексія Михайловича**  
 в навчальному процесі  
 кафедри цифрових технологій та проєктно-аналітичних рішень

Ми, що нижче підписалися, завідувачка кафедри цифрових технологій та проєктно-аналітичних рішень СМІРНОВА Ірина Іванівна, професор кафедри цифрових технологій та проєктно-аналітичних рішень САГАЙДА Павло Іванович, склали акт про те, що результати дисертаційної роботи КОНДРАТОВА Олексія Михайловича впроваджені в навчальний процес на кафедрі цифрових технологій та проєктно-аналітичних рішень.

Моделі та методи обчислювального інтелекту для моделювання, ідентифікації та оптимізації інформаційних управляючих систем впроваджено в дисципліну «Продуктивність офісних систем».

Технології дистанційної ідентифікації та моніторингу динамічних об'єктів на основі комп'ютерного зору для задач безпеки впроваджено в дисципліну «Основи кібербезпеки та захисту інформації».

Модель ідентифікації параметрів об'єкту з використанням DETection TRansformer (DETR) та Optical Flow (Оптичний потік) з графічною візуалізацією результатів впроваджено в дисципліну «Розробка GUI».

Методи обчислювального інтелекту та ансамблеві моделі для ідентифікації, прогнозування та підтримки прийняття рішень у системах інформаційної безпеки впроваджено в дисципліну «Технології захисту даних та інформаційної безпеки».

Завідувачка кафедри  
цифрових технологій та  
проєктно-аналітичних рішень

Ірина СМІРНОВА

Професор кафедри  
цифрових технологій та  
проєктно-аналітичних рішень

Павло САГАЙДА







ЗАТВЕРДЖУЮ

Проректор

Національного технічного  
університету «Харківський  
політехнічний інститут»

Руслан МИГУЩЕНКО

12 2025 р.

## АКТ

про використання результатів дисертаційної роботи  
**КОНДРАТОВА Олексія Михайловича**  
 в навчальному процесі кафедри інформаційних систем та технологій  
 Національного технічного університету «ХПІ»

Ми, що нижче підписалися, завідувачка кафедри інформаційних систем та технологій **НІКУЛІНА Олена Миколаївна**, професор кафедри інформаційних систем та технологій **МОСКАЛЕНКО Валентина Володимирівна**, доцент кафедри інформаційних систем та технологій **ХАЦЬКО Наталія Євгеніївна** склали акт про те, що результати дисертаційної роботи **КОНДРАТОВА Олексія Михайловича** впроваджені в навчальний процес на кафедрі інформаційних систем та технологій.

Модель інформаційних управляючих систем на основі обчислювального інтелекту дистанційної ідентифікації параметрів об'єктів впроваджено в дисципліну «Математичне моделювання та аналіз систем».

Технологія відстеження руху об'єктів впроваджено в дисципліну «Дослідження операцій».

Модель ідентифікації параметрів об'єкту з використанням DEtection TRansformer (DETR) та Optical Flow (Оптичний потік) впроваджено в дисципліну «Інтелектуальні системи».

Модель обчислювального інтелекту для моделювання, ідентифікації, прогнозування та класифікації впроваджено в дисципліну «Інтелектуальний аналіз даних».

Завідувачка кафедри  
 інформаційних систем та  
 технологій НТУ «ХПІ»

Олена НІКУЛІНА

Професор кафедри  
 інформаційних систем та  
 технологій НТУ «ХПІ»

Валентина МОСКАЛЕНКО

Доцент кафедри  
 інформаційних систем та  
 технологій НТУ «ХПІ»

Наталія ХАЦЬКО





ЗАТВЕРДЖУЮ

Директор радіоастрономічного інституту

НАН України

академік

В. В. Захаренко

2025 р.

## АКТ

про використання в радіоастрономічному інституті НАН України результатів дисертаційної роботи Кондратова Олексія Михайловича, виконаної на здобуття доктора філософії

Комісія у складі: голови комісії к.ф.-м.н. Бубнова І. М., провідного наукового співробітника радіоастрономічного інституту НАН України, та членів комісії – к.ф.-м.н. Шевчука М. В. та к.ф.-м.н. Антоненко Ю. В. склала цей акт щодо використання в науково-дослідній роботі радіоастрономічного інституту НАН України наступних результатів дисертаційної роботи Кондратова О. М. щодо «Моделі, методи та інформаційна технологія дистанційної ідентифікації параметрів динамічних об'єктів», у частині дослідження та аналізу динамічних процесів у сонячній атмосфері за даними дистанційних радіоспостережень, зокрема для виявлення та ідентифікації сонячних радіо сплесків. У процесі науково-дослідних робіт було використано такі результати дисертаційної роботи:

– моделі дистанційної ідентифікації параметрів динамічних об'єктів, які базуються на поєднанні методів комп'ютерного зору, оптичного потоку та трансформерів виявлення (DETR, Transformer-based models), що дозволяють розглядати сонячні радіо сплески як динамічні об'єкти з просторово-часовими ознаками та виконувати їх автоматизоване виявлення на спектрограмах радіовипромінювання;

– методи ідентифікації параметрів динамічних процесів, з використанням критеріїв оптимальності та ансамблевих методів обчислювального інтелекту, застосовані для оцінювання параметрів розвитку сонячних сплесків (швидкість дрейфу частоти, тривалість, інтенсивність) за експериментальними даними радіоастрономічних спостережень;

– інформаційна технологія дистанційної ідентифікації, що поєднує блоки попередньої обробки даних, виявлення сплесків, оптимізаційної ідентифікації параметрів та аналізу результатів, яка використовується для дослідження нестационарних і швидкоплинних процесів у сонячній плазмі;

– програмна реалізація методів і моделей, що ґрунтується на використанні неймережових архітектур (FlowNet, Optical Flow, YOLO, DETR, GeoNet) та дозволяє виконувати автоматизований пошук сонячних сплесків різних типів у режимі після обробки та експериментального аналізу.

Використання зазначених результатів дисертаційної роботи Кондратова О. М. дозволило скоротити час обробки великих обсягів радіоастрономічних даних, підвищити точність та надійність виявлення сонячних радіо сплесків як динамічних об'єктів, розширити можливості моделювання та дистанційної ідентифікації параметрів нестационарних процесів у сонячній атмосфері, а також забезпечити основу для подальшого розвитку інтелектуальних систем аналізу даних сонячної активності.

Провідний  
науковий співробітник,  
к.ф.-м.н., с.н.с.

 І. М. Бубнов

Провідний  
науковий співробітник,  
к.ф.-м.н., с.н.с.

 М. В. Шевчук

Вчений секретар,  
к.ф.-м.н.

 Ю. В. Антоненко