УДК 004.76

Г.О. ЩУР, І.Б. ДУМИН
Національний університет «Львівська політехніка»

ПІДВИЩЕННЯ ВИЯВЛЕННЯ НЕБЕЗПЕЧНИХ ОБ'ЄКТІВ НА ОСНОВІ ВИКОРИСТАННЯ ПРОМІЖНОГО ЗЛИТТЯ МУЛЬТИМОДАЛЬНИХ ДАНИХ

Надійне виявлення небезпечних об'єктів є критично важливим компонентом у багатьох сферах, зокрема в автономному транспорті, промисловій безпеці, моніторингових системах і робототехніці. У реальних умовах – за поганого освітлення, у складних погодних умовах, за часткового перекриття об'єктів або візуального шуму – системи, які базуються на одному типі сенсора, часто демонструють нестабільну роботу або хибні спрацьовування. З метою подолання цих обмежень у роботі запропоновано підхід, що базується на злитті мультимодальних даних із використанням механізмів уваги й оцінки критичності.

Запропонована система об'єднує інформацію з RGB-камери, LiDAR, тепловізора та (опційно) тактильних сенсорів, що дозволяє формувати цілісну картину навколишнього середовища. Кожен сенсорний потік проходить через власний енкодер ознак, після чого активується механізм уваги, який зважує вклад кожного джерела даних залежно від ситуації. Об'єднаний вектор ознак проходить через детектор, що формує ймовірність присутності об'єкта, а паралельно функціонує модуль оцінювання критичності, який аналізує просторову відстань, швидкість і орієнтацію виявленого об'єкта.

На основі зваженої комбінації обох оцінок – детекції та критичності – формується остаточне рішення про наявність загрози. Якщо комбіноване значення перевищує встановлений поріг, система спрацьовує: генерує тривогу, активує аварійне гальмування або передає сигнал оператору. У статті також наводиться гіпотетичний сценарій автономного водіння, у якому детально демонструється ефективність запропонованої архітектури за виявленні пішохода в умовах обмеженої видимості.

Система демонструє високу точність, гнучкість і адаптивність до змін середовища. Її перевагами є збереження просторових і семантичних ознак, можливість аналізу об'єктів у різних спектральних діапазонах, а також пріоритет загроз залежно від ступеня небезпеки. У перспективі дана архітектура може стати основою для реалізації безпечних автономних рішень у транспорті, промисловості та міській інфраструктурі.

Ключові слова: *злиття мультимодальних сенсорних даних, виявлення об'єктів, автономні системи, LiDAR, тепловізор, механізм уваги, критичність.*

H.O. SHCHUR, I.B. DUMYN
Lviv Polytechnic National University

ENHANCING DANGEROUS OBJECT DETECTION THROUGH INTERMEDIATE MULTIMODAL DATA FUSION

Reliable detection of dangerous objects is a critical component in numerous domains, including autonomous transportation, industrial safety, surveillance systems, and robotics. In real-world scenarios – such as poor lighting, adverse weather, occluded views, or noisy backgrounds – single-sensor detection systems often struggle to maintain accuracy and stability. To address these limitations, this paper proposes an approach based on multimodal data fusion, leveraging attention mechanisms and object criticality assessment.

The proposed system integrates information from an RGB camera, LiDAR, thermal imager, and optionally, tactile sensors. Each sensor stream is processed by a dedicated feature encoder. Attention weights are dynamically calculated based on the environmental context to determine the relevance of each modality. The resulting fused feature vector is passed through a detection module that produces a base confidence score, while a parallel criticality module evaluates spatial distance, relative velocity, and orientation of the detected object.

The final decision is formed by combining both the detection score and the criticality assessment using a weighted factor. If the resulting Sfinal value exceeds a predefined threshold, the system triggers a response – issuing a warning, initiating emergency braking, or notifying the operator. A hypothetical case study involving pedestrian detection in limited visibility illustrates the practical utility of the proposed architecture.

The system exhibits high accuracy, robustness, and contextual adaptability. Its main advantages include preservation of spatial and semantic features, the ability to process inputs across multiple sensing domains, and risk-based prioritization of threats. This architecture presents a strong foundation for the development of reliable, safety-oriented autonomous solutions in transportation, industrial monitoring, and urban environments.

Key words: *multimodal sensor data fusion, object detection, autonomous systems, LiDAR, attention mechanisms, feature extraction, risk assessment.*

https://doi.org/10.32782/mathematical-modelling/2025-8-1-27

Постановка проблеми

Існуючі методи виявлення, що базуються на окремих сенсорах (наприклад, RGB-камерах), мають обмежену надійність у реальних умовах. Основні труднощі виникають за поганого освітлення (ніч, тінь, затемнення), несприятливих погодних умов (дощ, туман, сніг), часткового або цілковитого перекриття об'єктів іншими об'єктами (наприклад, коли пішохід частково закритий деревом або автомобілем), складного фону (наприклад, густе міське середовище або промислові об'єкти з великою кількістю візуальних деталей), браку інформації про просторову глибину або температуру об'єктів.

Усе це знижує ефективність систем безпеки та створює ризики для автономного керування або моніторингу. У реальному світі жоден сенсор не є універсальним: камери надають кольорову картинку, але не бачать у темряві; LiDAR створює тривимірну карту, але не відображає теплову інформацію; тепловізори бачать об'єкти за температурою, але не дають чітких контурів ідентифікації.

Рішенням є мультимодальне злиття – об'єднання даних із кількох сенсорів, яке дозволяє компенсувати слабкості одного джерела інформації перевагами іншого. Нещодавні досягнення доводять ефективність такого підходу.

Проте об'єднання різних типів сенсорних даних потребує спеціальних підходів до синхронізації, вирівнювання, обробки та представлення ознак. Також актуальним є питання пріоритезації важливіших сигналів, особливо у критичних ситуаціях. Саме тому запропоновано модель проміжного злиття, що поєднує дані з кількох сенсорів і використовує механізм уваги для динамічного зважування інформації залежно від контексту. Такий підхід дає змогу досягти високої точності й надійності у виявленні небезпечних об'єктів у складних умовах навколишнього середовища.

Традиційно такі системи базувались на використанні одного типу сенсорів – зазвичай візуальних (RGB-камери), які надають детальні зображення навколишнього середовища. Проте ефективність цих підходів суттєво знижується у складних умовах: за слабкого освітлення, туману, дощу, а також у разі часткового перекриття об'єктів або складного фону. У таких ситуаціях важливо мати систему, яка здатна компенсувати слабкі сторони одного сенсора за допомогою іншого, що можливо лише за умови використання мультимодального підходу.

Мультимодальні системи інтегрують інформацію з кількох джерел, як-от RGB-зображення, лазерні сканери (LiDAR), тепловізори, а в деяких випадках – також тактильні сенсори або терагерцові сканери. Наприклад, камера забезпечує високоякісне зображення, тоді як LiDAR додає інформацію про глибину, а тепловізор допомагає ідентифікувати об'єкти вночі або крізь дим. Завдяки такій синергії система отримує повніше уявлення про навколишнє середовище.

У даній роботі запропоновано уніфіковану модель, що поєднує кілька сенсорних потоків за допомогою проміжного злиття, доповнену механізмом уваги. Такий підхід дозволяє не просто об'єднати дані, а й адаптивно зважувати внесок кожного сенсора залежно від контексту, що особливо важливо в динамічних, змінних середовищах.

Аналіз останніх досліджень і публікацій

У роботі [1] наведено традиційні системи виявлення об'єктів, які зазвичай покладаються лише на один тип датчика та часто мають проблеми за складних умов, як-от погане освітлення, оклюзії або несприятлива погода. Ці недоліки зумовили зростання інтересу до злиття мультимодальних датчиків [2]. Автори [3] доводять актуальність глобальної проблеми виявлення небезпечних об'єктів. Сучасні платформи моделювання дозволяють створювати реалістичні середовища, де можна відпрацьовувати сценарії розмінування, моделювати різноманітні загрози й адаптувати автономні алгоритми керування [4]. Для ефективного зберігання мультимодальних даних використовуються реляційні та нереляційні бази даних, графові структури та спеціалізовані платформи для потокової передачі великих обсягів даних [5].

https://doi.org/10.32782/mathematical-modelling/2025-8-1-27

Наукова спільнота демонструє успішні приклади застосування схожого підходу. Дослідження [6] показало, що злиття LiDAR і візуальних даних значно покращує класифікацію об'єктів у морських сценаріях. Автори [7] об'єднали RGB та тепловізійні дані для підвищення точності в умовах слабкої видимості. У роботі [8] досліджено злиття для покращення виявлення на великих відстанях.

Окрім того, Ф.С. Альсубаей зі співавторами [9] запропонували модель виявлення дрібних об'єктів на основі глибокого навчання, що знайшла застосування в системах управління відходами розумних міст – підкреслюючи ефективність глибоких нейронних мереж у нестандартних сценаріях.

А. Чеккареллі та Л. Монтеккі [10] акцентували увагу на важливості оцінювання помилкових виявлень (misdetection) з позиції безпеки та надійності, що є безпосередньо дотичним до критичних систем, подібних до запропонованої.

Значний інтерес становлять також роботи, присвячені ф'юзії з тактильними даними. Так, С. Табрік зі співавторами [11] продемонстрували, що візуальна й тактильна системи поділяють спільні риси в розпізнаванні об'єктів, тоді як Ґ. Рухафзай і А.-М. Крету [12] застосували глибинне навчання для обробки тактильних сигналів, підкріплених візуальним супроводом.

Д. Ахмад і А. Дель Буе [13] представили модель mmFUSION – архітектуру мультимодального злиття для 3D-детекції об'єктів, яка демонструє високу продуктивність на складних сценах.

На окрему увагу заслуговують підходи до виявлення небезпечної поведінки в реальному часі. Зокрема, Unsafe-Net, запропонований О. Онал і Е. Дандил [14], об'єднує YOLOv4 та ConvLSTM для фіксації небезпечних дій у робочому середовищі.

У контексті специфічних спектральних діапазонів С.А. Дансо зі співавторами [15] застосували глибинні моделі до обробки терагерцового зображення, що дозволило виявляти приховані небезпечні об'єкти з високою точністю.

Мета дослідження

Метою дослідження є створення системи, здатної надійно виявляти небезпечні об'єкти й оцінювати ступінь їхньої критичності для оперативного ухвалення рішень у режимі реального часу. Для цього в роботі запропоновано уніфіковану модель, яка поєднує кілька сенсорних потоків за допомогою проміжного злиття, доповненого механізмом уваги. Такий підхід забезпечує не лише інтеграцію даних, а й можливість адаптивного зважування внеску кожного сенсора відповідно до контексту, що є особливо важливим у динамічних і змінних умовах середовища.

Виклад основного матеріалу дослідження

Для ефективного виявлення небезпечних об'єктів у складному та змінному середовищі запропонована система використовує комбінацію кількох сенсорів, кожен із яких має унікальні властивості, що доповнюють один одного. Такий мультимодальний підхід дозволяє компенсувати обмеження окремих сенсорів і забезпечити надійне функціонування системи за різних умов. Нижче наведено перелік набору сенсорів для збору даних.

1. RGB-камера. Це базовий сенсор, що надає зображення високої роздільної здатності з детальною інформацією про колір, форму, текстуру та контури об'єктів. RGB-камери добре працюють у стандартних умовах освітлення і є незамінними для семантичного аналізу сцени, розпізнавання класів об'єктів і візуального супроводу. Водночас у разі поганого освітлення або великої кількості візуального шуму якість даних знижується.

2. LiDAR (Light Detection and Ranging). Лазерний сканер дозволяє створювати точну тривимірну модель середовища, що особливо важливо для визначення просторового положення об'єктів, їхніх розмірів і відстані до них. LiDAR не залежить від освітлення, що робить його незамінним у нічний час або в умовах обмеженої видимості. Його дані формують 3D-сітки зайнятості (occupancy grids), які потім перетворюються на воксельні представлення для подальшої обробки.

https://doi.org/10.32782/mathematical-modelling/2025-8-1-27

3. Тепловізор (інфрачервона камера). Цей сенсор фіксує теплове випромінювання, дозволяє ідентифікувати об'єкти за температурними характеристиками. Він особливо ефективний у темряві, під час дощу, туману або диму, коли звичайна камера втрачає здатність виявляти об'єкти. Тепловізори часто застосовуються для пошуку людей, тварин або перегрітих елементів обладнання, які можуть становити потенційну небезпеку.

4. Тактильні сенсори. Хоча вони використовуються рідше в автономних платформах, тактильні сенсори відіграють важливу роль у ситуаціях, де потрібно зафіксувати фізичну взаємодію з об'єктами. Вони дозволяють вимірювати тиск, вібрації, контактну температуру, жорсткість матеріалів та інші характеристики, які можуть бути критичними, наприклад, під час роботи з небезпечними або вибуховими об'єктами. У поєднанні з візуальними даними ці сенсори створюють багатошарову модель об'єкта.

Збір даних у системі відбувається синхронізовано – з урахуванням часових міток і просторової прив'язки. Це забезпечує узгодженість потоків інформації та дозволяє ефективно поєднувати їх у наступних етапах злиття. Під час тестування системи дані із сенсорів збиралися в контрольованих умовах із подальшим планом масштабування на реальні сценарії автономної навігації або промислового моніторингу.

У підсумку, комплексне використання вищеназваних сенсорів дозволяє побудувати гнучку, адаптивну та точну систему для виявлення небезпечних об'єктів із можливістю масштабування для різних галузей застосування. На рисунку 1 наведено повний потік даних від вхідного сигналу сенсора до виходу.

На рисунку 2 представлено приклади даних, отриманих від трьох ключових сенсорів, що використовуються в системі. Панель A (рисунок 2-A) демонструє RGB-зображення, отримане звичайною камерою в умовах денного освітлення. Таке зображення містить візуально насичену інформацію про колір, текстуру та форму об'єктів на перехресті, що дозволяє легко ідентифікувати транспортні засоби, пішоходів, дорожню розмітку та навколишню інфраструктуру. Панель B (рисунок 2-B) ілюструє хмару точок, згенеровану лазерним сканером LiDAR. Кожна точка в цій тривимірній репрезентації відповідає відображеному променю лазера, що формує просторову структуру сцени. Видно чітке позиціонування об'єктів у просторі, що особливо важливо для точного обчислення відстаней і взаємного розташування рухомих і статичних об'єктів. Панель C (рисунок 2-C) показує теплове зображення тієї ж сцени, яке візуалізує розподіл температур. Завдяки цьому можна виявити живі об'єкти (людей, тварин) або техніку, що працює, навіть у темряві чи візуально складних умовах. Яскраві ділянки вказують на теплові джерела, а холодні – на інертні поверхні.

Ці три зображення однієї сцени демонструють, наскільки різною може бути інформація про навколишнє середовище залежно від сенсора, та підкреслюють важливість їх об'єднання в мультимодальній системі. Комбіноване використання таких даних дозволяє моделі отримати повніше й точніше розуміння ситуації, що підвищує надійність виявлення небезпечних об'єктів.

Наступним етапом є попередня обробка зібраних даних. Дані кожного сенсора обробляються окремо (рисунок 3):

– зображення з камери покращуються для ясності;

– дані LiDAR перетворюються на воксельну 3D-сітку;

– теплові зображення вирівнюються для точного злиття.

Підготовлені й оброблені дані потрапляють у модуль проміжного злиття. Кожен потік даних кодується окремим енкодером:

$$f_i = E_i(S_i), \tag{1}$$

де S_i – вхід від сенсора, а E_i – енкодер.

https://doi.org/10.32782/mathematical-modelling/2025-8-1-27

Рис. 1. Повний потік даних від вхідного сигналу сенсора до виходу, критичного для безпеки

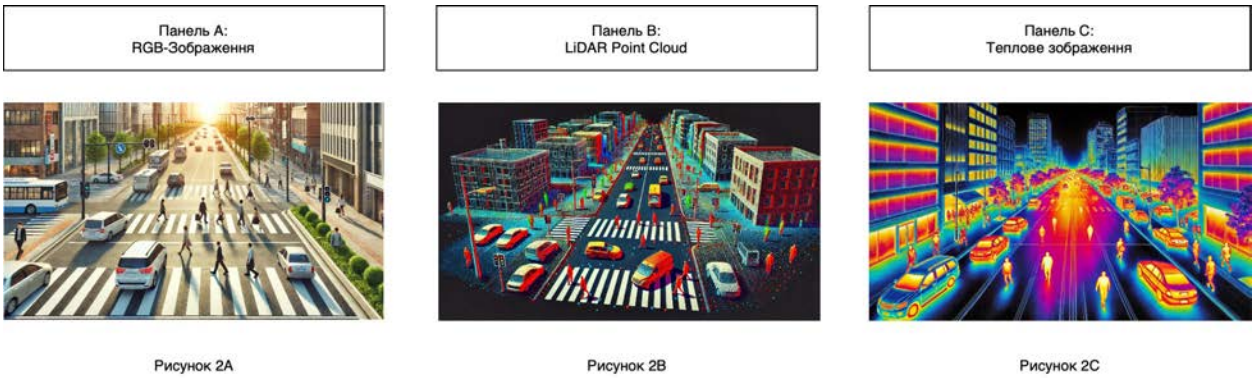

Рис. 2. Приклади даних датчиків

https://doi.org/10.32782/mathematical-modelling/2025-8-1-27

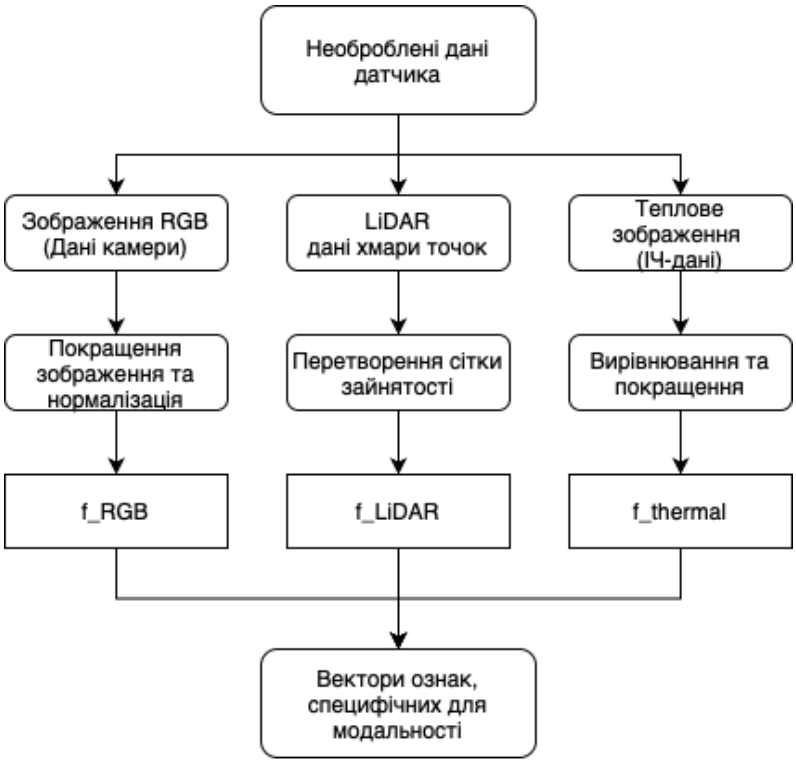

Рис. 3. Блок-схема попередньої обробки

Механізм уваги визначає ваги важливості для кожної модальності α_i, після чого формується об'єднаний вектор ознак F:

$$F = \sum_i a_i \cdot f_i. \tag{2}$$

Ця проміжна стратегія злиття розроблена для динамічного зважування важливості кожного датчика на основі поточних умов навколишнього середовища. На рисунку 4 наведено схему роботи модуля проміжного злиття мультимодальних даних.

Наступним етапом роботи запропонованої системи є декодування та виявлення об'єктів. Для цього вектор F подається в декодер D, який формує:

– рамки границь;
– класи об'єктів;
– оцінку впевненості.

Математично це представлено як:

$$D(F) = \{b, c, s\}, \tag{3}$$

де b позначає обмежувальні рамки, c – мітки класів, а s – упевненість у виявленні.

Для оцінювання критичності об'єктів для кожного об'єкта розраховується оцінка C за формулою:

$$C = f(d, v, \theta, a), \tag{4}$$

де d – відстань, v – відносна швидкість, θ – орієнтація, a – константи згасання.

Ця оцінка гарантує, що об'єкти з високим ризиком (як-от ті, що розташовані на курсі зіткнення) мають пріоритет (рисунок 5).

https://doi.org/10.32782/mathematical-modelling/2025-8-1-27

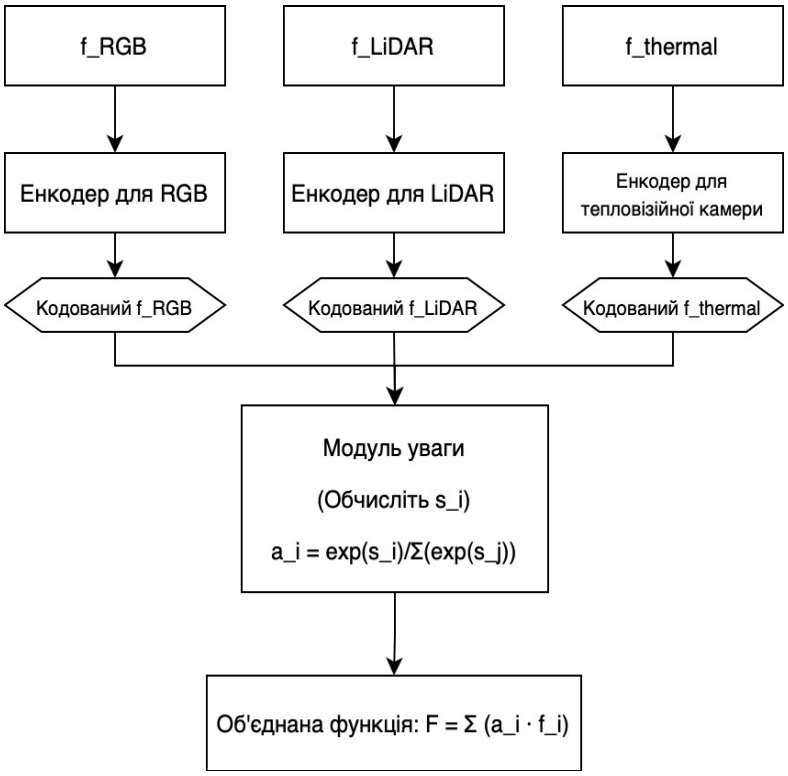

Рис. 4. Схема роботи модуля проміжного злиття

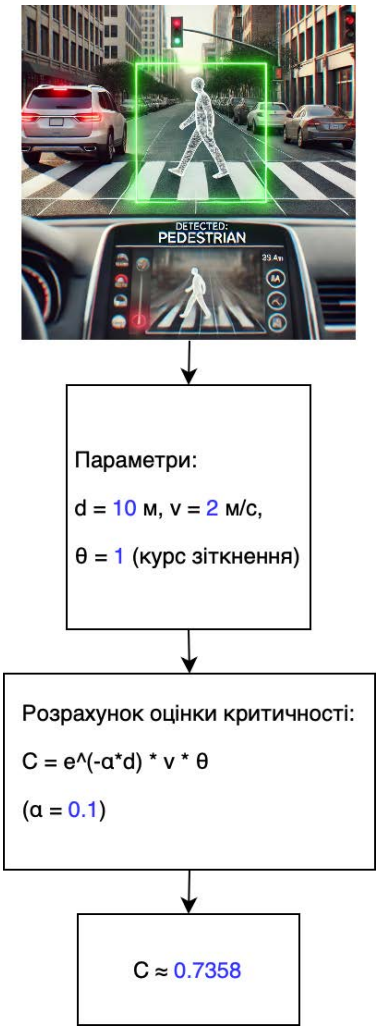

Рис. 5. Діаграма оцінювання критичності об'єкта

Остаточне рішення щодо виявлення об'єкта визначається за допомогою модуля ухвалення рішення та сповіщення і представлено як:

$$S_{final} = \beta \cdot s + (1-\beta) \cdot C, \tag{5}$$

де β – це ваговий коефіцієнт, що визначає пріоритетність між упевненістю в детекції та оцінкою потенційної небезпеки.

Значення β може змінюватися залежно від сценарію: наприклад, у середовищах з високим візуальним шумом β може бути зменшено, щоб більше покладатися на критичність об'єкта (C), тоді як у чітких умовах β може бути підвищено, віддаючи перевагу прямій оцінці впевненості.

Якщо S_{final} перевищує заздалегідь визначений пороговий рівень, система ініціює тривожний сигнал або інше захисне реагування, наприклад екстрене гальмування, зміну маршруту чи виведення повідомлення оператору.

Блок-схема на рисунку 6 наочно демонструє процес інтеграції даних із двох ключових джерел – оцінки достовірності детекції ($S_{detection}$) та оцінки критичності об'єкта (C). Обчислення остаточного результату (S_{final}) базується на параметрі β, який дозволяє балансувати між довірою до вхідних детекцій і ризик-орієнтованим підходом.

Візуалізація процесу ухвалення рішення допомагає краще зрозуміти логіку системи та її здатність до адаптації в реальному часі. Такий підхід особливо важливий у системах, де необхідне негайне реагування – наприклад, за автоматичного гальмування в разі наявності загрози зіткнення або інформування оператора про потенційну небезпеку.

Застосування чітко визначеного порогового значення забезпечує контроль над чутливістю системи та дозволяє налаштовувати її відповідно до конкретних сценаріїв використання – від промислових роботів до автономного транспорту. Отже, представлена блок-схема є не лише

https://doi.org/10.32782/mathematical-modelling/2025-8-1-27

технічним елементом, а й засобом забезпечення прозорості та надійності роботи системи у критичних ситуаціях.

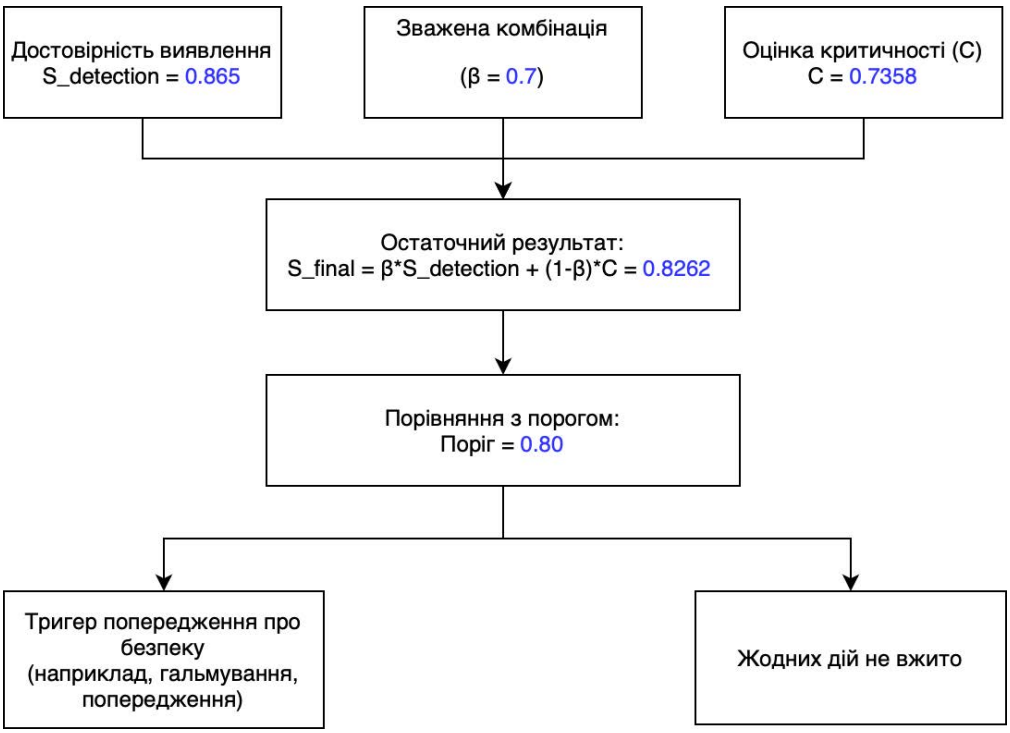

Рис. 6. Блок-схема ухвалення остаточного рішення про виявлення

Приклад роботи запропонованої системи

Рисунок 7 ілюструє покроковий процес обробки мультимодальних сенсорних даних у запропонованій системі. На першому етапі система отримує вхідну інформацію із трьох сенсорів: RGB-камери, LiDAR і тепловізора. Кожному з них присвоюється значення надійності, а також формується вектор ознак – f_{cam}, f_{LiDAR}, $f_{thermal}$ відповідно.

Дані проходять через етап попередньої обробки, де усуваються шуми, нормалізуються значення та виділяються ключові ознаки. Наступним кроком є обчислення ваг уваги (attention weights) – ці значення (наприклад, a_{cam} = 0,6) визначають вплив кожного сенсора на фінальне рішення, з урахуванням умов середовища.

У модулі проміжного злиття ознаки поєднуються в єдиний вектор F, який передається до об'єднаного декодера. Декодер формує оцінку ймовірності виявлення об'єкта $S_{detection}$, що в цьому прикладі становить 0,865.

Окремо проводиться оцінювання критичності об'єкта (C) на основі фізичних параметрів: відстані (d = 10 м), швидкості (v = 2 м/с) і орієнтації (θ = 1), що дозволяє системі оцінити потенційний рівень загрози. Значення C у прикладі – 0,7358.

Після цього розраховується фінальна оцінка S_{final} із використанням вагового коефіцієнта β = 0,7. У результаті отримано S_{final} = 0,8262, що перевищує поріг 0,8. Отже, система активує тригер безпеки, який може ініціювати, наприклад, автоматичне гальмування чи попередження оператора.

Цей рисунок наочно демонструє ефективність багаторівневої логіки виявлення та ухвалення рішень у системі, що поєднує точність детекції з аналізом ризиків у реальному часі.

Запропонована архітектура системи, яка базується на злитті мультимодальних сенсорних даних із механізмами уваги, демонструє низку ключових переваг, що роблять її перспективною для застосування у сфері автономних технологій, робототехніки, охорони праці та безпеки.

https://doi.org/10.32782/mathematical-modelling/2025-8-1-27

По-перше, це висока точність і стійкість системи. Завдяки інтеграції різних типів сенсорів, як-от RGB-камери, LiDAR і тепловізори, модель може ефективно функціонувати навіть у несприятливих умовах – за слабкого освітлення, в умовах дощу чи туману, або коли один із сенсорів дає нечітку або обмежену інформацію. Такий підхід значно перевершує односенсорні рішення, які часто втрачають ефективність у складних середовищах.

По-друге, динамічне зважування сенсорних потоків через механізм уваги дозволяє адаптувати поведінку системи в реальному часі. Це забезпечує гнучкість і розуміння контексту, що особливо важливо в переході між відкритими просторами, тунелями, міським середовищем або промисловими об'єктами.

По-третє, урахування рівня критичності виявлених об'єктів через спеціальний модуль оцінки дозволяє системі не просто фіксувати наявність об'єктів, а й аналізувати, наскільки вони становлять потенційну небезпеку. Це забезпечує кращу пріоритезацію рішень, що може бути критичним у ситуаціях з обмеженим часом на реакцію.

Четвертою перевагою є використання проміжного злиття, яка забезпечує збереження як просторових, так і семантичних ознак даних. Це дозволяє досягти балансу між раннім злиттям (яке може втратити специфічну інформацію) та пізнім (яке може ускладнити інтеграцію через відмінності у форматах вихідних даних).

Водночас запропонований підхід має деякі обмеження і виклики, які необхідно враховувати в його практичному впровадженні:

– високоточна калібровка сенсорів. Будь-які зсуви або помилки в узгодженні координат, часових міток або позиціювання між сенсорами можуть призводити до спотворення або втрати частини інформації, що суттєво впливає на результат злиття;

– висока обчислювальна складність. Застосування складних механізмів уваги, багаторівневих енкодерів і декодерів, а також обробка потоків даних із кількох сенсорів потребують значних ресурсів, як обчислювальних, так і енергетичних. Це може бути критичним для автономних пристроїв з обмеженими можливостями;

– обмежена доступність якісних мультимодальних датасетів. Для повноцінного навчання та тестування таких систем потрібні великі масиви даних, що включають синхронізовані записи з кількох типів сенсорів у реальних умовах. Нині такі датасети є рідкісними й обмеженими за варіативністю.

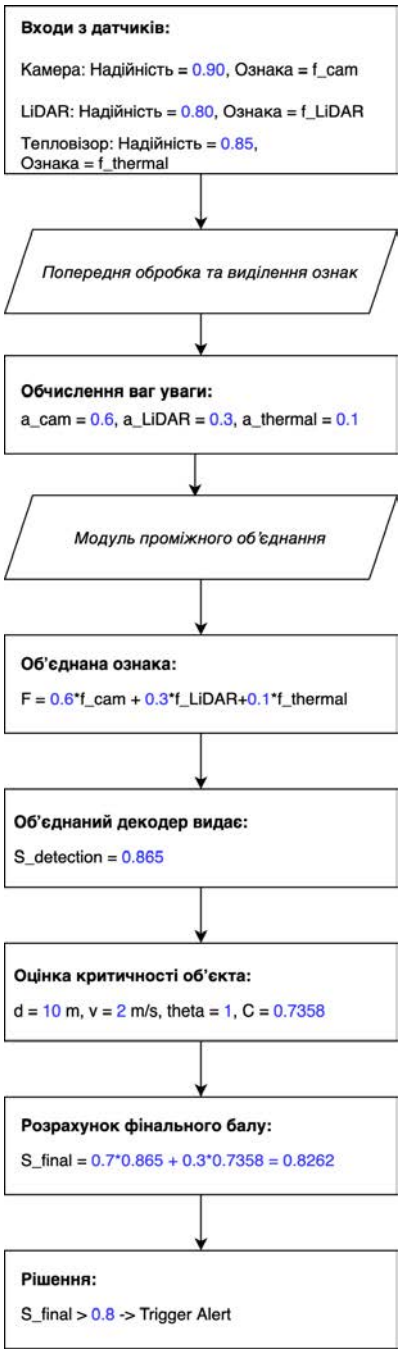

Рис. 7. Комплексна діаграма потоку даних для гіпотетичного прикладу

Отже, хоча система вже показує високі результати на етапі експериментальної перевірки, її подальше вдосконалення є необхідним для широкого впровадження в реальні автономні платформи.

Висновки

У даному дослідженні запропоновано уніфіковану систему виявлення небезпечних об'єктів, що базується на проміжному злитті мультимодальних сенсорних даних із використанням

https://doi.org/10.32782/mathematical-modelling/2025-8-1-27

механізмів уваги. Такий підхід дозволяє поєднувати переваги різних типів сенсорів – візуальних, тепловізійних, глибинних тощо – для створення цілісного й адаптивного уявлення про навколишнє середовище. Завдяки інтеграції механізмів уваги система може динамічно зважувати інформацію з кожного джерела залежно від поточних умов, що суттєво підвищує її стійкість і точність навіть у складних або нетипових сценаріях.

Окрім того, уведення модуля оцінювання критичності об'єктів дозволяє не лише ідентифікувати об'єкти, а й формувати їх пріоритет за ступенем потенційної небезпеки. Це особливо важливо для автономних систем, де швидкість і обґрунтованість ухвалення рішень безпосередньо впливають на безпеку.

Результати моделювання демонструють перспективність запропонованого підходу та його придатність до реального впровадження. Водночас дослідження виявило низку викликів, зокрема – потребу у високоточній калібровці сенсорів, зниження обчислювальної складності та розширення бази даних для тренування моделей.

У майбутньому доцільно зосередитися на розробленні алгоритмів автоматичної калібровки сенсорів, оптимізації архітектури нейромереж для роботи в режимі реального часу, а також на створенні більш масштабних і різнорідних мультимодальних датасетів. Вирішення цих завдань сприятиме подальшому розвитку надійних систем виявлення небезпечних об'єктів для автономних роботизованих рішень, транспортних засобів і критичних інфраструктур.

Список використаної літератури

1. Arya C., Tripathi A., Singh P., Diwakar M., Sharma K., Pandey H. Object detection using deep learning: a review. *Journal of Physics: Conference Series*. 2021. Vol. 1854, № 1. Article 012012. DOI: 10.1088/1742-6596/1854/1/012012.
2. Zhao F., Zhang C., Geng B. Deep Multimodal Data Fusion. *ACM Computing Surveys*. 2024. Vol. 56, № 9. Article 216. DOI: 10.1145/3649447.
3. Fedorenko H., Fesenko H., Kharchenko V. Analysis of methods and development of a concept for guaranteed detection and recognition of explosive objects. *Innovative Technologies and Scientific Solutions for Industries*. 2022. № 4 (22). P. 20–31.
4. Щур Г. Інтелектуальна система керування роботом-сапером у віртуальному середовищі. *Вісник Хмельницького національного університету.* 2024. № 3 (1). P. 326–329. DOI: 10.31891/2307-5732-2024-335-3-43.
5. Dumyn I., Basystiuk O., Dumyn A. Graph-based approaches for multimodal medical data processing. 2023.
6. Thompson D.J. Maritime Object Detection, Tracking, and Classification Using LiDAR and Vision-Based Sensor Fusion. 2017. [Technical report].
7. Vadidar M., Kariminezhad A., Mayr C., Kloeker L., Eckstein L. Robust Environment Perception for Automated Driving: A Unified Learning Pipeline for Visual-Infrared Object Detection. *Proceedings of the IEEE Intelligent Vehicles Symposium (IV)*. 2022. P. 367–374.
8. Bhown A. Improving Long-Range 3D Object Detection Methods for Autonomous Box Trucks Using Sensor Fusion. 2022. [Preprint].
9. Alsubaei F.S., Al-Wesabi F.N., Hilal A.M. Deep Learning-Based Small Object Detection and Classification Model for Garbage Waste Management in Smart Cities and IoT Environment. 2022. [Preprint].
10. Ceccarelli A., Montecchi L. Evaluating Object (mis)Detection from a Safety and Reliability Perspective: Discussion and Measures. *IEEE Access*. 2023. Vol. 11. P. 44952–44963.
11. Tabrik S., Behroozi M., Schlaffke L., Heba S., Lenz M., Lissek S., Güntürkün O., Dinse H.R., Tegenthoff M. Visual and Tactile Sensory Systems Share Common Features in Object Recognition. *eNeuro*. 2021. Vol. 8, № 5. DOI: 10.1523/ENEURO.0261-21.2021.

https://doi.org/10.32782/mathematical-modelling/2025-8-1-27

12. Rouhafzay G., Cretu A.-M. An Application of Deep Learning to Tactile Data for Object Recognition under Visual Guidance. *Sensors*. 2019. Vol. 19, № 7. Article 1534. DOI: 10.3390/s19071534.
13. Ahmad J., Del Bue A. mmFUSION: Multimodal Fusion for 3D Objects Detection. arXiv preprint arXiv:2311.04058. 2023. URL: https://arxiv.org/abs/2311.04058.
14. Önal O., Dandıl E. Unsafe-Net: YOLO v4 and ConvLSTM Based Computer Vision System for Real-Time Detection of Unsafe Behaviours in Workplace. *Multimedia Tools and Applications.* 2024. P. 1–27.
15. Danso S.A., Shang L., Hu D., Odoom J., Liu Q., Nyarko B. Hidden Dangerous Object Recognition in Terahertz Images Using Deep Learning Methods. *Applied Sciences*. 2022. Vol. 12, № 15. Article 7354.

References

1. Arya, C., Tripathi, A., Singh, P., Diwakar, M., Sharma, K., & Pandey, H. (2021). Object detection using deep learning: A review. *Journal of Physics: Conference Series*, 1854 (1), 012012. https://doi.org/10.1088/1742-6596/1854/1/012012 [in English].
2. Zhao, F., Zhang, C., & Geng, B. (2024). Deep multimodal data fusion. *ACM Computing Surveys*, 56 (9), Article 216. https://doi.org/10.1145/3649447 [in English].
3. Fedorenko, H., Fesenko, H., & Kharchenko, V. (2022). Analysis of methods and development of a concept for guaranteed detection and recognition of explosive objects. *Innovative Technologies and Scientific Solutions for Industries*, 4 (22), 20–31 [in English].
4. Shchur, H. (2024). Intelektualna systema keruvannia robotom-saperom u virtualnomu seredovyshchi [Intelligent system for demining robot control in a virtual environment]. *Visnyk Khmelnytskoho natsionalnoho universytetu – Herald of Khmelnytskyi National University. Technical Sciences*, 3 (1), 326–329. https://doi.org/10.31891/2307-5732-2024-335-3-43 [in Ukrainian].
5. Dumyn, I., Basystiuk, O., & Dumyn, A. (2023). Graph-based approaches for multimodal medical data processing. *Unpublished manuscript* [in English].
6. Thompson, D.J. (2017). *Maritime object detection, tracking, and classification using LiDAR and vision-based sensor fusion* [Technical report].
7. Vadidar, M., Kariminezhad, A., Mayr, C., Kloeker, L., & Eckstein, L. (2022). Robust environment perception for automated driving: A unified learning pipeline for visual-infrared object detection. In *Proceedings of the IEEE Intelligent Vehicles Symposium* (IV) (pp. 367–374) [in English].
8. Bhown, A. (2022). *Improving long-range 3D object detection methods for autonomous box trucks using sensor fusion* [Preprint] [in English].
9. Alsubaei, F.S., Al-Wesabi, F.N., & Hilal, A.M. (2022). *Deep learning-based small object detection and classification model for garbage waste management in smart cities and IoT environment* [Preprint] [in English].
10. Ceccarelli, A., & Montecchi, L. (2023). Evaluating object (mis)detection from a safety and reliability perspective: Discussion and measures. *IEEE Access*, 11, 44952–44963. https://doi.org/10.1109/ACCESS.2023.3282082 [in English].
11. Tabrik, S., Behroozi, M., Schlaffke, L., Heba, S., Lenz, M., Lissek, S., Güntürkün, O., Dinse, H.R., & Tegenthoff, M. (2021). Visual and tactile sensory systems share common features in object recognition. *eNeuro*, 8 (5). https://doi.org/10.1523/ENEURO.0261-21.2021 [in English].
12. Rouhafzay, G., & Cretu, A.-M. (2019). An application of deep learning to tactile data for object recognition under visual guidance. *Sensors*, 19 (7), 1534. https://doi.org/10.3390/s19071534 [in English].
13. Ahmad, J., & Del Bue, A. (2023). *mmFUSION: Multimodal fusion for 3D objects detection* [Preprint]. arXiv. Retrieved from: https://arxiv.org/abs/2311.04058 [in English].

14. Önal, O., & Dandıl, E. (2024). Unsafe-Net: YOLO v4 and ConvLSTM based computer vision system for real-time detection of unsafe behaviours in workplace. *Multimedia Tools and Applications*, 1–27. https://doi.org/10.1007/s11042-024-18555-z [in English].
15. Danso, S. A., Shang, L., Hu, D., Odoom, J., Liu, Q., & Nyarko, B. (2022). Hidden dangerous object recognition in terahertz images using deep learning methods. *Applied Sciences*, 12 (15), 7354. https://doi.org/10.3390/app12157354 [in English].

Щур Гліб Олександрович – аспірант кафедри систем штучного інтелекту Національного університету «Львівська політехніка». E-mail: hlib.o.shchur@lpnu.ua, ORCID: 0000-0002-3796-2866.

Думин Ірина Богданівна – к.т.н, старший викладач кафедри систем штучного інтелекту Національного університету «Львівська політехніка». E-mail: iryna.b,shvorob@lpnu.ua, ORCID: 0000-0001-5569-2647.

Shchur Hlib Oleksandrovych – Postgraduate Student at the Department of Artificial Intelligence Systems of the Lviv Polytechnic National University. E-mail: hlib.o.shchur@lpnu.ua, ORCID: 0000-0002-3796-2866.

Dumyn Iryna Bohdanivna – Candidate of Technical Sciences, Senior Lecturer at the Department of Artificial Intelligence Systems of the Lviv Polytechnic National University. E-mail: iryna.b.shvorob@lpnu.ua, ORCID: 0000-0001-5569-2647.

https://doi.org/10.32782/mathematical-modelling/2025-8-1-27