

2025 p.

Харківський національний університет радіоелектроніки

Факультет Електронної та біомедичної інженерії
(повна назва)

Кафедра Мікроелектроніки, електронних приладів та пристроїв
(повна назва)

Рівень вищої освіти другий (магістерський)

Спеціальність 171 Електроніка
(код і повна назва)

Тип програми освітньо-професійна
(освітньо-професійна або освітньо-наукова)

Освітня програма Електронні прилади та пристрої
(повна назва)

ЗАТВЕРДЖУЮ:

Зав. кафедри _____
(підпис)

« _____ » 20 ____ р.

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ

Здобувачеві Крохмалю Андрію Віталійовичу
(прізвище, ім'я, по батькові)

1. Тема роботи Розробка та удосконалення методів розпізнавання малорозмірних та малоконтрастних об'єктів в оптико-електронних системах технічного зору

затверджена наказом університету від 17 листопада 2025 р. № 199 Стз

2. Термін подання студентом роботи до екзаменаційної комісії 17 грудня 2025 р.

3. Вихідні дані до роботи параметри системи технічного зору: роздільна здатність матриці: 1600x1200 пікселі; лінійний розмір фотоприймача: 3.590x2.684 мм; чутливість фотоприймача: 0.6 В / Лк·с; напруга насичення: 2.7 В; темновий струм: 15 мВ/с при 60°; апертура об'єктиву: квадратна, 1x1 мм; фокусна відстань лінзи об'єктиву: 3.78 мм; фрагмент зображення для аналізу: 224x224 пікселі, максимальна кількість виявлених об'єктів: 5.

4. Перелік питань, що потрібно опрацювати в роботі Методи обробки сигналів в системах технічного зору (структурна схема та методи формування зображень в системах технічного зору; особливості формування, просторові характеристики сигналів та зображень, які реєструються системами технічного зору; методи та алгоритми для реєстрації об'єктів в системах технічного зору).

Розробка та удосконалення системи технічного зору для виявлення та підрахунку малорозмірних та малоконтрастних об'єктів (імітаційне моделювання відгуку багатоелементного фотоприймача з урахуванням умов спостереження; розробка комбінованого методу підрахунку та локації фіксованого числа об'єктів на основі глибокого навчання; апаратно-програмна реалізація системи технічного зору для вбудованої системи).

Експериментальні дослідження (моделювання вихідних сигналів фотоприймача в

складних умовах; визначення статистичних параметрів близько розташованих сигналів, тестування системи технічного зору, реалізованої у вигляді вбудованої системи).

5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп'ютерних ілюстрацій (п.5 включається до завдання за рішенням випускової кафедри)

Слайди 1-21

6. Консультанти розділів роботи (п.6 включається до завдання за наявності консультантів згідно з наказом, зазначеним у п.1)

Найменування розділу	Консультант (посада, прізвище, ім'я, по батькові)	Позначка консультанта про виконання розділу	
		підпис	дата

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Терміни виконання етапів роботи	Примітка
1	Одержання завдання на виконання роботи	17.11.2025	виконано
2	Укладання і погодження з керівником плану і етапів виконання роботи	19.11.2025	виконано
3	Узагальнення даних літературних джерел	21.11.2025	виконано
4	Аналіз шляхів виконання завдання. Вибір і погодження з керівником оптимального шляху виконання завдання	25.11.2025	виконано
5	Аналіз технічних засобів та існуючих систем	28.12.2025	виконано
6	Реалізація практичної частини завдання	05.12.2025	виконано
7	Виконання експериментальних досліджень	08.12.2025	виконано
8	Укладання, оформлення та погодження пояснювальної записки з керівником	15.12.2025	виконано
9	Підготовка слайдів презентації та доповіді	16.12.2025	виконано
10	Надання пояснювальної записки на кафедру	17.12.2025	виконано

Дата видачі завдання 17 листопада 2025 р.

Здобувач 
(підпис)

Керівник роботи _____ проф. Стрілкова Т. О.
(підпис) (посада, прізвище, ініціали)

РЕФЕРАТ

Пояснювальна записка містить: 108 с., 40 рис., 5 табл., 67 джерел.

СИСТЕМИ ТЕХНІЧНОГО ЗОРУ, ОПТИКО-ЕЛЕКТРОННІ СИСТЕМИ,
МЕТОДИ ОБРОБКИ СИГНАЛІВ, СТОХАСТИКО-ДЕТЕРМІНОВАНА ТЕОРІЯ
СИГНАЛІВ, ОПТИЧНА ЛОКАЦІЯ, ГЛИБОКЕ НАВЧАННЯ, ЗГОРТКОВІ
НЕЙРОННІ МЕРЕЖІ

Об'єкт дослідження – процеси приймання, реєстрації та обробки сигналів в оптико-електронних системах технічного зору.

Предмет дослідження – методи формування відгуку сенсору в оптоелектронній системі, метод і алгоритм обробки вихідного сигналу (виявлення та локації малорозмірних і малоконтрастних об'єктів).

Мета роботи – вдосконалення процесу виявлення та локації малорозмірних та малоконтрастних об'єктів в оптико-електронних системах технічного зору, що використовують методи розпізнавання на основі глибокого навчання.

Методологічна та теоретична основа дослідження: елементи хвильової теорії світла, теорії ймовірності та математичної статистики, теорії побудови оптико-електронних систем, теорії множин, теорії розпізнавання образів, аналіз і синтез, порівняння, абстрагування, моделювання, експеримент, загальні методи обробки сигналів, методи комп'ютерного зору, методи обробки «великих даних», методи глибинного навчання.

В роботі проаналізовано загальні принципи формування сигналів та зображень в системах технічного зору та методи виявлення сигналів. Розроблено імітаційну модель формування зображень в системах технічного зору, що враховує комплекс фізичних явищ (дифракцію, атмосферну турбулентність, шумові процеси) та інші особливості роботи фотоприймача. Розроблено комбінований метод виявлення та локації об'єктів, зокрема близько розташованих, малорозмірних та малоконтрастних, на основі глибоких

згорткових нейронних мереж. Реалізовано алгоритм виявлення та локації об'єктів для застосування у вбудованих та автономних системах технічного зору.

Отримані результати можуть бути використані в широкому спектрі наукових, технічних та прикладних застосувань, включаючи астрономічні спостереження, системи контролю у промисловості та логістиці, у військовій розвідці та системах наведення, сільському господарстві та медицині.

ABSTRACT

Explanatory note consists of: 108 p., 40 fig., 5 tables, 67 sources.

TECHNICAL VISION SYSTEMS, OPTOELECTRONIC SYSTEMS,
SIGNAL PROCESSING TECHNIQUES, STOCHASTIC-DETERMINISTIC
SIGNAL THEORY, OPTICAL LOCATION, PHOTODETECTOR RESPONSE,
DEEP LEARNING, CONVOLUTIONAL NEURAL NETWORKS

Object of Research – the processes of reception, registration, and processing of signals in optoelectronic technical vision systems.

Subject of Research – methods of sensor response formation in an optoelectronic system, as well as methods and algorithms for processing the output signal (detection and localization of small-sized and low-contrast objects).

Purpose of the Work – improvement of the processes of detection and localization of small-sized and low-contrast objects in optoelectronic technical vision systems employing recognition methods based on deep learning.

The methodological and theoretical foundation of the study integrates elements of wave optics, probability theory and mathematical statistics, optoelectronic system design, set theory, and pattern recognition, complemented by analytical and synthetic approaches, abstraction, modeling, experimentation, as well as contemporary methods of signal processing, computer vision, big data analysis, and deep learning.

The study analyzes the general principles of signal and image formation in technical vision systems and methods of signal detection. A simulation model of image formation in technical vision systems has been developed, which accounts for a complex of physical actors (diffraction, atmospheric turbulence, noise processes) and other specific aspects of photodetector operation. A combined method for the detection and localization of objects – particularly closely spaced, small-sized, and low-contrast ones – has been proposed based on deep convolutional neural networks. An algorithm

for object detection and localization has been implemented for use in embedded and autonomous technical vision systems.

The obtained results can be applied across a wide range of scientific, technical, and practical domains, including astronomical observations, industrial and logistics control systems, military reconnaissance and guidance systems, agriculture, and medicine.

ЗМІСТ

ВСТУП	10
1 МЕТОДИ ОБРОБКИ СИГНАЛІВ В СИСТЕМАХ ТЕХНІЧНОГО ЗОРУ ...	18
1.1 Структурна схема та методи формування зображень в системах технічного зору.....	18
1.2 Особливості формування, просторові характеристики сигналів та зображень, які реєструються системами технічного зору	21
1.3 Методи та алгоритми для реєстрації об'єктів в системах технічного зору	29
Висновки до розділу 1	45
1 РОЗРОБКА ТА УДОСКОНАЛЕННЯ СИСТЕМИ ТЕХНІЧНОГО ЗОРУ ДЛЯ ВИЯВЛЕННЯ ТА ПІДРАХУНКУ МАЛОРОЗМІРНИХ ТА МАЛОКОНТРАСТНИХ ОБ'ЄКТІВ.....	47
2.1. Імітаційне моделювання відгуку багатoelementного фотоприймача з урахуванням умов спостереження.....	47
2.2 Розробка комбінованого методу підрахунку та локації фіксованого числа об'єктів на основі глибокого навчання	59
2.3 Апаратно-програмна реалізація системи технічного зору для вбудованої системи	72
Висновки до розділу 2	81
2 ЕКСПЕРИМЕНТАЛЬНІ ДОСЛІДЖЕННЯ.....	83
3.1 Результати моделювання вихідних сигналів фотоприймача в складних умовах.....	83
3.2. Результати визначення статистичних параметрів близько розташованих сигналів	92
3.3. Результати тестування системи технічного зору, реалізованої у вигляді вбудованої системи	94
Висновки до розділу 3	98
ВИСНОВКИ.....	99

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ	101
Додаток А.....	Помилка! Закладку не визначено.
Додаток Б	Помилка! Закладку не визначено.

ВСТУП

Останні десятиліття характеризуються стрімким розвитком автоматизації та впровадженням інформаційних систем і технологій в усі сфери людської діяльності, що зумовлює зростання ролі інтелектуальних систем обробки інформації. Серед них особливе місце посідають системи технічного зору [1, 2].

Це пов'язано, в першу чергу, з тим, що переважна більшість інформації, яка отримується людиною про оточуюче середовище (до 80 % усієї інформації) приходить саме на візуальну.

Вони забезпечують сприйняття, аналіз та інтерпретацію візуальних даних у режимі реального часу, що є необхідною умовою функціонування автономних комплексів, зокрема систем підтримки прийняття рішень.

Сфери застосування систем технічного зору надзвичайно широкі. Системи технічного зору є ключовим елементом концепції Індустрії 4.0, що передбачає глибоку інтеграцію засобів автоматизації, сенсорів та технологій штучного інтелекту у виробничі процеси. У промисловості вони застосовуються для автоматичного контролю якості продукції, ідентифікації дефектів, керування технологічними операціями та навігації роботизованих маніпуляторів. Завдяки цьому забезпечується підвищення ефективності виробництва та зменшення впливу людського фактора [3, 4].

У військовій галузі системи технічного зору відіграють важливу роль у розвідувальних і спостережних комплексах. Зокрема, вони застосовуються для виявлення, розпізнавання та супроводу об'єктів у складних умовах низької контрастності, при наявності завад або маскування. Використання тепловізійних, мультиспектральних та гіперспектральних сенсорів дає змогу отримувати достовірну інформацію про обстановку в будь-яких погодних і часових умовах, що є критично важливим для підвищення ситуаційної обізнаності та ефективності прийняття рішень [5].

В транспортній сфері системи технічного зору становлять основу для реалізації систем допомоги водію (ADAS) та автономного транспорту. На основі даних із камер, лідарів та інших сенсорів формується тривимірна модель навколишнього середовища, що дозволяє транспортним засобам розпізнавати пішоходів, перешкоди, дорожню розмітку, сигнали світлофорів та інші об'єкти. Це сприяє значному підвищенню рівня безпеки дорожнього руху та відкриває нові можливості для впровадження безпілотних транспортних систем.

У медицині технології технічного зору використовуються для автоматизованої діагностики, аналізу зображень медичних сканерів та керування роботизованими системами під час хірургічних втручань. Сучасні методи виявлення забезпечують ідентифікацію патологій із високим рівнем достовірності, що сприяє підвищенню ефективності клінічних досліджень та оптимізації процесів медичного обслуговування [6].

В сільському господарстві системи технічного зору забезпечують моніторинг стану рослин, оцінку ступеня дозрівання плодів, виявлення бур'янів і шкідників, а також контроль якості зібраного врожаю. Інтеграція таких систем із безпілотними літальними апаратами дозволяє реалізувати концепцію точного землеробства, що базується на раціональному використанні ресурсів і зниженні негативного впливу на довкілля.

Крім того, системи технічного зору активно впроваджуються у сферах безпеки, охорони, екологічного моніторингу, розумних міст і наукових досліджень. Технічний зір, поєднуючи методи та засоби оптоелектроніки, комп'ютерного зору та штучного інтелекту, забезпечує можливість автоматизованого сприйняття та аналізу візуальної інформації, яка раніше була доступна лише людині-оператору [7].

Отже, дослідження, спрямовані на розвиток систем технічного зору, є актуальним науково-технічним завданням, яке має вагоме практичне значення для різних галузей людської діяльності. Необхідність розвитку систем технічного зору визначається кількома чинниками. По-перше, зростає потреба у високоточному та швидкому контролі якості в умовах масового виробництва.

По-друге, глобальна конкуренція стимулює підприємства до впровадження інноваційних технологій, що знижують витрати та підвищують ефективність. По-третє, розвиток алгоритмів машинного навчання, глибоких нейронних мереж і спеціалізованих апаратних рішень робить системи технічного зору доступнішими та функціонально потужнішими. У результаті вони стають не лише допоміжним інструментом, а й стратегічним елементом сучасних виробничих і сервісних систем.

Подальший прогрес у цій сфері пов'язаний із удосконаленням методів формування, обробки та інтерпретації зображень, створенням ефективних алгоритмів розпізнавання та підвищенням стійкості систем до зовнішніх факторів, що в цілому сприятиме підвищенню рівня автоматизації, безпеки та інтелектуалізації сучасних технічних та технологічних систем.

Синергетичне поєднання оптоелектронних засобів формування зображень та інтелектуальних алгоритмів розпізнавання образів утворює єдину систему технічного зору, здатну не лише фіксувати візуальну інформацію, а й здійснювати її семантичну інтерпретацію.

Ефективність функціонування систем технічного зору значною мірою визначається рівнем розвитку оптоелектронних систем, які забезпечують первинне перетворення оптичного випромінювання в електричні сигнали, що підлягають подальшій цифровій обробці. Оптоелектронна частина таких систем виконує роль сенсорного рівня сприйняття навколишнього середовища, формуючи вхідні дані для алгоритмів розпізнавання та інтерпретації зображень. Вона включає сукупність оптичних, фотоприймальних, підсилювальних і перетворювальних елементів, які працюють у взаємозв'язку з цифровими засобами обробки сигналів та програмними модулями аналізу інформації [8].

Апарат розпізнавання образів дозволяє отримувати класифікаційні рішення на основі аналізу широкого спектру різноформатних даних та є ефективним інструментом підтримки прийняття рішень у випадках, коли об'єкт розпізнавання представляє собою складну систему з певним ступенем стохастичності. У сучасних системах процеси інтелектуальної обробки сигналів

здійснюються з використанням нейронних мереж, алгоритмів глибокого навчання, статистичних і нечітких методів аналізу, що дозволяє досягати високої точності та ефективності [9].

Однак, однією з ключових проблем є обмежена роздільна здатність оптичної системи, що може бути недостатньою при спостереженні двох близько розташованих джерел випромінювання. Якщо відстань між ними менша за межу, визначену критерієм Релея, то зображення сигналів перекриваються настільки, що їх неможливо розрізнити навіть за умови високої якості сенсора. У таких випадках система втрачає здатність до точної локалізації об'єктів, що особливо критично для задач високоточного наведення, дистанційного зондування або визначення координат малих об'єктів. Також ефективність систем технічного зору в реальних умовах експлуатації істотно обмежується низкою факторів, пов'язаних із фізичною природою процесів формування зображення, а також з дією зовнішніх і внутрішніх завад. Ці чинники знижують точність просторового розділення, контрастність, стабільність та інформативність зображень, що ускладнює подальшу обробку і розпізнавання об'єктів [10].

Огляд сучасної літератури показує, що дослідники зосереджені на подоланні ключових складнощів, пов'язаних із роботою систем технічного зору в умовах високого рівня завад, турбулентної атмосфери та динамічних спотворень. Рішення в галузі оптоелектроніки спрямовані на підвищення сигнал-на-шум за рахунок апаратних удосконалень, зокрема використання фотоприймачів із підвищеною спектральною селективністю та оптичних фільтрів [11]. Паралельно з цим активно розвиваються наноінженерні підходи: використання метаматеріалів і квантових точок дозволяє підвищити чутливість сенсорів і компенсувати вплив зовнішніх перешкод. Цей напрям детально розкрито в роботі [12], де наведено методи підвищення чутливості оптоелектронних сенсорів за рахунок новітніх наноструктурованих матеріалів.

Для стабільної роботи оптичних систем у середовищі з абераціями й турбулентністю застосовуються методи адаптивної оптики, орієнтовані на

компенсацію спотворень хвильового фронту за допомогою сенсорів Шека–Гартмана і деформівних дзеркал [14].

У сфері цифрової обробки зображень ключову роль відіграють алгоритми адаптивного підвищення контрастності, вирівнювання гістограм і маскової фільтрації. Ці методи широко висвітлено у працях [15, 16]. Крім того, сучасні методи глибокого навчання відіграють важливу роль у підвищенні ефективності систем технічного зору: використання згорткових нейронних мереж і сучасних архітектур для компенсації руху забезпечує поліпшення точності та стійкості детекції [17].

Таким чином, сучасні підходи до вдосконалення систем технічного зору поділяються на такі, що:

- пропонують удосконалення оптоелектронних компонентів, фотоприймачів, джерел випромінювання та адаптивних оптичних систем для компенсації аберацій і турбулентності;
- базуються на інтелектуальних методах цифрової обробки зображень, машинному навчанні та нейронних мережах для фільтрації шумів, компенсації руху об'єктів і підвищення контрастності;
- передбачають комплексне застосування попередніх двох підходів для забезпечення підвищеної точності, завадостійкості та надійності функціонування систем технічного зору в умовах низької освітленості та нестабільних атмосферних факторів.

У даній випускній кваліфікаційній роботі було розглянуто другий підхід – інтелектуалізацію методів обробки сигналів та зображень задля розширення можливостей оптоелектронних систем. Отже, розробка та вдосконалення методів виявлення об'єктів, зокрема малорозмірних та малоконтрастних, в системах технічного зору є актуальною науково-технічною задачею.

Метою дослідження є вдосконалення процесу виявлення та локації малорозмірних та малоконтрастних об'єктів в оптико-електронних системах

технічного зору, що використовують методи розпізнавання на основі глибокого навчання.

Для досягнення зазначеної мети було визначено наступні задачі дослідження:

- аналіз методів формування сигналів та зображень в оптико-електронних системах, визначення просторових характеристик сигналів, які реєструються;
- огляд існуючих методів та засобів виявлення об'єктів в системах технічного зору, визначення особливостей їх застосування, переваг та недоліків;
- розробка імітаційної моделі поширення та реєстрації сигналів;
- генерація синтетичного набору даних за допомогою розробленої аналітичної моделі, включаючи такі, що імітують близьке розташування об'єктів та складні умови спостереження;
- розробка комбінованого методу виявлення та локації об'єктів, що використовує методи глибокого навчання, навчання та тестування нейронних мереж;
- огляд апаратних засобів для реалізації системи технічного зору, аналіз їх характеристик та можливостей, вибір платформи для розробки системи технічного зору;
- програмна реалізація алгоритму реєстрації об'єктів відповідно до обраного апаратно-програмного середовища;
- експериментальне дослідження результатів моделювання вихідних сигналів фотоприймача для об'єктів, розташованих на різних відстанях, та в складних умовах спостереження;
- експериментальна оцінка якості виявлення та локації об'єктів за допомогою розробленого методу;
- збірка тестового стенду, експериментальне визначення продуктивності системи з використанням обраного апаратного забезпечення.

Гіпотеза дослідження. При недостатній роздільній здатності матриці фотоприймача можуть виникнути складнощі з виявленням близько розташованих об'єктів за пороговими методами через невиконання критерію Релея. Якщо просторові характеристики сигналів відомі та незмінні, виявлення таких об'єктів стане можливим за іншими ознаками, зокрема за виглядом дифракційної картини. Такий метод виявлення має бути стійким до складних умов спостереження, таких як низька контрастність, атмосферна турбулентність та розмиття через рух об'єкта. Використання глибоких нейронних мереж у обробці сигналів має забезпечити високу здатність до узагальнення та стійкість до флуктуацій, що підвищить ефективність методу в складних умовах. Розробка аналітичної моделі відгуку фотоприймача дозволить генерувати синтетичні дані у довільному обсязі, що забезпечить якісне навчання глибоких ШНМ для покращеної системи технічного зору.

Об'єктом дослідження є процеси приймання, реєстрації та обробки сигналів в оптико-електронних системах технічного зору.

Предметом дослідження є методи формування відгуку сенсору в оптоелектронній системі, метод і алгоритм обробки вихідного сигналу (виявлення та локації малорозмірних і малоконтрастних об'єктів).

Методологічна та теоретична основа дослідження. Для побудови моделі реєстрації сигналу оптико-електронною системою було використано елементи хвильової теорії світла, теорії ймовірності та математичної статистики, теорії побудови оптико-електронних систем. Для розробки методу виявлення використовувались теорія множин та теорія розпізнавання образів. Методи дослідження – загальнонаукові: аналіз і синтез, порівняння, абстрагування та моделювання; конкретнонаукові: загальні методи обробки сигналів, методи комп'ютерного зору, методи машинного навчання, зокрема методи обробки «великих даних», методи глибинного навчання, експеримент.

Практичне значення отриманих результатів полягає у створенні ефективної системи технічного зору, що реалізує вдосконалений метод виявлення та локації об'єктів, зокрема малорозмірних та малоконтрастних, у

складних умовах спостереження. Вона може бути використана в широкому спектрі наукових, технічних та прикладних застосувань, включаючи астрономічні спостереження, системи контролю у промисловості та логістиці, у військовій розвідці та системах наведення, сільському господарстві та медицині.

Апробація результатів дослідження. Основні положення дослідження було висвітлено у доповідях на конференції «Електронна, лазерна та біотехнічна інженерія» у рамках 29-го Міжнародного молодіжного форуму «Радіoeлектроніка та молодь у ХХІ столітті» (16–19 квітня 2025 р., м. Харків) та міжнародній науково-технічній конференції «Інформаційно-комунікаційні технології та кібербезпека» ІКТК-2025 (4–5 грудня 2025 р., м. Харків).

1 МЕТОДИ ОБРОБКИ СИГНАЛІВ В СИСТЕМАХ ТЕХНІЧНОГО ЗОРУ

1.1 Структурна схема та методи формування зображень в системах технічного зору

Системи технічного зору (СТЗ) є складними багатокомпонентними комплексами, призначеними для автоматизованого сприйняття, аналізу та інтерпретації візуальної інформації. Типова структурна схема СТЗ включає такі основні елементи: джерело випромінювання, середовище поширення, оптичну ланку (об'єктиви, фільтри, ослаблювачі), фотоприймач (наприклад, ПЗЗ-матриці або фотодіоди), електронний тракт, блок обробки та накопичення даних, а також пристрій прийняття рішення (рис. 1.1).

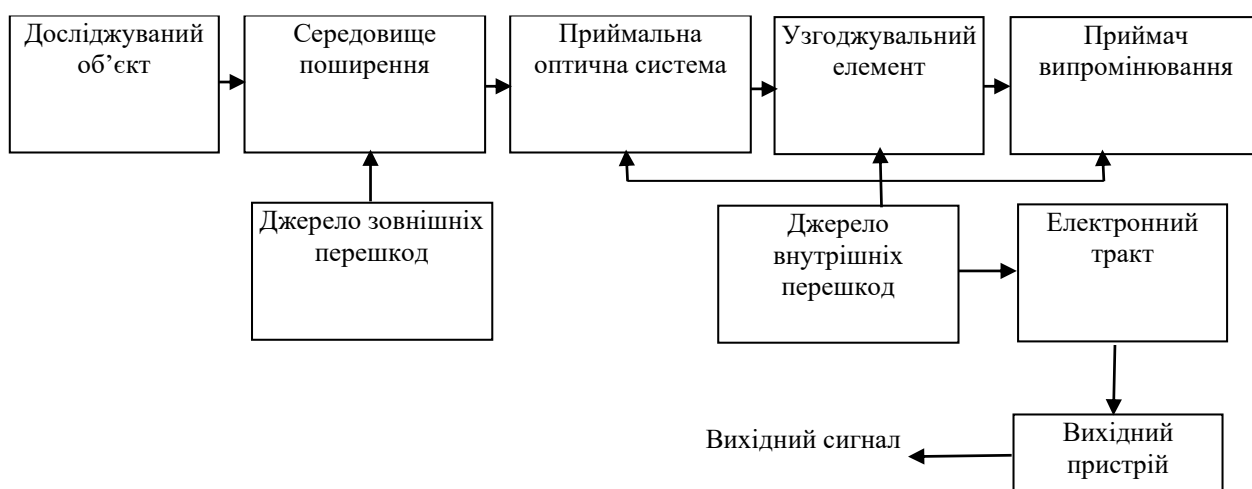


Рисунок 1.1 – Узагальнена схема СТЗ

Оптична система виконує функцію збору та фокусування електромагнітного випромінювання. Її якість визначається параметрами лінз, наявністю дефектів, неоднорідностей, а також впливом атмосферних явищ. Фотоприймач здійснює детектування світлового потоку, при цьому важливими характеристиками є квантова ефективність, спектральна чутливість та

динамічний діапазон. Електронний тракт забезпечує перетворення аналогового сигналу в цифровий, його підсилення та фільтрацію.

Обробка сигналів у СТЗ реалізується за допомогою міжкадрового та внутрішньокадрового накопичення, компресії (наприклад, MPEG, H.264), кореляційних, різницевих, порогових, рангових та wavelet-методів. Останні дозволяють ефективно зменшувати обсяг даних без втрати критичної інформації. Важливим аспектом є оптимізація часу експозиції, що впливає на якість зображення та ймовірність виявлення об'єктів [18].

Методи формування зображень у оптоелектронних системах базуються на фізичних моделях поширення та взаємодії електромагнітного випромінювання з середовищем і детекторною системою. У структурі моделей розрізняють три підходи, що доповнюють один одного: хвильову, корпускулярну та статистичну описові моделі. Кожен підхід дає свій набір апроксимацій, які визначають просторову роздільну здатність, контрастність та стохастичну поведінку кінцевого зображення.

Хвильова теорія описує формування інтерференційних та дифракційних ефектів, що визначають межі просторової роздільної здатності оптичних трактів СТЗ. Вхідні поля в апертурі системи моделюються як комплексні амплітуди; їх розповсюдження до площини фокуса описується інтегральними перетвореннями Фур'є та хвильовими функціями Гріна. Врахування фазових зсувів від оптичних аберацій і атмосферної турбулентності дозволяє прогнозувати змазування точкових джерел і формувати матрицю розмиття (PSF), яку використовують у процедурах деконволюції та підвищення роздільної здатності. Основний принцип моделювання оптичних систем за хвильовою теорією – представлення поля в компактній формі через перетворення Фур'є і взаємозв'язок між апертурою та зображенням у предметній площині. З цього випливає, що дифракційний межа залежить від розміру апертури й довжини хвилі, фазові спотворення призводять до деградації контрасту і зростання похибок локалізації об'єктів [19].

Корпускулярний підхід моделює фотонний потік як дискретний випадковий процес, у межах якого кількість фотонів, що потрапляє на детектор у заданому інтервалі часу, підпорядковується пуассонівській статистиці при незалежних подіях реєстрації. Такий опис важливий для моделювання СТЗ при низькому рівні освітлення або при роботі на межі чутливості, оскільки дискретність природи випромінювання визначає фундаментальний шум фоновому сигналу. При малих середніх числах фотонів дисперсія сигналу дорівнює його середньому, що обмежує можливості усереднення і змушує застосовувати статистичні процедури оцінки сигналу та байєсовські методи відновлення зображення. Це має бути враховане, насамперед при проєктуванні фотоприймачів: необхідність оптимізації квантової ефективності, часу інтеграції та режимів квантування для зниження похибок вимірювання при фонових і темнових флуктуаціях [20, 21].

Статистичні моделі ширше відображають реальні випадки, коли шуми і флуктуації не підпорядковуються простій гаусовій моделі. Для аналізу застосовують класичні та нетрадиційні розподіли:

- Гаусові моделі застосовуються для великих чисел фотонів і коли загальні шумові складові домінують; вони зручні для лінійної аналітики та оптимізації лінійних фільтрів;
- α -стійкі розподіли використовують для опису складних випадків і «вибухоподібних» флуктуацій, характерних для деяких середовищ та несприятливих атмосферних умов. Такі розподіли пояснюють появу великих рідкісних викидів шуму, що негативно впливають на ефективність класичних методів усереднення;
- статистика Бозе-Ейнштейна та квантіфікація кореляцій застосовуються при аналізі когерентних або частково когерентних джерел і в системах, де проявляється статистика квантових полів (наприклад, при використанні лазерного збудження або у системах з підсилювачами);
- моделі парнокорельованих потоків вводяться для опису ситуацій, коли фотони або пакети енергії приходять із кореляціями у часі або просторі;

це впливає на оцінювання локальної варіації інтенсивності та на методи виявлення слабких сигналів на фоні корельованого шуму.

Комбінований підхід поєднує хвильову і корпускулярну складові у рамках ітераційних методів відновлення зображення та статистичних оцінювачів. Реконструкція з урахуванням PSF і пуассонівського шуму реалізується через методи максимальної правдоподібності та апостеріорні байєсові алгоритми, які дозволяють отримати стійкі оцінки при низькому SNR [21]. Адаптивні фільтри і регуляризатори, що враховують нетипові розподіли шуму (α -стійкі), знижують вплив рідкісних викидів і зберігають краєві деталі об'єктів, покращуючи ідентифікацію малоконтрастних структур [22]. Гібридні підходи частіше використовують для застосувань реального часу, де аналітичні хвильові моделі задають грубу модель, а статистичні процедури локально коригують оцінки інтенсивності та ймовірності виявлення.

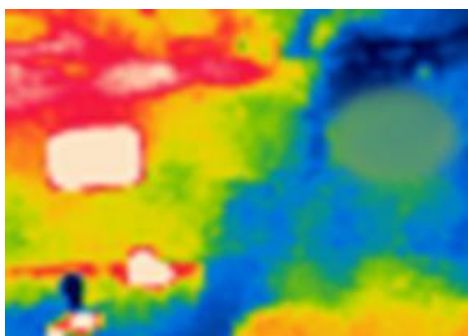
Таким чином, формування та реєстрація сигналів у системах технічного зору є складним процесом, що потребує врахування як фізичних, так і статистичних характеристик середовища, об'єкта та елементів системи.

Практичне застосування оптико-електронних систем охоплює такі галузі, як астрономія, відеоспостереження, телемедицина, спектральний аналіз, біологічна діагностика та військова техніка. Ефективність системи визначається не лише апаратними характеристиками, але й алгоритмічними рішеннями, що враховують стохастичний характер сигналів та обмеження фізичних процесів.

1.2 Особливості формування, просторові характеристики сигналів та зображень, які реєструються системами технічного зору

Формування зображень у системах технічного зору залежить від типу приймача оптичного сигналу та його робочого спектра. Тепловізійні системи реєструють інфрачервоне випромінювання об'єктів, перетворюють його в електричні сигнали і формують термограму – двовимірне поле інтенсивностей з

колірними або градаційними відображеннями, пропорційними локальному температурному розподілу (рис. 1.2 (а)). Телевізійні (видимі) системи виконують аналогічну послідовність перетворень у видимому діапазоні, де оптичне випромінювання фокусують на багатоелементному фотоприймачі, з якого отримують цифрове зображення після підсилення й аналого-цифрового перетворення (рис. 1.2 (б)).



а)



б)

Рисунок 1.2 – Термограма (а) та цифрове телевізійне зображення (б)

Формування і передача просторової інформації у системі описується послідовністю: об'єкт, оптична система, багатоелементний фотоприймач. У цьому ланцюзі геометрична оптика задає відповідність «точка об'єкта – точка зображення», при якій просторові координати об'єкта відображаються в координати матриці фотоприймача. До базових параметрів, що визначають просторові характеристики, належать лінійне поле зору, кутове поле зору та лінійний розмір елемента фотоприймача [18]. Значення цих величин визначають покриття сцени, кутове розділення та розмір відтворюваних об'єктів у пікселях, що, у сукупності з оптичною PSF, визначає розпізнавальні можливості системи.

Лінійна роздільна здатність фотоприймача визначається співвідношенням між лінійним розміром фотокатода (або активної площі приймача) і кількістю елементів у напрямку сканування. Формально:

$$R = \frac{L_{\phi n}}{N} \quad (1.1)$$

де $L_{\phi n}$ – лінійний розмір фотоприймача;

N – кількість елементів у відповідному напрямку. Кількість елементів, що фіксують зображення, оцінюється як

$$N = \frac{L_{зобр}}{L_{елем}}, \quad (1.2)$$

де $L_{зобр}$ – лінійний розмір зображення в площині фотоприймача;

$L_{елем}$ – розмір одного елемента.

Ці співвідношення використовуються для розрахунку параметрів матриць сенсорів та визначення просторової роздільності системи при заданих оптичних масштабах і геометричних розмірах оптики.

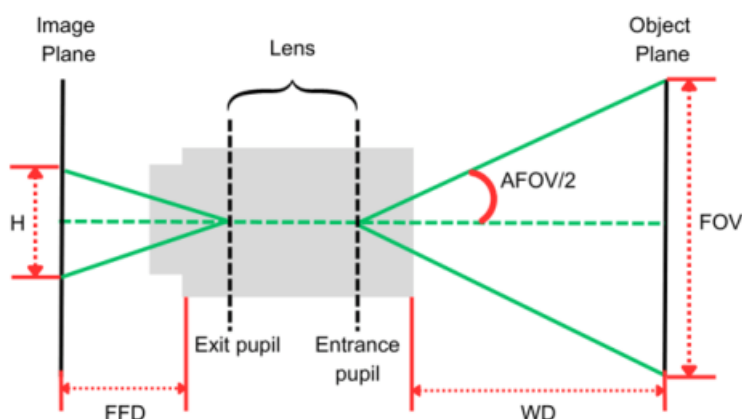


Рисунок 1.3 – Лінійне поле зору (FOV) та кутове поле зору (AFOV) телевізійної чи тепловізійної системи

Кутове поле зору (КПЗ) – це кутова ширина сцени, що охоплюється системою по відповідному напрямку, і є фундаментальною оптичною характеристикою, що залежить від розмірів апертури сенсора і фокусної відстані (рис. 1.2). Для прямокутного сенсора у напрямку вимірювання:

$$\theta = 2 \arctan\left(\frac{L_{\phi n}}{2f}\right), \quad (1.3)$$

де $L_{\phi n}$ – лінійний розмір активної області сенсора вздовж розглянутого виміру;

f – фокусна відстань об'єктива.

Для малих відношень $\frac{L_{\text{sensor}}}{2f}$ використовується лінійне наближення

$$\theta \approx L_{\phi n} / f .$$

Кутове поле зору пов'язане з кутовою роздільною здатністю системи – мінімальною кутовою відстанню між двома джерелами, які можуть бути розрізнені. Кутова роздільна здатність при дифракційному обмеженні апертури діаметром $D_{\text{ап}}$ оцінюється через критерій Релея:

$$\theta_{\text{дифр}} \approx 1,22 \frac{\lambda}{D_{\text{ап}}}, \quad (1.4)$$

де λ – робоча довжина хвилі.

Отже, реальна кутова роздільність системи визначається максимумом із двох факторів: оптичного (дифракційного/абераційного) та квантово-детекторного (дискретизація сенсора) [23].

КПЗ використовується для проектування кадрів, вибору об'єктива і визначення відношення між кутовими розмірами об'єктів на сцені та їхнім представленням на сенсорі; це критично для задач виявлення і геометричного вимірювання об'єктів.

Лінійне поле зору (ЛПЗ) визначається як лінійний розмір сцени, що проектується на площину фотоприймача при фіксованій відстані до об'єкта D (рис. 1.3). Для оптичної системи з фокусною відстанню f і габаритним лінійним розміром сенсора в напрямку вимірювання $L_{\phi n}$ (активна площа в цьому напрямку) лінійне поле зору на відстані D виражається через кутове поле зору θ або безпосередньо через геометрію проєкції:

$$L_{\phi n}(D) = 2D \tan\left(\frac{\theta}{2}\right), \quad (1.5)$$

де θ - кутове поле зору.

$$\theta = 2 \arctan\left(\frac{L_{\phi n}}{2f}\right). \quad (1.6)$$

Комбінуючи, одержимо корисне вираження через параметри сенсора та фокусної відстані:

$$L_{nz}(D) = 2D \tan\left(\arctan\left(\frac{L_{\phi n}}{2f}\right)\right) = 2D \frac{L_{\phi n}}{2f} / \sqrt{1 + \left(\frac{L_{\phi n}}{2f}\right)^2}, \quad (1.7)$$

що при малих кутових наближеннях ($L_{\phi n} \ll 2f$) спрощується до лінійної пропорційності:

$$L_{nz}(D) \approx D \frac{L_{\phi n}}{f}. \quad (1.8)$$

ЛПЗ служить базовою величиною для визначення покриття сцени й оцінки, скільки метрів сцени припадає на матрицю сенсора при заданій відстані до об'єкта; вона використовується для розрахунку лінійної роздільної здатності та для планування експериментів або вимірювальних процедур.

Прив'язка лінійного розміру пікселя до лінійної роздільності сцени на відстані D (ground sample distance, GSD) здійснюється через оптичний масштаб:

$$\text{GSD} = \frac{D}{f} p \quad (\text{при } D \gg f), \quad (1.9)$$

або, використовуючи лінійне поле зору L_{nz} :

$$\text{GSD} = \frac{L_{nz}(D)}{N} = \frac{L_{nz}(D)p}{L_{\phi n}}. \quad (1.10)$$

Цей вираз показує, яка лінійна частина сцени на відстані D відповідає одному пікселю; зменшення p або збільшення f поліпшує просторову характеристику, але при цьому слід враховувати дифракційні обмеження та співвідношення сигнал/шум.

Критерій дискретизації (Найквіста) для коректного відновлення просторових частот вимагає, щоб крок дискретизації (піксель) забезпечував не менше ніж два відліки на мінімально значиму просторову структуру, наприклад, на діаметр центрального максимуму плями розсіювання (Airy disk) у площині зображення. Діаметр центрального максимуму (по інтенсивності першого мінімуму) в площині зображення дорівнює:

$$d_{\text{розс}} = 2.44 \frac{\lambda f}{D_{an}}. \quad (1.11)$$

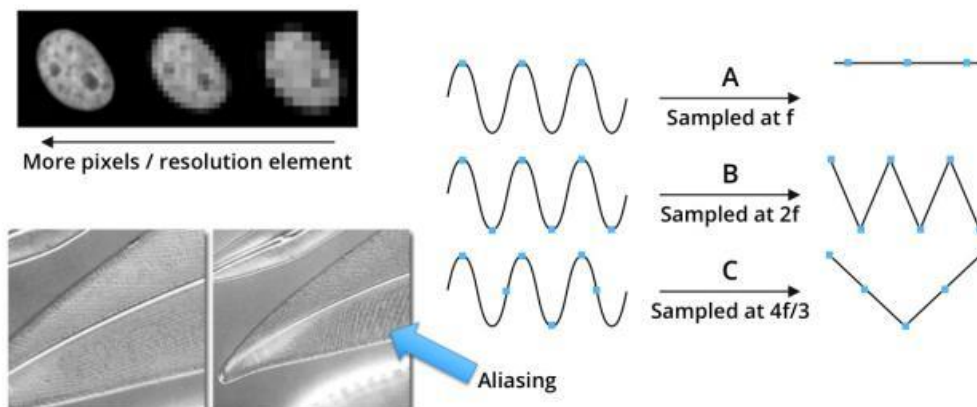


Рисунок 1.4 – Вплив дискретизації на вигляд зображення

Умова Найквіста формулюється як:

$$p \leq \frac{d_{\text{розс}}}{2} = 1.22 \frac{\lambda f}{D_{\text{ап}}}, \quad (1.12)$$

що гарантує, що оптична деталь не буде недодискретизована фотоприймачем. Невиконання цієї умови призводить до аліазингу й штучної втрати просторової інформації, а надто дрібні пікселі при недостатньому світловому потоці можуть, навпаки, погіршити SNR через розділення фотонів між великою кількістю елементів (рис. 1.4) [24].

Структура кінцевого сигналу, що надходить із фотоприймача, складається з кількох складових (рис. 1.5):

- фонові складова, яка може бути просторово змінною внаслідок неоднорідності сцени та фону оптики;
- сигнальна складова, що представлена корисним випромінюванням від об'єкта і відносно стабільна в часі в умовах стаціонарного спостереження;
- флуктуації фону та сигналу, які змінюються у часі і просторі і визначаються як статистикою випромінювання, так і характеристиками приймача.

Вихідний електричний сигнал при фотонній квантовій природі випромінювання є стохастичним процесом; у класичній постановці для

незалежних випадків реєстрації кількість фотоелектронів підпорядковується пуассонівському закону, при якому математичне сподівання рівне дисперсії і дорівнює середньому числу фотоелектронів за інтервал інтеграції [25]. Наслідком цього є фундаментальне обмеження SNR при низьких рівнях освітленості і специфічні правила побудови критеріїв виявлення та оцінювання похибок.

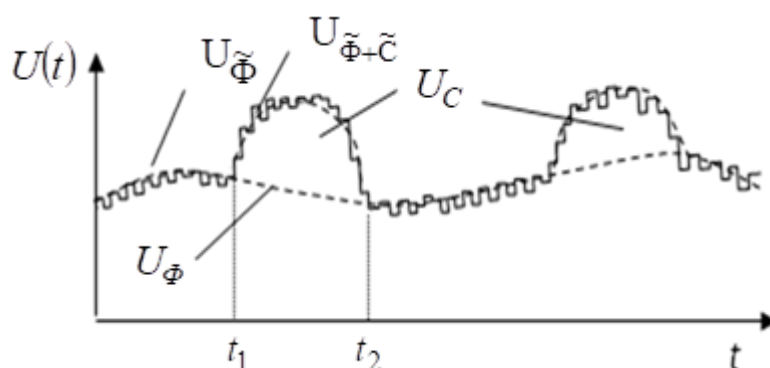


Рисунок 1.5 – Вихідний сигнал тепловізійних та телевізійних систем при спостереженні випромінювання, відбитого від об'єктів

Істотну роль відіграють внутрішні та зовнішні шуми, які знижують співвідношення сигнал/шум (SNR) на виході фотоприймального модуля (рис. 1.5). Інтенсивність внутрішніх шумів визначається фізико-технологічними параметрами фотоприймача, температурним режимом і швидкістю зчитування сигналу. Зовнішні шуми та завади, у свою чергу, можуть мати природне (фонове випромінювання, сонячні відблиски, атмосферні розсіяння) або техногенне походження (електромагнітні випромінювання, відбиття від конструктивних елементів, паразитні сигнали). Їх дія призводить до спотворення амплітудно-частотного спектра корисного сигналу, що ускладнює його подальшу обробку та інтерпретацію [26].

Джерела шуму в системах технічного зору включають:

- випромінювання фону сцени та навколишнього середовища, що додає рівномірний або просторово варіабельний фон;

- власне випромінювання середовища й оптичних елементів, що особливо важливо в ІЧ-діапазоні;
- шуми приймача та підсилювача, включно з термічним та електронним шумом схем обробки;
- флуктуації параметрів елементів системи (температурні, електричні, механічні), які вносять додаткову нестабільність у часі і просторі.

Додаткову складність створює рух об'єкта спостереження або рух самої системи технічного зору. Навіть незначне зміщення під час експозиції призводить до розмиття зображення та зниження просторової роздільної здатності. У високошвидкісних сценах спостерігається ефект динамічної інтерференції, який ускладнює точне визначення контурів об'єктів і суттєво знижує ефективність алгоритмів розпізнавання.

Суттєвий вплив на якість зображення має турбулентність атмосфери, що спричиняє випадкові флуктуації показника заломлення повітря. Це призводить до фазових спотворень фронту хвилі, виникнення блукань зображення (image dancing) та зміни контрасту внаслідок інтерференційних ефектів. Такі спотворення особливо помітні під час спостереження на великих відстанях у тепловому та видимому діапазонах і вимагають застосування адаптивної оптики або цифрових методів корекції фазових флуктуацій.

Іншими джерелами деградації сигналу є оптичні аберації, що зумовлені недосконалістю геометрії оптичних елементів або їх неправильним юстуванням. Сферичні, коматичні та астигматичні аберації спотворюють форму зображення точкових об'єктів, знижують контрастність та ускладнюють точну оцінку просторового розподілу інтенсивності.

Отже, для ефективної обробки сигналу тепловізійної чи телевізійної системи та виявлення об'єктів необхідно використовувати методи, що враховують стохастичну природу сигналу, його просторові характеристики та чинники, що спричиняють шуми.

1.3 Методи та алгоритми для реєстрації об'єктів в системах технічного зору

Сучасні системи технічного зору ґрунтуються на різних математичних, алгоритмічних і апаратних підходах, що дозволяють автоматично ідентифікувати, локалізувати, сегментувати та відстежувати об'єкти у зображеннях чи відеопотоці. Для виявлення слабоконтрастних або малопомітних об'єктів застосовуються квазіоптимальні виявники, методи просторово-частотної фільтрації, накопичення кадрів, а також оцінка SNR у реальному часі тощо. [27].

Розробка методів і алгоритмів оброблення сигналів, орієнтованих на покращення якості зображень рухомих і близько розташованих малорозмірних, мало-контрастних об'єктів, спирається на використання таких підходів, як хвильовий і корпускулярний опис сигналів, стохастико-детерміновані моделі, а також сучасні методи просторово-часової, порогової, кореляційної та сегментаційної обробки [28]. На сьогоднішній день універсального методу виявлення не існує, тому вибір методу здійснюється з урахуванням параметрів системи технічного зору та поставленої задачі. Умовно методи реєстрації об'єктів в системах технічного зору можна поділити на:

- а) порогові методи: класичні порогові алгоритми виявлення, порогові методи у поєднанні з просторовою фільтрацією;
- б) кореляційні методи, зокрема для слабких сигналів за наявності апіорної інформації;
- в) методи сегментації та виділення об'єктів: сегментація з використанням ознак, специфічних для об'єктів, різницеві методи для виділення рухомих об'єктів у послідовності кадрів, методи виділення об'єктів із використанням лінійної і нелінійної просторової фільтрації;
- г) методи просторово-часової обробки: міжкадрова та внутрішньокадрова обробка сигналів;

д) стохастико-детерміновані методи: фільтрація шумових складових на основі хвильового чи комбінованого хвильово-корпускулярного підходу, моделювання вихідного сигналу як некорельованого гаусового випадкового процесу, методи обробки сигналів із урахуванням додаткових характеристик зареєстрованих сигналів.

Порогові методи становлять основоположний клас алгоритмів сегментації і реєстрації об'єктів у цифровому зображенні. Вони забезпечують швидкий і простий спосіб відокремлення об'єкта від фону за інтенсивністю. Порогова обробка є математично детермінованим інструментом сегментації за інтенсивністю. Вхідним є зображення у градаціях сірого кольору, у якому інтенсивність $I(x,y)$ вказує на світлість чи яскравість пікселя. Обирається певний поріг T (або набір порогів у багаторівневих випадках). Простий поріг розділяє пікселі за критерієм:

$$I(x,y) \geq T \quad (1.13)$$

Таким чином, вихідним результатом є двійкова (або багатовимірна) маска, в якій відокремлено ті області, де інтенсивність перевищує поріг, від тих, що її не перевищують (рис. 1.6) [29].

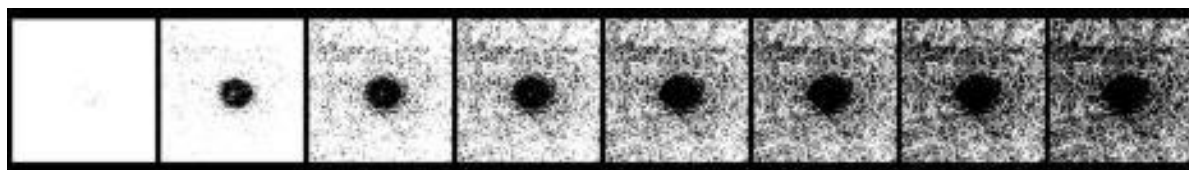


Рисунок 1.6 – Проста порогова обробка з різними значеннями порогу

Даний метод є ефективним якщо фон і об'єкт мають відносно чітко відмінні інтенсивності. Зокрема, він широко застосовується у задачах технічного зору для контролю якості, коли контраст між об'єктом і фоном високий. Однак коли умови складніші (неоднорідне освітлення, невелика різниця в інтенсивностях об'єкта і фону, наявність шуму), стають необхідними вдосконалення: автоматичний вибір порогу (наприклад, метод Метод Оцу), адаптивне порогоування, локальне порогоування тощо.

Задля того, щоб підвищити стійкість алгоритму до шуму і неоднорідності освітлення, перед застосуванням порогування часто застосовують попередню обробку зображення (розмиття, корекція освітлення, підсилення контрасту) або адаптацію порогу до локальних статистик (локальне середнє, стандартне відхилення).

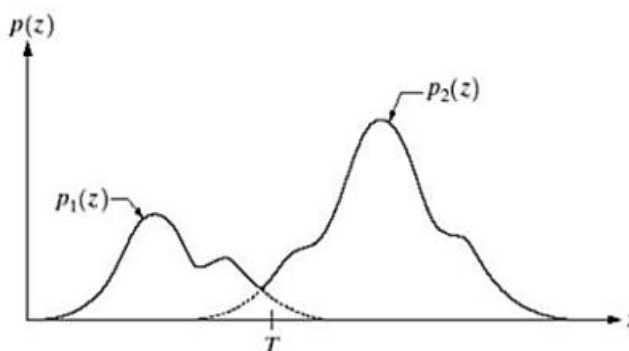


Рисунок 1.7 – Гістограма розподілу інтенсивності для зображення з фоновною та сигнальною складовими

На рисунку 1.7 зображено імовірнісні щільності яскравості двох класів пікселів зображення $p_1(z)$ - для фону та $p_2(z)$ для об'єкта. Уздовж осі z відкладаються рівні яскравості (градації сірого), а вісь $p(z)$ представляє щільність імовірності появи цих рівнів. Завдання порогового розділення полягає у виборі такого значення порогу T , при якому пікселі можна оптимально класифікувати:

$$\begin{cases} z < T \Rightarrow \text{піксель належить до класу } \omega_1, \\ z \geq T \Rightarrow \text{піксель належить до класу } \omega_2. \end{cases} \quad (1.14)$$

У цьому контексті ймовірність правильного виявлення (детекції) P_d визначається як імовірність того, що піксель, який належить до об'єкта (ω_2), буде класифікований правильно:

$$P_d = \int_T^{z_{\max}} p_2(z) dz. \quad (1.15)$$

Водночас ймовірність хибного виявлення (false alarm) P_{fa} визначає ймовірність того, що піксель фону (ω_1) помилково буде віднесено до об'єкта, тобто його значення яскравості також перевищить поріг T :

$$P_{fa} = \int_T^{z_{\max}} p_1(z) dz. \quad (1.16)$$

Обидві ці ймовірності залежать від вибраного порогу T . Зменшення T підвищує P_d , але одночасно збільшує P_{fa} . Навпаки, надмірно високий поріг зменшує кількість хибних спрацьовувань, проте може призвести до втрати об'єктів.

Оптимальний поріг T^* зазвичай вибирають шляхом мінімізації сумарної помилки класифікації або максимізації ймовірності правильного рішення. За рівноймовірної появи об'єкта і фону – $P(\omega_1) = P(\omega_2) = 0.5$. Тоді оптимальний поріг визначається розв'язанням рівняння:

$$p_1(T^*) = p_2(T^*), \quad (1.17)$$

тобто у точці перетину щільностей ймовірності обох класів. Цей підхід лежить в основі таких методів, як метод Отсу, який обчислює значення порога на основі мінімізації внутрішньокласової дисперсії.

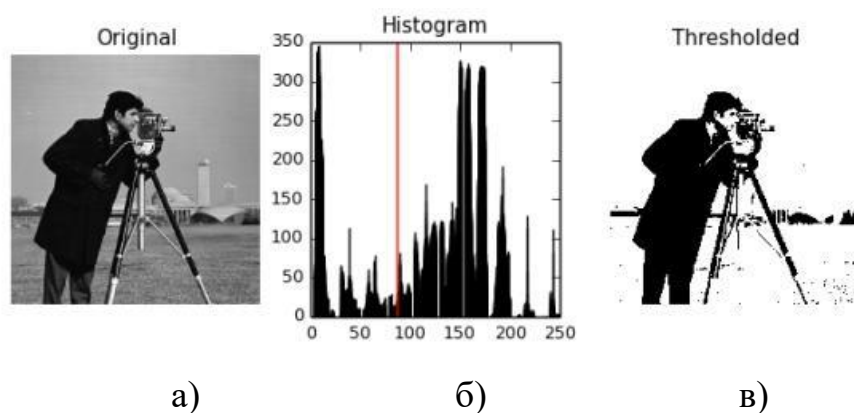
Метод Оцу [30] є класичним підходом до автоматичного вибору порогу сегментації, який ґрунтується на аналізі гістограми зображення (рис. 1.8). Його суть полягає у знаходженні такого порогового значення T^* , яке мінімізує внутрішньокласову дисперсію або, еквівалентно максимізує міжкласову дисперсію двох груп пікселів – фону та об'єкта. Нехай $\omega_1(T)$ і $\omega_2(T)$ – частки пікселів, що належать відповідно до класів з інтенсивністю менше і більше за поріг T , а $\mu_1(T)$ та $\mu_2(T)$ – їх середні інтенсивності. Тоді міжкласова дисперсія визначається як:

$$\sigma_B^2(T) = \omega_1(T) \omega_2(T) [\mu_1(T) - \mu_2(T)]^2. \quad (1.18)$$

Знаходження оптимального порогу здійснюється за правилом:

$$T^* = \arg \max_T \sigma_B^2(T), \quad (1.19)$$

що гарантує найкраще відокремлення двох класів за інтенсивністю. Тому метод Отцу ефективний для зображень із бімодальною гистограмою.



а) вхідне зображення (відтінки сірого), б) гистограма та порогове значення, в) бінарне зображення

Рисунок 1.8 – Визначення порогового значення за методом Отцу

У випадках, коли вибір єдиного (глобального) порогу не є ефективним через неоднорідність освітлення, може бути застосоване локальне порогоування, коли зображення поділяється на окремі регіони (блоки), і для кожного з них обчислюється власний поріг на основі локальної гистограми або середнього значення яскравості. Порогування в межах кожного блоку здійснюється незалежно, що дозволяє покращити якість сегментації при значних просторових варіаціях освітлення. На відміну від глобального методу, локальне порогоування не потребує однакової інтенсивності фону по всьому зображенню, проте його результат залежить від розміру вибраного вікна та ступеня перекриття між регіонами.

Адаптивне порогоування є узагальненням локального підходу з урахуванням статистичних характеристик околиці кожного пікселя. У цьому випадку порогове значення $T(x, y)$ може обчислюватися як

$$T(x, y) = \mu_{local}(x, y) + k \sigma_{local}(x, y), \quad (1.20)$$

де μ_{local} - локальне середнє значення яскравості у вікні навколо пікселя;

σ_{local} - локальне стандартне відхилення;

k - емпіричний коефіцієнт чутливості.

Такий підхід дозволяє компенсувати вплив неоднорідного освітлення й локальних змін контрасту, забезпечуючи стабільніше виділення об'єктів у складних умовах реєстрації (рис. 1.9). Адаптивне порогування є більш ефективним при наявності тіней, градієнтів освітлення чи часткових відблисків.



а)

б)

в)

а) вхідне зображення, б) застосування глобального порогу, в)
застосування адаптивного порогу

Рисунок 1.9 – Глобальне та локальне порогування

Після застосування порогової обробки може застосовуватись морфологічна пост-обробка, що забезпечує видалення шумових компонент, закриття «острівців», видалення дрібних компонентів, що не належать об'єкту.

Перевагами порогових методів є простота реалізації й низька обчислювальна складність (що важливо для систем реального часу). Вони працюють ефективно, коли об'єкт і фон мають чітку різницю інтенсивності та однорідне освітлення.

Основними обмеженнями є чутливість до неоднорідного освітлення (якщо фон або освітлення змінюються по сцені, глобальне порогування не дає

адекватного розділення), неефективність при низькому контрасті об'єкта і фону або при наявності шуму чи артефактів. Також методи, засновані на порогованні, не враховують просторової інформації – один піксель оцінюється лише за інтенсивністю, без урахування контексту. Також вони не є ефективними для сегментації складних сцен, особливо з декількома класами інтенсивностей чи сильно змінною освітленістю [3, 31].

У задачах реєстрації об'єктів цей підхід застосовується, як правило, коли об'єкт має достатній контраст до фону. Якщо ж контраст дуже низький або об'єкт дуже маленький, просте пороговання може виявитися неефективним. Однак в складних умовах реєстрації малоконтрастних, рухомих або близько розташованих об'єктів застосування даного підходу найчастіше доповнюється або замінюється більш складними методами (адаптивні пороги, комбіновані методи, статистичні чи кореляційні підходи).

Кореляційні методи базуються на вимірюванні ступеня подібності між шаблоном (еталоном) $T(x, y)$ і фрагментами аналізованого зображення $I(x, y)$ [32]. Ідея полягає в тому, щоб співставити шаблон з кожною ділянкою зображення і обчислити кореляційну (або крос-кореляційну) функцію, яка оцінює, наскільки вікно зображення співпадає із шаблоном (рис. 1.10). В області зображення, де цей коефіцієнт досягає високого значення, розташований відповідний об'єкт.

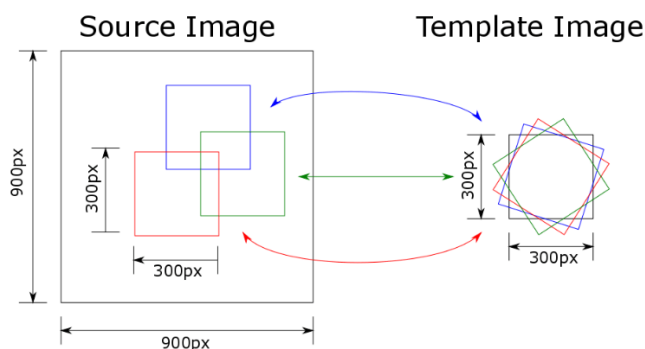


Рисунок 1.10 – Співставлення з еталоном згідно кореляційного підходу

Класична крос-кореляція (без нормалізації) між шаблоном $T(x, y)$ розміром $M \times N$ і фрагментом зображення $I(x+i, y+j)$ може бути записана як:

$$C(i, j) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} T(x, y) \cdot I(x+i, y+j). \quad (1.21)$$

Така форма чутлива до зміни амплітуди сигналу, до освітлення, контрасту. Тому застосовують нормалізовану крос-кореляцію (NCC):

$$NCC(i, j) = \frac{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} (T(x, y) - \bar{T}) (I(x+i, y+j) - \bar{I}_{i,j})}{\sqrt{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} (T(x, y) - \bar{T})^2 \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} (I(x+i, y+j) - \bar{I}_{i,j})^2}}, \quad (1.22)$$

де $\bar{T}$ - середнє значення шаблону;

$\bar{I}_{i,j}$ - середнє значення відповідного фрагмента зображення згідно зі зсувом (i, j) .

Значення NCC лежить в діапазоні $[-1, +1]$, ближче до $+1$ відповідає вищій подібності.

Після побудови матриці NCC обирається пік (значення максимуму) або поріг, при якому у точці (i^*, j^*) є об'єкт, якщо

$$NCC(i^*, j^*) \geq \tau, \quad (1.23)$$

де τ - заданий поріг.

Щоб врахувати зміни масштабу s і обертання θ , часто використовують модифіковану задачу:

$$NCC_{s,\theta}(i, j) = NCC(T_{s,\theta}, I_{s,\theta}(x+i, y+j)), \quad (1.24)$$

де $T_{s,\theta}$ - шаблон, змінений за масштабом та обертанням.

Такий підхід використання методів нормалізованої кореляції дозволяє враховувати зміни ракурсу (обертання) та масштабу шаблонів [33].

Кореляція в просторі може бути обчислена через множення в частотній області завдяки властивості перетворення Фур'є:

$$C = T * I \Leftrightarrow \mathcal{F}\{C\} = \mathcal{F}\{T\} \cdot \mathcal{F}\{I\}. \quad (1.25)$$

Таке представлення дозволяє значно зменшити обчислювальну складність при великих розмірах зображення чи шаблону та забезпечує високу стійкість до шуму [34].

Кореляційний підхід часто комбінують з рядом додаткових алгоритмів з метою оптимізації: зменшення розмірності даних (наприклад, через суб-семплінг чи використання меншої роздільної здатності), попередня фільтрація (наприклад, гаусового розмиття або підсилення контрасту) для зниження впливу шуму, застосування обмеженого пошуку (наприклад, лише в області інтересу). Це дозволяє підвищити швидкість алгоритму кореляційного пошуку та знизити хибні спрацьовування [35].

Кореляційні методи застосовуються в системах технічного зору для задач виявлення об'єктів із відомою формою чи текстурою. Кореляційний підхід особливо корисний, коли є апріорна інформація про форму об'єкта чи його текстуру, тобто шаблон заданий. При цьому метод є чутливим до змін масштабу, обертання, деформацій і освітлення, але сучасні модифікації дозволяють значно підвищити стійкість до цих факторів.

Перевагами кореляційних підходів є висока стійкість до шуму за умови нормалізації та попередньої обробки та можливість визначення положення (зсуву) шаблону у сцені. Основними обмеженнями є залежність від апріорної інформації (шаблон повинен бути відомий і адекватно співпадати з об'єктом), висока обчислювальна складність при великих розмірах зображення чи значних варіаціях, чутливість до змін форми, освітлення, перспективи або часткового накриття об'єкта. Відповідно, при рості розміру зображення ефективність класичних алгоритмів падає і потрібними стають додаткові оптимізації (наприклад, алгоритмічні прискорення, обмеження простору пошуку).

Сегментаційні методи спрямовані на розбиття зображення на логічно відокремлені області (сегменти), які відповідають окремим об'єктам, фону або текстурам. Ця задача передбачає, що кожна точка зображення (піксель або група пікселів) належить одному з класів, і мета алгоритму – визначити ці класи на основі характеристик (яскравості, кольору, текстури, просторового контексту).

У типових підходах до сегментації застосовуються: кластеризація, регіональний ріст, метод водорозділу, графові алгоритми, методи на основі штучних нейронних мереж (ШНМ). Ці підходи здатні задовольнити вимоги до виділення областей інтересу у складних сценах – нерівномірне освітлення, складні текстури, присутність шуму чи руху.

Кластеризація – метод, що групує пікселі в класи на основі їхньої подібності у просторі ознак. У методі кластеризації k-means (рис. 1.11 (а)) задається кількість кластерів K , після чого пікселі представлено як вектори ознак x_i . Розбиття на класи здійснюється шляхом мінімізації функції:

$$J = \sum_{i=1}^N \sum_{k=1}^K u_{ik} \|x_i - c_k\|^2, \quad (1.26)$$

де c_k - центр k-го кластера;

$u_{ik} = 1$ якщо x_i належить кластеру k , інакше 0 [36].

Метод Fuzzy C-Means використовує математичний апарат нечіткої логіки. В ньому присутня нечітка приналежність $u_{ik} \in [0,1]$ і мінімізується функція:

$$J_{FCM} = \sum_{i=1}^N \sum_{k=1}^K u_{ik}^m \|x_i - c_k\|^2, \quad (1.27)$$

де $m > 1$ - параметр нечіткості [37].

Метод регіонального росту (region growing) збільшує регіони, починаючи з початкових точок, приєднуючи суміжні пікселі, якщо вони задовольняють критеріям однорідності (рис. 1.3.6 (б)):

$$|I(x, y) - \mu_R| < T_r, \quad (1.28)$$

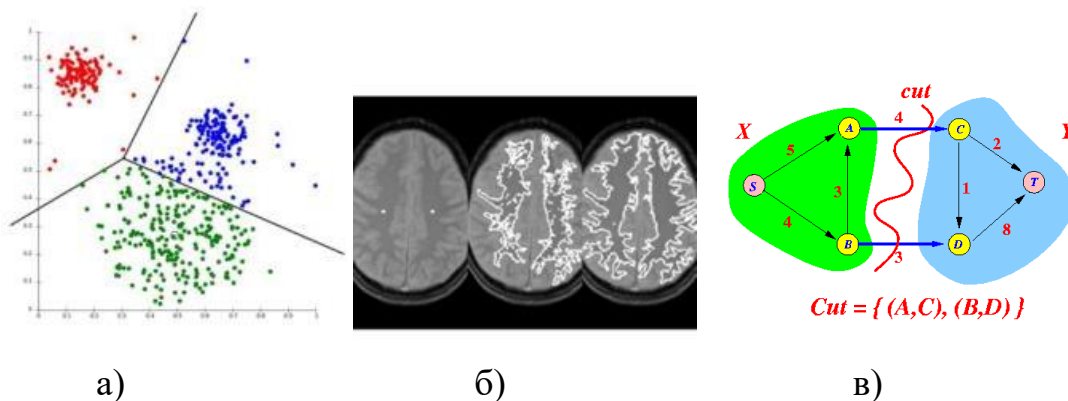
де μ_R - середнє значення пікселів у регіоні R ;

T - порогова межа [38].

Графові алгоритми (рис. 1.11 (в)), наприклад Normalized Cuts [39], засновані на представленні зображення у вигляді графа $G = (V, E)$, де пікселі / суперпікселі – вершини (V), ребра – зв'язки за подібністю (w_{ij}). Задача розбиття зводиться до мінімізації вартості розрізу між сегментами:

$$Ncut(A, B) = \frac{\sum_{i \in A, j \in B} w_{ij}}{\sum_{i \in A, j \in V} w_{ij}} + \frac{\sum_{i \in B, j \in A} w_{ij}}{\sum_{i \in B, j \in V} w_{ij}}, \quad (1.29)$$

де A і B – частини розбиття.



а) кластеризація k-means, б) регіонального росту, в) на основі розрізання графової моделі

Рисунок 1.11 – Методи сегментації

Сучасні нейромережеві підходи використовують згорткові моделі, такі як Fully Convolutional Network (FCN) – мережі, які приймають зображення довільного розміру і видають карта-маску сегментації піксель-до-пікселя. Навчання мережі може здійснюватись за допомогою бінарної крос-ентропічної функції втрат. Застосування згорткових нейронних мереж дозволяє автоматично здійснювати виділення об'єктів складної форми [40, 41].

Отже, сегментаційні методи становлять важливий блок сучасної реєстрації об'єктів в системах технічного зору – вони забезпечують високий рівень деталізації та можливість виділити об'єкти складної форми. Основними недоліками часто є висока обчислювальна складність (особливо для графових і нейромережевих методів), а також необхідність тонкого підбору параметрів (таких як кількість кластерів, пороги росту регіону, гіперпараметри ШНМ) і наявність достатніх даних для навчання (у випадку використання нейронних мереж).

Стохастико-детерміновані методи реєстрації об'єктів поєднують у собі дві концептуально різні парадигми, що доповнюють одна одну: ймовірнісне (стохастичне) моделювання та детерміновані підходи до аналізу сцени. Стохастична компонента описує невизначеність, шум, випадкові збурення та неповноту даних, тоді як детермінована частина формалізує відомі закономірності, фізичні властивості об'єктів або чітко задані правила прийняття рішень. Такий синтез дозволяє будувати алгоритми, здатні забезпечити стабільне виявлення та відстеження об'єктів у складних умовах: змінного освітлення, наявності значних шумів, низької контрастності, часткового перекриття об'єктів.

Стохастична частина алгоритмів базується на ймовірнісному описі стану об'єкта. У загальному вигляді процес спостереження можна подати як систему рівнянь:

$$\begin{cases} x_k = f(x_{k-1}) + w_k \\ z_k = h(x_k) + v_k \end{cases}, \quad (1.30)$$

де x_k - вектор стану об'єкта на кроці k ;

z_k - вектор вимірювань;

$f(\cdot)$, $h(\cdot)$ - детерміновані функції переходу стану та спостереження;

w_k і v_k - випадкові збурення, що моделюють шум системи та вимірювання.

Оцінка стану здійснюється за допомогою фільтра Калмана (у разі лінійних систем із гаусовим шумом) або його нелінійних модифікацій, таких як розширений фільтр Калмана чи частковий фільтр. Ймовірність справжнього стану оцінюється за формулою Байєса:

$$p(x_k | z_{1:k}) = \frac{p(z_k | x_k) p(x_k | z_{1:k-1})}{p(z_k | z_{1:k-1})}, \quad (1.31)$$

де $p(x_k | z_{1:k})$ - апостеріорна оцінка стану;

$p(z_k | x_k)$ - модель вимірювань;

$p(x_k | z_{1:k-1})$ - апостеріорна ймовірність попереднього стану.

Детермінована компонента полягає у формалізації правил прийняття рішення щодо виявлення чи підтвердження наявності об'єкта. Наприклад, використовується пороговий критерій

$$D(x, z) < T, \quad (1.32)$$

де $D(x, z)$ - міра відхилення між оціненим станом та спостереженням;

T - детермінований поріг, визначений на основі фізичної моделі чи апіорних даних.

У роботах [42, 43] розглянуто використання марковських випадкових полів (MRF) та байєсівських мереж для сегментації і виявлення об'єктів в умовах високого рівня шуму. Марковська модель формалізує просторові залежності між пікселями, визначаючи апіорну ймовірність стану пікселя через його сусідів:

$$P(x_i | X_{\text{сусід}}) = \frac{1}{Z} \exp \left(- \frac{U(x_i, X_{\text{сусід}})}{T} \right), \quad (1.33)$$

де U - енергетична функція;

T - параметр «температури»;

Z - нормуючий множник.

Актуальними є комбіновані алгоритми, у яких стохастична оцінка стану об'єкта з фільтра Калмана доповнюється детермінованим критерієм відбору кандидатів за морфологічними ознаками. Таким чином, ймовірнісна частина фільтрує шум та прогнозує рух об'єкта, а детермінована – уточнює рішення на основі геометричних параметрів [44]. Такий підхід виявився ефективним у задачах реєстрації об'єктів у складних сценах. Для кількісної оцінки використовуються показники ймовірності виявлення та ймовірності хибного виявлення:

$$P_d = \frac{N_{\text{виявлених істинно}}}{N_{\text{усіх істинних об'єктів}}}, \quad P_f = \frac{N_{\text{хибних спрацьовувань}}}{N_{\text{всього фоновому сигналу}}}, \quad (1.34)$$

У роботі [45] описано основи байєсівського оцінювання станів, що дозволяють інтегрувати випадкові та детерміновані компоненти в єдину модель. Оптимізаційна постановка комбінованих методів передбачає мінімізацію узагальненого функціоналу втрат:

$$L(\theta) = \mathbb{E} \left[\| z_k - h(f(x_{k-1}, \theta)) \|^2 \right] + \lambda \Phi(\theta), \quad (1.35)$$

де θ – параметри моделі;

перший доданок відповідає стохастичній похибці оцінки;

другий доданок відповідає детермінованому регуляризатору, що покращує стабільність моделі.

Таким чином, стохастико-детерміновані моделі забезпечують підвищену стійкість до шумів і змін умов спостереження, що робить їх незамінними для систем технічного зору. Вони ефективно компенсують недоліки суто детермінованих методів, які не враховують випадковість, і водночас уникають надмірної чутливості до шуму, властивій суто стохастичним моделям. Комбінування байєсівських підходів, марковських полів і детермінованих критеріїв оптимізації дає змогу побудувати гнучкі алгоритми з адаптивним балансом між точністю, обчислювальною складністю та стабільністю.

Просторово-часові методи належать до класу алгоритмів, які одночасно аналізують просторові (геометричні, текстурні, контурні) та часові (динамічні, кінематичні) характеристики об'єктів. На відміну від статичних методів, вони розглядають зміну стану сцени у часі, що дозволяє відслідковувати рух, деформації, появу чи зникнення об'єктів (рис. 1.12). У загальному випадку послідовність кадрів $I(x, y, t)$ описує тривимірне поле інтенсивностей у просторі (x, y) і часі t . Завдання реєстрації зводиться до оцінки функції переміщення, яка визначає вектор руху кожної точки між сусідніми кадрами:

$$\mathbf{v}(x, y, t) = (u(x, y, t), v(x, y, t)) \quad (1.36)$$

Основний принцип просторово-часового аналізу базується на умові сталості яскравості:

$$I(x, y, t) = I(x + \delta x, y + \delta y, t + \delta t), \quad (1.37)$$

що означає незмінність інтенсивності точки при її переміщенні. Розклавши праву частину у ряд Тейлора та відкинувши старші члени, отримують рівняння оптичного потоку:

$$I_x u + I_y v + I_t = 0, \quad (1.38)$$

де I_x, I_y, I_t - частинні похідні інтенсивності по координатах x, y та часу t ;

u, v - компоненти швидкості.



Рисунок 1.12 – Приклад виявлення рухомих об'єктів з використанням просторово-часових методів

У класичному підході Horn–Schunck (1981) [46] використовується мінімізація функціонала:

$$E(u, v) = \iint \left[(I_x u + I_y v + I_t)^2 + \alpha^2 (\|\nabla u\|^2 + \|\nabla v\|^2) \right] dx dy, \quad (1.39)$$

де перший доданок забезпечує узгодження з яскравістю;

другий доданок відповідає за згладжування швидкісного поля;

α - коефіцієнт регуляризації.

Локальні методи (наприклад, Lucas–Kanade) оцінюють рух у невеликому вікні через розв'язання системи лінійних рівнянь. Відомою є модель багатокомпонентного оптичного потоку, що враховує наявність кількох незалежних рухомих областей [47]. Вона ґрунтується на мінімізації енергетичного функціонала:

$$E = \sum_{x,y} \rho(I_x u + I_y v + I_t) + \lambda \sum_{x,y} \rho_s(\nabla u, \nabla v), \quad (1.40)$$

де $\rho()$ - робастна функція втрат, що зменшує вплив шуму та оклюзій.

У роботі [28] розглянуто просторово-часове моделювання у контексті стохастико-детермінованої обробки сигналів, де світловий потік накопичується у просторі та часі з урахуванням випадкових флуктуацій. Для випадкової складової сигналу $s(t)$ визначається очікуване значення накопиченої потужності:

$$\mathbb{E}[S^2] = \int_0^T \int_0^T R_s(\tau) d\tau_1 d\tau_2, \quad (1.41)$$

де $R_s(\tau)$ - автокореляційна функція сигналу.

Це дозволяє оцінювати ймовірність виявлення P_d і хибного спрацювання P_f з урахуванням часової акумуляції інформації:

$$P_d = 1 - \Phi\left(\frac{T_p - \mu_1}{\sigma_1}\right), \quad P_f = 1 - \Phi\left(\frac{T_p - \mu_0}{\sigma_0}\right), \quad (1.41)$$

де $\Phi()$ - функція Гауса;

T_p - поріг прийняття рішення;

μ_i, σ_i - математичне сподівання і стандартне відхилення сигналів при наявності (1) і відсутності (0) об'єкта.

Подальший розвиток просторово-часових методів пов'язаний із впровадженням глибоких нейронних мереж, які інтегрують часову інформацію безпосередньо в архітектуру [48]. Архітектурні рішення, такі як D-ConvNet, Two-Stream Networks, I3D, Non-local Networks здатні одночасно аналізувати просторові й часові ознаки у відеопотоці. У таких моделях обчислюється просторово-часовий тензор ознак $F(x, y, t)$, а його згортки здійснюються у тривимірному просторі, що дозволяє мережі кодувати динаміку рухів і взаємозв'язки між кадрами:

$$F'_{i,j,k} = \sum_{p,q,r} W_{p,q,r} \cdot F_{i+p,j+q,k+r}, \quad (1.42)$$

Окремим класом просторово-часових методів є різницеві методи [49], які ґрунтуються на порівнянні поточного кадру $I_t(x, y)$ з попереднім $I_{t-1}(x, y)$ або фоном $I_b(x, y)$:

$$D_t(x, y) = |I_t(x, y) - I_{t-1}(x, y)|. \quad (1.43)$$

Після порогування $D_t(x, y) > T$ формується маска рухомих областей. Одним з можливих покращень цього методу є застосування адаптивного оновлення фону, що підвищує стійкість до змін освітлення і фонових шумів [50]:

$$I_b^{(t)} = (1 - \alpha) I_b^{(t-1)} + \alpha I_t \quad (1.44)$$

Іншим напрямом удосконалення є поєднання методу різниці кадрів із методами матричного розкладання, наприклад, на основі стійкого головного компонентного аналізу (RPCA). Така комбінація забезпечує послаблення впливу фонових шумів і виділення динамічних компонент [51]

Отже, просторово-часові методи формують основу сучасних систем відеоаналізу. До цієї групи входять алгоритми оптичного потоку, трекінгу, просторово-часової сегментації, аналізу траєкторій тощо. Застосовуються для відеопотоків, динамічних сцен, розпізнавання дій.

Висновки до розділу 1

Даний розділ випускної кваліфікаційної роботи присвячений аналізу загальних принципів формування сигналів та зображень в системах технічного зору, визначення їх основних параметрів та просторових характеристик. Також здійснено огляд методів виявлення об'єктів у системах технічного зору, визначено специфіку їх застосування.

Просторові характеристики систем технічного зору визначаються геометрією оптичної схеми, параметрами фотоприймача та умовами дискретизації. Якість сигналів істотно обмежується дією різноманітних шумових факторів – від флуктуацій фону та власного випромінювання оптики до статистичних шумів приймача. Сукупність цих чинників формує стохастичну природу вихідного сигналу, що зумовлює необхідність застосування спеціальних методів аналізу та обробки, які враховують як просторову структуру зображення, так і статистичні властивості шумів, забезпечуючи підвищення достовірності реєстрації та ідентифікації об'єктів у складних умовах спостереження.

Ефективні методи виявлення малорозмірних та малоконтрастних об'єктів, розробляються на основі комплексного підходу, який поєднує хвильовий і корпускулярний опис сигналів, стохастико-детерміновані моделі, а також

сучасні методи просторово-часової, порогової, кореляційної та сегментаційної обробки. Кожен із них має свої особливості та специфіку застосування. На сьогоднішній день універсального методу виявлення не існує, тому вибір методу виявлення для системи технічного зору має бути обґрунтованим згідно поставленої задачі. Найбільш актуальними на сьогодні є гібридні методи, що поєднують стохастико-детерміновані підходи з інтелектуальними методами аналізу даних, зокрема машинним навчанням. Вони забезпечують підвищену стійкість до завад, змін освітлення, флуктуацій шуму та дозволяють ефективно виявляти об'єкти в складних динамічних сценах.

Таким чином, доцільним є розробка та удосконалення методів виявлення та локації об'єктів, насамперед – малорозмірних та малоконтрастних, що використовують сучасні підходи та технології штучного інтелекту, а саме – глибокого навчання штучних нейронних мереж.

1 РОЗРОБКА ТА УДОСКОНАЛЕННЯ СИСТЕМИ ТЕХНІЧНОГО ЗОРУ ДЛЯ ВИЯВЛЕННЯ ТА ПІДРАХУНКУ МАЛОРОЗМІРНИХ ТА МАЛОКОНТРАСТНИХ ОБ'ЄКТІВ

2.1 Імітаційне моделювання відгуку багатоеlementного фотоприймача з урахуванням умов спостереження

Для розробки та удосконалення електронної системи обробки зображень через розробку методу виявлення малорозмірних та малококонтрастних об'єктів виникає необхідність у синтезі штучних зображень, що формуються оптико-електронною системою відповідно до реальних умов спостереження. Для цього було розроблено імітаційну модель відгуку багатоеlementного фотоприймача оптико-електронної системи, що використовує хвильову теорію світла для моделювання поширення та реєстрації оптичного сигналу. Розроблена модель має враховувати: проходження оптичного сигналу через середовище, вплив оптичної ланки, формування відгуку фотоеlementа з урахуванням аналогової та цифрової обробки сигналів.

Першим етапом є обчислення оптичної передавальної функції (OTF) для заданої апертури. Оптична передавальна функція (OTF) – це комплексна функція, яка описує, як оптична система передає просторові частоти від об'єкта до зображення. Вона визначається як нормована автокореляція функції пропускання апертури у частотній області:

$$\text{OTF}(\nu_x, \nu_y) = \frac{\iint P(x, y) P^*(x + \xi, y + \eta) dx dy}{\iint |P(x, y)|^2 dx dy}, \quad (2.1)$$

де $P(x, y)$ – функція пропускання апертури (амплітудна маска);

(ν_x, ν_y) – просторові частоти.

Вираз у знаменнику виконує нормування так, щоб $\text{OTF}(0, 0) = 1$.

Модуль OTF називається модульною передавальною функцією (MTF):

$$\text{MTF}(\nu_x, \nu_y) = |\text{OTF}(\nu_x, \nu_y)|. \quad (2.2)$$

Аргументом OTF є фазова передавальна функція (PTF).

OTF дорівнює нулю для просторових частот, більших за граничну:

$$\nu_c = \frac{D}{\lambda F}, \quad (2.3)$$

де D - діаметр апертури;

λ - довжина хвилі;

F - фокусна відстань.

Для апертур з осьовою симетрією OTF є дійсною і невід'ємною. Отже, форма OTF залежить від геометрії апертури. Для прямокутної апертури розміром $a \times b$ вона розкладається на добуток одновимірних трикутників:

$$\text{OTF}_{\text{прямокут. аперт.}}(f_x, f_y) = \left(1 - \frac{|f_x|}{f_{c,x}}\right)_+ \cdot \left(1 - \frac{|f_y|}{f_{c,y}}\right)_+, \quad (2.4)$$

де $(\cdot)_+$ означає $\max(0, \cdot)$;

$$f_{c,x} = \frac{a}{\lambda}, f_{c,y} = \frac{b}{\lambda};$$

f_x, f_y — кутові/просторові частоти в циклах на радіан.

В площині зображення (фізична відстань), потрібно застосувати фокусну відстань F :

$$f_{c,x} = \frac{a}{\lambda F}, f_{c,y} = \frac{b}{\lambda F}. \quad (2.5)$$

Для круглої апертури діаметром D OTF має радіальну симетрію:

$$\text{OTF}_{\text{кругл. аперт.}}(\rho) = \frac{2}{\pi} \left[\arccos(\rho) - \rho \sqrt{1 - \rho^2} \right], \quad 0 \leq \rho \leq 1, \quad (2.6)$$

де $\rho = \frac{\nu}{\nu_c}$;

$$\nu_c = \frac{D}{\lambda F}.$$

Слід зазначити, що для круглої апертури $\text{OTF}(0)=1$, $\text{OTF}(1)=0$, тобто форма OTF плавно спадає від 1 до 0 (рис. 2.1).

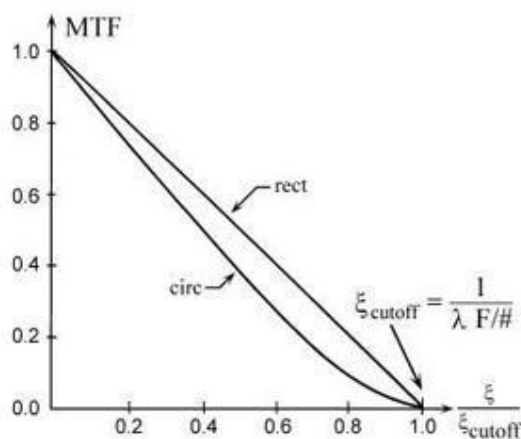


Рисунок 2.1 – Оптична передавальна функція для прямокутної та круглої апертури

Наступним етапом є обчислення OTF атмосфери. Використана модель довго-експозиційної оптичної передачі описує усереднений вплив турбулентності на поширення просторових частот від об'єкта до зображення. Вона безпосередньо пов'язана зі статистикою фазових викривлень, що виникають унаслідок флуктуацій показника заломлення повітря. Для моделі Колмогорова фазова структура функції має степеневий закон із показником $5/3$, що відображає каскад енергії від великих масштабів турбулентності до малих. OTF атмосфери залежить лише від радіальної частоти f і від параметра Фріда r_0 :

$$\text{OTF}_{\text{атм}}(f) = \exp \left[-3.44 \left(\frac{\lambda f}{r_0} \right)^{5/3} \right], \quad f = \sqrt{f_x^2 + f_y^2}. \quad (2.7)$$

де f - радіальна просторово-частотна координата;

λ - довжина хвилі випромінювання,

r_0 - параметр Фріда, що характеризує ефективний діаметр когерентності хвильового фронту.

Ця функція є ізотропною, тобто залежить лише від модуля просторової частоти, і монотонно спадає від одиниці при $f=0$ до нуля при високих частотах. Таким чином, за умови довгої експозиції атмосфера діє як низькочастотний фільтр, що послаблює передачу дрібних деталей зображення.

Залежність r_0 від довжини хвилі описується законом:

$$r_0(\lambda) = r_0(\lambda_0) \left(\frac{\lambda}{\lambda_0} \right)^{6/5}. \quad (2.8)$$

Це означає покращення умов спостереження на хвилях більшої довжини.

Перетворення Фур'є від $\text{OTF}_{\text{атм}}(f)$ дозволяє отримати довгоекспозиційну функцію розсіювання точки (PSF), яка має гаусоподібний профіль із розширеними «крилами». Її ширина на половині висоти визначає кутову роздільну здатність, обмежену атмосферою - $\theta_{\text{атм}}$. Аналітично ця величина виражається як:

$$\theta_{\text{атм}} \approx 0.98 \frac{\lambda}{r_0}. \quad (2.9)$$

При довжині хвилі $\lambda = 0,5\text{м}$ значення $\theta_{\text{атм}} = 1$ кут. с. відповідає $r_0 \approx 10,1$ см. Це співвідношення широко використовується в астрономії для оцінки якості атмосфери та планування спостережень [52].

Далі обчислюється OTF повної системи. Якщо системні впливи є лінійними й просторово-інваріантними (згортки для інтенсивності), то їхні OTF множаться. Тобто дифракція апертури і усереднена дія турбулентності комутують у частотній області як добуток:

$$\text{OTF}_{\text{сист}}(f_x, f_y) = \text{OTF}_{\text{аперт}}(f_x, f_y) \cdot \text{OTF}_{\text{атм}}(f). \quad (2.10)$$

Наступним кроком обчислюється функція розсіювання точки (PSF) системи – імпульсна характеристика системи, просторовий розподіл інтенсивності від точкового джерела. Отже, PSF визначає роздільну здатність системи: чим вужча функція, тим краща здатність системи відтворювати дрібні деталі.

Для когерентного випадку PSF дорівнює квадрату модуля перетворення Фур'є від функції пропускання апертури. Якщо апертура має осьову симетрію (наприклад, круга), то й PSF буде радіально симетричною.

Для інкогерентної системи (у даному випадку) PSF безпосередньо пов'язана з оптичною передавальною функцією (OTF) через обернене перетворення Фур'є:

$$\text{PSF}(x, y) = \mathcal{F}^{-1}\{\text{OTF}_{\text{сист}}(f_x, f_y)\}, \quad (2.11)$$

де $\mathcal{F}^{-1}$ - обернене перетворення Фур'є [53].

PSF завжди нормована так, щоб інтеграл по всій площині дорівнював одиниці:

$$\iint \text{PSF}(x, y) dx dy = 1. \quad (2.12)$$

Після розрахунку функції розсіювання точки оптичною системою можливим стає моделювання процесу формування зображення. У випадку інкогерентного освітлення, зображення визначається як двовимірна згортка початкового розподілу інтенсивності з PSF системи:

$$I_{\text{вихідн}}(x, y) = I_{\text{вихідн}}(x, y) * \text{PSF}(x, y), \quad (2.13)$$

де $I_{\text{вихідн}}$ - початковий розподіл інтенсивності;

* - 2D-згортка.

У неперервному випадку згортка записується у вигляді інтеграла:

$$I_{\text{вихідн}}(x, y) = \iint_{-\infty}^{\infty} I_{\text{вихідн}}(x', y') \text{PSF}(x - x', y - y') dx' dy'. \quad (2.14)$$

Цей вираз означає, що інтенсивність у кожній точці зображення є зваженою сумою внесків від усіх точок об'єкта, де ваговою функцією є PSF [54].

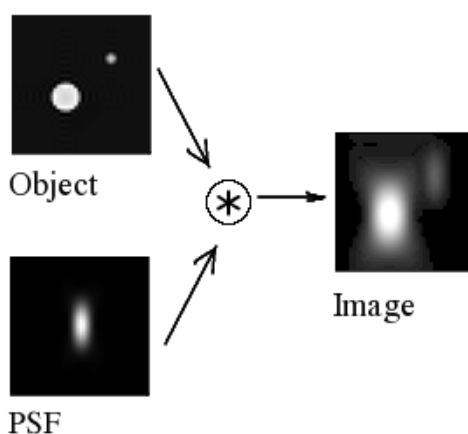


Рисунок 2.2 – Формування зображення з урахуванням функції розсіювання точки

Наступним етапом моделювання після визначення впливу оптичної ланки є формування відгуку оптичного сенсора.

Одним з важливих аспектів є оптичні втрати. У реальних сенсорах зареєстрована інтенсивність завжди менша за ту, що падає на вхідну апертуру. Це пояснюється низкою фізичних факторів [55]:

- відбиття від поверхонь: на межах поділу середовищ (скло–повітря, шари мікролінз, захисні покриття) частина світлового потоку відбивається назад;
- абсорбція у шарах: фотони поглинаються матеріалами сенсора або оптичних елементів, не досягаючи фоточутливого шару;
- квантова ефективність (QE): навіть ті фотони, що досягли активної області, не завжди генерують електрон; реальна ефективність завжди менша за 100%.

Внаслідок цього реєстрована інтенсивність масштабується коефіцієнтом оптичних втрат D :

$$I' = \frac{I}{D_{opt}} \quad (2.15)$$

де D_{opt} - коефіцієнт оптичних втрат.

Значення D_{opt} залежить від конструкції сенсора, якості антиблікових покриттів та довжини хвилі випромінювання.

Світло має квантову природу, тому надходження фотонів на сенсор є випадковим процесом, що описується статистикою Пуассона. Це означає, що навіть при сталому середньому потоці фотонів кількість зареєстрованих квантів у кожному пікселі коливається при сталому інтервалі експозиції, що обумовлює фотонний шум. Кількість зареєстрованих фотонів у пікселі:

$$N_{\text{фот}} \sim \text{Poisson}\left(\frac{I'}{C}\right) \quad (2.16)$$

де C - коефіцієнт перетворення (e-/ADU).

Після перетворення фотонів у електрони та подальшої оцифровки сигналу враховується коефіцієнт перетворення C :

$$I'' = N_{\text{фот}} \cdot C \quad (2.17)$$

Фотоприймач має скінчену ємність пікселя, яка визначає максимальну кількість носіїв заряду (електронів), що можуть бути накопичені під час експозиції. Ця межа відповідає напрузі насичення $V_{\text{нас}}$. При досягненні цієї межі подальше збільшення вхідної інтенсивності світла не призводить до зростання вихідного сигналу. Максимальне значення напруги досягається при:

$$I_{\text{нас}} = \frac{V_{\text{нас}}}{S \cdot t} \quad (2.18)$$

де $V_{\text{нас}}$ - напруга насичення;

S – чутливість фотоприймача;

t - експозиція.

При цьому виконується умова обмеження:

$$V_{\text{max}} \leq V_{\text{нас}}. \quad (2.19)$$

Далі застосовується нелінійна сенсорна функція відгуку, яка відображає інтенсивність у напругу:

$$V = f_{\text{сенси}}(I'', I_{\text{нас}}, V_{\text{нас}}, t) \quad (2.20)$$

До певної межі сенсор працює як лінійний фотодетектор. Біля межі насичення сигнал плавно переходить у плато, що відображає обмежену ємність пікселя (рис. 2.1.3). У робочій зоні (до насичення) сенсор поводить приблизно лінійно:

$$V_{\text{лін}}(I) = \frac{I}{I_{\text{нас}}} \cdot V_{\text{нас}}, \quad (2.21)$$

де I - вхідна світлова інтенсивність;

$I_{\text{нас}}$ - інтенсивність насичення;

$V_{\text{нас}}$ - напруга насичення [56].

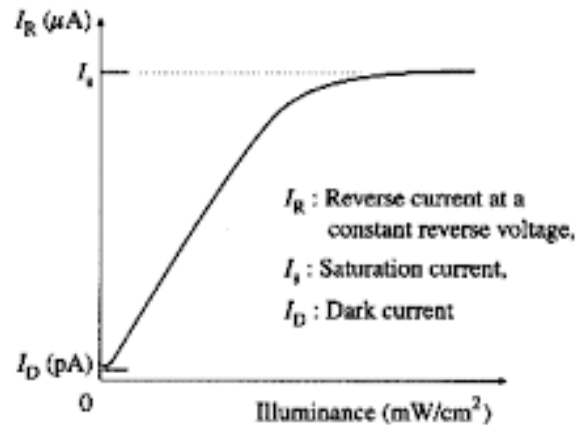


Рисунок 2.3 – Функція відгуку оптичного фотоприймача

У реальних сенсорах перехід у режим насичення відбувається не різко, а поступово. Для цього вводиться згладжуюча сигмоїдальна функція:

$$S(I) = \frac{1}{1 + \exp(-k(I - I_{sat}))}, \quad (2.22)$$

де параметр k визначає крутість переходу (великі значення k - різкий перехід (наближення до кліпінгу), малі значення k - плавний перехід).

Фінальна вихідна напруга обчислюється як змішування лінійної частини та плато насичення:

$$V(I) = V_{lin}(I) \cdot (1 - S(I)) + V_{sat} \cdot S(I) \quad (2.23)$$

Ця формула описує поведінку фотоприймача:

- при малих інтенсивностях $I \ll I_{sat}$: $S(I) \approx 0 \Rightarrow V(I) \approx V_{lin}(I)$ (вихідний сигнал описується лінійною моделлю);
- при великих інтенсивностях $I \gg I_{sat}$: $S(I) \approx 1 \Rightarrow V(I) \approx V_{sat}$ (вихідний сигнал виходить на плато).

Темновий струм виникає у фотодетекторах навіть за відсутності світла. Його джерелом є теплова генерація електрон-діркових пар у напівпровідниковому матеріалі сенсора. Ці носії заряду накопичуються у пікселі так само, як і фотоелектрони, створюючи паразитний сигнал [57]. Кількість темнових електронів, що накопичуються у пікселі за час експозиції t , визначається як:

$$N_{темн} = D(T) \cdot t, \quad (2.24)$$

де $D(T)$ - швидкість генерації темного струму (електрони/піксель/секунда);
 t - час експозиції.

Величина $D(T)$ експоненційно залежить від температури, що відображає активаційний характер процесу (рис. 2.4). Для моделювання використовується рівняння, яке враховує енергію активації та закон Больцмана. При відомій швидкості темного струму $D(T_{on})$ за опорної температури T_{on} , при іншій температурі T , вона масштабується експоненційно:

$$D(T) = D(T_{on}) \cdot \exp\left[-\frac{E_a}{k_B} \left(\frac{1}{T} - \frac{1}{T_{on}}\right)\right], \quad (2.25)$$

де $k_B = 8.617 \times 10^{-5}$ еВ/К - стала Больцмана;

E_a – енергія активації (енергетичний бар'єр для термічної генерації електронів).

Енергія активації $E_a \approx 0.65 \text{ eV}$ відповідає типовим значенням для кремнієвих сенсорів [58].

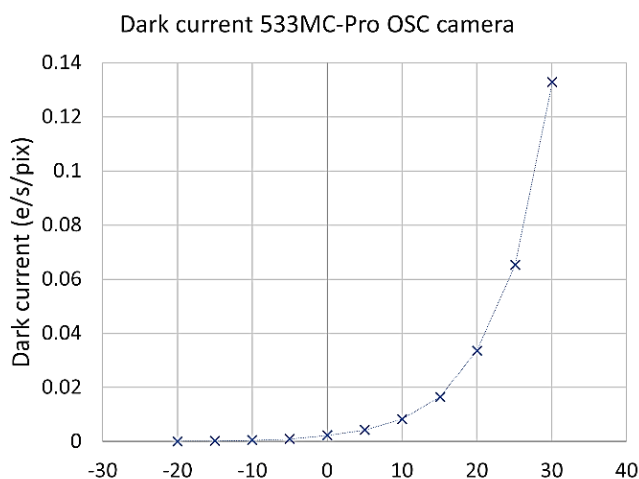


Рисунок 2.4 – Залежність темного струму від температури

Процес генерації електронів є випадковим та підпорядковується статистиці Пуассона. Це означає, що кількість темнових електронів у кожному пікселі коливається навколо середнього значення $N_{\text{темн}}$, а дисперсія цього шуму дорівнює середньому значенню:

$$N_{\text{темн.шум}} \sim \text{Poisson}(N_{\text{темн}}) \quad (2.26)$$

З урахуванням темнового струму вихідний сигнал у пікселі можна записати як:

$$V = V_{\text{сигн}} + N_{\text{темн.шум}} \quad (2.27)$$

Після зчитування заряд із пікселя перетворюється у напругу, яка далі проходить через аналоговий підсилювач. Це необхідно для узгодження рівня сигналу з діапазоном роботи АЦП. Математично процес описується як:

$$V' = G \cdot V, \quad (2.28)$$

де G - коефіцієнт підсилення.

Процес зчитування супроводжується додатковим шумом, який виникає у схемах підсилення та перетворення сигналу. Цей шум зазвичай моделюється як адитивний гаусівський шум із нульовим середнім:

$$N_{\text{зчит}} \sim \mathcal{N}(0, \sigma_{\text{зчит}}^2) \quad (2.29)$$

де $\sigma_{\text{зчит}}$ - стандартне відхилення шуму зчитування.

Вихідний сигнал після додавання шуму:

$$V'' = V' + N_{\text{зчит}} \quad (2.30)$$

Таким чином, навіть при відсутності світла на сенсорі, у зчитаному сигналі завжди присутні флуктуації, зумовлені електронними компонентами [59].

Вихідний аналоговий сигнал обмежується діапазоном напруг сенсора ($0 \dots V_{\text{нас}}$), після чого переводиться у цифрові одиниці (ADU):

$$V_{\text{вих}} = \text{clip}(V'', 0, V_{\text{нас}}), \quad (2.32)$$

де функція $\text{clip}()$ обмежує значення, що виходять за межі допустимого діапазону:

- якщо $V'' < 0$, то $V_{\text{вих}} = 0$;
- якщо $V'' > V_{\text{нас}}$, то $V_{\text{вих}} = V_{\text{нас}}$;
- якщо $0 < V_{\text{нас}}$, то $V_{\text{вих}} = V''$.

Аналоговий сигнал передається на аналогово-цифровий перетворювач (АЦП), який виконує дискретизацію за рівнями – перетворює безперервну напругу у кінцеву кількість цифрових значень.

Кількість доступних рівнів визначається кількістю бітів n , які виділені на представлення одного пікселя:

$$L = 2^n - 1, \quad (2.33)$$

де L – максимальний цифровий рівень (наприклад, для 8-бітного АЦП $L=255$).

Оскільки АЦП може представити лише дискретні рівні, кожне аналогове значення округлюється до найближчого рівня:

$$Q(V) = \text{round}(V_{ADC}) \quad (2.34)$$

Це призводить до появи помилки квантування (рис. 2.1.5), яка є різницею між реальним аналоговим значенням та його цифровим представленням:

$$\varepsilon_{\text{квант}} = V_{\text{вих}} - V_{\text{квант}}, \quad (2.35)$$

де $V_{\text{квант}}$ - округлене значення, яке відповідає одному з $L+1$ рівнів.

Шум квантування є детермінованим, але може бути представлений як рівномірно розподілений у межах одного кванту:

$$\varepsilon_{\text{квант}} \sim \mathcal{U}\left(-\frac{\Delta}{2}, \frac{\Delta}{2}\right), \quad (2.36)$$

де $\Delta = \frac{V_{\text{макс}}}{L}$ - розмір одного кванту.

Дисперсія шуму квантування

$$\sigma_{\text{квант}}^2 = \frac{\Delta^2}{12}. \quad (2.37)$$

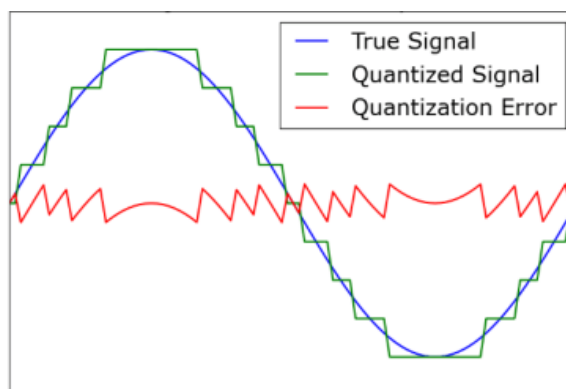


Рисунок 2.5 – Помилка квантування на етапі аналогово-цифрового перетворення

Якщо сигнал виходить за межі допустимого діапазону $[0, L]$, то в цьому випадку відбувається насичення АЦП:

$$Q_{\text{ADC}} = \text{clip}(Q(V), 0, L). \quad (2.38)$$

- якщо $Q(V) < 0$, то $Q_{\text{ADC}} = 0$;
- якщо $Q(V) > L$, то $Q_{\text{ADC}} = L$;
- якщо $0 < Q(V)$, то $Q_{\text{ADC}} = Q(V)$.

Після дискретизації кожен піксель має цілочисельне значення, яке можна зберігати у цифровому представленні. Типовими форматами є:

- 8-bit PNG - $n=8$, $L=255$: компактне представлення для обробки та візуалізації;
- 12-bit RAW - $n=12$, $L=4095$: дозволяє зберігати більший динамічний діапазон яскравості, використовується у професійних фотокамерах;
- 16-bit TIFF - $n=16$, $L=65535$: застосовується у науковій візуалізації та медичних зображеннях.

Таким чином, розроблена імітаційна модель формування зображень в оптоелектронних системах дозволяє моделювати вихідне зображення на основі інформації про далеко розташовані спостережувані об'єкти: інтенсивності та положення (в координатах у площині фотоприймача). Модель враховує такі фактори, як:

- дифракція Фраунгофера на апертурі довільної форми;
- довгоекспозиційна атмосферна турбулентність;
- зовнішні та внутрішні шуми, зокрема темновий струм, фотонний шум та шум зчитування;
- нелінійність світлової характеристики фотоприймача;
- температурна залежність темного струму;
- вплив аналогово-цифрового перетворювача.

Основними недоліками розробленої моделі є:

- можливість моделювання лише монохроматичних світлових сигналів (зовнішні шуми також моделюються як монохроматичні);

- неврахування відстані до спостережуваного об'єкта (об'єкти вважаються точковими);
- можливість моделювання впливу атмосферної турбулентності лише при довгій експозиції (коли час експозиції набагато більше періоду «замороженості» атмосфери);
- неврахування корпускулярних властивостей світла на етапі моделювання дифракційної плями.

Імітаційну модель було програмно реалізовано з використанням середовища Python 3.13 та бібліотек, наявних у відкритому доступі:

- numpy – для виконання векторних обчислень;
- scipy – для обчислення операції згортки;
- matplotlib – для візуалізації;
- Pillow – для збереження результатів у растрових цифрових форматах.

Розроблене програмне забезпечення використовує паралельні обчислення, що дозволяє ефективно виконувати моделювання великої кількості сигналів та зображень. Це дає можливість створення великих наборів синтетичних даних довільного розміру з довільними параметрами модельованої оптико-електронної системи та спостережуваних об'єктів. Такі набори даних можуть бути використані для подальшого аналізу та обробки, зокрема в задачах розробки та випробування методів реєстрації об'єктів.

2.2 Розробка комбінованого методу підрахунку та локації фіксованого числа об'єктів на основі глибокого навчання

Для реєстрації близько розташованих малорозмірних та малоконтрастних об'єктів було розроблено комбінований метод [60], що дозволяє визначити кількість об'єктів на зображенні, а також математичне сподівання координат центру для кожного з об'єктів (рис. 2.6). Він є комбінацією глибоких згорткових нейронних мереж, кожна з яких виконує специфічну задачу. За основу взято

архітектуру згорткової нейронної мережі MobileNet V1 (рис 2.7), що має суттєві переваги для застосування у вбудованих системах завдяки своїй обчислювальній та енергетичній ефективності. Ключовим елементом є використання роздільних згорток (depthwise separable convolutions), що дозволяє значно зменшити кількість параметрів і операцій порівняно з традиційними згортковими шарами, зберігаючи при цьому прийнятний рівень точності. Такий підхід забезпечує компактність моделі, низьке споживання пам'яті та можливість роботи в режимі реального часу навіть на пристроях з обмеженими ресурсами, таких як мобільні телефони, мікроконтролери чи інші обчислювальні платформи з обмеженим енергоспоживанням. Крім того, наявність коефіцієнтів масштабування глибини та ширини мережі дозволяє гнучко адаптувати архітектуру під конкретні апаратні обмеження, досягаючи оптимального балансу між швидкістю та точністю [61]. В даному випадку вони дорівнюють 1,0 та 0,25 відповідно.

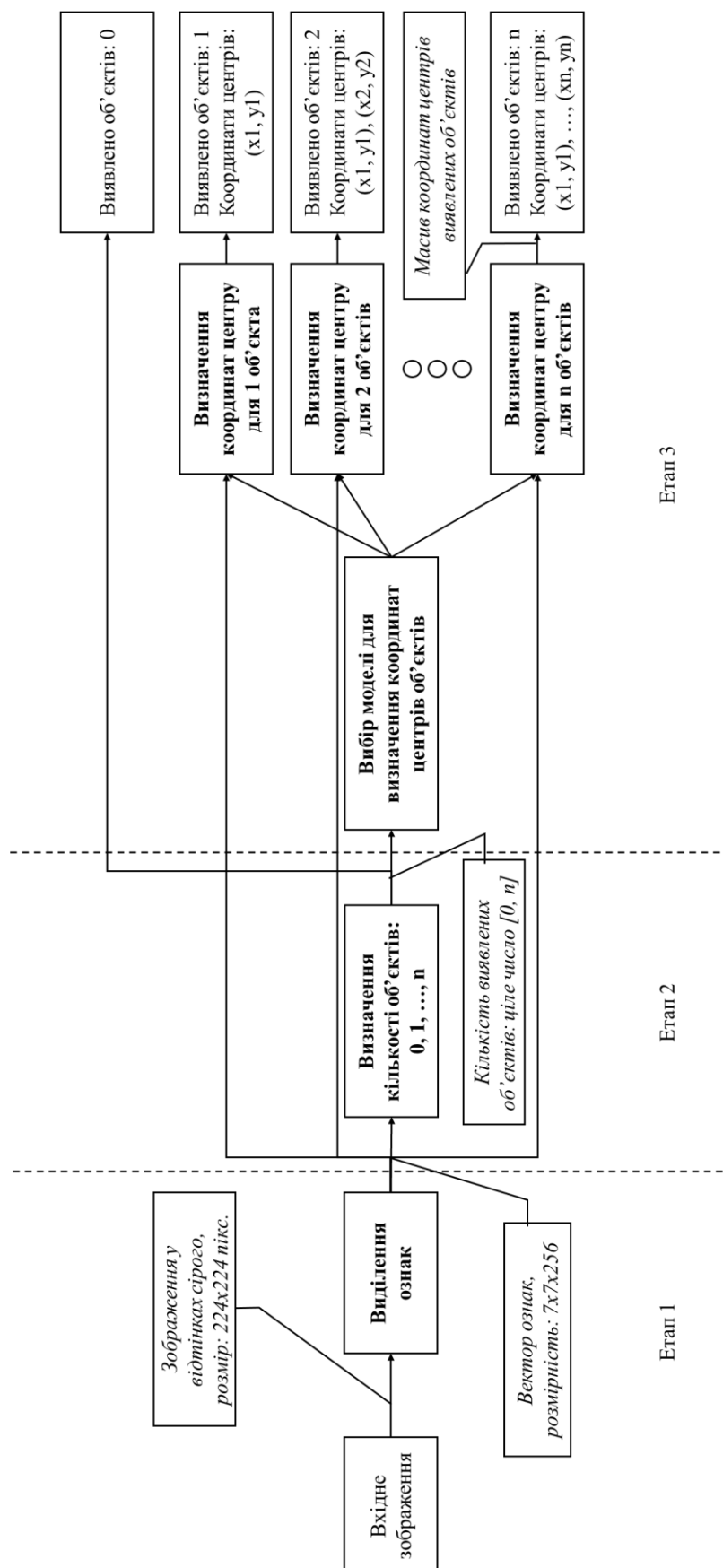


Рисунок 2.6 – Структурна схема розробленого методу реєстрації об'єктів

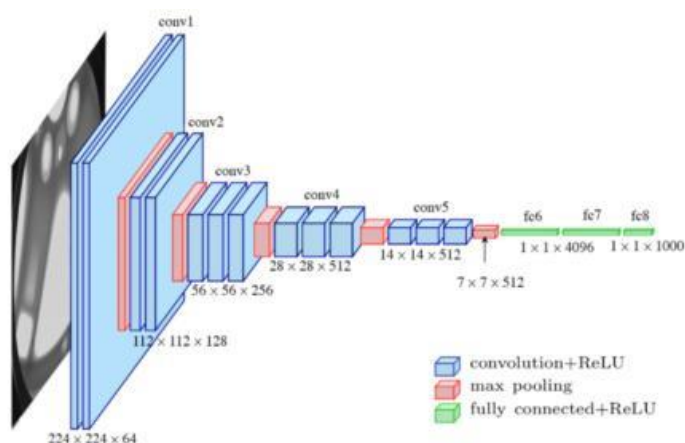


Рисунок 2.7 – Архітектура ШНМ MobileNet V1

Вхідними даними для запропонованого методу є необроблене цифрове зображення, отримане з фотоприймача оптико-електронної системи розмірністю 224×224 пікселі. Зображення має бути одноканальне (у відтінках сірого кольору) з глибиною дискретизації 8 біт на піксель. Якщо роздільна здатність багатoelementного фотоприймача є більшою, ніж необхідна, тоді вхідне зображення може бути отримане шляхом обрізки початкового. Зниження розміру інтерполяцією в даному випадку недопустиме, так як може призвести до втрати деталізації та спотворення дрібних деталей.

Першим етапом є виділення ознак, що має на меті зниження розмірності інформаційного поля завдяки виділенню інформативної складової вхідних даних та перетворення їх у представлення, що є більш ефективним для подальшої обробки. Враховуючи специфіку предметної області, екстрактор ознак має враховувати масштабно-незалежні інтенсивнісні та просторові характеристики зображення, а також мати високу стійкість до шумів. В даному випадку виділення ознак здійснюється шляхом вилучення даних з одного з прихованих шарів нейронної мережі, попередньо навченої для задачі класифікації зображень з набору ImageNet [62]. В якості архітектури ШНМ для виділення ознак обрано MobileNet V1, першу згортку якої було адаптовано для використання одноканального вхідного зображення шляхом підсумування вагових відповідних коефіцієнтів кожного з трьох каналів.

У такий спосіб формується вектор ознак сталої розмірності, який відображає узагальнені візуальні характеристики вхідного зображення, незалежні від класів початкового набору даних. Завдяки попередньому навчанню на великому та різноманітному наборі візуальних даних, мережа набуває здатності виділяти універсальні патерни: від низькорівневих контурів і текстур до більш абстрактних структурних ознак. Вибір шару, з якого вилучаються ознаки, безпосередньо визначає рівень їхньої абстракції та придатність для різних завдань: ранні шари забезпечують ширшу узагальненість і чутливість до локальних структур, тоді як пізні – вищу семантичну інформативність і дискримінативність. Для поточної задачі обрано останній шар 9-го згорткового блока (з 14) через універсальність отриманих даних як для задачі підрахунку об'єктів, так і для їх локації. Це робить отримане представлення придатним для задач подальшої класифікації, кластеризації, регресії тощо. Вхідне необроблене зображення у цифровому форматі займає 50176 байтів, тоді як отриманий вектор ознак – 12544 байтів, тобто в 4 рази менше. Виділені ознаки з тестового набору (40000 зразків) було візуалізовано, використовуючи метод t-SNE (рис. 2.8). Аналіз діаграми демонструє кореляцію між кількістю сигналів на зображенні та просторовим розташуванням точки, що відповідає даному зразку, особливо для випадку 0 та 1 сигналу. Це свідчить про коректність ознак та адекватність обраного методу їх виділення.

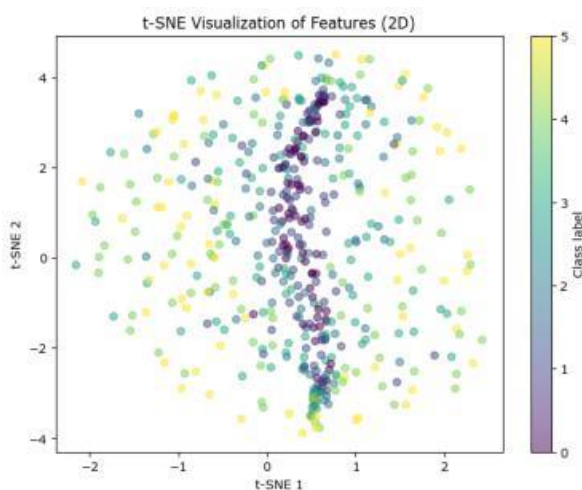


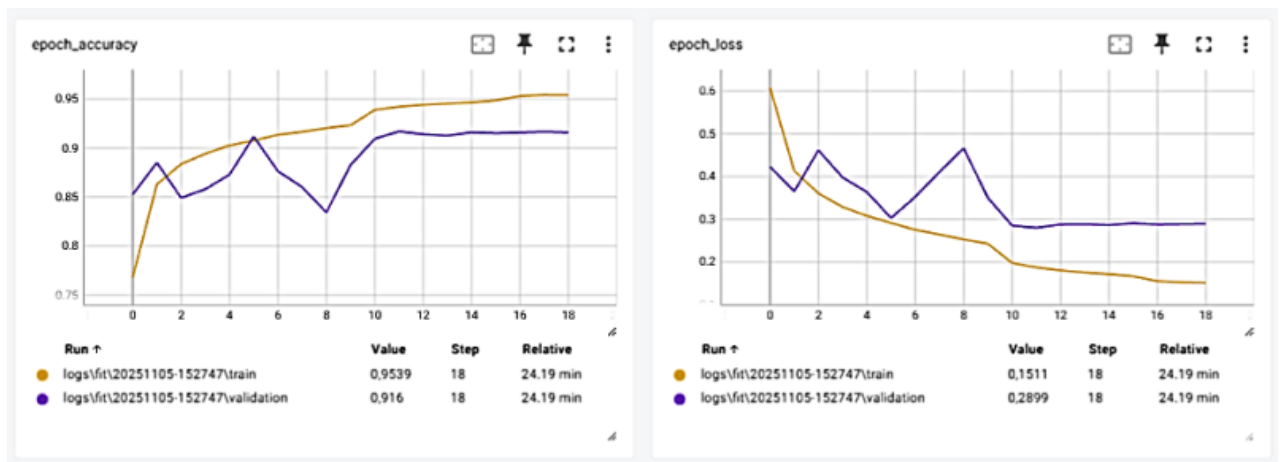
Рисунок 2.8 – Візуалізація виділених ознак за методом t-SNE

Другим етапом методу реєстрації об'єктів є визначення їх кількості на зображенні. Дана задача може бути формалізована як задача класифікації (з поміж класів: 0 – об'єктів не виявлено, 1 – виявлено 1 об'єкт і т.д.) або порядкової регресії (для передбачення цілого числа у фіксованому діапазоні). Класифікаційне рішення може бути прийнято на основі наявності ділянок з меншою інтенсивністю між зображеннями об'єкта (згідно критерію Релея) або за непрямыми ознаками, такими як форма дифракційної картини. Отже, класифікатор, що використовується в даному випадку, має враховувати обидва способи, тому він повинен бути стійким до широкої варіативності вхідних даних та мати високу здатність до узагальнення. Тому, враховуючи специфіку задачі, було обрано згорткову нейронну мережу, отриману шляхом використання шарів, що залишилися незадіяними на етапі виділення ознак.

Отже, на даному етапі, вектор ознак, взятий з певного прихованого шару ШНМ подається на наступний прихований шар, після чого розповсюджується через глибші згорткові шари до класифікаційної «голови», модифікованої під необхідну розмірність вихідного шару. Розмірність повнозв'язного шару також було знижено до 128 з огляду на малу розмірність вихідного шару. Для прискорення сходимості було використано метод перенесення навчання: перед навчанням ШНМ згорткові шари було проініціалізовано ваговими коефіцієнтами, взятими з відповідних шарів моделі MobileNet V1, попередньо навченої на наборі ImageNet.

Навчання здійснювалось з використанням набору синтетичних даних, згенерованих шляхом імітаційного моделювання процесу формування зображень в оптико-електронній системі, із можливою кількістю об'єктів від 0 до 5 (300000 зразків). Для навчання використано 80% зразків у наборі даних, решта (20%) – для валідації та тестування. Після тренування впродовж 19 епох модель показала високу загальну якість класифікації з точністю 0,917. Графіки функції втрат та точності наведені на рис. 2.9. Макро- та зважені середні показники precision, recall і F1 близькі до 0,92, що вказує на однорідну продуктивність між класами загалом і на відсутність сильного зсуву в бік однієї

з категорій. Матриця помилок та значення метрик precision, recall, fl-score та support наведені на рис. 2.10 та в табл. 2.1 відповідно.



а)

б)

а) точності, б) функції втрат

Рисунок 2.9 – Графіки метрик навчання ШНМ-класифікатора

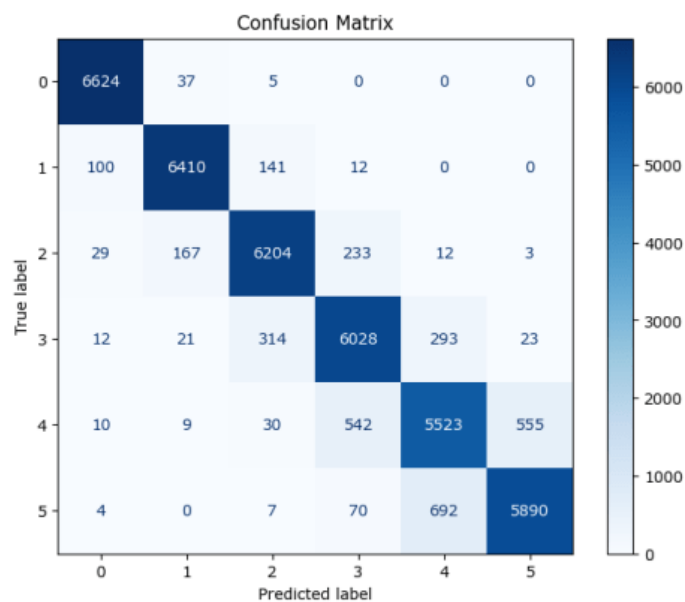


Рисунок 2.10 – Матриця помилок касифікації

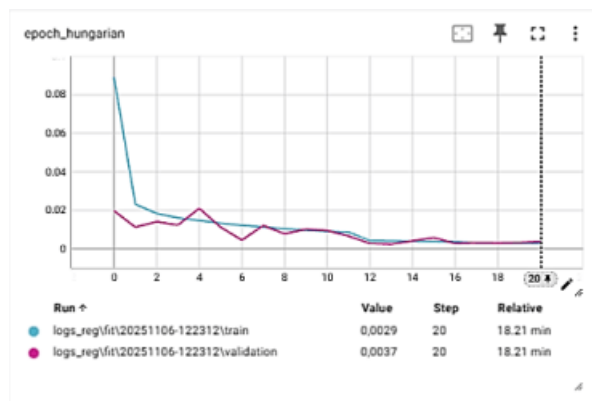
Таблиця 2.1 – Метрики тестування ШНМ-класифікатора

Клас	Метрика			
	precision	recall	f1-score	support
0	0,98	0,99	0,99	6666
1	0,96	0,96	0,96	6663
2	0,93	0,93	0,93	6648
3	0,88	0,90	0,89	6691
4	0,85	0,83	0,84	6669
5	0,91	0,88	0,90	6663

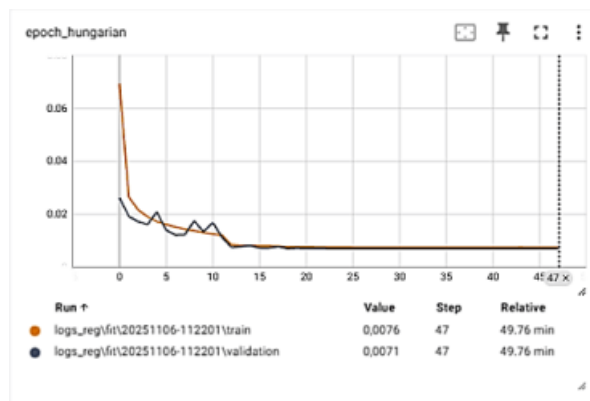
Третім етапом комбінованого методу реєстрації об'єктів є визначення математичного сподівання координат їх центрів. Дана задача є задачею регресії неупорядкованої множини та потребує пермутаційної інваріантності: перестановка елементів множини не повинна змінювати оцінки центрів. Так як розмірність вихідних даних залежить від кількості виявлених об'єктів, було прийнято рішення щодо використання окремої моделі-регресора, що виконує локацію лише заданої фіксованої кількості об'єктів. Наприклад, для визначення центру одного виявленого об'єкта необхідна модель з розмірністю вихідного шару 2, для n об'єктів розмірність становитиме $2n$. Відповідну модель для локації об'єкта буде обрано на основі класифікаційного рішення, отриманого на другому етапі, вхідними даними для моделі виступатиме той самий вектор ознак, отриманий на першому етапі. В якості архітектури моделі для здійснення регресії координат центрів було взято частину від MobileNet V1, що залишається незадіяною при виділенні ознак, аналогічно до архітектури моделі-класифікатора. Однак відмінністю буде адаптована «голова» моделі: вона є глибшою, ніж у класифікаційної моделі, та має два повнозв'язні шари. З огляду на різну складність завдань локації різної кількості об'єктів k , було обрано змінну ширину повнозв'язних шарів: $64k$ для першого та $32k$ для другого. Тренування кожної з п'яти регресійних моделей виконувалось з використанням вибірки

розміром 60000 зразків з того самого синтетичного набору даних. Аналогічно, з кожної вибірки для навчання використано 80% її обсягу. В якості пермутаційно-інваріантної функції втрат використано модифіковану середньоквадратичну функцію в поєднанні з методом співставлення Hungarian matching [63]. Графіки метрик в процесі навчання регресійних моделей (функція втрат) наведені на рис. 2.11. Діаграми парітетів за результатами тестування навчених моделей наведені на рис. 2.12. Значення метрик середньоквадратичної помилки (Mean Squared Error), середньої абсолютної помилки (Mean Absolute Error) та коефіцієнт детермінації (R^2 Score) наведені в табл. 2.2.

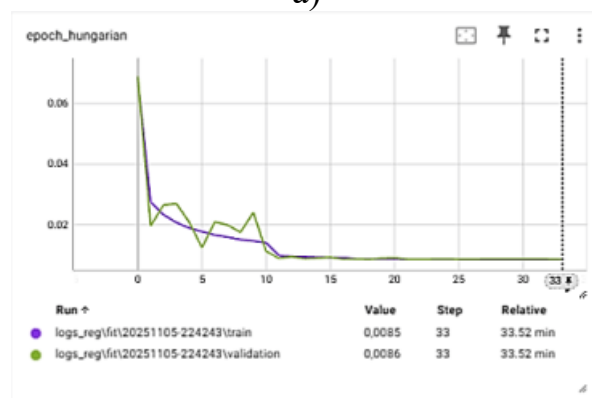
Отримані результати свідчать, що регресійна модель демонструє високу узгодженість та точність для одиничних сигналів, проте зі збільшенням кількості сигналів спостерігається монотонне погіршення за всіма метриками: зростання MSE/MAE та зниження R^2 . Така динаміка вказує на зростання складності задачі виявлення декількох сигналів, що зумовлена високою варіативністю взаємного розташування та потенційно обмеженням поточної архітектури. Виявлена закономірність має системний характер: найбільший деградаційний стрибок спостерігається при переході від одного до двох сигналів, через що стає можливим варіант їх близького розташування. Загалом, модель є адекватною для локації невеликої кількості об'єктів, але її масштабованість до високої щільності сигналів вимагає архітектурного та навчального удосконалення.



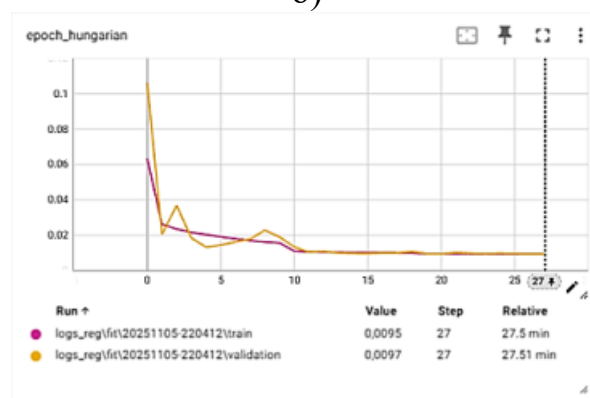
а)



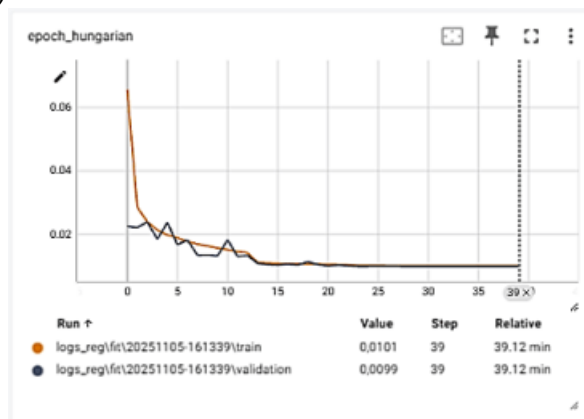
б)



в)



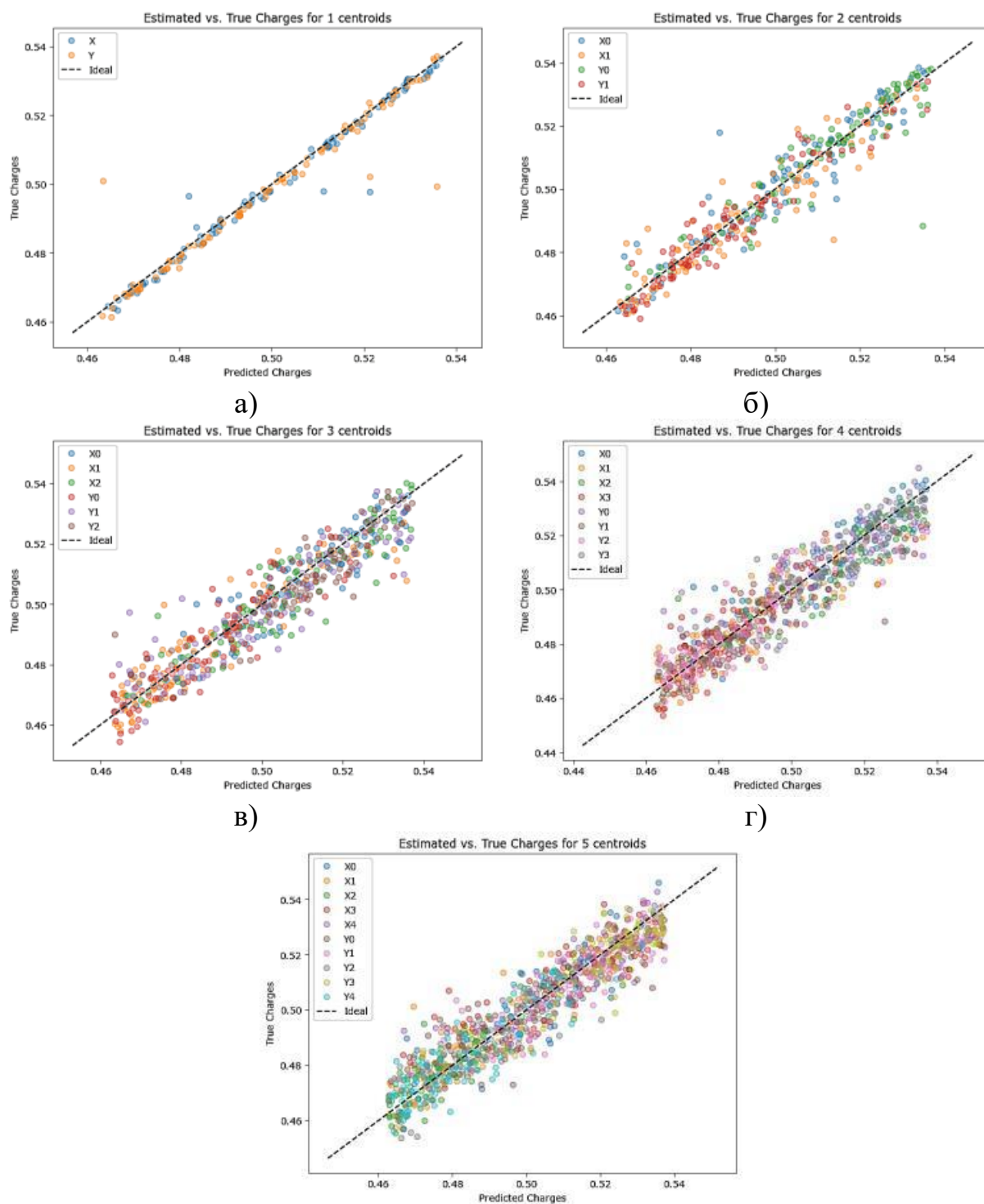
г)



д)

а) для 1 об'єкта, б) для 2 об'єктів, в) для 3 об'єктів, г) для 4 об'єктів, д)
для 5 об'єктів

Рисунок 2.11 – Метрики навчання регресійних моделей



а) для 1 об'єкта, б) для 2 об'єктів, в) для 3 об'єктів, г) для 4 об'єктів, д)
для 5 об'єктів

Рисунок 2.12 – Діаграми паритету результатів тестування регресійних моделей

Таблиця 2.2 – Метрики тестування ШНМ-регресорів

Кількість сигналів	Метрика		
	Mean Squared Error (MSE)	Mean Absolute Error (MAE)	R ² Score
1	0,000012	0,001565	0,9742
2	0,000039	0,004532	0,9147
3	0,000053	0,005465	0,8863
4	0,000065	0,006117	0,8602
5	0,000068	0,006281	0,8526

У даній роботі для розробки моделей глибокого навчання було використано фреймворки Keras та Tensorflow. Перший забезпечує простоту реалізації та високу швидкість прототипування через його високорівневий API. Його безшовна інтеграція з TensorFlow забезпечує доступ до розширених обчислювальних можливостей та підтримку масштабованих архітектур. TensorFlow має спеціалізований модуль TensorFlow Lite (TFLite), що дозволяє конвертувати нейронні мережі у формат, оптимізований для мобільних та вбудованих пристроїв. Завдяки цьому модулю можна зменшити розмір моделі та прискорити її виконання у середовищах з обмеженими апаратними ресурсами. Також даний фреймворк має власний засіб моніторингу метрик Tensorboard, що дозволяє проводити глибокий аналіз роботи моделі [64].

Результатом реєстрації об'єктів на зображенні (вихідні дані) є ціле число, що означає кількість виявлених об'єктів (від 0 до 5), а також множина векторів, що позначають математичне сподівання координат центру кожного з виявлених об'єктів (або порожня множина у випадку їх відсутності). Розроблений метод є адаптивним, так як може бути адаптованим під конкретні параметри оптико-електронної системи та спостережуваних об'єктів шляхом навчання (або тонкого налаштування, *fine-tuning*) нейронних мереж. Сильними сторонами розробленого методу реєстрації об'єктів є:

- можливість розрізнення близько розташованих об'єктів завдяки врахуванню форми дифракційної плями;
- ефективна робота у складних умовах: висока стійкість до шумів (внутрішніх та зовнішніх), незалежність від варіацій інтенсивностей сигнальної та фонові складових, врахування впливу атмосферної турбулентності;
- отримання відносно точних результатів в умовах обмежених обчислювальних ресурсів завдяки виконанню в три етапи (що дозволяє здійснювати обробку послідовно з метою зменшення обсягу оперативної пам'яті);
- модульність (на кожному етапі завантажуються лише необхідні моделі, що дозволяє зменшувати споживання ОЗП);
- використання специфічної архітектури MobileNet, що використовує спрощені згортки (які апроксимують стандартні згорткові операції за допомогою розділення на depthwise та pointwise компоненти).

Серед недоліків методу слід зазначити:

- низька роздільна здатність зображення, що аналізується (фіксована: 224x224 пікселі);
- можливість виявлення та локації фіксованої кількості об'єктів (від 0 до 5);
- обмежена ефективність в реальних умовах через використання синтетичних даних для навчання ШНМ.

Таким чином, розроблений метод забезпечує баланс між продуктивністю та ефективністю, дозволяючи реалізувати алгоритм реєстрації об'єктів у режимі реального часу навіть на апаратних платформах із невисокою обчислювальною потужністю, таких як мікроконтролери.

2.3 Апаратно-програмна реалізація системи технічного зору для вбудованої системи

Фінальним етапом розробки моделі ШНМ є її оптимізація для використання в режимі виведення, що є особливо актуальним для використання на таких платформах як мобільні пристрої чи вбудовані системи. Було здійснено посттренувальне квантування моделі (PTQ) за допомогою вбудованого засобу у фреймворку Tensorflow, що передбачає зменшення точності збереження числових параметрів мережі з 32-бітного представлення з плаваючою комою до формату 8-бітних цілих чисел. Основними перевагами такого підходу є значне зменшення розміру моделі та прискорення виконання операцій у режимі виведення. Перевага PTQ полягає у простоті застосування та відсутності необхідності повторного навчання моделі. На відміну від квантування під час навчання (QAT), яке дає вищу точність, PTQ є менш затратним і краще підходить для задач, де допустиме незначне зниження якості заради суттєвого виграшу в ресурсах [65]. В результаті розмір файлів, в яких зберігається конфігурація та параметри моделей, скоротився в понад 3,5 разів, обсяг ОЗП, що використовується під час виведення також має скоротитись у майже 4 рази.

Наступним етапом є вибір апаратних засобів для реалізації та розгортання розробленого методу реєстрації об'єктів. У сучасних вбудованих системах технічного зору застосування мікроконтролерів із різним рівнем продуктивності визначається балансом між обчислювальними можливостями, енергоспоживанням та наявністю периферійних інтерфейсів. У таблиці 2.3 наведено деякі з популярних мікроконтролерів з даної категорії швидкодії.

Представники серії STM32L4 на основі ядра Cortex-M4 орієнтовані, в першу чергу, на енергоефективність. Вони підтримують базові DSP-інструкції та мають обмежені обсяги оперативної пам'яті (64–320 KB), що робить їх придатними для реалізації простих алгоритмів обробки сигналів та невеликих квантованих моделей нейронних мереж.

Мікроконтролери на ядрі Cortex-M0+ (наприклад, ATSAM21) характеризуються низькою тактовою частотою (до 48 MHz) та порівняно малим обсягом пам'яті. Відсутність SIMD-інструкцій обмежує їх застосування у складних задачах, зокрема комп'ютерного зору, однак можуть використовуватися для отримання та базової попередньої обробки даних та керування периферією.

Мікроконтролери nRF52840 на основі ядра Cortex-M4 із тактовою частотою 64 MHz поєднують базові DSP-інструкції з інтегрованим Bluetooth Low Energy 5.0. Це робить платформу придатною для енергоефективних застосувань, де результати обробки даних передаються бездротовим каналом.

Мікроконтролери Raspberry Pi Pico (RP2040: має два ядра Cortex-M0+ (133 MHz) із 264 KB SRAM та RP2350: два ядра Cortex-M33 з 520 KB SRAM) вирізняються гнучкою системою програмованих інтерфейсів (PIO). Хоча відсутність SIMD-інструкцій обмежує швидкість обробки, даний мікроконтролер може використовуватися для прототипування та реалізації мінімальних моделей машинного навчання з невисокими вимогами щодо обсягу ОЗП.

Серія STM32F4, що має ядро Cortex-M4 із тактовою частотою до 180 MHz та підтримкою DSP-інструкцій забезпечує вищу продуктивність порівняно з L-серією. Завдяки більшому обсягу пам'яті (128–256 KB) можливе виконання виведення за допомогою оптимізованих нейронних мереж у реальному часі. Однак відсутність спеціалізованих апаратних засобів прискорення векторних обчислень залишається основним обмеженням.

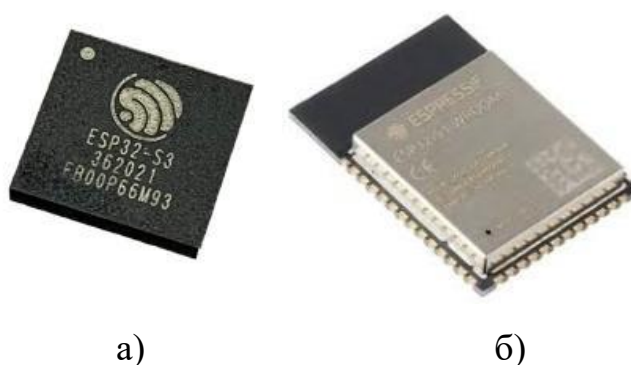
Мікроконтролери Ambiq Apollo3 характеризуються архітектурою Cortex-M4e, яка оптимізована для наднизького енергоспоживання. При тактовій частоті до 96 MHz та обсязі пам'яті близько 384 KB ці мікроконтролери орієнтовані на автономні IoT-системи, де ресурсоемні обчислення виконуються періодично з мінімальними витратами енергії.

Таблиця 2.3 – Порівняльна характеристика мікроконтролерів для застосування в системі технічного зору

Мікрокон- тролер	Ядро такт	Векторні інструкції	RAM (вбудов- ана)	Інтерфейс камери	Інтерфейси передачі даних	Орієнтовне енергоспоживання
STM32L4	Cortex M4, до 80–120 MHz	DSP інструкції (M4) без SIMD; FPU (в деяких)	64–320 KB	Немає (залежить від підсерії)	SPI/I2C/UART/ USB; деякі підтримують Ethernet	Активний режим при 80–100 MHz: 10–25 mA; Stop: 1–2 µA; Standby: 0.3–0.6 µA
ATSAMD 21	Cortex M0+, до 48 MHz	Немає SIMD/DSP	32–64 KB	Немає (можлива програмна реалізація за допомогою PIO)	SPI/I2C/UART/ USB	Активний режим при 48 MHz: 6–12 mA; Standby: 0.5–1.5 µA; Idle: 2–10 µ
nRF52840	Cortex M4, 64 MHz	DSP інструкції (M4) без SIMD; FPU відсутній	256 KB	Немає	USB, SPI/I2C/UART, BLE 5	CPU активний, 64 MHz: 5–7 mA; Radio RX: 4.6 mA; TX (0 dBm): 4.8 mA; System ON idle: 2–3 µA; System OFF: 0.3–0.5 µA
RP2040	2× Cortex M0+, 133 MHz	Немає SIMD/DSP	264 KB SRAM	Немає (можлива програмна реалізація за допомогою PIO)	USB, SPI/I2C/UART, PIO	Активний режим при 168–180 MHz: ~70–100 mA; Stop: 10–50 µA; Standby: 1–5 µA
STM32F4	Cortex M4, до 180 MHz	DSP інструкції (M4); початкові FPU	128– 256 KB	Немає в базових моделях	USB, SPI/I2C/UART, Ethernet (на деяких підсеріях)	Активний режим при 168–180 MHz: 70–100 mA; Stop: 10–50 µA; Standby: 1–5 µA
Ambiq Apollo3	M4e, до 48– 96 MHz	Оптимізований M4e для енергоефективно- сті	384 KB (варію- ється)	Немає	SPI/I2C/UART, Wi-Fi (на деяких підсеріях)	Активний режим 48 MHz: 1.0–1.5 mA; 96 MHz: 2.0–3.0 mA; Deep sleep: 1–5 µA
ESP32-S3	Xtensa LX7, до 240 MHz	SIMD / підтримка векторних інструкцій для MAC (із апаратною підтримкою для прискорення NN)	512 KB SRAM + зовніш- ня PSRAM	DVP (з підтрим- кою PCLK/HSY- NC/VSYN- C)	SPI/I2C/UART/ USB OTG, Wi-Fi 2.4 GHz, BLE	CPU активний: 80– 120 mA; Wi-Fi TX пік: 200– 240 mA; Modem-sleep: 3–10 mA; Deep sleep: 5–20 µA

Мікроконтролер ESP32-S3 має ядро Xtensa LX7 (до 240 MHz) доповнене SIMD-інструкціями та апаратною підтримкою MAC-операцій. Наявність інтерфейсу DVP для камер, зовнішньої PSRAM та інтегрованих модулів Wi-Fi/BLE робить платформу придатною для реалізації легких глибоких моделей ШНМ із можливістю бездротової передачі результатів.

Враховуючи розглянуті варіанти, у рамках даного дослідження в якості апаратної платформи було обрано мікроконтролер ESP32-S3 від Espressif Systems Co., Ltd., що є високопродуктивним і водночас економічно доступним рішенням для реалізації систем з елементами штучного інтелекту на вбудованих пристроях. Даний мікроконтролер представляє собою систему на кристалі (SoC), що поєднує високу обчислювальну потужність з підтримкою сучасних периферійних пристроїв, інтерфейсів та протоколів [66]. Важливою перевагою ESP32-S3 є наявність ефективної вбудованої SRAM до 8 МБ, що значно розширює можливості розгортання згорткових нейронних мереж порівняно з попередніми поколіннями цієї платформи. Ядро також підтримує інструкції векторної обробки, прискорюючи виконання матричних обчислень, характерних для виведення на ШНМ. Детальні характеристики мікроконтролера наведені в табл. 2.4, зовнішній вигляд – на рис. 2.13.



а) мікроконтролер у корпусі QFN56, б) модуль з мікроконтролером та обв'язкою

Рисунок 2.13 – Зовнішній вигляд SoC ESP32-S3

Таблиця 2.4 – Основні характеристики апаратної платформи ESP32-S3

Характеристика	Значення
Ядро ЦП	2× Xtensa® LX7, до 240 МГц
SRAM	До 8 МБ (вбудована + PSRAM)
Flash	До 16 МБ (зовнішня)
Wi-Fi	2.4 GHz, IEEE 802.11 b/g/n
Bluetooth	5.0 LE + Mesh
Периферія	GPIO, SPI, I2C, I2S, UART, ADC, DAC
SIMD інструкції	Підтримуються (для ML-операцій)
USB	USB OTG (Device/Host)
Підтримка ML	TensorFlow Lite Micro, emlearn тощо
MicroPython	Підтримується
Орієнтовна ціна	3–6 USD

У рамках даного дослідження було обрано MicroPython в якості програмної платформи для розгортання моделі глибокого навчання на мікроконтролерах, з огляду на її компактність, кросплатформеність та зручність інтеграції з бібліотеками для інференсу, зокрема emlearn. MicroPython – це легковагова реалізація мови програмування Python, спеціально розроблена для роботи на мікроконтролерах. Вона зберігає зручний синтаксис Python та забезпечує інтерпретоване середовище з підтримкою REPL, що значно прискорює цикл розгортання, інтеграції та тестування моделей без потреби у перекомпіляції. Завдяки підтримці таких платформ, як ESP32, STM32 та RP2040, можливе широке застосування в енергоефективних вбудованих системах. Це робить MicroPython придатним як для швидкого прототипування, так і для прикладних досліджень у сфері Embedded AI.

Для розгортання згорткової ШНМ в умовах вбудованої системи було обрано фреймворк emlearn [67]. Перш за все, він реалізований мовою C та дозволяє інтегрувати бібліотеку як у системи з RTOS (реального часу), так і у

«bare-metal» середовища, що характерно для більшості мікроконтролерів. Однією з ключових переваг emlearn є простота експорту моделі з Keras. Фреймворк emlearn надає інструменти для конвертації моделей у формат, придатний для вбудованих застосунків, що дозволяє уникнути потреби у повторній реалізації логіки мережі вручну. Крім того, підтримка квантування дозволяє суттєво зменшити обсяг пам'яті та прискорити обчислення без значної втрати точності, що критично для реалізації CNN на мікроконтролерах.

З погляду досліджень та швидкого прототипування, важливою перевагою є наявність біндінгів MicroPython. Це дозволяє розробникам використовувати emlearn в інтерактивному середовищі, проводити тестування й налаштування моделі на пристрої з мінімальними витратами часу на компіляцію та прошивку. Такий підхід є особливо корисним у сфері прикладних досліджень і навчання, де важлива оперативність і наочність. Так як основна логіка emlearn реалізована мовою C та скомпільована у машинний код, наявність інтерпретатора MicroPython не має суттєво вплинути на швидкість виведення.

Для проєктування системи використано відлагоджувальну плату, що дозволяє підключення оптичних сенсорів через інтерфейс DVP (рис. 2.14, 2.15).

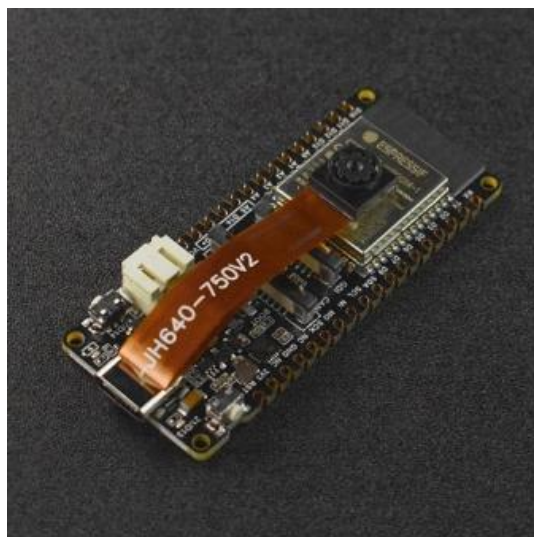


Рисунок 2.14 – Відлагоджувальна плата на основі ESP32-S3 з можливістю підключення оптичних сенсорів через інтерфейс DVP



Рисунок 2.15 – Блок-схема відлагоджувальної плати

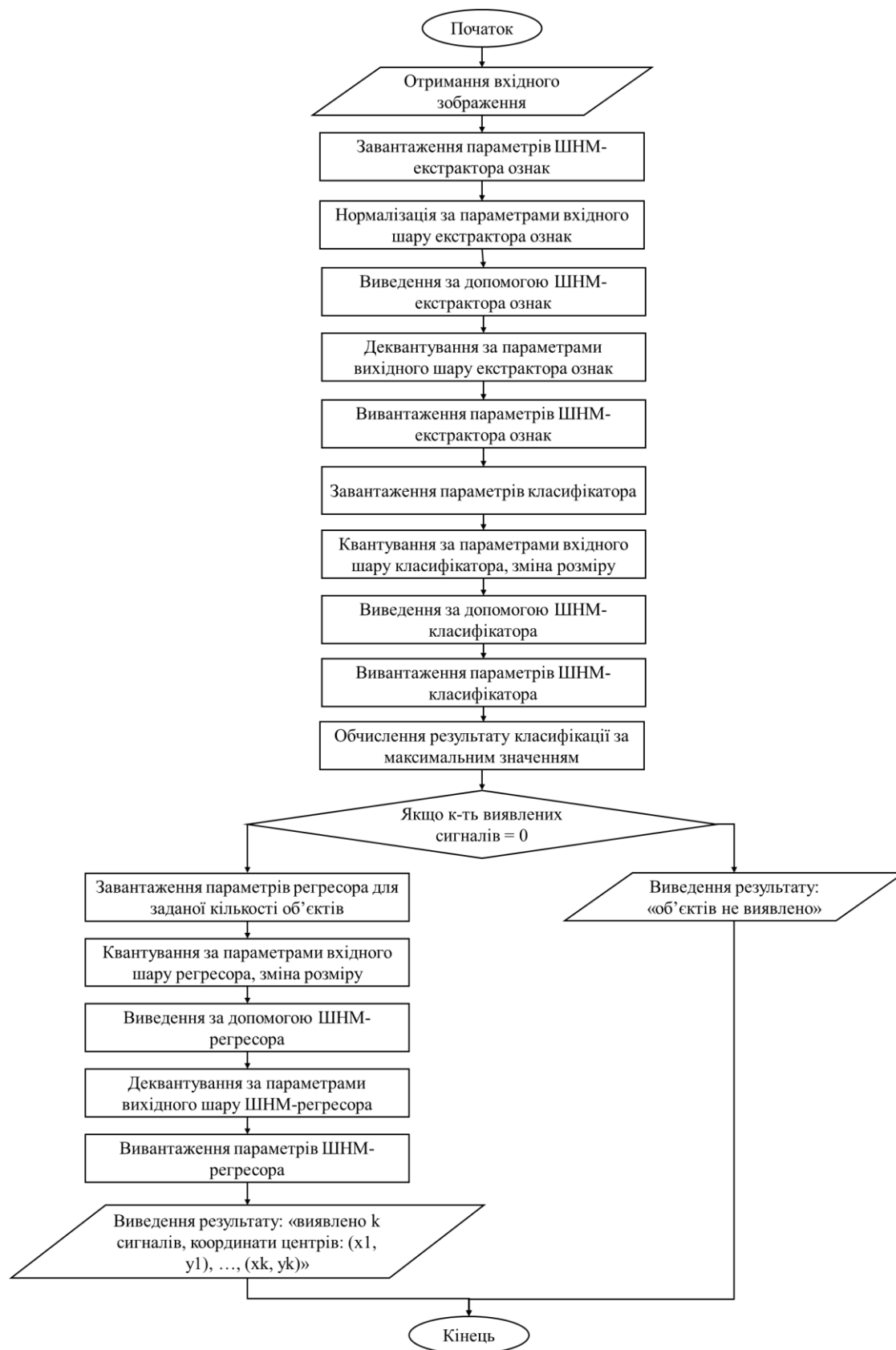


Рисунок 2.16 – Блок-схема алгоритму реєстрації об'єктів

На основі розробленого методу виявлення та локації об'єктів було створено алгоритм, блок-схема якого наведена на рис. 2.16, Ключовими особливостями алгоритму є:

- поетапне завантаження параметрів моделей у той момент, коли відповідні ШНМ задіяні в алгоритмі, з наступним звільненням пам'яті. Такий підхід дозволяє суттєво зменшити споживання ОЗП під час виконання програми, проте знижує швидкодію системи через затримку, що виникає при завантаженні даних. Тому доцільним є розміщення параметрів моделей (які сумарно займають приблизно 1Mb) в ПЗП з швидким доступом та високою пропускнуою здатністю інтерфейсу. В даному випадку задіяно вбудовану флеш-пам'ять з інтерфейсом QSPI (частота 80МГц);
- квантування вхідних даних перед виконанням виведення та деквантування вихідних даних. Так як моделі було квантовано у форматі цілого 8-ми розрядного числа зі знаком, то обчислення в процесі виведення проводяться у даному типі даних (що має діапазон значень $[-127, 128]$). Тому як вхідні, так і вихідні дані матимуть даний тип даних. Квантування вхідних даних – процедура, спрямована на приведення вхідних даних, що мають певний діапазон значень до діапазону типу даних `int8`. Деквантування є зворотною операцією, що має на меті відтворення початкового динамічного діапазону.

Розроблений алгоритм покладений в основу вбудованої програми, яка виконує приймання цифрового сигналу оптичного сенсора, формує зображення, виконує виявлення та просторову локацію об'єктів у кадрі, виводить результат через UART, а також здійснює мережеву комунікацію з сервером шляхом надсилання HTTP-запитів.

Висновки до розділу 2

Даний розділ випускної кваліфікаційної роботи присвячено розробці та вдосконаленню методу виявлення та локації об'єктів в оптико-електронній системі технічного зору.

Розроблено імітаційну модель відгуку фотоприймача оптико-електронної системи, що моделює вихідне зображення на основі інформації про об'єкти, що спостерігаються (інтенсивність та просторове розташування), параметрів системи (лінійний розмір фотоприймача, фокусна відстань об'єктива, форма апертури, інтервал експозиції, світлова характеристика сенсору тощо), а також умов спостереження (зовнішні шуми, атмосферна турбулентність). Модельовані об'єкти вважаються точковими та далеко розташовані, тому їх зображення являє собою дифракційну пляму, що формується за законами дифракції Фраунгофера. Вплив атомосферної турбулентності за довгого інтервалу експозиції визначається фільтрацією низькочастотних компонент зображення, його міру визначається теорією Колмогорова. Імітаційну модель було реалізовано у програмному забезпеченні, за допомогою якого здійснено генерацію синтетичного набору даних. Останній містить параметри об'єктів у відповідності до змодельованих зображень та був застосований для розробки та тестування методу реєстрації об'єктів.

Створено вдосконалений метод виявлення просторової локації об'єктів, що дозволяє реєструвати зокрема малорозмірні та малоконтрастні об'єкти, навіть при їх близькому розташуванні. Результатом реєстрації є кількість виявлених об'єктів у кадрі та математичне сподівання координат центру кожного об'єкта (за наявності хоча б одного). Розроблений метод здійснює обробку у три етапи: виділення ознак із вхідного зображення, визначення кількості об'єктів шляхом класифікації цих ознак та передбачення координат центрів шляхом регресії. Кожен з етапів використовує глибоку згорткову нейронну мережу на основі архітектури MobileNet V1. Використання методів глибокого навчання

робить процес реєстрації об'єктів стійким до шумів та інших інтенсивнісних флуктуацій. Метод є адаптивним та може бути переналаштований під параметри конкретної оптико-електронної системи для забезпечення максимальної точності виявлення. Завдяки поетапному виконанню даний метод дозволяє здійснювати виявлення та локацію з досить високою точністю за умов обмежених обсягів оперативної пам'яті та низької швидкодії процесора.

Запропонований метод було реалізовано у вигляді алгоритму реєстрації об'єктів з використанням в якості апаратної платформи мікроконтролера ESP32-S3. Завдяки наявності до 8 МБ вбудованої SRAM та підтримці векторних інструкцій він забезпечує ефективне виконання згорткових нейронних мереж у вбудованих системах. Попередньо навчені моделі було квантовано за методом PTQ для цільового типу даних int8, що дозволило знизити використані обсяги ОЗП та ПЗП у майже 4 рази. Програмною платформою обрано MicroPython та фреймворк emlearn, що полегшує швидке прототипування та прикладні дослідження у сфері Embedded AI. Для проєктування запропоновано використати відлагоджувальну плату з інтерфейсом DVP для підключення оптичних сенсорів. Розроблений алгоритм реалізовано у вбудованій програмі, яка приймає цифровий сигнал сенсора, формує зображення, виконує детекцію та локалізацію об'єктів, а результати передає через UART та надсилає на сервер за допомогою HTTP-запитів.

2 ЕКСПЕРИМЕНТАЛЬНІ ДОСЛІДЖЕННЯ

3.1 Результати моделювання вихідних сигналів фотоприймача в складних умовах

На рисунку 3.1 наведено результати моделювання характеристик оптичної системи у випадку ідеальних умов поширення світла, тобто за відсутності атмосферної турбулентності. Двовимірне представлення оптичної передавальної функції (OTF) демонструє симетричний розподіл амплітуд у частотній області. Центральна яскраво виражена зона відповідає низьким просторовим частотам, які передаються системою без суттєвих втрат. Одновимірне представлення демонструє спад амплітуди OTF зі зростанням просторової частоти. Це відображає фундаментальне обмеження роздільної здатності оптичної системи: високі частоти (дрібні деталі зображення) передаються з меншою ефективністю.

Двовимірне зображення функції розсіювання точки (PSF) ілюструє форму світлової плями, яку створює система при відтворенні точкового джерела. Центральний максимум відповідає основній енергії сигналу, а навколишні кільця є результатом дифракції на апертурі. Горизонтальний переріз показує характерний профіль для дифракційної картини: різкий центральний пік та осциляції, що поступово згасають із віддаленням від центру.

Отже, результати моделювання підтверджують, що за відсутності турбулентності оптична система зберігає симетрію OTF та PSF та передає просторові частоти відповідно до теоретичних обмежень, пов'язаних із дифракцією.

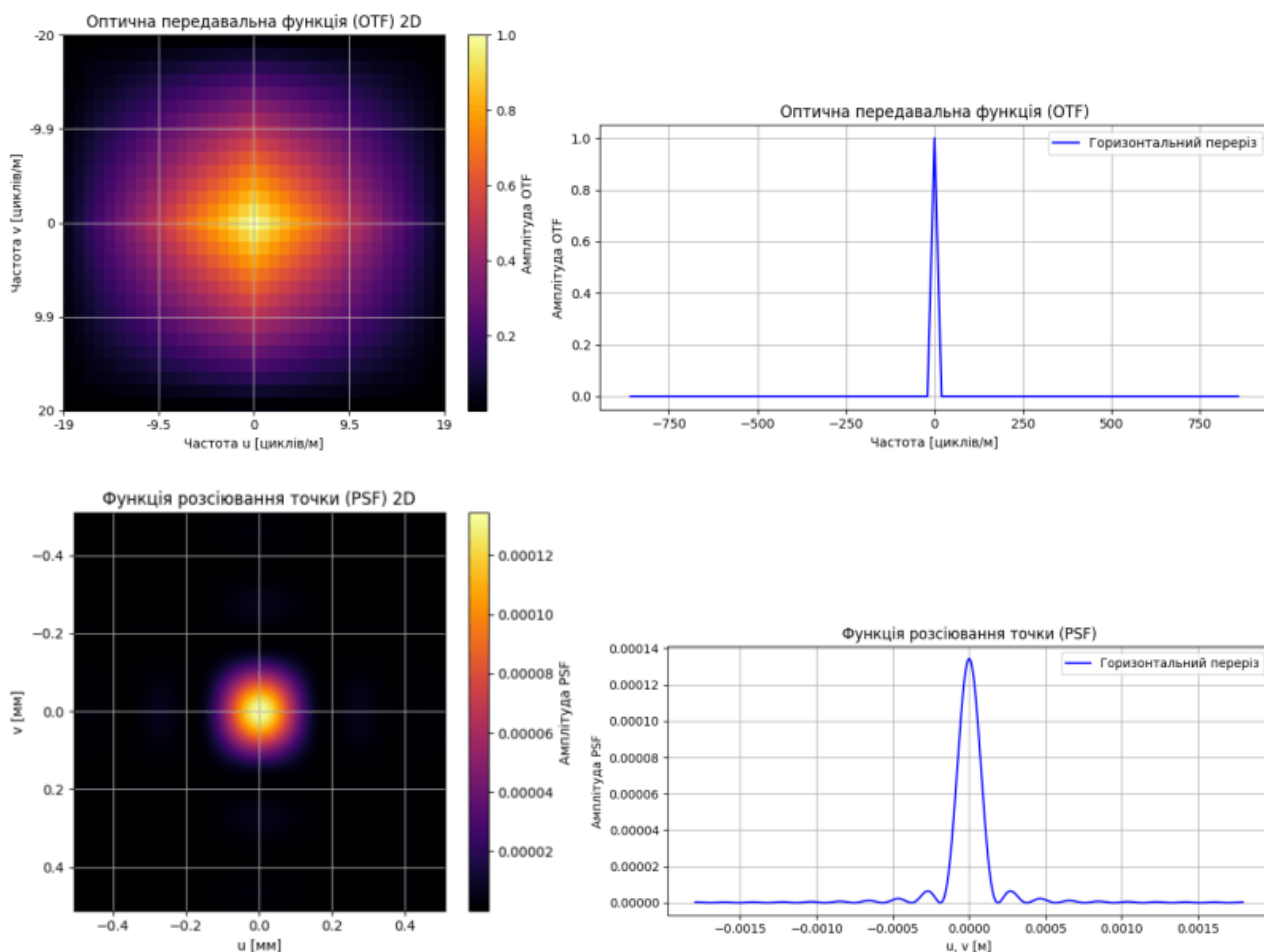


Рисунок 3.1 – Оптична передавальна функція та функція розсіювання точки оптичної системи за відсутності турбулентності

На рисунку 3.2 наведено результати моделювання вихідних сигналів фотоприймача за умов сильної атмосферної турбулентності ($r_0 = 1 \cdot 10^{-5}$ м). При моделюванні прийнято час експозиції набагато більший за інтервал замороженості атмосфери, тому результуючий вплив турбулентності є усередненим у часі та відповідає довгоекспозиційному режиму спостереження.

Оптична передавальна функція у двовимірному представленні демонструє вузький симетричний розподіл у просторовій частотній області. Центральна зона значно зменшена, амплітуди високих частот швидко згасають. Її горизонтальний переріз підтверджує більш різке падіння амплітуди на високих та середніх просторових частотах у порівнянні з випадком без турбулентності. Це свідчить про те, що дрібні деталі зображення значно гірше передаються системою.

Функція розсіювання точки ілюструє значне розширення центральної плями та зникнення дифракційних кілець. Горизонтальний переріз показує зниження контрасту, так як центральний пік стає менш вираженим, а навколишні осциляції згладжуються. Така зміна PSF означає втрату здатності системи до формування чітких зображень. Зображення точкового джерела перетворюється на розмиту пляму з низьким співвідношенням сигнальної та фонових складових.

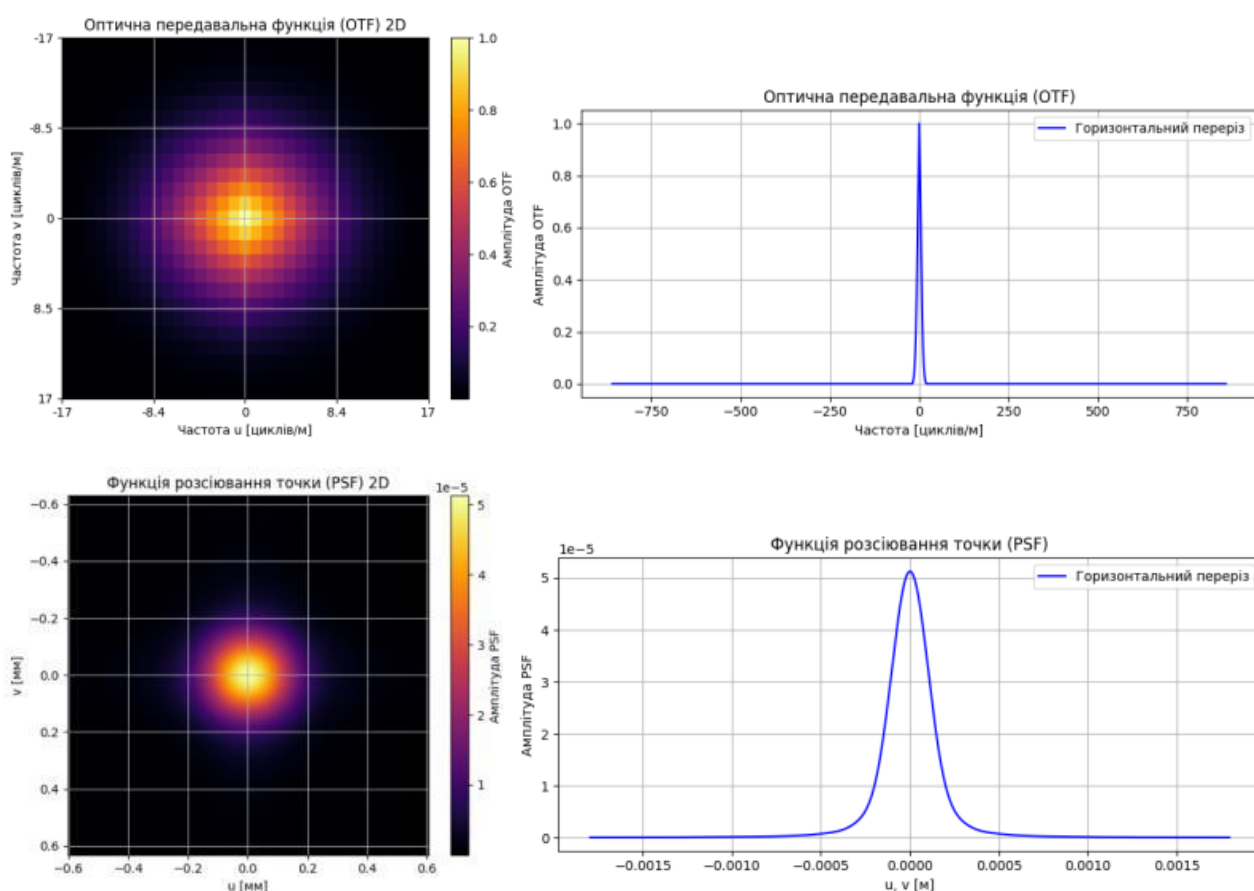
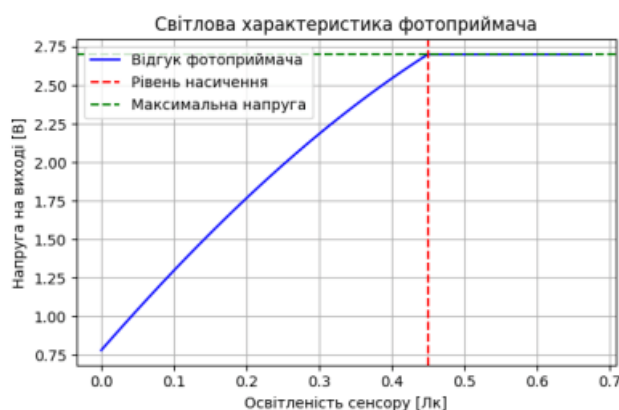


Рисунок 3.2 – Оптична передавальна функція та функція розсіювання точки оптичної системи за сильної турбулентності

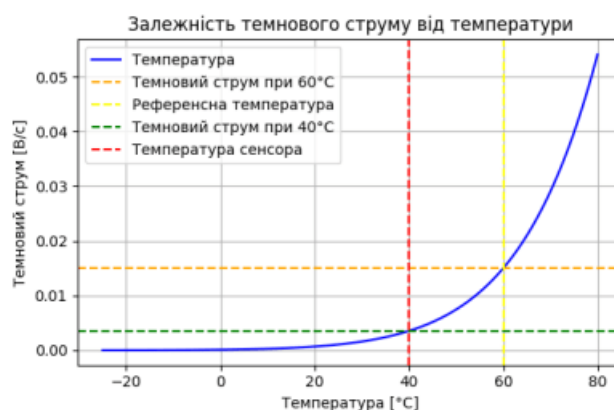
На графіку (рис. 3.3 (а)) проілюстровано змодельовану світлову характеристику фотоприймача, яка відображає залежність вихідного фотоструму від рівня освітленості. Спочатку спостерігається майже лінійне зростання струму зі збільшенням освітленості, що відповідає основному режиму роботи сенсора. Горизонтальна ділянка кривої демонструє насичення фотоприймача, після якого

подальше збільшення освітленості не призводить до пропорційного зростання сигналу.

Графік справа (рис. 3.3 (б)) ілюструє змодельовану залежність темнового струму від температури фотоприймача. Темновий струм зростає експоненційно зі збільшенням температури, що є наслідком термічної генерації носіїв заряду у фоточутливому матеріалі. Ця залежність визначає теплові обмеження роботи фотоприймача та показує, що для збереження низького рівня шумів необхідно забезпечувати низьку температуру.



а)



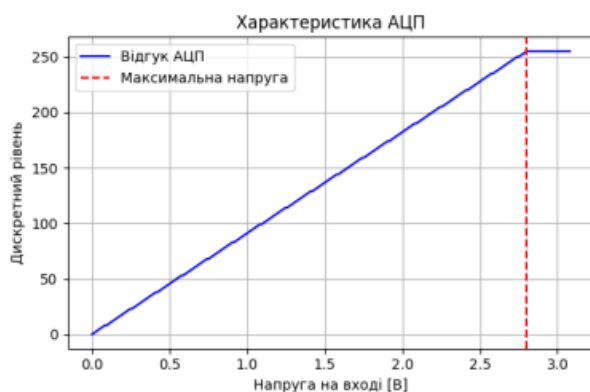
б)

а) світлова характеристика фотоприймача, б) залежність темнового струму від температури фотоприймача

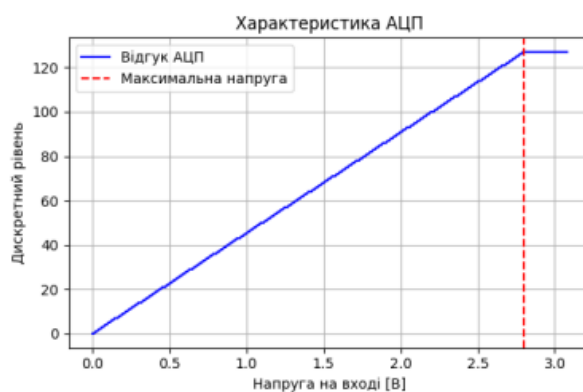
Рисунок 3.3 – Моделювання особливостей роботи оптичного приймача

На рисунку 3.4 наведено змодельовані залежності цифрового виходу від вхідної напруги для трьох варіантів аналого-цифрового перетворювача (АЦП) з різною глибиною квантування. Варіант 8-бітного АЦП має найбільшу кількість рівнів квантування (256), крива виглядає майже неперервною, що забезпечує досить високу точність відтворення аналогового сигналу: так як крок квантування мінімальний, похибка дискретизації візуально непомітна. У 7-бітного АЦП кількість рівнів зменшена до 128, вихідна характеристика стає «східчастою», що відображає більший крок квантування. 6-бітний АЦП має

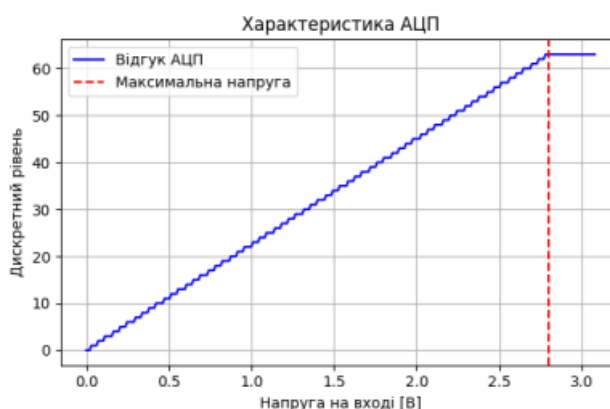
лише 64 рівні квантування, тому східчаста форма вихідної характеристики стає вираженою, похибка дискретизації значно зростає. Такий АЦП демонструє обмежену точність і може використовуватися лише там, де вимоги до роздільної здатності невисокі.



а)



б)



в)

а) 8 біт, б) 7 біт, в) 6 біт

Рисунок 3.4 – Залежність характеристики АЦП від його розрядності

На рисунку 3.5 показані результати моделювання відгуку оптико-електронної системи при реєстрації далекорозташованих об'єктів у двовимірному та тривимірному представленнях. У випадку відсутності турбулентності інтенсивність зображення зосереджена у компактній центральній області, що відповідає дифракційно обмеженому процесу формування зображення та забезпечує порівняно високу роздільну здатність.

Тривимірна гістограма демонструє два піки із значним проміжком між ними. Натомість за умов сильної турбулентності ($r_0 = 1 \cdot 10^{-5}$ м) спостерігається розмиття та розширення центральної плями, зниження контрасту та розмиття зображень об'єктів. Це зумовлено випадковими фазовими абераціями хвильового фронту, що інтегруються протягом експозиції та обмежують здатність системи до передавання високочастотних компонент зображення. Цей вплив призводить до часткового перекриття зображень двох далекорозташованих об'єктів, проміжок між піками інтенсивності стає менш вираженим, що створює складність при спробі реєстрації об'єктів.

Рисунок 3.6 ілюструє просторовий розподіл інтенсивності сигналу багатoelementного фотоприймача у випадку, коли в полі зору одночасно присутні кілька близькорозташованих джерел випромінювання з різними рівнями яскравості. Яскравіший об'єкт формує домінуючий максимум інтенсивності у центральній області, який визначає основний відгук системи. Об'єкт зі слабшою інтенсивністю, розташований поблизу, відтворюється менш виразно, так як його сигнал частково маскується «хвостами» функції розсіювання точки від сигналу сильнішого джерела. Взаємний вплив проявляється у вигляді асиметрії розподілу інтенсивності та зниження контрасту слабшого об'єкта, що ускладнює його реєстрацію та ідентифікацію. Результат моделювання підтверджує, що нерівноцінність яскравостей призводить до ефекту «пригнічення» слабкого сигналу сильним, що є типовим для оптико-електронних систем із обмеженою динамічною чутливістю.

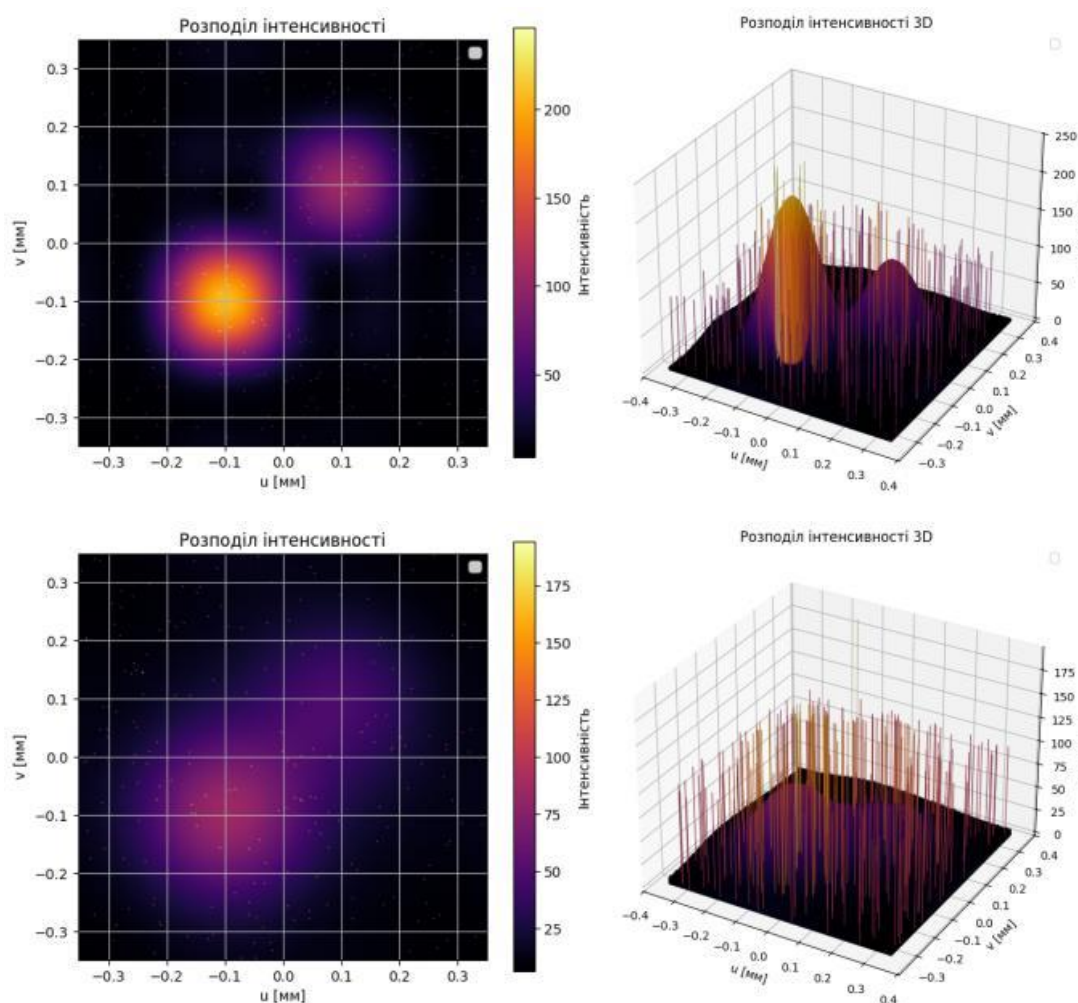


Рисунок 3.5 – Результат моделювання відгуку при реєстрації далекорозташованих об’єктів за умов сильної турбулентності (знизу) та її відсутності (згори)

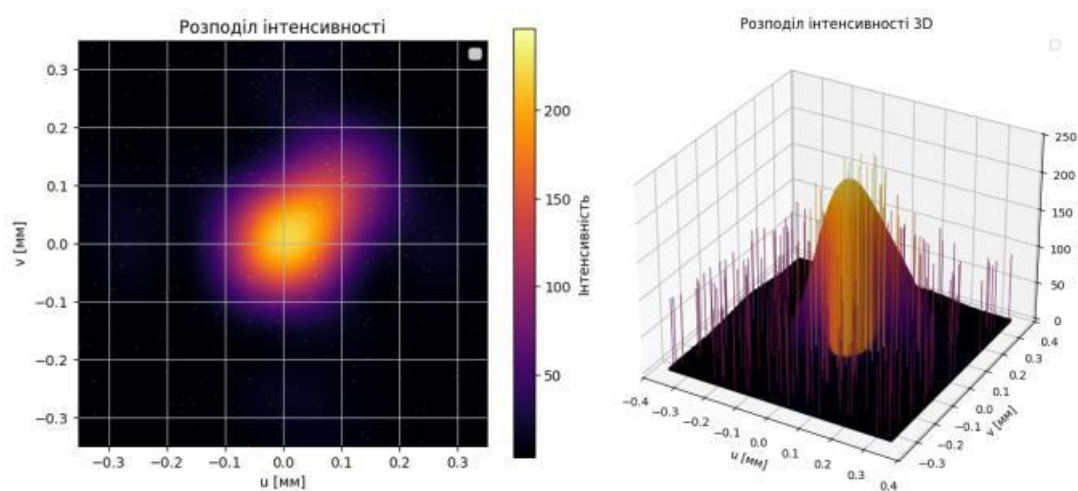


Рисунок 3.6 – Результат моделювання відгуку при реєстрації близькорозташованих об’єктів, нерівноцінних за яскравістю

Рисунок 3.7 ілюструє результати моделювання відгуку оптико-електронної системи при реєстрації близькорозташованих об'єктів за складних умов спостереження. На першій ілюстрації (згори) показано вплив сильної атмосферної турбулентності ($r_0 = 1 \cdot 10^{-5}$ м), яка призводить до розмиття центральних максимумів, розширення плям та зниження контрасту між об'єктами, аналогічно до попереднього випадку (реєстрації об'єктів на більшій відстані між ними).

Наступне зображення демонструє дію інтенсивних зовнішніх шумів та фонові складові. Для моделювання даної ситуації було додано адитивний Гаусовий шум із середнім значенням 1 та дисперсією 2,5. Відгук системи містить паразитні сигнали, які пригнічують об'єкти з низькою інтенсивністю, а яскравіші джерела втрачають чіткість контурів. Загалом, інтенсивна фонові складова та зовнішній шум знижують контрастність та співвідношення сигнал/шум отриманих зображень.

На останньому зображенні проілюстровано вплив сильних внутрішніх шумів фотоприймача, що проявляються у вигляді випадкових флуктуацій інтенсивності, зумовлених темновим шумом та шумом зчитування, що надають зернистої структури зображення. При моделюванні взято дисперсію шуму зчитування рівну 0,7; температуру сенсора взято рівною 50°C , що відповідає значенню темнового струму приблизно $8 \cdot 10^{-3}$ В/с. Такий характер сигналу ускладнює ідентифікацію об'єктів, оскільки інформація про їхню форму та яскравість значно спотворюється на рівні сенсора.

Отже, результати моделювання підтверджують, що кожен із розглянутих факторів: атмосферна турбулентність, зовнішні шуми та внутрішні шумові процеси, суттєво погіршує просторові характеристики системи та підвищує складність у задачах реєстрації близькорозташованих об'єктів.

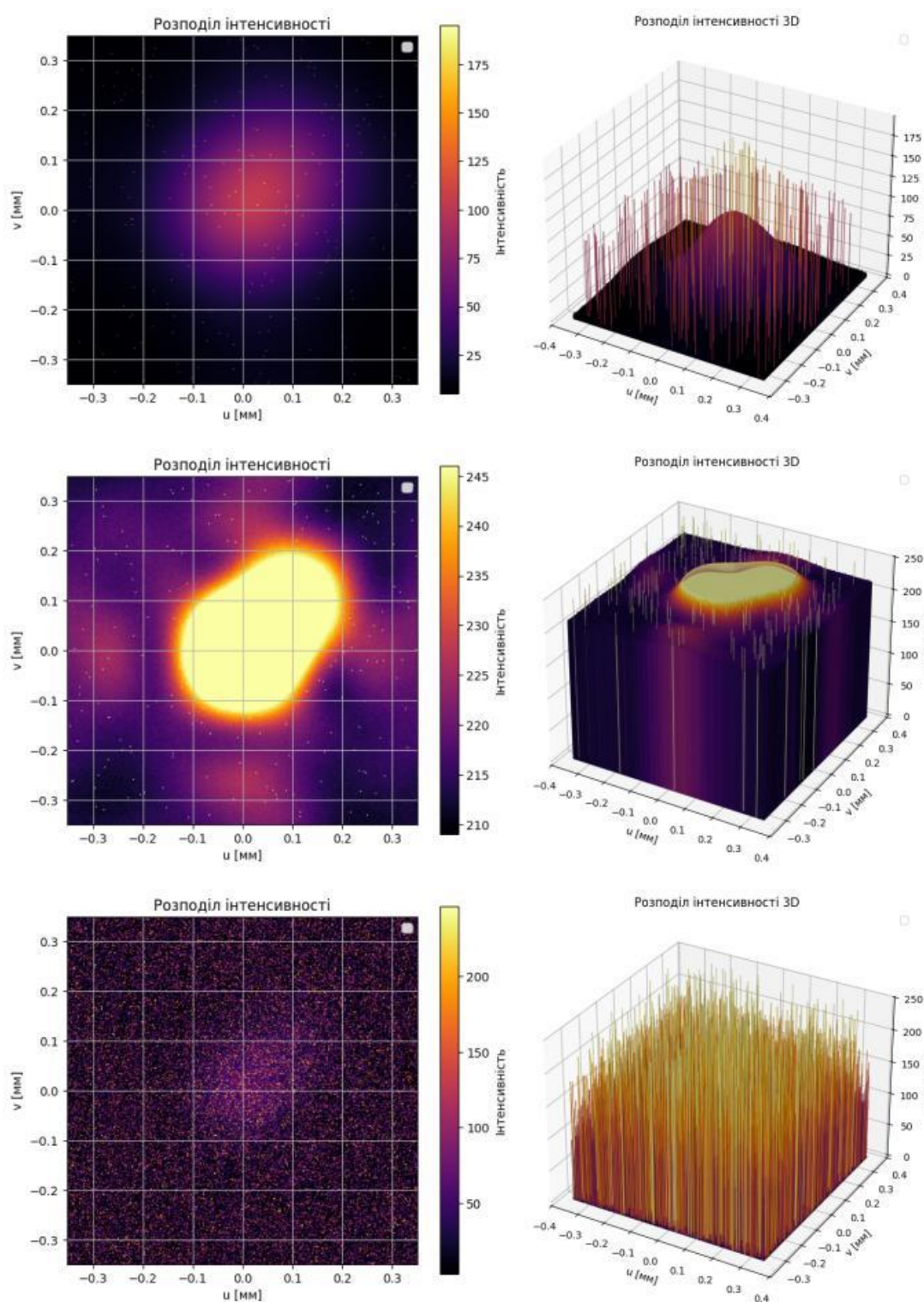


Рисунок 3.7 – Результати моделювання відгуку при реєстрації близькорозташованих об’єктів за складних умов спостереження, згори донизу: сильної атмосферної турбулентності; інтенсивних зовнішніх шумів та фоновій складовій; сильних внутрішніх шумів

3.2 Результати визначення статистичних параметрів близько розташованих сигналів

Для тестування точності методу реєстрації об'єктів було використано валідаційну вибірку, що містить зображення з різною кількістю об'єктів, розташованих на різних відстанях один від одного та за різних умов спостереження. Найбільшу дослідницьку цінність становлять саме випадки близькорозташованих об'єктів, оскільки вони створюють найбільш складні умови для алгоритму реєстрації. У таких ситуаціях дифракційні плями, що є зображеннями точок окремих об'єктів, перекриваються, що призводить до зростання похибок при визначенні координат їх центрів. Це ускладнює задачу регресії та вимагає від моделі здатності коректно розділяти сигнали навіть за умов їх високої щільності, шумів та низького контрасту.

У певних випадках, коли об'єкти розташовані на близьких відстанях, метод демонструє задовільні результати: центри визначаються з високою точністю. Такі приклади наведено на рис. 3.8, де видно узгодженість між істинними (позначені зеленим) та передбаченими (позначені червоним) координатами. Однак, виникають помилкові ситуації, де спостерігаються відхилення від істинних координат (рис. 3.9).

Отже, результати тестування підтверджують, що метод є ефективним для більшості ситуацій, включаючи реєстрацію близькорозташованих об'єктів, але має обмеження, пов'язані з неповнотою навчальних даних та особливостями архітектури.

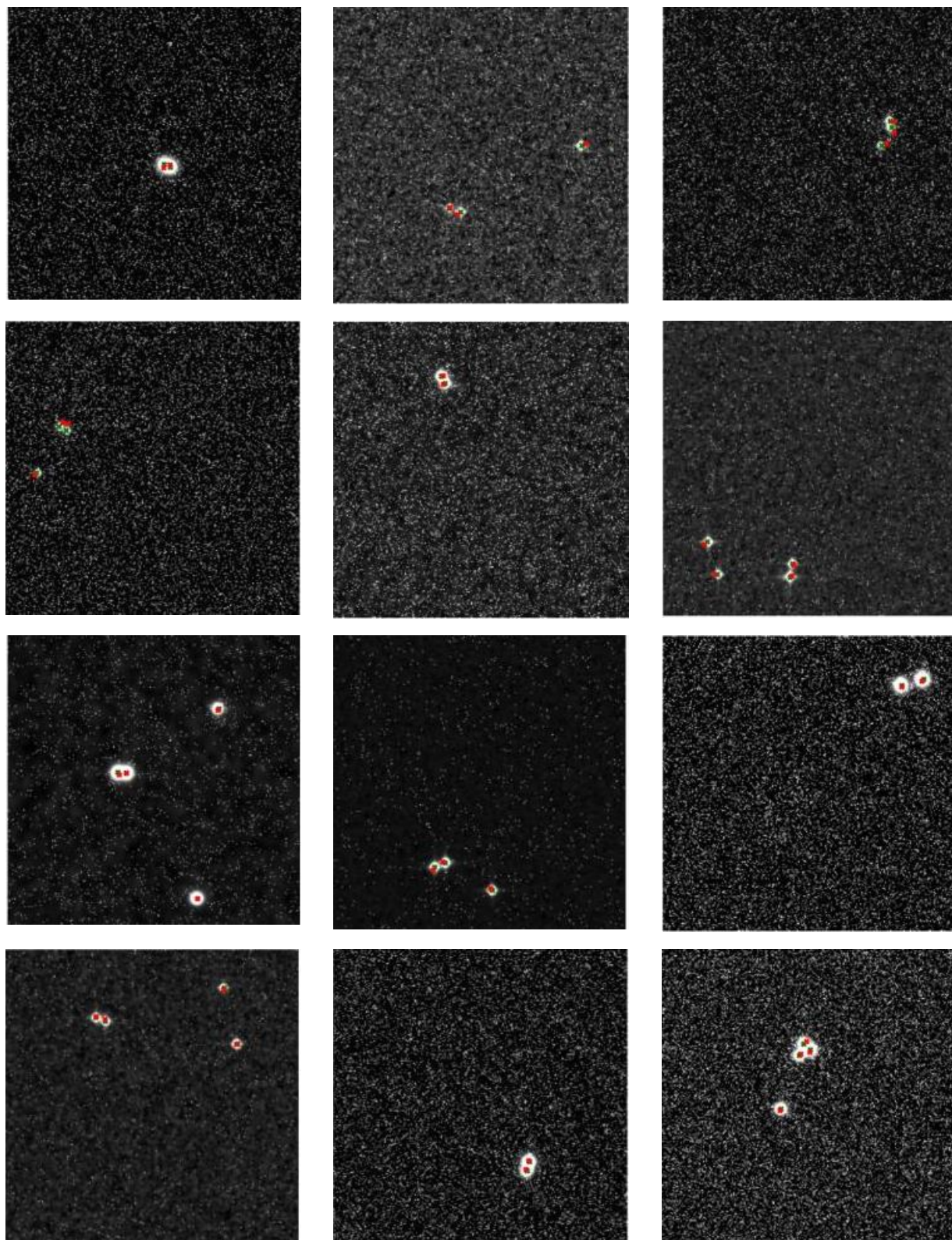


Рисунок 3.8 – Приклади задовільних результатів реєстрації
близькорозташованих об'єктів

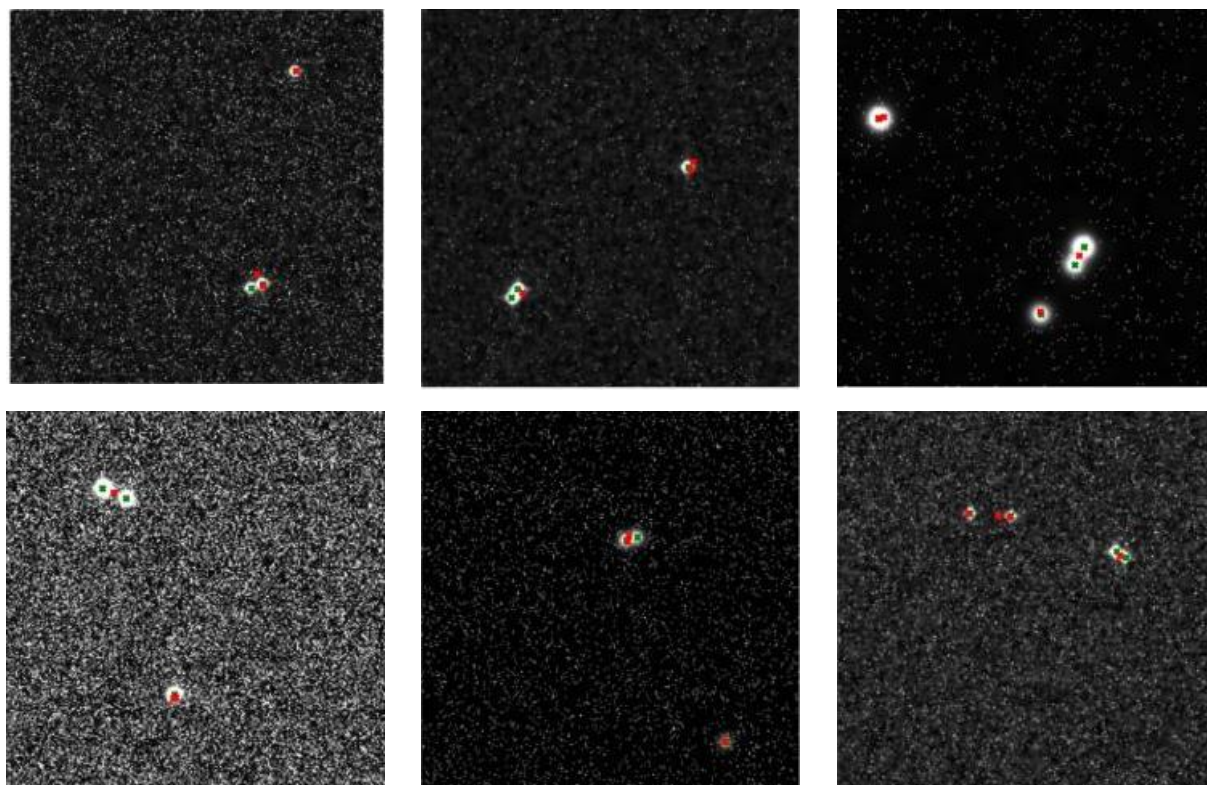


Рисунок 3.9 – Приклади помилок при реєстрації близькорозташованих об'єктів

3.3. Результати тестування системи технічного зору, реалізованої у вигляді вбудованої системи

Для тестування роботи системи технічного зору з використанням цільової апаратно-програмної платформи було зібрано тестовий стенд, зображений на рисунку 3.10. Він складається з відлагоджувальної плати на основі мікроконтролеру ESP32-S3 N16R8, блоку ПЗП у вигляді накопичувача SD-Card та блоку індикації з монохромним OLED-дисплеєм із роздільною здатністю 128x32 пікселі на базі контролеру SSD1306. Накопичувач було використано для збереження тестових зображень у растровому форматі BMP замість їх отримання за допомогою фотоприймача. Блок індикації використовується для відображення у текстовому поданні поточного номеру зображення, кількості виявлених об'єктів, а також часу обробки. В правій частині дисплею візуалізовано

просторове розташування виявлених об'єктів у вигляді засвічених пікселів всередині квадратної рамки, що позначає границі зображення (рис. 3.11).

Для тестування продуктивності системи було обрано вибірку з 20 зображень із різною кількістю об'єктів. Кожне з них було послідовно завантажено та оброблено за допомогою алгоритму виявлення, результат виведено на графічний дисплей, надіслано через інтерфейс UART та відправлено на віддалений сервер за допомогою HTTP-запиту.



Рисунок 3.10 – Зовнішній вигляд тестового станду

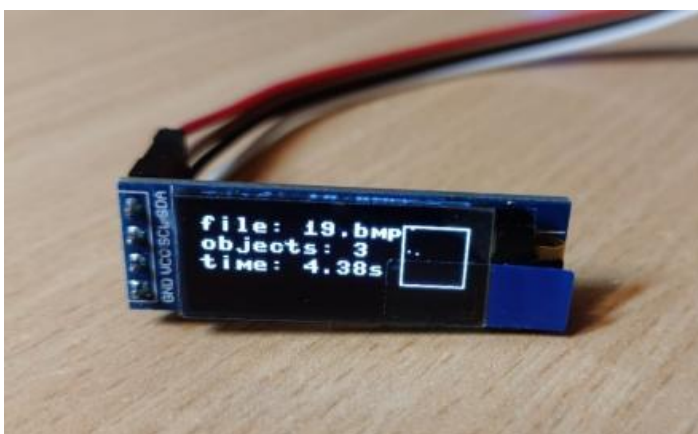


Рисунок 3.11 – Приклад результату реєстрації об'єктів

В таблиці 3.1 наведено результати тестування продуктивності системи технічного зору для кожного зображення з тестової вибірки із зазначенням часових міток та кількості виявлених об'єктів. Час, затрачений на ініціалізацію

периферійних підсистем, з'єднання з точкою доступу Wi-Fi та завантаження зображення з накопичувача не включені в дану оцінку. Очікувано, середній час обробки при відсутності виявлених сигналів є найменшим серед інших класів та складає 3309 мс. Це пов'язано з виконанням лише перших двох етапів реєстрації об'єктів: виділення ознак та визначення кількості сигналів. При наявності хоча б одного сигналу, додатково виконувався третій етап – визначення координат центрів, тому середній час обробки виріс та склав 4379 мс.

Таблиця 3.1 – Експериментальна оцінка часу виконання алгоритму

№ досліджу	Час початку, мс	Час завершення, мс	Час обробки, мс	Кількість виявлених об'єктів
1	152281	155589	3308	0
2	159716	164091	4375	1
3	168242	172619	4377	3
4	176771	181144	4373	4
5	185306	189687	4381	4
6	193867	198245	4378	4
7	202398	206792	4394	4
8	210945	215318	4373	5
9	219473	222782	3309	0
10	226916	231301	4385	1
11	235488	239869	4381	2
12	244022	248400	4378	3
13	252559	256936	4377	5
14	261094	265468	4374	4
15	269651	272960	3309	0
16	277089	281470	4381	1
17	285634	290007	4373	5
18	294171	298542	4371	2
19	302698	307092	4394	3
20	311243	315616	4373	4

Дана оцінка охоплює лише етап обробки зображення на стороні мікроконтролера та не враховує час ініціалізації камери, час зйомки (інтервал експозиції) та затримку, пов'язану з передачею отриманого зображення через інтерфейс та буферизацією. Отже, з використанням апаратної платформи ESP32-S3, розроблений алгоритм реєстрації об'єктів, може бути застосований для

рішень, де допустимий час отримання рішення має бути не більшим 5 секунд (без врахування вищезазначених факторів).

Додатково було здійснено експериментальну оцінку часу виконання основних етапів алгоритму реєстрації зображень. Аналіз розподілу часу виконання етапів алгоритму (рис. 3.12) показує, що найбільш ресурсомістким є процес виділення ознак, який займає 2030 мс і становить майже половину загального часу роботи системи. Це пояснюється складністю обчислень, пов'язаних із виконанням виведення за допомогою згорткової нейронної мережі з відносно великою розмірністю вхідних та вихідних даних, а також прихованого простору ознак. Отже, цей етап значною мірою визначає продуктивність алгоритму та є його основним «вузьким місцем». Етапи завантаження моделей для класифікації та регресії займають відповідно 460 мс і 597 мс, що свідчить про помірні витрати часу на передачу даних та ініціалізацію нейронної мережі. Подальше виконання класифікації (474 мс) та регресії (459 мс) має схожий рівень обчислювальної складності, що вказує на меншу складність цих процедур. Найменший внесок у загальний час має квантування ознак (136 мс) та інші процедури (деквантування результатів та керування виділенням пам'яті).

Отже, оптимізація системи має бути спрямована насамперед на скорочення часу виконання етапу виділення ознак, що дозволить підвищити загальну швидкодію та ефективність реєстрації об'єктів.

**Розподіл часу виконання етапів алгоритму
реєстрації об'єктів**

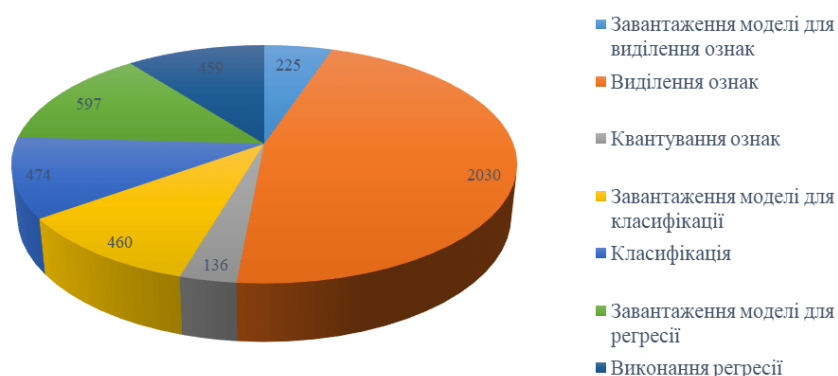


Рисунок 3.12 – Розподіл часу виконання етапів алгоритму

Висновки до розділу 3

Даний розділ кваліфікаційної роботи було присвячено експериментальним дослідженням окремих аспектів роботи системи технічного зору, зокрема методу реєстрації об'єктів за різних умов спостереження. За допомогою розробленої імітаційної моделі відгуку оптико-електронної системи було виконано моделювання характеристик оптичної системи за ідеальних умов поширення світла та за умов сильної атмосферної турбулентності. Проаналізовано зміни оптичної передавальної функції та функції розсіювання точки, визначено вплив турбулентності на роздільну здатність та контрастність зображень. Досліджено ефекти взаємного впливу близькорозташованих джерел випромінювання з різною яскравістю, а також вплив зовнішніх та внутрішніх шумів на якість сигналу.

Проведено тестування алгоритму реєстрації об'єктів на вибірці з різною кількістю об'єктів та різними умовами спостереження. Оцінено точність визначення координат центрів та виявлено помилкові ситуації.

Для перевірки продуктивності системи в умовах цільової апаратно-програмної платформи було створено стенд на базі мікроконтролера ESP32-S3, що виконує завантаження тестових зображень, їх обробку алгоритмом, відображення результатів на графічному дисплеї та їх передачу на віддалений сервер. Виконано експериментальну оцінку часу роботи основних етапів алгоритму, визначено найбільш ресурсомісткі операції та сформульовано напрямки подальшої оптимізації з метою підвищення швидкодії.

ВИСНОВКИ

В ході проведеного дослідження у рамках випускної кваліфікаційної роботи було розроблено вдосконалений метод реєстрації малорозмірних та малококонтрастних об'єктів для систем технічного зору, що використовує стохастико-детерміновану теорію оброблення сигналів та методи глибокого навчання. Він був реалізований у вигляді алгоритму виявлення та локації об'єктів для вбудованої системи. Таким чином, основна мета кваліфікаційної роботи – досягнута. Всі поставлені задачі були вирішені. В результаті проведеного дослідження було отримано наступні основні результати:

Проведено аналіз основних принципів формування сигналів і зображень в оптико-електронних системах, їх параметрів та просторових характеристик. Визначено основні чинники, що впливають на якість вихідного сигналу та достовірність ідентифікації об'єктів за складних умов спостереження.

Розглянуто найбільш поширені методи виявлення об'єктів в системах технічного зору, визначено їх особливості та специфіку застосування. Обґрунтовано важливість розробки та вдосконалення методів реєстрації об'єктів на основі гібридних підходів, що поєднують стохастико-детермінований опис сигналу та сучасні методи штучного інтелекту.

Розроблено імітаційну модель відгуку фотоприймача оптико-електронної системи, що дозволяє моделювати вихідне зображення на основі даних про просторове розташування точкових об'єктів, параметрів системи та умов спостереження. Модель враховує комплекс фізичних явищ: дифракцію на апертурі, атмосферну турбулентність, шумові процеси; особливості фотоприймача: світлову характеристику, темновий струм; особливості реєстрації сигналу: підсилення, аналого-цифрове перетворення.

Створено програмне забезпечення, що реалізує розроблену імітаційну модель. З його допомогою здійснено генерацію синтетичного набору даних, що

містить пари вхідних параметрів та відповідних змодельованих зображень, у широкому різноманітті просторового розташування та умов спостереження.

Розроблено вдосконалений метод реєстрації близькорозташованих малорозмірних та малококонтрастних об'єктів для систем технічного зору. Він дозволяє здійснювати виявлення до п'яти об'єктів одночасно та визначення математичного сподівання їх центрів. Завдяки модульній архітектурі та використанню глибоких згорткових нейронних мереж, він поєднує високу стійкість до шумів та достатню точність із низькими вимогами до обчислювальних ресурсів.

Здійснено аналіз апаратних засобів для реалізації системи технічного зору для запропонованої задачі. В результаті було обрано мікроконтролер ESP32-S3 з позицій компромісу між обчислювальними можливостями, енергоспоживанням та наявністю периферійних інтерфейсів. Відповідно до обраної апаратної платформи було визначено програмні засоби для реалізації вбудованої програми.

Було створено алгоритм, що реалізує розроблений метод реєстрації об'єктів. На його основі розроблено вбудовану програму, яка приймає цифровий сигнал сенсора, формує зображення, виконує детекцію та локалізацію об'єктів, передає результати на віддалений сервер.

Проведено тестування імітаційної моделі відгуку оптико-електронної системи шляхом дослідження вихідних зображень при різних вхідних параметрах. Здійснено тестування алгоритму реєстрації об'єктів за різних умов спостереження. Оцінено точність визначення координат центрів, знайдено помилкові відмітки виявлення. Виконано оцінку часу роботи основних етапів алгоритму системи в умовах цільової апаратно-програмної платформи. Результати експериментальних досліджень продемонстрували високу достовірність виявлення та прийнятні часові показники.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Tzampazaki M. Machine Vision – Moving from Industry 4.0 to Industry 5.0. *Applied Sciences (Appl. Sci.)*. 2024. Vol. 14, Iss. 4. Art. 1471. DOI: 10.3390/app14041471.
2. Haffner O., Kučera E., Rosinová D. Applications of Machine Learning and Computer Vision in Industry 4.0. *Applied Sciences*. 2024. Vol. 14, Iss. 6. Art. 2431. DOI: 10.3390/app14062431.
3. Javaid M. Exploring impact and features of machine vision for progressive industry 4.0 culture / Mohd Javaid *et al.* *Sensors International*. 2022. Vol. 3. Art. 100132. DOI: 10.1016/j.sintl.2021.100132.
4. Trigka M., Dritsas E. A comprehensive survey of deep learning approaches in image processing. *Sensors*. 2025. Vol. 25, No. 2. Art. 531. DOI: 10.3390/s25020531.
5. Grujić K. A Review of Thermal Spectral Imaging Methods for Monitoring High-Temperature Molten Material Streams. *Sensors*. 2023. Vol. 23. Art. 1130. DOI: 10.3390/s23031130.
6. Pinto-Coelho L. How Artificial Intelligence Is Shaping Medical Imaging: a review / L. Pinto-Coelho *et al.* *Frontiers*. 2023. URL: <https://www.ncbi.nlm.nih.gov/pmc/> (accessed on: 01.12.2025).
7. Ghazal S., Munir A., Qureshi W. S. Computer vision in smart agriculture and precision farming: Techniques and applications. *Artificial Intelligence in Agriculture*. 2024. Vol. 13. P. 64–83. DOI: 10.1016/j.aiia.2024.06.004.
8. Gounella R. A Review of Optical Sensors in CMOS / R. Gounella *et al.* *Electronics*. 2024. Vol. 13. Art. 691. DOI: 10.3390/electronics13040691.
9. Захожай О. І., Крохмаль А. В. Екстенціональний підхід до розпізнавання растрових зображень на основі їх символічних перетворень. *Наукові вісти Днівського університету*. 2023. № 25. DOI: 10.33216/2222-3428-2023-25-2.

10. Paúr M. Achieving the ultimate optical resolution / M. Paúr та ін. *Optica*. 2016. Vol. 3. P. 1144–1147.
11. Охрімчук О. Б. Оптиелектронні сенсори. *Перспективні технології та прилади*. 2023. № 22. Р. 16. DOI: 10.36910/6775-2313-5352-2023-22-16.
12. Verbytskyi V. H., Okhrimchuk O. B. Methods to Enhance the Sensitivity of Optoelectronic Sensors. *Open Information and Computer Integrated Technologies*. 2025. № 104. P. 121–132. DOI: 10.32620/oikit.2025.104.08.
13. Hampson K. M. Adaptive optics for high-resolution imaging / K. M. Hampson et al. *Nature Reviews Methods Primers*. 2021. Vol. 1. Art. 68. DOI: 10.1038/s43586-021-00068-5.
14. Rais M. Improving wavefront sensing with a Shack–Hartmann device / M. Rais та ін. *Applied Optics*. 2016. Vol. 55. P. 7836–7846.
15. Кобилін О. А., Творошенко І. С. Методи цифрової обробки зображень : навч. посіб. Харків. 2021. 248 с. DOI: 10.30837/978-966-659-295-1.
16. Андрікевич С. А., Тужанський С. Є. Застосування методу адаптивного вирівнювання гістограм для підвищення контрастності зображень очного дна. *Репозиторій ВНТУ*. 2025. URL: <https://ir.lib.vntu.edu.ua/handle/123456789/48219> (дата звернення: 01.12.2025).
17. Липовий А. Алгоритми покращення якості зображення за допомогою глибинних нейронних мереж: порівняльний аналіз сучасних методів. *Measuring and Computing Devices in Technological Processes*. 2025. № 82. Р. 37. DOI: 10.31891/2219-9365-2025-82-37.
18. Методичні вказівки до лабораторних робіт з дисципліни “Тепловізійні та телевізійні пристрої і системи ” для студентів денної і заочної форм навчання галузі знань 17 «Електроніка, автоматизація та електронні комунікації», спеціальні 171 «Електроніка», освітньо-професійної програми «Електронні прилади та пристрої» [Електронний ресурс] / упоряд: Т.О. Стрілкова, О.С. Калмиков - Електронне видання. - Харків: ХНУРЕ, 2023. - 49 с.

- 19.Стрілкова Т. О. Вплив моделі опису сигналів на розрахунки показників ефективності оптико-електронних систем / Т. О. Стрілкова та ін. *Східно-Європейський журнал передових технологій*. 2020. Vol. 4. № 5 (106). Р. 41–50.
- 20.Стрілкова Т. О. Розвиток стохастико-детермінованої теорії прийому та обробки сигналів в оптико-електронних системах : автореф. дис. д-ра техн. наук : 05.11.07 «Оптичні прилади та системи» / Стрілкова Тетяна Олександрівна ; М-во освіти і науки України, Київський політехнічний інститут ім. Ігоря Сікорського. – Київ, 2017. – 40 с.
- 21.Strelkova T. Theoretical methods of images processing in optoelectronic systems / T. Strelkova et al. *Developing and Applying Optoelectronics and Machine Vision*. 2016. Chapter 6. P. 181–206. URL: <http://www.igi-global.com/book/developing-applying-optoelectronics-machine-vision/147652> (accessed on 01.12.2025).
- 22.Ruderman D. L., Bialek W. Statistics of natural images: Scaling in the woods. *Physical Review Letters*. 1994. Vol. 73. No. 6. P. 814–817.
- 23.Колобродов В. Г., Лихоліт М. І. Проектування тепловізійних і телевізійних систем спостереження. Київ. 2007. 180 с.
- 24.Yang F. Bits from photons: Oversampled image acquisition using binary Poisson statistics / F. Yang et al. *IEEE Transactions on Image Processing*. 2012. Vol. 21. No. 4. P. 1421–1436.
- 25.Seitz P., Theuwissen A. J. P. Single-Photon Imaging. 2011.
- 26.Konnik M., Welsh J. High-level numerical simulations of noise in CCD and CMOS photosensors: review and tutorial. 2014.
- 27.Стрілкова Т. О., Крохмаль А. В. Формування зображень в оптико-електронних системах з урахуванням умов спостереження. 29-й Міжнародний молодіжний форум «Радіoeлектроніка та молодь у XXI столітті». Зб. матеріалів форуму. Т. 1. – Харків: ХНУРЕ. 2025. – 162 с.
- 28.Стрілкова Т. О. Розвиток стохастико-детермінованої теорії прийому та обробки сигналів в оптико-електронних системах : дис. д-ра техн. наук :

- 05.11.07 «Оптичні прилади та системи» / Стрілкова Тетяна Олександрівна ; М-во освіти і науки України, Київський політехнічний інститут ім. Ігоря Сікорського. – Київ, 2017. – 343 с.
29. Порогування (обробка зображень). *Wikipedia*. : веб- сайт. URL: [https://uk.wikipedia.org/wiki/%D0%9F%D0%BE%D1%80%D0%BE%D0%B3%D1%83%D0%B2%D0%B0%D0%BD%D0%BD%D1%8F_\(%D0%BE%D0%B1%D1%80%D0%BE%D0%B1%D0%BA%D0%B0_%D0%B7%D0%BE%D0%B1%D1%80%D0%B0%D0%B6%D0%B5%D0%BD%D1%8C\)](https://uk.wikipedia.org/wiki/%D0%9F%D0%BE%D1%80%D0%BE%D0%B3%D1%83%D0%B2%D0%B0%D0%BD%D0%BD%D1%8F_(%D0%BE%D0%B1%D1%80%D0%BE%D0%B1%D0%BA%D0%B0_%D0%B7%D0%BE%D0%B1%D1%80%D0%B0%D0%B6%D0%B5%D0%BD%D1%8C)) (дата звернення: 01.12.2025).
30. Otsu N. A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man, and Cybernetics*. 1979. Vol. 9. No. 1. P. 62–66. DOI: 10.1109/TSMC.1979.4310076.
31. Krokhmal V., Krokhmal A., Tarasov V., Rudniev Ye. Application of the Combined Method of Raster Image Recognition in the Computer Vision System of unmanned Vehicles in the Mining Industry. *Geo-Technical Mechanics*. 2024. № 171. DOI: 10.15407/geotm2024.171.098.
32. Крохмаль А. В. Підходи до побудови систем розпізнавання образів в контексті обробки візуальної інформації. *Цілі сталого розвитку: проблеми і можливості досягнення в Україні та світі: матеріали VIII Всеукраїнської наук.- практ. конф. здобувачів вищої освіти та молодих вчених, 23-24 жовтня 2024 р., м. Київ. - Київ: [Східноукр. нац. ун-т ім.В. Даля], 2024. - 376 с.*
33. Hashemi N. S. Template Matching Advances and Applications in Image Analysis / N. S. Hashemi et al. *American Scientific Research Journal for Engineering, Technology, and Sciences*. 2016. Vol. 26. No. 3. P. 91–108.
34. Han Y. Reliable Template Matching for Image Detection in Vision Sensor Systems. *Sensors*. 2021. Vol. 21. Art. 8176. DOI: 10.3390/s21248176.
35. Swaroop P., Sharma N. An Overview of Various Template Matching Methodologies in Image Processing. *International Journal of Computer Applications*. 2016. Vol. 153. P. 8–14. DOI: 10.5120/ijca2016912165.

- 36.Mittal H. A comprehensive survey of image segmentation: clustering methods, performance parameters, and benchmark datasets / H. Mittal et al. *Multimedia Tools and Applications*. 2022. Vol. 81. No. 24. P. 35001–35026. DOI: 10.1007/s11042-021-10594-9.
- 37.Wiharto W., Suryani E. The Comparison of Clustering Algorithms K-Means and Fuzzy C-Means for Segmentation Retinal Blood Vessels. *Acta Informatica Medica*. 2020. Vol. 28. No. 1. P. 42–47. DOI: 10.5455/aim.2020.28.42-47.
- 38.Hu X. Research on the Region-Growing and Segmentation Technology of Micro-Particle Microscopic Images Based on Color Features / X. Hu et al. *Symmetry*. 2021. Vol. 13. Art. 2325. DOI: 10.3390/sym13122325.
- 39.Malik J., Shi J.-B. Normalized Cuts and Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2000.
- 40.Reddy K. S. G. Unseggnet: Unsupervised image segmentation using graph neural networks / K. S. G. Reddy et al. 2024. arXiv:2405.06057. DOI: 10.48550/arXiv.2405.06057.
- 41.Захожай О.І., Крохмаль А.В. Метод і програмне забезпечення символічного перетворення растрових зображень. *Наукові вісті Дніпровського університету*. №26. 2024р. DOI: 10.33216/2222-3428-2024-26-1.
- 42.Bouman C. A. Markov Random Fields and Stochastic Image Models. Purdue University. 1995. URL: <http://dynamo.ecn.purdue.edu/~bouman> (accessed on 01.12.2025).
- 43.Kato Z. Markov Random Fields in Image Segmentation / Z. Kato, J. Zerubia. *Foundations and Trends in Signal Processing*. 2011. Vol. 5. DOI: 10.1561/20000000035.
- 44.Arulampalam M. A Tutorial on Particle Filters for Online Nonlinear/Non-Gaussian Bayesian Tracking / M. Arulampalam et al. *IEEE Transactions on Signal Processing*. 2002. Vol. 50. P. 174–188. DOI: 10.1109/78.978374.
- 45.Ghahramani Z. An introduction to hidden Markov models and Bayesian networks. *Hidden Markov Models: Applications in Computer Vision*. World Scientific Publishing Co., Inc., USA. 2001. P. 9–42.

46. Bruhn A. Lucas/Kanade Meets Horn/Schunck: Combining Local and Global Optic Flow Methods / A. Bruhn et al. *International Journal of Computer Vision*. 2005. Vol. 61. P. 211–231. DOI: 10.1023/B:VISI.0000045324.43199.43.
47. Black M., Anandan P. The Robust Estimation of Multiple Motions: Parametric and Piecewise-Smooth Flow Fields. *Computer Vision and Image Understanding*. 1996. Vol. 63. P. 75–104. DOI: 10.1006/cviu.1996.0006.
48. Simonyan K. Two-stream convolutional networks for action recognition in videos / K. Simonyan et al. *Proceedings of the 28th International Conference on Neural Information Processing Systems (NIPS'14)*. 2014. Vol. 1. Cambridge, MA, USA: MIT Press. P. 568–576.
49. Sharma R., Gupta S. A Survey on Moving Object Detection and Tracking Based On Background Subtraction. *The Oxford Journal of Intelligent Decision and Data Science*. 2018. P. 55–62. DOI: 10.5899/2018/ojids-00041.
50. Simsek E., Ozyer B. Selected Three Frame Difference Method for Moving Object Detection. *International Journal of Intelligent Systems and Applications in Engineering*. 2021. Vol. 9. No. 2. P. 48–54. DOI: 10.18201/ijisae.2021.233.
51. Zhang Q. L. Moving Object Detection Method Based on the Fusion of Online Moving Window Robust Principal Component Analysis and Frame Difference Method / Q. L. Zhang et al. *Neural Processing Letters*. 2024. Vol. 56. Art. 55. DOI: 10.1007/s11063-024-11463-w.
52. Ismail M. M. A Study on atmospheric turbulence with Shearing Interferometer wavefront sensor. *International Journal of Innovative Research in Science, Engineering and Technology*. 2015. Vol. 4. P. 991–999. DOI: 10.15680/IJIRSET.2015.0403039.
53. Perrin S. Fourier optics: basic concepts. *arXiv preprint*. 2018. DOI: 10.48550/arXiv.1802.07161.
54. Fiete R. D. Modeling the optical transfer function in the imaging chain / R. D. Fiete, B. D. Paul. *Optical Engineering*. 2014. Vol. 53. No. 8. Art. 083103. DOI: 10.1117/1.OE.53.8.083103.

55. Catrysse P., Wandell B. Optical efficiency of image sensor pixels. *Journal of the Optical Society of America A*. 2002. Vol. 19. P. 1610–1620. DOI: 10.1364/JOSAA.19.001610.
56. Bigas M. Review of CMOS image sensors / M. Bigas et al. *Microelectronics Journal*. 2006. Vol. 37. P. 433–451. DOI: 10.1016/j.mejo.2005.07.002.
57. Abarca A. A CMOS Image Sensor Dark Current Compensation Using In-Pixel Temperature Sensors / A. Abarca. *Sensors*. 2023. Vol. 23. Art. 9109. DOI: 10.3390/s23229109.
58. Kuusk J. Dark Signal Temperature Dependence Correction Method for Miniature Spectrometer Modules. *Journal of Sensors*. 2011. Art. 608157. 9 p. DOI: 10.1155/2011/608157.
59. Tian H. Noise Analysis in CMOS Image Sensors. Stanford: Stanford University, Department of Applied Physics. 2000. URL: <https://books.google.com.ua/books?id=tQAhAQAAIAAJ> (accessed on 01.12.2025).
60. Крохмаль А.В., Стрілкова Т.О. Інтелектуальна система виявлення близькорозташованих малорозмірних об'єктів: архітектура та реалізація. *Матеріали Міжнародної науково-технічної конференції «Інформаційно-комунікаційні технології та кібербезпека» (IKTK-2025)*. Харків, ХНУРЕ, 2025, С. 317-321.
61. Howard A. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications / A. Howard et al. 2017. DOI: 10.48550/arXiv.1704.04861.
62. Deng J. ImageNet: A Large-Scale Hierarchical Image Database / J. Deng et al. *IEEE Computer Vision and Pattern Recognition (CVPR)*. 2009.
63. Chen L. How to Backpropagate through Hungarian in Your DETR? / L. Chen et al. *ArXiv*. 2022. DOI: 10.48550/arXiv.2211.14448.
64. TensorFlow - An end-to-end platform for machine learning : веб- сайт. URL: <https://www.tensorflow.org> (дата звернення: 01.12.2025).

65. Model Quantization Methods In TensorFlow Lite. *Study Machine Learning*. : веб-сайт. URL: <https://studymachinelearning.com/model-quantization-methods-in-tensorflow-lite> (дата звернення: 01.12.2025).
66. ESP32-S3 Series Datasheet. Version 2. URL: https://www.espressif.com/sites/default/files/documentation/esp32-s3_datasheet_en.pdf (дата звернення: 01.12.2025).
67. Emlearn: Machine learning for microcontroller and embedded systems : веб-сайт. URL: <https://github.com/emlearn/emlearn> (дата звернення: 01.12.2025). DOI: 10.5281/zenodo.2589393.