

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет інформаційно-аналітичних технологій та менеджменту

(повна назва)

Кафедра прикладної математики

(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА
Пояснювальна записка

рівень вищої освіти другий (магістерський)

Математичні моделі та методи використання інфрачервоного
зображення для розпізнавання об'єктів

(тема)

Виконав:

студент 2 курсу, групи САУМ-22-1

Зошук М.М.

(прізвище, ініціали)

Спеціальність 124 Системний аналіз

(код і повна назва спеціальності)

Тип програми освітньо-професійна

(освітньо-професійна або освітньо-наукова)

Освітня програма Системний аналіз і управління

(повна назва освітньої програми)

Керівник доц. Єсілевський В.С.

(посада, прізвище, ініціали)

Допускається до захисту

Зав. кафедри ПМ

(підпис)

Сидоров М.В.

(прізвище, ініціали)

2024 р.

Харківський національний університет радіоелектроніки

Факультет інформаційно-аналітичних технологій та менеджменту

Кафедра прикладної математики

Рівень вищої освіти другий (магістерський)

Спеціальність 124 Системний аналіз

(код і повна назва)

Тип програми освітньо-професійна

(освітньо-професійна або освітньо-наукова)

Освітня програма Системний аналіз і управління

(повна назва)

ЗАТВЕРДЖУЮ:

Зав. кафедри ПМ _____

(підпис)

“ 06 ” листопада 2023 р.

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ

студентові Зошуку Миколі Миколайовичу

(прізвище, ім'я, по батькові)

1. Тема роботи Математичні моделі та методи використання інфрачервоного зображення для розпізнавання об'єктів

затверджена наказом по університету від 2 листопада 2023 р. № 1277 Ст

2. Термін подання студентом роботи до екзаменаційної комісії 10 січня 2024 р.

3. Вихідні дані до роботи математичні моделі та методи інфрачервоного зображення

4. Перелік питань, що потрібно опрацювати в роботі _____

1. Системний аналіз предметної області

2. Вибір і обґрунтування методу розв'язання

3. Програмна реалізація

4. Результати обчислювального експерименту

5. Аналіз можливих застосувань

5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп'ютерних ілюстрацій _____

1. Актуальність теми роботи _____

2. Постановка задачі _____

3. Системний аналіз предметної області _____

4. Метод чисельного аналізу _____

5. Результати обчислювального експерименту _____

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Терміни виконання етапів роботи	Примітка
1	Підбір та вивчення технічної літератури за темою роботи	6 – 12 листопада 2023 р.	виконано
2	Вибір та обґрунтування методу	13 – 26 листопада 2023 р.	виконано
3	Розробка алгоритму і програми	27 листопада – 10 грудня 2023 р.	виконано
4	Проведення аналітичних досліджень та розрахунків	11 грудня – 24 грудня 2023 р.	виконано
5	Робота над текстом пояснювальної записки	25 грудня 2023 р. – 9 січня 2024 р.	виконано
6	Представлення роботи на рецензію в ЕК	10 січня 2024 р.	виконано

Дата видачі завдання 6 листопада 2023 р.

Студент _____
(підпис)

Керівник роботи _____ доц. Єсілевський В.С
(підпис) (посада, прізвище, ініціали)

РЕФЕРАТ

Пояснювальна записка: 52 с., 1 табл., 19 рис., 1 дод., 15 джерел.

МАТЕМАТИЧНІ МОДЕЛІ, ІНФРАЧЕРВОНЕ ЗОБРАЖЕННЯ, МЕТОДИ ІНФРАЧЕРВОНОГО ЗОБРАЖЕННЯ, РОЗПІЗНАВАННЯ ОБ'ЄКТІВ.

Об'єкт дослідження – математичні моделі, об'єкти для розпізнавання за інфрачервоним зображенням.

Мета роботи – підвищення ефективності (за критеріями точності та швидкодії) розпізнавання об'єктів на інфрачервоних зображеннях та відеопослідовностях шляхом подальшого розвитку методу розпізнавання об'єктів та рішень його реалізації. Для досягнення поставленої мети необхідно виконати наступні завдання: проаналізувати проблеми, які виникають при виявленні об'єктів в оптикоелектронних пристроях; дослідити відомі запатентовані технічні рішення, які здійснюють виявлення об'єктів в інфрачервоному діапазоні; навести огляд сучасних методів виявлення об'єктів при роботі в реальному часі; удосконалити обраний метод виявлення та проведено оцінку отриманих результатів.

Методи дослідження – аналіз математичних моделей та програмні бібліотеки для роботи з інфрачервоним зображенням.

У даній роботі було розглянуто математичні моделі та розпізнавання математичних моделей на інфрачервоному зображенні, заснований на застосуванні технологією формування інфрачервоних зображень, що генерує сцену на основі різниці інфрачервоного випромінювання між об'єктом і фоном. Інфрачервоне зображення має низку переваг перед камерами та лідарами, оскільки вони отримують інформацію у видимому діапазоні спектра.

ABSTRACT

Introductory note: 52 pages, 1 tables, 19 figures, 1 appendixes, 15 sources.

MATHEMATICAL MODELS, INFRARED IMAGE, INFRARED IMAGING METHODS, OBJECT RECOGNITION.

Object of research – mathematical models, objects for infrared image recognition.

Purpose of work – to increase the efficiency (in terms of accuracy and speed) of object recognition in infrared images and video sequences by further developing the object recognition method and solutions for its implementation. To achieve this goal, it is necessary to perform the following tasks: analyze the problems that arise when detecting objects in optoelectronic devices; investigate known patented technical solutions that detect objects in the infrared range; provide an overview of modern methods of object detection in real time; improve the selected detection method and evaluate the results obtained.

Methods of research – analysis of mathematical models and software libraries for working with infrared images.

In this paper, we have considered mathematical models and the recognition of mathematical models in an infrared image based on the use of infrared imaging technology, which generates a scene based on the difference in infrared radiation between an object and the background. Infrared imaging has a number of advantages over cameras and lidars, as they receive information in the visible range of the spectrum.

ЗМІСТ

	С.
Перелік скорочень, умовних познач, одиниць і термінів	7
Вступ	8
1 Системний аналіз предметної області та постановка задач дослідження	10
1.1 Системний аналіз задачі моделювання інфрачервоного зображення	10
1.2 Аналіз сценаріїв вирішення задачі комп'ютерного моделювання	12
1.3 Змістовна та формальна постановка задачі	13
1.4 Постановка задач дослідження	14
2 Вибір та обґрунтування методу розв'язання	16
2.1 Аналіз сучасних методів згорткових мереж розпізнавання об'єктів	16
2.2 Математичні методи обробки інфрачервоного зображення.....	17
2.3 Експериментальне дослідження методів комплексування зображень з відеокамери та інфрачервоної камери в реальному часі	19
Висновки за розділом 2	29
3 Програмна реалізація	30
3.1 Python для вирішення задач лінійного програмування.....	30
3.2 Алгоритм розв'язання задачі математичного моделювання інфрачервоного зображення для розпізнавання об'єктів.....	30
3.3 Опис програми	38
Висновки за розділом 3	38
4 Результати обчислювального експерименту та їх аналіз	40
4.1 Структура обчислення системи розпізнавання класів повітряних об'єктів на інфрачервоному зображенні.....	40
4.2 Навчання нейромережевого класифікатора	45
Висновки за розділом 4	46
Висновки	48
Перелік джерел посилання	49
Додаток А Лістинг програми	51

ПЕРЕЛІК СКОРОЧЕНЬ, УМОВНИХ ПОЗНАК, ОДИНИЦЬ І ТЕРМІНІВ

ІЗ – інфрачервоне зображення;

ОС – оптична система;

ПО – повітряні об'єкти;

БПЛА – безпілотні літальні апарати;

ЛА – літальні апарати;

FD – дескриптор Фур'є.

ВСТУП

Актуальність теми. Актуальність роботи зумовлена технологією формування інфрачервоних зображень, що генерує сцену на основі різниці інфрачервоного випромінювання між об'єктом і фоном. Інфрачервоне зображення має низку переваг перед камерами та лідарами, оскільки вони отримують інформацію у видимому діапазоні спектра й гірше справляються з поганими погодними умовами та темними часами доби. Датчики захоплення зображень, що фіксують навколишню ситуацію систем безпеки, транспортних засобів або безпілотних літальних апаратів мають безпомилково розуміти оточення. Однак, виявлення об'єктів на інфрачервоних зображеннях є більш складною проблемою у порівнянні зі звичайними камерами видимого діапазону.

Об'єктом дослідження є процеси цифрової обробки зображень з використанням глибоких згортальних мереж для розпізнавання об'єктів.

Предметом дослідження є методи розпізнавання об'єктів на інфрачервоних зображеннях та відеопослідовностях і алгоритмічні рішення їх реалізації.

На сьогодні сучасні методи виявлення об'єктів на основі згорткових мереж забезпечують високу точність і продуктивність розпізнавання на високоякісних RGB зображеннях. В даному випадку, під ефективністю мається на увазі швидкість обробки зображень у реальному часі (вище 30 кадрів/с), особливо на вбудованих пристроях (наприклад, в автономних транспортних засобах, інтерфейсах доповненої реальності, оптикоелектронних системах відеоспостереження й безпеки тощо) та значення точності виявлення класів mAP вище за 85%. Ще одна складність, із якою можна зіткнутися під час навчання нейронної мережі – низький обсяг даних у порівнянні з поширеними зображеннями в RGB домені. Сукупність цих факторів визначає актуальність вирішуваного завдання.

На сьогоднішній день сучасні методи виявлення об'єктів на основі

згорткових нейронних мереж забезпечують високу точність і продуктивність розпізнавання на високоякісних RGB зображеннях.

Мета і завдання кваліфікаційної роботи. Метою кваліфікаційної роботи є підвищення ефективності (за критеріями точності та швидкодії) розпізнавання об'єктів на інфрачервоних зображеннях та відеопослідовностях шляхом подальшого розвитку методу розпізнавання об'єктів та рішень його реалізації. Для досягнення поставленої мети необхідно виконати наступні завдання:

- проаналізувати проблеми, які виникають при виявленні об'єктів в оптикоелектронних пристроях;
- дослідити відомі запатентовані технічні рішення, які здійснюють виявлення об'єктів в інфрачервоному діапазоні;
- навести огляд сучасних методів виявлення об'єктів при роботі в реальному часі.

Методи дослідження. У кваліфікаційній роботі застосовуються методи використання інфрачервоного зображення для розпізнавання об'єктів за допомогою комп'ютерного моделювання.

У цьому важливими є дослідження застосування таких нейронних мереж для конкретних завдань класифікації. Незважаючи на часто використовувані в таких завданнях глибоких згорткових мереж дуже затребуваними є методи, що вимагають менше обчислювальних ресурсів як на стадії навчання нейронних мереж, так і на стадії застосування. Отримані із застосуванням цих методів розв'язання завдання класифікації або використовують нейронні мережі, або поділяють на стадії навчання нейронної мережі контури об'єктів на класи, відповідні трьом різним проєкціям.

Результати, отримані під час проходження професійної практики, були представлені на II Міжнародній молодіжній науково-практичній конференції англійською мовою «Навчання і викладання: після війни та під час миру» (м. Харків, 10 листопада 2023 р.) [1].

1 СИСТЕМНИЙ АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ ТА ПОСТАНОВКА ЗАДАЧ ДОСЛІДЖЕННЯ

1.1 Системний аналіз задачі комп'ютерного моделювання інфрачервоного зображення

На сучасному етапі розвитку інформаційних технологій приділяється велика увага системам, які використовують машинний зір як основне джерело інформації. Це викликано необхідністю автоматизації та роботизації процесів у промисловості, науці, а також у побуті. Здебільшого це процеси пов'язані з монотонною роботою, або небезпечні для людини, або ж потребують швидкого прийняття рішень. Прикладами застосування таких систем можуть служити: біометрична ідентифікація, відео-спостереження, автономні транспортні системи, індексація зображень та відео у базах даних з урахуванням їх вмісту тощо. Також масове поширення мобільних телефонів, оснащених фото- та відеокамерами, доступність комп'ютерів зі значними обчислювальними потужностями призвело до стрімкого розвитку комп'ютерного зору. Комп'ютерний зір – теорія та технологія створення програм та машин, які можуть проводити виявлення, стеження та класифікацію об'єктів. Основним змістом комп'ютерного зору є добування інформації із зображень або їх послідовності. Тому тема роботи, що направлена на виявлення та розпізнавання об'єктів є актуальною. Варто зазначити, що розвиток комп'ютерного зору прискорився завдяки удосконаленню алгоритмів машинного зору, в тому числі і впровадженню штучних нейронних мереж. Значний внесок у вдосконалення таких алгоритмів здійснив проект OpenCV, запущений з ініціативи компанії Intel у 1999 році. Зараз проект OpenCV підтримується некомерційною організацією OpenCV.org. Результатом проекту стала поява однойменної бібліотеки. OpenCV (Open Source Computer Vision Library) – це бібліотека функцій та алгоритмів обробки зображень та комп'ютерного зору з відкритим кодом. Бібліотека містить понад 2500 оптимізованих алгоритмів, що включає в

себе вичерпний набір як класичних, так і сучасних алгоритмів комп'ютерного зору та машинного навчання. Може вільно використовуватися в академічних та комерційних цілях. Бібліотеку у своїх продуктах використовують такі відомі компанії, як Google, Yahoo, Microsoft, Intel, IBM, Sony, Honda, Toyota, а також безліч стартапів. Вибір стратегії залежить від умов отримання зображення. Моделювання фону застосовне тільки для “ідеальних” умов зйомки. Моделювання об'єкта – більш загальний підхід. До моделювання об'єкта для пошуку об'єкта на зображенні можна застосовувати різні підходи. Не можна стверджувати, що якийсь із методів ефективніший за інший. Вибір конкретного метода залежить від багатьох умов. Найпростішими методами виділення об'єкта на зображенні є колірні фільтри. Такі методи застосовуються, якщо об'єкт суттєво виділяється на фоні. Виділення країв та контурний аналіз будуть корисними у випадках, якщо об'єкт досить складний, але добре виділяється. Це дає змогу перейти від роботи з зображенням до роботи з об'єктами на цьому зображенні. Далі можна перевірити наявність на зображенні певних геометричних форм. Метод співставлення зі шаблоном (template matching) полягає у пошуку на зображенні ділянок, які співпадають з зображенням шуканого об'єкта. Якщо зображення об'єкта повернуте чи масштабоване відносно шаблону, то цей метод неефективний. Для таких випадків краще підійдуть методи засновані на так званих особливих точках. Особливі точки – це особливі характеристики об'єкта. Вони дозволяють співставити об'єкт сам з собою або зі схожими класами об'єктів. Існує кілька способів виділяти особливі точки. Деякі способи виділяють особливі точки на сусідніх кадрах, деякі – через великі проміжки часу та при різному освітленні, деякі дозволяють знайти особливі точки, навіть при повороті зображення. Найскладнішими випадками розпізнавання є пошук об'єктів певного класу. В таких випадках задачу виявлення і розпізнавання можна вирішити за допомогою побудови класифікатора на основі машинного навчання, який складається з метода виділення особливостей (feature extractor) та власне класифікатора. Методи виділення особливостей залежать від поставленої задачі. Для одного класу

задач це може бути навчання на позитивних і негативних наборах зображень, для інших – це виділення кластерів дескрипторів особливих точок і створення, так би мовити, словника дескрипторів. Не зважаючи на велику кількість відомих методів розпізнавання об'єктів не існує універсального набору для всіх умов розпізнавання. Задача розпізнавання може бути вирішена лише за конкретних умов.

1.2 Аналіз сценаріїв вирішення задачі комп'ютерного моделювання інфрачервоного зображення

Оптико-електронні системи (ОЕС), які використовуються для пошуку та виявлення наземних об'єктів і повітряних носіїв стають усе більш поширеними з кожним роком. Водночас, сучасні алгоритми й методи комп'ютерного зору привнесли величезне значення для систем, які можуть обробляти зображення та відеодані так само, як люди або навіть краще. Якщо об'єднати два цих ствердження, то отримаємо високоефективні пошуково-оглядові комплекси, які можуть допомогти реєструвати відеодані в автономних транспортних засобах та системах для допомоги водію, а також у системах безпеки і спостереження. Однак, наразі в основному використовуються камери видимого діапазону спектра, які мають певні проблеми при виявленні об'єктів у темний час доби і при поганих погодних умовах. І разом із цим, камери, які працюють в інфрачервоному діапазоні можуть істотно підвищити показник виявлення у порівнянні з камерами видимого діапазону в подібних умовах. Це досягається завдяки тому, що тепловізори сприймають випромінювання з довжинами хвиль у діапазоні від 1,4 мкм до 14 мкм, які не сприймає людина. Відповідно, ці зображення відрізняються від того, що ми бачимо. Оскільки видиме світло не потрібно для побудови сцени, можна з легкістю спостерігати за об'єктами у повній темряві, а також у різних погодних умовах і навіть у тумані. Залежно від технології формування зображення є різні системи, які мають свої переваги й

недоліки. Як показано на рис. 1.1, інфрачервоне випромінювання поділяється на NIR, SWIR, MWIR, LWIR та VLWIR, що мають довжини хвиль від ближнього діапазону спектра до понаддовгохвильового відповідно. В інфрачервоних ОЕС використовується світло короткохвильового діапазону (1,4 – 3) мкм, середнього діапазону (3 – 5) мкм та довгохвильового діапазону (8 – 14) мкм.

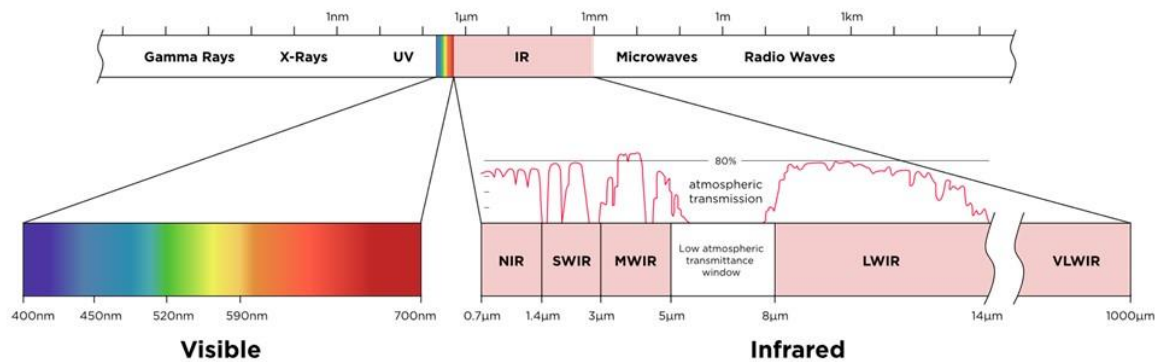


Рисунок 1.1 – Інфрачервоне випромінювання

1.3 Формальна та змістовна постановка задачі

Метою роботи є підвищення ефективності (за критеріями точності та швидкодії) розпізнавання об'єктів на інфрачервоних зображеннях та відеопослідовностях шляхом подальшого розвитку методу розпізнавання об'єктів на основі згорткових нейромереж та алгоритмічних і програмних рішень його реалізації.

Завдання класифікації, яке вирішується в цій роботі, є складовою апаратно-програмного комплексу виявлення рухомих повітряних об'єктів і визначення їх характеристик.

Датчики захоплення зображень, що фіксують навколишню ситуацію систем безпеки, транспортних засобів або безпілотних літальних апаратів мають безпомилково розуміти оточення.

1.4 Постановка задач дослідження

Всі тепловізори можна розділити на охолоджувальні й неохолоджувальні. Тепловізійні камери з охолоджувальними сенсором володіють деякими перевагами в порівнянні з неохолоджувальним, однак вони дорожче. Неоохолоджувальні системи мають термосенсор, який називається мікроболометр. Він реагує на зміни опору при нагріванні або охолодженні. Тобто, коли випромінювання потрапляє на цей сенсор, він нагрівається і відповідно його опір змінюється. Потім зміна опору перетворюється в цифровий вигляд, калібрується за температурою і залежно від температури присвоюється значення кольору для кольорового уявлення або присвоюється значення різних рівнів градацій сірого для чорно-білого залежно від камери. Переваги таких тепловізорів це – низьке енергоспоживання, низька вартість, помірна чутливість і відсутність механічного охолодження. А до недоліків можна віднести низьку роздільну здатність та повільну частоту кадрів. Охолоджувальні камери працюють іншим чином. Вони збирають фотони інфрачервоної енергії, які проходять через оптику. Ці фотони перетворюються в електрони і зберігаються в конденсаторі. Після певного періоду часу заряд зчитується в цифровий відлік і також калібрується по температурі і, відповідно, присвоюється значення кольору або шкали сірого. Завдяки такому підходу маємо високу частоту кадрів, гарну чутливість, здатність бачити і вимірювати дуже маленькі об'єкти і кращу роздільну здатність.

При виконанні завдання розпізнавання та ідентифікації об'єктів, важливу роль відіграє роздільна здатність зображення. Існує безліч досліджень на тему того, як може впливати роздільна здатність на якість алгоритмів виявлення. Якщо проаналізувати дані, які показують вплив різних чинників на необхідну роздільну здатність для успішного детектування, то можна виділити такі фактори:

а) фактори навколишнього середовища (опаді, стан повітря, вологість, температура);

б) візуальні перешкоди, які не потребують розпізнання (нерівномірний фон, або фон, що складається з геометрично схожих предметів);

в) геометричні особливості об'єкту (розміри та розташування в просторі).

Деякі фактори, які вносять додаткові труднощі візуального сприйняття, представлені в табл. 1.1.

Таблиця 1.1 – Фактори, що впливають на якість розпізнавання об'єктів

Фактор	Похибка виявлення, %	Коментар
Характеристика сцени	Від 50 до 250	Нерівномірний фон, щільність
Характеристика об'єкту	До 200	Розміри, розташування в просторі
Погодні умови	До 100	Опади, стан повітря, вологість, температура
Характеристика приладу	До 30	Роздільна здатність, різкість зображення, якість передачі.
Дисплей і спостерігач	До 100	Навколишні умови і кваліфікація спостерігача

Як видно з вищенаведеної таблиці, людський фактор дуже сильно може вплинути на розпізнавання і тому, на поточний момент розробляється велика кількість ОЕС в поєднанні з автоматизованими методами, які побудовані на нейромережевих технологіях.

2 ВИБІР ТА ОБҐРУНТУВАННЯ МЕТОДУ РОЗВ'ЯЗАННЯ

2.1 Аналіз сучасних методів згорткових нейронних мереж розпізнавання об'єктів

Виявлення об'єктів – це технологія, яка пов'язана з комп'ютерним зором і основне завдання якої, виявляти екземпляри семантичних об'єктів певного класу (наприклад людей, транспортні засоби, тварин) на відеопослідовностях або статичних зображеннях. Завдяки цій технології зараз можливе безпілотне водіння, відеоспостереження без оператора, виявлення обличчя та інші важливі галузі. Завдяки недавнім досягненням в XXI столітті з'явилося багато інноваційних та креативних методологій, які можуть бути, або з високими характеристиками, але ресурсоємні для використання в додатках реального часу, або навпаки менш точними, але давати нагоду користувачам використовувати виявлення об'єктів у компактних модульних структурах і навіть на мобільних пристроях.

У даній роботі не будуть розглянуті моделі з архітектурою Regions With Convolution Neural Networks (R-CNNs) та їхні похідні, які чудово справляються з завданням розпізнавання в автономному режимі, але абсолютно не підходять для реалізації в реальному часі. Такі алгоритми використовують двоетапний підхід – спочатку визначають області, де об'єкти повинні бути знайдені, а потім виявляють об'єкти тільки в цих, окремих регіонах, використовуючи згорткові нейронні мережі. З іншого боку, є алгоритми, які використовують повністю згортковий підхід, наприклад – Single Shot Detector (SSD), при якому мережа може знаходити всі об'єкти на зображенні за один прохід. Алгоритми з регіонами мають кращу точність, але повільніше в роботі, у той час, як алгоритми «одиночного пострілу» більш ефективні й мають таку ж високу точність.

2.2 Математичні методи обробки інфрачервоного зображення

Цифрова обробка зображень (з англ. Digital Image Processing або DIP) – це створення алгоритмів для обробки цифрових зображень використовуючи комп'ютерні технології. Вхідними даними для DIP є цифрові зображення, наприклад фото або відеокадри. Вихідними даними, після обробки, є такі ж цифрові зображення які мають цікаву для спостереження інформацію, в залежності від алгоритму. DIP використовують для підготовки даних до розпізнавання об'єктів, для виділення необхідної інформації або для підготовки даних до використання різних моделей, які працюють з даними зображеннями. Системи розпізнавання обличь все частіше і частіше використовуються в різних сферах. Такі системи використовуються для надання дозволу на проходження в місця з обмеженим доступом, або для систем спостереження, щоб визначити ціль та вилучити якомога більше інформації про людину. Зазвичай задачу по створенню системи розпізнавання обличь називають міжмодельним гетерогенним розпізнаванням обличь (з англ. Cross-model heterogeneous face recognition або HFR). HFR є однією з найскладніших областей досліджень біометрії та комп'ютерного зору.

Multi-scale transform методи, засновані на перетвореннях, є найбільш активно використовувани у злитті зображень і передбачають розбиття зображення на різні шари в різних зернах. Ці методи розкладають вихідні зображення на кілька рівнів, зливають відповідні шари з певними правилами і відповідно реконструюють цільові зображення. Популярні перетворення, що використовуються для розкладання та реконструкції, включають wavelet, pyramid, curvelet та їхні перероблені версії. Sparse representation методи працюють на основі можливого представлення зображень з лінійними комбінаціями розріджених базисом у надповних словниках, які є ключовими для їх гарної роботи. Neural network методи імітують поведінку людського мозку у сприйнятті нейронної інформації; Ця процедура має переваги хорошої пристосовуваності, відмовостійкості та шуму. Subspace методи, що базуються

на підпросторі, використовують подібні ідеї, але реалізують їх у просторах з повними базами, а не надзаповненими. Основними методами цієї категорії є аналіз основних компонентів Principal component analysis (PCA), неотрицательна матрична факторизація non-negative matrix factorization (NMF) та незалежний аналіз компонентів independent component analysis (ICA). Методи, засновані на виразності, засновані на тому, що візуальна увага людини часто захоплюється об'єктами або пікселями, які є більш значними, ніж їхні сусіди. Крім того, ці способи можуть зберігати цілісність основних областей об'єктів і покращувати візуальну якість поєднаних зображень. Вищезгадані способи об'єднання інфрачервоного і видимого зображення мають свої переваги і недоліки. Тим часом hybrid model поєднують переваги цих схем для поліпшення продуктивності злиття зображень. Інші методи базуються на повній total variation, fuzzy theory, entropy. Ці методи можуть призвести до нового бачення об'єднання інфрачервоного та видимого злиття зображень, що може стати товчком до проектування нових моделей.

Недоліки існуючих методів. Дослідження алгоритмів синтезу в теорії та застосуванні стрімко розвиваються в останні роки. Однак певні проблеми потребують уваги. Вибір методу вилучення інформації є ключовим фактором для забезпечення хорошої продуктивності синтезу. Деякі сучасні підходи, особливо багатовимірні перетворювальні методи, залежать від попередньо визначених перетворень і відповідних рівнів для декомпозиції і реконструкції. Однак для оцінки цих перетворень не використовуються жодні критерії і рівні. Отже, перетворення вибираються сліпо, тим самим погіршуючи продуктивність. Більш того, поточні правила злиття в основному надмірно прості, тому такі прості дефекти зображення як гало, вводяться в результати. Методи в інших категоріях також страждають від декількох проблем. Наприклад, побудова відповідних словників, які мають гарне представлення цільових даних з використанням розріджених методів, що базуються на репрезентації, досить важка задача. Проектування відповідних нейронних мереж та коригування відповідних параметрів залишаються складними. Крім

того, застосування сучасних методів глибокого навчання до інфрачервоного та видимого злиття зображень ще не вирішено. Знайти підпростір з потужною виразною здатністю в методах, що базуються на підпросторі, важко. Між тим, розробка методів націлювання на нішу, методів виявлення протишумової чутливості у відповідності з візуальною системою людини та поєднання переваг різних методів створення розширених моделей гібридного злиття є складним завданням та є відкритою проблемою.

Інфрачервоне і видиме об'єднання зображень має величезний функціональний потенціал, але сучасні методи не можуть повністю забезпечити отримання інформаційно точного об'єднаного зображення.

2.3 Експериментальне дослідження методів комплексування зображень з відеокамери та інфрачервоної камери в реальному часі

Проведені експериментальні дослідження по комплексуванню різноспектральних зображень, тобто зображень з відеокамери та інфрачервоної камери, з використанням програмного забезпечення, що було розроблено.

Первинні ображення з відеокамери «Камера №1» та інфрачервоної (тепловізійної) камери «Камера №2» показані на рис. 2.1 (а) та рис. 2.1 (б) відповідно. Камери були розташовані на спільній платформі одна поруч з іншою та спрямовувались на макет літака. Хоча камери фіксували об'єкт зі спільного напрямку та були розташовані близько одна до одної, в просторі вони були все ж таки рознесені і тому їх оптичні осі не співпадали. До того ж відрізнялась їх роздільна здатність та оптичні системи, в результаті чого не співпадали їх масштабування. Все це призводило до того, що перспектива та масштаб зображень з камер відрізнялись. В результаті при накладенні двох зображень, контури об'єкта (макета літака), що реєструвався камерами, не співпадали, що відображено (рис. 2.1 (в)). Ці розбіжності були усунені за допомогою геометричного перетворення методом гомографії зображення з

відеокамери «Камера №1» і його приведенні до зображення з інфрачервоної камери «Камера №2». Геометричне перетворення виконувалось за основі завдання чотирьох точок на зображенні макета літака з відеокамери та завдання тих же точок на зображенні макета літака з інфрачервоної камери. Тобто було реалізовано геометричне комплексування. Результат геометричного перетворення зображення з «Камери №1» показаний на рис. 2.1 (г).

Після цього були реалізовані комплексування по яркості пікселів зображень різними методами, що були описані вище.

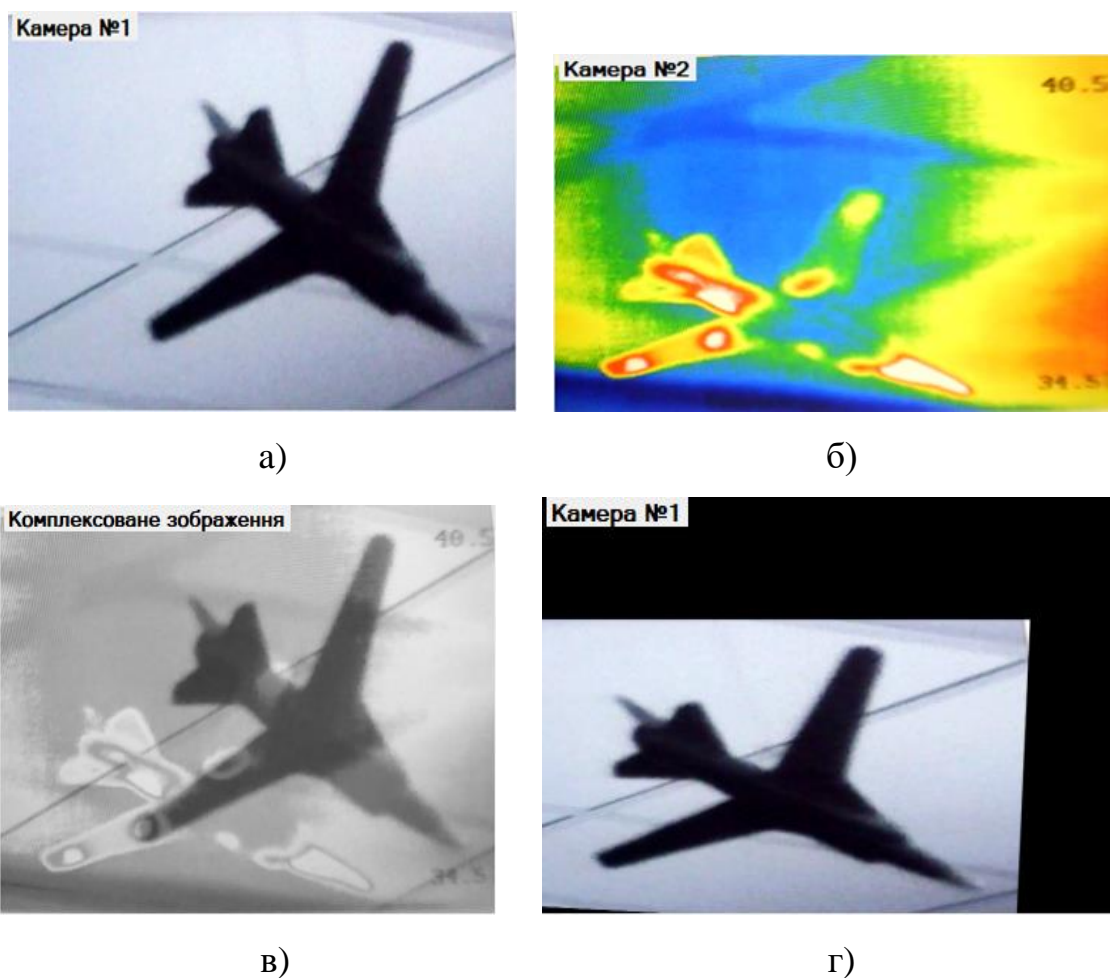
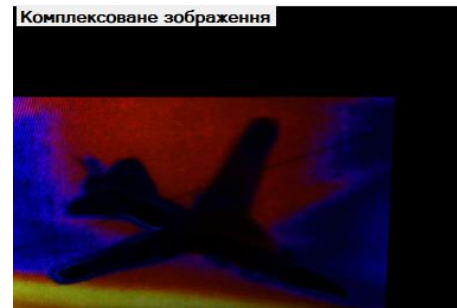


Рисунок 2.1 – Геометричне перетворення зображень: а) первинне зображення з відеокамери «Камера №1»; б) первинне зображення з інфрачервоної камери «Камера №2»; в) комплексоване зображення без геометричного перетворення; г) зображення з камери №1 після геометричного перетворення

Результати комплексуювання різноспектральних зображень з відеокамери та інфрачервоної камери представлені на рис. 2.2.



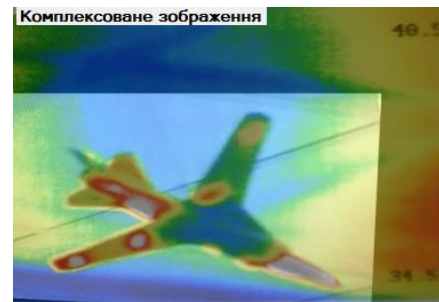
а)



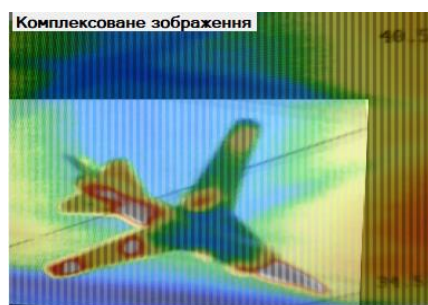
б)



в)



г)



д)

Рисунок 2.2 – Результати комплексуювання різноспектральних зображень

На рис. 2.2 (а) показаний результат комплексуювання методом вагової функції, на рис. 2.2 (б) – методом степеневого перетворення, на рис. 2.2 (в) – методом максимуму, на рис. 2.2 (г) – методом усереднення, на рис. 2.2 (в) – методом черезрядкового комплексуювання.

Аналіз зображень показує, що комплестовані зображення об'єднують інформативні компоненти зображень з двох камер. Це робить результуюче комплексоване зображення більш інформативним чим зображення з окремих камер, що дає змогу більш якісно робити захват на супроводження цілей.

Однак, методи глибокого навчання вимагають великого обсягу даних і великих витрат комп'ютерного часу на навчання мережі [5, 6]. Ці підходи застосовуються і в області розпізнавання повітряних об'єктів [7], хоча в цій предметній області виникає специфічна проблема, пов'язана з різницею плоских зображень тривимірних тіл, які мають 3 ступені свободи в просторовій орієнтації. При цьому тяжкість навчання глибоких мереж з нуля може бути частково компенсована використанням трансферного навчання [8], при якому попередньо навчена універсальна мережа доучується на специфічних зображеннях повітряних об'єктів, отриманих при різних кутах просторового положення.

Застосування трансферного навчання дозволяє обійти обмеження, пов'язані з часом навчання. Проте, залишаються відкритими деякі питання. По-перше, це питання про надмірності універсального підходу на основі глибоких нейронних мереж в разі, коли мова йде про класифікацію повітряних об'єктів. По-друге, питання про віддільність класів об'єктів, кожен з яких представлений дуже різними (за рахунок просторової орієнтації) образами об'єктів на плоскому зображенні.

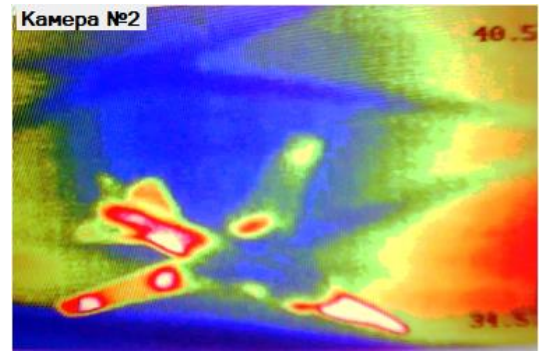
Систематизація результатів наведених досліджень дозволяє зробити висновок, що існуючі підходи до вирішення проблеми класифікації повітряних об'єктів спираються або на великовагові методи глибоких згортальних мереж.

Рішення проблеми класифікації плоских зображень повітряних об'єктів, що знаходяться в довільному просторовому положенні, на основі векторів ознак контуру, може бути спрощене шляхом використання комплекснозначних нейронних мереж і використанням специфічних методик їх навчання.

На рис. 2.3 представлені результати комплексування зображень з камер, що показані на рис. 2.1 з покращенням контрастності первинних зображень.



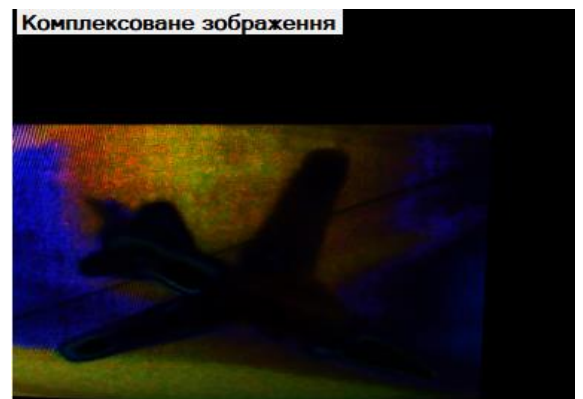
а)



б)



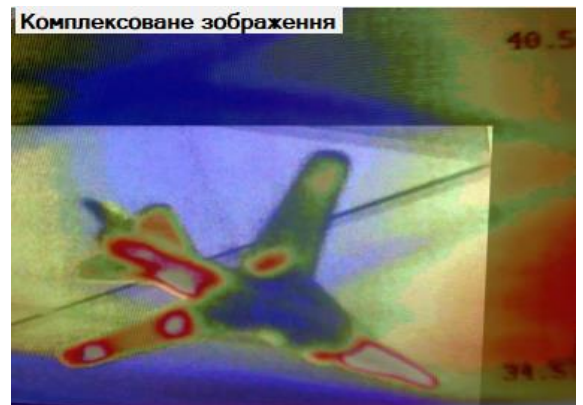
в)



г)



д)



е)

Рисунок 2.3 – Результати комплексування різноспектральних зображень

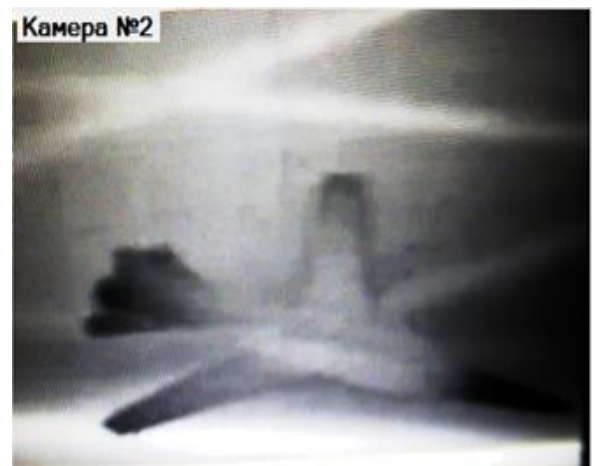
На рис. 2.3 (а) та рис. 2.3 (б) показані первинні зображення з попереднім покращенням контрастності, на рис. 2.3 (в) показаний результат комплексування методом вагової функції, на рис. 2.3 (г) – методом степеневого

перетворення, на рис. 2.3 (д) – методом максимуму, на рис. 2.3 (е) –методом усереднення.

На рис. 2.4 – рис. 2.6 показаний результат комплексування коли частина зображення однієї з камер перекрита. На рис. 2.4 показаний результат комплексування коли літак на обох зображеннях відкритий, тобто повністю видимий. Далі в експерименті послідовно була закрита частина об'єктива відеокамери («Камера №1») (рис. 2.5) та інфрачервоної камери («Камера №2») (рис. 2.6), що емітувало наявність, наприклад, хмар, що закривають частину літака.



а)



б)



в)

Рисунок 2.4 – Результат комплексування з відкритими зображеннями літака

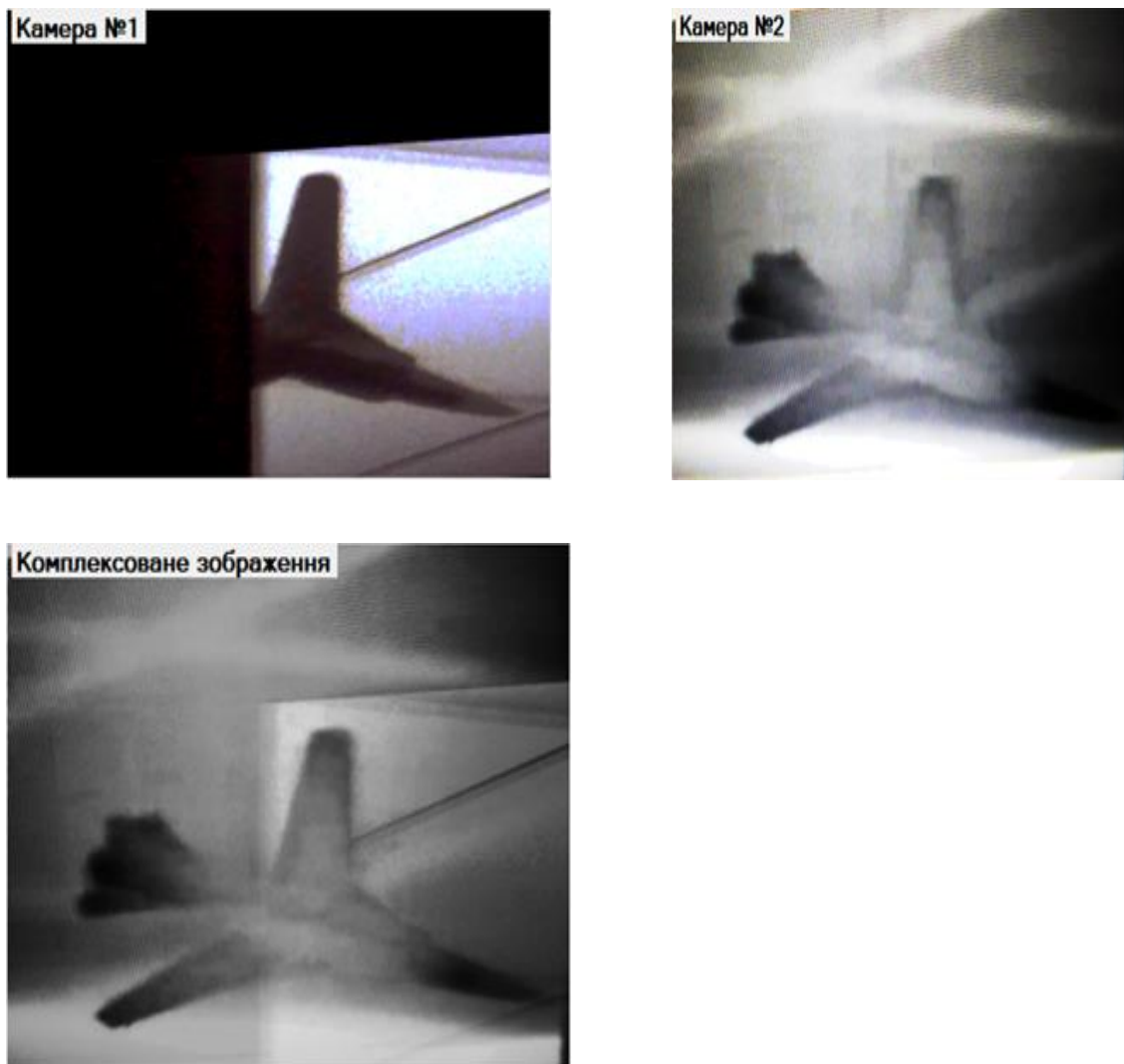


Рисунок 2.5 – Результат комплексування при закритій частині об’єктива відеокамери («Камера №1»)

Швидке і якісне вирішення задач виявлення і розпізнавання повітряних об’єктів має велике значення як для задач військового призначення, так для цивільного застосування в галузі управління повітряним рухом, моніторингу повітряної обстановки в аеропортах. В даний час перелік типів розпізнавання повітряних об’єктів (ПО) значно розширився за рахунок застосування безпілотних літальних апаратів (БПЛА), квадрокоптера, крилатих ракет, вертольотів.

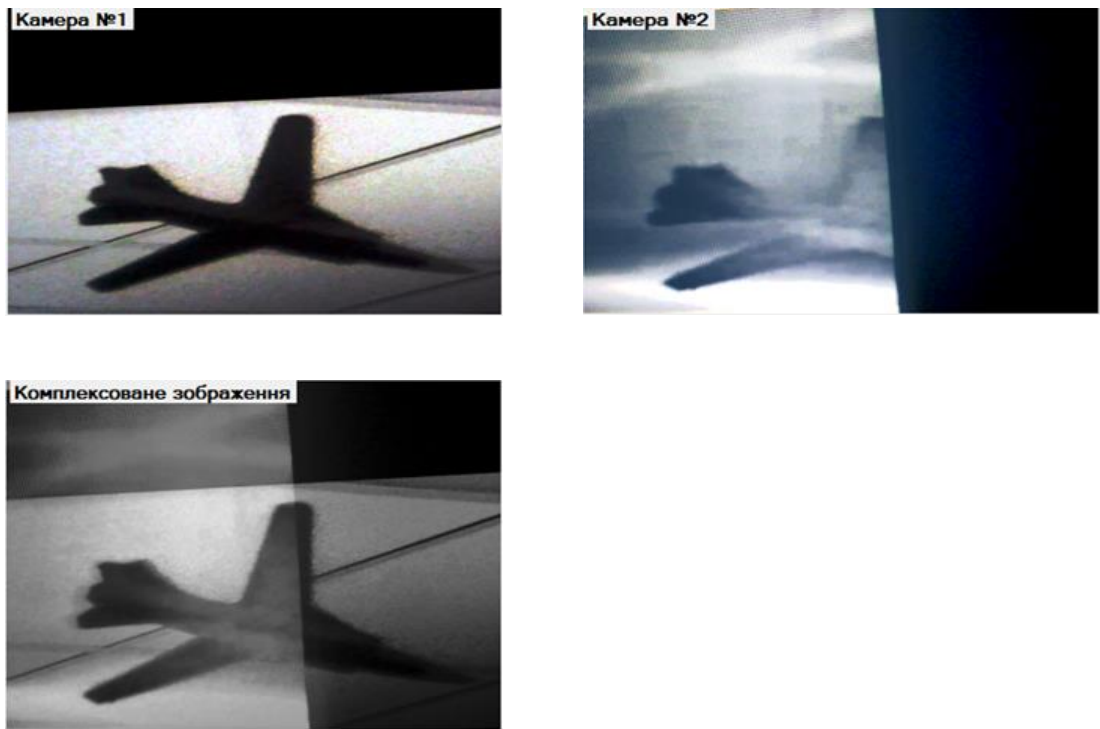


Рисунок 2.6 – Результат комплексування при закритій частині об'єктива інфрачервоної камери («Камера №2»)

Це принципово змінило діапазон виявляються параметрів ПО, починаючи від форми і розмірів, і закінчуючи динамічними характеристиками руху. Тому для виявлення повітряних об'єктів все частіше замість класичних радіолокаційних систем використовуються системи оптичного відеоспостереження. Залучення нових технологій переносить тяжкість задач розпізнавання повітряної обстановки в область обробки цифрового відеозображення і автоматичного розпізнавання об'єктів.

Для порівняння застосування зображення з інфрачервоної камери у чистому вигляді та комплексованого зображення було проведено експериментальне дослідження, представлене на рис. 2.7.

Обробка відеозображення включає в себе ряд задач, серед яких виявлення рухомого ПО, визначення типу ПО, визначення характеристик об'єкта. У запропонованому підході ми досліджуємо задачу визначення типу ПО, попередньо виявленого за ознаками руху, і локалізованого на цифровому зображенні. В області комп'ютерного зору такого роду задачі відносять до

задач класифікації зображень, і вирішують на основі підходів штучного інтелекту.

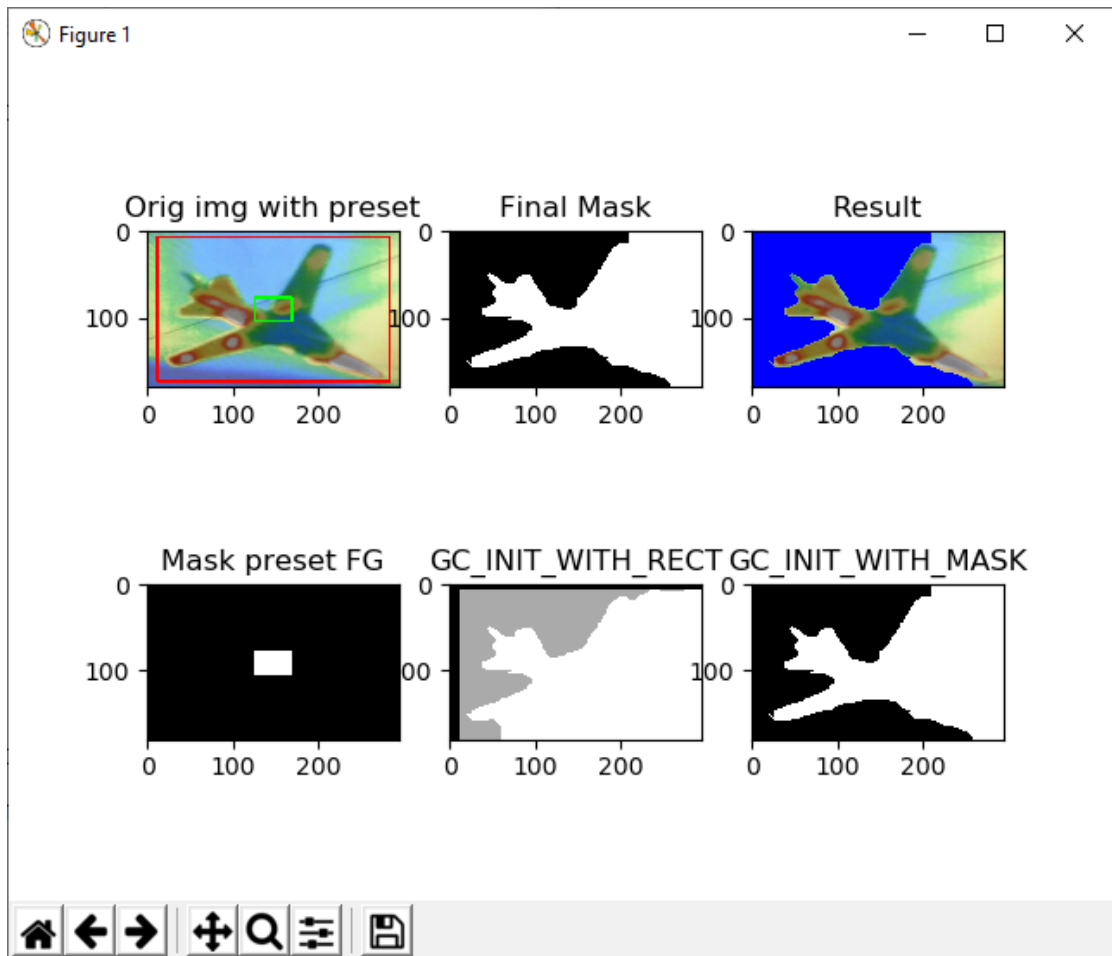


Рисунок 2.7 – Застосування зображення з інфрачервоної камери у чистому вигляді

Очевидно, що при візуальному розпізнаванні типу ПО його геометрична форма, представлена зовнішнім контуром, є мінімально достатнім ідентифікатором. При грубому розпізнаванні типу повітряних об'єктів, саме контури об'єкта є широко використовуваними для виділення певних ознак, які представляють зображення. Проблема полягає в тому, що характерні ознаки повинні бути інваріантними щодо геометричної деформації (зміщення, орієнтації, масштабу).

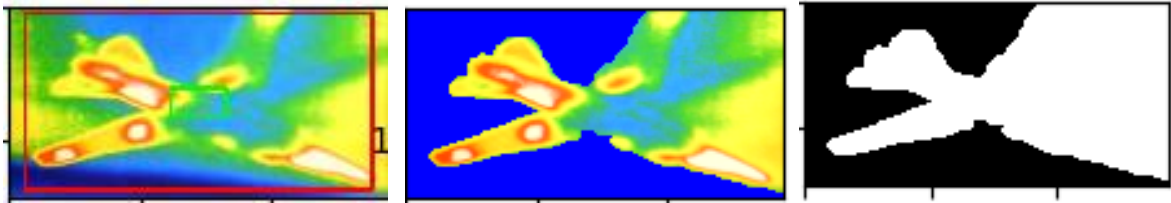


Рисунок 2.8 – Отримання бінарзованого зображення ПО, отриманого з інфрачервоної камери.

У зв'язку з цим важливими є дослідження застосовності таких нейронних мереж для конкретних задач класифікації. Незважаючи на те, що часто використовуються в таких задачах глибокі згортальні мережі, дуже затребуваними є методи, що вимагають менше обчислювальних ресурсів, як на стадії навчання нейронних мереж, так і на стадії застосування. Роботи [1 – 3], присвячені використанню методів розпізнавання об'єктів на основі саме векторів ознак, що отримуються в результаті контурного аналізу. Але, ці методи вирішення задачі класифікації або не використовують комплекснозначні нейронні мережі, або не поділяють на стадії навчання нейронної мережі контури об'єктів на класи, що відповідають трьом різним проекціям.

Задача аналізу повітряної обстановки засобами візуального спостереження являє собою ряд задач: виявлення рухомого ПО; визначення характеристик об'єкта (дальність, розмір, швидкість, маневреність і т. п.); визначення типу повітряних об'єктів. Ці задачі можуть вирішуватися як поетапно, так і одночасно. В даному дослідженні ми припускаємо, що об'єкт, який рухається, спочатку буде виявлений і локалізований на відеозображенні, а потім будуть вирішуватися задачі віднесення його до певного класу.

Висновки за розділом 2

Аналіз результатів, зображених на рис. 2.5 – рис. 2.6 показує, що при частковому перекритті об'єкта спостереження на одній з камер, результуюче комплексоване зображення об'єкта зберігає повний вид об'єкта і його контури не втрачаються.

3 ПРОГРАМНА РЕАЛІЗАЦІЯ

3.1 Python для вирішення задач лінійного програмування

Python – об'єктно-орієнтована мова програмування високого рівня із найсуворішою динамічною типізацією. Структури даних високого рівня разом з динамічною семантикою та динамічною зв'язкою роблять її привабливою для швидкої розробки програм. Python підтримує модулі та пакети модулів, що сприяє модульності та повторного використання коду. В екосистемі Python є багато бібліотек для вирішення лінійних задач програмування. Зауважимо, що майже всі широко використовувані бібліотеки лінійного програмування та змішано-цілочисельного лінійного програмування написані мовами Fortran, C або C++, адже лінійне програмування потребує інтенсивної обчислювальної роботи з матрицями, часто дуже великими. Відповідні інструменти Python – це просто зручні інтерфейси для роботи з низькорівневими бібліотеками – Солвер. У цій кваліфікаційній роботі було використано Python-бібліотеки: SciPy (універсальний пакет для наукових обчислень з Python; його внутрішній пакет `scipy.optimize` можна використовувати як для лінійної, так і нелінійної оптимізації), NumPy (бібліотека мови Python, що надає підтримку великим багатовимірним масивам та матрицям, разом із великою бібліотекою високорівневих математичних функцій для операцій із цими масивами) та Matplotlib (бібліотека двовимірної графіки для мови програмування Python, за допомогою якої можна створювати високоякісні малюнки різних форматів).

3.2 Алгоритм розв'язання задачі математичного моделювання інфрачервоного зображення для розпізнавання об'єктів

Швидке і якісне вирішення задач виявлення і розпізнавання повітряних об'єктів має велике значення як для задач військового призначення, так для

цивільного застосування в галузі управління повітряним рухом, моніторингу повітряної обстановки в аеропортах. В даний час перелік типів розпізнавання повітряних об'єктів (ПО) значно розширився за рахунок застосування безпілотних літальних апаратів (БПЛА), квадрокоптера, крилатих ракет, вертольотів. Це принципово змінило діапазон виявляються параметрів ПО, починаючи від форми і розмірів, і закінчуючи динамічними характеристиками руху. Тому для виявлення повітряних об'єктів все частіше замість класичних радіолокаційних систем використовуються системи оптичного відеоспостереження. Залучення нових технологій переносить тяжкість задач розпізнавання повітряної обстановки в область обробки цифрового відеозображення і автоматичного розпізнавання об'єктів.

Обробка відеозображення включає в себе ряд задач, серед яких виявлення рухомого ПО, визначення типу ПО, визначення характеристик об'єкта. У запропонованому підході ми досліджуємо задачу визначення типу ПО, попередньо виявленого за ознаками руху, і локалізованого на цифровому зображенні. В області комп'ютерного зору такого роду задачі відносять до задач класифікації зображень, і вирішують на основі підходів штучного інтелекту.

Обробку відеозображення наземними засобами стеження за повітряною обстановкою можна розбити на низку завдань. Серед них можна виділити такі завдання: виявлення рухомого об'єкта, що летить; визначення характеристик об'єкта (дальність, розмір, швидкість, маневреність тощо); визначення типу ЛА. Ці завдання можуть вирішуватися як поетапно, так і одночасно. Це може бути виявлення всіх об'єктів на одному кадрі з визначенням їх типу і подальшим їх супроводом у послідовності відеокadrів або виявлення рухомих об'єктів за відеорядом із подальшою ідентифікацією об'єктів на кожному кадрі. У запропонованому підході ми припускаємо, що рухомий об'єкт або об'єкти, що рухаються, спочатку будуть виявлені та локалізовані на відеозображенні, а потім буде вирішуватися завдання віднесення кожного об'єкта до певного класу.

У галузі розпізнавання образів на зображенні прийнято таку стійку термінологію для визначення завдань [1], представлених на рис. 3.1.



а)

б)

в)

Рисунок 3.1 – Різні завдання розпізнавання образів на зображенні:

а) класифікація, б) класифікація з локалізацією, в) детектування

Детектування – визначення класу кожного об'єкта на зображенні. Класифікація – визначення класу одного об'єкта на зображенні. Класифікація з локалізацією – класифікація одного об'єкта і зазначення його місця на зображенні.

Таким чином, у цій термінології ми будемо обговорювати методи розв'язання задачі класифікації ЛА на цифровому зображенні.

Класичні методи розпізнавання ЛА, що застосовувалися до появи глибоких нейронних мереж, були засновані на порівнянні невідомого зображення з бібліотекою відомих. Складність розв'язання такого завдання полягає в правильному визначенні невідомого ЛА незалежно від його орієнтації. Для подолання цих труднощів використовується підхід з обчисленням чисельних значень (моментів), інваріантних до повороту зображення на площині. До цих характеристик належать моменти Ху, моменти Зерніке і Вейвлет-моменти [2,3].

При цьому завдання класифікації розв'язували методами машинного навчання, як правило, з використанням класичних нейронних мереж прямого поширення з невеликою кількістю прихованих шарів [2].

Очевидно, що при візуальному розпізнаванні типу ПО його геометрична форма, представлена зовнішнім контуром, є мінімально достатнім ідентифікатором. При грубому розпізнаванні типу повітряних об'єктів, саме контури об'єкта є широко використовуваними для виділення певних ознак, які представляють зображення. Проблема полягає в тому, що характерні ознаки повинні бути інваріантними щодо геометричної деформації (зміщення, орієнтації, масштабу).

Цей підхід давав непогані результати для розпізнавання літаків на аерокосмічних знімках аеродромів, де варіації положення об'єкта відносилися до повороту на площині. Для розпізнавання літака, що летить, ці методи менш пристосовані. Для такого роду завдань використовуються різні методи розпізнавання тривимірних об'єктів [4,5].

Революційних успіхів у сфері аналізу зображень було досягнуто після застосування глибоких згорткових нейронних мереж, навчених на тестовому наборі зображень ImageNet [6]. Модифікації підходів, пов'язаних із глибокими нейронними мережами, були використані і для вирішення завдання ідентифікації типів літаків [7].

Під час навчання глибоких нейронних мереж розв'язання задачі класифікації з нуля виникає низка проблем. Насамперед – необхідність створення великого навчального набору розмічених зображень, що охоплює мільйони зображень як в ImageNet. І складною технічною проблемою є підключення великих обчислювальних ресурсів, що включають в себе кластери графічних відеокарт, щоб закінчити навчання в прийнятний час.

Розв'язанням описаних проблем визначення типу літального апарата є трансферне навчання глибоких нейронних мереж, що дає змогу використовувати попередньо навчену (навіть для інших класів об'єктів) мережу для побудови класифікаторів зображень ЛА.

Дослідженню цього підходу присвячено цю роботу.

У зв'язку з цим важливими є дослідження застосовності таких нейронних мереж для конкретних задач класифікації. Незважаючи на те, що часто використовуються в таких задачах глибокі згортальні мережі, дуже

затребуваними є методи, що вимагають менше обчислювальних ресурсів, як на стадії навчання нейронних мереж, так і на стадії застосування. Роботи [1 – 3], присвячені використанню методів розпізнавання об'єктів на основі саме векторів ознак, що отримуються в результаті контурного аналізу. Але, ці методи вирішення задачі класифікації або не використовують комплекснозначні нейронні мережі, або не поділяють на стадії навчання нейронної мережі контури об'єктів на класи, що відповідають трьом різним проекціям.

Це підтверджує, що тематика запропонованих досліджень, присвячених вдосконаленому методу розпізнавання повітряних об'єктів, що знаходяться в довільному просторовому положенні, на основі векторів ознак контуру з використанням комплекснозначних нейронних мереж є актуальною.

Значні досягнення в області застосування нейронних мереж для розпізнавання довільних зображень в даний час засновані на використанні глибоких згортальних нейронних мереж [4].

Однак, методи глибокого навчання вимагають великого обсягу даних і великих витрат комп'ютерного часу на навчання мережі [5, 6]. Ці підходи застосовуються і в області розпізнавання повітряних об'єктів [7], хоча в цій предметній області виникає специфічна проблема, пов'язана з різницею плоских зображень тривимірних тіл, які мають 3 ступені свободи в просторовій орієнтації. При цьому тяжкість навчання глибоких мереж з нуля може бути частково компенсована використанням трансферного навчання [8], при якому попередньо навчена універсальна мережа доучується на специфічних зображеннях повітряних об'єктів, отриманих при різних кутах просторового положення.

Пояснення процесу розпізнавання в теорії глибоких нейронних мереж засновано на виділенні ієрархії характерних ознак об'єктів на зображенні. У перших шарах мережі цими ознаками є особливі точки (в тому числі і точки контуру). У наступних шарах ці ознаки збираються в групи, які стають ознакою наступного рівня і т. д. Ієрархія ознак дозволяє сформувати чисельні

дескриптори об'єктів на зображенні. Ці дескриптори лежать в основі задачі класифікації останніми шарами мережі, які, по суті, являють собою класичний багат шаровий перцептрон. Це означає, що процес навчання мережі включати тренування мережі, в тому числі, і для формування ознак контурів, хоча ця задача може бути вирішена більш ефективно класичними методами обробки зображень [9].

Застосування трансферного навчання дозволяє обійти обмеження, пов'язані з часом навчання. Проте, залишаються відкритими деякі питання. По-перше, це питання про надмірності універсального підходу на основі глибоких нейронних мереж в разі, коли мова йде про класифікацію повітряних об'єктів. По-друге, питання про віддільність класів об'єктів, кожен з яких представлений дуже різними (за рахунок просторової орієнтації) образами об'єктів на плоскому зображенні.

Опис зображення об'єкта його контуром є достатнім для задач класифікації повітряних об'єктів та використовує набагато менший обсяг інформації, ніж при аналізі за допомогою глибоких нейронних мереж, що дозволяє отримати ряд переваг. Відомі різні методи математичного опису контуру [10, 11], в тому числі і для визначення типів повітряних об'єктів [12].

Існують у відкритому доступі як модулі для вбудовування у власне програмне забезпечення версії глибоких нейронних мереж, що добре себе зарекомендували для розв'язання задач штучного інтелекту: від розпізнавання зображень до оброблення природної мови. Ці моделі пропонують не тільки готову архітектуру мережі, а й це дуже важливо – вагові коефіцієнти попередньо натренованої мережі. Використовувану в цій роботі мережу Inception v3 [9] натреновано на одній з найбільших баз даних зображень ImageNet з більш ніж 10 мільйонами зображень, анотованих для 1000 класів об'єктів.

Сучасні моделі розпізнавання зображень – це глибокі згорткові нейронні мережі (Convolutional Neural Network, CNN) з численними шарами. Нижні шари такої CNN-мережі відповідають за розпізнавання низькорівневих елементарних

ознак, таких як краї, контури та частини зображення. Середні шари забезпечують розпізнавання груп низькорівневих ознак, які своєю чергою є елементарними ознаками для наступних шарів. Останній шар визначає категорію зображення.

Трансферне навчання у випадку класифікування зображень ЛА дає змогу повторно натренувати останній шар такої CNN-мережі за допомогою власного набору зображень у прийнятний час, не змінюючи ваги інших шарів та досягаючи необхідної точності.

Ще одна перевага трансферного навчання – це те, що для додаткового тренування останнього шару CNN-мережі необхідна не така велика кількість даних.

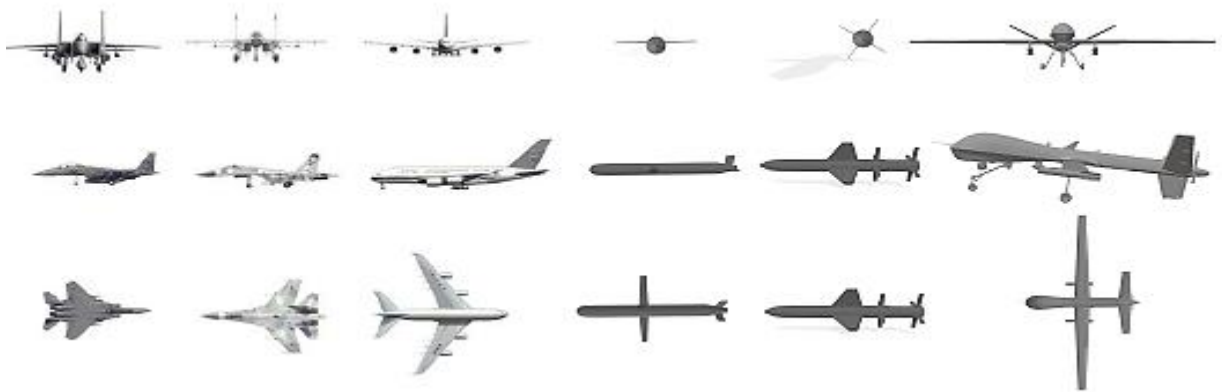


Рисунок 3.2 – Три ортогональні проекції для літаків Су-27, F15, A380, ракет BGM-109 Tomahawk, AVMT300 і БПЛА MQ-9 Predator

У розпізнаванні 3-х вимірних об'єктів (рисунок 3.2) по їх 2-х вимірному зображенню, що принципово важливо для розпізнавання ПО, виникає проблема отримання числового дескриптора, інваріантного щодо орієнтації об'єкта і його розмірів на зображенні. Для подолання цієї проблеми використовувалися різні характеристики контуру літака, отримані з допомогою моментів X_y , моментів Цернике і вейвлет моментів [1].

Для генерації вихідних зображень ЛА було використано сіткові тривимірні моделі ЛА як на рис. 3.3.

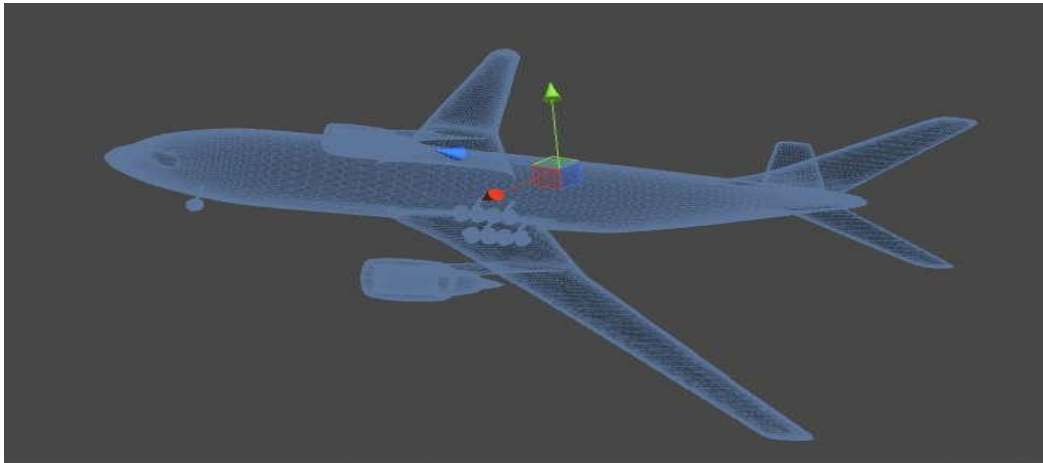


Рисунок 3.3 – Сіткова тривимірна модель літаків

Для цієї задачі використовувалися моделі для визначення одного з класів: літак, БПЛА, гелікоптер, квадрокоптер (рис. 3.4). Були використані різні види ЛА для кожного класу.



Рисунок 3.4 – Тривимірні моделі основних типів ЛА для класифікації

Ці моделі для анімації були включені в проєкт для Unity – середовища розробки комп'ютерних ігор. Програмне керування положенням ЛА в ігровому середовищі дало змогу підготувати кілька сотень зображень під різними кутами.

Процес оцінки якості класифікації і помилки в ході навчання наведені на рисунку 3.5.

Графік точності класифікації підтверджує правильність запропонованого підходу до вирішення задачі класифікації повітряних об'єктів.

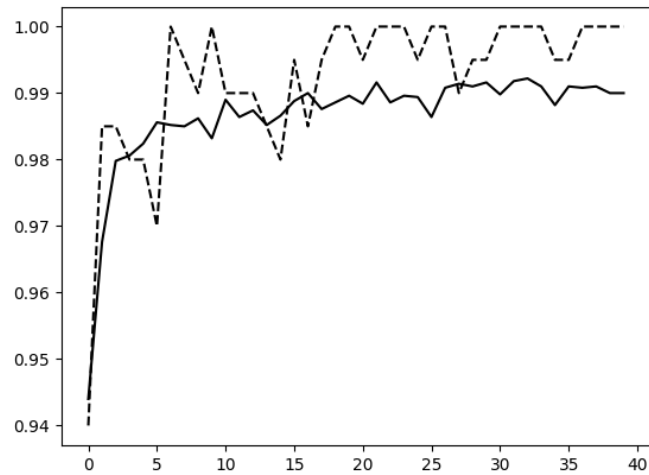


Рисунок 3.5 – Графік оцінки точності класифікації в процесі навчання (штрихована лінія – оцінка на тренувальному наборі, суцільна – на контрольному наборі)

3.3 Опис програми

Python – багатоцільова мова програмування, яка дозволяє писати код, що добре читається. Відносний лаконізм мови Python дозволяє створити програму, що буде коротшою свого аналога, написаного на іншій мові.

Python – багатоплатформова мова програмування. Це означає, що програми на Python можна запускати в різних операційних системах без будь-яких змін.

Ще однією перевагою Python є його стандартна бібліотека, яка встановлюється разом з Python і містить готові інструменти для роботи з операційною системою, веб-сторінками, базами даних.

Висновки за розділом 3

Висновком 3 розділу є дослідження можливості застосування нейронних мереж з комплекснозначним обчисленням для вирішення задачі класифікації з

комплексним входом, яке забезпечує визначення типів повітряних об'єктів за описом їх двомірного контуру на цифровому зображенні в системі відеовиявлення.

Запропонований підхід дає змогу отримати розв'язання поставленої задачі з досить високим ступенем точності на модельних даних. Для розв'язання задачі в реальних умовах цей спосіб може бути запропоновано як один із методів, який разом з іншими методами (наприклад, визначення динамічних характеристик польоту за відеорядом зображень) дасть змогу отримати правильне розв'язання поставленої задачі.

Для досягнення поставленої мети були сформульовані наступні задачі:

- розробити алгоритмічне та програмне забезпечення для виявлення контуру об'єкта на зображенні і отримати його числове представлення, придатне для використання в задачі класифікації;

- розробити програмне забезпечення для графічного моделювання зображення 3-х вимірних моделей повітряних об'єктів, які знаходяться в заданому положенні по відношенню до площини відеокадру, з метою отримання набору даних для дослідження варіації ознак і створення навчальної вибірки для задачі класифікації;

- дослідити залежність зміни різних ознак, сформованих на основі контурного аналізу зображень, і визначити структуру і розмір векторів ознак, заснованих на обробці цифрового уявлення контуру їх 2-х вимірної проекції, які найкращим чином підходили б для вирішення задачі класифікації досліджуваного набору типів повітряних об'єктів;

- дослідити можливість застосування комплекснозначної нейронної мережі і визначити її найбільш доцільний тип для вирішення задачі класифікації повітряних об'єктів в системі відеовиявлення по опису їх контурів за допомогою запропонованого набору ознак;

- провести комп'ютерне моделювання задачі навчання і тестування запропонованої нейронної мережі.

4 РЕЗУЛЬТАТИ ОБЧИСЛЮВАЛЬНОГО ЕКСПЕРИМЕНТУ ТА ЇХ АНАЛІЗ

4.1 Структура обчислення системи розпізнавання класів повітряних об'єктів на інфрачервоному зображенні

Задача класифікації, яка вирішується в цій роботі, є складовою частиною апаратно-програмного комплексу виявлення рухомих повітряних об'єктів та визначення їх характеристик. В рамках цієї системи задача виявлення вирішується таким чином, що на першому етапі об'єкт або об'єкти спочатку повинні бути виявлені як рухомі в послідовності відеокадрів і локалізовані на відеозображенні. На другому етапі вирішується задача віднесення кожного об'єкта на кадрі до певного класу.

В результаті аналізу різних підходів до задачі класифікації повітряних об'єктів, описаної вище, запропонований наступний алгоритм її вирішення.

Крок 1. Попередня обробка зображення для отримання оцифрованого контуру у вигляді послідовності координат.

Крок 2. Отримання послідовності дескрипторів Фур'є для контуру і усічення її до 32 комплексних значень.

Крок 3. Формування 15 комплексних нормалізованих дескрипторів Фур'є.

Крок 4. Розв'язання задачі класифікації за допомогою попередньо навченої комплекснозначної нейронної мережі, яка відносить 15-вимірний комплексний вектор до одного з 11 класів. Для кожного з 4 типів повітряних об'єктів порівнюються 3 класи контурів для 3 ортогональних проєкцій за винятком квадрокоптеру, у якого дві проєкції збігаються.

Для розпізнавання повітряного об'єкта, необхідний набір еталонних зображень збирається в єдину базу даних, з якої буде відбуватися порівняння реального зразку.

Для генерації вихідних зображень ПО були використані сіткові тривимірні моделі, як на рисунку 4.1.

Для даної задачі використовувалися моделі для визначення одного з класів: літак, БПЛА, вертоліт, квадрокоптер (рисунок 4.6). Були використані різні види об'єктів для кожного класу. При цьому для кожного об'єкту в якості еталонного плоского зображення використовувалися зображення, що відповідають трьом різним проекціям (рисунок 4.2).

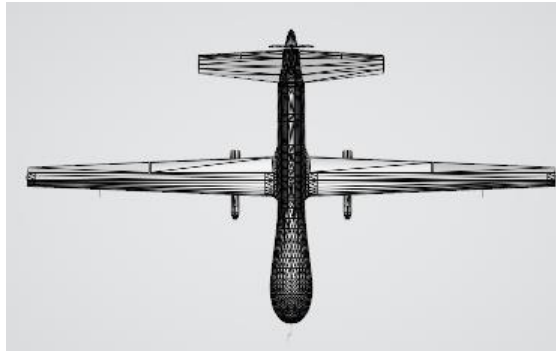


Рисунок 4.1 – Сіткова тривимірна модель БПЛА



Рисунок 4.2 – Тривимірні моделі основних класів ПО для тестування методу фазової кореляції



Рисунок 4.3 – Зображення, що відповідають трьом різним проекціям для БПЛА

Слід зазначити, що використання комплекснозначною нейронної мережі для вирішення задачі класифікації є новим напрямком у цій предметній області і вимагає додаткового дослідження. Перспективним є також напрямок

дослідження, пов'язаний з обчисленням параметрів афінного перетворення і визначенням реального просторового розташування повітряних об'єктів на окремому кадрі по знайденим дескрипторам, що дозволить визначати характер маневру повітряних об'єктів в відеопослідовності.

Для початку, нам знадобиться речова проєктивна площина P^2 . Проектна площина складається з кінцевих точок, виду $(x, y, 1)$ та ідеальних точок, виду $(x, y, 0)$. Точка $(0, 0, 0)$ не задає елемент P^2 . На P^2 задані класи еквівалентності – якщо домножити трійку чисел на ненульовий множник, то вони будуть задавати ту саму точку в P^2 .

$$\text{Тобто } (x, y, z) \sim (\omega x, \omega y, \omega z), \omega \neq 0.$$

Елементи проєктивної площини можна як кінцеві точки $(x, y, 1)$ на площині R^2 , доповнені ідеальними точками H_{il} – точками на нескінченності в напрямку (x, y) . Також елементи проектної площини можна сприймати як лінії на площині $(a, b, c) \sim ax + by + c = 0$. У всіх теоремах про проєктивну площину працює принцип двоїстості, коли можна поміняти місцями точки та лінії та отримати подвійну теорему.

P^2 також можна уявити як безліч променів (хоча точніше буде сказати прямих), вкладених в R^3 і проходять через початок координат. Промені, що перетинають площину $z = 1$, відповідають кінцевим точкам P^2 , а промені у площині $z = 0$ – ідеальним.

Гомографія H (проєктивне перетворення) – це перетворення $P^2 \rightarrow P^2$ яке прямі переводить у прямі. Для нас важливо, що гомографія задається множенням однорідних координат на оборотну матрицю (x, y, z) . Зверніть увагу, що гомографія ωH задає одне й те саме перетворення для всіх ненульових множників ω .

Розглянемо, як працює модель пінхол камери. Вона ставить те, як 3D точка на площині зображення:

$$u = f_x \frac{x}{z} + c_x, v = f_y \frac{y}{z} + c_y.$$

В цю же точку (u, v) проєктуються всі крапки на промені $w(x, y, z), w \neq 0$. Тобто модель камери зіставляє 2D-точку у площині зображення променю у 3D просторі.

У \mathbb{P}^2 ми зіставляємо точці $(x, y, 1)$ у площині $z = 1$ промінь (пряму)

$$w(x, y, 1), w \neq 0.$$

Тобто ототожнення точки на площині з променем у 3D вже мається на увазі і саме собою проєктування вже відбулося. Для того, щоб відповідати моделі, залишилося зробити розтяг і зміщення площини. Це можна зробити наступною гомографією:

$$H_p = \begin{pmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{pmatrix}.$$

Перевіримо. Нехай у нас задана (x, y, z) . Спроєктуємо її на площину точка зображення за допомогою гомографії H_p .

$$H_p(x, y, z)^T = (f_x x + c_x z, f_y y + c_y z, z) \sim (f_x \frac{x}{z} + c_x, f_y \frac{y}{z} + c_y, 1).$$

Те саме, що передбачається моделлю камери.

Для побудови плоскої панорами нам потрібно поєднати всі вихідні зображення в одній площині.

Для спрощення нашого алгоритму використовуватимемо площину першого зображення.

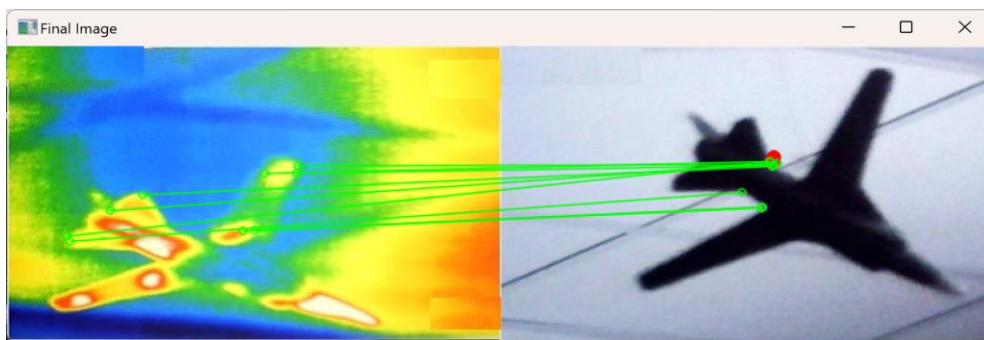


Рисунок 4.4 – Зображення співставлення точок

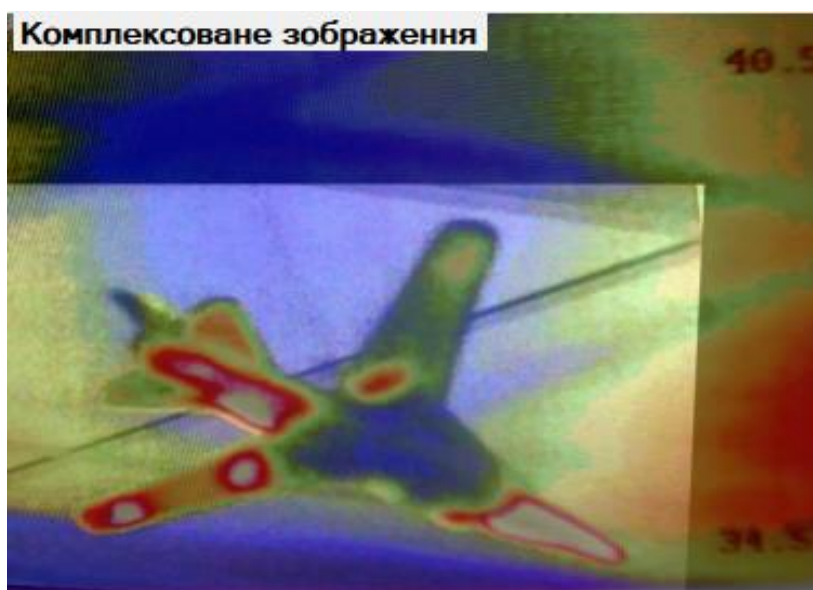


Рисунок 4.5 – Зображення після знаходження гомографій

Перше зображення можна перетворити саме на себе за допомогою identity гомографії $H_{11} = I_{3 \times 3}$. Для кожного з інших зображень потрібно знайти гомографію H_{i1} , яка переводить зображення у площину першого зображення. Це не завжди вдасться зробити. У найгіршому випадку наша панорама складатиметься лише з першого зображення.

Після знаходження гомографій, коли ми маємо набір зображень та відповідні їм матриці гомографії, ми можемо обчислити межі панорамного зображення та об'єднати вихідні фотографії в одне зображення.

4.2 Навчання нейромережевого класифікатора

Вирішенням описаних проблем визначення типу літального апарата є трансферне навчання глибоких нейронних мереж, що дає змогу використовувати попередньо навчену (навіть для інших класів об'єктів) мережу для побудови класифікаторів зображень ЛА. Як базову переднавчену мережу використовували мережу Inception v3.

Трансферне навчання передбачає розв'язання двох завдань: зміна архітектури НС і навчання зміненої мережі. Зміна архітектури полягала в заміні останнього класифікаційного шару на шар із 4 елементами (за кількістю типів ЛА). Числове значення виходів останнього шару визначає ймовірність віднесення вхідного зображення до відповідного класу.

У процесі навчання було "заморожено" ваги всіх нейронів попередньо навченої на зображеннях ImageNet моделі, крім нейронів останнього шару. 4000 епох навчання дали змогу отримати розв'язання задачі класифікації. Процес оцінювання якості класифікації та помилки під час навчання наведено на рис. 4.6.

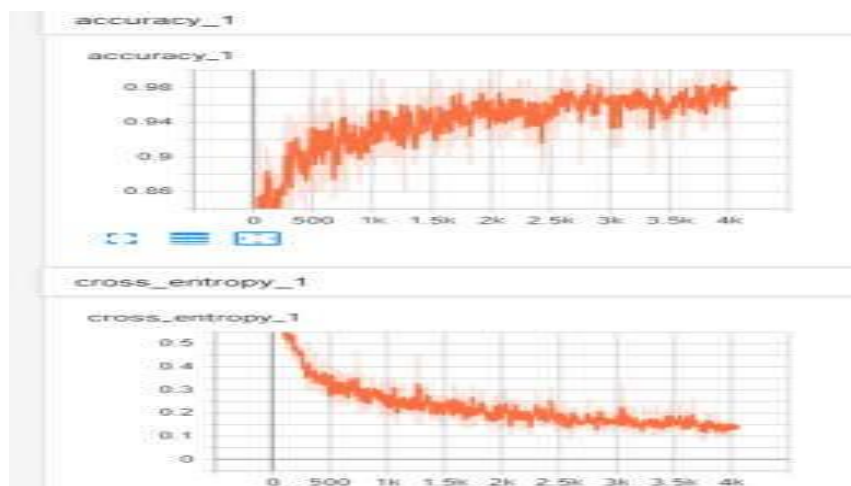


Рисунок 4.6 – Графік оцінки якості класифікації та помилки в процесі навчання

Результати роботи класифікатора на реальних зображеннях на рис 4.7.



Рисунок 4.7 – Результати роботи класифікатора на реальних зображеннях

Програмне забезпечення для проєкту розроблено в частині тривимірного моделювання на C# в Unity. Програми навчання мережі та її використання для класифікації написані на Python з використанням бібліотеки TensorFlow.

Запропонований підхід дає змогу отримати розв'язання поставленої задачі з досить високим ступенем точності на модельних даних. Для розв'язання задачі в реальних умовах цей спосіб може бути запропоновано як один із методів, який разом з іншими методами (наприклад, визначення динамічних характеристик польоту за відеорядом зображень) дасть змогу отримати правильне розв'язання поставленої задачі.

Висновки за розділом 4

Отримані результати збіжності процедури навчання (рис. 4.6) свідчать про те, що запропонований підхід до вибору нормалізованих дескрипторів контуру в якості векторів ознак об'єкта для вирішення задачі класифікації повітряних об'єктів за допомогою комплексно значною нейронної мережі, дозволяє будувати ефективну підсистему розпізнавання для комплексу виявлення рухомих повітряних об'єктів. Дійсно, з рисунку 3.2 видно, що контрольна точність, отримана на зображеннях, які не використовувалися для

тренування нейронної мережі досягає 99 %. При цьому слід зазначити, що ефективність розв'язання задачі класифікації досягається за рахунок поділу навчальної вибірки на класи, які відповідають трьом різним проекціям об'єкту. Це призводить до спрощення процедури навчання, бо нейронній мережі не доводиться відносити до одного класу несхожі зображення одного і того ж об'єкта в різних ракурсах і, отже, сильно відрізняються багатовимірні ознаки. Крім того вибір комплекснозначної нейронної мережі дозволяє провести процедуру навчання без використання градієнтних методів, що не приводить до ефекту застрягання в локальних мінімумах. Це пов'язано з вибором архітектури мережі, побудованої на комплексному обчисленні, що найбільшою мірою відповідає уявленню векторів ознак в комплексному просторі. Таке уявлення диктується комплексною формою методу перетворення Фур'є, що використовується для обчислення нормалізованих дескрипторів. Використання нормалізованих дескрипторів дозволяє уникнути проблем зазвичай використовуваних в контурному аналізі дескрипторів Фур'є, залежності від таких факторів як: спотворення форми контуру об'єкта через взаємне просторове розташування відеокамери і повітряних об'єктів, і вибору початкової точки контуру при обробці зображення. Саме тому запропонований варіант структури системи розпізнавання класів повітряних об'єктів слід визнати перспективним.

ВИСНОВКИ

Під час виконання роботи вирішено актуальну та важливу науково-прикладну задачу підвищення ефективності (за критеріями точності та швидкодії) розпізнавання об'єктів на інфрачервоних зображеннях та відеопослідовностях шляхом подальшого розвитку методу розпізнавання об'єктів на основі згорткових нейромереж та алгоритмічних і програмних рішень його реалізації. Під час виконання досліджень отримано наступні практичні та наукові результати.

Проведено аналіз існуючих проблем виявлення об'єктів в оптико-електронних пристроях, який показав, що найбільший негативний внесок у розпізнавання вносять візуальні перешкоди, нерівномірний фон, погодні умови й людський фактор спостерігача.

Досліджено відомі патентні технічні рішення, що використовують процедуру розпізнавання об'єктів в інфрачервоному діапазоні та встановлено відсутність у більшості пристроїв швидкісного завадостійкого каналу передачі відеоданих, що важливо при розміщенні джерела відеоінформації на певній відстані (до 15 м) від обчислювального блоку чи блоку виведення пристрою.

Проведено аналіз сучасних одноетапних методів виявлення об'єктів, призначених для роботи в реальному часі та обґрунтовано вибір за критеріями швидкості і точності методу в якості базового для дослідження.

Результати можна застосовувати для обчислення багатьох задач.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Зошук М.М. Система розпізнавання математичних моделей з використанням алгоритмів глибокого навчання. *II Міжнародна молодіжна науково-практична конференція «Навчання і викладання: після війни та під час миру»* : зб. матеріалів форуму (м. Харків, 9–13 травня 2023 р.). Т. 7. Харків : ХНУРЕ, 2023. С. 180-194.
2. See the full spectrum. URL: <https://www.infiniioptics.com/electromagnetic-spectrum> (дата звернення: 15.10.2023).
3. See Heat. URL: <https://www.infiniioptics.com> (дата звернення: 15.10.2023).
4. Thermal Imager Range: Predictions, Expectations, and Reality / Dragana P., Branko L., Miroslav P., Sasa V. 2019. 329 p.
5. SSD: Single Shot MultiBox Detector. *Computer Vision. ECCV 2016, Springer*. 2016. Vol. 9905. P. 21–37.
6. Ramachandran P., Zoph B., Le Q. V. Searching Activation Functions. 2017.
7. Misra D. Mish A. Self Regularized Non-Monotonic Activation Function. 2019. P. 35–49.
8. Distance-IoU Loss: Faster and Better Learning for Bounding Box Regression / Zheng Z., Wang P., Liu W., Li J., Ye R., Ren D. 2001. 310 p.
9. Piella G. A generation framework for multisolution image fusion: from pixels to regions. *Infrared Fusion*. 2003. 280 p.
10. Jyotismita C., Nilanjan D. A beginner's guide to image shape feature extraction techniques. New York: CRC Press, 2019. P. 147.
11. Mönning N., Manandhar S. Evaluation of complex-valued neural networks on real-valued classification tasks. York : Department of Computer Science University of York, 2018. 18 p. (A preprint. Department of Computer Science University of York ; November 30, 2018).
12. Multiplevalued threshold functions, Boolean complex-threshold functions and their generalization / Aizenberg N. N., Ivaskiv Yu. L., Pospelov D. A., Hudia-kov G. F. *Kibernetika (Cybernetics)*. 1971. №4. P. 44–51.

13. Nitta T. Ability of the 1-n-1 Complex-Valued Neural Network to Learn Transformations. *Computational Modeling and Simulation of Intellect: Current State and Future Perspectives*. 2011. № 22. P. 566–597.

14. Li W, Bebis G, Bourbakis NG. 3D object recognition using 2D Views: *IEEE Trans Image Process*, 2008. 165 p.

15. Weighted Marginal Fisher Analysis with Spatially Smooth for aircraft recognition Wei, Zhenzhong, Liu, Chang, Li, Nan: *Chinese Journal of Aeronautics*. February, 2014. P. 27.