

**НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
імені ІГОРЯ СІКОРСЬКОГО»**

Факультет інформатики та обчислювальної техніки

Кафедра інформаційних систем та технологій

До захисту допущено:

Завідувач кафедри

_____ Олександр РОЛІК

«__» _____ 2025 р.

**Дипломний проєкт
на здобуття ступеня бакалавра
за освітньо-професійною програмою «Інформаційне забезпечення
робототехнічних систем»
спеціальності 126 «Інформаційні системи та технології»
на тему: «Підсистема комп'ютерного зору мобільного робота для
покращення якості вхідного зображення шляхом усунення впливу
погодних явищ»**

Виконала:

студентка IV курсу, групи ІК-11

Качуринець Надія Юріївна _____

Керівник:

доцент кафедри ІСТ, к.т.н., доц.

Олійник Володимир Валентинович _____

Рецензент:

доцент кафедри ІПІ, к.т.н., доц.

Лісовиченко Олег Іванович _____

Засвідчую, що у цьому дипломному
проєкті немає запозичень з праць інших
авторів без відповідних посилань.

Студентка _____

Київ – 2025 року

Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»
Факультет інформатики та обчислювальної техніки
Кафедра інформаційних систем та технологій

Рівень вищої освіти – перший (бакалаврський)

Спеціальність – 126 «Інформаційні системи та технології»

Освітньо-професійна програма «Інформаційне забезпечення робототехнічних систем»

ЗАТВЕРДЖУЮ

Завідувач кафедри

_____ Олександр РОЛІК

«__» _____ 2025 р.

ЗАВДАННЯ
на дипломний проєкт студентці
Качуринець Надії Юрївні

1. Тема проєкту «Підсистема комп'ютерного зору мобільного робота для покращення якості вхідного зображення шляхом усунення впливу погодних явищ», керівник проєкту Олійник Володимир Валентинович, доцент кафедри ICT, к.т.н., затверджені наказом по університету від «23» травня 2025 р. № 1705-с

2. Термін подання студентом проєкту: 12 червня 2025 року

3. Вихідні дані до проєкту: таблиці з метриками оцінки моделей PSNR та SSIM, таблиці швидкодії моделей на CPU та GPU, вихідні зображення з наборів даних RESIDE-6K, SnowCityScapes, Rain1400, WeatherStream

4. Зміст пояснювальної записки:

1. Задача покращення якості вхідного зображення шляхом усунення впливу погодних явищ: постановка задачі, опис задачі й її особливостей, існуючі підходи до вирішення задачі, існуючі реалізації глибоких нейронних мереж для змішаних погодних умов, формулювання конкретних вимог до розробки.

2. Інформаційне забезпечення: модель даних та архітектурні схеми, алгоритмічні складові, набори даних.

3. Програмне забезпечення: вибір інструментів та технологій, використані апаратні ресурси, опис програмного продукту, структура програмного забезпечення, обробка даних для циклу тренування, опис циклу підготовки моделей машинного навчання, опис етапу валідації моделей, інтерфейсні компоненти.

4. Результати роботи та експериментальні дослідження: результати валідації моделей, результати тестування на апаратних пристроях, сфера застосування розробленої підсистеми, перспективи вдосконалення підсистеми.

5. Перелік графічного матеріалу (із зазначенням обов'язкових креслеників, плакатів, презентацій тощо)

1. Блок схема обробки даних у підсистемі комп'ютерного зору

2. Блок схема алгоритму підсистеми комп'ютерного зору

3. Схема структурна підсистеми комп'ютерного зору

4. Діаграма прецедентів підсистеми комп'ютерного зору

6. Дата видачі завдання 03.03.2025

Календарний план

№ з/п	Назва етапів виконання дипломного проєкту	Термін виконання етапів проєкту	Примітка
1	Аналіз предметної області	07.03.25 – 15.03.25	
2	Постановка мети, опис задачі та її особливостей	15.03.25 – 20.03.25	
3	Огляд існуючих реалізацій та підходів до вирішення задачі	20.03.25 – 27.03.25	
4	Огляд та вибір алгоритмів машинного навчання, пошук наборів даних	27.03.25 – 04.04.25	
5	Проектування архітектури підсистеми та написання програмного коду	04.04.25 – 18.04.25	
6	Тренування моделі та аналіз результатів	18.04.25 – 09.05.25	
7	Подальше дослідження натренованих моделей	09.05.25 – 16.05.25	
8	Розробка графічного інтерфейсу	16.05.25 – 24.05.25	

Студентка

Надія КАЧУРИНЕЦЬ

Керівник

Володимир ОЛІЙНИК

АНОТАЦІЯ

Підсистема комп'ютерного зору мобільного робота для покращення якості вхідного зображення шляхом усунення впливу погодних явищ.

Проект містить 83 с. тексту, 10 рисунків, 14 таблиць, посилання на 21 літературне джерело, додатки та 4 конструкторських документи.

ОЧИЩЕННЯ ЗОБРАЖЕННЯ, ПОГОДНІ ЯВИЩА, ГЛИБОКЕ НАВЧАННЯ, МАШИННЕ НАВЧАННЯ, КОМП'ЮТЕРНИЙ ЗІР, RESIDE-6K, RAIN1400, SNOWCITYSCAPES, WEATHERSTREAM, TANET, MWFORMER, PYTHON, PYTORCH.

Об'єктом розробки є підсистема комп'ютерного зору для покращення якості вхідного зображення шляхом усунення впливу погодних явищ.

Мета розробки — підвищення ефективності роботи підсистеми комп'ютерного зору мобільного робота шляхом підвищення якості вхідного зображення за умов туману, дощу та снігу із застосуванням сучасних методів глибокого навчання.

У дипломному проєкті досліджено моделі глибокого навчання TANet та MWFormer для очищення вхідних зображень від ефектів погодних явищ, а також на їх основі розроблено підсистему комп'ютерного зору для покращення візуальної якості вхідного потоку. Значну увагу було приділено створенню збалансованих навчальних датасетів, що включають як синтетичні, так і реальні зображення, а також налаштуванню циклу тренування для досягнення максимальної якості результатів. У роботі порівняно ефективність обох моделей за ключовими метриками (PSNR, SSIM) на різних типах атмосферних викривлень.

Отримані результати можуть бути корисними для впровадження в системах автономного транспорту, відеоспостереження, робототехніки, а також у будь-яких застосуваннях, де необхідно забезпечити стабільну якість зображення за несприятливих погодних умов.

SUMMARY

Computer vision subsystem of a mobile robot to improve the quality of the input image by eliminating the influence of weather phenomena.

The project contains 83 pages text, 10 figures, 14 tables, links to 21 literary sources, annexes and 4 design documents.

Keywords: image cleaning, weather phenomena, deep learning, machine learning, computer vision, RESIDE-6K, Rain1400, SnowCityScapes, WeatherStream, TANet, MWFormer, Python, PyTorch.

The object of development is a computer vision subsystem to improve the quality of the input image by eliminating the influence of weather phenomena.

The purpose of the development is to increase the efficiency of solving the problem of recognizing the environment and objects by a mobile robot by improving the quality of the input image in conditions of fog, rain, and snow through the use of modern deep learning methods.

The thesis project investigated the deep learning models TANet and MWFormer for cleaning input images from the effects of weather phenomena, and based on them, a computer vision subsystem was developed to improve the visual quality of the input stream. Considerable attention was paid to the creation of balanced training datasets, including both synthetic and real images, as well as to tuning the training cycle to achieve maximum quality of results. The work compared the effectiveness of both models by key metrics (PSNR, SSIM) on different types of atmospheric distortions.

The results obtained can be useful for implementation in autonomous transport systems, video surveillance, robotics, as well as in any applications where it is necessary to ensure stable image quality under adverse weather conditions.

№ рядка	Формат	Позначення	Найменування	Кіл. аркушів	№ екз.	Примітка
1			Документація загальна			
2						
3			Знову розроблена			
4						
5	A4	ІК11.120БАК.006 ПЗ	Підсистема комп'ютерного зору	83		
6			мобільного робота для			
7			покращення якості вхідного			
8			зображення шляхом усунення			
9			впливу погодних явищ.			
10			Пояснювальна записка			
11	A3	ІК11.120БАК.006 Д1	Підсистема комп'ютерного зору	1		
12			мобільного робота для			
13			покращення якості вхідного			
14			зображення шляхом усунення			
15			впливу погодних явищ.			
16			Блок схема обробки даних			
17	A3	ІК11.120БАК.006 Д2	Підсистема комп'ютерного зору	1		
18			мобільного робота для			
19			покращення якості вхідного			
20			зображення шляхом усунення			
21			впливу погодних явищ.			
22			Діаграма прецедентів			
23	A3	ІК11.120БАК.006 ДЗ	Підсистема комп'ютерного зору	1		
24			мобільного робота для			
25			покращення якості вхідного			
26			зображення шляхом усунення			
27			впливу погодних явищ.			
28			Блок схема алгоритму			
			ІК11.120БАК.006 ТП			
		№ докум.	Підпис			
Розробив	Качуринець			Літ.	Арк.	Аркушів
Перевірів	Олійник В.В.			Т	1	2
			Підсистема комп'ютерного зору мобільного робота для покращення якості вхідного зображення шляхом усунення впливу погодних явищ. Відомість дипломного проєкту	КПІ ім. Ігоря Сікорського Група ІК-11		
Затв.						

№ рядка	Формат	Позначення	Найменування	Кіл. аркушів	№ екз.	Примітка
29	A3	ІК11.120БАК.006 Д4	Підсистема комп'ютерного зору	1		
30			мобільного робота для			
31			покращення якості вхідного			
32			зображення шляхом усунення			
33			впливу погодних явищ.			
34			Структурна схема			
35						
36						
37						
38						
39						
40						
41						
42						
43						
44						
45						
46						
47						
48						
49						
50						
51						
52						
53						
54						
55						
56						
					ІК11.120БАК.006 ТП	
Зм.	Лист	№ докум.	Підпис	Дата		
					Арк. 2	

**Пояснювальна записка
до дипломного проєкту
на тему: «Підсистема комп'ютерного зору мобільного
робота для покращення якості вхідного зображення
шляхом усунення впливу погодних явищ»**

ЗМІСТ

ВСТУП	5
1 ЗАДАЧА ПОКРАЩЕННЯ ЯКОСТІ ВХІДНОГО ЗОБРАЖЕННЯ ШЛЯХОМ УСУНЕННЯ ВПЛИВУ ПОГОДНИХ ЯВИЩ	7
1.1 Постановка задачі покращення якості вхідного зображення	7
1.1.1 Актуальність проблеми усунення впливу погодних явищ	7
1.1.2 Об'єкт і предмет дослідження	7
1.1.3 Мета і завдання дослідження	7
1.1.4 Наукова новизна	8
1.1.5 Практичне значення результатів	9
1.1.6 Положення, винесені на захист	9
1.2 Опис задачі усунення атмосферних спотворень та її особливостей.....	9
1.2.1 Актуальність, суть задачі та предметна область.....	9
1.2.2 Особливості задачі усунення впливу погодних явищ.....	11
1.3 Існуючі підходи до вирішення задачі усунення впливу погодних явищ	13
1.3.1 Класичні методи усунення атмосферних спотворень	13
1.3.2 Методи машинного навчання та нейронні мережі	15
1.4 Існуючі реалізації глибоких нейронних мереж для змішаних погодних умов	17
1.5. Формулювання конкретних вимог до розробки підсистеми	19
1.5.1. Загальні вимоги до розробки підсистеми	19
1.5.2. Функціональні вимоги до розробки підсистеми.....	20
1.5.3 Нефункціональні вимоги до розробки підсистеми.....	20
Висновки до розділу 1	21
2 ІНФОРМАЦІЙНЕ ЗАБЕЗПЕЧЕННЯ.....	23
2.1 Модель даних та архітектурні схеми підсистеми	23

					ІК11.120БАК.006 ПЗ			
		№ докум.	Підпис					
Розробив	Качуринець				Підсистема комп'ютерного зору мобільного робота для покращення якості вхідного зображення шляхом усунення впливу погодних явищ. Пояснювальна записка	Літ.	Арк.	Аркушів
Перевірив	Олійник В.В.					Т	2	83
						КПІ ім. Ігоря Сікорського Група ІК-11		
Затв.								

2.1.1 Блок схема обробки відеоданих у підсистемі комп'ютерного зору	23
2.1.2 Діаграма прецедентів підсистеми комп'ютерного зору.....	24
2.1.3 Блок схема алгоритму підсистеми комп'ютерного зору	25
2.1.4 Структурна схема підсистеми комп'ютерного зору.....	26
2.2 Алгоритмічні складові підсистеми комп'ютерного зору.....	28
2.2.1 Triplet Attention Network (TANet).....	28
2.2.2 Multi Weather Transformer (MWFormer)	30
2.3 Набори даних.....	31
2.3.1 Набір даних SnowCityScapes.....	32
2.3.2 Набір даних Rain1400	32
2.3.3 Набір даних RESIDE-6K.....	32
2.3.4 WeatherStream Dataset.....	33
Висновки до розділу 2	34
3 ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ.....	35
3.1 Вибір інструментів та технологій.....	35
3.2 Використані апаратні ресурси	36
3.3 Опис програмного продукту	37
3.4 Структура програмного забезпечення підсистеми комп'ютерного зору	38
3.5 Обробка даних для циклу тренування	41
3.5.1 Попередня обробка даних для моделі TANet.....	41
3.5.2 Попередня обробка даних для моделі MWFormer.....	42
3.6 Опис циклу підготовки моделей машинного навчання	43
3.6.1 Цикл тренування TANet	44
3.6.2 Цикл тренування MWFormer	48
3.8 Опис етапу валідації моделей	54
3.8.1 Валідація для моделі TANet.....	54
3.8.2 Валідація для моделі MWFormer.....	55
3.9 Інтерфейсні компоненти.....	56
3.9.1 Графічний інтерфейс користувача (GUI)	56
Висновки до розділу 3	57

4 РЕЗУЛЬТАТИ РОБОТИ ТА ЕКСПЕРИМЕНТАЛЬНІ ДОСЛІДЖЕННЯ..	59
4.1 Результати валідації моделей.....	59
4.2 Результати тестування на апаратних пристроях.....	69
4.3 Сфера застосування розробленої підсистеми	74
4.3.1. Застосування моделі TANet	74
4.3.2 Застосування моделі MWFormer	75
4.4 Перспективи вдосконалення підсистеми.....	76
Висновки до розділу 4	78
ВИСНОВКИ.....	80
ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	82
ДОДАТОК А.....	84
ДОДАТОК Б	87
ДОДАТОК В.....	90
ДОДАТОК Г	93
ДОДАТОК Д.....	94

ВСТУП

У мобільних роботах, що виконують завдання поза лабораторією, якість вхідного зображення безпосередньо впливає на надійність і точність алгоритмів комп'ютерного зору. Погодні явища – туман, дощ, сніг – створюють додаткові перешкоди: розсіювання та поглинання світла призводять до втрати дрібних деталей, зниження контрастності й спотворення кольорів кадрів. Унаслідок цього страждають результати детекції, сегментації та класифікації об'єктів, що критично для автономної навігації, відеомоніторингу, аерофотозйомки та інших застосувань мобільних роботів.

Для вирішення цієї проблеми у робототехніці доцільно розробити підсистему попередньої обробки, здатну адаптивно усувати атмосферні спотворення та відновлювати природну якість зображення. Саме це і є предметом нашого дослідження: аналіз сучасних методів глибокого навчання на основі синтезованих і реальних даних із відкритих репозиторіїв та створення на їхній базі ефективної підсистеми корекції кадрів.

Сучасний стан розроблених технологій включає як класичні підходи, засновані на фізичних моделях розсіювання світла і статистичних пріорах, так і новітні рішення на основі згорткових нейронних мереж із механізмами уваги. Глибокі моделі демонструють високу здатність узагальнювати інформацію за різних умов, однак залишаються відкритими питання вибору оптимальних архітектур, які поєднують найкращу якість відновлення деталей і достатню швидкодію для роботи в реальному часі.

Обґрунтування необхідності проведення роботи полягає у зростаючих вимогах до автономних систем: застосування мобільних роботів у складних погодних умовах вимагає надійного попереднього усунення артефактів для забезпечення стійкої роботи алгоритмів комп'ютерного зору. Вихідними даними для дослідження стануть багаторівневі датасети, що включають синтезовані сцени з контрольованою інтенсивністю туману, дощу та снігу, а також реальні

зображення зі спеціалізованих репозиторіїв (наприклад, RESIDE, Foggy Cityscapes).

У межах роботи буде виконано:

- огляд і порівняльний аналіз фізичних і статистичних методів та сучасних архітектур глибинних мереж;
- тренування обраних моделей на зібраних датасетах із оцінкою за метриками PSNR, SSIM;
- тестування роботи підсистеми для цифрових зображень та відеопотоку на різних апаратних засобах.

Дипломний проєкт складається з таких розділів:

- вступ, з обґрунтуванням актуальності та постановки завдань;
- розділ 1: задача покращення якості вхідного зображення шляхом усунення впливу погодних явищ;
- розділ 2: інформаційне забезпечення підсистеми комп'ютерного зору;
- розділ 3: програмне забезпечення підсистеми комп'ютерного зору;
- розділ 4: результати роботи та експериментальні дослідження;
- висновки: підсумки та перспективи подальшої роботи;
- список використаних джерел;
- додатки: вихідний код, конфігурації, додаткові графіки та таблиці;
- графічний матеріал.

Очікуваний результат – адаптивний модуль передобробки, що суттєво зменшить атмосферні артефакти, відновить природну колірну гаму та контрастність без втрати ключових особливостей сцени. Практична значущість роботи полягає в підвищенні стійкості систем детекції й трасування об'єктів у мобільних роботах за умов змішаних погодних явищ, що розширить можливості їх застосування в промисловості, екологічному моніторингу та рятувальних операціях.

1 ЗАДАЧА ПОКРАЩЕННЯ ЯКОСТІ ВХІДНОГО ЗОБРАЖЕННЯ ШЛЯХОМ УСУНЕННЯ ВПЛИВУ ПОГОДНИХ ЯВИЩ

1.1 Постановка задачі покращення якості вхідного зображення

1.1.1 Актуальність проблеми усунення впливу погодних явищ

Актуальність теми зумовлена потребою підвищення надійності та безпеки мобільних роботів, що функціонують поза лабораторними умовами. Здатність алгоритмів комп'ютерного зору стабільно працювати за змінних погодних умов (туман, дощ, сніг) є критично важливою для розширення сфер їхнього застосування – зокрема, у промисловості, агросекторі, екологічному моніторингу та під час рятувальних операцій. Зменшення атмосферних спотворень у зображеннях сприятиме підвищенню точності виявлення об'єктів і загальної ефективності роботизованих систем[19].

1.1.2 Об'єкт і предмет дослідження

Об'єкт дослідження – цифрові зображення та відео, спотворені атмосферними явищами, отримані з відкритих джерел даних.

Предмет дослідження – алгоритми глибокого навчання для автоматичного усунення атмосферних артефактів, побудови карт корекції та покращення візуальної якості кадрів.

1.1.3 Мета і завдання дослідження

Мета дослідження – підвищення ефективності роботи підсистеми комп'ютерного зору мобільного робота шляхом підвищення якості вхідного зображення за умов туману, дощу та снігу із застосуванням сучасних методів глибокого навчання. Результатом має стати підсистема, здатна в реальному часі адаптивно усувати атмосферні спотворення без втрати якості зображення.

Критерії ефективності:

					ІК11.120БАК.006 ПЗ	Арк.
						7
Зм.	Лист	№ докум.	Підпис	Дата		

- метрики якості зображення (PSNR, SSIM) на тестових наборах;
- здатність обробляти відео з частотою не менше 25 кадрів/с на вбудованих і мобільних платформах;
- універсальність: адаптація до змішаних погодних умов без зниження якості.

Основні завдання дослідження:

- а) проаналізувати природу атмосферних спотворень (туман, дощ, сніг) і труднощі їх компенсації;
- б) провести систематизований огляд сучасних методів (фізичні моделі, статистичні пріори, глибокі нейронні мережі з механізмами уваги);
- в) сформувати збалансовані датасети із синтезованими і реальними зображеннями з різними рівнями атмосферних артефактів;
- г) сформулювати вимоги до підсистеми: точність, швидкодія, апаратні обмеження, інтерфейс для відеоконвеєра;
- д) обрати і налаштувати передові архітектури глибоких мереж з механізмами уваги, що дозволяють локальну адаптацію корекції;
- е) реалізувати прототип підсистеми за допомогою TensorFlow/PyTorch та OpenCV;
- ж) провести тестування на підготовлених наборах із вимірюванням PSNR, SSIM та швидкодії на вбудованих платформах у реальному часі та в постобробці;
- и) надати рекомендації щодо оптимального використання підсистеми залежно від сценарію застосування.

1.1.4 Наукова новизна

Наукова новизна полягає у проведенні комплексного порівняння сучасних нейромережових підходів до корекції атмосферних спотворень на єдиному наборі тестових даних з різними поєднаннями туману, дощу й снігу. Це дозволяє визначити найефективніші архітектури для використання як у реальному часі, так і в режимі постобробки.

					ІК11.120БАК.006 ПЗ	Арк.
						8
Зм.	Лист	№ докум.	Підпис	Дата		

1.1.5 Практичне значення результатів

Практична цінність роботи полягає у створенні адаптивного модуля передобробки зображень для мобільних роботів, що дозволяє:

- зменшити атмосферні артефакти;
- покращити візуальну якість кадрів;
- зберегти високу швидкодію обробки;
- забезпечити стійкість алгоритмів детекції та трасування об'єктів у складних погодних умовах.

1.1.6 Положення, винесені на захист

- комплексне дослідження методів глибокого навчання для корекції атмосферних артефактів у відео мобільного робота;
- порівняльна оцінка продуктивності згорткових мереж із механізмами уваги на реальних та синтезованих датасетах;
- рекомендації щодо використання підсистеми в реальному часі на вбудованих пристроях із обґрунтуванням вибору моделей.

1.2 Опис задачі усунення атмосферних спотворень та її особливостей

1.2.1 Актуальність, суть задачі та предметна область

Сучасні мобільні роботи дедалі частіше використовуються у сферах автономного транспорту[20], моніторингу довкілля, безпеки та рятувальних операцій. Однією з ключових складових таких систем є підсистема комп'ютерного зору, яка забезпечує сприйняття середовища через відео з камер. Проте у реальних умовах експлуатації, зокрема під час дощу, снігу або туману, якість зображень суттєво погіршується. Це негативно впливає на точність роботи алгоритмів виявлення, розпізнавання та відстеження об'єктів, що в свою чергу ставить під загрозу надійність і безпеку мобільного робота.

					ІК11.120БАК.006 ПЗ	Арк.
						9
Зм.	Лист	№ докум.	Підпис	Дата		

Предметна область дослідження охоплює розробку та впровадження методів глибинного навчання для адаптивного відновлення якості зображень, спотворених атмосферними явищами, з подальшою інтеграцією відповідного модуля в підсистему комп'ютерного зору мобільних роботів. Особлива увага приділяється роботі в режимі реального часу на пристроях з обмеженими обчислювальними ресурсами – дронах, вбудованих системах та автономних платформах.

Дослідження показують, що в умовах туману та дощу точність виявлення об'єктів може значно знижуватися. Наприклад, у роботі "Robust Object Detection in Challenging Weather Conditions"[1] зазначено, що такі погодні умови призводять до зменшення точності детекції через зниження видимості та появу шумів, спричинених погодними явищами.

У сфері дистанційного зондування земної поверхні туман може призводити до значних втрат корисної інформації. Зокрема, дослідження "Remote Sensing of Fog: Spatially and Temporally Continuous Estimation and Monitoring"[2] підкреслює, що туман суттєво впливає на якість даних, отриманих з віддалених сенсорів, що обмежує точність екологічного моніторингу та оцінки стану довкілля. Аналогічні проблеми виникають у міських системах відеоспостереження, де атмосферні опади викликають збої сенсорів та фальшиві тривоги.

Основне завдання полягає в автоматичному відновленні структури сцени та кольорової достовірності кадрів, спотворених розсіюванням і поглинанням світла атмосферними частинками. Такі спотворення призводять до зниження контрасту, появи шумоподібних артефактів від крапель дощу чи снігу, а також до «затуманення» важливих деталей на передньому та задньому планах. Як наслідок, критичні об'єкти можуть не розпізнаватися системами детекції та сегментації[21]. Ключова складність полягає в нелінійній природі атмосферних спотворень, які змінюються залежно від глибини сцени та складу аерозольних частинок. Через це фіксовані алгоритми обробки зображень або лінійні фільтри часто виявляються неефективними.

У зв'язку з цим дедалі більше дослідників звертаються до методів глибинного навчання – згорткових нейронних мереж і трансформерів, здатних відновлювати

					ІК11.120БАК.006 ПЗ	Арк.
						10
Зм.	Лист	№ докум.	Підпис	Дата		

чисті зображення з урахуванням локальних та глобальних контекстних ознак. Наприклад, архітектури на зразок ZeroScatter[3] ефективно поєднують синтетичні й реальні дані, дозволяючи моделі узагальнювати між різними умовами видимості.

У структурі системи комп'ютерного зору мобільного робота модуль адаптивної корекції атмосферних спотворень доцільно розміщувати перед блоками детекції та трекінгу. Це забезпечує стабільну якість вхідного відео навіть на ресурсно обмежених платформах, включно з безпілотниками, роботизованими платформами та вбудованими сенсорними системами.

Отже, впровадження глибинних нейромережових методів очищення зображень у підсистему комп'ютерного зору мобільного робота є актуальним і перспективним напрямом, що дозволяє суттєво підвищити безпеку, точність та надійність роботи таких систем у реальних складних умовах експлуатації.

1.2.2 Особливості задачі усунення впливу погодних явищ

Аналізуючи специфіку автоматичного відновлення зображень, спотворених атмосферними явищами, можна виокремити низку особливостей, які значно ускладнюють розробку універсального та ефективного рішення.

Нелінійність спотворень. Атмосферні ефекти (розсіювання, поглинання, заломлення світла) змінюють інтенсивність пікселів по різному залежно від їхнього положення та колірного спектра. Спроба описати ці процеси лінійними операціями призводить до значних похибок: наприклад, рівномірні фільтри часто “змивають” деталі на краях об'єктів або створюють артефакти в областях різкої зміни освітленості. Тому для адекватного моделювання спотворень необхідні нелінійні компоненти (наприклад, згортки з адаптивною ваговою маскою або механізми уваги), які враховують складність оптичних процесів у кожній точці зображення.

Контекстна залежність спотворень. Ефект атмосферних феноменів на пікселі визначається не лише загальними умовами (інтенсивність туману або дощу), але й відстанню сцени від камери, формою й матеріалом об'єктів у кадрі. Пікселі, що знаходяться далі, відображають меншу контрастність і синій відтінок, тоді як

ближчі можуть виглядати більш чіткими. Це вимагає від алгоритму здатності розрізняти локальні контексти та коригувати зображення диференційовано – наприклад, через багаторівневі мережі, які спочатку оцінюють глибину сцени або використовують додаткову карту передачі, а вже потім застосовують локальні корекційні фільтри.

Змішані погодні умови. У реальному світі туман часто поєднується з дощем, снігом, а іноді і з низькою видимістю через пил або смог. Ці явища можуть одночасно додавати до зображення як загальний розмитий “вуалевий” ефект, так і яскраві шумоподібні краплі або кристали. Алгоритму необхідно розрізняти й усувати кілька типів артефактів одночасно, не “гублячи” інформацію при переході між режимами (наприклад, щоб дощ не став лише “додатковим голубим завісом”). Найбільш гнучкі рішення використовують багатозадачні підмережі або трансформери з розгалуженою архітектурою, які одночасно адаптуються до кількох джерел спотворення.

Обмеження відеопотоку та апаратні ресурси. Багато застосувань (дрони, вбудовані пристрої, системи відеоспостереження) працюють із високошвидкісним відеопотоком та мають обмежений обчислювальний бюджет. Це накладає жорсткі вимоги на затримки – обробка кадру має відбуватися за десятки мілісекунд, щоб не створювати лагів у потоці. У зв’язку з цим архітектура мережі має бути збалансованою між глибиною (для якості) та швидкістю (для продуктивності), а також підтримувати легковагові операції, наприклад роздільні convolution-групи або квантування ваг для прискорення інференсу на CPU чи енергоефективному апаратному прискорювачі.

Різномірність даних. Синтезовані набори даних зі штучно доданим туманом чи дощем зазвичай мають однорідну статистику шуму та рівномірний рівень спотворень. Реальні кадри з відкритих репозиторіїв можуть містити неоднорідні, мінливі умови – як слабкий туман в одній частині кадру, так і сильний дощ в іншій, плюс різні джерела світла та відблиски. Це призводить до необхідності добре відпрацьованої стратегії доменного переносу: мережі мають вміти

узагальнюватися на реальних прикладах, а не лише на синтетичних, інакше вони демонструватимуть високу якість лише на навчальній вибірці.

Ці складні особливості зумовлюють потребу в гібридних рішеннях, що комбінують фізично обґрунтовані моделі оптичного розсіювання з гнучкими методами глибокого навчання та мультизадачними архітектурами. Оскільки в межах проєкту передбачається дослідження виключно моделей глибокого навчання, акцент буде зроблено на виборі та адаптації сучасних нейромережових архітектур, здатних ефективно відновлювати зображення в умовах атмосферних спотворень. Особливу увагу приділено архітектурам, що поєднують локальні й глобальні контекстні ознаки, забезпечують високу якість реконструкції та придатні для використання на обмежених апаратних ресурсах. Такий підхід дозволить досягти балансу між точністю, швидкістю та універсальністю підсистеми без залучення фізичних моделей.

1.3 Існуючі підходи до вирішення задачі усунення впливу погодних явищ

Нижче наведено короткий огляд існуючих підходів до вирішення задачі покращення якості вхідного зображення шляхом усунення впливу погодних явищ. Спочатку розглядаються класичні фізико-статистичні методи, потім – сучасні моделі машинного навчання та нейронні мережі, що дозволяють обрати оптимальний шлях реалізації підсистеми.

1.3.1 Класичні методи усунення атмосферних спотворень

Фізичні моделі розсіювання. Класичні підходи до усунення атмосферних спотворень базуються на моделі атмосферного розсіювання світла[4], яка описує зменшення яскравості прямого випромінювання та додавання розсіяного світла ("повітряного світла") на шляху від об'єкта до камери. Ця модель дозволяє оцінити карту передачі (transmission map), що відображає ступінь прозорості сцени.

					ІК11.120БАК.006 ПЗ	Арк.
						13
Зм.	Лист	№ докум.	Підпис	Дата		

Ранні методи, такі як підхід Фаттала, вимагали багатокадрових послідовностей або додаткової інформації про сцену, що обмежувало їх застосування в реальному часі. Удосконалені варіанти адаптують оцінку параметрів моделі локально, наприклад, коригуючи передачу через сегментацію неба чи інші просторові пріори.

Однак ці методи мають обмеження, пов'язані з припущеннями про однорідність атмосфери та постійність параметрів розсіювання, що не завжди відповідає реальним умовам.

Статистичні пріори. Статистичні пріори використовують емпіричні спостереження для оцінки атмосферних спотворень у зображеннях.

а) Dark Channel Prior (DCP). Запропонований Хе та ін. у 2009 році, цей підхід ґрунтується на спостереженні, що в більшості локальних областей безтуманних зображень існують пікселі з дуже низькою інтенсивністю в одному з кольорових каналів. Це дозволяє оцінити карту передачі та відновити зображення без туману. Однак DCP може призводити до артефактів на різких межах та має високі обчислювальні витрати при згладжуванні карти передачі.

б) Haze Lines Prior[5]. Цей метод аналізує колірні закономірності в просторі хроматичності, дозволяючи точніше розділити атмосферний ефект і структуру сцени. Haze Lines Prior може краще справлятися з різноманітними умовами освітлення та складними сценами, але також має обмеження, пов'язані з припущеннями про колірні властивості зображень.

Хоча статистичні пріори є простими у реалізації та не потребують навчання на великих наборах даних, вони часто не враховують складні варіації атмосферних умов та можуть бути менш ефективними в реальних сценаріях.

Ці класичні методи заклали основу для подальшого розвитку підходів до усунення атмосферних спотворень. Однак їх обмеження в обробці складних та варіативних умов сприяли переходу до використання методів глибокого навчання, які здатні адаптивно моделювати складні взаємозв'язки в зображеннях та забезпечувати більш точне відновлення якості зображень у реальному часі.

1.3.2 Методи машинного навчання та нейронні мережі

а) ранній етап: навчання простих CNN.

Однією з перших моделей, що застосовували згорткові нейронні мережі для видалення туману, є DehazeNet. Ця модель використовує компактну архітектуру CNN для оцінки карти пропускання середовища, яка потім використовується для відновлення зображення без туману. DehazeNet демонструє перевагу над класичними методами, такими як DCP, у швидкодії, але може поступатися в деталізації відновлених зображень.

б) сучасні CNN-архітектури та механізми уваги.

Сучасні моделі використовують механізми уваги для покращення якості відновлення зображень:

– ADMC2-Net[6]. Ця модель інтегрує інформацію з RGB та HSV кольорових просторів, використовуючи механізми уваги для збереження кольорової достовірності та деталей зображення.

– ODD-Net[7]. Гібридна архітектура, що поєднує попередньо навчені моделі A-Net і T-Net у рамках атмосферної моделі розсіювання. ODD-Net забезпечує високу якість відновлення, особливо в оцінці карти пропускання та глибини, перевершуючи сучасні методи за метриками PSNR та SSIM.

– Fine-scale Perception Network. Ця модель спрямована на відновлення високочастотних деталей завдяки багаторівневому вилученню ознак, що підвищує різкість меж об'єктів.

в) методи з контрастивною та фізичною регуляризацією.

Інтеграція фізичних моделей та контрастивного навчання дозволяє покращити інтерпретованість та стійкість моделей:

Curricular Contrastive Regularization [8]. Цей підхід вбудовує фізичний розсіювальний процес у глибинні ознаки, що дозволяє зберегти реалістичність відновленого зображення та покращує узагальнення на реальних кадрах.

г) генеративні змагальні мережі (GAN).

Генеративні змагальні мережі складаються з двох компонентів: генератора, який створює зображення, та дискримінатора, який оцінює їхню реалістичність. Ці мережі навчаються у змагальному режимі, де генератор прагне створити зображення, що не відрізняються від реальних, а дискримінатор намагається їх розпізнати.

Приклади застосування:

– Dehaze-GAN. Ця модель використовує GAN для видалення туману з зображень, зберігаючи текстурні деталі та покращуючи візуальну якість.

– FD-GAN[9]. Пропонує повністю кінець-до-кінця архітектуру для видалення туману, використовуючи дискримінатор злиття, який враховує частотну інформацію для зменшення артефактів та покращення кольорової достовірності.

– Perceptually Optimized GAN (POGAN)[10]. Ця модель оптимізує сприйняття зображень, використовуючи спеціальну функцію втрат, яка враховує особливості зорового сприйняття людини, що дозволяє отримувати більш реалістичні результати.

Загалом, GAN-підходи дозволяють отримувати високоякісні зображення з покращеною деталізацією та реалістичністю, що є важливим для систем комп'ютерного зору в умовах поганої видимості.

д) рекурентні нейронні мережі (RNN).

RNN, зокрема їх варіанти, такі як LSTM, ефективно обробляють послідовні дані, що робить їх придатними для обробки відео, де важливо враховувати часову послідовність кадрів: Dual-Stage Temporal Fusion Network (DS-TFN). Ця модель використовує двоетапний підхід для видалення туману з відео, поєднуючи попередню обробку з подальшим вдосконаленням, враховуючи часову інформацію для покращення якості зображення.

RNN-підходи дозволяють враховувати часову інформацію, що є критично важливим для обробки відео в реальному часі, забезпечуючи плавність та послідовність зображень.

Таким чином, сучасні методи машинного навчання, особливо глибокі нейронні мережі з механізмами уваги та контрастивною регуляризацією, демонструють значні переваги у відновленні зображень в умовах несприятливих погодних умов, забезпечуючи баланс між якістю відновлення та швидкістю. Цей аналіз показує, що класичні методи залишаються простими у впровадженні, але не витримують конкуренції за швидкістю та якістю з сучасними глибокими мережами. Гібридні архітектури, що поєднують фізичні знання з глибоким навчанням, дозволяють досягти оптимальних результатів у змінних погодних умовах.

1.4 Існуючі реалізації глибоких нейронних мереж для змішаних погодних умов

Нижче наведено огляд реальних реалізацій методів корекції атмосферних спотворень у зображеннях для різних погодних умов, із зазначенням їх ключових особливостей, переваг і недоліків.

TransWeather[11] – трансформерна архітектура для відновлення зображень, спотворених несприятливими погодними умовами. Вона використовує єдиний енкодер-декодер з внутрішньопатчовими трансформерами та вбудованими векторними представленнями типів погоди для адаптації до різних атмосферних умов. Переваги: універсальність для різних типів погодних спотворень (туман, дощ, сніг); ефективність завдяки єдиній архітектурі. Недоліки: може мати обмежену продуктивність на пристроях з обмеженими ресурсами через обчислювальну складність трансформерів.

PromptIR[12] – підхід до відновлення зображень, який використовує підказки (prompts) для кодування специфічної інформації про деградацію, що дозволяє моделі динамічно адаптуватися до різних типів і рівнів спотворень. Переваги: гнучкість у роботі з різними типами деградацій без потреби в попередньому знанні про них; ефективність завдяки легким підказкам. Недоліки: може вимагати ретельного налаштування підказок для оптимальної продуктивності.

NAFNet (Nonlinear Activation Free Network)[13] – модель відновлення зображень, яка досягає високої продуктивності без використання нелінійних функцій активації, таких як ReLU або GELU. Вона демонструє конкурентоспроможні результати з меншими обчислювальними витратами. Переваги: проста архітектура; висока ефективність; менші вимоги до обчислювальних ресурсів. Недоліки: може мати обмежену гнучкість у моделюванні складних нелінійних залежностей.

MPRNet (Multi-Stage Progressive Image Restoration Network)[14] – багатоступенева архітектура для прогресивного відновлення зображень. Вона поєднує енкодер-декодер для навчання контекстуалізованих ознак з високороздільною гілкою для збереження локальної інформації. Переваги: висока якість відновлення завдяки прогресивному підходу; ефективне поєднання глобальних і локальних ознак. Недоліки: складна архітектура; вищі вимоги до обчислювальних ресурсів.

FocalNet[15] – мережа для відновлення зображень, яка використовує механізм вибору в фокусі для акцентування на важливих областях зображення, таких як краї та складні регіони. Вона також інтегрує багатомасштабні рецептивні поля для покращення ефективності та продуктивності. Переваги: ефективне відновлення деталей у складних регіонах; покращена продуктивність завдяки багатомасштабному підходу. Недоліки: може бути складною для реалізації через інтеграцію кількох механізмів уваги.

WeatherDiffusion[16] – модель, яка використовує дифузійні процеси для відновлення зображень, спотворених погодними умовами. Вона застосовує патч-базовані дифузійні моделі для покращення якості зображень за різних погодних умов. Переваги: висока якість відновлення; здатність до генерації реалістичних деталей. Недоліки: високі обчислювальні витрати; складність у навчанні та реалізації.

TaNet (Triplet Attention Network for All-In-One Adverse Weather Image Restoration)[17] – архітектура із розгалуженими підмережами для одночасної оцінки карти передачі, освітлення та відновлення кадру, забезпечує пристосування

до різного рівня атмосферних розсіювачів. Переваги: чітке відділення зон із різним ступенем зашумленості та туману; збереження глобальної цілісності кольору. Недоліки: складна структура з кількома паралельними гілками; підвищена затримка через мультизадачність.

MWFormer (Multi-Weather Image Restoration Using Degradation-Aware Transformers)[18] – трансформерна модель для відновлення зображень, спотворених погодними умовами. Вона поєднує гіпермережу для витягування погодних ознак і адаптивну модифікацію параметрів основної мережі, що дозволяє моделі гнучко підлаштовуватись до різних чи комбінованих спотворень. Переваги: адаптивність без повторного навчання; висока якість відновлення завдяки FiLM-модулюванню. Недоліки: складна реалізація через використання гіпермережі; потреба в ресурсах для обробки зображень високої якості.

Моделі TANet та MWFormer були обрані для подальшого дослідження в дипломному проєкті через їхню здатність ефективно обробляти зображення, спотворені різними погодними умовами, та для порівняння їхньої швидкодії відносно якості, що є критично важливим для реальних застосувань у мобільних роботах. Далі будуть визначені вимоги до розробки підсистеми.

1.5. Формулювання конкретних вимог до розробки підсистеми

1.5.1. Загальні вимоги до розробки підсистеми

– реалізація на Python (версія не менше 3.8) із застосуванням одного з фреймворків глибинного навчання (TensorFlow 2.x або PyTorch 1.x) та OpenCV для обробки зображень;

– архітектура системи має бути модульною, містить окремі компоненти для завантаження даних та попередньої обробки, тренування та інференсу моделі, постобробки та інтерфейсу взаємодії.

1.5.2. Функціональні вимоги до розробки підсистеми

- вхідні формати підтримують статичні зображення (PNG, JPEG, BMP) та відеопотік (MP4, AVI, RTSP);
- попередня обробка кадрів включає зміну розміру до заданої роздільної здатності (224×224), нормалізацію значень пікселів;
- інфраструктура моделі використовує навчені ваги для генерації карти корекції атмосферних явищ для кожного кадру;
- постобробка включає застосування карти корекції до початкового зображення, відновлення колірного простору та контрастності;
- логування та моніторинг застосовується для збереження метрик оцінки, параметрів налаштування моделей, часу обробки кожного етапу та реєстрації помилок;
- режим тестування, що містить автоматичне обчислення PSNR та SSIM, використовуючи набір еталонних зображень;
- графічний інтерфейс з можливістю завантаження зображень та відео, налаштування параметрів корекції та перегляду результатів обробки через інтерфейс вебсторінки та у відповідній директорії.

1.5.3 Нефункціональні вимоги до розробки підсистеми

- продуктивність: обробка цифрових зображень на CPU Intel Core i5 (8 потоків) та обробка відеопотоку на GPU середнього рівня (NVIDIA GTX 1050Ti);
- стійкість: відсутність витоків пам'яті при безперервній роботі до 24 годин; відновлення після збоїв із мінімальними затримками;
- портативність: сумісність із вбудованими платформами (наприклад, NVIDIA Jetson Nano) без значних компромісів щодо продуктивності;
- безпека: обмеження доступу до файлової системи; валідація вхідних даних для запобігання виконання небезпечного коду;

					ІК11.120БАК.006 ПЗ	Арк.
						20
Зм.	Лист	№ докум.	Підпис	Дата		

– масштабованість: можливість розгортання кількох екземплярів сервісу для паралельної обробки різних відеопотоків.

Висновки до розділу 1

У цьому розділі було чітко сформульовано мету дослідження – покращення якості вхідного зображення шляхом усунення впливу погодних явищ за допомогою методів глибокого навчання. Об’єктом дослідження визначено цифрове відео, спотворене атмосферними умовами, а предметом – алгоритмічні методи його корекції з використанням нейромережевих моделей.

У межах опису задачі проведено аналіз характерних атмосферних артефактів, таких як туман, дощ і сніг. Було виокремлено основні складнощі, що виникають при обробці таких спотворень: нелінійність атмосферного впливу, наявність змішаних погодних умов у сцені та обмеження апаратних ресурсів для реального часу.

Розглянуто існуючі підходи до вирішення задачі, серед яких – класичні фізико-статистичні методи на основі моделей розсіювання та пріорів (наприклад, Dark Channel Prior), а також сучасні архітектури глибокого навчання з використанням механізмів уваги, багаторівневого вилучення ознак і контрастивної регуляризації.

Окрему увагу приділено огляду реальних реалізацій – таких, як DCP, DehazeNet, ODD-Net, C2PNet тощо. Визначено їхні сильні сторони (фізична обґрунтованість, здатність зберігати колірну структуру) та недоліки (артефакти на межах, затримки при обробці), що підкреслило необхідність універсального та ефективного рішення, придатного до інтеграції у вбудовані системи.

Формалізовано вимоги до розробки підсистеми, зокрема підтримку різних форматів відео та зображень, попередню обробку та аугментацію, застосування навчених моделей, постобробку результату, логування, обчислення метрик (PSNR, SSIM), а також забезпечення графічного інтерфейсу для взаємодії з користувачем. Усі вимоги узгоджено з метою досягнення адаптивності та модульності рішення.

					ІК11.120БАК.006 ПЗ	Арк.
						21
Зм.	Лист	№ докум.	Підпис	Дата		

На основі аналізу існуючих підходів обрано дві передові архітектури глибокого навчання – TaNet та MWFormer – для подальшого дослідження їх точності, швидкодії та придатності до використання в реальному часі у складних погодних умовах.

У результаті проведено системний аналіз проблеми, узагальнено сучасні підходи та реалізації, що дозволило сформулювати обґрунтоване технічне завдання для створення адаптивної підсистеми корекції атмосферних спотворень.

					ІК11.120БАК.006 ПЗ	Арк.
						22
Зм.	Лист	№ докум.	Підпис	Дата		

2 ІНФОРМАЦІЙНЕ ЗАБЕЗПЕЧЕННЯ

2.1 Модель даних та архітектурні схеми підсистеми

Для забезпечення ефективної роботи підсистеми комп'ютерного зору з покращення якості зображень за умов складних погодних явищ розроблено модель даних і архітектуру підсистеми, що відображають логіку обробки, обмін інформацією між компонентами та функціональні етапи обробки зображення або відеопотоку.

2.1.1 Блок схема обробки відеоданих у підсистемі комп'ютерного зору

Блок схема, що представлена на кресленику ІК11.120БАК.006 Д1 відображає послідовність обробки одного кадру від моменту надходження до його передачі далі у конвеєр відеоаналітики.

а) джерело відео: надходження зображень або відеопотоку з камери, відеофайлів;

б) завантаження: вибір джерела кадру, буферизація, перевірка формату зображення;

в) попередня обробка: корекція розміру, нормалізація інтенсивності, фільтрація шуму;

г) інференс нейромережі: завантаження ваг (TANet або MWFormer), обчислення карти корекції для кожного пікселя;

д) постобробка: накладання карти на оригінал, відновлення контрасту та колірної гами;

е) передача результату: оновлений кадр направляється далі до модулів детекції/сегментації або на вивід користувача, виконується логування результатів, оновлення UI.

2.1.2 Діаграма прецедентів підсистеми комп'ютерного зору

На кресленику ІК11.120БАК.006 Д2 зображено діаграму прецедентів підсистеми комп'ютерного зору, яка демонструє взаємодію між зовнішніми учасниками (актором-оператором та підсистемою) і функціональними можливостями, які надає система.

Опис акторів:

– оператор є безпосереднім користувачем, який взаємодіє з інтерфейсом для завантаження даних, налаштування моделі, ініціалізації обробки та перегляду результатів;

– підсистема виконує роль внутрішнього або зовнішнього модулю, що може ініціювати, логувати або аналізувати процес обробки автоматично.

Основні прецеденти (використання функціоналу):

а) завантажити вхідні дані: оператор надає відео або зображення, які потребують обробки;

б) налаштувати параметри моделі: налаштовуються гіперпараметри або вибирається модель (TANet чи MWFormer);

в) ініціювати обробку: запуск процесу обробки відповідно до заданих параметрів;

г) виконати покадрову обробку відео / зображення: основна функція, яка усуває атмосферні артефакти;

д) здійснити логування параметрів: зберігаються параметри запуску та проміжні результати;

е) оцінити якість результату: проводиться автоматична або ручна оцінка на основі метрик PSNR та SSIM;

ж) зберегти результат: результати обробки записуються у відповідному форматі;

и) переглянути результат: оператор може переглянути очищене зображення або відео;

					ІК11.120БАК.006 ПЗ	Арк.
						24
Зм.	Лист	№ докум.	Підпис	Дата		

к) відобразити результат у інтерфейсі: результат показується в графічному або текстовому вигляді;

л) Вивести звіт у консоль: система може виводити метрики або технічну інформацію про обробку.

Ця діаграма дозволяє наочно уявити загальну структуру підсистеми, описати межі її функціоналу та виявити ключові взаємодії між компонентами. Вона є фундаментом для подальшого проєктування та реалізації програмної архітектури.

2.1.3 Блок схема алгоритму підсистеми комп'ютерного зору

Блок схема алгоритму підсистеми, що зображена на кресленику ІК11.120БАК.006 ДЗ, ілюструє покрокову обробку вхідних зображень або відео для усунення атмосферних завад із використанням нейромереж TANet або MWFormer.

Робота підсистеми починається з вибору моделі глибокого навчання (TANet або MWFormer), після чого користувач задає шлях до вхідних даних, які можуть бути представлені у вигляді окремого зображення або відеофайлу. Якщо вхідним є відео, відбувається його буферизація та розбиття на окремі кадри.

Далі система здійснює перевірку цілісності завантажених даних, після чого виконується масштабування до необхідного розміру, нормалізація піксельних значень та перетворення вхідного зображення до тензора.

Отриманий тензор подається на вхід обраної нейромережі для обробки з метою усунення атмосферних завад (туман, дощ, сніг тощо). В результаті мережа формує прогноз очищеного зображення.

Далі система виконує перетворення тензора назад у формат зображення, після чого результат може бути збережено у файл, а у випадку обробки відео — відтворено у форматі відеоролика.

При наявності еталонних (GT) зображень проводиться порівняння отриманого результату з GT-даними. Паралельно система виконує облік часу інференсу, що дозволяє оцінити середній час обробки кадру.

На завершальному етапі підсистема виводить результати у консоль або зберігає їх у зазначену директорію, надаючи користувачу шлях до збережених зображень або відео.

Таким чином, блок-схема демонструє повну послідовність дій від завантаження вхідних даних до отримання очищеного результату, включаючи попередню та постобробку, логування часу роботи і збереження результатів у зручному форматі.

2.1.4 Структурна схема підсистеми комп'ютерного зору

Структурна схема демонструє взаємодію підсистеми покращення зображення з іншими компонентами в рамках більш широкої системи комп'ютерного зору. У даній роботі розглядаються два типових сценарії використання підсистеми, що представлені на кресленику ІК11.120БАК.006 Д4, – як частини системи реального часу (наприклад, в дроні), та як етапу постобробки в аналітичних системах (наприклад, моніторинг безпеки чи дорожнього трафіку).

У першому варіанті підсистема інтегрується безпосередньо в обчислювальний модуль дрона й працює як частина конвеєра обробки відеопотоку в реальному часі. Основна мета – зменшити вплив погодних артефактів (туман, сніг, дощ) перед передачею кадру до модуля аналізу (детекція, трекінг об'єктів, планування маршруту).

Структура взаємодії:

- а) модуль захоплення зображення – передає кадри в потоці;
- б) підсистема покращення зображення – обробляє кожен кадр, усуваючи атмосферні спотворення;
- в) модуль детекції об'єктів – отримує покращені зображення для точнішого розпізнавання;
- г) модуль навігації дрона – використовує результати детекції для ухвалення рішень;

д) система зберігання/логування – зберігає результати або відправляє на сервер для додаткового аналізу.

Особливості:

- обмежені ресурси (CPU/GPU);
- потреба в низькій затримці;
- адаптація до змішаних погодних умов у реальному часі.

У другому варіанті підсистема використовується для офлайн-обробки відеозаписів, які містять зображення, погіршені погодними умовами. Такий варіант застосовується в системах безпеки, моніторингу доріг, екологічного спостереження. Основна мета – покращити якість відео для аналітиків або для подальшої обробки комп’ютерним зором.

Структура взаємодії:

а) модуль завантаження відео з архіву – читає заздалегідь записаний матеріал;

б) підсистема покращення зображення – обробляє кадри послідовно (або пакетно);

в) модуль аналітики/детекції – аналізує відновлене зображення;

г) інтерфейс користувача – дозволяє оператору порівняти «до/після», позначити події, виявити об’єкти;

д) зберігання результатів – запис поліпшених відео та аналітики в базу даних.

Особливості:

- менш жорсткі вимоги до продуктивності;
- можливість використання більш важких моделей;
- краща якість зображень для ретроспективного аналізу.

Організаційна схема показує, що підсистема покращення зображення є гнучким компонентом, здатним інтегруватися в різні сценарії використання систем комп’ютерного зору. Вона може працювати як у реальному часі, забезпечуючи стабільну роботу автономних систем у складних умовах, так і в режимі постобробки, підвищуючи якість даних для подальшого аналізу в інших

предметних областях. Далі більш детально будуть розглянуті використані алгоритми глибоких нейронних мереж.

2.2 Алгоритмічні складові підсистеми комп'ютерного зору

У цьому підрозділі детально описуються ключові алгоритмічні компоненти обраних моделей TaNet та MWFormer, які лягають в основу підсистеми покращення якості зображень за умов складних погодних явищ.

2.2.1 Triplet Attention Network (TANet)

Triplet Attention Network (TANet) – це сучасна нейронна мережа, розроблена для одночасного усунення атмосферних спотворень, таких як туман, дощ і сніг, зображень у єдиній архітектурі. Модель поєднує три типи механізмів уваги в одному блоці Triplet Attention Block (TAB), що дозволяє ефективно відновлювати зображення, спотворені різними погодними умовами.

TANet реалізована як згорткова мережа типу енкодер-декодер. На початку використовуються два шари Feature Embedding Layers (FEL) для зменшення просторової роздільності вхідних ознак. Після цього слідує стек блоків Triplet Attention Block (TAB), кожен з яких включає три паралельні гілки уваги.

– Local Pixel-wise Attention (LPA). Цей механізм фокусується на локальних просторових залежностях, дозволяючи моделі ефективно обробляти короткодистанційні спотворення, спричинені атмосферними умовами;

– Global Strip-wise Attention (GSA). GSA включає горизонтальні та вертикальні стрічкові механізми уваги, що дозволяє моделі захоплювати глобальні просторові залежності та ефективно обробляти довгі лінійні артефакти, характерні для дощу та снігу;

– Global Distribution Attention (GDA). Цей механізм виконує нормалізацію ознак та зважає їх відповідно до розподілу атмосферних частинок, що дозволяє враховувати варіації контрасту та кольору в усьому зображенні.

					ІК11.120БАК.006 ПЗ	Арк.
						28
Зм.	Лист	№ докум.	Підпис	Дата		

Після обробки в TAB, ознаки конкатенуються та проходять через додатковий згортковий шар з залишковим з'єднанням, формуючи багатомасштабну карту ознак. На завершальному етапі два додаткові FEL шари використовуються для відновлення просторової роздільності та реконструкції очищеного зображення.

Для тренування моделі TANet було використано три синтетичні набори даних, кожен з яких відповідає певному типу атмосферного спотворення:

а) RESIDE (для усунення туману) включає:

- піднабір ITS із 110 000 пар зображень (внутрішні сцени);
- піднабір OTS із 313 950 пар (зовнішні сцени).

б) Rain1400 (для усунення дощу) містить 12 600 пар зображень;

в) Snow100K (для усунення снігу) включає 100 000 пар.

Щоб забезпечити збалансоване навчання, з кожного набору було рівномірно вибрано по 5 000 пар, які об'єднано в єдиний навчальний датасет обсягом 15 000 пар. На цій змішаній вибірці було оптимізовано TANet, а також усі порівнювані методи, що дозволяє провести коректну оцінку якості.

Для тестування ефективності моделі застосовано як синтетичні, так і реальні набори даних:

а) синтетичні тестові датасети:

- SOTS з 500 парами для внутрішніх і 500 для зовнішніх зображень;
- Rain1400 містить 1 400 пар;
- Snow100K-L містить 16 801 пару.

б) реальний тестовий набір WeatherStream:

- 4 500 зображень для усунення туману;
- 4 800 для дощу;
- 3 960 для снігу.

Таким чином, TANet була навчена на різноманітних погодних умовах і протестована як на штучно згенерованих, так і на реальних даних, що підтверджує її універсальність та здатність до генералізації.

2.2.2 Multi Weather Transformer (MWFormer)

MWFormer (Multi-Weather Image Restoration Using Degradation-Aware Transformers) – це сучасна трансформерна архітектура, розроблена для відновлення зображень, спотворених різними погодними умовами, такими як туман, дощ, сніг або їх комбінації. Модель поєднує гіпермережі та механізм модулювання параметрів для адаптивної обробки зображень без необхідності повторного навчання.

MWFormer складається з двох основних компонентів:

- гіпермережа для виділення погодних ознак. Ця легка мережа, яка може бути реалізована на основі CNN або трансформера, навчається за допомогою контрастивного навчання для виділення ознак, незалежних від вмісту, що кодують інформацію про типи погодних умов на зображенні.

- основна мережа відновлення зображень. Цей компонент включає енкодер, декодер і додаткові згорткові шари. Параметри певних шарів цієї мережі модулюються за допомогою механізму Feature-Wise Linear Modulation (FiLM), використовуючи вектор ознак, отриманий з гіпермережі. Це дозволяє адаптувати обробку до конкретних погодних умов.

Модель MWFormer тренувалася на стандартному мультипогодному бенчмарку, що поєднує три основні набори даних:

- RainDrop для задачі видалення крапель дощу з зображень;
- Outdoor-Rain для задач комбінованого усунення дощу та туману;
- Snow100K для видалення снігових артефактів.

Для тестування було використано відповідні тестові підмножини кожного датасету:

- RainDrop test set для перевірки ефективності видалення крапель;
- Test1 з Outdoor-Rain для оцінки продуктивності у складних погодних умовах (дощ і туман);
- Snow100K-L для перевірки відновлення зображень, спотворених снігом.

Такий підхід дозволив об'єктивно оцінити якість роботи MWFormer у завданнях відновлення зображень під впливом різних атмосферних умов.

Ключові компоненти:

- Вектор ознак погоди ($z \in \mathbb{R}^d$). Вектор, що інформує про поточні погодні умови, може містити інформацію про декілька типів одночасно;
- Feature-Wise Linear Modulation (FiLM). Механізм, який обчислює коефіцієнти γ та β для кожного шару, модулюючи вихід x як $\gamma \odot x + \beta$, де $\odot$ означає поелементне множення;
- Intra-Patch Transformer Blocks. Блоки, що обробляють дрібні підпатчі зображення, покращуючи здатність мережі враховувати дрібні артефакти, такі як краплі дощу чи сніжинки.

Режими роботи:

- Single Weather Mode. Використання фіксованого вектора ознак для відновлення зображень, спотворених одним типом погоди;
- Hybrid Weather Mode. Послідовна обробка зображення через MWFormer для усунення декількох типів спотворень без необхідності додаткового навчання.

Застосування TANet з її потужним Triplet Attention Block та MWFormer з гіпермережею і FiLM модулюванням створює гнучку та ефективну підсистему, здатну адаптуватися до широкого спектру атмосферних спотворень як у реальному часі, так і в режимі постобробки.

2.3 Набори даних

Для навчання та тестування моделей TANet і MWFormer були використані як синтетичні, так і реальний набір даних, що охоплюють різні погодні умови, зокрема сніг, дощ і туман. Нижче наведено опис кожного з них.

2.3.1 Набір даних SnowCityScapes

Цей набір даних створено на основі Cityscapes – великого датасету міських сцен із високоякісною розміткою. Використовується для задач видалення снігу. SnowCityScapes включає 2 000 пар зображень для навчання та 2 000 для тестування, кожне з яких має роздільну здатність 512×256 пікселів. Сніг накладається синтетично в трьох варіантах інтенсивності: легкий, середній та сильний. Це дозволяє моделювати різні погодні умови та забезпечує різноманітність для навчання моделей. У додатку Д, на рисунку Д.1, наведені приклади з тренувального набору даних, де є зображення з ефектом снігу та оригінальні.

2.3.2 Набір даних Rain1400

Rain1400 – синтетичний датасет, що застосовується для задач видалення дощу, створений шляхом накладання дощових смуг на 1 000 чистих зображень, відібраних з UCID, BSD та пошуку Google. Кожне зображення використано для генерації 14 варіантів із різною інтенсивністю та напрямком дощу, що в результаті дало 14 000 пар зображень. З них 12 600 пар використовуються для навчання, а 1 400 – для тестування. Цей набір даних широко застосовується для оцінки алгоритмів видалення дощу. У додатку Д, на рисунках Д.2–Д.3, наведені приклади з тренувального набору даних, де є зображення з ефектом дощу та оригінальні.

2.3.3 Набір даних RESIDE-6K

RESIDE-6K є підмножиною великого набору даних RESIDE (REalistic Single Image Dehazing) для задач видалення туману. Він містить 6 000 пар зображень для навчання та 1 000 для тестування, що охоплюють як сцени всередині приміщень, так і просто неба. Кожна пара складається з зображення з туманом та відповідного чистого зображення. Цей набір даних широко використовується для навчання та оцінки алгоритмів видалення туману. У додатку Д, на рисунках Д.4–Д.5, наведені

					ІК11.120БАК.006 ПЗ	Арк.
						32
Зм.	Лист	№ докум.	Підпис	Дата		

прикладі з тренувального набору даних, де є зображення з ефектом туману та оригінальні.

2.3.4 WeatherStream Dataset

Було використано також реальний набір даних WeatherStream, який є найбільшим на сьогодні датасетом для задачі видалення погодних спотворень з зображень. Його головна особливість – автоматизоване формування пар зображень із погодними ефектами (дощ, сніг, дощовий туман) та відповідних чистих версій цих же сцен. Це дозволяє уникнути обмежень симуляцій та ручного маркування, що характерні для попередніх підходів.

WeatherStream створено завдяки автоматизованому пайплайну, який базується на принципах світлотранспорту (light transport physics). Алгоритм автоматично фільтрує понад 99,6% непридатних відео і залишає лише ті, де присутні ідеальні умови для формування пар "чисте–зі спотворенням". Для цього застосовуються чотири основні критерії:

- відповідність фону;
- хроматичні варіації частинок (дощу, снігу);
- розмиття, спричинене розсіюванням;
- постійність освітлення.

У наборі даних наявні різноманітні зображення реальних сцен міст, вулиць, природних ландшафтів, тощо. Враховано різні погодні умови, рівні освітлення, типи камер та параметри зйомки. Джерело відео охоплює понад 5000 потоків YouTube Live, з яких зібрано понад 2 млн кандидатних відео, але лише незначний відсоток із них відповідає строгим вимогам. У додатку Д, на рисунках Д.6–Д.7, наведені приклади з тренувального набору даних, де є зображення з погодними ефектами та відповідні без.

Замість симуляцій або ручного підбору кадрів, як у попередніх роботах, WeatherStream використовує тайм-мультиплексування – відбір кадрів із відеопотоку в моменти, коли погодний ефект зникає або з’являється, без зміни

сцени чи освітлення. Це дозволяє отримувати майже ідеальні пари, близькі до справжніх "ground truth" зображень.

WeatherStream значно перевершує існуючі синтетичні датасети завдяки своїй реалістичності, масштабності та автоматизованій побудові. Його використання забезпечує кращу адаптацію моделей до реального середовища та дає змогу уникнути розриву між симульованими й реальними даними.

Усі зазначені набори даних є відкритими та доступні для завантаження, що сприяє відтворюваності експериментів та подальшим дослідженням у галузі комп'ютерного зору.

Висновки до розділу 2

У даному розділі було всебічно розкрито інформаційне забезпечення підсистеми мобільного робота для покращення зображень, зосереджене на завданнях комп'ютерного зору в умовах несприятливих погодних впливів. Розглянуті моделі, структурні та функціональні схеми, а також алгоритмічні та організаційні компоненти утворюють цілісну архітектуру, яка дозволяє ефективно вирішувати поставлені задачі в режимах реального часу або офлайн-обробки.

Створені схеми чітко відображають взаємозв'язки між основними модулями системи: від отримання вхідного сигналу до передачі результатів на подальші етапи обробки або взаємодії з користувачем. Інтеграція передових моделей глибокого навчання, зокрема TANet та MWFormer, дозволяє адаптувати алгоритмічне ядро під різні сценарії застосування.

Організаційна структура демонструє гнучкість системи в різних контекстах використання – як у автономних системах (наприклад, на дронах), так і в аналітичних комплексах відеоспостереження. Це підтверджує потенціал підсистеми до масштабування та інтеграції у складні програмно-апаратні рішення.

Таким чином, сформована інформаційна база є фундаментом для подальшої реалізації, тестування та адаптації системи до прикладних задач у сфері комп'ютерного зору.

					ІК11.120БАК.006 ПЗ	Арк.
						34
Зм.	Лист	№ докум.	Підпис	Дата		

3 ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ

3.1 Вибір інструментів та технологій

У рамках реалізації підсистеми комп'ютерного зору було створено спеціалізоване програмне забезпечення, призначене для підвищення якості візуальних даних, які зазнали погіршення через атмосферні умови, такі як туман, опади або низький контраст. Основний фокус полягав у поєднанні високої точності обробки з придатністю до інтеграції у прикладні системи технічного зору – як в автономних пристроях, так і в стаціонарних аналітичних комплексах.

Розробка підсистеми базувалася на актуальних технологічних рішеннях, адаптованих під специфіку завдання:

- мова програмування. Обрано Python версії 3.10 як основну мову завдяки її зручному синтаксису, великій спільноті та широкому набору бібліотек для комп'ютерного зору, обробки зображень і глибокого навчання;

- фреймворк глибокого навчання. Для реалізації нейронних мереж використовувався PyTorch версії 1.13, що забезпечує гнучкість при побудові індивідуальних архітектур, а також ефективну GPU-акселерацію, критично важливу для обробки даних у реальному часі;

- графічний інтерфейс користувача. Інтерфейс створено за допомогою бібліотеки Gradio. Вона забезпечує простий спосіб розгортання веб-інтерфейсу для демонстрації та взаємодії з моделями машинного навчання, дозволяючи користувачам легко завантажувати зображення або відео, змінювати параметри обробки та переглядати результати прямо в браузері;

- системи керування середовищем та версіями. Для керування залежностями використовувалася платформа Anaconda. Контроль версій і командна робота були організовані за допомогою системи Git, що забезпечує відстеження змін і зручну інтеграцію з хостинговими сервісами;

- інструменти логування й моніторингу. У процес навчання та тестування моделей було інтегровано платформу Weights & Biases (wandb). Вона надала можливість відстежувати ключові метрики якості (SSIM, PSNR), візуалізувати

процес тренування, порівнювати експериментальні запуски, контролювати динаміку функції втрат і зберігати параметри для подальшого аналізу. Це значно спростило вибір оптимальних налаштувань моделі та забезпечило прозорість дослідницького процесу.

Було реалізовано підтримку двох моделей глибокого навчання – TANet і MWFormer. Ці архітектури забезпечують якісну реконструкцію зображень при наявності атмосферних перешкод. TANet більш ефективна в умовах розсіювання та при невеликих наявних ресурсах обробки, тоді як MWFormer демонструє стабільність при комбінованих впливах, зберігаючи деталізацію.

3.2 Використані апаратні ресурси

Для реалізації експериментальної частини роботи було використано два типи апаратних ресурсів: віддалений сервер для тренування моделей та локальний комп'ютер для їх тестування і демонстрації.

Серверна частина (тренування та валідація моделей). Для виконання ресурсоємних операцій навчання нейронних мереж та валідації результатів за наявності еталонних (ground truth) зображень використовувалася відеокарта NVIDIA GeForce GTX 1080 Ti. Цей графічний процесор має 11 ГБ відеопам'яті GDDR5X, 3584 CUDA-ядер та забезпечує високу продуктивність при паралельних обчисленнях, що є критично важливим для глибокого навчання. Тренування моделей проводилося з використанням фреймворків PyTorch та підтримкою апаратного прискорення CUDA.

Локальна частина (тестування та демонстрація). Для тестування моделей у режимі без наявності ground truth (тобто на реальних або нових зображеннях) та для запуску демонстрації моделі у вебінтерфейсі використовувався персональний комп'ютер на базі процесора Intel Core i5. Оскільки на цьому етапі не відбувається навчання, обчислювальні вимоги значно нижчі. Для взаємодії користувача з моделлю використовувався фреймворк Gradio, що дозволяє швидко створити

інтуїтивний вебінтерфейс для завантаження зображень, перегляду результатів та взаємодії з нейронною мережею.

Таке розділення ресурсів дало змогу оптимально поєднати високу продуктивність під час навчання моделей та зручність локального тестування та демонстрації.

3.3 Опис програмного продукту

Створене програмне забезпечення являє собою десктопну програму з графічним інтерфейсом, яка дозволяє:

- завантажувати окремі зображення або відеопотік (послідовність кадрів);
- обрати один із запропонованих методів обробки (TANet або MWFormer);
- переглядати результат в режимі порівняння (до/після);
- зберігати оброблені зображення;
- аналізувати метрики якості (PSNR, SSIM, швидкодія).

Програмне забезпечення не потребує зовнішніх API або мережових підключень, та може працювати у режимі офлайн-обробки окремих кадрів або всього поданого відеопотоку. Основна логіка реалізована як окрема програма, що взаємодіє з обчислювальними модулями моделей через внутрішні виклики Python. Для інтеграції у сторонні системи передбачено підтримку імпорту бібліотеки як модуля Python (наприклад, для розширення у вбудованих системах чи робототехніці).

Графічний інтерфейс включає наступні основні елементи:

- поле завантаження зображень: із підтримкою drag-and-drop або вибору з файлової системи;
- меню вибору моделі обробки;
- область попереднього перегляду результату (розділене зображення до/після);
- інформаційне поле, що відображає час інференсу;
- системне повідомлення про прогрес обробки або помилки.

Інтерфейс виконано у стандартному стилі з акцентами на зручність для користувача: великі кнопки, читабельні шрифти, автоматична адаптація до роздільної здатності екрана.

3.4 Структура програмного забезпечення підсистеми комп'ютерного зору

Всю архітектуру програмного забезпечення даного проєкту можна описати розподілом на модулі, що зображений на рисунку 3.1.

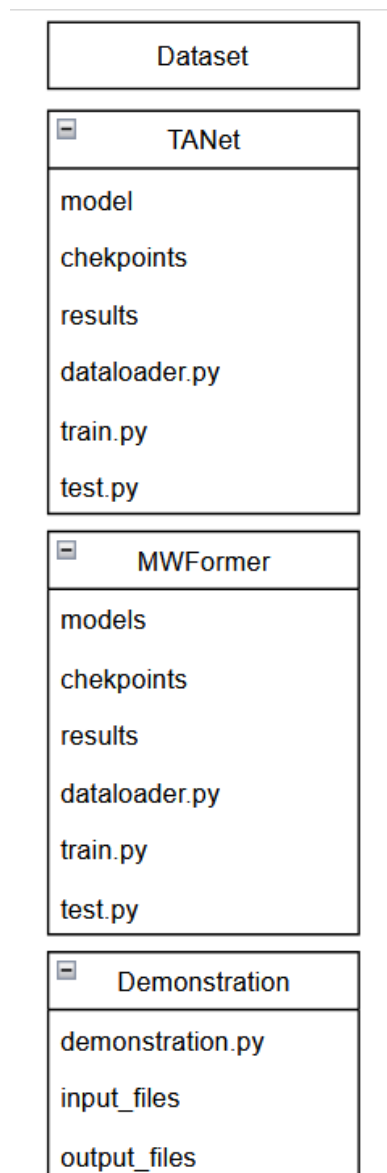


Рисунок 3.1 – Модульна структура програмного забезпечення

Перше, це набір даних (Dataset), що застосовується для тренування та валідації наявних моделей TANet та MWFormer. Він у свою чергу містить набір синтетичних даних, що складається із зображень з туманом, снігом, дощем, взятих з відкритих наборів даних RESIDE-6K, SnowCityScapes, Rain1400 відповідно. Також дана директорія містить реальний набір даних WeatherStream, Кожний з цих двох великих датасетів розбитий на папки для тренування та валідації, що у свою чергу розподілені відповідно до різних ефектів погодних явищ. Кожне зображення з наявним погодним ефектом має відповідне чисте зображення. Організацію даної директорії можна бачити на рисунку 3.2.

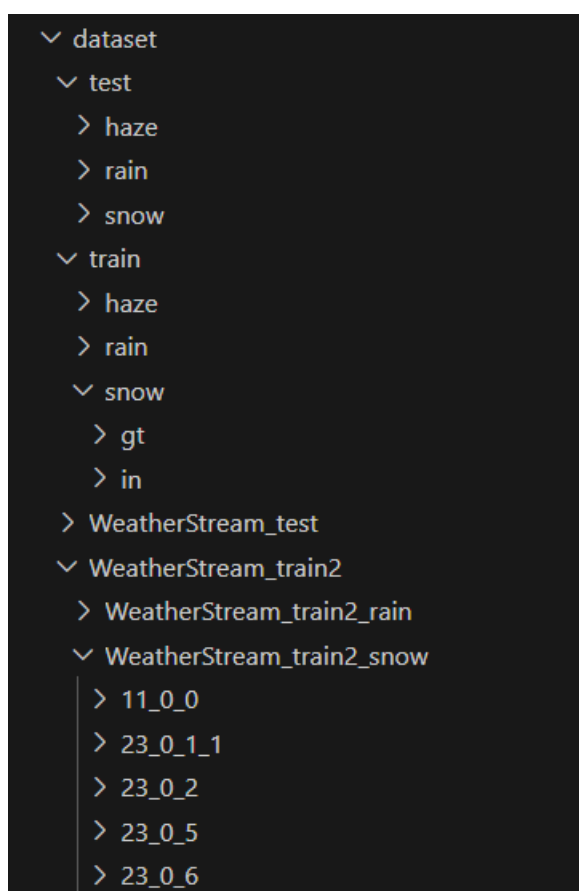


Рисунок 3.2 – Структура організації датасету

У наборі даних WeatherStream кожна директорія train або test містить ще папки. Кожна така папка являє собою окрему локацію, де були зафільмовані зображення.

Наступний модуль описує логіку роботи для моделі глибокого навчання TANet та має наступний вміст:

					ІК11.120БАК.006 ПЗ	Арк.
						39
Зм.	Лист	№ докум.	Підпис	Дата		

– model містить опис архітектури нейронної мережі TANet. У цьому каталозі зосереджено всі необхідні класи та функції, що реалізують структуру згорткових блоків, attention-механізмів та інших елементів моделі;

– checkpoints зберігає проміжні та фінальні ваги моделі, отримані після навчання при різних кількостях епох або на різних підмножинах даних (синтетичних чи реальних);

– у папку results автоматично зберігаються результати тестування у вигляді зображень: як вихід моделі, так і GT для візуального порівняння;

– dataloader.py містить усі необхідні функції попередньої обробки зображень та їх коректного завантаження у форматі, зручному для подачі до нейронної мережі;

– скрипт train.py є основним для навчання моделі. Реалізує цикл епох, функції втрат, логування процесу, а також збереження ваг моделі;

– test.py використовується для оцінювання моделі на тестовій вибірці. Завантажує зображення, передає їх через модель та зберігає результати.

Уся логіка реалізована у фреймворку PyTorch, що забезпечує гнучкість у зміні архітектури, ефективність тренування на GPU та легку інтеграцію з іншими модулями.

Модуль для моделі MWFormer має аналогічну структуру з відповідними папками model, checkpoints, results, а також скриптами train.py і test.py. Проте, через особливості архітектури MWFormer (наявність декількох блоків attention-перетворень та специфічної попередньої обробки), реалізація деяких компонентів, зокрема dataloader.py та внутрішніх частин моделі, відрізняється. Деталі реалізації цієї моделі описано у наступних розділах роботи.

Модуль демонстрації складається з наступних елементів:

– скрипт demonstration.py для запуску демонстраційного графічного інтерфейсу, реалізованого з використанням бібліотеки Gradio. Дозволяє користувачу завантажити нове зображення, обрати модель (TANet або MWFormer), виконати обробку та переглянути результат;

– папка `input_files` для збереження зображень, які завантажуються користувачем через інтерфейс;

– `output_files` містить результати обробки відповідних вхідних зображень.

Цей модуль забезпечує зручне тестування моделей на нових реальних даних без `ground truth`, що особливо корисно для практичної перевірки якості роботи алгоритмів у побутових умовах.

Повний код програмного забезпечення наведений у додатку Г на рисунку Г.1, що містить посилання на гітхаб.

3.5 Обробка даних для циклу тренування

3.5.1 Попередня обробка даних для моделі TANet

Для підготовки вхідних зображень перед навчанням моделі TANet використовувалась спеціально реалізована послідовність перетворень, метою якої є покращення узагальнюючої здатності моделі за рахунок аугментації даних. Ці перетворення реалізовано у вигляді окремих класів із перевизначеним методом `__call__()`, які послідовно застосовуються до пари зображень: вхідного (`input`) та відповідного еталонного (`gt`).

Основні етапи попередньої обробки:

а) випадкове обертання (`RandomRotate`). Зображення обертається на 0° , 90° , 180° або 270° градусів. Напрямок обертання обирається випадково, що дозволяє моделі побачити сцени під різними кутами та покращує її інваріантність до поворотів;

б) випадкове дзеркальне відображення (`RandomFlip`). Існує ймовірність, що зображення буде дзеркально відображено по горизонталі та/або вертикалі. Це підвищує стійкість моделі до змін орієнтації об'єктів у кадрі;

в) випадкове вирізання області (`RandomCrop`). Зі зображення вирізається випадкова підобласть фіксованого розміру (224×224). Це дозволяє зменшити вплив на навчання фонових або повторюваних частин зображення;

г) нормалізація (Normalize). Піксельні значення переводяться з діапазону [0, 255] у [-0.5, 0.5]. Це покращує стабільність і швидкість навчання нейронної мережі;

д) перетворення у тензори (ToTensor). Кожне зображення перетворюється у тензор PyTorch формату $C \times H \times W$ (канали, висота, ширина), що є стандартним представленням зображень у PyTorch.

Загальна послідовність перетворень реалізована за допомогою конструкції `transforms.Compose`, яка послідовно застосовує вказані кроки.

Завдяки такій системі попередньої обробки зображень, модель TANet тренується на більш різноманітному наборі вхідних даних, що підвищує її здатність до узагальнення та адаптації до нових умов освітлення, орієнтацій і ракурсів сцени.

3.5.2 Попередня обробка даних для моделі MWFormer

У моделі MWFormer попередня обробка зображень виконується без використання агресивних аугментацій, але з дотриманням суворих геометричних і колірних перетворень для забезпечення відповідності між вхідними та еталонними зображеннями. Вона включає наступні ключові етапи:

а) завантаження зображень. Із вказаного списку файлів завантажуються зображення з поганими погодними умовами (input) та відповідні чисті зображення (gt) за допомогою бібліотеки PIL (`Image.open()`);

б) масштабування за потребою. Якщо розміри оригінального зображення менші за розмір обрізки (`crop_width`, `crop_height`), обидва зображення (вхідне та GT) масштабуються до мінімально необхідного розміру з використанням згладжування (`Image.ANTIALIAS`). Це дозволяє уникнути помилок під час подальшого вирізання області;

в) випадкове вирізання області (Random Crop). Зображення обрізаються до фіксованого розміру з випадковим вибором координат верхнього лівого кута (x, y). Цей етап забезпечує варіативність навчальних прикладів, а також зменшує вплив фону або статичних областей. Вирізання здійснюється одночасно ідентично для вхідного та GT-зображення, щоб зберегти відповідність;

г) перетворення у тензори (ToTensor). Зображення переводяться у формат тензорів PyTorch, де канали переносяться на перший вимір ($C \times H \times W$), що є необхідною вимогою для більшості нейронних мереж;

д) нормалізація (Normalize). До вхідного зображення застосовується нормалізація піксельних значень за каналами з використанням середнього значення 0.5 та стандартного відхилення 0.5 для кожного каналу RGB. Це переводить значення до діапазону $[-1,1]$, що покращує стабільність і швидкість тренування. GT-зображення при цьому не нормалізується (лишається у діапазоні $[0,1]$).

Цей пайплайн обробки зображень дозволяє моделі MWFormer ефективно працювати з парами зображень навіть у випадках, коли вони мають різні розміри або розташування важливих деталей. Водночас, відсутність інтенсивної аугментації свідчить про орієнтацію моделі на точну реконструкцію деталей без суттєвого шуму або спотворень. Тепер можна перейти до циклу тренування моделей.

3.6 Опис циклу підготовки моделей машинного навчання

У межах даного дипломного проєкту обидві моделі машинного навчання – TANet та MWFormer – були підготовлені у двох окремих незалежних циклах навчання: на синтетичних та на реальних зображеннях.

Перший цикл передбачав тренування моделей на синтетично згенерованих парах зображень, де зображення з погодними артефактами були створені штучно на основі чистих сцен. Це дозволяло забезпечити контроль над вмістом зображень, а також точно відповідні ground-truth пари. Загалом було використано:

- 6300 пар з ефектами синтетичного дощу;
- 6000 пар з ефектами синтетичного туману;
- 6000 пар з ефектами синтетичного снігу.

Кожен тип погодного ефекту моделювався окремо, що дало змогу дослідити вплив кожного виду викривлення на роботу моделей.

Другий цикл навчання був виконаний окремо – на реальних зображеннях, отриманих із датасету WeatherStream. Було використано:

- 6000 пар з реальним снігом;
- 12000 пар з реальним дощем та туманом.

Тренування на таких даних дозволяло оцінити реальну здатність моделей до узагальнення на непередбачувані погодні умови та обробку натуральних сцен з усіма їхніми складнощами: варіативність об’єктів, освітлення, динаміка та шум.

3.6.1 Цикл тренування TANet

Початково визначаються шляхи до моделі, даних і параметри експерименту: розмір обрізки зображення (`crop_size = 224`), розмір батчу (`batch_size = 16`), кількість епох (`end_epoch = 50`), початкова та мінімальна швидкість навчання (`init_lr = 1e-4`, `min_lr = 1e-7`), а також частота збереження контрольних точок. Для збереження відтворюваності результатів було зафіксовано генератори випадкових чисел (`random`, `numpy`, `torch`). Навчання здійснювалось із можливістю продовження з останньої збереженої епохи.

Для підготовки даних використовується клас `Train_Loader`, що здійснює:

- зчитування зображень та GT;
- аугментацію (обрізання, фліпи, повороти);
- нормалізацію та перетворення у тензори.

Дані завантажуються за допомогою `torch.utils.data`, який відповідає за ефективну подачу тренувальних даних до моделі. Цей об’єкт формує батчі з набору `Train_set`, який містить зображення разом із відповідними мітками (`ground truth`) і реалізує логіку попередньої обробки кожного зразка. Параметр `batch_size` визначає кількість зображень у кожному батчі, що дозволяє ефективно використовувати ресурси GPU. Встановлення `shuffle=True` забезпечує випадкове перемішування зразків перед кожною епохою, що зменшує ризик перенавчання. Завдяки параметру `num_workers=cpu_count()` дані завантажуються паралельно за допомогою всіх доступних ядер CPU, що значно пришвидшує процес обробки. Нарешті,

					ІК11.120БАК.006 ПЗ	Арк.
						44
Зм.	Лист	№ докум.	Підпис	Дата		

drop_last=False означає, що навіть неповний батч наприкінці епохи не буде відкинутий, що дозволяє використати всі наявні зразки.

Модель TANet ініціалізується та обгортається в DataParallel для використання кількох GPU (у разі наявності). Оптимізатор – Adam, а планувальник швидкості навчання – CosineAnnealingLR, що поступово зменшує lr до мінімального значення.

Якщо у папці результатів присутній файл збереженого стану моделі (last_*.pth), він автоматично завантажується: ваги моделі, стан оптимізатора та планувальника lr.

Навчання відбувається в межах заданої кількості епох. Для кожного батчу дані проходять такі етапи:

а) зображення та відповідні еталонні дані (GT) передаються на графічний процесор (GPU);

б) модель виконує прямий прохід (forward pass) для отримання результату;

в) обчислюється значення функції втрат Reconstruction_criterion, яка поєднує дві складові:

– L1 loss, абсолютна похибка між значеннями пікселів вихідного та еталонного зображень, що сприяє точнішому відновленню основної структури зображення;

– FFT loss, відстань між спектрами зображень у частотній області, що дозволяє зберігати текстури та деталі високих частот.

Зважування здійснюється через коефіцієнт $\alpha = 0.03$, тобто частотна складова впливає на загальну функцію втрат із помірною вагою. Остаточна формула 3.1 має вигляд:

$$\mathcal{L}_{recon} = \mathcal{L}_{L1} + \alpha \cdot \mathcal{L}_{FFT} \quad (3.1)$$

г) проводиться зворотне поширення похибки (backpropagation) для обчислення градієнтів;

д) оптимізатор оновлює ваги моделі на основі обчислених градієнтів.

Для оцінки якості відновлення зображень під час навчання використовується метрика PSNR (Peak Signal-to-Noise Ratio) — вона вимірює співвідношення максимально можливої інтенсивності пікселя до рівня шуму у відновленому зображенні, представлена формулою 3.2. Вища PSNR вказує на кращу якість результату.

$$PSNR = 20 \cdot \log_{10} \left(\frac{PIXEL_{MAX}}{\sqrt{MSE}} \right), \quad (3.2)$$

де $PIXEL_MAX = 1$ для нормалізованих зображень,

MSE — середньоквадратична помилка (MSE) між реконструйованим результатом і GT.

Результати тренування логуються в реальному часі за допомогою WandB (Weights & Biases).

Після кожної епохи:

а) зберігається поточний стан моделі та оптимізатора (last_*.pth);

б) якщо номер епохи кратний check_point_epoch, створюється контрольна точка (epoch_*.pth);

в) після завершення тренування зберігається фінальна версія ваг (final_*.pth).

У процесі тренування моделі TANet було досягнуто характерних змін у функції втрат і метриці PSNR, які свідчать про поступове покращення здатності мережі до реконструкції зображень. На рисунку 3.3 зображено динаміку функції втрат (train loss) протягом 50 епох навчання. Помаранчева лінія відображає тренування на синтетичних даних, а зелена — на реальних. Видно, що модель навчається краще на синтетичних зразках: до останньої епохи значення функції втрат зменшилось до 0.08 на синтетичних та до 0.11 на реальних даних. Це свідчить про меншу середньоквадратичну помилку реконструкції при навчанні на контрольованих штучно згенерованих прикладах, що мають менше шуму й артефактів.

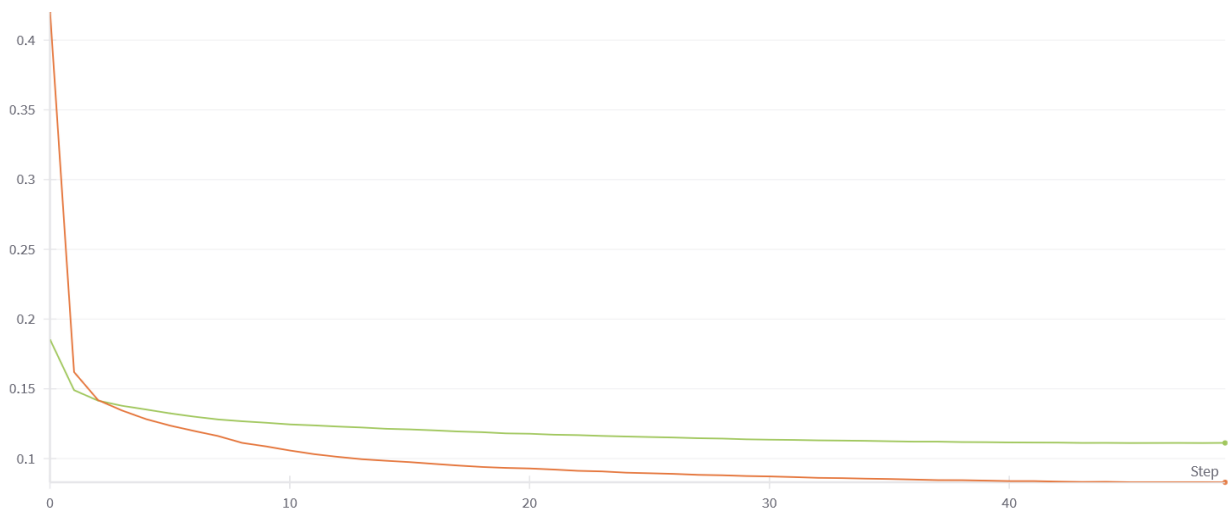


Рисунок 3.3 – Функція втрат моделі TANet

На рисунку 3.4 наведено зміну метрики PSNR (Peak Signal-to-Noise Ratio) протягом тих самих 50 епох. В останній епосі значення PSNR досягли 30.64 для синтетичних і 30.12 для реальних даних. Починаючи з приблизно 33-ї епохи, PSNR на синтетичних прикладах стабільно перевищує показники на реальних. Отже, модель краще відновлює зображення, коли має справу з менш варіативними синтетичними даними. Такі результати підтверджують узгодженість між функцією втрат і метрикою якості, демонструючи стабільне навчання моделі.

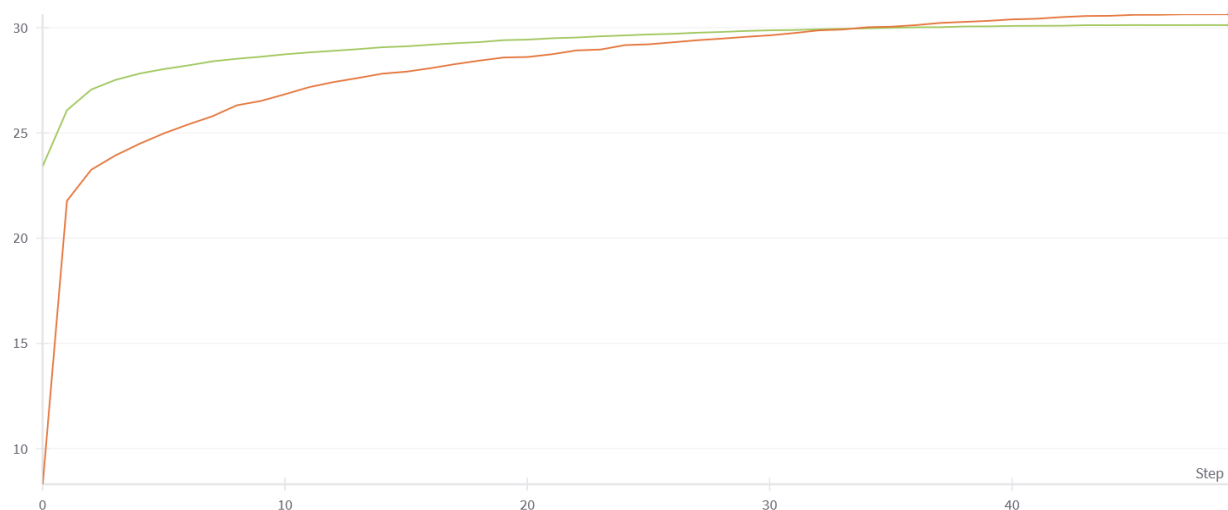


Рисунок 3.4 – Метрика якості PSNR моделі TANet

Модель TANet досить швидко навчилася розпізнавати основні особливості даних, досягаючи значного покращення метрик вже протягом перших 10 епох.

Після цього зростання PSNR та зниження функції втрат сповільнюються, і в останньому десятку епох ці значення майже не змінюються. Це свідчить про досягнення плато навчання, коли модель вичерпала потенціал для подальшого вдосконалення на даному наборі даних. Отже, тренування протягом 50 епох є достатнім, щоб забезпечити стабільні результати, і подальше збільшення кількості епох, ймовірно, не дасть значущого приросту якості.

3.6.2 Цикл тренування MWFormer

Тренування моделі MWFormer починається з ініціалізації параметрів через аргументи командного рядка, які задаються в функції `get_arguments()`. Вони включають такі ключові налаштування, як кількість робочих потоків CPU (`num_workers`) для паралельного завантаження даних, загальна кількість ітерацій тренування (`num_steps = 80000`), розмір батчу (`batch_size = 6`), початкова швидкість навчання (`learning_rate = 5e-5`), коефіцієнт ваги для перцептуальної частини функції втрат (`lambda_loss = 0.04`), шляхи до контрольних точок для відновлення ваг попередньо навчених моделей (`restore_from`, `restore_from_stylefilter`), а також директорії для збереження моделей і тренувальних даних.

На початку тренувального циклу MWFormer виконується підготовка та ініціалізація ключових компонентів. Для забезпечення відтворюваності результатів фіксуються випадкові зерна генераторів псевдовипадкових чисел у модулях `random`, `numpy` та `torch`. Одночасно ініціалізується система логування та моніторингу WandB (Weights & Biases), яка дозволяє відслідковувати прогрес навчання. Далі відбувається конфігурація GPU – обираються доступні графічні пристрої, після чого на них розгортається основна мережа `Network_top`. Якщо вказано шлях до чекпоінта, мережа ініціалізується з попередньо натренованими вагами. Після цього створюється стилевий фільтр `StyleFilter_Top`, який відповідає за витяг векторів стилю з вхідних зображень – він також завантажується з чекпоінта і готується до подальшого тренування. Для розрахунку перцептуальних втрат, які враховують не лише піксельні відмінності, а й структурно-текстурну схожість між

GT та результатом, ініціалізується модель VGG16, з якої беруться перші 16 шарів – ці параметри заморожуються, а сама мережа використовується у ролі фіксованого лос-функціоналу.

Дані для тренування розбиваються на три окремі датасети – для різних типів погодних шумів: краплі дощу (raindrop), дощ (rain) і сніг (snow). Кожен з них завантажується через DataLoader із заданим розміром батчу, кількістю потоків (num_workers) та випадковим перемішуванням. Також виділяється валідаційний набір із повного датасету. Оптимізатор був обраний Adam, що оновлює параметри як мережі відновлення зображень (Network_top), так і стилевого фільтру (StyleFilter).

Цикл виконується від start_iter до num_steps:

а) регулювання швидкості навчання: на 50 000 та 75 000 ітераціях швидкість навчання зменшується вдвічі;

б) отримання батчів: для кожної ітерації завантажуються батчі з трьох датасетів (raindrop, rain, snow). Якщо ітератор датасету досягає кінця, він перезапускається;

в) формування вхідних даних: батчі об'єднуються по вертикалі (concatenate) у великий батч для одночасного тренування;

г) обчислення ознак стилю: стильовий фільтр (StyleFilter) отримує на вхід вхідні зображення і генерує вектор стилю (feature vector);

д) прямий прохід: основна модель Network_top обробляє вхідні зображення разом із вектором стилю і видає відновлені зображення;

е) обчислення функції втрат за допомогою формули 3.3, що поєднує Smooth L1 втрату та перцептуальну втрату на основі попередньо натренованої VGG16. Такий підхід дозволяє моделі враховувати не лише піксельну різницю між передбаченим (pred) і істинним (gt) зображеннями, а й їхню візуальну подібність на високому рівні ознак, що сприяє кращій реконструкції структури та текстур. Smooth L1 Loss є поєднанням L1 та L2 втрат, що є менш чутливою до викидів, ніж звичайна L2:

$$\mathcal{L}_{smooth}(x, y) = \begin{cases} 0.5(x - y)^2, & \text{якщо } |x - y| < 1 \\ |x - y| - 0.5, & \text{інакше} \end{cases} \quad (3.3)$$

Перцептуальна втрата (формула 3.4) обчислюється за допомогою нейронної мережі VGG16 (перші 16 шарів), яка фіксується (`eval() + requires_grad = False`) і використовується як детектор ознак. Вона дозволяє порівнювати високорівневі ознаки (feature maps) між `pred` та `gt`, які витягуються за допомогою `LossNetwork(vgg_model)`:

$$\mathcal{L}_{perc}(x, y) = \sum_l \|\phi_l(x) - \phi_l(y)\|_1 \quad (3.4)$$

де ϕ_l – активації на шарі l мережі VGG16.

Загальна функція втрат:

$$\mathcal{L}_{total} = \mathcal{L}_{smooth} + \lambda \cdot \mathcal{L}_{perc}, \quad (3.5)$$

де λ (в коді `lambda_loss`) – ваговий коефіцієнт, який визначає вплив перцептуальної складової на загальну втрату;

ж) зворотне поширення і оновлення ваг: втрати обчислюються і виконується `loss.backward()`, після чого оптимізатор оновлює ваги моделей;

и) збереження логів: кожні `loss_save_step` (за замовчуванням 1000) ітерацій втрати записуються в масив і логуються в WandB;

к) збереження чекпоінтів: кожні `save_pred_every` (за замовчуванням 20000) ітерацій модель і стилевий фільтр зберігаються у вигляді контрольних точок;

л) валідація: при кожному кроці збереження також виконується оцінка якості на валідаційному наборі, вивчається PSNR і SSIM, результати зберігаються і логуються. Якщо якість покращилась (PSNR більший за попередній), модель зберігається як "найкраща". Метрика SSIM (Structural Similarity Index Measure), представлена у формулі 3.6, є значно точнішим показником візуальної якості зображення, ніж традиційні піксельні метрики (наприклад, MSE чи PSNR),

					ІК11.120БАК.006 ПЗ	Арк.
						50
Зм.	Лист	№ докум.	Підпис	Дата		

оскільки краще відображає те, як людина сприймає зображення. Високе значення SSIM означає, що відновлене зображення дуже подібне до оригінального з точки зору візуального сприйняття. Значення метрики належить до інтервалу $[0,1]$.

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}, \quad (3.6)$$

де μ_x, μ_y – середні значення пікселів зображень x та y ;

σ_x^2, σ_y^2 – дисперсії зображень;

σ_{xy} – коваріація між x та y ;

C_1, C_2 – малі сталі, які стабілізують обчислення при малих знаменниках;

i) раннє завершення: якщо досягнуто максимальної кількості ітерацій (num_steps_stop), тренування завершується і фінальна модель зберігається.

У процесі тренування моделі MWFormer спостерігалася динаміка зміни функції втрат та метрик якості відновлення, що представлена на рисунках 3.5–3.8.

На графіку функції втрат (рис. 3.5), який будувався протягом 80 000 ітерацій з логуванням кожні 500 ітерацій, можна побачити результати трьох варіантів тренування: модель, натренована лише на синтетичних даних (малинова лінія, ~63 000 ітерацій), модель на реальних даних (зелена лінія) та модель на змішаних даних у співвідношенні 1:2 (рожева лінія). Хоча всі графіки мають певну нестабільність і локальні коливання, загальна тенденція є спадною – втрати поступово зменшуються. Особливо це помітно в метриці Time-weighted EMA на рисунку 3.6, де тренування на синтетичних даних демонструє стабільно нижчі втрати протягом усього циклу. Наприкінці тренування спад втрат уповільнюється, однак не зупиняється, що вказує на потенціал подальшого вдосконалення моделі при розширенні кількості епох або обсягу даних.

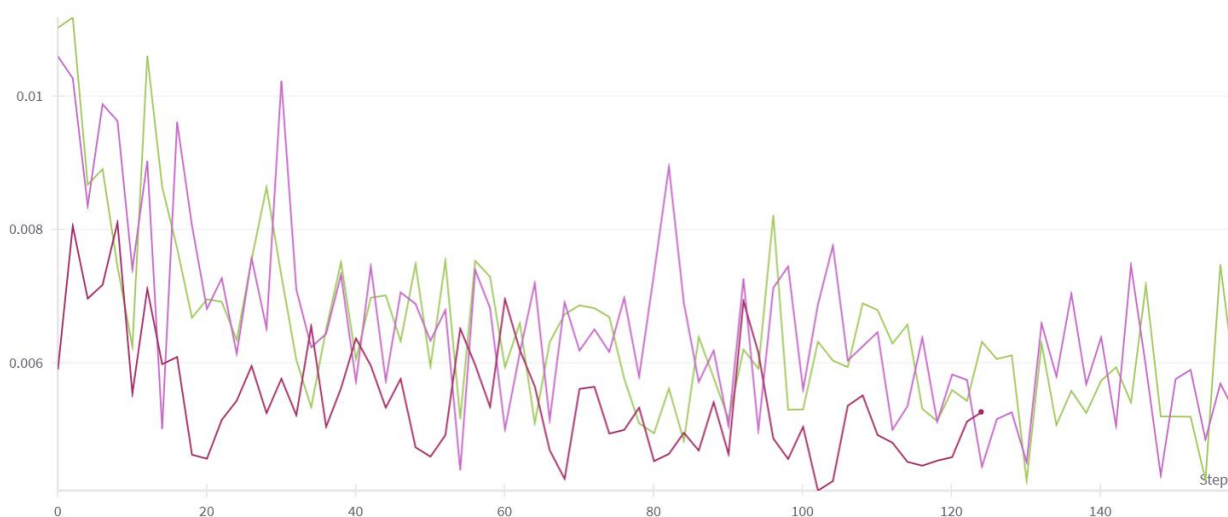


Рисунок 3.5 – Функція втрат моделі MWFormer

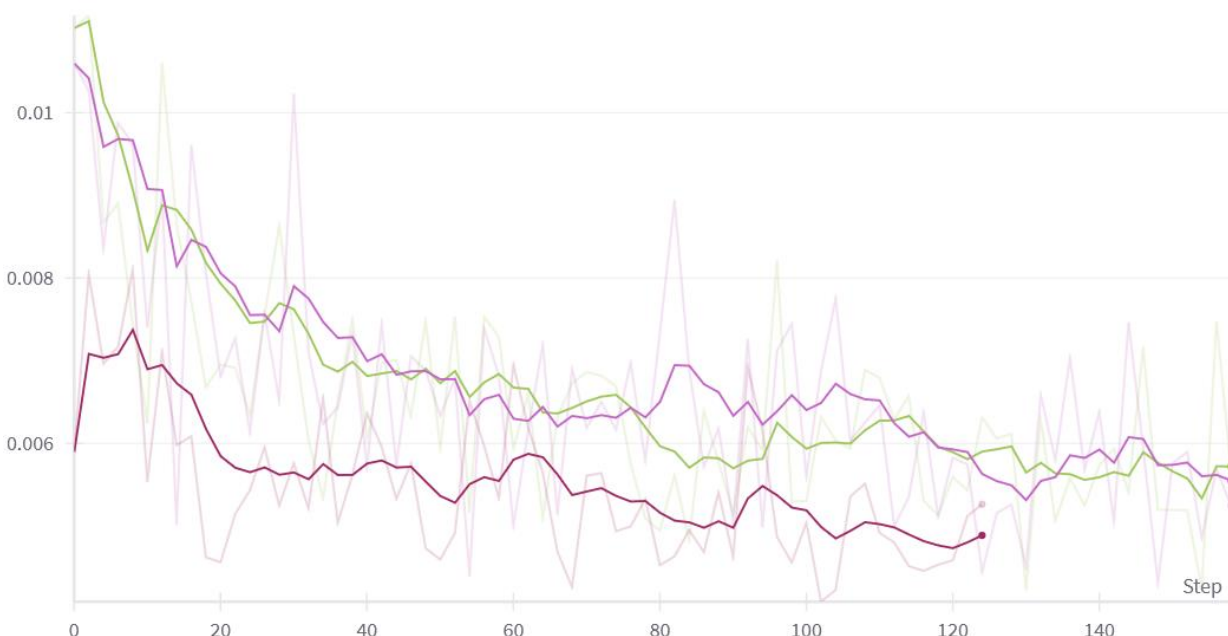


Рисунок 3.6 – Функція втрат моделі MWFormer згладжена

Оцінка метрик PSNR та SSIM на валідаційних даних (рис. 3.7–3.8) демонструє іншу картину. Модель, натренована на змішаних даних, показує дещо гірші результати: середній PSNR ~ 29 та SSIM ~ 0.83 , що свідчить про зниження якості реконструкції. Натомість при роздільному тренуванні на синтетичних і реальних даних, результати вищі — близько 33 PSNR у середньому. Найвищі значення SSIM (~ 0.97) спостерігаються на синтетичних даних, тоді як на реальних вони трохи нижчі (~ 0.95). Водночас, на змішаному датасеті метрики

Зм.	Лист	№ докум.	Підпис	Дата

характеризуються значним шумом і непослідовністю, що може бути пов'язано з різномірністю даних або невдалим співвідношенням їх у тренувальному процесі.



Рисунок 3.7 – Метрика якості PSNR моделі MWFormer

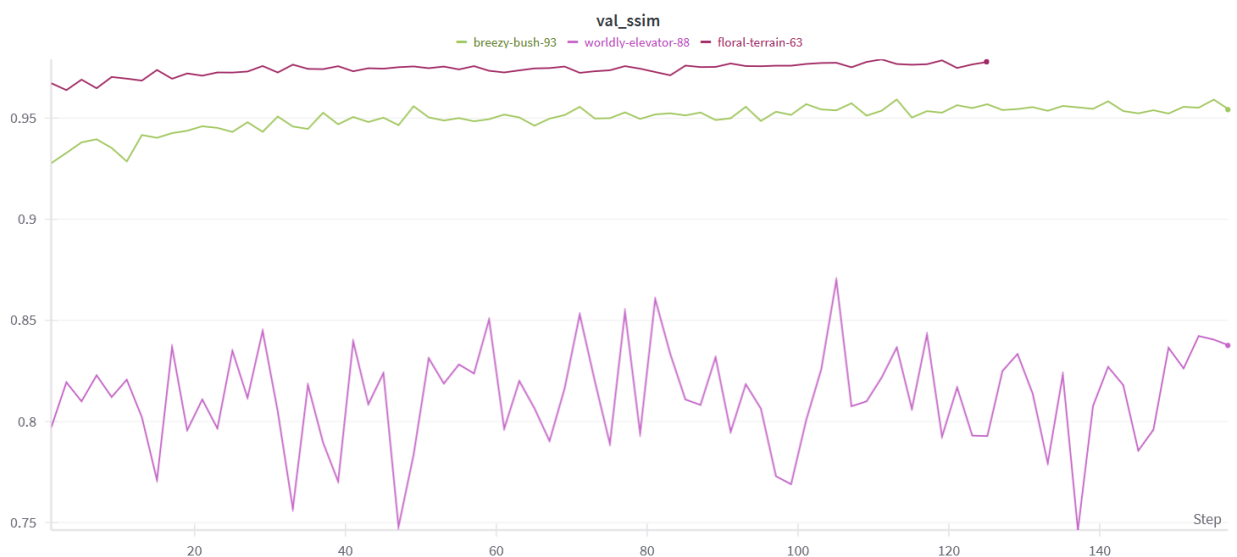


Рисунок 3.8 – Метрика якості SSIM моделі MWFormer

Модель MWFormer демонструє кращу здатність до відновлення при тренуванні на чистих синтетичних або реальних датасетах, особливо в плані структурної подібності та точності (SSIM і PSNR). Тренування на змішаних даних не дає таких стабільних результатів і має більше коливань, що вказує на потенційні складнощі з узагальненням. Незважаючи на це, загальне зниження функції втрат і збереження позитивної динаміки в кінці тренування підтверджують, що модель має

потенціал для подальшого поліпшення результатів при відповідній адаптації під набір даних.

3.8 Опис етапу валідації моделей

3.8.1 Валідація для моделі TANet

Етап валідації моделі TANet слугує для оцінки її здатності відновлювати зображення за різних погодних умов на основі тестового набору даних. У цьому процесі використовується попередньо натренована модель, яка завантажується разом зі збереженим станом оптимізатора та scheduler'a, що дозволяє забезпечити точну відповідність результатів останнього етапу тренування. Для оцінки обрається певний тестовий датасет, який подається у модель з фіксованим розміром батчу без перемішування, аби результати можна було зіставити із відповідними входами.

Перед подачею у модель вхідні зображення попередньо обробляються: якщо їх розміри не кратні 64, вони доповнюються паддінгом за допомогою відбиття. Це забезпечує сумісність із структурою моделі, що очікує вхід певного формату. Після обробки зображення передаються у модель TANet, яка реконструює чисту версію кожного кадру. Результати обрізаються до початкових розмірів, щоб виключити вплив доданих пікселів.

Якість реконструкції оцінюється за допомогою спеціальної функції втрат `Reconstruction_criterion`. Додатково обчислюється метрика PSNR. По завершенні проходження через весь датасет усі отримані значення втрат та PSNR агрегуються й обчислюються як середні значення. Вони логуються у сервіс WandB, а також виводяться у консоль для подальшого аналізу. Валідація демонструє, наскільки добре модель справляється із завданням очищення зображень при подачі на вхід нових, не бачених під час навчання даних, і допомагає виявити потенційні зони для подальшого поліпшення або перенавчання.

3.8.2 Валідація для моделі MWFormer

Етап валідації моделі MWFormer побудований як автономний процес, що дозволяє протестувати продуктивність попередньо навчених компонентів – основної відновлювальної мережі та фільтра стилів – на заданому тестовому наборі. Перед початком перевірки конфігурація налаштовується за допомогою аргументів командного рядка, які визначають шляхи до збережених ваг моделей, директорію з валідаційними даними, розмір батчу та інші параметри. Водночас ініціалізується WandB-сесія для збирання метрик і логування результатів. Для відтворюваності валідаційного експерименту встановлюється фіксоване зерно генератора випадкових чисел.

Для узгодженості з розмірами, прийнятними для нейронної мережі, зображення масштабуються так, щоб їхні розміри були кратні 16. Потім вони нормалізуються та перетворюються на тензори. Кожне вхідне зображення разом із GT подається в модель без перемішування порядку, оскільки ключовим є точне порівняння результату з оригіналом.

Валідація проводиться шляхом послідовної передачі кожного зображення через два компоненти: StyleFilter та Network_top. StyleFilter відповідає за витяг та узгодження стилістичних особливостей, що можуть бути релевантними для коректного очищення зображення, тоді як Network_top виконує безпосереднє відновлення. Обидві мережі попередньо завантажуються з файлів із збереженими вагами, при цьому всі параметри фільтра стилів блокуються для уникнення випадкових змін.

Після обробки кожного прикладу модель обчислює метрики якості: PSNR та SSIM, які є класичними показниками точності відновлення зображення. Усі результати агрегуються для формування середнього значення по всьому тестовому набору, після чого логуються у WandB для подальшого аналізу. Додатково фіксується час, витрачений на всю валідацію, що дозволяє оцінити ефективність роботи моделі в реальних умовах. У підсумку цей етап демонструє, наскільки добре MWFormer справляється з очищенням погодних артефактів на практиці.

3.9 Інтерфейсні компоненти

Для забезпечення зручної взаємодії користувача з підсистемою обробки відеоданих передбачено два ключові інтерфейсні блоки – графічний інтерфейс (GUI) разом з панеллю конфігурації. Це дозволяє гнучко керувати процесом обробки та оперативно змінювати параметри без необхідності прямого доступу до коду чи терміналу.

3.9.1 Графічний інтерфейс користувача (GUI)

Цей модуль слугує основним середовищем взаємодії для користувача. Його структура включає:

- кнопки для імпорту даних. Забезпечують завантаження зображень або відеофайлів з локальної файлової системи. Підтримуються різні формати (JPEG, PNG, MP4 тощо), з можливістю пакетного імпорту;

- випадаючий список вибору моделі. Дозволяє користувачу швидко обрати одну з доступних моделей для обробки – TANet або MWFormer. Перемикання між ними оновлює параметри обробки в системі;

- області візуалізації вхідного та обробленого контенту. Інтерфейс передбачає дві панелі попереднього перегляду: ліва відображає оригінальні кадри, а права – результати після проходження обробки. Це дає змогу здійснювати пряме порівняння «до» та «після» покадрово або у вигляді відеопотоку;

- система повідомлень. Для інформування користувача про статус обробки, помилки або успішне завершення процесів передбачені спливаючі інформативні повідомлення, які з'являються в нижній частині екрана. Також передбачений підрахунок часу обробки завантаженого файлу.

На рисунках 3.9 та 3.10 зображені знімки екрану наявного інтерфейсу програми перед початком роботи та після процесу оцінки зображення обраною моделлю.

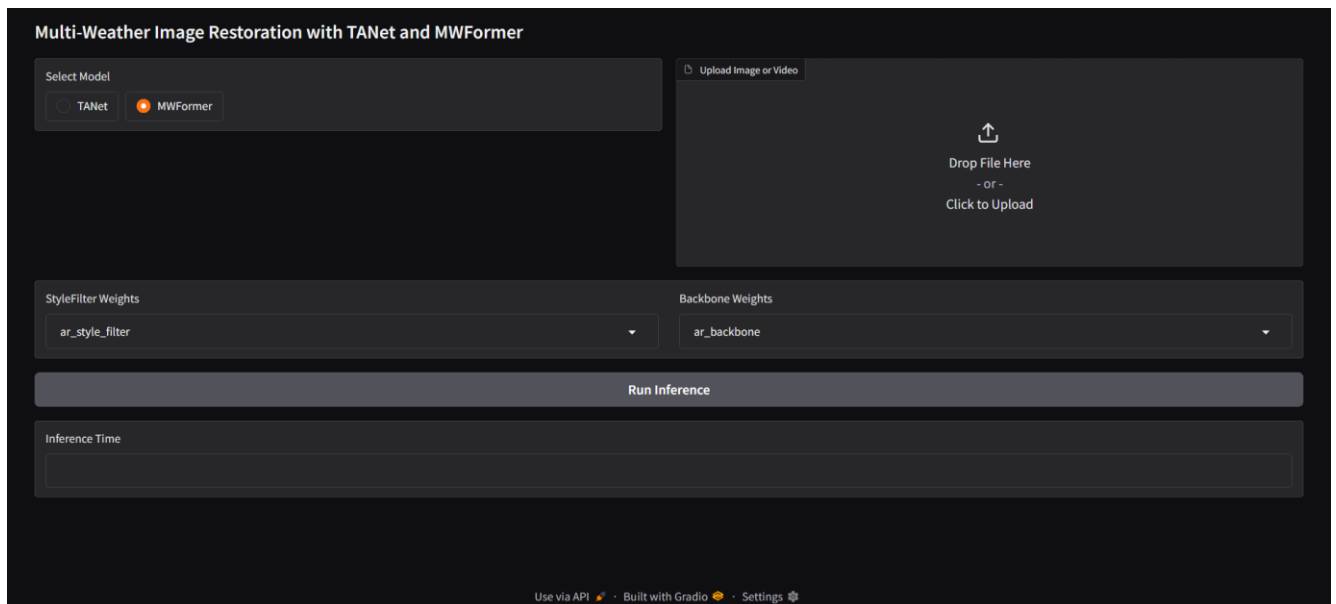


Рисунок 3.9 – Графічний інтерфейс для демонстрації роботи моделей

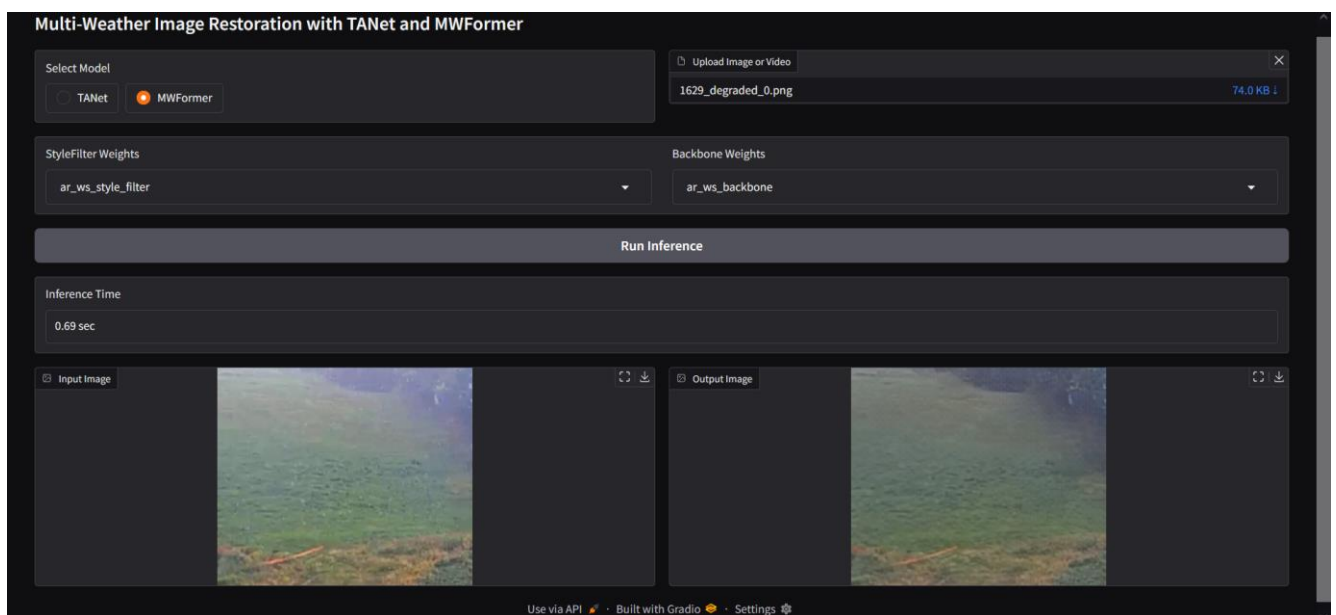


Рисунок 3.10 – Графічний інтерфейс після оцінки зображення

Висновки до розділу 3

У розділі 3 дипломної роботи було реалізовано повний цикл розробки програмного забезпечення для дослідження та порівняння двох моделей глибокого навчання – TANet і MWFormer – у задачі покращення зображень в складних погодних умовах. На початковому етапі було обґрунтовано вибір інструментів, бібліотек та мов програмування, з урахуванням специфіки задачі, продуктивності

					ІК11.120БАК.006 ПЗ	Арк.
Зм.	Лист	№ докум.	Підпис	Дата		57

обробки зображень та сумісності з сучасними фреймворками для машинного навчання. Крім того, було визначено апаратні ресурси, необхідні для ефективного тренування й тестування моделей.

Програмний продукт було детально описано як з функціональної, так і зі структурної точки зору. Було реалізовано модульну архітектуру, яка передбачає окреме оброблення вхідних даних, підготовку до навчання, запуск тренувального циклу та валідацію. Для кожної з моделей були реалізовані окремі етапи попередньої обробки даних, що враховують їхню архітектурну специфіку та вимоги до вхідних форматів. Це забезпечило коректне подання зображень у нейронні мережі та підвищило стабільність тренування.

У розділі також було детально описано алгоритм циклу навчання для кожної з моделей. Для TANet це був класичний підхід з використанням функції втрат, оптимізатора Adam, крокового зниження швидкості навчання та збереження контрольних ваг. У випадку MWFormer на окрему увагу заслуговують дві складові моделі – фільтр стилю та основна мережею відновлення, а також правильне завантаження ваг з попереднього навчання.

Фінальним етапом стало проведення валідації моделей на тестовому наборі зображень. Обидві моделі було оцінено за допомогою класичних метрик PSNR і SSIM, а результати було зафіксовано з використанням системи логування WandB. Це дозволило об'єктивно порівняти якість відновлення зображень кожною моделлю та зробити висновки про їхню ефективність. Таким чином, розділ програмного забезпечення заклав технічну основу для всього експериментального дослідження, забезпечивши повний цикл обробки, навчання та оцінки моделей.

4 РЕЗУЛЬТАТИ РОБОТИ ТА ЕКСПЕРИМЕНТАЛЬНІ ДОСЛІДЖЕННЯ

У цьому розділі представлено результати експериментальних досліджень, проведених із використанням розробленого програмного забезпечення. Експерименти включають опис методики тестування, отримані показники якості відновлених зображень, а також їх аналіз та порівняння. Основну увагу зосереджено на оцінці продуктивності натренованих моделей MWFormer та TANet.

Було проаналізовано характеристики кожної моделі. MWFormer складається з двох частин: основної мережі (backbone) з 170 мільйонами параметрів та модуля StyleFilter з понад 12 мільйонами параметрів. У порівнянні з цим, модель TANet є менш об'ємною, маючи приблизно 9.9 мільйона параметрів. Тренування MWFormer на GPU тривало 17 годин 41 хвилину на 80 тисячах ітерацій, тоді як TANet було навчено за 10 годин 49 хвилин протягом 50 епох.

Усі експерименти та тестування моделей були виконані із використанням графічного процесора, що забезпечило ефективну обробку даних. Підсумкові результати представлені у таблицях 4.1–4.10 та є основою для подальшого порівняльного аналізу моделей.

4.1 Результати валідації моделей

З отриманих результатів, що наведені у таблиці 4.1, видно, що модель TANet, навчена на синтетичних погодних даних, демонструє стабільні та передбачувані характеристики залежно від кількості епох. Найкращу якість видалення туману модель показала на 40-й епосі, досягнувши найвищого значення PSNR 30.91808 та найменшого значення функції втрат 0.10434. Це свідчить про те, що на цьому етапі модель досягла оптимального балансу між навченістю та генералізацією для даного типу атмосферного спотворення.

У випадку дощу та снігу найкращі результати були зафіксовані на 50-й епосі, де PSNR для дощу склав 31.40387, а для снігу – 34.12732. Це може свідчити про те, що для ефективного видалення цих типів завад модель потребує дещо більшої

кількості ітерацій навчання, ймовірно, через їхню більшу візуальну варіативність або складність у порівнянні з туманом.

Щодо реальних даних з датасету WeatherStream, різниця між 40-ю та 50-ю епохами є мінімальною, хоча найвищий PSNR (21.84585) досягнуто на 40-й епосі, а найменша втрата (0.28463) – на 50-й. Це свідчить про відносну стабільність моделі на практичних прикладах, а також про те, що приріст якості після 40-ї епохи вже не є суттєвим. Таким чином, для загальних задач очистки зображень від атмосферних спотворень можна вважати, що 40-а або 50-а епоха є прийнятним компромісом між ефективністю та часом тренування.

Таблиця 4.1 – Порівняльна таблиця моделі TAnet, що тренована на синтетичних даних

Model name	Epochs	Test set	Test psnr (10 imgs)	Test loss (10 imgs)
TANet_exp_001(synthetic data)	30	haze	30.56238	0.10776
TANet_exp_001(synthetic data)	30	rain	31.3694	0.16713
TANet_exp_001(synthetic data)	30	snow	33.76112	0.090435
TANet_exp_001(synthetic data)	30	stream	21.70473	0.28791
TANet_exp_001(synthetic data)	40	haze	30.91808	0.10434
TANet_exp_001(synthetic data)	40	rain	31.37947	0.16738
TANet_exp_001(synthetic data)	40	snow	33.76112	0.090435
TANet_exp_001(synthetic data)	40	stream	21.84585	0.28476
TANet_exp_001(synthetic data)	50	haze	30.5266	0.10684
TANet_exp_001(synthetic data)	50	rain	31.40387	0.16713
TANet_exp_001(synthetic data)	50	snow	34.12732	0.087478
TANet_exp_001(synthetic data)	50	stream	21.83539	0.28463

Як видно з таблиці 4.2, модель TAnet, натренована на синтетичних погодних даних, демонструє покращені результати при тестуванні на більшій вибірці реальних зображень. Зокрема, найкращі значення метрик досягнуто на 50-й епосі: PSNR склав 22.17922, а функція втрат – 0.26986. Це підтверджує припущення про те, що збільшення кількості епох сприяє кращому узагальненню моделі на практичні приклади.

Крім того, порівняно з попереднім експериментом на 10 зображеннях, цього разу було використано 39 зображень, і середні значення метрик дещо покращилися. Це свідчить про те, що модель добре масштабується при роботі з більшими обсягами даних, не втрачаючи якості оцінки. Час тестування на цій вибірці становив приблизно 1.98 секунди, що свідчить про достатню швидкодію моделі навіть при розширенні вибірки.

Таблиця 4.2 – Порівняльна таблиця моделі TAnet, що тренована на синтетичних даних та тестована на реальних даних

Model name	Epochs	Test set	Test psnr (10 imgs)	Test loss (10 imgs)	Test psnr (39 imgs)	Test loss (39 imgs)	Test time (39 imgs), sec
TANet_exp_001(synthetic data)	30	stream	21.70473	0.28791	21.98181	0.27315	1.98995
TANet_exp_001(synthetic data)	40	stream	21.84585	0.28476	22.15481	0.27001	1.96785
TANet_exp_001(synthetic data)	50	stream	21.83539	0.28463	22.17922	0.26986	1.98516

Як показано в таблиці 4.3, моделі TAnet, натреновані на різних типах даних, при однаковому часу навчання (30 епоха), демонструють чітко виражену залежність результатів від типу навчального набору. Зокрема, модель, натренована на синтетичних погодних умовах (TANet_exp_001), значно краще справляється із зображеннями, що відповідають синтетичним сценаріям, з високими значеннями PSNR для всіх типів атмосферних викривлень.

Натомість модель, навчена на реальних даних (TANet_exp_002), суттєво поступається при обробці синтетичних зображень. Наприклад, її результати за PSNR для синтетичного туману, дощу та снігу були нижчими відповідно на 10.44137, 8.68314 та 15.59995 одиниць порівняно з моделлю, навченою на синтетичних даних. У той же час, на реальному наборі weatherStream навпаки, вона показала кращу якість відновлення зображень, перевищивши модель на синтетичних даних на 0.51647 за PSNR.

Ці результати підтверджують, що якість відновлення моделі суттєво залежить від подібності між навчальними і тестовими даними. Модель краще узагальнює інформацію з тієї самої доменної області, на якій була натренована, і її продуктивність помітно знижується при переході до іншого типу вхідних даних.

Таблиця 4.3 – Порівняльна таблиця моделей TAnet, що треновані на синтетичних та реальних даних

Model name	Epochs	Test set	Test psnr (10 imgs)	Test loss (10 imgs)
TANet_exp_001(synthetic data)	30	haze	30.56238	0.10776
TANet_exp_001(synthetic data)	30	rain	31.3694	0.16713
TANet_exp_001(synthetic data)	30	snow	33.76112	0.090435
TANet_exp_001(synthetic data)	30	stream	21.70473	0.28791
TANet_exp_002 (real data)	30	haze	20.12101	0.27194
TANet_exp_002 (real data)	30	rain	22.68626	0.27242
TANet_exp_002 (real data)	30	snow	18.16127	0.36807
TANet_exp_002 (real data)	30	stream	22.2212	0.26962

Модель TAnet, натренована на реальних даних, демонструє найвищі результати за метриками PSNR і loss саме на 30-й епосі, незалежно від розміру вибірки – як на 10, так і на 39 зображеннях, як видно з таблиці 4.4. Це свідчить про оптимальне співвідношення між навченістю моделі та її здатністю до узагальнення на даному етапі.

Починаючи з 35-ї епохи, показники якості починають погіршуватися, що вказує на явище перенавчання. Модель надмірно адаптується до тренувального набору і втрачає ефективність при обробці нових, раніше не бачених прикладів. Це особливо критично для реальних даних, які мають високу варіативність та шум.

Таким чином, отримані результати підтверджують доцільність зупинки навчання TAnet, натренованої на реальних зображеннях, на 30-й епосі. Це дозволяє уникнути перенавчання й досягти кращої узагальнюваності моделі. Час тестування при цьому залишається стабільним — приблизно 1.96 секунди на 39 зображеннях.

Таблиця 4.4 – Порівняльна таблиця моделі TAnet, що тренована та тестована на реальних даних

Model name	Epochs	Test set	Test psnr (10 imgs)	Test loss (10 imgs)	Test psnr (39 imgs)	Test loss (39 imgs)	Test time (39 imgs), sec
TANet_exp_002 (real data)	30	stream	22.2212	0.26962	22.45452	0.26527	1.94127
TANet_exp_002 (real data)	35	stream	22.18736	0.27116	22.42762	0.2665	1.96204
TANet_exp_002 (real data)	40	stream	22.15543	0.27212	22.37545	0.26778	1.96578

Як показано в таблиці 4.5, результати порівняння з оригінальною моделлю TANet, описаною в статті, демонструють конкурентоспроможність моделей, натренованих у межах даного проєкту. Незважаючи на значно меншу кількість епох у процесі навчання (30 для моделі на реальних даних і 50 для моделі на синтетичних), вони демонструють вищу або порівнювану якість зображень після відновлення.

Зокрема, модель, натренована на синтетичних даних, досягає кращих результатів при обробці зображень із туманом – на 0.95564 більше за показником PSNR порівняно з оригінальною версією, і також демонструє нижчий рівень втрат (loss). При цьому, на зображеннях із дощем, оригінальна модель має перевагу – на 0.74049 за PSNR. Щодо снігу, варто враховувати, що оригінальна модель використовувала інший набір даних (Snow100K), тому отримана в рамках проєкту модель суттєво перевершила її за PSNR на 7.0625, ймовірно, через кращу адаптацію до власного датасету.

Важливо також зазначити, що оригінальна модель тренувалася протягом 500 епох, тоді як моделі з цього дослідження – лише до 30 та 50 епох відповідно. Це підкреслює якісну побудову тренувального процесу, що дозволило досягти високих результатів за значно менших ресурсних витрат.

Таблиця 4.5 – Порівняльна таблиця моделей TANet, що тестовані на синтетичних даних

Model name	Epochs	Test set	Test psnr (10 imgs)	Test loss (10 imgs)
TANet	500	haze	29.57096	0.11311
TANet	500	rain	32.14481	0.16016
TANet	500	snow	27.06482	0.15382
TNet_exp_001 (synthetic data)	50	haze	30.5266	0.10684
TNet_exp_001 (synthetic data)	50	rain	31.40387	0.16713
TNet_exp_001 (synthetic data)	50	snow	34.12732	0.087478
TANet_exp_002 (real data)	30	haze	20.12101	0.27194
TANet_exp_002 (real data)	30	rain	22.68626	0.27242
TANet_exp_002 (real data)	30	snow	18.16127	0.36807

При порівнянні результатів тестування на реальних зображеннях, наведених у таблиці 4.6, вони демонструють різні тенденції залежно від розміру вибірки. На меншій вибірці з 10 зображень оригінальна модель TANet, що була тренувана повністю на синтетичних даних і протягом 500 епох, забезпечує найкращі результати. Це може бути зумовлено тим, що при меншій кількості прикладів її здатність до генералізації залишається стабільною.

Однак на більшій вибірці з 39 зображень перевагу демонструють моделі, розроблені в межах проєкту. Найкращий результат показує модель, натренована на реальних даних, а на другому місці – модель, натренована на синтетичних даних. Оригінальна модель в цьому випадку поступається обом. Така зміна розстановки сил свідчить про те, що розроблені моделі краще адаптовані до реальних умов, і їх продуктивність зростає зі збільшенням обсягу тестових даних.

Зважаючи на те, що оригінальна модель тренувалася упродовж 500 епох, а моделі в проєкті – лише до 50, досягнуте перевершення на реальному наборі вказує на більш вдалий підбір даних для тренування та потенційно кращу здатність моделей до роботи в реальних умовах. Це є суттєвим досягненням, що підтверджує ефективність запропонованого підходу до навчання моделей машинного зору.

Таблиця 4.6 – Порівняльна таблиця моделей TANet, що тестовані на реальних даних

Model name	Epochs	Test set	Test psnr (10 imgs)	Test loss (10 imgs)	Test psnr (39 imgs)	Test loss (39 imgs)	Test time (39 imgs), sec
TANet	500	stream	22.27444	0.27877	21.97716	0.27364	1.9614
TANet_exp_001 (synthetic data)	50	stream	21.83539	0.28463	22.17922	0.26986	1.98516
TANet_exp_002 (real data)	30	stream	22.2212	0.26962	22.45452	0.26527	1.94127

Зображення, отримані в результаті оцінки на різних варіантах моделей TANet, наочно підтверджують наведені вище висновки. У кожній послідовності зображень спочатку представлено вхідне зображення зі значними атмосферними спотвореннями, після чого йдуть результати відновлення для моделей TANet, натренованих на синтетичних (TANet_001) та реальних (TANet_002) даних

відповідно. У додатку А (рис. А.1–А.5) подано результат оригінальної моделі TANet, представленої в статті, та еталонне зображення (Ground Truth) для порівняння. Таке візуальне представлення дозволяє не лише кількісно, а й якісно оцінити різницю в здатності моделей до усунення атмосферних спотворень.

Наступна таблиця 4.7 демонструє порівняння трьох варіантів моделі MWFormer, натренованих на різних типах даних: синтетичних (MWFormer_run1), реальних (MWFormer_run2) та на суміші синтетичних і реальних (MWFormer_run3). Аналіз проведено на вибірці з 10 зображень із різних датасетів, що включають синтетичні туман (RESIDE-6K), дощ (Rain1400), сніг (SnowCityScapes) та реальні дані з weatherStream.

Очікувано, модель, натренована на синтетичних даних, показує найкращі результати на відповідних синтетичних зображеннях. Водночас модель, що тренувалась на суміші даних, демонструє дуже близькі до неї показники, поступаючись по PSNR лише трохи більше ніж на одиницю. Для реальних даних кращим виявився MWFormer_run2 з PSNR 24.81987, хоча різниця з MWFormer_run3 невелика. Щодо метрики SSIM, найвищі значення зафіксовані у моделі MWFormer_run1 – 0.75798, друга позиція за MWFormer_run3 – 0.75148, і лише третя – за MWFormer_run2 із 0.74423.

Середній час оцінки для всіх тестових наборів склав близько 1.48 секунди на 10 зображень, що свідчить про ефективність реалізації моделі у контексті швидкодії. Отже, результати підтверджують важливість вибору відповідного тренувального набору даних для досягнення оптимальної якості очищення зображень у різних умовах.

Таблиця 4.7 – Порівняльна таблиця тренуваних моделей MWFormer, що тестовані на синтетичних та реальних даних

Model name	Test set	Test time (10 images), sec	Test ssim (10 imgs)	Test psnr (10 imgs)
MWFormer_run1 (synthetic data)	haze	1.94161	0.98196	33.4537
MWFormer_run1 (synthetic data)	rain	1.52515	0.91569	33.30014
MWFormer_run1 (synthetic data)	snow	1.31608	0.9702	35.42151
MWFormer_run1 (synthetic data)	stream	1.14472	0.75798	23.74173
MWFormer_run2 (real data)	haze	1.91691	0.85155	22.28771

Model name	Test set	Test time (10 images), sec	Test ssim (10 imgs)	Test psnr (10 imgs)
MWFormer_run2 (real data)	rain	1.53009	0.835	25.00298
MWFormer_run2 (real data)	snow	1.31759	0.47283	19.18795
MWFormer_run2 (real data)	stream	1.13566	0.74423	24.81987
MWFormer_run3 (synthetic+real data)	haze	1.95106	0.98003	32.74622
MWFormer_run3 (synthetic+real data)	rain	1.53754	0.91493	32.95792
MWFormer_run3 (synthetic+real data)	snow	1.32991	0.96518	34.59347
MWFormer_run3 (synthetic+real data)	stream	1.13425	0.75148	24.161

Ці відмінності також підтверджуються на прикладі результатів оцінки зображень, представлених нижче. Зокрема, можна помітити, що синтетичний сніг ефективно видаляється моделями MWFormer_run1 та MWFormer_run3, демонструючи якісне очищення. Водночас, реальний сніг усі моделі прибирають однаково недостатньо, що свідчить про складнішу і відмінну структуру шуму у реальних погодних умовах. Це підкреслює важливість врахування особливостей реальних даних при розробці та тренуванні моделей для задач очищення зображень.

Послідовність зображень для порівняння у додатку Б (рис. Б.1–Б.5) включає початкове зображення, результати роботи трьох моделей MWFormer: MWFormer_run1, MWFormer_run2 та MWFormer_run3, а також еталонне зображення (Ground Truth). Такий порядок дає змогу наочно оцінити якість очищення, виконаного кожною моделлю, і простежити ступінь наближення до ідеального результату.

У таблиці 4.8 наведено порівняння оригінальних моделей MWFormer зі статті, а також моделей MWFormer_real та MWFormer_L на різних наборах даних, включно із синтетичними (туман, дощ, сніг) та реальними даними з weatherStream, на вибірці з 10 зображень. Загалом, модель MWFormer_real демонструє кращі результати у порівнянні з MWFormer_L за показниками PSNR та SSIM. Особливо помітна перевага MWFormer_real на реальних даних, де вона перевищує MWFormer_L за PSNR на 2.55541, досягаючи значення 25.19491, а також має

нижчий показник loss – 0.76754. Ці результати свідчать про більш ефективну здатність моделі MWFormer_real до очищення зображень у реальних умовах.

Таблиця 4.8 – Порівняльна таблиця оригінальних моделей MWFormer, що тестовані на синтетичних та реальних даних

Model name	Test set	Test time (10 images), sec	Test ssim (10 imgs)	Test psnr (10 imgs)
MWFormer_real	haze	1.91353	0.89721	20.43392
MWFormer_real	rain	1.52286	0.87626	29.47276
MWFormer_real	snow	1.26693	0.89309	27.38661
MWFormer_real	stream	1.08503	0.76754	25.19491
MWFormer_L	haze	1.90295	0.87669	18.42528
MWFormer_L	rain	1.56185	0.87329	29.5604
MWFormer_L	snow	1.27679	0.88488	27.01041
MWFormer_L	stream	1.14255	0.76517	22.6395

На зображеннях, наведених у додатку В (рис. В.1–В.5), чітко видно, що моделі краще справляються з очищенням саме реальних зображень порівняно з синтетичними. Водночас, сніг на синтетичних даних також очищується досить якісно, що демонструє хорошу адаптацію моделей до цих умов. Особливо варто відзначити, що очищення снігу на реальних даних відбувається найкраще серед усіх протестованих моделей, що підтверджує їх високу ефективність у роботі з реальними погодними умовами. Це також може свідчити про більш вдало підібрані тренувальні набори та більшу кількість ітерацій під час навчання, що сприяє кращій узагальненості та точності моделей при роботі з цим видом шуму.

У таблиці 4.9 наведено порівняння моделей MWFormer, тренуваних у проєкті (MWFormer_run1, MWFormer_run2, MWFormer_run3), з оригінальними моделями MWFormer_real та MWFormer_L на реальних даних. Тестування виконувалося на вибірках з 10 та 39 зображень. Оригінальні моделі MWFormer_real та MWFormer_L були навчені на синтетичних наборах даних RainDrop, Outdoor-Rain і Snow100K із тонким налаштуванням протягом 190 тисяч ітерацій.

Результати показують, що модель MWFormer_real демонструє найкращі показники за метриками PSNR та SSIM серед усіх моделей на обох вибірках. На другому місці за PSNR (23.38809) та SSIM (0.74208) на вибірці з 39 зображень

знаходиться MWFormer_run1, яка була тренувана на синтетичних даних. Модель MWFormer_run2 посідає друге місце за PSNR, а MWFormer_run3 – друге за SSIM на тій же вибірці.

Середній час оцінки для 10 зображень становить 1.13 секунди, а для 39 зображень – близько 3.07 секунди. Ці результати свідчать про конкурентоспроможність розроблених моделей, особливо MWFormer_run1, попри меншу кількість ітерацій тренування у порівнянні з оригінальними моделями.

Таблиця 4.9 – Порівняльна таблиця тренуваних та оригінальних моделей MWFormer, що тестовані на реальних даних

Model name	Test set	Test time (10 images), sec	Test ssim (10 imgs)	Test psnr (10 imgs)	Test ssim (39 imgs)	Test psnr (39 imgs)	Test time (39 imgs), sec
MWFormer_run1 (synthetic data)	stream	1.14472	0.75798	23.74173	0.74208	23.38809	3.06527
MWFormer_run2 (real data)	stream	1.13566	0.74423	24.81987	0.71622	23.75962	3.08468
MWFormer_run3 (synthetic+real data)	stream	1.13425	0.75148	24.161	0.73292	23.35067	3.06989
MWFormer_real	stream	1.08503	0.76754	25.19491	0.74731	24.40169	3.06932
MWFormer_L	stream	1.14255	0.76517	22.6395	0.73606	21.90275	3.04306

У таблиці 4.10 наведено порівняння найкращих моделей TANet та MWFormer, тренуваних на різних наборах даних та протестованих на реальних зображеннях. Загалом спостерігається, що моделі MWFormer демонструють вищі показники за метрикою PSNR у порівнянні з TANet. Зокрема, середнє значення різниці PSNR між моделями MWFormer і TANet на вибірці з 39 зображень становить 1.521 на користь MWFormer.

Водночас TANet має перевагу у швидкості обробки: середній час тестування на вибірці з 39 зображень на GPU приблизно на 1.0988 секунди менший, ніж у MWFormer. Ці результати свідчать про те, що MWFormer забезпечує кращу якість очищення зображень, тоді як TANet є більш ефективним з точки зору швидкодії.

Таблиця 4.10 – Порівняльна таблиця моделей TANet та MWFormer, що тестовані на реальних даних

Model name	Epochs	Test set	Test psnr (10 imgs)	Test psnr (39 imgs)	Test time (39 imgs), sec
MWFormer_run1 (synthetic data)	best_all	stream	23.74173	23.38809	3.06527
MWFormer_run2 (real data)	best_all	stream	24.81987	23.75962	3.08468
MWFormer_run3 (synthetic+real data)	best_all	stream	24.161	23.35067	3.06989
MWFormer_real	backbone	stream	25.19491	24.40169	3.06932
TANet	500	stream	22.27444	21.97716	1.9614
TANet_exp_001 (synthetic data)	50	stream	21.83539	22.17922	1.98516
TANet_exp_002 (real data)	30	stream	22.2212	22.45452	1.94127

У результаті проведених експериментальних досліджень можна зробити кілька важливих висновків. Модель MWFormer демонструє вищу якість очищення зображень у порівнянні з TANet, проте її можна додатково вдосконалити шляхом тренування на більшому обсязі даних. Зокрема, для покращення розпізнавання снігу рекомендується використовувати набори даних із більшою представленістю цього явища, оскільки наявні моделі недостатньо ефективно справляються з очищенням реального снігу.

Щодо продуктивності, TANet показує кращі результати за швидкістю, що робить її більш придатною для застосувань з обмеженими обчислювальними ресурсами або де важлива оперативність обробки. Водночас MWFormer варто застосовувати там, де пріоритетом є якість результату. Таким чином, вибір моделі залежить від конкретних вимог до балансу між якістю очищення зображень та часом їх обробки.

4.2 Результати тестування на апаратних пристроях

Для оцінки швидкодії було зібрано дані часу виконання розроблених моделей TANet_exp_002 (35 епоха тренування) та MWFormer_run2 (краща ітерація) на різних апаратних платформах. Для тестування використовувалися зображення з популярних датасетів, що імітують умови туману (RESIDE-6K), дощу (Rain1400),

снігу (SnowCityScapes), а також реальні відео та зображення з набору weatherStream.

Оцінка продуктивності проводилася як на серверному обладнанні з графічним прискорювачем NVIDIA GeForce GTX 1080 (з підтримкою CUDA), так і на локальному комп'ютері з процесором Intel Core i5 без використання GPU. Для обробки відео застосовувалася методика з пропуском кадрів (frame skipping) з параметром 5, що дозволяє оптимізувати час обробки при збереженні якості результату.

Тестування моделі TANet на сервері з використанням графічного процесора показало стабільно низький час обробки окремих зображень. Час оцінки одного зображення коливався від 0.121 до 0.165 секунд, на вибірці з 5 зображень, зі середнім значенням близько 0.146 секунди.

Для оцінки продуктивності на відео були проведені експерименти з різною тривалістю та якістю відеопотоків. Результати наведені в таблиці 4.11. Час обробки одного кадру варіюється в межах від 0.0182 до 0.0217 секунди, що демонструє високу швидкодію моделі при роботі з відеоданими різної роздільної здатності та тривалості. Найшвидше відбувалась обробка відео з роздільною здатністю 240p, що логічно, враховуючи меншу кількість пікселів у кадрі.

Отже, модель TANet на GPU забезпечує швидке та ефективне виконання задачі видалення погодніх артефактів як для окремих зображень, так і для відео, що є важливим фактором для застосувань у реальному часі.

Таблиця 4.11 – Порівняльна таблиця моделі TANet для оцінки відео на GPU

Тривалість відео, сек	Якість відео, p	Тривалість оцінки, сек	Тривалість оцінки одного кадру, сек
65	720	8.049	0.0205
74	720	7.459	0.0201
74	360	7.006	0.0188
74	240	6.762	0.0182
28	720	3.694	0.0217
28	480	3.461	0.0204
28	360	3.499	0.0206

Модель MWFormer_run2, натренована на реальному наборі даних, також була протестована на сервері з GPU. При оцінці зображень час виконання варіювався від 0.116 до 0.169 секунди, а середній час склав 0.1412 секунди – трохи довше, ніж у TANet, що очікувано через більшу складність архітектури та кількість параметрів.

Для відео результати наведено в таблиці 4.12. Видно, що час оцінки одного кадру суттєво залежить від роздільної здатності відео. Наприклад, відео з роздільною здатністю 240p оброблялось зі швидкістю ~0.0216 секунди на кадр, що близько до показників TANet. Проте для відео у 720p час на кадр значно зростає — до 0.1439 секунди. Це свідчить про значно вищі обчислювальні витрати моделі MWFormer при роботі з відео високої якості.

Таким чином, хоча MWFormer забезпечує вищу якість результатів, він є менш ефективним з точки зору швидкості обробки, особливо на відео високої роздільної здатності. Це робить його менш придатним для задач реального часу, але виправданим у випадках, коли якість оцінки є критично важливою.

Таблиця 4.12 – Порівняльна таблиця моделі MWFormer для оцінки відео на GPU

Тривалість відео, сек	Якість відео, p	Тривалість оцінки, сек	Тривалість оцінки одного кадру, сек
65	720	56.55	0.1439
74	720	53.527	0.1439
74	360	12.985	0.0349
74	240	8.003	0.0216
28	720	24.715	0.1454
28	480	9.752	0.0574
28	360	9.803	0.0577

Для порівняння з результатами на сервері з GPU, було також проведено тестування моделі TANet_exp_002 (35 епоха) на локальному комп'ютері з процесором Intel Core i5, без використання апаратного прискорення (CUDA). Очікувано, час обробки зображень на CPU виявився значно більшим — від 0.529 до 2.309 секунди, із середнім значенням 1.529 секунди на зображення, що майже в 10 разів повільніше, ніж на GPU.

Результати оцінки відео представлені в таблиці 4.13. Час обробки одного кадру коливається від ~9 до ~14 секунд, залежно від роздільної здатності та довжини відео. Наприклад, для відео тривалістю 14 секунд у якості 720p загальний час оцінки становив 1244.980 секунди (більше 20 хвилин), тобто близько 14.8 секунд на кадр. Це ще раз демонструє, що модель TANet не є ефективною для відеообробки в реальному часі на CPU, хоча на GPU забезпечує високу швидкодію.

Таке порівняння підкреслює важливість вибору відповідної апаратної платформи для практичного використання моделі, зокрема у задачах обробки відео.

Таблиця 4.13 – Порівняльна таблиця моделі TANet для оцінки відео на CPU

Тривалість відео, сек	Якість відео, p	Тривалість оцінки, сек	Тривалість оцінки одного кадру, сек
14	720	1244.980	14.8212
9	720	545.967	10.1105
9	360	501.545	9.2879
9	240	556.578	10.3070
6	720	410.415	11.4004
6	480	337.482	9.3745
6	360	416.687	11.5746

Аналогічно було протестовано модель MWFormer_run2 на локальному комп'ютері з Intel Core i5, без використання GPU. Незважаючи на складність архітектури, середній час обробки одного зображення склав 0.608 секунди, що майже в 2.5 рази швидше, ніж у TANet на тій самій платформі (1.529 секунди). Це свідчить про кращу оптимізацію або ефективність реалізації обробки зображень у MWFormer.

При цьому, час обробки відео моделю MWFormer, згідно з Таблицею 4.14, залишається значним. Обробка одного кадру займає від 7.1 до 9.8 секунд, залежно від роздільної здатності та довжини відео. Наприклад, для відео довжиною 14 секунд у якості 720p загальна тривалість оцінки становила 605.471 секунди (близько 10 хвилин), що дає 7.208 секунди на кадр. Це істотно повільніше, ніж результати на GPU, але при цьому все ще майже вдвічі швидше, ніж у TANet на CPU.

Таким чином, MWFormer демонструє кращу продуктивність на CPU в порівнянні з TANet, як для обробки зображень, так і відео. Це робить його більш придатним для задач, де доступ до графічного прискорення обмежений, за умови, що допустимий більший час обробки, а пріоритетом є якість.

Таблиця 4.14 – Порівняльна таблиця моделі MWFormer для оцінки відео на CPU

Тривалість відео, сек	Якість відео, p	Тривалість оцінки, сек	Тривалість оцінки одного кадру, сек
14	720	605.471	7.2080
9	720	389.229	7.2079
9	360	464.308	8.5983
9	240	384.692	7.1239
6	720	354.551	9.8486
6	480	293.239	8.1455
6	360	303.082	8.4189

Проведене тестування на різних апаратних платформах (GPU та CPU) показало суттєву різницю в часі обробки між моделями TANet та MWFormer. Модель TANet демонструє вищу швидкодію, особливо на GPU, де час обробки зображення становить у середньому 0.146 с, а час обробки одного кадру при якості відео 720p у середньому лише 0.0208 с, що робить її зручною для задач реального часу. Натомість модель MWFormer, хоч і повільніша, забезпечує вищу якість очищення зображень, що підтверджується як метриками PSNR і SSIM, так і візуальною оцінкою.

На CPU модель MWFormer також виявилася продуктивнішою за TANet — як у випадку обробки окремих зображень, так і відео, що може бути критично важливим у середовищах з обмеженими обчислювальними ресурсами. Таким чином, вибір моделі залежить від вимог до продуктивності та якості обробки: TANet – для швидкості, MWFormer – для точності. Далі буде розглянуто практичні приклади застосування розробленої підсистеми в реальних умовах.

4.3 Сфера застосування розробленої підсистеми

Розроблена підсистема, що базується на двох сучасних моделях глибинного навчання – TANet та MWFormer, може бути адаптована до широкого спектра практичних застосувань, пов'язаних з покращенням якості зображень і відео в несприятливих погодних умовах, таких як туман, дощ чи сніг. Обидві моделі мають свої переваги, тому доцільно розглядати їх використання у відповідних контекстах.

4.3.1. Застосування моделі TANet

Модель TANet вирізняється високою швидкістю, особливо при використанні GPU (середній час обробки одного зображення – 0.146 с, відео – ~0.02 с/кадр). Це робить її ідеальним кандидатом для задач, де необхідна обробка в режимі реального часу або наближеного до нього.

Можливі сфери застосування TANet:

а) автомобільні системи підтримки водія (ADAS):

- очищення відеопотоку з камер в умовах дощу, снігу чи туману для поліпшення виявлення об'єктів, дорожніх знаків і розмітки;

- робота в реальному часі дозволяє інтегрувати модель у вбудовані системи з GPU-процесорами.

б) відеоспостереження та безпека:

- забезпечення кращої видимості на вулиці за складних погодних умов;
- мінімальний час обробки дозволяє підтримувати безперервну аналітику подій.

в) безпілотні літальні апарати (дрони):

- підвищення якості зображення з бортових камер під час польоту у тумані чи опадах;
- важливо для орієнтації в просторі й запобігання зіткненням.

г) мобільні та ресурсно обмежені пристрої:

– ефективність моделі дозволяє її запуск навіть на пристроях без високопродуктивного обладнання.

4.3.2 Застосування моделі MWFormer

MWFormer, хоча і поступається TANet у швидкості (середній час на GPU – 0.1412 с/зображення, 0.1439 с/кадр відео 720p), забезпечує суттєво вищу якість очищення. Дана модель краще справляється з видаленням складних атмосферних викривлень, таких як сніг чи сильний дощ, та загалом з роботою у більш складних погодних умовах. Це робить її придатною для застосувань, де якість результату має пріоритет над швидкістю.

Можливі сфери застосування MWFormer:

а) професійна обробка відео та зображень:

– використання в кіновиробництві, журналістиці чи архівуванні візуального контенту, де потрібно відновити якість зйомки;

– обробка великої кількості кадрів офлайн із збереженням найвищої точності.

б) геопросторові інформаційні системи (ГІС) та супутникова зйомка:

– покращення якості аерофотознімків, зокрема в хмарну або зимову погоду;

– актуально для моніторингу місцевості, аналізу ландшафту та змін навколишнього середовища.

в) наукові візуальні дослідження:

– застосування у випадках, коли необхідно прибрати артефакти або погодні завади з кадрів, що впливають на аналіз зображень.

г) системи відеоаналітики для транспорту та логістики:

– обробка відео з камер в аеропортах, на вокзалах, складах тощо — навіть за поганої видимості.

Таким чином, розроблена підсистема є гнучкою та масштабованою, і може адаптуватись до різних цілей — від високошвидкісної обробки відео в реальному часі (TANet), до глибокого очищення складних погодних викривлень (MWFormer).

4.4 Перспективи вдосконалення підсистеми

Розроблена підсистема, що використовує моделі TANet та MWFormer для очищення зображень та відео від атмосферних викривлень, уже демонструє високі результати якості оцінки, а також потенціал до застосування в реальних умовах. Проте існує низка напрямків, які дозволять суттєво вдосконалити її ефективність, адаптивність та універсальність.

У межах подальшого розвитку підсистеми важливим напрямом є розширення та вдосконалення навчальних даних, а також збільшення часу тренування моделей. Один із ключових висновків, отриманих під час дослідження, полягає в тому, що як MWFormer, так і TANet демонструють значно гірші результати на зображеннях із реальними сніговими умовами порівняно з синтетичними. У той же час, оригінальні моделі MWFormer, демонструють суттєво вищу ефективність. Це вказує на те, що структура шуму в реальних сніжних сценах значно відрізняється від синтетичних аналогів, і для досягнення високої точності очищення потрібно сформувати більш реалістичний, різноманітний та об'ємний набір прикладів снігу.

Окрім вдосконалення якості наборів даних, варто також розглянути збільшення кількості ітерацій під час тренування, зокрема для моделі MWFormer. Аналіз показав відсутність ознак перенавчання навіть на пізніх етапах навчання, що відкриває можливість для подальшого поліпшення результатів без шкоди для генералізації. Це дозволить моделі краще враховувати складні особливості атмосферних викривлень у реальних умовах, забезпечуючи вищу якість очищення.

Щодо моделі TANet, перспективним є тренування на змішаних наборах синтетичних та реальних даних, оскільки наявні експерименти охоплювали лише окреме навчання на кожному типі даних. Комбінування таких прикладів під час тренування дозволить моделі краще узагальнювати, враховувати як ідеальні

(синтетичні), так і варіативні реальні умови, і таким чином досягати балансу між високою якістю результатів та швидкою обробкою, яка є однією з головних переваг TANet.

Попри вже досягнуті результати, існує потенціал для удосконалення архітектур обох моделей. Зокрема, для TANet доцільним є впровадження легших, оптимізованих версій архітектури, наприклад, мобільних модифікацій або використання механізмів knowledge distillation. Це дозволить ще більше знизити затримку обробки без істотного погіршення якості очищення, що особливо важливо в умовах обмежених ресурсів або потреби в роботі в реальному часі. Для MWFormer перспективним напрямом є інтеграція сучасних attention-механізмів нового покоління, таких як Swin Transformer або EfficientFormer, які вже продемонстрували високу ефективність у задачах обробки зображень. Це може суттєво покращити якість результатів у складних погодних умовах, таких як щільний сніг або комбінація викривлень.

Ще одним напрямом розвитку є інтеграція механізмів динамічного вибору моделі. Аналіз результатів показав, що різні моделі демонструють кращу ефективність залежно від типу атмосферного шуму. Тому доцільним є впровадження автоматичного визначення типу викривлення (класифікації між туманом, дощем, снігом тощо) з подальшим динамічним перемиканням між моделями TANet та MWFormer. Такий підхід дозволить ефективніше використовувати ресурси та досягати оптимального балансу між якістю очищення і продуктивністю в реальному часі, залежно від конкретного сценарію використання.

Для забезпечення практичного застосування системи в реальних умовах особливо важливою є оптимізація продуктивності для мобільних і вбудованих систем. Це включає перенесення моделей у формати, оптимізовані для edge-обчислень, такі як ONNX, TensorRT, CoreML або TFLite. Крім того, використання технік квантування моделей (наприклад, FP16 або INT8) дозволить суттєво зменшити їх розмір та підвищити швидкодію, зберігаючи прийнятний рівень точності. Також перспективним є використання апаратних платформ на зразок

					ІК11.120БАК.006 ПЗ	Арк.
						77
Зм.	Лист	№ докум.	Підпис	Дата		

NVIDIA Jetson, які дозволяють реалізувати обробку відео в режимі реального часу навіть на вбудованих пристроях, що відкриває можливості для інтеграції підсистеми в автономні системи спостереження, транспорт, робототехніку тощо.

У подальшому функціонал підсистеми може бути розширений для виконання суміжних задач, які підвищать її корисність і універсальність. Зокрема, можлива реалізація механізмів виявлення атмосферних перешкод, таких як туман або дощ, що дасть змогу автоматично ідентифікувати умови зйомки. Також можна додати можливість оцінки щільності туману чи інтенсивності дощу, що буде корисним для адаптивного налаштування параметрів обробки або прийняття рішень в автономних системах. Крім того, на основі покращених зображень можлива побудова глибинних карт, що відкриє нові можливості для 3D-аналізу та розпізнавання просторових об'єктів у складних погодних умовах.

Таким чином, підсистема має значний потенціал розвитку — як у напрямі покращення якості, так і в напрямі адаптації до практичного використання. Подальші вдосконалення дозволять забезпечити її застосування в широкому спектрі задач, де потрібна стабільна візуальна інформація в складних погодних умовах, а також ефективна робота на різних апаратних конфігураціях.

Висновки до розділу 4

У цьому розділі було детально представлено результати досліджень, спрямованих на покращення якості вхідного зображення шляхом усунення впливу погодних явищ. Було наведено результати валідації моделей TANet та MWFormer, які показали ефективність у видаленні атмосферних перешкод, таких як туман, дощ та сніг, зокрема з різних тестових наборів як синтетичних, так і реальних даних. Порівняльний аналіз свідчить про те, що MWFormer забезпечує вищі якісні показники за метриками PSNR і SSIM, тоді як TANet відзначається кращою швидкістю, що робить її більш придатною для задач з обмеженими ресурсами.

У наступному підрозділі було проведено тестування розроблених моделей на різних апаратних платформах, зокрема на сервері з GPU NVIDIA GeForce GTX

					ІК11.120БАК.006 ПЗ	Арк.
						78
Зм.	Лист	№ докум.	Підпис	Дата		

1080 та локальній машині з CPU Intel Core i5. Результати продемонстрували суттєві відмінності у швидкодії: TANet забезпечує значно менший час обробки як зображень, так і відео, тоді як MWFormer, попри вищу точність, вимагає більших обчислювальних ресурсів і часу для обробки на графічному процесорі.

Було розглянуто варіанти впровадження моделей у різні практичні системи: від систем допомоги водіям (ADAS) і відеоспостереження до безпілотних літальних апаратів і геопросторових інформаційних систем. При цьому підкреслено, що вибір моделі залежить від конкретних вимог – для задач, де пріоритетом є якість очищення, доцільніше застосовувати MWFormer, тоді як у випадках, що потребують швидкої обробки, перевагу має TANet.

Нарешті, окреслено перспективи вдосконалення підсистеми, зокрема можливості подальшого розширення навчальних даних, збільшення кількості ітерацій тренування, а також архітектурних удосконалень. Запропоновано інтегрувати механізми динамічного вибору моделей залежно від умов зйомки, а також оптимізувати продуктивність для мобільних і вбудованих платформ. Окрім того, підсистема має потенціал для розширення функціоналу, що включає виявлення атмосферних явищ і побудову глибинних карт.

Таким чином, проведені дослідження підтверджують ефективність запропонованих підходів до покращення якості зображень у складних погодних умовах, а також відкривають широкі можливості для практичного застосування та подальшого розвитку розробленої системи.

ВИСНОВКИ

У даному проєкті було досліджено проблему автоматичного покращення якості вхідного зображення мобільного робота за умов атмосферних опадів – туману, дощу та снігу – з метою забезпечення надійного сприйняття та розпізнавання візуальної інформації в реальному часі. Об’єктом дослідження виступали цифрові зображення та відео, спотворені атмосферними явищами, а предметом — алгоритми глибокого навчання для усунення атмосферних артефактів та покращення якості зображення.

У межах роботи було реалізовано повноцінну підсистему комп’ютерного зору, яка інтегрує дві сучасні архітектури глибоких нейронних мереж – TANet і MWFormer. Кожну з моделей було адаптовано до роботи з відео та протестовано на синтетичних і реальних даних, включаючи складні погодні умови. Результати експериментів підтвердили, що MWFormer демонструє високу якість зображення за рахунок глибоких attention-механізмів, а TANet є доцільним вибором для мобільних систем завдяки значно меншому часу обробки кадру.

Під час дослідження досягнуто основних цілей: проаналізовано природу атмосферних спотворень, зібрано збалансовані набори даних із реальних та синтетичних джерел, визначено вимоги до системи, підібрано та модифіковано архітектури, реалізовано та протестовано прототип. Моделі успішно досягають підвищення метрик PSNR та SSIM на тестових наборах, а також забезпечують обробку відео з частотою понад 25 кадрів на секунду на GPU.

Запропоновано варіанти застосування розробленої підсистеми в реальних умовах – від систем допомоги водіям до дронів і відеоспостереження, з акцентом на практичну доцільність використання кожної моделі залежно від ресурсних обмежень та вимог до якості. Крім того, окреслено перспективи подальшого розвитку підсистеми: удосконалення навчальних даних, архітектурні покращення, інтеграція динамічного вибору моделей та адаптація до мобільних платформ.

Таким чином, робота зробила внесок у розвиток систем комп’ютерного зору для несприятливих погодних умов, довівши доцільність використання методів

					ІК11.120БАК.006 ПЗ	Арк.
						80
Зм.	Лист	№ докум.	Підпис	Дата		

глибокого навчання для підвищення надійності автономного функціонування мобільних роботів.

ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Robust object detection in challenging weather conditions. CVF Open Access. URL: https://openaccess.thecvf.com/content/WACV2024/papers/Gupta_Robust_Object_Detection_in_Challenging_Weather_Conditions_WACV_2024_paper.pdf (дата звернення: 08.05.2025).

2. Remote Sensing of Fog: Spatially and temporally continuous estimation and monitoring - Sudoor Manthan- Science Portal of IIRS, ISRO. Sudoor Manthan- Science Portal of IIRS, ISRO. URL: <https://science.iirs.gov.in/remote-sensing-of-fog-spatially-and-temporally-continuous-estimation-and-monitoring/>.

3. ZeroScatter: domain transfer for long distance imaging and vision through scattering media. arXiv.org. URL: <https://arxiv.org/abs/2102.05847>.

4. Multiple linear regression haze-removal model based on dark channel prior. arXiv.org. URL: <https://arxiv.org/abs/1904.11587>.

5. Single image dehazing using haze-lines. researchgate.net. URL: https://www.researchgate.net/publication/329082093_Single_Image_Dehtazing_Using_Haze-Lines.

6. An efficient attentional image dehazing deep network using two color space (admc2-net) - pubmed. PubMed. URL: <https://pubmed.ncbi.nlm.nih.gov/38276379/>.

7. ODD-Net: a hybrid deep learning architecture for image dehazing - Scientific Reports. Nature. URL: <https://www.nature.com/articles/s41598-024-82558-6>.

8. Curricular contrastive regularization for physics-aware single image dehazing. CVF Open Access. URL: https://openaccess.thecvf.com/content/CVPR2023/papers/Zheng_Curricular_Contrastive_Regularization_for_Physics-Aware_Single_Image_Dehtazing_CVPR_2023_paper.pdf.

9. FD-GAN: generative adversarial networks with fusion-discriminator for single image dehazing. arXiv.org. URL: <https://arxiv.org/abs/2001.06968>.

10. Perceptually optimized generative adversarial network for single image dehazing. arXiv.org. URL: <https://arxiv.org/abs/1805.01084>.

					IK11.120БАК.006 ПЗ	Арк.
						82
Зм.	Лист	№ докум.	Підпис	Дата		

11. TransWeather: transformer-based restoration of images degraded by adverse weather conditions. arXiv.org. URL: <https://arxiv.org/abs/2111.14813>.

12. PromptIR: prompting for all-in-one blind image restoration. arXiv.org. URL: <https://arxiv.org/abs/2306.13090>.

13. GitHub - megvii-research/nafnet: the state-of-the-art image restoration model without nonlinear activation functions. GitHub. URL: <https://github.com/megvii-research/nafnet>.

14. Multi-Stage progressive image restoration. arXiv.org. URL: <https://arxiv.org/abs/2102.02808>.

15. GitHub - c-yn/focalnet: [ICCV23] focal network for image restoration. GitHub. URL: <https://github.com/c-yn/FocalNet>.

16. Restoring vision in adverse weather conditions with patch-based denoising diffusion models. arXiv.org. URL: <https://arxiv.org/pdf/2207.14626>.

17. TANet: triplet attention network for all-in-one adverse weather image restoration. arXiv.org. URL: <https://arxiv.org/abs/2410.08177>.

18. MWFormer: multi-weather image restoration using degradation-aware transformers. arXiv.org. URL: <https://arxiv.org/abs/2411.17226>.

19. Oliinyk, V. An efficient face mask detection model for real-time applications / V. Oliinyk, A. Ryzhiy // Адаптивні системи автоматичного управління: міжвідомчий науково-технічний збірник. – 2022. – № 1 (40). – С. 54-64.

20. Oliinyk V. Autonomous car parking model for different types of parking lots using deep reinforcement learning / Oliinyk V., Danyliuk Y. // Адаптивні системи автоматичного управління: міжвідомчий науково-технічний збірник. – 2025. – № 1 (46). – С. 237-246.

21. Oliinyk V. Method for improving accuracy of mobile AR navigators. ISJ Industry 4.0.–2020. Vol. 5, Is. 1. – P. 21-22