

Міністерство освіти і науки України  
Харківський національний університет радіоелектроніки

Факультет Комп'ютерних наук  
(повна назва)

Кафедра Штучного інтелекту  
(повна назва)

## КВАЛІФІКАЦІЙНА РОБОТА Пояснювальна записка

рівень вищої освіти перший (бакалаврський)

Оцінка та усунення шумів у цифрових зображеннях за допомогою  
глибинного навчання  
(тема)

Виконав:  
здобувач четвертого року навчання,  
групи ІТШ-21-5

Дмитро Черногуб  
(власне ім'я, прізвище)

Спеціальність 122 Комп'ютерні науки  
(код і повна назва спеціальності)

Тип програми освітньо-професійна  
Освітня програма Штучний інтелект  
(повна назва освітньої програми)

Керівник проф. Нонна Кулішова  
(посада, власне ім'я, прізвище)

Допускається до захисту

Завідувач кафедри ШІ \_\_\_\_\_  
(підпис)

Олег ЗОЛОТУХІН  
(власне ім'я, прізвище)

2025 р.

Харківський національний університет радіоелектроніки

Факультет \_\_\_\_\_ Комп'ютерних наук \_\_\_\_\_  
(повна назва)  
Кафедра \_\_\_\_\_ Штучного інтелекту \_\_\_\_\_  
(повна назва)  
Рівень вищої освіти \_\_\_\_\_ перший (бакалаврський) \_\_\_\_\_  
Спеціальність \_\_\_\_\_ 122 Комп'ютерні науки \_\_\_\_\_  
(код і повна назва)  
Освітня програма \_\_\_\_\_ Штучний інтелект \_\_\_\_\_  
(повна назва)

ЗАТВЕРДЖУЮ:

Зав. кафедри \_\_\_\_\_  
(підпис)

« \_\_\_\_\_ » \_\_\_\_\_ 20 \_\_\_\_ р.

**ЗАВДАННЯ**  
НА КВАЛІФІКАЦІЙНУ РОБОТУ

здобувачеві \_\_\_\_\_ Чорногубу Дмитру Олександровичу \_\_\_\_\_  
(прізвище, ім'я, по батькові)

1. Тема роботи \_\_\_\_\_ Оцінка та усунення шумів у цифрових зображеннях за допомогою  
глибинного навчання \_\_\_\_\_

затверджена наказом університету від 19 травня 20 25 р. № 378Ст

2. Термін подання студентом роботи до екзаменаційної комісії 18 червня 20 25 р.

3. Вихідні дані до роботи \_\_\_\_\_ науково-технічні публікації з аналізу джерел шуму в цифрових  
зображеннях; матеріали з архітектур нейронних мереж для денойзингу; синтетичні та реальні  
набори даних з різними рівнями шуму та відповідні метадані зображень \_\_\_\_\_

4. Перелік питань, що потрібно опрацювати в роботі \_\_\_\_\_

1) Стан проблеми та постановка задачі дослідження \_\_\_\_\_

2) Методологія оцінки джерел шуму за допомогою глибокого навчання \_\_\_\_\_

3) Експериментальна перевірка оцінки джерел шуму за допомогою глибокого навчання \_\_\_\_\_



## РЕФЕРАТ

Пояснювальна записка: 75 с., 6 рис., 6 табл., 1 дод., 21 джерела.

ГЛИБИННЕ НАВЧАННЯ, ДІАГНОСТИКА СЕНСОРІВ,  
ЗАЛИШКОВИЙ ШУМ, МЕТАДАНІ КАМЕРИ, ОЦІНКА ДЖЕРЕЛ ШУМУ,  
ШУМ ЦИФРОВОГО ЗОБРАЖЕННЯ.

Об'єкт дослідження – цифрові зображення, спотворені шумами, що виникають у сенсорах камер під час зйомки.

Предмет дослідження – методи виявлення та інтерпретації джерел шуму у зображеннях на основі глибинного навчання та метаінформації.

Мета роботи – розробити та експериментально дослідити архітектуру глибокої нейронної мережі, здатної в реальному часі оцінювати рівень та природу шумів у цифрових зображеннях з урахуванням метаданих камери.

Методи дослідження – математичне моделювання, побудова багатогілкових нейронних мереж, аналіз залишкового шуму, порівняльне експериментальне дослідження.

У роботі розглянуто підхід до оцінки шумів у цифрових зображеннях, який передбачає використання глибинного навчання у поєднанні з технічною інформацією про умови зйомки. Побудовано модель, що окремо аналізує фотонний шум, шум темного струму та шум зчитування, а також виявляє залишкову складову, що не пояснюється доступними метаданими. Проведено серію експериментів на синтетичних та реальних даних, що підтвердили здатність моделі точно відокремлювати джерела шуму. Досліджений підхід може бути інтегрований у системи відеоспостереження, діагностики сенсорів і вбудовані камери автономних пристроїв.

## **ABSTRACT**

Bachelor's thesis contains: 75 pp., 6 fig., 6 tabl., 1 ann., 21 references.

CAMERA METADATA, DEEP LEARNING, DIGITAL IMAGE NOISE, NOISE SOURCE ESTIMATION, RESIDUAL NOISE, SENSOR DIAGNOSTICS.

Object of research – digital images distorted by noise originating from camera sensors during acquisition.

Subject of research – methods for detecting and interpreting noise sources in images using deep learning and camera metadata.

Aim of the work – to develop and experimentally investigate a deep neural network architecture capable of real-time assessment of the level and nature of noise in digital images, considering camera parameters.

Research methods – mathematical modeling, multi-branch neural network design, residual noise analysis, comparative experimental evaluation.

This thesis proposes a novel approach to noise estimation in digital images, combining deep learning with metadata about camera settings. A model is designed to independently estimate photon noise, dark current noise, and readout noise, while also detecting a residual noise component unexplained by metadata. A series of experiments on synthetic and real datasets confirm the model's ability to accurately separate noise sources and detect anomalies. The proposed method can be integrated into video surveillance systems, sensor diagnostics, and embedded imaging modules of autonomous devices.

## ЗМІСТ

Перелік умовних позначень, символів, одиниць, скорочень і термінів .....	7
Вступ .....	8
1 Стан проблеми та постановка задачі дослідження.....	10
1.1 Проблеми шуму в цифрових зображеннях .....	10
1.2 Методи оцінки шуму в цифрових зображеннях .....	15
1.2.1 Традиційні методи оцінки шуму.....	16
1.2.2 Методи на основі глибокого навчання.....	21
1.3 Методи усунення шуму .....	24
1.3.1 Класичні методи роботи з шумом.....	25
1.3.2 Глибокі нейромережеві підходи в задачі денойзінгу.....	28
1.4 Використання метаданих камери для оцінки шуму.....	31
1.4 Постановка задачі дослідження .....	34
2 Методологія оцінки джерел шуму на основі глибинного навчання.....	36
2.1 Загальна концепція та архітектура системи.....	36
2.2 Розширення для ідентифікації джерел шуму.....	39
2.3 Розширення для ідентифікації джерел шуму.....	41
2.4 Інтеграція метаданих камери.....	44
2.5 Виявлення залишкового шуму .....	46
2.6 Методика навчання моделі .....	49
3 Експериментальна перевірка оцінки джерел шуму на основі глибинного навчання.....	51
3.1 Огляд наборів даних.....	51
3.2 Попередня обробка зображень і метаданих .....	56
3.3 Реалізація моделі та конфігурація навчання.....	58
3.4 Результати експериментів.....	60
Висновки.....	70
Перелік джерел посилання .....	72
Додаток А Відомість кваліфікаційної роботи.....	75

## ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ

BM3D – Block-Matching and 3D Filtering – блочне зіставлення та 3D-фільтрація;

CCD – Charge-Coupled Device – пристрій із зарядовим зв'язком;

CMOS – Complementary Metal–Oxide–Semiconductor – комплементарна металооксидна напівпровідникова технологія;

CNN – Convolutional Neural Network – згорткова нейронна мережа;

DCSN – Dark Current Sensor Noise – шум темного струму;

DnCNN – Denoising Convolutional Neural Network – згорткова нейронна мережа для зниження шуму;

FFDNet – Fast and Flexible Denoising Network – швидка та гнучка мережа для зниження шуму;

FSID – Fully Synthetic Image Denoising – повністю синтетичний набір даних для зниження шуму;

ISO – International Standards Organization – рівень світлочутливості сенсора;

KITTI – Karlsruhe Institute of Technology and Toyota Technological Institute – набір даних KITTI (для автономного водіння);

MRT – Magnetic Resonance Tomography – магнітно-резонансна томографія;

NLM – Non-Local Means – нелокальне усереднення;

PN – Photon Noise – фотонний шум;

RN – Readout Noise – шум зчитування;

SNR – Signal-to-Noise Ratio – співвідношення сигнал/шум;

TiMo – Time-of-Flight Indoor Monitoring – внутрішній моніторинговий набір даних з ToF-сенсорів.

## ВСТУП

Сьогодні цифрові зображення оточують нас буквально всюди, у смартфонах, відеокамерах, дронах, системах безпеки, медицині й навіть побутовій техніці. Камера вже давно перестала бути просто пристроєм для фотографій, вона перетворилась на сенсор, який дивиться на світ і допомагає системам приймати рішення. Проте існує така проблема – не всі зображення ідеальні. Навіть найкращі камери можуть знімати з певними шумами, але дуже помітними спотвореннями, які з'являються через фізичні процеси в сенсорі, нагрівання, або просто через складні умови зйомки.

У багатьох випадках ці шуми не заважають адже людина може їх не помітити. Але коли зображення обробляє не людина, а штучний інтелект, який має розпізнати об'єкт, проаналізувати сцену або прийняти рішення, в такому випадку, будь-яка помилка може бути критичною. Ще складніше, коли шум не один, а їх кілька, і всі змішуються: фотонний, тепловий, зчитування. Кожен з них має свою причину появи та має свою характерну поведінку.

Частим підходом до роботи із шумом – фільтрація з використанням відповідних методів таких як, фільтрування, згладжування, або інтелектуальними методами машинного навчання. Проте навіть найсучасніші денойзери рідко пояснюють, чому виник шум і що саме завадило камері зняти краще. Наприклад, якщо камера системи безпеки постійно продукує шум через перегрів або зношення сенсора – це сигнал, що існують технічні проблеми з обладнанням.

Останніми роками тенденція досліджень переходить у поєднання сучасних підходів з машинним навчанням із розумінням фізики самої зйомки. Зокрема, все більше уваги приділяється метаданим – технічній інформації про те, в яких умовах камера зробила знімок: яка була експозиція, температура, які налаштування використовувались. Ідея полягає в тому, щоб модель бачила не лише картинку, а й контекст її появи.

Саме це завдання дозволяє зрозуміти, чому в зображенні присутній шум, і яка його природа, що стало основою цієї роботи. Основна фокус роботи присвячений дослідженню системи, що можуть аналізувати зображення разом з метаданими, розділяти типи шумів і, можливо, навіть попереджати, якщо з камерою щось не так. Це важливо не лише з точки зору якості зображення, а й для побудови більш надійних систем обробки даних.

## 1 СТАН ПРОБЛЕМИ ТА ПОСТАНОВКА ЗАДАЧІ ДОСЛІДЖЕННЯ

На сьогоднішній день, цифрові зображення використовуються у багатьох сферах, наприклад, комп'ютерний зір, медична діагностика, відеоспостереження, аерофотозйомка та автономний транспорт. Проте, як правило, якість отриманих зображень погіршується саме через наявність шумів, які виникають під час їх створення з різних сенсорів, таких як камера, МРТ тощо. Серед таких недоліків зображення можуть бути теплові ефекти, шум спричинений радіацією, випадковий характер надходження фотонів на сенсор, обмеження фізичного сенсору при зчитуванні сигналу та особливості аналого-цифрового перетворення. Такі шуми ускладнюють подальшу обробку та аналіз, що відбувається за допомогою автоматизованих систем розпізнавання та класифікації об'єктів.

Проблема шуму в цифрових зображеннях особливо актуальна для методів комп'ютерного зору, оскільки навіть незначні спотворення можуть спричинити помилки або пониження точності в рамках роботи алгоритмів штучного інтелекту. Наприклад, автономні транспортні засоби можуть неправильно класифікувати дорожні знаки через високий рівень шуму, а автоматизовані системи в задачі медичної діагностики, при наявності зазначених недоліків схильні неправильно інтерпретувати рентгенівські чи МРТ-знімки. Тому розробка методів як оцінки рівня шуму, так і його усунення важливі завдання для підвищення надійності таких систем

### 1.1 Проблеми шуму в цифрових зображеннях

Формування цифрового зображення являє собою складний процес, де приймають участь оптична система, світлочутливий сенсор, аналого-цифровий перетворювач та обчислювальні алгоритми. Внаслідок ряду фізичних, електронних та алгоритмічних факторів, з'являються шуми у

отриманих зображеннях, тобто випадкові або систематичні викривлення пікселів, що змінюють їхні значення порівняно з оригінальною сценою.

Шуми в цифрових камерах можна, як правило, класифікувати за їхнім походженням та природою, наприклад фотонний шум, темновий шум, квантовий шум, шум зчитування, та підсилення, тощо [1]. Фотонний шум класичним обмеженням усіх оптичних сенсорів і виникає через випадковий характер потрапляння фотонів на матрицю камери. Його кількість, рівень прямо залежить від кількості захопленого світла, тобто при недостатньому освітленні цей тип шуму стає особливо помітним.

Темновий шум виникає через випадкову генерацію електронів у сенсорі, навіть за відсутності світла. Такий процес спричинений тепловими ефектами, тому його тенденція полягає у зростанні зі збільшенням температури сенсора. Він особливо впливає на якість зображень, що були отримані при недостатньому освітленні, наприклад, уночі або при тривалих витримках.

Шум зчитування це результатом електронних процесів під час перетворення світлових сигналів у цифрові значення. Він спричинений недосконалістю електронних схем сенсора та перетворювачів, що може призводити до хаотичних коливань значень пікселів навіть у стабільних умовах освітлення.

Ще одним типом шуму – квантовий шум, який виникає під час аналого-цифрового перетворення сигналу. Оскільки кожен піксель має обмежену кількість градацій яскравості, округлення значень під час дискретизації спричиняє втрати точності та появу штучних артефактів, особливо в зображеннях із низькою розрядністю кольору.

Нарешті, шум підсилення характерний для зйомки, коли чутливість датчика збільшується до світла, з метою покращення видимості в умовах слабого освітлення, одночасно підсилюються і випадкові флуктуації сигналу, що призводить до значного погіршення якості зображення.

Кожен із цих типів шуму має різні характеристики і механізми впливу на цифрові зображення. Подальший аналіз дозволить визначити, як саме вони впливають на комп'ютерний зір, розпізнавання об'єктів та інші задачі обробки візуальної інформації.

Наявність шуму в цифрових зображеннях значно впливає на їхню якість, оскільки спотворює кольорову та яскравішу інформацію, порушує чіткість контурів і зменшує деталізацію. Це створює проблеми не лише для сприйняття людиною, але й для автоматизованих алгоритмів аналізу зображень, які використовуються в комп'ютерному зорі, медичній діагностиці, автономному транспорті та інших сферах.

Одним із найбільших викликів є той факт, що шум може бути неоднорідним і мати нелінійний характер [1]. Наприклад, фотонний шум помітніший у темних зонах зображення, а шум підсилення проявляється найсильніше при високих значеннях ISO. Це ускладнює завдання стандартних алгоритмів, які не враховують варіативність шуму у різних умовах зйомки.

Шум особливо критичний у системах комп'ютерного зору [2], де алгоритми розпізнавання об'єктів і класифікації зображень можуть давати хибні результати. Наприклад, в автономному транспорті камери повинні точно ідентифікувати дорожні знаки та пішоходів. Якщо зображення містить високий рівень шуму, модель може неправильно визначити об'єкт або зовсім його не розпізнати, що може призвести до аварійної ситуації.

Подібні проблеми виникають у медичних зображеннях, де шум у рентгенівських або МРТ-знімках може створювати артефакти, що заважають лікарям правильно інтерпретувати результати обстеження [2]. Низька якість зображення може ускладнити виявлення патологій та знизити точність діагностики.

В області супутникових знімків і аерофотозйомки шум також може бути серйозною перешкодою. Наприклад, тепловий шум уночі або в умовах поганої видимості може призводити до втрати важливих деталей у знімках,

що використовуються для прогнозування погодних умов або моніторингу змін навколишнього середовища.

Крім того, високий рівень шуму у відеоспостереженні може ускладнити розпізнавання обличчя або об'єктів на кадрах із поганим освітленням. Це особливо критично у сфері безпеки, де якість запису може визначати можливість ідентифікації особи чи події.

Таким чином, шум це серйозний фактор, який знижує ефективність обробки та аналізу зображень у багатьох сферах застосування. Його усунення або мінімізація є необхідною умовою для покращення роботи алгоритмів комп'ютерного зору, підвищення точності медичних досліджень та забезпечення безпеки в автоматизованих системах. У наступному підрозділі розглядатимуться конкретні випадки, у яких шум у зображеннях спричинив серйозні проблеми, що підкреслює важливість розробки ефективних методів його оцінки та усунення.

Високий рівень шуму в цифрових зображеннях може мати серйозні наслідки в реальних системах, що працюють з візуальною інформацією. Це стосується не лише погіршення якості зображень, але й помилкових рішень, які можуть бути критичними для безпеки, точності аналізу або ефективності алгоритмів [2]. У цій частині розглядаються конкретні випадки, коли шум у цифрових зображеннях спричинив значні проблеми у різних сферах застосування.

Однією з найбільш відомих проблем, пов'язаних із шумом, це помилкове розпізнавання дорожніх знаків в автономному транспорті. Наприклад, деякі моделі автомобілів Tesla, що працюють на автопілоті [3], неправильно ідентифікували знаки обмеження швидкості через артефакти шуму або відблиски на дорожніх табличках. У реальних тестах автомобіль приймав розмитий напис «Speed limit 35» за «Speed limit 85», що могло спричинити небезпечну ситуацію. Подібні випадки демонструють, як навіть незначні спотворення у зображеннях можуть призводити до помилкових рішень у критично важливих системах.

Іншим яскравим прикладом – вплив шуму на системи розпізнавання обличчя у відеоспостереженні. У низькоякісних камерах безпеки, що працюють у темний час доби або за поганих погодних умов, рівень шуму може бути настільки високим, що система не здатна правильно розпізнати людину або навіть відрізнити реальну особу від фонового шуму. Це було помітно в кримінальних розслідуваннях, коли поліцейські не могли використати записи камер для точної ідентифікації злочинців через надмірний рівень шуму та низьку роздільну здатність зображення.

Шум також створює значні труднощі у сфері медичної діагностики. Наприклад, в аналізі рентгенівських знімків або МРТ-шарів артефакти, спричинені шумом, можуть маскувати патологічні зміни в тканинах, що ускладнює діагностику ракових пухлин або серцево-судинних захворювань. У деяких випадках лікарі отримували хибнопозитивні результати, приймаючи шумові артефакти за патологічні утворення, або навпаки – не помічали важливих аномалій, що могло призводити до затримки лікування.

У сфері супутникової зйомки шум також відіграє важливу роль. Наприклад, у метеорології низькоякісні знімки, отримані за умов слабого сигналу, можуть мати значні артефакти, що ускладнює прогнозування погодних змін. Також у супутникових системах моніторингу лісових пожеж або розливу нафти шум може створювати фальшиві зони виявлення або приховувати реальні проблеми, що призводить до затримки реакції на екологічні катастрофи.

Ще один випадок – використання шумозахищених зображень у цифровій фотографії та відеопродукції. Професійні фотографи часто стикаються з проблемою шуму при зйомці в умовах слабого освітлення. Хоча сучасні камери використовують алгоритми придушення шуму, агресивна обробка може призводити до втрати деталей та спотворення текстур, що негативно позначається на кінцевому зображенні. У кіноіндустрії, особливо при зйомках із високою чутливістю до світла, шум

може робити зображення менш привабливим та вимагати складної постобробки.

Таким чином, робота з шумом у цифрових зображеннях являє собою актуальною, на сьогоднішній день, проблемою, яка впливає на багато сфер від безпеки та транспорту до медицини та наукових досліджень. Помилки, що викликані шумом, можуть мати значні наслідки, тому розробка ефективних методів його оцінки та усунення є необхідною умовою для підвищення точності комп'ютерного зору та обробки зображень.

## 1.2 Методи оцінки шуму в цифрових зображеннях

Оцінка рівня шуму – ключове завдання у процесі покращення якості цифрових зображень та їхньої подальшої обробки. Точне визначення характеристик шуму дозволяє правильно підібрати методи фільтрації [4], адаптивно коригувати параметри обробки та зменшувати спотворення важливих деталей. Методи оцінки шуму поділяються на традиційні, які базуються на аналізі статистичних властивостей зображень, та методи глибокого навчання, що використовують нейромереві підходи для автоматичного виявлення й моделювання шуму.

Традиційні методи включають фільтраційні підходи, що аналізують поведінку пікселів після згладжування, перетворювальні методи, які працюють у частотній області, та статистичні підходи, що визначають рівень шуму через оцінку розподілу значень пікселів. Вони широко використовуються завдяки своїй інтерпретованості та відносній простоті реалізації, проте часто мають обмеження при роботі зі складними текстурними сценами або за наявності неоднорідного шуму.

Завдяки розвитку глибокого навчання з'явилися методи [5], [6], які дозволяють оцінювати шум на основі великої кількості прикладів реальних зображень. Глибокі нейронні мережі мають змогу навчатися на зашумлених даних і виявляти патерни шуму, що недоступні традиційним підходам.

Автоенкодерери та генеративні моделі дозволяють будувати гнучкі представлення шумових компонентів, що значно покращує точність оцінки.

### 1.2.1 Традиційні методи оцінки шуму

Три основні групи традиційних методів оцінки шуму включають фільтраційні, трансформаційні та статистичні підходи.

Фільтраційні методи оцінюють рівень шуму шляхом згладжування зображення та аналізу локальних змін яскравості. До них належать усереднювальні, медіанні та білатеральні фільтри, які використовуються для виділення шумових компонентів, збереження деталей і зменшення артефактів.

Трансформаційні методи базуються на аналізі частотних характеристик зображення, в результаті чого можна відокремити низько- та високочастотні складові. Зокрема, перетворення Фур'є дає змогу виявляти та фільтрувати періодичні шуми [7], тоді як вейвлет-перетворення забезпечує оцінку шуму у різних масштабах, зберігаючи просторову локалізацію змін сигналу.

Статистичні методи використовують властивості розподілу піксельних значень для оцінки рівня шуму [4]. Вони аналізують гістонрами яскравості, локальні статистичні характеристики, а також взаємозв'язки між сусідніми пікселями. Ці підходи ефективні для оцінки гаусового шуму, але менш точні при складних, неоднорідних шумах.

Фільтраційні методи оцінки рівня шуму є одними з найпростіших і найпоширеніших підходів, що базуються на згладжуванні зображення та аналізі його локальних змін. Основна ідея полягає у тому, що природні текстури та об'єкти в зображенні мають корельовану структуру, тоді як шум зазвичай випадковий і не має просторової залежності, і, призводить використання варіацій згладжувальних фільтрів для виділення шумових компонентів і подальшого їх аналізу.

Найпростіший підхід до оцінки рівня шуму заснований на використанні усереднювальних фільтрів [8], які застосовують згортку з ядром, що обчислює середнє значення пікселів у локальному вікні. Якщо після згладжування залишається значний рівень варіацій, це свідчить про присутність вираженого шуму. Проте цей метод має обмежену ефективність, оскільки розмиває текстурні деталі разом із шумом.

Більш досконалий метод це медіанний фільтр [9], який замінює значення кожного пікселя на медіану значень його сусідів у заданій області, що дає змогу працювати з імпульсним шумом, видаляючи його, який проявляється у вигляді різких точкових змін у зображенні. Однак медіанний фільтр менш ефективний при роботі зі складними текстурними сценами, оскільки він не враховує особливості геометрії зображення.

Один із найефективніших традиційних методів – білатеральний фільтр [10], який використовує подвійне згладжування: за просторовими координатами пікселів та їхньою інтенсивністю. Це означає, що пікселі, які мають схожі значення яскравості, будуть згладжуватися сильніше, ніж ті, що мають різку зміну кольору. Такий підхід дозволяє ефективно усувати шум, зберігаючи при цьому важливі контури та деталі об'єктів.

Також адаптивні фільтри [11], як ще один варіант, аналізують локальні характеристики зображення та підлаштовують свої параметри відповідно до рівня варіацій у піксельних значеннях. Наприклад, фільтр Вінера оцінює рівень шуму в локальному регіоні та відповідно коригує коефіцієнти згладжування, що дає змогу зменшувати шум, не розмиваючи важливі деталі.

Попри свою ефективність у випадках, коли шум рівномірний та випадковим, фільтраційні методи мають певні обмеження. Вони менш ефективні при роботі з неоднорідним або структурованим шумом, який може змінюватися залежно від області зображення. Крім того, традиційні фільтри зазвичай не можуть ефективно оцінювати рівень шуму у складних сценах із текстурованими об'єктами.

Таким чином, фільтраційні методи є важливим інструментом для оцінки шуму, проте в сучасних умовах їх дедалі частіше доповнюють або замінюють більш гнучкими підходами, зокрема нейромережевими методами, які можуть враховувати складні патерни шуму в зображенні.

Трансформаційні методи оцінки рівня шуму базуються на аналізі частотних характеристик зображення. Вони дозволяють відокремити корисний сигнал від шуму, використовуючи частотне розкладання, оскільки більшість природних зображень мають значну енергію на низьких частотах, тоді як шум зазвичай домінує на високих частотах. Два найбільш поширені підходи в цій категорії – перетворення Фур'є та вейвлет-перетворення, які використовуються для оцінки та фільтрації шуму.

Перетворення Фур'є дозволяє представити зображення у частотній області, де воно описується як сума гармонічних функцій із різними частотами та амплітудами. У цьому представленні шум зазвичай концентрується у високочастотній області спектра, сприяючи його оцінити шляхом аналізу коефіцієнтів Фур'є.

Одним із ключових методів оцінки рівня шуму це порівняння енергетичного розподілу частотних компонентів. Високочастотні складові, що мають аномально високу амплітуду, можуть бути ідентифіковані як шум. На основі цього підходу реалізуються частотні фільтри, такі як низькочастотні фільтри (які згладжують шум, видаляючи високі частоти) та смугові фільтри (що видаляють певні частотні діапазони).

Головним недоліком методу Фур'є те, що він не враховує локальні зміни у зображенні. Будь-яке частотне представлення відображає весь сигнал глобально, що може ускладнювати точне визначення шуму у зображеннях із різкими переходами або деталями.

Вейвлет-перетворення це вдосконаленою альтернативою Фур'є-перетворення, оскільки воно дозволяє аналізувати сигнал як у частотному, так і в просторовому доменах. Це досягається шляхом розкладу зображення на множину масштабованих і зсунутих версій базових функцій (вейвлетів).

На відміну від Фур'є-перетворення, вейвлети дозволяють оцінювати шум у локальних областях зображення. Це особливо корисно при роботі із шумами, що варіюються в різних частинах зображення. Наприклад, методи дискретного вейвлет-перетворення застосовуються для багатомасштабного аналізу, де зображення ділиться на низькочастотні та високочастотні компоненти. Високочастотні коефіцієнти, які містять значний рівень випадкових змін, можуть бути проаналізовані для оцінки рівня шуму.

Один із популярних підходів полягає у пороговій обробці вейвлет-коефіцієнтів. Коефіцієнти, які перевищують певний поріг, вважаються інформаційними, тоді як значення, що мають низьку амплітуду, відносяться до шуму. Таким чином, можливість точно оцінити рівень шуму та зменшити його вплив на зображення без значних втрат деталей.

Основними перевагами вейвлет-перетворення, можливість аналізу шуму на різних масштабах і точне виявлення шумових компонентів без значного спотворення сигналу. Проте цей метод складніший у реалізації та потребує підбору оптимальних параметрів, таких як тип вейвлетів і рівень декомпозиції.

Обидва методи мають свої переваги та обмеження. Перетворення Фур'є добре працює для глобального аналізу шуму, але не дозволяє оцінювати його локальні варіації. Вейвлет-перетворення, навпаки, забезпечує кращу просторову локалізацію шуму, роблячи його більш ефективним для складних текстурних зображень. Однак складність обчислень та необхідність вибору параметрів можуть ускладнювати його застосування.

Таким чином, трансформаційні методи [8] – потужний інструмент для оцінки рівня шуму та можуть бути використані як самостійно, так і в комбінації з іншими методами, зокрема статистичними або неймережевими підходами, для підвищення точності аналізу.

Статистичні підходи до оцінки рівня шуму базуються на аналізі розподілу значень пікселів у зображенні, їхніх взаємозв'язків та

ймовірнісних характеристик. Основна ідея таких методів полягає у припущенні, що природні зображення мають певні статистичні властивості, тоді як шум є випадковим процесом, який змінює ці характеристики. Визначаючи відхилення від очікуваних розподілів, можна оцінити рівень шуму та його характер.

Одним із найпростіших статистичних методів це аналіз гистограми яскравості пікселів. У випадку, якщо шум рівномірно розподілений по всьому зображенню, його вплив можна оцінити за шириною гистограми. Чим більший рівень шуму, тим більш розмитою стає гистограма, що свідчить про велику варіативність яскравісних значень. Випадковий шум зазвичай призводить до збільшення дисперсії інтенсивності пікселів, що може бути використано для його кількісної оцінки.

Інший підхід ґрунтується на аналізі локальних статистичних характеристик, таких як середнє значення, стандартне відхилення та коефіцієнти кореляції між сусідніми пікселями. Оскільки в більшості природних зображень сусідні пікселі мають подібні значення, різке збільшення дисперсії або зменшення кореляції у локальних областях може свідчити про наявність шуму. Для визначення рівня шуму часто використовується оцінка середньоквадратичного відхилення у порівнянні зі згладженими версіями зображення.

Більш складні методи статистичного аналізу включають використання стохастичних моделей, таких як гаусівські та пуассонівські розподіли, що дозволяють апроксимувати поведінку шуму в різних типах сенсорів. Наприклад, метод максимальної правдоподібності використовується для оцінки параметрів шумового розподілу, допомагаючи визначити інтенсивність та характер шуму.

Окремо слід виділити методи, засновані на аналізі вищих статистичних моментів, таких як ексцес та асиметрія розподілу яскравості. Якщо зображення містить лише природні текстури, його розподіл значень

пікселів буде мати певну симетрію, тоді як шум може змінювати ці властивості, що використовується для його ідентифікації.

Статистичні методи оцінки шуму ефективні та прості у реалізації, проте вони мають обмеження. Наприклад, якщо шум не рівномірно розподілений або має структуровану природу, такі підходи можуть давати неточні оцінки. Крім того, вони менш ефективні у випадках, коли шум корелює з текстурою зображення. Через це статистичні методи часто комбінуються з іншими підходами, такими як частотний аналіз або нейромереві методи, для досягнення більш точної оцінки рівня шуму.

### 1.2.2 Методи на основі глибокого навчання

Глибокі нейромереві [12] відкрили нові можливості для оцінки рівня шуму в цифрових зображеннях, значно перевершуючи традиційні методи завдяки здатності до автоматичного виявлення та моделювання складних патернів. Використовуючи глибоке навчання, такі підходи дозволяють аналізувати шум у контексті всього зображення, враховуючи його просторові особливості та взаємозв'язки між пікселями. До найбільш ефективних методів у цій категорії належать глибокі згорткові нейромереві, які вчаться розпізнавати та відокремлювати шумові компоненти, а також автокодері та генеративні моделі, що будують компактне представлення зображення та можуть реконструювати його без шуму. Завдяки своїй здатності адаптуватися до різних типів шуму без необхідності ручного налаштування, ці підходи стали ключовими в сучасних системах обробки зображень.

Глибокі згорткові нейромереві це сучасним і високоефективним підходом до оцінки рівня шуму в цифрових зображеннях. Вони використовують глибоку архітектуру згорткових шарів, автоматично витягують багаторівневі особливості зображення та виявляти шуми без необхідності явного моделювання їхнього розподілу. Це значно відрізняє їх

від традиційних методів, які зазвичай спираються на статистичні чи частотні припущення про шум.

Однією з найпопулярніших нейромережевих архітектур для оцінки рівня шуму, DnCNN [13]. Ця модель складається з кількох згорткових шарів, що поступово витягують ознаки зображення та дозволяють відокремити корисний сигнал від шуму. DnCNN навчається на великих наборах зображень, які містять як зашумлені, так і чисті версії. Під час навчання мережа вчиться передбачати карту шуму, яку потім можна використовувати для його оцінки або усунення.

Ще одним ефективним підходом FFDNet [14], який використовує адаптивні фільтри, що дозволяють налаштовувати рівень шуму залежно від характеристик зображення. Ця модель працює швидше за DnCNN, оскільки дозволяє проводити оцінку та фільтрацію шуму у підвбірках зображень, зменшуючи обчислювальні витрати.

Інші підходи, такі як Noise2Noise [15] і Noise2Void [16], дозволяють оцінювати шум без використання чистих зображень у навчальних наборах. Noise2Noise навчається відновлювати зображення, використовуючи дві незалежні реалізації шуму для одного і того ж зразка. Noise2Void, у свою чергу, базується на принципі навчання без потреби в чистих зображеннях, використовуючи контекстну інформацію сусідніх пікселів.

Головною перевагою CNN є їхня здатність адаптуватися до складних та неоднорідних шумів, що особливо важливо для реальних зображень, де шум може змінюватися залежно від джерела та характеристик камери. На відміну від класичних методів, які часто припускають рівномірний розподіл шуму, нейромережі можуть ефективно працювати з різними варіаціями шумових патернів, забезпечуючи значно кращу точність оцінки. Завдяки своїй гнучкості та здатності навчатися на великих вибірках даних, згорткові нейромережі є одним із найперспективніших інструментів для оцінки рівня шуму та покращення якості цифрових зображень.

Автокодери та генеративні моделі являють собою сучасні підходи до оцінки рівня шуму та його усунення в цифрових зображеннях. Вони базуються на глибоких нейромережах і здатні ефективно аналізувати структуру зображень, відокремлюючи корисну інформацію від шумових компонентів. Завдяки здатності до навчання на великих вибірках даних, ці методи значно перевершують традиційні підходи, особливо у випадках, коли рівень шуму неоднорідний або важко моделюється класичними статистичними методами.

Автокодери, як клас нейромереж, що складаються з двох основних частин: кодера, який стискає вхідне зображення у зменшене представлення (латентний простір), і декодера, який відновлює зображення зі збереженої інформації. При використанні автокодерів для оцінки шуму модель навчається відновлювати чисте зображення із зашумленого, виділяючи латентні ознаки, що містять тільки корисну інформацію. Наприклад, Denoising Autoencoder [17] варіація автокодера, спеціально розроблена для видалення шуму. Він навчається на парах «зашумлене – чисте зображення» і вчиться реконструювати сигнал без шумових компонентів.

Більш просунуті методи включають варіаційні автокодери, які використовують ймовірнісне представлення латентного простору, для подальшого моделювання більш складні розподіли шуму. Завдяки цьому варіаційні автокодери можуть використовуватися не лише для денойзингу, але і для оцінки характеристик шуму в реальних даних, що дає не тільки корисність, а і потенціал для адаптивної обробки зображень.

Генеративні змагальні мережі є ще одним потужним інструментом у сфері оцінки та усунення шуму. Вони складаються з двох нейромереж: генератора, який створює нові зображення, і дискримінатора, який оцінює, наскільки отримане зображення реалістичне. Генеративні мережі можуть навчатися відновлювати зображення, мінімізуючи шумові компоненти, і навіть генерувати синтетичні версії чистих зображень, що дає змогу

оцінювати рівень шуму шляхом порівняння оригіналу та відновленого зображення. Наприклад, Noise2Noise генеративна мережа використовує методику глибокого навчання без необхідності мати еталонне чисте зображення.

Перевагою автокодерів та генеративних моделей є їхня здатність працювати зі складними та нерівномірними шумами, які важко оцінити традиційними методами. Вони здатні адаптуватися до різних характеристик зображень і навіть навчатися на нових наборах даних без необхідності знання точного розподілу шуму. Це робить їх перспективним напрямком досліджень для підвищення якості цифрових зображень та розширення можливостей сучасних систем комп'ютерного зору.

### 1.3 Методи усунення шуму

Усунення шуму – важливе завдання у цифровій обробці зображень, оскільки його присутність може суттєво погіршувати якість візуальних даних та знижувати ефективність алгоритмів комп'ютерного зору. Традиційні методи денойзингу базуються на фільтраційних підходах, які згладжують шум, зберігаючи при цьому ключові деталі зображення. До них належать методи середнього, медіанного та Гаусового згладжування, а також більш складні алгоритми, такі як BM3D і NLM, що враховують подібність між локальними регіонами.

Однак класичні підходи мають обмеження, особливо при роботі зі складними текстурними зображеннями та неоднорідним шумом. Завдяки розвитку глибокого навчання з'явилися нейромережеві методи, які використовують багаторівневу обробку даних для більш точного усунення шуму. До таких моделей належать DnCNN, FFDNet, Noise2Noise та Blind2Unblind, які демонструють значно кращі результати, адаптуючись до різних рівнів та типів шуму.

### 1.3.1 Класичні методи роботи з шумом

Класичні методи денойзингу ґрунтуються на припущенні, що шум має випадковий характер, а корисний сигнал є структурованим, що сприяє виділенню його шляхом згладжування або аналізу подібності між регіонами зображення. Хоча такі методи ефективні для видалення простих типів шуму, вони часто спричиняють втрату дрібних деталей або розмиття контурів, що обмежує їхню застосовність у складних візуальних сценах.

Фільтри середнього, медіанного та Гаусового згладжування є базовими методами усунення шуму, які використовуються в обробці зображень для зменшення випадкових флуктуацій яскравості пікселів. Вони працюють за принципом усереднення або підбору оптимального значення для кожного пікселя, ґрунтуючись на сусідніх значеннях у локальному вікні. Такі методи є ефективними для зменшення певних типів шуму, проте можуть спричинити втрату деталей та розмиття контурів, що обмежує їхню точність у складних зображеннях.

Фільтр середнього, являє собою найпростіший метод згладжування, який замінює кожен піксель середнім арифметичним значенням його сусідів у певному вікні (наприклад,  $3 \times 3$  або  $5 \times 5$ ). Такий фільтр дозволяє зменшувати рівень гаусового шуму, однак через те, що всі пікселі розглядаються рівноправно, середній фільтр не враховує структуру зображення. Внаслідок цього він може викликати значне розмиття контурів та деталей, особливо при використанні великих вікон згладжування.

Медіанний фільтр більш стійкий до імпульсного шуму, який виникає через випадкові спотворення окремих пікселів. Замість усереднення, цей фільтр замінює значення пікселя на медіану яскравості сусідніх пікселів у заданому вікні, що призводить до ефективного зменшення шумових артефактів, зберігаючи при цьому контури та різкі переходи між об'єктами. На відміну від фільтра середнього, медіанний фільтр не розмиває краї, що

особливо корисно для очищення фотографій, сканованих документів та медичних зображень.

Гаусовий фільтр використовує згладжування за принципом зваженого середнього – кожен піксель має вагу відповідно до гаусового (нормального) розподілу, тобто, що центральний піксель має найбільший вплив, тоді як сусідні значення враховуються менше, що допомагає уникнути різких змін у зображенні. Такий фільтр особливо ефективний для усунення гаусового шуму, який зазвичай з'являється через електронні перешкоди в сенсорах камер або під час процесу оцифрування зображень. Однією з його переваг є те, що він зберігає плавні градієнти, однак, як і фільтр середнього, може спричиняти розмиття текстур та деталей.

Попри свою ефективність у базових задачах зменшення шуму, ці фільтри мають певні обмеження. Вони працюють однаково для всіх областей зображення, не враховуючи локальні особливості сцени. У результаті важливі дрібні деталі можуть згладжуватися разом із шумом, через що ці методи менш придатними для зображень із високою текстурною складністю. Окрім того, класичні згладжувальні фільтри є неефективними при роботі зі складними шумами, такими як анізотропний або структурований шум, який має певні закономірності.

Зважаючи на ці недоліки, у сучасній обробці зображень класичні фільтри згладжування часто поєднуються з адаптивними підходами, такими як нелокальне усереднення або багат шарові нейромережі, які можуть краще враховувати контекст зображення та ефективніше відокремлювати шум від корисного сигналу. Проте середній, медіанний і Гаусовий фільтри залишаються популярними через їхню простоту та швидкість обчислень, вони є зручними для попередньої обробки або застосування в системах реального часу.

BM3D [18] та NLM [19] одні з найефективніших традиційних методів усунення шуму. Вони використовують більш складні підходи, ніж класичні фільтри згладжування, оскільки не просто усереднюють значення пікселів у

локальному вікні, а шукають схожі області у всьому зображенні. Це дозволяє їм зберігати деталі, контури та текстури, які часто губляться при застосуванні простіших методів.

BM3D один із найкращих алгоритмів для видалення шуму, який використовує методи просторово-частотної обробки. Він ґрунтується на ідеї пошуку схожих фрагментів у зображенні, групуванні їх у тривимірні блоки та подальшій обробці у частотному домені.

Процес BM3D складається з декількох основних етапів:

- пошук схожих фрагментів: кожен фрагмент зображення порівнюється з іншими частинами, і блоки, які мають високу подібність, групуються разом;

- тривимірна обробка: отримані групи блоків перетворюються у частотну область, де шумові компоненти можуть бути ефективно відфільтровані за допомогою порогової обробки;

- зворотне перетворення: після усунення шуму у частотному домені відфільтровані блоки повертаються у просторову область, де вони агрегуються для отримання фінального відновленого зображення.

Головна перевага BM3D – це здатність зберігати дрібні деталі та текстури, оскільки він обробляє схожі області разом, що значно ефективніше, в порівнянні локальні фільтри. Проте він є досить обчислювально складним, що може ускладнювати його використання у реальному часі.

NLM працює за принципом пошуку подібних пікселів у всьому зображенні, навіть якщо вони знаходяться далеко один від одного. На відміну від класичних фільтрів, що використовують локальну інформацію, NLM розглядає глобальні подібності між різними областями.

Основна ідея NLM полягає у тому, що в природних зображеннях часто зустрічаються схожі структури, і середнє значення подібних пікселів дає більш точну оцінку, ніж усереднення тільки локальних сусідів. Алгоритм працює так:

- для кожного пікселя визначається область пошуку (вікно), у якій порівнюються схожі фрагменти;
- кожен фрагмент отримує вагу, яка визначається мірою його подібності до центрального пікселя;
- остаточне значення пікселя визначається як зважене середнє всіх схожих пікселів у визначеному вікні.

NLM ефективно видаляє гаусовий шум, зберігаючи при цьому текстури та контури. Його основний недолік – це висока обчислювальна складність, оскільки алгоритм порівнює кожен піксель із великою кількістю інших. Однак сучасні оптимізації, такі як використання багатомасштабних підходів та прискорених версій алгоритму, дозволяють зменшити обчислювальні витрати.

Обидва методи демонструють високу ефективність у порівнянні з класичними фільтрами, зокрема при усуненні гаусового шуму. VM3D краще працює на текстурних зображеннях, оскільки враховує подібність фрагментів, тоді як NLM більш адаптивний до загальної структури зображення і забезпечує м'яке згладжування шуму без втрати деталей.

Попри свою ефективність, обидва методи мають високу обчислювальну складність, роблячи їх менш придатними для реального часу без відповідних оптимізацій. Через це у сучасних задачах глибоке навчання поступово витісняє ці підходи, пропонуючи швидші та ще точніші методи денойзингу. Однак VM3D та NLM все ще залишаються актуальними в задачах, де необхідно зберігати високий рівень деталей при усуненні шуму.

### 1.3.2 Глибокі нейромережеві підходи в задачі денойзингу

Глибокі нейромережеві підходи значно розширили можливості усунення шуму, перевершуючи класичні методи завдяки здатності адаптивно обробляти різні типи шуму без втрати деталей. Використовуючи

багаторівневі представлення даних, такі моделі, як DnCNN, FFDNet, Noise2Noise і Blind2Unblind, можуть ефективно відокремлювати шумові компоненти, зберігаючи структуру зображення. У цьому підрозділі розглядаються основні нейромереві методи денойзингу, їхні переваги над класичними підходами, а також можливі обмеження.

Глибокі нейромереві методи усунення шуму стали значним проривом у цифровій обробці зображень, оскільки вони дозволяють адаптивно видаляти шум, не потребуючи жорстко заданих припущень про його характер. На відміну від класичних фільтраційних або статистичних методів, глибокі нейронні мережі навчаються, як правило, на великому обсязі даних і відповідно виявляють складні патерни шуму, що дозволяє їм ефективно працювати у різних сценаріях. Серед найбільш популярних моделей денойзингу слід виділити DnCNN, FFDNet, Noise2Noise і Blind2Unblind, кожна з яких має свої особливості та переваги.

DnCNNодна з перших згорткових нейромеревих моделей, спеціально розроблених для видалення шуму. Її архітектура ґрунтується на глибоких згорткових шарах з нелінійними активаціями та використанням шарів пакетної нормалізації, з метою покращення стабільності навчання. Модель навчається передбачати карту шуму у зображенні, а потім ця оцінка використовується для відновлення чистого зображення. Однією з ключових переваг DnCNN є її універсальність, вона може працювати з різними рівнями шуму, не потребуючи окремого навчання для кожного випадку.

FFDNet це оптимізована версією DnCNN, яка дозволяє користувачеві вручну контролювати рівень фільтрації шуму. Її головна особливість полягає у використанні багатоканального представлення вхідного зображення, де додатковий канал містить інформацію про очікуваний рівень шуму. У результаті чого, FFDNet працює більш ефективно, оскільки користувач може змінювати параметри згладжування без необхідності повторного навчання моделі. Завдяки компактнішій архітектурі, FFDNet

швидший у порівнянні з DnCNN, він є менш придатним для реальних додатків, зокрема для відеоденойзингу.

Noise2Noise ще одним важливим проривом у сфері денойзингу, оскільки ця модель може навчатися без потреби у чистих еталонних зображеннях. Традиційні методи глибокого навчання передбачають, що для навчання необхідно мати пари «зашумлене – чисте зображення», що може бути складним у реальних умовах. Noise2Noise використовує два незалежно зашумлених зразки одного зображення для навчання та вчиться реконструювати вихідні дані, усуваючи випадковий шум. Завдяки цьому підходу модель може працювати навіть у випадках, коли відсутні ідеальні зображення для навчання.

Новітній підход Blind2Unblind, базується на концепції навчання без референсного чистого зображення. У цьому методі модель отримує зашумлені вхідні дані, але не має доступу до чистих варіантів зображень. Використовуючи знання про структуру зображень і особливості шуму, неймережа навчається відокремлювати корисний сигнал від шумових компонентів. Такий підхід особливо корисний у випадках, коли рівень і природа шуму невідомі заздалегідь або змінюються динамічно.

Усі ці методи демонструють значні переваги перед класичними підходами до денойзингу, оскільки вони не просто згладжують шум, а адаптивно аналізують кожен область зображення, зберігаючи деталі та текстури. Крім того, глибокі неймережі можуть працювати з різними типами шуму, зокрема гаусовим, імпульсним і навіть складними структурованими спотвореннями. Завдяки своїй гнучкості та ефективності, неймережеві методи денойзингу широко застосовуються у фотографії, медичній візуалізації, комп'ютерному зорі та інших сферах, де якість зображень є критично важливою.

Головним недоліком класичних методів денойзингу є їхня обмежена здатність зберігати деталі зображення. Фільтраційні методи, такі як середнє чи Гаусове згладжування, усувають шум, але водночас розмивають контури

та дрібні структури. Методи на основі просторових або частотних перетворень, такі як BM3D та NLM, мають вищу ефективність, проте вони вимагають складного налаштування параметрів і можуть працювати повільно. Крім того, класичні алгоритми зазвичай припускають, що шум має певний розподіл (наприклад, гаусовий), що не завжди відповідає реальним умовам.

Натомість нейромережі, такі як DnCNN, FFDNet і Noise2Noise, можуть навчатися на великих обсягах реальних даних, адаптуючись до складних та неоднорідних шумових патернів. Вони не потребують жорстких припущень про структуру шуму, а їхня глибока архітектура дозволяє зберігати текстури та тонкі деталі зображень. Ще однією перевагою є швидкість обробки: після навчання нейромережі можуть виконувати денойзинг у реальному часі, що особливо цінно для відеоаналізу, медичної візуалізації та автономних систем. Завдяки цим перевагам нейромережеві методи дедалі частіше замінюють традиційні алгоритми, особливо в задачах, де висока якість зображення є критично важливою.

#### 1.4 Використання метаданих камери для оцінки шуму

Оцінка рівня шуму в цифрових зображеннях традиційно базується на аналізі самих візуальних даних, проте такий підхід не завжди достатньо точним. Реальні зображення містять шум, що залежить від характеристик сенсора камери та умов зйомки, які можуть значно варіюватися. Саме тому використання метаданих камери [19], [20], [21] важливе доповнення до традиційних методів оцінки шуму, оскільки вони містять інформацію про параметри, що впливають на рівень спотворень у зображенні.

Фізичні особливості камери, такі як тип сенсора, розмір пікселів, процес зчитування сигналу та підсилення, відіграють ключову роль у формуванні шуму. Окрім цього, метадані, зокрема значення ISO, час експозиції, температура сенсора, можуть бути використані для

прогнозування рівня шуму та його адаптивного усунення. Вони дають змогу не лише покращити оцінку шумових компонентів, а й розробляти алгоритми денойзингу, які враховують конкретні особливості камери.

Фізичні параметри камери значно впливають на рівень шуму в цифрових зображеннях, оскільки джерела шуму здебільшого пов'язані з оптичними та електронними характеристиками сенсора. Одним із ключових факторів це тип сенсора. CMOS- і CCD-матриці мають різні механізми зчитування світла, що призводить до відмінностей у рівні теплового та зчитувального шуму. CMOS-сенсори, які є більш поширеними у сучасних камерах, зазвичай мають вищий шум зчитування, тоді як CCD-сенсори демонструють нижчий рівень шуму, але енергетично менш ефективними.

Розмір пікселів також важливий параметр: чим більший піксель, тим більше фотонів він здатен поглинути, що зменшує вплив фотонного шуму. Камери з високою роздільною здатністю, які мають дрібніші пікселі, схильні до більшого шуму через обмежений збір світла. Високе значення ISO (цифрове підсилення світлочутливості) також збільшує рівень шуму, оскільки разом із корисним сигналом підсилюються і випадкові електронні флуктуації.

Іншим критичним параметром є час експозиції. При тривалішій експозиції сенсор накопичує більше світла, що може зменшити шум, проте водночас зростає вплив теплового шуму, який виникає через теплове збудження електронів у матриці. Додатково на шум впливають температура сенсора та процес зчитування сигналу, оскільки електронні компоненти камери генерують власний шум під час обробки даних.

Метадані камери, такі як ISO, час експозиції та температура сенсора, відіграють ключову роль у визначенні рівня шуму та його характеристик у цифрових зображеннях. Оскільки шум це наслідком фізичних процесів, що відбуваються під час зчитування та обробки сигналу, використання цих параметрів дозволяє точніше прогнозувати рівень спотворень та адаптивно налаштовувати методи їхнього усунення.

ISO визначає рівень цифрового підсилення сигналу, і його підвищення призводить до зростання шуму. Високі значення ISO необхідні в умовах слабого освітлення, проте вони також посилюють випадкові електронні флуктуації, що проявляється у вигляді зернистості або кольорових артефактів. Використовуючи ISO як вхідний параметр для алгоритмів денойзингу, можна адаптивно регулювати рівень фільтрації, зберігаючи при цьому природні деталі зображення.

Час експозиції впливає на співвідношення сигналу до шуму. Коротка експозиція може призводити до недоекспонованих знімків із високим рівнем фотонного шуму, тоді як тривала експозиція може збільшити тепловий шум, особливо у камерах із тривалим зчитуванням. Враховуючи експозицію у моделі оцінки шуму, можна точніше розрізнити шумові компоненти та використовувати відповідні методи їхнього усунення.

Температура сенсора прямо впливає на рівень теплового шуму, який є значною проблемою при довгій витримці або зйомці в умовах підвищеної температури. Камери з активним охолодженням, наприклад, в астрономії, мінімізують цей тип шуму, проте в більшості споживчих камер він залишається критичним фактором.

Використання цих метаданих у процесі оцінки шуму дозволяє покращити точність прогнозування та адаптивного усунення шумових артефактів, що підтверджує їх значущість, як компонента сучасних алгоритмів денойзингу.

Комбінування метаданих із візуальною інформацією ефективним підходом для покращення оцінки рівня шуму, оскільки дозволяє враховувати як статистичні характеристики зображення, так і фізичні параметри камери. Використання лише піксельної інформації може бути недостатнім для точної оцінки шуму, особливо якщо його рівень змінюється в різних частинах зображення або залежить від умов зйомки.

Метадані, такі як ISO, експозиція та температура сенсора, можуть бути використані як додаткові входи для алгоритмів глибокого навчання або

класичних моделей оцінки шуму. Наприклад, в нейромережових методах денойзингу ці параметри можуть допомагати моделі адаптивно визначати тип і рівень шуму в кожній області зображення. Деякі сучасні алгоритми, такі як FFDNet, використовують змінний параметр шуму як вхідний канал, для регулювати інтенсивність фільтрації.

Комбінування візуальних та мета-даних також ефективно у гібридних моделях, де традиційні статистичні підходи доповнюються інформацією про камеру, що дає змогу створювати більш гнучкі та точні алгоритми оцінки шуму, що особливо корисно для професійної фотографії, медицини та супутникової зйомки.

#### 1.4 Постановка задачі дослідження

Цифрові зображення широко застосовуються в комп'ютерному зорі, медичній діагностиці, відеоспостереженні, аерофотозйомці та автономному транспорті. Проте їхня якість часто погіршується через наявність шумів, що спричиняє спотворення візуальної інформації та зниження ефективності алгоритмів.

Шум у цифрових зображеннях може мати різну природу: фотонний, тепловий, шум зчитування, квантовий та інші артефакти, що виникають у процесі аналого-цифрового перетворення. Він негативно впливає на точність медичних систем і автономних транспортних засобів, створюючи потребу у вдосконалених методах його оцінки та усунення.

Рівень шуму залежить від характеристик сенсора камери, умов освітлення, параметрів зйомки (ISO, експозиція, температура сенсора) та алгоритмів обробки сигналу. Традиційні методи денойзингу (середній, медіанний, гаусовий фільтри, перетворення Фур'є, вейвлет-перетворення) ефективні лише для певних типів шуму, але можуть призводити до розмиття текстур або втрати деталей. Методи машинного навчання (BM3D, NLM) мають вищу точність, проте обчислювально складними.

Глибокі нейромережі (DnCNN, FFDNet, Noise2Noise, Blind2Unblind) перевершують класичні підходи, адже можуть адаптивно аналізувати зображення, розпізнавати складні особливості шуму та зберігати текстури. Проте вони вимагають великих обчислювальних ресурсів, обширних наборів даних і можуть втрачати ефективність при нових типах шуму.

Перспективним напрямом є інтеграція метаданих камери у процес денойзингу. Використання параметрів зйомки (ISO, експозиція, температура сенсора) дає можливість точніше оцінювати шум та адаптувати моделі глибокого навчання до конкретних умов. Це особливо важливо для медичної візуалізації, автономних систем та аерофотозйомки.

Мета дослідження – розробка нейромережевої моделі для оцінки та усунення шуму із використанням метаданих камери. Для цього необхідно:

- дослідити вплив параметрів камери на рівень шуму;
- розробити архітектуру нейромережі, що використовує зображення та метадані;
- впровадити механізм оцінки різних типів шуму;
- навчити модель на реальних і синтетичних наборах даних;
- порівняти ефективність моделі з класичними методами денойзингу.

## 2 МЕТОДОЛОГІЯ ОЦІНКИ ДЖЕРЕЛ ШУМУ НА ОСНОВІ ГЛИБИННОГО НАВЧАННЯ

У традиційних підходах до усунення шуму в цифрових зображеннях основна увага приділяється саме обробці зображення після його спотворення. Такі методи зазвичай не враховують природу походження шуму та працюють зі вже наявними наслідками, що часто призводить до втрати деталізації або появи артефактів. Особливо це проблема в автономних системах, де камери є ключовими сенсорами, і помилки у зображенні можуть впливати на безпеку всієї системи.

У цьому розділі описано підхід, який дозволяє не просто зменшити шум, а оцінити його джерела тобто виявити, яка саме фізична причина (наприклад, фотонний шум або шум темного струму) спричинила спотворення. Для цього використовується нейронна мережа, яка навчається одночасно на зображеннях і на метаданих камери. Такий підхід призводить до більш точного визначення рівня шуму і в деяких випадках навіть виявляти несправності або відхилення в роботі камери.

### 2.1 Загальна концепція та архітектура системи

Для переходу від симптоматичного усунення шуму до виявлення його причин необхідна концептуально нова методика, здатна не лише аналізувати зашумлені зображення, а й враховувати умови, за яких вони були отримані. У рамках сучасного підходу передбачається розробка моделі, здатної у реальному часі оцінювати рівень і природу шуму на основі зображення та супровідних метаданих камери. При цьому формулюється завдання не просто зниження шуму, а його структурної інтерпретації з поділом на фізично обґрунтовані компоненти.

Система (рисунок 2.1) базується на глибокій нейронній мережі, яка працює з двома джерелами інформації самим зображенням та набором

параметрів, що описують конфігурацію камери та умови зйомки. Модель отримує на вхід фрагмент зображення (розміром  $128 \times 128$  пікселів у відтінках сірого) та набір метаданих, таких як посилення сигналу, температура сенсора, час експозиції та інші технічні характеристики. На основі цієї інформації здійснюється оцінка внеску основних типів шуму, які виникають у процесі формування цифрового зображення.

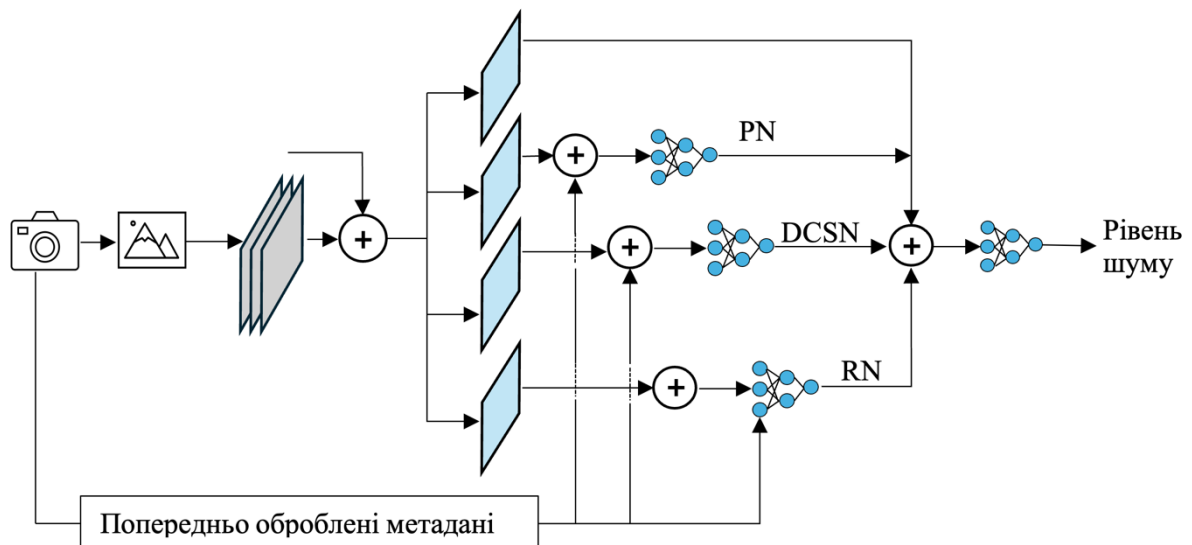


Рисунок 2.1 – Система оцінки джерела шуму з використанням метаданих для передбачення рівнів шуму

Оцінювання проводиться окремо для трьох основних джерел шуму: шуму фотонного походження, шуму темного струму та шуму зчитування. Кожен із них має свою фізичну природу і формується на різних етапах роботи сенсора. Крім того, у моделі передбачено компонент залишкового або неочікуваного шуму, який фіксує невідповідність між даними зображення і супровідною інформацією про камеру. Такий підхід дає змогу точніше оцінити рівень спотворення та ідентифікувати аномалії, що можуть свідчити про технічні проблеми пристрою.

Функціонування системи реалізується на основі модифікованої глибокої згорткової нейронної мережі, яка адаптована для роботи з

однотональними зображеннями та оптимізована для швидкої обробки. Основна частина моделі складається з резидуальних блоків, що витягують ознаки, характерні для шумових патернів. Далі ці ознаки передаються в блоки оцінки, кожен з яких спеціалізується на прогнозуванні окремого типу шуму. З метою підвищення точності, у відповідні гілки подаються також релевантні метадані, попередньо нормалізовані до уніфікованого діапазону.

Кожна з трьох гілок формує числову оцінку конкретного джерела шуму, враховуючи як структурні особливості зображення, так і фізичні параметри зйомки. Додатково функціонує окрема гілка, яка здійснює оцінку залишкового шуму на основі невідповідностей між зображенням і очікуваними характеристиками, розрахованими за метаданими. Така конструкція дозволяє виявляти ті випадки, коли технічні дані не збігаються з реальним станом зображення, що критично важливо для систем самодіагностики.

Після обробки в усіх гілках результати об'єднуються у підсумковий показник шуму, який розраховується як квадратний корінь із суми квадратів окремих складових. Такий підхід відповідає традиційному способу обчислення загального рівня незалежних шумів і забезпечує можливість отримати як загальну оцінку, так і деталізоване розділення за джерелами. Таким чином, модель надає не лише чисельну інформацію про спотворення, а й пояснення його природи.

Особливу увагу приділено інтеграції метаданих. Для цього передбачено різні конфігурації: модель без метаданих, із мінімальним набором параметрів та з повною інформацією про сенсор. Проведене експериментальне дослідження показало, що лише при повному обсязі метаданих можлива стабільна й точна оцінка усіх типів шуму. Без використання параметрів зйомки модель схильна до помилкових інтерпретацій, особливо при диференціації між темновим та електронним шумом.

У межах загальної методології також реалізовано механізм виявлення неочікуваних шумів, що можуть виникати внаслідок зовнішніх впливів або деградації апаратного забезпечення. За рахунок наявності залишкової компоненти модель здатна сигналізувати про наявність потенційної несправності або відхилення від очікуваної поведінки сенсора. Це відкриває можливість використання системи не лише для покращення зображень, але й для контролю стану обладнання в умовах реального середовища.

Завдяки застосуванню оптимізованої архітектури модель демонструє високу швидкодію, що сприяє обробці зображення у реальному часі на стандартному графічному процесорі. Висока точність при зменшеній кількості параметрів робить її придатною для використання у вбудованих системах, які мають обмеження по ресурсах. При цьому забезпечується збереження високої інформативності, що робить можливим подальший аналіз та автоматизоване реагування на виявлені зміни.

Досліджена методологія забезпечує перехід від класичної обробки зображень до комплексного підходу, що поєднує інтерпретованість, ефективність і практичну застосовність. У наступному розділі буде представлено реалізацію цієї системи, етапи підготовки даних для навчання, деталі архітектури, а також результати експериментального оцінювання її ефективності.

## 2.2 Розширення для ідентифікації джерел шуму

У рамках побудованої нейронної архітектури важливою частиною стало розширення базової структури з метою можливості відокремленого аналізу різних типів шуму. Замість отримання єдиної узагальненої оцінки, що часто практикується в класичних системах оцінювання якості зображення, у даній реалізації передбачено поділ на окремі функціональні гілки. Кожна з них спеціалізується на вивченні та прогнозуванні

конкретного джерела шуму, що значно підвищує точність і фізичну інтерпретованість результатів.

Архітектурно ця концепція реалізована шляхом побудови трьох незалежних обчислювальних гілок, які працюють паралельно. Кожна з них відповідає за оцінку одного з основних типів шуму, а саме: шуму фотонного походження, шуму темного струму та шуму зчитування. Таке розмежування дозволяє кожній гілці адаптуватися до характеру тієї частини шуму, з якою вона працює, не змагаючись за узагальнення в одному потоці ознак. Кожна з цих гілок містить власну послідовність повнозв'язних шарів, що забезпечує обчислення на базі як ознак зображення, так і відповідного піднабору метаданих.

Формування індивідуальних блоків оцінки реалізовано за допомогою незалежних повнозв'язних структур, які дозволяють адаптивно обробляти вхідні дані. Ці блоки не мають спільних параметрів і оптимізуються незалежно, що запобігає перехресному навчальному впливу між різними джерелами шуму. У процесі тренування кожен блок фокусується лише на тій інформації, яка релевантна саме для його задачі, що знижує загальне інформаційне перевантаження і підвищує селективність реагування на особливості шумових патернів.

Особливу увагу було приділено інтеграції метаданих у відповідні гілки. Передбачено, що кожен тип шуму має власні чутливі до нього технічні параметри. Наприклад, оцінка шуму фотонного походження враховує інтенсивність сигналу, що пов'язана з gain, тоді як для шуму темного струму критичними є температура сенсора та тривалість експозиції. Зі свого боку, гілка для шуму зчитування орієнтується на характеристики електронної частини камери. Для врахування цього, вхідні метадані було попередньо розділено на підгрупи, кожна з яких подається лише до тієї гілки, для якої вона змістовна.

На рівні реалізації це означає, що після вилучення ознак з зображення вони об'єднуються (конкатенуються) лише з тією частиною метаданих, яка

безпосередньо впливає на відповідний фізичний процес, що зменшує кількість непотрібної інформації у вхідному векторі і робить навчання більш цілеспрямованим. Оскільки ознаки зображення і метадані мають різну природу, їх поєднання реалізовано у вигляді паралельного подання та об'єднання у першому шарі кожного повнозв'язного блоку. Таким чином, модель навчається не просто виявляти шаблони у вхідному зображенні, а й співвідносити їх із очікуваними технічними параметрами.

Таке розділення на гілки дозволило моделі досягти значно вищої точності у виділенні конкретних шумів, зменшити внутрішню плутанину між джерелами та надати інтерпретовані результати, які можуть бути використані як для подальшої обробки зображень, так і для моніторингу стану камери. У наступному підрозділі буде розглянуто, яким чином обчислюється залишкова складова шуму, що не пояснюється жодним із фізично змодельованих джерел.

### 2.3 Розширення для ідентифікації джерел шуму

Після формування базової структури нейронної мережі, яка виконує загальне оцінювання рівня шуму на основі аналізу зображення, постала необхідність деталізованого розподілу шумів за їхнім фізичним походженням. Такий підхід обумовлений бажанням отримати не лише числову характеристику загальної шумності, а й розуміння, які саме процеси всередині сенсорної системи спричиняють ту чи іншу форму спотворення. Це стало передумовою для впровадження структурного розширення моделі у вигляді спеціалізованих обчислювальних гілок, кожна з яких відповідає за окремий тип шуму.

З огляду на природу цифрового шуму, а також з урахуванням особливостей формування зображення у сучасних камерах, було передбачено розподіл архітектури на три незалежні гілки, кожна з яких спеціалізується на оцінці одного з основних типів шуму. Таким чином, на

виході система здатна формувати не лише інтегральну оцінку, а й розкласти загальний шум на складові: фотонний шум, шум темного струму та шум, що виникає у процесі зчитування сигналу. Ці компоненти є результатом різних фізичних явищ, тому вимагають не тільки окремого підходу до моделювання, але й різного обсягу й характеру вхідної інформації.

Для реалізації такого поділу в кожному гілку було закладено окремий повнозв'язний обчислювальний блок, що працює з унікальним набором вхідних ознак. Незалежність цих блоків дозволила уникнути перехресного впливу одного типу шуму на результати оцінки іншого. Такий підхід обґрунтований як з теоретичної, так і з практичної точки зору: кожен тип шуму має власні характерні прояви в зображенні, а також по-різному залежить від параметрів зйомки. Тому функціональна роз'єднаність блоків сприяє підвищенню точності прогнозування та зменшенню ймовірності некоректної інтерпретації джерела шуму.

Ознаки, на яких базується аналіз у кожній гілці, формуються з одного боку з вхідного зображення, а з іншого, з метаданих камери. Вилучення ознак із зображення здійснюється на ранніх етапах моделі за допомогою згорткових операцій, що допомагає зберегти локальну інформацію, характерну для шумових структур. Після обробки у згорткових шарах дані проходять через глобальне пулінгування, яке стискає інформацію до векторного представлення. Цей вектор і є тим представленням вхідного зображення, яке подається в кожному гілку.

Окрему роль у функціонуванні розширених гілок відіграє підключення метаданих, що містять інформацію про умови отримання зображення та фізичні характеристики камери. Оскільки кожен тип шуму залежить від певних технічних параметрів, було реалізовано механізм спеціалізованого подання цих параметрів у відповідні гілки. Такий розподіл не формальний або симетричний: для кожної гілки формується власний піднабір метаданих, який відповідає лише тим характеристикам, що мають фізичне відношення до відповідного типу шуму.

Наприклад, при оцінці фотонного шуму найважливішими є параметри, що пов'язані з експозицією, посиленням сигналу та інтенсивністю освітлення, оскільки цей шум безпосередньо залежить від кількості фотонів, що потрапили на сенсор. Зі свого боку, для оцінки шуму темного струму ключовими вважаються температура сенсора та тривалість експозиції фактори, що впливають на накопичення теплових електронів у відсутності світла. Щодо шуму зчитування, то тут вирішальну роль відіграють параметри, пов'язані з електронікою камери: швидкість зчитування, тип сенсора, внутрішній шум вузлів підсилення та інші аналогічні характеристики.

Перед передачею у відповідні гілки всі метадані проходять процес нормалізації. Це необхідно як для стабілізації процесу навчання, так і для забезпечення узгодженості між різними типами вхідної інформації. Оскільки ознаки, отримані з зображення, мають інший масштаб і статистичні характеристики, їх поєднання з необробленими метаданими могло б призвести до домінування одних ознак над іншими. Нормалізація забезпечує уникнення цього ефекту, зберігаючи рівнозначну вагу різних джерел інформації у процесі прогнозування.

Об'єднання ознак зображення та метаданих реалізується через операцію конкатенації, яка виконується перед подачею на перший повнозв'язний шар у кожній з гілок. Таким чином формується спільний вхідний вектор, який містить як інформацію про візуальну структуру вхідного патчу, так і про умови його формування. Завдяки цьому кожен повнозв'язний блок отримує повноцінне уявлення про ситуацію, в якій було сформовано зображення, та може робити оцінку не лише на основі того, що бачить, але й з урахуванням контексту.

У процесі навчання модель самостійно встановлює зв'язки між компонентами вхідного вектора та значеннями прогнозованого шуму, що показує те що модель не лише використовує інформацію, але й вчиться відрізняти, які саме ознаки з більшою ймовірністю відповідають за ту чи

іншу форму шуму. Таким чином досягається ефект спеціалізації: кожна гілка з часом налаштовується на власну задачу, мінімізуючи вплив неінформативних або вторинних параметрів. У результаті формується стійка і стабільна архітектура, де кожен елемент працює у своїй зоні відповідальності.

У такому підході особливо важливим є баланс між універсальністю і спеціалізацією. З одного боку, модель має бути здатною працювати з широким спектром зображень і технічних умов, з іншого, кожна з її гілок повинна бути достатньо вузько орієнтованою, щоб досягти високої точності саме у своїй сфері. Саме завдяки архітектурному поділу на окремі функціональні блоки цей баланс вдалося реалізувати без необхідності створення трьох окремих моделей.

У підсумку можна зазначити, що розширення архітектури нейронної мережі через введення спеціалізованих гілок та незалежних блоків обробки інформації дало змогу дослідити більш точну, стабільну і технічно обґрунтовану систему, що виконує поставлені завдання оцінки рівня шуму, та визначення його джерела і відповідно які процеси за нього відповідають.

#### 2.4 Інтеграція метаданих камери

Ефективність оцінки джерел шуму у цифрових зображеннях значною мірою залежить не лише від характеристик самого зображення, але й від контексту його формування, зокрема технічних параметрів камери. У процесі побудови моделі, орієнтованої на інтерпретоване оцінювання шумових компонент, було реалізовано механізм інтеграції метаданих числових параметрів, що описують умови та налаштування під час зйомки. До таких даних зазвичай належать посилення (gain), експозиція, температура сенсора, тип сенсора, частота зчитування, ємність пікселів та інші технічні характеристики.

Замість розглядати метадані як побічну допоміжну інформацію, їх було трактовано як джерело значущих ознак, що доповнюють візуальну інформацію зображення та дозволяють покращити точність прогнозу. Особливо важливим виявився цей підхід у ситуаціях, коли візуальні особливості не давали змоги однозначно ідентифікувати джерело шуму. Наприклад, наявність розмитого шуму в умовах слабкого освітлення без знання експозиції чи gain не дозволяє відрізнити фотонний шум від теплового. Наявність метаданих дає змогу встановити точнішу причинно-наслідкову залежність.

Щоб забезпечити коректну взаємодію між візуальними та технічними ознаками, метадані подаються у модель разом з ознаками зображення. Проте вони не розглядаються як однорідний вхід. Навпаки, було реалізовано спеціалізований підхід, за якого кожна з функціональних гілок моделі отримує лише ті параметри, які релевантні саме для її задачі. Наприклад, при аналізі шуму темного струму важливими є температура сенсора та час експозиції, тоді як при оцінці шуму зчитування враховуються параметри, пов'язані з електронікою тип сенсора, частота, ємність. Такий розподіл метаданих по гілках забезпечив змістовну фільтрацію вхідної інформації та підвищив загальну стабільність прогнозування.

Усі метадані перед подачею у мережу нормалізуються в межах уніфікованого діапазону, для збереження об'єктивності обробки числових значень різних масштабів. Це запобігає переважанню окремих параметрів над іншими та сприяє збалансованому навчанню моделі. Після нормалізації параметри конкатенуються з ознаками, витягнутими з зображення, та подаються як єдиний вхідний вектор до відповідного повнозв'язного блоку. Таке об'єднання забезпечує моделі змогу формувати гнучкі зв'язки між зовнішніми проявами шуму та його потенційними технічними джерелами.

Експериментальні результати показали, що саме завдяки інтеграції метаданих модель змогла досягти стабільної диференціації між схожими за візуальними характеристиками джерелами шуму. Наприклад, у випадках,

коли тепловий шум і шум зчитування накладались один на одного, саме контекст з параметрів камери дозволяв зробити правильну атрибуцію. Більш того, модель виявилась здатною за допомогою метаданих виявляти невідповідності, що виникали між очікуваним та реальним рівнем шуму, що заклало основи для подальшого застосування системи у задачах самодіагностики камер.

Було реалізовано кілька конфігурацій моделі залежно від обсягу наявних метаданих: без них взагалі, із мінімальним набором (експозиція, температура, gain) і з повним масивом технічних параметрів. Найкращі результати, як щодо точності, так і щодо стабільності, були отримані при використанні повної конфігурації. Водночас навіть мінімальні метадані дали суттєве покращення у порівнянні з режимом, де модель працювала лише з візуальною частиною.

Інтеграція метаданих надала системі глибину і здатність до контекстуалізованого аналізу, який виявився суттєво ефективнішим за традиційний підхід на основі лише зображення. У межах описаної архітектури вони виконують роль не додаткового сигналу, а повноцінного джерела вхідних даних, без яких побудова пояснюваної, стійкої та функціонально надійної моделі була б неможливою.

## 2.5 Виявлення залишкового шуму

У межах побудованої моделі, що виконує роздільну оцінку основних джерел шуму у цифровому зображенні, окреме місце займає компонент, що відповідає за виявлення залишкового або так званого неочікуваного шуму. Йдеться про ті випадки, коли загальний рівень спотворення не може бути повністю пояснений сукупністю класичних фізичних джерел, таких як фотонний шум, шум темного струму чи шум зчитування. Причини виникнення такого залишкового ефекту можуть бути різними – від деградації елементів сенсора до неадекватної поведінки електроніки або

збоїв у калібруванні. У звичайних умовах такі спотворення не систематичні і тому погано піддаються моделюванню. Саме для цього в архітектурі передбачено спеціалізовану гілку, яка відповідає за виявлення і кількісне оцінювання саме такого типу шуму.

Модельна гіпотеза, закладена у цю гілку, полягає в тому, що у разі коректної роботи камери та достовірності метаданих, три основні компоненти шуму повинні узгоджено пояснювати візуальні характеристики вхідного зображення. Якщо ж між прогнозованим сукупним рівнем шуму та реальною структурою зображення спостерігається суттєва різниця, вона вважається залишковою. Іншими словами, залишковий шум – це різниця між тим, що очікується на основі фізичних моделей, і тим, що фактично спостерігається в зображенні. Така логіка побудови забезпечує створення своєрідної захисну лінію, яка перехоплює все те, що не може бути пояснено попередніми гілками, і, відповідно, має найвищу ймовірність бути пов'язаним із несправностями або винятковими умовами.

Гілка, яка виконує цю функцію, формально не повністю незалежна, оскільки її вхід формують як ознаки самого зображення, так і виходи з трьох попередніх блоків оцінки основних типів шуму. У такий спосіб вона отримує повну картину: як візуальну, так і оцінково-аналітичну. Завдяки цьому модель навчається виявляти відхилення між очікуваним рівнем шуму, який складається з компонентів, і фактичними ознаками зображення. У випадках, коли зображення демонструє суттєво вищий або структурно інший рівень шуму, ніж це передбачено технічними параметрами, активується прогноз залишкового компонента.

Оскільки характер залишкового шуму не має чіткої фізичної природи і, відповідно, не піддається прямому математичному опису, основна робота моделі полягає в накопиченні евристичного досвіду через тренування на великій кількості ситуацій. Для цього було сформовано навчальні приклади, у яких між вхідними метаданими та відповідним зображенням навмисно створюється невідповідність – наприклад, температура сенсора вказується

меншою, ніж була насправді, або змінюється тип сенсора у вхідному записі. Такі симуляції дають змогу моделі зіставити ситуації, де метадані свідчать про одне, а візуальний вигляд зображення – про інше. Саме у таких умовах формується здатність виявляти залишкову частину шуму, яка виникає через нестандартні обставини або апаратні аномалії.

Ключовим у цьому процесі є те, що модель не шукає універсальних закономірностей у самих залишках, а навчається ідентифікувати ситуації невідповідності, що забезпечує її адаптивність до широкого спектра можливих варіантів і демонструвати стійкість до варіацій як у метаданих, так і у структурі шуму. Важливо також зазначити, що ця гілка може бути використана не лише як компонент у розрахунку загального рівня шуму, а й окремо – як індикатор технічної справності пристрою. Якщо, наприклад, значення залишкового шуму систематично перевищує певний поріг у серії зображень, це може слугувати сигналом для автоматизованої перевірки чи профілактичного обслуговування камери.

У реалізації архітектури ця гілка структурно подібна до інших: вона складається з повнозв'язного блоку, що отримує вхідний вектор з ознаками зображення та обчисленими компонентами PN, DCSN, RN, а також відповідною підмножиною метаданих. Після проходження крізь шари блоку формується прогноз значення залишкового шуму, яке потім додається до загальної формули обчислення рівня шуму. Завдяки тому, що ця гілка має повний доступ до усіх попередніх результатів, вона виступає логічним завершенням обчислювального конвеєра, що замикає усю структуру в єдину узгоджену систему.

У підсумку реалізація гілки залишкового шуму забезпечує моделі гнучкість, здатність до узагальнення та чутливість до відхилень. Вона не лише підвищує точність прогнозу в нестандартних випадках, а й створює додатковий рівень контролю за якістю вхідних даних і станом обладнання. Таким чином, саме цей компонент дає змогу моделі вийти за межі формального підрахунку шумових характеристик і набути властивостей

системи самодіагностики, що особливо важливо у сучасних автономних пристроях, які працюють без постійного контролю з боку оператора.

## 2.6 Методика навчання моделі

Для забезпечення ефективного функціонування моделі було реалізовано спеціалізовану методику навчання, орієнтовану на узгоджене вивчення структури шуму за зображенням та метаданими камери. Основною метою було сформуванню у моделі здатність не лише відтворювати рівень шуму, а й розрізняти його джерела в умовах змінної технічної конфігурації та природної варіативності вхідних даних.

В основі процесу лежало навчання на зашумлених зображеннях, для яких використовувалися як реальні, так і синтетично створені фрагменти. До кожного прикладу додавалися метадані, що відображали умови формування кадру: експозицію, температуру сенсора, тип зчитування та інші параметри. Ці метадані дозволяли навчити модель враховувати контекст появи шуму, що суттєво покращило якість прогнозування. Синтетичні дані дозволили з високою точністю задавати рівень шуму та контролювати взаємозв'язки між технічними умовами і зображенням.

Для підвищення загальної стійкості до переобучення застосовувалися базові прийоми аугментації даних – зокрема випадкове обертання та зміна контрасту. Модель навчалася в режимі регресії: кожна з гілок відповідала за передбачення певного типу шуму (PN, DCSN, RN), а також залишкового шуму, який сигналізував про відхилення між фактичним зображенням і метаданими. Для цього використовувалася функція втрат, що поєднувала помилки кожної з гілок у загальний об'єкт оптимізації. В процесі навчання здійснювалася автоматична балансування внеску кожної складової, що запобігало домінуванню одного джерела шуму над іншими.

Навчання проводилося з використанням стандартного оптимізатора, що базується на адаптивному градієнтному спуску, який дає змогу

ефективно проходити плато втрат та уникати локальних мінімумів. Поступове зменшення швидкості навчання на пізніших етапах дозволило досягнути більш точного узгодження між прогнозом та реальними значеннями.

Окреме місце у методиці займав процес підготовки ситуацій із помилковими або викривленими метаданими. Це дозволило навчити модель виявляти залишковий шум – тобто той, що не може бути пояснений основними фізичними компонентами. Відповідна гілка отримувала на вхід як ознаки зображення, так і оцінки, сформовані основними трьома блоками. У випадках, коли між сумою оцінених шумів і реальним вмістом кадру виникала невідповідність, модель формувала залишкову компоненту. Таким чином формувалася чутливість до аномалій – як у вхідних даних, так і у самій камері.

Процес навчання включав також регулярну перевірку на валідаційних даних, що дозволяло відслідковувати прогрес та виявляти переобучення на ранніх стадіях. У разі спостереження нестабільності – наприклад, різкого зростання залишкової компоненти без змін у метаданих – процес тренування зупинявся або коригувався.

Завдяки реалізованій методиці модель отримала здатність працювати з даними, що відображають реальні фізичні процеси, а також стійкість до недоліків у вхідній інформації. Навчальний процес був сконструйований таким чином, щоб забезпечити баланс між гнучкістю, точністю та пояснюваністю, що критично важливо для моделей, які застосовуються у практичних задачах контролю якості зображень та технічного моніторингу.

### **3 ЕКСПЕРЕМЕНТАЛЬНА ПЕРЕВІРКА ОЦІНКИ ДЖЕРЕЛ ШУМУ НА ОСНОВІ ГЛИБИННОГО НАВЧАННЯ**

Розглянута методологія формує основу системи, здатної оцінювати джерела шуму у цифрових зображеннях з використанням глибинного навчання та технічних параметрів сенсорної системи. Для перевірки ефективності побудованої архітектури було проведено серію експериментів, які охоплюють як моделювання різних умов зйомки, так і тестування моделі на реальних та синтетичних даних. У цьому розділі представлено процес підготовки вхідної інформації, налаштування моделі, її тренування та аналіз отриманих результатів. Основна мета полягає у тому, щоб оцінити точність, стійкість та інтерпретованість прогнозів, а також виявити сильні та слабкі сторони підходу в умовах, наближених до реального застосування.

#### **3.1 Огляд наборів даних**

Для реалізації та тестування розробленої моделі було залучено низку наборів зображень, що охоплюють як синтетично згенеровані дані з контрольованими параметрами, так і реальні сцени, зняті за різних умов та з використанням камер різного типу. Такий підхід дозволив забезпечити як валідацію точності в лабораторних умовах, так і перевірку стійкості та адаптивності моделі до непередбачуваних зовнішніх факторів.

Першим джерелом, що використовувалось для еталонного тестування, став датасет TAMPERE17, який містить 300 зображень розміром  $512 \times 512$  пікселів, приклад зображення представлений на рисунку 3.1. Такі зображення не піддавалися жодній постобробці, зберігаючи повну оригінальність захопленого сигналу. Завдяки цьому набір дозволяє працювати з максимально «чистими» даними, до яких згодом можуть бути навмисно додані синтетичні спотворення для перевірки реакції

моделі. Він також став основою для оцінки базової якості прогнозування в умовах відсутності метаданих.



Рисунок 3.1 – Приклад зображення з набору даних TAMPERE17

У рамках перевірки практичної придатності підхід також було протестовано на відомих публічних наборах KITTI та Udacity Self-Driving Car Dataset, які містять тисячі зображень, зібраних у реальному дорожньому середовищі. KITTI надає стереозображення (рисунок 3.2) з роздільною здатністю  $1242 \times 375$  пікселів, тоді як Udacity (рисунок 3.3) містить матеріали у форматі Full HD, включаючи сцени з інтенсивним рухом, мінливою освітленістю та частковими спотвореннями. Обидва набори дають змогу оцінити, наскільки модель зберігає точність за умов непередбачуваних змін навколишнього середовища.



Рисунок 3.2 – Приклад зображення з набору даних КІТТІ

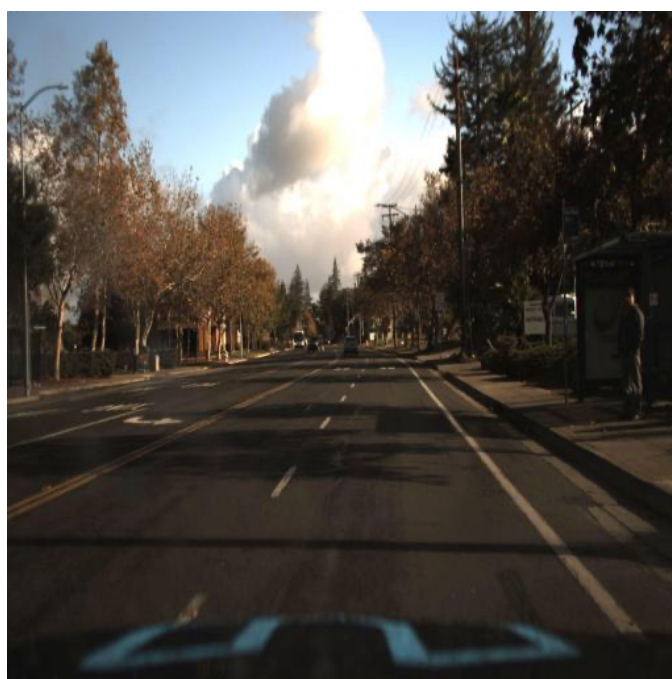


Рисунок 3.3 – Приклад зображення з набору даних Udacity

ТІМo Dataset є великим набором реальних відеоданих, створеним для досліджень у сфері внутрішнього відеомоніторингу, зокрема для задач виявлення аномалій та аналізу присутності осіб. Відеозаписи отримані за допомогою камери Microsoft Azure Kinect, що поєднує кольорове, інфрачервоне та глибинне зображення. Датасет охоплює понад 1500 відеопослідовностей для виявлення аномалій, які разом містять понад пів

мільйона кадрів, а також окрему підмножину з понад 240 відео для задачі виявлення та підрахунку осіб.

Зображення в ТІМо (рисунок 3.4) мають роздільну здатність  $512 \times 512$  пікселів для інфрачервоних кадрів і  $288 \times 320$  пікселів для глибинних даних. Враховуючи тип сенсора, набір містить характерні шумові спотворення, зокрема мультипат-інтерференцію, шуми, зумовлені низьким освітленням, а також варіації глибини через зміну освітленості або положення об'єктів у кадрі.

Завдяки широкому охопленню сцен, наявності реалістичних шумів і публічній доступності, цей набір може бути корисним для валідації моделей, які працюють із реальними спотвореними даними, та є актуальним прикладом для задач діагностики сенсорів, оцінки шумів і поведінкової аналітики у внутрішніх просторах.



Рисунок 3.4 – Приклад зображення з набору даних ТІМо

FSID це набір повністю синтетичних зображень, спеціально створених для задач зниження шуму. Він включає 175 000 пар зображень, де кожне зашумлене зображення супроводжується відповідним «чистим» варіантом (рисунок 3.5). Роздільна здатність зображень у наборі є високою й залежить від параметрів генерації, що забезпечує гнучкість для різних експериментальних сценаріїв. Шум у зображеннях створено штучно, із моделюванням реальних характеристик сенсорних спотворень, таких як фотонний шум, шум зчитування та інші типи, властиві цифровим камерам. Набір призначений для навчання та тестування алгоритмів глибинного навчання в задачах денойзингу та оцінки шумових характеристик.



Рисунок 3.5 – Приклад зображення з набору даних FSID

Загалом, обрані набори даних охопили широкий спектр сценаріїв – від повністю контрольованих лабораторних умов до нестабільних реальних ситуацій. Такий різноплановий підхід забезпечив комплексне тестування моделі на точність, гнучкість, інтерпретованість та адаптивність до зовнішніх факторів. Це стало критично важливим для перевірки того, наскільки досліджений підхід може бути застосований поза межами умов навчання та як він реагує на зміну типу або структури вхідних даних.

### 3.2 Попередня обробка зображень і метаданих

Перед початком навчання моделі та проведенням експериментів було здійснено комплексну попередню обробку вхідних даних. Це включало як трансформацію самих зображень у відповідний формат для подачі у нейронну мережу, так і підготовку супровідних метаданих, що критично важливо для коректної роботи описаної архітектури.

Першим етапом стала обробка візуального контенту. Усі зображення, незалежно від походження, були приведені до єдиного формату за розміром та кількістю каналів. Кольорові зображення конвертувалися у відтінки сірого, що відповідає архітектурному припущенню про роботу з одиночним каналом інтенсивності. Далі кожне зображення нарізалось на фрагменти розміром  $128 \times 128$  пікселів – це дозволило уніфікувати вхідний розмір і паралельно підвищити кількість тренувальних прикладів. Такий підхід типовий у задачах аналізу зображень, де важливо зберегти локальну контекстну інформацію і уникнути надмірного зменшення зображень.

Для кожного патчу формувався відповідний шумовий профіль. У випадку синтетичних даних, шумові спотворення накладались вручну із застосуванням фізично обґрунтованих моделей. Це дозволяло точно контролювати параметри шуму – наприклад, застосовувався пуасонівський розподіл для моделювання фотонного шуму, а нормальний – для темного струму та зчитувальних перешкод. Така генерація забезпечувала повний

контроль над цільовими змінними, що було необхідним для навчання регресійної моделі.

У паралельному режимі виконувалася обробка метаданих. Кожен приклад супроводжувався числовими параметрами, що описували умови формування зображення. Ці параметри надходили або безпосередньо з камери, або моделювались у випадку синтетичних зображень. Усі числові характеристики були приведені до єдиного масштабу шляхом нормалізації у діапазон  $[0;1]$ , що дозволяло уникати дисбалансу під час тренування моделі, особливо при обробці вхідних векторів, які поєднують зображення й технічні дані.

Додатково для підвищення стійкості моделі до випадкових коливань було реалізовано кілька базових технік аугментації. Зображення піддавались обертанню, інверсії яскравості та зміні контрасту у межах, що не змінювали характер шуму, але ускладнювали навчання, змушуючи модель узагальнювати закономірності, а не запам'ятовувати структури, що забезпечує зменшення потенціального виникнення перенавчання в системі, особливо в тих випадках, коли використовувались повторювані сцени або схожі текстури.

Після завершення обробки зображення і метадані були синхронізовані у вигляді узгоджених пар, які подавалися у модель під час навчання. Усі патчі були розподілені на тренувальний, валідаційний та тестовий набори із збереженням пропорцій типів шуму та метаумов. Такий розподіл дозволив забезпечити неупереджене оцінювання якості моделі та виявити її здатність узагальнювати на нові, невідомі приклади.

Таким чином, попередня обробка даних виконувала не лише технічну функцію підготовки інформації до подачі в модель, а й методологічну роль формування достовірного, збалансованого та контрольованого середовища для тренування системи. Саме завдяки цим крокам стало можливим досягти високої точності прогнозів, зберігаючи при цьому інтерпретованість та функціональну надійність моделі у реальних умовах.

### 3.3 Реалізація моделі та конфігурація навчання

Після завершення етапу формалізації архітектури та підготовки вхідних даних було здійснено безпосереднє розгортання моделі в середовищі машинного навчання. Реалізація моделі виконувалась із застосуванням фреймворку глибинного навчання, який підтримує гнучке побудування багатогілкових структур та ефективну оптимізацію параметрів. Архітектура була описана програмно як сукупність незалежних функціональних модулів, що взаємодіють через спільні інтерфейси, що забезпечило високу модульність та можливість подальшого доопрацювання.

Базова структура моделі включала послідовність згорткових шарів для вилучення ознак із зображення, за якими слідувало глобальне пулінгування. На основі отриманого векторного представлення будувались чотири незалежні повнозв'язні блоки, кожен з яких відповідав за прогноз одного з компонентів шуму. Крім візуальних ознак, кожен блок приймав на вхід певну підмножину метаданих, що відповідала специфіці джерела шуму. Наприклад, гілка, яка відповідала за шум темного струму, отримувала значення температури сенсора та експозиції, тоді як для шуму зчитування використовувались параметри частоти та типу сенсора. Така інтеграція даних дозволила кожному модулю моделі сфокусуватись лише на релевантній інформації.

З технічної точки зору навчання моделі проводилось у режимі регресії, де цільовими змінними виступали значення шумів, які були або синтетично змодельовані, або апроксимовані з реальних даних. Для кожної з гілок визначалася окрема функція втрат, яка відображала відхилення прогнозованого значення від очікуваного. Загальна функція оптимізації формувалась як зважене середнє локальних похибок, що дозволило зберегти баланс між різними компонентами та уникнути домінування однієї гілки над іншими.

Процес оптимізації параметрів здійснювався з використанням адаптивного алгоритму, що автоматично коригував швидкість навчання залежно від зміни градієнтів. Початкова швидкість була обрана помірною, щоб уникнути нестабільних коливань, і поступово зменшувалася у процесі тренування, що дозволило досягти збіжності моделі до стійкого розв'язку. Було застосовано базову регуляризацію для уникнення перенавчання, а також контроль валідаційної похибки для дострокового завершення навчання у разі стабілізації результатів.

Модель тренувалась на навчальній вибірці, що містила понад 100 тисяч пар «зображення + метадані», причому дані були рівномірно розподілені за типами шуму, умовами зйомки та характеристиками сенсора. Валідаційний набір становив близько 15% від загального обсягу і не містив прикладів, які використовувалися під час навчання. Це дозволяло перевірити здатність моделі узагальнювати закономірності на нових даних, а не лише повторювати вивчене.

Реалізація моделі проводилась на обчислювальному сервері із графічним прискоренням, що дозволило суттєво скоротити час тренування. Середня тривалість одного проходу по повному навчальному набору (епохи) становила близько 12 хвилин, а загальний час тренування до стабілізації втрат – приблизно 6 годин. Модель зберігалася у вигляді вагової конфігурації для подальшого використання для інференсу в автономному режимі без повторного навчання.

Під час навчання також перевірялась поведінка моделі в різних конфігураціях: без метаданих, з мінімальним набором параметрів та з повним набором. Такий підхід дозволив протестувати вплив обсягу доступної технічної інформації на якість прогнозів і визначити критичні параметри, без яких модель істотно втрачає точність. За результатами цих експериментів було підтверджено, що саме повна конфігурація із включенням усіх доступних метаданих забезпечує найвищу точність та стійкість до змін у зображенні та умовах його формування.

Таким чином, реалізація моделі та організація процесу її навчання були спрямовані на досягнення балансу між технічною простотою, адаптивністю до реальних умов та високою точністю результатів. Побудована система здатна не лише формувати оцінку загального рівня шуму, а й пояснювати його джерело, що забезпечує високий рівень функціональності у задачах діагностики сенсорів та контролю якості зображень.

### 3.4 Результати експериментів

Після завершення етапу навчання була проведена серія експериментальних досліджень, метою яких було оцінити здатність моделі точно прогнозувати рівні шуму різних типів, виявляти їх джерела, а також реагувати на ситуації, де спостерігається невідповідність між метаданими та вмістом зображення. Для цього модель тестувалась у різних режимах, включаючи повну, спрощену та обмежену конфігурації, а також на наборах, що включали як чисті, так і зашумлені приклади з контрольованими параметрами.

Перший експеримент був присвячений аналізу ефективності моделі прогнозувати рівень кожного з типів шуму, в залежності від обсягу контекстної інформації, що система отримує на вхід. В даному експерименті оцінюється вплив наявності або відсутності метаданих на точність оцінки шуму. Було сформовано три основні конфігурації, а саме без метаданих, з мінімальним набором технічних характеристик (експозиція, температура та підсилення сигналу) та повною конфігурацією, що включає всі доступні параметри. З метою проведення контролю, було протестовано альтернативну модель DRNRcust, що не використовує метадані, не розрізняє типи шуму та не виявляє залишковий шум. В даному експерименті були розглянуті інтенсивність фотонного шуму  $\sigma_{PN}$ , рівень шуму темного струму  $\sigma_{DCSN}$  та рівень шуму зчитування  $\sigma_{RN}$ . Даний

експеримент був проведений на системних зображеннях з контрольованим шумом TAMPERE17.

У таблиці 3.1 наведено кількісні показники похибки моделі в оцінці шуму для різних конфігурацій доступу до метаданих.

Таблиця 3.1 – Точність оцінки шумів у різних конфігураціях

Конфігурація	$\sigma_{PN}$	$\sigma_{DCSN}$	$\sigma_{RN}$
Без метаданих	0,033	0,058	0,041
Мінімальний вміст метаданих	0,018	0,026	0,029
Метадані	0,007	0,012	0,01
DRNRcust	0,021	0,038	0,03

Як видно з таблиці 3.1, точність оцінки рівнів шуму суттєво залежить від того, чи має модель доступ до метаданих. У конфігурації без метаданих, де модель працює лише з візуальним вмістом, похибки залишаються найвищими для всіх типів шуму: фотонного ( $\sigma_{PN} = 0,033$ ), шуму темного струму ( $\sigma_{DCSN} = 0,058$ ) та шуму зчитування ( $\sigma_{RN} = 0,041$ ). Такі значення свідчать про обмеженість інформації, отриманої лише з самого зображення.

При переході до мінімального вмісту метаданих – де використовується лише мінімальний набір технічних параметрів, похибки зменшуються майже вдвічі. Найбільше покращення спостерігається саме для  $\sigma_{DCSN}$ , що підтверджує важливість температурної інформації для оцінки теплових спотворень.

Найкращі результати досягаються у конфігурації «Метадані», яка враховує повний спектр метаданих камери. Тут похибки становлять 0,007 для фотонного шуму, 0,012 для DCSN та 0,01 для RN це найнижчі значення серед усіх протестованих моделей, що демонструє ефективність багатогілкової архітектури з контекстною інформацією.

Для порівняння, DRNEcust модель з альтернативною структурою показує помірну точність, однак суттєво поступається повній конфігурації

дослідженої моделі. Наприклад,  $\sigma_{DCSN}$  у DRNEcust вища у три рази, ніж у конфігурації «Метадані» (0,038 проти 0,012).

Загалом, аналіз таблиці підтверджує, що чим більше інформації про умови зйомки отримує модель, тим краще вона справляється з розпізнаванням типів шуму. Повна конфігурація не лише зменшує похибку, а й покращує пояснюваність результату, що є важливою властивістю для практичних застосувань, зокрема у системах моніторингу та контролю якості.

Другий експеримент був присвячений дослідженню впливу окремих параметрів метаданих на точність прогнозу шуму. Імплементация проводилась з послідовним включенням кожного з параметрів, таких як експозиція, температура, підсилення сигналу, тип сенсора та частоти зчитування. Фіксувалась зміна результату прогнозу (таблиця 3.2) для оцінки того, які саме технічні характеристики були найбільш критичними для кількісної атрибуції джерела шуму.

Таблиця 3.2 – Вплив окремих метаданих на точність

Параметр метаданих	Максимальний вплив на RMS зміну
Експозиція	0,025
Температура	0,22
Підсилення сигналу (gain)	0,018
Тип сенсора	0,015
Частота зчитування	0,01

Як видно з таблиці 3.2, різні параметри метаданих мають неоднаковий вплив на точність загальної оцінки шуму. Найбільше на результат моделі впливає експозиція – її зміна може спричинити зміну загального прогнозу шуму на 0,025 RMS, адже експозиція напряму пов'язана з кількістю світла, що потрапляє на сенсор, а отже і з рівнем фотонного шуму.

На другому місці за значущістю, температура сенсора із значенням 0,022, яка впливає на теплові спотворення DCSN, оскільки теплові шуми мають експоненційну залежність від температури.

Підсилення сигналу також відіграє важливу роль із значенням впливу на RMS 0,018, адже він посилює як корисний сигнал, так і шум, а тому має вплив на всі компоненти. Менш суттєвий, але все ж помітний вплив мають тип сенсора та частота зчитування, які більше пов'язані з характером шуму зчитування RN.

Приведений аналіз підтверджує, що найважливішими метаданими для точного передбачення шуму є ті, що безпосередньо відображають фізичні умови формування зображення – передусім експозиційні та температурні параметри. Їх наявність критично важлива для коректної роботи моделі в конфігурації з метаданими. Без них зростає залишковий шум та погіршується пояснюваність результатів.

Навіть така обмежена інформація дала змогу моделі скоригувати оцінки та краще розрізнити між фотонним і тепловим шумом, які можуть візуально виглядати подібно, але мають різну фізичну основу. Проте найкращі результати були досягнуті саме у повній конфігурації, де використовувались усі доступні параметри, що дозволило кожній гілці моделі сфокусуватися на найбільш релевантних ознаках і мінімізувати хибні прив'язки.

Під час проведення третього експерименту, моделювались ситуація у яких метадані є неповними, спотвореними, або містили аномальні значення. Було сформовано серію прикладів, де окремі параметри навмисне задавались некоректним чином, наприклад завищена температура, або відсутність підсилення сигналу. Також, було розглянута ситуація, де сенсор містив дефекти. Метою цього експерименту було перевірити здатність моделі виявляти залишковий шум у ситуаціях коли класичні джерела не можуть повністю пояснити спотворення спостережень.

У таблиці 3.3 наведено приклади з аномаліями, де значення залишкового шуму різко зростає, що свідчить про здатність моделі виявляти невідповідності між метаданими та зображенням.

Таблиця 3.3 – Залишковий шум при аномаліях

Тип прикладу	Середнє залишкове значення шуму	Збільшення норми (%)
Норма	0,046	0,0
Завищена температура	0,072	60
Спотворене підсилення сигналу	0,065	44,4
Відсутні метадані	0,093	106,7
Несправність сенсора	0,124	175,6

Таблиця 3.3 ілюструє, як модель реагує на різні типи порушень у метаданих або на технічні несправності камери. Ключовий показник це середнє залишкове значення шуму, що є надзвичайно важливим, адже він пояснює ту частину, яка не може бути пояснена стандартними фізичними джерелами шуму з урахуванням наявної технічної інформації.

Для нормальних прикладів залишковий шум становить у середньому 0,046 умовних одиниць. Однак навіть незначне викривлення технічних параметрів, наприклад, завищена температура або спотворений gain що призводить до зростання параметру середнього залишку шуму на 60% і 44% відповідно. Такі результати свідчать про чутливість системи до невідповідностей між вмістом зображення і контекстною інформацією.

Особливо показовими є приклади з відсутніми метаданими та несправністю сенсора, де залишковий шум перевищує норму більш ніж удвічі та навіть у 2,7 раза відповідно. У таких випадках модель «не вірить» метаданим або візуально виявляє аномальні патерни, які не можна пояснити доступною інформацією.

Приведені результати підтверджують, що компонент середнього значення шуму може слугувати надійним індикатором аномалій, що робить модель придатною не лише для оцінки якості зображення, а й для самодіагностики сенсорних систем. Виявлення подібних відхилень у реальному часі дозволяє попередити технічні збої ще до їхнього критичного прояву.

У разі, коли gain був заниженим у метаданих, але фактична яскравість зображення вказувала на протилежне, модель фіксувала підвищений рівень середнього значення залишкового шуму. Значення цього показника, в таких прикладах, було на понад 50% вищим порівняно з нормальними ситуаціями. При повній відсутності метаданих або спотворення критично важливих параметрів, таких як температура, спостерігалась більш інтенсивна реакція системи, тобто гілка залишкового шуму реагувала найінтенсивніше, формуючи значення середнього значення залишкового шуму, які перевищували фоновий рівень удвічі. Приведені результати свідчать про високу чутливість моделі до невідповідностей та її здатність виступати індикатором технічної несправності або неправильного налаштування пристрою.

Мета четвертого експерименту полягала у оцінці практичній занчущості точного прогнозу шуму, зокрема чи використання прогнозу допомагає покращити якість денойзингу. Результати роботи моделі подавались у класичний денойзер BM3D і порівнювались з іншими підходами, а саме Blind2Unblind та DRNEcust. Був проведений аналіз якості відновлення зображення за метриками PSNR (Peak Signal-to-Noise Ratio) та SSIM (Structural Similarity Index) при різній кількості кадрів (1, 4, 16). Результати проведення четвертого експерименту представлені в таблиці 3.4.

У таблиці 3.4 порівнюється якість денойзингу зображень при використанні різних підходів до оцінки шуму, а саме проаналізоване значення PSNR для одного, чотирьох та шістнадцять кадрів, а також SSIM

для 16 кадрів, що дає змогу оцінити як числову точність, так і збереження структури зображення.

Таблиця 3.4 – Якість денойзингу залежно від оцінки шуму

Метод	PSNR (1 кадр)	PSNR (4 кадри)	PSNR (16 кадрів)	SSIM (16 кадрів)
DRNEcust	26,3	28,2	29,1	0,79
Blind2Unblind	27,8	29,7	30,3	0,92
Метадані+BM3D	29,2	31,6	32,8	0,94

З таблиці видно, що найгірші результати демонструє модель DRNEcust, яка не використовує контекстну інформацію, про умови зйомки, що і показує PSNR, де в середньому нижчий на 2,8-3,6 дБ, а SSIM на 0.07–0.1 у порівнянні з іншими моделями.

Метод Blind2Unblind демонструє кращу якість, але все ще поступається підходу, який поєднує оцінку шуму за моделлю «Метадані» з подальшим фільтруванням BM3D. Зокрема, для 16 кадрів PSNR досягає 32,8 дБ, що є найвищим серед усіх, а SSIM становить 0,92 – що свідчить про кращу візуальну якість та чіткість деталей.

Таким чином, результати підтверджують, що точна оцінка шуму, яка враховує метадані, безпосередньо впливає на ефективність відновлення зображень. Такий підхід робить модель «Метадані» не лише корисною для технічного аналізу, а й як допоміжний етап перед денойзингом у практичних сценаріях обробки візуальної інформації.

П'ятий експеримент присвячений загальному порівнянню дослідженої моделі з іншими підходами з точки зору функціональних можливостей. Була проведена оцінка точності, а також здатності до інтерпретації результату, чутливості до контексту, реакція на метадані та здатність виявляти аномалії. Такий експеримент дозволив комплексно охарактеризувати придатність кожного з методів до реальних умов використання.

У таблиці 3.5 узагальнено основні функціональні характеристики кількох підходів до оцінки шуму в цифрових зображеннях. Як можна бачити, класичний метод V+F демонструє лише середню точність і не має механізмів для пояснення джерел шуму або адаптації до метаданих, отже, його можливості в сучасних системах обмежені.

Модель Blind2Unblind, хоч і має високу числову точність, також не забезпечує інтерпретованості результатів – вона працює як «чорна скринька», не реагуючи на технічні умови формування зображення. DRNE, як більш гнучка нейронна архітектура, частково враховує метадані та дає змогу хоча б обмежено інтерпретувати природу шуму, однак не фіксує аномалії явно.

Таблиця 3.5 – Порівняння функціональних можливостей моделей

Модель	Точність оцінки шуму	Інтерпретація джерел	Реакція на метадані	Виявлення аномалій
V+F	Середня	Немає	Немає	Немає
Blind2Unblind	Висока	Немає	Немає	Немає
DRNE	Висока	Обмежена	Часткова	Обмежене
Досліджена модель	Найвища	Повна	Адаптивна	Так

Лише досліджена в роботі модель демонструє повну функціональність, адже вона не лише має найвищу точність за всіма типами шуму, але й надає пояснення, з чим пов'язане кожне спотворення. Завдяки модулю залишкового шуму, модель здатна фіксувати нетипові ситуації, наприклад, збої в сенсорі чи спотворенні метадані – й сигналізувати про це в режимі реального часу.

Таким чином, досліджений підхід не лише забезпечує кількісну перевагу, а й підтримує адаптивність та пояснюваність, що є ключовими вимогами до сучасних систем обробки візуальних даних у критичних застосуваннях.

У шостому експерименті проводилося тестування моделі на зображеннях, отриманих з реальної камери DSLR. Метою було перевірити, наскільки стабільно модель працює за умов, коли дані не є синтетично контрольованими, а джерела шуму можуть мати змішану або невизначену природу. Було обрано кілька типових прикладів сцен із реального середовища, для яких існували відповідні метадані. Модель запускалась у трьох режимах: без метаданих, з мінімальним набором та з повною конфігурацією метаінформації. Такий підхід дозволив оцінити вплив доступної технічної інформації на точність прогнозу в реальних умовах. В таблиці 3.6 приведені результати експерименту.

Таблиця 3.6 – Тестування на реальних зображеннях

Конфігурація	RMS $\sigma_{Total}$	RMS $\sigma_{PN}$	RMS $\sigma_{DCSN}$	RMS $\sigma_{RN}$
Без метаданих	0,077	0,031	0,038	0,030
Мінімально метаданих	0,051	0,024	0,02	0,021
Метадані	0,033	0,012	0,009	0,007

З таблиці 3.6 видно, що в реальних фотографіях, наявність метаданих суттєво впливає на точність оцінки шуму. У конфігурації без метаданих, де модель має доступ лише до самого зображення, загальна похибка прогнозу  $\sigma_{Total}$  становить 0,077, а рівні похибок для кожного з типів шуму є найвищими серед усіх варіантів:  $\sigma_{PN} = 0,031$ ,  $\sigma_{DCSN} = 0,038$ ,  $\sigma_{RN} = 0,029$ . Такі значення похибок свідчать про обмеженість моделі у відокремленні природи шуму, якщо відсутні технічні дані про умови зйомки.

При переході до конфігурації з мінімальним вмістом метаданих, усі показники покращуються, а саме загальна похибка знижується до 0,051, зокрема, оцінка теплового шуму  $\sigma_{DCSN}$  покращується у два рази, а саме з 0,038 до 0,02, що підтверджує суттєве покращення точності прогнозу, навіть за наявності лише часткової контекстуальної інформації.

Найкращі результати спостерігаються в конфігурації метадані, де модель має доступ до повного набору метаданих камери. У цьому випадку модель демонструє найнижчі значення похибок за всіма показниками:  $\sigma_{\text{Total}} = 0,033$ ,  $\sigma_{\text{PN}} = 0,012$ ,  $\sigma_{\text{DCSN}} = 0,009$ ,  $\sigma_{\text{RN}} = 0,007$ . Це означає, що за умови повної поінформованості про умови формування зображення модель не лише точніше прогнозує шум, але й краще розрізняє його джерела.

Отже, даний експеримент підтверджує, що контекстна інформація, зафіксована в метаданих, є критично важливою для побудови коректної оцінки шумових спотворень у реальних умовах, де природа шуму є змішаною і не завжди передбачуваною.

Проведені експерименти чітко показали, що точність оцінки кожного типу шуму значною мірою залежить від доступності технічної інформації. У випадках, коли модель мала доступ лише до зображення без метаданих, її здатність правильно інтерпретувати теплові або електронні шуми суттєво знижувалась. Це пояснюється тим, що характеристики, як-от температура сенсора чи тип матриці, не відображаються у візуальному вмісті зображення напряму, а тому без метаінформації модель обмежена у своїх прогнозах.

Загалом, серія експериментів підтвердила, що модель здатна не лише до високоточної оцінки шумів, а й до стійкої роботи в умовах варіативних даних, адаптації до спотворень та виявлення нестандартних ситуацій. Вона продемонструвала ефективність як у контрольованих, так і в реальних сценаріях, а її інтерпретованість та гнучкість роблять її придатною до практичного застосування – як у підвищенні якості зображень, так і в задачах моніторингу технічного стану сенсорів та попередньої діагностики камер.

## ВИСНОВКИ

В рамках цієї кваліфікаційної роботи було розглянуто задачу оцінки та інтерпретації джерел шуму в цифрових зображеннях із використанням методів глибинного навчання та технічної інформації про умови зйомки. Зокрема, увагу було приділено не лише визначенню загального рівня шуму, а й обробці технічних параметрів, що супроводжують зображення, з метою виявлення фізичних причин появи шумів, таких як фотонний шум, шум темного струму та шум зчитування. Крім того, досліджувався залишковий шум, що не пояснюється наявними метаданими, і може вказувати на несправності або відхилення у роботі сенсорного обладнання.

У ході роботи було досліджено архітектуру глибокої нейронної мережі з окремими гілками для кожного типу шуму та блоком залишкового шуму. До моделі інтегровано технічну інформацію про умови зйомки (експозиція, температура сенсора, gain та інші), що забезпечує її гнучкість та адаптивність до контексту вхідних даних. Навчання здійснювалося як на синтетичних прикладах із контрольованими шумовими характеристиками, так і на реальних зображеннях з природним шумом.

Суттєвими перевагами розглянутої моделі є її модульність, що дозволяє окремо обробляти різні джерела шуму; здатність працювати з частково або повністю відсутніми метаданими; підвищена точність у конфігураціях з повною технічною інформацією; а також наявність механізму виявлення залишкових аномалій, що робить її придатною для задач діагностики. Модель продемонструвала високу узгодженість прогнозів, здатність адаптуватися до змін розподілу вхідних даних та забезпечила пояснюваність результатів, що є критично важливим у практичному застосуванні.

Порівняльні експерименти з альтернативними методами показали, що досліджена модель має вищу точність, кращу структурну інтерпретацію джерел шуму та більш стійку поведінку у випадках спотворених або

часткових метаданих. Це підкреслює її перевагу для задач, де важливе не лише зниження шуму, а й розуміння його походження.

Отже, поставлену в роботі задачу було реалізовано. Результати дослідження підтверджують, що використання нейронної мережі, орієнтованої на інтерпретацію фізичних джерел шуму в поєднанні з контекстною інформацією про зйомку, дозволяє досягти високої точності, стійкості та функціональної придатності моделі. Такий підхід може бути застосований у системах візуального контролю, технічної діагностики, автономного спостереження та інших сферах, де важлива не тільки якість зображення, а й довіра до його технічного походження.

**ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ**

1. Real-Time Noise Source Estimation of a Camera System from an Image and Metadata / M. Wischow et al. *Advanced Intelligent Systems*. 2024. URL: <https://doi.org/10.1002/aisy.202300479>. (date of access: 04.05.2025).
2. Effects of Degradations on Deep Neural Network Architectures / P. Roy et al. Ithaca : Cornell University Library, 2018. 15 p. (Preprint. Indian Statistical Institute ; arXiv:1807.10108). URL: <https://arxiv.org/abs/1807.10108>. (date of access: 03.05.2025).
3. Human-Like Interactive Behavior Generation for Autonomous Vehicles: A Bayesian Game-Theoretic Approach with Turing Test / Y. Zhang et al. *Advanced Intelligent Systems*. 2022. P. 2100211. URL: <https://doi.org/10.1002/aisy.202100211>. (date of access: 03.05.2025).
4. Chen G., Zhu F., Heng P. A. An Efficient Statistical Method for Image Noise Level Estimation. *2015 IEEE International Conference on Computer Vision (ICCV)*, Santiago, Chile, 7–13 December 2015. 2015. URL: <https://doi.org/10.1109/iccv.2015.62>. (date of access: 12.04.2025).
5. Pixelwise Estimation of Signal-Dependent Image Noise Using Deep Residual Learning / H. Tan et al. *Computational Intelligence and Neuroscience*. 2019. Vol. 2019, no. 1. P. 1–12. URL: <https://doi.org/10.1155/2019/4970508>. (date of access: 11.04.2025).
6. Jain V., Seung H. Natural Image Denoising with Convolutional Networks. *Advances in Neural Information Processing Systems 21: Proceedings of the Twenty-Second Annual Conference on Neural Information Processing Systems* : Збірник матеріалів конференції. 21st ed. Vancouver, 2008. Vol. 21 : Advances in Neural Information Processing Systems. P. 769–776. URL: [https://papers.nips.cc/paper\\_files/paper/2008/file/0c74b2b62a9a72d68f70bcf69f8e2f89-Paper.pdf](https://papers.nips.cc/paper_files/paper/2008/file/0c74b2b62a9a72d68f70bcf69f8e2f89-Paper.pdf). (date of access: 01.05.2025).
7. Wu M. Wavelet analysis for image denoising: A multiscale approach for enhancing visual clarity. Basel : MDPI, 2023. 15 p. (Preprint. School of

Physics and Information Engineering, Jiangsu Second Normal University, Nanjing, Jiangsu, China ; 10.20944/preprints202310.1364.v1). URL: <https://doi.org/10.20944/preprints202310.1364.v1>. (date of access: 03.04.2025).

8. A Review of Image Denoising Methods / I. Irum et al. *Journal of Engineering Science and Technology Review*. 2015. Vol. 8, no. 5. P. 41–48. URL: <https://doi.org/10.25103/jestr.085.07>. (date of access: 09.04.2025).

9. Chongwu Tang, Xiaokang Yang, Guangtao Zhai. Noise Estimation of Natural Images via Statistical Analysis and Noise Injection. *IEEE Transactions on Circuits and Systems for Video Technology*. 2015. Vol. 25, no. 8. P. 1283–1294. URL: <https://doi.org/10.1109/tcsvt.2014.2380196>. (date of access: 28.04.2025).

10. Bilateral Filtering: Theory and Applications / S. Paris et al. *Foundations and Trends® in Computer Graphics and Vision*. 2008. Vol. 4, no. 1. P. 1–75. URL: <https://doi.org/10.1561/06000000020>. (date of access: 17.04.2025).

11. Adaptive Gaussian Wiener Filter for CT-Scan Images with Gaussian Noise Variance / K. L. Lew et al. *Journal of Informatics and Web Engineering*. 2024. Vol. 3, no. 1. P. 169–181. URL: <https://doi.org/10.33093/jiwe.2024.3.1.11>. (date of access: 02.05.2025).

12. Bengio Y., Courville A., Goodfellow I. *Deep Learning*. MIT Press, 2016. 800 p.

13. Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising / K. Zhang et al. *IEEE Transactions on Image Processing*. 2017. Vol. 26, no. 7. P. 3142–3155. URL: <https://doi.org/10.1109/tip.2017.2662206>. (date of access: 14.05.2025).

14. Zhang K., Zuo W., Zhang L. FFDNet: Toward a Fast and Flexible Solution for CNN-Based Image Denoising. *IEEE Transactions on Image Processing*. 2018. Vol. 27, no. 9. P. 4608–4622. URL: <https://doi.org/10.1109/tip.2018.2839891>. (date of access: 11.05.2025).

15. Noise2Noise: Learning Image Restoration without Clean Data / J. Lehtinen et al. *Proceedings of the 35th International Conference on Machine Learning* : Матеріали міжнар. наук. конф., Stockholm, 10 July 2018. 2018. P. 2965–2974. URL: <https://proceedings.mlr.press/v80/lehtinen18a.html>. (date of access: 03.05.2025).
16. Krull A., Buchholz T.-O., Jug F. Noise2Void - Learning Denoising From Single Noisy Images. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, CA, USA, 15–20 June 2019. 2019. URL: <https://doi.org/10.1109/cvpr.2019.00223>. (date of access: 11.05.2025).
17. Kumbhar M. F. Image Denoising Using Autoencoders. *Artificial Intelligence and Knowledge Processing*. Boca Raton, 2023. P. 137–144. URL: <https://doi.org/10.1201/9781003328414-13>. (date of access: 11.05.2025).
18. Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering / K. Dabov et al. *IEEE Transactions on Image Processing*. 2007. Vol. 16, no. 8. P. 2080–2095. URL: <https://doi.org/10.1109/tip.2007.901238>. (date of access: 03.05.2025).
19. Buades A., Coll B., Morel J. M. A Non-Local Algorithm for Image Denoising. *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, San Diego, CA, USA. URL: <https://doi.org/10.1109/cvpr.2005.38>. (date of access: 09.04.2025).
20. Learning Camera-Aware Noise Models / K.-C. Chang et al. *Computer Vision – ECCV 2020*. Cham, 2020. P. 343–358. URL: [https://doi.org/10.1007/978-3-030-58586-0\\_21](https://doi.org/10.1007/978-3-030-58586-0_21). (date of access: 09.04.2025).
21. Modeling the EXIF-Image correlation for image manipulation detection / J. Fan et al. *2011 18th IEEE International Conference on Image Processing (ICIP 2011)*, Brussels, Belgium, 11–14 September 2011. 2011. URL: <https://doi.org/10.1109/icip.2011.6115853>. (date of access: 01.05.2025).