

Міністерство освіти і науки України
Харківський національний університет імені В. Н. Каразіна
Навчально-науковий інститут комп'ютерних наук та штучного інтелекту
Кафедра комп'ютерних систем та робототехніки

До захисту допущено
Кафедрою комп'ютерних систем та робототехніки
протокол № від грудня 2025р.

завідувач кафедри _____ Максим ХРУСЛОВ
(підпис)

«__» _____ 2025 р.

Кваліфікаційна робота
здобувача другого (магістерського) рівня вищої освіти
«МЕТОД РОЗПІЗНАВАННЯ ОБ'ЄКТІВ НА
ЗАШУМЛЕНИХ ЗОБРАЖЕННЯХ»

Спеціальність 123 – *Комп'ютерна інженерія*
Освітня програма *Комп'ютерна інженерія*

Виконавець _____ Данило ЯКОВЛЕВ
(підпис)

Науковий керівник _____ Вікторія СТРИЛЕЦЬ
(підпис)

Харків – 2025

АНОТАЦІЯ

Пояснювальна записка до магістерської кваліфікаційної роботи складається зі вступу, трьох розділів, висновків, списку використаних джерел і трьох додатків. Загальний обсяг роботи складає 66 сторінок, із яких 50 сторінок основної частини з 11 рисунками, 4 таблицями, 20 найменувань списку використаних джерел та трьома додатками.

Метою кваліфікаційної роботи є підвищення точності відновлення та розпізнавання об'єктів на зашумлених зображеннях шляхом дослідження та застосування сучасних методів обробки, включаючи традиційні, інтелектуальні та гібридні підходи.

Об'єкт дослідження – процес розпізнавання цифрових зображень, які містять різні типи шумів, що спотворюють візуальну інформацію та ускладнюють автоматичне розпізнавання об'єктів.

Предмет дослідження – методи та алгоритми обробки зашумлених зображень, включаючи традиційні фільтри, глибокі моделі та гібридні підходи.

У роботі розглядається проблема зниження точності автоматизованих систем розпізнавання через наявність різних типів шумів на зображеннях. Шумові спотворення, такі як гауссівський шум, імпульсний шум, спекл-шум, шум Пуассона, значно ускладнюють виділення корисних ознак та погіршують якість класифікації. Актуальність проблеми зумовлена широким використанням систем комп'ютерного зору в реальних умовах, де зашумленість даних є типовим явищем.

Сфера використання розроблених методів охоплює відеоспостереження, медичну діагностику, дистанційне зондування Землі, системи безпеки, транспортні системи, промисловий моніторинг та інші галузі, де критичною є якість вхідних зображень.

Ключові слова: *TensorFlow, Python, Machine Learning, CNN, Autoencoder, ResNet-18, шум, зашумлені зображення.*

ABSTRACT

The explanatory note to the master's qualification thesis consists of an introduction, three chapters, conclusions, a list of references, and three appendices. The total volume of the work is 66 pages, including 50 pages of the main part with 11 figures, 4 tables, 20 references, and three appendices.

The aim of the qualification thesis is to improve the accuracy of restoration and recognition of objects in noisy images by investigating and applying modern processing methods, including traditional, intelligent, and hybrid approaches.

Object of study – recognition process of the digital images that contain various types of noise which distort visual information and complicate automatic object recognition.

Subject of study – methods and algorithms for processing noisy images, including traditional filters, deep learning models and hybrid approaches.

The work addresses the problem of reduced accuracy in automatic recognition systems caused by the presence of various types of noise in images. Noise distortions, such as Gaussian noise, impulse noise, speckle noise, and Poisson noise, significantly complicate the extraction of meaningful features and deteriorate classification performance. The relevance of this problem is determined by the widespread use of computer vision systems in real-world conditions, where image degradation caused by noise is a common occurrence.

The scope of application for the developed methods includes video surveillance, medical diagnostics, remote sensing of the Earth, security systems, transportation systems, industrial monitoring, and any other fields where the quality of input images is the critical factor.

Keywords: *TensorFlow, Python, Machine Learning, CNN, Autoencoder, ResNet-18, noise, noisy images.*

ЗМІСТ

ВСТУП	6
РОЗДІЛ 1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ ТА ОГЛЯД ІНСУЮЧИХ МЕТОДІВ РОЗПІЗНАВАННЯ ОБ'ЄКТІВ НА ЗАШУМЛЕНИХ ЗОБРАЖЕННЯХ	8
1.1 Загальне значення шуму та зашумлених зображень	8
1.2. Види шумів на зображеннях	9
1.3. Класифікація методів обробки зашумлених зображень.	14
1.3.1. Традиційні методи обробки.....	14
1.3.2. Сучасні інтелектуальні методи	15
1.3.3. Гібридні підходи обробки зображень.....	18
Висновки до розділу 1	19
РОЗДІЛ 2. МЕТОДИ ТА МОДЕЛІ КЛАСИФІКАЦІЇ ОБ'ЄКТІВ НА ЗАШУМЛЕНИХ ЗОБРАЖЕННЯХ	20
2.1. Постановка задачі класифікації зображень	20
2.2. Класичні методи класифікації зображень та їх обмеження при роботі з шумом.....	21
2.3. Глибокі моделі класифікації та обробки зашумлених зображень	23
2.3.1. Згорткові нейронні мережі	24
2.3.2. Модель DnCNN.....	26
2.3.3. Модель ResNet-18.....	29
2.4. Автоенкодери та гібридні підходи для роботи із зашумленими зображеннями.....	32
Висновок до розділу 2	34
РОЗДІЛ 3. РОЗРОБКА ТА ДОСЛІДЖЕННЯ МОДЕЛІ КЛАСИФІКАЦІЇ ОБ'ЄКТІВ НА ЗАШУМЛЕНИХ ЗОБРАЖЕННЯХ	35
3.1. Постановка задачі	35
3.2. Підготовка даних для навчання моделей.....	35
3.3. Використані програмні засоби та інструменти.....	36
3.4. Побудова моделей класифікації зашумлених зображень.....	37

3.4.1. Базова згорткова нейронна мережа (CNN)	38
3.4.2. Побудова моделі ResNet-18	40
3.4.3. Гібридна модель Autoencoder + CNN	43
3.5. Аналіз результатів навчання і тестування моделей	45
3.5.1. Результати роботи базової моделі CNN	45
3.5.2. Результати роботи моделі ResNet-18	47
3.5.3. Результати роботи гібридної моделі	48
Висновок розділу 3	49
ВИСНОВКИ.....	51
ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ	53
ДОДАТКИ	53

ВСТУП

У сучасному світі цифрові зображення є невід’ємною частиною практично всіх сфер людської діяльності – від медицини й транспорту до промисловості, освіти та розваг. З розвитком технологій комп’ютерного зору та машинного навчання зображення стали не лише засобом передачі інформації, але й ключовим джерелом даних для інтелектуальних систем. Завдяки нейронним мережам комп’ютери сьогодні здатні ідентифікувати об’єкти, визначати їхні межі, розпізнавати сцени, виявляти аномалії та навіть генерувати нові візуальні дані.

Проте на практиці зображення часто містять спотворення: шум, артефакти, розмитість або інші завади. Ці спотворення можуть бути виявлені наприклад у процесі знімання, передавання або обробки фото та відео, через що вони зазнають спотворень. Шум є однією з ключових проблем, що ускладнює автоматичне розпізнавання об’єктів. Його виникнення може бути спричинене поганим освітленням, рухом об’єкта, низькою якістю сенсорів або технічними обмеженнями камер. Тому створення моделей, здатних ефективно працювати у зашумлених умовах, є важливою задачею сучасного машинного навчання.

Завдання з обробки таких зображень має особливе значення у сфері машинного навчання, адже якість вхідних даних безпосередньо впливає на точність моделі. Традиційні фільтри, такі як медіанний або гауссівський, мають обмежену ефективність і часто призводять до втрати важливих деталей. Тому сучасні дослідження зосереджуються на використанні глибоких нейронних мереж, здатних самостійно навчатися виділяти ознаки й відновлювати структуру зображення навіть при значному рівні шуму.

Отже, проблема покращення обробки зашумлених зображень є актуальною як з теоретичної, так і з практичної точки зору. Вона охоплює питання фільтрації, реконструкції, сегментації та класифікації даних у складних умовах.

Метою кваліфікаційної роботи є підвищення точності відновлення та розпізнавання об'єктів на зашумлених зображеннях шляхом дослідження та застосування сучасних методів обробки, включаючи традиційні, інтелектуальні та гібридні підходи.

Об'єкт дослідження – процес розпізнавання цифрових зображень, які містять різні типи шумів, що спотворюють візуальну інформацію та ускладнюють автоматичне розпізнавання об'єктів.

Предмет дослідження – методи та алгоритми обробки зашумлених зображень, включаючи традиційні фільтри, глибокі моделі та гібридні підходи.

Завданням роботи є аналіз сучасних методів машинного навчання, які використовуються для обробки зашумлених зображень, порівняння їх ефективності та визначення найперспективніших напрямів подальшого розвитку.

Задачі дослідження:

1. Провести аналіз теоретичних основ шумів у цифрових зображеннях та їх впливу на процес класифікації.
2. Дослідити сучасні методи фільтрації та очищення зображень, у тому числі традиційні та гібридні підходи.
3. Розібрати та побудувати нейромережеві моделі класифікації зашумлених зображень.
4. Реалізувати процес навчання моделей та підготувати набір зашумлених даних для експериментального аналізу.
5. Провести оцінювання продуктивності кожної моделі на різних типах шумів та сформулювати порівняльний аналіз результатів.
6. Визначити переваги, недоліки та практичну придатність кожного з підходів у задачі класифікації зашумлених зображень.

РОЗДІЛ 1

АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ ТА ОГЛЯД ІНСУЮЧИХ МЕТОДІВ РОЗПІЗНАВАННЯ ОБ'ЄКТІВ НА ЗАШУМЛЕНИХ ЗОБРАЖЕННЯХ

1.1 Загальне значення шуму та зашумлених зображень

Шум [14] – це небажані випадкові або систематичні зміни інтенсивності пікселів зображення, які не містять корисної інформації про об'єкт, але спотворюють його візуальне відображення.

Шум може виникати на будь-якому етапі формування або передавання зображення: під час зйомки, цифрової обробки, збереження чи передачі даних. Основними причинами шуму є [18]:

- недосконалість сенсорів цифрових камер (електронні коливання в матриці, термічний шум);
- недостатнє або надмірне освітлення сцени;
- рух об'єкта або камери під час зйомки;
- вплив зовнішніх електромагнітних перешкод;
- похибки під час кодування, стискання або передачі зображення.

Зашумлені зображення [18] – це зображення, у яких присутні сторонні елементи або завади, спричинені додаванням шуму різної природи. Їх можна представити формулою [14]:

$$I(x, y) = O(x, y) + N(x, y), \quad (1.1)$$

де $I(x, y)$ – зашумлене зображення, $O(x, y)$ – вихідне або чисте зображення, $N(x, y)$ – шумовий сигнал.

Наявність шумів на зображеннях призводить до їх спотворення: втрати дрібних деталей, зниження контрастності, зміни кольорової гами та розмиття контурів. Це ускладнює подальшу обробку, аналіз і розпізнавання об'єктів, оскільки алгоритми комп'ютерного зору можуть помилково інтерпретувати зашумлені ділянки як частину об'єкта або, навпаки, втрачати важливу інформацію.

Окрім шумів, що виникають безпосередньо під час зйомки, існують також спотворення, які з'являються на етапах збереження, обробки або передавання зображення. До них належать:

- помилки передавання даних у комп'ютерних або мережевих системах, що можуть спричиняти втрату окремих бітів або блоків інформації;
- артефакти стиснення, які з'являються через використання алгоритмів з втратами (наприклад, JPEG), коли частина даних безповоротно відкидається;
- пошкодження під час запису або зчитування з носіїв інформації внаслідок збоїв файлової системи або фізичних дефектів пристрою;
- спотворення при конвертації форматів, коли багаторазове перекодування призводить до деградації якості;
- перешкоди при передачі аналогового сигналу, які можуть виникати через електромагнітні впливи або низьку якість з'єднань.

1.2. Види шумів на зображеннях

У процесі формування або передавання зображень можуть виникати шуми різної природи. Кожен тип шуму має свої статистичні властивості, характер впливу на пікселі та способи усунення. Розглянемо найпоширеніші види шумів, що зустрічаються в цифрових зображеннях.

Гауссівський шум (Gaussian noise) [9, 14] описується нормальним розподілом і є одним із найпоширеніших у цифровій фотографії. Він виникає внаслідок електронних коливань сенсорів, підсилювачів або термічних процесів у матриці камери.

Гауссівський шум можна представити як випадкову величину $N(x,y)$, що додається до значень пікселів вихідного зображення $O(x,y)$:

$$I(x,y) = O(x,y) + N(x,y), \quad N(x,y) \sim N(\mu, \sigma^2), \quad (1.2)$$

де μ – середнє значення шуму, σ – стандартне відхилення, яке характеризує його інтенсивність.

На практиці гауссівський шум проявляється як слабка зернистість або “пелена”, що рівномірно покриває зображення (рис. 1.1). Цей тип шуму має критичний вплив на обробку зображень, оскільки навіть невеликі значення σ

можуть знижувати точність алгоритмів розпізнавання контурів і дрібних деталей.

Приклади виникнення:

1. Нічна фотографія з високою чутливістю ISO.
2. Медичні знімки при слабкому сигналі (рентген, МРТ).
3. Супутникові зображення при віддалених умовах освітлення.

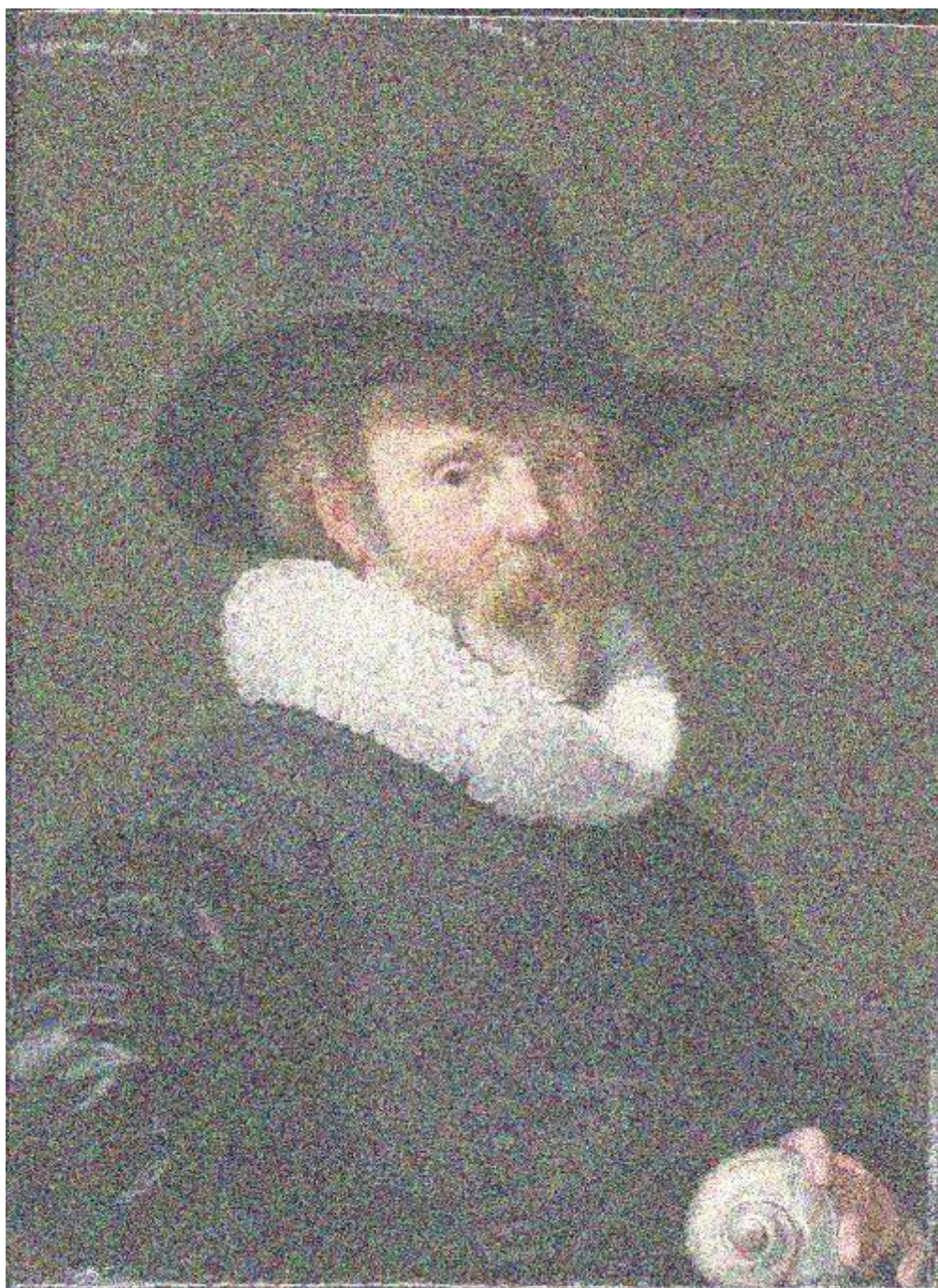


Рисунок 1.1 – Приклад гауссівського шуму на зображенні

Імпульсний шум (Salt-and-Pepper noise) [9] або “сіль і перець” шум виникає у вигляді випадкових білих і чорних пікселів, що різко контрастують із фоном (рис. 1.2). Основними причинами його появи є збої при передаванні даних, помилки в сенсорах або пошкодження під час збереження. Цей тип шуму є особливо руйнівним для методів розпізнавання контурів і дрібних деталей, оскільки поодинокі спотворення можуть сприйматися як об’єкти.

Імпульсний шум можна записати наступним чином:

$$I(x, y) = \begin{cases} 0, & \text{з ймовірністю } p/2, \\ 255, & \text{з ймовірністю } p/2, \\ O(x, y), & \text{з ймовірністю } 1 - p, \end{cases} \quad (1.3)$$

де $I(x, y)$ – значення пікселя у зашумленному зображенні, $O(x, y)$ – значення чистого пікселя, p – ймовірність появи імпульсу.

Така математична модель відображає несправність окремих пікселів, де значна частина зображення залишається незмінною, а лише окремі пікселі на ньому стають білими або чорними.



Рисунок 1.2 – Приклад імпульсивного шуму на зображеннях

Пуассонівський шум (Poisson noise, або фотонний) [10] виникає через статистичні флуктуації кількості фотонів, що потрапляють на сенсор під час зйомки (рис. 1.3). Такий шум притаманний зображенням, отриманим при слабкому освітленні або короткій витримці. Його інтенсивність залежить від яскравості пікселів: у світлих ділянках шум помітний менше, а в темних – сильніше.

Математичного його можна записати наступним чином [14]:

$$I(x, y) = \text{Poisson}(O(x, y)), \quad (1.4)$$

де $I(x, y)$ – значення пікселя у зашумленому зображенні, $O(x, y)$ – середнє очікуване значення пікселя інтенсивність фотонів, $\text{Poisson}(\lambda)$ – випадкова величина з розподілом Пуассона з параметром $\lambda = O(x, y)$.

Імовірність того, що піксель отримає значення k (кількість фотонів) описується формулою Пуассона:

$$P(k) = \frac{\lambda^k e^{-\lambda}}{k!}, \quad (1.5)$$

де k – випадкова величина, що показує кількість подій за фіксований проміжок часу або простору, λ – математичне очікування випадкової величини (середня кількість подій за фіксований проміжок часу).

У випадку з зображеннями, цю формулу можна представити у наступному вигляді:

$$P(I(x, y) = k) = \frac{O(x, y)^k * e^{-O(x, y)}}{k!}, \quad (1.6)$$

де $I(x, y)$ – випадкова величина, що показує кількість фотонів, зафіксованих сенсором у пікселі (x, y) , k – конкретне число фотонів, яке спостерігається, $O(x, y)$ – середня кількість фотонів (інтенсивність чистого пікселя).



Рисунок 1.3 – Приклад пуассонівського шуму на зображеннях

Спекл-шум (Speckle noise) [11] характерний для зображень, отриманих за допомогою когерентного випромінювання – наприклад, у лазерній або радарній зйомці (рис. 1.4). На відміну від гаусового чи імпульсного шуму, спекл-шум має мультиплікативний характер, тобто його інтенсивність пропорційна інтенсивності самого сигналу.

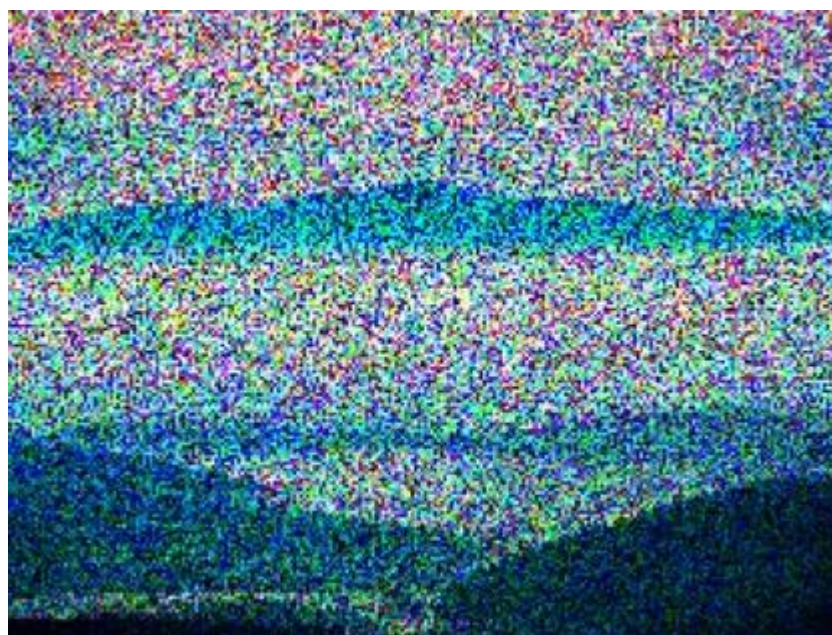


Рисунок 1.4 – Приклад Спекл-шуму на зображеннях

Математично спекл-шум можна описати як:

$$I(x, y) = O(x, y) + O(x, y) * N(x, y), \quad (1.7)$$

де $I(x, y)$ – зашумлене зображення, $O(x, y)$ – чисте зображення, $N(x, y)$ – випадкова величина, що описує шум.

Квантовий (або дискретизаційний) шум з'являється під час перетворення аналогового сигналу в цифровий через обмежену роздільну здатність АЦП (аналогово-цифрового перетворювача). Його ефект – спотворення кольорів і зниження плавності градієнтів.

1.3. Класифікація методів обробки зашумлених зображень.

Обробка зашумлених зображень спрямована на відновлення їхньої якості шляхом приглушення шумів без втрати важливої візуальної інформації. Існує велика кількість методів фільтрації та відновлення, які можна класифікувати за різними ознаками: за принципом дії, за типом використовуваних моделей, за рівнем адаптивності та за способом реалізації.

Умовно всі методи можна поділити на традиційні (детерміновані) та сучасні (інтелектуальні).

1.3.1. Традиційні методи обробки

Традиційні методи базуються на математичному аналізі зображення, використанні статистичних моделей і просторових залежностей між пікселями. Їх перевагою є простота реалізації, невисокі обчислювальні витрати та передбачуваний результат.

Лінійні фільтри [8] передбачають згладжування зображення за допомогою усереднення інтенсивностей сусідніх пікселів. Найбільш поширеними є середньозважені фільтри, фільтр Гаусса та фільтр Вінера. Вони ефективно зменшують гауссівський шум, але водночас можуть розмивати краї та дрібні деталі.

Нелінійні фільтри [8]. Найвідомішим представником є медіанний фільтр, який замінює значення пікселя на медіану значень у його локальному

оточенні. Такий підхід ефективний проти імпульсного (“сіть і перець”) шуму, оскільки зберігає контури об’єктів і не створює розмиття.

Адаптивні методи фільтрації [8] враховують локальні характеристики зображення, наприклад, різкість або контрастність у межах невеликої області. Фільтр Вінера є прикладом адаптивного підходу, який мінімізує середньоквадратичну помилку між відновленим і оригінальним сигналом.

Методи у частотній області ґрунтуються на перетвореннях Фур’є або Вейвлет-перетвореннях. Вони дозволяють виділяти й ослаблювати високочастотні компоненти, що відповідають шуму, зберігаючи при цьому основні структури зображення.

1.3.2. Сучасні інтелектуальні методи

З розвитком штучного інтелекту з’явилися нові підходи до відновлення зашумлених зображень, які базуються на машинному навчанні та глибоких нейронних мережах.

Такі методи не потребують точного аналітичного опису шуму, а навчаються самостійно визначати закономірності у даних.

Нейронні мережі [1, 4, 6]. Мережі типу Convolutional Neural Networks (CNN) здатні виявляти складні просторові залежності між пікселями та ефективно виділяти структуру об’єктів навіть за високого рівня шуму. Відомими прикладами є архітектури DnCNN, U-Net, FFDNet, які показують високу якість очищення без розмиття деталей.

Автоенкодери [3, 13] – це нейронні моделі, що навчаються стискати та відновлювати зображення. Denoising Autoencoder (DAE) – спеціальний тип автоенкодера, який тренується на парах “зашумлене–чисте” і вчиться самостійно усувати шум при реконструкції.

Глибокі генеративні моделі [1]. Моделі типу GAN (Generative Adversarial Networks) використовують два нейронні блоки – генератор і дискримінатор – для створення очищених зображень, максимально подібних до оригінальних. Такі підходи забезпечують природне відновлення текстур і кольорів, проте вимагають великих обсягів даних для навчання.

Методи з урахуванням попередніх знань (prior-based). Деякі сучасні алгоритми (наприклад, BM3D, NLM) використовують апріорні знання про структуру природних зображень – подібність локальних фрагментів, повторюваність текстур тощо.

У таблиці 1.1 узагальнено переваги, недоліки та особливості різних традиційних і новітніх методів.

Таблиця 1.1

Порівняння методів обробки зашумлених зображень

Метод	Переваги	Недоліки	Ефективність для типів шуму	Обчислювальні ресурси
Лінійні фільтри	Простота реалізації, висока швидкість	Розмивають контури та дрібні деталі	Гауссівський шум, слабкий рівень шуму	Низькі
Медіанні фільтри	Ефективні проти імпульсного шуму, зберігають контури	Менш ефективні для гаусового шуму	Імпульсний шум	Низькі
Адаптивні фільтри	Підлаштовуються під локальні характеристики зображення, ефективні для мультиплікативного шуму	Складніші у реалізації, чутливі до параметрів	Спекл-шум, Пуассонівський шум	Середні
Вейвлет-фільтри	Зберігають структуру зображення, ефективні при низьких рівнях шуму	Складність налаштування, потребують обчислень	Гауссівський, Пуассонівський, спекл	Середні
Нейронні мережі (CNN)	Висока точність, здатність навчатися на конкретних типах шуму	Потребують навчальної бази та ресурсів	Всі типи шуму, особливо змішані	Високі
Автоенкодери (DAE)	Навчання на парах “зашумлене–чисте”, добре відновлюють текстури	Потребують великих обсягів навчальних даних	Гауссівський, імпульсний, змішані	Високі
Генеративні моделі (GAN)	Природне відновлення текстур і кольорів, відновлюють деталі	Складність тренування, потребують	Складні змішані шуми	Дуже високі

		потужних ресурсів		
Методи з урахуванням апріорних знань (BM3D, NLM)	Використовують повторюваність локальних фрагментів, висока якість відновлення	Складність реалізації, потребують пам'яті	Гауссівський, спекл, Пуассонівський	Середні-високі

Згідно з таблицею 1.1, можна виділити чіткі тенденції в ефективності різних методів обробки зашумлених зображень. Традиційні методи відзначаються високою швидкістю роботи та простотою реалізації, що робить їх оптимальними для задач, де переважно присутній шум низької або середньої інтенсивності, та коли необхідно обмежити обчислювальні ресурси. Лінійні фільтри, наприклад, ефективні для гауссівського шуму, однак їхнє застосування часто призводить до розмиття контурів і втрати дрібних деталей, що може бути критичним для задач розпізнавання об'єктів. Медіанні та адаптивні фільтри частково компенсують ці недоліки, дозволяючи зберегти контури та текстури, але їхня ефективність залежить від правильного налаштування параметрів.

На сьогоднішній день використання традиційних методів залишається актуальним у тих випадках, коли необхідно забезпечити швидку обробку зображень у реальному часі або коли шум має відносно просту природу. Наприклад, у промислових системах контролю якості або у медичних пристроях із обмеженими обчислювальними ресурсами традиційні фільтри можуть бути єдиним доступним рішенням.

Сучасні інтелектуальні методи, навпаки, демонструють значно вищу ефективність при обробці складних шумів, змішаних типів шумових спотворень та високому рівні шуму. Нейронні мережі та автоенкодери здатні виявляти складні просторові та текстурні закономірності, що дозволяє відновлювати деталі та структури зображення, які втрачаються при застосуванні традиційних методів. Генеративні моделі (GAN) забезпечують природне відновлення текстур і кольорів, що особливо важливо для задач комп'ютерного зору, де візуальна достовірність грає ключову роль.

1.3.3. Гібридні підходи обробки зображень

Окрім стандартних та інтелектуальних методів обробки зашумлених зображень, існують гібридні підходи, які поєднують переваги обох класів алгоритмів. Основна ідея таких методів полягає в тому, щоб спершу застосувати прості та швидкі традиційні фільтри для попереднього зменшення шуму, а потім використовувати сучасні інтелектуальні моделі для відновлення дрібних деталей і текстур [15].

Гібридні підходи дозволяють досягати балансу між швидкістю обробки та якістю відновлення зображень, особливо у випадках складних шумових спотворень або змішаних типів шуму.

Структурно гібридні методи виглядають наступним чином:

1. *Попередня фільтрація.* На цьому етапі використовуються традиційні методи, такі як медіанні або адаптивні фільтри, для зниження рівня шуму та збереження основних контурів об'єктів.

2. *Інтелектуальне відновлення.* Після попередньої обробки застосовуються нейронні мережі, автоенкодери або генеративні моделі для відновлення дрібних деталей, текстур та кольорових відтінків.

3. *Комбінування результатів.* Результати різних етапів можуть поєднуватися, наприклад, шляхом зваженого усереднення або подачі виходу традиційного методу на вхід нейронної мережі.

Гібридні методи показують високу якість відновлення зображень при різних типах шуму, дозволяють зберігати контури та текстури навіть за високого рівня шуму, а також зменшують навантаження на складні нейронні моделі завдяки попередньому згладжуванню шуму традиційними фільтрами. Однак їхня реалізація є складнішою та потребує додаткових обчислювальних ресурсів порівняно з використанням лише традиційних методів.

Приклади гібридних методів:

- BM3D + CNN: BM3D зменшує гауссівський шум, після чого CNN відновлює дрібні деталі;

- Медіанний фільтр + DAE: медіанний фільтр усуває імпульсний шум, а Denoising Autoencoder покращує загальну якість зображення;
- Вейвлет-фільтрація + GAN: вейвлет-фільтри знижують високочастотний шум, а GAN відновлює текстури і кольори.

Висновки до розділу 1

У першому розділі проведено детальний теоретичний аналіз предметної області обробки та розпізнавання об'єктів на зашумлених зображеннях. Розглянуто природу шумів, їхні види, математичні моделі та вплив на візуальне відображення об'єктів. Виокремлено основні причини появи шумів під час зйомки, збереження, обробки та передачі зображень, а також їхній вплив на точність алгоритмів комп'ютерного зору.

Досліджено класифікацію методів обробки зашумлених зображень, включаючи традиційні (лінійні, медіанні, адаптивні, частотні), сучасні інтелектуальні (нейронні мережі, автоенкодера, генеративні моделі, методи з урахуванням апіорних знань) а також поєднання їх у гібридні методи. Показано їхні переваги та обмеження, а також ефективність у боротьбі з різними типами шуму.

РОЗДІЛ 2.

МЕТОДИ ТА МОДЕЛІ КЛАСИФІКАЦІЇ ОБ'ЄКТІВ НА ЗАШУМЛЕНИХ ЗОБРАЖЕННЯХ

2.1. Постановка задачі класифікації зображень

Задача класифікації зображень [1, 16] полягає у визначенні класу одного або кількох об'єктів, які присутні на зображенні. У класичному формулюванні модель класифікації отримує на вхід зображення I , обробляє його та повертає мітку класу, до якого належить розпізнаний об'єкт. Загальну модель класифікації можна представити формулою:

$$f(I) = y, \quad (2.1)$$

де

I – вхідне зображення,

f – модель класифікації,

y – відповідний клас.

На відміну від класичної задачі класифікації, де модель працює з чистими, а також добре структурованими зображеннями, наявність шуму суттєво може впливати та ускладнювати процес розпізнавання об'єктів на цих зображеннях. Спотворені шумом зображення мають менший обсяг інформації для їх ідентифікації. Шум змінює як низькорівневі характеристики (краї об'єкту, градієнт, текстури), так і високорівневі семантичні ознаки, що формуються моделями класифікації, наприклад, у глибинних шарах нейронної мережі. Як результат, модель отримує на обробку дані, які відрізняються від тих, на яких вона навчалась, що може призводити до падіння точності, як на кілька відсотків, так і у декілька разів, до помилкових класифікацій або в цілому до нестабільної роботи моделі [8, 14].

Шум впливає на роботу класифікації через такі фактори:

1) збільшення варіативності входів. Шум створює нові, не властиві об'єкту текстури, точки, плями, які не присутні у тренувальних даних;

2) спотворення контурів. Багато моделей залежать від контурів та країв, і шум може руйнувати їх або створювати хибні контури.

3) зміна статистики розподілу даних. Чисті дані та зашумлені дані мають різні розподіли ознак, що спричиняє падіння точності моделі.

Таким чином, *задача класифікації на зашумлених зображеннях* формулюється як створення моделі f , здатної точно визначати клас об'єкта, незважаючи на шумові спотворення або після попереднього відновлення зображення. Відповідно до наведеного визначення розглядають два підходи до побудови моделей класифікації зашумлених зображень.

1. Пряме навчання моделей класифікації на зашумлених зображеннях.
2. Попереднє очищення шуму за допомогою допоміжних інструментів з подальшою класифікацією об'єктів на даних зображеннях.

У розділі далі розглянемо існуючі моделі, а також підходи, які дозволяють досягти високої точності в умовах шуму.

2.2. Класичні методи класифікації зображень та їх обмеження при роботі з шумом

До появи глибоких нейронних мереж задачі класифікації зображень розв'язувалися за допомогою класичних алгоритмів машинного навчання. Вони передбачали використання вручну створених ознак (feature engineering), таких як текстурні оператори, гістограми градієнтів (HOG), контури, локальні бінарні патерни (LBP) та інші дескриптори. Однак у контексті зашумлених зображень такі методи виявляють суттєві обмеження.

Логістична регресія [1] є одним із базових методів класифікації, який моделює ймовірність належності об'єкта до певного класу за допомогою сигмоїдної або softmax-функції

Одна з основних переваг даної моделі полягає у простоті реалізації та високій швидкості роботи. Однак у контексті зображень цей алгоритм:

- вимагає попереднього виділення ознак;
- погано працює в умовах невизначеності та шуму;

- не здатний моделювати складні нелінійні залежності.

Попри простоту логістичної регресії, недоліки роблять даний метод нерелевантним у задачах класифікації зображень, особливо з наявним шумом.

Метод опорних векторів (Support Vectors Machines, SVM) [1] є одним з найефективніших класичних алгоритмів, завдяки здатності будувати нелінійні межі рішень з використанням ядерних функцій (RBF, polynomial kernels). Його переваги:

- хороша робота з даними середньої розмірності;
- стійкість до частини шумових спотворень;
- можливість оптимізувати межу класифікації за допомогою параметра регуляризації.

Однак у задачах класифікації зображень SVM має такі недоліки:

- сповільнення роботи на великих наборах даних (CIFAR-10, CIFAR-100);
- залежність від вручну розроблених дескрипторів;
- чутливість до високої кількості шуму або випадкових артефактів, що змінюють структуру ознак.

Таким чином, SVM може бути ефективним лише у поєднанні з добре підготовленими ознаками та фільтрами попередньої обробки.

Ансамблеві методи [1] є більш стійкими до шуму за рахунок побудови множини моделей та агрегування їхніх результатів. Найбільш використовуваними методами є випадкові ліси (Random Forest) та бустінг дерев (Gradient Boosting).

Random Forest базується на побудові великої кількості рішучих дерев, є менш чутливий до випадкових спотворень у даних і працює краще за SVM у високовимірних просторах.

Gradient Boosting навчає слабкі моделі таким чином, щоб вони компенсували помилки попередніх, при цьому має високу точність на структурованих даних.

Недоліки ансамблевих методів у задачах класифікації зображень:

- погана робота на сирих зображеннях без глибоких ознак;
- потреба в екстракції ознак типу HOG, LBP, SIFT;
- недостатня потужність для розпізнавання складних шумових спотворень.

Обмеження класичних методів для зашумлених зображень. Таким чином для класичних методів машинного навчання можна визначити наступні загальні недоліки:

- 1) залежність від обробки ознак вручну. Шум легко спотворює локальні дескриптори, роблячи класифікацію ненадійною;
- 2) погана масштабованість. Методи швидко втрачають ефективність при переході до високої роздільної здатності або великих наборів даних;
- 3) недостатня виразна здатність. Моделі не можуть автоматично вчитися складним шаблонам, як це роблять CNN.

Тому в сучасних системах класифікації зображень, особливо за наявності шумів, ключову роль відіграють глибокі нейронні мережі, а також спеціальні моделі для попереднього очищення зображень перед їх класифікацією.

2.3. Глибокі моделі класифікації та обробки зашумлених зображень

Глибокі нейронні мережі стали основним інструментом для розпізнавання об'єктів на зображеннях завдяки здатності автоматично виділяти складні просторові та текстурні ознаки. На відміну від класичних алгоритмів, вони не потребують ручного формування дескрипторів, а навчають багаторівневі представлення даних безпосередньо з навчальної вибірки. [1]

Особливо ефективними для задач класифікації та попереднього очищення шуму є згорткові нейронні мережі (CNN), спеціалізовані моделі для денойзингу (DnCNN) та глибокі резидуальні мережі (ResNet). Далі розглянемо їх детально.

2.3.1. Згорткові нейронні мережі

Згорткові нейронні мережі (Convolutional Neural Networks, CNN) [6, 12] є фундаментальною архітектурою для задач комп'ютерного зору. Їхня ключова особливість – використання згорткових шарів, які виділяють локальні ознаки на різних рівнях абстракції. Вони дозволяють моделі ефективно обробляти зображення завдяки просторовій ієрархії ознак: від простих країв до складних текстур та форм об'єктів.

CNN складається з трьох основних структурних елементів (рис. 2.1): згортковий, агрегувальний і повнозв'язний шари.

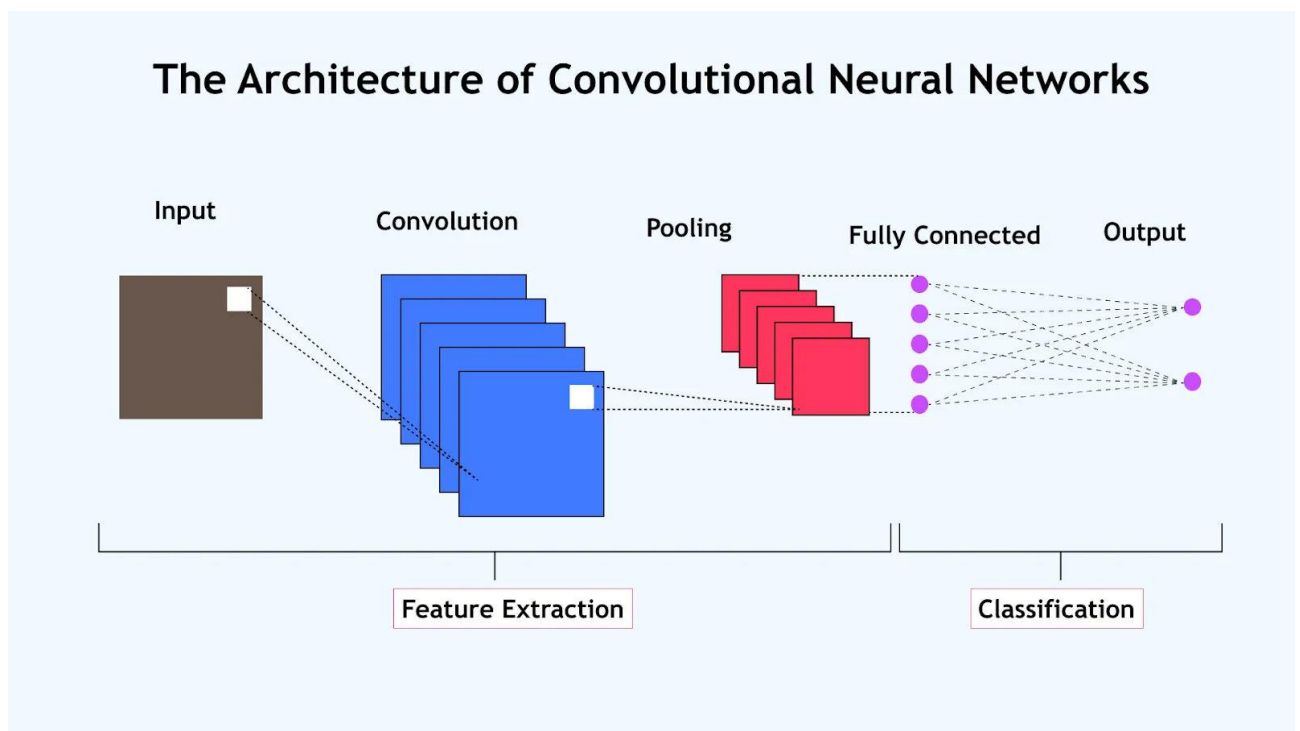


Рисунок 2.1 – Архітектура CNN

Згорткові шари (Convolutional layers) – це головний компонент CNN. Згортковий шар застосовує набір фільтрів (ядра згортки або фільтри), які “ковзають” по зображенню та виділяють характерні патерни: краї та кути (у перших шарах), текстури та форми (у середніх шарах), високорівневі семантичні ознаки (у глибоких шарах).

Згорткові шари скорочують кількість параметрів завдяки локальним зв'язкам, що підвищує ефективність навчання.

Агрегувальні шари (Pooling layers) використовуються для зменшення просторових розмірів карти ознак. Подібно до згорткового шару, pooling проходить фільтром (ядром) по вхідним даним, але на відміну від згорткового ядра, фільтр pooling не має вагових коефіцієнтів. Найчастіше застосовуються такі типи pooling:

- MaxPooling – вибір максимального значення у локальному вікні;
- AveragePooling – усереднення локальних значень;
- SumPooling – повертає суму всіх значень карти ознак.

Шари Pooling роблять модель менш чутливою до зсувів об'єкта, більш стійкою до шумів низького рівня та менш обчислювально затратною.

Наявність *функцій активації* (Activation layers) після згортки є критично важливим, оскільки вводять у нейронну мережу нелінійність і надають можливість моделювати складні ознаки. Використовують:

- ReLU – прискорює навчання і працює стабільніше за сигмоїду та тангенс;
- LeakyReLU – запобігає «вмиранню» нейронів;
- ELU, GELU – використовуються у деяких сучасних архітектурах.

Повнозв'язні шари (Fully Connected layers) розташовані у заключній частині моделі. Вони здійснюють класифікацію, об'єднуючи інформацію зі всіх попередніх згорткових шарів у вектор ознак. Вихідний повнозв'язний шар зазвичай використовує функцію активації *softmax*, яка перетворює вихід моделі у ймовірності для кожного класу.

Однією з найпоширеніших проблем, які виникають при розробці CNN, є перенавчання. Для його зменшення застосовують додаткові шари.

Нормалізація (Batch Normalization) є стандартом практично в усіх CNN. Цей шар нормалізує активації всередині шару та забезпечує:

- стабільніше та швидше навчання,

- зменшення чутливості до випадкового шуму,
- уникнення переобчислення масштабів активацій.

Dropout є методом регуляризації. Він випадковим чином вимикає частину нейронів під час навчання. Це є особливо важливим при роботі із зашумленими даними. Dropout зменшує перенавчання, підвищує стійкість мережі до шуму, робить модель більш узагальненою.

2.3.2. Модель DnCNN

Модель DnCNN (Denoising Convolutional Neural Network) [4] є однією з найефективніших згорткових нейронних мереж, спеціально розроблених для задачі очищення зашумлених зображень. На відміну від класичних фільтрів (медіанного, гауссівського, Вінера) та традиційних підходів у частотній області, DnCNN використовує глибоку структуру з нормалізацією та залишковим навчанням для точного відновлення зображення та пригнічення шуму різного типу.

Вона має архітектуру, яка складається з наступних шарів (рис. 2.2):

1. *Первинний згортковий шар* виконує базове виділення локальних ознак: Операція згортки (Conv), функція активації (ReLU). Цей шар не містить нормалізації, щоб зберегти вихідні статистичні характеристики зображення.

2. *Основна частина* складається з 12–18 однакових блоків, кожен із яких містить згортковий шар, шар нормалізації Batch Normalization та нелінійну функцію активації ReLU. Саме ці блоки відповідають за глибоке вилучення ознак та детальний аналіз структур шуму. Завдяки багатошаровому послідовному застосуванню згорток модель здатна виявляти як низькорівневі патерни шуму, так і його складні статистичні закономірності, що істотно підвищує точність реконструкції зображення.

3. *Вихідний згортковий* (фінальний) шар застосовує згортку розміром 3×3 з одним вихідним каналом (для зображень у градаціях сірого) або трьома каналами (для кольорових зображень у форматі RGB). На цьому етапі не використовується нелінійна функція активації, оскільки завданням є

отримання точного числового представлення шуму без додаткової трансформації.

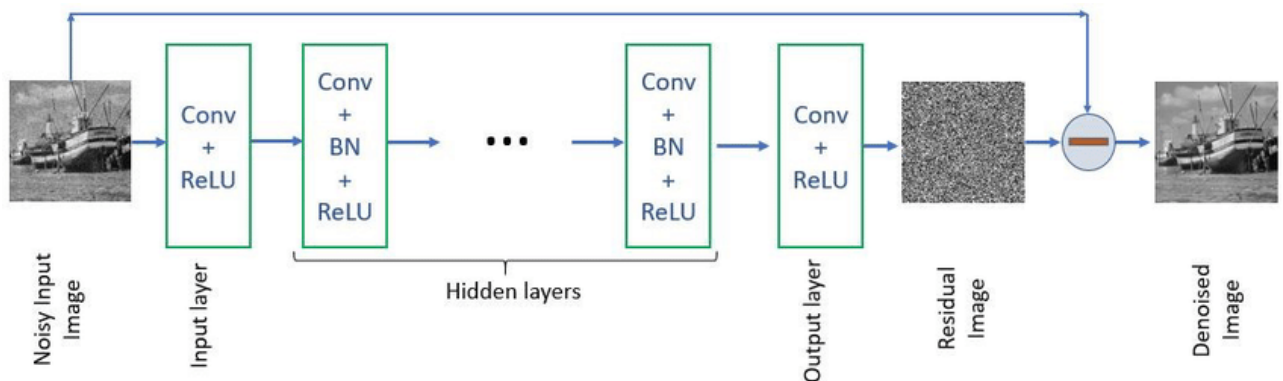


Рисунок 2.2 – Архітектура DnCNN

Вихідний шар формує оцінку шумової складової N , яку модель вважає присутньою на вхідному зашумленому зображенні. Відсутність активації дозволяє повернути значення шуму у широкому діапазоні, не обмежуючи їх позитивними або нормованими величинами. Завдяки цьому підходу мережа може точно відтворювати як позитивні, так і негативні коливання інтенсивності, що характерно для адитивних шумів. Отримана оцінка шуму далі віднімається від вхідного зображення, що дозволяє реконструювати максимально близьке до оригіналу, очищене від шуму зображення.

Принцип роботи DnCNN. Ключовою ідеєю DnCNN є передбачення саме шуму, а не очищеного зображення

$$N = f(I_{noisy}). \quad (2.2)$$

Після чого реконструйоване чисте зображення отримується як:

$$I_{clear} = I_{noisy} - N. \quad (2.3)$$

Перевагами такого підходу є те, що:

- шум зазвичай є простішим для моделювання, ніж чисте зображення;
- залишкове навчання прискорює збіжність моделі;
- підвищується точність реконструкції;
- глибока архітектура не страждає від затухання градієнтів.

DnCNN може працювати з різними *типами шумів*, такими як:

- гауссівський шум (найпоширеніший випадок);
- пуассонівський шум;
- спекл-шум (мультиплікативний);
- імпульсний шум (“сіль і перець”);
- змішані типи шуму.

Можливості моделі до адаптування до різних розподілів шуму пояснюється її здатністю вивчати статистичні закономірності шумових патернів у глибоких шарах. На відміну від класичних фільтрів, які мають фіксовані правила обробки, DnCNN навчається розпізнавати та пригнічувати навіть складні та нерівномірні спотворення

Використання DnCNN у гібридних підходах. DnCNN часто використовуються як частина двоетапної системи:

$$DnCNN(I_{noisy}) \rightarrow Classifier(I_{clean}). \quad (2.4)$$

Такі системи показують кращу стійкість до шуму, вищу точність класифікації, а також меншу кількість хибних спрацьовувань. Проте незважаючи на плюси даного підходу, він має і певні обмеження, зокрема уповільнення обробки через додавання етапу очищення, що може бути критичним у системах реального часу, де кожна секунда важлива.

Крім того, надмірна фільтрація може призводити до втрати важливих дрібних ознак, які класифікатор використовує для розпізнавання складних або подібних між собою класів. Якщо DnCNN занадто агресивно пригнічує шум, вона може одночасно згладжувати текстури, структури або локальні патерни, що негативно впливає на точність подальшої класифікації.

Ще одним недоліком є неузгодженість цілей між двома моделями. DnCNN навчається мінімізувати різницю між чистим і зашумленим зображенням, тоді як класифікатор оптимізується під максимальну точність розпізнавання класів. У певних випадках це призводить до того, що відновлене зображення не є оптимальним з точки зору класифікації.

Також варто відзначити, що гібридний підхід потребує окремого навчання двох моделей, що збільшує вимоги до обчислювальних ресурсів і часу тренування. Додатково ускладнюється процес розгортання системи, оскільки необхідно інтегрувати дві нейронні мережі та забезпечити коректну передачу даних між ними.

Попри ці обмеження, методи типу «DnCNN + класифікатор» залишаються одним із найбільш ефективних рішень у ситуаціях, коли рівень шуму значний або коли система повинна працювати зі зображеннями низької якості. Вони забезпечують суттєве покращення продуктивності порівняно з класичними класифікаторами, навченими лише на чистих даних, і є широко застосовуваними у складних практичних сценаріях.

2.3.3. Модель ResNet-18

ResNet-18 (Residual Neural Network, 18 шарів) [2] є однією з найпопулярніших та найефективніших архітектур для задач класифікації зображень, особливо коли йдеться про роботи з даними, що містять шум або різні форми спотворень. Модель належить до сімейства глибоких резидуальних мереж, запропонованих у 2015 році командою Microsoft Research, і стала важливим проривом у побудові дуже глибоких нейронних мереж.

Основна інновація ResNet полягає у використанні skip-конекцій або резидуальних зв'язків, які передають вихід деяких шарів напряму до дальших шарів, оминаючи проміжні.

Завдяки цьому глибина мережі може суттєво збільшуватися без проблем затухання чи вибуху градієнтів. Це дозволяє мережі навчатися значно ефективніше навіть при наявності складних спотворень або шумів.

Формально, резидуальний блок обчислюється за наступною формулою:

$$y = F(x) + x, \quad (2.5)$$

де $F(x)$ – перетворення, виконане згортковими шарами, x – початковий вхід блоку.

Архітектурні особливості ResNet-18. ResNet-18 складається з таких компонентів (рис. 2.3):

1. Початковий етап:
 - а. Великий згортковий шар 7x7.
 - б. Шар підвибірки (MaxPooling 3x3), котрий дозволяє швидко зменшити просторовий розмір карти ознак.

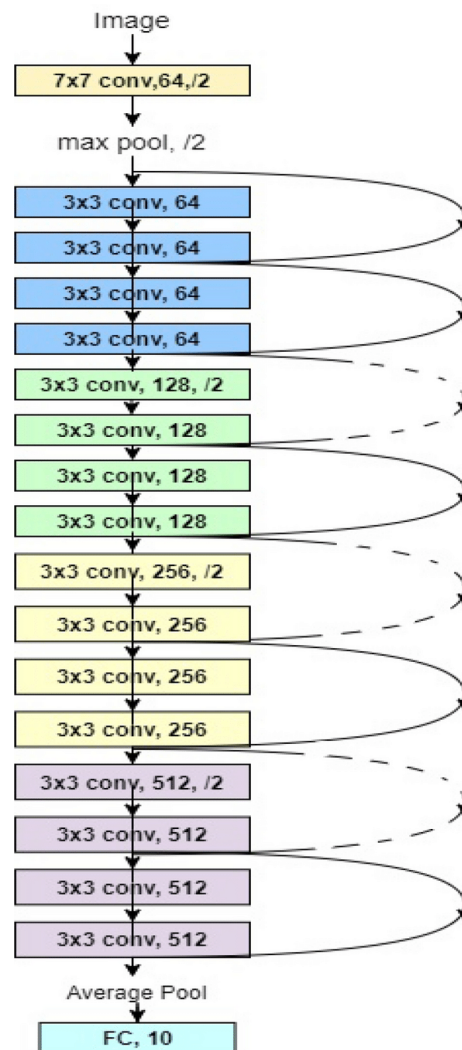


Рисунок 2.3 – Архітектура ResNet-18

2. Чотири групи резидуальних блоків.

Кожна група містить по 2 блоки, у яких виконуються дві згортки 3×3 та додавання пропущеного з'єднання. Кількість каналів зростає в міру поглиблення мережі: $64 \rightarrow 128 \rightarrow 256 \rightarrow 512$.

3. Глобальний усереднювальний пулінг займається перетворенням карти ознак у вектор, незалежно від розміру вхідного зображення.

4. Кінцевий повнозв'язний шар формує ймовірнісну оцінку класів.

Переваги ResNet-18 у контексті класифікації зашумлених зображень. ResNet-18 демонструє високу стійкість до шуму завдяки використанню резидуальних зв'язків, які дозволяють зберігати важливу структурну інформацію про зображення, навіть коли окремі шари піддаються спотворенням. Глибша частина мережі краще виявляє високорівневі семантичні ознаки, тому навіть при зашумленні низькорівневих фільтрів (таких як краї чи текстури) результуючі ознаки залишаються достатньо стабільними для коректної класифікації. Крім того, ResNet-18 характеризується швидким і стабільним процесом навчання порівняно з класичними згортковими мережами аналогічної глибини. Важливо й те, що модель має оптимальний баланс між точністю та швидкістю: вона достатньо компактна для використання у системах реального часу, одночасно забезпечуючи високу якість класифікації навіть у складних умовах.

ResNet-18 у гібридних системах. У поєднанні з моделями відновлення зображення (такими як DnCNN) ResNet-18 може працювати як ефективний класифікатор:

$$I_{clean} = DnCNN(I_{noisy}) \rightarrow ResNet - 18(I_{clean}). \quad (2.6)$$

Таке поєднання дозволяє зменшити вплив шуму на ранні шари моделі, стабілізувати високорівневі ознаки а також підвищити класифікації у задачах із значними шумовими забрудненнями

Недоліки ResNet-18 у задачах з шумом. Попри свої переваги, ResNet-18 не є ідеальною моделлю. Ранні шари мережі залишаються чутливими до сильного шуму, а точність може погіршуватися при обробці нестандартних або непередбачуваних типів шуму. Для ефективного навчання ResNet-18 потрібен достатньо великий обсяг даних, і складні шумові патерни інколи можуть інтерпретуватися мережею як значущі ознаки. У таких випадках більш

ефективним може бути використання гібридних підходів або спеціалізованих моделей, таких як DnCNN чи автоенкодера.

2.4. Автоенкодера та гібридні підходи для роботи із зашумленими зображеннями

Автоенкодера (Autoencoders, AE) [3, 17] – це тип нейронних мереж, призначений для навчання компактного представлення (latent representation) даних. Вони складаються з трьох основних частин (рис. 2.4):

- енкодер (encoder) – займається стисканням вхідного зображення у низькорозмірний вектор ознак;
- латентний простір (latent space) – є результатом роботи енкодера, містить найбільш важливі ознаки, необхідні для відновлення входу;
- декодер (decoder) – займається відновленням зображення з вектору ознак.

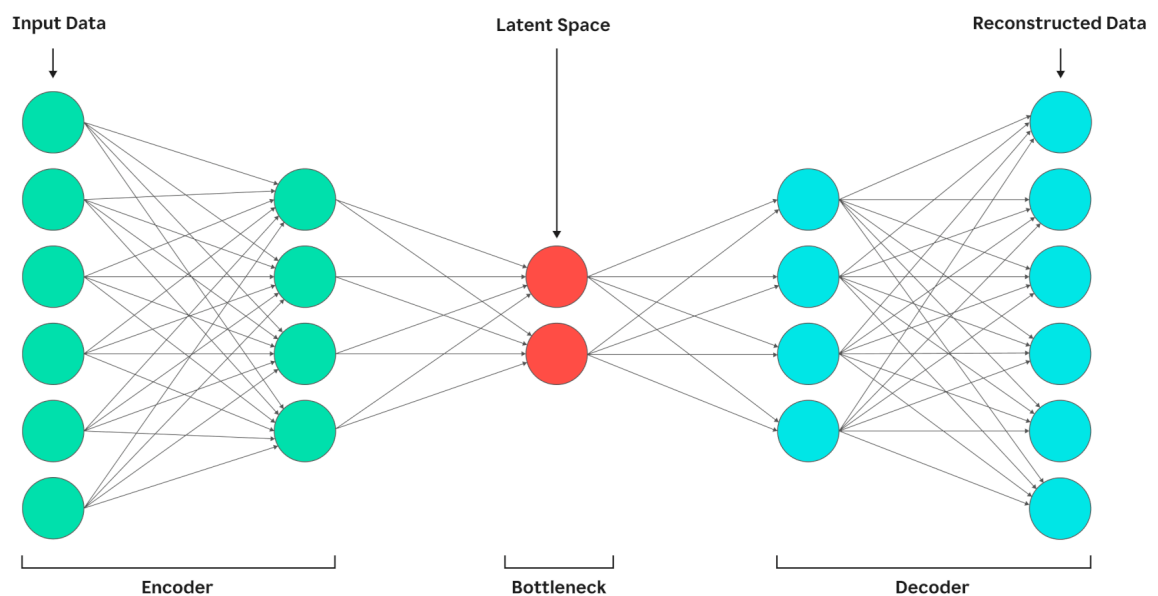


Рисунок 2.4 – Архітектура Autoencoder

У задачах очищення зображень застосовується денойзингові автоенкодері (Denoising Autoencoders, DAE). Основна їх ідея полягає у тому, що мережа навчається відновлювати чисте зображення із зашумленого. Їх можна зобразити наступною формулою

$$I_{clear} = Decoder(Encoder(I_{noisy})). \quad (2.7)$$

Використання автоенкодерів у задачах денойзингу має кілька важливих *переваг*. По-перше, мережа автоматично виділяє ознаки, навчаючись знаходити структури, що характеризують об'єкт, і одночасно відокремлювати шум. По-друге, автоенкодері демонструють гнучкість щодо різних типів шуму, адже навчання на різних шумових розподілах дозволяє моделі адаптуватися до нових спотворень. Нарешті, вони легко інтегруються у гібридні системи, де можуть виконувати роль попереднього етапу очищення перед класифікацією, підвищуючи точність та стабільність роботи основного класифікатора.

Використання автоенкодерів у гібридних підходах. У гібридних підходах автоенкодері зазвичай використовуються як етап попереднього очищення зображень перед передачею їх на класифікатор. Така комбінація дозволяє значно підвищити точність розпізнавання, оскільки автоенкодер усуває шум і стабілізує високорівневі ознаки, навіть якщо низькорівневі фільтри спотворені [15]. При цьому модель здатна працювати з різними типами шуму, включаючи суміші імпульсного та гауссівського, що робить систему більш універсальною. Водночас гібридні підходи мають певні обмеження: додатковий етап денойзингу подовжує час обробки, а надмірна фільтрація може призводити до втрати дрібних деталей, необхідних для розпізнавання схожих класів. Крім того, цілі автоенкодера і класифікатора не завжди збігаються: перший оптимізується на відновлення зображення, другий – на правильність класифікації. Незважаючи на це, поєднання автоенкодера і класифікатора залишається одним із найефективніших рішень для роботи із зашумленими зображеннями, особливо в умовах високого рівня шумового сигналу.

Висновок до розділу 2

У другому розділі було розглянуто методи та моделі класифікації об'єктів на зашумлених зображеннях, а також специфіку роботи з шумом у сучасних системах комп'ютерного зору. Було показано, що класичні методи, такі як логістична регресія, метод опорних векторів та ансамблеві підходи, суттєво обмежені при роботі із зашумленими даними через залежність від ручного виділення ознак, погану масштабованість та недостатню здатність моделювати складні шумові патерни.

На відміну від них, глибокі нейронні мережі, зокрема згорткові мережі (CNN), демонструють високу ефективність завдяки автоматичному виділенню ознак на різних рівнях абстракції. Спеціалізовані моделі для денойзингу, такі як DnCNN, дозволяють точно відновлювати зображення та пригнічувати різні типи шуму, а гібридні системи, що поєднують автоенкодери або DnCNN з класифікаторами (наприклад, ResNet-18), забезпечують значне підвищення точності класифікації та стабільності високорівневих ознак.

ResNet-18, як представник глибоких резидуальних мереж, показує високу стійкість до шуму завдяки використанню резидуальних блоків, які дозволяють зберігати структурну інформацію та ефективно виявляти високорівневі семантичні ознаки. Попри це, ранні шари залишаються чутливими до сильного шуму, точність може погіршуватися при нестандартних типах спотворень, а для ефективного навчання потрібен достатньо великий обсяг даних. У таких випадках більш ефективним є застосування гібридних підходів або спеціалізованих моделей, таких як DnCNN або автоенкодери.

Таким чином, аналіз показав, що для задач класифікації зашумлених зображень ключовими є використання сучасних глибоких архітектур та спеціалізованих методів очищення шуму. Найбільш оптимальні підходи поєднують денойзинг та потужні класифікатори, забезпечуючи баланс між точністю, стійкістю до шуму та обчислювальною ефективністю.

РОЗДІЛ 3.

РОЗРОБКА ТА ДОСЛІДЖЕННЯ МОДЕЛІ КЛАСИФІКАЦІЇ ОБ'ЄКТІВ НА ЗАШУМЛЕНИХ ЗОБРАЖЕННЯХ

3.1. Постановка задачі

Метою практичної частини є розробка та експериментальне дослідження моделей класифікації об'єктів на зашумлених зображеннях. Основним завданням є визначення того, наскільки різні архітектури глибоких нейронних мереж здатні протистояти шумовим спотворенням та зберігати високу точність класифікації в умовах деградації даних.

У рамках дослідження розглядається декілька підходів до класифікації зашумлених зображень. Перший підхід полягає у безпосередньому навчанні класифікаторів на зашумлених даних, що дозволяє оцінити базову стійкість моделей до шумів. Другий підхід передбачає використання окремої моделі для усунення шуму, після чого очищені зображення подаються на вхід класифікатору. Такий гібридний підхід дає можливість оцінити, наскільки попереднє відновлення покращує кінцеву точність.

Для порівняння ефективності використовуються такі моделі: базова згортова нейронна мережа (CNN), глибока резидуальна мережа ResNet-18, а також гібридна система, що поєднує автоенкодер для очищення зображень із CNN-класифікатором. Оцінювання проводиться за стандартними метриками класифікації, а також із урахуванням різних рівнів шуму, що дозволяє визначити залежність між інтенсивністю шумових спотворень та продуктивністю моделей.

3.2. Підготовка даних для навчання моделей

Для дослідження моделей класифікації зашумлених зображень було використано набір даних CIFAR-10 [5], який є одним із найпоширеніших відкритих датасетів для навчання та тестування моделей комп'ютерного зору. CIFAR-10 містить 60 000 зображень розміром 32×32 пікселі, поділених на 10

різних класів. З них 50 000 зображень використовуються для навчання, а 10 000 – для тестування. Усі зображення представлені у форматі RGB, що дозволяє оцінити роботу моделей у багатоканальному просторі.

Перед навчанням моделі дані було нормалізовано шляхом ділення значень пікселів на 255, що переводить їх у діапазон $[0, 1]$ та покращує стабільність навчання нейронних мереж. Також було виконано перетворення міток класів до цілочисельного формату.

Для дослідження впливу шуму на точність класифікації були штучно згенеровані кілька типів шумів, які часто зустрічаються у реальних умовах. Зокрема:

- гауссівський шум з різними значеннями стандартного відхилення ($\sigma = 0.1$ та 0.2). Цей шум моделює випадкові флуктуації інтенсивності пікселів та є одним із найпоширеніших у цифрових сенсорах;
- імпульсний шум типу “сіть і перець” з параметром $\text{amount} = 0.02$. Він проявляється поодинокими білими та чорними пікселями, що особливо ускладнює роботу контурних фільтрів;
- спекл-шум (multiplicative noise) зі стандартним відхиленням 0.1 . Такий шум залежить від інтенсивності пікселя та часто зустрічається в медичній або радарній зйомці;
- пуассонівський шум, який моделює статистичні флуктуації кількості фотонів. Використовується для моделювання шуму при низькій освітленості.

Функції генерації шуму реалізовані програмно та застосовуються до тестової вибірки, що дозволяє оцінити роботу моделі на одних і тих самих зображеннях у різних умовах зашумлення.

3.3. Використані програмні засоби та інструменти

Для реалізації експериментальної частини дослідження було використано набір сучасних інструментів та бібліотек, які забезпечують гнучкість, високу продуктивність обчислень та зручність у роботі з глибинними нейронними мережами.

Основним середовищем розробки виступала мова програмування Python, яка є стандартом де-факто у галузі машинного навчання та комп'ютерного зору. Для реалізації моделей, роботи з даними та проведення експериментів були використані:

1. TensorFlow [20] та Keras [19] бібліотеки застосовувалися для побудови, навчання та оцінки моделей нейронних мереж. Keras забезпечила високорівневий інтерфейс для швидкої розробки моделей CNN, ResNet-18 та автоенкодера, а TensorFlow виступав як бекенд для ефективних чисельних обчислень на CPU та GPU.

2. NumPy [21] бібліотека використовувалася для векторизованих обчислень, нормалізації зображень, роботи з багатовимірними масивами та реалізації функцій додавання шуму: гауссівського, імпульсного, спекл-шуму та пуассонівського.

3. Pandas [22] Застосовувалася для структурованого збереження результатів експериментів, формування таблиць точності та запису підсумкових даних у формат CSV.

4. Scikit-learn [23] Використовувався для розрахунку метрик точності (accuracy score) та аналізу результатів класифікації.

3.4. Побудова моделей класифікації зашумлених зображень

У цьому підрозділі представлено процес побудови та архітектурні особливості моделей, використаних для класифікації зашумлених зображень. Для дослідження було розроблено три різні підходи, що відрізняються рівнем складності та стійкістю до шумових спотворень.

Першим є базовий класифікатор на основі згорткової нейронної мережі (CNN), який дозволяє оцінити, як проста модель реагує на різні рівні шуму та наскільки її продуктивність погіршується у порівнянні з більш складними архітектурами. Другий підхід передбачає використання глибокої резидуальної мережі ResNet-18, відомої високою здатністю до узагальнення та стабільним навчанням завдяки резидуальним зв'язкам. Третім підходом є комбінація

автоенкодера та CNN, де автоенкодер виконує попереднє очищення зображень, а CNN здійснює подальшу класифікацію. Такий гібридний підхід дозволяє оцінити ефективність попереднього усунення шуму перед класифікацією.

3.4.1. Базова згорткова нейронна мережа (CNN)

Першою моделлю, що буде розглядатись для класифікації зашумлених зображень, є базова згорткова нейронна мережа (CNN). Вона була обрана як стартовий підхід, оскільки саме згорткові мережі є фундаментальною архітектурою для задач комп'ютерного зору та широко застосовуються для класифікації зображень різної складності. Вона дозволяє виділяти як низькорівневі, так і високорівневі ознаки, поступово формуючи інформативні представлення об'єктів у процесі навчання.

Архітектурна структура моделі.

1. *Вхідний шар*. Мережа приймає зображення розміром $32 \times 32 \times 3$, де 3 – це кількість каналів (RGB).

На цьому етапі дані не зазнають жодних змін, проте визначення вхідної форми є необхідним для побудови мережі.

2. *Перший згортковий шар* має конфігурацію: `Conv2D(32, (3,3), activation='relu', padding='same')`. Він використовує 32 фільтри з ядрами розміром 3×3 та падингом `same`, що дозволяє зберігати початкову просторову розмірність зображення після операції згортки. Основне завдання цього шару – початкове вилучення базових локальних ознак, таких як краї, кути та орієнтовані лінії. Саме тут формується первинна карта ознак, яка слугує основою для дедалі складніших патернів, що будуть виявлятися глибшими шарами мережі.

Розмір ядра 3×3 обрано не випадково – він є сучасним стандартом, адже забезпечує оптимальний баланс між виразністю моделі та кількістю параметрів. Більші ядра (наприклад, 5×5 чи 7×7) значно збільшують обчислювальні витрати, при цьому не даючи суттєвої переваги в якості. Функція активації ReLU додає моделі здатність до нелінійного відображення,

дозволяючи вивчати складні залежності та покращуючи швидкість збіжності під час навчання.

3. *Перший шар субдискретизації $\text{MaxPooling2D}((2,2))$* зменшує просторову розмірність карти ознак удвічі, вибираючи максимальні значення в межах кожного локального фрагмента. Завдяки цьому модель стає менш чутливою до незначних геометричних змін, таких як невеликі зміщення об'єкта. Крім того, субдискретизація зменшує кількість параметрів і обсяг обчислень у наступних шарах, а також відіграє роль у пригніченні дрібних випадкових флуктуацій – у тому числі й шуму низької інтенсивності.

4. *Другий згортковий шар, визначений як $\text{Conv2D}(64, (3,3), \text{activation}='relu', \text{padding}='same')$* , розширює глибину представлення, збільшуючи кількість фільтрів до 64. Застосування тих самих 3×3 ядер дозволяє шарові ефективно працювати з локальними патернами, однак тепер він уже аналізує не початкові пікселі, а карти ознак, створені попередніми шарами. На цьому рівні мережа виявляє складніші структури, такі як текстури, поєднання контурів та характерні фрагменти об'єктів (наприклад, очі тварин, колеса автомобілів, крила літаків). Збільшення кількості фільтрів – це стандартна практика, оскільки глибші шари повинні відображати більшу різноманітність структур.

5. *Другий шар $\text{MaxPooling2D}((2,2))$* знову зменшує карту ознак удвічі, цього разу до розміру 8×8 пікселів. На цьому етапі модель переходить від локальних структур до більш глобальних характеристик об'єкта. Це робить подальшу обробку значно ефективнішою та дозволяє мережі краще виявляти великомасштабні патерни, які є ключовими для класифікації.

6. *Шар $\text{Flatten}()$* перетворює двовимірні карти ознак у один довгий вектор, що необхідно для подальшої передачі даних у повнозв'язні шари. Цей етап не змінює значення елементів, але змінює структуру даних, переводячи їх у формат, з яким можуть працювати Dense-шари.

7. *Повнозв'язний шар $\text{Dense}(128, \text{activation}='relu')$* виконує узагальнення інформації, отриманої після всіх згорткових операцій. Він аналізує сукупність

виділених ознак і формує внутрішнє представлення об'єкта, необхідне для подальшої класифікації. Кількість нейронів – 128 – є оптимальною для датасету CIFAR-10, оскільки забезпечує баланс між виражальною здатністю та обчислювальною ефективністю. Активація ReLU сприяє навчанню складних взаємозв'язків між ознаками.

8. Шар *Dropout(0.3)* випадковим чином вимикає 30% нейронів під час навчання. Це серйозно знижує ризик перенавчання, дозволяючи моделі краще узагальнювати і не «запам'ятовувати» тренувальні дані. Крім того, Dropout робить мережу більш стійкою до шуму, оскільки вона вчиться працювати у ситуаціях, коли частина інформації недоступна – це імітує реальні умови зашумлення та погіршення якості зображень.

9. *Вихідний шар* – останній шар, *Dense(10, activation='softmax')*, повертає розподіл ймовірностей між десятима класами CIFAR-10: літак, автомобіль, птах, кіт, олень, собака, жаба, кінь, корабель і вантажівка. Softmax забезпечує нормування виходів так, щоб їх сума дорівнювала 1, що робить результати інтерпретованими як ймовірності. Це стандартний підхід для багатокласової класифікації.

3.4.2. Побудова моделі ResNet-18

ResNet-18 є однією з найпоширеніших глибоких резидуальних архітектур, що включає 18 навчальних шарів та використовує ключову ідею – резидуальні зв'язки (*skip connections*). На відміну від класичних згорткових мереж, ResNet-18 дозволяє ефективно навчати глибокі моделі без проблем затухання чи вибуху градієнтів. Це досягається завдяки тому, що кожен блок у мережі навчає не абсолютне перетворення входу, а залишок між вхідним та бажаним вихідним представленням. У результаті модель здатна краще зберігати інформацію та стійкіше працює навіть на зашумлених даних.

Архітектурно ResNet-18 складається з кількох послідовних модулів, кожен з яких виконує специфічну функцію в перетворенні та узагальненні ознак.

Архітектурна структура моделі:

1. *Початковий згортковий модуль.* Модель ResNet-18 починає обробку даних з вхідного зображення розміром $32 \times 32 \times 3$. Першим застосовується великий згортковий шар із ядром 7×7 , що охоплює більшу частину зображення в одному кроці. Такий шар забезпечує грубе вилучення низькорівневих ознак на широкому рецептивному полі: плавні контури, великі переходи інтенсивності та глобальні напрямки ліній. На відміну від класичних CNN, де переважають фільтри 3×3 , використання ядра 7×7 дозволяє отримати більш узагальнене та глобальне представлення на початковому етапі. Після цього застосовується шар підвибірки MaxPooling, який зменшує просторову роздільну здатність карти ознак, знижує обчислювальну складність і дає змогу моделі зосередитися на важливих структурних компонентах об'єкта.

2. *Резидуальні блоки першої групи.* Після первинної обробки модель переходить до першої серії резидуальних блоків, кожен з яких містить два згорткові шари 3×3 із функцією активації ReLU. Головною особливістю таких блоків є наявність резидуального зв'язку: вхідні дані блоку додаються до його виходу. Такий механізм дає змогу уникати затухання градієнтів та зберігати інформацію про попередні ознаки на всьому шляху прямування сигналу. На цьому рівні мережа навчається виділяти базові структури об'єктів – орієнтації ліній, прості контури та локальні текстури. Водночас резидуальні зв'язки гарантують, що навіть якщо деякі перетворення дещо спотворюють сигнал, його первинна форма не втрачається.

3. *Резидуальні блоки другої групи.* У другому наборі блоків кількість каналів збільшується до 128. Це означає перехід до аналізу складніших і більш різноманітних ознак. Коли кількість каналів розширюється, звичайне додавання входу та виходу неможливе через різні розмірності, тому у блоках застосовуються спеціальні проєкційні шари – згортки 1×1 . Вони вирівнюють кількість каналів перед операцією додавання, роблячи резидуальний зв'язок коректним. На цьому етапі модель уже здатна розпізнавати складніші локальні фрагменти об'єктів, такі як характерні деталі форми, текстури та комбінації контурів, які є типовими для певних категорій CIFAR-10.

4. *Резидуальні блоки третьої та четвертої групи.* У третій та четвертій групах кількість каналів збільшується до 256 та 512 відповідно – це найглибші рівні обробки в ResNet-18. Тут модель опрацьовує абстрактні та високорівневі представлення об'єктів. Саме в цих блоках формується комплексне семантичне уявлення: частини об'єкта поєднуються у цілісну структуру, яка дозволяє точніше розпізнавати клас. Застосування резидуальних зв'язків на цьому рівні особливо важливе, оскільки зростає ризик втрати частини інформації через багаторазові нелінійні перетворення. Skip-connections гарантують, що інформація про структуру зображення, отримана на ранніх етапах, не зникає і може бути використана на фінальних рівнях прийняття рішення. Це робить ResNet-18 значно стійкішою до зашумлення, оскільки модель може опиратися не лише на локальні патерни, які легко спотворюються шумом, а й на глобальні та стійкі ознаки.

5. *Глобальний усереднюючий пулінг.* Після проходження крізь усі резидуальні блоки модель отримує високорівневі карти ознак. Замість передачі цих карт у великий повнозв'язний шар, як це робиться у традиційних CNN, ResNet-18 використовує Global Average Pooling. Ця операція обчислює середнє значення кожного каналу ознак і перетворює карту ознак у компактний вектор. Такий метод значно зменшує кількість параметрів, робить систему менш схильною до перенавчання та підвищує її стійкість до шумових коливань. Крім того, global pooling робить модель менш чутливою до положення об'єкта в кадрі та до локальних спотворень.

6. *Вихідний класифікаційний шар.* Отриманий після глобального пулінгу вектор ознак подається у заключний повнозв'язний шар, який містить 10 нейронів – відповідно до кількості класів у датасеті CIFAR-10. На виході використовується функція активації Softmax, яка перетворює значення на ймовірнісні оцінки належності зображення до кожного класу. Цей етап завершує процес обробки та забезпечує формування підсумкового рішення класифікатора. Разом з попередніми рівнями ця структура дозволяє ResNet-18

досягати високої точності навіть при наявності значного шуму, оскільки модель опирається на стійкі високорівневі представлення об'єктів.

3.4.3. Гібридна модель Autoencoder + CNN

Останньою моделлю дослідження було обрано побудувати гібридний підхід, що поєднує автоенкодер та згорткову нейронну мережу (CNN). Такий метод є актуальним у задачах класифікації зашумлених зображень, оскільки дозволяє розділити процес на два незалежні етапи: очищення даних і їх подальшу класифікацію. Ідея полягає в тому, що автоенкодер виконує попереднє відновлення зображення, зменшуючи інтенсивність шуму та відтворюючи максимально чисту структуру, після чого CNN працює з уже обробленими даними, що підвищує точність класифікації.

Автоенкодер складається з двох основних частин – енкодера і декодера. Енкодер поступово зменшує просторовий розмір зображення, зберігаючи найбільш інформативні характеристики. У початкових шарах енкодер виконує базове виділення ознак за допомогою згортки 3×3 , після чого застосовується MaxPooling, що знижує розмірність та пригнічує випадкові спотворення. На глибших рівнях енкодер переходить до виділення складніших патернів, що дозволяє моделі формувати компактне, але змістовне внутрішнє представлення.

Декодер, навпаки, виконує відновлення просторової структури зображення. Розгортальні (Conv2DTranspose) шари по черзі збільшують розмір карт ознак, відтворюючи початкове зображення з фільтрованими шумовими компонентами. Фінальний шар з активацією sigmoid генерує реконструйоване RGB-зображення у діапазоні $[0, 1]$. Завдяки симетричній структурі енкодера і декодера, автоенкодер вчиться відокремлювати шум від корисних ознак, відтворюючи лише ті компоненти, які є важливими для подальшої класифікації.

На другому етапі очищені зображення передаються у базову згорткову мережу CNN, описану у розділі 3.4.1. CNN виконує традиційну класифікацію, працюючи вже не з зашумленими даними, а з результатом роботи

автоенкодера. Завдяки цьому мережа отримує на вхід більш стабільне, структуроване зображення, що позитивно впливає на точність розпізнавання. Такий підхід дозволяє моделі зменшити залежність від інтенсивності шуму та працювати значно ефективніше на складних або спотворених даних.

Архітектурна структура моделі Autoencoder:

1. *Початковий згортковий шар.* Робота автоенкодера починається з обробки вхідного зображення розміром $32 \times 32 \times 3$. Перший згортковий шар з 32 фільтрами розміром 3×3 виконує базове вилучення низькорівневих ознак – таких як контури, дрібні деталі та локальні зміни інтенсивності. Використання активації ReLU дозволяє моделі виявляти нелінійні патерни та запобігає проблемі затухання градієнтів. Падинг 'same' зберігає розмірність простору ознак, що важливо для подальшої реконструкції.

2. *Перший шар підвибірки MaxPooling.* Після первинного вилучення ознак дані проходять через MaxPooling із вікном 2×2 . Цей шар зменшує просторову роздільність зображення у два рази, знижуючи масштаб карти ознак до 16×16 . Така операція виконує дві ключові функції: по-перше, зменшує кількість параметрів і обчислень у подальших шарах, а по-друге, дозволяє моделі зосередитися на найбільш виражених елементах сигналу, придушуючи дрібний шум, що є критично важливим для задачі денойзингу.

3. *Другий згортковий шар.* На наступному етапі застосовується згортковий шар із 64 фільтрами розміром 3×3 . Збільшення кількості каналів дозволяє моделі вилучати складніші ознаки, які не були доступні на попередньому рівні: локальні текстури, комбінації контурів і характерні фрагменти об'єктів. Завдяки ReLU автоенкодер здатний моделювати нелінійні залежності між пікселями, що суттєво покращує здатність відокремлювати шум від корисного сигналу.

4. *Другий шар MaxPooling.* Другий етап субдискретизації знову зменшує розмірність карти ознак у два рази, приводячи її до 8×8 пікселів. На цьому рівні формується компактне просторове представлення зображення – свого роду «кодування» об'єкта, в якому зберігається ключова інформація про структуру,

але при цьому значна частина шуму відсікається. Для автоенкодерів така компресія є важливою, оскільки змушує модель навчитися відокремлювати суттєві ознаки від випадкових спотворень.

5. Перший транспонований згортковий шар. У декодері модель починає відновлення простору ознак. Перший транспонований згортковий шар із 64 фільтрами та кроком 2 виконує операцію, протилежну MaxPooling: збільшує розмірність карти ознак із 8×8 до 16×16 . Завдяки фільтрам 3×3 цей шар поступово відновлює загальні контури та структуру зображення, враховуючи інформацію, закодовану у компресованій формі.

6. Другий транспонований згортковий шар. Другий шар декодера також використовує транспоновану згортку – цього разу з 32 фільтрами. Він збільшує розмірність карти ознак до початкових 32×32 пікселів. На цьому рівні модель відновлює більш дрібні деталі, кольорові переходи та локальні елементи структури. Саме цей шар відтворює текстурні патерни, які могли бути втрачені у фазі кодування.

7. Вихідний згортковий шар. Фінальний шар автоенкодера застосовує звичайну згортку 3×3 з трьома каналами на виході – відповідно до кількості каналів в оригінальному зображенні RGB. На відміну від попередніх шарів, він використовує активацію sigmoid, яка обмежує значення кожного пікселя в діапазоні $[0,1]$ і забезпечує коректне формування вихідного зображення. На цьому етапі модель формує очищену реконструкцію, зменшуючи рівень шуму і зберігаючи ключові візуальні ознаки.

3.5. Аналіз результатів навчання і тестування моделей

У результаті побудови моделей були отримані наступні результати точності для кожного виду шуму:

3.5.1. Результати роботи базової моделі CNN

Проста модель CNN була попередньо навчена на чистих зображеннях CIFAR-10, після чого її було протестовано на різних типах зашумлених

зображеннях. Результати тестування моделі CNN наведені в табл. 3.1 та на рис. 3.1.

Таблиця 3.1

Показники точності для кожного типу шумів при використанні моделі CNN

Вид шуму	Показник точності (accuracy)
gaussian_0.1	0.5439
gaussian_0.2	0.3176
saltpepper_0.02	0.4886
speckle_0.1	0.6288
poisson	0.6423

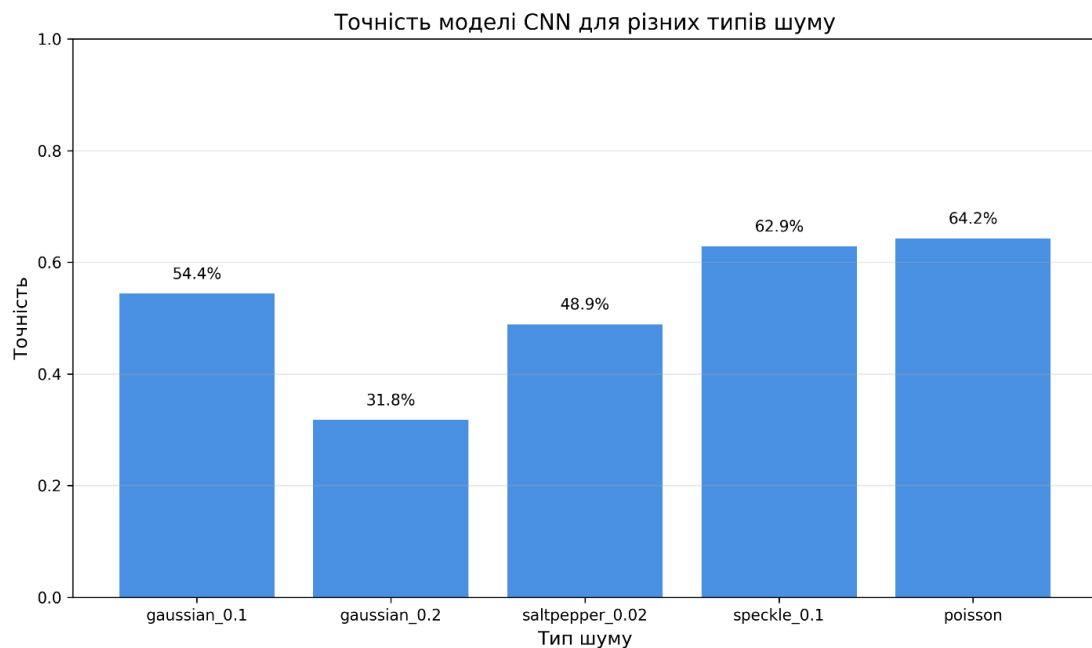


Рисунок 3.1 – Гістограма точності моделі CNN для різних типів шуму

Отримані результати свідчать про те, що базова модель CNN демонструє середню стійкість до шумів: найгірший показник спостерігається при гауссівському шумі високої інтенсивності, тоді як найкраще CNN справляється

зі speckle та poisson шумами. Це пов'язано з тим, що модель навчалася на чистих зображеннях і не має механізмів адаптації до сильних спотворень.

3.5.2. Результати роботи моделі ResNet-18

Модель ResNet-18, як і базова модель CNN також навчалася на чистій вибірці зображень, після чого була перевірена на зображеннях із додаванням шумів на них. Результати тестування моделі ResNet-18 наведені в табл. 3.2 та на рис. 3.2.

Таблиця 3.2

Показники точності для кожного типу шумів при використанні моделі ResNet-18

Вид шуму	Показник точності
gaussian_0.1	0.5281000137329102
gaussian_0.2	0.2903999984264374
saltpepper_0.02	0.415800005197525
speckle_0.1	0.6592000126838684
poisson	0.671999990940094

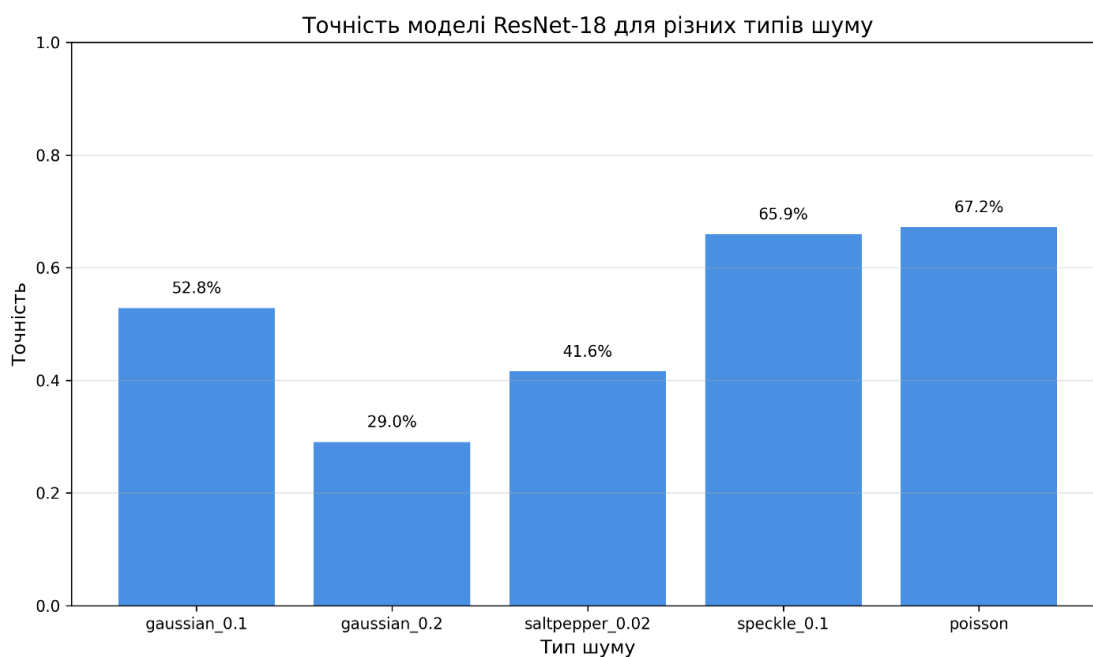


Рисунок 3.2 – Гістограма точності моделі ResNet-18 для різних типів шуму

Модель ResNet-18 показує схожу динаміку поведінки, однак у більшості випадків її точність є дещо нижчою за CNN. Незважаючи на глибшу архітектуру, сильний шум негативно впливає на результати, оскільки ResNet-18 також не була спеціально адаптована до шумових спотворень. Найкращий результат модель демонструє для speckle та poisson шумів, що узгоджується з поведінкою базової CNN.

3.5.3. Результати роботи гібридної моделі

Третя модель була гібридною Autoencoder&CNN. Першим чином було навчено автоенкодер прибирати з зображень різні типів шумів. У цей час модель CNN залишалась без змін, та була навчена на стандартних зображеннях. Тестування відбувалось наступним чином: зашумлене зображення передавалось до автоенкодера для прибирання шумів, після чого очищене зображення далі надходило до CNN, котра і проводила класифікацію.

Результати тестування гібридної моделі наведені в табл. 3.3 та на рис. 3.3.

Таблиця 3.3

Показники точності для кожного типу шумів при використанні гібридної моделі Autoencoder + CNN

Вид шуму	Показник точності
gaussian_0.1	0.2903999984264374
gaussian_0.2	0.415800005197525
saltpepper_0.02	0.6592000126838684
speckle_0.1	0.671999990940094
poisson	0.5281000137329102

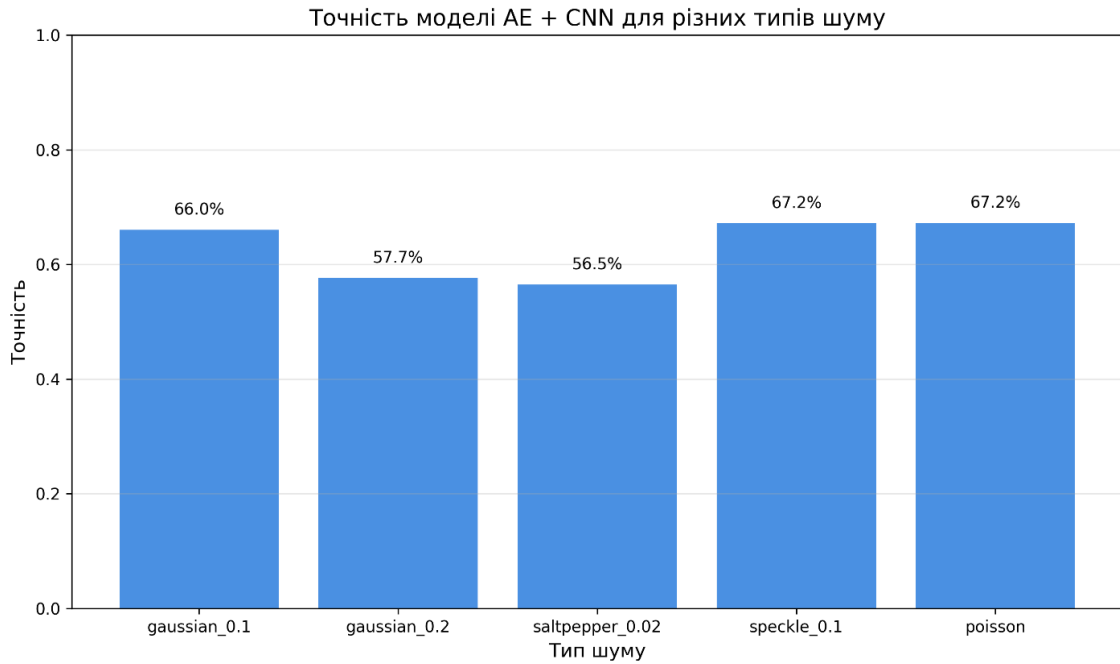


Рисунок 3.3 – Гістограма точності гібридної моделі Autoencoder + CNN для різних типів шуму

Гібридна модель Autoencoder + CNN суттєво покращує результати класифікації порівняно з попередніми моделями. Автоенкодер ефективно видаляє більшість типів шумів перед подачею зображення до класифікатора, що забезпечує значне зростання точності, особливо при salt-and-pepper, speckle та gaussian шумі високої інтенсивності. Це підтверджує перевагу використання попередньої очистки зображень у задачах класифікації зашумлених даних.

Висновок до розділу 3

У третьому розділі було проведено практичне дослідження роботи трьох підходів до класифікації зашумлених зображень: базової згорткової нейронної мережі (CNN), глибокої архітектури ResNet-18 та гібридної моделі Autoencoder + CNN. Метою експериментів було оцінити, наскільки кожна з моделей здатна зберігати точність класифікації за умов різних типів та інтенсивностей шуму, а також визначити, який із підходів є найбільш стійким до деградації вхідних даних.

Під час дослідження використовувалися п'ять типів шумів: гауссівський шум різної дисперсії, імпульсний "сіль і перець", спекл-шум та пуассонівський

шум, які моделюють різні варіанти спотворень, що можуть виникати в реальних умовах отримання зображення. Кожна модель тестувалася на однаковому наборі зашумлених даних, що забезпечує коректність порівняння результатів.

Отримані результати показали, що базова CNN демонструє відносно середню стійкість до шумів: вона краще працює зі спекл- та пуассонівським шумом (точність 0.6288 та 0.6423 відповідно), проте значно втрачає точність при збільшенні гауссівського шуму (до 0.3176). Це підтверджує висновки теоретичного розділу: неглибокі CNN добре вилучають локальні ознаки, але є чутливими до випадкових високочастотних спотворень.

Модель ResNet-18 у більшості випадків показала кращу або співставну точність порівняно з базовою CNN. Особливо добре вона впоралася зі спекл- та пуассонівським шумом (0.6592 та 0.6719 відповідно), що підтверджує перевагу глибших резидуальних архітектур, здатних виділяти високорівневі та стійкі ознаки. Проте ефективність ResNet-18 також суттєво падає при інтенсивному гауссівському шумі (0.2904), що свідчить про те, що навіть глибокі моделі не здатні повністю компенсувати зашумлення без спеціальних методів попередньої обробки.

Найвищу стійкість до шуму у загальному випадку продемонстрував гібридний підхід Autoencoder + CNN, який передбачає попереднє очищення зображень за допомогою автоенкодера. Особливо помітним є покращення при високих рівнях гауссівського шуму та шуму "сіль і перець" – точність зросла до 0.5767 та 0.5649 відповідно. Це свідчить про те, що автоенкодер ефективно пригнічує випадкові спотворення, зберігаючи ключові структури зображення. Однак, результати показали, що гібридний підхід не завжди є найкращим для всіх типів шуму: у випадку пуассонівського шуму точність знизилась до 0.5281. Це вказує на те, що автоенкодер, навчаючись переважно на гауссівському шумі, гірше узагальнює нестандартні або шумові патерни з іншою статистичною природою.

ВИСНОВКИ

У кваліфікаційній роботі було проведено комплексне дослідження проблеми класифікації об'єктів на зашумлених зображеннях, що є актуальним завданням сучасного комп'ютерного зору. Наявність шуму суттєво ускладнює процес виділення ознак, порушує статистику даних та знижує точність роботи моделей глибокого навчання. Тому пошук ефективних підходів, здатних забезпечити стійкість класифікації в умовах деградації зображень, має як практичне, так і наукове значення.

У першому розділі було систематизовано теоретичні основи задачі класифікації зашумлених зображень, розглянуто природу різних типів шумів (гауссівський, імпульсний, пуассонівський, спекл-шум), їх вплив на низькорівневі та високорівневі ознаки, а також окреслено основні підходи до їх обробки. Показано, що навіть незначне шумове спотворення може істотно змінити розподіл ознак, що негативно впливає на роботу класифікаторів.

У другому розділі було детально описано сучасні моделі глибокого навчання, застосовні до задачі класифікації: згорткові нейронні мережі (CNN), модель ResNet-18 із резидуальними зв'язками, а також гібридні підходи, що поєднують моделі очищення зображень (Autoencoder, DnCNN) із класифікаційними мережами. Особлива увага приділялась принципам побудови резидуальних блоків, механізму skip-connections та їх ролі у зменшенні втрат інформації під час глибокої обробки даних.

У третьому розділі проведено практичне дослідження трьох архітектур: базової CNN, моделі ResNet-18 та гібридної системи Autoencoder + CNN. Для експериментів використовувався датасет CIFAR-10, а тестування здійснювалося на зображеннях, спотворених різними типами шуму. За результатами експериментів встановлено, що:

- базова CNN демонструє середній рівень стійкості. Її точність різко знижується при збільшенні інтенсивності шуму, проте вона зберігає прийнятні результати для спекл- та пуассонівського шуму;

– ResNet-18 показує подібну залежність до CNN, однак при деяких видах шуму (зокрема speckle та poisson) демонструє вищу стабільність завдяки використанню резидуальних зв'язків;

– гібридна модель Autoencoder + CNN істотно підвищує точність класифікації у порівнянні з іншими підходами. Попереднє очищення зображень автоенкодером дозволяє компенсувати значну частину шумових спотворень, що позитивно впливає на роботу CNN-класифікатора.

Отримані результати показують, що комбінація автоенкодера та CNN може бути використана для підвищення стійкості систем комп'ютерного зору, які працюють у складі динамічних систем. Це дозволить покращити якість ідентифікації об'єктів у реальному часі навіть за умови значних спотворень оптичних сигналів, що є важливим для задач навігації, керування рухом, супроводу об'єктів та уникнення зіткнень.

ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Goodfellow, I., Bengio, Y., Courville, A. *Deep Learning*. MIT Press, 2016. 800 p. DOI: 10.1007/s10710-017-9314-z.
2. He, K., Zhang, X., Ren, S., Sun, J. Deep Residual Learning for Image Recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2016, pp. 770-778. doi: 10.1109/CVPR.2016.90.
3. Vincent, P., Larochelle, H., Lajoie, I., Bengio, Y., Manzagol, P.-A. Stacked Denoising Autoencoders: Learning Useful Representations in a Deep Network with a Local Denoising Criterion. *Journal of Machine Learning Research*. 2010, no. 11, pp. 3371-3408.
4. Zhang, K., Zuo, W., Chen, Y., Meng, D., Zhang, L. Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising. *IEEE Transactions on Image Processing*. 2017, vol. 26, no. 7, pp. 3142-3155. doi: 10.1109/TIP.2017.2662206.
5. Krizhevsky, A. *Learning Multiple Layers of Features from Tiny Images*. CIFAR-10 Technical Report, 2009. URL: <https://www.cs.toronto.edu/~kriz/learning-features-2009-TR.pdf>.
6. O'Shea, K., Nash, R. An Introduction to Convolutional Neural Networks. *arXiv. Computer Science*. 2015. <https://doi.org/10.48550/arXiv.1511.08458>.
7. Chollet, F. *Deep Learning with Python*. Manning Publications, 2018. 373 p.
8. Gonzalez, R. C., Woods, R. E. *Digital Image Processing*. Pearson Educational International, 2018. 976 p.
9. Image noise. *Wikipedia*. URL: https://en.wikipedia.org/wiki/Image_noise.
10. Shot noise. *Wikipedia*. URL: https://en.wikipedia.org/wiki/Shot_noise.
11. Speckle (interference). *Wikipedia*. URL: [https://en.wikipedia.org/wiki/Speckle_\(interference\)](https://en.wikipedia.org/wiki/Speckle_(interference)).

12. What is CNN for image processing. URL: <https://intelgic.com/what-is-cnn-for-image-processing#:~:text=CNNS%20process%20input%20images%20through,shapes%20and%20objects%20are%20recognized.>
13. Autoencoder for Computer Vision – the Complete Guide. URL: <https://viso.ai/deep-learning/autoencoder/>
14. Zargul, A. Understanding Image Noise: Exploring Different Kinds of Noises in Image Processing. *Medium*. 04.02.2025. URL: https://medium.com/@zar_373/understanding-image-noise-exploring-different-kinds-of-noises-in-image-processing-dcd26c6c9e5a.
15. Sudipta Singha Roy, Mahtab A., M. A. H. Akhand. Noisy image classification using hybrid deep learning methods. *Journal of Information and Communication Technology*. 2018, vol. 17(2), pp. 233-269. DOI: 10.32890/jict2018.17.2.4
16. *GeeksForGeeks*. Computer Vision. 03.10.2025. URL: <https://www.geeksforgeeks.org/computer-vision/what-is-image-classification/>.
17. *GeeksForGeeks*. Machine Learning. 03.10.2025. URL: <https://www.geeksforgeeks.org/machine-learning/auto-encoders/>
18. Understanding Image Noise: Exploring Different Kinds of Noises in Image Processing. URL: https://medium.com/@zar_373/understanding-image-noise-exploring-different-kinds-of-noises-in-image-processing-dcd26c6c9e5a
19. Keras Documentation URL: <https://keras.io>
20. TensorFlow Documentation URL: <https://www.tensorflow.org>
21. NumPy Documentation URL: <https://numpy.org/doc/>
22. Pandas Documentation URL: <https://pandas.pydata.org/docs/>
23. Scikit-learn Documentation URL: <https://scikit-learn.org/>

ДОДАТКИ

Додаток А

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Харківський національний університет імені В. Н. Каразіна

Навчально-науковий інститут комп'ютерних наук та штучного інтелекту
Кафедра комп'ютерних систем та робототехніки
Рівень вищої освіти (освітньо-кваліфікаційний рівень) Магістр
Галузь знань: 12 – Інформаційні технології
Спеціальність: 123 «Комп'ютерна інженерія»
Освітня програма «Комп'ютерна інженерія»

ЗАТВЕРДЖУЮ
Завідувач кафедри комп'ютерних
систем та робототехніки
к. ф.-м. н., доц. ХРУСЛОВ М. М.
«02» жовтня 2024 року

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ МАГІСТРА

ЯКОВЛЕВА Данила В'ячеславовича
(прізвище, ім'я, по батькові студента)

1. Тема роботи «МЕТОД РОЗПІЗНАВАННЯ ОБ'ЄКТІВ НА ЗАШУМЛЕНИХ
ЗОБРАЖЕННЯХ»

керівник роботи **СТРІЛЕЦЬ Вікторія Євгенівна, к.т.н., доцент кафедри**
комп'ютерних систем та робототехніки
(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

затверджені наказом по університету № 4101-5/3554 від 30 вересня 2025 року

2. Строк подання студентом роботи 30 листопада 2025 року

3. Перелік питань, які потрібно розробити

1. Аналіз існуючих методів і підходів до розпізнавання об'єктів на зашумлених зображеннях.
2. Аналіз типів шумів та методів їх опрацювання, а також впливу різних типів шумів на якість обробки та розпізнавання зображень.
3. Оцінка архітектур моделей нейронних мереж для розпізнавання зображень та їх стійкості до шумових спотворень.
4. Побудова та навчання моделей для класифікації зашумлених зображень.
5. Порівняння результатів роботи моделей і формування узагальнених висновків щодо їх точності та стійкості.

4. План роботи

№ з/п	Назви етапів роботи	Термін виконання етапів роботи
1	Затвердження теми роботи та керівника	02.09.2024 -
2	Ознайомлення з іноземними науково-інформаційними джерелами з питань методів обробки зашумлених зображень	03.09.2024 - 24.10.2024
3	Аналіз існуючих методів і підходів до розпізнавання об'єктів на зашумлених зображеннях.	25.10.2024 - 09.11.2024
4	Аналіз типів шумів та методів їх опрацювання, а також впливу різних типів шумів на якість обробки та розпізнавання зображень.	10.11.2024 - 24.11.2024
5	Оцінка архітектур моделей нейронних мереж для розпізнавання зображень та їх стійкості до шумових спотворень.	25.11.2024 - 08.12.2024
6	Побудова та навчання моделей для класифікації зашумлених зображень.	09.12.2024 - 29.01.2025
7	Тестування та оцінка ефективності розроблених моделей	30.01.2025- 31.02.2025
8	Порівняння результатів роботи моделей і формування узагальнених висновків щодо їх точності та стійкості.	01.03.2025- 01.04.2025
9	Підготовка і оформлення звітних матеріалів.	01.05.2025- 30.08.2025-
10	Оформлення звіту про науково-дослідну практику Написання статті за матеріалами кваліфікаційної роботи.	01.09.2025- 30.10.2025-
11	Підготовка і оформлення звітних матеріалів кваліфікаційної роботи. Оформлення списку літератури	10.10.2025- 30.10.2025-
12	Оформлення пояснювальної записки кваліфікаційної роботи відповідно вимогам до звітів про НДР.	10.10.2025- 30.11.2025
13	Підготовка і оформлення звітних матеріалів та додатків кваліфікаційної роботи.	01.11.2025- 30.11.2025
14	Оформлення звіту про переддипломну практику	24.11.2025 - 30.11.2025
15	Представлення кваліфікаційної роботи керівнику та рецензенту	24.11.2025 - 30.11.2025

5. Дата видачі завдання 02 жовтня 2025 року.

Студент

Д.В. Яковлев

ініціали, прізвище



підпис

Керівник роботи

В.Є. Стрілець

ініціали, прізвище



підпис

Додаток Б

Затверджую

«_____» _____ 2025 р.

Технічне завдання
на розробку програмного виробу «Метод розпізнавання об'єктів на
зашумлених зображеннях»

1.	Введення	1.1. Назва: Метод розпізнавання об'єктів на <u>зашумлених зображеннях</u> . 1.2. Галузь застосування: інформаційні технології, штучний інтелект, комп'ютерні системи
2.	Підстава для розробки	2.1. Навчальний план за спеціальністю 123 – Комп'ютерна інженерія 2.2. Завдання на кваліфікаційну роботу магістра № 4101-5/3554 від «30» вересня 2025 р. (представити як Додаток А до пояснювальної записки до кваліфікаційної роботи).
3.	Призначення розробки	3.1. Мета розробки: підвищити точність відновлення та розпізнавання об'єктів на <u>зашумлених зображеннях</u> шляхом дослідження та застосування сучасних методів обробки, включаючи традиційні, інтелектуальні та гібридні підходи. 3.2. Призначення розробки: розроблений програмний виріб призначений для автоматичного очищення та класифікації зображень, що містять різні види шумів. Модель забезпечує виявлення та розпізнавання об'єктів у складних візуальних умовах, що дозволяє застосовувати її у таких сферах, як відеоспостереження, аналіз медичних зображень, робототехніка, системи безпеки, транспортні системи та комп'ютерний зір. 3.3. Вихідні дані розробки: зображення, попередньо спотворені різними видами шумів
4.	Технічні вимоги до програмного виробу	4.1. Вимоги до функціональних характеристик. Програма має: забезпечувати завантаження зображень у форматах PNG, JPEG або BMP; підтримувати імпорт навчальних наборів даних із локальної директорії або з вказаних шляхів; повинна мати механізм додавання шуму різних типів (<u>Gaussian, Salt-and-Pepper, Poisson, Speckle</u>) для формування навчальних і тестових наборів; містити модуль побудови моделі та

		налаштування гіперпараметрів; забезпечувати автоматичне тестування моделі на зашумлених зображеннях; має бути реалізовано обчислення показників якості (наприклад, <u>ассигасу</u>). 4.2. Вимоги до надійності: програма має стабільно працювати з великими наборами даних, забезпечувати коректне оброблення зображень та зберігати стійкість при роботі з високими рівнями шумів. 4.3. Вимоги до умов експлуатації: доступ до глобальної мережі <u>Internet</u> (за необхідності отримання бібліотек). Програма повинна запускатися в умовах користувацької або серверної інфраструктури. 4.4. Вимоги до складу і параметрів технічних засобів: для виконання програми підходить ПК пристрій із будь-якою операційною системою (Windows, Linux, <u>Mac OS</u>). 4.5. Вимоги до інформаційної та програмної сумісності: підтримка сучасних ОС (Windows, Linux), встановлений Python 3.x, <u>TensorFlow</u>, <u>NumPy</u>, <u>Pandas</u>, <u>Matplotlib</u> та інші допоміжні бібліотеки. 4.6. Вимоги до маркування та упаковки: вимоги до маркування та упаковки не представляються. 4.7. Вимоги до транспортування і зберігання: вимоги до транспортування та зберігання не представляються, можливе зберігання на хмарних носіях або локальних дисках. 4.8. Спеціальні вимоги: результати роботи мають бути придатними до подальшого використання для аналітики, навчання студентів, дослідницьких експериментів та інтеграції у системи комп'ютерного зору.
5.	Вимоги до програмної документації	Програмною документацією до виробу «Метод розпізнавання об'єктів на <u>зашумлених зображеннях</u>» вважати: 1) Опис алгоритмів та програмного виробу (представити як розділ 3 пояснювальної записки до кваліфікаційної роботи) 2) Програму і методику випробувань розробленої системи (представити як додаток В до пояснювальної записки до кваліфікаційної роботи). 3) Опис програмного виробу (представити в розділі 3 пояснювальної записки до кваліфікаційної роботи).

6.	Вимоги до техніко-економічних показників	Програмною документацією до виробу «Метод розпізнавання об'єктів на <u>зашумлених зображеннях</u>» вважати: 1) Дане Технічне завдання на розробку програмного виробу (представити у вигляді Додатку Б до пояснювальної записки до кваліфікаційної роботи). 2) Опис програмного виробу (представити в розділі 3) 3 пояснювальної записки до кваліфікаційної роботи). 3) Джерела базової інформації.	
7.	Стадії і етапи розробки	Дата	Назва етапу
		від 30 вересня 2025 до 1 жовтня 2025	Затвердження теми роботи та керівника.
		від 1 жовтня 2025 до 3 жовтня 2025	Аналіз наукових джерел і методичних матеріалів щодо створення Telegram чат-ботів та застосування NLP-технологій.
		від 3 жовтня 2025 до 5 жовтня 2025	Вивчення існуючих підходів до класифікації об'єктів на <u>зашумлених зображеннях</u>
		від 5 жовтня 2025 до 15 жовтня 2025	Вибір інструментів і бібліотек для побудови моделей класифікації <u>зображень</u>
		від 16 жовтня 2025 до 25 жовтня 2025	Розробка загальних <u>архітектур</u> моделей для класифікації об'єктів на <u>зашумлених зображеннях</u>
		від 26 жовтня 2025 до 1 листопада 2025	Проведення функціонального тестування моделей та усунення проблем
		від 2 листопада 2025 до 8 листопада 2025	Оцінка точності моделей та ефективності їх роботи
		від 11 листопада 2025 до 15 листопада 2025	Оформлення звіту про науково-дослідну практику
		від 15 листопада 2025 до 18 листопада 2025	

		від 18 листопада 2025 до 25 листопада 2025 від 25 листопада 2025 до 27 листопада 2025 від 27 листопада 2025 до 29 листопада 2025 від 29 листопада 2025 до 30 листопада 2025	Написання статті за матеріалами кваліфікаційної роботи. Підготовка і оформлення звітних матеріалів кваліфікаційної роботи. Оформлення списку літератури. Оформлення пояснювальної записки кваліфікаційної роботи відповідно вимогам до звітів про НДР. Підготовка і оформлення звітних матеріалів та додатків кваліфікаційної роботи. Оформлення звіту про переддипломну практику Представлення кваліфікаційної роботи керівнику та рецензенту
8.	Порядок контролю і приймання програмного продукту (моделі)	1. Перевірку ходу розробки програми виконувати раз в 3 тижні. 2. Захист розробленої моделі провести на засіданні Атестаційної комісії. 3. Пояснювальну записку подати на паперових носіях в 1 примірнику і в електронному вигляді в 1 примірнику.	

Виконавець

студент групи КІ- 61

Яковлев Д.В.,



Замовник

~~канд. техн. наук~~

Стрілець В. Є.



Додаток В

Програма і методика випробувань
програмного виробу
«Метод розпізнавання об'єктів на
запумлених зображеннях»

1 Об'єкт випробувань

1. Назва програмного виробу: «Метод розпізнавання об'єктів на запумлених зображеннях».
2. Галузь застосування: Інформаційні технології.
3. Перераховані відомості запозичуються з відповідних розділів Технічного завдання.

2. Мета випробувань

Перевірка відповідності функціональності програмної реалізації системи заявленим функціональним можливостям в технічному завданні (Додаток Б до пояснювальної записки до кваліфікаційної роботи).

3. Загальні положення**1. Підстави для проведення випробувань**

Підставою для проведення випробувань є наказ про призначення атестаційної комісії.

2. Місце і тривалість випробувань

Приймальні (приймально-здавальні) випробування проводяться на базі комп'ютерного класу кафедри в період роботи атестаційної комісії.

3. Обсяг випробувань

Приймальні випробування програмного виробу проводяться в обсязі відповідному цієї програми і методики випробувань.

4. Організації, які беруть участь у випробуваннях

Приймальні випробування проводяться атестаційною комісією напередодні засідання (або в процесі засідання) за участю Замовника, Виконавця та інших осіб, присутніх на засіданні.

4. Вимоги до програми або програмного виробу

Модель має задовольняти наступним вимогам:

1. Вимоги до функціональних характеристик. Програма має: забезпечувати завантаження зображень у форматах PNG, JPEG або BMP; підтримувати імпорт навчальних наборів даних із локальної директорії або з вказаних шляхів; повинна мати механізм додавання шуму різних типів (Gaussian, Salt-and-Pepper, Poisson, Speckle) для формування навчальних і тестових наборів; містити модуль побудови моделі та налаштування гіперпараметрів; забезпечувати автоматичне тестування моделі на зашумлених зображеннях; має бути реалізовано обчислення показників якості (наприклад, accuracy).
2. Вимоги до надійності. Програма має стабільно працювати з великими наборами даних, забезпечувати коректне оброблення зображень та зберігати стійкість при роботі з високими рівнями шумів.
3. Вимоги до умов експлуатації: доступ до глобальної мережі Internet (за необхідності отримання бібліотек). Програма повинна запускатися в умовах користувацької або серверної інфраструктури.
4. Вимоги до складу і параметрів технічних засобів: для виконання програми підходить ПК пристрій із будь-якою операційною системою (Windows, Linux, Mac OS).
5. Вимоги до інформаційної та програмної сумісності: підтримка сучасних ОС (Windows, Linux), встановлений Python 3.x, TensorFlow, NumPy, Pandas, Matplotlib та інші допоміжні бібліотеки.
6. Вимоги до маркування та упаковки: вимоги до маркування та упаковки не представляються.
7. Вимоги до транспортування і зберігання: вимоги до транспортування та зберігання не представляються, можливе зберігання на хмарних носіях або локальних дисках.

8. Спеціальні вимоги: Результати роботи мають бути придатними до подальшого використання для аналітики, навчання студентів, дослідницьких експериментів та інтеграції у системи комп'ютерного зору.

5. Вимоги до програмної документації

Програмною документацією до виробу «Метод розпізнавання об'єктів на зашумлених зображеннях» вважати:

1. Опис алгоритмів та програмного виробу (представити як розділ 2 пояснювальної записки до кваліфікаційної роботи)
2. Програму і методику випробувань розробленої системи (представити як додаток В до пояснювальної записки до кваліфікаційної роботи).
3. Опис програмного виробу (представити в розділі 3 пояснювальної записки до кваліфікаційної роботи).

6. Засоби і порядок випробувань

6.1 Засоби випробувань

Для проведення випробувань необхідний ПК на якому встановлено Python та (за необхідністю) PyCharm.

6.2 Порядок проведення випробувань

Як правило, випробування проводяться в два етапи:

- ознайомчий (1-й етап);
- випробування програмного виробу (2-й етап).

Перелік перевірок, що проводяться на 1 етапі випробувань:

1. Перевірку комплектності програмної документації.
2. Перевірка комплектності складу програмної документації здійснюється за критерієм наявності зазначеної в ТЗ документації.
3. Перевірку комплектності складу технічних і програмних засобів.
4. Методику проведення перевірок на 1 етапі випробувань.
5. Якість програмної документації перевіряється на відповідність вимогам стандартів ЕСПД.

Перелік перевірок, що проводяться на 2 етапі випробувань:

1. Перевірку відповідності технічних характеристик програми вимогам технічного завдання.
2. Перевірку ступеня виконання функціональних вимог до програми.
3. Методику проведення перевірок, що входять до переліку по 2 етапу випробувань.
4. Запуск додатку PyCharm.
5. Обрати зображення для класифікації, котре співпадає з одним із класів об'єктів у списку classes.
6. У файлі `main.py` у змінну `image_path` вставити шлях до власного зображення.
7. Запустити програму.
8. Дочекались результатів класифікації.

Для проведення випробувань пропонується тест 1 та тест 2.

Тест 1

1. Перевірка виконання програми;
2. Перевірка навчання моделі класифікації та автоенкодера.

	experiment ∇	÷	noise_type ∇	÷	accuracy ∇	÷
1	AE + CNN		gaussian_0.1		0.6603	
2	AE + CNN		gaussian_0.2		0.5767	
3	AE + CNN		saltpepper_0.02		0.5649	
4	AE + CNN		speckle_0.1		0.672	
5	AE + CNN		poisson		0.6725	

Рисунок В.1 – Тест 1

Тест вважається успішним якщо після виконання програми модель показує задовільнені результати точності та стійкість до різного типу шумів

Тест 2

1. Перевірка виконання програми;
2. Перевірка класифікації одиничного зображення.

```
=== РЕЗУЛЬТАТ КЛАСИФІКАЦІЇ ДЛЯ ОДНОГО ЗОБРАЖЕННЯ ===  
Передбачений клас: 1 → automobile
```

Рисунок В.2 – Тест 2

Тест вважається успішним якщо програма зашумить штучним образом зображення, надасть його до автоенкодера для денойзинга та після цього модель розпізнає об'єкт на цьому зображенні як автомобіль.

Висновки: тест 1 успішно пройшов випробування і тест 2 успішно пройшов випробування. Випробування пройшло успішно.

Виконавець: студент групи КІ-61, Яковлев Д.В. _____

