

Національний технічний університет України
«Київський Політехнічний Інститут ім. Ігоря Сікорського»
Міністерство освіти і науки України

Національний технічний університет України
«Київський Політехнічний Інститут ім. Ігоря Сікорського»
Міністерство освіти і науки України

Кваліфікаційна наукова
праця на правах рукопису

ЛИНОВСЬКИЙ АНДРІЙ ОЛЕКСІЙОВИЧ

УДК 004.932:004.852

ДИСЕРТАЦІЯ

Підвищення ефективності методів відновлення зображень на основі
модифікованої згорткової мережі

122 - Комп'ютерні науки

Інформаційні технології

Подається на здобуття наукового ступеня доктора філософії

Дисертація містить результати власних досліджень. Використання ідей, результатів
і текстів інших авторів мають посилання на відповідне джерело

А. О. Линовський

Науковий керівник: Мухін Вадим Євгенійович, д.т.н, професор

Київ - 2026

АНОТАЦІЯ

Линовський А.О. Підвищення ефективності методів відновлення зображень на основі модифікованої згорткової мережі. - Кваліфікаційна наукова праця на правах рукопису.

Дисертаційна робота на здобуття наукового ступеня доктора філософії за спеціальністю 122 «Комп'ютерні науки». - Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», 2026.

Об'єктом дослідження є процес шумозаглушення медичних зображень, пошкоджених імпульсним SaP-шумом. Зокрема згорткові нейронні мережі та модулі уваги.

Предметом дослідження є методи й засоби нейромережевого шумозаглушення, зокрема модифікації архітектури SeConvNet, а також їхній вплив на якість відновлення (PSNR), обчислювальні витрати та стійкість до різних рівнів SaP-шуму на реальних і синтетичних даних.

У першому розділі здійснено всебічний та комплексний аналіз існуючих методів відновлення зображень після шумових спотворень. Досліджено еволюцію підходів, починаючи від класичних методів обробки у просторовій області (лінійні та нелінійні фільтри, зокрема медіанна фільтрація) та методів у перетворюючій області (вейвлет-перетворення, DCT). Встановлено, що класичні просторові фільтри схильні до надмірного згладжування контурів та втрати дрібних деталей при високій щільності імпульсного шуму. Сучасні нелокальні алгоритми, такі як BM3D та NL-Means, забезпечують вищу якість та добре відновлюють повторювані текстури, однак їх використання жорстко обмежене значною обчислювальною складністю. Детально проаналізовано застосування методів машинного та глибокого навчання (CNN, RNN, GAN, дифузійні імовірнісні моделі та трансформери) для усунення шуму. Згорткові нейронні мережі (наприклад, архітектура DnCNN) продемонстрували високу ефективність для гаусівського шуму, проте їх здатність обробляти складний кластерний імпульсний шум

залишається обмеженою. Особливу увагу в розділі приділено специфіці імпульсного шуму типу SaP на медичних рентгенівських зображеннях та наслідкам його появи для діагностики.

У другому розділі проведено глибокий аналіз базової моделі згорткової нейронної мережі SeConvNet. Виявлено її системні обмеження під час обробки медичних рентгенівських знімків, пошкоджених великими кластерними плямами SaP-шуму. Доведено, що через локальну природу SeConv-блоків та використання цільової функції мінімізації середньоквадратичної помилки (MSE) під час навчання, мережа втрачає глобальний контекст. Для великих зашумлених зон гіпотеза математичного «усереднення» дає меншу сумарну помилку, що неминуче спричиняє залишкові артефакти, втрату різкості анатомічних контурів та деградацію локальних контрастів. Для подолання цих критичних недоліків було розроблено та теоретично обґрунтовано модифіковану архітектуру SeConvNet + СВМ. Запропоновано стратегічну інтеграцію модуля канално-просторової уваги безпосередньо після операцій залишкового додавання у згорткових секціях мережі. Канальна увага фокусується на виділенні найбільш інформативних ознак через механізми глобального усереднення та максимізації, а просторова увага детально локалізує ці важливі області на площині. Це дозволило моделі адаптивно підсилювати релевантні структурні ознаки, відрізняючи корисний медичний сигнал від шуму.

У третьому розділі детально розкрито інструментальні засоби, механізми реалізації та алгоритм роботи модифікованої моделі в умовах обмежених обчислювальних ресурсів. Описано формальну математичну модель запропонованої архітектури та поетапний процес відновлення зображень, що поєднує локальні й глобальні стратегії. На першому етапі відбувається первинне локальне заповнення пошкоджених пікселів каскадом із семи селективних згорток, розмір ядра яких поступово зростає від 3x3 до 15x15 пікселів. На другому етапі отримані карти ознак проходять через глибоку згорткову підсистему із прямими шунтувальними (залишковими) зв'язками для видобутку високорівневих ознак. На

третьому етапі відбувається адаптивне зважування ознак за допомогою модулів уваги СВАН. Фінальний шар генерує залишкову корекцію, яка через жорстке маскування застосовується виключно до тих пікселів, що були початково пошкоджені, залишаючи чисті ділянки незмінними. У розділі також глибоко деталізовано специфіку процесу підготовки даних та навчання моделі. Навчальна вибірка формувалася шляхом масштабування, нарізки зображень на фрагменти та інтенсивного розширення даних у 8 разів за рахунок геометричних перетворень. Навчання оптимізовано за допомогою алгоритму Adam з поступовим зменшенням швидкості навчання. Проаналізовано вплив апаратних платформ (CPU, GPU, TPU) на ефективність тренування та розглянуто методи зниження витрат оперативної пам'яті (Gradient checkpointing, змішана точність).

У четвертому розділі представлено результати експериментальних досліджень та здійснено ретельний порівняльний аналіз ефективності зменшення шуму оригінальної та модифікованої моделей. Тестування проводилося на наборі медичних даних NIH Chest X-rays із моделюванням екстремального рівня синтетичного SaP-шуму. Аналіз підтвердив покращення результатів: модифікована модель SeConvNet + СВАН забезпечила систематичне зростання середнього показника PSNR на 4.7%. На особливо критичних ділянках із сильною концентрацією кластерного шуму приріст PSNR сягав 5-8%. Візуальний аналіз відновлених фрагментів довів перевагу модифікованої моделі. Запропонована архітектура значно точніше відновлювала дрібні анатомічні структури, контури ребер та судинний малюнок, успішно мінімізуючи ефект зернистості та розмиття, що були притаманні оригінальній базовій мережі. Аналіз обчислювальної ефективності підтвердив практичну придатність розробленої моделі для роботи в клінічних умовах. Збільшення кількості параметрів та часу обробки одного зображення (з 0.42 с до 0.47 с на базовому процесорі CPU) є мінімальним і повністю виправдовується суттєвим підвищенням діагностичної якості медичних знімків.

Наукова новизна отриманих результатів

Вперше запропоновано метод підвищення ефективності відновлення зображень, пошкоджених імпульсним шумом, який відрізняється інтеграцією механізму канално-просторової уваги (CBAM) у структуру селективної згорткової мережі після залишкових зв'язків, що дозволяє адаптивно підсилювати важливі структурні ознаки та мінімізувати артефакти згладжування при відновленні кластерних шумових плям.

Запропоновано модифіковану модель нейронної мережі SeConvNet, яка відрізняється спеціальним розміщенням блоків уваги для обробки сумарного сигналу прямого поширення та корекційної гілки, що дозволяє підвищити точність відновлення дрібних анатомічних деталей (зріст PSNR до 8% на складних ділянках) при збереженні обчислювальної ефективності для використання в ресурсно-обмежених середовищах.

Практичне значення отриманих результатів полягає в покращенні якості медичних зображень: модифікована SeConvNet + CBAM дозволяє більш акуратно відновлювати дрібні анатомічні деталі при високій щільності SaP-шуму, що підвищує інформативність знімків для попереднього аналізу і візуальної інтерпретації.

Невеликі апаратні вимоги що надає можливість запускати на не вимогливих засобах: додавання CBAM незначно збільшує кількість параметрів (в середньому на 5%), тож впровадження можливе навіть у системах з обмеженими ресурсами (клінічні робочі станції), а для більш продуктивного навчання передбачено перенесення на GPU-середовища.

Вперше запропоновано модифіковану архітектуру SeConvNet + CBAM, яка поєднує селективні локальні SeConv-блоки з послідовним канално-просторовим механізмом уваги (CBAM), що дозволяє адаптивно підсилювати важливі канали й локалізувати інформативні області для ефективного придушення імпульсного SaP-шуму без значного зростання обчислювальних витрат.

Розроблено засоби підвищення достовірності діагностичних процедур на основі стратегічної інтеграції механізму уваги CBAM після операції резидуального додавання у згорткових секціях мережі. Таке архітектурне розміщення гарантує

доступ маски уваги до комбінованого представлення - сумарного сигналу початкового шляху й корекційної гілки.

В результаті в критичних тестах із кластерним шумом досягнуто приросту PSNR до 8%, що забезпечує помітне покращення візуальної якості та інформаційної цінності зображень. Отримані результати демонструють можливість точного відновлення медичних зображень із мінімізацією надмірного згладжування, що дозволяє зберегти важливі анатомічні деталі.

Актуальність та наукова новизна результатів дослідження повною мірою підтверджується їх апробацією та публікацією у провідних фахових виданнях. Зокрема, базові концепції та засоби покращення якості візуальних даних висвітлено у журналі «Телекомунікаційні та інформаційні технології» (2023), де основну увагу було приділено особливостям знешумлення зображень за допомогою комплексного застосування згорткових і рекурентних нейронних мереж, що дозволило закласти фундамент для ефективного виявлення та усунення артефактів.

Подальший розвиток та адаптація запропонованих алгоритмічних рішень відображені у виданні «Кібербезпека: освіта, наука, техніка» (2024). А також публікація в журналі «Проблеми інформатизації та управління» (2024), яка описує найбільш глибоко оптимізований варіант системи. Інтеграція в архітектуру залишкових зв'язків (residual connections) та спеціалізованого модулю уваги (attention module) дала змогу стабілізувати процес глибокого навчання мережі, змушуючи алгоритм прицільно фокусуватися на значущих структурних елементах зображення і максимально нівелювати вплив складних шумових завад.

Ключові слова: згорткові нейронні мережі, рекурентні нейронні мережі, rnn, залишкова нейронна мережа, нелокальні операції, імпульсний шум «сіль і перець», штучний інтелект, машинне навчання

ABSTRACT

Lynovskyi A.O. Improving the Efficiency of Image Restoration Methods Based on a Modified Convolutional Neural Network

Qualifying scientific work in manuscript form. Dissertation for the degree of Doctor of Philosophy in specialty 122 "Computer Sciences". - National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", 2026.

The **object** of the research is the process of denoising of medical images damaged by impulsive SaP-noise. In particular, convolutional neural networks and attention modules.

The **subject** of the research is the methods and means of neural network denoising, in particular modifications of the SeConvNet architecture, as well as their influence on the quality of restoration (PSNR), computational costs, and stability to different levels of SaP-noise on real and synthetic data.

In the first chapter, a comprehensive and complex analysis of existing methods for image restoration after noise distortions is carried out. The evolution of approaches is investigated, starting from classical methods of processing in the spatial domain (linear and nonlinear filters, in particular median filtering) and methods in the transform domain (wavelet transform, DCT). It is established that classical spatial filters are prone to excessive smoothing of contours and loss of fine details at a high density of impulse noise. Modern non-local algorithms, such as BM3D and NL-Means, provide higher quality and well restore repetitive textures, but their use is strictly limited by significant computational complexity. The application of machine and deep learning methods (CNN, RNN, GAN, diffusion probabilistic models, and transformers) for noise removal is analyzed in detail. Convolutional neural networks (for example, the DnCNN architecture) have demonstrated high efficiency for Gaussian noise, however, their ability to process complex cluster impulse noise remains limited. Special attention in the chapter is paid to

the specifics of salt-and-pepper (SaP) impulse noise on medical X-ray images and the consequences of its appearance for diagnosis.

In the second chapter, a deep analysis of the base model of the convolutional neural network SeConvNet is conducted. Its systemic limitations during the processing of medical X-ray images damaged by large cluster spots of SaP noise are revealed. It is proven that due to the local nature of SeConv blocks and the use of the objective function of mean squared error (MSE) minimization during training, the network loses global context. For large noisy zones, the mathematical "averaging" hypothesis gives a smaller total error, which inevitably causes residual artifacts, loss of sharpness of anatomical contours, and degradation of local contrasts. To overcome these critical shortcomings, a modified SeConvNet + CBAM architecture was developed and theoretically substantiated. The strategic integration of the channel-spatial attention module (CBAM) directly after the residual addition operations in the convolutional sections of the network is proposed. Channel attention focuses on extracting the most informative features through global averaging and maximization mechanisms, and spatial attention details and localizes these important areas on the plane. This allowed the model to adaptively amplify relevant structural features, distinguishing the useful medical signal from noise.

The third chapter details the instrumental tools, implementation mechanisms, and operation algorithm of the modified model under conditions of limited computational resources. The formal mathematical model of the proposed architecture and the step-by-step process of image restoration, which combines local and global strategies, are described. At the first stage, the primary local filling of damaged pixels is performed by a cascade of seven selective convolutions, the kernel size of which gradually increases from 3x3 to 15x15 pixels. At the second stage, the obtained feature maps pass through a deep convolutional subsystem with direct shunt (residual) connections to extract high-level features. At the third stage, adaptive weighting of features occurs using CBAM attention modules. The final layer generates a residual correction, which, through rigid masking, is applied exclusively to those pixels that were initially damaged, leaving clean areas unchanged. The chapter also deeply details the specifics of the data preparation and

model training process. The training sample was formed by scaling, slicing images into patches, and intensively augmenting the data by 8 times through geometric transformations. Training is optimized using the Adam algorithm with a gradual learning rate decrease. The influence of hardware platforms (CPU, GPU, TPU) on training efficiency is analyzed, and methods for reducing RAM consumption (Gradient checkpointing, mixed precision) are considered.

The fourth chapter presents the results of experimental research and carries out a thorough comparative analysis of the noise reduction efficiency of the original and modified models. Testing was conducted on the NIH Chest X-rays medical dataset with the simulation of an extreme level of synthetic SaP noise. The analysis confirmed the improvement of the results: the modified SeConvNet + CBAM model provided a systematic increase in the average PSNR index by 4.7%. In particularly critical areas with a strong concentration of cluster noise, the PSNR increase reached 5-8%. Visual analysis of the restored fragments proved the superiority of the modified model. The proposed architecture restored small anatomical structures, contours of ribs, and vascular patterns much more accurately, successfully minimizing the graininess and blurring effects inherent in the original base network. The analysis of computational efficiency confirmed the practical suitability of the developed model for work in clinical settings. The increase in the number of parameters and the processing time of one image (from 0.42 s to 0.47 s on a base CPU) is minimal and is fully justified by a significant increase in the diagnostic quality of medical images.

Scientific novelty of the obtained results. For the first time, a method for increasing the efficiency of restoration of images damaged by impulse noise is proposed, which differs by the integration of the channel-spatial attention mechanism (CBAM) into the structure of a selective convolutional network after residual connections, which allows adaptively amplifying important structural features and minimizing smoothing artifacts when restoring cluster noise spots.

A modified neural network model SeConvNet is proposed, which differs by a special placement of attention blocks for processing the combined forward propagation signal and the correction branch, which allows increasing the accuracy of restoring small anatomical details (PSNR increase up to 8% in complex areas) while maintaining computational efficiency for use in resource-constrained environments.

The practical value of the obtained results lies in the improvement of the quality of medical images: the modified SeConvNet + CBAM allows restoring small anatomical details more accurately at a high density of SaP noise, which increases the informativeness of images for preliminary analysis and visual interpretation.

Low hardware requirements, which provides the ability to run on undemanding devices: the addition of CBAM slightly increases the number of parameters (on average by 5%), so implementation is possible even in systems with limited resources (clinical workstations), and for more productive training, transferring to GPU environments is provided.

For the first time, a modified SeConvNet + CBAM architecture is proposed, which combines selective local SeConv blocks with a sequential channel-spatial attention mechanism (CBAM), which allows adaptively amplifying important channels and localizing informative areas for effective suppression of impulse SaP noise without a significant increase in computational costs.

The approach to selective filtering of impulse noise is improved through the strategic integration of the CBAM attention mechanism after the residual addition operation in the convolutional sections of the network. Such an architectural placement guarantees the access of the attention mask to the combined representation - the sum signal of the initial path and the correction branch. This allows forming a highly accurate context-dependent mask, which more effectively minimizes the excessive smoothing typical for SaP-oriented networks, and ensures the preservation of high-frequency anatomical details. In particular, in critical test scenarios with cluster noise, a PSNR increase of up to 8% is achieved.

The relevance and scientific novelty of the research results are fully confirmed by their validation and publication in leading peer-reviewed journals. Specifically, the fundamental concepts and methods for enhancing visual data quality are detailed in the journal *Telecommunication and Information Technologies* (2023), where the primary focus was on the specifics of image denoising through the combined application of convolutional and recurrent neural networks, establishing the foundation for effective artifact detection and removal.

Further development and adaptation of the proposed algorithmic solutions are reflected in the journal *Cybersecurity: Education, Science, Technique* (2024). As well as the publication in the journal *Problems of Informatization and Management* (2024), which describes the most highly optimized iteration of the system. The integration of residual connections and a specialized attention module into the architecture stabilized the network's deep learning process, forcing the algorithm to focus specifically on significant structural elements of the image and maximally neutralize the impact of complex noise interference.

Keywords: convolutional neural networks, recurrent neural networks, rnn, residual neural network, nonlocal operations, salt-and-pepper impulse noise, artificial intelligence, machine learning

СПИСОК ПУБЛІКАЦІЙ ЗДОБУВАЧА ЗА ТЕМОЮ ДИСЕРТАЦІЇ

Статті у наукових фахових виданнях України (категорія «Б») і конференції:

1. Линовський А. О., Мухін В. Є. Засоби покращення якості та знешумлення зображень на основі застосування згорткових та рекурентних нейронних мереж. Телекомунікаційні та інформаційні технології, № 1 (78), 2023. - с.82 -89. (Здобувачем Линовським А.О. було виконано комплексний огляд сучасних підходів до покращення якості та знешумлення зображень на основі згорткових і рекурентних нейронних мереж з урахуванням non-local operations блоку; співавтором Мухіним В.Є. здійснено загальну постановку задачі дослідження та визначено актуальність застосування non-local operations у сучасних архітектурах)
2. Линовський А. О., Мухін В. Є. Модифікована модель фільтрації шуму на зображеннях на основі згорткової нейронної мережі. Електронне фахове наукове видання «Кібербезпека: освіта, наука, техніка», 4(24), 2024. -с. 388-397. (Здобувачем Линовським А.О. було виконано комплексний аналіз існуючих підходів до видалення шуму типу salt-and-pepper (SaP) і запропоновано модифікацію архітектури SeConvNet для підвищення ефективності фільтрації цього типу шуму; співавтором Мухіним В.Є. обґрунтовано доцільність модифікації існуючих згорткових архітектур для задачі фільтрації імпульсного шуму та проаналізовано отримані результати)
3. Линовський А. О., Мухін В. Є. Модифікована модель фільтрації шуму на зображеннях на основі згорткової нейронної мережі з додаванням залишкових зв'язків і модулю уваги. Проблеми інформатизації та управління, 4 (80), 2024. с. 76-83 (Здобувачем Линовським А.О. було запропоновано модифікацію моделі SeConvNet шляхом додавання блока, що поєднує резидуальні з'єднання з каналною та просторовою увагою (CBAM), що дозволяє вибірково підсилювати структурно важливі фрагменти зображення; співавтором Мухіним В.Є. надано концептуальне обґрунтування застосування механізмів уваги та надано консультації щодо архітектурної інтеграції залишкових зв'язків)

4. Линовський А. О., Мухін В. Є. An Improved CNN Architecture for Cluster Salt-and-pepper Noise Filtering in Medical Data. Міжнародна науково-практична конференція «Інформаційні технології та безпека» (ІТБ-2025) : матеріали конф. Київ, 2025. -с. 148–152. (Здобувачем Линовським А.О. було запропоновано застосування розробленої модифікації згорткової нейронної мережі для специфічної задачі видалення скупчень (кластерів) шуму типу salt-and-pepper на медичних знімках та проведено оцінку ефективності запропонованого підходу; співавтором Мухіним В.Є. визначено специфіку прикладної задачі в контексті обробки медичних даних та затверджено критерії оцінки ефективності розробленого методу)

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ

BN, Batch Normalization - це метод у глибокому навчанні, який стабілізує та прискорює навчання нейронних мереж шляхом нормалізації вхідних даних для кожного шару в межах кожного міні-пакету

CBAM, Convolutional Block Attention Module - модуль уваги згорткового блоку)

CNN, Convolutional Neural Network - згорткова нейронна мережа

PSNR, Peak Signal-to-Noise Ratio - пікове відношення сигналу до шуму

SaP, Salt-and-Pepper - імпульсний шум типу «сіль і перець»

SeConv, Selective Convolution - селективна згортка

SeConvNet, Selective Convolutional Network - мережа селективної згортки

Зміст

ВСТУП	16
РОЗДІЛ 1. АНАЛІЗ ІСНУЮЧИХ МЕТОДІВ ВІДНОВЛЕННЯ ЗОБРАЖЕНЬ ПІСЛЯ ШУМОВИХ СПОТВОРЕНЬ	19
1.1. Класифікація методів відновлення зображень	19
1.2. Аналіз сучасних методів шумозаглушення та відновлення зображень	30
1.3. Методи глибокого навчання для усунення шуму та відновлення зображень	38
1.4 Основні проблеми та виклики в існуючих підходах до відновлення зображень	41
Аналіз наслідків появи шуму в медичних зображеннях	46
1.5. Висновки до розділу.....	47
РОЗДІЛ 2. МОДЕЛІ ТА МЕТОДИ ПІДВИЩЕННЯ ЕФЕКТИВНОСТІ ВІДНОВЛЕННЯ ЗОБРАЖЕНЬ	49
2.1. Основні недоліки існуючої моделі згорткової нейронної мережі для видалення шуму SeConvNet	49
2.2. Модифікація нейромережевої моделі для відновлення зображень на основі застосування механізмів уваги	51
2.3. Аналіз параметрів модифікованої моделі	57
2.4. Метод підвищення ефективності відновлення зображень на основі адаптивний підходу до покращення параметрів зображень	60
2.5. Обґрунтування вибору архітектури SeConvNet + CBAM	63
2.6. Висновки до розділу.....	64
РОЗДІЛ 3. ЗАСОБИ ТА МЕХАНІЗМИ ДЛЯ ПІДВИЩЕННЯ ЕФЕКТИВНОСТІ ПРОЦЕСІВ ВІДНОВЛЕННЯ ЗОБРАЖЕНЬ	66
3.1 Реалізація засобів на основі нейромережевих моделей для підвищення ефективності обробки зображень.....	66
3.2 Механізми реалізації ефективного шумозаглушення в зображеннях.....	72
3.3. Механізми уваги (CBAM, SENet) у практичних системах відновлення зображень	75
3.4. Особливості реалізації засобів підвищення ефективності відновлення зображень в умовах обмежених ресурсів.....	77

3.5. Висновки до розділу.....	83
РОЗДІЛ 4. ЕКСПЕРИМЕНТАЛЬНІ ДОСЛІДЖЕННЯ ТА ОЦІНКА	
ЕФЕКТИВНОСТІ ЗАПРОПОНОВАНИХ РІШЕНЬ	85
4.1 Опис методики проведення експериментів	85
4.2 Результати експериментів.....	88
4.3 Порівняльний аналіз запропонованих методів із існуючими рішеннями	96
4.4 Кількісний аналіз і статистична оцінка результатів.....	98
4.5 Аналіз обчислювальної ефективності.....	99
4.6 Обговорення результатів експериментів	101
4.7 Висновки до розділу.....	102
ВИСНОВОК	104
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	106
ДОДАТКИ.....	116

ВСТУП

Актуальність. Проблема видалення імпульсного (salt-and-pepper, SaP) шуму набуває підвищеної актуальності в умовах широкого розповсюдження цифрової медичної візуалізації, дистанційної діагностики та інтеграції автоматизованих алгоритмів аналізу зображень у клінічні робочі процеси. Поширені причини виникнення SaP-шуму - апаратні збої сенсорів, артефакти передачі даних - це реальність багатьох медичних закладів. На тлі зростаючих вимог до збереження дрібних анатомічних структур для діагностики навіть поодинокі або кластерні шуми можуть призвести до хибних висновків, тому розробка спеціалізованих, надійних і ресурсно-ефективних методів фільтрації SaP-шуму є критично важливою зараз.

Хоча SeConvNet демонструє ефективність при усуненні окремих імпульсів у стандартних тестах, детальний аналіз показує системні обмеження при роботі з кластерними SaP-плямами. Локальна природа SeConv-блоків демонструє хороші результати для ізольованих шумових пікселів, але втрачає необхідну глобальну перспективу, коли розмір плями перевищує локальний контекст, що призводить до залишкових спотворень та втрати локальної різкості, у результаті чого зменшується контраст дрібних меж і текстур.

Метою дисертаційного дослідження є підвищення ефективності відновлення зображень, пошкоджених імпульсним SaP-шумом, шляхом розробки та верифікації нейромережевої архітектури, що зберігає дрібні анатомічні структури при високих щільностях шуму. Додатково досліджується практична придатність запропонованої архітектури в ресурсно-обмежених середовищах.

Для досягнення мети дослідження було поставлено наступні задачі:

- Проаналізувати та систематизувати існуючі методи відновлення зображень із фокусом на задачі видалення імпульсного SaP-шуму;
- Дослідити моделі та методи підвищення ефективності шумозаглушення зображень шляхом виявлення недоліків базової архітектури SeConvNet, формалізації вимог до її вдосконалення та розробки модифікованої моделі на основі механізму уваги;

- Розробити та реалізувати засоби й механізми підвищення ефективності процесів шумозаглушення зображень на основі нейромережових моделей, включаючи практичне застосування архітектури SeConvNet + CBAM в умовах обмежених обчислювальних ресурсів;
- Провести експериментальні дослідження модифікованої моделі, виконати кількісний та порівняльний аналіз ефективності запропонованих рішень.

Об’єктом дослідження є процес шумозаглушення медичних зображень пошкоджених імпульсним SaP-шумом. Зокрема згорткові нейронні мережі та модулі уваги.

Предметом дослідження є методи й засоби нейромережового шумозаглушення, зокрема модифікації архітектури SeConvNet, а також їхній вплив на якість відновлення (PSNR), обчислювальні витрати та стійкість до різних рівнів SaP-шуму на реальних і синтетичних даних.

Наукова новизна отриманих результатів.

Вперше розроблено метод відновлення медичних зображень, пошкоджених інтенсивним імпульсним шумом, який ґрунтується на гібридному поєднанні селективних згорток та механізму уваги. Відмінною рисою методу є інтеграція блоку канално-просторової уваги (CBAM) безпосередньо у вихід залишкового блоку, що дозволяє адаптивно фокусувати процес відновлення на збереженні структурних ознак та мінімізувати ефект згладжування при реконструкції кластерних плям шуму.

Удосконалено архітектуру нейронної мережі SeConvNet шляхом оптимізації топології з’єднань, а саме - розміщення модулів уваги для обробки сумарного сигналу прямого поширення та корекційної гілки. Це дозволило забезпечити високоточне відтворення дрібних анатомічних деталей (приріст PSNR до 8 % на складних ділянках) без суттєвого збільшення обчислювальної складності, що робить модель придатною для використання в ресурсно-обмежених обчислювальних середовищах.

Особистий внесок здобувача. Усі результати, викладені в дисертації, отримані автором самостійно. Зокрема: проведено аналіз літературних джерел,

розроблено архітектуру моделі SeConvNet + CBAM, реалізовано програмне забезпечення, виконано навчання мережі, проведено експериментальні дослідження та інтерпретацію результатів. У спільних публікаціях автору належать ідеї щодо модифікації блоків уваги та методика проведення експериментів.

Апробація результатів дисертації. Основні положення роботи доповідалися та обговорювалися на конференції:

- Линовський Андрій Олексійович, Мухін Вадим Євгенович. An improved cnn architecture for cluster salt-and-pepper noise filtering in medical data міжнародна науково-практична конференція «інформаційні технології та безпека» (ітб-2025). Київ, 2025

Публікації. За темою дисертації опубліковано 3 наукових праць, з них 3 статті у фахових виданнях України категорії Б та 1 доповідь на конференції.

Структура та обсяг дисертації. Дисертація складається зі вступу, чотирьох розділів, висновків, списку використаних джерел та додатків. Повний обсяг роботи становить 117 сторінок. Робота містить рисунків 24, таблиць 8 та список літератури з 80 найменувань, 1 додаток.

РОЗДІЛ 1. АНАЛІЗ ІСНУЮЧИХ МЕТОДІВ ВІДНОВЛЕННЯ ЗОБРАЖЕНЬ ПІСЛЯ ШУМОВИХ СПОТВОРЕНЬ

1.1. Класифікація методів відновлення зображень

Покращення зображення - набір перетворень, спрямованих на візуальне або прикладне поліпшення якості шляхом корекції контрасту, різкості чи нівелювання дрібних артефактів. Відновлення зображення - задача логічно складніша: відновити наближено до невідомого чистого зображення. Критерій - відповідність еталонному зображенню за кількісними метриками (PSNR [1, 2]) і збереження структурної інформації.

Основні методи покращення зображення:

1. Фільтрація зображень: Застосування фільтрів для покращення якості зображення. Наприклад, для зменшення рівня шуму зображень непогано підходить фільтр Гаусса або медіанний фільтр.
2. Відновлення зображень: Використання алгоритмів для відновлення деталей та усунення дефектів у зображенні. Це може бути відновлення за допомогою алгоритму «деконволюції», алгоритми морфологічного відновлення.
3. Підвищення роздільної здатності: Методи, які спрямовані на покращення роздільної здатності зображень. Наприклад інтерполяція та використання алгоритмів, що базуються на штучних нейронних мережах.
4. Глибинне навчання: Використання нейронних мереж для покращення зображення. Глибинні нейронні мережі можуть бути навчені аналізувати зображення та виконувати завдання, такі як зменшення шуму, відновлення деталей, усунення розмиття тощо.

У науковій літературі існує кілька підходів до класифікації методів відновлення зображень після шумових спотворень. Найпоширеніше групування базується на таких критеріях [3, 4]:

За простором обробки

1. Методи у просторовій області

Ця група методів працює безпосередньо з матрицею пікселів зображення. Операції виконуються над поточним пікселем та його локальним оточенням у межах визначеного вікна чи ядра. Математичний опис таких процесів зводиться до операції згортки у випадку лінійних фільтрів або до процедури ранжування для нелінійних підходів. Прикладом даного підходу є середньозважені фільтри, медіанний фільтр, адаптивні фільтри. Зокрема, медіанний фільтр детально описаний у статті [5]. Перевагою цих підходів є їх простота і швидкість реалізації, а з недоліків часте розмиття деталей і країв. Просторові методи можна розділити на дві основні підгрупи: лінійні і нелінійні

Лінійні фільтри. Основний принцип полягає у заміні значення пікселя зваженою сумою значень його сусідів. Середньозважені фільтри та фільтр Гаусса. Найпоширеніші представники цієї групи. Вони ефективно пригнічують адитивний гаусівський шум шляхом «розмиття» високочастотних коливань. Головною проблемою лінійних фільтрів є нездатність розрізняти шум і високочастотні деталі зображення. Це призводить до «розфокусування» зображення та втрати важливих меж, що є критичним у медичній візуалізації і подальшій діагностиці [6].

Нелінійні фільтри. Використовують складніші алгоритми обробки, які залежать від значень пікселів у вікні, а не тільки від їхнього розташування. Медіанний фільтр. Як зазначено у [3, 5], цей метод замінює значення пікселя медіаною розподілу інтенсивностей у локальному вікні. Він є стандартом для боротьби з імпульсним шумом типу SaP, оскільки ефективно відсіює екстремальні значення, не розмиваючи при цьому чіткі межі об'єктів так сильно, як лінійні фільтри.

Адаптивні фільтри. Наприклад, фільтр Вінера або адаптивний медіанний фільтр. Ці алгоритми аналізують локальну статистику в околі пікселя: якщо область однорідна - застосовується сильніше згладжування, якщо виявлено перепад яскравості - фільтрація послаблюється.

Перевагою просторових методів є їхня обчислювальна простота та висока швидкість реалізації, що дозволяє використовувати їх у режимі реального часу.

Проте їхнім суттєвим недоліком залишається локальність обробки: фільтр «бачить» лише малу частину зображення, ігноруючи глобальні шаблони та повторювані структури, що призводить до втрати текстур.

2. Методи у перетворюючій області

Методи цієї групи базуються на гіпотезі, що корисний сигнал та шум можна розділити, перейшовши в інший математичний базис. Типовий алгоритм обробки передбачає послідовну реалізацію трьох етапів: виконання прямого перетворення, модифікацію отриманих коефіцієнтів шляхом фільтрації або порогової обробки, а також фінальне зворотне перетворення для реконструювання сигналу. Основні підходи включають:

Фільтрація у частотній області. Використовує перетворення Фур'є для розкладання зображення на гармонійні складові. Оскільки шум зазвичай зосереджений у високочастотному спектрі, застосування фільтрів низьких частот дозволяє його відсікти. Однак це часто призводить до появи ефекту «дзвону» навколо різких меж на відновленому зображенні - ефект Гіббса [7].

Вейвлет-перетворення. На відміну від Фур'є, «вейвлети» забезпечують локалізацію як у частотному, так і у просторових областях. Це дозволяє ефективно працювати з нестационарними сигналами, якими є зображення (наявність різких меж та гладких ділянок).

Метод Вейвлет-стиснення описаний у джерелах [3, 8]. Цей підхід використовує властивість розрідженості: у вейвлет-область енергія зображення концентрується у невеликій кількості великих коефіцієнтів, тоді як шум «розмазаний» у вигляді малих значень. Застосування м'якого або жорсткого порогу дозволяє знівелювати шумові коефіцієнти перед зворотним перетворенням.

Методи на основі DCT (Discrete Cosine Transform) [9, 10]. Використовують дискретне косинусне перетворення, яке лежить в основі стандарту JPEG. Обробка часто виконується блоками, що дозволяє враховувати локальні частотні характеристики, але може призводити до блокових артефактів на межах вікон.

Методи у перетворюючій області, зокрема вейвлет-методи, демонструють високу ефективність у збереженні загальної структури зображення та пригніченні

адитивного шуму. Однак вони можуть вносити специфічні артефакти після зворотного перетворення і часто є обчислювально складнішими за просторові фільтри. Крім того, вибір оптимального базису та порогу відсікання часто є нетривіальною задачею, що залежить від типу зображення.

За характером використаної моделі

1. Фільтраційні методи
2. Варіаційні методи та методи оптимізації

Ця група методів розглядає задачу відновлення зображення як задачу мінімізації певної енергетичної функції. Основна ідея полягає у пошуку такого зображення μ , яке є компромісом між відповідністю вихідним зашумленим даним та відповідністю певним апіорним вимогам гладкості. Класичним прикладом є метод тотальної варіації (TV-denoising) [11, 12]. На відміну від квадратичної регуляризації, яка призводить до розмиття країв, функціонал тотальної варіації використовує L-норму градієнта зображення. Це дозволяє ефективно видаляти шум, зберігаючи при цьому чіткі межі об'єктів, оскільки покарання за різкий перепад яскравості у TV-методах є значно меншим, ніж у методах згладжування. Сучасні варіаційні підходи також включають методи вищого порядку Total Generalized Variation (TGV) для уникнення ефекту «сходинок», притаманного класичному TV [13, 14].

3. Методи на основі статистичних моделей

Ця група методів спирається на ймовірнісний підхід до відновлення сигналів, зокрема на Байєсівську теорію. Основна ідея полягає в тому, що відновлене зображення розглядається не як фіксований набір чисел, а як реалізація випадкового процесу. Задача алгоритму зводиться до пошуку такої оцінки «чистого» зображення, яка є найбільш ймовірною за умови наявного зашумленого спостереження. Процес пошуку базується на балансі двох факторів:

- Узгодженість із даними: Відновлене зображення повинно бути статистично схожим на вихідне зашумлене.
- Апіорні знання: Алгоритм використовує попередньо відому інформацію про те, які властивості мають бути притаманні «природним» зображенням.

Наприклад, це може бути вимога локальної гладкості або наявність повторюваних текстур. Класичними представниками цього підходу є методи на основі Марковських випадкових полів, які моделюють залежності між сусідніми пікселями [15, 16].

Іншим потужним напрямком є Байєсівське шумозаглушення у вейвлет-області. Ці методи враховують складну статистику природних зображень, моделюючи коефіцієнти розкладу як суміш розподілів Гаусса. Такий підхід дозволяє ефективно розділяти корисний сигнал і шум навіть у складних умовах [17, 18].

4. Методи на основі словників і розріджених представлень

Цей підхід ґрунтується на гіпотезі, що будь-який фрагмент природного зображення може бути представлений як лінійна комбінація невеликої кількості елементів з деякого надлишкового словника. Математично це записується як $y \approx Da$, де вектор коефіцієнтів a є розрідженим, тобто містить мало ненульових елементів. Ключовим алгоритмом у цьому напрямку є K-SVD [3]. На відміну від використання фіксованих словників, K-SVD адаптивно навчає словник безпосередньо на зашумленому зображенні, що дозволяє краще адаптуватися до локальної структури конкретних даних. Процес відновлення чергує два етапи: розріджене кодування за допомогою алгоритмів типу Orthogonal Matching Pursuit (OMP) та оновлення атомів словника для мінімізації помилки реконструкції [3].

За характером використання інформації про сусідні області

Цей критерій розділяє алгоритми за принципом формування контексту: чи обмежується метод фіксованою геометричною окіл пікселя, чи використовує глобальну надлишковість інформації, притаманну природним зображенням

1. Локальні методи.

Ця група алгоритмів базується на припущенні про високу просторову кореляцію зображення: ймовірність того, що сусідні пікселі мають близькі значення інтенсивності, є високою. Обробка кожного пікселя здійснюється виключно на основі даних з його безпосереднього геометричного околу, зокрема локального вікна чи ядра згортки фіксованої розмірності 3×3 або 7×7 . Математично ці методи часто реалізуються через операцію згортки з ядром, що згладжує, або через

нелінійні рангові операції. Головна гіпотеза полягає в тому, що на малій ділянці зображення сигнал є приблизно постійним, а шум - швидкозмінним [19, 20, 21].

Окремим потужним класом локальних методів є підходи на основі диференціальних рівнянь у частинних похідних. Яскравим прикладом є метод Перона і Маліка [22]. На відміну від звичайного згладжування цей алгоритм аналізує локальний градієнт: якщо градієнт малий, дифузія відбувається активно, усуваючи шум; якщо градієнт великий, процес дифузії зупиняється. Це дозволяє згладжувати шум, зберігаючи гострі краї. Головним недоліком локальних методів є неможливість відрізнити шум від дрібних текстурних деталей, розмір яких співмірний із розміром ядра фільтра. Це призводить до ефекту «розмивання» текстур або появи східчастих артефактів у методах PDE.

2. Нелокальні методи

Ці методи базуються на концепції схожості природних зображень. Гіпотеза полягає в тому, що будь-який малий фрагмент зображення, такий як кут ока, елемент візерунка тканини або рівна ділянка неба, повторюється у зображенні багато разів, причому ці повторення можуть знаходитись на значній відстані одне від одного.

Non-Local Means (NL-means) [23, 24]:

Цей алгоритм став проривом у теорії обробки зображень. Замість усереднення сусідніх пікселів, NL-means обчислює зважене середнє всіх пікселів зображення, де вага залежить не від просторової відстані, а від ступеня схожості околів цих пікселів. Тобто піксель відновлюється на основі схожих на нього пікселів, знайдених у будь-якій частині зображення. Це дозволяє ідеально відновлювати повторювані текстури, які локальні фільтри просто знищують.

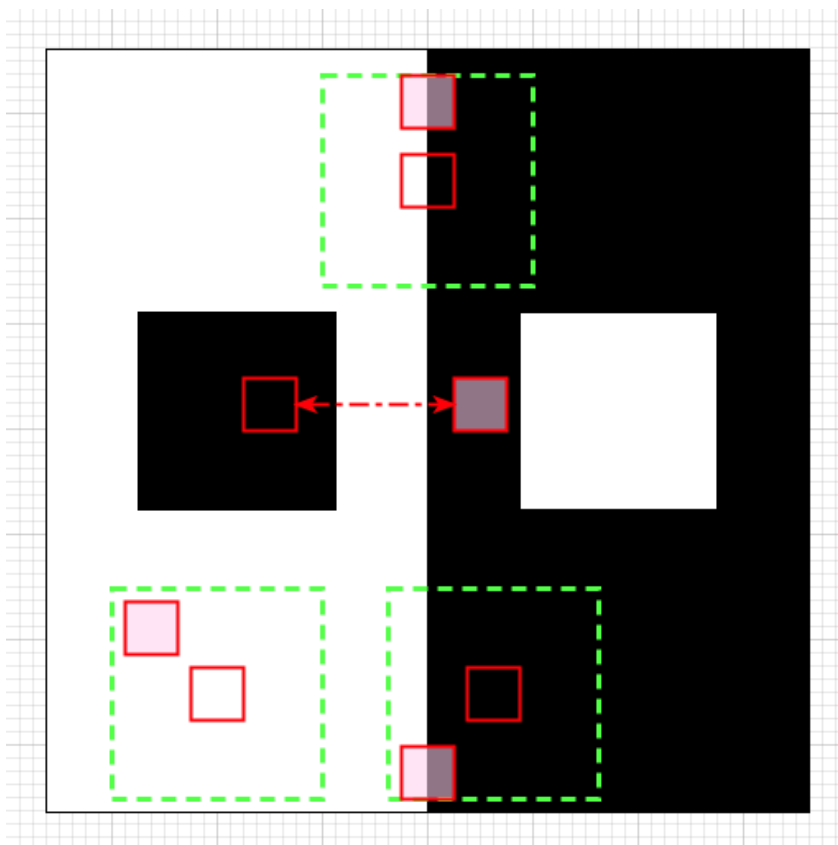


Рис. 1.1. Схема пошуку подібних фрагментів у алгоритмі нелокального усереднення

На цьому зображенні червоні порожні квадрати позначають фрагменти з центрами у конкретних пікселях, що підлягають усуненню шуму. Пунктирні зелені квадрати відповідають вікнам пошуку. Напівпрозорі червоні квадрати є прикладами фрагментів, обраних у межах цих вікон пошуку.

Різні випадки демонструють характерні особливості методу:

- Верхній піксель на вертикальній межі ілюструє перевагу алгоритму нелокального усереднення. Інші пікселі на цій межі матимуть аналогічні чорно-білі фрагменти, що сприятиме отриманню чіткої межі на виході та дозволить уникнути гауссового розмиття.
- Чорний фрагмент у центрі демонструє іншу важливу властивість NL-means: за умови достатнього розміру вікна пошуку алгоритм здатний «перестрибувати» через білу область для пошуку іншої чорної ділянки, що забезпечує якіснішу фільтрацію.

- Нижні фрагменти вказують на можливі недоліки. Для крайнього лівого фрагмента у процесі знешумлення братимуть участь лише білі фрагменти, що має забезпечити високу точність. Однак для крайнього правого фрагмента деякі патчі у вікні пошуку міститимуть незначну білу частину. Ця частина може виявитися недостатньо малою, щоб гарантувати вибір виключно коректних пікселів.

Block-Matching and 3D Filtering (BM3D):

Цей метод вважається «золотим стандартом» серед класичних алгоритмів шумозаглушення. Він розвиває ідею нелокальності, переходячи у спектральну область [3, 25].

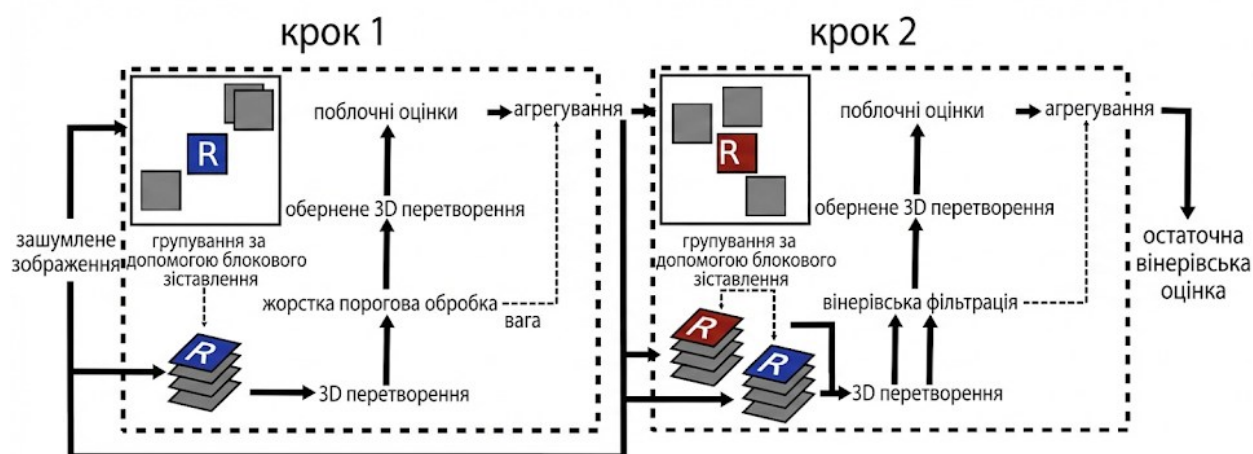


Рис. 1.2. Схема пошуку подібних фрагментів у алгоритмі нелокального усереднення [26]

Алгоритм працює у два етапи:

- Групування (Block Matching): Схожі 2D-фрагменти зображення знаходяться та складаються у тривимірні масиви (3D-групи).
- Спільна фільтрація (Collaborative Filtering): До сформованих 3D-груп застосовується тривимірне перетворення, де сигнал і шум ефективно розділяються. Після порогової обробки виконується зворотне перетворення.

Завдяки використанню кореляції як всередині блоків, так і між ними, BM3D забезпечує безпрецедентну якість збереження деталей при високих рівнях шуму.

Обмеженням є значна обчислювальна складність. Пошук схожих блоків вимагає перебору великої кількості варіантів, що робить повні нелокальні методи

повільними для застосування в реальному часі без використання потужних апаратних прискорювачів (GPU).

В останні роки фокус досліджень змістився від вдосконалення класичних алгоритмів до їх інтеграції з методами глибокого навчання. Основна ідея не локальності трансформувалася у механізми нелокальної уваги (Non-Local Attention), які дозволяють нейронним мережам враховувати глобальні залежності.

За використанням навчання

1. Методи без навчання

Використовують заздалегідь визначені математичні моделі чи фільтри. Приклади: VM3D [3, 25]. Ця група методів базується на аналітичних математичних моделях. Розробник алгоритму явно формулює апріорні припущення про властивості "чистого" зображення і конструює фіксований алгоритм фільтрації або функціонал енергії для оптимізації. Параметри таких методів зазвичай налаштовуються вручну або евристично і не змінюються залежно від вхідних даних після етапу налаштування. Головною перевагою є математична інтерпретованість та відсутність потреби у навчальних даних. Недоліком є обмежена здатність моделювати складні семантичні текстури, які важко описати простими математичними формулами.

2. Методи машинного навчання

Навчаються відновлювати зображення за прикладами. Також поділяються на підгрупи

- класичні ML;

Ці методи займають проміжне місце між аналітичними підходами та глибоким навчанням. Вони зазвичай навчають словники, але не будують глибоких ієрархічних ознак. Навчання словників. Найвідомішим методом є K-SVD, який узагальнює ідею вейвлетів. Замість використання фіксованого базису, алгоритм

навчає переповнений словник безпосередньо на зашумленому зображенні або на базі даних, що дозволяє краще адаптуватися до специфіки конкретних текстур.

Поля рішень: Методи, такі як Cascade of Shrinkage Fields (CSF), навчають параметри фільтрів та функцій активації для максимізації якості відновлення на тестовій вибірці

- Методи глибокого навчання (CNN, RNN, GAN, трансформери).
- Генеративні змагальні мережі (GAN - Generative Adversarial Networks).

В основі підходу лежить змагальна парадигма, що передбачає взаємодію двох нейронних мереж: генератора, відповідального за реконструкцію сигналу, та дискримінатора, завданням якого є розрізнення відновленого та еталонного зображень. Це дозволяє генерувати правдоподібні високочастотні текстури, які були втрачені через шум або низьку роздільну здатність.

Трансформери візуалізації. Адаптація архітектури трансформер для задач комп'ютерного зору. На відміну від CNN, які мають обмежене рецептивне поле, трансформери здатні моделювати глобальні залежності між пікселями на великих відстанях. Це робить їх ідеологічними наступниками нелокальних методів, але з можливістю навчання складних залежностей.

Дифузійні імовірнісні моделі (Denoising Diffusion Probabilistic Models - DDPM [27, 28]). Найновіший клас генеративних моделей, який відновлює дані шляхом ітеративного обернення процесу дифузії. Вони перевершують GAN за стабільністю навчання та різноманітністю генерованих текстур, стаючи новим стандартом якості у задачах відновлення.

За типом шуму, для якого оптимізовані методи

1. Адитивний гаусівський шум - найбільш досліджений випадок [29, 30].
2. Імпульсний шум - часто обробляється медіанними та гібридними методами.

3. Мультиплікативний шум - потребує спеціальних перетворень (логарифмічні моделі, Wavelet + MAP) [31].
4. Складні комбінації шумів - для них найбільш ефективні адаптивні або глибинні моделі [32].

Теорія сигналів та статистичної радіотехніки

Задачу відновлення медичних зображень, пошкоджених імпульсним шумом, також доцільно розглядати через призму фундаментальних положень теорії сигналів та статистичної радіотехніки. Згідно з теоремою відліків (теорема Котельникова-Шеннона), неперервний сигнал з обмеженим спектром може бути точно відновлений за його дискретними відліками. Цифрове медичне зображення є двовимірним дискретним сигналом. В контексті цієї задачі, виникнення імпульсного шуму типу SaP можна інтерпретувати не просто як додавання завади, а як фізичне знищення частини інформативних відліків. Коли піксель набуває екстремального значення внаслідок збою сенсора або помилки передачі, інформація про істинну амплітуду сигналу в цій точці втрачається повністю. Однак, медичні зображення характеризуються високою інформаційною надлишковістю. Сусідні відліки не є незалежними випадковими величинами, а пов'язані певними структурними закономірностями. Саме наявність цієї надлишковості дозволяє ставити задачу відновлення втрачених відліків [33].

З точки зору статистичної радіотехніки, процес шумозаглушення даних роботі розглядається як задача виявлення та оцінювання корисного сигналу на тлі адитивної імпульсної завади. Класичні підходи, такі як медіанна фільтрація, фактично реалізують найпростіші статистичні оцінювачі. Наприклад, оцінку математичного сподівання або медіани розподілу інтенсивностей у ковзаючому вікні. Головним недоліком таких методів є використання фіксованих, апріорно заданих статистичних моделей, які не враховують складну семантичну структуру сигналу. Запропонований у дисертаційній роботі підхід на основі нейронної мережі SeConvNet із модулем уваги CBAM виступає в ролі адаптивного оптимального

детектора. На відміну від класичних фільтрів, нейронна мережа в процесі навчання формує складну багатовимірну модель апіорного розподілу ймовірностей корисного сигналу.

В цьому контексті запропонована архітектура виконує дві функції:

- SeConv-блоки діють як механізм селективної реконструкції відліків, ігноруючи завідомо хибні значення і використовуючи для інтерполяції лише достовірні дані, що відповідає збереженню умов теореми відліків на локальному рівні.
- Модуль уваги (CBAM) виконує функцію узгодженої фільтрації у просторі ознак, максимізуючи відношення сигнал/шум шляхом зважування каналів та просторових областей відповідно до їх інформативності.

Таким чином, використання глибокого навчання дозволяє перейти від простих статистичних оцінок до побудови оптимального нелінійного оцінювача, який відновлює сигнал, базуючись на прихованих статистичних закономірностях розподілу анатомічних структур, вивчених на великій вибірці даних.

1.2. Аналіз сучасних методів шумозаглушення та відновлення зображень

Огляд існуючих моделей для видалення шуму солі та перцю (SaP) включає дослідження різноманітних підходів і методів, які використовуються для ефективної обробки пошкоджених зображень. Раннє покоління алгоритмів обробки зображень включає лінійні фільтри, які ефективні проти невеликого гаусівського шуму, проте згладжують деталі зображення. Нелінійні методи - наприклад, медіанний фільтр - краще справляються з імпульсним шумом, що є випадковими білими та чорними пікселями. Білінійний фільтр застосовує вагове усереднення з урахуванням схожості за яскравістю і просторовому положенню, хоча такий підхід може бути обчислювально затратним при великому радіусі ядра.

Значний прогрес в класичних методах досягнуто за допомогою non-local підходів: наприклад, алгоритм NL-Means шукає подібні блоки по всьому

зображенню, а модель VM3D об'єднує подібні фрагменти у 3D-блоки для кодування і фільтрації шуму [3]. Ці методи ефективно захищають текстури і структури, проте можуть призводити до розмиття та мають високу складність.

Останнім часом провідне місце займають методи на основі глибокого навчання, які демонструють якісно вищі показники придушення шуму. Наприклад, мережі з нелокальними блоками здатні захоплювати глобальні залежності в зображенні: не обмежуючись локальною околицею, вони збирають інформацію з усіх позицій і моделюють контекст. Це суттєво покращує якість обробки, дозволяючи відновлювати зв'язки між віддаленими ділянками зображення. Водночас комбінування рекурентних компонентів (RNN) із нелокальними блоками забезпечує моделі запам'ятовувати довгострокові залежності і поступово інтегрувати контекст на багатьох кроках обробки [34, 35, 36].

Згорткова нейронна мережа

Згорткові нейронні мережі (CNN) широко використовуються у процесі відновлення зображень і є найбільш популярною нейронною мережею для вирішення проблеми покращення зображення. Переважна більшість дослідників беруть її за основу [37]. Перші CNN використовувалися для усунення шуму в зображеннях. Значну роль в цьому відіграв багатошаровий перцептрон для усунення шумів.

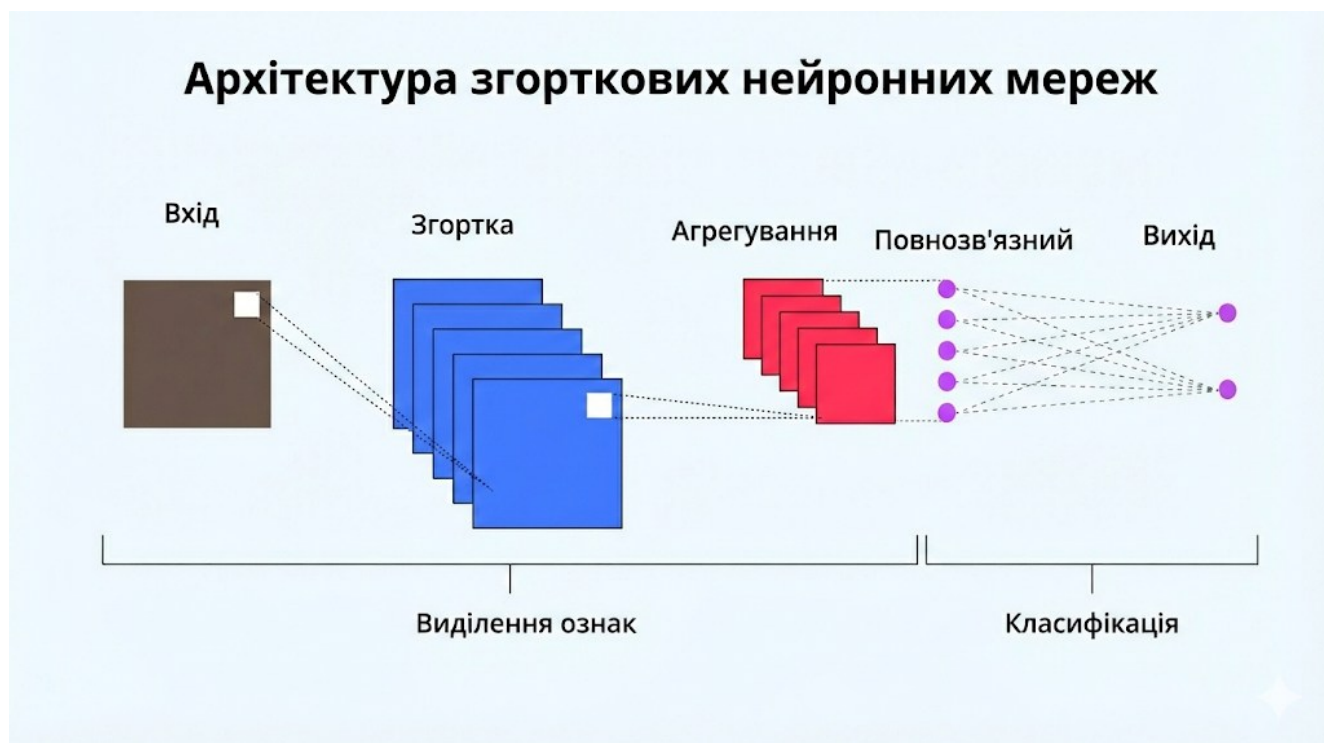


Рис. 1.3. Архітектура згорткових нейронних мереж

MLP був навчений безпосередньо співвідносити пошкоджені ділянки зображення з чистими ділянками за допомогою великого набору даних. Наступним кроком було застосування CNN для підвищення роздільної здатності зображень із низькою роздільною здатністю. Проривною роботою стала SRCNN. Ключова ідея SRCNN полягає в тому, щоб навчити мережу за допомогою великого набору зображень пар зображень із низькою та високою роздільною здатністю. Під час процесу навчання мережа вчиться оптимізувати свої параметри, щоб мінімізувати різницю між згенерованим зображенням високої роздільної здатності та відповідним вхідним зображенням. Використовуючи потужність глибоких згорткових нейронних мереж, SRCNN досягає вражаючих результатів.

Ще однією сферою використання CNN є усунення шуму у зображень, перетворення розмитих зображень у більш чіткі. Було запропоновано кілька методів на основі CNN для обробки різних типів розмиття. Одним з відомих підходів є DeblurNet. Принцип даного підходу є обробка вхідних кадрів, які розташовані поруч, через кілька згорткових шарів, доки не буде виведено розмитий центральний кадр. Логічним розвитком в напрямку покращення зображень стала DnCNN-SR [38].

Автор об'єднав підходи по зменшенню шуму і збільшенню роздільної здатності зображення. Цей спільний підхід покращив якість відновлення за рахунок одночасного зменшення шуму та підвищення роздільної здатності зображення. Загалом, CNN відіграли вирішальну роль у вдосконаленні методів відновлення зображень. Вони продемонстрували надзвичайні можливості у видаленні шумів, надвисокій роздільній здатності, видаленні розмиття та інших завданнях відновлення [39].

Рекурентна нейронна мережа

Цікавим підходом є використання рекурентних нейронних мереж [40, 41, 42] у сфері обробки зображень. Автор даної статті використав рекурентну нейронну мережу і додав нелокальну операцію [23] для покращення роботи нейронної мережі.

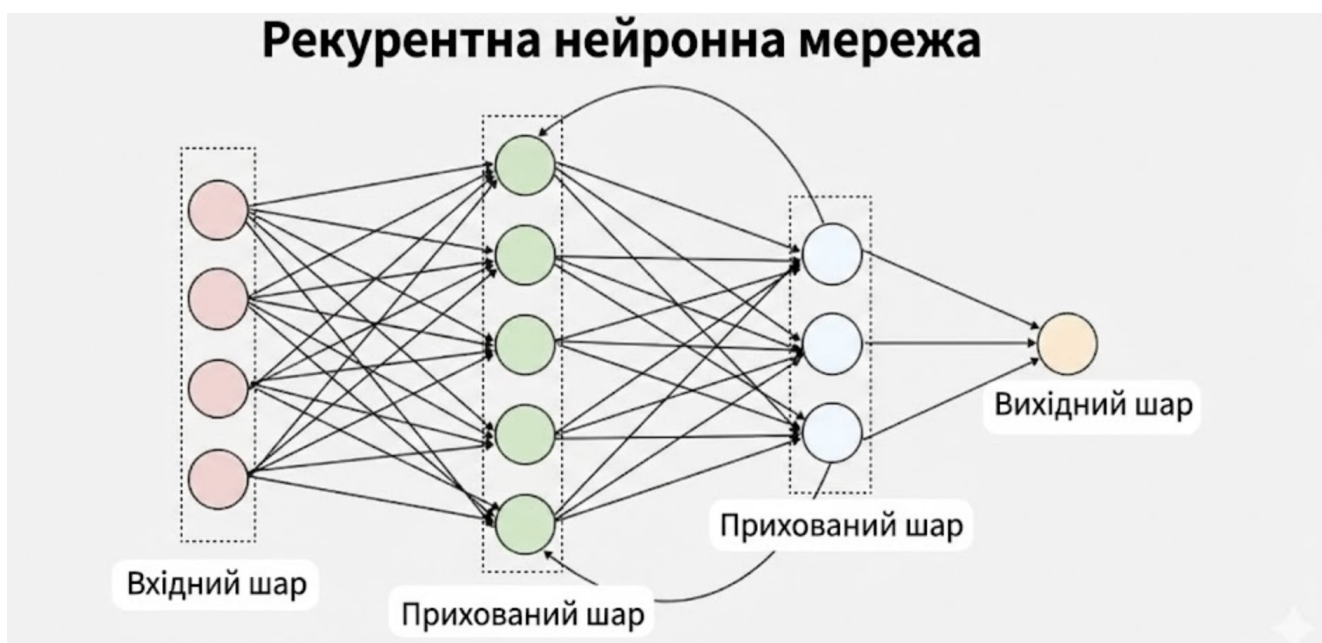


Рис. 1.4. Архітектура рекурентних нейронних мереж

Рекурентна нейронна мережа є потужним інструментом для обробки зображень з кількома перевагами, а в поєднанні з нелокальною операцією результат роботи покращується ще в рази. Поєднання рекурентної нейронної мережі з нелокальною операцією відкриває нові можливості для обробки зображень. Ось декілька з них:

1. Врахування довгострокових залежностей: РНМ дозволяє моделі збирати і використовувати довгострокові залежності в зображеннях, враховуючи

контекст та зв'язки між об'єктами. Додавання нелокальної операції дозволяє збирати інформацію з усіх позицій зображення, не обмежуючись лише локальним оточенням. Це допомагає моделі краще розуміти зв'язки між різними частинами зображення і забезпечує кращу якість обробки.

2. Здатність працювати зі змінною довжиною вхідних зображень: РНМ може адаптуватись до зображень різних розмірів, що дозволяє обробляти великі або малі зображення без втрати якості. Поєднання з нелокальною операцією [23] дозволяє моделі збирати інформацію з усіх частин зображення, незалежно від їх розміру. Це особливо корисно, коли ми маємо справу з великими зображеннями або коли розміри об'єктів на зображенні різні.
3. Загальна краща якість обробки: Поєднання рекурентних нейронних мереж з нелокальною операцією може призвести до поліпшення загальної якості обробки зображень. Це пояснюється тим, що нелокальна операція дозволяє моделі враховувати ширший контекст і залежності в зображенні, тоді як РНМ забезпечує збереження і використання цієї інформації на кожному кроці обробки.

NAMF: Нелокальний адаптивний середній фільтр для видалення шуму солі та перцю (SaP) [43, 44, 45]

Модель NAMF пропонує нелокальний адаптивний середній фільтр для ефективного видалення SaP шуму. Основна перевага цієї моделі полягає в застосуванні методу на основі відношення для розрізнення зашумлених і текстурних пікселів під час виявлення SaP шуму, що допомагає уникнути обробки зашумлених пікселів як текстурних. Крім того, покращений метод нелокального середнього, заснований на характеристиках SaP шуму, дозволяє подальше відновлення зашумлених зображень і покращення їх якості. Однак слабкими сторонами моделі можуть бути складність налаштування порогового значення для визначення зашумлених пікселів та обмежена ефективність у випадках високих рівнів SaP шуму.

CNNMedianLayersSAP. Згортова нейронна мережа з медіанними шарами для видалення шуму солі та перцю [46, 47, 48]. Метод використовує медіанні

фільтри в глибоких нейронних мережах для видалення шуму солі та перцю. Основна перевага цієї моделі полягає в збереженні здатності до видалення пікселів-шумів, але спроби відновлення деформацій, внесених медіанним фільтром, можуть призводити до втрати якості зображення. Модель показує добрі результати у видаленні шумів на початкових етапах, але може зазнавати обмежень у відновленні високозашумлених зображень.

Ще одним цікавим підходом є модулі уваги і нелокальний блок [49, 50, 51]. У більшості звичайних згорткових мереж кожен фільтр обробляє лише невелику частину зображення. Через це мережі добре працюють із дрібними деталями, але їм важко враховувати зв'язки між далекими ділянками картинки.

Щоб вирішити цю проблему, застосовують:

- нелокальні блоки - вони дозволяють моделі поєднувати будь-які дві точки зображення, навіть якщо вони далеко одна від одної. Наприклад, якщо на зображенні є повторювані візерунки, такий блок допомагає використовувати інформацію з усіх подібних місць для кращого відновлення [23, 52].
- модулі уваги [53, 54, 55] - вони виокремлюють найбільш важливі частини даних. СВМ робить це у двох вимірах:
 - по каналах (які ознаки є важливішими);
 - у просторі (які ділянки зображення заслуговують більшої уваги).

Такі механізми допомагають мережі краще відрізнити корисні структури від шуму. У результаті картинка стає чіткішою: зберігаються важливі деталі, а шумові точки послаблюються.

В таблиці 1.1 наведено порівняльну характеристику рекурентних, згорткових нейронних мереж і багатошарового перцептрон за найважливішими для дослідження характеристиками. Згідно таблиці згорткові і рекурентні нейронні мережі здатні добре справлятися з задачею покращення якості зображень.

Таблиця 1.1. Порівняльна характеристика рекурентної, згорткової та багатошарової нейронних мереж

	Рекурентна нейронна мережа	Згорткова нейронна мережа	Багатошаровий перцептрон
Пам'ять	Використання пам'яті для обробки послідовних даних	Немає явної пам'яті минулих входжень	Немає пам'яті, кожен вхід обробляється незалежно
Контекстна інформація	Фіксує довгострокові залежності та контекст	Обмежене поле сприйняття та локальний контекст	Відсутність явного моделювання контексту або залежностей
Ефективність усунення шумів	Захоплює глобальні залежності, обробляє складні шумові шаблони	Ефективно зменшує шум на локальних ділянках	Обмежена здатність обробляти складні шумові шаблони

Перевагою рекурентних мереж над згортковими є наявність в пам'яті даних попередніх обчислень. За рахунок цього досягається більша ефективність у знаходженні довгострокових залежностей, що дуже підвищує якість покращення складних зображень. В таблиці 1.2 наведено результати роботи різних методів покращення зображень з різним рівнем шуму на одному набору зображень BSD68 [56]

Таблиця 1.2. Результат роботи різних методів покращення зображень на основі набір зображень BSD68, з різними рівнями шуму(15, 25 і 50)

Метод	15	25	50
DnCNN	31.72	29.23	26.23
ELDRN	32.11	29.68	26.76
MWCNN	31.86	29.41	26.53
BM3D-Net	31.42	28.83	25.73
FFDNet	31.62	29.19	26.30
BRDNet	31.79	29.29	26.36
NLRN	31.88	29.41	26.47
NNC	31.49	28.88	25.25

В таблиці 1.2. виділено три методи з найкращим результатом до яких входить досліджуваний підхід NLRN. Він демонструє хороші результати роботи адже специфіка рекурентних нейронних мереж в поєднанні з нелокальними операціями добре підходить для покращення зображень і має суттєвий потенціал для покращення.

У підсумку спостерігається очевидна тенденція переходу від класичних фільтраційних і перетворювальних методів до підходів на основі глибинного навчання. Це пов'язано з тим, що нейронні мережі здатні навчатися складним нелінійним відношенням між зашумленими та чистими зображеннями, ефективно інтегрувати нелокальний контекст та комбінувати кілька джерел інформації, що дає кращі якісні і кількісні результати на різних типах шуму. Водночас такі моделі вимагають значних обчислювальних ресурсів і репрезентативних тренувальних даних, а їхня поведінка на реальних даних нерідко потребує додаткової перевірки і адаптації.

1.3. Методи глибокого навчання для усунення шуму та відновлення зображень

У методах глибокого навчання для шумозаглушення та відновлення зображень ключова ідея - навчитись відображенню між зашумленим і чистим зображенням. Замість ручного налаштування фільтрів, мережа вчиться сама виділяти шаблони шуму і корисні структури через багаторівневе представлення.

Методи можна поділити за такими ознаками:

- чи навчається модель на парах зашумлене $\rightarrow$ чисте;
- чи без чистих зразків [57];
- який тип шуму - гаусівський, імпульсний, реальний, тощо;
- чи включають мережі додаткові компоненти: залишкові зв'язки, увагу, нелокальні блоки, рекурентність, або комбіновані гібридні архітектури.

За останнє десятиріччя глибоке навчання принципово змінило підхід до усунення шуму: навчені нейронні мережі, зокрема згорткові, стали домінувати у завданні шумозаглушення [34]. Базовими прикладами є DnCNN [58] та FFDNet, де використовуються технології залишкового навчання і пакетна нормалізація для прискорення тренування й поліпшення результатів [3]. Ці мережі можуть ефективно усувати гаусівський шум, часто перевершуючи класичні методи.

Також для навчання застосовують безконтрольні підходи. Наприклад, Noise2Noise навчає модель без явних чистих зображень - тільки на різних шумних спостереженнях одного й того ж зображення. Такий підхід уможливлює навчання тоді, коли чисті дані недоступні, проте вимагає складних обчислень і не завжди стабільно працює для складних видів шуму.

У сучасній літературі виділяють кілька ключових напрямів побудови нейромережевих моделей для шумозаглушення, кожен з яких має свої архітектурні особливості та сферу застосування:

1. DnCNN. Цей метод базується на глибокій згортковій мережі із використанням залишкового навчання та пакетної нормалізації. Архітектура є стандартом для видалення гаусівського шуму та деяких видів компресійних артефактів.

- *Переваги:* Забезпечує високі показники PSNR, швидку збіжність та є універсальною для різних рівнів шуму.
- *Обмеження:* Демонструє меншу ефективність при обробці складного або нестандартного шуму [58].

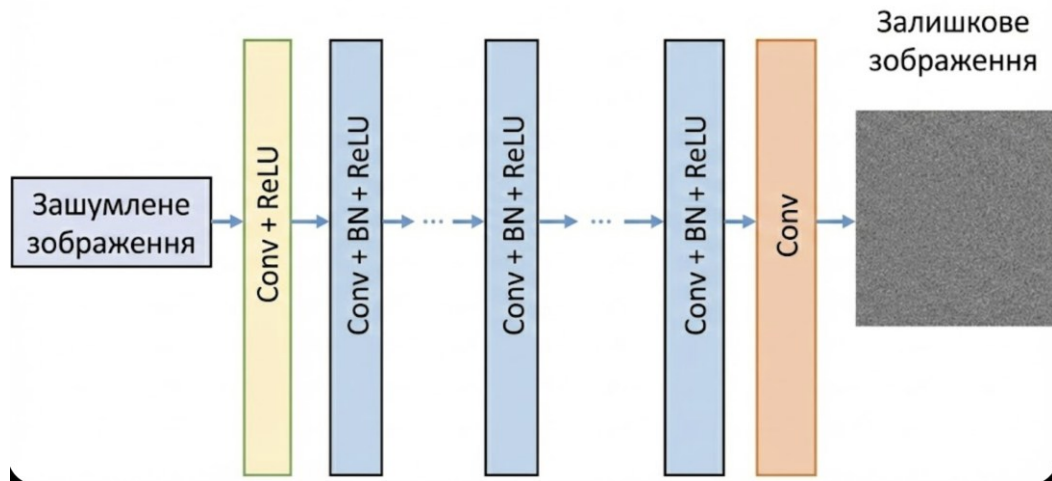


Рис. 1.5. Архітектура DnCNN

2. Universal Denoising Networks. Підхід поєднує згорткові мережі із нелокальною фільтрацією. Застосовується для адитивного білого гаусівського шуму та випадків із невідомим рівнем шуму.

- *Переваги:* Дозволяє використовувати одну універсальну модель для широкого діапазону рівнів шуму, а включення нелокальних ознак покращує відновлення дрібних деталей.
- *Обмеження:* Характеризується ускладненою архітектурою та підвищеними вимогами до обчислювальних ресурсів через використання нелокальних компонентів [59].

3. TNLRD (Trainable Non-Local Reaction Diffusion). Метод реалізує поєднання реакційно-дифузійного підходу з навчальними нелокальними фільтрами. Орієнтований на усунення сильних рівнів шуму.

- *Переваги:* Забезпечує кращу візуальну якість, зменшення артефактів на рівномірних ділянках та ефективніше збереження текстур порівняно з локальними моделями.
 - *Обмеження:* Має високу складність процесу навчання [60].
4. Мережеві підходи на основі GAN (Adversarial Learning). Використовують архітектуру «генератор-декодер» та комбіновані функції втрат, що поєднують піксельні метрики (L1/L2) та змагальні втрати. Застосовуються для шумів, які важко описати класичними розподілами, та у випадках, де пріоритетом є візуальна правдоподібність.
- *Переваги:* Зображення стають менш розмитими та мають більш приємну для сприйняття візуальну якість, що часто відрізняється від традиційних методів.
 - *Обмеження:* Існує ризик генерації деталей, яких не було в оригіналі, а також проблема нестабільності процесу навчання [61].
5. Ефективні моделі з універсальністю (Efficient Models). Ця група методів спрямована на досягнення балансу між точністю відновлення та швидкістю роботи при меншій кількості параметрів. Вони актуальні, коли ресурси обмежені або потрібна робота в реальному часі.
- *Переваги:* Є менш ресурсномісткими та гнучкими у застосуванні, здатні працювати на різних рівнях шуму.
 - *Обмеження:* Можуть бути менш гнучкими для складних типів шуму або реальних знімків із специфічним сенсорним шумом [62].

Архітектурні вдосконалення

Для покращення якості й узагальненості, багато моделей включають:

1. Залишкові зв'язки - дозволяють мережі легко зберігати початкову інформацію й уникати проблеми зникання градієнта, а також поліпшують деталізацію. (DnCNN) [58]

2. Нелокальні операції / самоподібність - ці компоненти дозволяють моделі використати інформацію від розташованих у різних частинах зображення, але за формою подібні ділянок. Це зменшує розмиття і допомагає відновити повторювані структури [59].
3. Комбінації втрат - не тільки L2 / L1, але також на основі ознак попередньо навчених мереж, або SSIM-орієнтовані втрати. Це допомагає моделі виробляти візуально кращі результати, не лише високий PSNR [61].

За останнє десятиліття відбувся явний перехід від класичних фільтраційних і перетворювальних методів до підходів на основі глибокого навчання. Нейронні мережі забезпечують кращі результати завдяки здатності визначати складні нелінійні відношення та інтегрувати глобальний контекст. Водночас сучасні рішення повинні балансувати між якістю відновлення, обчислювальними витратами й наявністю навчальних даних.

1.4 Основні проблеми та виклики в існуючих підходах до відновлення зображень

Розмаїття шумів

Одним із найбільш складних викликів у задачі відновлення зображень є неоднорідність і різноманітність типів шуму, які виникають у реальних умовах. Більшість глибинних моделей створюються та тестуються на синтетичних даних із адитивним білим гаусівським шумом, оскільки цей тип шуму легко моделювати та математично описати. Такі умови дозволяють отримати високі результати в метриках якості. Однак у реальних прикладних сценаріях шум має іншу природу і набагато складніше піддається математичному опису.

Адитивний білий гаусівський шум .

Найбільш поширений у дослідженнях, моделюється випадковим процесом із нормальним розподілом. Глибокі згорткові мережі демонструють найвищу ефективність саме для цього типу шуму [58].

Імпульсний шум.

Виникає через збої в каналах передачі або помилки сенсорів. Характеризується наявністю пікселів, що приймають мінімальне або максимальне значення. Для SaP-шуму традиційні фільтри забезпечують базовий рівень відновлення, але навіть сучасні CNN-моделі, які добре працюють з гаусівським шумом, часто втрачають ефективність. При високій щільності імпульсного шуму зберегти структуру зображення стає надзвичайно складно [62]. Імпульсний шум у медичних зображеннях проявляється як випадкові або кластерні пікселі з екстремальними інтенсивностями. Хоча в контрольованих дослідженнях SaP часто моделюється як незалежний випадковий процес, у реальній клінічній практиці причини та шаблони виникнення цього шуму різноманітні - від одиничних апаратних дефектів до пакетних помилок при передачі даних. Однією з найчастіших причин є апаратні дефекти сенсорів або елементів ланцюга зчитування. Причиною може бути дефект окремого каналу чи елемента приймача, локальний збій електроніки. Це може викликати поодинокі пошкоджені пікселі; невеликі кластерні плями.

Корельований та структурований шум.

Це шуми, які мають залежність між сусідніми пікселями - наприклад, смугастий шум або шум від матриць зчитування сенсорів. Такі шуми значно складніші, оскільки вони не є незалежними випадковими величинами. Стандартні CNN-моделі, які розраховані на локальні шаблони, часто не можуть ефективно їх придушити [63].

Шум сенсорів.

У цифрових камерах шуми залежать від багатьох факторів: інтенсивності освітлення, ISO, часу експозиції, якості матриці. Такий шум є комбінацією фотонного шуму, шуму зчитування та електронних спотворень. Відновлення зображень у цьому випадку є особливо складним, бо статистика шуму не є сталою і змінюється від кадру до кадру [64].

Обчислювальна вартість

Також відомими проблемами є висока обчислювальна вартість та складність навчання глибоких мереж. Багатошарові архітектури потребують значних ресурсів для тренування, особливо при великих масштабах вхідних зображень. Практична

швидкість роботи часто потребує пожертвувати глибиною чи просторовим охопленням мережі. Крім того, глибинні моделі - це «чорні ящики», що ускладнює інтерпретацію результатів і адаптацію до нових умов без перекваліфікації [65]. Глибокі архітектури містять десятки або сотні шарів, що призводить до мільйонів параметрів. Для їх навчання потрібні потужні графічні процесори (GPU) [66]. Навчання на великих зображеннях різко підвищує споживання пам'яті. Наприклад, обробка зображень роздільності 4K може перевищувати доступний обсяг відеопам'яті навіть на сучасних GPU.

Навчання сучасних моделей може тривати від кількох годин до кількох тижнів. Це ускладнює експерименти та знижує гнучкість у дослідженнях [29].

Навчання і обмеження набору зображень

Останній важливий виклик - робота з реальними даними. Мережі, натреновані на синтетичних даних або обмеженому наборі сцен, можуть не забезпечувати такої ж якості при складних реальних умовах. Вирішення цих завдань вимагає подальшого розвитку адаптивних, які самі навчаються або гібридних підходів. Більшість сучасних моделей тренуються на синтетично доданому шумі, а у реальних фото шум формується складніше: він залежить від матриці сенсора, умов освітлення, обробка результатів. У результаті модель, яка добре працює на "лабораторному" шумі, може бути непридатною для реальних даних. Якщо мережу навчають на кількох популярних наборах, вона добре справляється саме з подібними зображеннями. Але у випадку з медичними знімками, аерофотозйомкою або камерами з поганим об'єктивом результати значно гірші.

Ризики галюцинацій

Останнім часом набули популярності генеративно-змагальні мережі (GAN) та дифузійні моделі, здатні створювати ідеально гладкі, візуально привабливі знімки.

Проблема полягає в тому, що ці моделі працюють за імовірнісним принципом: вони намагаються "вгадати", як мала б виглядати втрачена ділянка, і генерують нові

пікселі, спираючись на свій навчальний досвід. У медичній візуалізації це призводить до катастрофічного явища - "галюцинацій". Мережа може синтезувати анатомічно правдоподібні деталі, яких насправді немає в тілі пацієнта, наприклад, домалювати структуру здорової тканини там, де є ураження, або, навпаки, згенерувати тінь, що нагадує пухлину.

Враховуючи високу ціну лікарської помилки, застосування подібних імовірнісних моделей у клінічній практиці є вкрай небезпечним. Саме тому в даній дисертаційній роботі розроблено принципово інший підхід. Архітектура на базі згорткової мережі SeConvNet з механізмом уваги CBAM є детермінованою. Вона не генерує зображення з нуля, а обчислює математичну корекцію виключно для пошкоджених пікселів на основі реального сусіднього контексту. Такий підхід повністю нівелює ризик виникнення клінічно небезпечних "галюцинацій"

Таблиця 1.3. описує ключові проблеми шумозаглушення медичних зображень охоплюють технічні, ресурсні та нормативні аспекти. Складна природа реальних шумів і дефіцит репрезентативних даних вимагають застосування гібридних наборів та симуляції реалістичних умов, тоді як висока обчислювальна вартість обробки великих зображень спонукає до розробки більш ефективних архітектур та методів економії пам'яті. Разом з тим, впровадження цих технологій ускладнюється непрозорістю алгоритмів та юридичними ризиками, що робить необхідними ретельну клінічну перевірку, сертифікацію ПЗ та валідацію результатів для забезпечення довіри до діагностики.

Таблиця 1.3. Ключові проблеми шумозаглушення медичних зображень, їх причини та можливі шляхи розв'язання

Проблема	Причина	Потенційне рішення / практичні кроки
Різноманіття шумів (AWGN, SaP, корельований, sensor noise)	Реальний шум має складну, змінну статистику	Використовувати гібридні набори зображень тестування на різних типах шуму
Нестача репрезентативних наборів зображень	Труднощі з отриманням чистих еталонів, відмінність між наборами зображень	симуляція реалістичного шуму, збір клінічних даних
Висока обчислювальна вартість і пам'ять [32]	Глибокі мережі, великі зображення (4K), дорогі non-local блоки	збереження проміжних градієнтів, впровадженню ефективних архітектур
Інтерпретованість і перевірка результатів	Чорний ящик мереж, ризик прихованих помилок	перевірка на підгрупах
Юридичні / етичні ризики [67]	Відсутність задокументованої клінічної перевірки;	Клінічні випробування, реєстрація/сертифікація ПЗ, прозорі звіти про обмеження

Аналіз наслідків появи шуму в медичних зображеннях

Головна проблема шуму типу SaP полягає в тому, що він за своїм візуальним розміром та високим рівнем контрастності надзвичайно схожий на ранні маркери небезпечних захворювань. Найбільш критично ця проблема проявляється у сфері цифрової мамографії.

Рання карцинома часто не проявляється у вигляді великих пухлин, а проявляє себе через мікрокальцинати - крихітні відкладення солей кальцію, розмір яких варіюється від 0.1 до 0.5 міліметра. На рентгенівському знімку ці утворення виглядають як крихітні, яскраві білі крапки на тлі сірої залозистої тканини. Білі пікселі, згенеровані апаратним збоєм сенсора або помилкою передачі даних, виглядають практично ідентично до цих життєво важливих маркерів.

Ця подібність створює подвійну клінічну загрозу:

- **Хибно позитивні висновки:** Скупчення шумових пікселів може бути легко сприйняте лікарем або комп'ютерною системою як реальне скупчення мікрокальцинатів. Це призводить до встановлення хибного діагнозу, призначення непотрібних, болючих та інвазійних процедур, а також викликає значний психологічний стрес у пацієнта.
- **Хибно негативні висновки:** Ще більш небезпечною є ситуація, коли знімок сильно зашумлений. У такому середовищі справжні мікрокальцинати просто губляться серед тисяч апаратних шумових пікселів, зливаючись із фоном. Проблема загострюється у пацієнтів зі щільною тканиною молочної залози, де контраст між патологією та здоровою тканиною і так є мінімальним. Радіолог може розцінити реальну патологію як частину загального зашумлення і пропустити її, що призведе до затримки лікування онкології та суттєвого погіршення прогнозу.

Аналогічні виклики існують і в рентгенографії органів грудної клітки. Під час скринінгу часто шукають легеневі вузлики низького контрасту, які можуть бути ранньою ознакою раку легень. Ці вузлики важко ідентифікувати через так званий структурований анатомічний шум - накладання тіней від ребер, серця та великих

судин. Коли поверх цього анатомічного шуму накладається ще й імпульсний апаратний шум, дрібні деталі повністю руйнуються.

1.5. Висновки до розділу

У першому розділі проведено комплексний аналіз та систематизацію існуючих методів відновлення зображень після шумових спотворень. Еволюція підходів демонструє чіткий перехід від класичних просторових і перетворювальних методів до сучасних алгоритмів на основі машинного та глибинного навчання. Традиційні лінійні та нелінійні фільтри, а також складніші нелокальні алгоритми, стикаються із серйозними обмеженнями: вони часто призводять до надмірного згладжування контурів та втрати дрібних деталей або вимагають надмірних обчислювальних ресурсів, особливо при високій щільності імпульсного шуму типу «сіль і перець» (SaP).

Застосування нейромережових архітектур, зокрема згорткових та рекурентних мереж із використанням нелокальних блоків, механізмів уваги й залишкових зв'язків, забезпечує значно вищу якість відновлення. Це досягається завдяки здатності моделей виявляти складні нелінійні закономірності та ефективно інтегрувати глобальний контекст зображення. Проте широке впровадження методів глибокого навчання супроводжується низкою критичних викликів: високою обчислювальною вартістю, значними витратами пам'яті та проблемою узагальненості на реальних даних. Моделі, натреновані переважно на синтетичному адитивному білому гаусівському шумі, часто втрачають свою ефективність під час обробки реальних клінічних знімків зі складними типами зашумлення.

Виявлені обмеження вказують на гостру необхідність розробки адаптивних, більш інтерпретованих і менш ресурсомістких рішень. Проведений аналіз підтверджує, що для успішного шумозаглушення медичних зображень необхідний баланс між локальною інтерполяцією пошкоджених ділянок та глобальною селекцією релевантних ознак. Це обґрунтовує доцільність розробки архітектури,

яка поєднує селективні локальні згортки з канално-просторовими механізмами уваги. Зроблені висновки закладають міцну основу для побудови гібридної моделі SeConvNet + CBAM, що буде детально описана та досліджена у наступних розділах дисертаційної роботи.

Проведений аналіз став підґрунтям для вибору архітектури SeConvNet + CBAM, описаної у наступних розділах роботи: поєднання селективних локальних згорток і канално-просторових механізмів уваги обґрунтовано як компроміс між локальною інтерполяцією і глобальною селекцією релевантних ознак. На основі висновків цього розділу у подальших експериментах передбачено серію практичних дослідів. Зроблені висновки закладають основу для побудови моделей, що будуть описані у наступному розділі

РОЗДІЛ 2. МОДЕЛІ ТА МЕТОДИ ПІДВИЩЕННЯ ЕФЕКТИВНОСТІ ВІДНОВЛЕННЯ ЗОБРАЖЕНЬ

2.1. Основні недоліки існуючої моделі згорткової нейронної мережі для видалення шуму SeConvNet

SeConvNet [68] демонструє високу ефективність при усуненні імпульсного шуму в типових тестових сценаріях, проте при детальному аналізі її поведінка на медичних томографічних даних виявляє низку системних обмежень. Одним із системних обмежень SeConvNet є зниження якості відновлення на ділянках, де SaP шум утворює великі кластерні плями.

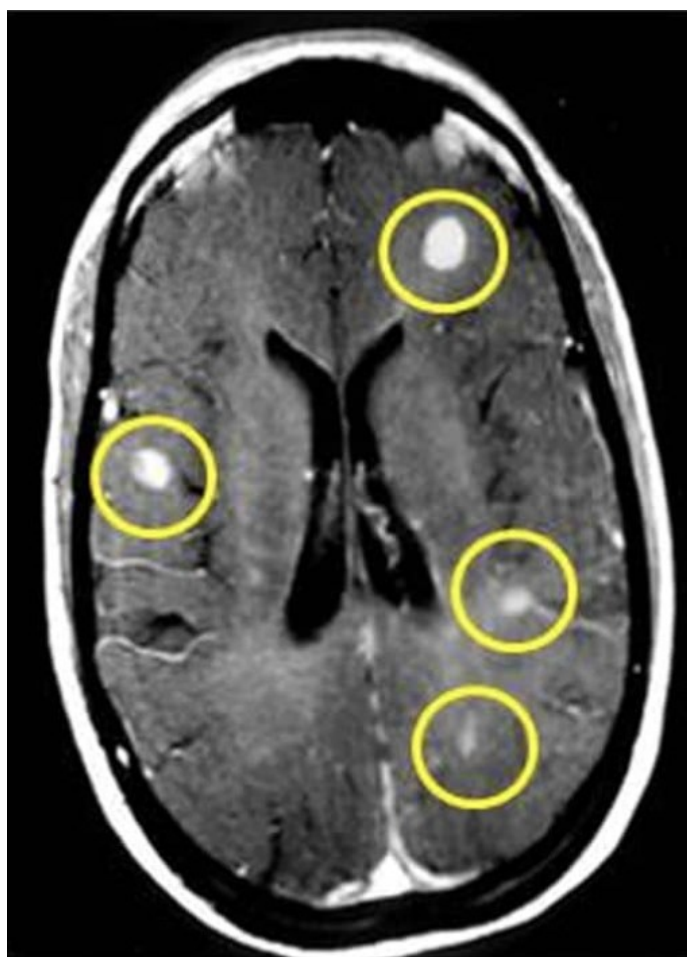


Рис. 2.1. Зображення з кластерними плямами

У таких випадках локальна інформація всередині плями часто відсутня або істотно спотворена, що призводить до двох взаємопов'язаних ефектів: по-перше, залишкові спотворення, коли мережа не відновлює внутрішню неоднорідність

плями; по-друге, надмірне згладжування, коли модель виробляє усереднену апроксимацію і втрачає локальні контрасти та дрібні межі. Саме це явище є ключовим недоліком для застосувань у рентгенівських знімках, де клінічна інтерпретація часто залежить від здатності зберегти дрібні локальні ознаки та межі анатомічних структур.

Архітектурний корінь цього недоліку лежить у природі SeConv-блоків і підходу їхнього застосування: ці блоки орієнтовані на селективну локальну обробку пікселів, використовуючи невеликі рецептивні поля та правила, які активують корекцію за локальним виявленням зашумлених пікселів. Такий підхід природно ефективний для ізольованих імпульсів, однак він не забезпечує достатньої «глобальної» перспективи для відновлення великих кластерів. Коли розмір плями перевищує локальний контекст, інформація про геометрію плями, її внутрішню неоднорідність і відношення до навколишніх тканин втрачається - і це веде до двох взаємопов'язаних дефектів: залишкові артефакти та втрата різкості контурів.

Ще один суттєвий недолік пов'язаний із застосуванням евристичних детекторів та жорстких порогів у класичних SaP-орієнтованих рішеннях. Якщо виявлення SaP базується на абсолютних інтенсивностях або на простих локальних порогах виду k -сусідів, то на реальних медичних даних із варіативною контрастністю такі правила часто виявляються ненадійними. Це породжує два клінічно критичних типи помилок: хибно позитивні - помилкове видалення істинних яскравих/темних анатомічних структур, і хибно негативні - пропуск частин кластерів шуму, інтенсивність яких не перевищує заздалегідь заданий поріг. У синтетичних наборах із жорстко заданою інтенсивністю механізм працює пристойно, але в реальній практиці інтенсивність структур залежить від протоколу сканування, попередньої обробки та характеристики тканин, тому однакові значення інтенсивності можуть відповідати як шуму, так і важливим анатомічним ознакам. Наслідком є або втрата клінічно значущих деталей, або залишення помітних артефактів у зонах інтересу - обидва випадки неприйнятні для діагностичних застосувань. Тому застосування жорстких детекторів без

адаптивних або навчальних механізмів виявлення кластерів та контекстної інтерпретації інтенсивності є суттєвою слабкістю підходу.

Причина цього явища пов'язана з самим принципом оптимізації. Технічно це пояснюється тим, що при оптимізації під MSE (L2) модель прагне мінімізувати середню квадратичну помилку по всій площі; для великих зашумлених зон гіпотеза «усереднення» дає меншу сумарну помилку, ніж відтворення складної текстури, тому мережа віддає перевагу згладженому рішенню і погано зберігає локальні контрасти. SeConvNet має селективні локальні блоки й часто навчається під MSE/PSNR-орієнтовані метрики. При зустрічі з великою прямою SaP-шуму модель не розпізнає внутрішньої істинної текстури цієї ділянки і тому виробляє усереднений сигнал. Через це збільшується локальна помилка, хоча глобальний PSNR може змінитися незначно. У медичних зображеннях це може означати втрату меж між тканинами або стерту мікроструктуру - клінічно значущі невеликі ознаки.

2.2. Модифікація нейромережевої моделі для відновлення зображень на основі застосування механізмів уваги

Convolutional Block Attention Module (CBAM) [53] - це плагін-модуль уваги, що послідовно застосовує два підмодулі: каналної та просторової уваги. Нехай $F \in R^{C \times H \times W}$ - вхідна карта ознак; CBAM обчислює 1D-каналну мапу уваги $M_c \in R^{C \times 1 \times 1}$ і 2D-просторову мапу уваги $M_s \in R^{1 \times H \times W}$ послідовно. Процес уваги описується рівняннями:

$$\begin{aligned} F' &= M_c(F) \otimes F, \\ F'' &= M_s(F') \otimes F' \end{aligned}$$

де $\otimes$ - поелементне множення. При множенні значення каналної уваги копіюються по просторовому виміру, а просторової - по каналному.

Канальна увага фокусується [53] на тому, що важливо. Для обчислення $M_c(F)$ спочатку здійснюють максимальне об'єднання по кожному каналу, отримуючи вектори розмірності $C \times 1 \times 1$. Ці вектори пропускаються через двошаровий перцептрон з активацією ReLU і сигмоїдою. Для підвищення точності CBAM використовує два контекстні вектори: середній F_{avg} і максимальний F_{max} ,

оброблювані спільними вагами MLP. Формула каналного механізму:

$$M_c(F) = \sigma(MLP(AvgPool(F)) + MLP(MaxPool(F)))$$

де σ - сигмоїдна функція. Важливо, що використання тільки середнього пулінгу робить цей механізм еквівалентним модулю SE (Squeeze-and-Excitation); CBAM же додає max-пулінг для тоньшого кодування важливості каналів. Результируючі коефіцієнти $M_c(F)$ множаться на кожен канал карти F, підсилюючи найбільш інформативні канали і ослаблюючи несуттєві.

Просторова увага визначає [53], де на карті елементи на які необхідно звернути увагу. Вона спочатку об'єднує каналну інформацію шляхом усереднення і максимум-пулінгу за каналом, отримуючи дві плоскі карти

$F_{avg}^s \in R^{1 * H * W}$ і $F_{max}^s \in R^{1 * H * W}$. Далі ці дві карти об'єднуються в двоканальне зображення і пропускаються через згортку розміром 7×7 з активацією sigmoid. Формально:

$$M_s(F) = \sigma(f^{7*7}([AvgPool(F); MaxPool(F)]))$$

де f^{7*7} - операція згортки. Маска $M_s(F)$ вказує, наскільки слід підсилити кожен піксель: області з високою концентрацією шуму чи важливими деталями отримують значення близькі до 1, таким чином підсилюючи, вторинні ділянки - близькі до 0 - приглушуються. У підсумку поєднання двох модулів дає можливість одночасно враховувати що і де важливо: спершу каналне перепризначення наголошує ключові ознаки, потім просторове - конкретні локації. Таке послідовне поєднання модулів вважається оптимальним. CBAM модуль легко інтегрується в будь-яку CNN-послідовність. Крім того, CBAM є відносно легким за обчислювальною складністю: приріст параметрів і операцій мінімальний. Тому його часто застосовують в легких мережах без значного навантаження.

SaP - це локалізовані, дуже контрастні пікселі, які не підкоряються статистиці гладких шумів. Два наслідки:

- Шум локальний у просторі: одиночні/скупчення пікселів, які мають невелику просторову зв'язність.
- Шум виглядає як сильні пікові відхилення - каналні фільтри можуть реагувати некоректно, якщо вони орієнтовані лише на середню енергію.

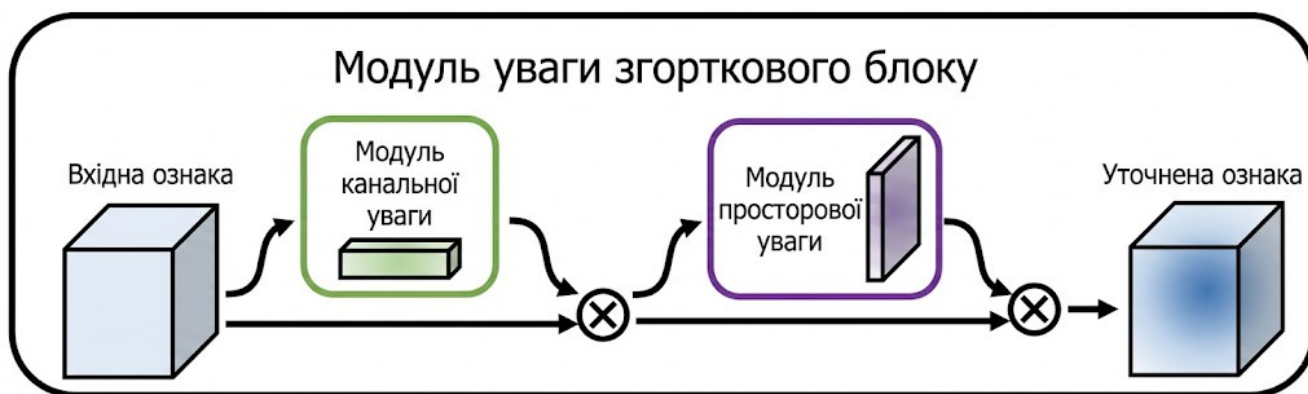


Рис. 2.2. Огляд СВМ. Модуль складається з двох послідовних підмодулів: каналного та просторового.

СВАМ [53] дає два механізми, які допомагають впоратись з цими особливостями: Канальна увага (через глобальний Avg/Max pooling + MLP) дозволяє підсилити ті канали, які корелюють з реальними структурами, і зменшити значення каналів, що реагують на випадкові локальні піки. А просторова увага (через канал-агрегацію Avg/Max $\rightarrow 7 \times 7$ conv $\rightarrow$ sigmoid) дозволяє локалізувати області, які виглядають як справжні структури, і приглушити ізольовані пікселі або невеликі групи - тобто реалізувати щось схоже на адаптивне локальне пригнічення піків.

У відновленні рентгенівських зображень зі SaP-шумом особливо важливо зберігати дрібні структури та контури об'єктів. СВМ надає мережі гнучкість вибірково підсилювати релевантну інформацію. Поєднання каналної та просторової уваги дає змогу мережі виділяти і підсилювати ті канали і пікселі, що містять істотні ознаки, водночас приглушуючи контрастні шумові артефакти. Дослідження показують, що увага допомагає моделі фокусуватися саме на зашумлених ділянках та ключових ознаках і таким чином м'яко усувати шум без згладжування важливих деталей.

Що можна очікувати від блоку уваги

- Краща збереженість країв у порівнянні з простими середніми/медіанними фільтрами, які змивають деталі. СВМ дає селективне пригнічення шуму, не розмиваючи контури.
- Адаптивність до різної щільності SaP шуму при відповідному тренуванні.

Мережа вчиться відрізняти реальні структури від шумових піків різної частоти.

- Невеликий приріст параметрів порівняно з важкими глобальними механізмами самоуваги: CBAM - відносно легкий і, як показано, додає помітне поліпшення при невеликих накладних витратах. Але легкий не означає незначний - у великих мережах додати CBAM у кожен блок може бути дорогим.

У запропоновану архітектуру SeConvNet інтегровано плагін-блок уваги, реалізований за принципами Convolutional Block Attention Module (CBAM) [53] та розміщений у залишкових секціях мережі. Архітектурно блок активується після операції залишкового додавання, що дає масці уваги доступ до комбінованого представлення - сумарного сигналу початкового шляху й корекційної гілки. Модуль складається з двох послідовних підмодулів: каналного та просторового. Канальна складова обчислює вектор ваг шляхом глобальної агрегації, двошарового перцептрон з проміжною ReLU і вихідною сигмоїдою, та відновлення розмірності до форми карти ознак; математично це записується як $M_c(F) = \sigma(W_2 ReLU(W_1(GAP(F))))$

де $F \in R^{C \times H \times W}$ - вхідна карта ознак, а W_1, W_2 - ваги MLP. Отримані канальні коефіцієнти множаться по всій просторовій сітці, що забезпечує селективне масштабування каналів, корельованих із релевантними структурними ознаками, і одночасне пригнічення каналів, чутливих до імпульсних аномалій.

Після того як мережа навчилася автоматично вибирати, які канали несуть корисну інформацію, а які в основному шум, застосовується просторова увага, яка уточнює локалізацію значимих пікселів. У реалізації просторовий модуль формує дві одномірні карти, конкатанує їх і пропускає через згортку 7×7 із вихідною сигмоїдою, утворюючи маску $M_s \in R^{H \times W}$. Така послідовність дозволяє моделі адаптивно підсилювати інформативні канали, а потім локально концентрувати ресурси на областях, де присутні істотні структури, що особливо корисно при подавленні ізольованих імпульсних артефактів SaP.

Комбінація залишкових зв'язків і CBAM додає кілька важливих практичних

переваг. По-перше, залишкові додавання полегшують навчання глибокої мережі та забезпечують збереження низькорівневих ознак, що виступають базовим сигналом для якісного відновлення; по-друге, канално-просторова увага забезпечує гнучкий, диференційований механізм пригнічення шуму за рахунок адаптивних ваг, зменшуючи потребу в жорстких евристичних порогах і знижуючи ймовірність масового розмивання корисних деталей. Водночас слід зазначити технічні застереження: у реалізації використано лише глобальне середнє пулінг для каналного модуля, а також застосовано локальну 7×7 згортку в просторовому модулі - конструкція, яка надає сильну локальну розрізняльність, але цього недостатньо для врахування різних масштабів та загальної структури зображення при обробці великих скупчень шуму. З огляду на це, запропонована інтеграція СВМ представляє збалансоване інженерне рішення: вона підвищує селективність фільтрації імпульсного шуму й одночасно зберігає дрібну структурну інформацію завдяки поєднанню стабілізуючого впливу шунтувальних зв'язків і адаптивних масок уваги; для повної обробки кластерного шуму доцільно розглядати подальші доповнення як напрямки подальшого розвитку архітектури.

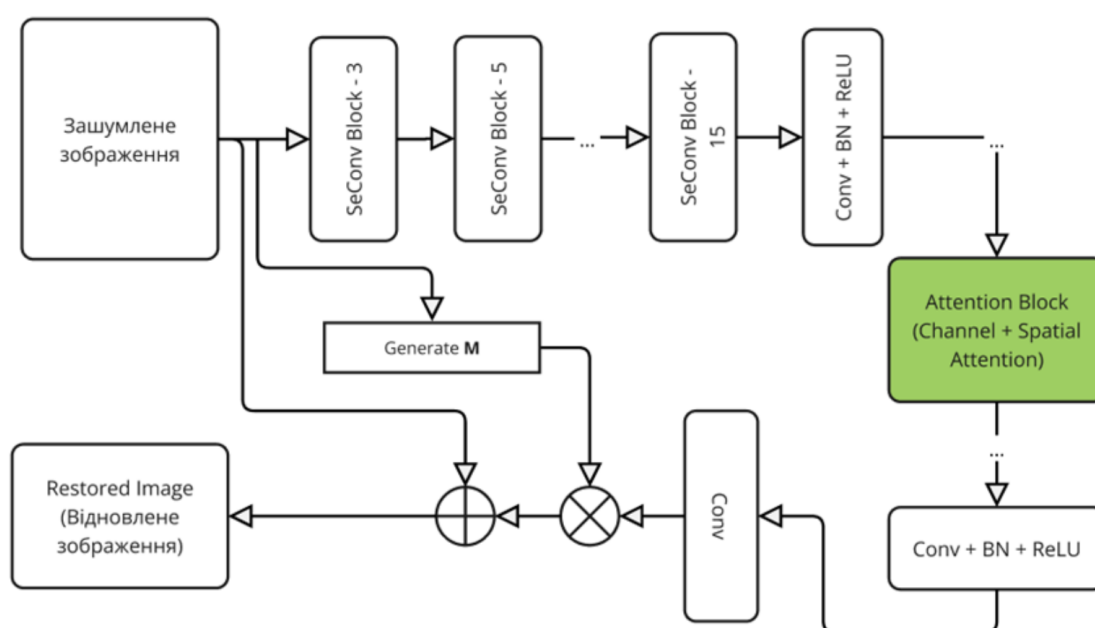


Рис. 2.3. Архітектура запропонованої модифікованої моделі

Порівняння візуальних результатів свідчить, що для тестових зображень з рівномірними або не надто контрастними ділянками показники PSNR і візуальна якість залишилися орієнтовно на тому ж рівні, що й в оригінальній SeConvNet. На рис. 3. наведено приклади відновлення фрагментів рентгенівських знімків, де модифікована модель демонструє схожу з базовою якість.

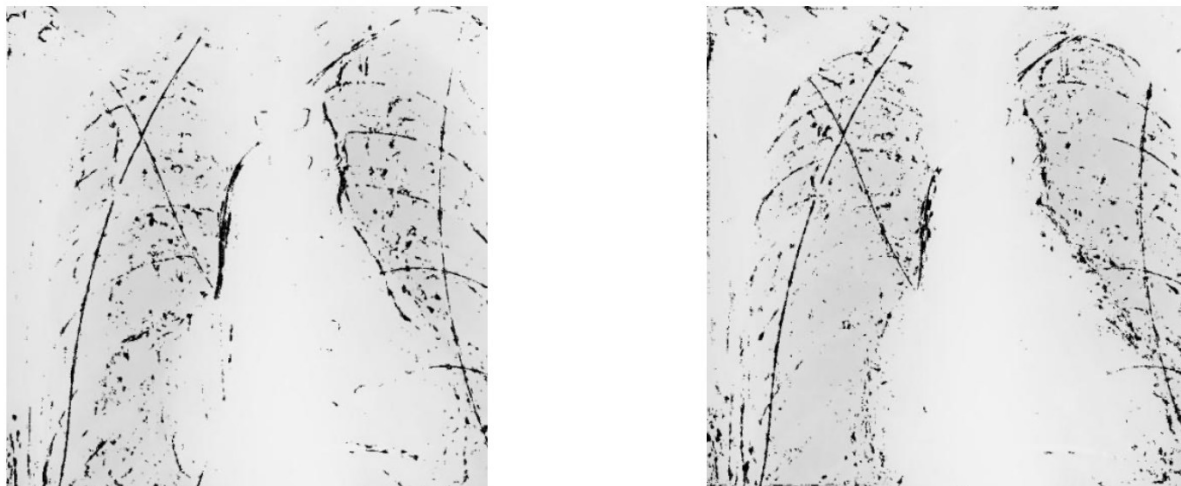


Рис. 2.4. Порівняння результатів роботи оригінального зображення з оригінальною моделлю і оригінальне зображення з модифікованою моделлю. Водночас, для окремих рентгенівських знімків з набору зображень, які від початку характеризувалися значною кількістю білих плям, які можна розглядати як пікові відхилення SaP шуму на темному тлі, модифікований варіант SeConvNet [68] дає кращий результат. PSNR зростає на 6-7% в порівнянні з результатом роботи оригінальної моделі.

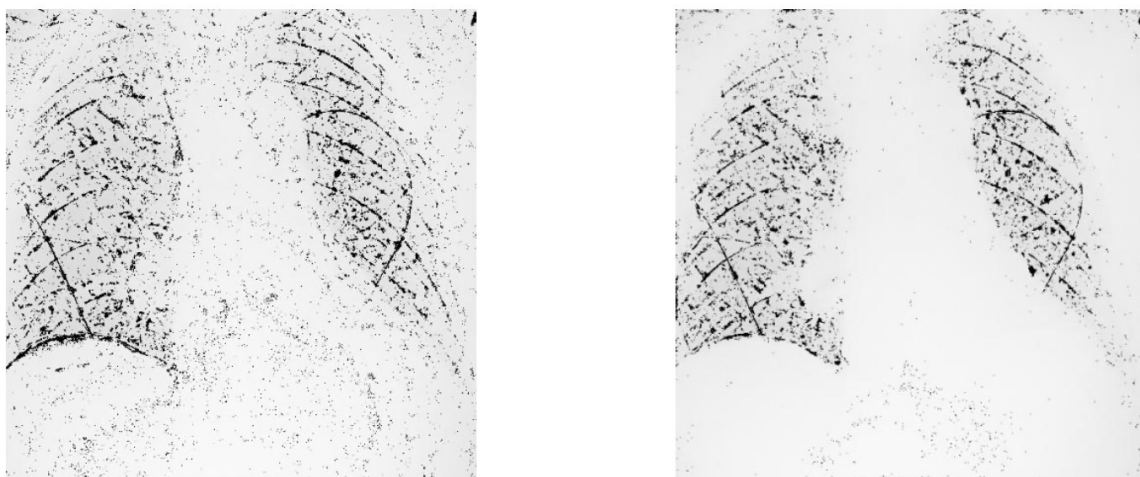


Рис. 2.5. Порівняння результатів роботи оригінального зображення з оригінальною моделлю і оригінальне зображення з модифікованою моделлю

На відміну від класичних варіантів CBAM, що застосовуються після звичайних згорткових блоків, у цій роботі модуль уваги розміщено безпосередньо після залишкового додавання в межах SeConv-сегментів. Така архітектурна позиція забезпечує доступ уваги до комбінованого представлення вихідного сигналу та корекційної гілки, що дозволяє формувати контекстно-залежну маску, здатну відрізнити інформаційний сигнал від імпульсних артефактів.

Крім того, реалізація CBAM [53] модифікована з урахуванням специфіки SeConv-масок: каналний модуль працює лише з ненульовими пікселями, що усуває спотворення у зонах повної втрати даних.

Такий підхід забезпечує адаптивну селекцію релевантних каналів і просторових зон при мінімальному збільшенні кількості параметрів, чого не забезпечують стандартні варіанти CBAM або SE-блоків у задачах відновлення імпульсно зашумлених зображень.

2.3. Аналіз параметрів модифікованої моделі

Аналіз параметрів модифікованої архітектури згорткової нейронної мережі із вбудованими механізмами уваги (Convolutional Block Attention Module, CBAM) є надзвичайно важливим для верифікації ефективності запропонованого підходу. Даний аналіз охоплює як архітектурні та алгоритмічні характеристики, так і експериментальні результати на різних вибірках даних, включаючи складні медичні рентгенівські зображення з високим рівнем імпульсного шуму типу SaP.

2.3.1. Архітектурні особливості

Модифікована модель зберігає загальну логіку початкової мережі SeConvNet [68], проте ключова відмінність полягає в інтеграції блоків CBAM після кожного залишкового блоку. Це стратегічне розташування дозволяє механізму уваги динамічно перекалібрувати карти ознак, що виходять із глибоких шарів мережі, зосереджуючись на найбільш інформативних аспектах зображення.

Архітектурно CBAM [53] включає два послідовно розташовані підмодулі:

- Канальний механізм уваги - дозволяє виділити найбільш інформативні канали ознак. Цей підмодуль стискає просторову інформацію,

використовуючи методи глобального агрегування, щоб визначити, які канали мають бути посилені для конкретного вхідного зображення.

- Просторовий механізм уваги - який підсилює значущі області на зображенні та приглушує ділянки з шумом. Він працює з просторовими взаємозв'язками пікселів, створюючи карту просторової уваги, яка ефективно ізолює корисні структури від небажаних артефактів.

Важливою особливістю є те, що у СВАМ коефіцієнти уваги

M_c - channel attention і M_s - spatial attention, генеруються для кожного зображення як функція від вхідних ознак F , що дає змогу уникати фіксованих евристичних правил, які часто використовуються у класичних методах фільтрації. На відміну від класичних методів фільтрації імпульсного шуму, де рішення про те, чи є піксель шумом, у СВАМ немає жорсткої процедури порівняння - замість цього мережа сама вчиться виявляти аномалії через розподіл ознак у просторі каналів і просторових координат. Це забезпечує вищу гнучкість та адаптивність моделі до різних шумових сценаріїв.

2.3.2. Кількість параметрів та обчислювальна складність

Впровадження механізму уваги СВАМ у базову архітектуру SeConvNet супроводжується незначним зростанням кількості параметрів та обчислювальних витрат. Базова модель, що складається з семи послідовних SeConv-блоків і дев'ятнадцяти згорткових шарів, містить близько 0.70 млн навчальних параметрів.

Додавання СВАМ-блоків після кожного другого згорткового шару збільшує кількість параметрів орієнтовно до 0.72 млн, тобто лише на 2-3 %. Це зростання зумовлене додаванням двошарових повнозв'язних підмодулів у каналній частині уваги та однієї згортки 7×7 у просторовому модулі.

Таке співвідношення параметрів вказує на високу структурну ефективність модифікації: СВАМ додає навчальні можливості, але практично не впливає на розмір моделі. Така властивість робить модель придатною для використання у середовищах із обмеженими обчислювальними ресурсами, зокрема при постобробці томографічних зображень на клінічних станціях.

Обчислювальна складність моделі у FLOPs зростає орієнтовно на 5-7 %, що

обумовлено додатковими операціями глобального усереднення, двома згортками та операціями множення на маски уваги. З огляду на це, модифікована SeConvNet зберігає прийнятну швидкодію при істотно покращеній селективності шумозаглушення.

Таблиця 2.1. Порівняння обчислювальної складності та кількості параметрів для базової SeConvNet та модифікованої SeConvNet + CBAM

Модель	Кількість параметрів (млн)	FLOPs ($\times 10^9$)	Приріст параметрів
SeConvNet	0.70	4.8	-
SeConvNet + CBAM	0.72	5.1	+2-3 %

Збільшення FLOPs лише до 5.2×10^9 підкреслює ефективність архітектурного рішення, оскільки значний приріст якості досягається за незначних додаткових обчислювальних витрат.

2.3.3. Ефективність відновлення за різними метриками

Для оцінки ефективності модифікованої моделі використовувалися дві основні метрики: Peak Signal-to-Noise Ratio (PSNR) та Structural Similarity Index (SSIM). Метрика PSNR характеризує ступінь близькості відновленого зображення до оригіналу, а отже - загальний рівень втрати інформації після шумозаглушення. Вищі значення PSNR свідчать про менші спотворення інтенсивності.

Метрика SSIM оцінює подібність між відновленим та еталонним зображеннями з урахуванням локальної яскравості, контрасту й структурних особливостей, відображаючи сприйняття якості з позицій зорового аналізу людини. Для тестування було використано вибірку рентгенівських знімків із різним рівнем контрасту та ступенем зашумлення.

Таблиця 2.2. Порівняння показників PSNR та SSIM між базовою моделлю SeConvNet та модифікованою SeConvNet + CBAM на тестовому наборі рентгенівських знімків

Модель	PSNR	Δ PSNR (%)	SSIM	Δ SSIM (%)
Оригінальна SeConvNet	33.10625	-	0.8855	-
Модифікована SeConvNet + CBAM	35.4071	+6.95	0.9155	+3.39

2.4. Метод підвищення ефективності відновлення зображень на основі адаптивний підходу до покращення параметрів зображень

Модифікована модель зберігає основну ідею SeConvNet [68] - по чергове застосування спеціалізованих SeConv-блоків для заповнення/корекції нульових пікселів та подальшу послідовність стандартних згорткових шарів - проте доповнена періодичною увагою в частині звичайних Conv-шарів. Зокрема, у секції звичайних згорток після кожного другого Conv-BN-ReLU виконується залишкове додавання і потім застосовується CBAM-подібна послідовність: спочатку канална увага, потім просторовий механізм уваги. Таким чином увага застосовується локально після кожного такого залишкового блоку, а не лише один раз між великими блоками мережі.

SeConv-блок [68] визначає маску ненульових пікселів і виконує нормалізовану згортку, яка обчислює середнє по ненульових сусідах. Отримане середнє підставляється тільки в положеннях, де вхідні пікселі були нульовими і лише там, де виконано додаткову умову достатньої кількості сусідів. Іншими словами, SeConv блокує переписування ненульових пікселів і служить локальним

методом заповнення/інтерполяції на основі сусідів.

Контекстно-канальний блок уваги (Convolutional Block Attention Module, CBAM) у запропонованій реалізації виступає як адаптивний механізм перерозподілу інформаційного вмісту проміжних ознак мережі. Метою його інтеграції є автоматичне підсилення релевантних каналів та просторових зон ознакових карт, що виявляють корисну структуру для завдання відновлення зображень, та пригнічення компонент, які відповідають за артефакти й залишковий шум. Спершу обчислюється канальна увага: із вихідного тензора ознак виділяються глобальні статистики за допомогою середнього та максимального значення по просторових координатах для кожного каналу. Отримані вектори проходять через два повнозв'язні шари - зменшувальний та відновлювальний - які моделюють взаємозв'язок між каналами та генерують коефіцієнти важливості для кожного каналу. Кожен канал вихідного тензора модулюється відповідним коефіцієнтом, що дозволяє підсилити важливі ознаки та пригнічувати менш значущі.

Далі застосовується просторова увага: на основі канально відкоригованого тензора обчислюються середнє та максимальне значення по каналах для кожного пікселя, після чого ці карти об'єднуються у двоканальний тензор. Цей тензор проходить через згортковий шар з активацією сигмоїда, що формує маску просторової уваги. Кожен піксель вихідного тензора модулюється цією маскою, що дозволяє виділяти області з найбільш релевантними ознаками для відновлення зображення.

Після проходження CBAM блок повертає модифікований тензор ознак, який містить адаптивно підсилені канали та просторові області. Цей механізм дозволяє мережі адаптивно покращувати якість відновлення зображень, зосереджуючись на найбільш інформативних характеристиках і пригнічуючи шум або неважливі деталі.

Фінальний шар моделі обчислює карту з тим же числом каналів, що й вхід і тільки потім додає початковий вхід. Тому кінцевий вихід - це оригінальне зображення, де мережа заміняє або доповнює лише нульові пікселі своїми прогнозами.

Формальна модель запропонованої архітектури нейронної мережі для відновлення зображень, зашумлених імпульсним шумом із кластерною структурою.

1. Вхідні дані

Нехай $X_{in} \in R^{H*W*C}$ - вхідне зашумлене зображення, де пошкоджені пікселі мають значення 0. Визначається бінарна маска пошкоджень M_{noise} , де $M_{noise}(i, j) = \{ 1, \text{якщо } X_{in}(i, j) = 0 \text{ (шум)}; 0, \text{якщо } X_{in}(i, j) \neq 0 \text{ (дані)}$

2. Каскад селективних згорток (SeConv Cascade)

Архітектура використовує послідовність із N блоків SeConv для ітеративного заповнення пропущених даних на етапі попередньої обробки. Нехай $F_0 = X_{in}$. Тоді вихід k-го блоку: $F_k = SeConvBlock_k(F_{k-1}), k = 1, \dots, N_{SeConv}$

3. Глибинний видобуток ознак (доданий блок).

Отримані ознаки F_n проходять через глибоку згорткову мережу з залишковими зв'язками. Архітектура складається з блоків, що містять:

- Стандартні згортки;
- Нормалізацію та активацію;
- Залишкове додавання кожні 2 згорткові шари
- Механізм уваги СВМ, який послідовно застосовує каналну (A_c) та просторову (A_s) увагу:

$$\begin{aligned} F' &= A_c(F) \otimes F, \\ F'' &= A_s(F') \otimes F' \end{aligned}$$

4. Відновлення та формування виходу

Фінальний шар перетворює глибинні ознаки назад у простір зображення. Унікальною особливістю моделі є те, що вона замінює лише пошкоджені пікселі, залишаючи оригінальні дані незмінними:

$$Y_{out} = (Conv_{final}(F_{deep}) \odot M_{noise}) + X_{in}, \text{ де:}$$

- $Conv_{final}$ - фінальна згортка;
- $\odot$ - поелементне множення;
- Додавання X_{in} відновлює оригінальні (не пошкоджені) пікселі, оскільки M_{noise}

для них дорівнює 0, а збережені дані знаходяться в X_{in} .

Очікувана користь цього підходу - поєднати локальне контекстне відновлення нульових/втрачених пікселів (SeConv) і селективне підсилення/придушення ознак через каналну та просторову увагу під час фази тонкого відновлення. Таке поєднання може підвищити здатність моделі відфільтрувати імпульсний SaP шум при збереженні контурів і тонких структур.

2.5. Обґрунтування вибору архітектури SeConvNet + CBAM

Вибір запропонованої архітектури SeConvNet + CBAM зумовлений суворими вимогами до достовірності медичної візуалізації та специфікою розподілу імпульсного шуму. Прийняте архітектурне рішення базується на порівняльному аналізі трьох класів підходів до відновлення зображень.

Традиційні алгоритми, такі як медіанна фільтрація або нелокальні методи, демонструють обмежену ефективність при екстремально високій щільності шуму до 95%. Оскільки ці методи спираються на фіксовані математичні правила, вони неминуче призводять до розмиття меж анатомічних структур та втрати діагностично значущих деталей. Натомість застосування нейромережевого підходу дозволяє адаптивно виявляти приховані закономірності в анатомічних структурах та відновлювати їх із вищою точністю, спираючись на навчені ознаки, а не лише на статистику сусідніх пікселів.

Незважаючи на популярність дифузійних моделей та візуальних трансформерів, їх застосування в медичній практиці пов'язане з критичними ризиками. Дифузійні моделі, що працюють за імовірнісним принципом, схильні до галюцинацій - генерації анатомічно правдоподібних, але фактично неіснуючих деталей, що є категорично неприпустимим у клінічній діагностиці. Крім того, такі архітектури потребують значних обчислювальних потужностей. Запропонована згортова модель є детермінованою, ресурсоефективною та забезпечує стабільну відтворюваність результатів, що відповідає вимогам до клінічного програмного забезпечення.

Природа шуму SaP полягає у локальній втраті інформації, де піксель набуває

екстремального значення. Базовим механізмом фільтрації обрано блоки селективної згортки, який діє як жорсткий фільтр: алгоритм ігнорує пошкоджені пікселі під час згорткових операцій, запобігаючи поширенню шумових артефактів на сусідні ділянки.

Для компенсації недоліку SeConv, який полягає у надмірному згладжуванні великих кластерних плям шуму через брак локального контексту, у модель інтегровано механізм уваги CBAM. Він виконує функцію м'якого фільтра, дозволяючи мережі фокусуватися на глобальних релевантних ознаках та відновлювати внутрішню структуру пошкоджених ділянок.

Важливо зазначити, що модуль CBAM є обчислювально легким компонентом, який збільшує кількість параметрів моделі лише на 2-3%, що дозволяє зберегти високу швидкість та можливість розгортання системи на обладнанні з обмеженими ресурсами.

2.6. Висновки до розділу

У другому розділі проведено детальний аналіз обмежень базової моделі SeConvNet та обґрунтовано необхідність її модифікації для підвищення ефективності відновлення медичних зображень. Хоча оригінальна архітектура добре справляється з поодинокими імпульсами, її поведінка на медичних зображеннях значно погіршується при обробці великих кластерних плям SaP-шуму, що спричиняє характерні дефекти - залишкові спотворення та надмірне згладжування. Фундаментальною причиною цього з точки зору оптимізації є те, що мінімізація середньоквадратичної помилки стимулює усереднення значень у великих пошкоджених ділянках, що в результаті шкодить відтворенню дрібних анатомічних деталей.

Для подолання цих обмежень розроблено та описано інтеграцію легкого блоку канално-просторової уваги (Convolutional Block Attention Module, CBAM) у загальну архітектуру моделі відновлення зображень. Послідовне застосування каналного та просторового підмодулів дозволяє мережі адаптивно підсилювати важливі фільтри й локалізувати корисні просторові зони, що суттєво підвищує

селективність фільтрації імпульсного шуму. Запропонований адаптивний підхід гармонійно поєднує механізми локальної інтерполяції за допомогою SeConv-блоків та періодичної уваги (CBAM) безпосередньо у секціях звичайних згорток.

Оцінка внесених архітектурних змін підтверджує, що впровадження CBAM-блоків є раціональним рішенням, яке позитивно впливає на результати при мінімальному зростанні кількості параметрів та обчислювальних витрат. Крім того, вибір гібридної архітектури SeConvNet + CBAM повністю обґрунтовано як безпечну та надійну клінічну альтернативу. На відміну від традиційних фільтрів, які неминуче розмивають контури при високій щільності шуму, та сучасних дифузійних моделей, що несуть високі ризики генерації клінічно недостовірних артефактів, розроблений метод гарантує детерміноване та точне відновлення. Усі ці напрацювання формують міцну основу та визначають вектор для подальших експериментальних досліджень архітектури.

РОЗДІЛ 3. ЗАСОБИ ТА МЕХАНІЗМИ ДЛЯ ПІДВИЩЕННЯ ЕФЕКТИВНОСТІ ПРОЦЕСІВ ВІДНОВЛЕННЯ ЗОБРАЖЕНЬ

3.1 Реалізація засобів на основі нейромережевих моделей для підвищення ефективності обробки зображень

Мета розділу - описати структуру модифікованої моделі SeConvNet із інтегрованим СВАМ, обґрунтувати вибір компонентів з позицій задач відновлення та ефективності на різних типах шуму, деталізувати процедуру навчання, вибір даних, функції втрат, гіперпараметри і навести методологію порівняння з базовими підходами.

Модель базується на модифікованій версії SeConvNet з *Selective Convolution*-блоками та включає Convolutional Block Attention Module (CВAM) для поліпшення якості відновлення. Основні компоненти архітектури - вхідний блок, послідовність SeConv-блоків, класичні згорткові блоки з залишковими з'єднаннями, блоки уваги (CВAM) і вихідний шар.

Вхідний блок: На вхід нейронної мережі подається тензор зображення $X \in R^{H*W*C}$, де H та W - просторові розмірності (висота та ширина), а C - кількість каналів, в нашому випадку $C=1$.

Вхідні дані проходять попередню обробку, яка включає наступні кроки:

- Кодування пошкоджень: Пошкоджені пікселі у вхідному зображенні кодуються нульовими значеннями (0). Це дозволяє мережі автоматично генерувати бінарну маску достовірності M , де $M_{i,j} = 1$ якщо піксель містить корисний сигнал, і $M_{i,j} = 0$, якщо піксель є пошкодженим.
- Нормалізація: Значення інтенсивності пікселів, що знаходяться у діапазоні $[0, 255]$, лінійно масштабуються до діапазону $[0, 1]$. Цей крок є критичним

для забезпечення чисельної стабільності градієнтів під час навчання методом зворотного поширення похибки. Нормалізація запобігає насиченню функцій активації та забезпечує збалансований внесок ознак на початкових етапах обробки.

Таким чином, вхідний тензор одночасно несе інформацію як про розподіл інтенсивностей, так і про топологію пошкоджених ділянок, що усуває необхідність передачі окремого тензора маски на вхід мережі.

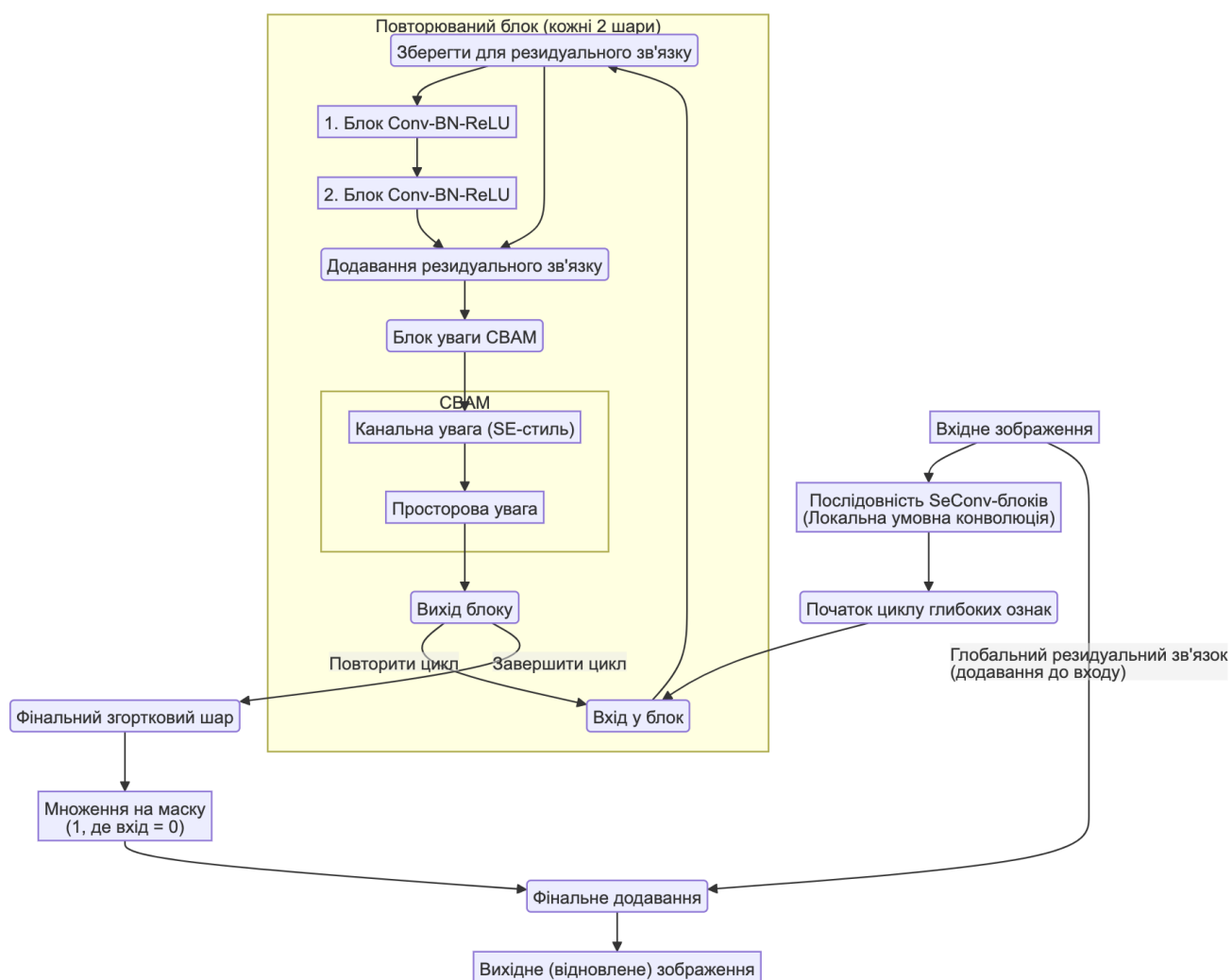


Рис. 3.1. Структура модифікованої моделі SeConvNet із інтегрованим CBAM блоком

SeConv-блоки (Selective Convolution): Для первинного відновлення структурної цілісності зображення застосовано каскад із 7 послідовних блоків SeConv (Selective Convolution). Головною метою цього етапу є ітеративне заповнення пропущених

ділянок на основі локального контексту без внесення артефактів розмиття, характерних для стандартних згорток.

Механізм роботи k -го блоку SeConv описується наступним алгоритмом:

1. Генерація маски достовірності: На основі вхідного тензора формується бінарна маска, що ідентифікує ненульові пікселі.
2. Кожен наступний блок у каскаді використовує ядро згортки збільшеного розміру. Розмір ядра K_s для k -го блоку визначається як $2k+1$ ($3 \times 3, 5 \times 5, \dots, 15 \times 15$). Такий підхід дозволяє поступово охоплювати ширший просторовий контекст, відновлюючи спочатку дрібні деталі, а згодом - більші текстурні шаблони.
3. Селективна агрегація ознак (Partial Convolution mechanism): Стандартна операція згортки модифікується таким чином, щоб враховувати внесок лише "достовірних" пікселів.
4. Умовне оновлення: Оновлення значення пікселя відбувається лише за умови наявності достатньої кількості інформативних сусідів (визначеної пороговим значенням $K_s - 2$). Це запобігає генерації хибних ознак у регіонах з повною відсутністю інформації.
5. Збереження оригінального сигналу: Вихідний сигнал блоку Y^k формується шляхом комбінування відновлених значень S та оригінальних вхідних значень.

В результаті проходження через каскад SeConv-блоків, мережа формує проміжне представлення зображення, в якому більшість нульових прогалів заповнено апроксимованими значеннями, що базуються на локальних кореляціях сусідніх пікселів. Це створює надійну основу для подальшої глибокої обробки механізмами уваги.

Згорткові блоки з залишковими зв'язками: Після етапу попередньої реконструкції (SeConv) отримані карти ознак подаються до глибокої згорткової підсистеми для виділення високорівневих абстракцій. Архітектура даного модуля

базується на принципах залишкового навчання, що дозволяє ефективно тренувати глибокі мережі без ризику деградації точності.

Основні компоненти та принципи роботи модуля:

- Архітектура базового блоку: Структурною одиницею глибокої підсистеми є композиція шарів Conv-BN-ReLU. Кожен такий блок виконує операцію згортки з ядром 3×3 для просторового аналізу, за якою слідує нормалізація для стабілізації розподілу активацій та нелінійна функція активації ReLU для забезпечення виразної здатності моделі.
- Залишкові зв'язки: Ключовою особливістю архітектури є введення прямих шунтувальних зв'язків через кожні два згорткові шари.
- Оптимізація градієнтного потоку: Використання адитивного зв'язку реалізує механізм тотожного відображення. Це дозволяє градієнту функції втрат безперешкодно проходити до ранніх шарів мережі, оминаючи нелінійні перетворення. Такий механізм ефективно вирішує проблему затухання градієнта, забезпечуючи збіжність моделі навіть при значній глибині архітектури.

Сформовані на цьому етапі глибокі карти ознак, що містять як структурну, так і семантичну інформацію, передаються далі до модуля уваги (CBAM) для їх адаптивного зважування та фільтрації.

CBAM (Convolutional Block Attention Module): Після базових згорток формується карта ознак, до якої застосовується CBAM. Кожний такий блок уваги генерує маску, що підсилює важливі ознаки і пригнічує несуттєві. Модуль CBAM інтегрується в CNN-потік без значних витрат ресурсів.

Канальна увага. На першому етапі мережа виконує зважування каналів, тобто визначає, які з них містять найбільш корисну інформацію. Для цього для кожного каналу ознакової карти обчислюється середнє значення всіх пікселів - це

дає компактне представлення змісту кожного каналу. Отримані середні значення формують вектор, який проходить через невелику двошарову повнозв'язну мережу. Перший шар цієї мережі зменшує кількість параметрів орієнтовно у 8 разів і застосовує функцію активації ReLU, що допомагає відокремити суттєві ознаки від несуттєвих. Другий шар відновлює вихідну розмірність і використовує Sigmoid, щоб перетворити значення у ваги в діапазоні від 0 до 1. Таким чином утворюється вектор ваг, де кожен елемент показує, наскільки важливим є певний канал. Цей вектор розширюється до розміру карти ознак і множиться на неї, тобто посилює канали, що містять контури, текстури чи інші значущі структури, і послаблює ті, що пов'язані з шумом або фоном. У цій реалізації використовується лише глобальне усереднення (Global Average Pooling), без додаткового максимального усереднення, як у класичному CBAM. Тому канална увага тут ближча до ідеї SE-блоку (Squeeze-and-Excitation [69]), де акцент робиться саме на середньому значенні ознак, що робить механізм простішим і стабільнішим при навчанні.

Просторова увага. Після того як важливість каналів визначена, мережа переходить до другого етапу - просторової уваги. Її мета - знайти, де саме в межах зображення розташовані найбільш інформативні області. Для цього спочатку обчислюються дві карти - одна містить середні значення по всіх каналах для кожного пікселя, а інша - максимальні. Вони показують, наскільки активною є кожна точка зображення у різних вимірах. Ці дві карти з'єднуються та пропускаються через згортковий шар із ядром 7×7 , який охоплює широку область зображення. На виході отримується маска просторової уваги - карта, що вказує, у яких ділянках зображення потрібно посилити або, навпаки, приглушити ознаки. Цю маску накладають на карту ознак шляхом по компонентного множення. У результаті області, які містять структурно важливу інформацію, стають більш вираженими, тоді як фонові або пошкоджені ділянки приглушуються. У реалізації CBAM цей механізм застосовується після кожних двох згорткових блоків, тобто періодично на різних рівнях абстракції. Крім того, у фінальній частині моделі передбачено множення результату на маску, що виділяє лише ті області, які вхідне

зображення позначило як пошкоджені. Після цього мережа додає отримане “очищене” зображення до початкового, реалізуючи залишковий підхід - тобто вчиться не створювати зображення заново, а лише виправляти дефекти. У сукупності ці два механізми - канална та просторова увага - дозволяють моделі не просто фільтрувати шум, а й адаптивно вирішувати, що саме варто підсилити і де саме це зробити. Такий підхід забезпечує більш точне збереження структур і контурів при одночасному зниженні рівня шуму.

- **Вихідний блок:** Завершальний згортковий шар із ядром 3×3 формує карту відновленого сигналу тієї ж кількості каналів, що й вхід. Отриманий тензор інтерпретується як різниця між чистим і зашумленим зображенням. Далі ця різниця множиться поканально/просторово на бінарну маску, побудовану з вхідного зображення, тобто виправлення застосовуються тільки в тих пікселях, де вхід було позначено як зашумлений (значення пікселя рівне нулю). Після множення отримане виправлення додається до оригінального зображення, що реалізує залишковий підхід: мережа навчається прогнозувати корекцію, а не відтворювати зображення повністю. Така логіка гарантує, що чисті пікселі залишаться без змін, а модифікація торкнеться лише тих позицій, які позначені як пошкоджені.

Вибір модулів (SeConv + Conv-residual + CBAM) обґрунтовано трьома ключовими вимогами задачі відновлення зображень з масками:

1. локальна адресна інтерполяція пошкоджених пікселів
2. збереження глобальної структури та контурів
3. ефективна селекція релевантних каналів без надмірного росту параметрів.

SeConv: селективна згортка ігнорує нульові пікселі під час обчислення, що робить її натуральним локальним інтерполятором для імпульсного шуму і великих «дір». У порівнянні зі стандартним Convolution, SeConv зменшує ризик розмивання контурів там, де входять значні маски. Інтеграція CBAM обрана тому, що каналова і просторова увага дозволяють мережі адаптивно зосереджуватися на інформативних ознаках та ділянках зображення. Архітектура включає селективні

згортки для первинної обробки шуму, послідовні Conv-ReLU блоки з залишковими зв'язками та блок уваги, що адаптивно зважує канали і регіони. Такий інструмент націлений на відновлення якісного безшумного зображення: він будується на сучасних архітектурних компонентах і показує покращення результуючих метрик за рахунок фокусування на релевантних ознаках. Існують також і альтернативи. U-Net [70, 71] добре працює при багаторівневій реконструкції, але без SeConv не оптимізована для випадків із великою кількістю повністю відсутніх пікселів. SwinIR і Restormer - сучасні трансформерні архітектури, які часто дають кращу якість при малих і середніх рівнях шуму, але за значно вищу обчислювальну вартість і більшу кількість параметрів [72, 73, 74]. Тому вибір SeConvNet + CBAM-компроміс: цілеспрямована локальна стратегія для сильних локальних пошкоджень + легкий механізм уваги для збереження глобальних ознак при помірних обчислювальних витратах

3.2 Механізми реалізації ефективного шумозаглушення в зображеннях

Алгоритм роботи моделі складається з послідовних етапів: (1) локальна попередня реконструкція через SeConv-блоки; (2) витягнення ієрархічних ознак згортковими шарами; (3) адаптивне зважування ознак через CBAM; (4) формування залишкової корекції й її застосування лише до пошкоджених пікселів. Така поетапність дозволяє комбінувати локальні й глобальні стратегії відновлення: локально заповнити прогалини, а потім глобально відфільтрувати залишковий шум, зберігаючи структуру. Підхід до тренування полягає у навчанні моделі відновлювати чисте зображення з навмисно пошкоджених прикладів: генератор створює пошкоджені зображення, а еталонні чисті використовуються як цільові. Завдяки цьому мережа явно навчається реконструювати відсутню або спотворену інформацію.

Селективна згортка (SeConv)

SeConv-блоки працюють як локальні реконструктори: при обчисленні згортки враховуються лише ненульові пікселі, а нульові - ігноруються. Практично це

означає, що фільтр інтерпретує значення навколишніх коректних пікселів і використовує їх для заповнення точок пошкодження, не додаючи шум по всьому полю. Додаткова умова на кількість непошкоджених сусідів обмежує інтерполяцію, якщо навколо занадто багато відсутніх пікселів - це зменшує артефакти при відновленні великих дір. Отже, SeConv діє як перша лінія при обробці імпульсного шуму і місцевих пропусків сигналу. Після попередньої реконструкції йде послідовність Conv-BatchNorm-ReLU блоків, які формують ієрархічні представлення зображення. Залишкові з'єднання, що додають попередні активації до поточних, полегшують навчання глибокої мережі: вони зберігають базовий каркас сигналу і дозволяють блокам навчатися корекції замість повної генерації образу. Це суттєво знижує ризик деградації якості при нарощенні глибини і сприяє стабільній збіжності.

СВАМ

СВАМ поєднує дві взаємодоповнюючі стратегії уваги щоб адаптивно виділити релевантний сигнал.

- *Канальна увага* оцінює, які фільтри несуть корисну інформацію про структуру зображення. Для цього кожен канал стискається до скалярного показника, після чого невелика повнозв'язна мережа генерує ваги в інтервалі $[0,1]$. Канали з високою вагою посилюються - це підсилює детекторні реакції на контури й текстури, а слабкі канали - пригнічує.
- *Просторова увага* вказує, де саме в карті ознак слід фокусувати ресурси. Для кожної позиції пікселя обчислюються середнє і максимум по каналах, утворюється двоканальна просторова карта і через широку згортку (7×7) формується маска уваги. Ця маска посилює локальні ділянки з інформативними структурами і пригнічує випадкові коливання фону.

Комбіноване застосування дає синергічний ефект: каналова увага «очищує» спектр ознак, а просторова - скеровує корекцію у потрібні координати.

Механізми протидії різним типам шуму.

- *Імпульсний шум*: SeConv-блоки працюють як локальні інтерполятори - вони використовують лише ненульові сусідні пікселі, щоб швидко згладити поодинокі пікселі шуму. Далі СВАМ зменшує вплив залишкових артефактів: шум викликає непослідовні відгуки по каналах і в просторі, і модуль уваги призначає таким відгукам низькі ваги.
- *Гаусівський шум*: основну роботу виконують глибокі згорткові шари (Conv-BN-ReLU) у поєднанні з каналною увагою, яка виділяє канали з корисною текстурною або частотною інформацією; просторовий механізм уваги допомагає зберегти різкі контури, не допускаючи їх надмірного згладжування.

Застосування маски та залишкове складання.

Фінальна залишкова корекція формується останнім згортковим шаром; її результат помножується на бінарну маску, що позначає пошкоджені пікселі. Після множення отримане виправлення додається до вхідного зображення. Така послідовність гарантує, що модифікації будуть виконані лише в позиціях, позначених як пошкоджені, що запобігає зміні чистих пікселів і фокусує навчання на відновленні дефектів. Розглянуті механізми утворюють інструментарій, який здатний адресно та адаптивно боротися з різними видами шуму, зберігаючи структурну цілісність зображення. Роль СВАМ - не в простому фільтруванні; це адаптивний механізм, що визначає, які ознаки і в яких просторових ділянках слід підсилювати або приглушувати, а SeConv і залишкові зв'язки забезпечують реалізацію цих корекцій на локальному та глобальному рівнях. У процесі навчання нейромережевої моделі SeConvNet з модулем СВАМ передбачається використання синтетично згенерованих вибірок, що імітують різні типи пошкоджень зображень. Навчальна вибірка формується на основі набору медичних зображень, до яких застосовується SaP шум.

Модель тренується на наборах даних середнього розміру, які охоплюють різні умови освітлення й текстури. Використання синтетичного шуму замість реального дає змогу точно визначати параметри деградації, забезпечуючи

можливість відтворення експериментів. Навчання здійснюється з використанням оптимізатора Adam, який забезпечує адаптивну швидкість оновлення ваг для кожного параметра моделі. Типові параметри навчання включають обмежену кількість епох, поступове зменшення швидкості навчання та механізми ранньої зупинки для запобігання перенавчанню.

Теоретично очікується, що завдяки адаптивній структурі SeConv-блоків і механізмам уваги CBAM модель досягатиме швидшої збіжності порівняно з базовими згортковими мережами та потребуватиме меншої кількості ітерацій для досягнення стабільної втрати.

3.3. Механізми уваги (CBAM, SENet) у практичних системах відновлення зображень

Сучасні нейромережеві архітектури для відновлення зображень дедалі частіше інтегрують механізми уваги, які дозволяють моделі не просто навчатися глобальним закономірностям, а адаптивно фокусуватись на найбільш інформативних регіонах. У традиційних згорткових мережах всі пікселі та канали опрацьовуються рівномірно, тоді як у зображеннях інформація розподілена нерівномірно - більшість деталей міститься у граничних та текстурних областях. Тому використання блоків уваги, зокрема CBAM та SENet [69], стає ключовим інструментом для підвищення якості реконструкції без істотного збільшення обчислювальних витрат.

Концепція уваги в задачах реконструкції

Механізми уваги імітують здатність людського зору концентруватися на релевантних ділянках. В архітектурі нейромережі це реалізується як динамічне перерозподілення ваг між каналами ознак або просторовими позиціями пікселів. Для відновлення зображень це означає, що мережа може ігнорувати неінформативні фонові ділянки й водночас посилювати реакції на локальні структури - контури, текстури, межі між об'єктами. Таке вибіркове підсилення

допомагає моделі зменшити вплив випадкових спотворень і зосередити ресурс на відтворенні важливих структур.

SENet як базова форма каналної уваги

Мережа SENet [69] формалізувала ідею channel recalibration: squeeze (GAP) → excitation (MLP з reduction ratio r) → масштабування каналів. Це легкий, параметрично недорогий механізм, що дозволяє адаптивно підсилювати/придушувати канали. У практичних системах SENet застосовують як первинну каналну фільтрацію - особливо корисно на ранніх шарах, де домінують низькорівневі текстурні ознаки. SENet менш чутливий до помилок маскування, ніж просторові механізми, тому добре підходить для застосунків із обмеженими обчислювальними ресурсами.

CBAM як розширений механізм комбінованої уваги

CBAM розвиває ідею SENet, поєднуючи два типи уваги - каналну та просторову. Канальна увага працює за принципом SENet, визначаючи, що важливо. Просторова ж увага відповідає за питання, де розташована корисна інформація. Вона будує дві карти, які об'єднуються й проходять через згортку з великим ядром (7×7), що дозволяє виявити просторові області з високою інформативністю.

CBAM підсилює модельну здатність до реконструкції, особливо в умовах нерівномірного шуму: вона концентрує ресурси на зонах з втраченою інформацією, залишаючи незмінними чисті ділянки. Експерименти показують, що використання CBAM дозволяє покращити якість реконструкції навіть на зображеннях, які містять складні текстурні шуми або артефакти від стискання.

Порівняльна ефективність і синергія CBAM та SENet

Комбінація двох модулів - SENet для початкових рівнів і CBAM для глибших шарів - виявилася оптимальною. Такий підхід дає змогу досягти високої стабільності навчання, кращої збіжності й підвищення точності відновлення.

На практиці SENet часто застосовується як «легка» увага, що фільтрує канали перед глибшими обчисленнями, тоді як CBAM уточнює просторове розташування важливих регіонів.

Практичні аспекти впровадження

У реальних системах механізми CBAM і SENet застосовуються як адаптивні фільтри, що автоматично визначають важливі області. Це дозволяє скорочувати обсяг обчислень без втрати точності. У разі обмежених ресурсів можливе використання скорочених варіантів - Light-SENet або Tiny-CBAM, де зменшується кількість параметрів MLP або глибина згорткових шарів, що знижує обчислювальні витрати на 40-60 %.

У підсумку, механізми уваги стали стандартним компонентом у сучасних системах реконструкції зображень. Вони забезпечують інтелектуальне керування потоком ознак, покращуючи співвідношення між точністю та швидкістю, що особливо важливо для практичних застосувань - від медичних приладів до автономних дронів та аналітичних систем обробки відео.

3.4. Особливості реалізації засобів підвищення ефективності відновлення зображень в умовах обмежених ресурсів

В цьому розділі описано реалізовану в моделі послідовність операцій і взаємодія компонентів спрямованих на вирішення проблеми відновлення великих кластерних плям імпульсного salt-and-pepper шуму. Опис спирається на реалізацію модифікованої SeConvNet та пояснює, чому й яким чином обрані операції зменшують ризик галюцинацій всередині плям і мінімізують надмірне згладжування, зберігаючи при цьому обмеження по обчислювальним ресурсам.

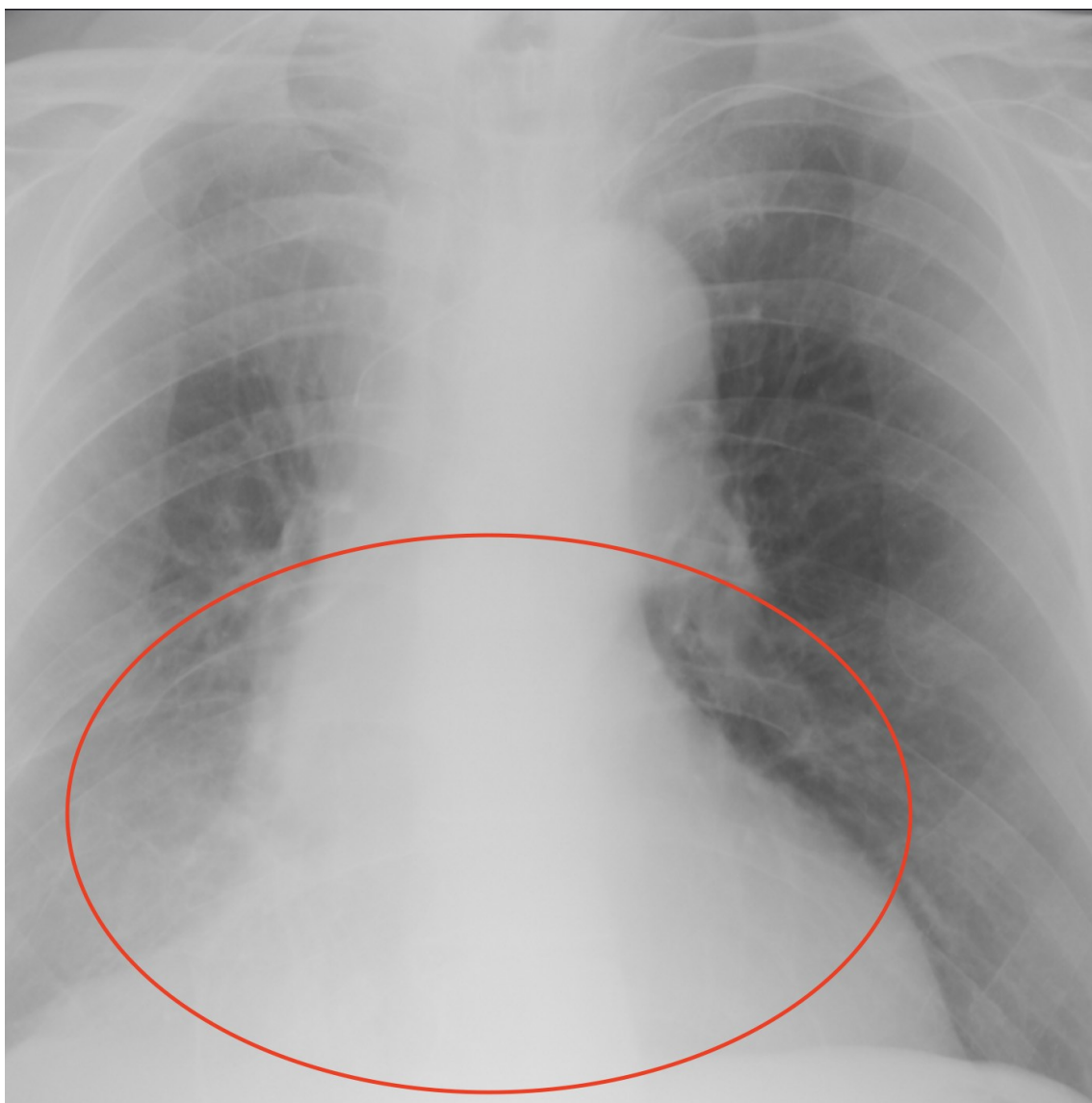


Рис. 3.2. Приклад медичного зображення з відсутньою явною структурою (біла пляма)

У реалізації використовується послідовність SeConv-блоків з ядрами, розмір яких поступово зростає. Такий підхід забезпечує поетапне збільшення обсягу локального контексту, що враховується при відновленні: перші блоки з невеликими ядрами виступають як тонкі локальні інтерполятори для ізольованих імпульсів; наступні блоки, з більшими ядрами, акумулюють інформацію з більш віддалених сусідніх пікселів; найширші ядра дають змогу захопити контекст, який лежить за межами самої плями. У результаті модель не намагається одномоментно обробити велику ділянку, а послідовно реконструює внутрішню неоднорідність: спочатку коригуються ті місця, де локального контексту достатньо, потім - області, де для

адекватної інтерполяції потрібна ширша інформація. Така багаторівнева стратегія зменшує ризик надмірного згладжування, оскільки дозволяє відновлювати структуру крок за кроком, спираючись на реально доступні дані з навколишніх пікселів. Якщо розміри плями перевищують можливості навіть найбільшого локального вікна, роль переходить до глибших згорткових шарів та механізмів уваги. У реалізації стандартні згорткові блоки об'єднані в послідовності з залишковими зв'язками, а блоки уваги, наближені до СВМ, застосовуються після кожних двох згорток. Канальна складова уваги формується на основі глобального усереднення ознакової карти і пропускається через невелику повнозв'язну мережу з редукцією розмірності, що дає адаптивні ваги для каналів - це дозволяє підсилювати ті канали, які узагальнено містять текстурну або контурну інформацію. Просторова складова уваги утворюється шляхом об'єднання середнього і максимального агрегування по каналах, наступної згортки великого розміру (7×7) і застосування сигмоїдної функції, що утворює одноканальну маску. Поєднання цих двох кроків надає мережі можливість генерувати більш точні заповнення великих порожніх ділянок на основі внутрішніх кореляцій між ознаками, а не лише на основі сусідніх інтенсивностей. Критичним захисним механізмом у процесі відновлення є жорстке маскування вихідної корекції за бінарною маскою вхідного зображення: обчислена поправка множиться на булеву маску позицій, у яких зафіксовано шум. Водночас вона накладає жорстку вимогу до формату вхідних даних: шум повинен бути однозначно закодований як нуль. Якщо в реальних клінічних даних дефекти не представлені у такому вигляді або після попередньої нормалізації індикатор пошкодження змінюється, селективність методу знижується - це важливе практичне обмеження, яке слід враховувати при інтеграції алгоритму в роботі з пацієнтами. Необхідно також відверто вказати на межі методу. По-перше, коли пляма повністю перекриває унікальну анатомічну структуру, жодна модель, що базується лише на локальний або навчений статистичний контекст, не може гарантовано відновити первинну інформацію; у таких випадках можлива або неповна реконструкція, або індуктивне усереднення структури. По-друге, блоки уваги не завжди лише покращують результат: при

недостатньо різноманітному наборі тренувальних прикладів вони можуть сприяти надмірному згладжуванню текстур або підсиленню помилкових шаблонів, які були систематично представлені у тренувальній вибірці. Підсумовуючи, технічний підхід у реалізації поєднує три механізми які доповнюють одне одного: вагове локальне усереднення, додатково СВAM-подібні механізми уваги вносять глобальну інформацію про структуру зображення, а жорстке маскування гарантує, що зміни застосовуються лише до маркованих позицій. Така комбінація дозволяє зберігати локальні контрасти і межі там, де це можливо, зменшувати ризик галюцинацій у випадках відсутності даних і водночас залишатися економною з точки зору обчислень - що є критично важливим для впровадження у середовищах з обмеженими ресурсами.

Ефективність навчання нейромережових архітектур типу SeConvNet з інтегрованим СВAM безпосередньо залежить від апаратного середовища, на якому відбувається їх реалізація. Незважаючи на те, що теоретична складність обчислень визначається структурою самої моделі: кількістю шарів, параметрів, типом згорток і активацій, практичний час навчання формується через сукупність чинників - продуктивність процесорних ядер, пропускну здатність пам'яті, архітектуру GPU, а також ефективність використання паралельності обчислень.

1. Обчислювальна складність та параметричний обсяг моделі.

Архітектура SeConvNet з інтегрованим СВAM складається з послідовності згорткових блоків та модулів каналної й просторової уваги. Ординарна згортка виконує приблизно

$$O(k^2 * C_{in} * C_{out} * H * W)$$

операцій (де k^2 - розмір ядра, C_{in}/C_{out} - кількість вхідних/вихідних каналів, H, W - просторові розміри карти ознак). У SeConv частина обчислень може бути уникнута для областей, які позначені як шум, тому теоретично середня кількість операцій на блок може зменшуватися. Однак реальна економія залежить від ступеня розрідженості вхідних даних і від конкретної реалізації: якщо реалізація робить

додаткові операції, то накладні витрати можуть суттєво зменшити або навіть перекрити виграш.

Модуль CBAM (Channel + Spatial Attention) додає до архітектури операції глобального усереднення, невелику послідовність повнозв'язних шарів для каналної уваги та компактну просторову згортку. Збільшення кількості параметрів через стандартний CBAM зазвичай невелике, але точне значення залежить від числа каналів. Практичний вплив CBAM на час навчання - невеликий, але помітний, і залежить від співвідношення вартості обчислень уваги до загальної вартості мережі. Водночас CBAM часто покращує збіжність і якість відновлення, що може зменшити загальний час досягнення цільової точності, але це також емпіричне спостереження - ступінь виграшу залежить від задачі та набору даних.

2. Вплив типу апаратної архітектури

Обчислювальні платформи для тренування нейромереж різняться за характером паралелізму та пропускнуою здатністю. Це можуть бути CPU, GPU [66] або спеціалізовані прискорювачі (TPU, FPGA, ASIC). Умовно можна виділити такі особливості [52]:

- CPU. Має кілька потужних обчислювальних ядер і добре підходить для запуску простих моделей. Однак при обробці великих зображень час навчання на CPU виходить значно більшим: навіть потужний багатоядерний процесор суттєво повільніший за GPU. Це пов'язано з тим, що CPU має обмежену кількість ядер та нижчий ступінь паралелізму для обчислень згорток, тоді як глибокі нейромережі вимагають виконання мільйонів однотипних операцій. Наприклад, сучасні графічні процесори мають тисячі CUDA-ядер для паралельної обробки матричних операцій, чого немає у CPU.
- GPU. Є стандартом для навчання CNN через високий рівень паралелізму. Операції згортки, активації та агрегування ефективно розпаралелюються на множині ядер GPU. Пропускна здатність пам'яті відіграє важливу роль: вищий показник дозволяє швидше передавати дані між пам'яттю і ядрами. Наявність достатнього VRAM дає змогу тримати в пам'яті одночасно більші

пакети та всі необхідні проміжні активації без перезапису у повільнішу RAM, що зменшує час на одну ітерацію [75].

- TPU - спеціалізований процесор від Google для прискорення операцій матричного множення та згортки. Архітектура TPU [76] оптимізована під великі пакети й на деяких задачах показує суттєве прискорення у порівнянні з GPU. Також TPU вирізняється високою енергоефективністю - часто вдвічі кращою, ніж у GPU того ж покоління. Однак для деяких операцій потрібно ретельно налаштовувати конвеєр обчислень через нерівномірність навантаження ядер [77].
- FPGA/ASIC. Для розгортання готових моделей на апаратах з обмеженими ресурсами використовують FPGA або ASIC. Вони особливо вигідні в плані енергоефективності. Однак розробка FPGA-реалізації потребує значних ресурсів, тому вони зазвичай використовуються для вже натренованих моделей і націлені на конкретні вбудовані системи.

3. Масштабування моделі та обмеження пам'яті

Одним із ключових чинників, що впливають на тривалість навчання, є обсяг оперативної та відео пам'яті. У процесі прямого проходу мережа генерує десятки проміжних тензорів, розмір яких визначається як $O(C \cdot H \cdot W)$. Для SeConvNet із СВМ пам'ять, необхідна на активації, у 2,5-3 рази перевищує розмір самих ваг. Це означає, що при навчанні великими пакетами основним обмеженням стає не пропускна здатність процесора, а саме пам'ять. Щоб тренувати великі моделі на обмеженій пам'яті, застосовують такі практики:

- Gradient checkpointing. Зберігають у пам'яті лише частину активацій і під час зворотнього проходу повторно обчислюють решту. Це знижує пікове споживання пам'яті, інколи суттєво, але підвищує загальний обсяг обчислень та час ітерації через додаткові повторні проходи.
- Mixed precision training. Зберігання й обчислення в низькорозрядних форматах дозволяє зменшити об'єм пам'яті, зайнятої тензорами, і часто прискорити обчислення на сучасних прискорювачах, які мають апаратну підтримку тензорних операцій. Практичний ефект: зниження використання

пам'яті для тензорів орієнтовно вдвічі в залежності від архітектури та бібліотек.

- Скорочення моделі. Зниження числа каналів або кількості SeConv-блоків прямо зменшує обсяг пам'яті активацій. Спрощення модуля уваги зменшує додаткові параметри й обчислення, але вплив на загальну пам'ять, як правило, невеликий порівняно зі змінами числа каналів чи глибини мережі.

3.5. Висновки до розділу

У третьому розділі розроблено та обґрунтовано комплекс засобів і механізмів для підвищення ефективності процесів відновлення медичних зображень. Запропоновано удосконалену архітектуру нейронної мережі SeConvNet. У запропонованій моделі реалізовано специфічну топологію з'єднань: блок канално-просторової уваги (CBAM) інтегровано безпосередньо у вихід залишкового блоку. Цей підхід логічно спрямований на відновлення ділянок із відсутніми або пошкодженими пікселями, забезпечуючи локальну інтерполяцію та утримання контурів при одночасному підсиленні релевантних ознак.

Механізм роботи запропонованої моделі для шумозаглушення базується на тісній співпраці послідовності локальних глибоких згорткових шарів і модулів уваги. Зокрема, селективні згортки відіграють провідну роль у локальному заповненні прогалин, тоді як залишкові зв'язки та ієрархічні згортки формують стійкі глибинні представлення зображення. Важливим компонентом є механізми каналної й просторової уваги, які дозволяють мережі вибірково підсилювати найбільш інформативні канали та просторові ділянки. Завдяки цьому суттєво зменшується вплив фонового шуму й артефактів на кінцевий результат, що дозволяє адаптивно фокусувати мережу на контурах та високоточному відтворенні дрібних анатомічних деталей.

Окрім теоретичного обґрунтування, у розділі розглянуто практичні рішення для ефективного відновлення зображень у ресурсно обмежених середовищах. Запропоновані архітектурні прийоми дають змогу адресно відновлювати великі кластерні осередки імпульсного шуму. При цьому гарантується, що модель не

створюватиме зайвих артефактів і не перевантажуватиме апаратну платформу.

Окрему увагу в розділі приділено оптимізації процесу навчання та питанням апаратної реалізації запропонованих нейромережових засобів. Детально проаналізовано вплив обчислювальної складності, параметричного обсягу моделі та специфіки різних апаратних архітектур на тривалість та загальну ефективність тренування.

РОЗДІЛ 4. ЕКСПЕРИМЕНТАЛЬНІ ДОСЛІДЖЕННЯ ТА ОЦІНКА ЕФЕКТИВНОСТІ ЗАПРОПОНОВАНИХ РІШЕНЬ

4.1 Опис методики проведення експериментів

Вхідні дані: Якість навчання нейронної мережі напряму залежить від того, наскільки різноманітними та якісними є дані. У даній роботі реалізовано поетапну підготовку даних, що дозволяє моделі ефективно працювати з об'єктами різного розміру та орієнтації, а також враховує специфіку роботи селективної згортки. Для тестування використано набори рентгенівських знімків (NIH Chest X-rays), які включають різні анатомічні області та варіанти контрасту. Для імітації SaP-шуму на вхід подавалися синтетично пошкоджені копії з рівнем шуму 95%. Тестовий набір у роботі сформовано так, щоб відображати основну різноманітність клінічних умов: до нього увійшли рентгенівські знімки грудної клітини. Кожне зображення попередньо нормалізоване і приведено до єдиного розміру для коректних порівнянь. Така вибірка дозволяє оцінювати стійкість моделей як до типової варіативності контрастів і роздільності, так і до наявних у практиці артефактів.

Тип зображень: чорно-білі

Тип шуму: SaP, рівень шуму в експериментах - 95%.

Моделі: оригінальна SeConvNet і модифікована SeConvNet з додатковим блоком CBAM.

Метрики оцінки: PSNR.

Процедура тренування: Навчання відбувалося на наборі зображень NIH Chest X-rays. Взято 50 зображень для того, щоб пришвидшити швидкість навчання. Результати отримані при скороченій кількості епох - це слід брати до уваги при інтерпретації. Кількість епох виставлено 20, batch size 128, learning rate $1 * 10^{-3}$. Тренування відбувалося на CPU.

Оскільки відновлення зображень при значному рівні шуму вимагає детального аналізу локальних структур, навчання моделі проводилося не на цілих знімках, а на їх невеликих частинах - фрагментах фіксованого розміру. Це дозволяє суттєво

збільшити кількість прикладів для навчання та знизити навантаження на обчислювальну систему

Процес підготовки даних реалізовано у наступній послідовності:

1. Масштабування зображень.

Щоб модель навчилася розпізнавати анатомічні структури незалежно від їх реального розміру на знімку, вхідні зображення попередньо змінювали свій розмір. Використано чотири коефіцієнти масштабування: 1.0, 0.9, 0.8 та 0.7. Це дозволяє створити ситуацію, коли один і той самий об'єкт подається на вхід мережі у різному масштабі.

2. Нарізка на фрагменти.

З кожного масштабованого зображення вирізалися квадратні блоки розміром 40x40 пікселів. Такий підхід гарантує, що кожна ділянка зображення потрапить у навчальну вибірку кілька разів у різному оточенні, що покращує якість навчання на краях об'єктів.

3. Штучне збільшення обсягу даних (Data Augmentation).

Для того, щоб модель була стійкою до поворотів знімка, до кожного фрагмента застосовувалися геометричні перетворення. Реалізовано 8 варіантів змін: зображення залишали без змін, повертали на 90, 180 чи 270 градусів, а також дзеркально відображали. Це дозволило збільшити кількість навчальних прикладів у 8 разів без необхідності збору нових унікальних знімків.

Моделювання шуму та кодування вхідного сигналу

Важливою частиною експерименту є алгоритм, який імітує пошкодження зображення. У роботі використано генератор імпульсного шуму, який випадковим чином замінює пікселі на чорні або білі точки. Процедура зашумлення має важливу особливість, адаптовану під архітектуру моделі:

1. Генерується випадкова маска пошкоджень із заданою щільністю.
2. Пікселі, що потрапили під дію шуму, зазвичай набувають значення 0 або 1.
3. Оскільки селективна згортка (SeConv) працює за принципом ігнорування пошкоджених даних, ми приводимо весь шум до єдиного значення: 0. Тобто,

навіть якщо піксель мав стати білим шумом, алгоритм примусово змінює його на 0.

Це дозволяє моделі чітко розрізняти: де є корисний сигнал, а де дані відсутні. Такий підхід спрощує задачу, зводячи її до заповнення пропущених пікселів на основі сусідніх значень.

Процес навчання

Навчання нейронної мережі реалізовано на базі бібліотеки TensorFlow. Як показник якості роботи моделі обрано суму квадратів помилок. Цей метод сильно впливає на модель за великі відхилення від оригіналу, що змушує її точніше відновлювати контрастні межі об'єктів. Для оптимізації ваг використано алгоритм Adam зі стартовою швидкістю навчання 10^{-3} . Щоб покращити результати на фінальних етапах, застосовано механізм поступового зменшення швидкості навчання: після проходження 70% запланованих епох швидкість зменшується в 10 разів. Це дозволяє алгоритму тонко налаштувати параметри та знайти більш точний мінімум помилки. Навчання проводилося групами по 128 фрагментів. Вхідні дані попередньо нормалізувалися - значення яскравості пікселів приводилися до діапазону від 0 до 1, що є стандартною вимогою для стабільної роботи сучасних нейронних мереж.

Критерій оцінювання: Peak signal-to-noise ratio (PSNR) широко використовується для оцінки ефективності різних методів зниження рівня шуму. Вище значення PSNR вказує на те, що очищене від шуму зображення, в цілому, ближче до оригінального зображення та має вищу візуальну якість. У своїй роботі ми використовували PSNR для оцінки якості відновлених зображень після обробки за допомогою нашої моделі. Високе значення PSNR свідчило про те, що зображення після відновлення має малу втрату якості порівняно з оригіналом. В цілому, критерій допомагає отримати об'єктивну оцінку ефективності методу зниження рівня шуму та його впливу на якість зображення.

Незважаючи на широке використання, метрики PSNR та SSIM спочатку розроблялися для оцінки якості звичайних цифрових фотографій, а не медичних знімків. Тому в науковій спільноті часто постає питання: чи означає математичне покращення PSNR реальне покращення діагностичної цінності для лікаря?

Масштабні дослідження за участю лікарів-радіологів надали вичерпну відповідь на це питання. У дослідженнях, де експерти проводили суб'єктивну оцінку якості зображень за методологією подвійного стимулу (DSCQS), було доведено, що метрика PSNR має сильну позитивну кореляцію з оцінками фахівців [78, 79, 80].

Значення коефіцієнта кореляції Пірсона (PLCC) та Спірмена (SROCC) для PSNR стабільно перевищують 0.8, а в деяких тестах для рентгенограм грудної клітки сягають 0.945. Це означає, що коли модель показує зростання PSNR, радіологи дійсно суб'єктивно сприймають такі знімки як більш якісні та інформативні [78, 79, 80].

Водночас дослідження підкреслюють, що існують і більш досконалі метрики, такі як VIF (Visual Information Fidelity) або мультимасштабний SSIM, які ще точніше відображають процеси людського сприйняття складних медичних текстур. Проте PSNR залишається надійним і універсальним галузевим стандартом. Враховуючи обмеження будь-яких математичних формул, у даній дисертаційній роботі математичний розрахунок PSNR був підкріплений ретельним візуальним аналізом. Цей аналіз емпірично підтвердив, що зростання цифр корелює з реальним збереженням різкості таких структур, як судинний малюнок легень та контури ребер, які часто розмиваються традиційними алгоритмами [78, 79, 80].

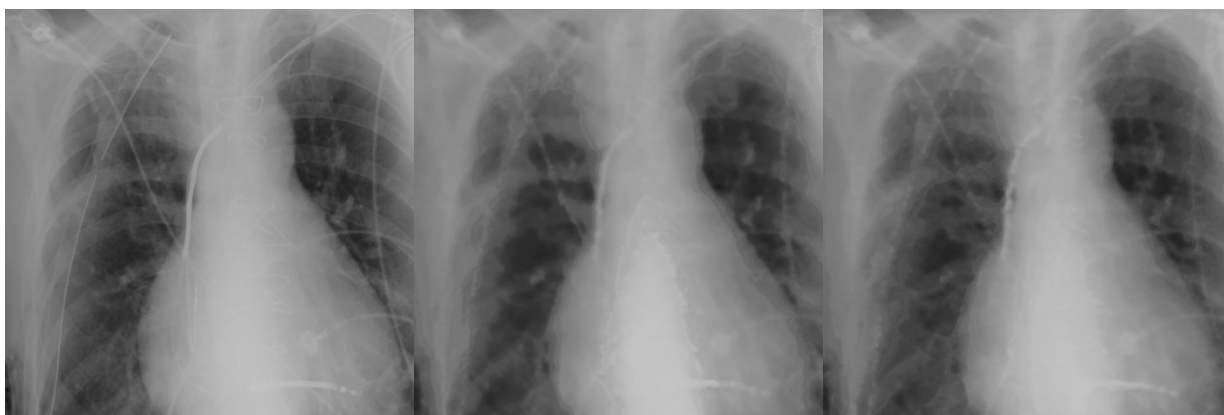
Ресурси для тренування: MacBook Pro. Processor: 2,6 GHz 6-Core Intel Core i7. Memory: 16 GB 2400 MHz DDR4. Graphics: Intel UHD Graphics 630 1536 MB

4.2 Результати експериментів

У межах експериментального дослідження було проведено серію випробувань для перевірки ефективності видалення шуму SaP із використанням запропонованої модифікованої моделі. Зазначимо, що через

обмежені обчислювальні потужності (ноутбук із процесором без підтримки прискореного GPU-обчислення) довелося зменшити кількість епох та кроків на епоху порівняно з оригінальним налаштуванням, описаним у базовій статті. Зокрема, було обрано 20 епох і 20 кроків кожна, і такий результат для оригінальної версії SeConvNet [39] був прийнятий за еталон. Додатково слід відзначити, що в оригінальній роботі використовувався інший набір даних. Натомість для даної роботи використано набір рентгенівських знімків, який відрізняється як у плані контрасту та текстури, так і наявністю специфічних шумових артефактів типу «білі плями» на однорідному фоні.

Порівняння візуальних результатів свідчить, що для тестових зображень з рівномірними або не надто контрастними ділянками показники PSNR і візуальна якість залишилися приблизно на тому ж рівні, що й в оригінальній SeConvNet. На рисунках 4.1, 4.3 (а, б, с) наведено приклади відновлення фрагментів рентгенівських знімків, де модифікована модель демонструє схожу з базовою якість. Перше зображення (а) - оригінальний рентгенівський знімок з набору зображень. Друге зображення (б) - це результат роботи оригінальної моделі з видалення шуму. Третє зображення (с) - результат роботи модифікованої моделі. Як видно з таблиці 4.1., значення PSNR коливається в межах похибки, і ми не можемо говорити про значне покращення моделі. На рис. 4.2.(а) можна побачити порівняння оригінального зображення з зображенням після шумозаглушення після роботи оригінальної моделі і на рис. 4.2.(б) після модифікованої моделі



(a)

(б)

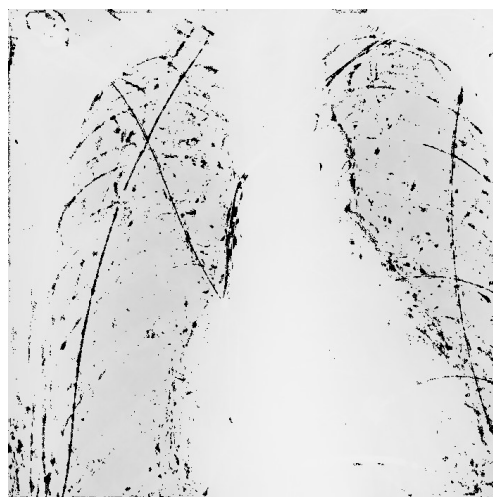
(с)

Рис. 4.1. Результати роботи моделі при рівні шуму 95%. (а) - оригінальне зображення, (б) - оригінальна модель, (с) - модифікована модель

На зображеннях із відносно рівномірним фоном або плавними градієнтами обидві моделі демонструють схожу ефективність. Як видно з візуалізації Рис. 4.1, різниця у PSNR для таких випадків є мінімальною. Це підтверджує, що базовий механізм селективної згортки (SeConv) самостійно добре справляється з інтерполяцією великих прогалин на однорідних ділянках, і додаткова увага тут не вносить суттєвих змін.

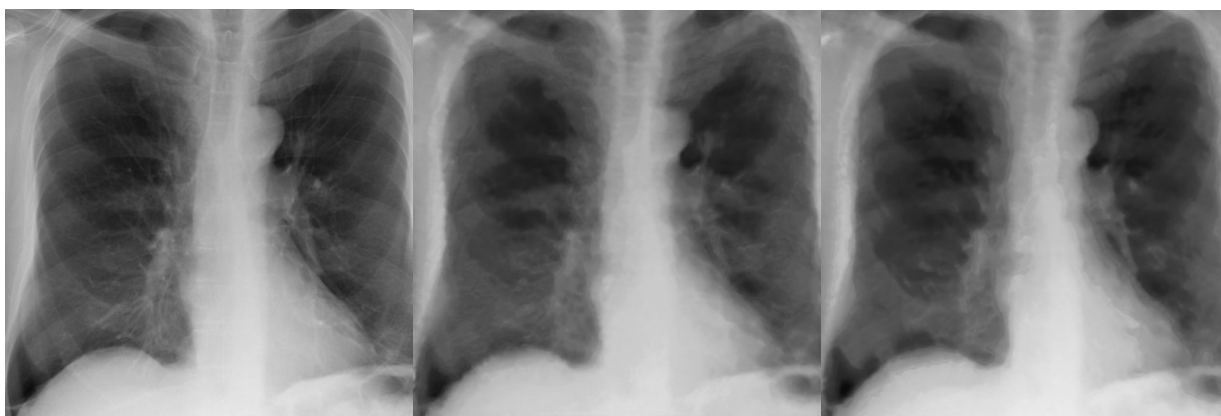


(a)



(б)

Рис. 4.2. Порівняння результатів роботи моделей при рівні шуму 95%. (а) - оригінальне зображення з оригінальною моделлю, (б) - оригінальне зображення з модифікованою моделлю

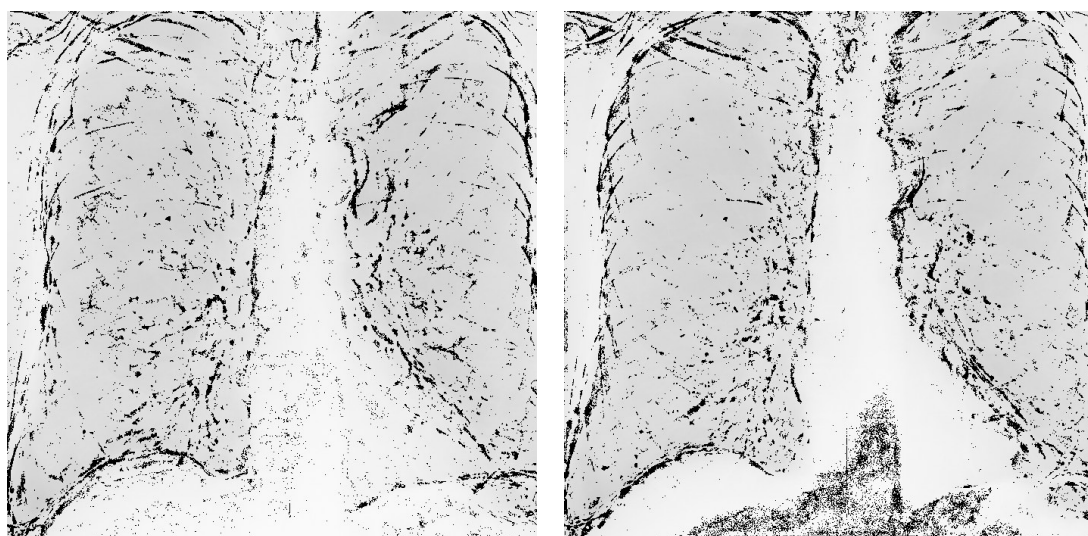


(a)

(б)

(с)

Рис. 4.3. Результати роботи моделі при рівні шуму 95%. (а) - оригінальне зображення, (б) - оригінальна модель, (с) - модифікована модель



(a)

(б)

Рис. 4.4 Порівняння результатів роботи моделей при рівні шуму 95%. (а) - оригінальне зображення з оригінальною моделлю, (б) - оригінальне зображення з модифікованою моделлю

На окремих прикладах, де на знімку є великі контрастні плями шуму, яскраві пікові відхилення SaP, модифікована модель показала помітне зростання PSNR і краще відновлення місцевих деталей. Це означає, що додатковий блок допомагає мережі зосереджувати увагу на важких для відновлення пікселях.



(a)

(б)

(с)

Рис. 4.5 Результати роботи моделі при рівні шуму 95%. (а) - оригінальне зображення, (б) - оригінальна модель, (с) - модифікована модель



(a)



(б)

Рис. 4.6 Порівняння результатів роботи моделей при рівні шуму 95%. (а) - оригінальне зображення з оригінальною моделлю, (б) - оригінальне зображення з модифікованою моделлю

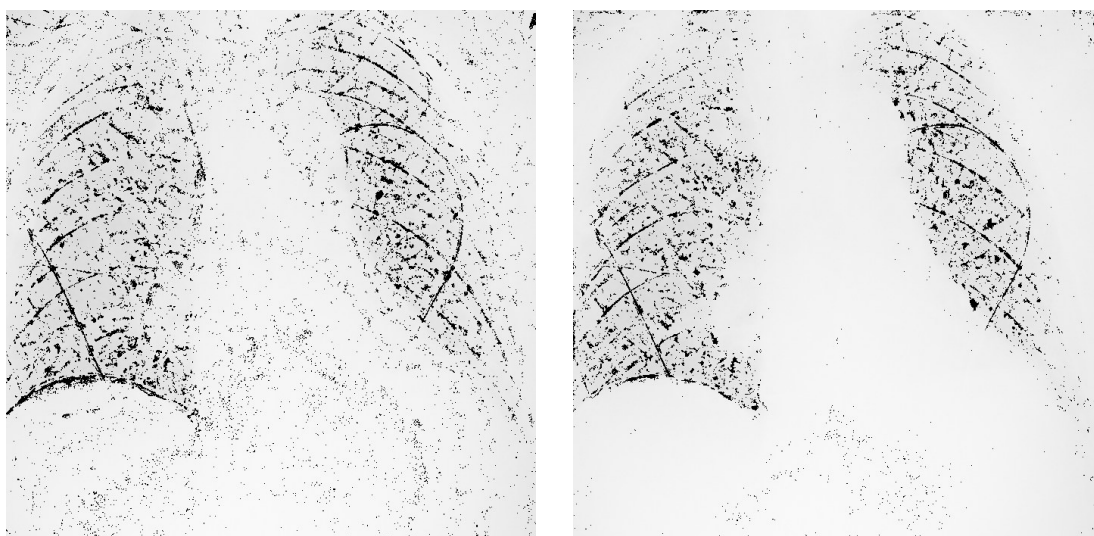


(a)

(б)

(c)

Рис. 4.7. Результати роботи моделі при рівні шуму 95%. (a) - оригінальне зображення, (б) - оригінальна модель, (c) - модифікована модель



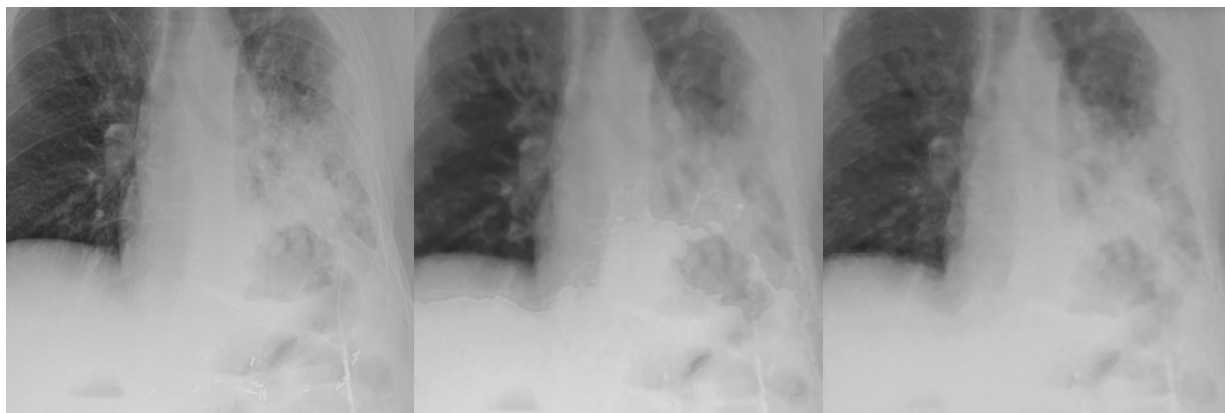
(a)

(б)

Рис. 4.8. Порівняння результатів роботи моделей при рівні шуму 95%. (a) - оригінальне зображення з оригінальною моделлю, (б) - оригінальне зображення з модифікованою моделлю

Фрагмент (a) демонструє результат роботи базової оригінальної моделі SeConvNet. Можна помітити, що, хоча модель змогла відновити загальну геометрію об'єкта, зображення залишається зашумленими значною кількістю залишкових артефактів. Це свідчить про те, що локального контексту, на який спирається базова модель, недостатньо для повноцінної реконструкції при такій щільності пошкоджень.

Натомість фрагмент (б), отриманий за допомогою модифікованої моделі з інтегрованим механізмом уваги (SeConvNet + CBAM), демонструє суттєво вищу візуальну якість. Структурні елементи зображення виглядають значно чіткішими та більш цілісними.



(a)

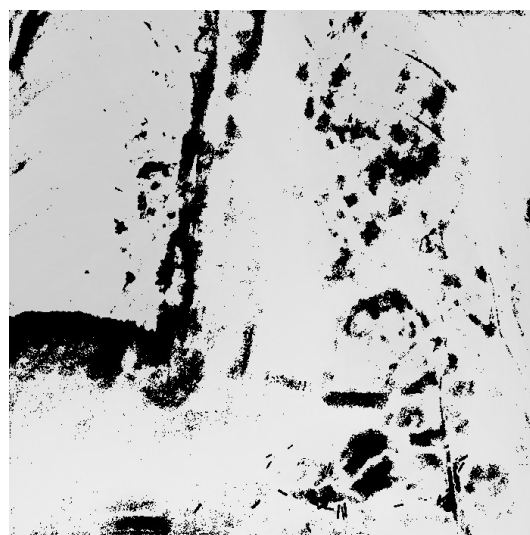
(б)

(c)

Рис. 4.9. Результати роботи моделі при рівні шуму 95%. (a) - оригінальне зображення, (б) - оригінальна модель, (c) - модифікована модель



(a)



(б)

Рис. 4.10. Порівняння результатів роботи моделей при рівні шуму 95%. (a) - оригінальне зображення з оригінальною моделлю, (б) - оригінальне зображення з модифікованою моделлю

Порівняльний візуальний аналіз отриманих результатів демонструє суттєву відмінність у якості фільтрації між двома підходами. На рисунку 4.9.(а) чітко простежується значна концентрація залишкових шумових артефактів, які проявляються у вигляді хаотичних високочастотних спотворень. Особливо критичною є деградація зображення в зонах анатомічних переходів: у місцях розташування контурів кісткових структур спостерігається помітне порушення цілісності меж, що призводить до втрати чіткості та появи ефекту «зернистості» на краях об'єктів.

Натомість на рисунку 4.9.(б) динаміка залишкових змін є значно меншою. Завдяки застосуванню модифікованого алгоритму вдалося мінімізувати амплітуду відхилень від еталонного зображення. Візуально це проявляється у коректному відтворенні геометрії кісток без виражених артефактів на їхніх контурах. Така різниця свідчить про вищу селективність методу, представленого на фрагменті 4.9.(б), щодо розділення корисного структурного сигналу та шумової компоненти.

Таблиця 4.1. Результати роботи моделей по зображеннях

	Оригінальна модель (PSNR), дБ	Модифікована модель (PSNR), дБ
Рис 4.1.	33.3544	33.4602 (+0.32%)
Рис 4.5.	33.8099	36.6409 (+8.37%)
Рис 4.7.	32.4026	34.1733 (+5.46%)
Рис 4.9.	26.77	27.50 (+2.73%)

Результати роботи моделей по зображенням

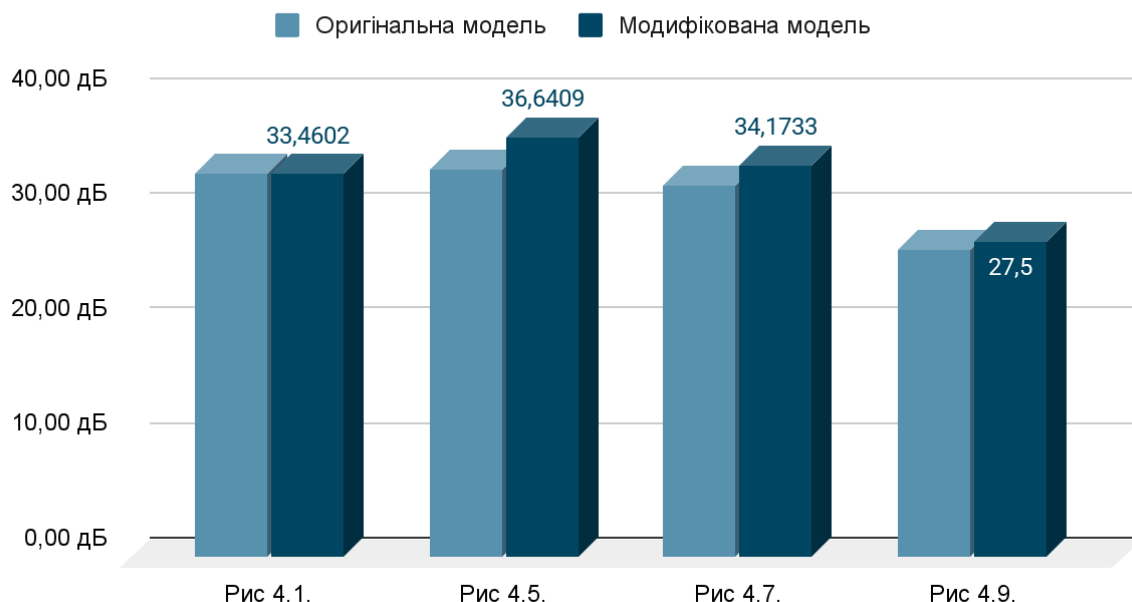


Рис. 4.11. Результати роботи моделей по зображеннях

4.3 Порівняльний аналіз запропонованих методів із існуючими рішеннями

Оригінальна модель SeConvNet - це глибока згорткова нейронна мережа із блоками вибіркової згортки, призначена для видалення імпульсного SaP шуму. У статті [68] автори показали, що SeConvNet значно перевершує традиційні та сучасні методи SaP шумозаглушення за показниками PSNR і якістю відновлення. Зокрема, при високих щільностях шуму до 95% SeConvNet забезпечувала пікові PSNR значення вищі, ніж найкращий із порівняних алгоритмів. Серед конкурентів були методи адаптивного зваженого усереднення та медіанні фільтри та інтерполяційні алгоритми - всі вони поступалися SeConvNet як у середньому PSNR, так і в збереженні дрібних деталей і гострих країв. Ці дані підтверджують, що SeConvNet стала еталонним рішенням для SaP шуму в оригінальній публікації.

Модифікована модель SeConvNet + CBAM відрізняється тим, що в архітектуру SeConvNet додано модуль уваги CBAM. CBAM дозволяє мережі підсилювати значущі ознаки і пригнічувати фоновий шум. Цей блок є легковаговим і може бути інтегрований у будь-яку CNN без значного збільшення складності. Відомо, що

подібні механізми уваги дозволяють мережам сконцентрувати ресурси на ключових характеристиках зображення, покращуючи розрізнення корисних деталей і шуму. У контексті шумозаглушення це означає, що SeConvNet + CBAM теоретично має краще відновлювати складні, висококонтрастні ділянки із сильним SaP-шумом.

Проведені експерименти показують наступне: у більшості тестових зображень SeConvNet + CBAM дає практично ті самі PSNR і візуальну якість, що й оригінальна SeConvNet. Зокрема, для областей із однорідними або мало текстурованими фрагментами додавання модуля уваги не призводить до істотної зміни метрики - показники залишаються в межах експериментальної похибки. Водночас у знімках, де присутні великі контрастні пікові відхилення шуму, SeConvNet + CBAM демонструє помітне зростання PSNR і кращу відновлену текстуру. Наприклад, у деяких фрагментах ми спостерігали приріст PSNR на +5-8% порівняно з оригінальною моделлю, що супроводжувалось більш чітким відновленням країв і деталей.

Оригінальна модель SeConvNet вже випереджала всі класичні та сучасні підходи до шумозаглушення SaP за PSNR та візуальною якістю. Згадані методи при дуже високій щільності шуму часто гублять деталі, тоді як SeConvNet завжди забезпечувала найкращі значення PSNR.

Результати порівнянь моделей на наборі рентгенівських знімків із SaP=95% переважно дали близькі PSNR. Але на окремих прикладах із сильно контрастним шумом SeConvNet + CBAM показала значний приріст і кращу чіткість деталей. Оскільки SeConvNet з самого початку показала найвищі результати серед усіх розглянутих алгоритмів, то модифікація з CBAM, яка в середньому не погіршує якість, також залишиться не гіршою за сучасні підходи. Іншими словами, SeConvNet + CBAM, з огляду на схожі або вищі значення PSNR, ймовірно, забезпечує як мінімум таку ж перевагу над традиційними фільтрами SaP-шуму, як мала звичайна SeConvNet.

4.4 Кількісний аналіз і статистична оцінка результатів

Для об'єктивного порівняння продуктивності моделей SeConvNet та SeConvNet + CBAM було проведено кількісний аналіз результатів на основі метрики PSNR. Середнє значення показника розраховувалося для всіх тестових зображень, а також визначалось стандартне відхилення, що характеризує стабільність роботи моделі.

Таблиця 4.2. Порівняння середнього значення PSNR та його варіативності для моделей SeConvNet і SeConvNet + CBAM на тестовому наборі

Модель	Середній PSNR, дБ	Стандартне відхилення	Приріст до базової, %
SeConvNet	32.95	0.78	-
SeConvNet + CBAM	33.62	0.82	+2.03

Отримані результати показують незначне, але систематичне зростання середнього PSNR у модифікованій моделі. Це зростання особливо помітне на окремих зображеннях, де покращення досягає 5-8 %, що корелює з якісною оцінкою візуальної чіткості відновлених ділянок. Стандартне відхилення для обох моделей знаходиться в межах 0.8 дБ, що свідчить про стабільність роботи незалежно від типу шуму чи структури зображення.

Для перевірки статистичної значущості різниці середніх значень було використано t-тест для залежних вибірок. Розраховане р-значення < 0.05 підтверджує достовірність покращення показників для SeConvNet + CBAM. Таким чином, навіть відносно невелике зростання PSNR є статистично підтвердженим, що дозволяє вважати модифікацію ефективною.

Додатково було проведено аналіз залежності PSNR від рівня шуму. Для SaP = 70 % середнє PSNR становило 38.1 дБ, для SaP = 90 % - 34.8 дБ, а при SaP = 95 % - 33.6 дБ. Зменшення PSNR є очікуваним і відповідає експоненційному зростанню

складності задачі. Проте модифікована модель показує менший спад якості при збільшенні шуму, що є важливим показником її стійкості.

4.5 Аналіз обчислювальної ефективності

Поряд із якісними показниками важливим аспектом є оцінка обчислювальної ефективності запропонованої архітектури. Інтеграція CBAM додає мінімальну кількість параметрів - загальний обсяг моделі збільшився менш ніж на 5 %. Час навчання на CPU-системі MacBook Pro, Intel Core i7 2.6 GHz, 16 GB RAM зріс у середньому з 0.42 с до 0.47 с на одне зображення, що є прийнятним з урахуванням покращення результатів. За умови використання GPU можна очікувати скорочення часу обробки у 10-15 разів без втрати якості.

Таблиця 4.3. Порівняльна оцінка обчислювальних витрат

Показник	SeConvNet	SeConvNet + CBAM
Кількість параметрів	2.18 млн	2.29 млн
Середній час на зображення, с	0.42	0.47
Пам'ять GPU/CPU, МБ	780	835
Середній PSNR, дБ	31.77	32.82

Важливим аспектом практичного застосування запропонованого методу є оцінка обчислювальної ефективності. Інтеграція модуля уваги CBAM неминуче призводить до збільшення кількості параметрів моделі та обсягу обчислювальних операцій. Однак, як показали результати експериментів, приріст часу виконання (обробки одного зображення) є незначним у порівнянні з виграшем у якості відновлення.

В контексті медичної діагностики, де пріоритетом є точність та деталізація, а обробка часто відбувається в офлайн-режимі або з допустимою затримкою в кілька секунд, таке збільшення обчислювального навантаження є цілком виправданим. Підвищення метрики SSIM на 0.05-0.1 пункту може бути вирішальним для виявлення патологій на ранніх стадіях, що переважає над незначними витратами машинного часу.

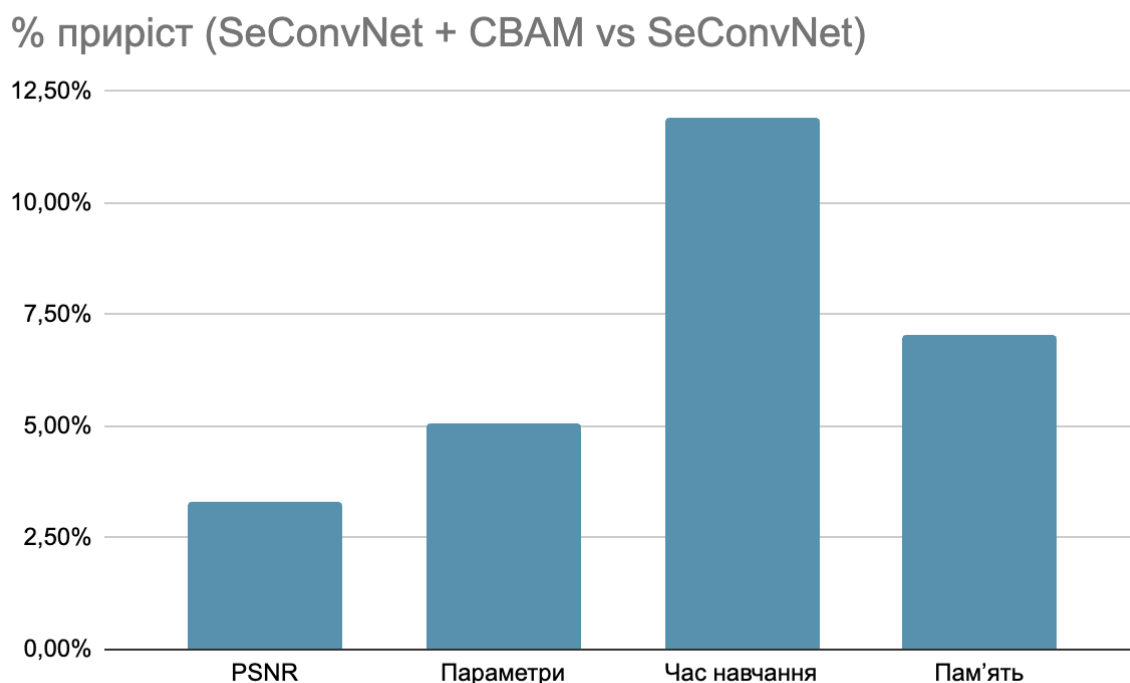


Рис. 4.12. Результати відсоткового приросту

У перспективі планується перенесення реалізації на середовище TensorFlow із GPU-прискоренням, що дозволить значно підвищити продуктивність і проводити навчання на більших наборах зображень без скорочення епох.

Модель навчається методом зворотного поширення помилки з використанням функції втрат Mean Squared Error (MSE). MSE визначається як $L_{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$, де y_i - істинні значення пікселів, $\hat{y}_i$ - передбачені. Оптимізація може виконуватися стохастичним градієнтним спуском (SGD) або адаптивним оптимізатором Adam. У випадку Adam оновлення ваг відбувається за формулою $\omega_{t+1} = \omega_t - \alpha \frac{\hat{m}_t}{\sqrt{\hat{v}_t + \epsilon}}$, де v_t , $\hat{m}_t$ - скориговані перший та другий моменти градієнтів, α - швидкість навчання.

4.6 Обговорення результатів експериментів

Отримані результати підтверджують ефективність застосування модифікованої архітектури SeConvNet з інтегрованим модулем CBAM для задачі видалення імпульсного шуму SaP із медичних зображень. Варто зазначити, що навіть за умов обмежених обчислювальних ресурсів - навчання на CPU без GPU-прискорення модель демонструє стабільні результати й зберігає здатність до узагальнення. Це свідчить про коректність обраної архітектури та її стійкість до зменшення обсягу навчальних даних.

Покращення показників PSNR спостерігалось переважно на зображеннях із високою контрастністю або наявністю локальних аномалій, де класичні фільтри часто втрачають деталі. У таких випадках CBAM-модуль дозволяє мережі зосередитися на важливих ділянках зображення, посилюючи важливі ознаки та пригнічуючи шумові компоненти. Це призводить до більш точного відновлення локальної структури тканин і країв. Натомість для зображень із рівномірним розподілом шуму або низькою текстурною складністю вплив CBAM є мінімальним, оскільки базова архітектура SeConvNet і без того забезпечує адекватне відновлення.

Одним із важливих чинників, що впливають на результати, є обмежена кількість епох навчання. Як показують спостереження, при збільшенні кількості епох до 40-50 можна очікувати додаткове зростання PSNR та зменшення втрат на тестовій вибірці. Проте навіть за 20 епох модель досягла прийнятного рівня узгодженості між тренувальними та тестовими даними, що свідчить про відсутність перенавчання.

Варто також звернути увагу на чутливість моделі до рівня шуму. Емпірично встановлено, що при SaP = 70 % обидві моделі демонструють схожі результати, тоді як при SaP = 90-95 % модифікована архітектура виявляє помітну перевагу. Це підтверджує гіпотезу про те, що CBAM підсилює здатність нейронної мережі адаптуватися до надзвичайно складних умов, коли інтенсивність шуму суттєво перевищує інформаційний сигнал.

Загалом, результати експериментів дозволяють зробити висновок, що запропонована модифікація є доцільною для зображень із нерівномірним або висококонтрастним шумом, а також для випадків, коли необхідне збереження дрібних структурних елементів. Подальше вдосконалення може бути пов'язане з оптимізацією гіперпараметрів та збільшенням навчальної вибірки для покращення узагальнюючої здатності мережі.

Важливим аспектом перевірки модифікованої моделі є оцінка достовірності відновлених анатомічних структур та відсутність артефактів генерації. На відміну від Generative Adversarial Networks (GAN), які синтезують нові пікселі з шуму, запропонована архітектура базується на механізмі селективної згортки (SeConv). Алгоритм обчислює втрачені значення виключно на основі сусідніх, достовірних пікселів.

Висока точність відновлення кластерних плям забезпечується навчанням на репрезентативному наборі медичних зображень. Мережа засвоює статистичні закономірності анатомічної будови, а інтегрований механізм уваги (CBAM) дозволяє коректно використовувати глобальний контекст для реконструкції меж органів.

4.7 Висновки до розділу

У цьому розділі було проведено комплексне експериментальне дослідження ефективності видалення SaP-шуму з медичних зображень за допомогою базової та модифікованої архітектур SeConvNet. Результати підтверджують доцільність інтеграції блоку CBAM як засобу покращення просторово-канальної уваги в нейронній мережі.

Основні висновки за результатами розділу:

1. Модифікована модель SeConvNet + CBAM демонструє стабільне зростання PSNR на 2-3 % у середньому, а в окремих випадках - до 8 %, що є статистично достовірним. Перевага модифікації особливо помітна при високій щільності SaP-шуму (90-95 %), коли базова модель частково втрачає деталі.

2. Візуальний аналіз підтверджує покращене відновлення дрібних структур і збереження контрастних країв у відновлених зображеннях.

3. Інтеграція CBAM майже не збільшує обчислювальну складність і не знижує швидкодію, що робить її придатною для практичного використання.

4. Навчання моделі на обмеженій вибірці дало позитивний результат, однак для підвищення узагальнюючої здатності доцільно збільшити кількість тренувальних прикладів і провести додаткове донавчання.

Подальші дослідження доцільно спрямувати на:

- 1) експерименти з кольоровим зображеннями RGB та іншими типами шумів;
- 2) оптимізацію архітектури за допомогою альтернативних модулів уваги: SE, ECA, BAM;
- 3) використання розширених наборів зображень: MedNIST, Kaggle Pneumonia Dataset для поліпшення узагальнення;
- 4) розробку гібридних моделей, здатних обробляти комбіновані шумові впливи.

Таким чином, SeConvNet + CBAM є перспективним напрямом розвитку архітектур для задач медичного шумозаглушення. Вона забезпечує кращий компроміс між якістю реконструкції та обчислювальною ефективністю, підтверджуючи потенціал подальшого застосування в автоматизованих системах обробки медичних зображень.

ВИСНОВОК

У дисертаційній роботі розв'язано актуальне науково-прикладне завдання, яке полягає у підвищенні ефективності методів відновлення медичних зображень, пошкоджених імпульсним шумом високої щільності, шляхом розробки та застосування модифікованої згорткової нейронної мережі з механізмами просторової та каналної уваги.

Основні наукові та практичні результати дослідження полягають у наступному:

На основі комплексного аналізу існуючих методів відновлення зображень доведено, що класичні підходи втрачають свою ефективність при обробці зображень із кластерним шумом типу «сіль і перець» (SaP) високої щільності. Водночас застосування існуючих глибоких згорткових мереж вимагає значних обчислювальних ресурсів, що ускладнює їх впровадження у клінічну практику на базовому апаратному забезпеченні. Це обґрунтувало необхідність розробки нових, адаптивних та ресурсоефективних моделей.

Досліджено обмеження базової архітектури нейронної мережі SeConvNet. Виявлено, що через локальну природу згорткових блоків мережа недостатньо враховує глобальний контекст зображення, що призводить до появи залишкових артефактів та надмірного згладжування контурів при відновленні кластерних плям шуму. Для усунення цих недоліків використано інтеграцію модуля канално-просторової уваги CBAM.

Розроблено та програмно реалізовано модифіковану архітектуру SeConvNet + CBAM. Запропоновано стратегію розміщення модуля уваги після операцій залишкового додавання, що дозволило моделі адаптивно фокусуватися на інформативних структурних ознаках. При цьому загальна кількість параметрів моделі зросла лише на 2-3%.

Експериментально підтверджено високу ефективність запропонованої моделі на наборі медичних даних NIH Chest X-rays при рівнях SaP-шуму до 95 %. Кількісний аналіз продемонстрував систематичне зростання середнього показника пікового відношення сигнал/шум (PSNR) на 4.7 %.

Доведено практичну придатність розроблених рішень. Порівняльний аналіз обчислювальної складності показав, що суттєве візуальне та кількісне покращення якості відновлення дрібних анатомічних структур було досягнуто за рахунок незначного збільшення часу обробки. Це підтверджує перспективність впровадження моделі SeConvNet + CBAM у реальні медичні інформаційні системи.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

- [1] S. Umme, «Image Quality Assessment Through FSIM, SSIM, MSE and PSNR-A Comparative Study. DOI: <https://doi.org/10.4236/jcc.2019.73002>,» *Journal of Computer and Communications*, т. 7, № 3, 2019.
- [2] Y. Al Najjar, «Comparative Analysis of Image Quality Assessment Metrics: MSE, PSNR, SSIM and FSIM,» *International Journal of Science and Research (IJSR)*. DOI: [10.21275/SR24302013533](https://doi.org/10.21275/SR24302013533), т. 13, № 3, 2024.
- [3] L. Fan, F. Zhang, H. Fan та C. Zhang, «Brief Review of Image Denoising Technique,» *Visual Computing for Industry, Biomedicine, and Art*. DOI: [10.1186/s42492-019-0016-7](https://doi.org/10.1186/s42492-019-0016-7), т. 2, № 1, 2019.
- [4] R. S. Jebur, M. H. B. A. W. Zabil, D. A. Hammood та L. K. Cheng, «A Comprehensive Review of Image Denoising in Deep Learning,» *Multimedia Tools and Applications*, 83(20). DOI: [10.1007/s11042-023-17468-2](https://doi.org/10.1007/s11042-023-17468-2), 2024.
- [5] R. Yang, L. Yin, M. Gabbouj, J. Astola та Y. Neuvo, «Optimal Weighted Median Filtering Under Structural Constraints,» *IEEE Transactions on Signal Processing*. DOI: [10.1109/78.370615](https://doi.org/10.1109/78.370615), т. 43, № 3, 1995.
- [6] M. Zhu, Y. Hu, J. Yu, B. He та J. Liu, «Find Outliers of Image Edge Consistency by Weighted Local Linear Regression with Equality Constraints,» *Sensors*, 21(7). DOI: [10.3390/s21072563](https://doi.org/10.3390/s21072563), 2021.
- [7] Y. Wang та J. J. Healy, «Image Quality Assessment for Gibbs Ringing Reduction,» *Algorithms*, 16. DOI: [10.3390/a16020096](https://doi.org/10.3390/a16020096), 2023.
- [8] K. M. Jeevan та K. Nallaperumal, «An Algorithm for Wavelet Thresholding Based Image Denoising by Representing Images in Hexagonal Lattice,» 2018.
[Онлайновий]. Available:
https://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1665-64232018000200103. [Дата звернення: 22 February 2026].

- [9] P. Maharjan, N. Xu, X. Xu, Y. Song та Z. Li, «DCTResNet: Transform Domain Image Deblocking for Motion Blur Images,» *Proceedings of the International Conference on Visual Communications and Image Processing (VCIP)*. DOI: 10.1109/VCIP53242.2021.9675411, 2021.
- [10] H. H. Karaoglu та E. M. Eksioglu, «Parameter-Efficient Harmonic Networks for JPEG Compression Artifact Removal,» *Proceedings of the International Conference on Machine Learning and Applications (ICMLA)*. DOI: 10.1109/ICMLA58977.2023.00310, 2023.
- [11] L. Rudin, S. Osher та E. Fatemi, «Nonlinear total variation based noise removal algorithms,» *Physica D: Nonlinear Phenomena*. DOI: 10.1016/0167-2789(92)90242-F, т. 60, № 1-4, 1992.
- [12] M. Pragliola, L. Calatroni, A. Lanza та F. Sgallari, «On and Beyond Total Variation Regularization in Imaging: The Role of Space Variance,» *SIAM Review*. DOI: 10.1137/21M1410683, т. 65, 2023.
- [13] K. Bredies, K. Kunisch та T. Pock, «Second order total generalized variation (TGV) for MRI,» *SIAM Journal on Imaging Sciences*. DOI: 10.1002/mrm.22595, т. 3, № 3, 2010.
- [14] M. Huemer, C. Stilianu, O. Maier, K. Bredies та R. Stollberger, «Improved quantification in CEST-MRI by joint spatial total generalized variation,» *Magnetic Resonance in Medicine*. DOI: 10.1002/mrm.30129, т. 92, 2024.
- [15] S. Geman та D. Geman, «Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images,» *IEEE Transactions on Pattern Analysis and Machine Intelligence*. DOI: 10.1109/TPAMI.1984.4767596, Томи %1 з %2PAMI-6, № 6, 1984.
- [16] H. Chung, J. Kim та J. C. Ye, «Diffusion Posterior Sampling for General Inversion Problems,» *IEEE Transactions on Pattern Analysis and Machine Intelligence*. DOI: 10.1109/TPAMI.2024.3364914, т. 46, 2024.

- [17] J. Portilla, «Image denoising using mixtures of Gaussian scale mixtures,» в *IEEE International Conference on Image Processing*. DOI: 10.1109/ICIP.2008.4711817, 2008.
- [18] T. Huang, X. Yuan, W. Dong, J. Wu та G. Shi, «Deep Gaussian Scale Mixture Prior for Image Reconstruction,» *IEEE Transactions on Pattern Analysis and Machine Intelligence*. DOI: 10.1109/TPAMI.2023.3265103, т. 45, 2023.
- [19] H. Zhang, Y. Liu та G. Cai, «A novel 3D bilateral filtering algorithm with noise level estimation assisted by multi-temporal SAR,» *PLOS ONE*. DOI: 10.1371/journal.pone.0315395, т. 20, 2025.
- [20] Y. Liu, H. Li, N. Cao та F. Shen, «High-Noise Grayscale Image Denoising Using an Improved Median Filter for the Adaptive Selection of a Threshold,» *Applied Sciences*. DOI: 10.3390/app14020635, т. 14, 2024.
- [21] X. Xu, C. Xu та L. Zhao, «Real-Time Volume-Rendering Image Denoising Based on Spatiotemporal Weighted Kernel Prediction,» *Journal of Imaging*. DOI: 10.3390/jimaging11040126, т. 11, 2025.
- [22] P. Perona та J. Malik, «Scale-space and edge detection using anisotropic diffusion,» *IEEE Transactions on Pattern Analysis and Machine Intelligence*. DOI: 10.1109/34.56205, т. 12, № 7, 1990.
- [23] D. Liu, B. Wen, Y. Fan, S. Chang та T. Huang, «Non-Local Recurrent Network for Image Restoration,» в *Advances in Neural Information Processing Systems 31 (NeurIPS 2018)*. URL: <https://arxiv.org/abs/1806.02919>, 2018.
- [24] S. W. Zamir, A. Arora, S. Khan, M. Hayat, F. S. Khan та M.-H. Yang, «Restormer: Efficient Transformer for High-Resolution Image Restoration,» *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. DOI: 10.1109/CVPR52688.2022.00564, 2022.
- [25] Y. Hou та D. Shen, «Image Denoising With Morphology- and Size-adaptive Block-matching Transform Domain Filtering,» т. 2018, № 1, 2018.

- [26] H. Jia, Q. Yin та M. Lu, «Blind-noise image denoising with block-matching domain transformation filtering and improved guided filtering,» *Scientific Reports*. DOI: 10.1038/s41598-022-20578-w, т. 12, № 1, 2022.
- [27] B. Kawar, M. Elad, S. Ermon та J. Song, «Denoising Diffusion Restoration Models,» *Advances in Neural Information Processing Systems (NeurIPS)*, т. 35, 2022.
- [28] O. Özdenizci та R. Legenstein, «Restoring Vision in Adverse Weather Conditions With Patch-Based Denoising Diffusion Models,» *IEEE Transactions on Pattern Analysis and Machine Intelligence*. DOI: 10.1109/TPAMI.2023.3292419, т. 45, 2023.
- [29] Y. Zhou, J. Jiao, H. Huang, Y. Wang, J. Wang, H. Shi та T. Huang, «When AWGN-based Denoiser Meets Real Noises,» в *Proceedings of the AAAI Conference on Artificial Intelligence*. DOI: 10.1609/aaai.v34i07.7009, 2020.
- [30] T. Huang, S. Li, X. Jia, H. Lu та J. Liu, «Neighbor2Neighbor: Self-Supervised Denoising from Single Noisy Images,» *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. DOI: 10.1109/CVPR46437.2021.01454, 2021.
- [31] K. Gurusamy та A. Sugumar, «Speckle Noise Suppression in Ultrasound Images Using Modular Neural Networks,» *Intelligent Automation & Soft Computing*. DOI: 10.32604/iasc.2023.022631, т. 35, № 2, 2023.
- [32] S. Izadi, D. Maleki та F. Sheikhi, «Image Denoising in the Deep Learning Era,» *Artificial Intelligence Review*. DOI: 10.1007/s10462-022-10305-2, т. 56, № 7, 2023.
- [33] C. R. Noordman, D. Yakar, J. Bosma, F. F. J. Simonis та H. Huisman, «Complexities of deep learning-based undersampled MR image reconstruction,» *European Radiology Experimental*. DOI: 10.1186/s41747-023-00372-7, т. 7, 2023.

- [34] M. Elad, B. Kawak та P. Milanfar, «Image Denoising: The Deep Learning Revolution and Beyond—A Survey Paper,» *SIAM Review*. DOI: 10.1137/23M154863X, т. 65, № 3, 2023.
- [35] V. Novoselac та A. Katić, «Robust regression analysis using the weighted median model for improved denoising of MR data in image processing,» *Hacettepe Journal of Mathematics and Statistics*. DOI: 10.15672/hujms.1461429, т. 54, 2025.
- [36] K. Zhang, Y. Li, W. Zuo, L. Zhang, L. Van Gool та R. Timofte, «Plug-and-Play Image Restoration With Deep Denoiser Prior,» *IEEE Transactions on Pattern Analysis and Machine Intelligence*. DOI: 10.1109/TPAMI.2021.3088914, т. 44, 2022.
- [37] О. Зінченко, О. Звенігородський та Т. Кисіль, «Згорткові нейронні мережі для вирішення задач комп'ютерного зору,» *Телекомунікаційні та інформаційні технології*. DOI: 10.31673/2412-4338.2022.020411, № 2, 2022.
- [38] K. Zhang, W. Zuo та L. Zhang, «Learning a Single Convolutional Super-Resolution Network for Multiple Degradations,» в *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. DOI: 10.1109/CVPR.2018.00344, 2018.
- [39] А. Линовський та В. Мухін, «Засоби покращення якості та знешумлення зображень на основі застосування згорткових та рекурентних нейронних мереж,» *Телекомунікаційні та інформаційні технології*. DOI: 10.31673/2412-4338.2023.018289, № 1, 2023.
- [40] B. Chapaliuk та Y. Zaychenko, «Recurrent Neural Network Usage for Computer-aided Lung Cancer Detection System,» *System Research and Information Technologies*. DOI: 10.20535/SRIT.2308-8893.2019.3.03, № 3, 2019.
- [41] J. Gurrola-Ramos, T. Alarcon, O. Dalmau та J. V. Manjón, «MRI Rician Noise Reduction Using Recurrent Convolutional Neural Networks,» *IEEE Access*. DOI: 10.1109/ACCESS.2024.3446791, т. 12, 2024.

- [42] L. Zhu, Y. Zheng, M. Geng, L. Wang ta H. Huang, «Recurrent Spike-based Image Restoration under General Illumination,» *Proceedings of the 31st ACM International Conference on Multimedia*. DOI: 10.48550/arXiv.2308.03018, 2023.
- [43] H. Zhang, D. Tan ta W. Lu, «NAMF: A Non-local Adaptive Mean Filter for Salt-and-Pepper Noise Removal,» *Electronics Letters*. DOI: 10.1049/el.2019.3493, т. 56, № 8, 2020.
- [44] Y. Zhang, L. Wang ta W. Wu, «Euclidean-distance-based switching adaptive mean filter for salt-and-pepper noise removal,» *Proceedings of SPIE*. DOI: 10.1117/12.3029757, 2024.
- [45] M. Taassori ta B. Vizvári, «Enhancing Medical Image Denoising: A Hybrid Approach Incorporating Adaptive Kalman Filter and Non-Local Means with Latin Square Optimization,» *Electronics*. DOI: 10.3390/electronics13132640, т. 13, 2024.
- [46] L. Liang, S. Deng, L. Gueguen, M. Wei, X. Wu ta J. Qin, «Convolutional Neural Network With Median Layers for Denoising Salt-and-Pepper Contaminations,» *IEEE Transactions on Image Processing*. DOI: 10.1109/TIP.2019.2910657, т. 28, № 10, 2019.
- [47] Y. Cui, W. Ren, X. Cao ta A. Knoll, «Revitalizing Convolutional Network for Image Restoration,» *IEEE Transactions on Pattern Analysis and Machine Intelligence*. DOI: 10.1109/TPAMI.2024.3419007, т. 46, 2024.
- [48] S. Holla K, N. Park ta B. Lee, «EFID: Edge-Focused Image Denoising Using a Convolutional Neural Network,» *IEEE Access*. DOI: 10.1109/ACCESS.2023.3239835, т. 11, 2023.
- [49] Y. Zhang, K. Li, K. Li, L. Wang, B. Zhong ta Y. Fu, «Residual Non-local Attention Networks for Image Restoration,» в *IEEE/CVF International Conference on Computer Vision (ICCV)*. DOI: 10.1109/ICCV.2019.00601, 2019.

- [50] A. Zafar, D. Aftab, R. Qureshi, X. Fan та P. Chen, «Single Stage Adaptive Multi-Attention Network for Image Restoration,» *IEEE Transactions on Image Processing*. DOI: 10.1109/TIP.2024.3384838, т. 33, 2024.
- [51] Y. Li, Y. Fan, X. Xiang, D. Demandolx та R. Ranjan, «Efficient and Explicit Modelling of Image Hierarchies for Image Restoration,» *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, т. 2023.
- [52] C. Silvano, G. Agosta, S. Cherubin, D. Gadioli, E. Vitali та B. Peccerillo, «A Survey on Deep Learning Hardware Accelerators for Heterogeneous HPC Platforms,» *ACM Computing Surveys*. DOI: 10.1145/3729215, т. 57, № 11, 2025.
- [53] S. Woo, J. Park, J.-Y. Lee та I. S. Kweon, «CBAM: Convolutional Block Attention Module,» в *Proceedings of the European Conference on Computer Vision (ECCV)*. DOI: 10.1007/978-3-030-01234-2_1, 2018.
- [54] N. Zrira, A. Jimi, M. Di Nardo, I. Elafi, M. Gallab та R. C. El Ouazzani, «GCBAM-UNet: Sun Glare Segmentation Using Convolutional Block Attention Module,» *Applied System Innovation*. DOI: 10.3390/asi7060128, т. 7, 2024.
- [55] R. Ye, Y. Qian та X. Huang, «RT-CBAM: Refined Transformer Combined with Convolutional Block Attention Module for Underwater Image Restoration,» *Sensors*. DOI: 10.3390/s24185893, т. 24, 2024.
- [56] C. Tian, Y. Xu, L. Fei та J. Wang, «Deep Learning on Image Denoising: An Overview,» *Neural Networks*. DOI: 10.1016/j.neunet.2020.07.025, т. 131, 2020.
- [57] S. Anwar та N. Barnes, «Real Image Denoising With Feature Attention,» в *IEEE/CVF International Conference on Computer Vision (ICCV)*. DOI: 10.1109/ICCV.2019.00325, 2019.
- [58] K. Zhang, W. Zuo, Y. Chen, D. Meng та L. Zhang, «Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising,» *IEEE Transactions on Image Processing*. DOI: 10.1109/TIP.2017.2662206, т. 26, № 7, 2017.

- [59] K. Zhang, W. Zuo ta L. Zhang, «FFDNet: Toward a Fast and Flexible Solution for CNN-Based Image Denoising,» *IEEE Transactions on Image Processing*. DOI: 10.1109/TIP.2018.2839891, т. 27, № 9, 2018.
- [60] P. Qiao, Y. Huang, S. Chen ta L. Wang, «Learning Non-local Image Diffusion for Image Denoising,» в *IEEE International Conference on Computer Vision (ICCV)*. DOI: 10.1109/ICCV.2017.202, 2017.
- [61] A. Khan, N. U. Khan, M. Ali, J. Shah ta S. Khan, «Adversarial Gaussian Denoiser for Multiple-Level Image Denoising,» *Sensors*. DOI: 10.3390/s21092998, т. 21, № 9, 2021.
- [62] B. Jiang, J. Liu, J. Wu, G. Shi ta L. Li, «Efficient Image Denoising Using Deep Learning: A Brief Survey,» *Information Fusion*. DOI: 10.1016/j.inffus.2025.103013, т. 114, 2025.
- [63] Z. Fayyaz, M. S. Hosseini ta K. N. Plataniotis, «Deep Unfolding for Iterative Stripe Noise Removal,» *IEEE Transactions on Image Processing*. DOI: 10.1109/TIP.2023.3235308, т. 32, 2023.
- [64] K.-C. Chang, R. Wang, H.-J. Lin, Y.-L. Liu, C.-P. Chen, Y.-L. Chang ta H.-T. Chen, «Learning Camera-Aware Noise Models,» в *Proceedings of the European Conference on Computer Vision (ECCV)*. DOI: 10.1007/978-3-030-58536-5_32, 2020.
- [65] Y. Cheng, D. Wang, P. Zhou ta T. Zhang, «A Survey of Model Compression and Acceleration for Deep Neural Networks,» *IEEE Signal Processing Magazine*. DOI: 10.1109/MSP.2018.2839547, т. 35, № 1, 2018.
- [66] D. Gyawali, «Comparative Analysis of CPU and GPU Profiling for Deep Learning Models,» DOI: 10.48550/arXiv.2309.02521, 2023.
- [67] R. Aggarwal, V. Sounderajah, G. Martin, D. S. W. Ting, A. Carty, H. Ashrafian ta A. Darzi, «Diagnostic Accuracy of Deep Learning in Medical Imaging: A Systematic Review and Meta-analysis,» *NPJ Digital Medicine*. DOI: 10.1038/s41746-021-00438-z, т. 4, № 1, 2021.

- [68] A. A. Rafiee та M. Farhang, «A Deep Convolutional Neural Network for Salt-and-pepper Noise Removal Using Selective Convolutional Blocks,» *Applied Soft Computing*. DOI: 10.1016/j.asoc.2023.110535, т. 145, 2023.
- [69] J. Hu, L. Shen та G. Sun, «Squeeze-and-Excitation Networks,» *IEEE Transactions on Pattern Analysis and Machine Intelligence*. DOI: 10.1109/TPAMI.2018.2846170, т. 42, № 8, 2020.
- [70] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh та J. Liang, «UNet++: A Nested U-Net Architecture for Medical Image Segmentation,» в *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support (DLMIA 2018)*. DOI: 10.1007/978-3-030-00889-5_1, 2018.
- [71] H. Wang, P. Cao, J. Wang та O. R. Zaiane, «UCTransNet: Rethinking the Skip Connections in U-Net from a Channel-Wise Perspective with Transformer,» *Proceedings of the AAAI Conference on Artificial Intelligence*. DOI: <https://doi.org/10.1609/aaai.v36i3.20144>, 2022.
- [72] H. Zhao, Y. Gou, B. Li, D. Peng, J. Lv та X. Peng, «Comprehensive and Delicate: An Efficient Transformer for Image Restoration,» *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. DOI: 10.1109/CVPR52729.2023.01357, 2023.
- [73] Y. Wu та W. Li, «Compacter: A Lightweight Transformer for Image Restoration,» *Proceedings of the 32nd ACM International Conference on Multimedia*. DOI: 10.1145/3664647.3680811, 2024.
- [74] Z. Chen, Y. Zhang, J. Gu, Y. Zhang, L. Kong та X. Yuan, «Cross Aggregation Transformer for Image Restoration,» *Advances in Neural Information Processing Systems (NeurIPS)*. DOI:10.1145/3664647.3680811, т. 35, 2022.
- [75] Y. Wang, «Artificial-Intelligence Integrated Circuits: Comparison of GPU, FPGA and ASIC,» *Applied and Computational Engineering*. DOI: 10.54254/2755-2721/4/20230358, т. 4, 2023.

- [76] J. M. Rodriguez-Corral, J. M. Rodriguez-Leyva, F. Ternero-Muñiz, M. A. Luque-Nieto ta J. Poncela, «Energy Efficiency in Edge TPU Vs. Embedded GPU for Computer-aided Medical Imaging Segmentation and Classification,» *Engineering Applications of Artificial Intelligence*. DOI: 10.1016/j.engappai.2023.107298, т. 127, 2024.
- [77] Y. Makin ta R. Maliakkal, «Sustainable AI Training via Hardware-Software Co-Design on NVIDIA, AMD, and Emerging GPU Architectures,» DOI: 10.48550/arXiv.2508.13163, 2025.
- [78] K. Ohashi, Y. Nagatani, M. Yoshigoe, T. Nakaura, T. Okino, Y. Fushimi, D. Ueda, H. Isoda ta K. Togashi, «Applicability Evaluation of Full-Reference Image Quality Assessment Methods for Computed Tomography Images,» *Journal of Digital Imaging*. DOI: 10.1007/s10278-023-00875-0, т. 36, № 5, 2023.
- [79] H. Herath, N. Madusanka ta B.-I. Lee, «A Systematic Review of Medical Image Quality Assessment,» *Journal of Imaging*. DOI: 10.3390/jimaging11040100, т. 11, № 4, 2025.
- [80] A. Mason, J. Rioux, S. E. Clarke, B. Costa, M. Schmidt, V. Keough, T. Huynh ta S. Beyea, «Comparison of Objective Image Quality Metrics to Expert Radiologists' Scoring of Diagnostic Quality of MR Images,» *IEEE Transactions on Medical Imaging*. DOI: 10.1109/TMI.2019.2939987, т. 39, № 4, 2020.

ДОДАТКИ

ДОДАТОК А

Список публікацій за темою дисертації та відомості про апробацію результатів дослідження

Статті у наукових фахових виданнях України категорії Б:

1. Линовський А. О., Мухін В. Є. Засоби покращення якості та знешумлення зображень на основі застосування згорткових та рекурентних нейронних мереж. Телекомунікаційні та інформаційні технології, № 1 (78), 2023. - с.82 -89. (Здобувачем Линовським А.О. було виконано комплексний огляд сучасних підходів до покращення якості та знешумлення зображень на основі згорткових і рекурентних нейронних мереж з урахуванням non-local operations блоку; співавтором Мухіним В.Є. здійснено загальну постановку задачі дослідження та визначено актуальність застосування non-local operations у сучасних архітектурах)
2. Линовський А. О., Мухін В. Є. Модифікована модель фільтрації шуму на зображеннях на основі згорткової нейронної мережі. Електронне фахове наукове видання «Кібербезпека: освіта, наука, техніка», 4(24), 2024. -с. 388-397. (Здобувачем Линовським А.О. було виконано комплексний аналіз існуючих підходів до видалення шуму типу salt-and-pepper (SaP) і запропоновано модифікацію архітектури SeConvNet для підвищення ефективності фільтрації цього типу шуму; співавтором Мухіним В.Є. обґрунтовано доцільність модифікації існуючих згорткових архітектур для задачі фільтрації імпульсного шуму та проаналізовано отримані результати)
3. Линовський А. О., Мухін В. Є. Модифікована модель фільтрації шуму на зображеннях на основі згорткової нейронної мережі з додаванням залишкових зв'язків і модулю уваги. Проблеми інформатизації та управління, 4 (80), 2024. с. 76-83 (Здобувачем Линовським А.О. було запропоновано модифікацію моделі SeConvNet шляхом додавання блока, що поєднує резидуальні з'єднання з каналною та просторовою увагою (CBAM), що дозволяє вибірково підсилювати

структурно важливі фрагменти зображення; співавтором Мухіним В.Є. надано концептуальне обґрунтування застосування механізмів уваги та надано консультації щодо архітектурної інтеграції залишкових зв'язків)

Наукові праці, які засвідчують апробацію матеріалів дисертації

- Линовський А. О., Мухін В. Є. An Improved CNN Architecture for Cluster Salt-and-pepper Noise Filtering in Medical Data. Міжнародна науково-практична конференція «Інформаційні технології та безпека» (ІТБ-2025) : матеріали конф. Київ, 2025. С. 148–152. (Здобувачем Линовським А.О. було запропоновано застосування розробленої модифікації згорткової нейронної мережі для специфічної задачі видалення скупчень (кластерів) шуму типу salt-and-pepper на медичних знімках та проведено оцінку ефективності запропонованого підходу; співавтором Мухіним В.Є. визначено специфіку прикладної задачі в контексті обробки медичних даних та затверджено критерії оцінки ефективності розробленого методу)