МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
УМАНСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
Факультет економіки, підприємництва та інформаційних технологій
Кафедра інформаційних технологій

Допущено до захисту
"_____" _____ 2026 р.
Зав. кафедри, к.т.н., доцент
_____ **Ліщук Р.І.**

КВАЛІФІКАЦІЙНА РОБОТА БАКАЛАВРА

РОЗРОБКА ПРОГРАМНОГО МОДУЛЯ ДЛЯ ВІДНОВЛЕННЯ КОЛЬОРУ НА ФОТОГРАФІЯХ З ВИКОРИСТАННЯМ МЕТОДІВ ГЛИБОКОГО НАВЧАННЯ

Виконав – студент 4 курсу, групи 41-кн
освітньої програми "Комп'ютерні науки"
спеціальності 122 Комп'ютерні науки
Якименко Євгеній Олександрович

Науковий керівник – к.т.н., доцент
Ліщук Роман Ігорович

Умань – 2026

Зміст

ВСТУП 3

РОЗДІЛ 1 ОБҐРУНТУВАННЯ ДОЦІЛЬНОСТІ РОЗРОБКИ ПРОГРАМНОГО МОДУЛЯ ДЛЯ ВІДНОВЛЕННЯ КОЛЬОРУ НА ЦИФРОВИХ ФОТОГРАФІЯХ 5

1.1. Аналіз сучасних технологій відновлення кольору на цифрових фотографіях 5

1.2. Аналіз програм аналогів розфарбовування цифрових фотографій .. 6

1.3. Постановка задачі 11

РОЗДІЛ 2 ПРОГРАМНОГО МОДУЛЯ РОЗФАРБОВУВАННЯ ЦИФРОВИХ ФОТОГРАФІЙ 14

2.1. Обґрунтування доцільності застосування нейромережевого методу обробки візуальних даних у задачі відновлення кольору монохромних знімків 14

2.2. Процес навчання нейронної мережі 24

2.3. Обґрунтування вибору колірної моделі 27

2.4. Розробка архітектури штучної нейронної мережі та алгоритмічної схеми для відновлення кольору цифрових фотографій 31

2.5. Обґрунтування вибору інструментальних засобів та середовища розробки 35

РОЗДІЛ 3 РЕАЛІЗАЦІЯ МОДУЛЯ РОЗФАРБОВУВАННЯ ЦИФРОВИХ ФОТОГРАФІЙ 40

3.1. Реалізація алгоритмів програмного модуля розфарбовування цифрових фотографій 40

3.2. Побудова UML-діаграми класів програмного модуля відновлення кольору цифрових фотографій 45

3.3. Тестування програмного модуля відновлення кольору цифрових фотографій та аналіз результатів його функціонування 48

ВИСНОВКИ 53

СПИСОК ЛІТЕРАТУРИ 55

ДОДАТКИ 59

ВСТУП

Процес відновлення кольору монохромних зображень є складною науково-практичною задачею. На сучасному етапі розвитку інформаційних технологій кількість програмних рішень, здатних повністю автоматизувати зазначену процедуру, залишається обмеженою. Традиційно відновлення кольору виконується мануальним методом, що передбачає цілеспрямовану модифікацію функції яскравості для кожної окремої локальної області зображення з метою досягнення необхідного хроматичного значення. Даний підхід характеризується високою трудомісткістю та значними часовими витратами. З огляду на це, задоволення високого ринкового попиту на послуги з відновлення кольору візуальних даних супроводжується значною економічною вартістю робіт. Показовим прикладом є ініційований у дві тисячі дев'ятому році процес відновлення кольору класичних монохромних кінострічок, де бюджет обробки першого відеоматеріалу сягнув близько мільйона доларів США та потребував багатомісячної праці колективу з понад шестисот фахівців.

В епоху до масового впровадження цифрових технологій відновлення кольору монохромних світлин здійснювалося виключно ручним способом. Еволюція обчислювальної техніки частково оптимізувала даний процес, однак він продовжує залежати від висококваліфікованої праці спеціалістів з комп'ютерної графіки. Основна проблематика частково автоматизованого підходу полягає у необхідності ручної сегментації кожного візуального кадру, що зумовлює високу собівартість кінцевого результату. Протягом подальшого розвитку галузі відбулася інтенсифікація досліджень у напрямку розробки методів автоматизованого відновлення кольору. Ключовою метою зазначених підходів стало не лише алгоритмічне виявлення границь монохромних сегментів, але й точне обчислення їхніх хроматичних характеристик. Перевагою такої парадигми є суттєве скорочення часових витрат на обробку візуальних даних, проте перші покоління відповідних алгоритмів демонстрували недостатній рівень якості результуючих зображень.

Відчутний прогрес у вирішенні задачі автоматизованого відновлення кольору було зафіксовано лише впродовж останніх років. Фундаментальною передумовою цього прориву стала стрімка популяризація та вдосконалення апарату штучних нейронних мереж, зокрема архітектур згорткового типу. Наведені фактори комплексно обґрунтовують актуальність розробки програмного модуля для відновлення кольору цифрових фотографій із застосуванням нейромережевих технологій.

Метою дослідження є підвищення якості відновлення кольору цифрових фотографій.

Об’єкт дослідження – процес відновлення кольору цифрових фотографій.

Предмет дослідження – програмні засоби відновлення кольору цифрових фотографій.

Реалізація визначеної мети дослідження зумовлює необхідність виконання комплексу таких науково-практичних завдань: Провести критичний огляд та систематизацію сучасних методів і технологічних підходів до відновлення кольору цифрових фотографій; Здійснити порівняльну характеристику існуючих програмних рішень та систем-аналогів у сфері відновлення кольору візуальних даних; Спроектувати загальну архітектуру та структурну організацію спеціалізованого програмного модуля для відновлення кольору цифрових фотографій; Синтезувати алгоритмічне забезпечення процесу автоматизованого відновлення кольору цифрових фотографій; Виконати практичну імплементацію розробленого алгоритмічного апарату у вигляді діючого програмного модуля відновлення кольору цифрових фотографій; Організувати експериментальну апробацію створеного програмного забезпечення з подальшим аналітичним оцінюванням результатів його функціонування.

РОЗДІЛ 1 ОБҐРУНТУВАННЯ ДОЦІЛЬНОСТІ РОЗРОБКИ ПРОГРАМНОГО МОДУЛЯ ДЛЯ ВІДНОВЛЕННЯ КОЛЬОРУ НА ЦИФРОВИХ ФОТОГРАФІЯХ

1.1. Аналіз сучасних технологій відновлення кольору на цифрових фотографіях

Дана робота присвячена вирішенню науково-практичної задачі відновлення кольору монохромних зображень. Актуальність обраної проблематики зумовлена наявністю значних обсягів ретроспективних фотографічних матеріалів у приватних колекціях, які потребують візуального збагачення шляхом застосування методів відновлення кольору. Інтеграція хроматичної інформації у структуру історичного фотодокумента сприяє більш глибокому та деталізованому семантичному аналізу зафіксованих сцен. Процедура відновлення кольору також виступає важливим інструментарієм для фахівців у галузі історичних досліджень під час опрацювання різноманітних архівних візуальних даних, що підтверджує безперечну актуальність даної проблематики на сучасному етапі.

До періоду масового розповсюдження обчислювальної техніки процедура відновлення кольору здійснювалася виключно мануальним способом. Подальша еволюція інформаційних технологій дозволила перевести цей процес у цифрове середовище з використанням базових інструментів графічної обробки. Незважаючи на це, рівень трудомісткості зазначеної процедури без залучення спеціалізованих засобів автоматизації залишається надзвичайно високим. Відповідно, додатковим мотивом даного дослідження є підвищення якісних показників процесу відновлення кольору у порівнянні з існуючими програмними рішеннями. Сучасні алгоритми автоматизованого відновлення кольору демонструють зростання швидкодії, однак результуюча якість оброблених зображень часто не відповідає необхідним критеріям.

Технологічний цикл автоматизованого відновлення кольору структурно поділяється на кілька послідовних стадій:

1. Сегментація, що являє собою процес декомпозиції цифрового зображення на окремі локальні області. Ключовою метою сегментації є оптимізація та трансформація візуального подання даних для спрощення їх подальшого аналітичного опрацювання [5]. Зазначена процедура традиційно застосовується для просторової локалізації об'єктів та ідентифікації контурів на площині зображення. У більш точному визначенні, сегментація зображення полягає у класифікації кожного елемента растра з присвоєнням відповідного ідентифікатора, що забезпечує групування пікселів зі спільними візуальними характеристиками.
2. Обчислення хроматичних параметрів кожного виділеного сегмента, базуючись на аналізі значень монохромного каналу та структурних особливостей самої області.
3. Безпосереднє відновлення кольору ідентифікованих сегментів згідно з розрахованими спектральними показниками.

На основі аналітичного огляду суміжних наукових розвідок було встановлено, що фундаментальним фактором, який негативно впливає на підсумкову якість відновлення кольору, є недостатня точність алгоритмічного визначення хроматичних характеристик окремих сегментів. З метою мінімізації впливу цієї проблеми та підвищення загальної точності системи пропонується імплементувати механізм інтерактивної взаємодії, який передбачає використання допоміжних вхідних даних від користувача у форматі локальних колірних маркерів.

1.2. Аналіз програм аналогів розфарбовування цифрових фотографій

Процес відновлення кольору монохромних фотографічних зображень становить собою нетривіальну науково-технічну задачу. На сучасному етапі розроблено обмежену кількість програмних рішень, здатних автоматизувати зазначені процедури. Переважно відновлення кольору здійснюється мануальним методом, за якого для кожної локальної області зображення модифікується функція яскравості відповідно до цільового хроматичного

значення. Дана процедура характеризується високою трудомісткістю та вимагає значних зусиль. З огляду на це, в сучасних ринкових умовах послуги з відновлення кольору монохромних зображень відзначаються значною вартістю при збереженні високого рівня попиту [6].

Серед існуючого програмного інструментарію для відновлення кольору монохромних зображень доцільно виділити такі продукти та сервіси: Adobe Photoshop, веб-ресурс ColorizePhoto, веб-сервіс Colorize-photos Algorithmia та Akvis Coloriage. Слід зауважити, що перші два рішення функціонують без застосування методів штучного інтелекту, тоді як два останні базуються на інтелектуальних алгоритмах.

Кожен із зазначених технічних засобів володіє специфічними перевагами та недоліками, проте базовий алгоритм відновлення кольору є універсальним і полягає в локалізації площі монохромного сегмента з її подальшим відокремленим опрацюванням. У межах даного дослідження передбачається автоматизація саме цієї діяльності.

Програмний продукт Adobe Photoshop являє собою графічний редактор, розроблений та підтримуваний корпорацією Adobe Systems. Зазначене рішення займає провідні позиції на ринку комерційних засобів редагування растрових зображень і є найбільш відомим продуктом компанії-розробника [7]. Здебільшого дане середовище застосовується для обробки цифрових фотографій та генерування растрової графіки, що проілюстровано на рис. 1.1. Оцінюючи це програмне забезпечення як інструмент для відновлення кольору, необхідно наголосити, що його ефективне використання вимагає наявності значного практичного досвіду обробки візуальних даних та високого рівня художньої кваліфікації фахівця. За умови поєднання згаданих факторів підсумковий результат відновлення кольору відзначається винятково високою якістю. Проте вказана перевага частково нівелюється значними часовими витратами на виконання операцій у порівнянні з іншими спеціалізованими програмами для відновлення кольору.



Рисунок 1.1 – Інтерфейс програмного середовища Adobe Photoshop

Сервіс ColorizePhoto являє собою спеціалізований веб-додаток, призначений для відновлення кольору без залучення технологій штучного інтелекту [8]. Базовий механізм обробки в даному середовищі є аналогічним до методів, що застосовуються у програмному продукті Adobe Photoshop. Проте функціональні можливості цього сервісу є вузькоспеціалізованими та спрямовані виключно на вирішення задачі відновлення кольору. Використання зазначеного інструменту дозволяє досягти високих якісних показників обробки візуальних даних, однак, як і у випадку з попередньо розглянутим програмним забезпеченням, ця процедура залишається вкрай трудомісткою та супроводжується значними часовими витратами, що проілюстровано на рисунку 1.2.

Альтернативним програмним рішенням виступає веб-сервіс Algorithmia Colorize Photos. Зазначена платформа була розроблена групою дослідників у галузі нейромережевих технологій з метою демонстрації прикладного потенціалу систем штучного інтелекту. У межах даного проекту алгоритми машинного навчання застосовуються для відновлення кольору монохромних фотографій. Користувачам надається можливість ініціювати процес хроматичної реконструкції шляхом введення гіперпосилання на цільове зображення, після чого обробка виконується в автоматизованому режимі [9].



Рисунок 1.2 – Інтерфейс програмного додатка ColorizePhoto

Попри те, що генерація результату триває лічені секунди, обчислювальне навантаження на систему штучного інтелекту залишається надзвичайно високим. Це зумовлено необхідністю комплексного аналізу візуального контенту, ідентифікації розпізнаваних об'єктів, екстракції відповідних спектральних характеристик із бази знань моделі та їх подальшої інтеграції у структуру вихідного зображення. Ефективність процесу відновлення кольору перебуває у прямій залежності від семантичних особливостей вхідних візуальних даних. Залежно від контексту, результуючі зображення можуть демонструвати як високий рівень хроматичної достовірності та варіативності, так і суттєві артефакти або некоректне відтворення спектра. Емпіричним шляхом встановлено, що досліджуваний сервіс забезпечує найвищу точність під час обробки таких класів об'єктів, як елементи людського тіла, гідрологічні об'єкти, рослинність, транспортні засоби та архітектурні споруди.

Підсумовуючи аналіз функціонування даного сервісу, варто зазначити, що, на противагу раніше розглянутим програмним продуктам, він забезпечує високу швидкодію процесу відновлення кольору, проте якість отриманого результату часто є недостатньою, що наочно продемонстровано на рисунку 1.3.



Рисунок 1.3 – Інтерфейс веб-сервісу Algorithmia Colorize Photos

Наступним об'єктом аналізу є програмний продукт AKVIS Coloriage. Зазначене рішення становить собою спеціалізований інструментарій, призначений для відновлення кольору монохромних фотографічних матеріалів та модифікації хроматичних параметрів наявних кольорових зображень [10].

Досліджуване програмне забезпечення демонструє стабільну ефективність під час обробки різноманітних візуальних жанрів, серед яких портретна та пейзажна зйомка, а також натюрморти. Функціональні можливості системи дозволяють здійснювати відновлення кольору ретроспективних оцифрованих монохромних фотографій, локальну заміну спектральних характеристик у цифрових файлах, а також відновлення кольору монохромних графічних ескізів та анімаційних матеріалів. Додатково, використання даного продукту є доцільним для вирішення прикладних задач із моделювання колірної гами об'єктів дизайну та експериментальної зміни зовнішнього вигляду об'єктів, що наочно проілюстровано на рисунку 1.4.

Архітектурно AKVIS Coloriage реалізовано з можливістю автономного функціонування у форматі окремого програмного застосунку, а також як модуль розширення для інтеграції у сторонні графічні редактори. За результатами комплексного порівняльного аналізу доцільно констатувати, що розглянуте програмне забезпечення вирізняється високим рівнем загальної

ефективності серед наведених аналогів завдяки низці вагомих технологічних переваг. Попри це, об'єктивна якість процедури відновлення кольору все ще не повною мірою задовольняє актуальні науково-практичні вимоги до обробки візуальних даних.

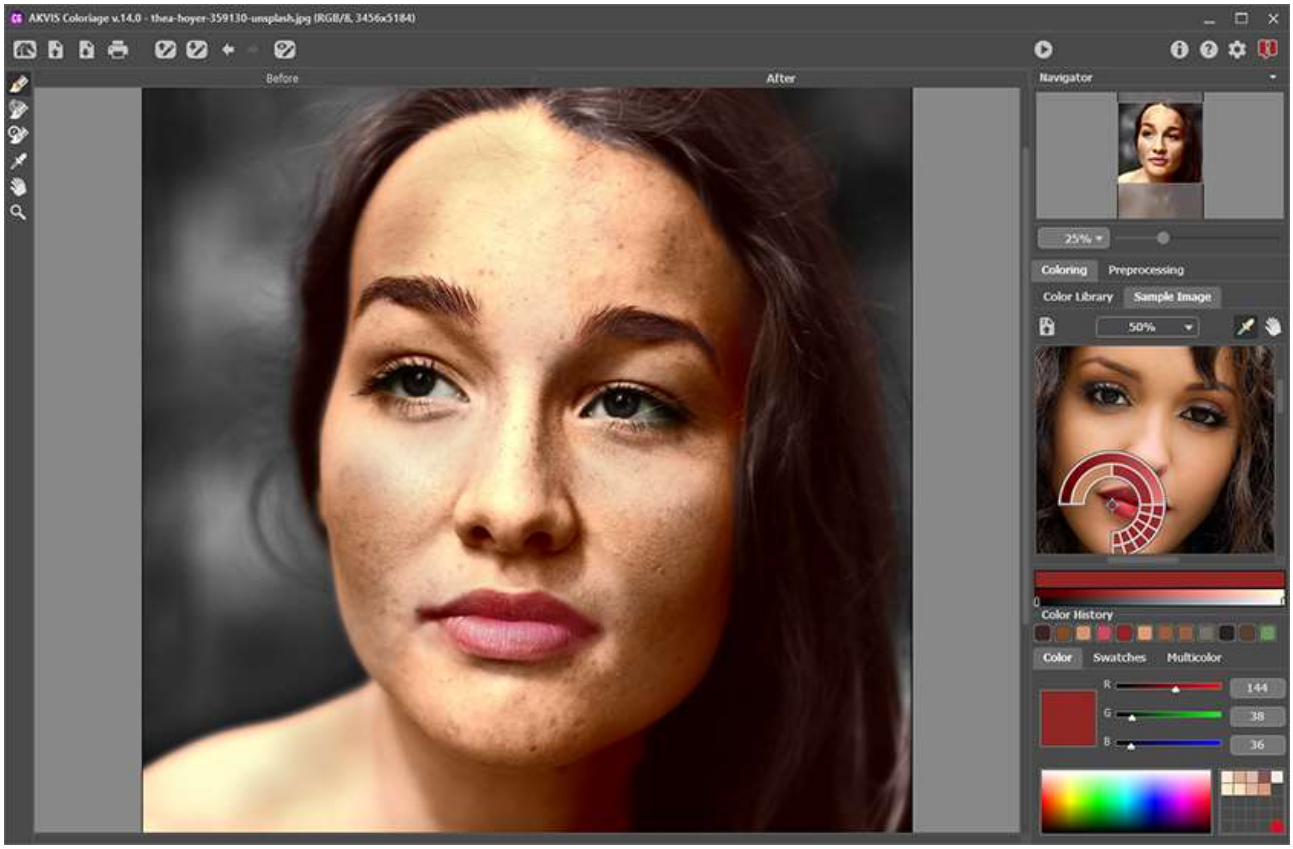


Рисунок 1.4 – Відновлення кольору в програмному середовищі AKVIS Coloriage

1.3. Постановка задачі

У процесі дослідження передбачається глибокий аналіз предметної області відновлення кольору цифрових фотографій. Наступним етапом є дослідження сучасного стану розвитку технологій колоризації, проведення критичного огляду існуючих систем-аналогів з метою обґрунтування доцільності розробки нового програмного продукту та формування відповідного технічного завдання. Крім того, вимагається здійснення обґрунтованого вибору інструментальних засобів розробки, зокрема операційної системи, мови програмування та інтегрованого середовища.

З точки зору функціональних вимог, проектований модуль має забезпечувати можливість завантаження та візуалізації монохромного зображення з подальшим відновленням кольору як в автоматичному режимі, так і за допомогою інтерактивних точкових маркерів. Відповідно до кожної заданої користувачем підказки, алгоритм повинен генерувати спектр імовірних відтінків для цільового сегмента. У разі невідповідності згенерованих варіантів вимогам користувача, система повинна надавати доступ до розширеної колірної палітри для здійснення мануальної корекції. Візуалізація результуючого зображення має здійснюватися паралельно з вихідним монохромним файлом, забезпечуючи динамічне оновлення екрана при кожній ітерації відновлення кольору. Додатковою вимогою є реалізація функції збереження оброблених цифрових матеріалів.

У математичному контексті проблему відновлення кольору доцільно сформулювати як задачу обчислення відповідних компонент тензора C на основі відомих значень елементів матриці G, яка описує вхідне монохромне зображення. Звідси випливає необхідність побудови такого відображення, яке задовольняє наступним критеріям:

1. Просторове розширення результуючого зображення має бути тотожним кількості пікселів вихідного масиву даних.
2. Візуальне сприйняття обробленого фотографічного матеріалу повинно характеризуватися високим ступенем реалістичності з позиції людського зору.

Якщо перший критерій піддається чіткій математичній формалізації, то другий носить виражений суб'єктивний характер і не може бути описаний строгими аналітичними залежностями. Зазначений дуалізм становить фундаментальну складність як у процесі вирішення поставленої задачі, так і під час об'єктивного оцінювання якості отриманих результатів.

Таким чином, здійснено ґрунтовний аналіз базових технологій відновлення кольору фотографій, підтверджено актуальність обраного напрямку дослідження та ідентифіковано ключові недоліки процесу відновлення кольору

монохромних візуальних даних, що переважно зводяться до обмежених можливостей користувача впливати на проміжні та кінцеві результати обробки. Аналітичний огляд існуючих систем-аналогів дозволяє стверджувати про відсутність комплексного програмного рішення, здатного забезпечити одночасно високу швидкість та належну якість відновлення кольору. За результатами порівняння встановлено, що найвищі показники серед досліджуваних систем демонструє програмний продукт AKVIS Coloriage, який вирізняється значною швидкодією. Попри загалом задовільну візуальну якість, на оброблених зображеннях спостерігається генерація специфічних артефактів. Врахування ідентифікованих недоліків дозволило коректно та обґрунтовано здійснити фінальну постановку задачі дослідження.

РОЗДІЛ 2 ПРОГРАМНОГО МОДУЛЯ РОЗФАРБОВУВАННЯ ЦИФРОВИХ ФОТОГРАФІЙ

2.1. Обґрунтування доцільності застосування нейромережевого методу обробки візуальних даних у задачі відновлення кольору монохромних знімків

Процес відновлення кольору передбачає просторову локалізацію областей монохромних зображень з ідентичними спектральними характеристиками. Зазначену підзадачу доцільно формалізувати як класичну проблему розпізнавання образів. У сучасній теорії обробки візуальної інформації виділяють три фундаментальні класи методів розпізнавання, до яких належать статистичні підходи, структурні методи та апарат штучних нейронних мереж.

Статистична методологія ґрунтується на застосуванні математичних вирішальних правил класифікації, синтезованих на основі теорії ймовірностей та математичної статистики. Даний підхід забезпечує побудову оптимального класифікатора за умови наявності апріорної інформації щодо щільності розподілу ймовірностей для всієї генеральної сукупності зображень та ймовірнісних характеристик належності об'єкта до певного класу. У межах цього підходу сутність, що підлягає ідентифікації, репрезентується у формі багатовимірного вектора елементарних ознак.

У парадигмі структурного розпізнавання об'єкт дослідження розглядається як складна ієрархічна система, утворена сукупністю базових примітивів, їхніх атрибутивних властивостей та просторово-логічних взаємозв'язків, які комплексно формують глобальні характеристики об'єкта. Ключовими етапами реалізації даного підходу є екстракція непохідних елементів зображення, їхня синтаксична інтеграція шляхом побудови граматики зображень, а також подальше виконання відповідних процедур синтаксичного аналізу.

Апарат штучних нейронних мереж становить високоефективну альтернативу класичним детермінованим алгоритмам розпізнавання образів. На сучасному етапі розвитку обчислювального інтелекту цей науковий напрям характеризується інтенсивною динамікою, що супроводжується розробкою нових архітектурних рішень та модифікацією існуючих топологій відповідно до

специфіки нових прикладних задач. Штучна нейронна мережа є математичною моделлю з відповідною програмною чи апаратною імплементацією, структурна організація та принципи функціонування якої наслідують біологічні нейронні структури живих організмів. Фундаментальні засади цього напряму були закладені в результаті досліджень когнітивних процесів мозку та спроб їх математичного моделювання, першою з яких стала формальна модель нейрона Воррена Маккаллока та Волтера Піттса [12].

До ключових переваг застосування нейромережевих технологій належать наступні аспекти:

1. Здатність до екстракції прихованих закономірностей. Завдяки властивості генералізації знань на основі навчальної вибірки, нейронна мережа здатна апроксимувати складні нелінійні залежності між вхідними та вихідними даними навіть за умови відсутності аналітичного опису предметної області. Використання традиційних математичних методів та експертних систем у подібних ситуаціях є неефективним.
2. Інваріантність до зашумлення вхідних даних. Нейромережеві моделі демонструють високу стійкість під час обробки інформації з наявністю значної кількості неінформативних шумових компонентів. Процедура попередньої фільтрації часто не є обов'язковою, оскільки алгоритм здатний самостійно ідентифікувати та нівелювати вплив таких перешкод під час обчислень.
3. Адаптивність до динаміки середовища. Нейронні мережі володіють здатністю підлаштовувати власні параметри відповідно до змін зовнішніх умов функціонування. Зокрема, моделі, налаштовані на певне середовище, піддаються оперативному донавчанню у разі флуктуації його характеристик. Крім того, для роботи в нестаціонарних умовах, де статистичні показники змінюються з часом, розробляються архітектури зі здатністю до навчання в режимі реального часу. Високий рівень адаптивності підвищує стабільність системи, проте потребує збалансованого підходу для уникнення втрати продуктивності через

надмірну чутливість до високочастотних збурень. Для повноцінного використання адаптивного потенціалу базові параметри системи повинні поєднувати стійкість до локальних шумів із гнучкістю реакції на глобальні зміни середовища [12].

4. Висока обчислювальна продуктивність. Архітектура нейронних мереж передбачає можливість масової паралельної обробки даних, що забезпечує значне підвищення швидкодії під час апаратної або оптимізованої програмної реалізації [12].
5. Апаратна відмовостійкість. Завдяки розподіленому характеру зберігання інформації у вигляді матриці синаптичних ваг, нейромережеві системи характеризуються високим рівнем стійкості до апаратних збоїв. Деградація продуктивності мережі у разі пошкодження окремих вузлів або зв'язків відбувається градієнтно, і лише критичні руйнування топології здатні суттєво вплинути на її загальну працездатність [12].

З огляду на вищенаведене, для імплементації процесу відновлення кольору монохромних зображень було обрано метод, що базується на застосуванні апарату штучних нейронних мереж.

На структурному рівні нейронна мережа є складною системою взаємопов'язаних обчислювальних елементів. Окремий штучний нейрон виконує базові математичні операції над отриманими сигналами та ретранслює результат іншим вузлам. Незважаючи на обчислювальну примітивність ізольованого нейрона, просторова інтеграція таких елементів у масштабні мережеві структури забезпечує здатність до вирішення високорівневих комплексних задач.

Фундаментальною відмінністю нейромереж від класичних алгоритмів є застосування парадигми навчання замість традиційного імперативного програмування. Процес навчання зводиться до ітеративної оптимізації вагових коефіцієнтів синаптичних зв'язків між нейронами. У результаті успішного навчання формується здатність моделі виявляти складні кореляції між вхідними та вихідними даними, а також генерувати правильні результати під час обробки

нових, неповних або частково спотворених візуальних даних, які не були представлені у початковій навчальній множині.

Топологічно нейронна мережа формується з множини нейронів, згрупованих у послідовні шари. Кожен обчислювальний вузол отримує вектор вхідних сигналів від елементів попереднього шару через відповідні синаптичні з'єднання, що проілюстровано на рисунках 2.1 та 2.2. У межах нейрона здійснюється обчислення зваженої суми всіх вхідних сигналів з подальшим застосуванням функції активації для формування кінцевого вихідного значення:

$$Y = F(S) \tag{2.1}$$

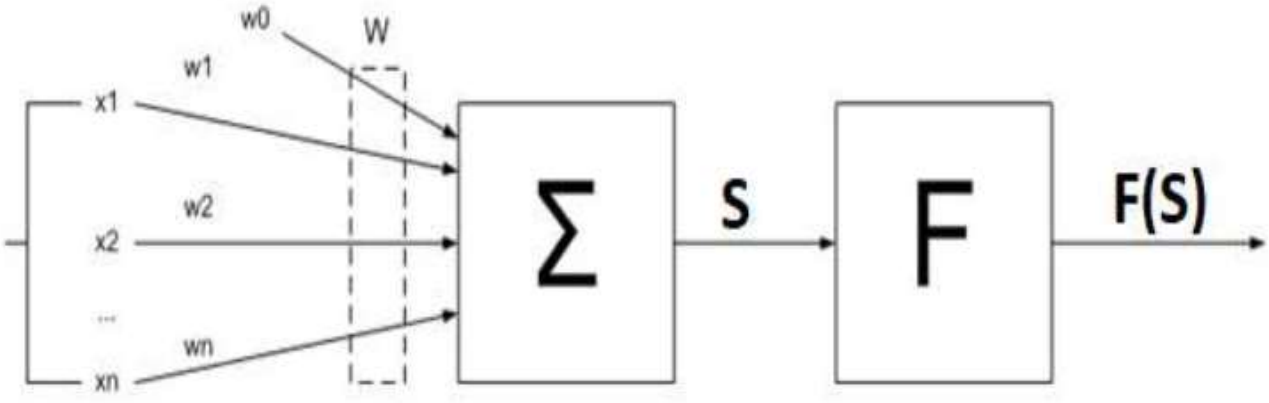


Рисунок 2.1 – Архітектура штучного нейрона

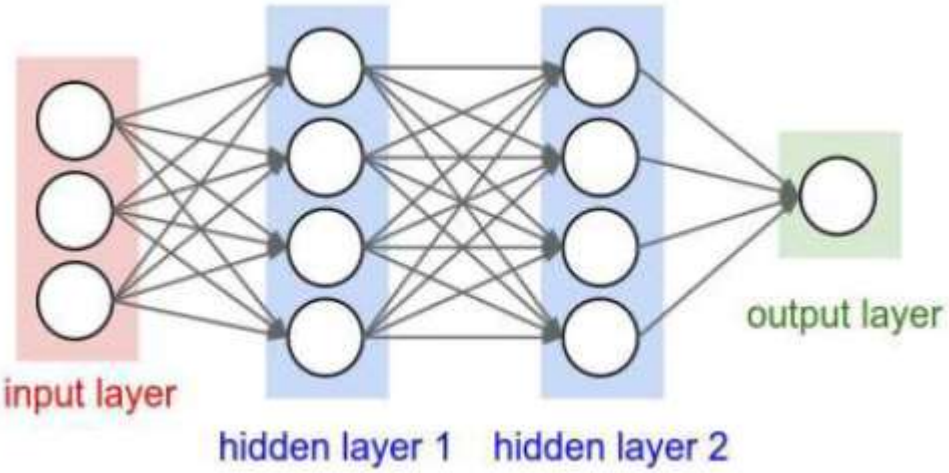


Рисунок 2.2 – Зв'язки між шарами у нейронній мережі

Теоретично функцією активації може виступати довільна математична функція, проте на практиці застосовується обмежений їх клас, що задовольняє низку специфічних критеріїв, виконання яких безпосередньо впливає на ефективність процесу навчання нейронних мереж:

Нелінійність. За умови використання нелінійної функції активації доводиться, що двошарова нейронна мережа набуває властивостей універсального апроксиматора. У протилежному випадку нарощування кількості шарів втрачає математичний сенс, оскільки суперпозиція лінійних функцій залишається лінійним відображенням. Як зазначалося раніше, задачу відновлення кольору доцільно формалізувати як проблему синтезу складного математичного відображення F. Здатність нейромережевих архітектур формувати універсальні наближення виступає фундаментальним обґрунтуванням їх застосування для вирішення поставленої задачі [11].

Безперервна диференційованість. Наявність цієї властивості є критичною математичною умовою для імплементації методів оптимізації параметрів нейронної мережі, що ґрунтуються на обчисленні градієнтів [12].

Обмеженість. Застосування необмежених функцій активації призводить до дестабілізації обчислювальних процесів під час використання оптимізаційних алгоритмів у ході навчання нейронної мережі.

Монотонність. Аналітично доведено, що монотонність функції активації забезпечує опуклість функції помилки в межах окремого ізольованого шару. Ця обставина суттєво розширює арсенал доступних методів вирішення задач оптимізації, оскільки теорія опуклого програмування є значно більш дослідженою [11].

Монотонність похідної. Теоретично та емпірично підтверджено, що застосування функцій активації з цією характеристикою сприяє суттєвому покращенню здатності моделі до генералізації [11].

Апроксимація тотожного відображення в околі нуля. Наявність зазначеної властивості гарантує ефективність процесу навчання за умови ініціалізації вагових коефіцієнтів малими випадковими значеннями. Відсутність цієї характеристики зумовлює необхідність проведення додаткових досліджень для визначення оптимальних стратегій початкового налаштування параметрів моделі [11].

Класичні архітектури нейронних мереж базуються виключно на повнозв'язних шарах. Зазначений підхід демонструє високу ефективність під час вирішення задач малої розмірності, однак при опрацюванні візуальних даних кількість параметрів моделі зростає експоненціально навіть для двошарових конфігурацій. З огляду на це, для алгоритмічної реалізації процесу відновлення кольору монохромних зображень було обрано архітектуру згорткової нейронної мережі.

Згорткові нейронні мережі являють собою клас глибоких штучних нейронних мереж прямого поширення, концептуалізований Яном Лекуном із метою оптимізації процесів ідентифікації та розпізнавання візуальних образів [2].

Термінологічна назва даної архітектури походить від базової математичної операції згортки, сутність якої полягає у поелементному множенні локальної області пікселів зображення на матрицю ваг або ядро згортки з подальшим підсумовуванням результатів та їх проекціюванням у відповідну просторову позицію вихідного тензора.

Принцип функціонування згорткової нейронної мережі традиційно інтерпретується як ієрархічний перехід від екстракції низькорівневих просторових ознак до формування абстрактних високорівневих семантичних концептів. У процесі обробки сигналу мережа алгоритмічно генерує необхідну

ієрархію ознак у вигляді послідовності карт активацій, здійснюючи фільтрацію дрібних деталей та фокусуючись на значущих елементах.

Наведена інтерпретація носить переважно ілюстративний характер. На практиці ознаки, що генеруються глибокими архітектурами, відрізняються високою складністю та важко піддаються прямій структурній інтерпретації. Відповідно, під час розробки прикладних систем не рекомендується здійснювати мануальне конструювання або коригування ознак. Натомість зусилля доцільно спрямовувати на оптимізацію загальної топології та архітектури мережі. На відміну від повнозв'язних архітектур, де кожен обчислювальний вузол з'єднаний з усіма елементами попереднього шару унікальними синаптичними зв'язками, у згорткових нейронних мережах застосовується матриця ваг обмеженої розмірності. Це ядро згортки ітеративно зміщується вздовж площини оброблюваного тензора, генеруючи активаційний сигнал для відповідного просторового нейрона наступного шару. Таким чином, ідентична матриця вагових коефіцієнтів використовується для множини нейронів вихідного шару. Цей механізм інтерпретується як просторове кодування певного візуального патерна, наприклад, наявності градієнтного переходу або контуру під заданим кутом. Шар, утворений у результаті застосування такої матриці, фіксує наявність шуканої ознаки та її просторові координати, формуючи карту ознак. У згортковій архітектурі функціонує ціла сукупність фільтрів, кожен з яких кодує окремі графічні елементи візуального образу. При цьому матриці ваг не задаються апріорі, а оптимізуються самостійно в процесі навчання моделі за допомогою класичного методу зворотного поширення помилки. Застосування множини ядер згортки генерує відповідну кількість незалежних карт ознак, що зумовлює багатоканальну структуру згорткового шару. Процес сканування вхідного тензора зазвичай здійснюється з кроком зміщення, меншим за розмірність самого ядра. Як

приклад, матриця ваг розмірністю 4 × 4 може переміщуватися з кроком в один або два пікселі. Візуалізацію математичної операції згортки із застосуванням ядра розмірністю 3 × 3 наведено на рисунку 2.3 [12].

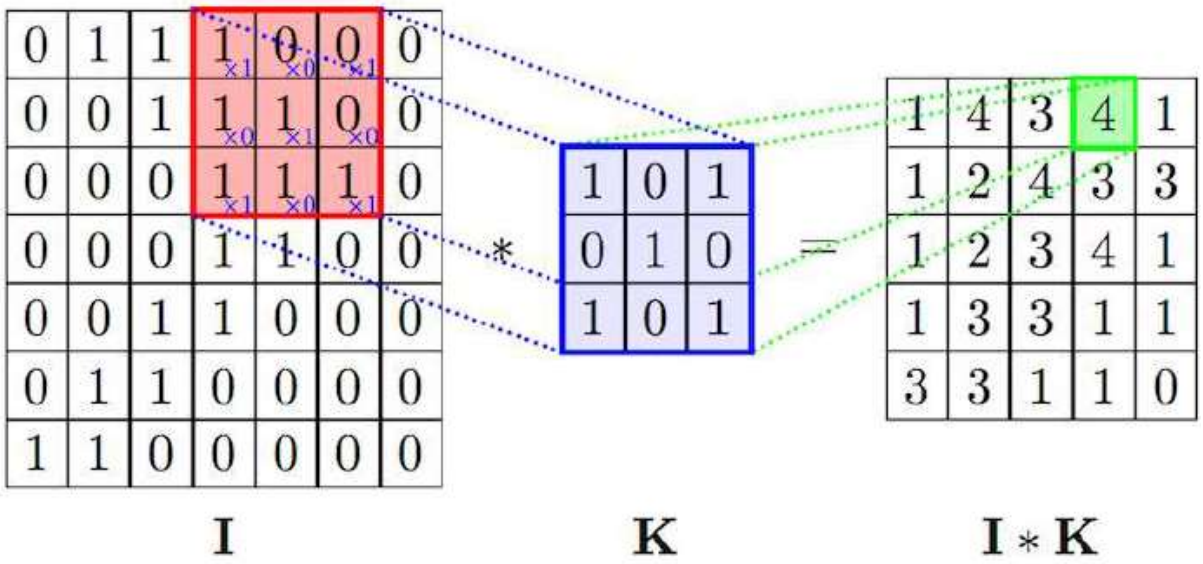


Рисунок 2.3 – Приклад виконання операції згортки

Операція субдискретизації зменшує просторову розмірність карт ознак, сформованих на виході згорткового шару. Базовий принцип даної мережевої архітектури полягає в тому, що факт ідентифікації цільової ознаки має вищий пріоритет порівняно з точним визначенням її просторових координат. З огляду на це, із локальної групи суміжних елементів карти ознак виокремлюється єдине максимальне значення, яке інтегрується до структури ущільненої карти. Застосування зазначеної операції забезпечує не лише прискорення подальших обчислювальних процесів, але й сприяє досягненню інваріантності моделі щодо масштабних трансформацій вхідного зображення [2].

Детальний аналіз типової структурної організації згорткової нейронної мережі дозволяє виділити її багатошарову топологію. Після проходження початкового рецепторного рівня сигнал поширюється через каскад згорткових шарів, де відбувається почергове застосування математичних операцій згортки та просторової субдискретизації. Систематичне чергування зазначених шарів забезпечує можливість синтезу комплексних ієрархічних функцій. Зі збільшенням глибини нейронної мережі просторова розмірність карт ітеративно

зменшується, що супроводжується пропорційним зростанням кількості спектральних каналів. У більшості практичних реалізацій, після проходження багатошарового каскаду, тензор ознак редукується до векторного або скалярного формату при одночасному генеруванні значної кількості таких карт. Традиційна архітектура передбачає інтеграцію кількох повнозв'язних шарів на фінальному етапі конвеєра обробки, на вхід яких транслюються результуючі карти ознак.

Згортковий шар виступає фундаментальним структурним компонентом згорткової нейромережевої архітектури. Даний елемент інкапсулює незалежний фільтр для кожного інформаційного каналу. Ядро згортки здійснює фрагментарне опрацювання вихідних даних попереднього шару шляхом агрегації результатів матричного множення для кожного локального рецепторного поля. Початкові значення вагових коефіцієнтів згорткових ядер є невідомими та підлягають ітеративній оптимізації в процесі навчання моделі [2]. Характерною властивістю згорткового шару є порівняно низька розмірність простору параметрів, що підлягають налаштуванню під час тренування. Як приклад, за умови наявності вхідного зображення розмірністю 100 на 100 пікселів у трьох колірних каналах та застосування згорткового шару з просторовим ядром 3 на 3 пікселі при шестиканальному виході, процедура навчання вимагає оптимізації лише 162 параметрів. Зазначений обсяг обчислень є незіставно меншим порівняно з повнозв'язною архітектурою, яка за аналогічних умов потребувала б налаштування близько 100 мільйонів вагових коефіцієнтів.

Шар субдискретизації реалізує процедуру нелінійного ущільнення карти ознак, під час якої локальна група пікселів редукується до єдиного елемента шляхом застосування специфічного нелінійного оператора. Найбільш поширеною математичною функцією для виконання цієї операції є обчислення максимуму, що проілюстровано на рисунку 2.4 [23]. Процес перетворення охоплює суміжні прямокутні або квадратні області, що не перетинаються, в межах яких здійснюється екстракція пікселя з максимальним значенням

інтенсивності. Застосування субдискретизації дозволяє суттєво зменшити просторовий об'єм оброблюваних даних. Семантична інтерпретація цієї операції зводиться до того, що після просторової локалізації певних візуальних ознак на етапі згортки, надлишкова деталізація втрачає свою інформативність для подальшого аналізу і підлягає редукції. Додатковою перевагою відсіювання малозначущих деталей є мінімізація ризику перенавчання моделі. Структурно шар субдискретизації традиційно інтегрується в архітектуру безпосередньо після згорткового шару та перед наступним каскадом згортки, як наочно продемонстровано на рисунку 2.4.

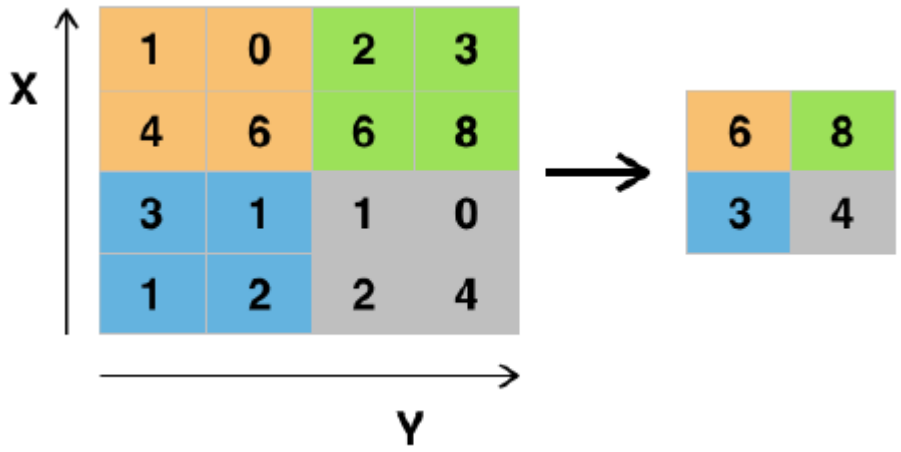


Рисунок 2.4 – Субдискретизація з кроком 2 та просторовим розміром вікна 2×2

Окремої уваги при формуванні архітектури мережі заслуговує математична операція транспонованої згортки, яку також термінологічно визначають як деконволюцію. За своєю функціональною сутністю зазначена процедура виступає оберненою до класичної операції згортки. Практичне застосування стандартної згортки з кроком зміщення, що перевищує одиницю, закономірно зумовлює зменшення просторової розмірності результуючого зображення. Натомість імплементація транспонованої згортки забезпечує просторове розширення вихідного тензора, що наочно продемонстровано на рисунку 2.5 [4].

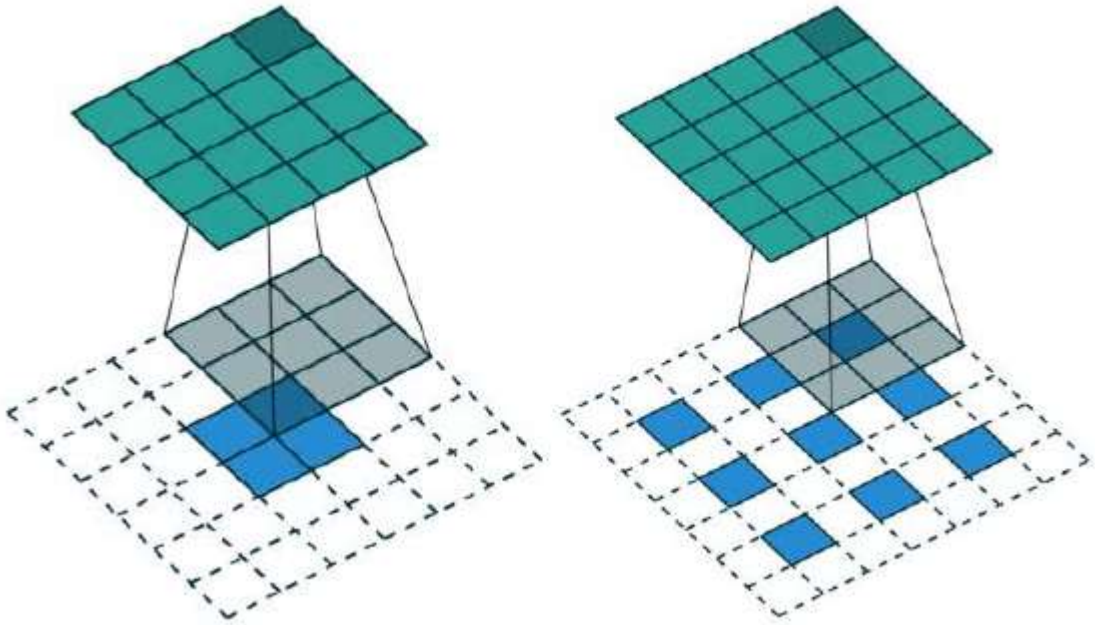


Рисунок 2.5 – Операція класичної та транспонованої згортки

У практичних обчислювальних реалізаціях математичний апарат транспонованої згортки переважно застосовується з метою реконструкції просторової розмірності тензора після його проходження через каскад шарів субдискретизації або стандартних згорткових перетворень. Зазначений алгоритмічний підхід набув широкого застосування в архітектурах генеративних моделей, функціональне призначення яких полягає у синтезі нових візуальних структур, просторові та семантичні характеристики яких відрізняються від параметрів початкових вхідних даних.

2.2. Процес навчання нейронної мережі

Алгоритм або правило навчання становить собою математичну модель або формалізований метод, спрямований на ітеративне підвищення обчислювальної продуктивності штучної нейронної мережі шляхом систематичної модифікації її параметрів.

У теорії машинного навчання виділяють три фундаментальні парадигми навчання нейромережевих структур: навчання з учителем, навчання без учителя та навчання з підкріпленням.

Для архітектур згорткових нейронних мереж найбільш доцільним і науково обґрунтованим є застосування парадигми навчання з учителем, що підтверджується результатами чисельних досліджень. Зазначений підхід базується на наявності відповідного цільового вихідного вектора для кожного вхідного вектора даних. Сукупність таких пар формує репрезентативну навчальну вибірку. Процедура навчання полягає в ітеративній подачі навчальних векторів на рецепторний шар мережі, обчисленні величини відхилення між фактичним і цільовим значеннями вихідного сигналу та відповідній корекції синаптичних ваг з метою мінімізації цієї похибки.

Кількісною мірою продуктивності нейронної мережі слугує функція втрат, яка підлягає оптимізації під час тренування моделі. За умови коректного вибору та налаштування цієї функції, системне зменшення її значення свідчить про підвищення точності вирішення поставленої задачі. Відповідно, процес навчання нейронної мережі математично формалізується як задача мінімізації цільової функції втрат, для розв'язання якої можуть застосовуватися різноманітні методи математичної оптимізації [2].

Попри те, що концептуальні засади штучних нейронних мереж були закладені на початку минулого століття, їхнє інтенсивне практичне застосування стало можливим відносно недавно. Каталізатором цього процесу стала розробка алгоритму зворотного поширення помилки [6]. Впровадження цього методу суттєво підвищило ефективність навчання багатошарових архітектур порівняно з попередніми математичними підходами. З огляду на високу результативність, даний алгоритм було обрано як базовий для оптимізації проектованої нейромережевої моделі.

Фундаментальний принцип алгоритму зворотного поширення полягає у трансляції сигналу помилки від вихідних шарів мережі до вхідних у напрямку, протилежному до прямого розповсюдження інформаційного сигналу. На кожній ітерації алгоритму здійснюється обчислення функції втрат та подальша модифікація вагових коефіцієнтів з метою підвищення точності апроксимації на елементах навчальної множини.

Для визначення напрямку та кроку коригування параметрів обчислюється градієнт функції втрат за всіма ваговими коефіцієнтами моделі. Це зумовлює строгу математичну вимогу щодо безперервної диференційованості функцій активації нейронів та наявності аналітичних виразів для їхніх похідних. За своєю математичною сутністю даний алгоритм є спеціалізованою модифікацією класичного методу градієнтного спуску [2].

Процедура мінімізації функції втрат становить собою складну багатовимірну задачу оптимізації. Це пояснюється наявністю значної кількості параметрів моделі, яка в сучасних обчислювальних реалізаціях може сягати мільйонів, а також складною топологією гіперповерхні цільової функції, що характеризується наявністю великої кількості локальних мінімумів.

До масового впровадження нейромережевих технологій існувала низка альтернативних алгоритмічних підходів до відновлення кольору візуальних даних. Одним із фундаментальних методів, що став основою для багатьох подальших досліджень, є формулювання процедури відновлення кольору як задачі умовної оптимізації. У межах цього підходу здійснюється розрахунок невідомих значень хроматичних каналів aта b, тоді як значення яскравості з каналу L вважається відомим і використовується як базове монохромне подання вхідного зображення.

У просторовому поданні кожен елемент піксельної сітки характеризується двовимірними координатами x та y. Відповідно, на вхід алгоритму подається розподіл інтенсивності $L = L(x, y)$, а результатом його роботи є згенеровані масиви хроматичних значень $a(x, y)$ та $b(x, y)$. Для спрощення математичного опису просторові координати доцільно позначити радіус-вектором r.

В основу зазначеного алгоритму покладено евристичне припущення, згідно з яким просторово суміжні пікселі з близькими значеннями інтенсивності повинні мати подібні хроматичні характеристики. Математично ця умова зводиться до необхідності мінімізації відхилення між значеннями колірних компонентів $a(r)$ і $b(r)$ та їхнім зваженим середнім значенням в околі суміжних елементів. Формалізований запис цієї різниці на прикладі хроматичного каналу a набуває відповідного аналітичного вигляду:

$$J(a) = \sum_{r} \left(a(r) - \sum_{s \in N(r)} w_{rs} a(s) \right)^2 \quad (2.2)$$

де w_{rs} — вагова функція, сума значень якої дорівнює одиниці. Зазначена функція набуває максимальних значень за умови наближення показника інтенсивності $L(r)$ до значення $L(s)$, та пропорційно зменшується у протилежному випадку;

$N(r)$ — множина просторово суміжних пікселів відносно елемента з координатою r.

Описаний математичний підхід визначено як базовий для практичної реалізації алгоритму в межах даної кваліфікаційної роботи.

2.3. Обґрунтування вибору колірної моделі

Колірна модель являє собою математичну абстракцію, призначену для формалізованого опису колірних відтінків за допомогою впорядкованих числових масивів, що переважно складаються з трьох або чотирьох елементів. Зазначені числові значення термінологічно визначаються як колірні компоненти або координати. Сукупність усіх допустимих значень, згенерованих у межах обраної математичної моделі, формує відповідний багатовимірний колірний простір.

Метод параметризації та алгоритмічного представлення хроматичної інформації здійснює фундаментальний вплив на ефективність вирішення задачі відновлення кольору. Зокрема, встановлено, що класичний колірний простір RGB демонструє значно нижчу результативність у більшості прикладних обчислювальних задач порівняно зі спеціалізованими, хоча й менш поширеними, альтернативними моделями.

Модель RGB, що базується на просторовій комбінації червоного, зеленого та синього базових каналів, належить до класу адитивних колірних систем. Спектральний базис даної системи обґрунтований фізіологічними та оптичними закономірностями функціонування зорового апарату людини. Адитивна

природа цього простору полягає у синтезі цільового відтінку шляхом пропорційного додавання інтенсивностей базових випромінювань до нульового фонового рівня, що відповідає абсолютно чорному кольору, як це наочно проілюстровано на рисунку 2.6.

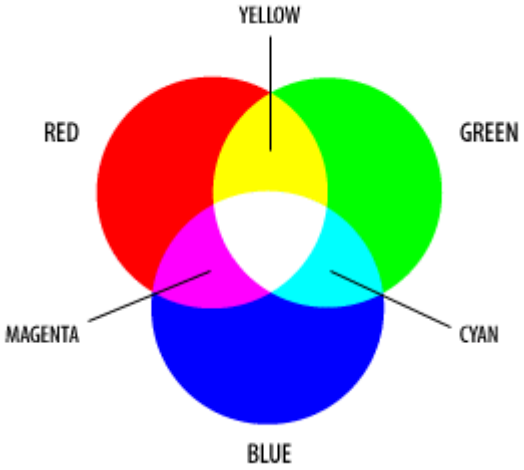


Рисунок 2.6 – Принцип адитивності колірної моделі RGB

Будь-який колірний відтінок у цій системі піддається математичній формалізації через вектор із трьох числових компонентів, значення яких належать до нормалізованого діапазону від нуля до одиниці. Відповідно, геометрична інтерпретація всієї сукупності колірного простору набуває форми одиничного тривимірного куба, що наочно продемонстровано на рисунку 2.7.

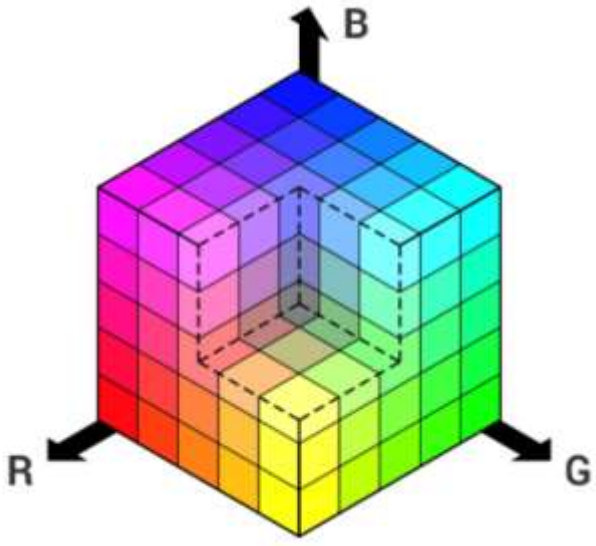


Рисунок 2.7 – Геометрична інтерпретація колірної моделі RGB у вигляді одиничного куба

До суттєвих недоліків колірної моделі RGB слід віднести наступні фактори:

Апаратна залежність.

Підсумковий хроматичний відтінок, синтезований шляхом адитивного змішування базових компонентів, детермінується фізико-технічними властивостями конкретного випромінювального елемента або люмінофора.

Обмеженість колірного охоплення.

Даний простір характеризується звуженим діапазоном відтворюваних відтінків, що обмежує точність передачі кольору на різних технічних пристроях незалежно від їхньої внутрішньої колірної гами.

Альтернативною системою просторового подання є колірна модель HSV, яка базується на трьох фундаментальних координатах:

- Колірний тон, який визначає безпосередньо хроматичний відтінок. Цей параметр традиційно вимірюється в кутових величинах від нуля до трьохсот шістдесяти градусів, або ж нормується в еквівалентних діапазонах від нуля до одиниці чи до ста відсотків;
- Насиченість, що характеризує інтенсивність або спектральну чистоту кольору і нормується в межах від нуля до ста відсотків або від нуля до одиниці. Наближення цього показника до максимального значення свідчить про високу спектральну чистоту, тоді як зниження параметра до нуля зумовлює деградацію кольору в бік ахроматичного сірого градієнта;
- Яскравість, яка визначає загальний рівень освітленості колірного вектора та використовує ідентичні шкали математичного нормування.

У просторовому відношенні модель HSV алгоритмічно репрезентується в циліндричній системі координат або візуалізується у геометричній формі конуса, що відображено на рисунку 2.8.

До базових недоліків колірної моделі HSV належать:

- Успадкування функціональних обмежень попередньої колірної моделі;
- Фундаментальна обмеженість метричного колірного простору.

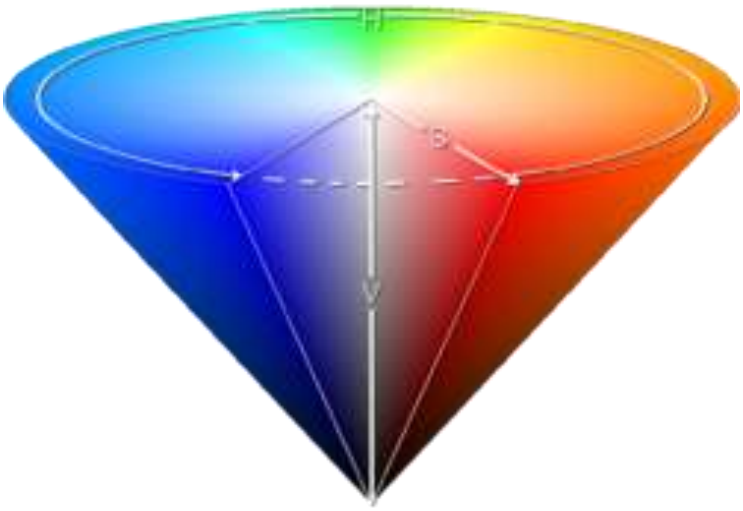


Рисунок 2.8 – Колірна модель HSV

Колірна модель CIELAB була синтезована з метою формування такого колірного простору, у якому градієнтна зміна спектральних характеристик характеризується лінійністю відносно особливостей людського візуального сприйняття. Відповідно, ідентичні кількісні флуктуації значень колірних координат у довільних сегментах цього простору зумовлюють еквівалентну суб'єктивну зміну відтінку [34]. Застосування зазначеної математичної абстракції дозволяє ефективно нівелювати нелінійний характер перцепції кольору зоровим апаратом людини.

Структурно дана модель містить канал світлоти L, функціональне призначення якого є тотожним компоненту Y у системі YUV, а також дві хроматичні координати: вісь a, що визначає спектральний діапазон від зеленого до червоного кольорів, та вісь b, яка описує перехід від синього до жовтого спектра, що наочно проілюстровано на рисунку 2.9.

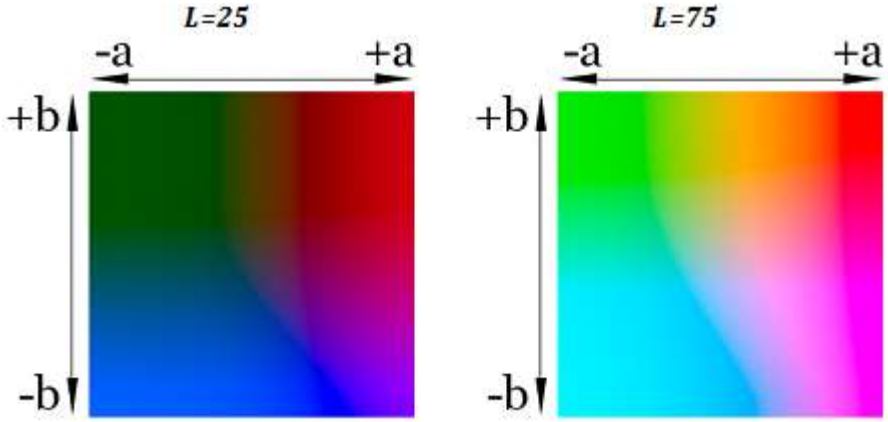


Рисунок 2.9 – Залежність колірної моделі CIELAB від значень хроматичних координат a та b за умови фіксованих показників світлоти L на рівні 0.25 та 0.75

До безперечних переваг застосування колірної моделі CIELAB належать [3]:

1. Абсолютна апаратна незалежність, що гарантує ідентичність відтворення спектральних характеристик на різних технічних пристроях;
2. Розширене колірне охоплення, яке суттєво перевищує діапазони відтворюваних відтінків у системах RGB та HSV;
3. Універсальність математичного апарату, що дозволяє використовувати параметри цієї системи як фундаментальний базис для точного обчислення координат у будь-яких інших колірних просторах.

Зважаючи на вищенаведені функціональні характеристики, зазначену математичну модель було визначено як оптимальну для реалізації алгоритмічного механізму інтеграції колірних маркерів до вихідного монохромного зображення в межах загальної процедури відновлення кольору.

2.4. Розробка архітектури штучної нейронної мережі та алгоритмічної схеми для відновлення кольору цифрових фотографій

Класична архітектура нейронної мережі для відновлення кольору базується на топології кодера та декодера, що структурно складається з каскаду послідовних згорткових шарів із подальшим застосуванням відповідних шарів транспонованої згортки [3]. Зазначена архітектурна концепція забезпечує

здатність моделі ідентифікувати комплексні просторові патерни аж до вузла максимального стиснення інформації. Даний функціональний сегмент нейронної мережі термінологічно визначається як кодер.

Суміжний сегмент мережевої структури виконує обернену функцію, яка полягає у генерації цільового вихідного масиву даних, що в контексті даного дослідження репрезентує зображення з відновленим кольором. Цей функціональний блок ідентифікується як декодер.

Проте для ефективного вирішення задачі відновлення кольору описаного базового підходу виявляється недостатньо. У процесі просторової редукції тензора ознак такою архітектурою відбувається незворотна втрата значного обсягу низькорівневої інформації, зокрема даних про просторові межі контурів або об'єкти складної морфології. З огляду на це, обґрунтовано застосування альтернативного підходу до проектування топології, відомого як архітектура U-net. Фундаментальна особливість цієї моделі полягає у використанні прокидних зв'язків, завдяки яким обчислювальні шари декодера отримують вхідні масиви даних не лише від безпосередньо попереднього рівня, але й від симетричного йому шару з боку кодера, що наочно проілюстровано на рисунку 2.10.

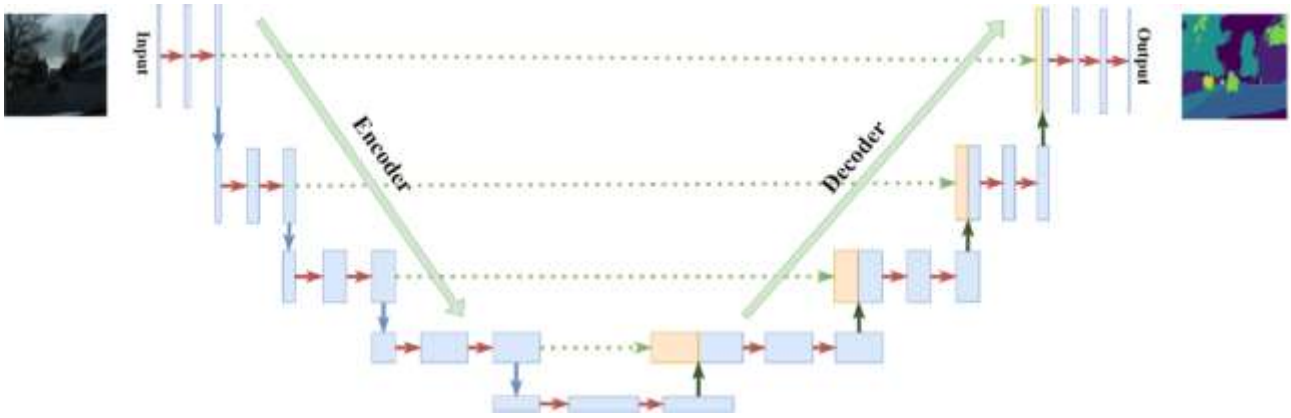


Рисунок 2.10 – Структурна організація нейромережевих архітектур типу кодер-декодер та U-net

Суттєвий вплив на якісні показники процесу навчання має застосування шарів нормалізації. У запропонованій архітектурі їх інтеграція передбачена безпосередньо після кожного згорткового шару.

Окремої уваги потребує питання застосування методів регуляризації L1 та L2. Як зазначалося раніше, їх використання в даному контексті є недоцільним, оскільки це провокує генерацію візуальних артефактів у вигляді розмитих контурів та надмірно згладжених колірних градієнтів. З огляду на це, механізми запобігання перенавчанню реалізуються виключно через застосування алгоритмів пакетної нормалізації та методу випадкового виключення нейронів.

Значною обчислювальною перевагою є синтез результуючих зображень у колірному просторі CIELAB замість традиційної моделі RGB. Це зумовлено тим, що канал світлоти L за своєю фізичною сутністю є абсолютним еквівалентом вхідного монохромного зображення. Завдяки цьому математичному спрощенню архітектура оптимізується для алгоритмічної генерації лише двох хроматичних каналів замість трьох. В якості теоретичної альтернативи розглядалася можливість використання простору HSV з інтеграцією вихідних монохромних значень у канал V, однак емпіричним шляхом було встановлено негативний вплив такого підходу на якість навчання моделей. Це пояснюється математичною сингулярністю даного простору, за якої при нульовій насиченості зміна колірного тону жодним чином не впливає на результуючий візуальний образ.

Фундаментальним недоліком більшості існуючих алгоритмів для автоматизованого відновлення кольору є відсутність дієвих механізмів інтерактивного впливу користувача на проміжний чи кінцевий результат, навіть за умови наявності достовірних апріорних відомостей щодо справжньої колірної гами окремих сегментів сцени. Наслідком цього є часта невідповідність згенерованих кольорів об'єктів їхнім реальним історичним або фізичним прототипам. Крім того, на основі наявної монохромної інформації для одного й того ж об'єкта може існувати ціла множина візуально правдоподібних хроматичних варіацій. Зважаючи на зазначену проблематику, у даній магістерській кваліфікаційній роботі імплементується метод відновлення кольору, що базується на просторовому поширенні локальних інтерактивних підказок користувача.

Базова архітектура згорткової нейронної мережі U-net, розроблена в межах поточного дослідження, була структурно модифікована шляхом інтеграції додаткової підмережі для обробки користувацьких маркерів, що наочно проілюстровано на рисунку 2.11.

Вхідним вектором даних слугує монохромне зображення у вигляді каналу світлоти L колірної моделі CIE Lab, бінарна просторова маска, де одиничні значення відповідають просторовим координатам розміщення користувацьких підказок, а також два хроматичні канали a та b, що безпосередньо містять кількісні спектральні значення цих маркерів.

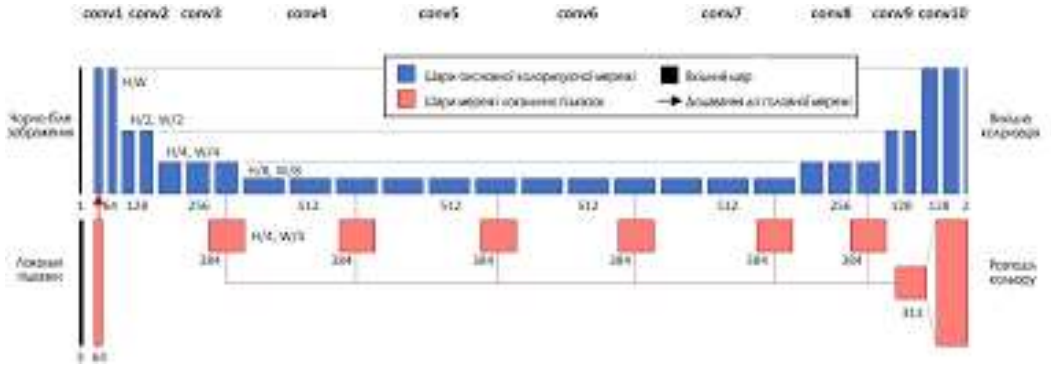


Рисунок 2.11 – Структурна організація нейронної мережі для відновлення кольору цифрових фотографій

На противагу традиційним методам обробки, паралельно із синтезом базових хроматичних каналів a та b, запропонована нейромережева архітектура здійснює генерацію розподілу колірних значень для кожного окремого пікселя растра, що надалі використовується для забезпечення функціональності інтерфейсу користувача.

Структурно нейронна мережа декомпозується на дві взаємодіючі складові: основну мережу відновлення кольору, імплементовану на базі архітектури U-net із застосуванням десяти послідовних згорткових блоків, та допоміжну підмережу обробки локальних інтерактивних маркерів.

Відповідно до визначених концептуальних засад, спроектовано загальну блок-схему алгоритму функціонування спеціалізованого програмного модуля відновлення кольору цифрових фотографій, яку наочно проілюстровано на рисунку 2.12.

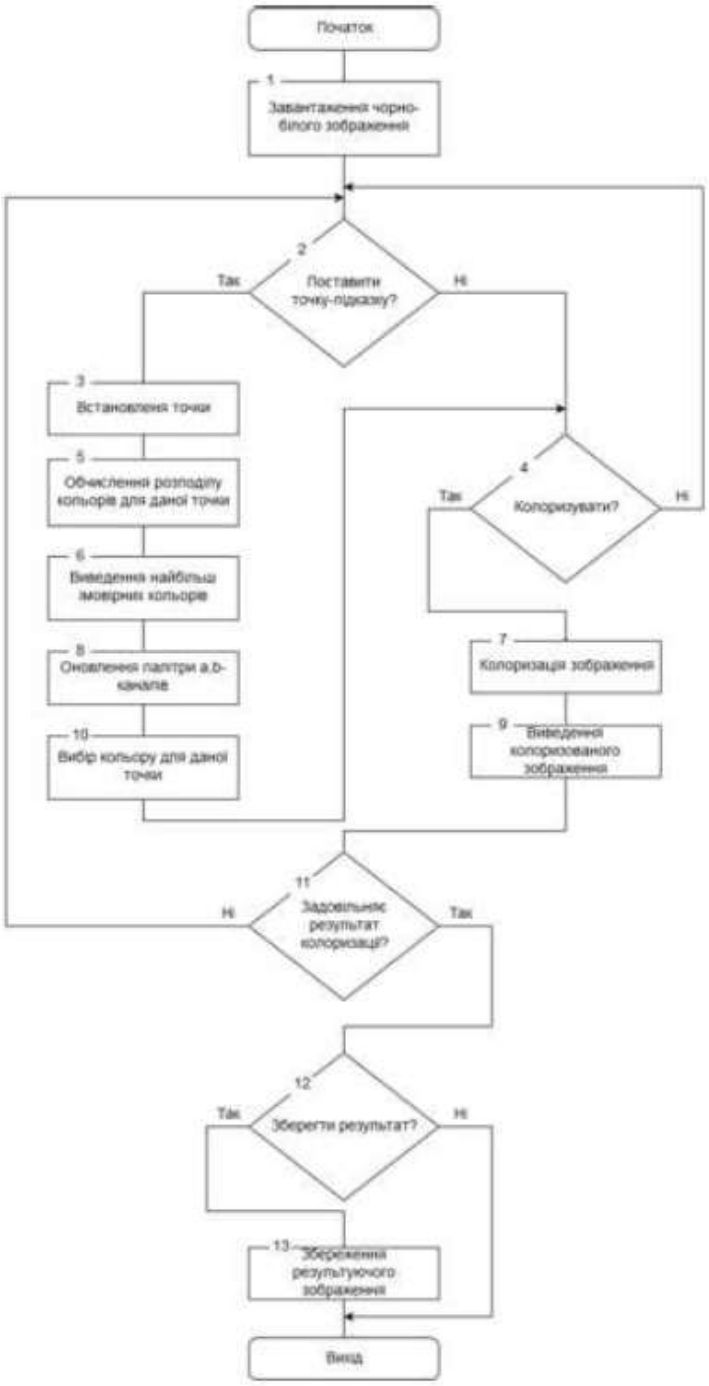


Рисунок 2.12 – Блок-схема алгоритму програмного модуля для відновлення кольору цифрових фотографій

2.5. Обґрунтування вибору інструментальних засобів та середовища розробки

У процесі вибору інструментальних засобів розробки програмного забезпечення виникає необхідність врахування комплексу різноманітних факторів, серед яких фундаментальне значення має вибір мови програмування,

оскільки саме вона детермінує спектр супутніх технологій та доступних інструментів. Зокрема, для реалізації графічного інтерфейсу користувача необхідна сумісна бібліотека візуальних компонентів, що забезпечує наявність готових елементів управління. Базовими критеріями вибору мови програмування в межах даного дослідження визначено загальну продуктивність процесу розробки, ефективність функціонування кінцевого програмного продукту, оптимізацію використання ресурсів оперативної пам'яті та рівень кросплатформної універсальності.

Зазначеним критеріям універсальності для створення складних програмних систем відповідають три сучасні мови програмування: Java, C# та Python.

Мова Java становить собою об'єктно-орієнтовану парадигму програмування, початково представлену корпорацією Sun Microsystems у тисяча дев'яносто п'ятому році як базовий компонент однойменної платформи. На сучасному етапі розвиток мови підтримується компанією Oracle. Синтаксична структура мови значною мірою успадкувала принципи C та C++. В офіційній реалізації вихідний код транслюється у проміжний байт-код, який згодом інтерпретується віртуальною машиною під конкретну апаратну платформу. Корпорація Oracle забезпечує доступ до компілятора та віртуальної машини, що відповідають відкритим специфікаціям.

C# є об'єктно-орієнтованою мовою програмування з безпечною системою типізації, розробленою спеціалістами під егідою Microsoft Research. Синтаксичні конструкції C# демонструють подібність до C++ та Java. Мова характеризується строгою статичною типізацією, підтримкою поліморфізму, перевантаженням операторів, наявністю вказівників на функції-члени класу, атрибутів, подій та властивостей. Акумулювавши передовий досвід попередників, розробники C# вилучили з архітектури потенційно конфліктні механізми проектування, зокрема множинне успадкування класів.

Python являє собою інтерпретовану високорівневу об'єктно-орієнтовану мову програмування з динамічною семантикою, розроблену в тисяча дев'ятисот дев'ностому році Гвідо ван Россумом. Наявність високорівневих структур даних у

поєднанні з динамічним зв'язуванням забезпечує високу ефективність для швидкого прототипування та розробки програмного забезпечення, а також для інтеграції існуючих програмних компонентів. Архітектура Python підтримує розгалужену модульну систему, що сприяє повторному використанню програмного коду. Інтерпретатор та стандартні бібліотеки є кросплатформними. Мова підтримує множинні парадигми програмування, включаючи процедурну, об'єктно-орієнтовану, функціональну та аспектно-орієнтовану.

Критичним недоліком C# у контексті даного дослідження є пропрієтарний характер ліцензування та обмеженість повноцінного функціоналу платформою Windows. Зважаючи на вимогу забезпечення повної підтримки UNIX-подібних операційних систем, доцільний вибір звужується до Java та Python. Швидкість розробки програмного коду на Python є суттєво вищою порівняно з Java завдяки механізмам динамічної типізації. Крім того, інтерпретована природа мови значно спрощує процедуру внесення модифікацій та конфігурування системи на етапі впровадження.

Вирішальним аргументом на користь Python стала наявність найширшої екосистеми спеціалізованих бібліотек для роботи зі штучними нейронними мережами, серед яких виділяються Theano, Pylearn2, Deepnet, Caffe та Keras. На основі наведеного аналізу мову Python визначено як базовий інструмент програмування.

Інтегроване середовище розробки становить собою комплексне програмне забезпечення, призначене для оптимізації процесу створення програмних продуктів. Стандартна архітектура такого середовища охоплює:

- текстовий редактор коду;
- компілятор або інтерпретатор;
- інструменти автоматизації процесів збирання;
- вбудований налагоджувач.

Сучасні середовища додатково інтегруються із системами контролю версій та містять спеціалізований інструментарій для проектування графічного інтерфейсу. Для підтримки об'єктно-орієнтованого програмування передбачено

засоби аналізу ієрархії класів та інспектори об'єктів. Незважаючи на існування універсальних рішень, що підтримують декілька мов, таких як Eclipse, Qt Creator чи Visual Studio, більшість середовищ оптимізується під конкретні технологічні стеки.

Головною метою використання інтегрованого середовища є максимізація продуктивності розробника шляхом об'єднання розрізнених інструментів в єдиний програмний комплекс зі зручним інтерфейсом користувача. Це мінімізує часові витрати на перемикання між різними етапами розробки. Зазвичай усередині єдиної програми здійснюється весь цикл створення, модифікації, компіляції, розгортання та налагодження коду. Наявність засобів статичного аналізу забезпечує миттєвий зворотний зв'язок щодо синтаксичних помилок, що суттєво підвищує загальну ефективність роботи.

Враховуючи раніше обґрунтовану перевагу екосистеми Python у сфері машинного навчання, для імплементації алгоритмів відновлення кольору монохромних зображень було обрано бібліотеку Caffe.

Розробка фреймворку Caffe була ініційована у дві тисячі тринадцятому році дослідниками Каліфорнійського університету в Берклі і наразі активно підтримується науковою спільнотою розробників. Програмний продукт розповсюджується на умовах відкритої ліцензії.

Ядро бібліотеки реалізовано мовою C++ із забезпеченням відповідних програмних інтерфейсів для Python та MATLAB. Офіційна підтримка функціонування охоплює операційні системи Linux та OS X. Для виконання ресурсоємних матричних та векторних обчислень система інтегрується з оптимізованими математичними бібліотеками. З метою прискорення обчислювальних процесів фреймворк забезпечує повноцінну підтримку апаратного прискорення на базі графічних процесорів із використанням технологій CUDA та спеціалізованої бібліотеки примітивів глибокого навчання cuDNN.

Архітектура Caffe надає вичерпний інструментарій для проектування, тренування та тестування нейромережевих моделей високого рівня складності. Даний фреймворк володіє низкою суттєвих переваг порівняно з аналогами:

- Оптимізована архітектура гарантує швидку імплементацію моделей. Топологія мереж визначається за допомогою декларативних конфігураційних файлів, що виключає необхідність жорсткого кодування параметрів. Це забезпечує безперешкодне перемикання обчислень між центральним та графічним процесорами, дозволяючи виконувати навчання на високопродуктивних робочих станціях із подальшим перенесенням на інші обчислювальні системи.
- Відкритий вихідний код надає можливості не лише для аудиту алгоритмів, але й для їх глибокої модифікації відповідно до специфічних вимог. Завдяки цій властивості довкола проекту сформувалася потужна екосистема незалежних розробників.
- Висока обчислювальна швидкодія робить даний інструмент оптимальним для вирішення практичних задач комерційного рівня. Навіть при використанні базових апаратних конфігурацій бібліотека здатна обробляти сотні мільйонів зображень щодоби. За показниками швидкодії серед архітектур згорткових нейронних мереж даний фреймворк займає провідні позиції.

Підсумовуючи результати проведеного системного аналізу, у роботі обґрунтовано вибір оптимальної колірної моделі для вирішення задачі відновлення кольору візуальних даних. Доведено доцільність застосування архітектури згорткових нейронних мереж як найбільш ефективного математичного апарату для виконання зазначеної процедури. Здійснено теоретичний вибір та обґрунтування алгоритмів навчання нейромережевої моделі.

У результаті дослідження синтезовано спеціалізовану топологію нейронної мережі для відновлення кольору цифрових фотографій, фундаментом якої слугує архітектура U-net. Додатково розроблено комплексну алгоритмічну схему функціонування проектованого програмного модуля відновлення кольору цифрових фотографій.

РОЗДІЛ 3 РЕАЛІЗАЦІЯ МОДУЛЯ РОЗФАРБОВУВАННЯ ЦИФРОВИХ ФОТОГРАФІЙ

3.1. Реалізація алгоритмів програмного модуля розфарбовування цифрових фотографій

Етапами імплементації методів машинного навчання виступають процеси тренування, валідації та тестування. Тренувальна стадія характеризується циклічністю, де кожна ітерація становить одну епоху, що передбачає повне проходження навчальної вибірки даних. Етап валідації ініціюється після завершення кожної навчальної епохи та призначений для об'єктивного оцінювання динаміки покращення прогностичних властивостей моделі на раніше невідомих даних із валідаційної множини, що проілюстровано на рисунку 3.1.

Алгоритм навчання та валідації формалізується у вигляді такої послідовності кроків:

1. Ініціалізація архітектури та відповідних підмодулів штучної нейронної мережі;
2. Стохастична ініціалізація синаптичних вагових коефіцієнтів;
3. Конфігурування цільової функції втрат;
4. Визначення та налаштування алгоритму оптимізації;
5. Формування навчальної та валідаційної вибірок даних;
6. Створення генераторів даних для ітеративного подання елементів тестової та валідаційної множин;
7. Реалізація епохи навчання, яка для кожного обчислювального міні-пакета з генератора навчальних даних включає:
    - обнулення градієнтів;
    - виконання прямого поширення сигналу;
    - обчислення значення функції втрат;
    - виконання алгоритму зворотного поширення помилки;
    - здійснення кроку оптимізації вагових коефіцієнтів;
8. Реалізація епохи валідації, яка для кожного міні-пакета з валідаційного генератора передбачає:

- виконання прямого поширення сигналу;
- обчислення показників функції втрат;

9. Проміжне збереження параметрів моделі;
10. Тонке налаштування гіперпараметрів нейронної мережі;
11. Експорт та фінальне збереження оптимізованої моделі.

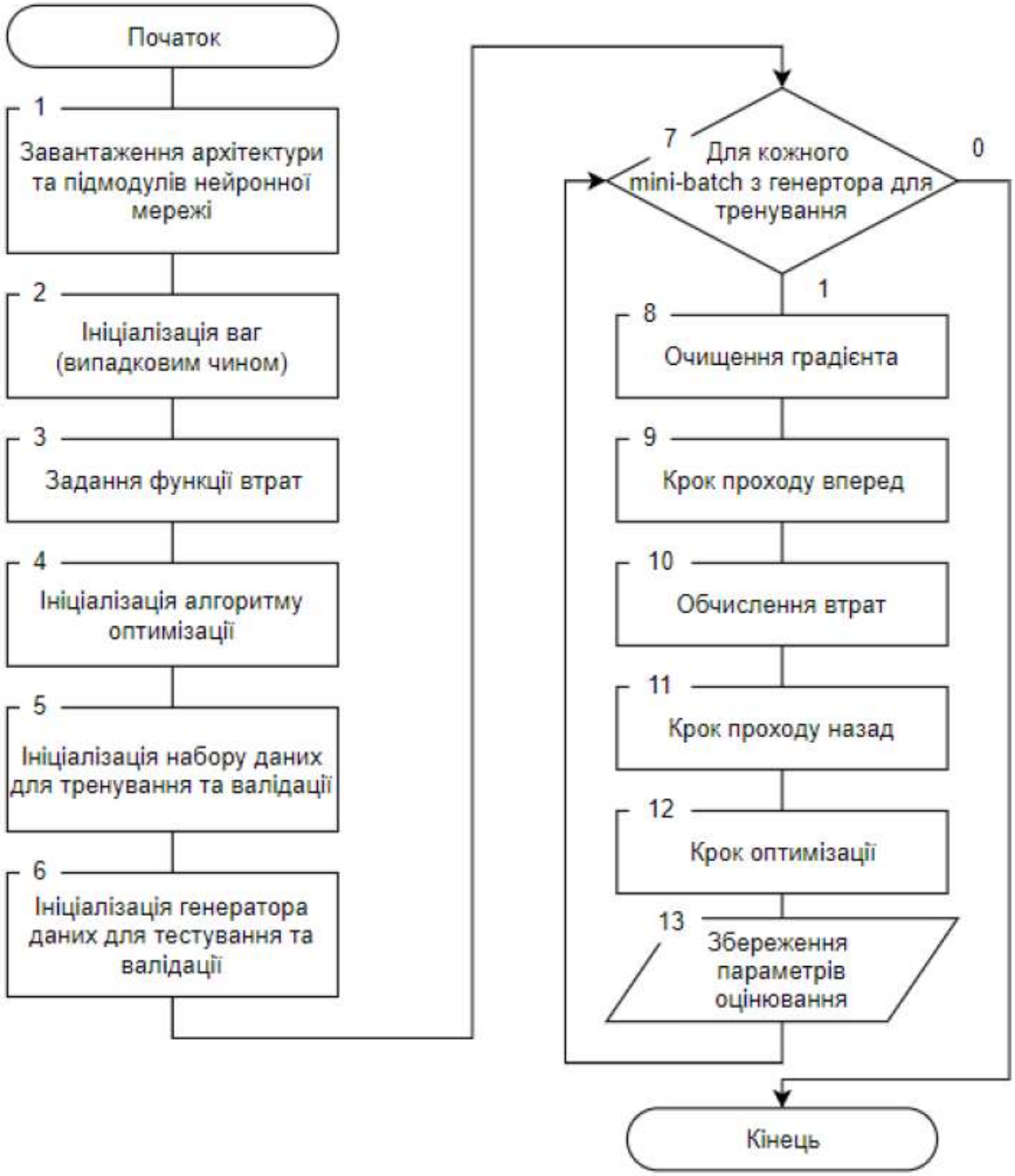


Рисунок 3.1 – Блок-схема алгоритму тренування та валідації нейронної мережі відновлення кольору цифрових фотографій

Алгоритм тестування нейронної мережі, наочно проілюстрований на рисунку 3.2, формалізується у вигляді такої послідовності етапів:

1. Завантаження оптимізованих параметрів моделі з відповідного файлу;
2. Ініціалізація базової архітектури та структурних підмодулів штучної нейронної мережі;
3. Інтеграція завантажених параметрів у структуру розгорнутої моделі;
4. Конфігурування цільової функції втрат;

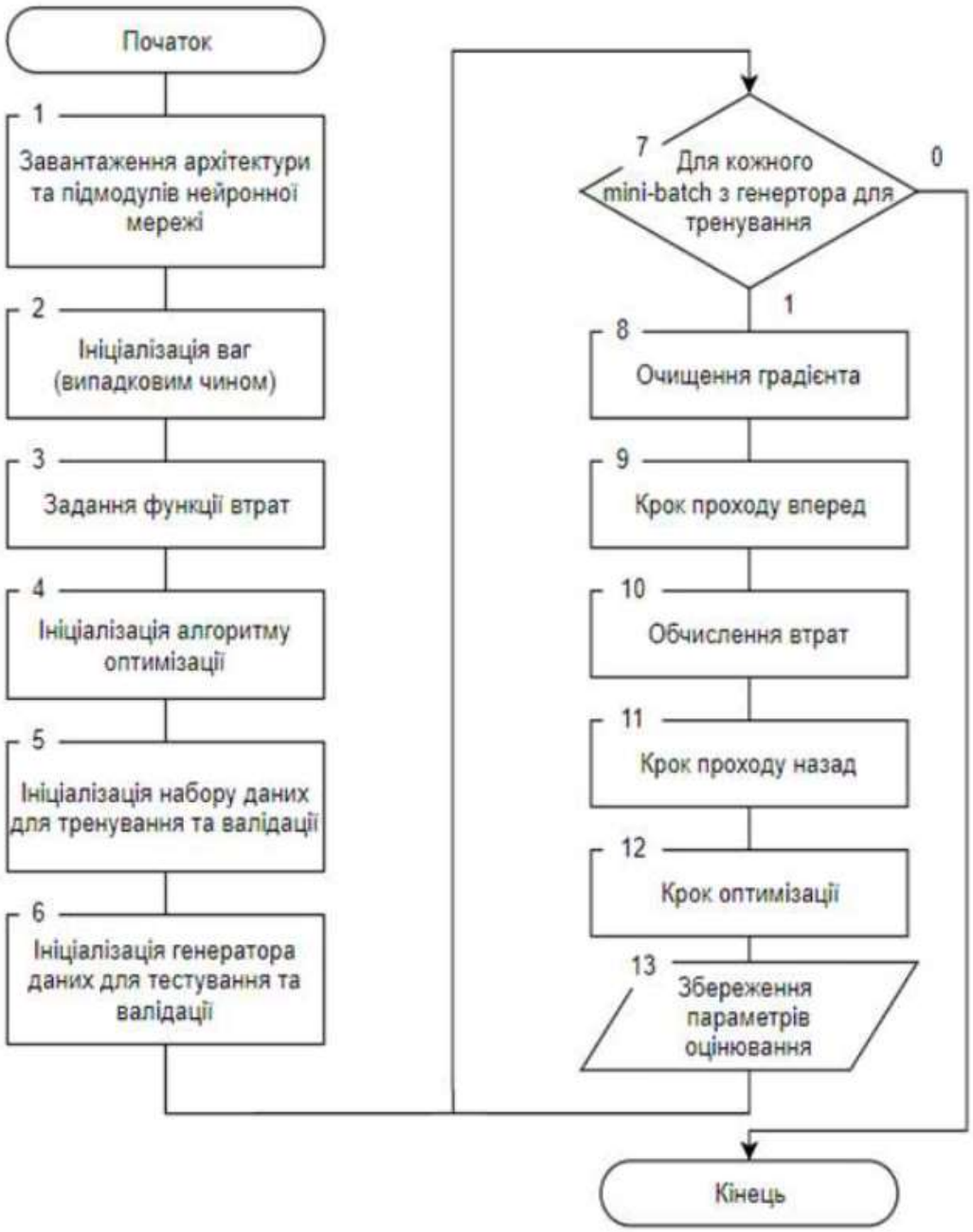


Рисунок 3.2 – Загальна блок-схема алгоритму тестування нейронної мережі відновлення кольору цифрових фотографій

5. Формування масиву даних для проведення процедури тестування;
6. Ініціалізація модуля завантаження даних для ітеративного опрацювання тестової вибірки;
7. Реалізація епохи тестування, що для кожного обчислювального міні-пакета з тестового завантажувача даних передбачає виконання таких процедур:
    - здійснення прямого поширення сигналу;
    - обчислення кількісних показників функції втрат;
    - збереження поточного стану моделі.

Опис програмної реалізації алгоритму тренування нейронної мережі для відновлення кольору цифрових фотографій

Наведені блок-схеми ілюструють деталізований алгоритм процесу тренування штучної нейронної мережі. Структурно алгоритм відображає класичний підхід до навчання моделей глибокого машинного навчання із застосуванням градієнтних методів оптимізації та розбиттям навчальної вибірки на міні-пакети.

Логіка виконання алгоритму декомпозується на наступні обчислювальні етапи:

Блоки 1-6 відповідають за етап конфігурування та ініціалізації програмного середовища. Відбувається завантаження топології нейромережевої моделі, ініціалізація синаптичних ваг випадковими значеннями для порушення симетрії, визначення цільової функції втрат та вибір алгоритму оптимізації. Також на цьому етапі формуються масиви даних та ініціалізуються ітератори для пакетного завантаження інформації.

Блок 7 репрезентує головний цикл навчання, який ітерується по всіх міні-пакетах, згенерованих з навчальної вибірки. Умова виходу з циклу спрацьовує, коли генератор вичерпує всі доступні пакети даних в межах поточної епохи.

Блоки 8-12 формують ядро математичної оптимізації моделі. Для кожного окремого міні-пакета виконується обнулення накопичених градієнтів з попереднього кроку, здійснюється пряме поширення сигналу через шари

мережі для отримання прогнозу, обчислюється значення функції втрат шляхом порівняння прогнозу з еталоном. Далі алгоритм зворотного поширення помилки розраховує градієнти для кожного параметра, після чого оптимізатор виконує оновлення вагових коефіцієнтів у напрямку антиградієнта.

Блок 13 забезпечує збереження оптимізованих параметрів моделі на постійний носій для можливості подальшого тестування, валідації або відновлення процесу тренування.

Нижче наведено програмну імплементацію описаного алгоритму мовою програмування Python із використанням сучасних фреймворків глибокого навчання, що структурно відповідає поданій блок-схемі:

```
import torch
import torch.nn as nn
import torch.optim as optim
from torch.utils.data import DataLoader

# Блок 1: Завантаження архітектури та підмодулів нейронної мережі
model = UNetColorizationModel()

# Блок 2: Ініціалізація ваг випадковим чином
def init_weights(m):
    if isinstance(m, nn.Conv2d) or isinstance(m, nn.ConvTranspose2d):
        nn.init.xavier_uniform_(m.weight)
        if m.bias is not None:
            nn.init.zeros_(m.bias)

model.apply(init_weights)

# Блок 3: Задання функції втрат
criterion = nn.MSELoss()

# Блок 4: Ініціалізація алгоритму оптимізації
optimizer = optim.Adam(model.parameters(), lr=0.0002)

# Блок 5: Ініціалізація набору даних для тренування та валідації
train_dataset = ColorizationDataset(mode='train')
val_dataset = ColorizationDataset(mode='validation')

# Блок 6: Ініціалізація генератора даних для тестування та валідації
train_loader = DataLoader(train_dataset, batch_size=32, shuffle=True)
val_loader = DataLoader(val_dataset, batch_size=32, shuffle=False)

num_epochs = 100

for epoch in range(num_epochs):
```

```
    model.train()

    # Блок 7: Для кожного mini-batch з генератора для тренування
    for batch_idx, (inputs, targets) in enumerate(train_loader):

        # Блок 8: Очищення градієнта
        optimizer.zero_grad()

        # Блок 9: Крок проходу вперед
        outputs = model(inputs)

        # Блок 10: Обчислення втрат
        loss = criterion(outputs, targets)

        # Блок 11: Крок проходу назад
        loss.backward()

        # Блок 12: Крок оптимізації
        optimizer.step()

    # Блок 13: Збереження параметрів оцінювання
    torch.save({
        'epoch': epoch,
        'model_state_dict': model.state_dict(),
        'optimizer_state_dict': optimizer.state_dict(),
        'loss': loss.item(),
    }, f'colorization_model_epoch_{epoch}.pth')
```

3.2. Побудова UML-діаграми класів програмного модуля відновлення кольору цифрових фотографій

Для архітектурної та програмної реалізації модуля відновлення кольору цифрових фотографій було спроектовано об'єктно-орієнтовану логічну структуру, що базується на взаємодії базових спеціалізованих класів, а саме: Main, ColorizeImage, LabGamut та GUIDesign. Графічну репрезентацію зазначеної ієрархії та зв'язків між об'єктами відображено на рисунку 3.3.

Клас Main виконує функцію точки входу та первинної ініціалізації програмного комплексу. Його ключовим завданням є забезпечення стартової активності системи з подальшим делегуванням управління класу GUIDesign.

Клас GUIDesign відповідає за управління підсистемою візуалізації. Цей компонент забезпечує рендеринг графічних елементів головного меню застосунку та комплексно реалізує механізми інтерактивної взаємодії з користувачем.

Клас ColorizeImage інкапсулює фундаментальне алгоритмічне забезпечення системи. У межах даного об'єкта зосереджена основна обчислювальна логіка безпосереднього процесу відновлення кольору цифрових фотографій.

Клас LabGamut виконує допоміжну аналітичну функцію в загальній архітектурі застосунку. Зазначений модуль відповідає за алгоритмічний розрахунок та оптимізацію хроматичних параметрів під час вибору колірної гами для генерування рекомендованих користувацьких підказок.

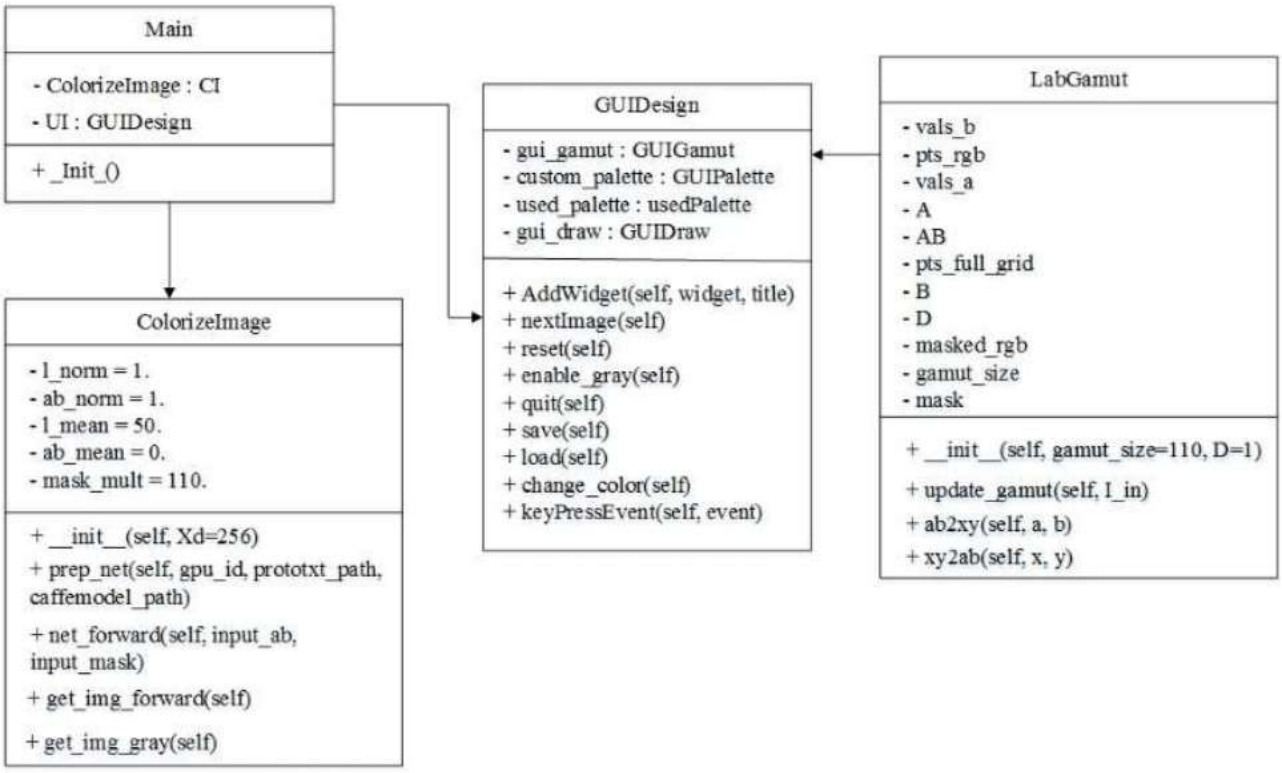


Рисунок 3.3 – UML-діаграма класів програмного модуля відновлення кольору цифрових фотографій

На наведеному рисунку зображено структурну UML-діаграму класів програмного модуля відновлення кольору цифрових фотографій. Архітектура системи базується на об'єктно-орієнтованій взаємодії чотирьох ключових класів: Main, ColorizeImage, GUIDesign та LabGamut.

Клас Main виконує функцію головного керуючого модуля ініціалізації. Він містить закриті атрибути для зберігання посилань на об'єкти інтерфейсу та алгоритмічної обробки зображень, а саме ColorizeImage типу CI та UI типу GUIDesign. Публічний інтерфейс класу представлено єдиним методом ініціалізації Init. Графічні зв'язки у вигляді спрямованих асоціацій вказують на те, що об'єкт

Main ініціює та управляє життєвим циклом екземплярів класів ColorizeImage та GUIDesign.

Клас ColorizeImage інкапсулює математичну та алгоритмічну частину взаємодії зі згортковою нейронною мережею. Його внутрішній стан визначається закритими константними атрибутами нормалізації та базових статистичних характеристик колірних каналів, серед яких l_norm, ab_norm, l_mean, ab_mean та mask_mult. Поведінка класу реалізується через набір публічних методів: конструктор ініціалізації init, метод конфігурування нейромережевої моделі та графічного процесора prep_net, метод обчислення прямого поширення сигналу net_forward, а також методи екстракції результуючого відновленого та вихідного монохромного зображень get_img_forward і get_img_gray відповідно.

Клас GUIDesign проектує логіку функціонування графічного інтерфейсу користувача. У його структурі інкапсульовані закриті атрибути, що репрезентують допоміжні візуальні компоненти палітри та рендерингу: gui_gamut, custom_palette, used_palette та gui_draw. Публічні методи даного класу забезпечують комплексну обробку користувацьких подій та управління візуальним станом. До їх числа входять методи динамічного додавання віджетів AddWidget, навігації по зображеннях nextImage, скидання параметрів обробки reset, управління монохромним режимом enable_gray, безпечного завершення роботи quit, дискових операцій вводу-виводу save та load, інтерактивної модифікації колірних маркерів change_color, а також низькорівневої обробки апаратних переривань keyPressEvent.

Клас LabGamut формує обчислювальний базис для метричних операцій із колірним простором CIELAB та оптимізації генерації інтерактивних підказок. Він акумулює масив закритих атрибутів для зберігання проміжних розрахунків, колірних сіток, матричних подань та бінарних масок, зокрема vals_b, pts_rgb, vals_a, A, AB, pts_full_grid, B, D, masked_rgb, gamut_size та mask. Функціонал класу абстраговано через публічний конструктор init, метод динамічного оновлення гіперспектрального простору update_gamut та методи афінних трансформацій двовимірних координат у просторі хроматичних каналів ab2xy і xy2ab. Спрямований зв'язок на діаграмі демонструє інтеграцію результатів обчислень класу LabGamut до структури керуючого класу GUIDesign.

3.3. Тестування програмного модуля відновлення кольору цифрових фотографій та аналіз результатів його функціонування

З метою верифікації працездатності та підтвердження коректності закладених алгоритмів було здійснено комплексне тестування розробленого програмного модуля відновлення кольору цифрових фотографій.

У межах етапу емпіричного дослідження було виконано сто контрольних ітерацій запуску програмного забезпечення. Під час цих сеансів здійснено перевірку всього спектра функціональних можливостей модуля, що забезпечило формування репрезентативної вибірки даних для подальшого об'єктивного та комплексного оцінювання отриманих результатів обробки візуальної інформації.

Безпосередня ініціалізація програмного комплексу супроводжується завантаженням та відображенням головного вікна графічного інтерфейсу користувача, візуальне подання якого проілюстровано на рисунку 3.4

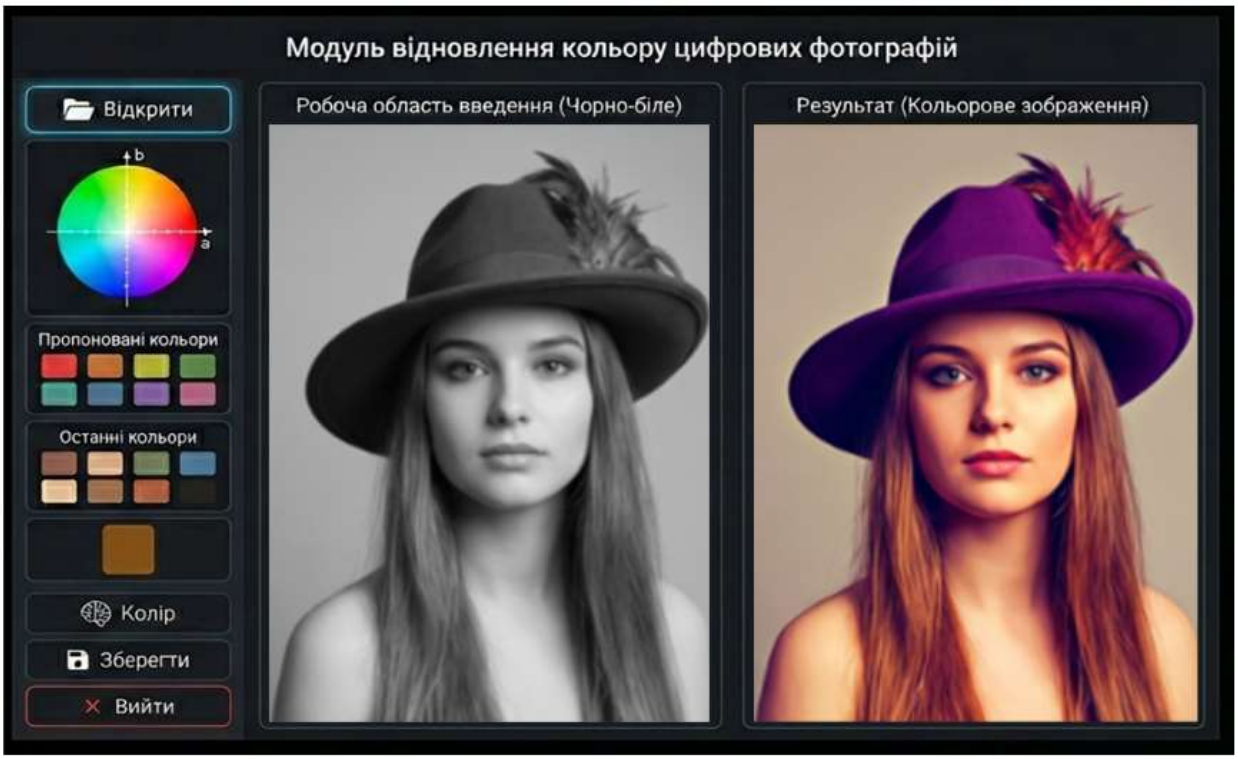


Рисунок 3.4 – Інтерфейс програмного модуля розфарбовування цифрових фотографій без використання користувацьких підказок

Головне вікно розробленого модуля відновлення кольору цифрових фотографій структурно складається з наступних елементів інтерактивної взаємодії з користувачем:

- Елемент управління «Відкрити», призначений для ініціалізації завантаження вихідного монохромного зображення, що підлягає обробці;
- Робоча область введення, яка забезпечує візуалізацію завантаженого монохромного зображення та слугує інтерактивним середовищем для просторового розміщення користувацьких хроматичних маркерів;
- Панель проекції хроматичної площини ab, що репрезентує допустиму колірну гаму каналів a та b простору CIELAB для вибраної просторової координати. Зазначений діапазон обчислюється на основі локального значення ахроматичного каналу L та надає можливість точного вибору спектральних характеристик для опорної точки;
- Блок рекомендованих відтінків «Пропоновані кольори», що генерує дискретну множину найбільш імовірних колірних рішень, синтезованих нейромережевою моделлю для конкретної локації, з можливістю їх подальшого використання як підказок;
- Панель історії «Останні кольори», що фіксує хронологію застосованих відтінків для забезпечення швидкого доступу та повторного використання;
- Індикатор поточного стану «Колірна панель», що здійснює безперервну візуалізацію активного вибраного відтінку;
- Елемент ініціалізації «Колір», активація якого запускає алгоритмічний процес відновлення кольору;
- Область виведення «Результат», призначена для відображення фінального або проміжного синтезованого зображення з відновленими хроматичними характеристиками. Цей компонент характеризується динамічним оновленням після кожної ітерації алгоритму;
- Елемент управління «Зберегти», що ініціює процедуру експорту обробленого візуального масиву на локальний носій інформації;

- Керуючий елемент «Вийти», який забезпечує коректне завершення сеансу роботи програмного комплексу.

Емпіричним шляхом встановлено, що результати автономного алгоритмічного відновлення кольору суттєво відрізняються від візуальних даних, згенерованих із застосуванням навіть одиничного інтерактивного маркера користувача.

У таблиці 3.1 систематизовано результати порівняльного аналізу ефективності функціонування спроектованого програмного модуля та існуючих альтернативних програмних рішень, орієнтованих на задачу відновлення кольору монохромних візуальних масивів. На рисунку 3.5 наочно продемонстровано зіставлення результатів обробки зображень розробленим алгоритмічним комплексом та його функціональним аналогом.



Рисунок 3.5 – Приклад відновлення кольору за допомогою розробленого програмного модуля та програми-аналога Akvis Coloriage

За результатами виконання серії тестових ітерацій функціонування програмного комплексу було сформульовано такі висновки: розроблена система здатна здійснювати відновлення кольору як в цілком автономному режимі, так і за умови інтеграції інтерактивних користувацьких маркерів. Емпірично встановлено, що візуальні матеріали, згенеровані із застосуванням просторових хроматичних підказок, демонструють вищий рівень суб'єктивної перцептивної якості порівняно з результатами роботи виключно автоматичного алгоритму.

Методика кількісного оцінювання якості базувалася на такому алгоритмі: оригінальне повноколірне зображення трансформувалося в монохромне подання, після чого піддавалося процедурі відновлення кольору за допомогою розробленого програмного модуля та комерційного аналога Akvis Coloriage. На наступному етапі здійснювався розрахунок метрики структурної відмінності DSSIM для об'єктивної оцінки ступеня подібності двох візуальних масивів [4]. Значення зазначеного показника нормуються в математичному діапазоні від мінус одиниці до одиниці. Обчислення даного критерію виконувалося для таких парних вибірок: еталонне повноколірне зображення та результат обробки розробленим модулем; еталонне зображення та результат роботи алгоритмів Akvis Coloriage.

Таблиця 3.1 – Порівняльний аналіз якості відновлення кольору цифрової фотографії

Програмне забезпечення	DSSIM
Розроблений модуль без уточнень	0,094
Розроблений модуль з 5 уточненнями	0,102
Розроблений модуль з 10 уточненнями	0,106
AKVIS Coloriage	0,087

Спроектований програмний модуль демонструє вищі показники якості відновлення кольору порівняно з існуючим програмним аналогом. Емпірично встановлено, що загальна ефективність відносно базового прототипу покращилася майже на 2,8%.

Таким чином, у межах дослідження синтезовано та формалізовано алгоритми тренування, валідації та тестування штучної нейронної мережі, призначеної для відновлення кольору цифрових фотографій. Спроектовано логічну UML-діаграму класів відповідного програмного забезпечення та здійснено його практичну алгоритмічну імплементацію.

За результатами проведеного експериментального тестування повністю підтверджено функціональну та обчислювальну коректність роботи розробленого модуля. Додатково виконано комплексний порівняльний аналіз створеної системи з наявним функціональним аналогом із застосуванням об'єктивного критерію оцінки якості відновлення кольору. Практично доведено, що метрика якості відновлення візуальних даних порівняно з обраним прототипом зросла майже на 2,8%.

ВИСНОВКИ

У межах виконання кваліфікаційної роботи було спроектовано та реалізовано спеціалізований програмний модуль для відновлення кольору цифрових фотографій.

Здійснено комплексний аналіз існуючих технологій відновлення кольору візуальних даних, що дозволило науково обґрунтувати актуальність проведеного дослідження. Виокремлено фундаментальні недоліки традиційних алгоритмів обробки монохромних фотографій, ключовим з яких є відсутність дієвих механізмів інтерактивного впливу користувача на кінцевий результат відновлення кольору. За результатами критичного огляду існуючих програмних аналогів встановлено відсутність комплексних рішень, здатних забезпечити одночасне досягнення високої швидкодії та належної якості відновлення кольору монохромних знімків. На основі виявленої проблематики формалізовано загальну постановку задачі дослідження.

Науково обґрунтовано вибір оптимальної колірної моделі для алгоритмічної обробки зображень. Доведено математичну та обчислювальну доцільність застосування апарату згорткових нейронних мереж як найбільш ефективного інструментарію для вирішення проблеми відновлення кольору. Теоретично обґрунтовано вибір оптимального методу навчання нейромережевої моделі. Синтезовано спеціалізовану архітектуру штучної нейронної мережі для відновлення кольору цифрових фотографій, структурним базисом якої виступає топологія U-net. Спроектовано комплексну блок-схему алгоритму функціонування відповідного програмного модуля відновлення кольору цифрових фотографій.

Наведено обґрунтування вибору інструментальних засобів розробки, включаючи мову програмування, інтегроване середовище та спеціалізовані нейромережеві фреймворки. Практичну реалізацію системи виконано з використанням мови програмування Python та бібліотеки Caffe, які визначені як високоефективні програмні інструменти для роботи з архітектурами згорткових нейронних мереж.

Розроблено та формалізовано алгоритми тренування, валідації та тестування нейронної мережі для відновлення кольору цифрових фотографій. Побудовано структурну UML-діаграму класів розроблюваного програмного модуля відновлення кольору візуальних масивів та успішно здійснено його практичну програмну імплементацію.

За результатами проведеного експериментального тестування створеного програмного забезпечення повністю підтверджено обчислювальну та алгоритмічну коректність його функціонування. Здійснено комплексний порівняльний аналіз розробленого модуля із наявним програмним аналогом на основі об'єктивних критеріїв оцінки якості відновлення кольору. Емпірично доведено, що показник якості відновлення кольору запропонованою системою перевищує результати базового прототипу на 2,8%.

СПИСОК ЛІТЕРАТУРИ

1. Goodfellow I., Bengio Y., Courville A. Deep Learning. MIT Press, 2016. 800 p.
2. Chollet F. Deep Learning with Python. 2nd Edition. Manning Publications, 2021. 504 p.
3. Bishop C. M. Pattern Recognition and Machine Learning. Springer, 2006. 738 p.
4. Шаховська Н. Б., Камінський Р. М. Системи штучного інтелекту: навчальний посібник. Львів: Видавництво Львівської політехніки, 2021. 392 с.
5. Ткаченко Р. О., Ізонін І. В. Нейромережеві методи обробки даних. Львів: Львівська політехніка, 2017. 256 с.
6. Бодянський Є. В., Руденко О. Г. Штучні нейронні мережі: архітектури, навчання, застосування. Харків: ТЕЛЕТЕХ, 2004. 369 с.
7. Любчик Л. М., Грінберг Г. Л. Штучні нейронні мережі: архітектура, алгоритми навчання, застосування. Харків: НТУ «ХПІ», 2018. 145 с.
8. Снітюк В. Є. Еволюційні технології прийняття рішень. Київ: Макмак, 2010. 300 с.
9. Грабовський В. В., Палагін О. В. Основи глибокого навчання: навчальний посібник. Київ: КПІ ім. Ігоря Сікорського, 2020. 210 с.
10. Gonzalez R. C., Woods R. E. Digital Image Processing. 4th Edition. Pearson, 2017. 1192 p.
11. Szeliski R. Computer Vision: Algorithms and Applications. 2nd Edition. Springer, 2022. 931 p.
12. Bradski G., Kaehler A. Learning OpenCV 3: Computer Vision in C++ with the OpenCV Library. O'Reilly Media, 2016. 1018 p.
13. Рашкевич Ю. М., Пелешко Д. Д., Ковальчук А. М. Алгоритми та методи обробки відеоданих: монографія. Львів: Видавництво Львівської політехніки, 2012. 180 с.

14. Гороховатський В. О. Системи штучного інтелекту та розпізнавання образів. Харків: Компанія СМІТ, 2020. 272 с.
15. Бабак В. П., Тарасенко В. П. Обробка сигналів та зображень: навчальний посібник. Київ: НАУ, 2004. 156 с.
16. Яворський І. М., Юзефович Р. М. Аналіз сигналів і зображень у системах комп'ютерного зору. Львів: ФМІ НАН України, 2015. 215 с.
17. Антонов В. М. Машинне бачення та обробка зображень: конспект лекцій. Одеса: ОНПУ, 2020. 120 с.
18. Zhang R., Isola P., Efros A. A. Colorful Image Colorization. European Conference on Computer Vision (ECCV). 2016. P. 649–666.
19. Iizuka S., Simo-Serra E., Ishikawa H. Let there be color! Joint end-to-end learning of global and local image priors for automatic image colorization with simultaneous classification. ACM Transactions on Graphics. 2016. Vol. 35, No. 4. P. 1–11.
20. Larsson G., Maire M., Shakhnarovich G. Learning representations for automatic colorization. European Conference on Computer Vision (ECCV). 2016. P. 577–593.
21. Ronneberger O., Fischer P., Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. Medical Image Computing and Computer-Assisted Intervention (MICCAI). 2015. P. 234–241.
22. He K., Zhang X., Ren S., Sun J. Deep Residual Learning for Image Recognition. Proceedings of the IEEE conference on computer vision and pattern recognition. 2016. P. 770–778.
23. LeCun Y., Bengio Y., Hinton G. Deep learning. Nature. 2015. Vol. 521, No. 7553. P. 436–444.
24. Krizhevsky A., Sutskever I., Hinton G. E. ImageNet classification with deep convolutional neural networks. Communications of the ACM. 2017. Vol. 60, No. 6. P. 84–90.

25. Su J., Chu H., Huang J. B. Instance-aware Image Colorization. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2020. P. 2720–2729.
26. Guadarrama S. et al. PixColor: Pixel Recursive Colorization. British Machine Vision Conference (BMVC). 2017.
27. Sharma G., Wu W., Dalal E. N. The CIEDE2000 color-difference formula: Implementation notes, supplementary test data, and mathematical observations. Color Research & Application. 2005. Vol. 30, No. 1. P. 21–30.
28. Poynton C. Digital Video and HD: Algorithms and Interfaces. Morgan Kaufmann, 2012. 732 p.
29. Fairchild M. D. Color Appearance Models. 3rd Edition. John Wiley & Sons, 2013. 472 p.
30. Снитюк В. Є. Моделювання просторових характеристик колірних систем. Київ: КНУ ім. Т. Шевченка, 2014. 135 с.
31. McKinney W. Python for Data Analysis. 2nd Edition. O'Reilly Media, 2017. 544 p.
32. Lutz M. Learning Python. 5th Edition. O'Reilly Media, 2013. 1600 p.
33. VanderPlas J. Python Data Science Handbook. O'Reilly Media, 2016. 548 p.
34. Rosebrock A. Deep Learning for Computer Vision with Python. PyImageSearch, 2017.
35. Марценюк В. П., Сверстюк А. С. Основи машинного навчання на базі Python: навчальний посібник. Тернопіль: ТДМУ, 2018. 176 с.
36. Кондратюк В. А. Машинне навчання та розпізнавання образів мовою Python. Тернопіль: ТНТУ, 2019. 210 с.
37. Jia Y., Shelhamer E., Donahue J., Karayev S. Caffe: Convolutional Architecture for Fast Feature Embedding. Proceedings of the 22nd ACM international conference on Multimedia. 2014. P. 675–678.
38. Paszke A. et al. PyTorch: An Imperative Style, High-Performance Deep Learning Library. Advances in Neural Information Processing Systems. 2019. P. 8024–8035.

39. Офіційна документація нейромережевої бібліотеки Caffe [Електронний ресурс]. URL: http://caffe.berkeleyvision.org/.
40. Офіційна документація бібліотеки комп'ютерного зору OpenCV [Електронний ресурс]. URL: https://docs.opencv.org/

ДОДАТКИ