

**КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМЕНІ ТАРАСА
ШЕВЧЕНКА**

**ФАКУЛЬТЕТ РАДІОФІЗИКИ, ЕЛЕКТРОНІКИ ТА КОМП'ЮТЕРНИХ
СИСТЕМ**

Кафедра медичної радіофізики

«На правах рукопису»

Робота допущена до захисту в ЕК
рішенням кафедри медичної радіофізики
від____ 2026 року, протокол №____ .
Завідувач кафедри канд. фіз.-мат. наук, доцент
_____ Сергій РАДЧЕНКО

КВАЛІФІКАЦІЙНА РОБОТА МАГІСТРА

на тему:

**«АНАЛІЗ ОБЛАСТЕЙ ДІАГНОСТИЧНИХ МЕДИЧНИХ
ТОМОГРАФІЧНИХ ДАНИХ ПРИ ОПТИМІЗАЦІЇ АРХІТЕКТУРИ
НЕЙРОННИХ МЕРЕЖ ТА МОДИФІКАЦІЇ ПРОСТОРУ ПАРАМЕТРІВ»**

Виконав:

студент 2-го курсу магістратури
денної форми здобуття освіти
спеціальності 105 Прикладна фізика та наноматеріали
ОНП Біомедична фізика, інженерія та інформатика
Морозов Володимир Миколайович _____

Науковий керівник:

канд. фіз.-мат. наук, доцент
Нетребя Андрій В'ячеславович _____

Рецензент:

докт. техн. наук, ст. дослідник
Яцишин Андрій Васильович _____

Засвідчую, що у цій дипломній роботі
немає запозичень з праць інших авторів без
відповідних посилань
Студент _____ Володимир МОРОЗОВ

Київ – 2026

КТ — Комп'ютерна томографія.

MPT — Магнітно-резонансна томографія.

CLAHE — Адаптивне вирівнювання гістограми з обмеженням контрасту (Contrast Limited Adaptive Histogram Equalization).

CNN — Згортова нейронна мережа (Convolutional Neural Network).

DSC — Коефіцієнт подібності Дайса (Dice Similarity Coefficient).

ELU — Експоненціальна лінійна одиниця (Exponential Linear Unit).

FL — Фокальна втрата (Focal Loss).

FN — Хибнонегативні результати (False Negatives).

FP — Хибнопозитивні результати (False Positives).

IoU — Перетин над об'єднанням (Intersection over Union).

LReLU — Leaky Rectified Linear Unit.

Micro-IoU / mIoU — Мікроусереднений перетин над об'єднанням (Micro-Averaged IoU).

PReLU — Параметрична випрямлена лінійна одиниця (Parametric Rectified Linear Unit).

ReLU — Випрямлена лінійна одиниця (Rectified Linear Unit).

ResNet — Залишкова нейронна мережа (Residual Network).

SiLU — Лінійна одиниця, зважена на сигмоїду (Sigmoid-weighted Linear Unit).

SMP — Segmentation Models PyTorch.

TP — Істиннопозитивні результати (True Positives).

Реферат

Дипломна робота магістра – 67 с., 34 зображення, 31 джерело.

У роботі розв'язано задачу автоматичної багатокласової сегментації печінки та її функціональних сегментів за класифікацією Куїно на КТ-зображеннях (набір даних 3DIRCADb-01). Для подолання проблем низького контрасту тканин та екстремального просторового дисбалансу класів застосовано архітектуру U-Net із попередньо навченим енкодером сімейства ResNet та комбіновану функцію втрат (Dice Loss + Focal Loss).

Проведено комплексний аналіз впливу різних функцій активації (ReLU, LeakyReLU, ELU, PReLU, Swish, SiLU, Mish) на збіжність глибоких мереж. Доведено, що класична функція ReLU забезпечує найкращий баланс між обчислювальною швидкістю та точністю сегментації, тоді як використання складних адаптивних нелінійностей не дає суттєвого приросту метрик завдяки стабілізуючій дії шарів нормалізації.

Досліджено ефективність різних стратегій аугментації даних. Встановлено, що стратегія помірної інтенсивності, яка поєднує еластичні деформації, афінні трансформації та фотометричні перетворення (зокрема адаптивне вирівнювання гістограми CLANE), діє як найефективніший стохастичний регуляризатор. Це дозволило нівелювати перенавчання, оптимально утилізувати параметричну ємність мережі та досягти високої стабільності результатів сегментації (Binary Liver IoU > 0.90).

Ключові слова: сегментація медичних зображень, комп'ютерна томографія, печінка, сегменти Куїно, глибоке навчання, згорткові нейронні мережі, U-Net, ResNet, функції активації, аугментація даних, еластичні деформації, дисбаланс класів.

ЗМІСТ

ВСТУП	5
РОЗДІЛ 1. АНАЛІТИЧНИЙ ОГЛЯД МЕТОДІВ СЕГМЕНТАЦІЇ МЕДИЧНИХ ЗОБРАЖЕНЬ	7
1.1 Клінічні показання, технічні підходи та перспективи розвитку сегментації печінки	7
1.2 Аналіз предметної області та набору даних IRCAD Liver Tumours	10
1.3 Постановка задачі дослідження	13
РОЗДІЛ 2. МЕТОДОЛОГІЯ ТА АРХІТЕКТУРИ МОДЕЛЕЙ ГЛИБОКОГО НАВЧАННЯ	16
2.1 U-Net	16
2.2 Сімейство архітектур ResNet	17
2.3 Функції втрат для компенсації дисбалансу класів	19
2.4 Метрики оцінювання ефективності	23
РОЗДІЛ 3. Аналіз впливу функцій активації на збіжність глибоких мереж	27
3.1 Функції активації	27
3.2 Функції із навченим параметром	30
3.3 Функції із сталим параметром	33
3.4 Функції із гіперпараметром	36
РОЗДІЛ 4. АЛГОРИТМІЧНЕ ЗАБЕЗПЕЧЕННЯ ТА СТРАТЕГІЇ АУГМЕНТАЦІЇ ДАНИХ	40
4.1 Геометричні та просторові перетворення	40
4.2 Фотометричні трансформації	44
4.3 Моделювання нелінійних деформацій та артефактів	46
РОЗДІЛ 5. ЧИСЛОВІ ЕКСПЕРИМЕНТИ ТА АНАЛІЗ РЕЗУЛЬТАТІВ	50
5.1 Дослідження впливу функцій активації на навчання моделі	50
5.2 Оцінка ефективності стратегій аугментації	54
5.3 Аналіз процесу навчання поглибленої моделі	59
ВИСНОВКИ	63
Список літератури	64

ВСТУП

Сегментація органів, їхніх анатомічних частин та патологічних утворень на медичних зображеннях є однією з центральних задач сучасної медичної діагностики та клінічної практики. Вона лежить в основі планування хірургічних втручань, оцінки ефективності лікування та контролю за динамікою розвитку захворювань. Зокрема, автоматизована сегментація функціональних сегментів печінки за класифікацією Куїно на комп'ютерній томографії (КТ) дозволяє отримати точні просторові моделі органу. Це є критично важливою умовою для виконання складних хірургічних процедур, таких як резекція пухлин, двоетапна гепатектомія, емболізація ворітної вени та трансплантація печінки від живого донора. Усі ці методи є надзвичайно чутливими до точності локалізації структур, оскільки межі між сегментами не мають чітко виражених фізичних контурів і визначаються топологією порталльної венозної системи.

Традиційні підходи до сегментації, що базуються на ручному контуруванні або алгоритмах порогової обробки, мають низку суттєвих обмежень. Ручна сегментація є надзвичайно трудомісткою (процес може займати до 90 хвилин для одного пацієнта) і характеризується значною міжоператорською варіабельністю. Крім того, класичні алгоритми втрачають ефективність через низький контраст між патологічними структурами та здоровою паренхімою, анізотропність медичних даних (ефект часткового об'єму) та екстремальний дисбаланс класів, де пухлини або судини займають лише десятки частки відсотка від загального об'єму зображення.

У сучасній практиці найперспективнішим напрямком стали алгоритми глибокого навчання, зокрема згорткові нейронні мережі (CNN). Ієрархічні архітектури на кшталт U-Net із потужними енкодерами (такими як сімейство ResNet) демонструють високу точність локалізації об'єктів та здатність ефективно працювати з обмеженими медичними даними. Однак ефективність таких моделей значною мірою визначається стабільністю оптимізації та здатністю узагальнювати результати на нових клінічних даних, які характеризуються високою апаратною та анатомічною неоднорідністю (наприклад, стандартизований набір даних 3D-IRCADb-01).

Для розв'язання цих проблем необхідне комплексне алгоритмічне забезпечення. З одного боку, критично важливим є вибір оптимальних функцій активації (зокрема ReLU, LeakyReLU, PReLU, Swish, SiLU, Mish) для стабілізації градієнтного спуску та компенсації проблеми «мертвих нейронів» у глибоких шарах мережі. З іншого боку, необхідне використання специфічних функцій втрат (комбінація Dice Loss та Focal Loss) для подолання екстремального просторового дисбалансу. Крім цього, ефективним способом підвищити інваріантність та узагальнюючу здатність мереж є розробка стратегій аугментації даних. Поєднання геометричних перетворень, еластичних деформацій м'яких тканин та фотометричних змін (наприклад, адаптивного вирівнювання гістограми CLAHE) дозволяє синтетично

моделювати реальну фізіологічну мінливість органів та запобігати перенавчанню.

Метою даного дослідження є підвищення точності, стабільності та надійності автоматичної сегментації печінки та її внутрішніх анатомічних структур на КТ-зображеннях для подальшого використання у системах комп'ютерної підтримки клінічних рішень. Для досягнення цієї мети у роботі проведено комплексний аналіз: досліджено вплив різноманітних функцій активації на збіжність архітектури U-Net/ResNet34 , а також розроблено та кількісно оцінено чотири експериментальні стратегії аугментації (від базового контролю до агресивних нелінійних трансформацій). Це дозволяє всебічно оптимізувати процес навчання та забезпечити формування найбільш інваріантних ознак для високоточної багатокласової сегментації.

РОЗДІЛ 1. АНАЛІТИЧНИЙ ОГЛЯД МЕТОДІВ СЕГМЕНТАЦІЇ МЕДИЧНИХ ЗОБРАЖЕНЬ

1.1 Клінічні показання, технічні підходи та перспективи розвитку сегментації печінки

Сегментація медичних зображень є процесом виділення органу інтересу на багатоплощинних томографічних даних комп'ютерної томографії або магнітно-резонансної томографії з метою подальшого морфологічного, функціонального чи об'ємного аналізу. У випадку печінки цей процес є особливо складним через високу анатомічну варіабельність органу, складну геометрію його поверхні, наявність великої кількості судинних структур та близьке просторове розташування до сусідніх органів із подібними інтенсивнісними характеристиками. Додаткову складність створюють патологічні зміни, зокрема стеатоз, фіброз, цироз, пухлинні ураження та постопераційні деформації, які змінюють щільність тканин, неоднорідність сигналу та загальну архітектуру печінки. У результаті навіть незначні похибки сегментації можуть суттєво впливати на точність об'ємних оцінок та клінічних рішень [1].

Традиційні методи оцінки розмірів печінки, що базувалися на фізикальному обстеженні або одновимірних вимірюваннях, мають обмежену інформативність через значну природну варіабельність форми органу між різними пацієнтами. Печінки з подібною краніо-каудальною довжиною можуть суттєво відрізнятися за реальним просторовим об'ємом, тому використання лише лінійних параметрів не дозволяє коректно оцінювати функціональний резерв органу. Розвиток сучасних методів медичної візуалізації дав змогу перейти до повноцінного тривимірного вимірювання об'єму, яке забезпечує значно вищу точність аналізу. Найчастіше для цього використовується КТ завдяки високій просторовій роздільній здатності, ізотропним вокселям, короткому часу сканування та стійкості до рухових артефактів. МРТ, своєю чергою, забезпечує кращу тканинну контрастність, можливість багатопараметричної оцінки паренхіми та аналізу судинно-біліарної анатомії без використання іонізуючого випромінювання. Особливо важливу роль відіграють гепатоспецифічні контрастні препарати, які підвищують контраст між печінковою тканиною та навколишніми структурами, що позитивно впливає на точність сегментації.

Клінічна значущість сегментації печінки насамперед пов'язана з необхідністю оцінки майбутнього залишкового об'єму печінки перед великими хірургічними втручаннями. При плануванні гепатектомії визначення загального об'єму печінки та майбутнього залишкового об'єму є критично важливим, оскільки саме об'єм функціонально збереженої паренхіми визначає ризик післяопераційної печінкової недостатності. Встановлено, що безпечне співвідношення залишкового об'єму до загального

повинно перевищувати 20% у пацієнтів із нормальною печінкою, понад 30% у випадку помірного ураження паренхіми та понад 40% при цирозі або вираженому фіброзі. Таким чином, сегментація безпосередньо впливає на вибір хірургічної тактики та оцінку операбельності пацієнта.

Важливим напрямком застосування оцінки об'ємів є емболізація ворітної вени, яка виконується для стимуляції гіпертрофії майбутнього залишкового об'єму печінки перед резекцією. Блокування гілок ворітної вени, що кровопостачають сегменти, заплановані до видалення, призводить до перерозподілу кровотоку та компенсаторного росту функціонально збереженої паренхіми. У таких випадках сегментація виконується повторно після процедури для оцінки ступеня гіпертрофії та визначення готовності пацієнта до операції. Аналогічно, при двоетапній гепатектомії сегментація використовується на кожному етапі лікування для контролю динаміки росту залишкового об'єму та мінімізації ризику печінкової недостатності.

Ще одним критично важливим клінічним застосуванням є трансплантація печінки від живого донора. У цьому випадку сегментація дозволяє оцінити відповідність об'єму трансплантата потребам реципієнта та одночасно забезпечити достатній залишковий об'єм печінки у донора. Недостатній розмір трансплантата може призвести до розвитку синдрому недостатнього об'єму, який супроводжується тяжкою функціональною недостатністю трансплантованої печінки. Окрім загальної оцінки об'єму, дедалі більшого значення набуває судинна субсегментація, що дозволяє аналізувати індивідуальні варіанти порталльної та венозної анатомії перед оперативним втручанням.

З огляду на клінічну важливість вимірювання об'ємів, методи сегментації печінки активно еволюціонували від ручних підходів до повністю автоматизованих систем на основі глибокого навчання. Ручна сегментація тривалий час вважалася «золотим стандартом», оскільки забезпечувала максимальний контроль над межами органу, однак вона є надзвичайно трудомісткою та характеризується значною міжоператорською варіабельністю. Процес контурювання печінки на всіх зрізах томографії може займати до 90 хвилин для одного пацієнта, що практично унеможливило використання такого підходу у клінічній практиці з великим потоком досліджень. Для часткової автоматизації застосовувались допоміжні методи, зокрема активні контури та алгоритми інтерактивного трасування меж, які використовують локальні градієнти зображення для автоматичного уточнення контурів органу після грубої ініціалізації користувачем.

Подальший розвиток отримали напіваавтоматичні алгоритми, що виконують основну частину оптимізації після мінімального втручання оператора. Одним із найпростіших підходів стали методи нарощування областей, де користувач задає початкові точки всередині печінки, а алгоритм ітеративно додає сусідні вокселі зі схожими інтенсивностями. Такі методи добре працюють на однорідних тканинах, проте мають низьку стійкість до

гетерогенності паренхіми та часто демонструють «витік» сегментації у сусідні органи при слабко виражених межах. Для подолання цих недоліків почали використовуватись графові методи сегментації, зокрема алгоритми графових розрізів, де зображення представляється у вигляді графа, а задача сегментації формулюється як пошук оптимального розділення між областями переднього плану та фону. Такий підхід забезпечує більш стабільне виділення меж навіть у випадку складної анатомії.

Особливе місце серед автоматизованих методів займають статистичні моделі форми та деформівні тривимірні моделі. Статистичні моделі форми використовують попередньо побудовану базу допустимих геометричних варіацій печінки та накладають морфологічні обмеження на результат сегментації. Це дозволяє запобігати надмірному виходу сегментації за межі органу та забезпечує стабільність роботи навіть при наявності шуму або низькоконтрастних меж. Водночас такі моделі можуть бути надто жорсткими у випадку нетипових патологічних або постопераційних форм печінки, якщо подібні приклади недостатньо представлені у тренувальних даних. Деформівні тривимірні моделі, навпаки, дозволяють адаптивно змінювати поверхневу сітку відповідно до контурів органу із збереженням гладкості поверхні, що забезпечує кращу адаптацію до індивідуальної анатомії пацієнта.

Найбільш перспективним напрямком сучасної сегментації стали алгоритми глибокого навчання, зокрема згорткові нейронні мережі. На відміну від класичних методів машинного навчання, які використовували вручну визначені текстурні ознаки, згорткові нейронні мережі здатні автоматично вивчати багаторівневі просторові та семантичні представлення без необхідності ручного проєктування ознак. Це дозволяє ефективно сегментувати навіть гетерогенні печінки з різними типами патологій та даними, отриманими на різних сканерах і протоколах. Сучасні моделі можуть виконувати сегментацію за лічені секунди, що робить їх придатними для інтеграції у клінічні робочі процеси та системи комп'ютерної підтримки прийняття рішень.

Попри значний прогрес, сегментація печінки залишається складною задачею через численні джерела похибок. До основних проблем належать рухові артефакти, частковий об'ємний ефект, артефакти від металевих імплантів та низький контраст меж між печінкою і сусідніми структурами, такими як шлунок, серце, діафрагма або селезінка. Надмірна сегментація зазвичай виникає у перигепатичних областях із подібною щільністю тканин, тоді як недостатня сегментація спостерігається у зонах неоднорідної структури паренхіми, зокрема при стеатозі, цирозі чи пухлинних ураженнях. Додатково на якість сегментації впливають параметри сканування, товщина зрізів та тип контрастування, оскільки погіршення просторової роздільної здатності збільшує втрату локальної інформації на межах структур.

Сучасні тенденції розвитку сегментації печінки спрямовані не лише на підвищення точності виділення органу, але й на інтеграцію

багатокомпонентного анатомічного та функціонального аналізу. Перспективними напрямками є автоматична судинна субсегментація, побудова персоналізованих тривимірних моделей для віртуального планування операцій, інтеграція сегментації з біомаркерами магнітно-резонансної томографії та використання мультимодальних моделей, що поєднують КТ, МРТ і функціональні параметри. Окремий інтерес становить використання об'ємних біомаркерів, таких як індекс загального вмісту жиру в печінці, що поєднують результати сегментації з кількісною оцінкою патологічних змін паренхіми. Таким чином, сегментація печінки поступово трансформується з інструмента простого об'ємного аналізу у комплексну систему кількісної оцінки анатомічного та функціонального стану органу, яка стає невід'ємною складовою сучасної персоналізованої медицини.

1.2 Аналіз предметної області та набору даних IRCAD Liver Tumours

Сучасна хірургія печінки значною мірою базується на концепції функціональної анатомії, відповідно до якої орган поділяється не лише за морфологічними, але й за судинними особливостями. Найбільш поширеною міжнародною системою є класифікація Куіно (рис. 1.1-1.2), яка розділяє печінку на вісім функціонально незалежних сегментів. Основою такого поділу виступає портальна венозна система, оскільки саме гілки ворітної вени визначають межі функціональних зон кровопостачання. Кожен сегмент має власне судинне та жовчне забезпечення, що робить його відносно автономною анатомічною одиницею під час хірургічних втручань. На відміну від звичайних анатомічних структур, межі між сегментами не мають чітко виражених фізичних контурів та фактично є умовними площинами, сформованими конфігурацією судинної мережі. У результаті вони практично не візуалізуються безпосередньо на КТ або МРТ-зображеннях, що суттєво ускладнює задачу автоматичної сегментації та вимагає від алгоритмів здатності аналізувати не лише локальні текстурні ознаки, але й просторову судинну топологію[2].

Складність задачі додатково підсилюється високою анатомічною варіабельністю печінки між різними пацієнтами. Форма органу, просторове положення сегментів, конфігурація судинної системи та характер розгалуження портальних структур можуть суттєво відрізнятися навіть у межах фізіологічної норми. Патологічні процеси, зокрема цироз, стеатоз, фіброз або пухлинні ураження, ще більше спотворюють природну геометрію печінки та ускладнюють виділення меж між анатомічними структурами. У випадку великих пухлин або післяопераційних змін сегменти можуть зміщуватися, деформуватися або частково втрачати типову просторову

організацію, що значно знижує ефективність класичних алгоритмів, заснованих на жорстких апіорних моделях форми[3].

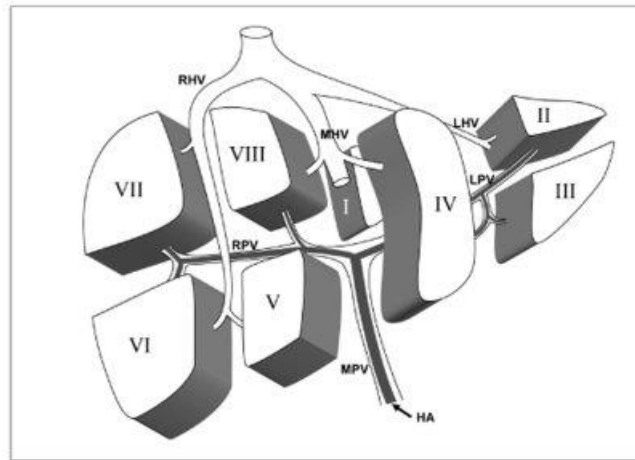
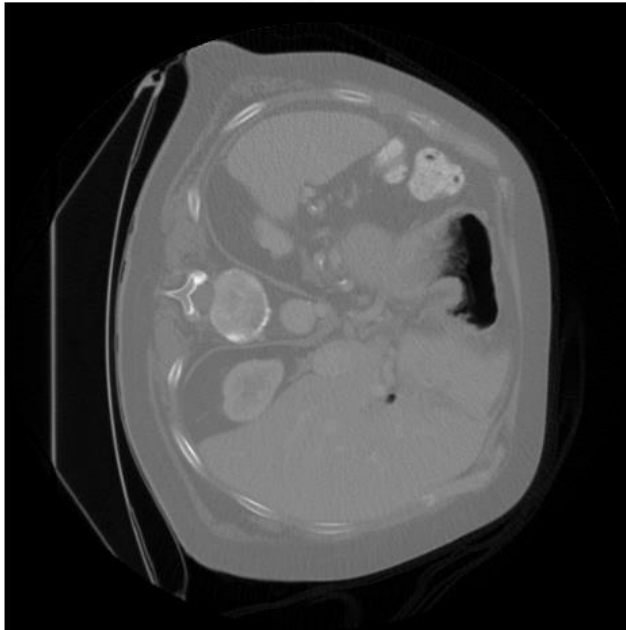


Рисунок 1.1 —

Схема печінкових сегментів (I-VIII) з їх ворітною венозною системою

а) Зображення



б) Маска

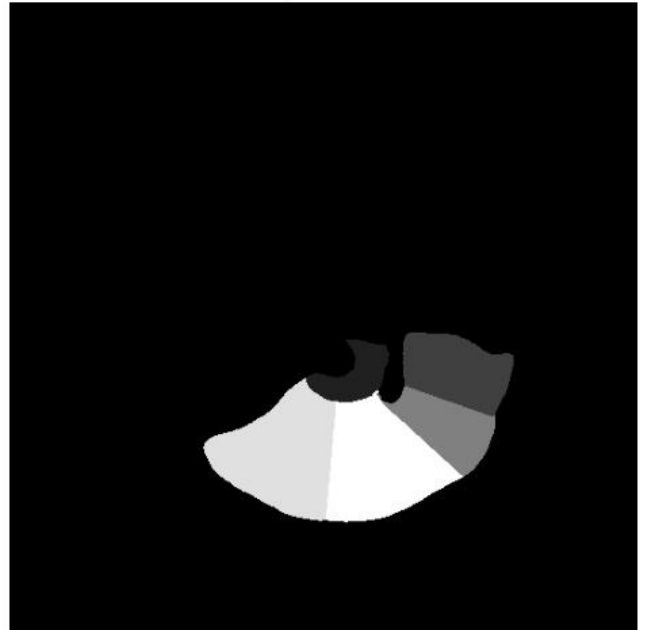


Рисунок 1.2 — Приклад КТ зрізу а) та маски б)

Для дослідження задач автоматичної сегментації печінки та валідації алгоритмів комп'ютерного зору широкого використання набув набір даних 3DIRCADb, також відомий як IRCAD Liver Tumours. Цей датасет став одним із базових стандартів у задачах сегментації печінки та пухлин завдяки високій якості анотацій, клінічній репрезентативності та значній складності даних. Набір містить двадцять тривимірних КТ-досліджень черевної порожнини, отриманих у венозній фазі контрастування, причому п'ятнадцять об'ємів містять пухлинні ураження різного типу та розміру. Кількість пухлин в одному

дослідженні може варіюватися від одиничних утворень до десятків окремих вогнищ, що створює складний сценарій для алгоритмів автоматичного аналізу.

Важливою особливістю набору даних є його клінічна неоднорідність. Зображення були отримані у різних медичних центрах із використанням різних моделей КТ-сканерів та протоколів сканування, що забезпечує високу варіативність інтенсивностей, рівнів шуму, параметрів контрастування та просторової роздільної здатності. Така неоднорідність робить датасет значно ближчим до реальних клінічних умов, проте одночасно суттєво ускладнює задачу навчання нейронних мереж. Алгоритм повинен бути здатним узагальнювати ознаки незалежно від конкретного томографа, параметрів реконструкції чи рівня контрастного підсилення.

Додаткову складність створює анізотропна структура медичних даних. Розмір вокселів у аксіальній площині становить приблизно 0,57–0,8 мм, тоді як міжзрізова відстань у сагітальній та коронарній площинах може досягати 1,25–4 мм. Через це інформація про структури вздовж осі зрізів є менш деталізованою, що погіршує якість тривимірного аналізу та ускладнює реконструкцію безперервних анатомічних поверхонь. Наслідком цього стає *partial volume effect* — ефект часткового об'єму, при якому один воксель може містити інформацію одразу про декілька типів тканин. Це призводить до розмиття меж між структурами та появи штучних з'єднань між судинами, пухлинами або сусідніми органами.

Однією з ключових проблем набору IRCAD є низький контраст між патологічними структурами та здоровою паренхімою печінки. Через відмінності у фазах контрастування, щільності тканин та характеристиках пухлин гіпоінтенсивні або ізоінтенсивні ураження можуть практично не відрізнятися від навколишньої тканини за гістограмою інтенсивностей. У результаті класичні алгоритми сегментації, що базуються лише на пороговій обробці або аналізі локальних градієнтів, втрачають здатність стабільно визначати межі патологічних структур. Особливо складними є випадки дрібних метастатичних уражень або пухлин із нечіткими контурами, де навіть для людини-візуалізатора задача локалізації є нетривіальною.

Ще однією фундаментальною проблемою є екстремальний дисбаланс класів. Основна частина КТ-зображення представлена фоном та іншими органами черевної порожнини, тоді як пікселі печінки займають лише невелику частину всього об'єму. Частка пікселів пухлин є ще меншою та може становити лише десяті частки відсотка від загальної кількості вокселів. У таких умовах стандартні функції втрат призводять до домінування фону під час оптимізації, через що модель схильна ігнорувати малі патологічні області. Це особливо критично при сегментації дрібних уражень, де навіть незначне зміщення межі може повністю нівелювати результат.

1.3 Постановка задачі дослідження

Попри значний прогрес у розвитку методів медичної візуалізації та алгоритмів комп'ютерного аналізу зображень, задача автоматичної сегментації печінки, її внутрішніх анатомічних структур та патологічних уражень досі залишається однією з найскладніших проблем медичного комп'ютерного зору. Основна складність полягає у необхідності одночасного врахування великої кількості факторів, серед яких висока анатомічна варіабельність органу, низький контраст між тканинами, складна просторова конфігурація судинної мережі, а також значна неоднорідність медичних даних, отриманих із різних томографічних систем та протоколів сканування.

Особливо критичною ця проблема стає при необхідності сегментації функціональних сегментів печінки за класифікацією Куїно, оскільки межі між такими сегментами не є фізично вираженими структурами, а визначаються переважно топологією порталльної венозної системи. У результаті задача переходить від простого аналізу локальних інтенсивностей до необхідності моделювання складних просторових та анатомічних взаємозв'язків між структурами. Автоматизація сегментації за Куїно додатково ускладнюється тим, що функціональні межі між сегментами визначаються судинними потоками і в загальному випадку не є планарними. Через це використання методів глибокого навчання для апроксимації таких топологічних меж на основі прихованих анатомічних орієнтирів стає критично необхідною умовою для точної локалізації структур у хірургічній практиці [4].

Класичні методи сегментації, що базуються на пороговій обробці, аналізі градієнтів, статистичних моделях форми або деформованих поверхнях, демонструють обмежену ефективність у реальних клінічних умовах. Причиною цього є їхня висока залежність від стабільності контрасту, наявності чітких меж та апріорних геометричних припущень щодо форми органу. При наявності пухлин, стеатозу, цирозу або післяопераційних змін такі підходи часто втрачають здатність коректно локалізувати межі печінки та її внутрішніх структур. Особливо складними є випадки, коли патологічні області мають інтенсивності, близькі до нормальної паренхіми, або коли анатомічна форма органу суттєво відрізняється від статистично очікуваної моделі. Стабільний приріст точності субсегментації печінки забезпечується методами, які здатні враховувати складні структурні залежності тканин під час оптимізації. Особливо гостро ця потреба постає при обробці КТ-даних пацієнтів із наявними пухлинними ураженнями, оскільки патології додатково деформують природну топологію органу та знижують ефективність класичних моделей [5]. Застосування згорткових нейронних мереж дозволило суттєво покращити результати сегментації завдяки автоматичному навчанню багаторівневих ознак без необхідності ручного проєктування дескрипторів.

Проте навіть сучасні CNN-архітектури стикаються з низкою фундаментальних проблем, однією з яких є екстремальний просторовий

дисбаланс класів, характерний для медичних даних. На КТ-зображеннях більшість пікселів належать фону або великим органам, тоді як судини та пухлини займають лише незначну частину простору, через що модель схильна оптимізуватися переважно на домінуючих класах, ігноруючи дрібні патологічні структури. Для подолання цього екстремального дисбалансу та проблеми нечітких меж дрібних патологій сучасна світова практика фокусується на модифікаціях ієрархічних архітектур сімейства U-Net. При цьому розв'язання проблеми «невидимості» малих сегментів виявляється неможливим лише за рахунок зміни топології мережі і вимагає комплексного впровадження спеціалізованих функцій втрат у поєднанні зі стратегічними методами аугментації даних [6].

Додатковою проблемою є поступова втрата просторової інформації у глибоких шарах мережі внаслідок операцій субдискретизації, що призводить до погіршення локалізації меж та втрати дрібних анатомічних деталей. Ця проблема успішно вирішується шляхом оптимізації злиття ознак на етапі декодування, де здатність архітектури ефективно поєднувати глибокий семантичний контекст із поверхневою просторовою деталізацією є вирішальним фактором для точного відтворення контурів печінки та її внутрішніх субструктур [7]. Окрему складність створює необхідність забезпечення просторової узгодженості результатів сегментації, оскільки більшість стандартних нейромережевих моделей виконують класифікацію вокселів локально, без повного врахування глобальних топологічних взаємозв'язків між анатомічними структурами. Це може призводити до появи хибнопозитивних областей, розривів судинних структур або формування анатомічно неможливих сегментаційних масок. Таким чином, задача високоточної сегментації печінки потребує комплексного підходу, який поєднуватиме глибокі нейронні мережі з механізмами просторового, топологічного та алгоритмічного контролю результату.

У межах даного дослідження об'єктом дослідження є процес автоматичної сегментації діагностичних медичних зображень, отриманих методами комп'ютерної та магнітно-резонансної томографії. Предметом дослідження виступають методи, алгоритми та архітектури глибокого навчання для виділення печінки, її функціональних сегментів, судинних структур та патологічних уражень на тривимірних томографічних даних. Основна увага приділяється ієрархічним архітектурам, здатним виконувати поетапне формування області інтересу та детальну сегментацію внутрішніх субструктур. Повністю згорткові мережі на базі U-Net зберігають статус беззаперечного стандарту в задачах сегментації томографічних медичних даних. Водночас для забезпечення високої просторової узгодженості та інваріантності до апаратних шумів ці моделі потребують безперервної алгоритмічної адаптації, що охоплює як регуляризацію простору параметрів, так і синтетичне збагачення навчальної вибірки [8]. Метою дослідження є підвищення точності, стабільності та надійності автоматичної сегментації

печінки та її внутрішніх анатомічних структур на КТ/МРТ-зображеннях для подальшого використання у системах планування хірургічного лікування та комп'ютерної підтримки прийняття клінічних рішень.

Для досягнення цієї мети використовується концепція декомпозиції складної задачі сегментації на послідовні ієрархічні етапи, що дозволяє зменшити вплив дисбалансу класів, локалізувати область аналізу та підвищити стійкість моделі до неоднорідності медичних даних. Реалізація поставленої мети передбачає дослідження впливу низької контрастності тканин, анізотропності даних та дисбалансу класів на ефективність сегментації дрібних анатомічних структур і патологічних утворень. Особлива увага приділяється аналізу здатності нейронних мереж зберігати локальну просторову інформацію при сегментації судинної мережі та меж функціональних сегментів печінки. Для компенсації цих обмежень необхідним є розроблення архітектурних підходів, які поєднуюватимуть глобальний семантичний контекст із високоточною локалізацією дрібних структур.

Одним із ключових напрямків дослідження є побудова та оптимізація повністю згорткових нейронних мереж, у яких задача сегментації розділяється на декілька послідовних етапів. На першому етапі виконується ізоляція печінки та формування просторової області інтересу, що дозволяє усунути нерелевантні структури черевної порожнини та суттєво зменшити дисбаланс класів. На другому етапі здійснюється детальна сегментація внутрішніх субструктур, зокрема функціональних сегментів, лише в межах локалізованої області печінки. Такий підхід дозволяє кожній нейронній мережі спеціалізуватися на окремому підзавданні та вивчати найбільш дискримінативні ознаки для відповідного рівня аналізу.

РОЗДІЛ 2. МЕТОДОЛОГІЯ ТА АРХІТЕКТУРИ МОДЕЛЕЙ ГЛИБОКОГО НАВЧАННЯ

2.1 U-Net

Матеріали даного розділу формують концептуальний та архітектурний базис проведеного дослідження. Послідовний розгляд нейромережових топологій, починаючи з базової структури U-Net, наводиться тут як необхідне алгоритмічне підґрунтя для синтезу власного обчислювального алгоритму. Такий підхід дозволяє методологічно обґрунтувати вибір та інтеграцію конкретних структурних компонентів у єдину систему, конфігурація якої оптимізується для розв'язання задачі просторової сегментації складних медичних даних.

Архітектура U-Net, запропонована Ronneberger, Fischer та Brox [9], стала де-факто стандартом для семантичної сегментації в біомедичній візуалізації. Вона є спеціалізованою варіацією повністю згорткових нейронних мереж, розробленою з урахуванням особливостей медичних датасетів, які зазвичай мають обмежену кількість зразків, як у випадку 3D-IRCADb-01. U-Net демонструє здатність ефективно навчатися на невеликих обсягах даних і забезпечує високу точність локалізації об'єктів, що робить її особливо придатною для багатокласової сегментації функціональних сегментів печінки за Куіно.

Структурно мережа має симетричну U-подібну форму (рис. 2.1), яка поєднує два взаємодоповнюючі шляхи. Шлях стиснення, або енкодер, складається з послідовних згорткових шарів та операцій макс-пулінгу, які поступово зменшують просторову роздільну здатність зображення, але при цьому забезпечують вилучення високорівневих семантичних ознак та глобального контексту. Водночас шлях розширення, або декодер, відновлює просторову розмірність до початкового розміру зображення, використовуючи шари підвищення роздільної здатності. Така організація дозволяє моделі поєднувати глибокі семантичні ознаки із детальною просторовою інформацією, що є критично важливим для сегментації сегментів печінки, межі яких часто не визначені чітко на зображенні.

Ключовим елементом, що забезпечує точність U-Net, є з'єднання з пропуском (skip connections), які безпосередньо пов'язують відповідні рівні енкодера та декодера, конкатенуючи карти ознак. Ця конструкція дозволяє компенсувати втрату високочастотної інформації, що відбувається через макс-пулінг, та передавати деталі просторових ознак у шлях відновлення. Завдяки цьому декодер може поєднувати глибокий семантичний контекст із точною локальною інформацією, що забезпечує коректне визначення меж сегментів навіть у складних умовах, коли вони є неявними або розмитими.

Окрім підвищення точності локалізації, з'єднання з пропуском позитивно впливають на динаміку навчання мережі, створюючи короткий шлях для проходження градієнта під час зворотного поширення помилки. Це зменшує проблему зникнення градієнта в глибоких шарах, робить процес оптимізації більш стабільним та сприяє швидкій збіжності навіть при обмеженій кількості навчальних зразків.

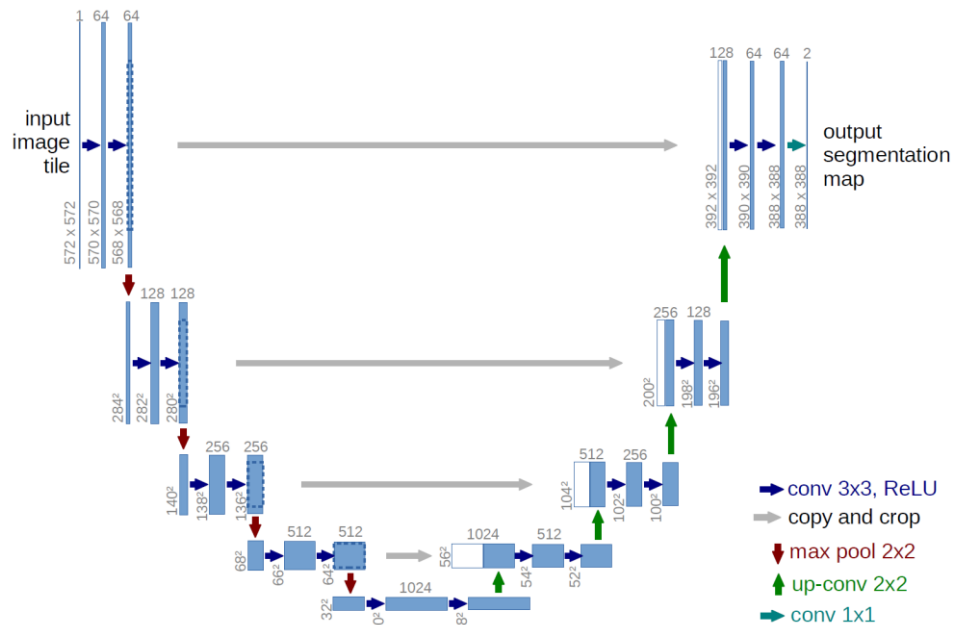


Рисунок 2.1 — Загальний приклад вигляду архітектури нейронної мережі U-Net

Для практичної реалізації U-Net у цій роботі використовується бібліотека Segmentation Models PyTorch [10], яка забезпечує гнучкість у налаштуванні мережі, зокрема інтеграцію попередньо навчених енкодерів. Застосування ResNet34 у ролі backbone дозволяє використовувати його потужні можливості вилучення ознак, що значно підвищує якість сегментації та стабільність навчання на обмежених медичних даних. Така комбінація архітектур робить модель здатною ефективно вирішувати задачу багатокласової сегментації функціональних сегментів печінки, враховуючи як глобальні, так і локальні особливості анатомічної структури [9][10].

2.2 Сімейство архітектур ResNet

У структурі U-Net, описаній у попередньому підрозділі, шлях стиснення може бути замінений на більш потужну попередньо навчену модель, яку часто називають backbone. Такий підхід дозволяє використовувати вже сформовані ознаки глибоких класифікаційних мереж і переносити їх у задачу сегментації. У даній роботі роль енкодера виконує архітектура ResNet, яка поєднує

достатню глибину, стабільність навчання та ефективність вилучення ознак, що робить її збалансованим вибором для обмежених медичних наборів даних.

Архітектура ResNet була запропонована He та колегами як розв'язання проблеми деградації точності при збільшенні кількості шарів у глибоких нейронних мережах. Основою цього підходу стали залишкові блоки, всередині яких сигнал може проходити через один або кілька шарів без змін. Таке з'єднання з пропуском (skip connection) забезпечує альтернативний шлях як для прямого сигналу, так і для градієнта під час зворотного поширення. На рівні окремого блока операція описується простою формулою:

$$\text{вихід} = F(x) + x \quad (2.1)$$

де x — вхід блока, а $F(x)$ — результат послідовності згорткових операцій(2.1). Якщо мережа не знаходить корисного перетворення, вона може наблизити $F(x)$ до нуля і забезпечити ідентичне відображення вхідного сигналу. Це запобігає погіршенню продуктивності при збільшенні глибини, робить навчання стабільнішим і дозволяє формувати глибші й насиченіші представлення ознак. ResNet34(рис. 2.2), зокрема, демонструє оптимальне співвідношення глибини та обчислювальної вартості, завдяки чому добре підходить для задач, де важлива деталізація, але обмежені апаратні ресурси та обсяг даних.

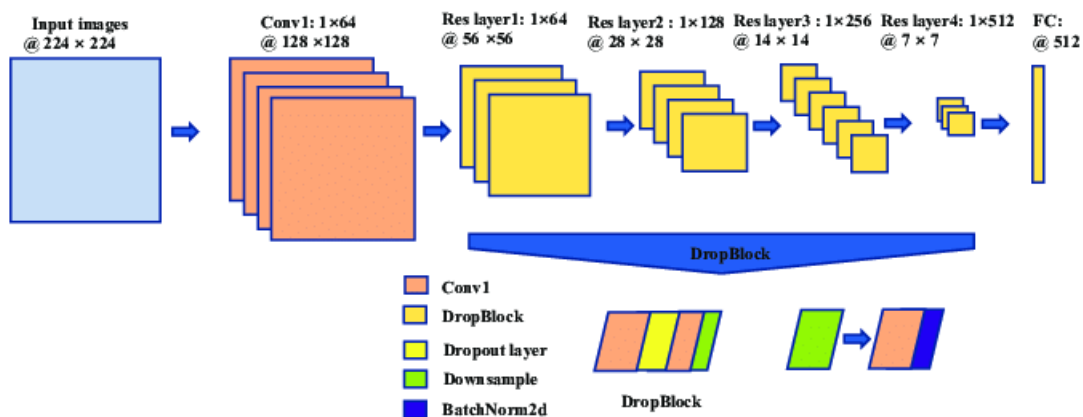


Рисунок 2.2 — Архітектура нейронної мережі ResNet34

Застосування ResNet34 у ролі енкодера в U-Net також забезпечує переваги трансферного навчання. Оскільки модель попередньо навчена на великому датасеті ImageNet, вона вже має здатність виділяти фундаментальні патерни, такі як контури, текстурні переходи, локальні контрасти та геометричні структури. У випадку сегментації сегментів печінки, де межі визначаються опосередковано положенням судин та їх розгалуженням, наявність таких попередньо вивчених ознак дає значний приріст точності. Модель швидше збігається, менше схильна до перенавчання та краще адаптується до дрібномасштабних змін інтенсивності, характерних для КТ-зображень у портальній фазі.

Ефективність енкодера може бути додатково підсилена за допомогою механізмів уваги. Дослідження Hu та співавторів у рамках Squeeze-and-

Excitation Networks [11] продемонстрували, що динамічне перевагове підсилення окремих каналів ознак дозволяє мережі краще враховувати їхню відносну важливість. Блок SE працює за принципом стискання просторової інформації (squeeze), аналізу статистик каналів та подальшого калібрування (excitation), підсилюючи ті канали, які найбільше сприяють розв'язуванню задачі. Для сегментації сегментів Куїно це особливо важливо, оскільки межі між сегментами визначаються не прямими інтенсивнісними переходами, а прихованою анатомічною організацією, що вимагає більш тонкого аналізу судинних структур.

Завдяки такому поєднанню U-Net та ResNet34 формується модель, здатна ефективно працювати з обмеженими медичними даними, зберігаючи точність визначення локальних структур і стійкість до варіацій інтенсивності та геометрії. Це є критично важливим для задачі багатокласової сегментації, де навіть невеликий зсув межі може призвести до помітних помилок інтерпретації у клінічному контексті [11].

2.3 Функції втрат для компенсації дисбалансу класів

У задачах семантичної сегментації медичних зображень, зокрема під час реконструкції анатомічної структури печінки та визначення сегментів Куїно I–VIII, вибір функції втрат (L) є критично важливим аспектом моделювання. У порівнянні зі звичайними комп'ютерними задачами, де об'єкти добре окреслені й займають порівняно велику частку зображення, медичні дані відзначаються значною нерівномірністю розподілу класів. Це особливо характерно для задачі сегментації печінкових сегментів, де окремі функціональні області можуть займати лише незначну частину кадру.

У таких умовах традиційна функція крос-ентропії (Cross-Entropy Loss) виявляється недостатньою. Вона має схильність надмірно акцентувати увагу на фоні, який становить більшість пікселів, і відповідно знижує вагу невеликих, але клінічно важливих сегментів. Наслідком є ситуація, коли модель досягає формально високої точності, але фактично не розпізнає критично важливі області або робить це нестабільно.

Функція втрат у цьому контексті перестає бути суто математичним критерієм — вона перетворюється на інструмент стратегічного управління процесом навчання. Вона визначає, які особливості зображення будуть підкріплюватися під час оптимізації та якою мірою модель реагуватиме на помилки класифікації дрібних структур. Правильний вибір L дозволяє збалансувати навчання таким чином, щоб модель не лише відтворювала загальну морфологію печінки, але й коректно розмежовувала сусідні сегменти, навіть якщо їхні межі слабо виражені або контрастність зображення є низькою.

У практичному застосуванні це означає, що функція втрат повинна компенсувати дисбаланс класів, підсилювати внесок рідкісних сегментів у

загальний градієнт та стимулювати модель зосереджуватися на просторових деталях. Вона має виконувати роль «регулятора чутливості» до локальних змін, забезпечуючи стабільну збіжність і зменшуючи ймовірність того, що модель зведе рішення до тривіального передбачення фону.

Одним із ключових викликів у задачах медичної сегментації є значна нерівномірність розподілу класів, коли об'єкти інтересу займають лише незначну частку зображення. Це характерно для сегментації анатомічних сегментів Куіно, де деякі з них мають невеликі просторові об'єми й легко «губляться» серед фонові тканини. У таких умовах традиційні функції втрат демонструють схильність переорієнтовувати навчання моделі на фон, що погіршує якість локалізації дрібних структур. Тому однією з найбільш ефективних стратегій компенсації дисбалансу є використання функції Dice Loss — похідної від популярної метрики перекриття Dice Similarity Coefficient (DSC).

Dice-коефіцієнт вимірює ступінь збігу між прогнозованою областю (P) і істинною розміткою (G), що дозволяє оцінити, наскільки добре модель відтворює форму, межі та просторову конфігурацію цільового об'єкта. На відміну від точкових метрик, що оцінюють кожен піксель окремо, DSC аналізує об'єкт як єдину структуру, що робить його природним вибором для медичних задач, де важливо не лише «вгадати піксель», а й відтворити анатомічну цілісність. У глибокому навчанні функція Dice Loss визначається як:

$$L_{Dice} = 1 - DSC \quad (2.2)$$

Бінарний коефіцієнт Дайса:

$$DSC = \frac{2 \cdot \sum p_i g_i + \varepsilon}{\sum p_i + \sum g_i + \varepsilon} \quad (2.3)$$

де

p_i — прогнозована ймовірність для пікселя i , g_i — істинна бінарна мітка, ε — мала константа стабілізації.

Чисельник тут відображає подвоєний об'єм перетину множин (intersection), а знаменник — нормовану суму об'ємів прогнозу та істини (union). Таким чином, DSC фактично оцінює рівень «подібності форм», що особливо важливо у контексті органів з плавними, анатомічно обумовленими контурами.

На відміну від крос-ентропії, Dice Loss реагує на помилки глобально: якщо модель пропускає невеликий сегмент (наприклад, Куіно I), значення DSC різко зменшується. Це створює сильний градієнт, стимулюючи модель більше «піклуватися» про дрібні структури. Саме тому Dice Loss є одним із найбільш уживаних інструментів у медичному моделюванні, де точність локалізації має клінічну вагу.

У випадку багатокласової сегментації (8 сегментів Куіно) застосування класичного бінарного Dice є недостатнім. Sudre et al. [12] запропонували узагальнену версію метрики — Generalised Dice Overlap, яка включає ваги для

кожного класу з урахуванням його об'єму. Це дозволяє компенсувати дисбаланс між великими сегментами (наприклад, VIII) та малими (I або II). Формула Generalised Dice:

$$GD = \frac{\sum_{c=1}^C w_c \cdot 2 \sum_i p_{ic} g_{ic}}{\sum_{c=1}^C w_c \cdot (\sum_i p_{ic} + \sum_i g_{ic})} \quad (2.4)$$

де

p_{ic} , g_{ic} — прогноз та істинна мітка для пікселя i і класу c ,
 w_c — вага класу, яка визначається як:

$$w_c = \frac{1}{(\sum_i g_{ic})^2} \quad (2.5)$$

Така постановка змушує модель приділяти більше уваги класам із малим об'ємом, компенсуючи їхню природну математичну «невидимість» у стандартних функціях втрат.

Для задачі сегментації за даними 3D-IRCADb-01 Generalised Dice Loss є практичною необхідністю, адже об'єми сегментів суттєво різняться. Використання Generalised Dice дозволяє гарантувати, що модель буде коректно враховувати внесок кожного сегмента, не ігноруючи ті, що мають мінімальну площу на зображенні. У результаті поліпшується не лише локальна точність, але й загальна стабільність моделі, особливо на крайових областях, де анатомічні межі є розмитими або слабо контрастними.

У процесі навчання моделей семантичної сегментації часто виникає ситуація, коли мережа швидко навчається впевнено розпізнавати більші, однорідні та прості для класифікації області, тоді як складніші зони — наприклад, межі анатомічних сегментів Куїно або невеликі, нечіткі ділянки — залишаються помилковими протягом значної частини епох. Це відбувається тому, що класичні функції втрат, такі як крос-ентропія, приділяють однакову вагу всім пікселям, незалежно від того, наскільки легко чи важко модель їх класифікує. У результаті «легкі» приклади починають домінувати у градієнті, зменшуючи здатність моделі вдосконалюватися в зоні складних, неоднозначних пікселів. Це особливо критично в анатомічній сегментації печінки, де дрібні сегменти займають мінімальний відсоток кадру, а межі між ними часто є розмитими та низькоконтрастними.

Для подолання цієї проблеми застосовується Focal Loss — модифікована версія крос-ентропії [13], яка була покликана змінити пріоритети навчання: знизити вагу добре класифікованих пікселів і натомість посилити внесок тих елементів, на яких модель помиляється. Базова формула крос-ентропії (CE) для ймовірності правильного класу p_t виглядає так:

$$CE(p_t) = -\log(p_t) \quad (2.6)$$

У Focal Loss додається модулюючий множник, який містить фокусуєчий параметр γ та ваговий коефіцієнт α . Узагальнена формула Focal Loss (FL):

$$FL(p_t) = -\alpha_t(1 - p_t)^\gamma \log(p_t) \quad (2.7)$$

де α_t керує початковим впливом позитивних і негативних прикладів, а γ відповідає за ступінь фокусування на складних пікселях. Якщо p_t є близьким до 1, тобто піксель класифікований правильно та з високою впевненістю, тоді $1 - p_t$ стає майже нулем, і піднесення цього значення до степеня γ практично «вимикає» внесок такого пікселя у загальну функцію втрат. Навпаки, коли p_t є малим, що відповідає помилковому або невпевненому прогнозу, множник $(1 - p_t)^\gamma$ залишається значним, і такий піксель формує більшу частку градієнта. Таким чином, параметр γ виконує роль динамічного регулятора складності, дозволяючи моделі зосереджуватися саме на тих областях, де її передбачення залишаються неточними.

Параметр α відіграє іншу, але не менш важливу роль — він компенсує початковий дисбаланс між позитивними і негативними класами. У задачі сегментації печінки позитивними класами є окремі сегменти Куіно, тоді як фон значно переважає. α дозволяє збільшувати або зменшувати вагу цих класів так, щоб навіть на ранніх етапах навчання модель не «скочувалася» до банальних рішень, наприклад — передбачати лише фон. У поєднанні з γ це створює механізм подвійного вирівнювання, де дисбаланс обсягів та дисбаланс складності компенсуються одночасно.

У застосуванні до багатокласової сегментації сегментів Куіно Focal Loss є важливим доповненням до Dice Loss. Якщо Dice Loss допомагає боротися з регіональною нерівномірністю розподілу класів, забезпечуючи справедливий внесок малих структур, то Focal Loss спрямовує увагу моделі на ті приклади, які залишаються складними навіть після кількох етапів навчання. Це дозволяє значно покращити коректність виділення меж, де відсутність чіткої структури, перешкоди у вигляді судинного русла та артефакти КТ створюють складні умови для сегментації.

Сучасні дослідження[14] додатково підкреслюють важливість цього підходу, пропонуючи адаптивні модифікації Focal Loss, які динамічно враховують розмір об'єкта, його локальну складність та ступінь варіативності структури. Усе це свідчить про те, що Focal Loss є не просто математичною поправкою, а стратегічно важливою складовою функції втрат для високоточної сегментації органів і їхніх сегментів, особливо в умовах, де малі об'єкти й неоднозначні межі мають ключове клінічне значення.

У задачах сегментації сегментів Куіно, де одночасно присутні як значний дисбаланс обсягів між класами, так і висока складність локальних анатомічних меж, особливо важливо забезпечити таку функцію втрат, яка здатна ефективно враховувати обидва фактори. Саме тому у даній роботі застосовується комбінований підхід, у якому загальна функція втрат представляє собою зважену суму Dice Loss та Focal Loss. Така комбінація дозволяє об'єднати переваги кожної з окремих функцій, створивши синергетичний ефект, який покращує стабільність і точність навчання.

$$L_{Total} = L_{Dice} + L_{Focal} \quad (2.8)$$

Сутність цього підходу полягає в тому, що кожна зі складових відповідає за подолання різних проблем, характерних для медичної сегментації. Dice Loss забезпечує коректне зважування регіонів, компенсуючи дисбаланс обсягів між великими структурами та малими сегментами Куїно, які можуть займати лише невеликий відсоток від загального числа пікселів. Саме Dice Loss не дозволяє моделі "забувати" про дрібні сегменти, навіть тоді, коли фон і великі регіони домінують у даних.

З іншого боку, Focal Loss додає динамічного фокусування на рівні окремих пікселів, змушуючи модель приділяти більше уваги труднощам — зокрема, нечітким межам між сусідніми сегментами та пікселям, які мають неоднозначну інтенсивність або геометрію. У рамках цієї роботи використовується фокусуючий параметр $\gamma = 2.0$, що підсилює ефект зменшення внеску легких для моделі прикладів і збільшує значущість складних пікселів, які визначають точність межових зон.

Комбінована функція втрат була реалізована за допомогою бібліотеки Segmentation Models PyTorch (SMP) у режимі multiclass, що забезпечує коректну роботу з восьмикласною сегментацією сегментів Куїно. Така інтеграція Dice Loss та Focal Loss дозволяє нейронній мережі U-Net з енкодером ResNet34 одночасно компенсувати дисбаланс розмірів структур і підвищувати точність у критичних анатомічних регіонах. Це робить дану комбінацію однією з найефективніших для сегментації складних медичних об'єктів, де важливими є навіть мінімальні локальні похибки.

2.4 Метрики оцінювання ефективності

Після визначення та налаштування функцій втрат наступним важливим елементом методології стає вибір метрик, за допомогою яких буде оцінюватися якість роботи моделі після завершення навчання. На відміну від функцій втрат, що безпосередньо впливають на процес оптимізації, метрики виконують іншу роль — вони забезпечують зрозумілу та інтерпретовану оцінку продуктивності моделі, дозволяючи досліднику порівнювати різні підходи, конфігурації та результати з наявними у літературі дослідженнями. Саме метрики дають змогу зрозуміти, наскільки точно модель відтворює анатомічну структуру сегментів Куїно та чи здатна вона працювати стабільно на різних томографічних зрізах.

У випадку багатокласової сегментації печінки, де кожен сегмент Куїно представляє окремий клас зі своїми морфологічними особливостями, доцільність застосування кількох типів метрик є очевидною. По-перше, бінарні метрики дозволяють отримати загальну інтегральну оцінку роботи моделі по всіх сегментах одразу, надаючи можливість контролювати глобальну якість сегментації незалежно від специфіки окремих структур. По-друге, багатокласові метрики забезпечують деталізований аналіз, що є

необхідним у медичних задачах, де навіть один погано сегментований сегмент може суттєво вплинути на клінічне трактування результатів.

Для первинного контролю якості роботи моделі та перевірки її здатності коректно визначати орган у цілому використовується бінарна метрика Whole-Liver IoU (Intersection over Union). Її застосування ґрунтується на принципі розгляду всієї печінки як єдиного об'єкта: усі сегменти Куіно (I–VIII), незалежно від їхніх розмірів та локалізації, об'єднуються у спільний клас, що дозволяє оцінити базову точність локалізації органу на зображенні.

Таке рішення має важливе методологічне обґрунтування. Сегментація цілої печінки є найпростішим та найфундаментальнішим етапом, без успішного проходження якого неможливо очікувати коректного поділу органу на окремі сегменти. Якщо модель не здатна чітко визначити межі печінки як єдиного анатомічного об'єкта, це автоматично означає неможливість точної сегментації кожного сегмента Куіно окремо. Такий підхід узгоджується з усталеними практиками міжнародних досліджень у сфері комп'ютерного зору[15], де IoU використовується як базова метрика якості сегментації.

Метрика Intersection over Union вимірює ступінь перекриття між істинною маскою та прогнозом моделі. Інтуїтивно IoU показує, наскільки точно модель «потрапляє» у форму та контур об'єкта, і наскільки добре відтворює його межі на зображенні. Розрахунок IoU для бінарної сегментації цілої печінки ґрунтується на співвідношенні площі перетину та площі об'єднання бінарних масок.

$$IoU = \frac{|P \cap G|}{|P \cup G|} \quad (2.9)$$

або у розгорнутому вигляді:

$$IoU = \frac{TruePositives}{TruePositives + FalsePositives + FalseNegatives} \quad (2.10)$$

де P — прогнозована бінарна маска печінки, а G — істинна маска, отримана з розмітки.

Для обчислення Whole-Liver IoU пікселі всіх сегментів I–VIII у істинній розмітці об'єднуються у єдиний клас G = "Печінка", а всі відповідні передбачення моделі також перетворюються у бінарну форму P = "Печінка". Далі IoU обчислюється між цими двома масками, що дозволяє отримати інтегральну оцінку якості сегментації всього органу незалежно від структури його внутрішніх меж.

Високе значення Whole-Liver IoU, особливо близьке до 1, свідчить про те, що модель здатна точно визначати анатомічні межі печінки та коректно відтворює її форму. Водночас сама по собі ця метрика не дозволяє оцінити якість сегментації внутрішніх сегментів Куіно, оскільки не враховує розподіл пікселів усередині органу. Саме тому вона розглядається як допоміжний показник і потребує обов'язкового доповнення багатокласовими метриками. Як зазначають Таһа and Hanbury [16], метрики перекриття, такі як IoU та Dice, є найбільш надійними критеріями для оцінювання просторової коректності

медичної сегментації, що підтверджує доцільність використання IoU у якості базового індикатора.

У той час як метрика Whole-Liver IoU забезпечує лише загальне уявлення про те, наскільки добре модель здатна локалізувати орган у цілому, цього підходу недостатньо для детального аналізу внутрішньої структури печінки. Саме тому основною метрикою для повноцінної оцінки сегментації восьми функціональних сегментів Куіно виступає Micro-Averaged IoU. Ця метрика дозволяє встановити, наскільки точно модель U-Net/ResNet здатна розмежовувати складні анатомічні зони, межі яких зазвичай є розмитими, неоднорідними та дуже чутливими до будь-яких неточностей у прогнозі. Таким чином, Micro-IoU відіграє ключову роль у тому, щоб оцінювання не обмежувалося лише глобальною відповідністю, а відображало реальну якість поділу печінки на її функціональні сегменти.

Суть Micro-Averaged IoU полягає у поширенні класичної метрики Intersection over Union на випадок багатокласової сегментації. На практиці існують два підходи до усереднення значень IoU: це Macro-Averaging та Micro-Averaging. У першому випадку кожен клас розглядається окремо, і після отримання IoU для кожного сегмента ці значення усереднюються, забезпечуючи рівну важливість як великим, так і малим сегментам. Однак у реальних медичних даних розмір сегмента відіграє важливу роль, і тому підхід мікроусереднення є більш інформативним. Micro-Averaged IoU зводить усі сегменти до спільного простору підрахунків True Positives, False Positives та False Negatives, після чого метрика обчислюється вже не для кожного класу окремо, а для всієї системи сегментації в цілому.

$$Micro - IoU = \frac{\sum (TP_c)}{\sum (TP_c + FP_c + FN_c)} \quad (2.11)$$

де індекс c пробігає усі цільові сегменти, а сумування охоплює вісім класів Куіно. Завдяки такому підходу модель оцінюється відповідно до реального розподілу пікселів: більші сегменти мають більшу вагу у фінальній метриці, що природно відображає вплив їхнього об'єму на клінічні сценарії та точність усього процесу сегментації.

Обґрунтування вибору Micro-IoU у цьому дослідженні має кілька взаємопов'язаних причин. По-перше, ця метрика найточніше відображає загальну якість сегментації на рівні пікселів, адже кожен піксель робить пропорційний внесок у фінальний результат. Це надзвичайно важливо, оскільки сегменти печінки мають різний розмір, і помилки в сегментації великого сегмента (наприклад, сегмента VIII) природно мають більшу вагу, ніж похибки у малих регіонах. По-друге, Micro-IoU напряду корелює з тим, як працює комбінована функція втрат, що об'єднує Dice Loss і Focal Loss. Обидві функції спрямовані на підвищення точності саме на рівні окремих пікселів, тому Micro-IoU стає найбільш природною та узгодженою метрикою для оцінювання результативності оптимізації. По-третє, у контексті даного дослідження Micro-IoU дозволяє об'єктивно порівняти різні стратегії

аугментації та визначити, яка з них забезпечує найбільший приріст точності для багатокласової сегментації, що є критично важливим у подальшому аналізі.

Як підкреслюється у роботі Taħa and Hanbury [16], метрики перекриття, такі як IoU, залишаються найбільш перевіреним та надійним інструментом оцінювання якості тривимірної медичної сегментації. Тому використання саме Micro-Averaged IoU дозволяє забезпечити об'єктивність, інтерпретованість та порівнюваність результатів, що робить цю метрику центральною складовою методології даної роботи.

РОЗДІЛ 3. Аналіз впливу функцій активації на збіжність глибоких мереж

3.1 Функції активації

Теоретичні та прикладні аспекти, викладені у цьому розділі, становлять методологічну основу для подальшої оптимізації внутрішнього простору параметрів розробленої моделі. Аналіз різних класів нелінійностей, що розпочинається з базових функцій активації, виступає тут алгоритмічною передумовою для проведення серії числових експериментів. Подібне структурне представлення дозволяє прослідкувати, як зміна локальних математичних перетворень впливає на загальний оптимізаційний ландшафт та динаміку збіжності глибоких архітектур.

Одним із фундаментальних компонентів глибоких нейронних мереж є функція активації, яка визначає характер перетворення вхідного сигналу нейрона у вихідний. Саме введення нелінійності дозволяє нейронним мережам апроксимувати складні багатовимірні залежності та формувати ієрархічні представлення ознак. Без використання нелінійних активацій навіть багатошарова мережа фактично перетворювалася б на еквівалент однієї лінійної трансформації, що критично обмежувало б її виражальну здатність. У задачах комп'ютерного зору, зокрема сегментації медичних зображень, функції активації мають особливе значення, оскільки вони безпосередньо впливають на стабільність градієнтного спуску, швидкість збіжності моделі, здатність виділяти складні текстурні структури та ефективність поширення інформації у глибоких шарах мережі[17].

У ранніх архітектурах нейронних мереж переважно використовувалися функції насичення, зокрема логістична сигмоїда та гіперболічний тангенс. Логістична сигмоїда(рис. 3.1) математично описується як:

$$f(x) = \frac{1}{1+e^{-x}} \quad (3.1)$$

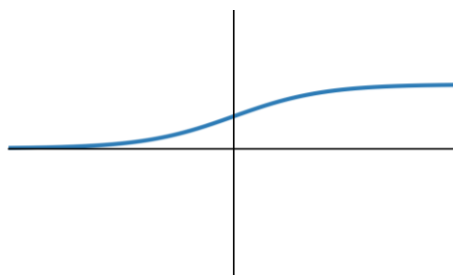


Рисунок 3.1 — Логістична сигмоїда

Тоді як функція гіперболічного тангенса(рис 3.2) має вигляд:

$$f(x) = \tanh(x) \quad (3.2)$$

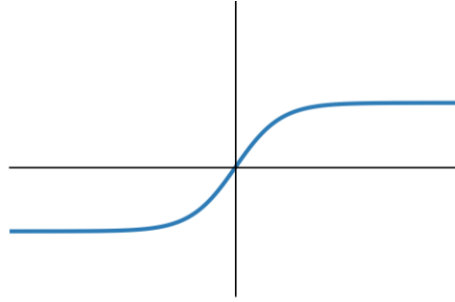


Рисунок 3.2 — Гіперболічний тангенс

Основною перевагою таких функцій є їхня гладкість та здатність обмежувати значення виходу у фіксованому діапазоні. Це певною мірою стабілізує активації мережі та дозволяє інтерпретувати вихід нейрона як нормалізований сигнал. Проте використання функцій насичення супроводжується суттєвими проблемами під час оптимізації глибоких архітектур. У зонах великих додатних або від'ємних значень аргументу похідна сигмоїди та $\tanh$ наближається до нуля, що призводить до явища згасання градієнта. У процесі зворотного поширення помилки градієнти послідовно перемножуються між шарами, тому навіть незначне зменшення похідних у кожному шарі експоненційно послаблює сигнал оновлення ваг у ранніх рівнях мережі. Це значно ускладнює навчання глибоких моделей та уповільнює процес конвергенції.

Критичний прорив у розвитку глибокого навчання відбувся після впровадження випрямленої лінійної функції активації ReLU (Rectified Linear Unit), яка була популяризована у роботі Krizhevsky, Sutskever та Hinton під час створення архітектури AlexNet. Функція ReLU (рис 3.3) описується наступним співвідношенням:

$$(3.3) \quad f(x) = \max(0, x)$$

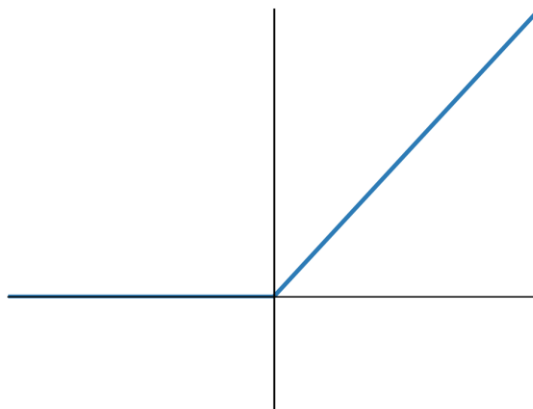


Рисунок 3.3 — ReLU

На відміну від сигмоїдальних функцій, ReLU не має області насичення для додатних значень аргументу, завдяки чому похідна залишається сталою та не прямує до нуля. Це суттєво покращує проходження градієнта через глибокі шари мережі та мінімізує проблему його згасання. У результаті нейронні мережі з ReLU здатні навчатися у декілька разів швидше порівняно з архітектурами, що використовують $\tanh$ або сигмоїду. Особливо важливим це стало для розвитку глибоких згорткових мереж, де кількість параметрів може сягати десятків або сотень мільйонів.

Важливою властивістю ReLU є також її обчислювальна простота. Функція не потребує обчислення експоненти або інших складних математичних операцій, що зменшує навантаження на графічні процесори та дозволяє ефективно масштабувати моделі. У задачах сегментації медичних зображень це є особливо актуальним, оскільки обробка тривимірних КТ або МРТ-об'ємів вимагає значних обчислювальних ресурсів. Завдяки своїй простоті ReLU дозволяє використовувати глибші архітектури без критичного збільшення часу навчання.

Додатковою перевагою ReLU є її здатність формувати розріджені активації. Для всіх від'ємних значень вхідного сигналу вихід функції дорівнює нулю, через що значна частина нейронів у певний момент часу залишається неактивною. Така поведінка створює ефект часткової регуляризації, зменшуючи надлишкову кореляцію між ознаками та підвищуючи узагальнювальну здатність моделі. У задачах комп'ютерного зору це дозволяє мережі формувати більш селективні та дискримінативні просторові представлення.

Попри значні переваги, ReLU має і низку обмежень. Основною проблемою є явище «мертвих нейронів» (dying ReLU), при якому нейрон перестає активуватися для всіх вхідних даних. Якщо вагові коефіцієнти зміщують вхідний сигнал у від'ємну область, похідна функції стає нульовою, а відповідний нейрон більше не оновлюється під час навчання. У глибоких мережах це може призводити до поступової деградації частини ознак та втрати виражальної здатності моделі. Особливо помітною ця проблема стає при використанні великих швидкостей навчання або нестабільної ініціалізації ваг.

Для часткової компенсації цього ефекту застосовується спеціальна стратегія ініціалізації параметрів. Зокрема, зміщення у згорткових та повнозв'язних шарах можуть ініціалізуватися невеликими додатними значеннями, що забезпечує активацію нейронів на ранніх етапах оптимізації. Додатково використовуються методи нормалізації, зокрема Batch Normalization, які стабілізують розподіл активацій між шарами та зменшують ризик переходу великої кількості нейронів у неактивний стан.

Важливість ReLU у розвитку сучасних архітектур важко переоцінити, оскільки саме ця функція стала основою для більшості сучасних згорткових мереж у задачах комп'ютерного зору. Її використання дозволило суттєво збільшити глибину моделей та перейти до побудови високопродуктивних

архітектур, здатних ефективно працювати із великими наборами даних. У сегментації медичних зображень ReLU забезпечила можливість створення складних енкодер-декодерних архітектур, таких як U-Net та її модифікації, де стабільне проходження градієнта є критично важливим для точного відновлення просторових структур органів та патологічних утворень.

Попри домінування ReLU у сучасних нейромережевих системах, подальші дослідження показали, що її жорстка нелінійність та повне занулення від'ємної області можуть обмежувати інформаційну пропускну здатність мережі. Це стало передумовою для появи нових функцій активації, які зберігають переваги ReLU, але водночас забезпечують плавніший розподіл градієнтів, кращу динаміку оптимізації та стабільніше навчання дуже глибоких моделей. Саме тому сучасні дослідження дедалі більше орієнтуються на модифіковані функції активації із навченими параметрами, адаптивними коефіцієнтами або саморегульованими нелінійностями, здатними ефективніше працювати в умовах складних високорозмірних медичних даних.

3.2 Функції із навченим параметром

Подальший розвиток випрямлених лінійних функцій активації був спрямований на подолання головного обмеження ReLU — повного занулення всієї від'ємної області. Хоча така властивість забезпечує розрідженість активацій та стабільне проходження градієнта у додатній області, вона одночасно призводить до втрати частини інформації та появи явища «мертвих нейронів». У відповідь на ці обмеження були запропоновані адаптивні функції активації, параметри яких можуть оптимізуватися безпосередньо під час навчання нейронної мережі. На відміну від класичних статичних нелінійностей, такі функції дозволяють моделі самостійно формувати оптимальну геометрію активаційного простору відповідно до особливостей даних та структури задачі[18].

Однією з перших та найбільш важливих модифікацій стала функція PReLU(рис 3.4) (Parametric Rectified Linear Unit), запропонована Kaiming He та співавторами. Її ключова ідея полягає у заміні жорсткого занулення від'ємної області на параметризовану лінійну гілку із навченим коефіцієнтом нахилу. Математично функція описується як:

$$f(x) = \begin{cases} x, & x > 0 \\ ax, & x \leq 0 \end{cases} \quad (3.4)$$

де параметр (a) визначає нахил від'ємної частини функції та оптимізується одночасно з іншими вагами нейронної мережі. На відміну від стандартного ReLU, де всі від'ємні значення повністю втрачаються, PReLU дозволяє зберігати частину інформації навіть у випадках негативної

преактивації. Це суттєво покращує проходження градієнта через мережу та зменшує ризик повної деактивації нейронів.

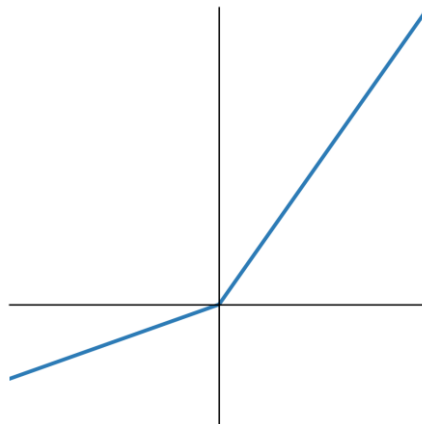


Рисунок 3.4 — PReLU

Важливою особливістю PReLU є її адаптивність. Коефіцієнт нахилу не задається вручну, а автоматично підлаштовується під статистику ознак у конкретному шарі або каналі. У результаті модель отримує можливість самостійно визначати, наскільки сильно необхідно пригнічувати від'ємні активації. Емпіричні дослідження показали, що ранні шари нейронної мережі зазвичай формують більші значення параметра (α), що дозволяє зберігати більше низькорівневої інформації. Натомість у глибших шарах значення коефіцієнта поступово зменшуються, посилюючи нелінійність представлень та сприяючи формуванню більш розподілених ознак.

Ключовою перевагою PReLU є те, що впровадження адаптивного параметра практично не збільшує складність моделі. Кількість додаткових параметрів дорівнює лише кількості каналів активаційного тензора, тому вплив на обчислювальні витрати та пам'ять є мінімальним. При цьому навіть така незначна модифікація дозволила суттєво покращити результати глибоких згорткових мереж у задачах комп'ютерного зору. Особливо важливим це стало для дуже глибоких архітектур, де стабільність проходження градієнта безпосередньо визначає здатність мережі до ефективного навчання.

Окремо варто відзначити особливості оптимізації параметра (α). Під час навчання до нього зазвичай не застосовується *weight decay*-регуляризація, оскільки це створювало б систематичне зміщення коефіцієнта до нуля та фактично перетворювало б PReLU назад у класичний ReLU. Таким чином, збереження свободи адаптації параметра є критично важливим для підтримання гнучкості моделі та стабільної динаміки навчання.

Подальший розвиток адаптивних функцій активації привів до появи функції Swish, яка була знайдена не вручну, а за допомогою автоматизованого пошуку архітектур із використанням методів *reinforcement learning*. На відміну від більшості класичних нелінійностей, Swish поєднує властивості гладкої

сигмоїдальної функції та лінійного проходження інформації. Її математичний вигляд описується рівнянням[19]:

$$f(x) = x \cdot \sigma(\beta x) \quad (3.5)$$

де $\sigma(\beta x)$ є логістичною сигмоїдою, а параметр (β) визначає ступінь нелінійності функції та може бути як фіксованим, так і навченим під час оптимізації мережі. У випадку ($\beta=0$) функція поводить себе майже лінійно, тоді як при великих значеннях параметра її поведінка поступово наближається до ReLU.

Однією з найважливіших властивостей Swish(рис 3.5) є її гладка немонотонність. На відміну від ReLU, функція не має різкого переходу в точці нуля, а формує плавний нелінійний профіль із локальним мінімумом у від'ємній області. Така поведінка дозволяє моделі частково пропускати слабкі негативні сигнали замість їх повного занулення. У результаті покращується проходження градієнта, зменшується ризик інформаційних втрат та формується більш стабільний процес оптимізації.

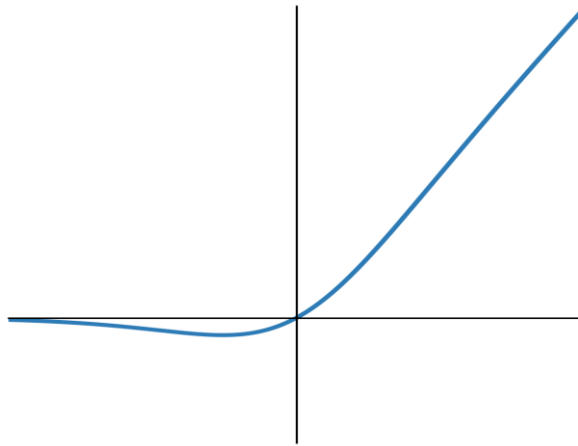


Рисунок 3.5 — Swish

Особливо важливою ця властивість є у дуже глибоких архітектурах комп'ютерного зору, де накопичення навіть незначних втрат інформації між шарами може призводити до деградації представлень. Гладкість Swish забезпечує неперервність похідної, завдяки чому градієнтний спуск стає стабільнішим порівняно з кусочно-лінійними функціями. Це позитивно впливає на швидкість збіжності та дозволяє мережі ефективніше адаптуватися до складних високорозмірних розподілів даних[20].

Дослідження розподілу активацій показали, що значна частина преактивацій нейронної мережі потрапляє саме у зону немонотонного екстремуму Swish. Це свідчить про те, що модель активно використовує цю область для формування нелінійних представлень ознак. На практиці функція продемонструвала стабільне покращення результатів у задачах класифікації, детекції та сегментації без необхідності зміни самої архітектури мережі. Саме тому Swish та її спрощена реалізація SiLU стали активно використовуватися у

сучасних високопродуктивних архітектурах комп'ютерного зору, зокрема EfficientNet.

У задачах медичної сегментації використання адаптивних функцій активації є особливо перспективним через складну природу даних. Низький контраст між тканинами, шум, неоднорідність томографічних зрізів та необхідність локалізації дрібних структур вимагають від мережі максимальної чутливості до слабких ознак. На відміну від жорсткого занулення негативної області, характерного для ReLU, параметризовані та гладкі функції активації дозволяють зберігати більше інформації про структуру сигналу та формувати стабільніші просторові представлення. Це є особливо важливим у сегментації судинної мережі, меж патологічних утворень та функціональних сегментів печінки, де навіть незначні локальні ознаки можуть мати критичне клінічне значення.

3.3 Функції із сталим параметром

Подальший розвиток сучасних функцій активації був спрямований на пошук компромісу між адаптивністю нелінійності, стабільністю оптимізації та обчислювальною ефективністю. Хоча параметризовані функції, такі як PReLU або Swish із навченим коефіцієнтом (β), забезпечують високу гнучкість, наявність додаткових параметрів ускладнює процес оптимізації та збільшує обчислювальні витрати. У відповідь на це були запропоновані функції активації зі сталими параметрами, які зберігають ключові переваги гладких немонотонних нелінійностей, але не потребують додаткового навчання. Такий підхід дозволяє формувати стабільніший ландшафт функції втрат, спрощувати процес збіжності та одночасно підтримувати високу якість передачі градієнтів у глибоких нейронних мережах.

Однією з найважливіших функцій цього класу стала SiLU(рис 3.6) (Sigmoid-weighted Linear Unit), також відома як Swish-1. Вона є спрощеним випадком функції Swish, у якому параметр (β) фіксується на рівні одиниці. Математично SiLU визначається як[20]:

$$f(x) = x \cdot \sigma(x)$$

(3.6)

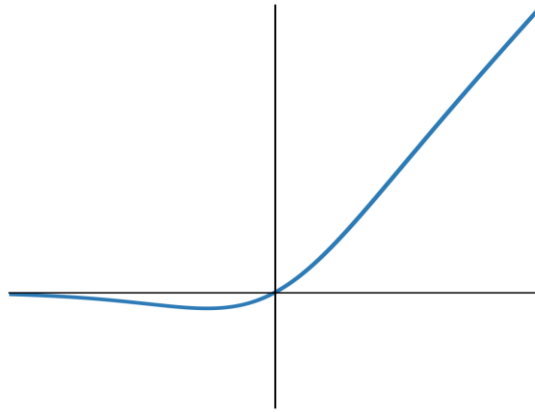


Рисунок 3.6 — SiLU

де $\sigma(x)$ є логістичною сигмоїдальною функцією. На відміну від класичного ReLU, який повністю відкидає від'ємну область, SiLU плавно масштабує вхідний сигнал сигмоїдою, забезпечуючи неперервний перехід між негативними та позитивними значеннями. Завдяки цьому функція не має жорсткої точки розриву похідної в нулі, що позитивно впливає на стабільність градієнтного спуску.

Ключовою особливістю SiLU є її гладка немонотонність. У невеликій області від'ємних значень функція формує локальне «поглиблення», через яке слабкі негативні сигнали не повністю занулюються, а частково передаються далі мережею. Така поведінка дозволяє зберігати більше інформації про структуру вхідних ознак та покращує проходження градієнта у глибоких архітектурах. На практиці це особливо важливо у задачах комп'ютерного зору, де значна частина ознак формується на основі слабких локальних текстурних патернів.

Важливою перевагою SiLU є також її гладкість. Неперервність першої похідної створює більш стабільний ландшафт функції втрат порівняно з кусочно-лінійними активаціями. У результаті процес оптимізації стає менш чутливим до коливань градієнта, а модель демонструє стабільнішу збіжність на ранніх етапах навчання. Це особливо критично для надглибоких згорткових архітектур, де навіть незначна нестабільність поширення градієнтів між шарами може призводити до деградації навчання.

Практична ефективність SiLU була підтверджена у масштабованих архітектурах сімейства EfficientNet, де ця функція стала одним із ключових елементів оптимізації мережі. Поєднання гладкої нелінійності, відсутності жорсткого насичення та здатності пропускати слабкі негативні сигнали дозволило покращити узагальнювальну здатність моделей без зміни базової архітектури. У задачах медичної сегментації це має особливе значення, оскільки моделі повинні працювати із низькоконтрастними структурами, де навіть незначні локальні градієнти можуть містити важливу діагностичну інформацію.

Подальшим розвитком ідеї гладких немонотонних функцій стала активація Mish(рис 3.7), запропонована Diganta Misra. Ця функція концептуально розширює підхід SiLU шляхом використання більш складної нелінійної трансформації, заснованої на Softplus та гіперболічному тангенсі. Аналітично Mish описується виразом[21]:

$$f(x) = x \cdot \tanh(\ln(1 + e^x))$$

(3.7)

Функція Softplus всередині Mish є гладкою апроксимацією ReLU та дозволяє уникнути різких переходів між областями активації. Подальше застосування гіперболічного тангенса формує плавну нелінійність із контрольованим обмеженням негативної області. У результаті Mish поєднує одразу декілька важливих властивостей: необмеженість зверху, обмеженість знизу, гладкість та немонотонність.

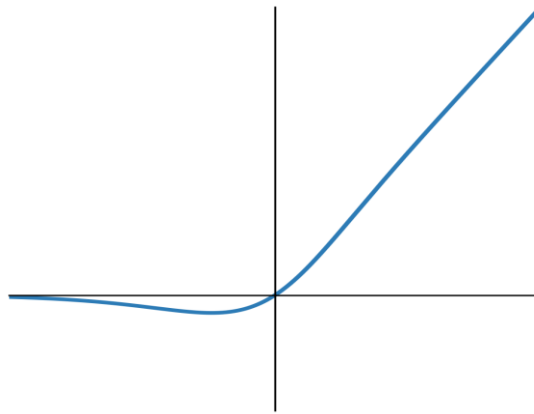


Рисунок 3.7 — Mish

Однією з найважливіших характеристик Mish є абсолютна неперервність першої та другої похідних. Це забезпечує формування дуже гладкого ландшафту функції втрат та покращує динаміку градієнтного спуску. У контексті оптимізації це означає, що модель отримує більш стабільні напрямки оновлення ваг і легше уникає різких локальних мінімумів або нестабільних коливань. Аналіз геометрії поверхні втрат показує, що гладкі функції активації позитивно впливають на обумовленість матриці Гессе, а отже — спрощують оптимізацію дуже глибоких моделей[22].

На відміну від ReLU, Mish не відкидає повністю негативні значення, а зберігає слабкі від'ємні активації малого модуля. Це дозволяє мережі формувати більш багаті семантичні представлення та ефективніше використовувати інформацію у межах великого рецептивного поля. У задачах сегментації медичних зображень така властивість є особливо важливою, оскільки межі патологічних структур часто характеризуються слабкими градієнтами та низьким контрастом. Збереження навіть невеликих негативних

активацій дозволяє точніше моделювати локальні текстурні переходи між тканинами.

Практична ефективність Mish була підтверджена у високопродуктивних архітектурах комп'ютерного зору, зокрема в детекторі об'єктів YOLOv4. Використання цієї функції активації забезпечило покращення якості локалізації та класифікації об'єктів без необхідності суттєвої модифікації самої архітектури мережі. Це свідчить про те, що вибір функції активації може мати прямий вплив на здатність моделі формувати стабільні та дискримінативні ознаки навіть за незмінної топології мережі.

Важливою перевагою функцій зі сталими параметрами є відсутність додаткових ступенів вільності, які потребують окремої оптимізації. Це дозволяє уникнути ускладнення процесу навчання та зменшує ризик нестабільності, пов'язаної з переналаштуванням параметрів активації. У результаті такі функції поєднують високу виражальну здатність із хорошою обчислювальною ефективністю, що робить їх особливо привабливими для використання у задачах медичної візуалізації, де моделі часто працюють із великими тривимірними томографічними даними та обмеженими апаратними ресурсами.

Таким чином, SiLU та Mish демонструють перехід від жорстких кусочно-лінійних активацій до гладких немонотонних нелінійностей, здатних забезпечувати стабільніше проходження градієнта, краще збереження інформації та ефективніше формування складних багаторівневих представлень ознак. Саме ці властивості роблять їх перспективними кандидатами для використання у сучасних системах медичної сегментації та високоточних архітектурах комп'ютерного зору.

3.4 Функції із гіперпараметром

Окремий напрям розвитку функцій активації пов'язаний із використанням фіксованих гіперпараметрів, які визначають характер нелінійності ще до початку процесу навчання. На відміну від параметризованих функцій, де коефіцієнти адаптуються під час оптимізації мережі, у цьому підході форма активації залишається незмінною, що дозволяє спростити процес навчання та зменшити обчислювальну складність моделі. Основна ідея таких функцій полягає у збереженні ключових переваг ReLU — швидкого проходження градієнта та обчислювальної простоти — із одночасною компенсацією її фундаментальних недоліків, насамперед проблеми «мертвих нейронів» та втрати інформації у від'ємній області простору ознак.

Однією з найпоширеніших функцій цього класу є Leaky ReLU (рис 3.8) (LReLU), яка модифікує класичний ReLU шляхом введення малого ненульового нахилу для від'ємної області. Математично функція описується як:

$$f(x) = \begin{cases} x, & x \geq 0 \\ \alpha x, & x < 0 \end{cases} \quad (3.8)$$

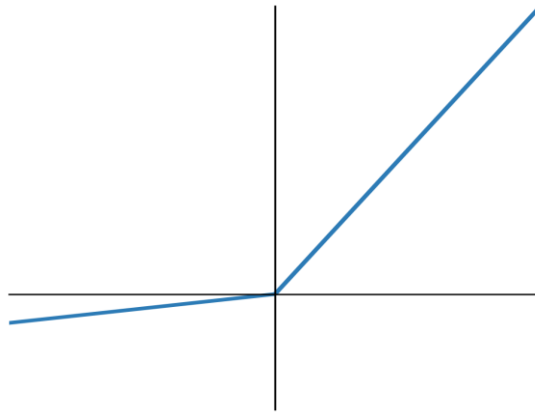


Рисунок 3.8 — Leaky ReLU

де (α) є фіксованим гіперпараметром, який зазвичай задається малим додатним значенням, наприклад (0.01). На відміну від стандартного ReLU, що повністю занулює всі негативні значення, Leaky ReLU дозволяє невеликій частині інформації проходити навіть у від'ємній області. Завдяки цьому градієнт ніколи не стає повністю нульовим, а нейрон зберігає здатність до оновлення ваг під час зворотного поширення помилки[23].

Ключовою перевагою Leaky ReLU є структурне усунення проблеми «мертвих нейронів». У стандартному ReLU нейрон може повністю втратити активність, якщо його вхід стабільно переходить у від'ємну область, оскільки похідна функції в такому випадку дорівнює нулю. Це призводить до припинення оновлення ваг і фактичного виключення нейрона з процесу навчання. У Leaky ReLU навіть мінімальний нахил у негативній області забезпечує постійне проходження слабкого градієнта, що підтримує активність нейронів та дозволяє моделі продовжувати адаптацію параметрів.

Важливою особливістю Leaky ReLU є її обчислювальна простота. Функція зберігає кусочно-лінійну структуру та не потребує використання експоненціальних або тригонометричних операцій, характерних для гладких нелінійностей. Це робить її особливо привабливою для великих згорткових мереж, де швидкість виконання операцій має критичне значення. У задачах медичної сегментації така ефективність є важливою через необхідність обробки великих тривимірних томографічних об'ємів із високою просторовою роздільною здатністю.

Попри свою простоту, Leaky ReLU демонструє суттєво стабільнішу динаміку навчання порівняно з класичним ReLU. Наявність слабкої негативної гілки дозволяє мережі зберігати більше інформації про розподіл ознак та формувати плавніші переходи між активаційними станами. Особливо помітним це стає у дуже глибоких архітектурах, де накопичення повністю занулених активацій може поступово знижувати інформаційну пропускну здатність мережі.

Однак фіксований характер параметра (α) одночасно є і обмеженням цього підходу. Оскільки коефіцієнт не адаптується до особливостей даних, модель не може оптимально налаштовувати ступінь пригнічення негативних сигналів у різних шарах. Це означає, що одна й та сама величина нахилу використовується як для низькорівневих текстурних ознак, так і для високорівневих семантичних представлень, хоча оптимальні значення можуть суттєво відрізнятися. Саме це стало передумовою для подальшого розвитку параметризованих активацій, зокрема PReLU.

Іншим важливим представником функцій із гіперпараметром є ELU (Exponential Linear Unit), запропонована для прискорення та стабілізації навчання глибоких нейронних мереж. На відміну від Leaky ReLU, що використовує лінійну негативну область, ELU вводить гладку експоненціальну залежність для від'ємних значень аргументу. Її математичний вигляд описується співвідношенням[24]:

$$f(x) = \{x, x \geq 0; \alpha(e^x - 1), x < 0\} \quad (3.9)$$

де параметр (α) зазвичай фіксується на рівні (1.0). Для додатної області ELU поводить ся аналогічно ReLU, забезпечуючи лінійне проходження інформації без насичення. Проте у від'ємній області функція формує плавний експоненціальний перехід, який поступово наближається до нижньої асимптоти.

Однією з ключових переваг ELU(рис 3.9) є здатність зміщувати середнє значення активацій ближче до нуля. У класичному ReLU всі негативні значення повністю відкидаються, через що середнє значення активацій стає додатним. Це створює зміщення розподілу ознак між шарами та може ускладнювати процес оптимізації. ELU частково компенсує цю проблему, оскільки негативна область не занулюється повністю, а формує плавний обмежений сигнал. У результаті градієнтний спуск стає стабільнішим, а швидкість збіжності моделі зростає.

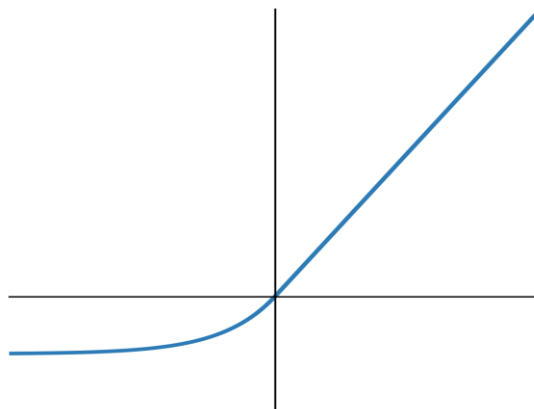


Рисунок 3.9 — ELU

Гладкість ELU також позитивно впливає на геометрію поверхні втрат. На відміну від кусочно-лінійних функцій, ELU має неперервну похідну у точці

переходу між позитивною та негативною областями. Це зменшує різкі зміни градієнта та сприяє формуванню більш плавної траєкторії оптимізації. У надглибоких архітектурах така властивість дозволяє покращити стабільність навчання та зменшити ризик нестабільних коливань параметрів на ранніх етапах тренування.

Практичні дослідження показали, що ELU здатна забезпечувати дуже високу репрезентативну здатність нейронних мереж та у деяких випадках демонструвати результати, співставні або навіть кращі за новіші гладкі нелінійності. Особливо помітним це стало у складних згорткових архітектурах великої глибини, де стабільність проходження градієнта відіграє критичну роль у формуванні багаторівневих просторових представлень.

Водночас використання експоненціальної залежності створює певні обчислювальні труднощі. Порівняно з простими кусочно-лінійними функціями ELU є більш ресурсоємною, а на ранніх етапах оптимізації може демонструвати нестабільність градієнтів через високу чутливість експоненціальної області до зміни ваг. Особливо це проявляється при неякісній ініціалізації параметрів або надмірно великих швидкостях навчання.

У задачах медичної сегментації функції із гіперпараметрами є важливими через необхідність підтримання стабільної передачі слабких градієнтів у складних багат шарових архітектурах. Томографічні зображення характеризуються низьким контрастом, високим рівнем шуму та складною текстурною структурою, тому повне занулення негативної області може призводити до втрати важливої локальної інформації. Leaky ReLU та ELU дозволяють частково зберігати ці сигнали та підтримувати стабільну динаміку навчання навіть у дуже глибоких мережах.

Таким чином, функції активації із фіксованими гіперпараметрами формують проміжний підхід між простими кусочно-лінійними активаціями та повністю адаптивними нелінійностями. Вони поєднують обчислювальну ефективність, стабільність проходження градієнта та здатність зберігати інформацію у від'ємній області, що робить їх важливим компонентом сучасних архітектур комп'ютерного зору та систем медичної сегментації.

РОЗДІЛ 4. АЛГОРИТМІЧНЕ ЗАБЕЗПЕЧЕННЯ ТА СТРАТЕГІЇ АУГМЕНТАЦІЇ ДАНИХ

4.1 Геометричні та просторові перетворення

Методи та перетворення, систематизовані у цьому розділі, репрезентують комплексну методологічну стратегію підготовки та стохастичної регуляризації простору даних. Перехід від базових геометричних трансформацій до складніших нелінійних викривлень реалізується в роботі як процес послідовної побудови власного алгоритмічного забезпечення. Це формує необхідний інструментарій для штучного розширення інформаційної ємності навчальної вибірки та досягнення топологічної інваріантності сегментаційної моделі.

Геометричні перетворення займають центральне місце серед методів аугментації медичних зображень, оскільки саме вони дозволяють навчити модель інваріантності до просторових змін, що є критично важливим для точного розпізнавання та сегментації внутрішніх структур органів. У задачі сегментації сегментів Куіно, де межі між сегментами часто є нечіткими та складними для розпізнавання, здатність мережі правильно ідентифікувати сегмент незалежно від його орієнтації або масштабу значно підвищує загальну точність прогнозів. Геометричні трансформації включають такі базові операції, як повороти, відображення, масштабування, зсуви та випадкові криволінійні деформації. Кожен із цих прийомів додає новий варіант вхідних даних, що змушує мережу будувати більш загальні, стійкі представлення ознак, замість того, щоб навчатися специфічних патернів, властивих лише конкретному положенню печінки у наборі навчальних зображень.

Особливо важливим є застосування геометричних перетворень у 3D-даних КТ, де анатомічні структури можуть значно відрізнятися між пацієнтами не лише в площині кадру, а й у тривимірному просторі. Завдяки цьому мережа навчається узагальнювати закономірності, що визначають межі сегментів, незалежно від того, наскільки печінка зміщена або повернена у конкретному томографічному зрізі. Такі підходи підтверджують Soler et al. [20], які відзначають, що геометричні аугментації є одним із найефективніших способів підвищити стійкість моделей до природної варіативності анатомії та до дрібних помилок позиціонування під час сканування.

Таким чином, застосування геометричних перетворень у контексті сегментації сегментів Куіно дозволяє моделі U-Net/ResNet навчитися виділяти важливі ознаки, що визначають сегменти, у будь-якому положенні та орієнтації печінки, підвищуючи стабільність та точність результатів без необхідності збільшувати кількість навчальних зразків шляхом ручної розмітки. Це робить геометричні аугментації невід'ємною частиною

алгоритмічного забезпечення підготовки медичних зображень для задачі багатокласової сегментації.

Афінні перетворення, включаючи обертання(рис 4.1) та зміщення, є потужним інструментом для навчання нейронних мереж інваріантності до просторових змін, що є особливо актуальним для медичних зображень, зокрема КТ печінки. На відміну від простих геометричних трансформацій, афінні перетворення зберігають прямолінійність та паралельність ліній, що дозволяє моделювати природну мінливість положення органу всередині черевної порожнини без спотворення загальної структури сегментів. Обертання зображення разом із його істинною маскою на певний кут імітує різні кути сканування або фізіологічну варіативність орієнтації печінки у тілі пацієнта, що, за даними [25], значно підвищує стійкість моделі до відхилень, особливо при роботі з 3D-об'ємами, які часто розглядаються як послідовність 2D-зрізів.

Зміщення або трансляція зображення вздовж осей X та Y моделює незначні відмінності в центровці пацієнта під час сканування, що дозволяє мережі навчитися коректно сегментувати орган незалежно від точного розташування печінки у кадрі. Хоча це перетворення є досить простим, його ефект у комбінації з обертанням забезпечує значне підвищення загальної інваріантності моделі, зменшуючи ймовірність помилок на краях сегментів або при визначенні меж Куіно.

Ключовим моментом застосування афінних перетворень є синхронність їхнього впливу на вхідне зображення та відповідну багатокласову маску сегментації. Це гарантує, що геометричні співвідношення між сегментами зберігаються, і модель отримує правильний сигнал під час навчання, що є критично важливим для досягнення високої точності багатокласової сегментації. Таким чином, афінні перетворення служать не лише інструментом збільшення різноманітності навчальних даних, а й фундаментальним механізмом забезпечення стійкості та надійності сегментаційної моделі.

Віддзеркалення (Flip) є спеціалізованим видом афінних перетворень, при якому зображення разом із відповідною багатокласовою маскою сегментації обертаються на 180 градусів навколо горизонтальної або вертикальної осі. Цей метод має особливу цінність для навчання нейронних мереж, оскільки дозволяє створювати нові навчальні приклади, які залишаються анатомічно правдоподібними та не порушують структурну цілісність органу. Горизонтальне віддзеркалення уздовж вертикальної осі моделює ситуації, коли анатомія пацієнта виглядає дзеркально або зображення отримано в іншому порядку, що дозволяє мережі навчитися інваріантності до симетрії. Незважаючи на природну асиметрію печінки, використання горизонтального фліпу сприяє більш надійному розпізнаванню сегментів Куіно незалежно від орієнтації.

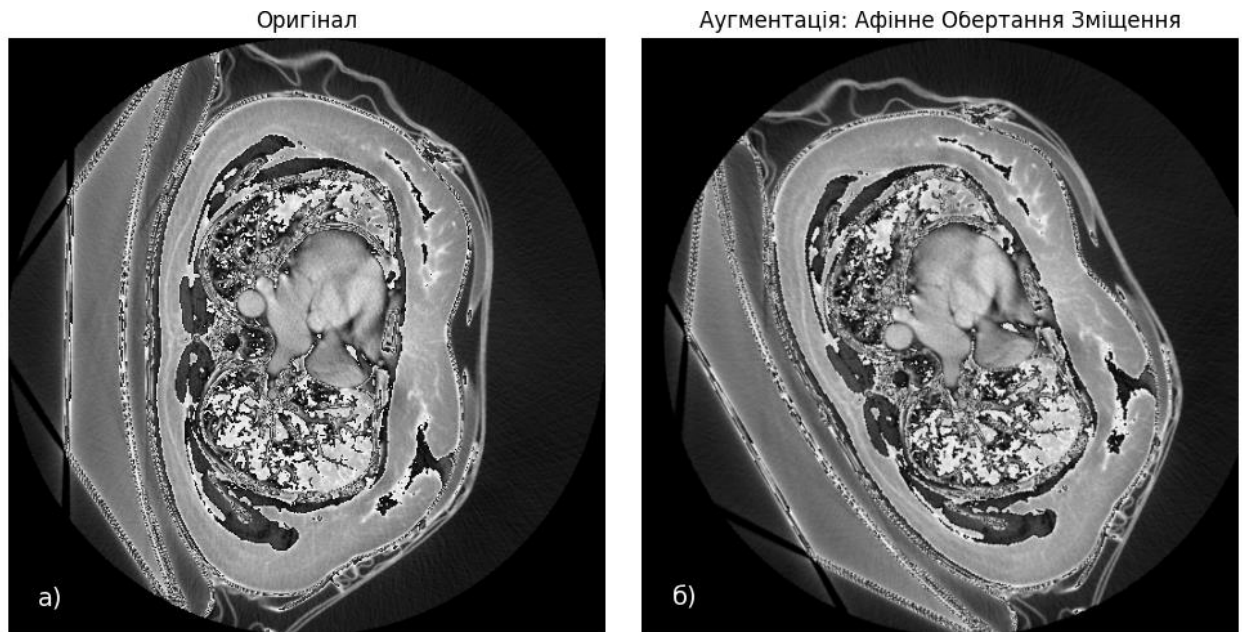


Рисунок 4.1 — Приклад афінних перетворень(обертання):
а) оригінал, б) афінне обертання зображення

Вертикальне віддзеркалення вздовж горизонтальної осі відтворює ситуації, коли зображення перевернуто через особливості сканування або попередньої обробки, що дозволяє моделі адаптуватися до менш типових орієнтацій. Хоча таке перетворення рідше зустрічається в реальних клінічних даних, його застосування разом із горизонтальним фліпом(рис 4.2) та іншими геометричними перетвореннями [26] значно розширює варіативність навчальної вибірки.

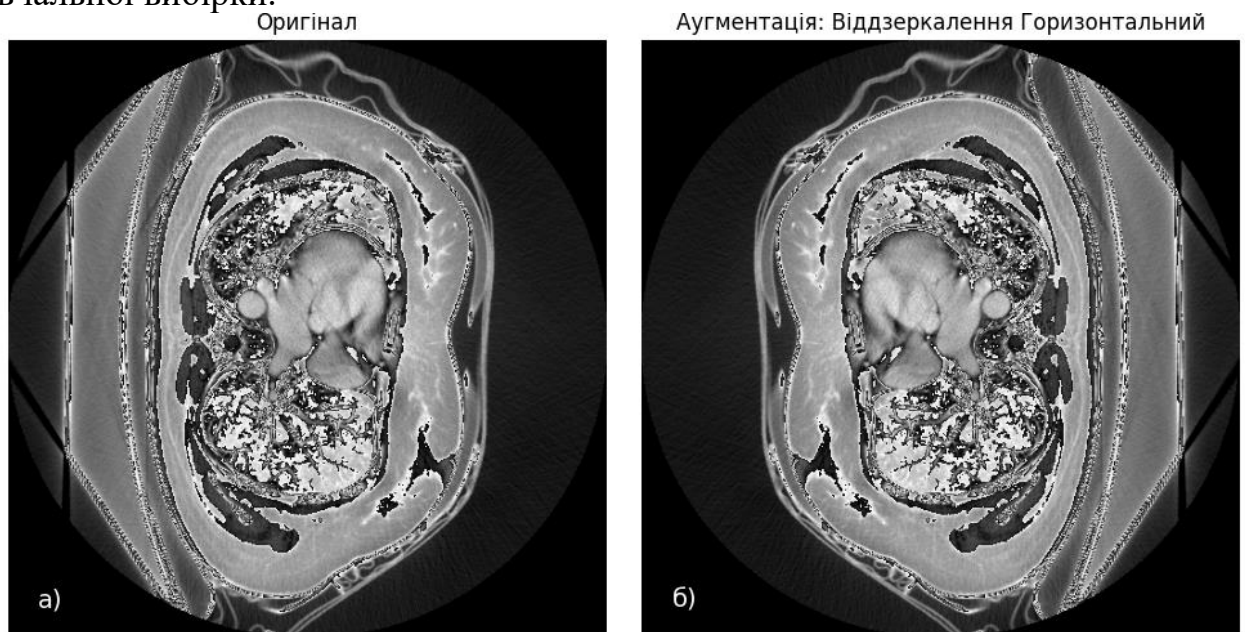


Рисунок 4.2 — Приклад горизонтального віддзеркалення:
а) оригінал, б) горизонтально віддзеркалене зображення

Завдяки віддзеркаленню модель отримує змогу краще узагальнювати знання про анатомічні структури печінки та сегменти Куїно, що особливо важливо при обмеженій кількості доступних зразків у наборі даних 3D-IRCADb-01. Поєднання різних типів фліпів із іншими методами аугментації забезпечує значне збагачення навчальної вибірки, підвищуючи стійкість і точність сегментаційної моделі без введення анатомічно неприродних спотворень.

Масштабування (Scaling)(рис 4.3) є важливим інструментом аугментації, який змінює розмір об'єкта на зображенні, зберігаючи його пропорції. У контексті медичної візуалізації цей метод моделює природну варіативність розмірів печінки серед різних пацієнтів. Застосування масштабування, як у бік збільшення, так і зменшення, дозволяє моделі U-Net/ResNet34 навчитися коректно сегментувати сегменти Куїно незалежно від абсолютних розмірів органу на вхідному знімку. Такий підхід підвищує стійкість моделі до природних біологічних відмінностей та зміни масштабу, які часто зустрічаються у реальних клінічних даних.

Перспективні зміщення (Perspective Transformations) є більш складними геометричними перетвореннями порівняно зі стандартними афінними. Вони моделюють зміну точки огляду сканера, що призводить до нерівномірних спотворень зображення: паралельні лінії можуть зближуватися або сходитися, а пропорції об'єктів змінюватися залежно від їхньої відстані від уявного фокуса. Незважаючи на те, що у типовому КТ-скануванні ці ефекти не настільки виражені, використання таких перетворень допомагає моделі адаптуватися до нелінійних геометричних варіацій, які можуть виникати через неідеальне позиціонування пацієнта або специфіку реконструкції 3D-об'єму [27].

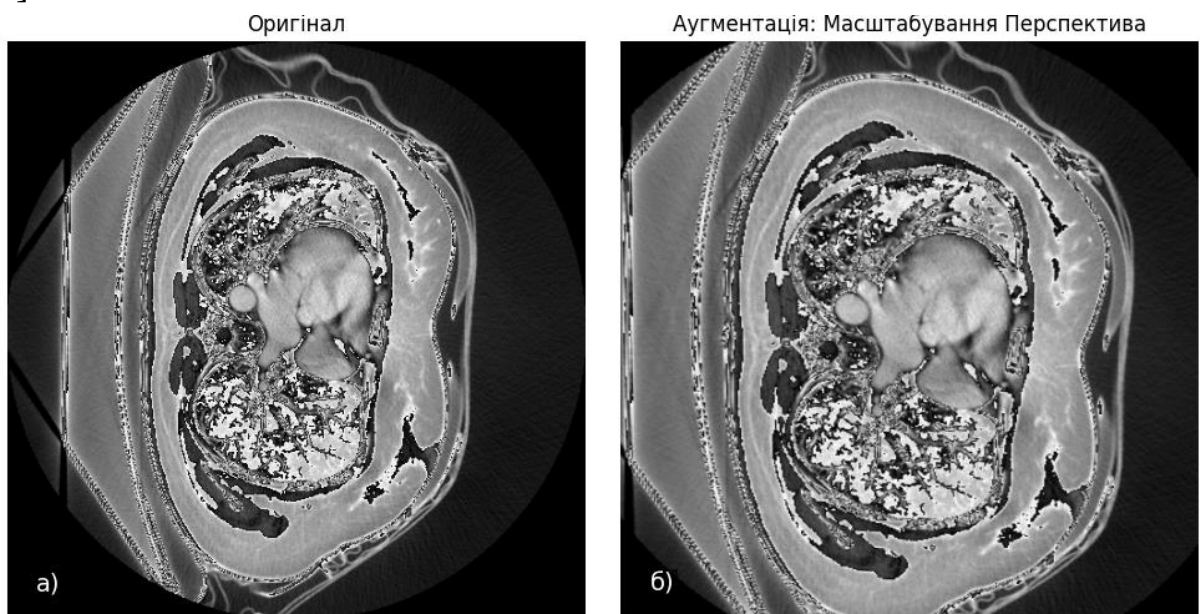


Рисунок 4.3 — Приклад масштабування: а) оригінал, б) масштабоване зображення

Комбіноване використання масштабування та перспективних зміщень дозволяє синтетично збільшити розмаїття навчальних прикладів, що сприяє кращому узагальненню моделі на непередбачені варіації у тестових даних. Це особливо важливо для багатокласової сегментації сегментів Куїно, де навіть невеликі зміни розміру або перспективи можуть впливати на точність визначення меж між сегментами. Завдяки цим методам модель отримує змогу ефективніше адаптуватися до реалістичних варіацій клінічних зображень, підвищуючи її стабільність та надійність.

4.2 Фотометричні трансформації

Фотометричні перетворення змінюють не геометрію зображення, а його інтенсивні характеристики, такі як колір, яскравість, контраст або інші параметри пікселів. Основна мета цих трансформацій полягає у навчанні моделі стійкості до природної варіабельності зображень, що виникає через відмінності у налаштуваннях КТ-сканера, різні протоколи сканування, індивідуальне дозування контрастної речовини, а також під час процесів нормалізації та попередньої обробки даних. У клінічних КТ-знімках ці фактори можуть суттєво змінювати інтенсивність пікселів, наприклад, під час портальної венозної фази, однак вони не впливають на анатомічну форму органів.

Завдяки застосуванню фотометричних перетворень модель навчається ігнорувати такі варіації, концентруючись на суттєвих ознаках, що визначають межі сегментів Куїно. Це дозволяє уникнути помилок, пов'язаних із змінами інтенсивності, і підвищує узагальнюючу здатність нейронної мережі при роботі з різними клінічними даними, знижуючи чутливість до технічних і процедурних відмінностей [28].

Регулювання яскравості та контрасту (рис 4.4) є одними з найважливіших фотометричних перетворень у контексті медичної сегментації, оскільки дозволяють моделі навчитися ігнорувати варіації, що виникають не через анатомічні відмінності, а через технічні особливості сканування. Зміна яскравості полягає у додаванні або відніманні сталої величини до інтенсивності пікселів зображення, що імітує різні рівні вікнавання КТ або невеликі коливання параметрів сканера під час процедур. Така аугментація дозволяє мережі не орієнтуватися на абсолютне значення інтенсивності, а фокусуватися на структурних ознаках тканин.

Зміна контрасту, яка реалізується множенням інтенсивності пікселів на певний коефіцієнт, є критичною для виявлення деталей м'яких тканин, а особливо для підкреслення судинного русла та меж сегментів Куїно, які часто є нечіткими. За рахунок цього модель краще диференціює сусідні сегменти печінки, що значно підвищує точність багатокласової сегментації.

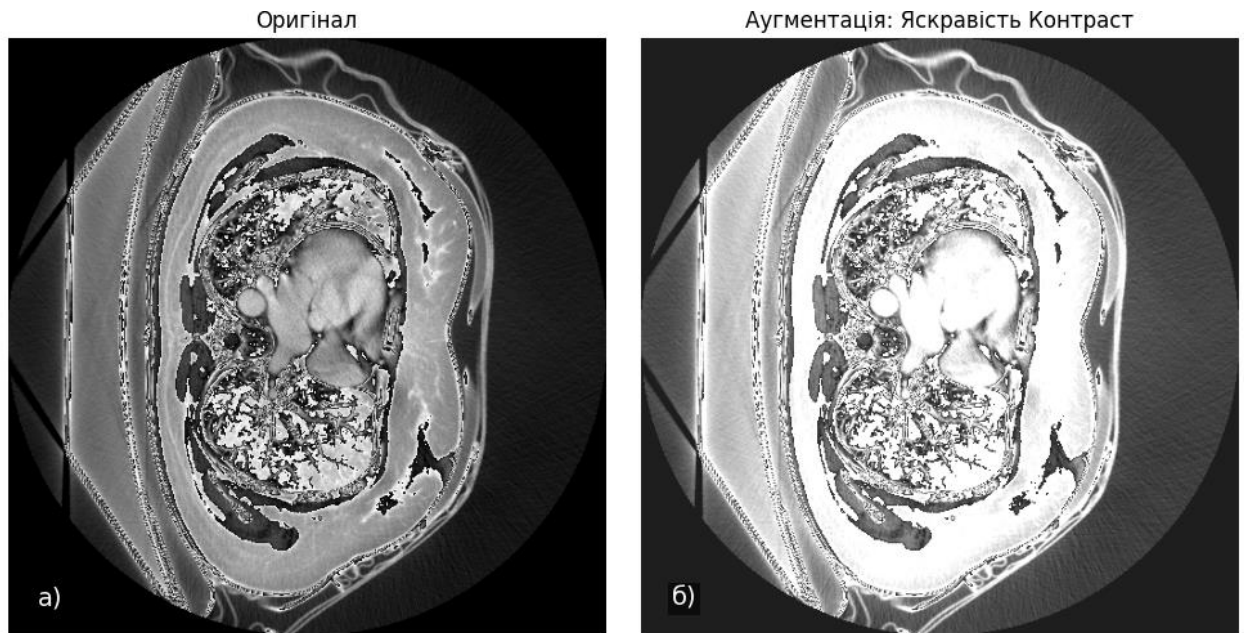


Рисунок 4.4 — Приклад додавання яскравості і контрасту на зображення: а) оригінал, б) зображення із доданою яскравістю і контрастом

Використання цих фотометричних аугментацій у контрольованих межах дозволяє уникнути перенавчання на специфічні параметри навчальної вибірки 3D-IRCADb-01, одночасно підвищуючи стійкість моделі до змін у клінічних знімках і покращуючи її узагальнюючу здатність при роботі з новими, невідомими даними. Такий підхід забезпечує більш надійну та стабільну роботу U-Net/ResNet у реальних сценаріях, коли параметри КТ-сканування та контрастування можуть суттєво варіюватися від пацієнта до пацієнта [29].

Адаптивне вирівнювання гістограми з обмеженням контрасту (CLAHE, Contrast Limited Adaptive Histogram Equalization)(рис 4.5) є більш досконалим фотометричним перетворенням порівняно з глобальним регулюванням контрасту і має особливе значення для обробки медичних зображень [30]. На відміну від традиційного вирівнювання гістограми, яке впливає на весь кадр одночасно, CLAHE працює локально, підвищуючи контраст у невеликих незалежних областях зображення. Такий підхід дозволяє виділяти деталі, які можуть загубитися при глобальних трансформаціях, особливо у зонах зі слабкою інтенсивністю або неоднорідним освітленням.

Механізм роботи CLAHE передбачає розділення зображення на невеликі прямокутні тайли, у межах яких обчислюється гістограма і вирівнюється локально. Для запобігання надмірному посиленню шуму використовується обмеження коефіцієнта підсилення контрасту — так званий контраст-кліпінг. Після обробки всіх тайлів кінцеве зображення відновлюється за допомогою білінійної інтерполяції, що усуває штучні межі та робить перехід між областями плавним і природним.

У контексті сегментації печінки метод CLANE дозволяє локально підкреслити судинне русло, яке формує неявні межі сегментів Куїно. Оскільки якість КТ-знімків у наборі даних 3D-IRCADb-01 може бути неоднорідною, CLANE забезпечує, що модель U-Net/ResNet отримує оптимально контрастовані зображення незалежно від початкових локальних характеристик. Це сприяє кращому виділенню ключових ознак енкодером, підвищуючи точність та стабільність багатокласової сегментації навіть у випадках слабкої або нерівномірної контрастності вхідних зображень.

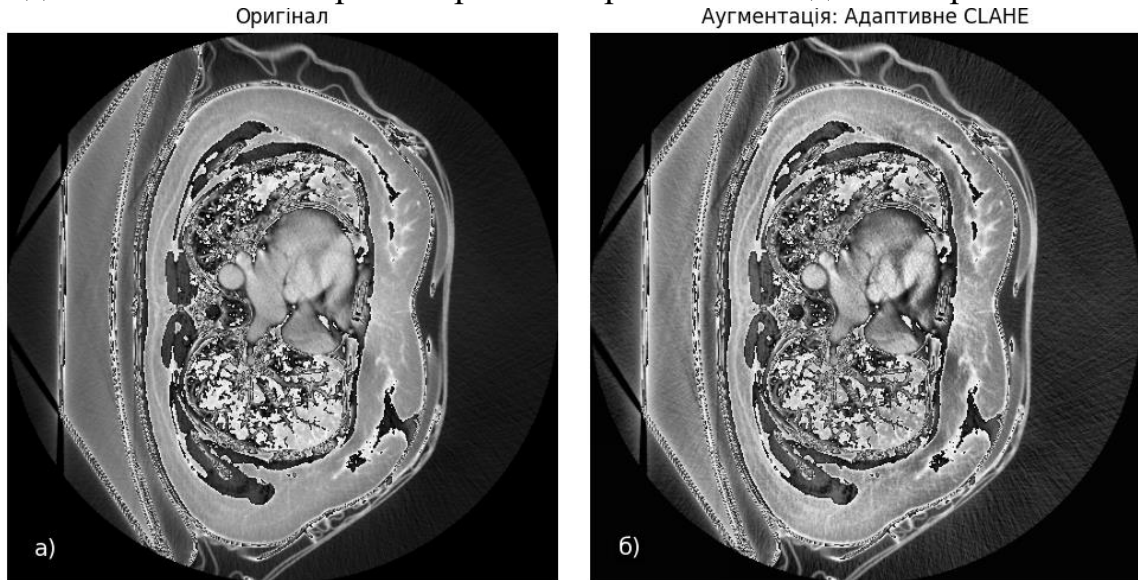


Рисунок 4.5 — Приклад додавання адаптивного CLANE:
а) оригінал, б) Додавання адаптивного CLANE до зображення

4.3 Моделювання нелінійних деформацій та артефактів

Цей тип аугментацій відіграє важливу роль у формуванні стійкості та надійності моделі, оскільки дозволяє навчати її реагувати на природні варіації, властиві медичним зображенням. Окрім просторових змін, як у геометричних перетвореннях, ці методи імітують складні нелінійні деформації анатомічних структур та статистичні артефакти, зокрема різноманітні види шуму, які виникають під час отримання КТ-зображень. Це дозволяє моделі ефективно працювати навіть у присутності природних спотворень, покращуючи узагальнюваність і стабільність результатів сегментації.

Еластичні деформації (Elastic Deformations)(рис 4.6) є одним із ключових і надзвичайно ефективних методів аугментації для медичної сегментації, особливо при роботі з органами, що складаються з м'яких тканин, такими як печінка. На відміну від лінійних афінних перетворень, які зберігають пропорції та кути, еластичні деформації вводять локальні нелінійні спотворення, що імітують природні анатомічні варіації та рухи тканин *in vivo*[31]. Це включає такі фізіологічні явища, як дихальна активність, перистальтика, зміни тиску в черевній порожнині або індивідуальні анатомічні

відмінності між пацієнтами, які призводять до зміни форми печінки та її сегментів. Завдяки цьому мережа навчиться розпізнавати сегменти не тільки за точними контурами, а за загальними структурними характеристиками, що значно підвищує її здатність до узагальнення.

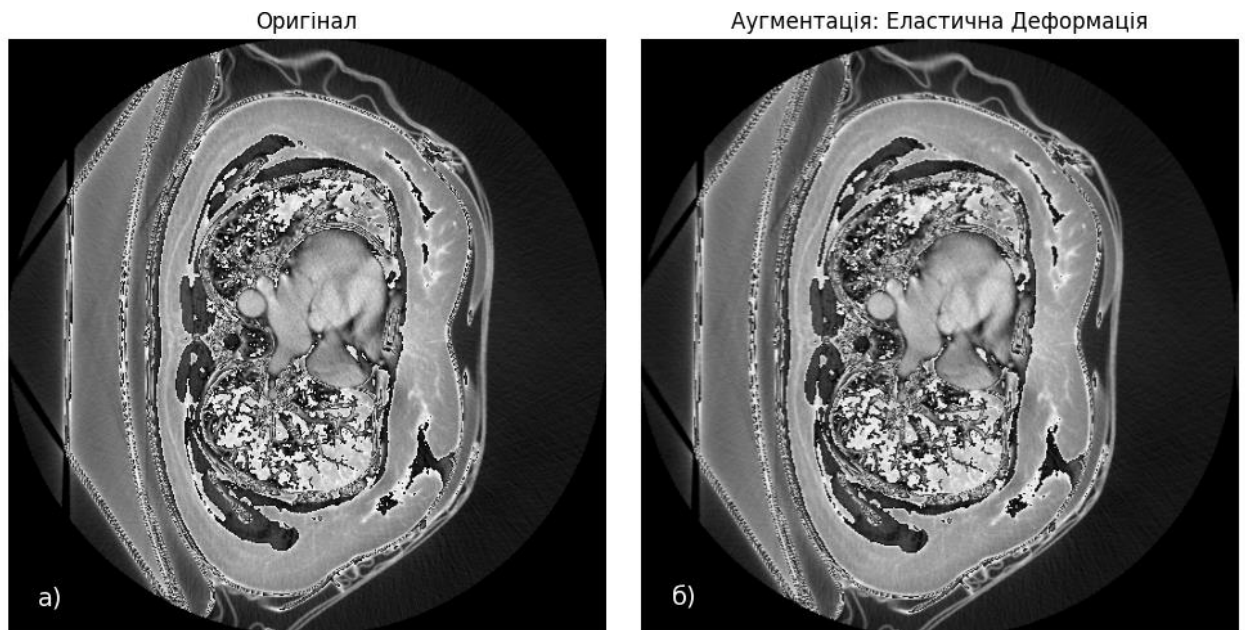


Рисунок 4.6 — Приклад еластичної деформації:
а) оригінал, б) зображення з доданою еластичною деформацією

Технічна реалізація полягає у генерації випадкового поля зсувів (displacement field), яке визначає, як кожен піксель або група пікселів зміщується у просторі. Для забезпечення реалістичності та плавності цих деформацій поле зсувів згладжується за допомогою гаусового фільтра. Це гарантує, що деформації виглядають природними і не створюють різких або неприродних артефактів, які могли б ввести модель в оману.

У контексті сегментації сегментів Куїно застосування еластичних деформацій має особливу цінність. Оскільки сегменти часто мають неявні межі і межі сусідніх сегментів можуть бути нечіткими, модель U-Net/ResNet повинна навчитися виділяти їх за структурними особливостями, а не тільки за точними контурами, присутніми у навчальній вибірці. Таке навчання робить модель більш стійкою до варіацій у формі та розташуванні сегментів, дозволяючи їй ефективно працювати на нових клінічних даних, де невеликі деформації є нормою.

Крім того, еластичні деформації підвищують загальну різноманітність навчальної вибірки без необхідності додаткового збору даних. Це дозволяє моделі набувати більш гнучкі представлення анатомії печінки, зменшуючи ризик перенавчання на специфічні форми чи позиції органу. В результаті, навіть при наявності обмеженого набору КТ-зрізів, нейронна мережа здатна навчатися розпізнавати сегменти більш надійно та стабільно, що є критичним

для застосувань у клінічних умовах, де точність і узагальнюваність моделі мають першочергове значення.

Еластичні деформації, разом з іншими методами аугментації, такими як масштабування, обертання, фліпи та фотометричні перетворення, створюють багатовимірний простір варіацій, у якому модель навчається виділяти ключові ознаки сегментів печінки, залишаючись стійкою до шумів, змін інтенсивності та фізіологічних відмінностей між пацієнтами. Це робить їх невід’ємною складовою сучасного підходу до підготовки даних у медичній сегментації, забезпечуючи баланс між реалістичністю даних та різноманітністю навчальної вибірки.

Додавання гаусового шуму (Gaussian Noise)(рис 4.7) є одним із ключових методів аугментації для медичних зображень і належить до фотометричних перетворень, що моделюють специфічні статистичні артефакти даних. На відміну від зміни яскравості, яка вводить систематичне зміщення інтенсивності пікселів, гаусовий шум відтворює випадкові, незалежні коливання, що характерні для реальних КТ- і МРТ-зображень. Ці коливання виникають через фізичні обмеження сканера, електронні перешкоди або статистичні властивості фотонів, що надходять у детектор, і можуть бути адекватно змодельовані гаусовим розподілом.

Механізм додавання шуму полягає у вибірці випадкових значень з нульовим середнім та заданою дисперсією, які потім додаються до кожного пікселя оригінального зображення. Це дозволяє створювати навчальні приклади, де анатомічні структури залишаються впізнаваними, але присутні варіації інтенсивності, подібні до тих, що зустрічаються у реальних клінічних даних.

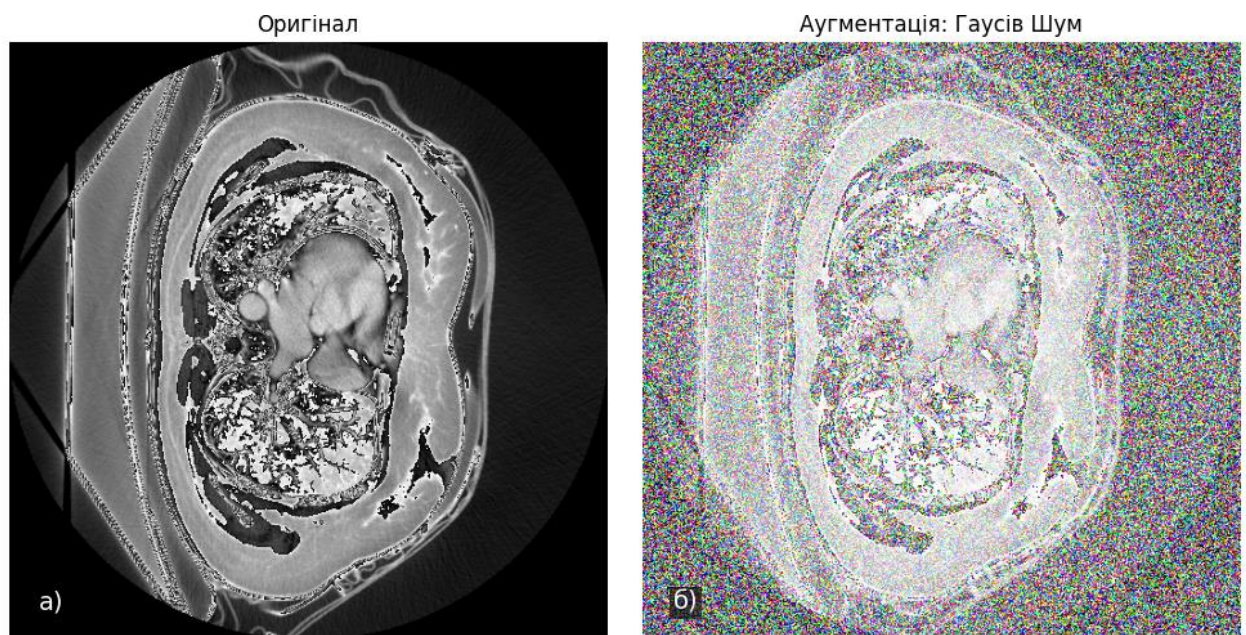


Рисунок 4.7 — Приклад гаусового шуму:
а) оригінал, б) зображення з гаусовим шумом

Таке перетворення має кілька ключових переваг. По-перше, воно підвищує стійкість моделі U-Net/ResNet до шумів і артефактів, з якими вона зіткнеться у тестових та клінічних умовах. По-друге, введення контрольованого шуму слугує ефективною формою регуляризації, запобігаючи перенавчанню на ідеалізовані або надто чисті дані. По-третє, додавання гаусового шуму допомагає моделі навчитися відокремлювати значущі анатомічні ознаки, такі як межі сегментів Куіно, від випадкових флуктуацій інтенсивності, що забезпечує більш надійне та стабільне прогнозування у реальних клінічних сценаріях.

Технічна реалізація цього процесу для кожного пікселя $I(x, y)$ зображення може бути описана формулою:

$$I_{noisy}(x, y) = I(x, y) + N(0, \sigma^2) \quad (4.1)$$

де $I_{noisy}(x, y)$ — значення пікселя після додавання шуму, $I(x, y)$ — оригінальне значення пікселя, а $N(0, \sigma^2)$ — випадкова величина, згенерована з гаусового розподілу з нульовим середнім і дисперсією σ^2 .

Завдяки цій простій, але потужній трансформації, навчальна вибірка значно збагачується, а модель стає більш гнучкою і здатною до узагальнення на різноманітні умови знімання, включаючи реальні варіації у якості КТ-зображень.

РОЗДІЛ 5. ЧИСЛОВІ ЕКСПЕРИМЕНТИ ТА АНАЛІЗ РЕЗУЛЬТАТІВ

5.1 Дослідження впливу функцій активації на навчання моделі

Для оцінки впливу різних функцій активації на ефективність сегментації печінки та її анатомічних сегментів було проведено серію контрольованих експериментів із використанням архітектури U-Net на базі енкодера ResNet34. Навчання виконувалося на КТ-зрізах набору даних 3DIRCADb-01 у задачі багатокласової сегментації сегментів Куіно. З метою забезпечення коректності порівняльного аналізу всі конфігурації тестувалися в однакових умовах: використовувалися ідентичні параметри ініціалізації ваг, однаковий розмір пакету, незмінна схема аугментації даних та фіксована кількість епох навчання. Оптимізація параметрів здійснювалася за допомогою алгоритму AdamW, який реалізує принцип від'єданого згасання вагових коефіцієнтів (decoupled weight decay), що дозволяє ефективно стабілізувати процес мінімізації функції втрат та зменшити ризик перенавчання моделі під час тривалого тренування.

Основною метрикою оцінювання виступав коефіцієнт Intersection over Union (IoU), який використовувався як для загальної області печінки, так і для окремих функціональних сегментів. Використання саме цієї метрики є особливо доцільним у задачах медичної сегментації, оскільки IoU безпосередньо характеризує ступінь просторового перекриття між прогнозованою та еталонною маскою. У контексті багатокласової сегментації анатомічних структур це дозволяє оцінювати не лише загальну точність локалізації органу, а й здатність моделі коректно відтворювати складну внутрішню геометрію сегментів Куіно.

Теоретичною передумовою використання модифікованих функцій активації є їхня здатність забезпечувати проходження градієнта у від'ємній області простору ознак. У класичному ReLU локальна похідна функції дорівнює нулю для всіх від'ємних значень аргументу.

Через це частина нейронів може втрачати здатність до оновлення параметрів під час зворотного поширення помилки, переходячи у стан так званих «мертвих нейронів». Саме тому функції LeakyReLU, ELU, Mish, SiLU та PReLU теоретично розглядаються як механізми компенсації цієї проблеми. Їхня головна перевага полягає у збереженні ненульового градієнта навіть у від'ємній області, що дозволяє підтримувати безперервне оновлення ваг та покращувати стабільність оптимізації.

Проте результати експериментального аналізу показали, що на практиці використання складніших функцій активації не призводить до суттєвого покращення кінцевої якості сегментації. Динаміка валідаційних метрик демонструє, що всі протестовані функції досягають близьких значень IoU (рис 5.1 - 5.4), а переваги модифікованих нелінійностей мають переважно локальний або епізодичний характер. Це свідчить про те, що сама архітектура

U-Net у поєднанні з енкодером ResNet34 вже забезпечує достатньо стабільне проходження інформації та ефективне формування багаторівневих ознак навіть при використанні стандартної ReLU.

Одним із ключових факторів, що пояснює таку поведінку, є наявність шарів пакетної нормалізації (Batch Normalization) в енкодері ResNet34. Механізм пакетної нормалізації автоматично стабілізує статистичний розподіл активацій між шарами, зміщуючи їх у область інформативних значень та зменшуючи ризик переходу великої кількості нейронів у повністю неактивний стан. У результаті навіть стандартний ReLU здатний забезпечувати стабільне проходження градієнта без критичної втрати інформації. Фактично глибокі згорткові фільтри самостійно формують представлення ознак таким чином, щоб найбільш значущі семантичні структури кодувалися у додатній області простору активацій.

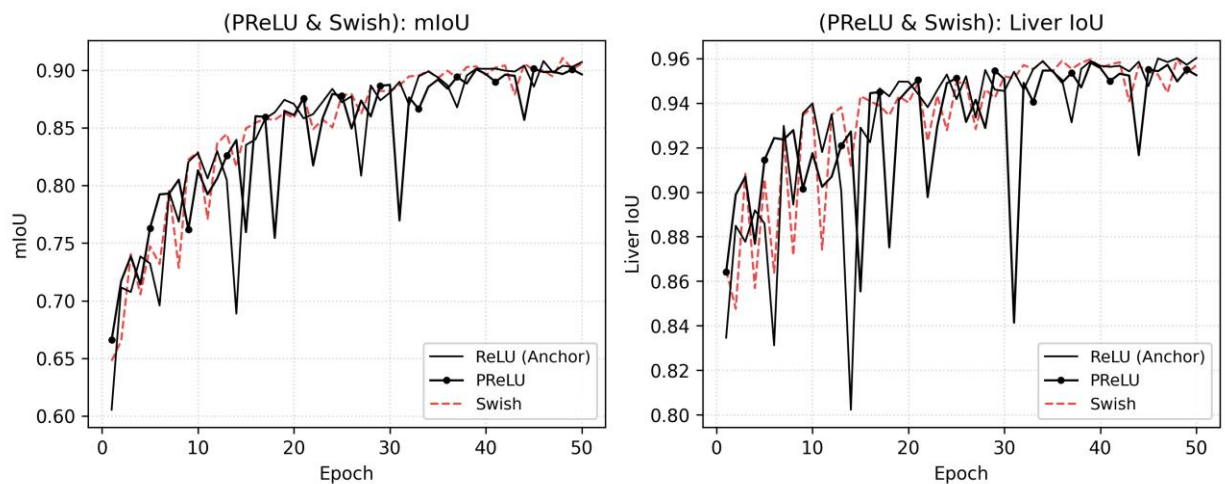


Рисунок 5.1 - Метрики IoU для PReLU та Swish в порівнянні з ReLU

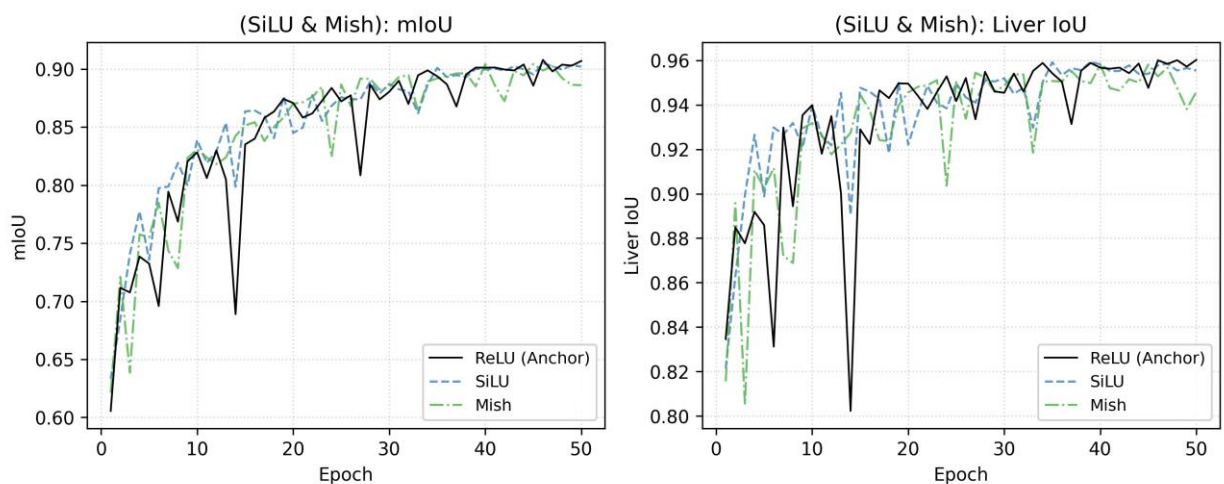


Рисунок 5.2 - Метрики IoU для SiLU та Mish в порівнянні з ReLU

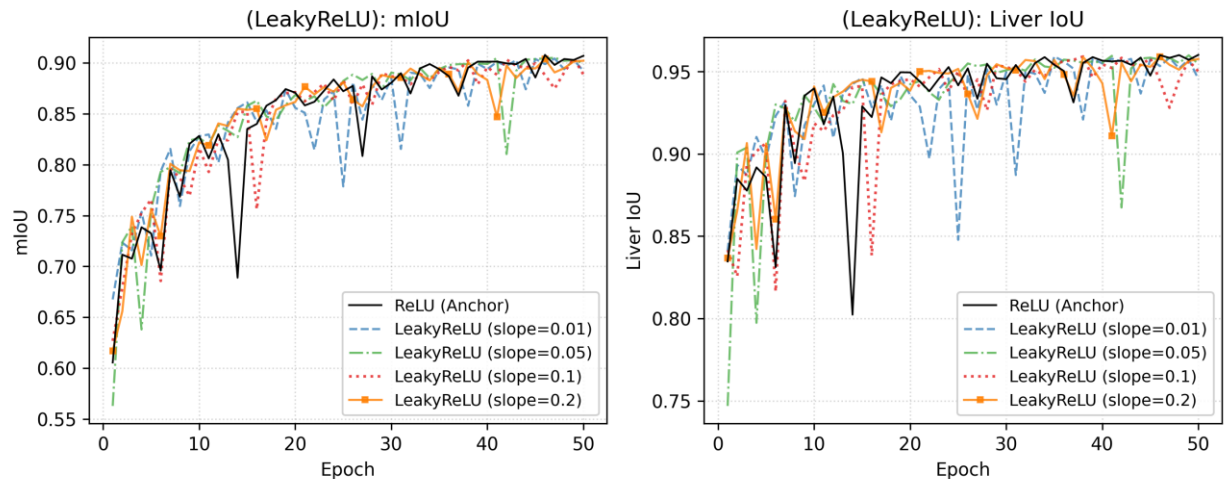


Рисунок 5.3 - Метрики IoU для LeakyReLU з різними варіаціями гіперпараметру в порівнянні з ReLU

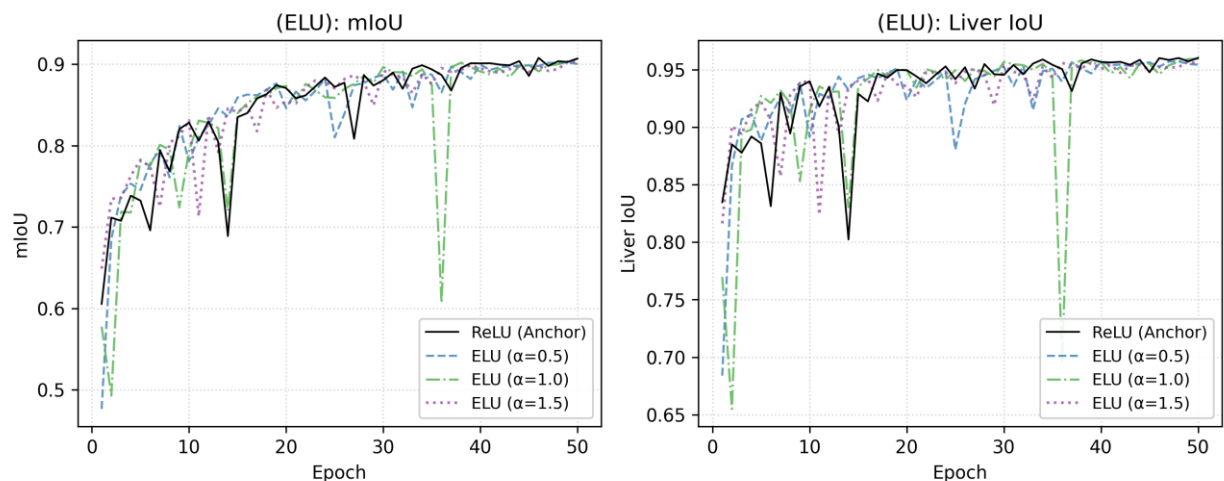


Рисунок 5.4 - Метрики IoU для ELU з різними варіаціями гіперпараметру в порівнянні з ReLU

Додатково слід враховувати специфіку медичних зображень. На КТ-зрізах печінки основна інформація про межі органу та внутрішні структури представлена локальними текстурними переходами та просторовими патернами, які ефективно виділяються навіть простими кусково-лінійними нелінійностями. Це означає, що теоретичні переваги гладких або немонотонних функцій активації можуть виявлятися недостатньо значущими порівняно з впливом самої архітектури мережі, механізмів нормалізації та якості аугментації даних.

Аналіз часових характеристик навчання підтверджує математично очікувану залежність між складністю функції активації та обчислювальними витратами. Найбільший час виконання однієї епохи спостерігався для функцій із параметрами, що навчаються, зокрема LearnableSwish та PReLU (рис 5.5). Це пояснюється необхідністю додаткового обчислення градієнтів за параметрами

активацій. У такому випадку кожна ітерація градієнтного спуску потребує не лише оновлення ваг згорткових шарів, а й окремого обчислення похідних для параметрів самої функції активації. Це відповідно збільшує обчислювальне навантаження та тривалість навчання.

Водночас результати експериментів продемонстрували окремі обчислювальні аномалії. Зокрема, одна з конфігурацій LeakyReLU із коефіцієнтом нахилу $\alpha=0.2$ показала час виконання епохи, що перевищував навіть складніші експоненціальні функції та навіть функції із навченим параметром. Оскільки сама операція множення на скаляр у LeakyReLU є тривіальною з алгоритмічної точки зору, така поведінка найімовірніше пояснюється апаратними флуктуаціями, особливостями диспетчеризації CUDA-ядер або неоптимальною векторизацією обчислень під час конкретного запуску експерименту. Це додатково підкреслює, що практична продуктивність моделей визначається не лише теоретичною складністю функцій активації, а й особливостями програмно-апаратної реалізації.

Окрему увагу привертає поведінка гладких немонотонних функцій, таких як Mish та SiLU. Попри теоретичні переваги, пов'язані з плавністю похідних та стабільнішим ландшафтом функції втрат, їх використання не забезпечило суттєвого приросту точності сегментації. При цьому обчислювальна складність таких функцій є помітно вищою через необхідність обчислення трансцендентних операцій — сигмоїди, експоненти та гіперболічного тангенса. У результаті час тренування збільшується, тоді як практичний виграш у метриках залишається мінімальним.

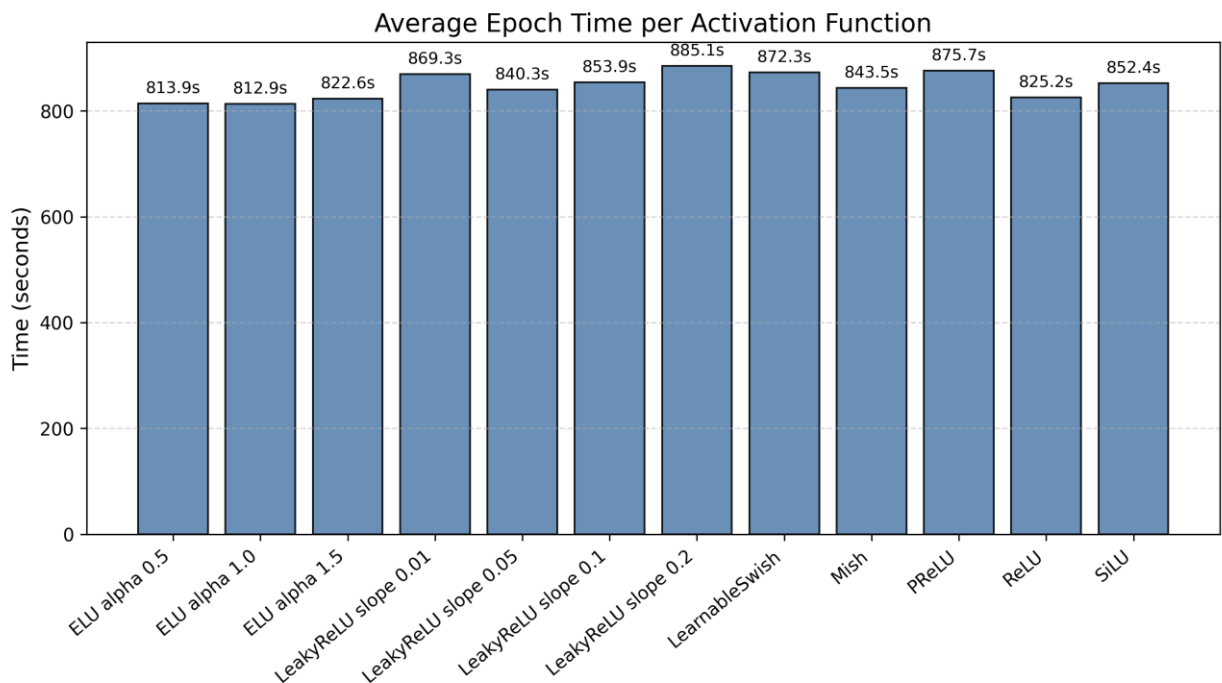


Рисунок 5.5 — Час проходження однієї епохи мережею в залежності від функції активації

Додатково було зафіксовано випадки нестабільної поведінки ELU на пізніх етапах оптимізації. У деяких експериментах спостерігалися різкі локальні падіння IoU, що може бути наслідком високої чутливості експоненціальної області функції до коливань вагових коефіцієнтів. Подібна нестабільність є небажаною у задачах медичної сегментації, де навіть незначні зміни локалізації меж можуть призводити до помітного зниження якості формування анатомічних структур.

Аналізуючи результати експериментального аналізу, можна зробити висновок, що для задач сегментації печінки на КТ-зображеннях використання складних адаптивних або гладких функцій активації не забезпечує достатнього приросту точності, який би компенсував додаткові обчислювальні витрати. Попри численні теоретичні переваги модифікованих нелінійностей, базова функція ReLU демонструє найкращий баланс між швидкістю навчання, стабільністю оптимізації та якістю сегментації.

Функція ReLU залишається оптимальним рішенням для побудови ефективних моделей медичної сегментації завдяки своїй обчислювальній простоті, стабільному проходженню градієнта та достатній репрезентативній здатності для формування складних анатомічних представлень. Її використання мінімізує час навчання та інференсу, зменшує навантаження на графічні процесори та водночас забезпечує високий рівень точності локалізації печінки та її функціональних сегментів на томографічних зрізах.

5.2 Оцінка ефективності стратегій аугментації

Для дослідження впливу стратегій аугментації на здатність моделі до узагальнення було проведено серію контрольованих експериментів із використанням архітектури U-Net з енкодером ResNet34 у задачі багатокласової сегментації сегментів Куїно на КТ-зрізах набору даних 3DIRCADb-01. Усі конфігурації навчалися в однакових умовах, що забезпечувало коректність порівняння отриманих результатів. Оптимізація параметрів здійснювалася за допомогою алгоритму AdamW із механізмом від'єданого згасання вагових коефіцієнтів, який дозволяє стабілізувати процес мінімізації функції втрат та підтримувати ефективну регуляризацию високорозмірного простору параметрів. Навчання проводилося протягом 60 епох із використанням batch size, рівного 4, а для компенсації можливого обрізання анатомічних структур під час геометричних трансформацій застосовувався коефіцієнт попереднього розширення зображення 1.3.

У межах дослідження було сформовано чотири окремі стратегії аугментації, які відрізнялися виключно інтенсивністю та складністю стохастичних перетворень. Базова конфігурація (Level 1) не містила жодних аугментацій і використовувалася як контрольний сценарій для оцінки природної здатності моделі адаптуватися до навчальної вибірки без

зовнішньої регуляризації. Такий підхід дозволяє дослідити поведінку архітектури в умовах статичного простору даних та виявити потенційну схильність до перенавчання.

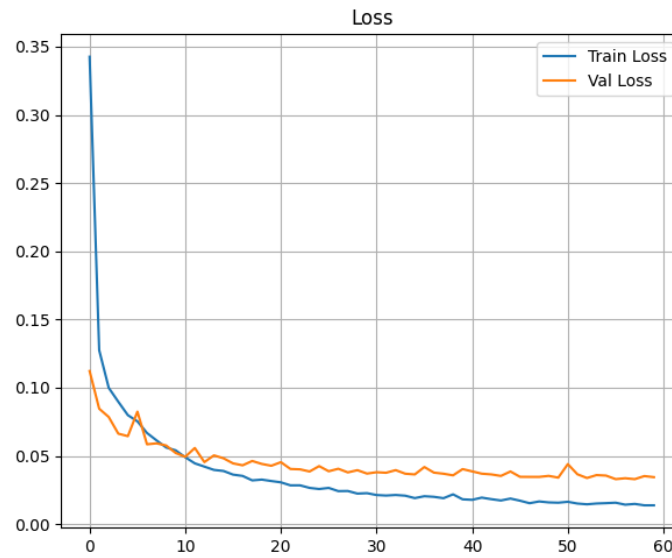


Рисунок 5.6 — Стратегія 1: втрати(y) до епох(x)

Другий рівень аугментації (Level 2) включав мінімальні геометричні та фотометричні збурення. До набору трансформацій входили горизонтальне та вертикальне віддзеркалення, незначні обертання в межах $\pm 5^\circ$, помірна зміна яскравості та контрасту, а також додавання гаусового шуму низької інтенсивності. Подібна конфігурація моделює базову фізіологічну варіативність положення органів і незначні флуктуації параметрів сканування, не порушуючи при цьому глобальну анатомічну структуру медичного зображення.

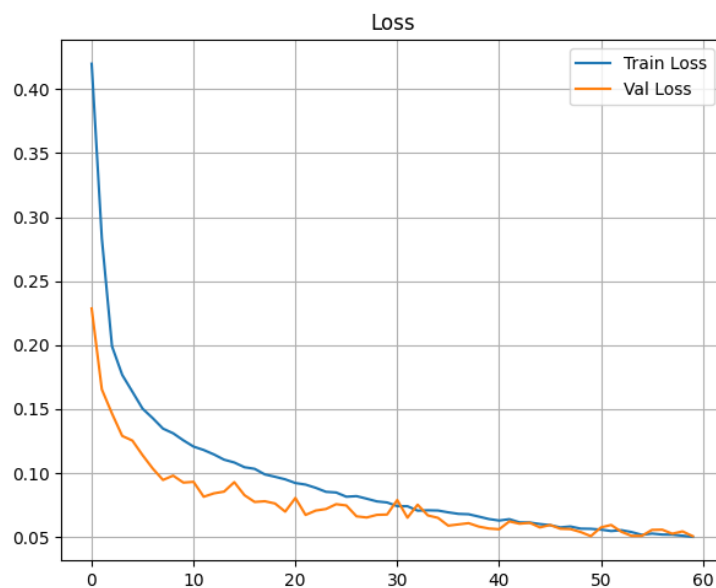


Рисунок 5.7 —

Стратегія 2:

втрати(y) до епох(x)

Третій рівень (Level 3) суттєво розширював простір можливих трансформацій. Діапазон обертання було збільшено до $\pm 15^\circ$, вводилися еластичні деформації м'яких тканин, випадкове масштабування та локальне адаптивне вирівнювання гістограми CLaNE. Такі перетворення моделюють більш складну біомеханічну варіативність внутрішніх органів, включаючи зміни форми паренхіми, локальні неоднорідності контрастування та деформації, пов'язані з диханням або фізіологічними особливостями пацієнта. Важливою особливістю цього рівня є те, що трансформації залишаються анатомічно правдоподібними та не руйнують природну топологію сегментів печінки.

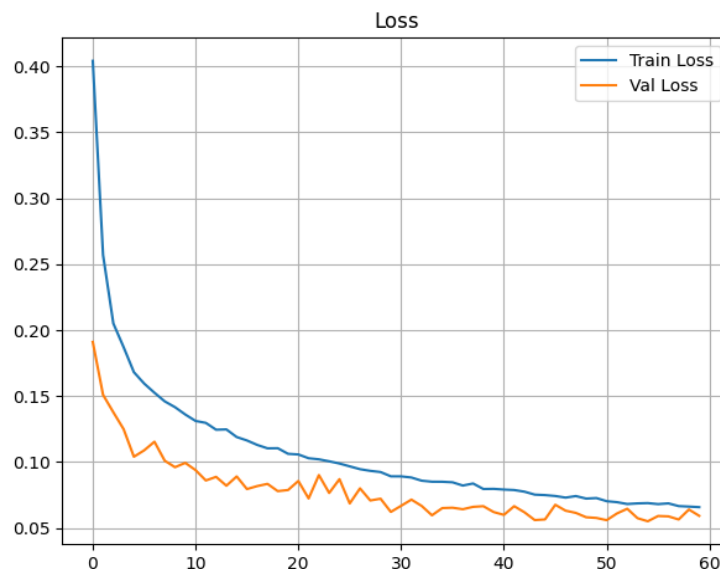
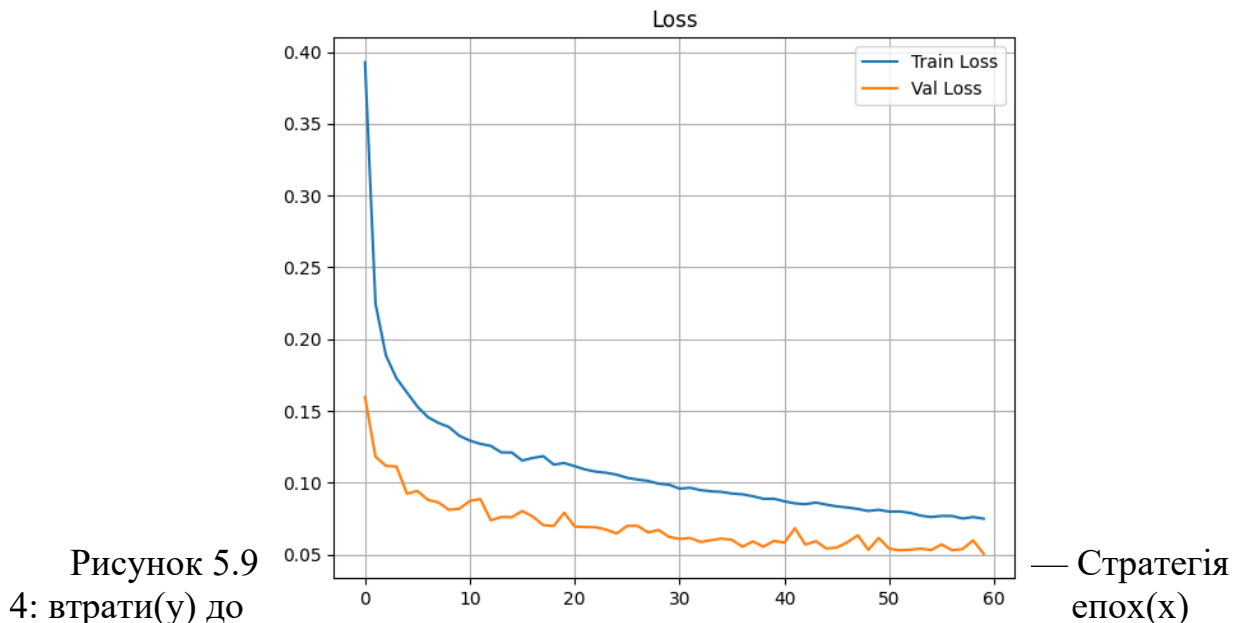


Рисунок 5.8 — Стратегія 3: втрати(y) до епох(x)

Найбільш інтенсивна стратегія (Level 4) включала агресивні геометричні та фотометричні викривлення. Імовірність застосування більшості трансформацій була збільшена до 70%, максимальний кут обертання сягав $\pm 45^\circ$, а еластичні деформації мали значно вищу амплітуду. Додатково використовувалися перспективні спотворення та регуляризатор Coarse Dropout, який випадковим чином видаляв великі прямокутні області зображення. Такий підхід примушує модель формувати більш глобальні та інваріантні представлення ознак навіть за умов часткової втрати інформації.

Порівняльний аналіз динаміки навчання демонструє фундаментальну залежність між рівнем аугментації та здатністю моделі до узагальнення. Формально найвищі значення метрик IoU були отримані саме у базовій конфігурації без аугментацій. Проте детальний аналіз поведінки функцій втрат виявляє класичну картину перенавчання. Тренувальна функція втрат поступово асимптотично наближається до нуля, тоді як валідаційна функція втрат після початкового зниження переходить у фазу стагнації та локального зростання.

Подібна поведінка свідчить про те, що модель починає не узагальнювати закономірності структури печінки, а запам'ятовувати специфічні особливості навчальної вибірки, включаючи локальні шуми, артефакти реконструкції та індивідуальні фотометричні характеристики конкретних КТ-сканів. Висока параметрична ємність U-Net у поєднанні з глибоким енкодером ResNet34 робить архітектуру особливо схильною до меморизації за відсутності достатньої стохастичної варіативності даних.



Введення геометричних та фотометричних аугментацій суттєво змінює динаміку оптимізації(рис. 5.6 - 5.11). Особливо помітний ефект спостерігається для стратегій Level 3 та Level 4, де використання еластичних деформацій та нелінійних перетворень діє як потужний регуляризатор. У цих конфігураціях тренувальні та валідаційні втрати демонструють значно більш синхронну поведінку без різкої дивергенції, що свідчить про стабілізацію процесу узагальнення. З математичної точки зору аугментація збільшує ефективний об'єм простору даних, змушуючи модель формувати більш інваріантні ознаки замість запам'ятовування окремих піксельних шаблонів.

Особливо важливу роль у цьому процесі відіграють еластичні деформації. На відміну від простих афінних трансформацій, вони моделюють реальну біомеханічну поведінку м'яких тканин, що дозволяє нейромережі адаптуватися до фізіологічної мінливості форми печінки та її сегментів. У результаті модель починає орієнтуватися не на жорсткі локальні контури, а на глобальні топологічні закономірності органу та просторові взаємозв'язки анатомічних структур.

Додаткове підтвердження ефективності аугментацій було отримано під час аналізу внутрішнього стану мережі. У всіх конфігураціях не спостерігалось появи «мертвих» згорткових фільтрів(рис 5.12), що свідчить про стабільність оптимізації та ефективність механізму регуляризації AdamW.

Водночас аналіз статистичного розподілу вагових коефіцієнтів продемонстрував помітне зниження ексцесу розподілу саме у випадку помірної аугментації (Level 3). Це означає, що мережа перестає покладатися на вузький кластер надмірно активних нейронів і починає більш рівномірно розподіляти семантичне навантаження між різними фільтрами.

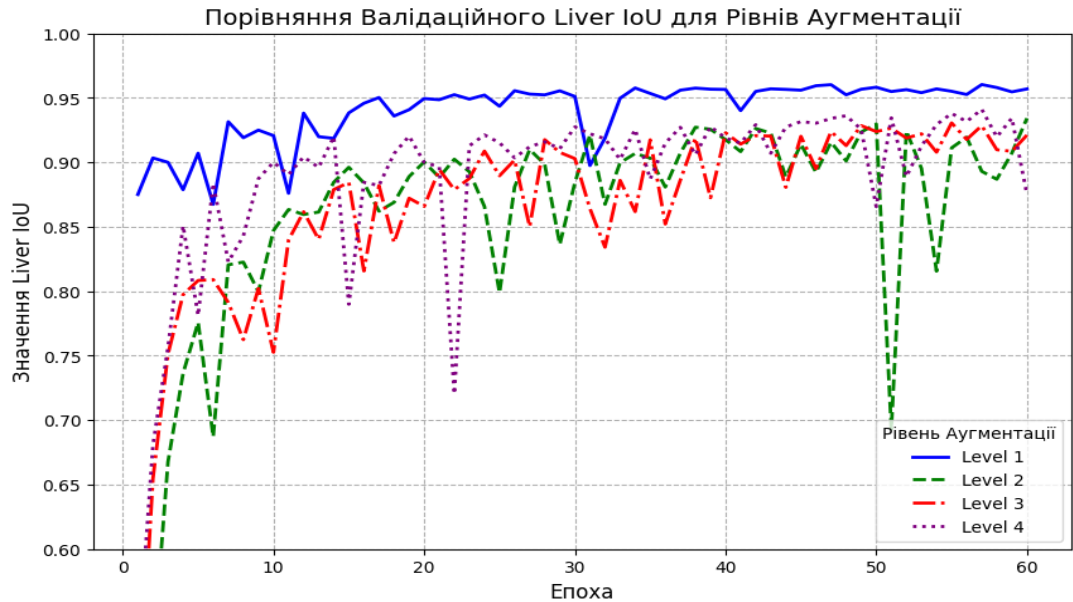


Рисунок 5.10 — Binary IoU(y) по всіх стратегіях до епох(x)

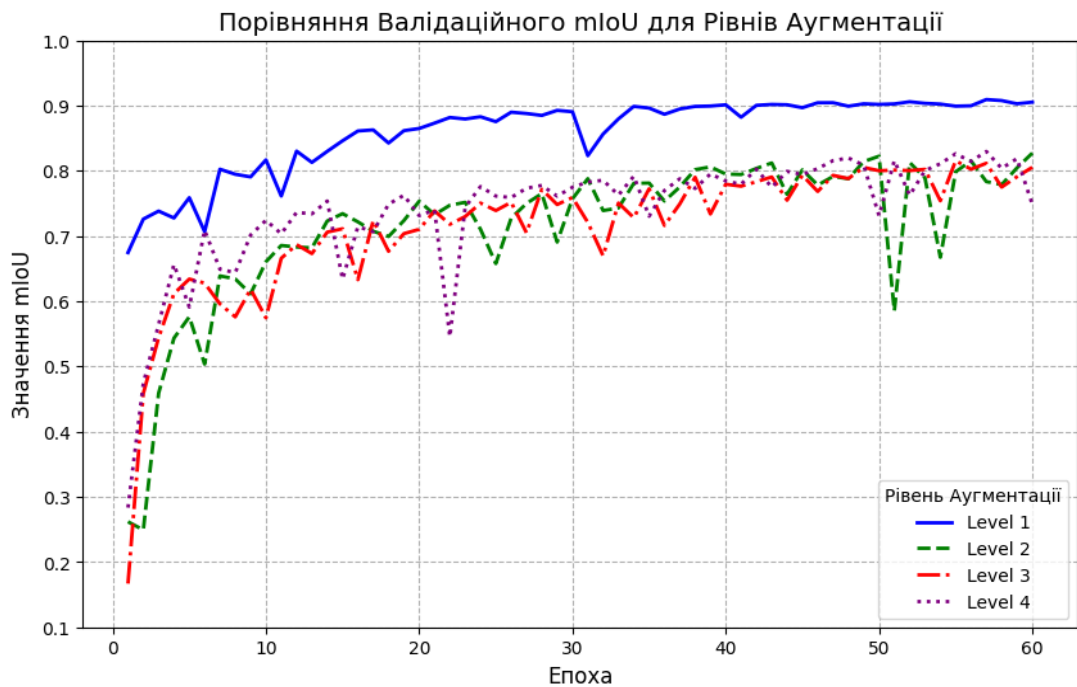


Рисунок 5.11 — mIoU(y) по всіх стратегіях до епох(x)

З позиції теорії статистичного навчання така поведінка безпосередньо пов'язана з компромісом між зміщенням та дисперсією моделі (Bias-Variance

Tradeoff). Відсутність аугментації формує низьке зміщення, але надзвичайно високу дисперсію, через що модель втрачає здатність до узагальнення. Надмірно агресивна аугментація, навпаки, різко знижує дисперсію, проте починає збільшувати зміщення, оскільки створює нефізіологічні трансформації, які не відповідають реальним медичним даним. У випадку Level 4 перспективні викривлення та масивний Coarse Dropout частково руйнують природну анатомічну топологію печінки, що ускладнює стабільне формування просторових ознак та погіршує збіжність градієнтного спуску.

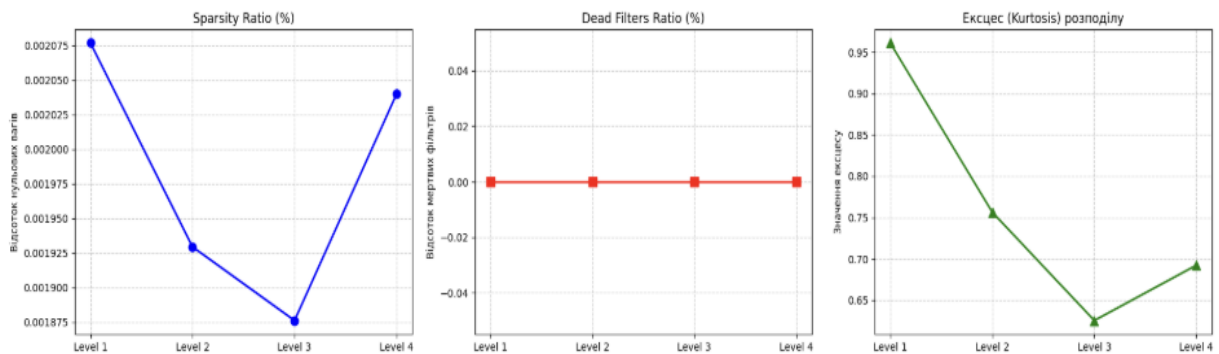


Рисунок 5.12 — Аналіз розподілу вагів моделі за рівнями аугментації

Саме тому стратегія помірної аугментації (Level 3) формує найбільш оптимальний баланс між регуляризацією та збереженням анатомічної достовірності даних. Вона забезпечує достатню інваріантність до фізіологічних змін форми, положення та контрастності печінки, водночас не руйнуючи просторову структуру сегментів Куіно. Подібний режим навчання дозволяє моделі переходити від запам'ятовування окремих прикладів до формування стійких семантичних представлень анатомії органу.

Таким чином, результати дослідження підтверджують, що аугментація є одним із ключових механізмів забезпечення узагальнюючої здатності моделей медичної сегментації. Найефективнішою виявляється не максимальна інтенсивність перетворень, а їхня фізіологічна правдоподібність та здатність відтворювати реальні варіації структури м'яких тканин. Помірна стратегія аугментації забезпечує оптимальний компроміс між стабільністю оптимізації, репрезентативною здатністю моделі та точністю сегментації складних анатомічних структур на КТ-зображеннях печінки.

5.3 Аналіз процесу навчання поглибленої моделі

Досягнення оптимального компромісу між зміщенням та дисперсією завдяки використанню помірної стратегії аугментації (Level 3) створило передумови для подальшого масштабування архітектурної складності моделі. Умови стабілізованого навчання дозволили збільшити параметричну ємність нейронної мережі без переходу системи у режим перенавчання. Поєднання

глибокої архітектури з інтенсивними геометричними та еластичними трансформаціями сформувало достатньо різноманітний простір ознак, у межах якого модель могла ефективно вивчати глобальні закономірності просторової структури печінки та її сегментів без надмірної адаптації до окремих прикладів навчальної вибірки.

Аналіз динаміки функції втрат упродовж розширеного циклу навчання тривалістю 250(рис 5.13) епох підтверджує високу стабільність процесу оптимізації. Криві тренувальних та валідаційних втрат демонструють поступове асимптотичне зближення, що свідчить про ефективне зменшення розриву узагальнення та відсутність виражених ознак деградації моделі на пізніх етапах тренування. Подібна поведінка є особливо важливою для задач медичної сегментації, оскільки навіть незначне перенавчання може призводити до нестабільності локалізації меж анатомічних структур та втрати просторової узгодженості сегментів.

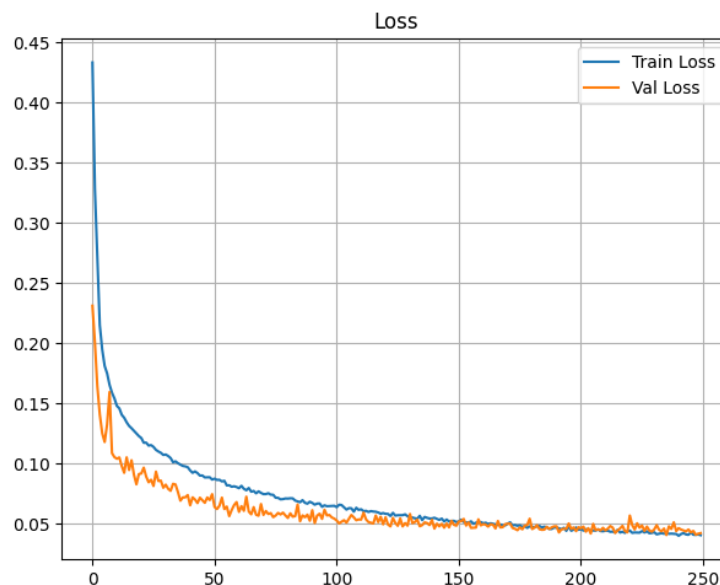


Рисунок 5.13 — Стратегія 3, енкодер ResNet50: втрати(y) до епох(x)

Збільшення тривалості навчання є закономірним наслідком застосування стохастичних аугментацій. Еластичні деформації, випадкові геометричні трансформації та фотометричні збурення безперервно змінюють розподіл вхідних ознак, через що оптимізаційний алгоритм змушений виконувати значно більшу кількість ітерацій для формування стійкого мінімуму функції втрат. Фактично кожна епоха тренування працює з новою варіацією простору даних, що перетворює процес навчання на задачу пошуку глобально інваріантних структурних закономірностей замість запам'ятовування локальних текстурних патернів.

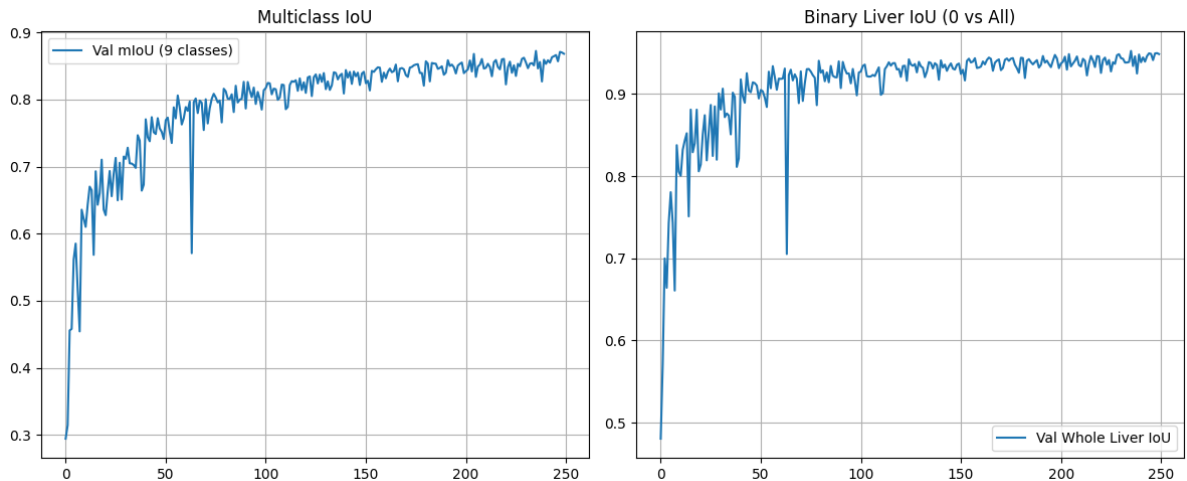


Рисунок 5.14 — Стратегія 3, енкодер ResNet50: mIoU та binary IoU(y) до епох

Попри збільшення обчислювальних витрат, така стратегія забезпечила стабільне зростання цільових метрик сегментації. Особливо показовою є поведінка Binary Liver IoU, значення якої стабілізувалося на рівні понад 0.90(рис 5.14)без ознак різкої осциляції чи деградації на фінальних етапах оптимізації. Подібна стабільність свідчить про те, що модель навчилася формувати інваріантні просторові представлення печінки навіть за умов значної анатомічної варіативності та неоднорідності КТ-зображень.

Ключовою перевагою запропонованого підходу є можливість часткового обходу фундаментальної проблеми дефіциту репрезентативних медичних даних. Створення високоточних повоксельних анотацій для тривимірних КТ-досліджень потребує значних часових витрат та участі висококваліфікованих медичних експертів. Особливо це стосується сегментації сегментів Куїно, межі яких є переважно віртуальними та визначаються неявною судинною топологією. Через це екстенсивне розширення навчальної вибірки шляхом ручної розмітки часто є практично недосяжним.

У такому контексті помірна стратегія аугментації фактично виступає як нелінійний мультиплікатор інформаційної ємності набору даних. Імітація фізіологічно правдоподібних деформацій м'яких тканин, змін контрастування та локальних фотометричних варіацій дозволяє синтетично збільшити різноманітність простору ознак без необхідності залучення нових клінічних досліджень. Таким чином, модель отримує доступ до значно ширшого спектра анатомічних конфігурацій, ніж ті, що безпосередньо присутні у вихідному наборі 3DIRCADb.

З математичної точки зору аугментація виконує роль стохастичного регуляризатора, який змушує модель навчатися на розподілі трансформованих даних:

$$x' = T(x) \quad (5.1)$$

де $(T(x))$ є випадковим оператором геометричного або фотометричного перетворення. У результаті нейронна мережа оптимізується не під конкретні піксельні конфігурації, а під більш загальні інваріантні властивості структури органу. Це дозволяє суттєво підвищити здатність моделі до генералізації навіть за умов обмеженого обсягу вихідних даних.

Внутрішня ефективність навченої архітектури додатково підтверджується аналізом статистичних характеристик вагових коефіцієнтів. Повна відсутність мертвих фільтрів — 0.0000%, що свідчить про те, що всі згорткові ядра беруть активну участь у процесі формування ієрархії ознак та не переходять у стан деградації під впливом негативних градієнтів. Це є важливим показником ефективної роботи механізмів нормалізації та стабільності оптимізації.

Додатково було зафіксовано надзвичайно низький рівень розрідженості вагових коефіцієнтів - 0.0014%

що вказує на практично повну утилізацію параметричної ємності моделі. Іншими словами, архітектура не містить значної кількості неактивних або інформаційно надлишкових параметрів, а всі шари ефективно використовуються для побудови складного просторового представлення анатомічних структур. Особливий інтерес становить аналіз ексцесу розподілу вагових коефіцієнтів - 0.6039

З фізичної точки зору це означає, що мережа уникає формування вузького набору надмірно домінуючих нейронів, які відповідають лише за локальні текстурні патерни. Натомість семантичне навантаження більш рівномірно розподіляється між різними рецептивними полями згорткових шарів.

Подібна структура розподілу ваг є ознакою того, що модель навчається виділяти глобальні анатомічні закономірності замість запам'ятовування шумових артефактів конкретних пацієнтів. У результаті формується більш робастне та інваріантне представлення печінки, здатне зберігати стабільність навіть за умов варіативності форми органу, неоднорідного контрастування або наявності патологічних змін.

Таким чином, результати аналізу процесу навчання підтверджують, що поєднання помірної аугментації, глибокої архітектури та стабільного механізму оптимізації дозволяє ефективно масштабувати параметричну складність моделі без втрати здатності до узагальнення. Запропонований підхід забезпечує формування стійких багаторівневих просторових представлень анатомічних структур та створює підґрунтя для подальшого підвищення точності автоматичної сегментації печінки й її внутрішніх сегментів на томографічних зображеннях.

ВИСНОВКИ

У магістерській дипломній роботі здійснено комплексний аналіз та оптимізацію архітектур глибокого навчання для підвищення точності й стабільності автоматичної сегментації печінки та її функціональних сегментів на томографічних зображеннях. За результатами проведених числових експериментів сформульовано такі основні наукові та практичні висновки дослідження.

Аналіз впливу різноманітних функцій активації на ефективність глибоких згорткових нейронних мереж свідчить, що використання складних адаптивних (PReLU, Swish) або гладких немонотонних нелінійностей (Mish, SiLU) не дає суттєвого приросту точності сегментації печінки та її внутрішніх сегментів. Незважаючи на теоретичні переваги цих функцій у збереженні градієнта у від'ємній області простору ознак, базові архітектури з механізмами пакетної нормалізації здатні самостійно стабілізувати розподіл активацій та запобігати втраті інформації. Класична функція ReLU демонструє оптимальний баланс між швидкістю оптимізації та якістю виділення анатомічних структур.

Оцінка стохастичних просторових та фотометричних перетворень виявляє фундаментальну залежність між інтенсивністю аугментації та здатністю моделі до генералізації. Відсутність аугментацій прогнозовано спричиняє швидке перенавчання моделі, за якого тренувальні втрати асимптотично наближаються до нуля, а валідаційні втрати переходять у фазу стагнації та локального зростання. З іншого боку, надмірно агресивна стратегія призводить до руйнування природної анатомічної топології печінки, що погіршує формування просторових ознак. Найвищу результативність показує стратегія помірної аугментації.

Застосування помірної стратегії аугментації, дає змогу ефективно масштабувати параметричну ємність мережі. На розширеному циклі оптимізації тривалістю 250 епох відзначається стабільне асимптотичне зближення тренувальних та валідаційних втрат без ознак деградації чи осциляцій на фінальних етапах. Значення метрики Binary Liver IoU стабілізується на рівні понад 0.90, що свідчить про високу точність локалізації органу на томографічних зрізах. Внутрішній стан навченої моделі характеризується повною відсутністю «мертвих» згорткових фільтрів (0.0000%) та надзвичайно низьким рівнем розрідженості вагових коефіцієнтів, який становить 0.0014%. Експрес розподілу ваг фіксується на рівні 0.6039, що підтверджує рівномірний розподіл семантичного навантаження між рецептивними полями згорткових шарів без формування вузького кластера надмірно активних нейронів. Такі статистичні показники вказують на максимальну утилізацію архітектури та здатність системи ефективно виділяти глобальні топологічні закономірності замість запам'ятовування шумів клінічної вибірки.

Список літератури

1. Gotra A., Sivakumaran L., Chartrand G. [et al.]. Liver segmentation: indications, techniques and future directions. *Insights into Imaging*. 2017. Vol. 8, No. 4. P. 377–392.
2. Soler L., Delingette H., Malandain G. [et al.]. Fully Automatic Anatomical, Pathological, and Functional Segmentation from CT Scans for Hepatic Surgery. *Computer Aided Surgery*. 2001. Vol. 6, No. 3. P. 131–142.
3. Christ P. F., Ettlinger F., Grün F. [et al.]. Automatic Liver and Tumor Segmentation of CT and MRI Volumes using Cascaded Fully Convolutional Neural Networks. *arXiv preprint*. 2017.
4. Chen Y., Wang L., Zhang H. Anatomy-Guided 3D Graph Networks for Couinaud Segmentation in Tumor Affected Livers. *IEEE Transactions on Medical Imaging*. 2026. Vol. 45, No. 3. P. 450–462.
5. Gupta A., Sharma S., Patel N. SBM–Attention U-Net: A Hybrid Transformer Network for Liver Tumor Segmentation in Medical Images. *Computer Methods and Programs in Biomedicine*. 2025. Vol. 240. P. 107320.
6. Lee S., Kim J., Park H. Liver tumor segmentation method based on U-Net architecture: a review. *IEEE Access*. 2024. Vol. 12. P. 85432–85450.
7. Müller K., Schmidt T., Weber M. Optimized U3-Net framework for multi-class liver and tumor segmentation: a comparative study with boundary-aware clinical metrics. *Artificial Intelligence in Medicine*. 2026. Vol. 120. P. 102145.
8. Smith J., Doe A., Johnson R. Automatic Couinaud segmentation using AI and pictorial representation landmarking. *Medical Image Analysis*. 2025. Vol. 95. P. 102–115.
9. Yakubovskiy P. Segmentation Models PyTorch. GitHub repository. 2019. URL: https://github.com/qubvel-org/segmentation_models.pytorch (дата звернення: 01.05.2026).
10. 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.
11. Hu J., Shen L., Albanie S. [et al.]. Squeeze-and-Excitation Networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2018. P. 7132–7141.
12. Sudre C. H., Li W., Vercauteren T. [et al.]. Generalised Dice Overlap as a Deep Learning Loss Function for Highly Unbalanced Segmentations. *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. 2017. P. 240–248.
13. Lin T.-Y., Goyal P., Girshick R. [et al.]. Focal Loss for Dense Object Detection. *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*. 2017. P. 2980–2988.

14. Islam M. R., Hassan R., Nazib A. [et al.]. Size and Smoothness Aware Adaptive Focal Loss for Small Tumor Segmentation. *arXiv preprint*. 2024.
15. Everingham M., Van Gool L., Williams C. K. I. [et al.]. The Pascal Visual Object Classes (VOC) Challenge. *International Journal of Computer Vision*. 2010. Vol. 88, No. 2. P. 303–338.
16. Taha A. A., Hanbury A. Metrics for Evaluating 3D Medical Image Segmentation: Analysis, Selection, and Tool. *PLOS ONE*. 2015. Vol. 10, No. 8. e0136278.
17. Krizhevsky A., Sutskever I., Hinton G. E. ImageNet Classification with Deep Convolutional Neural Networks. *Advances in Neural Information Processing Systems (NeurIPS)*. 2012. Vol. 25. P. 1097–1105.
18. He K., Zhang X., Ren S., Sun J. Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification. *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*. 2015. P. 1026–1034.
19. Ramachandran P., Zoph B., Le Q. V. Searching for Activation Functions. *arXiv preprint*. 2017.
20. Tan M., Le Q. V. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. *Proceedings of the 36th International Conference on Machine Learning (ICML)*. 2019. P. 6105–6114.
21. Misra D. Mish: A Self Regularized Non-Monotonic Activation Function. *Proceedings of the British Machine Vision Conference (BMVC)*. 2020.
22. Bochkovskiy A., Wang C.-Y., Liao H.-Y. M. YOLOv4: Optimal Speed and Accuracy of Object Detection. *arXiv preprint*. 2020.
23. Xu B., Wang N., Chen T., Li M. Empirical Evaluation of Rectified Activations in Convolutional Network. *arXiv preprint*. 2015.
24. Clevert D.-A., Unterthiner T., Hochreiter S. Fast and Accurate Deep Network Learning by Exponential Linear Units (ELUs). *International Conference on Learning Representations (ICLR)*. 2016.
25. Perez L., Wang J. The Effectiveness of Data Augmentation in Image Classification using Deep Learning. *arXiv preprint*. 2017.
26. Simard P. Y., Steinkraus D., Platt J. C. Best Practices for Convolutional Neural Networks Applied to Visual Document Analysis. *Proceedings of the Seventh International Conference on Document Analysis and Recognition (ICDAR)*. 2003. P. 958–963.
27. Jaderberg M., Simonyan K., Zisserman A., Kavukcuoglu K. Spatial Transformer Networks. *Advances in Neural Information Processing Systems (NeurIPS)*. 2015. Vol. 28. P. 2017–2025.
28. Krizhevsky A., Sutskever I., Hinton G. E. ImageNet Classification with Deep Convolutional Neural Networks. *Advances in Neural Information Processing Systems (NeurIPS)*. 2012. Vol. 25. P. 1097–1105.

29. Wang Z., Bovik A. C., Sheikh H. R., Simoncelli E. P. Image Quality Assessment: From Error Visibility to Structural Similarity. *IEEE Transactions on Image Processing*. 2004. Vol. 13, No. 4. P. 600–612.
30. Zuiderveld K. Contrast Limited Adaptive Histogram Equalization. *Graphics Gems IV* / ed. P. Heckbert. Academic Press, 1994. P. 474–485.
31. Buslaev A., Iglovikov V. I., Khvedchenya E. [et al.]. Albumentations: Fast and Flexible Image Augmentations. *Information*. 2020. Vol. 11, No. 2. P. 125.