

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет Комп'ютерних наук
(повна назва)

Кафедра Штучного інтелекту
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА

Пояснювальна записка

рівень вищої освіти другий (магістерський)

Аналіз алгоритмів сегментації пухлин за гібридними зображеннями
(тема)

Виконав:
студент 2 курсу, групи СШМ-22-2
Мандріков А.Д.
(прізвище, ініціали)

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

Тип програми освітньо-наукова
(освітньо-професійна або освітньо-наукова)

Освітня програма Системи штучного інтелекту
(повна назва спеціалізації)

Керівник доц. Авер'янова Л.О.
(посада, прізвище, ініціали)

Допускається до захисту

Зав. кафедри _____
(підпис)

В.О. Філатов
(прізвище, ініціали)

2024 р.

Харківський національний університет радіоелектроніки

Факультет Комп'ютерних наук
(повна назва)
Кафедра Штучного інтелекту
(повна назва)
Рівень вищої освіти другий (магістерський)
Спеціальність 122 Комп'ютерні науки
(код і повна назва)
Тип програми освітньо-наукова
(освітньо-професійна або освітньо-наукова)
Освітня програма Системи штучного інтелекту
(повна назва)

ЗАТВЕРДЖУЮ:

Зав. кафедри _____

(підпис)

«_____» _____ 20__ р.

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ

студентові Мандрікову Андрію Денисовичу
(прізвище, ім'я, по батькові)

1. Тема роботи Аналіз алгоритмів сегментації пухлин за гібридними зображеннями

затверджена наказом університету від 1 квітня 20 24 р. № 260Ст

2. Термін подання студентом роботи до екзаменаційної комісії 4 червня 20 24 р.

3. Вихідні дані до роботи Теоретичне дослідження рішення та опис технологій, пов'язаних із сегментацією пухлин на основі гібридних зображень, робота з потоком даних для аналізу та обробки медичних зображень пухлин, математична основа рішення, що використовується у вибраних алгоритмах сегментації, практична частина роботи, включаючи вибір та реалізацію підходів для сегментації пухлин на гібридних зображеннях, застосувані технології та методи програмної реалізації алгоритмів сегментації, попередня обробка та візуалізація даних для підготовки даних та аналізу результатів сегментації.

4. Перелік питань, що потрібно опрацювати в роботі _____

1) Огляд задачі сегментації медичних зображень

2) Огляд алгоритмів та метрик для сегментації пухлин

3) Особливості реалізації програмного застосунку

РЕФЕРАТ

Пояснювальна записка: 70 с., 23 рис., 1 табл., 1 дод., 17 джерел.

ГІБРИДНІ ЗОБРАЖЕННЯ, КОМП'ЮТЕРНИЙ ЗІР, НЕЙРОННІ МЕРЕЖІ, СЕГМЕНТАЦІЯ МЕДИЧНИХ ЗОБРАЖЕНЬ.

Об'єкт дослідження – аналіз та розробка методів сегментації пухлин за допомогою гібридних зображень.

Предмет дослідження – алгоритми сегментації пухлин на основі гібридних зображень у медичній практиці.

Мета роботи – оцінка ефективності та точності різних алгоритмів сегментації пухлин за допомогою гібридних зображень та вибір найбільш адекватного алгоритму для конкретних клінічних умов.

Методи дослідження – аналіз літературних джерел, експериментальне порівняння різних алгоритмів на тестових наборах медичних зображень, визначення показників ефективності та порівняння результатів.

ABSTRACT

Master's thesis contains: 70 p., 23 fig., 1 tabl., 1 ann., 17 references.

**HYBRID IMAGES, COMPUTER VISION, NEURAL NETWORKS,
MEDICAL IMAGE SEGMENTATION.**

The object of the research is the analysis and development of tumor segmentation methods using hybrid images.

The subject of the research – tumor segmentation algorithms based on hybrid images in medical practice.

The aim of the work is to evaluate the effectiveness and accuracy of different tumor segmentation algorithms using hybrid images and to select the most adequate one for specific clinical conditions.

Research methods include analysis of literature sources, experimental comparison of different algorithms on test sets of medical images, determination of performance indicators, and comparison of results.

ЗМІСТ

Перелік умовних позначень, символів, одиниць, скорочень і термінів	7
Вступ.....	8
1 Огляд задачі сегментації медичних зображень.....	12
1.1 Опис задачі сегментації.....	12
1.2 Постановка задачі сегментації.....	15
1.3 Огляд методів сегментації зображень.....	15
2 Огляд алгоритмів та метрик для сегментації пухлин.....	28
2.1 Застосування гібридних зображень у сегментації пухлин	28
2.2 Аналіз алгоритмів сегментації пухлин	33
2.3 Вибір метрик оцінки моделі сегментації пухлин.....	43
3 Особливості реалізації програмного застосунку	48
3.1 Огляд вибраного набору даних.....	48
3.2 Попередня обробка даних	56
3.3 Навчання моделі та аналіз метрик.....	58
Висновки	65
Перелік джерел посилання	67
Додаток А Відомість кваліфікаційної роботи	70

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ

ЗНМ – згорткова нейронна мережа;

ШІ – штучний інтелект;

DL – Deep Learning – глибинне навчання;

DNN – Deep Neural Network – глибока нейрона мережа;

IoU – Intersection over Union – показник якості сегментації, що визначає відповідність між сегментацією та реальним об'єктом на зображенні;

ROI – Region of Interest – область інтересу (частина зображення, яка підлягає аналізу).

ВСТУП

Актуальність досліджуваної теми визначається в контексті активного розвитку медичної науки та технологій. Зростання кількості випадків онкологічних захворювань та потреби у ранній діагностиці створюють підвищений попит на нові методи та інструменти для аналізу пухлин. Гібридні зображення, які об'єднують дані з різних медичних обладнань, відкривають нові можливості для точної та ефективною сегментації пухлин. Підвищена актуальність полягає в тому, що існуючі методи сегментації не завжди забезпечують необхідну точність та надійність, особливо при складних клінічних випадках.

Дослідження в цій області також є актуальним через постійне зростання обсягу медичних даних та потреби у їхньому аналізі. Швидкий розвиток обчислювальних технологій і штучного інтелекту відкриває нові можливості для розробки більш точних і автоматизованих методів сегментації пухлин. Важливість цієї теми полягає в тому, що вона спрямована на забезпечення швидкого та ефективного доступу до точної медичної інформації, що може покращити діагностику, лікування та прогнозування онкологічних захворювань.

Таким чином, актуальність дослідження алгоритмів сегментації пухлин за допомогою гібридних зображень є невід'ємною частиною стрімкого розвитку медичної науки та технологій, а також відповідає загальним вимогам сучасного медичного сектору щодо покращення точності та ефективності діагностики онкологічних захворювань.

Проблематика дослідження у сегментації пухлин за допомогою гібридних зображень охоплює кілька ключових аспектів, які варто враховувати:

– недостатня точність і надійність існуючих методів – багато з існуючих методів сегментації пухлин мають свої обмеження, особливо

при роботі з гібридними зображеннями, що поєднують дані з різних медичних систем. Це може призводити до неточностей та неповноцінності діагностики;

- складність обробки гібридних зображень. Гібридні зображення часто містять інформацію різної природи, таку як структурна та функціональна інформація, що може ускладнювати їх обробку. Розробка алгоритмів сегментації, які ефективно працюють з цими різномірними даними, є складним завданням;

- неоднорідність даних та варіативність пухлин. Пухлини можуть мати різну форму, розмір, місцезнаходження та внутрішню структуру. Ця варіабельність ускладнює розробку універсальних алгоритмів сегментації, які були б ефективними для різних типів пухлин та медичних умов;

- потреба у високій швидкості. Для впровадження алгоритмів сегментації в клінічну практику важливо мати методи, які працюють швидко та ефективно, забезпечують стабільність та точність навіть у складних умовах. Отже, проблематика дослідження включає в себе ряд складних викликів, які потребують аналізу та розробки нових методів та підходів для ефективної сегментації пухлин за допомогою гібридних зображень.

У ході роботи будуть вирішені наступні завдання:

- дослідження та опис актуального стану вивченості проблеми сегментації пухлин за допомогою гібридних зображень у медичній практиці;

- проведення огляду та аналізу основних методів сегментації пухлин за допомогою гібридних зображень;

- вибір найбільш оптимального алгоритму для сегментації таких зображень та порівняння його ефективності з іншими методами;

- практична реалізація обраного алгоритму та оцінка його точності та ефективності.

Метою роботи є оцінка ефективності та точності різних алгоритмів сегментації пухлин за допомогою гібридних зображень та вибір найбільш оптимального алгоритму для медичних досліджень.

Об'єктом дослідження є аналіз та розробка методів сегментації пухлин за допомогою гібридних зображень алгоритми сегментації пухлин, а предметом – алгоритми сегментації пухлин на основі гібридних зображень у медичній практиці.

Для реалізації дослідження з аналізу алгоритмів сегментації пухлин за гібридними зображеннями були використані різноманітні методи та підходи. Починаючи від огляду та аналізу наукових джерел, що розглядають сучасні тенденції у цій області, і завершуючи практичною реалізацією алгоритмів за допомогою спеціалізованих програмних засобів. Одним із ключових етапів було вивчення та опис актуального стану дослідження в галузі сегментації пухлин. Це дозволило визначити основні напрямки розвитку та перспективні методи, які варто розглядати у рамках власного дослідження.

Далі, з метою вибору найбільш оптимального алгоритму для сегментації пухлин на гібридних зображеннях, було проведено огляд і аналіз різних підходів. Це включало в себе оцінку їхньої ефективності, точності та можливостей застосування у медичній практиці. Після вибору оптимального алгоритму, наступним кроком була його практична реалізація та оцінка. За допомогою сучасних програмних засобів, таких як TensorFlow і Keras, була побудована модель, яка здатна ефективно сегментувати пухлини на гібридних зображеннях. Після цього проводилися експерименти та тестування для оцінки точності та ефективності розробленого алгоритму.

В майбутньому, на основі отриманих результатів, можливо подальше вдосконалення розробленого алгоритму сегментації пухлин за гібридними зображеннями. Це може включати в себе розширення функціональності

моделі для врахування різних типів пухлин та їхніх характеристик. Крім того, можливе поглиблення досліджень у напрямку використання новітніх методів машинного навчання та штучного інтелекту для поліпшення точності результатів.

Також, в майбутньому можливе розширення області застосування розроблених алгоритмів на інші захворювання, де сегментація зображень має велике значення для діагностики та лікування.

1 ОГЛЯД ЗАДАЧІ СЕГМЕНТАЦІЇ МЕДИЧНИХ ЗОБРАЖЕНЬ

1.1 Опис задачі сегментації

Сегментація зображень відіграє важливу роль у цифровій обробці зображень, спрямовуючись на поділ зображення на різні частини згідно їх унікальних властивостей. Головна мета цієї процедури полягає в спрощенні зображення для подальшого аналізу. Під час сегментації зображення воно розбивається на окремі сегменти, які мають схожі характеристики. Ці сегменти, які утворюються в результаті, називаються об'єктами зображення, і вони становлять основу для подальшого аналізу. Застосовуючи методи сегментації зображень, можна виділяти та групувати певні пікселі на зображенні, призначати їм мітки та класифікувати інші пікселі відповідно до цих міток. Таким чином, можна малювати межі, визначати контури та відокремлювати окремі об'єкти від інших компонентів на зображенні.

Одним з найпоширеніших використань сегментації є медична діагностика, де зображення може бути розділене на різні області, що відповідають окремим органам чи структурам у тілі пацієнта. Це допомагає лікарям точно визначати місця уражень, пухлин або інших патологічних змін, що дозволяє зробити більш точний та швидкий діагноз. Крім того, сегментація використовується у робототехніці для розпізнавання об'єктів та навігації роботів у середовищі. У відеоспостереженні та обробці зображень вона використовується для виявлення об'єктів, відстеження руху та використання у системах безпеки. Крім того, сегментація зображень важлива у галузі автомобільної промисловості для розпізнавання дорожньої смуги, пішоходів та інших об'єктів для систем автоматичного водіння та безпеки на дорозі. Вона також застосовується у виробництві для інспекції та контролю якості продукції, у сільському господарстві для аналізу земель та рослин, а також у багатьох інших галузях.

Сегментація зображень використовується для визначення місця розташування об'єктів та їх меж, таких як лінії та криві на зображеннях. Це процес присвоєння мітки кожному пікселю зображення таким чином, щоб пікселі з однаковою міткою мали певні спільні характеристики.

Завдання сегментації зображень можна поділити на три групи в залежності від кількості та типу інформації, яку вони передають:

– семантична сегментація відноситься до класифікації пікселів на зображенні на семантичні класи (рисунк 1.1). Пікселі, які належать до певного класу, просто класифікуються до цього класу без урахування іншої інформації або контексту. Цей підхід не ефективний, коли на зображенні присутні тісно згруповані кілька екземплярів одного класу;

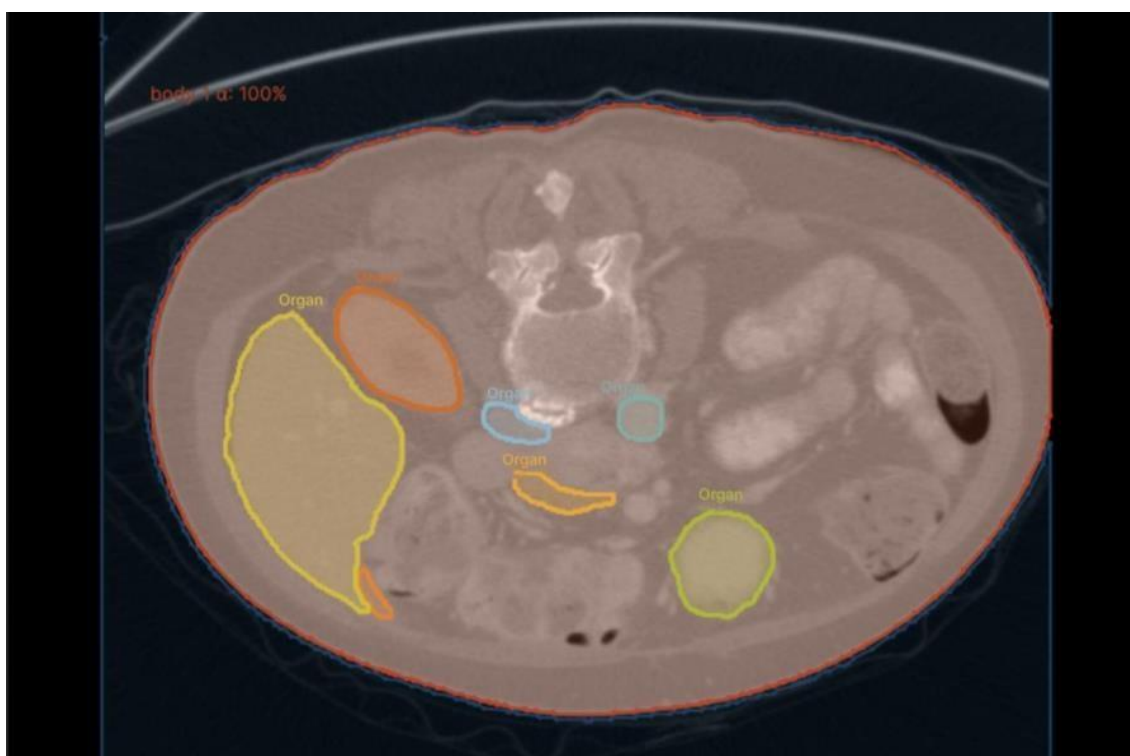


Рисунок 1.1 – Приклад семантичної сегментації

– сегментація екземплярів. Моделі сегментації екземплярів класифікують пікселі за категоріями на основі «екземплярів», а не класів.

Алгоритм сегментації екземплярів не має уявлення про клас, до якого належить класифікована область, але може відрізнити об'єкти, що перекриваються або дуже схожі області об'єктів на основі їх меж (рисунок 1.2);

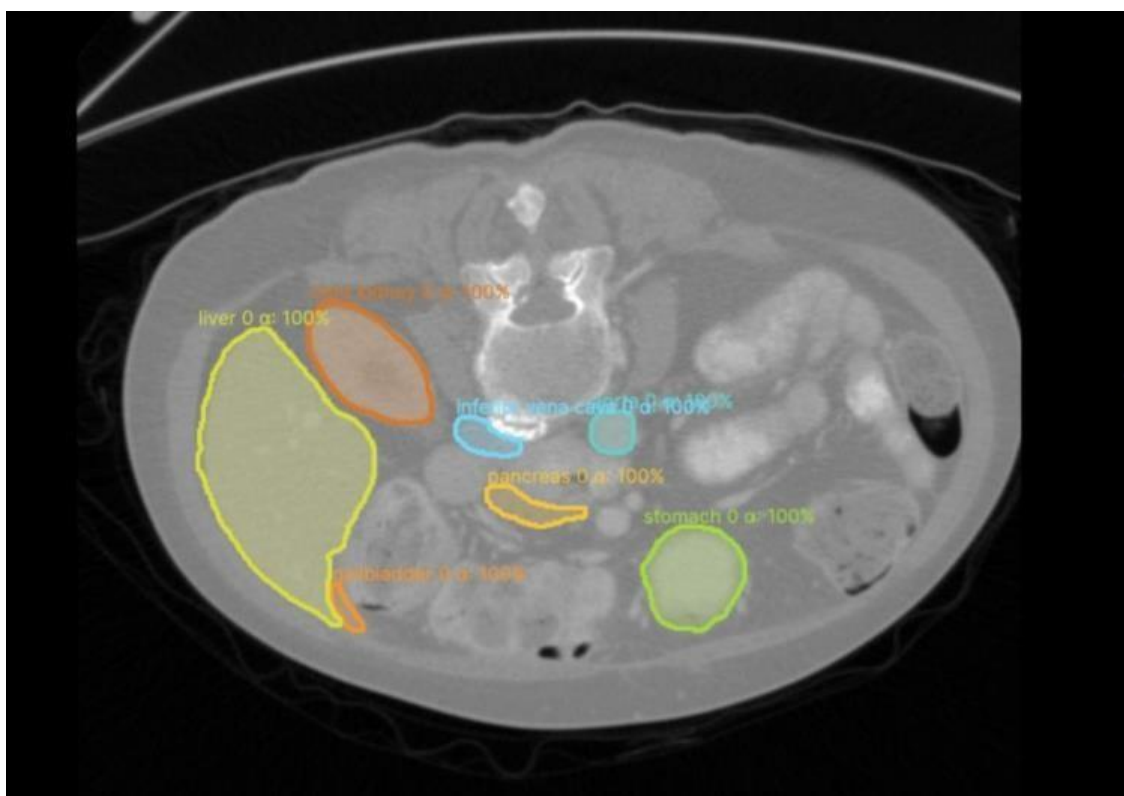


Рисунок 1.2 – Приклад сегментації екземплярів

– паноптична сегментація може бути виражена як комбінація семантичної сегментації та сегментації екземплярів, де кожен екземпляр об'єкта на зображенні відрізняється, а ідентичність об'єкта прогнозується. Цей тип знаходить широке застосування в популярних завданнях, таких як автоматизовані автомобілі, де потрібно захопити великий обсяг інформації про оточення за допомогою відеопотоку. Паноптична сегментація надає нам сегментні карти всіх об'єктів будь-якого конкретного класу, присутніх на рисунку 1.3.

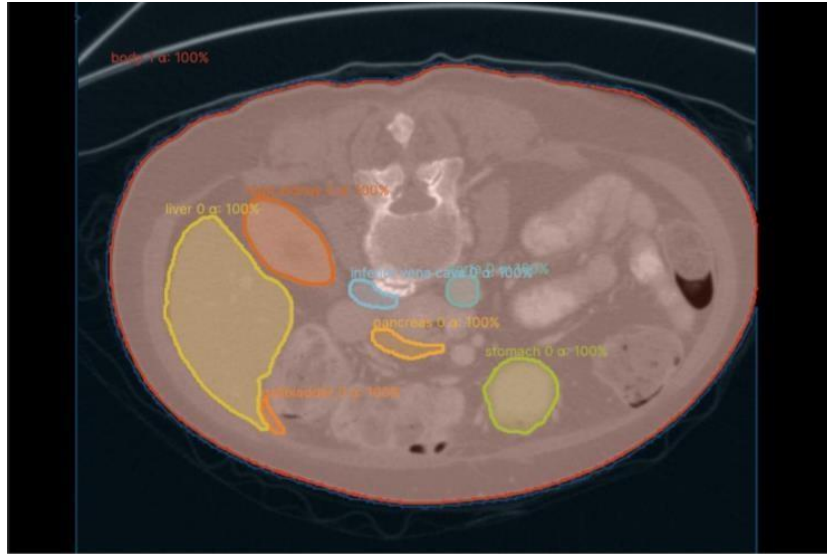


Рисунок 1.3 – Приклад паноптичної сегментації

1.2 Постановка задачі сегментації

Нехай задано цифрове зображення $I = \{p_{ij}\}$, де p_{ij} – піксель, який характеризується координатами (i, j) . Утворимо скінченний набір множин $\{S_k\}_{k=1}^N$, такий, що $\bigcup_{k=1}^N S_k = I$ і $S_{k_1} \cap S_{k_2} = \emptyset$ при $k_1 \neq k_2$. Знайдемо підмножину індексів $t = \{t_1, \dots, t_l\}$ таку, що довільні два пікселі p_{ij} та $p_{i'j'}$, де p_{ij} належить деякій множині S_q , $q \in t$, а $p_{i'j'}$ належить S_m , $m \notin t$, мають різні значення кольорових складових.

Необхідно зауважити, що у випадку, коли довільні два пікселі p_{ij} та $p_{i'j'}$, які належать деяким множинам S_q та S_l , де $q, l \in t$, то вони мають однакові значення кольорових складових.

1.3 Огляд методів сегментації зображень

Розглянемо основні типи методів сегментації зображень.

Порогова сегментація. Цей метод полягає у визначенні пікселів, які перевищують або не досягають певного порогового значення

інтенсивності. Він є простим та швидким у реалізації, але може бути неефективним в разі недостатньої чіткості контурів об'єктів або коли рівні інтенсивності змінюються.

У простій реалізації, якщо інтенсивність пікселів у рівнях сірого перевищує встановлений поріг T , їхня інтенсивність встановлюється на максимальне значення 255 – білий. У протилежному випадку, значення встановлюється на 0 – чорний. Таким чином, області зображення з інтенсивністю сірого більше порогу стають білими, а інші - навпаки, чорними.

Математично це можна записати так:

$$g(x, y) = \begin{cases} 255, & f(x, y) > T \\ 0, & f(x, y) \leq T \end{cases}, \quad T \in [0, 255]. \quad (1.1)$$

Даний алгоритм може бути ефективним для сегментації медичних зображень з високою контрастністю, наприклад, для сегментації рентгенівських зображень з чіткою різницею між тканинами.

Сегментація на основі країв. Виявлення країв – це процес визначення країв зображення, який є критичним для розуміння особливостей зображення. З точки зору сегментації, це процес класифікації пікселів зображення як крайових та виділення їх як окремий клас. Краї вважаються значущими ознаками та містять важливу інформацію. Це сприяє значному зменшенню розміру зображення, що буде оброблено далі, і фільтрації надлишкової інформації, зберігаючи увагу лише на важливих структурних аспектах зображення. Алгоритми виявлення країв в основному можна розділити на дві категорії: методи на основі градієнта та гістограми.

Алгоритми сегментації на основі країв функціонують для визначення перехрещень зображення на основі різниць у рівнях сірого кольору, текстури, яскравості, насиченості та контрасту. Для подальшого покращення результатів, необхідно застосувати додаткові обробки з метою об'єднання

всіх виявлених країв у ланцюжки, які найкраще відповідають межам зображення. Зазвичай це досягається за допомогою спеціальних фільтрів, які дозволяють виділити краї зображення під час їх згортки. Ці фільтри розраховуються за допомогою спеціальних алгоритмів, які працюють з визначенням градієнтів зображення в окремих координатах x та y . Оператори визначення меж, такі як оператор Собеля, Канні та Лапласіан, є основними складовими цих алгоритмів, які допомагають виявити перетин країв та визначити їх межі. Наприклад, алгоритм Канні є одним з найбільш популярних у цьому відношенні.

Оператор Собеля в основному використовується для виявлення країв і технічно є дискретним диференціальним оператором, який використовується для обчислення наближення градієнту функції яскравості зображення. Оператор Собеля є типовим оператором виявлення країв на основі першої похідної. У результаті застосування оператора введено подібну локальну операцію усереднення, тому шум має згладжуючий ефект і може ефективно усунути вплив шуму. Вплив оператора Собеля на позицію пікселя зважений, що краще, ніж оператори Прюїтта та Робертса. Оператор Собеля складається з двох наборів 3×3 матриць, які є поперечними та поздовжніми шаблонами, і наносяться на площину зображення, відповідно, для отримання різниці між горизонтальною та вертикальною різницею. Фактично використовуються наступні два шаблони для виявлення країв зображення:

– для виявлення горизонтальних країв (поперечний шаблон)

$$G_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}; \quad (1.2)$$

– для виявлення вертикальних країв (поздовжній шаблон)

$$G_Y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}. \quad (1.3)$$

Наближення горизонтального і вертикального градієнта для кожного пікселя зображення можна об'єднати, щоб обчислити розмір градієнта за такою формулою:

$$G = \sqrt{G_X^2 + G_Y^2}. \quad (1.4)$$

Потім градієнт можна розрахувати за наступною формулою:

$$\theta = \arctan\left(\frac{G_Y}{G_X}\right). \quad (1.5)$$

У наведеному вище прикладі, якщо вищевказаний кут θ дорівнює нулю, тобто зображення має поздовжній край, то лівий край темніший за правий.

Оператор Лапласа є ізотропним оператором, диференціальним оператором другого порядку. Він є більш доречним, коли цікавить лише положення краю, незалежно від різниці сірих відтінків пікселів навколо нього. Реакція оператора Лапласа на ізольовані пікселі є сильніша, ніж на край або лінію, і тому застосовується лише до зображень без шуму. За наявності шуму, оператор Лапласа повинен виконати низькочастотну фільтрацію перед виявленням краю. Тому звичайний алгоритм сегментації поєднує оператор Лапласа з оператором згладжування для створення нового шаблону.

Оператор Лапласа також є найпростішим ізотропним диференціальним оператором з обертальною інваріантністю. Перетворення Лапласа двовимірної функції зображення має вигляд:

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}. \quad (1.6)$$

Також оператор Лапласа може бути виражений у вигляді наступного шаблону:

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}. \quad (1.7)$$

Приклад застосування оператора Собеля та Лапласа наведено на рисунку 1.4.

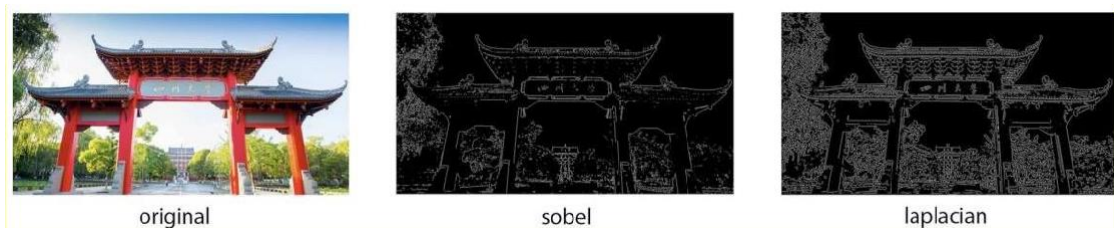


Рисунок 1.4 – Приклад застосування оператора Собеля та Лапласа

При застосуванні до набору зображень, типового для медичної візуалізації, отримані контури можна використовувати для створення тривимірних реконструкцій за допомогою алгоритмів інтерполяції.

Отже, цей метод використовує інформацію про зміни інтенсивності між пікселями для визначення контурів об'єктів. Він ефективний у виявленні деталей та контурів, але може бути схильний до впливу шуму та неоднорідності освітлення.

Межі та краї областей в зображеннях є важливими для сегментації, оскільки часто спостерігається значний перепад яскравості на кордонах областей. Таким чином, методи виділення країв використовуються як ключова складова для інших методів сегментації. Цей підхід часто

застосовується спільно з методом Grabcut. Зазвичай, він використовується після ручної сегментації для точного визначення контурів об'єктів або виправлення помилок, що можуть виникнути під час роботи з методом Grabcut.

Це допомагає виокремити області зображення, де відбувається значний перепад яскравості, що в подальшому може бути використано для визначення контурів об'єктів.

Даний алгоритм може бути корисним для сегментації деталей на медичних зображеннях, таких як контури органів, але не завжди ефективним для сегментації пухлин та інших об'єктів з нечіткими межами.

Сегментація на основі регіонів. Алгоритми сегментації на основі регіонів оперують шляхом пошуку схожостей між сусідніми пікселями та їх групування в загальні класи, де пікселі мають схожі характеристики. Зазвичай процес сегментації розпочинається з певних пікселів або невеликих областей зображення, встановлених як початкові елементи, і алгоритм працює, визначаючи безпосередні межі початкових пікселів і класифікуючи їх як схожі чи різні.

Потім використовуються певні підходи для додавання більше пікселів до початкових точок або ще більшого зменшення початкової точки до менших сегментів, які потім об'єднуються з іншими меншими початковими точками. Після цього найближчі сусіди розглядаються як початкові, і ці кроки повторюються до тих пір, поки не буде сегментовано все зображення. На основі цього методу існують дві основні методики.

Перший – це метод зростання регіонів. Він працює знизу вгору, коли сегментація розпочинається з меншого набору пікселів і поступово накопичує або ітеративно об'єднує їх на основі певних заздалегідь визначених обмежень схожості. Цей алгоритм починається з вибору довільного початкового пікселя на зображенні та порівняння його з сусідніми пікселями. Якщо є збіг або схожість у сусідніх пікселях, то вони

додаються до початкового пікселя, збільшуючи розмір області. Коли зростання цієї області не може продовжуватися далі, алгоритм вибирає інший початковий піксель, який обов'язково не належить до жодного з попередньо визначених регіонів, і починає процес знову. Методи зростання регіонів часто досягають ефективної сегментації, яка добре відповідає краям. Однак іноді, коли алгоритм дозволяє регіону зростати, перш ніж спробувати інші вихідні дані, це зазвичай зміщує сегментацію на користь регіонів, які сегментовані першими. Щоб протидіяти цьому ефекту, більшість алгоритмів спочатку вимагають вводити користувачем інформацію про схожість пікселів, так щоб жоден регіон не домінував та повністю не зростав, а декільком регіонам дозволялося розвиватися одночасно.

Методи сегментації, засновані на розділенні та об'єднанні, використовують дві основні методики, які виконуються разом – розділення регіонів та об'єднання регіонів для сегментації зображення. Розбиття включає ітераційне поділ зображення на області, що мають подібні характеристики, а об'єднання використовує об'єднання сусідніх областей, які дещо схожі один на одного. Після виконання процесу поділу є багато однаково позначених областей, розкиданих по всіх пікселях зображення. Щоб завершити процес, необхідно виконати об'єднання, яке після кожного розділення порівнює сусідні регіони, і, за необхідності, на основі ступенів схожості, об'єднує їх. Такі алгоритми також називаються *split-merge*.

Отже, цей підхід групує сусідні пікселі з подібними характеристиками в один об'єкт або регіон. Він ефективний у визначенні об'єктів з однорідними властивостями, але може бути неефективним у випадку перекриваючих областей та великого контрасту.

Даний алгоритм може бути ефективний для сегментації об'єктів на медичних зображеннях, якщо вони мають чітко визначені ознаки регіонів.

Сегментація вододілу. Вододіл – це метод на основі регіонів, який дотримується концепції топологічної інтерпретації. Розглядається аналогія з географічним ландшафтом з хребтами та долинами для різних компонентів зображення. Нахил і висота згаданої топографії чітко визначаються сірими значеннями відповідних пікселів, які називаються величиною градієнта. На основі цього тривимірного зображення, яке зазвичай використовується для ландшафтів Землі, перетворення вододілу розкладає зображення на області, які називаються басейнами водозбору. Для кожного локального мінімуму водозбірний басейн містить усі пікселі, чий шлях найкрутішого спуску сірих значень закінчується на цьому мінімумі.

Простими словами, алгоритм розглядає пікселі як локальну топографію, або висоту, часто ініціалізуючи себе з визначених користувачем маркерів. Потім алгоритм визначає так звані басейни, які є точками мінімуму, і, отже, басейни «заливаються» з маркерів до тих пір, поки басейни не зустрінуться на лініях водорозділу. Вододіли, які тут так утворені, відокремлюють басейни один від одного. Таким чином, зображення сегментується, тому що ми маємо пікселі, призначені кожному такому регіону або водорозділу.

Отже, цей метод використовує поняття мінімального та максимального рівня серед пікселів для розділення зображення на регіони. Він ефективний у визначенні об'єктів з градієнтами інтенсивності, але може бути чутливим до шуму та нерівномірності освітлення.

Даний алгоритм може бути корисним для сегментації складних структур, таких як судини або нервові волокна, але може не підходити для використання при сегментації пухлин.

Алгоритми сегментації на основі кластеризації. Алгоритми кластеризації відносяться до неконтрольованих методів, оскільки вони не вимагають попередньо визначеного набору функцій, класів або груп, на відміну від алгоритмів класифікації. Вони допомагають отримати основну,

приховану інформацію з даних, таких як структури, кластери та групи, які зазвичай невідомі з евристичної точки зору.

Методи, що ґрунтуються на кластеризації, проводять сегментацію зображень на кластери або групи пікселів зі схожими характеристиками. Завдяки основним властивостям кластеризації, елементи даних розподіляються на кластери таким чином, що елементи в одному кластері більш схожі один на одного, ніж в інших кластерах.

Сучасні процедури сегментації, що базуються на обробці зображень, часто використовують алгоритми кластеризації. Вони працюють краще за свої аналоги та можуть забезпечити досить хороший поділ за невеликий проміжок часу. Популярні алгоритми, такі як К-середніх, MeanShift та CamShift, є неконтрольованими методами, які працюють шляхом групування пікселів із загальними атрибутами, що належать до певного сегмента.

Алгоритм К-середніх (рисунок 1.5) є одним з найпростіших алгоритмів навчання без учителя, який може вирішити проблеми кластеризації загалом. Він класифікує дане зображення за допомогою певної кількості кластерів, фіксованих попередньо, та розділяє простір зображення на пікселі, що представляють центроїди групи. Потім кожен об'єкт призначається до групи на основі його відстані від кластера. Такий процес повторюється, доки центроїди не перестануть зміщуватися.

Алгоритм Fuzzy C Means, у свою чергу, дозволяє об'єднувати пікселі в більш ніж один кластер, дозволяючи кожному пікселю належати до більш ніж однієї групи з рівним рівнем схожості.

Отже, цей підхід групує пікселі зображення на кластери відповідно до подібності їхніх характеристик. Він ефективний у визначенні груп пікселів з подібними характеристиками, але може вимагати точного визначення кількості кластерів та параметрів кластеризації.

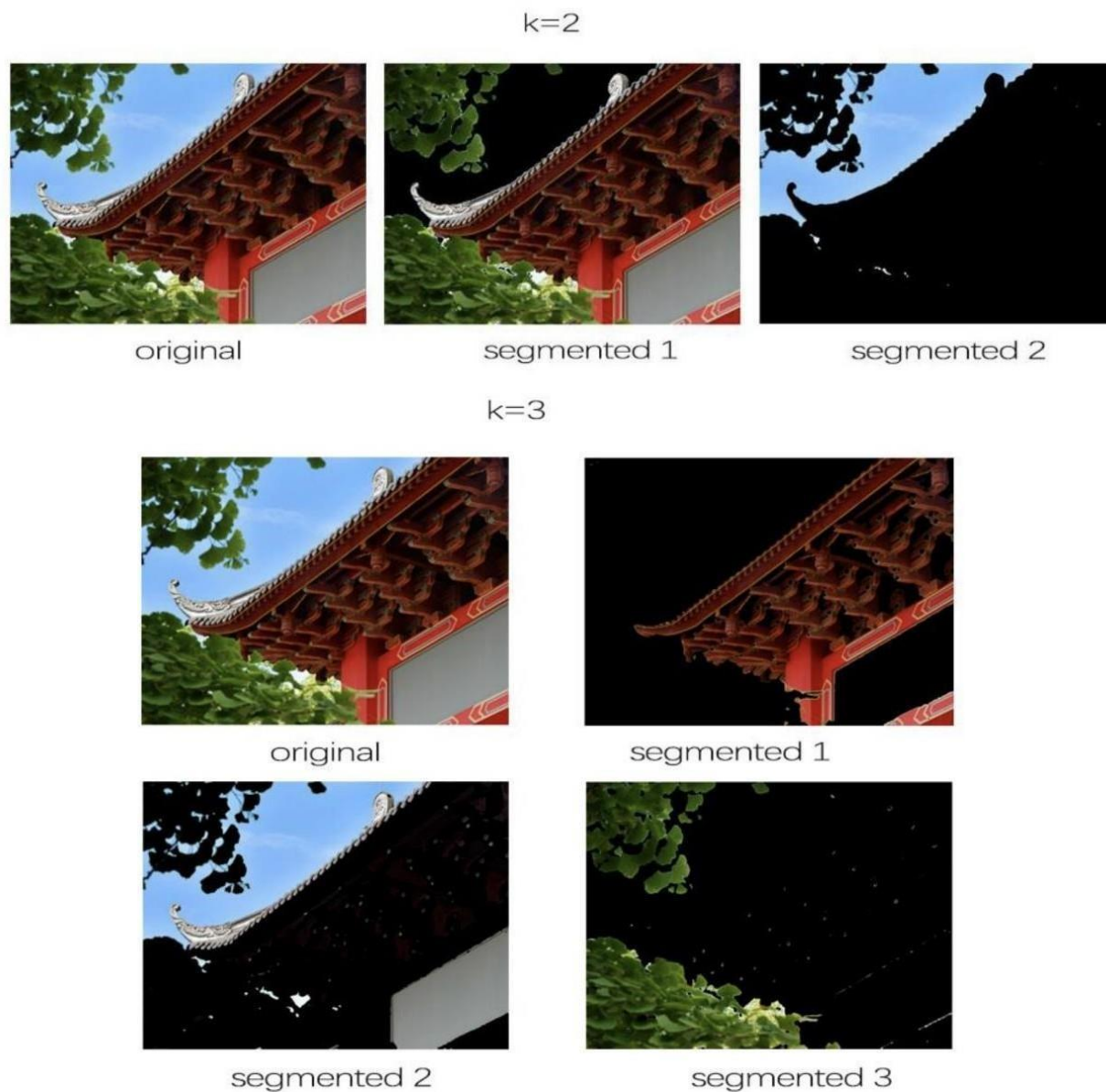


Рисунок 1.5 – Приклади зображень сегментованих методом К-середніх ($k = 2$, $k = 3$)

Даний алгоритм є стійким до шуму, що дозволяє ефективно відокремлювати об'єкти від фону навіть у випадку наявності шуму на зображенні.

Однак, для коректної сегментації потрібно визначити кількість кластерів, що може бути складно у випадку, коли кількість об'єктів на зображенні невідома або змінюється. А також, якщо об'єкти на зображенні

мають нерівномірний розподіл або перекриваються, алгоритми кластеризації можуть давати неправильні результати, що ускладнює сегментацію медичних зображень, де об'єкти можуть мати складну форму та різну інтенсивність.

Нейронні мережі для сегментації. Глибинне навчання широко використовується для обробки зображень, і наразі існує багато завдань у цифровій обробці зображень, для яких використовуються підходи на основі глибинного навчання. Цей метод представляє собою набір технік машинного навчання, які використовують глибокі штучні нейронні мережі – алгоритми, що беруть за приклад структуру та функціонування мозку. Завдяки глибинному навчанню можна створювати моделі, які постійно покращують свої результати навчаючись на прикладі. Останнім часом ця технологія привертає багато уваги через досягнення результатів у багатьох різних сферах, які раніше були неможливими.

Нейронна мережа (також її називають штучною нейронною мережею) – це адаптивна система, яка навчається за допомогою взаємопов'язаних вузлів або нейронів у шаруватій структурі, що нагадує людський мозок. Нейронні мережі можуть навчитися на даних – тому їх часто використовують для таких практичних задач, як розпізнавання закономірностей, класифікування даних чи прогнозування майбутніх подій.

Повнозв'язна нейронна мережа (рисунок 1.6) приймає на вхід вектор значень, тоді, коли зображення є матрицею значень пікселів (а у випадку кольорового зображення вона є тривимірною). Щоб передати зображення до такої моделі, матрицю треба перетворити в вектор, тобто записати вектори матриці один під одним.

Такий підхід застосовувався деякий час, однак він був здатний вирішувати відносно нескладні задачі. Для аналізу зображень більш успішно використовуються згорткові нейронні мережі.

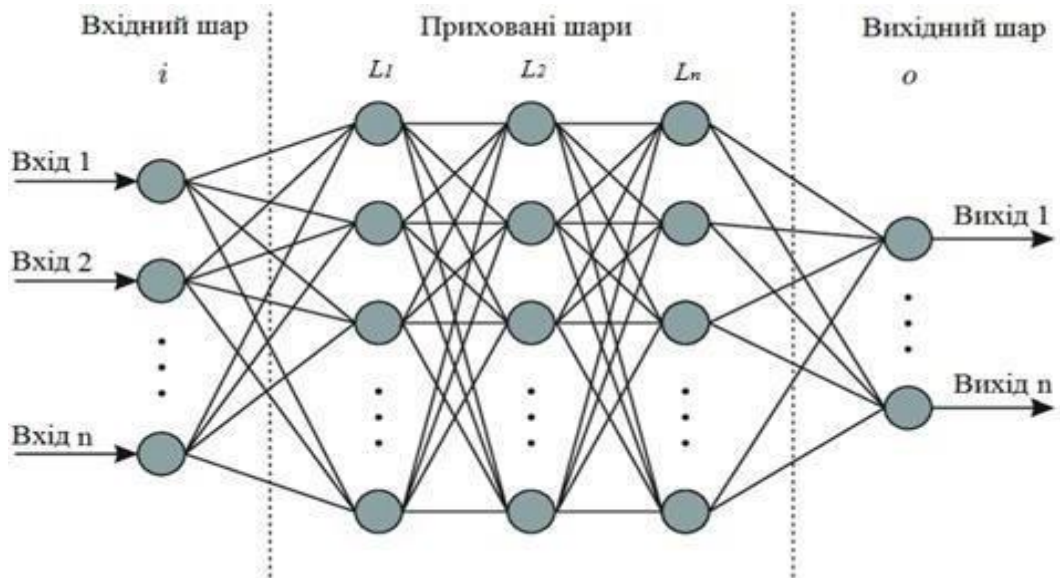


Рисунок 1.6 – Структура повнозв'язної нейронної мережі

Типова згорткова нейронна мережа складається з великої кількості шарів, серед яких чергуються шари згортки та шари субдискретизації. Згорткові шари застосовують операцію згортки до вхідної матриці, матриця ядра якої складається з невідомих ваг, які потім уточнюються у процесі навчання. У згортковому шару визначають розмір ядра, зсув, на який ядро зміщується при виконанні операції згортки, та функцію активації. Ідея полягає у побудові такого фільтра, застосування якого до всього зображення дає можливість знаходити деякі ознаки. Можна використовувати декілька різних ядер, в такому випадку на виході буде отримуватись тривимірна матриця. Це застосовується для знайдення більш складних ознак.

Початкові згорткові шари мережі звичайно тренуються знаходити прості ознаки, наприклад, межі об'єктів, тоді коли згорткові шари наприкінці мережі вже здатні вилучати більш складні, абстрактні ознаки.

Після використання шару згортки використовується шар субдискретизації. Він відповідає за зменшення просторового розміру матриці, отриманої на попередньому шарі згортки та здатний вилучати домінуючі риси, фільтруючи непотрібні деталі.

Отже, цей підхід використовує нейронні мережі для автоматичного виявлення та класифікації об'єктів на зображеннях. Він може виявляти складні зразки та структури, але потребує великої кількості даних для тренування та обчислювальних ресурсів.

Нейронні мережі широко використовуються в медичній сегментації через їхню здатність до автоматичного виявлення складних зразків та структурна зображеннях, які часто зустрічається в медичних дослідженнях.

2 ОГЛЯД АЛГОРИТМІВ ТА МЕТРИК ДЛЯ СЕГМЕНТАЦІЇ ПУХЛИН

2.1 Застосування гібридних зображень у сегментації пухлин

Гібридні зображення відіграють важливу роль у сегментації пухлин. Ці зображення отримуються шляхом комбінування різних модальностей зображень, таких як МРТ, КТ та інші, що надає більше інформації про внутрішню структуру пухлини та її оточуюче середовище. Гібридні зображення можуть покращити якість сегментації, дозволяючи алгоритмам точніше виділяти та аналізувати пухлину та суміжні тканини. Це особливо корисно для визначення розміру, форми та типу пухлини, а також для виявлення можливих ускладнень та оцінки ефективності лікування. Таким чином, застосування гібридних зображень у сегментації пухлин є важливим етапом у покращенні діагностики та лікування онкологічних захворювань.

Підходи до сегментації пухлин з використанням гібридних зображень надають можливість вдосконалити різні аспекти діагностики та лікування. Одним із ключових переваг використання гібридних зображень є підвищення контрастності та роздільної здатності, що сприяє кращій візуалізації структури пухлини та її взаємодії з навколишніми тканинами. Наприклад, поєднання МРТ та КТ дозволяє отримати комплексне зображення, яке відображає як м'які, так і тверді тканини, що може бути важливим при оцінці характеристик пухлини та її розмірів.

Крім того, застосування гібридних зображень в сегментації пухлин може сприяти виявленню різних структур та патологічних змін у тканинах, які можуть бути недостатньо видимими на окремих модальностях зображень. Наприклад, виявлення пухлини на МРТ-зображенні може бути покращене шляхом додавання інформації з КТ, що дозволяє збільшити чутливість та специфічність сегментації.

Більшість алгоритмів сегментації використовують глибоке навчання, яке є потужним інструментом для роботи з гібридними зображеннями.

Завдяки своїм можливостям автоматичного визначення зразків та властивостей пухлин, моделі глибокого навчання можуть ефективно адаптуватися до різних модальностей зображень та різних характеристик пухлин. Такі підходи можуть покращити якість сегментації та зменшити час, потрібний для ручної обробки зображень медичними фахівцями.

Для розуміння ефективності методів сегментації з використанням гібридних зображень необхідно також досліджувати вплив різних алгоритмів та параметрів обробки на якість результатів. Для цього проводяться порівняльні аналізи різних підходів та їхніх варіацій, що дозволяє визначити найбільш оптимальні стратегії сегментації для конкретних умов та типів пухлин.

Один із основних викликів у застосуванні гібридних зображень полягає у складності їх обробки та аналізу. Оскільки гібридні зображення поєднують інформацію з різних джерел, потрібно розробляти ефективні методи сегментації, які можуть адаптуватися до цієї різноманітності та забезпечити точні та надійні результати. Також важливо враховувати специфічні особливості кожного типу гібридного зображення та розробляти індивідуалізовані підходи до їх обробки та аналізу.

Наприклад, використання методів машинного навчання та штучних нейронних мереж може забезпечити ефективний та автоматизований процес сегментації пухлин на гібридних зображеннях. Шляхом навчання моделей на великому обсязі даних, вони можуть виявляти складні взаємозв'язки та патерни на зображеннях, що допомагає підвищити точність сегментації та знизити кількість помилок.

Крім того, розвиток нових методів обробки та аналізу гібридних зображень може включати в себе використання технік комп'ютерного зору,

обробки сигналів та інші підходи, спрямовані на виокремлення корисної інформації та відхилення від шуму та артефактів на зображеннях.

У загальному, використання гібридних зображень у сегментації пухлин є обіцяним напрямком у дослідженні та практиці медичної обробки зображень. Розвиток ефективних методів сегментації на основі цих зображень може покращити якість діагностики та лікування онкологічних захворювань, сприяти розвитку нових технологій та підвищити ефективність медичної практики.

Отже, гібридна візуалізація – це злиття двох (або більше) методів візуалізації для створення нової техніки. Завдяки синергетичному поєднанню вроджених переваг об'єднаних технологій візуалізації зазвичай з'являється нова, більш потужна модальність.

Існуючі гібридні методи візуалізації включають: ПЕТ-КТ; ОФЕКТ-КТ; ПЕТ-МРТ; ОФЕКТ-МРТ; УЗД та МРТ; УЗД та КТ; МРТ та КТ; ангиографія та КТ. Розглянемо детальніше деякі із наведених вище модальностей та яким чином вони отримуються.

Позитронно-емісійна томографія (ПЕТ) – це метод візуалізаційного дослідження, який використовує радіоактивні речовини, щоб отримати детальні зображення внутрішніх органів та тканин. Під час ПЕТ сканування пацієнту вводиться речовина-маркер, яка містить радіоактивний ізотоп. Цей ізотоп розпадається, видаючи позитрони, які зіштовхуються з електронами в тілі пацієнта, утворюючи гамма-кванти. Детектори реєструють ці гамма-кванти, а потім комп'ютер обробляє інформацію, щоб створити детальне зображення внутрішніх органів та тканин, включаючи їхню функціональну активність. ПЕТ-сканування широко використовується в онкології для виявлення ракових пухлин, оцінки їх розміру, розподілу та відслідковування ефективності лікування.

Комп'ютерна томографія (КТ) – це метод візуалізаційного дослідження, який використовує рентгенівські промені та комп'ютерну

обробку для отримання детальних зображень внутрішніх органів та тканин. Під час КТ сканування пацієнт лежить на столі, який потім переміщується через отвір у томографічному апараті. Під час цього рентгенівські промені обертаються навколо пацієнта, отримуючи ряд зображень різних площин тіла. Ці зображення потім обробляються комп'ютером, що дозволяє отримати детальні перетинні зрізи внутрішніх органів та тканин. КТ широко використовується для діагностики різноманітних захворювань, включаючи травми, захворювання легенів, мозку, органів черевної порожнини та інших.

Однофотонна емісійна комп'ютерна томографія (ОФЕКТ) – це метод візуалізаційного дослідження, що використовує радіоактивні речовини, що емітують гамма-кванти зсередини тіла людини, що дає можливість отримати детальну інформацію про функціональну активність тканин та органів в організмі. У процесі ОФЕКТ реєструються гамма-кванти, які випромінюються радіонуклідами, що вводяться в пацієнта. Це дозволяє створити тривимірне зображення внутрішньої структури об'єкта, що досліджується, та виявити будь-які аномалії або патологічні процеси. ОФЕКТ широко використовується в онкології, кардіології, неврології та інших галузях медицини для діагностики захворювань і планування лікування.

МРТ (магнітно-резонансна томографія) – це нетравматичний метод візуалізаційного дослідження, який використовує магнітні поля та радіохвилі для створення детальних зображень внутрішніх органів та тканин у тілі. Під час МРТ пацієнт поміщається у спеціальний трубчастий апарат, де відбувається взаємодія магнітних полів і радіохвиль. Зображення, отримані за допомогою МРТ, дозволяють отримати докладну інформацію про структуру та функції органів, а також виявити патологічні зміни, такі як пухлини, запальні процеси, травми та інші захворювання. МРТ широко

використовується в онкології, неврології, ортопедії, кардіології та інших галузях медицини для точної діагностики та планування лікування.

УЗД (ультразвукове дослідження) – це метод візуалізаційного дослідження, який використовує високочастотні акустичні хвилі для створення зображень внутрішніх органів та тканин у тілі. Під час УЗД лікар за допомогою спеціального пристрою (трансдюсера) сканує тіло у площині під ним. Акустичні хвилі проникають у тіло, відбиваються від органів та тканин та повертаються до трансдюсера, де перетворюються на зображення. Отримані зображення дозволяють лікарю оцінити структуру органів, виявити патологічні зміни, які можуть бути пухлинами, каменями, запальними процесами та іншими відхиленнями. УЗД є безболісним, безпечним та неінвазивним методом діагностики, що широко використовується для скринінгу та оцінки різних захворювань, таких як вагітність, захворювання органів черевної порожнини, суглобів, м'язів та інших структур.

Ангіографія – це метод візуалізаційного дослідження, що використовує рентгеноконтрастні або магнітоконтрастні речовини для отримання детальної інформації про стан судин в організмі. Цей метод застосовується для діагностики захворювань судин, таких як атеросклероз, тромбози, аневризми та інші патологічні стани, а також для планування хірургічних втручань та інтервенцій на судинах. Ангіографія може бути проведена в різних частинах тіла, включаючи мозкові, серцеві, ниркові та інші судинні системи.

Загальні переваги гібридної візуалізації включають підвищену точність діагностики, подальший крок до індивідуалізованої медицини, точний моніторинг інтервенційних процедур, зменшення променевого навантаження, наприклад, динамічне ультразвукове дослідження після отримання карти КТ.

2.2 Аналіз алгоритмів сегментації пухлин

Сегментація медичних зображень є однією з задач комп'ютерного зору, що включає в себе поділ медичного зображення на різні сегменти, де кожен із них представляє певний об'єкт або структуру інтересу. Основною метою сегментації таких зображень є точне виявлення таких об'єктів та структур інтересу для постановки діагнозу та планування лікування.

Розвиток ефективних та дієвих методів сегментації медичних зображень відкриває безліч нових можливостей у сфері клінічної практики:

- персоналізація та оптимізація сканування. Завдяки точній індивідуалізації області інтересу на медичних зображеннях для кожного пацієнта, ми можемо зменшити ризик передавання зайвих доз опромінення, не втрачаючи при цьому цінної інформації. Ефективне виявлення та аналіз органів на попередніх зображеннях дозволяє оптимізувати процес сканування, зменшуючи час, вартість та ризики для пацієнтів;

- автоматизована діагностика та планування лікування. Автоматизована сегментація медичних зображень значно спрощує роботу лікарів під час діагностики та планування лікувальних заходів. Швидке та точне визначення уражених регіонів дозволяє прискорити процес встановлення діагнозу, знижуючи вартість медичних досліджень та підвищуючи ефективність лікування;

- кількісна діагностика та дослідження. Автоматизована сегментація зображень сприяє швидкому та точному вимірюванню обсягів органів, що є важливим для кількісної діагностики хвороб. Забезпечуючи лікарів необхідними даними та параметрами, сегментація сприяє подальшому розвитку нових методів та моделей дослідження.

Відтак, розвиток технологій візуалізації медичних даних вимагає постійного вдосконалення методів аналізу, зокрема, застосування новітніх підходів глибокого навчання та сегментації зображень.

Одним із типів сегментації медичних зображень є сегментація пухлин. Завдання сегментації пухлин полягає в визначенні її просторового розташування на медичному зображенні. У такому типі сегментації кожен піксель класифікується як той, який містить пухлину, чи ні.

Останнім часом медична сегментація зображень зазнала великого прогресу за допомогою глибокого навчання. Зокрема, згорткові нейронні мережі (CNN), особливо повністю згорткові мережі (FCN), виявили себе як домінуючі у медичній сегментації. З розвитком медичної сегментації зображень серед різних варіантів моделей U-Net став фактичним вибором, що складається з симетричної мережі енкодера-декодера із з'єднаннями для збереження деталей. На основі цієї нейронної мережі можуть автоматично виділятися характеристики зображення, які можуть бути використані в задачах сегментації. Декілька моделей глибокого навчання вже були використані з відмінними результатами у медичній сегментації зображень, такі як U-Net, UNet, 3D U-Net, V-Net, Attention-UNet, TransUNet та Swin-UNet.

Розглянемо одну із найбільш популярних мережевих структур для сегментації, а саме U-Net.

Архітектура U-Net – це архітектура згорткових нейронних мереж (англ. CNN), яка використовується для задач сегментації зображень. Спочатку розроблена для біомедичної сегментації зображень, вона знайшла застосування в різних сферах. Архітектура U-Net обробляє локальну та глобальну інформацію і має U-подібну структуру енкодера-декодера [10].

Архітектура мережі показана на рисунку 2.1. Загалом вона складається із трьох частин: стиснення, вузького місця та розширення. Частина стиснення відповідає типовій архітектурі згорткової мережі. Кожен блок приймає вхідні дані, застосовує два згорткових шари 3×3 , за якими слідує випрямлений лінійний вузол (англ. ReLU) та максимізаційне агрегування (англ. max pooling) розміром 2×2 для зменшення вибірки.

Найнижчий шар є посередником між частиною стиснення та розширення. Він складається із двох згорткових шарів 3×3 , за яким слідує шар збільшення розширення (англ. up-convolution) 2×2 . В частині розширення використовуються шари розгортки (англ. up-convolution). Кожен шар передає дані двом згортковим шарам 3×3 , за якими слідує випрямлений лінійний вузол та шар збільшення розширення 2×2 . Кількість блоків в частині розширення відповідає кількості блоків в частині стиснення зображення, і до кожного блоку в частині розширення на вхід надсилається карта ознак (англ. feature map) з відповідного блоку в частині стиснення. Такий процес забезпечує те, що функції, які були вивчені під час стиснення зображення будуть використанні під час його розширення, тобто реконструкції.

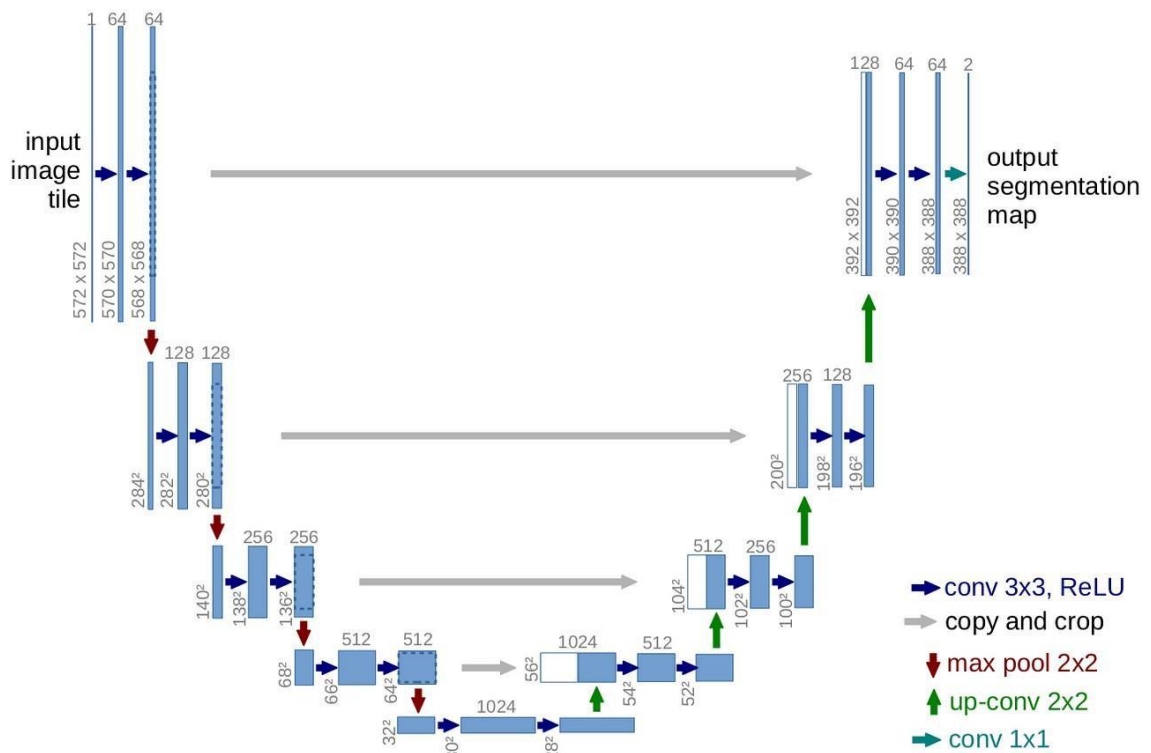


Рисунок 2.1 – Архітектура U-Net

В останньому шарі використовується згортка 1×1 для відображення кожного 64-компонентного вектора ознак на потрібну кількість класів. Загалом мережа містить 23 згорткові шари.

На останньому шарі використовується згортка 1×1 для відображення кожного 64-компонентного вектора ознак на потрібну кількість класів. Загалом мережа має 23 згорткових шари [14].

Насправді, архітектура U-Net має кілька переваг для задач сегментації зображень. По-перше, її U-подібна конструкція дозволяє поєднувати низькорівневі та високорівневі функції, що забезпечує кращу локалізацію об'єктів. По-друге, з'єднання між блоками стискування та розширення допомагають зберегти просторову інформацію, що забезпечує більш точну сегментацію. Нарешті, архітектура U-Net має відносно невелику кількість параметрів, що робить її більш ефективною з точки зору обчислень, ніж інші архітектури. Тому, можна зробити висновок, що мережа U-Net є кращою ніж згорткова нейронна мережа (англ. CNN), оскільки через наведені вище переваги дозволяє краще сегментувати дрібні деталі, навіть з обмеженим обсягом даних.

Також варто зазначити, що можна виділити декілька модифікацій мережі U-Net, наприклад, такі як 3D U-Net, U-Net++, R2U-Net, Graph U-Net.

Розглянемо детальніше мережу 3D U-Net та чим вона відрізняється від мережі U-Net.

Як можна побачити з рисунку 2.2 архітектура мережі 3D U-Net дуже схожа до мережі U-Net. Відмінність даних мереж полягає в тому, що 3D U-Net може обробляти 3D зображення з тією ж U-мережевою структурою.

Переваги використання даної мережі полягають найперше в тому, що, об'ємні дані часто зустрічаються в біомедичному аналізі даних. Анотування таких даних за допомогою сегментаційних міток викликає труднощі, оскільки на екрані комп'ютера можна показати лише 2D-зрізи. Таким чином, анотування великих об'ємів даних у режимі «зріз за зрізом» є дуже

виснажливим. Це також неефективно, оскільки сусідні зрізи показують майже однакову інформацію. Особливо для підходів, заснованих на навчанні, які вимагають значної кількості анотованих даних, повне анотування 3D-об'ємів не є ефективним способом створення великих і багатих навчальних наборів даних, які б добре узагальнювали.

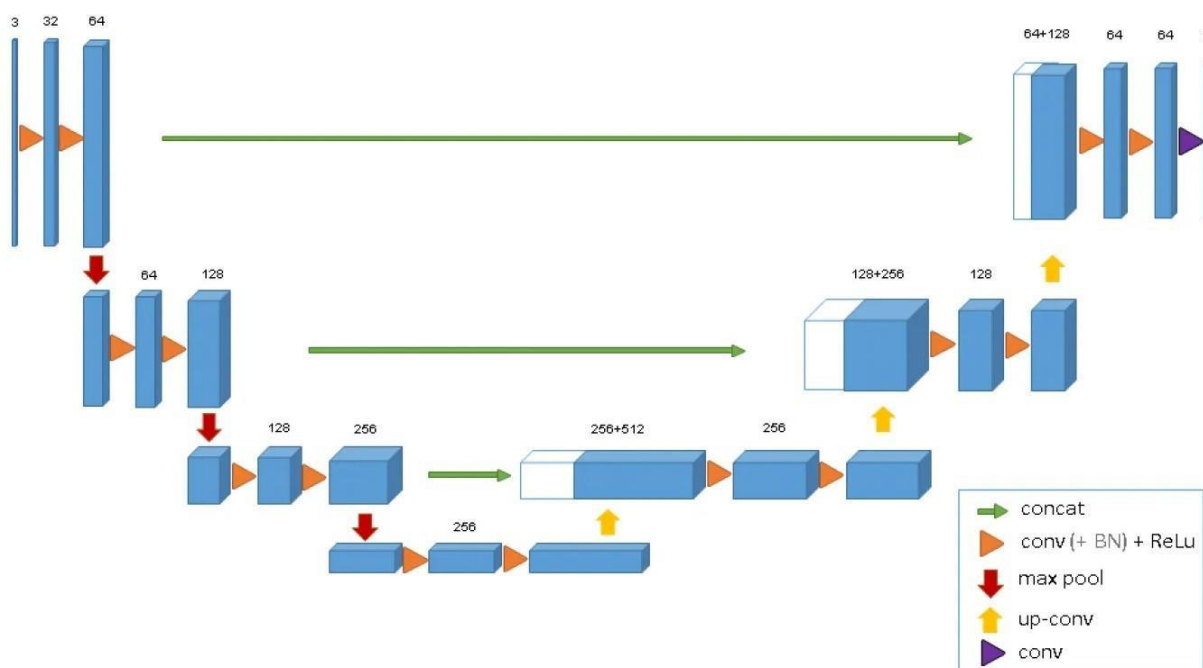


Рисунок 2.2 – Архітектура 3D U-Net

Як вже було зазначено вище, мережа базується на архітектурі U-Net, яка складається із стискаючої частини, тобто енкодера, для аналізу всього зображення, та послідовної розширюючої частини, тобто декодера, для створення сегментації з повною роздільною здатністю. В той час, як мережа U-Net є повністю 2D-архітектурою, мережа 3D U-Net приймає на вхід тривимірні об'єми і обробляє їх відповідними 3D-операціями, зокрема, 3D-згортками, шарами максимізаційного агрегування та 3D розгортки. Крім того, в архітектурі даної мережі використовується нормалізація, а саме Batch normalization для швидшої збіжності.

У багатьох біомедичних додатках для навчання мережі, яка досить добре узагальнює, потрібно дуже мало зображень. Це пов'язано з тим, що кожне зображення вже містить повторювані структури з відповідними варіаціями. Для об'ємних зображень цей ефект ще більш виражений, тому можна навчити мережу лише на двох об'ємних зображеннях, щоб узагальнити третє. Зважена функція втрат і спеціальна попередня обробка даних дозволяють навчати мережу лише на невеликій кількості анотованих вручну зрізів, тобто для навчання мережі буде використано якомога менше даних.

Також розглянемо більш детально архітектуру мережі UNet++. Архітектура мережі UNet++ була запропонована Чжоу у 2018 році, ввівши концепцію щільних з'єднань в мережу U-Net. Структура моделі зображена на рисунку 2.3.

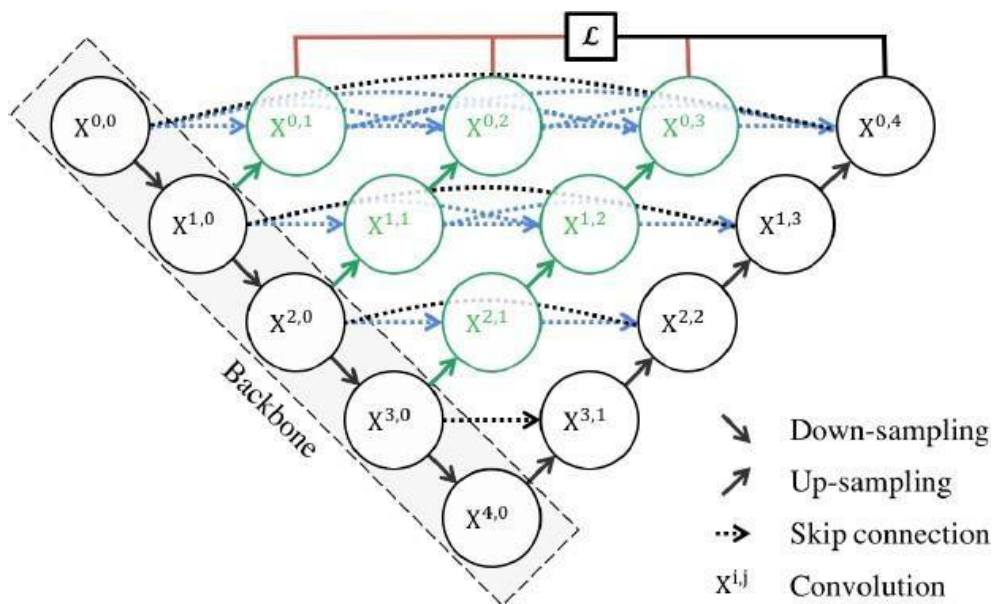


Рисунок 2.3 – Архітектура U-Net++

Загалом, UNet++ базується на моделі U-Net, зберігаючи довгі з'єднання з 4-ма пропусками та додаючи більше шляхів з короткими пропусками та блоки з шарами розгортки для формування нових рівнів

енкодерів. U-подібна структура з'єднання в UNet++ досягається шляхом об'єднання кожного енодера в декодері з іншими енодерами на тому ж рівні.

Зокрема, кожен енодер отримує карти ознак того ж масштабу від інших енодерів і об'єднує їх для отримання більш дискримінативного представлення ознак. Крім того, пізніше був доданий механізм уваги до UNet++, який покращує конкатенацію карт ознак, фокусування та вилучення важливих ознак під час злиття енодерів. UNet++ фіксує ознаки з різних рівнів, вводючи щільні зв'язки, що дозволяє витягувати інформацію про ознаки з різних шарів і масштабів. Ці ознаки інтегруються в остаточний прогноз для підвищення точності сегментації.

Загалом, ідея щільних зв'язків виникла в мережі DenseNet, до виникнення таких зв'язків еволюція згорткових нейронних мереж зазвичай передбачала збільшення глибини або ширини мереж. DenseNet представив нову структуру шляхом повторного використання функцій, що не тільки полегшило проблему зникнення градієнта, але й зменшило кількість параметрів моделі.

Наступна мережа яку б було варто розглянути – це TransUNet.

Основний недолік CNN те, що вони можуть використовувати тільки інформацію найближчого радіуса дії через невеликий розмір ядра згортки, що не дозволяє ефективно досліджувати дані в задачах, які містять компоненти із залежностями на великих відстанях. Тому в даній мережі було запропоновано використання трансформерів, які часто використовуються для обробки природної мови (англ. NLP), та можуть ефективно досліджувати інформацію на далеких відстанях, проте вони не є такими надійними при роботі з даними поблизу, як це наприклад роблять CNN.

Отже, для того, щоб використовувати можливості CNN для компенсації недоліків трансформерів Ченом та іншими у 2021 році була

запропонована архітектура мережі TransUNet як мережа сегментації зображень на основі трансформера. Також варто додати те, що автори впевнилися в доцільності використання CNN та трансформерів, спочатку намагавшись використовувати чисту архітектуру трансформера для сегментації зображення. Проте дана архітектура не показала такі ж хороші результати як архітектура з додаванням CNN, оскільки трансформери не такі хороші під час експлуатації локальних можливостей як CNN.

Структура моделі зображена на рисунку 2.4. TransUNet базується на мережі U-Net і аналогічним чином включає в себе кодер та декодер для кодування та декодування інформації про зображення з метою отримання сегментації.

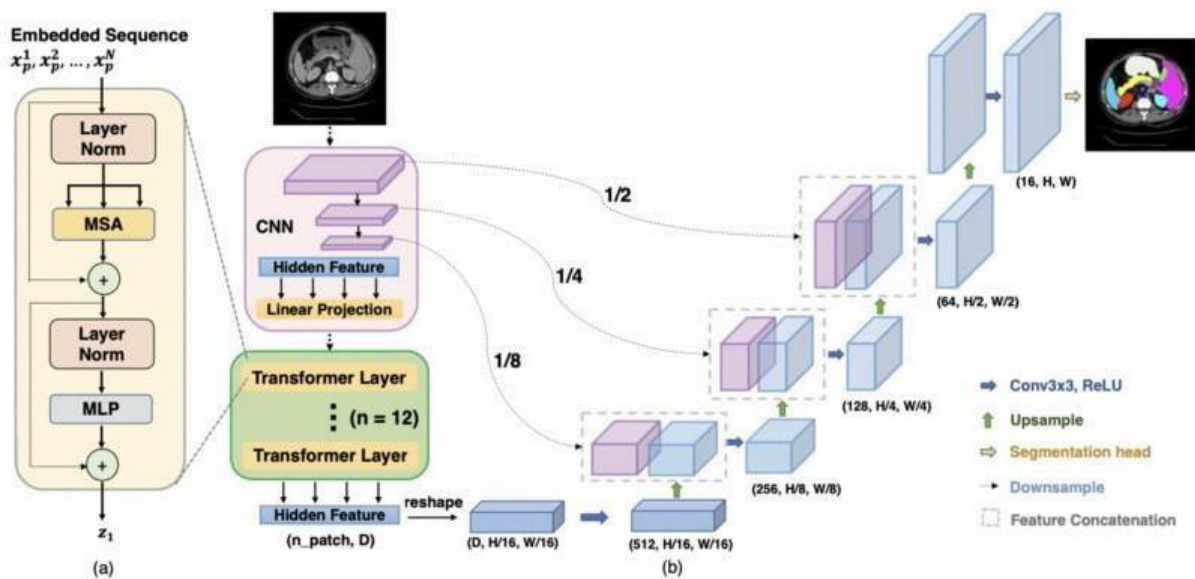


Рисунок 2.4 – Архітектура TransUNet

Різниця даної мережі із мережею U-Net полягає в тому, що TransUNet використовує гібридну архітектуру CNN-трансформер в якості кодера для отримання інформації щодо просторових параметрів високої роздільної здатності від CNN, так і відомостей глобального контексту від трансформерів.

Загалом в частині кодування використовується CNN для вилучення ознак при створенні їхньої карти для вхідного сигналу, як показано в рожевій рамці на рисунку 2.4.

Для кожного рівня екстрактора ознак їхня вихідна карта (яка кодує проміжні карти ознак високого рівня) конкатенується в маршрут декодера того самого рівня, як показано пунктирними стрілками на тому ж рисунку. Потім карта властивостей токенизується (векторизується) у двовимірний ембеддінг форми (n_patch, D) шляхом лінійного проектування, де D – його повна довжина. Ембеддінги будуть зберігати позиційну інформацію карти ознак.

Після отримання ембеддінгів вони подаються на 12 шарів трансформера для кодування меншої кількості інформації на малих відстанях і більшої кількості на великих дистанціях. Кожен шар використовує модулі багатоголової самоуважності (англ. multi-head self-attention) та багат шарового перцептрона (англ. multi-layer perceptron). Модуль багатоголової самоуважності є основним будівельним блоком для трансформерів, а багат шаровий перцептрон складається з декількох повністю зв'язаних шарів. Нарешті, щоб бути готовим до апсемплінгу, тобто до частини розширення, вихідний сигнал перетворюється в $(D, \frac{H}{16}, \frac{W}{16})$, де $\frac{H}{16}$ та $\frac{W}{16}$ означають, що висота і ширина вже зменшилися в 16 разів внаслідок попередніх операцій.

Архітектура частини розширення є достатньо простою. По-перше, вхідні дані кодера CNN-трансформера обробляються згортковим шаром 3×3 використовуючи активацію ReLU, проходять через шар розширення, а потім конкатенуються з виходом екстрактора ознак CNN третього рівня. Отримані карти ознак знову проходять через згортку 3×3 із шаром активації ReLU, і результати конкатенуються з даними, отриманими за допомогою екстрактора властивостей CNN другого рівня.

Такий крок повторюється знову, лише тепер на виході отримується маска у формі (C, H, W) , де C – кількість об'єктивних класів, H – висота зображення, W – ширина зображення. [16].

Однак у порівнянні з традиційними U-подібними моделями, які використовують згорткові нейронні мережі, TransUNet вводить стек з 12 модулів Transformer, що значно збільшує кількість параметрів, що збільшує складність навчання моделі.

Розглянемо також іншу модифікацію мережі, яка називається Swin-Unet.

Архітектура мережі була запропонована Цао та іншими у 2023 році. Структура даної мережі зображена на рисунку 2.5. На відміну від TransUNet, який замінює згорткові 6 блоків в кодери U-Net на блоки Transformer, Swin-Unet використовує блоки Swin Transformer для вилучення ієрархічних ознак із вхідного зображення. Swin-Unet є першою U-подібною архітектурою, що базується виключно на трансформерах. Swin Transformer розширює одновимірну послідовність традиційного трансформера до двовимірних блоків зображення і використовує ієрархічний механізм уваги для захоплення ознак на більшому рецептивному полі.

Ця структура схожа на ієрархічну структуру в згорткових нейронних мережах і слугує для виділення ознак. Крім того, Swin Transformer вводить механізм зсуву вікон поверх механізму самоуваги. Обмежуючи розрахунок уваги вікнами, розташованими поблизу поточної області, Swin-Unet краще зберігає просторову інформацію і ще більше покращує продуктивність моделі. У Swin-Unet Swin Transformer застосовується в модулях кодування, пошуку вузьких місць та декодування.

Важливо, що стиснення характеристик кожного шару в Swin-Unet є меншим, ніж у TransUNet. Замість того, щоб додавати додаткові модулі

трансформації, Swin-Unet замінює модулі згортки на модулі трансформації, ефективно зменшуючи кількість параметрів моделі [17].

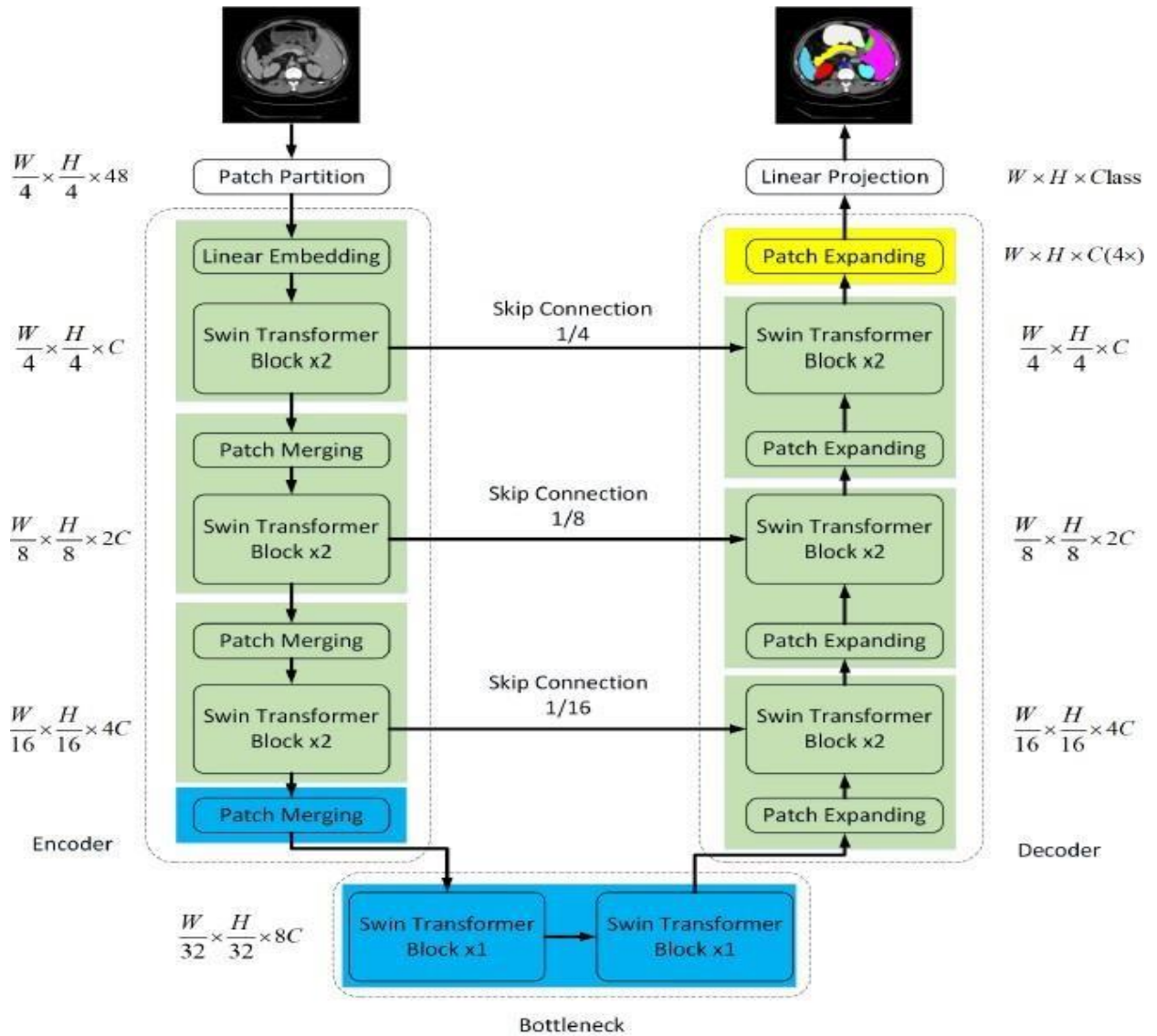


Рисунок 2.5 – Архітектура Swin-Unet

2.3 Вибір метрик оцінки моделі сегментації пухлин

Усі найбільш поширені метрики оцінки моделей сегментації медичних зображень, включаючи також сегментацію пухлин, базуються на обчисленні матриці помилок для бінарної маски сегментації (рисунок 2.6).

		Реальні значення	
		Область з пухлиною	Фон
Прогноз моделі	Область з пухлиною	TP	FP
	Фон	FN	TN

Рисунок 2.6 – Матриця помилок

Така маска сегментації складається з пікселів, де кожен піксель відображений як 1, якщо містить область з пухлиною, та 0 – якщо ні.

Розглянемо детальніше кожен елемент матриці помилок.

TP (True Positive – істинно позитивні) відповідає кількості пікселів, які були правильно класифіковані як такі, що містять область з пухлиною.

FP (False Positive – хибно позитивні) відповідає кількості пікселів, які представляють фон, та були хибно класифіковані як такі, що містять область з пухлиною.

FN (False Negative – хибно негативні) відповідає кількості пікселів, які містять область з пухлиною, та було хибно класифіковані як такі, що представляють фон.

TN (True Negative – істинно негативні) відповідає кількості пікселів, які були правильно класифіковані як такі, що представляють фон.

Розглянемо метрики оцінки, які можна використовувати для оцінки сегментації пухлин.

Точність обчислюється як відношення кількості правильно класифікованих пікселів, які належать до області з пухлиною, до загальної кількості пікселів, що були класифіковані як область з пухлиною:

$$Precision = \frac{TP}{TP+FP}. \quad (2.1)$$

При використанні точності ми оцінюємо, наскільки правильно модель визначає пікселі як такі, що належать області з пухлиною. Це особливо корисно, коли нам потрібна висока точність у визначенні самої пухлини, незалежно від того, чи модель визначає деякі пікселі, що входять в область з пухлиною як фон.

Повнота визначається як відношення кількості правильно класифікованих пікселів, які належать до області з пухлиною, до загальної кількості пікселів, які дійсно належать до області з пухлиною:

$$Recall = \frac{TP}{TP+FN}. \quad (2.2)$$

При використанні повноти ми оцінюємо, наскільки ефективно модель виявляє всі пікселі, що належать до області з пухлиною. Це важливо, коли ми хочемо, щоб модель виявляла якомога більше пікселів, що належать області з пухлиною, навіть якщо при цьому будуть включені пікселі, які її не містять.

Оцінка точності (або індекс Ренда) – кількість правильно класифікованих пікселів як таких, що містять область з пухлиною або фон, поділена на загальну кількість пікселів зображення.

$$Accuracy = \frac{TP+TN}{TP+TN+FN+FP}. \quad (2.3)$$

Однак використання оцінки точності при аналізі результатів медичних зображень не рекомендується, оскільки в таких зображеннях спостерігається значна незбалансованість між класами об'єктів інтересу та фоном. Інакше кажучи, на цих зображеннях лише незначний відсоток

пікселів відноситься до областей інтересу, тоді як решта пікселів відноситься до фону. Оскільки оцінка точності включає істинно негативні результати, то така міра завжди буде відображати штучно завищені результати.

F -міра – одна із найпоширеніших метрик, використовуваних для аналізу результатів моделей комп'ютерного зору та сегментації медичних зображень. На основі F -міри виділяють декілька широко використовуваних метрик.

Коефіцієнт подібності Дайса, також відомий як міра $F1$, найбільш використовувана метрика у більшості наукових публікацій для оцінки сегментації медичних зображень;

Міра $F1$ є середнім гармонійним точності та повноти. Іншими словами, ця міра обчислюється як подвоєний перетин передбаченої сегментації та дійсної, поділений на загальну кількість пікселів в обох зображеннях:

$$Dice = \frac{2TP}{2TP+FN+FP}. \quad (2.4)$$

Дана формула надає однакову вагу точності і повноти, тому F -міра буде спадатиме однаково при зменшенні як точності, так і повноти.

Можливо розрахувати F -міру надавши різну вагу точності і повноті для того, щоб збільшити пріоритет однієї з метрик в порівнянні з іншою:

$$F_{\beta} = \frac{(1+\beta^2) \cdot TP}{(1+\beta^2) \cdot TP + \beta^2 \cdot FN + FP}, \quad (2.5)$$

де β приймає значення в діапазоні $0 < \beta < 1$ якщо пріоритет віддається точності, а при $\beta > 1$ пріоритет віддається повноті;

При $\beta = 1$ формула зводиться до попередньої і отримується збалансована F -міра.

Індекс Жаккара (коефіцієнт подібності Жаккара) – це перетин передбаченої сегментації та дійсної поділений на їх об'єднання;

$$IoU = \frac{TP}{TP+FN+FP}. \quad (2.6)$$

Коефіцієнт подібності Дайса та індекс Жаккара є найбільш поширеними метриками для сегментації медичних зображень, оскільки обидві метрики штрафують хибні позитивні результати, що є типовою ознакою для високо незбалансованих наборів даних, до яких і відноситься сегментація медичних зображень. Вибір між цими двома метриками залежить від конкретної задачі, проте варто зауважити, що коефіцієнт подібності Жаккара штрафує недосегментацію та перевищену сегментацію більше, ніж коефіцієнт подібності Дайса.

3 ОСОБЛИВОСТІ РЕАЛІЗАЦІЇ ПРОГРАМНОГО ЗАСТОСУНКУ

3.1 Огляд вибраного набору даних

Для початку коротко опишемо реалізований застосунок.

В результаті дослідження алгоритму сегментації пухлин за гібридними зображеннями було реалізовано програмний застосунок, що сегментує цифрові зображення МРТ головного мозку розроблено в інтегрованому середовищі розробки PyCharm з використанням мови програмування python та бібліотеки Keras для роботи з нейронними мережами та бібліотеки TensorFlow для реалізації та навчання глибоких нейронних мереж. Мінімальні системні вимоги: 64-бітна операційна система Windows, Python 3.6.0.

Вхідні дані: цифрові зображення формату NIfTI.

Вихідні дані: сегментовані цифрові зображення МРТ головного мозку.

Для реалізації моделі сегментації пухлин був використаний набір даних BraTS2020 (Brain Tumor Segmentation 2020).

Цей набір даних містить МРТ-сканування пухлин головного мозку, а саме кожен пацієнт у BraTS2020 представлений чотирма різними МРТ-скануваннями (модальностями), які називаються T1, T1CE, T2 та FLAIR відповідно.

Усі набори зображень були сегментовані вручну одним або кількома оцінювачами, використовуючи один й той самий протокол анотації, а їхні анотації були схвалені досвідченими нейрорадіологами. Кожне сегментоване зображення містить від 1 до 4 класів, проте третій клас відсутній. Тому четвертий клас має бути перейменований на третій під час обробки даних цього набору.

Кожному класу відповідає певний колір як показано на рисунку 3.1.

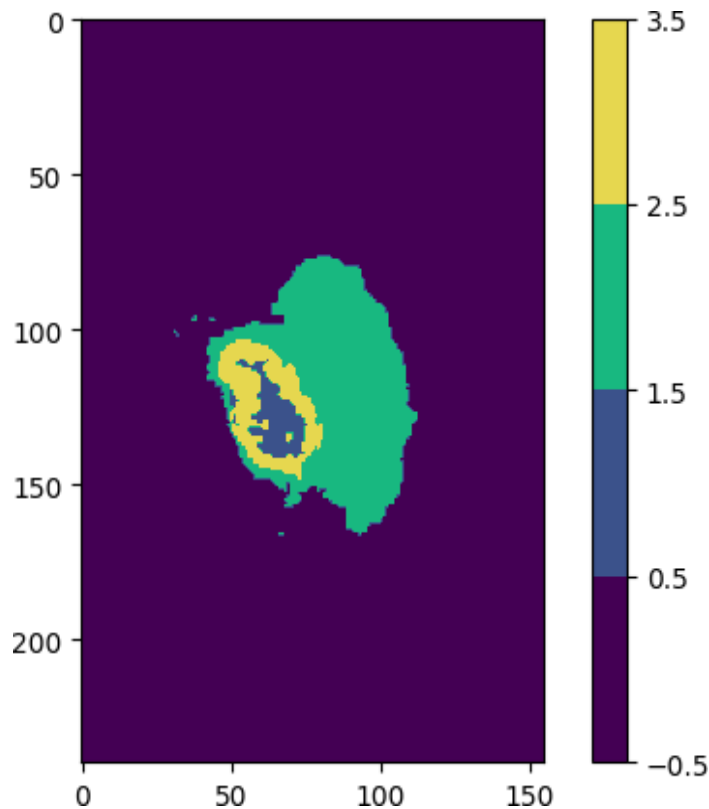


Рисунок 3.1 – Відображення сегментації МРТ-сканування пацієнта

Розглянемо детальніше кожний з класів:

- клас 0 – представлений фіолетовим кольором і відображає область, що відповідає непухлинній, тобто або фону зображення, або здоровій зоні;
- клас 1 – представлений синім кольором і відображає область, що містить некротичну пухлину, яка не збільшується;
- клас 2 – представлений зеленим кольором і відображає область, що містить перитуморальний набряк;
- клас 4 – представлений жовтим кольором і відображає область, що відповідає зоні підвищення пухлини.

Отже, анотації включають вище наведені класи, як описано в статті BraTS 2012-2013 року в MIT та в останній підсумковій статті BraTS. Надані дані спільно реєструються в одному анатомічному шаблоні, інтерполюються до тієї ж роздільної здатності (1 мм³) і позбавляються мозкового шару.

Кожна з модальностей МРТ-сканування (рисунок 3.2) відображає різні аспекти, а саме:

– T1: послідовність імпульсів T1-зваженого зображення є ключовою у магнітно-резонансній томографії і відображає відмінності в сигналі на основі внутрішнього часу релаксації T1 різних тканин. Клінічно T1-зважені зображення, як правило, найкраще показують нормальну анатомію і можуть виділяти патологію, особливо за використання контрасту гадолінієм. Наприклад, жир має високу інтенсивність сигналу на T1-зважених зображеннях, тоді як рідина має низьку інтенсивність сигналу;

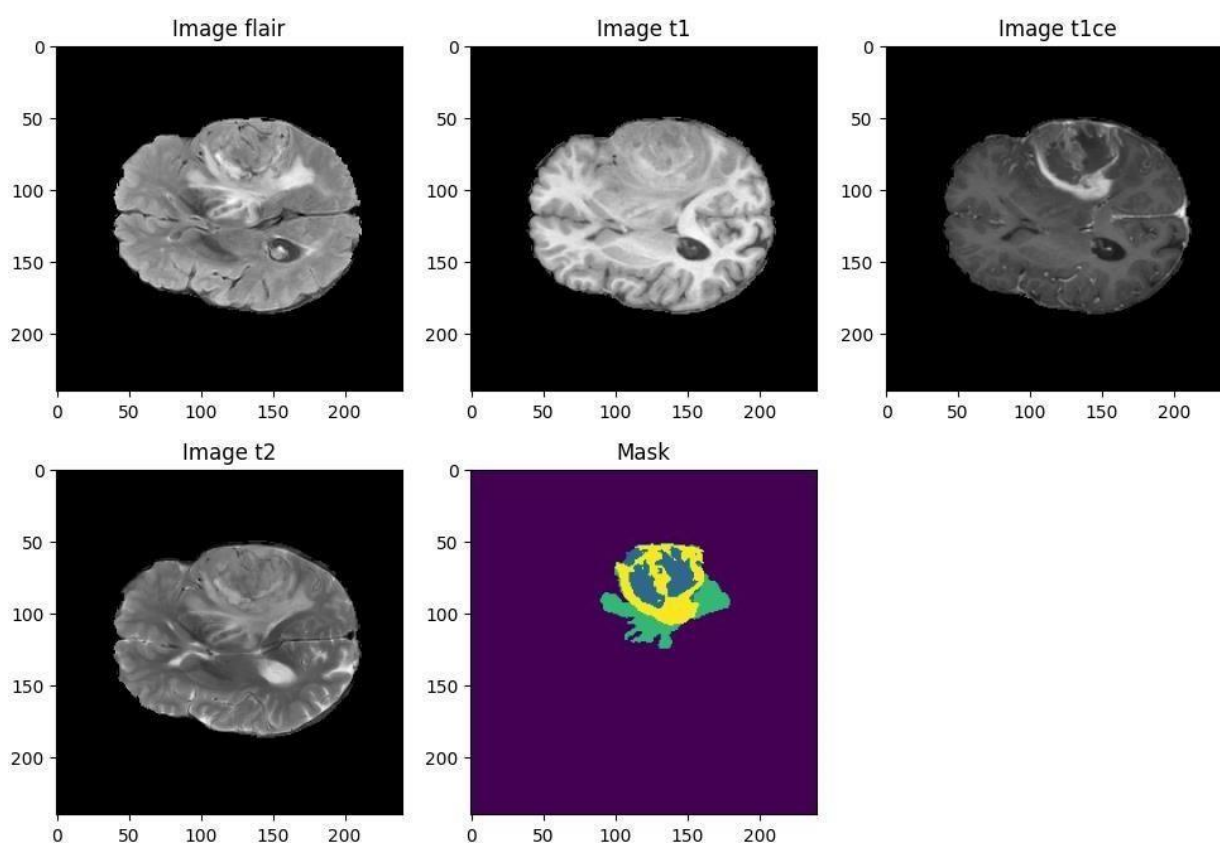


Рисунок 3.2 – Відображення 4-х модальностей МРТ-сканування пацієнта та відповідної сегментації

– T1CE: T1CE зображення схоже на T1 зображення, але відзначається підсиленою видимістю аномалій за рахунок введення контрастного агента;

– T2: T2-зважене зображення є основною імпульсною послідовністю в магнітно-резонансній томографії, що відображає відмінності у часі релаксації T2 для різних тканин. Клінічно, T2 -зважені зображення забезпечують найефективніше зображення захворювань, оскільки більшість тканин, що втягнуті в патологічний процес, мають вищий вміст води, що призводить до того, що уражені ділянки виглядають більш яскравими на T2-зважених зображеннях;

– FLAIR: FLAIR зображення приглушує вміст рідини для того, щоб краще ідентифікувати пухлини та ураження, які може бути не чітко видно на зображеннях модальностей T1 та T2.

Для експерта корисно мати всі ці чотири модальності МРТ-сканування пацієнта для більш точного аналізу та підтвердження присутності або відсутності пухлини. Проте для навчання моделі використання тільки двох модальностей може зробити процес більш швидким та ефективним, оскільки це дозволить зменшити обсяг оброблюваних даних і, відповідно, зменшити вимоги до обчислювальної потужності та пам'яті. Тому з даного набору даних були виключені модальності T1 та T2 МРТ-сканувань. Модальність T1CE є фактично покращеною версією модальності T1, а модальність T2 відображає рідини, що можуть тільки погіршити результат сегментації. Замість неї краще використати модальність FLAIR, де ці рідини видаляються, а області ураження, навпаки, краще виділяються.

Набір даних BraTS2020 містить 369 файлів у форматі NIfTI (інформаційна технологія нейровізуалізації). Формат NIfTI – формат файлів для зберігання медичних зображень, таких як магнітно-резонансна томографія (МРТ) та ком'ютерна томографія (КТ). Зображення у такому форматі є цифровим представленням тривимірного об'єкта, у випадку з набором даних BraTS2020 – головного мозку. У цьому наборі даних кожна

з модальностей МРТ-сканування та відповідні анотації представлені у тривимірній формі розмірності (240, 240, 155).

У форматі NIfTI перші три виміри використовуються для визначення тривимірних просторових координат – x , y та z , а четвертий вимір призначений для визначення часових точок – t . Решта розмірів, від п'ятого до сьомого, використовуються для інших цілей.

У наборі даних BraTS2020 зображення мають 3 виміри. Кожен вимір складається з серії двовимірних зображень, відомих як зрізи, які містять однакову кількість пікселів і складаються разом, щоб створити тривимірне представлення. Ці 3 виміри відповідають трьом просторовим вимірам зображення: ширині, висоті та глибині. У медичній візуалізації вони називаються аксіальною/поперечною, корональною та сагітальною площинами, що відповідають трьом основним орієнтаціям людського тіла, а отже, і людського мозку (рисунок 3.3).

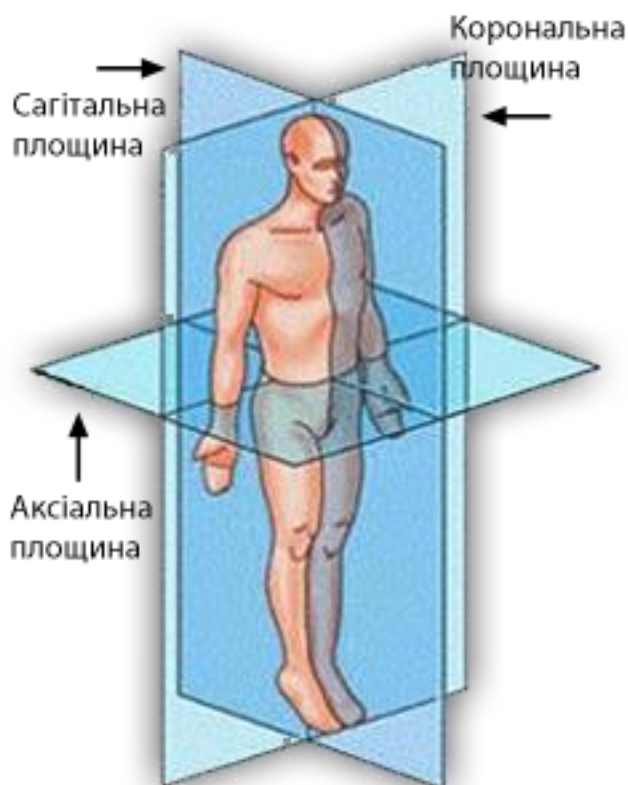


Рисунок 3.3 – Анатомічні площини тіла

На рисунку 3.3 показано кожну з площин, розглянемо детальніше кожну з них:

– корональна площина – це площина, що поділяє тіло на передню і задню частини та часто називається площиною «бік у бік»;

– сагітальна площина – це площина, що поділяє тіло на ліву та праву частини і часто називається площиною «спереду-назад»;

– аксіальна або поперечна площина – це площина, що поділяє тіло на верхню та нижню частини та часто називається площиною «голова-п'яти».

Кожну модальність МРТ-сканування пацієнта можна відобразити у наведених вище трьох площинах. Наприклад, на рисунку 3.4 відображені сагітальна, корональна та аксіальна площини зображення модальності T1.

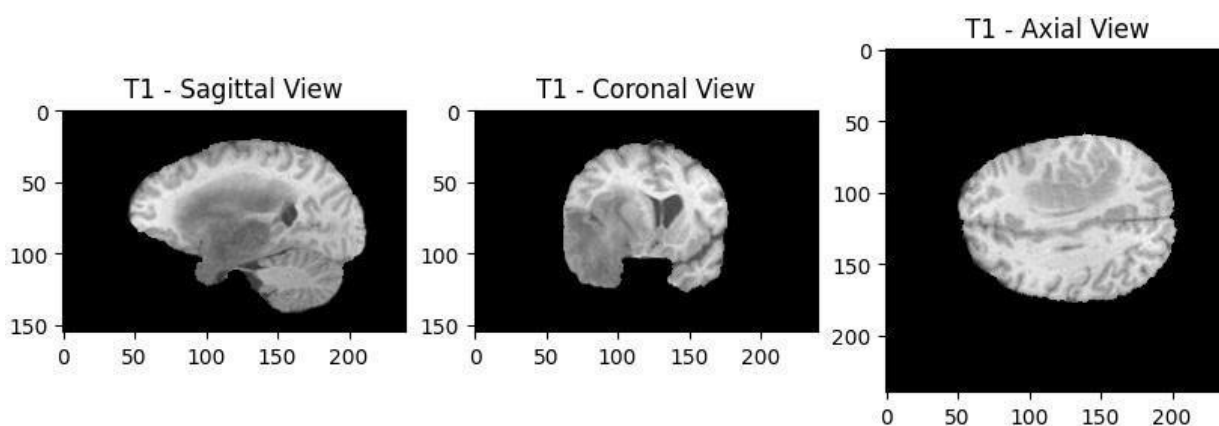


Рисунок 3.4 – Зображення МРТ-сканування головного мозку модальності T1 в трьох різних площинах

Для того, щоб краще зрозуміти з яких саме тривимірних зображень складається набір даних BraTS2020, представимо перше в даному наборі даних МРТ-сканування для модальності T1. Приклад візуалізації NIfTI файлу для цього МРТ-сканування представлено на рисунку 3.5. Загалом даний файл складається з 155 зрізів, кожний з яких має розмір 240x240 пікселів.

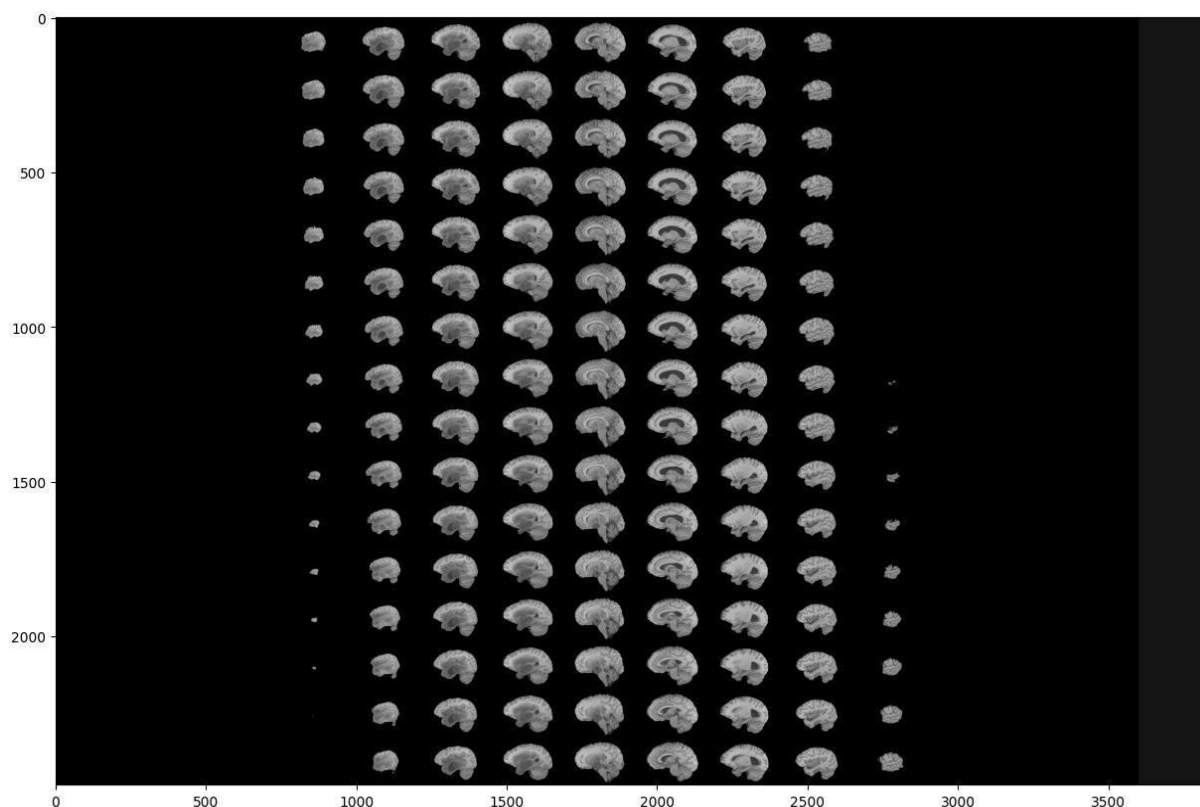


Рисунок 3.5 – Відображення всіх зрізів МРТ-сканування для модальності T1 для першого пацієнта

З рисунку 3.5 можна побачити, що велика частина зрізів не містить багато інформації, оскільки вони поступово відтворюють мозок. Крім того, на даному рисунку видно, що з обох його боків присутні зображення, які повністю чорного кольору.

Якщо розглянути відповідні сегментовані зрізи, зображені на рисунку 3.6, можна помітити, що експерти не змогли сегментувати зрізи, які містять мало інформації. Отже, для того, щоб зменшити кількість зображень та пришвидшити процес навчання моделі, ми виключимо певні зрізи із розгляду, залишивши лише ті, що належать діапазону 60:135. Це дозволить зосередитися на більш значущих та інформативних зрізах, що сприятиме покращенню якості сегментації та ефективності навчання моделі. Такий діапазон відображений на рисунку 3.7.

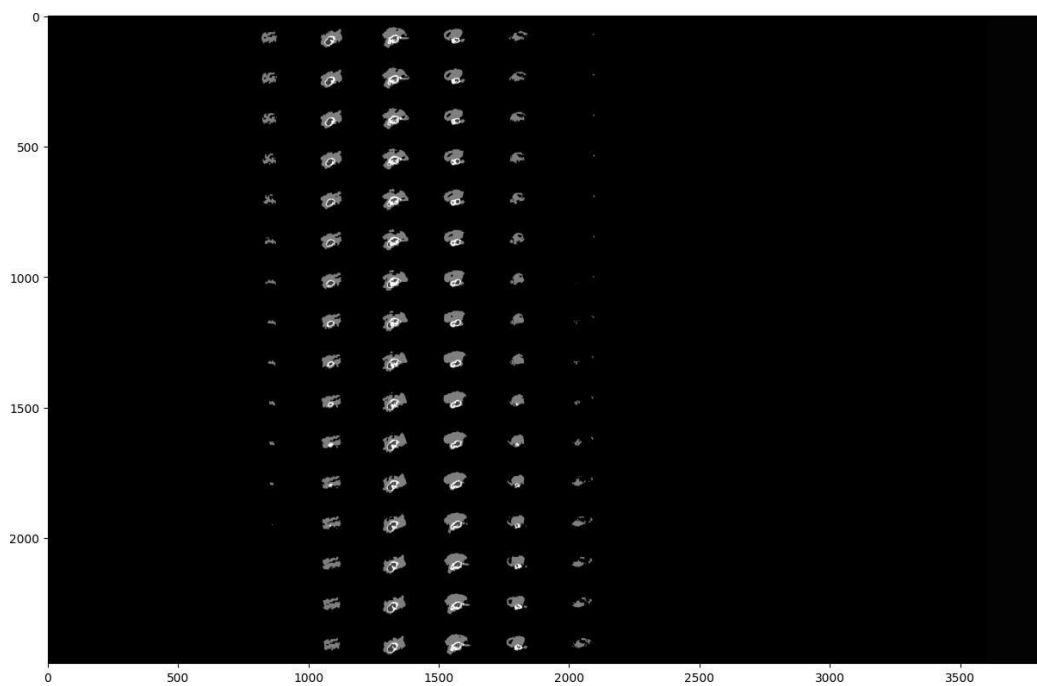


Рисунок 3.6 – Відображення сегментації всіх зрізів МРТ-сканування для модальності Т1 для першого пацієнта

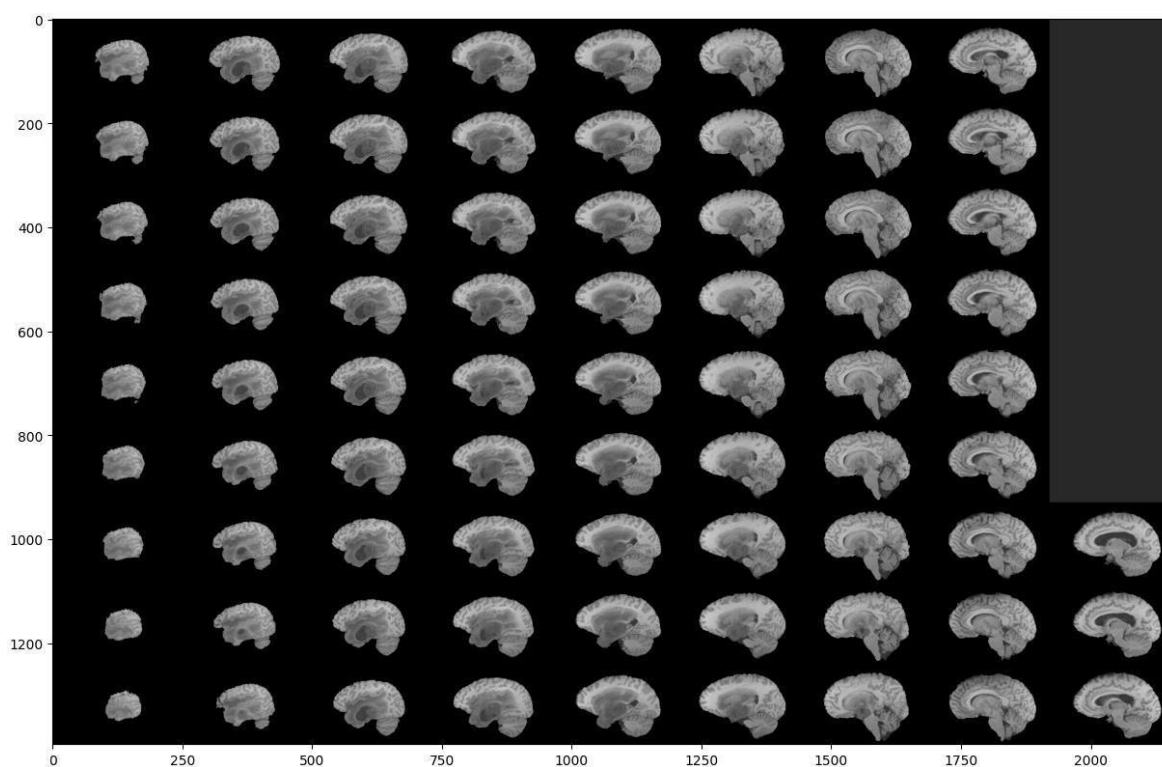


Рисунок 3.7 – Відображення діапазону 60:135 всіх зрізів МРТ-сканування для модальності Т1 для першого пацієнта

3.2 Попередня обробка даних

Загалом якість моделі визначається її здатністю робити точні прогнози на нових, небачених даних. Щоб досягти цього, важливо розділити дані на три набори: навчальний, валідаційний та тестовий.

Навчальний набір використовується для навчання моделі. Під час навчання модель отримує навчальні дані та підлаштовує свої параметри так, щоб мінімізувати похибку між її прогнозами та істинними значеннями.

Валідаційний набір – це підмножина даних, які використовуються для об'єктивної оцінки відповідності моделі навчальним даним при зміні гіперпараметрів. Метою використання даного набору даних є порівняння різних гіперпараметрів та вибір найкращої конфігурації для моделі.

Тестовий набір даних – це підмножина навчального набору даних, яка використовується для об'єктивної оцінки моделі після її навчання, щоб побачити, наскільки добре вона працює на небачених даних.

Набір даних BraTS2020 вже містить дві папки: навчальну та валідаційну, однак в папці з валідаційними файлами у форматі NIfTI не має сегментованих даних, тому дані даної папки не можуть бути використані для валідації. Тому такі зразки були проігноровані, а валідаційний набір був створений використовуючи дані з навчальної папки.

На рисунку 3.8 зображено розподіл даних, де навчальний набір даних складає 68% від всіх даних у наборі даних BraTS2020, валідаційний набір складає 20%, а тестовий – 12%.

Для того, щоб модель навчилася сегментувати об'єкти, в нашому випадку знаходити область з пухлиною на МРТ-скануванні, необхідно надати їй вхідні дані (X) та істинні значення, тобто сегментовані зображення (y). Отримуючи ці два типи даних, нейронна мережа може навчитися розпізнавати локалізацію пухлини.

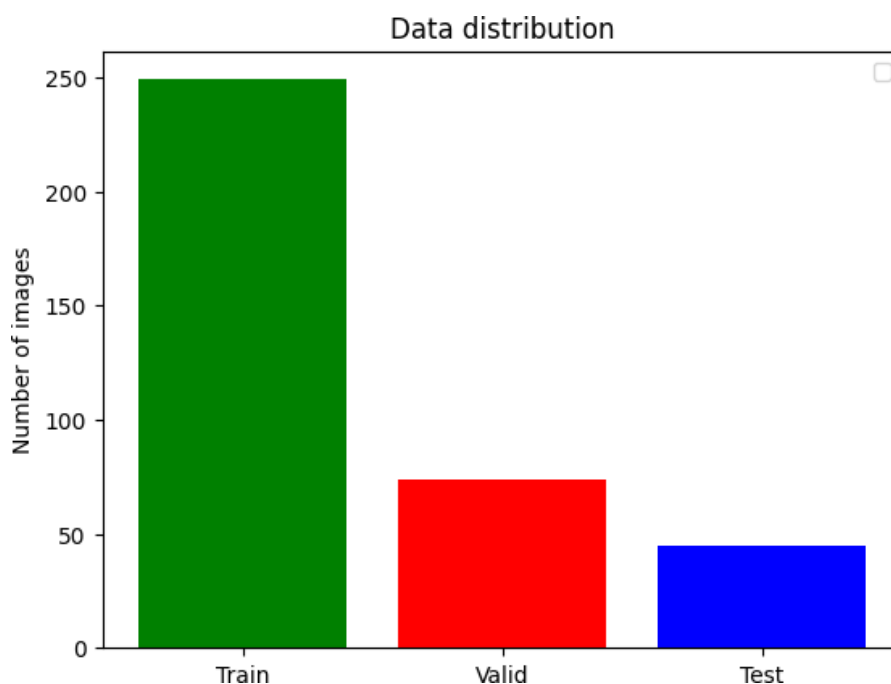


Рисунок 3.8 – Розподіл даних датасету BraTS2020

Проте, мультимодальні зображення та відповідні сегментації набору даних BraTS2020 не можуть бути безпосередньо передані до моделі, оскільки такі дані є тривимірними зображеннями, і вони не тільки перевантажать пам'ять нашого середовища, але і призведуть до помилок через невідповідність розмірності. Тому перш ніж передавати дані для навчання моделі, необхідно попередньо обробити їх.

Отже, як було зазначено вище для навчання моделі було використано тільки T1CE та FLAIR модальності та використано зрізи із діапазону 60:135. Також розмір кожний зрізу був зменшений з 240×240 пікселів до 128×128 пікселів. Зміна розміру необхідна через те, що в згортковій нейронній мережі буде використаний шар MaxPooling2D, який використовується для зменшення розмірів вхідних даних, при цьому зберігаючи найважливішу інформацію та відкидаючи менш важливі деталі.

Оскільки шар MaxPooling2D дозволяє зменшити розмірність зображення в 2 рази, то розмірність зображення повинна бути степенем двійки, тобто її можна представити у вигляді 2^n , де n – ціле число.

При цьому розмір зображень був не змінений до розмірності 256×256 пікселів, хоча вона також є степенем 2 і є ближчою до 240, ніж 128, для того, щоб модель мала менше параметрів, що дозволить зменшити її час навчання та вимоги до пам'яті.

Також, для масиву u , що містить сегментовані зображення, було застосоване однократне кодування, так як метою є визначення регіонів, що відповідають наведеним вище трьом класам. Однократне кодування дозволить перетворити категоріальні змінні (класи) у формат, який може бути використаний побудованою нейронною мережею. Категоріальні змінні повинні бути замінені на етапі попередньої обробки на числові вхідні змінні. Оскільки з математичної точки зору, надсилання масиву u в неперетвореному вигляді означатиме, що деякі класи є вищими за інші, тоді як між ними немає жодного зв'язку переваги. Наприклад, клас 1 є нижчим за клас 4, оскільки $1 < 4$. Але використовуючи однократне кодування дозволить маніпулювати числовими даними у вигляді нулів та одиниць.

3.3 Навчання моделі та аналіз метрик

Оскільки набір даних BraTS2020 складається з 3D зображень, кожне з яких складається з декількох 2D зрізів у трьох ортогональних площинах, загалом можна використати або 2D U-Net, або 3D U-Net.

Для вирішення даного завдання сегментації буде використана архітектура 2D U-Net оскільки вона добре підходить для сегментації біомедичних зображень і особливо добре підходить для задач сегментації, де області інтересу невеликі та мають складну форму (наприклад, пухлини на МРТ-знімках).

Теоретично, 3D U-Net є кращим варіантом для даного випадку, так як вона може використовувати тривимірний просторовий контекст зображень, а це означає, що вона може зменшити кількість хибно позитивних та хибно

негативних, спричинених через часткову або неповну інформацію в окремих 2D-зрізах.

Однак 2D U-Net працює швидше і вимагає менше пам'яті, ніж 3D U-Net, що може бути перевагою при роботі з великими наборами даних або обмеженими обчислювальними ресурсами.

На практиці часто корисно спробувати обидві архітектури 2D і 3D U-Net і порівняти їхню продуктивність для конкретного завдання сегментації.

Також для навчання моделі важливо правильно обрати функцію втрат, адже саме така функція дозволить порівняти прогнозовані пікселі з істинними для кожного пацієнта. Метою кожної епохи є оновлення ваг моделі таким чином, щоб мінімізувати цю функцію втрат, а отже, підвищити точність її прогнозів.

Для задач багатокласової класифікації зазвичай використовують категоріальну перехресну ентропію, яка вимірює різницю між прогнозованим розподілом ймовірностей кожного пікселя і реальним значенням однократно закодованим істинним значенням. Також часто використовують функцію втрат Дайса для моделей сегментації.

Щоб отримати розподіл ймовірностей між різними класами для кожного пікселя, ми застосовуємо функцію активації softmax до вихідного шару нейронної мережі. Це означає, що під час навчання нейронна мережа підлаштовує свої ваги так, щоб мінімізувати функцію втрат, яка порівнює прогнозовані ймовірності, задані функцією softmax, з ймовірностями сегментації на основі істинних значень.

Під час навчання також будуть використані колбеки, а саме:

– `ReduceLROnPlateau`: цей зворотний виклик зменшує швидкість навчання, коли метрика перестала покращуватися. Швидкість навчання зменшується в 0.2 рази, терпіння встановлюється на 2 епохи, а мінімальна швидкість навчання встановлюється на 0.000001;

– ModelCheckpoint: зберігає найкращі ваги моделі (модель, яка отримала найменші втрати при валідації протягом різних епох). Збереження моделі дозволяє повторно використовувати її пізніше без необхідності перенавчати її з нуля.

Також, варто визначити метрики, що будуть аналізуватися при навчанні моделі. Також важливо відстежувати ефективність моделі за допомогою оціночних метрик. Звичайно, ми будемо використовувати точність, яка є популярним показником. Однак ця метрика може вводити в оману при роботі з незбалансованими наборами даних, такими як BraTS2020. Щоб вирішити цю проблему, ми будемо використовувати інші метрики, такі як перетин над об'єднанням (IoU), коефіцієнт Дайса, точність, чутливість та специфічність. Навчання даної моделі триватиме 35 епох і його результати показані на рисунку 3.9.

```

Epoch 31/35
250/250 [=====] - ETA: 0s - loss: 0.0193 - accuracy: 0.9931 - mean_io_u_6: 0.8138 - dice_
coef: 0.6243 - precision: 0.9937 - sensitivity: 0.9914 - specificity: 0.9979
Epoch 31: val_loss did not improve from 0.02533
250/250 [=====] - 116s 464ms/step - loss: 0.0193 - accuracy: 0.9931 - mean_io_u_6: 0.8138
- dice_coef: 0.6243 - precision: 0.9937 - sensitivity: 0.9914 - specificity: 0.9979 - val_loss: 0.0254 - val_accu
acy: 0.9916 - val_mean_io_u_6: 0.8167 - val_dice_coef: 0.6098 - val_precision: 0.9922 - val_sensitivity: 0.9897 -
val_specificity: 0.9974 - lr: 1.6000e-06
Epoch 32/35
250/250 [=====] - ETA: 0s - loss: 0.0193 - accuracy: 0.9931 - mean_io_u_6: 0.8137 - dice_
coef: 0.6242 - precision: 0.9937 - sensitivity: 0.9914 - specificity: 0.9979
Epoch 32: ReduceLRonPlateau reducing learning rate to 1e-06.

Epoch 32: val_loss did not improve from 0.02533
250/250 [=====] - 101s 404ms/step - loss: 0.0193 - accuracy: 0.9931 - mean_io_u_6: 0.8137
- dice_coef: 0.6242 - precision: 0.9937 - sensitivity: 0.9914 - specificity: 0.9979 - val_loss: 0.0254 - val_accu
acy: 0.9916 - val_mean_io_u_6: 0.8166 - val_dice_coef: 0.6099 - val_precision: 0.9922 - val_sensitivity: 0.9897 -
val_specificity: 0.9974 - lr: 1.6000e-06
Epoch 33/35
250/250 [=====] - ETA: 0s - loss: 0.0193 - accuracy: 0.9931 - mean_io_u_6: 0.8137 - dice_
coef: 0.6244 - precision: 0.9937 - sensitivity: 0.9914 - specificity: 0.9979
Epoch 33: val_loss did not improve from 0.02533
250/250 [=====] - 86s 345ms/step - loss: 0.0193 - accuracy: 0.9931 - mean_io_u_6: 0.8137
- dice_coef: 0.6244 - precision: 0.9937 - sensitivity: 0.9914 - specificity: 0.9979 - val_loss: 0.0254 - val_accu
acy: 0.9916 - val_mean_io_u_6: 0.8166 - val_dice_coef: 0.6100 - val_precision: 0.9922 - val_sensitivity: 0.9897 -
val_specificity: 0.9974 - lr: 1.0000e-06
Epoch 34/35
250/250 [=====] - ETA: 0s - loss: 0.0193 - accuracy: 0.9931 - mean_io_u_6: 0.8137 - dice_
coef: 0.6244 - precision: 0.9937 - sensitivity: 0.9914 - specificity: 0.9979
Epoch 34: val_loss did not improve from 0.02533
250/250 [=====] - 86s 344ms/step - loss: 0.0193 - accuracy: 0.9931 - mean_io_u_6: 0.8137
- dice_coef: 0.6244 - precision: 0.9937 - sensitivity: 0.9914 - specificity: 0.9979 - val_loss: 0.0254 - val_accu
acy: 0.9916 - val_mean_io_u_6: 0.8166 - val_dice_coef: 0.6100 - val_precision: 0.9922 - val_sensitivity: 0.9897 -
val_specificity: 0.9974 - lr: 1.0000e-06
Epoch 35/35
250/250 [=====] - ETA: 0s - loss: 0.0193 - accuracy: 0.9931 - mean_io_u_6: 0.8137 - dice_
coef: 0.6245 - precision: 0.9937 - sensitivity: 0.9914 - specificity: 0.9979
Epoch 35: val_loss did not improve from 0.02533
250/250 [=====] - 87s 349ms/step - loss: 0.0193 - accuracy: 0.9931 - mean_io_u_6: 0.8137
- dice_coef: 0.6245 - precision: 0.9937 - sensitivity: 0.9914 - specificity: 0.9979 - val_loss: 0.0254 - val_accu
acy: 0.9916 - val_mean_io_u_6: 0.8166 - val_dice_coef: 0.6101 - val_precision: 0.9922 - val_sensitivity: 0.9897 -
val_specificity: 0.9974 - lr: 1.0000e-06

```

Рисунок 3.9 – 31–35 епохи навчання моделі

Після того як модель була навчена можна проаналізувати її метрики навчання, які показані на рисунку 3.10. На першому графіку відображено, що точність навчання, так і точність валідації зростають з епохами і досягають плато.

Це свідчить про те, що модель навчається на даних (навчальна вибірка) і добре узагальнює нові дані (валідаційна вибірка). Так як обидві метрики покращуються, не можна сказати, що модель перенавчається. З другого графіка, можна побачити, що модель навчається на навчальних даних, оскільки втрати зменшуються з часом. З наступного графіка можна побачити, що найкраща версія моделі досягається приблизно на 26 епохі, оскільки коефіцієнт Дайса досягає свого максимального значення як на навчальній, так і на валідаційній вибірці.

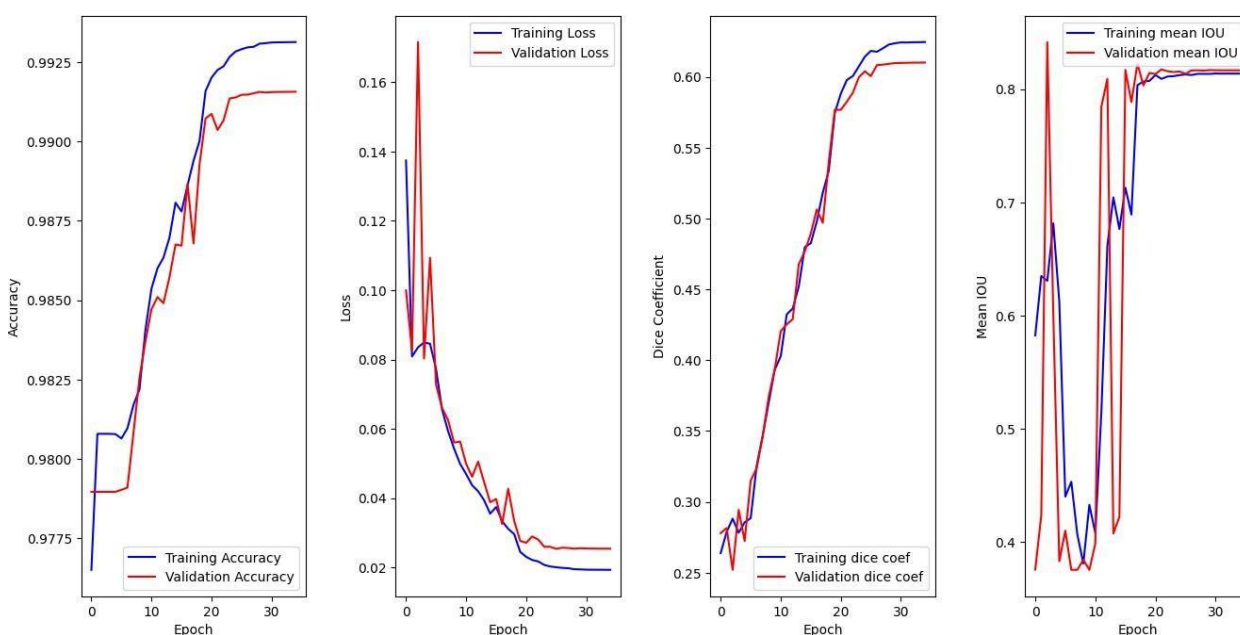


Рисунок 3.10 – Метрики навчання моделі

Після навчання моделі варто перевірити як добре вона сегментує зображення на тестовому наборі даних. Для цього використовуючи функцію `evaluate()`, отримуємо результати основних метрик на тестовій вибірці, які представленні у таблиці 3.1.

Таблиця 3.1 –Значення метрик на тестовому наборі даних

Метрика	Значення
Категоріальна крос-ентропійна втрата	0.0206
Точність	0.9935
Індекс Жаккара	0.8176
Коефіцієнт Дайса	0.6008
Прецизійність	0.9938
Чутливість	0.9922
Специфічність	0.9979

Можна зробити висновок, що модель дуже добре сегментує зображення, оскільки на тестовому наборі даних досягла низьких втрат (0.0206), а також хороших результатів за іншими метриками, які вказують на те, що модель має хороші показники узагальнення на небачених даних.

Щоб краще зрозуміти, що стоїть за цими показниками, відобразимо декілька результатів сегментування на тестовому наборі даних для рандомно вибраних пацієнтів, ці результати продемонстровані на рисунку 3.11.

Прогнозована сегментація є досить точною, але є декілька хибнопозитивних результатів, які виявляють пухлину там, де її насправді немає (наприклад, 3-ій ряд зображень рисунку 3.11).

У такій галузі як медична, такого виду помилки можуть бути неприпустимими, оскільки якщо ШІ буде повідомляти, що у пацієнта є пухлина, хоча він є здоровим, це може призвести до непотрібного лікування, що негативно вплине на його здоров'я та якість життя.

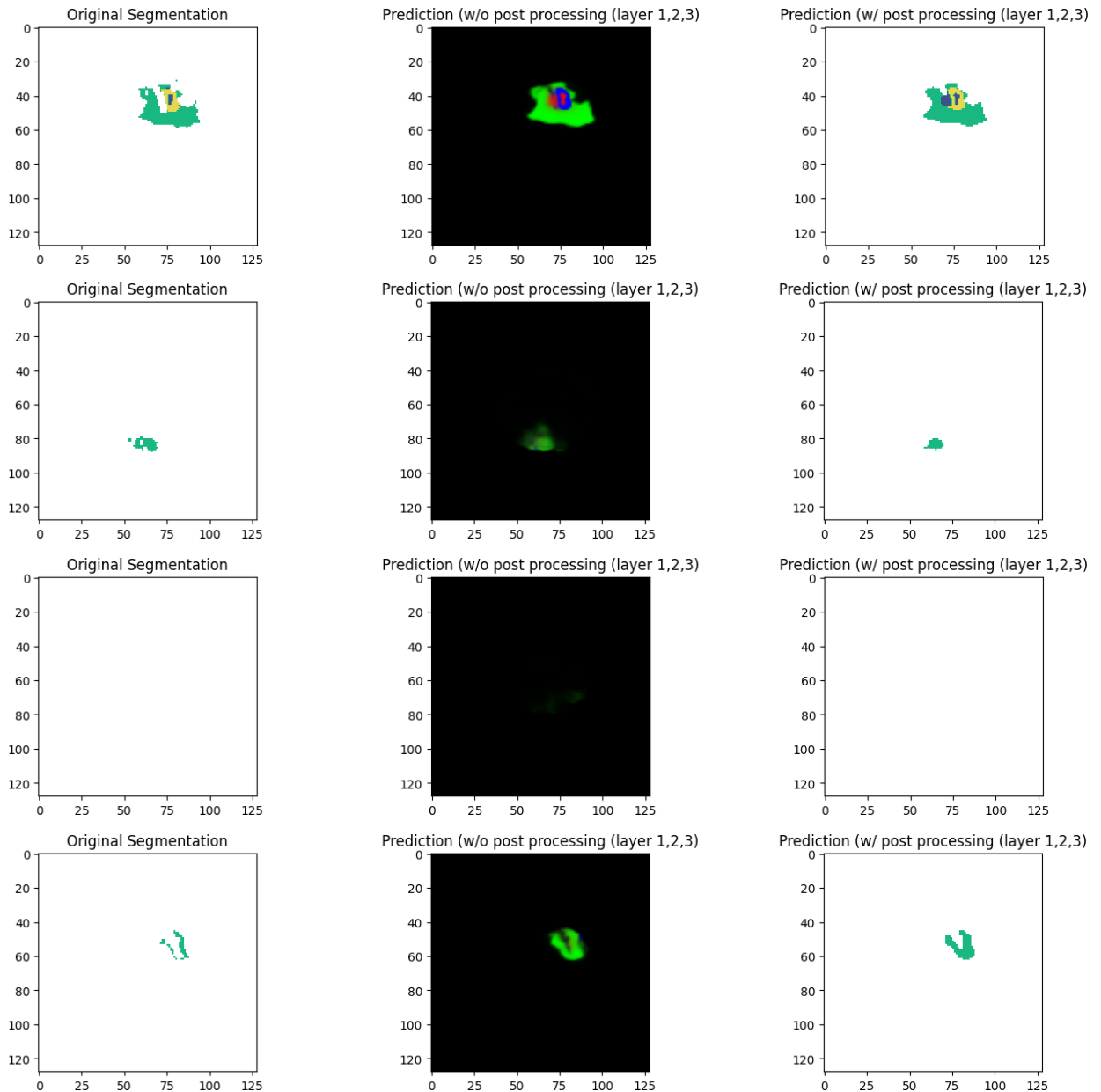


Рисунок 3.11 – Результати сегментування на тестовому наборі даних

Для того, щоб уникнути цієї проблеми можна використати декілька варіантів, а саме:

- змінити архітектуру моделі, наприклад спробувати використати 3D CNN;
- навчити моделі на більшому наборі даних;

– вибрати для навчання моделі інший діапазон зрізів, або використати інші модальності;

– застосувати методи постобробки, такі як встановлення порогових значень, `argmax`-декодування, морфологічні операції для уточнення прогнозованих сегментацій.

Як можна побачити з рисунку 3.11, такого виду помилкові прогнози є майже непомітними, тому над прогнозованими сегментаціями була виконана постобробка, для того, щоб перетворити ймовірності, отримані за допомогою функції `softmax` для кожного класу в один клас, який має найбільшу ймовірність. Для цього було застосовано функцію `Argmax()`, яка також дозволила видалити деякі хибнопозитивні випадки, та відтворити кольори оригінальної сегментації для легшого порівняння прогнозованої сегментації з оригінальною.

ВИСНОВКИ

У результаті проведеного дослідження алгоритмів сегментації пухлин за допомогою гібридних зображень можна зробити наступні висновки:

– було проведено огляд та реалізацію різних варіантів архітектури UNet, таких як UNet++, SwishUNet тощо, а також традиційна архітектура UNet. Кожен варіант було оцінено за його ефективність у завданнях сегментації пухлин;

– було виявлено, що використання гібридних зображень у сегментації пухлин дозволяє покращити точність та надійність результатів порівняно з традиційними методами. Гібридні зображення можуть включати різні модальності, такі як зображення з рентгенівською, магнітною резонансною томографією тощо, що дозволяє отримувати додаткову інформацію про пухлину та навколишні тканини;

– було розроблено та навчено 2D модель UNet на наборі даних, що містить зображення пухлин. Навіть за умови того, що набір даних містив зображення різних модальностей, для навчання використовувалися лише дві з них. Результати навчання були виявлені задовільними;

– незважаючи на успішні результати 2D моделі UNet, має сенс розглянути можливість реалізації 3D моделі UNet у майбутньому. Це особливо важливо, оскільки 3D моделі можуть ефективніше враховувати просторову інформацію, що є ключовим у медичних завданнях, де структури можуть мати складні 3D форми;

– було проведено аналіз метрик, які використовуються для оцінки якості моделей на тренувальних та тестових даних. Метрики, такі як коефіцієнт Дайса, перетин з'єднання (IoU), чутливість, специфічність та точність, надають можливість глибшого розуміння ефективності моделі у сегментації пухлин;

– крім числових результатів, був наданий детальний опис та візуалізацію результатів сегментації. Це включає в себе відображення сегментованих областей пухлин на оригінальних зображеннях, виділення областей точної сегментації, а також виявлення потенційних помилок моделі. Ці висновки свідчать про важливість та перспективність використання глибокого навчання у сегментації пухлин на медичних зображеннях та підкреслюють потребу в подальших дослідженнях у цій області для досягнення ще кращих результатів.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

- 1) Hussain S., Anwar S. M., Majid M. Segmentation of glioma tumors in brain using deep convolutional neural network, *Neurocomputing*, 2018. Vol. 282. P. 248–261. URL: <https://doi.org/10.1016/j.neucom.2017.12.032> (дата звернення: 24.04.2024).
- 2) Deep learning model integrating features and novel classifiers fusion for brain tumor segmentation / S. Iqbal et al. *Microscopy research and technique*. 2019. Vol. 82, no. 8. P. 1302–1315. URL: <https://doi.org/10.1002/jemt.23281> (date of access: 17.04.2024).
- 3) Zachow, Stefan, Michael Zilske, and Hans-Christian Hege. 3D reconstruction of individual anatomy from medical image data: Segmentation and geometry processing. *Proceedings of the CADFEM Users Meeting*. 2007.
- 4) Pal N. R., Pal S. K. A review on image segmentation techniques. *Pattern recognition*. 1993. Vol. 26, no. 9. P. 1277–1294. URL: [https://doi.org/10.1016/0031-3203\(93\)90135-j](https://doi.org/10.1016/0031-3203(93)90135-j) (date of access: 20.05.2024).
- 5) Kaur D., Kaur Y. Various image segmentation techniques: a review. *International Journal of Computer Science and Mobile Computing*. 2014. Vol. 3, Issue 5. P. 809–814.
- 6) Pham D. L., Xu C., Prince J. L. Current methods in medical image segmentation. *Annual review of biomedical engineering*. 2000. Vol. 2, no. 1. P. 315–337. URL: <https://doi.org/10.1146/annurev.bioeng.2.1.315> (date of access: 15.04.2024).
- 7) Heaton J. Ian Goodfellow, Yoshua Bengio and Aaron Courville: deep learning. *Genetic programming and evolvable machines*. 2017. Vol. 19, no. 1-2. P. 305–307. URL: <https://doi.org/10.1007/s10710-017-9314-z> (date of access: 12.04.2024).

8) Yuheng S., Hao Y. Image segmentation algorithms overview. URL: <https://arxiv.org/ftp/arxiv/papers/1707/1707.02051.pdf> (дата звернення: 16.04.2024).

9) Brain tumour image segmentation using deep networks / M. Ali et al. *IEEE access*. 2020. Vol. 8. P. 153589–153598. URL: <https://doi.org/10.1109/access.2020.3018160> (date of access: 19.04.2024).

10) Zhang J. UNet – line by line explanation. *Towards Data Science* 18, 2019. URL: <https://towardsdatascience.com/unet-line-by-line-explanation-9b191c76baf5> (дата звернення: 27.04.2024).

11) Tsang Sik-Ho. Review: 3D U-Net - Volumetric Segmentation (Medical Image Segmentation). *Medium*, Towards Data Science, 2019. URL: <https://towardsdatascience.com/review-3d-u-net-volumetric-segmentation-medical-image-segmentation-8b592560fac1> (дата звернення: 01.05.2024).

12) 3D U-Net: learning dense volumetric segmentation from sparse annotation / Çiçek Ö. et al. *Medical Image Computing and Computer-Assisted Intervention–MICCAI*. 2016. URL: <https://arxiv.org/pdf/1606.06650.pdf> (дата звернення: 24.04.2024).

13) From CNN to transformer: A review of medical image segmentation models / Yao W. et al. *Journal of Imaging Informatics in Medicine*. 2023. URL: <https://arxiv.org/pdf/2308.05305.pdf> (дата звернення: 13.04.2024).

14) Ronneberger O., Fischer P., Brox T. U-net: convolutional networks for biomedical image segmentation. *Medical image computing and computer-assisted intervention–MICCAI*, 2015. URL: <https://arxiv.org/abs/1505.04597.pdf> (дата звернення: 10.05.2024).

15) Unet++: A nested U-Net architecture for medical image segmentation / Zongwei Z. et al. *Springer International Publishing*, 2018. URL: <https://arxiv.org/pdf/1807.10165.pdf> (дата звернення: 19.04.2024).

16) TransUNet: Transformers make strong encoders for medical image segmentation / Chen J. et al. 2021. URL: <https://arxiv.org/pdf/2102.04306v1.pdf> (дата звернення: 23.04.2024).

17) Swin-Unet: Unet-like pure transformer for medical image segmentation / Cao H. et al. *Springer Nature Switzerland*, 2021. URL: <https://arxiv.org/pdf/2105.05537v1.pdf> (дата звернення: 27.04.2024).