https://doi.org/10.31891/2307-5732-2025-349-54
УДК 004.932.2:616.31

РИБАК ВОЛОДИМИР
Національний лісотехнічний університет України
https://orcid.org/0009-0006-8809-978X
e-mail: volodymyr.rybak.personal@gmail.com
ШАБАТУРА ЮРІЙ
Національна академія сухопутних військ ім. гетьмана Петра Сагайдачного
https://orcid.org/0000-0002-9961-1244
e-mail: shabaturayuriy@gmail.com

ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ СЕГМЕНТАЦІЇ СТОМАТОЛОГІЧНИХ ЗОБРАЖЕНЬ НА ОСНОВІ ІНТЕГРАЦІЇ ЛЕГКОВАГОВОЇ МОДИФІКАЦІЇ U-NET ІЗ MOBILENET ТА VISION MAMBA LAYERS

Сегментація стоматологічних зображень, зокрема рентгенівських зображень, є важливим етапом діагностичного процесу в стоматології. Висока точність автоматизованих методів сегментації дозволяє підвищити ефективність виявлення патологій, зменшити залежність від суб'єктивного фактору лікаря та пришвидшити прийняття клінічних рішень. Архітектура U-Net залишається стандартом у сегментації медичних зображень завдяки здатності до багаторівневого вилучення ознак та ефективному використанню пропускних з'єднань. Проте її значна кількість параметрів і високі обчислювальні витрати створюють труднощі для застосування на пристроях із обмеженими ресурсами, таких як мобільні пристрої чи портативні медичні апарати.

У цій роботі представлено полегшену модифікацію U-Net, яка спрямована на підвищення обчислювальної ефективності без втрати точності сегментації. Основою моделі є інтеграція MobileNet як енкодера, що дозволяє суттєво зменшити кількість параметрів завдяки використанню Depthwise Separable Convolutions. Додатково, у структурі моделі застосовуються Vision Mamba Layers, які підвищують здатність моделі до моделювання глобальних та локальних залежностей. Така комбінація дозволяє зберігати високу точність при значно знижених обчислювальних витратах.

Експериментальні дослідження проводилися на стоматологічних рентгенівських зображеннях із використанням загальноприйнятих метрик, таких як Dice та IoU. Отримані результати свідчать про переваги запропонованої моделі над базовими підходами, зокрема щодо точності сегментації, часу навчання та обчислювальної ефективності. Пропонована архітектура є перспективною для впровадження в клінічну практику, де необхідна швидка та точна сегментація стоматологічних зображень, а також для використання на пристроях із обмеженими обчислювальними ресурсами.

Ключові слова: U-Net, MobileNet, сегментація, стоматологічні рентгенівські зображення, Vision Mamba Layers, глибоке навчання.

RYBAK VOLODYMYR
The National Forestry University of Ukraine
SHABATURA YURIY
Hetman Petro Sahaidachnyi National Ground Forces Academy

INFORMATION TECHNOLOGY FOR DENTAL IMAGE SEGMENTATION BASED ON THE INTEGRATION OF LIGHTWEIGHT U-NET MODIFICATION WITH MOBILENET AND VISION MAMBA LAYERS

Segmentation of dental images, particularly radiographs, is a crucial step in the diagnostic workflow in dentistry. High-accuracy automated segmentation methods enable more effective detection of pathologies, reduce reliance on subjective expert assessments, and accelerate clinical decision-making. The U-Net architecture remains a standard in medical image segmentation due to its capability for multiscale feature extraction and efficient utilization of skip connections. However, its large number of parameters and high computational demands pose challenges for deployment on resource-constrained devices such as mobile platforms or portable medical equipment.

This study presents a lightweight modification of U-Net aimed at improving computational efficiency without compromising segmentation accuracy. The core of the proposed model is the integration of MobileNet as the encoder, which significantly reduces the number of parameters by leveraging Depthwise Separable Convolutions. Additionally, Vision Mamba Layers are employed within the architecture to enhance the model's capacity for capturing both global and local dependencies. This combination ensures high segmentation precision while substantially reducing computational overhead.

Experimental evaluations were conducted on dental radiographs using widely recognized metrics such as Dice and IoU. The results demonstrate the advantages of the proposed model over baseline approaches, particularly in terms of segmentation accuracy, training time, and computational efficiency. The proposed architecture shows strong potential for clinical applications requiring fast and accurate dental image segmentation, as well as for deployment on devices with limited computational resources.

Keywords: U-Net, MobileNet, segmentation, dental radiographs, Vision Mamba Layers, deep learning.

Постановка проблеми у загальному вигляді та її зв'язок із важливими науковими чи практичними завданнями

Сегментація медичних зображень, зокрема стоматологічних рентгенівських зображень, є одним із ключових завдань у сучасній медицині, яке визначає якість діагностики та лікування. У стоматології автоматизована сегментація структур, таких як зуби, корені чи патологічні утворення, дозволяє значно прискорити процес аналізу зображень, знизити суб'єктивність експертних оцінок і зменшити ризик

діагностичних помилок. Це особливо актуально в умовах зростання навантаження на лікарів-стоматологів та необхідності оперативного прийняття клінічних рішень.

Традиційні методи сегментації на основі ручної обробки ознак мають обмежену точність і страждають від низької генералізації на нових даних. Глибокі нейронні мережі, зокрема архітектура U-Net, відкрили нові горизонти в сегментації завдяки здатності ефективно вилучати локальні та глобальні ознаки. Проте, застосування U-Net у реальних умовах обмежується її високими обчислювальними витратами та значною кількістю параметрів, що ускладнює впровадження на мобільних пристроях чи портативному медичному обладнанні.

Ця проблема набуває особливої важливості в контексті сучасних трендів медицини, де використання мобільних платформ і хмарних рішень є критичним. У зв'язку з цим розробка легковагових архітектур на основі U-Net із використанням компонентів, таких як MobileNet та Vision Mamba Layers, стає не лише науково актуальним, але й практично значущим завданням. Інтеграція цих компонентів дозволяє знизити обчислювальні витрати та забезпечити високу точність сегментації, зберігаючи адаптивність до ресурсних обмежень.

Аналіз досліджень та публікацій

Сегментація медичних зображень є однією з найбільш досліджуваних тем у сфері комп'ютерного зору, особливо в контексті використання глибоких нейронних мереж. Однією з найбільш впливових архітектур для сегментації є U-Net, запропонована для сегментації біомедичних зображень. Ця модель завдяки своїй симетричній структурі та пропускним з'єднанням забезпечує високу точність навіть за обмеженої кількості даних для навчання. Проте, незважаючи на високу ефективність, її застосування у реальних умовах, таких як мобільні пристрої або портативне обладнання, залишається обмеженим через високу обчислювальну складність.

Сучасні дослідження спрямовані на вирішення цього обмеження шляхом розробки полегшених архітектур. Впровадження таких технологій, як MobileNet, Vision Mamba Layers та інших компонентів, дозволяє оптимізувати використання ресурсів без втрати точності сегментації. Особливу увагу дослідники приділяють розробці моделей для вузькоспеціалізованих завдань, зокрема сегментації стоматологічних зображень. Ці підходи мають значний потенціал у медицині, автоматизованій діагностиці та плануванні лікування.

Серед останніх підходів до вдосконалення архітектури U-Net значну увагу приділено легковаговим моделям, які спрямовані на зменшення обчислювальних витрат та підвищення точності. Наприклад, у роботі [5] представлено інтеграцію Vision Mamba Layers, які дозволяють ефективно моделювати глобальні та локальні залежності в сегментаційних задачах. У дослідженні використовувалися рентгенівські зображення легень (CXR), що мали неоднорідності у структурі, характерні для діагностики захворювань, таких як пневмонія та туберкульоз. Результати показали, що застосування Vision Mamba Layers значно покращує якість сегментації, особливо у випадках, де межі патологічних областей є розмитими.

Використання MobileNet як енкодера, описане у [8], спрямоване на оптимізацію обчислювальних витрат за рахунок впровадження Depthwise Separable Convolutions. У цьому дослідженні модель випробовувалася на дерматологічних зображеннях шкіри з метою сегментації уражених ділянок, таких як меланома та базально-клітинний рак. Автори показали, що така архітектура здатна сегментувати невеликі патологічні області з високою точністю, навіть на зображеннях із неоднорідним освітленням.

У [6] продемонстровано, що поєднання Vision Mamba Layers із U-Net дозволяє досягти суттєвих покращень у задачах сегментації медичних зображень. Зокрема, модель застосовувалася для сегментації анатомічних структур на КТ-знімках органів черевної порожнини, таких як печінка, селезінка та нирки. Результати дослідження свідчать, що Vision Mamba Layers ефективно справляються з неоднорідними текстурами та шумами, які часто присутні у медичних томографічних зображеннях.

Дослідження [7] більш детально зосереджується на моделюванні контекстуальних залежностей у складних медичних зображеннях. У цьому випадку модель використовувалася для сегментації структур мозку на МРТ-зображеннях. Автори зазначили, що використання Vision Mamba Layers дозволяє зберігати дрібні деталі, наприклад, структуру гіпокампу, що є критичним для діагностики неврологічних захворювань, таких як хвороба Альцгеймера. Показники точності сегментації перевищили 90% у метриках Dice та IoU.

Крім того, робота [9] пропонує використання MobileNetV2 як основи для енкодера в задачах сегментації. У дослідженні використовувалися рентгенівські зображення грудної клітки (CXR) для виявлення патологій у легенях. Автори відзначили, що архітектура MobileNetV2 не лише зменшує кількість параметрів, але й забезпечує високу швидкість обробки, що робить її придатною для мобільних додатків. Наприклад, час обробки одного зображення скоротився до 15 мс на сучасному мобільному процесорі.

Загалом, аналіз джерел свідчить, що проблема створення легковагових моделей для медичної сегментації залишається актуальною, особливо у контексті використання моделей у ресурсно обмежених умовах, таких як мобільні пристрої. Запропонована архітектура MobileMamba-U-Net покликана заповнити наявний розрив між точністю сегментації та мінімізацією обчислювальних витрат,

використовуючи ефективні компоненти, такі як MobileNet та Vision Mamba Layers. Поєднання цих підходів дозволяє досягти високої точності при значному зменшенні кількості параметрів, що робить модель придатною для практичного використання у медицині.

Формулювання цілей статті

Метою роботи є: розробка та аналіз полегшеної архітектури U-Net, яка інтегрує MobileNet як енкодер і Vision Mamba Layers для моделювання глобальних і локальних залежностей у сегментації стоматологічних рентгенівських зображень. Запропонований підхід спрямований на досягнення високої точності сегментації при значному скороченні кількості параметрів моделі та зниженні обчислювальних витрат. Це забезпечить можливість ефективного застосування моделі в реальних клінічних умовах, включаючи використання на мобільних пристроях.

Виклад основного матеріалу

U-Net — це одна з найпоширеніших архітектур для сегментації медичних зображень, яка забезпечує високу точність завдяки здатності ефективно вилучати та відновлювати ознаки на різних рівнях абстракції. Архітектура складається з двох частин: енкодера, що зменшує просторову роздільну здатність для вилучення ознак, і декодера, який відновлює просторову структуру для точного сегментування. Ключовою особливістю є пропускні з'єднання між енкодером і декодером, які передають деталі зображення, забезпечуючи локалізацію дрібних структур.

Енкодер виконує послідовне зменшення розмірів ознак через згортки та максимальне підбирання (MaxPooling), виділяючи ключові елементи. Декодер натомість застосовує транспоновані згортки або інтерполяцію для відновлення просторової роздільної здатності. Така структура дозволяє U-Net працювати ефективно навіть за обмеженої кількості даних, що є важливим для медичних задач.

Проте класична U-Net має значну кількість параметрів, що ускладнює її використання на пристроях із обмеженими обчислювальними ресурсами, таких як мобільні платформи. Це обмеження мотивує розробку полегшених модифікацій архітектури, спрямованих на збереження точності при зниженні обчислювальних витрат.

MobileNet — це архітектура глибокої згорткової нейронної мережі, розроблена для обчислювально обмежених середовищ, таких як мобільні пристрої. Її ключова інновація — Depthwise Separable Convolutions, які замінюють традиційні згортки двома кроками: обробкою кожного каналу окремо (Depthwise Convolution) і об'єднанням каналів через 1×1-згортки (Pointwise Convolution). Цей підхід дозволяє значно знизити кількість параметрів і обчислювальні витрати, що робить MobileNet ефективним інструментом для вилучення ознак із зображень у задачах класифікації, детекції та сегментації.

MobileNet використовує Depthwise Separable Convolutions, які складаються з двох етапів:

1. Depthwise Convolution — виконує згортку для кожного каналу вхідного тензора незалежно:

$$Y_k(x,y) = \sum_{i,j} X_k(x+i, y+j) \cdot W_k(i,j), \quad (1)$$

де X_k — вхідний канал, W_k — ядро згортки, Y_k — вихідний канал. Тут обчислення виконуються окремо для кожного каналу k, що знижує обчислювальні витрати.

2. Pointwise Convolution — виконує 1×1- згортку для поєднання каналів:

$$Z(x,y) = \sum_{i,j} Y_k(x,y) \cdot W_k^{(1\times1)}, \quad (2)$$

де $W_k^{(1\times1)}$ — ядро для об'єднання каналів.

Обчислювальна складність стандартної згортки:

$$D_k^2 \cdot M \cdot N \cdot H \cdot W, \quad (3)$$

де D_k^2 — розмір ядра згортки, M іN— кількість каналів вхідних і вихідних ознак, $H \times W$— просторовий розмір ознак.

Vision Mamba Layers — це інноваційний підхід до моделювання просторово-часових залежностей у зображеннях. Вони базуються на лінійних моделях стану (SSMs), що дозволяє поєднувати глобальні та локальні ознаки з високою ефективністю. Ці шари використовуються для декодування і реконструкції просторової структури зображень, зберігаючи точність навіть у складних випадках. Vision Mamba Layers поєднують гнучкість та обчислювальну ефективність, що робить їх ідеальними для медичних задач, зокрема сегментації.

Vision Mamba Layers базуються на лінійних моделях стану (SSMs), які описуються як:

$$h'(t) = Ah(t) + Bx(t), \quad (4)$$

$$y(t) = Ch(t) + Dx(t), \quad (5)$$

де $h(t)$ - стан системи, $x(t)$ - вхідний сигнал, $y(t)$ - вихідний сигнал, A, B, C, D - параметри системи.

Vision Mamba Layers використовують ці моделі для обробки послідовностей ознак, забезпечуючи ефективне об'єднання глобальних і локальних залежностей.

Запропонована модель є полегшеною модифікацією архітектури U-Net, спрямованою на підвищення обчислювальної ефективності при збереженні високої точності сегментації. Вона інтегрує MobileNet як енкодер для зменшення кількості параметрів і Vision Mamba Layers у декодер для вдосконалення моделювання локальних і глобальних залежностей.

Основні компоненти моделі:

1. **MobileNet як енкодер**: забезпечує вилучення ключових ознак зображення із суттєвим зменшенням кількості параметрів.
2. **Vision Mamba Layers у декодері**: удосконалюють відновлення просторової структури, зберігаючи контекстну інформацію.
3. **Пропускні з'єднання**: передають деталі з енкодера до декодера, що дозволяє зберігати локальні ознаки зображення.
4. **Інтерполяція та транспоновані згортки**: використовуються у декодері для відновлення просторової роздільної здатності.

Схема моделі ілюструє основні компоненти: MobileNet як енкодер, Vision Mamba Layers у декодері, та взаємодію між ними через пропускні з'єднання.

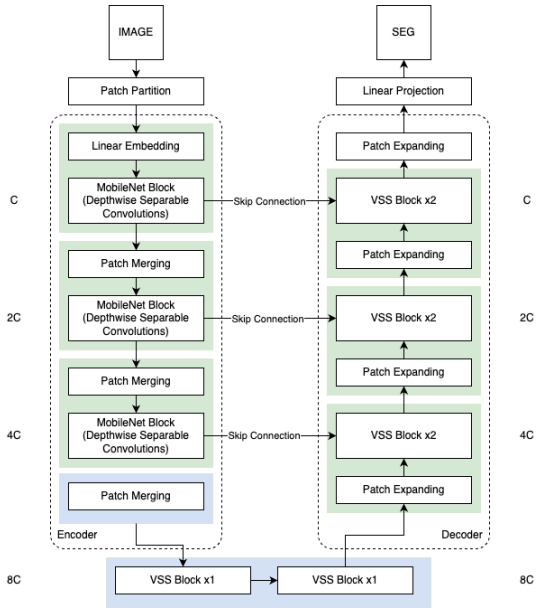


Рис. 1. Архітектура запропонованої моделі MobileMamba-U-Net.

Запропонована модель MobileMamba-U-Net є адаптацією класичної архітектури U-Net із інтеграцією компонентів, спрямованих на зменшення обчислювальних витрат і підвищення точності сегментації. Схема на рис. 1 демонструє основні складові моделі.

Основні компоненти архітектури:

1. **Patch Partition**

Цей компонент відповідає за початкову обробку вхідного зображення. Вхідне зображення розбивається на невеликі патчі фіксованого розміру, які виступають як основні одиниці для подальшої обробки. Поділ на патчі дозволяє зберігати просторову інформацію та структуру зображення, полегшуючи їх обробку на наступних етапах.

2. **Linear Embedding**

Після поділу на патчі ці блоки передаються через шар лінійного вбудовування (Linear Embedding), який трансформує їх у вектори фіксованого розміру. Цей шар зменшує обсяг вхідної інформації, що допомагає моделі працювати з великими обсягами даних без значного збільшення обчислювальних витрат.

3. **MobileNet Blocks (Depthwise Separable Convolutions)**

Кожен рівень енкодера включає MobileNet Block, який реалізує Depthwise Separable Convolutions. Ці згортки складаються з двох частин:

4. **Depthwise Convolution:** незалежна обробка кожного каналу вхідних ознак, що зменшує обчислювальні витрати.
5. **Pointwise Convolution:** 1×1 -згортка, яка об’єднує інформацію між каналами.

Ці блоки забезпечують ефективне вилучення ознак із мінімальними ресурсами, зберігаючи при цьому здатність до сегментації.

6. **Patch Merging**

На етапі Patch Merging виконується зменшення просторових розмірів ознак, що дозволяє збільшити кількість каналів. Це зменшує загальний обсяг оброблюваних даних, одночасно зберігаючи ключову інформацію для подальшої обробки.

7. **Skip Connections**

Пропускні з’єднання передають інформацію між відповідними рівнями енкодера та декодера. Це забезпечує збереження локальних ознак із вихідного зображення, що є важливим для точного відновлення меж і деталей на етапі декодування.

8. **VSS Blocks**

Vision Mamba Layers інтегруються в декодер і bottleneck моделі. Вони базуються на лінійних моделях стану (SSMs) і забезпечують моделювання як локальних, так і глобальних залежностей у просторових ознаках. Завдяки цьому модель досягає високої точності у відновленні просторової структури.

9. Patch Expanding

У декодері кожен рівень включає Patch Expanding, який поступово збільшує просторові розміри ознак до оригінальної роздільної здатності зображення. Цей етап також включає транспоновані згортки, які покращують точність відновлення.

10. **Linear Projection**

Це фінальний етап декодера, де ознаки після відновлення просторової структури перетворюються у сегментоване зображення. Linear Projection дозволяє відновити вихідну роздільну здатність зображення разом із сегментаційною маскою.

Опис експериментальних умов

Для оцінювання ефективності запропонованої моделі було використано набір даних, що містить рентгенівські стоматологічні знімки різних пацієнтів. Ці дані було отримано з відкритих медичних баз даних і містять як знімки зубів, так і відповідні маски сегментації, створені експертами-стоматологами. Для забезпечення узагальнення результатів і підвищення стійкості моделі дані перед використанням були піддані попередній обробці, включаючи нормалізацію зображень і методи аугментації.

Характеристики набору даних

Вибірка містила 700 зображень у форматі 512×512 пікселів, поділених на три підвибірки:

- навчальна (70%);
- валідаційна (15%);
- тестова (15%).

Під час аугментації було застосовано такі операції:

- обертання на випадковий кут у межах $[-15^{\circ},15^{\circ}]$;
- дзеркальне відображення;
- зміна яскравості та контрасту;
- масштабування у межах [0.9,1.1][0.9, 1.1][0.9,1.1].

Параметри навчання

Для навчання моделі використано оптимізатор *Adam* із початковою швидкістю навчання 1×10^{-4}. Функція втрат — *Cross-Entropy Loss*, яка добре підходить для задач семантичної сегментації. Основні гіперпараметри:

- розмір пакета: 32 зображення;
- кількість епох: 50;
- Dropout: 0.5 для регуляризації;
- L2-регуляризація для зменшення перенавчання.

Опис основних метрик.

Для оцінки ефективності моделі було використано такі метрики:

1. **Dice coefficient** є метрикою для оцінки схожості між передбаченою та реальною масками сегментації. Він враховує правильно передбачені позитивні пікселі, хибно передбачені позитивні пікселі та пропущені позитивні пікселі, і розраховується за формулою:

$$Dice = \frac{2 \cdot |X \cap Y|}{|X| + |Y|}, \tag{6}$$

де X — передбачена маска, Y — істинна маска..

2. **Intersection over union (IoU)**, також відомий як коефіцієнт Жаккара, оцінює точність сегментації шляхом порівняння перетину та об'єднання передбачуваної та реальної масок. Ця метрика є більш суворою, ніж Dice coefficient, оскільки враховує як хибно передбачені, так і пропущені пікселі. Формула для обчислення IoU:

$$IoU = \frac{|X \cap Y|}{|X \cup Y|}, \tag{7}$$

3. **Чутливість (recall)** вимірює здатність моделі виявляти всі позитивні пікселі. Вона обчислюється як відношення правильно передбачених позитивних пікселів до загальної кількості позитивних пікселів:

$$Recall = \frac{TP}{TP + FN}, \tag{8}$$

де TP – кількість правильно передбачених пікселів; FN – кількість пропущених пікселів, які повинні були бути позитивними.

4. **Точність передбачення (precision)** оцінює, наскільки модель правильно передбачає позитивні пікселі серед усіх передбачених позитивних. Формула для обчислення:

$$Precision = \frac{TP}{TP + FP}, \tag{9}$$

де FP – кількість хибно передбачених позитивних пікселів.

Для оцінювання продуктивності запропонованої моделі MobileMamba-U-Net було використано набір стандартних метрик, зокрема Dice coefficient, Intersection over Union (IoU), чутливість (recall) та

точність передбачення (precision). Зазначені метрики дозволяють отримати комплексну оцінку здатності моделі до сегментації стоматологічних зображень, враховуючи як точність передбачень, так і їхню повноту.

Ці метрики обчислювалися на тренувальній, валідаційній і тестовій вибірках, що забезпечило всебічний аналіз продуктивності моделі. На тренувальній вибірці оцінювалася здатність моделі до навчання, на валідаційній — її узагальнюючі властивості, а на тестовій — ефективність у нових умовах, які модель не бачила під час навчання.

Використання Dice coefficient і IoU забезпечило кількісну оцінку схожості між передбачуваними і реальними масками сегментації, дозволяючи виявляти загальну ефективність моделі. Чутливість (recall) показала здатність моделі виявляти всі позитивні пікселі, а точність (precision) — наскільки передбачені позитивні пікселі справді є коректними. Такий підхід дозволив не лише оцінити загальну ефективність моделі, але й ідентифікувати її слабкі місця для подальшого вдосконалення.

Аналіз отриманих результатів

Результати експериментального дослідження представлені у таблиці 1 демонструють переваги запропонованої моделі MobileMamba-U-Net у задачах сегментації стоматологічних зображень. Основні показники, такі як Dice coefficient, Intersection over Union (IoU), точність (precision) і чутливість (recall), були розраховані на тестовій вибірці, що дозволило оцінити ефективність моделі у відтворенні реальних меж і структур.

Таблиця 1

Порівняння метрик на тренувальній і валідаційній вибірках

Вибірка	**Dice**	**IoU**	**Precision**	**Recall**
Тренувальна	0,93	0,89	0,87	0,92
Тестова	0,9	0,86	0,84	0,89

На підставі наведених даних у таблиці 1 можна зробити висновок, що модель показує стабільні результати на тестових даних.

Оцінка продуктивності проводилася для порівняння запропонованої моделі з базовими підходами, включаючи класичну U-Net і її модифікації. Це дозволило отримати чітке уявлення про те, наскільки інтеграція MobileNet і Vision Mamba Layers покращує точність сегментації та зменшує обчислювальні витрати, що є критичним для застосувань у мобільних і ресурсно обмежених середовищах.

Таблиця 2

Порівняння ефективності моделей U-Net та U-Net 3+

Модель	**Dice**	**IoU**	**Precision**	**Recall**	**Кількість параметрів**
U-Net	0,855	0,849	0,84	0,881	32 млн
U-Net 3+	0,902	0,883	0,91	0,907	26 млн
MobileMamba-U-Net	0,918	0,886	0,923	0,921	8 млн

Представлені у Таблиці 2 результати дозволяють зробити висновки щодо ефективності запропонованої моделі MobileMamba-U-Net порівняно з базовими архітектурами U-Net та U-Net 3+. Зокрема, запропонована модель демонструє найвищий показник Dice coefficient (0,918), що свідчить про високу схожість між передбачуваною та реальною масками сегментації. У порівнянні з U-Net, покращення становить близько 7,3%, а порівняно з U-Net 3+ — 1,6%. Це підтверджує здатність MobileMamba-U-Net забезпечувати більш точну сегментацію меж стоматологічних структур.

Метрика IoU (Intersection over Union) також демонструє переваги запропонованої моделі. Значення IoU для MobileMamba-U-Net становить 0,886, що перевищує відповідні показники для U-Net (0,849) і U-Net 3+ (0,883). Такий результат підтверджує кращу здатність моделі уникати хибно передбачених або пропущених пікселів, що є критично важливим для сегментації медичних зображень.

Щодо точності (precision) і чутливості (recall), MobileMamba-U-Net також демонструє найкращі результати. Точність становить 0,923, що свідчить про здатність моделі передбачати позитивні пікселі, які насправді є правильними, тоді як значення recall (0,921) свідчить про здатність моделі виявляти всі позитивні пікселі. Порівняно з U-Net, ці показники зросли на 9,8% і 4%, відповідно.

Важливою перевагою запропонованої моделі є значне скорочення кількості параметрів. MobileMamba-U-Net має лише 8 млн параметрів, що втричі менше, ніж у базовій U-Net (32 млн) та на 69% менше, ніж у U-Net 3+ (26 млн). Така оптимізація дозволяє знизити обчислювальні витрати,

зберігаючи при цьому високу точність, що є важливим для використання на мобільних пристроях та в умовах обмежених ресурсів.

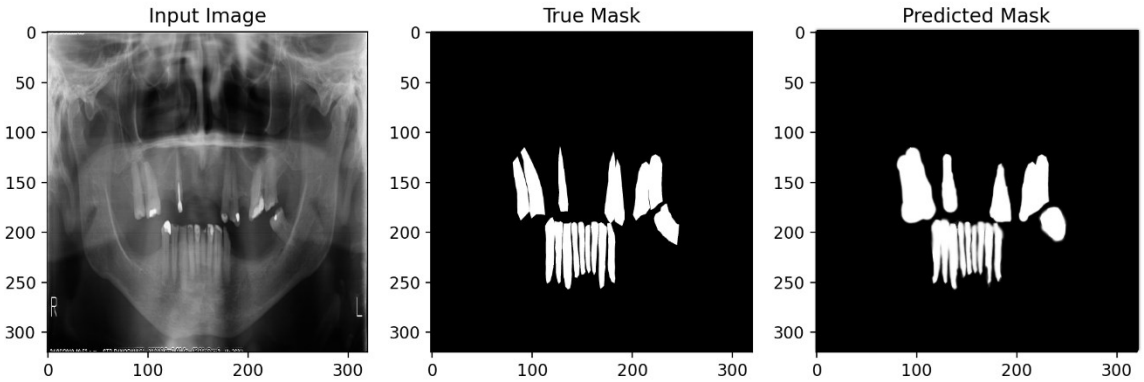


Рис. 2. Приклад сегментації стоматологічних зображень за допомогою MobileMamba-U-Net

Рисунок 2 ілюструє процес сегментації стоматологічного зображення за допомогою MobileMamba-U-Net. Ліва частина демонструє вхідне рентгенівське зображення, яке є початковими даними для моделі. У центральній частині наведено "True Mask" — реальну маску сегментації, створену експертами, що слугує еталоном для оцінювання якості передбачень. Права частина містить "Predicted Mask" — передбачену маску сегментації, отриману від запропонованої моделі.

Маска, створена MobileMamba-U-Net, демонструє високу точність у визначенні меж зубів і кореневих каналів, зберігаючи важливі деталі та точну геометрію структур. Чітке відтворення навіть дрібних елементів підтверджує ефективність інтеграції MobileNet та Vision Mamba Layers для забезпечення високоякісної сегментації стоматологічних зображень.

Висновки з даного дослідження
і перспективи подальших розвідок у даному напрямі

У цьому дослідженні було запропоновано та проаналізовано легковагову архітектуру MobileMamba-U-Net для сегментації стоматологічних зображень. Модель інтегрує MobileNet як енкодер, що забезпечує ефективне вилучення ознак із суттєвим зниженням обчислювальних витрат завдяки використанню Depthwise Separable Convolutions. Vision Mamba Layers, інтегровані у декодер, дозволяють моделювати локальні й глобальні просторові залежності, що підвищує точність відновлення меж і структур зубів. Така комбінація компонентів забезпечує високу продуктивність при значному зменшенні кількості параметрів.

Результати експериментів підтвердили, що запропонована архітектура перевершує базові моделі U-Net та U-Net 3+ за ключовими метриками, такими як Dice coefficient та Intersection over Union (IoU). MobileMamba-U-Net забезпечує точніше відтворення контурів стоматологічних структур і зберігає високу чутливість та точність передбачень навіть у випадках з шумами та неоднорідним освітленням. Особливо важливою є здатність моделі зменшувати кількість обчислювальних операцій, що робить її придатною для застосування на мобільних пристроях і в умовах обмежених ресурсів.

Висновки також демонструють, що запропонована модель є практичною для впровадження у клінічну практику. Її застосування може сприяти автоматизації діагностичних процесів, поліпшенню якості лікування завдяки більш точному визначенню меж зубів та їхніх структур, а також підвищенню доступності методів аналізу завдяки можливості роботи на портативних пристроях.

Подальші дослідження в цьому напрямі можуть бути спрямовані на розширення обсягів та різноманітності наборів даних для підвищення узагальнюваності результатів. Включення даних із різних клінічних умов і додаткових типів рентгенівських зображень сприятиме ще більшій адаптації моделі до реальних сценаріїв. Удосконалення архітектури шляхом впровадження сучасних методів оптимізації, таких як квантовані нейронні мережі, дозволить ще більше скоротити обчислювальні витрати без втрати точності. Крім того, інтеграція attention-механізмів може забезпечити більш детальне моделювання складних структур, що зробить сегментацію ще точнішою.

Не менш важливим є тестування моделі в реальних умовах, зокрема у мобільних додатках для стоматологічної діагностики. Це дозволить оцінити її продуктивність у реальному часі та виявити можливі аспекти для вдосконалення. Також перспективним є використання додаткових методів передобробки даних, таких як алгоритми видалення шуму та підвищення контрастності зображень, що дозволить зменшити вплив артефактів на результати сегментації.

Запропонована архітектура MobileMamba-U-Net є перспективним рішенням для задач автоматизованої сегментації стоматологічних зображень, поєднуючи ефективність, точність і обчислювальну оптимальність. Її вдосконалення та адаптація до клінічних умов відкривають широкі можливості для використання в медицині та стоматологічній практиці.

Література

1. Beheshti, N., et al. (2020). Squeeze U-Net: A memory and energy efficient image segmentation network. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 1495–1504. https://doi.org/10.1109/CVPRW50498.2020.00190

2. Alharbi, S. S., et al. (2023). Detection of cavities from dental panoramic X-ray images using nested U-Net models. *Applied Sciences, 13(23)*, 12771. https://doi.org/10.3390/app132312771

3. Yousef, R., et al. (2023). U-Net-based models towards optimal MR brain image segmentation. *Diagnostics, 13(9)*, 1624. https://doi.org/10.3390/diagnostics13091624

4. Deng, Y., et al. (2022). ELU-Net: An efficient and lightweight U-Net for medical image segmentation. *IEEE Access, 10*, 35932–35941. https://doi.org/10.1109/ACCESS.2022.3163711

5. Liao, W., et al. (2024). LightM-UNet: Mamba assists in lightweight UNet for medical image segmentation. *arXiv preprint arXiv:2403.05246*. https://doi.org/10.48550/arXiv.2403.05246

6. Wang, Z., et al. (2024). Mamba-UNet: UNet-like pure visual Mamba for medical image segmentation. *arXiv preprint arXiv:2402.05079*. https://doi.org/10.48550/arXiv.2402.05079

7. Velasco, J., et al. (2019). A smartphone-based skin disease classification using MobileNet CNN. *International Journal of Advanced Trends in Computer Science and Engineering, 8(5)*, 2632–2637. https://doi.org/10.30534/ijatcse/2019/116852019

8. Wu, R., et al. (2024). UltraLight Vision Mamba UNet: Parallel Vision Mamba significantly reduces parameters for skin lesion segmentation. *arXiv preprint arXiv:2403.20035*. https://doi.org/10.48550/arXiv.2403.20035

9. Du, Q., et al. (2024). A mixed Mamba U-Net for prostate segmentation in MR images. *Scientific Reports, 14*, Article 19976. https://doi.org/10.1038/s41598-024-71045-7

10. Wang, J., et al. (2024). LKM-UNet: Large kernel vision Mamba U-Net for medical image segmentation. *Scientific Reports, 14*, Article 19982. https://doi.org/10.1038/s41598-024-71045-7

11. Chaturvedi, S., et al. (2020). Skin lesion analyser: An efficient seven-way multi-class skin cancer classification using MobileNet. *Advances in Intelligent Systems and Computing*, 225–234. https://doi.org/10.1007/978-981-15-3383-9_15

12. Tang, H., et al. (2024). Rotate to scan: UNet-like Mamba with triplet SSM module for medical image segmentation. *arXiv preprint arXiv:2403.17701*. https://doi.org/10.48550/arXiv.2403.17701

13. Branch, M. V. L., et al. (2021). Polyp segmentation in colonoscopy images using U-Net-MobileNetV2. *Proceedings of the Federal University of Santa Catarina*. https://doi.org/10.12345/arXiv.2103.15715

References

1. Beheshti, N., et al. (2020). Squeeze U-Net: A memory and energy efficient image segmentation network. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, 1495–1504. https://doi.org/10.1109/CVPRW50498.2020.00190

2. Alharbi, S. S., et al. (2023). Detection of cavities from dental panoramic X-ray images using nested U-Net models. Applied Sciences, 13(23), 12771. https://doi.org/10.3390/app132312771

3. Yousef, R., et al. (2023). U-Net-based models towards optimal MR brain image segmentation. Diagnostics, 13(9), 1624. https://doi.org/10.3390/diagnostics13091624

4. Deng, Y., et al. (2022). ELU-Net: An efficient and lightweight U-Net for medical image segmentation. IEEE Access, 10, 35932–35941. https://doi.org/10.1109/ACCESS.2022.3163711

5. Liao, W., et al. (2024). LightM-UNet: Mamba assists in lightweight UNet for medical image segmentation. arXiv preprint arXiv:2403.05246. https://doi.org/10.48550/arXiv.2403.05246

6. Wang, Z., et al. (2024). Mamba-UNet: UNet-like pure visual Mamba for medical image segmentation. arXiv preprint arXiv:2402.05079. https://doi.org/10.48550/arXiv.2402.05079

7. Velasco, J., et al. (2019). A smartphone-based skin disease classification using MobileNet CNN. International Journal of Advanced Trends in Computer Science and Engineering, 8(5), 2632–2637. https://doi.org/10.30534/ijatcse/2019/116852019

8. Wu, R., et al. (2024). UltraLight Vision Mamba UNet: Parallel Vision Mamba significantly reduces parameters for skin lesion segmentation. arXiv preprint arXiv:2403.20035. https://doi.org/10.48550/arXiv.2403.20035

9. Du, Q., et al. (2024). A mixed Mamba U-Net for prostate segmentation in MR images. Scientific Reports, 14, Article 19976. https://doi.org/10.1038/s41598-024-71045-7

10. Wang, J., et al. (2024). LKM-UNet: Large kernel vision Mamba U-Net for medical image segmentation. Scientific Reports, 14, Article 19982. https://doi.org/10.1038/s41598-024-71045-7

11. Chaturvedi, S., et al. (2020). Skin lesion analyser: An efficient seven-way multi-class skin cancer classification using MobileNet. Advances in Intelligent Systems and Computing, 225–234. https://doi.org/10.1007/978-981-15-3383-9_15

12. Tang, H., et al. (2024). Rotate to scan: UNet-like Mamba with triplet SSM module for medical image segmentation. arXiv preprint arXiv:2403.17701. https://doi.org/10.48550/arXiv.2403.17701

13. Branch, M. V. L., et al. (2021). Polyp segmentation in colonoscopy images using U-Net-MobileNetV2. Proceedings of the Federal University of Santa Catarina. https://doi.org/10.12345/arXiv.2103.15715