



В. Р. Рибак¹, Ю. В. Шабатура²

¹ Національний лісотехнічний університет України, м. Львів, Україна

² Національна академія сухопутних військ ім. гетьмана Петра Сагайдачного, м. Львів, Україна

ОЦІНЮВАННЯ ЕФЕКТИВНОСТІ ЗАСТОСУВАННЯ МОДЕЛІ U-NET 3+ У ПРОЦЕСІ СЕГМЕНТАЦІЇ АНАТОМІЧНИХ СТРУКТУР РЕНТГЕНІВСЬКИХ ЗНІМКІВ ЗУБІВ

Автоматизація процесів аналізу цифрових зображень є важливим завданням сучасності. Особливо вона є актуальною для медичної сфери, де це сприяє підвищенню ефективності дослідження та зменшенню впливу людського чинника. Досліджено особливості застосування архітектури моделі U-Net 3+ для здійснення процедури автоматизованої сегментації цифрових стоматологічних рентгенівських зображень. Встановлено, що використання класичних методів сегментації зображень, таких як порогові методи та методи кластеризації, методи активних контурів і регіонального зростання, мають обмежену ефективність під час оброблення складних зображень, що зумовлено як складністю анатомічних структур, так і наявністю областей із подібним контрастом, що ускладнює точне виділення меж між об'єктами. Оцінено вплив застосування глибокого навчання архітектури типу U-Net та її модифікацій на якість сегментації складних стоматологічних структур. Охарактеризовано ефективність використання моделі U-Net 3+, яка завдяки своїм повношкальним зв'язкам забезпечує інтеграцію інформації з різних рівнів мережі та дає змогу зберігати як глобальний, так і локальний контексти. З'ясовано, що модель U-Net 3+ демонструє покращену здатність до сегментації структур зубів на рентгенівських зображеннях порівняно з класичною архітектурою типу U-Net завдяки її архітектурним удосконаленням, досягаючи значень Dice Coefficient та IoU на високому рівні. Оцінено точність сегментації на цифрових зображеннях дрібних анатомічних структур, таких як корені зубів і зони зі складною морфологією, що часто зумовлюють труднощі під час використання традиційних методів. Досліджено можливість безперервного навчання моделі U-Net 3+ на нових даних без втрати точності на попередніх вибірках, що є важливим для її довготривалого використання в динамічному середовищі. Встановлено, що модель U-Net 3+ має високу стійкість до перенавчання і до змін якості вхідних зображень, в т.ч. шуми та варіації освітлення, та демонструє стабільні результати на валідаційних вибірках, що підтверджує її ефективність для впровадження та використання у реальних умовах. Описано перспективи застосування моделі для автоматизованого аналізу стоматологічних знімків, у т.ч. інтеграцію у системи підтримки прийняття рішень, що дасть змогу значно підвищити швидкість і точність діагностики.

Ключові слова: глибоке навчання; сегментація зображень; комп'ютерний зір; стоматологічні рентгенівські зображення; автоматизація діагностики.

Вступ / Introduction

Цифрова рентгенографія є потужним інструментом отримання об'єктивної первинної інформації в медичній практиці. Цифрові рентгенівські зображення мають важливе значення у стоматології. Вони дають можливість стоматологам виявляти приховані патології, які неможливо діагностувати візуальним оглядом, зокрема карієс на початкових стадіях, захворювання ясен, кісти, переломи та інші проблеми з кістковою тканиною. Однак, інтерпретація цих зображень вимагає високої кваліфікації та досвіду, оскільки анатомічні особливості пацієнта, а також варіації якості зображень можуть ускладнювати їхній аналіз.

Варто зауважити, що окрім технічних характеристик обладнання, на процес діагностики впливають і людські чинники. Фізіологічний стан лікаря, такий як втома чи перевантаження, може істотно знижувати здатність фо-

кусуватися на деталях зображення. Психологічний тиск, стрес або інтенсивний графік роботи також впливають на здатність лікаря приймати правильні рішення, збільшуючи ризик діагностичних помилок.

Отже, цифрові рентгенівські зображення, хоча і є важливим інструментом для встановлення діагнозу, водночас, стають джерелом додаткового навантаження та можливих діагностичних помилок стоматологів, що потребує пошуку нових підходів до їх оброблення та аналізу.

Сучасні технології штучного інтелекту, зокрема комп'ютерний зір, активно трансформують методи аналізу цифрових зображень, що дає змогу автоматизувати складні завдання, які раніше вимагали тільки ручної праці людини. Комп'ютерний зір дає змогу розпізнавати та інтерпретувати зображення, що є основою для багатьох інноваційних рішень зокрема й у сфері охорони здоров'я. Штучний інтелект не тільки здатен виявляти ознаки

Інформація про авторів:

Рибак Володимир Романович, магістр, аспірант, кафедра інженерії програмного забезпечення.

Email: volodymyr.rybak@nltu.edu.ua; <https://orcid.org/0009-0006-8809-978X>

Шабатура Юрій Васильович, д-р техн. наук, професор, кафедра електромеханіки та електроніки.

Email: shabaturayuriy@gmail.com; <https://orcid.org/0000-0002-9961-1244>

Цитування за ДСТУ: Рибак В. Р., Шабатура Ю. В. Оцінювання ефективності застосування моделі U-Net 3+ у процесі сегментації анатомічних структур рентгенівських знімків зубів. Науковий вісник НЛТУ України. 2024, т. 34, № 7. С. 145–152.

Citation APA: Rybak, V. R., & Shabatura, Yu. V. (2024). Evaluation of the U-Net 3+ model efficiency in segmenting anatomical structures on dental x-ray images. *Scientific Bulletin of UNFU*, 34(7), 145–152. <https://doi.org/10.36930/40340718>

захворювань, але й робить це з високою точністю та надійністю, зменшуючи вплив людського чинника, який може призводити до помилок через втому або перевантаження лікарів.

Одним із ключових завдань комп'ютерного зору в медицині є сегментація цифрових зображень – процес поділу зображення на окремі області, що дає змогу ідентифікувати важливі структури для їх подальшого аналізу. З розвитком глибокого навчання, зокрема згорткових нейронних мереж, стало можливим створення моделей, здатних автоматизувати цей процес. Важливим проривом у цій галузі стало впровадження архітектур, таких як U-Net, які спеціально розроблені для сегментації цифрових медичних зображень.

Модель U-Net 3+ є однією з найсучасніших і найбільш потужних версій цієї архітектури, що демонструє її високу ефективність у вирішенні завдань сегментації зображень. Завдяки своїй здатності працювати з багатоплановими зображеннями та відновлювати складні контури об'єктів, вона має потенціал значно покращити діагностичні процеси у стоматології.

Об'єкт дослідження – сегментація анатомічних структур рентгенівських знімків зубів.

Предмет дослідження – методи та підходи до застосування моделі U-Net 3+ для автоматичної сегментації анатомічних структур на рентгенівських знімках зубів, які забезпечують підвищення точності та ефективності діагностики стоматологічних захворювань.

Мета роботи – дослідити можливості використання моделі U-Net 3+ для автоматичної сегментації цифрових стоматологічних рентгенівських зображень зубів, що дасть змогу підвищити точність діагностики їх стоматологічних патологій, зокрема карієсу, й отримати стабільність результатів.

Для досягнення зазначеної мети визначено такі основні завдання дослідження:

- проаналізувати сучасні підходи до автоматизованої сегментації медичних зображень, включно з архітектурами на підставі глибокого навчання типу U-Net та її модифікації, що дасть змогу визначити оптимальні методи для підвищення точності сегментації, адаптовані до специфічних медичних завдань, таких як аналіз стоматологічних рентгенівських знімків, і зменшити ризик помилкових діагнозів;
- розробити та адаптувати модель U-Net 3+ для сегментації цифрових стоматологічних знімків, які допоможуть підвищити точність діагностики та пришвидшать процес оброблення зображень;
- експериментально дослідити продуктивності моделі U-Net 3+ на реальних клінічних даних, використовуючи відкриті набори зображень, що дасть можливість оцінити її продуктивність за метриками точності Dice та IoU;
- порівняти ефективність використання моделі U-Net 3+ із класичною архітектурою типу U-Net, що дасть змогу визначити їх можливості для подальшого вдосконалення.

Аналіз останніх досліджень та публікацій. З розвитком технологій глибокого навчання, зокрема архітектур типу U-Net, було досягнуто значного прогресу в автоматизації сегментації медичних зображень. Архітектура U-Net та її модифікації стали популярними завдяки своїй здатності ефективно сегментувати складні анатомічні структури на медичних зображеннях, в т.ч. рентгенівські зображення у стоматології. Їх висока ефективність забезпечується використанням симетричних зв'язків між шарами та можливістю працювати з обме-

женими наборами даних, що робить ці підходи універсальними для медичних застосувань.

Одним із значних проривів є розроблення моделі U-Net 3+, яка поєднує в собі багатообіцяючі ознаки своїх попередників і покращує точність за допомогою глибоких з'єднань і функцій багаторівневого навчання. Наприклад, дослідження Хуїміна Хуана [10] та його колеги вказує на те, що модель U-Net 3+ підвищує ефективність завдяки оптимізованим підходам до зменшення кількості її параметрів.

Попередні дослідження вказують на ефективність архітектур на підставі архітектури U-Net, але також вказують на потребу подальших вдосконалень для роботи з малими об'єктами та складними текстами. Наприклад, у роботі [3] пропонують модель SE-RegUNet для сегментації коронарних судин. Завдяки використанню RegNet-енкодерів і оптимізації Ranger21 модель демонструє високу точність діагнозу із Dice Score 0.76 та точністю 0.97, що робить її перспективною для клінічного застосування.

Окрім цього, полегшені архітектури, як-от архітектура ELU-Net [8], скорочують споживання пам'яті та обчислювальні витрати без втрати точності. Моделі на зразок Squeeze U-Net [2] адаптуються для застосування на пристроях з обмеженими ресурсами, що є важливим для мобільних додатків.

Застосування моделі U-Net 3+ для сегментації кров'яних клітин [1, 16] підтвердило її ефективність. Використання оптимізації Байєса сприяло підвищенню точності моделі, запобігаючи перенавчанню, що підтверджується метриками IoU (0,93) та Dice Score (0,96).

Останніми роками також розширюється застосування мережі GAN (англ. *Generative Adversarial Networks*) для збільшення вибірки даних у медичній сегментації зображень [18]. Це дає змогу підвищити стабільність роботи моделей навіть за обмежених обсягів навчальних даних.

Загалом розвиток архітектур типу U-Net, зокрема моделі U-Net 3+ та її модифікацій, забезпечує покращення точності сегментації зображень та продуктивності відповідної процедури, сприяючи розвитку автоматизованих підходів у медичній діагностиці.

Матеріали та методи дослідження. У дослідженні застосовано модель U-Net 3+ для сегментації стоматологічних зображень завдяки її здатності інтегрувати інформацію з різних рівнів мережі, що забезпечує високу точність. Навчальний набір даних DENTEX містить панорамні рентгенівські знімки та супровідні маски уражених областей. Усі дані були підготовлені через розширення вибірки та їх нормалізацію.

Результати дослідження та їх обговорення / Research results and their discussion

Теоретичні особливості медичної сегментації зображень. Сегментація зображень є однією з ключових задач комп'ютерного зору і відіграє важливу роль у різних галузях, зокрема в медицині. Глибоке навчання, зокрема згорткової нейронної мережі CNN (англ. *Convolutional Neural Networks*), стало основним інструментом для автоматизації складних завдань сегментації зображень завдяки здатності навчатися на великих обсягах даних і витягувати істотні характеристики без ручного втручання. У медичній діагностиці, де точність і швидкість аналізу мають вирішальне значення, глибоке нав-

чання значно підвищило ефективність аналізу медичних зображень, таких як рентгенівські знімки, комп'ютерна томографія (КТ) і магнітно-резонансна томографія (МРТ).

Класичні підходи до сегментації зображень містили алгоритми на підставі порогових значень, кластеризацію або методи активних контурів. Однак, ці підходи мали обмежену ефективність у разі складної анатомії або низької якості зображень, що часто трапляється в медичних даних. Розвиток глибокого навчання штучних нейронних мереж привів до створення архітектур, здатних автоматично навчатися та адаптуватися до специфічних особливостей медичних зображень. Однією з найважливіших архітектур для сегментації медичних зображень є архітектура U-Net, розроблена для біомедичних завдань і відома своєю здатністю сегментувати складні структури навіть на невеликих наборах даних.

Архітектура U-Net відкрила нові горизонти для медичних досліджень, даючи можливість точніше виділяти анатомічні структури на різних типах зображень. Однак, розвиток цієї архітектури не зупинився на базовій версії – з'явилося декілька модифікацій, таких як моделі U-Net++ і U-Net 3+, які продовжують покращувати результати сегментації зображень завдяки глибоким та інноваційним архітектурним рішенням [20].

Отже, сегментація зображень за допомогою глибокого навчання стала критично важливою в медичній діагностиці, даючи можливість автоматизувати процеси, які раніше потребували значних зусиль лікарів, і знижуючи ризик помилок, пов'язаних із людським чинником.

Архітектуру U-Net вперше представила у 2015 р. команда дослідників під керівництвом Олафа Роннеберґера для задач біомедичної сегментації зображень [20]. Основна ідея цієї моделі полягає у симетричній архітектурі, яка складається з двох частин: контракційного шляху (downsampling) та експансійного шляху (upsampling). Контракційний шлях виконує згортки та по-

етапне зменшення розмірів зображення, виділяючи основні характеристики, тоді як експансійний шлях відновлює просторову інформацію та забезпечує точну локалізацію об'єктів на зображенні.

Основною особливістю архітектури U-Net є з'єднання між шарами, що дає змогу передавати інформацію з ранніх шарів до пізніх, запобігаючи втраті важливих деталей на етапі downsampling. Це робить модель особливо ефективною для медичних зображень, де важливо зберегти деталі малих анатомічних структур.

Архітектура U-Net стала стандартом у медичній сегментації зображень завдяки своїй здатності обробляти відносно невеликі набори даних, що є типовою ситуацією в біомедичних дослідженнях. Це стало можливим завдяки використанню глибоких згорткових шарів і стратегії поетапного відновлення зображення, яка забезпечує високу точність сегментації навіть для зображень зі складними структурами.

Архітектура типу U-Net виявилася ефективною у багатьох біомедичних задачах, таких як сегментація клітин, органів і патологій на різних типах медичних зображень. Вона дала змогу автоматизувати процес аналізу складних біомедичних даних, скоротивши час, необхідний для діагностики, і підвищила точність виявлення захворювань.

Модель U-Net 3+ є важливою модифікацією архітектури типу U-Net, яка об'єднує ключові удосконалення з попередніх версій і додає нові інноваційні рішення для ще точнішого сегментування зображень. Головною особливістю моделі U-Net 3+ є використання повношкільних зв'язків, які забезпечують інтеграцію інформації з різних рівнів мережі, від найглибших до найповерхневіших. Це дає змогу моделі ефективно зберігати й використовувати як глобальну, так і локальну інформацію під час сегментації зображень, що особливо важливо для медичних зображень, де необхідно зберегти деталі на всіх масштабах [10].

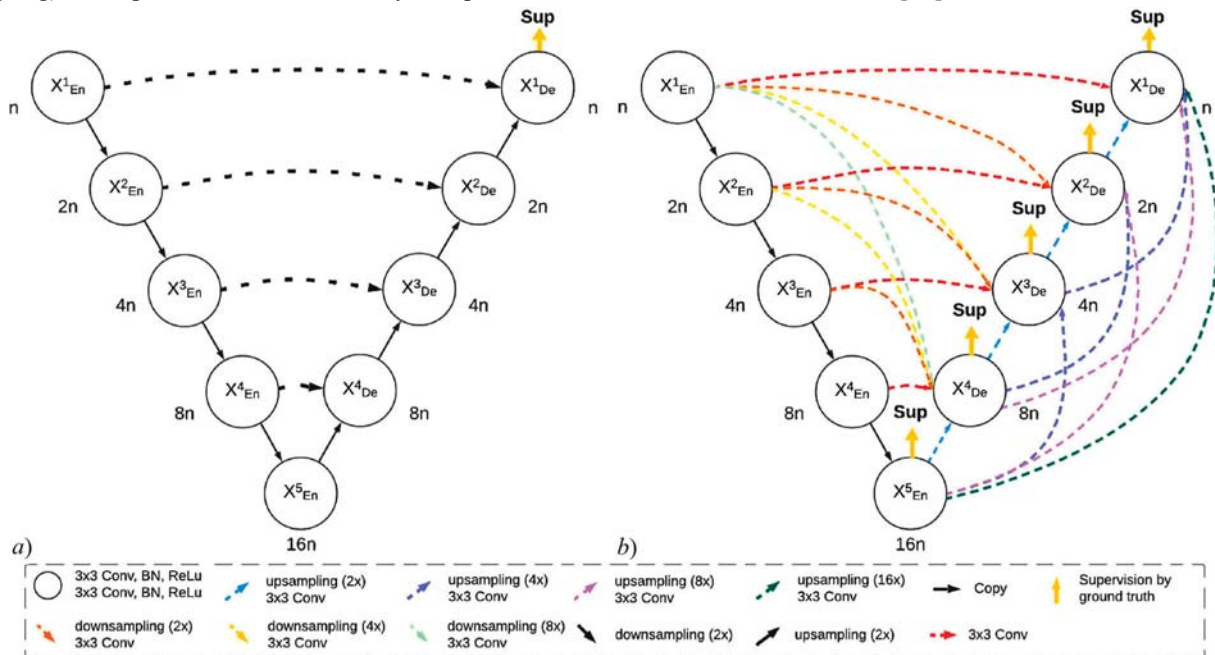


Рис. 1. Архітектури моделей U-Net (a) та U-Net 3+ (b) із різними типами пропускових зв'язків між рівнями мережі / Architecture of U-Net and U-Net 3+ models with different types of skip connections between network levels [14]

На рис. 1 наведено архітектуру моделей U-Net та U-Net 3+ [14], на яких використано такі позначення:

- X^n_{En} – ознаки на n -му рівні енкодера;

- X^n_{De} – ознаки на n -му рівні декодера;
- $3 \times 3 \text{ Conv}$ – згортковий шар із фільтром 3×3 ;
- BN – пакетна нормалізація (англ. *Batch Normalization*);

- *ReLU* – функція активації (англ. Rectified Linear Unit);
- Downsampling (2x, 4x, 8x, 16x) – зменшення розмірності з відповідним коефіцієнтом;
- Upsampling (2x, 4x, 8x, 16x) – збільшення розмірності;
- Сору – операція копіювання ознак між рівнями;
- Supervision by ground truth – порівняння з еталонними даними для навчання.

Модель U-Net використовує прості пропускні зв'язки між відповідними рівнями енкодера та декодера, що дає змогу передавати локальні ознаки, але може втрачати глобальний контекст. Модель U-Net 3+ усуває цю проблему завдяки повномасштабним пропускним зв'язкам, що агрегують ознаки з усіх рівнів мережі, забезпечуючи кращу сегментацію.

Повношкільні зв'язки в моделі U-Net 3+ дають змогу зменшити проблему втрати просторової інформації, яка може виникати під час багатоетапного зменшення дискретизації. Завдяки цьому модель може краще зберігати контексти, необхідні для точного сегментування малих або тонких структур, таких як судини чи нерви, що робить її особливо корисною для складних медичних завдань, зокрема в стоматології [20].

Ще однією важливою перевагою моделі U-Net 3+ є зменшення кількості параметрів моделі, що робить її ефективнішою з погляду обчислювальних ресурсів, але без втрати точності. Це дає змогу використовувати модель на пристроях із обмеженими ресурсами або в умовах, де обчислювальна потужність обмежена [20].

Модель U-Net 3+ також інтегрує різні типи блоків для оброблення багатосарової інформації, що робить її адаптивною до різних типів даних. Це робить модель універсальною для багатьох медичних завдань, зокрема для сегментації зображень зубів й інших анатомічних структур на рентгеновських знімках, де важливо виділити як великі, так і дрібні елементи [10]. Ці вдосконалення дають можливість моделі U-Net 3+ перевершувати попередні моделі, такі як модель U-Net і U-Net++, у задачах сегментації зображень, де важлива як точність, так і ефективність [10, 20].

Опис експериментальних умов. Для цього дослідження було використано набір даних DENTEX, що містить панорамні рентгеновські знімки, отримані з трьох різних стоматологічних установ. Ці дані було отримано за стандартними клінічними умовами, але з використанням різного обладнання та протоколів зображення, що дало можливість забезпечити варіативність набору даних і підвищити універсальність тренувальної моделі. Набір даних містив різні випадки стоматологічних захворювань, таких як карієс та інші патології зубів і тканин, що їх оточують.

Перед тренуванням модель U-Net 3+ пройшла процес підготовки даних, що містив декілька етапів. Спочатку проводилася нормалізація зображень для забезпечення єдиного масштабу яскравості та контрасту. Потім використовувалися методи розширення вибірки, зокрема горизонтальне і вертикальне перевертання, обертання, а також дзеркальне відображення зображень. Ці методи дали змогу істотно збільшити обсяг даних для тренування моделі та зробити її стійкішою до різноманітних умов зображення.

Для навчання моделі U-Net 3+ було обрано такі основні параметри [20]:

- оптимізатор: використовувався Adam (англ. *Adaptive Moment Estimation*) із початковою швидкістю навчання 1e-4;

- функція втрат: використовувалася cross-entropy loss, яка оптимізує модель на підставі ймовірнісної класифікації пікселів;
- кількість епох: модель тренувалася протягом 30 епох;
- розмір пакета: було використано розмір пакета 32 зображення на ітерацію;
- dropout: застосовано dropout із ймовірністю 0.5 для регуляризації та покращення здатності до генералізації;
- L2-регуляризація: використовувалася для зменшення ризику перенавчання.

Дані були розділені на навчальну, валідаційну та тестову вибірку у співвідношенні 70:15:15. Валідаційна вибірка використовувалася для оцінювання продуктивності моделі під час навчання, а тестова вибірка слугувала для остаточної перевірки ефективності моделі на нових даних, що дає змогу оцінити її надійність і практичну цінність у реальних умовах використання.

Опис основних метрик. Для оцінювання ефективності моделі U-Net 3+ було використано декілька ключових метрик, які забезпечують комплексне оцінювання якості сегментації зображень:

1. Dice coefficient – одна з найважливіших метрик для оцінювання якості сегментації зображень. Dice coefficient розраховує схожість між передбачуваною маскою і реальною, зважаючи на правильність пікселів сегментації зображень, обчислюють за такою формулою:

$$Dice = \frac{2 \cdot TP}{2 \cdot TP + FP + FN} \quad (1)$$

2. Intersection over union (IoU), або коефіцієнт Жаккара, оцінює співвідношення між перетином та об'єднанням передбачуваної та реальної масок. Цю метрику також використовують для оцінювання точності сегментації зображень, але вона більш строга, ніж метрика Dice coefficient, оскільки враховує як пропущені, так і хибно передбачені пікселі. Формула для обчислення IoU має такий вигляд:

$$IoU = \frac{TP}{TP + FP + FN} \quad (2)$$

3. Точність (accuracy) вимірює загальну правильність передбачуваних пікселів, обчислюють за формулою:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (3)$$

4. Чутливість (recall) оцінює здатність моделі виявляти всі позитивні пікселі. Її обчислюють як відношення правильно передбачених позитивних пікселів до загальної кількості позитивних пікселів, а саме:

$$Recall = \frac{TP}{TP + FN} \quad (4)$$

5. Точність передбачення (precision) оцінює здатність моделі передбачати позитивні пікселі, які насправді є позитивними, обчислюють за такою формулою:

$$Precision = \frac{TP}{TP + FP} \quad (5)$$

де: *TP* – кількість правильно передбачених пікселів; *FP* – кількість хибно передбачених позитивних пікселів; *FN* – кількість пропущених пікселів, які повинні були бути позитивними; *TN* – кількість правильно передбачених негативних пікселів.

Хоча метрика точність передбачення є й корисною, однак, вона може бути упередженою у випадках, коли кількість негативних пікселів значно перевищує кількість позитивних (наприклад, для завдань сегментації зображень, де більшість пікселів є фоновими).

Усі зазначені вище метрики було використано для оцінювання продуктивності моделі U-Net 3+ на трену-

вальної, валідаційної і тестовій вибірках, що дало можливість авторам дослідження отримати повний огляд того, наскільки ефективно модель справляється із завданнями сегментації стоматологічних зображень.

Аналіз результатів на тестовій вибірці. Результати оцінювання ефективності моделі U-Net 3+ було наведено у вигляді графіків точності та втрат для навчальної та валідаційної вибірок (рис. 2). На графіку можна спостерігати, що модель демонструє стабільне зростання точності на навчальній вибірці, досягаючи значення близько 0.91 після 20 епох. Водночас, точність на валідаційній вибірці стабілізувалася на рівні 0.86 вже після 10 епох, що свідчить про узгодженість роботи моделі та відсутність істотного перенавчання.

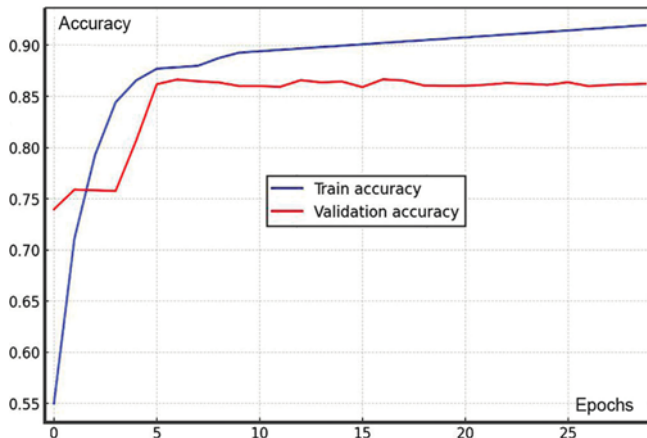


Рис. 2. Графічне подання результатів оцінювання точності тренувальної та валідаційної вибірок / Graphic presentation of the results of accuracy assessment of the training and validation samples

На підставі наведених у табл. 1 даних можна зробити висновок, що модель показує стабільні результати на валідаційних даних і не схильна до значного перенавчання, що є позитивним показником.

Табл. 1. Порівняння метрик на тренувальній і валідаційній вибірках / Comparison of metrics on training and validation samples

Вибірка	Accuracy	Dice	IoU	Precision	Recall
Навчальна	0,91	0,88	0,82	0,84	0,91
Валідаційна	0,86	0,82	0,79	0,81	0,88

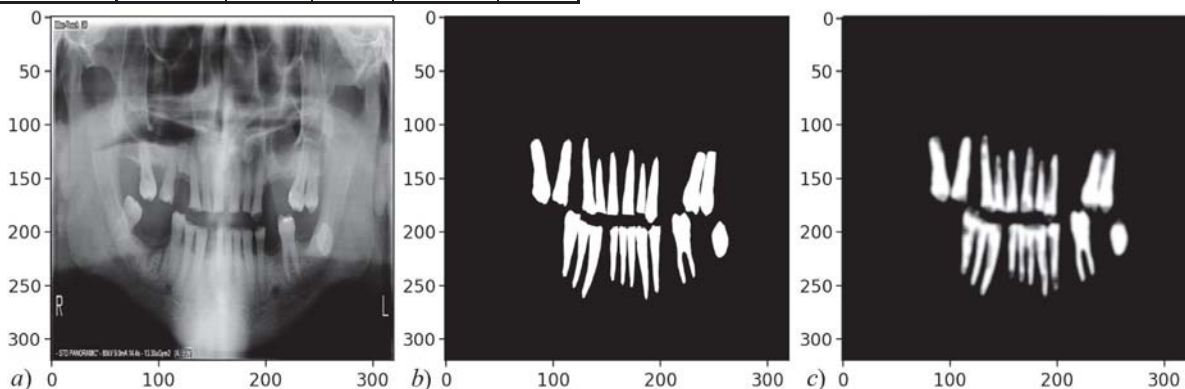


Рис. 3. Порівняння вхідного зображення, істинної маски та передбаченої маски для стоматологічного рентгенівського знімка / Comparison of the input image, true mask, and predicted mask for a dental X-ray image

На рис. 3 зображено три етапи оброблення стоматологічного рентгенівського знімка за допомогою моделі U-Net 3+ [14, 20]:

- input image (вхідне зображення): оригінальне рентгеновське зображення ротової порожнини, яке використовувалося як вхідний зразок для моделі. Зображення містить як м'які тканини, так і кісткові структури щелепи,

Як видно з графіка, модель досягає швидкої стабілізації на валідаційних даних, а різниця між навчальною та валідаційною точністю залишається в межах 0.05. Це свідчить про добру узгодженість моделі з результатами валідаційної вибірки, що є ознакою збалансованої архітектури та відповідно налаштованих гіперпараметрів.

Одним із аспектів цього дослідження є порівняння продуктивності моделі U-Net 3+ із базовою архітектурою U-Net. Модель U-Net стала стандартом для сегментації медичних зображень завдяки своїй здатності ефективно виділяти анатомічні структури навіть на невеликих наборах даних. Проте, вона має певні обмеження у здатності передавати інформацію з різних рівнів мережі, що може призводити до втрати деталей у разі складної анатомії.

Для оцінювання переваг моделі U-Net 3+ було проведено експериментальне порівняння двох моделей на однаковому наборі даних із використанням тих самих гіперпараметрів. Результати показали, що модель U-Net 3+ перевершує базову модель U-Net за всіма ключовими метриками.

Табл. 2. Порівняння ефективності моделей U-Net та U-Net 3+ / Comparison of the efficiency of U-Net and U-Net 3+ models

Модель	Dice	IoU	Precision	Recall	Параметри
U-Net	0,863	0,759	0,85	0,876	31 млн
U-Net 3+	0,91	0,833	0,9	0,918	25 млн

Згідно з даними табл. 2, модель U-Net 3+ демонструє значне покращення точності, особливо у показниках Dice Coefficient та IoU, що вказує на кращу здатність моделі розпізнавати межі анатомічних структур. Крім цього, завдяки використанню повношкільних зв'язків, модель U-Net 3+ ефективніше передає контекстну інформацію між рівнями, що дає змогу зберігати деталі анатомічних структур під час оброблення зображень з малими об'єктами.

Також варто відзначити, що модель U-Net 3+ вимагає менше параметрів для досягнення аналогічних або кращих результатів, що робить її більш ефективною з погляду обчислювальних ресурсів. Незважаючи на те, що тривалість тренування моделі дещо збільшується, це не є критичним фактором за врахування покращення точності сегментації зображень.

що ускладнює сегментацію зубів через наявність подібних за яскравістю і формою областей;

- true mask (істинна маска): еталонна маска (ground truth), яку використовують для оцінювання роботи моделі. На цій масці білим кольором виділено області, що відповідають зубам, а чорним кольором – фоніві області. Істинну маску було створено вручну експертами-стомато-

логами та використовують для порівняння з результатами сегментації зображень моделі;

- predicted mask (передбачена маска): маска, передбачена моделлю U-Net 3+. Як видно з рис. 3, передбачена моделлю маска загалом відповідає істинній масці, хоча в деяких областях є незначні відхилення. Модель успішно виділила основні структури зубів, але спостерігаються певні розмиті області, що може вказувати на потребу подальшого вдосконалення моделі для точнішого виділення контурів.

Зображення на рис. 3 демонструє ефективність використання моделі U-Net 3+ для сегментації стоматологічних знімків і вказує на потребу подальшої роботи над покращенням точності у складних зонах, таких як корені зубів або зони, де зуби щільно розташовані.

Результати проведених експериментів свідчать про значне покращення точності та ефективності сегментації зображень за допомогою моделі U-Net 3+ порівняно з класичною моделлю U-Net. Завдяки використанню повношкільних зв'язків, які інтегрують інформацію з різних рівнів мережі, модель U-Net 3+ змогла досягти вищих значень Dice Coefficient, IoU та Precision, що вказує на кращу здатність моделі зберігати важливі деталі під час сегментації складних структур зубів.

Одним із головних покращень, яке було відзначено в моделі U-Net 3+, є більш ефективне збереження просторової інформації завдяки комбінованій архітектурі з багатоступінним підсумовуванням та передачею контекстних даних між різними рівнями. Це дало можливість моделі краще впоратися з неоднорідністю стоматологічних знімків, де присутні дрібні деталі та складні структури, такі як корені зубів і стики між ними. Порівняно, класична модель U-Net продемонструвала певні обмеження у виділенні тонких деталей, що призвело до втрати контурів на складних ділянках.

Проте, варто відзначити, що попри загальне покращення точності, модель U-Net 3+ все ще демонструє певні труднощі під час сегментації тонких структур і малих об'єктів, особливо у випадках, де є часткове накладання зубів або деякі зуби мають низьку контрастність відносно фону. Це може бути спричинено недостатньою кількістю даних для навчання, а також потребою в більш специфічних методах після оброблення результатів. Окрім цього, певне розмиття контурів у передбачених масках може вказувати на потребу тоншого налаштування гіперпараметрів моделі.

Ще однією важливою особливістю моделі U-Net 3+ є зменшення кількості параметрів, що забезпечує вищу обчислювальну ефективність без втрати якості сегментації зображень. Це робить модель придатною для використання в умовах обмежених обчислювальних ресурсів або на мобільних пристроях, де класична архітектура типу U-Net може виявитися занадто важкою для реалізації.

Порівняльний аналіз результатів на навчальній та валідаційній вибірках показав, що модель U-Net 3+ має високу стабільність і не схильна до значного перенавчання. Однак, деяка різниця у точності між навчальною та валідаційною вибірками може свідчити про відмінності у структурі даних між вибірками або недостатнє узгодження з певними типами аномалій.

Обговорення результатів дослідження. Сегментація зображень із використанням сучасних моделей глибокого навчання, таких як архітектура типу U-Net та її модифікації, є предметом активних досліджень через їх

високу ефективність у вирішенні складних задач розпізнавання та аналізу цифрових зображень. Різні дослідники демонструють, що такі архітектури здатні забезпечувати отримання точних результатів у областях, де традиційні підходи, зокрема – порогові методи або методи активних контурів, виявляються малоефективними через неоднорідність даних або складність їх структури. Особливу увагу дослідження приділяють медичній візуалізації та аналізу біомедичних даних, що підтверджує актуальність використання глибоких мереж для автоматизації подібних процесів і підвищення точності та швидкості аналізу.

У дослідженні [17] автори розробили глибокий навчальний пайплайн для автоматизованої сегментації серцевих МРТ-даних, зокрема для пацієнтів із єдиною шлуночковою фізіологією. Використання моделі U-Net 3+ дало можливість досягти високої точності сегментації зображень, що підтверджують коефіцієнтами Dice на рівні 0.91 для кінцево-діастолічного та 0.86 – для кінцево-систолічного об'ємів. Система показала стабільність результатів на даних із різних лікарень і МРТ-пристроїв, знижуючи варіабельність, що часто виникає за ручної сегментації зображень. Цей підхід підтверджує ефективність автоматизації сегментації зображень у медичній сфері, що співзвучно з цілями нашого дослідження, яке спрямоване на покращення точності аналізу складних анатомічних структур на стоматологічних рентгенівських знімках.

У дослідженні [11] було наведено концепцію повністю згорткових мереж, що стали проривом у галузі семантичної сегментації зображень. Основна ідея полягала в адаптації наявних архітектур для піксельної сегментації зображень, що забезпечило передбачення для зображень будь-якого розміру. Використання skip connections дало можливість поєднувати семантичну інформацію з глибоких шарів із дрібномасштабною інформацією, що значно підвищило точність сегментації зображень. Дослідження проклало шлях до розробки більш складних архітектур, зокрема U-Net та її модифікацій, таких як модель U-Net 3+, які використовуються у медичній діагностиці, в т.ч. сегментацію стоматологічних рентгенівських знімків.

У роботі [20] запропоновано архітектуру U-Net++, яка вдосконалює класичну архітектуру типу U-Net внаслідок використання вкладених щільних зв'язків між енкодером та декодером. Це рішення дає змогу зменшити семантичний розрив між ознаками на різних рівнях мережі та підвищити точність сегментації зображень. Автори відзначають, що модель U-Net++ перевершила традиційну архітектуру типу U-Net, досягнувши на 3.9 % вищих значень IoU при розпізнаванні анатомічних структур та патологій на медичних зображеннях. Цей підхід узгоджують з нашим дослідженням, де використання моделі U-Net 3+ із повношкільними зв'язками дало можливість досягти високої точності у сегментації стоматологічних рентгенівських знімків.

У дослідженні [12] було наведено архітектуру Attention U-Net, яка розширює можливість класичної архітектури типу U-Net шляхом інтеграції механізму уваги (attention mechanism). Цей підхід дає змогу моделі автоматично виділяти найбільш релевантні області зображення під час сегментації зображень, що особливо важливо під час оброблення складних анатомічних структур. Автори зосереджують увагу на ефективності засто-

сування архітектури Attention U-Net для сегментації органів і патологій, зокрема підшлункової залози, де точне визначення контурів є вкрай важливим для діагностики. Результати дослідження підтверджують ефективність використання глибокого навчання архітектур із механізмом уваги для задач медичної сегментації зображень, що узгоджують з нашим підходом, де модель U-Net 3+ також забезпечує високу точність під час оброблення складних стоматологічних зображень.

Розвиток архітектур типу U-Net продовжується через вдосконалення пропускових з'єднань та поліпшення роботи з багаторівневими ознаками. Модель U-Net v2 [13] застосовує з'єднання, що збагачують кожен рівень ознак як семантичною інформацією з високих рівнів, так і деталями з низьких, що дає змогу краще визначати границі об'єктів і зменшувати кількість хибних сегментацій. Впровадження модуля SDI (англ. *Semantic and Detail Infusion*) забезпечує збалансоване поєднання низькорівневих і високорівневих ознак, покращуючи якість сегментації зображень. Експерименти на наборах даних ISIC 2017 та 2018 рр. продемонстрували перевагу моделі U-Net v2 над такими моделями, як модель UNet++ та модель SANet, з підвищеними метриками Dice та IoU. Окрім цього, модель оптимізує використання обчислювальних ресурсів і пришвидшує оброблення зображень, що робить її придатною для практичних медичних застосувань, вказуючи на актуальність комплексного підходу до роботи з ознаками в сучасних методах сегментації зображень.

Отже, внаслідок виконаної роботи можна сформулювати такі наукову новизну та практичну значущість результатів дослідження.

Наукова новизна отриманих результатів дослідження – удосконалено методику сегментації складних стоматологічних структур на рентгенівських знімках зубів через адаптацію архітектури U-Net 3+ до особливостей структури таких знімків, що сприяє підвищенню точності розпізнавання малих анатомічних структур і покращенню якості діагностики стоматологічних патологій пацієнтів.

Практична значущість результатів дослідження – результати дослідження можна застосувати для розроблення автоматизованих систем аналізу стоматологічних рентгенівських зображень з подальшою інтеграцією у системи підтримки прийняття рішень для підвищення точності та швидкості діагностики, що дає змогу мінімізувати вплив людського чинника на підготовку рішень та оптимізувати процеси оброблення зображень у клінічній практиці.

Висновки / Conclusions

Досліджено можливість використання моделі U-Net 3+ для автоматичної сегментації цифрових стоматологічних рентгенівських зображень зубів, що дало змогу підвищити точність діагностики їх стоматологічних патологій, зокрема карієсу, й отримати стабільність результатів. За результатами проведеного дослідження можна зробити такі основні висновки:

1. На підставі проведеного аналізу сучасних методів сегментації цифрових медичних зображень, побудованих на використанні класичних архітектур нейронних мереж U-Net та її модифікації, таких як моделі U-Net++ і U-Net 3+, з'ясовано, що вони мають різний ступінь ефективності у виділенні складних анатомічних структур.

2. Обґрунтовано переваги використання моделі U-Net 3+ у процесі сегментації цифрових стоматологічних рентгенівських зображень завдяки застосуванню багатошарових зв'язків, які забезпечують інтеграцію інформації з різних рівнів мережі.
3. Виконано навчання моделі U-Net 3+ на загальнодоступних наборах стоматологічних рентгенівських даних та підтверджено її ефективність для автоматичної сегментації навіть на обмежених наборах зображень. Це дає змогу скоротити тривалість підготовки даних та забезпечити стабільну продуктивність на різних валідаційних наборах.
4. Підтверджено перевагу моделі U-Net 3+ порівняно з класичною архітектурою типу U-Net внаслідок зменшення кількості параметрів моделі без втрати якості сегментації зображень, що робить її придатною для застосування в умовах обмежених обчислювальних ресурсів.
5. Виконане дослідження може стати основою для подальшого удосконалення методів сегментації цифрових зображень у стоматології, в т.ч. розширення наборів даних із залученням рідкісних патологій та застосуванням додаткових модулів, таких як механізми уваги, для підвищення адаптивності моделі до надскладних зображень.

References

1. Alruily, M., Ezz, M., Elmezain, M., & Ali, M. A. (2023). Breast ultrasound images augmentation and segmentation using GAN with identity block and modified U-Net 3+. *Sensors*, 23(8599), 1–16. <https://doi.org/10.3390/s23208599>
2. Beheshti, N., & Johnsson, L. (2020). Squeeze U-Net: A memory and energy efficient image segmentation network. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 1495–1504. <https://doi.org/10.1109/CVPRW50498.2020.00190>
3. Chang, S.-S., Yang, T.-H., Lin, W.-H., Huang, J.-W., & Lin, M.-C. (2024). Optimizing ensemble U-Net architectures for robust coronary vessel segmentation in angiographic images. *Scientific Reports*, 14(1), 1–11. <https://doi.org/10.1038/s41598-024-57198-5>
4. Chen, L.-C., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2018). Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(4), 834–848. <https://doi.org/10.1109/TPAMI.2017.2699184>
5. Chen, L.-C., Papandreou, G., Schroff, F., & Adam, H. (2017). Rethinking atrous convolution for semantic image segmentation. <https://doi.org/10.48550/arXiv.1706.05587>
6. Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F., & Adam, H. (2018). Encoder-decoder with atrous separable convolution for semantic image segmentation. In Proceedings of the European Conference on Computer Vision (ECCV), 801–818. https://doi.org/10.1007/978-3-030-01234-2_49
7. de Boer, P.-T., Kroese, D. P., Mannor, S., & Rubinstein, R. Y. (2005). A tutorial on the cross-entropy method. *Annals of Operations Research*, 134(1), 19–67. <https://doi.org/10.1007/s10479-005-5724-z>
8. Deng, Y., Hou, Y., Yan, J., & Zeng, D. (2022). ELU-Net: An efficient and lightweight U-Net for medical image segmentation. In *IEEE Access*, Vol. 10, 35932–35941. <https://doi.org/10.1109/ACCESS.2022.3163711>
9. Hrytsiuk, Yu. I. (2022). Software quality management system. *Ukrainian Journal of Information Technology*, 4(1), 01–20. <https://doi.org/10.23939/ujit2022.01.001>
10. Huang, H., Lin, L., Tong, R., Hu, H., Zhang, Q., Iwamoto, Y., Han, X., Chen, Y.-W., & Wu, J. (2020). UNet 3+: A full-scale connected UNet for medical image segmentation. In ICASSP 2020-45th International Conference on Acoustics, Speech, and Signal Processing, 1055–1059. <https://doi.org/10.1109/ICASSP40776.2020.9053405>
11. Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. *Proceedings of the IEEE*

- Conference on Computer Vision and Pattern Recognition (CVPR)*, 3431–3440. <https://doi.org/10.48550/arXiv.1411.4038>
12. Oktay, O., Schlemper, J., Folgoc, L. L., Lee, M., Heinrich, M. P., Misawa, K., Mori, K., McDonagh, S., Hammerla, N. Y., Kainz, B., Glocker, B., & Rueckert, D. (2018). Attention U-Net: Learning where to look for the pancreas. <https://doi.org/10.48550/arXiv.1804.03999>
 13. Peng, Y., Sonka, M., & Chen, D. Z. (2024). U-Net v2: Rethinking the skip connections of U-Net for medical image segmentation, 1–25. <https://doi.org/10.48550/arXiv.2311.17791>
 14. Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, 234–241. Springer. https://doi.org/10.1007/978-3-319-24574-4_28
 15. Tran, D. Q., Boulton, A., Lim, Y. J., & Morita, A. (2022). Improving segmentation accuracy by incorporating domain knowledge into neural networks. *Frontiers in Bioinformatics*, 2. <https://doi.org/10.3389/fbinf.2022.999700>
 16. Ulutaş, H. (2024). A deep learning-based U-Net 3+ technique for segmentation of blood cells. *Turkish Journal of Science and Technology*, 19(2), 485–495. <https://doi.org/10.55525/tjst.1404899>
 17. Yao, T., St. Clair, N., Miller, G. F., Dorfman, A. L., Fogel, M. A., Ghelani, S., Krishnamurthy, R., Lam, C. Z., Quail, M., Robinson, J. D., Schidlow, D., Slesnick, T. C., Weigand, J., Steeden, J. A., Rathod, R. H., & Muthurangu, V. (2024). A deep learning pipeline for assessing ventricular volumes from a cardiac MRI registry of patients with single ventricle physiology. *Radiology: Artificial Intelligence*, 6(1). <https://doi.org/10.1148/ryai.230132>
 18. Yousef, R., Khan, S., Gupta, G., Siddiqui, T., Albahlal, B. M., Alajlan, S. A., & Haq, M. A. (2023). U-Net-based models towards optimal MR brain image segmentation. *Diagnostics*, 13(9). <https://doi.org/10.3390/diagnostics13091624>
 19. Zhao, H., Shi, J., Qi, X., Wang, X., & Jia, J. (2017). Pyramid scene parsing network. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2881–2890. <https://doi.org/10.1109/CVPR.2017.308>
 20. Zhou, Z.-W., Siddiquee, M. M. R., Tajbakhsh, N., & Liang, J. M. (2018). UNet++: A nested U-Net architecture for medical image segmentation. In *Deep learning in medical image analysis and multimodal learning for clinical decision support*, 3–11. Springer. https://doi.org/10.1007/978-3-030-00889-5_1

V. R. Rybak¹, Yu. V. Shabatura²

¹ Ukrainian National Forestry University, Lviv, Ukraine

² Hetman Petro Sahaidachnyi National Army Academy, Lviv, Ukraine

EVALUATION OF THE U-NET 3+ MODEL EFFICIENCY IN SEGMENTING ANATOMICAL STRUCTURES ON DENTAL X-RAY IMAGES

Automation of digital image analysis processes is a crucial task in modern technology. It is especially relevant in the medical field, where it enhances the efficiency of research and reduces the impact of human error. This study investigates the application of the architecture of the U-Net 3+ model for the automated segmentation of digital dental radiographs. In the course of research, it has been established that traditional segmentation methods, such as thresholding, clustering, active contours, and region-growing techniques, show limited effectiveness when processing complex images. This is due to the intricacy of anatomical structures and presence of regions with similar contrast, which complicate the accurate delineation of object boundaries. The impact of applying deep learning architecture of the U-Net type and its modifications on the quality of complex dental structure segmentation has been evaluated. The study characterizes the efficiency of the U-Net 3+ model, which, through its full-scale skip connections, integrates information from different network levels and allows for preserving both global and local contexts. It has been found that the U-Net 3+ model demonstrates enhanced segmentation capabilities for segmentation of tooth structures in radiographic images compared to the classical U-Net architecture, due to its architectural improvements, achieving high-level values of Dice Coefficient and IoU. The accuracy of segmentation on digital images of small anatomical structures, such as tooth roots and areas with complex morphology, which often present challenges for traditional methods, has been assessed. The study also explored the possibility of continuous learning of the U-Net 3+ model on new data without losing accuracy on previous datasets, which is essential for its long-term use in dynamic environments. It has been revealed that the U-Net 3+ model demonstrates high resilience to overfitting and changes in input image quality, including noise and lighting variations, showing stable results on validation datasets. This confirms the model effectiveness for deployment and use in real-world conditions. Some prospects for applying the model to automated analysis of dental images are outlined, including integration into decision support systems, which will significantly improve the speed and accuracy of diagnosis.

Keywords: deep learning; image segmentation; computer vision; dental radiographs; diagnostic automation.