

Шифр «VEGETATION»

Інтелектуальна система діагностики інфекційного
ендокардиту на основі результатів обробки
ехокардіографії серця

АНОТАЦІЯ

Робота складається зі: вступу, 3 розділів, висновків, списку літературних джерел, який налічує 36 джерел. Загальний обсяг роботи – 20 сторінок (без врахування додатків та списку використаних джерел). Містить 2 таблиці, 11 рисунків.

Актуальність

Актуальність роботи зумовлена складністю аналізу ехокардіографії та виявлення аномалій клапанів, які призводить до помилково негативних діагнозів інфекційного ендокардиту (ІЕ). З огляду на високий рівень смертності (~30%) та важливості оперативного хірургічного втручання, розробка автоматичної та незалежної ідентифікації аномалій без людської похибки стала насущною. Аналіз сучасних робіт в цьому напрямку показав, що є прогалина в сфері комп'ютерного зору та ІЕ, так як більшість робіт фокусується або на загальній структурі серця, або на клінічних показниках ІЕ.

Мета роботи

Розробка методів та підходів, а також програмного забезпечення, для автоматичної інтелектуальної діагностики інфекційного ендокардиту на основі обробки зображень ехокардіографії серця людини.

Завдання

- 1) аналіз існуючих інтелектуальних методів аналізу ІЕ;
- 2) дослідження сучасних методів обробки зображень для задачі сегментації: згорткові нейронні мережі та візуальні трансформери;
- 3) побудова власної гібридної топології нейронної мережі для сегментації зображень ехокардіографії;
- 4) попередня обробка датасету вегетацій на ехокардіографії, який був наданий Інститутом серця МОЗ України;
- 5) навчання побудованої мережі та інших існуючих архітектур мереж на підготовленому датасеті;
- 6) аналіз отриманих результатів, порівняння з іншими архітектурами;
- 7) підбір програмної архітектури для реалізації програмного забезпечення;

8) розробка програмного забезпечення діагностики ІЕ на основі власних розроблених методів та мереж.

Використана методика

У роботі вперше були досліджені методи сегментації на основі глибоких згорткових нейронних мереж (CNN) та Vision Transformer (ViT) в рамках задачі сегментації вегетацій на знімках ехокардіографії з метою подальшого аналізу. Ці моделі використовують інноваційні підходи, такі як трансформерні методи та надглибокі нейронні мережі. Також були досліджені ансамблеві методи поєднання мереж, в тому числі запропоновано новий підхід нейронної мета-моделі з назвою *pointwise weighted channel aggregation*. Для побудови програмного забезпечення були використані методи розробки програмного забезпечення та підбору програмних архітектур (монолітна, серверлесс, клієнт-сервер).

Ключові слова

згорткові нейронні мережі, візуальні трансформери, семантична сегментація, ехокардіографія, інфекційний ендокардит, вегетації, інтелектуальна система

ЗМІСТ

РОЗДІЛ 1 ОГЛЯД ЛІТЕРАТУРНИХ ДЖЕРЕЛ	5
1.1 Особливості інфекційного ендокардиту	5
1.2 Ехокардіографія	6
1.3 Діагностика ІЕ	6
1.4 Актуальність дослідження	7
1.5 Огляд інтелектуальних систем обробки зображень ехокардіографії	8
РОЗДІЛ 2 ТЕОРЕТИЧНІ ДОСЛІДЖЕННЯ.....	11
2.1 Матеріали	11
2.2 Постановка задачі	13
2.3 Сучасні підходи. CNN та ViT.	14
2.4 Запропонований підхід	15
РОЗДІЛ 3 ПРАКТИЧНІ РЕЗУЛЬТАТИ ТА ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ ..	17
3.1 Baseline моделі для порівняння	17
3.2 Використані метрики	19
3.3 Особливості навчання.....	19
3.4 Результати запропонованого підходу	20
3.2 Аналіз результатів.....	20
3.6 Програмне забезпечення	21
Висновки	25
Використані джерела	26

РОЗДІЛ 1

ОГЛЯД ЛІТЕРАТУРНИХ ДЖЕРЕЛ

1.1 Особливості інфекційного ендокардиту

Ендокардит – захворювання серце-судинної системи яке виражається в запаленні ендокардіальної поверхні серця. В залежності від причини виникнення розрізняють інфекційний та неінфекційний. Відповідно до назви, під час інфекційного утворюється запалення в результаті потрапляння бактерій або ж грибків в кровообіг, що згодом осідають на клапанах серця або інших частинах ендокарда, в свою чергу неінфекційний може бути спричинений аутоімунними реакціями, тривалим запаленням, утворенням тромбів або іншими неінфекційними процесами.

Інфекційний ендокардит (ІЕ) зазвичай виникає по краях серцевих клапанів. Ураження, які називаються **вегетаціями**, являють собою маси, що складаються з фібрину, тромбоцитів і інфікуючих мікроорганізмів, які утримуються разом агглютинуючими антитілами, що виробляються бактеріями. Оскільки запалення продовжується, виразка може призвести до ерозії або перфорації стулок клапана, що призводить до недостатності клапана, пошкодження провідного шляху (якщо в ділянці перегородки) або розриву синуса Вальсальви (якщо в ділянці аорти) [1].

Інфекційний ендокардит є небезпечною для життя серцевою інфекцією і схильний до виникнення у деяких осіб із захворюваннями кількох серцевих клапанів, із щорічною частотою виявлення ~4.6 на 100 000 осіб з певною лінійною прогресією [2]. Зростання показників стійкості до антибіотиків, атипових мікробів і навіть грибків нині зачіпають значну частину всіх пацієнтів з ІЕ. Поєднання цих несприятливих факторів господаря, джерела та збудника призводить до непропорційно високого рівня смертності (~30%), зазвичай раптової та короткочасної, що контрастує з частотою виздоровлення при інших серцево-судинних захворюваннях [3]. Тобто **інфекційний ендокардит** представляє собою досить **розповсюджену** та надзвичайно **летальну** хворобу, яку надто часто виявляють лише під час аутопсії.

1.2 Ехокардіографія

Сучасна методологія виявлення ІЕ включає в себе модифіковані критерії Дюка, наріжним каменем серед яких є ехокардіографія — використання ультразвуку для дослідження серця. Це різновид медичної візуалізації, яка використовує стандартне ультразвукове або доплерівське ультразвукове дослідження [4]. Візуальне зображення, сформоване за допомогою цієї техніки, називається ехокардіограмою або ж серцевим ехо.

Ехокардіографія регулярно використовується для діагностики, лікування та спостереження за пацієнтами з будь-якими підозрюваними або відомими захворюваннями серця. Це один із найбільш використовуваних методів візуалізації в кардіології. Він може надати велику кількість корисної інформації, включаючи розмір і форму серця (кількісне визначення розміру внутрішньої камери), потужність насоса, розташування та ступінь будь-якого пошкодження тканини та оцінку клапанів. Знімок отриманий з ехокардіографа представляє собою покадровий запис функціонування серцевого м'яза в різних його фазах.

Покращення роздільної здатності датчика та розвиток кольорової доплерівської технології твердо встановили ехокардіографію як основний діагностичний метод у кардіології. Ехокардіографічна лабораторія часто є останньою діагностичною зупинкою пацієнта перед хірургічним втручанням, що вимагає максимально всеосяжного анатомічного та фізіологічного опису серцево-судинної системи та безпрецедентної деталізації ехокардіографічної оцінки. Запізнілі або неточні діагнози можуть призвести до несприятливого або ж летального результату [5].

1.3 Діагностика ІЕ

Ехокардіографія відіграє все більш важливу роль в оцінці та лікуванні ендокардиту. Розвиток двовимірної, а пізніше трансезофагеальної ехокардіографії значно покращив неінвазивне виявлення вегетацій. Крім того, ехо-доплерівські дослідження надають клінічно важливу інформацію про наявність і ступінь деструкції стулок та їх гемодинамічні наслідки, а також про наявність перивальвеолярної інфекції. Діагностична стратегія, запропонована

Дюраком і колегами [6] (критерії Дюка), поєднувала ехокардіографічні результати з клінічними та мікробіологічними даними. Основними критеріями діагностики ендокардиту вважали три ехокардіографічні знахідки: (1) наявність вегетацій, визначених як рухливі ехощільні маси, імплантовані в клапанний або муральний ендокард на шляху струменя регургітації або імплантовані в протезний матеріал без альтернативного анатомічного пояснення; (2) наявність абсцесів; або (3) наявність нового розшарування протеза клапана. Аномальні ехокардіографічні дані, які не відповідають цим визначенням, вважалися другорядними критеріями [7].

Модифікований критерій Дюка є нині рекомендованим діагностичним алгоритмом при підозрі на інфекційний ендокардит (ІЕ) [8, 9]. Розподіл на основні та другорядні критерії дозволяє легко застосовувати його в клінічній практиці, але може не оптимально використовувати індивідуальну інформацію про пацієнта. На противагу цьому, детальна статистична оцінка множинних характеристик з використанням штучного інтелекту покращила прогнозування різних серцево-судинних захворювань порівняно з традиційними клінічними стратегіями.

1.4 Актуальність дослідження

Попри надзвичайно велику компетентність лікарів прикрі похибки трапляються навіть з професіоналами. За даними дослідження [10] з журналу «Cardiovascular Ultrasound» частота помилково негативних перевірок знімків на наявність аномалій клапанів досягає 14.5%. Основними причинами виявились атипічне положення та малий розмір вегетацій. Діагностування подібних випадків потребує дійсно ветеранів своєї справи, нажаль не на всіх пацієнтів вони знайдуться. Автори роботи [11] зазначають, що хоча класична клінічна класифікація на синдроми гострого або підгострого ендокардиту не повністю втратила свою корисність, сучасні клінічні форми змінилися відповідно до глибоких епідеміологічних змін, що спостерігаються в розвинених країнах. Це вимагає тісної співпраці **мультидисциплінарної** команди, яка включає експертів з критичної допомоги пацієнтам з тяжким сепсисом або септичним

шоком, а також впровадження новітніх технологій та розробок в сфері ШІ. Звичайна людська похибка може зіграти свою роль під час будь-яких візуальних перевірок, тому необхідним сучасним рішенням для підтримки прийняття рішень є інтеграція моделей комп'ютерного зору, які здатні нівелювати людський фактор та підвищити якість діагностики в тандемі з лікарем-кардіологом.

Вчасна діагностика захворювання здатна значно скоротити смертність пацієнтів. За даними [12] загальна смертність становила 24%, 42%, 50% і 56% через 1, 5, 10 і 20 років відповідно. Мова йде не тільки про термінову, миттєву допомогу, якої потребує пацієнт доставлений з приступом, але й про діагностику людей в групі ризику в цілому. Розвиток інструментів ШІ аналізу ехокардіограм та точної діагностики захворювань здатен корінним чином змінити та упростити весь процес виявлення захворювань, що само по собі зменшить вартість необхідних процедур для пацієнтів. Це відкриває двері платної медицини для малозабезпечених верств населення, які найбільш страждають від відсутності вчасної діагностики. Попри здоровий сенс, за даними [13] не кожна літня людина здатна себе забезпечити обстеженнями. Дешева та доступна медицина здатна скоротити смертність населення та підвищити рівень життя в країні.

1.5 Огляд інтелектуальних систем обробки зображень ехокардіографії

Робота [14] с журналу Journal of Imaging дослідила роль штучного інтелекту в аналізі ехокардіографії. Автори підтверджують, що досягнення в галузі застосування ШІ в кардіології та ехокардіографії швидко розширюються, що призводить до революційних змін у лікуванні пацієнтів. Алгоритми ШІ допомагають у виявленні, класифікації, діагностиці та прогнозуванні серцевих патологій. Вони обіцяють підвищену ефективність робочого процесу, поліпшену відтворюваність і вищу точність діагностики, а також можуть стати економічно ефективним інструментом для задоволення зростаючого попиту на кардіологічну візуалізацію. Існує багато перешкод, які необхідно подолати, щоб дозволити використання ШІ в клінічній практиці, в тому числі недостатня кількість даних, що стосуються ШІ і клінічних результатів. Тому, підсумовують

автори, необхідні подальші дослідження для визначення їх точності та ефективності, а також того, як ШІ може впливати на клінічні результати.

Розробки останніх 3 років в сфері глибокого навчання та ехокардіографії були сфокусовані загалом на сегментації та ідентифікації серцевої структури глибокими пірамідальними нейронними мережами локальної уваги [15], глибокими мережами EchoNet [16], виявлення фаз серцевого ритму глибокими згортковими мережами [17], класифікація відео ехокардіографій розподіленими в часі згортковими мережами та двопотоковими мережами [18], створення 3D-моделей серця на основі сегментації згортковими мережами [19]. Оглядові роботи [20] також загалом концентрувались на загальних задачах сегментації та класифікації звичайними CNN внутрішньої частини серця.

В роботі [8] автори перевірили гіпотезу про те, що нейронні мережі та логістичний регресійний аналіз забезпечать краще прогнозування ІЄ порівняно з модифікованим критерієм Дюка. Результати на незалежній вибірці з 261 пацієнта показали, що логістичний регресійний аналіз і нейронні мережі забезпечують краще прогнозування ІЄ порівняно з клінічно встановленою модифікацією критерія Дюка. Цей результат підтвердив доцільність та потужність використання штучного інтелекту для діагностики ІЄ та необхідність в подальшій розробці та модифікації таких алгоритмів. Але особливістю та недоліком цієї роботи є використання лише клінічних показників пацієнтів, в той час як зображення ехокардіографії оцінював вручну кардіолог у лабораторії.

Через велику смертність (~30%) від ІЄ автори роботи [21] розробили надійну модель оцінку ризиків на основі машинного навчання для прогнозування ранньої смертності після операції з приводу ІЄ, що може допомогти у прийнятті клінічних рішень та покращити наслідки. Всього було включено 476 послідовних пацієнтів з ІЄ, які були прооперовані в 2 центрах. Тренувальна вибірка складалася з 276 пацієнтів. З 89 потенційних предикторів для навчання моделі XGBoost було відібрано вісім змінних, які стали вхідними даними для моделі прогнозування. Автори також включили в їх модель ознаки вегетацій, хоча не запропонували автоматичний їх розрахунок та сегментацію. Тим не менш, на

внутрішній тестовій когорті AUC становила 0,813 (95% ДІ, 0,670-0,933), а в зовнішній тестовій когорті AUC становила 0,812 (95% ДІ, 0,606-0,956). AUC авторів була значно вищою, ніж у інших моделях ансамблевого навчання, моделі логістичної регресії та Європейської системи оцінки кардіологічного операційного ризику II.

При імплементації інтелектуальної системи використовувались матеріали з роботи [22].

На даний час нами не було знайдено робіт, які б ставили на меті сегментацію та аналіз вегетації на зображеннях ехокардіографії для діагностики ІЕ, що підтверджує актуальність дослідження. Тому запропонована робота є унікальною і критично важливою для автоматичного розрахунку та ідентифікації вегетації для подальшого їх використання в оцінці ризиків та діагностиці.

РОЗДІЛ 2

ТЕОРЕТИЧНІ ДОСЛІДЖЕННЯ

2.1 Матеріали

Набір даних використаний для дослідження представляє собою 668 кадрів з кардіографа з зареєстрованою спеціалістами аномалією (вегетациєю та абсцесом) та 632 «чистих» кадри, загалом 1300 в парастернальній та апікальній проекціях. Зображення були вилучені з 20-ти ехокардіограм в форматі DICOM. Набір даних був наданий Інститутом серця МОЗ України.

Початковий формат зображень відповідно був 708 на 1016 пікселів з трьома каналами RGB, згодом оброблений до одноканального сірого зображення розміром 512 на 512 пікселів з зафіксованою цільовою областю.

Аугментація та попередня обробка не використовувались, з огляду на те, що знімки проводять під тими ж самими ракурсами та за однакових умов.

Приклад зображень з датасету наведено на Рис. 1 та Рис. 2.

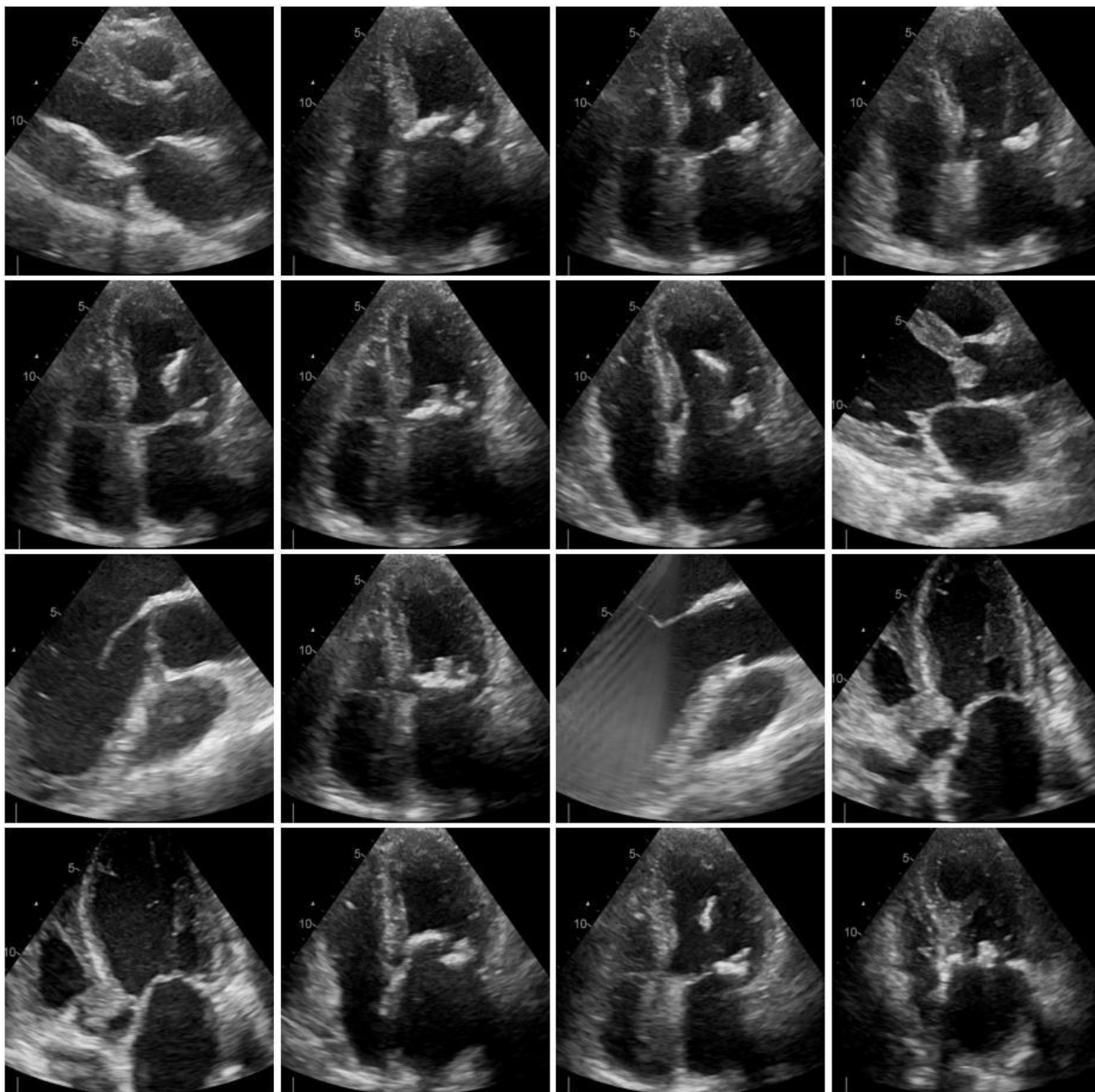


Рис. 1. Приклади знімків ехокардіографа

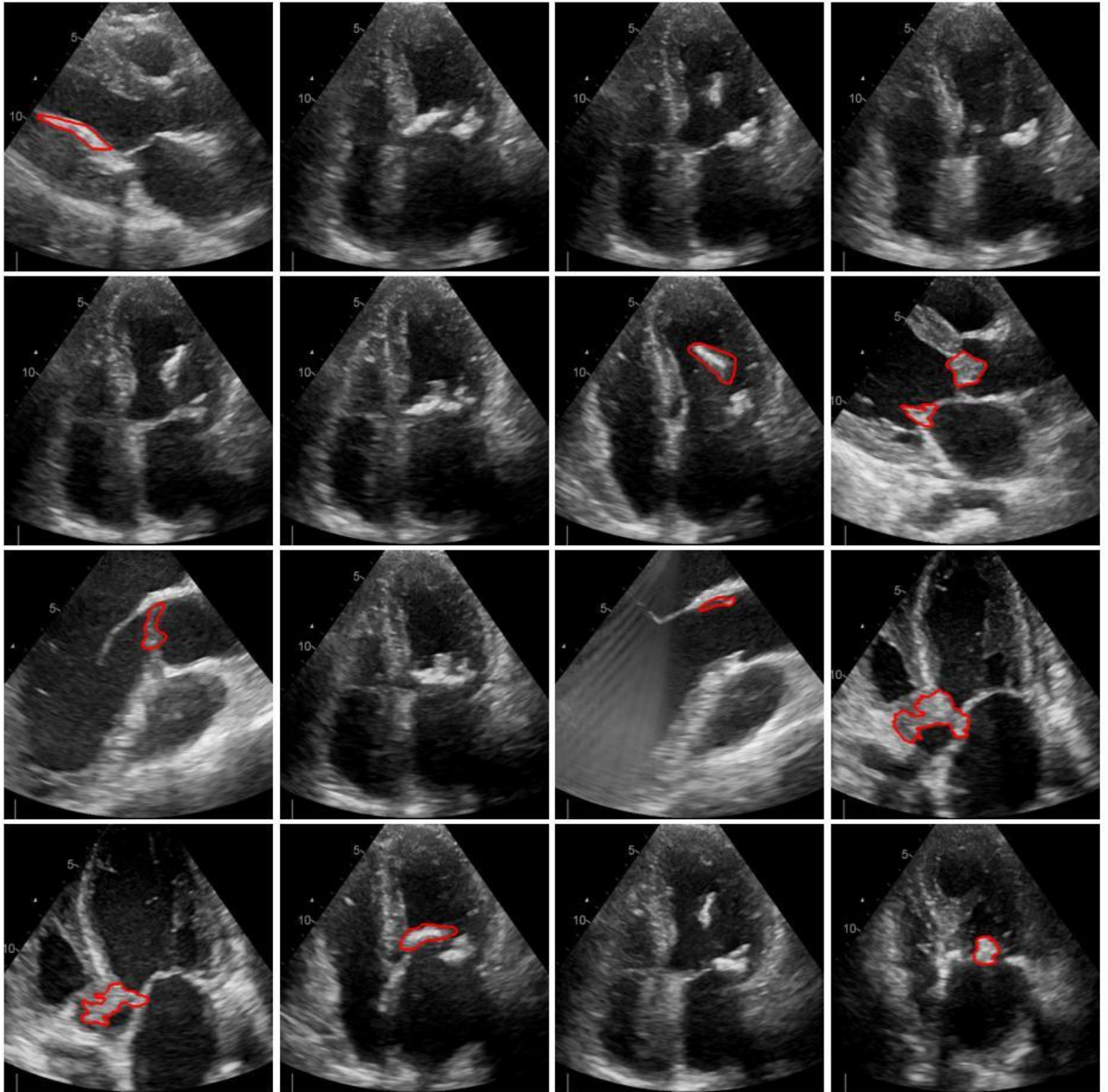


Рис. 2. Аномалії виявлені спеціалістами на знімках

2.2 Постановка задачі

Вирішення задачі діагностики ІЕ полягає у пошуку активних вогнищ (вегетацій) на зображеннях ехокардіографії серця пацієнтів. Основним завданням є семантична сегментація зображень на два класи: вегетація та фон.

Для вибірки $(X_1, Y_1), \dots, (X_l, Y_l)$, де $X_i \in \mathbb{R}^{n \times n}$ – це матриця зображення, а $Y_i \in \{0, 1\}^{n \times n}$ – це бінарна маска того ж розміру, що і оригінальне зображення, де елемент матриці 0 означає відсутність, а 1 присутність вегетації на даному

пікселі зображення, створити класифікатор $f(X)$, який коректно передбачає бінарну матрицю $Y \in \{0, 1\}^{n \times n}$ нової матриці зображення $X \in \mathbb{R}^{n \times n}$.

Посилаючись на отриманий результат можна оцінити наскільки значна множина аномальних пікселів та виставити прогноз щодо здоров'я тканин на знімку (Рис. 3).

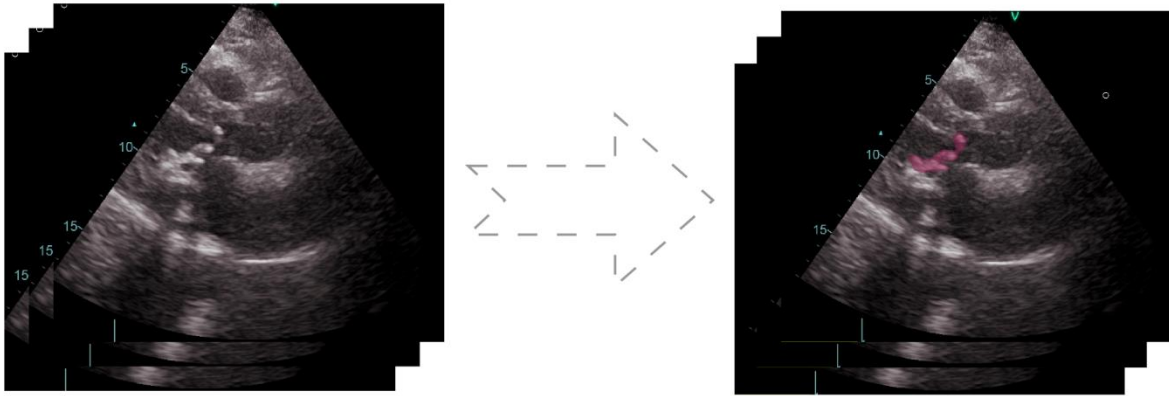


Рис. 3 Розпізнавання аномалій тканин ендокарда

2.3 Сучасні підходи. CNN та ViT.

Задачу такого типу в комп'ютерному зорі класифікують як семантичну сегментацію. На відміну від екземплярної сегментації, даний тип задач виділяється надзвичайно високою точністю і такими ж високими вимогами до обчислюваних ресурсів. Навіть найпростіші моделі потребують гігабайти пам'яті графічного процесора та для дослідження таких моделей потрібна певна інфраструктура, тому наразі системи аналізу візуальної інформації на основі штучного інтелекту в медицині зустрічаються досить рідко.

Сучасними класичними методами вирішення задачі семантичної сегментації є глибокі згорткові моделі архітектури encoder-decoder (наприклад, UNet [23], DeepLabV3+ [24]). Однак дослідження останніх двох років показали, що нові архітектури на основі трансформерів, а точніше візуальних трансформерів (Visual Transformers (ViT)), можуть досягти порівняно кращих результатів, аніж класичні згорткові моделі [25, 26], включаючи задачу семантичної сегментації [27].

2.4 Запропонований підхід

Серед нині існуючих масивів моделей сегментації, в даній роботі пропонується використання лише CNN-based, Transformer-based та їх гібридних методів з огляду на потребу в точності прогнозів [28]. Трансформерні та згорткові архітектурні конфігурації архітектур можуть бути занадто спеціалізованими на певних класах зображень та погано підхоплювати інші. Операція згортки підхоплює лише інформативні признаки з певного околу і в процесі глибокого навчання агрегує їх на високому рівні. Трансформери же на кожному блоці Multihead-Attention будують карту релевантності для всього масиву пікселів в відношенні один одного [29]:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V, \text{ де} \quad (1)$$

- Q - вектор запитів (query)
- K - вектор ключів (key)
- V - вектор значень (value)
- d_k корінь з розмірності вектора ключів (для нормалізації ваг уваги)

Це дозволяє моделям мати різний погляд на одне і те ж зображення, не враховуючи архітектурних особливостей окремих моделей.

Дієвий та ефективний шлях для комбінації розрахунків декількох моделей — використання ансамблів. В роботах [30, 31] ансамблеві методи проявили себе найкращим чином на задачі сегментації медичних зображень. Проте класичні ансамблі не завжди можуть покрити і справедливо врахувати результати моделей, тому було прийнято рішення використовувати ансамбль на основі мета-моделі, агрегації середнього та максимуму.

Загальна структура ансамбля представлена на рис. 4.

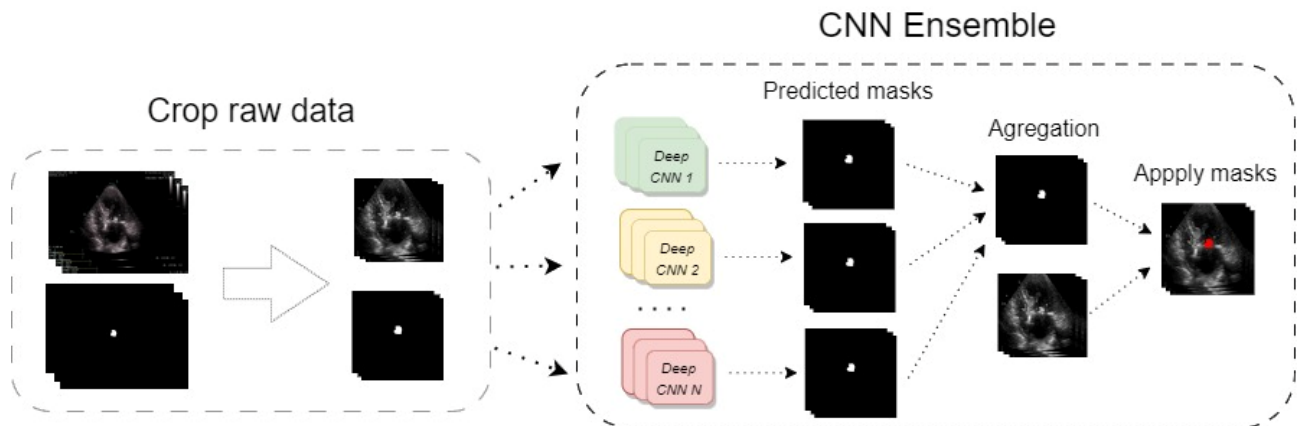


Рис. 4. Pipeline агрегації результатів ансамблевого методу

Запропонований підхід Pointwise weighted channel aggregation – як спосіб агрегації для мета-моделі використати модель машинного навчання з одним шаром, кожен вхід якої представляє собою сукупність значень від кожної окремої моделі-учасниці ансамблю. Тобто в даній конфігурації ми задаємо власні ваги для кожного пікселя виходу однієї з мереж. Результат отримуємо виконуючи динамічне попіксельне зважування результатів. На рис. 5 представлена структура подібної агрегації.

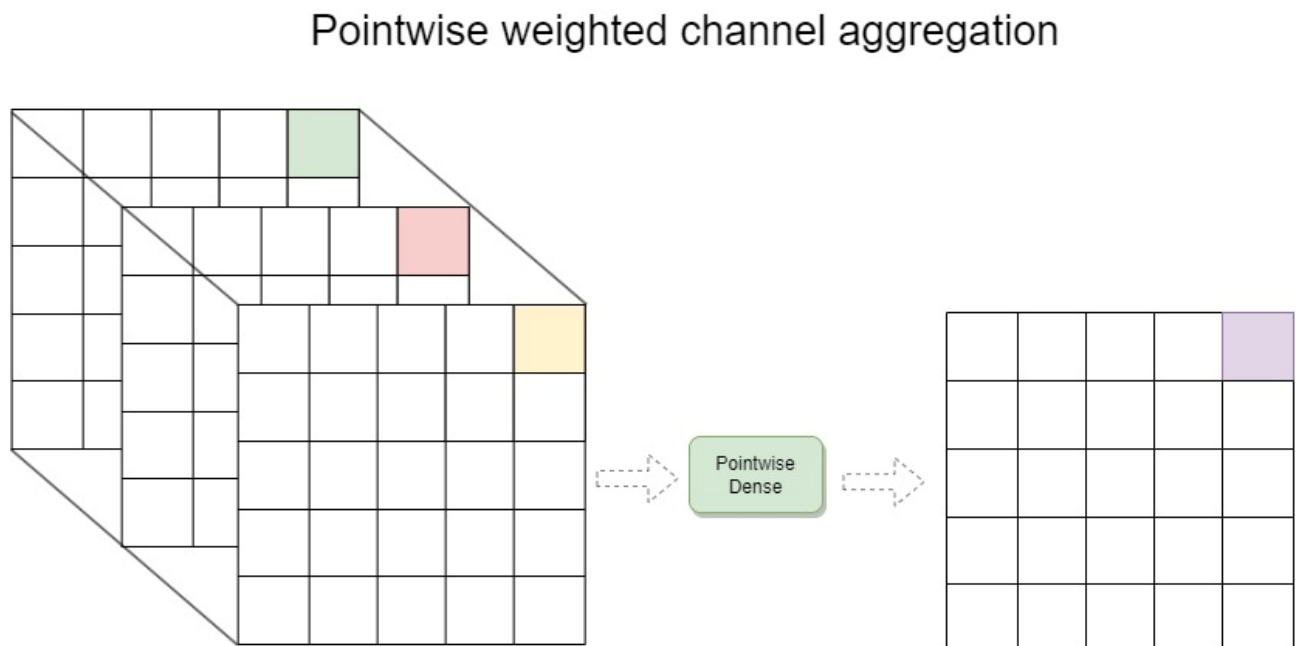


Рис. 5 Структура Pointwise-агрегації

РОЗДІЛ 3

ПРАКТИЧНІ РЕЗУЛЬТАТИ ТА ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ

3.1 Baseline моделі для порівняння

В якості моделей для порівняння результатів використовувались як моделі CNN, так і сучасні моделі трансформерів. Загалом було використано 6 baseline моделей:

- 1) **Unet++** — це архітектура на основі U-Net. Завдяки використанню густо з'єднаних вкладених підмереж декодера він покращує обробку виділених ознак і, як дослідили його автори [32], перевершує U-Net у завданнях медичної сегментації зображень електронної мікроскопії (ЕМ), клітин, ядер, пухлин головного мозку, печінки та легень. Архітектура складається зі звужувального шляху (він же шлях зменшення дискретизації, кодувальник), де ширина та висота карт функцій зменшуються, а канал розширюється в 2 рази, поки не досягне 1024 (як правило, максимальний рекомендований рівень для CNN), вузьке місце як «точка повороту» та шлях розширення (він же шлях підвищення дискретизації, декодер), де ширина та висота карт функцій розширюються до розміру маски.
- 2) **DeepLabv3+** — базується на DeepLabv3 та перевершує його наявністю простого, але ефективного модуля декодування [24]. Багаторазове зменшення дискретизації CNN призведе до зменшення роздільної здатності карти ознак, що призведе до зниження точності передбачення та втрати граничної інформації в семантичній сегментації. Подібним чином, агрегування контексту навколо функції допомагає краще її сегментувати, що досягається за допомогою атральних звивин. DeepLabv3+ допомагає вирішити ці проблеми.
- 3) В основі **SegTransformer** лежить трансформер, який зазвичай застосовується для завдань обробки послідовностей, таких як машинний переклад або моделювання тексту. У випадку сегментації, вхідні зображення розглядаються як послідовності пікселів, і блок самоуваги трансформера оцінює релевантність пікселів в відношенні один одного [33]. Для підвищення

ефективності обчислень SegTransformer використовує паралельну обробку зображень. Замість обробки кожного пікселя послідовно, він ділить зображення на блоки і обчислює взаємодії всередині кожного блоку та між блоками паралельно. Однією з ключових ідей є використання механізму уваги для моделювання взаємодії між об'єктами на зображенні. Це допомагає моделі визначати контекст об'єктів та їхніх взаємовідносин для точнішої сегментації.

- 4) **UNETR**, або UNet Transformer, — це заснована на трансформерах архітектура для сегментації медичних зображень, яка використовує чистий трансформатор як енкодер для вивчення представлень послідовності вхідного набору даних [34]. Кодер-трансформер підключається безпосередньо до декодера через skip-з'єднання різної розмірності, як U-Net, для обчислення остаточного виходу семантичної сегментації.
- 5) **MANET** — мережа механізму множинної уваги, основою якої є використання кількох механізмів уваги, що дозволяє мережі ефективніше обробляти важливі області зображення, покращуючи тим самим точність сегментації. Ці механізми можуть включати просторову увагу (яка концентрується на певних областях зображення) та каналну увагу (яка концентрується на певних каналах ознак) та увагу на рівні пікселів [35]. Попри те, що на момент написання немає робіт з використанням трансформерів в якості енкодера для мережі MANET, уже існують подібні реалізації. Тому подальшому в даній роботі подібна конфігурація буде іменуватись MANETR.
- 6) **PAN** — поєднує механізм уваги та просторову піраміду, для отримання точних та стислих характеристик для маркування пікселів замість складної розширеної згортки та штучно розроблених мереж декодера. Зокрема в статті [36] було представлено модуль Feature Pyramid Attention як реалізація структури просторової піраміди уваги на високорівневому виході та комбінування глобального пулінгу для отримання кращого представлення властивостей.

3.2 Використані метрики

Для тестування точності моделей бінарної сегментації використовувалися такі метрики:

- 1) Mean intersection over union (mIoU)
- 2) F1 (Dice) score
- 3) F2 score
- 4) Precision
- 5) Recall

Формула для IoU:

$$IoU = \frac{TP}{TP + FP + FN}$$

Precision і Recall були обраховані наступним чином:

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

F1 score був використаний в наступній формі з $\beta = 1$:

$$F1 = \frac{(1 + \beta^2) \cdot TP + \epsilon}{(1 + \beta^2) \cdot TP + \beta^2 \cdot FN + FP + \epsilon'}$$

де TP – кількість істинно-позитивних міток (пікселів у випадку сегментації), FN – кількість хибнонегативних міток, FP – кількість хибнопозитивних міток, ϵ – дуже мала константа, щоб уникнути ділення на нуль, β – коефіцієнт важливості Recall над Precision. Для задачі бінарної сегментації TP, FN, FP обчислювалися попіксельно.

Відповідно F2 score обчислювався для коефіцієнта $\beta = 2$.

3.3 Особливості навчання

Початкова вибірка з 1300 зображень була рівномірно розділена на три вибірки: тренувальна (800 зображень, ~61.5%), валідаційна (200 зображень, ~15.5%), тестова (300 зображень, ~23%).

Для навчання моделей використовувалась графічна карта RTX 4090. Навчання проводилось батчами по 16 зображень з оптимізатором Adam. Як

функція втрат використовувався Dice loss. Зменшення learning rate проводилось при відсутності зміни в validation loss. Навчання проводилось до 200 епох при зупинці, якщо validation loss не падає більш ніж 10 епох.

3.4 Результати запропонованого підходу

В таблиці 1 представлені результати запропонованої архітектури у порівнянні з іншими підходами на тестовій вибірці.

Таблиця 1. Результати підходів на тестовій вибірці

Підхід	Енкодер	IoU	F1 (Dice) score	F2 score	Precision	Recall	Кількість параметрів архітектурної конфігурації
UNet++	ResNet34	0.6043	0.7532	0.8153	0.8632	0.6691	26.1 М
DeepLabV3+	ResNet50	0.7107	0.8308	0.8504	0.864	0.8003	26.7 М
SegTransformer	mit_b2	0.5718	0.7274	0.7696	0.8015	0.6686	27.3 М
UNETR	mit_b2	0.6093	0.7569	0.8018	0.835	0.6928	27.5 М
MANETR	mit_b2	0.5984	0.7487	0.79	0.8207	0.6897	35.0 М
PAN	ResNet50	0.6868	0.8143	0.8423	0.8625	0.7724	24.3 М

В таблиці 2 представлені результати отримані ансамблями мереж на тестовій вибірці з використанням моделей DeepLabV3+ та PAN.

Таблиця 2. Результати агрегації моделей в ансамблі

Тип агрегації	IoU	F1 (Dice) score	F2 score	Precision	Recall
Average	0.6975	0.8218	0.7961	0.8688	0.7796
Maximum	0.6961	0.8207	0.8443	0.7844	0.8607
Мета-модель pointwise weighted channel aggregation	0.7822	0.886	0.8891	0.8912	0.8810

3.2 Аналіз результатів

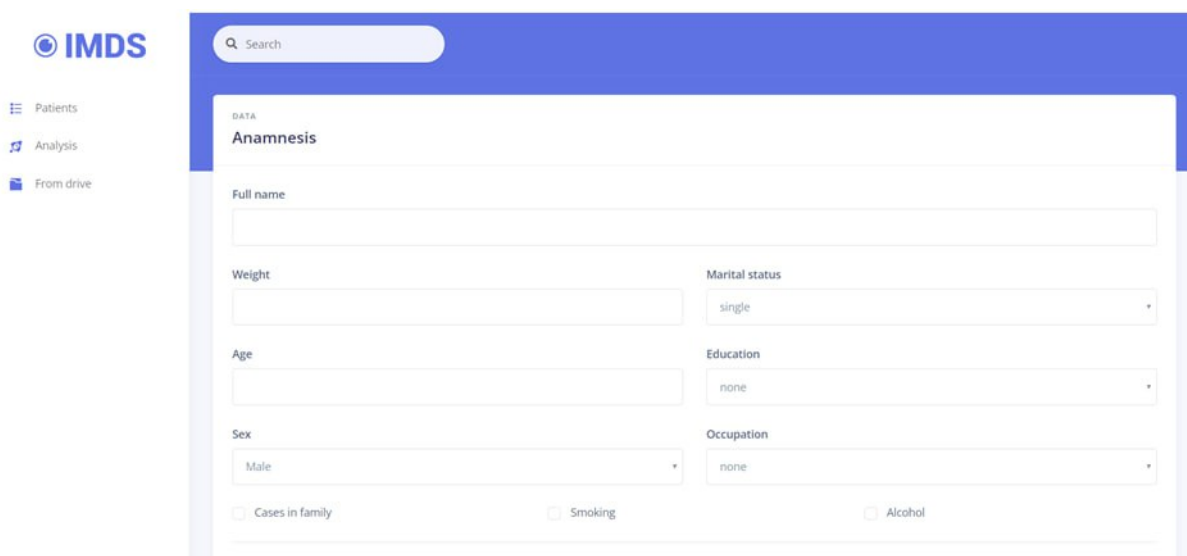
Результати в таблиці 1 та 2 показали суттєву домінацію класичних методів сегментації на основі згорткових нейронних мереж понад гібридними мережами, які включають в себе трансформерну архітектуру.

Можна зробити висновок, що трансформерні підходи потребують серйозного допрацювання та розвитку для їх поширення на сферу охорони здоров'я. Більшість медичних знімків вимагають чітку увагу мережі до текстурних особливостей вхідних даних, а з огляду на дослідження, трансформерні конфігурації моделей до подібного слабо адаптовані.

Ансамблеві агрегації в даній роботі суттєво не вплинули на результат та показали спірну ефективність. Але запропонована нами мета-модель показала великий буст результатів та її релевантність.

3.6 Програмне забезпечення

На Рис. 6-11 наведені скріншоти реалізованого програмного забезпечення запропонованої системи. Були реалізовані монолітна та клієнт-серверна архітектура с serverless-backend на платформі AWS.



The screenshot displays the user interface of the IMDS (Intelligent Medical Decision Support) system. On the left, there is a navigation menu with options: 'Patients', 'Analysis', and 'From drive'. The main content area is titled 'DATA Anamnesis' and contains a form for entering patient information. The form includes the following fields:

- Full name: A text input field.
- Weight: A text input field.
- Age: A text input field.
- Sex: A dropdown menu with 'Male' selected.
- Marital status: A dropdown menu with 'single' selected.
- Education: A dropdown menu with 'none' selected.
- Occupation: A dropdown menu with 'none' selected.
- Three checkboxes at the bottom: 'Cases in family', 'Smoking', and 'Alcohol', all of which are currently unchecked.

Рис. 6. Верхня частина форми внесення загальних даних про пацієнта

Search

DATA

Blood tests

Total protein	White blood cells	Platelet distribution width
Alkaline phosphatase	Granulocytes	Erythrocyte sedimentation rate
Aspartate aminotransferase	Lymphocytes	Acidity
Gamma glutamyl transpeptidase	Monocytes	Prothrombin ratio
Alanine aminotransferase	Red blood cells	Plasma recalcification time
Glucose	Hemoglobin	Fibrinogen

Рис. 7. Форма вводу параметрів аналізу крові

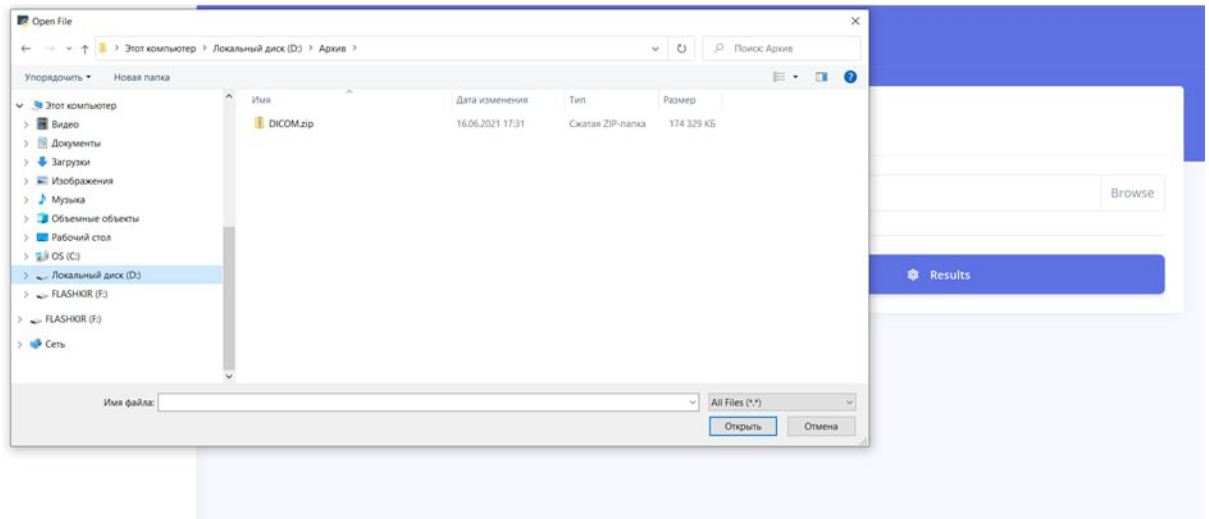


Рис. 8. Файловий діалог вибору файла архіву сканів ехокардіографії

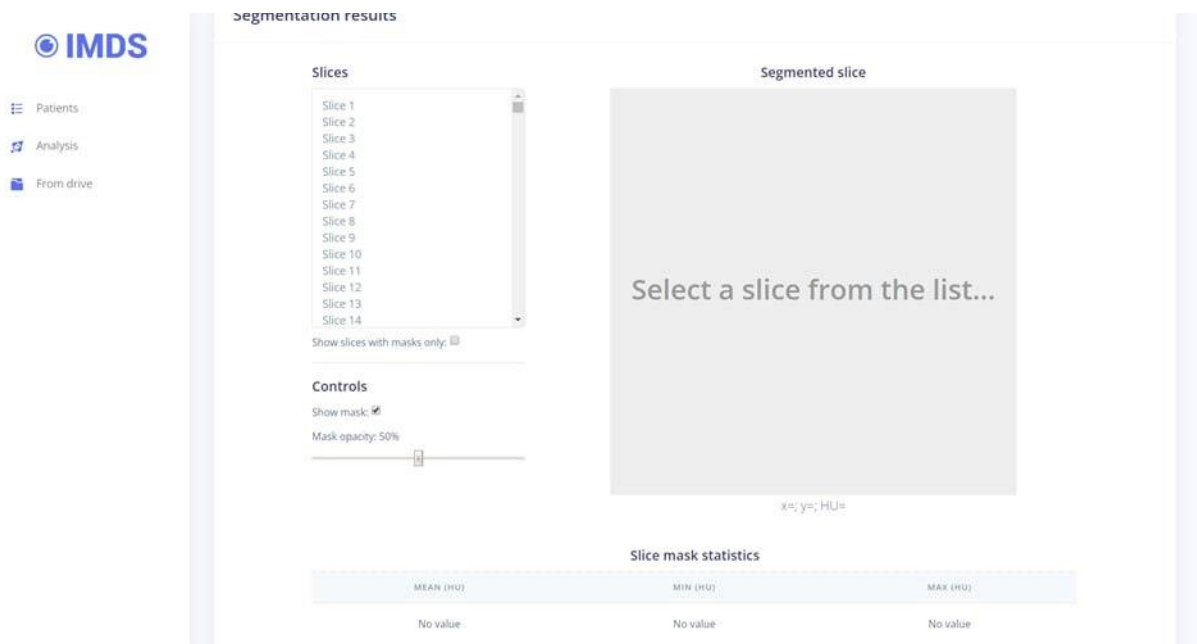


Рис. 9. Форма результатів Segmentation results

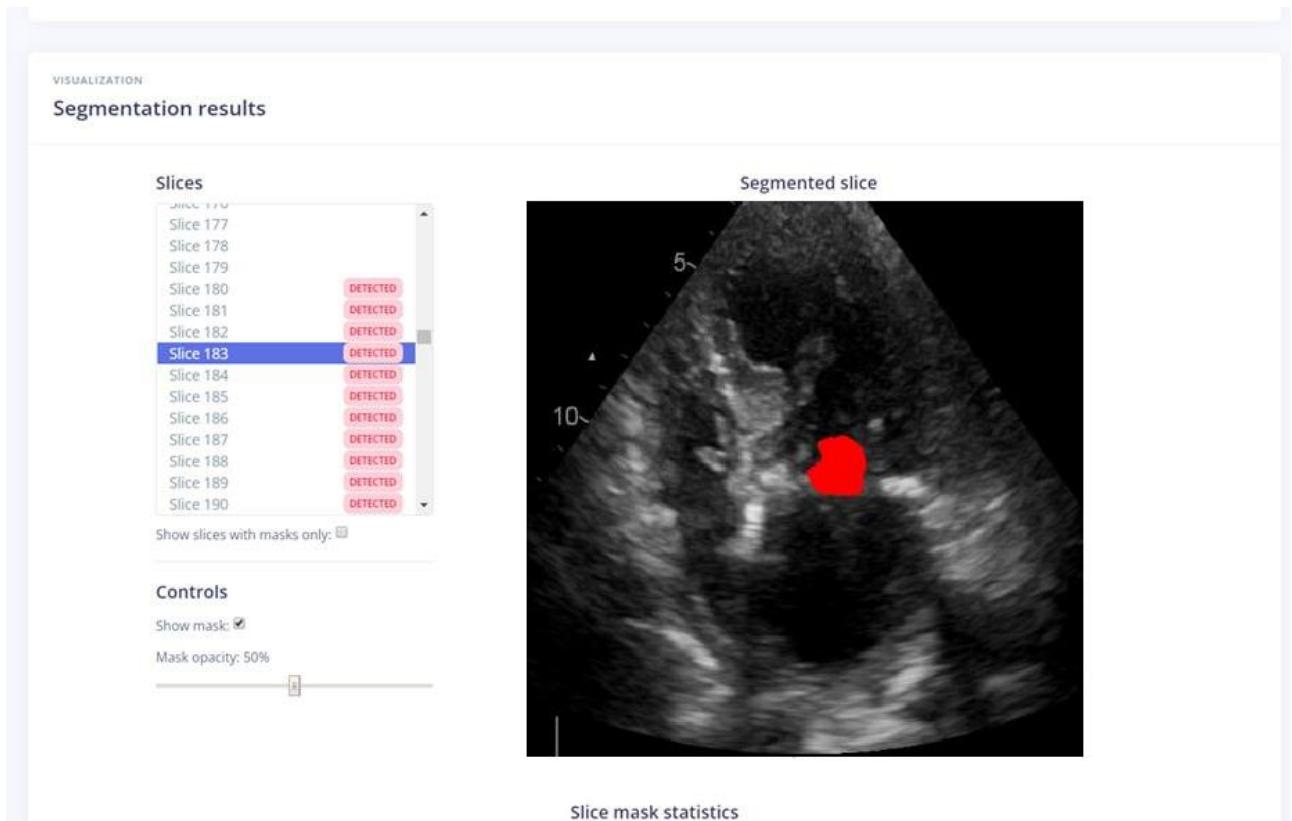


Рис. 10. Форма з результатами сегментації сканів серця

The screenshot displays the IMDS web application interface. On the left, there is a navigation menu with three items: 'Patients', 'Analysis', and 'From drive'. The main content area features a search bar at the top. Below it, the 'OUTPUT' section is titled 'Diagnostics results' and contains a table with patient information. The table has two columns: 'FIELD' and 'VALUE'. The data rows are as follows:

FIELD	VALUE
Patient ID	3
Full name	Іванов Іван Іванович
Age	22
Weight	56.89
Sex	male
Doctor Check	-

Below the table, the 'VISUALIZATION' section is titled 'Segmentation results'.

Рис. 11. Форма результатів Diagnostics results

ВИСНОВКИ

В даній роботі була розроблена інтелектуальна система діагностики інфекційного ендокардиту. Проблема цього захворювання полягає в складності аналізу ехокардіографії людиною, що призводить до помилково негативних перевірок знімків на наявність аномалій клапанів, яка досягає 14.5%. Основними причинами складності є атипичне положення та малий розмір вегетацій. У поєднанні з рівнем смертності в ~25-30% та критичної залежності від оперативного хірургічного втручання необхідність в автоматичній та незалежній ідентифікації аномалій та їх оцінці зростає. Аналіз сучасних наукових робіт показав, що наразі проводяться медичні дослідження для діагностики ІЕ, але майже немає мультидисциплінарних робіт використання штучного інтелекту та комп'ютерного зору для діагностики цього захворювання, роблячи цю сферу відкритою для досліджень.

Нами були дослідженні гібридні моделі нейронних мереж на основі трансформерів для сегментації вегетацій на знімках ехокардіографії для їх подальшого аналізу. Для випробуваних методів була досягнена наступна точність в сегментації iou 0.7822, dice score 0.886.

Аналіз результатів показав значну перевагу у точності моделі DeepLabV3+ у порівнянні з іншими моделями CNN та ViT. Інші конфігурації моделей хоча і відстають в даному дослідженні, також мають значний потенціал та потребують ретельних досліджень. Запропонований нами метод агрегації pointwise weighted channel aggregation показав найкращі результати щодо ансамблювання.

Отримані результати свідчать про великий потенціал розробленої діагностичної системи. Хоча система потребує допрацювань, її результати уже були впроваджені в компанії ONVY HealthTech Group GmbH та плануються впровадити в Інститут серця МОЗ України. Імплементация цієї системи на наявних клінічних ресурсах здатна покращити ефективність та точність виявлення ІЕ на різних стадіях захворювання, сприяючи швидшому та більш точному лікуванню пацієнтів.

ВИКОРИСТАНІ ДЖЕРЕЛА

- [1] Cardiology Explained. Ashley EA, Niebauer J. London: Remedica; 2004. <https://www.ncbi.nlm.nih.gov/books/NBK2204/>
- [2] Cresti A, Chiavarelli M, Scalese M, Nencioni C, Valentini S, Guerrini F, D'Aiello I, Picchi A, De Sensi F, Habib G. Epidemiological and mortality trends in infective endocarditis, a 17-year population-based prospective study. *Cardiovasc Diagn Ther.* 2017 Feb;7(1):27-35. doi: 10.21037/cdt.2016.08.09. PMID: 28164010; PMCID: PMC5253443.
- [3] Richard Rezar, Michael Lichtenauer, Markus Haar, Georg Hödl, Jan Marco Kern, Zhichao Zhou, Thomas Wuppinger, Johannes Kraus, Bernhard Strohmer, Uta C. Hoppe, Bernhard Wernly, Infective endocarditis – A review of current therapy and future challenges, *Hellenic Journal of Cardiology*, Volume 62, Issue 3, 2021, Pages 190-200, ISSN 1109-9666, <https://doi.org/10.1016/j.hjc.2020.10.007>.
- [4] Adams, D., Forsberg, E. (2009). Conducting a Cardiac Ultrasound Examination. In: Nihoyannopoulos, P., Kisslo, J. (eds) *Echocardiography*. Springer, London. https://doi.org/10.1007/978-1-84882-293-1_2
- [5] Moradian M. Diagnostic errors in echocardiography: review of five interesting pediatric cases. *J Tehran Heart Cent.* 2012 Winter;7(1):33-6. Epub 2012 Feb 28. PMID: 23074633; PMCID: PMC3466877. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3466877/>
- [6] Durack DT, Lukes AS, Bright DK. New criteria for diagnosis of infective endocarditis: utilization of specific echocardiographic findings. Duke Endocarditis Service. *Am J Med.* 1994 Mar;96(3):200-9. doi: 10.1016/0002-9343(94)90143-0. PMID: 8154507.
- [7] Evangelista A, Gonzalez-Alujas MT. Echocardiography in infective endocarditis. *Heart.* 2004 Jun;90(6):614-7. doi: 10.1136/hrt.2003.029868. PMID: 15145856; PMCID: PMC1768290.
- [8] L Vogel, I Dykun, M Totzeck, T Rassaf, A Mahabadi, Neuronal nets and logistic regression analysis provide improved prediction of infective endocarditis compared to

the modified Duke Score: a post-hoc analysis of the prospective PRO-ENDOCARDITIS study, *European Heart Journal*, Volume 43, Issue Supplement_2, October 2022, ehac544.2776, <https://doi.org/10.1093/eurheartj/ehac544.2776>

[9] Li JS, Sexton DJ, Mick N, Nettles R, Fowler VG Jr, Ryan T, Bashore T, Corey GR. Proposed modifications to the Duke criteria for the diagnosis of infective endocarditis. *Clin Infect Dis*. 2000 Apr;30(4):633-8. doi: 10.1086/313753. Epub 2000 Apr 3. PMID: 10770721.

[10] Ren, Z., Zhang, J., Chen, H. *et al*. Preoperative false-negative transthoracic echocardiographic results in native valve infective endocarditis patients: a retrospective study from 2001 to 2018. *Cardiovasc Ultrasound* **19**, 2 (2021). <https://doi.org/10.1186/s12947-020-00229-8>

[11] Cuervo G, Escrhuella-Vidal F, Gudiol C, Carratalà J. Current Challenges in the Management of Infective Endocarditis. *Front Med (Lausanne)*. 2021 Feb 22;8:641243. doi: 10.3389/fmed.2021.641243. PMID: 33693021; PMCID: PMC7937698.

[12] Netzer RO, Altwegg SC, Zollinger E, Täuber M, Carrel T, Seiler C. Infective endocarditis: determinants of long term outcome. *Heart*. 2002 Jul;88(1):61-6. doi: 10.1136/heart.88.1.61. PMID: 12067947; PMCID: PMC1767177.

[13] Nishi T, Babazono A, Maeda T. Association between income levels and irregular physician visits after a health checkup, and its consequent effect on glycemic control among employees: A retrospective propensity score-matched cohort study. *J Diabetes Investig*. 2019 Sep;10(5):1372-1381. doi: 10.1111/jdi.13025. Epub 2019 Mar 7. PMID: 30758145; PMCID: PMC6717811.

[14] Barry, T.; Farina, J.M.; Chao, C.-J.; Ayoub, C.; Jeong, J.; Patel, B.N.; Banerjee, I.; Arsanjani, R. The Role of Artificial Intelligence in Echocardiography. *J. Imaging* 2023, 9, 50. <https://doi.org/10.3390/jimaging9020050>

[15] Liu F, Wang K, Liu D, Yang X, Tian J. Deep pyramid local attention neural network for cardiac structure segmentation in two-dimensional echocardiography. *Med Image Anal*. 2021 Jan;67:101873. doi: 10.1016/j.media.2020.101873. Epub 2020 Oct 18. PMID: 33129143.

- [16] Ghorbani, A., Ouyang, D., Abid, A. et al. Deep learning interpretation of echocardiograms. *npj Digit. Med.* 3, 10 (2020). <https://doi.org/10.1038/s41746-019-0216-8>
- [17] Farhad, M., Masud, M.M., Beg, A. et al. Cardiac phase detection in echocardiography using convolutional neural networks. *Sci Rep* 13, 8908 (2023). <https://doi.org/10.1038/s41598-023-36047-x>
- [18] Howard JP, Tan J, Shun-Shin MJ, Mahdi D, Nowbar AN, Arnold AD, Ahmad Y, McCartney P, Zolgharni M, Linton NWF, Sutaria N, Rana B, Mayet J, Rueckert D, Cole GD, Francis DP. Improving ultrasound video classification: an evaluation of novel deep learning methods in echocardiography. *J Med Artif Intell.* 2020 Mar 25;3:4. doi: 10.21037/jmai.2019.10.03. PMID: 32226937; PMCID: PMC7100611.
- [19] T Shirakawa, Y Koyama, R Shibata, S Fukui, M Tatsuoka, M Yoshitatsu, K Toda, I Fukuda, Y Sawa, Automated heart segmentation using a convolutional neural network accelerates 3D model creation for cardiac surgery, *European Heart Journal - Cardiovascular Imaging*, Volume 22, Issue Supplement_1, January 2021, jeaa356.353, <https://doi.org/10.1093/ehjci/jeaa356.353>
- [20] Krittanawong C, Omar AMS, Narula S, Sengupta PP, Glicksberg BS, Narula J, Argulian E. Deep Learning for Echocardiography: Introduction for Clinicians and Future Vision: State-of-the-Art Review. *Life (Basel).* 2023 Apr 17;13(4):1029. doi: 10.3390/life13041029. PMID: 37109558; PMCID: PMC10145844.
- [21] Luo L, Huang SQ, Liu C, Liu Q, Dong S, Yue Y, Liu KZ, Huang L, Wang SJ, Li HY, Zheng S, Wu ZK. Machine Learning-Based Risk Model for Predicting Early Mortality After Surgery for Infective Endocarditis. *J Am Heart Assoc.* 2022 Jun 7;11(11):e025433. doi: 10.1161/JAHA.122.025433. Epub 2022 Jun 3. PMID: 35656984; PMCID: PMC9238722.
- [22] Recommender Systems Based on Reinforced Learning
article in *Electronic and Control Systems №2 (76) 2023* Victor Sineglazov
Andrii Sheruda <https://doi.org/10.18372/1990-5548.76.17668>

- [23] Ronneberger, O., Fischer, P., Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab, N., Hornegger, J., Wells, W., Frangi, A. (eds) Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. MICCAI 2015. Lecture Notes in Computer Science(), vol 9351. Springer, Cham. https://doi.org/10.1007/978-3-319-24574-4_28
- [24] Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation Liang-Chieh Chen, Yukun Zhu, George Papandreou, Florian Schroff, and Hartwig Adam
<https://doi.org/10.48550/arXiv.1802.02611>
- [25] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, Jakob Uszkoreit, & Neil Houlsby (2021). An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. In International Conference on Learning Representations.
<https://doi.org/10.48550/arXiv.2010.11929>
- [26] Wu, B., Xu, C., Dai, X., Wan, A., Zhang, P., Yan, Z., Tomizuka, M., Gonzalez, J., Keutzer, K., & Vajda, P. (2020). Visual Transformers: Token-based Image Representation and Processing for Computer Vision. arXiv e-prints, arXiv:2006.03677.
<https://doi.org/10.48550/arXiv.2006.03677>
- [27] S. Zheng et al., "Rethinking Semantic Segmentation from a Sequence-to-Sequence Perspective with Transformers," 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, TN, USA, 2021, pp. 6877-6886, doi: 10.1109/CVPR46437.2021.00681.
- [28] I. Ahmed, M. Ahmad, F. A. Khan and M. Asif, "Comparison of Deep-Learning-Based Segmentation Models: Using Top View Person Images," in IEEE Access, vol. 8, pp. 136361-136373, 2020, doi: 10.1109/ACCESS.2020.3011406.
- [29] Vaswani, Ashish, Noam M. Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser and Illia Polosukhin. "Attention is All you Need." Neural Information Processing Systems (2017).

<https://doi.org/10.48550/arXiv.1706.03762>

[30] DivergentNets: Medical Image Segmentation by Network Ensemble Vajira Thambawitaa Steven A. Hicksa Pål Halvorsena,b and Michael A. Rieglera
<https://doi.org/10.48550/arXiv.2107.00283>

[31] An Analysis on Ensemble Learning optimized Medical Image Classification with Deep Convolutional Neural Networks Dominik Müller, Iñaki Soto-Rey and Frank Kramer
<https://doi.org/10.48550/arXiv.2201.11440>

[32] UNet++: A Nested U-Net Architecture for Medical Image Segmentation Zongwei Zhou, Md Mahfuzur Rahman Siddiquee, Nima Tajbakhsh, and Jianming Liang doi:
[10.1007/978-3-030-00889-5_1](https://doi.org/10.1007/978-3-030-00889-5_1)

[33] SegFormer: Simple and Efficient Design for Semantic Segmentation with Transformers Enze Xie, Wenhai Wang, Zhiding Yu, Anima Anandkumar, Jose M. Alvarez, Ping Luo
<https://doi.org/10.48550/arXiv.2105.15203>

[34] UNETR: Transformers for 3D Medical Image Segmentation Ali Hatamizadeh Yucheng Tang Vanderbilt University Vishwesh Nath Dong Yang
<https://doi.org/10.48550/arXiv.2103.10504>

[35] Multi-Attention-Network for Semantic Segmentation of Fine-Resolution Remote Sensing Images Rui Li, Shunyi Zheng, Ce Zhang, Chenxi Duan, Jianlin Su, Libo Wang, and Peter M Atkinson DOI:[10.1109/TGRS.2021.3093977](https://doi.org/10.1109/TGRS.2021.3093977)

[36] G. Gendy, H. Mohammed, N. Sabor and G. He, "A Deep Pyramid Attention Network for Single Image Super-resolution," 2021 9th International Japan-Africa Conference on Electronics, Communications, and Computations (JAC-ECC), Alexandria, Egypt, 2021, pp. 14-19, doi: 10.1109/JAC-ECC54461.2021.9691443