

УДК 519.6:004.93'1:616-073.763

DOI [https://doi.org/10.24144/2616-7700.2026.48\(1\).206-217](https://doi.org/10.24144/2616-7700.2026.48(1).206-217)**В. М. Самусь**

ДВНЗ «Ужгородський національний університет»,
аспірант кафедри системного аналізу та теорії оптимізації
vasyl.samus@uzhnu.edu.ua
ORCID: <https://orcid.org/0000-0002-1682-689X>

ГІБРИДНА МОДЕЛЬ АНАЛІЗУ СИГНАЛУ ЕКГ НА ОСНОВІ ІНЖЕНЕРНИХ ОЗНАК ТА ГЛИБИННОГО НАВЧАННЯ

У статті представлено гібридну модель аналізу сигналу електрокардіограми (ЕКГ), що поєднує інженерні ознаки та глибинне навчання. Запропоновано архітектуру з адаптивним злиттям ознак і сирого сигналу за допомогою Clinical Gating Unit. Розглянуто математичні моделі побудови ЕКГ, основні часові, морфологічні, статистичні та геометричні ознаки. Наведено приклад реалізації архітектури з використанням PyTorch. Обґрунтовано переваги комбінованого підходу щодо точності, узагальнюваності та інтерпретованості результатів.

Ключові слова: математична модель, ЕКГ, HRV, нейромережа, глибинне навчання, CNN, LSTM, інженерні ознаки, Clinical Gating Unit.

1. Вступ. Електрокардіограма (ЕКГ) є одним із ключових інструментів діагностики серцево-судинних захворювань. Вона відображає електричну активність серця у вигляді сигналу, що містить інформацію як про ритм, так і про морфологію скорочень міокарда. Завдяки своїй неінвазивності, доступності та високій інформативності, ЕКГ широко використовується у клінічній практиці, моніторингу та скринінгу пацієнтів. Аналіз ЕКГ може здійснюватися як візуально (лікарем), так і автоматизованими алгоритмами. Автоматизація аналізу дозволяє зменшити вплив людського фактору, прискорити обробку даних та підвищити точність виявлення патологій. Історично методи обробки сигналу ЕКГ ґрунтувалися на виділенні ключових ознак — тривалості інтервалів, амплітуд хвиль, статистик варіабельності серцевого ритму (HRV). Водночас ці ознаки є агрегованими й можуть не повністю відображати морфологічні особливості сигналу. З розвитком глибинного навчання виникли моделі, здатні безпосередньо аналізувати сирій ЕКГ-сигнал та виявляти приховані патерни. Зокрема, згорткові нейронні мережі (CNN) ефективно розпізнають морфологію хвиль, а рекурентні моделі (RNN, LSTM, GRU) — часові залежності. Проте й ці підходи мають обмеження: вони можуть бути чутливими до шуму, артефактів та погано узагальнюватися без залучення доменної інформації. Сучасною тенденцією є поєднання обох підходів — інженерних ознак та глибинного представлення — у гібридні моделі. Такі архітектури, зокрема двовежові (dual-tower), дозволяють одночасно враховувати як клінічно інтерпретовані ознаки, так і сирі дані. Додаткове вдосконалення можливе через адаптивне злиття (fusion) з урахуванням клінічного контексту, зокрема із застосуванням модулів типу *Clinical Gating Unit*.

Постановка проблеми Сигнали електрокардіограми є складними біомедичними часовими рядами, які відображають електричну активність серця. Їх аналіз дозволяє виявляти порушення ритму, ішемічні зміни, електролітні зрушення та інші патології. Незважаючи на активний розвиток методів машинного

навчання, проблема надійного, інтерпретованого та стійкого до шуму автоматичного аналізу ЕКГ залишається відкритою. Традиційні підходи базуються на виділенні експертних (інженерних) ознак, таких як інтервали між зубцями, HRV-показники, морфологічні характеристики. Вони добре інтерпретуються та узагальнюються, однак мають обмежену здатність враховувати складні морфологічні патерни. З іншого боку, глибинні моделі (CNN, LSTM) здатні виявляти складні особливості без явного виокремлення ознак, але страждають від слабкої пояснюваності та чутливості до зашумлених або неповних даних.

Таким чином, виникає потреба у створенні моделі, яка:

- поєднує переваги експертних ознак та глибинного навчання;
- забезпечує адаптивне злиття різних джерел інформації залежно від клінічного контексту;
- зберігає інтерпретованість результатів та є стійкою до шуму й артефактів.

Для реалізації цього підходу у роботі пропонується архітектура з двома обчислювальними вежами (feature tower та signal tower), що зливаються за допомогою модуля *Clinical Gating Unit*, який керує вагуванням представлень на основі клінічних маркерів (наприклад, SDNN, QTc, ST).

2. Аналіз останніх досліджень і публікацій. Аналіз електрокардіограми традиційно базувався на морфологічних ознаках, що відображають часові, амплітудні та просторові характеристики електричної активності серця. З появою цифрових обчислень у 1970-х роках з'явилися перші алгоритми автоматизованого аналізу, спрямовані на фільтрацію шумів, детекцію *R*-піків та сегментацію сигналу. Згодом розроблено методи обчислення широкого спектра інженерних ознак — показників варіабельності серцевого ритму (HRV), ентропійних коефіцієнтів та морфологічних параметрів, що мають клінічну інтерпретованість. Попри це, такі ознаки не здатні повністю охопити складну морфологію сирого ЕКГ-сигналу. З розвитком глибинного навчання з'явилися архітектури, здатні автоматично вилучати релевантні патерни: згорткові мережі (CNN) ефективно виявляють локальні морфологічні структури, а рекурентні (RNN, LSTM, GRU) моделюють ритміку та часову динаміку сигналу. Інженерні ознаки є компактними та пояснюваними, однак можуть втрачати тонкі морфологічні відхилення; натомість CNN/LSTM-моделі виявляють складні закономірності, проте мають нижчу інтерпретованість і чутливість до артефактів. Це зумовило появу гібридних архітектур, які поєднують інженерні та глибинні представлення, що забезпечує баланс між точністю, узагальнюваністю та пояснюваністю. В низці сучасних робіт підтверджено ефективність такого комбінованого підходу. Zhang та співавт. [1] об'єднали вейвлет-ознаки з CNN через теорію Демпстера–Шейфера, досягнувши підвищення точності класифікації. Rawshani та ін. [2] продемонстрували двопоточну архітектуру, де один канал обробляє сирий ЕКГ-сигнал, а другий — HRV та демографічні ознаки, що суттєво покращило виявлення фібриляції передсердь. Galli та співавт. [3] застосували Gaussian feature streaming для стабільного моніторингу сигналів у реальному часі. Giordano та співавт. [6] показали переваги багаторівневого злиття (*mid-level fusion*) при поєднанні ручних і глибинних ознак у задачах класифікації аритмій. Таким чином, сучасна література демонструє послідовний рух у напрямі гібридних моделей з адаптивним злиттям інформації з різних джерел. Подальші дослідження спрямовані на підвищення стійкості до шумів та покращення інтерпретованості

рішень. У цьому контексті запропонована модель з модулем *Clinical Gating Unit* є перспективним кроком до адаптивного поєднання клінічних, морфологічних та сигнальних ознак у єдиній нейромережевій системі аналізу ЕКГ.

Це підкреслює актуальність запропонованого підходу з *Clinical Gating Unit* для адаптивного злиття клінічних, морфологічних та сигнальних ознак у єдиному фреймворку.

3. Математична модель вхідного сигналу. Математична модель — це опис певного реального об'єкта, явища чи процесу за допомогою математичних понять, символів, формул і співвідношень. Вона дозволяє відобразити суттєві властивості об'єкта та досліджувати його поведінку без необхідності експериментувати безпосередньо з самим об'єктом.

Основні риси математичної моделі

- Абстракція — відкидання другорядних деталей і виділення головних характеристик.
- Формалізація — опис у вигляді рівнянь, функцій, матриць, ймовірностей тощо.
- Адекватність — модель має відтворювати суттєві властивості реального процесу.

3.1. Формалізація вхідного сигналу. Побудову моделі доцільно розпочати з формалізації вхідного сигналу ЕКГ, який містить кілька каналів, відповідних до системи відведень, що відображають електричну активність серця з різних проєкцій. Загальноприйнято виділяти три основні групи: стандартні біполярні відведення (I — між правою і лівою рукою, II — між правою рукою і лівою ногою, III — між лівою рукою і лівою ногою), підсилені однополюсні від кінцівок (aVR, aVL, aVF), а також грудні (прекардіальні) відведення V1–V6, які розміщуються на грудній клітці відповідно до міжреберних проміжків і анатомічних ліній. Кожне відведення забезпечує унікальний «кут огляду» на серцеву діяльність, що дозволяє моделі враховувати просторову неоднорідність сигналу.

Кожне відведення — це різниця потенціалів між двома точками тіла, де встановлені електроди. Нехай

$\varphi_i(t)$ — електричний потенціал у точці i ,

$L(t)$ — сигнал у відведенні.

Тоді відведення визначається як:

$$L(t) = \varphi_a(t) - \varphi_b(t),$$

де a, b — точки вимірювання.

Для посиленних відведень використовується різниця потенціалів між відповідної кінцівки та середнім потенціалом кінцівок:

$$V_W(t) = \frac{\varphi_{RA}(t) + \varphi_{LA}(t) + \varphi_{LL}(t)}{3}.$$

Якщо узагальнити підхід, то електрична активність серця в кожен момент часу може бути зведена до єдиного еквівалентного електричного диполя. Цей диполь $p(t) = [p_x(t), p_y(t), p_z(t)]$ постійно змінює величину й напрям у часі. Кожне відведення k має свій вектор-спостерігач I_k , який визначається положенням електродів. Тоді сигнал у відведенні є проєкцією вектора диполя на напрям цього відведення:

$$s_k(t) = \mathbf{l}_k \cdot \mathbf{p}(t) = |\mathbf{p}(t)| |\mathbf{l}_k| \cos \theta_k,$$

де θ_k — кут між диполем і напрямком відведення.

Цей результат можна представити у матричній формі:

$$\mathbf{s}(t) = L p(t) + \mathbf{b}(t) + \boldsymbol{\eta}(t).$$

Для кожного відведення сигнал можна розділити на хвилі, які фізіологічно відповідають скороченню серцевого м'язу. Хвилю можна представити як суму компонент:

$$ECG(t) = P(t) + QRS(t) + T(t) + U(t) + baseline(t) + noise(t),$$

де

- $P(t)$ — хвиля передсердної деполяризації,
- $QRS(t)$ — деполяризація шлуночків,
- $T(t)$ — реполяризація шлуночків,
- $U(t)$ — додаткова реполяризація,
- $baseline(t)$ — дрейф ізолінії,
- $noise(t)$ — шумові компоненти.

Також кожен хвилю можна подати аналітично, наприклад використовуючи гаусіан

$$g(t) = A \cdot e^{-\frac{(t-\mu)^2}{2\sigma^2}},$$

де

- A — амплітуда (висота хвилі),
- μ — центр (часова позиція максимуму),
- σ — ширина (визначає тривалість хвилі).

Хвилі Р та Т мають м'яку та "дзвоникоподібну" форму і добре моделюються окремим гаусіаном. Комплекс QRS різкіший, тому його можна апроксимувати кількома гаусіанами (Q — негативний, R — позитивний, S — негативний). В результаті отримуємо модель серцевого циклу через гаусіан:

$$ECG(t) = \sum_{m \in \{P, Q, R, S, T\}} A_m \cdot e^{-\frac{(t-\mu_m)^2}{2\sigma_m^2}}.$$

У своїх працях М. А. Awal[1] та А. Galli[2] дослідили, що модель сигналу ЕКГ може бути представлена як

$$ECG_{model} = \sum_{i \in \{P, Q, R, S, T\}}^{j=2} A_{i,j} e^{-\left[\frac{(t-t_{i,j})^2}{\sigma_{i,j}^2}\right]} + c_i,$$

де j вказує на кількість необхідних гаусіанів. Для вибору параметрів використовуємо метод мінімізації середньоквадратичної похибки (RMSE). Математично це можна записати як:

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^N (ECG_{Model}(t) - ECG_{real}(t))^2} < \epsilon$$

4. Виділення і моделювання основних ознак ЕКГ. Інтервальні ознаки базуються на послідовності RR інтервалів і є основою як для класичного аналізу варіабельності серцевого ритму (HRV), так і для виявлення патологій. Базовою

ознакою є RR -інтервал — час між сусідніми R-піками Нехай $\{t_i\}_{i=1}^N$ — часові координати послідовних R-піків на ЕКГ.

Тоді $RR_i = t_{i+1} - t_i$, $i = 1, 2, \dots, N - 1$.

Визначимо індикатор "нормальності" інтервалу NN (Normal-to-Normal):

$$\delta_i = \begin{cases} 1, & \text{якщо } RR_i \text{ належить до нормального циклу,} \\ 0, & \text{якщо } RR_i \text{ виключається (екстрасистола або артефакт).} \end{cases}$$

Тоді:

$$NN = \{RR_i \cdot \delta_i \mid i = 1, \dots, N - 1 \text{ і } \delta_i = 1\}.$$

Оцінка варіабельності серцевого ритму ґрунтується, зокрема, на аналізі змін ΔRR -інтервалів, з акцентом на абсолютні відхилення між сусідніми значеннями:

$$\Delta RR_i = RR_{i+1} - RR_i, \quad i = 1, 2, \dots, N - 2,$$

так і відносні:

$$\Delta RR^{rel}_i = \frac{RR_{i+1} - RR_i}{RR_i}.$$

Для варіації RR інтервалів обчислюють показник:

$$CVRR = \frac{SDNN}{mean(RR)},$$

де $SDNN$ — стандартне відхилення NN -інтервалів.

4.1. Морфологічні ознаки. До морфологічних ознак належать часові параметри, зокрема тривалості хвиль і комплексів, таких як T_P , T_{QRS} , T_{QT} , T_{ST} . Їх доповнюють амплітудні характеристики — висоти зубців P , R , S , T . Важливими є також інтегральні показники, зокрема площі ділянок, обмежених відповідними фрагментами сигналу. Серед геометричних ознак варто виокремити нахили сегментів, ступінь асиметрії T -хвилі, а також кути між фазовими відрізками комплексу.

4.2. Статистичні ознаки HRV.

- $MeanRR$ — середня тривалість RR .
- $SDNN$ — стандартне відхилення NN -інтервалів (загальна варіабельність).
- $RMSSD$ — корінь з середнього значення квадрату різниць сусідніх NN :

$$RMSSD = \sqrt{\frac{1}{M-1} \sum_{i=1}^{M-1} (NN_{i+1} - NN_i)^2}.$$

- $NN50$ — кількість пар сусідніх NN , які відрізняються більш ніж на 50 мс:
 $NN50 = \sum_{i=1}^{M-1} (|NN_{i+1} - NN_i| > 50 \text{ мс}).$
- $pNN50$ — відсоток сусідніх NN , які відрізняються > 50 мс (ознака «нерегулярності»).

$$pNN50 = \frac{NN50}{M-1} \times 100\%.$$

- ентропія $ApEn(m, r, N) = \phi_m(r) - \phi_{m+1}(r)$
- $MedianRR$, $IQR(RR)$ — робастні характеристики.

4.3. Часові залежності. . Дозволяють визначити ознаки в локальних вікнах.

- $SDANN$ — стандартне відхилення середніх RR у 5-хвилинних вікнах.
- $SDNNi$ — середнє стандартних відхилень у коротких сегментах (локальна мінливість).

4.4. Геометричні та ентропійні ознаки. Діаграма Лоренца — це графічний метод аналізу варіабельності серцевого ритму, який відображає взаємозв'язок між сусідніми NN-інтервалами. Це множина точок.

$$P = \{(NN_i, NN_{i+1}) \mid i = 1, 2, \dots, M - 1\}.$$

На діаграмі формується хмарка точок, яку можна апроксимувати еліпсом. Основні показники:

- поперечна дисперсія (ширина еліпса): $SD1 = \sqrt{\frac{1}{2}} \cdot RMSSD$
- поздовжня дисперсія (довжина еліпса): $SD2 = \sqrt{2 \cdot SDNN^2 - \frac{1}{2} RMSSD^2}$
- $SD1/SD2$ — баланс короткочасної та довготривалої варіабельності.

Певні ознаки або їх комбінації можуть свідчити про ектопічну природу серцевої активності. Зокрема, підвищена дисперсія RR -інтервалів, високе значення ентропії та зниження періодичності ритму є характерними для фібриляції передсердь (AF). Натомість регулярні осциляції RR -інтервалів, що часто корелюють із дихальним циклом, є типовими для синусової аритмії.

5. Побудова гібридної математичної моделі аналізу ЕКГ.

5.1. Розглянемо архітектуру гібридної моделі. Вхідний потік даних складається з 2 частин:

- Вхід 1 (features): $x_f \in \mathbb{R}^d$ — інженерні ознаки ($SDNN, RMSSD, QTc \dots$).
- Вхід 2 (signal): $x_s \in \mathbb{R}^{C \times T}$ — сирий фрагмент ЕКГ (C =кількість відведень, T =довжина вікна).

На його основі будуємо дві “башти”:

- Башта ознак (MLP): $h_f = \text{MLP}(x_f) \in \mathbb{R}^p$.
- Башта сигналу (CNN/LSTM): $\tilde{x}_s = \text{CNN}(x_s) \rightarrow \text{BiLSTM}(\cdot) \rightarrow \text{pooling} \Rightarrow h_s \in \mathbb{R}^q$.

далі відбувається злиття та кінцева класифікація:

- Злиття: $z = [h_f \parallel h_s] \in \mathbb{R}^{p+q}$.
- Голова: $y = \text{Classifier}(z)$.

5.2. Стратегії злиття. У двохвежових (hybrid) моделях ЕКГ існує кілька стратегій злиття даних з ознак та сирого сигналу:

- **Раннє злиття.** Основною ідеєю є об'єднати інженерні ознаки та сирі ознаки та вже потім подати все разом на входи нейромережі. Ця проста ідея має суттєвий недолік, який полягає в різній природі даних, що призводить до проблем з балансуванням масштабу для інженерних ознак.
- **Проміжне злиття.** Тут проводимо об'єднання представлень середнього рівня. Наприклад, CNN витягує морфологічні ознаки з ЕКГ, MLP аналізує статистику HRV і їхні приховані вектори зливаються на проміжному шарі. Дану ідею можна розширити додавши додаткові вежі. Серед недоліків є більша обчислювальна вартість.
- **Пізнє злиття.** Ідея полягає в тому, що кожна вежа повертає свою імовірність належності до класів, і ці ймовірності надалі об'єднуються. Кожна модель тренується окремо, і це є плюсом, однак моделі практично не взаємодіють на етапі тренування і це є недоліком.

Одним з ефективних підходів до інтеграції інженерних та глибинних ознак є використання *проміжного злиття* (mid-level fusion). На відміну від простого об'єднання вихідних рішень або повного злиття на рівні сирих даних, цей підхід дозволяє поєднувати латентні представлення обох «веж» — ознак та сирого сигналу.

Класична схема проміжного злиття ґрунтується на рівноправній конкатенації латентів з «feature tower» h_f та «signal tower» h_s . Проте на практиці ці представлення можуть бути різного масштабу, що вимагає адаптивного визначення важливості кожного з компонентів. З цією метою вводиться параметр ваги $\alpha \in [0, 1]$, який визначає ступінь довіри до ознак або до сирого сигналу при злитті:

$$z = \alpha \cdot h_f + (1 - \alpha) \cdot h_s.$$

Щоб надати вагам клінічної інтерпретованості, пропонується обчислювати їх на основі вектору клінічних ознак $f_{\text{clin}} \in \mathbb{R}^k$. Для цього застосовується сигмоїдна активація:

$$\alpha = \sigma(W f_{\text{clin}} + b),$$

де W та b — параметри, що навчаються під час оптимізації. Таким чином, ступінь впливу кожного з джерел інформації визначається контекстно — відповідно до специфіки клінічної ситуації. Запропонований механізм ми назвемо CGU (Clinical Gating Unit), що формалізує принцип “клінічного керування” злиттям ознак. На рисунку 1 наведемо схему архітектури вдосконаленого механізму злиття:

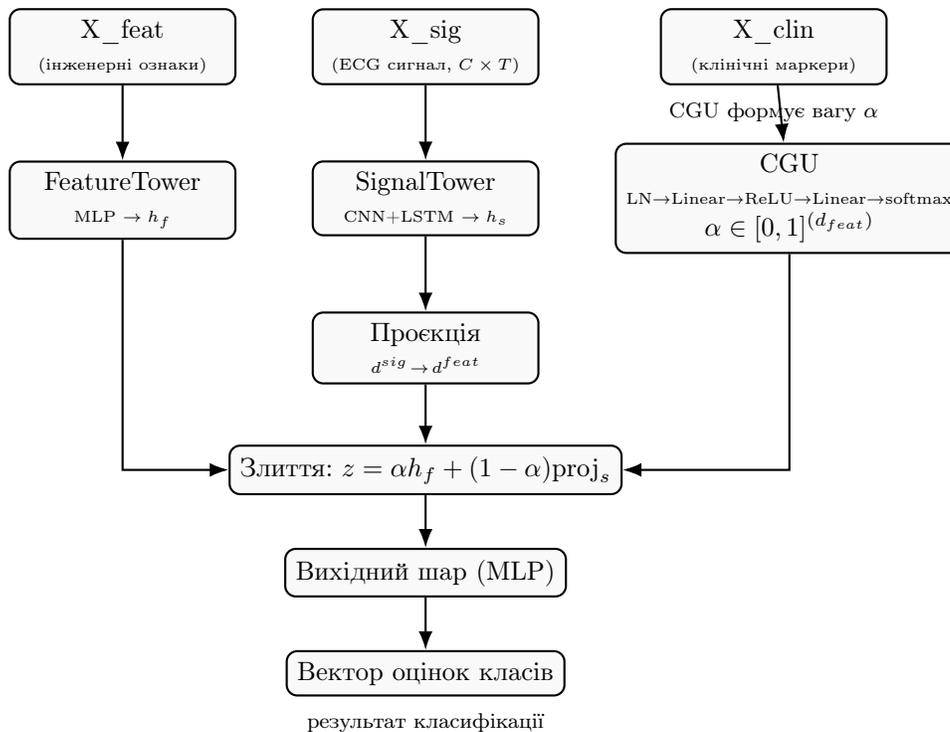


Рис. 1. Схема моделі ECG-ACGF. *FeatureTower* обробляє інженерні ознаки, *SignalTower* — ECG-сигнал, *CGU* використовує клінічні показники для обчислення ваги α , яка визначає внесок обох представлень у спільний латент z .

Описану схему реалізовано програмно. У лістингу нижче наведено приклад реалізації основних класів

FeatureTower. MLP-вежа для інженерних ознак. Перетворює вектор $x_{\text{feat}} \in \mathbb{R}^{d_{\text{in}}}$ у латентне представлення $h_f \in \mathbb{R}^{d_{\text{lat}}}$. Архітектура: LayerNorm \rightarrow Linear(256) \rightarrow ReLU \rightarrow Dropout \rightarrow Linear(d_{lat}) \rightarrow ReLU.

SignalTower. Комбінація згорткових і рекурентних шарів (CNN + BiLSTM) для обробки сирого сигналу ЕКГ. Отримує $x_{\text{sig}} \in \mathbb{R}^{B \times C \times T}$ і формує латент сигналу $h_s \in \mathbb{R}^{B \times d_{\text{lat}}}$. Послідовність операцій: Conv1d \rightarrow ReLU \rightarrow MaxPool \rightarrow BiLSTM \rightarrow Global Average Pooling \rightarrow Linear \rightarrow ReLU.

ClinicalGatingUnit (CGU). Модуль, який із клінічних ознак x_{clin} обчислює вагу $\alpha \in [0, 1]$, що визначає баланс між двома потоками ознак. Структура: LayerNorm \rightarrow Linear(hidden) \rightarrow ReLU \rightarrow Linear(1) \rightarrow Sigmoid.

ECG_ACGF. Основна модель (*Adaptive Clinical-Guided Fusion*), яка об'єднує інформацію з трьох джерел — інженерних ознак, сирого сигналу та клінічних параметрів:

- 1) $h_f = \text{FeatureTower}(x_{\text{feat}})$
- 2) $h_s = \text{SignalTower}(x_{\text{sig}})$
- 3) $\text{proj}_s = W_{\text{proj}} h_s$ — проєкція сигналу у простір ознак
- 4) $\alpha = \text{CGU}(x_{\text{clin}})$
- 5) Злиття: $z = \alpha h_f + (1 - \alpha) \text{proj}_s$
- 6) Класифікаційна голова (MLP) \rightarrow логіти y

Вихід моделі: (logits, α), де α відображає вагу клінічної довіри між двома потоками ознак.

Лістинг 1: Python example

```
import torch
import torch.nn as nn
class FeatureTower(nn.Module):
    def __init__(self, d_in, d_lat=128, p_drop=0.3):
        super().__init__()
        self.net = nn.Sequential(
            nn.LayerNorm(d_in),
            nn.Linear(d_in, 256), nn.ReLU(), nn.Dropout(p_drop),
            nn.Linear(256, d_lat), nn.ReLU()
        )
    def forward(self, x_feat):
        return self.net(x_feat)

class SignalTower(nn.Module):
    def __init__(self, c_in, d_lat=256):
        super().__init__()
        self.cnn = nn.Sequential(
            nn.Conv1d(c_in, 32, kernel_size=7, padding=3), nn.ReLU(),
            nn.Conv1d(32, 64, kernel_size=5, padding=2), nn.ReLU(),
            nn.MaxPool1d(2),
            nn.Conv1d(64, 128, kernel_size=5, padding=2), nn.ReLU(),
            nn.MaxPool1d(2)
        )
        self.lstm = nn.LSTM(
            input_size=128, hidden_size=128,
            num_layers=1, batch_first=True, bidirectional=True
        )
        self.proj = nn.Linear(256, d_lat)

    def forward(self, x_sig):
        f = self.cnn(x_sig)
```

```

        f = f.transpose(1, 2)
        out, _ = self.lstm(f)
        h = out.mean(dim=1)
        return F.relu(self.proj(h))

class ClinicalGatingUnit(nn.Module):
    def __init__(self, d_clin, hidden=32):
        super().__init__()
        self.net = nn.Sequential(
            nn.LayerNorm(d_clin),
            nn.Linear(d_clin, hidden), nn.ReLU(),
            nn.Linear(hidden, 1)
        )

    def forward(self, x_clin):
        alpha_logits = self.net(x_clin)
        alpha = torch.sigmoid(alpha_logits)
        return alpha

class ECG_ACGF(nn.Module):
    def __init__(self, d_feat, c_sig, d_clin, n_classes,
                 d_lat_feat=128, d_lat_sig=256, vector_gate=False):
        super().__init__()
        self.ft = FeatureTower(d_feat, d_lat_feat)
        self.st = SignalTower(c_sig, d_lat_sig)
        self.vector_gate = vector_gate
        if vector_gate:
            self.cgu_f = nn.Sequential(
                nn.LayerNorm(d_clin), nn.Linear(d_clin, d_lat_feat)
            )
            self.cgu_s = nn.Sequential(
                nn.LayerNorm(d_clin), nn.Linear(d_clin, d_lat_sig)
            )
        else:
            self.cgu = ClinicalGatingUnit(d_clin)
        self.head = nn.Sequential(
            nn.Linear(d_lat_feat, 128), nn.ReLU(), nn.Dropout(0.3),
            nn.Linear(128, n_classes)
        )

    def forward(self, x_feat, x_sig, x_clin):
        h_f = self.ft(x_feat)
        h_s = self.st(x_sig)
        proj_s = nn.functional.linear(
            h_s,
            torch.eye(h_s.size(-1), device=h_s.device)[:h_f.size(-1)]
        )
        if not hasattr(self, "proj_to_feat"):
            self.proj_to_feat = nn.Linear(h_s.size(-1), h_f.size(-1)).to(h_s.device)
        proj_s = self.proj_to_feat(h_s)
        if self.vector_gate:
            alpha_f = torch.sigmoid(self.cgu_f(x_clin))
            alpha_s = torch.sigmoid(self.cgu_s(x_clin))
            alpha = torch.sigmoid(alpha_f)
            z = alpha * h_f + (1 - alpha) * proj_s
        else:
            alpha = self.cgu(x_clin)
            z = alpha * h_f + (1 - alpha) * proj_s
        logits = self.head(z)
        return logits, alpha

```

6. Результати експериментів. Для базового оцінювання використано підвибірку бази даних MIT-BIH Arrhythmia. Як інженерні ознаки було обрано ширину, потужність та енергію комплексу QRS, рівень і базову лінію сегмента ST, а також тривалості та амплітуди хвиль. До клінічних ознак віднесено зсув ST (маркер ішемії), скоригований інтервал QT (характеризує відновлення міокарда), локальні RR-інтервали (варіабельність та іррегулярність серцевого ритму) та миттєву частоту серцевих скорочень (HR).

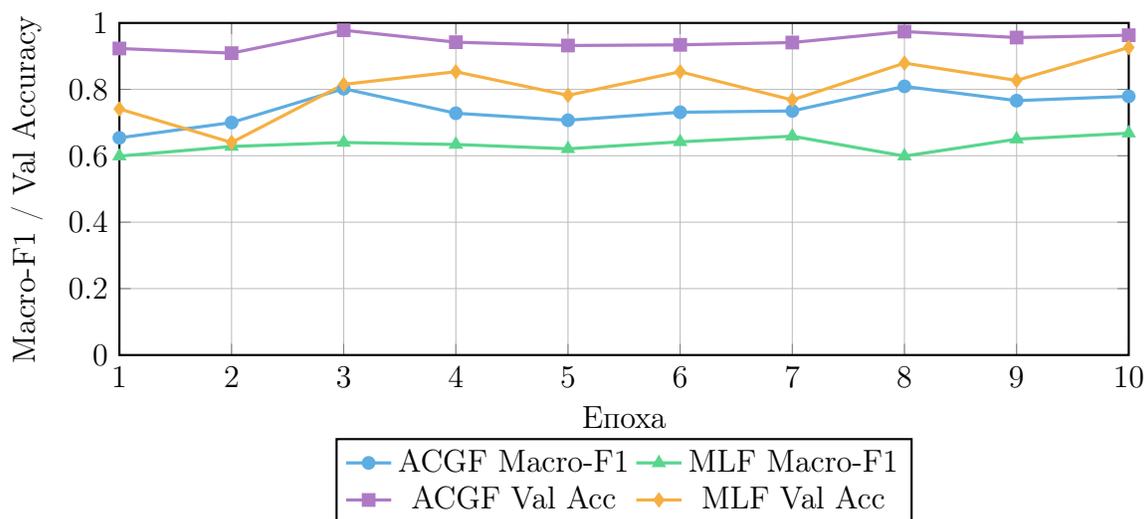


Рис. 2. Порівняння динаміки навчання Adaptive Clinical-Guided Fusion (ACGF) та Mid-Level Fusion (MLF). ACGF демонструє стабільно вищі показники Macro-F1 та точності, особливо після 3-ї епохи.

Експерименти проводилися з різною кількістю вхідних записів та вихідних класів. Вхідні дані розбивалися на тренувальну та тестову вибірки, після чого нормалізувалися і подавалися на вхід як класичної моделі MidLevelFusion, так і розробленої архітектури ECG-ACGF, що дозволяло здійснити коректне порівняння результатів на ідентичних наборах даних.

ECG-ACGF поєднує дві вежі: *FeatureTower* (інженерні/морфологічні ознаки) та *SignalTower* (CNN-LSTM, що обробляє сирій сегмент ЕКГ). Їхній внесок у фінальне рішення зважується блоком *Clinical Gating Unit (CGU)* за клінічними маркерами (X_{clin}). У базовій моделі MidLevelFusion використано статичне злиття представлень шляхом конкатенації:

$$[h_f, \text{proj}(h_s), h_f - \text{proj}(h_s), h_f \odot \text{proj}(h_s)],$$

після чого застосовується MLP для класифікації.

За результатами серії експериментів встановлено, що при невеликій кількості вихідних класів (наприклад, *normal*, *extrasystole*) обидва підходи демонструють високі показники точності, з незначною перевагою моделі з CGU. Зі збільшенням кількості класів та обсягу вхідних даних точність класифікації зменшується, однак деградація якості у ECG-ACGF є істотно меншою порівняно з MidLevelFusion.

Для демонстрації різниці в ефективності було використано 12 записів (22 929 фрагментів) ЕКГ із п'ятьма класами: *normal*, *atrial*, *supraventricular*, *ventricular*, *fusion*. Порівняння проводилося за показниками *validation accuracy* та *macro-F1*.

Порівняння з класичним MidLevelFusion відображено в таблиці 1 та на рисунку 2

Отримані результати демонструють стале покращення показника *macro-F1* для моделі ECG-ACGF (~ 0.78 – 0.81) порівняно з базовою архітектурою MidLevelFusion (~ 0.63 – 0.66), а також суттєву перевагу в точності класифікації

(приблизно 95% проти 82%). Середнє значення коефіцієнта злиття $\alpha \approx 0.45$ з дисперсією $\sigma \approx 0.30$ свідчить про активну адаптацію механізму *Clinical Gating Unit*, який динамічно змінює вагу між морфологічними (FeatureTower) та сигнальними (SignalTower) представленнями залежно від клінічних показників пацієнта.

Таблиця 1.

Порівняння ефективності злиття ознак і сигналу

Модель	Тип об'єднання	Val Acc	Macro-F1	α (mean \pm std)
ECG-ACGF	Адаптивне клінічно кероване злиття	0.95	0.78	0.45 \pm 0.30
MidLevelFusion	Статичне mid-level злиття	0.82	0.63	—

7. Висновки. Запропонований модуль Clinical Gating Unit (CGU) має низку переваг порівняно з класичними підходами до злиття ознак і сигналів у задачах аналізу ЕКГ.

По-перше, він забезпечує адаптивність: вага між сирим сигналом і інженерними ознаками змінюється залежно від клінічного контексту — наприклад, подовження інтервалу QTc , зсуву ST чи зниження варіабельності серцевого ритму ($SDNN$). Завдяки цьому модель динамічно підлаштовується під фізіологічний стан пацієнта.

По-друге, CGU підвищує стійкість до шуму. У разі зашумленого сигналу більша частка ваги автоматично надається стабільним статистичним та морфологічним показникам, тоді як при неточній детекції інтервалів вагу отримує згорткова або рекурентна частина (CNN/LSTM), яка безпосередньо аналізує форму сигналу.

Ще однією суттєвою перевагою є інтерпретованість. Завдяки контролю ваг α можна пояснити, чому модель прийняла певне рішення: наприклад, через низький $SDNN$ вона надала перевагу HRV-ознакам, або через подовжений QTc — сигналові. Такий підхід вводить у модель клінічно обґрунтоване упередження (clinical prior), що зменшує ризик покладання на випадкові патерни чи артефакти.

Крім того, розділення функцій вагування та класифікації спрощує навчання — кінцевий класифікатор фокусується лише на прогнозуванні, не вирішуючи одночасно задачу балансування джерел інформації.

Нарешті, модульність CGU робить систему гнучкою та розширюваною: до неї легко інтегрувати додаткові клінічні параметри — від SpO_2 і артеріального тиску до демографічних чи лабораторних показників.

Модуль Clinical Gating Unit забезпечує більш стійке, пояснюване й клінічно обґрунтоване злиття інформації, ніж традиційні методи mid-level fusion.

Конфлікт інтересів

Автор заявляє, що не має конфлікту інтересів щодо даного дослідження, включаючи фінансовий, особистий, авторський або будь-який інший, який міг би вплинути на дослідження, а також на результати, представлені в даній статті.

Фінансування

Дослідження здійснено в рамках кафедральної науково-дослідної роботи

«Моделі і методи системного аналізу в міждисциплінарних дослідженнях» (державний обліковий номер 0125U003246).

Доступність даних

Усі дані доступні в цифровій або графічній формі в основному тексті рукопису.

Використання штучного інтелекту

Автор підтверджує, що при створенні даної роботи вони не використовувались технології штучного інтелекту.

Авторські права ©



(2026). Самусь В. М. Ця робота ліцензується відповідно до Creative Commons Attribution 4.0 International License.

Список використаної літератури

1. Zhang, J., Wu, C., Ruan, C., Zhang, R., Zhao, Z., & Cheng, X. (2021). ECG Signal Classification Based on Fusion of Hybrid CNN and Wavelet Features by D-S Evidence Theory. *Journal of Healthcare Engineering*, 2021, 4222881. <https://doi.org/10.1155/2021/4222881>
2. Rawshani, A., Smith, G., Boren, J., Bhatt, D. L., Börjesson, M., Engdahl, J., Kelly, P., Louca, A., Ramunddal, T., Andersson, E., Omerovic, E., Mandalenakis, Z., & Gupta, V. (2025). Integrating deep learning with ECG, heart rate variability and demographic data for improved detection of atrial fibrillation. *Open Heart*, 12(1), e003185. <https://doi.org/10.1136/openhrt-2025-003185>
3. Galli, A., Giorgi, G., & Narduzzi, C. (2023). Accurate ECG monitoring by Gaussian feature streaming. *Measurement*. <https://doi.org/10.1016/j.measurement.2023.113757>
4. Awal, M. A., Mostafa, S. S., Ahmad, M., Alahe, M. A., Rashid, M. A., Kouzani, A. Z., & Mahmud, M. A. P. (2021). Design and Optimization of ECG Modeling for Generating Different Cardiac Dysrhythmias. *Sensors*, 21, 1638. <https://doi.org/10.3390/s21051638>
5. Fira, M., Gordan, C., & Ionescu, B. (2022). An Explainable Deep Learning Framework for Multilead ECG Classification. *Biomedical Signal Processing and Control*, 73, 103402. <https://doi.org/10.1016/j.bspc.2021.103402>
6. Giordano, D., & Laurito, M., & Spampinato, C. (2021). Fusion of handcrafted and deep features for ECG classification. *Computers in Biology and Medicine*, 135, 104552. <https://doi.org/10.1016/j.combiomed.2021.104552>

Samus V. M. Hybrid ECG Signal Analysis Model Based on Engineered Features and Deep Learning.

This paper presents a hybrid model for electrocardiogram (ECG) signal analysis that combines handcrafted features with deep learning. An adaptive fusion architecture is proposed, integrating feature- and signal-based inputs using a Clinical Gating Unit. Mathematical models of ECG generation are reviewed alongside temporal, morphological, statistical, and geometrical characteristics. A PyTorch-based implementation of the proposed architecture is provided. The hybrid approach is shown to improve classification accuracy, robustness, and interpretability.

Keywords: mathematical model, ECG, HRV, neural network, deep learning, CNN, LSTM, handcrafted features, Clinical Gating Unit.

Отримано: 04.11.2025

Прийнято: 24.11.2025

Опубліковано: 29.01.2026