

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ “ЛЬВІВСЬКА ПОЛІТЕХНІКА”

ХОМА Юрій Володимирович



УДК 621.382: 681.58:681.32

**ТЕОРІЯ І МЕТОДИ КОМП'ЮТЕРНОГО ОПРАЦЮВАННЯ
БІОСИГНАЛІВ НА ОСНОВІ МАШИННОГО НАВЧАННЯ**

05.13.05 – Комп'ютерні системи та компоненти

Автореферат
дисертації на здобуття наукового ступеня
доктора технічних наук

Львів – 2020

Дисертацією є рукопис.

Робота виконана на кафедрі інформаційно-вимірювальних технологій Національного університету «Львівська політехніка» Міністерства освіти і науки України

Науковий консультант:

доктор технічних наук, професор

Микийчук Микола Миколайович,

Національний університет «Львівська політехніка»,

директор Інституту комп'ютерних технологій, автоматики і метрології

Офіційні опоненти:

член-кор. НАН України, доктор технічних наук, професор

Бабак Віталій Павлович,

заступник директора Інституту технічної теплофізики НАН України, м. Київ

доктор технічних наук, професор

Саченко Анатолій Олексійович

завідувач кафедри інформаційно-обчислювальних систем і управління

Західноукраїнського національного університету, м. Тернопіль

доктор технічних наук, професор

Русин Богдан Павлович

завідувач відділом методів та систем дистанційного зондування

Фізико-механічного інституту ім. Г.В. Карпенка НАН України, м. Львів

Захист відбудеться 26 лютого 2021 р. о 12⁰⁰ год. на засіданні спеціалізованої вченої ради Д 35.052.08 у Національному університеті «Львівська політехніка» за адресою: 79013, Львів-13, вул. С. Бандери, 12, ауд. 226 головного корпусу.

З дисертацією можна ознайомитись у бібліотеці Національного університету «Львівська політехніка».

Автореферат розісланий “ ____ ” січня 2021 р.

Учений секретар спеціалізованої
вченої ради, д.т.н., професор



Луцик Я.Т.

ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

Актуальність роботи. Біосигнали є носіями цінної біомедичної інформації, а їх застосування у медичній діагностиці має тривалу історію. Ще у 1924 році нідерландський фізіолог Віллем Ейнтговен був нагороджений Нобелівською премією за відкриття механізму електрокардіограми. На цей час існує апаратура для реєстрації і опрацювання десятків видів біосигналів. Більше того, опрацювання біосигналів поширилося на вирішення нових задач у таких сферах як клінічний моніторинг, біометрика, афективна інформатика, людино-машинна взаємодія.

Проте різнорідність біосигналів та різноплановість задач біоінформатики спричинили вироблення спеціалізованих рішень для видобування біомедичних показників, що ускладнює перенесення знань і методик із однієї сфери на іншу. Крім того, завершальний етап досліджень, пов'язаний із інтерпретацією результатів виконується спеціалістами відповідного фаху.

Природна мінливість біосигналів і їх сприйнятливості до різних дестабілізуючих впливів обмежує ефективність традиційних методів комп'ютерного опрацювання біосигналів, що спираються на формальні моделі. Для цих цілей видається доцільним застосувати інструменти машинного, зокрема глибинного навчання, які довели свою ефективність під час вирішення подібних проблем, наприклад, розпізнаванні мовних сигналів.

Відмінною особливістю машинного навчання є застосування нечітких (розмитих) моделей для опрацювання даних. Побудова таких моделей відбувається за допомогою наперед підготовленого (розміченого) навчального набору даних. Адекватність так збудованої моделі, а, отже, і достовірність результатів опрацювання напряму залежить від репрезентативності навчального набору даних. Тому передумовою застосування машинного і особливо глибинного навчання є наявність великого обсягу даних, а їх підготовка є одним із ключових завдань.

З погляду практичного застосування, на цей час існує потреба як в уніфікації вимірювальних перетворень для підвищення достовірності результатів досліджень шляхом синхронного опрацювання різних видів біосигналів, так і у спрощенні спеціалізованих рішень, щоб, наприклад, забезпечити доступність для персоніфікованого використання засобів біоінформатики у побутових умовах. Разом з тим у теорії відсутня цілісна, науково-обґрунтована методологія побудови комп'ютерних систем для опрацювання біосигналів, яка б враховувала, з одного боку, їх специфічність, а також різний характер задач сучасної біоінформатики, а з іншого, можливості сучасних обчислювальних платформ та інструментів машинного навчання. Отже, існує об'єктивне протиріччя між станом теорії та потребами практики, що утворює науково-прикладну проблему, вирішення якої і є предметом даного дисертаційного дослідження.

Таким чином, розроблення нових, а також вдосконалення відомих підходів і методів підвищення точності та обчислювальної ефективності програмно-апаратних засобів комп'ютерного опрацювання біосигналів і даних на основі поєднання методів цифрового опрацювання сигналів і технологій машинного навчання – є актуальною науково-прикладною проблемою, що має теоретичне і практичне значення.

Зв'язок роботи з науковими програмами, планами, темами. Дисертаційна робота виконана в рамках пріоритетних напрямків розвитку науки і техніки в Україні, а також в межах держбюджетної науково-дослідної роботи, що виконується у Національному університеті «Львівська політехніка» за тематичними планами Міністерства освіти і науки України “Розроблення програмно-апаратного забезпечення для інструментального методу реабілітації хворих облітеруючим атеросклерозом судин кінцівок шляхом in-suit електростимуляції з урахуванням серцевого ритму” за номером держреєстрації № 0120U102205, 2020 р.

Мета і задачі дослідження. Метою дисертаційної роботи є створення теоретичних засад для удосконалення методологічної, алгоритмічної і програмно-технічної бази комп'ютерних систем опрацювання біосигналів і даних на основі широкого використання штучних нейронних мереж та технологій глибинного навчання.

Для досягнення поставленої мети необхідно розв'язати такі **основні задачі досліджень**:

1. Проаналізувати сучасний стан та перспективи розвитку комп'ютерних систем аналізу і опрацювання біомедичних даних. Окреслити невирішені завдання, сформулювати науково-прикладну проблему та обґрунтувати доцільні підходи щодо її вирішення, передовсім оцінити ступінь придатності різних методів штучного інтелекту та глибинного навчання до застосування у комп'ютерних біоінформатичних системах.

2. Розвинути науково-методологічну базу комп'ютерного опрацювання біосигналів із урахуванням потреб практики, шляхом широкого використання технологій машинного навчання, а також удосконалення методів цифрового опрацювання біосигналів.

3. Проаналізувати зміст етапів перетворення біосигналів у різних застосуваннях з урахуванням факторів, що ускладнюють видобування біомедичних показників та інтерпретацію результатів досліджень. На основі комплексного критичного аналізу актуальних вимог до комп'ютерного опрацювання біосигналів та оцінювання можливостей сучасних програмно-апаратних засобів обґрунтувати узагальнену структуру комп'ютерної вимірювальної системи для задач біоінформатики.

4. Розробити концепцію покращення характеристик комп'ютерних систем опрацювання біосигналів для широкого спектру задач біоінформатики, на основі поєднання методів цифрового опрацювання сигналів і машинного навчання, імплементувати розроблену концепцію у відповідний фреймворк, а також обґрунтувати методологію виконання дисертаційних досліджень і вибрати критерії оцінювання ефективності визначення біомедичних показників.

5. Розробити нові й удосконалити відомі підходи до підвищення точності та обчислювальної ефективності методів і засобів первинного опрацювання біосигналів, які б послужили підґрунтям уніфікації аналогової частини вимірювального каналу.

6. Дослідити можливості покращення характеристик первинних перетворювачів біоімпеданса, зокрема:

- застосувати нейронну мережу для коригування динамічних похибок перетворювачів автокомпенсаційного типу, що дасть змогу розширити діапазон робочих частот аналізатора біоімпеданса;

- побудувати на основі незрівноваженої мостової схеми перетворювач змін адмітанса підвищеної чутливості для контролю електродермальної активності в задачах афективної інформатики;

- розробити метод підвищення обчислювальної ефективності цифрового фазоквадратурного перетворювача сигналів імпедансної спектроскопії.

7. Побудувати та виконати комплексне дослідження системи біометричної ідентифікації за ЕКГ-сигналами на основі глибоких нейронних мереж:

- розробити новий підхід до виявлення і коригування залишкових аномалій у біосигналах, який базується на застосуванні нейромережових автоенкодерів для нелінійної фільтрації завад, зосереджених у тій самій смузі, що й корисний сигнал;

- удосконалити метод автоматичного пошуку оптимальних значень гіперпараметрів багатопарового нейрокласифікатора із використанням прогностичної моделі збудованої за допомогою обчислювально простого алгоритму машинного навчання;

- розробити метод програмного кондиціонування біосигналів, що забезпечує інваріантність комп'ютерної системи біометричної ідентифікації до параметрів біосенсорів і апаратури відбору біосигналів;

- виконати комплексну апробацію ефективності розроблених методів націлених на забезпечення надійності електрокардіограми як біометричного маркера.

8. Удосконалити метод інтелектуального опрацювання сигналів віброартрографії за рахунок суміщення технологій цифрового оброблення сигналів і машинного навчання:

- виконати хвилькову декомпозицію ВАГ-сигналу, вибрати рівні на скалограмі, що містять релевантну інформацію;

- виконати пошук дескрипторів, які репрезентують діагностичні ознаки ступеня артроскопічного захворювання суглобів;

- збудувати нечітку модель класифікатора системи підтримки прийняття рішення для проведення медичної діагностики стану колінних суглобів.

9. Удосконалити структуру вимірювального реографічного каналу завдяки:

- застосуванню принципу компенсації базового біоімпеданса за допомогою кодированого потенціометра (тримера), причому керівні дії формуються мікроконтролером у контурі від'ємного зворотного зв'язку, що дало змогу покращити метрологічні характеристики пристрою;

- модифікації схеми джерела струму Хавленда, що уможливило компенсацію базового біоімпеданса і розширює функціональні можливості реографічних досліджень;

- використанню сформованих на засадах прямого цифрового синтезу ортогональних зондувальних сигналів, що дає змогу удвічі підвищити швидкодію.

10. Навести приклади практичної реалізації результатів дисертаційних досліджень:

- програмно-апаратну реалізацію комп'ютерної системи біометричної ідентифікації на базі глибоких нейронних мереж;
- біо-машинного інтерфейсу управління за сигналами електроенцефалограми з використанням глибоких нейронних мереж;
- систему підтримки прийняття рішень медичної діагностики остеоартриту колінних суглобів на основі сигналів віброартрографії;
- комп'ютерної системи для побудови нечіткої моделі на базі EMR-даних (записів електронних медичних карт) з метою прогнозування захворювань.

Об'єктом дослідження є процеси відбору і опрацювання біосигналів, а також інтерпретації результатів біомедичних досліджень.

Предметом дослідження є принципи і методи підвищення точності та обчислювальної ефективності програмно-апаратних засобів комп'ютерного опрацювання біосигналів і даних на основі технологій цифрового опрацювання сигналів і машинного навчання.

Методи дослідження. Виконані дисертаційні дослідження базуються на методах теорії систем і системного аналізу (для дослідження особливостей функціонування комп'ютерних біоінформатичних систем та виділення системних рівнів), теорії вимірювань (для планування біомедичних вимірювальних експериментів і оцінювання похибок вимірювання), теорії нечітких множин (для структурно параметричного синтезу неформальних моделей нейрокласифікаторів, автоенкодерів, нейромережевого еквалайзера), теорії ймовірностей та математичної статистики (для аналізу часових рядів, розробки методу виявлення аномальних відхилень, формування за методом Монте-Карло навчальної вибірки для побудови моделі нейромережевого еквалайзера, опрацювання результатів експериментальних досліджень та імітаційного моделювання), методів машинного навчання (для вибору гіперпараметрів глибокої нейронної мережі, алгоритмів навчання і верифікації точності роботи класифікаторів, автоенкодерів, еквалайзерів), а також імітаційного моделювання (для дослідження розроблених моделей).

Наукова новизна одержаних результатів полягає у такому:

1. Одержала подальший розвиток теорія комп'ютерного опрацювання біомедичної інформації, в рамках якої обґрунтовано широке, комплексне використання на всіх структурних рівнях штучних нейронних мереж і технологій глибокого навчання у поєднанні із методами цифрового опрацювання сигналів, що слугує універсальною платформою для розширення функціональних можливостей і покращення характеристик комп'ютерних систем для ефективного вирішення різнопланових завдань у сфері біоінформатики.

2. Вперше представлено концепцію трьох системних рівнів комп'ютерного опрацювання біосигналів, в основу якої покладено ідею чіткого розмежування функцій системних рівнів від методів і засобів, що застосовуються для реалізації цих функцій; таке розділення сприяє структуризації знань у предметній галузі, уможливорює зіставлення і оцінювання ефективності різних методів, а також вибір кращих рішень на етапі проектування біоінформатичних систем під конкретні завдання із урахуванням умов і сценаріїв їх використання.

3. Вперше розроблено метод пошуку оптимальних значень гіперпараметрів багатопараметричного нейрокласифікатора, особливістю якого є використання

прогностичної моделі збудованої за допомогою обчислювально простого алгоритму машинного навчання, а його застосування на 4 порядки скорочує час пошуку порівняно із повним перебором в просторі можливих значень.

4. Вперше запропоновано і апробовано підхід до виявлення та коригування залишкових аномалій в біосигналах, який базується на застосуванні нейромережових автоенкодерів для нелінійної фільтрації завад, зосереджених в тій самій частині спектру, що й корисний сигнал, а його застосування у 5-7 разів зменшує похибку ідентифікації.

5. Вперше застосовано штучну нейромережу для коригування динамічних похибок у вимірювальному каналі частотного аналізатора біоімпеданса, що дало змогу на три порядки розширити смугу робочих частот та дещо покращити точність порівняно із відомим формальним методом алгоритмічної корекції.

6. Вперше розроблено метод програмного кондиціонування біосигналів, який передбачає уніфікацію параметрів цифрових записів біосигналів із відкритих баз шляхом передискретизації, унормування за рівнем і тривалістю, а його застосування вирішує проблему нестачі даних для навчання глибокої нейромережі.

7. Удосконалено метод інтелектуального опрацювання сигналів віброартрографії, який відрізняється виділенням діагностичних ознак, одержаних статистичними методами на вибраних рівнях хвилькової декомпозиції ВАГ-сигналу, а його застосування забезпечило високу точність і чутливість класифікатора системи медичної діагностики попри малий обсяг даних.

8. Вперше виконано комплексне дослідження ефективності різних методів побудови класифікаторів систем біометричної ідентифікації на основі електрокардіограми, що дало можливість вибрати методи, які зберігають інваріантність до параметрів системи відбору ЕКГ-сигналу та придатні до масштабування (число користувачів, тривалість записів, неоднорідність даних, перекошеність класів і т.д.).

9. Вперше розроблено метод реалізації цифрового фазоквадратурного перетворювача сигналів імпедансної спектроскопії, який на відміну від відомих відрізняється високою обчислювальною ефективністю завдяки заміні операцій акумуляції добутків арифметичного перемноження вибірок аналізованої напруги і опорних сигналів синуса і косинуса на просте підсумовування вибірок аналізованої напруги упродовж інтервалів, які визначаються логічними операціями «&» від функції знаку опорних сигналів.

10. Удосконалено структуру вимірювального реографічного каналу завдяки:

- застосуванню принципу компенсації базового біоімпеданса за допомогою кодированого потенціометра (тримера), причому керівні дії формуються мікроконтролером у контурі від'ємного зворотного зв'язку, що дало змогу покращити метрологічні характеристики пристрою;

- модифікації схеми джерела струму Хавленда, що уможливило компенсацію базового біоімпеданса і розширює функціональні можливості реографічних досліджень;

- використанню сформованих на засадах прямого цифрового синтезу ортогональних зондувальних сигналів, що дає змогу удвічі підвищити швидкодію.

Практичне значення отриманих результатів підтверджується таким:

1. Запропоновано репрезентувати розроблені моделі та методи за трьома профілями, що спрощує аналіз і проектування біоінформатичних систем з огляду на їх складність і багатоваріантність:

– за видами біосигналів і даних (електрокардіограма, енцефалограма, реограма, імпедансна спектрограма, віброартрограма, електронні медичні карти і т.д.);

– за системними рівнями, на яких виконується опрацювання (сенсорний/інструментальний рівень відбору і аналогових перетворень біосигналів, рівень кондиціонування сигналів методами цифрової обробки і машинного навчання, рівень визначення та інтерпретації біомедичних параметрів для конкретного застосування);

– за характером прикладних задач біоінформатики (біометрична ідентифікація, медична діагностика, клінічний моніторинг, афективна інформатика, людино-машинна взаємодія).

2. Розроблено фреймворк для проектування комп'ютерних систем біоінформатики, застосування якого пришвидшує процес проектування, забезпечує гнучкість налаштування на різні задачі біоінформатики, уможливорює вибір лише потрібних елементів, з урахуванням умов та сценаріїв використання системи, ефективну імплементацію на різних обчислювальних платформах залежно від специфіки і складності обчислювальних задач (зокрема і на етапі навчання), а також забезпечує гнучкість і раціональний розподіл програмно-апаратних ресурсів, відкриває можливість масштабування системи, модифікації її обчислювальних компонент, підвищує надійність і спрощує верифікацію системи.

3. Виконано комплексне дослідження впливу варіативності параметрів системи відбору ЕКГ-сигналу (різних наборів даних, частоти дискретизації, тривалості запису, обсягу бази) на точність біометричної ідентифікації, результати якого показали достатню для потреб практики стійкість і надійність електрокардіограми як біометричного маркера та підтвердили можливість його застосування у реальних системах біометричної ідентифікації.

4. Розроблено нову високочутливу схему первинного перетворювача біоімпедансу для системи контролю електродермальної активності в задачах афективної інформатики.

5. Обґрунтовано методологію оцінювання ефективності комп'ютерних систем біометричної ідентифікації, що враховує природу біосигналів та специфіку інструментів штучного інтелекту, а її застосування дає змогу дослідити надійність біометричних маркерів, а також провести порівняльний аналіз відомих та розроблених методів ідентифікації.

6. Результати експериментальних досліджень підтверджують правильність наукових положень щодо розширення функціональних можливостей і покращення характеристик комп'ютерних систем біоінформатики на основі широке використання штучних нейронних мереж і технологій глибокого навчання у поєднанні із методами цифрового опрацювання сигналів.

7. Теоретичні та практичні результати роботи впроваджено у діяльність ДП НДІ "Система" (м. Львів) та ТОВ "СВІФТ СОЛЮШНС" (м. Харків), а також використано в навчальному процесі НУ «Львівська політехніка».

Обґрунтованість і достовірність наукових положень, висновків і рекомендацій. Наукові положення, висновки і рекомендації дисертації обґрунтовуються коректним використанням математичного апарату, та доцільними допущеннями, які не суперечать відомим підходам та забезпечують адекватність застосованих моделей. Достовірність наукових положень, висновків і рекомендацій підтверджується узгодженістю результатів теоретичних досліджень та імітаційного моделювання, практичною імплементацією розроблених моделей і методів, результатами експериментів на реальних біосигналах і даних, які узгоджуються із даними відомих досліджень, а також практичним впровадженням результатів дисертаційного дослідження

Особистий внесок здобувача. Основні положення і результати дисертаційної роботи, що виносяться на захист, здобувач отримав особисто. В роботах, опублікованих одноосібно автором, отримано такі результати:

- дослідження інваріантності до типу і параметрів апаратури реєстрації електрокардіограми, тривалості записів, числа користувачів у базі даних на точність біометричної ідентифікації [4];

- обґрунтування надійних дескрипторів на основі дискретного хвилькового перетворення ВАГ-сигналу та дослідження інструментів машинного навчання, які класифікують хрящові розлади в колінному суглобі за загальноприйнятими діагностичними критеріями [5];

- представлення ієрархічно-модульного принципу побудови комп'ютерних біоінформатичних систем та його застосування на прикладі кількох архітектур, призначених для реалізації різних завдань біоінформатики, за різних сценаріїв та у різних умовах [14].

У роботах, опублікованих у співавторстві, автору належать:

- ідея і практичне застосування штучної нейронної мережі на інструментальному рівні для коригування похибок вимірювального каналу частотного аналізатора імпеданса [1,24,26];

- принцип застосування глибоких нейронних мереж для розпізнавання людей за їх електрокардіограмами без виділення дескрипторів [2,23,25];

- концептуальні засади комплексного використання технологій глибокого навчання у поєднанні із методами цифрового опрацювання сигналів, як універсальної платформи для забезпечення гнучкості та розширення функціональності засобів біоінформатики [3,27,30];

- аналіз проблеми варіативності ЕКГ-сигналу з позицій біометрики та метод програмного кондиціонування біосигналів шляхом передискретизації, унормування за рівнем і тривалістю [6];

- підхід до виявлення та коригування промахів у біосигналах, який базується на застосуванні нейромережових автоенкодерів для нелінійної фільтрації залишкових аномалій [7,29,32];

- ідея і концепція чіткого розмежування функцій нижнього, середнього та верхнього системних рівнів від методів і засобів, що застосовуються для реалізації цих функцій [8];

- шляхи уніфікації вимірювальних перетворень біосигналів різної природи та синтез структури аналогової частини комп'ютерної вимірювальної системи [9];

- опрацювання нового підходу до побудови біосенсора, як базового компонента портативного частотного аналізатора, а також синтез цифрового еквалайзера [10];

- розроблення структури та програмно-апаратна імплементація біометричної системи автентифікації на основі ЕКГ-сигналу [11,28];

- опрацювання методики проведення експериментів та визначення критеріїв порівняння ефективності алгоритмів коригування промахів розробленими методами [12];

- синтез апаратного і програмного забезпечення для лікувальної фізкультури шляхом електростимуляції м'язів in-suit з урахуванням серцевого ритму [13];

- архітектура інтелектуальної системи опрацювання ЕЕГ-сигналів для нейрокомп'ютерного інтерфейсу управління колісним роботом, а також підбір гіперпараметрів для згорткової мережі і багатошарового перцептрона [15];

- методологія комплексного дослідження надійності електрокардіограми як біометричного маркера і ступеня інваріантності до параметрів апаратури відбору, методів кондиціонування біосигналів, тривалості записів, перекошеності класів [16];

- підходи до удосконалення структури вимірювального реографічного каналу, обчислювально ефективний спосіб фазоквадратурного розділення [17];

- комплексне дослідження ефективності різних методів побудови класифікаторів систем біометричної ідентифікації з погляду інваріантності до параметрів системи відбору ЕКГ-сигналу та масштабованості [18];

- розроблення підходу та імплементація відповідного алгоритму для дескриптивної статистики, обчислення цільового параметра дослідження (Hazard ratio, графік Kaplan-Meier) та їх довірчих інтервалів для наборів даних IMS French EMR, IMS Ambulatory та IMS P-Plus [19];

- розроблення і реалізація методу для побудови та навчання прогностичних моделей на основі алгоритмів машинного навчання для прогнозування ефективності лікування різновиду Неоджкінської лімфоми (Diffuse Large B-Cell Lymphoma) на основі записів медичної історії пацієнтів (Electronic Medical Records) [20];

- удосконалення алгоритмів коригування динамічних похибок біоімпедансного спектроаналізатора формальними методами [21];

- розроблення концепції коригування промахів та аномалій ЕКГ-сигналу у системах біометричної ідентифікації [22];

- метод автоматичного пошуку оптимальних значень гіперпараметрів багатошарового нейрокласифікатора, з використанням прогностичної моделі на основі простого алгоритму машинного навчання [31].

З робіт, що опубліковані у співавторстві, у роботі використовуються виключно результати, отримані особисто здобувачем.

Апробація результатів дисертації. Основні положення дисертаційної роботи і конкретні наукові результати досліджень доповідались, обговорювались та отримали схвальний відгук на таких конференціях:

- V Міжнародна науково-технічна конференція «Захист інформації і безпека інформаційних систем» (2–3 червня 2016 року, Львів 2016, Україна);

- XLVIII Międzynarodowa Konferencja Metrologów MKM 2016, Akademia Górniczo-Hutnicza w Krakowie, (5-7 вересня 2016 року, Краків, Польща);
- VI Міжнародна науково-технічна конференція «Захист інформації і безпека інформаційних систем» (1–2 червня 2017 року, Львів 2017, Україна);
- The 9th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (21-23 вересня, 2017 року, Бухарест, Румунія);
- 59th Ilmenau Scientific Colloquium, Technische Universität Ilmenau (11 – 15 вересня 2017 року, Ільменау, Німеччина);
- The 3rd International Scientific Conference on Brain-Computer Interfaces (13–14 березня 2018 року, Ополе, Польща);
- The 23rd International Conference on Methods and Models in Automation and Robotics (27-30 серпня 2018 року, Мендзиздрое, Польща);
- The 2018 IEEE 4th International Symposium on Wireless Systems within the International Conferences on Intelligent Data Acquisition and Advanced Computing Systems (20-21 вересня 2018 року, Львів, Україна);
- V Всеукраїнська науково-технічна конференція у царині метрології «Technical Using of Measurement – 2018» (13 - 18 лютого 2018 року, Славське, Україна);
- V Всеукраїнська науково-технічна конференція у царині метрології «Technical Using of Measurement – 2019» (29 січня - 2 лютого 2019 року, Славське, Україна);
- VII Міжнародна науково-технічна конференція “Захист інформації і безпека інформаційних систем” (30-31 травня 2019, Львів, Україна);
- Міжвідомчі міжрегіональні семінари Наукової Ради НАН України «Технічні засоби захисту інформації» (14 березня 2019 року, 14 травня 2020 року, Львів, Україна);
- V Всеукраїнська науково-технічна конференція у царині метрології «Technical Using of Measurement – 2020» (4 - 7 лютого 2020 року, Славське, Україна).

Публікації. За результатами виконаних у дисертаційній роботі досліджень опубліковано 34 наукові праці (3 одноосібні), зокрема:

- 15 статей у наукових фахових виданнях України [1-15];
- 7 статей [16–22] у міжнародних рецензованих періодичних виданнях, що індексуються в наукометричних базах даних *Web of Science*, *Scopus* (відповідно до класифікації SCImago Journal and Country Rank дві статті віднесено до квартилю Q1, дві - до квартилю Q2, три – до квартилю Q3);
- 12 статей та матеріалів доповідей на міжнародних, всеукраїнських та регіональних конференціях і семінарах [23-34], з яких 3 включено до наукометричних баз даних *Web of Science*, *Scopus*.

Структура та обсяг дисертації. Дисертація складається з анотації, змісту, переліку умовних позначень, вступу, шести розділів, загальних висновків, додатків, списку використаних джерел та має 282 сторінки основного тексту, 96 рисунків, 25 таблиць, 22 сторінки додатків. Список використаних джерел містить 352 найменування. Загальний обсяг дисертаційної роботи становить 379 сторінок.

ОСНОВНИЙ ЗМІСТ РОБОТИ

У **вступі** подано загальну характеристику роботи, обґрунтовано актуальність теми дисертації, сформульовано мету і завдання досліджень, відображено наукову новизну й практичну цінність отриманих результатів, наведено дані щодо їх апробації та впровадження.

У **першому розділі** проаналізовано сучасний стан, сфери застосування та перспективи розвитку комп'ютерних систем аналізу і опрацювання біомедичних даних. Розглянуто питання взаємозв'язку біосигналів із психофізіологічними процесами в організмі людини та біомедичними показниками. Розширено класифікацію біосигналів за способом їх формування, в рамках якої виділено три класифікаційні групи: біосигнали електричної природи (такі як ЕКГ, ЕЕГ), біосигнали, що відображають неелектричні процеси (напр., віброартрографія) і біосигнали як реакція на зовнішні стимули (біоімпедансометрія, фотоплетизмографія і т.п.). Наведено стисло характеристику найпоширеніших методів біомедичних досліджень.

Природна мінливість біосигналів та їх сприйнятливність до різних дестабілізуючих впливів знижує ефективність відомих методів цифрового опрацювання сигналів (ЦОС), що спираються на формалізовані моделі. Класичні алгоритми машинного навчання (МН) застосовуються лише для реалізації неформальної процедури із інтерпретації результатів дослідження. Проте, зазвичай, це стосується якоїсь конкретної задачі. У зв'язку з цим, обґрунтовано доцільність застосування технологій машинного навчання до вирішення різнопланових завдань, особливо у таких «немедичних» сферах як біометрика, афективна інформатика та людино-машинна взаємодія.

Різноманітність біосигналів, з одного боку, і різноплановість цілей з іншого позначилися на виробленні спеціалізованих рішень щодо програмно-інструментального забезпечення. Це ускладнює обмін знаннями між спеціалістами різних напрямків і перенесення методик з однієї підгалузі в іншу, а відтак стримує розвиток біоінформатики.

Є багато різних біосигналів, причому шляхом їх синергії (сумісного опрацювання) завдяки нададитивному ефекту можна підвищити точність і достовірність результатів виконання однієї конкретної задачі (рис. 1,а). Такий підхід у сучасній біоінформатиці одержав назву злиття сенсорів (Sensor Fusion). Наприклад, синхронно опрацьовуючи такі сигнали як ЕКГ, ЕЕГ, ІПГ вдається значно покращити достовірність роботи поліграфа. Для реалізації синергетичного підходу, потрібно максимально уніфікувати вимірювальні перетворення, передовсім у вхідних аналогових колах вимірювального каналу.

Шляхом відповідного опрацювання одного виду біосигналу, для прикладу, ЕКГ розв'язують різнопланові задачі (рис. 1,б) – медична діагностика (виявлення захворювання), біометрика (ідентифікація людини), клінічний моніторинг (цілодобове спостереження за роботою серця).

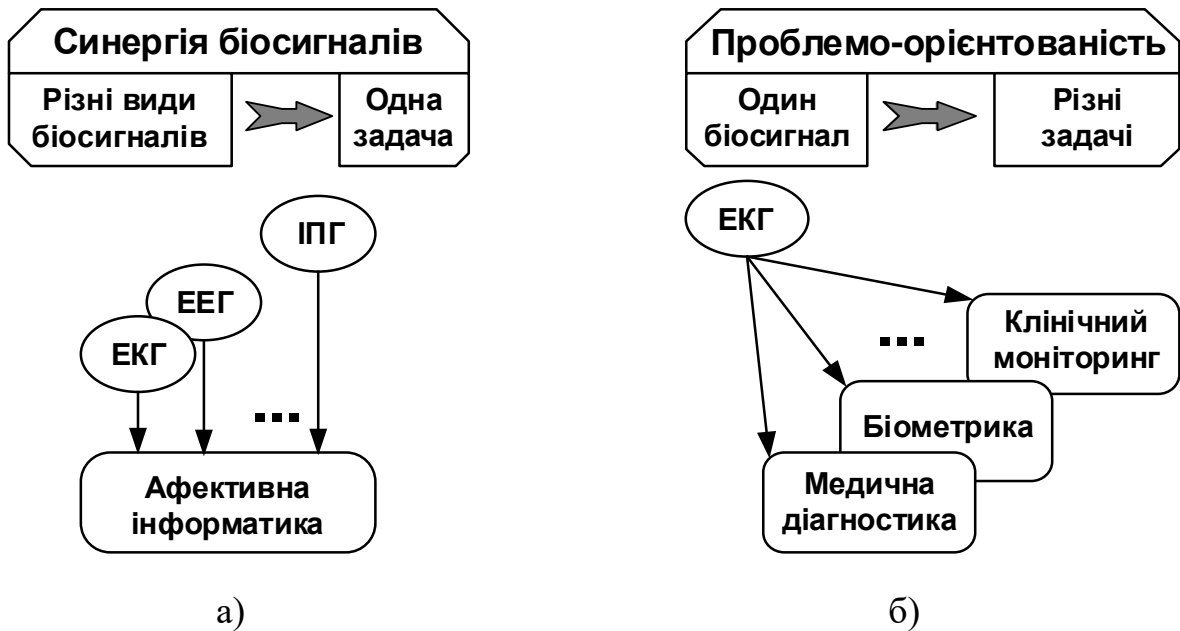


Рис. 1. Контroversійні підходи до опрацювання біосигналів: а – синергетичний, б – багатоцільовий

На відміну від попереднього випадку, акцент вирішення проблеми багатоцільового (Multi-Task) підходу зміщується із вхідних кіл біосенсорів, на вихід системи комп'ютерного опрацювання. Тобто стосується витягування інформативних ознак й інтерпретації результатів. Оскільки для багатоцільового підходу важливо забезпечити гнучкість переналаштування системи із однієї задачі на іншу, то для цієї мети виглядає доцільним застосувати технології машинного, зокрема глибокого навчання.

На основі виконаного аналізу сформульовано науково-прикладну проблему та завдання дисертаційного дослідження.

Другий розділ присвячено розвитку теоретичних засад комп'ютерного опрацювання біомедицинської інформації та удосконаленню концепції побудови комп'ютерної біоінформатичної системи із покращеними характеристиками.

Розкрито сутність етапів перетворення і опрацювання біосигналів у різних біомедицинських застосуваннях та охарактеризовано фактори, які ускладнюють це завдання та спонукають застосовувати машинне навчання:

- складність і різноплановість задач біоінформатики;
- природна мінливість біосигналів,
- специфічні спотворення біологічного походження (інтерференція сторонніх біопроектів, екранування корисного біосигналу).

Відомі підходи комп'ютерного опрацювання біосигналів (рис. 2,а) спираються на низку перетворень (переважно на основі алгоритмів ЦОС). Істотними обмеженнями таких підходів є специфічність методів і засобів як щодо виду біосигналу на первинних етапах опрацювання, так і щодо вирішуваної задачі біоінформатики на етапах формування дескрипторів для класичних алгоритмів МН. Це ускладнює реалізацію і синергетичного, і проблемно-орієнтованого підходів.

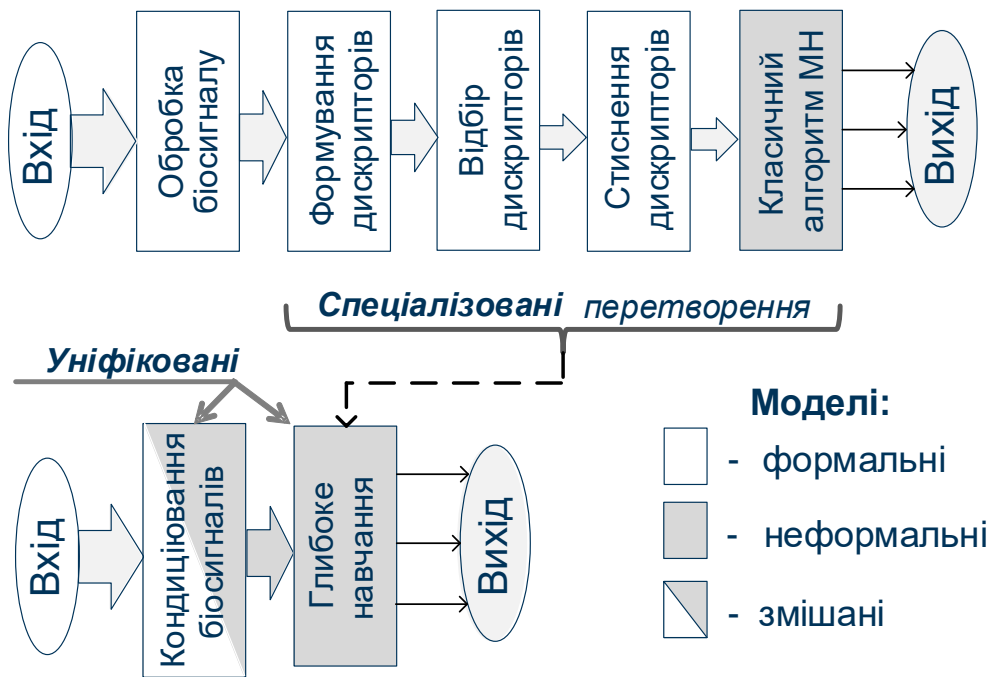


Рис. 2. Підходи до комп'ютерного опрацювання біосигналів: відомі (вгорі), пропонувані (знизу)

У роботі запропоновано розширити рамки теорії комп'ютерного опрацювання біосигналів за рахунок:

- застосування технологій глибинного навчання для інтелектуального аналізу біосигналів;
- поєднання неформальних методів машинного навчання із формальними методами цифрового опрацювання сигналів на різних етапах опрацювання біосигналів;
- мінімізації вимірювальних перетворень, що виконуються у аналоговому вигляді.

Глибинне навчання кардинально змінює парадигму комп'ютерного опрацювання біосигналів. Оскільки на вхід багатосарової штучної нейронної мережі (ШНМ), що виконує роль класифікатора, подається повний релевантний біосигнал, тому відпадає потреба у застосуванні формальних моделей для попереднього виділення інформативних ознак, їх відбору і стиснення, тобто скорочується число етапів перетворення біосигналів (рис. 2,б). Більше того, заміна вирішуваної задачі біоінформатики вимагає лише перенавчання нейромережевого класифікатора на основі розмічених даних, які відповідають цьому завданню. Цим досягається інваріантність архітектури нейронної мережі до задач біоінформатики.

Для реалізації глибинного навчання актуальним є:

- оптимізація гіперпараметрів та архітектури багатосарової штучної нейронної мережі;
- формування вхідних даних шляхом кондиціонування біосигналу;
- великий обсяг даних для синтезу нечіткої нейромережевої моделі.

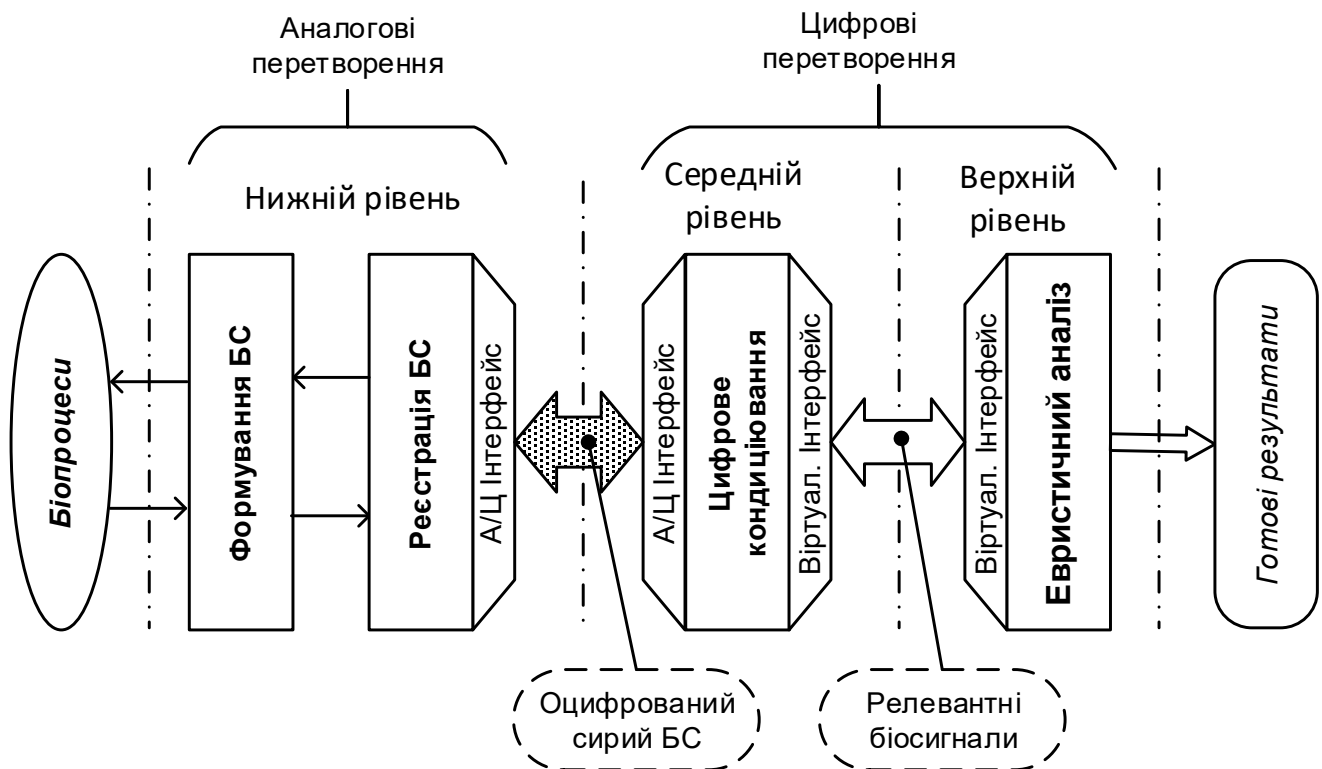


Рис. 3. Системні рівні та міжрівневі інтерфейси

У результаті теоретичних узагальнень дисертаційних досліджень:

- одержала подальший розвиток теорія комп'ютерного опрацювання біомедичної інформації, в рамках якої обґрунтовано широке, комплексне використання на всіх структурних рівнях штучних нейронних мереж і технологій глибинного навчання у поєднанні із методами ЦОС, що послужило універсальною платформою для розширення функціональних можливостей і покращення характеристик комп'ютерних систем для ефективного вирішення різнопланових завдань у сфері біоінформатики;

- запропоновано концепцію трьох системних рівнів комп'ютерного опрацювання біосигналів, в основу якої покладено ідею чіткого розмежування функцій системних рівнів від методів і засобів, що застосовуються для реалізації цих функцій (рис. 3); таке розділення сприяє структуризації знань у предметній галузі, уможливорює зіставлення і оцінювання ефективності, а також вибір кращих рішень на етапі проектування біоінформатичних систем під конкретні завдання із урахуванням умов і сценаріїв їх використання.

Цю концепцію реалізовано у вигляді програмно-апаратного каркасу (англ. framework), а також обґрунтовано вибір архітектури обчислювальних засобів за використання машинного навчання для різних задач біоінформатики і сценаріїв їх розгортання. Також проведено оцінювання придатності до об'єднання функцій структурних елементів системи комп'ютерного опрацювання біосигналів (див. табл. 1).

Досліджено особливості програмно-апаратного забезпечення алгоритмів навчання глибоких ШНМ. Проведено порівняльний аналіз програмних фреймворків для глибинного навчання, показано особливості імплементації цих

алгоритмів на основі графічних процесорів та оцінено перспективи реалізації алгоритмів глибинного навчання на базі програмованих логічних матриць.

Таблиця 1. Придатність до об'єднання функцій структурних елементів системи

Функціональний модуль	Системний рівень	Оцінка
Електроди, сенсори, актуатори	Нижній рівень	Індивідуальні апаратні модулі
Інструментальний підсилювач, трансїмпедансний підсилювач, мостова вимірювальна схема і т.д., формувач фотостимулу, генератор струму Хавленда і т.д.		
Аналоговий кондиціонер сигналу		Групові апаратні модулі
АЦП, ЦАП		
Апаратні прискорювачі: DDS, FFT, DPD і т.п.	Середній рівень	Групові апаратно-програмні модулі
Нейромережевий еквалайзер, автоенкодер		Групові програмні модулі
Програмні модулі: хвилькова декомпозиція, автокореляція, виявлення R-піків, сегментація, амплітудна і часова нормалізація		Спільні програмні модулі
Формальні алгоритми, алгоритми машинного навчання, штучні нейронні мережі	Верхній рівень	Спільні програмні модулі

Одним з ключових аспектів побудови комп'ютерних систем опрацювання біосигналів є питання ефективної імплементації алгоритмів ЦОС і МН на обчислювальні платформи. При виборі архітектури системи і обчислювальної платформи важливо врахувати такі фактори:

- характер та можливі специфічні сценарії прикладної задачі, а також реальні умови використання системи;

- особливості та технічні характеристики існуючих обчислювальних компонент, передовсім, з погляду ефективної імплементації алгоритмів МН.

Запропоновано три архітектури систем комп'ютерного опрацювання біосигналів, що покривають вимоги як типових біомедичних задач, так і специфічних сценаріїв сучасної біоінформатики: стаціонарну (Stand-Alone), хмарну (Cloud) та носиму (Wearable).

З огляду на складність і багатоваріантність структур комп'ютерних біоінформатичних систем результати дисертаційних досліджень представлено за трьома напрямками - вид біосигналу, тип завдання, системний рівень. Це спрощує аналіз змісту опрацювання біомедичних сигналів і даних та дає змогу розкрити переваги розробленої концепції побудови комп'ютерних систем для вирішення різнопланових завдань біоінформатики.

В третьому розділі розглянуто питання підвищення точності та обчислювальної ефективності методів і засобів первинного опрацювання біосигналів. Проаналізовано особливості відбору електричних біопотенціалів на тлі перешкоджаючих факторів, які маскують інформативні компоненти. Показано особливості вимірювання біоімпеданса для таких важливих задач біоінформатики, як дослідження кровонаповнення окремих органів тіла методом імпедансної плетизмографії у медичній діагностиці, визначення складу тіла (процентного вмісту жирових клітин, міжклітинної та внутрішньоклітинної рідини) у клінічному моніторингу, контроль електродермальної активності, що проявляється, зокрема, як зміна імпедансу шкіри і використовується в одному із каналів поліграфа у афективній інформатиці.

У цьому розділі реалізовано концепцію побудови комп'ютерної біоінформатичної системи у частині мінімізації та уніфікації аналогової частини вимірювального каналу. Розширену функціональність й інваріантність первинного вимірювального перетворення біосигналів досягнуто завдяки швидкодіючому багаторозрядному аналого-цифровому перетворювачу.

У цьому розділі представлено дослідження, які містять такі нові науково-практичні результати:

- модифікована схема джерела струму Хавленда, що уможливорює компенсацію базового біоімпеданса і розширює функціональні можливості реографічних досліджень;

- новий високочутливий первинний перетворювач біоімпедансу для системи контролю електродермальної активності в задачах афективної інформатики;

- новий спосіб реалізації цифрового фазоквадратурного перетворення сигналів імпедансної спектроскопії, який на відміну від відомих відрізняється високою обчислювальною ефективністю завдяки заміні операцій акумуляції добутків арифметичного перемноження вибірок аналізованої напруги і опорних сигналів синуса і косинуса на просте підсумовування вибірок аналізованої напруги упродовж інтервалів, які визначаються логічними операціями «&» від функції знаку опорних сигналів.

Щоб використовувати генератор струму на різних діапазонах вимірювання біоімпедансу (напр., 100, 200, 500 чи 1000 Ом) зберігаючи високі параметри вихідного імпеданса, застосовано буферний повторювач БП напруги для зменшення впливу на резистори у колі додатного зворотного зв'язку операційного підсилювача ОП (рис. 4).

Спад напруги, зумовлений протіканням струму через біоімпеданс, вимірюється за допомогою інструментального підсилювача ПІ2

$$U_B = I_L \times Z_B = Z_B \times U_P / R_5. \quad (1)$$

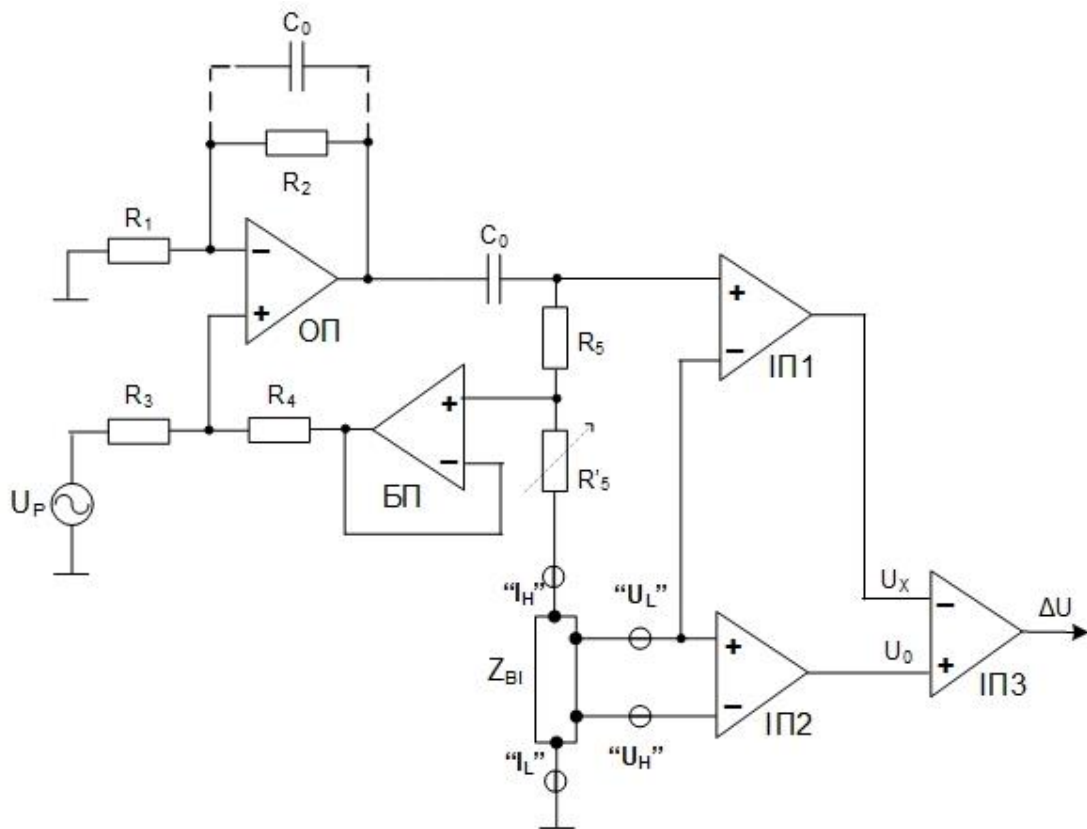


Рис. 4. Багатодіапазонний перетворювач біоімпедансу на основі генератора струму Хавленда для реографічної системи

Аби виокремити на тлі домінуючого базового біоімпеданса Z_B його пульсуючу складову $\Delta Z(t) \ll Z_B$, що відображає кровонаповнення досліджуваної ділянки тіла, застосовано принцип компенсації напруги U_B . Для цього за допомогою інструментального підсилювача ІП₁ контролюють спад напруги U_K одночасно на двох резисторах - резисторі піддіапазону R_5 і кодкерованому резисторі (цифровому тримері) $R_{ЦТ}$.

Якщо опір резистора піддіапазону вибирають спираючись на апріорні відомості про орієнтовне значення базового біоімпедансу досліджуваної ділянки тіла, то налаштування цифрового тримера відбувається шляхом низки вимірювальних перетворень у каналі реографа. Інформативний сигнал (реограма) формується на виході інструментального підсилювача ІП₃ як різниця напруг на ІП₂ і ІП₁:

$$\Delta U = U_B - U_K = I_L \times Z_B - I_L \times (r + R_T) = I_L \times \Delta Z(t). \quad (2)$$

Конденсатор C_k малої ємності запобігає втраті стійкості генератора Хавленда, у якому використано не лише від'ємний, але і додатний зворотний зв'язок. Роздільний конденсатор C_p запобігає проходженню постійного струму через біооб'єкт. Для компенсації частотних похибок, які виникають через неідеальність параметрів застосованих операційних підсилювачів, можна застосувати принцип вирівнювання частотної характеристики за допомогою цифрового еквалайзера.

Для виділення слабкого корисного сигналу на тлі інтенсивних завад застосовують синхронні (фазочутливі) детектори (Lock-In Amplifier). Відомі

цифрові реалізації фазочутливих детекторів вимагають виконання $2 \times N$ перемножень і $2 \times N$ додавань на одну вибірку біоімеданса. Ці арифметичні операції можуть бути реалізовані на сигнальних процесорах. Задля підвищення обчислювальної ефективності та розширення частотного діапазону запропоновано нову структуру цифрового фазочутливого перетворювача (рис. 5).

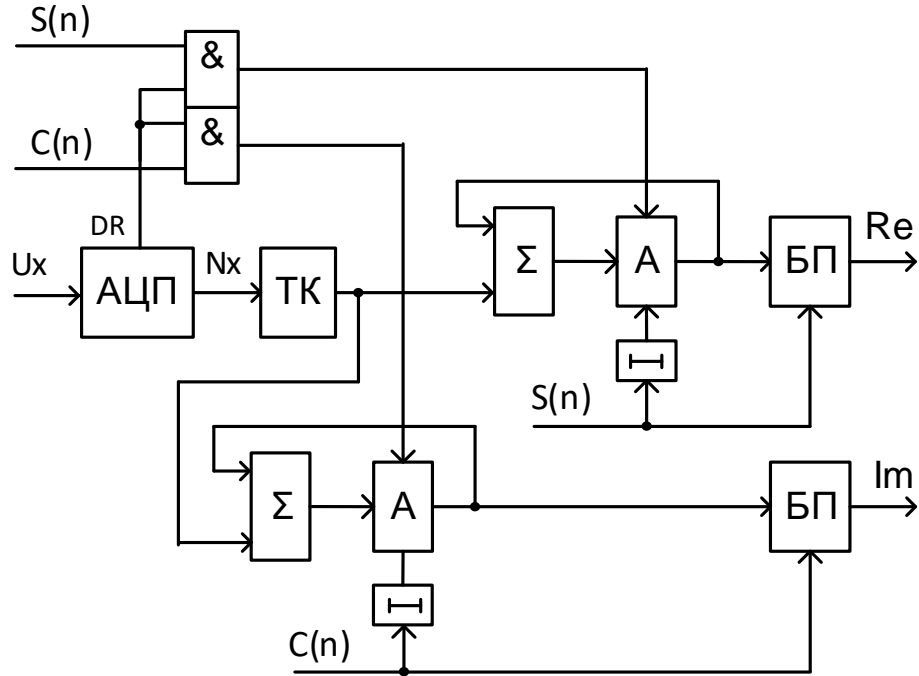


Рис. 5. Структура обчислювально-ефективного цифрового фазоквадратурного перетворювача

Суть нового способу реалізації фазоквадратурного перетворення полягає у відмові перемноження оцифрованих вибірок напруги $U_X(n)$ із миттєвими значеннями ортогональних опорних сигналів $\sin(n)$ і $\cos(n)$ на користь їх цифрового інтегрування упродовж точно визначених інтервалів часу:

$$M_{\text{Re}} = \sum_{n=0}^{K \times N - 1} U_X(n) \times S(n) = \sum_{n=0}^{K \times N - 1} U_X(n) \times \text{sign}[\sin(n)] \quad (3a)$$

$$M_{\text{Im}} = \sum_{n=0}^{K \times N - 1} U_X(n) \times C(n) = \sum_{n=0}^{K \times N - 1} U_X(n) \times \text{sign}[\cos(n)] \quad (36)$$

де $N = F_S / F_O$ – число вибірок сигналу за один період, K – ціле число періодів аналізованого сигналу.

Цифровий інтегратор збудовано на основі суматора Σ і акумулятора A , ємність якого має становити:

$$M_A = m + \log_2 N, \quad (4)$$

де m – розрядність (число біт) аналого-цифрового перетворювача.

Щоб процес нагромадження вибірок у акумуляторі A відбувався коректно, застосовано транскодер TK , завданням якого є перетворення вихідного коду АЦП до формату доповнення до 2. Якщо дані на виході АЦП представлено у форматі

зміщеного коду, то функція транскодера зводиться до інвертування найстаршого біту (біт знаку).

Інтервали часу цифрового інтегрування (нагромадження вибірок) задаються керівними сигналами $S(n)$ і $C(n)$, які формуються як функція знаку опорних сигналів $\sin(n)$ і $\cos(n)$ генератора прямого цифрового синтезу. По завершенню часу інтегрування старші розряди акумуляторів M_{Re} і M_{Im} переписуються до буферів результату БР і відображають, наприклад, дійсну і уявну складові біоімпеданса.

Існує можливість ефективного підвищення завадостійкості цифрового фазоквадратурного перетворювача шляхом збільшення часу цифрового інтегрування. З цією метою застосовують двійкові лічильники у колах керування онуленням акумуляторів і запису результатів у буфери результату, тому час інтегрування збільшується

$$\tau_{DI} = \frac{K}{F_0} = \frac{K \times N}{F_S} \quad (5)$$

де $K = 2^i$ ($i=1,2,3,\dots$) - коефіцієнт ділення двійкового лічильника, який визначає ціле число періодів аналізованого сигналу.

Це вимагає збільшити ємність акумуляторів до

$$M_A = m + \log_2 N + \log_2 K \quad (6)$$

Дослідження показали, коли добуток $K \times N$ перевищує 1024, відношення «сигнал/шум» на виході цифрового синхронного детектора покращується понад 60 дБ.

У четвертому розділі представлено науково-прикладні результати, які стосуються застосування глибоких нейронних мереж для комп'ютерного опрацювання ЕКГ-сигналів у біометричних системах та ЕЕГ-сигналів у нейрокомп'ютерному інтерфейсі управління роботом.

У цьому контексті досліджено особливості електрокардіограми як біометричного маркера, представлено структуру і функціонування комп'ютерної системи біометричної ідентифікації за сигналом ЕКГ (рис. 6).

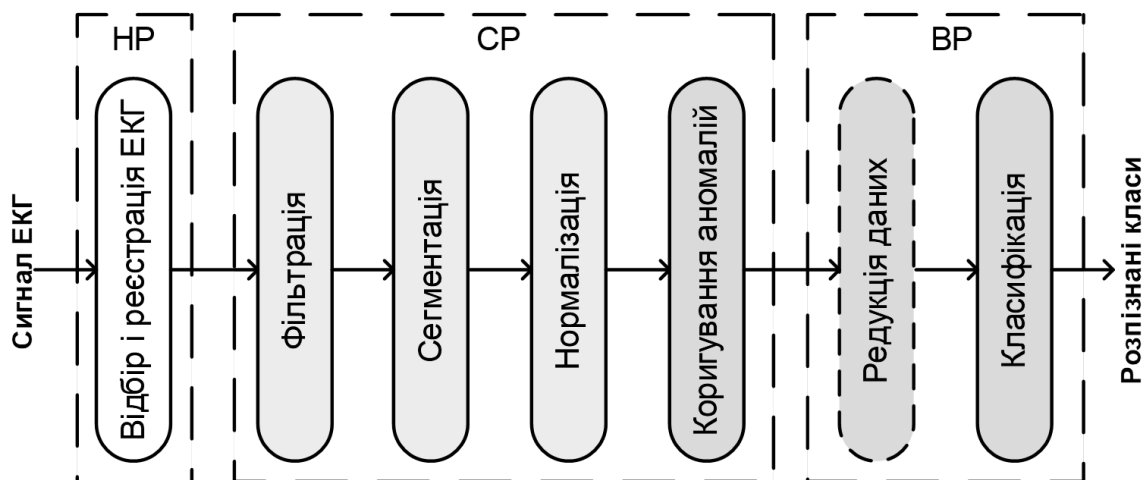


Рис. 6. Комп'ютерна система біометричної ідентифікації за сигналом ЕКГ на основі глибокого навчання: НР, СР і ВР –відповідно нижній, середній і верхній рівні

Відбір електрокардіограм відбувається із пальців правої і лівої руки (спрощена схема I відведення Ейнтговена). Сирий оцифрований ЕКГ-сигнал із нижнього надходить на середній рівень, де відбувається його кондиціонування за допомогою алгоритмів ЦОС (фільтрація, сегментування і нормалізація за амплітудою і тривалістю). Обмежена ефективність формальних алгоритмів ЦОС зумовлює потребу застосування інструментів глибинного навчання для коригування залишкових аномалій. На верхньому рівні багатосаровий нейрокласифікатор розпізнає релевантні сегменти ЕКГ, що відповідають окремому серцебиттю, відносячи їх до одного із класів, сформованих на етапі навчання.

Виконано комплексне дослідження розробленої біометричної системи ідентифікації як на основі власне створеної бази записів електрокардіограм LBDS, так і з використанням таких доступних баз як Physionet ECG-ID, Physionet MIT-BIH, Physionet QT. Результати дослідження дають підстави стверджувати, що ЕКГ-сигнал є надійним біометричним маркером, дуже стійким до апаратних варіацій, наявності шуму та артефактів, стабільним у часі, масштабованим щодо досить значного числа користувачів (>90). Система забезпечує високу швидкість розпізнавання (порядку 5 сек.) і точність ідентифікації 95-98%.

Проста апаратура і спосіб відбору електрокардіограми позначаються на дуже низькій якості записів. Тому для виявлення і коригування аномалій в сигналах ЕКГ запропоновано застосувати нечіткі моделі спеціального виду багатосарових ШНМ – автоенкодерів. На рис. 7 наведено дві моделі розроблених автоенкодерів.

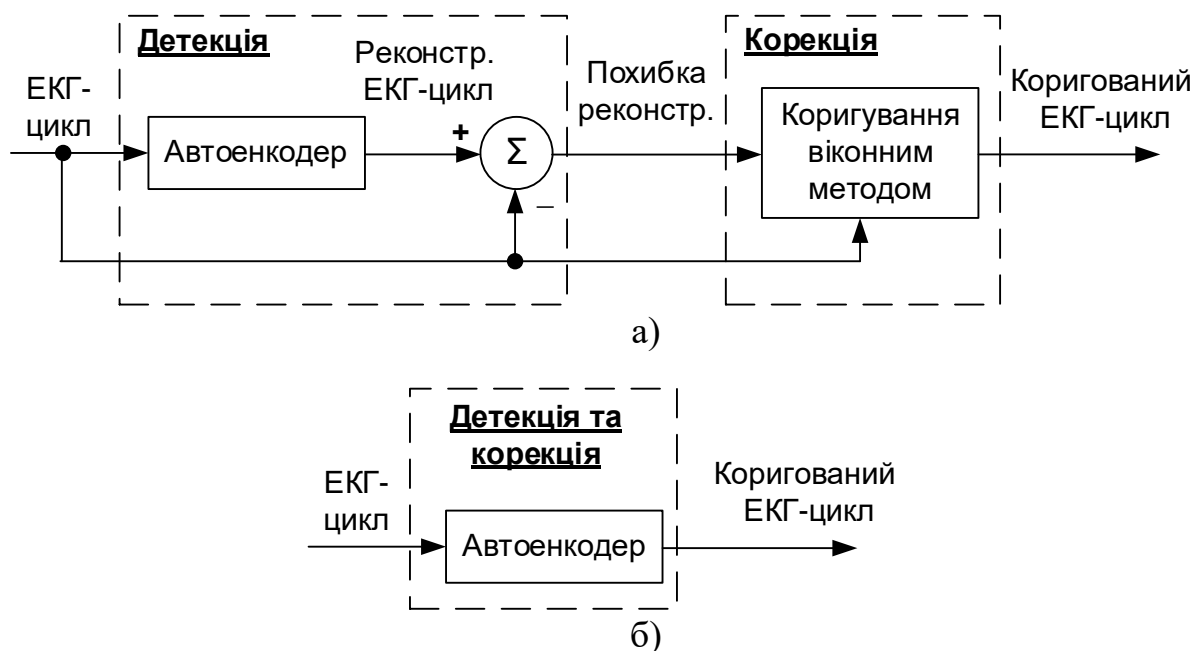


Рис. 7. Моделі автоенкодерів для виявлення і коригування аномалій у сигналах ЕКГ

Перша з них (рис. 7, а) функціонує так:

1. Модель навчається компресувати та реконструювати сигнал ЕКГ

$$r = g(h) = \sigma'(W'\sigma(Wx + b) + b') \quad (7)$$

де W та b – матриця ваг та вектор зміщення моделі, σ – функція активації, x та r – вхідний та реконструйований сигнали.

2. Різниця між вхідним x та реконструйованим r сигналами в межах деякого ковзного вікна використовується для виявлення аномалій

$$L(x, r) = \|x - r\|^2 = \|x - \sigma'(W' \sigma(Wx + b) + b')\|^2 \quad (8)$$

3. Усі аномальні вибірки замінюються відповідними значеннями реконструйованого сигналу.

Основна відмінність другої моделі автоенкодера (рис. 7, б) полягає у використанні реконструйованого сигналу для подальшої ідентифікації замість аналізу та обробки помилки реконструкції. Таким чином, фаза детекції промахів не представлена як окрема трансформація, оскільки всі промахи коригуються на етапі реконструкції вхідного сигналу автоенкодера.

Рис. 8 ілюструє вигляд сегментів ЕКГ на вході та виході автоенкодера.

Достовірність роботи багат шарових нейронних мереж залежить не лише від релевантності і репрезентативності навчальних наборів даних, але також є чутливою до значень гіперпараметрів, тобто сталих параметрів моделі, що використовується для керування процесом навчання (напр. число шарів та нейронів, тип активаційних функцій, коефіцієнт регуляризації та ін.). Тому однією з проблем підвищення ефективності машинного навчання є пошук та вибір оптимальних значень гіперпараметрів, причому ці значення залежить як від структури аналізованих даних, так і від завдань класифікації.

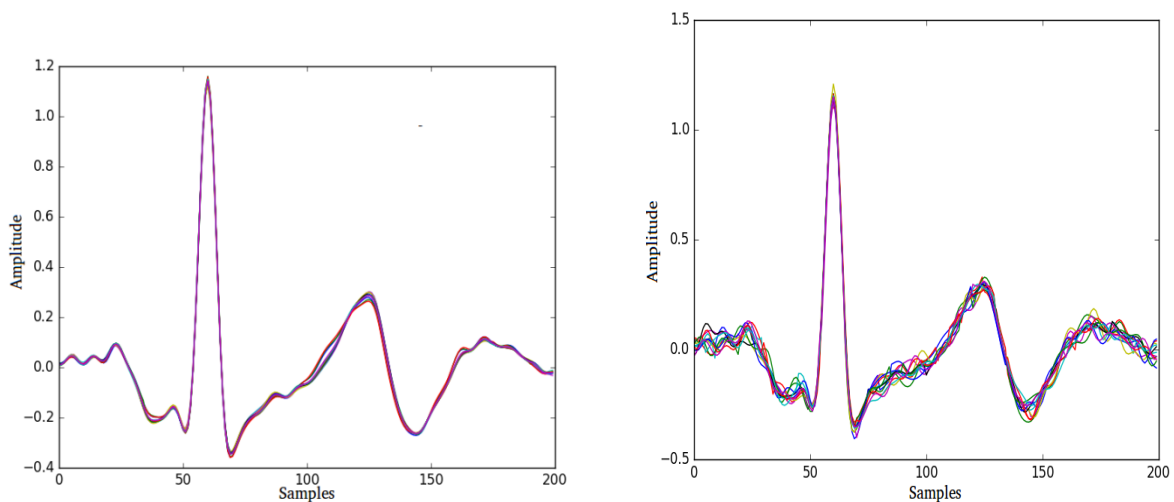


Рис. 8. Результати коригування автоенкодером промахів: серцеві цикли ЕКГ до (а) та після (б) коригування

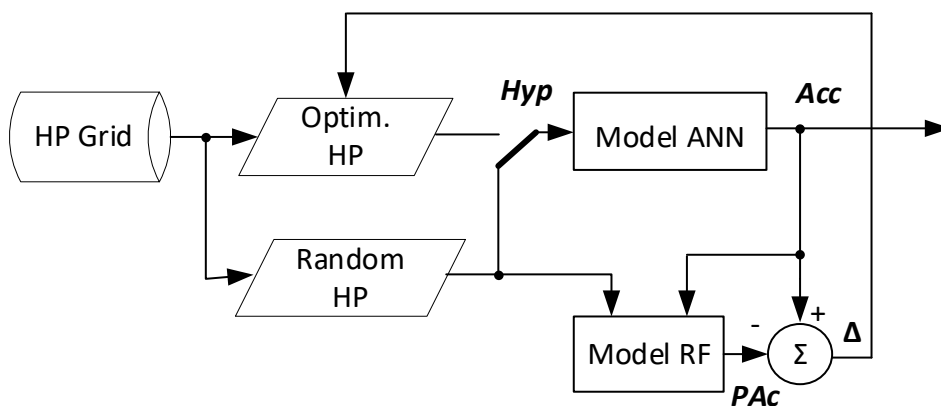
Повний перебір можливих значень гіперпараметрів з метою пошуку кращих є нездійсненний через обчислювальну складність. Тому запропоновано автоматизацію пошуку оптимальних значень гіперпараметрів нейромережевого класифікатора системи біометричної ідентифікації. Особливістю цього підходу є використання прогностичної моделі збудованої за допомогою обчислювально простого алгоритму машинного навчання (рис. 9,а).

Процес автоматичного пошуку оптимальних гіперпараметрів класифікатора на основі багат шарової ШНМ складається із таких етапів:

1. Підготовка фіксованої сітки значень гіперпараметрів (HP Grid).
2. Рандомізація вибірок гіперпараметрів (Random HP) - із сітки (простору) десятків тисяч можливих значень вибрано сотні комбінацій.
3. Навчання та тестування штучної нейронної мережі (Model ANN), коли на її вхід подаються вибрані значення гіперпараметри (H_{up}), а на виході оцінюється точність класифікації (Acc).
4. Навчання з використанням так одержаних наборів даних H_{up} і Acc простої моделі машинного навчання, наприклад, збудованої за алгоритмом випадкового лісу (Random Forest).
5. Швидке прогнозування простою моделлю із п.4 очікуваної точності PAc для всіх можливих значень гіперпараметрів (десятки тисяч).
6. Вибір найбільш перспективних гіперпараметрів за критерієм найвищої очікуваної точності PAc .
7. Повторення етапу 3 для перевірки ефективності методу автоматизованого пошуку з використанням результатів етапу 6 – оцінюється розбіжність очікуваної точності та її реального значення $\Delta = |PAc - Acc|$.

Як видно з наведеної гістограми (рис. 9,б) для переважної більшості випадків розбіжність актуальної точності моделі і прогнозованих результатів не перевищує 5%. Крім того, запропонований підхід на 4 порядки скорочує час пошуку порівняно із повним перебором в просторі можливих значень. Це означає можливість ефективного вибору оптимальних значень гіперпараметрів нечітких моделей багат шарових нейромереж для різних типів біосигналів та завдань досліджень.

Крім того, у цьому розділі розроблено програмну архітектуру та проведено порівняльний аналіз двох варіантів реалізацій інтерфейсу «мозок-комп'ютер» для керування колісним роботом на базі технологій глибинного навчання, а саме згорткової нейронної мережі і багат шарового перцептрона. Нейрокласифікатор на згортковій мережі забезпечує вищу точність $\sim 94\%$, але в аспекті обчислень є складнішим, а час його навчання майже втричі триваліший порівняно із багат шаровим перцептроном.



a)

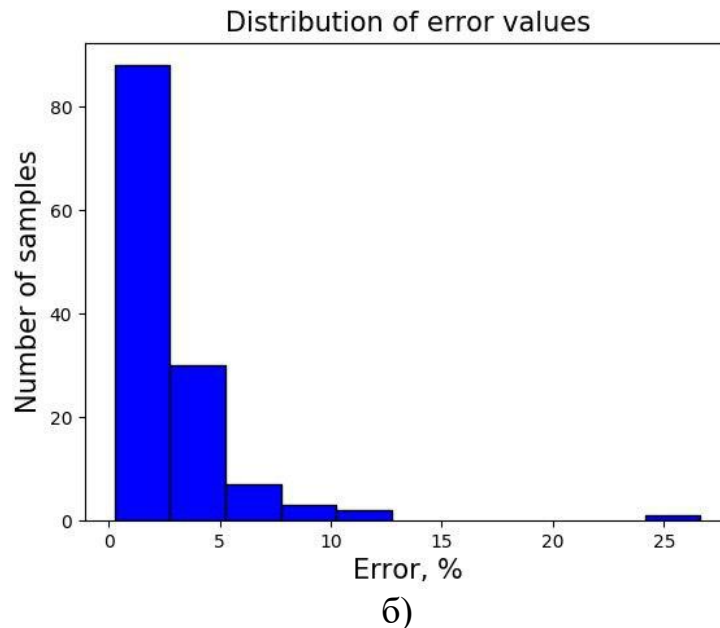


Рис. 9. Модель автоматичного пошуку оптимальних гіперпараметрів нейромережевого класифікатора (а) та гістограма розподілу похибки прогнозованої і реальної точності ідентифікації (б)

В п'ятому розділі наведено результати досліджень, які стосуються застосування технологій машинного навчання у галузі медичної діагностики на прикладах таких двох прикладах:

- суміщення технологій цифрового опрацювання сигналів і машинного навчання для встановлення ступеня хрящових розладів в колінному суглобі за загальноприйнятими діагностичними критеріями;
- використання алгоритмів машинного навчання для побудови прогностичних моделей на базі даних електронних медичних записів (EMR).

Віброартроскопія є новим і перспективним методом оцінювання ступеня дизфункції суглобу, що базується на реєстрації і аналізі віброакустичних сигналів під час його згинання/розгинання. Подано короткі відомості про використану у дослідженнях базу віброартрографічних (ВАГ) сигналів колінного суглобу. Основним обмеженням у даному дослідженні був малий обсяг (по одному запису із 4-ма циклами згинання/розгинання для 105 різних пацієнтів, віднесених до 5-ти категорій, залежно від ступеня захворювання). Це унеможливило застосування технологій глибокого навчання.

ВАГ-сигнали мають нестационарний характер, тому для їх аналізу застосовано дискретне хвилькове перетворення, яке надає можливості вивчення не лише частотного вмісту сигналу, але і його локалізацію щодо часу. Аналіз скалограми 6-ти рівневого хвилькового розкладу дав змогу виділити смугу від 78 до 781 Гц, де зосереджена корисна діагностично значуща інформація (рис. 10). Реконструкція сигналу у вказаній смузі забезпечила усунення дестабілізуючих впливів дрейфу ізоляції та високочастотних шумів (рис. 11).

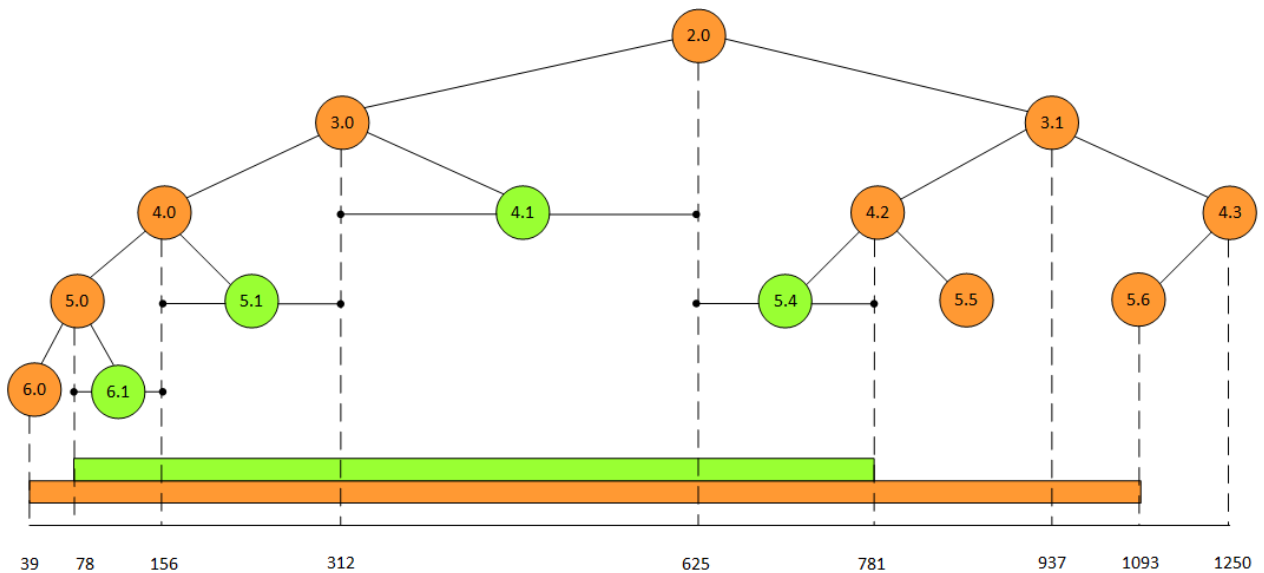


Рис. 10. Дерево хвилькової декомпозиції віброартрографічного сигналу (зеленим виділено область релевантного сигналу)

На рис. 11 показано вигляд сирого та відновленого релевантного сигналу, а також артефакти (дрейф ізолінії та високочастотні шуми).

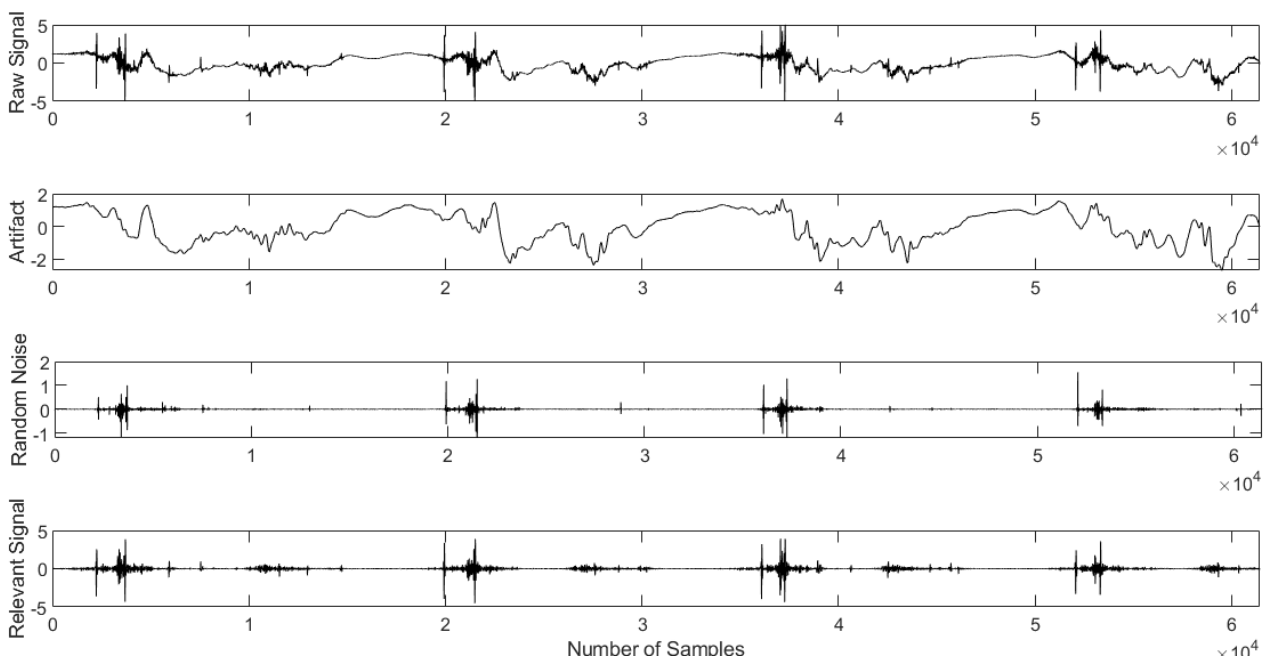


Рис. 11. Вигляд сирого віброартрографічного сигналу (а), дрейфу ізолінії (б), шумів (в) та релевантного сигналу (г): по вертикальній осі амплітуда у Вольтах

На основі статистичних параметрів коефіцієнтів хвилькового розкладу у вузлах хвилькової декомпозиції 4.1; 5.1; 5.4; 6.1 збудовано 12 дескрипторів, які концентровано відображають діагностичні ознаки (рис. 12). На основі таких дескрипторів, витягнутих із розмічених даних, натреновано класифікатор системи підтримки прийняття рішення, що визначає дизфункцію суглобу.

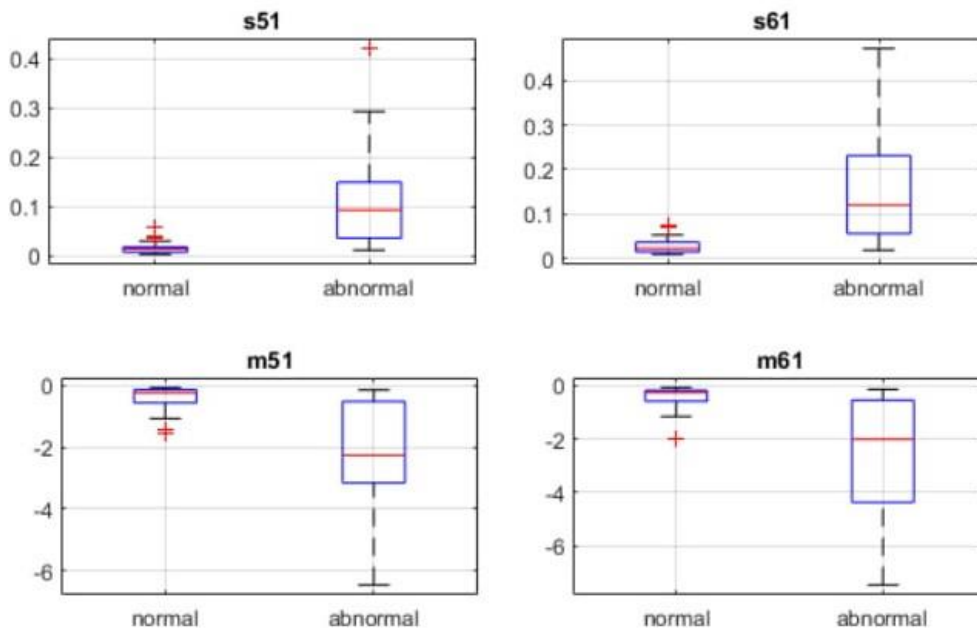


Рис. 12. Приклад статистичних оцінок деяких дескрипторів ВГА-сигналу

Виконано верифікацію інтелектуальної системи медичної діагностики у варіантах дво- та п'ятикласового розпізнавання даних із тестового набору. Результати верифікації для двох груп показали, що більшість метрик демонструють високі, або дуже високі результати, в той час як для п'яти груп результати класифікації були відносно посередні. Проте їх можна покращити за умови формування складніших дескрипторів та гнучкого їх налаштування у межах класу.

Також у цьому розділі розглянуто і апробовано метод інтелектуального аналізу електронних медичних реєстрів, шляхом побудови прогностичних моделей на основі низки алгоритмів машинного навчання, передовсім глибоких нейронних мереж. Розроблені методи застосовано для прогнозування ризику виникнення набряку Квінке (ангіоневротичний набряк) при медикаментозному лікуванні пацієнтів з епілепсією та прогнозування ефективності лікування і найбільш імовірних варіантів розвитку одного з найскладніших онкологічних захворювань – Негоджкінської лімфоми.

Результати досліджень показали придатність методів до аналізу ризику виникнення побічних ефектів і оцінювання ймовірної дієвості за різних стратегій лікування. Зокрема можна передбачати реакції пацієнтів на стандартне лікування і завчасно ідентифікувати резистивних пацієнтів, тобто осіб, організм яких, з високою імовірністю не продемонструє позитивної динаміки лікування за використання стандартного підходу.

Шостий розділ присвячено застосуванню технологій цифрового опрацювання сигналів і машинного навчання у задачах біоінформатики, що базуються на вимірюваннях біоімпеданса.

Динамічні похибки є серйозним бар'єром на шляху побудови малозатратного частотного аналізатора біоімпеданса. З огляду на це у роботі запропоновано інноваційний підхід до розширення діапазону робочих частот аналізатора біоімпеданса, який не потребує ні апаратної, ні часової надлишковості. Його суть

полягає у вирівнюванні частотної характеристики вимірювального тракту сенсора біоімпеданса шляхом програмного коригування впливу динамічних похибок. Реалізація цифрового еквалайзера вимагає лише залучення відповідної обчислювальної потужності, якою володіють сучасні вимірювальні пристрої.

Запропонований метод коригування впливу динамічних похибок використано для побудови сенсорної частини аналізатора біоімпеданса, структура якого наведена на рис. 13. Відмінною особливістю структури сенсора біоімпеданса порівняно з відомими рішеннями є одноканальний тракт вимірювання, що є важливо і зручно для побудови простих портативних пристроїв.

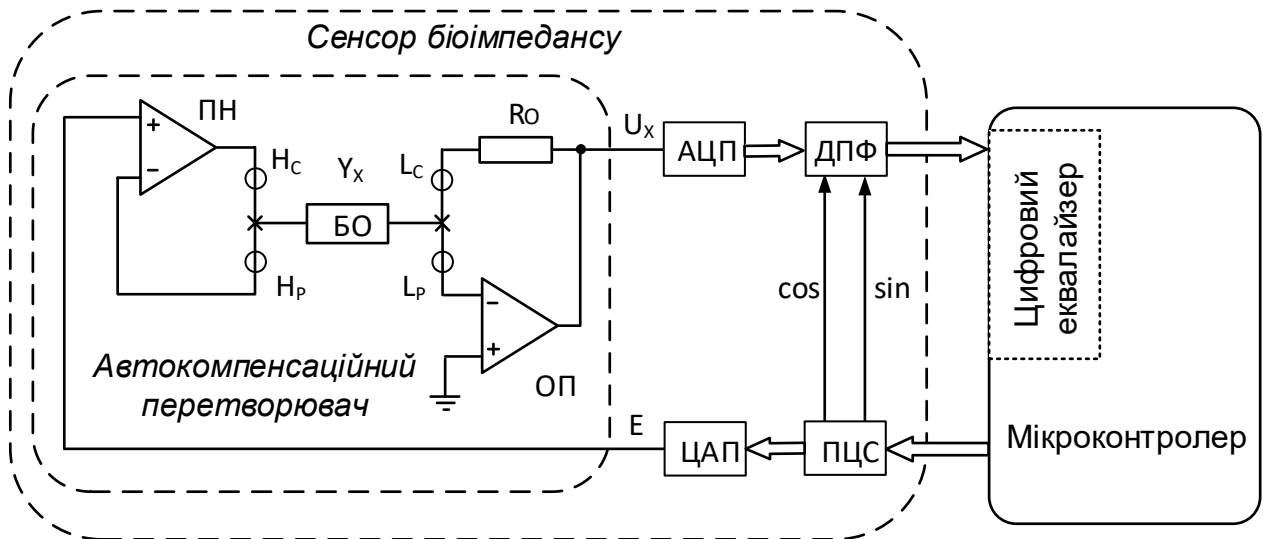


Рис. 13. Структура сенсора біоімпеданса на основі методу автобалансування

Рівняння перетворення сенсора біоімпеданса на основі методу автобалансування є нелінійним та доволі складним:

$$H = -\frac{U_x}{E} = -\frac{Y \cdot R_0 - \frac{Y \cdot R_{OUT}}{A}}{1 + \frac{1}{A} \left[\left(1 + Y \cdot R_0 + \frac{R_0}{Z_p} \right) \left(1 + \frac{R_{OUT}}{Z_L} \right) + Y \cdot R_{OUT} + \frac{R_{OUT}}{Z_p} \right]} = G_1 + jB_1 \quad (9)$$

де G_1 і B_1 – величини, що у першому наближенні відображають активну і реактивну складові адмітанса Y (оберненої до біоімпеданса Z).

Проте за певних допущень аналітичним способом можна синтезувати формальну модель цифрового еквалайзера для одержання відкоригованих значень складових G_2 і B_2 біоімпеданса:

$$G_2 = \frac{a \cdot c - b \cdot d}{a^2 + b^2}; \quad B_2 = \frac{b \cdot c + a \cdot d}{a^2 + b^2}, \quad (10)$$

$$\text{де } a = 1 + \frac{B_1}{k}; \quad b = \frac{G_1}{k}; \quad c = G_1 - G_1 \frac{C}{K} - \frac{B_1}{K}; \quad d = B_1 - B_1 \frac{C}{K} + \frac{G_1}{K}.$$

Розроблено альтернативний метод, який використовує концепцію машинного навчання для побудови нечіткої моделі нейромережевого еквалайзера. Процес

розроблення цифрового еквалайзера на базі ШНМ складається за таких етапів: підготовка даних, ініціалізація, навчання (калібрування), тестування (оцінювання точності).

На першому етапі підготовлено навчальний і тестовий набори даних. Обидва набори мають ідентичну структуру і складаються з пар вхідних та цільових (вихідних) векторів. Кожен вхідний вектор містить сім величин, що входять до рівняння перетворення (9): $X_1 = f$, $X_2 = P$, $X_3 = Q$, $X_4 = f_T$, $X_5 = C_{in}$, $X_6 = R_0$, $X_7 = R_{OUT}$. Кожен цільовий (вихідний) вектор містить дві величини G_2 і B_2 . Тренувальний набір включає в себе 10 пакетів по 10 тисяч зразків, а тестувальний набір – один пакет з 1 мільйона зразків.

Значення величин f , f_T , C_{in} , R_{OUT} , R_0 , R , X згенеровано за методом Монте-Карло випадковим чином за рівномірним законом розподілу в діапазонах допусків параметрів, взятих із документації операційного підсилювача та технічних характеристик проєктованого аналізатора біоімпеданса. Величини G_1 і B_1 розраховані відповідно до виразу (9).

На етапі ініціалізації створюється ШНМ відповідного типу та архітектури. Початкові значення коефіцієнтів – випадкові числа. Також встановлюються значення всіх параметрів необхідних для етапів навчання і тестування.

В ході експериментів використовувалася ШНМ із двох прихованих шарів по 50 нейронів із функцією активації сигмоїд. Результати наведено в табл. 2.

Таблиця 2. Порівняння результатів коригування частотних похибок сенсора біоімпеданса

	Довірча похибка ($P_{дов} = 0,997$), %		Гранична похибка, %	
	ШНМ-модель	Аналітична модель	ШНМ-модель	Аналітична модель
R	0,0022	0,0037	0,027	0,024
X	0,0016	0,0032	0,013	0,015

Дослідження показали, що частотні похибки найбільше спотворюють меншу за значенням ємнісну складову біоімпеданса. Ефективність нейромережевого еквалайзера ілюструють наведені на рис. 14 графіки частотної залежності ємнісної складової біоімпеданса без вирівнювання (згори) та з вирівнюванням частотного відгуку АКП (внизу). Також досліджено ефективність нейромережевого еквалайзера на прикладі реєстрації діаграм Коула.

Крім того, розроблено структуру портативного реографічного модуля із покращеними характеристиками, придатного до застосування не лише для медичної діагностики, але і у задачах афективної інформатики. Удосконалення структури вимірювального реографічного каналу досягнуто завдяки принципу компенсації базового біоімпеданса за допомогою кодированого потенціометра

(цифрового тримера), що входить до модифікованої схеми джерела струму Хавленда.

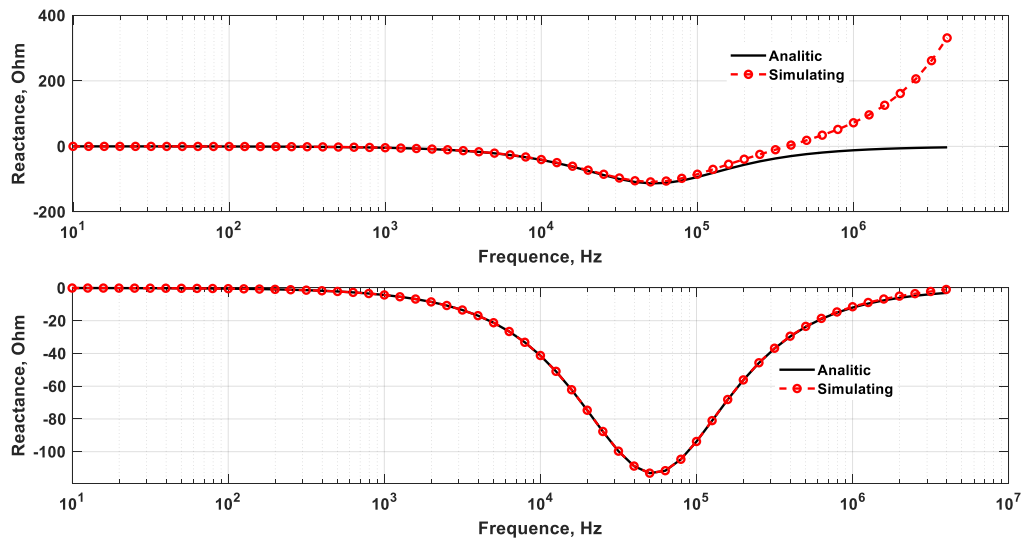


Рис. 14. Вимірювання реактивності біоімпеданса без вирівнювання частотного відгуку АКП (вгорі) та з вирівнюванням (внизу)

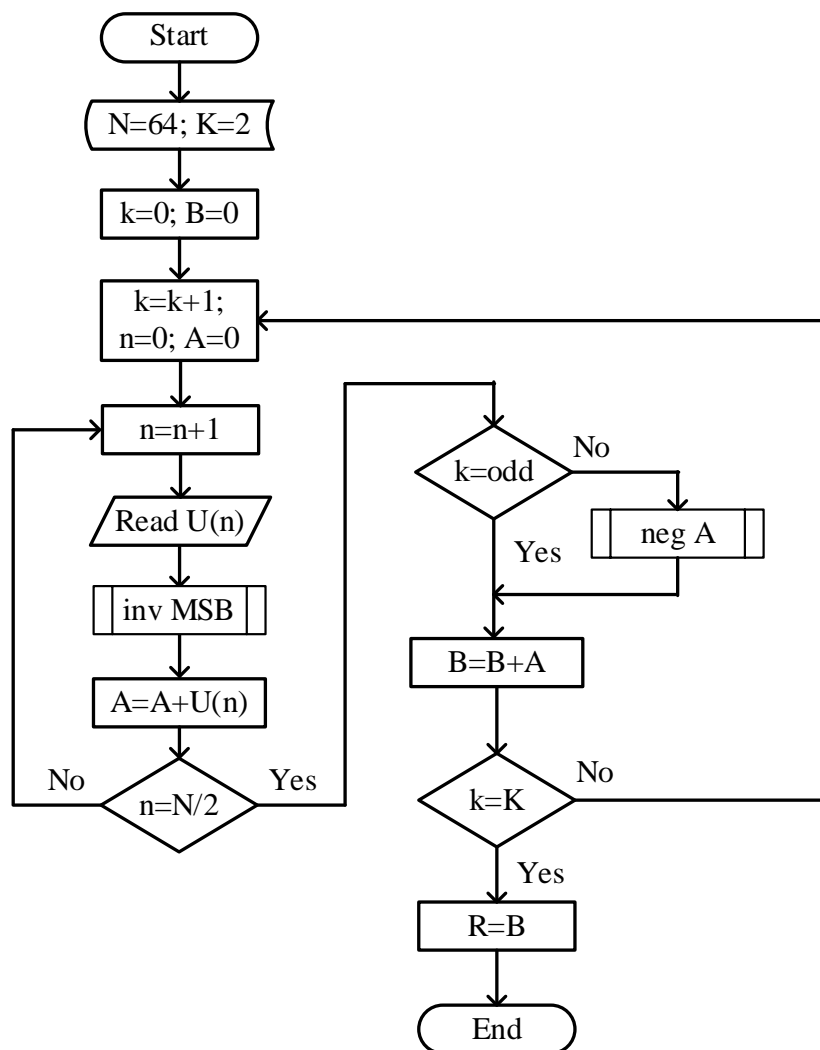


Рис. 15. Блок-схема алгоритму синхронного детектування

Важливою особливістю вимірювального тракту реографа є використання програмної реалізації синхронного детектування, реалізованого за інноваційним рішенням (рис. 15). Основними перевагами використання синхронного детектування в реографічних системах є збільшення завадостійкості і виокремлення пульсацій активного опору (а не імпедансу, як у існуючих реографах).

Для покращення характеристик реографа також використано сформовані на засадах прямого цифрового синтезу ортогональні зондувальні сигнали із точним підтриманням квадратури. Це дало можливість розширити діагностичну здатність реографічних досліджень і удвічі підвищити швидкодію, завдяки паралельній реєстрації реограм, одержаних квадратурними каналами на одній і тій самій частоті із симетричних ділянок тіла.

У **висновках** сформульовано основні результати дисертаційних досліджень.

Додатки містять акти впровадження результатів роботи, опис програмного забезпечення комп'ютерної системи біометричної ідентифікації на основі електрокардіограми, коди програм у пакетах MatLab і Python для математичного моделювання ключових структурних елементів вимірювального каналу.

ВИСНОВКИ

У дисертаційній роботі вирішено актуальну науково-прикладну проблему у галузі інструментального забезпечення біоінформатики – створення теоретичних засад для удосконалення методологічної, алгоритмічної і програмно-технічної бази комп'ютерних систем опрацювання біосигналів і даних на основі широкого використання штучних нейронних мереж та технологій глибинного навчання. Основні наукові і практичні результати полягають у такому.

1. Проаналізовано сучасний стан та перспективи розвитку комп'ютерних систем аналізу і опрацювання біомедичних даних, що дало змогу окреслити невирішені завдання, сформулювати науково-прикладну проблему та обґрунтувати доцільні підходи щодо її вирішення, передовсім оцінити ступінь придатності різних методів штучного інтелекту та глибинного навчання до застосування у комп'ютерних біоінформатичних системах.

2. Розвинуто науково-методологічну базу комп'ютерного опрацювання біосигналів із урахуванням потреб практики, шляхом широкого використання технологій машинного навчання, а також удосконалення методів цифрового опрацювання біосигналів.

3. Обґрунтовано узагальнену структуру комп'ютерної вимірювальної системи для задач біоінформатики на підставі аналізу змісту етапів перетворення біосигналів у різних застосуваннях із урахуванням факторів, що ускладнюють видобування біомедичних показників та інтерпретацію результатів досліджень, а також на основі комплексного критичного аналізу актуальних вимог до комп'ютерного опрацювання біосигналів та оцінювання можливостей сучасних програмно-апаратних засобів.

4. Удосконалено концепцію побудови комп'ютерних біоінформатичних систем, яка передбачає ієрархічний і модульний принцип імплементації розроблених методів аналогового перетворення і цифрового опрацювання

біосигналів, а також алгоритмів машинного навчання з урахуванням особливостей сучасних програмно-апаратних засобів, специфіки і складності обчислювальних задач (зокрема і на етапі навчання), забезпечуючи гнучкість і раціональний розподіл програмно-апаратних ресурсів, масштабованість системи, можливість модифікації її обчислювальних компонент, а також підвищення надійності і спрощення верифікації системи.

5. Імплементовано розроблену концепцію у відповідний фреймворк, застосування якого пришвидшує процес проектування комп'ютерної системи біоінформатики, забезпечує ефективну реалізацію запропонованих методів комп'ютерного опрацювання біосигналів на різних обчислювальних платформах, а також гнучкість налаштування на різні задачі біоінформатики та вищий рівень інформативності шляхом злиття біосенсорів.

6. У зв'язку зі складністю і багатоваріантністю структур комп'ютерних систем опрацювання біосигналів репрезентацію розроблених моделей і методів запропоновано здійснювати за трьома напрямками (профілями):

– за напрямом видів біосигналів і даних (електрокардіограма, енцефалограма, реограма, імпедансна спектрограма, віброартрограма, електронні медичні карти і т.д.);

– за напрямом рівнів опрацювання (сенсорний/інструментальний рівень відбору і аналогових перетворень біосигналів, рівень кондиціонування сигналів методами цифрової обробки і машинного навчання, системний рівень визначення та інтерпретації біомедичних параметрів для конкретного застосування);

– за напрямом прикладних задач біоінформатики (біометрична ідентифікація, медична діагностика, клінічний моніторинг, афективна інформатика, людино-машинна взаємодія).

7. Удосконалено метод пошуку оптимальних значень гіперпараметрів багатопарового нейрокласифікатора, особливістю якого є використання прогностичної моделі збудованої за допомогою обчислювально простого алгоритму машинного навчання, а його застосування дає змогу автоматизувати і на 4 порядки скоротити час пошуку порівняно із повним перебором в просторі можливих значень.

8. Запропоновано і апробовано новий підхід до виявлення та коригування залишкових аномалій в біосигналах, який базується на застосуванні нейромережових автоенкодерів для нелінійної фільтрації завад, зосереджених в тій самій частині спектру, що й корисний сигнал, а його застосування у 5-7 разів зменшує похибку ідентифікації, а порівняно із відомими підходами також забезпечує збереження обсягу набору даних.

9. Розроблено новий підхід із забезпечення інваріантності класифікатора системи до параметрів біосенсорів і апаратури відбору біосигналів, особливістю якого є уніфікація параметрів цифрових записів цих сигналів за допомогою програмних кондиціонерів, що здійснюють передискретизацію, а також унормування за рівнем і тривалістю наборів даних.

10. Виконано комплексне дослідження впливу варіативності параметрів системи відбору ЕКГ-сигналу (різних наборів даних, частоти дискретизації, тривалості запису, обсягу бази) на точність біометричної ідентифікації, результати

якого показали достатню для потреб практики стійкість і надійність електрокардіограми як біометричного маркера, що підтвердило можливість його застосування у реальних системах біометричної ідентифікації.

11. Проведено порівняльний аналіз ефективності різних методів побудови класифікаторів систем біометричної ідентифікації на основі електрокардіограми, що дало можливість вибрати методи, які зберігають інваріантність до параметрів системи відбору ЕКГ-сигналу та придатні до масштабування (збільшення числа користувачів, неоднорідність даних, перекошеність класів і т.д.).

12. Розроблено новий підхід до коригування динамічних похибок у автокомпенсаційному перетворювачі біоімпедансу, зміст якого полягає у формуванні на основі нелінійної математичної моделі перетворювача навчального набору даних, застосуванні цього набору для навчання штучної нейромережі, яка у подальшому коригує результати вимірювання складових біоімпедансу, забезпечуючи майже на 3 порядки розширення діапазону робочих частот аналізатора біоімпедансу.

13. Удосконалено комп'ютерну систему підтримки прийняття рішень для проведення медичної діагностики стану колінних суглобів, основу якої складає новий метод інтелектуального опрацювання сигналів віброартрографії, що відрізняється описом діагностичних ознак за дескрипторами, одержаними статистичними методами на вибраних рівнях хвилькової декомпозиції ВАГ-сигналу, а його застосування забезпечило високу точність і чутливість класифікатора системи медичної діагностики.

14. Розроблено новий метод реалізації цифрового фазоквадратурного перетворювача сигналів імпедансної спектроскопії, який завдяки заміні операцій акумуляції добутків арифметичного перемноження вибірок аналізованої напруги і опорних сигналів на просте підсумовування вибірок аналізованої напруги упродовж визначених інтервалів не менше ніж вдвічі збільшує обчислювальну ефективність.

15. Розширено функціональні можливості, покращено метрологічні характеристики та вдвічі збільшено швидкодію цифрового реографа завдяки модифікації схеми джерела струму Хавленда, компенсації базового біоімпедансу за допомогою кодированого потенціометра (тримера) та використанню сформованих на засадах прямого цифрового синтезу ортогональних зондувальних сигналів.

16. Запропоновано застосувати схему розбалансованого моста змінного струму для вимірювання електродермальної активності в задачах афективної інформатики, застосування якого дало змогу істотно підвищити чутливість первинного перетворювача адмітанса.

17. Виконано апробацію результатів дисертаційних досліджень, на прикладах програмно-апаратної реалізації комп'ютерної системи біометричної ідентифікації, біо-машинного інтерфейсу управління за сигналами електроенцефалограми, мобільної системи розпізнавання емоцій, нечіткої моделі прогнозування захворювань на основі EMR-даних (записів електронних медичних карт).

СПИСОК ОПУБЛІКОВАНИХ ПРАЦЬ ЗА ТЕМОЮ ДИСЕРТАЦІЇ

Наукові праці, в яких опубліковано основні наукові результати дисертації

1. Стадник Б., Хома В., Хома Ю. Застосування штучних нейронних мереж для коригування похибок вимірювання імпедансу // Метрологія та прилади. 2016. № 02. с. 47-52.
2. Хома В., Хома Ю., Герасименко В., Сабодашко Д. ЕКГ-ідентифікація з використанням глибинних нейронних мереж // Вісник НУ «Львівська політехніка» – «Автоматика, вимірювання та керування». – 2017. №880. с. 67-72.
3. Хома Ю.В., Стадник Б.І., Микийчук М.М., Фріш С. Методи і засоби вимірювання та комп'ютерного опрацювання біосигналів // Вимірювальна техніка та метрологія. 2018. Вип. 79 (3). с. 5-16.
4. Хома Ю.В. Дослідження придатності електрокардіограм для біометричної ідентифікації // Сучасна спеціальна техніка. – 2018. 4(55). с. 104-114.
5. Хома Ю.В. Класифікація віброартрографічних сигналів з використанням хвилькового перетворення і технологій машинного навчання // Вісник НУ «Львівська політехніка» – «Інформаційні системи та мережі». – 2019. №5. с. 40-52.
6. Дудикевич В.Б., Хома В.В., Чекурін В.Ф., Хома Ю.В., Сабодашко Д.В. Нормалізація сигналів ЕКГ для застосування в системах біометричної ідентифікації // Вчені записки Таврійського національного університету імені В.І. Вернадського. Серія: Технічні науки. – 2019. Том 30 (69), ч. 1, № 4, с. 49-56.
7. Хома В.В., Хома Ю.В., Сабодашко Д.В., Хома П.П. Автоенкодера для опрацювання промахів сигналів ЕКГ у системі біометричної автентифікації // Штучний інтелект. – 2019. №1-2. с. 101-110.
8. Khoma Y., Szmajda M., Pelc M. Development of scientific-methodological approaches of machine learning application in biosignals processing // Herald of Advanced Information Technology. 2020, Vol. 3, No.1. p. 383–394.
9. Khoma V., Khoma Y., Kochan O. Unification of the analog part of the biosignal processing channel // Scientific Journal of TNTU. – 2020. Vol. 97, No 1. p. 97–109.
10. Хома Ю.В., Хома В.В. Розроблення цифрового еквалайзера для біосенсора портативного аналізатора біоімпедансу // Сучасні інформаційні системи. – 2020. 4 (2). с. 100-109.
11. Сабодашко Д.В., Хома Ю.В., Хома В.В. Дослідження часової стійкості сигналу ЕКГ як біометричного маркера в системі автентифікації // Вчені записки Таврійського національного університету імені В.І. Вернадського. Серія: Технічні науки. – 2020. Том 31(70), №2. с. 170-180.
12. Хома Ю.В., Хома В.В., Юн С., Кочан О.В. Аналіз ефективності методів коригування промахів у системах біометричної ідентифікації на підставі електрокардіограми // Науковий вісник НЛТУ України. – 2020. 30(3). с. 99-105.
13. Yatsyshyn S., Stadnyk V., Khoma Y., Mykyuchuk M., Mihunov D., Froehlich T., Mastylo R. Development of software and hardware for the method of rehabilitation of patients with obliteration of atherosclerosis of the limb vessels // Вимірювальна техніка та метрологія. – 2019. Том 80, Вип. 1. с. 24-26.

14. Хома Ю.В. Ієрархічно-модульний принцип побудови комп'ютерних біоінформатичних систем // Комп'ютерні технології друкарства. – 2020. № 1(43). с. 6-25.
15. Хома Ю.В., Микитин І.П., Артемук С.І., Бридінський В.А. Нейрокомп'ютерна система керування на основі платформи Open VCI та глибоких нейронних мереж // Вчені записки Таврійського національного університету імені В.І. Вернадського. Серія: Технічні науки. – 2020. – Том 31(70), №4. с. 127-135.
16. Pelc M., Khoma Y., Khoma V. ECG Signal as Robust and Reliable Biometric Marker: Datasets and Algorithms Comparison // Sensors. – 2019. 19(10), 2350. p. 1-8. (Web of Science, Scopus, Q2)
17. Smajda M., Khoma V., Khoma Y., Otenko V. Zastosowanie technologii cyfrowego przetwarzania sygnałów w nowoczesnych układach reograficznych wbudowanych // Przegląd Elektrotechniczny (Electrical Review). – 2019. R. 95, Nr 11. s. 231-237. (Web of Science, Scopus, Q3)
18. Su Jun, Szmajda M., Khoma V., Khoma Y., Sabodashko D., Kochan O., Jinfei Wang. Comparison of methods for correcting outliers in ECG-based biometric identification // Metrology and measurement systems. – 2020. Vol. 27(3). – p. 387–398. (Scopus, Q2)
19. Duke J., Ryan P., Suchard M., Hripcsak G., Jin Peng, Reich C., Schwalm M.-S., Khoma Y., Wu Yonghui, Xu Hua, Shah N., Banda J., Schuemie M. Risk of Angioedema Associated with Levetiracetam Use: Findings of the Observational Health Data Sciences and Informatics // Epilepsia. – 2017. Vol. 58, Issue 8. – p. e101–e106. (Web of Science, Scopus, Q1)
20. Galaznik A., Reich C., Klebanov G., Khoma Y., Allakhverdiiev E., Hather G., Yaping Shou. Predicting Outcomes in Patients With Diffuse Large B-Cell Lymphoma Treated With Standard of Care // Cancer Informatics. – Vol. 18. – p. 1–16. (Web of Science, Scopus, Q3)
21. Khoma Y., Khoma V., Podpora M., Kawala-Sterniuk A., Martinek R., Ozana S. Synthesis of Dynamic Errors Correcting Algorithm for Auto–Balancing Bridge Methods // IEEE Access. – 2020. Vol. 8. p. 165307-165318. (Web of Science, Scopus, Q1)
22. Khoma V., Pelc M., Khoma Y., Sabodashko D. Outlier Correction in ECG-Based Human Identification. // International Scientific Conference Brain Computer Interface 2018 Opole, Poland, 13-14 March 2018. In: Hunek W., Paszkiel S. (eds) Biomedical Engineering and Neuroscience. Advances in Intelligent Systems and Computing. 2018. Vol 720. p. 11-22. Springer, Cham. (Scopus, Q3)

Праці, які засвідчують апробацію матеріалів дисертації

23. Хома Ю., Герасименко В., Сабодашко Д. ECG identification using deep neural networks // Матеріали VI Міжнародної науково-технічної конференції «Захист інформації і безпека інформаційних систем». Львів, 1–2 червня 2017 р. – с. 53-54.
24. Wrzuszczak M., Khoma Y. Wykorzystanie sztucznych sieci neuronowych do zmniejszenia błędów przetworników impedancji // Materiały XLVIII Międzyuczelnianej

Konferencji Metrologów. Kraków, 5-7 września 2016 r. Zeszyty Naukowe Wydziału Elektrotechniki i Automatyki PG. 2016. Nr 49. s. 133-136.

25. Wieclaw L., Khoma Y., Falat P., Sabodashko D., Herasymenko V. Biometric Identification From Raw ECG Signal Using Deep Learning Techniques // In Proc.: The 9th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications. Romania, Bucharest, 21-23 September, 2017. p. 129-133. (Web of Science, Scopus)

26. Stadnyk B., Fröhlich T., Khoma Y., Herasymenko V., Chaban O. Impedance analyser error correction using artificial neural networks // In: Engineering for a Changing World: Proceedings; The 59th Ilmenau Scientific Colloquium, Technische Universität Ilmenau, 11-15 September, 2017, Ilmenau, Germany. – 2017. Vol. 59, No. 1.2.P1.

27. Гапоненко М. В., Брийовський Р. М., Артемук С. І., Бридіньський В. А., Назаркевич І. Б., Микитин І. П., Хома Ю. В. Керування моделлю автомобілю за допомогою мозку: програмна реалізація // Technical Using of Measurement-2018: тези доповідей IV Всеукраїнської науково-технічної конференції молодих вчених у царині метрології (Славське, 13–18 лютого 2018 р.). – 2018. – с. 114.

28. Khoma V., Pelc M., Khoma Y. Artificial Neural Network Capability for Human Being Identification based on ECG // Proceedings of the 23rd International Conference on Methods and Models in Automation and Robotics. Poland, Miedzyzdroje, 27-30 August 2018. p. 479-482. (Web of Science, Scopus)

29. Karpinski M., Khoma V., Dudykevych V., Khoma Y., Sabodashko D. Autoencoder Neural Networks for Outlier Correction in ECG- Based Biometric Identification / Proceedings of the 2018 IEEE 4th International Symposium on Wireless Systems within the International Conferences on Intelligent Data Acquisition and Advanced Computing Systems (IDAACS-SWS). Lviv, 20-21 Sept. 2018. p. 210- 215. (Web of Science, Scopus)

30. Стадник Б.І., Хома Ю.В. Застосування технологій глибокого навчання для опрацювання біосигналів // V Всеукраїнська науково-технічна конференція у царині метрології «Technical Using of Measurement – 2019». Славське, Україна, 29 січня - 2 лютого 2019 р.: Тези доповідей. – 102 с. ISBN 978-617-7363-88-9

31. Khoma P., Ruda K., Khoma Y., Khoma V. Automating of artificial neural network hyperparameters selection in cybersecurity // Тези доповідей VII міжнародної науково-технічної конференції “Захист інформації і безпека інформаційних систем”. Львів, 30-31 травня 2019. с. 142-143.

32. Khoma V., Khoma Y., Sabodashko D., Shereha V. Outlier Correction using Autoencoder Neural Networks for Human Being Identification based on ECG // Тези доповідей VII міжнародної науково-технічної конференції “Захист інформації і безпека інформаційних систем”. Львів, 30-31 травня 2019. с. 128–129.

33. Хома Ю., Сабодашко Д. Біометрична ідентифікація за допомогою електрокардіограми // Захист інформації і безпека інформаційних систем : матеріали V Міжнародної науково-технічної конференції, 2–3 червня 2016 р., Львів. 2016. С. 146–147.

34. Гапоненко М. В., Брийовський Р. М., Артемук С. І., Бридіньський В. А., Назаркевич І. Б., Микитин І. П., Хома Ю. В. Керування моделлю автомобілю за

допомогою мозку: програмна реалізація // Technical Using of Measurement-2018: тези доповідей IV Всеукраїнської науково-технічної конференції молодих вчених у царині метрології (Славське, 13–18 лютого 2018 р.). – 2018. – с. 114.

АНОТАЦІЯ

Хома Ю.В. Теорія і методи комп'ютерного опрацювання біосигналів на основі машинного навчання. – *На правах рукопису.*

Дисертація на здобуття наукового ступеня доктора технічних наук за спеціальністю 05.13.05 – комп'ютерні системи та компоненти – Національний університет «Львівська політехніка», МОН України, Львів, 2020.

У дисертаційній роботі вирішено актуальну науково-прикладну проблему у галузі інструментального забезпечення біоінформатики – розвиток теоретичних засад і нових підходів до удосконалення комп'ютерних систем опрацювання біосигналів на основі широкого використання штучних нейромереж і технологій глибинного навчання.

Представлено концепцію трьох системних рівнів комп'ютерного опрацювання біосигналів, що базується на чіткому розмежуванні функцій системних рівнів від методів і засобів, що їх реалізують. Таке розділення сприяє структуризації знань, уможливорює оцінювання ефективності різних методів і вибір кращих проектних рішень із урахуванням специфіки завдань, умов і сценаріїв.

Автоматизовано пошук оптимальних значень гіперпараметрів багат шарового нейрокласифікатора шляхом використання простої прогностичної моделі машинного навчання. Це дає змогу на 4 порядки скоротити час пошуку порівняно із повним перебором в просторі можливих значень.

Розроблено і апробовано підхід до виявлення і коригування залишкових аномалій в біосигналах, який базується на застосуванні нейромережових автоенкодерів для нелінійної фільтрації завад, зосереджених в тій самій частині спектру, що й корисний сигнал. Застосування підходу у 5-7 разів зменшує похибку ідентифікації.

Результати роботи можуть бути застосовані до різних прикладних задач, у таких сферах як кібербезпека та системи доступу (біометрична ідентифікація), робототехніка (нейромережові інтерфейси управління) і афективна інформатика (аналіз психоемоційного стану), а також у медицині (діагностика, клінічні дослідження тощо).

Ключові слова: комп'ютерна біоінформатична система, біосигнали, біометрична ідентифікація, нейрокомп'ютерні інтерфейси, вимірювальні перетворювачі імпедансу, цифрове опрацювання сигналів, цифровий еквалайзер, алгоритми коригування аномалій, машинне навчання, класифікація, штучні нейронні мережі, автоенкодери.

АННОТАЦИЯ

Хома Ю.В. Теория и методы компьютерной обработки биосигналов на основе машинного обучения. – *На правах рукописи.*

Диссертация на соискание ученой степени кандидата технических наук по

специальности 05.13.05 – компьютерные системы и компоненты. – Национальный университет «Львовская политехника», МОН Украины, Львов, 2020.

В диссертационной работе решена актуальная научно-прикладная проблема в области инструментального обеспечения биоинформатики - развитие теоретических основ и новых подходов к совершенствованию компьютерных систем обработки биосигналов на основе широкого использования искусственных нейронных сетей и технологий глубинного обучения.

Представлена концепция трех системных уровней компьютерной обработки биосигналов, основанная на четком разграничении функций системных уровней от методов и средств, которые используются для их реализации. Такое разделение способствует структуризации знаний, позволяет оценить эффективность различных методов и выбрать лучшие проектные решения с учетом специфики задач, условий и сценариев.

Автоматизирован поиск оптимальных значений гиперпараметров многослойного нейроклассификатора путем использования простой прогностической модели машинного обучения. Это позволяет на 4 порядка сократить время поиска по сравнению с полным перебором в пространстве всех возможных значений.

Разработан и апробирован подход к выявлению и коррекции остаточных аномалий в биосигналах, основанный на применении нейросетевых автоэнкодеров для нелинейной фильтрации помех, сосредоточенных в той же части спектра, и полезный сигнал. Применение коррекции аномалий в 5-7 раз уменьшает погрешность идентификации.

Результаты работы могут быть применены в различных направлениях, таких как компьютерная безопасность и системы доступа (биометрическая идентификация), робототехника (нейросетевые интерфейсы управления) и аффективная информатика (анализ психоэмоционального состояния), а также в медицине (диагностика, клинические исследования и т.п.).

Ключевые слова: компьютерная биоинформационная система, биосигналы, биометрическая идентификация, нейрокомпьютерные интерфейсы, измерительные преобразователи импеданса, цифровая обработка сигналов, цифровой эквалайзер, алгоритмы коррекции аномалий, машинное обучение, классификация, искусственные нейронные сети, автоэнкодеры.

ANNOTATION

Khoma Yuriy. Theory and methods of computer processing of biosignals based on machine learning. – *Manuscript.*

A thesis prepared for obtaining a degree of Doctor of Science (S.D.) in technical sciences on a specialty 05.13.05 - computer systems and components – Lviv Polytechnic National University, Ministry of Education and Science of Ukraine, Lviv, 2020.

The thesis solves a scientific problem in the field of instrumental support of bioinformatics - the development of theoretical foundations, improvement of methodological, algorithmic, software, and technical basis of the computer systems for

processing of biosignals and data based on the extensive use of artificial neural networks and deep learning technologies.

Current state and future perspectives of machine learning usage in the computer bioinformatics systems are analyzed in the thesis. It is shown that the heterogeneousness of data and a wide range of bioinformatics tasks influenced the development of specialized solutions for each separate domain or application. This complicates the possibility to compare the effectiveness of certain methods as well as the usage of the best system design variants for the new tasks.

A novel framework related to the development of principles for the design of the biosignals computer processing systems involving a combination of machine learning techniques and digital signal processing is presented in the thesis.

The expediency of separation of the system levels within the process of biosignals processing is reasoned, and their functions are outlined. Innovativeness of the suggested approach lies in the separation of functions of the lower, middle, and upper levels from methods with the help of which they are realized, as well as from the implementation variants for these methods based on the hardware and software components. The middle system level is significantly invariable both in regards to the task to be solved and to the biosignal type. At the same time, the upper level is specific as to the final application, and the lower level is specific as to the type of biosignal. Distinct outlining of functions for each system level and the inter-level interfaces opens perspectives for information structuring during the analysis of the known decisions, which simplifies the analysis and comparison of the effectiveness of these solutions. The design process of the computer system for the specific tasks gets simplified and potentially quickens due to the possibility of transferring the best results among the related tasks. On the basis of the developed three system levels concept the range of tasks related to machine learning application and biosignals processing on all the system levels was studied and analyzed.

A novel method of optimal hyperparameters selection for a multilayer neural network classifier based on the Monte Carlo method and predictive modeling was developed and introduced in the thesis, new algorithms for detection and correction of anomalies in ECG-signals were presented as well. This resulted in reduction of hyperparameters optimization time by 4 orders compared to a grid search approach in the entire hyperspace of possible values.

A new approach to the detection and correction of residual anomalies in biosignals was developed and tested. This approach is based on the use of neural network autoencoders for nonlinear filtering of distortions that are located in the same spectral band as the useful signal and its application reduces identification error by 5-7 times.

A method of software conditioning of biosignals was developed, which enabled parameters unification of digital records of biosignals from open databases by resampling, rescaling, time normalization, etc. and its application improves the issue of low data volume while deep neural networks training.

A comprehensive study of the influence of the variability of ECG signal acquisition systems (different data sets, sampling frequency, recording duration, data volume) on the accuracy of biometric identification was performed. The obtained results proved sufficient stability and reliability of the electrocardiogram as a biometric marker and confirmed the possibility of its real-world application in biometric identification systems.

Neural network equalizer was designed for dynamical error correction of bioimpedance sensors, which resulted in expanding the operating frequency band up to 100 times.

Functionality was expanded, metrological characteristics were improved and the speed of the digital rheograph was doubled by improving the Howland current pump by compensating the basic bioimpedance with a coded potentiometer (trimmer) and using a direct digital synthesis of orthogonal probing signals.

Intelligent processing of vibroarthrography signals based on a combination of machine learning algorithms and wavelet decomposition was developed, which allowed to achieve better accuracy in terms of knee joint disorders analysis.

The results of the work can be used in various applications, such as cybersecurity and access systems (biometric identification), robotics (bio-machine control interfaces), and affective informatics (psycho-emotional state analysis), as well as a medical domain (diagnostics, clinical trials).

***Key terms:* computer bioinformatics system, biosignals, biometric identification, brain computer interfaces, impedance measuring converters, digital signal processing, digital equalizer, anomaly correction algorithms, machine learning, classification, artificial neural networks, autoencoders.**