МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ «ЛЬВІВСЬКА ПОЛІТЕХНІКА»

Кваліфікаційна наукова
праця на правах рукопису

САБОДАШКО ДМИТРО ВОЛОДИМИРОВИЧ

УДК 681.32: 004.93

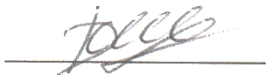
ДИСЕРТАЦІЯ

ВДОСКОНАЛЕННЯ МЕТОДІВ І ЗАСОБІВ БІОМЕТРИЧНОЇ АВТЕНТИФІКАЦІЇ НА ОСНОВІ ЕЛЕКТРОКАРДІОГРАМИ

125 – «Кібербезпека»
12 – «Інформаційні технології»

Подається на здобуття наукового ступеня доктора філософії

Дисертація містить результати власних досліджень. Використання ідей, результатів і текстів інших авторів мають посилання на відповідне джерело.



Д.В. Сабодашко

Науковий керівник:
Хома Володимир Васильович
доктор технічних наук, професор

Львів - 2021

АНОТАЦІЯ

Сабодашко Д.В. **Вдосконалення методів і засобів біометричної автентифікації на основі електрокардіограми.** – Кваліфікаційна наукова праця на правах рукопису.

Дисертація на здобуття наукового ступеня доктора філософії за спеціальністю 125 «Кібербезпека» (12 – Інформаційні технології). – Національний університет «Львівська політехніка», Львів, 2021.

У дисертаційній роботі розв'язано актуальну науково-прикладну задачу у галузі кібербезпеки - покращення характеристик системи біометричної автентифікації за сигналом електрокардіограми на основі раціонального поєднання технологій цифрового оброблення сигналів і машинного навчання, що підвищує рівень захищеності ресурсів на об'єктах інформаційної діяльності.

Дисертаційна робота складається зі вступу, чотирьох розділів, висновків, списку літературних джерел та додатків.

В ***першому розділі*** розглянуто основні режими роботи і подано порівняльну характеристику сучасних систем біометричної автентифікації. Проведено порівняння найпоширеніших біометричних маркерів за допомогою формалізованих критеріїв (універсальність, унікальність, постійність, вимірюваність, продуктивність, прийнятність, стійкість до обману, ціна, тощо). Представлено детальний опис електрокардіограми (ЕКГ) як біометричного маркера в системах розпізнавання, показано його переваги і проблеми на шляху практичного застосування в системах автентифікації. Проаналізовано відомі підходи опрацювання ЕКГ-сигналу на основі виділення характерних точок (fiducial points) та без такого виділення (non-fiducial point), тобто на основі інтелектуального аналізу повного набору вибірок ЕКГ-сигналу. Сформульовано завдання дисертаційного дослідження.

У ***другому розділі*** розглянуто особливості процесу автентифікації за ЕКГ-сигналом. Формалізовано структуру біометричної системи розпізнавання.

Наведено детальний опис і функції кожного із структурних елементів. Розглянуто перспективні підходи до покращення технічних і експлуатаційних характеристик біометричної системи ЕКГ-автентифікації. Передовсім, обґрунтовано доцільність введення в ланцюг опрацювання електрокардіограми двох додаткових компонент:

• компонента темпоральної нормалізації ЕКГ-сигналу, що покликана забезпечити інваріантність результатів автентифікації до зміни частоти серцевого ритму, тим самим підвищивши достовірність роботи біометричної системи;

• компонента виявлення та коригування артефактів у ЕКГ-сигналі, яка за допомогою інструментарію статистики або машинного навчання підвищує точність і швидкодію системи біометричної автентифікації.

Подано методики оцінювання ефективності методів і засобів біометричної автентифікації на основі ЕКГ-сигналу. Представлено сформовану автором базу записів електрокардіограм (Lviv Biometric Dataset), яка на момент написання дисертації містила 1809 записів виміряних у 115 суб'єктів на часовому горизонті понад два роки. Базу Lviv Biometric Dataset викладено у відкритий доступ, поряд із іншими базами ЕКГ-записів.

Третій розділ спрямовано на розроблення моделей та методів для покращення характеристик біометричних систем автентифікації на основі ЕКГ.

Спрощений і зручний для систем автентифікації відбір ЕКГ-потенціалів із пальців лівої і правої рук призводить до зниження якості запису. Частотні смуги корисного сигналу і завад перекриваються, тому після цифрової фільтрації у ЕКГ-записах спостерігаються залишкові артефакти. Описані у літературних джерелах підходи спираються на виявлення і відкидання фрагментів ЕКГ з аномальними відхиленнями.

У роботі вперше запропоновано не відкидати, а виправляти фрагменти з аномальними відхиленнями, що важливо для збереження необхідного обсягу даних для класифікатора і скорочення часу відбору ЕКГ. Розроблено підхід до

виправлення залишкових артефактів у ЕКГ-сигналах, який складається із трьох етапів:

1. формування референційного образу біометричного маркера;
2. виявлення фрагментів ЕКГ-сигналу із промахами;
3. заміна цих фрагментів на відповідні значення із референційного образу.

Застосування такого підходу дає змогу суттєво підвищити достовірність результатів автентифікації.

Запропоновано і досліджено два методи формування референційного образу ЕКГ-маркера для виправлення залишкових артефактів у ЕКГ-сигналах у системі біометричної автентифікації:

- на основі невимогливої до обчислювальних ресурсів формальної статистичної моделі;

- на базі нечіткої нейромережевої моделі, що дає змогу зменшити похибки автентифікації першого і другого роду відповідно у 4 та 3 рази.

Для етапу виявлення фрагментів із промахами виконано дослідження впливу гіперпараметрів (тривалість ковзного вікна і поріг допустимого відхилення вибірки) на точність автентифікації, що дало змогу знайти оптимальні їх значення за критерієм мінімальної похибки.

Розроблено та апробовано інструментарій для темпоральної нормалізації ЕКГ-сигналу. Серед сучасних методів класифікації (метод опорних векторів, лінійний дискримінантний аналіз, k-найближчих сусідів, дерева рішень, нейронні мережі, тощо) здійснено вибір оптимального для побудови системи автентифікації. Досліджено придатність біометричної системи автентифікації до масштабування, а саме визначено вплив збільшення числа користувачів на точність автентифікації.

У ***четвертому розділі*** імплементовано біометричну систему автентифікації з покращеними характеристиками на основі використання розроблених автором моделей і методів. Зокрема, щоб показати можливість імплементації біометричної системи ЕКГ-автентифікації на пристроях з обмеженими

обчислювальними ресурсами імплементовано біометричну систему на базі мікрокомп'ютера Raspberry Pi 3B. Проведено дослідження швидкодії імплементованої біометричної системи, за результатами якого сформовано рекомендації для імплементації біометричних систем автентифікації як на основі персональної робочої станції, так і на базі мікрокомп'ютера Raspberry Pi.

Досліджено часову стабільність ЕКГ-сигналів на довготривалих проміжках часу (роки, місяці), а також визначено ступінь впливу варіативності інформативних ознак електрокардіограми на точність автентифікації. Результати досліджень засвідчили, що ЕКГ є стабільним маркером і може застосовуватися у реальних системах автентифікації, причому система здатна адекватно розпізнавати користувачів упродовж тривалого часу без необхідності проміжного калібрування. Таким чином, доведено високий потенціал і перспективність електрокардіограми, як біометричного маркера, для побудови надійних систем автентифікації.

Наведено перелік можливих сфер застосування та опис прикладних застосувань для біометричних систем автентифікації за ЕКГ-сигналом.

Ключові слова: біометрична автентифікація, електрокардіограма, цифрове оброблення сигналу, машинне навчання, нейронні мережі, автоенкодери, корекція аномалій, темпоральна нормалізація, Raspberry Pi.

ABSTRACT

Sabodashko D.V. **Improvement of methods and means of biometric authentication based on electrocardiogram. –** Manuscript.

Thesis paper for achievement of the scientific degree Doctor of Philosophy in the specialty 125 “Cybersecurity” (12 – Information technologies). – Lviv Polytechnic National University, Lviv, 2021.

The thesis paper solved the relevant scientific and applied task in the sphere of cybersecurity – improvement of characteristics in biometric authentication systems based on electrocardiogram signal via the rational combination of digital signal processing technologies and machine learning, which increases the resources protection level on the information activity objects.

The thesis paper consists of introduction, four chapters, conclusions, list of references and appendixes.

The first chapter examines the main working modes and provides the comparative characteristics of modern biometric authentication systems. Comparison of the most widely spread biometric markers using the formalized criteria (scalability, uniqueness, steadiness, productivity, acceptability, stability, forge resistance, price, etc.) was performed. Detailed description of electrocardiogram (ECG) as a biometric marker in the recognition systems was presented, its advantages and problems on the way of practical application in the authentication systems were revealed. Known approaches of ECG-signal processing based on fiducial points extraction and without such an extraction (non-fiducial point), thus on the basis of the intellectual analysis of the full set of ECG-signal data selection were analyzed. The task of the thesis research was stated.

In ***the second chapter*** the peculiarities of authentication process based on ECG-signal are studied. The structure of biometric recognition system is formalized. The detailed description and functions of each of the structure elements are provided. Perspective approaches to the improvement of the technical and operating

characteristics for ECG-authentication biometric system are studied. First of all, the expediency of introduction into the electrocardiogram processing pipeline of two additional components is reasoned:

• component of ECG-signal temporal normalization, which is supposed to provide invariability of the authentication results as to the change of the heart rate, thus increasing the authenticity of the biometric system work;

• component detecting and correcting anomalies in the ECG-signal, which using the statistical or machine learning instruments increase the accuracy and the speed of the biometric authentication.

Techniques evaluating efficiency of methods and means of biometric authentication based on ECG-signal are provided. Electrocardiogram records dataset (Lviv Biometric Dataset) formed by the author is presented, which at the moment the thesis was written contained 1809 records measured from 115 subjects on the time horizon of over two years. Lviv Biometric Dataset is open sourced together with other ECG-records datasets

The third chapter is directed onto development of models and methods for improvement of ECG-based biometric authentication systems.

The simplified and friendly to authentication systems selection of ECG-signals from the fingers of the left and right hands lead to decrease in record quality. Frequency bands of the useful signal and obstacles are overlaid, thus after the digital filtering the residual artifacts remain the in the ECG-signal. The approaches described in the references are drawn upon the detection and removal of ECG fragments with anomaly deviations.

It was first suggested in the work not to remove, but to correct the fragments with the anomaly deviations, which is important for preservation of the data amount for the classifier and shortening of time for ECG selection. Approach for correction of the residual artifacts in the ECG-signals was developed, it consists of three stages:

1) formation of reference pattern for the biometric marker;
2) detection in the ECG-signal of fragments with anomalies;

3) substitution of these fragments with the corresponding values from the reference pattern.

Application of such an approach allows to substantially increase the validity of the authentication results.

Two methods for reference pattern formation of the ECG-marker were suggested for correction of the residual artefacts in the ECG-signals for systems of biometric authentication:

- based on the undemanding as to the calculation resources formal statistical model;

- based on the fuzzy neural network model allowing to decrease the type I and type II authentication errors in 4 and 3 times correspondingly.

For the purposes of stage detecting fragments with anomalies the research on influence of hyperparameters was performed (duration of the sliding window and the threshold of the acceptable deviation of the dataset) for accuracy of the authentication, which allowed to find their optimal values according to the criterion of the minimal error.

Instruments for temporal normalization of ECG signal were designed and tested. The selection was performed among the modern classification methods (support vector machines, linear discriminant analysis, k-nearest neighbors algorithm, decision trees, neural networks, etc.) for optimal design of the authentication system. Scalability of the biometric authentication system was researched, namely the influence of users increase on the authentication accuracy was studied.

Biometric authentication system with improved characteristics based on usage of the models and methods designed by the author was implemented in the ***fourth chapter.*** Among others, to show the implementation possibility of the biometric system with ECG-authentication on the devices with limited calculation resources the biometric system based on Raspberry Pi 3B was implemented. The speed of the implemented biometric authentication system was verified, and based on the obtained results the recommendations for implementation of the biometric authentication systems on the stand-alone PC as well as on the Raspberry Pi microcomputer were formed.

Long term (years, months) temporal stability of the ECG-signals was studied, as well as the influence of the variability of the informative features from electrocardiogram on the authentication accuracy was defined. Research results certified that the ECG was a stable marker and might be applied in real authentication systems, at this the system was able to adequately recognize the users during the long time without the necessity of the intermediate calibration. Therefore, the high potential and prospects of the electrocardiogram as a biometric marker for the design of the reliable authentication systems are proved.

The list of possible application spheres and the description of the applied usage for biometric authentication systems based on the ECG-signal are provided.

Keywords: biometric authentication, electrocardiogram, digital signal processing, machine learning, neural networks, autoencoders, anomaly correction, temporal normalization, Raspberry Pi.

СПИСОК ПРАЦЬ ОПУБЛІКОВАНИХ ЗА ТЕМОЮ ДИСЕРТАЦІЇ

Наукові праці, в яких опубліковано основні наукові результати дисертації

1. 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.

2. 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.

3. Хома В.В., Хома Ю.В., Сабодашко Д.В., Хома П.П. Автоенкодери для опрацювання промахів сигналів ЕКГ у системі біометричної автентифікації // Штучний інтелект. – 2019. №1-2. С. 108-117.

4. 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.

5. Хома Ю.В., Хома В.В., Сабодашко Д.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.

6. 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)*

7. 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)*

8. Sabodashko D. Normalizacja temporalna sygnału EKG w systemie identyfikacji biometrycznej // Przetwarzanie, transmisja i bezpieczeństwo informacji: monografia / Akademia Techniczno-Humanistyczna. Bielsko-Biała, 2019. T. 2. S. 313–322.

Праці, які засвідчують апробацію матеріалів дисертації

1. 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.

2. 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)*

3. 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)*

4. 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.

5. 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.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ

АЦП	аналого-цифрове перетворення
АЧХ	амплітудно-частотна характеристика
ДНК	дезоксирибонуклеїнова кислота
ЕКГ	електрокардіограма
ФЧХ	фазо-частотна характеристика
AUC	area under curve
FAR	false accept rate
FRR	false reject rate
FN	false negatives
FP	false positives
EER	equal error rate
ICA	independent component analysis
KNN	k-nearest neighbors
LDA	linear discriminant analysis
PCA	principal component analysis
ROC	receiver operating characteristic
SVD	singular value decomposition
SVM	support vector machine
TN	true negatives
TP	true positives

ВСТУП

Актуальність теми. Стрімкий розвиток інформаційних технологій призвів до їх проникнення практично в усі сфери людської діяльності. Відтак одним з першочергових завдань є забезпечення надійності й захищеності інформаційних систем, що вводить кібербезпеку і захист даних в ранг найважливіших та найактуальніших проблем сьогодення.

Впродовж тривалого часу на ринку спостерігається стабільне та істотне зростання попиту на біометричні системи контролю доступу, які базуються на розпізнаванні користувачів за унікальними фізіологічними чи поведінковими ознаками. Ці ознаки, або як їх ще називають - біометричні маркери, можуть походити зі сталих фізіологічних характеристик, таких як відбиток пальця, геометрія долоні, райдужна оболонка ока або видобуватися із процесу, що відображає особливості роботи органів і систем організму, наприклад, спектральне забарвлення голосового тону, почерк на письмі чи стиль роботи на клавіатурі.

Біометрична система може застосовуватися у двох режимах – автентифікації (верифікації) чи ідентифікації. У режимі автентифікації біометрична система дає відповідь чи є допустимою розбіжність зареєстрованих характеристичних показників із «образом», що зберігається в пам’яті системи і прив’язаний до заявленого в ідентифікаторі користувача.

Для забезпечення надійного розпізнавання людей біометричні маркери повинні відповідати низці вимог, передусім таких як:

- універсальність (кожна людина повинна мати таку ознаку);
- унікальність (ознака кожної людини має бути унікальною);
- довговічність (з плином часу ознака повинна бути незмінною);
- вимірюваність (ознака має бути придатною для вимірювання);
- ефективність (можлива ефективна програмно-апаратна імплементація біометричного методу розпізнавання за обраною ознакою).

Проте ключовою вимогою до вибору біометричного маркера є стійкість до підміни (фальсифікації). Саме цій вимозі не завжди відповідають традиційно поширені у біометриці технології. Наприклад, відбитки пальців можна підмінити за допомогою гелевих накладок на пальці, а райдужну оболонку ока можна сфальсифікувати контактними лінзами. Тому тривають пошуки нових біометричних маркерів, які у комплексі відповідають зазначеним вимогам. Однією із таких є електрокардіограма (ЕКГ). Цей біометричний маркер є внутрішнім, тому його складно імітувати. Крім того, електрокардіограма завжди доступна і походить із життєво важливого органу.

Проте такі особливості ЕКГ-сигналу як мінливість, низький рівень, вразливість до завад знижують ефективність їх аналізу методами цифрового оброблення сигналів, що спираються на формальні моделі. Тому відносно хороші результати автентифікації людей за ЕКГ-сигналом можна отримати лише за допомогою інструментів машинного навчання, що стало можливим лише в останні роки.

На цей час опубліковано десятки наукових публікацій, присвячених застосуванню електрокардіограми як біометричного маркера. Досліджено методи автентифікації, що базуються як на виділенні характерних точок (fiducial points) електрокардіограми, так і на основі безпосереднього опрацювання вибірок ЕКГ-сигналу. Застосовано класифікатори, що працюють за різними алгоритмами машинного навчання, а одержані результати показали досить високу точність розпізнавання. Проте на шляху застосування електрокардіограми у реальних біометричних системах автентифікації залишалися недослідженими низка важливих питань, зокрема стабільність ЕКГ як біомаркера на довготривалому часовому інтервалі, вплив варіабельності серцевого ритму на точність системи ідентифікації. З погляду практики важливо також розробити методи кондиціювання сигналів ЕКГ, щоб забезпечити можливість застосування простих способів відбору біопотенціалів.

Таким чином, пошук шляхів покращення технічних і експлуатаційних характеристик біометричних систем автентифікації за ЕКГ-сигналом є актуальною науково-практичною задачею галузі кібербезпеки.

Зв'язок роботи з науковими програмами, планами, темами. Дисертацію виконано на кафедрі захисту інформації Національного університету «Львівська політехніка». Тема дисертації відповідає науковому напрямку кафедри.

Дисертаційні дослідження виконувалися в межах держбюджетних науково-дослідних робіт:

- «Розвиток теоретичних засад створення комплексних систем безпеки автоматизованих і комунікаційних систем» (№ державної реєстрації 0115U006722; терміни виконання - 2018-2020 рр.);
- «Розроблення та удосконалення методів і засобів захисту інформації для протидії несанкціонованому доступу в інформаційно-комунікаційних мережах» (№ державної реєстрації 0119U101690; терміни виконання - 2020-2022 рр.).

Мета і завдання дослідження. Метою дисертаційного дослідження є покращення характеристик системи біометричної автентифікації за сигналом електрокардіограми на основі раціонального поєднання технологій цифрового оброблення сигналів і машинного навчання, що забезпечить підвищення рівня захищеності ресурсів на об'єктах інформаційної діяльності.

Для досягнення зазначеної мети необхідно виконати такі завдання:

- проаналізувати сучасні методи і засоби біометричної автентифікації та сформулювати завдання дисертаційного дослідження;
- розробити концепцію покращення характеристик біометричних систем автентифікації;
- опрацювати, імплементувати і апробувати методи виявлення та коригування артефактів у ЕКГ-записах;
- синтезувати та верифікувати метод темпоральної нормалізації ЕКГ-сигналів;

- дослідити часову інваріантність електрокардіограми як біометричної характеристики;
- реалізувати біометричну систему ЕКГ-автентифікації на основі мікрокомп’ютера Raspberry Pi;
- розробити рекомендації для практичного використання імплементованої системи.

Об’єктом дослідження є процеси відбору та опрацювання ЕКГ-сигналу в системах біометричної автентифікації.

Предметом дослідження є методи та засоби підвищення точності й обчислювальної ефективності системи біометричної автентифікації за ЕКГ-сигналом.

Методи дослідження. Для розв’язання сформульованих задач застосовано методи: системного і порівняльного аналізу, теорії систем і цифрового оброблення сигналів, машинного навчання, математичної статистики, імітаційного моделювання. Методами системного та порівняльного аналізу досліджено моделі і структури сучасних систем біометричної автентифікації з метою виявлення їхніх переваг та недоліків. Теорію систем, методи цифрового оброблення сигналів і машинного навчання застосовано для розроблення нових і вдосконалення відомих підходів опрацювання ЕКГ-сигналів. Інструментарій математичної статистики використано для виявлення і коригування аномальних відхилень, а імітаційне моделювання - для дослідження розроблених моделей і алгоритмів.

Наукова новизна отриманих результатів полягає в розробленні та удосконаленні методів опрацювання ЕКГ-сигналу для підвищення ефективності систем біометричної автентифікації:

1. Вперше розроблено підхід до виправлення залишкових артефактів у ЕКГ-сигналах, який складається із трьох етапів - формування референційного образу біометричного маркера, виявлення фрагментів із промахами та, власне, заміна

цих фрагментів на відповідні значення із референційного образу, застосування якого дає змогу підвищити достовірність результатів автентифікації.

2. Вперше для етапу виявлення фрагментів із промахами виконано дослідження впливу гіперпараметрів (тривалість ковзного вікна і поріг допустимого відхилення вибірки) на точність автентифікації, що дало змогу знайти оптимальні їх значення за критерієм мінімальної похибки.

3. Вперше запропоновано метод формування референційного образу ЕКГ-маркера на основі статистичної моделі, який є невимогливим до обчислювальних ресурсів та може бути використаний для виправлення залишкових артефактів у ЕКГ-сигналах у системі біометричної автентифікації.

4. Вперше запропоновано метод формування референційного образу ЕКГ-маркера на базі нечіткої нейромережевої моделі, а його застосування для виправлення залишкових артефактів у ЕКГ-сигналах дає змогу зменшити похибки автентифікації першого і другого роду відповідно у 4 та 3 рази.

5. Розроблено метод темпоральної нормалізації серцевого ритму, який здійснює часову трансформацію ЕКГ-сигналу з приведення тривалості циклу до наперед встановленого значення, а його застосування дає змогу підвищити точність автентифікації.

6. Досліджено короткотривалу та довготривалу стійкість сигналу ЕКГ багатьох користувачів, за результатами якого доведено можливість практичного застосування електрокардіограми як біометричного маркера в реальних системах автентифікації.

Практичне значення одержаних результатів полягає у розробленні компонентів структурної схеми систем біометричної автентифікації, а саме:

- Досліджено швидкодію систем біометричної автентифікації імплементованих на основі персонального комп'ютера та платформи Raspberry Pi 3.

- На основі розробленого методу виявлення та виправлення промахів у ЕКГ-сигналах розроблено програмний модуль, застосування якого підвищує точність систем автентифікації на понад 7%.
- На основі розробленого методу темпоральної нормалізації ЕКГ-сигналів розроблено програмний модуль, який забезпечує стійкість нейромережевого автентифікатора до перенавчання, та підвищує його точність на 8%.
- Розроблено кожен із компонентів структурної схеми систем біометричної автентифікації та біометричну систему в цілому. Перевірено ефективність системи біометричної автентифікації імплементованої на платформі Raspberry Pi 3
- Зібрано та розміщено у відкритому доступі власний набір даних, який містить понад 1800 ЕКГ-записів від 115-ти осіб. Даний набір даних використано для перевірки ефективності розроблених методів.
- Теоретичні та практичні результати роботи впроваджено у діяльність ТОВ "СВІФТ СОЛЮШНС" (м. Харків) та 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 обґрунтованими допущеннями, які не суперечать відомим підходам та забезпечують адекватність застосованих моделей. Достовірність наукових положень, висновків і рекомендацій підтверджується узгодженістю теоретичних досліджень з результатами імітаційного моделювання, практичною імплементацією розроблених моделей і методів, результатами експериментів на реальних електрокардіограмах і даних, які узгоджуються з даними відомих досліджень, а також практичним впровадженням результатів дисертаційного дослідження.

Особистий внесок здобувача полягає у формулюванні мети та основних завдань досліджень, обґрунтуванні наукових положень. Автором проаналізовано

літературні джерела за темою дисертації, обґрунтовано напрями досліджень, розроблено моделі і методи темпоральної нормалізації, а також виявлення і коригування аномальних відхилень, виконано експериментальні дослідження, систематизовано і узагальнено отримані результати. Робота містить прикладні положення та висновки, сформульовані дисертантом особисто. Ідеї, положення чи гіпотези інших авторів, які присутні в дисертації, мають відповідні посилання і використані лише для підкріплення ідей та результатів здобувача. Постановка завдань та їхнє обговорення здійснено під керівництвом д.т.н., проф. Хоми В.В.

Апробація результатів дисертації. Результати дисертаційного дослідження апробовано на міжнародних наукових та науково-практичних конференціях, наукових школах та консорціумах, семінарах:

- V Міжнародна науково-технічна конференція «Захист інформації і безпека інформаційних систем» (2–3 червня 2016 року, Львів 2016, Україна);
- The 9th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (21-23 вересня, 2017 року, Бухарест, Румунія);
- VI Міжнародна науково-технічна конференція «Захист інформації і безпека інформаційних систем» (1–2 червня 2017 року, Львів 2017, Україна);
- The 3rd International Scientific Conference on Brain-Computer Interfaces (13–14 березня 2018 року, Ополе, 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 2018року, Львів, 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)

- IX Międzynarodowa Konferencja Studentów oraz Doktorantów „Inżynier XXI wieku” (6 грудня 2019, Бельсько-Бяла, Польща).

Публікації. Основні положення дисертації опубліковано у 13 наукових працях, з яких: 1 розділ у колективній монографії, 1 статтю у науковому періодичному виданні іншої держави, що включене до міжнародної наукометричної бази даних (Scopus), 1 статтю у серійному науковому виданні іншої держави, що включене до міжнародної наукометричної бази даних (Scopus), 5 статей у наукових фахових виданнях України з технічних наук та 5 наукових публікацій у збірниках матеріалів та тез конференцій, з яких 2 включені до наукометричних баз даних Web of Science, Scopus.

Структура й обсяг дисертації. Дисертаційна робота викладена на 204 сторінках та складається з анотації, змісту, переліку скорочень, вступу, чотирьох основних розділів, в яких міститься 53 рисунки та 35 таблиць, списку використаних джерел з 101 найменування, а також 6 додатків. За структурою, мовою та стилем викладення дисертація відповідає вимогам МОН України. Робота написана грамотною українською мовою з використанням сучасної наукової термінології, а стиль викладення матеріалу є послідовним та логічним.

РОЗДІЛ 1. ОГЛЯД ЛІТЕРАТУРНИХ ДЖЕРЕЛ ЗА ТЕМАТИКОЮ РОБОТИ

1.1. Порівняльна характеристика сучасних біометричних систем автентифікації

Біометрика - це наука про визначення особистості на основі фізіологічних чи поведінкових характеристик людини [1]. Біометричний захист ефективніший ніж такі традиційні методи автентифікації як, використання смарт-карток, паролів, PIN-кодів. Використовуючи біометричні дані, можна встановити особистість на основі того, ким ви є, а не на основі того, що вам належить (посвідчення особи, ID-картка) чи того, що ви пам’ятаєте (наприклад пароля) [1]. Біометрика забезпечує неймовірну зручність для користувачів (оскільки від користувачів більше не вимагається запам'ятовувати велику кількість довгих та складних паролів, що часто змінюються), зберігаючи при цьому досить високий ступінь безпеки.

Зазвичай біометричні системи працюють в трьох режимах роботи (рис. 1.1) [2]:

- Режим реєстрації даних: під час цього режиму роботи біометрична система збирає фізіологічні та/або поведінкові характеристики (наприклад, ЕКГ, зображення обличчя чи райдужної оболонки), виконує перевірку якості отриманих даних та зберігає їх у бази даних системи. У даному режимі додатково може здійснюватися витягнення інформативних ознак із виміряного сигналу.
- Режим ідентифікації: у цьому режимі система здійснює розпізнавання особи шляхом пошуку відповідності у всій базі даних шаблонів. Система проводить порівняння *один до багатьох* для встановлення особистості. Мета цього режиму роботи – відповісти на запитання: кому належать виміряні біометричні дані?
- Режим автентифікації: у цьому режимі система здійснює порівняння виміряних біометричних характеристик особи із біометричним шаблоном особи, який зберігається в базі даних системи. Мета цієї операції – відповісти на запитання: чи є користувач тим, ким він чи вона себе видає?

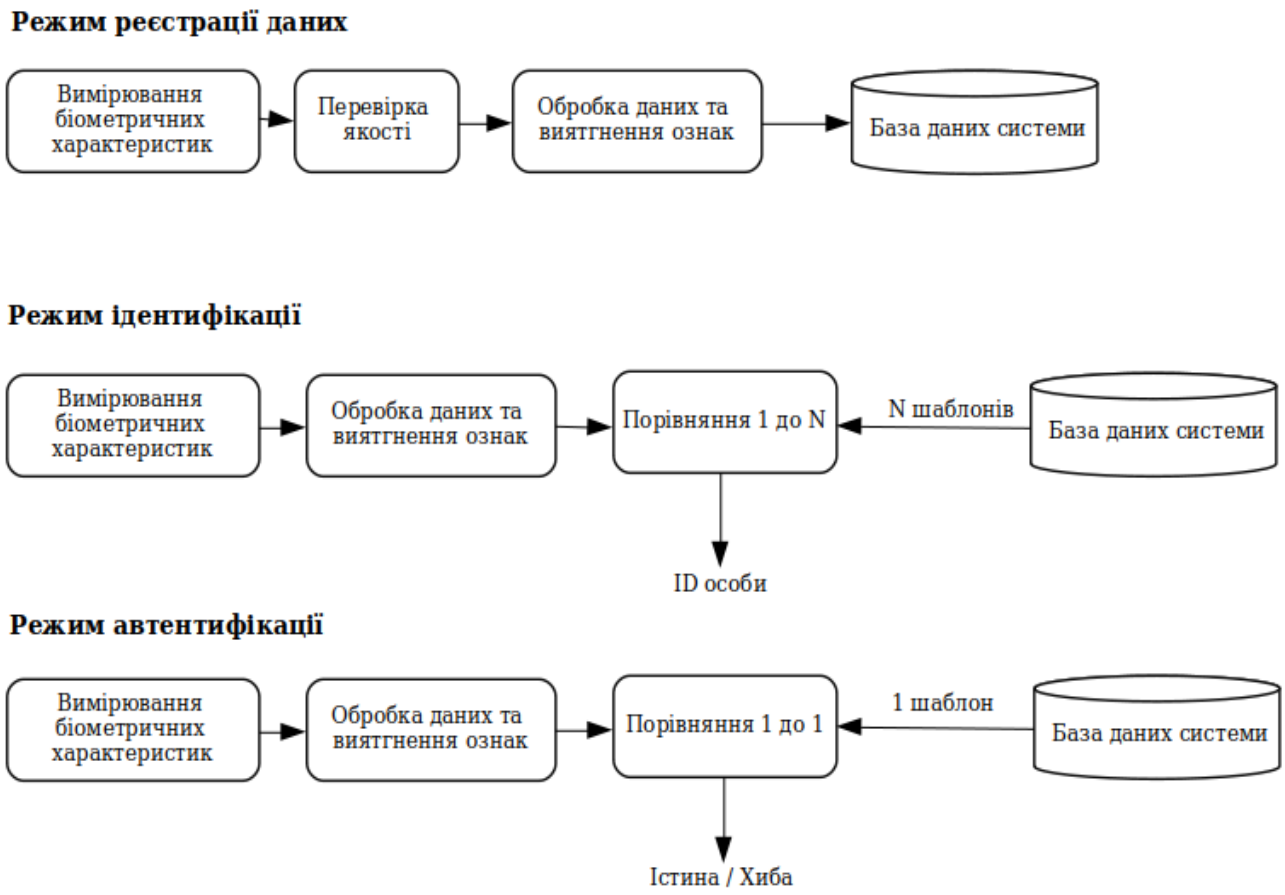


Рис. 1.1. Режими роботи біометричних систем

Фізіологічні характеристики – це біометричні дані, які безпосередньо вимірюються з тіла людини. До фізіологічних характеристик належать (рис. 1.2):

ДНК (Дезоксирибонуклеїнова кислота). Дані отримують із клітин, що містять ДНК, таких як шкіра, слизова оболонка рота або коріння волосся. Для розпізнавання використовується частина ДНК послідовності у вигляді унікального одновимірного масиву. В даний час ДНК використовується здебільшого у криміналістичних розслідуваннях для розпізнавання осіб. Основними недоліками застосування ДНК-коду в біометричних системах є те, що його можна легко викрасти у людей, а також складність та дороговизна імплементації біометричної системи [3].

Обличчя – це основна біометрична характеристика, яку люди використовують для розпізнавання один одного [4], і одна з найбільш універсальних та потужних біометричних характеристик. У роботі [5]

стверджується, що біометрія обличчя є другою найпоширенішою біометричною характеристикою після відбитків пальців, а також згадується, що, на відміну від відбитків пальців, для розпізнавання не потрібна згода. Обличчя можна записати здалеку, непомітно і без згоди особи. Тому розпізнавання обличчя може використовуватися не тільки для автентифікації, але також для спостереження та відстеження. Це також може бути загрозою для систем автентифікації на основі облич, оскільки зловмисники можуть отримати несанкціонований доступ до зображень або 3D-моделей облич та здійснювати спуфінг-атаки. Існує широкий вибір підходів до розпізнавання облич, включаючи системи на основі 2D або, де дані отримують із нерухомих зображень або послідовностей зображень у видимому чи інфрачервоному спектрі. Чжао та інші [6] поділяють системи розпізнавання облич на підходи на основі локалізованих ознак (рот, очі, ніс, тощо) та цілісні підходи на основі глобальної морфології обличчя. Однак фактори навколишнього середовища, такі як освітлення, можуть негативно впливати на роботу 3D-зображень відповідних систем. Крім того, старіння та прояви емоцій можуть суттєво впливати на точність біометричної системи [7].

***Відбитки пальців*:** це одна з найпоширеніших біометричних характеристик. До появи біометричних інструментів саме відбитки пальців (зафіксовані на папері з використанням чорнил) широко використовувались у криміналістиці для ідентифікації та перевірки злочинців. З появою новітніх технологій, відбитки пальців фіксуються за допомогою оптичних, ємнісних або ультразвукових датчиків, які вимірюють хребти, долини та острови на кінцівках пальців [8]. Незважаючи на те, що відбитки пальців легко імітувати отримавши їх із поверхонь, до яких раніше торкалися або за допомогою камер високої роздільної здатності, вважається, що відбитки пальців забезпечують високий рівень безпеки. Вони використовуються для ідентифікації та автентифікації на особистому (контроль доступу на смартфонах), корпоративному та державному (використовується у паспортах, виданих Європейським Союзом з 2009 р. [9] та паспортах України з 2015 р.[10]) рівнях.

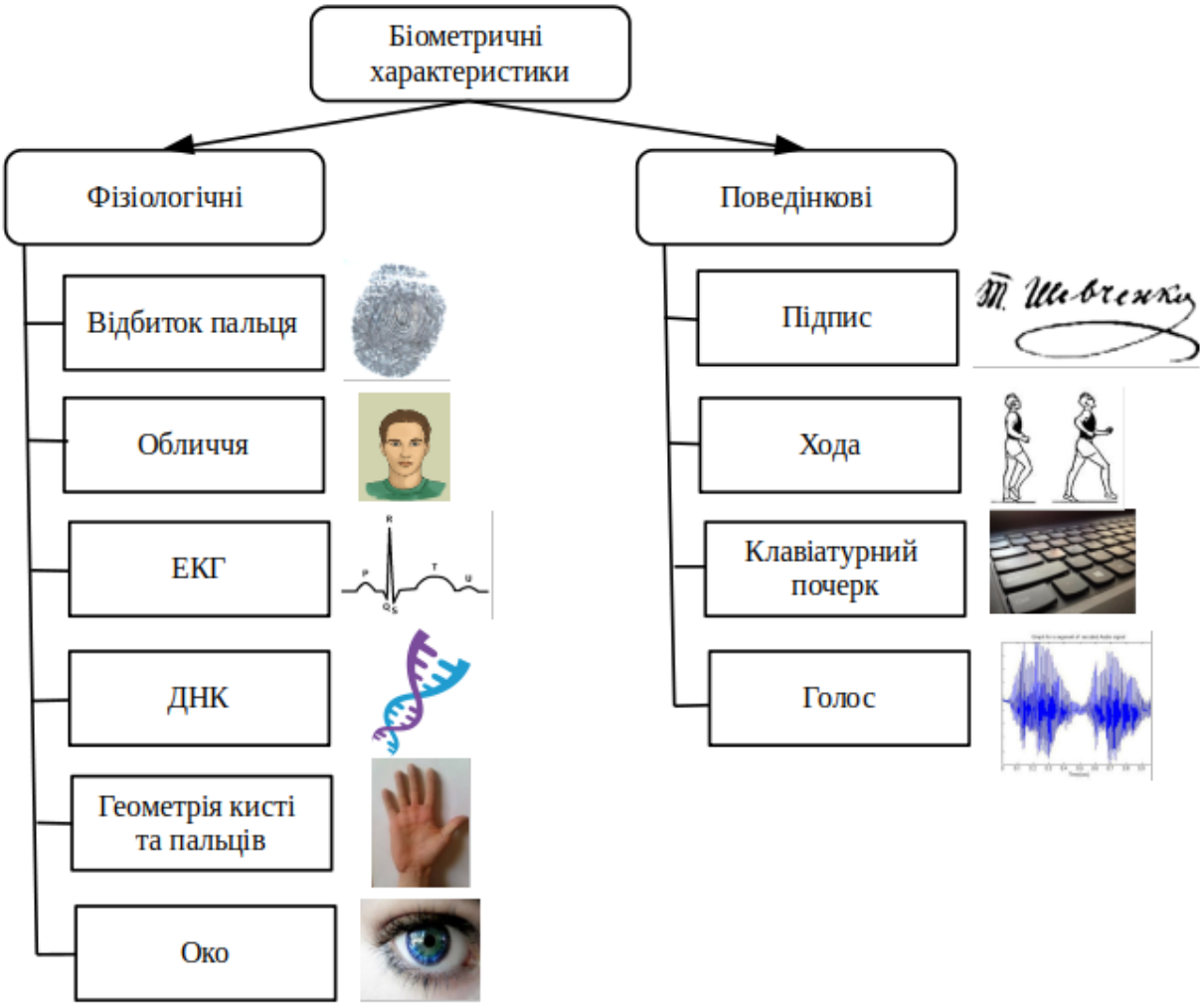


Рис. 1.2. Приклади фізіологічних та поведінкових біометричних характеристик

Геометрія кисті та пальців: Системи розпізнавання на основі геометрії рук засновані на ряді вимірювань людської руки, включаючи її форму, розмір долоні, довжину та ширину пальців, тощо. Геометрія кисті та пальців не дуже відмінна і не може бути використана для систем, що вимагають розпізнавання великої кількості персон [11].

Око. Разом з візерунками сітківки та склери, візерунки райдужної оболонки утворюють групу очної біометрії. Хоча вони розташовані дуже близько один до одного, процедура запису, а також сфера їх використання відрізняються одна від одної. Склера та райдужна оболонка лежать на зовнішній стороні ока і тому можуть реєструватися у видимому спектрі світла, наприклад за допомогою камери смартфона. Сітківка криється на внутрішній частині ока, що ускладнює

запис. Око просвічується інфрачервоним світлом, а камера повинна бути розташована безпосередньо перед оком. Сканування сітківки ока є порівняно надокучливим процесом, потребує спеціального обладнання, а тому використання обмежується військовими та подібними цілями. Незручність запису даних є недоліком і найсильнішим активом водночас. Отримати несанкціонований доступ до сітківки дуже важко і майже неможливо при цьому залишатися непоміченим. Зображення райдужної оболонки ока або склери можна порівняно легко отримати за допомогою камер високої роздільної здатності на відстані декількох метрів [12].

Електрокардіограма - це добре відома технологія для медичної діагностики серцево-судинної системи, яка детально описана в наступному підрозділі. Форма ЕКГ-сигналу залежить від індивідуальної фізіології і є дискримінаційною біометричною характеристикою, що було продемонстровано в багатьох різних дослідженнях [13-18]. Вимірювання ЕКГ-сигналів можна вважати ненав'язливим, оскільки вимірювальні електроди можна інтегрувати у різні системи, наприклад можна вимірювати ЕКГ тримаючи смартфон у руках чи за допомогою різноманітних браслетів. У підрозділі 4.4 представлені можливі сфери застосування біометричних систем на основі ЕКГ.

Поведінкові характеристики - це біометричні дані, які формуються підсвідомими рухами у процесі відтворення якої-небудь дії. До поведінкових характеристик належать (рис. 1.2):

Хода. Біометричні системи на основі ходи здійснюють розпізнавання на основі даних акселерометра або відеозаписів. Системи на основі відеозаписів вимагають камер, які моніторять область, де слід виконувати розпізнавання ходи. Системи на основі акселерометра краще підходять для мобільного та постійного використання (вони більш гнучкі та доступні, оскільки більшість смартфонів вже мають акселерометри). Розпізнавання ходи на основі акселерометра є ненав'язливим та ефективним [4, 7]. За оцінками [12] рівень точності систем автентифікації на основі ходи складає понад 90%, але згадується, що поведінкові

атрибути, такі як хода, не містять достатньої дискримінаційної інформації для надійної автентифікації. Це пов’язано з тим, що на них впливає емоційний стан, стан здоров’я, харчові звички та старіння. Крім того, дана біометрична характеристика, як і будь-яка поведінкова характеристика, може бути зімітованою зловмисником.

Голос належить до поведінкових біометричних характеристик, хоча голосовий тракт людини є фізіологічною особливістю. Це пов’язано з тим, що для автентифікації використовується згенерований голосовим трактом сигнал. Людський голос зазнає постійних змін. Він змінюється з часом доби, а також з часом року або віком. На це також впливає стан здоров’я та емоційний стан ораторів [4, 19].

Клавіатурний почерк. Проведено багато досліджень у галузі динаміки натискання клавіш. Для розпізнавання використовуються такі параметри як час утримання клавіш, час між натисканням клавіш, час введення часто вживаних у мові послідовностей букв, тощо [20, 21]. Системи, які здійснюють розпізнавання із сенсорних екранів, мають у своєму розпорядженні ще більше параметрів, наприклад значення тиску на екран, прискорення та швидкість жестів, тощо [22]. Для традиційної одноразової автентифікації (наприклад, щоб отримати доступу до системи) клавіатурний почерк вважається недостатньо безпечним [4, 23]. Такі системи ефективно працюють у фоновому режимі та можуть використовуватися у комплексі з іншими біометричними характеристиками.

Рукописний підпис також можна використовувати для побудови систем автентифікації. Він використовувався для підписання документів, листів та контрактів задовго до того, як автентичність могла бути визначена алгоритмічно [24].

Для успішного використання сучасних біометричних систем у повсякденному житті такі системи повинні відповідати різноманітним вимогам. Основні категорії вимог до біометричних систем наведено нижче:

● ***Продуктивність***: продуктивність розпізнавання представляє найбільший інтерес при виборі біометричних систем. Біометрична система схильна до помилок, зазвичай кількість помилок ідентифікації і автентифікації є рівною.

● ***Вартість***: Вартість розгортання біометричної системи часто оцінюється від її прямих і непрямих компонентів. Прямий компонент включає в себе апаратні складові (сенсор, процесор, пам'ять) і програмні модулі. Непрямі витрати часто включають в себе установку системи, навчання/вимоги до технічного обслуговування і прийнятності для користувачів [25].

● ***Зручність користування***: Біометрична система повинна бути зручною для користувачів. Сенсори мають бути гігієнічними та не повинні впливати на здоров’я користувачів [25].

● ***Функціональна сумісність***: Оскільки біометричні системи все частіше використовують в широкому діапазоні застосувань, необхідно, щоб система була сумісна між різними біометричними технологіями (сенсори / алгоритми / апаратура). Біометричні система не буде ефективно функціонувати, якщо буде очікувати, що один і той же сенсор, одні і ті ж алгоритми, або ж робочі умови, завжди будуть доступні протягом всього терміну служби [25].

● ***Безпека***: Біометричні системи уразливі до потенційних порушень безпеки від фальсифікації та різноманітних атак зловмисників. Біометричні системи повинні забезпечувати високий рівень захисту до різних вразливостей, які виникають в результаті власних помилок або атак зловмисників [1].

Виходячи з вищевказаних міркувань, біометричні системи повинні бути простими у використанні, бути дешевими, легко інтегруватись у наявні системи безпеки, повинні бути захищеними та мати високу точність. Для надійного використання біометричні характеристики повинні підкорятися набору вимог [11]:

● ***Універсальність***: кожна особа повинна мати характеристику;

● ***Унікальність***: характеристика повинна бути різною для будь-яких двох осіб;

● ***Постійність***: характеристика повинна бути інваріантна в часі;

● ***Вимірюваність***: повинна бути можливість кількісного виміру характеристики.

Якщо система використовується як персональна система розпізнавання, додатково існують наступні види вимог [11]:

● ***Продуктивність***: швидкодія та точність біометричної системи, а також стійкість цих показників при різних екологічних та експлуатаційних факторах;

● ***Прийнятність***: прийнятність стосується бажання людей щоденно використовувати систему;

● ***Обман***: стосується можливостей хакерів-професіоналів, які використовують шахрайські методи, обманути біометричну систему.

У таблицях 1.1 та 1.2 резюмуються параметри головних біометричних характеристик відносно описаних вище вимог. Таблиці показують відмінність між основними типами біометрії. Вибір певної ознаки залежить від конкретного варіанту застосування.

Таблиця 1.1.

Порівняння біометричних характеристик, частина 1

	Око (райдужна оболонка)	**Відбиток пальця**	**Геометрія кисті та пальців**	**Підпис**	**Клавіатурний почерк**
Універсальність	Висока	Висока	Висока	Середня	Середня
Унікальність	Висока	Висока	Середня	Висока	Середня
Постійність	Висока	Висока	Середня	Низька	Середня
Вимірюваність	Середня	Висока	Висока	Низька	Висока
Продуктивність	Висока	Середня	Середня	Середня	Середня
Прийнятність	Низька	Висока	Висока	Висока	Висока
Обман	Висока	Низька	Низька	Низька	Низька
Ціна	Висока	Низька	Середня	Середня	Низька

Таблиця 1.2.

Порівняння біометричних характеристик, частина 2

	Голос	**ЕКГ**	**ДНК**	**Обличчя**	**Хода**
Універсальність	Висока	Висока	Висока	Висока	Середня
Унікальність	Висока	Висока	Висока	Висока	Середня
Постійність	Середня	Висока	Висока	Середня	Середня
Вимірюваність	Середня	Висока	Низька	Середня	Середня
Продуктивність	Середня	Висока	Висока	Середня	Середня
Прийнятність	Середня	Середня	Низька	Середня	Висока
Обман	Середня	Висока	Висока	Низька	Низька
Ціна	Середня	Низька	Висока	Середня	Середня

Сигнал електрокардіограми є однією із найновітніших та найперспективніших біометричних характеристик. ЕКГ відображає електричну активність серця в часі. Багато досліджень підтверджують ефективність використання цього сигналу для задач ідентифікації та автентифікації [26, 27]. ЕКГ як біометрична характеристика має ряд переваг порівняно із іншими характеристиками:

• ***Універсальність***. Універсальність забезпечується тим, що фізіологічно люди відрізняються один від одного та відповідно електричні сигнали, які проходять через тіло, будуть відрізнятися.

• ***Постійність***. Вимога до постійності ЕКГ також задовольняється, оскільки основна структура таких сигналів є інваріантною протягом тривалого періоду часу. Підтвердження постійності електрокардіограми в часі є частиною дисертаційного дослідження і детально висвітлено у підрозділі 4.3. Крім того, серце дуже добре захищене в організмі людини, тому фактори навколишнього середовища не можуть мати значного впливу на його діяльність, на відміну від інших біометричних характеристик.

• ***Обман***. Ще однією суттєвою перевагою застосування ЕКГ у біометричних системах є їх стійкість проти застосування фальсифікованих даних. Формою сигналу електрокардіограми керує автономна нервова система, під впливом комбінації симпатичних та парасимпатичних факторів. Це свідчить про те, що сигнал кожного серцебиття є відносно різним, тому його важко імітувати або відтворити. Більше того, у порівнянні з іншими біометричними характеристиками, такими як відбитки пальців, набагато складніше викрасти чиюсь електрокардіограму.

• ***Низькі обчислювальні затрати***. Біометричні системи на основі ЕКГ-сигналу можуть бути імплементовані на малопотужних пристроях (наприклад, смартфонах).

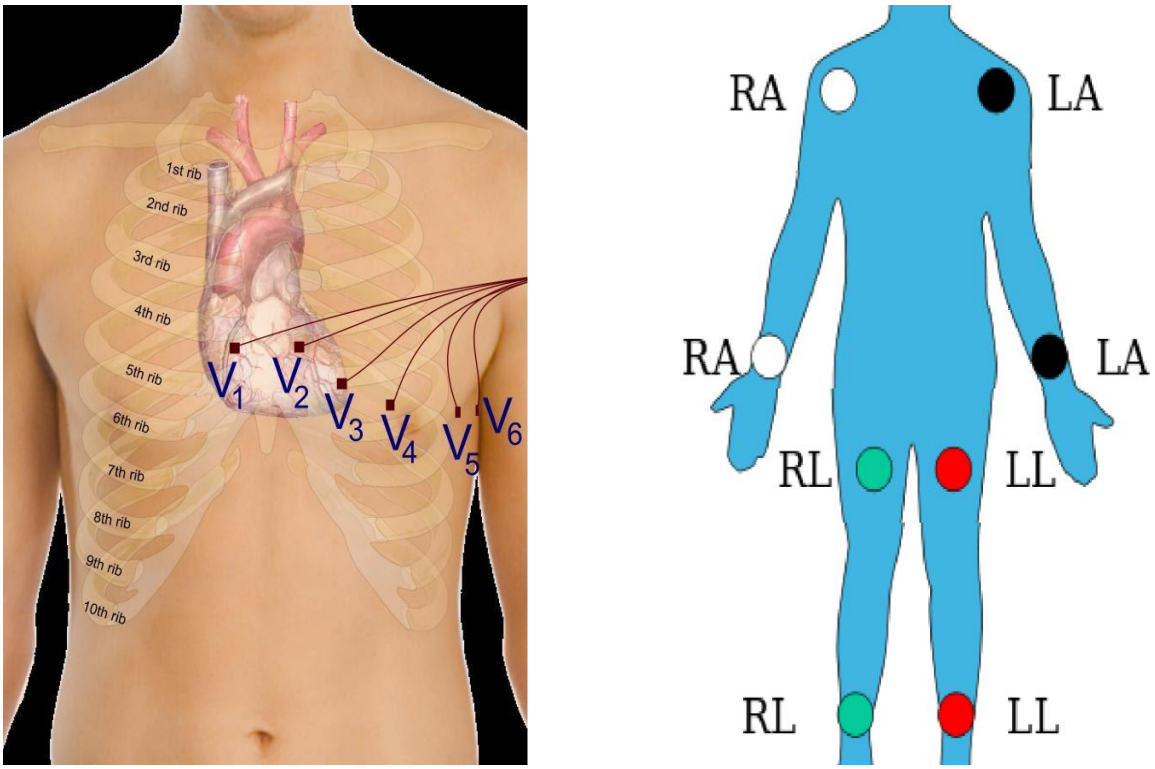
• ***Вимірюваність***. У зв’язку з останніми досягненнями в області біомедичної апаратури, вимірювання ЕКГ-сигналів може здійснюватися на грудях, використовуючи сорочку з вбудованою електронікою, на шиї, використовуючи намисто з підвіскою, на пальцях рук і долоні з допомогою давачів (наприклад, браслету).

• ***Виявлення життєздатності***. Біометричні системи на основі ЕКГ можуть тривіально перевіряти життєздатність суб'єкта. Інші біометричні характеристики, такі як райдужна оболонка ока чи відбиток пальця, потребують додаткової обробки, щоб встановити чи належить виміряний сигнал живому суб'єкту. Крім того, біометричні системи на основі ЕКГ можуть здійснювати медичну діагностику, виявляти чи передбачати різноманітні серцеві захворювання.

1.2. Електрокардіограма як біометрична характеристика

Електричні потенціали, які генеруються серцем розповсюджуються по всьому тілу. Різниця потенціалів може бути визначена шляхом вимірювання напруги між електродами розміщеними на поверхні тіла. Для вимірювання ЕКГ-слід використовувати малі струми (в ідеалі взагалі не подавати струм, оскільки струм деформує електричне поле, що створює різницю потенціалів) [28]. Різниця

потенціалів вимірюється при розташуванні двох електродів на різних лініях електричного поля продукованого серцем. Різні пари електродів розміщені в різних місцях на тілі дають різні напруги, що зумовлено просторовою залежністю електричного поля [29]. Отже, важливо мати певні стандартні позиції для вимірювання ЕКГ. Кінцівки є прекрасними орієнтирами розміщення електродів для вимірювання ЕКГ. У таблиці 1.3 представлено різноманітні позиції розміщення електродів на тілі, напругу між якими можна використовувати в якості біометричної характеристики. Як видно з рис. 1.3, зазвичай на тілі розміщують десять електродів: шість на грудній клітці і по одному на кожній кінцівці.



(а) розміщення електродів в міжребер'ї. (б) розміщення електродів на тілі

Рис. 1.3. Основні позиції розміщення електронів [29]

Ділянка між електродами двох кінцівок називається відведенням. Розрізняють: стандартні (I, II, III), стандартні посилені та грудні відведення. Зв'язки між електродами стандартних відведень утворюють трикутник Ейнтговена (рис.1.4), де серце являє собою нульову точку.

Таблиця 1.3.

Позиції електродів

Назва електроду	Розміщення електроду
RA	На правій руці, уникаючи товстих м'язів.
LA	У тому ж місці, де поміщений RA, але на лівій руці.
RL	На правій нозі, уникаючи литкового м'яза.
LL	У тому ж місці, де поміщений RL, але на лівій нозі.
V_1	У четвертому міжребер'ї (між ребрами 4 і 5) правіше грудної клітки.
V_2	У четвертому міжребер'ї (між ребрами 4 і 5) лівіше грудної клітки.
V_3	Між електродами V_2 та V_4.
V_4	У п'ятому міжребер'ї (між ребрами 5 і 6)
V_5	На рівні V_4, в лівій передній пахвовій лінії.
V_6	На рівні V_4 та V_5, в середній пахвовій лінії.

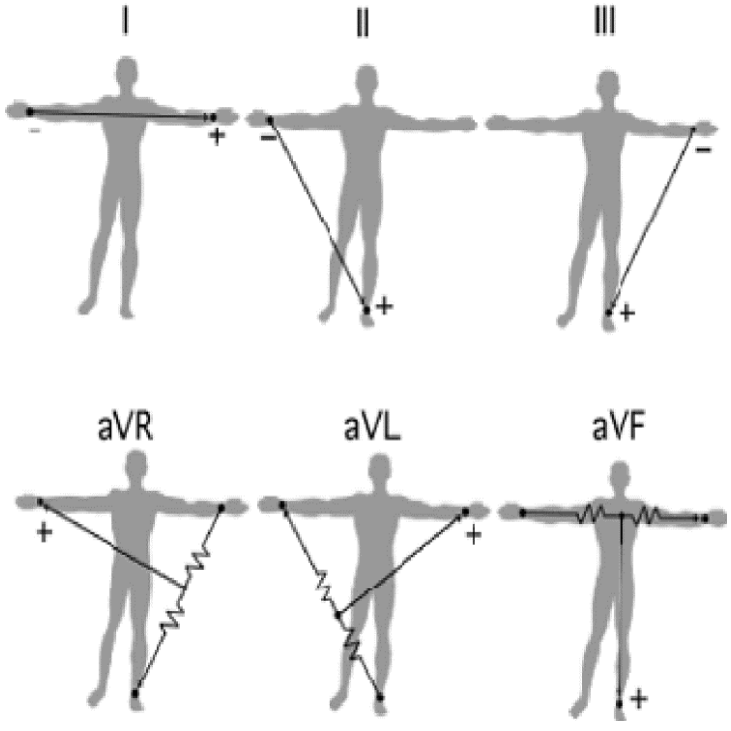


Рис. 1.4. Розміщення електродів на кінцівках для запису стандартних відведень (вгорі) та доповнених стандартних відведень (знизу)

Сучасні методики показують зменшення числа електродів необхідних для діагностики людини. Оскільки дана робота сфокусована на використанні ЕКГ для

побудови системи автентифікації особистості, а не проведення клінічних досліджень, то для автентифікації достатньо використовувати ЕКГ-сигнал I стандартного відведення Ейнтговена [16], отриманий з допомогою трьох електродів (рис.1.5).

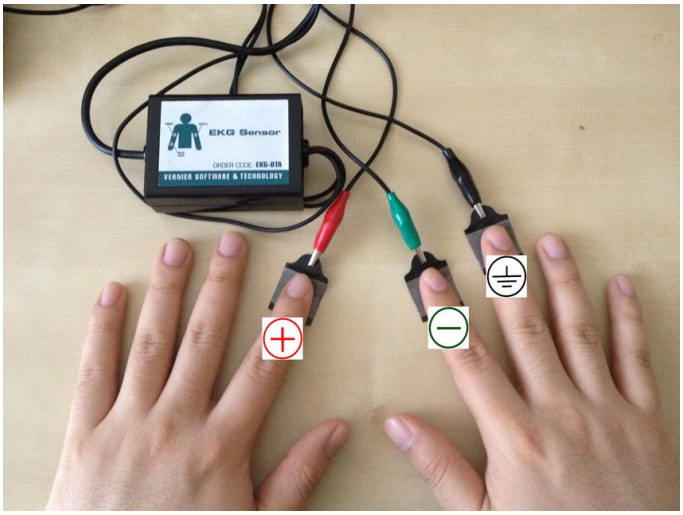


Рис. 1.5. Розміщення електродів при розпізнаванні

На ЕКГ-сигнал впливають різноманітні фізичні, емоційні та навіть зовнішні фактори. Ці обмеження роблять розпізнавання особистості по ЕКГ-характеристиці непростим завданням. Не зважаючи на це, електрокардіограма постійна і завжди доступна характеристика.

На рис. 1.6 наведено форму стандартної ЕКГ-хвилі. Кожна ЕКГ-хвиля репрезентує активність протягом одного серцебиття. Форма ЕКГ-хвилі при синхронному записі з різних ділянок тіла буде різною. Зубці та хвилі ЕКГ характеризують величину, напрямок та локалізацію потенціалів серця. Відрізки ЕКГ, що знаходяться між зубцями, називаються сегментами, а відрізки, які складаються з сегмента та наступного за ним зубця – інтервалами [30].

Кожному сегменту хвилі відповідає одна фаза серцебиття. Кількість серцебиттів у хвилину зазвичай називають частотою серцебиття або пульсом. Тривалість між послідовними серцевими скороченнями відома як RR-інтервал. Для здорових людей проміжок часу між послідовними серцебиттями змінюється навіть в умовах відпочинку. Це називається варіабельністю серцевого ритму і продемонстровано на рис. 1.7.

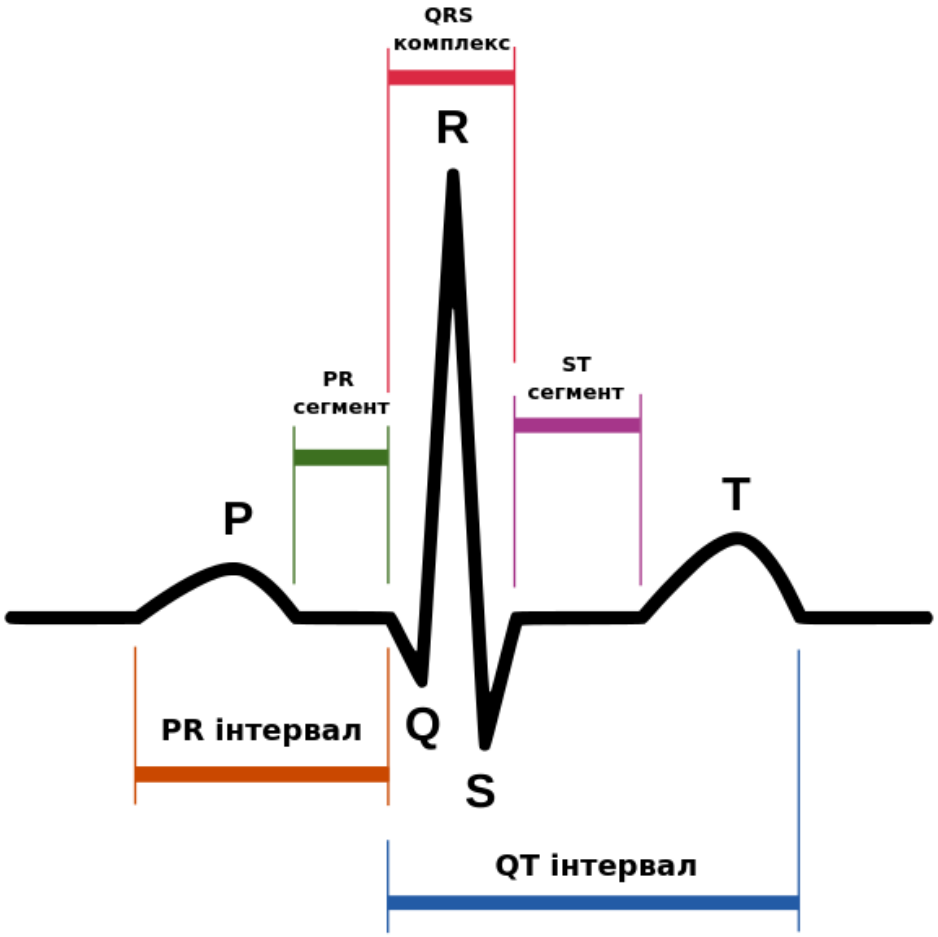


Рис. 1.6. Форма стандартної ЕКГ-хвилі [33]

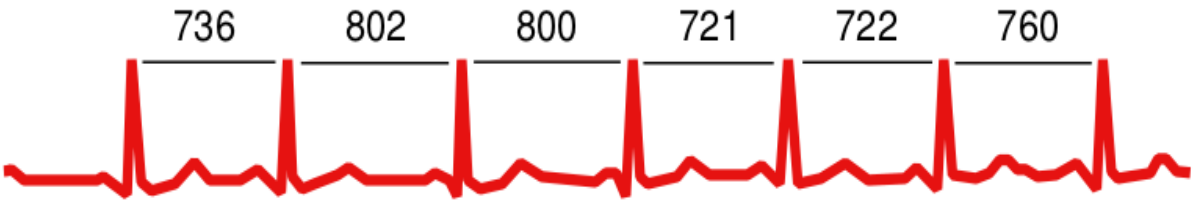


Рис. 1.7. Ілюстрація ЕКГ-сигналу здорової людини. RR-інтервали вказані у мілісекундах. Частота серцевих скорочень варіюється

P-зубець. Серцебиття ініціюється синоатріальним вузлом (його ще називають водієм серцевого ритму), котрий нав’язує свій ритм нижче лежачим відділам серця. Він спонтанно генерує електричний імпульс, який при розповсюдженні через серце викликає серцеве скорочення. Згідно з [31] та [32], тривалість P-зубця, як правило, триває менше 120 мс і його спектр зосереджений на проміжку від 10 Гц до 15 Гц.

QRS-комплекс. Після Р-зубця зазвичай можна спостерігати три піки підряд. Зубці Q, R і S відповідають деполяризації шлуночків, які ініціюють скорочення серця. Як показано на малюнку 1.6, від'ємний зубець Q передує додатному зубцю R, за яким слідує від'ємний зубець S. Тривалість QRS - комплексу коливається від 60 до 100 мс. Його спектр становить від 10 Гц до 40 Гц [31].

Т-зубець відповідає процесам припинення збудження шлуночків. Він триває близько 160 мс і, залежно від частоти серцевих скорочень, з'являється за 80 - 120 мс після QRS - комплексу. Він може бути додатним, від'ємним та двофазним. У роботі [31] стверджується, що ST - сегмент стає коротшим зі збільшенням частоти серцевих скорочень.

U-зубець. Інколи за Т-зубцем може слідувати U-зубець – змінюваний і невеликий.

1.3. Аналіз відомих підходів для побудови біометричної системи автентифікації на основі ЕКГ-сигналу

Біометричні системи розпізнавання на основі ЕКГ-сигналу почали активно розвиватися на початку 2000-их років [34]. Цьому сприяв стрімкий розвиток комп'ютерного обладнання та програмних алгоритмів. В даному підрозділі здійснено огляд наукових робіт, які демонструють різноманітні підходи до побудови систем розпізнавання. Загалом, залежно від способу екстракції ознак, розрізняють підходи на основі характерних точок (fiducial points) та підходи на основі не характерних точок (non-fiducial point).

У підходах на основі характерних точок, ознаки, як правило, виділяються із зубців та інтервалів ЕКГ-сигналу зображеного на рис. 1.6. Ознаками можуть бути відстані між точками, кути, площі під кривими, амплітуди, тощо. Вилучення ознак з характерних точок є найбільш інтуїтивним підходом, оскільки саме так у медицині лікарі оцінюють сигнали ЕКГ [35]. На рис. 1.8 показані деякі з ознак, котрі використовувались у дослідженні [36].

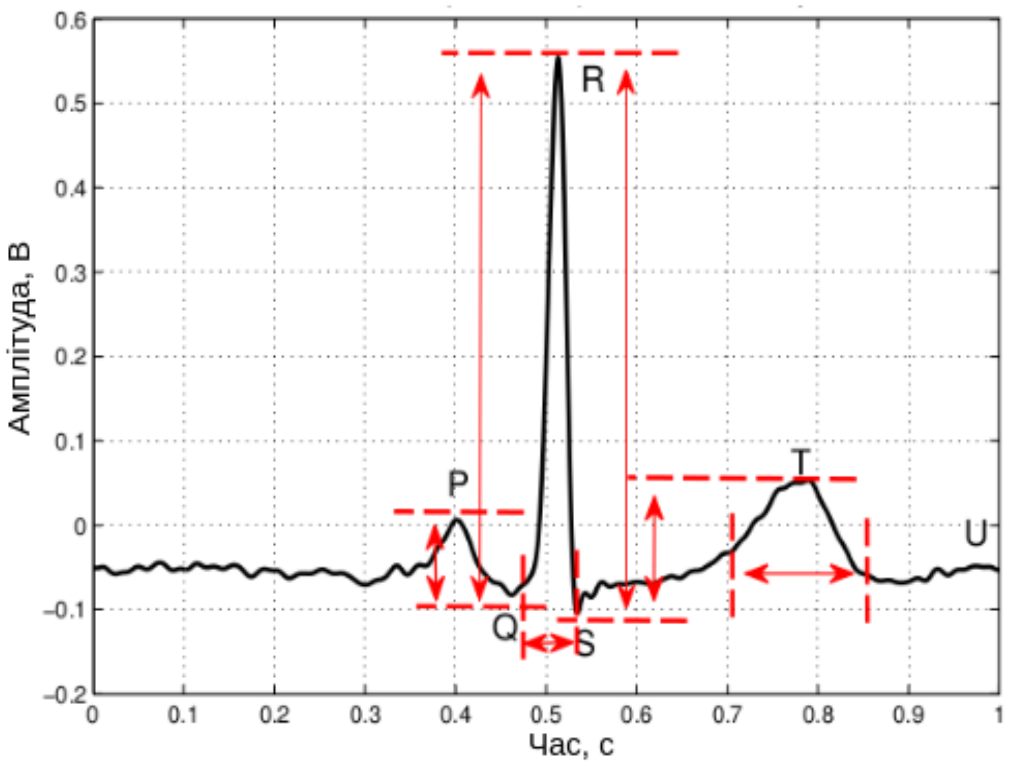


Рис. 1.8. Деякі з ознак використаних у [36], котрі демонструють підхід на основі характерних точок

Однією з перших робіт у галузі біометрики на основі ЕКГ-сигналу є робота [34]. Автори представили можливість високоточної класифікації 20 осіб. У роботі використовувались сигнали з 12-ти відведень ЕКГ. З кожної ЕКГ-хвилі виділялося 30 ознак, які описують її зубці та інтервали (P, Q, R, S, T та U). Загалом для класифікації використовувалось 360 ознак (по 30 ознак для кожного з 12-ти відведень). Також в даній роботі висунуто припущення, що одного відведення достатньо для розпізнавання особистості. Це припущення базується на тому, що існує сильна кореляція між ознаками з різних відведень. Найкращий результат був отриманий, коли використовувалося лише 10 ознак із 360.

У роботі [17] було застосовано стадію попередньої обробки для усунення низькочастотних флуктуацій, перешкод електромережі та високочастотних наведень. Інформативні ознаки було вилучено з Q, R, S, T-хвиль. Щоб врахувати варіабельність серця, ЕКГ-хвилі нормалізувались за тривалістю QT-інтервалу. Розпізнавання проводилось у два етапи: перевірка шаблонів та класифікація за допомогою нейронної мережі на основі рішень. Перевірка шаблонів базувалася

на обчисленні кореляції. Класифікація за допомогою нейронної мережі проводилась, якщо коефіцієнт кореляції > 0,85. Автори припустили, що цей підхід є ефективним з точки зору обчислювальних ресурсів, має прийнятний час навчання та має етап попереднього скринінгу за допомогою перевірки шаблонів, що підвищує точність.

У роботі [14] для знешумлення ЕКГ-сигналу використовувався смуговий фільтр із частотами зрізу 1,1-40 Гц. Характерні точки було вилучено з P, Q, R, S і T-хвиль. Додатково вилучались інші чотири точки: T'(кінець хвилі T), S'(кінець хвилі S), P'(кінець хвилі P) і L'(початок хвилі P). Отже, автори загалом використовували дев'ять характерних точок. З цих дев'яти характерних точок було обчислено п'ятнадцять відрізків для створення вектора ознак. Кожен з відрізків описує ЕКГ-сигнал на темпоральній площині. Дванадцять з п’ятнадцяти ознак використовувались у більшості експериментів. Класифікація здійснювалася за допомогою лінійного дискримінантного аналізу.

У роботі [37] досліджувалися темпоральні та спектральні характеристики ЕКГ-сигналу. Дискретне хвилькове перетворення застосовувалося для усунення шумів та обчислення тридцяти чотирьох характерних точок. Для усунення ефекту порушення серцевого ритму застосовувалась нормалізація амплітудних та темпоральних характеристик. За допомогою аналізу головних компонент розмірність ознак зменшувалась до тринадцяти. Автори стверджують, що вони покращили продуктивність та зменшили складність, використовуючи зменшення розмірності. Крім того, для класифікації використовувався метод опорних векторів.

У роботі [36] для знешумлення ЕКГ-сигналу використовувався смуговий фільтр Баттерворта з частотами зрізу 1-40 Гц. Для екстракції характерних точок використовувався вирівняний за R-піком QRS-комплекс. QRS-комплекс розглядався як вікно тривалістю 800 мілісекунд із центром у піку R. Використовувались темпоральні атрибути, подібно до [14], проте ознаки було нормалізовано за тривалістю P’T’. Разом із темпоральними ознаками

використовувалось шість амплітудних ознак. Класифікація здійснювалася за допомогою лінійного дискримінантного аналізу. У статті висловлено припущення, що шість амплітудних атрибутів значно покращили результати.

У роботі [38] для зменшення розмірності ознак використовувались різні методи: аналіз головних компонент, лінійний дискримінантний аналіз, коефіцієнт отримання інформації, тощо. Для попередньої обробки ЕКГ-сигналу застосовувався смуговий фільтр Баттерворта з частотами зрізу 0,5-40 Гц. Детальна інформація про екстракцію характерних точок представлена в роботі [39]. Двадцять вісім ознак були отримані з характерних точок, де 19 з них були темпоральними, шість з них були амплітудними, та три з них - кутовими. До темпоральних ознак застосовувалась нормалізація за тривалістю серцебиття. До амплітудних ознак застосовувалась нормалізація відносно амплітуди R - піку. Класифікація здійснювалася за допомогою нейронних мереж із радіальними базисними функціями. Автори дійшли висновку, що амплітудні та кутові характеристики були більш дискримінаційними, ніж темпоральні.

Основним недоліком використання підходів на основі характерних точок є необхідність знаходити характерні точки в ЕКГ. Помилки при визначенні характерних точок саботують ефективність біометричної системи.

Підходи на основі не характерних точок є комплексними підходами, які розглядають ЕКГ-сигнал або окреме серцебиття в цілому. В даних підходах можуть використовуватися характерні точки (наприклад, R піки) для вирівнювання та сегментації сигналу ЕКГ, проте для розпізнавання використовується "сирий" ЕКГ-сигнал [40].

Підходи на основі не характерних точок можуть бути розділені на два основні потоки: ті, які потребують сегментації та вирівнювання ЕКГ-сигналу [41, 42, 43] та ті, які не вимагають жодної інформації про ЕКГ-сигнал [31, 44].

У роботі [41] для розпізнавання використовувалось одне з дванадцяти відведень ЕКГ. Для класифікації використовувався спектр ЕКГ-сигналу.

Класифікатор було побудовано на основі нейронної мережі зворотного поширення похибки.

У роботі [42] для знешумлення ЕКГ-сигналу використовувалось дискретне хвилькове перетворення. Авторами показано, що частотно-часове представлення працює краще ніж інші методи фільтрації. Коефіцієнти автокореляції використовувались як ознаки для класифікації. Також застосовувались лінійні та нелінійні методи зменшення розмірності ознак. Зокрема, використовувались метод головних компонент, аналіз основних компонент та лінійний дискримінантний аналіз. Зменшений за розмірами вектор ознак подавався на вхід класифікатора на основі методу опорних векторів.

Дискретне хвилькове перетворення також використовувалось у роботі [43]. Попередня обробка сигналу базувалася на високочастотних та низькочастотних фільтрах. Високочастотний фільтр з частотою зрізу 0,5 Гц було використано для усунення дрейфу базової лінії, тоді як фільтр низьких частот з частотою зрізу 45 Гц застосовувався для усунення високочастотних шумів, таких як перешкоди лінії електропередач, тощо. Детекцію QRS - комплексу було здійснено за допомогою методу, описаного в [45]. Сегмент серцебиття ЕКГ розглядався протягом 128 вибірок із даних, виміряних із частотою дискретизації 250 Гц. Кожному сегменту відповідали 44 вибірки до R - піку та 84 вибірки після. Кожні чотири послідовні сегменти використовувались для обчислення ознак. Хвилькові коефіцієнти обчислювались для кожного з чотирьох сегментів. Для розрахунку хвилькових коефіцієнтів було використано метод Хаара з дев'ятьма рівнями розкладання. Для класифікації використовувалася евклідова відстань.

На відміну від попередніх робіт, де для сегментації використовувалися характерні точки ЕКГ-сигналу, у роботі [44] не використовувалась інформація про характерні точки. Автори визнали, що не існує остаточного і загальновизнаного правила для визначення початку та кінця ЕКГ-серцебиття. У даній роботі вперше для фільтрації було застосовано смуговий фільтр з частотами зрізу 0,5-40 Гц. Далі обчислювалась автокореляція з довжиною вікна

більшою ніж тривалість одного серцебиття. Автокореляція нормалізовувалась за максимальним значенням для усунення факторів зміщення. Для зменшення розмірності ознак до коефіцієнтів автокореляції застосовувалось дискретне косинусне перетворення. Класифікація здійснювалась за допомогою методу максимальної правдоподібності та евклідової відстані.

Робота [31] також орієнтована в тому ж напрямку, що і робота [44]. У ній використовувався фільтр Баттерворта з частотами зрізу 1-40 Гц. ЕКГ-сигнал розбивався на сегменти за допомогою вікон, що не перекриваються. Для цих вікон розраховувалась автокореляція. Для зменшення розмірності використовувались дискретне косинусне перетворення та лінійний дискримінантний аналіз. Класифікація здійснювалась за допомогою евклідової відстані.

Для побудови біометричних систем також може бути використано гібридний підхід. При такому підході для класифікації одночасно використовуються набори ознак отримані з характерних точок та з “сирого” ЕКГ-сигналу.

Загалом у описаних вище системах розпізнавання спільними є наступні структурні елементи: фільтрація, нормалізація, екстракція ознак, зменшення розмірності ознак та класифікація. У таблиці 1.4 представлено короткий огляд розглянутих вище робіт.

У даному дисертаційному дослідженні для побудови біометричної системи автентифікації на основі ЕКГ-сигналу буде використано підхід на основі не характерних точок (non-fiducial points). Стрімкий розвиток алгоритмів машинного навчання та обчислювальної техніки зробив можливим використання “сирого” ЕКГ-сигналу для побудови високоточних та продуктивних систем розпізнавання людини. При такому підході ми також уникаємо помилок, які можуть виникнути при екстракції характерних точок.

Таблиця 1.4.

Коротка характеристика представлених вище публікацій

Автори	**Екстракція ознак**	**Класифікація**	**Кількість суб'єктів**	**Точність, %**
S. A. Israel та інші [14]	Темпоральні ознаки отримані з характерних точок	Лінійний дискримінантний аналіз	29	100
T. Shen та інші [17]	Темпоральні та амплітудні ознаки отримані з характерних точок	Перевірка шаблонів та нейронні мережі	20	100
F. Agrafioti та інші [31]	Коефіцієнти автокореляції	Евклідова відстань	27	100
L. Biel та інші [34]	Темпоральні та амплітудні ознаки отримані з характерних точок	М'яке незалежне моделювання аналогії класів	20	100
Y. Wang та інші [36]	Темпоральні та амплітудні ознаки отримані з характерних точок	Лінійний дискримінантний аналіз	29	100
A. Kaveh та інші [37]	Спектральні та темпоральні ознаки отримані з характерних точок	Метод опорних векторів	18	98,7
M. Tantawi та інші [38]	Темпоральні, кутові та амплітудні ознаки отримані з характерних точок	Нейронні мережі	14	96±2
C. Chen та інші [41]	Спектральні коефіцієнти	Нейронні мережі	19	≥90
M. Hejazi та інші [42]	Коефіцієнти автокореляції	Метод опорних векторів	52	88.18± 0.82
C.-C. Chiu та інші[43]	Коефіцієнти хвилькового перетворення	Евклідова відстань	45	100
K. N. Plataniotis та інші[44]	Коефіцієнти автокореляції	Метод максимальної правдоподібності та евклідова відстань	14	100

1.4. Формулювання задач дисертаційного дослідження

Аналіз сучасних методів та засобів біометричної автентифікації показав, що електрокардіограма є однією із найновітніших та найперспективніших біометричних характеристик. Вона має ряд переваг порівняно із іншими біометричними характеристиками, серед яких: універсальність, постійність, стійкість до обману, простота у вимірюванні та низькі обчислювальні затрати при розпізнаванні.

Одним із завдань дисертаційного дослідження є розроблення концепції покращення характеристик для біометричних систем описаних у попередньому підрозділі.

На шляху до масового використання біометричних систем на основі електрокардіограми стоять такі бар'єри, як якість виміряного сигналу та варіабельність серцевого ритму. Щоб зробить біометричну систему зручною у використанні, ЕКГ-сигнал доводиться вимірювати сухими електродами із пальців лівої та правої руки. Такий спосіб реєстрації ЕКГ-сигналу негативно впливає на його якість.

Серцевий ритм – це характеристика фізіологічного стану людини, яка розраховується як число скорочень серця за хвилину. Дана характеристика є змінною у часі та залежить від різноманітних фізіологічних та психологічних чинників. Варіабельність серцевого ритму безпосередньо впливає на форму ЕКГ-сигналу, а також може бути використана у якості дискримінативної ознаки на етапі класифікації.

Для подолання представлених вище бар'єрів необхідно вирішити наступні завдання:

- імплементувати та апробувати алгоритми для виявлення та виправлення артефактів у ЕКГ-сигналах;
- здійснити опрацювання та верифікацію алгоритмів темпоральної нормалізації ЕКГ-сигналу.

Апробація та валідація розроблених компонентів біометричної системи потребує репрезентативного набору електрокардіограм. Огляд та аналіз існуючих наборів ЕКГ-сигналів показав, що такі набори, зазвичай, відсутні у публічному доступі та містять достатньо велику підмножину електрокардіограм хворих людей (оскільки є “перекваліфікованими” для вирішення проблем розпізнавання людини із задач розпізнавання захворювань). Тому одним із завдань дисертаційного дослідження є вимірювання та підготовка власного набору ЕКГ-сигналів, який буде використано у даній роботі.

Не менш важливою є демонстрація постійності ЕКГ-сигналу в часі, для цього необхідно здійснити дослідження часової інваріантності електрокардіограми як біометричної характеристики.

Щоб показати можливість імплементації біометричної системи ЕКГ-автентифікації на пристроях з невеликими обчислювальними ресурсами необхідно реалізувати біометричну систему на основі мікрокомп'ютера Raspberry Pi. Верифікацію імплементованої біометричної системи необхідно здійснити дослідивши її швидкодію.

Наостанок, необхідно розробити рекомендації для практичного використання біометричних систем автентифікації на основі ЕКГ-сигналу.

Висновки до розділу 1

1. Проведено огляд та дослідження літературних джерел за темою дисертаційної роботи, зокрема, проаналізовано основні режими роботи і подано порівняльну характеристику сучасних систем біометричної автентифікації. Також проведено порівняння найпоширеніших біометричних маркерів за допомогою формалізованих критеріїв.

2. Представлено детальний опис електрокардіограми як біометричного маркера в системах розпізнавання, показано його переваги і проблеми на шляху практичного застосування в системах автентифікації.

3. Проаналізовано відомі підходи опрацювання ЕКГ-сигналу на основі виділення характерних точок та на основі інтелектуального аналізу повного набору вибірок ЕКГ-сигналу.

4. За результатами проведеного аналізу сформульовано науково-прикладну проблему та визначено відповідні завдання дисертаційного дослідження, щодо удосконалення методів та засобів біометричної автентифікації на основі електрокардіограми.

РОЗДІЛ 2. КОНЦЕПЦІЯ ПОКРАЩЕННЯ ХАРАКТЕРИСТИК СИСТЕМ БІОМЕТРИЧНОЇ АВТЕНТИФІКАЦІЇ ЗА ЕКГ -СИГНАЛОМ

2.1. Особливості процесу автентифікації за сигналом ЕКГ

Ґрунтуючись на огляді існуючих підходів побудови біометричних системи автентифікації на основі ЕКГ-сигналу було узагальнено структурну схему таких систем (рис. 2.1).

Вимірювання ЕКГ-сигналу здійснюється електродами та допоміжною електронікою, за допомогою яких вимірюється зміна різниці потенціалів пальців правої та лівої руки, що відображають роботу серця. Цей спосіб реєстрації електрокардіограми відповідає першому стандартному відведенню, яке використовують в медичній діагностиці.

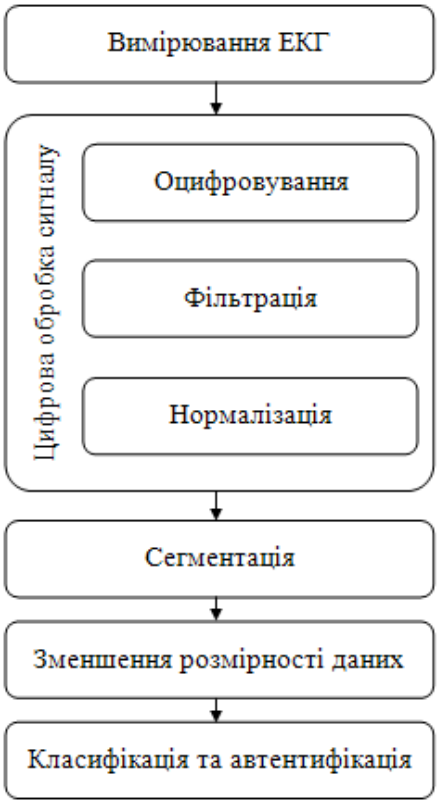


Рис. 2.1. Етапи опрацювання ЕКГ-сигналу в системі біометричної автентифікації

У ланці цифрової обробки сигналу виміряний ЕКГ-сигнал оцифровується і піддається фільтрації та нормалізації за амплітудою.

Далі оброблений сигнал піддається сегментації, застосовується зменшення розмірності даних та здійснюється автентифікація за допомогою одного з алгоритмів класифікації.

Нижче подано детальний опис для кожного із елементів наведеної вище структурної схеми.

2.1.1. Вимірювання ЕКГ-сигналу

Вимірювання ЕКГ-сигналу здійснюється за допомогою мікросхем Arduino та e-Health Sensor Shield V2.0, які вимірюють зміну різниці потенціалів пальців правої та лівої руки (рис. 2.2). Цей спосіб реєстрації електрокардіограми відповідає першому стандартному відведенню, яке використовують в медичній діагностиці.

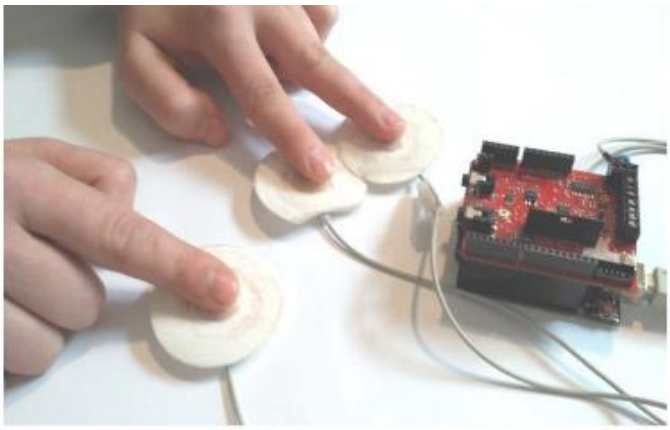


Рис. 2.2. Апаратура для вимірювання ЕКГ

Мікросхема Arduino — апаратна обчислювальна платформа, основними компонентами якої є плата мікроконтролера з елементами вводу/виводу та середовище розробки Processing/Wiring на мові програмування, що є спрощеною підмножиною C/C++. Arduino може використовуватися як для створення автономних інтерактивних об'єктів, так і підключатися до програмного забезпечення, яке виконується на комп'ютері. [46]

Мікросхема e-Health Sensor Shield V2.0 дає змогу використовувати Arduino в біометричних і медичних цілях. Моніторинг тіла можна здійснювати за допомогою 10 різних датчиків: пульсу, рівня кисню в крові, витрати повітря, температури тіла, вимірювати електрокардіограму (ЕКГ), рівень глюкози в крові,

шкірно-гальванічну реакцію, артеріальний тиск (тонометром), положення пацієнта (акселерометр) і м’язову активність за допомогою електроміограми (рис. 2.3).

Рис. 2.3. Мікросхема e-Health Sensor Shield V2.0 з підключеними давачами

Ця інформація може бути використана для моніторингу стану пацієнта в режимі реального часу або для отримання даних для подальшого аналізу для медичного діагностування. Зібрану біометричну інформацію можна бездротово надіслати за допомогою будь-якого з 6 доступних варіантів підключення: Wi-Fi, 3G, GPRS, Bluetooth, 802.15.4 та ZigBee залежно від архітектури системи. [47].

На рис. 2.4 наведено підключення електродів до мікросхеми.

Рис. 2.4. Підключення електродів до мікросхеми e-Health Sensor Shield V2.0

2.1.2. Цифрова обробка сигналу

Аналоговий сигнал ЕКГ виміряний мікросхемою e-Health Sensor Shield V2.0 передається на аналогові входи мікросхеми Arduino Uno. За допомогою вбудованого модуля аналого-цифрового перетворення (АЦП) здійснюється оцифрування аналогового сигналу ЕКГ. 10-ти бітний АЦП здійснює оцифрування аналогового сигналу в межах 0 - 5 В, роздільна здатність такого АЦП 4.9 мВ. Також на Arduino задається значення частоти дискретизації – параметр, який визначає скільки вибірок буде отримано за одну секунду вимірювань (в побудованій системі частота дискретизації дорівнює 277 Гц).

Далі оцифрований сигнал ЕКГ (рис. 2.5) піддається фільтрації. Виміряний сигнал ЕКГ може містити спотворення такого характеру:

- Низькочастотні флуктуації, викликані рухами пальців по електродах під час вимірювання ЕКГ;
- Мережеву заваду частотою 50 Гц та її гармоніки;
- Високочастотні складові зумовлені різноманітними наведеннями на мікросхемах.

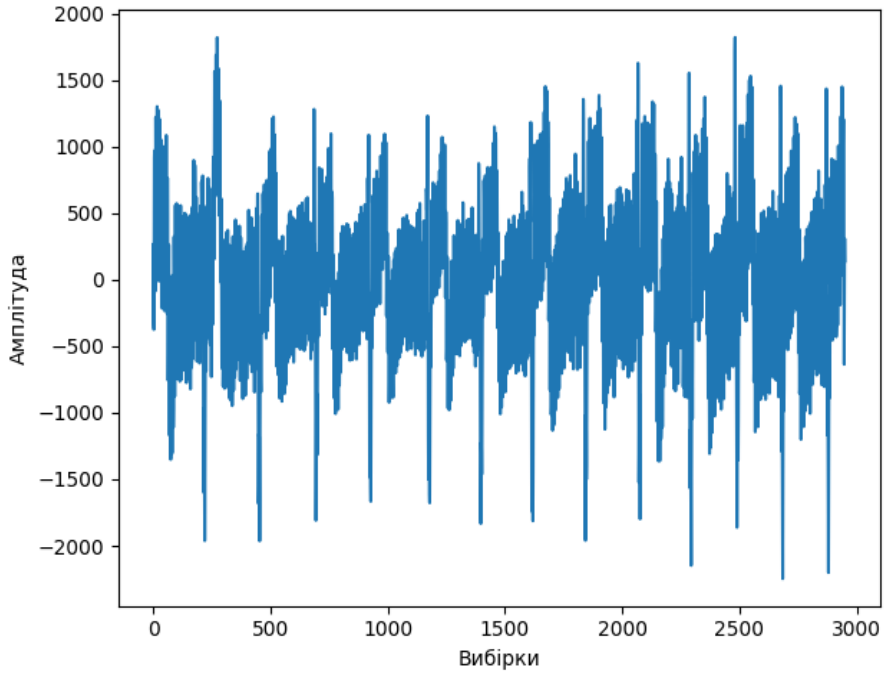


Рис. 2.5. Оцифрований сигнал ЕКГ

Після аналізу спектру оцифрованих сигналів ЕКГ (рис. 2.6) було прийнято проектувати смуговий фільтр Баттерворта із смугою пропускання 4 - 35 Гц, таким чином вдасться відфільтрувати як низькочастотні флуктуації, так і мережеві та високочастотні наведення.

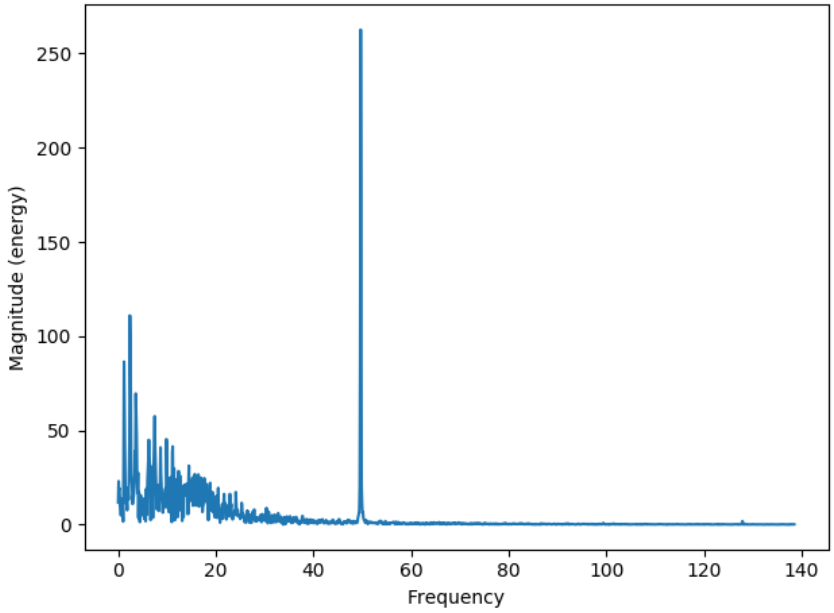


Рис. 2.6. Спектр оцифрованого вхідного сигналу

Основними характеристиками будь-якого фільтру є його амплітудно-частотна (АЧХ) та фазо-частотна (ФЧХ) характеристики. Вони показують, який вплив фільтр робить на амплітуду і фазу різних гармонік оброблюваного сигналу.

Для збереження форми сигналу було спроектовано фільтр зі скінченною імпульсною характеристикою. Такий фільтр ще називають нерекурсивним через відсутність зворотного зв'язку. Знаменник передавальної функції такого фільтра — це константа.

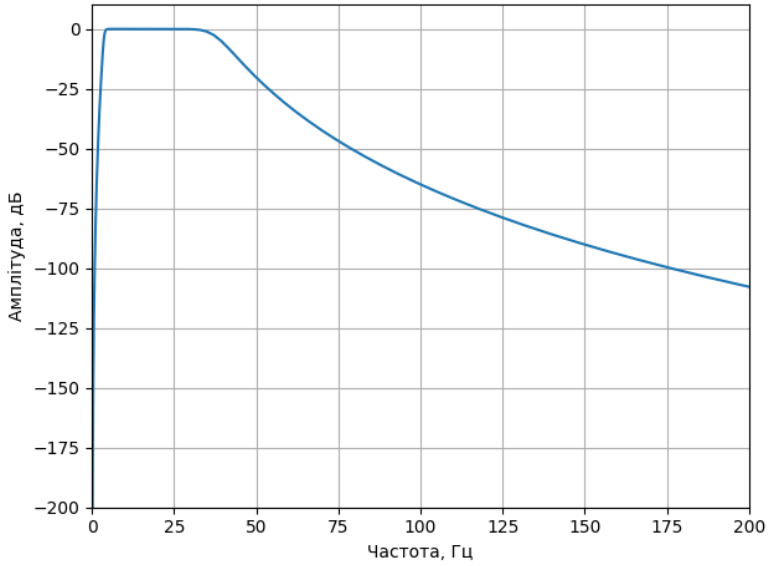


Рис. 2.7. АЧХ спроектованого фільтра

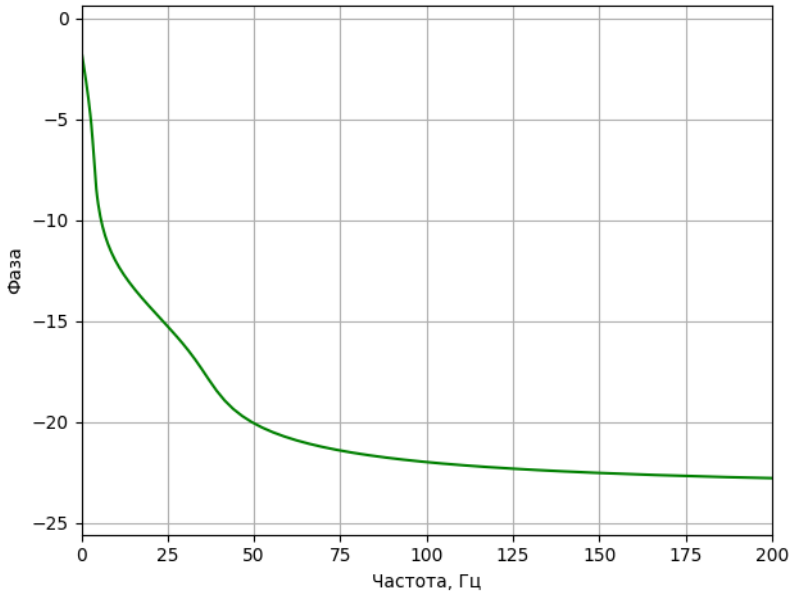


Рис. 2.8. ФЧХ спроектованого фільтра

На рис. 2.7 та 2.8 наведено АЧХ та ФЧХ спроектованого фільтра. Фільтр задовольняє поставлені при проектуванні вимоги та може бути використаним в біометричній системі автентифікації.

На рис. 2.9 наведено часові діаграми вхідного зашумленого та вихідного сигналу фільтра. Сигнал з виходу фільтра піддається нормалізації.

а) б)

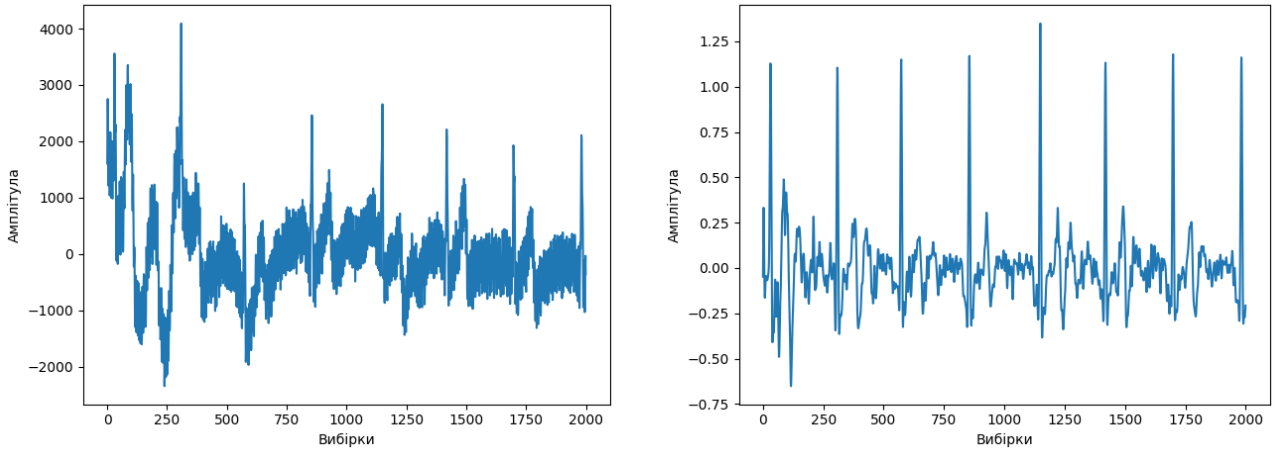


Рис. 2.9. ЕКГ-сигнал до (а) та після (б) фільтрації

Нормалізація є загальноприйнятою технікою попередньої обробки, що використовується у багатьох підходах машинного навчання [48]. В даній системі використовується z-score нормалізація. Нормалізація ознак перемасштабовує значення кожної ознаки таким чином, щоб ознаки мали властивості стандартного нормального розподілу із нульовим середнім значенням та одиничною дисперсією. Z-score нормалізацію можна здійснити шляхом знаходження стандартних оцінок кожної ознаки, як зазначено у формулі 2.1.

$$z_i = \frac{(x_i - \mu_i)}{\sigma_i} \tag{2.1}$$

де x_i - і-та вхідна ознака, μ_i - середнє і-ої ознаки та σ_i - її стандартне відхилення.

Загалом, нормалізація гарантує, що всі ознаки мають однаковий масштаб і, таким чином, однаково впливають на прогнози, зроблені алгоритмом машинного навчання. Нормалізація має позитивний ефект на наступні алгоритми:

● K-найближчих сусідів з евклідовою відстанню. Нормалізація гарантує, що всі ознаки однаково впливають на результат класифікації.

● Алгоритми, що використовують градієнтний спуск для оптимізації витрат. Без нормалізації деякі ваги можуть оновлюватися швидше, ніж інші.

● Алгоритми зменшення розмірності (наприклад, аналіз основних компонентів). Без нормалізації на алгоритм впливатимуть ознаки, що мають більшу дисперсію, спричинену різницею в масштабі.

Через це нормалізація сигналу проводиться після оцифрування та фільтрації та перед застосуванням будь-яких подальших трансформацій.

2.1.3. Сегментація

Сегментація сигналу ЕКГ здійснюється на основі алгоритму Хамілтона для детекції QRS-комплексу описаного у роботі [49]. Даний алгоритм базується на алгоритмі детекції QRS-комплексів представленому у роботі [50]. Перевагами алгоритму Хамілтона для виявлення QRS-комплексів є те, що він ефективний і легко модифікується для різних частот дискретизації.

На рис. 2.10 наведено структурну схему алгоритму детекції QRS- комплексів. Детектор QRS-комплексів фільтрує сигнал ЕКГ, щоб створити локальну оцінку потужності в смузі пропускання QRS-комплексів. Фільтри цього детектора описані в [51]. Оскільки всі фільтри базуються на ковзаючих усереднюючих вікнах, для їх адаптації до різних частот дискретизації потрібна лише зміна числа вибірок в ковзаючих вікнах.

Алгоритм Хамілтона відрізняється від описаного у [50] використанням випрямленого, а не квадратичного сигналу та вікна усереднення 80 мс, а не вікна усереднення 150 мс. Також у роботі [50] продемонстровано, що використання випрямлення замість піднесення до квадрату покращує ефективність детектора серцебиттів.

Рис. 2.10. Детектор QRS-комплексів

Нижче наведено основні правила детектора серцебиттів:

1. Ігнорувати всі піки, що передують чи слідують за більшими піками, менше ніж за 200 мс.

2. Якщо виявлено пік, необхідно перевірити, чи містить сигнал ЕКГ як позитивні, так і негативні нахили, якщо ні, пік являє собою базовий зсув.

3. Якщо пік виник протягом 360 мс від попереднього виявлення і мав максимальний нахил менше половини максимального нахилу попереднього виявлення, припускаємо, що це Т-хвиля.

4. Якщо пік перевищує поріг виявлення - це QRS-комплекс.

5. Якщо з останньої детекції минув інтервал, рівний 1,5 кратному середньому інтервалу між R піками, та в межах цього інтервалу спостерігався пік, який перевищував половину порога виявлення, і пік слідував за попереднім виявленням принаймні на 360 мс, класифікуємо цей пік як QRS-комплекс.

Поріг виявлення, використаний у 4 та 5 правилах, обчислюється з використанням оцінок піку QRS-комплексу та висоти піку шуму. Кожен раз, коли пік класифікується як QRS-комплекс, він додається до буфера, що містить вісім останніх піків QRS. Кожного разу, коли виникає пік, який не класифікується як комплекс QRS, він додається до буфера, що містить вісім останніх піків, що не є QRS-комплексом (піки шуму). Поріг виявлення встановлюється між середнім значенням або медіаною пікового шуму та пікових буферів QRS-комплексів за формулою:

$$\text{Detection\_Threshold} = \text{Average\_Noise\_Peak} + \text{TH} * (\text{Average\_QRS\_Peak} - \text{Average\_Noise\_Peak})$$

де TH - пороговий коефіцієнт (зазвичай між 0,3125 і 0,475).

Аналогічно, оцінка інтервалу між R піками, використана в 5-ому правилі, обчислюється як медіана або середнє значення останніх восьми інтервалів між R піками.

Для сегментації ЕКГ-сигналу також можна використовувати алгоритми Христова [52] чи Зеленберга [53]. Для побудови системи біометричної автентифікації було обрано алгоритм Хамілтона через його високу точність та помірну затратність обчислювальних ресурсів.

На рис. 2.11 наведено сигнал ЕКГ із задетектованими R піками. Далі сигнал розбивається на сегменти довжиною 138 вибірок таким чином щоб R пік кожного із сегментів знаходився на 40-ій вибірці (рис. 2.12).

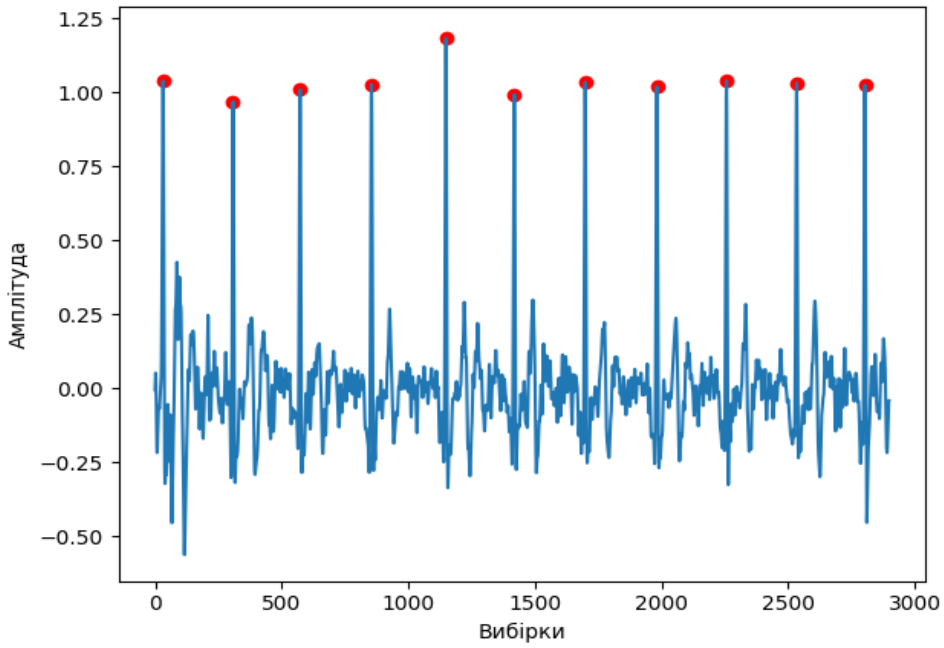


Рис. 2.11. Сигнал ЕКГ із задетектованими R-піками (червоні точки)

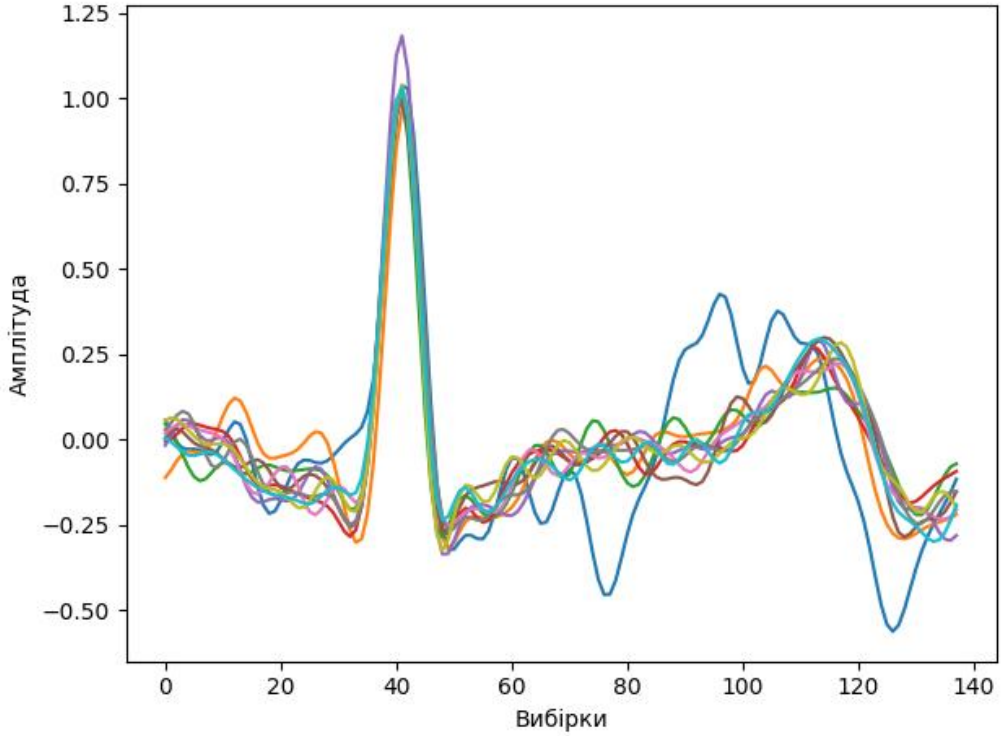


Рис. 2.12. Сегментований ЕКГ-сигнал

2.1.4. Зменшення розмірності даних

Підхід зменшення розмірності даних часто використовується для зменшення розмірності великих наборів даних, перетворюючи великий набір ознак у менший, який все ще містить більшу частину інформації з великого набору.

Зменшення кількості ознак для набору даних, природно, відбувається за рахунок точності, але хитрість у зменшенні розмірності полягає в нехтуванні невеликою точністю для простоти. Оскільки менші набори даних легше досліджувати, візуалізувати та опрацьовувати, а аналіз даних набагато простіший та швидший для алгоритмів машинного навчання без обробки сторонніх ознак.

Отже, підводячи підсумок, ідея зменшення розмірності даних проста - зменшити кількість ознак набору даних, зберігаючи якомога більше інформації. Даний елемент структури біометричної системи автентифікації повинен зменшити кількість вибірок у кожному з сегментів.

Для зменшення розмірності даних можна використати один з наступних методів:

- Аналіз основних компонент (англ. Principal component analysis (PCA))
- Аналіз незалежних компонент (англ. Independent Component Analysis (ICA))
- Лінійний дискримінантний аналіз (англ. Linear Discriminant Analysis (LDA))
- Сингулярний розклад матриці (англ. Singular Value Decomposition (SVD))
- Автоенкодер (англ. Autoencoder)

PCA – це один із найпоширеніших методів зменшення розмірності даних. Мета PCA полягає у отриманні нових змінних, які є лінійними комбінаціями вхідних змінних і є некорельованими. Він знаходить меншу групу базових змінних, що описують дані. PCA проектує n-вимірні дані на нижчий d-вимірний підпростір таким чином, щоб мінімізувати суму квадратів помилок, або максимально збільшити дисперсію, і отримати некорельовані спроектовані розподіли [54]. У більшості випадків основна структура даних буде розрідженою.

Але PCA часто генерує щільні вирази, що ускладнює інтерпретацію. Він обчислюється шляхом виконання власного розкладу на матриці коваріації даних (∑) або шляхом обчислення власних векторів і власних чисел коваріаційної матриці початкових даних. Матрицю коваріації ∑ можна розкласти як,

$$\sum = U \wedge U_T$$

де ∧ - діагональна матриця, що містить власні значення, а U - матриця, що містить відповідні власні вектори. Отримані власні вектори є схожими на основні осі підпростору максимальної дисперсії, власні значення представляють дисперсію вхідних даних уздовж основних осей, а кількість значущих власних значень позначає розрахункову розмірність. Розмір матриці коваріації пропорційний розмірності даних, що робить власне розкладання обчислювально дорогим для даних з дуже великими розмірами. PCA простий в обчисленні і гарантовано забезпечує точне представлення низьких розмірів даних. Але це не дає високої точності для некорельованих даних .

ICA припускає, що приховані змінні є взаємонезалежними та називаються незалежними компонентами спостережуваних даних. Це поверхово пов'язане з аналізом основних компонентів та більш потужною методикою. ICA добре підходить для розділення накладених сигналів. Він застосовує лінійне перетворення для розкладання вихідних даних на компоненти, які є максимально незалежними один від одного. Для зменшення розмірностей, аналіз незалежних компонентів знаходить k компонент, які ефективно відтворюють варіативність вихідних даних. Він розкладає матрицю даних розміром t × d на дві матриці, такі, що:

$$A_{t\times d} = C_{t\times k} * F_{k\times d},$$

де C - матриця коефіцієнтів, а F містить незалежні компоненти.

ICA гарантує точність у випадку некорельованих даних, але отримані незалежні компоненти можуть бути нерелевантними.

LDA – техніка зменшення розмірності, яка найчастіше використовується на етапі попередньої обробки в машинному навчанні. Завдання полягає в проекції

даних на простори менших розмірностей з хорошою роздільною здатністю для класів, щоб уникнути перенавчання, а також зменшити обчислювальні витрати. Це допомагає зменшити розмірність даних і водночас намагається зберегти суттєві особливості кожного з класів.

LDA – це алгоритм навчання з вчителем, який використовує розмічені дані (анотації) під час зменшення розмірності. Алгоритм шукає новий простір ознак, що забезпечує максимальну відокремленість класів, використовуючи підхід, дуже подібний до того, який використовується в PCA.

PCA – це статистична процедура, яка перетворює набір можливо корельованих змінних в набір лінійно некорельованих ознак, званих головними компонентами. По суті, це відкидає найменш важливі ознаки, зберігаючи цінні, знаходячи основні осі компонентів, уздовж яких дисперсія даних висока.

LDA намагається максимізувати поділ між класами, максимізуючи відстань між центроїдами класів, і одночасно мінімізувати дисперсію всередині класу, щоб утворилися добре розділені кластери, що не перекриваються. Мінімізація розбіжностей між класами призводить до створення компактних, менш рознесених класів.

LDA застосовує спектральний розклад до набору даних, а обчислені власні вектори зберігаються у наборі матриць розсіювання (матриці розсіювання між класами та матриці розсіювання всередині класу). Відповідні власні значення позначають довжину або величину власних векторів. Якщо спостерігається, що всі власні значення мають однакову величину, то можна зробити висновок, що дані проектуються на хорошому просторі ознак. Загалом вибираються власні вектори, пов'язані з найбільшими власними значеннями, оскільки вони передають значну інформацію про розподіл даних.

LDA виконується в 5 етапів [55]:

1. Обчислюються d-вимірні середні вектори, m_i для різних класів із набору даних.

2. Обчислюються матриці розсіювання:

(a) матриця розсіювання в межах класу:

$S_w = \sum_{i=1}^{c} S_i$, де S_i - матриця розсіювання для кожного з класів

(b) матриця розсіювання між класами:

$S_B = \sum_{i=1}^{c} N_i(m_i - \overline{m})(m_i - \overline{m})^T$, де $\overline{m}$ - загальне середнє значення, а N_i - розмір кожного класу.

3. Розв'язується узагальнена задача власних значень для матриці $S_w^{-1}S_B$

Власні вектори та власні значення виражають інформацію про спотворення лінійного перетворення. Власні вектори представляють напрямок, а власні значення позначають величину спотворення. Результуючі власні вектори утворюють нові осі нового простору ознак.

4. Вибираються лінійні дискримінанти для нового простору об’єктів. Це робиться шляхом сортування власних векторів за спаданням власних значень та вибору власних векторів з найбільшими власними значеннями, тим самим будується матриця власних векторів $W_{d \times k}$.

5. Трансформація зразків у новий підпростір за допомогою рівняння $Y = X \times W$, де $X_{n \times d}$ - вхідна матриця, і-й рядок відповідає і-му зразку, а $Y_{n \times k}$ - трансформована матриця.

Автоенкодер. Нейронні мережі глибокого навчання можуть бути використані для зменшення розмірності. Автоенкодер – це штучна нейронна мережа без вчителя, яка стискає дані в нижчу розмірність, а потім реконструює вхід назад. Автоенкодер знаходить спосіб представлення даних у нижчому вимірі, більше зосереджуючись на важливих характеристиках, позбавляючись від шуму та надмірності. Він заснований на архітекторі Кодувальник-Декодувальник, де кодувальник кодує дані великого розміру в нижчі розмірності, а декодувальник приймає дані в нижчій розмірності та намагається відновити вхідні дані великого розміру.

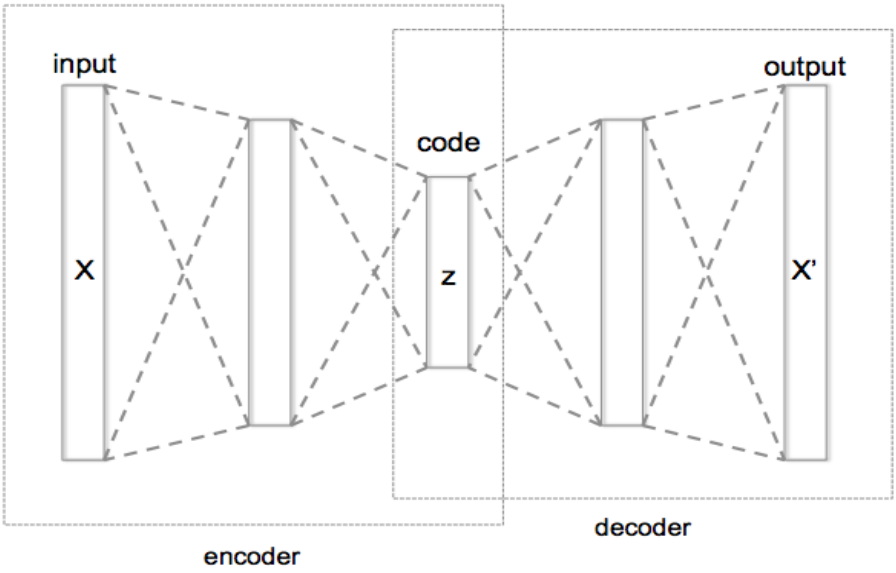


Рис. 2.13. Архітектура автоенкодера

На наведеній вище схемі (рис. 2.13) X - вхідні дані, z - вхідні дані представлені в нижчій розмірності, а X’- відновлені вхідні дані. Відображення вищих та нижчих розмірностей може бути лінійним або нелінійним залежно від вибору функції активації.

SVD. Цей підхід дозволяє отримати точне представлення будь-якої матриці, а також полегшує визначення та усунення менш важливих частин цього представлення для отримання приблизного представлення з будь-якою бажаною кількістю вимірів. Звичайно, чим менше вимірів ми виберемо, тим менш точним буде представлення.

Розглянемо матицю M розмірності m × n, а ранг матриці M позначимо як r. Встановлено, що ранг матриці - це найбільша кількість рядків (або еквівалентно стовпців), для яких ми можемо вибрати лінійну комбінацію рядків, що дорівнює нульовому вектору 0 (ми говоримо, що набір таких рядків або стовпців є незалежним). Тоді ми можемо легко знайти матриці U, Σ і V, як показано на рис. 2.14, враховуючи наступні властивості, перелічені нижче:

1. U розглядається як m × r ортогональна матриця стовпців; тобто кожен із його стовпців є одиничним вектором, а скалярний добуток будь-яких двох стовпців дорівнює 0.

2. V розглядається як n × r ортонормальна матриця стовпців. Відзначається, що V використовується у транспонованій формі, тому саме рядки V^T є ортонормальними.

Σ є діагональною матрицею, оскільки всі елементи не знаходяться на головній діагоналі. Елементи Σ називаються сингулярними значеннями матриці M. [56]

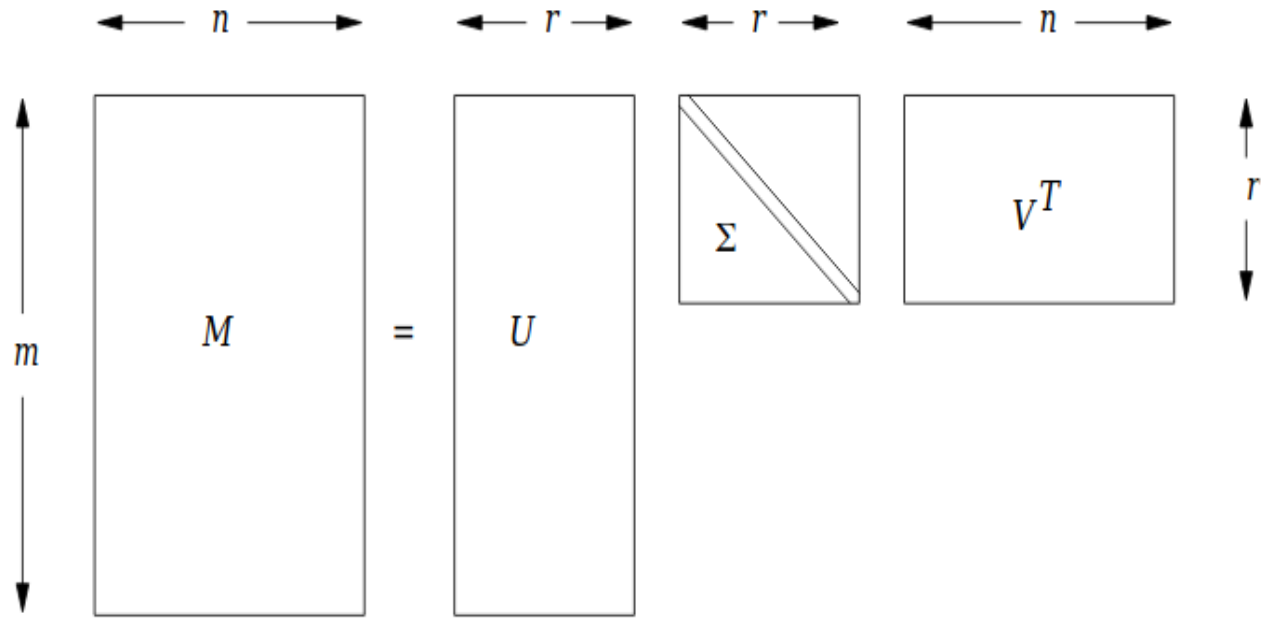


Рис. 2.14. Форма SVD

Таблиця 2.1

Порівняння різних методів зменшення розмірності [57]

	PCA	**ICA**	**LDA**	**SVD**	**Авто-енкодер**
Попередня обробка даних	не потрібна	не потрібна	не потрібна	потрібна	потрібна
Відмовостійкість	слабо чутливий до помилок	чутливий до помилок	слабо чутливий до помилок	слабо чутливий до помилок	чутливий до помилок
Важливі параметри	Ортогональне лінійне перетворення	Статистичні трансформації S=Wx	-	Сингулярні значення	Ваги та зв'язки
Можливість обробки великого набору даних	хороша	хороша	хороша	погана	середня
Обробка багатовимірних даних	хороша	хороша	хороша	погана	погана
Перенавчання	проблема для великого набору даних	проблема для великих розмірних даних із недостатнім обсягом вибірки	трапляєтьс я, коли вихідний набір даних невеликий	тільки для окремих значень	трапляється рідко
Навчання	потрібне	не потрібне	не потрібне	не потрібне	потрібне
Час навчання	великий	середній	менший ніж в PCA	середній	великий

На основі порівняльного аналізу можна зробити висновок, що для нашого набору даних підійде метод PCA як один із найпоширеніших та ефективних алгоритмів зменшення розмірності даних.

2.1.5. Класифікація та автентифікація

Автентифікація суб'єкта здійснюється за допомогою одного з алгоритмів класифікації. Загалом є кілька варіантів побудови системи автентифікації на основі класифікатора:

● класифікація один проти одного, коли для кожного суб'єкта тренується окремий класифікатор. Завдання класифікатора - відповісти чи належить вхідний вектор ознак відповідному суб'єкту. Плюсами такого підходу є можливість швидко та легко масштабувати систему біометричної автентифікації, коли додаються нові суб'єкти (потрібно просто натренувати додаткові моделі класифікатора для нових суб'єктів). Мінусом є складність побудови навчальних наборів для моделей кожного з класифікаторів, оскільки такі набори повинні містити набори з ‘хибними’ ознаками, які не належать суб'єкту відповідної моделі. Складністю є формування таких наборів з достатньою варіативністю ознак.

● класифікація один проти багатьох, коли тренується загальна модель для всіх об'єктів і рішення автентифікації будується на тому чи спрогнозував класифікатор клас об'єкта. Плюсом такого підходу є простота, оскільки для всіх класів тренується одна модель. Також немає проблем з формуванням варіативного тренувального набору (його формують всі ЕКГ-записи). Мінусами є те, що при додаванні нового класу або при перетренуванні існуючого потрібно перетреновувати цілу модель. В даному дисертаційному дослідженні буде використано саме цей підхід, через простоту імплементації, представлення параметрів та результатів експериментів у порівнянні з іншими підходами Проте не виключається можливість використання будь-якого з представлених алгоритмів для імплементації реальної системи автентифікації на основі ЕКГ.

● класифікація на основі векторних ЕКГ-вкладень (ECG embeddings). Відповідно до даного підходу на великому наборі даних (тисячі - сотні тисяч класів) тренується модель, яка здійснює трансформацію вхідного сигналу ЕКГ у векторні ЕКГ вкладення. Дистанція між векторами того ж класу повинна бути

мінімальною, в той час як дистанція між векторами різних класів повинна бути якомога більшою. Ідея даного підходу запозичена з області розпізнавання облич, де багато алгоритмів будується на його основі. Плюсами такої системи є те, що не потрібно тренувати нові моделі при додаванні нових користувачів до системи автентифікації, адже ми використовуємо існуючу модель, яка поверне вектори ЕКГ-вкладень для нових користувачів, які будуть додані в базу даних. В подальшому використанні автентифікація здійснюється на основі обчислення дистанції між виміряними ЕКГ векторами та векторами ЕКГ з бази. Мінусом такого підходу є складність імплементації через відсутність відповідного набору даних. Даний підхід неодмінно буде реалізовано в моїх подальших наукових роботах.

В даному дисертаційному дослідженні буде використано наступні архітектури для побудови класифікатора системи біометричної автентифікації на основі ЕКГ-сигналу:

- Метод опорних векторів (англ. Support Vector Machine (SVM));
- Лінійний дискримінантний аналіз (англ. Linear Discriminant Analysis (LDA));
- K найближчих сусідів (англ. K-nearest neighbors (KNN));
- Дерева рішень (англ. Decision trees);
- Нейронні мережі (англ. Neural networks).

SVM - це потужний та гнучкий алгоритми машинного навчання з учителем, який використовуються як для задач класифікації, так і для регресії. Але загалом вони використовуються в задачах класифікації. У 1960-х роках SVM був вперше представлений, а пізніше в 1990 році був удосконалений. SVM має свій унікальний спосіб імплементації порівняно з іншими алгоритмами машинного навчання. Останнім часом SVM надзвичайно популярний завдяки своїй здатності обробляти безліч безперервних і категоріальних змінних.

Модель методу опорних векторів - це, в основному, представлення різних класів на гіперплощині в багатовимірному просторі. Гіперплощина генерується

ітераційним способом SVM, таким чином щоб можна було мінімізувати помилку класифікації чи регресії. Мета SVM - розділити набори даних на класи, щоб знайти максимальну розділову гіперплощину (рис. 2.15).

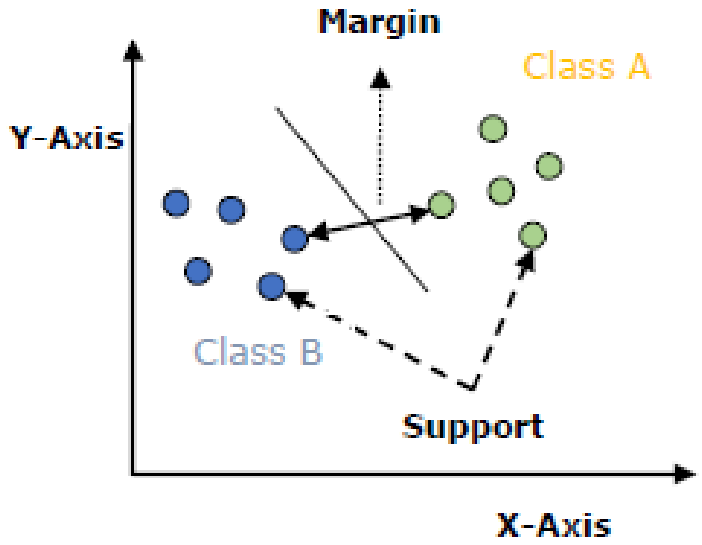


Рис. 2.15. Пояснення принципу роботи SVM на двовимірному просторі

Метод опорних векторів використовує:

Опорні вектори - точки даних, які є найближчими до гіперплощини, називаються опорними векторами. Розділювальна лінія визначається за допомогою цих точок даних.

Гіперплощина - Як ми бачимо на наведеній вище схемі, це площина прийняття рішень або простір, який розділений між наборами об’єктів, що мають різні класи.

Розділення - може бути визначено як інтервал між двома опорними векторами з різних класів. Велике значення розділення характеризує хорошу роздільну здатність і відповідно якість класифікації чи регресії, а мале - як погану.

Переваги SVM:

- ефективність для даних представлених в багатовимірному просторі;
- залишаються ефективними коли кількість вхідних ознак є більшою за кількість вибірок;

● можуть моделювати нелінійні розділові гіперплощини, а також на вибір є багато ядер;

● стійкість до перенавчання, особливо у багатовимірному просторі.

Недоліки SVM:

● SVM потребує пам'яті, а також його складніше налаштовувати через важливість вибору правильного ядра;

● класифікатори на основі SVM погано працюють із накладанням класів.

LDA використовується як інструмент в задачах класифікації, зменшення розмірності та візуалізації даних. Незважаючи на свою простоту, LDA часто дає надійні, стійкі та зрозумілі результати класифікації. Вирішуючи реальні проблеми класифікації, LDA часто є першим та базовим методом порівняння з іншими більш складними підходами.

Лінійний дискримінантний аналіз – це метод пошуку лінійної комбінації змінних, що найкращим чином розділяє деяку множину об’єктів на два або більше класів. Даний підхід детально описаний у попередньому пункті 2.1.4.

Переваги підходу LDA:

● використання для побудови прототипів, оскільки використовується дистанція до середини класу, то його прогнози просто інтерпретувати;

● оскільки це лінійна модель - її просто реалізувати, а класифікація надійна;

● зменшення розмірності є частиною класифікатора, також це корисно для візуалізації даних;

● обчислювальна ефективність, порівняно із більш складними алгоритмами.

Недоліки підходу LDA:

● оскільки це лінійна модель то вона може не адекватно розділяти класи;

● вимагає нормального розподілу вхідних даних та погано працює з категорійними даними.

KNN - це алгоритм машинного навчання з учителем, який часто використовується в задачах класифікації. Він працює на простому припущенні, що «Яблуко від яблуні недалеко падає», що означає, що подібні речі завжди

знаходяться в безпосередній близькості. Цей алгоритм працює шляхом класифікації точок даних на основі класифікації сусідів, класи яких уже відомі. Будь-яка нова вибірка класифікується на основі показника подібності всіх наявних вибірок. Технічно, алгоритм класифікує невідомий елемент, переглядаючи k його вже класифікованих, найближчих сусідів, виявляючи більшістю голосів до якого класу його віднести.

KNN – це алгоритм ледачого навчання, оскільки він не має етапу навчання як такого, а просто запам’ятовує набір навчальних даних. Усі обчислення затримуються до здійснення класифікації.

Також KNN використовує необроблені навчальні екземпляри з проблемної області для прогнозування і часто згадується як алгоритм навчання на основі конкретного випадку. Навчання на основі конкретних випадків означає, що KNN не вивчає явно модель. Швидше він запам'ятовує навчальні випадки, які потім використовуються як "знання" для фази прогнозування. Враховуючи вхідні дані, коли ми просимо алгоритм зробити прогнозування, він використовуватиме запам’ятовані навчальні екземпляри, щоб дати відповідь (рис. 2.16).

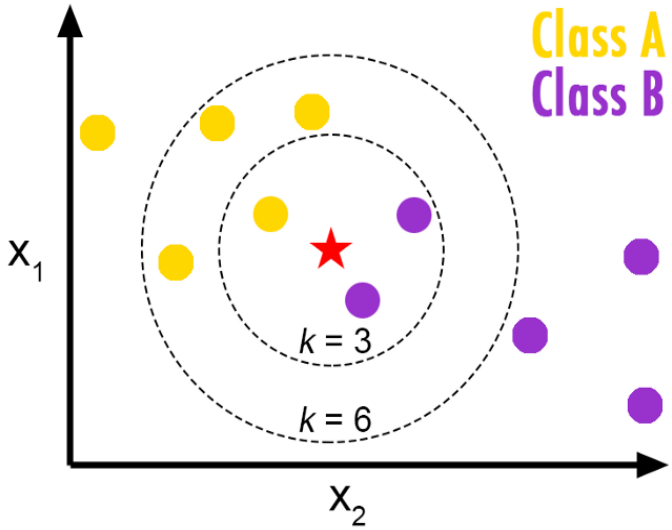


Рис. 2.16. Пояснення принципу роботи KNN. При k=3 метод класифікує невідомий елемент як клас B, при k=6 - як клас A

KNN це непараметричний метод, незалежно від обсягу даних, єдиним невідомим параметром є K, який задає кількість сусідів на основі яких робиться класифікація.

Переваги підходу KNN:

- простий у реалізації;
- чудово працює на наборах даних з багатьма класами;
- KNN є непараметричним алгоритмом і не вимагає жодних припущень щодо розподілу даних. Це надає KNN додаткових переваг у певних налаштуваннях, де дані вкрай незвичні. Це є причиною того, що KNN є першим вибором, коли немає попередніх знань або дуже мало знань про розподіл даних;
- обчислювальна ефективність, порівняно із більш складними алгоритмами.

Недоліки підходу KNN:

- високі обчислювальні затрати – оскільки алгоритм зберігає всі навчальні дані;
- ефективність алгоритму дуже швидко падає із зростанням набору даних;
- він страждає від перекосу розподілу класів, якщо певний клас часто трапляється у навчальному наборі, то, швидше за все, він буде домінувати в більшості голосів при класифікації.

Дерева рішень – це популярний алгоритм машинного навчання з учителем для задач класифікації та регресії. Дерево рішень можна використовувати для візуального та зрозумілого представлення прийнятих рішень.

Дерева рішень надзвичайно популярні з різних причин, а їх інтерпретабельність, мабуть, є їх найважливішою перевагою. Їх можна дуже швидко навчити і їх легко зрозуміти, відслідкувавши шлях використаний для прогнозування.

Дерева рішень зазвичай складаються з наступних елементів (рис. 2.17):

- Кореневий вузол. Цей вузол верхнього рівня представляє кінцеву мету або важливе рішення, яке ви намагаєтеся прийняти.

● Гілки. Гілки, що виходять з кореня, представляють різні варіанти – або варіанти дій, доступні при прийнятті певного рішення.

● Листковий вузол. Листкові вузли, які прикріплені в кінці гілок, представляють можливі результати для кожної дії. Зазвичай існує два типи листкових вузлів: квадратні листові вузли, що вказують на інше рішення, яке потрібно прийняти, та круглі листкові вузли, які вказують на прийняте рішення.

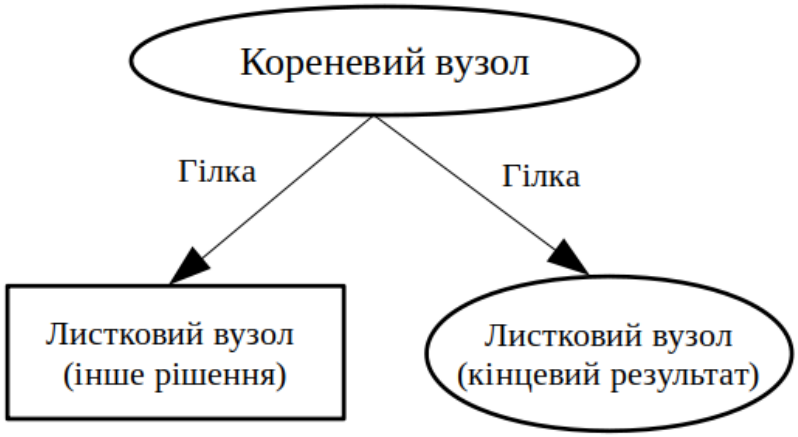


Рис. 2.17. Структура дерева рішень

Дерева рішень використовують підхід зверху вниз, вони намагаються згрупувати схожі спостереження та шукати найкращі правила, що розділяють неподібні спостереження. Вони використовують процес розбиття по рівнях, коли на кожному рівні намагаються розділити дані на дві або більше груп, щоб дані, що потрапляють в одну групу, були найбільш схожими між собою (однорідність), а групи максимально відрізнялися одна від одної (неоднорідність).

Переваги підходу на основі дерев рішень:

● простий для розуміння, інтерпретації та візуалізації;

● може обробляти як числові, так і категоріальні дані. Добре працює на великих наборах даних;

● працює швидко.

Недоліки підходу на основі дерев рішень:

● ймовірність перенавчання;

● моделі дерев рішень часто схильні до розбиття ознак на велику кількість рівнів;

● невеликі зміни в навчальних даних можуть призвести до значних змін в логіці прийняття рішень.

Нейронні мережі – це набір алгоритмів, функціонування яких відтворює роботу людського мозку. Коли ви відкриваєте очі, те що ви бачите називається даними і обробляється нейронами вашого мозку, таким чином ви розпізнаєте об'єкти навколо вас. Нейронні мережі беруть великий набір даних, обробляють їх (витягують з даних приховані закономірності) та повертають результат.

Нейронна мережа складається з трьох типів важливих шарів (рис. 2.18):

● Шар входу – шар, на який надходять ознаки, які подаються на вхід нейронної мережі;

● Прихований шар - між шаром входу та шаром виходу повинен бути один або кілька прихованих шарів, в яких проводяться основні обчислювальні операції, знаходяться закономірності в вхідних даних;

● Шар виходу - вхідні дані проходять серію перетворень через прихований шар, результат яких доставляється через цей шар.

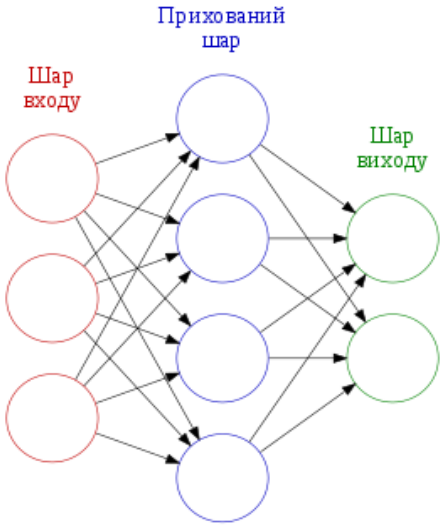


Рис. 2.18. Структура штучної нейронної мережі

Функціонування нейронних мереж розглянемо на принципі роботи перцептрона. Перцептрон – це одношарова нейронна мережа, яка використовується для класифікації лінійних даних. Він має 4 важливі компоненти (рис. 2.19):

- Входи;
- Ваги та зміщення;
- Підсумовуюча функція;
- Активаційна функція.

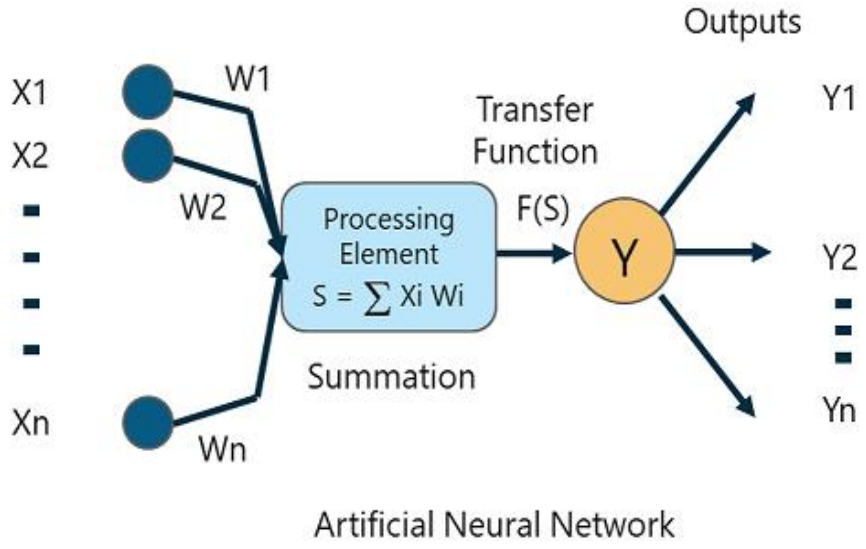


Рис. 2.19. Будова перцептрону

Перцептрон працює наступним чином – вхідні дані (x), отримані від вхідного рівня, множаться на присвоєні їм ваги w. Далі застосовується зміщення та відбувається підсумовування. Після чого зважена сума входів передається на вхід відповідної функції активації. Функція активації відображає вхід на відповідний вихід (y).

Ваги та зміщення. Вага кожної ознаки вхідних даних вказує, наскільки ця вхідна інформація важлива для прогнозування результату. Параметр зміщення, навпаки, дозволяє регулювати криву функції активації таким чином, щоб досягти точного виходу.

Підсумовуюча функція. Після того, як входам присвоєно деяку вагу, береться добуток відповідного входу та ваги. Додавання всіх цих добутків дає нам зважену суму. Це робиться за допомогою підсумовуючої функції.

Активаційна функція. Метою активаційної функції є введення нелінійності у вихід нейрона. Це важливо, оскільки більшість даних у реальному світі нелінійні і ми хочемо, щоб нейрони вивчали ці нелінійні закономірності. Кожна функція активації приймає одне число і виконує над ним певну фіксовану математичну операцію. Існує кілька функцій активації, з якими ви можете зіткнутися на практиці: нормована експоненційна функція, сигмоїд, випрамляч, тощо.

Переваги підходу на основі нейронних мереж:

- нейронні мережі здатні вивчати нелінійні та складні взаємозв'язки;
- нейронні мережі підходять для неструктурованих наборів даних, таких як зображення, аудіо та текст, і вони погано працюють на структурованих наборах даних.

Недоліки підходу на основі нейронних мереж:

- Великий час навчання;
- Потрібні високі обчислювальні потужності;
- Природа чорної коробки - ми не можемо сказати як і чому модель зробила певне прогнозування.

Кожну з наведених вище архітектур буде використано у експериментальній частині дисертаційного дослідження, як для перевірки ефективності розроблених методів для покращення експлуатаційних характеристик біометричної системи, так і для вибору оптимальної архітектури для побудови компонента класифікації біометричної системи автентифікації.

2.2. Вибір перспективних підходів щодо покращення технічних і експлуатаційних характеристик біометричної системи ЕКГ- автентифікації

Сигнал ЕКГ утворюється електричними імпульсами, що надходять від мозку до серця. Кожен імпульс стимулює різні частини серцевих м'язів (міокарда), щоб виконати повний цикл серцебиття. Цей процес повторюється циклічно, а тому сигнал ЕКГ має квазіперіодичний характер. Крім того, можна спостерігати розподіл електричного поля по всьому тілу, а вимірювання різниці потенціалів із різних частин на поверхні тіла і є сутністю формування ЕКГ-сигналу [58].

Рівень сигналу ЕКГ є низьким, а тому сприйнятливий до різного роду спотворень. Щоб забезпечити якість зареєстрованого ЕКГ-сигналу в клінічній практиці використовуються такі методи і засоби, як гель для підвищення провідності контакту «тіло-електрод», багатоканальний запис біопотенціалів із різних точок грудної клітки і кінцівок (зазвичай 12 відведень), комфортні умови обстеження людини (зручне лежання на кушетці), зупинка дихання на момент запису ЕКГ. Вочевидь, цей спосіб отримання ЕКГ не може бути використаний у задачах біометрики, зокрема, у системах контролю доступу. У таких застосуваннях, зазвичай, потрібно забезпечити якомога простішу реєстрацію ЕКГ-сигналу, наприклад, сухими електродами із пальців лівої і правої руки (одноканальний варіант відомий як перше відведення). Звісно, якість сигналу при цьому сильно постраждає.

Але не лише число каналів і якість сигналу ЕКГ різняться у медичній діагностиці та біометрії – зовсім різними є цілі дослідження: для медицини – це встановити має місце чи ні відхилення електрокардіограми від норми, а для біометрики розрізнити ЕКГ, що належать різним суб'єктам.

Вважається, що на форму ЕКГ-сигналу мають вплив будова міокарду, його розташування у грудній клітці та інші фізіологічні особливості. Численні дослідження показали, що електрокардіограма є надійним біометричним маркером [59-63], але зовсім не зручним для сприйняття людиною на зір або слух, на відміну від розпізнавання облич чи голосу. Малоефективними виявилися і

класичні алгоритми цифрового оброблення сигналів, передовсім, через значну варіативність ЕКГ-сигналу. Лише системи, засновані на машинному навчанні можуть впоратися із задачею розпізнавання людей за електрокардіограмою, причому істотним для біометрики є те, що форма хвилі серцебиття відрізняється від суб'єкта до суб'єкта [59].

Проте на шляху практичного застосування технологій машинного навчання у реальних біометричних системах на базі ЕКГ виникають істотні проблеми, однією з яких є залежність від частоти серцевих скорочень. Така варіативність серцевого ритму може спостерігатися не лише між різними суб'єктами, але і для однієї і тієї ж людини у різні моменти часу (рис. 2.20).

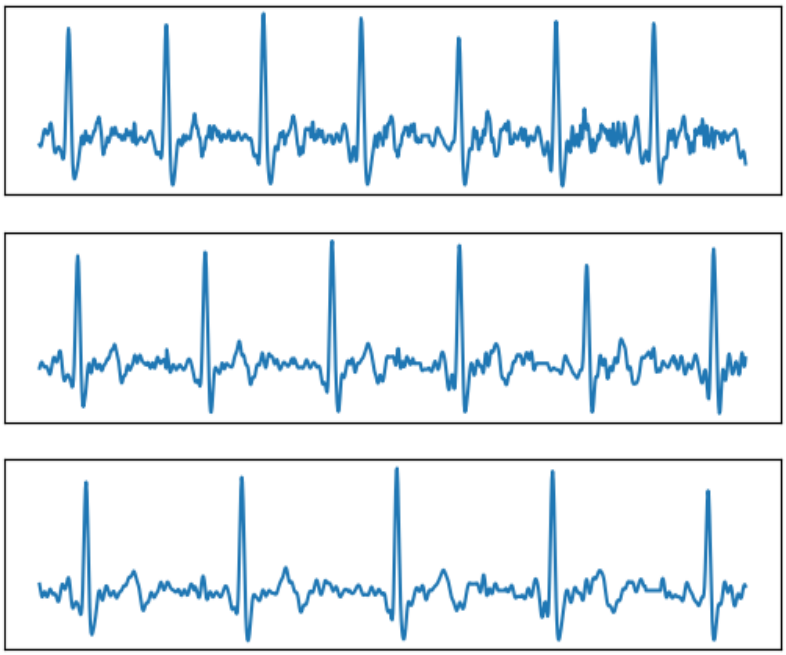


Рис. 2.20. Фрагменти сигналу ЕКГ із різною тривалістю серцевого ритму

У повсякденному використанні система біометричної ідентифікації має бути інваріантною до зміни серцевого ритму зумовленої емоційними, фізичними чи іншими чинниками. Проблема наборів даних, які використовуються для навчання класифікаторів полягає у тому, що вимірювання ЕКГ-сигналу зазвичай проводилось в один день впродовж короткого проміжку часу. Як наслідок, значення серцевого ритму в одержаних записах є доволі сталими. Дослідження

показали, що класифікатори з поміж іншого формували характерні ознаки, які базуються на тривалості серцевого ритму. Це, звичайно, є хибним, оскільки, у застосуваннях автентифікації збільшуватиме число відмов у доступі справжнім користувачам (зростають помилки 2-го роду).

Тому для сигналу ЕКГ, як функції у 2D-просторі, окрім нормалізації за амплітудою, потрібно застосувати нормалізацію в часі. Суть запропонованого підходу полягає у тому, щоб масштабувати кожен сегмент електрокардіограми до заданого стандартного вікна, максимально зберігаючи форму ЕКГ-сигналу, зумовлену істотними фізіологічними особливостями кожного суб'єкта. Це забезпечить інваріантність системи розпізнавання до стиснення/розтягнення ЕКГ-сигналу як по вертикалі, так і по горизонталі.

Також одним із найважливіших завдань розробки біометричної системи на основі ЕКГ є отримання сигналу належної якості. Це може бути проблематично, оскільки інформативна форма хвилі змішується з різними перешкодами (шум, дихання, м’язова активність тощо). У багатьох випадках форма окремих ударів серця може суттєво відрізнятися від сусідніх сегментів, а відтак трактуватися як аномалія чи артефакт. Вдалим рішенням для розуміння специфіки промахів на ЕКГ є візуальна оцінка за допомогою накладання сегментів ЕКГ-сигналу. Для цього оригінальний сигнал розбивається на відтинки, що відповідають окремим скороченням серця з подальшим вирівнюванням по R зубцю. Приклади промахів в ЕКГ-сигналі наведено на рис. 2.21.

Класична стратегія для визначення артефактів ґрунтується на оцінці евклідової відстані між сегментами ЕКГ-сигналу. Якщо деякі сегменти не вписуються в певний діапазон, то їх ідентифікують, як промахи та відкидають. Основним недоліком такого підходу є те, що він заснований на інтегральній оцінці, яка не завжди враховує локальні особливості артефактів. У багатьох випадках невеликі артефакти зосереджені в певній частині сегменту (наприклад P, або T хвилі). Відтак їх загальний вплив на евклідову відстань може бути

відносно не значний, але при цьому цей зразок матиме спантеличувальний ефект, і як наслідок може призвести до некоректних результатів автентифікації.

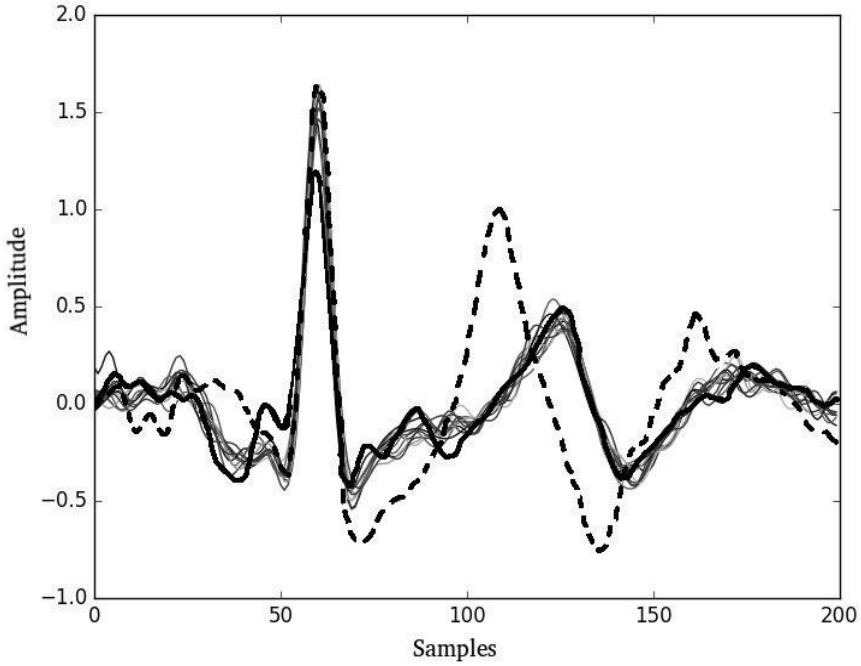


Рис. 2.21. Артефакти у ЕКГ-сигналах: невеликий (жирна крива) та значний (пунктирна крива)

Отже, для покращення технічних та експлуатаційних характеристик біометричної системи ЕКГ-автентифікації пропонується включити в її структурну схему наступні компоненти:

- компонент темпоральної нормалізації ЕКГ-сигналу, який повинен зробити біометричну систему інваріантною до зміни серцевого ритму, тим самим підвищивши її точність.
- компонент виявлення та виправлення артефактів у ЕКГ-сигналі, який за допомогою інструментарію на основі статистики та машинного навчання повинен підвищити точність та якість біометричної системи, а також підвищити її відмовостійкість.

2.3. Розроблення структури біометричної системи ЕКГ-автентифікації із покращеними характеристиками

В попередньому підрозділі було розглянуто підходи за допомогою яких можна домогтися покращення технічних та експлуатаційних характеристик біометричної системи ЕКГ-автентифікації. Для розв'язання поставлених завдань було вирішено включити в структурну схему (рис. 2.1) біометричної автентифікації на основі ЕКГ компонент темпоральної нормалізації та компонент виявлення та виправлення артефактів (рис. 2.22).

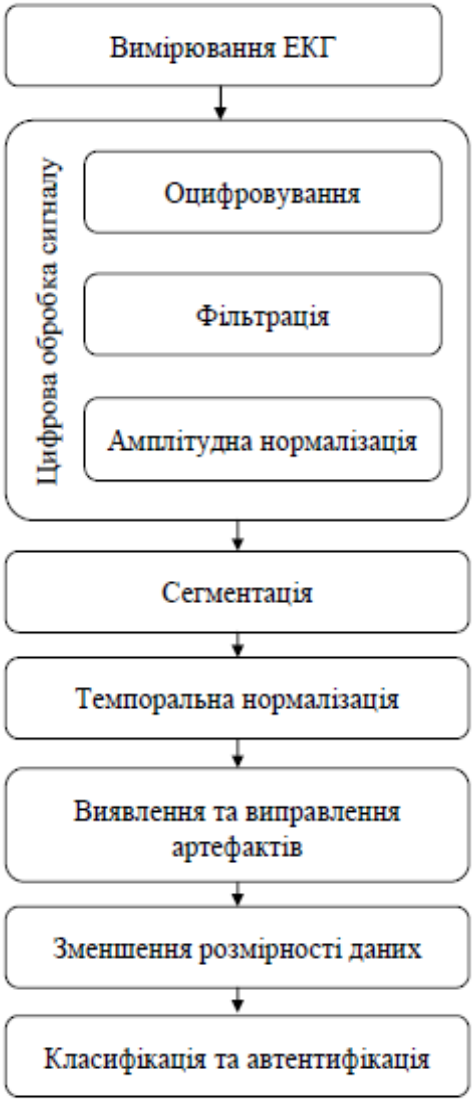


Рис. 2.22. Структурна схема біометричної системи автентифікації із покращеними характеристиками.

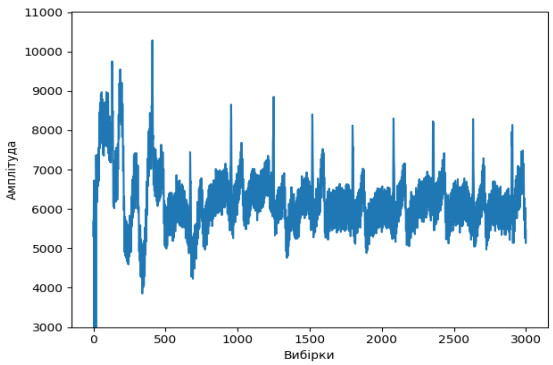
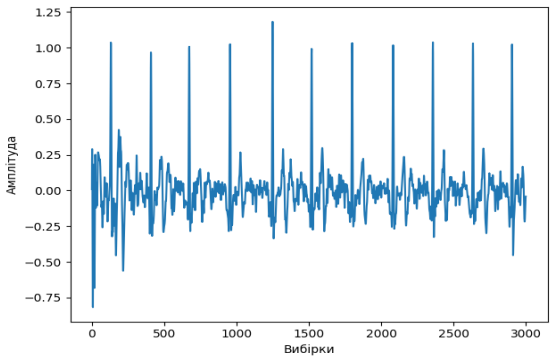
Алгоритм темпоральної нормалізації працює на сегментному рівні, тому його доцільно розмістити одразу після компоненту сегментації.

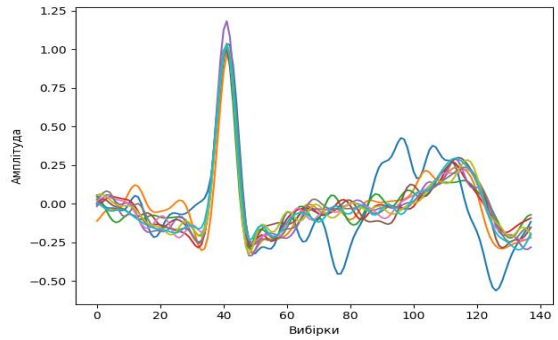
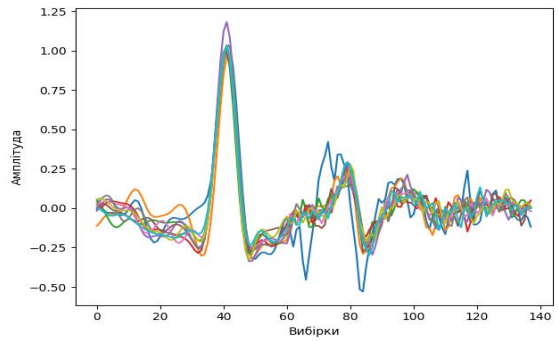
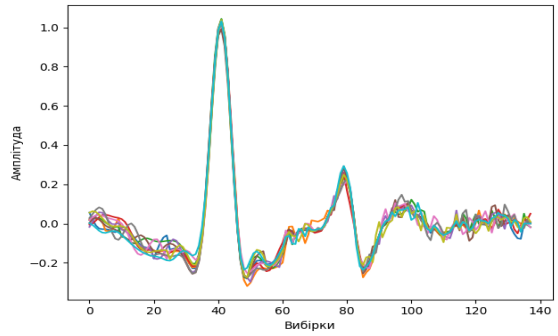
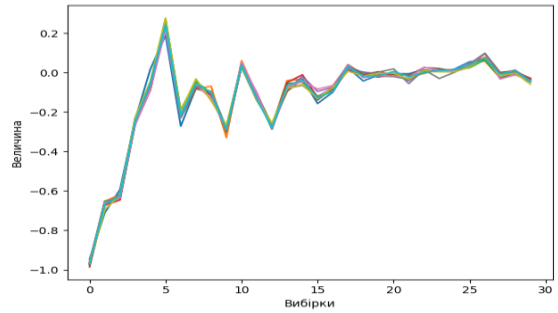
Алгоритм виправлення та виправлення артефактів також працює на сегментному рівні. Його доцільно розмістити після компоненти темпоральної нормалізації, перед компонентою зменшення розмірності даних, оскільки алгоритм зменшення розмірності даних здійснює проекцію ознак в іншу гіперплощину, де алгоритм виправлення та виправлення артефактів може бути неефективним.

Щоб краще зрозуміти як функціонує біометрична система автентифікації на основі ЕКГ нижче наведено таблицю із виглядом та описом сигналів які циркулюють всередині системи автентифікації (таблиця 2.2)

Таблиця 2.2.

Опис сигналів біометрична система автентифікації на основі ЕКГ

Компонент	Вигляд сигналу на виході	Опис сигналу на виході
Вимірювання ЕКГ		аналоговий сигнал, тривалістю ~10 с.
Цифрова обробка сигналу		цифровий сигнал, підданий фільтрації та нормалізації

Сегментація		цифровий сигнал, матриця розміру Nx138, де N - кількість сегментів або серцебиттів
Темпоральна нормалізація		цифровий сигнал, матриця розміру Nx138, підданий темпоральній нормалізації
Виявлення та виправлення артефактів		цифровий сигнал, матриця розміру Nx138, артефакти піддані корекції
Зменшення розмірності даних		цифровий сигнал, матриця розміру Nx30, дані піддані проекції з гіперплощини Nx138 розміром в гіперплощину розміром Nx30 із збереженням корисної інформації
Класифікація		цифровий сигнал, бінарний, відповідь чи належить кардіограма заявленому користувачу

2.4. Методика оцінювання ефективності методів і засобів біометричної автентифікації на основі ЕКГ-сигналу

Вибір правильної методики оцінювання є невід'ємною частиною проектування життєздатної біометричної системи. Метою даного підрозділу є огляд найпоширеніших метрик для оцінювання ефективності проектованої біометричної системи та вибір оптимальних метрик, які будуть використані у наступних розділах.

Accuracy. У найпростішому метрикою може бути частка шаблонів для яких класифікатор прийняв правильне рішення.

$$Accuracy = \frac{P}{N} \tag{2.2}$$

де P – кількість шаблонів за якими класифікатор прийняв правильне рішення; N – розмір навчальної вибірки.

Очевидне рішення, на якому для початку можна зупинитися.

Проте, у цієї метрики є одна особливість, яку необхідно враховувати. Вона привласнює всім шаблонам однакову вагу, що може бути не коректно в разі, якщо розподіл шаблонів у навчальній вибірці сильно зміщений в бік якогось одного або декількох класів. В цьому випадку у класифікатора є більше інформації про ці класи, і, відповідно, в рамках цих класів він буде приймати більш адекватні рішення. На практиці це призводить до того, що ви маєте точність, скажімо, 80%, але при цьому в рамках якогось конкретного класу класифікатор працює погано не визначаючи правильно навіть третину шаблонів.

Єдиним виходом з ситуації є навчати класифікатор на спеціально підготовленому, збалансованому комплексі шаблонів.

Інший вихід полягає в зміні підходу до формальної оцінки якості.

Матриця неточностей. Класифікатор системи автентифікації чутливий до багатьох факторів, таких як вибір алгоритму класифікації, обсяг навчальних даних, якість даних, ефективність екстракції ознак, тощо. Ці фактори впливають на показники ефективності, що розраховуються для кожного класифікатора. У таблиці 2.3 наведені різні типи метрик, які можуть бути використані для

оцінювання будь-якого класифікатора. В таблиці продемонстровано результати двокласної класифікації: прогнози класифікатора знаходяться у стовпцях, а істинні відомі класи у рядках. Діагональ з верхнього лівого кута до правого нижнього показує кількість правильно класифікованих шаблонів. TP (True Positives) та TN (True Negatives) спостерігаються, коли класифікатор прогнозує той самий результат, як і істинне значення класу для шаблону. FN (False Negatives) та FP (False Positives) – це коли класифікатор дає протилежний результат відомій класифікації.

Таблиця 2.3.

Матриця неточностей для двокласної класифікації

		Прогнозований клас	
		Позитивне	Негативне
Істинний клас	Позитивне	TP	FP
	Негативне	FN	TN

False Accept Rate (FAR) – також відома як помилка другого роду. FAR відображає відсоток випадків, коли автентифікацію пройшов хибний користувач. Якщо FN – число випадків хибно автентифікованих користувачів, а N – загальна кількість спроб автентифікації, то FAR може бути обчислена за формулою (2.3).

$$FAR = FN / N \quad (2.3)$$

False Reject Rate (FRR) – також відома як помилка першого роду. FRR відображає відсоток випадків, коли користувача помилково не автентифіковано. Якщо FP – число випадків хибно не автентифікованих користувачів, а N – загальна кількість спроб автентифікації, то FRR може бути обчислена за формулою (2.4).

$$FRR = FP / N \quad (2.4)$$

Equal Error Rate (EER) - коефіцієнт рівної ймовірності помилок першого та другого роду, що відповідає величині співпадіння значень FRR та FAR (рис. 2.23). Якісна та надійна біометрична система повинна мати низький рівень EER. У

деяких системах існує можливість регулювання порогу чутливості, що дозволяє гнучко налаштовувати їх відповідно до вимог безпеки. Не слід забувати, що збільшення чутливості систем (і, як наслідок, зменшення ймовірностей хибного доступу - FAR) одночасно призводить до збільшенням часу автентифікації та підвищення ймовірності хибної відмови - FRR. Необхідно досягти компромісу між значеннями FRR та FAR.

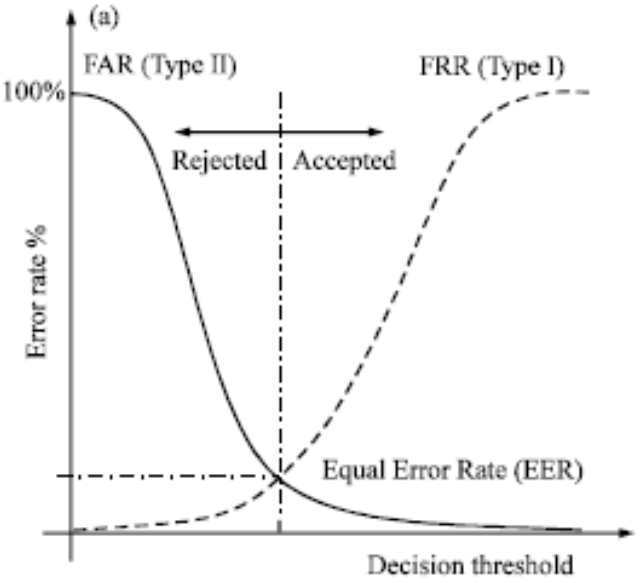


Рис. 2.23. Визначення EER на основі кривих FRR та FAR

EER також можна визначити, побудувавши ROC криву для класифікатора, як детально описано нижче, та визначивши його абсцису, побудувавши діагональну лінію від верхнього лівого до правого нижнього кутів та спостерігаючи, де дві лінії перетинаються.

ROC крива (англ. receiver operating characteristic, робоча характеристика приймача), як видно з рисунку 2.24, показує взаємозв'язок між FAR та True Accept Rate (TAR), тобто відсоток шаблонів, які було вірно автентифіковано. Крива ROC показує загальну ефективність моделей класифікації. Чим ближче лінія підходить до лівого верхнього кута графіку, тим кращою є система автентифікації.

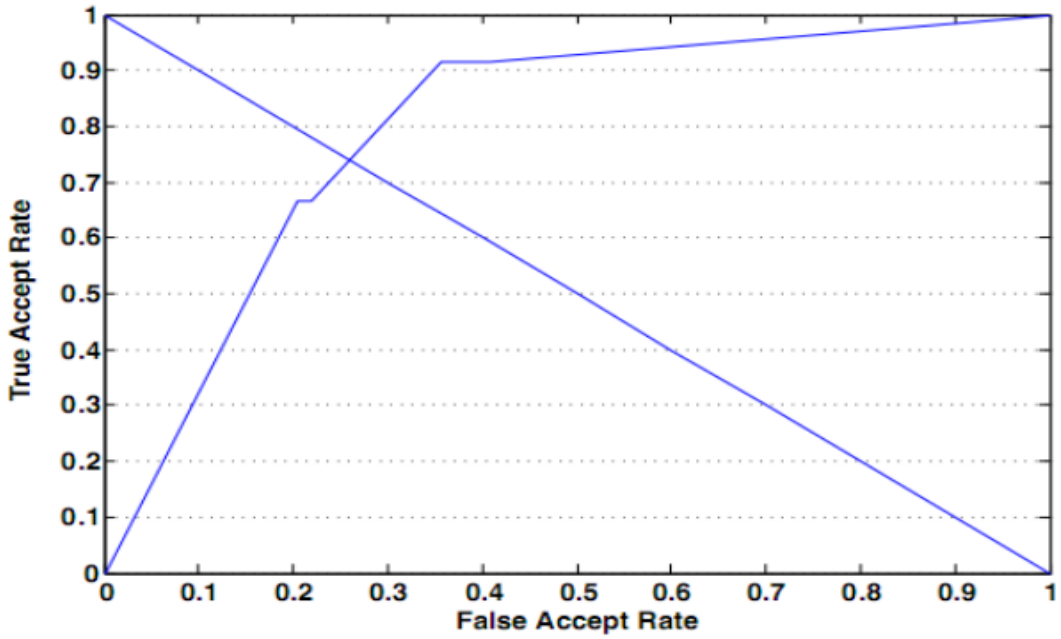


Рис. 2.24. Приклад ROC кривої. AUC для цієї кривої становить ~80%. EER (26%) – це точка перетину двох ліній

Area Under Curve (AUC) - це міра площі під ROC кривою для певного класифікатора та певного користувача. Це представлення ймовірності вірної відповіді при класифікації даних - випадковий класифікатор матиме значення AUC 0.5, а ідеальний класифікатор - AUC 1.0. AUC - це числова репрезентація ROC кривої.

Для оцінки якості роботи алгоритмів автентифікації на кожному з користувачів окремо введемо метрики precision (2.5) і recall (2.6):

$$\text{precision} = \frac{TP}{TP+FP} \tag{2.5}$$

$$\text{recall} = \frac{TP}{TP+ FN} \tag{2.6}$$

Precision можна інтерпретувати як частку шаблонів, названих класифікатором позитивними і які є правильно класифікованими, а recall показує, яку частину від усіх об’єктів позитивного класу було коректно класифіковано.

F score. Precision і recall об’єднують в єдину метрику під назвою F міра (F score) чи її частковий випадок — міру F1, особливо якщо вам потрібен простий спосіб порівняння двох класифікаторів. F1 міра — це середнє гармонійне

precision і recall, що дає можливість об’єктивно вибрати найкращу модель з найкращими значеннями precision та recall водночас. У результаті класифікатор отримує високу міру F1, тільки якщо високими є precision і recall (2.7).

$$F_1 = 2 * \frac{\text{precision} * \text{recall}}{\text{precision} + \text{recall}} \quad (2.7)$$

Дана формула надає однакову вагу precision і recall, тому F-міра буде падати однаково при зменшенні і precision, і recall. Можна розраховувати F-міру надавши різну вагу precision і recall, якщо ви свідомо віддаєте пріоритет однієї з цих метрик при розробці алгоритму [64, 65].

$$F = (\beta^2 + 1) \frac{\text{precision} * \text{recall}}{\beta^2 * \text{precision} + \text{recall}} \quad (2.8)$$

де β приймає значення в діапазоні 0 <β <1 якщо ви хочете віддати пріоритет precision, а при β> 1 пріоритет віддається recall. При β = 1 формула зводиться до попередньої (2.7) і ви отримуєте збалансовану F-міру.

2.5. Опис ЕКГ наборів даних

В даному підрозділі наведено огляд та співставлення існуючих наборів ЕКГ-даних. Разом із розвитком та популяризацією алгоритмів штучного інтелекту, галузь біометричної ідентифікації/автентифікації отримала новий поштовх у розвитку. Сьогодні з'являється дедалі більше наборів даних призначених для побудови біометричних систем на основі ЕКГ.

Один з таких наборів вдалося зібрати автору даної роботи у співпраці з колегами із Національного університету “Львівська політехніка” Хомою Володимиром Васильовичем та Хомою Юрієм Володимировичем. Назва цього набору - Lviv Biometric Dataset. Вимірювання ЕКГ-записів здійснювалося за допомогою апаратури, яку детально описано у 2.1.1. В основному суб'єктами вимірювання були студенти та науково-педагогічний склад Національного університету “Львівська політехніка”.

На момент проведення експериментальної частини дисертації набір даних Lviv Biometric Dataset містив 1809 записів виміряних у 115 суб'єктів. Вимірювання ЕКГ, залежно від суб'єкта, здійснювалося у 1-11 сеансів. Різниця у

часі між першим та останнім вимірювань для підмножини суб'єктів може становити понад 2 роки. Детальні характеристики даного набору даних наведено у додатку Б.

На думку автора уваги також заслуговують наступні набори даних, більшість з яких автор використовував у своїх наукових публікаціях:

MIT-BIH Arrhythmia: Набір електрокардіограм хворих на аритмію [66], один з найбільш популярних наборів у дослідженнях біометричних систем на основі ЕКГ. Даний набір містить 48 сигналів тривалістю по 30 хвилин. Сигнали було виміряно у 47 суб'єктів.

MIT-BIH Normal Sinus Rhythm: Даний набір сформовано з уривків ЕКГ-сигналів набору MIT-BIH Arrhythmia, які виміряно у 18 суб'єктів. У даних уривках відсутні прояви аритмії чи інших захворювань.

ECG-ID [67]: Даний набір є призначеним винятково для біометричних застосувань. Містить сигнали виміряні у 90 суб'єктів (44 чоловіки та 46 жінок віком від 13 до 75 років). Кількість записів для кожного з суб'єктів варіюється від 2 (виміряних протягом одного дня) до 20 (вимірюваних періодично впродовж шести місяців). ЕКГ-сигнал вимірювався зі зап’ястків.

PTB [68]: Набір даних містить 549 записів виміряних у 290 суб'єктів. Містить ЕКГ як здорових суб'єктів, так і хворих на різноманітні серцеві захворювання (такі як інфаркт міокарда, аритмія, гіпертрофія чи серцева дисфункція). Кількість записів для кожного з суб'єктів варіюється від 1 до 5, тривалістю від 38.4 до 104.2 секунд. Записи містять усі 12 стандартних медичних каналів.

QT: Набір даних QT спрямований на сприяння розробці автоматичних методів вимірювання QT хвиль [69]. Даний набір є збіркою сигналів із публічних наборів даних, яка містить сигнали 105 суб'єктів тривалістю 15 хвилин.

UofTDB [70]: Набір даних університету Торонто (The University of Toronto ECG Database) був спеціально створений для біометрії та розкриває кілька важливих критеріїв для ретельної оцінки біометричних показників. Набір містить ЕКГ-записи виміряні “сухими” електродами у 1019 суб'єктів з пальців рук. Кожен

з суб’єктів містить до шести ЕКГ-записів виміряних протягом шести місяців у різних позах: лежачи на спині, після виконання вправ, сидячи та стоячи.

AHA [71]**:** Набір даних підготовлений американською асоціацією серця (American Heart Association) для підготовки медичних працівників з діагностики аритмії. Він містить 154 ЕКГ-записи реальних пацієнтів, подарованих різними установами. Тривалість кожного із записів складає 3 години.

CYBHi (Check Your Biosignals Here initiative) [72]: Набір даних виміряних двома “сухими” електродами з долонь та двома електродами з пальців рук. Складається з двох підмножин: перша містить ЕКГ виміряні у 65 волонтерів впродовж однієї вимірювальної сесії; друга містить ЕКГ виміряні у 63 суб'єктів впродовж двох сесій із перервою у три місяці. Тривалість одного вимірювання – 5 хвилин. Протягом вимірювання суб'єкти переглядали відеоматеріали, котрі повинні були викликати різноманітні емоції.

Відповідно до таблиці 2.4 велика частина з наборів ЕКГ-даних була “перекваліфікована” для вирішення проблеми розпізнавання людини із задач розпізнавання захворювань. Такі набори даних містять достатньо велику підмножину електрокардіограм хворих людей, що може мати негативний вплив при побудові біометричної системи автентифікації, оскільки деякі ознаки захворювання можуть служити певною “підказкою” для автентифікатора.

Не менш важливим є доступність таких наборів даних. Деякі з них є приватними, що ускладнює науковцям їх використання.

Для експериментальної частини дисертаційного дослідження було вирішено використовувати Lviv Biometric Dataset. Оскільки він містить ЕКГ-записи виключно здорових людей. А розміщення електродів при вимірюванні відповідає розміщенню електродів на етапі експлуатації біометричної системи автентифікації.

Таблиця 2.4

Порівняння характеристик наборів ЕКГ даних

Назва	**Доступ**	**Кількість суб'єктів**	**Розміщення електродів**	**Частота дискре-тизації**	**Стан здоров'я**	**Тривалість одного вимірювання**
UofTDB	приватний	1020	пальці	200	здорові	2-5 хв.
PTB	публічний	290	груди та кінцівки	1000	різні захворювання	38.4 - 104.2 с.
AHA	приватний	154	груди	250	різні захворювання	3 год.
CYBHi	публічний	128	долоні та пальці	1000	здорові	до 5 хв.
LBDS	**публічний**	**115**	**пальці**	**277**	**здорові**	**10 с**
QT	публічний	105	груди	250	різні захворювання	15 хв.
ECG-ID	публічний	90	зап’ястя	500	здорові	20 с
MIT-BIH Arrhythmia	публічний	47	груди	360	аритмія та інші	30 хв.
MIT-BIH Normal Sinus Rhythm	публічний	18	груди	360	здорові	-

Також підмножина даного набору даних (кілька вимірювальних сесій протягом двох років) дозволяє здійснити дослідження стійкості електрокардіограми як біометричної характеристики в часі та здійснити ефективне оцінювання алгоритму темпоральної нормалізації.

Висновки до розділу 2

1. Розглянуто особливості процесу автентифікації за ЕКГ- сигналом. Формалізовано структуру біометричної системи розпізнавання. Наведено детальний опис і функції кожного із структурних елементів.

2. Розроблено перспективні підходи до покращення технічних і експлуатаційних характеристик біометричної системи ЕКГ-автентифікації.

Передовсім, обґрунтовано доцільність введення в ланцюг опрацювання електрокардіограми двох додаткових компонент:

• компонента темпоральної нормалізації ЕКГ-сигналу, що покликана забезпечити інваріантність результатів автентифікації до зміни частоти серцевого ритму, тим самим підвищивши достовірність роботи біометричної системи;

• компонента виявлення та виправлення артефактів у ЕКГ-сигналі, яка за допомогою інструментарію статистики або машинного навчання підвищує точність і швидкодію системи біометричної автентифікації.

3. Представлено структурну схему біометричної системи ЕКГ-автентифікації із покращеними технічними і експлуатаційними характеристиками. Продемонстровано вигляд та опис сигналів, які циркулюють всередині системи автентифікації.

4. Подано методики оцінювання ефективності методів і засобів біометричної автентифікації на основі ЕКГ-сигналу.

5. Наведено огляд та співставлення існуючих наборів ЕКГ даних. Представлено сформовану автором базу записів електрокардіограм (Lviv Biometric Dataset). Обґрунтовано використання власного набору електрокардіограм у подальших дослідженнях.

РОЗДІЛ 3. РОЗРОБЛЕННЯ МОДЕЛЕЙ ТА МЕТОДІВ ДЛЯ ПОКРАЩЕННЯ ХАРАКТЕРИСТИК СИСТЕМ БІОМЕТРИЧНОЇ АВТЕНТИФІКАЦІЇ ЗА ЕКГ-СИГНАЛОМ

3.1. Поєднання методів виявлення та виправлення артефактів ЕКГ-сигналу

Одним із способів покращення технічних та експлуатаційних характеристик біометричної системи ЕКГ-автентифікації є використання методів виявлення та виправлення артефактів у виміряних ЕКГ-сигналах. Як зазначено у п. 2.2, якість ЕКГ-сигналу може серйозно деградувати під впливом низки різноманітних причин, наприклад, через м’язовий шум, дихальні завади, зміщення електродів тощо. Деякі з артефактів не піддаються фільтрації, тому потребують опрацювання за допомогою додаткових методів для виявлення та виправлення артефактів у квазіперіодичних сигналах.

Приклади артефактів наведено на рис. 3.1. Зазвичай ЕКГ-сегменти зі значимими артефактами виявити не складно. Найчастіше для цього використовують евклідову відстань та просто відкидають сегмент, навіть якщо артефакт зайняв тільки невелику частину вибірок. Виявлення незначних артефактів є не менш важливим, оскільки вони можуть опинитися в “інформативних” для класифікатора частинах сегменту [73].

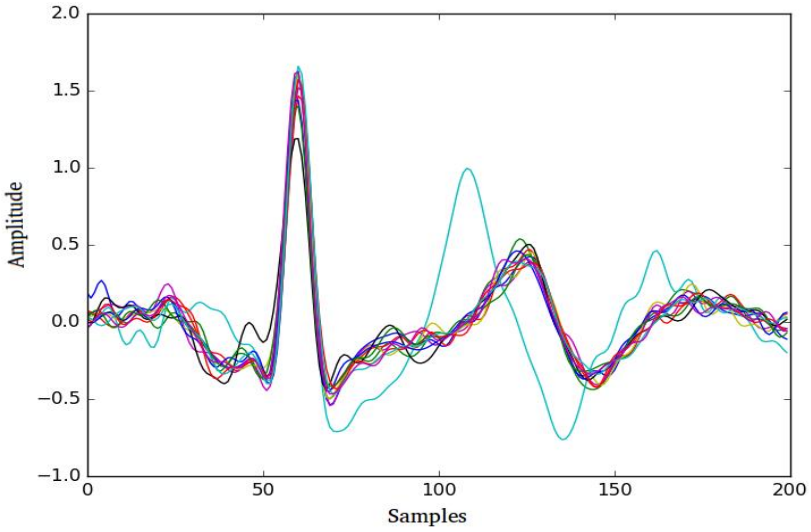


Рис. 3.1. Артефакти у ЕКГ-сигналах: невеликий (блакитна крива) та значний (чорна крива)

В даній роботі розроблено підхід для виявлення та виправлення залишкових артефактів у ЕКГ-сигналах (рис. 3.2), який складається із трьох етапів:

- формування референційного образу біометричного маркера;
- виявлення фрагментів ЕКГ-сигналу із промахами;
- заміна цих фрагментів на відповідні значення із референційного образу.

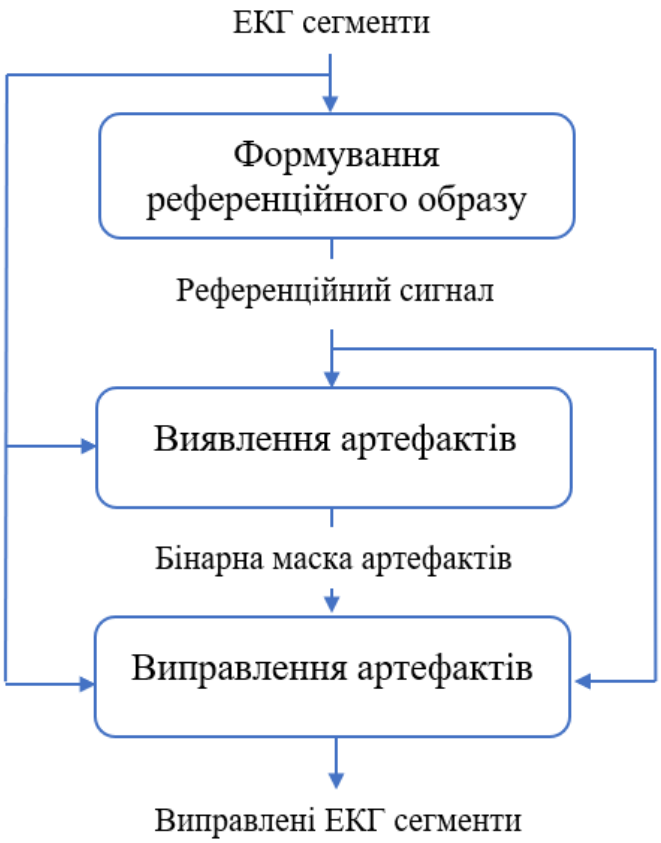


Рис. 3.2. Візуалізація підходу до виявлення та виправлення залишкових артефактів у ЕКГ-сигналах

Пропонований підхід дозволяє виявляти як значні, так і незначні артефакти, а також виправляти артефакти не відкидаючи сегмент. Нижче наведено методи, які розроблено на основі представленого вище підходу.

3.1.1. Статистичний метод для виявлення та виправлення артефактів

Статистичний метод виявлення та виправлення артефактів (рис. 3.3) складається з наступних етапів:

1. Формування референційного образу. Полягає у знаходженні опорного сигналу, який буде використано для виявлення артефактів у вхідних ЕКГ-сегментах. Для даного методу – це вектор усереднених значень для кожної вибірки.

2. Виявлення артефактів. На даному етапі за допомогою референційного образу знаходяться проміжки у кожному сегменті, у яких варіація для хоча б однієї вибірки перевищує визначене порогом значення (variance gain). На виході отримуємо бінарну матрицю з маскою артефактів. Розмірність такої матриці така ж як і розмірність матриці ЕКГ-сегментів.

3. Корекція артефактів. Полягає у заміні проміжків знайдених на другому етапі на еквівалентні проміжки референційного образу.

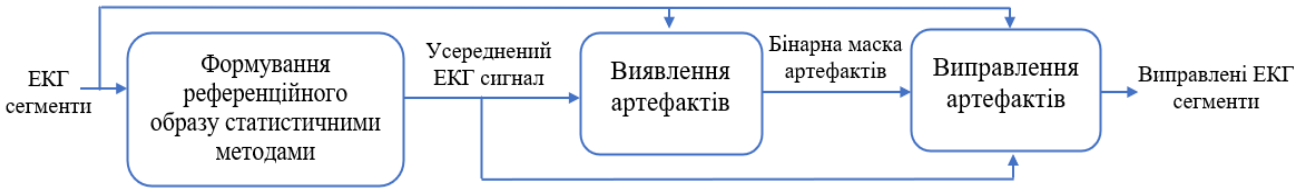


Рис. 3.3. Структура статистичного методу виявлення та виправлення артефактів

Розглянемо детальніше принцип роботи даного методу. На початковому етапі визначаємо вектор усереднених значень для кожної вибірки, що виконуватиме роль референційного образу:

$$\bar{x}(n) = \frac{1}{K}\sum_{k=1}^{K} x(k,n) \quad (3.1)$$

де $x(k,n)$ - елемент матриці ЕКГ-сегментів $X(K,N)$; $k \in 1 \div K$- рядки, які репрезентують ЕКГ-сегменти; $n \in 1 \div N$- колонки, які репрезентують вибірки кожного з ЕКГ-сегментів.

На другому етапі обчислюється вектор стандартних відхилень за допомогою формули (3.2)

$$std(n) = \frac{1}{K-1}\sqrt{\sum_{k=1}^{K}[x(k,n) - \bar{x}(n)]^2} \quad (3.2)$$

Далі знаходимо усереднене значення стандартного відхилення для матриці $X(K,N)$:

$$\overline{std} = \frac{1}{N}\sum_{n=1}^{N} std(n) \tag{3.3}$$

Тепер знаходимо маску артефактів для матриці ЕКГ-сегментів за допомогою формули (3.4)

$$o(k,n) = \frac{[x(k,n) - \bar{x}(n)]^2}{\overline{std}} > variance\ gain \tag{3.4}$$

де $o(k,n)$ - елемент бінарної матриці $O(K,N)$, у якій кожен не нульовий елемент репрезентує виявлений артефакт (у відповідній вибірці відповідного сегменту).

На етапі виправлення артефактів здійснюється наступна трансформація:

$$x(k,n{:}\,n+L) = \begin{cases} \bar{x}(n{:}\,n+L), if\ any\ in\ o(k,n{:}\,n+L) == 1 \\ x(k,n{:}\,n+L), if\ all\ in\ o(k,n{:}\,n+L) == 0 \end{cases} \tag{3.5}$$

де $\bar{x}(n{:}\,n+L)$ – вектор усереднених вибірок в межах вікна L,

$x(k,n{:}\,n+L)$ – вектор вибірок сегменту k в межах вікна L,

$o(k,n{:}\,n+L)$ – вектор маски артефактів для сегменту k в межах вікна L.

Описаний вище метод використовує два гіперпараметри – поріг варіації та розмір вікна корекції. Хоча для всіх ЕКГ-записів використовуються однакові значення гіперпараметрів, метод є адаптивним для кожного з записів, оскільки інформативним параметром для детекції артефактів є внутрішня варіативність вибірок.

Вибір гіперпараметрів здійснюється наступним чином:

1. Грубе налаштування - здійснюється візуально на частині набору даних, результатом є здатність методу знаходити та виправляти значні артефакти. Знайдені на цьому етапі значення гіперпараметрів використовуються як базові на етапі точного налаштування.

2. Точне налаштування — здійснюється разом із грубим налаштуванням, і його мета – знайти деякі невеликі відхилення грубо налаштованих параметрів, що призводить до підвищення точності класифікації.

Варто підкреслити, що точність описаного методу залежить від того, наскільки вдало були обрані гіперпараметри. У випадку, якщо поріг варіації обрано занадто великим, метод коригування може ігнорувати аномальні вибірки в ЕКГ-сегментах. З іншого боку, якщо поріг варіації обрано занадто малим, більшість вибірок буде розпізнано як аномалії та піддано коригуванню, що еквівалентно простому усередненню вибірок у межах запису (див. рис. 3.4). Отже, неточний вибір параметрів призводить до неправильної роботи всього методу виявлення та виправлення артефактів.

На рис. 3.4 наведено результати роботи описаного вище методу виявлення та виправлення артефактів над сегментами електрокардіограми з рис. 3.1. Кожен з вихідних скоректованих сегментів може використовуватися як окремий шаблон для класифікатора.

Згідно з експериментами найчастіше підлягають коригуванню P і T-хвилі сигналу ЕКГ. Кількість виправлень є варіативною, для деяких записів може становити в середньому 2-3 виправлення на 10 сегментів, тоді як на інших записах це може бути навіть 18-20 виправлень.

Для оцінки ефективності методу виявлення та виправлення артефактів було імплементовано біометричну систему автентифікації описану в розділі 2, здійснено тестування системи з та без включення методу виявлення та виправлення артефактів, а також підібрано оптимальні значення гіперпараметрів для статистичного методу виявлення та виправлення артефактів.

Для зменшення розмірності даних використано метод головних компонент (пункт 2.1.4), який трансформує оброблені сегменти у гіперплощину, де їх довжина становить 30 вибірок.

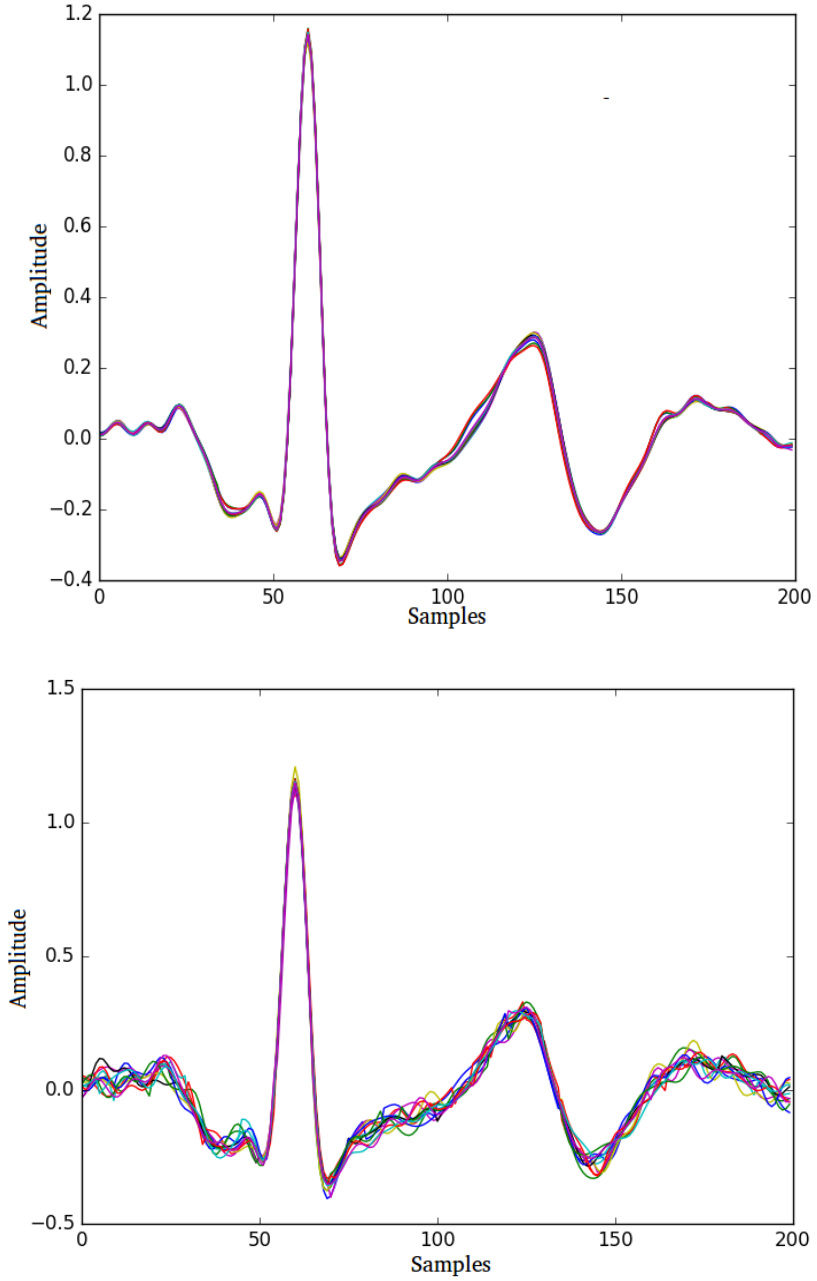


Рис. 3.4. Результат роботи методу виявлення та виправлення артефактів для випадків коли обрано надто малий поріг варіації (верхній) так коли обрано оптимальні значення гіперпараметрів (нижній)

Класифікатор системи автентифікації побудовано на основі лінійного дискримінантного аналізу (пункт 2.1.5). Він ідеально підходить для проведення таких експериментів, оскільки даний класифікатор дає надійні, стійкі та зрозумілі результати, а також він швидко тренується.

Для експерименту використано набір записів ЕКГ Lviv Biometric Dataset, який детально описано у підрозділі 2.5. Записи було розділено на навчальний та тестовий набори у пропорції 70% і 30%. Навчальний набір використано для навчання класифікатора системи біометричної автентифікації. Тестовий набір репрезентує небачені класифікатором дані, на яких здійснюється оцінка його ефективності. Конфігурація навчального та тестового наборів наведена у додатку В.

Для оцінювання ефективності автентифікації було використано наступні метрики:

- точність (англ. Accuracy***)***;
- помилка першого роду (FRR);
- помилка другого роду (FAR).

Експерименти проводились на робочій станції з наступними параметрами: процесор Intel® Core™ i7-4790 CPU @ 3.60GHz × 8, оперативна пам’ять 32Гб, операційна система Ubuntu 18.04.

Даний метод виявлення та виправлення артефактів, як і інші компоненти біометричної системи було імплементовано за допомогою мови програмування Python 3.6. Також було використано такі пакети мови Python: SciPy, NumPy, matplotlib, scikit-learn, biosppy.

Імплементована система біометричної автентифікації без застосування методів виявлення та виправлення артефактів на тестовому наборі даних продемонструвала такі результати:

- accuracy = 89.31%
- FAR = 2.07×10^{-3}
- FRR = 61.09×10^{-3}

Результати роботи біометричної системи із застосуванням статистичного методу для виявлення і виправлення артефактів у ЕКГ-сигналі продемонстровані у табл. 3.1.

Таблиця 3.1

Вибір оптимальних гіперпараметрів для статистичного методу виявлення та виправлення артефактів

Accuracy, % FAR, 10^{-3} FRR, 10^{-3}		Розмір вікна, вибірки			
		3	5	10	15
поріг варіації	0.05	93.19 1.49 46.28	93.37 1.33 42.96	93.37 1.48 45.72	92.45 1.53 49.93
	0.5	95.58 0.67 30.10	95.58 0.66 30.96	95.95 0.63 27.89	95.58 0.72 30.56
	1	95.95 0.58 27.56	95.95 0.61 27.89	95.77 0.52 29.81	95.95 0.50 28.18
	5	96.50 0.44 23.30	**96.69** **0.45** **22.62**	96.13 0.47 25.95	96.32 0.45 25.03
	10	96.32 0.46 23.99	96.32 0.46 23.99	96.50 0.42 24.53	96.50 0.42 24.53
	20	95.95 0.60 26.68	95.76 0.63 27.08	96.13 0.56 27.37	96.13 0.48 27.30

Також у таблиці представлено результати вибору оптимальних гіперпараметрів для описаного вище методу. Результати в таблиці демонструють чутливість методу до кожного з гіперпараметрів.

При довжині вікна 5 вибірок та порозі варіації 5 статистичний метод для виявлення та виправлення артефактів продемонстрував найкращий результат. Застосування даного методу дозволило підвищити точність біометричної системи автентифікації на 7,38%.

3.1.2. Метод виявлення та виправлення артефактів за допомогою нейронних мереж

Даний метод [74 – 77] базується на використанні алгоритмів машинного навчання для детекції аномальних вибірок у ЕКГ-сигналах. Для цього завдання існує безліч технік, але найбільш придатними є автоенкодерні нейронні мережі, або просто автоенкодери [78].

Автоенкодер – це спеціальний тип нейронної мережі, для якої вхідний і вихідний вектор збігаються. Зазвичай, автоенкодери застосовуються для виділення характерних ознак, нелінійної фільтрації та знешумлення даних, зменшення розмірності, тощо. Формально їх можна описати наступним чином:

$$g[f(x)] = x, \tag{3.6}$$

де f – функція енкодера,

g – функція декодера,

x – вхідний вектор.

Не зважаючи на складну назву, нейронні мережі на основі автоенкодерів є відносно простими алгоритмами машинного навчання. Внутрішньо автоенкодер складається з двох частин: кодера та декодера. Алгоритм кодера стискає вхідні значення і представляє їх у формі латентного простору («коду»), тоді як декодер намагається реконструювати вхідні дані з латентного простору. Виконуючи цю процедуру ітераційно, модель знайде оптимальні параметри (ваги), які будуть зберігати реконструйовані значення близькими до вхідних, наскільки це можливо. Цей метод працює, оскільки розмірність вектору в латентному просторі значно менша, ніж вхідного вектора в початковому просторі. Таким чином, кодується найбільш важлива, відтворювана інформація, тоді як шум та артефакти автоматично видаляються. Після декодування вихідний вектор схожий на вхідний, але в ньому вже майже відсутні викривлення та аномалії. В даній роботі було використано так званий «ванільний» автоенкодер з найпростішою архітектурою, яку показано на рис. 3.5 [74].

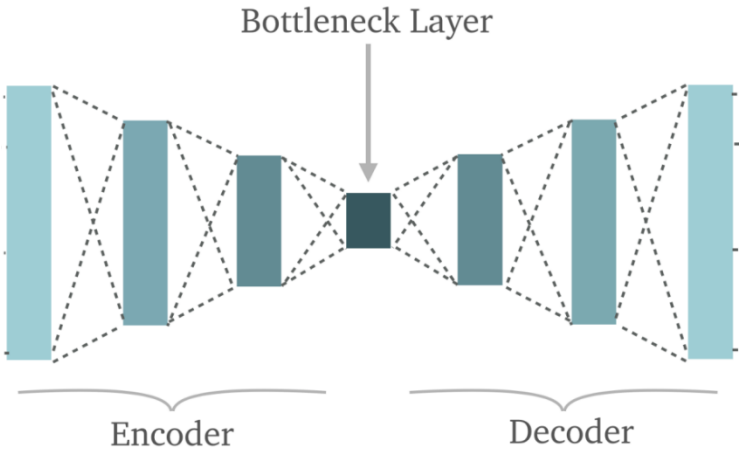


Рис. 3.5. Базова архітектура автоенкодера [78]

Основними перевагами нейронних мереж на основі автоенкодера є те, що вони легко адаптуються до даних, прості в імплементації, швидко навчаються, і відзначаються хорошою обчислювальною ефективністю. При цьому слід звернути увагу на такі моменти:

- автоенкодери є чутливими до даних, а отже модель працюватиме коректно лише з тими даними, на яких вона була натренована (або на даних з схожим розподілом);

- автоенкодери реалізують стиснення із втратами, тобто завжди декомпресовані дані будуть дещо спотворені, порівняно з оригіналом [74].

Даний метод виявлення та виправлення промахів (рис. 3.6) працює так:

1. На етапі формування референційного образу на кожному окремому сегменті ЕКГ тренується модель автоенкодера. Під час навчання підбираються ваги, які мінімізують загальну похибку реконструкції на навчальному наборі даних.

2. Виявлення артефактів здійснюється за допомогою порівняння оригінального та реконструйованого сигналу. На основі помилки реконструкції знаходяться потенційно аномальні вибірки.

4. На етапі виправлення проміжки з аномаліями замінюються середніми значеннями відповідних проміжків решти сегментів, використовуючи ту ж процедуру, що і у попередньому методі для виправлення артефактів (пункт 3.1.1).

Рис. 3.6. Структура методу виявлення та виправлення артефактів за допомогою нейронних мереж

Деякі зразки ЕКГ-сигналів з артефактами до та після коригування візуалізовано на рис. 3.7. Відповідно до експериментів, найчастіше коригування виконується для P та T-хвиль. Комплекс QRS коригується не так часто. Кількість корекцій варіюється для різних користувачів (для деяких корекція не потрібна взагалі, а декому необхідні десятки). З графіків також випливає, що коригування робить серцеві удари більш схожими, але все одно залишається багато відмінностей. Це вважається великою перевагою, так як не зважаючи на те, що сигнали в межах одного класу стають більш схожими між собою, між ними все одно залишаються характерні розбіжності, що дозволяє натренувати більш надійний та стійкий класифікатор [74].

Даний метод має наступні гіперпараметри: поріг помилки реконструкції та довжина вікна корекції.

Для оцінки ефективності даного методу виявлення та коригування артефактів було імплементовано біометричну систему автентифікації, здійснено тестування системи, а також підібрано оптимальні значення гіперпараметрів для методу виявлення артефактів за допомогою автоенкодерів та корекції їх ковзним вікном.

Як і в попередньому експерименті, для зменшення розмірності даних використано метод головних компонент, а класифікатор системи автентифікації побудовано на основі LDA.

Також для експерименту використано набір записів ЕКГ Lviv Biometric Dataset, в такій ж конфігурації навчального та тестового сетів, як і для попереднього експерименту.

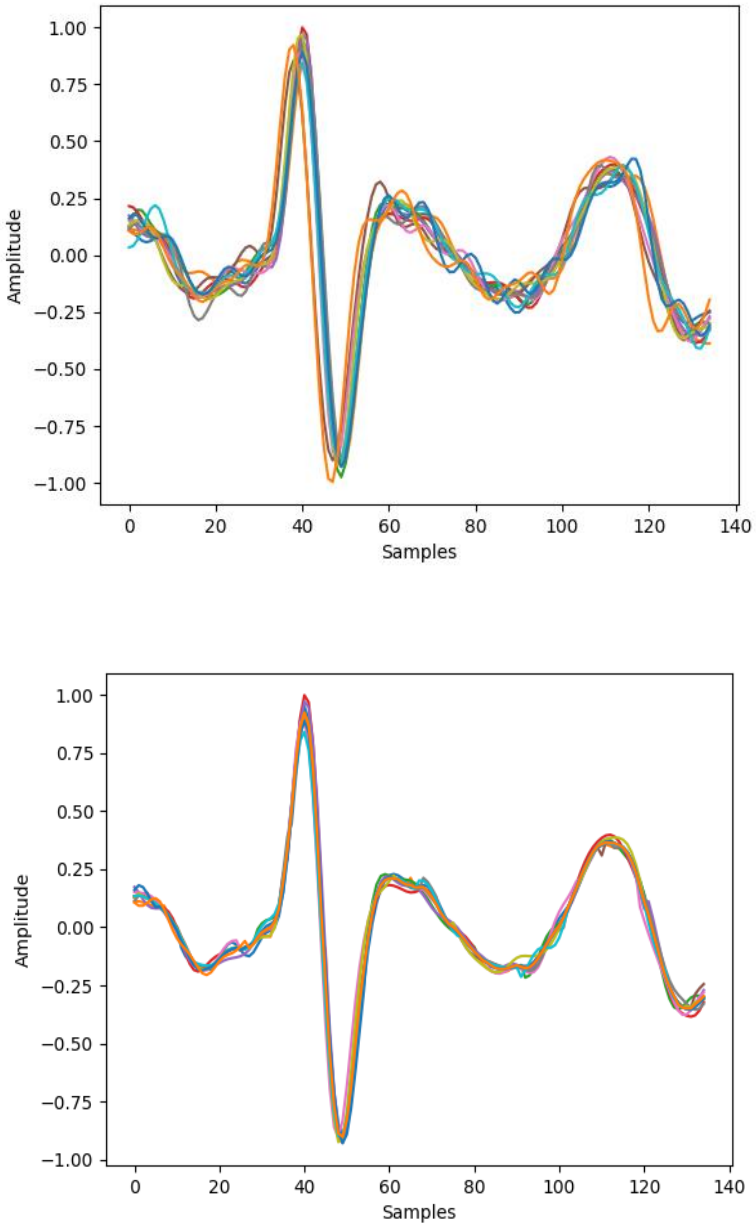


Рис. 3.7. ЕКГ-сегменти до (вгорі) та після (внизу) корекції

Експерименти проводились на робочій станції з наступними параметрами: процесор Intel® Core™ i7-4790 CPU @ 3.60GHz × 8, оперативна пам'ять 32Гб, операційна система Ubuntu 18.04.

Даний метод виявлення та виправлення артефактів, як і інші компоненти біометричної системи було імплементовано за допомогою мови програмування Python 3.6. Також було використано наступні пакети мови Python: SciPy, NumPy,

matplotlib, scikit-learn, biosppy. Модель автоенкодера було побудовано за допомогою модуля H2OAutoEncoderEstimator з бібліотеки h2o.

За допомого оптимізаційного методу пошуку, решіткою було обрано оптимальні гіперпараметри моделі автоенкодера (таблиця 3.2).

Таблиця 3.2

Оптимальні гіперпараметри моделі автоенкодера

Конфігурація прихованих шарів	[100, 100, 100, 100]
Активаційна функція	ReLU (rectified linear unit)
Регуляризація	Dropout
Ініціалізація ваг	UniformAdaptive
l1	1e-5
l2	1e-5

Результати роботи біометричної системи із застосуванням методу виявлення артефактів за допомогою автоенкодерів та їх коригування ковзним вікном, представлено у таблиці 3.3.

При довжині вікна 5 вибірок та порозі помилки реконструкції 80% метод виявлення артефактів за допомогою автоенкодерів з ковзним вікном продемонстрував найкращий результат. Застосування даного методу дозволило підвищити точність біометричної системи автентифікації на 7,19%.

Крім того, було розроблено метод, котрий полягає у використанні реконструйованого сигналу для подальшої автентифікації замість аналізу та обробки помилки реконструкції. Таким чином, етапи виявлення та виправлення артефактів не представляється як окреме перетворення, оскільки всі артефакти будуть виправлені на етапі реконструкції автоенкодера. На відміну від двох попередніх, цей метод не вимагає жодних гіперпараметрів.

Таблиця 3.3

Вибір оптимальних гіперпараметрів для методу виявлення артефактів за допомогою автоенкодерів та їх коригування віконним методом

Accuracy, % FAR, 10^{-3} FRR, 10^{-3}		Розмір вікна, вибірки			
		3	5	10	15
поріг, %	75	95.58 0.67 29.25	96.13 0.54 26.25	96.31 0.54 24.44	95.40 0.70 31.26
	80	95.76 0.63 28.55	**96.50** **0.53** **24.72**	95.95 0.58 27.71	95.40 0.72 31.39
	85	95.40 0.65 35.24	95.58 0.64 31.71	95.40 0.65 32.23	95.40 0.68 31.20
	90	94.66 0.85 35.14	94.84 0.80 35.99	94.84 0.75 36.69	94.84 0.76 36.45
	95	94.29 0.96 37.14	94.29 0.89 39.62	94.11 0.90 40.70	94.66 0.84 39.37

Даний метод працює наступним чином:

- для кожного ЕКГ-запису на множині усіх його сегментів навчається модель автоенкодера;

- через натреновану модель автоенкодера пропускають сегменти ЕКГ-запису та отримують реконструйовані сегменти, які далі використовуються для автентифікації.

Для оцінки ефективності даного методу виявлення та виправлення артефактів було імплементовано біометричну систему автентифікації та здійснено її тестування.

Як і в попередніх експериментах для зменшення розмірності даних використано метод головних компонент, а класифікатор системи автентифікації побудовано на основі LDA.

Також для експерименту використано набір записів ЕКГ Lviv Biometric Dataset, в такій ж конфігурації навчального та тестового сетів як і для попередніх експериментів.

Використано такі ж гіперпараметри моделі автоенкодера як і в попередньому методі (таблиця 3.2).

Імплементована система біометричної автентифікації із застосуванням методу виявлення та виправлення артефактів за допомогою автоенкодерів на тестовому наборі даних продемонструвала наступні результати:

- accuracy = 96.69%
- FAR = 0.48×10^{-3}
- FRR = 23.98×10^{-3}

Застосування даного методу дозволило підвищити точність біометричної системи автентифікації на 7,38%.

3.1.3. Порівняння методів виявлення та виправлення артефактів

У табл. 3.4 наведено порівняння методів виявлення та виправлення артефактів у ЕКГ-сигналах, які детально описано у попередніх пунктах [79].

Як видно з таблиці 3.4, кожен із описаних методів виконав поставлену мету - зробити систему біометричної автентифікації стійкішою до аномальних вибірок і відповідно точнішою. Кожен із методів підвищує точність біометричної системи на понад 7%.

Таблиця 3.4

Порівняння методів виявлення та виправлення артефактів.

Метод	без виявлення та корекції	статистичний метод (3.1.1)	автоенкодер з віконною корекцією (3.1.2)	реконструйований сигнал автоенкодера (3.1.2)
accuracy, %	89.31	96.69	96.50	96.69
FAR, 10^{-3}	2.07	0.45	0.53	0.48
FRR, 10^{-3}	61.09	22.63	24.72	23.98
час, мс	-	10	365	324

Також на основі аналізу наведених результатів можна зробити висновок, що кожен з розроблених методів може бути використаний залежно від задач, які стоять перед проектантом біометричної системи автентифікації. Наприклад, якщо будується система на малопотужних пристроях та час обробки сигналу є критичним – слід обрати статистичний метод для виявлення та виправлення артефактів. Оскільки він є відчутно швидшим порівняно з іншими, проте вимагає підбору оптимальних гіперпараметрів під час тренування.

Коли ж час обробки сигналу не є критичним - слід обрати метод для виявлення та виправлення артефактів на основі автоенкодерів. Він є повільнішим, проте не потребує підбору оптимальних гіперпараметрів.

3.2. Темпоральна нормалізація ЕКГ-сигналу

Покращення технічних та експлуатаційних характеристик біометричної системи ЕКГ-автентифікації також досягається при використанні алгоритмів темпоральної нормалізації ЕКГ-сигналу.

У повсякденному використанні система біометричної автентифікації має бути інваріантною щодо зміни серцевого ритму, зумовленої емоційними, фізичними чи іншими чинниками. Проблема наборів даних, які використовуються для навчання класифікаторів, полягає у тому, що вимірювання ЕКГ-сигналу зазвичай проводилось в один день впродовж короткого проміжку часу. Як наслідок, значення серцевого ритму в одержаних записах є доволі сталими. Дослідження показали, що класифікатори з поміж іншого формували характерні ознаки, які базуються на тривалості серцевого ритму. Це, зазвичай, є хибним, оскільки, у застосуваннях автентифікації збільшуватиме ймовірність помилкових відмов [80, 81].

Тому для сигналу ЕКГ, як функції у 2D-просторі, окрім нормалізації за амплітудою, потрібно застосувати нормалізацію в часі. Суть запропонованого підходу полягає у тому, щоб масштабувати кожен сегмент електрокардіограми до заданого стандартного вікна, максимально зберігаючи форму ЕКГ-сигналу, залежну від специфічних фізіологічних особливостей суб’єкта. Це забезпечить толерантність системи розпізнавання щодо стиснення/розтягнення ЕКГ-сигналу як у вертикальному, так і у горизонтальному напрямках [81].

Для забезпечення інваріантності класифікатора щодо частоти серцевих скорочень було розроблено та досліджено метод унормування тривалості кожного сегменту. Особливістю цього методу є приведення ЕКГ-сигналів від усіх людей до певного, наперед заданого та однакового для всіх значення ритму. На вхід алгоритму (рис. 3.8) подається набір ЕКГ-сегментів еквівалентних одному серцевому циклу.

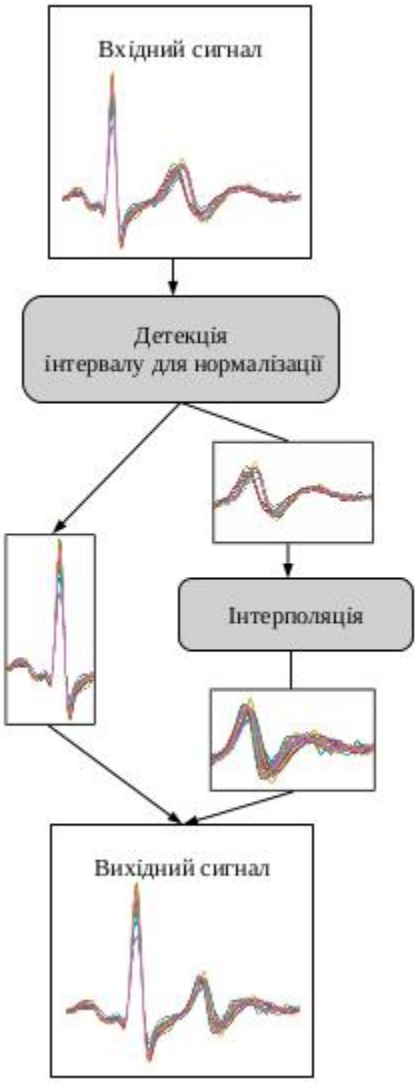


Рис. 3.8. Структурна схема алгоритму темпоральної нормалізації ЕКГ-сигналу

На ЕКГ серцевий цикл розділений на зубці та інтервали, кожен з яких відповідає певній фазі розповсюдження хвилі збудження у міокарді (рис. 3.9).

Інтервал від початку Р-зубця до кінця QRS-комплексу практично залишається незмінним, тобто мало залежить від зміни серцевого ритму, а тому на рис. 3.9 зазначений, як сталий інтервал. Отже, завдання блоку детекції для темпоральної нормалізації полягає у знаходженні та виділенні із сигналу кожного циклу серцебиття змінного інтервалу від початку ST-сегмента до кінця U-зубця.

Вузол інтерполяції лінійно інтерполює інтервали для нормалізації таким чином, щоб їх тривалість забезпечувала задане значення серцевого ритму. На виході

алгоритму формується набір сегментів ЕКГ-сигналу з нормалізованим серцевим ритмом, шляхом сполучення сталих інтервалів із відповідними інтерпольованими.

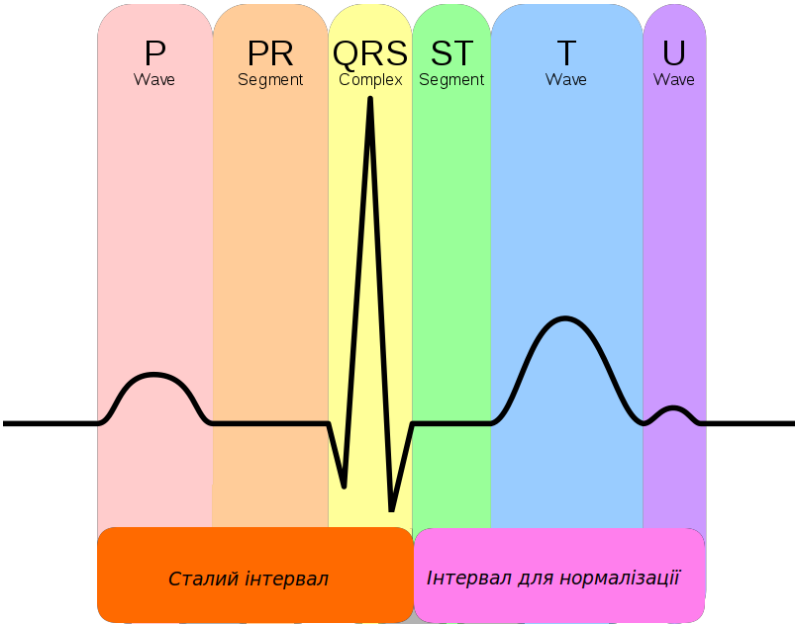


Рис. 3.9. Структура ЕКГ-сигналу в межах серцевого циклу

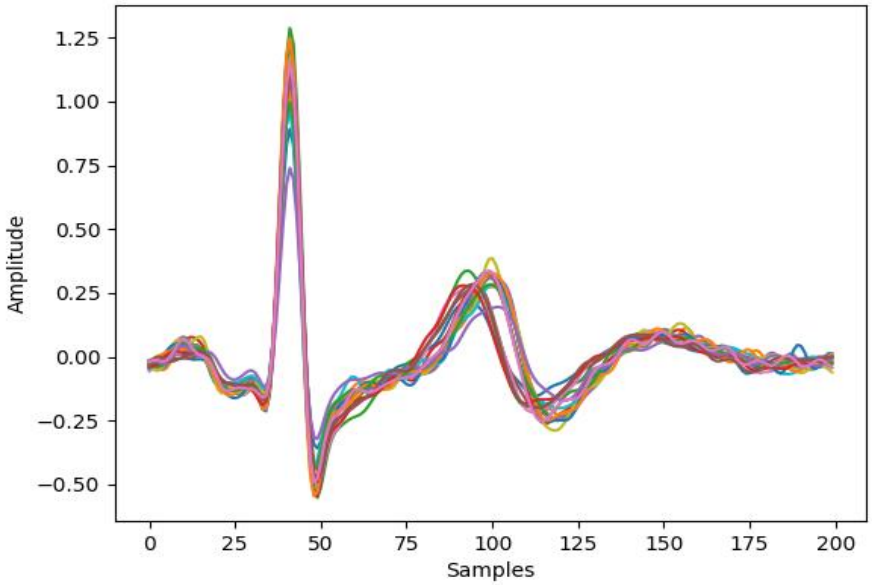
Вхідний набір ЕКГ-сигналів сформовано з результатів двох вимірювань, виконаних з інтервалом в один місяць (рис. 3.10 а). Під час першого вимірювання отримано 6 циклів ЕКГ-сигналу із серцевим ритмом 90 уд/хв (ударів за хвилину), під час другого – 10 із серцевим ритмом 75 уд/хв. На виході алгоритму темпоральної нормалізації усі сигнали приведено до серцевого ритму 120 уд/хв. На рис. 3.10 б можна візуально побачити результати роботи алгоритму, але важливо дослідити та кількісно оцінити, як темпоральна нормалізація впливає на точність автентифікації.

Для оцінки ефективності описаного вище алгоритму темпоральної нормалізації було імплементовано біометричну систему автентифікації описану в розділі 2, а також блоки, які покращують технічні та експлуатаційні характеристики такої системи.

Експерименти проводилися в таких конфігураціях:

– без темпоральної нормалізації;
– з темпоральною нормалізацією до значення еквівалентного 90 ударам серця за хвилину;
– з темпоральною денормалізацією.

а)



б)

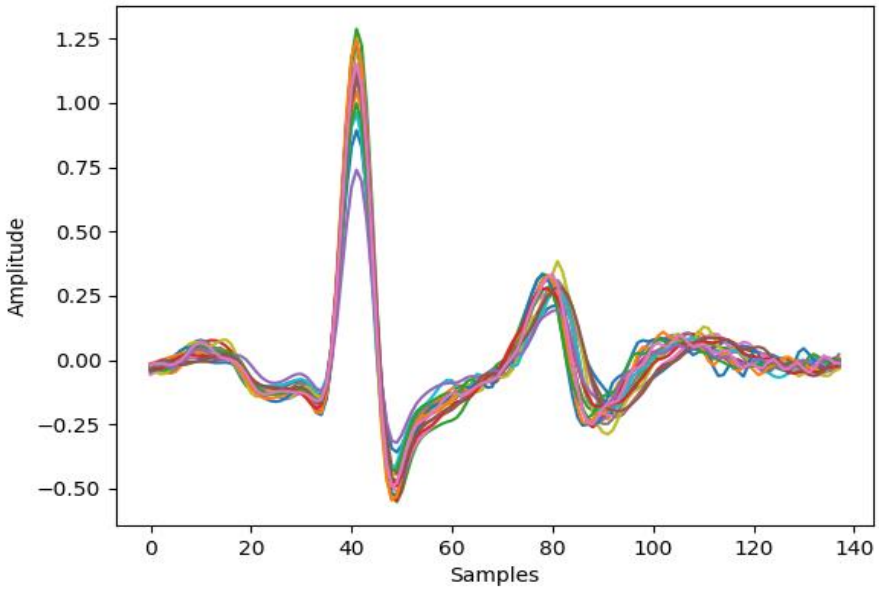


Рис. 3.10. Сегментований ЕКГ-сигнал до (а) та після (б) темпоральної нормалізації

Оскільки у базі даних переважно містяться електрокардіограми зареєстровані впродовж короткого періоду часу, значення серцевого ритму у цих вимірюваннях є сталим. Тестування алгоритму темпоральної нормалізації ЕКГ-сигналу на цих вимірюваннях не репрезентує його ефективності. Тому додатково було імплементовано темпоральний денормалізатор ЕКГ-сигналу, який для кожного вимірювання змінює значення серцевого ритму на випадкове, що міститься на проміжку 60-120 ударів за хвилину.

Також для наведених вище конфігурацій, пропонується зробити опціональним використання алгоритму коригування артефактів, оскільки він значною мірою підвищує точність, як системи з темпоральною нормалізацією ЕКГ-сигналу, так і системи без нормалізації. Це повинно підвищити репрезентативність застосування алгоритму темпоральної нормалізації у сучасних системах, оскільки виправлення артефактів є їх невіддільною складовою частиною. Для виявлення та виправлення артефактів використано статистичний метод, який детально описано у попередньому підрозділі.

Для зменшення розмірності даних використано метод головних компонент. Класифікатор системи автентифікації побудовано на основі лінійного дискримінантного аналізу. Він ідеально підходить для проведення таких експериментів, оскільки даний класифікатор дає надійні, стійкі та зрозумілі результати, а також він швидко тренується.

Для експерименту використано набір записів ЕКГ Lviv Biometric Dataset. Набір було розділено на навчальний та тестовий набори у пропорції 70 / 30%. Конфігурація навчального та тестового наборів наведена у додатку В.

Для оцінювання ефективності автентифікації було використано наступні метрики:

– точність (Accuracy);
– помилка першого роду (FRR);
– помилка другого роду (FAR).

Експерименти проводились на робочій станції з наступними параметрами: процесор Intel® Core™ i7-4790 CPU @ 3.60GHz × 8, оперативна пам'ять 32Гб, операційна система Ubuntu 18.04.

Точність розпізнавання електрокардіограм із тестового набору наведено в табл. 3.5.

Таблиця 3.5.

Продуктивність алгоритму нормалізації

Конфігурація	**accuracy, %** **FAR, 10^{-3}** **FRR, 10^{-3}**
Нормалізований серцевий ритм	86.00 2.28 106.28
Нормалізований серцевий ритм + коректор артефактів	95.95 0.63 25.54
Без нормалізації серцевого ритму	89.31 2.07 61.09
Без нормалізації серцевого ритму + коректор артефактів	96.69 0.45 22.63
Денормалізований серцевий ритм	69.43 8.97 181.08
Денормалізований серцевий ритм + коректор артефактів	87.63 1.87 83.42

Дані, наведені у таблиці 3.5, на перший погляд, показали негативний результат застосування темпоральної нормалізації – точність розпізнавання знизилася (приблизно на 1%), якщо порівняти з варіантом без нормалізації серцевого ритму. Це пояснюється тим, що класифікатор системи автентифікації без нормалізації сигналу ЕКГ використовує значення серцевого ритму як

інформативний параметр. У системах автентифікації з нормалізацією значення серцевого ритму приводиться до заданого, однакового для всіх значення, тому так навчений класифікатор допускає помилку 2-роду, відхиляючи кардіограми справжніх суб'єктів. Застосування алгоритму темпоральної нормалізації усуває описану вище проблему. Точність систем з темпоральною нормалізацією ЕКГ-сигналу є приблизно на 10% вищою ніж у систем, на вхід яких подається денормалізований сигнал. Це підтверджує, що алгоритм темпоральної нормалізації є важливим і необхідним компонентом у реальній системі біометричної автентифікації.

Особливістю доступних наборів електрокардіограм є те, що їх запис для кожного суб'єкта виконувався впродовж короткого інтервалу часу. Значення серцевого ритму у цих вимірюваннях є практично однаковим. Класифікатори біометричних систем використовують цю особливість, як інформаційну характеристику, за якою вони здійснюють автентифікацію. Тому, у реальних сценаріях зміна серцевого ритму призводить до помилок другого роду. Цю проблему можна вирішити одним із трьох підходів: розширенням обсягу даних, використанням темпоральної нормалізації ЕКГ-сигналу або застосуванням темпоральної денормалізації наборів даних для тренування класифікатора.

3.3. Вибір оптимального алгоритму класифікації для побудови системи автентифікації

Після того як сформовано компоненти та параметри біометричної системи автентифікації з покращеними експлуатаційними та технічними характеристиками, можна здійснити вибір оптимального алгоритму класифікації для побудови системи автентифікації.

В даному дослідженні було використано наступні архітектури (детально описані в пункті 2.1.5) для побудови класифікатора системи біометричної автентифікації на основі ЕКГ-сигналу:

- Метод опорних векторів;
- Лінійний дискримінантний аналіз;

- K найближчих сусідів;
- Дерева рішень;
- Нейронні мережі.

Для кожної з архітектур було визначено набір важливих гіперпараметрів та їх можливі значення. Гіперпараметри архітектури – це параметри, які не беруть участь у навчанні, а безпосередньо описують архітектуру. Наприклад, для нейронних мереж - це кількість прихованих прошарків та нейронів у них чи тип активаційної функції. За допомогою методу оптимізації гіперпараметрів методом пошуку ґраткою було здійснено вибір оптимальних гіперпараметрів для кожної з архітектур.

Для експериментів використано набір записів ЕКГ Lviv Biometric Dataset, який детально описано у підрозділі 2.5. Набір було розділено на навчальний та тестовий набори у пропорції 70 / 30%. Навчальний набір використано для навчання класифікатора системи біометричної автентифікації. Тестовий набір репрезентує небачені класифікатором дані на яких здійснюється оцінка його ефективності. Конфігурація навчального та тестового наборів наведена у додатку В.

Для даного дослідження імплементовано біометричну систему біометричної автентифікації [82], яка детально описана в попередньому розділі. В дану систему включено компонент виявлення та корекції артефактів на основі статистичного методу, а також компонент, який здійснює темпоральну нормалізацію.

Кожен з класифікаторів імплементовано за допомогою інструментарію пакету sklearn. Для пошуку ґраткою використано GridSearchCV[83] з пакету sklearn.

Для оцінювання ефективності автентифікації було використано наступні метрики:

- точність (Accuracy);
- помилка першого роду (FRR);
- помилка другого роду (FAR).

Експерименти проводились на робочій станції з наступними параметрами: процесор Intel® Core™ i7-4790 CPU @ 3.60GHz × 8, оперативна пам'ять 32 Гб, операційна система Ubuntu 18.04.

Додатково, окрім гіперпараметрів автентифікатора в експериментах обирався гіперпараметр методу головних компонент (алгоритму зменшення розмірності даних) - кількість компонент (pca_n_components).

Метод опорних векторів [84].

Для методу опорних векторів було виділено наступні гіперпараметри (таблиця 3.5), серед яких було здійснено пошук оптимальних для побудови автентифікатора біометричної системи:

- **C** - параметр регуляризації. Ефективність регуляризації є обернено пропорційною C. Повинен мати тільки позитивні значення;
- **kernel** - ядро методу опорних векторів. Можливі значення - лінійне (linear), поліноміальне (poly), на основі радіальної базисної функції (rbf) чи сигмоїдальне (sigmoid). Для лінійно не роздільних даних ефективно використовувати поліноміальне ядро або ядро на основі радіальної базисної функції

Таблиця 3.5.

Гіперпараметри та їх можливі значення для класифікатора на основі методу опорних векторів

Гіперпараметр	**Можливі значення**
pca_n_components	20, 30, 40, 50, 60
C	0.5, 1, 1.5
kernel	linear, poly, rbf, sigmoid

У таблиці 3.6. наведено три найкращі конфігурації гіперпараметрів, отримані за допомогою пошуку ґраткою. Найвищу точність одержано при конфігурації C=1.5, kernel=linear, pca_n_components=30.

Таблиця 3.6.

Оптимальні конфігурації гіперпараметрів для класифікатора на основі методу опорних векторів

Конфігурація	**accuracy**, %	**FAR**, 10^{-3}	**FRR**, 10^{-3}
C=1.5, kernel=linear, pca_n_components=30	98.16	0.19	13.35
C=1.5, kernel=rbf, pca_n_components=50	97.97	0.23	16.99
C=1, kernel=linear, pca_n_components=50	97.97	0.18	17.45

К найближчих сусідів [85]

Для методу K найближчих сусідів було виділено наступні гіперпараметри (таблиця 3.7), серед яких було здійснено пошук оптимальних для побудови автентифікатора біометричної системи:

- **n_neighbors** - кількість сусідів, на основі яких робиться класифікація.
- **weights -** функція ваг, яка використовується для прогнозування. Можливі значення - 'uniform' (рівномірні ваги, усі сусіди мають однаковий вплив на прогнозування) та 'distance' (вага, на основі оберненої дистанції до сусіда, сусіди, які знаходяться ближче матимуть більший вплив на прогнозування);
- **algorithm** - варіант імплементації алгоритму, який використовується для обчислення найближчих сусідів.

Таблиця 3.7.

Гіперпараметри та їх можливі значення для класифікатора на основі методу К найближчих сусідів

Гіперпараметр	**Можливі значення**
pca_n_components	20, 30, 40, 50, 60
n_neighbors	3, 5, 8, 10
weights	uniform, distance
algorithm	auto, ball_tree, kd_tree, brute

У таблиці 3.8. наведено три найкращі конфігурації гіперпараметрів отримані за допомогою пошуку ґраткою. Найвищу точність одержано при конфігурації algorithm=auto, n_neighbors=3, weights=distance, pca_n_components=20.

Таблиця 3.8.

Оптимальні конфігурації гіперпараметрів для класифікатора на основі методу К найближчих сусідів

Конфігурація	**accuracy**, %	**FAR**, 10^{-3}	**FRR**, 10^{-3}
algorithm=auto, n_neighbors=3, weights=distance, pca_n_components=20	97.61	0.25	17.93
algorithm=brute, n_neighbors=5, weights=distance, pca_n_components=20	97.42	0.31	19.31
algorithm=kd_tree, n_neighbors=8, weights=distance, pca_n_components=30	97.24	0.34	20.69

Дерева рішень [86]

Для методу на основі дерев рішень було виділено наступні гіперпараметри (таблиця 3.9), серед яких було здійснено пошук оптимальних для побудови автентифікатора біометричної системи:

- **max_depth** - максимальна голубина дерева;
- **min_samples_split** - мінімальна кількість вибірок, необхідних для розділення внутрішнього вузла.
- **criterion** - функція оцінки якості розбиття. Можливі критерії - gini на основі коефіцієнту Джині [87] та entropy на основі міри інформації.

Таблиця 3.9.

Гіперпараметри та їх можливі значення для класифікатора на основі дерев рішень

Гіперпараметр	**Можливі значення**	**Гіперпараметр**	**Можливі значення**
pca_n_components	20, 30, 40, 50, 60	min_samples_split	2, 3, 4
max_depth	3, 5, 8, 10	criterion	gini, entropy

У таблиці 3.10. наведено три найкращі конфігурації гіперпараметрів отримані за допомогою пошуку ґраткою. Найвищу точність одержано при конфігурації criterion=entropy, max_depth=10, min_samples_split=3, pca_n_components=20.

Таблиця 3.10.

Оптимальні конфігурації гіперпараметрів для дерев рішень

Конфігурація	**accuracy**, %	**FAR**, 10^{-3}	**FRR**, 10^{-3}
criterion=entropy, max_depth=10, min_samples_split= 3, pca_n_components=20	85.28	2.08	103.36
criterion=entropy, max_depth=10, min_samples_split=2, pca_n_components=20	84.72	2.18	117.00
criterion='entropy, max_depth=10, min_samples_split= 2, pca_n_components=30	84.35	2.15	110.03

Лінійний дискримінантний аналіз [88]

Для методу на основі лінійного дискримінантного аналізу було виділено наступний гіперпараметр (таблиця 3.11), серед значень якого було здійснено пошук оптимальних для побудови автентифікатора біометричної системи:

solver - розв'язувальний алгоритм. Можливі значення: svd (сингулярний розклад матриці. Не обчислює матрицю коваріації, тому цей алгоритм рекомендується для даних з великою кількістю ознак), eigen (базується на оптимізації коефіцієнта на основі відношення величини розкиду між класами до величини розкиду в середині класу) та не менш ефективний lsqr (least squares solution).

Таблиця 3.11.

Гіперпараметри та їх можливі значення для класифікатора на основі лінійного дискримінантного аналізу

Гіперпараметр	**Можливі значення**
pca_n_components	20, 30, 40, 50, 60
solver	svd, lsqr, eigen

У таблиці 3.12. наведено три найкращі конфігурації гіперпараметрів, отримані за допомогою пошуку ґраткою. Найвищу точність одержано при конфігурації solver=svd, pca_n_components=20.

Таблиця 3.12.

Оптимальні конфігурації гіперпараметрів для класифікатора на основі методу лінійного дискримінантного аналізу

Конфігурація	**accuracy**, %	**FAR**, 10^{-3}	**FRR**, 10^{-3}
solver=svd, pca_n_components=40	96.50	0.56	19.83
solver=lsqr, pca_n_components=50	96.32	0.48	23.24
solver=svd, pca_n_components=20	96.13	0.53	25.17

Нейронні мережі [89]

Для методу на основі нейронних мереж було виділено наступні гіперпараметри (таблиця 3.13), серед яких було здійснено пошук оптимальних для побудови автентифікатора біометричної системи:

- **hidden_layer_sizes** - кількість прихованих шарів у нейронній мережі та кількість нейронів у них;
- **activation -** тип активаційної функції;
- **alpha** - параметр регуляризації;
- **solver** - алгоритм оптимізації ваг. Можливі значення: lbfgs (оптимізатор із сімейства квазіньютонівських методів), sgd (відноситься до стохастичного градієнтного спуску) та adam (відноситься до стохастичного оптимізатора на основі градієнта, запропонованого Kingma, Diederik та Jimmy Ba[90]).

Таблиця 3.13.

Гіперпараметри та їх можливі значення для класифікатора на основі нейронних мереж

Гіперпараметр	**Можливі значення**
pca_n_components	20, 30, 40, 50, 60
hidden_layer_sizes	(256), (256, 128), (512, 128), (256, 64), (256, 128, 32), (256, 64, 32), (128, 64)
activation	logistic, tanh, relu
solver	lbfgs, sgd, adam
alpha	0.0001, 0.001

У таблиці 3.14. наведено три найкращі конфігурації гіперпараметрів отримані за допомогою пошуку ґраткою. Найвищу точність одержано при конфігурації pca_n_components=30, hidden_layer_sizes=(256, 128), solver=adam, activation=tanh, alpha=0.0001.

Таблиця 3.14.

Оптимальні конфігурації гіперпараметрів для класифікатора на основі нейронних мереж

Конфігурація	**accuracy**, %	**FAR**, 10^{-3}	**FRR**, 10^{-3}
pca_n_components=30, hidden_layer_sizes=(256, 128), solver=adam, activation=tanh, alpha=0.0001	98.90	0.18	7.32
pca_n_components=30, hidden_layer_sizes=(256, 64), solver=sgd, activation=tanh, alpha=0.0001	98.34	0.19	12.09
pca_n_components=40, hidden_layer_sizes=(256, 128), solver=adam, activation=relu, alpha=0.0001	98.16	0.20	10.59

На основі результатів наведених у таблиці 3.15, можна зробити висновок, що кожен із запропонованих методів, за винятком методу на основі дерев рішень, для побудови класифікатора біометричної системи автентифікації демонструє хороші результати.

Найвищий результат продемонструвала імплементація класифікатора на основі нейронних мереж (точність - 98,90%). Дана конфігурація буде використана для тестування можливостей масштабування біометричної системи автентифікації та дослідження стійкості ЕКГ-сигналу в часі.

Таблиця 3.15.

Зведена таблиця найкращих конфігурації для кожної із архітектур.

Архітектура	**Конфігурація гіперпараметрів**	**accuracy**, %	**FAR**, 10^{-3}	**FRR**, 10^{-3}
Метод опорних векторів	C=1.5, kernel=linear, pca_n_components=30	98.16	0.19	13.35
Метод К найближчих сусідів	algorithm=auto, n_neighbors=3, weights=distance, pca_n_components=20	97.61	0.25	17.93
Метод на основі дерев рішень	criterion=entropy, max_depth=10, min_samples_split= 3, pca_n_components=20	85.28	2.08	103.36
Лінійний дискримінантний аналіз	solver=svd, pca_n_components=40	96.50	0.56	19.83
Нейронні мережі	pca_n_components=30, hidden_layer_sizes=(256, 128), solver=adam, activation=tanh, alpha=0.0001	**98.90**	**0.18**	**7.32**

У таблиці 3.16. наведено набір метрик, детально описаних у підрозділі 2.4, для найкращої конфігурації гіперпараметрів архітектури на основі нейронних мереж .

Таблиця 3.16.

Значення метрик для найкращої архітектури.

Accuracy, %	98.90	**Precision, %**	99.05
FAR	0.00018	**Recall, %**	98.90
FRR	0.00732	**F1 score, %**	98.97

В таблиці 3.17 наведено типові значення параметрів FAR та FRR для сучасних методів біометричної автентифікації. З таблиці випливає, що імплементована біометрична система автентифікації на основі ЕКГ є конкурентоспроможною з системами на основі наведених біометричних параметрів.

Таблиця 3.17 .

Типові значення точності систем автентифікації [91]

Біометричний параметр	**FRR**	**FAR**
Відбиток пальця	0,01 - 0,0001	0,002 - 0,0001
Геометрія кисті руки	0,001	0,000001
Райдужна оболонка ока	0,009	0,000001
Тривимірне зображення обличчя	0,103	0,0047

3.4. Масштабування системи біометричної автентифікації

Метою даного експерименту є дослідження здатності до масштабування біометричної системи автентифікації на основі ЕКГ-сигналу, а саме визначення впливу від збільшення кількості користувачів на точність автентифікації.

Для даного дослідження було імплементовано біометричну систему автентифікації, яку детально описано у попередньому розділі. В дану систему включено компонент виявлення та виправлення артефактів на основі статистичного методу, а також компонент, який здійснює темпоральну нормалізацію. Для зменшення розмірності даних використано метод головних компонент, який зменшує розмірність кожного з сегментів до 30 вибірок. Відповідно до підрозділу 3.3, класифікатор побудовано на основі штучних нейронних мереж з наступними гіперпараметрами: hidden_layer_sizes=(256, 128), solver=adam, activation=tanh, alpha=0.0001.

Для оцінювання ефективності автентифікації було використано наступні метрики: точність (Accuracy***)***, помилка першого роду (FRR) та помилка другого роду (FAR).

Для дослідження здатності біометричної системи автентифікації до масштабування було натреновано моделі класифікатора на підмножинах персон з набору записів ЕКГ Lviv Biometric Dataset. Конфігурація даних підмножин наведена у додатку Г. Дане дослідження дозволяє визначити як змінюється точність біометричної системи автентифікації, при збільшенні кількості її користувачів (таблиця 3.18).

Таблиця 3.18.

Точність біометричної системи автентифікації при різній кількості персон

Кількість персон	**FAR, 10^{-3}**	**FRR, 10^{-3}**	**accuracy, %**
115	0.179	7.315	98.90
110	0.126	6.113	98.97
100	0.113	5.969	99.04
90	0.115	6.311	99.02
80	0.121	6.228	99.24
70	0.123	5.856	99.38
60	0.105	2.244	99.66
50	0.000	0.000	100.00
40	0.000	0.000	100.00
30	0.000	0.000	100.00
20	0.000	0.000	100.00
10	0.000	0.000	100.00

На рисунках 3.11 та 3.12 відображено вплив збільшення кількості користувачів біометричної системи автентифікації на її основні метрики. Імплементована система працює безпомилкова на множинах користувачів обмежених 50 персонами. Далі зі збільшенням користувачів значення метрик поступово погіршується.

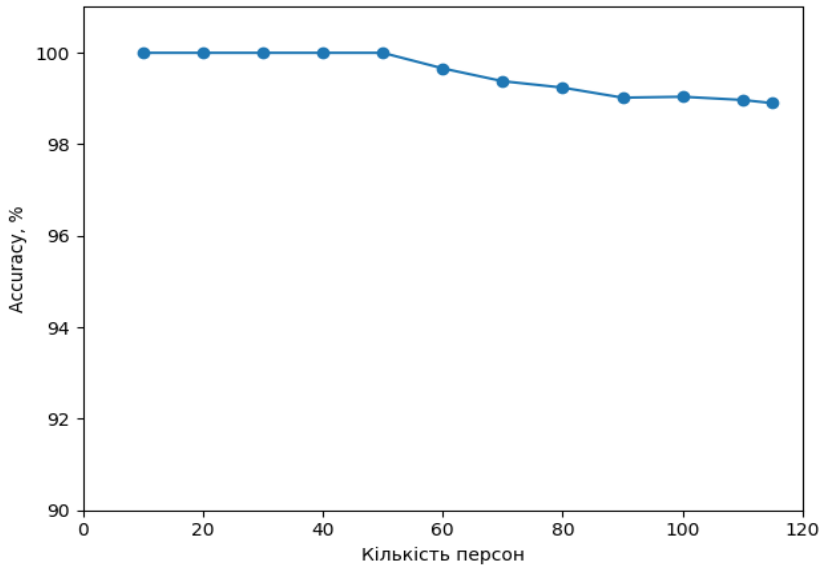


Рис. 3.11. Вплив кількості персон на точність біометричної системи автентифікації

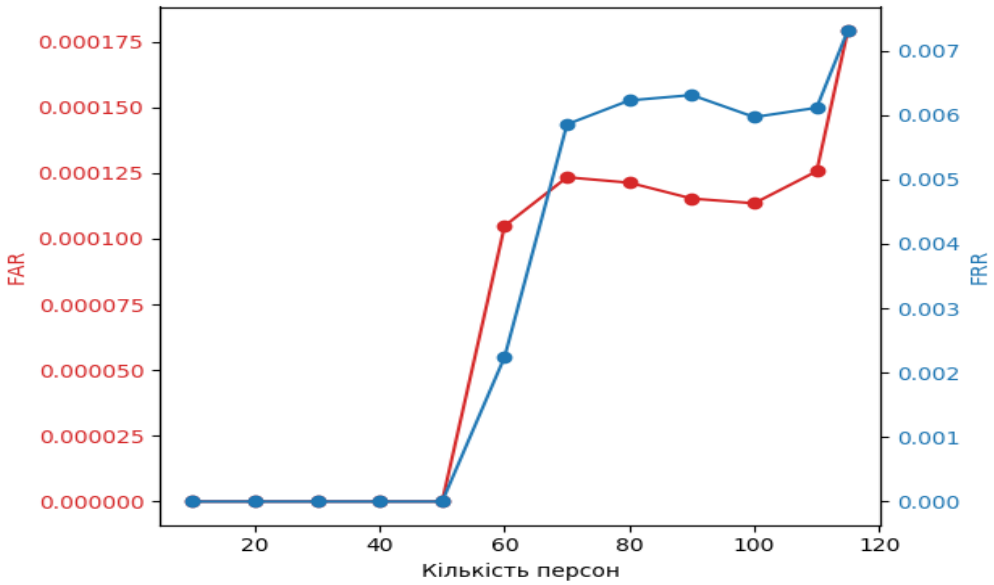


Рис. 3.12. Вплив кількості персон на значення FAR та FRR біометричної системи автентифікації

Accuracy порівняно із FAR та FRR описується більш лінійно. Саме тому Accuracy було використано для екстраполяції точності на проміжку [120, 300] персон (рис. 3.13). За допомогою екстраполяції встановлено, що точність біометричної системи для автентифікації 200 користувачів становитиме 97.71%, для 300 - 96.31%. Це достатньо хороші результати як для біометричної системи автентифікації на основі ЕКГ.

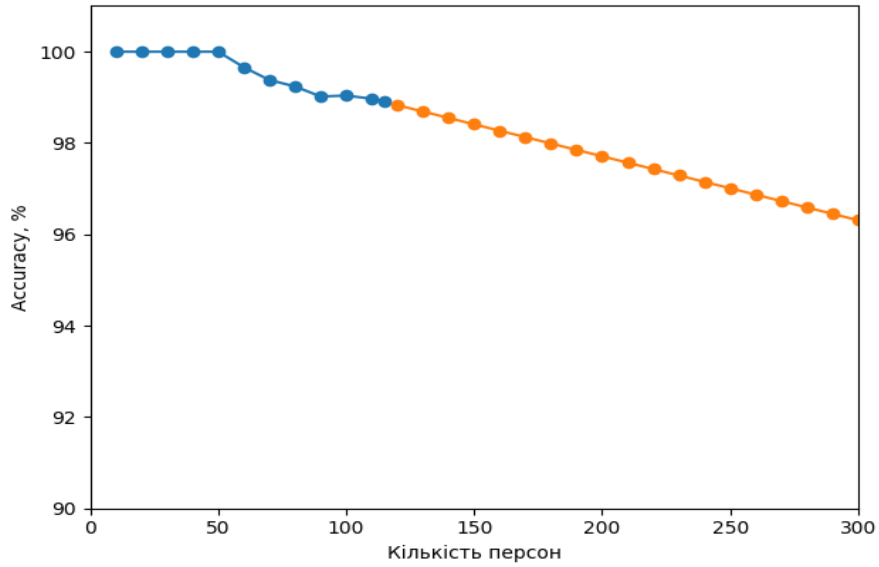


Рис. 3.13. Екстрапольований вплив кількості персон на точність біометричної системи автентифікації (помаранчева крива - екстрапольовані дані, блакитна - експериментальні дані)

Варто також відзначити, що на точність біометричної системи впливає не тільки кількість користувачів чи якість їх записів, але й кількість записів для кожного з користувачів. В майбутньому автор планує здійснити дослідження здатності до масштабування біометричної системи автентифікації одночасно у двох площинах - кількість персон та кількість записів для персони. Для цього планується збільшувати не тільки кількість персон у Lviv Biometric Dataset, але й кількість записів для існуючих персон.

Висновки до розділу 3

1. Розроблено такі методи для виявлення та виправлення артефактів у ЕКГ-сигналах:

- статистичний метод для виявлення та виправлення артефактів, за яким коригування застосовується до тих вибірок у межах ковзного вікна, статистичні характеристики яких перевищують встановлений поріг;
- виявлення артефактів за допомогою автоенкодера та їх коригування віконним методом, який базується на використанні алгоритмів машинного навчання для детекції аномальних вибірок у ЕКГ-сигналах. Виправлення артефактів, як і для попереднього методу здійснюється віконним методом;
- виявлення та виправлення артефактів за допомогою автоенкодера, котрий полягає у використанні реконструйованого автоенкодером сигналу для подальшої автентифікації замість застосування коригування до вхідного сигналу.

2. Застосування кожного із описаних вище методів робить систему біометричної автентифікації стійкішою до аномальних вибірок і, відповідно, точнішою. Використання методів для виявлення та виправлення артефактів підвищує точність біометричної системи на понад 7%.

3. Розроблено метод темпоральної нормалізації серцевого ритму, який здійснює часову трансформацію ЕКГ-сигналу з приведення тривалості серцевого циклу до наперед встановленого значення, а його застосування забезпечує стійкість нейромережевого автентифікатора до перенавчання та підвищує його точність розпізнавання на 8%.

4. Серед сучасних методів класифікації (метод опорних векторів, лінійний дискримінантний аналіз, k-найближчих сусідів, дерева рішень, нейронні мережі, тощо) здійснено вибір оптимального для побудови системи автентифікації. Досліджено придатність біометричної системи автентифікації до масштабування, а саме визначено вплив збільшення числа користувачів на точність автентифікації.

РОЗДІЛ 4. ЕКСПЕРИМЕНТАЛЬНІ ДОСЛІДЖЕННЯ ЕФЕКТИВНОСТІ РОЗРОБЛЕНИХ МЕТОДІВ ТА ЇХ ПРАКТИЧНА РЕАЛІЗАЦІЯ НА СУЧАСНИХ ОБЧИСЛЮВАЛЬНИХ ПЛАТФОРМАХ

4.1. Імплементація біометричної системи автентифікації

Для демонстрації можливості реалізації біометричної системи автентифікації на основі електрокардіограми було вирішено здійснити її імплементацію на базі мікрокомп'ютера Raspberry Pi 3B.

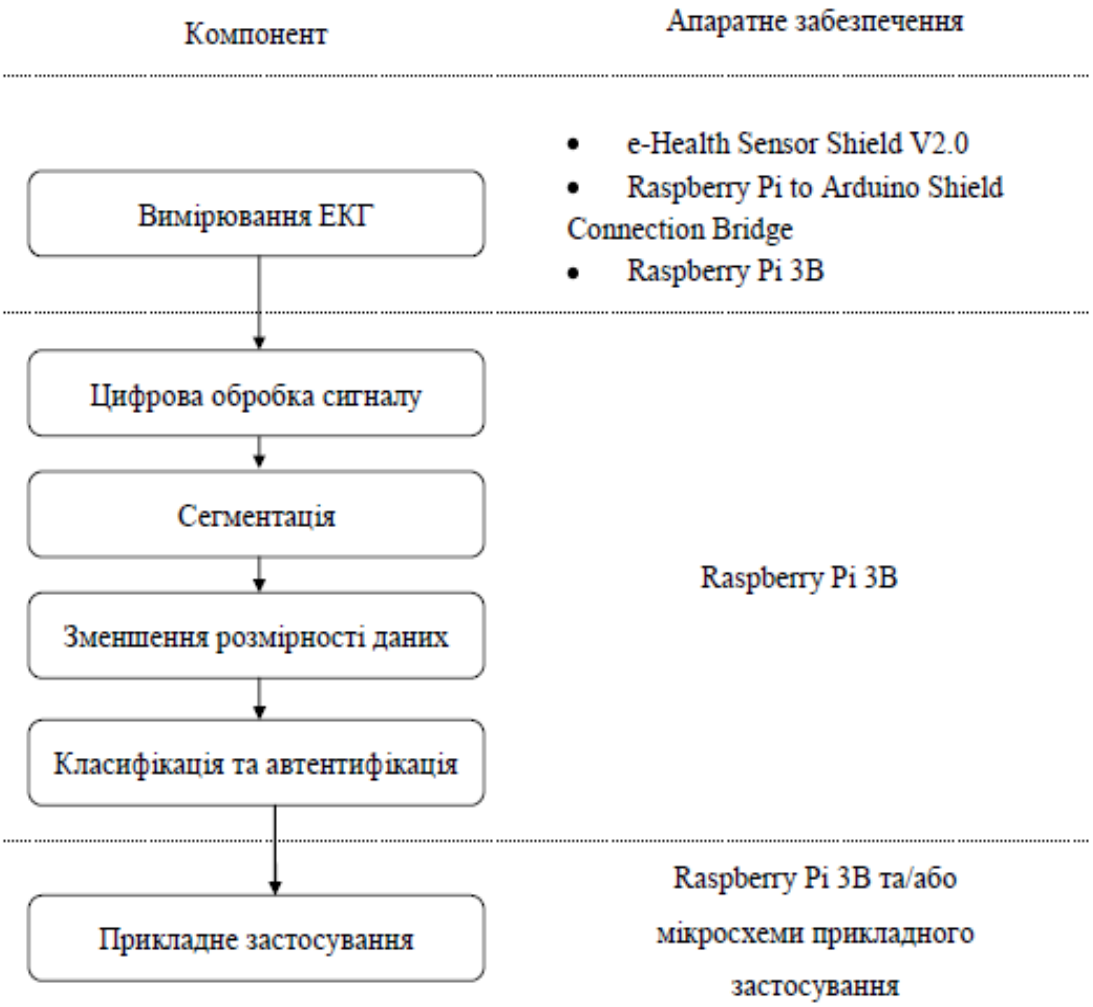


Рис. 4.1. Апаратне забезпечення для імплементації біометричної системи автентифікації

На рис. 4.1 наведено апаратне забезпечення необхідне для кожної із компонент структурної схеми біометричної автентифікації описаної у підрозділі

2.3. Отже, для імплементації біометричної системи автентифікації на основі електрокардіограми необхідно наступні апаратні компоненти:

- мікросхема e-Health Sensor Shield V2.0 для вимірювання ЕКГ;
- мікросхема Raspberry Pi to Arduino Shields Connection Bridge для під'єднання мікросхеми e-Health Sensor Shield V2.0 до Raspberry Pi 3B.
- мікрокомп’ютер Raspberry Pi 3B, на базі якого буде здійснюватися обробка ЕКГ та власне автентифікація.

Мікросхема e-Health Sensor Shield V2.0 (рис. 4.2) призначена для вимірювання біомедичних сигналів. Дана мікросхема детально описана у пункті 2.1.1. Разом із мікросхемою Arduino використовувалась для вимірювання ЕКГ-сигналів набору Lviv Biometric Dataset.

Рис. 4.2. Мікросхема e-Health Sensor Shield V2.0

Мікросхема Raspberry Pi to Arduino Shields Connection Bridge дозволяє приєднати до Raspberry Pi будь-які мікросхеми, плати та модулі розроблені для Arduino. Дана мікросхема також дозволяє підключення цифрових та аналогових датчиків до Raspberry Pi, використовуючи інтерфейси Arduino.

Для повної сумісності існує бібліотека arduPi [92], яка дозволяє використовувати компоненти на Raspberry з тим самим кодом, що і на Arduino.

Дана мікросхема разом із бібліотекою arduPi дозволяє:

● підключати до мікрокомп'ютера Raspberry будь-який бездротовий модуль Arduino (XBee 802.15.4/XBee ZigBee, RFID, NFC, Bluetooth, Bluetooth Pro, Wifi, GPRS, 3G)

● підключати будь-який датчик (аналоговий 0-5 В, цифровий) до Raspberry Pi за допомогою вбудованого АЦП.

● підключати будь-яку Arduino сумісну компоненту до мікрокомп'ютера Raspberry.

● підключати будь-який електронний модуль, що працює через i2C, SPI, UART [93]

Рис. 4.3. Мікросхема Raspberry Pi to Arduino Shields Connection Bridge

Raspberry Pi - це серія невеликих одноплатних комп’ютерів, розроблених у Великобританії фондом Raspberry Pi Foundation. На початку проект Raspberry Pi призначався для вивчення у школах базових комп'ютерних навичок. Пізніше оригінальна модель стала набагато популярнішою, ніж передбачалося, знайшовши широке використання у робототехніці. Зараз мікрокомп'ютер широко використовується у багатьох областях, через низьку вартість та високу портативність. [94]

Для імплементації біометричної системи автентифікації було використано модель Raspberry Pi 3B (рис. 4.4), технічні характеристики якої наведені у таблиці 4.1

Рис. 4.4. Мікрокомп'ютер Raspberry Pi 3B

Таблиця 4.1.

Характеристики мікрокомп'ютера Raspberry Pi 3B

Процесор	Broadcom BCM2837B0 quad-core A53 (ARMv8) 64-bit @ 1.4GHz
Графічний процесор	Broadcom Videocore-IV
RAM	1GB LPDDR2 SDRAM
Мережа	Gigabit Ethernet, 2.4GHz and 5GHz 802.11b/g/n/ac Wi-Fi
Bluetooth	Bluetooth 4.2, Bluetooth Low Energy (BLE)
Сховище даних	Micro-SD
Інтерфейси входу-виходу	40 пінів
Порти	HDMI, 3.5mm analogue audio-video jack, 4x USB 2.0, Ethernet, Camera Serial Interface (CSI), Display Serial Interface (DSI)
Розміри	82мм x 56мм x 19.5мм, 50г

Для Raspberry Pi існує величезна кількість різноманітних дистрибутивів, як офіційних, що підтримуються Raspberry Pi Foundation, так і не офіційних (розроблених сторонніми компаніями чи спільнотами користувачів).

Для побудови біометричної системи автентифікації було використано дистрибутив Raspberry Pi OS. Raspberry Pi OS (попередня назва Raspbian) – це операційна система для Raspberry Pi на основі Debian. З 2015 року вона офіційно надається фондом Raspberry Pi Foundation як основна операційна система для сімейства компактних одноплатних комп'ютерів Raspberry Pi. Оригінальна ОС Raspbian була створена Майком Томпсоном та Пітером Гріном як самостійний проект. Початкова збірка була завершена в червні 2012 року. Операційна система знаходиться на стадії активної розробки. Raspbian оптимізована для низькопродуктивних процесорів ARM, що використовуються в лінійці комп'ютерів Raspberry Pi. [95]

Raspberry Pi OS є безкоштовною операційною системою і постачається в наступних комплектаціях:

● Raspberry Pi desktop and recommended software - версія з графічною оболонкою PIXEL та повним набором програмного забезпечення (BlueJ, Geany, Greenfoot, Mathematica, mu, Node-RED, Scratch, Sense HAT Emulator, Sonic Pi, Thonny Python IDE, Wolfram, SmartSim, LibreOffice, Chromium, Claws, VNC Viewer, VLC, Minecraft Pi).

● Raspberry Pi with desktop - версія з графічною оболонкою PIXEL та мінімальним набором програмного забезпечення (Geany, Chromium, VLC, Image Viewer, Calculator, PDF Viewer, SD Card Copier).

● Raspberry Pi Lite - версія без графічної оболонки.

Raspberry Pi OS є оптимальною операційною системою для побудови біометричної системи автентифікації, оскільки вона дозволяє встановити та виконати Python застосунки, які використовувались для експериментів у попередніх розділах.

На рис. 4.5 наведено апаратне забезпечення використане для побудови біометричної системи автентифікації на основі ЕКГ. Варто відзначити невеликі габарити імплементованої системи. Вона може поміститись у корпус розмірами 85мм х 60мм х 60мм.

Також до Raspberry Pi через відповідні інтерфейси можливо підключити сенсорний дисплей, який дозволить спростити комунікацію з біометричною системою автентифікації.

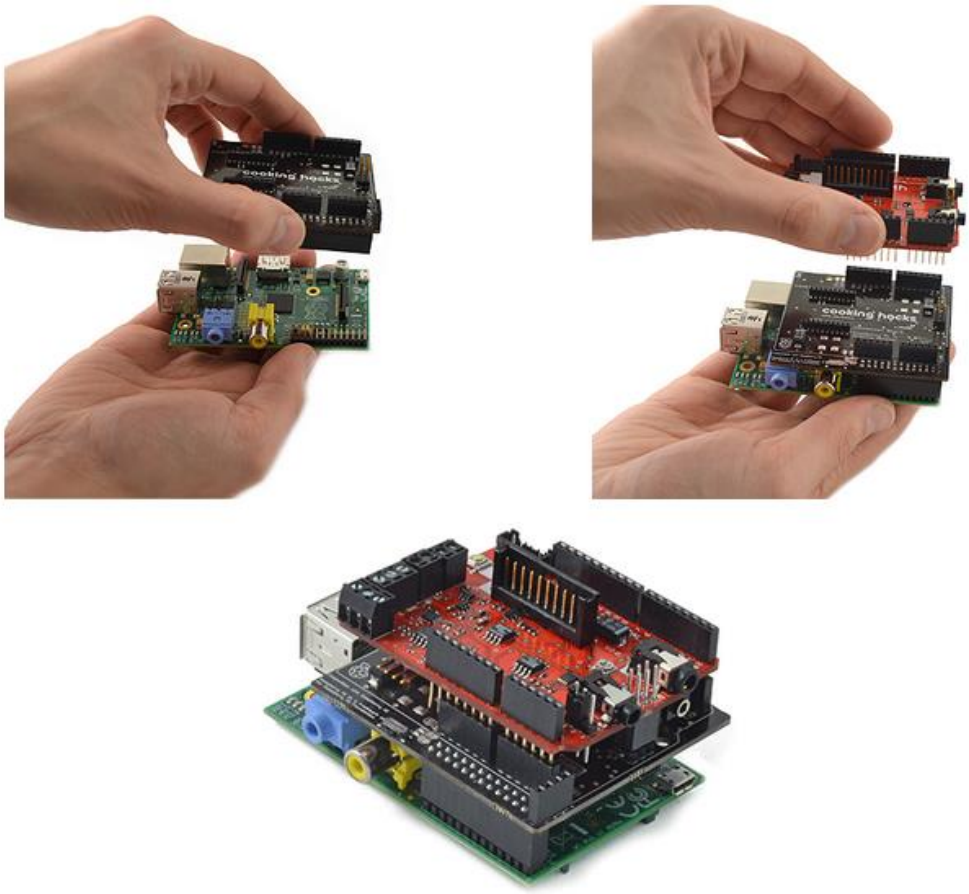


Рис. 4.5. Апаратне забезпечення біометричної системи автентифікації

При успішно пройденій автентифікації біометрична система повинна надати доступ до інформації яка міститься на мікрокомп'ютері Raspberry Pi або за допомогою апаратури прикладного застосування надати доступ до інформації за межами мікрокомп'ютера Raspberry Pi. Наприклад, це може бути система контролю доступу, яка при успішній автентифікації надаватиме доступ до приміщення.

4.2. Дослідження швидкодії імплементованої біометричної системи

Для дослідження швидкодії біометричної систем автентифікації було імплементовано систему на основі мікрокомп'ютера Raspberry Pi, яка детально описана у попередньому підрозділі. Також для співставлення було імплементовано біометричну систему на основі робочої станції з наступними параметрами: процесор Intel® Core™ i7-4790 CPU @ 3.60GHz × 8, оперативна пам'ять 32 Гб.

Таблиця 4.2.

Часові затрати кожної компоненти із структурної схеми біометричної системи імплементованої на основі персональної робочої станції.

Назва компонента		**Середнє значення, мс**	**Стандартне відхилення, мс**
Вимірювання ЕКГ		10000.00	1.72
Фільтрація		1.48	0.64
Амплітудна нормалізація		0.23	0.04
Сегментація		7.58	0.71
Темпоральна нормалізація		0.83	0.16
Виявлення та виправлення артефактів	Статистичний метод	10.25	2.04
	Автоенкодер із коригуванням віконним методом.	365.65	98.78
	Реконструйований сигнал автоенкодера	324.20	93.66
Зменшення розмірності даних		0.07	0.02
Класифікація	Метод опорних векторів	8.69	1.64
	Метод К найближчих сусідів	0.51	0.06
	Дерева рішень	0.07	0.01
	Лінійний дискримінантний аналіз	0.09	0.04
	Нейронні мережі	0.38	0.16

Обидві біометричні системи імплементовано на базі структурної схеми описаної в підрозділі 2.3. Вибір оптимальних гіперпараметрів для кожного із компонентів наведено в розділі 3. Продуктивність імплементованих біометричних систем в площині точності є однаковою та відповідає значенням наведеними у підрозділі 3.3.

Метою даного дослідження є оцінити продуктивність імплементованих біометричних систем в площині часу та обрати оптимальну конфігурацію структурної схеми для кожної з імплементацій.

Таблиця 4.3.

Часові затрати кожної компоненти із структурної схеми біометричної системи імплементованої на основі мікрокомп’ютера Raspberry Pi.

Назва компонента		**Середнє значення, мс**	**Стандартне відхилення, мс**
Вимірювання ЕКГ		10000.00	3.68
Фільтрація		17.19	5.14
Амплітудна нормалізація		2.09	0.67
Сегментація		101.84	28.81
Темпоральна нормалізація		10.77	3.75
Виявлення та виправлення артефактів	Статистичний метод	149.50	53.48
	Автоенкодер із коригуванням віконним методом.	13287.74	757.76
	Реконструйований сигнал автоенкодера	12345.99	725.65
Зменшення розмірності даних		1.39	0.32
Класифікація	Метод опорних векторів	219.98	42.27
	Метод К найближчих сусідів	11.44	3.53
	Дерева рішень	0.95	0.19
	Лінійний дискримінантний аналіз	1.43	0.30
	Нейронні мережі	13.32	3.65

У таблиці 4.2 наведено часові затрати для кожної компоненти із структурної схеми біометричної системи автентифікації імплементованої на основі персональної робочої станції. Як видно з таблиці найбільш затратним є вимірювання ЕКГ-сигналу яке в середньому триває 10с. Вимірювання ЕКГ-сигналу є обов'язковим та не може бути пришвидшеним за допомогою потужнішого обладнання. Наступним часозатратним компонентом є компонент виявлення та виправлення артефактів, який, залежно від конфігурації, витрачає від 10 до 400 мс на обробку артефактів. Однак це на порядок менше ніж затрати на вимірювання ЕКГ-сигналу.

У таблиці 4.3 наведено часові затрати для кожної компоненти із структурної схеми біометричної системи автентифікації імплементованої на основі мікрокомп'ютера Raspberry Pi. Тут кожен із компонентів структурної схеми, які здійснюють обробку виміряного ЕКГ-сигналу, витрачають на порядок більше часу ніж компоненти системи імплементованої на основі робочої станції. А компонент виявлення та коригування артефактів побудований на основі реконструйованого сигналу автоенкодера витрачає на обробку артефактів майже 13 с. Якщо враховувати час витрачений на вимірювання ЕКГ, це робить імплементацію на основі мікрокомп'ютера Raspberry Pi у двічі повільнішою, ніж імплементація на основі персональної робочої станції.

Згідно з пунктом 3.1.4 для імплементації на основі мікрокомп'ютера Raspberry Pi можна використати, не втративши в точності, компонент виявлення та виправлення артефактів на основі статистичного методу. Це значно пришвидшить час обробки артефактів у ЕКГ-сигналах.

Згідно з таблиць 4.2 та 4.3 час витрачений на класифікацію є прийнятним для комфортного використання імплементованих біометричних систем. Проте не менш важливим аргументом при виборі оптимального підходу для побудови класифікатора біометричної системи автентифікації є час його навчання. У таблиці 4.4 наведено інформацію про тривалість навчання кожного з методів класифікації. Відповідно до даної таблиці, навчання класифікатора на

персональній робочій станції здійснюється за менш ніж одну хвилину, це достатньо швидко та є комфортним для адміністрування біометричної системи. Навчання класифікатора на основі нейронних мереж на мікрокомп'ютері Raspberry Pi тривало приблизно 45 хвилин, що не є прийнятним. Тому для імплементації біометричної системи на основі мікрокомп'ютера Raspberry Pi рекомендується використовувати класифікацію на основі методу опорних векторів, оскільки він працює з прийнятною швидкістю та, відповідно до таблиці 3.15, забезпечує достатньо високу точність автентифікації

Таблиця 4.4.

Тривалість навчання класифікатора біометричної системи автентифікації

Платформа	**Персональна робоча станція**		**Raspberry Pi**	
Тривалість навчання	**Середнє значення, с**	**Стандартне відхилення, с**	**Середнє значення, с**	**Стандартне відхилення, с**
Метод опорних векторів	1.23	0.01	24.06	0.15
Метод К найближчих сусідів	0.06	0.01	1.20	0.04
Дерева рішень	1.66	0.01	182.40	0.17
Лінійний дискримінантний аналіз	0.12	0.02	2.42	0.06
Нейронні мережі	53.95	0.20	2644.60	88.41

Підсумовуючи вище написане, оптимальна структурна схема для імплементації біометричної системи автентифікації на основі персональної робочої станції містить компонент виявлення та виправлення артефактів на основі реконструйованого сигналу автоенкодера і компонент класифікації на

основі нейронних мереж. Середній час автентифікації одного суб'єкта - 10.33 с (таблиця 4.5).

Оптимальна структурна схема для імплементації біометричної системи автентифікації на основі мікрокомп'ютера Raspberry Pi містить компонент виявлення та виправлення артефактів на основі статистичного методу і компонент класифікації на основі методу опорних векторів. Середній час автентифікації одного суб'єкта - 10.5 с.

Таблиця 4.5.

Тривалість здійснення автентифікації

Платформа	**Персональна робоча станція**		**Raspberry Pi**	
Конфігурація структурної схеми	**Середнє значення, с**	**Стандартне відхилення, с**	**Середнє значення, с**	**Стандартне відхилення, с**
виявлення та виправлення артефактів на основі статистичного методу; класифікація на основі методу опорних векторів	10.03	0.01	10.50	0.11
виявлення та виправлення артефактів на основі реконструйованого сигналу автоенкодера; класифікація на основі нейронних мереж	10.33	0.09	22.55	0.76

Дані оптимальних конфігурації структурних схем забезпечують певний компроміс між тривалістю та точністю автентифікації в залежності від обчислювальних потужностей платформи на якій буде побудована біометрична система автентифікації.

Також варто відзначити, що тривалість автентифікації може бути зменшена за рахунок зменшення тривалості вимірювання ЕКГ-сигналу. Проте зменшення тривалості виміряного сигналу може негативно вплинути на якість виявлення та виправлення артефактів та точність автентифікації. Вибір оптимальної тривалості вимірювання ЕКГ-сигналу виходить за межі дисертаційного дослідження, та буде здійснено у майбутніх дослідженнях автора.

4.3. Дослідження часової інваріантності ЕКГ-сигналу

На шляху до практичного застосування систем біометричної автентифікації на основі ЕКГ важливо дослідити стабільність основних дискримінаційних ознак на довгих проміжках часу (місяці-роки), за якими класифікатор приймає рішення про надання чи відмову у доступі. Аналіз літературних джерел показав на відсутність публікацій із результатами подібних досліджень.

Метою даного підрозділу є дослідження часової стабільності ЕКГ-сигналів на довготривалих проміжках часу, а також оцінювання ступеня впливу варіативності інформативних ознак електрокардіограми на точність автентифікації [96].

Для даного дослідження було імплементовано біометричну систему автентифікації, яку детально описано у другому розділі. У дану систему включено компонент виявлення та виправлення артефактів на основі статистичного методу, а також компонент, який здійснює темпоральну нормалізацію [96]. Для зменшення розмірності даних використано метод головних компонент, який зменшує розмірність кожного з сегментів до 30 вибірок. Відповідно до підрозділу 3.3, класифікатор побудовано на основі штучних нейронних мереж.

Для оцінювання ефективності автентифікації було використано наступні метрики: точність (Accuracy), помилка першого роду (FRR) та помилка другого роду (FAR), а також для детальнішої інтерпретації результатів досліджень використано матрицю помилок (confusion matrix).

У дослідженнях використовувалася база даних електрокардіограм Lviv Biometric Data Set (LBDS), яка детально описана у підрозділі 2.5. Даний набір даних дозволяє здійснити дослідження часової стабільності ЕКГ-сигналів у двох конфігураціях експериментів:

1. Перевірка стабільності ЕКГ-сигналів протягом одного - двох місяців.
2. Перевірка стабільності ЕКГ-сигналів протягом двох років.

Розпочнемо із першого експерименту. Для цього з бази даних електрокардіограм було виокремлено підмножину ЕКГ-записів, інтервал між якими становив один-два місяці. Отриманий набір даних містить записи ЕКГ 20-ти осіб. Його було розділено на навчальний та тестовий набори так, щоб проміжок між останнім вимірюванням навчального набору та першим вимірюванням тестового становив як мінімум один місяць. Для кожної особи у табл. 4.6 наведено дані про число записів у навчальному і тестовому наборах.

Результати першого експерименту наведено у табл. 4.7. Для демонстрації ефекту темпоральної нормалізації експерименти проведено в наступній конфігурації:

− без застосування алгоритму темпоральної нормалізації. У навчальному та тестовому наборі збережено оригінальне значення серцевого ритму;

− з застосування алгоритму темпоральної нормалізації. Значення серцевого ритму усіх ЕКГ-записів навчального та тестового наборів приведено до значення, еквівалентного 90 ударам серця за хвилину.

Таблиця 4.6.

Опис набору даних першого експерименту для перевірки стабільності ЕКГ-сигналів протягом одного - двох місяців.

Назва суб'єкта	**Навчальний набір**		**Тестовий набір**		**Інтервал, дні**
	Кількість вимірювань	**Дата останнього вимірювання**	**Кількість вимірювань**	**Дата першого вимірювання**	
user4	11	2017-12-01	11	2018-02-12	73
user9	11	2017-10-25	11	2017-11-22	28
user13	10	2017-12-01	10	2018-02-12	73
user14	11	2017-12-01	11	2018-02-12	73
user15	11	2017-12-01	2	2018-02-12	73
user54	17	2017-12-01	11	2018-02-12	73
user55	11	2017-12-19	11	2018-02-01	44
user73	9	2017-10-25	10	2017-11-22	28
user74	8	2017-10-25	17	2017-11-22	28
user75	9	2017-10-25	11	2017-11-22	28
user76	11	2017-10-25	11	2017-11-22	28
user78	11	2017-10-25	11	2017-11-22	28
user80	21	2017-10-25	22	2017-11-22	28
user81	11	2017-10-25	11	2017-11-22	28
user82	11	2017-10-25	10	2017-11-22	28
user84	11	2017-10-25	5	2018-01-24	91
user85	11	2017-11-22	7	2018-01-24	63
user86	11	2017-11-22	6	2018-01-24	63
user87	17	2017-11-22	8	2018-01-24	63
user88	11	2017-11-22	9	2018-01-24	63

У багатьох дослідженнях здійснюють випадковий поділ на навчальний і тренувальний набори даних. Це зумовлено тим, що зазвичай набори даних, які використовуються у цих роботах, містять невеликий обсяг вимірювань, виконаних впродовж короткого проміжку часу. Набір даних, який використовується у даному дослідженні, дає можливість розділити тренувальний і тестовий набори за часом

вимірювання (це повинно репрезентувати використання біометричної системи в реальних умовах повсякденного використання). Тому в даному дослідженні також представлено порівняння цих двох варіантів поділу на навчальний і тестовий набори даних: випадковий та за часом вимірювання.

Таблиця 4.7.

Результати експерименту для перевірки стабільності ЕКГ-сигналів протягом одного - двох місяців.

	Без темпоральної нормалізації		Із темпоральною нормалізацією	
Конфігурація навчального та тестового наборів ЕКГ	Розподілені випадковим чином	Відповідно до табл. 4.6	Розподілені випадковим чином	Відповідно до табл. 4.6
Розмір тест сету	205			
Кількість помилок	3	24	2	9
Точність, %	98.54	88.29	99.02	**95.61**
FAR	0.0009	0.0106	0.0007	**0.0031**
FRR	0.011	0.0399	0.0092	**0.0376**

Вища точність автентифікації за випадкового розподілу пояснюється тим, що як у навчальному, так і тестовому наборі можуть міститись ЕКГ-записи виміряні протягом одного вимірювального сеансу. Результати у варіанті часового поділу наборів ЕКГ-записів, тобто коли тестовий набір виміряно на один-два місяці пізніше, ніж навчальний, також є достатньо високими за умови проведення темпоральної нормалізації серцевого ритму (див. останню колонку табл. 4.7). Для згаданої конфігурації системи автентифікації на рис. 4.6 наведено матрицю помилок, де головна діагональ матриці репрезентує випадки вірно автентифікованих користувачів за їх ЕКГ-сигналами. Як видно за нормалізації серцевого ритму, мало місце 9 помилок на 205 ЕКГ-записах (проти 24 без такої

нормалізації), а також відповідно є кращими результати за метриками точності, FAR і FRR.

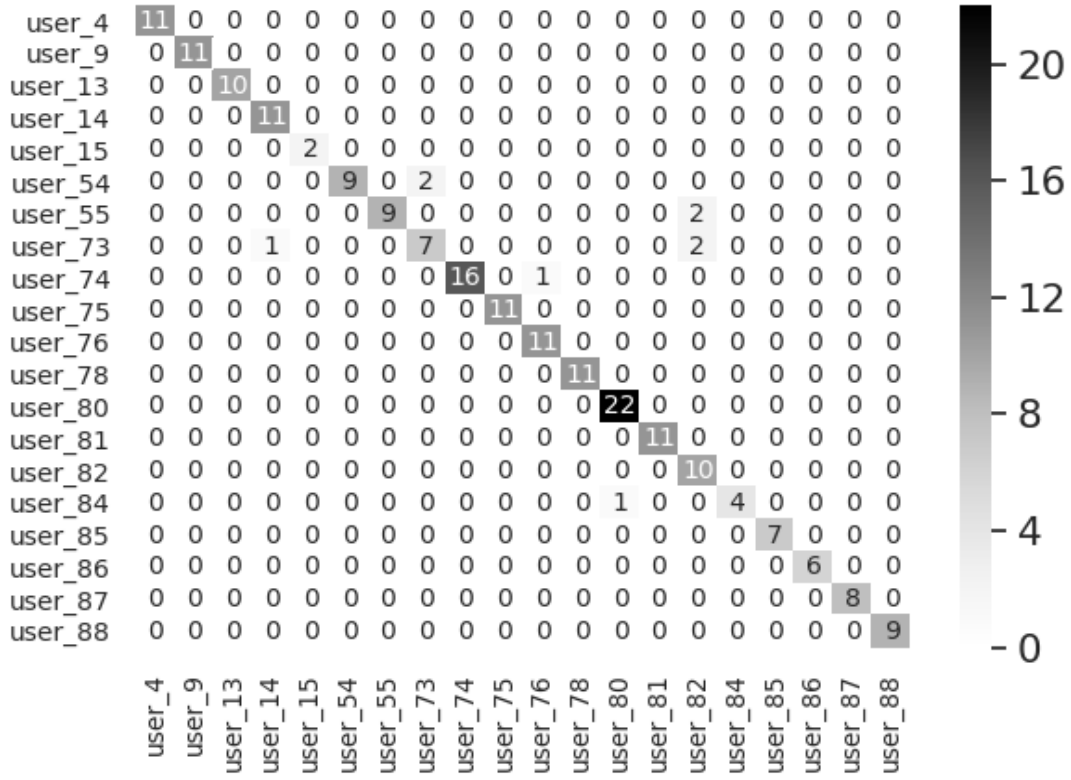


Рис. 4.6. Матриця помилок першого експерименту для конфігурації наборів ЕКГ із табл. 4.6

Варто також зазначити, що система здатна розпізнати кожного користувача (відсутні нульові елементи на головній діагоналі матриці помилок). Це свідчить про те, що імплементована система сконфігурована належним чином та може достовірно автентифікувати її користувачів.

Метою другого експерименту було – дослідити стійкість електрокардіограми як біометричного маркера на тривалішому часовому горизонті. Для цього з бази електрокардіограм було виокремлено підмножину записів, інтервал між якими становив приблизно два роки. Отриманий набір даних містить ЕКГ-записи 20-ти осіб (як і у першому експерименті). Його було розділено на навчальний і тестовий набори так, щоб проміжок між останнім вимірюванням навчального набору та першим вимірюванням тестового становив два роки. У табл. 4.8 наведено відомості про число ЕКГ-записів у навчальному і тестовому наборах даних для

кожної особи. Зазначимо, що користувачі у першому і другому експериментах не збігаються на 100% через відсутність у базі даних їх ЕКГ-записів за обидвома періодами.

Таблиця 4.8.

Опис набору даних другого експерименту для перевірки стабільності ЕКГ-сигналів протягом двох років.

Назва суб'єкта	**Навчальний набір**		**Тестовий набір**		**Інтервал, дні**
	Кількість вимірювань	**Дата останнього вимірювання**	**Кількість вимірювань**	**Дата першого вимірювання**	
user4	32	2018-02-12	15	2019-12-27	683
user9	35	2017-11-22	15	2019-12-20	758
user13	34	2018-02-12	15	2019-11-01	627
user14	33	2018-02-12	11	2019-12-27	683
user15	17	2018-02-12	6	2019-11-08	634
user17	21	2018-02-12	10	2019-12-09	665
user73	25	2018-01-24	9	2019-12-20	695
user74	31	2018-01-24	7	2019-12-20	695
user75	20	2017-11-22	10	2019-12-20	758
user78	26	2018-01-24	12	2019-12-20	695
user80	50	2018-01-24	6	2019-12-20	695
user81	30	2018-01-24	9	2019-12-20	695
user84	16	2018-01-24	7	2019-12-20	695
user85	18	2018-01-24	10	2019-12-20	695
user86	17	2018-01-24	10	2019-12-20	695
user87	25	2018-01-24	10	2019-12-20	695
user88	20	2018-01-24	6	2019-12-20	695
user96	7	2018-01-24	8	2019-12-20	695
user97	10	2018-02-12	9	2019-12-27	683
user99	11	2018-02-12	14	2019-12-27	683

Результати другого експерименту наведено у табл. 4.9. Помітно, що при випадковому розподілі ЕКГ-записів на навчальний та тестовий набори система також показала високу точність автентифікації. Але і результати автентифікації за тестовим набором ЕКГ-сигналів, що виміряні два роки пізніше за навчальний набір, також є достатньо високими. Як видно із рис. 4.7, мало місце 10 помилок на 199 ЕКГ-записах (проти 45 без нормалізації серцевого ритму).

Таблиця 4.9.

Результати експерименту для перевірки стабільності ЕКГ-сигналів протягом двох років

	Без темпоральної нормалізації		Із темпоральною нормалізацією	
Конфігурація навчального та тестового наборів ЕКГ	Розподілені випадковим чином	Відповідно до табл. 4.8	Розподілені випадковим чином	Відповідно до табл. 4.8
Розмір тест сету	199			
Кількість помилок	2	45	0	10
Точність, %	98.99	77.39	100.00	94.97
FAR	0.0006	0.0177	0.0	0.0025
FRR	0.0096	0.0839	0.0	0.0433

Таким чином, на основі результатів двох експериментів можна зробити висновок, що електрокардіограма є стабільним у часі біометричним маркером – точність автентифікації за тестовими записами ЕКГ-сигналів, зробленими через місяць та через два роки є на одному рівні – відповідно 95,61% і 94,97%. Близькими також є значення помилок: FAR - 0.003/0.0025 і FRR - 0.0376/0.0433.

Проведені дослідження засвідчили, що ЕКГ є стабільним маркером для побудови систем автентифікації, на основі якого система здатна адекватно

розпізнавати користувачів упродовж тривалого часу без необхідності проміжних перекалібрувань системи.

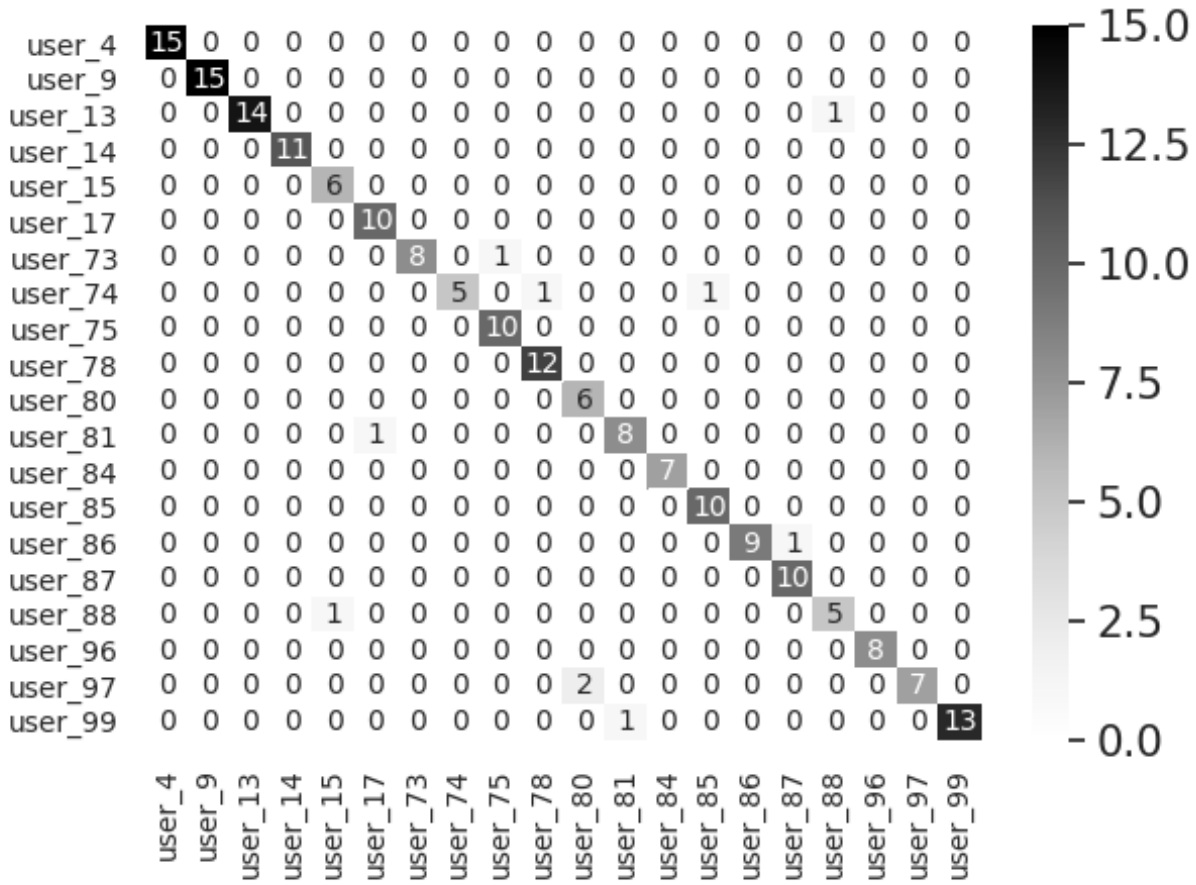


Рис. 4.7. Матриця помилок другого експерименту для конфігурації наборів ЕКГ із табл. 4.8

Крім того, слід відзначити, що застосування алгоритму темпоральної нормалізації ЕКГ-сигналу засвідчило свою ефективність. Його застосування дало змогу підвищити точність автентифікації на 7% (з 88,29% до 95,61%) – на горизонті в один місяць та аж на 17% (з 77,39% до 94,97%)– на горизонті в два роки.

Також дане дослідження продемонструвало, що випадковий розділ даних на тестовий та навчальний набори даних є неефективним, через явище перенавчання класифікатора [97]. Тому для ефективного оцінювання таких систем, необхідно накопичувати якомога більше даних, щоб здійснити розподіл на тренувальний та тестовий набори даних у часовій площині.

4.4. Сфери застосування біометричних систем ЕКГ-автентифікації

Кожна біометрична метрика має свої переваги і недоліки, а її вибір залежить від сфери застосування та обмежень накладених при побудові біометричних систем для певного застосування. Обґрунтованість використання ЕКГ для біометричного розпізнавання підтверджується тим фактом, що фізіологічні і геометричні відмінності серця у різних людей демонструють певну унікальність в їх ЕКГ-сигналах [98].

Результати, отримані в попередніх розділах підтверджують виразність і стабільність ЕКГ в якості біометричної характеристики. У порівнянні з іншими біометричними характеристиками, ЕКГ є більш універсальною та її важко імітувати. Біометричні системи з ЕКГ автентифікацією підходять широкому колу осіб, включаючи людей із обмеженими можливостями (ампутовані кінцівки, порушення зору і т.д.).

Важливою перевагою біометричних систем на основі ЕКГ є низькі обчислювальні затрати. Розпізнавання людини можна здійснювати на малопотужних пристроях та смартфонах за умови наявності вимірювальних електродів. Це усуває необхідність передавання біосигналів на потужніші сервери, що запобігає можливим атакам чи викраденню інформації. Зокрема, в попередніх підрозділах продемонстровано можливість імплементації біометричних систем на мікрокомп'ютери Raspberry Pi 3.

Окрім типового використання в системах контролю доступу, біометричні системи на основі ЕКГ потенційно можуть розпізнавати солдатів на полі бою, ідентифіковувати водія автомобіля, перевіряти пасажирів в аеропортах, автентифіковувати пацієнтів над яким здійснюється віддалений моніторинг, перевіряти особистість в фінансових установах, автентифіковувати власників смартфонів чи захищених носіїв, тощо (рис.4.8).

Автомобільний сектор може отримувати вигоду від цього дослідження, оскільки біометричні системи на основі ЕКГ зможуть ідентифіковувати водія автомобіля, використовуючи кермо чи сидіння для знімання ЕКГ-сигналу, ще до

того як запрацює двигун. Крім того, система може моніторити та, в разі необхідності, інформувати водія про стан серцевої діяльності, щоб уникнути нещасних випадків зі смертельними наслідками через фізіологічні чинники, такі як нестача сну, нерегулярні серцеві ритми чи інші захворювання серця. Система може знерухомити транспортний засіб з метою безпеки. Також такі біометричні системи можуть використовувати транспортні компанії чи служби таксі для того щоб автентифіковувати та моніторити робочий час водіїв.

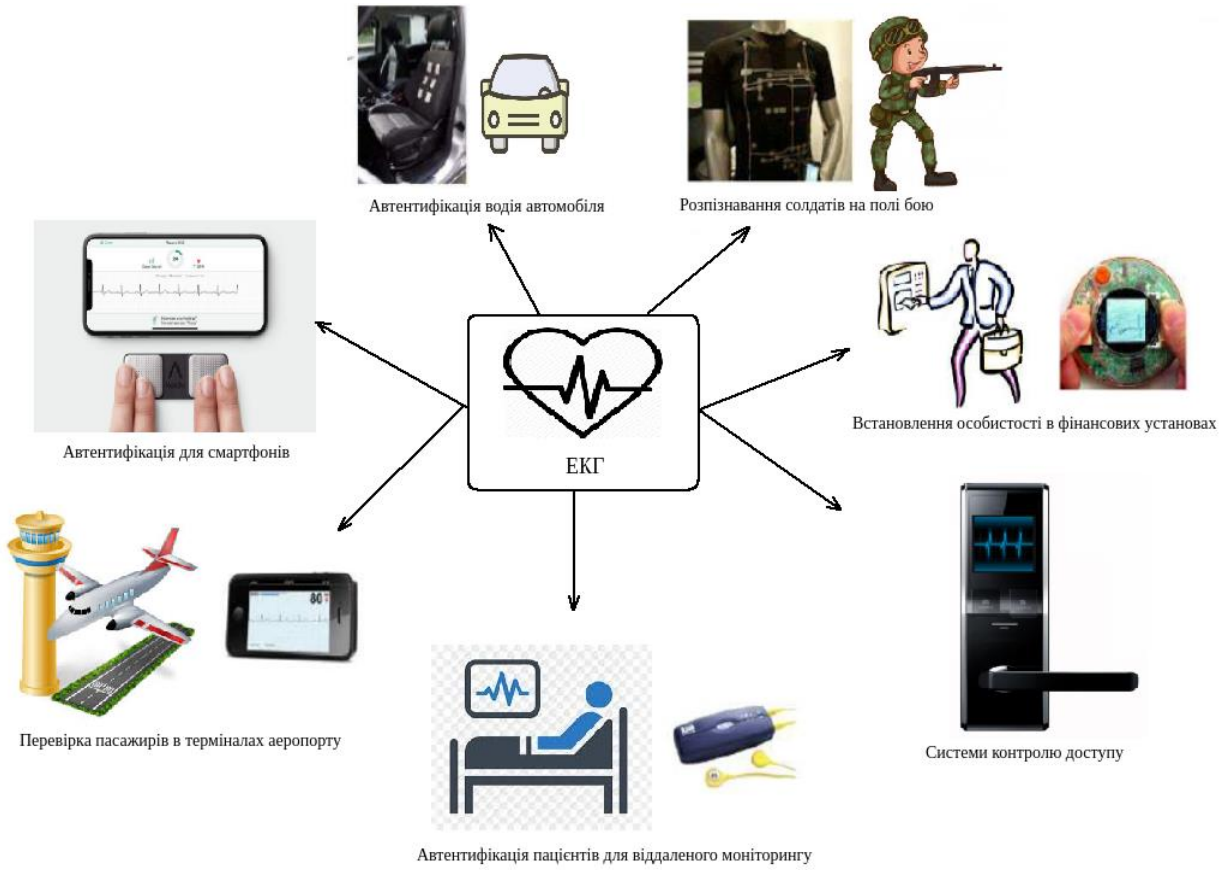


Рис. 4.8. Потенційні застосування біометричної системи автентифікації на основі ЕКГ

ЕКГ-автентифікація в системах віддаленого моніторингу може автентифіковувати пацієнта перед відправленням даних на сервер лікарні для встановлення діагнозу. Ця процедура дозволить запобігти атакам на відмову в обслуговуванні та усуває загрози модифікації персональних чи медичних даних пацієнта.

При використанні біометричної автентифікації на основі ЕКГ в терміналах аеропортів або фінансових установ, можливо порівнювати ЕКГ людини з ЕКГ-сигналами, які записані в його документах, що засвідчують особу чи перебувають у володінні аеропорту або банківських установ. Навіть, якщо людина має інвалідність її можливо автентифікувати по ЕКГ. На основі результатів біометричної автентифікації, влада може скасувати дозвіл в'їзду в країну або банк може відмовити у виконанні фінансових операцій.

Біометричну автентифікацію на основі ЕКГ можливо використовувати в смартфонах, як альтернативу існуючим біометричним маркерам (відбиткам пальців, зображенню лиця, тощо). В попередніх підрозділах продемонстровано, що цілком можливо імплементувати біометричну систему автентифікації на малопотужних пристроях. Єдиним обмеженням для сучасних смартфонів є необхідність вимірювальної апаратури для ЕКГ. На даний час існують зовнішні пристрої для вимірювання ЕКГ (можуть бути у вигляді браслетів, кілець, тощо), які передають зашифрований виміряний сигнал через безпровідний канал зв'язку на смартфон. В майбутньому вимірювальні електроди можуть бути вбудованими у корпус смартфона.

Оборонний комплекс матиме можливість ідентифікувати солдатів на полі бою за допомогою сенсорів ЕКГ вмонтованих в їхню форму. Це дозволить закрити лазівки для ворогів, які маскуються або видають себе за солдатів з протилежного боку, через їх однаковий зовнішній вигляд чи викрадену форму. Крім того можна безперервно контролювати стан серця військового персоналу, та в разі необхідності надавати медичну допомогу.

Додатковою перевагою електрокардіограми як біометричні метрики є можливість використовувати її в діагностичних цілях. Разом із виконанням автентифікації особи, біометрична система при потребі може здійснювати діагностику чи тривалий моніторинг серцевих захворювань, може виявляти їх на ранніх стадіях. За допомогою ЕКГ можливо діагностувати різного роду аритмії, хворобу “збільшеного серця” (кардіомегалію), інфаркт міокарда, кардіоміопатію,

серцеву недостатність, тощо. [99, 100]. Зокрема, якщо жирові відкладення звузили одну або кілька коронарних артерій настільки, що перешкоджають припливу крові до серця, ЕКГ зможе це виявити [101]. При виявленні ознак описаних вище захворювань біометрична системи рекомендуватиме людині звернутися до лікаря.

Біометрична система автентифікації вимагає від людини бути присутньою при розпізнаванні. Втрата біометричних характеристик, таких як пальці, руки, голос, райдужна оболонка ока і обличчя ставить під загрозу біометричну функціональність індивідуума. Як правило, інвалідність не впливає на ЕКГ. Це один з аспектів, який довів перевагу і універсальність ЕКГ у порівнянні з іншими біометричними характеристиками. Таким чином, це дослідження відкриває абсолютно нову галузь знань, яка, як очікується, в майбутньому отримає високий попит.

Висновки до розділу 4

1. Продемонстровано можливість імплементації біометричної системи ЕКГ-автентифікації на малопотужних обчислювальних пристроях. Зокрема, здійснено імплементацію біометричної системи на основі мікрокомп’ютера Raspberry Pi 3B. Наведено огляд апаратних компонент необхідних для побудови біометричної системи.

2. Здійснено тестування продуктивності біометричних систем автентифікації імплементованих на основі мікрокомп’ютера Raspberry Pi 3B та персональної робочої станції з наступними параметрами: процесор Intel® Core™ i7-4790 CPU @ 3.60GHz × 8, оперативна пам'ять 32 Гб, операційна система Ubuntu 18.04. На основі аналізу отриманих результатів було обрано оптимальні конфігурації структурних схем для кожної з імплементації. Для біометричної системи автентифікації на основі персональної робочої станції, відповідно до структурної схеми описаної у 2.3, рекомендовано використовувати компонент виявлення та виправлення артефактів на основі реконструйованого сигналу автоенкодера і компонент класифікації на основі нейронних мереж (середній час автентифікації

одного суб'єкта - 10.33 с). Для імплементації на основі мікрокомп'ютера Raspberry Pi, рекомендовано використовувати компонент виявлення та виправлення артефактів на основі статистичного методу і компонент класифікації на основі методу опорних векторів (середній час автентифікації одного суб'єкта - 10.50 с).

3. Крім того, у даному розділі проведено дослідження часової стабільності ЕКГ-сигналів на довготривалих проміжках часу. Дане дослідження підтвердило, що ЕКГ є стабільним біометричним маркером, на основі якого біометрична система здатна адекватно розпізнавати користувачів впродовж років без необхідності проміжних перекалібрувань системи.

4. Наведено перелік можливих сфер застосування та опис прикладних застосувань для біометричних систем автентифікації на основі ЕКГ. Продемонстровано високий потенціал ЕКГ як біометричного маркера для побудови сучасних біометричних систем.

ВИСНОВКИ

У дисертаційній роботі вирішено актуальну науково-прикладну проблему у галузі кібербезпеки - покращення характеристик системи біометричної автентифікації за сигналом електрокардіограми на основі раціонального поєднання технологій цифрового оброблення сигналів і машинного навчання, що підвищує рівень захищеності ресурсів на об'єктах інформаційної діяльності. Висновки дисертації можна сформулювати у вигляді наукових положень, рекомендацій, фундаментальних та прикладних результатів досліджень:

1. Проведено огляд літературних джерел за темою дисертаційної роботи, зокрема, проаналізовано основні режими роботи і подано порівняльну характеристику сучасних систем біометричної автентифікації. Проведено порівняння найпоширеніших біометричних маркерів за допомогою формалізованих критеріїв.

2. Представлено детальний опис електрокардіограми як біометричного маркера в системах розпізнавання, показано його переваги і проблеми на шляху практичного застосування в системах автентифікації.

3. На основі аналізу процесу автентифікації за ЕКГ- сигналом, формалізовано структуру біометричної системи розпізнавання. Розроблено перспективні підходи до покращення технічних і експлуатаційних характеристик біометричної системи ЕКГ-автентифікації.

4. Імплементовано та апробовано методи для виявлення та виправлення артефактів у ЕКГ-записах. Застосування розроблених методів робить систему біометричної автентифікації стійкішою до аномальних вибірок і підвищує точність біометричної системи на понад 7%.

5. Розроблено метод темпоральної нормалізації серцевого ритму, який здійснює часову трансформацію ЕКГ-сигналу з приведення тривалості серцевого циклу до наперед встановленого значення, а його застосування забезпечує

стійкість нейромережевого автентифікатора до перенавчання, та підвищує його точність розпізнавання на 8%.

6. Серед сучасних методів класифікації здійснено вибір оптимального для побудови системи автентифікації. Досліджено придатність біометричної системи автентифікації до масштабування, а саме визначено вплив збільшення числа користувачів на точність автентифікації.

7. Продемонстровано можливість імплементації біометричної системи ЕКГ-автентифікації на малопотужних обчислювальних пристроях. Зокрема, здійснено імплементацію біометричної системи на основі мікрокомп'ютера Raspberry Pi 3B. Наведено огляд апаратних компонент необхідних для побудови біометричної системи.

8. Здійснено тестування продуктивності біометричних систем автентифікації імплементованих на основі мікрокомп'ютера Raspberry Pi 3B та персональної робочої станції. На основі аналізу отриманих результатів було сформовано рекомендації щодо конфігурацій структурних схем для кожної з імплементації.

9. Проведено дослідження часової стабільності ЕКГ-сигналів на довготривалих проміжках часу. Дане дослідження підтвердило, що ЕКГ є стабільним біометричним маркером, на основі якого біометрична система здатна адекватно розпізнавати користувачів впродовж років без необхідності проміжних перекалібрувань системи.

10. Наведено перелік можливих сфер застосування та опис прикладних застосувань для біометричних систем автентифікації на основі ЕКГ. Продемонстровано високий потенціал ЕКГ як біометричного маркера для побудови сучасних біометричних систем.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Jain A.K., Flynn P., Ross A.A. (2008) Handbook of Biometrics. Springer, 556 p. ISBN: 978-0-387-71040-2.
2. Wayman, James L. (2001) “Fundamentals of biometric authentication technologies.” International Journal of Image and Graphics, vol. 01, no. 01, pp. 93–113., doi:10.1142/s0219467801000086.
3. Korte U. et al. (2008) “A cryptographic biometric authentication system based on genetic fingerprints”. Sicherheit, Schutz und Zuverlässigkeit. Beiträge der 4. Jahrestagung des Fachbereichs Sicherheit der Gesellschaft für Informatik e.V. (GI), pp. 263–276.
4. Delac K., Grgic M. (2004) “A survey of biometric recognition methods”. In: Electronics in Marine, Proceedings Elmar 2004. 46th International Symposium, pp. 184–193.
5. Abate, Andrea F., et al. (2007) “2D And 3D Face Recognition: A Survey.” Pattern Recognition Letters, vol. 28, no. 14, pp. 1885–1906., doi:10.1016/j.patrec.2006.12.018.
6. W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld. (2003). Face recognition: A literature survey. ACM Comput. Surv. 35, 4 (December 2003), 399–458. DOI: https://doi.org/10.1145/954339.954342
7. Ross A., Nandakumar K., Jain A.K. (2006) Handbook of multibiometrics, volume 6. Springer Science & Business Media. ISBN: 978-0-387-22296-7.
8. Maltoni, D., Maio, D., Jain, A. K., & Prabhakar, S. (2009). Handbook of Fingerprint Recognition. doi:10.1007/978-1-84882-254-2
9. European Union. Integration of biometric features in passports and travel documents. 2016. [Електронний ресурс]. – Режим доступу: http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:l14154
10. Постанова КМУ від 26 листопада 2014 р. № 669 “Про затвердження Порядку отримання, вилучення з Єдиного державного демографічного реєстру та знищення відцифрованих відбитків пальців рук особи”; [Електронний ресурс]. – Режим доступу: https://zakon.rada.gov.ua/laws/show/669-2014-%D0%BF#Text
11. Jain A., Bolle R., Pankanti S. (2006) Biometrics: personal identification in networked society, volume 479. Springer Science & Business Media, ISBN: 978-0-387-28539-9
12. Unar, J.A., et al. (2014) “A Review of Biometric Technology along with Trends and Prospects.” Pattern Recognition, vol. 47, no. 8, pp. 2673–2688., doi:10.1016/j.patcog.2014.01.016.

13. Guennoun, Mouhcine, et al. (2009) “Continuous Authentication by Electrocardiogram Data.” 2009 IEEE Toronto International Conference Science and Technology for Humanity (TIC-STH), doi:10.1109/tic-sth.2009.5444466.
14. Israel, Steven A., et al. (2005) “ECG to Identify Individuals.” Pattern Recognition, vol. 38, no. 1, pp. 133–142., doi:10.1016/j.patcog.2004.05.014.
15. Labati, Ruggero Donida, et al. (2013) “ECG Biometric Recognition: Permanence Analysis of QRS Signals for 24 Hours Continuous Authentication.” 2013 IEEE International Workshop on Information Forensics and Security (WIFS), doi:10.1109/wifs.2013.6707790.
16. Odinaka, Ikenna, et al. (2012) “ECG Biometric Recognition: A Comparative Analysis.” IEEE Transactions on Information Forensics and Security, vol. 7, no. 6, pp. 1812–1824., doi:10.1109/tifs.2012.2215324.
17. Shen, T.W., et al. (2002) “One-Lead ECG for Identity Verification.” Proceedings of the Second Joint 24th Annual Conference and the Annual Fall Meeting of the Biomedical Engineering Society] [Engineering in Medicine and Biology, doi:10.1109/iembs.2002.1134388.
18. Da Silva, Hugo Placido, et al. (2013) “Finger ECG Signal for User Authentication: Usability and Performance.” 2013 IEEE Sixth International Conference on Biometrics: Theory, Applications and Systems (BTAS), doi:10.1109/btas.2013.6712689.
19. Campbell, J.P. (1997) “Speaker Recognition: a Tutorial.” Proceedings of the IEEE, vol. 85, no. 9, pp. 1437–1462., doi:10.1109/5.628714.
20. Balagani, Kiran S., et al. (2011) “On the Discriminability of Keystroke Feature Vectors Used in Fixed Text Keystroke Authentication.” Pattern Recognition Letters, vol. 32, no. 7, pp. 1070–1080., doi:10.1016/j.patrec.2011.02.014.
21. Monrose, Fabian, and Aviel Rubin. (1997) “Authentication via Keystroke Dynamics.” Proceedings of the 4th ACM Conference on Computer and Communications Security - CCS '97, doi:10.1145/266420.266434.
22. Frank, Mario, et al. (2013) “Touchalytics: On the Applicability of Touchscreen Input as a Behavioral Biometric for Continuous Authentication.” IEEE Transactions on Information Forensics and Security, vol. 8, no. 1, pp. 136–148., doi:10.1109/tifs.2012.2225048.
23. Maple, C., and P. Norrington. (2006) “The Usability and Practicality of Biometric Authentication in the Workplace.” First International Conference on Availability, Reliability and Security (ARES'06), doi:10.1109/ares.2006.133.
24. Nalwa, V.S. (1997) “Automatic on-Line Signature Verification.” Proceedings of the IEEE, vol. 85, no. 2, pp. 215–239., doi:10.1109/5.554220.

25. Ioannidis, Dimosthenis, et al. (2012) “Gait and Anthropometric Profile Biometrics: A Step Forward.” The International Library of Ethics, Law and Technology, pp. 105–127., doi:10.1007/978-94-007-3892-8_5.
26. Agrafioti, Foteini, et al. (2011) “Heart Biometrics: Theory, Methods and Applications.” Biometrics, doi:10.5772/18113.
27. da Silva H, Lourenço A., et al. (2013) ECG Biometrics: Principles and Applications. Proceedings of the International Conference on Bio-Inspired Systems and Signal Processing. doi:10.5220/0004243202150220.
28. Webster J. G. (2009) "Basic Concepts of Medical Instrumentation", in Medical Instrumentation Application and Design. Wiley, 4th ed.
29. Neuman M. R. (2009) "Biopotencial Amplifiers", in Medical Instrumentation – Application and Design. Wiley, 4th ed.
30. Злепко С.М., Павлов С.В., Коваль Л.Г., Тимчик І.С. Основи біомедичного радіоелекторнного апаратободування. навч. посіб.: Вінниця : ВНТУ, 2011.– 132 с. ISBN 978-966-641-426-0
31. Agrafioti, Foteini, and Dimitrios Hatzinakos. (2008) “ECG Based Recognition Using Second Order Statistics.” 6th Annual Communication Networks and Services Research Conference (Cnsr 2008), doi:10.1109/cnsr.2008.38.
32. Frank G. Yanowitz. Characteristics of the Normal ECG. 2016 [Електронний ресурс]. – Режим доступу: http://ecg.utah.edu/lesson/3
33. Anthony Atkielski. SinusRhythmLabels. 2016 [Електронний ресурс]. – Режим доступу:
https://commons.wikimedia.org/w/index.php?title=File:SinusRhythmLabels.svg
&oldid=194734208.
34. Biel, L., et al. (2001) “ECG Analysis: a New Approach in Human Identification.” IEEE Transactions on Instrumentation and Measurement, vol. 50, no. 3, pp. 808–812., doi:10.1109/19.930458.
35. Bugdol, Marcin D., and Andrzej W. Mitas. (2014) “Multimodal Biometric System Combining ECG and Sound Signals.” Pattern Recognition Letters, vol. 38, pp. 107–112., doi:10.1016/j.patrec.2013.11.014.
36. Wang, Yongjin, et al. (2007) “Analysis of Human Electrocardiogram for Biometric Recognition.” EURASIP Journal on Advances in Signal Processing, vol. 2008, no. 1, doi:10.1155/2008/148658.
37. Kaveh, Anthony, and Wayne Chung. (2013) “Temporal and Spectral Features of Single Lead ECG for Human Identification.” 2013 IEEE Workshop on Biometric

Measurements and Systems for Security and Medical Applications, doi:10.1109/bioms.2013.6656143.
38. Tantawi, M. M., et al. (2012) “Fiducial Feature Reduction Analysis for Electrocardiogram (ECG) Based Biometric Recognition.” Journal of Intelligent Information Systems, vol. 40, no. 1, pp. 17–39., doi:10.1007/s10844-012-0214-7.
39. Tantawi, Manal M., et al. (2012) “An Evaluation of the Generalisability and Applicability of the PhysioNet Electrocardiogram (ECG) Repository as Test Cases for ECG-Based Biometrics.” International Journal of Cognitive Biometrics, vol. 1, no. 1, p. 66., doi:10.1504/ijcb.2012.046515.
40. Luz, Eduardo José, et al. (2014) “Evaluating the Use of ECG Signal in Low Frequencies as a Biometry.” Expert Systems with Applications, vol. 41, no. 5, pp. 2309–2315., doi:10.1016/j.eswa.2013.09.028.
41. Ching-Kun Chen, et al. (2014) “A Chaotic Theoretical Approach to ECG-Based Identity Recognition [Application Notes].” IEEE Computational Intelligence Magazine, vol. 9, no. 1, pp. 53–63., doi:10.1109/mci.2013.2291691.
42. Hejazi, Maryamsadat, et al. (2016) “ECG Biometric Authentication Based on Non-Fiducial Approach Using Kernel Methods.” Digital Signal Processing, vol. 52, pp. 72–86., doi:10.1016/j.dsp.2016.02.008.
43. Chiu, Chuang-Chien, et al.(2008) “A Novel Personal Identity Verification Approach Using a Discrete Wavelet Transform of the ECG Signal.” 2008 International Conference on Multimedia and Ubiquitous Engineering (Mue 2008), doi:10.1109/mue.2008.67.
44. Plataniotis, Konstantinos N., et al. (2006) “ECG Biometric Recognition Without Fiducial Detection.” 2006 Biometrics Symposium: Special Session on Research at the Biometric Consortium Conference, doi:10.1109/bcc.2006.4341628.
45. So, H.H., and K.L. Chan. (1997)“Development of Qrs Detection Method for Real-Time Ambulatory Cardiac Monitor.” Proceedings of the 19th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 'Magnificent Milestones and Emerging Opportunities in Medical Engineering' (Cat. No.97CH36136), doi:10.1109/iembs.1997.754529., pp. 289–292.
46. ARDUINO UNO REV3 [Електронний ресурс] // Режим доступу: https://store.arduino.cc/arduino-uno-rev3
47. e-Health Sensor Platform V2.0 for Arduino and Raspberry Pi [Biometric / Medical Applications] [Електронний ресурс] – Режим доступу: https://www.cooking-hacks.com/documentation/tutorials/ehealth-biometric-sensor-platform-arduino-raspberry-pi-medical.html

48. Sebastian Raschka. About feature scaling and normalization and the effect of standardization for machine learning algorithms, 2014. [Електронний ресурс] – Режим доступу: https://sebastianraschka.com/Articles/2014_about_feature_scaling.html
49. Hamilton, P. (20020 “Open Source ECG Analysis.” Computers in Cardiology, doi:10.1109/cic.2002.1166717.
50. Pan, Jiapu, and Willis J. Tompkins. “A Real-Time QRS Detection Algorithm.” IEEE Transactions on Biomedical Engineering, BME-32, no. 3, 1985, pp. 230–236., doi:10.1109/tbme.1985.325532.
51. Hamilton, Patrick S., and Willis J. Tompkins. (1986) “Quantitative Investigation of QRS Detection Rules Using the MIT/BIH Arrhythmia Database.” IEEE Transactions on Biomedical Engineering, BME-33, no. 12, pp. 1157–1165., doi:10.1109/tbme.1986.325695.
52. Christov, Ivaylo I. (2004) “Real Time Electrocardiogram QRS Detection Using Combined Adaptive Threshold.” BioMedical Engineering OnLine, vol. 3, no. 1, doi:10.1186/1475-925x-3-28.
53. Engelse WAH, Zeelenberg C (1979) A single scan algorithm for QRS-detection and feature extraction. Proc IEEE Computers in Cardiology 6: 37–42.
54. Tsai, Flora S. (2011) “Dimensionality Reduction Techniques for Blog Visualization.” Expert Systems with Applications, vol. 38, no. 3, pp. 2766–2773., doi:10.1016/j.eswa.2010.08.067.
55. S. Raschka, “Linear discriminant analysis bit by bit.” Blog, August 2014. [Електронний ресурс] – Режим доступу: https://sebastianraschka.com/Articles/2014_python_lda.html
56. YongchangWang, and Ligu Zhu. (2017) “Research and Implementation of SVD in Machine Learning.” 2017 IEEE/ACIS 16th International Conference on Computer and Information Science (ICIS), 2017, doi:10.1109/icis.2017.7960038.
57. Huang, Xuan, et al. (2019) “A Review on Dimensionality Reduction Techniques.” International Journal of Pattern Recognition and Artificial Intelligence, vol. 33, no. 10, 2019, p. 1950017., doi:10.1142/s0218001419500174.
58. Jenkins, D., Gerred, S. (2011): ECGs by Example, 3rd edn., 238 p. Elsilver, ISBN-13: 9780702042287.
59. Ribeiro Pinto, Joao, et al. (2018) “Evolution, Current Challenges, and Future Possibilities in ECG Biometrics.” IEEE Access, vol. 6, pp. 34746–34776., doi:10.1109/access.2018.2849870.

60. Pelc M., Khoma Y., Khoma V. (2019) “ECG Signal as Robust and Reliable Biometric Marker: Datasets and Algorithms Comparison.” Sensors, vol. 19, no. 10, p. 2350., doi:10.3390/s19102350.
61. Хома В., Хома Ю., Герасименко В., Сабодашко Д. (2017) ЕКГ-ідентифікація з використанням глибинних нейронних мереж // Вісник НУ «Львівська політехніка» – «Автоматика, вимірювання та керування». №880. с. 67-72.
62. Хома Ю., Герасименко В., Сабодашко Д. (2017) 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.
63. Хома Ю., Сабодашко Д. (2016) Біометрична ідентифікація за допомогою електрокардіограми // Захист інформації і безпека інформаційних систем : матеріали V Міжнародної науково-технічної конференції, 2–3 червня 2016 р., Львів. С. 146–147.
64. Ravi S. The study on multi-modal biometrics incorporating face, fingerprint and iris recognition techniques, 2016 [Електронний ресурс] – Режим доступу:: https://shodhganga.inflibnet.ac.in/bitstream/10603/235475/12/12_chapter%203.pdf
65. Оценка классификатора (точность, полнота, F-мера) [Електронний ресурс] – Режим доступу: http://bazhenov.me/blog/2012/07/21/classification-performance-evaluation.html
66. Moody, G.B., and R.G. Mark. (1990) “The MIT-BIH Arrhythmia Database on CD-ROM and Software for Use with It.” Proceedings Computers in Cardiology, .,vol. 17. Sep. 1990, pp. 185–188, doi:10.1109/cic.1990.144205.
67. Nemirko A.P., Lugovaya T.S. (2005) Biometric human identification based on electrocardiogram. Proc. XII-th Russian Conference on Mathematical Methods of Pattern Recognition, Moscow, MAKS Press, pp. 387-390. ISBN 5-317-01445-X.
68. Bousseljot, R., et al. (2009) “Nutzung Der EKG-Signaldatenbank CARDIODAT Der PTB Über Das Internet.” Biomedizinische Technik/Biomedical Engineering, pp. 317–318., doi:10.1515/bmte.1995.40.s1.317.
69. Laguna, P., et al. (1997) “A Database for Evaluation of Algorithms for Measurement of QT and Other Waveform Intervals in the ECG.” Computers in Cardiology ., vol. 24, Sep. 1997,pp. 673–676. doi:10.1109/cic.1997.648140.
70. UofT ECG Database [Електронний ресурс] // Режим доступу: https://www.comm.utoronto.ca/~biometrics/databases.html
71. American Heart Association ECG database. [Електронний ресурс] // Режим доступу: https://www.ecri.org/components/Pages

72. Da Silva, Hugo Plácido, et al. (2014)“Check Your Biosignals Here: A New Dataset for off-the-Person ECG Biometrics.” Computer Methods and Programs in Biomedicine, vol. 113, no. 2, pp. 503–514., doi:10.1016/j.cmpb.2013.11.017.
73. Khoma V., Pelc M., Khoma Y., Sabodashko D. (2018) 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. Vol 720. p. 11-22. Springer, Cham, doi:10.1007/978-3-319-75025-5_2.
74. Хома В.В., Хома Ю.В., Сабодашко Д.В., Хома П.П. (2019) Автоенкодери для опрацювання промахів сигналів ЕКГ у системі біометричної автентифікації // Штучний інтелект. №1-2. С. 108-117.
75. Karpinski M., Khoma V., Dudykevych V., Khoma Y., Sabodashko D. (2018) 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. doi:10.1109/idaacs-sws.2018.8525836.
76. Khoma V., Khoma Y., Sabodashko D., Shereha V. (2019) 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.
77. Хома Ю.В., Хома В.В., Сабодашко Д.В., Юн С., Кочан О.В. (2020) Аналіз ефективності методів коригування промахів у системах біометричної ідентифікації на підставі електрокардіограми // Науковий вісник НЛТУ України. 30(3). с. 99-105. doi:10.36930/40300317.
78. Dertat, A. Applied Deep Learning - Part 3: Autoencoders [Електронний ресурс] // Режим доступу: https://towardsdatascience.com/applied-deep-learning-part-3
79. Su Jun, Szmajda M., Khoma V., Khoma Y., Sabodashko D., Kochan O., Jinfei Wang. (2020) Comparison of methods for correcting outliers in ECG-based biometric identification // Metrology and measurement systems. Vol. 27(3). – p. 387–398. DOI: 10.24425/MMS.2020.132784
80. Дудикевич В.Б., Хома В.В., Чекурін В.Ф., Хома Ю.В., Сабодашко Д.В. (2019) Нормалізація сигналів ЕКГ для застосування в системах біометричної ідентифікації // Вчені записки Таврійського національного університету імені В.І. Вернадського. Серія: Технічні науки. Том 30 (69), ч. 1, № 4, с. 49-56.

81. Sabodashko D., Khoma V., (2019) Normalizacja Temporalna Sygnału Ekg w Systemie Identyfikacji Biometrycznej, Przetwarzanie, transmisja i bezpieczeństwo informacji Tom 2, Bielsko – Biała, p.313-322
82. Wieclaw L., Khoma Y., Falat P., Sabodashko D., Herasymenko V. (2017) 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. doi:10.1109/idaacs.2017.8095063.
83. Exhaustive search over specified parameter values for an estimator. [Електронний ресурс] // Режим доступу: https://scikit-learn.org/stable/modules/generated/ sklearn.model_selection.GridSearchCV.html
84. C-Support Vector Classification. [Електронний ресурс] // Режим доступу: *https://scikit-learn.org/stable/modules/generated/sklearn.svm.SVC.html#sklearn.svm.SVC*
85. Classifier implementing the k-nearest neighbors vote. [Електронний ресурс] // Режим доступу:*https://scikit-learn.org/stable/modules/generated/sklearn.neighbors.KNeighborsClassifier.html#sklearn.neighbors.KNeighborsClassifier*
86. A decision tree classifier. [Електронний ресурс] // Режим доступу: *https://scikit-learn.org/stable/modules/generated/sklearn.tree.DecisionTreeClassifier.html#sklearn.tree.DecisionTreeClassifier*
87. Gini coefficient. . [Електронний ресурс] // Режим доступу: https://en.wikipedia.org/wiki/Gini_coefficient
88. Linear Discriminant Analysis. [Електронний ресурс] // Режим доступу: *https://scikit-learn.org/stable/modules/generated/sklearn.discriminant_analysis.LinearDiscriminantAnalysis.html*
89. Multi-layer Perceptron classifier. [Електронний ресурс] // Режим доступу: *https://scikit-learn.org/stable/modules/generated/sklearn.neural_network.MLPClassifier.html#sklearn.neural_network.MLPClassifier*
90. Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.
91. Олешко І.В. (2014). Моделі та методи оцінки захищеності механізмів багатофакторної автентифікації від несанкціонованого доступу. Автореферат дисертації на здобуття наукового ступеня кандидата технічних наук, Харків
92. Raspberry Pi Resources. [Електронний ресурс] // Режим доступу: https://raspberry-projects.com/pi/arduino/using-arduino-shields/ardupi

93. Raspberry Pi to Arduino Shield Connection Bridge [Електронний ресурс] // Режим доступу: https://www.openhacks.com/page/productos/id/1604/title/ Raspberry-Pi-to- Arduino-Shield-Connection-Bridge
94. Raspberry Pi. [Електронний ресурс] // Режим доступу: https://en.wikipedia.org/ wiki/Raspberry_Pi
95. Raspberry Pi OS. [Електронний ресурс] // Режим доступу: https://en.wikipedia. org/wiki/Raspberry_Pi_OS
96. Сабодашко Д.В., Хома Ю.В., Хома В.В. (2020) 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.
97. Overfitting. [Електронний ресурс] // Режим доступу: https://en.wikipedia.org/ wiki/Overfitting
98. Hoekema, R., et al. (2001) “Geometrical Aspects of the Interindividual Variability of Multilead ECG Recordings.” IEEE Transactions on Biomedical Engineering, vol. 48, no. 5, pp. 551–559., doi:10.1109/10.918594.
99. Haque, Emranul & Ahmed, Feroz. (2018). ECG Signal Based Heart Disease Detection System for Telemedicine Application.
100. Huang, R., & Zhou, Y. (2015). Disease Classification and Biomarker Discovery Using ECG Data. BioMed research international, 2015, 680381. https://doi.org/10.1155/2015/680381
101. Tests for hidden heart disease. [Електронний ресурс] // Режим доступу: https://www.health.harvard.edu/heart-health/tests-for-hidden-heart-disease

ДОДАТКИ

Додаток А

ЗАТВЕРДЖУЮ
Директор
ТОВ “СВІФТ СОЛЮШНС”
ЄДРПОУ 38879401

_____ Агєєва І.Б.
61001, м. Харків,
Майдан Повстання, буд. 7/8
“__” _____ 2021 р.

АКТ
про впровадження результатів дисертації
Сабодашка Дмитра Володимировича
“Вдосконалення методів і засобів біометричної автентифікації на основі електрокардіограми” представленої на здобуття наукового ступеня доктора філософії за спеціальністю 125 "Кібербезпека"

Цим актом підтверджується використання результатів дисертації Сабодашка Дмитра Володимировича “Вдосконалення методів і засобів біометричної автентифікації на основі електрокардіограми”, зокрема:

- алгоритми темпоральної нормалізації ЕКГ-сигналів та їх імплементація мовою Python використано у системах розпізнавання квазіперіодичних сигналів, чим істотно підвищено їх точність;
- підходи для виявлення та виправлення артефактів та їх імплементація у вигляді програмної компоненти використовуються для опрацювання артефактів у модулях цифрової обробки сигналів, що покращило якість сигналів, а відтак точність систем розпізнавання на основі алгоритмів машинного навчання;
- реалізацію та верифікацію описаних вище алгоритмів на мікрокомп'ютері Raspberry Pi використано для ефективної імплементації компонентів цифрової обробки сигналів на пристроях з обмеженими обчислювальними ресурсами.

Даний акт не є підставою для фінансових розрахунків.

Директор
ТОВ “СВІФТ СОЛЮШНС” Агєєва І.Б.

ЗАТВЕРДЖУЮ

Проректор з науково-педагогічної роботи та соціального розвитку Національного університету «Львівська політехніка»

д.т.н., доцент _____ Корж Р.О.

«__» _____ 2021 р.

АКТ

про використання у навчальному процесі Національного університету «Львівська політехніка» результатів досліджень та розробок, одержаних при виконанні дисертаційної роботи Сабодашка Дмитра Володимировича “Вдосконалення методів і засобів біометричної автентифікації на основі електрокардіограми” представленої на здобуття наукового ступеня доктора філософії за спеціальністю 125 "Кібербезпека"

Комісія Національного університету «Львівська політехніка» у складі:

Голова комісії: голова науково-методичної ради інституту комп’ютерних технологій, автоматики та метрології, д.т.н., проф. Байцар Р.І.

Члени комісії: завідувач кафедри захисту інформації, д.т.н., проф. Дудикевич В.Б., професор кафедри захисту інформації, д.т.н., проф. Опірський Р.І., доцент кафедри захисту інформації, к.т.н., доцент Горпенюк А.Я.

даним актом підтверджує, що наукові дослідження Сабодашка Д.В. виконувалися ним на кафедрі захисту інформації Національного університету «Львівська політехніка». Основні положення та результати дисертаційної роботи "Вдосконалення методів і засобів біометричної автентифікації на основі електрокардіограми" впроваджено у навчальний процес кафедри захисту інформації Національного університету «Львівська політехніка» при вивченні дисциплін: "Системи автентифікації та управління доступом" та "Програмування скриптовими мовами" для підготовки фахівців з галузі знань 12 "Інформаційні технології" за спеціальністю 125 "Кібербезпека".

Голова комісії,
голова науково-методичної ради ІКТА,
д.т.н., професор _____ Байцар Р.І.

Члени комісії:
зав. каф. ЗІ, д.т.н., професор _____ Дудикевич В.Б.

проф. каф. ЗІ, д.т.н., професор _____ Опірський Р.І.

доц. каф. ЗІ, д.т.н., доцент _____ Горпенюк А.Я.

Додаток Б

Характеристика набору ЕКГ даних Lviv Biometric Dataset

N	Назва суб'єкта	Кількість вимірювань	Кількість вимірювальних сесій	Дата першого вимірювання	Дата останнього вимірювання
1	user1	9	4	26-02-2016	10-03-2016
2	user2	6	3	26-02-2016	10-03-2016
3	user3	11	1	26-04-2016	26-04-2016
4	user4	72	11	15-03-2016	27-12-2019
5	user5	7	3	26-02-2016	10-03-2016
6	user6	9	1	27-04-2016	27-04-2016
7	user7	4	2	26-02-2016	29-02-2016
8	user8	11	1	27-04-2016	27-04-2016
9	user9	50	4	26-04-2016	20-12-2019
10	user10	8	3	26-02-2016	10-03-2016
11	user11	10	1	27-04-2016	27-04-2016
12	user12	10	1	26-04-2016	26-04-2016
13	user13	49	5	02-12-2016	01-11-2019
14	user14	48	5	02-12-2016	27-12-2019
15	user15	23	4	02-12-2016	08-11-2019
16	user16	5	1	26-04-2016	26-04-2016
17	user17	31	7	28-02-2016	09-12-2019
18	user18	6	2	29-02-2016	10-03-2016
19	user20	8	1	29-05-2017	29-05-2017
20	user21	11	1	29-05-2017	29-05-2017
21	user22	3	1	29-05-2017	29-05-2017
22	user23	14	1	29-05-2017	29-05-2017

N	Назва суб'єкта	Кількість вимірювань	Кількість вимірювальних сесій	Дата першого вимірювання	Дата останнього вимірювання
23	user24	14	1	29-05-2017	29-05-2017
24	user25	13	1	29-05-2017	29-05-2017
25	user26	6	1	29-05-2017	29-05-2017
26	user27	14	1	29-05-2017	29-05-2017
27	user28	14	1	29-05-2017	29-05-2017
28	user29	14	1	29-05-2017	29-05-2017
29	user30	12	1	29-05-2017	29-05-2017
30	user31	14	1	29-05-2017	29-05-2017
31	user32	9	1	29-05-2017	29-05-2017
32	user33	14	1	29-05-2017	29-05-2017
33	user34	15	1	29-05-2017	29-05-2017
34	user35	10	1	02-06-2017	02-06-2017
35	user36	12	1	02-06-2017	02-06-2017
36	user37	12	1	02-06-2017	02-06-2017
37	user38	13	1	02-06-2017	02-06-2017
38	user39	12	1	02-06-2017	02-06-2017
39	user40	11	1	02-06-2017	02-06-2017
40	user41	12	1	02-06-2017	02-06-2017
41	user42	11	1	02-06-2017	02-06-2017
42	user43	12	1	02-06-2017	02-06-2017
43	user44	13	1	02-06-2017	02-06-2017
44	user45	12	1	02-06-2017	02-06-2017
45	user46	14	1	02-06-2017	02-06-2017
46	user47	12	1	02-06-2017	02-06-2017
47	user48	11	1	02-06-2017	02-06-2017

N	Назва суб'єкта	Кількість вимірювань	Кількість вимірювальних сесій	Дата першого вимірювання	Дата останнього вимірювання
48	user49	12	1	02-06-2017	02-06-2017
49	user50	12	1	02-06-2017	02-06-2017
50	user51	13	1	02-06-2017	02-06-2017
51	user52	13	1	02-06-2017	02-06-2017
52	user53	13	1	02-06-2017	02-06-2017
53	user54	39	3	01-12-2017	12-02-2018
54	user55	32	3	01-12-2017	01-02-2018
55	user56	9	1	01-12-2017	01-12-2017
56	user58	10	1	23-10-2017	23-10-2017
57	user59	6	1	23-10-2017	23-10-2017
58	user60	12	1	23-10-2017	23-10-2017
59	user61	10	1	23-10-2017	23-10-2017
60	user62	10	1	23-10-2017	23-10-2017
61	user63	10	1	23-10-2017	23-10-2017
62	user64	10	1	23-10-2017	23-10-2017
63	user65	10	1	23-10-2017	23-10-2017
64	user66	10	1	23-10-2017	23-10-2017
65	user67	9	1	23-10-2017	23-10-2017
66	user68	10	1	23-10-2017	23-10-2017
67	user69	10	1	23-10-2017	23-10-2017
68	user70	14	1	23-10-2017	23-10-2017
69	user71	8	1	23-10-2017	23-10-2017
70	user72	10	1	23-10-2017	23-10-2017
71	user73	37	4	25-10-2017	20-12-2019
72	user74	41	4	25-10-2017	20-12-2019

N	Назва суб'єкта	Кількість вимірювань	Кількість вимірювальних сесій	Дата першого вимірювання	Дата останнього вимірювання
73	user75	30	3	25-10-2017	20-12-2019
74	user76	28	3	25-10-2017	24-01-2018
75	user77	11	1	25-10-2017	25-10-2017
76	user78	38	4	25-10-2017	20-12-2019
77	user79	8	1	25-10-2017	25-10-2017
78	user80	57	4	25-10-2017	20-12-2019
79	user81	39	4	25-10-2017	20-12-2019
80	user82	30	3	25-10-2017	24-01-201
81	user83	21	1	25-10-2017	25-10-2017
82	user84	27	3	25-10-2017	20-12-2019
83	user85	28	3	22-11-2017	20-12-2019
84	user86	27	3	22-11-2017	20-12-2019
85	user87	35	3	22-11-2017	20-12-2019
86	user88	29	3	22-11-2017	20-12-2019
87	user89	11	1	22-11-2017	22-11-2017
88	user90	11	1	22-11-2017	22-11-2017
89	user92	10	1	22-11-2017	22-11-2017
90	user93	5	1	22-11-2017	22-11-2017
91	user94	11	1	19-12-2017	19-12-2017
92	user95	4	1	26-12-2017	26-12-2017
93	user96	16	2	24-01-2018	20-12-2019
94	user97	22	2	12-02-2018	27-12-2019
95	user98	10	1	12-02-2018	12-02-2018
96	user99	26	2	12-02-2018	27-12-2019
97	user100	10	1	12-02-2018	12-02-2018

N	Назва суб'єкта	Кількість вимірювань	Кількість вимірювальних сесій	Дата першого вимірювання	Дата останнього вимірювання
98	user101	10	1	20-12-2019	20-12-2019
99	user102	10	1	20-12-2019	20-12-2019
100	user103	8	1	20-12-2019	20-12-2019
101	user104	10	1	20-12-2019	20-12-2019
102	user105	10	1	20-12-2019	20-12-2019
103	user106	10	1	20-12-2019	20-12-2019
104	user107	10	1	20-12-2019	20-12-2019
105	user108	10	1	20-12-2019	20-12-2019
106	user109	11	1	20-12-2019	20-12-2019
107	user110	9	1	20-12-2019	20-12-2019
108	user111	9	1	20-12-2019	20-12-2019
109	user112	10	1	20-12-2019	20-12-2019
110	user113	10	1	20-12-2019	20-12-2019
111	user114	10	1	20-12-2019	20-12-2019
112	user115	10	1	27-12-2019	27-12-2019
113	user116	10	1	27-12-2019	27-12-2019
114	user117	11	1	27-12-2019	27-12-2019
115	user118	15	1	27-12-2019	27-12-2019

Додаток В

Конфігурація навчального та тестового наборів сформованих із записів Lviv Biometric Dataset.

N	Назва суб'єкта	Записи навчального набору	Записи тестового набору
1	user1	user1_2016_02_26_16_42_00.npy user1_2016_02_29_18_37_00.npy user1_2016_03_09_18_16_00.npy user1_2016_03_10_11_16_00.npy user1_2016_03_10_11_17_00.npy user1_2016_03_10_15_10_00.npy	user1_2016_02_29_18_38_00.npy user1_2016_03_10_11_14_00.npy user1_2016_03_10_15_09_00.npy
2	user2	user2_2016_03_09_18_15_00.npy user2_2016_03_10_16_02_00.npy user2_2016_03_10_16_03_00.npy user2_2016_03_10_16_04_00.npy	user2_2016_02_26_16_57_00.npy user2_2016_03_09_18_06_00.npy
3	user3	user3_2016_04_26_23_39_21.npy user3_2016_04_26_23_39_54.npy user3_2016_04_26_23_40_26.npy user3_2016_04_26_23_40_42.npy user3_2016_04_26_23_40_58.npy user3_2016_04_26_23_42_06.npy user3_2016_04_26_23_42_21.npy user3_2016_04_26_23_42_37.npy	user3_2016_04_26_23_39_38.npy user3_2016_04_26_23_40_10.npy user3_2016_04_26_23_41_42.npy
4	user4	user4_2016_03_15_12_56_00.npy user4_2016_03_15_12_57_00.npy user4_2016_03_15_12_59_00.npy user4_2016_04_25_20_24_00.npy user4_2016_04_25_20_25_00.npy user4_2016_04_25_23_34_00.npy user4_2016_04_25_23_36_00.npy user4_2017_11_22_08_44_18.npy user4_2017_11_22_08_45_47.npy user4_2017_12_01_13_33_26.npy user4_2017_12_01_13_34_06.npy user4_2017_12_01_13_34_21.npy user4_2017_12_01_13_34_34.npy user4_2017_12_01_13_34_47.npy user4_2017_12_01_13_34_59.npy user4_2017_12_01_13_35_25.npy user4_2017_12_01_13_35_51.npy user4_2017_12_19_16_56_42.npy user4_2017_12_19_16_57_00.npy user4_2017_12_19_16_57_33.npy user4_2017_12_19_16_58_21.npy user4_2017_12_19_16_58_37.npy user4_2017_12_19_16_58_52.npy user4_2017_12_19_16_59_07.npy user4_2017_12_19_16_59_34.npy user4_2017_12_26_10_32_41.npy user4_2017_12_26_10_33_31.npy user4_2018_01_23_18_16_22.npy user4_2018_01_23_18_17_27.npy user4_2018_01_23_18_17_46.npy user4_2018_01_23_18_18_04.npy user4_2018_02_01_18_59_01.npy	user4_2016_03_15_12_55_00.npy user4_2016_03_15_12_58_00.npy user4_2016_04_25_20_33_00.npy user4_2016_04_25_20_34_00.npy user4_2016_06_13_21_02_52.npy user4_2017_12_01_13_33_53.npy user4_2017_12_01_13_35_12.npy user4_2017_12_01_13_35_38.npy user4_2017_12_19_16_57_17.npy user4_2017_12_19_16_58_02.npy user4_2017_12_26_10_33_12.npy user4_2017_12_26_10_33_54.npy user4_2018_01_23_18_17_08.npy user4_2018_02_12_17_14_12.npy user4_2018_02_12_17_15_14.npy user4_2018_02_12_17_16_51.npy user4_2019_12_27_11_17_14.npy user4_2019_12_27_11_18_38.npy user4_2019_12_27_11_21_18.npy user4_2019_12_27_11_22_57.npy user4_2019_12_27_11_23_15.npy user4_2019_12_27_15_48_30.npy

		user4_2018_02_12_17_14_42.npy user4_2018_02_12_17_14_59.npy user4_2018_02_12_17_15_31.npy user4_2018_02_12_17_15_47.npy user4_2018_02_12_17_16_03.npy user4_2018_02_12_17_16_18.npy user4_2018_02_12_17_16_33.npy user4_2018_02_12_17_17_13.npy user4_2019_12_27_11_17_44.npy user4_2019_12_27_11_18_18.npy user4_2019_12_27_11_18_57.npy user4_2019_12_27_11_19_16.npy user4_2019_12_27_11_20_40.npy user4_2019_12_27_11_20_59.npy user4_2019_12_27_11_21_36.npy user4_2019_12_27_11_21_55.npy user4_2019_12_27_11_22_18.npy user4_2019_12_27_11_22_38.npy	
5	user5	user5_2016_02_29_18_23_00.npy user5_2016_02_29_18_25_00.npy user5_2016_03_10_15_06_00.npy user5_2016_03_10_15_07_00.npy user5_2016_03_10_15_08_00.npy	user5_2016_02_26_16_51_00.npy user5_2016_02_29_18_24_00.npy
6	user6	user6_2016_04_27_09_56_50.npy user6_2016_04_27_09_58_07.npy user6_2016_04_27_09_58_23.npy user6_2016_04_27_09_58_54.npy user6_2016_04_27_09_59_46.npy user6_2016_04_27_10_00_02.npy	user6_2016_04_27_09_57_12.npy user6_2016_04_27_09_57_28.npy user6_2016_04_27_09_58_38.npy
7	user7	user7_2016_02_26_16_44_00.npy user7_2016_02_29_18_31_00.npy user7_2016_02_29_18_34_00.npy	user7_2016_02_29_18_32_00.npy
8	user8	user8_2016_04_27_10_01_25.npy user8_2016_04_27_10_01_41.npy user8_2016_04_27_10_02_28.npy user8_2016_04_27_10_02_56.npy user8_2016_04_27_10_03_41.npy user8_2016_04_27_10_04_14.npy user8_2016_04_27_10_04_29.npy user8_2016_04_27_10_05_01.npy	user8_2016_04_27_10_01_57.npy user8_2016_04_27_10_03_57.npy user8_2016_04_27_10_04_46.npy
9	user9	user9_2016_04_26_10_28_35.npy user9_2016_04_26_10_28_58.npy user9_2016_04_26_10_29_40.npy user9_2016_04_26_10_29_57.npy user9_2016_04_26_10_30_13.npy user9_2016_04_26_10_30_29.npy user9_2016_04_26_10_30_45.npy user9_2016_04_26_10_31_01.npy user9_2016_04_26_10_31_34.npy user9_2016_04_26_10_31_50.npy user9_2016_04_26_10_32_06.npy user9_2017_10_25_09_54_31.npy user9_2017_10_25_09_54_45.npy user9_2017_10_25_09_55_26.npy user9_2017_10_25_09_55_39.npy user9_2017_10_25_09_56_05.npy user9_2017_11_22_09_43_36.npy user9_2017_11_22_09_43_51.npy user9_2017_11_22_09_44_06.npy	user9_2016_04_26_10_31_17.npy user9_2016_04_26_10_34_07.npy user9_2017_10_25_09_53_50.npy user9_2017_10_25_09_54_04.npy user9_2017_10_25_09_54_18.npy user9_2017_10_25_09_54_57.npy user9_2017_10_25_09_55_12.npy user9_2017_10_25_09_55_52.npy user9_2017_11_22_09_44_21.npy user9_2017_11_22_09_46_07.npy user9_2019_12_20_11_19_18.npy user9_2019_12_20_11_19_53.npy user9_2019_12_20_11_43_54.npy user9_2019_12_20_11_44_10.npy user9_2019_12_20_11_44_43.npy

		user9_2017_11_22_09_44_36.npy user9_2017_11_22_09_44_51.npy user9_2017_11_22_09_45_06.npy user9_2017_11_22_09_45_22.npy user9_2017_11_22_09_45_37.npy user9_2017_11_22_09_45_52.npy user9_2019_12_20_11_19_35.npy user9_2019_12_20_11_20_09.npy user9_2019_12_20_11_20_27.npy user9_2019_12_20_11_20_44.npy user9_2019_12_20_11_42_34.npy user9_2019_12_20_11_42_50.npy user9_2019_12_20_11_43_06.npy user9_2019_12_20_11_43_22.npy user9_2019_12_20_11_43_38.npy user9_2019_12_20_11_44_27.npy	
10	user10	user10_2016_02_26_16_30_00.npy user10_2016_02_29_18_11_00.npy user10_2016_03_10_14_54_00.npy user10_2016_03_10_14_57_00.npy user10_2016_03_10_16_48_00.npy user10_2016_03_10_16_50_00.npy	user10_2016_02_29_18_12_00.npy user10_2016_03_10_14_56_00.npy user10_2016_03_10_16_49_00.npy
11	user11	user11_2016_04_27_09_51_43.npy user11_2016_04_27_09_52_02.npy user11_2016_04_27_09_52_18.npy user11_2016_04_27_09_52_35.npy user11_2016_04_27_09_52_56.npy user11_2016_04_27_09_53_13.npy user11_2016_04_27_09_53_47.npy	user11_2016_04_27_09_53_31.npy user11_2016_04_27_09_54_04.npy user11_2016_04_27_09_54_22.npy
12	user12	user12_2016_04_26_11_39_08.npy user12_2016_04_26_11_40_33.npy user12_2016_04_26_11_40_49.npy user12_2016_04_26_11_42_25.npy user12_2016_04_26_11_42_40.npy user12_2016_04_26_11_43_12.npy user12_2016_04_26_11_43_29.npy	user12_2016_04_26_11_40_17.npy user12_2016_04_26_11_41_53.npy user12_2016_04_26_11_42_56.npy
13	user13	user13_2016_12_02_14_11_00.npy user13_2017_12_01_16_22_02.npy user13_2017_12_01_16_22_17.npy user13_2017_12_01_16_22_48.npy user13_2017_12_01_16_23_03.npy user13_2017_12_01_16_23_19.npy user13_2017_12_01_16_23_34.npy user13_2017_12_01_16_23_49.npy user13_2017_12_01_16_24_20.npy user13_2017_12_01_16_24_35.npy user13_2017_12_19_17_14_21.npy user13_2017_12_19_17_14_37.npy user13_2017_12_19_17_14_51.npy user13_2017_12_19_17_15_06.npy user13_2017_12_19_17_15_51.npy user13_2017_12_19_17_16_06.npy user13_2017_12_19_17_16_23.npy user13_2017_12_19_17_16_38.npy user13_2018_02_12_17_22_05.npy user13_2018_02_12_17_22_36.npy user13_2018_02_12_17_23_06.npy user13_2018_02_12_17_23_21.npy user13_2018_02_12_17_24_21.npy	user13_2016_12_02_14_12_00.npy user13_2016_12_02_14_13_00.npy user13_2017_12_01_16_22_33.npy user13_2017_12_19_17_15_22.npy user13_2017_12_19_17_15_36.npy user13_2017_12_19_17_16_53.npy user13_2018_02_12_17_22_21.npy user13_2018_02_12_17_23_36.npy user13_2018_02_12_17_23_51.npy user13_2018_02_12_17_24_06.npy user13_2019_11_01_16_14_29.npy user13_2019_11_01_16_17_25.npy user13_2019_11_01_16_17_44.npy user13_2019_11_01_16_19_02.npy user13_2019_11_01_16_24_18.npy

		user13_2018_02_12_17_24_37.npy user13_2019_11_01_16_13_23.npy user13_2019_11_01_16_16_44.npy user13_2019_11_01_16_17_06.npy user13_2019_11_01_16_18_07.npy user13_2019_11_01_16_18_25.npy user13_2019_11_01_16_18_45.npy user13_2019_11_01_16_23_32.npy user13_2019_11_01_16_23_54.npy user13_2019_11_01_16_24_34.npy user13_2019_11_01_16_24_51.npy	
14	user14	user14_2016_12_02_14_44_00.npy user14_2016_12_02_14_46_00.npy user14_2016_12_02_14_48_00.npy user14_2017_12_01_13_21_11.npy user14_2017_12_01_13_21_26.npy user14_2017_12_01_13_22_09.npy user14_2017_12_01_13_22_22.npy user14_2017_12_01_13_22_36.npy user14_2017_12_01_13_22_49.npy user14_2017_12_01_13_23_02.npy user14_2017_12_19_17_17_42.npy user14_2017_12_19_17_17_57.npy user14_2017_12_19_17_18_11.npy user14_2017_12_19_17_18_26.npy user14_2017_12_19_17_18_41.npy user14_2017_12_19_17_19_56.npy user14_2017_12_19_17_20_11.npy user14_2018_02_12_17_43_43.npy user14_2018_02_12_17_44_00.npy user14_2018_02_12_17_44_15.npy user14_2018_02_12_17_44_29.npy user14_2018_02_12_17_44_44.npy user14_2018_02_12_17_44_59.npy user14_2018_02_12_17_45_15.npy user14_2018_02_12_17_45_33.npy user14_2018_02_12_17_45_50.npy user14_2018_02_12_17_46_05.npy user14_2019_12_27_11_24_28.npy user14_2019_12_27_11_24_45.npy user14_2019_12_27_11_25_19.npy user14_2019_12_27_11_26_11.npy user14_2019_12_27_11_26_48.npy user14_2019_12_27_11_27_05.npy user14_2019_12_27_11_27_21.npy	user14_2016_12_02_14_50_00.npy user14_2017_12_01_13_20_56.npy user14_2017_12_01_13_21_40.npy user14_2017_12_01_13_21_54.npy user14_2017_12_01_13_23_15.npy user14_2017_12_19_17_18_56.npy user14_2017_12_19_17_19_11.npy user14_2017_12_19_17_19_26.npy user14_2017_12_19_17_19_40.npy user14_2018_02_12_17_46_20.npy user14_2019_12_27_11_25_02.npy user14_2019_12_27_11_25_35.npy user14_2019_12_27_11_25_52.npy user14_2019_12_27_11_26_29.npy
15	user15	user15_2016_12_02_14_31_00.npy user15_2016_12_02_14_33_00.npy user15_2016_12_02_14_34_00.npy user15_2017_12_01_13_37_13.npy user15_2017_12_01_13_37_27.npy user15_2017_12_01_13_37_40.npy user15_2017_12_01_13_37_54.npy user15_2017_12_01_13_38_21.npy user15_2017_12_01_13_38_48.npy user15_2017_12_01_13_39_01.npy user15_2018_02_12_17_51_18.npy user15_2018_02_12_17_51_34.npy user15_2019_11_08_15_56_18.npy user15_2019_11_08_15_56_34.npy user15_2019_11_08_15_57_07.npy	user15_2016_12_02_14_32_00.npy user15_2017_12_01_13_37_00.npy user15_2017_12_01_13_38_07.npy user15_2017_12_01_13_38_34.npy user15_2017_12_01_13_39_15.npy user15_2019_11_08_15_55_56.npy user15_2019_11_08_15_57_42.npy

		user15_2019_11_08_15_57_24.npy	
16	user16	user16_2016_04_26_12_24_00.npy user16_2016_04_26_12_24_36.npy user16_2016_04_26_12_26_39.npy user16_2016_04_26_12_26_55.npy	user16_2016_04_26_12_24_20.npy
17	user17	user17_2016_02_28_16_19_00.npy user17_2016_02_29_18_20_00.npy user17_2016_03_09_18_02_00.npy user17_2016_03_09_18_10_00.npy user17_2016_03_11_16_22_00.npy user17_2018_02_12_17_47_45.npy user17_2018_02_12_17_48_16.npy user17_2018_02_12_17_48_31.npy user17_2018_02_12_17_48_47.npy user17_2018_02_12_17_49_17.npy user17_2018_02_12_17_49_33.npy user17_2018_02_12_17_49_48.npy user17_2018_02_12_17_50_03.npy user17_2018_02_12_17_50_18.npy user17_2019_12_09_14_05_02.npy user17_2019_12_09_14_05_19.npy user17_2019_12_09_14_05_38.npy user17_2019_12_09_14_06_01.npy user17_2019_12_09_14_06_22.npy user17_2019_12_09_14_06_40.npy user17_2019_12_09_14_07_14.npy user17_2019_12_09_14_07_33.npy	user17_2016_02_29_18_16_00.npy user17_2016_02_29_18_21_00.npy user17_2016_03_09_18_04_00.npy user17_2016_03_09_18_09_00.npy user17_2016_03_10_11_13_00.npy user17_2018_02_12_17_48_01.npy user17_2018_02_12_17_49_01.npy user17_2019_12_09_14_04_43.npy user17_2019_12_09_14_06_57.npy
18	user18	user18_2016_03_10_15_58_00.npy user18_2016_03_10_15_59_00.npy user18_2016_03_10_16_55_00.npy user18_2016_03_10_16_56_00.npy	user18_2016_02_29_18_40_00.npy user18_2016_03_10_16_00_00.npy
19	user20	user20_2017_05_29_11_22_26.npy user20_2017_05_29_11_23_40.npy user20_2017_05_29_11_24_03.npy user20_2017_05_29_11_24_22.npy user20_2017_05_29_11_24_43.npy user20_2017_05_29_11_25_03.npy	user20_2017_05_29_11_22_54.npy user20_2017_05_29_11_23_17.npy
20	user21	user21_2017_05_29_11_27_18.npy user21_2017_05_29_11_28_00.npy user21_2017_05_29_11_28_19.npy user21_2017_05_29_11_28_59.npy user21_2017_05_29_11_29_20.npy user21_2017_05_29_11_29_44.npy user21_2017_05_29_11_30_03.npy user21_2017_05_29_11_30_24.npy	user21_2017_05_29_11_26_33.npy user21_2017_05_29_11_27_39.npy user21_2017_05_29_11_28_38.npy
21	user22	user22_2017_05_29_11_32_33.npy user22_2017_05_29_11_32_54.npy	user22_2017_05_29_11_32_12.npy
22	user23	user23_2017_05_29_11_34_36.npy user23_2017_05_29_11_35_17.npy user23_2017_05_29_11_35_37.npy user23_2017_05_29_11_35_56.npy user23_2017_05_29_11_37_10.npy user23_2017_05_29_11_37_31.npy user23_2017_05_29_11_37_50.npy user23_2017_05_29_11_38_13.npy user23_2017_05_29_11_38_33.npy user23_2017_05_29_11_38_54.npy	user23_2017_05_29_11_34_56.npy user23_2017_05_29_11_36_17.npy user23_2017_05_29_11_36_36.npy user23_2017_05_29_11_39_27.npy
23	user24	user24_2017_05_29_11_41_35.npy	user24_2017_05_29_11_41_12.npy

		user24_2017_05_29_11_41_54.npy user24_2017_05_29_11_42_19.npy user24_2017_05_29_11_42_39.npy user24_2017_05_29_11_42_59.npy user24_2017_05_29_11_43_19.npy user24_2017_05_29_11_44_17.npy user24_2017_05_29_11_44_36.npy user24_2017_05_29_11_45_41.npy user24_2017_05_29_11_46_00.npy	user24_2017_05_29_11_43_38.npy user24_2017_05_29_11_43_58.npy user24_2017_05_29_11_44_55.npy
24	user25	user25_2017_05_29_11_47_28.npy user25_2017_05_29_11_47_48.npy user25_2017_05_29_11_48_08.npy user25_2017_05_29_11_48_27.npy user25_2017_05_29_11_48_47.npy user25_2017_05_29_11_49_08.npy user25_2017_05_29_11_50_07.npy user25_2017_05_29_11_50_26.npy user25_2017_05_29_11_51_03.npy	user25_2017_05_29_11_47_02.npy user25_2017_05_29_11_49_27.npy user25_2017_05_29_11_49_47.npy user25_2017_05_29_11_50_44.npy
25	user26	user26_2017_05_29_11_52_43.npy user26_2017_05_29_11_53_39.npy user26_2017_05_29_11_54_18.npy user26_2017_05_29_11_54_38.npy	user26_2017_05_29_11_53_18.npy user26_2017_05_29_11_53_58.npy
26	user27	user27_2017_05_29_11_56_13.npy user27_2017_05_29_11_57_44.npy user27_2017_05_29_11_58_03.npy user27_2017_05_29_11_58_44.npy user27_2017_05_29_11_59_02.npy user27_2017_05_29_11_59_22.npy user27_2017_05_29_12_00_00.npy user27_2017_05_29_12_00_19.npy user27_2017_05_29_12_00_59.npy user27_2017_05_29_12_01_18.npy	user27_2017_05_29_11_57_03.npy user27_2017_05_29_11_57_23.npy user27_2017_05_29_11_58_23.npy user27_2017_05_29_11_59_40.npy
27	user28	user28_2017_05_29_12_02_33.npy user28_2017_05_29_12_02_52.npy user28_2017_05_29_12_03_30.npy user28_2017_05_29_12_04_08.npy user28_2017_05_29_12_04_30.npy user28_2017_05_29_12_04_49.npy user28_2017_05_29_12_05_08.npy user28_2017_05_29_12_05_27.npy user28_2017_05_29_12_05_46.npy user28_2017_05_29_12_06_06.npy	user28_2017_05_29_12_02_14.npy user28_2017_05_29_12_03_10.npy user28_2017_05_29_12_03_49.npy user28_2017_05_29_12_06_26.npy
28	user29	user29_2017_05_29_12_07_32.npy user29_2017_05_29_12_07_51.npy user29_2017_05_29_12_08_11.npy user29_2017_05_29_12_09_25.npy user29_2017_05_29_12_09_42.npy user29_2017_05_29_12_09_59.npy user29_2017_05_29_12_10_15.npy user29_2017_05_29_12_10_32.npy user29_2017_05_29_12_11_05.npy user29_2017_05_29_12_11_22.npy	user29_2017_05_29_12_08_30.npy user29_2017_05_29_12_08_49.npy user29_2017_05_29_12_10_49.npy user29_2017_05_29_12_11_39.npy
29	user30	user30_2017_05_29_12_12_29.npy user30_2017_05_29_12_14_16.npy user30_2017_05_29_12_14_33.npy user30_2017_05_29_12_14_50.npy user30_2017_05_29_12_15_23.npy user30_2017_05_29_12_15_39.npy user30_2017_05_29_12_16_05.npy	user30_2017_05_29_12_13_08.npy user30_2017_05_29_12_13_26.npy user30_2017_05_29_12_13_43.npy user30_2017_05_29_12_13_59.npy

		user30_2017_05_29_12_16_21.npy	
30	user31	user31_2017_05_29_12_18_00.npy user31_2017_05_29_12_18_16.npy user31_2017_05_29_12_18_33.npy user31_2017_05_29_12_18_49.npy user31_2017_05_29_12_19_21.npy user31_2017_05_29_12_19_38.npy user31_2017_05_29_12_19_54.npy user31_2017_05_29_12_20_10.npy user31_2017_05_29_12_20_28.npy user31_2017_05_29_12_20_44.npy	user31_2017_05_29_12_17_08.npy user31_2017_05_29_12_17_25.npy user31_2017_05_29_12_17_43.npy user31_2017_05_29_12_19_05.npy
31	user32	user32_2017_05_29_12_21_50.npy user32_2017_05_29_12_22_23.npy user32_2017_05_29_12_22_42.npy user32_2017_05_29_12_22_59.npy user32_2017_05_29_12_23_16.npy user32_2017_05_29_12_23_49.npy	user32_2017_05_29_12_22_06.npy user32_2017_05_29_12_23_33.npy user32_2017_05_29_12_24_06.npy
32	user33	user33_2017_05_29_12_25_45.npy user33_2017_05_29_12_26_01.npy user33_2017_05_29_12_26_19.npy user33_2017_05_29_12_26_36.npy user33_2017_05_29_12_27_36.npy user33_2017_05_29_12_27_53.npy user33_2017_05_29_12_28_11.npy user33_2017_05_29_12_28_28.npy user33_2017_05_29_12_29_06.npy user33_2017_05_29_12_29_40.npy	user33_2017_05_29_12_26_53.npy user33_2017_05_29_12_27_10.npy user33_2017_05_29_12_28_45.npy user33_2017_05_29_12_29_23.npy
33	user34	user34_2017_05_29_12_31_38.npy user34_2017_05_29_12_31_55.npy user34_2017_05_29_12_32_11.npy user34_2017_05_29_12_32_29.npy user34_2017_05_29_12_33_01.npy user34_2017_05_29_12_33_18.npy user34_2017_05_29_12_33_51.npy user34_2017_05_29_12_34_08.npy user34_2017_05_29_12_34_24.npy user34_2017_05_29_12_34_41.npy	user34_2017_05_29_12_30_39.npy user34_2017_05_29_12_30_56.npy user34_2017_05_29_12_31_22.npy user34_2017_05_29_12_32_45.npy user34_2017_05_29_12_33_34.npy
34	user35	user35_2017_06_02_11_01_17.npy user35_2017_06_02_11_01_47.npy user35_2017_06_02_11_02_33.npy user35_2017_06_02_11_04_42.npy user35_2017_06_02_11_05_02.npy user35_2017_06_02_11_05_20.npy user35_2017_06_02_11_07_29.npy	user35_2017_06_02_11_02_53.npy user35_2017_06_02_11_05_39.npy user35_2017_06_02_11_07_09.npy
35	user36	user36_2017_06_02_11_11_19.npy user36_2017_06_02_11_11_58.npy user36_2017_06_02_11_12_51.npy user36_2017_06_02_11_13_08.npy user36_2017_06_02_11_13_26.npy user36_2017_06_02_11_13_44.npy user36_2017_06_02_11_14_18.npy user36_2017_06_02_11_14_35.npy	user36_2017_06_02_11_11_39.npy user36_2017_06_02_11_12_15.npy user36_2017_06_02_11_12_33.npy user36_2017_06_02_11_14_00.npy
36	user37	user37_2017_06_02_11_15_30.npy user37_2017_06_02_11_15_47.npy user37_2017_06_02_11_18_02.npy user37_2017_06_02_11_18_20.npy user37_2017_06_02_11_18_36.npy user37_2017_06_02_11_18_52.npy user37_2017_06_02_11_19_09.npy	user37_2017_06_02_11_16_04.npy user37_2017_06_02_11_16_22.npy user37_2017_06_02_11_17_11.npy user37_2017_06_02_11_19_41.npy

		user37_2017_06_02_11_19_25.npy	
37	user38	user38_2017_06_02_11_21_03.npy user38_2017_06_02_11_21_20.npy user38_2017_06_02_11_22_11.npy user38_2017_06_02_11_22_27.npy user38_2017_06_02_11_22_46.npy user38_2017_06_02_11_23_03.npy user38_2017_06_02_11_23_22.npy user38_2017_06_02_11_23_38.npy user38_2017_06_02_11_23_56.npy	user38_2017_06_02_11_20_45.npy user38_2017_06_02_11_21_37.npy user38_2017_06_02_11_21_54.npy user38_2017_06_02_11_24_12.npy
38	user39	user39_2017_06_02_11_24_58.npy user39_2017_06_02_11_25_14.npy user39_2017_06_02_11_25_31.npy user39_2017_06_02_11_25_50.npy user39_2017_06_02_11_26_57.npy user39_2017_06_02_11_27_31.npy user39_2017_06_02_11_27_48.npy user39_2017_06_02_11_28_05.npy	user39_2017_06_02_11_26_07.npy user39_2017_06_02_11_26_24.npy user39_2017_06_02_11_26_40.npy user39_2017_06_02_11_27_14.npy
39	user40	user40_2017_06_02_11_28_55.npy user40_2017_06_02_11_29_12.npy user40_2017_06_02_11_29_28.npy user40_2017_06_02_11_30_08.npy user40_2017_06_02_11_30_45.npy user40_2017_06_02_11_31_02.npy user40_2017_06_02_11_31_35.npy user40_2017_06_02_11_32_11.npy	user40_2017_06_02_11_29_50.npy user40_2017_06_02_11_30_28.npy user40_2017_06_02_11_31_18.npy
40	user41	user41_2017_06_02_11_32_54.npy user41_2017_06_02_11_33_11.npy user41_2017_06_02_11_33_27.npy user41_2017_06_02_11_34_33.npy user41_2017_06_02_11_35_07.npy user41_2017_06_02_11_35_24.npy user41_2017_06_02_11_35_56.npy user41_2017_06_02_11_36_13.npy	user41_2017_06_02_11_33_44.npy user41_2017_06_02_11_34_00.npy user41_2017_06_02_11_34_16.npy user41_2017_06_02_11_34_50.npy
41	user42	user42_2017_06_02_11_37_54.npy user42_2017_06_02_11_38_11.npy user42_2017_06_02_11_38_27.npy user42_2017_06_02_11_38_45.npy user42_2017_06_02_11_39_04.npy user42_2017_06_02_11_39_21.npy user42_2017_06_02_11_39_57.npy user42_2017_06_02_11_40_13.npy	user42_2017_06_02_11_37_18.npy user42_2017_06_02_11_37_37.npy user42_2017_06_02_11_39_39.npy
42	user43	user43_2017_06_02_11_50_08.npy user43_2017_06_02_11_50_30.npy user43_2017_06_02_11_51_24.npy user43_2017_06_02_11_51_41.npy user43_2017_06_02_11_52_03.npy user43_2017_06_02_11_52_20.npy user43_2017_06_02_11_52_54.npy user43_2017_06_02_11_53_11.npy	user43_2017_06_02_11_50_51.npy user43_2017_06_02_11_51_07.npy user43_2017_06_02_11_52_37.npy user43_2017_06_02_11_53_28.npy
43	user44	user44_2017_06_02_11_54_53.npy user44_2017_06_02_11_55_11.npy user44_2017_06_02_11_55_30.npy user44_2017_06_02_11_55_49.npy user44_2017_06_02_11_56_07.npy user44_2017_06_02_11_56_25.npy user44_2017_06_02_11_57_01.npy user44_2017_06_02_11_57_37.npy user44_2017_06_02_11_57_54.npy	user44_2017_06_02_11_54_34.npy user44_2017_06_02_11_56_43.npy user44_2017_06_02_11_57_19.npy user44_2017_06_02_11_58_13.npy

44	user45	user45_2017_06_02_12_03_33.npy user45_2017_06_02_12_03_51.npy user45_2017_06_02_12_04_25.npy user45_2017_06_02_12_04_46.npy user45_2017_06_02_12_05_03.npy user45_2017_06_02_12_05_19.npy user45_2017_06_02_12_06_09.npy user45_2017_06_02_12_06_32.npy	user45_2017_06_02_12_03_17.npy user45_2017_06_02_12_04_07.npy user45_2017_06_02_12_05_35.npy user45_2017_06_02_12_05_52.npy
45	user46	user46_2017_06_02_12_07_45.npy user46_2017_06_02_12_08_02.npy user46_2017_06_02_12_08_36.npy user46_2017_06_02_12_08_53.npy user46_2017_06_02_12_09_31.npy user46_2017_06_02_12_09_47.npy user46_2017_06_02_12_10_22.npy user46_2017_06_02_12_11_04.npy user46_2017_06_02_12_11_22.npy user46_2017_06_02_12_11_43.npy	user46_2017_06_02_12_08_19.npy user46_2017_06_02_12_09_11.npy user46_2017_06_02_12_10_05.npy user46_2017_06_02_12_10_41.npy
46	user47	user47_2017_06_02_12_13_00.npy user47_2017_06_02_12_13_17.npy user47_2017_06_02_12_13_34.npy user47_2017_06_02_12_14_28.npy user47_2017_06_02_12_15_03.npy user47_2017_06_02_12_15_41.npy user47_2017_06_02_12_16_06.npy user47_2017_06_02_12_16_23.npy	user47_2017_06_02_12_13_50.npy user47_2017_06_02_12_14_08.npy user47_2017_06_02_12_14_46.npy user47_2017_06_02_12_15_23.npy
47	user48	user48_2017_06_02_12_18_57.npy user48_2017_06_02_12_19_50.npy user48_2017_06_02_12_20_08.npy user48_2017_06_02_12_20_24.npy user48_2017_06_02_12_20_41.npy user48_2017_06_02_12_21_13.npy user48_2017_06_02_12_21_30.npy user48_2017_06_02_12_21_47.npy	user48_2017_06_02_12_19_14.npy user48_2017_06_02_12_19_31.npy user48_2017_06_02_12_20_57.npy
48	user49	user49_2017_06_02_12_22_57.npy user49_2017_06_02_12_23_47.npy user49_2017_06_02_12_24_04.npy user49_2017_06_02_12_24_38.npy user49_2017_06_02_12_24_55.npy user49_2017_06_02_12_25_28.npy user49_2017_06_02_12_25_44.npy user49_2017_06_02_12_26_01.npy	user49_2017_06_02_12_23_14.npy user49_2017_06_02_12_23_31.npy user49_2017_06_02_12_24_21.npy user49_2017_06_02_12_25_11.npy
49	user50	user50_2017_06_02_12_26_35.npy user50_2017_06_02_12_27_09.npy user50_2017_06_02_12_27_26.npy user50_2017_06_02_12_27_43.npy user50_2017_06_02_12_28_03.npy user50_2017_06_02_12_28_38.npy user50_2017_06_02_12_29_12.npy user50_2017_06_02_12_29_29.npy	user50_2017_06_02_12_26_52.npy user50_2017_06_02_12_28_20.npy user50_2017_06_02_12_28_55.npy user50_2017_06_02_12_29_46.npy
50	user51	user51_2017_06_02_12_31_04.npy user51_2017_06_02_12_31_55.npy user51_2017_06_02_12_32_15.npy user51_2017_06_02_12_33_18.npy user51_2017_06_02_12_33_39.npy user51_2017_06_02_12_33_58.npy user51_2017_06_02_12_34_24.npy user51_2017_06_02_12_35_01.npy user51_2017_06_02_12_35_57.npy	user51_2017_06_02_12_32_40.npy user51_2017_06_02_12_34_43.npy user51_2017_06_02_12_35_20.npy user51_2017_06_02_12_35_39.npy

51	user52	user52_2017_06_02_12_37_38.npy user52_2017_06_02_12_38_08.npy user52_2017_06_02_12_38_27.npy user52_2017_06_02_12_39_55.npy user52_2017_06_02_12_40_18.npy user52_2017_06_02_12_40_57.npy user52_2017_06_02_12_41_34.npy user52_2017_06_02_12_41_52.npy user52_2017_06_02_12_42_12.npy	user52_2017_06_02_12_38_47.npy user52_2017_06_02_12_39_33.npy user52_2017_06_02_12_40_36.npy user52_2017_06_02_12_41_16.npy
52	user53	user53_2017_06_02_12_44_28.npy user53_2017_06_02_12_44_46.npy user53_2017_06_02_12_45_21.npy user53_2017_06_02_12_45_41.npy user53_2017_06_02_12_45_58.npy user53_2017_06_02_12_46_18.npy user53_2017_06_02_12_46_37.npy user53_2017_06_02_12_46_57.npy user53_2017_06_02_12_47_15.npy	user53_2017_06_02_12_43_34.npy user53_2017_06_02_12_43_52.npy user53_2017_06_02_12_44_09.npy user53_2017_06_02_12_45_04.npy
53	user54	user54_2017_12_01_12_27_34.npy user54_2017_12_01_12_27_48.npy user54_2017_12_01_12_28_01.npy user54_2017_12_01_12_28_55.npy user54_2017_12_01_12_29_36.npy user54_2017_12_01_12_30_07.npy user54_2017_12_01_12_30_21.npy user54_2017_12_01_12_30_34.npy user54_2017_12_01_12_31_14.npy user54_2017_12_01_12_31_27.npy user54_2017_12_01_12_31_40.npy user54_2017_12_01_12_31_53.npy user54_2017_12_01_12_32_06.npy user54_2017_12_18_14_51_37.npy user54_2017_12_18_14_51_58.npy user54_2017_12_18_14_52_16.npy user54_2017_12_18_14_53_13.npy user54_2017_12_18_14_53_29.npy user54_2017_12_18_14_53_50.npy user54_2017_12_18_14_54_05.npy user54_2017_12_18_14_54_19.npy user54_2018_02_12_17_18_21.npy user54_2018_02_12_17_18_36.npy user54_2018_02_12_17_19_20.npy user54_2018_02_12_17_19_49.npy user54_2018_02_12_17_20_03.npy user54_2018_02_12_17_20_32.npy	user54_2017_12_01_12_28_42.npy user54_2017_12_01_12_29_23.npy user54_2017_12_01_12_29_54.npy user54_2017_12_01_12_30_47.npy user54_2017_12_18_14_52_31.npy user54_2017_12_18_14_52_46.npy user54_2017_12_18_14_54_34.npy user54_2018_02_12_17_18_51.npy user54_2018_02_12_17_19_05.npy user54_2018_02_12_17_19_34.npy user54_2018_02_12_17_20_18.npy user54_2018_02_12_17_20_47.npy
54	user55	user55_2017_12_01_13_24_47.npy user55_2017_12_01_13_25_00.npy user55_2017_12_01_13_25_39.npy user55_2017_12_01_13_25_53.npy user55_2017_12_01_13_26_07.npy user55_2017_12_01_13_26_20.npy user55_2017_12_01_13_26_33.npy user55_2017_12_19_17_11_13.npy user55_2017_12_19_17_11_29.npy user55_2017_12_19_17_11_44.npy user55_2017_12_19_17_11_59.npy user55_2017_12_19_17_12_58.npy user55_2017_12_19_17_13_27.npy user55_2017_12_19_17_13_42.npy	user55_2017_12_01_13_25_13.npy user55_2017_12_01_13_25_26.npy user55_2017_12_01_13_26_46.npy user55_2017_12_19_17_12_14.npy user55_2017_12_19_17_12_29.npy user55_2017_12_19_17_12_44.npy user55_2017_12_19_17_13_13.npy user55_2018_02_01_19_00_25.npy user55_2018_02_01_19_00_38.npy user55_2018_02_01_19_01_45.npy

		user55_2018_02_01_19_00_11.npy user55_2018_02_01_19_00_52.npy user55_2018_02_01_19_01_05.npy user55_2018_02_01_19_01_18.npy user55_2018_02_01_19_01_32.npy user55_2018_02_01_19_01_58.npy user55_2018_02_01_19_02_12.npy user55_2018_02_01_19_02_25.npy	
55	user56	user56_2017_12_01_13_28_30.npy user56_2017_12_01_13_30_52.npy user56_2017_12_01_13_31_18.npy user56_2017_12_01_13_31_31.npy user56_2017_12_01_13_31_45.npy user56_2017_12_01_13_32_27.npy	user56_2017_12_01_13_30_11.npy user56_2017_12_01_13_30_38.npy user56_2017_12_01_13_31_05.npy
56	user58	user58_2017_10_23_11_48_41.npy user58_2017_10_23_11_49_19.npy user58_2017_10_23_11_49_38.npy user58_2017_10_23_11_49_56.npy user58_2017_10_23_11_51_14.npy user58_2017_10_23_11_51_32.npy user58_2017_10_23_11_52_08.npy	user58_2017_10_23_11_49_01.npy user58_2017_10_23_11_50_54.npy user58_2017_10_23_11_51_51.npy
57	user59	user59_2017_10_23_11_55_42.npy user59_2017_10_23_11_56_02.npy user59_2017_10_23_11_56_39.npy user59_2017_10_23_11_57_57.npy	user59_2017_10_23_11_56_21.npy user59_2017_10_23_11_57_00.npy
58	user60	user60_2017_10_23_11_58_49.npy user60_2017_10_23_11_59_08.npy user60_2017_10_23_11_59_28.npy user60_2017_10_23_11_59_46.npy user60_2017_10_23_12_00_59.npy user60_2017_10_23_12_01_19.npy user60_2017_10_23_12_01_54.npy user60_2017_10_23_12_02_11.npy	user60_2017_10_23_12_00_04.npy user60_2017_10_23_12_00_23.npy user60_2017_10_23_12_00_41.npy user60_2017_10_23_12_01_36.npy
59	user61	user61_2017_10_23_12_02_55.npy user61_2017_10_23_12_03_15.npy user61_2017_10_23_12_03_34.npy user61_2017_10_23_12_04_12.npy user61_2017_10_23_12_04_32.npy user61_2017_10_23_12_04_49.npy user61_2017_10_23_12_05_05.npy	user61_2017_10_23_12_03_52.npy user61_2017_10_23_12_05_21.npy user61_2017_10_23_12_05_36.npy
60	user62	user62_2017_10_23_12_06_35.npy user62_2017_10_23_12_06_52.npy user62_2017_10_23_12_07_27.npy user62_2017_10_23_12_07_43.npy user62_2017_10_23_12_08_01.npy user62_2017_10_23_12_08_16.npy user62_2017_10_23_12_08_48.npy	user62_2017_10_23_12_07_09.npy user62_2017_10_23_12_08_32.npy user62_2017_10_23_12_09_04.npy
61	user63	user63_2017_10_23_12_10_03.npy user63_2017_10_23_12_10_19.npy user63_2017_10_23_12_10_36.npy user63_2017_10_23_12_10_52.npy user63_2017_10_23_12_11_08.npy user63_2017_10_23_12_11_24.npy user63_2017_10_23_12_12_17.npy	user63_2017_10_23_12_09_47.npy user63_2017_10_23_12_11_39.npy user63_2017_10_23_12_11_56.npy
62	user64	user64_2017_10_23_12_13_27.npy user64_2017_10_23_12_13_45.npy user64_2017_10_23_12_14_21.npy user64_2017_10_23_12_14_40.npy	user64_2017_10_23_12_13_10.npy user64_2017_10_23_12_14_02.npy user64_2017_10_23_12_15_48.npy

		user64_2017_10_23_12_14_57.npy user64_2017_10_23_12_15_15.npy user64_2017_10_23_12_15_32.npy	
63	user65	user65_2017_10_23_12_17_03.npy user65_2017_10_23_12_17_19.npy user65_2017_10_23_12_17_52.npy user65_2017_10_23_12_18_09.npy user65_2017_10_23_12_18_26.npy user65_2017_10_23_12_18_42.npy user65_2017_10_23_12_18_59.npy	user65_2017_10_23_12_16_45.npy user65_2017_10_23_12_17_35.npy user65_2017_10_23_12_19_16.npy
64	user66	user66_2017_10_23_12_20_40.npy user66_2017_10_23_12_20_56.npy user66_2017_10_23_12_21_29.npy user66_2017_10_23_12_21_47.npy user66_2017_10_23_12_22_05.npy user66_2017_10_23_12_22_22.npy user66_2017_10_23_12_22_56.npy	user66_2017_10_23_12_20_22.npy user66_2017_10_23_12_21_13.npy user66_2017_10_23_12_22_39.npy
65	user67	user67_2017_10_23_12_23_46.npy user67_2017_10_23_12_24_03.npy user67_2017_10_23_12_24_21.npy user67_2017_10_23_12_24_37.npy user67_2017_10_23_12_25_28.npy user67_2017_10_23_12_25_45.npy	user67_2017_10_23_12_24_53.npy user67_2017_10_23_12_25_10.npy user67_2017_10_23_12_26_01.npy
66	user68	user68_2017_10_23_12_27_28.npy user68_2017_10_23_12_27_46.npy user68_2017_10_23_12_28_18.npy user68_2017_10_23_12_28_51.npy user68_2017_10_23_12_29_21.npy user68_2017_10_23_12_29_41.npy user68_2017_10_23_12_30_00.npy	user68_2017_10_23_12_26_55.npy user68_2017_10_23_12_27_11.npy user68_2017_10_23_12_30_19.npy
67	user69	user69_2017_10_23_12_35_21.npy user69_2017_10_23_12_35_38.npy user69_2017_10_23_12_35_55.npy user69_2017_10_23_12_36_11.npy user69_2017_10_23_12_36_28.npy user69_2017_10_23_12_37_01.npy user69_2017_10_23_12_37_51.npy	user69_2017_10_23_12_36_45.npy user69_2017_10_23_12_37_18.npy user69_2017_10_23_12_37_35.npy
68	user70	user70_2017_10_23_12_41_43.npy user70_2017_10_23_12_42_17.npy user70_2017_10_23_12_42_50.npy user70_2017_10_23_12_43_07.npy user70_2017_10_23_12_43_23.npy user70_2017_10_23_12_43_39.npy user70_2017_10_23_12_44_14.npy user70_2017_10_23_12_44_48.npy user70_2017_10_23_12_45_07.npy user70_2017_10_23_12_45_38.npy	user70_2017_10_23_12_42_01.npy user70_2017_10_23_12_42_33.npy user70_2017_10_23_12_43_57.npy user70_2017_10_23_12_44_30.npy
69	user71	user71_2017_10_23_12_46_32.npy user71_2017_10_23_12_47_05.npy user71_2017_10_23_12_47_23.npy user71_2017_10_23_12_48_12.npy user71_2017_10_23_12_49_04.npy user71_2017_10_23_12_49_21.npy	user71_2017_10_23_12_46_49.npy user71_2017_10_23_12_47_56.npy
70	user72	user72_2017_10_23_12_50_03.npy user72_2017_10_23_12_50_20.npy user72_2017_10_23_12_50_52.npy user72_2017_10_23_12_51_14.npy user72_2017_10_23_12_51_30.npy	user72_2017_10_23_12_50_36.npy user72_2017_10_23_12_51_47.npy user72_2017_10_23_12_52_37.npy

		user72_2017_10_23_12_52_03.npy user72_2017_10_23_12_52_20.npy	
71	user73	user73_2017_10_25_09_06_51.npy user73_2017_10_25_09_07_05.npy user73_2017_10_25_09_07_32.npy user73_2017_10_25_09_07_45.npy user73_2017_10_25_09_08_13.npy user73_2017_10_25_09_08_26.npy user73_2017_10_25_09_08_57.npy user73_2017_11_22_09_07_12.npy user73_2017_11_22_09_07_49.npy user73_2017_11_22_09_08_22.npy user73_2017_11_22_09_08_36.npy user73_2017_11_22_09_08_51.npy user73_2017_11_22_09_09_05.npy user73_2017_11_22_09_09_20.npy user73_2017_11_22_09_09_49.npy user73_2018_01_24_16_05_36.npy user73_2018_01_24_16_05_50.npy user73_2018_01_24_16_06_03.npy user73_2018_01_24_16_06_17.npy user73_2018_01_24_16_06_30.npy user73_2018_01_24_16_06_43.npy user73_2019_12_20_10_39_27.npy user73_2019_12_20_10_40_00.npy user73_2019_12_20_10_40_49.npy user73_2019_12_20_10_41_05.npy user73_2019_12_20_10_41_21.npy	user73_2017_10_25_09_07_19.npy user73_2017_10_25_09_08_40.npy user73_2017_11_22_09_05_23.npy user73_2017_11_22_09_05_38.npy user73_2017_11_22_09_08_07.npy user73_2017_11_22_09_10_04.npy user73_2019_12_20_10_39_06.npy user73_2019_12_20_10_39_43.npy user73_2019_12_20_10_40_16.npy user73_2019_12_20_10_40_33.npy user73_2019_12_20_10_41_38.npy
72	user74	user74_2017_10_25_09_10_23.npy user74_2017_10_25_09_10_50.npy user74_2017_10_25_09_11_05.npy user74_2017_10_25_09_12_23.npy user74_2017_11_22_09_27_02.npy user74_2017_11_22_09_27_17.npy user74_2017_11_22_09_27_46.npy user74_2017_11_22_09_28_01.npy user74_2017_11_22_09_28_15.npy user74_2017_11_22_09_29_01.npy user74_2017_11_22_09_29_16.npy user74_2017_11_22_09_29_30.npy user74_2017_11_22_09_30_17.npy user74_2017_11_22_09_30_31.npy user74_2017_11_22_09_30_46.npy user74_2017_11_22_09_31_01.npy user74_2017_11_22_09_31_16.npy user74_2018_01_24_16_03_50.npy user74_2018_01_24_16_04_30.npy user74_2018_01_24_16_04_43.npy user74_2018_01_24_16_04_57.npy user74_2019_12_20_10_21_44.npy user74_2019_12_20_10_22_07.npy user74_2019_12_20_10_22_39.npy user74_2019_12_20_10_22_56.npy user74_2019_12_20_10_23_13.npy user74_2019_12_20_10_23_31.npy user74_2019_12_20_10_23_50.npy user74_2019_12_20_10_24_06.npy	user74_2017_10_25_09_11_20.npy user74_2017_10_25_09_11_35.npy user74_2017_10_25_09_12_06.npy user74_2017_10_25_09_12_54.npy user74_2017_11_22_09_27_31.npy user74_2017_11_22_09_28_46.npy user74_2017_11_22_09_29_46.npy user74_2017_11_22_09_30_01.npy user74_2018_01_24_16_04_03.npy user74_2018_01_24_16_04_17.npy user74_2019_12_20_10_21_22.npy user74_2019_12_20_10_22_22.npy
73	user75	user75_2017_10_25_09_14_25.npy user75_2017_10_25_09_14_40.npy	user75_2017_10_25_09_13_58.npy user75_2017_10_25_09_14_54.npy

		user75_2017_10_25_09_15_08.npy user75_2017_10_25_09_15_22.npy user75_2017_10_25_09_15_50.npy user75_2017_10_25_09_16_04.npy user75_2017_10_25_09_16_17.npy user75_2017_11_22_09_14_55.npy user75_2017_11_22_09_15_10.npy user75_2017_11_22_09_16_01.npy user75_2017_11_22_09_16_17.npy user75_2017_11_22_09_16_32.npy user75_2017_11_22_09_16_48.npy user75_2017_11_22_09_17_03.npy user75_2017_11_22_09_17_18.npy user75_2019_12_20_11_21_30.npy user75_2019_12_20_11_21_48.npy user75_2019_12_20_11_22_04.npy user75_2019_12_20_11_22_21.npy user75_2019_12_20_11_22_39.npy user75_2019_12_20_11_22_56.npy	user75_2017_11_22_09_15_25.npy user75_2017_11_22_09_15_41.npy user75_2017_11_22_09_17_33.npy user75_2019_12_20_11_23_12.npy user75_2019_12_20_11_23_29.npy user75_2019_12_20_11_23_44.npy user75_2019_12_20_11_24_00.npy
74	user76	user76_2017_10_25_09_18_46.npy user76_2017_10_25_09_19_01.npy user76_2017_10_25_09_19_14.npy user76_2017_10_25_09_19_28.npy user76_2017_10_25_09_19_41.npy user76_2017_10_25_09_19_55.npy user76_2017_10_25_09_20_08.npy user76_2017_10_25_09_20_23.npy user76_2017_10_25_09_20_37.npy user76_2017_10_25_09_20_50.npy user76_2017_11_22_09_18_56.npy user76_2017_11_22_09_19_27.npy user76_2017_11_22_09_20_12.npy user76_2017_11_22_09_20_27.npy user76_2017_11_22_09_20_41.npy user76_2017_11_22_09_20_55.npy user76_2018_01_24_16_02_26.npy user76_2018_01_24_16_02_39.npy user76_2018_01_24_16_03_06.npy user76_2018_01_24_16_03_19.npy	user76_2017_10_25_09_21_05.npy user76_2017_11_22_09_18_26.npy user76_2017_11_22_09_18_41.npy user76_2017_11_22_09_19_12.npy user76_2017_11_22_09_19_42.npy user76_2017_11_22_09_19_57.npy user76_2018_01_24_16_02_12.npy user76_2018_01_24_16_02_52.npy
75	user77	user77_2017_10_25_09_22_23.npy user77_2017_10_25_09_22_37.npy user77_2017_10_25_09_22_51.npy user77_2017_10_25_09_23_32.npy user77_2017_10_25_09_23_47.npy user77_2017_10_25_09_24_14.npy user77_2017_10_25_09_24_27.npy user77_2017_10_25_09_24_43.npy	user77_2017_10_25_09_23_04.npy user77_2017_10_25_09_23_18.npy user77_2017_10_25_09_24_00.npy
76	user78	user78_2017_10_25_09_25_58.npy user78_2017_10_25_09_26_11.npy user78_2017_10_25_09_26_25.npy user78_2017_10_25_09_26_54.npy user78_2017_10_25_09_27_22.npy user78_2017_10_25_09_27_35.npy user78_2017_10_25_09_27_49.npy user78_2017_10_25_09_28_02.npy user78_2017_11_22_09_11_40.npy user78_2017_11_22_09_11_55.npy user78_2017_11_22_09_12_10.npy user78_2017_11_22_09_12_39.npy	user78_2017_10_25_09_26_38.npy user78_2017_10_25_09_27_08.npy user78_2017_10_25_09_28_15.npy user78_2017_11_22_09_12_24.npy user78_2017_11_22_09_13_25.npy user78_2017_11_22_09_13_56.npy user78_2019_12_20_11_52_08.npy user78_2019_12_20_11_53_17.npy user78_2019_12_20_11_54_09.npy user78_2019_12_20_11_54_27.npy user78_2019_12_20_11_55_17.npy

		user78_2017_11_22_09_12_54.npy user78_2017_11_22_09_13_09.npy user78_2017_11_22_09_13_41.npy user78_2017_11_22_09_14_13.npy user78_2018_01_24_16_09_38.npy user78_2018_01_24_16_09_51.npy user78_2018_01_24_16_10_04.npy user78_2018_01_24_16_10_18.npy user78_2019_12_20_11_52_25.npy user78_2019_12_20_11_52_43.npy user78_2019_12_20_11_53_00.npy user78_2019_12_20_11_53_34.npy user78_2019_12_20_11_53_50.npy user78_2019_12_20_11_54_44.npy user78_2019_12_20_11_55_01.npy	
77	user79	user79_2017_10_25_09_29_31.npy user79_2017_10_25_09_29_46.npy user79_2017_10_25_09_30_01.npy user79_2017_10_25_09_30_15.npy user79_2017_10_25_09_30_56.npy user79_2017_10_25_09_31_09.npy	user79_2017_10_25_09_30_42.npy user79_2017_10_25_09_31_22.npy
78	user80	user80_2017_10_25_09_32_39.npy user80_2017_10_25_09_32_52.npy user80_2017_10_25_09_33_20.npy user80_2017_10_25_09_33_33.npy user80_2017_10_25_09_34_00.npy user80_2017_10_25_09_34_13.npy user80_2017_10_25_09_34_26.npy user80_2017_10_25_09_34_40.npy user80_2017_10_25_09_34_56.npy user80_2017_10_25_09_35_23.npy user80_2017_10_25_09_35_36.npy user80_2017_10_25_09_36_31.npy user80_2017_10_25_09_36_44.npy user80_2017_10_25_09_36_57.npy user80_2017_10_25_09_37_10.npy user80_2017_11_22_08_47_25.npy user80_2017_11_22_08_47_40.npy user80_2017_11_22_08_48_10.npy user80_2017_11_22_08_48_25.npy user80_2017_11_22_08_48_40.npy user80_2017_11_22_08_48_58.npy user80_2017_11_22_08_49_32.npy user80_2017_11_22_08_49_48.npy user80_2017_11_22_08_50_09.npy user80_2017_11_22_08_50_39.npy user80_2017_11_22_08_50_55.npy user80_2017_11_22_08_51_12.npy user80_2017_11_22_08_51_30.npy user80_2017_11_22_08_52_16.npy user80_2017_11_22_08_53_24.npy user80_2018_01_24_16_12_38.npy user80_2018_01_24_16_12_51.npy user80_2018_01_24_16_13_04.npy user80_2018_01_24_16_13_18.npy user80_2018_01_24_16_13_31.npy user80_2018_01_24_16_13_44.npy user80_2018_01_24_16_13_57.npy user80_2019_12_20_10_43_06.npy user80_2019_12_20_10_43_26.npy	user80_2017_10_25_09_32_25.npy user80_2017_10_25_09_33_46.npy user80_2017_10_25_09_35_09.npy user80_2017_10_25_09_35_50.npy user80_2017_10_25_09_36_04.npy user80_2017_10_25_09_36_17.npy user80_2017_11_22_08_47_55.npy user80_2017_11_22_08_49_16.npy user80_2017_11_22_08_51_45.npy user80_2017_11_22_08_52_00.npy user80_2017_11_22_08_52_33.npy user80_2017_11_22_08_52_48.npy user80_2017_11_22_08_53_03.npy user80_2019_12_20_10_42_50.npy user80_2019_12_20_10_43_42.npy user80_2019_12_20_10_43_58.npy user80_2019_12_20_10_44_30.npy

		user80_2019_12_20_10_44_46.npy	
79	user81	user81_2017_10_25_09_38_18.npy user81_2017_10_25_09_38_32.npy user81_2017_10_25_09_38_45.npy user81_2017_10_25_09_38_59.npy user81_2017_10_25_09_39_12.npy user81_2017_10_25_09_39_25.npy user81_2017_10_25_09_39_52.npy user81_2017_10_25_09_40_06.npy user81_2017_11_22_08_55_13.npy user81_2017_11_22_08_55_28.npy user81_2017_11_22_08_56_01.npy user81_2017_11_22_08_56_17.npy user81_2017_11_22_08_56_32.npy user81_2017_11_22_08_56_47.npy user81_2018_01_24_16_16_32.npy user81_2018_01_24_16_16_46.npy user81_2018_01_24_16_16_59.npy user81_2018_01_24_16_17_26.npy user81_2018_01_24_16_17_39.npy user81_2018_01_24_16_17_52.npy user81_2018_01_24_16_18_06.npy user81_2019_12_20_10_45_36.npy user81_2019_12_20_10_45_54.npy user81_2019_12_20_10_46_15.npy user81_2019_12_20_10_47_12.npy user81_2019_12_20_10_47_27.npy user81_2019_12_20_10_47_43.npy	user81_2017_10_25_09_38_04.npy user81_2017_10_25_09_39_38.npy user81_2017_10_25_09_40_19.npy user81_2017_11_22_08_54_13.npy user81_2017_11_22_08_54_28.npy user81_2017_11_22_08_54_43.npy user81_2017_11_22_08_54_58.npy user81_2017_11_22_08_55_42.npy user81_2018_01_24_16_17_12.npy user81_2019_12_20_10_46_31.npy user81_2019_12_20_10_47_59.npy user81_2019_12_20_10_48_15.npy
80	user82	user82_2017_10_25_09_41_21.npy user82_2017_10_25_09_41_47.npy user82_2017_10_25_09_42_13.npy user82_2017_10_25_09_42_27.npy user82_2017_10_25_09_42_40.npy user82_2017_10_25_09_42_53.npy user82_2017_10_25_09_43_31.npy user82_2017_11_22_09_39_01.npy user82_2017_11_22_09_39_31.npy user82_2017_11_22_09_39_46.npy user82_2017_11_22_09_40_00.npy user82_2017_11_22_09_40_30.npy user82_2017_11_22_09_40_45.npy user82_2017_11_22_09_41_01.npy user82_2017_11_22_09_41_17.npy user82_2018_01_24_16_14_39.npy user82_2018_01_24_16_14_52.npy user82_2018_01_24_16_15_06.npy user82_2018_01_24_16_15_19.npy user82_2018_01_24_16_15_46.npy user82_2018_01_24_16_15_59.npy	user82_2017_10_25_09_41_34.npy user82_2017_10_25_09_42_00.npy user82_2017_10_25_09_43_06.npy user82_2017_10_25_09_43_19.npy user82_2017_11_22_09_39_16.npy user82_2017_11_22_09_40_15.npy user82_2017_11_22_09_41_33.npy user82_2018_01_24_16_14_26.npy user82_2018_01_24_16_15_32.npy
81	user83	user83_2017_10_25_09_44_56.npy user83_2017_10_25_09_45_11.npy user83_2017_10_25_09_45_24.npy user83_2017_10_25_09_45_37.npy user83_2017_10_25_09_45_51.npy user83_2017_10_25_09_46_04.npy user83_2017_10_25_09_46_17.npy user83_2017_10_25_09_46_30.npy user83_2017_10_25_09_46_43.npy user83_2017_10_25_09_47_17.npy	user83_2017_10_25_09_44_43.npy user83_2017_10_25_09_47_58.npy user83_2017_10_25_09_48_13.npy user83_2017_10_25_09_48_47.npy user83_2017_10_25_09_49_01.npy user83_2017_10_25_09_49_27.npy

		user83_2017_10_25_09_47_31.npy user83_2017_10_25_09_47_44.npy user83_2017_10_25_09_48_32.npy user83_2017_10_25_09_49_13.npy user83_2017_10_25_09_49_40.npy	
82	user84	user84_2017_10_25_09_50_40.npy user84_2017_10_25_09_50_54.npy user84_2017_10_25_09_51_07.npy user84_2017_10_25_09_51_33.npy user84_2017_10_25_09_51_50.npy user84_2017_10_25_09_52_02.npy user84_2017_10_25_09_52_16.npy user84_2017_10_25_09_52_30.npy user84_2017_10_25_09_53_02.npy user84_2018_01_24_15_54_07.npy user84_2018_01_24_15_54_37.npy user84_2018_01_24_15_54_51.npy user84_2018_01_24_15_55_06.npy user84_2018_01_24_15_55_20.npy user84_2019_12_20_10_25_27.npy user84_2019_12_20_10_25_44.npy user84_2019_12_20_10_26_37.npy user84_2019_12_20_10_27_13.npy user84_2019_12_20_10_27_50.npy	user84_2017_10_25_09_51_20.npy user84_2017_10_25_09_52_49.npy user84_2018_01_24_15_53_52.npy user84_2018_01_24_15_54_22.npy user84_2019_12_20_10_25_08.npy user84_2019_12_20_10_26_18.npy user84_2019_12_20_10_26_57.npy user84_2019_12_20_10_27_31.npy
83	user85	user85_2017_11_22_08_58_33.npy user85_2017_11_22_08_58_49.npy user85_2017_11_22_08_59_05.npy user85_2017_11_22_08_59_36.npy user85_2017_11_22_08_59_51.npy user85_2017_11_22_09_00_06.npy user85_2017_11_22_09_00_21.npy user85_2017_11_22_09_00_36.npy user85_2018_01_24_15_58_13.npy user85_2018_01_24_15_58_27.npy user85_2018_01_24_15_58_42.npy user85_2018_01_24_15_58_57.npy user85_2018_01_24_15_59_12.npy user85_2018_01_24_15_59_27.npy user85_2018_01_24_15_59_42.npy user85_2019_12_20_10_28_53.npy user85_2019_12_20_10_29_10.npy user85_2019_12_20_10_29_26.npy user85_2019_12_20_10_29_59.npy user85_2019_12_20_10_30_49.npy	user85_2017_11_22_08_58_00.npy user85_2017_11_22_08_58_17.npy user85_2017_11_22_08_59_20.npy user85_2019_12_20_10_29_43.npy user85_2019_12_20_10_30_16.npy user85_2019_12_20_10_30_33.npy user85_2019_12_20_10_31_06.npy user85_2019_12_20_10_31_22.npy
84	user86	user86_2017_11_22_09_01_42.npy user86_2017_11_22_09_01_59.npy user86_2017_11_22_09_02_15.npy user86_2017_11_22_09_02_48.npy user86_2017_11_22_09_03_18.npy user86_2017_11_22_09_04_02.npy user86_2017_11_22_09_04_17.npy user86_2018_01_24_16_00_33.npy user86_2018_01_24_16_00_47.npy user86_2018_01_24_16_01_00.npy user86_2018_01_24_16_01_27.npy user86_2018_01_24_16_01_40.npy user86_2019_12_20_10_32_11.npy user86_2019_12_20_10_32_45.npy user86_2019_12_20_10_33_02.npy	user86_2017_11_22_09_02_32.npy user86_2017_11_22_09_03_03.npy user86_2017_11_22_09_03_33.npy user86_2017_11_22_09_03_47.npy user86_2018_01_24_16_01_13.npy user86_2019_12_20_10_32_29.npy user86_2019_12_20_10_33_52.npy user86_2019_12_20_10_34_08.npy

		user86_2019_12_20_10_33_18.npy user86_2019_12_20_10_33_36.npy user86_2019_12_20_10_34_24.npy user86_2019_12_20_10_34_40.npy	
85	user87	user87_2017_11_22_09_22_10.npy user87_2017_11_22_09_22_25.npy user87_2017_11_22_09_22_40.npy user87_2017_11_22_09_22_55.npy user87_2017_11_22_09_23_10.npy user87_2017_11_22_09_23_25.npy user87_2017_11_22_09_23_40.npy user87_2017_11_22_09_23_54.npy user87_2017_11_22_09_24_24.npy user87_2017_11_22_09_25_15.npy user87_2017_11_22_09_25_44.npy user87_2017_11_22_09_26_13.npy user87_2018_01_24_15_55_56.npy user87_2018_01_24_15_56_25.npy user87_2018_01_24_15_56_40.npy user87_2018_01_24_15_57_09.npy user87_2018_01_24_15_57_38.npy user87_2019_12_20_10_35_38.npy user87_2019_12_20_10_35_55.npy user87_2019_12_20_10_36_27.npy user87_2019_12_20_10_37_14.npy user87_2019_12_20_10_37_31.npy user87_2019_12_20_10_37_47.npy user87_2019_12_20_10_38_02.npy	user87_2017_11_22_09_24_09.npy user87_2017_11_22_09_24_38.npy user87_2017_11_22_09_25_00.npy user87_2017_11_22_09_25_29.npy user87_2017_11_22_09_25_58.npy user87_2018_01_24_15_56_10.npy user87_2018_01_24_15_56_54.npy user87_2018_01_24_15_57_24.npy user87_2019_12_20_10_36_11.npy user87_2019_12_20_10_36_42.npy user87_2019_12_20_10_36_58.npy
86	user88	user88_2017_11_22_09_32_23.npy user88_2017_11_22_09_32_38.npy user88_2017_11_22_09_33_09.npy user88_2017_11_22_09_33_25.npy user88_2017_11_22_09_34_10.npy user88_2018_01_24_15_51_02.npy user88_2018_01_24_15_51_18.npy user88_2018_01_24_15_51_32.npy user88_2018_01_24_15_51_47.npy user88_2018_01_24_15_52_01.npy user88_2018_01_24_15_52_16.npy user88_2018_01_24_15_52_31.npy user88_2018_01_24_15_52_45.npy user88_2018_01_24_15_53_00.npy user88_2019_12_20_10_57_21.npy user88_2019_12_20_10_57_38.npy user88_2019_12_20_10_58_12.npy user88_2019_12_20_10_58_29.npy user88_2019_12_20_10_59_17.npy user88_2019_12_20_10_59_35.npy	user88_2017_11_22_09_32_08.npy user88_2017_11_22_09_32_52.npy user88_2017_11_22_09_33_40.npy user88_2017_11_22_09_33_55.npy user88_2017_11_22_09_34_25.npy user88_2017_11_22_09_34_41.npy user88_2019_12_20_10_57_04.npy user88_2019_12_20_10_58_45.npy user88_2019_12_20_10_59_01.npy
87	user89	user89_2017_11_22_09_35_19.npy user89_2017_11_22_09_35_49.npy user89_2017_11_22_09_36_07.npy user89_2017_11_22_09_36_23.npy user89_2017_11_22_09_36_37.npy user89_2017_11_22_09_36_54.npy user89_2017_11_22_09_37_10.npy user89_2017_11_22_09_37_55.npy	user89_2017_11_22_09_35_34.npy user89_2017_11_22_09_37_24.npy user89_2017_11_22_09_37_40.npy
88	user90	user90_2017_11_22_10_32_21.npy user90_2017_11_22_10_32_41.npy user90_2017_11_22_10_33_28.npy	user90_2017_11_22_10_31_58.npy user90_2017_11_22_10_32_56.npy user90_2017_11_22_10_33_12.npy

		user90_2017_11_22_10_33_43.npy user90_2017_11_22_10_33_58.npy user90_2017_11_22_10_34_14.npy user90_2017_11_22_10_34_29.npy user90_2017_11_22_10_34_45.npy	
89	user92	user92_2017_11_22_10_41_19.npy user92_2017_11_22_10_42_02.npy user92_2017_11_22_10_42_15.npy user92_2017_11_22_10_42_30.npy user92_2017_11_22_10_42_58.npy user92_2017_11_22_10_43_28.npy user92_2017_11_22_10_43_42.npy	user92_2017_11_22_10_41_48.npy user92_2017_11_22_10_42_43.npy user92_2017_11_22_10_43_15.npy
90	user93	user93_2017_11_22_10_47_33.npy user93_2017_11_22_10_48_14.npy user93_2017_11_22_10_49_08.npy user93_2017_11_22_10_49_21.npy	user93_2017_11_22_10_48_27.npy
91	user94	user94_2017_12_19_17_01_50.npy user94_2017_12_19_17_02_06.npy user94_2017_12_19_17_02_38.npy user94_2017_12_19_17_02_52.npy user94_2017_12_19_17_03_07.npy user94_2017_12_19_17_03_37.npy user94_2017_12_19_17_04_00.npy user94_2017_12_19_17_04_16.npy	user94_2017_12_19_17_02_23.npy user94_2017_12_19_17_03_22.npy user94_2017_12_19_17_04_31.npy
92	user95	user95_2017_12_26_10_36_57.npy user95_2017_12_26_10_37_12.npy user95_2017_12_26_10_37_27.npy	user95_2017_12_26_10_37_41.npy
93	user96	user96_2018_01_24_16_10_44.npy user96_2018_01_24_16_11_24.npy user96_2018_01_24_16_11_38.npy user96_2018_01_24_16_12_04.npy user96_2019_12_20_10_50_14.npy user96_2019_12_20_10_50_32.npy user96_2019_12_20_10_50_50.npy user96_2019_12_20_10_51_10.npy user96_2019_12_20_10_51_26.npy user96_2019_12_20_10_51_42.npy user96_2019_12_20_10_52_32.npy	user96_2018_01_24_16_10_58.npy user96_2018_01_24_16_11_11.npy user96_2018_01_24_16_11_51.npy user96_2019_12_20_10_52_00.npy user96_2019_12_20_10_52_16.npy
94	user97	user97_2018_02_12_17_26_00.npy user97_2018_02_12_17_26_15.npy user97_2018_02_12_17_26_45.npy user97_2018_02_12_17_27_00.npy user97_2018_02_12_17_27_46.npy user97_2018_02_12_17_28_25.npy user97_2018_02_12_17_28_41.npy user97_2019_12_27_11_48_12.npy user97_2019_12_27_11_48_30.npy user97_2019_12_27_11_50_14.npy user97_2019_12_27_11_51_33.npy user97_2019_12_27_11_51_51.npy user97_2019_12_27_11_52_25.npy user97_2019_12_27_11_52_58.npy user97_2019_12_27_11_53_15.npy	user97_2018_02_12_17_26_29.npy user97_2018_02_12_17_27_32.npy user97_2018_02_12_17_28_01.npy user97_2019_12_27_11_48_58.npy user97_2019_12_27_11_49_15.npy user97_2019_12_27_11_49_33.npy user97_2019_12_27_11_49_57.npy
95	user98	user98_2018_02_12_17_30_06.npy user98_2018_02_12_17_30_54.npy user98_2018_02_12_17_31_25.npy user98_2018_02_12_17_32_10.npy user98_2018_02_12_17_32_29.npy user98_2018_02_12_17_32_45.npy	user98_2018_02_12_17_30_39.npy user98_2018_02_12_17_31_09.npy user98_2018_02_12_17_31_55.npy

		user98_2018_02_12_17_33_01.npy	
96	user99	user99_2018_02_12_17_33_58.npy user99_2018_02_12_17_34_29.npy user99_2018_02_12_17_35_21.npy user99_2018_02_12_17_35_36.npy user99_2018_02_12_17_35_52.npy user99_2018_02_12_17_36_41.npy user99_2019_12_27_14_54_19.npy user99_2019_12_27_14_54_37.npy user99_2019_12_27_14_54_55.npy user99_2019_12_27_14_55_14.npy user99_2019_12_27_14_55_31.npy user99_2019_12_27_14_55_48.npy user99_2019_12_27_14_56_22.npy user99_2019_12_27_14_56_55.npy user99_2019_12_27_14_57_28.npy user99_2019_12_27_14_57_45.npy user99_2019_12_27_14_58_01.npy user99_2019_12_27_14_58_17.npy	user99_2018_02_12_17_34_13.npy user99_2018_02_12_17_34_50.npy user99_2018_02_12_17_35_05.npy user99_2018_02_12_17_36_26.npy user99_2018_02_12_17_36_56.npy user99_2019_12_27_14_56_04.npy user99_2019_12_27_14_56_38.npy user99_2019_12_27_14_57_12.npy
97	user100	user100_2018_02_12_17_39_34.npy user100_2018_02_12_17_39_49.npy user100_2018_02_12_17_40_04.npy user100_2018_02_12_17_40_34.npy user100_2018_02_12_17_41_04.npy user100_2018_02_12_17_41_19.npy user100_2018_02_12_17_41_53.npy	user100_2018_02_12_17_40_19.npy user100_2018_02_12_17_40_49.npy user100_2018_02_12_17_41_37.npy
98	user101	user101_2019_12_20_10_53_45.npy user101_2019_12_20_10_54_01.npy user101_2019_12_20_10_54_18.npy user101_2019_12_20_10_55_09.npy user101_2019_12_20_10_55_42.npy user101_2019_12_20_10_55_57.npy user101_2019_12_20_10_56_14.npy	user101_2019_12_20_10_54_34.npy user101_2019_12_20_10_54_52.npy user101_2019_12_20_10_55_26.npy
99	user102	user102_2019_12_20_11_09_43.npy user102_2019_12_20_11_09_59.npy user102_2019_12_20_11_10_32.npy user102_2019_12_20_11_10_48.npy user102_2019_12_20_11_11_04.npy user102_2019_12_20_11_11_52.npy user102_2019_12_20_11_12_08.npy	user102_2019_12_20_11_10_16.npy user102_2019_12_20_11_11_19.npy user102_2019_12_20_11_11_35.npy
100	user103	user103_2019_12_20_11_14_13.npy user103_2019_12_20_11_14_30.npy user103_2019_12_20_11_15_03.npy user103_2019_12_20_11_15_37.npy user103_2019_12_20_11_15_53.npy user103_2019_12_20_11_16_09.npy	user103_2019_12_20_11_14_46.npy user103_2019_12_20_11_15_21.npy
101	user104	user104_2019_12_20_11_25_07.npy user104_2019_12_20_11_25_38.npy user104_2019_12_20_11_25_54.npy user104_2019_12_20_11_26_25.npy user104_2019_12_20_11_26_42.npy user104_2019_12_20_11_26_58.npy user104_2019_12_20_11_27_15.npy	user104_2019_12_20_11_24_51.npy user104_2019_12_20_11_25_23.npy user104_2019_12_20_11_26_10.npy
102	user105	user105_2019_12_20_11_28_17.npy user105_2019_12_20_11_28_33.npy user105_2019_12_20_11_29_09.npy user105_2019_12_20_11_29_42.npy user105_2019_12_20_11_29_58.npy user105_2019_12_20_11_30_15.npy	user105_2019_12_20_11_28_49.npy user105_2019_12_20_11_30_38.npy user105_2019_12_20_11_30_55.npy

		user105_2019_12_20_11_31_11.npy	
103	user106	user106_2019_12_20_11_32_04.npy user106_2019_12_20_11_32_36.npy user106_2019_12_20_11_32_52.npy user106_2019_12_20_11_33_07.npy user106_2019_12_20_11_33_23.npy user106_2019_12_20_11_33_39.npy user106_2019_12_20_11_34_11.npy	user106_2019_12_20_11_32_20.npy user106_2019_12_20_11_33_55.npy user106_2019_12_20_11_34_26.npy
104	user107	user107_2019_12_20_11_35_51.npy user107_2019_12_20_11_36_08.npy user107_2019_12_20_11_36_24.npy user107_2019_12_20_11_36_56.npy user107_2019_12_20_11_37_12.npy user107_2019_12_20_11_37_29.npy user107_2019_12_20_11_37_45.npy	user107_2019_12_20_11_35_18.npy user107_2019_12_20_11_35_34.npy user107_2019_12_20_11_36_40.npy
105	user108	user108_2019_12_20_11_48_41.npy user108_2019_12_20_11_49_13.npy user108_2019_12_20_11_49_30.npy user108_2019_12_20_11_50_06.npy user108_2019_12_20_11_50_22.npy user108_2019_12_20_11_50_38.npy user108_2019_12_20_11_50_53.npy	user108_2019_12_20_11_48_57.npy user108_2019_12_20_11_49_49.npy user108_2019_12_20_11_51_12.npy
106	user109	user109_2019_12_20_11_56_17.npy user109_2019_12_20_11_56_50.npy user109_2019_12_20_11_57_09.npy user109_2019_12_20_11_57_24.npy user109_2019_12_20_11_58_12.npy user109_2019_12_20_11_58_28.npy user109_2019_12_20_11_58_44.npy user109_2019_12_20_11_59_00.npy	user109_2019_12_20_11_56_34.npy user109_2019_12_20_11_57_40.npy user109_2019_12_20_11_57_56.npy
107	user110	user110_2019_12_20_11_59_56.npy user110_2019_12_20_12_00_13.npy user110_2019_12_20_12_00_30.npy user110_2019_12_20_12_00_47.npy user110_2019_12_20_12_01_37.npy user110_2019_12_20_12_02_26.npy	user110_2019_12_20_12_01_20.npy user110_2019_12_20_12_01_53.npy user110_2019_12_20_12_02_10.npy
108	user111	user111_2019_12_20_12_03_34.npy user111_2019_12_20_12_03_50.npy user111_2019_12_20_12_04_06.npy user111_2019_12_20_12_04_38.npy user111_2019_12_20_12_05_25.npy user111_2019_12_20_12_05_52.npy	user111_2019_12_20_12_04_22.npy user111_2019_12_20_12_04_53.npy user111_2019_12_20_12_05_09.npy
109	user112	user112_2019_12_20_12_07_02.npy user112_2019_12_20_12_07_18.npy user112_2019_12_20_12_07_34.npy user112_2019_12_20_12_08_07.npy user112_2019_12_20_12_08_23.npy user112_2019_12_20_12_08_40.npy user112_2019_12_20_12_08_58.npy	user112_2019_12_20_12_07_50.npy user112_2019_12_20_12_09_15.npy user112_2019_12_20_12_09_32.npy
110	user113	user113_2019_12_20_12_10_35.npy user113_2019_12_20_12_10_51.npy user113_2019_12_20_12_11_08.npy user113_2019_12_20_12_11_42.npy user113_2019_12_20_12_12_06.npy user113_2019_12_20_12_12_42.npy user113_2019_12_20_12_12_59.npy	user113_2019_12_20_12_10_18.npy user113_2019_12_20_12_11_24.npy user113_2019_12_20_12_12_25.npy
111	user114	user114_2019_12_20_12_13_55.npy user114_2019_12_20_12_14_12.npy	user114_2019_12_20_12_14_29.npy user114_2019_12_20_12_15_21.npy

		user114_2019_12_20_12_14_45.npy user114_2019_12_20_12_15_04.npy user114_2019_12_20_12_15_56.npy user114_2019_12_20_12_16_13.npy user114_2019_12_20_12_16_29.npy	user114_2019_12_20_12_15_40.npy
112	user115	user115_2019_12_27_11_33_56.npy user115_2019_12_27_11_34_17.npy user115_2019_12_27_11_34_36.npy user115_2019_12_27_11_35_07.npy user115_2019_12_27_11_35_24.npy user115_2019_12_27_11_35_40.npy user115_2019_12_27_11_36_14.npy	user115_2019_12_27_11_35_57.npy user115_2019_12_27_11_36_49.npy user115_2019_12_27_11_37_06.npy
113	user116	user116_2019_12_27_11_43_16.npy user116_2019_12_27_11_43_53.npy user116_2019_12_27_11_44_28.npy user116_2019_12_27_11_44_45.npy user116_2019_12_27_11_45_03.npy user116_2019_12_27_11_46_01.npy user116_2019_12_27_11_46_38.npy	user116_2019_12_27_11_43_35.npy user116_2019_12_27_11_44_11.npy user116_2019_12_27_11_45_39.npy
114	user117	user117_2019_12_27_11_38_17.npy user117_2019_12_27_11_38_56.npy user117_2019_12_27_11_39_39.npy user117_2019_12_27_11_39_56.npy user117_2019_12_27_11_40_13.npy user117_2019_12_27_11_40_38.npy user117_2019_12_27_11_41_40.npy user117_2019_12_27_11_42_26.npy	user117_2019_12_27_11_37_56.npy user117_2019_12_27_11_38_38.npy user117_2019_12_27_11_39_21.npy
115	user118	user118_2019_12_27_15_42_45.npy user118_2019_12_27_15_43_33.npy user118_2019_12_27_15_44_10.npy user118_2019_12_27_15_44_28.npy user118_2019_12_27_15_45_04.npy user118_2019_12_27_15_45_22.npy user118_2019_12_27_15_46_04.npy user118_2019_12_27_15_46_24.npy user118_2019_12_27_15_46_42.npy user118_2019_12_27_15_47_22.npy user118_2019_12_27_15_47_41.npy	user118_2019_12_27_15_43_05.npy user118_2019_12_27_15_44_46.npy user118_2019_12_27_15_47_00.npy user118_2019_12_27_15_47_59.npy

Додаток Г

Конфігурація експериментів для дослідження масштабованості біометричної системи автентифікації.

Кількість персон	Список персон
115	user1, user2, user3, user4, user5, user6, user7, user8, user9, user10, user11, user12, user13, user14, user15, user16, user17, user18, user20, user21, user22, user23, user24, user25, user26, user27, user28, user29, user30, user31, user32, user33, user34, user35, user36, user37, user38, user39, user40, user41, user42, user43, user44, user45, user46, user47, user48, user49, user50, user51, user52, user53, user54, user55, user56, user58, user59, user60, user61, user62, user63, user64, user65, user66, user67, user68, user69, user70, user71, user72, user73, user74, user75, user76, user77, user78, user79, user80, user81, user82, user83, user84, user85, user86, user87, user88, user89, user90, user92, user93, user94, user95, user96, user97, user98, user99, user100, user101, user102, user103, user104, user105, user106, user107, user108, user109, user110, user111, user112, user113, user114, user115, user116, user117, user118
110	user1, user3, user4, user5, user6, user7, user8, user9, user10, user11, user12, user13, user15, user16, user17, user18, user20, user21, user22, user23, user24, user25, user26, user27, user28, user29, user30, user31, user32, user33, user34, user36, user37, user38, user39, user40, user41, user42, user43, user44, user45, user46, user47, user48, user49, user50, user51, user52, user53, user54, user55, user56, user58, user59, user60, user62, user63, user64, user65, user66, user67, user68, user69, user70, user71, user72, user73, user74, user75, user76, user77, user78, user79, user80, user81, user82, user83, user84, user85, user86, user87, user88, user89, user90, user92, user93, user94, user95, user96, user97, user98, user99, user100, user101, user102, user103, user104, user105, user106, user107, user108, user109, user110, user111, user112, user113, user114, user115, user117, user118
100	user1, user3, user5, user6, user7, user9, user10, user11, user12, user13, user16, user17, user18, user20, user21, user22, user23, user24, user25, user26, user27, user28, user29, user30, user31, user32, user33, user34, user37, user38, user39, user40, user41, user42, user43, user44, user46, user47, user48, user49, user50, user52, user53, user55, user56, user58, user59, user60, user62, user63, user64, user65, user66, user67, user69, user70, user71, user72, user73, user74, user75, user76, user77, user78, user79, user80, user81, user82, user83, user84, user85, user86, user87, user88, user89, user92, user93, user94, user95, user96, user97, user98, user99, user100, user101, user102, user104, user105, user106, user107, user108, user109, user110, user111, user112, user113, user114, user115, user117, user118

Кількість персон	Список персон
90	user3, user5, user6, user7, user9, user10, user12, user13, user16, user17, user18, user20, user21, user22, user23, user24, user25, user26, user27, user30, user31, user32, user33, user34, user37, user38, user39, user40, user41, user42, user43, user44, user46, user47, user49, user50, user52, user53, user55, user56, user58, user59, user60, user62, user64, user66, user67, user69, user70, user71, user72, user73, user74, user75, user76, user77, user78, user79, user80, user81, user82, user83, user84, user85, user86, user87, user89, user92, user93, user94, user95, user96, user97, user98, user99, user100, user101, user102, user104, user105, user106, user107, user108, user109, user111, user112, user114, user115, user117, user118
80	user3, user6, user7, user9, user10, user12, user13, user16, user17, user18, user20, user21, user23, user24, user25, user26, user30, user31, user32, user33, user34, user37, user38, user39, user40, user41, user43, user44, user46, user47, user49, user50, user52, user53, user55, user56, user58, user60, user62, user64, user66, user69, user70, user71, user73, user74, user75, user77, user78, user79, user80, user81, user82, user83, user84, user85, user86, user87, user89, user92, user93, user94, user95, user96, user97, user98, user99, user100, user101, user102, user104, user105, user106, user107, user108, user109, user111, user112, user114, user117
70	user3, user7, user9, user12, user13, user16, user17, user18, user20, user23, user24, user25, user26, user30, user32, user33, user37, user38, user39, user40, user41, user43, user44, user46, user47, user50, user52, user53, user55, user56, user58, user60, user62, user64, user66, user69, user70, user71, user73, user74, user75, user77, user78, user80, user81, user82, user83, user84, user85, user86, user87, user89, user92, user93, user94, user95, user96, user97, user98, user99, user100, user101, user105, user106, user107, user108, user109, user111, user112, user114
60	user3, user7, user12, user13, user16, user17, user18, user20, user23, user24, user25, user26, user30, user32, user33, user37, user38, user39, user41, user44, user46, user47, user50, user53, user55, user56, user58, user60, user62, user66, user69, user70, user71, user73, user74, user77, user78, user80, user81, user82, user83, user84, user85, user86, user87, user89, user92, user93, user95, user97, user98, user99, user100, user101, user105, user106, user107, user108, user109, user112
50	user7, user12, user13, user17, user18, user20, user23, user24, user25, user26, user30, user32, user33, user38, user39, user44, user47, user50, user53, user55, user56, user58, user60, user62, user66, user69, user70, user73, user77, user78, user80, user81, user82, user83, user84, user85, user86, user87, user89, user92, user93, user95, user97, user98, user99, user101, user105, user107, user108, user109

Кількість персон	Список персон
40	user7, user13, user17, user18, user20, user25, user26, user32, user33, user38, user39, user44, user47, user50, user53, user55, user56, user60, user62, user66, user69, user73, user77, user78, user80, user81, user82, user83, user84, user86, user87, user93, user95, user97, user98, user101, user105, user107, user108, user109
30	user13, user18, user20, user26, user32, user33, user39, user44, user47, user50, user53, user56, user60, user62, user66, user69, user73, user77, user78, user80, user82, user83, user86, user87, user95, user97, user98, user105, user107, user109
20	user18, user26, user32, user39, user44, user47, user50, user56, user60, user62, user69, user77, user80, user82, user83, user86, user97, user98, user105, user109
10	user26, user32, user44, user47, user60, user69, user77, user82, user83, user97

Додаток Д

Програмний код розроблених компонентів

Назва файлу: authentication_model.py
Короткий опис: імплементація компонентів для зменшення розмірності даних та класифікації
Вихідний код:

```
from sklearn.decomposition import PCA
from sklearn.pipeline import Pipeline
from sklearn.discriminant_analysis import LinearDiscriminantAnalysis
from sklearn.neural_network import MLPClassifier
from sklearn.preprocessing import LabelEncoder

from models.general_model import GeneralModel
from authentication_model_config import config as model_params


class AuthenticationModel(GeneralModel):
   def __init__(self, model_name):
       super().__init__(model_name, model_params.pipeline)
       self.model_path = model_params['model_path']
       self.label_encoder = LabelEncoder()

   def _build_pipeline(self, params):
       self.pipeline = Pipeline([
           ('pca', PCA(n_components=30)),
           ('clf', MLPClassifier(hidden_layer_sizes=(256, 128), tol=1e-6,
                                 shuffle=True,
                                 activation='tanh'))
       ])
       self.pipeline.set_params(**params)

   def _build_lda_pipeline(self, params):
       self.pipeline = Pipeline([
           ('pca', PCA(n_components=30)),
           ('clf', LinearDiscriminantAnalysis())
       ])
       self.pipeline.set_params(**params)

   def train(self, features, targets):
       classes = list(set(map(lambda x: x['user_name'], targets)))
       self.label_encoder.fit(classes)
       targets = self.label_encoder.transform(list(map(lambda x:
x['user_name'], targets)))
       self.pipeline.fit(features, targets)

   def predict(self, features):
       return self.label_encoder.inverse_transform(
self.pipeline.predict(features))
```

Назва файлу: outliers_correction.py
Короткий опис: утиліти для виявлення та виправлення артефактів
Вихідний код:

```
import h2o
import numpy as np
from h2o.estimators.deeplearning import H2OAutoEncoderEstimator

h2o.init()

def statistical_outliers_correction(input_beats, config):
   beats = input_beats
   windows_size = config['windows_size']
   (n_beats, len_beats) = beats.shape

   # expansion beats signal in order to fit the whole windows number
   if (len_beats % windows_size) != 0:
       beats = np.hstack((beats, np.zeros((n_beats, windows_size -
(len_beats % windows_size)))))

   # w - arrays with windows coordinates in the beats
   w = np.arange(0, len_beats, windows_size)
   w = np.vstack((w, w + windows_size))

   # calculating the average std's threshold
   beats_mean = np.mean(beats, axis=0)
   i_thr = 0
   for i in range(n_beats):
       i_thr = i_thr + np.sum((beats[i, :] - beats_mean) ** 2)
   i_thr = i_thr / (len_beats * n_beats)

   # if the error is exceeds the threshold within the window,
   # then replace the relevant part on the averaged part
   for k in range(n_beats):
       e = (i_thr / (beats[k, :] - beats_mean) ** 2) >=
config['variance_gain']
       for p in range(w.shape[1]):
           stw = w[0, p]
           enw = w[1, p]
           if np.all(e[stw: enw]):
               continue
           else:
               idx = [x for x in range(n_beats) if x != k]
               beats[k, stw:enw] = np.mean(beats[idx, stw:enw], axis=0)
   return beats[:, :len_beats]

def autoencoders_outliers_detection_with_window_correction(input_beats):
   ecg_frame = h2o.H2OFrame(input_beats)
   model = H2OAutoEncoderEstimator(
           activation="rectifier_with_dropout",
           initial_weight_distribution="UniformAdaptive",
           adaptive_rate=True,
           max_w2=10.0,
           hidden=list([100, 100, 100, 100, ]),
           l1=1e-5,
           l2=1e-5,
```

```
            score_interval=0,
            epochs=100
        )

    model.train(x=ecg_frame.names, training_frame=ecg_frame)
    reconstruction_error = model.anomaly(ecg_frame, per_feature=True)
    df = reconstruction_error.as_data_frame()

    # quantile
    rmse = df.values
    outputs = np.zeros(rmse.shape)
    rmse_quantile = np.asarray(df.quantile(np.linspace(.01, 1, 99)))
    for i in range(rmse.shape[1]):
        for j in range(rmse.shape[0]):
            outputs[j, i] = np.interp(rmse[j, i], rmse_quantile[:, i],
np.linspace(.01, 1, 99))
    anomalies = outputs > 0.95
    for i in range(anomalies.shape[1]):
        if np.all(anomalies[:, i]):
            anomalies[:, i] = np.logical_not(anomalies[:, i])

    w = 5
    for i in range(anomalies.shape[1]):
        st = i - w if (i-w) >= 0 else 0
        en = i + w if (i + w) < anomalies.shape[1] else anomalies.shape[1]
        input_beats[anomalies[:, i], st:en] =
np.mean(input_beats[np.logical_not(anomalies[:, i]), st:en], axis=0)
    return input_beats


def autoencoders_outliers_correction(input_beats):
    ecg_frame = h2o.H2OFrame(input_beats)
    model = H2OAutoEncoderEstimator(
            activation="rectifier_with_dropout",
            initial_weight_distribution="UniformAdaptive",
            adaptive_rate=True,
            max_w2=10.0,
            hidden=list([100, 100, 100, 100, ]),
            l1=1e-5,
            l2=1e-5,
            score_interval=0,
            epochs=100
        )

    model.train(x=ecg_frame.names, training_frame=ecg_frame)
    return model.predict(ecg_frame).as_data_frame().values
```

Назва файлу: segmentation.py
Короткий опис: утиліти для сегментації та темпоральної нормалізації ЕКГ-сигналу
Вихідний код:

```
import numpy as np
from biosppy.signals.ecg import hamilton_segmenter

def r_peak_detection(data, sampling_rate, config):
   if config['r_peak_detection_type'] == 'hamilton':
       r_peaks_tuple = hamilton_segmenter(signal=data,
sampling_rate=sampling_rate)
       r_peaks_ind = r_peaks_tuple['rpeaks']
       # expected range of R peak (in samples)
       r_peak_neighbourhood = config['r_peak_detection_neighbourhood']
       for i in range(r_peaks_ind.shape[0]):
           start = np.maximum(r_peaks_ind[i] - r_peak_neighbourhood, 0)
           stop = np.minimum(r_peaks_ind[i] + r_peak_neighbourhood,
len(data))
           ind = np.argmax(data[start: stop])
           r_peaks_ind[i] = start + ind
       r_peaks_val = data[r_peaks_ind]
   else:
       # unsupported mode
       r_peaks_val = None
       r_peaks_ind = None
   return r_peaks_val, r_peaks_ind

def segmentation(data, r_peaks_ind, config):
   left_r_peak_offset = int(config['r_peak_position'] *
config['segment_length'])
   n_peak = r_peaks_ind.shape[0]
   ecg_segm = np.zeros([n_peak - 1, config['segment_length']])
   r_peak_dist = np.diff(r_peaks_ind)
   median_dist = np.median(r_peak_dist)
   i_segm = 0  # segment counter
   for i in range(n_peak - 1):
       start = int(r_peaks_ind[i] - left_r_peak_offset)
       if config['use_temporal_normalization']:
           stop = int(start + r_peak_dist[i])
       else:
           stop = int(start + np.minimum(r_peak_dist[i],
config['segment_length']))
       if config['use_temporal_normalization']:
           if data[start:stop].shape[0] <= config['pqrs_length'] * 1.2:
               continue
           len_data_to_interp = int(median_dist) - config['pqrs_length']
           data_to_interp = np.zeros((len_data_to_interp,))
           tmp_data =
data[start:stop][config['pqrs_length']:][:len_data_to_interp]
           data_to_interp[:tmp_data.shape[0]] = tmp_data
           interp_data = np.interp(np.linspace(0, 1,
config['segment_length'] - config['pqrs_length']), np.linspace(0, 1,
data_to_interp.shape[0]), data_to_interp)
```

```
            ecg_segm[i_segm, :] =
np.hstack((data[start:stop][:config['pqrs_length']], interp_data))
        else:
            ecg_segm[i_segm, 0:stop - start] = data[start:stop]
        i_segm += 1  # increment segment counter
    return ecg_segm[~np.all(ecg_segm == 0, axis=1)]
```

Назва файлу: filtration.py
Короткий опис: імплементація цифрового фільтра
Вихідний код:

```
from scipy import signal

def filtrate(data, sampling_rate, config):
    # create bandpass filter
    freq_pass = config['freq_pass'] / (sampling_rate / 2.0)
    freq_stop = config['freq_stop'] / (sampling_rate / 2.0)
    filt_order, cut_freq = signal.buttord(
        wp=freq_pass,
        ws=freq_stop,
        gpass=config['gain_pass'],
        gstop=config['gain_stop'])
    param = signal.butter(filt_order, cut_freq, btype='bandpass')
    # input signal filtering
    filt_data = signal.filtfilt(param[0], param[1], data)
    # input signal normalization
    norm_gain = (max(filt_data) - min(filt_data)) / 2
    norm_data = filt_data / norm_gain
    return norm_data
```

Назва файлу: calculate_features.py
Короткий опис: опрацювання ЕКГ-сигналу
Вихідний код:

```
from features.features_config import config
from features.modules.filtration import filtrate
from features.modules.segmentation import r_peak_detection, segmentation
from features.modules.outliers_correction import *
from features.modules.normalization import normalize

def calculate_features(ecg_record, calc_config=config):
    filtered_ecg_record = filtrate(
        ecg_record, calc_config['sampling_rate'],
calc_config['filtration'])
    r_peaks_val, r_peaks_ind = r_peak_detection(
        filtered_ecg_record, calc_config['sampling_rate'],
calc_config['segmentation'])
    ecg_segments = segmentation(filtered_ecg_record, r_peaks_ind,
calc_config['segmentation'])
    ecg_segments = statistical_outliers_correction(ecg_segments,
calc_config['outliers_correction'])
    ecg_features = normalize(ecg_segments)
    return ecg_features
```

Додаток Е

Список публікацій здобувача за темою дисертації та відомості про апробацію результатів дисертації

Наукові праці, в яких опубліковано основні наукові результати дисертації

1. 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. *Особистий внесок здобувача: здійснено імплементацію розробленої системи та досліджено продуктивність системи при різній кількості класів.*

2. 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. *Особистий внесок здобувача: розроблено алгоритм темпоральної нормалізації ЕКГ-сигналу та верифіковано його ефективність.*

3. Хома В.В., Хома Ю.В., Сабодашко Д.В., Хома П.П. Автоенкодери для опрацювання промахів сигналів ЕКГ у системі біометричної автентифікації // Штучний інтелект. – 2019. №1-2. С. 108-117. *Особистий внесок здобувача: здійснено імплементацію компоненти для коригування промахів ЕКГ-сигналів. Здійснено її інтеграцію із компонентою для виявлення промахів на основі автоенкодера.*

4. 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. *Особистий внесок здобувача: здійснено дослідження стійкості ЕКГ-сигналу, продемонстровано ефективність алгоритмів темпоральної нормалізації для забезпечення ефективної роботи біометричної системи впродовж тривалої експлуатації без проміжних перекалібрувань.*

5. Хома Ю.В., Хома В.В., Сабодашко Д.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. *Особистий внесок здобувача: підготовлено набори даних для експериментальної частини дослідження, здійснено вибір оптимальних гіперпараметрів для методів виявлення та виправлення промахів.*

6. 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) Особистий внесок здобувача: здійснено імплементацію статистичного методу для виявлення та виправлення артефактів, зроблено вибір оптимальних гіперпараметрів та досліджено ефективність імплементованого методу.*

7. 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) Особистий внесок здобувача: розроблено компоненту для виявлення артефактів у ЕКГ-сигналах. Здійснено її інтеграцію із компонентою котра здійснює коригування артефактів у ЕКГ-сигналах.*

8. Sabodashko D. Normalizacja temporalna sygnału EKG w systemie identyfikacji biometrycznej // Przetwarzanie, transmisja i bezpieczeństwo informacji: monografia / Akademia Techniczno-Humanistyczna. Bielsko-Biała, 2019. T. 2. S. 313–322. *Особистий внесок здобувача: сформовано конфігурації експериментів та здійснено дослідження стійкості ЕКГ-сигналу як біометричного маркера для задач розпізнавання.*

Праці, які засвідчують апробацію матеріалів дисертації

1. 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. *Особистий внесок здобувача: розроблено та апробовано програмні компоненти котрі здійснюють цифрову обробку ЕКГ-сигналу.*

2. 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) Особистий внесок здобувача: імплементовано нейромережевий класифікатор на основі якого здійснювалась ідентифікація, здійснено його верифікацію.*

3. 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) Особистий внесок здобувача: здійснено вибір оптимальної архітектури автоенкодера, досліджено його продуктивність.*

4. 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. *Особистий внесок здобувача: імплементовано програмні компоненти біометричних систем, проведено експериментальну частину дослідження.*

5. 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. *Особистий внесок здобувача: представлено порівняльну характеристику сучасних біометричних систем, проведено порівняння найпоширеніших біометричних маркерів за допомогою формалізованих критеріїв.*