

М. С. Єфремов, Ю. В. Крак

ПРОГРАМНЕ РІШЕННЯ ДЛЯ ЗБЕРІГАННЯ ТА АНАЛІЗУ ЕЛЕКТРОКАРДІОГРАМ

Розробка програмних рішень для роботи у медичній сфері з кожним роком стає популярнішою завдяки розвитку інформаційних технологій. За останні десятиліття алгоритми знаходження окремих ділянок на електрокардіограмі (ЕКГ), виявлення та класифікація аритмій та інших патологій на ЕКГ досягли значної точності. У даній роботі запропоновано підхід до інтеграції таких алгоритмів у єдину систему шляхом розробки програмного рішення, яке аналізує наявні ЕКГ з різних джерел, приводить їх до спільного формату та надає звітність щодо пацієнтів. Реалізація здійснена у вигляді веб-застосунку, що може надати користувачеві інформацію із кардіограм, обробляти ЕКГ залежно від доданих алгоритмів, зберігати потрібні ділянки ЕКГ, результати діагностування захворювань тощо. В систему імplementовано новий алгоритм знаходження R-піків та візуалізація ЕКГ-сигналу обробленим алгоритмом з анотаціями, отриманими як автоматично, так і за участі спеціалістів-кардіологів.

Ключові слова: електрокардіограма, оброблення сигналів, програмні рішення, база даних, формати даних

M.S. Yefremov, Yu.V. Krak

SOFTWARE SOLUTION FOR STORING AND ANALYZING ELECTROCARDIOGRAMS

Software solutions for medical and healthcare fields are becoming increasingly popular due to advancements in information technology. Over the past decades peak detection algorithms as well as detection of other segments on ECG, arrhythmias detection and other pathologies detection algorithms achieved significant accuracy. This study proposes an approach for integrating such algorithms into a unified systems by developing a software solution that analyzes available ECGs from various sources, standardizes them into a common format, and provides detailed reporting that is easily understandable for medical experts or patients. The implementation is carried out as a web application that can provide users with information about cardiograms, process ECGs depending on the integrated algorithms, store necessary ECG segments, diagnostic results, and more. The system incorporates a newly developed algorithm for detecting R-peaks and visualizing ECG signals processed by the algorithm with annotations obtained both automatically and with the participation of cardiology specialists.

Key words: electrocardiogram, signal processing, software solutions, database, data formats.

Вступ

Електрокардіограма є важливим інструментом діагностики роботи серця людини, на основі якого можна встановити діагноз виявлених захворювань серця та визначити профілактику кардіологічних захворювань. Аналіз ЕКГ дозволяє лікарям виявляти різноманітні серцеві

проблеми, такі як аритмії різного рівня, ішемії та інші порушення серцевого ритму. У контексті ЕКГ для аналізу сигналу використовують алгоритми автоматичного знаходження піків, (наприклад, R-піків), інтервалів (P-QRS-T) та інших характеристик ЕКГ (див. рисунок 1).

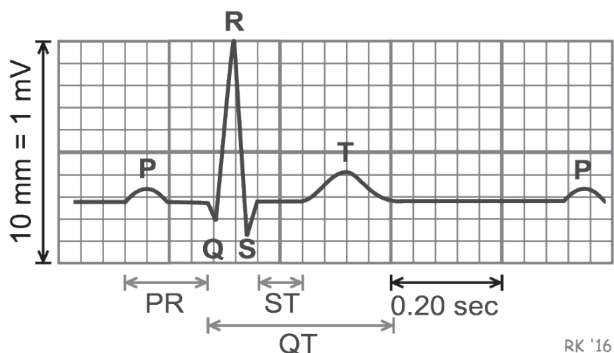


Рис. 1. Типовий аналіз одного серцевого ритму

Для оцінки точності та надійності цих алгоритмів необхідні великі та добре анотовані бази даних ЕКГ. Такі бази даних забезпечують стандартизовані набори даних, що включають різні типи серцевих ритмів і патологій, які можна використовувати для навчання та тестування алгоритмів. Такі бази стали незамінними ресурсами для вчених та розробників медичних програмних рішень.

Метою цієї статті є розроблення способів інтегрування алгоритмів обробки ЕКГ у системи, які можна використовувати у медичних закладах для аналізу та моніторингу станів пацієнтів. Для того, аби алгоритми аналізу працювали, треба щоб дані з холтерів або звичайних апаратів зняття ЕКГ можна було зчитати та перетворити у математичні моделі, зручні для роботи з алгоритмами. Відзначимо, що різні пристрої зняття ЕКГ-сигналів мають свої влаштовані формати подання даних, і приведення цих даних до певних стандартизованих форматів є надзвичайно важливою проблемою.

1. Огляд існуючих та стандартизованих форматів зберігання кардіограм

Для того, щоб здійснювати обробку ЕКГ та розробку застосунків, які працюють із кардіограмами важливо зрозуміти, в яких форматах зберігаються кардіограми, та які з них масово використовуються спеціалістами. Протягом останніх десятиліть дані з різних холтерів та кардіографів було збережено у базах даних, які мають відкритий доступ. Звернувшись до найпопулярніших

з таких баз, можна зробити висновки, які формати широко використовуються для зберігання і дослідження ЕКГ-сигналів. Наприклад, такі бази можна знайти на ресурсах Kaggle та PhysioNet [1]. Ці бази не просто мають вільний доступ, а були зібрані з кардіограм знятих з реальних пацієнтів. Розглянемо деякі бази із цих ресурсів. На ресурсі PhysioNet знаходиться велика кількість баз ЕКГ [1].

З основних можна виділити:

- MIT-BIH Arrhythmia Database[1].

Це одна з найвідоміших баз даних для аналізу серцевих аритмій, яка використовується у більшості досліджень, оскільки має анотовані сигнали різних випадків серцевих аритмій;

- European ST-T Database[1].

Ця база даних створена для оцінки алгоритмів виявлення та аналізу ST-сегменту і Т-хвилі. Вона містить записи від пацієнтів з різними серцевими станами;

- QT Database[1].

Вона містить записи ЕКГ з детальними анотаціями інтервалу QT, що важливо для дослідження інтервалу QT і пов'язаних з ним серцевих умов;

- Physikalisch-Technische Bundesanstalt (PTB) Diagnostic ECG Database[1]

Ця база даних містить багатоканальні записи ЕКГ від пацієнтів з різними серцевими захворюваннями, зокрема, інфарктом міокарда.

Що ж до ресурсу Kaggle, то на ньому є кілька баз даних ЕКГ, які використовуються, у тому числі як основні дані для проведення змагань для стимулювання розробки алгоритмів для автоматичного аналізу ЕКГ. База даних включає записи ЕКГ від китайських пацієнтів з різними типами аритмій. Однією з таких баз є база даних

- China Physiological Signal Challenge (CPSC) 2018[2].

Ця база даних містить записи ЕКГ з анотаціями, які можна використовувати для розробки і тестування алгоритмів розпізнавання ритмів і яка була наповнена 2018 року.

Наведемо інші бази з ресурсу Kaggle:

- China Physiological Signal Challenge (CPSC) 2020 Database[3].

Ця база даних була використана у змаганні CPSC 2020 і містить ширший набір даних для аналізу ЕКГ.

- Xinjiang Medical University ECG Data.

Ця база даних містить записи ЕКГ від пацієнтів з різними серцевими станами, надані Сінцзянським медичним університетом [4]

Для пошуку цих баз даних на Kaggle можна скористатися пошуком на платформі, ввівши відповідні ключові слова, наприклад, "ECG", "China", "CPSC". Kaggle також надає інструменти для роботи з даними та проведення змагань, що може бути корисним для тестування і розробки алгоритмів.

Розглянувши відкриті бази з сигналами ЕКГ, можна розділити файли сигналів ЕКГ на три категорії:

1. Файли у форматах стандартизованих ААМІ [5] такі як edf, bdf[6] тощо.
2. Бінарні файли у форматах, які були створені різними дослідницькими інститутами під себе, але, враховуючи їхню популярність, більшість програмних рішень, як-от wfdb[7], MATLAB мають бібліотеки та інструменти для роботи з ними.
3. Файли, які ніяк не шифруються, зазвичай це ЕКГ, зняті за короткий проміжок часу.

2. Декодування файлів ЕКГ та приведення до загального формату

Під час створення застосунку для аналізу кардіограм з метою його ширшого використання як у наукових дослідженнях, так і у прикладних застосуваннях, потрібно врахувати сумісність з якомога більшою кількістю форматів сигналу. Відзначимо, що, хоча описані вище формати й часто використовуються, підтримка будь-якого конкретного з них не гарантується у кардіографах та холтерах, особливо в пристроях, створених до стандартизації ЕКГ. Також постає проблема, пов'язана з тим, що здебільшого прилади, які повертають результати вимірювання, закодовані так, що їх можна прочитати лише використовуючи застосунки, розроблені авторами пристрою. Наприклад, холтер від фірми Philips[8], є комерційним продуктом, який часто використовується у медичних закладах, тому його формат .zhr[8] є поширеним та навіть підтримується wfdb. Навіть застосунок який іде разом з холтером має можливість конвертувати цей формат у стандартизовані edf, MIT тощо. Але розробники не надають файл метаданих, в якому зберігається інформація про розшифрування сигналу, тому постає проблема використання цих даних для розробників інших програмних засобів. Одним із способів вирішення цієї проблеми є метод підбору. Оскільки всі файли з кардіографів кодуються у бінарний формат, то підібрати функцію, що його декодує можливо, хоча ця процедура займає багато часу і під час декодування можливі втрати фрагментів сигналу. Так, під час роботи з холтером SDM-23 була спроба самостійно декодувати сигнал. Як результат сигнал втрачав деякі фрагменти (див. рисунок 2). Це зумовлено двома факторами: по-перше, відсутні метадані, де вказано параметри сигналу, які методом підбору не вийшло підібрати; по-друге, різні частини сигналу можуть бути закодовані різними методами. В результаті, окрім підбору алгоритму декодування, є потреба перебирати частинами, що є недоцільною витратою часу, оскільки заздалегідь невідомо, які саме методи та варіанти були закодовані.

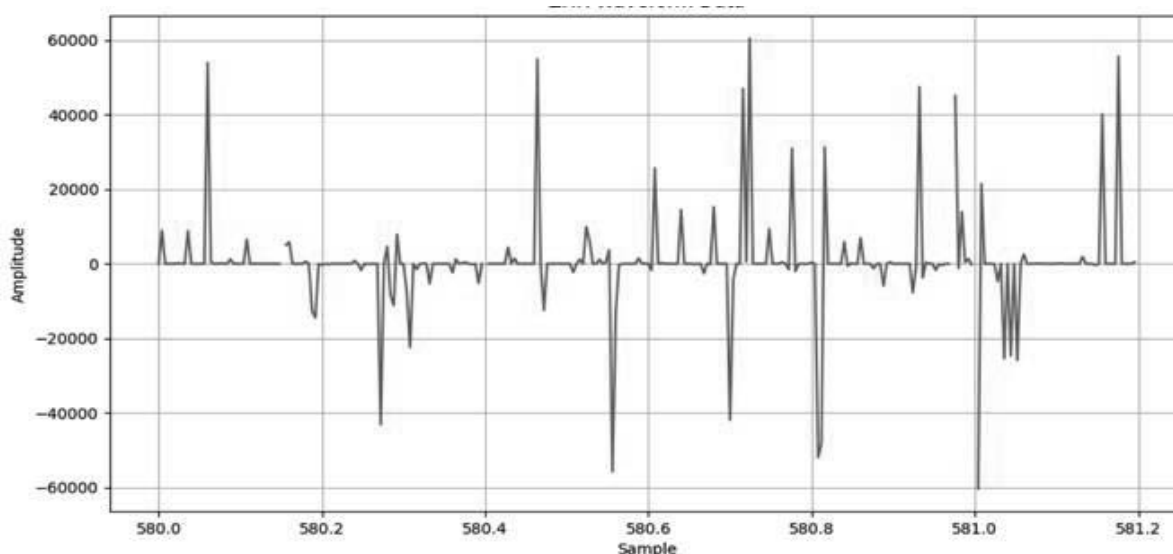


Рис 2. Невдале декодування сигналу ЕКГ

У таких випадках, коли сигнал неможливо якісно декодувати – єдиний вихід є напряму звертатись до розробників апаратури, або працівників медичних закладів та, за можливості, домовлятися про співпрацю.

Таким чином, з наведеної інформації можна зробити висновки, що під час обробки та аналізу ЕКГ даних у різних форматах важливою є розробка застосування, який давав можливість враховувати ці формати і декодувати інформацію у доступні формати. Зважаючи на це було підібрано архітектуру, яка дозволяє використовувати формати розробки як доступні та які широко використовуються, так і формати локальних пристроїв. Використавши паттерн-програмування «фабрика» створюється контролер форматів, який буде відповідати за модулі для кожного нового формату, підтримка якого додаватиметься до системи.

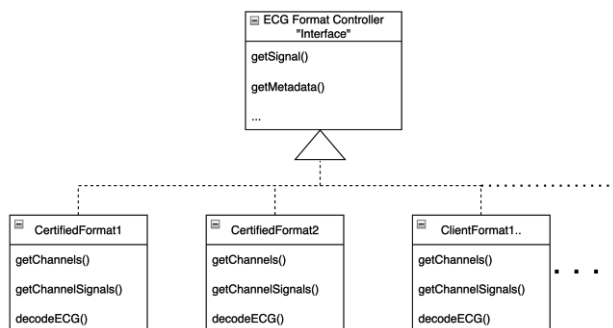


Рис 3. Фабрика декодування форматів ЕКГ

На рисунку 3 зображено реалізацію фабричного методу для декодування ЕКГ.

Основними компонентами для системи, що відповідає за декодування ЕКГ форматів, є контролер, який залежно від формату повертає з фабрики класів обробників потрібний обробник форматів. Контролер форматів визначає тип вхідних даних і направляє їх до відповідного обробника, здатного обробити конкретний формат. Фабрика обробників створює правильний парсер для даного формату, що витягує числові канали та метадані з вхідних даних, а також може додавати додаткову інформацію. Фабрика обробників є ключовою, оскільки вона забезпечує централізовану точку управління для створення обробників. Це дозволяє розширювати систему, додаючи нові формати без необхідності змінювати існуючий код, що підвищує модульність і підтримувальність (maintainability) системи. Кожен формат обробляється окремим обробником, що спрощує тестування та налагодження. Такий підхід особливо корисний у лікарнях, де використовуються власні формати ЕКГ, оскільки він дозволяє інтегрувати нові формати без значних змін у системі.

3. Архітектура рішення для роботи

Для повноцінної роботи системи, що здійснює обробку даних, пов'язаних з медициною варто враховувати, що такі дані містять конфіденційну інформацію про пацієнтів, яка охороняється законами про захист персональних даних, такими як GDPR у Європі [9] або HIPAA у США [10]. Забезпечення анонімності даних та їхній захист від несанкціонованого доступу є важливим, але складним завданням. Щоб не зберігати персональні дані клієнта, у запропонованому застосунку для роботи з ЕКГ, будемо підключати до нього реляційну базу даних, яка слугуватиме сховищем деанонізованим файлам. У цій базі будемо зберігати як закодовані файли, так й анонімізовані декодовані для подальшої їх обробки. Також до бази можна буде зберігати файли як через безпосередні застосунки поза даним сервісом, так й через

певний інтерфейс для користувачів цього застосунку.

Декодовані файли, без метаданих пацієнта вже можна передавати для їхньої подальшої обробки алгоритмами ЕКГ-аналізу. Так, дані з холтерів та кардіографів, які були отримані сервісом управління форматами, передаються на аналіз іншому сервісу, в якому реалізовані алгоритми обробки ЕКГ. Відзначимо, що був імплементований сервіс із знаходження R-піків, описаний у роботах [11,12] Цей алгоритм був протестований на вищезазначеній базі, а саме на базі з ресурсу [1], що підтвердило його ефективність та точність. Повна архітектура з усіх сервісів, які наразі доступні, можна побачити на рисунку 4. Така архітектура є досить гнучкою, та дозволяє інтегрувати його до різних лікарень без суттєвих проблем.

Для того, щоб будь-яка система мала успіх – важливою частиною є ті сервіси системи, які напряду комунікують із користувачем. У даному випадку - це є

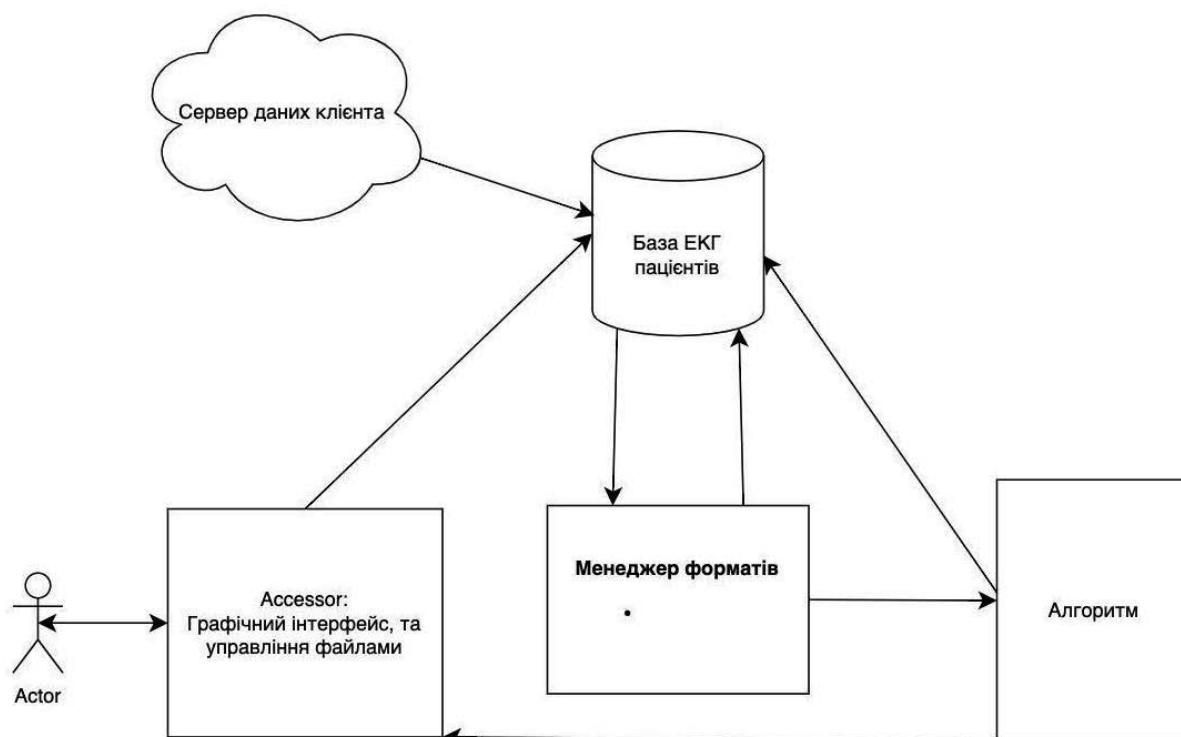


Рис. 4. Архітектура сервісів застосунку аналізу ЕКГ.

графічний інтерфейс, через який «актор» (лікар) завантажує файли з ЕКГ та отримує результати аналізу сигналу. Назвемо такий сервіс «Accesor», оскільки серед його задач є не лише відобразити, зчитати файл та занотувати проаналізований файл, а й упевнитись, що доступ до системи мають лише певні кваліфіковані люди. Такий доступ до системи можна організувати, як звичайним SSO провайдером, так і простими методами авторизації. Перед початком розробки GUI частини, що відображує сигнал було проаналізовано візуалізацію анотованих файлів бази МІТ-ВІН на ресурсі PhysioNet. Для зручності, сигнал розбивається на задані проміжки, переходячи між якими, можна роздивитись сигнал по частинах, наблизивши певні підозрілі ділянки.

На рисунку 5. зображено приклад, як виглядає графічний інтерфейс застосунку. Тепер, маючи представлення, з яких сервісів буде складатись застосунок, перейдемо до вибору технічних рішень, які можна використовувати для його реалізації. Для підтримки крос-платформленості було обрано реалізацію як веб застосунок. Таке рішення не лише уможливило збереження крос-платформленості, а й зберігає час потенційного користувача, якому не треба буде робити інсталяцію. Розгляда-

ючи тенденції мов програмування сьогодні, запропоноване рішення можна розробляти або використовуючи технологію nodeJS, або Flask/Django[13]. Було вирішено використовувати Flask, оскільки він написаний мовою Python, яка є популярною серед науковців, що займаються проблемами ЕКГ, та вже має готові бібліотеки роботи зі стандартизованими та поширеними форматами.

Відзначимо ще одну проблему, яку вирішує можливість авторизації.

У реальних умовах у великих лікарнях за день можуть приходити десятки ЕКГ для аналізу різних пацієнтів, прив'язаних до різних лікарів. Авторизація дозволяє відфільтрувати, які ЕКГ доступні користувачеві (в нашому випадку це лікар), що дозволить кожному лікарю бачити лише своїх пацієнтів та уникати плутанини.

Після проведення аналізу результатів будь-яких медичних вимірювань пацієнт зазвичай очікує опис від лікаря та звіт про зафіксовані патології або аномалії. Тому невід'ємною частиною застосунку, який проводить аналіз ЕКГ, є генерація звітності. Логічно цю задачу делегувати застосунку «Accesor», оскільки він комунікує з базою даних та отримує результати обробки ЕКГ алгоритмами аналізу.

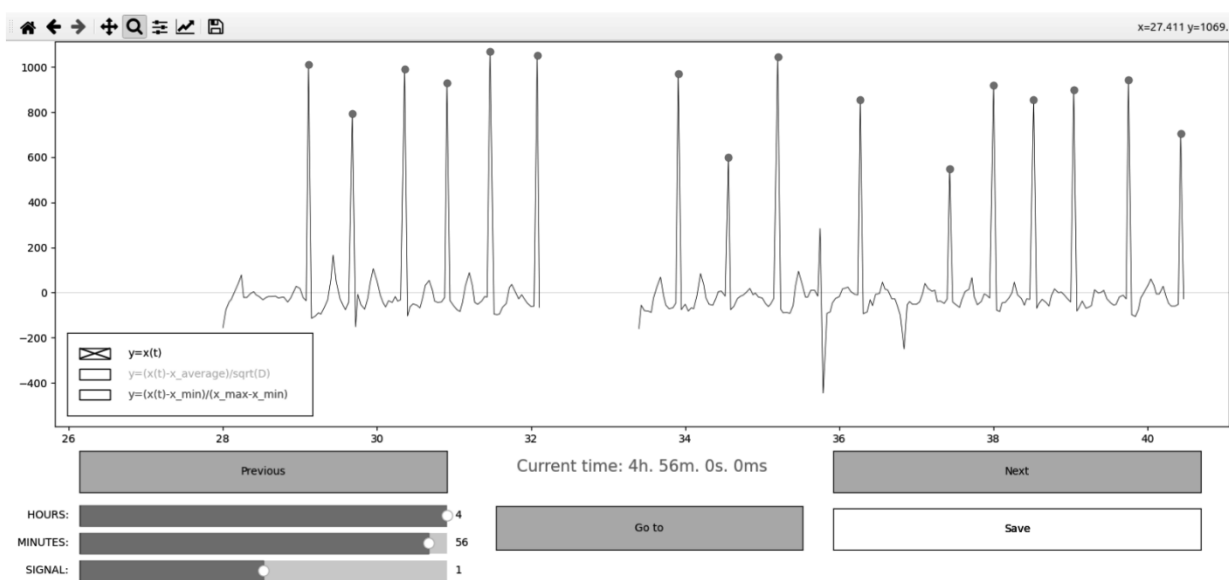


Рис. 5. Система візуалізації кардіограми з знайденими піками

Зазвичай звіт зручно генерувати або у PDF або .doc документи. Звісно, генерація таких документів є вже задачею складнішою за відображення, сортування файлів та внесення додаткової інформації, тому для цього варто створити окремий сервіс, який би з отриманих результатів обробки ЕКГ та додатково доданих лікарем анотацій та коментарів, генерував документи для пацієнтів. Підкреслимо, що декодовані ЕКГ є звичайними масивами з чисел, які можна зберігати у реляційній базі даних, то самі файли ЕКГ є складними бінарними структурами, тому треба використовувати сервіс зберігання даних. Це може бути як хмарне рішення, так і FTP сервер тощо. Оскільки декодуємо й отримуємо ЕКГ як масив чисел, з якого потім можна відтворити сигнал, то зберігати основні файли довгостроково немає сенсу. Такий підхід дозволяє використовувати сховища даних невеликих обсягів, які час від часу вичищаються. Поповнення таких сховищ можна реалізувати як через «Accesor», так і напряду з серверів даних лікарень, якщо у лікарні такі є. Беручи до уваги описані сервіси разом, – така мікро-сервісна архітектура дозволяє значно прискорити та автоматизувати роботу кардіологів, що приведе до покращення роботи з пацієнтами.

Висновки

Розроблено та запропоновано підхід для автоматизації роботи кардіологів у лікарнях шляхом розробки системи, яка дозволяє інтегрувати у зручному для лікарів форматі алгоритми аналізу ЕКГ у їх роботу опису кардіограми. Система дозволяє не лише здійснювати аналіз ЕКГ у пришвидшеному темпі, а й систематизувати файли різних форматів, які можуть бути наявні у лікарнях. Запропонований підхід дозволяє обробляти різні типи кардіограм та конвертувати їх у єдиний формат, що спрощує їхнє подальше використання алгоритмами роботи з ЕКГ. Розроблено крос-платформний застосунок, який дозволяє продивлятися фрагменти кардіограми, з детальною візуалізацією та алгоритмічно доданими нотатками, з мож-

ливістю додавання анотацій лікарями напряду, якщо алгоритм аналізу ЕКГ пропускає деякі ділянки. Система підтримує роботу із стандартизованими форматами ЕКГ та має можливість інтегрувати будь-який формат, якщо відома його специфікація. Архітектура побудована так, що дозволяє інтегруватись у готові рішення, які наявні у лікарнях, що дозволяє отримувати повну історію пацієнтів. Такий підхід забезпечує більш ефективне управління даними пацієнтів, що сприяє зниженню кількості помилок та підвищенню якості надання медичних послуг.

Даний сервіс був протестований на розробленому алгоритмі пошуку R-піків, для демонстрації його точності у порівнянні з іншими алгоритмами. Використання описаного сервісу не лише пришвидшує роботу кардіологів, а й систематизує зберігання ЕКГ у зручному вигляді та започатковує вирішення проблеми автоматизації аналізу медичних вимірювань, дозволяючи додавати алгоритми обробки та підтримку будь-яких форматів, не змінюючи основну архітектуру системи.

References

1. Goldberger, A.L. et al. “PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals.” *Circulation*. Vol. 101, 23 (2000): E215-20. doi:10.1161/01.cir.101.23.e215
2. Liu, F. et al. An open access database for evaluating the algorithms of electrocardiogram rhythm and morphology abnormality detection. *Journal of Medical Imaging and Health Informatics*, 8(7), (2018): 1368-1373. doi:10.1166/jmihi.2018.2442
3. Cai, Z. et al. An Open-Access Long-Term Wearable ECG Database for Premature Ventricular Contractions and Supraventricular Premature Beat Detection, *Journal of Medical Imaging and Health Informatics*, 10(11), (2020): 2663–2667. DOI: 10.1166/jmihi.2020.32892663
4. Mu, H. et al. Electrocardiogram Minnesota codings from 30 000 adult cases with Kazakh ethnicity in Xinjiang, China. *Zhonghua liuxingbingxue zazhi*. 31(4), (2010): 451-454 PMID: 20513295.

5. ANSI/AAMI EC11:1991/(R) (2001) *Diagnostic Electrocardiographic Devices, 2ed*
6. Kemp, B., Olivan, J. European data format 'plus' (EDF+), an EDF alike standard format for the exchange of physiological data. *Clinical Neurophysiology*, 14(9),(2003):1755-1761. [https://doi.org/10.1016/S1388-2457\(03\)00123-8](https://doi.org/10.1016/S1388-2457(03)00123-8).
7. Xie, C. et al. Waveform Database Software Package (WFDB) for Python (version 4.1.0). (2023). *PhysioNet*. doi/10.13026/9njx-6322
8. Philips Healthcare SERVICE BULLETIN. SB86000222A: 860292, 860323, 989803190581 - *Release Holter 3.0.3*. 2017.
9. Health Insurance Portability and Accountability Act [HIPAA] of 1996, Pub. L. No. 104-191.
10. Consolidated text: Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).
11. Krak, Iu., Pashko A., Stelia O., Barmak, O., Pavlov, S. Selection Parameters in the ECG Signals for Analysis of QRS Complexes. *1st International Workshop on Intelligent Information Technologies and Systems of Information Security, InteIITSIS 2020*, Khmelnytskyi, 10-12 June 2020. Vol. 2623. 2020, pp.1-13. <https://ceur-ws.org/Vol-2623/paper1.pdf>
12. Pashko, A., Krak, Iu., Stelia, O., Khorozov, O. Isolation of informative features for the analysis of QRS complex in ECG signals. *Advances in Intelligent Systems and Computing*, Vol. 1246 AISC, 2021. pp. 409 – 422. DOI: 10.1007/978-3-030-54215-3_26.
13. Grinberg, M. (2018). *Flask web development: developing web applications with python. x27*; Reilly Media, Inc.

Одержано: 02.09.2024

Внутрішня рецензія отримана: 11.09.2024

Зовнішня рецензія отримана: 13.09.2024

Про авторів:

^{1,2}Єфремов Микола Сергійович,
аспірант, асистент кафедри
<https://orcid.org/0000-0001-8698-3957>

^{1,2}Крак Юрій Васильович,
член-кореспондент НАН України,
д.ф.-м. н., професор,
завідувач кафедри
<https://orcid.org/0000-0002-8043-0785>

Місце роботи авторів:

¹Київський національний університет імені
Тараса Шевченка,

²Інститут кібернетики імені В.М.Глушкова
НАН України

Тел.: +380 44 239-33-33

e-mail: Iurii.krak@knu.ua, Yefremov@knu.ua