

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Харківський національний університет радіоелектроніки

Центр післядипломної освіти
Кафедра Програмної інженерії

КВАЛІФІКАЦІЙНА РОБОТА Пояснювальна записка

другий (магістерський)

(рівень вищої освіти)

Дослідження методів попередньої обробки датасетів сигналів ЕЕГ

Виконав:

студент 2 курсу групи ІПЗзДМ-21-1

Кісельгов Є.М.

(прізвище, ініціали)

Спеціальність 121 – Інженерія програмного

забезпечення

(код і повна назва спеціальності)

Тип програми Освітньо-наукова

Керівник проф. Смеляков С.В.

(посада, прізвище, ініціали)

Допускається до захисту

Зав. кафедри, проф. _____

Дудар З.В.

2023 р.

Харківський національний університет радіоелектроніки

Факультет _____ Центр післядипломної освіти _____
Кафедра _____ Програмної Інженерії _____
Рівень вищої освіти _____ другий (магістерський) _____
Спеціальність _____ 121 – Інженерія програмного забезпечення _____
(код і повна назва спеціальності)
Тип програми _____ Освітньо-наукова _____
Освітня програма _____ Інженерія програмного забезпечення _____

ЗАТВЕРДЖУЮ:

Зав. кафедри _____
(підпис)

« _____ » _____ 2023 р.

ЗАВДАННЯ**НА КВАЛІФІКАЦІЙНУ РОБОТУ**

студента _____ Кісельгова Євгена Миколайовича _____

1. Тема роботи «Дослідження методів попередньої обробки датасетів сигналів ЕЕГ» _____затверджена наказом по університету від «03» квітня 2023 р. № 83Стз2. Термін подання студентом роботи до екзаменаційної комісії «12» травня 2023р.3. Вихідні дані до роботи аналіз методів обробки датасетів ЕЕГ сигналів, розробка програмної системи очищення ЕЕГ від артефактів, застосування методу аналізу незалежних компонент, фреймворк розробки .Net, мова C# _____4. Перелік питань, що потрібно опрацювати в роботі мета роботи, аналіз предметної галузі та постановка задачі, дослідження існуючих технологій попередньої обробки ЕЕГ, ІСА алгоритм, формування вимог до програмної системи, розробка алгоритму очищення ЕЕГ та програмної системи, експериментальна перевірка системи _____

КАЛЕНДАРНИЙ ПЛАН

Номер	Назва етапу роботи	Термін виконання етапу роботи	Примітка
1	Аналіз предметної галузі	03.04.2023 - 10.04.2023	виконано
2	Постановка задачі	03.04.2023 - 10.04.2023	виконано
3	Формування вимог	11.04.2023 - 12.04.2023	виконано
4	Проведення експериментів	11.04.2023 - 15.04.2023	виконано
5	Проектування	13.04.2023 - 15.04.2023	виконано
6	Реалізація системи	16.04.2023 - 02.05.2023	виконано
7	Проведення тестування	02.04.2023 - 03.05.2023	виконано
8	Підготовка звіту до кваліфікаційної роботи	03.05.2023 - 06.05.2023	виконано
9	Перевірка на плагіат	08.05.2023	виконано
10	Перевірка на нормоконтроль	09.05.2023	виконано
11	Попередній захист	12.05.2023	виконано
12	Захист кваліфікаційної роботи	16.05.2023	виконано

Дата видачі завдання « 03 » квітня 2023 р.

Студент _____

(підпис)

_____ Кісельгова Є. М.

Керівник роботи _____

(підпис)

_____ проф. Смеляков С.В.

РЕФЕРАТ / ABSTRACT

Пояснювальна записка з роботи за темою дослідження «Дослідження методів попередньої обробки датасетів сигналів ЕЕГ» студента групи ПЗЗдм-21-1 Кісельгов Є.М., 74 стор., 33 рис., 4 табл., 26 джерел.

АРТЕФАКТ, АНАЛІЗ НЕЗАЛЕЖНИХ КОМПОНЕНТ,
ЕЛЕКТРОЕНЦЕФАЛОГРАМА, ФІЛЬТРАЦІЯ,

Об'єкт розробки – датасети сигналів ЕЕГ.

Мета розробки – проведення попередньої обробки сигналів ЕЕГ з метою пригнічення артефактів.

Метод рішення – автоматичне знаходження ділянок запису з можливими артефактами, декомпозиція їх на незалежні компоненти, виявлення артефактних складових у компонентах, їх пригнічення та зворотна композиція.

В роботі проведено аналіз існуючих методів попередньої обробки сигналів ЕЕГ з метою пригнічення артефактів та запропоновано рішення автоматичної попередньої обробки з усуненням артефактів з ЕЕГ сигналу.

ARTIFACT, BLIND SOURCE SEPARATION,
ELECTROENCEPHALOGRAM, FILTERING, INDEPENDENT COMPONENT
ANALYSIS

The object of study is a dataset of EEG signals.

The purpose of the work is to preprocess EEG signals with the aim to suppress artifact components in it, so cleaned EEG can be used for medical analysis.

The research method is performing automatic artifact range detections, ICA decomposition of this EEG ranges, detection of artifacts symptoms in component, suppress them and restore original EEG without artifacts.

The result of this scientific research practice is an ICA-based algorithm proposed for the automatic detection and cleaning of artifacts in EEG record.

Умови публікації пояснювальної записки

Я, Кісельгов Євген Миколайович

(прізвище, ініціали)

студент групи ІІЗЗдм-21-1 здобувач вищої освіти на другому (магістерському) рівнікафедра програмної інженерії

(повна назва кафедри)

заявляю: моя кваліфікаційна робота на тему «Дослідження методів попередньої обробки датасетів сигналів ЕЕГ», що буде представлена до ЕК для публічного захисту, виконана самостійно, в ній не містяться елементи плагіату і вона може бути опублікована в електронному архіві відкритого доступу EIArKhNURE. Всі запозичення з друкованих та електронних джерел мають відповідні посилання.

Я ознайомлений з діючим положенням «Про протидію академічному плагіату в ХНУРЕ», згідно з яким виявлення плагіату є підставою для відмови в допуску кваліфікаційної роботи до захисту та застосування дисциплінарних заходів.

ЗМІСТ

Вступ.....	8
1. Аналіз предметної галузі	10
1.1 Опис предметної галузі.....	10
1.2 Первинна обробка ЕЕГ сигналу	14
1.3 Виявлення проблеми очищення ЕЕГ сигналу.....	16
1.4 Аналіз підходів та методів рішення	17
1.5 Аналіз існуючих систем	19
1.6 Постановка задачі та функціональні вимоги.....	21
2. Дослідження та аналіз методів розробки.....	22
2.1 Аналіз незалежних компонент (ІСА).....	22
2.2 Знаходження артефактів у сигналах джерел	27
2.3 Джерела електродних артефактів	28
2.5 Джерела окулографічних артефактів	30
2.6 Виділення ділянок підозрілих на артефактну активність	33
2.7 Формулювання алгоритму вирішення задачі	35
3. Моделювання та програмна реалізація системи	36
3.1 Принципи роботи програмного продукту	36
3.2 Вимоги до функціональних характеристик сервісного модулю системи	37
3.3 Вимоги до інтерфейсу програмного додатку	37
3.4 Вимоги до алгоритму роботи й структури програми.....	39
3.5 Вимоги до інформаційної і програмної сумісності	39
3.6 Компоненти системи та діаграма пакетів.....	39
3.7 Декомпозиція системи на класи та діаграма послідовності алгоритму	41

	7
4. Планування та проведення експерименту	45
4.1 План проведення експерименту	45
4.1 Результати експерименту	45
4.1.1 Оцінка роботи системи по синтетичним даним.....	46
4.1.2 Оцінка роботи системи на реальному ЕЕГ датасету	50
4.2 Аналіз отриманих результатів	53
Висновки	55
Перелік джерел посилання	56
Додаток А Перелік джерел посилання за науковими напрямками керівника та науковців кафедри програмної інженерії	59
Додаток Б Апробація роботи	60
Додаток В Слайди презентації.....	64
Додаток Г Звіт результатів перевірки на унікальність тексту в базі ХНУРЕ.....	72
Додаток Д Експертний висновок результатів перевірки кваліфікаційної роботи на відповідність оформлення вимогам ДСТУ 3008: 2015	73
Додаток Е Листинг коду	74

ВСТУП

Ціллю роботи є дослідження та аналіз методів попередньої обробки сигналів електроенцефалограми (ЕЕГ) та проектування програмного забезпечення здатного виконувати таку обробку. Аналіз якості проведеної обробки.

Електроенцефалографія – це метод реєстрації електричної активності головного мозку, відведеної з поверхні шкіри голови, зі скальпу голови. Така електрична активність має назву – біопотенціали головного мозку. Її реєстрація є одним з поширених методів функціональної діагностики стану людини [1, 2]. Вона дозволяє проаналізувати фізичний стан і зрілість головного мозку, виявити на попередньому етапі наявність осередків уражень мозку, загальмований розлад тощо [3]. Можливість запису електричної активності мозку з поверхні скальпу було вперше продемонстровано у дослідних роботах В.В. Правдич-Немінського (1913р. лабораторія фізіології Київського Університету св. Володимира) та у роботах німецького психіатра Ганса Берга який вперше зареєстрував електроенцефалограму людини (ЕЕГ) у 20х роках ХХ сторіччя.

Методи реєстрації ЕЕГ постійно розвивалися та вдосконалювалися. У сучасному світі найпоширеніший метод – реєстрація за допомогою програмно-апаратних комплексів на базі персонального комп'ютеру. Цифрова ЕЕГ отримана таким чином є багатоканальний сигнал біопотенціалів мозку зареєстрованих зазвичай у різних станах пацієнта: спокій да деякі види когнітивного або подразнюючого навантаження, що формує датасет досить великого об'єму якій треба обробити та проаналізувати. Сучасні комплекси здатні якісно провести накопичення сигналу (реєстрацію) запобігаючи появі у запису електричних та гальванічних завад. Але комплекси не здані усунити причину появи фізичних перешкод у запису – самого пацієнта. Навіть спокійному пацієнту важко або повністю розслабитися або сконцентруватися у потрібні моменти реєстрації ЕЕГ та уникнути зайвих рухів або скорочень м'язів. Навіть незначне переміщення пацієнта або рух очних м'язів призводить до появи перешкод у ЕЕГ.

Такі перешкоди у EEG датасетах мають назву «артефакти». Артефакти у записах з'являються досить часто, тим шкідливіший вони коли у часі співпадають з моментом появи патологічної або корисної активності мозку яку важко повторити без повторення запису з початку. Зазвичай фізіолог, лікар-функціоналіст (лікар функціональної діагностики) або інша людина яка проводить запис EEG витрачає відносно багато часу у спробах усунути артефакти, тобто проводить складні операції попередньої обробки з очищення датасетів.

Мета роботи – дослідити та проаналізувати існуючі методи очищення EEG датасетів та створення програмного забезпечення яке дозволяє полегшити цей процес або повністю його автоматизувати.



Рисунок 1 – Реєстрація EEG

1. АНАЛІЗ ПРЕДМЕТНОЇ ГАЛУЗІ

1.1 Опис предметної галузі

Електроенцефалографія – реєстрації електричної активності головного мозку (біопотенціалів) з поверхні скальпу голови у стані спокою або функціонального навантаження пацієнта. Отриманий багатоканальний цифровий сигнал має назву електроенцефалограми (ЕЕГ) [1]. Реєстрація цифрового ЕЕГ проводиться за допомогою комп'ютерних електроенцефалографічних комплексів які складаються з кількох блоків:

- блок реєстрації сигналу,
- блок фільтрів та підсилювач,
- аналого-цифровий перетворювач (АЦП),
- комп'ютерна програма реєстрації, зберігання та обробки ЕЕГ.

Реєстрацію біопотенціалів мозку проводять з накладанням на поверхню скальпу металевих електродів з хлорсрібним покриттям. Електроди розташовані на поверхні скальпу таким чином щоб охопити усю доступну зону реєстрації з кори мозку. У деяких випадках наукового дослідження електроди вживлюють у поверхню кори мозку. Цифрова ЕЕГ є датасетом який сформований з наступних реєстрацій [2]:

- ЕЕГ у стані спокою пацієнта,
- ЕЕГ у стані функціонального навантаження пацієнта:
 - фото стимуляція: подразнення фотоспалахами різної частоти та тривалості,
 - звукова стимуляція: подразнення звуками різної частоти та тривалості,
 - гіпервентиляція: стимуляція кисневим голодуванням мозку шляхом гіпервентиляції легень.
- ЕЕГ у стані когнітивного навантаження пацієнта (викликані потенціали ВП): зорові, звукові, когнітивні Р300, біологічний зворотній зв'язок та багато інших видів [4, 5].

- Т – скроневі зони,
- С – центральна зона,
- Р – тім'яна зона
- О – потилична (окципітальна) зона.

Усі варіації схеми «10-20» - це так звані «монополярні» схеми оскільки реєструють потенціал сигнал відносно базового «референтного» електроди який зазвичай розташовують біля вуха з обох боків або на щелепі. Інший варіант реєстрації – «біполярний», коли потенціал електроду реєструється відносно сусіднього електроду або симетричного йому електроду у радіальній площині на скальпі. Оскільки будь-яка біполярна схема може бути математично розрахована з повної монополярної схеми то зазвичай для реєстрації вони не використовуються, а розраховуються при необхідності.

ЕЕГ є гомогенним цифровим сигналом малої амплітуди (у стані спокою від декількох мікрвольтів до десятків мілівольтів) зареєстрованим з постійною високою частотою дискретизації від 128 Гц до 1-2 КГц. Так звана корисна частина ЕЕГ – це біосигнал згенерований корою мозку зазвичай поділяють у частотній площині на такі діапазони:

Таблиця 1 – Частотні діапазони ЕЕГ коливань

Діапазон	Частота	Стан
дельта	до 4 Гц	коливання амплітудою 20-30 мкВ можуть зустрічатися у ЕЕГ здорової притомної людини; наявність коливань більш високої амплітуди (40-300мкВ) у ЕЕГ притомної людини є патологічною ознакою (мозкові пухлини); дельта-коливання стають вираженими під час фаз природного сну, наркотичного сну або у стані коми

Кінець таблиці 1

Діапазон	Частота	Стан
тета	4-8 Гц	коливання амплітудою до 40 мкВ можуть зустрічатися у ЕЕГ здорової притомної людини, зростання їх частки є ознакою емоційної активації та інших типів мозкової активності; наявність тета-коливань у більших кількостях пов'язана із патологічними станами або ж зміненими станами свідомості (сон, медитація та ін.)
альфа	8-13 Гц	синусоїдальні коливання амплітудою до 100 мкВ, амплітуда яких зростає у лобно-потиличному напрямку, є найбільш вираженим у ЕЕГ здорової притомної людини із закритими очима, у формі вираженого ритму реєструється у 80-90 % людей, пригнічується при відкриванні очей, переході до активної діяльності, аналізу інформації
бета	13-35 Гц	коливання амплітудою 5-30 мкВ, наявність яких у ЕЕГ пов'язана із активним когнітивним функціональним станом мозку, зростання рівня активації головного мозку здебільшого супроводжується зменшенням частки альфа-коливань і зростанням частки бета-коливань; з віком частка альфа коливань зменшується, а бета – зростає
гамма	30-100 Гц	вважається ознакою когнітивних процесів і свідомості, вище 15 мкВ - патологічна активність

Довжина одного запису датасету дуже варіюється від виду навантажень, якщо у стану спокою це може бути біля 40-60 секунд, то в стані гіпервентиляції або когнітивної стимуляції – до 10 хвилин. Існують моніторингові методи запису ЕЕГ (протягом значного часу) коли запис проводиться протягом годин. Такі коливання часу реєстрації потребують обробки відносно великого об'єму даних коли датасет у завантаженому стані може займати декілька гігабайт оперативної пам'яті. Що

потребує наявність засобів прискорення або автоматизації первісної обробки даних.

1.2 Первинна обробка ЕЕГ сигналу

Сучасні ЕЕГ реєстратори здатні записати цифровий сигнал ЕЕГ без електричних завад або з мінімальною їх кількістю, що вже є досить складним завданням оскільки амплітуда корисного сигналу ЕЕГ зазвичай у 100-1000 разів менша ніж завади електромереж. Але мала амплітуда ЕЕГ та значне підсилення сигналу у АЦП реєстратора робить увесь комплекс вкрай чутливим до перешкод які генеруються самим пацієнтом. Навіть незначні рухи, м'язові скорочення лица або зовсім непомітне коливання погано закріпленого електроду призводить до появи перешкод які досить сильно спотворюють сигнал ЕЕГ.

Такі перешкоди мають назви артефактів та у більшості випадків поділяються на наступні види.

Електродний артефакт. Пов'язаний з погано закріпленням електродом, висиханням контактної гелю, порушенням роботи хлорсрібного покриття. За частотою співпадає з альфа, бета та гамма діапазонами ЕЕГ. Зазвичай призводить до пошкодження усього сигналу зареєстрованого з електроду.

Артефакт референтного електроду. Має ті ж самі причини як і звичайний електродний артефакт але оскільки проявляється у референтному електроді монополярної схеми – пошкоджує сигнал зареєстрований з усіх електродів.

Окулографічний артефакт (або зоровий артефакт). Пов'язаний з роботою очних м'язів або блимання. Найбільше проявляється у фронтальних та центральних електродах має високу амплітуду, амплітуда спадає від фронтальних до потиличних електродів. У частотному діапазоні співпадає з більшістю ЕЕГ діапазонів що вкрай затрудняє його позбавлення.

Міографічний артефакт ковтання. Високочастотний артефакт пов'язаний з роботою м'язів голови та шиї.

Кардіо артефакт. При близькому розташуванні кровоносних судини до електродів можливі випадки прояви кардіограми у ЕЕГ запису. Зустрічається рідко.

Артефакти дихання. Глибоке дихання пацієнта може призводити до незначних зміщень потенціалів ЕЕГ в сусідніх групах електродів. Частотно співпадає з коливаннями у тета діапазоні. Може бути усунений фільтрацією.

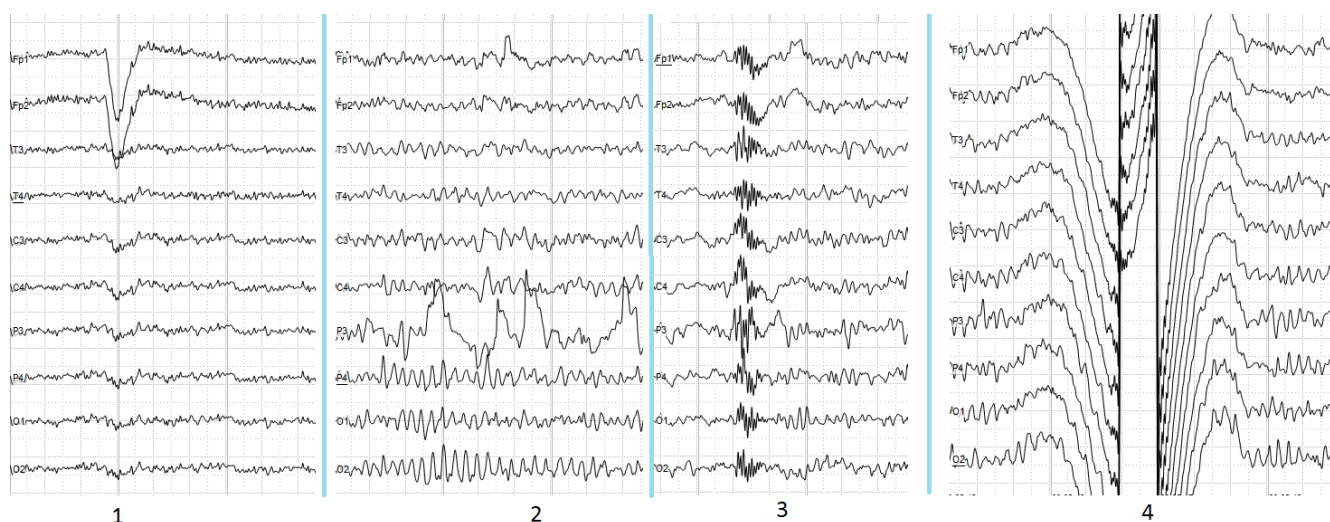


Рисунок 4 – Приклади артефактів: 1 – окулографічний артефакт, 2 – електродний артефакт у R3, 3 - міографічний артефакт ковтання, 4 – артефакт референтного електроду.

Отже первинна обробка ЕЕГ складається з етапів:

- фільтрація ЕЕГ для усунення електричних завад електромережі (50 Гц або 60 Гц). Зазвичай виконується автоматично програмним забезпеченням комплексу реєстрації або самою апаратурою реєстрації,
- очищення ЕЕГ:
 - перегляд та знаходження артефактів,
 - усунення артефактів.

Після первинної обробки ЕЕГ може бути переданий засобам автоматичного або ручного аналізу ЕЕГ для виявлення патологічної активності, або у разі викликаних потенціалів мозку – виявлення когнітивної активності ділянок мозку.

1.3 Виявлення проблеми очищення ЕЕГ сигналу

Коли ЕЕГ датасет складається з декількох записів великого об'єму знаходження артефактів може займати значний час у випадку ручного перегляду усього запису. Наприклад у середньому час запису одного звичайного функціонального неврологічного дослідження складає 20 хв. фонові ЕЕГ + ЕЕГ функціонального навантаження [8]. Зазвичай лікар-функціоналіст витрачає у майже у двічі більше часу на перегляд запису для знаходження та усунення артефактів у ручному режимі, це додаткові 30-40 хв. Усі обробка ЕЕГ у цьому випадку може займати до години. Якщо запис проводиться у моніторинговому режимі тобто декілька годин, попередня обробка, тобто очищення ЕЕГ сигналу займе дуже багато часу. Очевидно що ця операція потребує автоматизації, критерії якої дуже складно сформулювати якщо знаходження артефактів проводити по вихідному ЕЕГ сигналу без спеціальної додаткової математичної обробки.

Зі статистики ЕЕГ досліджень знаємо що артефакти різних типів мають різну вірогідність появи у ЕЕГ сигналі. Для більшості з них можна мінімізувати вірогідність появи грамотним накладанням електродів та переглядаючи ЕЕГ сигнал в режимі реального часу до моменту коли буде натиснута кнопка «почати запис». Тобто гарний досвід лікаря-функціоналіста є запорукою низької вірогідності появи таких артефактів як кардіо, артефакт дихання, постійний електродний артефакт (поганий контакт з поверхнею, «сухий» електрод) тощо.

Але існує три види артефактів для яких це твердження невірне:

- блукаючий електродний артефакт (коливання електроду внаслідок руху пацієнта)
- артефакт референтного електроду (коливання референтного електроду внаслідок руху пацієнта)
- окулографічний артефакт (синхронний з функціональним навантаженням пацієнта, а тому і з корисними ділянками запису).

Саме виявлення цих артефактів та очищення ЕЕГ сигналу від них і є головною проблемою аналізу методів рішення якою присвячена ця робота.

1.4 Аналіз підходів та методів рішення

Як було зазначено при опису артефактів, більшістю вони у частотному діапазоні співпадають з корисним ЕЕГ сигналом, що не дає змогу виділяти їх методами спектрального аналізу (перетворення Фур'є).

Методи статистичного аналізу дуже залежать від амплітуди та виду артефакту оскільки після попередньої фільтрації математичне сподівання корисного сигналу співпадає з математичним сподіванням артефактів. Дисперсія артефактів дуже відрізняється від дисперсії корисного ЕЕГ сигналу, але не у випадках патологічної активності, коли з'являються сплески ЕЕГ ритмів високої амплітуди. А саме таку патологічну активність слід виявити у результаті аналізу ЕЕГ і буде дуже погано якщо така активність буде «очищена» як прояв артефакту.

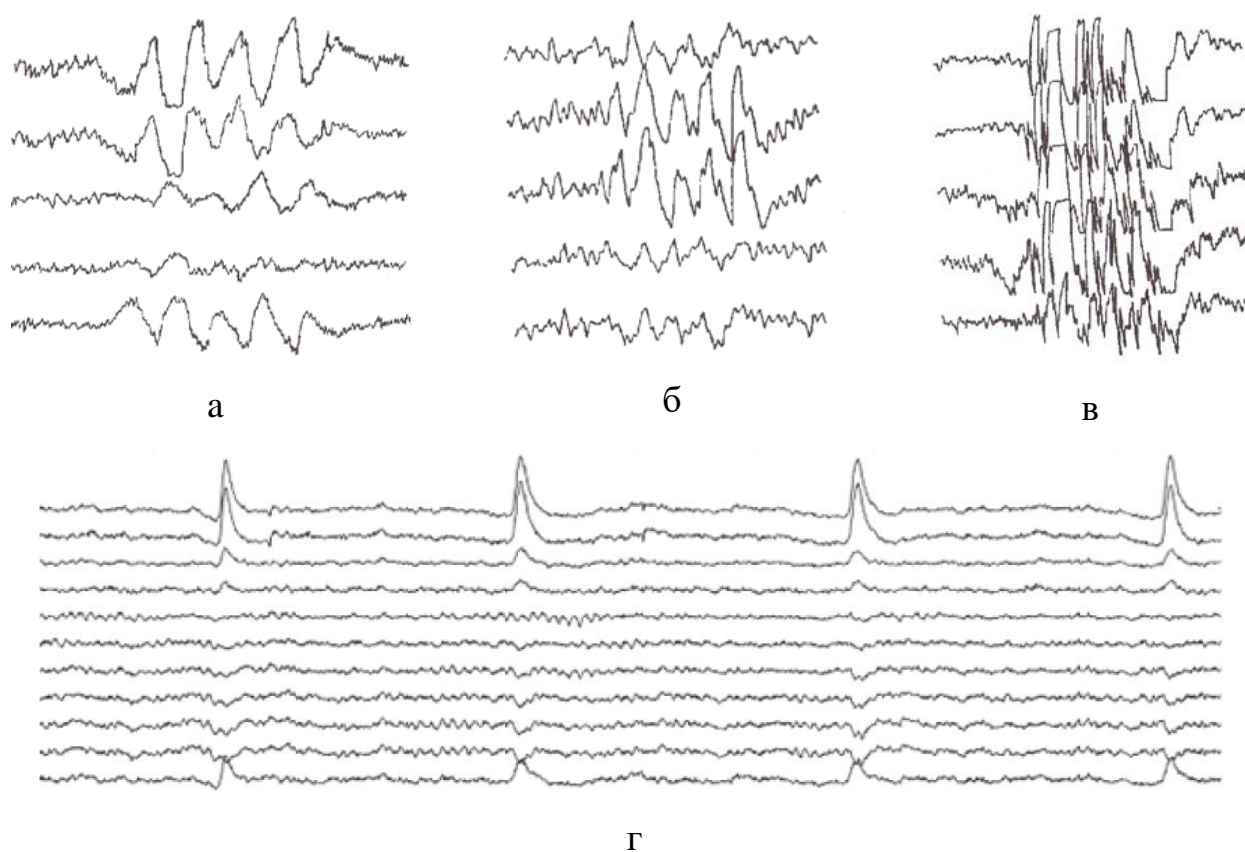


Рисунок 5 – Порівняння патологічної активності (а – спалах дельта хвиль, б – спалах тета хвиль, в – спайкова активність) та оклугорафічних артефактів (г).

Найбільш поширена практика - застосування методів контурного аналізу. Але ці методи дуже залежать від амплітудно/часового співвідношення артефакту

та корисного сигналу. Наприклад окулографічний артефакт якій змінює свій вплив на ЕЕГ по мірі віддалення від фронтальної зони дуже складно виявити у потиличних електродах [9]. Але контурний аналіз досі залишається одним з успішних способів виявлення якщо границі ділянок сигналу підозрілого на артефакти були попередньо знайдений за допомогою адаптивних моделей та алгоритмів пошуку країв які наприклад широко застосовуються при обробці зображень [10].

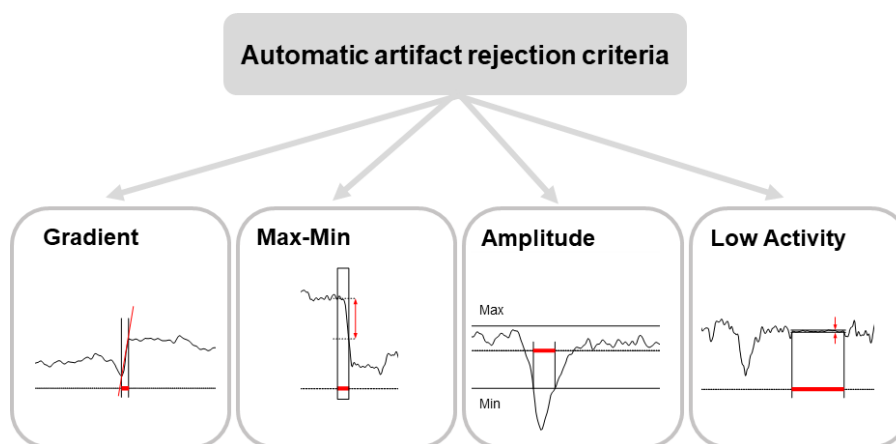


Рисунок 6 – Методи контурного аналізу у BrainVision Analyzer 2

Гарні результати для виявлення окулографічного артефакту показують методи вейвлет-перетворення. Завдяки досить стабільній формі цього артефакту він може бути знайдений та видалений за допомогою вейвлету Хаара або Добеши [11].

Методи імовірнісного аналізу які засновані на імовірнісному розділенні ЕЕГ сигналу найбільш перспективні на поточний момент [9, 12, 13]. Серед них виділяють метод головних компонент (PCA) та метод аналізу незалежних компонент (ICA) [14]. Останній метод найбільш популярний оскільки дозволяє не лише проводити пошук артефактів а також вирішувати так звану «зворотню задачу ЕЕГ» [15].

Методи нейронних мереж (штучний інтелект) є багатообіцяючими але поки що не застосовуються широко у ЕЕГ. Методи класифікації звичайних зображень на основі згорткових нейронних мереж [16] мають значні перспектив і у знаходженні

артефактів у ЕЕГ, але поки не знайдено модель яка б дозволила видалити артефакту складову з ЕЕГ сигналу.

Описані вище методи можуть дати відповідь де знаходиться артефактна ділянка описати клас артефакту, але не можуть видалити сам артефакт (окрім частково рішення оснований на вейвлет-перетворені та методів РСА та ІСА про що буде далі). Тому після знаходження артефакту найбільш розповсюджений спосіб його усунення – помітити ділянку запису як мусорну (артефактну) та у подальшому аналізі ігнорувати її віртуально склеюючи запис як би артефактної ділянки не було взагалі. Цей метод має назву – розмітка ЕЕГ запису. Головний його недолік очевидний з природи артефактів: більшість ЕЕГ записів – це записи з навантаженнями пацієнта, а такі артефакти як електродні та окулографічні, проявляються як раз у часи навантаження. У результаті:

- щоб зменшити співвідношення артефактних ділянок до загальної довжини запису її подовжують,
- чим більша тривалість запису – тим більша втома пацієнта, менша його концентрація на якості запису – вища вірогідність появи нового артефакту.

Це замкнуте коло зводить нанівець зусилля у спробі отримати якісний ЕЕГ запис за короткий проміжок часу. Потрібен інший метод який дозволить не лише знайти артефакт, але й видалити його – очистити запис від артефактів.

1.5 Аналіз існуючих систем

Як правило досить поширене явище коли лікар-функціоналіст проводить первину обробку самотужки по усьому датасету або корегуючи результат автоматичного пошуку артефактів. Чим краще автоматичний пошук знаходить артефакти – тим швидший буде результат. Якщо програмна реалізація ЕЕГ комплексу дозволяє не тільки знайти але й видалити артефакт з запису – це додає їй конкурентної переваги у порівнянні з іншими комплексами.

Серед широкого спектру існуючих систем слід виділити такі що дозволяють автоматизувати або значно полегшити пошук та видалення артефактів, розглянемо їх.

Програмно-апаратні комплекси функціональної діагностики «Нейроком» та «Нейролаб», розробник – «ХАІ-Медика» [17]. Комплекс реалізує автоматичну розмітку артефактних ділянок запису. Дозволяє аналізувати їх за ICA алгоритмом, полегшує ідентифікацію та видалення артефакту за допомогою зворотної ICA композиції.

Програмний застосунок BrainVision Analyzer 2, розробник Brain Products [18]. Реалізує знаходження та видалення артефактів на основі статистичних методів та контурного аналізу EEG сигналу. Підтримує ICA декомпозиції та зворотну композицію у ручному режимі.

EEGLab. Це Matlab toolbox розроблений Swartz Center for Computational Neuroscience UCSD [19]. Додаток містить найбільш сучасні методи обробки EEG сигналу: ICA, PCA, фільтри, частотний аналіз, карти різним видів та інше. Головне функціональне навантаження додатку – дослідницька діяльність у сфері EEG та ER.

BESA Research, розробник BESA GmbH [20]. Професійне пропріетарне програмне забезпечення яке дозволяє автоматично виділити ділянки запису які містять артефакти базуючись на контурному аналізі або пошуку за шаблонами. Окулографічні артефакти можуть бути видалені після знаходження за допомогою PCA алгоритму.

Переглянуті програмні системи є або спеціалізованими для наукових досліджень або досить кошовними системами які потребують певної кваліфікації для застосування їх у якості систем автоматичного очищення EEG датасетів від артефактів. Спеціалізованої програмної системи яка б працювала з EEG датасетами у відкритих стандартах та автоматично вирішувала б поставлену задачу не існує.

1.6 Постановка задачі та функціональні вимоги

Виходячи з описаної вище проблематики очищення ЕЕГ датасету від артефактів можливо сформулювати задачу та вимоги до програмної системи наступним чином.

Головним завданням системи – є здатність виділити ділянки запису які містять ймовірний артефакти. Слід зосередитися на головних видах артефактів які треба знаходити: електродний артефакт, артефакт референтного електроду, окулографічний артефакт. Ділянки ЕЕГ датасету повинні бути проаналізовані та у разі виявлення артефактів – видалити їх з ЕЕГ запису. Скорегованих ЕЕГ сигнал – є результатом роботи системи.

Функціонально система повинна працювати в автоматичному та напівавтоматичному режимі (ручний прохід кожного з етапів виявлення та видалення артефактів з допомогою системи). Система приймає на вхід ЕЕГ сигнал з датасету у одному з відкритих стандартів, обробляє його та повертає очищений ЕЕГ сигнал разом зі звітом які саме артефакти були знайдені та видалені, з якою вірогідністю та на яких ділянках запису були розташовані.

Автоматичний режим роботи системи працює за принципом чорної скриньки: запуск програмного додатку з параметром який містить посилання на ЕЕГ датасет для аналізу, результат – очищений ЕЕГ сигнал та звіт.

У напівавтоматичному режим система повинна реалізувати інтерфейс взаємодії з користувачем який дозволяю провести усі операції виявлення та видалення артефактів самостійно у ручному режимі.

2. ДОСЛІДЖЕННЯ ТА АНАЛІЗ МЕТОДІВ РОЗРОБКИ

2.1 Аналіз незалежних компонент (ICA)

Як було зазначено у попередньому розділі при аналізі існуючих систем, найбільш просунути програмні системи обробки ЕЕГ сигналу застосовують алгоритм незалежних компонент (ICA) як один з засобів обробки ЕЕГ. Головним чином – для знаходження локалізації (міста розташування) джерел патологічної активності у головному мозку. Але цей метод можна застосовувати для знаходження та пригнічення артефактів, він виконує роль просторового фільтру який працює у амплітудно-часовому вимірі.

Метод аналізу незалежних компонент ICA був розроблений у 90х роках ХХ ст. як засіб вирішення задачі «сліпого» розділення джерел. Класичним прикладом якого є «ефект коктейльної вечірки» – здатність мозку людини зосередитись на сприйнятті інформації лише з одного джерела в умовах одночасної реєстрації сигналів з декількох різних джерел навіть коли вони мають однакову з головним або більшу потужність сигналу.

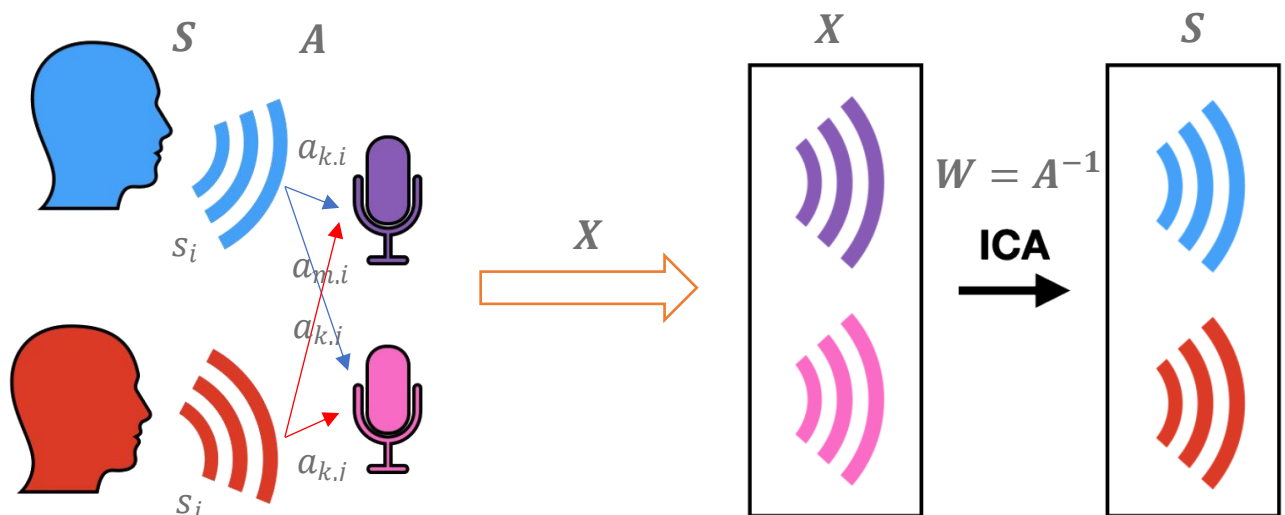


Рисунок 7 - Композиція джерел та «сліпе» розділення джерел за ICA

Кожен з учасників «коктейльної вечірки» є джерелом (наприклад за номерами i та j) власного сигналу (s_i та s_j). Також існують реєстратори сигналів джерел (k або m) це мікрофони або вуха людини як у прикладі вечірки або ж у

нашому варіанті – реєстратори ЕЕГ сигналів (електроди). Усі джерела змішуються на кожному реєстраторі за невідомим але лінійним законом формуючи на кожному реєстраторі сигнали суміші:

$$x_k = a_{k,i} * S_i + a_{k,j} * S_j$$

$$x_m = a_{m,i} * S_i + a_{m,j} * S_j$$

де a – невідомі коефіцієнти змішування сигналів джерел які обумовлені у ЕЕГ просторовою провідністю мозку людини та зовнішніх тканин.

Ці сигнали суміші і є багатоканальний вихідний запис який розглядається як суперпозиція декількох джерел сигналів з невідомим законом змішування що можна записати у матричному виді:

$$\mathbf{X} = \mathbf{A} \cdot \mathbf{S}$$

де \mathbf{X} – матриця вихідного сигналу ЕЕГ (суміш джерел), сигнал кожного електроду це окремий рядок матриці,

\mathbf{S} – матриця джерел, за розміром так ж як матриця \mathbf{X} ,

\mathbf{A} – квадратична матриці закону змішування, кожен стовпець матриці – вагові коефіцієнти з якими джерело з матриці \mathbf{S} та номером стовпця матриці \mathbf{A} змішується у вихідний сигнал ЕЕГ \mathbf{X} .

«Сліпе розділення» це знаходження матриці розділення \mathbf{W} оберненої до матриці \mathbf{A} яка відновлює матрицю джерел з матриці суміші:

$$\mathbf{S} = \mathbf{W} \cdot \mathbf{X}$$

Метод ІСА – це узагальнення методу РСА, який вирішують задачу «сліпого» розділення, а мірою оцінки розділення вихідної матриці \mathbf{X} на джерела є – повна статистична незалежність сигналів \mathbf{S} . У той час як РСА можна розглядати як

рішення задачі «сліпого» розділення коли мірою розділення слугує матриця декореляції сигналів.

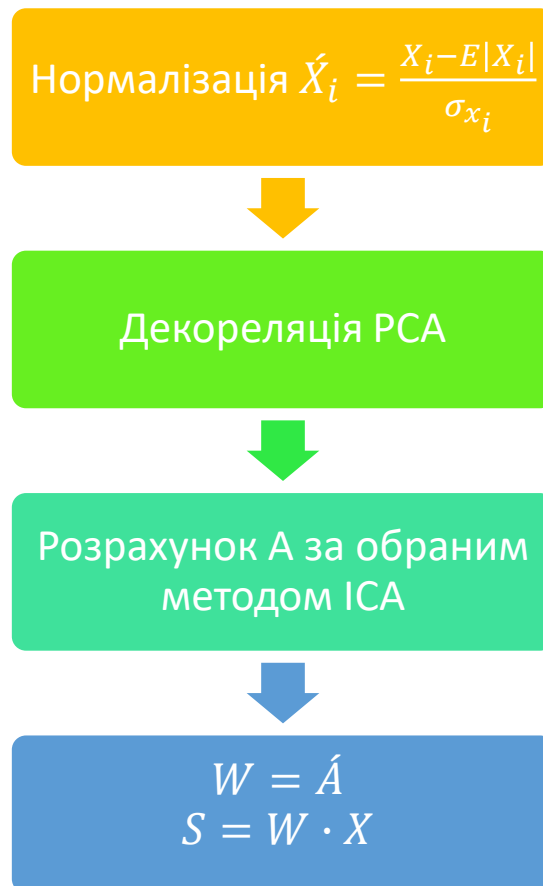


Рисунок 8 – Послідовність рішення задачі розділення X на S .

Існує декілька способів оцінки статистичної незалежності сигналів джерел:

- мінімізація взаємної інформації (максимізація ентропії)
- максимізація негаусовості сигналів (негетропія, ексцес, емпіричні оцінки як $\log \cosh(x)$).

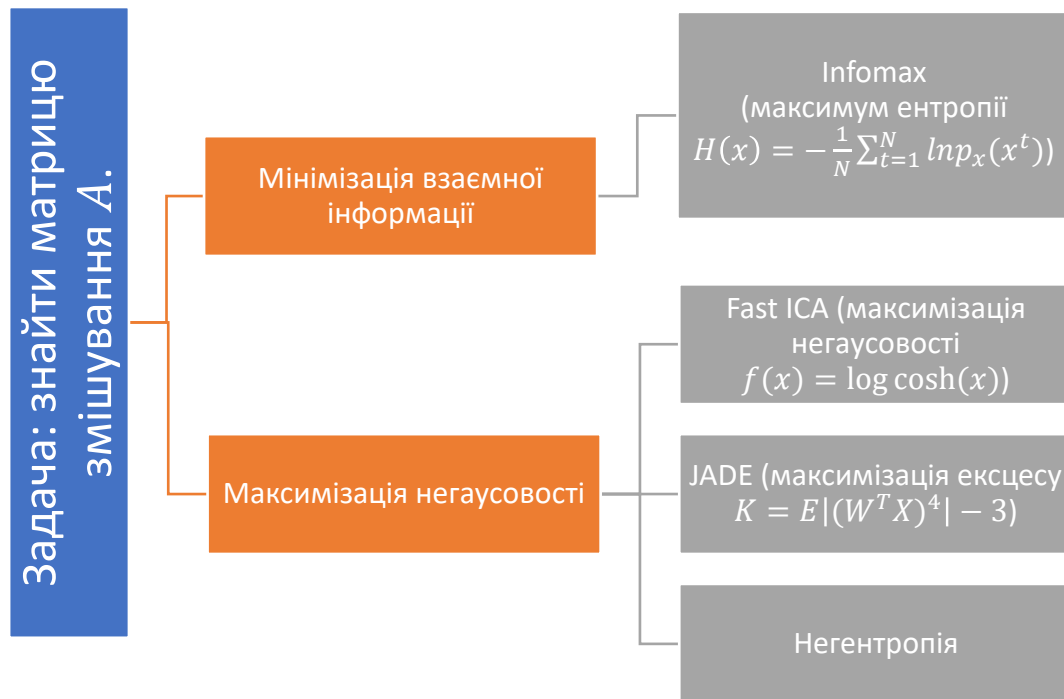


Рисунок 9 – Види ІСА

Як показали у своїх роботах Ті-Ван Лі, Т. Янг, Тони Белл, Теренс Сейновські [12, 13, 14] сигнали електродних, окулографічних та міографічних артефактів мають інші статистичні закони щільності розподілення вірогідності ніж фонова активність ділянок мозку ЕЕГ, а отже можуть бути виділені ІСА алгоритмом у окремі джерела (компоненти). Це стало початком застосування ІСА у багатьох задачах ЕЕГ.

Виділення артефакту у окремий компонент значно покращує співвідношення сигнал-фонова ЕЕГ активність що дозволяє усунути проблеми застосування статистичних та контурних методів у аналізі ЕЕГ. Крім того ІСА додає просторову інформацію для кожної компоненти ЕЕГ сигналу у вигляді карти вагових коефіцієнтів що дозволяє локалізувати зону артефакту.

Оберненість лінійного перетворення ІСА (здатність до просторової фільтрації) дозволяє «зібрати» ЕЕГ з компонент викинувши артефакти, що усуває всі недоліки методу розмітки ЕЕГ:

ICA decomposition

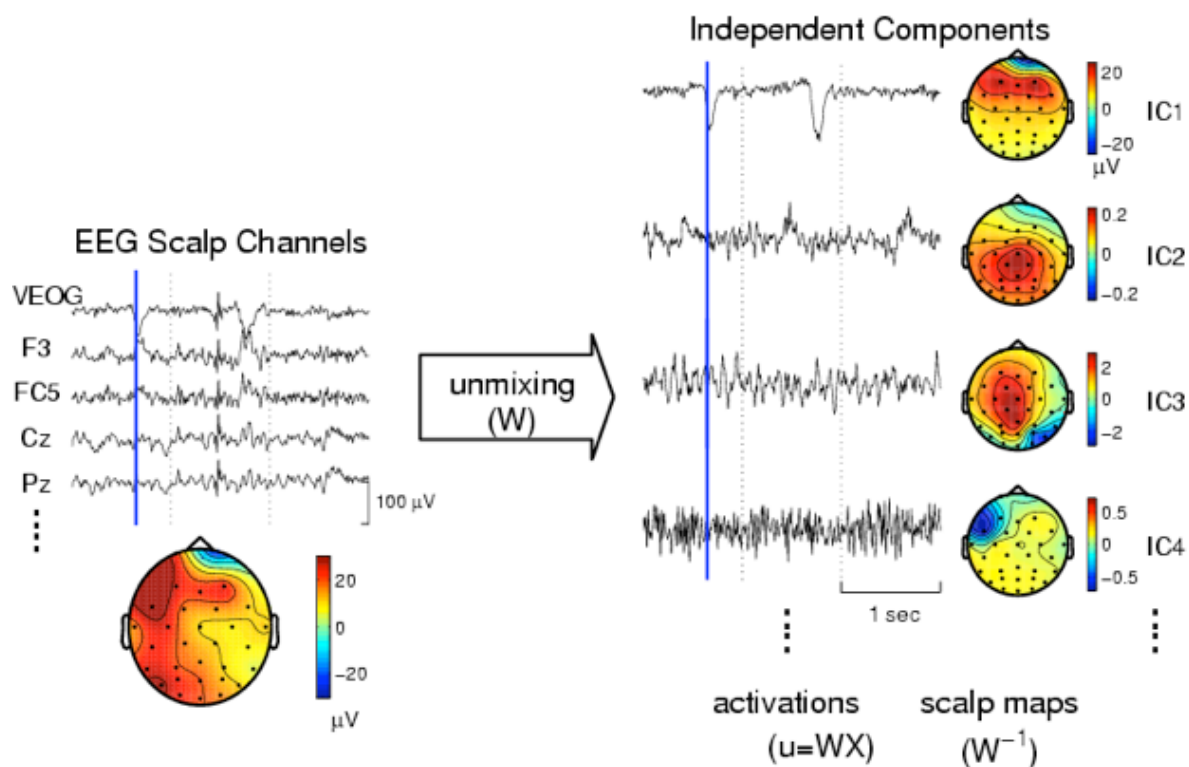


Рисунок 10 – Фрагмент ІСА декомпозиція EEG запису: IC1 - окулографічний артефакт, IC4 – електродний артефакт.

Summed Projection of Selected Components

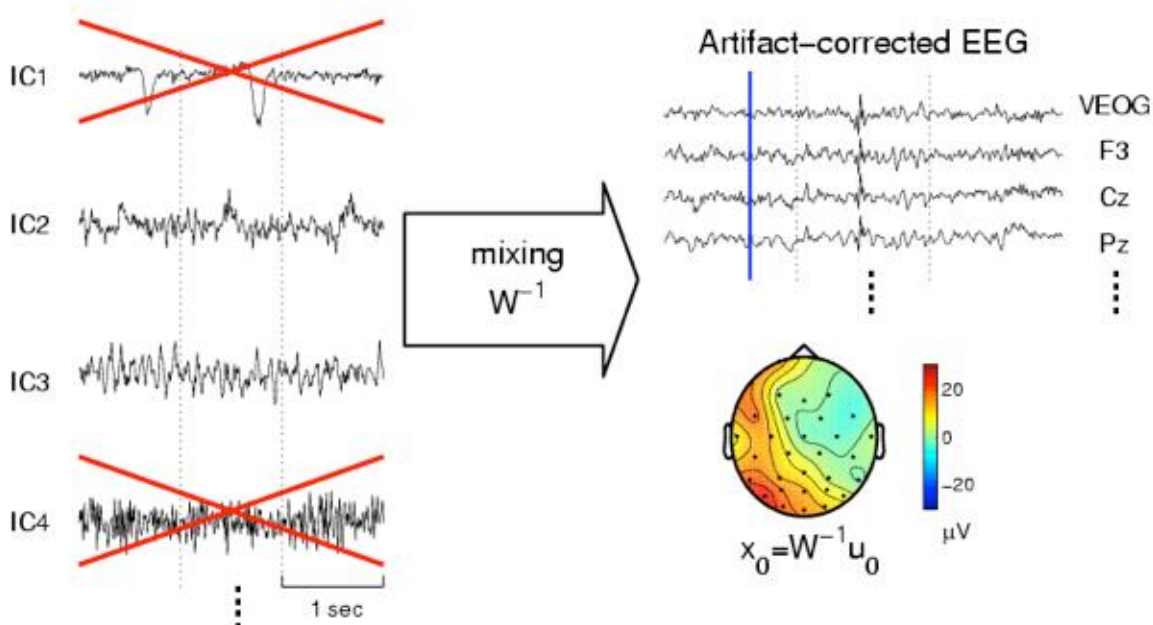


Рисунок 11 – Зворотна композиція компонентів у EEG без урахування артефактів

2.2 Знаходження артефактів у сигналах джерел

Отже виконувачи ІСА декомпозицію і отримавши сигнали джерел \mathbf{S} ми можемо відновити ЕЕГ сигнал без артефактів, тобто провести очищення. Але залишається виявити які саме сигнали джерел є артефактами а які ні.

Можемо застосовувати будь які методи оскільки після розділення ЕЕГ суміші на джерела у кожному з них маємо найкраще співвідношення сигналу джерела до залишків суміші ЕЕГ або завад. З точки зору високої продуктивності краще застосовувати:

- методи контурного аналізу сигналу,
- методи просторової локалізації сигналу джерел.

Звернімо увагу на просторову локалізацію. Отже для кожного джерела s_i після ІСА декомпозиції маємо співвіднесений до нього вектор-стовпець з матриці змішування \mathbf{A} :

$$\bar{a}_i = \begin{bmatrix} a_{i,1} \\ a_{i,2} \\ \dots \\ a_{i,m} \end{bmatrix}$$

де $a_{i,m}$ – ваговий коефіцієнт змішування сигналу джерела s_i у суміші на електроді з номером m ,

m – кількість електродів у суміші \mathbf{X} .

Тобто кожен ваговий коефіцієнт джерела має співвіднесений з ним електрод суміші, а знаючи просторову локалізацію всіх електродів у схемі «10-20» ми можемо будувати карти топографічного розподілення сигналу джерела на поверхні скальпу, тобто ЕЕГ карти джерел. Топографічні карти дають можливість візуально та алгоритмічно знайти просторову локалізацію джерела та зробити висновок про його «природу».



Рисунок 12 – Приклад ЕЕГ карт для джерел за максимальними ваговими коефіцієнтами у тім'яної зоні (а) та окципітальній зоні (б)

2.3 Джерела електродних артефактів

Виходячи з природи звичайного електродного артефакту ми знаємо що він проявляється лише в одному з каналів суміші, а отже його топографічна карта локалізації вагових коефіцієнтів має досить простий вид:

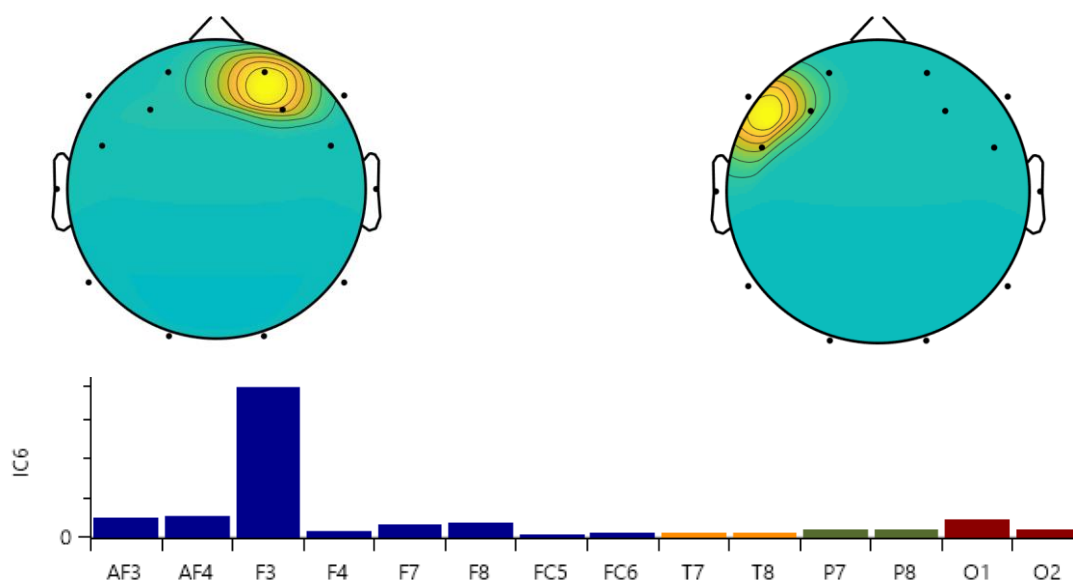


Рисунок 13 – Приклад ЕЕГ карт локалізації електродних артефактів та діаграми вагових коефіцієнтів джерела (для останньої карти).

Для знаходження такої локалізації пропонується побудувати для кожного електроду модельне розподілення (сигнатура артефакту) – вектор-стовпець \widehat{a}_k з ваговими коефіцієнтами

$$\widehat{a}_{k,j} = \begin{cases} 1; & k = j \\ 0.001; & k \neq j \end{cases}$$

де k – індекс електроду для якого будується модель,
 j – індекси інших електродів від 0 до m .

Якщо вектор-стовпець вагових коефіцієнтів джерела \bar{a}_l є «подібний» до сигнатури (моделі) якогось артефакту \widehat{a}_k , тоді це джерело – сигнал електродного артефакту. В якості міри «подібності» будемо застосовувати кореляційний коефіцієнт Пірсона, значення $r_{\widehat{a}_k, \bar{a}_l} \geq 0.9$ позначає високу достовірність артефакту:

$$r_{\widehat{a}_k, \bar{a}_l} = \frac{cov_{\widehat{a}_k, \bar{a}_l}}{\sigma_{\widehat{a}_k} \cdot \sigma_{\bar{a}_l}}$$

Знаходження артефакту референтного електроду є подібним до електродного артефакту в одному каналі з тієї різниці що артефакт в референтному електроді – це джерело яка вносить однаковий вклад в усі електроди, а отже розподілення його модельних коефіцієнтів (сигнатура) \widehat{a}_k – майже константа

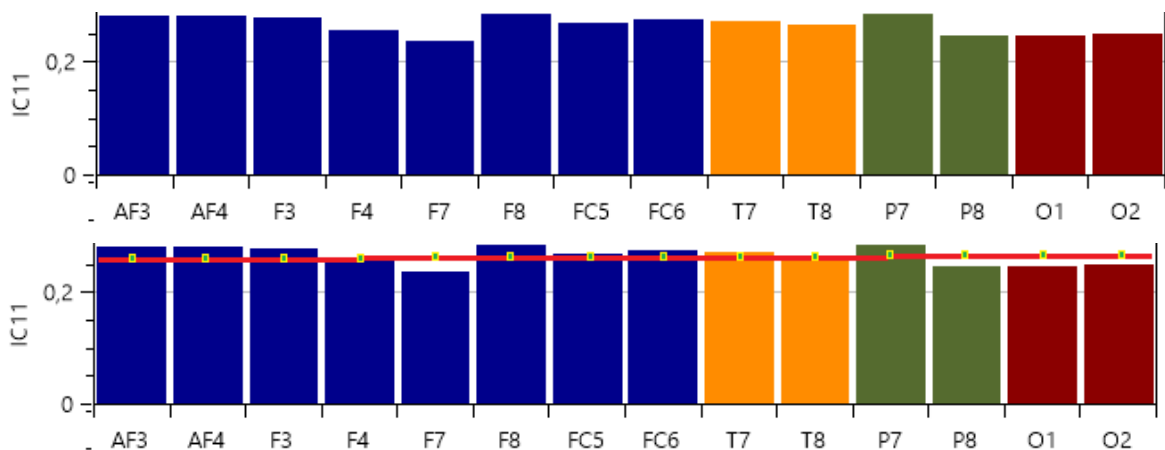


Рисунок 14 – Діаграма вагових коефіцієнтів джерела артефакту у референтному електроді без та з нанесеною лінійною регресією.

Для знаходження такої локалізації пропонується побудувати лінійну регресію по ваговим коефіцієнтам діаграми:

$$F(\bar{a}_i) = a_0 + a_1 * j$$

де a_0 та a_1 – коефіцієнти регресії,

j – номер електроду у діаграмі (від 1 до m).

Якщо за регресійна модель достовірна за t -статистикою Стюдента або фактором значущості коефіцієнтів, та кут нахилу регресії за модулем не більше 1° :

$$\arctg(a_1) \leq 1^\circ$$

тоді це джерело – сигнал електродного артефакту у референтному електроді.

2.5 Джерела окулографічних артефактів

Розподілення вагових коефіцієнтів джерела окулографічного артефакту не має такого простого закону як для електродних артефактів:

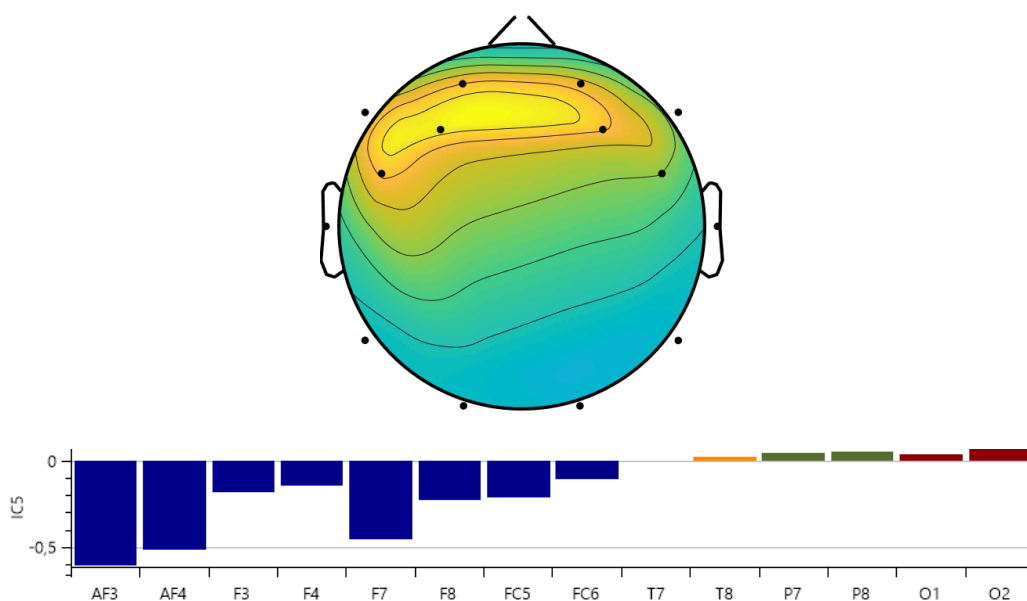


Рисунок 15 – Карта локалізації окулографічного артефакту та діаграма її вагових коефіцієнтів

Але з природи артефакту ми знаємо що він формується при роботі очних м'язів, а отже має чітку локалізацію у передній частині голови нижче лобних електродів. Крім того активність цих м'язів, як і активність більшої частини мозку у мить генерації спайків, параксизмів або спалахів ритмів, можна описати «дипольною моделлю» [21].

Сутність «дипольної моделі» полягає у описанні просторового закону розподілення сигналу джерела активності мозку як електричного диполю.

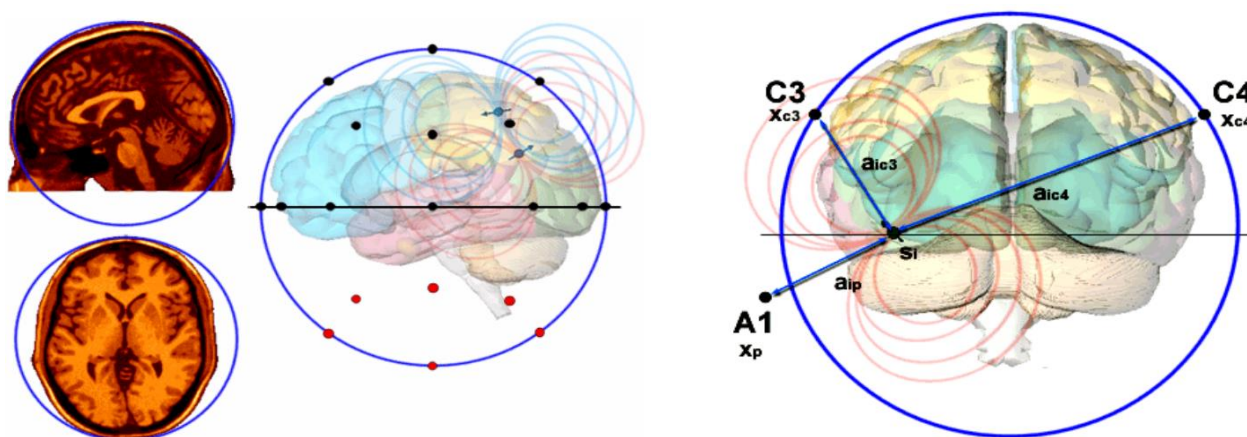


Рисунок 16 – Ілюстрація дипольної сферичної моделі

Якщо поверхню скальпу на яку накладаються електроди спрощено моделювати сферою [22, 23], тоді розподілення вагових коефіцієнтів амплітуди диполю який моделює джерело сигналу спрощено можна описати формулою:

$$A(r) = \frac{1}{4\pi\epsilon} * \left(\frac{p \cdot r}{|r|^3} \right)$$

де $A(r)$ – значення вагового коефіцієнту за напрямком r ,

r – вектор напрямку від центра диполя до точки на поверхні скальпу де розраховується коефіцієнт (точка локалізації електроду),

p – дипольний момент (вектор),

ϵ – діелектрична константа тканин мозку, змінюється від 0,1...50 в залежності від частоти сигналу джерела, для сигналів ЕЕГ вона приблизно дорівнює 0,3 у спрощених моделях.

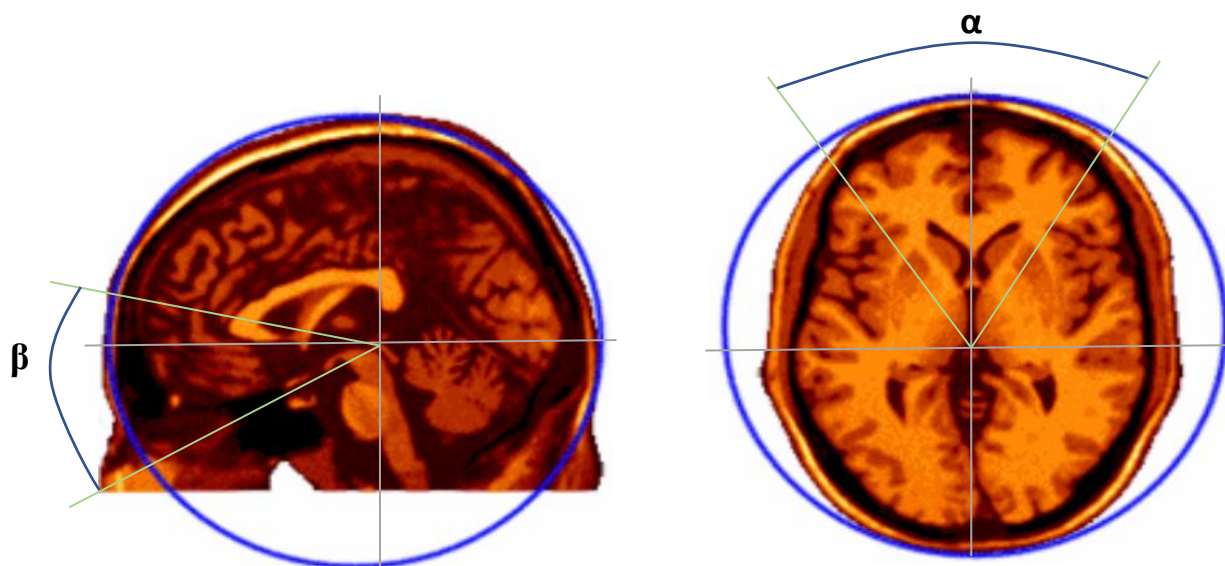


Рисунок 17 – Полярні координати пошуку диполю моделюючого окулографічний артефакт

Задача пошуку моделі окулографічного артефакту зводиться до пошуку для кожного джерела s_i та його вагових коефіцієнтів \bar{a}_i електричного диполю розташованого у полярних координатах по куту азимуту α від -30° до $+30^\circ$, та з куту піднесення β від -30° до $+10^\circ$ та відносною довжиною радіус вектору дипольного моменту від 0,9 до 0,5, де 1 – це поверхня сфери дипольної моделі, 0 – центр сфера або цент полярної системи координат. Пошук дипольної моделі – це оптимізаційна задача пошуку максимуму функції з трьома параметрами яка вирішуються градієнтними методами.

Якщо таку модель можна знайти з та коефіцієнт кореляції Пірсона між модельними коефіцієнтами та ваговими коефіцієнтами джерела $r_{A(r),\bar{a}_i} \geq 0.85$ – це джерело з високою достовірністю описує окулографічний артефакт.



Рисунок 18 – Приклад карти артефакту (а) та дипольної моделі (б) що його описує $\alpha = 4^\circ$, $\beta = 10^\circ$

2.6 Виділення ділянок підозрілих на артефактну активність

Залишилося визначити питання як знайти ділянки запису на яких будемо проводити ІСА декомпозицію для пошуку артефактів. Слід зауважити що для гарного пошуку окулографічних артефактів достатньо проводити ІСА декомпозиції відразу по всій довжині ЕЕГ запису. Окулографічні артефакти мають високу енергетичну складову, постійно локалізуються в одному й тому ж місті, а отже завжди будуть виділятися ІСА алгоритмом в окремий компонент – джерело. Проте електродні артефакти так знайти буде важко, а отже треба виділяти фрагменти запису в яких проводити декомпозицію та пошук артефактів.

Загальною характеристикою усіх трьох артефактів – різка зміна закону розподілення амплітуд сигналу в час прояву артефакту: зростає амплітудна варіативність. Зазвичай це характеризує й патологічну активність також, а не лише артефакти. Але оскільки ми шукаємо лише ділянки де артефакти можливі – цього буде достатньо. Ця характеристика добре описується статистичним показником – дисперсія розподілення амплітуди сигналу σ . Для порівняння сплеску дисперсії треба мати опорну величину – дисперсію корисного ЕЕГ сигналу яку також можна знайти якщо розділити сигнал за частотою у 1,5-2Гц: весь сигнал у частотному діапазоні вище цієї частоти зрізу слугує для розрахунку опорної дисперсії сигналу ЕЕГ σ .

Алгоритм пошуку фрагментів ризької зміни дисперсії (можливі артефактні ділянки) наступний:

Крок 1. Вихідний сигнал профільтрувати смуговим фільтром 0,5 ... 8 Гц (усунення зайвих коливань).

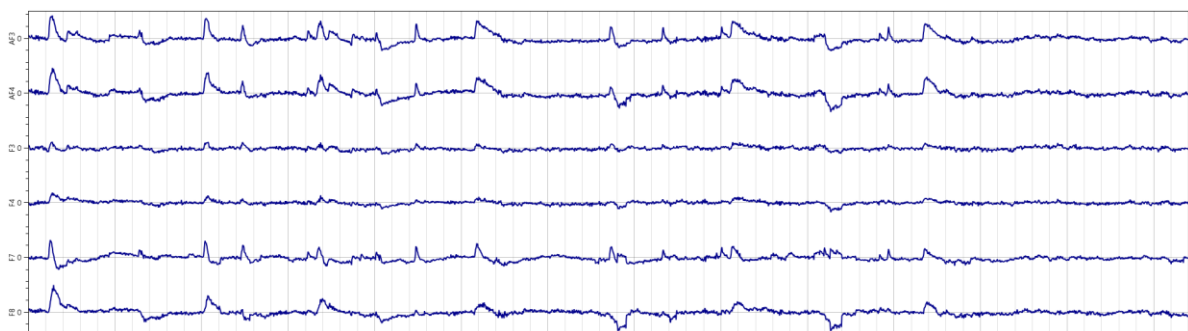
Крок 2. Для кожної складової ЕЕГ запису (сигналу з кожного електроду):

Крок 3. Профільтрувати фільтром високих частот з частотою зрізу 1,5 Гц та розрахувати опорну дисперсію σ .

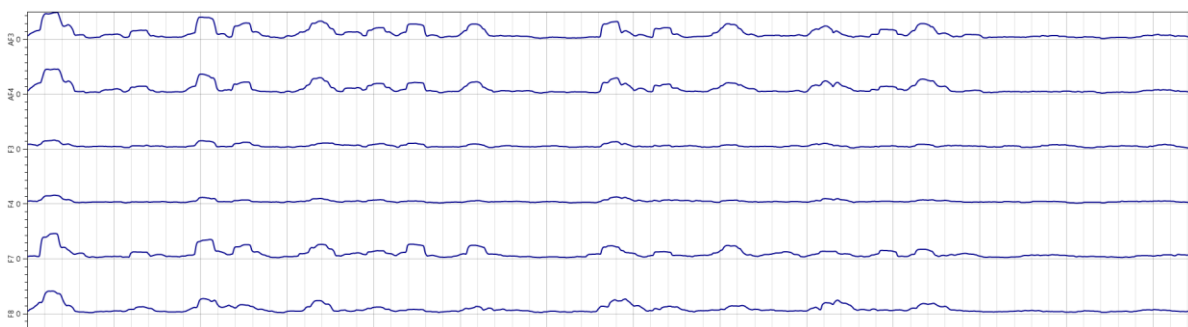
Крок 4. Методом ковзного вікна шириною 2 секунди розрахувати функцію зміни дисперсії σ сигналу.

Крок 5. Знайти усі ділянки де функція зміни дисперсії σ більша опорної дисперсії у три рази (пороговий коефіцієнт) та додати по 1 секунді зліва та справа до кожної ділянки. Повернутись на крок 2 якщо є необроблені складові.

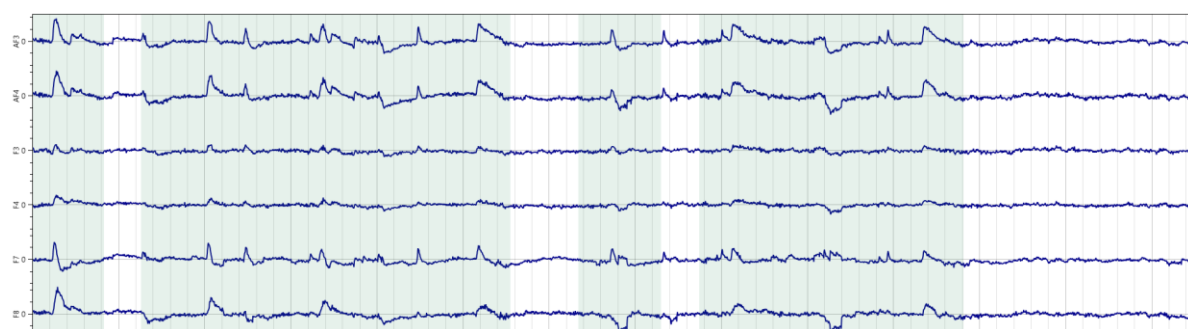
Крок 6. Ділянки усіх сигналів ЕЕГ що перекриваються у часовому просторі об'єднати у одну ділянку.



а



б



в

Рисунок 19 – Приклад роботи алгоритму: а – вихідний запис, б – функція зміни дисперсії σ з вікном у 2 секунди, в – результуючі ділянки

2.7 Формулювання алгоритму вирішення задачі

Отже, ІСА – це вдалий метод попередньої обробки датасету якій дозволяє провести якісне очищення датасету від артефактів запису.

Сумуючи вище сказане отримаємо алгоритм рішення завдання з попередня обробка EEG датасету.

Крок 1. розділення EEG датасету на фрагменти зручні для незалежної обробки (один фрагмент – EEG запис).

Крок 2. Фільтрація EEG запису смуговим відокремлюючим амплітудно-частотним фільтром 0.1 – 50 Гц (BandPass filter). Верхня границя фільтру може зміщуватись в залежності від виду запису та потреб. Якщо вона більша 50 Гц, потрібна додаткова фільтрація від завади електромережи в 50/60 Гц (BandStop filter).

Крок 3. Виявлення ділянок запису з імовірними артефактами.

Крок 4. Проведення ІСА декомпозиції поточної ділянки.

Крок 5. Аналіз кожної компоненти на її «артефактність».

Крок 6. Придушення сигналу компоненти в залежності від типу знайденого артефакту.

Крок 7. Зворотна ІСА композиція для отримання очищеної ділянки, повернення на крок 3 якщо є необроблені ділянки фрагменту.

Крок 4. Обробка наступної ділянки.

Крок 5. Обробка наступного фрагменту.

Крок 6. Завершення.

3. МОДЕЛЮВАННЯ ТА ПРОГРАМНА РЕАЛІЗАЦІЯ СИСТЕМИ

На основі попереднього аналізу та запропонованих алгоритмів можна сформулювати вимоги до програмної системи здатної проводити обчислення з попередньої обробки та очищення ЕЕГ сигналу.

3.1 Принципи роботи програмного продукту

Програмна система повинна складатися з таких частин як сервісна частина (ядро програмної системи), модуль тестування та UI додаток для запуску на персональному комп'ютері.

Сервісна частина – це головний модуль який реалізує всю бізнес логіку програмної системи але не містить UI логіку. Завдання цього модулю проведення усіх обчислень та експортування публічного API для використання його в будь якій сумісній системі. Бізнес логіка сервісної частини повинна реалізовувати автоматичний сценарій обробки ЕЕГ сигналу.

Модуль тестування – це модуль системного тестування сервісної частини. Головне завдання модулю перевіряти адекватність роботи алгоритмів які експортуються сервісною частиною за допомогою публічного API.

UI додаток для запуску на персональному комп'ютері – це програма яка є UI оболонкою над сервісною частиною. Вона реалізує кінцевий інтерфейс програмної системи та дозволяє користувачу проводити усі етапи попередньої обробки ЕЕГ сигналів використовуючи сервісну частину у автоматичному та ручному режимах. При необхідності сервісна частина може бути віддалена від програмного застосунку, а інтерфейс їх взаємодії організований за принципами RPC (виклик віддалених методів).

3.2 Вимоги до функціональних характеристик сервісного модулю системи

Сервісний модуль повинний містити наступні публічні функції (інтерфейси API):

- завантаження та збереження даних EEG запису з відкритих стандартів датасетів. Пропонується стандарт ARFF (Attribute-Relation File Format) [19]. Реалізація усіх найпоширеніших стандартів EEG записів не є темою цієї роботи,
- виконання амплітудно-частотної фільтрації EEG сигналу чи сигналу ICA компонентів,
- реалізація ICA декомпозиції та ICA композиції. У якості ICA алгоритму треба застосовувати FastICA метод,
- реалізація алгоритмів пошуку найбільш шкідливих видів артефактів: окулографічний, електродний, артефакт референтного електроду з наданням кількісної оцінки якості проведеного аналізу,
- реалізація додаткових алгоритмів допоміжних у знаходженні артефактів: знаходження ділянок можливої активності артефактів; дипольна модель EEG зі сферичним моделюванням скальпу; пошук найкращого диполю з обмеженнями розташування.

Сервісний модуль повинен бути реалізований за допомогою кросплатформених засобів програмування для легкого його переносу на сервісну платформу реалізовану на найпоширеніших середовищах виконання: Windows, Linux, MacOS-X.

3.3 Вимоги до інтерфейсу програмного додатку

Оскільки UI додаток для запуску системи на персональному комп'ютері не є головною частиною системи, а лише засіб наглядної візуалізації роботи – до нього не застосовуються окремі вимоги інтерфейсу. Цілком достатня реалізація UI у

інтерфейсному стилі (дизайн) операційної системи на якій він буде виконуватись з виконанням ергономічних характеристик.

Для створення інтерфейсу користувача (UI) необхідно враховувати користувацький досвід та дотримуватися загальних вимог до проектування інтерфейсів.

До загальних вимог до проектування відносяться наступні:

- всі елементи інтерфейсу повинні бути структуровані,
- всі елементи інтерфейсу повинні бути логічно згруповані,
- всі елементи інтерфейсу повинні бути вирівняні,
- всі елементи інтерфейсу та сторінки повинні бути виконані в єдиному стилі.

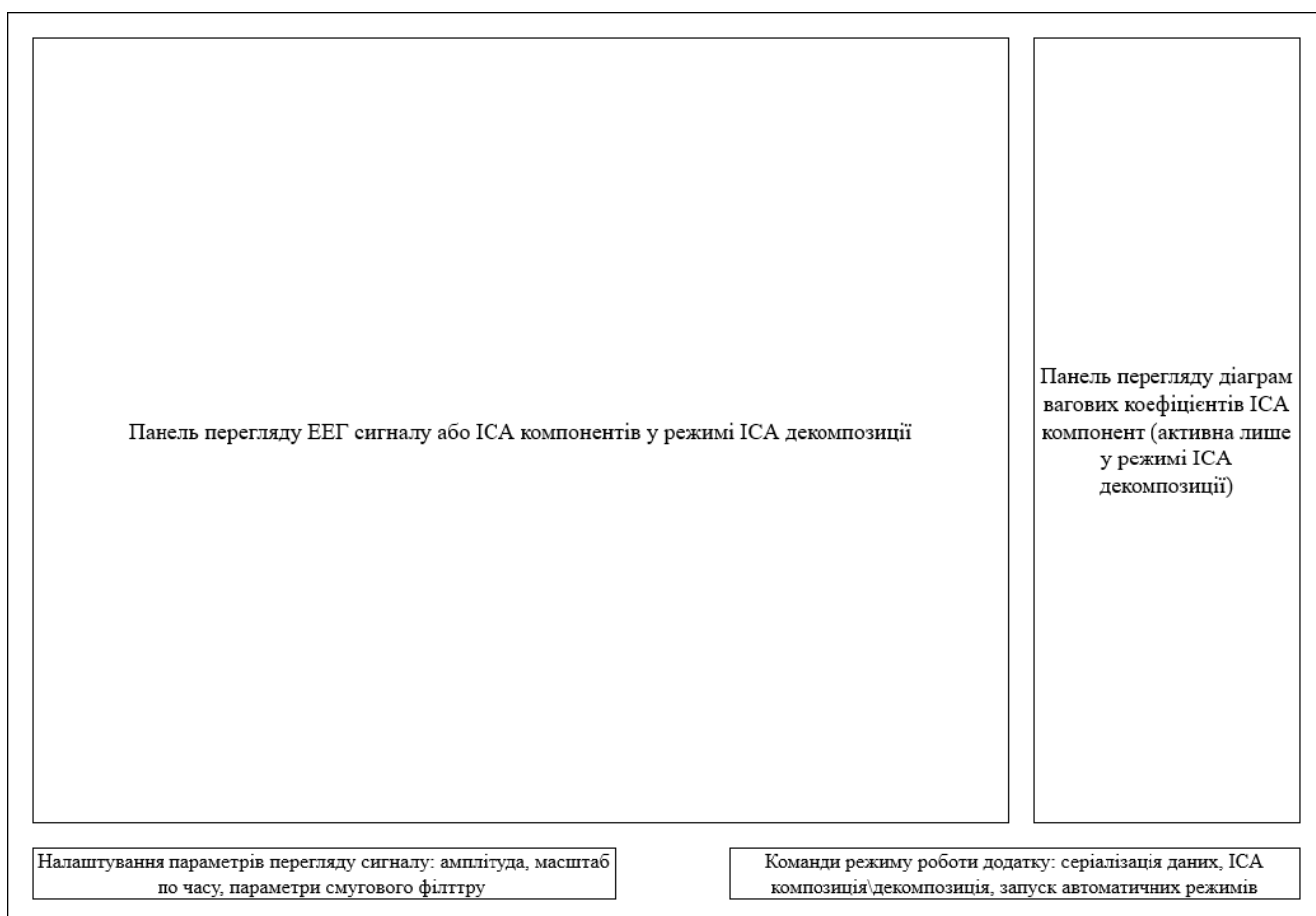


Рисунок 20 – Макет розташування основних елементів інтерфейсу користувача у програмному додатку

3.4 Вимоги до алгоритму роботи й структури програми

Алгоритм функціонування сервісного додатку повинен реалізувати етапи описані у пунктах «1.6 Постановка задачі» та «2.7 Формулювання алгоритму вирішення задачі». Вихідний код системи повинен бути реалізований за принципами об'єктно-орієнтованого програмування та розділений на функціональні блоки за принципами декомпозиції на класи які задовольняють SOLID критеріям. Вихідний код всіх модулів системи повинен бути логічно згрупований. API зовнішніх класів сервісного додатку повинні бути перевірені за допомогою модуля тестування.

3.5 Вимоги до інформаційної і програмної сумісності

Оскільки потрібно реалізувати кросплатформений підхід для сервісної частини та модулю тестування, у якості вимог програмної сумісності логічно обрати .NET екосистему:

- мова програмування: C#, фреймворк .Net Core,
- кодування: UTF-8 для вихідного коду та файлів зберігання ЕЕГ сигналів,
- IDE: Visual Studio.

Вимоги до UI додатку є логічним продовженням вимог до сервісної частини:

- мова програмування: C#, фреймворк .Net Core,
- кодування: UTF-8 для вихідного коду та файлів зберігання ЕЕГ сигналів,
- UI фреймворк: WinForm оскільки він є прямим адаптером між Windows Desktop API та фреймворком .Net Core,
- IDE: Visual Studio.

3.6 Компоненти системи та діаграма пакетів

Проведемо декомпозиції програмної системи на компоненти згідно з попередніх вимог:

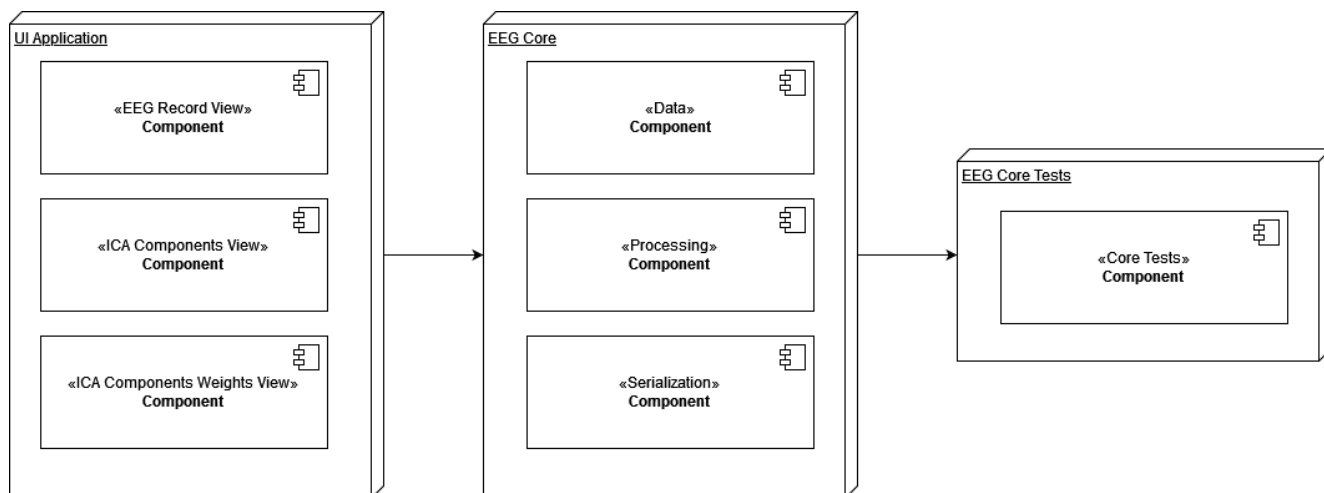


Рисунок 21 – Діаграма компонентів

Сервісний модулі «EEG Core» повинен містити три головних компоненти які реалізують три види сутності бізнес логіки: модель даних, контролер та модель серіалізації даних. Модель перегляду відсутня оскільки цей модуль зосереджений лише на методах попередньої обробки та очищені EEG даних.

Компонент моделі даних «Data» реалізує логіку представлення класів даних у пам'яті комп'ютера, це такі дані як EEG сигнал, незалежні компоненти тощо.

Компонент контролеру «Processing» відповідає за бізнес логіку обробки даних: фільтрація, ICA композиція\декомпозиція, аналіз артефактів тощо.

Компонент моделі серіалізації «Serialization» відповідає за збереження-завантаження даних з постійної пам'яті комп'ютера: ARFF серіалізація.

Модуль тестування «EEG Core Tests» містить лише один компонент – «Core Tests», це набір тестів для компонентів модуля «EEG Core» та вихідні данні для цих тестів.

UI модуль «UI Application» містить компоненти перегляду EEG сигналів у різних станах: EEG запис, перегляд ICA компонентів, перегляд просторової локалізації ICA компонентів (вагові коефіцієнти компонентів).

Діаграма пакетів значною мірою дублює діаграму компонентів, оскільки містить ту ж суму логіки декомпозиції:

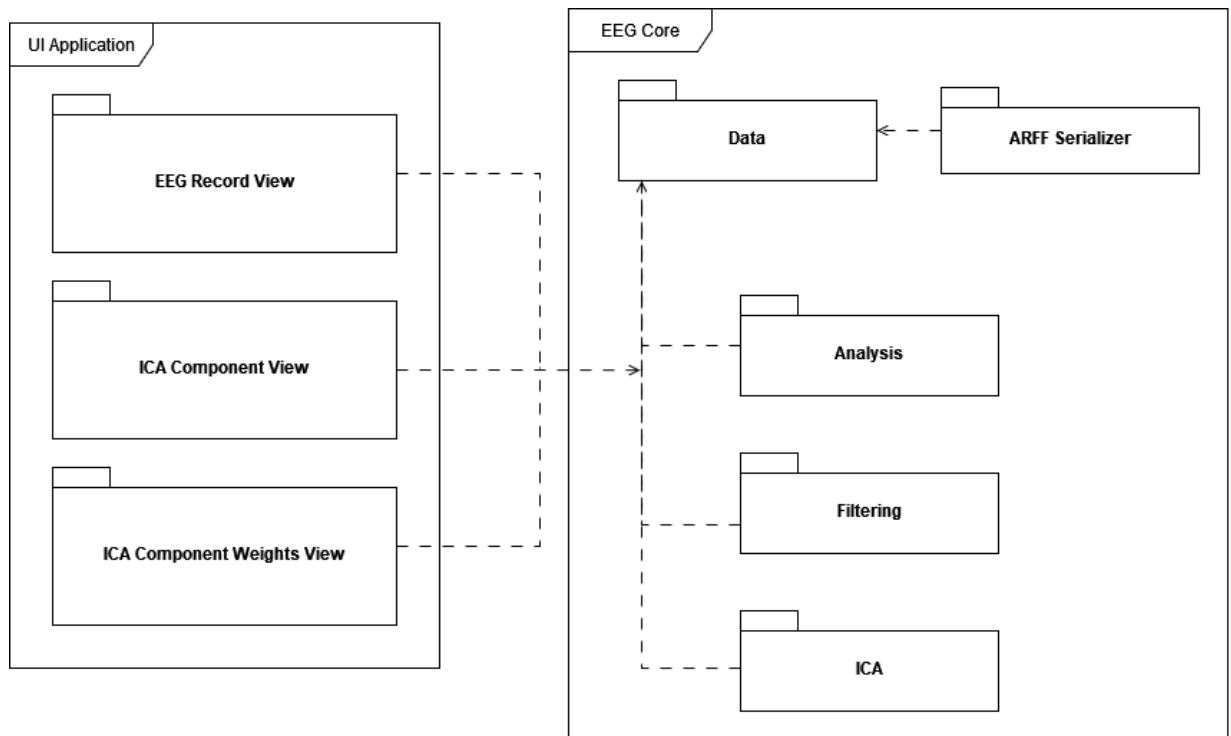


Рисунок 22 – Діаграма пакетів. Пакет модуля тестування незображений оскільки не є суттєвим на даній діаграмі

3.7 Декомпозиція системи на класи та діаграма послідовності алгоритму

Проведемо декомпозицію системи на класи згідно вимог. Оскільки головна частина програмної системи – це сервісний модуль «EEG Core», будемо розглядати саме його.

Компонент моделі даних «Data» складається з наступних класів:

- Lead – базовий клас одного каналу сигналу,
- Record – базовий клас логіки зберігання сету багатоканальних сигналів, зв'язаний композицією з набором класів Lead (один до багатьох),
- EEGLead – реалізація базового класу Lead для EEG сигналу з одного електроду,
- ComponentLead – реалізація базового класу Lead для незалежної компонент,

- EEGRecord – реалізація базового класу Record для EEG запису,
- ICARRecord – реалізація базового класу Record для результату ICA декомпозиції.

Компонент моделі Serialization складається з класів:

- RecordFactory – фабрика об'єктів класу Record яка створює необхідну реалізацію класу, необхідну реалізацію класу серіалізації та пов'язана з ним агрегацією через інтерфейси серіалізації,
- IRecordSerializer та IRecordDeserializer – інтерфейси алгоритму серіалізації даних. Алгоритм розділений на два інтерфейси для проведення збереження та завантаження незалежно один від іншого,
- ARFFSerializer – клас серіалізації у стандарт ARFF, реалізує IRecordSerializer та IRecordDeserializer.

Компонент контролеру Processing поділений на пакети аналізу, фільтрації та ICA декомпозиції:

- Фільтрація (Filtering):
 - FilterFactory – фабрика об'єктів класу фільтрації FIRFilter, відповідає за створення фільтру та налаштування параметрів фільтрації: фільтр мережевих завад, BandPass фільтр чи його варіанти HiPass або LoPass фільтри. Повертає інтерфейс налаштованого фільтру IFilter,
 - IFilter – інтерфейс фільтрації,
 - FIRFilter – реалізує інтерфейс фільтрації IFilter та є імплементацією KIX (FIR) фільтру.
- ICA:
 - ICA – клас-міст реалізації ICA композиції та декомпозиції, асоційований з класами Record та ICARRecord, для виконання ICA алгоритму клас звертається до «драйверного» абстрактного класу ICAEngine,
 - ICAEngine – абстрактний клас реалізації ICA алгоритму,
 - FastICA – клас реалізації ICA алгоритму на основі FastICA методу максимізації міри негаусовості джерел.

- Аналіз (Analysis):
 - AnalyzerBase – базовий абстрактний клас аналізу джерела ІСА декомпозиції на наявність артефактних ознак. Клас асоційований з класом ArtefactInfo який описує вид артефакту в даному джерелі та з'єднаний з класом джерела ComponentLead яка агрегація;
 - ElectrodeArtifactAnalyzer – клас-реалізація для AnalyzerBase, проводить пошук електродного або референтного артефактів,
 - EyeArtifactAnalyzer - клас-реалізація для AnalyzerBase, проводить пошук окулографічного артефакту
 - ArtifactCandidateRangeDetector - клас-реалізація для AnalyzerBase, проводить пошук ділянок ЕЕГ запису за алгоритмом аналізу зміни дисперсії сигналу.

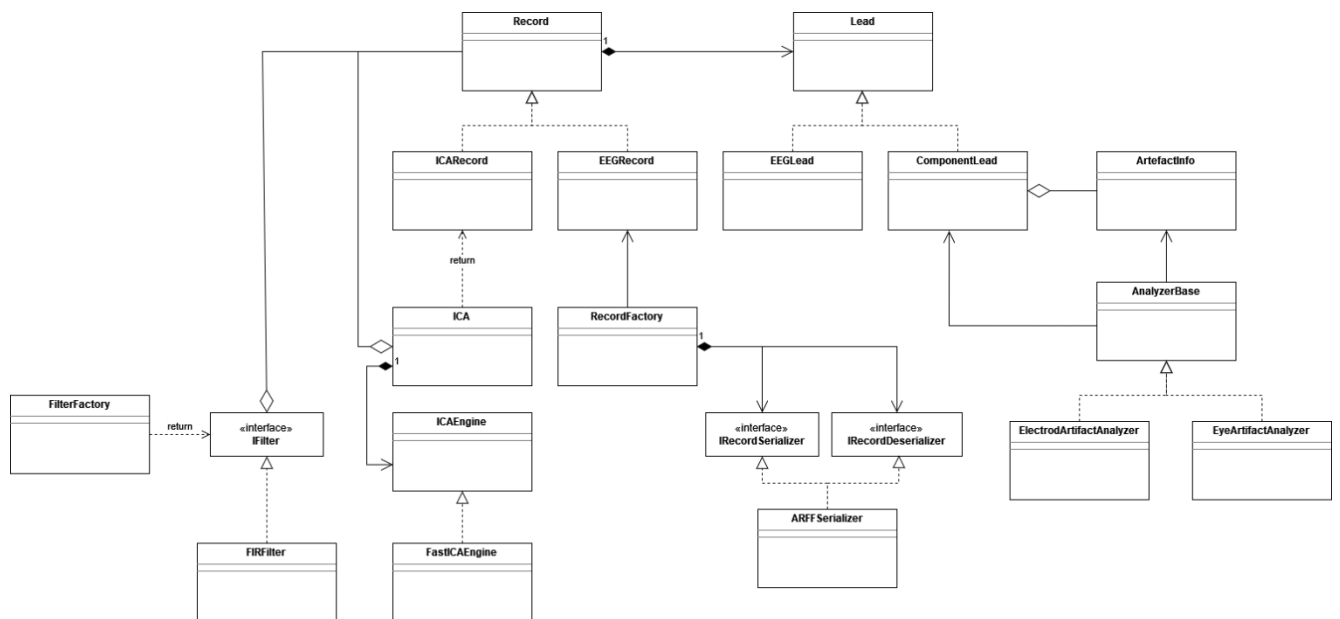


Рисунок 23 – Діаграма основних класів сервісного модулю

На основі декомпозиції на класи та алгоритму попередньої обробки ЕЕГ сигналів описаному у другому розділі сформуємо діаграму послідовності дій сервісного компоненту:

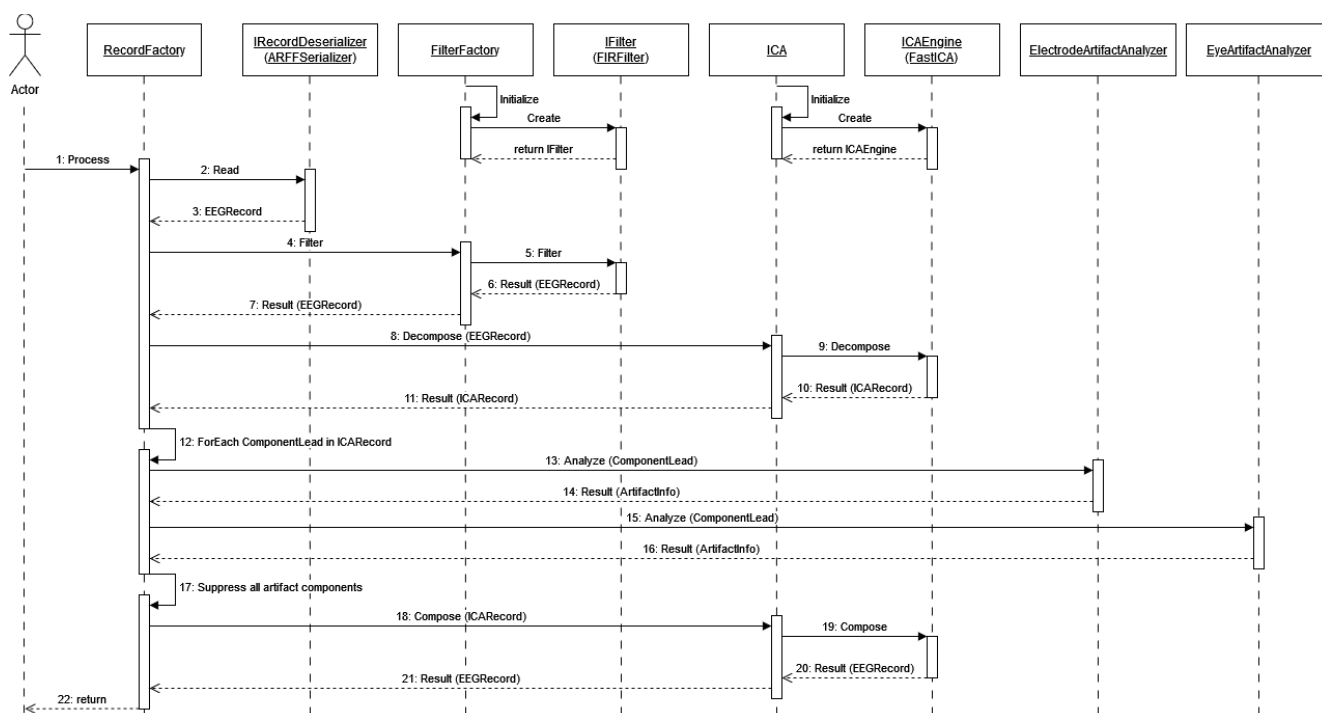


Рисунок 24 – Sequence diagram для основної задачі сервісного модулю (автоматичний режим)

Ця діаграма є зображенням послідовності дій як результат запиту на повну обробку даних одного фрагменту ЕЕГ сигналу. Оскільки всі описані класи є елементами публічного API сервісного модулю, то ця обробка може бути виконана послідовно крок за кроком за допомогою UI додатку який є складовою частиною програмної системи.

4. ПЛАНУВАННЯ ТА ПРОВЕДЕННЯ ЕКСПЕРИМЕНТУ

4.1 План проведення експерименту

Проведемо експеримент з оцінки якості попередньої обробки ЕЕГ сигналу та очищення його від артефактів. Для цього у якості вихідних даних візьмемо відкриті ЕЕГ датасети з [25] та [26], вони містять ЕЕГ датасети як з артефактами так і без.

Мета проведення експерименту – оцінити якість роботи автоматичного режиму програмної системи. Для досягнення цієї мети проведемо експеримент у два етапи:

- оцінка роботи програмної системи по синтетичним даним у тестах,
- оцінка роботи програмної системи на реальному ЕЕГ датасету з артефактами.

За результатами експерименту по кожному з етапів зробимо аналіз.

4.1 Результати експерименту

Оцінка результатів роботи системи по кожному з етапів потребує запуску системи різними способами.

Так оцінка системи по синтетичним даним з наявністю опорних даних у які штучно вносяться артефакти проводиться у повністю автоматичному режимі та є складовою тестів модуля тестування «EEG Core Tests».

Оцінка системи по реальному ЕЕГ датасету потребує запуску системи через програмний додаток з інтерфейсом користувача, запуску автоматичного очищення ЕЕГ сигналу через кнопку «Auto» та візуальну оцінку результату роботи.

4.1.1 Оцінка роботи системи по синтетичним даним

У якості опорного (еталонного) ЕЕГ сигналу будемо застосовувати чистий ЕЕГ запис з [25].

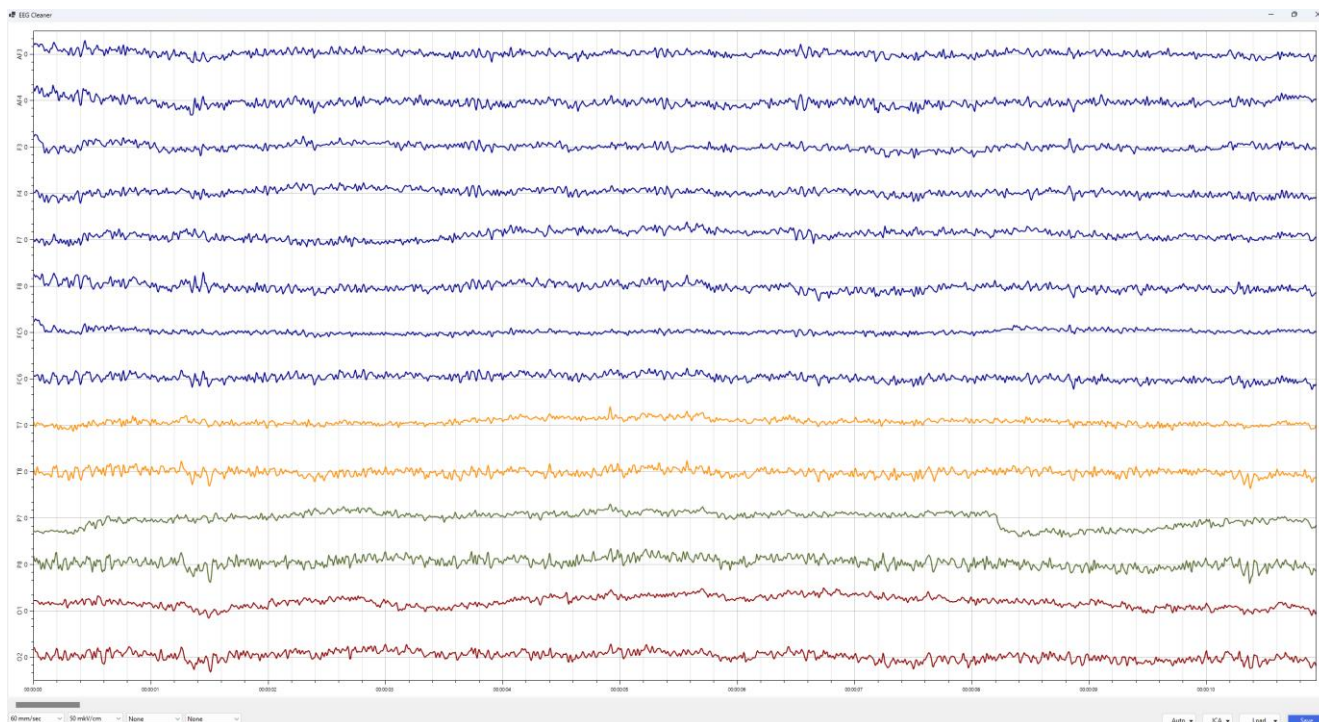


Рисунок 25 – Еталонна ЕЕГ без артефактів

Почнемо з перевірки роботи алгоритмів знаходження та видалення окулографічних артефактів. Для цього запустимо автоматичні тести пошуку артефактів «Test1_EyeArtifactAutoRemove» та «Test2_EyeArtifactAutoRemove», це тести з набору тестів у модулі тестування «EEG Core Tests». Вони виконують дві перевірки: тестування з синтетичними моделями окулографічних ЕЕГ артефактів низької частоти - 1Гц, та високої (для окулографічного артефакту) частоти – 2Гц.

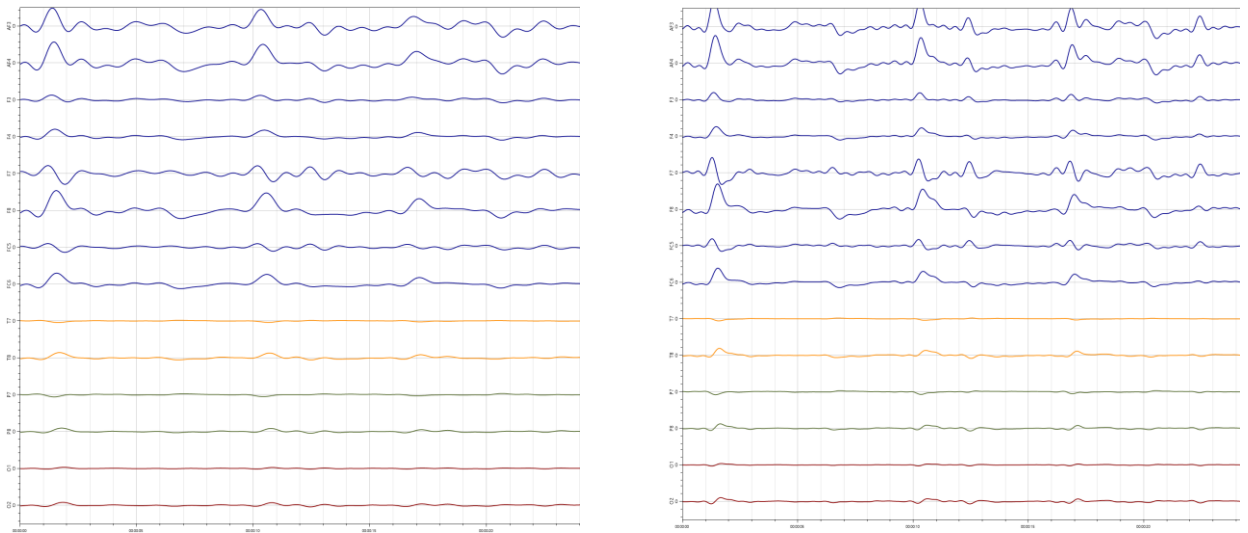


Рисунок 26 – Моделювання окулографічного артефакту частотними коливаннями 1Гц та 2Гц

Змодельований сигнал окулографічного артефакту змішується у тесті з сигналом еталонної ЕЕГ та запускається автоматичний режим пошуку та видалення артефактів. У автоматичному режимі програма працює згідно з алгоритмом «2.7 Формулювання алгоритму вирішення задачі». Результат роботи алгоритму, очищена ЕЕГ, порівнюється з еталонною та розраховується коефіцієнт кореляції Пірсона для порівняння їх подібності: абсолютні значення сигналу та його медіана могли змінитися внаслідок обробки ІСА алгоритмом, тому нам важливо збереження форми та частотної характеристики сигналу.

У результаті тестування отримали наступні величини:

Таблиця 2 – Оцінка якості очищення ЕЕГ сигналу від змодельованого артефакту

Тест	$r_{x_{src}, x_{est}}$	$\Delta r_{x_{src}, x_{est}}$
1Гц модель	92,9%	[75% ... 98%]
2Гц модель	88,2%	[68% ... 95%]

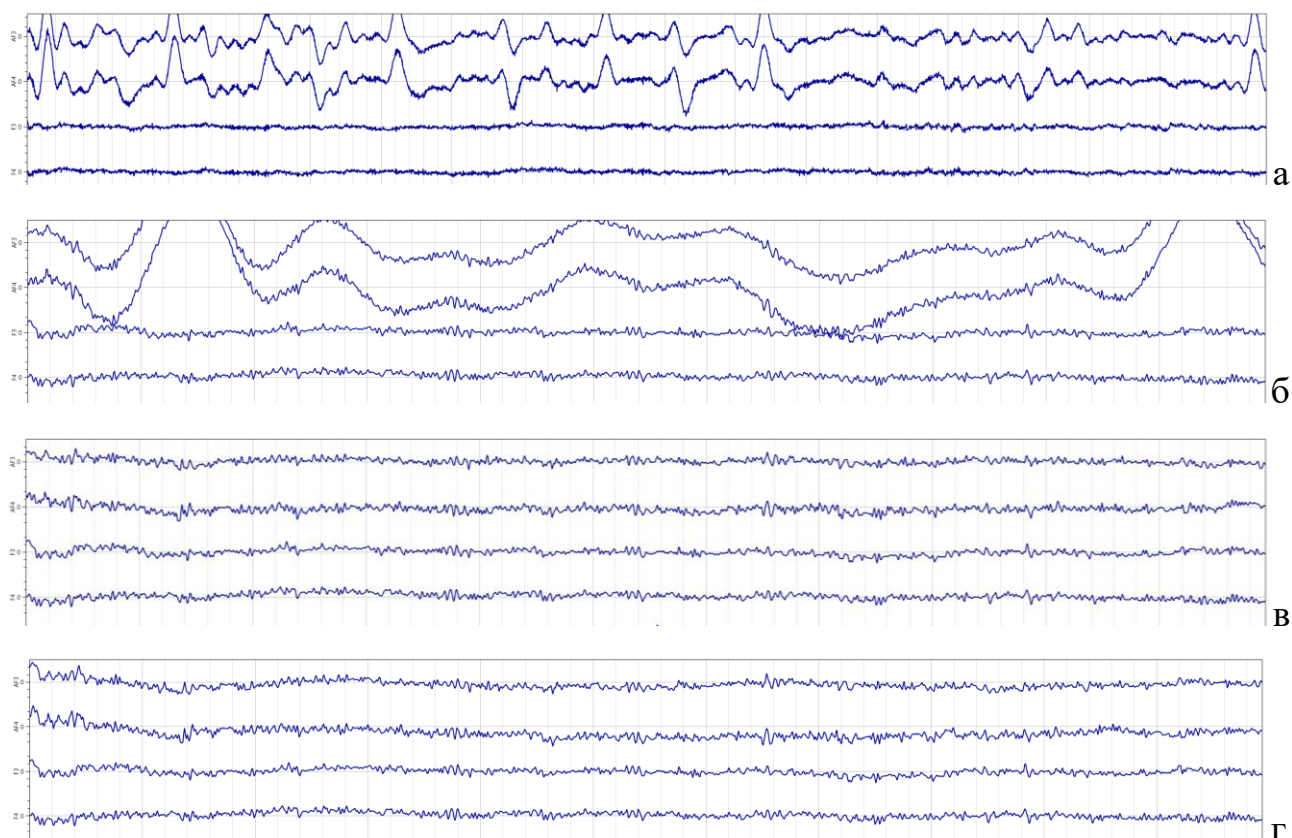


Рисунок 27 – Візуальне порівняння ЕЕГ з підмішаним артефактом (а, б) у різних масштабах, еталонної ЕЕГ (в) та її очищеного варіанту (г)

Аналогічно проведемо експеримент з виявлення та видалення артефакту референтного електроду. У якості еталонного сигналу застосовуємо той самий ЕЕГ датасет. Запустимо автоматичний тест пошуку артефактів референтного електроду «Test3_ReferenceArtifactAutoRemove». Цей тест моделює активність артефакту референтного електроду синусоїдальним сигналом амплітудою 40мкВ та частотою 10Гц. Змодельований артефакт підмішується до еталонного сигналу та запускається алгоритм пошуку та видалення артефактів. Результат оцінюється аналогічно до попереднього тесту

Таблиця 3 – Оцінка якості очищення ЕЕГ сигналу від змодельованого артефакту

Тест	$r_{x_{src}, x_{est}}$	$\Delta r_{x_{src}, x_{est}}$
10Гц модель артефакту	75,9%	[71% ... 85%]

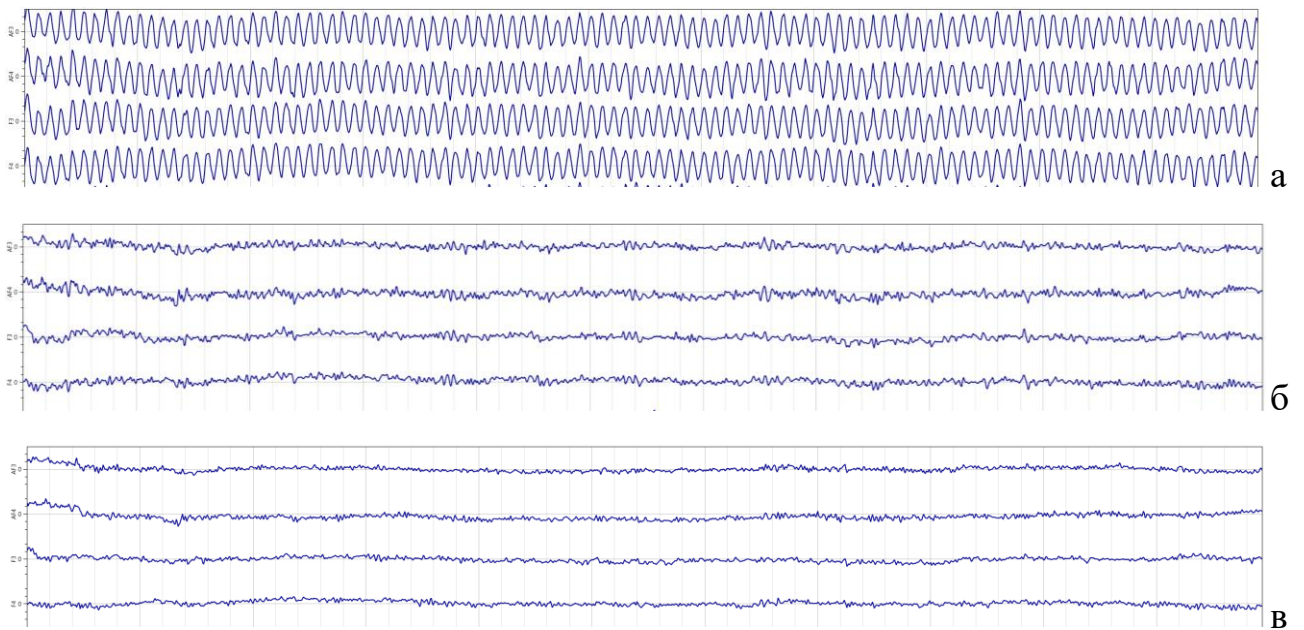


Рисунок 28 – Візуальне порівняння ЕЕГ з підмішаним артефактом референтного електроду (а), еталонної ЕЕГ (в) та її очищеного варіанту (г)

Аналогічно виконуємо останній тест «Test4_SingleArtifactAutoRemove» – перевірка якості знаходження та видалення електродного артефакту у одному електроді. Цей тест моделює електродний артефакт синусоїдальним сигналом амплітудою 40мкВ та частотою 13Гц, частота обрана відмінною від тесту «Test3_ReferenceArtifactAutoRemove» з метою перевірити роботу алгоритму на різних частотних діапазонах сигналу артефактів. Змодельований артефакт домішується лише до сигналу від одного електроду (другий канал в ЕЕГ датасеті, AF4, фронтальна частина скальпу). Інші аспекти роботи тесту - аналогічні попереднім тестам.

Таблиця 4 – Оцінка якості очищення ЕЕГ сигналу від змодельованого артефакту

Тест	$r_{x_{src}, x_{est}}$	$\Delta r_{x_{src}, x_{est}}$
13Гц модель артефакту	96,84%	[93% ... 98%]

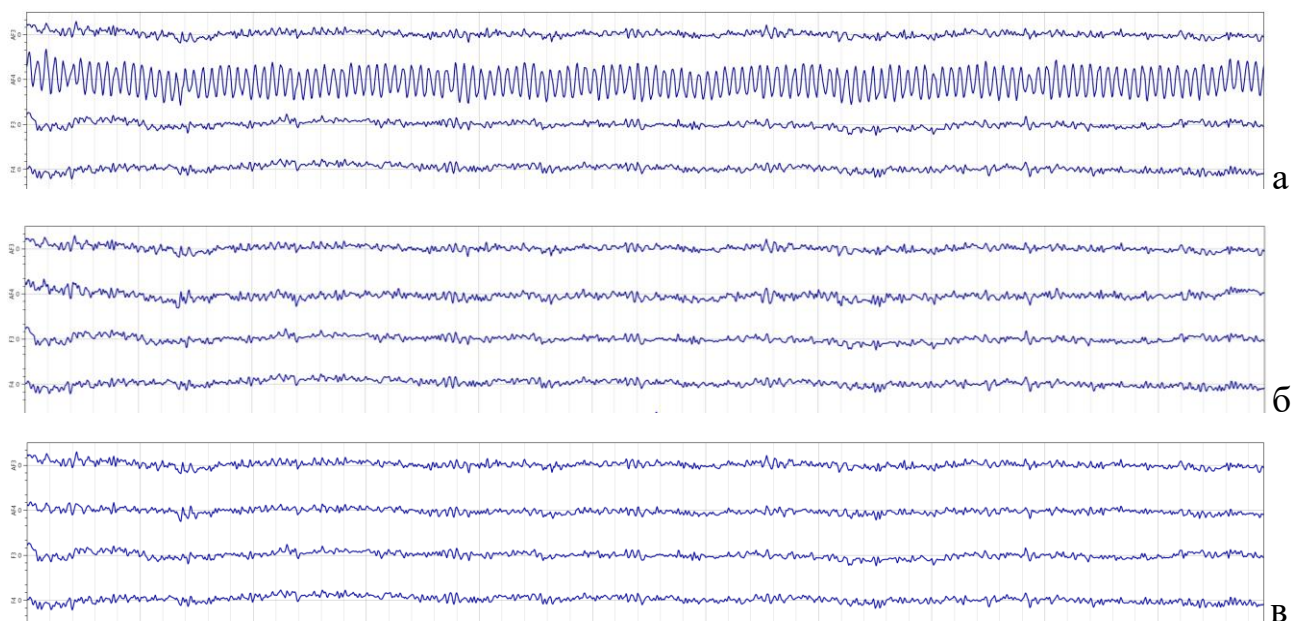


Рисунок 29 – Візуальне порівняння EEG з підмішаним артефактом референтного електроду (а) для електроду AF4, еталонної EEG (в) та її очищеного варіанту (г)

За синтетичними тестами бачимо що алгоритм працює якісно, очищений варіант EEG сигналу збігається з еталонним на 75 – 95% в залежності від тесту.

4.1.2 Оцінка роботи системи на реальному EEG датасету

У якості вихідних даних візьмемо один з EEG датасетів з наборів [25, 26] де присутня розмітка артефактних ділянок від лікаря-функціоналіста (Рисунок 30). Як видно цей EEG сигнал містить 22 окулографічних артефактів у різний час запису, та більшість з них розмічена як артефактні ділянки лікарем.

З опису EEG запису знаємо що реєстрація EEG проводилася одночасно з відеозаписом стану пацієнта і всі моменти де лікар помітив відкривання та закривання очей – були помічені як артефактні ділянки. На запису у діапазоні 1хв. 20с. – 1хв. 35с. добре помітний один артефакт референтного електроду який за формою сигналу дуже збігається з окулографічним артефактом, але проявляється в усіх каналах одночасно, що вказує на його референтну природу. Цей артефакт не був розмічений лікарем, його вірогідна причина – рух пацієнта у цей момент.

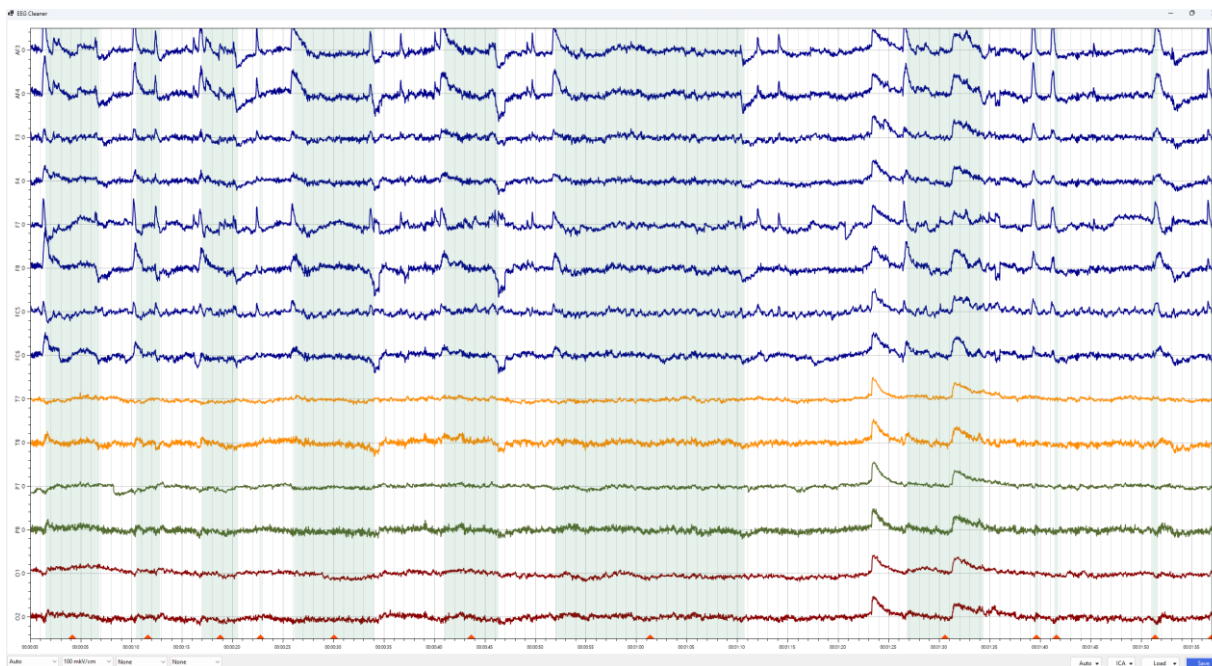


Рисунок 30 – ЕЕГ сигнал з полі-наявними окулографічними артефактами та артефактом референтного електроду у діапазоні 1хв. 20с. – 1хв. 35с.

Запускаємо автоматичне виявлення ділянок з можливою артефактною активністю: «Auto» → «Auto Ranges».

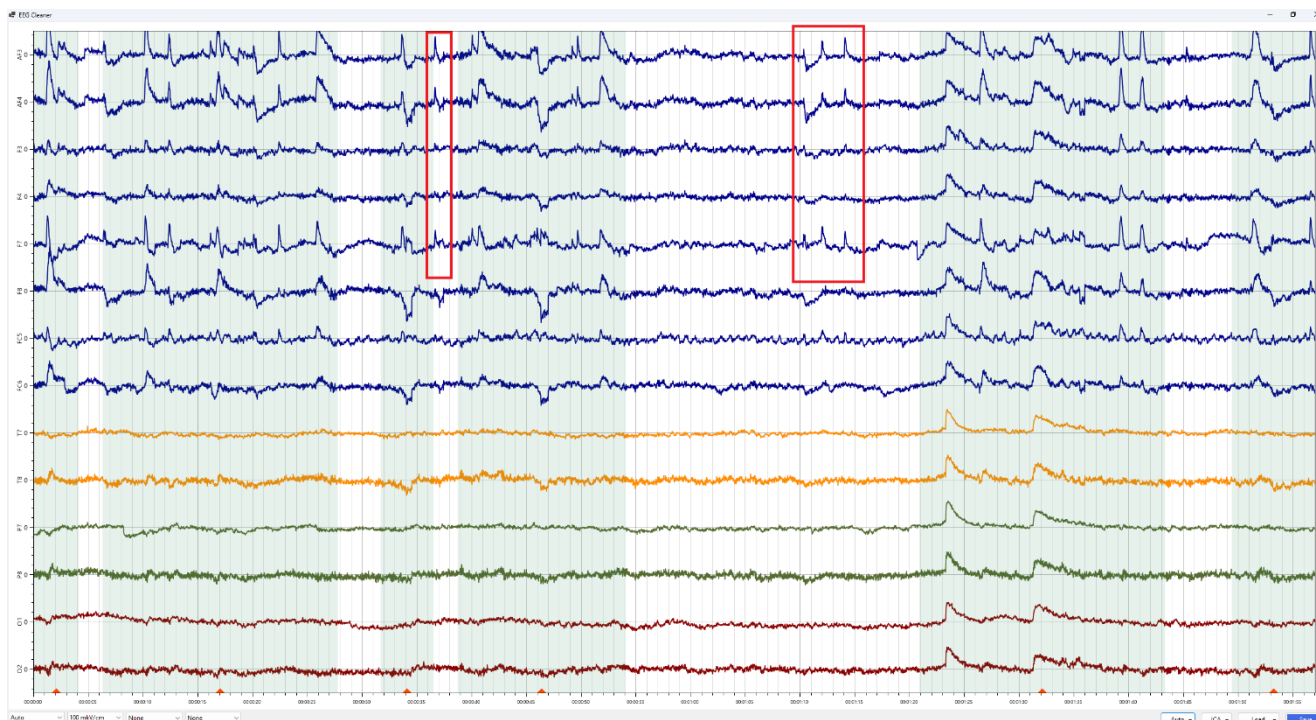


Рисунок 31 – Результат роботи алгоритму автоматичного пошуку ділянок запису для подальшого аналізу на наявність артефактів

Як бачимо розмітка змінилася, вона повністю містить у собі ділянки які позначив лікар, а також ділянку артефакту референтного електроду. Але є дві ділянки окулографічного артефакту (помічені червоним кольором) які не були знайдені – це недолік роботи порогового алгоритму. Проводимо автоматичне очищення запису: «Auto» → «Auto Clean».

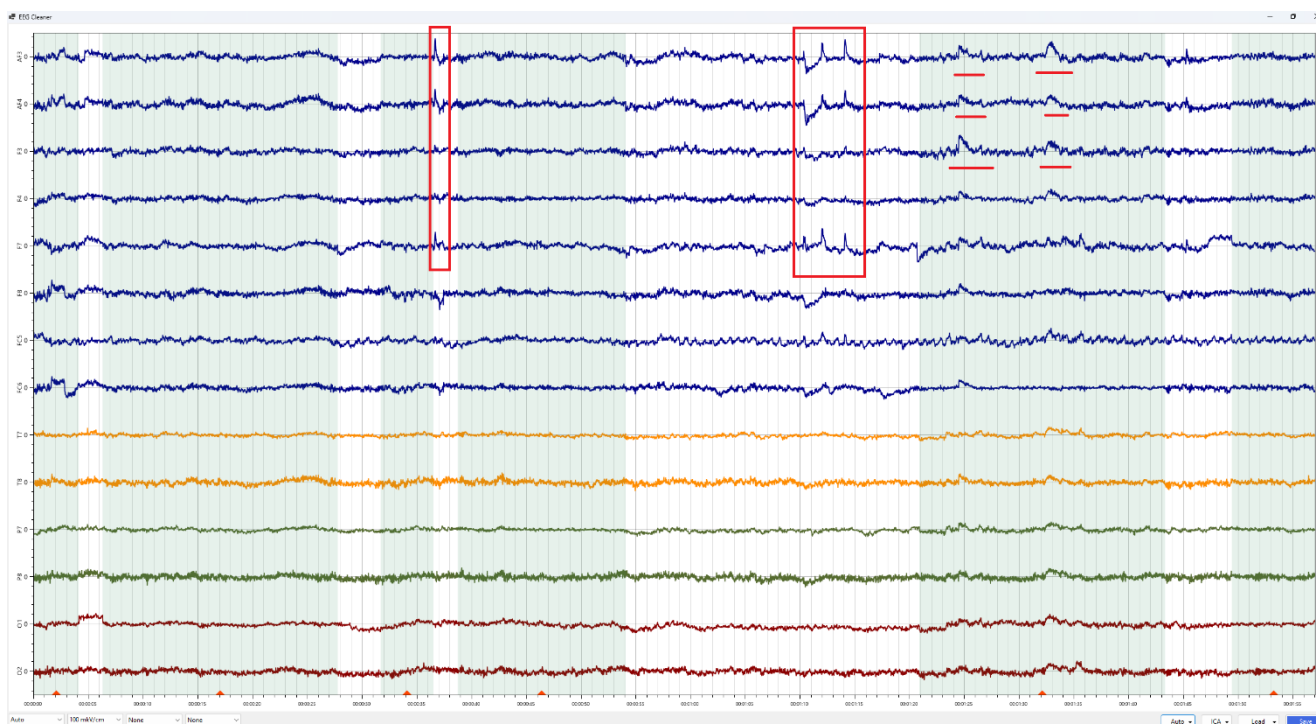


Рисунок 32 – Результат автоматичного очищення ЕЕГ сигналу від артефактів

Алгоритм пропустив 4 артефакти на ділянках запису які не потрапили до автоматичної розмітки, а отже не видалив їх. На ділянці де був присутній артефакт референтного електроду присутні його незначні залишки, що повністю збігається з результатами синтетичного тесту – тест автоматичного видалення референтного електроду показав найгірший результат. Але у цілому можна зробити висновок що система показала гарний результат у автоматичному режимі – видалено 82% артефактів, а на ділянках де артефакти ще залишилися можна додатково проаналізувати та видалити їх у ручному режимі.

4.2 Аналіз отриманих результатів

Отримані дані експерименту показують що результат не є стабільною величиною, можна лише говорити про вірогідність коректного очищення ЕЕГ сигналу. Це обумовлено принципом роботи ІСА алгоритму який є ключовим у розділенні ЕЕГ сигналу на сигнали джерел. ІСА - ймовірнісний алгоритм: він побудований на градієнтних методах оптимізації, оцінює цільову функцію ймовірнісними показниками, у якості початку пошуку застосовую функція випадкових величин. Таким чином оптимальна точка збіжності ІСА декомпозиції досягається лише з якоюсь ймовірністю. Синтетичні тести по еталону показали що еталонний сигнал очищується від артефактів з якістю у діапазоні 75-95%. Тест по реальному ЕЕГ датасету з апіорною інформацією про наявність артефактів показав якість у 82% що збігається з показником синтетичних тестів.

У якості факторів що впливають на якість роботи алгоритму знаходження артефактів слід зауважити наступне:

- вплив виду алгоритму ІСА досить суттєвий. У програмній системі реалізовано FastICA з оцінкою негаусовості компонентів через функцію $f(x) = \log \cosh(x)$ яка добре виділяє більшість видів артефактів з сигналами спайкового типу. Але існують інші варіанти ІСА (JADE, InfoMax ICA) які можуть дати гарний результат на інших видах артефактів. Тому досить цікавим напрямком дослідження було б застосувати комбінацію різних варіантів ІСА для кращої декомпозиції;
- знаходження артефактів у ІСА компонентах можна покращити застосуванням нейронних мереж;
- автоматичний пошук діапазонів з наявністю артефактів слід покращити, реалізований пороговий алгоритм може пропускати ділянки артефактів що видно у тесті по реальному ЕЕГ датасету;
- промахи при знаходженні ділянок артефактів можуть бути усунені у ручному режимі через аналіз цих ділянок у системі: після ручного вибору

ділянки та запуск ІСА режиму – система пропонує джерела до пригнічення які є підозрілими на артефакт за своєю локалізацією.

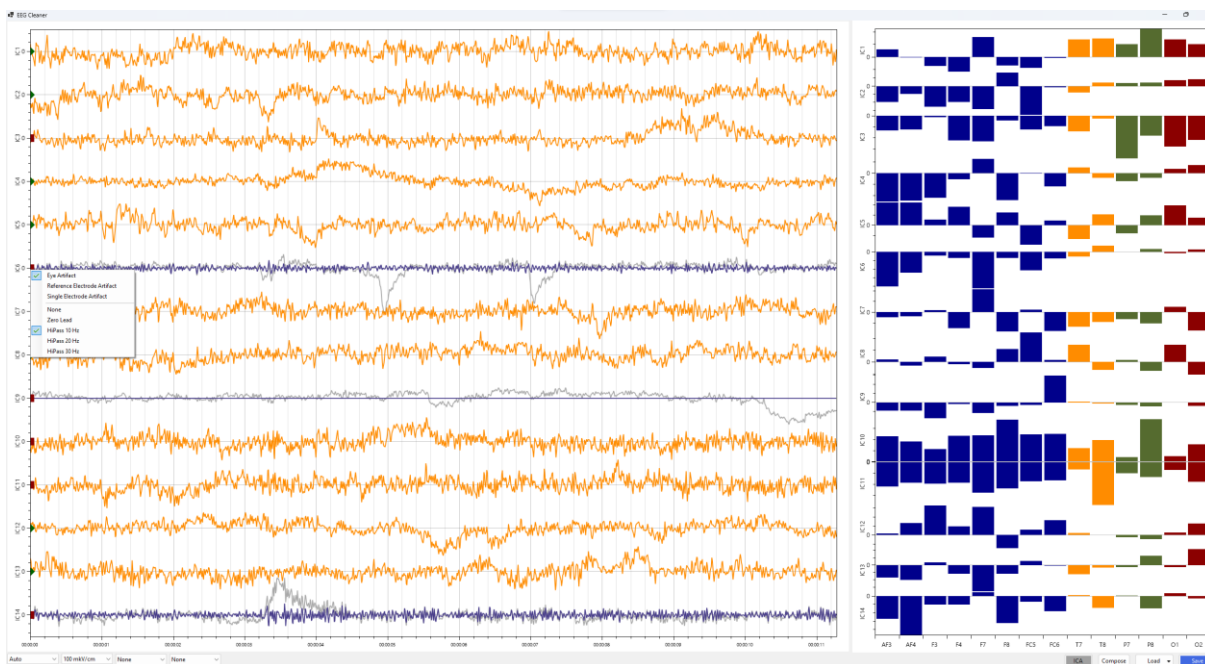


Рисунок 33 – Ручний (полу-автоматичний) режим видалення артефактів з допомогою ІСА декомпозиції та автоматичному аналізу ІСА компонентів (джерел)

Окремо слід зауважити що застосування паралельного обчислення при написанні сервісного модулю програми прискорює роботу алгоритму, так на автоматичну обробку EEG запису довжиною біля 2хв система витратила трохи більше 17 секунд. Така реалізація дозволяє спростити обробку EEG сигналів великого розміру знімаючи непотрібне навантаження на лікаря-функціоналіста.

Особливість автоматичного режиму полягає у тому, що він може бути запущений послідовно декілька разів поки артефакти «існують» у EEG датасеті. Після кожного запуску функція варіації дисперсії сигналу змінюється разом з опорної дисперсією корисного EEG сигналу, а отже зростає вірогідність знаходження на наступному кроку артефактний ділянок які були пропущені на попередньому. Але цей процес не слід використовувати бездумно – збіжність порогових алгоритмів низька якщо їх запускати ітеративно та без контролю. Можлива ситуація коли алгоритм завжди буде щось знаходити та видаляти з запису приводячи до деградації корисної складової EEG сигналу.

ВИСНОВКИ

У ході роботи були досліджені головні вимоги до попередньої обробки датасетів ЕЕГ сигналів, проаналізовані головні причини помилки при обробці ЕЕГ сигналів без попередньої обробки. Виявлено що головною завадою до коректної обробки ЕЕГ є артефактні прояви у сигналі природа яких пов'язана з наявністю неконтрольованої зміни у стані пацієнта. Серед артефактів обрані найбільш значущі у впливі на якість обробки ЕЕГ сигналів. Побудовані вимоги до усунення цих артефактів та проаналізовані існуючі методи досягання цих вимог, а також існуючі програмні засоби які реалізують повністю або частково описані методи. Обрано найкращий з оглянутих методів – аналіз незалежних компонент (ІСА), запропонований підхід для автоматичного застосування методу аналізу незалежних компонент для виявлення та видалення артефактів. Запропоновано алгоритм виявлення ділянок запису з підозрою на артефакти. Розроблені вимоги до програмної системи попередньої обробки ЕЕГ сигналів. Проведено проектування програмної системи – сервісного модулю програмної системи, додатку з інтерфейсом користувача, модулю системного тестування сервісної частини.

Після реалізації програмної системи був проведений експеримент з перевірки якості роботи систем на 4х синтетичних тестах та реальному ЕЕГ датасеті з попередньо розміченими лікарем ділянками артефактів по відеозапису стану пацієнта під час реєстрації ЕЕГ. Програмна система показала високу продуктивність в автоматичному режимі на синтетичних тестах. На реальному ЕЕГ датасеті програмна система допустила ~12% промахів у знаходженні та видаленні артефактів, що є гарним результатом. Артефакти що залишилися можуть бути видалені у ручному режимі з напівавтоматичними підказками системи.

Робота може бути продовжена для усунення існуючих недоліків системи шляхом удосконалення алгоритму пошуку ділянок артефактів та алгоритмів автоматичного визначення типу артефактів по результату ІСА декомпозиції. Перспективним напрямком було б застосування згорткових нейронних мереж для ідентифікації артефактів у ІСА декомпозиції.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Чернінський А. О., Крижановський С. А., Зима І. Г. Електрофізіологія головного мозку людини: методичні рекомендації до практикуму – К. : Видавець В. С. Мартинюк, 2011. 49 с.
2. Гусельников В. І., Електрофізіологія головного мозку. – Вища школа, 1976.
3. Зенков Л. Р. Клінічна електроенцефалографія з елементами епілептології — Вища школа, 2002
4. Luck S. J. An Introduction to the Event-Related Potential Technique – The MIT Press, 2005. 388 p.
5. Lang D. H., Inbar G. F. Modern technique in ERP research // Modern Techniques in Neuroscience Research. – NY : Springer, 1999. P. 997-1023.
6. Ivanitsky A. M., Nikolaev A. R., Ivanitsky G. A. Electroencephalography // Modern Techniques in Neuroscience Research. – NY : Springer, 1999. P. 971-996.
7. 10-20 system EEG // Wikipedia – URL: [https://en.wikipedia.org/wiki/10-20_system_\(EEG\)](https://en.wikipedia.org/wiki/10-20_system_(EEG)) (дата звернення: 01.02.2023).
8. Волошина П. В., Лінського І. В., Марути Н. О., Волошиної Н. П., Міщенко В. М., Дубенка А. Є. Клінічні рекомендації з надання медичної допомоги пацієнтам з неврологічними, психічними та поведінковими розладами — Харків : Видавець Строков Д. В., 2021. — 376 с.
9. Fernando Cross Villasana. Getting to know EEG artifacts and how to handle them in BrainVision Analyzer 2 // Brain Products – URL: <https://pressrelease.brainproducts.com/eeg-artifacts-handling-in-analyzer/> (дата звернення: 10.03.2023).
10. Smelyakov, K., Smelyakov, S., Chupryna, A. Adaptive Edge Detection Models and Algorithms // Studies in Computational Intelligence, 2020, 876, P. 1–51.
11. Krishnaveni V., Jayaraman S., Aravind S., Hariharasudhan V., Ramadoss K. Automatic Identification and Removal of Ocular Artifacts from EEG using Wavelet Transform // Measurement Science Review VI, 2005. P. 45-57.

12. Jung T-P., Makeig S., Humphries C., Lee TW., McKeown MJ., Iragui V., Sejnowski TJ. Removing Electroencephalographic Artifacts by Blind Source Separation. // *Psychophysiology* 37, 2000. P. 163-78.

13. Jung T-P., Humphries C., Lee TW., Makeig S., McKeown MJ., Iragui V., Sejnowski TJ. Removing Electroencephalographic Artifacts : Comparison between ICA and PCA // In: *Neural Networks for Signal Processing VIII*, 1998. P. 63-72.

14. Makeig S., Bell AJ., Jung T-P., Sejnowski TJ., Independent component analysis of Electroencephalographic data. // *Advances in Neural Information Processing Systems* 8, 1996. P. 145-151.

15. Gnezditsky V.V. Inverse EEG problem and clinical electroencephalography (mapping and localization of brain electrical activity sources) / *MEDPRESS INFORM*, 2004. 624 p.

16. Smelyakov, K., Honchar, Y., Bohomolov, O., Chupryna, A. Machine Learning Models Efficiency Analysis for Image Classification Problem // *CEUR Workshop Proceedings*, 2022, 3171, P. 942–959

17. Медичні комплекси функціональної діагностики // XAI-Медика – URL: <https://xai-medica.com/ua/equipments.html> (дата звертання 01.05.2023).

18. BrainVision Analyzer // Brain Products – URL: <https://www.brainproducts.com/downloads/analyzer/> (дата звертання 01.05.2023).

19. EEGLab // Swartz Center for Computational Neuroscience – URL: <https://sccn.ucsd.edu/eeglab/index.php> (дата звертання 01.05.2023).

20. BESA Research // BESA GmbH – URL: <https://www.besa.de/products/besa-research/besa-research-overview/> (дата звертання 01.05.2023).

21. Paul Nunez, Ramesh Srinivasan. Electric Fields of the Brain: The neurophysics of EEG // Oxford University Press, 2006. 611 p.

22. Solveig Næss, Chaitanya Chintaluri, Torbjørn V. Ness. Corrected Four-Sphere Head Model for EEG Signals // *Frontiers in Human Neuroscience* V.11, 2017 P. 1-7.

23. Кисельгов Е. Н. Наближене рішення проблеми референтного електроду в EEG аналізі // *Радіоелектронні і комп'ютерні системи*. – 2006. № 1. С. 115–123.

24. ARFF (Attribute-Relation File Format) // Weka Wiki – URL: https://waikato.github.io/weka-wiki/formats_and_processing/arff/ (дата звернення 24.02.2023).

25. EEG Eye State Data Set // UCI Machine Learning Repository – URL: <https://archive.ics.uci.edu/ml/datasets/EEG+Eye+State> (дата звернення 25.12.2022)

26. Detti, P. Siena Scalp EEG Database // PhysioNet, 2020 – URL: <https://doi.org/10.13026/5d4a-j060> (дата звернення 15.03.2023)