

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет навчально-науковий центр заочної форми навчання
(повна назва)

Кафедра електронних обчислювальних машин
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА
Пояснювальна записка

Рівень вищої освіти другий (магістерський)

Система аналізу напрямку погляду для
розпізнавання емоційного стану людини

(тема)

Виконав:

студент II курсу, групи КСМзм-23-1
Заброда І.С.
(прізвище, ініціали)

Спеціальність 123 «Комп'ютерна інженерія»
(код і повна назва спеціальності)

Тип програми освітньо-професійна
(освітньо-професійна або освітньо-наукова)

Освітня програма Комп'ютерні системи та мережі
(повна назва освітньої програми)

Керівник: доц. Барковская О.Ю.
(посада, прізвище, ініціали)

Допускається до захисту

Зав. кафедри ЕОМ

(підпис)

Коваленко А.А.

(прізвище, ініціали)

2025 р.

Харківський національний університет радіоелектроніки

Факультет навчально-науковий центр заочної форми навчання

Кафедра електронних обчислювальних машин

Рівень вищої освіти другий (магістерський)

Спеціальність 123 «Комп'ютерна інженерія»
(код і повна назва)

Тип програми освітньо-професійна
(освітньо-професійна або освітньо-наукова)

Освітня програма Комп'ютерні системи та мережі
(повна назва)

ЗАТВЕРДЖУЮ:

Зав. кафедри _____
(підпис)

“ _____ ” _____ 20__ р.

ЗАВДАННЯ

НА КВАЛІФІКАЦІЙНУ РОБОТУ

студенту Заброді Івану Сергійовичу
(прізвище, ім'я, по батькові)

1. Тема роботи Система аналізу напрямку погляду для розпізнавання емоційного стану людини

затверджена наказом по університету від “ 13 ” листопада 2024 р. № 189 Стз

2. Термін подання студентом роботи до екзаменаційної комісії 20 січня 2025 р.

3. Вхідні дані до роботи _____

1) Вхідні зображення очей пацієнта

2) База проявів тривожних станів за поглядом та станом очей

3) Програмний стек для рішення задач – Python, OpenCV, Dlib

4) Апаратний стек для рішення задачі – ноутбук Lenovo ThinkPad T470, зовнішня вебкамера lmx290 Starlight

4. Перелік питань, що потрібно опрацювати у роботі _____

Аналіз існуючих методів детектування області ока та виділення зіниць.

Огляд фізіологічних проявів психоемоційних станів людини за поглядом та станом очей.

Підготовка градуйованого за рівнем зашумленості сцени датасету.

Створення адаптивної нейромережевої моделі системи для визначення напрямку погляду.

Розробка методології проведення досліджень;

Проведення експериментів для виявлення залежності точності детектування зіниці від

Аналіз отриманих результатів.

5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп'ютерних ілюстрацій (слайдів) 15 слайдів

6. Консультанти розділів роботи (заповнюється за наявності консультантів згідно з наказом, зазначеним у п.1)

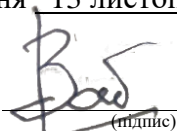
Найменування розділу	Консультант (посада, прізвище, ім'я, по батькові)	Позначка консультанта про виконання розділу	
		підпис	дата

КАЛЕНДАРНИЙ ПЛАН

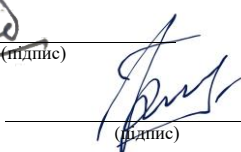
№	Назва етапів роботи	Термін виконання етапів роботи	Примітка
1	Огляд існуючих методів детектування області ока та виділення зіниць	14.11.24-23.11.24	
2	Підготовка градуйованого дасету очей	25.11.24-30.11.24	
3	Оцінка ефективності існуючих систем	02.12.24-07.12.24	
4	Розробка адаптивної нейромережевої моделі системи для визначення напрямку погляду	09.12.24-19.12.24	
5	Проведення експериментів	23.12.24-03.01.25	
6	Оформлення матеріалів кваліфікаційної роботи	04.01.25-07.01.25	
7	Подання кваліфікаційної роботи керівникові та її попередній захист	08.01.25-11.01.25	
8	Подання кваліфікаційної роботи на рецензування	13.01.25-17.01.25	

Дата видачі завдання 13 листопада 2024 р.

Студент


(підпис)

Керівник роботи


(підпис)

доц. Барковська О.Ю.

(посада, прізвище, ініціали)

РЕФЕРАТ

Пояснювальна записка кваліфікаційної роботи: 81 с., 20рис., 14 табл., 2 дод., 28джерел.

НАПРЯМ ПОГЛЯДУ, ЗІР, МІОЗ, МІДРІАЗ, DLІВ, ОРЕНСІV, МЕТОД ХОГА, МЕТОД ГАФА, RANSAC.

Метою кваліфікаційної роботи є створення системи аналізу напрямку погляду для розпізнавання емоційного стану людини при різних умовах освітленості та ступеню шумів.

У ході виконання кваліфікаційної роботи проведено аналіз існуючих методів детектування об'єктів, аналіз проблем області, а саме реакції зіниці ока на зовнішні психоемоційні подразники створення моделі системи для визначення напрямку погляду.

Подальшими кроками буде розробка та реалізація експертної системи для визначення психо-емоційного стану.

ABSTRACT

Master's thesis: 81 pages, 20 figures, 14 tables, 2 appendices, 28 sources.

DIRECTION OF GLASSES, VISION, MIOS, MYDRIASIS, DLIB, OPENCV, HOGA METHOD, GAFA METHOD, RANSAC.

The aim of the qualification work is to create a gaze direction analysis system to recognize a person's emotional state under varying lighting conditions and levels of noise.

During the execution of the qualification work, an analysis of existing object detection methods was conducted, along with an examination of the problems in the field, specifically the reaction of the pupil to external psycho-emotional stimuli, and the creation of a model for the system to determine the gaze direction.

The next steps will be the development and implementation of an expert system to determine the psycho-emotional state.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ	8
ВСТУП	9
1 ОГЛЯД ПРОБЛЕМНОЇ ОБЛАСТІ	11
1.1 Обґрунтування актуальності обраної теми	11
1.2 Фізіологічні аспекти зору людини	12
1.3 Місце та роль інтелектуальних методів для аналізу емоційного стану людини та виявлення ПТСР	19
1.3.1 Мікровирази обличчя.....	20
1.3.2 Положення тіла та напруження м'язів	21
1.3.3 Remote photoplethysmography	21
1.3.4 Розширення зіниць	23
1.3.5 Поведінкові зміни	23
1.4 Огляд існуючих досліджень виявлення ПТСР за допомогою комп'ютерного зору. Недоліки та обмеження розглянутих рішень.	24
1.5 Мета та задачі дослідження	29
2 ОБҐРУНТУВАННЯ ТА МЕТОДОЛОГІЧНЕ ПІДҐРУНТЯ ДОСЛІДЖЕННЯ	31
2.1 Математичний апарат, необхідний для аналізу психоемоційного стану.....	31
2.1.1 Random forest	31
2.1.2 Cognitive Emotion Fusion Network (CEFNet)	33
2.1.3 Faster Region-based Convolutional Neural Networks (IFR-CNN)	34
2.1.4 Deep Convolutional Neural Networks (DCNNs)	34
2.1.5 Deep Belief Networks (DBNs)	36
2.1.6 Bird's Nest Learning Analogy (BNLA)	36

2.2 Огляд можливостей фреймворків, бібліотек та апаратного забезпечення для вирішення поставленої задачі	38
3 МАТЕМАТИЧНА МОДЕЛЬ СИСТЕМИ ВИЗНАЧЕННЯ ЕМОЦІЙНОГО СТАНУ ЗА ДОПОМОГОЮ.....	47
3.1 Модель аналізу напрямку погляду для розпізнавання емоційного стану людини.....	47
3.2 Модель підсистеми комп'ютерного зору без використання нейромережових методів.....	48
3.3 Модель підсистеми комп'ютерного зору без використання нейромережових методів.....	51
3.4 Метод детектування зіниці на основі контурних методів	54
3.5 Детектування зіниці на основі методів, що аналізують градацію	55
4 РІШЕННЯ ПОСТАВЛЕНОЇ ЗАДАЧІ	57
4.1 Визначення математичного апарату для детектування очей.....	58
4.2 Вплив методів постпроцесінгу на точність визначення стану зіниці.....	62
4.3 Демонстрація працездатності запропонованої системи у реальних умовах зйомки	63
ВИСНОВКИ.....	68
ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ	69
ДОДАТОК А Графічний матеріал кваліфікаційної роботи.....	73

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ
І ТЕРМІНІВ

ЕЕГ – електроенцефалограма

ПТСР – посттравматичний стресовий розлад

СНС – симпатична нервова система

CNN – згорткова нейронна мережа

ML – машинне навчання (англ., machine learning)

PSNR – пікове співвідношення сигналу до шуму (англ., peak signal-to-noise ratio)

NLP – обробка текстів на природній мові (англ. Natural Language Processing)

rPPG – (англ. Remote Photoplethysmography)

ROI – (англ. region of interest)

ВСТУП

На сьогодні Україна переживає не легкі часи, сповнені низкою надзвичайних та екстремальних подій. Повномасштабне вторгнення вже відбувається понад два роки та близько десяти років ведуться бойові дії на території Донбасу та анексованого Криму. Всі ці події змінили відчуття безпеки та призвели до стресу, психологічні наслідки якого проявлятимуться в майбутньому.

Стрес несе негативні наслідки для організму людини. Наприклад, хронічний стрес змушує м'язи тіла перебувати в стані напруги. Що може викликати інші реакції організму і навіть сприяти виникненню розладів, пов'язаних зі стресом. Наприклад, головний біль напруги і головний біль мігрені пов'язані з хронічним напруженням м'язів в області плечей, шиї та голови. При стресі та сильних емоціях можуть проявлятися респіраторними симптомами, такими як задишка і прискорене дихання, оскільки дихальні шляхи між носом і легеньями звужуються. Гострий стрес – короткочасний або миттєвий стрес призводить до підвищення частоти серцевих скорочень і сильніших скорочень серцевого м'яза, що супроводжується виділенням гормонів стресу – адреналіну, норадреналіну та кортизолу – які є опосередковувачами цих ефектів. І навпаки, хронічний стрес, який визначається як постійний стрес, який відчувається протягом тривалого часу, може призвести до довгострокових ускладнень, що впливають на серце та кровоносні судини. [1]

Для деякого стрес може стати причиною розвитку посттравматичного стресового розладу, який несе руйнівний вплив на організм. Посттравматичний стресовий розлад (ПТСР), є серйозним захворюванням, яке потребує професійної допомоги. Виникає через надзвичайно травматичний вплив на психіку людини і проявляється багатьма психіатричними симптомами. [3, с. 584]

Важливо відстежувати зміни в психоемоційному стані учнів чи студентів та вчасно надавати психологічну допомогу.

Але військові дії, що тривають, і підвищена ракетна небезпека в деяких частинах нашої країни заважають дітям навчатися офлайн у школі, ускладнюючи вербальне спілкування між учнями та вчителями. За таких умов викладання вчасно виявити проблему та порекомендувати звернутися до відповідного спеціаліста за допомогою виявляється складно.

Для розв'язання даної проблеми можна застосовувати методи штучного інтелекту, зокрема комп'ютерного зору.

1 ОГЛЯД ПРОБЛЕМНОЇ ОБЛАСТІ

1.1 Обґрунтування актуальності обраної теми

В сучасних реаліях війни психологічна ситуація в Україні сьогодні є складною, це зумовлено військовими діями, які тривають з 2014 року і значно загострилися після повномасштабного вторгнення росії у 2022 році. Всі ці події мають серйозний вплив на емоційний стан людей в різних сферах.

Системи аналізу напрямку погляду та розпізнавання емоційного стану людини мають широкий спектр застосувань у різних галузях, що підкреслює їхню актуальність.

Розпізнавання емоційного стану людини може сприяти покращенню соціальних взаємодій, зокрема у сферах освіти, охорони здоров'я, психології та комунікацій. Наприклад, такі системи можуть допомогти викладачам розуміти емоційний стан студентів, що сприятиме покращенню освітнього процесу.

Аналіз погляду й емоційного стану використовується у діагностиці психічних розладів, моніторингу стану пацієнтів та реабілітації. Наприклад, ПТСР або гострий стресовий розлад. Це може покращити якість надання медичних послуг.

У транспортній індустрії системи аналізу погляду застосовуються для визначення рівня уваги водія, що знижує ризики аварійних ситуацій. У сфері безпеки вони можуть використовуватись для виявлення потенційно небезпечних осіб або визначення рівня стресу в критичних умовах.

Тема аналізу напрямку погляду та емоційного стану перебуває на стику декількох наукових дисциплін, таких як штучний інтелект, психологія та нейротехнології. Це сприяє розвитку міждисциплінарних підходів і створенню нових методик.

Таким чином, розробка системи аналізу напрямку погляду для розпізнавання емоційного стану людини є не лише актуальною, але й необхідною для вирішення сучасних соціальних, економічних та технічних викликів.

1.2 Фізіологічні аспекти зору людини

Сприйняття різноманітних предметів, кольорів, форм, розмірів, відстаней, просторового положення, рухів живих і неживих організмів у природі забезпечує зорова сенсорна система. Око являється частиною, зорової сенсорної системи. Воно має форму еліпсоїда, наближеного до сфери. Діаметр залежить від багатьох факторів, таких як вік, комплекція, вага та інші. Розмір ока дорослої людини приблизно 25 мм. Форма може відрізнятися залежно від раси. У представників негроїдів - витягнута, еліпсоїдна, у європеїдній раси вона більш сферична. Форма очного яблука підтримується скловидним тілом – прозорою, щільною, желеподібною речовиною. На рисунку 1.1 показана схема, яка відображає сучасне наукове уявлення про будову ока людини.



Рисунок 1.1 – Будова ока людини

Очне яблуко складається з трьох оболонок. Склера являє собою непрозору оболонку волокнистої структури, молочно-білого кольору, товщиною близько 1 мм. Завдяки своїй міцності склера захищає внутрішню частину ока від пошкоджень і деформацій. На передній, більш опуклій частині склери розташована рогівка – прозоре тіло, схоже на меніск кришталика, через яке в око потрапляє світло. Середня оболонка – це кровоносна судина або судинна оболонка, названа так тому, що в ній містяться кровоносні судини, які забезпечують живлення очного яблука. Він також входить в циліарне тіло і непрозору перегородку - райдужку.

Циліарне тіло разом із війковим смугом підтримує кришталик – еластичну внутрішню лінзу ока. Еластичність лінзи необхідна для зміни форми її зовнішньої поверхні. Від цього регулюється оптична сила Рогівка Райдужка Зіниця Водяниста волога Лінза М'язи війкового тіла Склера Судинна оболонка Зоровий нерв Сітківка Війковий смужок Склоподібне тіло Макула (macula) 15 оптичної системи ока, використовується для фокусування зображення об'єкта на сітківка. Регулювання форми кришталика забезпечується м'язами циліарного тіла. Райдужка має отвір, зіницю, діаметр якого регулюється. Регулюванням діаметра зіниці досягається необхідна освітленість сітківки, достатня для забезпечення нормальної роботи органу зору та захисту сітківки від надмірного світлового потоку.

Між рогівкою та кришталиком знаходиться передня камера, яка містить рідину, яка називається водянистою вологою. Задня камера розташована за кришталиком і заповнена щільною гелеподібною речовиною, яка називається склоподібним тілом. Він прозорий і досить оптично однорідний. Завдяки своїм механічним властивостям склоподібне тіло підтримує форму очного яблука. Слизова оболонка очного яблука покриває сітківку – систему світлочутливих рецепторів, що є гілкою зорового нерва. Він призначений як світлочутливий датчик мозку. Цей датчик дозволяє мозку

отримувати оптичну інформацію про предмети та явища в просторі перед очима.

Реакція зіниць – фізіологічна реакція при якій змінюється розмір зіниці через зоровий та окоруховий черепно-мозковий нерв.

Звужувальна реакція (міоз) [4] – звуження зіниці, відбувається, коли круговий м'яз, керований парасимпатичною нервовою системою (ПНС), стискається. Може бути спричинене наркотиками, такими як опіати/опіоїди або ліки проти гіпертонії. Звуження зіниці

Розширена реакція (мідріаз) – це розширення зіниці, відбувається, коли гладкі клітини променевого м'яза, керовані симпатичною нервовою системою (НС), скорочуються. Це може бути викликано адреналіном, антихолінергічними засобами або такими наркотиками, як екстазі, кокаїн, амфетаміни, диссоціативи та деякі галюциногени.

Реакції можуть виникати з різних причин, починаючи від мимовільного рефлексу до впливу подразників, таких як світло. В умовах недостатнього освітлення розширена зіниця забезпечує більший приплив світла в око. Крім того, розширення зіниці може вказувати на інтерес до об'єкта зосередженості, хвилювання, сексуальне збудження, невпевненість, конфлікти при прийнятті рішень, помилки або підвищений когнітивний попит. Існує сильна кореляція між цими реакціями та активністю нейромедіаторної системи locus coeruleus. Крім того, зіниці скорочуються безпосередньо перед початком швидкого сну. Крім того, реакцію зіниці можна навмисно обумовити як реакцію Павлова на певні подразники.

Латентність реакції зіниць (час, за який вона потрібна) зростає з віком. Вживання стимуляторів центральної нервової системи та деяких галюциногенних препаратів може спричинити розширення зіниці.

В офтальмології інтенсивні дослідження реакції зіниць проводяться за допомогою відеопупілометрії.

Анізокорія – це стан однієї зіниці, яка більше розширена, за іншу.

В таблиці 1.1 наведено порівняльну характеристику реакцій зіниці на подразники та на рисунку 1.2 приведені ілюстрації реакцій на подразники.

Таблиця 1.1 – Характеристика реакції зміни діаметра зіниці на подразники

Реакція зіниці	Світловий рефлекс	Зміна розміру зіниці	Інші фізіологічні причини
Міоз	Посилення освітлення	Звуження	Рефлекс акомодатії
Мідріаз	Зниження освітлення	Розширення	Бойова реакція



а)



б)

Рисунок 1.2 – Реакції зіниць: а) мідріаз, б) міоз

Середні розміри зіниць залежать від віку, статі, кольору райдужної оболонки, стану сітківки ока та зорового нерву. Але одним з найважливіших факторів, що впливає на розмір – рівень освітленості.

Синя лінія – це розмір зіниці, в певний момент часу в залежності від світлового подразника. Світловий подразник в нульовий момент часу призводить до швидкого зменшення діаметра зіниці. Латентність розраховується як час, що минув між початком світла та початком звуження. Потім зіниця швидко звужується (максимальна швидкість звуження; MCV) від базового (D_0) діаметра зіниці до мінімального (D_{min}) діаметра зіниці; час звуження (t_c) і максимальна амплітуда звуження (MCA) обчислюються як

інтервал часу і різниця розмірів між цими двома значеннями відповідно. При зміщенні світлового подразника або при тривалій світловій стимуляції зіниця переживає період швидкого розширення або «втечі» зіниці в частково звужений стан. Згодом зіниця повільно повертається до базового діаметру. На рисунку 1.3 графічно представлено динаміку зміни розміру зіниць при зміні освітлення.

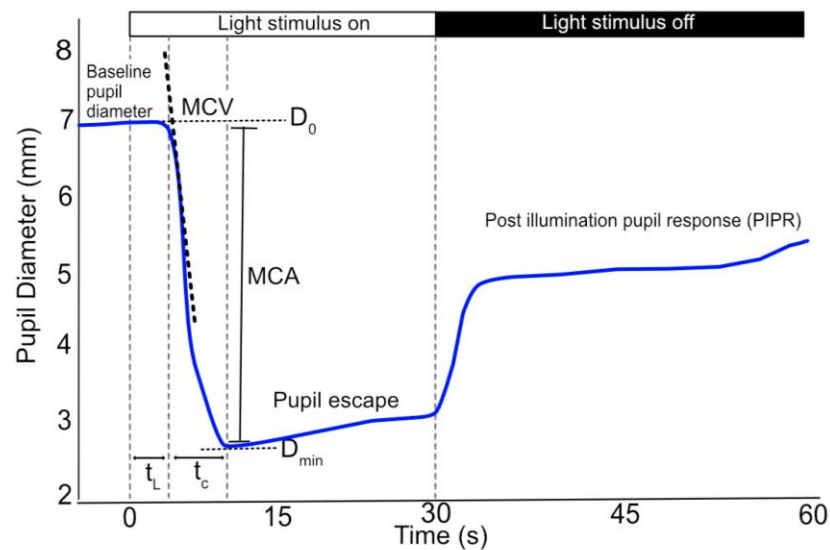


Рисунок 1.3 – Зміна діаметру зіниці в залежності від освітлення

Динамічні зміни розміру зіниці ока також можуть свідчити про емоційне збудження [9]. Симпатична система відповідає за розширення зіниць і контролює радіальні м'язи-розширювачі, активуючи систему стимуляції симпатичних нервів. За скорочення зіниці відповідає парасимпатична система, яка контролює сфінктер райдужки. Спектр емоцій людини досить широкий, і кожна емоція має свої особливості, що ускладнює завдання ідентифікації емоцій за допомогою візуального відображення. Наприклад, дослідники [9] запропонували штучну нейронну мережу, яка може виявляти лише чотири емоції, такі як: нейтральна, огида, сміх та інтерес. Запропоноване рішення для визначення емоційних станів базується на даних, зібраних візуальними датчиками, які стимулювали емоції учасників експерименту.

Крім аналізу розміру зіниць, не менш важливим є рух очей, який пов'язаний з фізіологічними процесами та когнітивною діяльністю. До них можна віднести:

- саккади – швидкі, стрибкоподібні рухи очей для зміни фокусу між об'єктами;
- мікростакади – це невеликі, мимовільні рухи очей, які відбуваються під час фіксації погляду на об'єкті;
- фіксації – моменти, коли очі затримуються на певному об'єкті для його детального розгляду;
- плавні рухи стеження – рухи, які дозволяють стежити за рухомим об'єктом;
- вестибуло-окулярний рефлекс – стабілізація погляду під час руху голови.

Діагностика за допомогою аналізу руху очей (офтальмологія, неврологія та психологія) дозволяє виявляти різні порушення, які можуть бути пов'язані із захворюваннями очей, нервової системи або навіть психічним здоров'ям. В таблиці 1.1. наведені основні порушення які можна діагностувати при аналізі руху очей пацієнта з використання методів комп'ютерного зору.

Таблиця 1.1 – Основні порушення, які можна діагностувати при аналізі руху очей пацієнта

Вид порушення	Діагноз
1	2
Неврологічні порушення	
Порушення центральної нервової системи	Інсульт (зміни в рухах очей, відсутність реакції зіниць). Черепно-мозкові травми (порушення фіксації або рухів очей). Розсіяний склероз (зміни в слідкувальному русі). Паркінсонізм (затримка рухів очей або мікросаккади).

Продовження таблиці 1.1

1	2
Порушення функції черепних нервів	Параліч очних м'язів (неможливість руху очей в певному напрямку). Косоокість (дискоординація рухів очей).
Офтальмологічні захворювання	
Порушення функцій очей	Ністагм (несвідоме тремтіння очей). Глаукома (обмеження периферійного зору). Катаракта (порушення здатності очей до фокусування). Дегенерація сітківки (зміни у здатності очей слідкувати за рухомими об'єктами).
Психічні та когнітивні розлади	
Порушення когнітивної діяльності	Синдром дефіциту уваги з гіперактивністю (СДУГ) (нестабільність фіксації погляду). Аутизм (зміни в русі очей при соціальній взаємодії). Шизофренія (нестабільність слідкувальних рухів очей). Тривожність або депресія (повільний або непослідовний рух очей).
Судинні та метаболічні захворювання	
Порушення в метаболізмі	Діабетична ретинопатія (зміни у зоровій функції, включаючи рух очей). Артеріальна гіпертензія (порушення мікроциркуляції сітківки, що впливає на рух очей).
Токсичні та медикаментозні впливи	Вживання алкоголю, наркотиків чи ліків (зміни в координації рухів очей, реакції зіниць).

1.3 Місце та роль інтелектуальних методів для аналізу емоційного стану людини та виявлення ПТСР

Щоб провести інтелектуальний аналіз проявів ПТСР треба визначити систему маркерів. Їх можна умовно поділити групи на інформаційні, візуальні, фізіологічні та акустичні.

До інформаційних відносяться: записи в медичних картках (діагнози, скарги), локації проживання, події і т. д. Це можуть бути структуровані бази даних або не структуровані. Приклад таких даних може бути проживання людини в зоні активних бойових дій. Такі бази даних можуть збиратися з слів людини та для їх аналізу використовувати методи обробки текстів на природній мові (Natural Language Processing, NLP).

В дослідженні «Diagnosing post-traumatic stress disorder using electronic medical record data» (2021 рік) описана модель аналізу структурованих та не структурованих даних нотатків в Electronic Medical Records для ідентифікації пацієнтів з діагнозом ПТСР [12]. Варто зазначити англійська мова більш структурована, ніж українська .

Візуальні дані охоплюють всю інформацію, яку можна зібрати за допомогою різних візуальних інструментів. Це включає в себе поведінкові фактори, напрямок погляду, аналіз мікросакад, мідріаз, розмір і послідовність рухів зіниць, rPPG, ЕЕГ, дослідження мікровиразів обличчя та зміни емоційних реакцій на подразники. Важливо підкреслити, що певні візуальні маркери вимагають спеціального обладнання; наприклад, оклюдери потрібні для оцінки розміру зіниці.

В дослідженні «Deep learning-based classification of posttraumatic stress disorder and depression following trauma utilizing visual and auditory markers of arousal and mood» (2020 рік) використовують поєднання комп'ютерного зору, семантичного та акустичного аналізу для виявлення ПТСР [13].

В статті «Detecting PTSD Using Neural and Physiological Signals: Recommendations from a Pilot Study» (2023 рік) у роботі описане дослідження

використання дефіциту когнітивних функцій для класифікації людей з (ПТСР) і без нього. Досліджують, чи можна використовувати нейронні та фізіологічні сигнали, такі як активність префронтальної кори, частота серцевих скорочень, дихання та електрошкірна активність, записані у поєднанні з виконанням когнітивних завдань, для покращення класифікації ПТСР. Результати вказують на те, що завдання на робочу пам'ять можуть досягти балу F1 0,80 при класифікації осіб з посттравматичним стресовим розладом, який можна додатково покращити до 0,91 шляхом поєднання мультимодальної інформації з нейрофізіологічних сигналів [14].

Діагностування ПТСР можна пов'язати з діагностуванням депресивних станів.

Акустичні – характеристики голосу та зміни швидкості мови, тональності, наявність тремору голосу, зміна гучності при інтерв'ю або аналіз змісту.

Було проведене дослідження по діагностиці ПТСР за допомогою аналізу голосу ветеранів . Згідно з дослідженням, що інструмент штучного інтелекту може розрізняти – з 89-відсотковою точністю – голоси людей з посттравматичним стресовим розладом або без нього. В дослідженні дослідники використовували алгоритм Random Forest. Дослідники спочатку записали стандартні багатогодинні діагностичні інтерв'ю під назвою «Шкала посттравматичного стресового розладу» (CAPS) 53 ветеранів Іраку та Афганістану з посттравматичним стресовим розладом, пов'язаним з військовою службою, а також 78 ветеранів без цього захворювання.

1.3.1 Мікровирази обличчя

Аналіз виразів обличчя (Facial Expression Analysis) вебкамери можуть використовуватися для аналізу мікровиразів обличчя, які є непомітними для людського ока, але можуть вказувати на психологічні стани, включаючи тривогу, страх, або стрес. За допомогою алгоритмів розпізнавання обличчя (наприклад, з використанням бібліотек OpenFace або Affectiva), вебкамери можуть

автоматично фіксувати зміни міміки, що відображають емоційний стан пацієнта.

Камера фіксує обличчя пацієнта під час експозиції на певні стимули або під час бесіди, а потім програмне забезпечення аналізує рухи м'язів обличчя, щоб визначити емоційні реакції.

1.3.2 Положення тіла та напруження м'язів

Однією з найчастіших фізичних реакцій при ПТСР є хронічне м'язове напруження. Стресові гормони, такі як адреналін і кортизол, які часто підвищені в осіб з ПТСР, змушують м'язи постійно знаходитися в стані готовності до дії. Це напруження може викликати біль у шиї, плечах, спині, головний біль та загальне почуття фізичної втоми.

1.3.3 Remote photoplethysmography

Аналіз частоти серцевих скорочень за допомогою вебкамери (Remote Photoplethysmography – rPPG) Сучасні алгоритми дозволяють вебкамерам вимірювати частоту серцевих скорочень пацієнта на основі аналізу зміни кольору обличчя, яке є наслідком пульсації крові під шкірою. Це може бути додатковим показником стресу або тривоги, оскільки частота серцебиття підвищується у стресових ситуаціях.

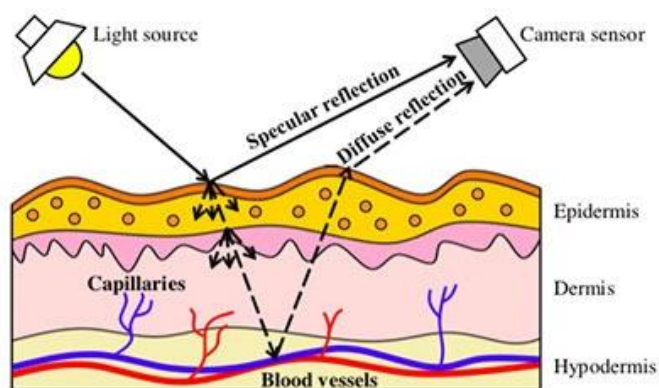


Рисунок 1.4 – Дифузне відбиття

Він вимірює варіативність відбиття червоного, зеленого та синього світла від шкіри, оскільки контраст між дзеркальним відбиттям і розсіяним відбиттям.

Дзеркальне відбиття – це чисте відбиття світла від шкіри. Дифузне відбиття – це відображення, яке залишається від поглинання та розсіювання в тканині шкіри, яке змінюється залежно від змін об'єму крові (рисунок 1.5).

Підхід до віддаленого виявлення полягає в наступному: обличчя на знятому зображенні з вебкамери визначається та моделюється з метою визначення орієнтирів на обличчі та орієнтації голови. Згодом в якості області інтересу вибирають приблизно верхні дві третини обличчя, де зосереджена велика частина кровоносних судин.

Середнє арифметичне кольорів кожного пікселя (червоний, зелений, синій) області вимірюється в часі (обидва дзеркальні + дифузні відображення).

Шум від рухів голови визначається шляхом підгонки моделі обличчя, а потім створюється частота серцевих скорочень без шуму.

Виявляючи піки, вимірюють міжударні інтервали, а потім оцінюють частоту серцевих скорочень і варіабельність серцевого ритму (рисунок 1.6).

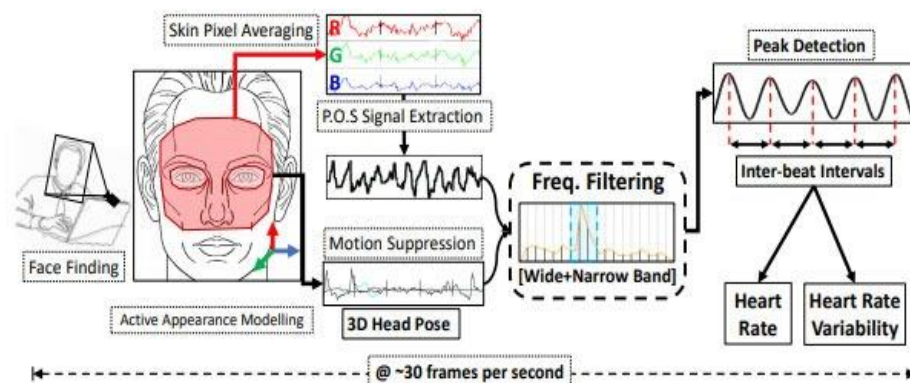


Рисунок 15 – Ілюстрація роботи RPPG

Основним недоліком даного методу, що під час споживання їжі відмінності між вимірюваннями RPPG та ECG були більшими під час споживання, що вказує на те, що RPPG недооцінював вищу частоту серцевих

скорочень. Рух обличчя під час споживання їжі порушив вимірювання RPPG. З іншого боку, ВІОРАС ECG і PPG показали високу узгодженість.

1.3.4 Розширення зіниць

Розширення зіниць при посттравматичному стресовому розладі (ПТСР) є одним із фізіологічних проявів, що можуть виникати у відповідь на стресові чи тривожні стимули. Цей ефект обумовлений активацією симпатичної нервової системи, яка готує організм до «боротьби або втечі». Основні аспекти розширення зіниць при ПТСР:

Розширення зіниць пов'язане з активацією симпатичної нервової системи, що відбувається при відчутті тривоги чи страху. При ПТСР такі реакції можуть бути більш вираженими і частішими навіть при незначних тригерах. Пацієнти з ПТСР можуть мати підвищену чутливість до стресових факторів. Це може призводити до миттєвого розширення зіниць при виникненні асоціативних чи нагадуючих стимулів.

Дослідження показують, що розширення зіниць може бути маркером суб'єктивного рівня тривоги. У людей із ПТСР підвищена чутливість нервової системи сприяє посиленню цієї реакції.

Розширення зіниць вивчається як один із можливих біомаркерів для діагностики ПТСР, оскільки це об'єктивний показник, який можна виміряти за допомогою спеціального обладнання.

Хоча розширення зіниць — це лише одна з фізіологічних реакцій, воно може допомогти у виявленні тригерів, що провокують симптоми ПТСР, а також у визначенні рівня тривожності чи стресу, пов'язаних з цим розладом.

1.3.5 Поведінкові зміни

Згорткові нейронні мережі можуть навчитися виявляти аномалії в поведінці людини, відстежувати зміни в звичках або реакції на навколишнє

середовище. Використовуючи ЗНМ у поєднанні з рекурентними мережами (наприклад, LSTM), можна аналізувати послідовності дій у часі. Це важливо для розпізнавання довгострокових змін у поведінці.

ЗНМ використовуються для аналізу виразу обличчя та рухів тіла, що допомагає визначати емоційний стан, рівень стресу або схильність до певних станів (наприклад, депресія).

Мережі також можуть адаптуватися до індивідуальних змін у поведінці користувача. Це корисно у додатках для навчання, коли система підлаштовується під ритм і потреби студента, а також у рекомендаційних системах.

Дослідження «Abnormal Behavior Detection in Surveillance CNN» для виявлення аномальної поведінки в системах відеоспостереження, таких як виявлення агресії та інцидентів у реальному часі. Модель використовує CNN разом із LSTM для аналізу послідовності дій (IEEE, 2022) [28].

1.4 Огляд існуючих досліджень виявлення ПТСР за допомогою комп'ютерного зору. Недоліки та обмеження розглянутих рішень.

Французькими вченими було опубліковане дослідження «PTSD–MDNN : Fusion tardive de réseaux de neurones profonds multimodaux pour la détection du trouble de stress post–traumatique» (2024) в ньому було проведено дослідження діагностики ПТСР та описана модель, яка об'єднує дві унімодальні згорткові нейронні мережі, яка використовує прості відео– та аудіо записи пацієнтів [15]. Варто відмітити основні недоліки – необхідність якісних датасетів для навчання та якісне обладнання.

В статті «PTSD in the Wild: A Video Database for Studying Post–Traumatic Stress Disorder Recognition in Unconstrained Environments» [16] (2022 рік) зібрили, анотували та підготували для публічного розповсюдження відеобазу даних для автоматичної діагностики ПТСР. База даних «природну» та велику варіативність умов придбання з різною позою, виразом обличчя, освітленням, фокусом, роздільною здатністю, віком, статтю, расою,

оклюзіями та походженням. Запропонований набір даних може бути використаний для бінарного навчання та класифікації ПТСР. Однак він не може бути використаний для оцінки посттравматичного стресового розладу, оскільки не проводиться опитувальник самооцінки, як тест на посттравматичний стресовий розлад (PCL). PCL-5 – це широко використовуваний самозвіт DSM-кореспондента з 20 пунктів, який оцінює 20 симптомів посттравматичного стресового розладу DSM-5.

Пропонують базову модель розпізнавання ПТСР з використанням набору даних PTSD in the-wild. Об'єкт – візуальні, звукові та текстові базові моделі показують дуже перспективні та високоефективні результати.

Однак наша базова візуальна модель дуже залежить від ручного вибору кадру співрозмовником. Продуктивність базової аудіомоделі сильно залежала від падіння двох категорій відео про авіакатастрофу та теракт через сильне втручання інтерв'юера під час інтерв'ю, а також погану якість відео та високе співвідношення сигналу до шуму. Три модальності зображень обличчя, звукових сигналів і мови дуже точно дозволяють розпізнавати і розрізняти патерни ПТСР під час інтерв'ю зі здоровими контрольними суб'єктами і суб'єктами ПТСР, які пережили травматичну подію. Високі показники базових моделей можна пояснити великою різницею між змістом двох класів: (1) перший клас інтерв'ю осіб, які пережили травматичну подію і які мають симптоми посттравматичного стресового розладу та говорять про травматичну подію та демонструють сильний емоційний стрес або фізичні реакції при нагадуванні про це, та (2) другий клас інтерв'ю на випадкові теми з акторами та громадськими діячами, прояв позитивних впливів, думок і установок. Виникає питання: чи можна розрізняти людину з ПТСР і людину з негативним афектом (наприклад, сумну людину), або людину з тривожним розладом, або також депресивну людину?

Дослідження, проведене вітчизняними науковцями ДУ «Інститут очних хвороб і тканинної терапії імені В.П. Філатова НАН України» у 2019 році, було присвячене стану учнів як показнику ПТСР серед осіб, які проживають

в АТО. зона. Отримані дані показали, що в осіб із зони АТО порівняно з контрольною групою спостерігається більший діаметр зіниці, більша відстань до найближчої точки конвергенції та значно менший об'єм акомодатії. Крім того, була виявлена пряма позитивна кореляція між діаметром зіниці та найближчою точкою конвергенції зі значенням 0,71 ($p < 0,05$) [17].

У цьому дослідженні пропонується використовувати аналізатор згорткової нейронної мережі для виявлення та оцінки діагностичного критерію посттравматичного стресового розладу, зокрема реакції очей на подразники. Це включає нестійкість фіксації погляду на предметах, що свідчить про підвищену збудливість пацієнта, а також гіперреактивність і підвищену чутливість симпатичної нервової системи до стресорів та інших подразників, таких як звуження зіниці [17].

Запропонована модель складається з двох окремих модулів: модуля автоматичної автентифікації, який використовує послідовність рухів очей людини, і модуля керування напрямком погляду. Початковий модуль спрямований на оцінку готовності користувача до процесу тестування в даний момент. Для досягнення цього введено динамічний метод автентифікації, заснований на послідовності рухів очей, викликаних командами, що відображаються на екрані монітора. Подальший аналіз візуального фокусу користувача передбачає відповідь на отриману інформацію, яка може включати або обхід запитання, або необхідність повторної авторизації [18].

Ensemble ML є перспективним підходом для покращення прогнозування розвитку ПТСР після госпіталізації у відділеннях невідкладної допомоги. Прогрес у визначенні більш корисних предикторів, які можна збирати в рутинному режимі, та усунення непотрібних, може підвищити ефективність, одночасно покращуючи точність алгоритмів, які можуть керувати прийняттям рішень пацієнтами та лікарями, які

розглядають можливість проведення цільового профілактичного втручання після травми.

В дослідженні «Діагностика посттравматичного стресового розладу за даними електронної медичної картки» дослідження спрямоване на розробку прогностичної моделі для точної ідентифікації ПТСР та пов'язаних з ним симптомів з використанням даних EMR на рівні громади первинної медичної допомоги. Модель, яка поєднує як структуровані, так і неструктуровані нотатки про зустріч у довільному тексті. Знайшли серійну модель, яка оцінювала як структуровані, так і неструктуровані дані для виявлення ПТСР, мала найвищу чутливість, F-показник та AUC [12].

Дослідження описує Cognitive Emotion Fusion Network (CEFNet), гібридну модель розпізнавання мовних емоцій (SER), яка об'єднує покращені та швидші згорткові нейронні мережі на основі регіонів (IFR-CNN), глибокі згорткові нейронні мережі (DCNNs), мережі глибоких переконань (DBNs) та аналогію навчання (BNLA). Прагнучи подолати обмеження традиційних моделей, CEFNet зосереджується на точній інтерпретації нюансованих емоційних виразів, використовуючи методи машинного навчання та всебічне вилучення ознак. Оцінено з використанням наборів даних EMODB та RAVDESS, CEFNet продемонстрували чудову продуктивність, досягнувши точності 98,11% та 91,17% на цих наборах даних відповідно, перевершуючи існуючі моделі за точністю та F1-оцінкою. Це дослідження знаменує значний внесок у SER, особливо в додатках для психічного здоров'я, пропонуючи надійну основу для розпізнавання емоцій у мовленні. Це відкриває шляхи для майбутніх удосконалень, включаючи ширше застосування в різних мовах і культурних контекстах, оптимізацію для середовища з обмеженими ресурсами та інтеграцію з іншими модальностями для більш цілісного розпізнавання емоцій.

В таблиці 1.2 наведено порівняння досліджень в сфері діагностування ПТСР.

Як видно з таблиці, більшість досліджень використовують мультимодальні дані.

Таблиця 1.2 – Порівняння досліджень в сфері діагностування ПТСР

Дослідження	Модальність	Які ознаки аналізує	Набори даних	Недоліки
1	2	3	4	5
«Diagnosing post-traumatic stress disorder using electronic medical record data» (2021)	Унімодальна	Записи в медичних картках	Electronic Medical Records	Недостатньо факторів для аналізу
«Діагностика посттравматичного стресового розладу за даними електронної медичної картки» (2000 рік)	Унімодальна	Записи в медичних картках	Electronic Medical Records	Недостатньо факторів для аналізу
«Deep learning-based classification of posttraumatic stress disorder and depression following trauma utilizing visual and auditory markers of arousal and mood» (2020 рік)	Мультимодальна	Відео, аудіо, семантичний, акустичний аналіз	Не визначено	Треба багато ресурсів
«Detecting PTSD Using Neural and Physiological Signals: Recommendations from a Pilot Study» (2023 рік)	Мультимодальна	Використовує нейронні та фізіологічні сигнали, такі як активність префронтальної кори, частота серцевих скорочень, дихання та електрошкірна активність, записані у поєднанні з виконанням когнітивних завдань	Не визначено	Треба спеціалізовані пристрої
«PTSD-MDNN : Fusion tardive de réseaux de neurones profonds multimodaux pour la détection du trouble de stress post-traumatique» (2024)	Дві унімодальні нейронні мережі	Відео– та аудіо записи	Не визначено	

Продовження таблиця 1.2

1	2	3	4	5
«PTSD in the Wild: A Video Database for Studying Post-Traumatic Stress Disorder Recognition in Unconstrained Environments» (2022)	Мультимодальна	Відео– та аудіо записи	Не визначено	
ДУ “Інститут очних хвороб і тканинної терапії ім. В.П. Філатова НАМН України” (2019)	нейромережевого аналізатора згорткового	Аналіз розміру зіниць, рухів очей людини та модуль контролю напрямку погляду	Не визначено	Аналіз проводився за допомогою високовартівського обладнання

ДУ “Інститут очних хвороб і тканинної терапії ім. В.П. Філатова НАМН України” проводив свої дослідження в умовах клініки та на високовартівському обладнанні. В роботі поставлена мета розробити модель інформаційної системи, яка буде використовувати не високовартівське обладнання та адаптована для роботи в сфері освіти. Визначити фактори які впливають на якість детектування розміру зіниці та напрямку погляду.

1.5 Мета та задачі дослідження

Методи машинного навчання добре підходять для виявлення знань і прогнозування результатів для захворювань, які мають складну етіологію та багатогранні прояви, такі як посттравматичний стресовий розлад.

Метою роботи є створення системи аналізу напрямку погляду для розпізнавання емоційного стану людини при різних умовах освітленості та ступеню зовнішніх шумів.

Для досягнення поставленої мети мають бути вирішені наступні задачі:

- аналіз існуючих методів детектування області ока та виділення зіниці;
- огляд фізіологічних проявів психоемоційних станів людини за поглядом та станом очей;
- підготовка градуйованого за рівнем зашумленості сцени датасету;

- створення адаптивної нейромережевої моделі системи для визначення напрямку погляду;
- розробка методології проведення досліджень;
- проведення експериментів для виявлення залежності точності детектування зіниці від зовнішніх умов;
- аналіз отриманих результатів.

2 ОБҐРУНТУВАННЯ ТА МЕТОДОЛОГІЧНЕ ПІДҐРУНТЯ ДОСЛІДЖЕННЯ

2.1 Математичний апарат, необхідний для аналізу психоемоційного стану

Методи інтелектуального аналізу для аналізу психологічного стану можна класифікувати за кількома категоріями, залежно від типу завдань, які вони вирішують, та методів, що використовуються (рисунок 2.1).

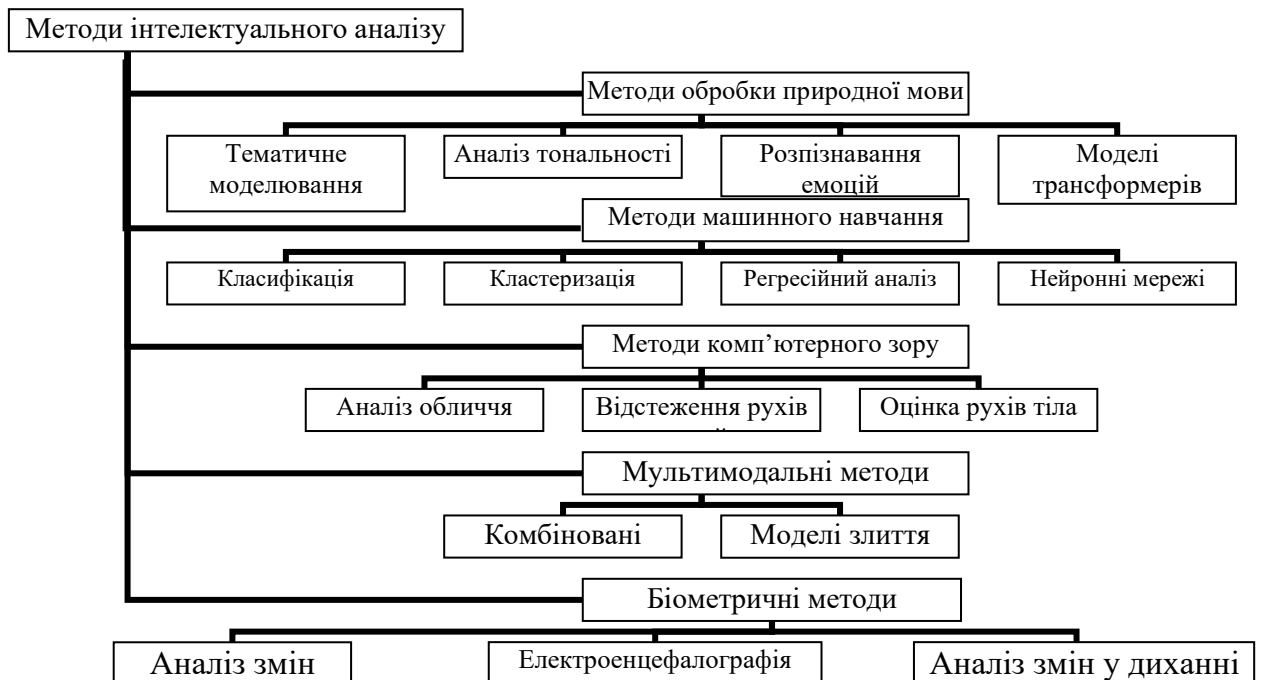


Рисунок 2.1 – Класифікація методів інтелектуального аналізу

В роботі для досягнення поставлених цілей використовуються методи комп'ютерного зору для відстежування руху очей та зміни діаметра зіниць.

2.1.1 Random forest

Random forest – це широко використовуваний алгоритм машинного навчання, розроблений Лео Брейманом та Адель Катлер, який поєднує

вихідні дані кількох дерев рішень для досягнення одного результату. Простота використання та гнучкість сприяли його прийняттю, оскільки він вирішує як проблеми класифікації, так і регресії.

Random Forest (RF) може бути корисним для класифікації психоемоційного стану, особливо коли є набір характеристик, що відображають фізіологічні або поведінкові патерни, пов'язані з різними емоційними станами.

Зазвичай, для аналізу психоемоційного стану використовуються дані про фізіологічні показники (пульс, варіабельність серцевого ритму, рівень кортизолу) або поведінкові дані (голос, вираз обличчя, текстові дані).

Наприклад, це можуть бути категорії наприклад "щасливий", "сумний", "тривожний" або шкали, що оцінюють рівень стресу.

Алгоритм Random Forest

Крок 1: Виберіть випадкові точки даних K з тренувальної вибірки.

Крок 2: Побудуйте дерева рішень, пов'язані з вибраними точками даних (підмножинами).

Крок 3: Виберіть число N для дерев рішень, які ви хочете побудувати.

Крок 4: Повторіть кроки 1 і 2.

Крок 5: Для нових точок даних знайдіть прогнози кожного дерева рішень і призначте нові очки даних категорії, яка набрала більшість голосів.

Random Forest добре справляється з шумами, тому він ефективний для даних, які можуть бути складними, неоднорідними або мати корельовані змінні.

Стаття присвячена методу вибору ознак, який спрямований на вилучення найбільш ефективних акустичних особливостей для покращення продуктивності розпізнавання емоцій. Запропоновано підбір емоційних ознак мовно-незалежного мовлення на основі Random Forest, який дозволяє прибрати зайві ознаки, що мають високі кореляції між собою. Проведено експеримент з розпізнавання мовних емоцій на основі Support Vector Machines (SVM), де для розпізнавання емоцій використовуються незалежні

від мовця ознаки, вибрані пропозицією та кореляційним аналізом Спірмена відповідно. Результати експериментів показують, що пропозиція досягає в середньому 78,6% впізнаваності, що на 2,2% вище, ніж при використанні Spearman. [19]

2.1.2 Cognitive Emotion Fusion Network (CEFNet)

У сфері розпізнавання мовних емоцій (SER), що розвивається і є важливою для розуміння та вирішення проблем психічного здоров'я, звичайні моделі часто зазнають невдач в інтерпретації складних емоційних станів, особливо тих, що пов'язані з психічними розладами, такими як посттравматичний стресовий розлад. Це дослідження представляє Cognitive Emotion Fusion Network (CEFNet), нову гібридну модель SER, яка об'єднує покращені та швидші згорткові нейронні мережі на основі регіонів (IFR–CNN), глибокі згорткові нейронні мережі (DCNNs), мережі глибоких переконань (DBNs) та аналогію навчання «Пташине гніздо» (BNLA). Прагнучи подолати обмеження традиційних моделей, CEFNet зосереджується на точній інтерпретації нюансованих емоційних виразів, використовуючи передові методи машинного навчання та всебічне вилучення ознак. Оцінений з використанням наборів даних EMODB і RAVDESS, CEFNet продемонстрував чудову продуктивність, досягнувши точності 98,11% і 91,17% на цих наборах даних відповідно, перевершивши існуючі моделі за точністю і оцінками Формули–1. Це дослідження знаменує значний внесок у SER, особливо в додатках для психічного здоров'я, пропонуючи надійну основу для розпізнавання емоцій у мовленні. Це відкриває шляхи для майбутніх удосконалень, включаючи ширше застосування в різних мовах і культурних контекстах, оптимізацію для середовища з обмеженими ресурсами та інтеграцію з іншими модальностями для більш цілісного розпізнавання емоцій.

2.1.3 Faster Region-based Convolutional Neural Networks (IFR-CNN)

Faster R-CNN – це одноступенева модель, яка навчається від початку до кінця. Він використовує нову мережу регіональних пропозицій (RPN) для генерації пропозицій регіонів, що економить час у порівнянні з традиційними алгоритмами, такими як вибіркового пошук. Він використовує шар ROI Pooling для вилучення вектора функцій фіксованої довжини з кожної регіональної пропозиції.

Одним з недоліків, який ми побачили у Faster R-CNN, є те, що для RPN всі якорі в міні-партії витягуються з одного зображення. Оскільки всі вибірки з одного зображення можуть бути корельовані (тобто їх характеристики схожі), мережі може знадобитися багато часу, перш ніж досягти конвергенції.

З огляду на це, Faster R-CNN є сучасною моделлю виявлення об'єктів. Маска R-CNN з тих пір була побудована на основі Faster R-CNN для повернення масок об'єктів для кожного виявленого об'єкта. [22]

2.1.4 Deep Convolutional Neural Networks (DCNNs)

Згорткові нейронні мережі є першими моделями глибокого навчання, які привернули багато уваги завдяки своїй вражаючій продуктивності в додатках комп'ютерного зору.

Основна ідея згорткових нейронних мереж полягає у вилученні локальних особливостей із даних. У згортковому шарі обчислюється схожість між маленькими ділянками зображення і деякими вивченими ядрами. Потім у шарі об'єднання значення пікселів, які знаходяться поблизу, групуються та об'єднуються в один піксель.

Перевага DCNN полягає в їх шаруватості. DCNNs використовує тривимірну нейронну мережу для одночасної обробки червоних, зелених і синіх елементів зображення. Це значно зменшує кількість штучних нейронів,

необхідних для обробки зображення, порівняно з традиційними нейронними мережами прямого зв'язку.

Глибокі згорткові нейронні мережі отримують зображення як вхідні дані та використовують їх для навчання класифікатора. У мережі замість множення матриць використовується спеціальна математична операція, яка називається «згорткою».

Архітектура згорткової мережі зазвичай складається з чотирьох типів шарів: згортки, об'єднання, активації та повністю підключених (рисунок 2.2).

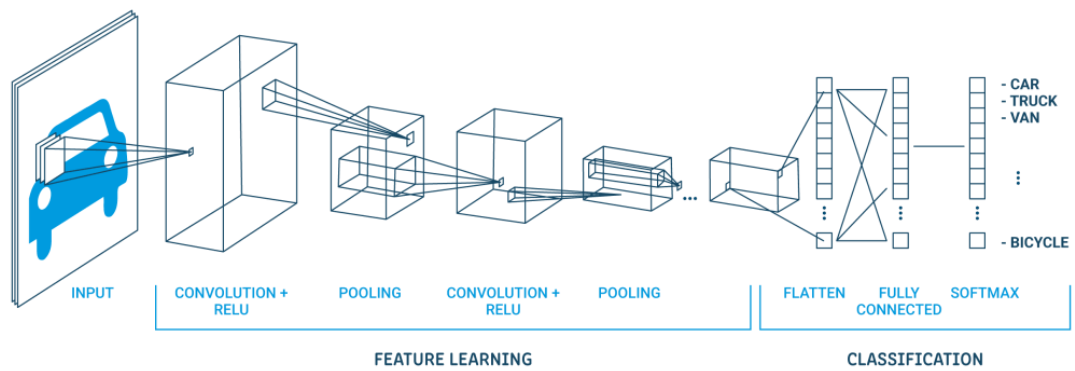


Рисунок 2.2 – Структура мережі

Це зменшує складність даних, що призводить до зменшення обсягу обчислень і вибору функцій, стійких до невеликих змін. Шари об'єднання і згорткові шари чергуються. Значення одного пікселя після першого шару об'єднання представляє цілу партію пікселів у вихідному зображенні, а після двох шарів об'єднання кожне значення представляє партію пакетів пікселів у вихідному зображенні. Перша пара шарів згортки та об'єднання поєднує пікселі зображення з локальними об'єктами, а наступні пари шарів згортки та об'єднання поєднують ці локальні ознаки з все більш і більш глобальними об'єктами. Ці глобальні ознаки потім використовуються для регресії. Як згадувалося вище, не кожен регресор може бути використаний, тому що нам потрібен сигнал помилки від регресора для адаптації ядер, що використовуються в згорткових шарах.

2.1.5 Deep Belief Networks (DBNs)

Глибока мережа переконань (DBN) – це генеративна графічна модель, або, як альтернатива, клас глибоких нейронних мереж, що складається з кількох шарів латентних змінних («прихованих одиниць»), зі зв'язками між шарами, але не між одиницями всередині кожного шару.[20]

При навчанні на наборі прикладів без нагляду DBN може навчитися імовірно реконструювати свої входи. Потім шари діють як детектори ознак.[20] Після цього етапу навчання ДБН може бути додатково навчений під наглядом для виконання класифікації.[21]

ДБН можна розглядати як склад простих, неконтрольованих мереж, таких як машини Больцмана з обмеженим доступом (RBM) або автоенкодера, де прихований шар кожної підмережі служить видимим шаром для наступного. RBM – це некерована генеративна модель на основі енергії з «видимим» вхідним шаром і прихованим шаром і зв'язками між шарами, але не всередині. Ця композиція призводить до швидкої, пошарової тренувальної процедури без нагляду, де контрастна розбіжність застосовується до кожної підмережі по черзі, починаючи з «найнижчої» пари шарів (найнижчий видимий шар – це тренувальна вибірка).

Спостереження про те, що DBN можна тренувати жадібно, по одному шару за раз, призвело до одного з перших ефективних алгоритмів глибокого навчання [21]. В цілому, існує багато привабливих впроваджень і застосувань DBN в реальних додатках і сценаріях (наприклад, електроенцефалографія, відкриття ліків).

2.1.6 Bird's Nest Learning Analogy (BNLA)

У сфері розпізнавання мовних емоцій (SER), що розвивається і є важливою для розуміння та вирішення проблем психічного здоров'я, звичайні моделі часто зазнають невдач в інтерпретації складних емоційних станів,

особливо тих, що пов'язані з психічними розладами, такими як посттравматичний стресовий розлад. Це дослідження представляє Cognitive Emotion Fusion Network (CEFNet), нову гібридну модель SER, яка об'єднує покращені та швидші згорткові нейронні мережі на основі регіонів (IFR–CNN), глибокі згорткові нейронні мережі (DCNNs), мережі глибоких переконань (DBNs) та аналогію навчання «Пташине гніздо» (BNLA). Прагнучи подолати обмеження традиційних моделей, CEFNet зосереджується на точній інтерпретації нюансованих емоційних виразів, використовуючи передові методи машинного навчання та всебічне вилучення ознак. Оцінений з використанням наборів даних EMODB і RAVDESS, CEFNet продемонстрував чудову продуктивність, досягнувши точності 98,11% і 91,17% на цих наборах даних відповідно, перевершивши існуючі моделі за точністю і оцінками F–1. Це дослідження знаменує значний внесок у SER, особливо в додатках для психічного здоров'я, пропонуючи надійну основу для розпізнавання емоцій у мовленні. Це відкриває шляхи для майбутніх удосконалень, включаючи ширше застосування в різних мовах і культурних контекстах, оптимізацію для середовища з обмеженими ресурсами та інтеграцію з іншими модальностями для більш цілісного розпізнавання емоцій.

В таблиці 2.1 проведено порівняння методів інтелектуального аналізу, які можна використовувати для досягнення поставлених цілей

Таблиця 2.1 – Порівняння методів

Модель	Призначення	Структура	Точність	Переваги	Обмеження
1	2	3	4	5	6
Random Forest	Класифікація, регресія	Множина дерев рішень	Висока	Висока стійкість до перенавчання, гнучкість	Складність для інтерпретації, потребує більше обчислювальних ресурсів

Продовження таблиці 2.1

1	2	3	4	5	6
CEFNNet	Глибоке навчання для виявлення об'єктів	Шари зворотного зв'язку	Висока	Ефективність у розпізнаванні образів	Високі вимоги до даних та обчислювальної потужності
IFR–CNN	Інтелектуальна класифікація з фокусом на деталях	Шари згортки з увагою	Дуже висока	Висока точність на деталізованих зображеннях	Часові витрати на навчання та оптимізацію
DCNNs	Класифікація, розпізнавання образів	Багатошарові згорткові шари	Висока	Стабільність у розпізнаванні складних образів	Вразливість до зміни масштабу та кута зображення
DBNs	Генеративне моделювання, відбір ознак	Шари обмежених машин Больцмана	Середня	Здатність до вивчення складної структури даних	Труднощі з навчанням, час навчання довший
BNLA	Аналіз часових рядів, прогнозування	Нейронні мережі з баєсовим навчанням	Висока	Прогнозування з урахуванням невизначеностей	Складність реалізації, потребує багато даних

2.2 Огляд можливостей фреймворків, бібліотек та апаратного забезпечення для вирішення поставленої задачі

Для вирішення даної задачі треба використовувати апаратне та програмне забезпечення. Апаратне забезпечення – це окулометр,

Окулометр – це прилад, який використовується для відстеження рухів очей. Його також називають ай–трекером (від англ. eye tracker). Основна мета окулометрів – вивчення поведінки зорової системи людини як потенційного індикатора психічного стану. Використовуючи технології, схожі на ті, що застосовуються в тракінгу очей (eye–tracking), окулометри дозволяють отримати

точні та кількісні дані, які допомагають у діагностиці ПТСР. Такі пристрої використовуються в різних галузях, наприклад, у нейромаркетингу, психології, медичних дослідженнях та у сфері розробки користувацьких інтерфейсів.

Оклуметри можуть працювати на основі камер та інфрачервоного світла для точного відстеження напрямку погляду, часу фіксації на об'єктах і рухів очей по поверхні (наприклад, екрані або тексту).

Оклуметри використовують кілька методів для вимірювання параметрів зору, таких як трекінг (рухи) очей, вимірювання діаметра зіниць, фіксація погляду та електроокулографія (EOG).

Рухи очей можуть свідчити про наявність тривожних думок або емоційних реакцій на специфічні стимули. Наприклад, при перегляді зображень або відео, пов'язаних із травматичними подіями, очі пацієнта можуть реагувати нестабільними або хаотичними рухами. Діаметр зіниць змінюється в залежності від емоційного та психічного стану людини. Під час стресу або тривоги зіниці можуть розширюватися, що є результатом активації симпатичної нервової системи. Оклуметри можуть точно вимірювати ці зміни та аналізувати їх для діагностики ПТСР. Затримка погляду на певних об'єктах або пунктах на зображенні може свідчити про емоційний зв'язок або переживання, які викликають ті чи інші стимули. Наприклад, пацієнти з ПТСР можуть довше фіксувати погляд на зображеннях, що нагадують травматичну подію, навіть не усвідомлюючи цього. Електроокулографія – вимірює електричні потенціали, які виникають в очах при їх рухах. Вона дозволяє дуже точно відслідковувати навіть незначні зміни в рухах очей.

Згорткові нейронні мережі можуть бути корисні у поєднанні з оклуметрами для аналізу руху очей, реакцій на стимули та діагностики психологічних станів, таких як посттравматичний стресовий розлад (ПТСР). Вони допомагають ефективно обробляти складні патерни у візуальних даних, що робить їх ідеальними для аналізу інформації, яку можуть надавати оклуметри [22].

В таблиці 2.2 наведено порівняння бібліотек для роботи з пристроями.

Таблиця 2.2. – Порівняльна характеристика бібліотек

Назва бібліотеки	Опис	Можливості	Мови програмування	Платформи	Обладнання	Ліцензія
1	2	3	4	5	6	7
Tobii Pro SDK	бібліотека розроблена компанією Tobii, одним із провідних виробників систем трекінгу очей. Вона дозволяє розробникам інтегрувати дані трекінгу очей у свої додатки.	Отримання даних траєкторії рухів очей. Вимірювання положення зіниць та діаметра. Підтримка реального часу для інтерактивних додатків	Python, C#, C++, MATLAB	Windows, macOS	Оклендтри	
Pupil Labs (Pupil SDK)	Платформа для трекінгу очей із підтримкою апаратного забезпечення та програмних інструментів для досліджень.	Відстеження рухів очей у реальному часі. Налаштовувані модулі для збору й аналізу даних	Python, C++	Windows, macOS, Linux	Різні сенсори	LGPL–3.0, GPL–3.0
EyeLink (SR Research)	система для трекінгу очей, яка пропонує SDK для інтеграції із сторонніми додатками	Висока точність вимірювання рухів очей. Підтримка досліджень із використанням зорових стимулів. Можливість аналізу даних у реальному часі.	C, C#, Python, MATLAB	Windows, macOS	Різні сенсори	

Продовження таблиці 2.2

1	2	3	4	5	6	7
OpenFace	Хоча OpenFace більше спрямована на аналіз міміки обличчя, вона також підтримує тракінг очей. Це відкрита бібліотека для вивчення емоцій та виразів обличчя, а також рухів очей.	Відстеження рухів очей. Аналіз виразів обличчя та емоційних реакцій. Підтримка відео в реальному часі.	Python, C++.	Windows, Linux, macOS	Різні сенсори	MIT
PyGaze	бібліотека спеціалізована на тракінгу очей і використовується для наукових досліджень.	Простий доступ до даних тракінгу очей. Інтеграція з апаратними рішеннями для різних пристроїв, таких як Tobii, SMI та інші.	Python	Windows, macOS, Linux	Різні сенсори	GPL–3.0
OpenCV	Бібліотека комп'ютерного зору, обробки зображень і чисельних алгоритмів загального призначення.	Бібліотека для комп'ютерного зору, дозволяє виконувати базовий тракінг очей і аналіз міміки обличчя за допомогою веб-камери.	Python, C++, C, Java,	Windows, macOS, Linux	Різні камери	Apache License 2.0

Продовження таблиці 2.2

DeepFace		Інструмент для аналізу обличчя, який використовує глибоке навчання для розпізнавання емоцій.				
Affectiva	Платформа для аналізу емоцій	Платформа для аналізу емоцій на основі міміки обличчя, яка дозволяє виявляти стресові стани. Може бути використана разом із веб-камерами для дистанційного моніторингу пацієнтів				

Веб-камери стають доступним і поширеним інструментом для детекції посттравматичного стресового розладу (ПТСР), особливо у поєднанні з алгоритмами машинного навчання та аналізом зображень. Сучасні веб-камери, попри свою відносну простоту, можуть використовуватися для фіксації таких даних, як міміка обличчя, рухи очей і зміни в положенні голови та тіла, що можуть вказувати на емоційні та психологічні стани. Завдяки програмному забезпеченню та спеціалізованим алгоритмам, веб-камери стають потужним інструментом для аналізу стресу та тривожних розладів, таких як ПТСР.

Відстеження рухів очей (Eye-Tracking) Хоча спеціалізовані камери для трекінгу очей зазвичай використовуються для високоточної фіксації рухів

очей, звичайні веб–камери також можуть виконувати базовий трекінг очей за допомогою спеціального програмного забезпечення. Це дозволяє визначити, куди спрямований погляд пацієнта, чи відбуваються типові рухи, пов'язані зі стресом або униканням, а також аналізувати діаметр зіниць як маркер стресових реакцій.

Веб-камера фіксує позицію очей і рухи з допомогою алгоритмів, таких як Dlib або OpenCV. На основі цих даних можна оцінити, як пацієнт реагує на візуальні стимули.

Варто відмітити, що менша точність у порівнянні зі спеціалізованими окуметрами, але достатня для базового аналізу у контексті телемедицини.

Моніторинг змін в поведінці (Behavioral Analysis) Веб-камери можуть використовуватися для аналізу загальної поведінки пацієнта під час сеансу. Зокрема, відстеження рухів голови, зміни позиції тіла або жестів можуть дати додаткову інформацію про стан тривоги чи стресу.

Вебкамера фіксує рухи голови та тіла, після чого алгоритми обробки зображень аналізують дані для виявлення потенційних ознак нервовості або стресу (наприклад, часті зміни пози, рухи голови чи рук).

Переваги: Доступність і можливість використання в умовах обмеженого бюджету.

Вебкамера записує відео обличчя пацієнта, а алгоритми rPPG аналізують зміни в яскравості шкіри для оцінки частоти серцевих скорочень.

Переваги: Безконтактний спосіб вимірювання фізіологічних параметрів, який може бути використаний для моніторингу в режимі реального часу.

Вебкамера активно використовуються у телемедичних консультаціях для моніторингу психічного здоров'я пацієнтів. Вони дозволяють лікарям спостерігати за емоційними реакціями та поведінкою пацієнта під час дистанційних сесій, що особливо актуально для людей із ПТСР, які можуть відчувати труднощі з відвідуванням терапевта особисто.

Як працює: Під час сеансів телемедицини вебкамера фіксує вирази обличчя та загальну поведінку пацієнта. Лікар може в реальному часі аналізувати стан пацієнта та за необхідності коригувати лікування.

Переваги використання веб-камер у детекції ПТСР в їхній доступності. Веб-камери є повсюдними та доступними як на комп'ютерах, так і на мобільних пристроях, що дозволяє використовувати їх для масового моніторингу. Веб-камери ідеально підходять для дистанційної терапії, що дозволяє пацієнтам з ПТСР отримувати лікування вдома.

Зі спеціальними алгоритмами веб-камери можуть фіксувати й аналізувати емоційні реакції та фізіологічні показники в реальному часі.

Обмеженість в оцінці фізіологічних параметрів: Незважаючи на можливість аналізу частоти серцевих скорочень через rPPG, веб-камери все ще обмежені в точності у порівнянні з медичними сенсорами.

Веб-камери є важливим інструментом для дистанційного моніторингу та діагностики ПТСР, особливо в умовах телемедицини. Завдяки сучасним алгоритмам обробки зображень та аналізу міміки, веб-камери можуть забезпечити корисну інформацію для оцінки емоційних станів пацієнтів. Незважаючи на деякі обмеження щодо точності та умов освітлення, цей інструмент стає все більш поширеним для моніторингу психічного здоров'я

Тепловізійні камери (Thermal Cameras) можуть фіксувати незначні зміни в температурі обличчя, які пов'язані з емоційними реакціями людини. Такі зміни можуть бути маркерами стресу, тривоги чи інших емоційних станів, характерних для пацієнтів з ПТСР. Ці камери використовуються для:

Вимірювання стресової реакції: Підвищення температури на деяких ділянках обличчя, таких як ніс або лоб, може вказувати на підвищений рівень стресу.

Аналіз змін у кровообігу: Стрес і тривожні стани часто впливають на кровообіг у верхній частині обличчя, що можна спостерігати через теплові камери.

FLIR Systems – лінійка тепловізійних камер, що можуть бути використані в психологічних дослідженнях для вимірювання змін у температурі шкіри під час експозиції на стресові стимули.

Таблиця 2.3 – Порівняння технічних засобів

Пристрій	Що можуть відслідковувати	Вартість	Обмеження	Переваги
Окулометр	трекінг (рухи) очей, вимірювання діаметра зіниць, фіксація погляду та електроокулографія	Висока	апаратне забезпечення може мати низьку роздільну здатність або недостатню точність, що призводить до похибок у відстеженні погляду. У деяких системах може виникати затримка, що призводить до неточностей у вимірюваннях.	Можна якісно відслідковувати кліпання очей
Веб-камери	міміка обличчя, рухи очей і зміни в положенні голови та тіла	Не велика	Звичайні веб-камери мають меншу точність, ніж спеціалізовані медичні камери або окулометри. Якість даних може залежати від умов освітлення, що впливає на точність аналізу виразів обличчя та рухів очей.	Не потребує спеціального обладнання; можна використовувати в домашніх умовах для телемедичних консультацій
Тепловізійні камери	Можна використовувати для температури шкіри	Середня	Вплив умов навколишнього середовища	Можна якісно відслідковувати температури
Інфрачервоні камери	Захоплення інфрачервоного спектра	Середня	Залежність від освітлення, чутливість до перешкод	Якісне відслідковування змін
RGB-D камери	Аналіз рухів тіла та голови Вимірювання відстані та пози	Середня	Чутливість до умов освітлення, обмеження у приміщеннях	Якісне відслідковування змін в поведінці

Камери для розпізнавання міміки використовуються для аналізу емоційних виразів обличчя пацієнтів під час тестування або в стресових

ситуаціях. Зміни міміки обличчя можуть бути показовими для діагностики ПТСР, оскільки певні емоційні вирази пов'язані з тривогою або стресом. Ці камери аналізують: мікрОВИрази обличчя, зміни в емоційних реакціях

RGB-D камери можуть збирати як колірну інформацію (RGB), так і дані про глибину. Це дозволяє більш точно відслідковувати рухи тіла, голови та обличчя, що може бути корисно для вивчення реакцій на стресові ситуації або візуальні стимули, що використовуються в діагностиці ПТСР.

Microsoft Kinect – це приклад глибинної камери, яка може використовуватися для відстеження рухів тіла та обличчя в реальному часі під час досліджень із детекції стресу та тривожності.

3 МАТЕМАТИЧНА МОДЕЛЬ СИСТЕМИ ВИЗНАЧЕННЯ ЕМОЦІЙНОГО СТАНУ ЗА ДОПОМОГОЮ

3.1 Модель аналізу напрямку погляду для розпізнавання емоційного стану людини

Аналіз напрямку погляду для розпізнавання емоційного стану являється між дисциплінарним завданням, таких як комп'ютерний зір, машинне навчання, нейрофізіологію та психологію.

Вхідними даними для аналізу напрямку погляду є потік відеокадрів зі швидкістю не нижче 30 кадрів за секунду. Апаратними пристроями можуть бути камери або модулі трекінгу погляду. В даній моделі передбачається що респондент сидить навпроти камери, яка встановлена на моніторі.

Загальна модель аналізу напрямку погляду для розпізнавання емоційного стану людини складається з двох основних блоків. Підсистема комп'ютерного зору забезпечує підготовку кадру для подальшого визначення напрямку погляду та часу фіксації. Що є підготовкою даних для роботи експертної (медичної) підсистеми прийняття рішень. Результатом якої є діагностований емоційний стан (дивись рисунок 3.1).

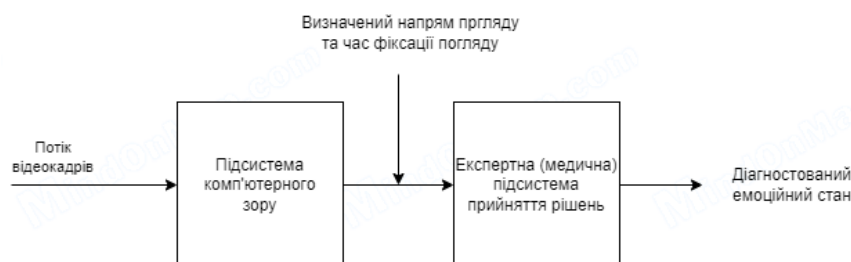


Рисунок 3.1 – Загальна модель аналізу напрямку погляду для розпізнавання емоційного стану людини

Існує безліч підходів для виділення ознак для детектування об'єктів, зокрема методи, які працюють на градієнтах та контурах. Методи що

працюють на основі градієнтів визначають зміни інтенсивності пікселів у зображенні, знаходячи області з різкими перепадами яскравості. На основі контурів виділяють замкнуті області або межі об'єктів на зображенні.



а)



б)

Рисунок 3.2 Методи виділення ознак. а) контури, б) градієнти

Контурні методи мають кілька переваг порівняно з градієнтними підходами, особливо в задачах оптимізації та аналізу функцій. Вони дозволяють більш ефективно досліджувати глобальну структуру цільової функції, оскільки використовують рівневі лінії або поверхні, що допомагає уникати локальних екстремумів і нестійких напрямків градієнта. Крім того, контурні методи можуть бути менш чутливими до початкових умов і часто забезпечують більш стабільну збіжність, особливо в задачах з нерівномірними або складними ландшафтами функцій.

3.2 Модель підсистеми комп'ютерного зору без використання нейромережових методів

Підсистема комп'ютерного зору складається з наступних етапів: попередня обробка даних, виділення ROI, покращення виділеної області, виділення ROI', пост процесинг (дивись рисунок 3.1).

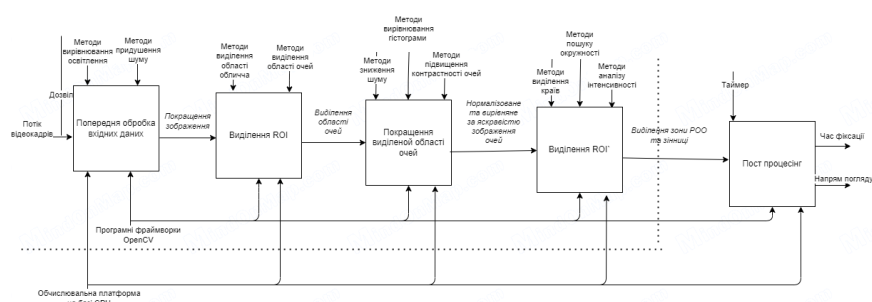


Рисунок 3.3 – Модель підсистеми комп'ютерного зору

На етапі попередньої обробки вхідних даних виконується визначається рівень шуму та освітленості кадру. Для визначення освітленості зображення переводимо кадр в колірну модель HSV та як ознаку беремо параметр V. На даному етапі нам важливо покращити освітлення на сцені тому ми використовуємо CLANE.

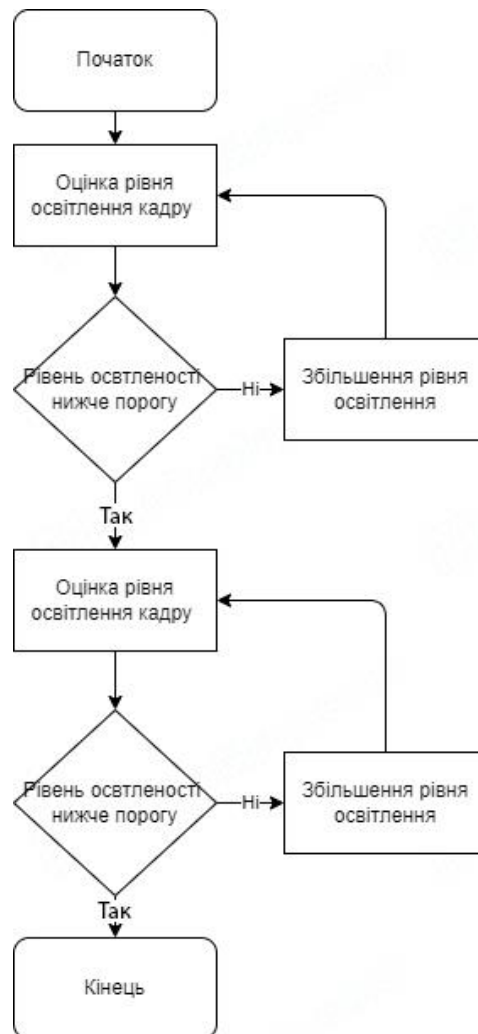


Рисунок 3.4 – Процес попередньої обробки даних

Рівень шуму визначаємо за допомогою пікового співвідношення сигнал/шум (PSNR). Вона дозволяє оцінити процес нормалізації кадру.

Об'єкт, схожий за глобальною ознакою, складається з темних і світлих областей. Він дає єдине значення, беручи різницю суми інтенсивностей темних областей і суми інтенсивностей світлих областей. Це робиться для

вилучення корисних елементів, необхідних для ідентифікації об'єкта. Особливості, методом Віоли-Джонса [23], дивись рисунок 3.5.

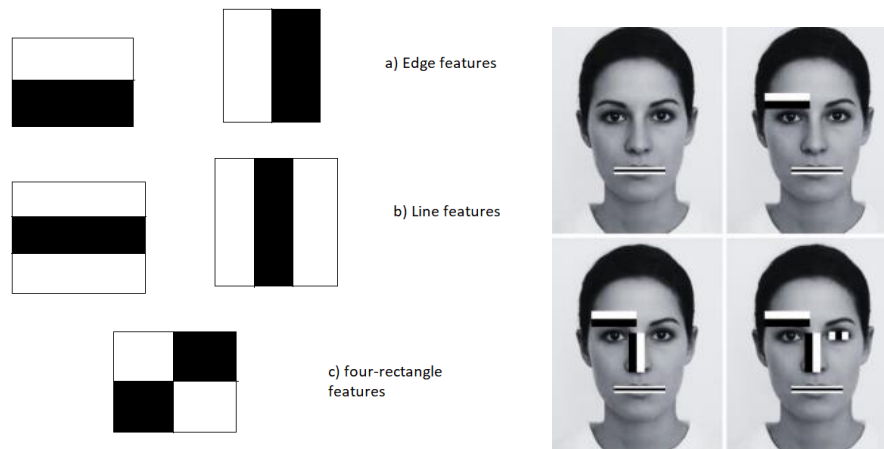


Рисунок 3.5 – Виділення областей за Віолою і Джонсом [7]

Етап виділення ROI – це процес виділення зі сцени обличчя та зони очей. Спочатку виділяємо обличчя за допомогою каскадного класифікатора, який працює за підходом Віоли-Джонса.

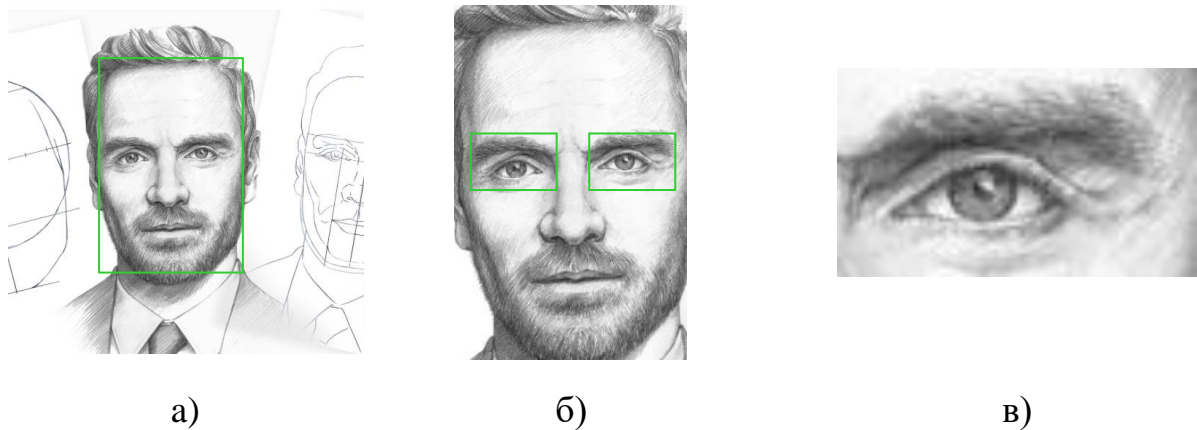


Рисунок 3.6 – Процес виділення очей: а) виділення із сцени зони обличчя; б), в) виділення зони очей

Метод Віоли-Джонса є одним із найефективніших підходів для швидкого виявлення облич у зображеннях завдяки використанню каскадної

класифікації на основі хаар-подібних ознак. Його головними перевагами є висока швидкість роботи, оскільки алгоритм застосовує інтегральне зображення для швидкого обчислення ознак, та ефективність, забезпечена каскадною структурою, яка відсіює більшість нерелевантних областей.

3.3 Модель підсистеми комп'ютерного зору без використання нейромережових методів

Модуль підготовки зображень призначений для попередньої обробки, тому вхідні зображення можуть бути зроблені в різних умовах освітленості та різного рівня зашумленості. Перед роботою обробкою визначається рівень шуму та освітленості. Якщо параметри нижче порогових значень, то виводимо повідомлення про неможливість роботи моделі в цих умовах освітленості чи зашумленості.

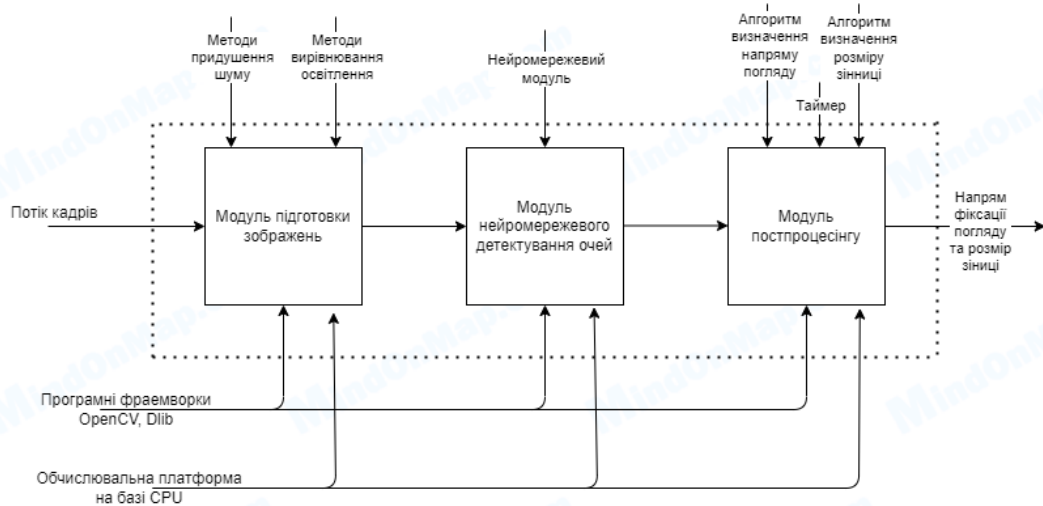


Рисунок 3.7 – Загальна модель аналізу напрямку погляду для розпізнавання емоційного стану людини

Для рівня освітлення визначається такими параметрами: Grayscale Brightness, HSV Brightness (V), LAB Brightness (L). Grayscale Brightness – це середнє значення яскравості зображення у відтінках сірого. Визначається як:

$$B_{gray} = \frac{1}{N} \sum_{i=1}^N I(i), \quad (3.1)$$

де, $I(i)$ – інтенсивність пікселів у градаціях сірого,

N – загальна кількість пікселів.

HSV Brightness (V) – це компонент «Value» у просторі кольорів HSV, який визначає максимальне значення інтенсивності серед каналів R, G і B:

$$V = \max(R, G, B), \quad (3.2)$$

де, R,G,B – компоненти червоного, зеленого та синього кольорів (нормалізовані до $[0,1]$ або 0-255).

Відображає найсильніше світлове випромінювання серед усіх кольорових складових.

LAB Brightness (L) – це компонент «L» у просторі LAB, який моделює людське сприйняття яскравості. Визначається як перетворення з простору RGB через XYZ-простір. Значення L змінюється від 0 (чорний) до 100 (білий).

Для аналізу якості зображення використовується пікове співвідношення сигналу до шуму (PSNR), вираховане через середньоквадратичне відхилення (MSE) для двох монохромних зображень.

$$PSNR = 20 \log_{10} \left(\frac{MAX_I}{\sqrt{MSE}} \right), \quad (3.3)$$

де MAX_I – це максимальне значення, яке приймається пікселем зображення. Коли пікселі мають розрядність 8 біт, $MAX_I = 255$. В загальному випадку, коли значення сигналу представлені лінійно (PCM) з 8 бітами на кожне значення, максимально можливе значення MAX_I буде $8B-1$.

Для того, щоб покращити точність детектування необхідно застосовувати методи боротьби з шумами на зображенні застосовуються для

поліпшення якості зображень у фотографії. Шум може бути спричинений сенсорами камери, передачею даних. Для розв'язання даної задачі досить ефективний Гаусівський фільтр.

Методи вирівнювання освітленості на зображенні застосовуються для усунення нерівномірного освітлення, покращення видимості деталей та підготовки зображення до подальшої обробки. CLAHE (Контрастно-обмежене адаптивне гістограмне вирівнювання) – це вдосконалений метод корекції освітленості, який має кілька переваг перед стандартним глобальним гістограмним вирівнюванням (HE) та адаптивним HE (AHE). CLAHE ефективний, бо дозволяє адаптивно покращити контраст, не спотворюючи зображення та не підсилюючи шум. Це робить його ідеальним вибором для застосувань у комп'ютерному зорі, медичній візуалізації та інших сферах обробки зображень.

Модуль неймережевого детектування лиця працює на основі методу Histogram of Oriented Gradients (HOG) + Support Vector Machine (SVM). Він добре працює для фронтальних облич, але менш ефективний для поворотів або частково закритих облич. Але для рішення поставленої задачі підходить.

Була обрана навчина попередньо модель `shape_predictor_68_face_landmarks` [24], що містить 68 ключових точок обличчя (facial landmarks). Вона включає в себе точки контурів обличчя, брів, очей, носу, губ.

Модуль постпроцесінгу складається з двох функцій визначення напрямку погляду та визначення розміру зінниць. Напрямок погляду виначається через положення зінниці. Знаючи що зінниця має круглу форму, визначення розміру зінниці відбувається за допомогою методу RANSAC та CircleModel.

При мікросміні погляду записуємо дані в лог для подальшого аналізу.

3.4 Метод детектування зінниці на основі контурних методів

Етап 1. Виділення контуру ока. Контур визначається як послідовність пікселів, що лежать на границі об'єкта (області з ненульовими значеннями). Це можна представити у вигляді множини точок:

$$C_k = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}, k = 1, \dots, K,$$

де K – кількість знайдених контурів, а кожен контур C_k складається з набору координат точок.

Об'єднуємо всі контури у один великий масив, де кожен рядок – це точка з контуру.

$$P = \bigcup_{k=1}^K C_k$$

Етап 2. Визначаємо райдужну оболонку

Нехай $P = \{p_1, p_2, \dots, p_m\}$ – набір точок, де кожна точка має координати $p_i = (x_i, y_i)$.

Коло в можна представити рівнянням:

$$(x - x_c)^2 + (y - y_c)^2 = r^2,$$

де (x_c, y_c) – центр кола, а r – його радіус.

Якщо знаходимо 1-ше коло, обрізаємо зображення по розміру кола повторюємо процес. 2-ге коло буде нашою зінницею.

3.5 Детектування зінниці на основі методів, що аналізують градацію кольорів

Етап 1. Адаптивна порогова обробка обчислює локальний поріг $T(x, y)$ для кожного пікселя $I(x, y)$ за формулою:

$$T(x, y) = \frac{1}{S} \sum_{(i,j) \in W} I(i, j) - C,$$

де: W – вікно розміру, центроване навколо (x, y) ,

$G(i, j)$ – коефіцієнти Гауссового ядра,

S – сума всіх коефіцієнтів ядра $G(i, j)$,

C – від'ємна корекційна константа (у коді $C=2$).

Бінаризоване зображення $B(x, y)$ отримує значення за правилом:

$$B(x, y) = \begin{cases} 255, & \text{якщо } I(x, y) < T(x, y) \\ 0, & \text{інакше} \end{cases}.$$

Етап 2.

Коло в можна представити рівнянням:

$$(x - x_c)^2 + (y - y_c)^2 = r^2$$

де (x_c, y_c) – центр кола;

а r – його радіус.

Вхідними даними для алгоритму RANSAC є набір спостережуваних значень даних, модель для відповідності спостереженням, а також деякі довірчі параметри, що визначають викиди. RANSAC досягає своєї мети, повторюючи наступні кроки:

Вибераємо випадкову підмножину вихідних даних. Назвемо цю підмножину гіпотетичними вкладниками.

До набору гіпотетичних вкладишів підганяється модель.

Потім усі дані порівнюються з встановленою моделлю. Всі точки даних (вихідних даних), які добре підходять до оціненої моделі, відповідно до деякої функції втрат для конкретної моделі, називаються консенсусною множиною (тобто множиною індикаторів для моделі).

Оцінена модель є досить хорошою, якщо достатня кількість точок даних була класифікована як частина консенсусного набору.

Модель може бути вдосконалена шляхом її переоцінки за допомогою використання всіх членів консенсусної множини. Якість апроксимації як міра того, наскільки добре модель вписується в консенсусний набір, буде використовуватися для підвищення чіткості примірки моделі в міру продовження ітерацій (наприклад, шляхом встановлення цього показника як критерію якості підгонки на наступній ітерації).[11]

Варто зазначити що, RANSAC є ітеративним алгоритмом для оцінки параметрів моделі в умовах значної кількості викидів. Він стійкий до шумів та працює досить швидко. Але має ряд недоліків: результат залежить від випадкового вибору підмножини точок, що може призводити до варіативності результатів; Може знайти лише "достатньо хороше" рішення, а не глобально оптимальне.

RANSAC є потужним інструментом для пошуку відповідностей у зображеннях, особливо в умовах великої кількості шуму та викидів. Однак його ймовірнісний характер і залежність від параметрів можуть вимагати додаткового налаштування для конкретних завдань.

4 РІШЕННЯ ПОСТАВЛЕНОЇ ЗАДАЧІ

Для реалізації моделі та проведення експериментів було обрано мову програмування Python, із-за простоти та швидкості написання програмного забезпечення та багатим набором потужних бібліотек для роботи з комп'ютерним зором та машинним навчанням.

В проведених експериментах використовувалися наступні бібліотеки:

`numpy` – бібліотека для числових обчислень у Python. Надає підтримку масивів і матриць, а також велику кількість математичних функцій для роботи з ними.

`opencv-python` – Обгортка Python для бібліотеки OpenCV, яка використовується для обробки зображень та комп'ютерного зору.

`Dlib` – це високопродуктивна бібліотека машинного навчання та комп'ютерного зору, написана мовою C++ з обгорткою для Python. Вона широко використовується для задач, пов'язаних з обробкою зображень, комп'ютерним зором та глибоким навчанням.

В дослідженнях використовувалося дві камери: середньостатистична, яка встановлена на більшості ноутбуків (вбудована в ноутбук), зовнішня з покращеними характеристиками. Порівняльну характеристику дивись в таблиці.

Таблиця 4.1 – Порівняльна характеристика камер

Параметр	Вбудована в ноутбук	Зовнішня
Роздільна здатність	720p	Full-hd
Частота кадрів	30 FPS	60 FPS
Апаратна обробка кадру	ні	так

Характеристики зовнішньої камери значно переважають камеру ноутбука.

4.1 Визначення математичного апарату для детектування очей

Для визначення якості детектування очей було взято частину (50) датасету BioID Face Database [24]. Це загальнодоступна база даних зображень облич для досліджень у сфері біометрії, комп'ютерного зору та розпізнавання облич. Вона відображає реальні умови з різними варіантами освітлення.

Як показано в таблиці 4.1, метод що має в собі основу нейронну мережу на 6% показав себе краще.

Таблиця 4.2 – Результати дослідження точності детектування очей на вибірці BioID Face Database

Метод	Кількість вірних знаходжень обох очей	У відсотковому співвідношенні
Точність детектування очей без неронної мережі	47 / 50	94%
Точність детектування очей з використанням НМ	50 / 50	100%

Таким чином, результати дослідження, проведені на датасеті BioID Face Database показали, що використання нейромережевих методів детектування очей дають перевагу у 6 %. Це обумовлює подальші дослідження впливу методів препроцесінгу лише за умов детектування очей на основі НМ методів.

Для перевірки точності детектування зінниці в різних умовах зашумленості було підготовано маркований датасет з різною градацією зашумленості (10%, 20%, 50%, 80%)(рисунок 4.1).

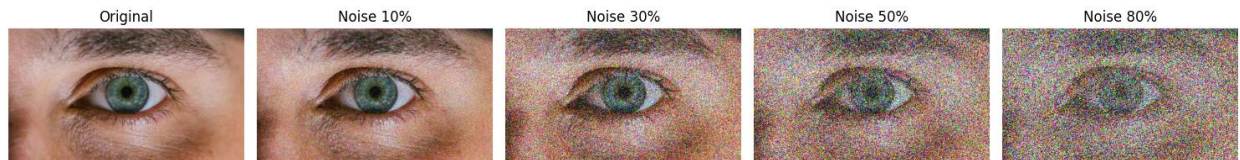


Рисунок 4.1 – Приклад градуїованого датасету

В таблиці 4.3 представлено результати досліджень детектування зіниці за допомогою градуїованого дата сету. Дослідження проводилося в два етапи: без шумопригнічення, з пригніченням шуму.

Таблиця 4.3 – Точність детектування зіниці з шумом на прикладі підготовленого градуїованого за ступенем шуму датасету

Наявність методів шумопригнічення	Без шумо пригнічення				З пригніченням шуму			
	10%	30%	50%	80%	10%	30%	50%	80%
Ступінь зашумленості								
Препроцесінг на основі градієнтів	100%	100%	-	-	100%	100%	100%	-
Препроцесінг на основі контурів	100%	100%	100%	-	100%	100%	100%	100%

Препроцесінг на основі контурів має більшу ефективність, тому що менше відчутливий до шумів. На рисунку 4.2 преведений приклад зобрвження зі шумом в 50% до пригнічення шуму та після.



а)



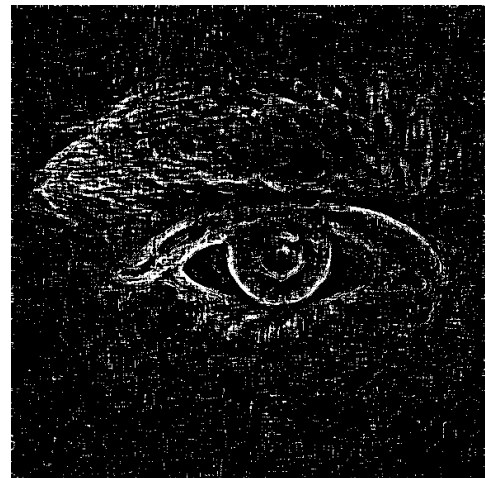
б)

Рисунок 4.2 – Приклад шумопригнічення: а) – зображення з шумом 50%; б) після шумопригнічення

Вхідним зображенням (рисунок 4.2) для виділення ознак для подальшого детектування. На рисунку 4.3 показані підготовлені зображення для детектування.



а)



б)

Рисунок 4.3 – Підготоване зображення різних методів: а) – градієнтів); б) - контурів

Точність детектування зінниці залежить від обраного методу препроцесінгу. В дослідженнях найкраще себе показав метод на основі контурів, оскільки менше піданий шумам.

Для визначення точності детектування очей в реальних умовах, було проведено дослідження. Було зроблено 20 кадрів при різному положенню обличчя. Дослідження проводилось для оцінки впливу методів препроцесінгу на основі градієнтів та контурів із подальшим нейромережовим детектуванням на точність детектування зінниць.

Таблиця 4.4 – Результати точності детектування очей різними відеореєстраторами (тестування на реальних вхідних даних)

Метод	Кількість вірних детектувань на камеру ноутбука	У відсотковому співвідношенні	Кількість вірних детектувань на зовнішню камеру	У відсотковому співвідношенні
Точність детектування очей без нейронної мережі	12 / 20	60%	16 / 20	80%
Точність детектування очей з нейронною мережею	14 / 20	70%	19 / 20	95%

В результаті проведеного дослідження для визначення точності детектування очей в реальних умовах, зовнішня камера показала на 15% кращі результати.



Рисунок 4.4 – Порівняння точності детектування очей різними відеореєстраторами

Це свідчить про ефективність використання нейронних мереж у задачах детектування очей, особливо при використанні якіснішої зовнішньої камери.

4.2 Вплив методів постпроцесінгу на точність визначення стану зіниці

Під постпроцесінгом в межах запропонованої системи розуміємо визначення радіусу зіниці та напрямку погляду.

Для визначення точності детектування розміру зіниці було проведено дослідження з використанням маркованого датасету та методу на основі градієнту та контурів.

Таблиця 4.5 – Точність детектування зміни розміру зіниці з шумом на прикладі підготовленого градуйованого за ступенем шуму датасету на основі НМ методу детектування зіниці

Метод препроцесінгу	10% шуму	30% шуму	50% шуму	80% шуму
На основі градієнтів	93%	55%	45%	10%
На основі контурів	98%	74%	55%	27%

Таким чином, препроцесінг впливає на точність детектування, а метод на основі контурів є більш стійким шуму.

На рисунку 4.1 показано приклад знаходження зіниці ока

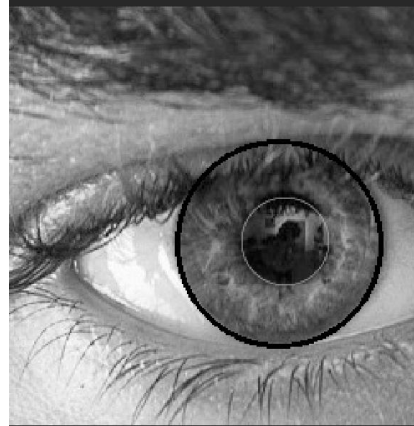


Рисунок 4.5 – Приклад знаходження зіниці та визначення діаметру

Для детектування стану зіниці, що метод, який використовує метод препроцесінгу контурів RANSAC дає точність вище на 11% чим метод, оснований на методі Хога.


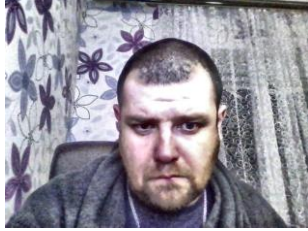
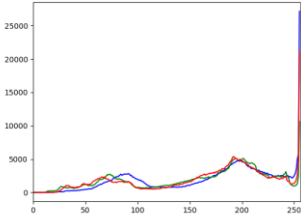
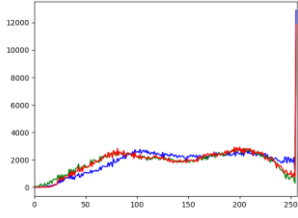


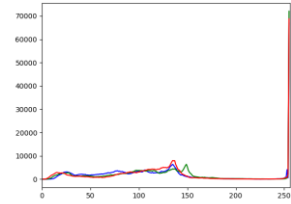
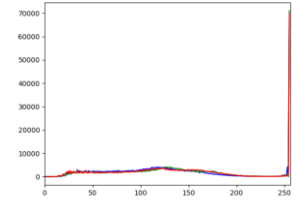
4.3 Демонстрація працездатності запропонованої системи у реальних умовах зйомки

Метою першого експерименту є визначення оптимальних порогів для попередньої обробки даних та чинників на якість вхідних даних. Об'єктом є показники рівня освітленості. Дані отримані за різних умов зйомки (тип освітлення – природній та штучний) камерою ноутбука.

В таблиці 4.6 представлені кадри, що були зроблені камерою ноутбука та етапи попередньої обробки кадру

Як видно з результатів дослідів, що освітлення є досить значимим фактором для отримання якісних результатів. Важливим показником є LAB Brightness (L) так як він відображає рівень освітлення сцени.

Таблиця 4.6 – Процес вирівнювання яскравості при реальних умовах зйомки

Назва	До вирівнювання яскравості	Після вирівнювання яскравості
Вхідний кадр		
Характеристика освітленості сцени	Grayscale Brightness: 166., HSV Brightness (V)': 178.78, LAB Brightness (L)': 171.02	Grayscale Brightness': 145.17, HSV Brightness (V)': 157.46, LAB Brightness (L)': 150.47
Гістограма до обробки		
Рівень шуму за параметром PSNR	39.56	35.15
Вхідний кадр		
Характеристика освітленості сцени		
Гістограма до обробки		
Характеристика освітленості сцени	Grayscale Brightness: 124.38, HSV Brightness (V)': 130.41, LAB Brightness (L): 130.28	Grayscale Brightness': 134.65, HSV Brightness (V)': 140.82, LAB Brightness (L): 140.82
Рівень шуму за параметром PSNR	47.92	46.22

Метою другого експерименту показати працездатність запропонованої системи на реальних даних, отриманих з камери ноутбука при різних кутах нахилу голови (таблиця 4.7). Вхідними даними є попередньо оброблені дані.

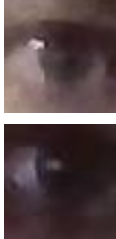
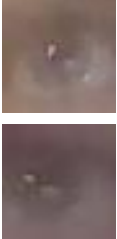

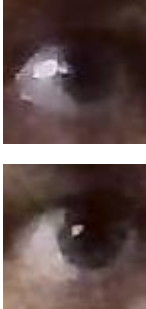
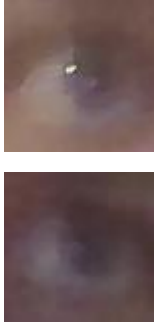
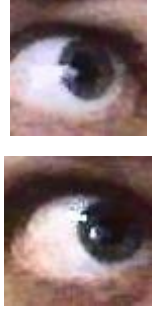
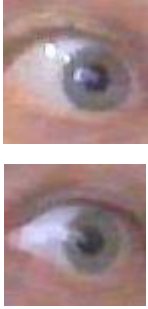
Таблиця 4.7 – Дослідження максимального кута при якому детектується обличчя

Вхідне зображення	Оцінка при різних кутах нахилу голови
	Прямий погляд. Вимкнений препроцесінг
	Погляд з боку. Вимкнений препроцесінг
	Коли не все обличчя. Увімкнений препроцесінг.

Підхід поетапним (без використання нейронної мережі) детектуванням містить недолік, при не повному попаданні в сцену унеможлиблює детектування очей та зіниць. Як покращити.

Третій експеримент проведено для вибору апаратного забезпечення для реєстрації зображення очей, які впливають на якість детектування зіниць.

Таблиця 4.8 – Апаратне забезпечення для реєстрації зображення очей

Параметри, при яких виділялася область ока	Вебкамера ноутбука		Камера з покращеними показниками	
	З вирівнюванням яскравості	Без вирівнювання яскравості	З вирівнюванням яскравості	Без вирівнювання яскравості
Приклади зображень				
Розміри зони обробки	40x40 px	40x40 px	>70x70px	>70x70px
Розмір приклад зображень після встановлення порогу				

Емпіричним шляхом було визначено, що розмір області для подальшої обробки має бути не менше 200x200 пікселів. Розмір області може залежати від якості камери та відстані обличчя до камери.

Висновок: Треба встановити поріг для зони до зони обробки, в разі меншої зони виводити повідомлення «Змініть камеру або наблизьтесь до камери».

Дослідження 4. Відстежування напрямку погляду та розміру зіниці в реальному часі. Для проведення дослідження було розроблено систему (рисунок 4.3).

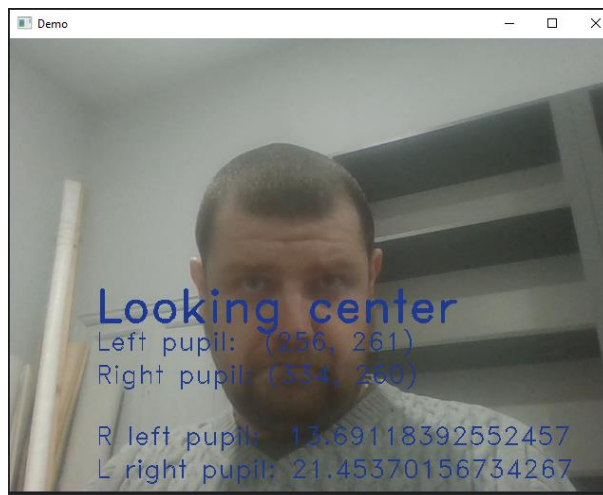


Рисунок 4.6 – Ілюстрація роботи системи

При використанні веб-камери ноутбука багато пропусків. Після проведеного аналізу було виявлено, тому що регіон очей не чіткий. Синею лінією – позначено зміну діаметру зіниці лівого ока, помпранчивою – правого.

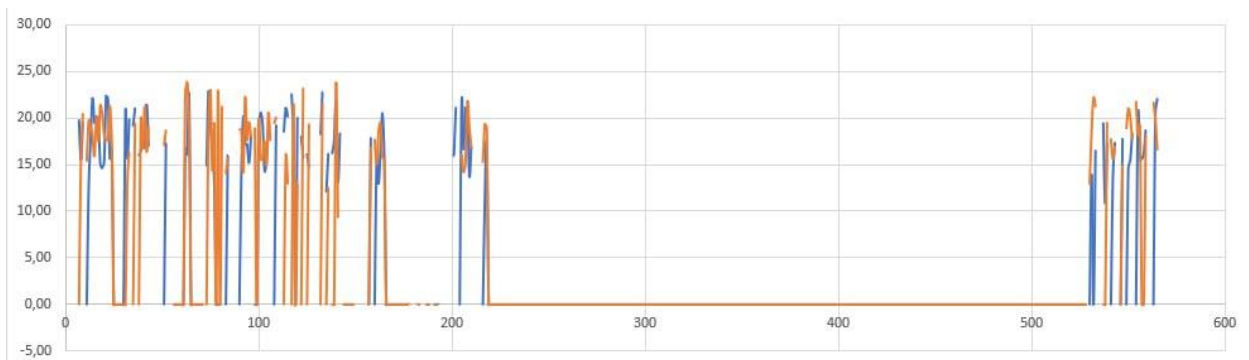


Рисунок 4.7 – Зміна розміру зіниці в реальному часі

В ході дослідження було виявлено ряд факторів що напряду впливають на детектування: якість камери, метод виділення ознак. В майбутньому планується розробка медичної експертної системи.

ВИСНОВКИ

Було проведено огляд основних психоемоційних станів та визначено основні порушення які можна діагностувати при аналізі руху очей пацієнта з використання методів комп'ютерного зору.

Була розроблена модель системи для детектування психо-емоційного стану на основі неронної мережі. Проведено аналіз існуючих методів детектування області ока та виділення зіниці. Проведено порівняння градієнтних та контурних методів.

Розроблено підсистему комп'ютерного зору складається з наступних етапів: попередня обробка даних, виділення обличчя, покращення виділеної області, виділення ROI, пост процесинг

Розроблено методологію проведення дослідження. Послідовність проведення досліджень: виділення обличчя з сцени, очей, зіниць, напрям погляду, розміру зіниці. Було підготовано маркований датасет з різною градацією зашумленості (10%, 20%, 50%, 80%) та проведені дослідження

Препроцесінг на основі контурів має більшу ефективність, тому що менш чутливий до шумів. Результати дослідження, які були проведені на датасеті BioID Face Database показали, що використання нейромережових методів детектування очей дають перевагу у 6 %. В результаті проведеного дослідження для визначення точності детектування очей в реальних умовах, зовнішня камера показала на 15% кращі результати. В дослідженнях з визначення стану зіниці показано, що препроцесінгу на основі методу контурів RANSAC дає точність вище на 11% чим метод, оснований на методі Хога. Тому що, RANSAC більш шумостійкий.

Було розроблено модуль нейромережевого детектування обличчя, який працює на основі методу Histogram of Oriented Gradients (HOG) у поєднанні із Support Vector Machine (SVM).

В майбутньому планується розробка медичної експертної системи.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Лисечко М. В. Вплив стресу на організм [Електронний ресурс] / Маргарита Валеріївна Лисечко – Режим доступу до ресурсу: <https://kdpu.edu.ua/sotsialno-hromadska-diialnist/psykholohichna-sluzhba/rekomendatsii/18060-vplyv-stresu-na-orhanizm.html>.
2. Оніщенко Н.В. Екстрена психологічна допомога постраждалим в умовах надзвичайної ситуації: теоретичні та прикладні аспекти: монографія. Х.: Право, 2014. 584 с
3. Посттравматичний стресовий розлад (ПТСР) і супутні з ним депресивні симптоми в осіб з регіонів, що значно постраждали від агресії РФ [Електронний ресурс]. – 2024. – Режим доступу до ресурсу: <https://uk.e-medjournal.com/index.php/psp/article/view/517>.
4. Ellis CJ (November 1981). The pupillary light reflex in normal subjects (PDF). *The British Journal of Ophthalmology*. 65 (11): 754—9
5. I. Loewenfeld, "The Pupl: Anatomy, Physiology, and Clinical Applications," Butterworth-Heinemann, no. 2, 1999.
6. C. A. Hall and R. P. Chilcott, "Eyeing up the Future of the Pupillary Light Reflex in Neurodiagnostics," *Diagnostics*, vol. 8, no. 1, 2018.
7. X. Fan, L. Hearne, B. Lei, J. Miles, N. Takahashi and G. Yao, "Weak gender effects on transient pupillary light reflex," *Autonomic Neuroscience: Basic Clinical*, no. 1-2, pp. 9-13, 2009.
8. O. Bergamin, A. Schoetzau, K. Sugimoto и M. Zulauf, «The influence of iris color on the pupillary light reflex,» *Graefe's Arch Clin Exp Ophtalmol*, pp. 567-570, 1998.
9. V. Raudonis, A. Paulauskaite, A. Vilkauskas, G. Dervinis and G. Kersulyte, "Evaluation of human emotion from eye motions," *International Journal of Advanced Computer Science and Applications*, vol. 4, no. 8, pp. 79-84, 2013.

10. Anuva Goyal Viola Jones Algorithm: A Classical Approach for Face Detection DOI:10.13140/RG.2.2.34263.07849

11. RANSAC URL: https://en.wikipedia.org/wiki/Random_sample_consensus

12. Zafari H, Kosowan L, Zulkernine F, Signer A. Diagnosing post-traumatic stress disorder using electronic medical record data. *Health Informatics J.* 2021;27(4):14604582211053259. doi:10.1177/14604582211053259

13. Schultebrucks K, Yadav V, Shalev AY, Bonanno GA, Galatzer-Levy IR. Deep learning-based classification of posttraumatic stress disorder and depression following trauma utilizing visual and auditory markers of arousal and mood. *Psychological Medicine.* 2022;52(5):957-967. doi:10.1017/S0033291720002718

14. M. Kalanadhabhatta, S. Roy, T. Grant, A. Salekin, T. Rahman and D. Bergen-Cico, "Detecting PTSD Using Neural and Physiological Signals: Recommendations from a Pilot Study," *2023 11th International Conference on Affective Computing and Intelligent Interaction (ACII)*, Cambridge, MA, USA, 2023, pp. 1-8, doi: 10.1109/ACII59096.2023.10388200

15. Long Nguyen-Phuoc, Renald Gaboriau, Dimitri Delacroix, Laurent Navarro, PTSD–MDNN : Fusion tardive de réseaux de neurones profonds multimodaux pour la détection du trouble de stress post–traumatique. doi: 10.48550/arXiv.2403.10565

16. Moctar Abdoul Latif Sawadogo, Furkan Pala, Gurkirat Singh, Imen Selmi, Pauline Puteaux, Alice Othmani. PTSD in the Wild: A Video Database for Studying Post–Traumatic Stress Disorder Recognition in Unconstrained Environments. doi: 10.48550/arXiv.2209.14085

17. Науменко В. А. та ін. Особливості зорових функцій в осіб, які проживають у зоні АТО, як маркер посттравматичного стресового розладу //Український медичний журнал. – 2019. – №. 1 (2). – С. 43–44

18. Я. А. Ляпін, О.Ю. Барковська Модель моніторингу напрямку погляду в системі комп'ютерного контролю знань – Режим доступу до

ресурсу: <https://repository.kpi.kharkov.ua/server/api/core/bitstreams/158bfa8b-1483-4b53-8d0b-d6b5c56ee068/content>

19. W. -H. Cao, J. -P. Xu and Z. -T. Liu, "Speaker-independent speech emotion recognition based on random forest feature selection algorithm," 2017 36th Chinese Control Conference (CCC), Dalian, China, 2017, pp. 10995-10998, doi: 10.23919/ChiCC.2017.8029112.

20. Ross Girshick, Jeff Donahue, Trevor Darrell, Jitendra Malik "Rich feature hierarchies for accurate object detection and semantic segmentation." doi:arXiv:1311.2524

21. Girshick, Ross. "Fast r-cnn." Proceedings of the IEEE international conference on computer vision. 2015. doi: arXiv:1504.08083

22. Ren, Shaoqing, et al. "Faster r-cnn: Towards real-time object detection with region proposal networks." Advances in neural information processing systems. doi: arXiv:1506.01497

23. P. Viola and M. J. Jones, «Rapid Object Detection using a Boosted Cascade of Simple Features», proceedings IEEE Conf. on Computer Vision and Pattern Recognition (CVPR2001), 2001 (Kauai, HI, 2001), pp. 511–518

24. shape_predictor_68_face_landmarks.dat | Kaggle
URL:<https://www.kaggle.com/datasets/sergiovirahonda/shape-predictor-68-face-landmarksdat>

25. Забрoда І.С., Барковська О.Ю. Маркери птср для інтелектуального аналізу // Проблеми інформатизації : XII міжнародна науково-технічна конференція. - 21-22 листопада 2024. -с.74. doi: <https://doi.org/10.32620/PI.24.t2>

26. Barkovska, O., Oliinyk, D., Sorokin, A., Zabroda, I., & Sedlaček, P. (2024). A system for monitoring the progress of rehabilitation of patients with musculoskeletal disorder. *Advanced Information Systems*, 8(3), 13-24. <https://doi.org/10.20998/2522-9052.2024.3.02>

27. Забрoда І. С. Система визначення ризиків розвитку ПТСР під час дистанційного навчання з використанням згорткових нейронних мереж / І. С.

Заброда ; наук. керівник к. т. н., доц. О. Ю. Барковська // *Радіоелектроніка та молодь у XXI столітті* : матеріали 28-го Міжнар. молодіж. форуму, 16–18 квітня 2024 р. – Харків : ХНУРЕ, 2024. – Т. 5. – С. 21–22. – DOI : <https://doi.org/10.30837/IYF.PCEIP.2024.021>.

28. Abnormal Behavior Recognition using CNN-LSTM with Attention Mechanism. URL: <https://ieeexplore.ieee.org/abstract/document/8974824>