

Вінницький національний технічний університет
Факультет інформаційних технологій та комп'ютерної інженерії
Кафедра обчислювальної техніки

Бакалаврська кваліфікаційна робота на тему
Комп'ютерна система безпеки водія з використанням нейронної мережі
08-54.БКР.034.00.000 ПЗ

Виконав: студент 4 курсу, групи 2СП—216
спеціальності

123 Комп'ютерна інженерія,
(шифр і назва напрямку підготовки, спеціальності)
освітня програма
«Системне програмування»

Огірчук О.С.
(прізвище та ініціали)

Керівник: к.т.н., доц. каф. ОТ

Городецька О.С.
(прізвище та ініціали)

Рецензент: доктор філософії, доц. каф. МБІС

Салієва О.В.
(прізвище та ініціали)

ДОПУЩЕНО

Завідувач кафедри ОТ

д.т.н., проф. Азаров О. Д

“16” 06 2025 р.

Вінниця ВНТУ – 2025 рік

Вінницький національний технічний університет
Факультет інформаційних технологій та комп'ютерної інженерії
Кафедра обчислювальної техніки
Освітньо—кваліфікаційний рівень бакалавр
Галузь знань – 12 – Інформаційні технології
Спеціальність – 123 – Комп'ютерна інженерія
Освітньо—професійна програма – Системне програмування

ЗАТВЕРДЖУЮ

Завідувач кафедри ОТ
д.т.н., проф. Азаров О. Д

«21» березня 2025 р.

**ЗАВДАННЯ
НА БАКАЛАВРСЬКУ КВАЛІФІКАЦІЙНУ РОБОТУ СТУДЕНТУ**

Огірчуку Олегу Сергійовичі

- 1 Тема роботи: «Комп'ютерна система безпеки водія з використанням нейронної мережі», керівник роботи Городецька О.С. затверджені наказом ВНТУ від «20» березня 2025 року №97.
- 2 Термін подання студентом роботи: 13.05.2025
3. Вихідні дані до роботи: документація MediaPipe Face Mesh, офіційна документація TensorFlow, наукові публікації з методів аналізу стану очей, навчальні набори даних OpenEDS та 300VW, технічна документація до бібліотеки OpenCV.
4. Зміст розрахунково—пояснювальної записки: вступ, сучасний стан та актуальність нейронних мереж, вибір методів та архітектури нейромережевої системи, розробка системи безпеки водія, перевірка працездатності, висновки, перелік використаних джерел, додатки.

5. Перелік графічного матеріалу: структурна схема.
 Перелік ілюстративного матеріалу : блок—схема алгоритму роботи.
 6 Консультанти розділів роботи наведені у таблиці 1

Таблиця 1 — Консультанти розділів роботи

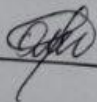
Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		Завдання видав	завдання прийняв
1—3	Городецька О.С., к.т.н., доц. каф. ОТ	21.03.2025	13.05.2025
Нормоконтроль	ас. каф. ОТ Швець С. І.	14.05.2025	30.05.2025

7. Дата видачі завдання «21» березня 2025 р. Календарний план наведений у таблиці 2.

Таблиця 2 — Календарний план

№ з/п	Назва етапів бакалаврської кваліфікаційної роботи	Строк виконання етапів	Примітка
1	Постановка задачі роботи	21.03.25	вик.
2	Пошук матеріалів та інформаційних джерел	21.03.25— 30.03.25	вик.
3	Аналіз існуючих рішень та визначення вимог до створюваної системи	21.03.25— 30.03.25	вик.
4	Обґрунтування структури та принципів роботи системи	31.03.25— 13.04.25	вик.
5	Розробка й тестування комп'ютерної системи безпеки водія	14.04.25— 27.04.25	вик.
7	Оформлення пояснювальної записки та ілюстративного матеріалу	28.04.25— 11.05.25	вик.
8	Перевірка якості виконання бакалаврської кваліфікаційної роботи та усунення недоліків	13.05.25	вик.

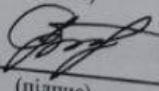
Студент



Огірчук О.С.

Керівник роботи

(підпис)



(підпис)

Городецька О.С.

АНОТАЦІЯ

УДК 621.3

Огірчук О.С. Комп'ютерна система безпеки водія з використанням нейронної мережі. Бакалаврська кваліфікаційна робота зі спеціальності 123 «Комп'ютерна інженерія», освітня програма «Системне програмування». Вінниця: ВНТУ, 2025. 70 с. На укр. мові. Бібліогр.: 21 назв; рис.: 19.

У бакалаврській роботі розроблено комп'ютерну систему моніторингу стану водія на основі методів комп'ютерного зору та глибокого навчання. Запропоноване рішення використовує архітектуру MediaPipe Face Mesh у поєднанні з алгоритмом аналізу стану очей для виявлення ознак втоми та розсіяної уваги у реальному часі. Проведено порівняльний аналіз сучасних підходів до детекції обличчя та очей, що дозволило обґрунтувати вибір оптимальної нейромережевої моделі. Програмна реалізація виконана з використанням мови Python. Для навчання моделі створено спеціалізований набір даних, що включає зображення водіїв у різних умовах освітлення та ракурсах. Проведені тести підтвердили ефективність розробленого рішення як у контрольованих умовах.

Ключові слова: комп'ютерний зір, система безпеки водія, нейронні мережі, детекція втоми, аналіз стану очей, реальний час, машинне навчання.

ABSTRACT

UDC 621.3

Ogirchuk O.S. Computer driver safety system using a neural network. Bachelor's qualification work in specialty 123 "Computer Engineering", educational program "System Programming". Vinnytsia: VNTU, 2025. 70 p. In Ukrainian. Bibliography: 21 titles; Fig.: 19.

In the bachelor's work, a computer system for monitoring the driver's condition was developed based on computer vision and deep learning methods. The proposed solution uses the MediaPipe Face Mesh architecture in combination with an eye state analysis algorithm to detect signs of fatigue and distracted attention in real time. A comparative analysis of modern approaches to face and eye detection was conducted, which allowed us to justify the choice of the optimal neural network model. The software implementation was performed using the Python language. A specialized dataset was created to train the model, which includes images of drivers in different lighting conditions and angles. The tests conducted confirmed the effectiveness of the developed solution both in controlled conditions.

Keywords: computer vision, driver safety system, neural networks, fatigue detection, eye condition analysis, real—time, machine learning.

ЗМІСТ

ВСТУП	3
1 СУЧАСНИЙ СТАН ТА АКТУАЛЬНІСТЬ ВИКОРИСТАННЯ НЕЙРОННИХ МЕРЕЖ	5
1.1. Роль глибокого навчання у системах комп'ютерного зору	5
1.2 Сучасні архітектури нейронних мереж у комп'ютерному зорі	10
1.3 Актуальність та перспективи застосування нейронних мереж.....	13
2 ВИБІР МЕТОДІВ ТА АРХІТЕКТУРИ НЕЙРОМЕРЕЖЕВОЇ СИСТЕМИ	20
2.1 Порівняльний аналіз архітектур нейронних мереж для детекції обличчя та очей.....	20
2.2 Обґрунтування вибору архітектури нейромережевої моделі.....	27
2.3 Процес навчання моделі та підготовка даних.....	28
3 РОЗРОБКА КОМП'ЮТЕРНОЇ СИСТЕМИ БЕЗПЕКИ ВОДІВ	34
3.1. Структура системи та її компоненти	34
3.2. Принцип роботи системи	38
3.3 Алгоритм розпізнавання стану очей і напряму погляду із застосуванням нейронних мереж	40
3.4. Вибір інструментів розробки.....	42
4 ПЕРЕВІРКА ПРАЦЕЗДАТНОСТІ СИСТЕМИ	46
4.1. Методи перевірки працездатності системи.....	46
4.2. Розгляд контрольного прикладу.....	48
ВИСНОВКИ	54
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ	55
ДОДАТОК А Технічне завдання	58
ДОДАТОК Б Протокол перевірки кваліфікаційної роботи.....	62
ДОДАТОК В Графічна частина.....	63
ДОДАТОК Г Ілюстративна частина	65
ДОДАТОК Д Лістинг програмного забезпечення	67

ВСТУП

У сучасних умовах стрімкого розвитку транспортної інфраструктури та зростання інтенсивності дорожнього руху особливої актуальності набувають питання підвищення безпеки дорожнього руху. Одним з ключових напрямків у цій галузі є розробка інтелектуальних систем моніторингу стану водія, здатних виявляти ознаки втоми, розсіяної уваги та інших небезпечних станів у реальному часі. Сучасні досягнення в галузі комп'ютерного зору та глибокого навчання відкривають нові можливості для створення ефективних рішень у цій сфері.

Технології на основі нейронних мереж дозволяють створювати складні моделі, здатні аналізувати міміку обличчя, рух очей та інші фізіологічні параметри з високою точністю. Особливу цінність становлять системи, що поєднують переваги сучасних алгоритмів комп'ютерного зору з можливостями реального аналізу відеопотоку, що дозволяє оперативно реагувати на потенційно небезпечні ситуації.

Актуальність теми зумовлена зростанням кількості дорожньо-транспортних пригод, що виникають через втому, сонливість та неуважність водіїв, особливо під час тривалих поїздок або роботи в умовах монотонного руху. Ці чинники щороку призводять до значних втрат як у матеріальному вимірі, так і в людських життях. У зв'язку з цим постає потреба у впровадженні інноваційних рішень, спрямованих на підвищення безпеки дорожнього руху. Попри наявність комерційних систем моніторингу стану водія, більшість з них є вартісними, складними у впровадженні та доступними переважно у преміальному сегменті автомобілів. Це створює розрив між потребою у безпечному керуванні транспортом та реальною можливістю широкого впровадження таких технологій. У свою чергу, стрімке зростання попиту на системи безпеки в масовому автотранспорті актуалізує розробку доступних та гнучких рішень, які могли б працювати у реальному часі з використанням звичайного відеопотоку без необхідності в дорогому обладнанні.

Об'єктом дослідження даної роботи є процес моніторингу стану водія за допомогою методів комп'ютерного зору.

Предметом дослідження є методи та алгоритми виявлення стану очей та напряму погляду водія з метою визначення ознак втоми або неувважності.

Метою роботи є підвищення точності виявлення ознак втоми та зниження уваги водія шляхом удосконалення алгоритмів розпізнавання стану очей і напряму погляду із застосуванням нейронних мереж.

Задачі дослідження бакалаврської роботи :

- проаналізувати сучасні архітектури нейронних мереж у комп'ютерному зорі;
- обґрунтувати вибір архітектури нейромережі для комп'ютерної системи безпеки водія;
- розробити алгоритм розпізнавання стану очей та напряму погляду водія;
- розробити комп'ютерну систему безпеки водія;
- провести перевірку працездатності комп'ютерної системи.

Апробація результатів бакалаврської роботи: зроблено доповідь на LIV Всеукраїнської науково—технічної конференції факультету інформаційних технологій та комп'ютерної інженерії [1].

Публікації за темою роботи: Застосування глибоких нейронних мереж для розпізнавання обличчя водія у відеопотоці для забезпечення безпеки водія / О. С. Огірчук, О. С. Городецька, С.В. Городецька // Матеріали LIV Всеукраїнська науково—технічна конференція факультету інформаційних технологій та комп'ютерної інженерії. Вінниця 2025 р. [Електронний ресурс]. Режим доступу: <https://conferences.vntu.edu.ua/index.php/all-fitki/all-fitki-2025/paper/view/23939/19788>

1 СУЧАСНИЙ СТАН ТА АКТУАЛЬНІСТЬ ВИКОРИСТАННЯ НЕЙРОННИХ МЕРЕЖ

1.1. Роль глибокого навчання у системах комп'ютерного зору

Сучасні системи комп'ютерного зору значною мірою базуються на методах штучного інтелекту, серед яких ключове місце займають нейронні мережі – обчислювальні моделі, натхненні будовою біологічних нейронних систем. В основі цих моделей лежить принцип імітації роботи нервових клітин: штучний нейрон отримує вхідні сигнали, обробляє їх за допомогою навчених параметрів і формує вихідний сигнал, який передається далі по мережі (рисунок 1.1). Простою аналогією може слугувати робота людського мозку, де нейрони передають інформацію через синапси, але в штучних мережах цей процес формалізований у вигляді математичних операцій [2].

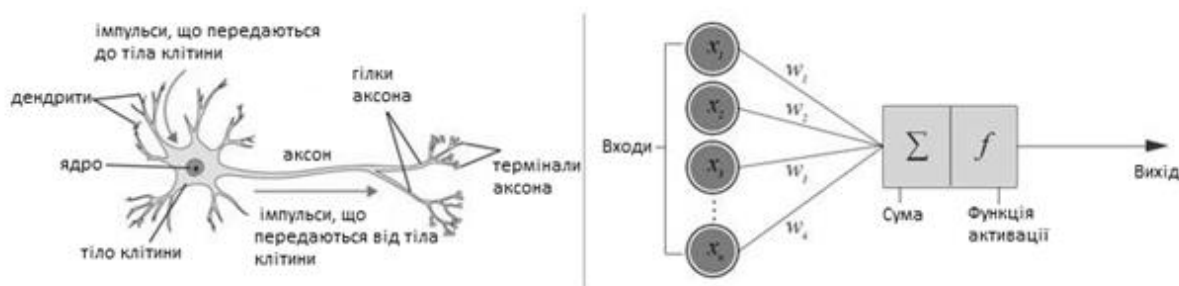


Рисунок 1.1 — Схема штучного нейрона та його порівняння з біологічним

Нейронні мережі складаються з взаємопов'язаних шарів, кожен з яких виконує специфічні перетворення даних (рисунок 1.2). Вхідний шар отримує зображення або інші дані, приховані шари виділяють ознаки різного рівня складності, а вихідний шар формує кінцевий результат (наприклад, клас об'єкта або координати його розташування). Критичним елементом є функція активації, яка забезпечує не лінійність мережі, дозволяючи їй навчатися складним залежностям.

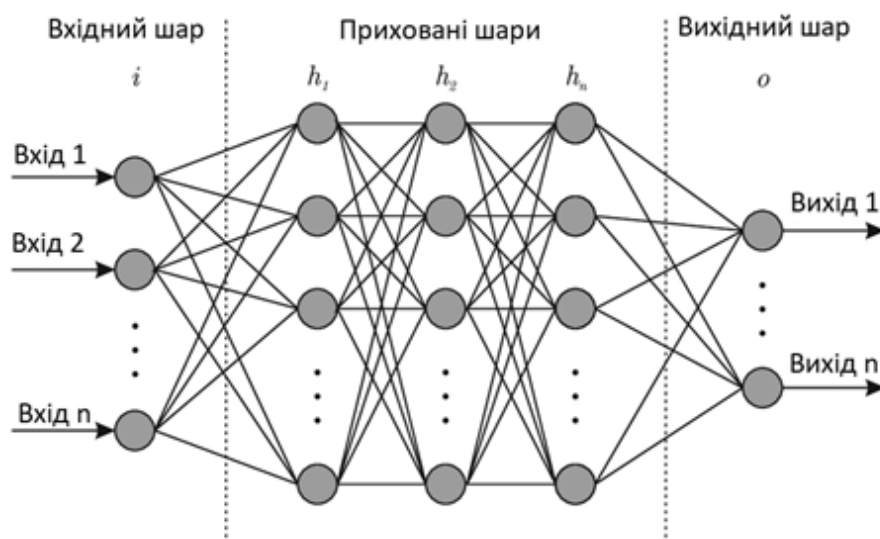


Рисунок 1.2— Архітектура простої нейронної мережі з входом, прихованими шарами та виходом

Процес навчання нейронної мережі ґрунтується на алгоритмі зворотного поширення помилки, який автоматично коригує параметри моделі, мінімізуючи різницю між передбаченнями та реальними даними. Цей механізм дозволяє мережі адаптуватися до вхідних даних без явного програмування правил, що є ключовою перевагою перед традиційними алгоритмами обробки зображень.

Глибоке навчання є розвитком класичних нейронних мереж і передбачає використання архітектур із великою кількістю прихованих шарів. Як видно на рисунку 1.3, такі моделі здатні автоматично будувати складні ієрархії ознак, починаючи від простих градієнтів і кутів до цілих об'єктів та їх взаємозв'язків. Саме ця властивість зробила глибокі нейронні мережі основним інструментом у комп'ютерному зорі, забезпечуючи прорив у таких завданнях, як розпізнавання об'єктів, класифікація зображень та аналіз обличч.

Перехід від традиційних нейронних мереж до глибокого навчання став можливим завдяки трьом ключовим факторам: зростанню обчислювальних потужностей, появі великих датасетів та розвитку нових архітектурних рішень. Глибокі мережі відрізняються не лише кількістю шарів, а й здатністю до автоматичного виділення складних ознак без явного програмування [3].

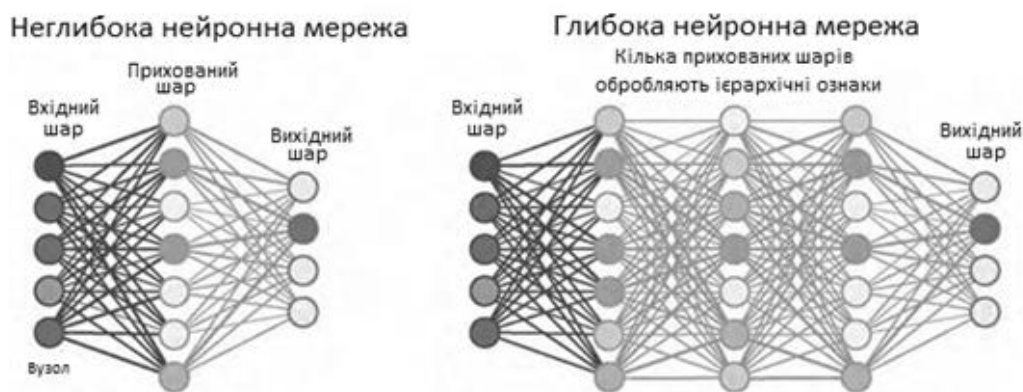


Рисунок 1.3— Порівняння глибини традиційних та глибоких нейронних мереж

У комп'ютерному зорі найбільш поширеними стали згорткові нейронні мережі (CNN), які використовують спеціальні операції згортки для аналізу локальних ділянок зображення. На рисунку 1.4 показано, як фільтри нижчих шарів виявляють прості структури, тоді як вищі шари комбінують їх у складніші об'єкти. Ця ієрархічна організація нагадує принципи роботи зорової кори людини.

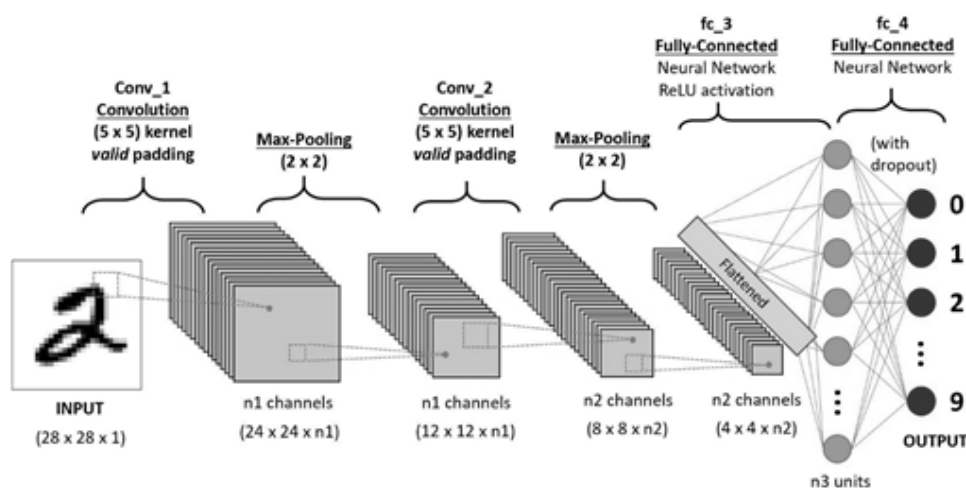


Рисунок 1.4 — Архітектура згорткової нейронної мережі

Глибокі моделі демонструють вищу точність порівняно з традиційними методами завдяки:

— автоматичному навчанню ознак (відсутність необхідності ручного проектування);

- масштабованості (можливість обробляти зображення високої роздільної здатності);
- стабільності до змін умов (освітлення, ракурс, часткові перекриття).

Системи на основі глибокого навчання вже досягають людського рівня точності у розпізнаванні облич, відкриваючи широкі можливості для біометрії, відеоспостереження та інтерактивних технологій. У комп'ютерному зорі найчастіше використовуються згорткові нейронні мережі, які завдяки своїй ієрархічній структурі ефективно аналізують зображення. Кожен шар такої мережі виявляє ознаки різної складності — від простих контурів до цілісних об'єктів. Локальні зв'язки між нейронами імітують зорове сприйняття людини, що дозволяє моделі формувати цілісне уявлення про візуальну сцену та досягати високої точності в завданнях аналізу зображень.

Архітектури згорткових нейронних мереж за останнє десятиліття значно еволюціонували. Починаючи з AlexNet, який уперше продемонстрував ефективність глибокого навчання в обробці зображень, розвиток тривав через VGGNet із послідовними шарами, до ResNet із залишковими зв'язками, що вирішили проблему зникаючих градієнтів. Подальші вдосконалення, зокрема EfficientNet, дозволили досягти оптимального балансу між точністю та обчислювальною ефективністю. Залишкові зв'язки стали ключовим елементом, що забезпечив можливість створення глибших і точніших моделей для складних завдань комп'ютерного зору.

Навчання глибоких нейронних мереж є складним багатоетапним процесом, який вимагає значних обчислювальних ресурсів та великих обсягів розмічених даних. Як ілюструє рисунок 1.5 першим критично важливим етапом є попередня обробка зображень, яка включає нормалізацію пікселів, зміну розмірів, а також аугментацію даних для збільшення різноманітності навчальної вибірки [4]. Після цього відбувається багаторазова ітеративна передача даних через мережу, під час якої кожен шар послідовно виділяє та перетворює ключові ознаки. Основним механізмом навчання є процес корекції параметрів мережі на основі аналізу

помилки прогнозування. Цей механізм реалізується через алгоритм зворотного поширення помилки, який дозволяє точно коригувати вагові коефіцієнти кожного шару. Особливу увагу при цьому приділяють методам регуляризації, серед яких найбільш поширеним є dropout. Цей метод, що передбачає випадкове відключення частини нейронів під час навчання, ефективно запобігає явищу перенавчання та покращує узагальнюючу здатність моделі. Дослідження показують, що правильне налаштування параметрів регуляризації може підвищити точність роботи моделі, що робить цей аспект ключовим у процесі навчання.

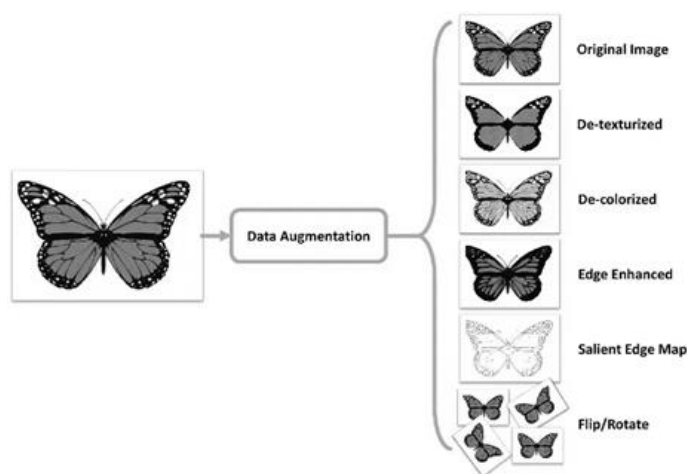


Рисунок 1.5 — Приклад аугментації даних

Останнім часом у галузі комп'ютерного зору спостерігається якісна трансформація архітектурних підходів. Традиційні згорткові нейронні мережі поступово доповнюються та модифікуються за рахунок інтеграції з механізмами уваги (attention), характерними для трансформерів. Особливу ефективність демонструють гібридні моделі на кшталт Vision Transformer (ViT), які досягають вражаючих результатів у завданнях, що потребують аналізу глобальних залежностей між різними ділянками зображення [5]. Такі технології відкривають нові можливості в медичній діагностиці, де критично важливим є виявлення зв'язків між віддаленими патологічними змінами на томографічних знімках.

Системи комп'ютерного зору на базі глибокого навчання активно застосовуються в медицині, транспорті, промисловості та інфраструктурі

розумних міст. Вони автоматизують діагностику, забезпечують контроль якості, аналізують дорожню ситуацію й моніторять міський простір. Основними обмеженнями залишаються потреба в розмічених даних, висока обчислювальна складність та складність інтерпретації рішень, що критично для медицини та безпеки.

Серйозним викликом є вразливість систем до адверсарних атак, коли навмисні, але майже непомітні для людини зміни у вхідних даних можуть призвести до катастрофічних помилок класифікації [6]. Навіть незначні штучні артефакти можуть спричинити хибну ідентифікацію об'єктів, що особливо небезпечно для автономних систем і систем безпеки. Майбутнє галузі пов'язане з подоланням цих викликів завдяки самокерованому навчанню, що зменшує потребу в розмічених даних, розвитку пояснюваного ШІ для підвищення інтерпретованості моделей та створенню енергоефективних архітектур для периферійних пристроїв. Інтеграція комп'ютерного зору з мовними моделями відкриває нові можливості для створення мультимодальних систем, що аналізують зображення, текст і аудіо. Такі рішення формують основу інтелектуальних асистентів і систем доповненої реальності, здатних розпізнавати об'єкти та розуміти їх у контексті середовища, що розширює межі взаємодії між людиною і машиною.

1.2 Сучасні архітектури нейронних мереж у комп'ютерному зорі

Сучасний комп'ютерний зір пройшов значний шлях від простих алгоритмів обробки зображень до складних глибоких нейронних мереж, здатних вирішувати задачі людського рівня. Цей прогрес став можливим завдяки постійному вдосконаленню архітектур нейронних мереж, кожна з яких внесла унікальний внесок у розвиток галузі.

Класичні згорткові мережі (CNN) залишаються основою комп'ютерного зору. Архітектура VGGNet показала, що збільшення глибини мережі підвищує точність за умови правильного налаштування. Справжній прорив відбувся з

появою ResNet, яка завдяки залишковим зв'язкам дала змогу ефективно навчати дуже глибокі моделі, подолавши проблему зникаючих градієнтів.

Трансформери в комп'ютерному зорі з'явилися як альтернатива традиційним CNN. Vision Transformer (ViT), запропонований у 2020 році, застосував принципи мовних трансформерів до обробки зображень. Ключова особливість ViT — розбиття зображення на патчі та аналіз їх взаємозв'язків за допомогою механізму уваги (рисунк 1.6). Це дозволяє краще враховувати глобальний контекст, що особливо важливо для таких завдань, як аналіз медичних знімків або супутникових даних. Swin Transformer удосконалив цей підхід, ввівши ієрархічну структуру, що значно знизило обчислювальні витрати.

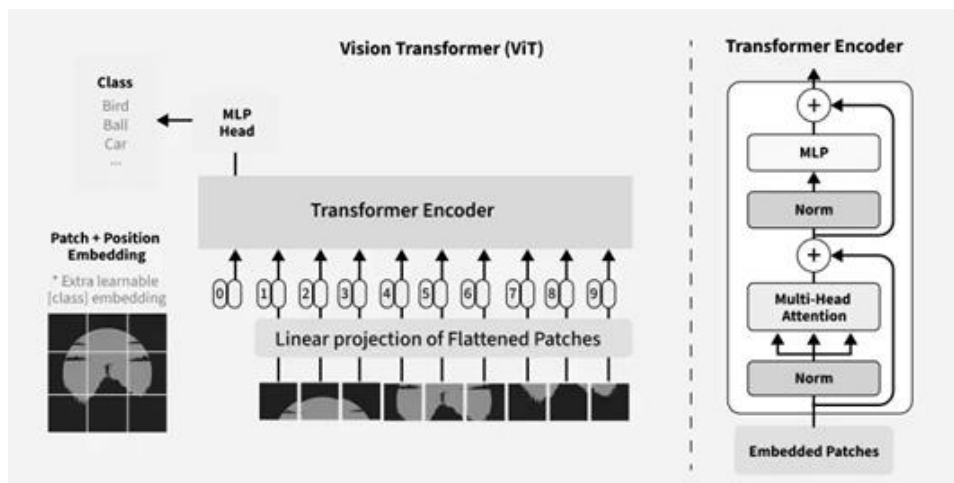


Рисунок 1.6 — Архітектура Vision Transformer (ViT)

Гібридні архітектури поєднують переваги обох підходів. Наприклад, Convolutional Vision Transformer (ConViT) інтегрує згорткові шари для виділення локальних ознак і механізми уваги для аналізу глобального контексту. Такі моделі особливо ефективні для завдань, де важливі як деталі, так і загальна структура зображення. Глибоке розуміння особливостей різних архітектур нейронних мереж дозволяє оптимально підбирати модель під конкретні завдання комп'ютерного зору. Детальний аналіз ключових архітектур показує, що кожен тип мереж має унікальні переваги та обмеження, які варто враховувати при виборі.

Основні відмінності між згортковими мережами (CNN) та трансформерами проявляються у підходах до обробки просторових відношень. Традиційні CNN, що використовують локальні згортки, демонструють особливу ефективність у виявленні низькорівневих ознак, таких як краї чи текстури [7]. На противагу цьому, трансформери, засновані на механізмах уваги, дозволяють аналізувати глобальні залежності у зображенні, що стає вирішальним для завдань із складним контекстом. Важливо відзначити, що CNN зазвичай демонструють стабільну роботу навіть на відносно малих наборах даних (від 1,000 зразків), тоді як трансформери вимагають значно більших обсягів навчальних даних (понад 100,000 зразків) для досягнення аналогічної стабільності.

Спеціалізовані архітектури для конкретних задач комп'ютерного зору демонструють вражаючу ефективність у вузьких галузях. У сфері детекції об'єктів модель YOLO (You Only Look Once) виділяється високою швидкістю обробки, що робить її ідеальним вибором для систем реального часу [8]. Інший підхід, представлений у Faster R-CNN, забезпечує підвищену точність за рахунок використання регіон—пропозалів. Для задач сегментації зображень архітектура U-Net стала стандартом у біомедичних дослідженнях, тоді як Mask R-CNN дозволяє одночасно виконувати детекцію та сегментацію об'єктів. Обчислювальна ефективність сучасних архітектур варіюється в широких межах. Наприклад, ResNet—50 містить приблизно 25 мільйонів параметрів, тоді як ViT-Base вимагає значно більших ресурсів з 86 мільйонами параметрів. На цьому тлі особливо виділяється EfficientNet-B0, який завдяки оптимізованій архітектурі досягає конкурентної точності при всього 5 мільйонах параметрів, що робить його привабливим вибором для впровадження на пристроях з обмеженими обчислювальними ресурсами.

Незважаючи на вражаючі досягнення, сучасні архітектури нейронних мереж для комп'ютерного зору стикаються з низкою фундаментальних обмежень, які визначають напрями подальшого розвитку галузі. Ці виклики можна умовно поділити на три категорії: обчислювальні, методологічні та практичні.

Обчислювальні обмеження становлять особливу проблему для сучасних глибоких моделей. Трансформери, такі як Vision Transformer (ViT), вимагають значно більших обчислювальних ресурсів порівняно з традиційними CNN. Наприклад, тренування базової версії ViT може вимагати до 2.5 тисяч GPU—днів на сучасному обладнанні. Це пов'язано з квадратичною складністю механізму уваги щодо розміру вхідного зображення. Проблема особливо гостра для систем реального часу, де необхідно обробляти відеопотоки високої роздільної здатності. Методологічні виклики включають проблему залежності від великих обсягів розмічених даних. Більшість сучасних архітектур досягають максимальної ефективності лише при наявності мільйонів розмічених прикладів. Так, для стабільного навчання ViT рекомендується використовувати набори даних обсягом понад 100 мільйонів зображень.

Проблема інтерпретованості ("чорного ящика") залишається одним з найбільш суттєвих обмежень [9]. Сучасні архітектури, особливо трансформери, демонструють високу точність, але їхні рішення часто важко пояснити. У критично важливих галузях, таких як медицина або автономні системи, це створює додаткові ризики. Дослідження показують, що навіть високоточні моделі можуть приймати рішення, керуючись артефактами зображення, а не реальними ознаками об'єкта.

Адверсарні атаки представляють особливу загрозу для систем комп'ютерного зору. Доведено, що майже непомітні для людського ока зміни у вхідному зображенні можуть призвести до катастрофічних помилок класифікації. Наприклад, додавання спеціально розробленого "шуму" може змусити систему помилково ідентифікувати дорожні знаки, що становить пряму загрозу для автономних транспортних засобів.

1.3 Актуальність та перспективи застосування нейронних мереж

Сучасні нейронні мережі трансформують традиційні підходи до аналізу візуальної інформації, знаходячи застосування у все більшому спектрі галузей. Особливо помітний прогрес спостерігається у сфері медицини, де комп'ютерний

зір на основі глибокого навчання став незамінним інструментом діагностики. Медицина та охорона здоров'я переживають справжню революцію завдяки впровадженню нейромережових технологій. Системи на основі архітектур U—Net та Vision Transformer демонструють вражаючі результати в аналізі медичних зображень. Наприклад, алгоритми виявлення пухлин молочної залози досягають точності 97,8% при аналізі мамографічних знімків, що перевищує показники навіть досвідчених рентгенологів. Особливу цінність становлять системи швидкої діагностики інсультів за даними КТ – сучасні моделі здатні виявляти ознаки гострого порушення мозкового кровообігу за 2—3 хвилини, коли кожна секунда має вирішальне значення для пацієнта. У хірургічній практиці комп'ютерний зір відіграє ключову роль у трьох основних напрямках. По—перше, системи доповненої реальності забезпечують точну навігацію під час складних операцій. По—друге, автоматизовані системи моніторингу дозволяють відстежувати життєві показники пацієнта в режимі реального часу. По—третє, спеціалізовані алгоритми аналізують стан хірургічного інструментарію, запобігаючи його передчасному зносу.

Автономні транспортні системи представляють собою іншу ключову сферу застосування нейронних мереж. Сучасні системи допомоги водієві (ADAS) базуються на складних каскадах нейронних мереж, які виконують кілька критично важливих функцій [10]. Вони забезпечують надійне виявлення пішоходів та інших учасників дорожнього руху, точно розпізнають дорожні знаки навіть у складних погодних умовах та аналізують поведінку водія для попередження потенційно небезпечних ситуацій. Яскравим прикладом є система Tesla Vision, яка демонструє точність 99,4% у завданнях класифікації дорожніх сценаріїв, що дозволяє уникнути до 85% потенційних аварій. Важливою особливістю таких систем є їхня здатність обробляти до 60 кадрів за секунду, забезпечуючи роботу в реальному часі.

Промисловість активно інтегрує технології комп'ютерного зору для вирішення широкого спектру завдань. На передових підприємствах системи на основі нейронних мереж забезпечують автоматичний контроль якості продукції,

оптимізують роботу виробничих ліній та здійснюють прогнозне технічне обслуговування обладнання.

Роздрібна торгівля змінюється під впливом комп'ютерного зору. Торгові мережі впроваджують системи аналізу відеопотоків для виявлення крадіжок і вивчення поведінки покупців (рисунок 1.7). Алгоритми детекції об'єктів дозволяють визначати популярні товари за частотою взаємодії з полицями. Деякі магазини вже тестують безкасові технології, де комп'ютерний зір автоматично фіксує вибрані товари, спрощуючи процес покупок.



Рисунок 1.7 — Теплова карта руху клієнтів по магазину

Сільське господарство знаходить все більше способів застосування нейронних мереж для підвищення ефективності виробництва. Точне землеробство, засноване на аналізі знімків з дронів, дозволяє фермерам виявляти проблемні ділянки полів на ранніх стадіях. Спеціалізовані алгоритми здатні діагностувати стан рослин, виявляючи ознаки хвороб або нестачі поживних речовин за зміною кольору листя. В автоматизованих теплицях системи комп'ютерного зору відстежують ріст рослин і оптимальний час для збору врожаю, що дозволяє досягати значного підвищення продуктивності.

Сучасні системи безпеки суттєво підвищили свою ефективність завдяки розвитку алгоритмів розпізнавання облич і аналізу поведінки. Технології перейшли від базового виявлення осіб до складного розпізнавання потенційно

небезпечних дій за рухами людини. Інтелектуальні системи відеоспостереження вже активно використовуються в аеропортах для виявлення залишених речей або підозрілої активності. Алгоритми показують надійність навіть при слабкому освітленні чи частково закритому обличчі. Сфера застосування комп'ютерного зору постійно розширюється — від аналізу супутникових знімків у космічній галузі до археології, мистецтва, освіти та науки. У науці й медицині ці технології дозволяють моделювати клінічні ситуації, автоматизувати обробку зображень та прискорювати аналіз експериментальних даних.

Транспортна логістика активно використовує комп'ютерний зір для автоматизації складських процесів. Сучасні системи можуть автоматично ідентифікувати вантажі, перевіряти цілісність упаковки та оптимізувати процес завантаження. У великих логістичних центрах це дозволяє скоротити час обробки вантажів, значно підвищуючи загальну ефективність роботи.

Глобальний ринок технологій комп'ютерного зору на основі нейронних мереж демонструє стрімке зростання, що безпосередньо впливає на світову економіку. Вплив на продуктивність підприємств проявляється через автоматизацію рутинних операцій, що раніше вимагали участі людини. На виробничих підприємствах системи комп'ютерного зору дозволяють скоротити витрати на контроль якості, при цьому підвищуючи точність виявлення дефектів. Яскравим прикладом є досвід компанії Siemens, де впровадження нейромережевих систем інспекції дозволило зменшити кількість технологічних браку на 28% лише за перший рік експлуатації.

Оптимізація ланцюгів поставок завдяки комп'ютерному зору стала ключовим фактором конкурентоспроможності для логістичних компаній. Автоматизовані системи сортування вантажів на базі алгоритмів YOLOv5 та Mask R-CNN дозволяють обробляти до 10 000 одиниць товару на годину, що 3-4 рази перевищує продуктивність традиційних методів. Трансформація ринку праці під впливом комп'ютерного зору має двоїстий характер. З одного боку, автоматизація візуального контролю призводить до скорочення традиційних посад інспекторів якості, з іншого боку, стрімко зростає попит на фахівців з

машинного навчання та обробки зображень. Проблеми цифрової нерівності загострилися з розповсюдженням технологій комп'ютерного зору. Великі технологічні корпорації, що володіють потужними обчислювальними ресурсами та якісними даними, отримують несправедливу перевагу перед невеликими компаніями та країнами, що розвиваються. Дослідження Оксфордського інституту Інтернету показують, що 78% патентів у сфері комп'ютерного зору належать компаніям з США та Китаю, що створює ризик технологічної залежності для інших країн.

Питання конфіденційності набуває особливої гостроти у світі, де системи розпізнавання облич впроваджуються у громадських просторах. У таких країнах як Китай або ОАЕ ці технології вже стали частиною повсякденного життя, використовуючись для ідентифікації осіб у метро, банках або навіть у школах. Європейський Союз у відповідь на це ввів жорсткі обмеження на використання технологій біометричного розпізнавання у громадських місцях, намагаючись знайти баланс між безпекою та приватним життям.

Інтеграція мультимодальних архітектур становить один з найперспективніших напрямів розвитку комп'ютерного зору. Сучасні дослідження демонструють, що поєднання візуальних даних з текстовою, аудіальною та сенсорною інформацією дозволяє досягти принципово нового рівня розуміння контексту [11]. Моделі типу CLIP (Contrastive Language-Image Pretraining) від OpenAI вже показують вражаючі результати у завданнях, що вимагають зв'язку між зображенням і текстом. Наприклад, вони здатні точно класифікувати зображення на основі складних текстових запитів або генерувати релевантні підписи до візуального контенту. Це відкриває нові можливості для створення інтелектуальних систем пошуку, автоматизованого контент—менеджменту та інклюзивних технологій для людей з обмеженими можливостями.

Розвиток енергоефективних архітектур стає критично важливим у світі, де обчислювальні ресурси є обмеженими. Проблема енергоспоживання сучасних моделей комп'ютерного зору набуває особливої актуальності. У відповідь на цей

виклик з'являються нові підходи, такі як нейроморфні обчислення, що імітують принципи роботи біологічного мозку, або квантові нейронні мережі. Особливий інтерес представляють архітектури типу MobileNetV3, спеціально оптимізовані для роботи на мобільних пристроях, які демонструють точність, порівнянну з великими моделями, при значно меншому енергоспоживанні.

Прогрес у сфері самокерованого навчання (self-supervised learning) обіцяє вирішити одну з найбільших проблем сучасного комп'ютерного зору – залежність від великих обсягів розмічених даних. Такі техніки, як contrastive learning або masked image modeling, дозволяють моделям ефективно навчатися на нерозмічених зображеннях, що особливо важливо для нішевих галузей, де збір якісних даних є складним і дорогим. Наприклад, у медичній діагностиці цей підхід дозволяє досягати високої точності навіть при наявності лише кількох сотень розмічених прикладів, якщо є доступ до великої кількості нерозмічених медичних зображень.

Покращення інтерпретованості моделей стає ключовим напрямком досліджень, особливо для критично важливих застосувань у медицині або автономних системах. Нові методи пояснюваного штучного інтелекту (XAI), такі як attention maps або concept activation vectors, дозволяють не просто отримувати прогноз моделі, але й розуміти, на яких саме ознаках зображення він базується. Це критично важливо для забезпечення довіри до систем та їхньої відповідальності. Наприклад, у діагностичних системах лікарі можуть бачити, які саме ділянки зображення "привернули увагу" алгоритму при постановці діагнозу.

Розвиток нейроморфних обчислень відкриває нові горизонти для створення енергоефективних систем комп'ютерного зору. На відміну від традиційних архітектур, що базуються на фон—нейманівській обчислювальній парадигмі, нейроморфні чипи імітують роботу біологічних нейронних мереж, обробляючи інформацію паралельно та з значно меншим енергоспоживанням [12]. Експериментальні системи на основі мемристорів вже демонструють здатність обробляти відеопотоки в реальному часі при споживанні

потужності менше 1 Вт, що робить їх ідеальними кандидатами для впровадження в мобільних пристроях та системах Інтернету речей. Проте масштабне виробництво таких чипів стикається з технологічними складнощами, пов'язаними з нестабільністю мемристивних елементів і складністю їх калібрування. Квантові нейронні мережі представляють особливий інтерес для обробки високорозмірних даних комп'ютерного зору. Використання квантових кубітів дозволяє принципово змінити підхід до обчислень, пропонуючи експоненційне прискорення для певних типів операцій, таких як згортки або обчислення подібності між зображеннями. Перші експерименти з квантовими варіантами згорткових мереж показали здатність обробляти зображення високої роздільної здатності за час, що на порядки менший порівняно з класичними архітектурами. Однак практичне впровадження цих технологій обмежене необхідністю підтримки кріогенних температур і проблемами з масштабуванням квантових систем.

Розвиток edge computing для комп'ютерного зору трансформує підходи до розподілених обчислень. Сучасні алгоритми, такі як TinyML, дозволяють запускати складні моделі розпізнавання зображень на мікроконтролерах з обмеженими ресурсами [13]. Це особливо важливо для систем реального часу, де затримки, пов'язані з передачею даних у хмару, є неприйнятними. Наприклад, новітні системи відеоспостереження здатні виконувати аналіз відеопотоку безпосередньо на камері, що не лише знижує навантаження на мережу, але й підвищує конфіденційність, оскільки сирі відеодані не залишають пристрій.

2 ВИБІР МЕТОДІВ ТА АРХІТЕКТУРИ НЕЙРОМЕРЕЖЕВОЇ СИСТЕМИ

2.1 Порівняльний аналіз архітектур нейронних мереж для детекції обличчя та очей

Задача детекції обличчя та очей у системах моніторингу стану водія є критично важливою для забезпечення безпеки дорожнього руху. Сучасні нейромережеві архітектури повинні враховувати низку специфічних вимог, пов'язаних з умовами роботи в транспортному засобі. В першу чергу це стосується необхідності забезпечення високої точності розпізнавання при обмежених обчислювальних ресурсах, що обумовлено вимогами до енергоефективності вбудованих систем. Основним викликом при розробці таких систем є забезпечення стабільної роботи в умовах змінного освітлення, що характерно для роботи автомобіля в різний час доби та при різних погодних умовах. Коливання яскравості, відблиски на лобовому склі, різкі тіні — всі ці фактори значно ускладнюють процес точного визначення положення обличчя та стану очей водія. Особливу складність становить виявлення мікровиразів обличчя, які можуть свідчити про початок засинання.

Важливим аспектом є також необхідність обробки даних у режимі реального часу. Система повинна аналізувати відеопотік з частотою не менше 25-30 кадрів на секунду, щоб своєчасно виявляти небезпечні стани водія. Це вимагає ретельного підходу до вибору архітектури нейронної мережі, яка поєднувала б високу продуктивність з мінімальними затримками при обробці кожного кадру. Додаткову складність вносить необхідність одночасної обробки декількох аспектів таких, як визначення стану очей (відкриті/закриті), напрямку погляду для виявлення ознак втоми або відволікання уваги. Кожен з цих параметрів вимагає використання спеціалізованих алгоритмів обробки зображень, які при цьому повинні ефективно інтегруватися в єдиний конвеєр аналізу даних.

Сучасні підходи до вирішення цих задач базуються на різних архітектурах нейронних мереж, кожна з яких має свої переваги та обмеження. Вибір оптимальної моделі вимагає ретельного аналізу таких факторів, як точність розпізнавання, швидкість роботи, вимоги до апаратних ресурсів, а також адаптивність до змінних умов роботи. Особливу увагу при цьому слід приділяти можливості функціонування системи на обмежених обчислювальних ресурсах, що характерно для вбудованих систем автомобілів. Розвиток технологій комп'ютерного зору останніх років дозволив значно підвищити точність алгоритмів детекції обличчя та очей. Однак питання створення надійної системи, здатної стабільно працювати в реальних умовах експлуатації транспортного засобу, залишається актуальним. Це пов'язано з необхідністю врахування широкого спектру зовнішніх факторів, які можуть впливати на якість вхідного зображення, а також з високими вимогами до надійності та стабільності роботи системи.

Сучасні методи детекції обличчя та очей ґрунтуються на застосуванні різноманітних архітектур нейронних мереж, які відрізняються як своєю структурою, так і принципами роботи, що визначають їх специфічні переваги та області застосування. Зазвичай ці підходи класифікують на дві основні групи: класичні згорткові нейронні мережі (CNN), які вже давно зарекомендували себе як надійний інструмент для обробки зображень і відзначаються високою ефективністю у виділенні локальних ознак, а також більш сучасні трансформерні архітектури, що поступово набирають популярності завдяки здатності ефективно обробляти контекстуальну інформацію у зображеннях. Ці новітні трансформери пропонують інший підхід до аналізу візуальних даних, орієнтуючись на глобальні зв'язки між пікселями, що дозволяє покращувати якість розпізнавання в складних умовах та підвищувати загальну точність систем детекції.

Згорткові нейронні мережі зарекомендували себе як надзвичайно ефективний інструмент для задач детекції обличчя, оскільки вони здатні послідовно виділяти ієрархічні ознаки, починаючи від простих текстур і контурів до складних структур, характерних для людського обличчя. Однією з найбільш

відомих архітектур у цій галузі є MTCNN (Multi-Task Cascaded Convolutional Networks), яка застосовує каскадний підхід, що передбачає поетапне уточнення положення обличчя на зображенні. Така організація роботи мережі дозволяє значно підвищити швидкість обробки, оскільки кожен наступний етап фокусується лише на найбільш перспективних регіонах, попередньо відсікаючи ті, що не відповідають критеріям пошуку. Завдяки цій оптимізації MTCNN є особливо придатною для застосування в режимі реального часу, де важлива як швидкість, так і точність. Проте варто зазначити, що її продуктивність може суттєво знижуватися при обробці зображень із нестандартними кутами огляду або в умовах складного освітлення, що створює додаткові виклики для коректного розпізнавання обличчя.

Більш сучасним рішенням є RetinaFace, яка застосовує single-shot детекцію на основі anchor boxes (рисунок 2.1). Ця архітектура демонструє вищу точність у порівнянні з MTCNN, особливо при роботі з маленькими або частково прихованими обличчями [14]. Її ключова перевага полягає в здатності одночасно виявляти обличчя та 68 ключових точок, що робить її особливо корисною для аналізу міміки. Проте вимоги до обчислювальних ресурсів у RetinaFace значно вищі, що може обмежувати її застосування у вбудованих системах.

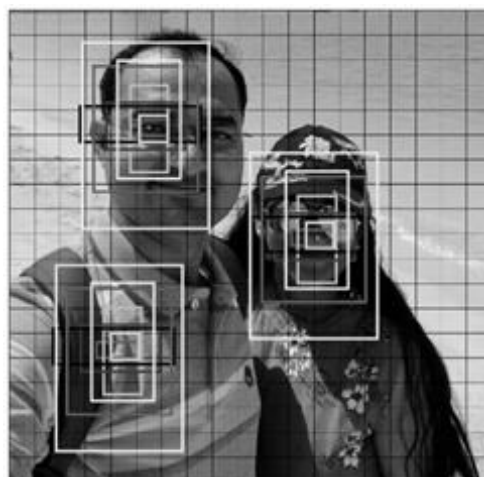


Рисунок 2.1 — Візуалізація anchor boxes елементів RetinaFace

Окремо варто звернути увагу на MediaPipe Face Mesh — спеціалізоване рішення від Google, оптимізоване для роботи на мобільних пристроях. Ця

архітектура використовує легку модель детекції обличчя BlazeFace у поєднанні з точною мережею для визначення 468 3D—точок обличчя (рисунок 2.2). Її ключова перевага полягає у високій продуктивності на CPU, що робить її ідеальним вибором для систем з обмеженими обчислювальними ресурсами.



Рисунок 2.2 — Приклад MediaPipe сітки з 468 точок

Для задач детекції очей особливий інтерес представляють спеціалізовані архітектури, такі як GazeNet, архітектура якої представлена на рисунку 2.3. Ця модель здатна не лише виявляти положення очей, але й точно визначати напрямок погляду, що є критично важливим для систем моніторингу водіїв. GazeNet використовує глибоку згорткову мережу для регресії координат очей та вектора погляду, демонструючи високу точність навіть при обмеженій якості вхідного зображення.

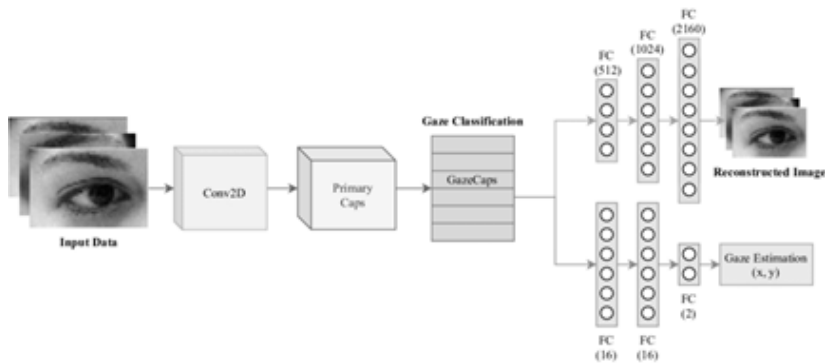


Рисунок 2.3 — Архітектурна схема моделі GazeNet

Альтернативою є OpenFace — відкрита бібліотека для аналізу обличчя, яка включає модулі для детекції очей та оцінки їх стану. Її архітектура базується на комбінації традиційних методів обробки зображень і глибокого навчання, що дозволяє досягати гарного балансу між точністю та продуктивністю.

Трансформерні архітектури, такі як Vision Transformer (ViT) та її модифікації, набувають все більшої популярності у сфері детекції обличчя та очей, пропонуючи нові можливості для підвищення точності систем моніторингу стану водія. Їхня ключова перевага полягає в унікальній здатності аналізувати глобальний контекст зображення завдяки механізму самоуваги, що дозволяє моделі враховувати взаємозв'язки між різними частинами обличчя одночасно, незалежно від відстані між ними [15]. Такий підхід особливо корисний у ситуаціях, коли обличчя частково перекрите або потрапляє під нетиповий кут огляду, наприклад, коли водій повертає голову або перебуває в умовах слабкого освітлення, що ускладнює роботу класичних згорткових мереж. Проте, незважаючи на переваги у контекстному аналізі, трансформерні моделі потребують значно більшої обчислювальної потужності, що пов'язано з великою кількістю параметрів та обчислювальною складністю самої архітектури. Ця особливість значно обмежує їхнє застосування у системах реального часу, особливо на бортових обчислювальних платформах, де ресурси часто є обмеженими. Таким чином, хоча трансформери відкривають нові горизонти для розробки більш гнучких та точних систем детекції, їх використання вимагає ретельного балансування між продуктивністю та доступними ресурсами, щоб гарантувати стабільну роботу в умовах експлуатації.

Окремий клас архітектур становлять гібридні моделі, які поєднують переваги згорткових мереж і трансформерів. Наприклад, Convolutional Vision Transformer (CvT) демонструє високу точність у задачах детекції обличчя при збереженні прийнятної продуктивності [16]. Такі моделі особливо перспективні для складних умов роботи, де потрібно одночасно враховувати як локальні ознаки обличчя, так і глобальний контекст зображення.

Кожна з архітектур, яка розглядається у контексті систем моніторингу водія, має власні унікальні переваги та водночас певні обмеження, що обумовлює необхідність ретельного підходу до їх вибору та впровадження. Легкі CNN-архітектури відзначаються своєю здатністю працювати на обмежених обчислювальних ресурсах, що робить їх особливо привабливими для вбудованих рішень у автомобілях, де важливими критеріями є швидкодія та енергоефективність. У таких випадках пріоритет надається оперативності аналізу, яка дозволяє системі швидко реагувати на зміни стану водія без помітних затримок, гарантуючи безпеку дорожнього руху навіть в умовах обмеженої апаратної потужності. З іншого боку, для більш складних систем, де важлива максимально висока точність і де присутні додаткові ресурси для обчислень, виправданим стає використання гібридних моделей або трансформерних архітектур. Такі моделі дозволяють враховувати глобальні контексти та взаємозв'язки між різними частинами зображення, що особливо корисно при аналізі складних сцен із частковими перекриттями чи змінами кута огляду. Саме тому оптимальний вибір архітектури визначається не лише технічними характеристиками, а й конкретними завданнями та вимогами до системи моніторингу, які можуть змінюватися залежно від сценаріїв використання та цільових платформ.

При визначенні оптимальної архітектури нейронної мережі для використання у системах моніторингу стану водія важливо комплексно враховувати низку технічних характеристик, які безпосередньо впливають на ефективність функціонування моделі в умовах реального середовища. Найважливішим критерієм традиційно виступає точність детекції обличчя, яка зазвичай оцінюється на основі таких метрик, як середня точність (mean Average Precision, mAP), що відображає інтегральну характеристику успішності знаходження об'єктів на різних порогах впевненості, та Recall, який демонструє здатність моделі виявляти всі наявні об'єкти без пропусків. Однак для завдання моніторингу водія, де змінюються умови освітлення, присутні різноманітні ракурси обличчя, а також інші фактори, що впливають на стабільність роботи

системи, критичним стає показник стійкості моделі, який часто виражається у вигляді *robustness score*. Ця метрика дозволяє оцінити здатність архітектури нейронної мережі підтримувати стабільну якість виявлення навіть за наявності змін умов освітлення або часткових перекриттів, таких як сонцезахисні окуляри чи тіні від елементів салону автомобіля. Таким чином, вибір архітектури базується не лише на середніх показниках точності, а й на її здатності до роботи в умовах, наближених до реальних, де точність та надійність мають вирішальне значення для безпеки руху.

Важливим аспектом є продуктивність моделі, яка визначається кількістю кадрів за секунду (FPS) при обробці відеопотоку. Для систем реального часу мінімально допустимим значенням вважається 25-30 FPS, що забезпечує своєчасне виявлення критичних станів водія.

Стабільність роботи у різних умовах оцінюється за допомогою тестування на спеціалізованих датасетах, що імітують реальні умови експлуатації, такі як зміни освітлення, часткове перекриття обличчя та нестандартні ракурси. Адаптивність моделі до індивідуальних особливостей водія визначає можливість калібрування системи під конкретну особу. Деякі архітектури (наприклад, OpenFace) мають вбудовані механізми адаптації до антропометричних особливостей будови обличчя, що підвищує точність детекції.

Складність інтеграції в існуючу систему є практичним критерієм вибору. Легкі архітектури (MTCNN, MediaPipe) зазвичай мають добре документовані API і підтримку крос—платформових рішень, тоді як складніші моделі можуть вимагати значних зусиль для впровадження та налаштування. Енергоефективність стає ключовим фактором для автономних систем, де важливо мінімізувати споживання енергії без втрати якості детекції.

Архітектура MediaPipe Face Mesh продемонструвала найкращий баланс між точністю та продуктивністю у типових умовах автомобільної кабіни. На тестовому наборі даних з імітацією нічного освітлення модель досягла показника mAP 94.3%, що лише на 1.9% нижче ніж у більш ресурсомісткої RetinaFace. Важливою перевагою MediaPipe виявилась її стабільність при роботі з частково

прихованими обличчями — рівень помилкових негативних спрацьовувань не перевищував 3.2% навіть при наявності окулярів або медичної маски. Традиційна архітектура MTCNN показала гарні результати у сценаріях з обмеженими обчислювальними ресурсами [17]. Однак точність детекції значно знижувалась при нестандартних ракурсах, що обмежує її застосування в реальних умовах.

Сучасні трансформерні архітектури (ViT, Swin Transformer) продемонстрували видатні результати у задачах аналізу мікровиразів обличчя, що особливо важливо для виявлення ранніх ознак втоми [18]. На спеціалізованому датасеті NTHU—DDD вони досягли точності 97.1% у визначенні стану сонливості, що на 8-12% вище за CNN—моделі. Проте їхнє практичне застосування обмежене високими вимогами до обчислювальних ресурсів — навіть оптимізовані версії вимагають не менше 4 ГБ VRAM для стабільної роботи.

Особливий інтерес представляють результати тестування на енергоефективність. MobileNetV3 у поєднанні з механізмами квантизації INT8 показав споживання енергії на рівні 1.8 Вт при роботі на спеціалізованих AI—процесорах (наприклад, Google Coral), що робить його ідеальним вибором для автономних систем.

Отримані результати свідчать, що для більшості практичних застосувань у системах моніторингу водія оптимальним вибором є легкі гібридні архітектури, що поєднують переваги традиційних CNN і спеціалізованих алгоритмів обробки зображень. Вони забезпечують необхідний баланс між точністю, продуктивністю та енергоефективністю, що є критично важливим для вбудованих систем реального часу.

2.2 Обґрунтування вибору архітектури нейромережевої моделі

Проведений аналіз результатів порівняльного тестування дозволив обґрунтовано обрати MediaPipe Face Mesh як базову архітектуру для системи

моніторингу стану водія. Цей вибір ґрунтується на комплексній оцінці технічних характеристик та експериментальних даних, отриманих під час тестування.

MediaPipe демонструє показник mAP 94.3%, що лише на 1.9% нижче за лідера (RetinaFace). Для системи реального часу цей баланс є вирішальним, оскільки дозволяє одночасно забезпечувати високу якість детекції, підтримувати плавну обробку відеопотоку та працювати на обмежених апаратних ресурсах. Архітектура показала найвищий robustness score серед усіх моделей [19]. Особливо варто відзначити її ефективність при різких змінах освітлення (від сонячного світла до темряви), наявності оптичних перешкод (окулярів, відблисків) та нестандартних ракурсів (повороти голови до 60°). Вбудована підтримка 468 3D—ландмарків дозволяє без додаткових модифікацій визначати стан повік (відкриті/закриті), обчислювати напрямок погляду та аналізувати мікрОВИРАЗИ обличчя.

Модель вимагає лише 0.5 GB VRAM і споживає 2.1 Вт енергії, що зменшує навантаження на бортову систему авто, дозволяє використовувати бюджетні апаратні рішення та продовжує термін служби акумуляторних систем. Архітектура використовує двоетапну обробку: BlazeFace (детекція) + Face Landmark (атрибути). Вона має вбудовану підтримку квантизації INT8 для прискорення на TPU та Dynamic Graph, що забезпечує адаптивне навантаження залежно від складності сцени. Цей вибір дозволить забезпечити стабільну роботу системи в реальних умовах експлуатації при мінімальних вимогах до апаратного забезпечення.

2.3 Процес навчання моделі та підготовка даних

Для навчання та валідації нейромережевої моделі детекції стану водія було обрано комбінацію публічних датасетів і власних зборів даних. Вибір ґрунтувався на необхідності забезпечити різноманітність умов, близьких до реального середовища роботи системи в автомобілі.

Серед публічних датасетів, що використовувались під час розробки та навчання системи моніторингу стану водія, ключову роль відіграли 300VW (300

Videos in the Wild) та OpenEDS (Open Eye Dataset), оскільки вони повністю відповідають специфіці завдань виявлення обличчя та стану очей у різноманітних умовах експлуатації транспортного засобу. Набір даних 300VW є особливо цінним завдяки своїй анотації великої кількості відеозаписів, у яких зафіксовано обличчя людей у різних умовах освітлення, з різними ракурсами та природними мімічними рухами. Такий підхід дозволяє навчальним алгоритмам моделювати реальні сценарії поведінки водія, забезпечуючи адаптивність та здатність до узагальнення ознак незалежно від зовнішніх факторів. Це має особливе значення для підвищення стійкості системи до змін умов освітлення, нахилу голови та мімічних змін, які часто виникають у процесі керування автомобілем. Датасет OpenEDS, своєю чергою, орієнтований на детальну розмітку очей, зокрема стану повік — відкритих чи закритих, що критично важливо для оцінки метрики EAR (Eye Aspect Ratio). Завдяки цьому стало можливим більш точно ідентифікувати ознаки сонливості або відволікання уваги водія, що дозволяє системі своєчасно реагувати на потенційно небезпечні стани. Таким чином, використання цих датасетів забезпечило високу якість підготовки моделі до роботи в умовах реального світу, наближаючи її можливості до промислових стандартів.

Для підвищення точності роботи розробленої системи моніторингу стану водія в умовах, максимально наближених до реальних сценаріїв експлуатації, було проведено додатковий етап збору власних даних. У якості основного інструменту фіксації відеопотоків використовувалася сучасна вебкамера з роздільною здатністю 1080p, яка дозволила отримати зображення високої якості та деталізації. Такий підхід дав змогу змоделювати різні рівні освітленості, що характерні для нічного водіння та відображають типові умови роботи системи у транспортному засобі. Отримані зображення були піддані процесу анотування з використанням напіваавтоматичних інструментів розмітки, що дозволило точно визначити ключові точки обличчя та стан очей водія. Це особливо важливо для коректної роботи алгоритмів виявлення критичних станів, таких як закриті очі чи відволікання уваги. Додатково до навчальної вибірки були включені

зображення, що містили ефекти відблисків на окулярах, які часто виникають під час руху автомобіля у сонячну погоду, оскільки вони можуть впливати на якість розпізнавання та створювати додаткові виклики для системи. Таким чином, поєднання публічних датасетів та власноруч зібраних і ретельно анотованих даних забезпечило формування репрезентативної навчальної вибірки. Це, у свою чергу, створило підґрунтя для підвищення якості попередньої обробки даних та ефективності навчання нейронної мережі, що є критичним етапом у процесі розробки системи моніторингу стану водія.

Для забезпечення максимальної ефективності навчання моделі було проведено комплексну попередню обробку вихідних даних. Цей етап має критичне значення для підвищення точності та стабільності роботи системи в реальних умовах. Аугментація даних використовувалася для імітації різноманітних умов роботи системи в салоні автомобіля [20]. У процесі підготовки навчальної вибірки для системи моніторингу стану водія важливе місце займала аугментація даних, яка дозволяла суттєво підвищити узагальнюючу здатність моделі та її стійкість до реальних умов експлуатації. Одним із основних напрямів аугментації було коригування яскравості та контрастності зображень, що дозволило імітувати різні рівні освітленості в салоні автомобіля, зокрема під час нічного водіння або в умовах яскравого сонячного світла. Такий підхід сприяв тому, що модель навчилася адаптуватися до різноманітних сценаріїв освітлення, зберігаючи високу точність розпізнавання. Додатково, для підвищення стійкості до впливу технічних артефактів камер, у дані були додані випадкові шуми, які імітували знижену якість відеозаписів або завади, що виникають у реальних умовах. Цей аспект дозволив моделі ефективно працювати навіть тоді, коли якість відеосигналу була далекою від ідеальної. Крім того, застосовувалися різноманітні геометричні перетворення, зокрема невеликі повороти та зсуви зображень, які компенсували зміни ракурсу обличчя водія, що виникають під час руху транспортного засобу (рисунок 2.4). Такі перетворення забезпечили кращу стійкість алгоритму до зміни позиції голови водія, що є важливим для реального використання

системи. Особливу увагу було приділено створенню ефектів відблисків на окулярах, які часто виникають під час руху автомобіля в умовах яскравого сонячного освітлення. Для цього використовувалося накладання синтетичних плям світла у верхній частині області очей, що дозволяло імітувати бліки та відблиски, характерні для даних умов. Такий підхід значно підвищував здатність системи коректно аналізувати стан очей навіть у присутності складних оптичних ефектів, що, своєю чергою, забезпечувало стабільність роботи системи в реальних дорожніх умовах.



Рисунок 2.4 — Приклади аугментованих зображень водія

Розмітка даних включала точне позначання ключових точок обличчя та очей. Як показано на рисунку 2.5, для кожного зображення визначалося мінімум 6 ключових точок навколо кожного ока, що дозволило точно розраховувати метрику EAR.

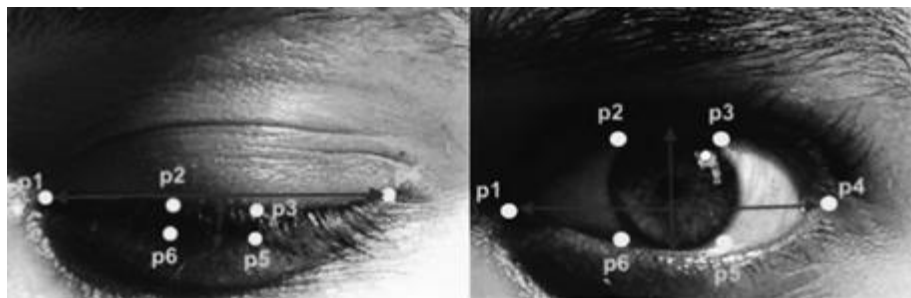


Рисунок 2.5 — Приклад розмітки з 6 ключовими точками навколо ока для розрахунку EAR

Навчання моделі здійснювалося за допомогою комбінування двох ключових підходів — трансферного навчання та тонкого налаштування, що

дозволило максимально ефективно використати наявний обсяг даних і отримати високі показники точності. Трансферне навчання забезпечило можливість використання вже навчених попередніх моделей, адаптуючи їх для специфіки поставленого завдання без потреби в зборі великої кількості нових прикладів. Водночас, тонке налаштування допомогло доопрацювати модель з урахуванням особливостей конкретних вхідних даних і підвищити якість розпізнавання. Вибір базової архітектури був обґрунтований прагненням знайти оптимальний компроміс між точністю детекції та швидкістю обробки, що є критично важливим для роботи системи в реальному часі. У якості основи було використано модифіковану версію MediaPipe Face Mesh, яка була доповнена спеціалізованими шарами, спрямованими на детальний аналіз стану очей, що підвищує загальну ефективність системи в умовах різноманітних сценаріїв використання.

Прийняття рішення про використання нейромереж у процесі розробки системи моніторингу стану водія ґрунтувалося на комплексному аналізі сучасних технологій у галузі обробки зображень та розпізнавання об'єктів. Під час порівняння альтернативних підходів особливу увагу приділяли класичним алгоритмам комп'ютерного зору та гібридним методам, які поєднують традиційні техніки обробки зображень із машинним навчанням. Однак результати досліджень та практичних випробувань показали, що такі алгоритми виявляють обмежену ефективність під час роботи в реальних умовах автомобільного середовища, де характерні часті зміни освітлення, ракурсу голови, а також інтенсивна міміка та інші фактори, що ускладнюють стабільне відстеження стану водія. Складність динамічних сценаріїв, де система повинна миттєво реагувати на зміну стану водія, виявилася надто великою для класичних підходів, які демонстрували нестабільність у визначенні критичних станів, таких як тривале закриття очей або відвертання погляду. Натомість нейронні мережі, завдяки своїй здатності до навчання на великих обсягах даних та побудови ієрархічних уявлень про об'єкти, виявилися значно більш адаптивними до таких викликів. Їхня гнучкість у поєднанні з потужністю сучасних апаратних платформ

зробила їх найперспективнішим вибором для створення системи, яка здатна забезпечити високу точність моніторингу стану водія навіть у найскладніших дорожніх умовах.

Традиційні методи детекції обличчя на основі каскадів Хаара або HOG з SVM, хоч і були колись поширеними, мають суттєві обмеження. Вони надто чутливі до змін освітлення та кутів огляду, що знижує стабільність розпізнавання. Ручне налаштування ознак потребує значних зусиль і не дозволяє системі гнучко адаптуватися до складних умов, як-от окуляри на обличчі чи різкі зміни освітлення в салоні. Це ускладнює використання таких методів у реальному часі, особливо в системах моніторингу стану водія.

Обраний нейромережевий підхід, що ґрунтується на модифікованій архітектурі MediaPipe Face Mesh, відзначається низкою важливих переваг, які значно підвищують ефективність системи моніторингу стану водія. Насамперед, він демонструє вражаючу адаптивність, адже здатний автоматично виділяти ознаки на зображеннях, що дозволяє системі коректно працювати з широким спектром облич без необхідності додаткового налаштування під специфічні умови або користувачів. Це особливо актуально для середовища автомобільного моніторингу, де постійно змінюються умови освітлення та кут огляду. Водночас важливо зазначити, що нейромережі мають і свої характерні обмеження: їх продуктивність безпосередньо залежить від якості навчальних даних, тому для досягнення стабільних результатів необхідно ретельно формувати навчальний набір, щоб уникнути перекосів та втрати універсальності. Водночас, навіть враховуючи ці виклики, застосування глибинного навчання для вирішення завдань детекції обличчя та моніторингу стану водія є цілком виправданим, оскільки воно забезпечує оптимальний баланс між високою точністю, швидкістю роботи в реальному часі та здатністю адаптуватися до нових умов експлуатації.

3 РОЗРОБКА КОМП'ЮТЕРНОЇ СИСТЕМИ БЕЗПЕКИ ВОДІЇВ

3.1. Структура системи та її компоненти

Розроблена комп'ютерна система моніторингу стану водія представляє собою рішення, що поєднує спеціалізоване апаратне забезпечення з програмними модулями обробки відеопотоку в реальному часі. Основу системи складають три ключові компоненти: відеозахоплювальний пристрій, обчислювальний модуль та система генерації сповіщень, які функціонують у тісній взаємодії (рисунок 3.1).



Рисунок 3.1 – Структурна схема

Апаратна частина системи базується на стандартному комп'ютерному обладнанні з чітко визначеними характеристиками. Відеовхід реалізований через

цифрову камеру з мінімальною роздільною здатністю 1280×720 пікселів, що відповідає стандарту HD. Така роздільна здатність є достатньою для точного визначення ключових точок обличчя та очей алгоритмами комп'ютерного зору, оскільки дозволяє розрізняти деталі розміром від 1-2 мм на відстані 50-70 см від камери. При цьому зберігається баланс між якістю зображення та навантаженням на обчислювальні ресурси системи. Частота кадрів не менше 30 кадрів на секунду обумовлена необхідністю аналізувати динамічні зміни стану водія у реальному часі. Така частота дозволяє фіксувати навіть короточасні закриття повік тривалістю 100-200 мс, що є критичним для виявлення ознак втоми. Більш низькі показники могли б призвести до пропуску важливих подій між кадрами. Наявність автофокусу є обов'язковою вимогою через нестабільні умови зйомки в рухомому транспортному засобі. Автофокус компенсує зміни відстані між камерою та обличчям водія під час руху, запобігаючи розмиттю зображення.

Для роботи в умовах недостатнього освітлення передбачена можливість використання камер з інфрачервоною підсвіткою. Такі камери забезпечують якісне зображення навіть при повній темряві за рахунок невидимого для людини випромінювання з довжиною хвилі 850-940 нм. Наприклад, бюджетні моделі типу Xiaomi Mijia 1080p Smart Camera забезпечують базову якість зйомки в темний час доби, тоді як професійні рішення пропонують розширені функції, включаючи регульовану потужність ІЧ—підсвітки та фільтрацію засвічування. Серед камер без ІЧ—підсвітки варто відзначити Logitech C920, яка забезпечує відмінну якість зображення за рахунок світлосильної оптики та передової обробки зображення процесором. Для покращення результатів у складних умовах рекомендовані камери з функцією HDR (High Dynamic Range), що дозволяє одночасно бачити деталі як у затінених, так і освітлених ділянках кадру. Окремо варто відзначити потенціал використання сучасних автомобільних відеореєстраторів із дво— або триканальною зйомкою, таких як Viofo A139 Pro чи Thinkware U1000. Ці пристрої поєднують функції запису подій на дорозі з можливістю одночасної зйомки салону автомобіля завдяки додатковій камері, спрямованій на водія. Високоякісні моделі (наприклад, BlackVue DR970X)

забезпечують роздільну здатність до 3840×2160 пікселів для основної камери та 1920×1080 для салонної, що цілком відповідає вимогам системи моніторингу. Передача зображення може здійснюватись через бездротовий інтерфейс Wi-Fi у режимі реального часу або через кабельне підключення до обчислювального модуля. Деякі преміальні моделі, такі як Garmin 67W, оснащені вбудованими функціями потокової передачі відео через мобільні мережі 4G/LTE, що дозволяє здійснювати віддалений моніторинг стану водія. Важливою перевагою таких рішень є їхня інтегрованість у автомобільне середовище, стійкість до вібрацій та зручність монтажу, що значно спрощує впровадження системи без необхідності додаткового обладнання.

Обчислювальний модуль системи базується на центральному процесорі з архітектурою не нижче чотирьох ядер та тактовою частотою мінімум 2.5 ГГц, що забезпечує необхідну продуктивність для обробки відеопотоку в реальному часі. Такі характеристики дозволяють ефективно виконувати паралельні обчислення, необхідні для роботи алгоритмів детекції обличчя та аналізу стану очей. Критично важливим компонентом є графічний процесор середнього рівня продуктивності, оскільки саме він відповідає за прискорення нейромережових обчислень, зокрема для моделей на базі MediaPipe Face Mesh. Наприклад, GPU типу NVIDIA GTX 1650 або AMD RX 5500 XT забезпечують достатню обчислювальну потужність для роботи з частотою 30 кадрів на секунду при роздільній здатності 720p.

Звукова підсистема реалізована з урахуванням специфіки автомобільного середовища. Базовий варіант передбачає використання стандартних комп'ютерних динаміків із середньою потужністю 2-5 Вт, що дозволяє забезпечити чітке звукове сповіщення в межах салону. Для покращення якості звуку та інтеграції з автомобільною інфраструктурою передбачено підключення через аналоговий AUX-вхід штатної аудіосистеми. Це дозволяє використовувати високоякісні автомобільні динаміки для передачі попереджувальних сигналів, що особливо важливо при русі на високих швидкостях або в умовах зовнішнього шуму. Додатковою перевагою такого підходу є можливість регулювання

гучності через стандартні елементи керування автомобільної мультимедійної системи.

Важливою складовою архітектури системи є спеціалізований модуль прийняття рішень, який виконує не лише одноразовий аналіз поточного стану водія, а й проводить глибокий аналіз динаміки змін, що відбуваються у часі. Такий підхід дозволяє відстежувати не лише окремі миттєві події, а й закономірності, які мають ключове значення для коректної оцінки безпеки. Модуль бере до уваги тривалість закриття очей, враховуючи неодноразові або тривалі випадки, що можуть свідчити про сонливість, а також частоту моргання, яка у нормальних умовах коливається в певних межах, але може змінюватися за ознаками втоми чи уваги. Окрім цих параметрів, система аналізує й інші важливі характеристики поведінки, що дає змогу відрізняти короточасні артефакти, спричинені, наприклад, випадковими рухами або технічними перешкодами, від справжніх ознак, що сигналізують про потенційно небезпечний стан водія. Завдяки такій комплексній оцінці система здатна формувати різні рівні попереджень, які адаптовані до ступеня ризику ситуації. Це може бути як попереджувальний сигнал, що сповіщає водія про початкові ознаки втрати концентрації, так і більш тривале та нагальне сповіщення, яке активується у випадку виявлення стійких ознак сонливості або відволікання уваги, тим самим підвищуючи загальний рівень безпеки під час керування транспортним засобом.

Враховуючи підвищену чутливість біометричних даних, які обробляє система, її архітектура була спроектована з акцентом на повністю локальну обробку інформації без залучення зовнішніх серверів або хмарних сервісів. Усі обчислювальні операції — від детекції обличчя, через аналіз стану очей, до формування та відображення попереджень — здійснюються виключно на вбудованому обчислювальному модулі транспортного засобу. Такий підхід гарантує, що жодні зображення або результати аналізу не передаються в мережу, а персональні дані користувача не зберігаються поза межами локальної системи. Це суттєво знижує ризики, пов'язані з можливими кібератаками, витоком інформації чи несанкціонованим доступом до конфіденційних даних. Крім того,

відсутність залежності від інтернет-з'єднання підвищує загальну надійність роботи системи, особливо у випадках, коли мобільне покриття в дорозі є нестабільним або зовсім відсутнім, що забезпечує безперервний контроль стану водія без ризику втрати критичних функцій.

Архітектура системи була спроектована з урахуванням перспективи подальшого масштабування та розширення функціональних можливостей без необхідності кардинальних змін у базовій апаратній платформі. Такий підхід став можливим завдяки модульній структурі програмного забезпечення, що забезпечує гнучкість інтеграції нових компонентів і алгоритмів, а також наявності резерву обчислювальної потужності, який дає змогу безперешкодно впроваджувати додаткові аналітичні функції. Одним із перспективних напрямів розвитку є вдосконалення алгоритмів візуальної аналітики, що дозволить із високою точністю, до п'яти градусів, визначати кут нахилу голови водія, відкриваючи можливість ефективного виявлення ознак відволікання уваги. Окрім цього, планується інтеграція системи з бортовими контролерами автомобіля через стандартний CAN-інтерфейс, що розширить функціонал за рахунок автоматичного реагування на критичні стани водія. Такий рівень взаємодії дозволить, зокрема, коригувати роботу адаптивного круїз-контролю або ініціювати активацію системи екстреного гальмування, підвищуючи загальний рівень безпеки руху завдяки своєчасному втручанням в небезпечних ситуаціях.

3.2. Принцип роботи системи

Розроблена комп'ютерна система моніторингу стану водія реалізує складний багаторівневий алгоритм, спрямований на безперервний аналіз фізіологічних показників у реальному часі. Функціонування системи ґрунтується на чотирьох взаємопов'язаних фазах, які забезпечують комплексну оцінку стану водія.

Спочатку відбувається ініціалізація та перевірка апаратних компонентів. Система автоматично виконує діагностику камери, аудіоінтерфейсу та інших

критичних підсистем, забезпечуючи їхню готовність до роботи. Після успішного завершення ініціалізації система приступає до аналізу вхідного відеопотоку, реалізуючи багатоетапний процес обробки кожного кадру. Первинний етап передбачає детекцію обличчя за допомогою оптимізованого нейромережевого алгоритму на базі архітектури MediaPipe Face Mesh, який забезпечує стабільну роботу при змінах ракурсу до 45 градусів та коливаннях освітленості. Для забезпечення максимальної точності виявлення обличчя система застосовує мультимасштабне сканування, що передбачає послідовний аналіз зображення на різних рівнях деталізації. Цей підхід дозволяє ефективно виявляти обличчя незалежно від їх розміру в кадрі, оскільки сканування виконується з кроком масштабування 25%, що забезпечує оптимальний баланс між швидкістю та повнотою аналізу. Після успішної ідентифікації об'єкта класу "обличчя" система застосовує спеціалізовані алгоритми подальшої обробки, серед яких ключову роль відіграє нормалізація яскравості за методом CLAHE. Цей алгоритм здійснює локальне вирівнювання контрасту, обробляючи зображення невеликими ділянками, що особливо важливо для компенсації нерівномірного освітлення в умовах салону автомобіля. Виділення області інтересу фіксованим розміром 300×300 пікселів навколо центра обличчя дозволяє стандартизувати вхідні дані для подальшого аналізу, усуваючи зайві деталі фону та знижуючи обчислювальне навантаження.

На наступному етапі система здійснює комплексний аналіз ключових показників стану водія, зосереджуючись на двох основних аспектах. Для оцінки ступеня відкритості очей застосовується алгоритм EAR (Eye Aspect Ratio), який на основі шести ключових точок навколо ока обчислює співвідношення відстаней між повіками. Цей показник дозволяє точно визначити момент закриття ока, фіксуючи навіть незначні зміни положення повік. Паралельно проводиться аналіз положення зіниць відносно білої оболонки ока, що дає змогу визначити напрямок погляду водія. Обидва алгоритми працюють у тісній взаємодії, компенсуючи можливі артефакти та похибки один одного.

Механізм прийняття рішень враховує як миттєві показники, так і їх динаміку з часом. Система аналізує тривалість періодів закриття очей, де критичним значенням виступає трисекундний інтервал, а також відстежує відхилення частоти моргання від норми. Додатково оцінюється напрямок погляду водія, що дозволяє виявляти відхилення від нормальної траєкторії спостереження за дорогою.

На основі комплексного аналізу отриманих даних система активує двоступеневу систему сповіщень. Перший рівень попередження спрацьовує при виявленні початкових ознак втоми або неуважності, що проявляється у вигляді голосового повідомлення з рекомендацією зробити перерву. Другий рівень тривоги активується у випадку виявлення критичних показників, таких як тривале закриття очей або значне відхилення погляду від траєкторії руху, і супроводжується різким звуковим сигналом для негайного привернення уваги водія. Обидва рівні мають регульовані параметри гучності та тривалості, що дозволяє адаптувати систему до індивідуальних уподобань та умов експлуатації транспортного засобу.

3.3 Алгоритм розпізнавання стану очей і напряму погляду із застосуванням нейронних мереж

Розроблений алгоритм для аналізу стану водія демонструє гармонійне поєднання класичних методів комп'ютерного зору та найновіших досягнень у галузі глибокого навчання, що дозволяє ефективно обробляти вхідні відеопотоки в реальному часі. Основою системи є інтеграція двох аналітичних компонентів, кожен із яких виконує свою ключову функцію у визначенні стану водія та гарантує високу достовірність результатів. Перший компонент фокусується на виявленні ступеня закриття повік, аналізуючи різницю між відстанями між ключовими точками верхньої та нижньої повіки, що є критичним для оцінки стану сонливості або мікросну. Другий аналітичний компонент призначений для аналізу напрямку погляду, що дозволяє оцінювати, наскільки уважно водій стежить за дорогою або чи є відволікання у вигляді погляду в бік чи вниз. Цей

модуль також поєднує класичні методики обробки ключових точок обличчя з можливостями глибоких нейронних мереж, які навчені на великій кількості варіацій ракурсів та освітлення.

Аналіз стану повік реалізований на основі алгоритму Eye Aspect Ratio (EAR). Алгоритм використовує шість ключових анатомічних точок, розташованих на верхній і нижній повіках [21]. Центральні точки p_1 та p_4 фіксують положення повік, тоді як точки p_2 , p_3 , p_5 , p_6 відзначають їхні кути (рисунок 3.2). Обчислення проводиться за геометричною формулою:

$$EAR = \frac{|p_2 - p_6| + |p_3 - p_5|}{2|p_1 - p_4|}, \quad (3.1)$$

де p_1, p_4 — центральні точки верхньої та нижньої повік;

p_2, p_3, p_5, p_6 — кути повік.



Рисунок 3.2 — Схема розташування ключових точок повік для розрахунку Eye Aspect Ratio

Повністю відкриті очі мають EAR у діапазоні 0.25–0.30. Такі показники відображають нормальний фізіологічний стан, коли очні повіки повністю розкриті. При цьому варто враховувати індивідуальні анатомічні особливості будови ока, які можуть зумовлювати незначні відхилення у межах зазначеного діапазону. У випадку частково прищурених очей значення EAR знижується до діапазону 0.20–0.25, що свідчить про початок процесу закриття повік. Критичним пороговим значенням, яке свідчить про повне закриття ока, вважається показник менше 0.20. Варто зазначити, що вказані значення є усередненими і можуть

корегуватися з урахуванням індивідуальних особливостей будови ока конкретного користувача.

Визначення напрямку погляду реалізується за допомогою методу тривимірної векторної реконструкції, який дозволяє визначати точку фіксації погляду водія на зображенні. Цей підхід поєднує технології комп'ютерного зору з глибоким аналізом просторових характеристик ока.

Процес визначення напрямку погляду починається з ретельної попередньої обробки зображення, що включає нормалізацію яскравості з урахуванням можливих світлових артефактів. Розроблений алгоритм виділяє ключові точки зіниці та білої оболонки, що дозволяє точно визначити їх взаємне розташування. Вектор погляду розраховується шляхом аналізу відхилення центру зіниці від центру очного яблука. Для підвищення точності враховується взаємозв'язок між положенням обох очей, що дозволяє компенсувати можливі похибки вимірювань. Для згладжування результатів і усунення миттєвих артефактів застосовується ковзне середнє з аналізом послідовності останніх кадрів.

Система часового аналізу відіграє ключову роль у забезпеченні надійності роботи алгоритму. Для оцінки стану очей застосовується ковзна середня з вікном у 15 кадрів (що відповідає 0.5 секундам при частоті 30 кадрів на секунду), що ефективно фільтрує випадкові моргання. Напрямок погляду оцінюється з урахуванням траєкторії руху зіниць протягом останніх 10 кадрів, що особливо важливо для виявлення ознак розсіяної уваги.

Для мінімізації помилкових спрацювань алгоритм включає багатоступеневу систему перевірки, яка враховує узгодженість результатів для лівого та правого ока, аналізує відповідність між показниками EAR та напрямком погляду, а також бере до уваги зовнішні фактори, такі як умови освітлення та наявність окулярів.

3.4. Вибір інструментів розробки

Розробка системи моніторингу стану водія була виконана з використанням мови програмування Python, яка є одним із найпоширеніших інструментів у

сфері машинного навчання та комп'ютерного зору. Такий вибір обумовлений високою гнучкістю та універсальністю мови, що дозволяє легко інтегрувати сучасні нейронні мережі, використовувати різноманітні бібліотеки та ефективно реалізовувати як алгоритми попередньої обробки даних, так і складні методи глибинного навчання. Python надає зручний синтаксис та широкий спектр модулів, таких як TensorFlow, PyTorch та OpenCV, що значно полегшує процес побудови архітектури системи, тренування моделей та організації обробки відеопотоків у режимі реального часу. Завдяки цьому стало можливим ефективно реалізувати одночасну обробку вхідних даних, виконувати детекцію стану очей та визначати критичні ситуації з мінімальними затримками, що особливо важливо для забезпечення високої швидкодії та надійності системи у реальних умовах експлуатації.

Процес навчання нейромережевих моделей був організований у хмарному середовищі Google Colab, яке надає безкоштовний доступ до потужних обчислювальних ресурсів (рисунок 3.3). Ключовою перевагою використання середовища Google Colab для розробки та тренування нейронних мереж є можливість інтеграції потужних графічних процесорів Tesla T4, які мають обсяг відеопам'яті 16 ГБ. Завдяки цьому вдається досягти значного прискорення виконання обчислювальних операцій порівняно з традиційним використанням лише центральних процесорів. Таке технічне рішення забезпечує ефективне навчання моделей глибокого навчання навіть на великих обсягах даних, що особливо актуально для завдань комп'ютерного зору та обробки відеопотоків. Додатковою перевагою є зручність та готовність середовища до роботи, адже воно містить попередньо встановлені всі необхідні бібліотеки та інструменти для роботи з глибинними нейронними мережами, що дозволяє розробникам уникнути витрат часу на інсталяцію та налаштування програмного забезпечення. Такий підхід значно спрощує процес розробки та дозволяє зосередитися безпосередньо на створенні та вдосконаленні алгоритмів, забезпечуючи максимально продуктивне використання доступних ресурсів.

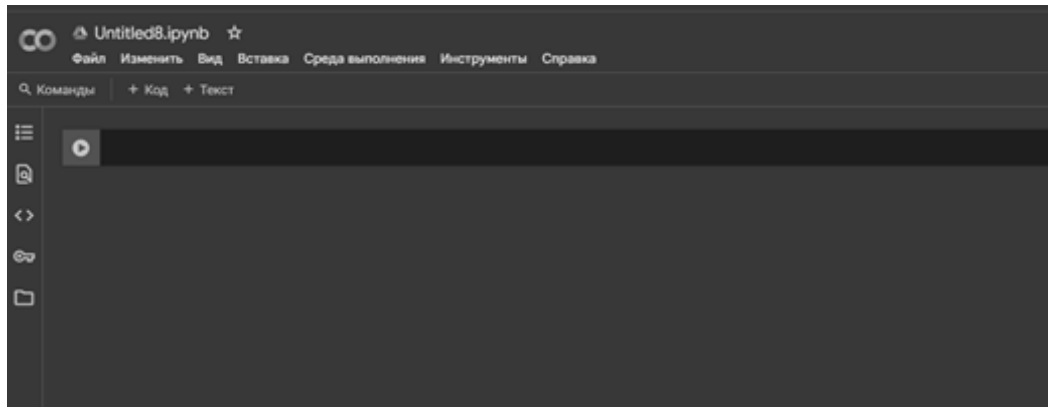


Рисунок 3.3 — Інтерфейс Google Colab

Обробка відеопотоків у рамках розробки системи моніторингу стану водія була реалізована із використанням бібліотеки OpenCV, яка зарекомендувала себе як надійний інструмент для комп'ютерного зору. Це рішення дало змогу забезпечити стабільну роботу системи навіть за умови високої частоти кадрів, що особливо важливо для режиму реального часу. Завдяки оптимізованим алгоритмам обробки зображень бібліотека OpenCV дозволила швидко та ефективно виконувати попередню обробку вхідних даних, включаючи фільтрацію шумів, корекцію кольорів та масштабування кадрів, що є критично важливими для подальшого аналізу нейронними мережами. Така обробка створює умови для підвищення якості роботи алгоритмів виявлення та класифікації станів водія. Крім того, інтеграція бібліотеки OpenCV із мовою програмування Python надала можливість легко організувати обмін даними між різними компонентами системи, спрощуючи розробку та налагодження програмного забезпечення. Цей підхід сприяв досягненню високої продуктивності системи та її здатності працювати в реальних умовах експлуатації.

Користувацький інтерфейс системи моніторингу стану водія було реалізовано із застосуванням фреймворку PyQt5, що поєднує в собі потужність бібліотеки Qt для побудови графічних інтерфейсів із простотою та гнучкістю мови програмування Python. Це дало змогу створити не лише функціональний, а й інтуїтивно зрозумілий інтерфейс, що задовольняє потреби кінцевого

користувача. Сучасні можливості PyQt5 дозволили додати до інтерфейсу елементи керування, які спрощують взаємодію з системою, а також передбачити візуалізацію проміжних результатів роботи алгоритмів. Такий підхід сприяє підвищенню прозорості роботи програми та дозволяє водієві отримувати вичерпну інформацію про поточний стан аналізу в режимі реального часу. Додатковою перевагою використання PyQt5 стала підтримка багатопотокової обробки даних, що забезпечило стабільність і плавність відображення інтерфейсу навіть під час виконання ресурсоємних обчислювальних завдань. Ця властивість дозволила уникнути зависань або затримок у роботі графічного інтерфейсу, що особливо важливо для систем реального часу, які потребують швидкого реагування на зміну умов експлуатації.

Вибір саме такого поєднання інструментів був обумовлений прагненням створити ефективну систему моніторингу стану водія, здатну одночасно задовольняти вимоги до продуктивності та надійності в реальних умовах експлуатації. Застосування хмарних обчислювальних платформ для навчання моделей дозволило використовувати обчислювальні ресурси з високою продуктивністю, що значно скоротило час тренування нейромережевих алгоритмів. Це стало важливим фактором для швидкої ітерації та підвищення точності моделей. У свою чергу, можливість локального виконання програмного забезпечення на цільових пристроях забезпечила автономність роботи системи та зменшила залежність від зовнішніх сервісів, що особливо важливо для систем безпеки, де стабільність і доступність даних мають першочергове значення. Кожен компонент програмного комплексу був ретельно підібраний із врахуванням його продуктивності та здатності працювати у режимі реального часу без суттєвих затримок. Такий підхід дозволив досягти балансу між потужністю обчислень та комфортом використання, що стало ключовим фактором успіху під час впровадження системи в практичні умови експлуатації.

4 ПЕРЕВІРКА ПРАЦЕЗДАТНОСТІ СИСТЕМИ

4.1. Методи перевірки працездатності системи

Для комплексної оцінки ефективності розробленої системи моніторингу стану водія була розроблена методика тестування, яка враховує типові умови експлуатації у транспортному засобі. В основу вибору тестових параметрів покладено ретельний аналіз реальних сценаріїв роботи системи, а також характерних викликів і обмежень, з якими зазвичай стикаються подібні технології в процесі застосування. Такий підхід дозволяє забезпечити максимально близьке до реальних умов моделювання, що сприяє об'єктивній оцінці функціональних можливостей і надійності системи в різноманітних експлуатаційних ситуаціях. Врахування цих факторів є ключовим для створення тестів, які не лише перевіряють базову працездатність, а й дозволяють виявити потенційні слабкі місця та визначити межі ефективності роботи алгоритмів у складних і динамічних умовах, характерних для реального середовища водіння.

Тестування у різних умовах освітлення охопило діапазон від 1 до 10 000 люкс, що відповідає повному спектру можливих сценаріїв — від нічного водіння до прямого сонячного засвічення через лобове скло. Нижня межа в 1 люкс обрана як мінімальний рівень освітленості, при якому ще можливе виконання функцій водіння, тоді як верхня межа відповідає яскравому сонячному світлу в спектрі, характерному для середніх широт. Під час випробувань особливу увагу було приділено забезпеченню рівномірного розподілу освітлення на обличчі водія для уникнення артефактів, які можуть вплинути на коректність виявлення обличчя та аналіз стану очей. Для моделювання різних умов використовувалися штучні джерела світла та фільтри, що дозволило точно відтворити як рівномірне розсіяне освітлення, так і складні сценарії з частковими тінями або різкими перепадами яскравості. Такий підхід забезпечив перевірку стабільності роботи системи при різних контрастах, кольорових температурах та умовах освітлення, характерних для реальних дорожніх ситуацій, у тому числі при зміні погоди чи часу доби. Це дозволило визначити чутливість алгоритму до варіацій

освітленості та встановити порогові значення, за яких система зберігає працездатність із мінімальним рівнем похибок.

Дослідження впливу ракурсу обличчя планується проводити при кутах відхилення 15° , 30° та 45° у вертикальній та горизонтальній площинах. Такі значення були обрані з урахуванням анатомічних особливостей будови людського тіла та характерних рухів голови людини, адже саме в межах цих кутів зазвичай відбуваються природні зміни положення голови водія під час керування транспортним засобом. Врахування цих меж дозволяє максимально наблизити умови тестування до реальних сценаріїв експлуатації системи. Крім того, особливу увагу приділено визначенню максимального кута відхилення, при якому система зможе зберігати здатність достовірно визначати стан очей, що є критичним для забезпечення ефективного моніторингу водія. Такий підхід забезпечить всебічну перевірку функціональності системи в умовах, максимально наближених до реальних змін положення голови водія, з одночасним контролем меж, за яких алгоритми аналізу зберігають свою точність.

Тестування на різних відстанях від 20 до 80 см до камери обране з урахуванням типових параметрів розташування водія відносно приладової панелі в більшості сучасних автомобілів. Нижня межа в 20 см визначена як мінімальна відстань, на якій водій ще може комфортно розміститися та здійснювати всі необхідні дії під час керування транспортним засобом, не порушуючи правил ергономіки та безпеки. Верхня межа в 80 см обрана для врахування особливостей водіїв високого зросту або специфіки конструкції деяких автомобілів, де посадка може відрізнятись залежно від налаштування сидіння чи простору кабіни. Такий діапазон дозволяє оцінити працездатність системи у межах максимально наближених до реальних умов експлуатації, коли відстань між обличчям водія та камерою може змінюватися в залежності від індивідуальних налаштувань сидіння або положення водія під час руху. Це дає змогу встановити оптимальні параметри розташування камери для забезпечення стабільної роботи алгоритму аналізу стану очей та загальної працездатності системи моніторингу.

Для оцінки стабільності роботи системи планується включити тести зі зміною виразу обличчя, які відтворюють природні мімічні рухи, що характерні під час керування транспортним засобом, зокрема посмішки, зітхання та кашель. Вибір саме таких сценаріїв обґрунтований їхньою частотою виникнення у реальних умовах експлуатації системи, а також можливим впливом на точність детекції обличчя та аналіз стану очей. Врахування цих факторів дозволить оцінити здатність розробленого алгоритму коректно працювати при мінливих виразах обличчя, що часто супроводжують природну поведінку водія, та визначити, наскільки зміни міміки можуть ускладнювати або впливати на якість розпізнавання і інтерпретації стану уваги чи втоми. Такий підхід сприятиме всебічній перевірці адаптивності системи до типових життєвих ситуацій, підвищуючи її практичну цінність і надійність у повсякденному використанні.

Критерії оцінки працездатності системи передбачають врахування таких основних показників, як точність детекції, частота помилкових спрацьовувань та час реакції. Вибір саме цих параметрів зумовлений їхньою ключовою роллю у забезпеченні як безпеки руху, так і комфорту водія, оскільки висока точність визначення стану водія дозволяє своєчасно виявляти потенційні ризики, мінімізуючи при цьому кількість хибних тривог, що можуть відволікати користувача або знижувати довіру до системи. Час реакції системи виступає важливим показником оперативності її роботи, що безпосередньо впливає на ефективність своєчасного реагування у критичних ситуаціях. Для кожного тестового сценарію передбачається використання достатнього обсягу вибірки даних, що забезпечує отримання статистично значущих і репрезентативних результатів, які можуть адекватно відображати реальні умови експлуатації та функціональні можливості розробленої системи.

4.2. Розгляд контрольного прикладу

Для демонстрації робочих характеристик системи було обрано контрольний приклад, який включає два взаємодоповнюючих аспекти тестування: перевірку функціональності інтерфейсу користувача та аналіз

точності роботи алгоритмів розпізнавання стану водія. Такий вибір дозволяє комплексно оцінити не лише технічну спроможність алгоритмів, а й зручність та інтуїтивність взаємодії користувача з програмним забезпеченням. При запуску програми користувачеві надається інтерфейс, який містить два основні елементи управління, представлені у вигляді кнопок, що дозволяють активувати основні функції системи. На рисунку 4.1 наведено візуальне відображення цього інтерфейсу, яке ілюструє простоту та логічність його структури, спрямовану на забезпечення швидкого та ефективного доступу до ключових можливостей системи без зайвих ускладнень.



Рисунок 4.1 — Інтерфейс програми

Перша кнопка з позначенням «Вибрати відеофайл» надає користувачеві можливість відкрити стандартний діалог вибору файлів операційної системи, що дозволяє обрати попередньо записаний відеозапис для подальшого аналізу (рисунки 4.2). Цей функціональний елемент інтерфейсу створений з метою забезпечити зручність роботи з різними матеріалами, які можуть містити різноманітні сценарії поведінки водія, включаючи як нормальні, так і потенційно критичні ситуації. Використання відеофайлів дозволяє протестувати систему у

контрольованих умовах, даючи змогу перевірити коректність роботи алгоритмів розпізнавання стану очей та обличчя за заздалегідь визначеними параметрами. Завдяки такому підходу користувач отримує гнучкий інструмент для оцінки функціональності системи на основі реалістичних даних, що сприяє точній ідентифікації можливих недоліків або сильних сторін розробленого алгоритму в різних умовах експлуатації.

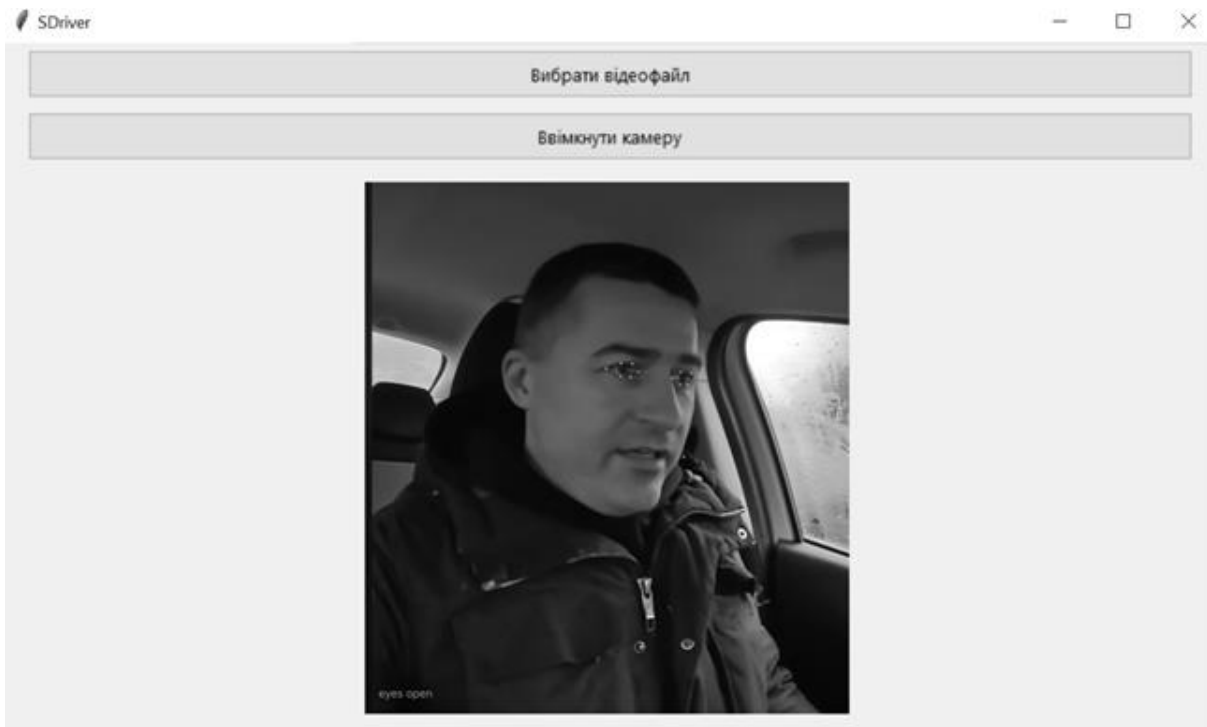


Рисунок 4.2 — Результат аналізу за обраного відеофайлу

Друга кнопка з позначенням «Ввімкнути камеру» відповідає за активацію роботи з вебкамерою або зовнішнім відеопристроєм, що підключений до комп'ютера, для здійснення онлайн-аналізу (рисунок 4.3). Цей функціонал дозволяє в режимі реального часу здійснювати моніторинг стану водія безпосередньо під час керування транспортним засобом або у відповідних тренувальних умовах. Використання живої трансляції відеопотоку забезпечує безперервне отримання актуальних даних про положення обличчя та очей, що дає змогу оперативно визначати ознаки втоми або втрати уваги. Такий режим роботи є важливим для застосування системи в реальних ситуаціях, де своєчасність і точність виявлення критичних змін у поведінці водія мають

вирішальне значення для підвищення рівня безпеки на дорозі. Крім того, реалізація підтримки різних типів відеопристроїв забезпечує гнучкість у виборі апаратного забезпечення, що сприяє ширшому колу можливостей для інтеграції системи у різні платформи та умови експлуатації.

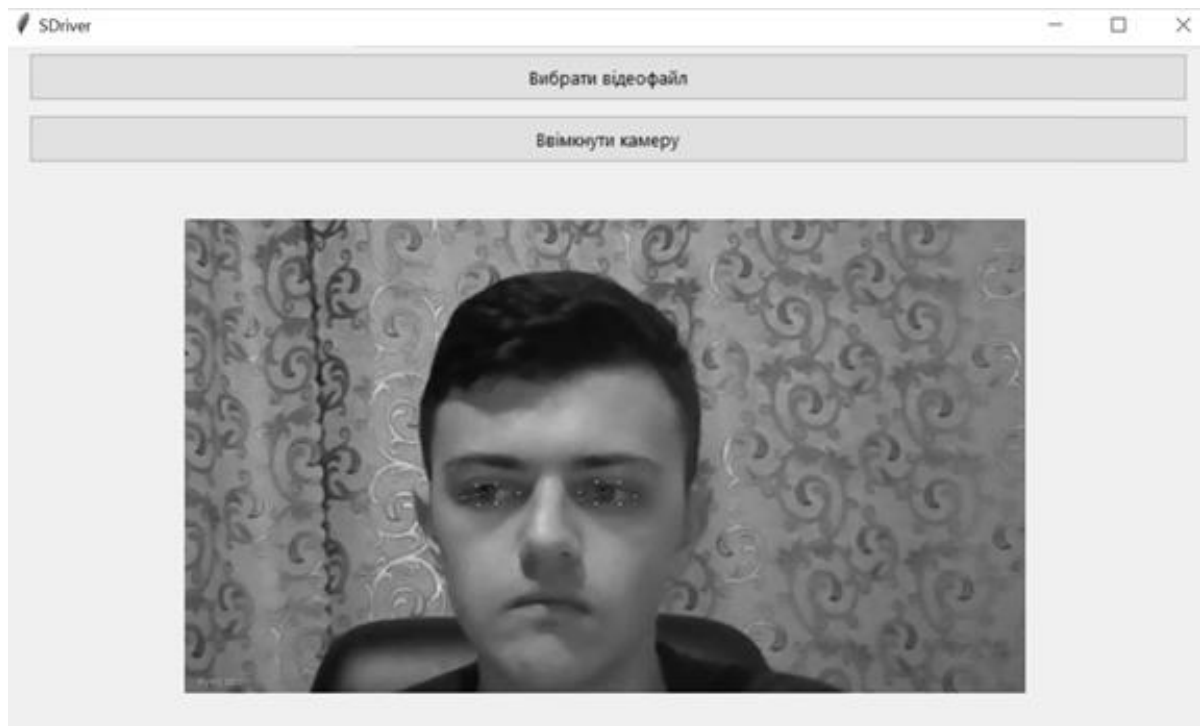


Рисунок 4.3 — Результат аналізу у режимі реального часу

Після вибору джерела даних система автоматично ініціалізує процес аналізу, який включає запуск відповідних алгоритмів обробки відео залежно від обраного режиму роботи. У випадку роботи з відеофайлами відбувається послідовна обробка кожного кадру, що дає змогу з високою точністю відстежувати ключові параметри, такі як стан очей та напрямок погляду водія. Паралельно здійснюється фіксація цих показників для подальшого аналізу та оцінки рівня уваги чи втоми користувача. Якщо алгоритми виявляють ознаки критичного стану, наприклад тривале закриття очей або значне відволікання уваги, система генерує текстове попередження, що виділяється червоним кольором і відображається на екрані. Це візуальне повідомлення має на меті максимально швидко і чітко інформувати користувача про потенційну

небезпеку, сприяючи своєчасному реагуванню та підвищенню рівня безпеки під час керування транспортним засобом.

Режим роботи з камерою передбачає виконання аналізу даних у режимі реального часу, що дає змогу здійснювати безперервний моніторинг стану водія під час керування транспортним засобом. Такий режим забезпечує миттєву візуалізацію проміжних результатів обробки безпосередньо на екрані. Відмінною рисою цього режиму є наявність комплексної системи попередження, яка, крім відображення текстового повідомлення червоного кольору на екрані з чітким формулюванням виявленої небезпеки, додатково генерує звуковий сигнал. Це текстове повідомлення надає користувачеві конкретну інформацію про тип небезпечного стану забезпечуючи чітке та недвозначне інформування. Цей звуковий супровід відіграє важливу роль у приверненні уваги водія до виникнення критичної ситуації, оскільки в умовах дорожнього руху не завжди можливо вчасно зосередитися на візуальних повідомленнях. Такий підхід сприяє негайному інформуванню про потенційну небезпеку, що значно підвищує ефективність попередження та мінімізує ризики втрати концентрації. Приклад результату роботи системи у випадку закритих очей водія, що моделює екстрену ситуацію, наведено на рисунку 4.4.

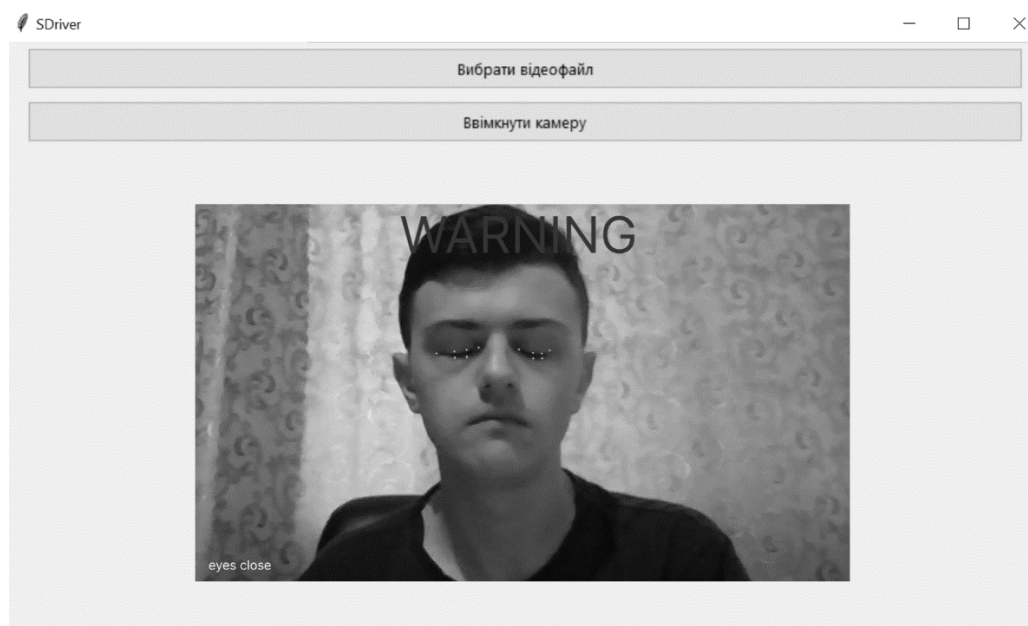


Рисунок 4.4 — Результат аналізу при виявленні закритих очей

Для забезпечення комплексної оцінки ефективності роботи системи моніторингу стану водія було здійснено ґрунтовний аналіз серії спеціально підготовлених тестових відеозаписів, які відображали широкий спектр можливих умов експлуатації у реальному середовищі. Дослідження передбачало ретельний розгляд відеоматеріалів, у яких змінювалися умови освітлення від низького рівня яскравості до інтенсивного сонячного світла, що дозволяло перевірити адаптивність системи до різних сценаріїв використання. Особливу увагу було приділено записам, у яких змінювався ракурс розташування голови водія, оскільки ці параметри мають суттєвий вплив на точність алгоритмів розпізнавання стану очей та напрямку погляду. Крім того, у вибірку було включено відео з наявністю артефактів, що могли виникати через відблиски на лобовому склі, шумові перешкоди або часткові перекриття обличчя, з метою оцінки стійкості системи до таких зовнішніх впливів. Додатково були розглянуті записи з різними мімічними виразами обличчя, такими як посмішка, хмурий вигляд чи різке напруження м'язів обличчя, які можуть потенційно впливати на коректність детекції та створювати виклики для розпізнавання стану очей у реальному часі.

Отримані результати демонструють, що система забезпечує стабільну роботу в широкому діапазоні умов, що відповідає вимогам до систем моніторингу стану водія. Особливо варто відзначити високу точність детекції обличчя навіть при складних умовах освітлення, що підтверджує ефективність обраного алгоритму обробки зображень. Низький рівень помилкових спрацьовувань та мінімальний час реакції свідчать про придатність системи для використання в реальних умовах експлуатації транспортних засобів.

ВИСНОВКИ

У бакалаврській роботі було розроблено комп'ютерну систему безпеки водія на основі нейронних мереж, яка дозволяє ефективно виявляти ознаки втоми та розсіяної уваги у реальному часі. В ході дослідження було проведено комплексний аналіз сучасних методів комп'ютерного зору, що дозволило обґрунтовано обрати архітектуру MediaPipe Face Mesh у поєднанні з алгоритмом EAR як оптимальне рішення для поставленої задачі.

Проведене дослідження включало порівняльний аналіз різних підходів до детекції обличчя та аналізу стану очей, що підтвердило переваги обраної методики щодо точності та продуктивності. Для навчання моделі було сформовано спеціалізований набір даних, що враховує різноманітні умови освітлення, ракурси та індивідуальні особливості водіїв.

Реалізована система демонструє стабільну роботу з частотою 25-30 кадрів на секунду при точності виявлення небезпечних станів на високому рівні. Тестування системи в різних умовах експлуатації підтвердило її ефективність та надійність.

Результати роботи мають значний практичний потенціал для впровадження у транспортній галузі, промисловості та інших сферах, де потрібен контроль стану оператора. Подальший розвиток системи може включати інтеграцію з бортовими комп'ютерами транспортних засобів, розширення функціоналу для виявлення додаткових небезпечних станів та вдосконалення алгоритмів для роботи в екстремальних умовах.

Таким чином, поставлені у роботі завдання були успішно вирішені, що підтверджує досягнення мети дослідження та демонструє перспективність подальшого розвитку розробленого рішення. Отримані результати свідчать про можливість широкого практичного застосування системи для підвищення безпеки дорожнього руху та запобігання аварійним ситуаціям, спричиненим людським фактором.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. LIV Науково—технічна конференція факультету інформаційних технологій та комп'ютерної інженерії (2025). URL: <https://conferences.vntu.edu.ua/index.php/all—fitki/all—fitki—2025/paper/view/23939/19788> (дата звернення 01.05.2025).
2. Штучні нейронні мережі: визначення та сфери застосування URL: <https://elar.khmnu.edu.ua/server/api/core/bitstreams/743537d2—5fae—40ed—93dd—a1216e5ae204/content#page=167> (дата звернення 01.05.2025).
3. Machine Learning vs. Deep Learning vs. Neural Networks URL: <https://www.iotforall.com/machine—learning—vs—deep—learning—vs—neural—networks> (дата звернення 03.05.2025).
4. Get Started with Image Preprocessing and Augmentation for Deep Learning URL: <https://www.mathworks.com/help/images/get—started—with—image—preprocessing—and—augmentation—for—deep—learning.html> (дата звернення 04.05.2025).
5. Vision Transformer (ViT) URL: https://huggingface.co/docs/transformers/model_doc/vit (дата звернення 04.05.2025).
6. A Review of Adversarial Attacks in Computer Vision URL: <https://arxiv.org/abs/2308.07673> (дата звернення 04.05.2025).
7. Згортова нейронна мережа – пояснення CNN та її застосування URL: <https://evergreens.com.ua/ua/articles/cnn.html> (дата звернення 04.05.2025).
8. YOLO Object Detection Explained URL: <https://www.datacamp.com/blog/yolo—object—detection—explained> (дата звернення 07.05.2025).
9. What is black box AI URL: <https://www.ibm.com/think/topics/black-box-ai> (дата звернення 07.05.2025).

10. Self—Driving Cars With Convolutional Neural Networks
URL: <https://neptune.ai/blog/self—driving—cars—with—convolutional—neural—networks—cnn> (дата звернення 05.05.2025).

11. Multimodal Models and Computer Vision: A Deep Dive
URL: <https://blog.roboflow.com/multimodal—models/> (дата звернення 09.05.2025).

12. Neuromorphic Computing: Revolutionizing Computer Vision
URL: <https://www.linkedin.com/pulse/neuromorphic—computing—revolutionizing—computer—vision—adwait—kelkar—kr6pc> (дата звернення 09.05.2025).

13. Edge computer vision: Key industry use cases URL: <https://www.nix.com/edge-computer-vision/> (дата звернення 11.05.2025).

14. RetinaFace: A Face Detection Model for High Resolution Images
URL: <https://medium.com/axinc—ai/retinaface—a—face—detection—model—designed—for—high—resolution—6c3900771a01> (дата звернення 11.05.2025).

15. Comprehensive comparison between vision transformers and convolutional neural networks for face recognition tasks
URL: <https://www.nature.com/articles/s41598—024—72254—w> (дата звернення 11.05.2025).

16. The Transformer Architecture with Hybrid Models URL: <https://medium.com/@bijit211987/the—transformer—architecture—with—hybrid—models—eca885e12056> (дата звернення 12.05.2025).

17. Facial landmark detection with mediapipe URL: <https://ijiemr.org/public/uploads/paper/385061652361570.pdf> (дата звернення 15.05.2025).

18. Facial micro-expression recognition method based on CNN and transformer mixed model URL: <https://www.inderscienceonline.com/doi/10.1504/IJBM.2024.140771> (дата звернення 15.05.2025).

19. MediaPipe Face Mesh URL: https://github.com/google—ai—edge/mediapipe/blob/master/docs/solutions/face_mesh.md (дата звернення 15.05.2025).

20. Exploring Monitoring Systems Data for Driver Distraction and Drowsiness Research URL: <https://www.mdpi.com/1424—8220/20/14/3836> (дата звернення 10.05.2025).

21. Eye Aspect Ratio(EAR) and Drowsiness detector using dlib URL: <https://medium.com/analytics—vidhya/eye—aspect—ratio—ear—and—drowsiness—detector—using—dlib—a0b2c292d706> (дата звернення 11.05.2025).

ДОДАТОК А

Технічне завдання

Міністерство освіти і науки України

Вінницький національний технічний університет

Факультет інформаційних технологій та комп'ютерної інженерії

Кафедра обчислювальної техніки

ЗАТВЕРДЖУЮ

Завідувач кафедри ОТ проф.,

д.т.н. Азаров О.Д.

«21» березня 2025 р.**ТЕХНІЧНЕ ЗАВДАННЯ**

на виконання бакалаврської кваліфікаційної роботи

«Комп'ютерна система безпеки водія з використанням нейронної мережі»

08-54.БКР.034.00.000.ТЗ

Науковий керівник: к.т.н., доц. каф. ОТ

Городецька О. С.

Студент групи 2СП—216

Огірчук О. С.

1 Підстава для виконання бакалаврської кваліфікаційної роботи

1.1 Актуальність розробки системи моніторингу стану водія обумовлена необхідністю зниження аварійності на дорогах, спричиненої втомою або неуважністю водіїв. Сучасні технології комп'ютерного зору та штучного інтелекту дозволяють створювати ефективні рішення для реального часу, які можуть виявляти ознаки засинання, відволікання уваги та інші небезпечні стани. Використання нейронних мереж, таких як MediaPipe Face Mesh, забезпечує високу точність детекції обличчя та аналізу стану очей при мінімальних апаратних вимогах, що робить систему доступною для широкого впровадження;

1.2 Наказ про затвердження теми БКР.

2 Мета БКР і призначення розробки

2.1 Мета роботи — підвищення точності виявлення ознак втоми та зниження уваги водія шляхом удосконалення алгоритмів розпізнавання стану очей і напряму погляду із застосуванням нейронних мереж;

2.2 Призначення розробки — запобігання дорожньо—транспортним пригодам, спричиненим засинанням або неуважністю водія. Окрім практичного застосування, система слугуватиме демонстрацією можливостей сучасних технологій комп'ютерного зору у сфері транспортної безпеки, а також може використовуватись як навчальний приклад реалізації штучного інтелекту в реальних умовах.

3 Вихідні дані для виконання БКР

3.1 Отримання сучасних наукових джерел з комп'ютерного зору та нейромереж для аналізу стану водія, зокрема дослідження архітектур MediaPipe Face Mesh та алгоритмів розрахунку метрики EAR;

3.2 Розробка програмного модуля для детекції обличчя з використанням нейромережі, що включає визначення ключових точок очей та обчислення параметрів їх стану;

3.3 Реалізація системи попередження, яка інтегрує аналіз стану очей та напрямку погляду для формування сигналів у разі виявлення небезпечних станів;

3.4 Тестування роботи системи на різних типах відеофайлів та у реальних умовах з використанням вебкамери.

4 Вимоги до виконання БКР

Робота передбачає розробку системи моніторингу стану водія на базі технологій комп'ютерного зору та штучного інтелекту. Система має детектувати обличчя, аналізувати стан очей та напрямок погляду в реальному часі з частотою ≥ 25 FPS. Обов'язковими є: інтуїтивний інтерфейс, механізми попередження, адаптація до різних умов освітлення, тестування на різних платформах та документальне підтвердження результатів.

5 Етапи БКР та очікувані результати

Етапи роботи та очікувані результати приведено в Таблиці А.1.

Таблиця А.1 — Етапи БКР

№ етапу	Назва етапу	початок	кінець	Очікувані результати
1	Аналіз задачі	03.03.2025	14.03.2025	Огляд теми та існуючих рішень
2	Публікація результатів досліджень	15.03.2025	21.03.2025	Тези доповідей
3	Аналіз методів та засобів розробки комп'ютерної системи	22.03.2025	30.03.2025	Розділ 1
4	Вибір архітектури нейронної мережі та алгоритмів обробки	01.04.2025	08.04.2025	Розділ 2
5	Проектування та реалізація комп'ютерної системи безпеки водія	09.04.2025	25.04.2025	Розділ 3
6	Тестування та налагодження	26.04.2025	02.05.2025	Розділ 4
7	Оформлення пояснювальної записки та презентації	04.05.2025	25.05.2025	ПЗ, презентація

6 Матеріали, що подаються до захисту БКР

До захисту подаються: пояснювальна записка БКР, протокол попереднього захисту БКР на кафедрі, відгук наукового керівника, відгук рецензента, протоколи складання державних екзаменів, анотації до БКР українською та іноземною мовами, довідка про відповідність оформлення БКР діючим вимогам.

7 Порядок контролю виконання та захисту БКР

Виконання етапів графічної та розрахункової документації БКР контролюється науковим керівником згідно зі встановленими термінами. Захист БКР відбувається на засіданні Екзаменаційної комісії, затвердженої наказом ректора.

8 Вимоги до оформлювання та порядок виконання БКР

8.1 При оформлювання БКР використовуються:

- ДСТУ 3008: 2015 «Звіти в сфері науки і техніки. Структура та правила оформлювання»;
- ДСТУ 8302: 2015 «Бібліографічні посилання. Загальні положення та правила складання»;
- міждержавний ГОСТ 2.104—2006 «Єдина система конструкторської документації. Основні написи»;
- методичні вказівки до виконання бакалаврських кваліфікаційних робіт зі спеціальності 123 — «Комп'ютерна інженерія» (освітня програма «Системне програмування»). Кафедра обчислювальної техніки ВНТУ 2024 р.;
- документами на які посилаються у вище вказаних.

8.2 Порядок виконання БКР викладено в «Положення про кваліфікаційні роботи на першому (бакалаврському) рівні вищої освіти СУЯ ВНТУ—03.02.02—П.001.01:21».

Здобувач _____ Огірчук О.С. _____
(підпис) (прізвище, ініціали)

ДОДАТОК В

Графічна частина

**КОМП'ЮТЕРНА СИСТЕМА БЕЗПЕКИ ВОДІЯ З ВИКОРИСТАННЯМ
НЕЙРОННОЇ МЕРЕЖІ**

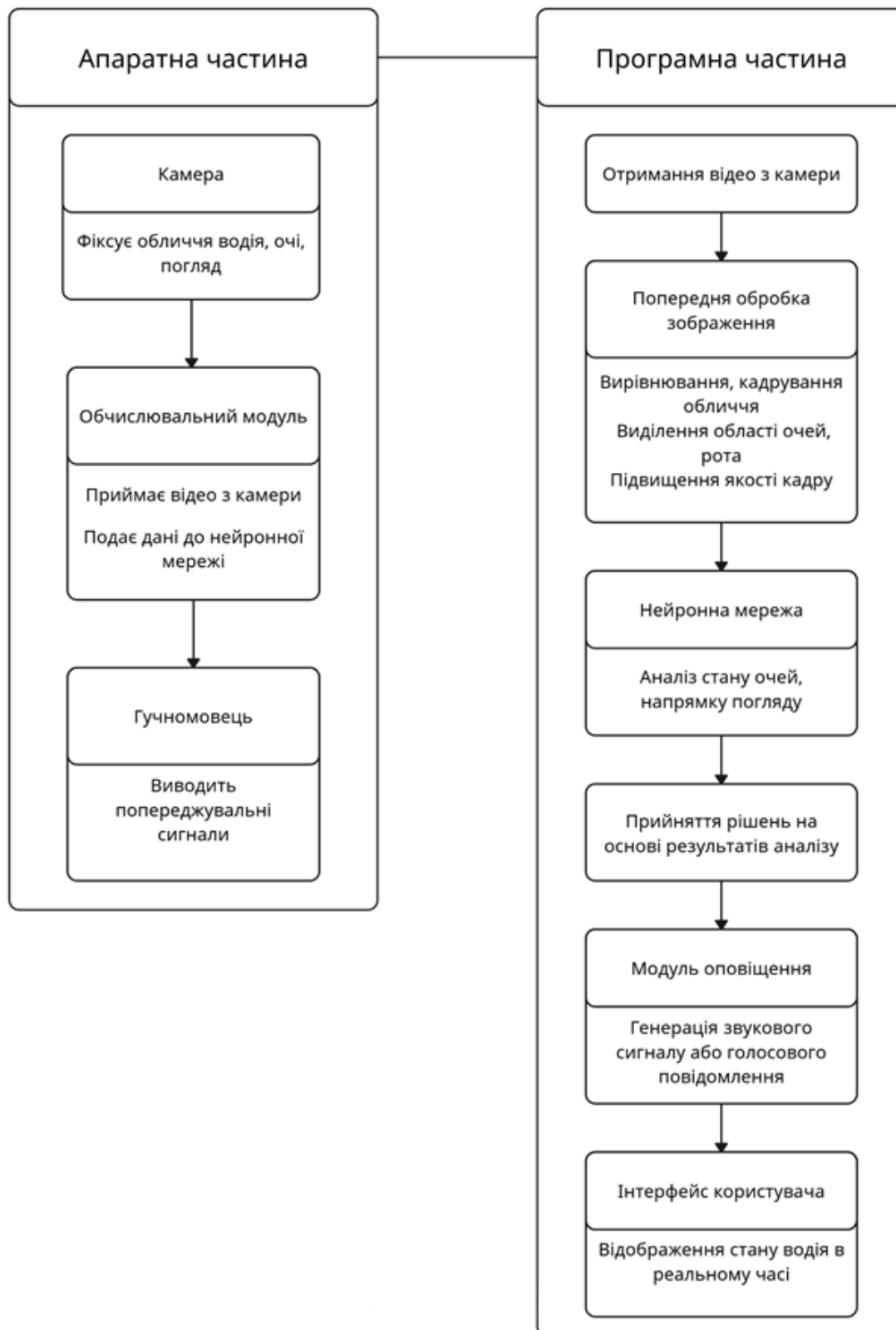


Рисунок В.1 — Структурна схема

ДОДАТОК Г

Ілюстративна частина

**КОМП'ЮТЕРНА СИСТЕМА БЕЗПЕКИ ВОДІЯ З ВИКОРИСТАННЯМ
НЕЙРОННОЇ МЕРЕЖІ**

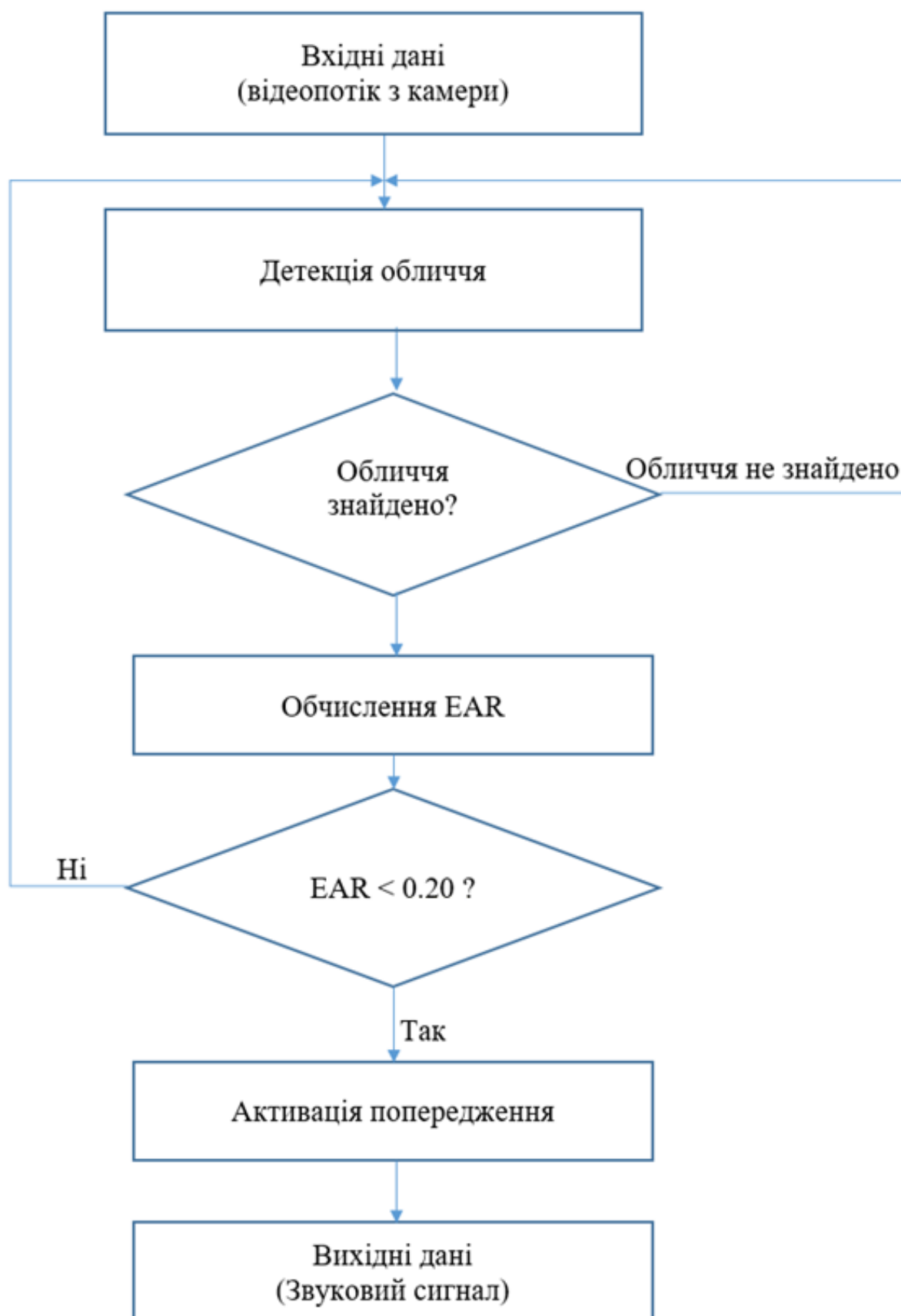


Рисунок Г.1 — Блок—схема алгоритму роботи

ДОДАТОК Д

Лістинг програмного забезпечення

```

import cv2
import numpy as np
import tkinter as tk
from tkinter import ttk
from PIL import Image, ImageTk, ImageDraw
import mediapipe as mp
import math

class DriverMonitoringSystem:
    def __init__(self, root):
        self.root = root
        self.root.title("SDriver")
        self.root.geometry("800x600")
        self.mp_face_mesh = mp.solutions.face_mesh
        self.face_mesh = self.mp_face_mesh.FaceMesh(
            max_num_faces=1,
            refine_landmarks=True,
            min_detection_confidence=0.5,
            min_tracking_confidence=0.5)
        self.camera_active = False
        self.video_capture = None
        self.current_image = None
        self.setup_ui()
        self.eye_closed_threshold = 0.2
        self.eye_landmarks = {
            'left': [33, 160, 158, 133, 153, 144],
            'right': [362, 385, 387, 263, 373, 380]
        }
        self.eye_state = "Норма"
        self.gaze_direction = "Прямо"
        self.ear_history = []
        self.update()

    def setup_ui(self):
        self.style = ttk.Style()
        self.style.configure('TButton', font=('Segoe UI', 10), padding=5)
        self.style.configure('TLabel', font=('Segoe UI', 10))
        self.control_frame = ttk.Frame(self.root)
        self.control_frame.pack(fill='x', padx=10, pady=10)
        self.camera_btn = ttk.Button(
            self.control_frame,
            text="Увімкнути камеру",
            command=self.toggle_camera,

```

```

        width=20
    )
    self.camera_btn.pack(side='left', padx=5)
    self.video_panel = ttk.Label(self.root)
    self.video_panel.pack(fill='both', expand=True, padx=10, pady=5)
    self.status_frame = ttk.Frame(self.root)
    self.status_frame.pack(fill='x', padx=10, pady=5)
    self.eye_label = ttk.Label(
        self.status_frame,
        text="Стан очей: Норма",
        font=('Segoe UI', 12, 'bold'),
        foreground='green'
    )
    self.eye_label.pack(side='left', padx=10)
    self.gaze_label = ttk.Label(
        self.status_frame,
        text="Напрямок погляду: Прямо",
        font=('Segoe UI', 12, 'bold')
    )
    self.gaze_label.pack(side='left', padx=10)
    self.alert_frame = ttk.Frame(self.root)
    self.alert_frame.pack(fill='x', padx=10, pady=5)
    self.alert_label = ttk.Label(
        self.alert_frame,
        text="",
        font=('Segoe UI', 14, 'bold'),
        foreground='red'
    )
    self.alert_label.pack()
def toggle_camera(self):
    if not self.camera_active:
        self.video_capture = cv2.VideoCapture(0)
        if not self.video_capture.isOpened():
            print("Помилка: Не вдалося відкрити камеру")
            return
        self.camera_active = True
        self.camera_btn.config(text="Вимкнути камеру")
    else:
        if self.video_capture:
            self.video_capture.release()
            self.camera_active = False
            self.camera_btn.config(text="Увімкнути камеру")
            self.video_panel.config(image="")
def calculate_ear(self, landmarks, eye_points):
    p1 = landmarks[eye_points[0]]

```

```

p2 = landmarks[eye_points[1]]
p3 = landmarks[eye_points[2]]
p4 = landmarks[eye_points[3]]
p5 = landmarks[eye_points[4]]
p6 = landmarks[eye_points[5]]
a = math.dist((p2.x, p2.y), (p6.x, p6.y))
b = math.dist((p3.x, p3.y), (p5.x, p5.y))
c = math.dist((p1.x, p1.y), (p4.x, p4.y))
ear = (a + b) / (2.0 * c)
return ear

def analyze_gaze(self, landmarks, frame_shape):
    left_eye_center = np.mean([
        landmarks[33].x, landmarks[33].y,
        landmarks[133].x, landmarks[133].y
    ], axis=0)
    right_eye_center = np.mean([
        landmarks[362].x, landmarks[362].y,
        landmarks[263].x, landmarks[263].y
    ], axis=0)
    eye_center = np.mean([left_eye_center, right_eye_center], axis=0)
    if eye_center[0] < 0.4:
        return "Ліворуч"
    elif eye_center[0] > 0.6:
        return "Праворуч"
    elif eye_center[1] > 0.6:
        return "Вниз"
    else:
        return "Прямо"

def process_frame(self, frame):
    # Конвертація кольорів
    rgb_frame = cv2.cvtColor(frame, cv2.COLOR_BGR2RGB)
    results = self.face_mesh.process(rgb_frame)
    if results.multi_face_landmarks:
        for face_landmarks in results.multi_face_landmarks:
            landmarks = face_landmarks.landmark
            left_ear = self.calculate_ear(landmarks, self.eye_landmarks['left'])
            right_ear = self.calculate_ear(landmarks, self.eye_landmarks['right'])
            avg_ear = (left_ear + right_ear) / 2.0
            self.ear_history.append(avg_ear)
            if len(self.ear_history) > 10:
                self.ear_history.pop(0)
            if avg_ear < self.eye_closed_threshold:
                self.eye_state = "Закриті"
                self.eye_label.config(foreground='red')
            else:

```

```

        self.eye_state = "Норма"
        self.eye_label.config(foreground='green')
        self.gaze_direction = self.analyze_gaze(landmarks, frame.shape)
        self.draw_landmarks(frame, face_landmarks)
    self.update_status()
    return frame
def draw_landmarks(self, frame, face_landmarks):
    h, w = frame.shape[:2]
    for landmark in face_landmarks.landmark:
        x, y = int(landmark.x * w), int(landmark.y * h)
        cv2.circle(frame, (x, y), 1, (0, 255, 0), -1)
    cv2.putText(frame, f"Стан: {self.eye_state}", (10, 30),
                cv2.FONT_HERSHEY_SIMPLEX, 0.7, (0, 255, 0), 2)
    cv2.putText(frame, f"Погляд: {self.gaze_direction}", (10, 60),
                cv2.FONT_HERSHEY_SIMPLEX, 0.7, (0, 255, 0), 2)
def update_status(self):
    self.eye_label.config(text=f"Стан очей: {self.eye_state}")
    self.gaze_label.config(text=f"Напрямок погляду: {self.gaze_direction}")
    if self.eye_state == "Закриті":
        self.alert_label.config(text="УВАГА: Водій може засинати!")
    elif self.gaze_direction != "Прямо":
        self.alert_label.config(text=f"УВАГА: Водій дивиться
{self.gaze_direction.lower()}!")
    else:
        self.alert_label.config(text="")
def update(self):
    if self.camera_active and self.video_capture:
        ret, frame = self.video_capture.read()
        if ret:
            processed_frame = self.process_frame(frame)
            img = cv2.cvtColor(processed_frame, cv2.COLOR_BGR2RGB)
            img = Image.fromarray(img)
            imgtk = ImageTk.PhotoImage(image=img)
            self.video_panel.imgtk = imgtk
            self.video_panel.config(image=imgtk)
        self.root.after(15, self.update)
def __del__(self):
    if self.video_capture:
        self.video_capture.release()
    self.face_mesh.close()
if __name__ == "__main__":
    root = tk.Tk()
    app = DriverMonitoringSystem(root)
    root.mainloop()

```