

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет _____ Комп'ютерних наук
(повна назва)
Кафедра _____ Штучного інтелекту
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА
Пояснювальна записка

рівень вищої освіти _____ перший (бакалаврський)

_____ Моделювання та оцінювання втоми людини за допомогою
_____ глибинного аналізу візуальних біомаркерів
(тема)

Виконав:
здобувач _____ четвертого _____ року навчання,
групи _____ ІТШ-22-4

_____ Гліб РЯЗАНЦЕВ
(власне ім'я, прізвище)

Спеціальність 122 Комп'ютерні науки

(код і повна назва спеціальності)

Тип програми _____ освітньо-професійна

Освітня програма _____ Штучний інтелект

(повна назва освітньої програми)

Керівник _____ доц. Олександра ВІТЬКО
(посада, власне ім'я, прізвище)

Допускається до захисту

Завідувач кафедри ШІ

_____ Лариса ЧАЛА
(підпис) (власне ім'я, прізвище)

2026 р.

Харківський національний університет радіоелектроніки

Факультет _____ Комп'ютерних наук _____

Кафедра _____ Штучного інтелекту _____

Рівень вищої освіти _____ перший (бакалаврський) _____

Спеціальність _____ 122 Комп'ютерні науки _____

(код і повна назва)

Тип програми _____ освітньо-професійна _____

Освітня програма _____ Штучний інтелект _____

(повна назва)

ЗАТВЕРДЖУЮ:

Зав. кафедри _____

(підпис)

« _____ » _____ 20 ____ р.

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ

здобувачеві _____ Рязанцеву Глібу Андрійовичу _____

(прізвище, ім'я, по батькові)

1. Тема роботи _____ Моделювання та оцінювання втоми людини за допомогою глибинного аналізу візуальних біомаркерів _____

затверджена наказом університету від 5 червня 2026 р. № 488Ст

2. Термін подання здобувачем роботи до екзаменаційної комісії 8 липня 2026 р.

3. Вихідні дані до роботи Науково-технічні публікації та монографії з питань комп'ютерного зору, аналізу зображень, машинного навчання, дані інтернет-джерел та відомих наукових проєктів у сфері визначення втоми людини, відкриті набори даних із зображеннями облич та ознаками сонливості водіїв, документації до мови програмування Python, бібліотек PyTorch, Torchvision, Pandas, NumPy, Scikit-learn та OpenCV. _____

4. Перелік питань, що потрібно опрацювати в роботі _____

1) Аналіз предметної галузі та постановка задачі _____

2) Розробка підходу до моделювання та оцінювання втоми людини _____

3) Практична реалізація системи _____

КАЛЕНДАРНИЙ ПЛАН

[illegible]

Дата видачі завдання 8 червня 20 26 р.

Здобувач _____
(підпис)

Керівник роботи _____
(підпис)

доц. Олександра ВІТЬКО
(посада, власне ім'я, прізвище)

РЕФЕРАТ

Пояснювальна записка: 62 с., 6 рис., 1 дод., 39 джерел.

АНАЛІЗ ЗОБРАЖЕНЬ, ВІЗУАЛЬНІ БІОМАРКЕРИ, ВТОМА ЛЮДИНИ, ГЛИБИННЕ НАВЧАННЯ, КОМП'ЮТЕРНИЙ ЗОР, CNN, DEEP LEARNING, DETECTION, FACE ANALYSIS, FATIGUE DETECTION.

Об'єкт дослідження – процес оцінювання функціонального стану людини та визначення рівня втоми за допомогою аналізу візуальних ознак обличчя.

Предмет дослідження – методи та засоби автоматизованого виявлення ознак втоми людини із використанням технологій комп'ютерного зору та моделей глибинного навчання.

Мета роботи – дослідити сучасні підходи до моделювання та оцінювання втоми людини на основі аналізу візуальних біомаркерів, обґрунтувати доцільність використання методів глибинного навчання та сформулювати концепцію інтелектуальної системи для автоматизованого визначення стану втоми людини за зображеннями обличчя.

Методи дослідження – аналіз наукових джерел та існуючих програмних рішень у сфері визначення втоми людини, порівняльний аналіз методів комп'ютерного зору та моделей глибинного навчання, аналіз наборів даних для задач класифікації стану людини.

У роботі розглянуто підходи до оцінювання втоми людини, проаналізовано існуючі методи визначення стану оператора на основі візуальних ознак, досліджено можливості використання технологій комп'ютерного зору та глибинного навчання для автоматизованого виявлення втоми, а також запропоновано практичну реалізацію.

ABSTRACT

Bachelor's thesis contains: 62 pp., 6 fig., 1 ann., 39 references.

CNN, COMPUTER VISION, DEEP LEARNING, DETECTION, FACE ANALYSIS, FATIGUE DETECTION, HUMAN FATIGUE, IMAGE ANALYSIS, VISUAL BIOMARKERS.

The object of research is the process of assessing the human functional state and determining the level of fatigue through the analysis of facial visual features.

The subject of research is methods and tools for automated detection of human fatigue indicators using computer vision technologies and deep learning models.

The purpose of the work is to investigate modern approaches to modelling and assessment of human fatigue based on the analysis of visual biomarkers, substantiate the feasibility of applying deep learning methods, and develop the concept of an intelligent system for automated fatigue state recognition from facial images.

Research methods include the analysis of scientific publications and existing software solutions in the field of human fatigue detection, comparative analysis of computer vision techniques and deep learning models, as well as the analysis of datasets intended for human state classification tasks.

The thesis examines modern approaches to human fatigue assessment, analyzes existing methods for determining an operator's state based on visual indicators, investigates the possibilities of applying computer vision and deep learning technologies for automated fatigue detection, and also offers a practical implementation.

ЗМІСТ

Перелік умовних позначень, символів, одиниць, скорочень і термінів	7
Вступ	8
1 Аналіз предметної галузі та постановка задачі	11
1.1 Поняття втоми людини та її вплив на ефективність діяльності	11
1.2 Існуючі методи оцінювання втоми людини	14
1.3 Використання методів комп'ютерного зору та глибинного навчання для визначення втоми людини	18
1.4 Візуальні біомаркери втоми людини та особливості їх аналізу	22
1.5 Аналіз наборів даних та постановка задачі дослідження	25
2 Розробка підходу до моделювання та оцінювання втоми людини	30
2.1 Обґрунтування вибору методів комп'ютерного зору та глибинного навчання	30
2.2 Розробка підходу до практичної реалізації	33
3 Практична реалізація системи	36
3.1 Формування та підготовка набору даних для навчання моделі	36
3.2 Формування інтегрованого набору даних для навчання моделі	38
3.3 Підготовка даних та організація процесу навчання	41
3.4 Розробка та навчання нейромережевої моделі оцінювання втоми людини	44
3.5 Аналіз результатів експериментального дослідження	47
Висновки	54
Перелік джерел посилання	57
Додаток А Відомість кваліфікаційної роботи	62

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ

ІІІ – штучний інтелект;

AI – Artificial Intelligence – штучний інтелект;

CNN – Convolutional Neural Networks – згорткова нейронна мережа;

CSV – Comma-Separated Values – формат файлів для зберігання табличних даних;

EAR – Eye Aspect Ratio – коефіцієнт співвідношення геометричних параметрів ока для оцінювання ступеня його відкриття;

LSTM – Long Short-Term Memory – рекурентна нейронна мережа з механізмом довготривалої пам'яті;

MAR – Mouth Aspect Ratio – коефіцієнт співвідношення геометричних параметрів рота для оцінювання ступеня його відкриття;

PERCLOS – Percentage of Eye Closure – показник відсотка часу, протягом якого очі залишаються закритими або частково закритими;

RGB – Red, Green, Blue – колірна модель подання зображень на основі червоного, зеленого та синього каналів.

ВСТУП

У сучасних умовах стрімкого розвитку інформаційних технологій, автоматизованих систем керування та інтелектуальних систем підтримки прийняття рішень особливого значення набувають задачі моніторингу функціонального стану людини. Однією з найбільш актуальних проблем у даній сфері є проблема своєчасного виявлення втоми людини, оскільки зниження концентрації уваги, сонливість та погіршення когнітивних функцій можуть призводити до значного зростання кількості помилок, аварійних ситуацій та зниження ефективності діяльності в різних сферах людської діяльності.

Особливо гостро проблема втоми проявляється у сферах, пов'язаних із підвищеним рівнем відповідальності та необхідністю тривалого утримання уваги. До таких сфер належать транспортна галузь, системи диспетчерського керування, промислове виробництво, медична сфера, військова діяльність та операторська робота. Значна частина дорожньо-транспортних пригод, виробничих аварій та критичних помилок операторів пов'язана саме зі зниженням рівня пильності та появою ознак втоми. У зв'язку з цим виникає потреба у створенні ефективних систем автоматизованого контролю функціонального стану людини, здатних оперативно визначати появу небезпечних ознак втоми та сонливості.

Традиційні методи оцінювання стану людини часто базуються на використанні спеціалізованих сенсорів, фізіологічних датчиків або експертного аналізу. Подібні підходи можуть забезпечувати достатньо високу точність, однак вони мають ряд суттєвих недоліків: потребують дорогого обладнання, є незручними для тривалого використання та не завжди придатні для повсякденного застосування в реальних умовах. Унаслідок цього все більшої популярності набувають безконтактні методи оцінювання втоми людини, засновані на аналізі зображень обличчя та поведінкових характеристик людини.

Сучасний розвиток технологій комп'ютерного зору та глибинного навчання створив широкі можливості для автоматизованого аналізу візуальних біомаркерів втоми. До найбільш поширених візуальних ознак, що можуть свідчити про зниження рівня концентрації та появу сонливості, належать частота моргання, тривале закриття очей, позіхання, зміна положення голови, напрямок погляду та інші особливості міміки. Використання моделей глибинного навчання дозволяє автоматизувати процес виявлення подібних ознак та забезпечити достатньо високу точність класифікації стану людини навіть у складних умовах спостереження.

На даний момент існує значна кількість досліджень та програмних рішень, присвячених задачам *fatigue detection* та *drowsiness detection*. Однак більшість існуючих систем орієнтована на використання окремих наборів даних або спеціалізованих умов зйомки, що обмежує можливість їх ефективного використання в реальних умовах. Крім того, багато існуючих рішень мають обмежену стійкість до змін освітлення, положення обличчя та індивідуальних особливостей користувачів. У зв'язку з цим актуальним є дослідження підходів до формування більш універсальних моделей оцінювання втоми людини на основі комбінованих наборів даних, сформованих із декількох відкритих джерел.

У рамках даної роботи досліджується предметна галузь, пов'язана з автоматизованим визначенням рівня втоми людини за допомогою аналізу візуальних біомаркерів обличчя. Особлива увага приділяється аналізу сучасних методів комп'ютерного зору та моделей глибинного навчання, що використовуються для задач класифікації стану людини. Також розглядаються існуючі набори даних для задач *fatigue detection* та підходи до їх об'єднання з метою підвищення узагальнювальної здатності майбутньої моделі.

Метою роботи є дослідження сучасних підходів до моделювання та оцінювання втоми людини на основі аналізу візуальних біомаркерів, аналіз існуючих методів і засобів розв'язання даної задачі та формування

концепції інтелектуальної системи для автоматизованого визначення стану втоми людини із використанням технологій глибинного навчання.

Практичне значення роботи полягає у можливості використання результатів дослідження під час розробки інтелектуальних систем моніторингу стану людини у транспортній сфері, системах операторського контролю, промислових комплексах, а також у сферах, де своєчасне виявлення втоми є критично важливим для забезпечення безпеки та ефективності діяльності.

1 АНАЛІЗ ПРЕДМЕТНОЇ ГАЛУЗІ ТА ПОСТАНОВКА ЗАДАЧІ

1.1 Поняття втоми людини та її вплив на ефективність діяльності

Втома людини є складним психофізіологічним станом організму, який виникає внаслідок тривалого фізичного або розумового навантаження та супроводжується тимчасовим зниженням працездатності, швидкості реакції, концентрації уваги та загальної ефективності діяльності. Виникнення втоми є природною реакцією організму на перевантаження та виконує захисну функцію, сигналізуючи про необхідність відпочинку та відновлення ресурсів організму.

У сучасних умовах розвитку суспільства проблема втоми набуває особливої актуальності через зростання інтенсивності праці, збільшення обсягів інформації, необхідність тривалого підтримання концентрації уваги та широкого використання автоматизованих систем керування. Втома може виникати у представників різних професійних сфер, однак особливо небезпечними її прояви є у діяльності операторів транспортних систем, диспетчерів, медичних працівників, військовослужбовців та інших спеціалістів, робота яких пов'язана з підвищеним рівнем відповідальності та необхідністю швидкого прийняття рішень, таких як пілоти.

У науковій літературі виділяють декілька основних видів втоми, зокрема фізичну, розумову, емоційну та сенсорну втому [1]. Фізична втома пов'язана зі зниженням функціональних можливостей м'язової системи внаслідок тривалого фізичного навантаження. Розумова втома характеризується погіршенням когнітивних функцій, зниженням концентрації уваги та сповільненням процесів обробки інформації. Емоційна втома виникає внаслідок тривалого психологічного напруження та стресу, а сенсорна втома пов'язана з перевантаженням органів чуття при тривалому сприйнятті однотипної інформації.

Особливе місце серед різновидів втоми займає сонливість, яка характеризується насамперед зниженням рівня пильності та підвищеною потребою у сні. Сонливість є одним із найбільш небезпечних факторів у процесі керування транспортними засобами та операторської діяльності, оскільки вона безпосередньо впливає на швидкість реакції людини, здатність до прийняття рішень та контроль над ситуацією. Згідно з дослідженнями, стан сонливості може призводити до ефекту, подібного до впливу алкогольного сп'яніння, знижуючи точність виконання завдань та погіршуючи координацію рухів [2].

Основними причинами виникнення втоми є недостатня тривалість сну, тривала монотонна діяльність, високий рівень психоемоційного навантаження, порушення режиму праці та відпочинку, а також тривале перебування у стресових умовах. Значний вплив на розвиток втоми мають також зовнішні фактори, серед яких недостатнє освітлення, підвищений рівень шуму, несприятливі температурні умови та тривала робота за комп'ютером.

Фізіологічні прояви втоми можуть мати різноманітний характер та проявлятися як на фізичному, так і на поведінковому рівнях. До найбільш поширених ознак втоми належать зниження концентрації уваги, сповільнення реакції, часте моргання, позіхання, погіршення координації рухів, зниження точності виконання дій, а також мікрозасинання. Крім того, у стані втоми у людини можуть спостерігатися зміни міміки, руху очей, положення голови та інших поведінкових характеристик, які можуть бути використані як візуальні біомаркери під час автоматизованого аналізу стану людини.

Втома суттєво впливає на когнітивні функції людини. У стані втоми знижується швидкість обробки інформації, погіршується здатність до концентрації уваги, зростає кількість помилок та збільшується час реакції на зовнішні подразники. Особливо небезпечним є те, що людина не завжди

здатна об'єктивно оцінити власний рівень втоми, що може призводити до прийняття помилкових рішень та виникнення аварійних ситуацій.

Однією з найбільш критичних сфер впливу втоми є транспортна галузь. За даними міжнародних досліджень, значна частина дорожньо-транспортних пригод пов'язана саме зі зниженням уваги та засинанням водіїв за кермом [3]. Особливу небезпеку становлять тривалі поїздки, нічний час доби та монотонні умови руху, які сприяють зниженню рівня пильності людини. Унаслідок цього задача своєчасного виявлення ознак втоми водія є важливою складовою забезпечення безпеки дорожнього руху.

Проблема втоми також є актуальною для операторів складних технічних систем та працівників промислових підприємств. Умови тривалого спостереження за інформаційними панелями, виконання однотипних дій та високий рівень відповідальності можуть призводити до зниження продуктивності праці та виникнення критичних помилок. У сучасних автоматизованих системах керування все більшого значення набувають засоби моніторингу функціонального стану оператора, здатні своєчасно виявляти ознаки перевтоми та попереджати виникнення небезпечних ситуацій.

Традиційні методи оцінювання втоми людини часто базуються на використанні фізіологічних датчиків, електроенцефалографії, аналізу серцевого ритму або експертного спостереження [4]. Подібні методи можуть забезпечувати високу точність оцінювання стану людини, однак вони мають ряд суттєвих обмежень, серед яких складність використання, необхідність спеціалізованого обладнання та низька зручність для повсякденного застосування [5]. У зв'язку з цим останніми роками активно розвиваються безконтактні методи оцінювання втоми, засновані на використанні технологій комп'ютерного зору та аналізу візуальних ознак поведінки людини.

Розвиток методів машинного та глибинного навчання створив широкі можливості для автоматизації процесу визначення втоми людини.

Використання моделей комп'ютерного зору дозволяє аналізувати зображення обличчя людини у режимі реального часу та виявляти характерні ознаки сонливості й зниження рівня концентрації. Завдяки цьому системи автоматизованого моніторингу втоми можуть бути інтегровані у транспортні засоби, системи операторського контролю, виробничі комплекси та інші інформаційні системи, де своєчасне виявлення ознак втоми має критичне значення для забезпечення безпеки та ефективності діяльності.

1.2 Існуючі методи оцінювання втоми людини

Проблема своєчасного визначення втоми людини протягом тривалого часу залишається актуальною у багатьох сферах діяльності, пов'язаних із підвищеним рівнем відповідальності та необхідністю постійного контролю функціонального стану оператора. У зв'язку з цим у науковій літературі та прикладних дослідженнях запропоновано значну кількість методів і підходів до оцінювання рівня втоми, які відрізняються за принципом роботи, точністю, складністю реалізації та сферою застосування [6].

Існуючі методи оцінювання втоми людини умовно можна поділити на декілька основних груп: фізіологічні методи, поведінкові методи, методи на основі аналізу продуктивності діяльності та комбіновані підходи. Кожна із зазначених груп має власні переваги та недоліки, а також характеризується різним рівнем придатності для використання у реальних умовах експлуатації.

Фізіологічні методи базуються на аналізі біологічних сигналів організму людини та є одними з найбільш точних способів визначення рівня втоми. До основних фізіологічних параметрів, що використовуються для оцінювання функціонального стану людини, належать електроенцефалограма, електрокардіограма, частота серцевих скорочень,

електроміографія, температура тіла та інші показники фізіологічної активності [7].

Одним із найбільш поширених підходів у межах фізіологічних методів є використання електроенцефалографії. Електроенцефалограма дозволяє реєструвати електричну активність головного мозку та аналізувати зміни мозкових ритмів, які виникають у стані втоми або сонливості. Дослідження показують, що при виникненні втоми спостерігається зміна активності альфа-, бета- та тета-ритмів мозку, що може використовуватися для автоматизованого визначення рівня пильності людини [8].

Незважаючи на високу точність, методи на основі електроенцефалографії мають ряд суттєвих недоліків. Використання EEG-систем потребує спеціалізованого обладнання, контакту датчиків зі шкірою людини та створення комфортних умов для проведення вимірювань. Крім того, подібні системи є складними у використанні в реальних умовах експлуатації, особливо у транспортній сфері або під час тривалої роботи оператора.

Ще одним поширеним напрямом є використання методів аналізу серцевого ритму та варіабельності серцевих скорочень. Дослідження показують, що стан втоми супроводжується змінами частоти серцевих скорочень та показників варіабельності серцевого ритму [9]. Аналіз подібних параметрів дозволяє оцінювати рівень навантаження на організм та прогнозувати виникнення сонливості. Однак для реалізації подібних методів також необхідне використання контактних сенсорів, що обмежує зручність їх практичного застосування.

Окрему групу складають методи, засновані на аналізі поведінкових характеристик людини. На відміну від фізіологічних підходів, подібні методи не потребують безпосереднього контакту із користувачем та можуть бути реалізовані за допомогою звичайних відеокамер або систем комп'ютерного зору. Поведінкові методи ґрунтуються на тому, що стан

втомі супроводжується характерними змінами поведінки людини, зокрема рухів очей, положення голови, міміки та частоти позіхання.

Одним із найбільш поширених підходів у межах поведінкових методів є аналіз стану очей людини. Дослідження показують, що під час втоми збільшується тривалість закриття очей, змінюється частота моргання та зростає кількість мікрозасинань. На основі цих ознак було запропоновано значну кількість алгоритмів оцінювання рівня сонливості водіїв та операторів [10].

Для автоматизованого аналізу стану очей часто використовуються методи визначення *facial landmarks* та розрахунок спеціалізованих коефіцієнтів, серед яких одним із найбільш відомих є коефіцієнт *Eye Aspect Ratio (EAR)*. Значення *EAR* дозволяє оцінювати ступінь відкриття очей на основі геометричних характеристик обличчя та використовується у багатьох сучасних системах *fatigue detection* [11].

Крім аналізу очей, у поведінкових системах оцінювання втоми широко застосовується аналіз позіхання та рухів голови. Часте позіхання, нахили голови та тривале утримання погляду в одній точці можуть свідчити про зниження рівня концентрації та появу сонливості. Використання декількох ознак одночасно дозволяє підвищити точність визначення стану людини та зменшити кількість помилкових спрацьовувань [12].

Окрему групу складають методи оцінювання втоми на основі аналізу продуктивності діяльності людини. Подібні системи оцінюють швидкість реакції оператора, точність виконання дій, стабільність керування транспортним засобом або інші характеристики виконання робочих завдань. Наприклад, у транспортній сфері можуть аналізуватися траєкторія руху автомобіля, частота корекції керма та характер прискорення транспортного засобу [13].

Методи на основі аналізу продуктивності мають перевагу у вигляді відсутності необхідності прямого спостереження за людиною, однак їх ефективність значною мірою залежить від зовнішніх умов та специфіки

виконуваної діяльності. Крім того, подібні методи часто дозволяють виявити втому лише на пізніх етапах, коли її вплив на діяльність людини вже є значним.

У сучасних дослідженнях все більшого поширення набувають комбіновані системи оцінювання втоми, які поєднують декілька різних типів методів одночасно. Наприклад, у межах однієї системи можуть одночасно використовуватися аналіз руху очей, оцінювання серцевого ритму та моніторинг поведінки оператора. Подібний підхід дозволяє підвищити надійність визначення стану людини та компенсувати недоліки окремих методів.

Останніми роками особливого розвитку набули методи оцінювання втоми, засновані на використанні технологій комп'ютерного зору та глибинного навчання. Використання згорткових нейронних мереж дозволяє автоматично виділяти інформативні ознаки із зображень обличчя людини та виконувати класифікацію стану користувача без необхідності ручного визначення параметрів. Подібні підходи демонструють високу точність навіть у складних умовах освітлення та при наявності часткових перекриттів обличчя [14].

Системи на основі комп'ютерного зору мають ряд суттєвих переваг. Вони є безконтактними, можуть працювати у режимі реального часу, не потребують використання дорогих фізіологічних сенсорів та можуть бути інтегровані у звичайні відеокамери або мобільні пристрої. Саме тому подібні системи розглядаються як один із найбільш перспективних напрямів розвитку технологій моніторингу функціонального стану людини [15].

Попри значний прогрес у сфері автоматизованого визначення втоми, проблема створення універсальних та стійких систем залишається актуальною. Значний вплив на точність роботи систем мають умови освітлення, положення обличчя, індивідуальні особливості користувачів, якість відеопотоку та інші зовнішні фактори [16]. У зв'язку з цим важливим напрямом сучасних досліджень є вдосконалення алгоритмів глибинного

навчання, а також використання комбінованих наборів даних для підвищення узагальнювальної здатності моделей.

1.3 Використання методів комп'ютерного зору та глибинного навчання для визначення втоми людини

Розвиток інформаційних технологій, збільшення обчислювальних можливостей сучасних комп'ютерних систем та активне впровадження методів штучного інтелекту сприяли значному поширенню технологій комп'ютерного зору у різних сферах діяльності людини. На даний момент методи computer vision активно використовуються у системах відеоспостереження, біометричної ідентифікації, медичної діагностики, автономного транспорту, робототехніки та інтелектуальних системах моніторингу функціонального стану людини [17].

Комп'ютерний зір являє собою напрям штучного інтелекту, основною задачею якого є автоматизований аналіз та інтерпретація зображень або відеоданих. Основною метою систем комп'ютерного зору є отримання змістовної інформації із візуальних даних та прийняття рішень на основі результатів їх аналізу. У задачах визначення втоми людини методи комп'ютерного зору використовуються для аналізу зображень обличчя, відстеження руху очей, оцінювання міміки, визначення положення голови та виявлення інших поведінкових ознак, що можуть свідчити про появу сонливості або зниження концентрації уваги.

Однією з ключових переваг використання комп'ютерного зору у системах оцінювання втоми є можливість реалізації безконтактного моніторингу стану людини. На відміну від фізіологічних методів, заснованих на використанні сенсорів або електродів, системи комп'ютерного зору можуть працювати за допомогою звичайних камер, не створюючи додаткового дискомфорту для користувача [18]. Це робить подібні системи більш придатними для практичного використання у

транспортній сфері, виробничих комплексах та системах операторського контролю.

Перші системи автоматизованого визначення втоми людини на основі комп'ютерного зору базувалися переважно на використанні класичних алгоритмів обробки зображень. Подібні системи використовували методи виділення контурів, аналізу яскравості, геометричних характеристик обличчя та порогових алгоритмів для виявлення очей або рота людини. Однак ефективність таких підходів значною мірою залежала від умов освітлення, якості відеопотоку та положення обличчя відносно камери.

Одним із важливих етапів розвитку систем комп'ютерного зору стало використання алгоритмів детекції обличчя. Одним із найбільш відомих методів виявлення обличчя є алгоритм Viola-Jones, запропонований у 2001 році. Даний підхід базується на використанні каскадів Haar-ознак та дозволяє виконувати виявлення обличчя у режимі реального часу [19]. Завдяки своїй ефективності алгоритм Viola-Jones тривалий час використовувався у системах моніторингу водіїв та автоматизованого визначення втоми.

Подальший розвиток технологій комп'ютерного зору привів до поширення методів facial landmark detection. Подібні алгоритми дозволяють визначати ключові точки обличчя людини, зокрема контури очей, рота, носа та інших елементів. На основі координат facial landmarks можуть розраховуватися спеціалізовані параметри для аналізу стану людини, серед яких найбільш поширеними є Eye Aspect Ratio та Mouth Aspect Ratio [20].

Використання facial landmarks дозволило значно підвищити точність аналізу поведінкових ознак втоми. Наприклад, на основі аналізу геометричних співвідношень між ключовими точками очей можливо визначати частоту моргання, тривалість закриття очей та засинання. Аналогічно, аналіз області рота дозволяє виявляти позіхання, яке є одним із характерних проявів сонливості.

Значний вплив на розвиток систем комп'ютерного зору здійснив розвиток методів машинного навчання та нейронних мереж. На відміну від класичних алгоритмів обробки зображень, методи машинного навчання дозволяють автоматично виявляти закономірності у даних та формувати моделі класифікації на основі прикладів [21]. У задачах визначення втоми це дозволило перейти від жорстко заданих правил до більш гнучких та адаптивних систем аналізу стану людини.

Особливого поширення у сфері комп'ютерного зору набули згорткові нейронні мережі. Convolutional Neural Networks являють собою клас глибинних нейронних мереж, спеціально адаптованих для аналізу зображень. Основною перевагою CNN є здатність автоматично виділяти інформативні ознаки без необхідності ручного визначення характеристик зображення [22].

Застосування CNN у задачах fatigue detection дозволило значно підвищити точність класифікації стану людини. Подібні моделі можуть аналізувати повне зображення обличчя або окремі його області та визначати наявність ознак сонливості навіть у складних умовах спостереження. Крім того, глибинні нейронні мережі демонструють високу стійкість до змін освітлення, масштабу зображення та положення голови користувача.

У сучасних системах визначення втоми широко використовуються архітектури глибинних нейронних мереж, такі як AlexNet, VGGNet, ResNet, MobileNet та EfficientNet [23]. Кожна з цих архітектур має власні особливості та може використовуватися залежно від вимог до швидкодії, точності та обчислювальної складності системи. Наприклад, архітектури MobileNet та EfficientNet часто використовуються у системах реального часу через відносно невеликий розмір моделей та високу ефективність роботи.

Одним із важливих напрямів розвитку сучасних систем комп'ютерного зору є використання transfer learning. Даний підхід передбачає використання попередньо навчених моделей, які були

натреновані на великих наборах зображень, для розв'язання нових задач класифікації. Використання transfer learning дозволяє зменшити потребу у великих обсягах навчальних даних та скоротити час навчання нейронної мережі [24].

Для задач визначення втоми людини також активно використовуються бібліотеки та програмні засоби комп'ютерного зору. Одним із найбільш поширених інструментів є OpenCV – бібліотека для обробки зображень та відео, що містить значну кількість алгоритмів детекції об'єктів, обробки кадрів та аналізу зображень [25]. Крім того, широкого поширення набули бібліотеки TensorFlow та PyTorch, які використовуються для побудови та навчання моделей глибинного навчання [26].

Останніми роками значної популярності набули системи на основі MediaPipe Face Mesh, що дозволяють виконувати високоточне відстеження ключових точок обличчя у режимі реального часу [27]. Використання подібних технологій дає можливість більш точно аналізувати рухи очей, положення голови та інші поведінкові ознаки втоми.

Попри значні досягнення у сфері комп'ютерного зору та глибинного навчання, проблема створення універсальних систем визначення втоми людини залишається актуальною. Ефективність роботи моделей може суттєво залежати від якості наборів даних, освітлення, ракурсу зйомки, наявності окулярів або часткових перекриттів обличчя. У зв'язку з цим сучасні дослідження спрямовані на підвищення узагальнювальної здатності моделей та створення більш стійких систем автоматизованого аналізу стану людини.

Таким чином, використання методів комп'ютерного зору та глибинного навчання є одним із найбільш перспективних напрямів розвитку систем автоматизованого визначення втоми людини. Подібні технології дозволяють реалізувати безконтактний моніторинг стану користувача у режимі реального часу та забезпечують високу точність аналізу візуальних біомаркерів втоми.

1.4 Візуальні біомаркери втоми людини та особливості їх аналізу

Одним із ключових напрямів сучасних досліджень у сфері автоматизованого визначення функціонального стану людини є аналіз візуальних біомаркерів втоми. Під візуальними біомаркерами розуміють сукупність зовнішніх поведінкових та фізіологічних ознак, які можуть бути зафіксовані за допомогою відеокамер або систем комп'ютерного зору та використовуються для оцінювання рівня сонливості, концентрації уваги або загального психофізіологічного стану людини.

На відміну від контактних фізіологічних методів, аналіз візуальних біомаркерів не потребує використання спеціалізованих сенсорів або безпосереднього контакту із користувачем. Завдяки цьому подібні підходи є більш зручними для практичного застосування у транспортних системах, операторських комплексах та системах моніторингу людини у режимі реального часу.

Дослідження показують, що стан втоми супроводжується рядом характерних змін поведінки людини, які можуть бути зафіксовані за допомогою аналізу зображень обличчя. До найбільш поширених візуальних ознак втоми належать частота моргання, тривалість закриття очей, позіхання, зміна напрямку погляду, положення голови, особливості міміки та мікроруки обличчя [18].

Одним із найбільш інформативних візуальних біомаркерів є стан очей людини. Дослідження у сфері *fatigue detection* демонструють, що при виникненні почуття сонливості спостерігається збільшення тривалості закриття очей, зниження швидкості реакції повік та зростання частоти моргання [10]. У деяких випадках можуть виникати епізоди мікрозасинання – короткочасні епізоди втрати уваги, під час яких очі людини залишаються закритими протягом декількох секунд.

Для оцінювання стану очей у сучасних системах комп'ютерного зору часто використовується показник *Eye Aspect Ratio*. Даний параметр

розраховується на основі співвідношення відстаней між ключовими точками очей та дозволяє оцінювати ступінь їх відкриття. При закриванні очей значення EAR зменшується, що дозволяє автоматизовано визначати моменти моргання або тривалого закриття очей.

Однією з найбільш відомих методик оцінювання сонливості є PERCLOS – показник, що характеризує відсоток часу, протягом якого очі людини залишаються закритими. Дослідження показують, що PERCLOS є одним із найбільш надійних індикаторів зниження рівня пильності та використовується у системах моніторингу стану водіїв [28].

Важливим візуальним біомаркером втоми також є позіхання. Збільшення частоти позіхання часто пов'язується зі зниженням рівня активності центральної нервової системи та появою сонливості. Для автоматизованого виявлення позіхання у системах комп'ютерного зору використовується аналіз області рота та визначення його відкриття.

Для оцінювання відкриття рота часто використовується показник Mouth Aspect Ratio, який аналогічно до EAR розраховується на основі геометричних співвідношень між ключовими точками області рота. Під час позіхання значення MAR значно збільшується, що дозволяє використовувати даний параметр для автоматизованого виявлення ознак сонливості.

Окрім аналізу очей та рота, значна увага приділяється аналізу напрямку погляду людини. Зниження концентрації уваги часто супроводжується тривалим фокусуванням погляду в одній точці або відведенням погляду від області спостереження. У транспортній сфері подібні зміни можуть свідчити про втрату контролю над дорожньою ситуацією та підвищення ризику виникнення аварійної ситуації.

Для визначення напрямку погляду використовуються методи eye tracking та gaze estimation. Подібні системи аналізують положення зіниць, рухи очей та координати facial landmarks для оцінювання напрямку погляду

користувача. Використання gaze tracking дозволяє реалізувати системи моніторингу рівня уваги водія або оператора у режимі реального часу.

Ще одним важливим поведінковим біомаркером є положення голови людини. У стані втоми можуть спостерігатися нахили голови, повільні рухи шиї або тривале утримання неприродного положення. Подібні зміни часто використовуються як додаткові ознаки сонливості у комбінованих системах fatigue detection.

Для аналізу положення голови застосовуються методи head pose estimation, які дозволяють оцінювати просторову орієнтацію обличчя відносно камери. На основі аналізу кутів нахилу голови можливо виявляти характерні зміни поведінки користувача, пов'язані зі зниженням рівня пильності.

У сучасних дослідженнях також розглядаються більш складні візуальні ознаки, пов'язані з аналізом міміки, мікровиразів обличчя та динаміки руху лицевих м'язів. Дослідження демонструють, що у стані втоми можуть змінюватися характеристики м'язової активності обличчя, інтенсивність емоційних реакцій та загальна виразність міміки [12]. Однак автоматизований аналіз подібних ознак потребує використання складних моделей машинного навчання та значних обсягів навчальних даних.

Ефективність систем аналізу візуальних біомаркерів значною мірою залежить від якості відеопотоку та умов спостереження. На точність визначення стану людини можуть впливати недостатнє освітлення, низька роздільна здатність камери, часткове перекриття обличчя, наявність окулярів, бороди або інших елементів зовнішності. Крім того, значний вплив мають індивідуальні особливості користувачів, зокрема форма обличчя, особливості міміки та поведінки.

Для підвищення точності роботи сучасні системи fatigue detection часто використовують комбінований аналіз декількох біомаркерів одночасно. Наприклад, система може одночасно оцінювати стан очей, частоту позіхання, положення голови та напрямок погляду. Подібний підхід

дозволяє зменшити ймовірність помилкових спрацьовувань та підвищити стійкість системи до зовнішніх факторів.

Значний розвиток методів глибинного навчання сприяв переходу від ручного виділення ознак до автоматизованого аналізу візуальних біомаркерів. Сучасні згорткові нейронні мережі здатні самостійно виділяти інформативні характеристики із зображень обличчя та виконувати класифікацію стану людини без необхідності ручного розрахунку окремих параметрів. Подібні підходи демонструють високу ефективність у задачах визначення сонливості та моніторингу стану водіїв.

Таким чином, візуальні біомаркери є важливою складовою сучасних систем автоматизованого визначення втоми людини. Використання технологій комп'ютерного зору та глибинного навчання дозволяє реалізувати безконтактний аналіз поведінкових ознак людини та забезпечити ефективний моніторинг функціонального стану у режимі реального часу.

1.5 Аналіз наборів даних та постановка задачі дослідження

Ефективність сучасних систем глибинного навчання значною мірою залежить від якості та обсягу навчальних даних. Для задач автоматизованого визначення втоми людини особливе значення мають набори зображень та відеоданих, що містять приклади різних станів користувача, зокрема ознак сонливості, зниження концентрації уваги та поведінкових проявів втоми. Саме тому одним із важливих етапів розробки систем fatigue detection є аналіз доступних наборів даних та визначення найбільш доцільного підходу до їх використання.

У сучасних дослідженнях для задач визначення втоми людини використовується значна кількість відкритих датасетів, які відрізняються за структурою, типом даних, умовами зйомки та способом анотації [29]. Більшість подібних наборів даних містять фотографії або відеозаписи облич

людей у різних станах пильності та використовуються для навчання моделей класифікації.

Одним із найбільш важливих критеріїв якості набору даних є різноманітність прикладів. Для забезпечення стійкості моделей глибинного навчання необхідно використовувати зображення, отримані в різних умовах освітлення, при різних положеннях голови, наявності окулярів, часткових перекриттів обличчя та інших факторів, що можуть впливати на якість розпізнавання. Недостатня різноманітність навчальних даних може призводити до перенавчання моделі та погіршення її здатності працювати з новими прикладами.

Для задач даної роботи було проаналізовано декілька відкритих наборів даних, розміщених на платформі Kaggle та присвячених визначенню сонливості або втоми людини. Дані набори містять зображення обличч користувачів у станах «втомлений» та «невтомлений», а також приклади закритих та відкритих очей, позіхання та інших поведінкових ознак сонливості.

Серед проаналізованих наборів даних можна виділити fatigue_detection [30], Fatigue Dataset [31], Drowsiness Prediction Dataset [32], Driver Drowsiness Dataset (DDD) [33], Drowness Detection [34], Drowsy Detection Dataset [35], Dataset_Drowsiness [36], driver_drowsiness_data [37], Driver_Drowsiness [38], Motorcycle-Driver-Drowsiness [39]. Більшість із них мають схожу структуру та містять попередньо розподілені категорії зображень, що значно спрощує процес підготовки даних для навчання моделей глибинного навчання.

Важливою особливістю розглянутих наборів даних є те, що вони були сформовані різними авторами та містять приклади, отримані в різних умовах зйомки. Це дозволяє розглядати можливість комбінування декількох датасетів в єдиний набір даних для підвищення різноманітності навчальної вибірки та покращення узагальнювальної здатності моделі.

У сучасних дослідженнях поєднання декількох наборів даних є досить поширеним підходом [29]. Використання комбінованих датасетів дозволяє збільшити кількість навчальних прикладів, зменшити вплив особливостей окремого джерела даних та підвищити стійкість моделей до різних умов експлуатації. Особливо актуальним це є для задач комп'ютерного зору, де різноманітність зображень безпосередньо впливає на ефективність навчання нейронних мереж.

Разом із перевагами комбінування наборів даних існують і певні складності. Основними проблемами є відмінності у форматах зображень, розмірах кадрів, схемах анотації та балансі класів. У різних датасетах можуть використовуватися різні критерії віднесення прикладів до категорії «втомлений» або «невтомлений», що може впливати на якість формування єдиної навчальної вибірки.

Для вирішення подібних проблем зазвичай використовуються методи попередньої обробки даних. До основних етапів *preprocessing* належать зміна розміру та пропорцій зображень, нормалізація піксельних значень, видалення пошкоджених файлів, балансування класів та аугментація даних [29]. Використання аугментації дозволяє штучно збільшити обсяг навчальної вибірки шляхом створення модифікованих копій зображень, зокрема із використанням поворотів, масштабування, зміни яскравості та інших трансформацій.

У задачах визначення втоми людини важливе значення має баланс між класами. Якщо кількість прикладів одного класу значно перевищує інший, модель може демонструвати упереджені результати та гірше розпізнавати менш представлений клас. Тому при формуванні єдиного набору даних необхідно враховувати співвідношення між категоріями «втомлений» та «невтомлений».

Для побудови сучасних систем *fatigue detection* широко використовуються згорткові нейронні мережі. CNN-моделі демонструють високу ефективність у задачах класифікації зображень та дозволяють

автоматично виділяти інформативні ознаки без необхідності ручного визначення параметрів. У межах даної роботи доцільним є використання готових архітектур глибинного навчання з можливістю подальшого донавчання на підготовленому наборі даних.

Одним із перспективних підходів є використання transfer learning, при якому застосовуються попередньо навчені моделі, натреновані на великих наборах зображень. Подібний підхід дозволяє суттєво зменшити обсяг необхідних навчальних даних та скоротити час навчання моделі. Крім того, transfer learning дозволяє досягати високої точності навіть при використанні відносно невеликих спеціалізованих датасетів.

У межах даної роботи передбачається дослідження можливості використання комбінованого набору даних, сформованого на основі декількох відкритих датасетів fatigue detection. Планується виконати попередню обробку та уніфікацію даних, сформувати збалансовану вибірку та провести навчання моделі глибинного навчання для автоматизованої класифікації стану людини за візуальними біомаркерами втоми.

Основною задачею дослідження є розробка та оцінювання моделі, здатної автоматично визначати ознаки втоми людини на основі аналізу зображень обличчя. Для досягнення поставленої мети необхідно вирішити ряд задач, серед яких:

- виконати аналіз предметної галузі та сучасних методів fatigue detection;
- дослідити особливості використання методів комп'ютерного зору та глибинного навчання;
- проаналізувати доступні набори даних для визначення втоми людини;
- сформувати та підготувати єдиний набір даних для навчання моделі;
- обрати архітектуру нейронної мережі та провести її навчання;
- оцінити ефективність роботи моделі на тестових даних;

– проаналізувати переваги, недоліки та перспективи подальшого розвитку систем автоматизованого визначення втоми людини.

Таким чином, проведений аналіз предметної галузі свідчить про актуальність задачі автоматизованого визначення втоми людини та доцільність використання методів комп'ютерного зору і глибинного навчання для її вирішення. Використання комбінованих наборів даних та сучасних архітектур нейронних мереж дозволяє створити систему, здатну ефективно виконувати аналіз візуальних біомаркерів та визначати ознаки сонливості у режимі реального часу.

2 РОЗРОБКА ПІДХОДУ ДО МОДЕЛЮВАННЯ ТА ОЦІНЮВАННЯ ВТОМИ ЛЮДИНИ

2.1 Обґрунтування вибору методів комп'ютерного зору та глибинного навчання

Проведений вище аналіз предметної галузі показав, що проблема автоматизованого визначення втоми людини є актуальною для багатьох сфер діяльності, пов'язаних із необхідністю тривалого підтримання концентрації уваги. Особливої важливості ця задача набуває у транспортній галузі, системах моніторингу операторів складного обладнання, а також у сучасних інформаційних системах підтримки безпеки.

У ході дослідження існуючих наукових підходів було встановлено, що для оцінювання функціонального стану людини можуть використовуватися різні групи методів. До них належать підходи, засновані на аналізі фізіологічних сигналів, поведінкових характеристик та візуальних ознак. Кожна з наведених груп має власні переваги та обмеження, що необхідно враховувати під час розробки практичних програмних рішень.

Виконаний огляд сучасних досліджень продемонстрував, що останніми роками найбільшого поширення набули системи, побудовані на основі комп'ютерного зору та методів глибинного навчання. Причиною цього є можливість автоматичного аналізу зображень обличчя без використання спеціалізованого обладнання та додаткових контактних датчиків. Подібні системи здатні працювати лише на основі даних, отриманих із звичайних цифрових камер, що значно спрощує їх практичне використання. Водночас проведений аналіз показав, що універсального підходу до вирішення поставленої задачі наразі не існує. Це створює передумови для розробки власного програмного рішення, яке б поєднувало переваги сучасних методів комп'ютерного зору з можливістю навчання на даних із декількох незалежних джерел.

Під час розробки системи моделювання та оцінювання втоми людини було розглянуто декілька підходів до автоматичного аналізу функціонального стану користувача. Аналіз проведених досліджень показав, що найбільш перспективним напрямком є використання методів комп'ютерного зору та моделей глибинного навчання, здатних працювати без застосування додаткових сенсорів або спеціалізованого обладнання.

На відміну від підходів, що базуються на фізіологічних сигналах, таких як електроенцефалографія, електрокардіографія чи аналіз рухів голови за допомогою спеціалізованих датчиків, використання візуальної інформації дозволяє проводити оцінювання стану людини безконтактним способом. Це значно спрощує практичне впровадження системи та розширює можливі сфери її застосування.

Основними візуальними ознаками, що характеризують стан втоми людини, є ступінь відкриття очей, частота та тривалість моргань, положення повік, відкриття рота під час позіхання, а також загальні зміни миміки обличчя. Дані ознаки можуть бути отримані безпосередньо із зображень або відеопотоку та використовуватись як інформативні біомаркери функціонального стану людини.

На початкових етапах розвитку систем визначення втоми переважно використовувались алгоритми ручного виділення ознак. Такі підходи вимагали попереднього пошуку ключових точок обличчя та розрахунку різноманітних геометричних характеристик, зокрема показників EAR, MAR та PERCLOS. Незважаючи на відносну простоту реалізації, подібні методи мають низку обмежень. Їхня ефективність суттєво залежить від якості освітлення, положення голови користувача, наявності окулярів та інших факторів зовнішнього середовища.

У зв'язку з цим у даній роботі було прийнято рішення використати підхід на основі згорткових нейронних мереж. Головною перевагою таких моделей є можливість автоматичного виділення інформативних ознак без необхідності їх явного програмного опису. Під час навчання мережа

самостійно формує багаторівневе представлення вхідного зображення та виявляє характерні особливості, що дозволяють розрізняти різні функціональні стани людини.

Для побудови моделі було розглянуто архітектури згорткових нейронних мереж, що підтримують transfer learning. Використання попередньо навчених моделей дозволило скоротити час навчання та підвищити точність класифікації навіть при використанні обмежених обсягів навчальних даних. Як потенційні архітектури розглядалися MobileNet, EfficientNet або ResNet, які широко використовуються у задачах комп'ютерного зору.

Важливою особливістю обраного підходу стало використання технології перенесення навчання. Замість навчання моделі з нуля використовувалась попередньо навчена нейронна мережа, що вже містила базові знання про структуру та особливості візуальних об'єктів. Такий підхід дозволив суттєво скоротити тривалість навчання та підвищити якість отриманих результатів навіть за умов використання неоднорідних наборів даних.

Для забезпечення високої узагальнювальної здатності моделі було прийнято рішення використовувати об'єднаний набір даних, сформований на основі декількох незалежних датасетів. Такий підхід дозволяє враховувати різноманітні умови зйомки, відмінності у якості зображень, ракурсах обличчя та особливостях розмітки даних. У результаті модель отримує можливість навчатися на більш репрезентативній вибірці та демонструвати стабільні результати при роботі з новими даними.

Таким чином, проведений аналіз показав доцільність використання методів комп'ютерного зору та глибинного навчання як основи для побудови системи оцінювання втоми людини. Обрані технології забезпечують необхідну точність класифікації, високу адаптивність до різних умов використання та можливість подальшого розвитку системи без суттєвої зміни її архітектури.

2.2 Розробка підходу до практичної реалізації

Для вирішення задачі автоматизованого визначення втоми людини у межах даної роботи розглядалася розробка системи, заснованої на використанні методів комп'ютерного зору та глибинного навчання для аналізу візуальних біомаркерів користувача. Основною метою системи є автоматизоване визначення ознак сонливості або втоми на основі аналізу зображень обличчя людини.

Запропонований підхід передбачає використання безконтактного методу моніторингу стану людини за допомогою набору статичних зображень. На відміну від фізіологічних систем, що потребують використання спеціалізованих датчиків та сенсорів, запропоноване рішення не створює дискомфорту для користувача та було реалізоване із застосуванням стандартного обладнання.

Оснoву системи складає модель глибинного навчання, навчена на зображеннях обличч людей із різними рівнями втоми. Передбачено, що модель виконує класифікацію стану користувача на декілька категорій, наприклад «втомлений» та «невтомлений». Для реалізації даної задачі доцільним було використання згорткових нейронних мереж, які демонструють високу ефективність у задачах аналізу зображень.

На першому етапі роботи системи виконано процедури підготовки даних. Для цього було використано комбінований набір даних, сформований на основі декількох відкритих датасетів. Об'єднання декількох наборів даних дозволило збільшити кількість навчальних прикладів та забезпечити більшу різноманітність умов зйомки, що позитивно вплинуло на стійкість моделі до зовнішніх факторів.

Після об'єднання наборів даних було виконано попередню обробку зображень. Даний етап включає зміну розміру зображень, нормалізацію піксельних значень, перевірку коректності даних, балансування класів та аугментацію. Використання аугментації дозволило штучно збільшити

різноманітність навчальної вибірки шляхом створення модифікованих копій зображень із використанням поворотів, масштабування, зміни яскравості та інших трансформацій.

Для реалізації системи було обрано архітектуру MobileNetV3 Small. Вибір даної моделі обумовлений поєднанням високої точності класифікації та відносно невеликої обчислювальної складності. На відміну від більш важких архітектур, MobileNetV3 потребує менше обчислювальних ресурсів та може використовуватись навіть на персональних комп'ютерах без високопродуктивних графічних прискорювачів.

Передбачено, що навчання моделі виконується із використанням фреймворків PyTorch. Даний програмний засіб забезпечує ефективну реалізацію алгоритмів глибокого навчання, а також дозволяє використовувати графічні процесори для прискорення процесу навчання нейронних мереж.

У межах роботи також було досліджено можливість використання додаткового аналізу окремих візуальних біомаркерів, зокрема стану очей та рота користувача. Для цього були розглянуті методи facial landmark detection та алгоритми визначення Eye Aspect Ratio і Mouth Aspect Ratio. Комбінування результатів класифікації нейронної мережі та аналізу окремих поведінкових ознак дозволяє підвищити надійність роботи системи.

Після завершення навчання моделі було проведено оцінювання ефективності її роботи на тестовій вибірці. Для аналізу якості класифікації було використано показники accuracy, precision, recall та F1- score. Також доцільним є аналіз confusion matrix для визначення характеру помилок класифікації.

Однією з переваг запропонованого рішення є його гнучкість та можливість подальшого масштабування. У майбутньому система може бути доповнена підтримкою відеоаналізу, використанням часових моделей,

наприклад LSTM або Transformer-архітектур, а також інтеграцією із мобільними пристроями або вбудованими системами.

Таким чином, запропоноване рішення передбачило створення інтелектуальної системи автоматизованого визначення втоми людини, заснованої на аналізі візуальних біомаркерів за допомогою методів комп'ютерного зору та глибинного навчання. Використання сучасних архітектур нейронних мереж та комбінованих наборів даних дозволило створити систему, здатну ефективно виконувати класифікацію стану користувача та забезпечити можливість подальшого практичного застосування у реальних умовах. Таким чином було розроблено структуру системи для практичної реалізації запропонованого рішення (рисунок 2.1).



Рисунок 2.1 – Структура системи оцінювання втоми людини

3 ПРАКТИЧНА РЕАЛІЗАЦІЯ СИСТЕМИ

3.1 Формування та підготовка набору даних для навчання моделі

Одним із ключових етапів розробки системи автоматизованого оцінювання втоми людини є формування якісного набору даних, який буде використовуватися для навчання та тестування моделі глибинного навчання. Оскільки метою даної роботи є створення універсальної системи, здатної працювати з різними типами зображень обличчя людини, було прийнято рішення використовувати не один окремий датасет, а об'єднати декілька незалежних наборів даних у єдиний інформаційний простір.

Для проведення дослідження було відібрано десять відкритих датасетів, присвячених задачам виявлення сонливості, втоми водіїв, розпізнавання закритих очей та визначення позіхання. Обрані набори даних відрізнялися між собою як за структурою організації файлів, так і за типами представлених зображень. Частина датасетів містила лише області очей, інші включали повні зображення обличчя, а деякі набори містили фотографії людей у реальних умовах експлуатації з наявністю навколишнього фону.

Загальний обсяг доступних даних після об'єднання всіх джерел склав понад 463 тисячі зображень. Такий масштаб набору даних значно перевищує обсяг більшості окремих датасетів, що використовуються у подібних дослідженнях, та дозволяє отримати більш стійку до перенавчання модель.

На початковому етапі було проведено автоматизований аналіз структури всіх наборів даних. Для кожного датасету виконувалась перевірка доступності файлів, визначення форматів зображень, аналіз розмірів кадрів та кольорових моделей. Результати дослідження показали значну неоднорідність вихідних даних. У різних джерелах зустрічалися як невеликі зображення розміром 48×48 пікселів, так і фотографії з роздільною

здатністю понад 1900×1900 пікселів. Частина наборів використовувала кольорові RGB-зображення, тоді як інші містили лише одноканальні напівтонові кадри.

Окремою задачею стало приведення різних систем розмітки до єдиної структури класів. Кожен із використаних датасетів був створений незалежними авторами, тому назви категорій суттєво відрізнялися. Наприклад, для позначення стану бадьорості використовувалися назви «normal», «active», «natural», «open-eyes», «alert» та інші аналогічні варіанти. Аналогічна ситуація спостерігалася для класів сонливості та втоми.

Для забезпечення сумісності даних було розроблено власну систему уніфікації класів. Усі наявні категорії були приведені до трьох узагальнених станів:

- Alert (користувач перебуває у нормальному стані бадьорості);
- Drowsy (користувач демонструє ознаки сонливості або втоми);
- Yawn (користувач здійснює позіхання).

Подібний підхід дозволив не лише об'єднати різноманітні джерела інформації, але й сформулювати більш універсальну задачу класифікації, яка відповідає практичним сценаріям використання системи моніторингу стану людини.

Після завершення процедури уніфікації було сформовано централізований каталог даних, який містив інформацію про шлях до кожного файлу, джерело його походження та відповідний клас. Для цього було створено спеціальний метафайл у форматі CSV, що містив понад 463 тисячі записів. Такий підхід дозволив працювати з великим набором даних без необхідності фізичного копіювання зображень між директоріями та значно спростив подальшу організацію експериментів.

На наступному етапі було проведено додатковий аналіз розподілу даних між джерелами. Отримані результати показали, що різні датасети мають суттєво відмінний внесок у загальний обсяг вибірки. Найбільші

набори містили понад 150 тисяч зображень, тоді як окремі джерела налічували лише декілька тисяч файлів.

Подібна ситуація створювала ризик виникнення ефекту домінування великих датасетів під час навчання моделі. У випадку прямого використання всіх доступних даних нейронна мережа могла б орієнтуватися переважно на особливості найбільших наборів, ігноруючи менш представлені джерела. Це негативно вплинуло б на здатність системи узагальнювати знання на нові дані.

Для перевірки потенційних проблем було проведено додатковий аналіз можливих дублікатів між датасетами. Було реалізовано механізм порівняння перцептивних хешів випадкових зображень з різних джерел. Результати дослідження показали наявність лише незначної кількості збігів відносно загального обсягу даних, що підтвердило можливість використання всіх датасетів у межах єдиного набору без суттєвого ризику дублювання інформації.

У результаті виконаних дій було сформовано єдиний консолідований набір даних, який став основою для подальших експериментів з побудови та навчання нейронної мережі. Отримана структура дозволила поєднати різноманітність джерел інформації з можливістю централізованого керування даними та створила необхідні умови для проведення наступних етапів дослідження.

3.2 Формування інтегрованого набору даних для навчання моделі

Одним із ключових етапів реалізації системи стала підготовка навчальної вибірки. На відміну від багатьох досліджень, у яких використовується один окремий датасет, у даній роботі було прийнято рішення об'єднати декілька незалежних наборів даних у єдиний інтегрований масив зображень. Такий підхід дозволяє підвищити різноманітність навчальних прикладів, зменшити залежність моделі від

особливостей конкретного джерела даних та покращити здатність до узагальнення на нових зображеннях.

Для формування навчального набору було використано десять відкритих датасетів, присвячених задачам виявлення сонливості, втоми водіїв та аналізу поведінкових ознак людини. Вихідні набори даних відрізнялися між собою за структурою каталогів, форматами зображень, роздільною здатністю кадрів, умовами зйомки та системою маркування класів.

Загальна кількість зображень у всіх вихідних наборах даних становила 463585 файлів. При цьому різні датасети використовували власні системи класифікації. Наприклад, окремі набори містили класи «Open-Eyes» та «Close-Eyes», інші використовували категорії «Active», «Sleep» або «Yawn», а деякі поділяли зображення на стани «alert», «non_vigilant» та «tired».

Безпосереднє об'єднання таких наборів даних неможливе через відмінність семантики класів. Тому перед формуванням інтегрованого датасету було розроблено процедуру уніфікації розмітки.

У рамках роботи всі початкові категорії були приведені до трьох узагальнених класів: клас Alert, якому відповідають зображення, на яких людина перебуває у нормальному стані бадьорості, клас Drowsy, якому відповідають зображення, на яких спостерігаються ознаки сонливості або втоми, клас Yawn, якому відповідають зображення, на яких зафіксовано процес позіхання.

До класу Alert були віднесені категорії normal, Non Drowsy, Open-Eyes, Active, Active Subj, NATURAL, NonFatigue, alert та 0, тобто класи, що характеризують активний стан людини.

До класу Drowsy були віднесені категорії sleep, Drowsy, Close-Eyes, Sleep, FatigueSubj, DROWSY, Fatigue, tired, non_vigilant та 1, тобто стани, що відповідають проявам втоми або сонливості.

До класу Yawn були віднесені категорії yawp та Yawn, тобто ті, що пов'язані з позіханням.

У результаті уніфікації було отримано наступний розподіл даних: клас Alert складався з 204087 зображень, клас Drowsy складався з 209496 зображень, клас Yawn складався з 50002 зображення.

Отримана статистика показала наявність значного дисбалансу між класами. Хоча класи Alert та Drowsy містили приблизно однакову кількість прикладів, клас Yawn виявився майже у чотири рази меншим. Крім того, внесок окремих датасетів також був нерівномірним. Найбільші набори містили понад 150 тисяч зображень, тоді як деякі спеціалізовані датасети налічували лише декілька тисяч файлів.

Додатковою проблемою стало те, що різні датасети представляли різні типи вхідних даних. Частина зображень містила лише область очей, інші включали повне обличчя, а деякі кадри відображали людину майже повністю разом із навколишнім середовищем. Роздільна здатність варіювалася від 48×48 до 2560×2560 пікселів. Також використовувалися як кольорові RGB-зображення, так і напівтонові зображення у відтінках сірого.

З метою перевірки якості об'єднання було виконано серію додаткових досліджень структури даних. Було проведено автоматичний аналіз випадкових вибірок з кожного датасету, під час якого оцінювалися формати файлів, кольорові моделі, роздільна здатність зображень, цілісність файлів та можливі дублікати між наборами даних.

Перевірка показала відсутність пошкоджених файлів серед усіх проаналізованих вибірок. Для виявлення можливих дублікатів використовувався метод перцептивного хешування зображень. Аналіз продемонстрував наявність певної кількості повторюваних кадрів між окремими датасетами, що є очікуваним результатом для відкритих наборів даних, сформованих на основі схожих джерел. Проте кількість таких збігів виявилася незначною порівняно із загальним обсягом вибірки та не могла суттєво вплинути на результати навчання.

Після завершення етапу аналізу було сформовано єдиний метаопис усіх доступних зображень у вигляді структурованої таблиці. Для кожного файлу зберігався шлях до зображення, назва початкового датасету та відповідний уніфікований клас.

Однак використання повного набору із 463585 зображень для навчання на центральному процесорі потребувало б надмірних обчислювальних ресурсів та значного часу навчання. Тому було реалізовано механізм квотованого відбору прикладів.

На відміну від випадкового скорочення датасету, яке могло б призвести до втрати окремих джерел даних, було використано пропорційне представлення всіх наборів даних усередині кожного класу. Для кожного з трьох класів формувалася вибірка однакового розміру, а кількість прикладів з конкретного датасету визначалася відповідно до його частки всередині класу.

Такий підхід дозволив одночасно забезпечити баланс між класами, зберегти інформацію з усіх доступних датасетів, уникнути домінування найбільших наборів даних та скоротити обсяг навчальної вибірки до прийняттого для експериментів розміру.

У результаті було сформовано збалансований інтегрований датасет, що містив 45000 зображень: по 15000 прикладів для кожного класу. Саме цей набір даних використовувався для подальшого навчання та тестування моделей глибокого навчання.

Сформований датасет став основою всіх подальших експериментів та дозволив досліджувати роботу моделі на даних різного походження, різної якості та різних умов зйомки, що наближає результати до реальних умов експлуатації системи оцінювання втоми людини.

3.3 Підготовка даних та організація процесу навчання

Після формування інтегрованого набору даних було виконано підготовку вибірки до навчання нейромережевої моделі. Основною метою даного етапу стало забезпечення коректного подання вхідних даних, зменшення впливу випадкових факторів та створення умов для об'єктивної оцінки якості роботи системи.

Оскільки інтегрований датасет був сформований шляхом об'єднання десяти незалежних джерел даних, окремі зображення суттєво відрізнялися між собою за роздільною здатністю, пропорціями сторін, кольоровими характеристиками та якістю зйомки. Для забезпечення сумісності зі згортковою нейронною мережею було необхідно привести всі зображення до єдиного формату.

У процесі завантаження кожне зображення автоматично перетворювалося до кольорового простору RGB незалежно від початкового формату. Такий підхід дозволив уніфікувати обробку як кольорових, так і напівтонових зображень, що були присутні у вихідних датасетах.

Наступним етапом стало масштабування зображень до розміру 224×224 пікселів. Вибір саме такого розміру обумовлений тим, що він є стандартним для більшості сучасних архітектур комп'ютерного зору та використовується під час попереднього навчання моделей на великих наборах даних. Крім того, даний розмір забезпечує прийнятний баланс між швидкістю навчання та збереженням інформативних ознак зображення.

Для підвищення узагальнювальної здатності моделі було використано процедуру аугментації даних. Її основна мета полягає у створенні додаткового різноманіття навчальних прикладів без фактичного збільшення кількості файлів у датасеті.

Під час навчання до зображень випадковим чином застосовувалися такі перетворення, як горизонтальне віддзеркалення, незначні повороти зображення, випадкові зміни масштабу, випадкові зміщення області кадру, незначні зміни яскравості та незначні зміни контрастності.

Застосування аугментації дозволяє моделі навчатися на ширшому спектрі можливих варіацій зовнішнього вигляду людини та умов зйомки. Особливо важливим це є для даної задачі, оскільки в реальних умовах положення голови, освітлення, якість камери та відстань до об'єкта можуть суттєво відрізнятися від умов, представлених у навчальних даних.

Після завершення етапу попередньої обробки було виконано поділ вибірки на навчальну, валідаційну та тестову частини.

Для забезпечення об'єктивності оцінювання використовувався стратифікований розподіл даних. Це означає, що співвідношення класів Alert, Drowsy та Yawn зберігалось однаковим у всіх підмножинах датасету.

У результаті поділу були отримані три вибірки: навчальна вибірка, що складалася з 31500 зображень, валідаційна вибірка, що складалася з 6750 зображень та тестова вибірка, що складалася з 6750 зображень.

Таким чином, близько 70% даних використовувалося для навчання моделі, тоді як по 15% даних було виділено для контролю якості під час навчання та незалежного підсумкового тестування.

Валідаційна вибірка використовувалася після завершення кожної епохи навчання. На основі її результатів оцінювалася поточна якість моделі та визначався момент досягнення найкращої узагальнювальної здатності. Збереження ваг моделі виконувалося відповідно до найкращого значення точності на валідаційних даних.

Тестова вибірка не брала участі ані в навчанні, ані в підборі параметрів моделі. Вона використовувалася лише після завершення всього процесу навчання для отримання фінальних показників ефективності системи. Такий підхід дозволяє уникнути завищення результатів та забезпечує більш реалістичну оцінку якості роботи моделі на нових даних.

Реалізація програмного комплексу виконувалася мовою Python із використанням бібліотек PyTorch, Torchvision, NumPy, Pandas, Pillow та Scikit-learn. Вибір PyTorch обумовлений широким поширенням даного

фреймворку в задачах глибокого навчання, високою гнучкістю та наявністю великої кількості готових архітектур нейронних мереж.

Навчання проводилося на персональному комп'ютері без використання дискретного графічного прискорювача. Усі обчислення виконувалися центральним процесором Intel Core i5-13420H. Незважаючи на відсутність спеціалізованого обладнання для машинного навчання, використання компактної архітектури MobileNetV3 дозволило виконати повний цикл експериментів за прийнятний час та отримати високі показники точності класифікації.

Особливістю організації експерименту стало прагнення максимально наблизити умови навчання до практичного застосування майбутньої системи. Для цього у навчальній вибірці одночасно використовувалися зображення різних типів: окремі області очей, повні обличчя та фотографії людини з частиною навколишнього середовища. Завдяки цьому модель навчалася працювати не лише з ідеалізованими даними, але й з більш складними випадками, характерними для реальних умов експлуатації.

У результаті виконаних підготовчих процедур було сформовано повністю готове середовище для проведення експериментального дослідження та навчання нейромережевої моделі оцінювання втоми людини за візуальними біомаркерами. Наступним етапом стало безпосереднє навчання моделі та аналіз отриманих результатів класифікації.

3.4 Розробка та навчання нейромережевої моделі оцінювання втоми людини

Після завершення підготовки даних було виконано розробку та навчання нейромережевої моделі, призначеної для автоматичного визначення стану людини за візуальними ознаками. Основною задачею моделі є віднесення вхідного зображення до одного з трьох класів: Alert, Drowsy або Yawn.

Під час вибору архітектури нейронної мережі враховувалися декілька важливих факторів. По-перше, модель повинна забезпечувати високу точність класифікації на різнорідних даних. По-друге, вона має залишатися достатньо компактною для практичного використання на звичайних персональних комп'ютерах або вбудованих пристроях. По-третє, бажано мінімізувати час навчання та виконання прогнозів без суттєвої втрати якості.

З урахуванням зазначених вимог було обрано архітектуру MobileNetV3 Small. Дана архітектура належить до сімейства компактних згорткових нейронних мереж та спеціально розроблена для задач комп'ютерного зору, де необхідно поєднати високу швидкодію та достатню точність класифікації.

На відміну від класичних глибоких мереж, які можуть містити десятки мільйонів параметрів, MobileNetV3 використовує механізм *depthwise-separable convolutions*, що дозволяє суттєво скоротити кількість обчислень. Крім того, в архітектурі застосовуються оптимізовані блоки Mobile Inverted Bottleneck Convolution та модулі уваги Squeeze-and-Excitation, які покращують якість виділення інформативних ознак.

Використання MobileNetV3 є особливо доцільним для даної роботи, оскільки майбутня система потенційно може використовуватися на ноутбуках, персональних комп'ютерах або мобільних пристроях без потужних графічних прискорювачів.

Для прискорення навчання було застосовано метод перенесення навчання (Transfer Learning). Замість навчання мережі з нуля використовувалися вагові коефіцієнти, попередньо отримані під час навчання моделі на великому наборі природних зображень ImageNet.

Такий підхід дозволяє використати вже сформовані низькорівневі та середньорівневі ознаки, які нейронна мережа навчилася виділяти під час початкового навчання. До них належать контури, текстури, градієнти, локальні форми та інші візуальні характеристики. Завдяки цьому модель

швидше адаптується до нової предметної галузі та потребує меншого обсягу навчальних даних.

Оскільки вихідна версія MobileNetV3 призначена для класифікації тисячі класів, фінальний класифікаційний шар було модифіковано відповідно до поставленої задачі. Вихідний класифікатор було замінено новим повнозв'язним шаром, що формує три вихідні значення, які відповідають класам Alert, Drowsy та Yawn.

Під час навчання використовувалася функція втрат Cross Entropy Loss. Дана функція є стандартним вибором для задач багатокласової класифікації та дозволяє оцінювати відмінність між прогнозованим розподілом ймовірностей та правильними мітками класів.

Для оновлення параметрів мережі використовувався оптимізатор Adam. Основною перевагою Adam є автоматичне адаптування швидкості навчання для кожного параметра моделі, що забезпечує швидшу збіжність порівняно з класичним стохастичним градієнтним спуском.

Початкове значення коефіцієнта навчання було встановлено на рівні 0,001. Такий вибір забезпечував стабільний процес навчання без різких коливань значення функції втрат.

Навчання виконувалося протягом двадцяти епох. Після завершення кожної епохи проводилася оцінка моделі на валідаційній вибірці. Це дозволяло контролювати процес навчання та відстежувати зміни якості класифікації.

У рамках експерименту для кожної епохи фіксувалися: значення функції втрат на навчальній вибірці, точність класифікації на валідаційній вибірці, тривалість виконання епохи та поточний стан моделі.

Крім того, реалізовано автоматичне збереження найкращої версії моделі. Якщо після завершення чергової епохи точність на валідаційній вибірці перевищувала попередній максимум, поточні ваги мережі зберігалися окремо як найкращий результат навчання.

Такий підхід дозволяє уникнути ситуації, коли після досягнення максимуму якості подальше навчання починає погіршувати узагальнювальну здатність моделі внаслідок перенавчання.

Аналіз журналу навчання показав стабільний процес оптимізації параметрів мережі. Уже на перших епохах модель демонструвала високий рівень точності, що підтверджує ефективність використання перенесення навчання та якість сформованого набору даних.

У ході навчання спостерігалось поступове зменшення значення функції втрат та одночасне зростання точності класифікації на валідаційній вибірці. Відсутність різких коливань або деградації результатів свідчить про коректний вибір параметрів навчання та стабільність архітектури.

Особливий інтерес становить той факт, що модель успішно навчалася на інтегрованому датасеті, який містив зображення різних типів. Під час навчання мережа одночасно отримувала як крупні фрагменти очей, так і повні обличчя та фотографії людей із навколишнім середовищем. Незважаючи на таку різноманітність даних, процес оптимізації проходив стабільно, що підтверджує здатність обраної архітектури виділяти універсальні ознаки втомі незалежно від типу вхідного зображення.

Після завершення двадцятої епохи було отримано фінальну версію моделі, яка продемонструвала найкращі показники на валідаційній вибірці та була використана для подальшого тестування на незалежному наборі даних.

Отримана модель стала основою розробленої системи оцінювання втомі людини та дозволила перейти до етапу кількісного аналізу результатів класифікації, дослідження помилок і оцінювання практичної ефективності запропонованого рішення.

3.5 Аналіз результатів експериментального дослідження

Після завершення навчання нейромережевої моделі було проведено комплексне експериментальне дослідження її ефективності. Основною метою даного етапу стала оцінка якості класифікації стану людини за візуальними біомаркерами та перевірка практичної придатності розробленої системи.

Для тестування використовувалася незалежна тестова вибірка, яка не брала участі ані в навчанні моделі, ані у процесі підбору параметрів. Такий підхід дозволяє отримати об'єктивну оцінку здатності системи працювати на нових даних.

Під час навчання було виконано двадцять епох оптимізації параметрів нейронної мережі. Для кожної епохи фіксувалися значення функції втрат та точність класифікації на валідаційній вибірці. Отримані результати показали стабільний характер навчання та поступове покращення якості роботи моделі.

На першій епосі точність класифікації на валідаційній вибірці становила 93,53%, що вже свідчить про достатньо високий рівень узагальнення завдяки використанню попередньо навчених ваг. У подальшому спостерігалось поступове покращення результатів практично на всьому проміжку навчання.

Максимальне значення точності на валідаційній вибірці було досягнуто на двадцятій епосі та склало 97,76%. Для наочності зміни точності класифікації в процесі навчання побудовано графік залежності точності від номера епохи (рисунок 3.1).

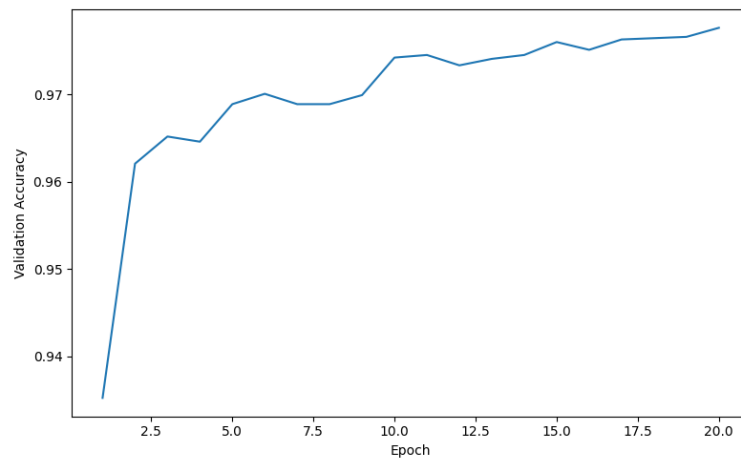


Рисунок 3.1 – Графік точності класифікації на валідаційній вибірці під час навчання моделі

Аналіз графіка показує, що після швидкого початкового покращення відбувається поступове наближення до максимуму. При цьому явних ознак нестабільності або деградації результатів не спостерігається. Це свідчить про правильний вибір параметрів навчання та достатній обсяг навчальних даних.

Додатково було проаналізовано поведінку функції втрат. Протягом усього процесу навчання спостерігалось її стабільне зменшення від початкових значень до фінального рівня. Графік функції втрат наведено нижче (рисунок 3.2).

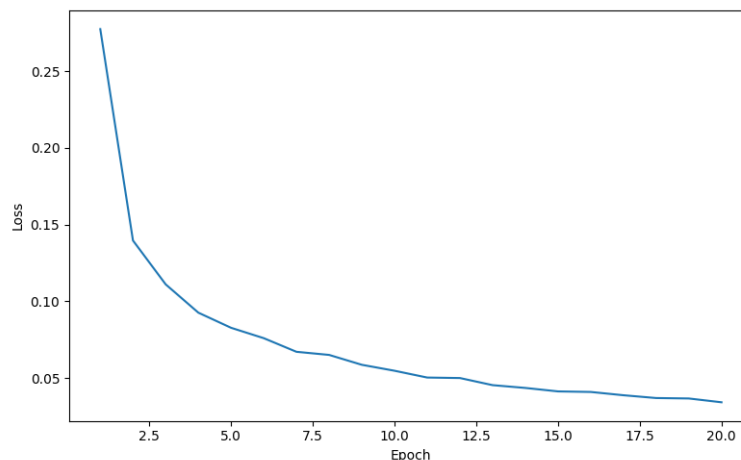


Рисунок 3.2 – Графік функції втрат під час навчання моделі

Отримана залежність підтверджує успішну збіжність алгоритму навчання та відсутність проблем, пов'язаних із перенавчанням або нестабільністю процесу оптимізації.

Після завершення навчання найкраща версія моделі була протестована на незалежній тестовій вибірці, що містила 6750 зображень.

За результатами тестування отримано високі значення основних показників ефективності, а саме: показник Ассигасу склав 97,67%, показник Precision дорівнював 97,68%, показник Recall склав 97,67%, а показник F1-score відповідно 97,67% (рисунок 3.3).

```
=====
TEST RESULTS
=====
Accuracy : 0.9767407407407407
Precision: 0.9767702353042695
Recall   : 0.9767407407407407
F1-score : 0.9767376184503178

Best epoch: 20
Best val accuracy: 0.9776296296296296
```

Рисунок 3.3 – Консольний вивід результатів тестування

Отримані значення свідчать про високу якість роботи розробленої системи. Особливо важливим є той факт, що всі чотири показники мають майже однакові значення. Це говорить про відсутність суттєвого перекосу моделі в бік окремих класів та про збалансованість навчання. Для більш детального аналізу результатів було побудовано матрицю помилок класифікації (рисунок 3.4).

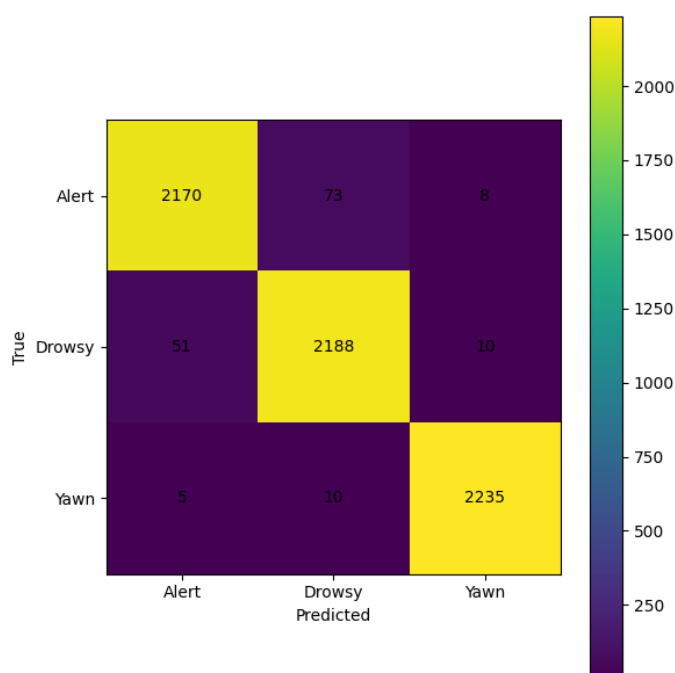


Рисунок 3.4 – Матриця помилок класифікації для тестової вибірки

Аналіз матриці показав розподіл результатів, що відповідає високому рівню точності класифікації.

Для класу Alert було правильно класифіковано 2170 з 2251 зображення. При цьому 73 зображення були помилково віднесені до класу Drowsy, а 8 зображень – до класу Yawn.

Для класу Drowsy правильно визначено 2188 із 2249 зображень. Помилково до класу Alert було віднесено 51 зображення, а до класу Yawn – 10 зображень.

Найкращі результати були отримані для класу Yawn. Із 2250 тестових зображень правильно класифіковано 2235. Лише 15 прикладів було віднесено до інших класів, а саме 5 до класу Alert та 10 до класу Drowsy.

Отримані результати дозволяють зробити важливий висновок. Позіхання є найбільш вираженим візуальним біомаркером серед трьох розглянутих категорій. Саме тому модель практично не допускає помилок при його виявленні.

Водночас основна кількість помилок виникає між класами Alert та Drowsy. Це є цілком очікуваним результатом, оскільки межа між станом бадьорості та початковими проявами сонливості часто є нечіткою навіть для людини-експерта.

Для детальнішої оцінки роботи моделі було проаналізовано показники для кожного класу окремо, а також показники швидкодії (рисунок 3.5).

```
=====
              precision    recall  f1-score   support

   Alert         0.97        0.96        0.97        2251
   Drowsy        0.96        0.97        0.97        2249
   Yawn          0.99        0.99        0.99        2250

 accuracy              0.98              6750
 macro avg           0.98        0.98        0.98        6750
weighted avg           0.98        0.98        0.98        6750
=====
Average ms/image: 0.7888
FPS: 1267.82
```

Рисунок 3.5 – Консольний вивід показників окремих класів та швидкодії

Для класу Alert значення F1-score склало 0,97. Для класу Drowsy значення F1-score також становило 0,97. Для класу Yawn показник F1-score досягнув 0,99. Середнє значення F1-score для всіх класів дорівнює 0,98, що свідчить про високий рівень узагальнення моделі та її здатність коректно працювати з різними типами вхідних даних.

Також було проведено оцінювання швидкодії розробленої системи. Вимірювання виконувалися для одного зображення після завершення навчання. Середній час обробки одного кадру становив 0,79 мс. Відповідна пропускна здатність системи склала приблизно 1268 кадрів за секунду.

Отриманий результат суттєво перевищує вимоги більшості практичних застосувань. Навіть при використанні стандартної вебкамери з частотою 30 кадрів за секунду продуктивність моделі є більш ніж достатньою для роботи в режимі реального часу.

Важливою особливістю проведеного дослідження є те, що модель навчалася не на одному спеціалізованому датасеті, а на інтегрованому

наборі даних, сформованому з десяти незалежних джерел. У процесі аналізу було встановлено, що модель успішно класифікує як зображення окремих очей, так і повні зображення обличчя та фотографії людей з навколишнім середовищем. Це підтверджує правильність обраного підходу до формування інтегрованого датасету та демонструє високу узагальнювальну здатність розробленого рішення.

Таким чином, результати проведеного експериментального дослідження підтверджують ефективність запропонованого підходу до оцінювання втоми людини за допомогою глибинного аналізу візуальних біомаркерів. Розроблена система забезпечує високу точність класифікації, стабільність роботи та достатню швидкодію для використання в режимі реального часу. Отримані результати дозволяють розглядати запропоновану модель для створення практичних систем моніторингу стану людини у транспортних, виробничих, навчальних та інших прикладних середовищах.

ВИСНОВКИ

У рамках виконання кваліфікаційної роботи було вирішено задачу моделювання та оцінювання втоми людини за допомогою методів глибинного навчання та аналізу візуальних біомаркерів. Проведене дослідження охопило як теоретичні аспекти проблеми виявлення втоми людини, так і практичну розробку програмного рішення для автоматичної класифікації стану людини за зображеннями.

У першій частині роботи було виконано аналіз предметної галузі та сучасного стану досліджень у сфері автоматизованого виявлення втоми та сонливості. Розглянуто основні візуальні біомаркери, що використовуються для оцінювання функціонального стану людини, зокрема ступінь відкриття очей, позіхання, особливості положення голови та зміни міміки. Проведено аналіз сучасних підходів до побудови систем моніторингу втоми, їх переваг, недоліків та обмежень практичного застосування.

На основі проведеного аналізу було сформульовано постановку задачі та обґрунтовано доцільність використання методів комп'ютерного зору і згорткових нейронних мереж для автоматизації процесу оцінювання втоми людини. Запропоновано концепцію системи, здатної працювати з різними типами вхідних зображень та здійснювати класифікацію стану користувача за трьома категоріями: Alert, Drowsy та Yawn.

У процесі реалізації практичної частини роботи було сформовано інтегрований набір даних на основі десяти незалежних відкритих датасетів. Загальний обсяг вихідних даних склав понад 460 тисяч зображень. Для забезпечення сумісності між різними джерелами було виконано уніфікацію системи розмітки та приведення вихідних категорій до єдиної трирівневої класифікації станів людини.

Проведено аналіз структури вихідних даних, досліджено особливості різних датасетів, виконано перевірку цілісності файлів та оцінено рівень можливого дублювання між окремими джерелами. Для усунення

дисбалансу класів та надмірного впливу окремих наборів даних було розроблено механізм квотованого формування збалансованої навчальної вибірки. У результаті сформовано підсумковий набір даних, що містить 45000 зображень із рівним представленням усіх трьох класів.

Для розв'язання поставленої задачі було реалізовано модель на основі архітектури MobileNetV3 Small із використанням технології перенесення навчання. Виконано підготовку даних, налаштування процесу навчання та проведено серію експериментів із використанням навчальної, валідаційної та тестової вибірок.

У результаті навчання моделі було досягнуто високих показників якості класифікації. Найкраща точність на валідаційній вибірці склала 97,76%, а під час тестування на незалежних даних отримано точність 97,67%. Значення Precision, Recall та F1-score також перевищили 97%, що свідчить про високу збалансованість та надійність розробленої системи.

Проведений аналіз матриці помилок показав, що модель ефективно розрізняє всі три досліджувані класи. Найкращі результати було отримано для категорії Yawn, яка характеризується найбільш вираженими візуальними ознаками. Основна частина помилок спостерігалася між класами Alert та Drowsy, що відповідає особливостям предметної галузі та складності розмежування ранніх проявів сонливості й нормального стану бадьорості.

Оцінювання швидкодії моделі продемонструвало можливість її використання в системах реального часу. Середній час обробки одного зображення склав менше однієї мілісекунди, що забезпечує продуктивність понад 1200 кадрів за секунду та значно перевищує вимоги більшості практичних сценаріїв використання.

Практична цінність роботи полягає у створенні програмного рішення, здатного автоматично оцінювати ознаки втоми людини за візуальними даними. Розроблений підхід може бути використаний у системах

моніторингу стану водіїв транспортних засобів, операторів виробничого обладнання, працівників відповідальних об'єктів, а також у навчальних та дослідницьких проєктах, пов'язаних із аналізом поведінки людини.

Отримані результати підтверджують ефективність використання методів глибинного навчання для задач оцінювання втоми людини та демонструють перспективність застосування інтегрованих наборів даних, сформованих із декількох незалежних джерел. Подальший розвиток роботи може бути пов'язаний із використанням відеопослідовностей замість окремих зображень, аналізом часової динаміки поведінки користувача, розширенням набору візуальних біомаркерів та адаптацією системи для роботи на мобільних і вбудованих пристроях.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

- 1) Бабак С. Фізіологічні та біохімічні механізми розвитку втоми при фізичних навантаженнях. *Вісник Національного університету оборони України*. 2023. С. 17–24. URL: <https://doi.org/10.33099/2617-6858-2023-71-1-17-24> (дата звернення: 13.05.2026).
- 2) Williamson A. M. Moderate sleep deprivation produces impairments in cognitive and motor performance equivalent to legally prescribed levels of alcohol intoxication. *Occupational and environmental medicine*. 2000. Vol. 57, no. 10. P. 649–655. URL: <https://doi.org/10.1136/oem.57.10.649> (date of access: 13.05.2026).
- 3) The role of driver sleepiness in car crashes: a systematic review of epidemiological studies / J. Connor et al. *Accident analysis & prevention*. 2001. Vol. 33, no. 1. P. 31–41. URL: [https://doi.org/10.1016/s0001-4575\(00\)00013-0](https://doi.org/10.1016/s0001-4575(00)00013-0) (date of access: 13.05.2026).
- 4) Lal S. K. L., Craig A. A critical review of the psychophysiology of driver fatigue. *Biological psychology*. 2001. Vol. 55, no. 3. P. 173–194. URL: [https://doi.org/10.1016/s0301-0511\(00\)00085-5](https://doi.org/10.1016/s0301-0511(00)00085-5) (date of access: 13.05.2026).
- 5) Sahayadhas A., Sundaraj K., Murugappan M. Detecting driver drowsiness based on sensors: a review. *Sensors*. 2012. Vol. 12, no. 12. P. 16937–16953. URL: <https://doi.org/10.3390/s121216937> (date of access: 13.05.2026).
- 6) A survey on state-of-the-art drowsiness detection techniques / M. Ramzan et al. *IEEE access*. 2019. Vol. 7. P. 61904–61919. URL: <https://doi.org/10.1109/access.2019.2914373> (date of access: 13.05.2026).
- 7) Hu S., Zheng G. Driver drowsiness detection with eyelid related parameters by Support Vector Machine. *Expert systems with applications*. 2009. Vol. 36, no. 4. P. 7651–7658. URL: <https://doi.org/10.1016/j.eswa.2008.09.030> (date of access: 13.05.2026).

8) Klimesch W. EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain research reviews*. 1999. Vol. 29, no. 2-3. P. 169–195. URL: [https://doi.org/10.1016/s0165-0173\(98\)00056-3](https://doi.org/10.1016/s0165-0173(98)00056-3) (date of access: 13.05.2026).

9) Drowsiness detection using heart rate variability / J. Vicente et al. *Medical & biological engineering & computing*. 2016. Vol. 54, no. 6. P. 927–937. URL: <https://doi.org/10.1007/s11517-015-1448-7> (date of access: 13.05.2026).

10) Driver fatigue detection method based on eye states with pupil and iris segmentation / Q. Zhuang et al. *IEEE access*. 2020. Vol. 8. P. 173440–173449. URL: <https://doi.org/10.1109/access.2020.3025818> (date of access: 13.05.2026).

11) Youwei L. Real-time eye blink detection using general cameras: a facial landmarks approach. *International science journal of engineering & agriculture*. 2023. Vol. 2, no. 5. P. 1–8. URL: <https://doi.org/10.46299/j.isjea.20230205.01>. (date of access: 13.05.2026).

12) Yawning detection using embedded smart cameras / M. Omidyeganeh et al. *IEEE transactions on instrumentation and measurement*. 2016. Vol. 65, no. 3. P. 570–582. URL: <https://doi.org/10.1109/tim.2015.2507378> (date of access: 13.05.2026).

13) Driver inattention monitoring system for intelligent vehicles: a review / Y. Dong et al. *IEEE transactions on intelligent transportation systems*. 2011. Vol. 12, no. 2. P. 596–614. URL: <https://doi.org/10.1109/tits.2010.2092770> (date of access: 13.05.2026).

14) Albadawi Y., Takruri M., Awad M. A review of recent developments in driver drowsiness detection systems. *Sensors*. 2022. Vol. 22, no. 5. P. 2069. URL: <https://doi.org/10.3390/s22052069> (date of access: 13.05.2026).

15) Deng W., Wu R. Real-Time driver-drowsiness detection system using facial features. *IEEE access*. 2019. Vol. 7. P. 118727–118738. URL: <https://doi.org/10.1109/access.2019.2936663> (date of access: 13.05.2026).

16) Driver gaze tracking and eyes off the road detection system / F. Vicente et al. *IEEE transactions on intelligent transportation systems*. 2015. Vol. 16, no. 4. P. 2014–2027. URL: <https://doi.org/10.1109/tits.2015.2396031> (date of access: 13.05.2026).

17) Forsyth D. A., Ponce J. Computer vision: a modern approach. Pearson Education, Limited, 2012. 792 p.

18) Ji Q. Real-Time eye, gaze, and face pose tracking for monitoring driver vigilance. *Real-Time imaging*. 2002. Vol. 8, no. 5. P. 357–377. URL: <https://doi.org/10.1006/rtim.2002.0279> (date of access: 14.05.2026).

19) Viola P., Jones M. Rapid object detection using a boosted cascade of simple features. *2001 IEEE computer society conference on computer vision and pattern recognition. CVPR 2001*, Kauai, HI, USA. URL: <https://doi.org/10.1109/cvpr.2001.990517> (date of access: 14.05.2026).

20) Yawning detection using embedded smart cameras / M. Omidyeganeh et al. *IEEE transactions on instrumentation and measurement*. 2016. Vol. 65, no. 3. P. 570–582. URL: <https://doi.org/10.1109/tim.2015.2507378> (date of access: 14.05.2026).

21) Bengio Y., Courville A., Goodfellow I. Deep learning. MIT Press, 2016. 800 p.

22) Gradient-based learning applied to document recognition / Y. Lecun et al. *Proceedings of the IEEE*. 1998. Vol. 86, no. 11. P. 2278–2324. URL: <https://doi.org/10.1109/5.726791> (date of access: 14.05.2026).

23) EfficientNet-Based robust recognition of peach plant diseases in field images / T. Ahmad et al. *Computers, materials & continua*. 2022. Vol. 71, no. 1. P. 2073–2089. URL: <https://doi.org/10.32604/cmc.2022.018961> (date of access: 14.05.2026).

24) Pan S. J., Yang Q. A survey on transfer learning. *IEEE transactions on knowledge and data engineering*. 2010. Vol. 22, no. 10. P. 1345–1359. URL: <https://doi.org/10.1109/tkde.2009.191> (date of access: 14.05.2026).

25) Kaehler A., Bradski G. Learning OpenCV: Computer Vision with the OpenCV Library. O'Reilly Media, Incorporated, 2008.

26) Waqas M. T. A comparative analysis of pytorch vs. tensorflow for developing convolutional neural networks. ResearchHub Technologies, Inc., 2026. URL: <https://doi.org/10.55277/researchhub.eqwrgs8v> (date of access: 14.05.2026).

27) Development of a real-time driver's drowsiness detection system using mediapipe face mesh / S. Bau et al. *International journal of engineering and manufacturing*. 2025. Vol. 15, no. 5. P. 46–57. URL: <https://doi.org/10.5815/ijem.2025.05.04> (date of access: 14.05.2026).

28) Sommer D., Golz M. Evaluation of PERCLOS based current fatigue monitoring technologies. *2010 32nd annual international conference of the IEEE engineering in medicine and biology society (EMBC 2010)*, Buenos Aires, 31 August – 4 September 2010. 2010. URL: <https://doi.org/10.1109/iembs.2010.5625960> (date of access: 14.05.2026).

29) Shorten C., Khoshgoftaar T. M. A survey on image data augmentation for deep learning. *Journal of big data*. 2019. Vol. 6, no. 1. URL: <https://doi.org/10.1186/s40537-019-0197-0> (date of access: 14.05.2026).

30) Fatigue_detection. *Kaggle: The World's AI Proving Ground*. URL: <https://www.kaggle.com/datasets/timmjy/fatigue-detection> (date of access: 14.05.2026).

31) Fatigue dataset. *Kaggle: The World's AI Proving Ground*. URL: <https://www.kaggle.com/datasets/rihabkaci99/fatigue-dataset> (date of access: 14.05.2026).

32) Drowsiness prediction dataset. *Kaggle: The World's AI Proving Ground*. URL: <https://www.kaggle.com/datasets/rakibuleceruet/drowsiness-prediction-dataset> (date of access: 14.05.2026).

33) Driver drowsiness dataset (DDD). *Kaggle: The World's AI Proving Ground*. URL: <https://www.kaggle.com/datasets/ismailnasri20/driver-drowsiness-dataset-ddd> (date of access: 14.05.2026).

34) Drownness_Detection. *Kaggle: The World's AI Proving Ground*.
URL: <https://www.kaggle.com/datasets/norannabil/drownness-detection> (date of access: 14.05.2026).

35) Drowsy detection dataset. *Kaggle: The World's AI Proving Ground*.
URL: <https://www.kaggle.com/datasets/yasharjebraeily/drowsy-detection-dataset> (date of access: 14.05.2026).

36) Dataset_Drowsiness. *Kaggle: The World's AI Proving Ground*.
URL: https://www.kaggle.com/datasets/tintrnc/dataset-drowsiness?select=Dataset_drowsiness (date of access: 14.05.2026).

37) Driver_drowsiness_data. *Kaggle: The World's AI Proving Ground*.
URL: <https://www.kaggle.com/datasets/abinkjohn/driver-drowsiness-data> (date of access: 14.05.2026).

38) Driver_Drowsiness. *Kaggle: The World's AI Proving Ground*.
URL: <https://www.kaggle.com/datasets/anhelayoub/driver-drowsiness> (date of access: 14.05.2026).

39) Motorcycle-Driver-Drowsiness. *Kaggle: The World's AI Proving Ground*.
URL: <https://www.kaggle.com/datasets/arvinantobahtiar/motorcycle-driver-drowsiness> (date of access: 14.05.2026).