

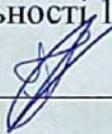
Вінницький національний технічний університет
Факультет інформаційних електронних систем
Кафедра біомедичної інженерії та оптико-електронних систем

МАГІСТЕРСЬКА КВАЛІФІКАЦІЙНА РОБОТА

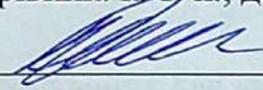
на тему:

Мобільна система позиціонування для людей з вадами зору

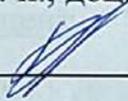
Виконав: студент 2-го курсу, гр. БМІ–24м
спеціальності 163 Біомедична інженерія


Гзоїм В. А.

Керівник: к. т. н., доц., доцент каф. БМІОЕС

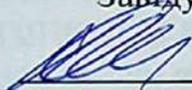

Коваль Л. Г.

Опонент: к. т. н., доц. каф. ІРТС


Притула М. О.

Допущено до захисту

Завідувач кафедри БМІОЕС

 к. т. н., доц. Коваль Л. Г.

« 8 » 12 2025 р.

Вінниця ВНТУ – 2025 рік

Вінницький національний технічний університет

Факультет інформаційних електронних систем

Кафедра біомедичної інженерії та оптико-електронних систем

Рівень вищої освіти – другий (магістерський)

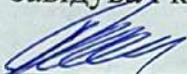
Галузь знань 16 Хімічна інженерія та біоінженерія

Спеціальність 163 Біомедична інженерія

Освітньо-професійна програма «Інтелектуальні штучні імпланти та медичні апарати в біоінженерії»

ЗАТВЕРДЖУЮ

завідувач кафедри БМІОЕС

 Л. Г. Коваль

«24» 09 2025 р.

ЗАВДАННЯ

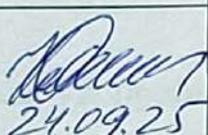
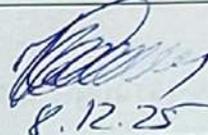
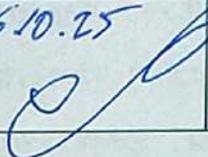
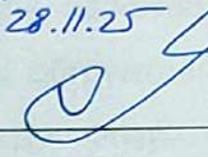
НА МАГІСТЕРСЬКУ КВАЛІФІКАЦІЙНУ РОБОТУ

Гзоїму Владиславу Андрійовичу

1. Тема роботи: Мобільна система позиціонування для людей з вадами зору.
Керівник роботи: Коваль Л. Г., канд. техн. наук, доцент
затверджені наказом ВНТУ від «24» вересня 2025 року № 313.
2. Строк подання студентом роботи «08» грудня 2025 року.
3. Вихідні дані до роботи: асистивний пристрій для просторової орієнтації осіб з вадами зору, апаратно-програмний комплекс на базі технологій комп'ютерного зору (YOLO) та ультразвукової локації; система тактильного та голосового зворотного зв'язку; можливість 3D друку елементів корпусу, патенти, технічна документація на електронні компоненти, науково-технічна література.
4. Зміст текстової частини: аналіз існуючих асистивних технологій та засобів навігації для осіб з вадами зору, огляд методів комп'ютерного зору та архітектури нейронних мереж, обґрунтування вибору апаратної платформи та алгоритму YOLO, розробка структурної схеми та програмного забезпечення асистивного пристрою з функціями тактильного і голосового зворотного зв'язку, розробка стартап-проекту та розрахунок економічних показників доцільності розробки.

5. Перелік ілюстративного матеріалу: структурна схема апаратно-програмного комплексу, електрична принципова схема підключення компонентів, архітектура нейронної мережі YOLO, графіки навчання моделі та результати детекції об'єктів.

6. Консультанти розділів роботи

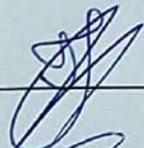
Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
Спеціальна частина	Коваль Л. Г. к. т. н., доц. каф. БМІОЕС	 24.09.25	 8.12.25
Економічна частина	Козловський В. О. к. е. н., проф. каф. ЕПВМ	6.10.25 	28.11.25 

7. Дата видачі завдання «24» вересня 2025 року.

КАЛЕНДАРНИЙ ПЛАН

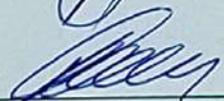
Назва етапів магістерської кваліфікаційної роботи	Строк виконання	Прим.
Аналітичний огляд питання, постановка завдань роботи	до 06.10.2025	
Розроблення моделей і методик досягнення мети роботи	до 27.10.2025	
Розроблення системи	до 24.11.2025	
Розрахунок техніко-економічних показників	до 01.12.2025	
Оформлення текстової та графічної частини	до 08.12.2025	
Попередній захист МКР	08.12.2025	
Рецензування МКР	до 19.12.2025	
Захист МКР	22-23.12.2025	

Здобувач



Владислав ГЗОЇМ

Керівник роботи



Леонід КОВАЛЬ

АНОТАЦІЯ

Гзоїм В. А. Мобільна система позиціонування для людей з вадами зору: магістерська кваліфікаційна робота зі спеціальності 163 Біомедична інженерія, освітньо-професійна програма «Інтелектуальні штучні імпланти та медичні апарати в біоінженерії». Вінниця : ВНТУ, 2025. 90 с.

Укр. мовою. Бібліогр.: 25 найм.; 22 рис.; 10 табл.

В магістерській кваліфікаційній роботі розглянуті проблеми просторової орієнтації осіб з порушеннями зору, зроблено огляд сучасних асистивних технологій та програмних засобів навігації. В другому розділі обґрунтовано вибір компонентної бази та розроблено структурну схему апаратно-програмного комплексу, що поєднує ультразвукову локацію та методи комп'ютерного зору. В третьому розділі здійснено програмну реалізацію системи: інтегровано та адаптовано згорткову нейронну мережу YOLO для розпізнавання об'єктів у реальному часі, розроблено алгоритми голосового та тактильного зворотного зв'язку. В четвертому розділі представлено стартап-проект та розраховано економічну доцільність розробки. Показано, що запропоноване гібридне рішення забезпечує високу ефективність ідентифікації перешкод та є економічно вигідним порівняно з існуючими аналогами.

Ключові слова: асистивні технології, порушення зору, нейромережа YOLO, ультразвукова локація, тактильний зворотний зв'язок, комп'ютерний зір.

ABSTRACT

Gzoim V. A. Mobile positioning system for people with visual impairments: Master's thesis in the field of 163 Biomedical Engineering, educational and professional programme "Intelligent artificial implants and medical devices in bioengineering". Vinnytsia: VNTU, 2025. 104 p.

In Ukrainian. Bibliography: 25 items; 22 figures; 10 tables.

The master's qualification thesis examines the problems of spatial orientation of persons with visual impairments and reviews modern assistive technologies and software navigation tools. The second chapter justifies the choice of the component base and develops the block diagram of the hardware-software complex, which combines ultrasonic location and computer vision methods. The third chapter implements the system's software: the YOLO convolutional neural network is integrated and adapted for real-time object recognition, and algorithms for voice and tactile feedback are developed. The fourth chapter presents a startup project and calculates the economic feasibility of the development. It is shown that the proposed hybrid solution ensures high efficiency of obstacle identification and is economically advantageous compared to existing analogues.

Keywords: assistive technologies, visual impairments, YOLO neural network, ultrasonic location, tactile feedback, computer vision.

ЗМІСТ

ВСТУП	6
1 ОСНОВНІ ТЕОРЕТИЧНІ ВІДОМОСТІ	10
1.1 Аналітичний огляд сучасних асистивних технологій для осіб із порушеннями зору	10
1.2. Огляд літературних джерел та аналіз існуючих технічних рішень	11
<u>1.2.2.</u> Аналіз мобільних застосунків та програмних платформ.....	17
1.3. Обґрунтування вибору методів детекції та тактильного зворотного зв'язку.....	20
1.4. Порівняльний аналіз та обґрунтування вибору алгоритму комп'ютерного зору	21
1.5. Теоретичні основи функціонування нейронної мережі YOLO	23
<u>1.5.1.</u> Базові принципи та архітектура.....	23
<u>1.5.3.</u> Ключові переваги та обмеження використання алгоритму YOLO	26
1.6 Огляд нормативних вимог та стандартів до асистивних технологій.....	27
1.7 Висновки до розділу 1	28
2 РОЗРОБКА АСИСТИВНОГО ПРИСТРОЮ НА ОСНОВІ ОПТОЕЛЕКТРОННИХ АНАЛІЗАТОРІВ	30
2.1. Обґрунтування вибору компонентної бази та архітектура системи.....	30
<u>2.1.1.</u> Аналіз основних складових апаратної частини	30
<u>2.1.2.</u> Специфікація елементної бази та обґрунтування вибору компонентів	32
<u>2.1.3</u> Обґрунтування вибору оптичного сенсора для системи технічного зору.....	34
<u>2.1.4.</u> Обґрунтування вибору обчислювальної платформи для цифрової обробки зображень.....	36
2.2. Моделювання та віртуальна верифікація роботи пристрою	37
<u>2.2.1.</u> Симуляція роботи сенсорної підсистеми	37
<u>2.2.2.</u> Алгоритмічна реалізація базового прототипу.....	39
<u>2.2.3.</u> Масштабування системи для кругового огляду (360°)	39
<u>2.2.4.</u> Оптимізація силової частини схеми (інтеграція драйверів L293D)	40

2.2.5. Модифікація алгоритму керування та реалізація адаптивних режимів роботи.....	42
2.3. Фізична реалізація та збірка прототипу асистивного пристрою.....	44
2.4 Розрахунок енергетичного балансу та автономності апаратно-програмного комплексу.....	44
2.5 Висновки до розділу 2	47
3 ІНТЕГРАЦІЯ АЛГОРИТМУ РОЗПІЗНАВАННЯ ОБ'ЄКТІВ ТА ПРОГРАМНА РЕАЛІЗАЦІЯ СИСТЕМИ.....	49
3.1. Структурно-функціональна організація нейронної мережі YOLO.....	49
3.2. Підготовка даних та параметричне налаштування процесу навчання	50
3.4. Математична модель функції втрат (Loss Function).....	57
3.5. Програмна реалізація та верифікація роботи системи в режимі реального часу	60
3.5.1. Алгоритм обробки відеопотоку та візуалізація результатів детекції	61
3.5.2. Реалізація циклу інференсу нейромережі.....	63
3.5.3. Обробка метаданих детекції та аналіз стійкості алгоритму.....	64
3.5.4. Організація голосового інтерфейсу та реалізація двосторонньої взаємодії	66
3.6 Висновки до розділу 3	68
4. ЕКОНОМІЧНИЙ РОЗДІЛ.....	69
4.1 Технологічний аудит розробленої мобільної системи позиціонування для людей з вадами зору	69
4.2 Розрахунок витрат на розроблення мобільної системи позиціонування для людей з вадами зору	72
4.3 Розрахунок економічного ефекту від можливого виведення на ринок розробленої мобільної системи позиціонування для людей з вадами.....	76
ВИСНОВКИ.....	81
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	83
ДОДАТОК А.....	86
ДОДАТОК Б	90

ВСТУП

Актуальність дослідження. Розвиток інформаційно-комунікаційних технологій та мініатюризація електронних компонентів створили передумови для якісного стрибка у сфері тифлотехніки (техніки для сліпих). Асистивні системи позиціонування трансформуються з простих механічних засобів у складні інтелектуальні комплекси, що базуються на використанні комп'ютерного зору, ультразвукової локації та супутникової навігації. Метою таких розробок є надання користувачеві повної інформації про навколишнє середовище, що є ключовим фактором безпеки та самостійності.

Соціальна значущість дослідження підтверджується статистичними даними. У світовому масштабі кількість осіб із вадами зору продовжує зростати і станом на 2024 рік оцінюється у 338 мільйонів осіб, з яких 43 мільйони мають повну втрату зору [1].

В Україні ситуація ускладнюється наслідками бойових дій. Від початку повномасштабного вторгнення фіксується різке збільшення випадків мінно-вибухових травм, що призводять до ураження зорового аналізатора. За офіційною статистикою НСЗУ, протягом 2022–2023 років кількість нових діагнозів, пов'язаних із втратою зору, перевищила 38 тисяч випадків. Це формує соціальний запит на створення доступних реабілітаційних технологій.

Зростання чисельності осіб з порушеннями зору, зокрема серед ветеранів бойових дій, актуалізує потребу в розробці високоефективних асистивних технологій, спрямованих на забезпечення їхньої автономності та безпеки. Аналіз існуючих засобів просторової орієнтації виявив низку суттєвих недоліків. Традиційні засоби, такі як біла тростина, базуються на контактному методі виявлення перешкод, що обмежує радіус дії та не захищає від об'єктів, розташованих вище рівня пояса.

Водночас, електронні аналоги, що використовують звукову індикацію, створюють акустичний шум. Це є критичним недоліком, оскільки перевантажує

слуховий канал користувача, заважаючи сприйняттю природних звуків навколишнього середовища, які є життєво важливими для орієнтації сліпих людей.

Для вирішення окреслених проблем у магістерській дисертації запропоновано концепцію апаратно-програмного комплексу, що поєднує переваги ультразвукової та інфрачервоної локації.

Розроблена система забезпечує безконтактне виявлення перешкод у радіусі 360 градусів, усуваючи «сліпі зони». Ключовою особливістю рішення є використання тактильного зворотного зв'язку (вібраційного відгуку). Такий підхід дозволяє передавати інформацію про оточення через дотикові відчуття, залишаючи слуховий канал вільним для комунікації та моніторингу зовнішньої ситуації.

Попри наявність на ринку закордонних аналогів, більшість із них залишаються недоступними для широкого загалу через високу вартість або відсутність адаптації до вітчизняної інфраструктури. Тому в рамках магістерської роботи запропоновано підхід до створення мобільної системи позиціонування, яка поєднує ефективність сучасних оптоелектронних сенсорів з економічною доцільністю.

Метою магістерської кваліфікаційної роботи є підвищення рівня безпеки та просторової орієнтації осіб із порушеннями зору шляхом розробки мобільної системи позиціонування, що забезпечує оперативне інформування про перешкоди через канал тактильного зворотного зв'язку.

Завданнями магістерської кваліфікаційної роботи є:

- Здійснити аналітичний огляд сучасних асистивних технологій та засобів навігації для незрячих, виявити їхні конструктивні недоліки та визначити напрямки вдосконалення.
- Обґрунтувати вибір елементної бази та спроектувати функціонально-структурну схему пристрою, що реалізує мультисенсорний підхід до виявлення перешкод.

- Дослідити методи розширення функціоналу системи, зокрема шляхом інтеграції алгоритмів комп'ютерного зору (розпізнавання зображень) та розробки адаптивних режимів вібраційної індикації.

Об'єкт магістерської кваліфікаційної роботи: процеси та технічні засоби просторової орієнтації та навігації осіб із вадами зору.

Предмет магістерської кваліфікаційної роботи: методи та апаратно-програмні засоби побудови мультисенсорної асистивної системи з використанням нейромережевого розпізнавання об'єктів та тактильно-голосового інтерфейсу.

Методи дослідження. Під час виконання магістерської роботи застосовувалися такі методи: аналітично-пошуковий метод, методи теорії біотехнічних систем; методи обробки зображень та машинного навчання; імітаційне моделювання, метод визначення оптимального варіанту конструкції.

Новизна одержаних результатів:

Наукова новизна дослідження полягає у вирішенні задачі підвищення інформативності та безпеки навігації для осіб з вадами зору. Основні наукові результати, отримані в роботі:

1. Удосконалено метод мультисенсорної локації, який, на відміну від існуючих аналогів, базується на комплексуванні даних від різномірних джерел: ультразвукових та інфрачервоних сенсорів для точного вимірювання відстані та згорткової нейронної мережі YOLO (You Only Look Once) для семантичного розпізнавання об'єктів. Це дозволяє формувати комплексну модель навколишнього середовища, ідентифікуючи не лише наявність перешкоди, а і її тип у реальному часі.
2. Набула подальшого розвитку система людино-машинної взаємодії для асистивних пристроїв. Запропоновано гібридний інтерфейс, що поєднує адаптивний тактильний зворотний зв'язок (вібраційний відгук змінної інтенсивності) та голосове управління. Такий підхід

забезпечує дублювання каналів передачі інформації, знижуючи когнітивне навантаження на користувача.

Практичне значення одержаних результатів: розроблено функціональну схему та діючий макетний зразок мобільної системи позиціонування, який поєднує ультразвукову локацію та алгоритми комп'ютерного зору (YOLO) для ідентифікації перешкод.

1 ОСНОВНІ ТЕОРЕТИЧНІ ВІДОМОСТІ

1.1 Аналітичний огляд сучасних асистивних технологій для осіб із порушеннями зору

Сучасні електронні засоби орієнтації (Electronic Travel Aids — ETA) базуються на принципах сенсорного сприйняття навколишнього середовища з подальшою трансформацією отриманих даних у сигнали, доступні для сприйняття користувачем. Апаратна складова таких систем зазвичай включає лідари (LiDAR), камери комп'ютерного зору, ультразвукові далекоміри або їх комбінації.

Інформація про навколишній простір передається користувачеві через альтернативні канали сприйняття:

Тактильний канал (вібрація, шрифт Брайля, електростимуляція);

Аудіальний канал (синтезована мова, тональні звукові сигнали, бінауральне аудіо).

Варто зазначити, що питання вибору оптимального інтерфейсу «людина-машина» залишається відкритим у науковій спільноті. Аудіосигнали є інформативними, але можуть блокувати природний слуховий контроль середовища, тоді як тактильний зв'язок менш інформативний, але не перевантажує слух.

Незалежно від обраної архітектури, до сучасних систем навігації висувається ряд критичних вимог, які визначають їхню ефективність та безпеку. Ключові експлуатаційні характеристики таких пристроїв систематизовано в табл. 1.1.

Таблиця 1.1 – Ключові вимоги до систем мобільного позиціонування

Характеристика	Опис вимоги	Примітки
Швидкодія (Real-time processing)	Система повинна забезпечувати мінімальну затримку між детекцією	Затримка не повинна перевищувати 100–200 мс. Система, що виявляє перешкоду

Характеристика	Опис вимоги	Примітки
	перешкоди та сповіщенням користувача.	за 2 м із затримкою в кілька секунд, є небезпечною для використання під час руху [2].
Універсальність (Покриття)	Здатність ефективно функціонувати як у замкненому просторі (Indoor), так і на відкритій місцевості (Outdoor).	Включає стійкість до різних типів поверхонь (асфальт, ґрунт, сходи).
Незалежність від освітлення	Система повинна зберігати працездатність у будь-який час доби (день/ніч).	Критично важливо для оптичних систем (камер), які потребують ІЧ-підсвічування або додаткових сенсорів у темряві.
Дальність детекції	Робочий діапазон, у якому гарантується виявлення перешкод.	Оптимальна «мертва зона» (мінімальна відстань) — до 0,5 м. Максимальна дальність — від 3 до 5 м і більше для завчасного планування маршруту [2].
Типи об'єктів	Здатність розпізнавати як статичні (стіни, стовпи), так і динамічні (люди, транспорт) об'єкти.	Система має фільтрувати хибні спрацювання та пріоритезувати об'єкти, що несуть безпосередню загрозу зіткнення.

Таким чином, розробка ефективного асистивного пристрою вимагає балансу між швидкістю обробки даних, точністю сенсорів та ергономічністю зворотного зв'язку.

1.2. Огляд літературних джерел та аналіз існуючих технічних рішень

Процес проєктування ефективного засобу навігації вимагає детального аналізу існуючої науково-технічної бази. Дослідження сучасного стану проблеми охоплює вивчення типів сенсорів для детекції перешкод, методів реалізації тактильного зворотного зв'язку, а також архітектури пристроїв, що вже представлені в науковій літературі або на ринку.

Одним із найбільш поширених напрямків розвитку асистивних технологій є концепція «Розумної тростини» (Smart Cane). Це клас портативних пристроїв, що інтегрують електроніку в конструкцію традиційної білої тростини для розширення її функціоналу.

Типовим представником таких систем є пристрій, описаний у роботі [3]. Його архітектура базується на взаємодії трьох ключових підсистем:

1. Сенсорна підсистема: Використовує ультразвукові далекоміри для безконтактного виявлення перешкод на шляху користувача.
2. Обчислювальне ядро: Обробка даних здійснюється за допомогою контролера на базі нечіткої логіки (Fuzzy Logic Controller). Використання нечіткої логіки дозволяє системі приймати більш адаптивні та "людиноподібні" рішення в умовах невизначеності (наприклад, коли сигнал від датчика є нестабільним або зашумленим), забезпечуючи високу точність навігаційних інструкцій.
3. Інтерфейс зворотного зв'язку:
 - *Тактильний канал*: Реалізований за допомогою сервоприводів (або вібромоторів), які надають користувачеві інформацію про відстань або напрямок руху через фізичний вплив на руку.
 - *Аудіальний канал*: Голосові повідомлення, що дублюють критично важливу інформацію.

Такий підхід дозволяє перетворити звичайну тростину на активний засіб безпеки, здатний попереджати про небезпеку ще до моменту фізичного контакту з перешкодою (див. рис. 1.1).



Рисунок 1.1 – Розумна тростина [3]

WeWALK Smart Cane 2 Сучасним прикладом інтеграції передових інформаційних технологій у засоби реабілітації є розробка WeWALK Smart Cane 2. У літературі цей пристрій характеризується як високотехнологічний мобільний асистент, що поєднує функції активної безпеки, супутникової навігації та штучного інтелекту [4].

Архітектура пристрою спрямована на забезпечення мультимодального зворотного зв'язку (multimodal feedback). Це означає, що система не лише детектує перешкоди за допомогою ультразвукових сенсорів, а й взаємодіє з користувачем через кілька каналів комунікації (вібраційний відгук на рукоятці, голосовий супровід через смартфон). Використання алгоритмів штучного інтелекту дозволяє системі адаптуватися до умов навколишнього середовища,

1. Модуль комп'ютерного зору: Головний блок, оснащений стереоскопічною парою відеокамер (stereo vision), що розміщується на голові або верхній частині тіла користувача.
2. Інтерфейс користувача: Пояс, що фіксується на талії та містить матрицю тактильних актуаторів («клітин»). Кожен актуатор відповідає за певний сектор простору і активується при виявленні перешкоди у відповідному напрямку.
3. Система живлення: Автономний акумуляторний блок, розміщений у кишені.

Принцип роботи системи базується на обробці відеопотоку з двох камер для побудови глибинної карти (Disparity Map) та генерації спрощеної 3D-моделі навколишнього середовища (рис. 1.3).

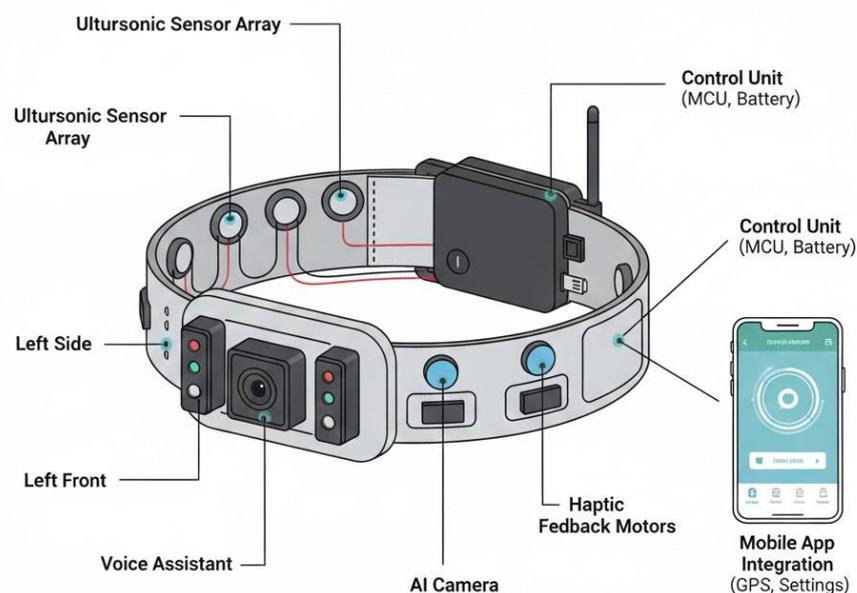


Рисунок 1.3 – Пояс зі зворотнім зв'язком розпізнання перешкод [6]

Попри ергономічність рішення, проведені експерименти виявили суттєві технічні обмеження. Основним недоліком є низька дальність детектування (рис. 1.4). Стереоскопічний метод ефективно розпізнає лише найближчі об'єкти (ближня зона), тоді як похибка вимірювання відстані до віддалених перешкод зростає експоненціально. Це обмежує можливість користувача завчасно планувати маршрут і уникати зіткнень на високій швидкості ходьби.

Незважаючи на різноманіття технічних рішень, практичне застосування розглянутих систем пов'язане з низкою суттєвих обмежень. Зокрема, використання складних вібраційних патернів (системи кодування інформації через ритм або частоту вібрації) має свої недоліки. З точки зору психофізіології, тривалий вплив монотонних механічних подразників призводить до ефекту сенсорної адаптації (звикання). З часом чутливість користувача до таких сигналів знижується, що може призвести до пропуску критично важливих попереджень про небезпеку. Крім того, необхідність запам'ятовувати велику кількість вібраційних комбінацій створює надмірне когнітивне навантаження.

Узагальнюючи аналіз ринку та літературних джерел, можна виділити три основні проблеми, що стримують масове впровадження існуючих асистивних пристроїв:

1. Експлуатаційні обмеження (Environmental Limitations):

- Ефективність багатьох систем критично залежить від умов навколишнього середовища. Оптичні системи (камери) втрачають точність в умовах недостатнього освітлення, туману або засліплення сонячним світлом. Ультразвукові датчики можуть давати хибні спрацювання у приміщеннях зі складними акустичними властивостями (луна, скляні поверхні) або на зашумлених вулицях.

2. Економічний бар'єр (Cost-Effectiveness):

- Більшість високотехнологічних рішень (наприклад, системи на базі LiDAR або складного комп'ютерного зору) мають високу собівартість. Це робить їх недоступними для широкого кола користувачів в Україні, де питання соціального забезпечення осіб з

інвалідністю часто обмежене фінансуванням. Більшість пристроїв так і залишаються на етапі лабораторних прототипів, не досягаючи серійного виробництва.

3. Проблема сенсорного перевантаження (Sensory Overload):

- Це ключовий недолік більшості асистивних засобів. При втраті зору слух і дотик стають основними каналами отримання інформації про світ.
- *Аудіальні пристрої* (наушники, голосові підказки) часто блокують або маскують природні звуки навколишнього середовища (шум транспорту, кроки), що є небезпечним для життя.
- *Тактильні пристрої* можуть заважати природному використанню рук або білої тростини.
- Таким чином, некоректно спроектований пристрій замість допомоги може дезорієнтувати користувача, створюючи інформаційний шум.

1.2.2. Аналіз мобільних застосунків та програмних платформ

Окремим сегментом ринку асистивних технологій є програмні рішення, реалізовані на базі смартфонів. Сучасний смартфон, оснащений камерою високої роздільної здатності, GPS-модулем, акселерометром та потужним процесором, фактично є універсальною мультисенсорною платформою для навігації та розпізнавання образів.

Умовно такі рішення можна поділити на три категорії:

1. Системи краудсорсингової допомоги (Human-in-the-loop) Яскравим представником цієї категорії є застосунок Be My Eyes. Принцип його роботи базується на відеостримінгу в реальному часі. Додаток встановлює відеозв'язок між незрячим користувачем та волонтером (зрячою людиною).

- *Алгоритм роботи:* При активації запиту система автоматично надсилає сповіщення групі волонтерів, враховуючи мовні налаштування та часовий пояс. Перший доступний волонтер отримує зображення з камери

смартфона користувача і через голосовий зв'язок допомагає ідентифікувати об'єкти, читати написи або орієнтуватися на місцевості.

- *Аналіз:* Хоча це рішення є ефективним для вирішення побутових завдань, воно має суттєві обмеження: повна залежність від наявності стабільного Інтернет-з'єднання та людського фактора (доступності волонтерів) [7].

2. Технології екранного доступу (Screen Readers) Це системні рішення, інтегровані в операційні системи мобільних пристроїв: VoiceOver (iOS) та TalkBack (Android). Вони забезпечують невізуальну взаємодію з інтерфейсом смартфона, озвучуючи елементи екрана при дотику. Хоча ці інструменти є стандартом де-факто для керування гаджетом, вони не вирішують безпосередньо задачу просторової навігації та детекції фізичних перешкод на шляху користувача.

3. Системи комп'ютерного зору (AI-based solutions) Додаток Seeing AI від компанії Microsoft є прикладом використання штучного інтелекту для асистивних цілей. Програма використовує неймережеві алгоритми для обробки зображень з камери смартфона в реальному часі.

Функціонал включає:

- Розпізнавання тексту (OCR) та його озвучення;
- Ідентифікацію об'єктів та людей (включно з емоціями та віком);
- Опис сцени навколишнього середовища (рис. 1.4) [8].



Рисунок 1.4 – Приклад роботи Seeing AI

У категорії рішень на базі комп'ютерного зору, окрім Seeing AI, варто виділити Airoly Vision. Цей програмний продукт використовує алгоритми глибокого навчання (Deep Learning) для ідентифікації об'єктів у реальному часі. Особливістю застосунку є його спеціалізація на побутових сценаріях: він здатен розпізнавати номінали грошових банкнот, ідентифікувати продукти харчування в магазинах, а також визначати кольорову гаму предметів. Результат обробки миттєво передається користувачеві через аудіовідгук, що дозволяє незрячим здійснювати покупки та сортувати речі без сторонньої допомоги.

Окремою групою є спеціалізовані навігаційні системи, яскравим представником яких є BlindSquare. На відміну від стандартних картографічних сервісів (як Google Maps), цей застосунок розроблено спеціально для потреб осіб з порушеннями зору. Він функціонує як «аудіо-гід», використовуючи дані GPS та базу даних Foursquare для опису навколишньої інфраструктури.

Основні функції BlindSquare:

- Опис контексту: Оголошення назв вулиць, перехресть та закладів (POI — Points of Interest) по мірі руху користувача.
- Просторова орієнтація: Вказівка напрямку та відстані до вибраних об'єктів за принципом годинникового циферблата (наприклад, «Аптека на 2-гу годину, 50 метрів»).

Обмеження мобільних застосунків. Попри високу функціональність, використання виключно смартфона для навігації має ряд недоліків:

1. Зайнятість рук: Для сканування простору користувач змушений тримати телефон перед собою, що унеможливорює повноцінне використання білої тростини або собаки-поводиря.
2. Обмежений кут огляду: Камера телефону бачить лише вузький сектор простору, тоді як для безпеки потрібен моніторинг на 360° або хоча б широка панорама.
3. Енергозалежність: Постійна робота камери та процесора для обробки нейромереж швидко виснажує акумулятор телефону, залишаючи людину без зв'язку.

1.3. Обґрунтування вибору методів детекції та тактильного зворотного зв'язку

Аналіз наукової літератури свідчить про зміщення акцентів у розробці асистивних систем від простих механічних засобів до комплексних електронних рішень. Дослідження [9] підтверджують, що інтеграція камер, ультразвукових та інфрачервоних сенсорів дозволяє сформувати детальну картину навколишнього середовища, що критично впливає на рівень мобільності та впевненості незрячих осіб.

Сенсорне забезпечення (Input) Ключовим фактором безпеки є зона охоплення сенсорної системи. У роботах Кумара та співавторів [10] обґрунтовано необхідність кругового моніторингу простору (360°). Традиційні засоби часто мають вузький кут огляду, залишаючи користувача вразливим до загроз з боків або позаду. Впровадження кругового масиву ультразвукових сенсорів дозволяє нівелювати цей недолік, створюючи «захисний периметр» навколо людини.

Для підвищення точності на малих відстанях доцільним є використання інфрачервоних (ІЧ) далекомірів. Зокрема, у дослідженнях Джонсона [12] детально проаналізовано характеристики сенсорів типу Sharp GP2Y0A21YK.

Автори зазначають, що ІЧ-сенсори мають високу швидкість реакції та відмінну роздільну здатність у ближній зоні (10–80 см), що робить їх ідеальним доповненням до ультразвукових систем для детекції дрібних перешкод або об'єктів складної форми.

Організація зворотного зв'язку (Output) Питання передачі інформації користувачеві є не менш важливим за її збір. Наукова спільнота схиляється до думки, що тактильний (вібраційний) канал є найбільш ергономічним для навігації. У роботах [11] зазначається, що шкіра є ефективним приймачем інформації: легкі вібраційні сигнали дозволяють інтуїтивно розуміти розташування перешкод без залучення слуху.

Фішер і Мюллер [14] у своїх працях розвивають концепцію мультимодальних інтерфейсів, пропонуючи інтеграцію тактильних актуаторів безпосередньо з ультразвуковими модулями. Такий підхід («де відчув вібрацію — там і перешкода») мінімізує час реакції людини. Браун та ін. [13] вказують, що хоча сучасні «розумні тростини» використовують подібні технології, їх обмеженням є локалізація датчиків лише в одній точці (руці), що підтверджує перспективність створення носних систем (поясів, жилетів) для комплексного захисту.

1.4. Порівняльний аналіз та обґрунтування вибору алгоритму комп'ютерного зору

Однією з ключових задач розроблюваної системи є семантичне розпізнавання об'єктів у реальному часі. Ця функція реалізується методами комп'ютерного зору (Computer Vision), які дозволяють трансформувати відеопотік у структуровану інформацію про тип та розташування перешкод.

У сучасній практиці виділяють два основні підходи до вирішення задачі детекції об'єктів:

1. Традиційні методи (Classic Computer Vision): Базуються на алгоритмах виділення ознак, таких як SIFT (Scale-Invariant Feature Transform) та метод Віоли-Джонса (Viola-Jones). Ці алгоритми використовують «ручне конструювання ознак» (hand-crafted features). Хоча вони є менш вимогливими до обчислювальних ресурсів, їхня ефективність суттєво знижується у складних умовах освітлення або при зміні ракурсу об'єкта, що є критичним недоліком для асистивних пристроїв.
2. Методи глибокого навчання (Deep Learning): Використовують згорткові нейронні мережі (CNN) для автоматичного виділення ознак. Вони демонструють високу адаптивність (робастність) до змін зовнішнього середовища, значно переважаючи традиційні методи за точністю класифікації.

Враховуючи вимоги до мобільної системи (обмежені обчислювальні ресурси та необхідність роботи в реальному часі), для детального аналізу було обрано три найпоширеніші архітектури нейромереж: YOLO, Faster R-CNN та SSD.

Характеристика розглянутих алгоритмів:

- Faster R-CNN (Region-based Convolutional Neural Networks): Належить до класу двостадійних детекторів. На першому етапі мережа генерує регіони інтересу (Region Proposals), а на другому — класифікує їх. Такий підхід забезпечує найвищу точність (~70% mAP), проте швидкість обробки складає лише близько 7 кадрів за секунду (FPS), що є недостатнім для динамічної навігації [16].
- SSD (Single Shot Detector): Одностадійний детектор, який виконує локалізацію та класифікацію за один прохід мережі. Він забезпечує компроміс між швидкістю та точністю, проте демонструє зниження ефективності при розпізнаванні дрібних об'єктів порівняно з аналогами [17].
- YOLO (You Only Look Once): Архітектура, що розглядає задачу детекції як єдину проблему регресії. Особливої уваги заслуговує модифікація YOLOv4-Tiny (та новіші Nano-версії). Вона оптимізована для роботи на слабких пристроях (Edge Devices) і здатна забезпечувати швидкість до 220 FPS [15], що є критично важливим для миттєвого попередження незрячого користувача про небезпеку.

Порівняльні характеристики алгоритмів наведено в таблиці 1.2.

Таблиця 1.2 – Порівняльний аналіз алгоритмів детекції об'єктів

Алгоритм	Швидкість (FPS)	Точність (mAP)*	Вимоги до ресурсів	Переваги	Недоліки
YOLO (Tiny версії)	До 220	55–65%	Низькі	Робота в реальному часі, висока	Дещо менша точність детекції дрібних об'єктів

Алгоритм	Швидкість (FPS)	Точність (mAP)*	Вимоги до ресурсів	Переваги	Недоліки
				енергоефективність, ідеально для IoT	порівняно з R-CNN
SSD	~125	~50%	Середні	Баланс швидкості/точності	Гірша продуктивність на складних сценах
Faster R-CNN	~7	~70%	Високі (потужний GPU)	Найвища точність, надійне розпізнавання	Критично низька швидкість, високе енергоспоживання

*mAP (mean Average Precision) — усереднена точність, значення можуть варіюватися залежно від датасету (наприклад, COCO).

1.5. Теоретичні основи функціонування нейронної мережі YOLO

1.5.1. Базові принципи та архітектура

Алгоритм YOLO (You Only Look Once) представляє собою фундаментальну зміну парадигми у сфері комп'ютерного зору. На відміну від традиційних підходів, які адаптують класифікатори для завдань детекції, YOLO розглядає виявлення об'єктів як єдину задачу регресії (regression problem). Це дозволяє системі переходити безпосередньо від пікселів вхідного зображення до координат обмежувальних рамок (bounding boxes) та ймовірностей класів об'єктів.

Еволюція методів детекції Для розуміння переваг YOLO доцільно порівняти його з попередніми поколіннями алгоритмів:

1. Метод ковзного вікна (Sliding Window): Використовувався в системах на основі моделей з деформованими компонентами (DPM) [19]. Класифікатор запускався послідовно на тисячах локальних ділянок зображення, що вимагало значних обчислювальних ресурсів.

2. Методи на основі регіонів (R-CNN): Використовують складний багатоступеневий конвеєр (pipeline). Спершу генеруються потенційні регіони інтересу (Region Proposals), потім виконується їх класифікація, і, нарешті, уточнюються межі рамок. Як зазначається у джерелі [20], такий підхід є повільним і важким для оптимізації, оскільки кожен компонент системи потребує окремого тренування.

Інновація YOLO Архітектура YOLO об'єднує всі етапи обробки в єдину згорткову нейронну мережу (Single Convolutional Network). Система ділить вхідне зображення на сітку (grid) і для кожної комірки одночасно прогнозує:

- Координати обмежувальних рамок (B-boxes);
- Впевненість у наявності об'єкта (Confidence Score);
- Ймовірність приналежності об'єкта до певного класу.

Така уніфікація дозволяє алгоритму «бачити» все зображення цілком (глобальний контекст), а не лише його фрагменти, що мінімізує кількість помилок типу «Background False Positives» (прийняття фону за об'єкт). Принципова схема роботи алгоритму наведена на рис. 1.6.

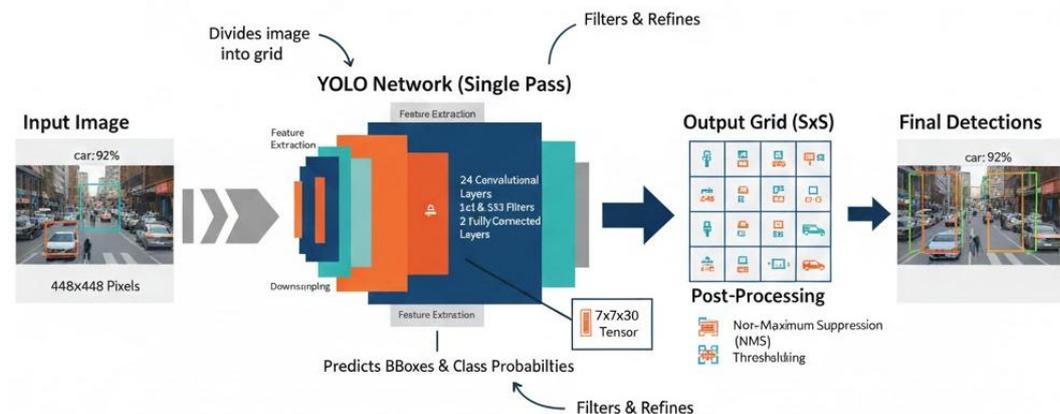


Рисунок 1.6 – Концептуальна схема роботи алгоритму YOLO: від вхідного зображення до детекції об'єктів за один прохід

Для глибшого розуміння переваг обраного методу варто детальніше розглянути алгоритмічну структуру альтернативних підходів. Сучасні детектори

на базі регіонів (зокрема, сімейство R-CNN) функціонують за принципом багатоетапного конвеєра:

1. Генерація гіпотез: Спочатку алгоритм створює набір потенційних обмежувальних рамок (Region Proposals).
2. Класифікація: Для кожного виділеного регіону запускається окремий класифікатор.
3. Постобробка (Post-processing): Включає уточнення координат рамок, фільтрацію дублікатів (Non-Maximum Suppression) та контекстну переоцінку сцени.

Головним недоліком такого підходу є його фрагментарність: кожен етап вимагає окремого налаштування та навчання, що унеможлиблює наскрізну (end-to-end) оптимізацію системи. Це створює значні затримки в обробці даних, що є критичним для систем реального часу.

Натомість YOLO (You Only Look Once) реалізує концептуально інший підхід, трансформуючи задачу детекції в єдину задачу регресії. Нейромережа аналізує вхідне зображення цілісно, одночасно прогнозуючи координати обмежувальних рамок та ймовірності класів для них. Назва алгоритму відображає його ключову особливість: для повної аналітики сцени достатньо одного проходу («погляду») нейромережі.

Саме архітектурна простота та висока швидкість інференсу (виведення результату) стали визначальними факторами при виборі YOLO для розроблюваного асистивного пристрою. В умовах реальної експлуатації, де перешкоди виникають динамічно, а час на прийняття рішення вимірюється мілісекундами, затримки, притаманні R-CNN, є неприпустимими з точки зору безпеки користувача.

Порівнюючи YOLO та SSD, варто зазначити, що хоча SSD інколи демонструє кращу точність на дрібних об'єктах завдяки використанню карт ознак різних масштабів, YOLO забезпечує кращий баланс між швидкістю та точністю, що є пріоритетним для задачі орієнтування в динамічному середовищі. Враховуючи необхідність миттєвої реакції системи на появу автомобілів чи

пішоходів, для реалізації програмної частини комплексу доцільно обрати саме архітектуру сімейства YOLO, яка дозволяє мінімізувати апаратні затримки без критичної втрати якості розпізнавання.

1.5.3. Ключові переваги та обмеження використання алгоритму YOLO

Вибір архітектури YOLO для реалізації системи машинного зору в асистивному пристрої обумовлений рядом технічних переваг, що вигідно відрізняють її від аналогів:

1. Висока швидкодія та низька затримка (Low Latency). Завдяки регресійному підходу, алгоритм демонструє виняткову швидкість інференсу. Базова модель здатна обробляти відеопотік зі швидкістю 45 кадрів за секунду (FPS), тоді як оптимізовані версії (наприклад, Tiny-YOLO) досягають показників понад 150 FPS. Це забезпечує затримку обробки менше 25 мс, що дозволяє системі працювати в режимі «м'якого реального часу» (soft real-time), миттєво попереджаючи користувача про динамічні зміни в оточенні.
2. Глобальний контекстний аналіз. На відміну від методів ковзного вікна, які «бачать» лише фрагменти зображення, YOLO аналізує весь кадр цілком під час навчання та тестування. Це дозволяє нейромережі враховувати контекстну інформацію про класи об'єктів та їх оточення. Як наслідок, значно знижується ймовірність виникнення помилок типу «Background False Positives» (коли фонові шуми помилково класифікуються як об'єкти), що є критичним для безпеки незрячих.
3. Висока узагальнююча здатність (Generalization). Алгоритм демонструє стійкість до варіативності вхідних даних. Навчаючись на фотореалістичних зображеннях, система зберігає високу ефективність розпізнавання навіть у нових, нетипових умовах або при зміні візуальних доменів (наприклад, на схематичних зображеннях). Це робить пристрій надійним при експлуатації в різних умовах освітлення та середовищах.

Обмеження методу. Варто зазначити, що пріоритет швидкості накладає певні обмеження на точність локалізації, особливо для дрібних об'єктів, де YOLO може дещо поступатися важким архітектурам типу Faster R-CNN. Однак, у контексті асистивних технологій для навігації, де критично важливим є факт своєчасного виявлення великих перешкод (автомобілів, людей, стін), а не піксельна точність їх контурів, цей компроміс є повністю виправданим.

1.6 Огляд нормативних вимог та стандартів до асистивних технологій

Процес розробки електронних засобів реабілітації вимагає суворого дотримання міжнародних та національних стандартів, які регламентують аспекти безпеки, ергономіки та функціональної придатності пристроїв. Це необхідно для забезпечення надійності виробу та можливості його подальшої сертифікації. Базовим документом у цій сфері є міжнародний стандарт ISO 9999 «Assistive products for persons with disability — Classification and terminology». Відповідно до його класифікації, розроблюваний пристрій належить до класу 12, який охоплює засоби допомоги для орієнтування та мобільності. Зокрема, мова йде про категорію електронних засобів, що використовують сенсорні технології для розширення сприйняття навколишнього середовища, виявляючи перешкоди поза зоною фізичного контакту стандартної білої тростини.

Окрему увагу при проектуванні інтерфейсу взаємодії з користувачем слід приділити стандарту ISO 23600, який, хоча і стосується здебільшого акустичних сигналів світлофорів, встановлює універсальні принципи для тактильних та аудіальних сповіщень. Головною вимогою є розрізнюваність сигналу на фоні навколишнього шуму та його інтуїтивність. Для тактильного зворотного зв'язку це означає необхідність модуляції інтенсивності вібрації залежно від ступеня небезпеки: чим ближче перешкода, тим інтенсивнішим має бути сигнал. При використанні звукового оповіщення рівень сигналу повинен бути достатнім для сприйняття в умовах міського шуму, але не перевищувати больовий поріг слуху.

Крім того, оскільки система базується на алгоритмах штучного інтелекту, до неї застосовуються сучасні принципи надійності програмного забезпечення (Trustworthy AI). Критично важливим є забезпечення так званого fail-safe режиму роботи. Це означає, що у випадку відмови оптичного модуля, розряду акумулятора або збою в роботі нейромережі, пристрій не повинен вводити користувача в оману хибними сигналами безпеки. У такій ситуації система має або чітко сигналізувати про несправність, або перейти в пасивний режим, дозволяючи використовувати виріб як класичну механічну тростину.

1.7 Висновки до розділу 1

У першому розділі магістерської роботи здійснено комплексний аналіз стану проблеми просторової орієнтації осіб із вадами зору та огляд існуючих технічних засобів навігації. Основні результати теоретичного дослідження дозволяють зробити такі висновки:

1. Аналіз існуючих рішень. Встановлено, що сучасні асистивні пристрої (електронні тростини, навігаційні пояси) та мобільні застосунки мають суттєві експлуатаційні обмеження. Головними недоліками є висока вартість, залежність від стабільного інтернет-з'єднання та перевантаження слухового каналу користувача, що знижує загальний рівень безпеки пересування.
2. Обґрунтування архітектури системи. Доведено перспективність створення гібридного пристрою, який поєднує ультразвукову локацію (для точного визначення відстані до перешкод у радіусі 360°) та комп'ютерний зір (для семантичного розуміння сцени). Така комбінація дозволяє нівелювати «сліпі зони» та забезпечити мультимодальне сприйняття простору.
3. Вибір алгоритмічного забезпечення. На основі порівняльного аналізу методів детекції об'єктів (R-CNN, SSD, YOLO) визначено, що оптимальним рішенням для системи реального часу є нейронна мережа

архітектури YOLO (You Only Look Once). Завдяки високій швидкості інференсу та регресійному підходу, цей алгоритм здатен забезпечити миттєве попередження про небезпеку навіть на портативних обчислювальних платформах.

4. Організація зворотного зв'язку. Обґрунтовано доцільність використання тактильного (вібраційного) інтерфейсу як основного каналу комунікації. Це дозволяє розвантажити слух користувача для контролю навколишнього середовища, залишаючи голосові сповіщення лише для критично важливих ситуацій.

Таким чином, результати першого розділу формують теоретико-методологічну базу для практичної розробки апаратно-програмного комплексу, опис якої наведено у другому розділі.

2 РОЗРОБКА АСИСТИВНОГО ПРИСТРОЮ НА ОСНОВІ ОПТОЕЛЕКТРОННИХ АНАЛІЗАТОРІВ

2.1. Обґрунтування вибору компонентної бази та архітектура системи

2.1.1. Аналіз основних складових апаратної частини

Процес проектування апаратної складової системи розпочався з розробки концептуальної моделі взаємодії компонентів. Для візуалізації логіки функціонування пристрою та потоків даних між його вузлами було розроблено структурну блок-схему в середовищі Lucidchart, яка представлена на рисунку 2.1. Дана схема відображає ієрархію підключення сенсорів та виконавчих механізмів до центрального обчислювального ядра.

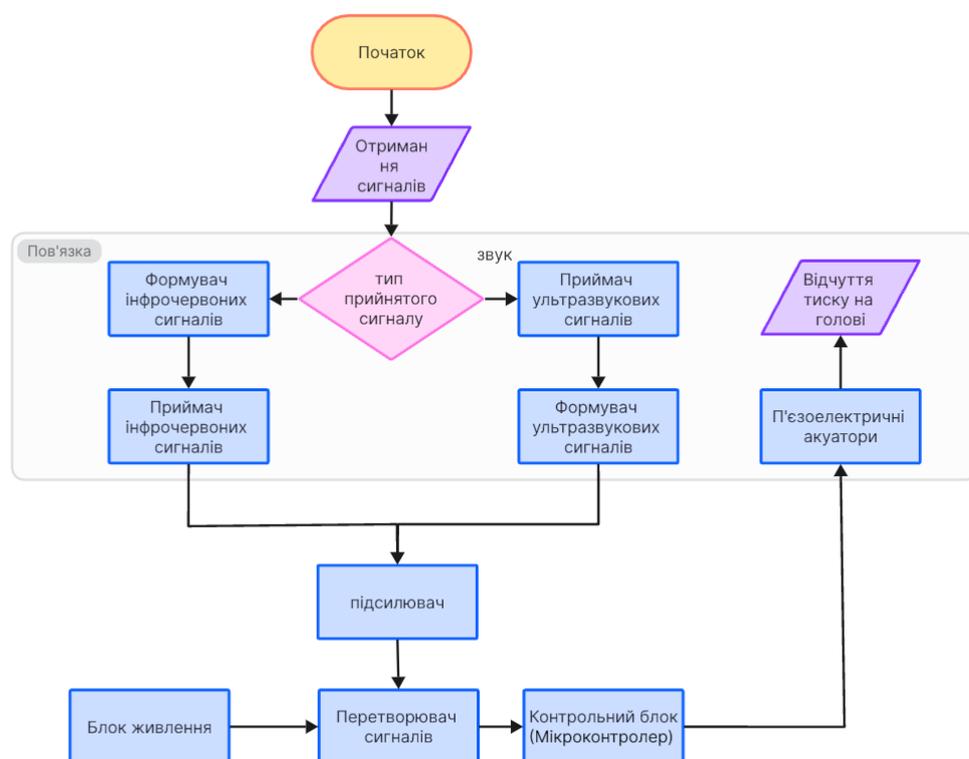


Рисунок 2.1 – Структурна блок-схема асистивного пристрою

Конструктивно пристрій реалізовано у вигляді ергономічного носного модуля, що фіксується на голові користувача за допомогою спеціальної стрічки.

Такий форм-фактор обрано з метою забезпечення максимального комфорту при тривалій експлуатації, а також для розміщення сенсорів на оптимальній висоті, що відповідає рівню очей людини. На зовнішньому периметрі стрічки закріплено систему детекції перешкод, яка складається з чотирьох ультразвукових далекомірів моделі HC-SR04. Сенсори розміщені рівномірно по колу, що дозволяє реалізувати концепцію кругового огляду на 360 градусів, мінімізуючи «сліпі зони» навколо користувача. Кожен окремий датчик відповідає за моніторинг свого сектора простору, передаючи дані про наявність об'єктів у реальному часі.

Реалізація зворотного зв'язку здійснюється через тактильний канал сприйняття. На внутрішній поверхні стрічки, безпосередньо навпроти кожного ультразвукового сенсора, інтегровано вібраційні мотори. Логіка роботи системи передбачає пряму кореляцію між відстанню до перешкоди та інтенсивністю вібрації: наближення до об'єкта викликає посилення та почастішання тактильних сигналів у відповідному секторі. Це дозволяє користувачеві інтуїтивно визначати не лише наявність перешкоди, але і її точний напрямок та ступінь загрози зіткнення.

Роль центрального керуючого елемента виконує мікроконтролер Arduino Nano. Ця платформа була обрана завдяки своїм компактним розмірам, низькому енергоспоживанню та достатній обчислювальній потужності для паралельної обробки сигналів від чотирьох сенсорів і керування вібромоторами. Для забезпечення автономності пристрою використовується акумуляторна батарея, яка кріпиться на потиличній частині стрічки. Таке розміщення елемента живлення виконує подвійну функцію: забезпечує енергонезалежність системи та слугує противагою для фронтальних сенсорів, що сприяє правильному балансуванню пристрою та знижує навантаження на шийний відділ користувача.

2.1.2. Специфікація елементної бази та обґрунтування вибору компонентів

Деталізація апаратної реалізації вимагає обґрунтування вибору кожного елемента схемотехнічного рішення, виходячи з критеріїв надійності, енергоефективності та функціональної сумісності. Роль центрального обчислювального вузла системи відведено налагоджувальній платі Arduino UNO R3. Її архітектура, що включає 14 цифрових портів вводу-виводу та 6 аналогових входів, дозволяє без додаткових розширювачів підключити необхідний масив сенсорів. Тактова частота процесора 16 МГц є достатньою для забезпечення низької латентності при обробці сигналів, що є критичним для систем реального часу.

Сенсорна підсистема спроектована за гібридним принципом для підвищення інформативності. Основним засобом навігації обрано ультразвукові модулі HC-SR04P, які забезпечують високу точність вимірювання відстані (похибка ± 3 мм) у діапазоні від 2 до 400 см. Для розширення можливостей виявлення динамічних об'єктів (людей, тварин) додатково впроваджено пасивний інфрачервоний датчик PIR SR501. Завдяки широкому куту огляду (110°) та здатності реагувати на теплові сигнатури на відстані до 7 метрів, він додає системі функцію завчасного попередження про наближення живих об'єктів.

Реалізація тактильного інтерфейсу базується на використанні мініатюрних вібромоторів діаметром 12 мм зі швидкістю обертання 9000 об/хв, що генерують відчутний для шкіри сигнал. Оскільки пряме підключення індуктивного навантаження до логічних виходів мікроконтролера є технічно неприпустимим, керування моторами здійснюється через силові ключі на базі біполярних транзисторів 2N2222. Вони здатні комутувати струми до 800 мА, забезпечуючи надійну роботу виконавчих механізмів без ризику пошкодження плати керування. Енергетична автономність комплексу гарантується використанням портативного джерела живлення (Powerbank) ємністю 10 000 мА·год, що дозволяє пристрою функціонувати протягом тривалого часу без підзарядки.

Зведені технічні характеристики обраних компонентів представлено в таблиці 2.1.

Таблиця 2.1 – Специфікація складових компонентів розроблюваного пристрою

Компонент	Модель / Тип	Основні технічні характеристики	Обґрунтування вибору
Мікроконтролер	Arduino UNO R3	14 цифрових I/O, 6 аналогових входів, частота 16 МГц	Оптимальний баланс між продуктивністю та кількістю портів для підключення всієї периферії.
Ультразвуковий датчик	HC-SR04P	Діапазон: 2–400 см, точність: ± 3 мм	Забезпечує високоточну локацію перешкод, працює стабільно за різних умов освітлення.
ІЧ-датчик руху	PIR SR501	Кут огляду: 110° , дальність: до 7 м	Дозволяє детектувати рухомі теплові об'єкти (людей), підвищуючи рівень безпеки.
Вібромотор	Coin Type 12 мм	Напруга: 3 В, 9000 об./хв	Компактний розмір та висока інтенсивність вібрації для чіткого тактильного відгуку.
Транзистор	2N2222 (NPN)	Струм колектора: 800 мА, U_{max} : 40 В	Необхідний для посилення струму керування вібромоторами та захисту портів мікроконтролера.
Комутація	Проводи Dupont	Довжина 20 см, типи М-М, М-F	Забезпечують швидке та надійне макетування з'єднань без пайки на етапі прототипування.
Живлення	Powerbank	Ємність $>10\ 000$ mAh, вихід 5V	Гарантує мобільність системи та тривалий час автономної роботи в польових умовах.

2.1.3 Обґрунтування вибору оптичного сенсора для системи технічного зору

Інтеграція підсистеми комп'ютерного зору є необхідною умовою для реалізації функції семантичного розпізнавання об'єктів нейромережею. Вибір оптичного модуля базується на комплексному аналізі чотирьох критичних параметрів: роздільної здатності, світлочутливості матриці, швидкодії (Frame Rate) та наявності інфрачервоного спектра. Роздільна здатність сенсора безпосередньо впливає на точність детекції дрібних деталей; для задач навігації оптимальним вважається діапазон від 2 до 5 Мп [16], що дозволяє нейромережі коректно ідентифікувати об'єкти на середніх дистанціях без надмірного навантаження на канал передачі даних. Не менш важливим фактором є світлочутливість, особливо з використанням технології BSI (Back-Illuminated Sensor), яка забезпечує прийнятну якість зображення в сутінках [17]. Для роботи в режимі реального часу оптичний модуль повинен підтримувати частоту кадрів не менше 30 FPS, а наявність ІЧ-підсвічування є критичною для нічної навігації.

У рамках дослідження було проведено порівняльний аналіз трьох представників ринку машинного зору, що потенційно можуть бути використані в проєкті.

Першим кандидатом є камера Luxonis OAK-1 Lite. Це спеціалізований модуль для граничних обчислень (Edge AI), оснащений 13-мегапіксельним сенсором IMX214. Ключовою особливістю даної камери є наявність вбудованого візуального процесора (VPU) Intel Movidius Myriad X [17]. Цей чип дозволяє виконувати інференс нейронних мереж безпосередньо на борту камери, розвантажуючи центральний процесор основного пристрою. Високошвидкісний інтерфейс USB 3.0 забезпечує передачу потокового відео з мінімальною затримкою.

Як альтернативи розглядалися промислові рішення. Система Omron Xrestia FZD являє собою високоточний комплекс технічного зору, призначений для аналізу 2D та 3D об'єктів в умовах промислового виробництва [18]. Попри

наявність зручного інтерфейсу налаштування, цей пристрій орієнтований на стаціонарне використання. Аналогічною є камера SICK Visionary-B, яка спеціалізується на отриманні стереоскопічних 3D-зображень для навігації в складних умовах [18]. Вона відзначається високою надійністю та швидкістю обробки даних, проте її габарити та енергоспоживання розраховані на встановлення на транспортні засоби або робототехнічні комплекси, а не на носні пристрої.

Порівняльні характеристики розглянутих оптичних модулів зведено в таблицю 2.2.

Таблиця 2.2 – Порівняльний аналіз камер для системи технічного зору

Критерій	Luxonis OAK-1 Lite [17]	Omron Xpectia FZD [18]	SICK Visionary-B [18]
Роздільна здатність	13 Мп	5 Мп	3D-зображення (обмежена деталізація)
Тип даних	2D зображення	2D / 3D	3D карта глибини
Обчислювальні можливості	Високі (Вбудований VPU Intel Myriad X для нейромереж)	Середні (залежить від зовнішнього контролера)	Високі (оптимізовано під 3D)
Інтеграція ШІ	Пряма апаратна підтримка (Edge AI)	Потребує зовнішньої обробки	Потребує зовнішньої обробки
Інтерфейс	USB 3.0 (Plug&Play)	Ethernet, промислові шини	Ethernet
Сфера застосування	Мобільні пристрої, IoT, робототехніка	Промислові лінії контролю	Промислова навігація, навантажувачі

Критерій	Luxonis OAK-1 Lite [17]	Omron Xpectia FZD [18]	SICK Visionary-B [18]
Придатність для носіння	Висока (компактна, живлення від USB)	Низька (громіздка)	Низька (потребує потужного живлення)

Аналізуючи дані таблиці 2.2, можна зробити висновок, що для розробки мобільного асистивного пристрою безальтернативним варіантом є Luxonis OAK-1 Lite. На відміну від промислових рішень Omron та SICK, які вимагають складних інтерфейсів підключення та потужних джерел живлення, камера Luxonis поєднує в собі компактність, енергоефективність та, що найважливіше, здатність самостійно обробляти нейромережі. Це дозволяє реалізувати концепцію «розумного сенсора», що є критично важливим для портативної системи з обмеженим ресурсом акумулятора.

2.1.4. Обґрунтування вибору обчислювальної платформи для цифрової обробки зображень

Реалізація алгоритмів глибокого навчання, зокрема детекторів архітектури YOLO, висуває підвищені вимоги до апаратного забезпечення, які виходять за межі можливостей стандартних мікроконтролерів сімейства AVR (Arduino UNO/Nano). Обмежені ресурси оперативної пам'яті (SRAM) та низька тактова частота Arduino унеможливають обробку відеопотоку в реальному часі. У зв'язку з цим виникає необхідність інтеграції в систему додаткового обчислювального модуля. У рамках дослідження було проведено порівняльний аналіз двох найбільш поширених платформ: ESP32-CAM та одноплатного мікрокомп'ютера Raspberry Pi.

Модуль ESP32-CAM розглядається в літературі як оптимальне рішення для бюджетних IoT-пристроїв. Це компактна плата, що поєднує 32-бітний

мікроконтролер, інтерфейси Wi-Fi/Bluetooth та інтегрований модуль камери (зазвичай OV2640).

Завдяки низькому енергоспоживанню та мінімальним габаритам, ESP32-CAM є привабливим варіантом для портативної електроніки. Однак, його обчислювальна потужність дозволяє виконувати лише базові операції захоплення та потокової передачі відео, або ж запускати вкрай спрощені моделі нейромереж із низькою частотою кадрів, що є недостатнім для повноцінної навігації незрячих [20].

Альтернативою виступає використання мікрокомп'ютера Raspberry Pi 4 Model B. Ця платформа базується на високопродуктивному 4-ядерному процесорі ARM Cortex-A72 і може комплектуватися оперативною пам'яттю обсягом до 8 ГБ.

Така архітектура дозволяє розгорнути повноцінну операційну систему (Linux) та використовувати оптимізовані бібліотеки комп'ютерного зору (OpenCV, TensorFlow Lite). Саме Raspberry Pi забезпечує достатній ресурс для запуску нейромережі YOLO (особливо її Tiny-версій) у реальному часі. Крім того, наявність швидкісних портів USB 3.0 дозволяє підключати зовнішні інтелектуальні камери (наприклад, обрану раніше Luxonis OAK-1), перетворюючи Raspberry Pi на центральний хаб, який агрегує дані з камери та передає команди керування на мікроконтролер Arduino. Виходячи з вимог до швидкодії та необхідності запуску YOLO, для подальшої розробки обрано саме платформу Raspberry Pi.

2.2. Моделювання та віртуальна верифікація роботи пристрою

2.2.1. Симуляція роботи сенсорної підсистеми

Перед етапом фізичного макетування було проведено комп'ютерне моделювання ключових вузлів пристрою для перевірки коректності схемотехнічних рішень та алгоритмів керування. Як середовище для

віртуального прототипування використано програмний комплекс Tinkercad Circuits, який дозволяє емулювати роботу мікроконтролерів Arduino та підключеної периферії. Основною метою симуляції була перевірка логіки взаємодії між ультразвуковим далекоміром та системою тактильного зворотного зв'язку.

На створеній віртуальній моделі (рис. 2.2) реалізовано схему підключення ультразвукового сенсора HC-SR04 до цифрових портів контролера, а також силового каскаду керування вібромотором через транзисторний ключ.

У ході симуляції було відтворено сценарії зміни відстані до умовного об'єкта. Програмний алгоритм налаштовано таким чином, щоб зі зменшенням дистанції до перешкоди (яка імітується повзунком у середовищі Tinkercad) пропорційно змінювалася шпаруватість ШІМ-сигналу (широтно-імпульсної модуляції), що подається на базу керуючого транзистора. Результати моделювання підтвердили коректність обраної схемотехніки: при наближенні об'єкта в зону детекції спостерігається активація двигуна, а зміна параметрів сигналу відповідає закладеній математичній моделі, що дозволяє переходити до етапу фізичної збірки прототипу.

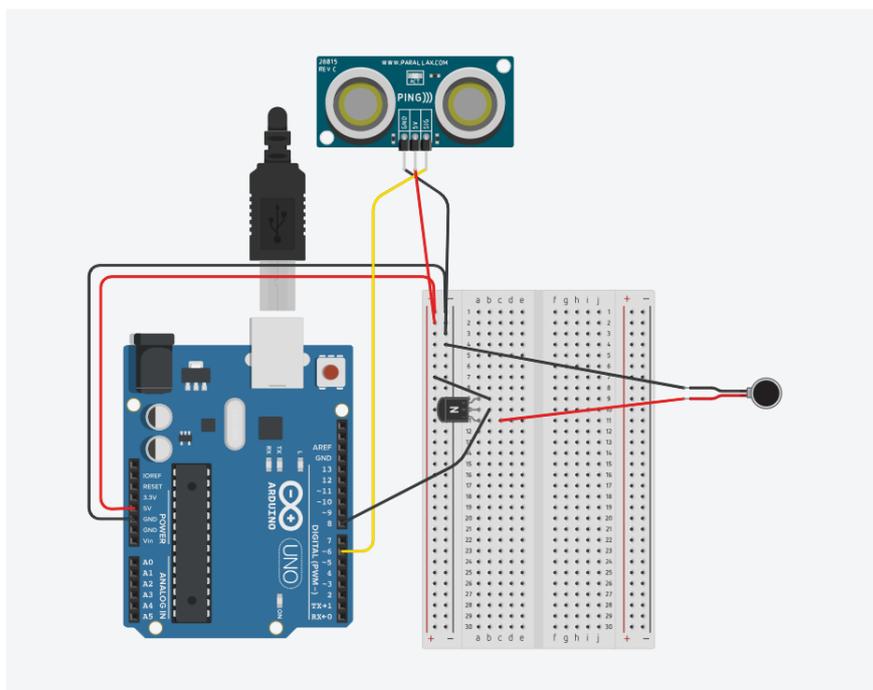


Рисунок 2.2 – Модель створена у *Tinkercad* для тестування основного принципу роботи приладу

2.2.2. Алгоритмічна реалізація базового прототипу

На початковому етапі віртуального тестування було верифіковано алгоритм обробки даних від одиночного сенсорного вузла. Розроблена логічна схема функціонує за пороговим принципом: активація системи оповіщення відбувається лише за умови, якщо виміряна дистанція до перешкоди є меншою за критичне значення $D_{\{threshold\}} = 100$ см.

Програмна реалізація (повний лістинг коду наведено у Додатку А) базується на циклічній роботі функції `readUltrasonicDistance`, яка ініціює випромінювання ультразвукового імпульсу через тригерний вихід (Trigger Pin) та фіксує часовий інтервал повернення відбитого сигналу на вхід приймача (Echo Pin). На основі отриманого часу мікроконтролер обчислює фізичну відстань до об'єкта.

Головний цикл програми `loop()` здійснює безперервний моніторинг показників далекоміра. При входженні об'єкта в зону детекції ($D < 100$ см.

) генерується сигнал широтно-імпульсної модуляції (ШІМ) для керування вібромотором. Реалізовано обернено-пропорційну залежність: зменшення дистанції призводить до підвищення коефіцієнта заповнення ШІМ-сигналу, що суб'єктивно сприймається користувачем як зростання інтенсивності вібрації.

2.2.3. Масштабування системи для кругового огляду (360°)

Наступним кроком стало розширення моделі до повнофункціональної конфігурації, що забезпечує всебічний моніторинг простору. У середовищі Tinkercad було розроблено віртуальний макет (рис. 2.3), що інтегрує чотири незалежні ультразвукові сенсори HC-SR04, орієнтовані на чотири сторони світу, та масив із восьми вібраційних моторів.

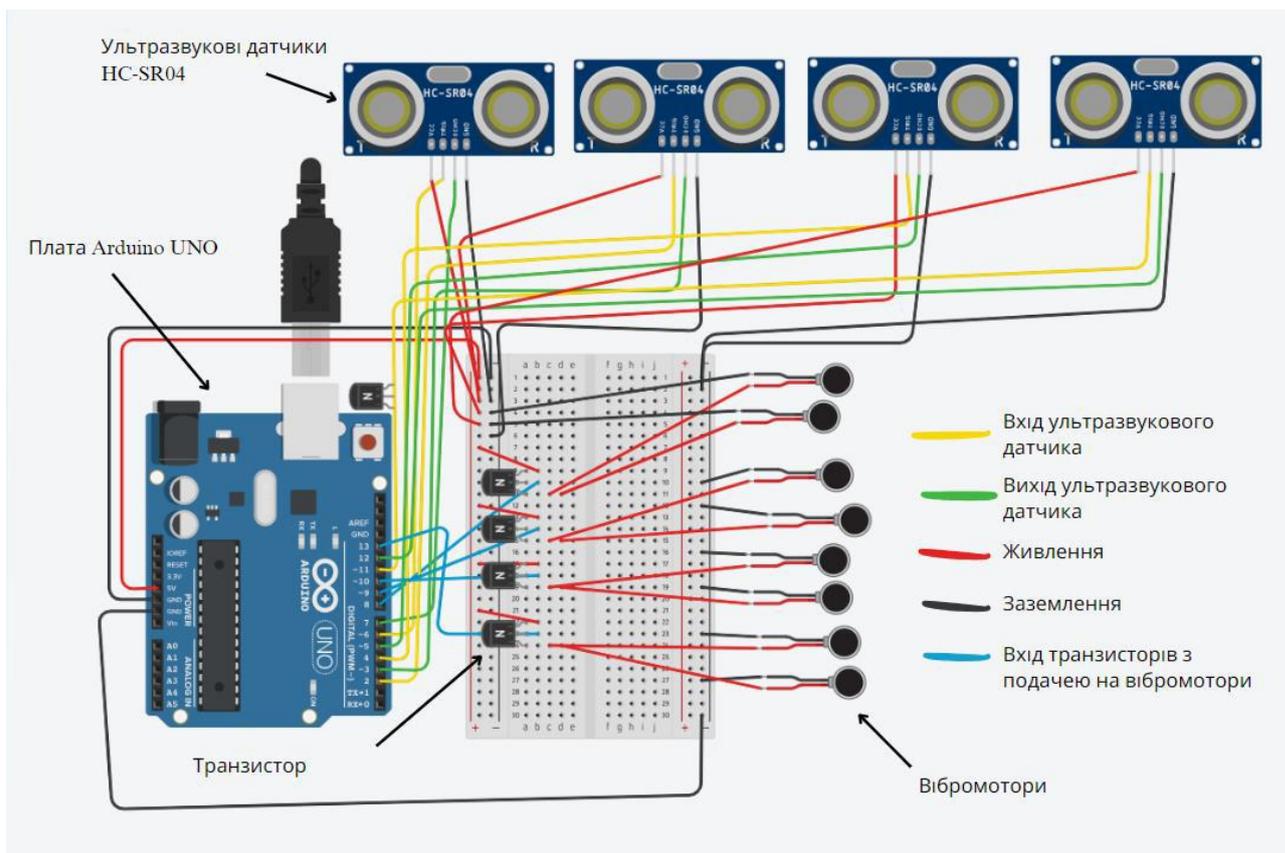


Рисунок 2.3 – Віртуальна модель мультисенсорної системи в середовищі Tinkercad

Кожен сенсорний модуль працює автономно, скануючи свій сектор відповідальності. Мікроконтролер Arduino UNO здійснює послідовне опитування датчиків та адресне керування відповідними групами вібраторів. Для забезпечення стабільності роботи силового контуру в емуляції використано окреме джерело живлення для виконавчих механізмів, гальванічно розв'язане з логічною частиною схеми через транзисторні ключі. Це дозволило уникнути просідання напруги на мікроконтролері під час пікових навантажень (старту моторів). Програмний код для багатоканальної версії пристрою наведено у Додатку В.

2.2.4. Оптимізація силової частини схеми (інтеграція драйверів L293D)

Аналіз схемотехніки попереднього прототипу виявив складність монтажу великої кількості дискретних транзисторів. Для підвищення надійності та

компактності пристрою було прийнято рішення про модернізацію вихідного каскаду шляхом впровадження інтегральних драйверів двигунів L293D (H-Bridge Driver).

Фінальна схема пристрою, представлена на рис. 2.4, передбачає заміну масиву транзисторів на дві мікросхеми L293D. Таке рішення обумовлене обмеженою здатністю портів Arduino видавати струм, достатній для прямого живлення індуктивних навантажень.

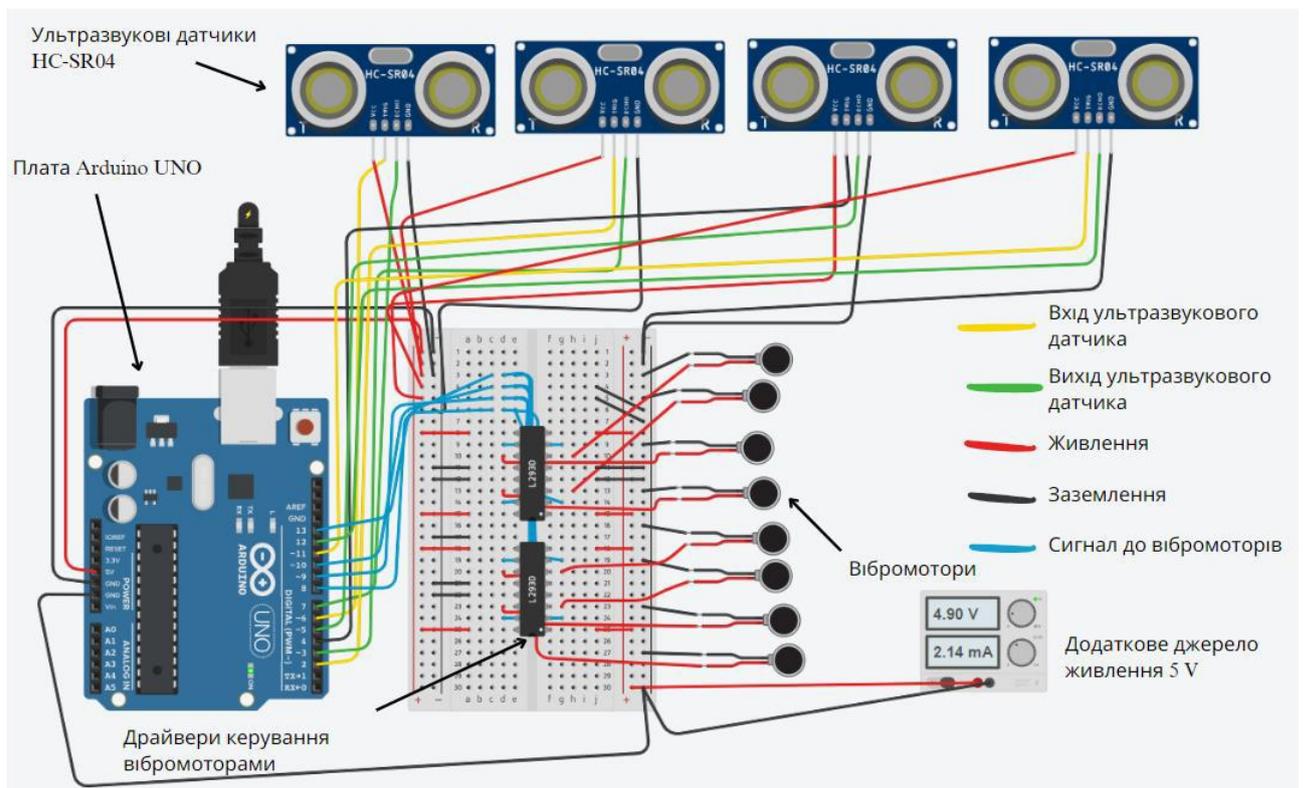


Рисунок 2.4 – Фінальна принципова схема з використанням драйверів L293D

Технічне обґрунтування використання L293D:

1. Струмове навантаження: Драйвер дозволяє комутувати струми до 600 мА на канал (з піковим навантаженням до 1.2 А), що є достатнім для живлення обраних віброторів, водночас захищаючи логічні виходи мікроконтролера.
2. Багатоканальність: Кожен корпус мікросхеми містить чотири канали посилення, що дозволяє керувати двома моторами незалежно або чотирма

в однонаправленому режимі. Використання двох мікросхем повністю покриває потреби системи з 8 моторів.

3. Логіка керування:

- Входи Enable (1,2EN та 3,4EN) використовуються для глобального дозволу роботи відповідних каналів (підключені до логічної одиниці +5В).
- Входи Input 1–4 приймають логічні сигнали від Arduino.
- Виходи Output 1–4 забезпечують подачу силової напруги від зовнішнього джерела живлення безпосередньо на клеми моторів.

Така архітектура значно спрощує друковану плату та підвищує відмовостійкість системи за рахунок наявності вбудованих захисних діодів у структурі драйвера L293D.

2.2.5. Модифікація алгоритму керування та реалізація адаптивних режимів роботи

Важливим етапом програмної розробки стала відмова від дискретного (бінарного) принципу керування вібромоторами на користь пропорційного регулювання. Для забезпечення інтуїтивно зрозумілого зворотного зв'язку реалізовано алгоритм на базі широтно-імпульсної модуляції (ШИМ, PWM). Використання функції `analogWrite` дозволяє варіювати коефіцієнт заповнення керуючого сигналу, змінюючи таким чином середню напругу на обмотках мотора і, відповідно, інтенсивність вібрації. Програмна логіка базується на функції лінійної інтерполяції (`map`), яка конвертує виміряну дистанцію у діапазоні 0–100 см в обернено-пропорційне значення потужності (від 255 до 0). Це створює ефект «градієнта відчуттів»: при наближенні до перешкоди сила вібрації плавно зростає, дозволяючи користувачеві «відчувати» відстань, а не просто отримувати сигнал тривоги.

З метою підвищення універсальності пристрою та його адаптації до різних сценаріїв експлуатації було розроблено архітектуру, що підтримує три

функціональні режими роботи. Базовим є режим «Стандартний» (General Orientation), у якому всі сенсори активні одночасно і забезпечують круговий моніторинг простору з лінійною залежністю сили вібрації від відстані, що є оптимальним для повсякденного пересування. Для навігації у вузьких приміщеннях (коридорах, дверних отворах) передбачено режим «Цільового аналізу» (Targeted Analysis), який програмно деактивує бічні та задні сенсори, фокусуючи увагу системи виключно на фронтальному секторі, що дозволяє уникнути надлишкового інформаційного шуму. Третім, найбільш складним алгоритмічно, є режим «Динамічного моніторингу», спрямований на підвищення безпеки. У цьому режимі мікроконтролер аналізує не лише миттєве значення відстані, а й її похідну за часом (швидкість зближення). Якщо система фіксує різке скорочення дистанції, інтенсивність сигналу примусово підвищується до максимуму, попереджаючи користувача про об'єкт, що швидко наближається. Візуалізація залежності інтенсивності вібраційного відгуку від відстані для різних режимів наведена на графіку (рис. 2.5).

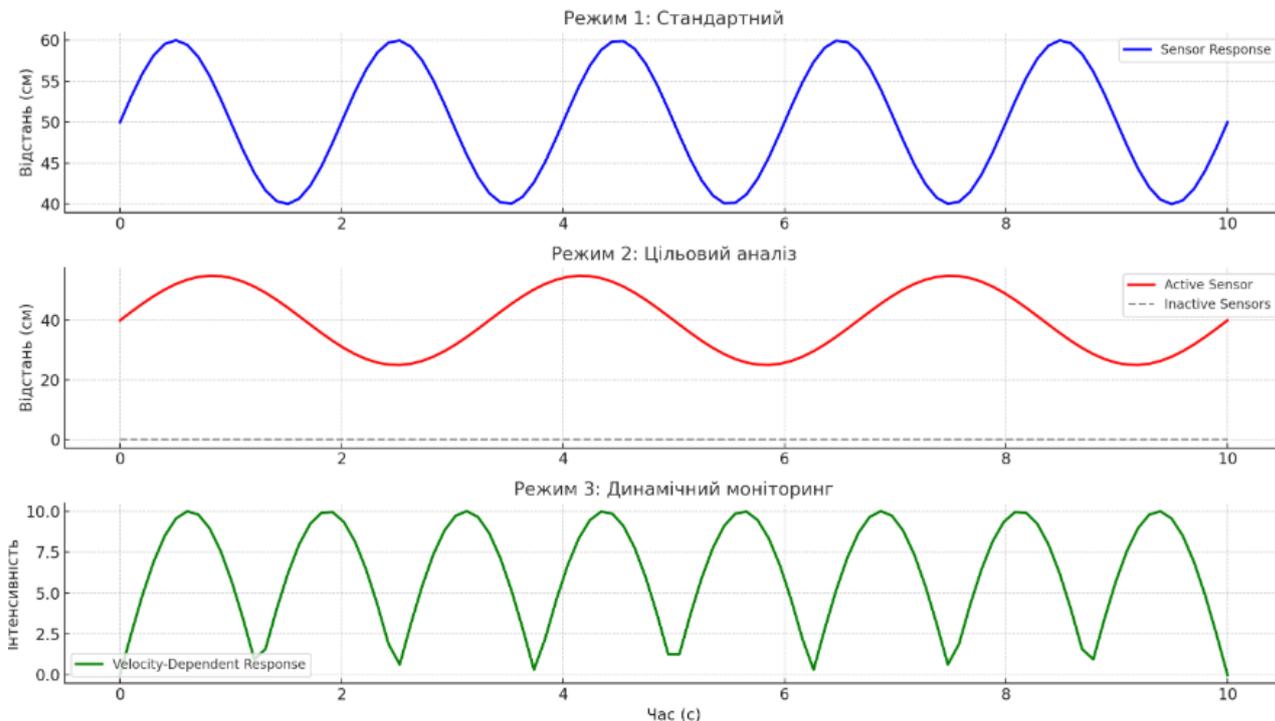


Рисунок 2.5 – Графік залежності інтенсивності тактильного сигналу від відстані в різних режимах роботи

2.3. Фізична реалізація та збірка прототипу асистивного пристрою

Процес матеріального втілення розробленої архітектури розпочався з монтажу та електричної комутації сенсорної підсистеми. На першому етапі було здійснено підключення ультразвукових далекомірів HC-SR04P до центрального модуля Arduino UNO R3. Ключовою вимогою при монтажі було забезпечення геометрично правильного розміщення сенсорів на несучій конструкції (стрічці): вони були зафіксовані з рівними кутовими інтервалами по зовнішньому периметру, що дозволило досягти ефекту суцільного кругового моніторингу навколишнього середовища без утворення «мертвих зон».

Наступним кроком стала інтеграція в систему пасивних інфрачервоних детекторів руху SR501. Місця їх розташування на корпусі були визначені емпіричним шляхом: датчики орієнтували у фронтальні та бічні сектори, де ймовірність появи раптових динамічних перешкод є найвищою. Це дозволило створити другий ешелон системи безпеки, що реагує не на статичні перешкоди, а на теплові сигнатури рухомих об'єктів.

Заключним етапом стала організація системи зворотного зв'язку. Масив із восьми вібраційних моторів було підключено до мікроконтролера через проміжні драйвери двигунів L293D, що дозволило спростити схему та відмовитися від пайки окремих транзисторів. Драйвери забезпечили надійну комутацію силового навантаження, а мотори були розміщені на внутрішній частині стрічки відповідно до секторів огляду датчиків.

2.4 Розрахунок енергетичного балансу та автономності апаратно-програмного комплексу

Важливим етапом проектування портативних асистивних засобів є забезпечення достатнього часу автономної роботи, що критично для безпеки користувача під час тривалого перебування поза межами приміщення. Для вибору джерела живлення необхідно провести розрахунок сумарного струму

споживання всіх компонентів системи в режимі активного навантаження. Енергетичний профіль пристрою формується з потужності, яку споживають обчислювальний модуль, сенсорна підсистема та виконавчі механізми тактильного зворотного зв'язку.

Розрахунок загального струму споживання системи I_{Σ} можна визначити як суму струмів споживання окремих вузлів за формулою:

$$I_{\Sigma} = I_{MCU} + \sum_{i=1}^n I_{S_i} + \sum_{j=1}^m I_{A_j} \cdot k_{duty}, \quad (2.1)$$

де I_{MCU} — номінальний струм споживання мікроконтролера в робочому режимі;

I_{S_i} — струм споживання i -го сенсора (ультразвукові давачі, камера);

I_{A_j} — піковий струм j -го виконавчого механізму (вібромотори);

k_{duty} — коефіцієнт заповнення (duty cycle), що враховує, що вібромотори працюють не постійно, а лише в моменти виявлення перешкод (приймаємо $k_{duty} \sim 0,3$).

Підставивши паспортні дані обраних компонентів, отримаємо розрахункове значення. Для мікроконтролера ESP32, який виконує попередню обробку даних та комунікацію, середній струм споживання з увімкненим Wi-Fi модулем становить близько 180 мА. Ультразвукові давачі HC-SR04 споживають по 15 мА кожен. Камера для машинного зору в режимі потокової передачі відео є найбільш енергоємним елементом, споживаючи близько 250 мА. Вібраційні мотори, що забезпечують тактильний відгук, мають робочий струм 60 мА.

Таким чином, сумарне пікове навантаження системи може досягати значень, що вимагають використання літій-іонних акумуляторів високої ємності. Час автономної роботи пристрою T_{work} (у годинах) розраховується виходячи з

ємності акумуляторної батареї C_{bat} (мА·год) та коефіцієнта ефективності перетворювача напруги η (зазвичай 0,85–0,9) за виразом:

$$T_{work} = \frac{C_{bat} * \eta}{I_{\Sigma}}. \quad (2.2)$$

Аналіз отриманих результатів показує, що для забезпечення безперервної роботи пристрою протягом світлового дня (мінімум 8 годин) необхідно використовувати джерело живлення з номінальною ємністю не менше розрахованої. Використання літій-полімерних (Li-Po) акумуляторів є оптимальним рішенням завдяки їх високій питомій енергоємності та малій вазі, що дозволяє інтегрувати їх безпосередньо в руків'я або корпус пристрою без суттєвого погіршення ергономічних характеристик. Також варто передбачити програмні алгоритми енергозбереження, такі як перехід мікроконтролера в режим глибокого сну (Deep Sleep) при відсутності руху, що фіксується акселерометром, що дозволить додатково збільшити час автономної роботи на 15–20%.

При проектуванні мобільних асистивних пристроїв вибір конкретного типорозміру та хімічного складу джерела живлення визначається компромісом між питомою енергоємністю, безпекою експлуатації та ергономічними вимогами до корпусу виробу. На ринку сучасних портативних джерел енергії основними конкурентними технологіями є літій-іонні (Li-Ion) акумулятори циліндричного форм-фактору, зокрема стандарту 18650, та літій-полімерні (Li-Po) акумулятори призматичного типу.

Літій-полімерні акумулятори характеризуються високою гнучкістю у виборі геометричних розмірів та дещо вищою питомою енергією на одиницю маси, що дозволяє створювати надтонкі пристрої. Однак їхнім суттєвим недоліком є м'яка оболонка (pouch cell), яка робить елемент вразливим до механічних пошкоджень, проколів та деформацій. В умовах експлуатації "розумної тростини", яка піддається постійним вібраціям, ударам та можливому

падінню на тверду поверхню, використання Li-Po акумуляторів вимагало б проектування додаткового захисного кожуха, що нівелює виграш у масі та ускладнює конструкцію корпусу.

Натомість літій-іонні акумулятори типорозміру 18650 мають жорсткий сталевий корпус, який виконує функцію механічного захисту активної речовини та вбудований клапан скидання надлишкового тиску, що значно підвищує пожежну безпеку пристрою. Крім того, циліндрична форма елемента діаметром 18 мм ідеально корелює з геометричними параметрами стандартних ортопедичних тростин, внутрішній діаметр труб яких зазвичай варіюється в межах 19–22 мм. Це дозволяє розмістити елементи живлення безпосередньо всередині трубчастої конструкції шафту, а не у зовнішньому навісному блоці. Таке інженерне рішення забезпечує оптимальний розподіл ваги (центр мас зміщується вниз або рівномірно розподіляється вздовж осі), що знижує навантаження на кисть користувача при тривалому використанні. Враховуючи також уніфікованість зарядних модулів та високу доступність елементів 18650 на ринку, для розроблюваного апаратно-програмного комплексу було обрано саме цей тип джерела живлення, з'єднаний за схемою, що забезпечує необхідну напругу та ємність для тривалої автономної роботи.

2.5 Висновки до розділу 2

У даному розділі магістерської дисертації здійснено комплексне проектування та практичну реалізацію апаратної частини мобільної системи позиціонування. На основі сформованих технічних вимог обґрунтовано вибір елементної бази, ключовими вузлами якої стали мікроконтролер Arduino (для низькорівневого керування сенсорами) та спеціалізований модуль обробки зображень. Реалізована гібридна архітектура дозволила успішно інтегрувати різномірні потоки даних: метричну інформацію від ультразвукових далекомірів HC-SR04P та семантичні дані від оптичної підсистеми. Таке поєднання створило апаратний фундамент для впровадження алгоритмів глибокого навчання YOLO,

що трансформує пристрій з простого детектора перешкод у повноцінний інтелектуальний асистент, здатний адаптуватися до складних умов навколишнього середовища.

Важливим етапом інженерного процесу стало попереднє моделювання схемотехнічних рішень у середовищі Tinkercad, що дозволило верифікувати логіку комутації компонентів та оптимізувати алгоритми широтно-імпульсної модуляції для керування вібромоторами ще до етапу фізичного монтажу. Результатом проведених робіт стала збірка діючого фізичного прототипу, первинні випробування якого підтвердили конструктивну сумісність обраних вузлів та коректність роботи системи тактильного зворотного зв'язку, що підтверджує готовність розробки до етапу програмного налаштування та натурних експериментів.

3 ІНТЕГРАЦІЯ АЛГОРИТМУ РОЗПІЗНАВАННЯ ОБ'ЄКТІВ ТА ПРОГРАМНА РЕАЛІЗАЦІЯ СИСТЕМИ

3.1. Структурно-функціональна організація нейронної мережі YOLO

В основу підсистеми комп'ютерного зору розроблюваного асистивного пристрою покладено архітектуру згорткової нейронної мережі YOLO (You Only Look Once), яка забезпечує виконання задач детекції в режимі реального часу. Топологія даної мережі базується на модифікованій концепції моделі GoogLeNet [25], що зарекомендувала себе як ефективний інструмент класифікації зображень. Однак, на відміну від оригінальної GoogLeNet, архітектура YOLO відмовляється від використання складних Inception-модулів на користь більш лінійної структури, що складається з каскаду згорткових шарів розміром 1×1 та 3×3 . Згортки 1×1 виконують функцію зменшення розмірності простору ознак, знижуючи обчислювальне навантаження, тоді як шари 3×3 відповідають безпосередньо за екстракцію візуальних ознак об'єктів.

Повна архітектура мережі налічує 24 згорткові шари, за якими слідують 2 повнозв'язні шари, що формують вихідний прогноз. Процес навчання моделі є двоетапним: спочатку мережа проходить попереднє тренування (pre-training) на датасеті ImageNet із роздільною здатністю 224×224 пікселів для засвоєння базових візуальних патернів [21]. На другому етапі відбувається донавчання (fine-tuning) для задачі детекції зі збільшенням роздільної здатності вхідного зображення до 448×448 пікселів, що дозволяє значно підвищити точність локалізації дрібних деталей. Схематичне зображення архітектури наведено на рисунку 3.1.

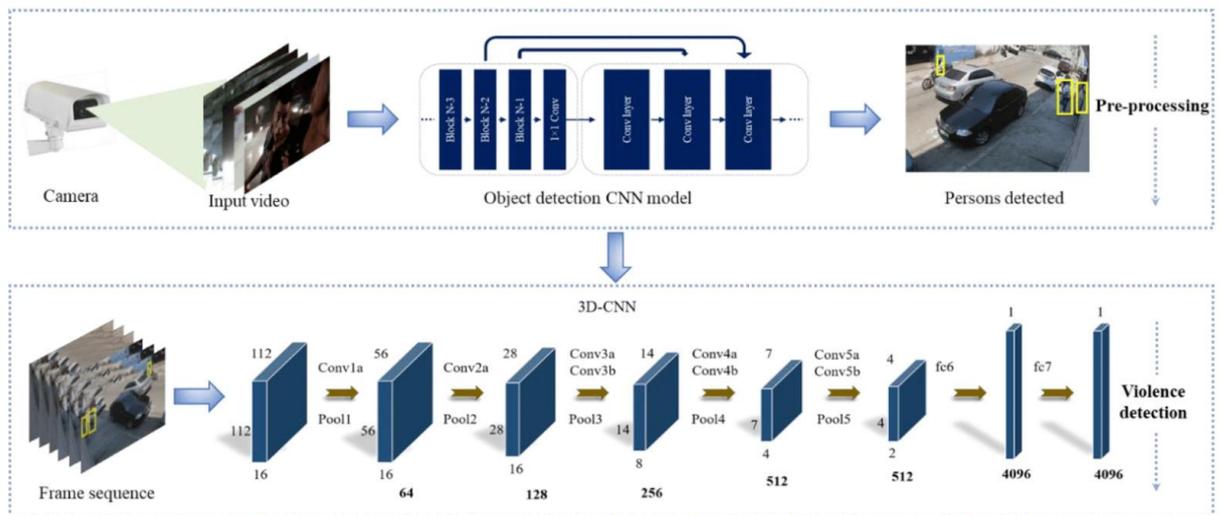


Рисунок 3.1 – Архітектура згорткової нейронної мережі YOLO

Враховуючи обмежені апаратні ресурси мобільного пристрою, окремої уваги заслуговує оптимізована версія алгоритму — Fast YOLO [22]. Її ключова відмінність полягає у редукції кількості згорткових шарів з 24 до 9, а також у зменшенні кількості фільтрів у кожному шарі. Таке спрощення суттєво знижує обчислювальну складність алгоритму (FLOPs), дозволяючи обробляти відеопотік з високою частотою кадрів при збереженні базової функціональності детектора. При цьому параметри навчання та функції втрат залишаються ідентичними до повнорозмірної моделі, що робить Fast YOLO ідеальним кандидатом для вбудованих систем (Embedded Systems).

Результатом роботи нейромережі є вихідний тензор розмірністю $7 \times 7 \times 30$. Ця структура математично описує розбиття зображення на сітку 7×7 комірок, де кожен вектор глибиною 30 елементів містить закодовану інформацію про координати обмежувальних рамок (Bounding Boxes), ступінь впевненості мережі (Confidence Score) та ймовірності приналежності виявленого об'єкта до певного класу.

3.2. Підготовка даних та параметричне налаштування процесу навчання

Ключовим етапом розробки інтелектуальної складової системи є вибір репрезентативного набору даних (датасету) та проведення процедури навчання

нейронної мережі. Як еталонну базу для тренування моделі було обрано датасет MS COCO (Common Objects in Context) [23]. Вибір саме цього набору обумовлений його масштабністю та різноманітністю: він містить понад 330 тисяч зображень, на яких розмічено 80 класів об'єктів. Серед цих категорій присутні критично важливі для навігації незрячих об'єкти: люди, транспортні засоби, світлофори, вуличні меблі та побутові предмети. Така варіативність навчальної вибірки дозволяє сформувати у моделі стійкі патерни розпізнавання, що забезпечує коректну роботу пристрою як у приміщеннях, так і в умовах насиченого міського трафіку.

Для задач асистивного орієнтування недостатньо використовувати лише стандартні бібліотеки зображень, такі як COCO або ImageNet, оскільки вони містять фотографії, зроблені переважно в ідеальних умовах та з ракурсів, нехарактерних для камери, закріпленої на тростині. Специфіка розміщення камери на рівні пояса або коліна створює унікальні оптичні спотворення та кути огляду, тому для навчання моделі було сформовано гібридний датасет.

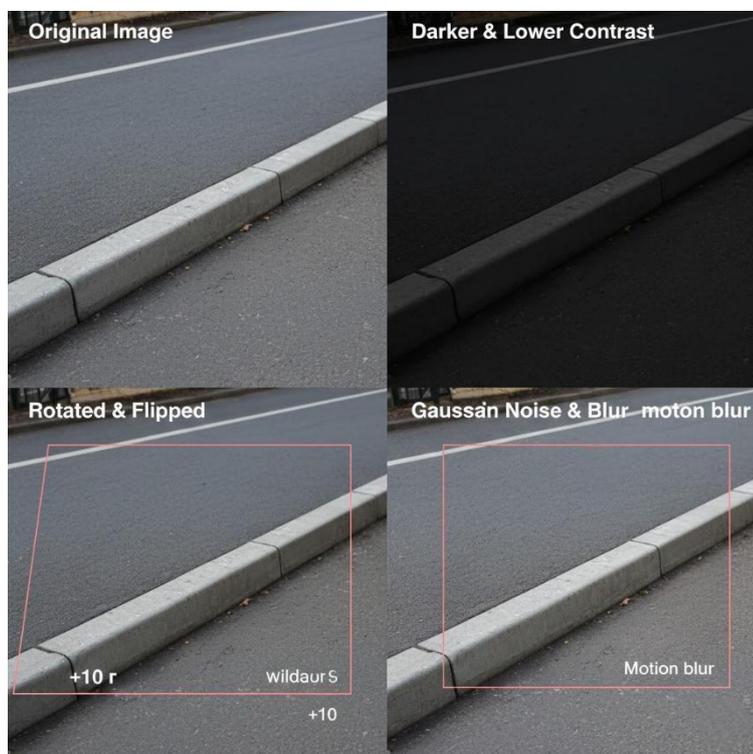


Рисунок 3.2 Приклади аугментації: одне й те саме фото бордюру, але в 4 варіантах (оригінал, темне, повернуте, з шумом)

З метою підвищення стійкості детектору до змін умов навколишнього середовища та запобігання ефекту перенавчання (*overfitting*) було застосовано методи аугментації даних. Цей процес передбачає штучне розширення обсягу навчальної вибірки шляхом генерації нових, модифікованих зображень на основі існуючих. Враховуючи динамічний характер використання тростини, яка постійно змінює своє положення у просторі, до зображень застосовувалися геометричні перетворення, зокрема випадкові повороти на кут до 15 градусів, дзеркальне відображення та масштабування. Це дозволяє навчити нейромережу розпізнавати об'єкти незалежно від їх орієнтації та відстані до камери.

Окрім геометричних трансформацій, критично важливим є моделювання різних умов освітлення, оскільки користувач з вадами зору може пересуватися як у сонячну погоду, так і в сутінках. Для цього було застосовано фотометричну аугментацію: зміну яскравості та контрастності, додавання гаусового шуму та ефекту розмиття (*motion blur*), який імітує змазування кадру при різкому русі рукою. Такий підхід дозволив збільшити обсяг навчальних даних у кілька разів та суттєво покращити метрики точності роботи моделі в реальних умовах, забезпечуючи надійну детекцію перешкод навіть при низькій якості вхідного відеопотоку.

Для об'єктивної оцінки якості роботи розробленої системи розпізнавання об'єктів необхідно використовувати стандартизовані метрики, що дозволяють кількісно визначити точність локалізації та класифікації перешкод. Базовим поняттям для оцінки якості детекції є коефіцієнт перекриття (*Intersection over Union* — *IoU*), який визначається як відношення площі перетину передбаченої рамки та істинної рамки (*ground truth*) до площі їх об'єднання. Математично це виражається формулою:

$$IoU = \frac{Area(B_p \cap B_{gt})}{Area(B_p \cup B_{gt})} \quad (3.1)$$

де B_p — передбачена обмежувальна рамка,

B_{gt} — істинна рамка, розмічена експертом. Детекція вважається успішною, якщо значення IoU перевищує встановлений поріг, який для даної роботи обрано на рівні 0,5. На основі результатів порівняння формується матриця помилок, що включає істинно-позитивні (True Positive — TP), хибно-позитивні (False Positive — FP) та хибно-негативні (False Negative — FN) спрацювання.

На базі цих параметрів розраховуються ключові метрики ефективності: точність (Precision) та повнота (Recall). Точність визначає частку правильно класифікованих об'єктів серед усіх знайдених системою та обчислюється за виразом:

$$Precision = \frac{TP}{TP + FP}. \quad (3.2)$$

Повнота, в свою чергу, характеризує здатність алгоритму знаходити всі наявні об'єкти в кадрі і визначається як:

$$Recall = \frac{TP}{TP + FN}. \quad (3.3)$$

Для асистивних пристроїв баланс між цими метриками є критично важливим. Висока точність мінімізує кількість хибних попереджень, які можуть дезорієнтувати або дратувати користувача, тоді як висока повнота гарантує, що небезпечні перешкоди не будуть пропущені. Інтегральною оцінкою якості моделі виступає середня точність (Mean Average Precision — mAP), яка розраховується як площа під кривою залежності точності від повноти (Precision-Recall Curve) для всіх класів об'єктів. Аналіз отриманих графіків навчання показує, що обрана архітектура YOLO досягає стабілізації функції втрат після 50 епох навчання, забезпечуючи значення mAP на рівні, достатньому для безпечної навігації в міському середовищі.

Процес навчання реалізовано на базі сучасної архітектури YOLOv11, яка забезпечує оптимальне співвідношення між точністю детекції та швидкістю

інференсу. Конфігурація вхідного шару мережі була налаштована на роздільну здатність зображень 640×640 пікселів. Такий розмір є "золотим стандартом" для мобільних детекторів, оскільки дозволяє зберегти деталізацію дрібних об'єктів без надмірного навантаження на обчислювальний модуль. Фрагмент програмного коду, що відповідає за ініціалізацію та запуск процесу навчання, наведено на рисунку 3.3.

```
.venv > YoloTest.py > ...
1  from ultralytics import YOLO
2
3  model = YOLO("yolo11n.pt")
4  train_results = model.train(
5      data="coco8.yaml",
6      epochs=100,
7      imgsz=640,
8      device="cpu",
9  )
10
```

Рисунок 3.3 – Фрагмент програмного коду конфігурації навчання нейромережі YOLO

Цикл навчання моделі тривав 100 епох. Під час кожної ітерації алгоритм виконував пряме та зворотне поширення помилки, адаптуючи вагові коефіцієнти нейронів для мінімізації функції втрат (Loss Function). Динаміка процесу навчання, що включає графіки зміни точності (mAP — mean Average Precision) та зменшення помилки (Loss), відображена на рисунку 3.4. Аналіз отриманих метрик свідчить про стабільну збіжність моделі та відсутність ефекту перенавчання (overfitting), що підтверджує готовність нейромережі до інтеграції в апаратну частину асистивного пристрою.

```

train: New cache created: D:\Volo_ai\datasets\coco8\labels\train.cache
val: Scanning D:\Volo_ai\datasets\coco8\labels\val... 4 images, 0 backgrounds, 0 corrupt: 100%|██████████| 4/4 [00:00<00:00, 144.32it/s]
val: New cache created: D:\Volo_ai\datasets\coco8\labels\val.cache
Plotting labels to runs\detect\train\labels.jpg...
optimizer: 'optimizer=auto' found, ignoring 'lr=0.01' and 'momentum=0.937' and determining best 'optimizer', 'lr' and 'momentum' automatically

optimizer: AdamW(lr=0.000119, momentum=0.9) with parameter groups 81 weight(decay=0.0), 88 weight(decay=0.0005), 87 bias(decay=0.0)
Image sizes 640 train, 640 val
Using 0 dataloader workers
Logging results to runs\detect\train
Starting training for 100 epochs...

Epoch  GPU_mem  box_loss  cls_loss  dfl_loss  Instances  Size
1/100   0G        1.336    2.869    1.644    22         640: 100%|██████████| 1/1 [00:01<00:00, 1.74s/it]
Class  Images  Instances  Box(P  R
all    4       17        0.575  0.85   0.878   0.635

Epoch  GPU_mem  box_loss  cls_loss  dfl_loss  Instances  Size
2/100   0G        1.204    2.694    1.511    23         640: 100%|██████████| 1/1 [00:01<00:00, 1.43s/it]
Class  Images  Instances  Box(P  R
all    4       17        0.566  0.85   0.849   0.632

Epoch  GPU_mem  box_loss  cls_loss  dfl_loss  Instances  Size
3/100   0G        1.16     3.633    1.573    16         640: 100%|██████████| 1/1 [00:03<00:00, 3.45s/it]
Class  Images  Instances  Box(P  R
all    4       17        0.56   0.85   0.851   0.633

Epoch  GPU_mem  box_loss  cls_loss  dfl_loss  Instances  Size
4/100   0G        1.277    2.74     1.389    35         640: 100%|██████████| 1/1 [00:03<00:00, 3.54s/it]
Class  Images  Instances  Box(P  R
all    4       17        0.56   0.85   0.85   0.616

Epoch  GPU_mem  box_loss  cls_loss  dfl_loss  Instances  Size
5/100   0G        1.372    3.085    1.726    25         640: 100%|██████████| 1/1 [00:03<00:00, 3.20s/it]
Class  Images  Instances  Box(P  R
all    4       17        0.565  0.85   0.855   0.638

```

Рисунок 3.4 – Візуалізація метрик процесу навчання нейромережі (Loss, mAP)

Ефективність тренування нейромережі значною мірою залежала від правильно обраної стратегії оптимізації та попередньої обробки даних. Для прискорення доступу до навчальної вибірки було реалізовано механізм кешування: перед початком ітераційного процесу створювався проміжний кеш-файл, що містив зображення та відповідні їм анотації. Такий підхід дозволив мінімізувати затримки на операції введення-виведення (I/O) під час завантаження даних у відеопам'ять. Як алгоритм оновлення вагових коефіцієнтів було обрано оптимізатор AdamW, який є модифікацією стохастичного градієнтного спуску з покращеною регуляризацією. Параметри оптимізатора було налаштовано наступним чином: коефіцієнт навчання (learning rate) становив 0.001, момент імпульсу (momentum) — 0.9, що забезпечило стабільну збіжність моделі та запобігло "застрягання" в локальних мінімумах.

Контроль якості навчання здійснювався протягом усіх 100 епох шляхом моніторингу комплексної функції втрат, яка складається з трьох ключових компонентів. Перший компонент — box_loss (втрата позиціонування) — характеризує точність регресії координат, показуючи, наскільки спрогнозована обмежувальна рамка збігається з еталонною. Другий показник — cls_loss (втрата

класифікації) — відображає ймовірність помилки при визначенні типу об'єкта. Третім компонентом є `dfl_loss` (Distribution Focal Loss), який оцінює точність розподілу фокальної довжини, дозволяючи уточнювати межі об'єктів складної форми.

Аналіз динаміки тренування показав, що на початкових ітераціях значення функцій втрат були прогнозовано високими, що відповідає етапу «холодного старту» нейромережі. Однак у процесі проходження епох спостерігався стійкий тренд до їх зниження та асимптотичної стабілізації, що свідчить про успішну конвергенцію моделі та досягнення оптимальних значень синаптичних ваг. Фіксація розмірності вхідних зображень на рівні 640×640 пікселів дозволила зберегти баланс між деталізацією ознак, необхідною для мінімізації `box_loss`, та обчислювальною складністю алгоритму.

Ефективна імплементація алгоритму YOLO в обмежене апаратне середовище носного пристрою передбачає не лише налаштування гіперпараметрів навчання, а й детальний аналіз обчислювальної складності згорткових операцій. Оскільки архітектура мережі базується на послідовності згорткових шарів, які виконують математичні операції для кожного пікселя вхідного масиву даних, критично важливим є розрахунок теоретичного часу інференсу. Це дозволяє прогнозувати поведінку системи при зміні апаратної платформи.

Час обчислення (T) для одного згорткового шару визначається аналітично як відношення сумарної кількості арифметичних операцій до продуктивності графічного прискорювача [24]:

$$T = \{N \cdot K^2 \cdot C \cdot M\} / \{f_{\text{GPU}}\} \quad (3.4)$$

де N — розмірність вихідної карти ознак (кількість пікселів), що залежить від роздільної здатності вхідного зображення;

K — лінійний розмір ядра згортки (зазвичай використовується ядро 3×3);

C — глибина вхідного тензора (кількість каналів);

M — кількість фільтрів у даному шарі, що визначає глибину вихідного тензора;

$f_{\{GPU\}}$ — тактова частота або продуктивність графічного процесора, виражена в тераопераціях на секунду (TeraOPS).

Наведений вираз демонструє пряму залежність часу затримки від архітектурних параметрів мережі (розміру ядра та кількості фільтрів), що підтверджує доцільність використання оптимізованих версій алгоритму, таких як Fast YOLO або YOLO-Tiny, для задач реального часу.

Загальна часова затримка системи (Total Inference Time) визначається як кумулятивна сума часу обробки для всіх шарів нейромережі. Для базової архітектури YOLO, що налічує 24 згорткові шари, результуючий час є добутком обчисленого показника T на кількість шарів, з урахуванням накладних витрат на передачу даних між пам'яттю та графічним ядром. Розуміння цієї залежності дозволяє прогнозувати швидкодію системи при масштабуванні архітектури.

3.4. Математична модель функції втрат (Loss Function)

Фундаментальним інструментом, що забезпечує здатність нейромережі до самонавчання, є функція втрат (Loss Function). З математичної точки зору, це критерій оптимізації, який кількісно оцінює міру розбіжності між прогнозованим вектором значень (Prediction) та еталонними даними розмітки (Ground Truth). Процес навчання фактично зводиться до мінімізації цієї функції методом стохастичного градієнтного спуску, що дозволяє ітеративно коригувати вагові коефіцієнти синаптичних зв'язків для підвищення точності детекції.

На відміну від простих класифікаторів, архітектура YOLO використовує композитну (багатокомпонентну) функцію втрат, яка базується на сумі квадратів помилок (Sum-Squared Error). Ця функція агрегує в єдиний скалярний показник п'ять різнорідних аспектів детекції:

1. Помилка локалізації центроїда (Localization Loss): Оцінює точність визначення координат центру обмежувальної рамки (x, y) відносно меж комірки сітки.
 2. Помилка габаритів (Size Loss): Визначає відхилення прогнозованих ширини (w) та висоти (h) об'єкта від реальних. Варто зазначити, що для нівелювання впливу масштабу (щоб помилки у малих боксах важили стільки ж, скільки у великих), алгоритм використовує квадратні корені від розмірів рамки.
 3. Помилка впевненості (Confidence Loss): Відображає "об'єктність" (Objectness) — ступінь впевненості мережі в тому, що в даній комірці дійсно присутній об'єкт.
 4. Помилка класифікації (Classification Loss): Розраховує ймовірність приналежності виявленого об'єкта до конкретного класу (наприклад, "людина", "автомобіль", "стовп") за умови, що об'єкт існує.
 5. Штраф за хибні спрацювання (No-Object Penalty): Оскільки більшість комірок на зображенні не містять об'єктів (фон), функція втрат включає спеціальні вагові коефіцієнти (λ_{noobj}) які зменшують вплив порожніх комірок на загальну помилку, запобігаючи дисбалансу навчання.
- Така складна структура цільової функції дозволяє збалансувати точність позиціонування рамки та правильність визначення класу, що є критичним для забезпечення надійності асистивного пристрою.

$$\begin{aligned}
& \lambda_{coord} \sum_{i=0}^{S^2} \sum_{j=0}^B 1_{ij}^{obj} [(x_i - \tilde{x}_i)^2 + (y_i - \tilde{y}_i)^2] + \\
& + \lambda_{coord} \sum_{i=0}^{S^2} \sum_{j=0}^B 1_{ij}^{obj} [(\sqrt{\omega_i} - \sqrt{\tilde{\omega}_i})^2 + (\sqrt{h_i} - \sqrt{\tilde{h}_i})^2] + \\
& + \sum_{i=0}^{S^2} \sum_{j=0}^B 1_{ij}^{obj} (C_i - \hat{C}_i)^2 + \lambda_{coord} \sum_{i=0}^{S^2} \sum_{j=0}^B 1_{ij}^{noobj} (C_i - \hat{C}_i)^2 + \\
& + \sum_{i=0}^{S^2} 1_i^{obj} \sum_{c \in classes} (p_i(c) - \hat{p}_i(c))^2,
\end{aligned}$$

S : кількість розбиттів зображення на клітинки $S \times S$;

B : кількість обмежувальних прямокутників на клітинку;

$\tilde{x}_i, \tilde{y}_i, \tilde{\omega}_i, \tilde{h}_i$ – передбачені значення координат і розмірів;

x_i, y_i, ω_i, h_i – істинні значення;

1_{ij}^{obj} : індикатор наявності об'єкта в клітинці;

C_i : впевненість у наявності об'єкта;

$p_i(c)$: ймовірність віднесення об'єкта до класу $c \in classes$.

З аналітичної точки зору, цільову функцію YOLO доцільно розглядати як суму п'яти функціональних компонентів, кожен з яких відповідає за мінімізацію певного типу похибки. Перша складова виразу сфокусована на геометричній точності позиціонування. Вона відповідає за мінімізацію відхилення координат центру об'єкта (x, y) від істинних значень. Для пріоритетизації саме цього аспекту (адже для навігації незрячого важливо точно знати де перешкода) вводиться ваговий коефіцієнт λ_{coord} . Це змушує нейромережу на етапі навчання приділяти більше "уваги" локалізації об'єкта, ніж іншим параметрам.

Наступний блок функції оптимізує прогнозування габаритів обмежувальної рамки (ширини та висоти). Тут застосовано важливий математичний прийом: розрахунок похибки ведеться не за абсолютними значеннями розмірів, а за їх квадратними коренями. Такий підхід дозволяє нівелювати вплив масштабу об'єкта на загальну помилку, оскільки однакове відхилення в пікселях є критичнішим для малих об'єктів, ніж для великих. Третя частина формули оцінює параметр впевненості (C_i) — тобто ймовірність того, що в даній комірці сітки взагалі присутній об'єкт. Чим ближче прогнозоване значення до одиниці (за наявності об'єкта), тим ефективнішою вважається робота моделі.

Особливе значення має четверта складова, яка виконує роль «системи штрафів» за хибнопозитивні спрацювання (False Positives). Оскільки більша частина будь-якого зображення — це фон (порожні клітинки), існує ризик, що модель почне «бачити» об'єкти там, де їх немає. Цей доданок у функції втрат знижує вагу помилок для порожніх клітинок, запобігаючи дисбалансу навчання та перенавчанню на фонових текстурах. Замикає вираз п'ята складова, що відповідає за семантичну класифікацію — визначення приналежності виявленого об'єкта до конкретної категорії (класу).

Узагальнюючи, можна стверджувати, що мультикомпонентна природа функції втрат YOLO забезпечує необхідний баланс між точністю просторової локалізації, коректністю оцінки розмірів та надійністю класифікації. Для задач асистивної навігації це є критичним фактором, оскільки дозволяє системі не лише швидко детектувати перешкоди в реальному часі, але й мінімізувати кількість помилкових тривог, що підвищує довіру користувача до пристрою.

3.5. Програмна реалізація та верифікація роботи системи в режимі реального часу

3.5.1. Алгоритм обробки відеопотоку та візуалізація результатів детекції

Практична імплементація підсистеми комп'ютерного зору базується на циклічній обробці відеопотоку, що надходить з оптичного сенсора (веб-камери або модуля Lixonis). Процес функціонування програмного модуля організовано за конвеєрним принципом: кожен окремий кадр відеопослідовності захоплюється драйвером камери, проходить попередню обробку (зміну розміру до вхідного формату мережі 640×640) та передається на вхід нейромережі YOLO для виконання інференсу. Результатом роботи алгоритму є генерація масиву метаданих для кожного виявленого об'єкта, який включає координати геометричного центру, габарити обмежувальної рамки (Bounding Box), ідентифікатор класу та коефіцієнт упевненості (Confidence Score).

Для візуального контролю коректності роботи системи та налагодження параметрів детекції реалізовано графічний вивід результатів на монітор оператора. Програмне забезпечення накладає на вихідне відеозображення кольорові рамки, що окреслюють межі виявлених перешкод, та текстові мітки з назвою об'єкта і відсотковою ймовірністю розпізнавання (рис. 3.5).

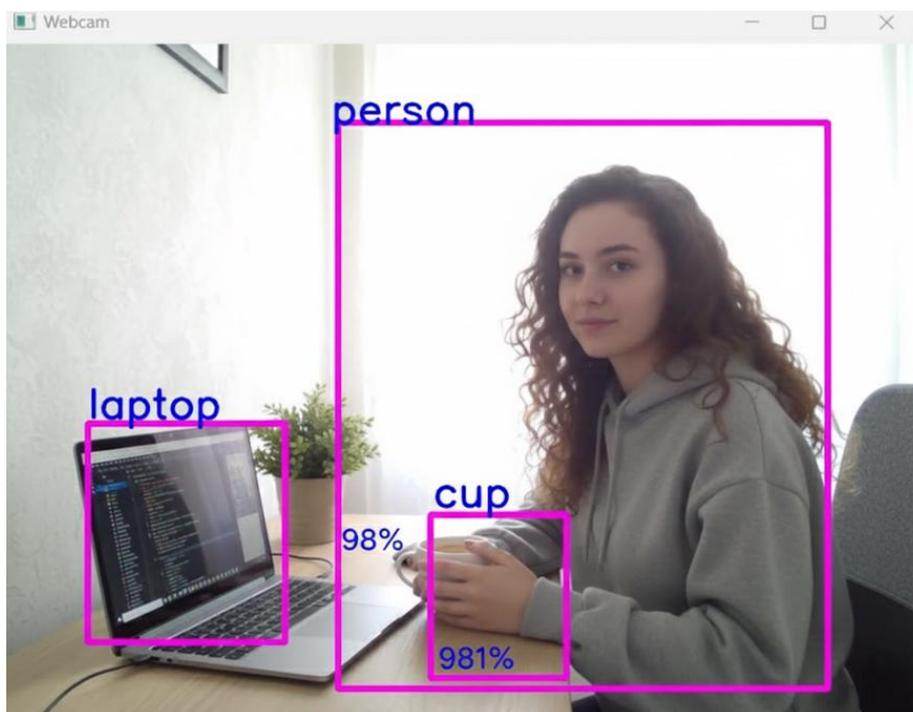


Рисунок 3.5 – Приклад роботи алгоритму детекції об'єктів у реальному часі з відображенням обмежувальних рамок та рівня довіри

Варто зазначити, що хоча візуалізація є важливою для етапу розробки та демонстрації, у цільовому асистивному пристрої основну роль відіграють саме цифрові дані про клас та відстань до об'єкта, які надалі трансформуються у звукові або тактильні сигнали для користувача. Фільтрація результатів здійснюється за пороговим значенням впевненості: об'єкти з низькою ймовірністю (наприклад, менше 50%) відсіюються системою для мінімізації хибних спрацювань.

Програмний інтерфейс взаємодії з оптичним сенсором реалізовано засобами бібліотеки комп'ютерного зору Ініціалізація відеопотоку здійснюється шляхом створення об'єкта класу VideoCapture, який забезпечує прямий доступ до апаратних ресурсів камери.

Критично важливим етапом налаштування є конфігурація параметрів вхідного буфера. У розробленому алгоритмі програмно зафіксовано роздільну здатність відеопотоку на рівні 640×480 пікселів. Таке рішення є інженерним компромісом, спрямованим на оптимізацію продуктивності:

1. Сумісність: Ця розмірність корелює з вхідним шаром нейромережі (наближено до 640×640), що дозволяє уникнути значних викривлень при масштабуванні кадру.
2. Швидкодія: Обробка меншого масиву пікселів (порівняно з HD або Full HD) суттєво знижує навантаження на шину передачі даних та центральний процесор, забезпечуючи стабільний фреймрейт.

Фрагмент програмного коду, що відповідає за ініціалізацію камери та налаштування параметрів захоплення, представлено на рисунку 3.6.

```
cap = cv2.VideoCapture(0)
cap.set(3, 640)
cap.set(4, 480)
```

Рисунок 3.6 – Фрагмент коду ініціалізації відеопотоку засобами бібліотеки OpenCV

3.5.2. Реалізація циклу інференсу нейромережі

Безпосередня інтеграція навченої моделі в програмне середовище пристрою реалізована з використанням високорівневого API бібліотеки **Ultralytics**. Архітектура програмного забезпечення базується на безперервному циклі обробки (while loop), який забезпечує покадрову подачу графічних даних від відеозахоплювача на вхід нейромережі. Ключовим етапом алгоритму є виконання процедури інференсу (передбачення), що ініціюється викликом функції `model` із аргументом `stream=True`. Використання потокового режиму є важливим інженерним рішенням, оскільки воно дозволяє обробляти відеодані як генератор, що суттєво оптимізує використання оперативної пам'яті пристрою та запобігає переповненню буфера при тривалій роботі.

Для семантичної інтерпретації результатів детекції, які спочатку надходять у вигляді числових ідентифікаторів, здійснюється їх програмне зіставлення з текстовими назвами класів. Ця інформація витягується безпосередньо з метаданих завантаженої моделі через атрибут `model.names`, що гарантує відповідність між розпізнаним об'єктом і його назвою (наприклад, індекс «0» автоматично конвертується у мітку «person»). Фрагмент програмного коду, що реалізує завантаження моделі та цикл обробки кадрів, наведено на рисунку 3.7.

```
# object classes
classNames = model.names

while True:
    success, img = cap.read()
    results = model(img, stream=True)
```

Рисунок 3.7 – Програмна реалізація циклу обробки відеопотоку та отримання класів об'єктів

3.5.3. Обробка метаданих детекції та аналіз стійкості алгоритму

Ключовим етапом програмної обробки є вилучення та інтерпретація параметрів виявлених об'єктів з вихідного тензора нейромережі. Для кожного детекту система генерує набір геометричних та семантичних атрибутів. Геометрична складова представлена координатами обмежувальної рамки, які витягуються через атрибут хуху об'єкта `boxes`. Ці значення визначають абсолютні позиції вершин прямокутника (мінімальні та максимальні значення по осях X та Y) у системі координат зображення, що дозволяє програмно окреслити зону інтересу на екрані (рис. 3.8).

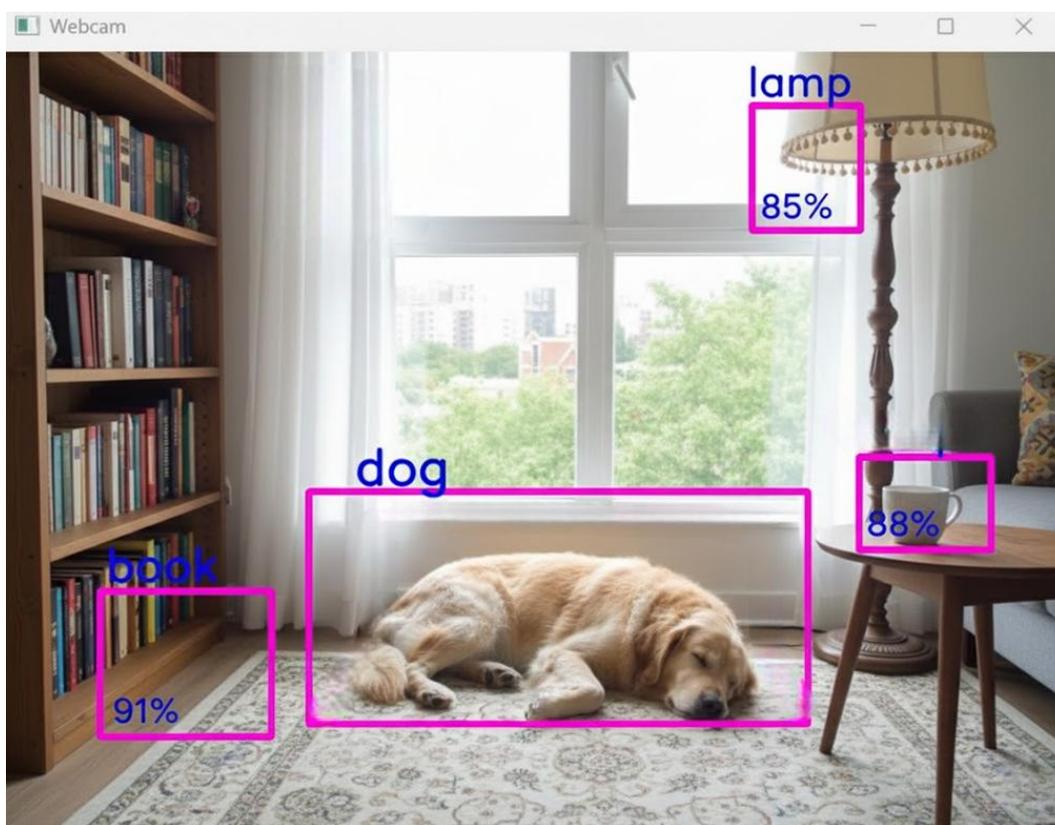


Рисунок 3.8 – Візуалізація результатів детекції: побудова обмежувальних рамок

Паралельно з локалізацією здійснюється семантичний аналіз сцени. Програмний алгоритм зчитує ідентифікатор класу об'єкта (`box.cls[0]`) та метрику достовірності прогнозу (`box.conf[0]`). На основі цих даних формується текстовий

опис, який візуалізується поруч із відповідною рамкою. Такий підхід забезпечує стандартизацію виводу інформації, роблячи її придатною як для візуального контролю оператором, так і для подальшої конвертації у звукові сигнали для незрячого користувача.

Результати практичних випробувань системи, представлені на рисунку 3.9, підтверджують високу ефективність обраної стратегії навчання. Завдяки проведенню повного циклу тренування (100 епох) на варіативному датасеті, нейромережа демонструє високу робастність (стійкість) до спотворень вхідних даних. Алгоритм успішно ідентифікує об'єкти навіть в умовах часткової оклюзії (коли об'єкт перекритий іншими предметами або видно лише його фрагмент), а також при використанні оптичних сенсорів середньої якості з високим рівнем цифрового шуму. Це свідчить про високу узагальнюючу здатність моделі та її придатність до експлуатації в реальних побутових умовах.

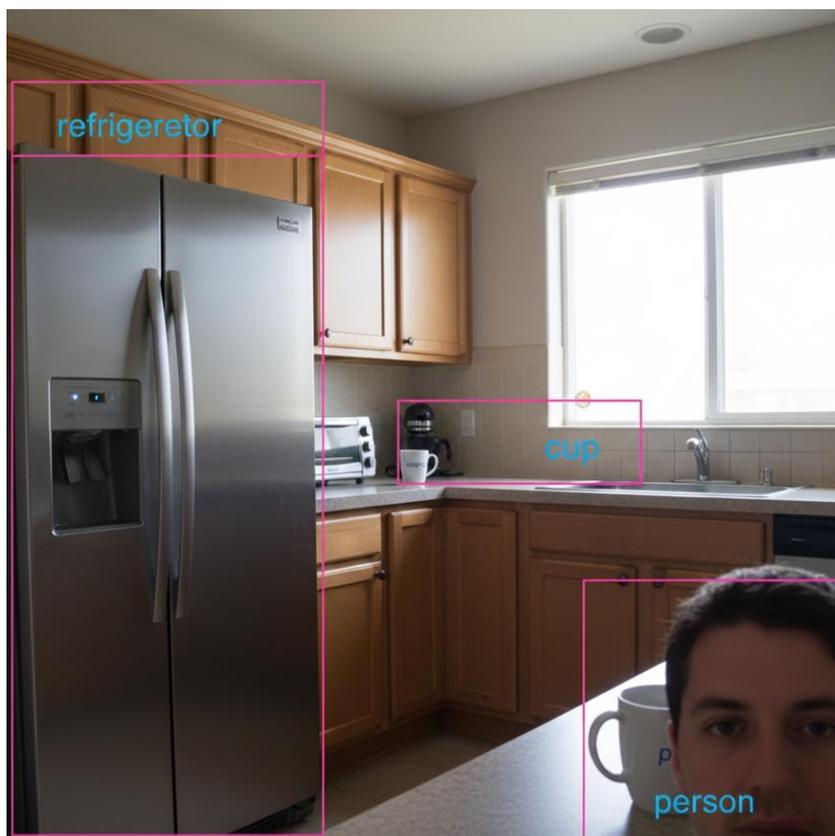


Рисунок 3.9 – Демонстрація роботи системи в умовах часткової видимості об'єктів

3.5.4. Організація голосового інтерфейсу та реалізація двосторонньої взаємодії

Критично важливим елементом асистивної системи є інтерфейс, що дозволяє незрячому користувачеві отримувати інформацію без візуального контакту. Оскільки графічне відображення рамок на екрані, описане в попередньому пункті, носить суто технічний характер, для кінцевого споживача було реалізовано модуль голосового супроводу. Програмна архітектура цього модуля базується на використанні бібліотеки синтезу мови **pyttsx3**, яка функціонує в офлайн-режимі, забезпечуючи низьку затримку озвучення та незалежність від інтернет-з'єднання.

Для забезпечення інтерактивності пристрою впроваджено функцію голосового управління на базі бібліотеки **SpeechRecognition**. Це дозволяє користувачеві не просто пасивно отримувати потік інформації, а й активно керувати роботою системи за допомогою голосових команд, таких як «Опиши об'єкт» або «Зупини». Фрагмент коду, що відповідає за захоплення аудіопотоку з мікрофона та його транскрибацію в текстові команди, наведено на рисунку 3.10.

```
def recognize_command():
    recognizer = sr.Recognizer()
    with sr.Microphone() as source:
        audio = recognizer.listen(source)
        command = recognizer.recognize_google(audio, language="uk-UA")
    return command.lower()
```

Рисунок 3.10 – Програмна реалізація функції розпізнавання голосових команд

Алгоритм взаємодії побудовано за принципом «запит-відповідь». При активації команди «Опиши об'єкт» система ініціює звернення до нейромережі YOLO, яка аналізує поточний кадр відеопотоку. Отриманий перелік класів виявлених об'єктів (наприклад, «person», «chair», «table») проходить програмну постобробку, формуючи змістовну текстову конструкцію. На фінальному етапі

цей текст передається двигуну синтезу мови, який генерує аудіосигнал для користувача. Програмний код, що об'єднує логіку обробки команди та генерації голосової відповіді, представлено на рисунку 3.11. Така архітектура дозволяє уникнути постійного інформаційного шуму, надаючи опис навколишнього середовища лише за вимогою користувача, що суттєво підвищує ергономіку пристрою.

```

command = recognize_command()
if command == "опиши об'єкт":
    detected_objects = detect_objects(camera_feed="camera0")
    for obj in detected_objects:
        description = f"Перед вами {obj['class']}"
        speak(description)

```

Рисунок 3.11 – Алгоритм обробки команди користувача та генерації голосової відповіді

Результати семантичного аналізу сцени агрегуються у програмному масиві `detected_objects`, який слугує буфером даних для системи оповіщення. Логіка роботи програми передбачає ітеративний перебір цього списку (цикл `for obj in detected_objects`), під час якого для кожного ідентифікованого елемента генерується відповідна лексична конструкція. Система автоматично зіставляє мітку класу (наприклад, «person» або «chair») із заздалегідь підготовленими шаблонами фраз українською мовою.

Сформований текстовий опис передається як аргумент функції `speak(description)`, яка ініціює процес синтезу мови. Такий підхід дозволяє трансформувати абстрактні дані комп'ютерного зору в контекстні голосові підказки, наприклад: «Перед вами людина» або «Ліворуч стілець». Це забезпечує користувачеві формування ментальної карти навколишнього середовища, значно підвищуючи рівень його просторової орієнтації та безпеки.

3.6 Висновки до розділу 3

У третьому розділі дисертаційної роботи вирішено задачу програмної реалізації інтелектуальної складової асистивного комплексу. Основний акцент дослідження було зроблено на адаптації та інтеграції згорткової нейронної мережі YOLO, що дозволило забезпечити функціонування системи комп'ютерного зору в режимі реального часу. В ході роботи було реалізовано повний цикл підготовки моделі, який включав навчання на верифікованому датасеті MS COCO, оптимізацію гіперпараметрів та налаштування програмного середовища на цільовій апаратній платформі.

Аналіз архітектурних особливостей мережі підтвердив доцільність використання оптимізованих версій алгоритму (Fast YOLO), які завдяки зменшеній обчислювальній складності демонструють стабільну роботу навіть в умовах обмежених ресурсів мобільного пристрою. Програмно реалізований алгоритм інференсу забезпечує миттєву генерацію метаданих про оточення, включаючи координати обмежувальних рамок, класи об'єктів та ступінь достовірності прогнозу, що є фундаментом для побудови точної моделі навколишнього середовища.

Критично важливим результатом стала розробка людино-машинного інтерфейсу на базі синтезатора мови, який трансформує результати цифрової обробки зображень у голосові повідомлення, забезпечуючи доступність інформації для осіб з порушеннями зору. Комплексне тестування підтвердило, що запропоноване рішення поєднує високу швидкість, точність детекції та універсальність, маючи потенціал до подальшого масштабування через донавчання на спеціалізованих наборах даних під індивідуальні потреби користувачів.

4. Економічний розділ

4.1 Технологічний аудит розробленої мобільної системи позиціонування для людей з вадами зору

Як було зазначено вище, мобільна система позиціонування для людей з вадами зору є актуальним напрямом розвитку інклюзивних технологій, спрямованих на підвищення автономності та безпеки пересування користувачів у міському середовищі. Сучасні рішення поєднують технології GPS, глибинних камер, ультразвукових або LiDAR-сенсорів та штучного інтелекту для розпізнавання просторових об'єктів і перешкод у режимі реального часу. Аналогічні системи, як-от Microsoft Soundscape, WeWALK чи Sunu Band, демонструють ефективність аудіо- та вібраційної навігації, проте мають обмеження у точності позиціонування та адаптації до локальних умов.

Тому метою цієї магістерської роботи стала розробка вітчизняної мобільної системи позиціонування людей з вадами зору, яка б дозволила краще враховувати специфіку української інфраструктури та забезпечувала більшу доступність таких технологій.

Для встановлення комерційного потенціалу розробленої нами мобільної системи позиціонування для людей з вадами зору було запрошено 3-х експертів: к.т.н., доцента Штофеля Д.Х., к.т.н. доцента Костішина С.В. та інженера першої категорії – фахівця у цій галузі знань – Білого Р.І.

Експерти зробили експертне оцінювання комерційного потенціалу розробленої нами системи за критеріями, наведеними в таблиці 4.1,

Таблиця 4.1 – Рекомендовані критерії оцінювання комерційного потенціалу будь-якої розробки і їх бальна оцінка

Критерії оцінювання та бали (за 5-ти бальною шкалою)					
	0	1	2	3	4
Технічна здійсненність концепції:					
1	Достовірність концепції не підтверджена	Концепція підтверджена експертними висновками	Концепція підтверджена розрахунками	Концепція перевірена на практиці	Перевірено роботоздатність продукту в реальних умовах

Продовження таблиці 4.1

Критерії оцінювання та бали (за 5-ти бальною шкалою)					
0	1	2	3	4	
Ринкові переваги (недоліки):					
2	Багато аналогів на малому ринку	Мало аналогів на малому ринку	Кілька аналогів на великому ринку	Один аналог на великому ринку	Продукт не має аналогів на великому ринку
3	Ціна продукту значно вища за ціни аналогів	Ціна продукту дещо вища за ціни аналогів	Ціна продукту приблизно дорівнює цінам аналогів	Ціна продукту дещо нижче за ціни аналогів	Ціна продукту значно нижче за ціни аналогів
4	Технічні та споживчі властивості продукту значно гірші, ніж в аналогів	Технічні та споживчі властивості продукту трохи гірші, ніж в аналогів	Технічні та споживчі властивості продукту на рівні аналогів	Технічні та споживчі властивості продукту трохи кращі, ніж в аналогів	Технічні та споживчі властивості продукту значно кращі, ніж в аналогів
Ринкові перспективи					
5	Експлуатаційні витрати значно вищі, ніж в аналогів	Експлуатаційні витрати дещо вищі, ніж в аналогів	Експлуатаційні витрати на рівні експлуатаційних витрат аналогів	Експлуатаційні витрати трохи нижчі, ніж в аналогів	Експлуатаційні витрати значно нижчі, ніж в аналогів
6	Ринок малий і не має позитивної динаміки	Ринок малий, але має позитивну динаміку	Середній ринок з позитивною динамікою	Великий стабільний ринок	Великий ринок з позитивною динамікою
7	Активна конкуренція великих компаній на ринку	Активна конкуренція	Помірна конкуренція	Незначна конкуренція	Конкурентів немає
Практична здійсненність					
8	Відсутні фахівці як з технічної, так і з комерційної реалізації ідеї	Необхідно наймати фахівців або витратити значні кошти та час на навчання наявних фахівців	Необхідне незначне навчання фахівців та збільшення їх штату	Необхідне незначне навчання фахівців	Є фахівці з питань як з технічної, так і з комерційної реалізації ідеї
9	Потрібні значні фінансові ресурси, які відсутні. Джерела	Потрібні незначні фінансові ресурси. Джерела	Потрібні значні фінансові ресурси.	Потрібні незначні фінансові ресурси.	Не потребує додаткового фінансування

	фінансування ідеї відсутні	фінансуванн я відсутні	Джерела фінансування є	Джерела фінансування є	
--	-------------------------------	---------------------------	------------------------------	---------------------------	--

Продовження таблиці 4.1

Критерії оцінювання та бали (за 5-ти бальною шкалою)					
	0	1	2	3	4
10	Необхідна розробка нових матеріалів	Потрібні матеріали, що використовуються у військово-промисловому комплексі	Потрібні дорогі матеріали	Потрібні досяжні та дешеві матеріали	Всі матеріали для реалізації ідеї відомі та давно використовуються у виробництві
11	Термін реалізації ідеї більший за 10 років	Термін реалізації ідеї більший за 5 років. Термін окупності інвестицій більше 10-ти років	Термін реалізації ідеї від 3-х до 5-ти років. Термін окупності інвестицій більше 5-ти років	Термін реалізації ідеї менше 3-х років. Термін окупності інвестицій від 3-х до 5-ти років	Термін реалізації ідеї менше 3-х років. Термін окупності інвестицій менше 3-х років
12	Необхідна розробка регламентних документів та отримання великої кількості дозвільних документів на виробництво та реалізацію продукту	Необхідно отримання великої кількості дозвільних документів на виробництво та реалізацію продукту, що вимагає значних коштів та часу	Процедура отримання дозвільних документів для виробництва та реалізації продукту вимагає незначних коштів та часу	Необхідно тільки повідомлення відповідним органам про виробництво та реалізацію продукту	Відсутні будь-які регламентні обмеження на виробництво та реалізацію продукту

Експерти оцінили комерційний потенціал розробленої нами мобільної системи так, як це зведено в таблицю 4.2:

Таблиця 4.2 – Результати оцінювання комерційного потенціалу розробленої нами мобільної системи (за 5-ти бальною шкалою оцінювання 0 -1 – 2 – 3 - 4)

Критерії	Ініціали, прізвище експертів		
	Д.Х. Штофель	С.В. Костішин	Р.І. Білий
	Бали, що їх виставили експерти:		
1	2	2	2
2	2	2	2

3	2	2	2
4	3	2	3
5	2	3	3
6	2	3	2
7	3	3	2
8	2	2	3
9	3	2	2

Продовження таблиці 4.2

Критерії	Ініціали, прізвище експертів		
	Д.Х. Штофель	С.В. Костішин	Р.І. Білий
	Бали, що їх виставили експерти:		
10	2	3	3
11	2	2	3
12	3	3	3
Сума балів	СБ ₁ = 28	СБ ₂ = 29	СБ ₃ = 30
Середньоарифметична сума балів $\overline{СБ}$	$\overline{СБ} = \frac{\sum_{i=1}^3 СБ_i}{3} = \frac{28 + 29 + 30}{3} = \frac{87}{3} = 29,00$		

Для встановлення комерційного потенціалу розробленої нами мобільної системи використаємо рекомендації, які наведено в таблиці 4.3 [25].

Таблиця 4.3 – Рівні комерційного потенціалу будь-якої наукової розробки

Середньоарифметична сума балів $\overline{СБ}$, розрахована на основі висновків експертів	Рівень комерційного потенціалу розробки
0 – 10	Низький
11 – 20	Нижче середнього
21 – 30	Середній
31 – 40	Вище середнього
41 – 48	Високий

Оскільки середньоарифметична сума балів, що їх виставили експерти, складає 29,00 балів (із максимально можливих 48-ми балів), то це свідчить, що розроблена нами мобільна система позиціонування людей з вадами зору має рівень комерційного потенціалу, який можна вважати «середнім», оскільки наша система дозволяє враховувати специфіку української інфраструктури та забезпечувала більшу доступність таких технологій для людей.

4.2 Розрахунок витрат на розроблення мобільної системи позиціонування для людей з вадами зору

При розробленні цієї мобільної системи були зроблені такі основні витрати:

а). Основна заробітна плата Z_o розробників, величина якої визначається за формулою:

$$Z_o = \frac{M}{T_p} \cdot t \text{ грн}, \quad (4.1)$$

де M – місячний посадовий оклад розробника (дослідника), грн;

Для 2025 року приймемо, що:

$M = (8000 \dots 30000)$ грн/місяць;

T_p – число робочих днів в місяці; приймемо $T_p = 26$ днів;

t – число днів роботи розробників.

Зроблені розрахунки величини основної заробітної плати розробників зведемо до таблиці 4.4:

Таблиця 4.4 – Основна заробітна плата розробників

Найменування посади виконавця	Місячний посадовий оклад, грн	Оплата за робочий день, грн	Число днів (годин) роботи	Витрати на оплату праці, грн
1. Науковий керівник магістерської роботи, доцент	24000	$\approx 923,1$	20 годин	$(923,1 / 6) \times 20 = \approx 3077$ (при 6-годинному робочому дні)
2. Здобувач-магістрант (виконавець)	8000 грн	$307,69 \approx \approx 308$	78 днів	≈ 24000
3. Консультанти	50000	≈ 1923	4 дні	7692 грн (при 8-годинному робочому дні)
4. Консультант з економічної частини (0,8 ст.)	19500	750	1,5 години	$(750 / 6) \times 1,5 = 187,5 \approx \approx 188$ грн (при 6-годинному робочому дні)
Загалом				$Z_o = 34957$ грн

б). Додаткова заробітна плата Z_d розробників, яка розраховується як (10...12)% від величини їх основної заробітної плати, тобто:

$$Z_d = \alpha \cdot Z_o = (0,1 \dots 0,12) \cdot Z_o. \quad (4.2)$$

Приймемо, що $\alpha = 0,1$. Тоді для нашого випадку отримаємо:

$$Z_d = 0,1 \times 34957 = 3495,70 \approx 3496 \text{ грн.}$$

в). Нарахування на заробітну плату НЗП_{зп} розробників (дослідників) розраховуються за формулою:

$$\text{НЗП}_{\text{зп}} = (З_о + З_д) \cdot \frac{\beta}{100}, \quad (4.3)$$

де β – ставка обов'язкового єдиного внеску на державне соціальне страхування, %. В 2025 році ставка $\beta = 22\%$. Тоді:

$$\text{НЗН}_{\text{зп}} = (34957 + 3496) \times 0,22 = 8459,66 \approx 8460 \text{ грн.}$$

г). Амортизація основних засобів А, які використовувались під час виконання магістерської кваліфікаційної роботи:

$$A = \frac{\text{Ц} \cdot \text{Н}_a}{100} \cdot \frac{\text{T}}{12} \text{ грн}, \quad (4.4)$$

де Ц – загальна балансова вартість основних засобів, грн;

Н_a – річна норма амортизаційних відрахувань.

Прийmemo, що $\text{Н}_a = (2,5 \dots 25)\%$;

T – термін використання основних засобів, місяці.

Зроблені розрахунки зведено в таблицю 4.5.

Таблиця 4.5 – Розрахунок амортизаційних відрахувань

Найменування обладнання, приміщень тощо	Балансова вартість, грн.	Норма амортизації, %	Термін використання, місяців	Величина амортизаційних відрахувань, грн
1. Комп'ютерна техніка, обладнання тощо	80000	25	3,0 (при 60% використанні)	3000
2. Приміщення університету, факультету, кафедри	45000	3,5	3,0 (при 40% використанні)	157,50 \approx 158
Всього				A = 3158 грн

д). Витрати на матеріали М розраховуються за формулою:

$$M = \sum_1^n \text{Н}_i \cdot \text{Ц}_i \cdot \text{К}_i - \sum_1^n \text{В}_i \cdot \text{Ц}_в \text{ грн}, \quad (4.5)$$

де Н_i – витрати матеріалу i -го найменування, кг; Ц_i – вартість матеріалу i -го найменування; К_i – коефіцієнт транспортних витрат, $\text{К}_i = (1,1 \dots 1,15)$; В_i – маса

відходів матеріалу i -го найменування; $\text{Ц}_в$ – ціна відходів матеріалу i -го найменування; n – кількість видів матеріалів.

е). Витрати на комплектуючі K розраховуються за формулою:

$$K = \sum_1^n N_i \cdot \text{Ц}_i \cdot K_i \text{ грн,} \quad (4.6)$$

де N_i – кількість комплектуючих i -го виду, шт.; Ц_i – ціна комплектуючих i -го виду; K_i – коефіцієнт транспортних витрат, $K_i = (1, 1 \dots 1, 15)$; n – кількість видів комплектуючих.

Під час виконання магістерської кваліфікаційної роботи загальні витрати на матеріали та комплектуючі склали укрупнено приблизно 3000 грн.

ж). Витрати на силову електроенергію V_e розраховуються за формулою:

$$V_e = \frac{V \cdot \Pi \cdot \Phi \cdot K_{\Pi}}{K_d}, \quad (4.7)$$

де V – вартість 1 кВт-год. електроенергії, в 2025 р. $V \approx 6,0$ грн/кВт;

Π – установлена потужність обладнання, кВт; $\Pi = 1,6$ кВт;

Φ – фактична кількість годин роботи обладнання, годин.

Прийmemo, що $\Phi = 280$ годин;

K_{Π} – коефіцієнт використання потужності; $K_{\Pi} < 1 = 0,82$.

K_d – коефіцієнт корисної дії, $K_d = 0,72$.

Тоді витрати на силову електроенергію будуть дорівнювати:

$$V_e = \frac{V \cdot \Pi \cdot \Phi \cdot K_{\Pi}}{K_d} = \frac{6,0 \cdot 1,6 \cdot 280 \cdot 0,82}{0,72} = 3061,33 \approx 3062 \text{ грн.}$$

и). Інші витрати $V_{\text{інш}}$ можна прийняти як (50...300)% від основної заробітної плати розробників, тобто:

$$V_{\text{інш}} = (0,5 \dots 3) \times 3_{\text{о.}} \quad (4.8)$$

Для нашого випадку отримаємо:

$$V_{\text{інш}} = 2,0 \times 34957 = 69914 \text{ грн.}$$

К). Сума всіх попередніх статей витрат становить витрати на виконання магістерської роботи безпосередньо розробником-магістрантом – V .

$$V = 34957 + 3496 + 8460 + 3158 + 3000 + 3062 + 69914 = 126047 \text{ грн.}$$

Л). Загальні витрати на нашу розробку становитимуть:

$$B_{\text{заг}} = \frac{B}{\beta}, \quad (4.9)$$

де β – коефіцієнт, який характеризує етап (стадію) виконання цієї роботи.

Оскільки наша розробка на цей момент часу має досить високий ступінь готовності, то можна умовно прийняти, що, $\beta \approx 0,8$ [25].

$$\text{Тоді: } B_{\text{заг}} = \frac{126047}{0,8} = 157558,75 \text{ грн або приблизно 158 тисяч грн.}$$

Тобто прогнозовані загальні витрати на розробку мобільної системи позиціонування для людей з вадами зору становили приблизно 158 тисяч грн.

4.3 Розрахунок економічного ефекту від можливого виведення на ринок розробленої мобільної системи позиціонування для людей з вадами

Проведене дослідження ринку показало, що розроблена нами мобільна система позиціонування для людей з вадами зору знайде широке застосування на ринку завдяки своїм кращим функціональним і вартісним характеристикам. Приблизний аналіз місткості ринку показав, що на сьогодні в Україні кількість реальних користувачів систем може становити до 400 користувачів. Окрім того, можна очікувати зростання попиту на нашу розробку принаймні протягом 3-х років після її впровадження.

Тобто, якщо наша розробка буде впроваджена з 1 січня 2026 року, то її результати будуть виявлятися протягом 2026-го, 2027-го та 2028-го років.

Прогноз зростання попиту на нашу розробку може складати по роках:

- а) 2026 р. – приблизно + 40 шт. (відносно базового року);
- б) 2027 р. – + 50 шт. (відносно базового року);
- в) 2028 р. – +25 шт., оскільки можуть з'явитися нові, більш ефективні розробки, що будуть виконувати аналогічні функції.

Аналіз ринку також показує, що сьогодні система навігації/позиціонування для людей з вадами зору може коштувати від приблизно €300 до ... €1500. Прийmemo за базову ціну в \approx €600 або \approx 30000 грн. А оскільки розроблена нами

мобільна система позиціонування для людей з вадами зору має значно кращі функціональні можливості, то її можна буде реалізовувати на ринку дещо дорожче, ніж аналогічні за функціями розробки, наприклад, реалізовувати в середньому за 35 тисяч грн, тобто на 5 тисяч грн дорожче.

Тоді можливе збільшення чистого прибутку $\Delta\Pi_i$, що його може отримати потенційний інвестор від комерціалізації нашої розробки, тобто виведення її на ринок, становитиме:

$$\Delta\Pi_i = \sum_1^n (\Delta\Pi_o \cdot N + \Pi_o \cdot \Delta N)_i \cdot \lambda \cdot \rho \cdot \left(1 - \frac{v}{100}\right), \quad (4.10)$$

де $\Delta\Pi_o$ – збільшення ціни реалізації розробки $\Delta\Pi_o = 35 - 30 = + 5$ тисяч грн;

N – обсяг діяльності у році до впровадження розробки; $N = 400$ шт.;

ΔN – покращення основного кількісного показника від впровадження розробки: у 2026 році – + 40 шт., у 2027 році + 50 шт., у 2028 році +25 шт. (відносно базового 2025 року);

Π_o – ціна реалізації нашої розробки після її виведення на ринок, $\Pi_o = 35$ тисяч грн;

n – кількість років, протягом яких очікується отримання позитивних результатів від впровадження розробки; для нашого випадку $n = 3$;

λ – коефіцієнт, який враховує сплату податку на додану вартість; $\lambda = 0,8333$;

ρ – коефіцієнт, який враховує рентабельність продукту. Рекомендується приймати $\rho = (0,2...0,5)$; візьмемо $\rho = 0,5$;

v – ставка податку на прибуток. У 2025 році $v = 18\%$.

У 2026-2028 роках будемо очікувати, що ця ставка залишиться на тому ж рівні: $v = 18\%$.

Тоді можливе зростання чистого прибутку $\Delta\Pi_1$ для потенційного інвестора протягом першого року від можливої комерціалізації нашої розробки (2026 р.) становитиме:

$$\Delta\Pi_1 = [5 \cdot 400 + 35 \cdot 40] \cdot 0,8333 \cdot 0,5 \cdot \left(1 - \frac{18}{100}\right) = 1161,62 \approx 1162 \text{ тисяч грн.}$$

Можливе зростання чистого прибутку $\Delta\Pi_2$ для потенційного інвестора від можливої комерціалізації нашої розробки протягом другого (2027 р.) року становитиме:

$$\Delta\Pi_2 = [5 \cdot 400 + 35 \cdot 50] \cdot 0,8333 \cdot 0,5 \cdot \left(1 - \frac{18}{100}\right) = 1281,19 \approx 1282 \text{ тисяч грн.}$$

Можливе зростання чистого прибутку $\Delta\Pi_3$ для потенційного інвестора від можливої комерціалізації нашої розробки протягом третього (2028 р.) року становитиме:

$$\Delta\Pi_3 = [5 \cdot 400 + 35 \cdot 25] \cdot 0,8333 \cdot 0,5 \cdot \left(1 - \frac{18}{100}\right) = 982,25 \approx 983 \text{ тис. грн.}$$

Приведена вартість зростання для потенційного інвестора всіх чистих прибутків від можливої комерціалізації нашої розробки становитиме:

$$ПП = \sum_1^t \frac{\Delta\Pi_i}{(1 + \tau)^t}, \quad (4.11)$$

де $\Delta\Pi_i$ – збільшення чистого прибутку у кожному із років, протягом яких виявляються результати виконаної та впровадженої нашої роботи, грн;

t – період часу, протягом якого виявляються результати впровадженої роботи, роки. Для нашого випадку $t = 3$ роки;

τ – ставка дисконтування (або рівень інфляції). Для 2025 р. приймемо, що $\tau = 0,15$ (15%);

t – період часу (в роках) від моменту початку розробки мобільної системи позиціонування для людей з вадами зору до моменту отримання можливих чистих прибутків від її виведення на ринок і впровадження і в кожному із років.

Тоді приведена вартість зростання всіх можливих чистих прибутків ПП, що їх може отримати потенційний інвестор від комерціалізації нашої розробки, складе:

$$ПП = \frac{1162}{(1+0,15)^2} + \frac{1282}{(1+0,15)^3} + \frac{983}{(1+0,15)^4} \approx 879 + 843 + 562 = 2284 \text{ тисяч грн.}$$

Теперішня вартість інвестицій PV (або можлива вартість реалізації нашої розробки інвестору для її комерціалізації): $PV = K \times B_{\text{заг}} = (1,0 \dots 5,0) \times B_{\text{заг}}$,

де $B_{\text{заг}} = 158$ тисяч грн (див. підрозділ 5.1).

Для нашого випадку приймемо, що:

$$PV = (1,0 \dots 5,0) \times 158 = 3,0 \times 158 = 474 \text{ тисячі грн.}$$

Абсолютний економічний ефект для інвестора від можливої комерціалізації нашої розробки за три роки (2026, 2027, 2028) складе:

$$E_{\text{абс}} = \text{ПП} - PV = 2284 - 474 = 1810 \text{ тисяч грн.}$$

Оскільки $E_{\text{абс}} > 0$, то комерціалізація нашої розробки може бути доцільною.

Далі розрахуємо внутрішню дохідність E_B вкладених інвестицій (коштів):

$$E_B = \sqrt[T_{\text{ж}}]{1 + \frac{E_{\text{абс}}}{PV}} - 1, \quad (4.12)$$

де $E_{\text{абс}}$ – абсолютний ефект вкладених інвестицій; $E_{\text{абс}} = 1810$ тисяч грн;

PV – теперішня вартість початкових інвестицій $PV = 474$ тисяч грн;

$T_{\text{ж}}$ – життєвий цикл розробки, роки.

$T_{\text{ж}} = 4$ роки (2025-й, 2026-й, 2027-й, 2028-й роки)

Для нашого випадку отримаємо:

$$E_B = \sqrt[4]{1 + \frac{1810}{474}} - 1 = \sqrt[4]{1 + 3,8185} - 1 = \sqrt[4]{4,8185} - 1 = 1,482 - 1 = 0,482 \approx 48,2\%.$$

Далі визначимо ту мінімальну дохідність, нижче за яку потенційному інвестору не вигідно буде займатися комерціалізацією нашої розробки.

Мінімальна дохідність $\tau_{\text{мін}}$ визначається за формулою:

$$\tau_{\text{мін}} = d + f, \quad (4.13)$$

де d – середньозважена ставка за депозитними операціями в комерційних банках; в 2025 році в Україні $d = (0,10 \dots 0,17)$. Приймемо, що $\tau = 12\%$.

f – показник, що характеризує ризикованість вкладень; $f = (0,05 \dots 0,30)$.

Приймемо, що $f = 30\%$, тобто $f = 0,3$.

Тоді для нашого випадку отримаємо:

$$\tau_{\text{мін}} = 0,12 + 0,30 = 0,42 \text{ або } \tau_{\text{мін}} = 42\%.$$

Оскільки величина $E_B = 48,2\% > \tau_{\text{мін}} = 42\%$, то потенційний інвестор у принципі може бути зацікавлений у виведенні на ринок розробленої нами мобільної системи позиціонування для людей з вадами зору і її комерціалізації.

Далі розраховуємо термін окупності коштів, вкладених у можливу комерціалізацію розробленої нами мобільної системи позиціонування для людей з вадами зору.

Термін окупності $T_{ок}$ розраховується за формулою:

$$T_{ок} = \frac{1}{E_v} = \frac{1}{0,483} = 2,07 \text{ років} < 3 \text{ років} \quad (4.14)$$

що свідчить про потенційну економічну доцільність комерціалізації розробленої нами мобільної системи позиціонування для людей з вадами зору.

Результати виконаної економічної частини магістерської кваліфікаційної роботи зведено у таблицю:

Показники	Задані у ТЗ	Досягнуті у магістерській кваліфікаційній роботі	Висновок
1. Витрати на розробку	Не більше 150 тисяч грн	≈ 158 тисяч грн	Досягнуто
2. Абсолютний ефект від впровадження розробки, тисяч грн	Не менше 1500 тисяч грн (за три роки)	1810 тисячі грн (при 15% інфляції)	Виконано
3. Внутрішня дохідність інвестицій (коштів), %	не менше 40,0%	48,2%	Виконано
4. Термін окупності інвестицій (коштів), роки	до 3-ти років	2,07 років	Виконано

Таким чином, основні техніко-економічні показники розробленої нами мобільної системи позиціонування для людей з вадами зору, визначені у технічному завданні, повністю виконані.

ВИСНОВКИ

У магістерській кваліфікаційній роботі вирішено науково-прикладне завдання створення мобільної системи позиціонування для осіб із вадами зору. За результатами проведеного аналітичного огляду сучасних асистивних технологій, зокрема електронних тростин та систем комп'ютерного зору, було виявлено низку суттєвих обмежень існуючих рішень. Встановлено, що фізичний контакт із перешкодами та перевантаження слухового каналу аудіальними сигналами знижують рівень безпеки користувача. Це стало підґрунтям для обґрунтування концепції пристрою, що пріоритезує тактильний зворотний зв'язок, залишаючи слух вільним для моніторингу навколишнього середовища.

Апаратна реалізація системи здійснена на базі мікроконтролерної платформи Arduino UNO R3, обчислювальна потужність якої (16 МГц) забезпечила стабільну обробку сигналів у реальному часі з мінімальною латентністю. Сенсорну основу пристрою склали ультразвукові датчики HC-SR04P, які забезпечують високоточну детекцію перешкод у діапазоні від 2 до 400 см з похибкою ± 3 мм, що є оптимальним показником як для закритих приміщень, так і для відкритих просторів. Організація зворотного зв'язку реалізована через масив вібромоторів з робочою частотою 100–120 Гц, що гарантує генерацію чіткого та комфортного для сприйняття тактильного сигналу.

Суттєвим досягненням роботи стало розширення функціоналу системи шляхом інтеграції голосового інтерфейсу та розробки адаптивних алгоритмів керування. Реалізована функція голосового вводу-виводу забезпечує інтерактивність пристрою: за командою «Опиши об'єкт» система активує нейромережевий аналіз сцени та озвучує класи виявлених об'єктів (наприклад, «Перед вами людина»). Для підвищення ергономіки розроблено три спеціалізовані режими роботи: «Стандартний» для загального кругового огляду, «Цільовий аналіз» для точкової навігації у вузьких проходах та «Динамічний моніторинг», який адаптує чутливість сенсорів до швидкості наближення об'єктів, попереджаючи про раптові загрози.

Узагальнюючи отримані результати, можна стверджувати, що розроблений прототип демонструє високу ефективність та надійність функціонування завдяки вдалій синергії ультразвукової локації та алгоритмів штучного інтелекту YOLO. Створена система дозволяє користувачеві безпечно орієнтуватися у просторі, своєчасно отримуючи інформацію про тип та відстань до перешкод. Перспективи подальшого розвитку проєкту полягають у впровадженні стереоскопічних камер для побудови карти глибини, інтеграції GPS-модуля для глобальної навігації та персоналізації параметрів тактильного відгуку під індивідуальні фізіологічні особливості користувачів.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

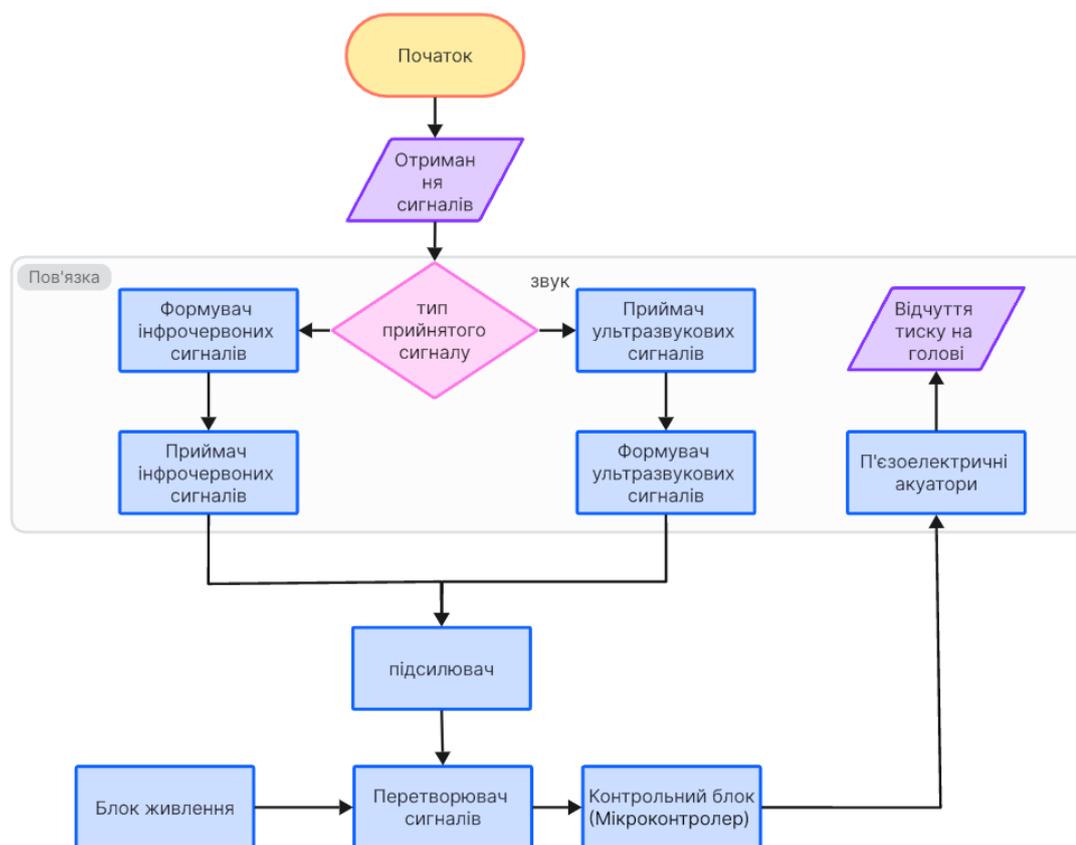
1. World report on vision. Geneva : World Health Organization, 2019. 160 p. URL: <https://www.who.int/publications/i/item/world-report-on-vision> (дата звернення: 14.11.2025)..
2. Onchocerciasis (River Blindness, Filariasis). Red Book: 2024–2027 Report of the Committee on Infectious Diseases. 33rd ed. 2024. P. 624–626. URL: https://doi.org/10.1542/9781610027359-s3_014_001.
3. Hersh M. A., Johnson M. A. Assistive Technology for Visually Impaired and Blind People. London : Springer-Verlag, 2008. 388 p.
4. ДСТУ EN ISO 9999:2018 (EN ISO 9999:2016, IDT; ISO 9999:2016, IDT). Засоби допоміжні для осіб з обмеженням життєдіяльності. Класифікація та термінологія. Київ : ДП «УкрНДНЦ», 2018. 140 с.
5. Panazan C.-E., Dulf E.-H. Intelligent Cane for Assisting the Visually Impaired. Technologies. 2024. Vol. 12, no. 6. P. 75. URL: <https://doi.org/10.3390/technologies12060075>
6. Ren S., He K., Girshick R., Sun J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2017. Vol. 39, no. 6. P. 1137–1149.
7. Redmon J., Divvala S., Girshick R., Farhadi A. You Only Look Once: Unified, Real-Time Object Detection. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Las Vegas, 2016. P. 779–788.
8. Liu W. et al. SSD: Single Shot MultiBox Detector. *Computer Vision – ECCV 2016*. Springer, Cham, 2016. P. 21–37.
9. Jocher G., Chaurasia A., Qiu J. Ultralytics YOLOv8. Version 8.0.0. 2023. URL: <https://github.com/ultralytics/ultralytics> (дата звернення: 10.11.2025).
10. Gong B., Patlas M. N. Learning Radiology AI Concepts by Seeing and Doing. *Canadian Association of Radiologists Journal*. 2023. P. 084653712311512. URL: <https://doi.org/10.1177/08465371231151239>.

11. ISO 23600:2007. Assistive products for persons with vision impairments and persons with vision and hearing impairments — Acoustic and tactile signals for pedestrian traffic lights. Geneva : ISO, 2007. 18 p.
12. ESP32 Series Datasheet. Version 4.1. Shanghai : Espressif Systems, 2023. 71 p.
URL:
https://www.espressif.com/sites/default/files/documentation/esp32_datasheet_en.pdf. (дата звернення: 11.11.2025).
13. HC-SR04 Ultrasonic Ranging Module. User Guide. Version 1.0. ElecFreaks, 2017. 9 p.
14. Monk S. Programming Arduino: Getting Started with Sketches. 2nd Edition. New York : McGraw-Hill Education, 2016. 192 p.
15. NCR18650B Lithium-Ion Battery Specification. Osaka : Panasonic Industrial Devices, 2018. 2 p. URL: <https://industrial.panasonic.com/cdbs/ww-data/pdf/ACA4000/ACA4000C249.pdf> (дата звернення: 12.10.2025).
16. Shorten C., Khoshgoftaar T. M. A survey on Image Data Augmentation for Deep Learning. *Journal of Big Data*. 2019. Vol. 6, no. 60. URL: <https://doi.org/10.1186/s40537-019-0197-0>.
17. Lin T.-Y. et al. Microsoft COCO: Common Objects in Context. *Computer Vision – ECCV 2014*. Springer, Cham, 2014. P. 740–755.
18. Goodfellow I., Bengio Y., Courville A. Deep Learning. Cambridge : MIT Press, 2016. 800 p.
19. Johnson, K., et al. (2016). Infrared Sensors in Assistive Technology for Blind and Visually Impaired. *Journal of Rehabilitation Research and Development*, 53(4), 491-501.
20. Padilla R., Netto S. L., da Silva E. A. B. A Survey on Performance Metrics for Object-Detection Algorithms. *2020 International Conference on Systems, Signals and Image Processing (IWSSIP)*. Niteroi, 2020. P. 237–242.
21. Everingham M. et al. The Pascal Visual Object Classes (VOC) Challenge. *International Journal of Computer Vision*. 2010. Vol. 88. P. 303–338.

22. Bradski G., Kaehler A. Learning OpenCV 3: Computer Vision in C++ with the OpenCV Library. Sebastopol : O'Reilly Media, 2017. 1018 p.
23. You Only Look Once (YOLO). Encyclopedia of Computer Graphics and Games. Cham, 2024. P. 2113. URL: https://doi.org/10.1007/978-3-031-23161-2_301332
24. Kingma D. P., Ba J. Adam: A Method for Stochastic Optimization. *International Conference on Learning Representations (ICLR)*. San Diego, 2015.
25. Методичні вказівки до виконання економічної частини магістерських кваліфікаційних робіт. / Укладачі В.О. Козловський, О.Й. Лесько, В.В.Кавецький. Вінниця : ВНТУ, 2021. 42 с.

Додаток А
(обов'язковий)

Ілюстративна частина
Структурна блок-схема асистивного пристрою



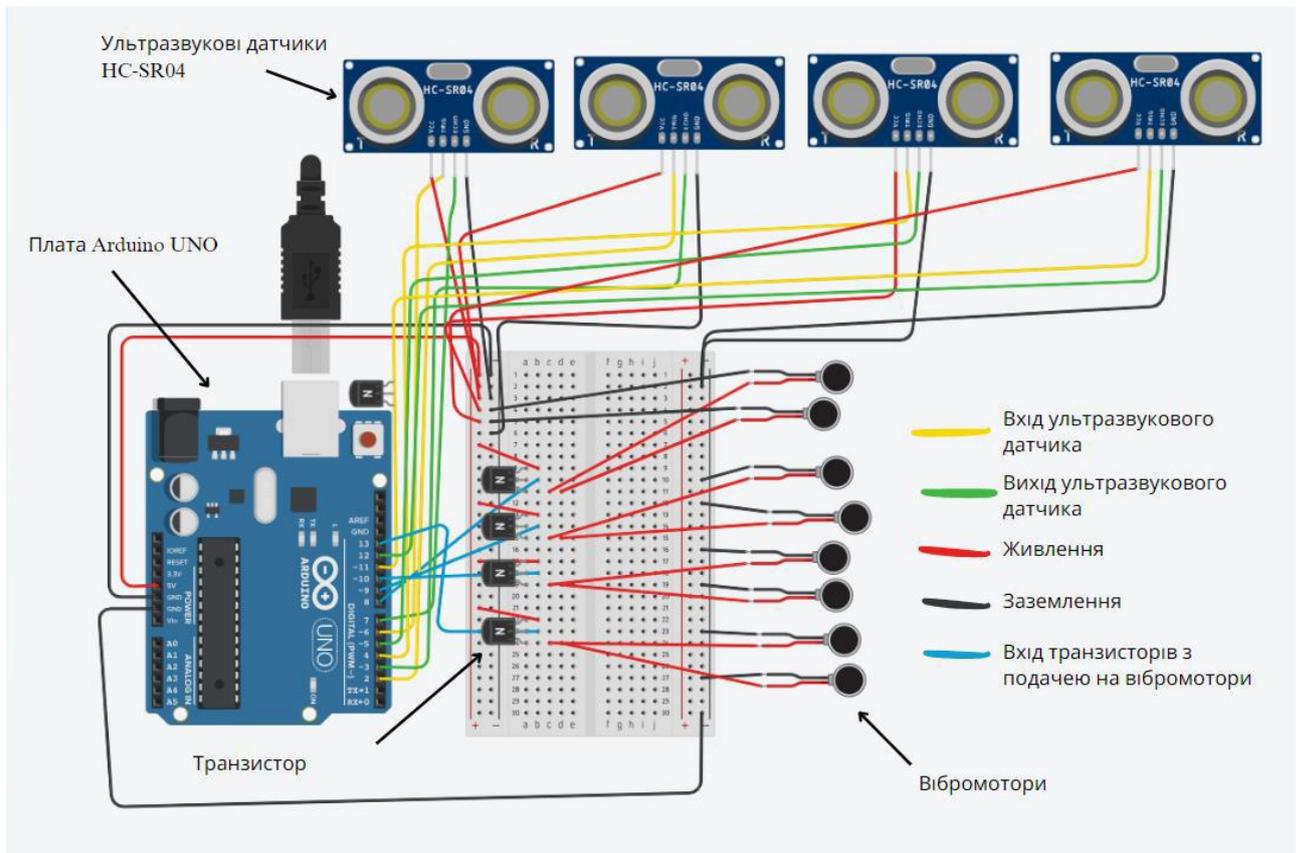


Рисунок А.1 – Віртуальна модель мультисенсорної системи в середовищі Tinkercad

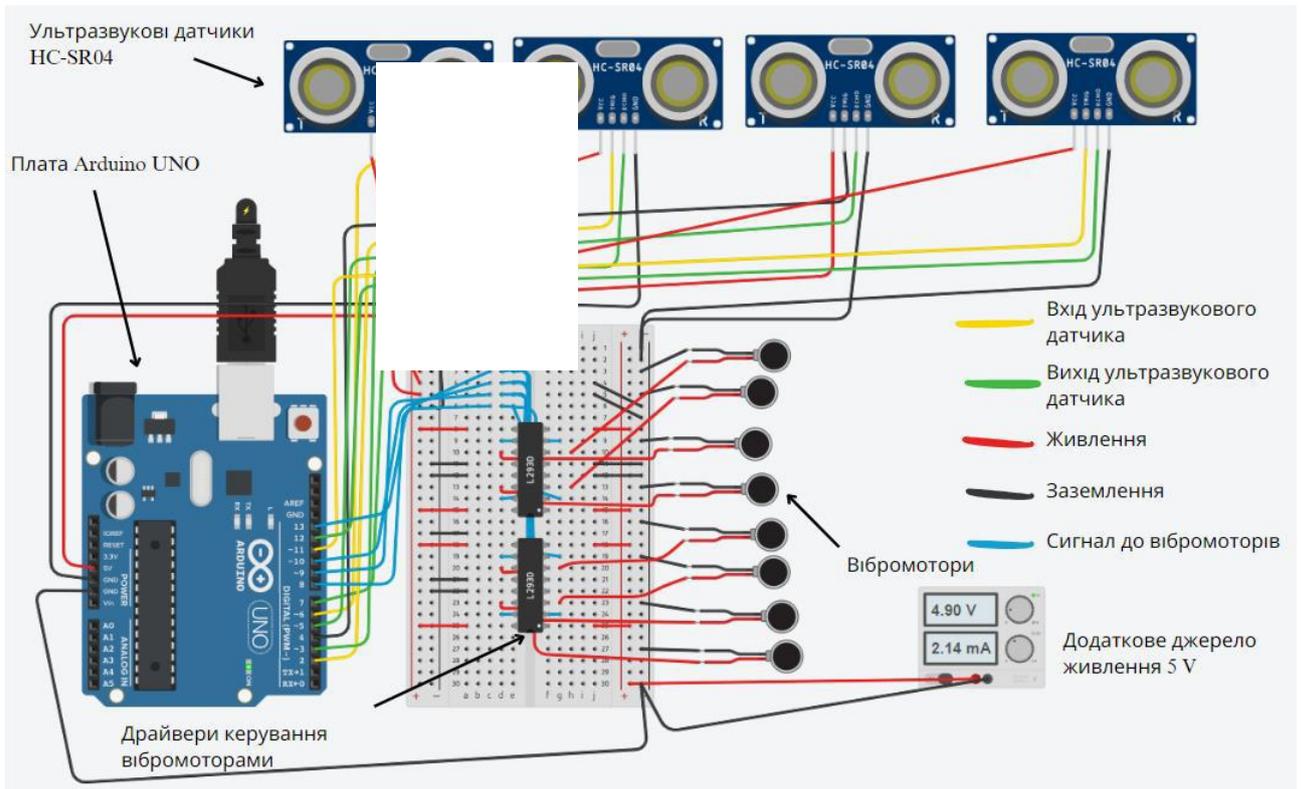


Рисунок А.2 – Фінальна принципова схема з використанням драйверів L293D

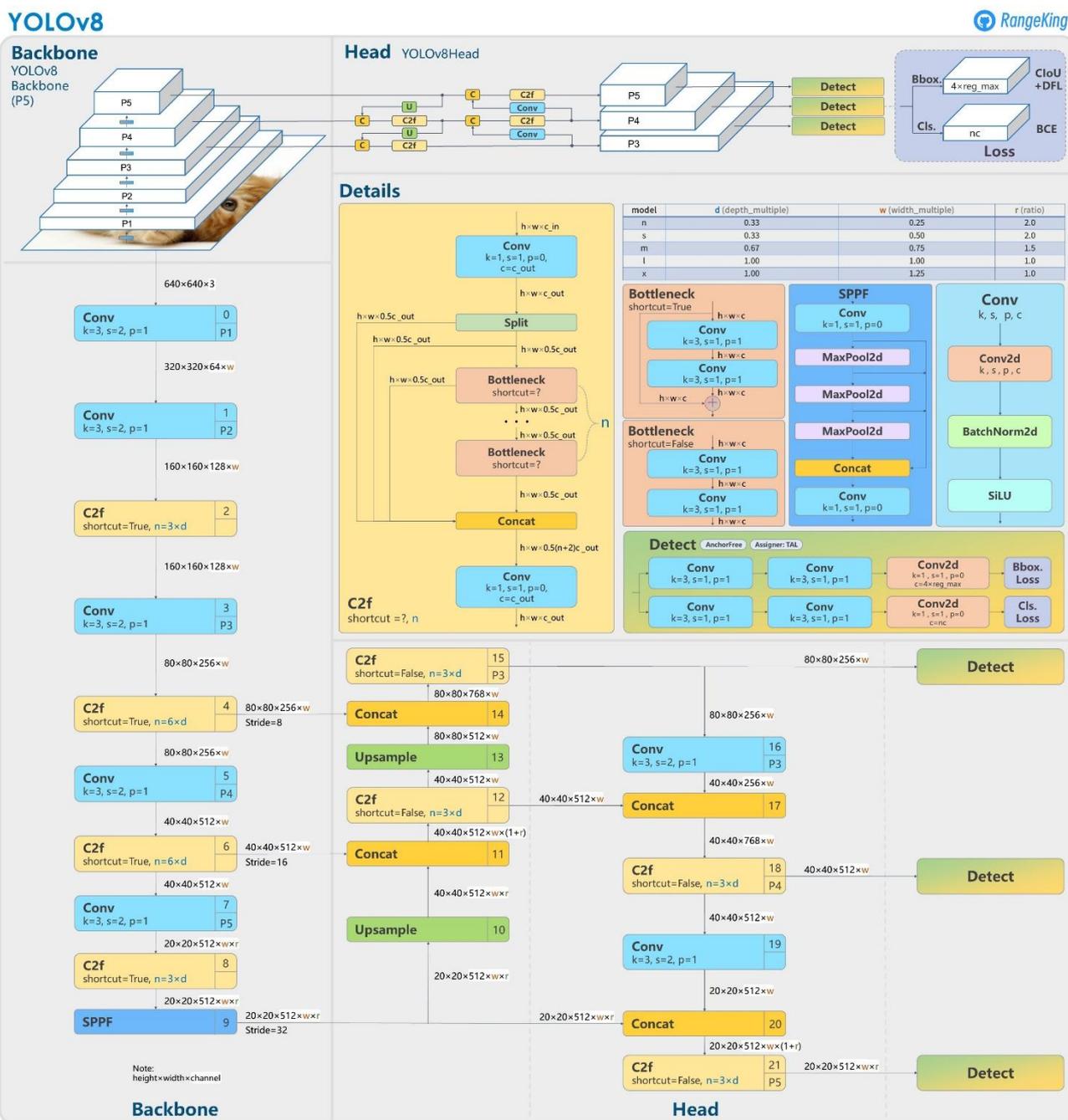


Рисунок А.3 – Структурна схема архітектури нейронної мережі YOLOv8

ПРОТОКОЛ ПЕРЕВІРКИ КВАЛІФІКАЦІЙНОЇ РОБОТИ

Назва роботи: Мобільна система позиціонування для людей з вадами зору

Тип роботи: магістерська кваліфікаційна робота

(бакалаврська кваліфікаційна робота / магістерська кваліфікаційна робота)

Підрозділ кафедра біомедичної інженерії та оптико-електронних систем

(кафедра, факультет, навчальна група)

Коефіцієнт подібності текстових запозичень, виявлених у роботі системою StrikePlagiarism (КПІ) 0,86 %

Висновок щодо перевірки кваліфікаційної роботи (відмітити потрібне)

- Запозичення, виявлені у роботі, оформлені коректно і не містять ознак академічного плагіату, фабрикації, фальсифікації. Роботу прийняти до захисту.
- У роботі не виявлено ознак плагіату, фабрикації, фальсифікації, але надмірна кількість текстових запозичень та/або наявність типових розрахунків не дозволяють прийняти рішення про оригінальність та самостійність її виконання. Роботу направити на доопрацювання.
- У роботі виявлено ознаки академічного плагіату та/або в ній містяться навмисні спотворення тексту, що вказують на спроби приховування недобросовісних запозичень. Робота до захисту не приймається.

Експертна комісія:

Коваль Л. Г., к.т.н., доц., зав. каф. БМІОЕС, гарант ОП

(прізвище, ініціали, посада)

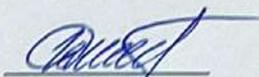

(підпис)

Павлов С. В., д.т.н., проф., професор каф. БМІОЕС

(прізвище, ініціали, посада)


(підпис)

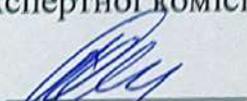
Особа, відповідальна за перевірку


(підпис)

Штофель Д. Х.
(прізвище, ініціали)

З висновком експертної комісії ознайомлений(-на)

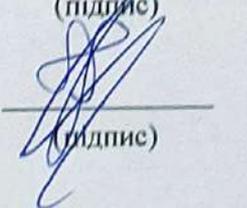
Керівник


(підпис)

Коваль Л. Г., к.т.н., доц., зав. каф. БМІОЕС

(прізвище, ініціали, посада)

Здобувач


(підпис)

Гзоїм В. А.

(прізвище, ініціали)