

ІНТЕЛЕКТУАЛЬНА СИСТЕМА СУПРОВОДУ ЛЮДЕЙ З ВАДАМИ ЗОРУ

Н.М. Омельчук¹, Я.Ю. Дорогий²

¹ Department of Information Systems and Technologies, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, Kyiv, Ukraine

² Department of Applied Mathematics and Informatics, Donetsk National Technical University, Drohobych, Ukraine

E-mail: omelchuk.nazarii@iit.kpi.ua

Отримано 01.11.2024

Прийнято до публікації 01.12.2024

Опубліковано 31.12.2024

АНОТАЦІЯ

Стаття присвячена аналізу сучасних технологій, що використовуються для створення інтелектуальних систем супроводу людей з вадами зору. Актуальність цієї теми зумовлена необхідністю розробки ефективних та доступних рішень, які сприяють полегшенню повсякденного життя людей з обмеженими можливостями, зокрема тих, хто має порушення зору. Метою статті є всебічний аналіз існуючих технологій, вивчення переваг та недоліків наявних рішень, а також визначення можливих напрямків удосконалення таких систем. У статті розглянуті провідні рішення у цій сфері, зокрема Aira, Be My Eyes, BlindSquare, Seeing AI, що дозволяють забезпечити навігацію та супровід для людей з вадами зору за допомогою мобільних пристроїв, камер та інших сенсорних систем.

Автори також досліджують потреби користувачів цих систем, визначаючи основні виклики, з якими стикаються люди з обмеженим зором, зокрема виявлення перешкод, надання своєчасної інформації про навколишнє середовище та забезпечення зручності використання. На основі цього аналізу сформульовані вимоги до майбутніх систем супроводу, а також розроблена архітектура інтелектуальної системи, що включає компоненти для розпізнавання об'єктів, перешкод, навігації та голосового супроводу.

Особливу увагу приділено інтеграції компонентів розпізнавання об'єктів та перешкод, що здійснюється за допомогою технологій комп'ютерного зору, таких як OpenCV та нейронні мережі. Окрім того, розглянуто можливість інтеграції голосових асистентів для забезпечення зручного та доступного зворотного зв'язку. У статті також наведено методи, що застосовуються для розробки та тестування таких систем, включаючи структурний і порівняльний аналіз існуючих рішень, а також експериментальні методи для перевірки ефективності програмного забезпечення.

Дослідження та розробка інтелектуальних систем супроводу людей з вадами зору є важливим кроком до створення доступного і безбар'єрного середовища для осіб з обмеженими можливостями, що відкриває нові можливості для покращення якості їхнього життя.

Ключові слова: інтелектуальна система супроводу, допомога для людей з вадами зору, веб-застосунок, комп'ютерний зір, розпізнавання об'єктів, YOLOv3, голосовий супровід, навігація, адаптивні технології.

ВСТУП

Інтелектуальні системи супроводу є важливим інструментом для полегшення повсякденного життя людей з вадами зору, забезпечуючи їм більшу автономність та самостійність у навігації по навколишньому середовищу. Завдяки використанню новітніх технологій, таких як комп'ютерний зір, штучний інтелект, системи глобального позиціонування (GPS), а також голосові асистенти, ці системи можуть допомогти користувачам орієнтуватися в просторі, визначати перешкоди, надавати інформацію про навколишні об'єкти та забезпечувати безпеку руху.

Основною метою таких систем є створення безбар'єрного середовища, яке дозволяє людям з обмеженими можливостями ефективно виконувати повсякденні завдання, зокрема, самостійно переміщатися в просторі. Для цього використовуються технології комп'ютерного зору, які здатні розпізнавати об'єкти, виявляти перешкоди, а також надавати точну інформацію про навколишнє середовище в реальному часі. Це дає можливість людям з вадами зору отримувати необхідну інформацію для прийняття рішень при пересуванні, що значно покращує їх якість життя.

Особливо важливим є використання алгоритмів глибинного навчання та штучного інтелекту для розпізнавання об'єктів у реальному часі, що дозволяє створювати високоефективні системи, здатні адаптуватися до змінюваних умов навколишнього середовища. Такі системи використовують зображення, отримані з камер або інших сенсорних пристроїв, для виявлення перешкод і визначення їхніх характеристик. Ці технології дають змогу створювати рішення, що дозволяють не тільки виявляти об'єкти, але й забезпечувати інтуїтивно зрозуміле сприйняття цих даних користувачем за допомогою голосових або візуальних повідомлень.

Важливим аспектом є також інтеграція системи супроводу з іншими технологіями, такими як глобальне позиціонування, що дозволяє точно визначати місцезнаходження користувача і здійснювати навігацію на основі реальних координат. Така інтеграція відкриває нові можливості для забезпечення безпеки та зручності пересування людей з вадами зору в складних міських умовах.

АНАЛІЗ ЛІТЕРАТУРНИХ ДЖЕРЕЛ ТА ПОСТАНОВКА ПРОБЛЕМИ

Розглянемо декілька досліджень, які дозволяють досить детально зрозуміти стан справ у досліджуваній сфері та можливі напрямки подальших досліджень.

У статті [1] аналізується глобальна поширеність сліпоти та порушень зору на близьку та дальню відстань. Це дослідження виявляє важливі статистичні тенденції щодо того, як проблеми зору поширюються по всьому світу, що стає основою для розробки інклюзивних технологій підтримки людей з порушеннями зору. Ось чому інноваційні технології, що використовують штучний інтелект і мобільні системи для допомоги таким особам, є надзвичайно актуальними.

У роботі [7] Б. П. Борківський та В. М. Теслюк описують використання нейромережевих технологій для розпізнавання об'єктів у мобільних системах для людей з порушеннями зору. Це дослідження також розглядає, як ці технології можуть допомогти у розв'язанні важливих задач, таких як виявлення перешкод і надання візуальних підказок користувачам в реальному часі, що значно підвищує автономність та безпеку цих людей.

Дослідження [8] пропонує автоматизовану систему для покращення мобільності осіб з порушеннями зору. Це рішення використовує інтегровані сенсори для надання користувачам простих навігаційних вказівок, що покращує їх здатність орієнтуватися у просторі без сторонньої допомоги. Ця технологія використовує штучний інтелект для точного аналізу навколишнього середовища та адаптується до різних умов.

У статті [9] описано, як технології штучного інтелекту можуть бути використані для розвитку розумної навігації для людей з порушеннями зору. У роботі розглядаються методи використання таких технологій для розпізнавання навколишнього середовища, а також надання користувачам необхідної інформації для покращення їх орієнтації і забезпечення зручності в повсякденному житті.

Робота [10] надає огляд системи, заснованої на технологіях AIoT (Artificial Intelligence of Things), для людей з порушеннями зору. Ця система пропонує інтеграцію сенсорних технологій та штучного інтелекту, що дозволяє створювати більш точні та ефективні рішення для навігації та орієнтації людей з обмеженим зором. Застосування таких систем є перспективним у контексті

розвитку інтелектуальних помічників для людей з обмеженими можливостями.

У роботі [11] автори досліджують використання інформаційних технологій для проектування інерціальних датчиків, що застосовуються в системах орієнтації та навігації. Така технологія є основою для створення систем підтримки для осіб з порушеннями зору, що дозволяють їм ефективно орієнтуватися в просторі без потреби у допомозі від сторонніх осіб.

У джерелі [12] наведено огляд GPS-систем позиціонування, що мають велике значення для точності навігаційних систем, призначених для людей з порушеннями зору. GPS-системи, як частина глобальних навігаційних технологій, стають основою для створення таких інклюзивних рішень, де точність позиціонування і орієнтація є критично важливими для безпеки користувачів.

У статті [13] розглядаються основи роботи GPS-систем, які використовуються в мобільних пристроях для підтримки орієнтації людей з обмеженим зором. Дана стаття дає розуміння технічних аспектів роботи таких систем та їхньої взаємодії з іншими інтелектуальними технологіями для покращення мобільності осіб з порушеннями зору.

Технологічні рішення відіграють важливу роль у підвищенні рівня самостійності людей із вадами зору. Одним із таких інструментів є мобільний застосунок Aira [2], що використовує технології відеозв'язку для надання візуальної допомоги. Цей застосунок забезпечує користувачам можливість взаємодіяти зі спеціально підготовленими операторами, які надають інформаційну підтримку в реальному часі.

Мета статті полягає в аналізі сучасних технологій, які використовуються для створення інтелектуальних систем супроводу людей з вадами зору, зокрема, вивченні їхніх переваг, недоліків та можливостей удосконалення. У статті здійснено аналіз існуючих рішень, таких як Aira, Be My Eyes, BlindSquare, Seeing AI, а також вивчено потреби користувачів, на основі яких сформульовано вимоги до майбутньої системи. Крім того, розроблено архітектуру інтелектуальної системи супроводу, яка включає основні функціональні компоненти, зокрема, системи розпізнавання об'єктів, перешкод, навігації та голосового супроводу.

Основні завдання дослідження включають:

- аналіз існуючих технічних рішень у сфері інтелектуальних систем супроводу людей з вадами зору;
- визначення технічних вимог до майбутньої системи та створення її архітектури;

- інтеграція компонентів системи, таких як розпізнавання об'єктів і голосова взаємодія.

Очікувані результати роботи матимуть важливе практичне значення. Запропонована система сприятиме підвищенню рівня незалежності осіб із вадами зору та забезпечить їм більшу свободу дій у повсякденному житті. Крім того, система має потенціал для комерційного впровадження, що дозволить забезпечити її доступність для широкого кола користувачів у всьому світі.

МАТЕРІАЛИ ТА МЕТОДИ ДОСЛІДЖЕНЬ

У роботі для досягнення поставлених цілей використовувались методи структурного і порівняльного аналізу, що дозволили всебічно оцінити існуючі технології та рішення для створення інтелектуальних систем супроводу людей з вадами зору.

Для вивчення наукової та методичної літератури, а також онлайн-ресурсів було застосовано інформаційний і аналітичний підхід. Це дозволило отримати глибоке розуміння проблеми, виявити існуючі тенденції та порівняти переваги і недоліки різних підходів і рішень, таких як Aira [2], Be My Eyes [3-4], BlindSquare [5] та інших.

З метою перевірки ефективності теоретичних положень і розроблених технічних рішень були застосовані експериментальні методи, зокрема для тестування програмного коду системи розпізнавання об'єктів. Це дало змогу оцінити працездатність запропонованих підходів у реальних умовах, а також виявити можливі проблеми і шляхи їх усунення.

РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ

Призначення та область застосування. Інтелектуальна система супроводу для осіб із вадами зору призначена для забезпечення допомоги у просторовій орієнтації та навігації шляхом використання універсального програмного забезпечення. Система є багатоплатформним рішенням, яке може працювати на різних пристроях, зокрема смартфонах, планшетах, комп'ютерах та інших гаджетах, забезпечуючи доступність для широкого кола користувачів. У її основі лежать технології штучного інтелекту, комп'ютерного зору та машинного навчання, що дозволяють розпізнавати об'єкти та перешкоди в реальному часі та надавати користувачам голосові інструкції, спрямовані на безпечніше й незалежне пересування.

Система може бути корисною для орієнтації на вулиці завдяки розпізнаванню дорожніх знаків, перехресть, громадського транспорту та інших об'єктів міської інфраструктури, що сприятиме підвищенню мобільності

користувачів у міських умовах. У великих громадських просторах, таких як торговельні центри, вокзали, лікарні чи офісні будівлі, система сприятиме ідентифікації дверей, ліфтів, сходів та інших елементів, полегшуючи самостійне пересування.

Під час участі у громадських або соціальних заходах система допоможе користувачам взаємодіяти з навколишнім середовищем, включаючи людей, об'єкти та простори, завдяки інтерактивним підказкам та голосовому супроводу. Зокрема, система забезпечуватиме визначення маршрутів, розпізнавання станцій і транспортних засобів, надаючи інформацію про розклади, що значно полегшить пересування в транспортній мережі.

Багатофункціональність програмного забезпечення забезпечує його гнучкість та адаптацію до різних умов використання. Система може застосовуватися як для особистих потреб, так і в межах соціальних програм, спрямованих на підтримку осіб із вадами зору. Інтеграція з різними типами пристроїв дозволяє врахувати індивідуальні потреби користувачів, сприяючи їхній активній участі в соціальному житті, підвищенню рівня незалежності та якості життя.

Ця робота має важливе значення з огляду на необхідність глобального поліпшення умов для осіб із порушеннями зору. На основі сучасних технологій створення багатофункціональної інтелектуальної системи супроводу є ключовим кроком до підвищення мобільності, незалежності та соціальної інтеграції таких людей. Система сприятиме не лише безпечнішій навігації в навколишньому середовищі, але й забезпечить оперативний доступ до важливої інформації у реальному часі. Її впровадження дозволить суттєво покращити якість життя користувачів, зменшити залежність від сторонньої допомоги та підвищити рівень їхньої соціальної активності.

МОБІЛЬНИЙ ЗАСТОСУНОК AIRA

Aira [2] є важливим кроком уперед у впровадженні технологій, спрямованих на підтримку людей із порушеннями зору. Вона дозволяє користувачам отримувати миттєву допомогу, що значно розширює їхні можливості для самостійного пересування, виконання повсякденних завдань і соціальної інтеграції. Водночас для забезпечення максимального ефекту від використання цієї технології важливо вирішити питання доступності та зменшення залежності від зовнішніх факторів, таких як інтернет-з'єднання чи оператори (Рис. 1).



Рис. 1. Мобільний застосунок Aira [2]

Мобільний застосунок Aira пропонує низку інноваційних функцій, серед яких:

- *відеоінтерфейс* – за допомогою відеозв'язку Aira підключає користувача до оператора, який отримує зображення з камери смартфона або спеціальних окулярів, оснащених вбудованою камерою. Це дозволяє операторам бачити навколишнє середовище користувача та надавати йому необхідну інформацію;
- *допомога від операторів* – оператори сприяють орієнтації в нових місцях, розпізнають тексти, ідентифікують об'єкти, прокладають маршрути та забезпечують безпеку. Такі можливості роблять застосунок універсальним рішенням для різних життєвих ситуацій;
- *підтримка на багатьох платформах* – Aira доступна для мобільних пристроїв на базі iOS та Android, що забезпечує її зручність і доступність для користувачів;
- *модель підписки* – застосунок працює на основі гнучкої системи тарифних планів. У деяких випадках користувачі можуть отримувати безкоштовний доступ завдяки підтримці партнерських організацій чи фондів;
- *безпека та конфіденційність* – дані передаються у зашифрованому вигляді, що забезпечує конфіденційність. Оператори проходять навчання аби гарантувати дотримання етичних стандартів.

До основних переваг застосунку Aira належать:

- забезпечення доступу до допомоги в реальному часі,
 - універсальність використання в різних ситуаціях – від навігації в нових місцях до виконання повсякденних завдань;
 - індивідуалізована підтримка, яка враховує специфічні потреби користувачів.
- Разом із перевагами Aira має і певні недоліки, які слід враховувати:
- залежність від інтернет-з'єднання, яка може бути критичною у віддалених регіонах чи під час подорожей;

- модель підписки, яка може бути фінансово недоступною для частини користувачів, особливо в країнах із низьким рівнем доходу;
- неавтономність роботи системи, оскільки функціонування застосунку залежить від наявності операторів.

МОБІЛЬНИЙ ЗАСТОСУНОК *BE MY EYES*

Мобільний застосунок *Be My Eyes* [3-4] є ефективним інструментом для забезпечення візуальної допомоги людям із вадами зору, орієнтованим на спрощення їхнього повсякденного життя. Завдяки платформі користувачі мають змогу підключатися до волонтерів із різних країн світу через відеозв'язок, що дозволяє отримувати оперативну підтримку в найрізноманітніших ситуаціях. Механізм роботи застосунку передбачає передачу зображень оточення від користувача до волонтера, який, базуючись на отриманій візуальній інформації, надає необхідну допомогу.

Основні функції та можливості *Be My Eyes* наступні:

- забезпечення відеозв'язку між користувачами та волонтерами;
- широка база волонтерів із різних країн, що гарантує доступність підтримки в будь-який час;
- інтуїтивно зрозумілий інтерфейс, що спрощує користування застосунком;
- багатомовна підтримка, яка робить додаток доступним для широкої аудиторії;
- можливість отримання допомоги у виконанні повсякденних завдань;
- інтеграція з компаніями та організаціями, які спеціалізуються на допомозі людям із порушеннями зору;
- безкоштовна основа використання, що забезпечує доступність для користувачів незалежно від їхнього фінансового стану.

Приклад інтерфейсу мобільного застосунку *Be My Eyes* наведено на Рис.2.

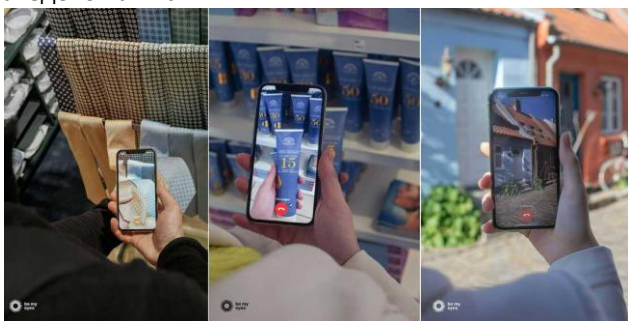


Рис. 2. Мобільний застосунок *Be My Eyes* [3]

Основними перевагами застосунку є оперативність отримання допомоги, що забезпечується завдяки великій базі активних волонтерів. Можливість користування додатком у будь-якому місці за наявності мобільного зв'язку значно розширює його доступність. Крім того, платформа сприяє соціальній інтеграції, встановлюючи взаємодію між користувачами та волонтерами, що зміцнює соціальні зв'язки. Безкоштовний доступ до функцій додатка робить його популярним і придатним для використання широким колом людей із вадами зору.

Незважаючи на значні переваги, *Be My Eyes* має певні обмеження. Доступність підтримки може залежати від кількості активних волонтерів у конкретний момент, що іноді призводить до затримок у наданні допомоги. Хоча в більшості випадків час очікування є незначним, можливість затримки залишається фактором ризику. Іншою проблемою є забезпечення конфіденційності, оскільки користувачі передають стороннім особам інформацію про своє оточення, що потребує обачності у використанні застосунку.

МОБІЛЬНИЙ ЗАСТОСУНОК *BLINDSQUARE*

Мобільний застосунок *BlindSquare* [5] є сучасним інструментом навігації, розробленим для людей із порушеннями зору. Використовуючи GPS і технології доповненої реальності, програма надає голосові вказівки, допомагаючи користувачам орієнтуватися в міському середовищі та знаходити важливі об'єкти, такі як кафе, магазини, транспортні зупинки тощо. *BlindSquare* інтегрується з картографічними сервісами для забезпечення точних даних про місцезоположення користувача та навколишнє середовище.

Однією з ключових функцій програми є використання GPS для визначення точного місця розташування користувача та навігації до потрібних об'єктів. *BlindSquare* отримує інформацію з картографічних сервісів, таких як *OpenStreetMap*, і надає користувачам голосові інструкції. Наприклад, програма може повідомити про кафе на відстані 200 метрів та запропонувати маршрут до нього.

Іншою важливою функцією є інтеграція технологій доповненої реальності. Використовуючи камеру смартфона, користувачі можуть активувати режим AR (augmented reality), який візуалізує розташування об'єктів у полі зору. Це додає додатковий рівень інформації про навколишнє середовище, сприяючи кращій орієнтації.

BlindSquare також підтримує функцію спільноти, яка дозволяє користувачам залишати відгуки про об'єкти, з якими вони взаємодіяли. Ця інформація надає актуальні дані про доступність, якість обслуговування та інші

аспекти, що можуть бути корисними для інших користувачів. Застосунок активно співпрацює з організаціями, які спеціалізуються на питаннях доступності об'єктів для людей із обмеженими можливостями.

Програма підтримує гнучке налаштування – користувачі можуть вибирати мову голосових підказок, частоту сповіщень, а також типи об'єктів, про які вони хочуть отримувати інформацію. Це забезпечує персоналізований досвід, враховуючи індивідуальні потреби користувачів.

BlindSquare надає наступні переваги:

- точну навігацію завдяки інтеграції з картографічними сервісами;
- голосові інструкції для зручності використання;
- багатофункціональність завдяки доповненій реальності та функціям спільноти;
- підтримку налаштувань для персоналізації досвіду;
- інтеграцію з організаціями, що працюють над покращенням доступності.

Незважаючи на вищезазначені переваги, програма має деякі обмеження:

- залежність від стабільного GPS-з'єднання, що може бути проблемою в умовах слабого сигналу або в приміщеннях;
- необхідність використання сучасного смартфона, що може бути фінансово недоступним для окремих категорій користувачів;
- відсутність функції виявлення перешкод, таких як бордюри чи інші фізичні об'єкти, що може створювати труднощі під час пересування.

МОБІЛЬНИЙ ЗАСТОСУНОК SEEING AI

Seeing AI [6], розроблений компанією Microsoft, є безкоштовним мобільним застосунком, що перетворює візуальну інформацію на аудіоповідомлення. Його мета – забезпечити людей із вадами зору можливістю "бачити" навколишній світ за допомогою камери смартфона та штучного інтелекту. Завдяки широкому функціоналу користувачі можуть орієнтуватися у просторі, розпізнавати текст, об'єкти, обличчя, емоції та виконувати інші важливі завдання.

Мобільний застосунок Seeing AI представлений на Рис. 3.

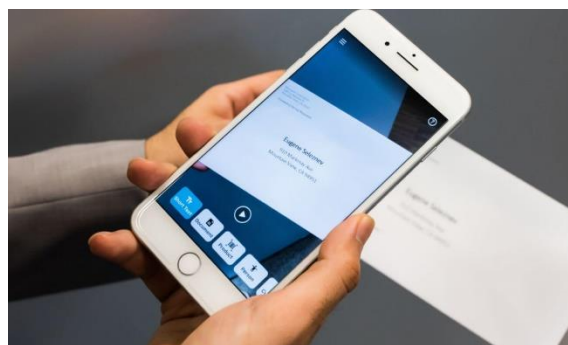


Рис. 3. Мобільний застосунок Seeing AI [6]

До основних функцій та можливостей Seeing AI належать наступні:

- розпізнавання тексту – програма зчитує друкований і рукописний текст із документів, книг або вивісок. У режимі "Документ" застосунок автоматично визначає межі сторінок для точнішого зчитування;
- штрих-коди та короткі тексти – Seeing AI допомагає знайти штрих-коди через звукові підказки та озвучує інформацію про товар. Крім того, програма миттєво зчитує короткі тексти, наприклад, з етикеток чи плакатів;
- опис сцени – застосунок аналізує зображення перед камерою, описуючи об'єкти, людей і їх дії, що допомагає користувачам краще орієнтуватися у просторі;
- розпізнавання обличчя та емоцій – програма ідентифікує знайомих людей та повідомляє про їхній емоційний стан;
- додаткові можливості – розпізнавання кольорів, ідентифікація валют та оцінка яскравості в приміщенні.

Перевагами використання Seeing AI є:

- доступність – застосунок безкоштовний та підтримує багато мов, що забезпечує глобальне використання;
- інтуїтивність – простий у використанні інтерфейс робить програму доступною для широкої аудиторії.
- широкий функціонал – від текстового розпізнавання до ідентифікації об'єктів, програма вирішує різноманітні завдання, що виникають у повсякденному житті.

Серед обмежень Seeing AI можна виділити наступні:

- технічна залежність – точність роботи залежить від якості камери смартфона, а для більшості функцій потрібне стабільне підключення до Інтернету;
- неточність алгоритмів – штучний інтелект може іноді надавати неправильні описи сцен;
- навігаційні обмеження – застосунок не виявляє фізичних перешкод, таких як бордюри або стовпи.

КОМПАРАТИВНИЙ АНАЛІЗ МОБІЛЬНИХ ЗАСТОСУНКІВ

Далі наведено компаративний аналіз розглянутих мобільних застосунків, розроблених для людей з вадами зору. Для забезпечення об'єктивної оцінки та порівняння різних програм було обрано низку критеріїв, що дозволяють врахувати їх основні функції, доступність та ефективність в реальних умовах використання. Ось ключові критерії, що використовуються для порівняння мобільних застосунків:

1. Основна функціональність

Цей критерій визначає, яку основну допомогу надає застосунок користувачам. Залежно від призначення програми, функції можуть варіюватися від надання допомоги через відеозв'язок з волонтерами (наприклад, *Be My Eyes* або *Aira*) до автономних функцій, таких як навігація на основі GPS або розпізнавання текстів і об'єктів за допомогою штучного інтелекту (як у *BlindSquare* чи *Seeing AI*).

2. Інтерфейс та зручність використання

Важливим аспектом є зручність і простота інтерфейсу. Оскільки мобільні застосунки для людей з вадами зору повинні бути інтуїтивно зрозумілими, критично важливо, щоб користувачі могли легко орієнтуватися в функціях програми. Простота навігації та доступність основних функцій забезпечують ефективне використання додатка.

3. Технічні вимоги та доступність

Під цим критерієм розуміється оцінка технічних ресурсів, необхідних для запуску програми. Зокрема,

важливими є вимоги до апаратного забезпечення, наявність певних функцій (наприклад, GPS або камера високої роздільної здатності), а також вимоги до операційної системи та програмного забезпечення.

4. Особливості взаємодії та персоналізація

Даний критерій оцінює, як користувачі можуть налаштовувати застосунок під свої потреби, чи є можливості для персоналізації (наприклад, вибір мови або налаштування частоти повідомлень). Важливими є також додаткові можливості, такі як інтеграція з соціальними мережами, можливість спілкування з волонтерами або іншими користувачами.

5. Переваги

Оцінка основних переваг, які надає застосунок. Це можуть бути такі аспекти, як швидкість надання допомоги, точність функцій (наприклад, розпізнавання об'єктів чи тексту), наявність безкоштовного доступу або можливість використовувати додаток без значних технічних обмежень.

6. Обмеження та недоліки

Важливим критерієм є оцінка обмежень, які можуть вплинути на ефективність застосунку. Це можуть бути технічні недоліки, необхідність стабільного Інтернет-з'єднання, обмеження в точності виконання функцій або залежність від зовнішніх факторів, таких як наявність активних волонтерів.

Ці критерії дозволяють провести порівняння мобільних застосунків на основі основних характеристик, що впливають на їх функціональність та ефективність для людей з вадами зору (Табл. 1).

Табл. 1. Компаративний аналіз застосунків

Критерій	<i>Be My Eyes</i>	<i>BlindSquare</i>	<i>Seeing AI</i>	<i>Aira</i>
Основна функціональність	Допомога через відеозв'язок з волонтерами	GPS-навігація, доповнена реальність	Розпізнавання текстів, об'єктів, емоцій	Допомога через відео та аудіозв'язок з асистентами
Інтерфейс	Простий, інтуїтивно зрозумілий	Інтуїтивний, потребує налаштувань	Простий, з налаштуваннями для персоналізації	Простий, інтеграція з віртуальними асистентами
Технічні вимоги	Мінімальні, потрібен стабільний Інтернет	GPS, картографічні сервіси	Камера високої якості, Інтернет	Стабільний Інтернет, доступ до відеозв'язку
Особливості досвіду	Взаємодія з волонтерами, соціальна підтримка	Автономна навігація, доповнена реальність	Самостійність у взаємодії з оточенням	Пряма допомога в реальному часі від асистентів
Переваги	Миттєва допомога, безкоштовний, доступний	Точна навігація, зручний для мобільності	Широкий функціонал, безкоштовний	Професійна допомога у реальному часі,

				доступ до допомоги в будь-який час
Обмеження	Залежить від волонтерів, питання конфіденційності	Залежність від GPS, потребує специфічного смартфона	Залежність від камери, можливі неточності	Потрібно стабільне з'єднання, платні послуги

ВИМОГИ ДО СИСТЕМИ

Перед початком розробки системи важливо чітко визначити її функціональні та нефункціональні вимоги, враховуючи результати аналізу існуючих рішень. Це забезпечить створення архітектури, яка задовольнить потреби користувачів та ефективно виконуватиме покладені на систему завдання. Уся функціональність

повинна бути доступна користувачам, включаючи виявлення об'єктів та перешкод у реальному часі, розпізнавання тексту з зображень, отримання навігаційних інструкцій на основі GPS, передача інформації через голосові підказки, а також можливість відправлення повідомлень про допомогу з визначенням місцезнаходження. Варіанти використання усіх функцій системи зображено на діаграмі прецедентів на Рис. 4.

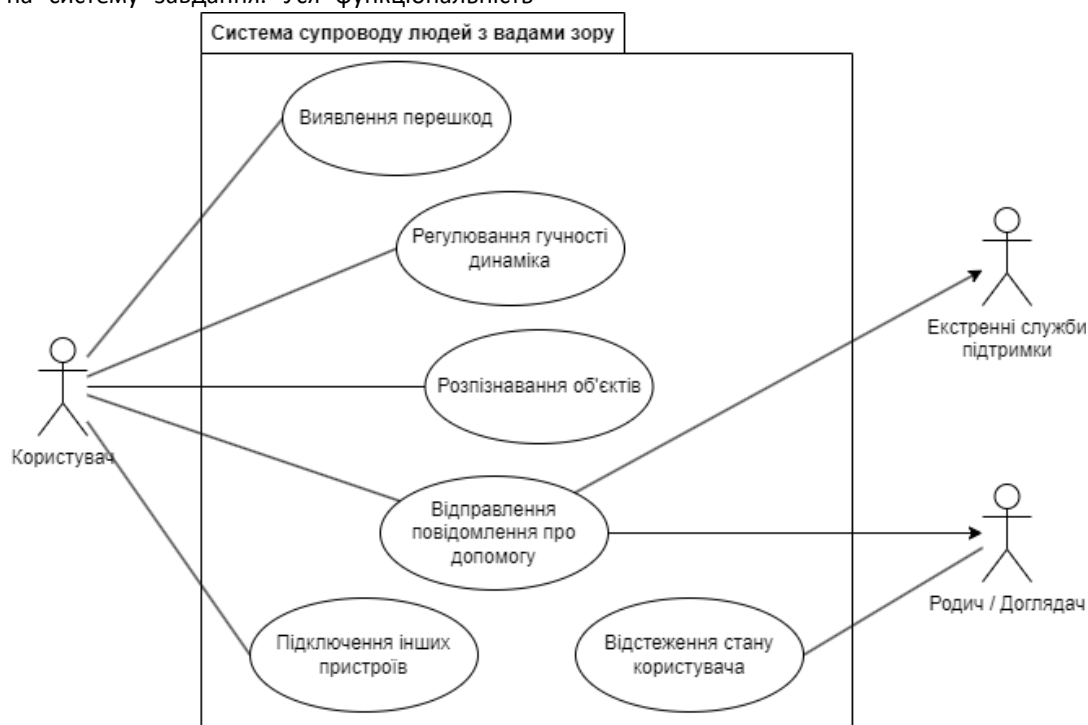


Рис. 4. Діаграма прецедентів

Функціональні вимоги системи. Враховуючи основні потреби користувачів з вадами зору, а також існуючі технології, були визначені такі функціональні вимоги до системи:

- **виявлення об'єктів та перешкод** – система повинна мати можливість ідентифікувати об'єкти в полі зору камери, а також попереджати користувача про перешкоди на шляху. Це включає розпізнавання таких об'єктів, як транспортні засоби, пішоходи, дорожні знаки та інші важливі об'єкти;

- **розпізнавання тексту** – система повинна бути здатна зчитувати текст, який може зустрічатися в

навколишньому середовищі – на вивісках, дорожніх знаках, документах тощо;

- **навігаційні можливості** – система повинна надавати точні навігаційні інструкції на основі GPS. Вона повинна забезпечувати як загальні маршрути в містах, так і локальні орієнтири в приміщеннях;

- **аудіовізуальний інтерфейс** – система повинна підтримувати голосові підказки та візуальні сповіщення, забезпечуючи чітке та своєчасне надання інформації про оточення, зміни в ньому та інструкції для безпечного пересування;

- **інтеграція з іншими системами** – система повинна мати можливість інтегруватися з іншими

застосунками та сервісами, такими як транспортні або навігаційні програми, через API, що дозволить додатково отримувати дані про транспортні маршрути, дорожній рух тощо.

Нефункціональні вимоги системи. Нефункціональні вимоги визначають важливі характеристики, зокрема продуктивність, надійність, зручність використання та інші аспекти, які впливають на роботу системи в реальних умовах. Система повинна забезпечувати наступні нефункціональні вимоги:

- **продуктивність** – система повинна працювати в реальному часі, з мінімальними затримками при обробці зображень та наданні зворотного зв'язку. Час реакції на виявлені об'єкти і перешкоди має бути не більше кількох секунд;

- **надійність і стійкість до помилок** – система повинна бути стійкою до збоїв апаратного забезпечення, таких як втрата GPS-сигналу або тимчасова втрата зображення з камери, і бути здатною відновлюватися після таких помилок без втрат в ефективності;

- **масштабованість** – система повинна адаптуватися до різних апаратних платформ, зокрема працювати на мобільних пристроях з обмеженими ресурсами, таких як смартфони або планшети;

- **безпека та конфіденційність** – система повинна гарантувати захист особистих даних користувачів, зокрема інформації про місцеположення, навколишнє середовище, та забезпечувати шифрування всіх даних, що передаються через мережу;

- **енергоефективність** – система повинна бути оптимізована для мінімального енергоспоживання, щоб забезпечити тривалу роботу на мобільних пристроях, враховуючи використання GPS, камери та інших ресурсів;

- **зручність у користуванні** – інтерфейс повинен бути інтуїтивно зрозумілим і доступним для людей з обмеженими можливостями, зокрема з вадами зору. Голосові підказки мають бути чіткими та легкими для

сприйняття, а візуальні елементи інтерфейсу – простими для взаємодії.

Обмеження та припущення. У процесі розробки системи необхідно враховувати такі обмеження та припущення:

1. Система повинна працювати у змінних умовах освітлення (наприклад, в темний час доби або при яскравому сонячному світлі).

2. Продуктивність системи може залежати від якості з'єднання з Інтернетом і точності GPS, що може бути обмежено в деяких регіонах або приміщеннях.

3. Користувачі системи можуть мати різний рівень технічної підготовки, тому необхідно забезпечити адаптацію інтерфейсу та функцій як для досвідчених користувачів, так і для новачків.

АРХІТЕКТУРА ТА СТРУКТУРА ІНТЕЛЕКТУАЛЬНОЇ СИСТЕМИ СУПРОВОДУ ЛЮДЕЙ З ВАДАМИ ЗОРУ

Далі представлена архітектура інтелектуальної системи супроводу людей з вадами зору, яка використовує штучний інтелект для виявлення об'єктів, перешкод, розпізнавання тексту, навігації та надання голосових підказок. Система працює на сервері, а користувач взаємодіє з нею через мобільний застосунок, отримуючи весь необхідний функціонал.

Основні підсистеми системи:

- система розпізнавання об'єктів та перешкод;
- система розпізнавання тексту;
- навігаційна система;
- система голосового оповіщення;
- система передачі сигналу SOS та визначення місцезнаходження;
- система управління пристроєм.

Архітектура системи і взаємодія її компонентів проілюстрована на Рис. 5.

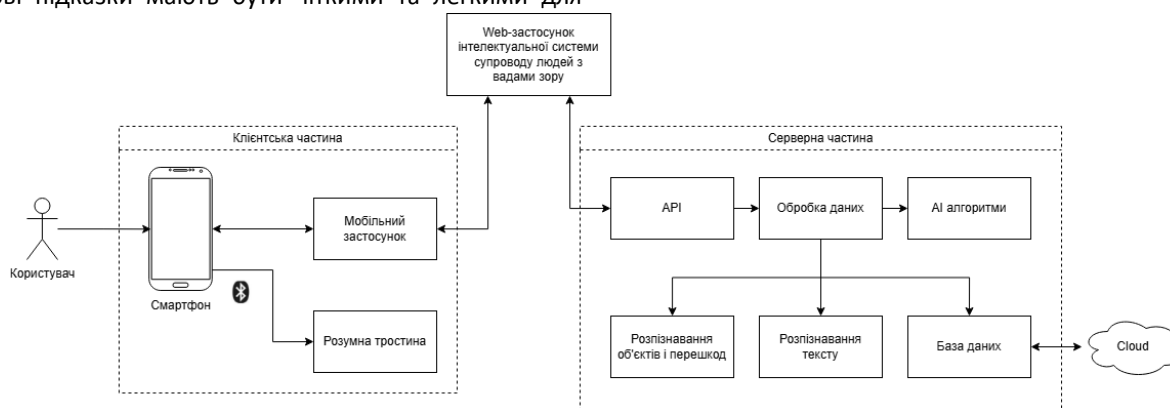


Рис. 5. Архітектура інтелектуальної системи супроводу людей з вадами зору

Система розпізнавання об'єктів та перешкод. Ця система забезпечує ідентифікацію об'єктів і перешкод у реальному часі, отримуючи відеопотік з камери смартфона користувача, який передається на сервер для обробки. В системі використовуються алгоритми

комп'ютерного зору, такі як CNN (Convolutional Neural Network), для розпізнавання об'єктів і перешкод та моделі глибокого навчання, такі як YOLO (You Only Look Once) або SSD (Single Shot Multibox Detector), для виявлення об'єктів на зображеннях (Рис. 6).

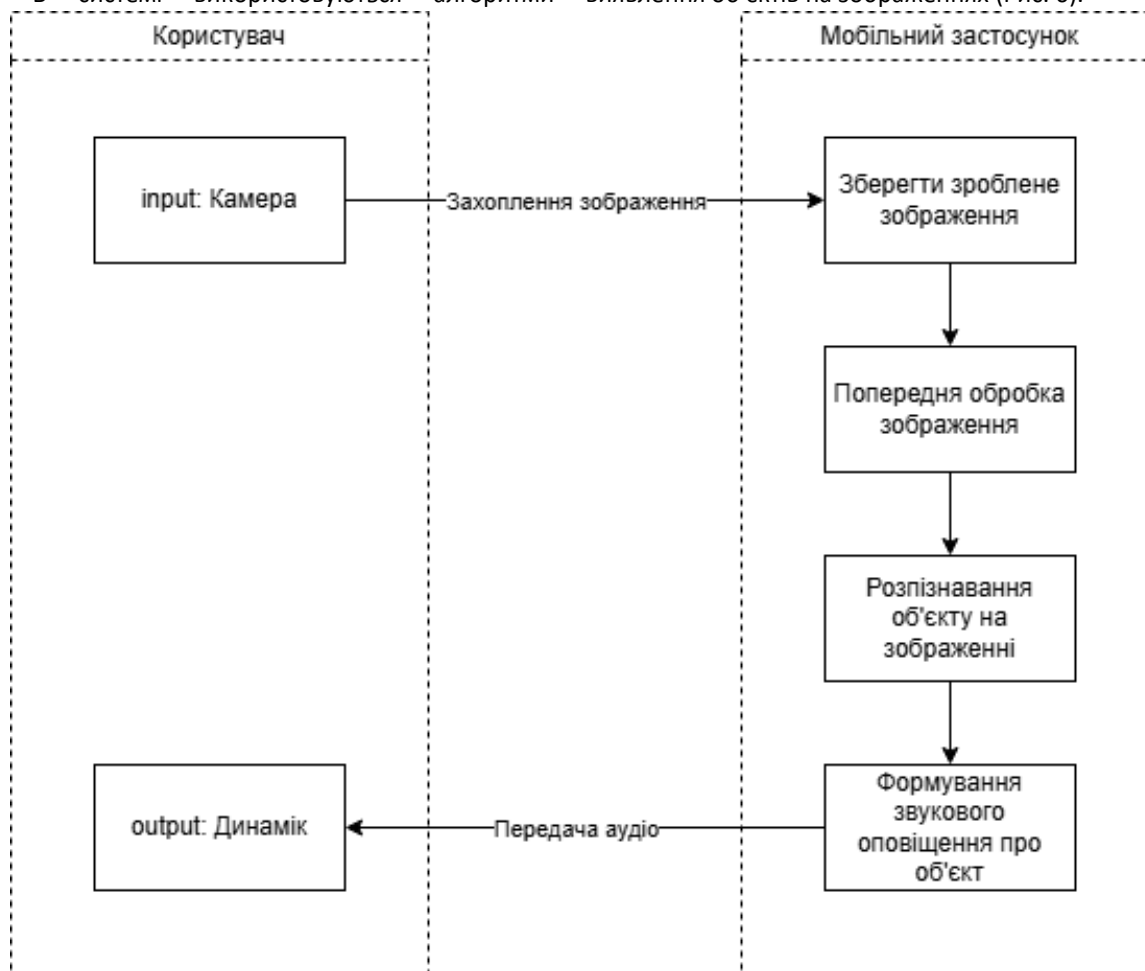


Рис. 6. Структурна схема системи розпізнавання об'єктів та перешкод

Після обробки на сервері, результати (у вигляді текстових або голосових підказок) повертаються до користувача через мобільний застосунок.

Основні компоненти системи наступні:

- сервер – обробляє зображення за допомогою алгоритмів глибокого навчання;
- камера смартфона – захоплює відеопотік для передачі на сервер.
- мобільний застосунок – взаємодіє з користувачем, відображаючи результати обробки.

Система розпізнавання тексту. Ця система дозволяє користувачеві отримувати текстову інформацію з навколишнього середовища (наприклад, дорожні знаки, вивіски або документи). Відеопотік з текстовими даними передається на сервер, де він обробляється за допомогою

OCR (оптичне розпізнавання символів), з використанням таких бібліотек як Tesseract, або сторонніх хмарних сервісів, таких як Google Cloud Vision API.

Після обробки текст передається користувачу для озвучення або відображення на екрані.

Навігаційна система. Навігаційна система допомагає користувачеві орієнтуватися в просторі, надаючи інформацію про його поточне місцезнаходження та маршрути. Вона використовує дані з GPS для прокладання маршрутів, а також сторонні API (наприклад, Google Maps API або OpenStreetMap) для обробки даних про навколишнє середовище і генерування інструкцій.

Основні компоненти системи наступні:

- GPS-модуль смартфона – визначає місцезнаходження користувача;

- сервер – обробляє запити на маршрутизацію та генерує інструкції;

- мобільний застосунок – надає голосові інструкції користувачеві.

Система голосового оповіщення. Система голосового оповіщення генерує голосові підказки на основі текстової інформації, що надходить із серверу. Технологія TTS (Text-to-Speech) використовується для перетворення тексту в мовлення. Сервіси на зразок Google TTS дозволяють генерувати голосові повідомлення, що допомагають користувачу орієнтуватися в навколишньому середовищі.

Основні компоненти системи:

- сервер – генерує текстові повідомлення для озвучення;

- мобільний застосунок – приймає і відтворює аудіо через динаміки смартфона.

Система передачі сигналу SOS та визначення місцезнаходження. Ця система дозволяє користувачеві передавати сигнал SOS в екстрених ситуаціях, разом із поточним місцезнаходженням. Сигнал передається через мобільну мережу за допомогою SMS або через сторонні сервіси, що використовують GSM-модуль смартфона.

Система управління пристроєм. Система управління пристроєм є основою для ефективної взаємодії користувача з системою. Вона забезпечує управління всіма компонентами і процесами, що взаємодіють між собою.

Основні компоненти системи:

- сервер – для обробки всіх даних і управління процесами.

- мобільний застосунок – для взаємодії користувача з системою.

Така структура забезпечує інтеграцію всіх компонентів, що дозволяє створити комплексну та ефективну систему для підтримки людей з вадами зору.

РОЗРОБКА СИСТЕМИ РОЗПІЗНАВАННЯ ОБ'ЄКТІВ

Далі розглядається процес розробки системи розпізнавання об'єктів для Android-застосунку з використанням камери (модуль ESP32-CAM) та алгоритму YOLOv3. Для реалізації проекту були застосовані такі програмні інструменти: інтегроване середовище розробки PyCharm, бібліотека OpenCV, алгоритм YOLOv3 та фреймворк Kivy.

PyCharm є популярним середовищем розробки для мови програмування Python, що було обрано для написання коду Android-застосунку з використанням фреймворку Kivy. Python є однією з основних мов програмування в галузі машинного навчання та

комп'ютерного зору, що забезпечує доступ до потужної екосистеми бібліотек, що значно спрощує розробку складних алгоритмів і додатків. Використання Kivy, відкритого фреймворку для створення багатоплатформених графічних інтерфейсів, дозволяє створювати застосунки, що можуть працювати на різних пристроях, включаючи Android.

YOLOv3 (You Only Look Once v3) є потужним алгоритмом для розпізнавання об'єктів у реальному часі, який базується на глибоких нейронних мережах. На відміну від інших методів, YOLOv3 здійснює повний аналіз зображення, розбиваючи його на сітку та прогнозує об'єкти і їхні класи для кожної клітинки. Цей підхід забезпечує високу швидкість і ефективність при вирішенні завдань розпізнавання об'єктів.

Модуль ESP32-CAM використовується для захоплення зображень і передачі їх на сервер або комп'ютер, де виконується алгоритм YOLOv3. Обробка відеопотоку та розпізнавання об'єктів на сервері дозволяє уникнути обмежень ресурсів мобільного пристрою, забезпечуючи високу продуктивність і точність розпізнавання.

Основні етапи процесу розпізнавання об'єктів у програмному коді включають такі дії: через фреймворк Kivy створюється інтерфейс користувача для Android-застосунку; завантажуються назви класів об'єктів з файлу `coco.names`; підключається конфігураційний файл та вагові коефіцієнти для моделі YOLOv3.

Кожні 1/60 секунди викликається функція `update_image()`, яка оновлює зображення, завантажує його з URL і перетворює в масив NumPy для подальшої обробки з використанням OpenCV. Зображення конвертується у формат blob, що подається на вхід нейронній мережі. Після проходження через мережу отримуються результати, що містять інформацію про виявлені об'єкти. Алгоритм non-maximum suppression застосовується для видалення зайвих прямокутників, які можуть перекривати один одного. На фінальному етапі на зображенні малюються рамки навколо об'єктів і додаються підписи до них.

Розроблений програмний застосунок забезпечує розпізнавання всіх об'єктів, що містяться у файлі `coco.names`. Цей файл містить класи об'єктів набору даних COCO, що використовуються для навчання моделі.

Для роботи алгоритму YOLOv3 необхідні файли `coco.names`, `yolov3.cfg` і `yolov3.weights`, які зберігаються в одній папці разом з головним файлом програми `main.py`. Ці файли є ключовими для коректної роботи нейронної мережі YOLOv3.

Завершений тестовий проєкт у середовищі PyCharm має чітку структуру, в якій всі необхідні файли та конфігурації організовано таким чином, що забезпечується успішна робота системи розпізнавання об'єктів.

Результатом функціонування системи розпізнавання об'єктів є відтворення відеопотоку з модуля ESP32-CAM в реальному часі. Після надходження відео з камери система аналізує кожен окремий кадр за допомогою алгоритму YOLOv3, який виконує обробку зображення, визначаючи об'єкти на основі завантаженої моделі. Для кожного виявленого об'єкта система накладає рамку і додає відповідну мітку, що позначає клас цього об'єкта.

З'єднання між мобільним пристроєм і модулем ESP32-CAM здійснюється через загальну Wi-Fi мережу. Завдяки мобільному застосунку, розробленому на базі фреймворку Kivy, результати роботи системи виявлення об'єктів відображаються на екрані телефону. Користувач має можливість переглядати відео з камери ESP32-CAM, де виявлені об'єкти підсвічуються рамками і позначаються відповідними класами.

Таким чином, розроблена система дозволяє користувачеві здійснювати моніторинг в реальному часі, переглядаючи відеопотік з камери на своєму мобільному пристрої, доповнений інформацією про виявлені об'єкти у вигляді рамок та міток (Рис. 7).

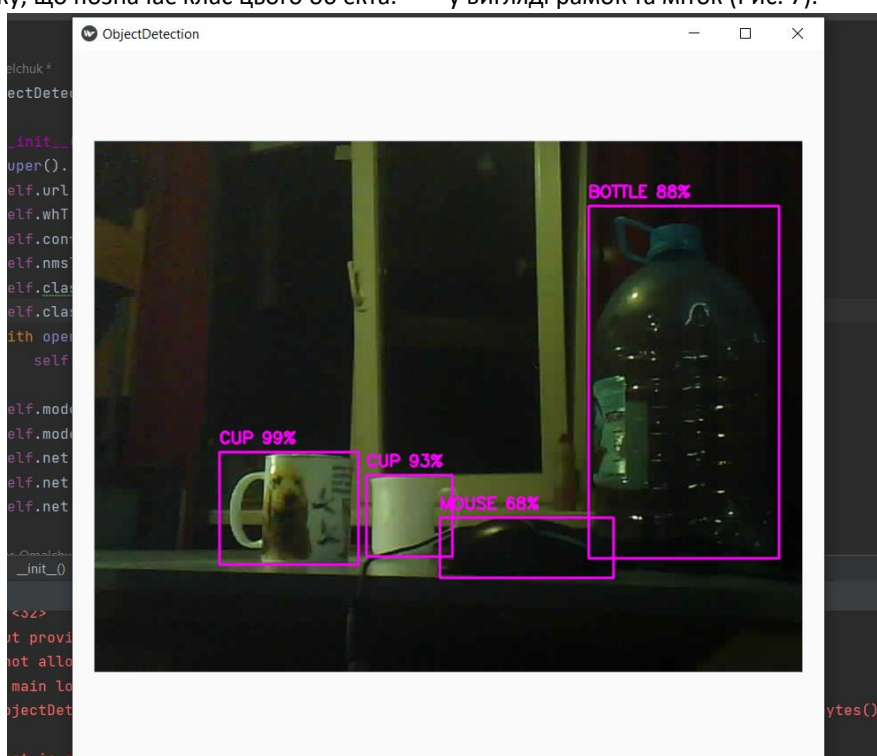


Рис. 7. Результат роботи системи розпізнавання об'єктів

ВИСНОВКИ

У статті було здійснено аналіз низки існуючих рішень для підтримки людей з вадами зору, таких як Aira, Be My Eyes, BlindSquare, Seeing AI, з метою визначення їх переваг, недоліків та можливостей для покращення. Оцінено ефективність цих рішень, зокрема їх здатність вирішувати основні завдання користувачів, як-то точне виявлення перешкод, навігація та надання інформації про навколишнє середовище.

Крім того, проведено глибокий аналіз потреб користувачів з вадами зору для розробки функціональних і нефункціональних вимог до майбутньої системи. Зокрема, були визначені основні виклики, з якими

стикаються ці користувачі, такі як необхідність точного виявлення перешкод, своєчасна інформація про навколишнє середовище, а також забезпечення простоти та інтуїтивної зрозумілості у використанні системи. На основі вивчених потреб було сформульовано вимоги до системи та розроблено її архітектуру, що включає опис основних підсистем, їх взаємодію та функціональні можливості.

Окрему увагу було приділено реалізації однієї з ключових підсистем – системи розпізнавання об'єктів. Ця система використовує алгоритм YOLOv3 для реального часу розпізнавання об'єктів і інтеграції з іншими компонентами, що забезпечує ефективну роботу та взаємодію з користувачем. Результати реалізації цієї

підсистеми підтверджують можливість інтеграції сучасних технологій розпізнавання об'єктів у систему підтримки людей з вадами зору.

Таким чином, виконана робота є основою для подальшого розвитку та вдосконалення інтелектуальних систем для людей з обмеженими можливостями, що включають в себе інноваційні методи обробки зображень та інтеграцію новітніх технологій для забезпечення доступності та підтримки.

ДОСТУПНІСТЬ ДАНИХ

Програмний код викладено для вільного доступу за посиланням: <https://github.com/Nazar-Omelchuk/ObjectDetectionAppForAndroid>.

ЛІТЕРАТУРА

- [1] Magnitude, temporal trends, and projections of the global prevalence of blindness and distance and near vision impairment: a systematic review and meta-analysis. [Онлайн]. URL: <http://surl.li/epnzmb>. Дата звернення: 28.11.2024.
- [2] We're Aira, a Visual Interpreting Service. [Онлайн]. URL: <https://www.aira.io/>. Дата звернення: 28.11.2024.
- [3] Be My Eyes, "Providing vision to blind and visually impaired people," [Онлайн]. URL: <http://surl.li/omxdxl>. Дата звернення: 28.11.2024.
- [4] A popular app is experimenting with AI. [Онлайн]. URL: <http://surl.li/qisnml>. Дата звернення: 28.11.2024.
- [5] What is BlindSquare? [Онлайн]. URL: <http://surl.li/jhkesl>. Дата звернення: 28.11.2024.
- [6] Seeing AI – Talking Camera for the Blind. [Онлайн]. URL: <https://www.seeingai.com/>. Дата звернення: 28.11.2024.
- [7] В. Р. Borkivskiy and V. M. Tesliuk, "Vykorystannia neiromerezhovykh zasobiv dlia rozpiznavannia ob'ektiv u mobil'nykh systemakh z obkhadom pereshkod," Scientific Bulletin of UNFU, vol. 33, no. 4, pp. 84-89, 2023. [Онлайн]. URL: <https://doi.org/10.36930/40330412>. Дата звернення: 28.11.2024.
- [8] M. Milon Islam, M. S. Sadi, and T. Bräunl, "Automated walking guide to enhance the mobility of visually impaired people," IEEE Trans. Med. Robotics Bionics, vol. 2, no. 3, pp. 485-496, Aug. 2020, doi: 10.1109/TMRB.2020.3011501.
- [9] M. Sadi, T. Bräunl, and M. Islam, "Smart Navigation for Visually Impaired People Using Artificial Intelligence," ITM Web of Conferences, vol. 46, p. 03053, 2022, doi: 10.1051/itmconf/20224603053. URL: <http://surl.li/tpjxmi>. Дата звернення: 28.11.2024.
- [10] M. R. N. Abdullah, M. M. Rahman, M. A. Fattah, and S. A. S. N. Islam, "An AIoT-Based Assistance System for Visually Impaired People," Electronics, vol. 12, no. 18, p. 3760, Sep. 2023, doi: 10.3390/electronics12183760.
- [11] S. Myronenko, S. A. Murakhovskiy, and O. A. Skidchenko, "Information technologies in the design of inertial sensors for orientation and navigation systems," Visnyk KPI, vol. 59, no. 1, pp. 63–70, 2020, doi: 10.20535/1970.59(1).2020.210019.
- [12] What is GPS: Types of Positioning Systems, How it Works, and Its Future. [Онлайн]. URL: <http://surl.li/jltkpe>. Дата звернення: 28.11.2024.
- [13] GPS – Wikipedia. [Онлайн]. URL: <http://surl.li/poogzq>. Дата звернення: 28.11.2024.

INTELLIGENT SYSTEM FOR ASSISTING PEOPLE WITH VISUAL IMPAIRMENTS

Nazar Omelchiuk, Iaroslav Dorogy

The article is dedicated to the analysis of modern technologies used to create intelligent systems for supporting people with visual impairments. The relevance of this topic is determined by the need to develop effective and accessible solutions that facilitate the daily lives of people with disabilities, particularly those with vision impairments. The aim of the article is to provide a comprehensive analysis of existing technologies, examine the advantages and disadvantages of available solutions, and identify potential ways to improve such systems. The article discusses leading solutions in this field, such as Aira, Be My Eyes, BlindSquare, and Seeing AI, which enable navigation and support for people with visual impairments through mobile devices, cameras, and other sensory systems.

The authors also explore the needs of the users of these systems, identifying the main challenges faced by people with limited vision, such as obstacle detection, providing timely information about the surrounding environment, and ensuring ease of use. Based on this analysis, the requirements for future support systems have been formulated, and the architecture of the intelligent system, which includes components for object recognition, obstacle detection, navigation, and voice assistance, has been developed.

Particular attention is given to the integration of object and obstacle recognition components, which are implemented using computer vision technologies such as OpenCV and neural networks. Additionally, the possibility of integrating voice assistants to provide convenient and accessible feedback is considered. The article also outlines the methods used for developing and testing such systems, including structural and comparative analysis of existing solutions, as well as experimental methods to assess the effectiveness of the software.

The research and development of intelligent systems for supporting people with visual impairments is an important step towards creating an accessible and barrier-free environment for people with disabilities,

opening up new opportunities to improve their quality of life.

Keywords: *intelligent support system, assistance for visually impaired individuals, web application, computer vision, object recognition, YOLOv3, voice assistance, navigation, adaptive technology.*

REFERENCES

- [1] Magnitude, temporal trends, and projections of the global prevalence of blindness and distance and near vision impairment: a systematic review and meta-analysis. [Online]. URL: <http://surl.li/epnzmb>. Accessed: Nov. 28, 2024.
- [2] We're Aira, a Visual Interpreting Service. [Online]. URL: <https://www.aira.io/>. Accessed: Nov. 28, 2024.
- [3] Be My Eyes, "Providing vision to blind and visually impaired people," [Online]. URL: <http://surl.li/omxdxl>. Accessed: Nov. 28, 2024.
- [4] A popular app is experimenting with AI. [Online]. URL: <http://surl.li/qisnml>. Accessed: Nov. 28, 2024.
- [5] What is BlindSquare? [Online]. URL: <http://surl.li/jhkesl>. Accessed: Nov. 28, 2024.
- [6] Seeing AI – Talking Camera for the Blind. [Online]. URL: <https://www.seeingai.com/>. Accessed: Nov. 28, 2024.
- [7] B. P. Borkivskyi and V. M. Tesliuk, "Vykorystannia neiromerezhovyykh zasobiv dlia rozpiznavannia ob'ektiv u mobil'nykh systemakh z obkodom pereshkod," Scientific Bulletin of UNFU, vol. 33, no. 4, pp. 84-89, 2023. [Online]. URL: <https://doi.org/10.36930/40330412>. Accessed: Nov. 28, 2024.
- [8] M. Milon Islam, M. S. Sadi, and T. Bräunl, "Automated walking guide to enhance the mobility of visually impaired people," IEEE Trans. Med. Robotics Bionics, vol. 2, no. 3, pp. 485-496, Aug. 2020, doi: 10.1109/TMRB.2020.3011501.
- [9] M. Sadi, T. Bräunl, and M. Islam, "Smart Navigation for Visually Impaired People Using Artificial Intelligence," ITM Web of Conferences, vol. 46, p. 03053, 2022, doi: 10.1051/itmconf/20224603053. URL: <http://surl.li/tpjxmi>. Accessed: Nov. 28, 2024.
- [10] M. R. N. Abdullah, M. M. Rahman, M. A. Fattah, and S. A. S. N. Islam, "An AIoT-Based Assistance System for Visually Impaired People," Electronics, vol. 12, no. 18, p. 3760, Sep. 2023, doi: 10.3390/electronics12183760.
- [11] S. Myronenko, S. A. Murakhovskiy, and O. A. Skidchenko, "Information technologies in the design of inertial sensors for orientation and navigation systems," Visnyk KPI, vol. 59, no. 1, pp. 63-70, 2020, doi: 10.20535/1970.59(1).2020.210019.
- [12] What is GPS: Types of Positioning Systems, How it Works, and Its Future. [Online]. URL: <http://surl.li/jltkpe>. Accessed: Nov. 28, 2024.
- [13] GPS – Wikipedia. [Online]. URL: <http://surl.li/poogzq>. Accessed: Nov. 28, 2024.