

УДК 004.8

Дидо Р.А., Собко О.В.

Хмельницький національний університет

МЕТОД ПОБУДОВИ АУДІОПОТОКУ ДОПОВНЕНОЇ РЕАЛЬНОСТІ ДЛЯ ПОКРАЩЕННЯ ОРІЄНТАЦІЇ ТА БЕЗПЕКИ ЛЮДЕЙ ІЗ ПРОБЛЕМАМИ ЗОРУ

Розглянуто метод побудови аудіопотоку доповненої реальності для покращення орієнтації та безпеки людей із проблемами зору. Метод ґрунтується на використанні моделі YOLOv8S для виявлення об'єктів у відеопотоці та згорткової нейронної мережі для класифікації названих осіб, що дає змогу забезпечити пріоритизоване озвучування небезпечних об'єктів та персоналізоване інформування користувача в реальному часі. Розроблений програмний прототип продемонстрував високі показники точності під час експериментального дослідження, підтвердивши ефективність методу як інструменту підвищення автономності та безпеки людей із зоровими порушеннями.

The method for constructing an augmented reality audio stream to improve orientation and safety for people with vision problems is considered. The method is based on the use of the YOLOv8S model for detecting objects in a video stream and a convolutional neural network for classifying named individuals, which allows for prioritized voicing of dangerous objects and personalized user information in real time. The developed software prototype demonstrated high accuracy rates during experimental research, confirming the effectiveness of the method as a tool for increasing the autonomy and safety of people with visual impairments.

У дослідженні представлено науково обґрунтований метод побудови аудіопотоку доповненої реальності для покращення орієнтації та безпеки людей із проблемами зору шляхом аналізу відеопотоку в реальному часі та озвучування пріоритетної інформації про об'єкти довкілля й ідентифікованих осіб. Зростання глобальної кількості людей із порушеннями зору, складність самостійної навігації, недостатня адаптованість інфраструктури [1] та обмежені можливості існуючих цифрових асистивних технологій формують потребу у створенні інтегрованих систем, здатних синтезувати зорову інформацію у зручний аудіоформат [2, 3]. Такий підхід має особливе значення в умовах нестабільного чи небезпечного середовища, де швидкість та точність повідомлень впливають на безпеку користувача [4].

Зростання інтересу до технологій доповненої реальності у поєднанні з інтелектуальними системами аналізу даних відкриває нові можливості для підтримки людей із порушеннями зору [5]. Одним із перспективних напрямів є формування аудіопотоку доповненої реальності, у якому вербальні описи, попередження та контекстні підказки синтезуються на основі багаторівневої

обробки інформації про довкілля [6]. Розвиток технологій комп'ютерного зору відкриває нові можливості для створення систем доповненої реальності, орієнтованих на підтримку людей із порушеннями зору. Актуальність цієї тематики зумовлена потребою забезпечити безпечну та інформативну взаємодію таких користувачів з оточенням у режимі реального часу [7]. Комп'ютерний зір дозволяє автоматично аналізувати просторову структуру сцени, визначати ключові об'єкти, оцінювати їхню відстань, тип, поведінкові характеристики та потенційну небезпеку [8]. Для людей з порушеннями зору ці можливості є критично важливими, оскільки формують основу для побудови аудіопотоку доповненої реальності, який компенсує брак візуальної інформації.

Сучасні підходи, засновані на глибинному навчанні, забезпечують високу точність розпізнавання об'єктів [9], виявлення перешкод [10], семантичної сегментації [11] та відстеження рухомих елементів сцени [12]. Це дозволяє не лише описувати статичне оточення, а й розпізнавати динамічні ситуації, що особливо важливо для навігації у міському середовищі, громадських будівлях чи транспортних вузлах [13]. Отримані комп'ютерним зором дані можуть бути конвертовані у структуровані текстові описи, які далі трансформуються в аудіопідказки – від коротких попереджень до детальних інструкцій, залежно від рівня складності ситуації та індивідуальних потреб користувача.

Перспективи розвитку таких систем пов'язані з інтеграцією мультимодальних моделей [14], здатних одночасно обробляти відеопотоки [16], просторові карти оточення та контекстні дані [15], що дозволить генерувати ще точніші та релевантні мовні повідомлення [17]. Такі технології здатні суттєво підвищити рівень мобільності та безпеки людей із порушеннями зору, створюючи інтелектуальний аудіонавігаційний простір, що природно інтегрується у їхню повсякденну взаємодію із навколишнім світом.

Наукова новизна дослідження полягає у поєднанні пріоритизованої ідентифікації об'єктів та персоніфікованої класифікації осіб із подальшим перетворенням цієї інформації на аудіопотік доповненої реальності. Такий підхід забезпечує адаптивну взаємодію користувача із середовищем, розширює можливості асистивних технологій, а також створює умови для більшої автономності осіб із порушеннями зору. Розроблений метод сприяє реалізації Цілей сталого розвитку ООН №3, №10 та №11 [18], оскільки підтримує збереження здоров'я, зменшує соціальну нерівність і сприяє формуванню інклюзивних міських середовищ.

Практична значущість роботи полягає у створенні програмного прототипу, який доводить застосовність методу в реальному режимі. Такий підхід може бути інтегрований в мобільні застосунки, системи навігації та інклюзивні інформаційні сервіси.

Запропонований метод базується на використанні моделі YOLOv8S для виявлення об'єктів у відеопотоці та згорткової нейронної мережі для класифікації названих осіб. Його структурна організація охоплює попередню обробку

відеоданих, виявлення та фільтрацію об'єктів за порогом довіри, персоніфіковану ідентифікацію та формування аудіопотоку з урахуванням пріоритетності потенційно небезпечних об'єктів. Ключовою особливістю є встановлення динамічної ієрархії класів, що визначає порядок озвучування залежно від рівня загрози, яку конкретні об'єкти можуть створювати для користувача. Це дозволяє оперативно привертати увагу до факторів ризику, підтримуючи безпечну траєкторію руху. Додатково, можливість класифікації названих осіб забезпечує персоніфіковану взаємодію з оточенням, що є важливим як у повсякденних ситуаціях, так і в контексті соціальної інтеграції.

Експериментальні дослідження виконано на основі створеного програмного прототипу (рисунок 1), що включає підсистеми попередньої обробки відеопотоку, виявлення об'єктів, класифікації названих осіб, навчання моделей і користувацького інтерфейсу [19].



Рисунок 1 – Програмний прототип запропонованого методу

Отримані значення макропоказників для моделі YOLOv8S підтвердили її ефективність, зокрема Precision 0.77, Recall 0.62, mAP@0.5 0.73 та mAP@0.5:0.95

0.5604. Аналіз мікропоказників для об'єктів із підвищеним рівнем небезпеки засвідчив переважно високу точність моделі, хоча для окремих класів виявлено доцільність подальшого донавчання для підвищення стабільності. Водночас CNN-модель для класифікації названих осіб продемонструвала практично безпомилкову роботу, що підтверджують макропоказники Accuracy 0.97, Precision 0.98, Recall 0.97 та F₁-score 0.98. Високі результати для класів «Person 1» та «Person 2» свідчать про надійність моделі, тоді як оптимізація вибірки для класу «Other» може підвищити повноту розпізнавання.

Подальший розвиток дослідження передбачає впровадження новітніх нейромережових архітектур, зокрема YOLO12, розширення та балансування датасетів для високопріоритетних класів, створення мобільного застосунку з адаптивною конфігурацією та розроблення підходів до динамічного контекстного аналізу середовища, що дозволить формувати більш точні й ситуаційно релевантні аудіоповідомлення.

Отримані результати підтверджують ефективність запропонованого методу як перспективного рішення для систем підтримки безпеки та орієнтації осіб із порушеннями зору. Очікується, що запропонований метод створюватиме умови для безпечнішої інтеграції в соціальний простір, підвищуватиме якість життя користувачів і закладає основу для розвитку нових інтелектуальних асистивних технологій.

Перелік посилань

1. World Health Organization. Blindness and Visual Impairment. WHO. URL: <https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment>
2. BBC News Ukraina. Yak suchasni tekhnolohii dopomahaiut liudiam z invalidnistiu. URL: https://www.bbc.com/ukrainian/science/2016/02/160202_tech_disability_ko
3. Omelchuk N., Dorohyi I. Intelligent System of Support for People with Visual Impairments. Scientific Papers of Donetsk National Technical University. Series: Computer Engineering and Automation. 2024. Pp. 4–17.
4. Dembitska S., Sievert I. Digital Accessibility in Education: Challenges and Prospects. Pedagogy of Security. 2024. Vol. 9, No. 2. Pp. 57–63.
5. Дидо Р.А., Мазурець О.В. Метод ідентифікації особистості на основі розпізнавання обличчя в реальному часі для систем кібербезпеки. Інформаційна, функційна і кібербезпека СКІФіК2024. Матеріали IV Всеукраїнської науково-технічної конференції. 29-30 листопада 2024. Харків. 2024. с. 36-37.
6. Kok I.A., Kadynska V.D., Zalutska O.O., Mazurets O.V. Object-Oriented Intelligent System for Automated Control of Smoking by Video Data. Current scientific goals, approaches and challenges. Proceedings of IV International Scientific and Theoretical Conference. June 13, 2025. Dresden, Federal Republic of Germany. Pp. 156-164.
7. Hladun O., Zalutska O., Klimenko V., Mazurets O. Research on the effectiveness of classifying the remains of destroyed buildings using MobileNetV3 neural network architecture. Innovations in Science: From Theoretical Foundations to Practical Impact. Proceedings 1st International Scientific and Practical Conference. May 12-14, 2025. Antwerp, Belgium. Pp. 158-162.

8. Mushtyn O., Sobko O., Molchanova M., Mazurets O. Convolutional Neural Network Architecture for Image-Based Architectural Style Recognition. *Evolving Science: Theories, Discoveries and Practical Outcomes. Proceedings of 4th International Scientific and Practical Conference.* June 9-11, 2025. Zurich, Switzerland. Pp. 130-143.
9. Malaydakh V., Molchanova M., Shevchuk P., Mazurets O. Deep learning neural network architecture for determining sunflower growth stage from visual data. *Modern Scientific Research: Theoretical and Practical Aspects. Proceedings II International Scientific and Practical Conference.* May 26-28, 2025. Riga, Latvia. Pp. 143-148.
10. Didur V., Molchanova M., Mazurets O. Research on the effectiveness of neural network detection of plots with the destroyed buildings remains. *Modern technologies and science: problems, new and relevant developments. Proceedings XXI International Scientific and Practical Conference.* May 26, 2025. Zaragoza, Spain. Pp. 245-251.
11. Hladun O.V., Molchanova M.O., Zalutska O.O., Mazurets O.V. Effectiveness research of using ViT neural network architecture for classifying the destroyed buildings remains. *Achievements of Science and Applied Research. Proceedings of 2nd International Scientific and Theoretical Conference.* May 19-21, 2025. Dublin, Ireland. Pp. 96-100.
12. Kharysh I., Sobko O., Mazurets O. Designing CNN Neural Network Model for Detecting Fractures of Lower Extremities by X-ray Images. *The Impact of Scientific Research on the Development of the Modern World. Proceedings of the XLIV International scientific and practical conference.* October 23-25, 2024. Dubrovnik, Croatia. 2024. Pp. 91-96.
13. Zharnovskiy O., Mazurets O., Sobko O. Approach to Identification of Artificial Intelligence-Generated People Images by Means of Machine Learning. *Key Aspects of the Development of Scientific Research in Modern Conditions. Proceedings of the XLV International scientific and practical conference.* October 30 – November 1, 2024. Constanta, Romania. 2024. Pp. 69-73.
14. Mazurets O. V., Klimenko V. I., Molchanova M. O., Sultanov A. V. Object-Oriented Intelligent System for Neural Network Detection of Sugar Crystallization Zones. *Global Science: Prospects and Innovations. Proceedings of the 10th International scientific and practical conference.* Cognum Publishing House. Liverpool, United Kingdom. 2024. Pp. 198-207.
15. Pokhytun A., Mazurets O., Molchanova M., Tyschenko O. Method for Neural Network Detecting Changed Images of People's Faces Using CNN. *New Horizons in Scientific Research: Challenges and Solutions. Proceedings of the 1st International scientific and practical conference.* October 21-23, 2024. Marseille, France. 2024. Pp. 35-40.
16. Похитун А.В., Мазурець О.В., Дидо Р.А., Молчанова М.О. Програмна архітектура для нейромережевого виявлення модифікованих фотографій облич людей. *Науковий журнал «Вісник Хмельницького національного університету» серія: Технічні науки.* Хмельницький, 2025. №3. Т.2. С. 493-500.
17. Dydo R., Sobko O., Klimenko V., Mazurets O. Datalogic structure for intelligent system for areas localization in photos with the remains of buildings using neural network. *Modern Scientific Research: Theoretical and Practical Aspects. Proceedings II International Scientific and Practical Conference.* May 26-28, 2025. Riga, Latvia. Pp. 123-127.
18. Gyimah P., Appiah K. O., Appiagyei K. Unraveling Contemporary Trends on United Nations Sustainable Development Goals: A New Global Bibliometric and Literature Review Analysis. *Sustainable Development.* 2025. Vol. 33, No. 2. Pp. 2579–2598.
19. Мазурець О.В., Собко О.В., Дидо Р.А., Молчанова М.О., Тищенко О.О. Просктування структури бази даних для інтелектуальної системи нейромережевого виявлення та локалізації на фотозображеннях ділянок із залишками зруйнованих будівель. *Науковий журнал «Наука і техніка сьогодні».* Київ, 2025. №4 (45). С. 1242-1258.