

УДК 004.93:004.89

АНАЛІЗ ОСОБЛИВОСТЕЙ СУЧАСНОГО МОБІЛЬНОГО ЗАСТОСУНКУ SEEING AI ДЛЯ РОЗПІЗНАВАННЯ ОБ'ЄКТІВ НА ЗОБРАЖЕННЯХ

Тютюнник Я.О.

e-mail: yaroslav.tiutiunnyk@nure.ua

Науковий керівник – к.т.н., доц. Творошенко І.С.

Харківський національний університет радіоелектроніки, каф. ІНФ
м. Харків, Україна

This work is devoted to the analysis of the features of the Seeing AI mobile application, designed for object recognition in images. The relevance of assistive technologies for visually impaired users is described, highlighting the importance of artificial intelligence in enhancing accessibility. The study examines the key functionalities of Seeing AI, including text recognition, object identification, and facial recognition. The challenges associated with accuracy, real-time processing, and user experience are discussed. Some recommendations are provided for improving the application's efficiency and expanding its capabilities in the future.

Сучасні мобільні застосунки, що використовують технології штучного інтелекту [1–5], відіграють важливу роль у забезпеченні доступності для людей із порушеннями зору. Один із таких застосунків – Seeing AI, розроблений компанією Microsoft, який дозволяє розпізнавати об'єкти, текст та обличчя на зображеннях у реальному часі.

Seeing AI – це безкоштовна програма, яка описує голосом навколишній світ. Це дослідницький проєкт, розроблений для незрячих людей і людей зі слабким зором за їх участю, який використовує можливості штучного інтелекту для відкриття цим людям візуального світу шляхом опису людей, тексту та об'єктів, що знаходяться поблизу.

Seeing AI містить засоби, які допомагають справлятися з різними повсякденними завданнями:

– озвучування короткого тексту. Застосунок вимовляє текст, коли він з'являється перед камерою;

– допомога в скануванні документів. Інструментальний засіб надає аудіоінструкції для зйомки друкованої сторінки та розпізнає текст з урахуванням вихідного форматування. Застосунок надає можливість поставити запитання про вміст, щоб допомогти легко знайти необхідну інформацію;

– сканування продуктів. Мобільний застосунок здійснює сканування штрих-кодів або QR-кодів з підтримкою спеціальних можливостей за допомогою звукових сигналів в якості покажчиків і дозволяє прослуховувати назви продуктів та відомості з упаковки, якщо вони доступні;

– аналіз навколишніх сцен. Застосунок надає можливість прослухати загальний опис зафіксованої сцени. Щоб прослухати докладніший опис, необхідно натиснути кнопку «Додаткові відомості». У користувача також є можливість вивчити фотографію, переміщуючи палець по екрану, щоб почути інформацію про розташування різних об'єктів;

– аналіз облич людей. Інструментальний засіб зберігає обличчя людей, допомагаючи користувачу впізнавати їх, а також приблизно визначає їх вік, стать і вираз обличчя;

– сканування валюти. Мобільний застосунок може розпізнавати банкноти при наведенні камери на купюри;

– розпізнавання кольорів. Застосунок вміє визначати кольори при наведенні на них камери користувачем;

– сканування рукописного тексту. Інструментальний засіб читає та відтворює голосом рукописний текст, наприклад, у вітальних листівках (доступно кількома мовами);

– аналіз яскравості навколишнього середовища. Застосунок видає звуковий сигнал, який відповідає яскравості навколишнього середовища;

– аналіз зображень в інших програмах. Інструментальний засіб може розпізнавати та аналізувати фотографії з інших застосунків, наприклад, з «Галереї» чи соціальних мереж.

Переваги застосунку Seeing AI:

– озвучування тексту в реальному часі. Seeing AI дозволяє швидко розпізнавати та озвучувати текст із документів, етикеток, плакатів, рукописних записів тощо. Це дає змогу людям із порушеннями зору легко отримувати інформацію з друкованих матеріалів;

– розпізнавання об'єктів та опис сцен. Застосунок може ідентифікувати предмети та озвучувати їхній опис. Наприклад, він може сказати: «Перед вами стілець» або «На столі лежить книга». Це значно спрощує орієнтацію у просторі;

– розпізнавання банкнот. Seeing AI допомагає розпізнавати купюри різних валют, що є особливо корисним у повсякденному житті, коли потрібно розрахуватися готівкою;

– ідентифікація осіб. Застосунок може розпізнавати обличчя знайомих людей, зберігати їхні імена та навіть повідомляти про емоційний стан співрозмовника, що допомагає у соціальній взаємодії;

– сканування штрихкодів. Seeing AI може сканувати штрихкоди на товарах та озвучувати їхню назву й опис, що дозволяє людям із порушеннями зору самостійно здійснювати покупки;

– робота без підключення до Інтернету. Деякі функції, такі як розпізнавання короткого тексту, працюють автономно, що дуже зручно в умовах відсутності стабільного інтернет-з'єднання;

– безкоштовне використання. Seeing AI доступний безкоштовно для користувачів iOS, що робить його доступним для широкої аудиторії.

Недоліки застосунку Seeing AI:

– обмежена доступність на Android. Наразі Seeing AI офіційно доступний лише для пристроїв на iOS, що є суттєвим недоліком для користувачів Android;

– не завжди точне розпізнавання. Хоча штучний інтелект працює досить добре, він може помилятися при ідентифікації об'єктів, людей або сцен, особливо в умовах поганого освітлення;

– проблеми з рукописним текстом. Seeing AI може розпізнавати рукописний текст, але не завжди точно, особливо якщо почерк нерозбірливий або нестандартний;

– обмеження у розпізнаванні мов. Деякі функції, такі як розпізнавання рукописного тексту або опис сцен, можуть працювати менш точно для рідкісних мов або діалектів;

– високе енергоспоживання. Оскільки застосунок активно використовує камеру та штучний інтелект, він може швидко розряджати акумулятор смартфона;

– питання конфіденційності. Seeing AI передає деякі дані на сервери Microsoft для обробки, що може викликати побоювання щодо збереження особистої інформації.

Список використаних джерел:

1. Yakovleva, O., Matúšová, S., Tvoroshenko, I., and Isaiev, Y. (2024). Visitor counting based on video stream analysis from surveillance cameras to solve various business problems. *Verejná správa a regionálny rozvoj ekonómia, manažment a marketing*, XX (1), 67-87.

2. Gorokhovatskyi V., and Tvoroshenko I. (2024) An effective method for transforming an image description into a compact vector for classification. *Information Technology and Implementation (Satellite): Conference Proceedings, November 21, 2024, Kyiv, Ukraine / Ministry of Education and Science of Ukraine, Taras Shevchenko National University of Kyiv and [etc]; Vitaliy Snytyuk (Editor). – Kyiv: Publishing House «Caravela», pp. 25-28.*

3. Daradkeh Y.I., Gorokhovatskyi V., Tvoroshenko I., and Zeghid M. (2024) Improving the effectiveness of image classification structural methods by compressing the description according to the information content criterion, *Computers, Materials & Continua*, vol. 80, no. 2, pp. 3085-3106.

4. Gorokhovatskyi V., Tvoroshenko I., and Yakovleva O. (2024) Transforming image descriptions as a set of descriptors to construct classification features, *Indonesian Journal of Electrical Engineering and Computer Science*, 33 (1), pp. 113-125.

5. Gorokhovatskyi V., Tvoroshenko I., Yakovleva O., Hudáková M., and Gorokhovatskyi O. (2024) Application a committee of Kohonen neural networks to training of image classifier based on description of descriptors set, *IEEE Access*, vol. 12, pp. 73376-73385.