UDC 004.93 **DOI: https://doi.org/10.18664/ikszt.v30i3.351322**

BARKOVSKA O.Y.
Candidate of Technical Sciences, Associate Professor at the Department of Electronic Computers, Kharkiv National University of Radio Electronics
NECHTAYLO O.V.
Master's student of the Department of Electronic Computers, Kharkiv National University of Radio Electronics
SERDECHNYI V.S.
PhD student of the Department of Electronic Computers, Kharkiv National University of Radio Electronics
BATURIN O.O.
Master's student of the Department of Electronic Computers, Kharkiv National University of Radio Electronics

Software Application for Assisting Visually Impaired Persons at Public Transport Stops

Abstract. *The paper addresses the urgent problem of social adaptation of visually impaired individuals in Ukraine, where the number of people with visual disabilities has significantly increased due to the consequences of military actions. The research presents the development of an intelligent software application designed to assist visually impaired persons in independent orientation at public transport stops. The system integrates deep learning, computer vision, and text-to-speech synthesis to automatically detect public transport vehicles (bus, trolleybus, tram) and recognize their route numbers in real time. The YOLOv8 model was employed for object detection, Tesseract OCR for route number recognition, and pyttsx3 for offline speech synthesis. Data preprocessing included dataset annotation, augmentation with simulated weather conditions, and the use of OpenCV-based filters to enhance OCR accuracy. Testing under both real and simulated conditions (fog, rain, snow, blur, low lighting) demonstrated consistently high detection accuracy (100%) and acceptable classification performance, though recognition robustness decreased under low-visibility scenarios. The results confirm the practical value of the proposed approach, while further improvements will focus on expanding real-world datasets, enhancing preprocessing methods, and integrating stronger deep learning models. The system holds promise as a foundation for wearable assistive technologies aimed at improving inclusivity and mobility for visually impaired users in urban environments.*

Keywords: *vision impairment; computer vision; YOLOv8; OCR; Tesseract; synthesis; object detection; assistive technologies; deep learning*

Introduction

The relevance of the problem of rehabilitation and social adaptation of people with visual impairments in Ukraine has significantly increased due to the consequences of military actions, which lead to damage to the organs of vision in both the military and the civilian population, requiring constant attention to technologies that help people adapt and integrate into society [1, 2]. According to the results of studies conducted in recent years in Ukraine, there were 19,000 citizens with visual impairments in 2023.

In 2021, this number was 17,000, which indicates an increase in the number of people with visual impairments [3, 4].

Large cities usually have an extensive and extensive system of urban roads for motor vehicles (intersections, bridges, tunnels, highways). There are various solutions for crossing all these road elements, such as traffic lights, underground/overground crossings, and pedestrian crossings. It should be noted that most of these measures are focused on visual identification of the road crossing location.

One of the busiest streets in Kharkiv, Prospekt Nauki, has six intersections per kilometre, only three of which are regulated but do not have audible crossing signals (figure 1).

Fig. 1. Number of intersections per 1 km on Nauky Avenue

This shows that the conditions of the modern urban environment remain insufficiently inclusive, as state resources are mainly directed towards the needs of the defence sector. The issue of orientation at public transport stops, which is a key element of infrastructure for visually impaired people who are unable to use private transport, is particularly complex.

Traditional means of assistance, such as guide dogs [5], do not provide complete autonomy, as they are unable to provide a person with information about the type of vehicle and route number. One promising approach to solving this problem is the use of computer vision technologies in combination with deep learning algorithms and speech synthesis systems. Such approaches allow for automatic recognition of vehicles and route numbers in real time and broadcast the results in the form of voice prompts. Some technological tools have already been created to help the visually impaired, but they have their drawbacks, which are described in the next section of this paper.

Therefore, the development of a proprietary software application using these technologies can significantly increase the level of inclusion in urban spaces, promote the principle of accessibility, and provide visually impaired people with more comfortable and safer conditions for travelling by public transport.

Analysis of recent studies and publications

Modern approaches to solving the problem of object detection in urban environments show a tendency to move from multi-step cascading architectures to single-step fully convolutional models. Traditional architectures, such as R-CNN, Fast R-CNN, and Faster R-CNN, have achieved a significant breakthrough in accuracy by integrating feature extraction and region generation mechanisms through regional proposals (RPN) [6]. However, their use in real-time systems, especially on resource-constrained embedded devices, remains problematic due to their significant computational complexity.

In the context of mobile or wearable devices, single-stage detectors such as SSD and YOLO family models are preferred. In particular, YOLOv8 provides high throughput and real-time performance by directly predicting classes and window coordinates in parallel [7]. This property makes it suitable for use in wearable or autonomous systems to support people with visual impairments.

In the sub-task of recognising public transport route numbers, two paradigms remain relevant: classical optical character recognition (OCR) methods, such as Tesseract, and deep neural networks focused on scene text recognition [8–9]. The Tesseract system, described by Smith at the ICDAR conference (2007), involves step-by-step segmentation, binarisation, character classification and labelling language. This makes it effective for processing clearly structured signs and labels with a fixed font.

In contrast, modern approaches based on CRNN (Convolutional Recurrent Neural Network) in combination with the CTC (Connectionist Temporal Classification) mechanism allow direct sequential text recognition without prior character segmentation. This significantly improves accuracy in conditions of perspective distortion, motion, or low video stream quality. The CRNN architecture combines convolutional layers for feature extraction, recurrent layers (usually BiLSTM) for sequence modelling, and a transcription layer, allowing for flexible processing of variable text structures in a frame.

The choice between Tesseract and neural network OCR models in systems with limited computing budgets is often driven by a trade-off between accuracy and resource consumption. In stable conditions, Tesseract demonstrates low latency and ease of integration, while

CRNN provides higher resistance to noise and distortions but requires a more powerful computing platform.

Stable video stream transmission remains a key factor in system reliability. To achieve this, video data is pre-processed and buffered using the OpenCV library, which allows you to adjust the frame rate, perform size alignment, focus, and apply filters necessary to improve the quality of input images. In practice, this is implemented through the VideoCapture API (Python/C++), which is currently the standard for video integration in real-time computer vision systems.

The final stage is the vocalisation of the results. In accordance with the requirements for the autonomy of systems for visually impaired people, it is recommended to use offline speech synthesis engines, in particular pyttsx3, which supports work with system TTS engines. This approach minimises the delay in starting synthesis and does not require a stable internet connection, which is critical for outdoor use. Cloud-based TTS services, such as gTTS, although they provide higher voice quality, are inferior to offline options in terms of speed and reliability in field conditions [10-11].

The above shows that a number of technological tools have been developed to support people with visual impairments. For example, GPS-based mobile applications (Seeing Eye GPS, LazarilloApp) provide navigation in open spaces, but they have limited accuracy indoors and in the absence of a stable satellite signal. Other solutions are based on ultrasonic sensor technology that signals obstacles (e.g., SmartCane), but such devices often do not take context into account (e.g., moving objects) and require additional user training. Computer systems using computer vision (e.g., Microsoft Seeing AI) are also becoming more widespread, providing object and text recognition, but they require significant computing resources and a constant internet connection, which reduces their practicality.

Work [12] is a systematic review of modern technologies designed to support individuals with visual impairments when navigating both indoor and outdoor environments. The authors reviewed 191 publications from 2011 to 2020, classifying approaches, hardware, and software solutions. Particular attention is paid to electronic canes, robotic assistants, systems based on GPS, RFID, Kinect, ultrasonic and infrared sensors. The analysis showed that the development of traditional electronic aids is gradually shifting towards the use of IoT solutions, computer vision algorithms and multimodal technologies.

However, the results have several limitations. The study covers only six major scientific libraries, which may lead to the omission of important works. Only the decade 2011–2020 was considered, so earlier or more recent developments were not taken into account. Most of the systems analyzed were tested under controlled conditions, often with a small number of users, and without a sufficient assessment of energy consumption, operating time, or user acceptance. Insufficient attention has also been paid to algorithms for selecting the shortest route and integration into urban infrastructure.

Work [13] presents a system of "smart glasses" for people with visual impairments, combining image enhancement in low light, DETR-based object recognition, tactile feedback, and text-to-speech. Experiments on the Low-Light and ExDark datasets showed competitive results in low-light conditions. However, the system has limitations, including reduced accuracy for small objects, dependence on a client-server architecture, and incomplete testing in real-world conditions.

Article [14] presents the ARIANNA+ system, which is based on augmented reality and computer vision technologies and is designed to support the mobility of people with visual impairments. The system allows you to build virtual routes without the need for physical landmarks, and also provides interaction with the environment through object recognition and access to digital content. Experiments have demonstrated the feasibility of reliable navigation in both indoor and outdoor environments using smartphones, achieving up to 90% accuracy in identifying cultural heritage sites.

Among the limitations of the application, it is worth noting its dependence on the accuracy of ARKit algorithms and the performance of the smartphone's camera, which complicates its use in conditions of sudden lighting changes or rapid movements. Navigation accuracy decreases when the device is positioned incorrectly, and the object recognition process relies on a limited amount of data collected in a specific cultural context, which raises concerns about the scalability of the solution. In addition, the tests were conducted under controlled conditions, and there is a lack of extensive testing involving users with visual impairments, which limits the practical assessment of effectiveness.

Thesis [15] describes a "smart cane" system that integrates Raspberry Pi 4B, a camera, an ultrasonic sensor, and Arduino. The system combines visual recognition (Viola–Jones, TensorFlow Object Detection) with distance measurement to obstacles and audible warning via a buzzer. The results showed an accuracy of approximately 91%, positioning the device as a compact and inexpensive solution for everyday use by people with visual impairments.

The following limitations are inherent in the development. First, the algorithms (in particular, Viola–Jones) remain sensitive to noise and camera vibrations, which reduces accuracy for small objects and in low-quality images. Second, the system relies primarily on simple computer vision methods, which are inferior to modern neural models in terms of stability and scalability.

Third, the tests were conducted mostly in controlled scenarios, which calls into question the effectiveness of the device in real urban environments with dynamic obstacles.

Paper [16] presents a mobile assistance system for visually impaired people that combines an OAK-D sensor with stereo vision, edge AI accelerators (NCS2) and optimised models (OpenVINO, TensorFlow Lite) for object detection, recognition of road signs, traffic lights, crosswalks, and changes in surface elevation. The system is compact, has relatively low power consumption, and a convenient voice interface, and its testing has shown promise in real urban environments.

However, the work has certain limitations. The use of lightweight models, such as SSD-MobileNet, results in a trade-off between speed and accuracy, with an increased likelihood of false positives and false negatives. Segmentation models operate with significant delays (0.4–2 fps), which complicates their use in real-time and may confuse the distinction between the road and the pavement. An additional challenge is the inability to reliably determine height changes based solely on OAK-D due to its low depth resolution. Therefore, the authors resort to machine learning on RGB and depth images, acknowledging that traditional point cloud processing methods could yield more robust results.

Thus, the review results confirm the feasibility of the selected software configuration for vehicle detection, route number recognition, and real-time voice prompt generation tasks. The use of YOLOv8 as the base detection model, flexible selection of OCR kernels depending on the context, an OpenCV-based video pipeline, and offline TTS provide the necessary balance between performance, accuracy, and affordability for implementation in systems that assist visually impaired users.

Purpose and objectives of the work

The goal of the work is to develop a software application to help visually impaired people independently navigate public transport stops, which provides automatic recognition of the type of transport and route number, as well as the generation of voice prompts for the user.

The following types of input data are required to implement the system:

- images of vehicles taken from a personal device (frames in RGB format);
- labelled datasets containing the coordinates of objects (bounding boxes) for transport types (bus, trolleybus, tram) and route numbers;
- augmented data with simulated weather effects: fog, rain, night, twilight, etc.;
- audio signals broadcast through a speaker in response to recognition.

Therefore, to achieve the set goal, a number of tasks must be performed:

- collection and labelling of data for deep learning of the YOLO model for recognising different types of vehicles and their route numbers;
- augmenting the dataset by simulating low visibility conditions, creating artificial weather conditions (fog, rain, night, etc.) at different times of the day (daylight, dusk, darkness);
- research and implementation of machine learning methods for the task of automated vehicle recognition in real time;
- testing and adaptation of the system in virtual conditions.

Further development of the project involves expanding functionality to work with smart stops (IoT infrastructure), integrating with a mobile application to personalize settings, and utilizing artificial intelligence technologies to adapt the system to different users.

Scientific results obtained. Discussion of results

The study describes the software component of a software-hardware complex designed to support the mobility of people with visual impairments. The hardware part is implemented on the basis of a single-board computer Raspberry Pi 5, to which a Logitech Carl Zeiss Tessar 2.0/3.7 (2 MP) webcam and an external Logitech G Pro X audio card with headphones are connected, providing real-time message playback. Power is supplied by a 10,000 mAh portable battery, which guarantees several hours of autonomous operation. At the same time, it is the software algorithms — image pre-processing, vehicle detection, and route number recognition — that determine the effectiveness of the complex. The efficiency of the system at various stages of its operation is ensured by the YOLOv8 framework, Tesseract, and the pyttsx3 library.

At the first stage of the system's operation, the incoming video stream from the camera is received, as well as accompanying information about the external environment, in particular, the level of illumination and weather conditions. This data is transmitted to the environment analysis module, which, based on the detected conditions, initiates the application of appropriate image pre-processing filters (e.g., contrast correction, noise reduction, visibility improvement in rain or at night).

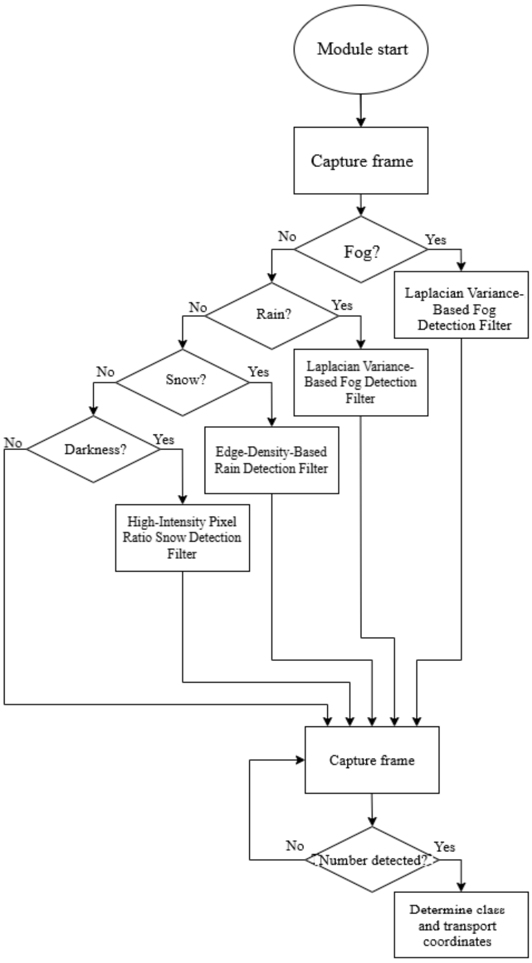


Fig. 2. Block diagram of the transport detection module

The environment analysis module performs a preliminary assessment of the shooting conditions based on the incoming video stream. To do this, it analyses the brightness characteristics of the frame (determining low light or high contrast), assesses the image noise level, and detects artefacts characteristic of rain or fog. Based on these characteristics, the system selects the appropriate pre-processing methods: brightness enhancement (gamma correction, CLAHE) in the dark, noise filtering (bilateral filtering) in the presence of interference. The use of defogging and deraining methods in case of deteriorating weather conditions are further steps necessary for implementation. Thus, the module ensures the operation of the entire system, improving the quality of input data before the object detection stage.

After that, the prepared image is sent to the computer vision module, implemented on the basis of the YOLOv8 architecture, where transport objects are detected. If the vehicle is successfully detected, the system proceeds to localise the area where the route number is located. The selected area is transferred to the optical character recognition (OCR) module, built on the Tesseract library.

If the route number is successfully read, a text message is generated, which is then sent to the speech synthesis module (pyttsx3) to voice the information received. If the OCR does not recognise the characters correctly, the system generates a voice message indicating only the type of vehicle (e.g. "bus", "trolleybus", etc.).

The general principle of the system is illustrated in detail in the block diagram (figure 2), which shows the sequence of video stream processing steps and the adaptability of the algorithm to changes in the external environment.

Analysis of the visual characteristics of route numbers has shown that they usually have a clearly defined rectangular shape and are placed on a contrasting background – black or white, sometimes with backlighting. This greatly facilitates their visual identification by humans and, at the same time, indicates the potential effectiveness of their automatic detection using neural network models of object detection.

Taking these features into account, it was decided to retrain the YOLOv8 model not only for the task of classifying vehicle types, but also for the task of localising the route number area on the vehicle body. This approach allows the area that is most likely to contain relevant text information to be selected directly from the frame.

The area selected using YOLOv8 is transferred to the optical character recognition (OCR) module, implemented on the basis of the Tesseract library. Given the peculiarities of its algorithms, which work more stably with black-and-white or grayscale images, contrast enhancement and color noise reduction methods are first applied to the localised area.

These operations are performed using the OpenCV library, which allows for the implementation of a number of filters: conversion to grayscale, adaptive binarisation, histogram equalisation, or contour enhancement. The result is an image optimised for further text recognition.

The general processing sequence and an example of image transformation before recognition are shown in figure 3.

The type of transport and its number are announced until the transport leaves the device's field of view, so that the user can understand that the transport is still in front of them.

To ensure that the system operates in real time, the voice announcement module is implemented using an asynchronous approach. The voice message is generated in the background, which does not interrupt the video frame processing cycle and minimises delays. This approach allows for continuous operation regardless of the duration of the voice announcement.

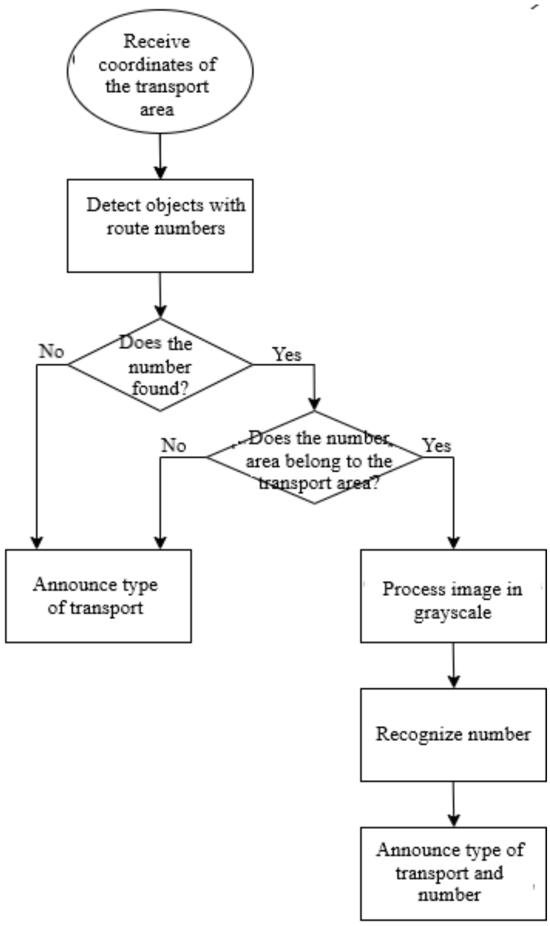


Fig. 3. Block diagram of the route number detection module

The message generation process is a key stage in the system's interaction with the user. A generalised description of the algorithm for voicing the detection results includes a sequence of actions from receiving the classified result to initiating speech synthesis, and also provides for the processing of exceptional situations, in particular cases where no route number is recognised.

In order to ensure the feasibility of the proposed approaches, a specialised dataset was prepared and labelled, including the following classes of objects:

- bus;
- trolleybus;
- tram;
- route number.

The annotation was performed in YOLO format, where the coordinates of the bounding box and the corresponding class label are specified for each image.

The initial sample consisted of 756 images, which were supplemented with augmentation to increase the model's resistance to changing external conditions. The following methods were used:

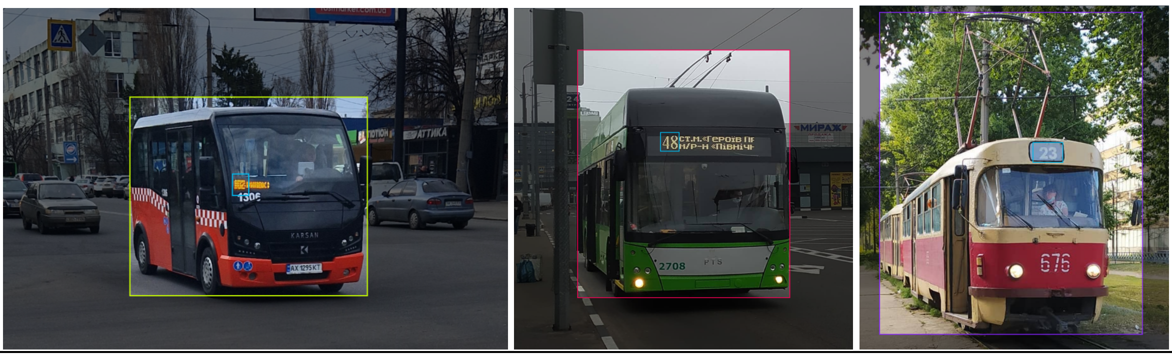
- change of lighting (brightness, contrast);
- adding noise (Gaussian, Salt & Pepper);
- simulation of rain;
- simulation of fog;
- simulation of night conditions (reduction of global brightness + local glare).

A specialised dataset was created for training the system, consisting of 756 source images of vehicles (buses, trolleybuses, trams) and route numbers. The clas-ses are distributed approximately evenly: 260 examples of buses, 248 trolleybuses and 248 trams. To improve the generalisation ability of the model, augmentation was applied, resulting in an almost twofold increase in the size of the dataset. The proportion of real data in the final set was about 60%, and the proportion of syn-thetically generated images (simulation of fog, rain, night conditions, etc.) was 40%.

A classic distribution scheme was used to form the samples: 70% of the data was used for training, 15% for validation, and 15% for testing. This division ensured the representativeness of the test sample even with a relatively small number of examples and prevented the "spillover" of augmented data from the training into the test subset.

This expansion of the training dataset is necessary to increase its representa-tiveness and ensure reliable model performance in a real-world environment.

It should be noted that route number markings require special attention — these areas are usually located on a contrasting rectangular background with back-lighting, and they have a characteristic shape that allows them to be effectively iden-tified using the YOLOv8 model, provided that they are correctly annotated during training (figure 4).



a) b) c)

Fig. 4. Example of frame marking for different vehicles: a) for a bus, b) for a trolleybus, c) for a tram

At the current stage, the system is focused on the route number formats represented in the dataset. Contextual constraints (number location zone, characteristic string length) are used to reduce the number of false positives. To avoid errors, a simple contextual filter is used: the location of the digits in the frame (for example, a characteristic area on the windscreen or on the transport display) and the format (1–3 digits, sometimes with a letter) are checked. Thus, the system does not announce numbers that fall outside these parameters (e.g., 4-digit on-board codes). In case of changes in the formats of route displays, the dataset is expanded and the model is retrained. Vehicles are classified based on visual characteristics specific to each class: the presence of poles and contact wires (trolleybuses), rail chassis (trams), and the absence of these elements (buses).

The Roboflow platform was used to label the images, which provides a complete set of functionality for annotation, dataset version management, and export in the required formats, in particular for YOLO models.

During the formation of the dataset, situations that often occur in real conditions were taken into account, such as the presence of several vehicles in one frame, particularly at public transport stops. Such scenes are often accompanied by partial overlap of vehicles, which complicates the detection task.

In such cases, it is worth performing an approximate annotation of their boundaries, taking into account the shape of the object and the context of the scene (figure 5).

This approach improves the generalisation of the model and ensures its stable operation in complex situations, particularly in urban environments.

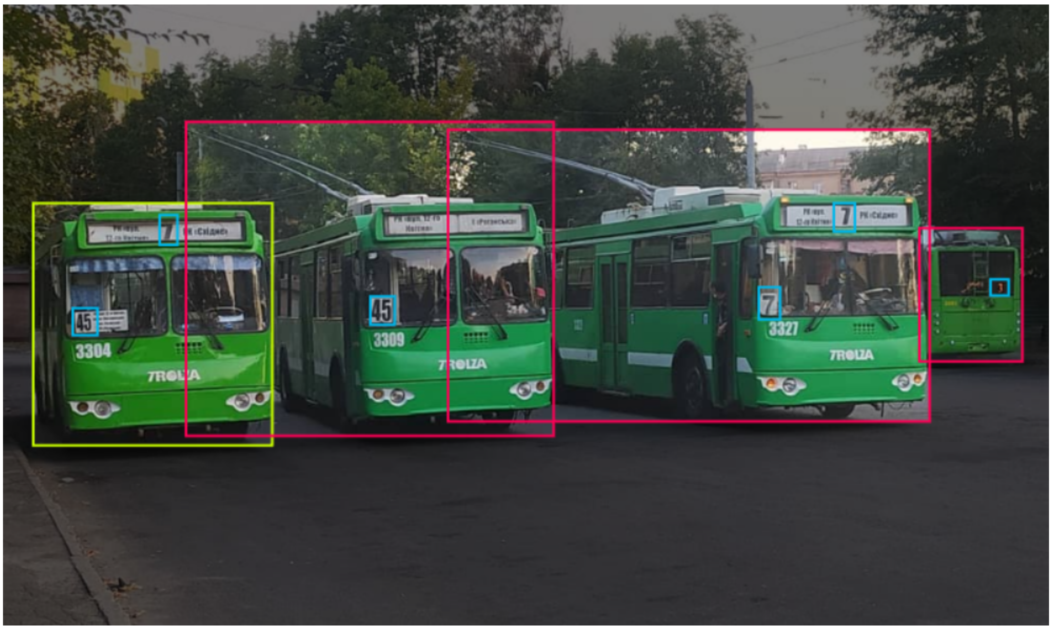


Fig. 5. Example of marking with several objects overlapping each other

To improve the model's generalisation ability, synthetic image augmentation was implemented. This allows the model to learn to recognise public transport objects in different weather conditions and in low light (figure 6).

a)

b)

c)

d)

Fig. 6. Conditions for dataset augmentation: a) rain; b) snow; c) darkness; d) blurring

To assess the functional suitability of the developed software and hardware complex, testing was carried out in conditions as close as possible to the real operating environment. The main goal of the experiment was to determine the accuracy of vehicle detection and route number recognition.

To simulate changing environmental conditions (fog, rain, night shooting), the OBS (Open Broadcaster Software) programme with the OBS Virtual Camera function was used, which allowed a video stream with simulated scenes, including buses, trolleybuses and trams, to be fed into the system.

The model was trained using the Adam optimiser, with a batch size of 16, 200 epochs, and an early stopping mechanism applied if no improvement was observed on the validation sample for 5 epochs. On the training part of the dataset, the average accuracy was 97%, which is consistent with the test results.

The use of virtual video streaming made it possible to test the system in scenarios that were difficult to implement physically during the summer. This approach allowed us to evaluate the adaptability of algorithms to complex visual conditions without losing accuracy.

Table 1. – *Table of test results in real and simulated conditions*

Experimental conditions	Number of frames, pcs.	Object recall (TPR), %	Precision of transport type, %	Accuracy of route number recognition, %
Simulation, daylight (100–10,000 lx)	30	10	95	89
Simulation, fog	30	98	95	84
Imitation, rain	30	100	100	85
Simulation, snow	30	100	100	90
Simulation, blurring (reduced sharpness)	30	96	95	80

Based on the results obtained, it can be concluded that the system demonstrates consistently high accuracy in detecting vehicles (mainly 100%), which indicates the reliability of the YOLOv8 algorithm in the task of object detection. Only in foggy conditions (98%)

and blurred images (96%) did the indicators decrease slightly, which is associated with a decrease in the contrast and sharpness of the input data (figure 7).

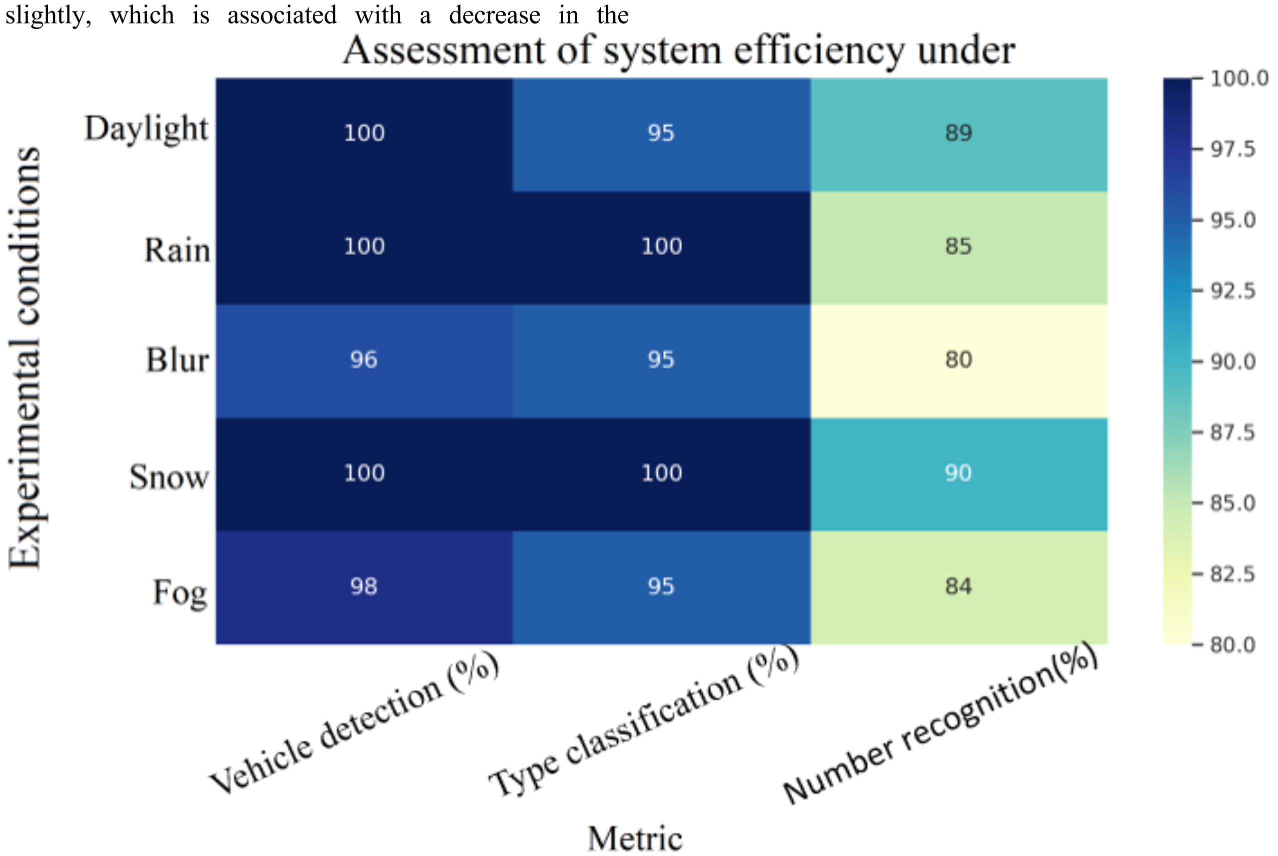


Fig. 7. Heat map of the results of experimental testing of the developed system

The accuracy of vehicle type classification ranged from 95 to 100%, with the best results observed in rainy and snowy conditions, where the contours of objects remained sufficiently clear. At the same time, in cases of fog and glare, the accuracy decreased to 95%, indicating the sensitivity of the model to blurred contours and loss of local features.

The higher accuracy in adverse conditions (100% in rain/snow) compared to daylight (95%) is explained by the uneven distribution of test examples: the number of examples in "difficult" conditions was smaller, which could artificially increase accuracy. In future work, we plan to balance the dataset and conduct a series of new experiments. It is also worth noting that some of the "adverse conditions" (rain, snow, fog) were generated using augmentation methods and reflect simulation scenarios. This could have affected the results and given higher accuracy compared to daylight conditions. The actual performance of the system in similar weather conditions may vary, which also requires further field experiments using real data.

The most problematic stage was the recognition of route numbers, where the accuracy ranged from 80% (blurred image) to 90% (snow, daylight). This is because the OCR system is highly dependent on the clarity of the characters and the quality of the frame pre-processing. Thus, further development of the system should be aimed at improving image pre-processing algorithms to increase the reliability of OCR in difficult visibility conditions.

Conclusions

During development, the software part of the intelligent module was created, combining deep learning, computer vision, and speech synthesis methods. To detect public transport, the YOLOv8n model was used, which demonstrated high accuracy in daytime conditions and acceptable speed on a single-board computer. Route number recognition was implemented using the Tesseract OCR module, which allowed the necessary text information to be obtained from video frames. The pyttsx3 library was used for rapid voice output of results, which ensures offline operation without an Internet connection and minimal delays.

Testing of the software component confirmed the effectiveness of the chosen solutions: the system consistently recognises vehicles and their numbers and correctly announces the results. At the same time, analysis showed that in low visibility conditions (twilight, darkness, blurred images), the accuracy of OCR and transport type classification decreases. This determines the directions for further work: improving image pre-processing algorithms, expanding the dataset with real examples from different weather conditions, and using more powerful deep learning models. As a result, the software component has proven its practical value and can be the basis for full integration into wearable systems to assist the visually impaired.

The tests showed that the software solutions provide acceptable performance in daylight conditions, but in twilight and with reduced lighting quality, the accuracy decreases, which necessitates further optimisation of the pre-processing methods. Thus, the results of the work confirm the practical significance of the created software as a key component of the software and hardware complex.

References

1. Berad, A. S. (2025). Emerging Technologies for Enhancing Vision in People with Blindness: Innovations, Applications, and Future Directions. *Central India Journal of Medical Research*, *4*(01), 18-20. https://doi.org/10.58999/cijmr.v4i01.218

2. Serdechnyi, V., Barkovska, O., Kovalenko, A., Martovytskyi, V., & Havrashenko, A. (2025). Knowledge systematization about the characteristics of existing technological means for assisting people with visual impairments. *Advanced Information Systems*, *9*(1), 13-23. https://doi.org/10.20998/2522-9052.2025.1.02.

3. Нечітайло О. В. Інтелектуальний модуль навігації для реабілітації осіб з вадами зору на зупинках громадського транспорту / О. В.Нечітайло, А. О. Гаврашенко // Сучасні напрями розвитку інформаційно-комунікаційних технологій та засобів управління : тези доп. 15-ї міжнар. наук.-техн. конф., 24-25 квітня 2025 р., Баку–Харків–Жиліна : [у 4 т.]. Т. 3 : секція 3,4 / Нац. ун-т оборони Азербайджанської Республіки [та ін.]. – Харків : Impress., 2025. – С. 123. https://doi.org/10.32620/ICT.25.t3

4. Tulashvili Y. Humanization of society through integration and inclusive education: Challenges and prospects for individuals with visual impairments in Ukraine //Forum for Education Studies. – 2024. – Т. 2. – №. 3. – С. 279-279. https://doi.org/10.59400/fes.v2i3.279

5. Khan, S., Nazir, S., & Khan, H. U. (2021). Analysis of navigation assistants for blind and visually impaired people: A systematic review. *IEEE access*, *9*, 26712-26734. https://doi.org/10.1109/ACCESS.2021.3052415.

6. Hmidani, O., & Alaoui, E. I. (2022, December). A comprehensive survey of the R-CNN family for object detection. In *2022 5th International Conference on Advanced Communication Technologies and Networking (CommNet)* (pp. 1-6). IEEE. https://doi.org/10.1109/CommNet56067.2022.9993862.

7. Ali, M. L., & Zhang, Z. (2024). The YOLO framework: A comprehensive review of evolution, applications, and benchmarks in object detection. *Computers*, *13*(12), 336. https://doi.org/10.3390/computers13120336.

8. Avyodri, R., Lukas, S., & Tjahyadi, H. (2022, September). Optical character recognition (ocr) for text recognition and its post-processing method: A literature review. In *2022 1st International Conference on Technology Innovation and Its Applications (ICTIIA)* (pp. 1-6). IEEE. https://doi.org/10.1109/ICTIIA54654.2022.9935961.

9. Thorat, C., Bhat, A., Sawant, P., Bartakke, I., & Shirsath, S. (2022). A detailed review on text extraction using optical character recognition. *ICT Analysis and Applications*, 719-728. https://doi.org/10.1109/I2CT61223.2024.10544007.

10. Kumar, Y., Koul, A., & Singh, C. (2023). A deep learning approaches in text-to-speech system: a systematic review and recent research perspective. *Multimedia Tools and Applications*, *82*(10), 15171-15197. https://doi.org/10.1007/s11042-022-13943-4.

11. Praveen Borra. ‘A Survey of Google Cloud Platform (GCP): Features, Services, and Applications’. International Journal of Advanced Research in Science, Communication and Technology, Jun. 2024, pp. 191–99. DOI.org (Crossref), https://doi.org/10.48175/IJARSCT-18922.

12. S. Khan, S. Nazir and H. U. Khan, "Analysis of Navigation Assistants for Blind and Visually Impaired People: A Systematic Review," in IEEE Access, vol. 9, pp. 26712-26734, 2021, doi: 10.1109/ACCESS.2021.3052415

13. Mukhiddinov, M.; Cho, J. Smart Glass System Using Deep Learning for the Blind and Visually Impaired. Electronics 2021, 10, 2756. https://doi.org/10.3390/electronics10222756

14. Lo Valvo, A.; Croce, D.; Garlisi, D.; Giuliano, F.; Giarré, L.; Tinnirello, I. A Navigation and Augmented Reality System for Visually Impaired People. Sensors 2021, 21, 3061. https://doi.org/10.3390/s21093061

15. U. Masud, T. Saeed, H. M. Malaikah, F. U. Islam and G. Abbas, "Smart Assistive System for Visually Impaired People Obstruction Avoidance Through Object Detection and Classification," in IEEE Access, vol. 10, pp. 13428-13441, 2022, https://doi.org/10.1109/ACCESS.2022.3146320.

16. J. K. Mahendran, D. T. Barry, A. K. Nivedha and S. M. Bhandarkar, "Computer Vision-based Assistance System for the Visually Impaired Using Mobile Edge Artificial Intelligence," 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Nashville, TN, USA, 2021, pp. 2418-2427, https://doi.org/10.1109/CVPRW53098.2021.00274.

Програмний модуль розпізнавання маршрутних номерів у реальному часі для вбудованих систем допомоги слабозорим

Abstract.. *У роботі розглядається актуальна проблема соціальної адаптації осіб із порушенням зору в Україні, кількість яких суттєво зросла внаслідок військових дій. Представлено розробку інтелектуального програмного додатку, призначеного для допомоги слабозорим особам в самостійній орієнтації на зупинках громадського транспорту. Система поєднує методи глибинного навчання, комп'ютерного зору та синтезу мовлення для автоматичного виявлення транспортних засобів (автобус, тролейбус, трамвай) та розпізнавання їхніх номерів маршрутів у реальному часі. Для детекції об'єктів застосовано модель YOLOv8, для розпізнавання номерів маршрутів – Tesseract OCR, а для озвучення результатів – офлайн синтезатор pyttsx3. Попередня обробка даних включала анотацію датасету, аугментацію із моделюванням погодних умов та використання фільтрів на основі OpenCV для підвищення точності OCR. Тестування у реальних і змодельованих умовах (туман, дощ, сніг, розмиття, низьке освітлення) показало стабільно високу точність виявлення (100%) та прийнятну класифікаційну продуктивність, хоча стійкість розпізнавання знижувалася за умов обмеженої видимості. Отримані результати підтверджують практичну цінність запропонованого підходу, а подальші вдосконалення зосереджуватимуться на розширенні реальних датасетів, удосконаленні методів попередньої обробки та інтеграції більш потужних моделей глибинного навчання. Запропонована система може стати основою для створення переносних асистивних технологій, спрямованих на підвищення інклюзивності та мобільності слабозорих користувачів в урбанізованому середовищі.*

Ключові слова: *порушення зору; комп'ютерний зір; YOLOv8; OCR; Tesseract; синтез; детекція об'єктів; асистивні технології; глибинне навчання*

About the authors

Барковська Олеся Юріївна, кандидат технічних наук, доцент кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна. E-mail: olesia.barkovska@nure.ua, ORCID ID http://orcid.org/0000-0001-7496-4353

Olesia Barkovska, Candidate of Technical Sciences, Associate Professor at the Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine.

E-mail: olesia.barkovska@nure.ua, ORCID ID http://orcid.org/0000-0001-7496-4353

Нечітайло Олександр Віталійович, магістрант кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна.

E-mail: oleksandr.nechitailo@nure.ua, ORCID ID https://orcid.org/0009-0005-7299-2875

Oleksandr Nechitailo, master's student of the Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine.

E-mail: oleksandr.nechitailo@nure.ua, ORCID ID https://orcid.org/0009-0005-7299-2875

Сердечний Віталій Сергійович, аспірант кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна.

E-mail: vitalii.serdechnyi@nure.ua, ORCID ID https://orcid.org/0009-0007-8828-5803

Vitalii Serdechnyi, PhD student of the Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine.

E-mail: vitalii.serdechnyi@nure.ua, ORCID ID https://orcid.org/0009-0007-8828-5803

Батурін Олексій Олександрович, магістрант кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна.

E-mail: oleksii.baturin@nure.ua, ORCID ID https://orcid.org/0009-0009-2815-3061

Baturin Oleksii, master's student of the Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine.

E-mail: oleksii.baturin@nure.ua, ORCID ID https://orcid.org/0009-0009-2815-3061