**About the data set**

Title: VIP-Mobility360: From RNIB Voices to Real-World Data – A Comprehensive Video Dataset of Outdoor Mobility Obstacles for People with Visual Impairments

Creators:

Dr Wasiq Khan (Corresponding author), Cameron Kinch, Dr Luke Topham

School of Computer Science and Mathematics, Liverpool John Moores University

Year of data collection: 2022

Year of publication submission: 2025

**Description**

we propose a primary and first-of-its-kind dataset designed specifically for Visually Impaired People (VIP) outdoor mobility along with several potential uses in similar domains. Our dataset comprises specific objects identified as obstacles to outdoor mobility for VIPs, based on findings from the Royal National Institute of Blind People (RNIB) survey and other relevant sources. We captured 100 video recordings for 8 types of obstacles that include wheelie-bins, bollards, e-scooters, garbage/trash-bags, advertising-boards, construction/roadwork signs, potholes, and puddles. These objects are captured in diverse and dynamic environmental conditions (e.g., weather, background, ground texture etc.), orientations, placements, and other aspects of variations. The dataset is intended to aid in developing and evaluation of object identification systems that better capture the diversity of public areas and the complexity of urban sidewalk environments.

**Contact details**

PI: Dr Wasiq Khan ([W.Khan@ljmu.ac.uk](mailto:W.Khan@ljmu.ac.uk))

Contact information for forwarding contact: Dr Luke Topham ([L.Topham@ljmu.ac.uk](mailto:L.Topham@ljmu.ac.uk))

**Terms of use**

This dataset, including all images, annotations, and associated documentation, is released under the Creative Commons Attribution 4.0 International License (CC BY 4.0).

Users are free to Share — copy and redistribute the material in any medium or format and Adapt — remix, transform, and build upon the material for any purpose, even commercially. You must give appropriate credit, provide a link to the license, cite the associated publication, and indicate if changes were made.

Citation/s:

[1] Wasiq Khan, Abir Hussain, Bilal Muhammad Khan, Keeley Crockett, *Outdoor mobility aid for people with visual impairment: Obstacle detection and responsive framework for the scene perception during the outdoor mobility of people with visual impairment*, Expert Systems with Applications, Volume 228, 2023, DOI: <https://doi.org/10.1016/j.eswa.2023.120464>

[2] Wasiq Khan, Luke Topham, Cameron Kinch, Hagar Elbatanouny, Abir Hussain, *VIP-Mobility360: From RNIB Voices to Real-World Data – A Comprehensive Video Dataset of Outdoor Mobility Obstacles for People with Visual Impairments,* NeuroIPS, 2025 [ready to submit]

**Project and funding information**

NA

**Contents**

A handful efforts are made to pre-process the raw dataset, mainly to follow the ethical protocols mainly to ensure the identity anonymisation where needed within video frames. The dataset is then organised into appropriate structured folders as can be seen in Figure 1. In the first layer, each subfolder contains specific category of obstacle for instance e-Scooter, bollard, trash bag and so on. The second layer of folders comprises 100 videos for each obstacle within a single category (e.g., 100 videos of E-Scooters), totalling ~1000 video clips for overall 8 objects. The third layer of folders comprises images (i.e. images) extracted from the video clips and corresponding annotations (labels) representing the bounding boxes around the target obstacle within each image frame. In total, there are approximately 150,000 images with corresponding annotations (i.e., 150,000 text files for corresponding labels).

A diagram of a data flow

AI-generated content may be incorrect.

Figure1: Dataset organisation and flow of video processing and annotation

**Methods**

The dataset is collected using mobile camera Samsung S22, with digital camera. Sample image frames demonstrate the diversity of our dataset in terms of camera perspective (360o), scale, size, varying background, weather conditions, and other dynamics where appropriate. We used the normal camera device mainly because ordinary street surveillance devices and available techs e.g., in urban areas are based on this resolution. Likewise, the processing of course-to-fine images can be realistic for assistive technologies. It should be noticed that consistency has been maintained throughout the procedure which might be useful for proposed case scenario (i.e. VIP mobility aid) as well as other cross sector applications.

While there is no identifiable information and/or data, we followed the ethical protocols and application procedure to capture this dataset. The entire dataset is reviewed by the team to ensure no identifiable information is published (e.g., human walking in the background etc.).