Personal and Ubiquitous Computing (2026) 30:2
https://doi.org/10.1007/s00779-026-01854-2

ORIGINAL PAPER

WLRI-AD: assistive device dataset for daily living automation

Katrin-Misel Ponomarjova[1] · Anke Fischer-Janzen[1] · Thomas M. Wendt[1] · Kristof Van Laerhoven[2]

Received: 12 March 2025 / Accepted: 30 January 2026

Abstract

Depending on the degree of disability, simple tasks of daily living can be challenging for people with physical disabilities, such as picking up and placing objects, eating, or reaching for a cup to drink independently. Pervasive technologies such as robotic arms can be used to assist with these daily tasks, allowing patients to regain independence while reducing the need for care. Specialized devices, such as assistive forks or spoons, can facilitate these tasks. Image datasets of everyday objects such as MS COCO do not contain assistive devices, which tend to look different from their non-assistive counterparts. We present the dataset WLRI-AD (Work-Life Robotics Institute–Assistive Devices) to enable a robot to interact with devices in assisted living homes. The benefits of including assistive devices are demonstrated by comparing versions of the dataset with each other and to a baseline. Initial results show an improvement in the detection of assistive devices by training a YOLOv8 model on the assistive devices.

Keywords Assistive technology · Object detection · YOLO · Dataset

1 Introduction

Enabling people with physical disabilities to perform daily activities independently can improve their quality of life [5, 8]. Assistive robotic arms can help achieve this goal. Currently, a variety of control modalities such as joystick, chin, tongue, eye tracking control, and brain computer interfaces (BCIs) are available [1, 11, 19]. However, direct manipulation of a robot can be time-consuming and difficult to learn. Therefore, shared control designs and algorithms are being developed in which the robot autonomously realizes a certain part of the control to reduce the complexity for the user [3, 18]. For example, the robot is able to predict the user's attention, plan and execute trajectories, and grasp objects autonomously. As input, the robot needs information about the task and the environment. Nowadays, cameras and the robot's internal sensor system are used as input [4, 23]. Information from the camera and depth sensors is used to calculate the object position, and point cloud information forms the basis for most grasp planning algorithms. Therefore, methods for locating the desired object within the camera scene must be integrated. One approach is object detection using the algorithm You Only Look Once (YOLO) [17].

The need for a model capable of recognizing assistive devices was identified in the context of an academic project aimed at developing a multimodal user interface that can be used by people with motor disabilities to control a robotic arm in everyday tasks [16]. The tasks included picking and placing objects, pouring water from a bottle into a cup, and serving a drink to individuals with physical disabilities. In the long term, further tasks to implement include serving food and performing personal hygiene, due to their importance to the community [2, 20, 22]. The intended setup involves attaching the robot to an electric wheelchair to assist in a variety of daily tasks.

The approach is based on a shared control where the user selects the object and the corresponding task by using a touch screen [16]. Object detection is performed using YOLOv8 [17]. One of the limitations faced in this project was the difficulty of existing models, such as YOLO pre-trained on the Microsoft Common Objects in Context dataset (continuously named MS COCO) [9], to accurately

✉ Anke Fischer-Janzen
anke.fischer-janzen@hs-offenburg.de

Katrin-Misel Ponomarjova
katrin-misel.ponomarjova@hs-offenburg.de

[1] Work-Life Robotics Institute, Offenburg University of Applied Sciences, Max-Planck-Str. 1, 77656 OffenburgBaden-Württemberg, Germany

[2] Ubiquitous Computing, University of Siegen, Hölderlinstr. 3, 57076 SiegenNordrhein-Westfalen, Germany

Fig. 1 Assistive devices as available from manufacturers. The difference to regular cutlery is their curved and padded handle facilitating gripping the object. Cups might have funnels, handles, and straw applications to facilitate holding and drinking from it. Bigger straps make it easier to grasp scissors

detect the most commonly used objects in assistive environments, such as cups, forks, spoons, and bottles. A review of the available literature revealed that no open source assistive device dataset containing these classes is currently available. Datasets such as MS COCO already do contain classes such as fork, spoon, and cup. Models trained on such datasets are often robust in detecting non-assistive objects. However, assistive objects differ visually from non-assistive objects, resulting in inaccurate detection, as shown in Fig. 1.

We argue that non-assistive devices could be used by the robot, but this would exclude some patients who are not completely paralyzed from cooperating with the robot, for example, when the robot should hand over an object to the user. In addition, it would be more practical to use the system directly in the assisted living environments without having to modify the existing devices. Detecting assistive devices could improve robot grasp planning, since the design is made to facilitate the patient's grasping ability by providing gripping points with thicker, coated handles to reduce slippage.

The aim of this paper is to show that the detection of assistive devices is improved by implementing the WLRI-AD dataset in the training process. Two datasets have been created and will be compared. One of the datasets contains only images of assistive devices (hereafter referred to as WLRI-AD-A, available at [14]). The second dataset is a combination of assistive and non-assistive devices (hereafter referred to as WLRI-AD-B, available at [15]).

The main contribution of this work includes the creation of an open-source dataset containing 762 images of assistive devices. We use the approach to supplement the assistive dataset with images of non-assistive devices, applying data augmentation techniques and using background colors, textures, and varying image angles to improve model performance.

2 Methods

2.1 Data collection

The images included in WLRI-AD-A were taken by the authors at assistive device vendors on multiple occasions after receiving permission from each store. The photographs used single objects as well as more cluttered environments and different fabric patterns and colors to vary the surface underneath the objects (see Fig. 3 for example images). The objects were randomly arranged on the surface. The surfaces were either table tops in white and wood, cloths with floral print, and monochrome colors. Images were taken from randomized freehand positions in the room. Photographers were advised to minimize the time between changing positions and taking pictures to avoid finding scenic views, which would strongly bias the randomness of the dataset. They were also advised to take pictures from higher and lower viewpoints, moving around the scene to capture images with occluded objects and varying illumination. Illumination was further varied by taking pictures in different spots at the locations and at different times of day. The number of objects in each image varied from one to 13 (occurrences (images in dataset): 0 occ. (11 images), 1 (430), 2–4 (77), 5–7 (147), 8–10 (96), and 11–13 (1)). The objects were moved and rotated randomly between shots to randomize their appearance in the scene.

A need for generalization was discovered when the performance of the model trained on WLRI-AD-A (hereafter referred to as Model A) was found to be insufficiently

Table 1 Class annotation distribution

Class	Number of annotations WLRI-AD-A	Annotations included from COCO	Resulting number of annotations WLRI-AD-B
Cup	521	262	783
Spoon	367	156	523
Fork	327	208	535
Scissors	229	182	411
Total	1444	808	2252
WLRI-AD-B	920	1840	2760

accurate in detecting assistive devices in images from other companies or countries when the devices differed greatly from those in WLRI-AD-A. It was decided to manually filter the MS COCO dataset for images of non-assistive cups, spoons, forks, and scissors that fit the purpose of the dataset (e.g., by removing wine glasses from the cup class) and add the images to WLRI-AD-A, creating WLRI-AD-B. The subset was created using custom code provided in [13].

It was hypothesized that giving the model access to images of non-assistive devices during training increases its robustness to variations in assistive devices by allowing it to focus on universal features, such as the “prongs” of a fork or the “bowl” of a spoon.

A total of 567 images from the MS COCO dataset [12] were added to append WLRI-AD-A, creating WLRI-AD-B.

2.2 Dataset characteristics

The model was trained on the classes cup, fork, spoon, and scissors. These objects were available from the two assistive device vendors that the authors were able to solicit for the images. They are also present in the MS COCO dataset, allowing us to compare models and test the hypothesis that supplementing a purely assistive dataset with non-assistive devices will improve detection accuracy. Additionally, classes were considered to add, such as plates, detachable handles, and dressing aids. In the case of assistive plates, adding pictures from large datasets is more preferable, since assistive plates and common plates only vary slightly in appearance.

An objective of the project was to detect bottles for realizing a pouring task. However, since assistive bottles commonly have detachable handles for a better grip, setting them apart from regular bottles, we decided not to take additional images of bottles for this dataset. This decision was made since detachable handles might constitute a separate class. Including them in the bottle labels would result in inconclusive classes. Therefore, the class “bottles” were added to the dataset by including pictures from the MS COCO dataset, as described above. The dataset is available on Roboflow [15]. The images are of the type.jpg with a resolution of 5152 × 3864 px, 350 DPI.

2.3 Ethical considerations

Ethical standards are ensured by excluding all images from WLRI-AD-A that contained faces. In the combined dataset, people may be visible in images from the MS COCO dataset, but as the dataset is open source, there are no ethical issues in using these images. Images in WLRI-AD-A were taken by the authors with full publication rights. Some images were provided by manufacturers and vendors who gave their consent. Due to photo usage rights in Germany, images without consent could not be included in the dataset. This could lead to a bias towards other designs of assistive devices. In reviewing the dataset for potential misuse, the authors did not find any cases that would be ethically challenging.

2.4 Annotation process

The Roboflow annotation tool [6] was used to upload, annotate, and export the WLRI-AD. This allowed the authors to collaborate and make the dataset publicly available. An object detection project was created. The dataset was uploaded and initially annotated by two team members. To ensure high accuracy, completeness, and consistency, all annotations were reviewed and corrected as necessary by the first author. Approximately 20% of cases required annotation correction, primarily due to slight adjustments to the bounding boxes to more precisely enclose the object. The annotated classes were named “cup,” “spoon,” “fork,” and “scissors.” The distribution of annotations is shown in Table 1. Objects that were partially occluded or visible only in reflections or through transparent objects were annotated without restrictions.

2.5 Algorithm selection

The datasets were used to train an object detection algorithm to compare the reliability of the WLRI-AD. For object detection, we decided to use YOLO [17] because of its real-time

Table 2 Distribution of pictures included in each dataset

Dataset	Training set	Augmentation train set	test set	Validation set	Total	Total w/ augmentation
WLRI-AD-A	536	1608	75	151	762	1834
WLRI-AD-B	920	1840	270	139	1329	2249

capabilities, specifically YOLOv8, which is pre-trained on the 81 classes of the MS COCO [12] dataset, one of the largest object detection datasets.

2.6 Training process

Augmentation was performed on both datasets, WLRI-AD-A and WLRI-AD-B. Table 1 shows the number of objects included for each class. On this data, enhancement techniques were applied to the training dataset to address color variation and illumination issues, such as cups of different colors not being detected or being detected with lower confidence in bright environments:

- Grayscale applied to 15% of the images
- Brightness adjustment in the range of –25% and +25%
- Blur up to 2.5 pixels
- Noise up to 2% of pixels

The images were resized to 640 × 640 px to reduce the preprocessing time during training. Both datasets were split into 70% training, 20% validation, and 10% test subsets as recommended and set as default by Roboflow [24]. Random split provided by Roboflow was applied for each version. This ensured an even representation of all classes in the training, validation, and test sets.

For each image in the training set of WLRI-AD-A, three variants were created using augmentation. This resulted in a total number of 1834 images. Since WLRI-AD-B is larger, only two variants were created using augmentation, resulting in a total number of images of 2249. Table 2 gives a detailed overview of the distribution between training, augmentation, test, and validation sets.

Due to the size of both datasets, the model was trained for 50 epochs. Patience was set to 10 epochs to ensure that ongoing improvements were stored. The training strategy for Model A was to fine-tune the YOLOv8 model to specialize it on the assistive devices. Model B was retrained with additional COCO images to generalize the ability to detect the classes and to see how the model performed on assistive and non-assistive objects. For both models, the architecture parameters were held identically. Numbers of layers (355 layers, including Conv, C2f blocks, and SPPF layers) and optimizers were set to auto (AdamW (auto-selected by Ultralytics), learning rate ~ 0.00125, momentum 0.9). Batch size was set to 16. The loss weights were set to box = 7.5, cls = 0.5, and dfl = 1.5.

Table 3 Baseline pretrained YOLOv8 evaluated on WLRI-AD-A

Class	Precision	Recall	mAP50	mAP50-95	Inference (ms)
All	0.0119	0.0695	0.00734	0.00592	5.1

A cross comparison was performed, where the test datasets from A and B were tested on the models trained with datasets WLRI-AD-A and WLRI-AD-B to test if this strategy was applicable to real-world scenarios. This resulted in four comparisons: WLRI-AD-A tested on Training Set A, WLRI-AD-A tested on Training Set B, WLRI-AD-B tested on Training Set A, and WLRI-AD-B tested on Training Set B.

As a performance parameter, the mean average precision (mAP50 and mAP50-90) was chosen. It measures the average precision of the model over several IoU thresholds. Precision and recall were used to understand the trade-off between false positives and false negatives. Finally, inference time was evaluated to measure the time it takes the model to perform predictions. This can be critical in assistive technology applications where real-time detection may be required.

3 Results

Initial tests with the YOLO model, pretrained on the MS COCO dataset, showed that the assistive devices were detected with low confidence or in some cases even with an incorrect class prediction, as shown in Fig. 3 and Table 3. According to [12], a natural class imbalance due to object frequency in real-world imagery exists and impacts model priors. In the baseline test, the classes of “person,” “bicycle,” “car,” and “motorcycle” were detected in the dataset images. However, the classes of “cup,” “fork,” “spoon,” and “scissors” were never detected, even though they existed in MS COCO. Therefore, the model was retrained with the assistive device dataset.

Tables 4, 5, 6, and 7 show the results of the trained and tested datasets. In general, Model A and Model B show sufficient performance to be used as models to detect assistive devices in the presented classes. However, a cross

Table 4 Trained on WLRI-AD-A, evaluated on WLRI-AD-A

Class	Precision	Recall	mAP50	mAP50-95	Inference (ms)
All	0.975	0.925	0.979	0.920	19.58
Cup	0.989	0.924	0.969	0.911	19.16
Fork	0.958	0.902	0.970	0.865	16.78
Scissors	0.964	0.957	0.989	0.920	16.63
Spoon	0.991	0.916	0.987	0.913	17.22

Table 5 Trained on WLRI-AD-A, evaluated on WLRI-AD-B

Class	Precision	Recall	mAP50	mAP50-95	Inference (ms)
All	0.959	0.657	0.724	0.676	22.67
Cup	0.926	0.659	0.735	0.687	21.99
Fork	0.991	0.674	0.741	0.687	22.88
Scissors	0.965	0.575	0.648	0.606	20.39
Spoon	0.955	0.718	0.774	0.723	21.10

Table 6 Trained on WLRI-AD-B, evaluated on WLRI-AD-B

Class	Precision	Recall	mAP50	mAP50-95	Inference (ms)
All	0.909	0.785	0.865	0.758	19.01
Cup	0.912	0.836	0.894	0.820	18.20
Fork	0.898	0.775	0.865	0.744	19.38
Scissors	0.916	0.816	0.861	0.723	20.52
Spoon	0.911	0.713	0.841	0.745	17.18

Table 7 Trained on WLRI-AD-B, evaluated on WLRI-AD-A

Class	Precision	Recall	mAP50	mAP50-95	Inference (ms)
All	0.995	0.987	0.995	0.945	20.22
Cup	0.987	0.993	0.994	0.960	19.88
Fork	1.000	0.969	0.995	0.921	16.48
Scissors	1.000	0.995	0.995	0.962	16.36
Spoon	0.991	0.992	0.995	0.937	16.60

comparison was performed to determine their generalizability when presented with new data.

When WLRI-AD-A is evaluated on test data A (Table 4) and WLRI-AD-B is evaluated on test data B (Table 6), they give comparable results, except for a drop in recall, mAP50, and mAP50-95 for WLRI-AD-B. This could be caused by the included images from MS COCO, which include a wider range of backgrounds and object designs, which when not presented are not detectable for the model. This causality is confirmed by the results of the cross-validation of WLRI-AD-A on test data B and WLRI-AD-B on test data A. While the evaluation of WLRI-AD-B on WLRI-AD-A (Table 7) shows a strong increase in performance parameters, vice versa, WLRI-AD-A struggles to predict the objects strongly (Table 5).

Significant differences between classes were tested using Levene's test and independent samples *t*-test with SPSS. No significant differences between classes were found ($p > 0.5$, tmin(df = 6) = −0.307, tmax(df = 6) = 1.360). This indicates similar performance for all classes. A detailed overview of the comparison can be found in Appendix Table 8. The confusion matrices in Fig. 2 illustrate the performance visualization. The trained model with WLRI-AD-A performs worse on the WLRI-AD-B test data (best score of 0.75), while the reverse is true for the WLRI-AD-B model (best score of 1.00). This indicates that adding pictures enhances the robustness of the predictions. As shown in the confusion matrix visualizing the performance of model B on test data, model B performs better overall, though not perfectly. In particular, the classes of fork and spoon tend to be misclassified. This is still a great improvement compared to the baseline performance shown in Table 3.

Figure 4 shows examples of the detection of the assistive devices. Compared to Fig. 3, the confidence level was strongly increased up to 0.98, which leads to a reliable detection for the presented assistive devices. In the lower left corner, there is an example that general cutlery and cups can still be detected with high confidence (confidence range = 0.82 to 0.95).

4 Discussion

4.1 Application in robotics

WLRI-AD classes represent assistive devices needed in everyday tasks. It was created for the purpose of allowing robots to grasp objects that are very different from common designs, as seen in Fig. 1. The goal in using this dataset was to establish a robust method for human-robot interaction, as presented in previous work [16]. This approach uses object detection to generate interactable objects in the robot's camera scene, allowing for robust task selection. Based on the received object data, the framework calculates the optimal grasp using its internal gripper model.

In this context, we discussed different approaches that might facilitate interacting with everyday objects. One approach is adapting the robot's end-effector to implement various cutlery attachments, which would eliminate the need for grasping forks and spoons. Another approach is adding form-fit or soft grippers to facilitate grasping as shown in [9, 10].

These alternatives have limitations when translated into everyday applications. Cutlery end-effectors attached to feeding robots have limited mobility due to their size, and their appearance may cause prejudice in public places

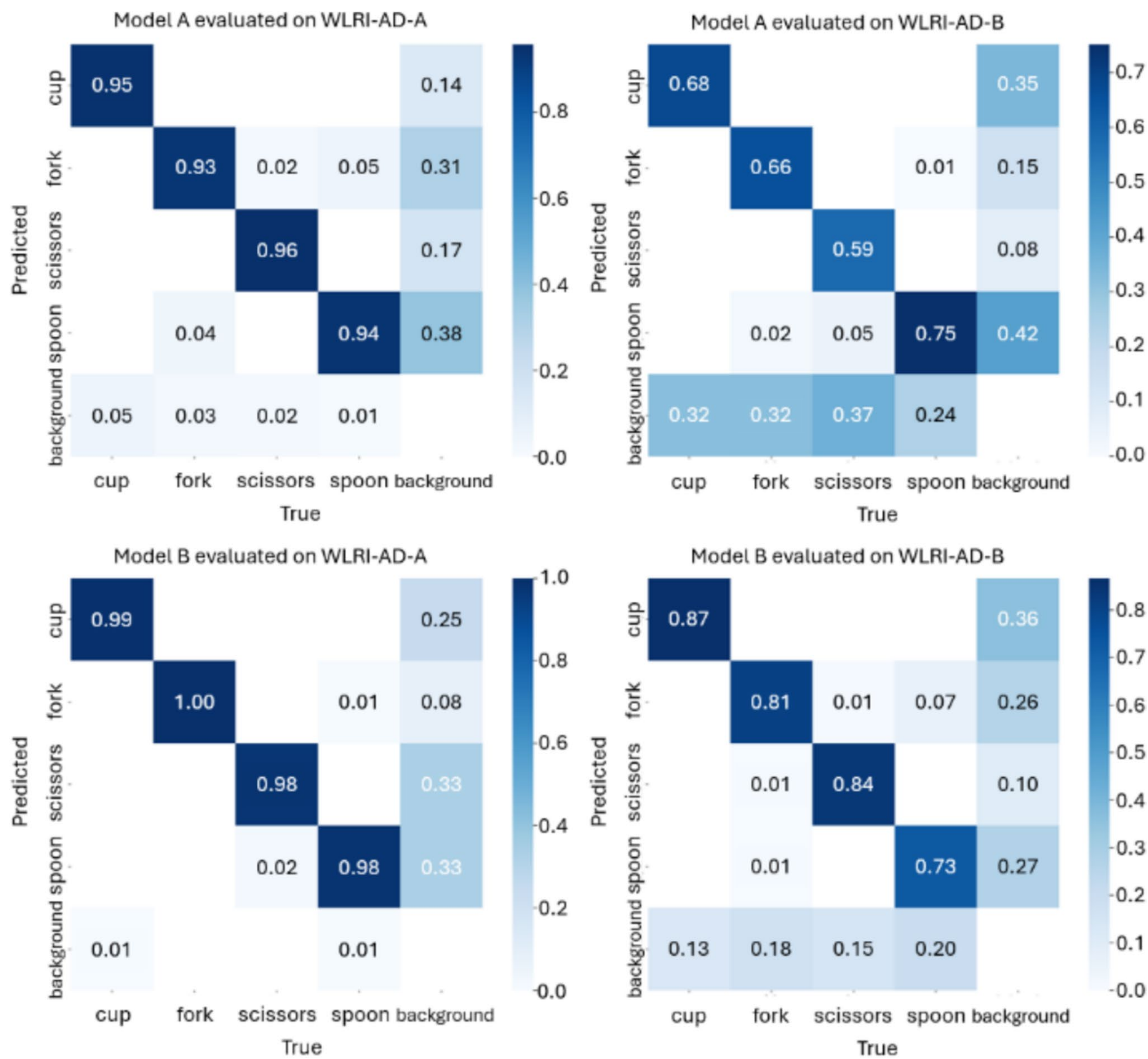

Fig. 2 Confusion matrices of model comparison

such as restaurants. Switching systems that change grippers require space to store the end-effectors or another person to help switch them periodically. This can result in time-consuming changes and, depending on the location of a tool station, hygiene risks. Finally, gripper customization is quite complex. Soft materials must be selected to match the shape to the wide variety of objects available. For example, with a standard two-finger gripper, assistive forks and spoons with padded handles are easier to grip because they do not lie flat on the surface. They need to be rearranged to facilitate grasping, as done in [21]. In this context, this dataset allows robots to interact with a wider variety of objects tailored to the needs of the community and offering advantages over common objects as shown in Fig. 4.

4.2 Limitations and opportunities

The dataset is limited by the small number of available classes. The reasons for this are the novelty of the approach of training a robot to perform a wide range of daily tasks. The research was initially focused on a reduced set of available tasks. Since eating independently is one of the most important tasks for people with physical disabilities, we

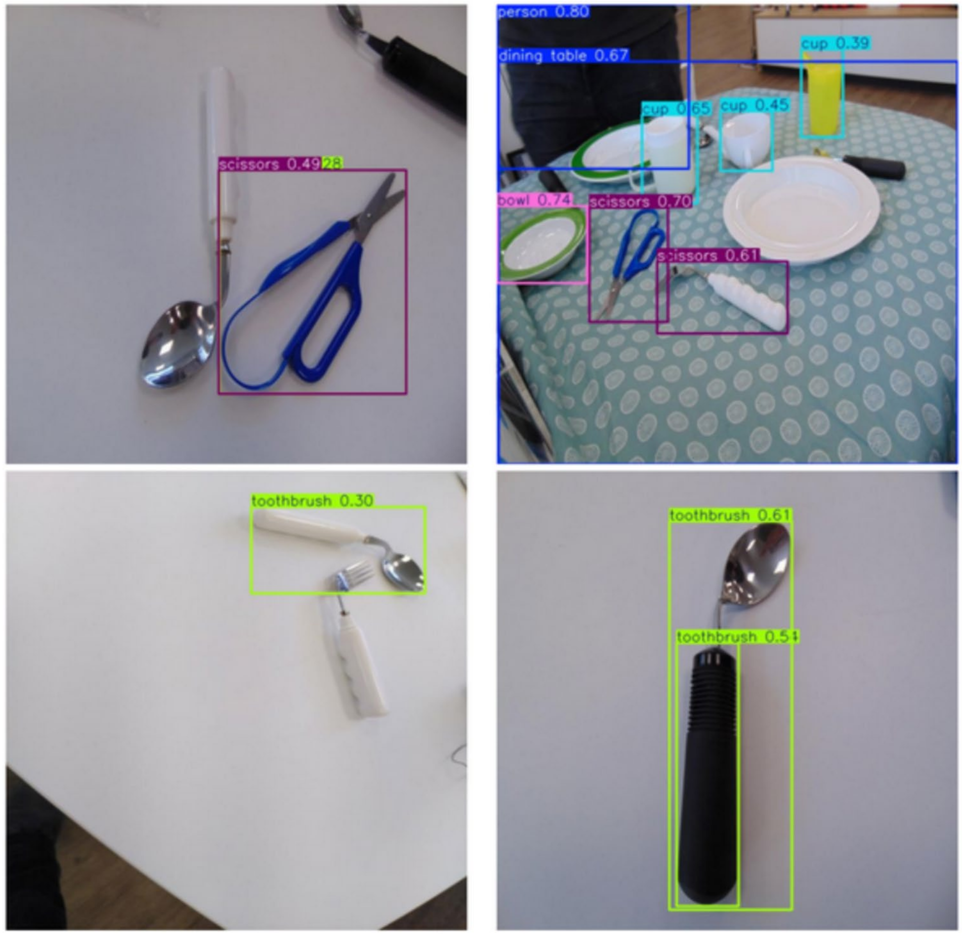

Fig. 3 YOLO pretrained on MS COCO, predictions on the test set of WLRI-AD

focused on object classes used for drinking, pouring, and eating.

This work focused on using objects designed for easier grasping, which resulted in a limited range of assistive objects with varying designs. A variety of games with inclusive design are available, but were beyond the scope of this study. Bowls vary slightly to common designs, most often indicated by a raised rim or a recess to make it easier to pick up food. Since these are only slight variations on common designs, there is a high probability that they will be recognized by the model trained on the MS COCO dataset alone, e.g., as the class "bowl" [7]. Additionally, bottles and other commonly designed objects can be equipped with detachable handles for easier grasping. While adding a class called "handles" could benefit distinguish different assistive devices, it could also lead to a worse performance in object detection since the resulting bounding boxes (bottle handle) would overlap. An online search was conducted to assess whether the variety of objects was sufficient in terms of design differences. Assistive cutlery differs from non-assistive cutlery, due to foam pads, larger handles, or stripes to achieve easier handling. In some cases, an inclined angle between the handle and the tip of the cutlery reduces the need for wrist movements. The objects in the dataset have these features and vary only in color, which is covered by the image augmentation.

Having each object appear in the images a maximum of once could lead to issues in future training. We attempted to mitigate this issue by incorporating images from MS COCO that depict multiple instances of the same object class within a single image. Resampling was not applied to the data split to minimize bias due to the sufficient performance in cross comparison, as shown in Fig. 3. We recommend that researchers using this dataset apply resampling as needed.

Because of this dataset's classic approach, the labeled data can be used for zero-shot and few-shot object detection. In this work, we used YOLOv8n because it is well-established and performs well in the described user interface.

In conclusion, the WLRI-AD performs well with the selected assistive device classes. The question is whether the dataset is applicable to assistive devices in other countries. For example, the "chopstick" class would benefit, as assistive chopsticks do exist, but due to availability, there were not enough varying designs to ensure sufficient detection. Due to data regulations in Germany and limited consent from manufacturers to use their images

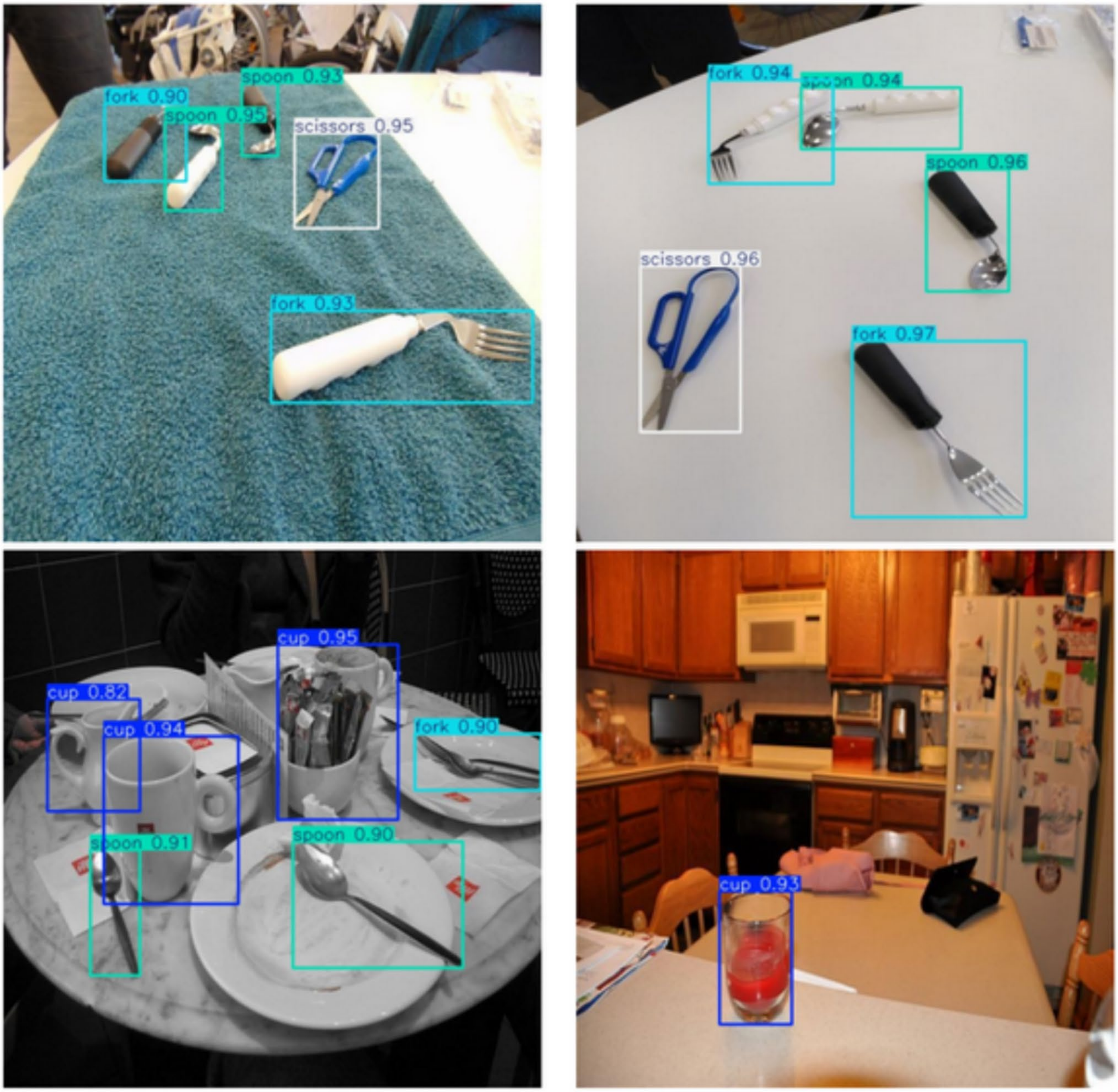

Fig. 4 YOLO pretrained on WLRI-AD-B, predictions on the validation set of WLRI-AD-B

from websites, no other images from online stores could be added to the dataset.

Other areas of research, such as generative AI, can benefit from this dataset by gaining new insights into cutlery design choices. Future work includes testing whether the recognition works with more complex designs, such as belts added to the handle of a shape, or animal designs such as those found in toddler forks.

5 Conclusion

Adding contextually relevant but non-target objects can improve object detection in specialized domains, which could influence future dataset design and model training practices in assistive technologies. In this paper, we present an assistive device dataset. To our knowledge, this is the first open-source dataset that focuses on assistive devices. By cross-comparing the two datasets, it was shown that adding additional images from MS COCO greatly improves the performance and confidence of the model. It allows the detection of assistive devices such as cups, forks, spoons, and scissors. In ongoing work, it is able to recognize objects for robotic grasping to assist people who rely on a robotic arm to perform daily tasks.

Appendix

In Table 8, all parameters are given to compare classes.

Table 8 Statistical evaluation of classes in cross comparison with groups

Comparison	Dependent variable	Levene's F	Levene's sig	$t(\mathrm{df}=6)$	Two-sided p	Mean diff
Cup-fork	Recall	0.000	1.000	0.236	0.822	0.023
	mAP50	0.022	0.886	0.064	0.951	0.005
	mAP50-90	0.005	0.948	0.500	0.635	0.040
	Inference	1.707	0.239	0.549	0.603	0.927
	Precision	0.003	0.955	−0.269	0.797	−0.008
Cup-scissors	Recall	0.303	0.602	0.145	0.890	0.0172
	mAP50	0.418	0.542	0.248	0.813	0.024750
	mAP50-90	10.508	0.265	0.405	0.699	0.041750
	Inference	30.1920	0.124	0.952	0.378	10.3325
	Precision	10.043	0.347	−0.292	0.780	−0.007750
Cup-spoon	Recall	0.108	0.753	0.181	0.863	0.018250
	mAP50	0.054	0.825	−0.016	0.988	−0.001250
	mAP50-90	0.023	0.884	0.184	0.860	0.015000
	Inference	0.332	0.586	10.360	0.223	10.78250
	Precision	0.340	0.581	−0.307	0.769	−0.008500
Fork-scissors	Recall	0.391	0.555	−0.050	0.962	−0.005750
	mAP50	0.329	0.587	0.196	0.851	0.019500
	mAP50-90	0.693	0.152	0.015	0.988	0.001500
	Inference	0.141	0.720	0.216	0.836	0.40500
	Precision	0.442	0.531	0.017	0.987	0.000500
Fork-spoon	Recall	0.221	0.655	−0.049	0.962	−0.004750
	mAP50	0.005	0.948	−0.082	0.938	−0.006500
	mAP50-90	0.147	0.715	−0.327	0.755	−0.025250
	Inference	0.637	0.455	0.473	0.653	0.85500
	Precision	0.096	0.767	−0.008	0.994	−0.000250
Scissors-spoon	Recall	0.162	0.701	0.008	0.994	0.001000
	mAP50	0.350	0.576	−0.266	0.799	−0.026000
	mAP50-90	0.656	0.154	−0.266	0.799	−0.026750
	Inference	0.684	0.440	0.292	0.780	0.45000
	Precision	0.196	0.674	−0.029	0.978	−0.000750

Acknowledgements We thank Schwendemann GmbH and Schaub Sanitätshaus for the access to the assistive devices and their hospitality. We thank our colleagues at Cadida Software GmbH Basian Kayser and Tobias Kraus for their feedback in the model usage.

Author contribution All authors contributed to the study conception, design and writing. Material preparation, data collection and analysis were performed by K.-M. Ponomarjova and A. Fischer-Janzen. The first draft of the manuscript was written by K.-M. Ponomarjova and A. Fischer-Janzen. T.M. Wendt and K. Van Laerhoven reviewed and edited the manuscript and took care of funding acquisition and supervision. All authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding Open Access funding enabled and organized by Projekt DEAL.

Data availability Datasets are available in Roboflow. The link is available in references [14] and [15].

Code availability The script used to train the model can be found in Github. The link is available in reference [13].

Declarations

Ethics approval We declare that the study has been reviewed for ethical claims. As a result, no ethics committee review was requested, as there were no ethically critical cases. For more information on the review, please refer to Sect. 2.3.

Competing interests The research leading to these results received fundings from the Central Innovation Programme (ZIM) under Grand Agreement Acronym HIRAC.

References

1. Awuah WA, Ahluwalia A, Darko K, Sanker V, Tan JK, Tenkorang PO, Ben-Jaafar A, Ranganathan S, Aderinto N et al (2024) Bridging minds and machines: the recent advances of brain-computer interfaces in neurological and neurosurgical applications. World Neurosurg 189:138–153. https://doi.org/10.1016/j.wneu.2024.05.104
2. Chang PH, Park HS (2003) Development of a robotic arm for handicapped people: a task-oriented design approach. Auton Robot 15:81–92. https://doi.org/10.1023/A:1024449018826
3. Chen TL et al (2013) Robots for humanity: using assistive robotics to empower people with disabilities. IEEE Robot Autom Mag 20(1):30–39. https://doi.org/10.1109/MRA.2012
4. Cio YL et al (2019) Proof of concept of an assistive robotic arm control using artificial stereovision and eye-tracking. IEEE Trans Neural Syst Rehabil Eng 27(12):2344–2352. https://doi.org/10.1109/TNSRE.2019.2950619
5. Clarke S et al (2001) Assessing individual quality of life in amyotrophic lateral sclerosis. Qual Life Res 10:149–58
6. B. Dwyer et al. Roboflow (Version 1.0) [Software] (2024) https://roboflow.com/. Accessed 10 Feb 2025
7. J. Gallagher. Microsoft COCO Classes Reference List (2024) https://blog.roboflow.com/microsoft-coco-classes/. Accessed 6 Feb 2025
8. Hammell KW (2004) Exploring quality of life following high spinal cord injury: a review and critique. Spinal Cord 42:491–502
9. Hesse S (2011) Greifertechnik. Effektoren f\"ur Roboter und Automaten. Carl Hanser Verlag, Munich
10. Maya Keely, Brandon Franco, Casey Grothoff, Rajat Kumar Jenamani, Tapomayukh Bhattacharjee, Dylan P. Losey, Heramb Nemlekar (2025) Kiri-Spoon: a Kirigami utensil for robot-assisted feeding. arXive, https://doi.org/10.48550/arXiv.2501.01323.
11. Kyrarini M, Lygerakis F, Rajavenkatanarayanan A, Sevastopoulos C, Nambiappan H et al (2021) A survey of robots in healthcare. Technologies 9(1):8. https://doi.org/10.3390/technologies9010008
12. Lin T et al (2014) Microsoft COCO: common objects in context. Computer Vision - ECCV 8693. https://doi.org/10.1007/978-3-319-10602-148.
13. Ponomarjova K (2025) Assistive training. URL: https://github.com/katrinmisel/assistive training. Accessed 06 Feb 2025
14. Ponomarjova K, and Fischer-Janzen A (2024) Assistive devices Computer Vision Project v8. https://universe.roboflow.com/katrinmisel-ponomarjova/wlri-ad. Accessed 07 Mar 2025
15. Ponomarjova K, and Fischer-Janzen A (2024) HIRAC combined Computer Vision Project v3. https://universe.roboflow.com/hochschule-offenburg/hirac-combined/dataset/3. Accessed 06 Feb 2025
16. Ponomarjova K et al (2023) Enhancing independence through intelligent robotics: an AI- driven assistive robotics interface. UR-AI2023 Artificial Intelligence for time series, robotics and beyond. Mulhouse, pp. 50–58
17. Redmon J et al (2016) You only look once: unified, real-time object detection. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 779–788. https://doi.org/10.1109/CVPR.2016
18. Selvaggio M et al (2021) Autonomy in physical human-robot interaction: a brief survey. IEEE Robot Autom Lett 6(4):7989–7996. https://doi.org/10.1109/LRA.2021.3100603
19. Siean AI, and Vatavu RD (2021) Wearable interactions for users with motor impairments: systematic review, inventory, and research implications. Proceedings of the 23rd International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '21) 7: 1–15. https://doi.org/10.1145/3441852.3471212
20. Stanger CA et al (1994) Devices for assisting manipulation: a summary of user task priorities. IEEE Trans Rehabil Eng 2(4):256–265. https://doi.org/10.1109/86.340872
21. Vermelho R, and Alexandre LA (2022) Grasping and sorting cutlery in an unconstrained environment with a 6 DoF robotic arm and an RGB+D camera. 2022 IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC). pp. 3–8. https://doi.org/10.1109/ICARSC55462.2022.9784780
22. WHO (2010) The ICF: an overview. https://www.cdc.gov/nchs/data/icd/icfoverviewfinalforwho10sept.pdf. Accessed 15 Jan 2025
23. Yang Bo et al (2021) Head-free, human gaze-driven assistive robotic system for reaching and grasping. 2021 40th Chinese Control Conference (CCC). pp. 4138–4143. https://doi.org/10.23919/CCC52363.2021.9549800
24. Jacob Solawetz (2020) Train, Validation, Test Split for Machine Learning. Roboflow Blog: https://blog.roboflow.com/traintest-split/