InVar-100: Industrial Objects in Varied Contexts Dataset

Jan Lehr Vivek Chavan Paul Koch Marian Schlueter Clemens Briese

Fraunhofer IPK, Berlin, Germany

{firstname.lastname}@ipk.fraunhofer.de

1 Introduction

This section outlines the various files aggregated with our dataset.

- **InVar-100 Dataset:** The image data folder contains 100 classes. Each class contains four subclasses/subcategories.
- **InVar-100 Metadata:** The metadata file includes various properties of the objects digitised in the InVar dataset. This includes the weight, length, breadth, height, superclass, material, shape, colour and additional properties. These tags and descriptors allow for further general research work, including modality fusion.
- **Datasheet:** This datasheet document describes the particulars of the dataset and provides more context on the objects, subcategories and related statistics.

2 Dataset

The Industrial Objects in Varied Contexts (InVar) Dataset was internally produced by our team and contains 100 objects in 20800 total images (208 images per class). The objects consist of common automotive, machine and robotics lab parts. Each class contains 4 sub-categories (52 images each) with different attributes and visual complexities. **White background (D_{wh}):** The object is against a clean white background and the object is clear, centred and in focus. **Stationary Setup (D_{st}):** These images are also taken against a clean background using a stationary camera setup, with uncentered objects at a constant distance. The images have lower DPI resolution with occasional cropping. **Handheld (D_{ha}):** These images are taken with the user holding the objects, with occasional occluding. **Cluttered background (D_{cl}):** These images are taken with the object placed along with other objects from the lab in the background and no occlusion. The dataset was produced by our staff at different workstations and labs in Berlin. More details regarding the objects used for digitisation are available in the metadata file.

Training Data	Validation Data			
	D_{wh}	D_{st}	D_{ha}	D_{cl}
D_{wh}	98.6%	3.1%	3.4%	3.6%
D_{st}	4.1%	93.1%	1.5%	1.4%
D_{ha}	31.7%	2.0%	89.4%	13.5%
D_{cl}	35.2%	1.2%	14.6%	88.1%

Table 1: Joint learning accuracy matrix for the subsets of the InVar-100 dataset. The results show that, in order to correctly recognise the object in a given context, it is necessary to introduce the context while training.

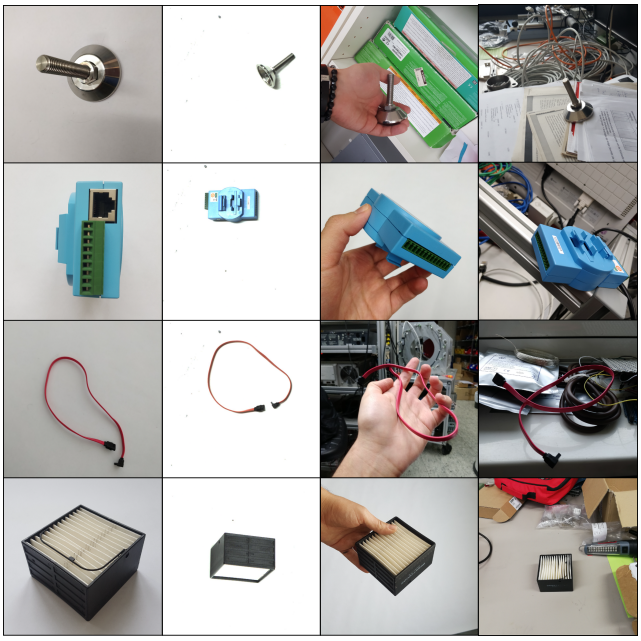
There are other larger datasets on industrial objects, such as the ABC dataset [2], MECCANO [3] and the MCB project [1]. While other datasets contain a higher number of classes and images, the four subcategories in our dataset simulate the different visual contexts in which industrial objects are generally digitised during inference time. The context of the images changes, but the underlying features of the target object remain constant, making it ideal for our investigation. Figure 1 shows sample images for the four subcategories.

3 Details

Table 1 shows the performance of each subcategory on the other subcategories as the validation data. While the objects being recognised remain the same, the differing contexts make it extremely challenging for the model to accurately identify the object. Figure 3 shows the 100 objects from the InVar dataset clustered in 2D. Figure 2 gives histogram plots for the objects in the dataset based on the superclass, the weight, and the length. The attached file contains more properties for the dataset that can be used for further research.

Attribute	D_{wh}	D_{st}	D_{ha}	D_{cl}
Object is centered	✓	✓*	✗	✗
Object in focus	✓	✓	✗	✗
High Resolution	✓	✗	✓	✓
Cropping	✗	✓*	✗	✗
Occlusion	✗	✗	✓*	✗
Clutter	✗	✗	✓*	✓
Blur	✗	✗	✓*	✓*

Table 2: Details on the InVar-100 dataset(* means only a fraction of images have the attribute)



(a) White (b) Stationary (c) Handheld (d) Cluttered

Figure 1: Example of images from the InVar-100 dataset with the subcategories

References

[1] Sangpil Kim, Hyung-gun Chi, Xiao Hu, Qixing Huang, and Karthik Ramani. A large-scale annotated mechanical components benchmark for classification and retrieval tasks with deep neural networks. In *Proceedings of 16th European Conference on Computer Vision (ECCV)*, 2020. 1

[2] Sebastian Koch, Albert Matveev, Zhongshi Jiang, Francis Williams, Alexey Artemov, Evgeny Burnaev, Marc Alexa, Denis Zorin, and Daniele Panozzo. Abc: A big cad model dataset for geometric deep learning. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2019. 1

[3] Francesco Ragusa, Antonino Furnari, Salvatore Livatino, and Giovanni Maria Farinella. The meccano dataset: Understanding human-object interactions from egocentric videos in an industrial-like domain. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pages 1569–1578, January 2021. 1

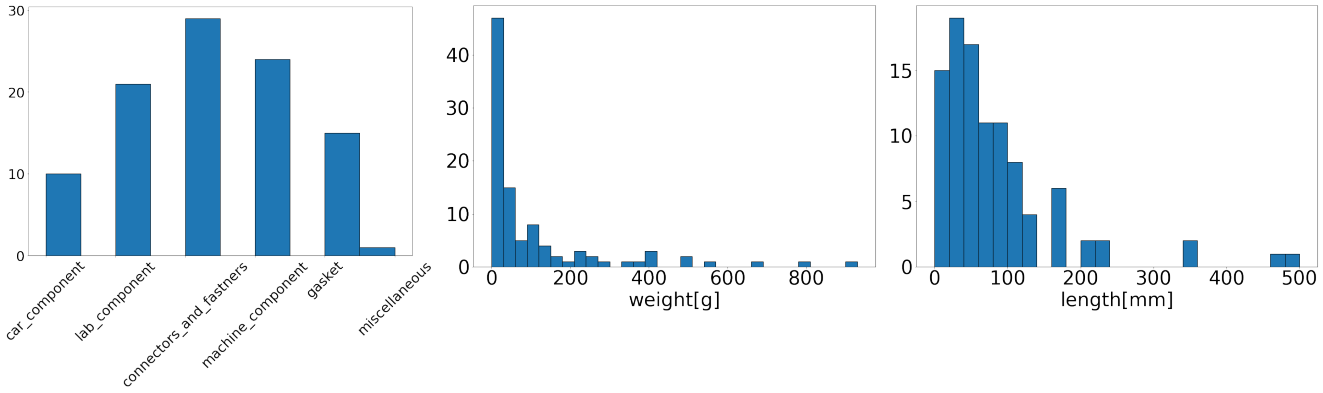


Figure 2: **Left:** A histogram of the Superclasses for the InVar-100 Dataset. **Middle:** Weight distribution between the objects. **Right:** Lengths of the objects.

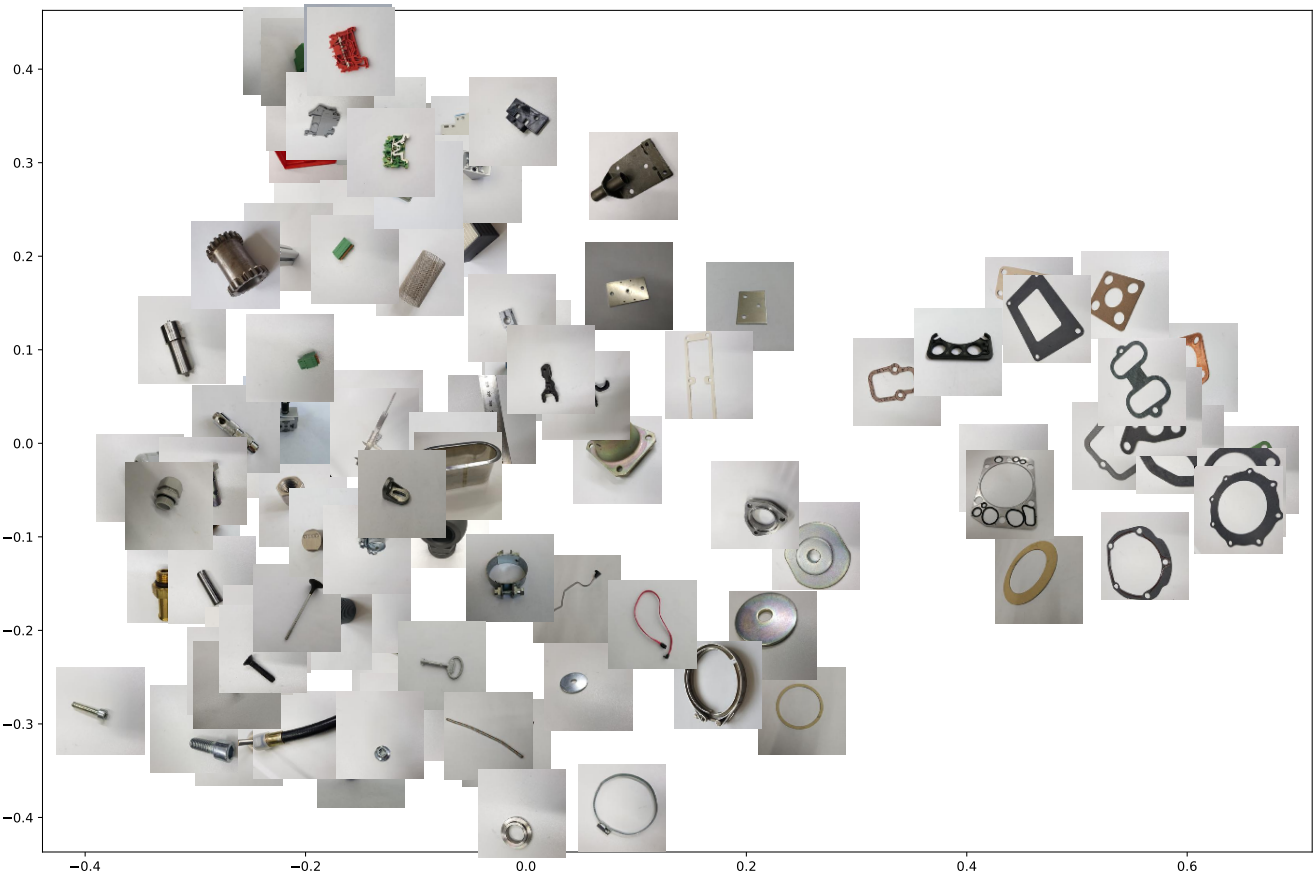


Figure 3: The 100 objects in InVar dataset. Please note that many similar components are placed closely and cannot be seen clearly in the figure.