Supplementary Material: Towards Realistic Evaluation of Industrial Continual Learning Scenarios with an Emphasis on Energy Consumption and Computational Footprint

Vivek Chavan* Paul Koch Marian Schlüter Clemens Briese

Fraunhofer IPK, Berlin, Germany

1 Experimental Setup

Table 1 gives the details of the setup for the experiments. We use the same setup for all our ML trainings for a fair and unbiased comparison. The variable increment sequences used in this study are meant to simulate a realistic project scenario.

Parameter	Value
Train-Val Split	80/20
Optimizer	SGD
lr start	0.1
weight decay	Taken from original implementations
Batch Size	16
Transforms: Train	Resize: (224, 224), RandomHorizontalFlip
Num epochs	Taken from original implementations Pod-AANet: 160 epochs
Other hyperparameters	Taken from original implementations
Num Classes: **Task 0**	Experiment 1: 10 Experiment 2: 10 (Sequence 1), 7 (Sequence 2) Experiment 3: 20 Experiment 4: 10
Experiment 1 Increment Sequences	5 classes per new task 18 total new tasks
Experiment 2	Sequence 1: [10, 10, 10, 10, 10, 4, 7, 2, 14, 5, 3, 9, 6] Sequence 2: [7, 4, 2, 5, 3, 9, 6, 14, 10, 10, 10, 10, 10]
Experiment 3	[20, 4, 10, 2, 9, 1, 8, 16, 7, 5, 8, 1, 6, 3]
Experiment 4 (6 Months)	[10,1,2,5,4,3,6,1,2,4,7,8,2, + 1,3,6,5,4,4,2,1,3,7,4,2,3]
System Memory	16GB
CPU Cores	8
GPU Count	1
GPU type	NVIDIA GeForce GTX 1070
Python version	3.8.12

Table 1: Details of the parameters and setup used for the experiments presented in the paper.

2 Industrial Data Issues

Figure 1 shows some images from a digitalisation station at the sorting facility during research with the EIBA project[1]. The images were taken by operators under strict time constraints while working on large batches of components. This can result in poor image quality and has further knock-on effects for Machine Learning applications. The InVar-100 dataset simulates such environments and highlights the issues in industrial setups.

3 InVar-100 Dataset

The dataset was produced by our staff at different workstations and labs in Berlin. Human subjects, when present in the images, (e.g. holding the object) remain anonymised. We label the subcategories of the dataset as White background: D_{wh}, Stationary: D_{st}, Handheld: D_{ha} and Cluttered background: D_{cl}. Table 2 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 4 shows the 100 objects from the InVar dataset clustered in 2D. Figure 2 shows some objects from the dataset that are visually similar, which introduces additional challenges w.r.t. fine-grained classification and continual learning. Figure 3 gives histogram plots for the objects in the dataset based on the superclass, the weight, and the length. The dataset web page contains a metadata file with additional properties and a datasheet [2] that can be used for further research. The metadata includes properties of the objects digitised in the

*Correspondence: vivek.chavan@ipk.fraunhofer.de
Dataset: http://dx.doi.org/10.24406/fordatis/266.2
DOI: 10.24406/fordatis/266.2
Code: https://github.com/Vivek9Chavan/RECIL

[1]The work (EIBA project 033R226) is funded by the German Federal Ministry of Education and Research (BMBF) in the ReziProK program over the FONA platform for sustainable research.

Figure 1: Sample images taken in by operators at the EIBA project location depicting miscellaneous issues such as clutter, dirt, cropping, occlusion, and blur. Our dataset and the visual contexts address such issues with industrial data collection.

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 2: 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.

Figure 2: An example of visually similar components from the InVar-100 dataset. This adds additional complexity and introduces challenges during incremental learning.

InVar dataset, viz., the weight, length, breadth, and height, along with the superclass, material, shape, colour and additional properties. These tags and descriptors allow for further general research work, including modality fusion.

4 Energy Consumption and Computational Footprint

We used the Shelly smart plug S [3] to log and monitor the task-wise energy consumption of the experiments. We establish a baseline power consumption for the idle workstation and subtract that value from the logged energy consumption during ML training. Algorithm 1 provides more details on our implementation. Additionally, we also log the model sizes and the computational complexities (#FLOPs) for the experiments.

Algorithm 1: Energy Consumption tracking for Incremental Learning Experiments

Input : Class Incremental Learning framework, System Utilisation metrics
Output: Power consumption log, Training time log, other metrics
Power Consumption Log = [ ]
Prompt to close all other applications
Start logging Shelly plug [3] power output
for *t in range(0, 1800)* **do**
    log.append(power,t)
end for
Baseline Power Consumption = sum(log)/1800 Watt
From Dataset get number of classes = N_0
while *task* $T \leftarrow 0$ **do**
    Initial Joint Training with N_0 classes
    Perform model update operations
    for *t in range(0, log size)* **do**
        Updated Log = [Log - Baseline Power Consumption] for all i
    end for
    Compute average power consumption
    Energy Consumption = Power Avg * 144/1000 kWh
    $EnergyLog$ ←Energy Consumption
    $TimeLog$ ←Time Consumption
end while
N ←N_0
Get N_1 From Dataset get the current number of classes N_1
if $N_1 > N_0$ **then**
    Incremental training loop for T
    $EnergyLog$ ←Energy Consumption
    $TimeLog$ ←Time Consumption
end if
Result ← *Allmetrics*
Back up logs

5 Additional Details on Experiments

Figure 5 expands on the results for Experiment 1 (Figure 6 from the main paper) and gives the performance for old and new classes for the four CIL methods. Similarly, Figure 6 provides additional context to Experiment 3 (Figure 10 in the paper) under the varying task increments for the

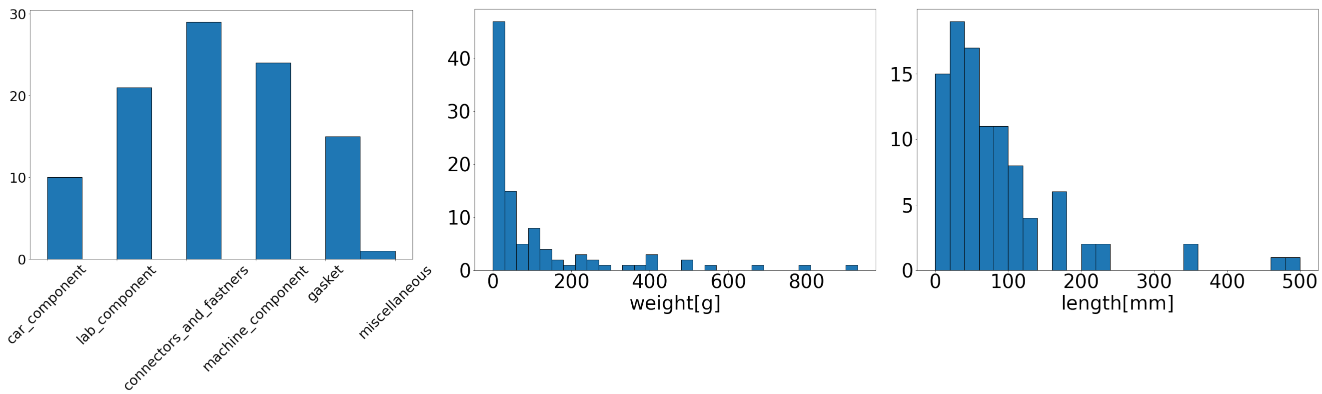

Figure 3: **Left:** A histogram of the Superclasses for the InVar-100 Dataset. **Middle:** Weight distribution between the objects. **Right:** Lengths of the objects.

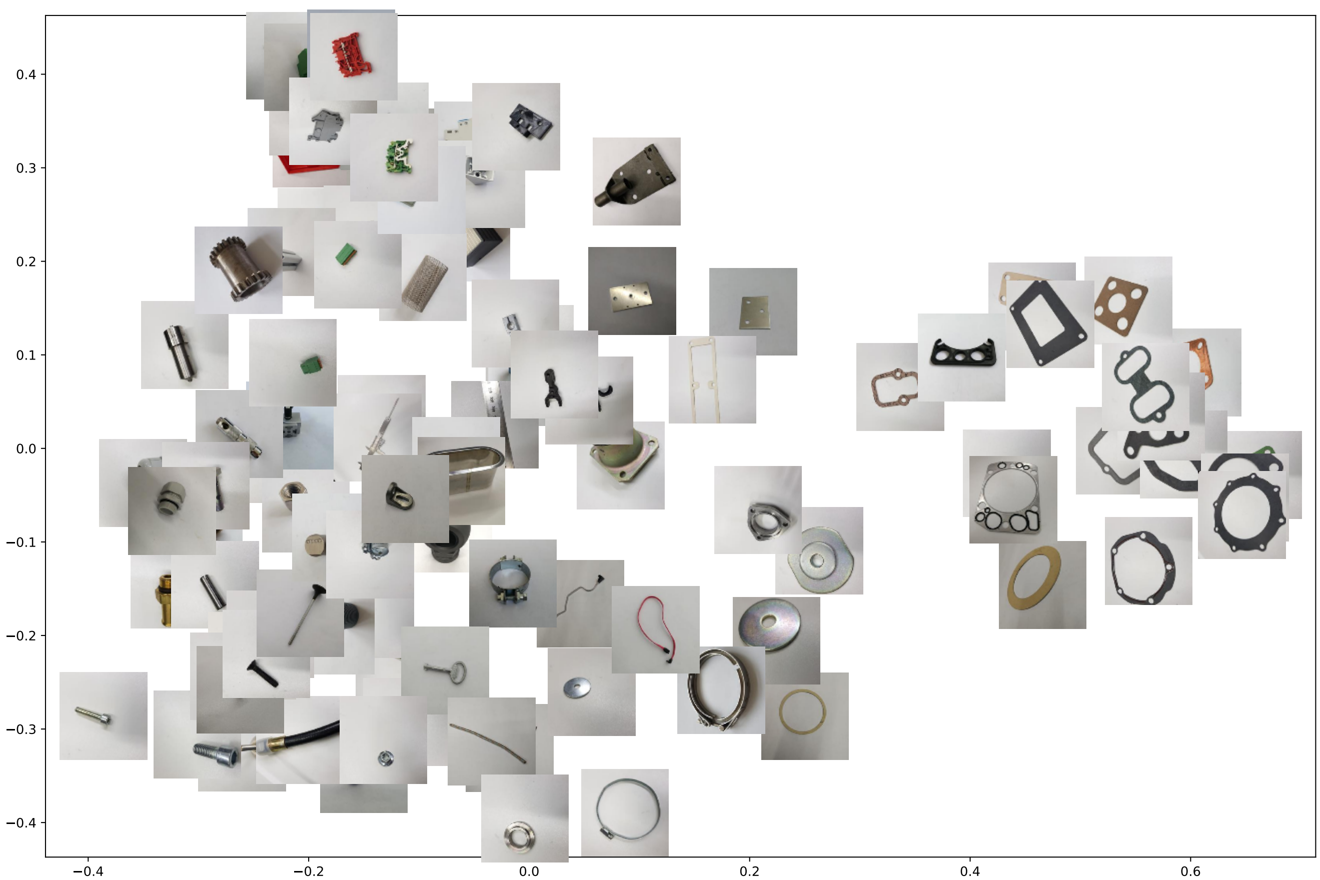

Figure 4: The 100 objects in InVar dataset, arranged according to downscaled PCA distribution (for D_{wh}). We use feature embeddings extracted from DeiT [4] pretrained using DINO [1]. Please note that many similar components are placed closely and cannot be seen clearly in the figure.

different subcategories of the dataset.

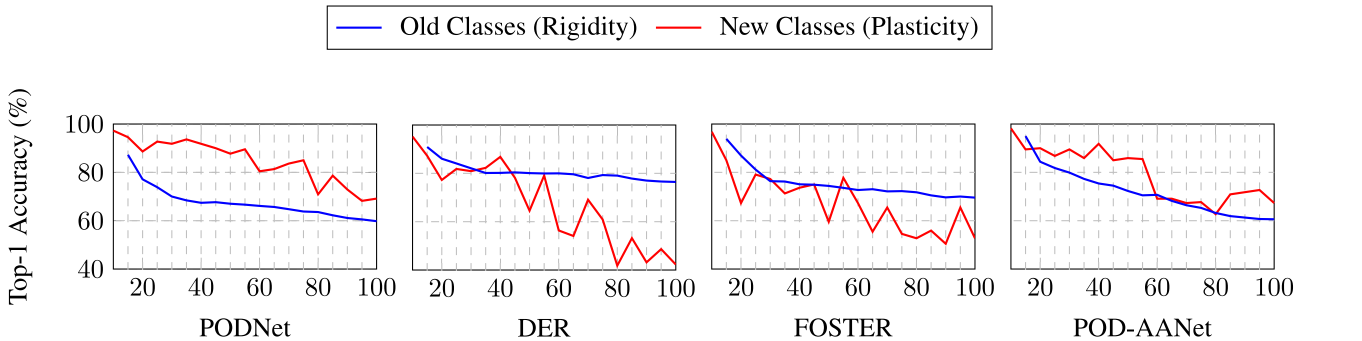

Figure 5: Performance on old and new classes for Experiment 1 (constant increment sizes) with InVar-100.

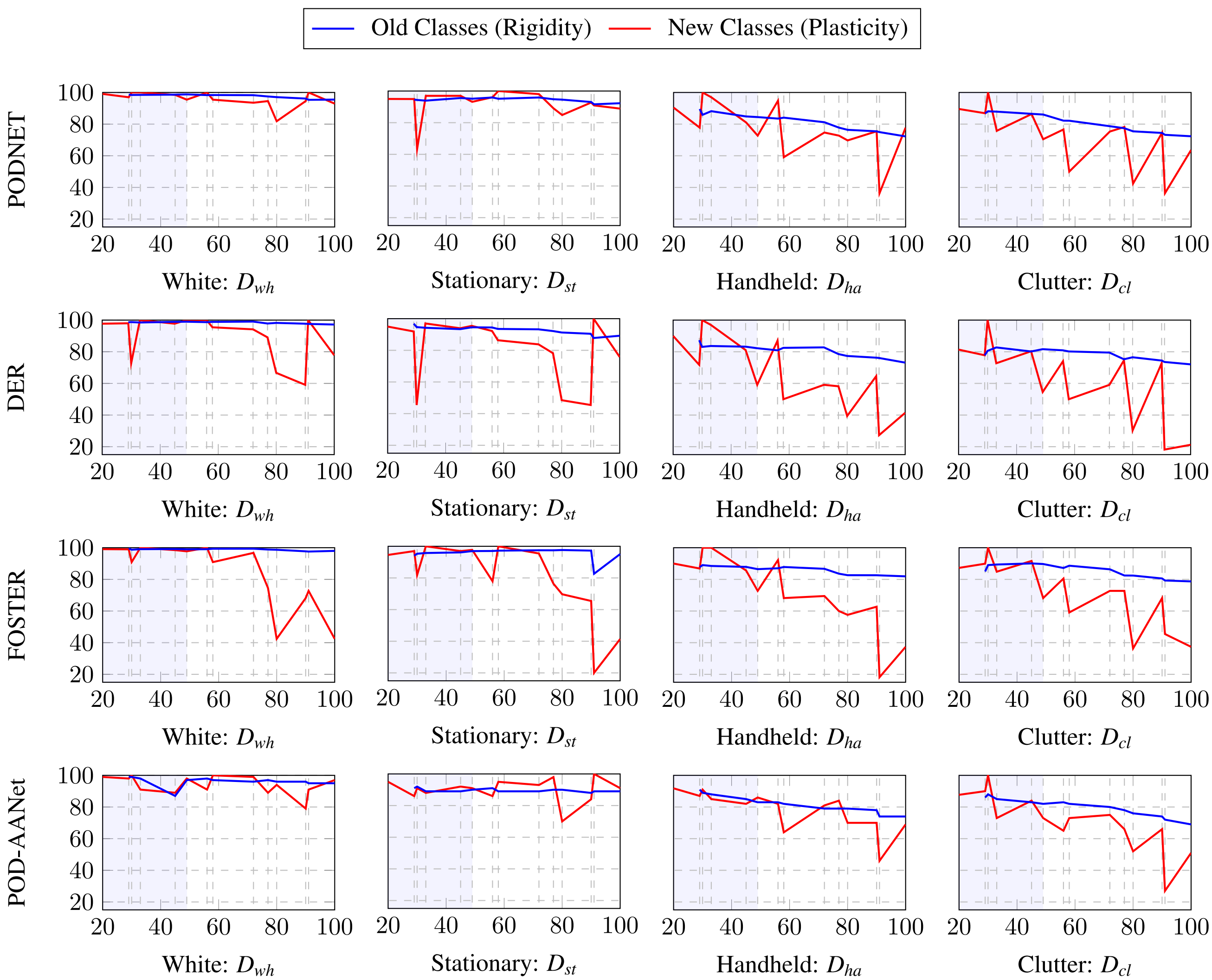

Figure 6: Top-1 accuracy performance on old and new classes for Experiment 3 (varying increment sizes and increased rehearsal memory) for each subcategory of the InVar-100 dataset. Earlier increments (shaded region) have access to all old data as exemplars.

References

[1] Mathilde Caron, Hugo Touvron, Ishan Misra, Hervé Jégou, Julien Mairal, Piotr Bojanowski, and Armand Joulin. Emerging properties in self-supervised vision transformers. In *Proceedings of the International Conference on Computer Vision (ICCV)*, 2021. 3

[2] Timnit Gebru, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna Wallach, Hal Daumé III, and Kate Crawford. Datasheets for datasets. *Commun. ACM*, 64(12):86–92, nov 2021. 1

[3] Allterco Robotics. Shelly plug. `https://shelly-api-docs.shelly.cloud/`, accessed 2022. 2

[4] Hugo Touvron, Matthieu Cord, Matthijs Douze, Francisco Massa, Alexandre Sablayrolles, and Herve Jegou. Training data-efficient image transformers amp; distillation through attention. In Marina Meila and Tong Zhang, editors, *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pages 10347–10357. PMLR, 18–24 Jul 2021. 3