Active Data Collection and Management for Real-World Continual Learning via Pretrained Oracle

Vivek Chavan*[1] Paul Koch[1] Marian Schlüter[1] Clemens Briese[1] Jörg Krüger[1,2]

[1] Fraunhofer IPK [2] Technical University of Berlin, Germany

Abstract

Incremental Learning (IL) deals with learning from continuous streams of data while minimising catastrophic forgetting. This field of Machine Learning (ML) research has introduced several novel approaches and methodologies for varying configurations. However, academic Continual Learning setups generally work with well-curated datasets under predefined conditions, which do not hold for practical applications. In real-world scenarios, the problem of ML starts with data collection and curation. Depending on the application, different challenges are posed w.r.t. data management, such as similar objects, unbalanced data containing sparse samples, visual artefacts, digitisation, and camera setup. This becomes an incrementally compounding issue in Continual Learning projects with data drift and varying conditions. We propose Active Data Collection and Management (ADCM), a straightforward and effective general framework for data collection, coreset/exemplar selection, and analysis. A pretrained Oracle model provides ground truth distribution for the other model that learns incrementally. We couple ADCM with traditional ML/IL setups and demonstrate its suitability for real-world tasks, such as fine-grained classification and anomaly detection. A baseline implementation of ADCM for Class-IL matches state-of-the-art exemplar selection strategies, providing an improvement in average incremental accuracy of 1.5% with Dynamically Expandable Representation (DER) and 4.1% with PODNet against Herding, and 0.8% on old class data against Reinforced Memory Management (RMM); and shows improved performance for general coreset selection. Our code is available at: `https://github.com/Vivek9Chavan/ADCM`

1. Introduction

There have been several rapid and significant advancements in the fields of ML and Computer Vision [22, 32, 36]. As a result, the performance of the current state-of-the-art methods surpasses traditional approaches for classification, semantic segmentation, and 3D applications, among many others [13]. In large part, the availability of large amounts of data and computational resources is responsible for these breakthroughs [20, 59]. When it comes to industrial and real-world applications, acquiring and processing large quantities of data poses several challenges, especially in flexible environments, such as manufacturing [21, 30], reverse logistics [4, 57] or collecting user-centric data [52].

A major purpose of ML systems in such applications is to produce accurate results with minimal system downtime [48]. In that regard, the data must be collected and curated to capture meaningful features of the target objects [47]. Controlled industrial applications often deal with visually similar objects, making classification challenging. Additionally, data collected by non-experts under strict time constraints tend to contain persistent issues, such as visual and contextual clutter, poor lighting, occlusion, object truncation, redundancies, and missing context. Getting the data annotated and sorted by experts proves to be time-consuming and expensive [34]. Moreover, the availability of data depends on several intractable factors in the sup-

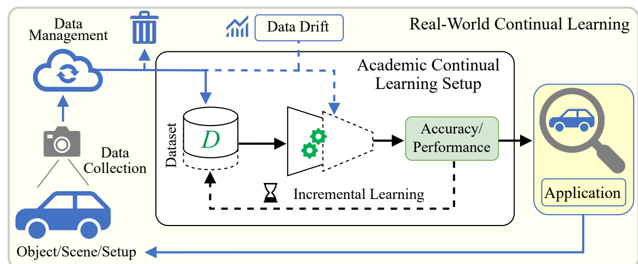

Figure 1. An overview of academic vs. real-world continual learning scenarios. In this work, we address aspects of data collection and management, data drift, coreset and exemplar selection, along with application-specific considerations and the overall efficacy of implementation.

*Correspondence: vivek.chavan@ipk.fraunhofer.de

ply chain [2, 4, 15, 57]. ML models, after being trained on curated data, need to be retrained after the newer data is appended to the dataset. This can either be in the form of new object classes (Class Incremental Learning- Class-IL/CIL) or reprocessed data for older objects (Online Learning or Domain-IL) [58, 66]. This further compounds the challenges w.r.t. data assessment and management. Recent works argue that academic continual learning setups do not translate well to real-world applications, with varying data streams and computational constraints [14, 28].

This paper brings together the fields of continual learning, Active Learning and coreset selection to propose a framework for iterative data collection and curation that can also be applied to coreset/exemplar selection for industrial IL (Figure 1). Like Active Learning approaches, we employ an *Oracle*. However, in our case, the Oracle is a pretrained encoder (leveraging the recent breakthroughs in Computer Vision) that yields an accurate distribution of the collected data, which is taken as the ground truth. The Oracle, in this case, is *blind*, i.e. unlike a human annotator, it does not have access to the class labels, metadata or intended application. We pair the Oracle with the incrementally trained ML model which provides details about the task, performance, labels and tags. Such a setup allows extending the Active Learning principle to not just data selection, but also to digitisation and data collection in controlled industrial setups, allowing better model training with limited data. We demonstrate the applicability of our approach by testing our hypotheses on relevant general-purpose and application-specific public datasets, addressing different real-world scenarios.

2. Related Work

Continual Learning is an active field of research that deals with expanding the capabilities of Neural Network (NN) based architectures to learn from new streams of data [17]. NN models often suffer from catastrophic forgetting when they are retrained on newer data [26, 29]. Van de Ven et al. [66] classify IL problems into three categories: Task, Domain, and Class Incremental learning. They argue that CIL is the most challenging scenario, which often requires representative exemplars to be stored in a memory buffer for retraining. Different approaches have been proposed for enabling CIL; including model expansion [39, 68, 69], regularisation [23, 40, 55], few-shot learning [18, 64, 73], among others. Conventional IL research mainly centres around Convolutional Neural Networks [36, 37] and Residual Networks (ResNet/RN) [32]. Recently, Transformer architectures have been proposed for a multitude of ML applications [67], including vision [22] and have found their way into IL research as well [24, 74]. Modern works incorporate supervised as well as Self-Supervised Learning (SSL) for IL. Research on the application of SSL to continual streams of data has been growing [27, 33, 63]. SSL pretraining has also been shown to be effective in continual learning frameworks [10, 25]. Several works on CIL follow the setup as discussed by Rebuffi et al. [55]. However, computational efficiency and real-world setups for IL have recently gained increased attention [14, 28, 31, 53].

Coreset Selection and Data pruning have been discussed in numerous contexts for ML [49, 50, 60, 70, 76]. In the context of CIL, approaches for selecting exemplars include Herding [55], Mnemonics [41] and RMM [42]. Recent works have also focused on improving memory efficiency for IL [1, 44, 75]. A related field is **Active Learning** , which aims to develop ML algorithms that selectively choose the training data, thereby requiring fewer samples to achieve good performance [7, 8, 54, 61]. Traditionally, such approaches involve an *Oracle* (usually a human annotator), who responds to the unlabeled *queries* and establishes ground truth for training.

In this paper, we focus on the research gaps between traditional continual learning setups and their real-world adoption. We argue that the optimisation process for ML should start during the data collection process itself, which is often not addressed in research [38, 45]. Additionally, we argue that depending on the industrial application, the data management process can be streamlined for object digitisation, data pruning, coreset selection and analysis. We include realistic scenarios with varying task sizes for diverse and fine-grained industrial problems.

3. Methods

Datasets and Setup: We use publicly available domain-specific industrial datasets for our study. The MVIP [35] dataset includes multiview image data and application-specific metadata. The InVar-100 [14] dataset contains industrial objects captured in varying visual contexts. The DIMO [56] dataset contains several metal objects taken with varying orientations, and lighting conditions, and also contains synthetically generated data for each class. The MVTec AD dataset [9] includes several cases of anomalies and defects in industrial setups.

3.1. Real-world Continual Learning Scenarios

Figure 1 gives an overview of an ML implementation in the industry. ML has emerged as a useful decision-making tool for part identification, anomaly detection, sorting and processing. Data collection is a major challenge for real-world applications and poses a significant overhead in terms of operating costs and time. Continual Learning hasn't seen widespread adoption in such use cases [14] but it promises to be an effective approach w.r.t. computational efficiency [28, 53]. This requires data-centric and application-oriented development, involving data acquisition, analysis, and processing in a quick and scalable manner.

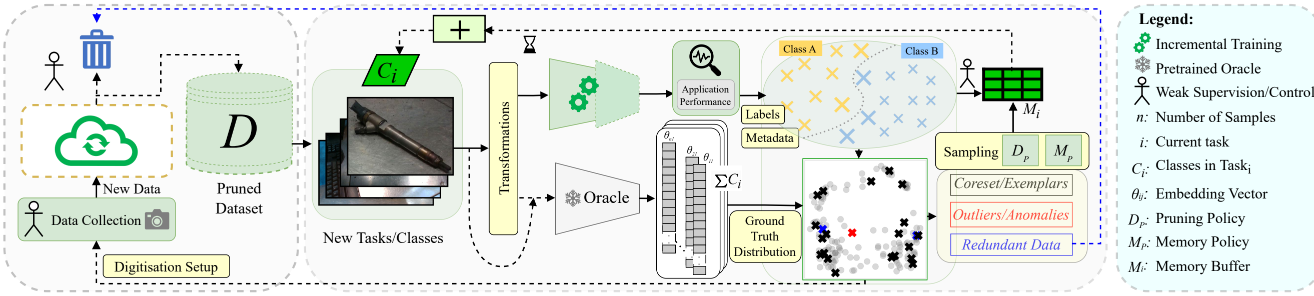

Figure 2. A summary of the ADCM framework. The digitisation and data collection process is analysed w.r.t. the ground truth distribution from the Oracle. The pruned Coreset D is then used for training. The feature embeddings of the new data are incrementally analysed by the Oracle, and outliers and redundancies are removed based on the pruning policy (D_p). The IL model provides additional information about application-based performance and decision boundaries between challenging classes, which is used to append the partial knowledge of the Oracle. The exemplars for retraining are selected depending on the Memory Policy (M_p) and available buffer M_i. We provide further details regarding digitisation and memory management in Figures 4 and 8 respectively.

We define *data analysis* as the process of identifying general patterns and issues in the data with limited or no human supervision. *Data pruning* involves removing *redundant data* (e.g. similar images) and potential *outliers* (poor quality data). Similarly, *coreset selection* deals with curating a smaller representative subset of the larger dataset while maintaining the essential characteristics and structure [49, 60]. Continual Learning in industrial applications includes Task, Domain and Class-IL. Following the observations by van de Ven et al. [66] and the current standard practice in IL [55], we focus on CIL with non-overlapping classes as the most challenging scenario. We take image classification as the benchmarking task. The objective is to continually obtain a pruned dataset (coreset) D from the collected data. With CIL, new non-overlapping classes are introduced sequentially over training tasks *0, 1...T*, each containing new classes C_0, C_1...C_T. The model must be able to classify the test data from all available classes $\sum_{i=0}^{t} C_i$ at a given phase t (t$\leq$T) of the project.

Conventionally, CIL involves learning from new data and assessing the feature distribution of all the available data using proposed algorithms (e.g. Herding) to select the most relevant samples, which are reintroduced during new tasks. Depending on the method, the model may struggle to learn from new data or may find it challenging to retain learning from older tasks [14]. The problem is exacerbated for fine-grained problems and data with clutter or other artefacts.

3.2. Active Data Collection and Management

We decouple the incremental training problem from the representation learning problem, thereby mitigating the effects of catastrophic forgetting on the latter. This allows a separate, impartial encoder to be used for coreset/exemplar selection and analysis. Furthermore, conventional Active Learning works with the available data to select the samples to label for obtaining the best possible accuracy. We take this further and argue that there is significant scope for influencing the data acquisition process itself. In industrial setups, this may involve designing digitisation stations and making decisions for choosing camera positioning, lighting setup and other specifications, which have a significant effect on the downstream application. Our approach, ADCM (Active Data Collection and Management) is elaborated in Figure 2. In other cases, where the data may already be available can also benefit from our implementation w.r.t. data pruning and analysis.

Oracle: Conventionally, the role of the Oracle is to label a subset of the dataset. We modify the role of the Oracle to encode visual data for accurate analysis in relation to other data samples. The Oracle in this case is a strong pretrained model, which allows repeatable use on vast amounts of data for quick processing. The *query* posed to the Oracle is *to provide a mathematical relation between a given sample and every other sample in high dimensional latent space*.

Features learnt using SSL have been shown to generalise better to other downstream tasks [11, 16, 62]. We analyse the features learnt by different state-of-the-art SSL approaches including MoCoV3 [16], SwAV [11], Barlow Twins [71], DINO [12], VICReg [5] and VICRegL [6] on the different datasets. We also study the effect of different model architectures for a given pretraining method, including RN50 [32], Data efficient image Transformer (DeiT) [65], ViT-B [22], RN50w5 and RN200x2.

Idrissi et al. [34] have shown that ViTs, especially with SSL pretraining can be robust to transformations and variances in data. Of particular interest to our work were DINO and VICRegL. DINO is more sensitive to texture, shape and occlusion by humans while being more robust against lighting, object size, and occlusion by other objects [34]. We verify this on the InVar-100 and DIMO datasets and find

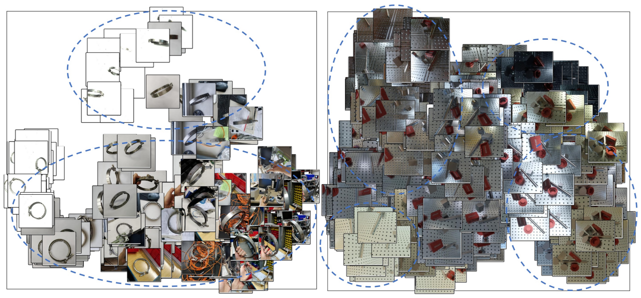

Figure 3. Intra-Class distribution of feature vectors from objects in the InVar-100 (left) and DIMO (right) datasets extracted using pretrained DeiT-S [12] and downsampled using PCA. The embeddings are clustered based on the visual context and object background. Notably, images containing similar object orientations also appear closer which is not observed in corresponding embeddings from supervised models.

that embeddings obtained from DINO cluster the images based on the object orientation, shape and lighting. Figure 3 shows the intra-class distribution for images under different backgrounds and orientations. The images are clustered not only based on the background context but also based on the object pose. For instance, the images gradually vary from a flat/oval object profile to a circular profile along the Y-axis. The embeddings from VICRegL also closely match these findings, however, we find the DINO features to be objectively superior. Caron et al. [12] show that SSL ViT features contain information about the scene layout and object boundaries, and serve as excellent inputs for zero-shot k-NN classification. Similarly, VICRegL has been shown to be effective for classification as well as segmentation tasks. As a baseline, we take DeIT-S pretrained using DINO as the pretrained Oracle for our work in this paper **($\mathbf{ADCM_0}$)**.

Data Collection and Digitisation: We demonstrate the adoption of ADCM for actively controlling data collection by utilising the shape and texture sensitivity of DINO. The generalised approach is shown in Figure 4. The objects of interest may be digitised on a multi-view digitisation station or via handheld cameras. The collected data is then studied and sampled based on the saliency of the embeddings with weak supervision. Adjusting the camera positions, lighting and other parameters, the Oracle is queried for understanding the relative feature distribution and the impact of the changes. Thus, the digitisation setup can be actively controlled to influence the ML pipeline. Our tests on multi-view table-mounted and handheld devices with various focal lengths (9mm to 50mm) showed identical sensitivity and adaptability to adjustments. Industrial anomaly and defect detection is an application where the setup and digitisation environments are controlled to effectively flag divergence from a preset output. Studying the sample distribution for MVTec-AD dataset [9] from ADCM_0, we see a noticeable

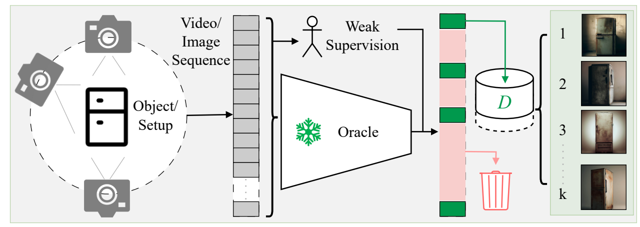

Figure 4. Multiview stationary object digitisation, data sampling and Coreset selection. The operator digitises the object via a digitisation station or a handheld device. The video/image frames are sampled based on the saliency of feature embeddings using the pretrained Oracle. Analysis of the sampled data is available to incorporate explainability into the pruning process. Weak supervision is optional and depends on the availability of additional information to the operator.

clustering of *good samples* and clear separation of *anomalies*. As illustrated in Algorithm 1, the Oracle is queried: *If the lighting, camera position, etc, were to be altered by a factor of X, what is its impact on data collected within the same setup? And how do the different setups compare, given the same object samples and labels?*

Algorithm 1 ADCM for Anomaly Detection

Require: A set of labelled images $\{I_1, I_2, ..., I_N\}$ from C camera perspectives, labelled as Good (G) or Anomalous (A).

$\Theta \leftarrow \{\text{Oracle}(I_1), \text{Oracle}(I_2), ..., \text{Oracle}(I_N)\}$
$S_{\text{max}} \leftarrow 0$
$c_{\text{optimal}} \leftarrow \text{null}$
for $c \in \{1, ..., C\}$ **do**
  $\Theta_c \leftarrow \text{PCA}(\Theta, c, 32)$ ▷ Reduce dimensions for c
  $\mu_G \leftarrow \frac{1}{|G|} \sum_{i \in G} \Theta_{c_i}$ ▷ Centroid for G
  $\mu_A \leftarrow \frac{1}{|A|} \sum_{i \in A} \Theta_{c_i}$ ▷ Centroid for A
  $D_G \leftarrow \frac{1}{|G|} \sum_{i \in G} \|\Theta_{c_i} - \mu_G\|^2$ ▷ Dispersion for G
  $D_A \leftarrow \frac{1}{|A|} \sum_{i \in A} \|\Theta_{c_i} - \mu_A\|^2$ ▷ Dispersion for A
  $S_c \leftarrow \|\mu_G - \mu_A\|_2 - \alpha(D_G + D_A)$ ▷ Adjusted separation for camera c
  if $S_c > S_{\text{max}}$ **then**
    $S_{\text{max}} \leftarrow S_c$
    $c_{\text{optimal}} \leftarrow c$
  end if
end for
Output: Camera c_{optimal} with separation S_{max}.

Sampling: We use sampling as a general term for coreset selection, exemplar selection and data pruning based on the application and performance of the IL model. As a baseline, we utilise the feature embeddings θ from the Oracle for the analysis. We use the Euclidean distance matrix to study the interrelation between all image pairs within an object

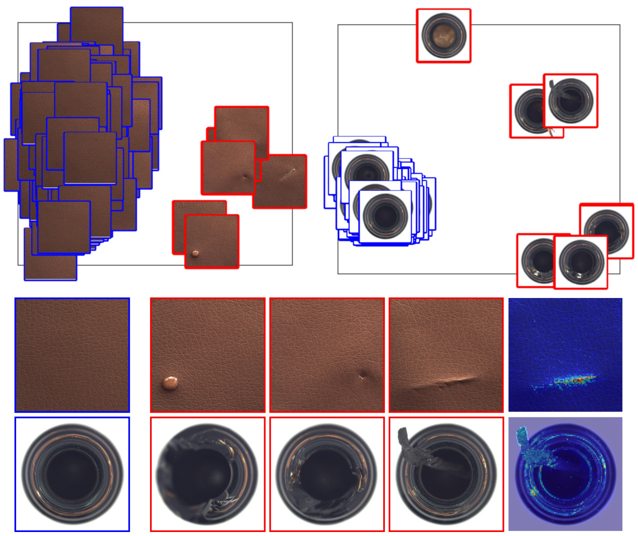

Figure 5. ADCM$_0$ implementation applied to study industrial data collection and flag issues w.r.t. anomalies/defects. MVTec AD dataset is used as the example dataset. This shows how our approach can be used to actively identify and collect data that is relevant to the ML application in controlled industrial environments. **Top:** A 2D downsampled plot showing the embedding distribution for the classes *leather* and *bottle*. **Bottom:** Examples of good and anomalous objects. A self-attention map for the last anomalous image for each class is also shown.

class in the dataset and flag similar image pairs that fall under a predefined threshold.

$$\forall i,j \quad \text{Dist}(\boldsymbol{\theta_i}, \boldsymbol{\theta_j}) = \sqrt{\sum_{n=1}^{dim_{(Oracle)}} (\boldsymbol{\theta_{i,n}} - \boldsymbol{\theta_{j,n}})^2} \tag{1}$$

We downscale the feature vectors to reduce processing speed. Based on our observation, reducing the DeiT-S outputs from `384` to `32` gave optimal results. The number of exemplars to be selected for IL depends on the memory policy and the number of old classes. Similarly, the coreset can be selected based on the computational requirements. K-means clustering is employed to identify intra-class clusters and samples closest to the cluster centre are taken as the representative, as shown in Figure 6. This approach is useful when a significant data reduction is needed. Otherwise, the clustering would be unnecessary and expensive, and the progressive elimination of redundant data (with an increasing threshold) is more suitable.

Variable Sampling: In case of feature imbalance in the dataset, we compute a class-wise weight (w_i) that is proportional to the feature spread for the class. Using, w_i, a variable cluster assignment K_i is used to select an appropriate number of exemplars from a given class i. Figure 7 shows an example.

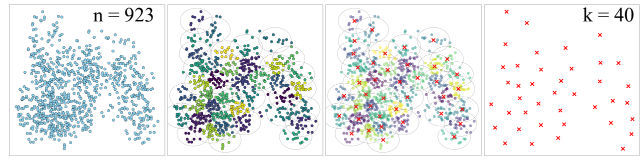

Figure 6. Encoded feature vectors for an object from the DIMO dataset. PCA downsamples the embeddings to 32 dimensions, but 2D plots are shown here for visualisation. Using K-means, optimal intra-class clusters are identified. The sample point closest to the cluster centre is taken as the representative image for that cluster. Thus, the 923 images in the class are reduced to 40 representatives.

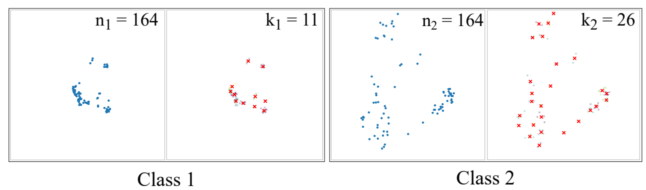

Figure 7. Variable sampling for datasets with feature imbalance for InVar-100 ^{U} (an artificially augmented (colour jitter, saturation and rotation) variant of InVar-100 containing feature imbalance between the subcategories). Both classes have the same number of original images, but different numbers of exemplars are sampled depending on the distribution. Related results are in Figure 13.

$$w_i = \frac{\sum_j \theta_j}{\sum_i \sum_j \theta_{ij}} \tag{2}$$

This class weight governs the number of exemplars/samples that shall be stored from the given class. *Memory* represents the total budget for storing old class exemplars. During the K-means clustering, the k value is given as:

$$K_i = w_i \times \frac{memory}{num_{classes}} \tag{3}$$

Data Quality: The problem of selecting the correct subset of data to improve model performance and robustness in different scenarios is non-trivial. Our assessment shows that image data with objects against a clean background yields the highest accuracy in incremental training over long timelines; this is in agreement with [14]. We classify this as *good data*. Depending on the scope of the project and the image data, cluttered, out-of-distribution or *challenging data* may be unsuitable for training or may represent an underrepresented visual context. In a continual learning scenario, when new data is successively introduced, we propose an expansion of the memory policy to avoid the entanglement of features and removal of relevant data as shown in Figure 8. The data which is not used for IL is archived and introduced periodically for model testing. The data that

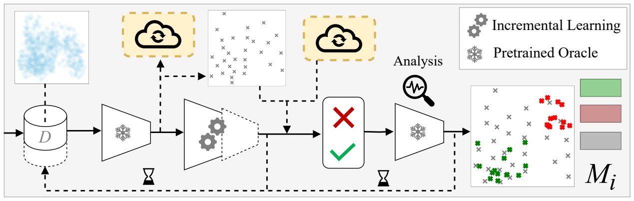

Figure 8. Proposed solution for incremental training and data curation during long project timelines. The memory budget M_i comprises of an ensemble of *good data* , *challenging data* , and *sampled exemplars* based on part identification performance, visual context and additional clustering and analysis.

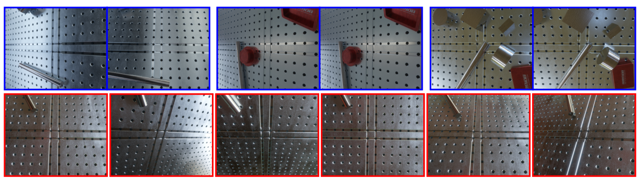

Figure 9. Examples from DIMO dataset with three identified redundant image pairs and outliers using our approach. The image on the bottom left was provided as a seed outlier image based on which, the other outliers are flagged. These may optionally be reviewed by the operator or directly sorted.

the model fails to correctly classify is flagged and reintroduced for analysis. Similarly, *good data* may be added to the memory based on the requirements. The memory budget M_i thus contains a mixture of different data attributes, viz. the *sampled exemplars* (that represents the coreset), the *good data*, and the *challenging data*. We take the unnormalised distribution from the pretrained encoder as the ground truth. This approach maintains the standard coreset and also reintroduces the data that may be incorrectly removed, or can potentially improve model performance and robustness. Similarly, the distribution change between the previous data streams (or stored exemplars) and the new data can be compared for understanding *data drift* and related changes.

Human in the Loop: Human control and feedback is essential for practical implementations. Operators can function as meta-Oracles to shape the decision-making process. E.g. the user may flag certain examples of outliers (weak supervision), based on which, data with similar issues can be pruned (Figure 9). The redundant samples and outliers may also be identified and flagged for review.

4. Experiments and Discussion

Setup: We use FACIL [46] and PyCIL [72] toolboxes along with the open-source CIL implementations with varying task sizes for our Experiments. A dedicated workstation is used for all experiments. We use an 80/20 train/val split for all datasets and use consistent hyperparameters.

Data Collection: The MVIP dataset contains metadata, camera tags and details for a multi-view setup. We aggregate the embeddings from each camera for different objects, as shown in Figure 10. It shows the image embeddings for original uncropped data as well as corresponding region of interest (ROI) crops. Next, we study the effect of object rotations and orientation [35] on the meaningful *information gain* using feature distribution for each view. The underlying hypothesis is that clusters with greater variance carry greater information regarding the object. Based on this, we prune the dataset to contain only 3 views per class (views 1, 4 and 8) and reduce the data down to 36 images per class from 120 (MVIP Coreset). *If we were to design the dataset with a limited storage budget, the Camera positions 1, 4 and 8 would be ideal for capturing the most meaningful aspects of the objects.*

Dataset	Coreset Size	Coreset Accuracy	Random Selection	Full Training Set
MVIP	11088/38000	**77.48%**	43.39%	89.4%
DIMO	37600/424000	**68.91%**	52.11%	94.2%

Table 1. Top-1 accuracy on coreset selected using ADCM with *weak supervision* based on knowledge of data collection, compared against random sampling and the full dataset. ResNet-18 was trained from scratch on the data. The coresets significantly outperform their random counterparts.

Coreset Selection with Weak Supervision: With the DIMO dataset, we employ our approach to curate a coreset containing an approximately equal number of real and synthetic images. Additionally, the data was weakly supervised to contain images from different lighting conditions, poses and orientations. The resulting coreset contains only 8.8% of the original training data (DIMO Coreset). A ResNet18 model was trained from scratch on the MVIP and DIMO Coresets. For comparison, we randomly select another subset of the datasets of the same size. The average test accuracy for the two cases is given in Table 1, along with results on the full training set. The results on the coresets are 34.09% and 16.8% better than a randomly chosen subset.

Exemplar Selection for IL: Figure 11 shows the comparison of our approach against Herding [55] and Figure 12 compares it against RMM [42], both for InVar-100. We test the approaches using PODNet [23], FOSTER [68] and DER [69] for Herding and using POD-AANet [40] for RMM. Constant as well as variable task sizes were introduced to simulate realistic scenarios. A constant limit of 20 exemplars per class was applied for all tasks for an impartial comparison. Our approach outperforms Herding by a statistically significant margin in all tested use cases. ADCM$_0$ outperforms Herding by 1.4% on the DER implementation and by 4.1% on PODNet. It outperforms RMM by 0.8% w.r.t. performance on old classes. The improvement is due to the

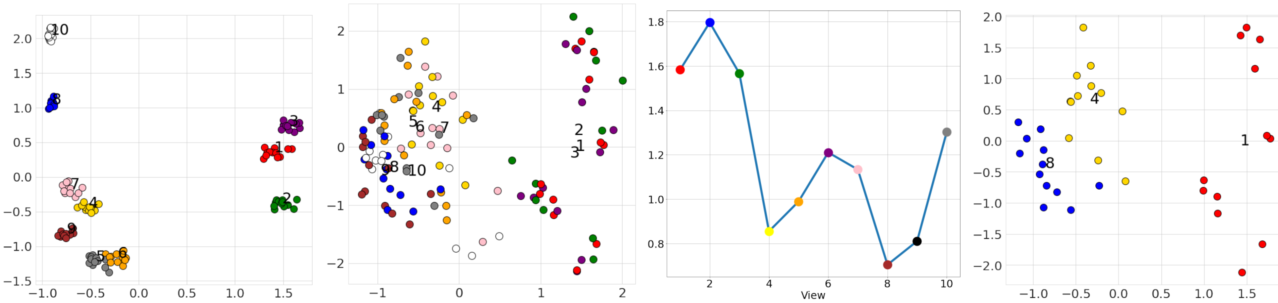

Figure 10. An example of data collection and curation using available metadata: Analysis of data from different views of the MVIP dataset. Each colour represents a different camera view. (1.) Uncropped image embeddings. The label locations correspond to the cluster centroids. (2.) Corresponding ROI crop embeddings (3.) Analysis of the variance of object rotations captured by various views. (4.) Coreset with 3 views. Corresponding comparison between the coreset, random selection and full dataset is given in Table 1.

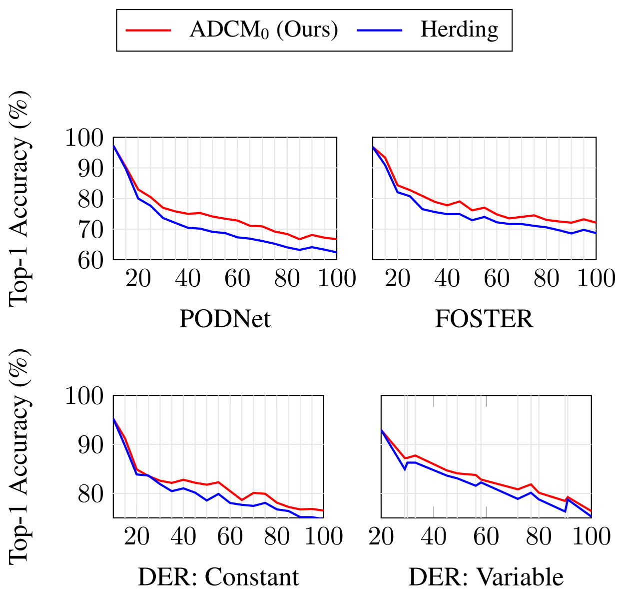

Figure 11. **Top:** Comparison of our approach (ADCM$_0$) against Herding using PODNet and FOSTER CIL implementations. **Bottom:** Our approach vs. Herding for constant (left) and varying (right) task sizes using the DER implementation.

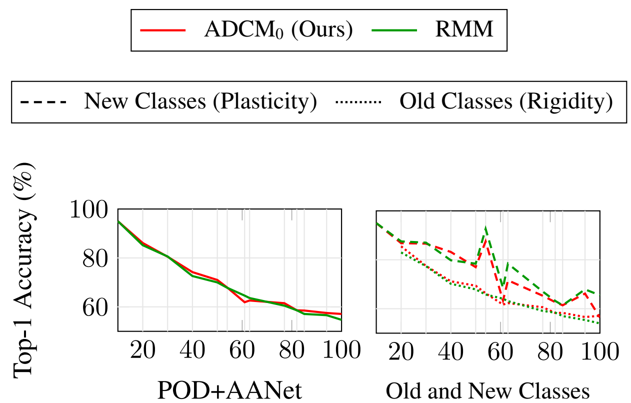

Figure 12. Comparison of our exemplar selection approach against RMM using POD-AANet implementation for all classes (left) and a breakdown of performance on old and new classes (right).

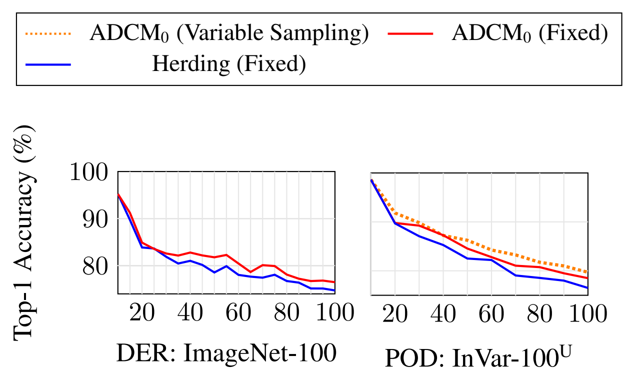

Figure 13. **Left:** Comparison of our exemplar selection approach against Herding using DER implementation for ImageNet-Subset. **Right:** Comparison of our variable exemplar election approach against Herding on InVar-100^U with PODNet. ADCM$_0$ yields state-of-the-art results for varying scenarios and setups.

exemplars being more representative of the underlying class distribution. In Figure 13, we compare the performance on ImageNet-Subset using DER and on InVar-100^U using PODNet via variable sampling strategy. ADCM$_0$ matches and outperforms the other approaches because the Oracle provides more accurate feature embeddings throughout the incremental tasks, whereas Herding and RMM use the IL-trained model.

Anomaly Detection. We study ADCM$_0$ for data collected on a digitisation station in our lab. The setup had 7 RealSense cameras mounted at different locations. 16 similar metal plate objects were digitised and annotated (8 good, 8 anomalous), with fine scratches or marks serving as anomalies. Using the process outlined in Algorithm 1, we identify Camera 3 as the optimal configuration, given the

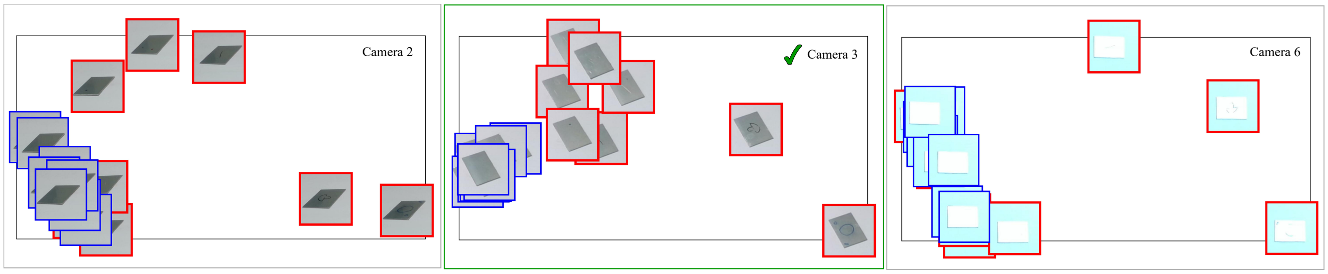

Figure 14. A comparison of the data collected from different perspectives in a challenging real-world environment. We use the approach outlined in Algorithm 1 to accurately identify the optimal perspective (Camera 3). The images show a 2D downscaled representation of the embeddings for good and anomalous objects. The data collected for this experiment will be made available.

data distribution, lighting conditions and available labels. Figure 14 shows the distribution from different setups.

5. Ablation Study and Comparison

Pretrained Oracle: We limit the focus of this paper to smaller model architectures (equivalent to ResNet50) as an Oracle to accommodate diverse deployment scenarios (e.g. on-device learning, resource-constrained systems). Additionally, the pretraining data is limited to ImageNet-1K for ADCM$_0$. Recently, several works have proposed larger model architectures [3, 19, 43, 51] with more pretraining data [6, 51]; and generally report better results on downstream tasks. Table 2 gives a comparison of different Oracles for coreset selection using k-means clustering. A ResNet18 model was trained from scratch on the sampled data to study the efficacy of each Oracle. Larger models or greater pretraining data do not offer significant improvement over ADCM$_0$. In general, we note that the pretraining approach has a greater impact on the quality of feature distribution produced by the Oracle than the model architecture or size. For instance, ResNet50 pretrained using DINO on ImageNet-1K also shows shape and texture sensitivity as shown in Figure 3 for DeIT-S.

Computational Footprint: K-means clustering tends to be the computationally intensive factor, especially when implemented in higher dimensions for ADCM. For our implementation, we use k-means for 100 iterations on downsampled data, with a computational complexity of $O(100 * k * n * d + min(p^2 n, p^3))$, k being the number of clusters (and exemplars), n is the number of data points, d is the number of dimensions (32), and p as the original feature dimension (384 for DeiT-S). Our approach was comparable to Herding in terms of the time required to process the data and compute requirements.

6. Conclusion

We introduced a general framework for Active Data Collection and Management for real-world IL. The approach lends itself to data collection and pruning, coreset/exemplar selection and analysis. The pretrained Oracle provides ground truth feature distribution for the data, using which, we decouple the incremental training from the representation learning task. A simplified baseline established in this paper- ADCM$_0$ (DeIT-S Pretrained using DINO as the Oracle, using k-means for sampling) matches the state-of-the-art performance w.r.t. exemplar selection for CIL, and proves to be effective for general corset selection. The conjecture was that optimising the feature distribution from the Oracle embeddings would also lead to better *meaningful information gain* and translate to better learning and performance. We demonstrated the efficacy of this approach for controlled industrial cases, such as fine-grained categorisation, anomaly detection, camera setup, and general data management. ADCM can be extended and modified to include other Oracle setups and sampling strategies. Our work encourages further research and adoption of IL.

Pretraining Method	Oracle	#Param (M)	#GMAC	Stanford Cars	MVIP	InVar -100	DIMO
Random	-	-	-	55.0	35.1	48.7	27.0
Supervised	RN50	25.6	4.1	51.3	45.2	52.3	34.3
BT [71]	RN50	25.6	4.1	58.1	46.9	54.4	36.8
VICRegL [6]	RN50	23.5	4.1	52.9	48.9	**61.8**	38.9
VICRegL [6]	CXL*	23.5	4.1	57.5	48.5	58.32	39.0
DINO [12]	DeiT-S	21.7	4.2	**61.6**	48.6	60.2	39.7
DINOV2 [51]	DeiT-S	21.7	4.2	61.2	47.9	60.5	**40.5**
SwAV [11]	RN50w5*	586.5	99.9	58.08	**49.0**	60.5	38.1
VICReg [5]	RN200x2*	250.1	59.9	57.41	43.1	52.6	36.4
MoCo v3 [16]	ViT-B*	85.7	16.9	59.23	44.6	58.8	35.2
Full Dataset	-	-	-	88.1	89.4	90.8	94.2

Table 2. Comparison of Top-1 accuracy (%) results obtained by selected samples using different pretrained Oracles (20 images per class). ResNet18 model was trained from scratch on the coreset. (* denotes a larger encoder)

Acknowledgments

This work is funded by the German Federal Ministry of Education and Research (BMBF) and the German Aerospace Center (DLR) under the KIKERP project 0118S23055C in the KI4KMU program.

References

[1] Shubhra Aich, Jesus Ruiz-Santaquiteria, Zhenyu Lu, Prachi Garg, K J Joseph, Alvaro Fernandez Garcia, Vineeth N Balasubramanian, Kenrick Kin, Chengde Wan, Necati Cihan Camgoz, Shugao Ma, and Fernando De la Torre. Data-free class-incremental hand gesture recognition. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 20958–20967, 2023. 2

[2] Ann-Louise Andersen, Kjeld Nielsen, and Thomas Ditlev Brunoe. Prerequisites and barriers for the development of reconfigurable manufacturing systems for high speed ramp-up. *Procedia Cirp*, 51:7–12, 2016. 2

[3] Mahmoud Assran, Quentin Duval, Ishan Misra, Piotr Bojanowski, Pascal Vincent, Michael Rabbat, Yann LeCun, and Nicolas Ballas. Self-supervised learning from images with a joint-embedding predictive architecture. *arXiv preprint arXiv:2301.08243*, 2023. 8

[4] Chunguang Bai and Joseph Sarkis. Flexibility in reverse logistics: a framework and evaluation approach. *Journal of Cleaner Production*, 47:306–318, 2013. 1, 2

[5] Adrien Bardes, Jean Ponce, and Yann LeCun. Vicreg: Variance-invariance-covariance regularization for self-supervised learning. *CoRR*, abs/2105.04906, 2021. 3, 8

[6] Adrien Bardes, Jean Ponce, and Yann LeCun. Vicregl: Self-supervised learning of local visual features. In *NeurIPS*, 2022. 3, 8

[7] Eden Belouadah, Adrian Popescu, Umang Aggarwal, and Léo Saci. Active class incremental learning for imbalanced datasets. In *Computer Vision–ECCV 2020 Workshops: Glasgow, UK, August 23–28, 2020, Proceedings, Part VI*, pages 146–162. Springer, 2021. 2

[8] J. Bengar, J. van de Weijer, L. Fuentes, and B. Raducanu. Class-balanced active learning for image classification. In *2022 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pages 3707–3716, Los Alamitos, CA, USA, 2022. IEEE Computer Society. 2

[9] Paul Bergmann, Michael Fauser, David Sattlegger, and Carsten Steger. Mvtec ad — a comprehensive real-world dataset for unsupervised anomaly detection. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 9584–9592, 2019. 2, 4

[10] Lucas Caccia and Joelle Pineau. Special: Self-supervised pretraining for continual learning. *CoRR*, abs/2106.09065, 2021. 2

[11] Mathilde Caron, Ishan Misra, Julien Mairal, Priya Goyal, Piotr Bojanowski, and Armand Joulin. Unsupervised learning of visual features by contrasting cluster assignments. 2020. 3, 8

[12] 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, 4, 8

[13] Junyi Chai, Hao Zeng, Anming Li, and Eric W.T. Ngai. Deep learning in computer vision: A critical review of emerging techniques and application scenarios. *Machine Learning with Applications*, 6:100134, 2021. 1

[14] Vivek Chavan, Paul Koch, Marian Schlüter, and Clemens Briese. Towards realistic evaluation of industrial continual learning scenarios with an emphasis on energy consumption and computational footprint. In *Proceedings of the International Conference on Computer Vision (ICCV)*, 2023. 2, 3, 5

[15] Toly Chen, Yi-Chi Wang, and Horng-Ren Tsai. Lot cycle time prediction in a ramping-up semiconductor manufacturing factory with a som–fbpn-ensemble approach with multiple buckets and partial normalization. *International Journal of Advanced Manufacturing Technology*, 42, 2009. 2

[16] X. Chen, S. Xie, and K. He. An empirical study of training self-supervised vision transformers. In *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 9620–9629, Los Alamitos, CA, USA, 2021. IEEE Computer Society. 3, 8

[17] Zhiyuan Chen, Bing Liu, Ronald Brachman, Peter Stone, and Francesca Rossi. *Lifelong Machine Learning*. Morgan & Claypool Publishers, 2nd edition, 2018. 2

[18] Ali Cheraghian, Shafin Rahman, Pengfei Fang, Soumava Kumar Roy, Lars Petersson, and Mehrtash Harandi. Semantic-aware knowledge distillation for few-shot class-incremental learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2534–2543, 2021. 2

[19] Mostafa Dehghani, Josip Djolonga, Basil Mustafa, Piotr Padlewski, Jonathan Heek, Justin Gilmer, Andreas Steiner, Mathilde Caron, Robert Geirhos, Ibrahim Alabdulmohsin, Rodolphe Jenatton, Lucas Beyer, Michael Tschannen, Anurag Arnab, Xiao Wang, Carlos Riquelme, Matthias Minderer, Joan Puigcerver, Utku Evci, Manoj Kumar, Sjoerd van Steenkiste, Gamaleldin F. Elsayed, Aravindh Mahendran, Fisher Yu, Avital Oliver, Fantine Huot, Jasmijn Bastings, Mark Patrick Collier, Alexey Gritsenko, Vighnesh Birodkar, Cristina Vasconcelos, Yi Tay, Thomas Mensink, Alexander Kolesnikov, Filip Pavetić, Dustin Tran, Thomas Kipf, Mario Lučić, Xiaohua Zhai, Daniel Keysers, Jeremiah Harmsen, and Neil Houlsby. Scaling vision transformers to 22 billion parameters, 2023. 8

[20] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pages 248–255, 2009. 1

[21] Uwe Dombrowski, Jonas Wullbrandt, and Philipp Krenkel. ”industrie 4.0 in production ramp-up management”. *Procedia Manufacturing*, 17:1015–1022, 2018. 28th International Conference on Flexible Automation and Intelligent Manufacturing (FAIM2018), June 11-14, 2018, Columbus, OH, USAGlobal Integration of Intelligent Manufacturing and Smart Industry for Good of Humanity. 1

[22] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, Jakob Uszkoreit, and Neil Houlsby. An image is worth 16x16 words: Transformers for image recognition at scale. In *International Conference on Learning Representations*, 2021. 1, 2, 3

[23] Arthur Douillard, Matthieu Cord, Charles Ollion, Thomas Robert, and Eduardo Valle. Podnet: Pooled outputs distillation for small-tasks incremental learning. In *Computer Vision - ECCV 2020 - 16th European Conference, Glasgow, UK, August 23-28, 2020, Proceedings, Part XX*, pages 86–102. Springer, 2020. 2, 6

[24] Arthur Douillard, Alexandre Ramé, Guillaume Couairon, and Matthieu Cord. Dytox: Transformers for continual learning with dynamic token expansion. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022. 2

[25] Enrico Fini, Victor G Turrisi da Costa, Xavier Alameda-Pineda, Elisa Ricci, Karteek Alahari, and Julien Mairal. Self-supervised models are continual learners. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022. 2

[26] Robert French. Catastrophic interference in connectionist networks: Can it be predicted, can it be prevented? In *Advances in Neural Information Processing Systems*. Morgan-Kaufmann, 1993. 2

[27] Jhair Gallardo, Tyler L. Hayes, and Christopher Kanan. Self-supervised training enhances online continual learning, 2021. 2

[28] Y Ghunaim, A Bibi, K Alhamoud, M Alfarra, HAAK Hammoud, A Prabhu, PHS Torr, and B Ghanem. Real-time evaluation in online continual learning: a new hope. 2023. 2

[29] Ian J. Goodfellow, Mehdi Mirza, Xia Da, Aaron C. Courville, and Yoshua Bengio. An empirical investigation of catastrophic forgeting in gradient-based neural networks. In *2nd International Conference on Learning Representations, ICLR 2014, Banff, AB, Canada, April 14-16, 2014, Conference Track Proceedings*, 2014. 2

[30] Martin Haller, Andreas Peikert, and Josef Thoma. Cycle time management during production ramp-up. *Robotics and Computer-Integrated Manufacturing*, 19(1-2):183–188, 2003. 1

[31] Md Yousuf Harun, Jhair Gallardo, Tyler L. Hayes, and Christopher Kanan. How efficient are today's continual learning algorithms?, 2023. 2

[32] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016. 1, 2, 3

[33] Dapeng Hu, Qizhengqiu Lu, Lanqing Hong, Hailin Hu, Yifan Zhang, Zhenguo Li, Alfred Shen, and Jiashi Feng. How well self-supervised pre-training performs with streaming data? *CoRR*, abs/2104.12081, 2021. 2

[34] Badr Youbi Idrissi, Diane Bouchacourt, Randall Balestriero, Ivan Evtimov, Caner Hazirbas, Nicolas Ballas, Pascal Vincent, Michal Drozdzal, David Lopez-Paz, and Mark Ibrahim. Imagenet-x: Understanding model mistakes with factor of variation annotations, 2022. 1, 3

[35] Paul Koch, Marian Schlüter, Clemens Briese, and Vivek Chavan. Mvip: A dataset for industrial part recognition, 2023. Fraunhofer Fordatis. 2, 6

[36] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems*. Curran Associates, Inc., 2012. 1, 2

[37] Yann Lecun and Y. Bengio. Convolutional networks for images, speech, and time-series. 1995. 2

[38] Yang Liu and Luiz G. Hafemann. A scale-invariant trajectory simplification method for efficient data collection in videos. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, pages 5129–5138, 2023. 2

[39] Yaoyao Liu, Bernt Schiele, and Qianru Sun. Adaptive aggregation networks for class-incremental learning. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2544–2553, 2020. 2

[40] Yaoyao Liu, Bernt Schiele, and Qianru Sun. Meta-aggregating networks for class-incremental learning. *CoRR*, abs/2010.05063, 2020. 2, 6

[41] Yaoyao Liu, Yuting Su, An-An Liu, Bernt Schiele, and Qianru Sun. Mnemonics training: Multi-class incremental learning without forgetting. In *Proceedings of the IEEE/CVF conference on Computer Vision and Pattern Recognition*, pages 12245–12254, 2020. 2

[42] Yaoyao Liu, Bernt Schiele, and Qianru Sun. Rmm: Reinforced memory management for class-incremental learning. In *Advances in Neural Information Processing Systems*, pages 3478–3490. Curran Associates, Inc., 2021. 2, 6

[43] Zhuang Liu, Hanzi Mao, Chao Yuan Wu, Christoph Feichtenhofer, Trevor Darrell, and Saining Xie. A convnet for the 2020s. In *Proceedings - 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2022*, pages 11966–11976. IEEE Computer Society, 2022. Funding Information: Acknowledgments. We thank Kaiming He, Eric Mintun, Xingyi Zhou, Ross Girshick, and Yann LeCun for valuable discussions and feedback. This work was supported in part by DoD including DARPA's XAI, LwLL, and/or SemaFor programs, as well as BAIR's industrial alliance programs. Publisher Copyright: © 2022 IEEE.; 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2022 ; Conference date: 19-06-2022 Through 24-06-2022. 8

[44] Zilin Luo, Yaoyao Liu, Bernt Schiele, and Qianru Sun. Class-incremental exemplar compression for class-incremental learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 11371–11380, 2023. 2

[45] Rafid Mahmood, James Lucas, Jose M. Alvarez, Sanja Fidler, and Marc Law. Optimizing data collection for machine learning. In *Advances in Neural Information Processing Systems*, pages 29915–29928. Curran Associates, Inc., 2022. 2

[46] Marc Masana, Xialei Liu, Bartłomiej Twardowski, Mikel Menta, Andrew D Bagdanov, and Joost van de Weijer. Class-incremental learning: survey and performance evaluation on image classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022. 6

[47] Mark Mazumder, Colby Banbury, Xiaozhe Yao, Bojan Karlaš, William Gaviria Rojas, Sudnya Diamos, Greg Diamos, Lynn He, Douwe Kiela, David Jurado, David Kanter, Rafael Mosquera, Juan Ciro, Lora Aroyo, Bilge Acun,

Sabri Eyuboglu, Amirata Ghorbani, Emmett Goodman, Tariq Kane, Christine R. Kirkpatrick, Tzu-Sheng Kuo, Jonas Mueller, Tristan Thrush, Joaquin Vanschoren, Margaret Warren, Adina Williams, Serena Yeung, Newsha Ardalani, Praveen Paritosh, Ce Zhang, James Zou, Carole-Jean Wu, Cody Coleman, Andrew Ng, Peter Mattson, and Vijay Janapa Reddi. Dataperf: Benchmarks for data-centric ai development, 2022. 1

[48] Daniele Mazzei and Reshawn Ramjattan. Machine learning for industry 4.0: A systematic review using deep learning-based topic modelling. *Sensors*, 22(22), 2022. 1

[49] Baharan Mirzasoleiman, Jeff A. Bilmes, and Jure Leskovec. Data sketching for faster training of machine learning models. *CoRR*, abs/1906.01827, 2019. 2, 3

[50] Baharan Mirzasoleiman, Jeff Bilmes, and Jure Leskovec. Coresets for data-efficient training of machine learning models. In *Proceedings of the 37th International Conference on Machine Learning*. JMLR.org, 2020. 2

[51] Maxime Oquab, Timothée Darcet, Théo Moutakanni, Huy Vo, Marc Szafraniec, Vasil Khalidov, Pierre Fernandez, Daniel Haziza, Francisco Massa, Alaaeldin El-Nouby, Mahmoud Assran, Nicolas Ballas, Wojciech Galuba, Russell Howes, Po-Yao Huang, Shang-Wen Li, Ishan Misra, Michael Rabbat, Vasu Sharma, Gabriel Synnaeve, Hu Xu, Hervé Jegou, Julien Mairal, Patrick Labatut, Armand Joulin, and Piotr Bojanowski. Dinov2: Learning robust visual features without supervision, 2023. 8

[52] Xiaqing Pan, Nicholas Charron, Yongqian Yang, Scott Peters, Thomas Whelan, Chen Kong, Omkar Parkhi, Richard Newcombe, and Yuheng (Carl) Ren. Aria digital twin: A new benchmark dataset for egocentric 3d machine perception. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 20133–20143, 2023. 1

[53] A Prabhu, HAAK Hammoud, P Dokania, P Torr, SN Lim, B Ghanem, and A Bibi. Computationally budgeted continual learning: what does matter? IEEE, 2023. 2

[54] Hiranmayi Ranganathan, Hemanth Venkateswara, Shayok Chakraborty, and Sethuraman Panchanathan. Deep active learning for image classification. In *2017 IEEE International Conference on Image Processing (ICIP)*, pages 3934–3938, 2017. 2

[55] Sylvestre-Alvise Rebuffi, Alexander Kolesnikov, Georg Sperl, and Christoph H Lampert. icarl: Incremental classifier and representation learning. In *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, pages 2001–2010, 2017. 2, 3, 6

[56] Peter De Roovere, Steven Moonen, Nick Michiels, and Francis Wyffels. Dataset of industrial metal objects, 2022. 2

[57] Marian Schlüter, Hannah Lickert, Katharina Schweitzer, Pinar Bilge, Clemens Briese, Franz Dietrich, and Jörg Krüger. Ai-enhanced identification, inspection and sorting for reverse logistics in remanufacturing. *Procedia CIRP*, 98: 300–305, 2021. The 28th CIRP Conference on Life Cycle Engineering, March 10 – 12, 2021, Jaipur, India. 1, 2

[58] Marian Schlüter, Robert Schimanek, Paul Koch, Clemens Briese, Vivek Chavan, Pinar Bilge, Franz Dietrich, and Jörg Krüger. Green incremental learning - energy efficient ramp-up for ai-enhanced part recognition in reverse logistics. *Procedia CIRP*, 116:414–419, 2023. 30th CIRP Life Cycle Engineering Conference. 2

[59] Christoph Schuhmann, Romain Beaumont, Richard Vencu, Cade Gordon, Ross Wightman, Mehdi Cherti, Theo Coombes, Aarush Katta, Clayton Mullis, Mitchell Wortsman, Patrick Schramowski, Srivatsa Kundurthy, Katherine Crowson, Ludwig Schmidt, Robert Kaczmarczyk, and Jenia Jitsev. Laion-5b: An open large-scale dataset for training next generation image-text models, 2022. 1

[60] Ozan Sener and Silvio Savarese. Active learning for convolutional neural networks: A core-set approach. In *International Conference on Learning Representations*, 2018. 2, 3

[61] Burr Settles. Active learning literature survey. Computer Sciences Technical Report 1648, University of Wisconsin–Madison, 2009. 2

[62] Ravid Shwartz-Ziv, Randall Balestriero, and Yann LeCun. What do we maximize in self-supervised learning?, 2022. 3

[63] Ben Sorscher, Robert Geirhos, Shashank Shekhar, Surya Ganguli, and Ari S. Morcos. Beyond neural scaling laws: beating power law scaling via data pruning, 2022. 2

[64] Xiaoyu Tao, Xiaopeng Hong, Xinyuan Chang, Songlin Dong, Xing Wei, and Yihong Gong. Few-shot class-incremental learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 12183–12192, 2020. 2

[65] Hugo Touvron, Matthieu Cord, Matthijs Douze, Francisco Massa, Alexandre Sablayrolles, and Herve Jegou. Training data-efficient image transformers & distillation through attention. In *Proceedings of the 38th International Conference on Machine Learning*, pages 10347–10357. PMLR, 2021. 3

[66] Gido M van de Ven, Tinne Tuytelaars, and Andreas S Tolias. Three types of incremental learning. *Nature Machine Intelligence*, 4:1185–1197, 2022. 2, 3

[67] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Ł ukasz Kaiser, and Illia Polosukhin. Attention is all you need. In *Advances in Neural Information Processing Systems*. Curran Associates, Inc., 2017. 2

[68] Fu-Yun Wang, Da-Wei Zhou, Han-Jia Ye, and De-Chuan Zhan. Foster: Feature boosting and compression for class-incremental learning. In *Computer Vision–ECCV 2022: 17th European Conference, Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part XXV*, pages 398–414. Springer, 2022. 2, 6

[69] S. Yan, J. Xie, and X. He. Der: Dynamically expandable representation for class incremental learning. In *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3013–3022, Los Alamitos, CA, USA, 2021. IEEE Computer Society. 2, 6

[70] Yu Yang, Hao Kang, and Baharan Mirzasoleiman. Towards sustainable learning: Coresets for data-efficient deep learning. In *In Proceedings of the 40th International Conference on Machine Learning*, 2023. 2

[71] Jure Zbontar, Li Jing, Ishan Misra, Yann LeCun, and Stéphane Deny. Barlow twins: Self-supervised learning via redundancy reduction. In *International Conference on Machine Learning*, 2021. 3, 8

[72] Da-Wei Zhou, Fu-Yun Wang, Han-Jia Ye, and De-Chuan Zhan. Pycil: A python toolbox for class-incremental learning, 2021. 6

[73] Da-Wei Zhou, Fu-Yun Wang, Han-Jia Ye, Liang Ma, Shiliang Pu, and De-Chuan Zhan. Forward compatible few-shot class-incremental learning. In *CVPR*, 2022. 2

[74] Da-Wei Zhou, Qi-Wei Wang, Zhi-Hong Qi, Han-Jia Ye, De-Chuan Zhan, and Ziwei Liu. Deep class-incremental learning: A survey, 2023. 2

[75] Da-Wei Zhou, Qi-Wei Wang, Han-Jia Ye, and De-Chuan Zhan. A model or 603 exemplars: Towards memory-efficient class-incremental learning. In *ICLR*, 2023. 2

[76] Xiao Zhou, Renjie Pi, Weizhong Zhang, Yong Lin, Zonghao Chen, and Tong Zhang. Probabilistic bilevel coreset selection. In *Proceedings of the 39th International Conference on Machine Learning*, pages 27287–27302. PMLR, 2022. 2