First Session Adaptation: A Strong Replay-Free Baseline for Class-Incremental Learning: Supplement

Aristeidis Panos
University of Cambridge
ap2313@cam.ac.uk

Yuriko Kobe
University of Cambridge
yk384@cam.ac.uk

Daniel Olmeda Reino
Toyota Motor Europe
daniel.olmeda.reino@toyota-europe.com

Rahaf Aljundi
Toyota Motor Europe
rahaf.al.jundi@toyota-europe.com

Richard E. Turner
University of Cambridge
ret26@cam.ac.uk

I. Datasets

The exact numbers of training samples and classes for each dataset used in the experiments of Section 4 in the main paper are given in Table 1. For datasets with more than 120k training instances in VTAB+, due to hardware limitations, we randomly sample 120k images and the associated labels and we consider this subset as the full training set. For instance, when we use dSprites-location dataset in a 50-shot setting, we first randomly sample 120k examples and then we randomly pick 50 images for each one of the 16 classes. For DomainNet and iNaturalist we apply a different procedure (see Section II for details).

For evaluation, we consider the protocol used in [30] on the 19 datasets in VTAB, where a balanced dataset of 2k images is created by randomly sampling images from the full test dataset. For FGVC-Aircraft, Cars, and Letters the full test dataset is utilized. We also use the full test dataset for evaluation in the high-shot Class-Incremental Learning (CIL) setting.

II. Dataset Information for the Class-Incremental Learning Experiments

In Table 2, we present detailed information about the Class-Incremental Learning Experiments (CIL) experiments that are in Section 4.3 in the main paper, such as the number of total sessions, number of train instances, and classes per session. Next, we discuss the exact setup for DomainNet and iNaturalist.

DomainNet & iNaturalist. DomainNet and iNaturalist are the only datasets for which we follow a different pre-processing procedure from the one described in Section I in order to create a few-shot CIL scenario similar to the ones considered in the literature. This is due to the large number of classes (iNaturalist has $10,000$ classes) and different domains (DomainNet includes images from 6 domains) these datasets have.

DomainNet is a large-scale dataset of ~ 0.6M images lying in 6 different domains (clipart, infograph, painting, quickdraw, real, sketch) and categorized into 365 distinct classes. These classes can be grouped into 24 superclasses: furniture, mammal, tool, cloth, electricity, building, office, human body, road transportation, food, nature, cold-blooded, music, fruit, sport, tree, bird, vegetable, shape, kitchen, Water transportation, sky transportation, insects, and others. In our CIL experiments, we use 60 classes from the superclasses with an adequate number of instances (> 150): furniture, mammal, tool, cloth, electricity, and road transportation. To the best of our knowledge, this is the first time such a dataset is considered for CIL problems. Table 3 summarizes the DomainNet classes we use for the CIL experiments. To build the 50-shot CIL setting of Section 4.3, we randomly sample 50 images per class and the rest of the images are used for evaluation.

The iNaturalist is another large-scale dataset. comprising ~ 2.7 million images of $10,000$ species. The species can be divided into 10 general categories: amphibians, animalia, arachnids, birds, fungi, insects, mammals, mollusks, plants, and reptiles. Due to the dataset's large size, we have opted to use the "mini" version of the training dataset[1] which has 50 images per class, and thus, this is the only dataset from the few-shot CIL experiments that we do not repeat for 5 times

[1] We use the data from the 2021 competition, available at https://github.com/visipedia/inat_comp/tree/master/2021.

Datasets	# Classes	# Train instances	ALL	CIL
Caltech101 [6]	102	$3,060$	✓	✗
CIFAR100 [14]	100	$50,000$	✓	✓
Flowers102 [20]	102	$1,020$	✓	✗
Pets [21]	37	$3,680$	✓	✗
Sun397 [28]	397	$76,127$	✓	✗
SVHN [19]	10	$73,257$	✓	✓
DTD [4]	47	$1,880$	✓	✗
EuroSAT [8]	10	$27,000$	✓	✗
Resics45 [3]	45	$31,500$	✓	✗
Patch Camelyon [26]	2	$262,144$	✓	✗
Retinopathy [11]	5	$35,126$	✓	✗
CLEVR-count [10]	8	$70,000$	✓	✗
CLEVR-dist [10]	6	$70,000$	✓	✗
dSprites-loc [18]	16	$737,280$	✓	✓
dSprites-ori [18]	16	$737,280$	✓	✗
SmallNORB-azi [15]	18	$24,300$	✓	✗
SmallNORB-elev [15]	9	$24,300$	✓	✗
DMLab [1]	6	$65,550$	✓	✗
KITTI-dist [7]	4	$6,347$	✓	✗
FGVC-Aircraft [17]	100	$6,667$	✓	✓
Cars [13]	196	$8,144$	✓	✓
Letters [5]	62	$74,107$	✓	✓
DomainNet [23]	60 ($345^\dagger$)	$569,010$	✗	✓
iNaturalist [25]	100 ($10,000^\dagger$)	$500,000$	✗	✓
Core50 [16]	50	$119,894$	✓	✓
CUB200 [27]	200	$11,788$	✗	✓

Table 1. Information concerning all datasets used in the experiments. † denotes the number of classes of the original dataset before they are modified for the continual learning scenarios (see Section II for more details). The first 22 datasets form the VTAB+ collection. We also indicate whether a dataset has been used in the offline experiments in Section 4.2 of the main paper which use all the available training data (ALL). Similarly, we indicate which datasets are considered for the Class-Incremental Learning settings in Section 4.

since the (mini) train dataset is already in a 50-shot setting. For evaluation, we use the validation data with 10 images per class. The number of classes considered for the CIL experiments is reduced from $10,000$ to 100; 10 classes per superclass (10 superclasses/sessions). Specific details are given in Tables 4 and 5.

III. Extra Training Details

Due to the large number of experiments and datasets we tried to keep the hyperparameter tuning to a minimum by choosing a set of hyperparameters that works fairly well across all datasets and settings. We have not used any data augmentation in our experiments and all images have been scaled to 224×224 pixels. The only exception is the experiments on CIFAR100, and CUB200 under the few-shot+ setting. There, for comparability reasons, we followed the exact experimental settings as in [31] where standard data augmentation techniques (e.g. random flips and crops) were utilized. Moreover, when we used ResNet-20 for CIFAR100 we maintained the original image size (32×32).

Computing Infrastructure Details & Code. All the experiments of Section 4 have been carried out on a Linux machine with a single NVIDIA-A100 (80GB memory) GPU. Our PyTorch-based code will be made available via a public repository after the review period.

Optimization Details. In all experiments, we train the models using a batch size of 256. Apart from GDumb, for the rest of the methods, EfficientNet-B0

CIL setting	Datasets	S	N_1	$\lvert\mathcal{Y}_1\rvert$	N_s	$\lvert\mathcal{Y}_s\rvert$
High-shot	CIFAR100	10	5k	10	5k	10
	SVHN	5	$\sim$ 19k	2	$\sim$ 14k	2
	dSprites-loc	7	24k	4	12k	2
	FGVC-Aircraft	10	667	10	$\sim$ 670	10
	Cars	10	652	15	$\sim$ 830	20
	Letters	11	$\sim$ 11k	12	$\sim$ 5k	5
	Core50	9	$\sim$ 24k	10	$\sim$ 12k	5
Few-shot+	CIFAR100	9	30k	60	25	5
	CUB200	11	3k	100	50	10
Few-shot	CIFAR100	9	1k	20	500	10
	SVHN	5	100	2	100	2
	dSprites-loc	7	200	4	100	2
	FGVC-Aircraft	9	1k	20	500	10
	Cars	9	1484	36	$\sim$ 830	20
	Letters	11	600	12	250	5
	DomainNet	9	600	12	300	6
	iNaturalist	9	1k	20	500	10

Table 2. Detailed CIL settings for the experiments of Section 4.3. We report the total number of sessions (S), the number of train instances (N_1), and the number of classes ($|\mathcal{Y}_1|$) of the first session and the rest of the sessions ($N_s, |\mathcal{Y}_s|, s > 1$).

backbones are optimized with the Adam optimizer [12] while for ResNet architectures we opt for SGD with momentum set to 0.9. For GDumb, we follow [24] and we use SGD with momentum. For FACT [31] and FSA with pre-trained EfficientNet-B0 backbone, we set the initial learning rate to 0.0001 with scheduled decays by a factor of 0.5 every 50 epochs while for FSA-FiLM, we set it to 0.005. We train all full-body adaptation methods for 200 epochs and the FSA-FiLM for 150 epochs (except for the high-shot setting where we use 200 epochs for fair time comparisons). For the few-shot+ CIL scenario, we follow the training setup of [31]. The weights of the pre-trained EfficientNet-B0 have been obtained from `https://github.com/lukemelas/EfficientNet-PyTorch` while for the pre-trained weights of ResNet-18 and ConvNext, we use the following repository `https://github.com/rwightman/pytorch-image-models`.

Competitors. We found empirically that the recommended hyperparameter values (learning rates, cutmix parameters, SGDR schedule) for GDumb in [24] work well in practice and we use these throughout the experiments. Similarly, for FACT, we use the default values $\alpha = 0.5, \gamma = 0.01, V =$ number of new classes in total [31]. For ALICE, following [22], the projection head is a two-layer MLP with a hidden feature size of 2048 and ReLU as the activation function. All the other hyperparameters (scale factor s, margin m, etc.) are set as in [22]. For E-EWC+SDC, a triplet loss [9] is used as in [29] and the final embeddings of 640 dimensions are normalized.

IV. Additional Results

In this section, we provide tables with the exact accuracies for each one of the datasets used in the experiments under different settings. We have run extra experiments on VTAB+ using meta-learned FiLM adapters in the offline setting and we report accuracies. Additionally, we perform a comparison between different backbones in the offline setting: EfficientNet-B0 and ResNet-18. For the high-shot setting, apart from the four datasets utilized in the main paper, we also deploy the methods on SVHN and present accuracies by session. Finally, accuracies at each session for all three CIL settings are provided.

Head Comparison. Here we provide the exact accuracies for each dataset based on Section 4.2 and Figure 1. Tables 6, 7, 8, and 9 give the offline accuracies for the no adaptation (NA) method for 5, 10, 50 shots, and all training data, respectively. Similar information for the FiLM adaptation method (A-FiLM) is given in Tables 10, 11, 12, and 13. Finally, Tables 14, 15, 16, and 17 provide the corresponding accuracies for the full-body adaptation method (A-FB).

Meta-learned FiLM Adapters. We consider experiments in the offline setting with meta-learned FiLM adapters. We use the meta-trained FiLM adapters as presented in [2]. The results for meta-learned FiLM adapters, as well as for no-adaptation (NA), FiLM (fine-tuned) adaptation (FiLM), and full body adaptation methods, are summarized in Tables 21, 22, and 23, for 5, 10, and 50 shots, respectively. We observe that the meta-trained FiLM adapters work better than NA in all cases, but they fail to compete with the fine-tuned FiLM adapters. Notice that as the number of shots increases, the accuracy difference between meta-learned and fine-tuned FiLM adapters also increases.

FiLM Adaptation: EfficientNet-B0 vs ResNet-18. To assess how different backbone architectures affect the performance of the no-adaptation and FiLM adaptation method, we compare ResNet-18 and EfficientNet-B0 (EN) backbones in Tables 18, 19, and 20, for 5, 10, and 50 shots, respectively. All tables demonstrate the superiority of EfficientNet-B0, regardless of the adaptation method. The tables also show that, regardless of backbone architecture and number

Domain	Clipart	Infograph	Painting	Quickdraw	Real	Sketch
Superclass	Furniture	Mammal	Tool	Cloth	Electricity	Road Transportation
Classes	Clipart	Infograph	Painting	Quickdraw	Real	Sketch
	Furniture	Mammal	Tool	Cloth	Electricity	Road Transportation
	Couch (1)	Cat (11)	Anvil (21)	Belt (31)	Calculator (41)	Ambulance (51)
	Fence (2)	Dolphin (12)	Basket (22)	Camouflage (32)	Computer (42)	Bus (52)
	Streetlight (3)	Squirrel (13)	Rifle (23)	Eyeglasses (33)	Fan (43)	Motorbike (53)
	Table (4)	Zebra (14)	Axe (24)	Helmet (34)	Oven (44)	Train (54)
	Toothbrush (5)	Cow (15)	Dumbbell (25)	Necklace (35)	Dishwasher (45)	Bicycle (55)
	Vase (6)	Elephant (16)	Pliers (26)	Rollerskates (36)	Headphones (46)	Car (56)
	Bed (7)	Pig (17)	Saw (27)	Sock (37)	Microwave (47)	Truck (57)
	Fireplace (8)	Tiger (18)	Skateboard (28)	Underwear (38)	Radio (48)	Bulldozer (58)
	Teapot (9)	Dog (19)	Bandage (29)	Bowtie (39)	Stereo (49)	Firetruck (59)
	Lantern (10)	Rabbit (20)	Paint Can (30)	Crown (40)	Toaster (50)	Tractor (60)

Table 3. DomainNet classes (class id in parentheses) used for the few-shot CIL setting.

	Superclass				
Session	**Amphibians** **1**	**Animalia** **2**	**Arachnids** **3**	**Birds** **4**	**Fungi** **5**
Classes	Ascaphus truei	Lumbricus terrestris	Eratigena duellica	Accipiter badius	Herpothallon rubrocinctum
	Bombina orientalis	Sabella spallanzanii	Atypoides riversi	Accipiter cooperii	Chrysothrix candelaris
	Bombina variegata	Serpula columbiana	Aculepeira ceropegia	Accipiter gentilis	Apiosporina morbosa
	Anaxyrus americanus	Spirobranchus cariniferus	Agalenatea redii	Accipiter nisus	Acarospora socialis
	Anaxyrus boreas	Hemiscolopendra marginata	Araneus bicentenarius	Accipiter striatus	Physcia adscendens
	Anaxyrus cognatus	Scolopendra cingulata	Araneus diadematus	Accipiter trivirgatus	Physcia aipolia
	Anaxyrus fowleri	Scolopendra heros	Araneus marmoreus	Aegypius monachus	Physcia millegrana
	Anaxyrus punctatus	Scolopendra polymorpha	Araneus quadratus	Aquila audax	Physcia stellaris
	Anaxyrus quercicus	Scutigera coleoptrata	Araneus trifolium	Aquila chrysaetos	Candelaria concolor
	Anaxyrus speciosus	Ommatoiulus moreleti	Araniella displicata	Aquila heliaca	Cladonia chlorophaea

Table 4. Classes used from iNaturalist to create the few-shot CIL setting. (table continues to Table 5).

of shots, FiLM adaptation provides significant performance benefits.

High-shot CIL: Accuracies per Session. We provide detailed accuracies for each incremental session for all baselines in the high-shot CIL setting. The accuracies for CIFAR100, CORE50, SVHN, dSPrites-loc, FGVC-Aircraft, Cars, and Letters can be found in Tables 24, 25, 26, 27, 28, 29, and 30, respectively. For GDumb, we provide results with a memory buffer of size 1k and 5k.

Few-shot+ CIL: Accuracies per Session. We provide detailed accuracy for each incremental session for all baselines in the few-shot+ CIL setting. The accuracies for CIFAR100 and CUB200 can be found in Tables 31 and 32, respectively.

Few-shot CIL: Accuracies per Session. We provide detailed accuracy (+ error bars) for each incremental session for all baselines in the few-shot CIL setting. The accuracies for CIFAR100, SVHN, dSprites-location, FGVC-Aircraft, Letters, DomainNet, and iNaturalist can be found in Tables 33, 34, 35, 36, 37, 38, and 39, respectively.

FSA-FiLM vs GDumb The trade-off between accuracy and training time for different continual learning methods on CIFAR100 and CORE50 is illustrated in Fig. 1. Several different memory sizes are used for GDumb. FSA-FiLM attains the highest accuracy (and the lowest PPDR) ≈13.5x faster than GDumb with a 5k memory buffer on CIFAR100 while on CORE50, GDumb requires at least a 5K memory buffer to outperform FSA-FiLM and ≈3x more training time than FSA-FiLM. Notice that FACT is unable to perform well under this setting due to the small number of available classes in the first session.

Session	Superclass: Insects 6	Mammals 7	Mollusks 8	Plants 9	Reptiles 10
Classes	Aptera fusca	Antilocapra americana	Ensis leei	Bryum argenteum	Alligator mississippiensis
	Panchlora nivea	Balaenoptera acutorostrata	Clinocardium nuttallii	Rhodobryum ontariense	Caiman crocodilus
	Pycnoscelus surinamensis	Megaptera novaeangliae	Dinocardium robustum	Leucolepis acanthoneura	Crocodylus acutus
	Blatta orientalis	Aepyceros melampus	Tridacna maxima	Plagiomnium cuspidatum	Crocodylus moreletii
	Periplaneta americana	Alcelaphus buselaphus	Donax gouldii	Plagiomnium insigne	Crocodylus niloticus
	Periplaneta australasiae	Antidorcas marsupialis	Donax variabilis	Rhizomnium glabrescens	Crocodylus porosus
	Periplaneta fuliginosa	Bison bison	Dreissena polymorpha	Dicranum scoparium	Sphenodon punctatus
	Pseudomops septentrionalis	Bos taurus	Mya arenaria	Ceratodon purpureus	Acanthocercus atricollis
	Arrhenodes minutus	Boselaphus tragocamelus	Cyrtopleura costata	Leucobryum glaucum	Agama atra
	Agrilus planipennis	Bubalus bubalis	Geukensia demissa	Funaria hygrometrica	Agama picticauda

Table 5. Classes used from iNaturalist to create the few-shot CIL setting. The first part can be found in Table 4.

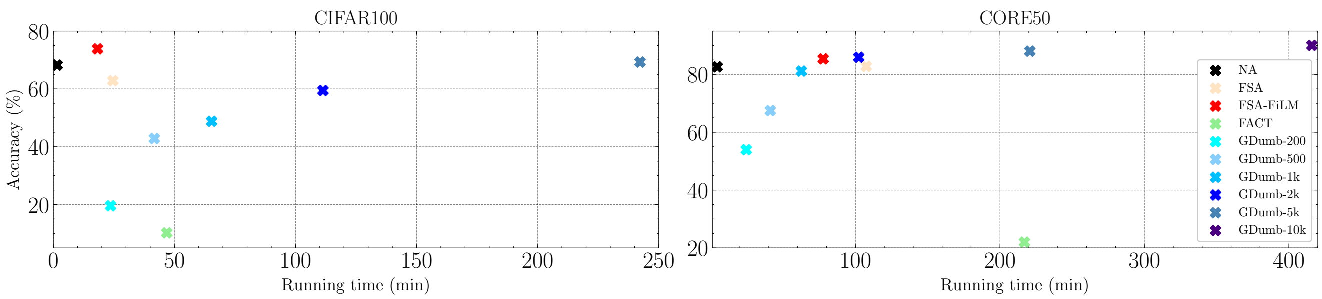

Figure 1. Last session's test accuracy (↑) and run time (↓) for the "high-shot CIL" setting. GDumb-m refers to memory buffer sizes $m \in \{200, 500, 1k, 2k, 5k, 10k^*\}$. We use a memory buffer of 10k images only for CORE50.

Dataset	NCM	LDA	Linear
Caltech101	85.7 ±0.9	**88.2 ±0.8**	87.2 ±0.7
CIFAR100	40.2 ±1.5	**42.7 ±1.6**	42.3 ±1.3
Flowers102	71.5 ±0.6	**76.1 ±0.4**	75.7 ±0.8
Pets	**83.0 ±1.5**	82.4 ±1.6	82.5 ±1.5
Sun397	41.0 ±0.9	41.9 ±1.0	**42.9 ±0.7**
SVHN	13.5 ±1.6	**16.5 ±1.1**	15.0 ±1.4
DTD	48.2 ±1.0	48.9 ±1.7	**49.2 ±1.8**
EuroSAT	72.1 ±1.9	**76.3 ±1.8**	74.6 ±2.0
Resics45	55.2 ±1.2	58.8 ±1.4	**58.9 ±1.2**
Patch Camelyon	59.6 ±8.7	**59.8 ±7.2**	59.0 ±6.4
Retinopathy	**26.3 ±2.5**	25.6 ±1.6	24.3 ±1.7
CLEVR-count	22.2 ±0.8	**23.1 ±1.1**	22.6 ±0.7
CLEVR-dist	23.0 ±2.7	**24.5 ±2.3**	24.4 ±2.2
dSprites-loc	7.5 ±0.7	**8.5 ±0.6**	7.3 ±0.6
dSprites-ori	13.1 ±1.2	**16.2 ±0.8**	14.8 ±1.1
SmallNORB-azi	6.8 ±0.6	**9.3 ±0.8**	8.8 ±1.0
SmallNORB-elev	13.0 ±1.5	**15.1 ±0.6**	14.5 ±0.9
DMLab	22.2 ±0.8	22.1 ±1.3	**23.8 ±0.8**
KITTI-dist	50.0 ±1.3	51.4 ±2.7	**51.9 ±3.1**
FGVC-Aircraft	19.4 ±0.6	22.1 ±1.0	**22.7 ±0.6**
Cars	20.0 ±0.6	**22.6 ±0.6**	22.0 ±0.7
Letters	28.3 ±1.8	**36.1 ±2.1**	34.0 ±1.3
Average acc	37.4	**39.5**	39.0

Table 6. Accuracy comparison between NCM, LDA, and Linear head without using any adaptation (**NA** method). The reported results are based on **5** shots and averaged over 5 runs (mean±std). A pre-trained EfficientNet-B0 is used as a backbone in all cases.

Dataset	NCM	LDA	Linear
Caltech101	88.5 ±0.9	**90.0 ±0.8**	89.6 ±0.7
CIFAR100	45.8 ±1.0	**50.1 ±1.5**	49.2 ±1.1
Flowers102	77.2 ±0.2	**83.9 ±0.3**	81.9 ±0.2
Pets	85.8 ±1.0	**86.4 ±0.7**	86.1 ±0.5
Sun397	47.1 ±1.3	49.0 ±1.3	**49.7 ±0.8**
SVHN	15.5 ±2.3	**19.4 ±2.9**	18.1 ±2.6
DTD	53.8 ±0.3	55.6 ±0.6	**56.3 ±0.7**
EuroSAT	76.6 ±1.2	**82.1 ±0.9**	80.2 ±0.9
Resics45	60.7 ±1.3	65.5 ±1.1	**65.9 ±1.2**
Patch Camelyon	63.0 ±5.3	**66.5 ±3.7**	65.3 ±4.7
Retinopathy	**27.5 ±3.2**	27.1 ±2.2	26.1 ±1.8
CLEVR-count	24.0 ±0.4	**25.7 ±0.6**	24.9 ±0.6
CLEVR-dist	24.2 ±1.2	**26.3 ±1.1**	26.3 ±1.4
dSprites-loc	7.5 ±0.5	**8.7 ±0.3**	7.9 ±0.3
dSprites-ori	14.2 ±0.9	**18.2 ±0.7**	16.4 ±1.0
SmallNORB-azi	8.4 ±0.4	9.5 ±1.1	**9.5 ±0.7**
SmallNORB-elev	13.6 ±0.8	**16.5 ±1.1**	16.2 ±0.7
DMLab	25.1 ±1.1	25.7 ±1.2	**27.2 ±1.3**
KITTI-dist	50.1 ±0.8	**52.9 ±1.5**	52.0 ±1.8
FGVC-Aircraft	23.1 ±0.4	**28.5 ±0.4**	27.0 ±0.4
Cars	25.4 ±0.5	30.4 ±0.5	**30.7 ±0.4**
Letters	34.2 ±0.8	**45.6 ±0.8**	45.2 ±1.4
Average acc	40.5	**43.8**	43.3

Table 7. Accuracy comparison between NCM, LDA, and Linear head without using any adaptation (**NA** method). The reported results are based on **10** shots and averaged over 5 runs (mean±std). A pre-trained EfficientNet-B0 is used as a backbone in all cases.

Dataset	**NCM**	**LDA**	**Linear**
Caltech101	90.4 ±0.6	91.9 ±0.5	**93.0 ±0.4**
CIFAR100	52.0 ±0.9	57.4 ±1.0	**60.9 ±1.0**
Flowers102	77.2 ±0.2	**83.9 ±0.3**	81.9 ±0.2
Pets	88.2 ±0.3	89.5 ±0.4	**89.9 ±0.6**
Sun397	52.6 ±1.1	55.9 ±1.0	**58.4 ±0.8**
SVHN	19.3 ±1.7	28.3 ±1.1	**28.9 ±1.1**
DTD	58.5 ±0.0	61.1 ±0.0	**65.4 ±0.2**
EuroSAT	81.7 ±0.6	87.7 ±0.9	**88.2 ±0.7**
Resics45	66.6 ±0.7	73.5 ±0.9	**78.2 ±0.7**
Patch Camelyon	70.9 ±5.4	**76.2 ±1.1**	76.2 ±1.5
Retinopathy	29.2 ±2.2	**32.9 ±2.2**	32.9 ±1.9
CLEVR-count	26.3 ±1.5	30.1 ±1.0	**31.2 ±1.0**
CLEVR-dist	27.8 ±0.8	**32.2 ±1.1**	31.7 ±1.0
dSprites-loc	9.3 ±0.8	**11.9 ±0.4**	9.9 ±0.6
dSprites-ori	14.9 ±0.5	**20.1 ±1.1**	18.7 ±2.2
SmallNORB-azi	9.5 ±0.6	**12.3 ±0.8**	12.1 ±1.1
SmallNORB-elev	15.2 ±1.3	19.1 ±0.7	**20.0 ±0.5**
DMLab	29.1 ±0.3	30.6 ±0.3	**32.3 ±0.7**
KITTI-dist	53.2 ±0.9	**61.4 ±2.3**	60.7 ±3.2
FGVC-Aircraft	30.9 ±0.5	41.0 ±0.7	**46.1 ±0.6**
Cars	33.7 ±0.0	43.3 ±0.0	**46.5 ±0.2**
Letters	43.1 ±1.0	57.6 ±0.8	**65.9 ±1.1**
Average acc	44.5	49.9	**51.3**

Table 8. Accuracy comparison between NCM, LDA, and Linear head without using any adaptation (**NA** method). The reported results are based on **50** shots and averaged over 5 runs (mean±std). A pre-trained EfficientNet-B0 is used as a backbone in all cases.

Dataset	**NCM**	**LDA**	**Linear**
Caltech101	90.4 ±0.6	91.9 ±0.5	**93.0 ±0.4**
CIFAR100	53.5 ±1.6	68.2 ±1.7	**68.3 ±0.8**
Flowers102	77.2 ±0.2	**83.9 ±0.3**	81.9 ±0.3
Pets	88.7 ±0.2	89.9 ±0.2	**90.7 ±0.3**
Sun397	53.9 ±1.1	56.9 ±1.2	**58.4 ±0.6**
SVHN	24.5 ±0.4	36.8 ±0.6	**40.8 ±0.8**
DTD	58.5 ±0.0	61.1 ±0.0	**65.3 ±0.3**
EuroSAT	82.4 ±0.3	88.1 ±0.1	**93.2 ±0.2**
Resics45	67.1 ±0.2	74.6 ±0.2	**81.8 ±0.4**
Patch Camelyon	72.9 ±0.0	79.1 ±0.0	**79.7 ±0.5**
Retinopathy	33.0 ±0.3	40.4 ±0.3	**46.9 ±0.4**
CLEVR-count	28.9 ±0.3	40.1 ±0.4	**50.3 ±0.3**
CLEVR-dist	29.2 ±0.5	38.7 ±0.4	**45.5 ±1.3**
dSprites-loc	14.6 ±0.6	20.9 ±0.7	**30.8 ±1.4**
dSprites-ori	15.5 ±0.4	22.5 ±0.6	**33.3 ±0.9**
SmallNORB-azi	11.5 ±0.5	14.1 ±0.7	**14.2 ±0.7**
SmallNORB-elev	19.3 ±0.4	24.1 ±0.5	**26.3 ±0.8**
DMLab	35.4 ±0.4	39.7 ±0.3	**44.4 ±0.4**
KITTI-dist	53.4 ±0.0	66.7 ±0.0	**69.8 ±0.4**
FGVC-Aircraft	31.8 ±0.0	41.3 ±0.0	**45.8 ±0.4**
Cars	33.7 ±0.0	43.3 ±0.0	**46.6 ±0.3**
Letters	44.9 ±1.3	59.7 ±0.5	**69.5 ±0.5**
Average acc	46.4	53.4	**58.0**

Table 9. Accuracy comparison between NCM, LDA, and Linear head without using any adaptation (**NA** method). The reported results are based on the full training dataset and averaged over 5 runs (mean±std). A pre-trained EfficientNet-B0 is used as a backbone in all cases.

Dataset	**NCM**	**LDA**	**Linear**
Caltech101	86.6 ± 0.5	**89.0** $\pm \mathbf{0.6}$	88.5 ± 0.4
CIFAR100	47.4 ± 1.2	**51.8** $\pm \mathbf{1.3}$	50.3 ± 1.4
Flowers102	80.2 ± 0.4	**85.0** $\pm \mathbf{0.8}$	83.5 ± 0.5
Pets	81.8 ± 1.5	81.8 ± 1.7	**82.6** $\pm \mathbf{1.6}$
Sun397	40.9 ± 0.7	**40.9** $\pm \mathbf{0.7}$	38.1 ± 1.0
SVHN	28.6 ± 3.8	**31.7** $\pm \mathbf{3.7}$	30.1 ± 4.5
DTD	49.6 ± 1.5	50.2 ± 0.9	**50.8** $\pm \mathbf{1.3}$
EuroSAT	75.8 ± 1.5	78.1 ± 1.2	**78.3** $\pm \mathbf{1.2}$
Resics45	62.7 ± 1.1	64.7 ± 1.2	**65.8** $\pm \mathbf{0.7}$
Patch Camelyon	64.7 ± 5.8	**64.9** $\pm \mathbf{6.6}$	62.9 ± 5.4
Retinopathy	**27.4** $\pm \mathbf{3.1}$	26.0 ± 2.0	25.2 ± 2.4
CLEVR-count	**24.0** $\pm \mathbf{1.3}$	23.4 ± 1.4	23.2 ± 0.7
CLEVR-dist	23.1 ± 1.4	23.1 ± 1.1	**24.0** $\pm \mathbf{1.3}$
dSprites-loc	19.5 ± 1.8	**19.8** $\pm \mathbf{2.0}$	16.7 ± 5.9
dSprites-ori	20.6 ± 1.6	**26.5** $\pm \mathbf{0.8}$	25.5 ± 1.4
SmallNORB-azi	9.0 ± 1.0	10.1 ± 0.6	**10.3** $\pm \mathbf{0.6}$
SmallNORB-elev	14.6 ± 0.9	15.4 ± 0.7	**15.5** $\pm \mathbf{0.8}$
DMLab	23.4 ± 2.3	23.3 ± 1.9	**24.8** $\pm \mathbf{0.9}$
KITTI-dist	**55.7** $\pm \mathbf{3.6}$	52.7 ± 3.5	53.2 ± 1.8
FGVC-Aircraft	28.7 ± 0.7	32.6 ± 0.9	**33.1** $\pm \mathbf{1.3}$
Cars	22.7 ± 0.5	**28.1** $\pm \mathbf{0.4}$	27.2 ± 0.7
Letters	52.2 ± 2.9	55.9 ± 2.5	**56.4** $\pm \mathbf{3.2}$
Average acc	42.7	**44.3**	43.9

Table 10. Accuracy comparison between NCM, LDA, and Linear head using FiLM adaptation (**A-FiLM** method). The reported results are based on **5** shots and averaged over 5 runs (mean±std). A pre-trained EfficientNet-B0 is used as a backbone in all cases.

Dataset	**NCM**	**LDA**	**Linear**
Caltech101	89.9 ± 0.2	**91.5** $\pm \mathbf{0.4}$	91.1 ± 0.6
CIFAR100	59.2 ± 1.7	**62.8** $\pm \mathbf{1.1}$	60.6 ± 0.5
Flowers102	86.2 ± 0.5	91.1 ± 0.3	**91.2** $\pm \mathbf{0.3}$
Pets	85.8 ± 1.0	86.3 ± 0.8	**86.6** $\pm \mathbf{0.8}$
Sun397	48.5 ± 0.8	**49.3** $\pm \mathbf{0.6}$	44.8 ± 1.2
SVHN	40.3 ± 2.5	**45.3** $\pm \mathbf{3.0}$	43.9 ± 3.3
DTD	58.3 ± 0.8	59.1 ± 0.8	**59.2** $\pm \mathbf{1.1}$
EuroSAT	81.9 ± 1.4	**84.4** $\pm \mathbf{1.6}$	83.5 ± 0.7
Resics45	70.2 ± 1.1	73.0 ± 1.3	**73.5** $\pm \mathbf{0.7}$
Patch Camelyon	**69.2** $\pm \mathbf{5.2}$	68.5 ± 6.1	67.1 ± 4.8
Retinopathy	26.7 ± 1.3	**26.9** $\pm \mathbf{1.1}$	25.6 ± 1.0
CLEVR-count	**30.1** $\pm \mathbf{1.9}$	27.6 ± 1.6	29.2 ± 1.2
CLEVR-dist	25.3 ± 1.7	26.2 ± 1.3	**26.5** $\pm \mathbf{0.7}$
dSprites-loc	**26.2** $\pm \mathbf{11.4}$	26.1 ± 10.6	24.2 ± 14.7
dSprites-ori	26.7 ± 2.4	**34.0** $\pm \mathbf{2.6}$	33.8 ± 2.4
SmallNORB-azi	11.0 ± 0.8	**11.7** $\pm \mathbf{1.2}$	11.3 ± 1.4
SmallNORB-elev	15.6 ± 0.7	16.3 ± 0.4	**16.6** $\pm \mathbf{0.3}$
DMLab	27.2 ± 1.9	26.6 ± 1.5	**29.3** $\pm \mathbf{1.7}$
KITTI-dist	**56.7** $\pm \mathbf{3.7}$	55.4 ± 3.7	56.2 ± 4.5
FGVC-Aircraft	37.5 ± 0.7	43.3 ± 1.1	**43.4** $\pm \mathbf{1.5}$
Cars	36.3 ± 0.8	**43.1** $\pm \mathbf{1.0}$	42.1 ± 1.0
Letters	64.0 ± 1.5	67.5 ± 1.3	**68.1** $\pm \mathbf{1.4}$
Average acc	48.8	**50.7**	50.4

Table 11. Accuracy comparison between NCM, LDA, and Linear head using FiLM adaptation (**A-FiLM** method). The reported results are based on **10** shots and averaged over 5 runs (mean±std). A pre-trained EfficientNet-B0 is used as a backbone in all cases.

Dataset	**NCM**	**LDA**	**Linear**
Caltech101	93.4 ±0.7	**93.8 ±0.5**	93.5 ±1.0
CIFAR100	72.6 ±0.7	**73.8 ±0.9**	73.7 ±0.5
Flowers102	86.2 ±0.5	91.1 ±0.3	**91.2 ±0.3**
Pets	89.3 ±0.6	**89.9 ±0.7**	89.9 ±0.7
Sun397	58.5 ±0.7	59.7 ±0.6	**60.8 ±0.7**
SVHN	73.9 ±1.1	**77.2 ±0.8**	76.8 ±1.0
DTD	66.8 ±0.3	68.4 ±0.2	**68.7 ±0.7**
EuroSAT	91.0 ±0.6	93.0 ±0.6	**93.2 ±0.6**
Resics45	81.7 ±0.2	83.4 ±0.6	**85.3 ±0.6**
Patch Camelyon	**78.5 ±2.0**	77.9 ±2.4	78.1 ±2.4
Retinopathy	34.2 ±1.6	**35.2 ±1.2**	33.5 ±1.9
CLEVR-count	**56.8 ±0.9**	46.6 ±1.1	53.1 ±1.1
CLEVR-dist	40.2 ±1.8	38.8 ±1.0	**41.2 ±1.5**
dSprites-loc	83.6 ±5.4	**83.7 ±5.6**	81.6 ±5.9
dSprites-ori	41.2 ±0.8	52.1 ±1.3	**53.8 ±1.4**
SmallNORB-azi	17.0 ±0.8	16.8 ±0.8	**17.5 ±1.0**
SmallNORB-elev	23.1 ±1.2	22.9 ±0.5	**24.0 ±0.6**
DMLab	35.0 ±0.6	34.6 ±0.6	**35.8 ±0.4**
KITTI-dist	**67.3 ±2.1**	66.8 ±3.0	66.5 ±2.7
FGVC-Aircraft	60.6 ±0.9	65.1 ±0.7	**68.0 ±0.6**
Cars	60.9 ±0.2	67.9 ±0.2	**74.5 ±0.4**
Letters	76.7 ±0.5	79.7 ±0.4	**81.9 ±0.8**
Average acc	63.1	64.5	**65.6**

Table 12. Accuracy comparison between NCM, LDA, and Linear head using FiLM adaptation (**A-FiLM** method). The reported results are based on **50** shots and averaged over 5 runs (mean±std). A pre-trained EfficientNet-B0 is used as a backbone in all cases.

Dataset	**NCM**	**LDA**	**Linear**
Caltech101	93.6 ±0.4	**94.4 ±0.3**	94.1 ±0.6
CIFAR100	77.4 ±1.1	78.2 ±1.1	**82.1 ±1.1**
Flowers102	88.6 ±0.6	**91.2 ±0.4**	90.6 ±0.5
Pets	90.2 ±0.2	90.8 ±0.3	**91.0 ±0.4**
Sun397	61.4 ±0.5	62.1 ±0.3	**63.7 ±0.9**
SVHN	92.9 ±0.4	93.1 ±0.4	**95.1 ±0.4**
DTD	64.8 ±0.2	66.8 ±0.5	**67.6 ±0.6**
EuroSAT	96.5 ±0.3	97.2 ±0.1	**98.1 ±0.3**
Resics45	88.0 ±0.3	89.4 ±0.4	**94.2 ±0.4**
Patch Camelyon	84.3 ±1.5	**85.9 ±1.2**	85.8 ±1.6
Retinopathy	52.3 ±0.7	52.6 ±0.8	**59.5 ±0.8**
CLEVR-count	94.5 ±0.2	93.5 ±0.7	**95.3 ±0.6**
CLEVR-dist	79.3 ±1.4	80.1 ±2.3	**84.5 ±2.0**
dSprites-loc	98.0 ±0.5	98.5 ±0.6	**99.3 ±0.4**
dSprites-ori	69.8 ±2.4	80.0 ±0.8	**90.7 ±0.8**
SmallNORB-azi	**26.1 ±1.2**	24.4 ±0.6	23.5 ±0.5
SmallNORB-elev	47.0 ±0.9	**47.9 ±1.2**	47.9 ±1.4
DMLab	60.2 ±0.6	61.3 ±0.5	**65.8 ±1.0**
KITTI-dist	78.5 ±0.9	**80.1 ±1.1**	79.7 ±0.4
FGVC-Aircraft	63.3 ±0.2	67.5 ±0.4	**71.5 ±0.5**
Cars	60.1 ±0.1	67.3 ±0.3	**73.6 ±0.3**
Letters	77.1 ±0.3	81.7 ±0.4	**85.2 ±0.3**
Average acc	74.7	76.6	**79.0**

Table 13. Accuracy comparison between NCM, LDA, and Linear head using FiLM adaptation (**A-FiLM** method). The reported results are based on the full training dataset and averaged over 5 runs (mean±std). A pre-trained EfficientNet-B0 is used as a backbone in all cases.

Dataset	**NCM**	**LDA**	**Linear**
Caltech101	87.1 ±0.6	**89.4 ±0.7**	86.1 ±0.9
CIFAR100	48.1 ±0.7	**49.3 ±1.1**	49.2 ±1.0
Flowers102	81.5 ±0.8	81.7 ±0.6	**83.6 ±0.8**
Pets	**80.9 ±1.7**	80.8 ±1.7	71.1 ±1.2
Sun397	35.5 ±0.3	35.2 ±0.5	**37.2 ±0.3**
SVHN	19.1 ±1.5	**19.6 ±1.2**	19.2 ±1.8
DTD	48.4 ±1.3	**49.0 ±1.3**	41.6 ±0.8
EuroSAT	75.4 ±4.1	76.7 ±2.9	**78.5 ±1.2**
Resics45	61.7 ±2.3	62.3 ±2.5	**63.5 ±2.5**
Patch Camelyon	59.2 ±4.9	59.4 ±4.9	**60.5 ±7.0**
Retinopathy	24.6 ±2.7	24.5 ±2.5	**26.1 ±2.4**
CLEVR-count	23.9 ±2.9	23.5 ±3.0	**24.2 ±3.1**
CLEVR-dist	25.1 ±3.3	25.6 ±3.5	**25.6 ±2.4**
dSprites-loc	26.1 ±2.7	**27.1 ±1.6**	25.5 ±3.6
dSprites-ori	18.3 ±1.6	**19.9 ±1.5**	15.6 ±1.9
SmallNORB-azi	10.0 ±0.7	**10.4 ±0.8**	10.3 ±0.9
SmallNORB-elev	15.8 ±1.1	**16.2 ±1.1**	15.6 ±0.8
DMLab	21.3 ±1.6	22.1 ±1.3	**22.6 ±0.9**
KITTI-dist	51.1 ±2.5	**52.8 ±2.5**	52.1 ±2.0
FGVC-Aircraft	23.8 ±0.7	23.6 ±0.8	**25.1 ±1.0**
Cars	23.1 ±0.3	23.5 ±0.3	**25.3 ±0.5**
Letters	35.5 ±3.3	35.7 ±3.3	**37.0 ±3.2**
Average acc	40.7	**41.3**	40.7

Table 14. Accuracy comparison between NCM, LDA, and Linear head using full-body adaptation (**A-FB** method). The reported results are based on **5** shots and averaged over 5 runs (mean±std). A pre-trained EfficientNet-B0 is used as a backbone in all cases.

Dataset	**NCM**	**LDA**	**Linear**
Caltech101	90.4 ±0.8	**91.7 ±0.7**	89.3 ±0.8
CIFAR100	58.8 ±0.4	**59.5 ±0.3**	59.5 ±0.8
Flowers102	89.9 ±0.6	90.0 ±0.5	**91.3 ±0.6**
Pets	85.5 ±0.9	**85.7 ±0.5**	78.4 ±1.7
Sun397	45.2 ±0.7	45.0 ±0.4	**46.3 ±0.3**
SVHN	26.4 ±3.4	**27.1 ±3.8**	26.4 ±3.6
DTD	55.2 ±0.8	**56.7 ±1.0**	48.0 ±0.8
EuroSAT	83.9 ±1.5	84.9 ±1.5	**86.3 ±1.5**
Resics45	72.6 ±1.2	72.8 ±1.3	**74.3 ±1.1**
Patch Camelyon	61.3 ±4.5	61.2 ±4.5	**63.0 ±5.1**
Retinopathy	25.1 ±4.2	24.6 ±3.3	**27.5 ±2.0**
CLEVR-count	28.0 ±1.9	27.7 ±2.0	**28.7 ±2.1**
CLEVR-dist	**30.1 ±4.1**	30.0 ±4.2	29.9 ±3.6
dSprites-loc	42.9 ±2.7	**44.8 ±1.3**	42.2 ±3.1
dSprites-ori	26.4 ±6.5	**29.6 ±6.8**	22.9 ±10.4
SmallNORB-azi	**12.2 ±0.6**	11.5 ±0.8	11.7 ±0.6
SmallNORB-elev	18.2 ±1.3	**18.3 ±1.5**	17.3 ±1.1
DMLab	25.3 ±1.2	25.6 ±1.3	**26.1 ±1.1**
KITTI-dist	53.5 ±2.3	**54.9 ±1.3**	52.2 ±2.1
FGVC-Aircraft	37.4 ±0.5	36.7 ±0.6	**38.6 ±0.4**
Cars	43.8 ±0.8	43.5 ±0.6	**46.9 ±0.8**
Letters	55.8 ±1.5	55.2 ±1.3	**56.6 ±1.4**
Average acc	48.5	**49.0**	48.3

Table 15. Accuracy comparison between NCM, LDA, and Linear head using full-body adaptation (**A-FB** method). The reported results are based on **10** shots and averaged over 5 runs (mean±std). A pre-trained EfficientNet-B0 is used as a backbone in all cases.

Dataset	NCM	LDA	Linear
Caltech101	92.8 ± 0.3	**93.9** $\pm$**0.3**	92.6 ± 0.5
CIFAR100	72.9 ± 0.9	**73.1** $\pm$**0.7**	72.7 ± 0.8
Flowers102	89.9 ± 0.6	90.0 ± 0.5	**91.3** $\pm$**0.6**
Pets	88.7 ± 0.5	**89.6** $\pm$**0.6**	84.9 ± 1.1
Sun397	59.9 ± 0.6	60.4 ± 0.6	**62.4** $\pm$**0.4**
SVHN	63.9 ± 1.5	64.5 ± 1.3	**64.6** $\pm$**1.6**
DTD	60.9 ± 0.3	**64.4** $\pm$**0.2**	57.4 ± 0.7
EuroSAT	93.8 ± 0.5	**94.1** $\pm$**0.6**	93.9 ± 1.0
Resics45	87.4 ± 0.8	87.6 ± 0.7	**88.0** $\pm$**0.8**
Patch Camelyon	71.6 ± 1.7	72.2 ± 1.8	**74.4** $\pm$**1.8**
Retinopathy	31.1 ± 3.3	31.3 ± 2.9	**31.9** $\pm$**2.2**
CLEVR-count	**46.5** $\pm$**1.0**	45.2 ± 1.1	45.4 ± 1.9
CLEVR-dist	44.0 ± 2.1	44.7 ± 2.2	**45.7** $\pm$**2.2**
dSprites-loc	85.3 ± 4.0	**87.1** $\pm$**2.4**	85.2 ± 3.2
dSprites-ori	42.6 ± 3.2	**44.8** $\pm$**2.3**	39.3 ± 4.7
SmallNORB-azi	**19.6** $\pm$**0.6**	18.3 ± 0.7	19.1 ± 1.1
SmallNORB-elev	**31.4** $\pm$**1.4**	31.3 ± 1.7	31.3 ± 2.1
DMLab	32.6 ± 1.4	32.7 ± 1.6	**34.0** $\pm$**0.9**
KITTI-dist	65.8 ± 2.1	**66.9** $\pm$**2.2**	65.8 ± 1.8
FGVC-Aircraft	74.2 ± 0.8	73.5 ± 0.4	**74.6** $\pm$**0.3**
Cars	79.3 ± 0.1	79.4 ± 0.1	**81.5** $\pm$**0.2**
Letters	82.1 ± 0.7	82.3 ± 0.9	**83.1** $\pm$**0.6**
Average acc	64.4	**64.9**	64.5

Table 16. Accuracy comparison between NCM, LDA, and Linear head using full-body adaptation (**A-FB** method). The reported results are based on **50** shots and averaged over 5 runs (mean±std). A pre-trained EfficientNet-B0 is used as a backbone in all cases.

Dataset	NCM	LDA	Linear
Caltech101	94.2 ± 0.5	94.6 ± 0.3	**94.8** $\pm$**0.3**
CIFAR100	84.2 ± 1.2	84.3 ± 0.9	**85.0** $\pm$**1.1**
Flowers102	90.3 ± 0.3	90.3 ± 0.4	**91.4** $\pm$**0.5**
Pets	89.7 ± 0.4	89.5 ± 0.3	**90.0** $\pm$**0.4**
Sun397	65.9 ± 0.3	66.1 ± 1.0	**66.7** $\pm$**0.4**
SVHN	**95.6** $\pm$**0.2**	95.3 ± 0.5	95.5 ± 0.3
DTD	67.6 ± 0.8	68.1 ± 0.4	**68.5** $\pm$**0.5**
EuroSAT	98.1 ± 0.2	98.5 ± 0.3	**98.6** $\pm$**0.2**
Resics45	95.3 ± 0.1	95.5 ± 0.2	**95.9** $\pm$**0.1**
Patch Camelyon	81.1 ± 2.2	85.0 ± 0.7	**86.5** $\pm$**0.9**
Retinopathy	55.8 ± 0.9	56.1 ± 1.4	**57.7** $\pm$**0.7**
CLEVR-count	98.5 ± 0.3	**98.7** $\pm$**0.3**	98.3 ± 0.3
CLEVR-dist	89.0 ± 0.6	89.0 ± 0.6	**89.4** $\pm$**1.5**
dSprites-loc	99.7 ± 0.3	**99.8** $\pm$**0.1**	99.6 ± 0.4
dSprites-ori	89.2 ± 1.0	**94.0** $\pm$**0.7**	93.0 ± 1.1
SmallNORB-azi	**29.8** $\pm$**1.0**	28.7 ± 0.6	28.9 ± 0.8
SmallNORB-elev	74.3 ± 4.3	**81.8** $\pm$**3.1**	77.2 ± 4.6
DMLab	64.8 ± 0.6	**65.7** $\pm$**0.4**	65.6 ± 0.7
KITTI-dist	78.2 ± 0.7	82.1 ± 0.6	**82.3** $\pm$**1.1**
FGVC-Aircraft	76.0 ± 0.5	75.8 ± 0.7	**76.7** $\pm$**0.8**
Cars	79.1 ± 0.1	78.9 ± 0.2	**81.3** $\pm$**0.2**
Letters	86.0 ± 0.5	85.7 ± 0.5	**87.2** $\pm$**0.3**
Average acc	81.0	82.0	**82.3**

Table 17. Accuracy comparison between NCM, LDA, and Linear head using full-body adaptation (**A-FB** method). The reported results are based on the full training dataset and averaged over 5 runs (mean±std). A pre-trained EfficientNet-B0 is used as a backbone in all cases.

Dataset	**NA(RN)**	**FiLM(RN)**	**NA(EN)**	**FiLM(EN)**
Caltech101	80.9 ± 0.7	81.1 ± 0.7	88.2 ± 0.8	89.0 ± 0.6
CIFAR100	40.4 ± 0.5	42.2 ± 0.9	42.7 ± 1.6	51.8 ± 1.3
Flowers102	72.6 ± 0.8	79.4 ± 1.4	76.1 ± 0.4	85.0 ± 0.8
Pets	76.9 ± 1.2	74.5 ± 1.4	82.4 ± 1.6	81.8 ± 1.7
Sun397	35.1 ± 0.9	29.1 ± 0.7	41.9 ± 1.0	40.9 ± 0.7
SVHN	20.9 ± 1.3	28.8 ± 2.7	16.5 ± 1.1	31.7 ± 3.7
DTD	43.2 ± 1.7	42.2 ± 0.7	48.9 ± 1.7	50.2 ± 0.9
EuroSAT	75.1 ± 1.9	79.3 ± 1.2	76.3 ± 1.8	78.1 ± 1.2
Resics45	56.8 ± 1.3	57.0 ± 1.2	58.8 ± 1.4	64.7 ± 1.2
Patch Camelyon	62.4 ± 5.5	64.6 ± 7.2	59.8 ± 7.2	64.9 ± 6.6
Retinopathy	23.0 ± 2.5	23.6 ± 1.7	25.6 ± 1.6	26.0 ± 2.0
CLEVR-count	21.6 ± 1.7	23.0 ± 1.3	23.1 ± 1.1	23.4 ± 1.4
CLEVR-dist	22.9 ± 1.4	24.4 ± 1.3	24.5 ± 2.3	23.1 ± 1.1
dSprites-loc	13.0 ± 1.0	15.9 ± 1.0	8.5 ± 0.6	19.8 ± 2.0
dSprites-ori	14.4 ± 0.6	22.9 ± 0.8	16.2 ± 0.8	26.5 ± 0.8
SmallNORB-azi	9.4 ± 0.8	9.8 ± 1.1	9.3 ± 0.8	10.1 ± 0.6
SmallNORB-elev	15.8 ± 0.7	15.9 ± 0.7	15.1 ± 0.6	15.4 ± 0.7
DMLab	21.6 ± 1.5	22.1 ± 1.8	22.1 ± 1.3	23.3 ± 1.9
KITTI-dist	54.3 ± 2.8	54.0 ± 2.5	51.4 ± 2.7	52.7 ± 3.5
FGVC-Aircraft	19.1 ± 0.9	20.3 ± 0.7	22.1 ± 1.0	32.6 ± 0.9
Cars	14.8 ± 0.5	13.9 ± 0.3	22.6 ± 0.6	28.1 ± 0.4
Letters	32.4 ± 1.9	45.5 ± 2.4	36.1 ± 2.1	55.9 ± 2.5
Average acc	37.6	39.5	39.5	44.3

Table 18. Accuracy comparison between NA and FiLM methods in offline mode using either a pre-trained ResNet-18 (RN) or a pre-trained EfficientNet-B0 (EN) backbone. We use an LDA head. The reported results are based on **5** shots. Results are averaged over 5 runs (mean$\pm$std).

Dataset	**NA(RN)**	**FiLM(RN)**	**NA(EN)**	**FiLM(EN)**
Caltech101	85.0 ± 0.6	86.6 ± 0.6	90.0 ± 0.8	91.5 ± 0.4
CIFAR100	48.5 ± 0.4	52.2 ± 0.8	50.1 ± 1.5	62.8 ± 1.1
Flowers102	81.2 ± 0.7	87.1 ± 0.2	83.9 ± 0.3	91.1 ± 0.3
Pets	82.3 ± 0.7	80.5 ± 1.1	86.4 ± 0.7	86.3 ± 0.8
Sun397	42.8 ± 0.9	36.3 ± 1.0	49.0 ± 1.3	49.3 ± 0.6
SVHN	24.6 ± 1.7	35.8 ± 4.1	19.4 ± 2.9	45.3 ± 3.0
DTD	51.9 ± 0.9	50.5 ± 0.6	55.6 ± 0.6	59.1 ± 0.8
EuroSAT	82.0 ± 0.7	85.4 ± 0.9	82.1 ± 0.9	84.4 ± 1.6
Resics45	64.8 ± 1.5	67.2 ± 2.0	65.5 ± 1.1	73.0 ± 1.3
Patch Camelyon	66.9 ± 3.7	68.8 ± 3.7	66.5 ± 3.7	68.5 ± 6.1
Retinopathy	25.5 ± 1.3	25.8 ± 3.7	27.1 ± 2.2	26.9 ± 1.1
CLEVR-count	23.8 ± 0.7	25.6 ± 2.0	25.7 ± 0.6	27.6 ± 1.6
CLEVR-dist	24.9 ± 0.7	27.2 ± 1.2	26.3 ± 1.1	26.2 ± 1.3
dSprites-loc	14.7 ± 0.4	25.4 ± 1.5	8.7 ± 0.3	26.1 ± 10.6
dSprites-ori	16.6 ± 0.9	29.7 ± 1.3	18.2 ± 0.7	34.0 ± 2.6
SmallNORB-azi	10.4 ± 0.9	12.2 ± 1.1	9.5 ± 1.1	11.7 ± 1.2
SmallNORB-elev	16.9 ± 0.8	17.2 ± 1.1	16.5 ± 1.1	16.3 ± 0.4
DMLab	24.6 ± 1.8	25.6 ± 1.4	25.7 ± 1.2	26.6 ± 1.5
KITTI-dist	53.0 ± 2.0	55.9 ± 3.5	52.9 ± 1.5	55.4 ± 3.7
FGVC-Aircraft	25.9 ± 0.8	29.5 ± 0.8	28.5 ± 0.4	43.3 ± 1.1
Cars	21.5 ± 0.5	24.0 ± 0.2	30.4 ± 0.5	43.1 ± 1.0
Letters	41.5 ± 1.2	62.7 ± 1.9	45.6 ± 0.8	67.5 ± 1.3
Average acc	42.2	46.0	43.8	50.7

Table 19. Accuracy comparison between NA and FiLM methods in offline mode using either a pre-trained ResNet-18 (RN) or a pre-trained EfficientNet-B0 (EN) backbone. We use an LDA head. The reported results are based on **10** shots. Results are averaged over 5 runs (mean$\pm$std).

Dataset	**NA(RN)**	**FiLM(RN)**	**NA(EN)**	**FiLM(EN)**
Caltech101	88.0 ±0.3	87.7 ±0.7	91.9 ±0.5	93.8 ±0.5
CIFAR100	58.2 ±0.9	61.8 ±0.7	57.4 ±1.0	73.8 ±0.9
Flowers102	81.2 ±0.7	87.1 ±0.2	83.9 ±0.3	91.1 ±0.3
Pets	86.9 ±0.4	86.8 ±0.6	89.5 ±0.4	89.9 ±0.7
Sun397	51.4 ±1.3	47.8 ±0.9	55.9 ±1.0	59.7 ±0.6
SVHN	37.3 ±0.8	74.5 ±0.9	28.3 ±1.1	77.2 ±0.8
DTD	59.9 ±0.0	59.3 ±0.6	61.1 ±0.0	68.4 ±0.2
EuroSAT	88.3 ±0.5	92.6 ±0.5	87.7 ±0.9	93.0 ±0.6
Resics45	73.7 ±1.1	77.5 ±1.1	73.5 ±0.9	83.4 ±0.6
Patch Camelyon	76.3 ±0.9	77.4 ±1.1	76.2 ±1.1	77.9 ±2.4
Retinopathy	29.2 ±2.1	30.4 ±1.0	32.9 ±2.2	35.2 ±1.2
CLEVR-count	29.5 ±1.3	39.9 ±1.4	30.1 ±1.0	46.6 ±1.1
CLEVR-dist	31.7 ±0.8	45.3 ±2.6	32.2 ±1.1	38.8 ±1.0
dSprites-loc	21.8 ±0.6	70.2 ±2.0	11.9 ±0.4	83.7 ±5.6
dSprites-ori	20.8 ±0.4	52.1 ±1.5	20.1 ±1.1	52.1 ±1.3
SmallNORB-azi	13.9 ±0.5	16.4 ±0.7	12.3 ±0.8	16.8 ±0.8
SmallNORB-elev	21.4 ±1.0	25.1 ±0.6	19.1 ±0.7	22.9 ±0.5
DMLab	29.7 ±0.5	31.6 ±0.9	30.6 ±0.3	34.6 ±0.6
KITTI-dist	62.0 ±3.5	66.0 ±2.7	61.4 ±2.3	66.8 ±3.0
FGVC-Aircraft	38.9 ±0.5	53.1 ±0.3	41.0 ±0.7	65.1 ±0.7
Cars	34.4 ±0.0	49.5 ±0.2	43.3 ±0.0	67.9 ±0.2
Letters	56.1 ±1.2	77.7 ±1.3	57.6 ±0.8	79.7 ±0.4
Average acc	49.6	59.5	49.9	64.5

Table 20. Accuracy comparison between NA and FiLM methods in offline mode using either a pre-trained ResNet-18 (RN) or a pre-trained EfficientNet-B0 (EN) backbone. We use an LDA head. The reported results are based on **50** shots. Results are averaged over 5 runs (mean±std).

Dataset	**NA**	**Meta-Learn**	**FiLM**	**Full-body**
Caltech101	88.2 ±0.8	86.7 ±0.9	89.0 ±0.6	**89.4 ±0.7**
CIFAR100	42.7 ±1.6	42.1 ±1.3	**51.8 ±1.3**	49.3 ±1.1
Flowers102	76.1 ±0.4	78.2 ±0.4	**85.0 ±0.8**	81.7 ±0.6
Pets	82.4 ±1.6	**83.8 ±1.0**	81.8 ±1.7	80.8 ±1.7
Sun397	**41.9 ±1.0**	40.2 ±0.6	40.9 ±0.7	35.2 ±0.5
SVHN	16.5 ±1.1	27.4 ±3.6	**31.7 ±3.7**	19.6 ±1.2
DTD	48.9 ±1.7	**50.5 ±1.7**	50.2 ±0.9	49.0 ±1.3
EuroSAT	76.3 ±1.8	75.2 ±1.5	**78.1 ±1.2**	76.7 ±2.9
Resics45	58.8 ±1.4	62.7 ±1.0	**64.7 ±1.2**	62.3 ±2.5
Patch Camelyon	59.8 ±7.2	64.2 ±7.3	**64.9 ±6.6**	59.4 ±4.9
Retinopathy	25.6 ±1.6	**26.8 ±3.5**	26.0 ±2.0	24.5 ±2.5
CLEVR-count	23.1 ±1.1	22.6 ±0.8	23.4 ±1.4	**23.5 ±3.0**
CLEVR-dist	24.5 ±2.3	23.8 ±1.1	23.1 ±1.1	**25.6 ±3.5**
dSprites-loc	8.5 ±0.6	8.9 ±0.5	19.8 ±2.0	**27.1 ±1.6**
dSprites-ori	16.2 ±0.8	19.2 ±0.7	**26.5 ±0.8**	19.9 ±1.5
SmallNORB-azi	9.3 ±0.8	8.7 ±1.0	10.1 ±0.6	**10.4 ±0.8**
SmallNORB-elev	15.1 ±0.6	15.4 ±0.5	15.4 ±0.7	**16.2 ±1.1**
DMLab	22.1 ±1.3	**24.9 ±1.5**	23.3 ±1.9	22.1 ±1.3
KITTI-dist	51.4 ±2.7	**55.0 ±1.5**	52.7 ±3.5	52.8 ±2.5
FGVC-Aircraft	22.1 ±1.0	31.9 ±0.6	**32.6 ±0.9**	23.6 ±0.8
Cars	22.6 ±0.6	22.8 ±0.4	**28.1 ±0.4**	23.5 ±0.3
Letters	36.1 ±2.1	46.5 ±3.2	**55.9 ±2.5**	35.7 ±3.3
Average acc	39.5	41.7	**44.3**	41.3

Table 21. Accuracy comparison between different adaptation methods in offline mode using a pre-trained EfficientNet-B0 backbone. We use an LDA head. The reported results are based on **5** shots and averaged over 5 runs (mean±std).

Dataset	NA	Meta-Learn	FiLM	Full-body
Caltech101	90.0 ±0.8	89.1 ±0.3	91.5 ±0.4	**91.7 ±0.7**
CIFAR100	50.1 ±1.5	50.1 ±1.2	**62.8 ±1.1**	59.5 ±0.3
Flowers102	83.9 ±0.3	84.4 ±0.3	**91.1 ±0.3**	90.0 ±0.5
Pets	86.4 ±0.7	**86.8 ±0.2**	86.3 ±0.8	85.7 ±0.5
Sun397	49.0 ±1.3	46.3 ±0.9	**49.3 ±0.6**	45.0 ±0.4
SVHN	19.4 ±2.9	33.0 ±2.2	**45.3 ±3.0**	27.1 ±3.8
DTD	55.6 ±0.6	57.7 ±1.4	**59.1 ±0.8**	56.7 ±1.0
EuroSAT	82.1 ±0.9	81.2 ±0.7	84.4 ±1.6	**84.9 ±1.5**
Resics45	65.5 ±1.1	68.4 ±1.2	**73.0 ±1.3**	72.8 ±1.3
Patch Camelyon	66.5 ±3.7	67.5 ±5.5	**68.5 ±6.1**	61.2 ±4.5
Retinopathy	**27.1 ±2.2**	26.9 ±0.4	26.9 ±1.1	24.6 ±3.3
CLEVR-count	25.7 ±0.6	24.3 ±1.1	27.6 ±1.6	**27.7 ±2.0**
CLEVR-dist	26.3 ±1.1	25.5 ±0.8	26.2 ±1.3	**30.0 ±4.2**
dSprites-loc	8.7 ±0.3	8.9 ±0.4	26.1 ±10.6	**44.8 ±1.3**
dSprites-ori	18.2 ±0.7	20.4 ±1.2	**34.0 ±2.6**	29.6 ±6.8
SmallNORB-azi	9.5 ±1.1	10.5 ±0.2	**11.7 ±1.2**	11.5 ±0.8
SmallNORB-elev	16.5 ±1.1	15.8 ±0.6	16.3 ±0.4	**18.3 ±1.5**
DMLab	25.7 ±1.2	**27.8 ±1.7**	26.6 ±1.5	25.6 ±1.3
KITTI-dist	52.9 ±1.5	**56.4 ±1.8**	55.4 ±3.7	54.9 ±1.3
FGVC-Aircraft	28.5 ±0.4	39.0 ±0.8	**43.3 ±1.1**	36.7 ±0.6
Cars	30.4 ±0.5	29.8 ±0.1	43.1 ±1.0	**43.5 ±0.6**
Letters	45.6 ±0.8	54.5 ±1.5	**67.5 ±1.3**	55.2 ±1.3
Average acc	43.8	45.7	**50.7**	49.0

Table 22. Accuracy comparison between different adaptation methods in offline mode using a pre-trained EfficientNet-B0 backbone. We use an LDA head. The reported results are based on **10** shots and averaged over 5 runs (mean±std).

Dataset	NA	Meta-Learn	FiLM	Full-body
Caltech101	91.9 ±0.5	91.0 ±0.4	93.8 ±0.5	**93.9 ±0.3**
CIFAR100	57.4 ±1.0	58.0 ±0.9	**73.8 ±0.9**	73.1 ±0.7
Flowers102	83.9 ±0.3	84.4 ±0.3	**91.1 ±0.3**	90.0 ±0.5
Pets	89.5 ±0.4	89.6 ±0.3	**89.9 ±0.7**	89.6 ±0.6
Sun397	55.9 ±1.0	53.7 ±0.8	59.7 ±0.6	**60.4 ±0.6**
SVHN	28.3 ±1.1	47.3 ±1.4	**77.2 ±0.8**	64.5 ±1.3
DTD	61.1 ±0.0	63.8 ±0.0	**68.4 ±0.2**	64.4 ±0.2
EuroSAT	87.7 ±0.9	85.7 ±0.6	93.0 ±0.6	**94.1 ±0.6**
Resics45	73.5 ±0.9	75.5 ±1.0	83.4 ±0.6	**87.6 ±0.7**
Patch Camelyon	76.2 ±1.1	**78.0 ±1.4**	77.9 ±2.4	72.2 ±1.8
Retinopathy	32.9 ±2.2	31.6 ±1.2	**35.2 ±1.2**	31.3 ±2.9
CLEVR-count	30.1 ±1.0	28.7 ±1.1	**46.6 ±1.1**	45.2 ±1.1
CLEVR-dist	32.2 ±1.1	30.5 ±1.7	38.8 ±1.0	**44.7 ±2.2**
dSprites-loc	11.9 ±0.4	12.2 ±0.5	83.7 ±5.6	**87.1 ±2.4**
dSprites-ori	20.1 ±1.1	24.7 ±1.9	**52.1 ±1.3**	44.8 ±2.3
SmallNORB-azi	12.3 ±0.8	12.4 ±0.5	16.8 ±0.8	**18.3 ±0.7**
SmallNORB-elev	19.1 ±0.7	18.9 ±1.0	22.9 ±0.5	**31.3 ±1.7**
DMLab	30.6 ±0.3	32.6 ±0.8	**34.6 ±0.6**	32.7 ±1.6
KITTI-dist	61.4 ±2.3	62.6 ±2.0	66.8 ±3.0	**66.9 ±2.2**
FGVC-Aircraft	41.0 ±0.7	50.9 ±0.7	65.1 ±0.7	**73.5 ±0.4**
Cars	43.3 ±0.0	40.1 ±0.0	67.9 ±0.2	**79.4 ±0.1**
Letters	57.6 ±0.8	64.2 ±0.6	79.7 ±0.4	**82.3 ±0.9**
Average acc	49.9	51.7	64.5	**64.9**

Table 23. Accuracy comparison between different adaptation methods in offline mode using a pre-trained EfficientNet-B0 backbone. We use an LDA head. The reported results are based on **50** shots and averaged over 5 runs (mean±std).

Method	Accuracy (%) in each session (↑)									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
NA	93.2	87.1	81.9	80.2	76.8	74.3	72.8	70.6	70.1	68.2
E-EWC+SDC	**97.2**	70.5	63.6	46.4	40.2	40.7	38.8	35.5	33.9	32.4
FACT	96.6	48.2	32.8	24.2	19.3	16.4	13.9	12.5	11.3	10.2
ALICE	96.6	80.2	73.7	69.4	64.8	61.7	58.1	55.7	54.3	52.4
FSA	96.0	86.3	80.5	77.7	74.2	70.9	68.2	65.8	64.2	62.8
FSA-LL	96.4	84.9	79.1	75.4	71.6	68.5	66.4	64.1	62.7	60.5
FSA-FiLM	96.4	**90.4**	**86.8**	**84.7**	**82.0**	**79.8**	**78.2**	**76.1**	**75.7**	**73.8**
GDumb-1k	94.17	86.2	81.0	76.1	70.8	64.3	62.0	59.7	57.1	54.5
GDumb-5k	97.0	91.6	88.1	85.1	81.8	77.9	75.6	73.2	71.8	69.3

Table 24. Detailed accuracy for each incremental session on **CIFAR100** under the **high-shot CIL** setting. The best results across all methods per session are in bold while the best results across the no-memory methods are underlined. A pre-trained EfficientNet-B0 on Imagenet-1k is used as a backbone for all methods.

Method	Accuracy (%) in each session (↑)								
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
NA	96.3	94.5	91.6	89.9	87.7	84.8	82.0	82.8	82.6
E-EWC+SDC	98.7	89.7	72.6	65.4	43.0	40.8	35.1	26.6	21.7
FACT	98.5	66.9	50.5	40.4	33.8	28.9	25.3	23.9	22.0
ALICE	98.1	92.2	87.0	84.4	80.4	76.4	72.6	73.4	72.8
FSA	98.0	93.6	89.8	88.8	86.8	84.4	81.7	82.3	82.8
FSA-LL	97.8	92.1	87.3	85.6	83.6	81.4	78.5	78.6	79.0
FSA-FiLM	**98.5**	**96.0**	**92.6**	**90.6**	**89.5**	**87.3**	**85.0**	**85.6**	**85.4**
GDumb-1k	96.9	94.3	91.9	90.6	88.2	85.4	80.7	82.3	82.4
GDumb-5k	97.5	96.7	94.6	92.7	91.8	90.3	89.0	90.0	90.0

Table 25. Detailed accuracy for each incremental session on **CORE50** under the **high-shot CIL** setting. The best results across all methods per session are in bold while the best results across the no-memory methods are underlined. A pre-trained EfficientNet-B0 on Imagenet-1k is used as a backbone for all methods.

Method	Accuracy (%) in each session (↑)				
	1	**2**	**3**	**4**	**5**
NA	81.5	61.5	48.5	42.5	39.9
E-EWC+SDC	99.0	55.5	48.1	44.7	39.5
FACT	**99.3**	63.5	47.4	38.7	33.8
ALICE	99.3	73.5	56.4	49.9	46.1
FSA	97.2	86.4	78.2	73.0	71.3
FSA-LL	96.7	82.7	72.6	67.4	64.6
FSA-FiLM	99.1	**89.0**	**81.7**	**77.6**	**75.9**
GDumb-1k	97.4	91.7	87.3	83.8	78.3
GDumb-5k	98.6	97.3	95.8	93.7	93.2

Table 26. Detailed accuracy for each incremental session on **SVHN** under the **high-shot CIL** setting. The best results across all methods per session are in bold while the best results across the no-memory methods are underlined. A pre-trained EfficientNet-B0 on Imagenet-1k is used as a backbone for all methods.

Method	Accuracy (%) in each session (↑)						
	1	**2**	**3**	**4**	**5**	**6**	**7**
NA	44.9	37.4	31.6	27.1	23.8	21.5	20.6
E-EWC+SDC	99.5	58.1	29.9	33.0	19.4	21.7	18.6
FACT	**100.0**	16.6	12.7	10.1	8.4	7.2	6.4
ALICE	**100.0**	92.6	77.6	69.6	65.6	69.8	68.3
FSA	**100.0**	**95.4**	**92.3**	**87.7**	**89.9**	**90.7**	**91.5**
FSA-LL	99.8	94.0	93.9	90.5	91.5	91.4	91.3
FSA-FiLM	99.6	89.6	84.6	78.7	77.5	77.0	76.9
GDumb-1k	91.2	85.5	85.6	83.8	76.3	78.1	79.5
GDumb-5k	99.4	99.5	99.6	98.5	99.4	98.4	99.4

Table 27. Detailed accuracy for each incremental session on **dSprites-loc** under the **high-shot CIL** setting. The best results across all methods per session are in bold while the best results across the no-memory methods are underlined. A pre-trained EfficientNet-B0 on Imagenet-1k is used as a backbone for all methods.

Method	Accuracy (%) in each session (↑)									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
NA	42.0	38.0	29.9	37.0	43.9	41.3	40.7	42.4	41.2	41.2
E-EWC-SDC	58.0	35.3	27.3	27.7	27.4	28.2	23.5	25.4	25.7	25.6
FACT	58.7	30.6	24.9	21.4	19.1	19.0	17.6	16.9	15.5	14.7
ALICE	**61.3**	43.5	36.7	39.5	41.5	41.8	41.0	41.6	40.0	39.8
FSA	54.2	44.4	39.7	45.9	52.2	49.6	48.9	51.4	51.1	50.8
FSA-LL	58.0	40.5	37.0	41.6	44.3	44.9	44.9	46.1	45.1	45.4
FSA-FiLM	52.9	**46.2**	**44.7**	**50.3**	**53.3**	**55.0**	**54.5**	**56.3**	**55.5**	**55.9**
GDumb-1k	59.8	47.3	42.6	46.5	51.8	43.7	43.4	43.0	39.2	38.4
GDumb-5k	58.6	43.0	36.9	40.1	41.0	36.1	30.3	30.3	29.5	25.3

Table 28. Detailed accuracy for each incremental session on **FGVC-Aircraft** under the **high-shot CIL** setting. The best results across all methods per session are in bold while the best results across the no-memory methods are underlined. A pre-trained EfficientNet-B0 on Imagenet-1k is used as a backbone for all methods.

Method	Accuracy (%) in each session (↑)									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
NA	72.7	49.9	49.9	46.4	47.1	47.2	44.5	44.7	44.6	43.3
E-EWC+SDC	80.2	47.3	38.7	37.2	34.8	33.5	31.0	31.6	31.2	30.0
FACT	79.8	2.9	1.9	1.4	1.1	0.9	0.8	0.7	0.6	0.6
ALICE	82.7	52.9	49.6	45.1	42.8	41.1	39.0	38.8	38.1	36.4
FSA	79.3	55.0	55.9	54.0	53.3	52.7	51.5	51.5	51.4	50.3
FSA-LL	**81.2**	49.3	50.5	49.9	49.0	46.9	46.8	46.9	46.6	45.7
FSA-FiLM	80.1	**59.5**	**60.0**	**59.4**	**58.9**	**58.2**	**56.8**	**57.3**	**56.7**	**55.9**
GDumb-1k	81.5	59.9	54.0	47.3	40.5	33.9	32.7	27.6	22.8	18.1
GDumb-5k	82.1	65.8	59.0	51.8	46.7	38.0	37.6	32.3	28.7	24.2

Table 29. Detailed accuracy for each incremental session on **Cars** under the **high-shot CIL** setting. The best results across all methods per session are in bold while the best results across the no-memory methods are underlined. A pre-trained EfficientNet-B0 on Imagenet-1k is used as a backbone for all methods.

Method	Accuracy (%) in each session (↑)										
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
NA	90.12	84.3	82.4	80.1	78.7	78.0	76.0	75.3	72.8	71.6	68.4
E-EWC+SDC	99.9	83.4	64.5	59.9	47.8	54.1	42.6	40.4	31.1	30.0	33.6
FACT	99.9	69.9	53.9	43.4	36.3	32.7	29.3	27.0	24.4	22.4	20.9
ALICE	**99.9**	96.1	93.4	89.5	88.7	87.9	85.8	83.7	81.1	79.3	75.7
FSA	99.8	96.4	94.6	91.3	90.3	89.6	87.9	86.3	83.4	82.0	78.4
FSA-LL	99.8	95.9	94.0	90.4	89.0	88.3	86.3	85.3	82.4	81.0	77.2
FSA-FiLM	99.6	**96.0**	**94.4**	**92.0**	**91.1**	**90.6**	**88.5**	**87.7**	**85.0**	**83.4**	**79.7**
GDumb-1k	96.0	92.2	89.4	86.7	85.7	83.9	81.1	80.3	76.2	75.2	70.1
GDumb-5k	99.4	98.3	97.2	95.2	94.6	94.3	92.2	91.3	88.5	86.9	82.6

Table 30. Detailed accuracy for each incremental session on **Letters** under the **high-shot CIL** setting. The best results across all methods per session are in bold while the best results across the no-memory methods are underlined. A pre-trained EfficientNet-B0 on Imagenet-1k is used as a backbone for all methods.

Method	Backbone	Accuracy (%) in each session (↑)								
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Decoupled-Cos*	RN-20	74.6	67.4	63.6	59.6	56.1	53.8	51.7	49.7	47.7
CEC*	RN-20	73.1	68.9	65.3	61.2	58.1	55.6	53.2	51.3	49.1
FACT*	RN-20	74.6	72.1	67.6	63.5	61.4	58.4	56.3	54.2	52.1
FSA	RN-20	75.1	71.2	67.5	63.3	60.0	57.6	55.5	54.2	52.0
NA	RN-18	68.9	65.4	62.4	58.7	57.2	54.7	53.3	51.9	50.4
FACT	RN-18	75.8	71.0	66.3	62.5	59.1	56.3	54.1	51.8	49.5
ALICE†	RN-18	79.0	70.5	67.1	63.4	61.2	59.2	58.1	56.3	54.1
FSA-FiLM	RN-18	73.0	69.7	66.3	63.2	61.9	59.3	58.3	57.2	55.2
FSA	RN-18	82.0	78.2	74.8	70.22	68.7	66.2	65.3	63.8	61.4
NA	EN-B0	74.4	70.4	67.4	63.4	62.4	59.8	58.4	56.9	55.2
FACT	EN-B0	86.4	80.6	75.6	71.1	67.6	64.4	61.8	59.2	56.5
ALICE	EN-B0	87.7	83.3	78.7	74.4	72.1	69.6	67.4	65.4	62.7
FSA-FiLM	EN-B0	79.6	75.6	72.9	68.8	68.2	65.4	64.9	63.9	61.8
FSA	EN-B0	**87.6**	**83.5**	**79.7**	**75.4**	**73.8**	**70.9**	**70.2**	**68.8**	**66.1**

Table 31. Detailed accuracy for each incremental session on **CIFAR100** under the **few-shot+ CIL** setting. Asterisk (*) indicates that the reported results of a method are from [31] and † that the reported results of a method are from [22]. We use three different backbones, EfficientNet-B0 (EN-B0) and ResNet-18/20 (RN-18/20); EN-B0 and RN-18 are pre-trained on Imagenet-1k.

Method	Backbone	Accuracy (%) in each session (↑)										
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
NA		70.7	66.7	63.4	59.0	58.2	56.4	54.0	52.3	50.5	50.5	50.0
Decoupled-Cos*		75.5	71.0	66.5	61.2	60.9	56.9	55.4	53.5	51.9	50.9	49.3
CEC*		75.9	71.9	68.5	63.5	62.4	58.3	57.7	55.8	54.8	53.5	52.3
FACT*	RN-18	75.9	73.2	70.8	66.1	65.6	62.2	61.7	59.8	58.4	57.9	56.9
ALICE†		77.4	72.7	70.6	67.2	65.9	63.4	62.9	61.9	60.5	60.6	60.1
FSA-FiLM		72.7	68.2	64.9	60.8	60.2	58.1	55.4	54.8	53.5	53.4	52.7
FSA		76.1	72.6	69.6	65.0	64.6	62.3	61.6	59.6	58.2	58.2	57.6
NA		78.6	75.8	73.4	69.5	69.2	67.3	66.5	64.3	62.7	63.1	63.2
FACT		**82.0**	**77.5**	74.4	70.0	69.3	66.6	66.2	64.7	**64.0**	63.3	62.9
ALICE	EN-B0	81.6	77.1	**75.1**	**71.9**	**70.5**	**67.8**	**66.8**	**65.7**	**64.1**	**64.0**	**63.5**
FSA-FiLM		79.0	75.3	72.7	69.5	68.3	66.5	65.3	64.1	62.8	62.9	62.9
FSA		80.2	77.1	74.2	69.3	69.3	66.9	66.4	64.8	63.6	63.8	63.4

Table 32. Detailed accuracy for each incremental session on **CUB200** under the **few-shot+ CIL** setting. Asterisk (*) indicates that the reported results of a method are from [31] and † that the reported results of a method are from [22]. We use two different backbones, EfficientNet-B0 (EN-B0) and ResNet-18 (RN-18); EN-B0 and RN-18 are pre-trained on Imagenet-1k.

Method	Accuracy (%) in each session (↑)								
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
NA	80.2 ± 0.9	76.2 ± 0.7	72.4 ± 1.5	68.7 ± 1.5	65.2 ± 1.4	63.5 ± 1.3	60.8 ± 1.3	59.5 ± 1.9	57.4 ± 1.0
GDumb	83.9 ± 1.4	80.1 ± 0.5	77.6 ± 1.6	71.7 ± 2.2	67.5 ± 1.8	64.0 ± 1.5	59.4 ± 1.2	58.6 ± 1.9	55.7 ± 1.0
FACT	83.0 ± 1.2	54.5 ± 1.4	40.7 ± 1.0	32.6 ± 1.0	27.6 ± 0.9	23.8 ± 0.7	20.8 ± 0.5	18.7 ± 0.4	16.8 ± 0.3
FSA	82.7 ± 2.1	75.4 ± 1.7	72.5 ± 1.8	69.5 ± 1.6	66.8 ± 1.8	64.5 ± 1.4	63.1 ± 1.5	62.7 ± 1.5	60.3 ± 1.3
FSA-FiLM	**89.2 ± 0.9**	**85.8 ± 1.3**	**84.3 ± 1.3**	**81.0 ± 1.3**	**77.9 ± 1.7**	**75.9 ± 1.0**	**74.3 ± 1.4**	**74.2 ± 1.1**	**70.9 ± 1.0**

Table 33. Detailed accuracy for each incremental session on **CIFAR100** under the **few-shot CIL** setting. GDumb is the only memory-based method used for comparisons; we use a buffer size equal to the first session's number of images N_1. The best results across all methods are in bold while the best results across the no-memory methods are underlined. A pre-trained EfficientNet-B0 on Imagenet-1k is used as a backbone for all methods.

Method	Accuracy (%) in each session (↑)				
	1	**2**	**3**	**4**	**5**
NA	73.5 ± 4.6	53.3 ± 3.7	37.3 ± 1.8	33.6 ± 1.5	28.3 ± 1.1
GDumb	78.2 ± 4.9	46.7 ± 5.1	35.2 ± 1.6	23.3 ± 3.8	21.0 ± 2.1
FACT	71.3 ± 1.0	46.6 ± 3.5	34.0 ± 1.9	27.7 ± 2.5	24.1 ± 2.0
FSA	70.7 ± 2.9	50.8 ± 3.9	38.4 ± 3.2	35.7 ± 1.6	32.9 ± 1.0
FSA-FiLM	**90.7 ± 1.8**	**70.4 ± 1.4**	**60.5 ± 1.5**	**55.5 ± 2.3**	**51.3 ± 2.1**

Table 34. Detailed accuracy for each incremental session on **SVHN** under the **few-shot CIL** setting. GDumb is the only memory-based method used for comparisons; we use a buffer size equal to the first session's number of images N_1. The best results across all methods are in bold while the best results across the no-memory methods are underlined. A pre-trained EfficientNet-B0 on Imagenet-1k is used as a backbone for all methods.

Method	Accuracy (%) in each session (↑)						
	1	**2**	**3**	**4**	**5**	**6**	**7**
NA	35.7 ± 1.4	26.4 ± 1.7	22.1 ± 1.7	18.1 ± 1.2	15.4 ± 0.8	13.5 ± 0.7	11.9 ± 0.4
GDumb	36.4 ± 7.9	29.9 ± 6.3	20.3 ± 3.6	22.1 ± 5.5	13.1 ± 2.1	11.7 ± 3.2	16.4 ± 2.6
FACT	32.6 ± 1.4	22.7 ± 2.2	18.4 ± 2.1	14.4 ± 1.9	12.4 ± 1.5	11.9 ± 1.6	11.7 ± 1.7
FSA	57.4 ± 2.3	44.8 ± 2.8	39.3 ± 1.6	34.8 ± 2.1	32.9 ± 1.9	33.2 ± 2.6	33.7 ± 1.7
FSA-FiLM	**62.7 ± 2.1**	**50.2 ± 2.4**	**46.9 ± 2.3**	**40.1 ± 2.2**	**37.3 ± 2.5**	**36.1 ± 2.2**	**35.7 ± 2.1**

Table 35. Detailed accuracy for each incremental session on **dSprites-position** under the **few-shot CIL** setting. GDumb is the only memory-based method used for comparisons; we use a buffer size equal to the first session's number of images N_1. The best results across all methods are in bold while the best results across the no-memory methods are underlined. A pre-trained EfficientNet-B0 on Imagenet-1k is used as a backbone for all methods.

Method	Accuracy (%) in each session (↑)								
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
NA	35.5 ± 0.9	28.7 ± 0.7	36.4 ± 0.8	42.8 ± 0.4	40.0 ± 0.4	39.4 ± 0.6	41.4 ± 0.7	40.3 ± 0.9	41.0 ± 0.7
GDumb	**51.1 ± 1.7**	**45.4 ± 1.2**	46.9 ± 1.8	52.2 ± 1.0	45.4 ± 1.7	42.5 ± 1.6	41.8 ± 0.9	39.5 ± 1.9	38.6 ± 1.0
FACT	41.4 ± 0.6	25.9 ± 0.9	19.6 ± 0.4	16.3 ± 0.4	13.7 ± 0.4	11.3 ± 0.5	10.4 ± 0.5	9.4 ± 0.6	8.3 ± 0.6
FSA	42.9 ± 2.6	39.5 ± 2.0	45.5 ± 1.5	51.6 ± 1.7	48.2 ± 1.8	47.3 ± 1.4	49.8 ± 1.5	49.1 ± 1.7	50.1 ± 1.5
FSA-FiLM	46.6 ± 1.9	44.8 ± 1.1	**49.4 ± 0.8**	**52.9 ± 1.6**	**54.0 ± 0.7**	**53.5 ± 1.0**	**55.8 ± 0.7**	**55.2 ± 0.5**	**55.8 ± 0.6**

Table 36. Detailed accuracy for each incremental session on **FGVC-Aircraft** under the **few-shot CIL** setting. GDumb is the only memory-based method used for comparisons; we use a buffer size equal to the first session's number of images N_1. The best results across all methods are in bold while the best results across the no-memory methods are underlined. A pre-trained EfficientNet-B0 on Imagenet-1k is used as a backbone for all methods.

Method	Accuracy (%) in each session (↑)										
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
NA	82.1 ± 0.8	76.6 ± 0.9	73.0 ± 1.5	69.6 ± 1.0	69.0 ± 1.3	67.6 ± 1.3	65.8 ± 1.2	64.1 ± 1.0	60.4 ± 0.6	59.0 ± 0.8	57.6 ± 0.8
GDumb	91.3 ± 1.6	**91.8 ± 1.2**	80.0 ± 1.4	72.0 ± 1.9	69.2 ± 3.0	63.5 ± 1.5	59.9 ± 1.1	54.5 ± 0.3	48.0 ± 0.4	44.3 ± 1.9	41.2 ± 1.7
FACT	84.3 ± 1.4	72.0 ± 1.2	68.0 ± 2.2	63.2 ± 1.0	62.3 ± 1.0	59.5 ± 1.2	58.0 ± 1.0	55.8 ± 1.0	52.8 ± 0.7	51.7 ± 0.8	49.8 ± 0.8
FSA	87.0 ± 1.4	79.6 ± 1.1	76.4 ± 0.9	72.7 ± 0.7	73.0 ± 0.9	71.3 ± 0.4	69.7 ± 0.4	68.4 ± 0.4	64.7 ± 0.5	62.9 ± 0.4	62.2 ± 0.4
FSA-FiLM	**94.3 ± 0.9**	90.6 ± 0.3	**88.6 ± 1.0**	**85.1 ± 0.6**	**84.9 ± 0.4**	**84.0 ± 0.4**	**82.5 ± 0.7**	**81.1 ± 0.4**	**76.8 ± 0.4**	**75.0 ± 0.4**	**73.4 ± 0.4**

Table 37. Detailed accuracy for each incremental session on **Letters** under the **few-shot CIL** setting. GDumb is the only memory-based method used for comparisons; we use a buffer size equal to the first session's number of images N_1. The best results across all methods are in bold while the best results across the no-memory methods are underlined. A pre-trained EfficientNet-B0 on Imagenet-1k is used as a backbone for all methods.

Method	Accuracy (%) in each session (↑)								
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
NA	83.3 ± 1.0	62.6 ± 0.5	60.9 ± 1.0	61.5 ± 0.3	68.3 ± 1.2	71.1 ± 0.6	72.0 ± 0.7	70.8 ± 1.0	69.0 ± 0.3
GDumb	**88.3 ± 0.4**	63.6 ± 0.8	58.6 ± 0.6	56.4 ± 1.0	66.0 ± 1.1	68.7 ± 0.4	68.9 ± 0.6	65.8 ± 1.0	63.2 ± 1.1
FACT	84.3 ± 0.6	53.6 ± 0.6	43.0 ± 0.5	35.9 ± 0.6	28.3 ± 0.6	24.2 ± 0.3	22.6 ± 0.3	22.0 ± 0.5	20.6 ± 0.2
FSA	85.2 ± 0.3	63.3 ± 0.3	61.4 ± 0.7	61.6 ± 0.5	68.5 ± 0.9	71.2 ± 1.1	72.2 ± 0.5	71.2 ± 0.7	70.3 ± 0.4
FSA-FiLM	87.7 ± 0.3	**68.5 ± 0.6**	**66.9 ± 0.4**	**66.7 ± 0.9**	**73.7 ± 0.6**	**76.0 ± 0.7**	**76.0 ± 0.5**	**75.0 ± 0.6**	**74.0 ± 0.3**

Table 38. Detailed accuracy for each incremental session on **DomainNet** under the **few-shot CIL** setting. GDumb is the only memory-based method used for comparisons; we use a buffer size equal to the first session's number of images N_1. The best results across all methods are in bold while the best results across the no-memory methods are underlined. A pre-trained EfficientNet-B0 on Imagenet-1k is used as a backbone for all methods.

Method	Accuracy (%) in each session (↑)								
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
NA	51.9	52.8	44.9	46.8	49.3	51.7	54.4	53.8	49.7
GDumb	56.4	50.1	36.3	47.5	44.2	44.7	46.4	40.9	40.4
FACT	54.9	29.9	24.6	23.8	20.4	17.8	15.7	16.4	14.3
FSA	52.1	53.6	40.0	47.8	49.2	51.3	55.1	55.1	51.5
FSA-FiLM	**61.8**	**61.6**	**52.0**	**56.5**	**57.0**	**59.4**	**61.8**	**61.2**	**58.8**

Table 39. Detailed accuracy for each incremental session on **iNaturalist** under the **few-shot CIL** setting. GDumb is the only memory-based method used for comparisons; we use a buffer size equal to the first session's number of images N_1. The best results across all methods are in bold while the best results across the no-memory methods are underlined. A pre-trained EfficientNet-B0 on Imagenet-1k is used as a backbone for all methods.

References

[1] Charles Beattie, Joel Z Leibo, Denis Teplyashin, Tom Ward, Marcus Wainwright, Heinrich Küttler, Andrew Lefrancq, Simon Green, Víctor Valdés, Amir Sadik, et al. Deepmind lab. *arXiv preprint arXiv:1612.03801*, 2016. 2

[2] John Bronskill, Daniela Massiceti, Massimiliano Patacchiola, Katja Hofmann, Sebastian Nowozin, and Richard Turner. Memory efficient meta-learning with large images. *Advances in Neural Information Processing Systems*, 34:24327–24339, 2021. 3

[3] Gong Cheng, Junwei Han, and Xiaoqiang Lu. Remote sensing image scene classification: Benchmark and state of the art. *Proceedings of the IEEE*, 105(10):1865–1883, 2017. 2

[4] Mircea Cimpoi, Subhransu Maji, Iasonas Kokkinos, Sammy Mohamed, and Andrea Vedaldi. Describing textures in the wild. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3606–3613, 2014. 2

[5] T. E. de Campos, B. R. Babu, and M. Varma. Character recognition in natural images. In *Proceedings of the International Conference on Computer Vision Theory and Applications, Lisbon, Portugal*, February 2009. 2

[6] Li Fei-Fei, Robert Fergus, and Pietro Perona. One-shot learning of object categories. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(4):594–611, 2006. 2

[7] Andreas Geiger, Philip Lenz, Christoph Stiller, and Raquel Urtasun. Vision meets robotics: The kitti dataset. *The International Journal of Robotics Research*, 32(11):1231–1237, 2013. 2

[8] Patrick Helber, Benjamin Bischke, Andreas Dengel, and Damian Borth. Eurosat: A novel dataset and deep learning benchmark for land use and land cover classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 12(7):2217–2226, 2019. 2

[9] Elad Hoffer and Nir Ailon. Deep metric learning using triplet network. In *Similarity-Based Pattern Recognition: Third International Workshop, SIMBAD 2015, Copenhagen, Denmark, October 12-14, 2015. Proceedings 3*, pages 84–92. Springer, 2015. 3

[10] Justin Johnson, Bharath Hariharan, Laurens Van Der Maaten, Li Fei-Fei, C Lawrence Zitnick, and Ross Girshick. Clevr: A diagnostic dataset for compositional language and elementary visual reasoning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2901–2910, 2017. 2

[11] Kaggle and EyePacs. Kaggle diabetic retinopathy detection. `https://www.kaggle.com/c/diabetic-retinopathy-detection/data`, 2015. 2

[12] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014. 3

[13] Jonathan Krause, Michael Stark, Jia Deng, and Li Fei-Fei. 3d object representations for fine-grained categorization. In *Proceedings of the IEEE International Conference on Computer Vision Workshops*, pages 554–561, 2013. 2

[14] Alex Krizhevsky, Geoffrey Hinton, et al. Learning multiple layers of features from tiny images. 2009. 2

[15] Yann LeCun, Fu Jie Huang, and Leon Bottou. Learning methods for generic object recognition with invariance to pose and lighting. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, volume 2, pages II–104. IEEE, 2004. 2

[16] Vincenzo Lomonaco and Davide Maltoni. Core50: A new dataset and benchmark for continuous object recognition. In *Conference on Robot Learning*, pages 17–26. PMLR, 2017. 2

[17] Subhransu Maji, Esa Rahtu, Juho Kannala, Matthew Blaschko, and Andrea Vedaldi. Fine-grained visual classification of aircraft. *arXiv preprint arXiv:1306.5151*, 2013. 2

[18] Loic Matthey, Irina Higgins, Demis Hassabis, and Alexander Lerchner. dsprites: Disentanglement testing sprites dataset. `https://github.com/deepmind/dsprites-dataset/`, 2017. 2

[19] Yuval Netzer, Tao Wang, Adam Coates, Alessandro Bissacco, Bo Wu, and Andrew Y Ng. Reading digits in natural images with unsupervised feature learning. 2011. 2

[20] Maria-Elena Nilsback and Andrew Zisserman. Automated flower classification over a large number of classes. In *2008 Sixth Indian Conference on Computer Vision, Graphics & Image Processing*, pages 722–729. IEEE, 2008. 2

[21] Omkar M Parkhi, Andrea Vedaldi, Andrew Zisserman, and CV Jawahar. Cats and dogs. In *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pages 3498–3505. IEEE, 2012. 2

[22] Can Peng, Kun Zhao, Tianren Wang, Meng Li, and Brian C Lovell. Few-shot class-incremental learning from an open-set perspective. In *Computer Vision–ECCV 2022: 17th European Conference, Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part XXV*, pages 382–397. Springer, 2022. 3, 16, 17

[23] Xingchao Peng, Qinxun Bai, Xide Xia, Zijun Huang, Kate Saenko, and Bo Wang. Moment matching for multi-source domain adaptation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1406–1415, 2019. 2

[24] Ameya Prabhu, Philip HS Torr, and Puneet K Dokania. GDumb: A simple approach that questions our progress in continual learning. In *European Conference on Computer Vision*, pages 524–540. Springer, 2020. 3

[25] Grant Van Horn, Oisin Mac Aodha, Yang Song, Yin Cui, Chen Sun, Alex Shepard, Hartwig Adam, Pietro Perona, and Serge Belongie. The iNaturalist Species Classification and Detection Dataset. In *Proceedings of the IEEE Conference/CVF on Computer Vision and Pattern Recognition*, pages 8769–8778, 2018. 2

[26] Bastiaan S Veeling, Jasper Linmans, Jim Winkens, Taco Cohen, and Max Welling. Rotation equivariant cnns for digital pathology. In *International Conference on Medical image computing and computer-assisted intervention*, pages 210–218. Springer, 2018. 2

[27] C. Wah, S. Branson, P. Welinder, P. Perona, and S. Belongie. Technical Report CNS-TR-2011-001, California Institute of Technology, 2011. 2

[28] Jianxiong Xiao, James Hays, Krista A Ehinger, Aude Oliva, and Antonio Torralba. Sun database: Large-scale scene recognition from abbey to zoo. In *2010 IEEE Conference on Computer Vision and Pattern Recognition*, pages 3485–3492. IEEE, 2010. 2

[29] Lu Yu, Bartlomiej Twardowski, Xialei Liu, Luis Herranz, Kai Wang, Yongmei Cheng, Shangling Jui, and Joost van de Weijer. Semantic drift compensation for class-incremental learning. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 6982–6991, 2020. 3

[30] Xiaohua Zhai, Joan Puigcerver, Alexander Kolesnikov, Pierre Ruyssen, Carlos Riquelme, Mario Lucic, Josip Djolonga, Andre Susano Pinto, Maxim Neumann, Alexey Dosovitskiy, et al. A large-scale study of representation learning with the visual task adaptation benchmark. *arXiv preprint arXiv:1910.04867*, 2019. 1

[31] 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 *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 9046–9056, 2022. 2, 3, 16, 17