Supplementary Material: Lightweight Learner for Shared Knowledge Lifelong Learning

Yunhao Ge[1] *yunhaoge@usc.edu*
Yuecheng Li[1*] *liyueche@usc.edu*
Di Wu[1*] *dwu92983@usc.edu*
Ao Xu[1*] *aoxu@usc.edu*
Adam M. Jones[2] *adammj@usc.edu*
Amanda Sofie Rios[3] *amanda.rios@intel.com*
Iordanis Fostiropoulos[1] *fostirop@usc.edu*
Shixian Wen[4] *sx.wen@siat.ac.cn*
Po-Hsuan Huang[2] *pohsuanh@usc.edu*
Zachary William Murdock[2] *zmurdock@usc.edu*
Gozde Sahin[1] *gsahin@usc.edu*
Shuo Ni[1] *shuoni@usc.edu*
Kiran Lekkala[1] *klekkala@usc.edu*
Sumedh Anand Sontakke[1] *ssontakk@usc.edu*
Laurent Itti[1,2,5] *itti@usc.edu*

[1] *Thomas Lord Department of Computer Science, University of Southern California*
[2] *Neuroscience Graduate Program, University of Southern California*
[3] *Intel Labs*
[4] *Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences*
[5] *Dornsife Department of Psychology, University of Southern California*
[*] *Equal contribution as second author*

Reviewed on OpenReview: *https://openreview.net/forum?id=Jjl2c8kWUc*

A Dataset subsampling details

Our SKILL-102 dataset comprises 102 distinct tasks that were obtained from previously published datasets. SKILL-102 is freely available for download on the project website: https://github.com/gyhandy/Shared-Knowledge-Lifelong-Learning.

Here, we subsampled the source datasets slightly, mainly to allow some of the baselines to converge in a reasonable amount of time. For dataset sampling, the following rules were used:

- For iNaturalist Insecta, since it contains a lot of classes, 500 classes were randomly sampled.
- For all other tasks, all classes are kept.
- For all tasks, round(54000/c) training images and round(6000/c) validation images and round(6000/c) test images are used for each class. If a class does not contain enough images, then all images for that class are used.
- The exact datasets as we used them in our experiments will be made available online after publication, to allow other researchers to reproduce (or beat!) our results.

The sequence of datasets and number of images in each dataset are shown in Fig. S5.

B GMMC number of clusters

Fig. S1 shows the GMMC performance with different numbers of clusters.

# of clusters	2	5	10	15	20	25	30	35	40
Accuracy	74.89%	76.98%	78.27%	82.06%	82.08%	82.38%	82.24%	81.79%	81.77%

Figure S1: On a small subset of tasks, we found that $k = 25$ GMMC clusters provided the best compromise between generalization and overfitting.

C Mahalanobis training MACs

The slope of MACs/image is higher until the number of training samples reaches 4,000. After that, the slope does not change. If we use 5 images per class to train, then the number of training samples would reach 4,000 after task 12. So for the majority of the tasks, the average MACs per image for training the Mahalanobis distance is around 250k.

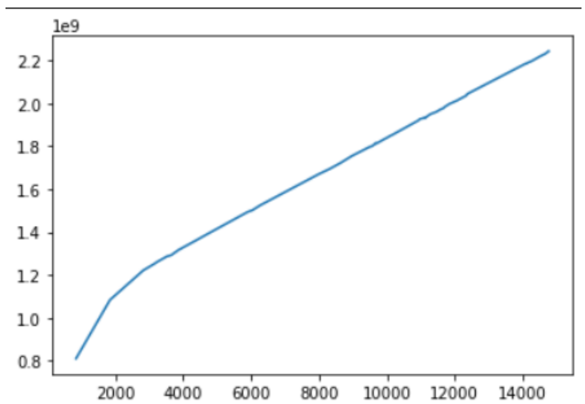

Figure S2: MACs for Mahalanobis training (vertical axis) as a function of number of training images (horizontal axis).

D CPU analysis

We compute everything in terms of MACs/image processed. There are a few caveats:

- Data sharing does not occur per training image, but rather per task (e.g., share 25 GMMC cluster means+diagonal covariances per task). Hence we first compute communication bytes/task and then convert that to "MACs equivalent" by assuming that sharing 1 byte takes the equivalent of 1,000 MACs. This value is a hyper-parameter than can be tuned depending on network type. Over wired Ethernet, it corresponds to 1.5 million MACs per packet (with MTU of 1500 bytes).
- Mahalanobis training time increases with the number of tasks received to date, as shown in Fig. S2.
- ER training increases over time as more tasks are added:
  - We first train task 1 using the whole task 1 training set (subsampled version described above).
  - Then train task 2 using the whole task 2 training set + 10 images/class of task 1 (chosen randomly). In what follows we use γ to represent this fraction of data used for rehearsing of old tasks, and we denote by S a nominal dataset size per task (2,500 images on average). Hence, for task 2, the episodic buffer method uses $S \times (1 + \gamma)$ images. With a normalized training time of 1 to learn one task, learning task 2 for this baseline takes normalized time $1 + \gamma$.

- Then train task 3 using the whole task 3 training set + 10 images/class of task 1 + 10 images/class of task 2. Normalized training time $1 + 2\gamma$.
- Then train task 4 using the whole task 4 training set + 10 images/class of task 1 + 10 images/class of task 2 + 10 images/class of task 3. Normalized training time $1 + 3\gamma$.
- etc. So the total normalized training time for N tasks is $(1) + (1 + \gamma) + (1 + 2\gamma) + (1 + 3\gamma) + ... + (1 + (N - 1)\gamma) = N + \gamma(1 + 2 + ... + N - 1) = N + \gamma(N - 1)(N - 2)/2$. With $N = 102$, the total training time for all tasks is $N + 5050\gamma$. In our experiments, our subsampled training sets averaged 254 images/class and hence $\gamma = 10/254 = 0.04$ on average, leading to a total normalized training time of 304 (broken down as a cost of 102 to learn the from 102 datasets, plus 202 to rehearse old tasks as we learn new tasks).
- This is for $\gamma = 0.04$ but performance is low, so using a higher γ is warranted for the episodic buffer approach. This is very costly, though. In the limit of retaining all images, which would give best performance, the training time of this approach is $102 + 5050 = 5152$ times the time it takes to learn one task. So, while the single-agent will require anywhere between $304 \times T$ and $5152 \times T$ to learn 102 tasks sequentially, our approach will learn all 102 tasks in parallel during just T.

Additional details used for our computations are in Fig. S3.

Teacher	Student	Runtime	
7.58E+11	0.00E+00	8.44E+09	unfrozen xception MACs/image
0.00E+00	0.00E+00	8.44E+09	frozen backbone MACs/image
1.21E+07	0.00E+00	1.01E+05	last layer only MACs/image
1.08E+09	0.00E+00	3.58E+07	BB only MACs/image
4.96E+06	0.00E+00	8.55E+07	GMMC MACs/image
2.50E+05	2.50E+05	6.07E+08	Mahalanobis MACs/image
8.56E+11	0.00E+00	8.44E+09	EWC MACs/image
3.07E+11	0.00E+00	3.42E+09	PSP MACs/image
2.22E+12	0.00E+00	8.44E+09	ER MACs/image
4.95E+11	0.00E+00	4.15E+09	SUPSUP MACs/image
7.61E+11	0.00E+00	8.44E+09	EWC-ONLINE MACs/image
7.65E+11	0.00E+00	8.44E+09	LwF MACs/image
1.02E+12	0.00E+00	8.44E+09	SI MACs/image
1.01E+12	0.00E+00	8.44E+09	MAS MACs/image
7.58E+11	0.00E+00	8.48E+09	BB backbone MACs/image

Parameters	
# tasks (=N)	102
MACs/byte transmitted	1000
bytes/parameter	4
bytes/image	268203
# xception weights	20884814
xception latent dims	2048
xception forward MACs	8.44E+09
# GMMC clusters	25
GMMC params/cluster	4096
avg classes/task	49.34
median classes/task	12.00
avg train img/task	20011.95
avg test img/task	2386.93
Mahalanobis img/class	5
# BB biases	50312
# training epochs (avg)	30
# train images	2041219
# test images	243467
ER replay Factor	2.931372549

Computed from parameters	
GMMC shared bytes/task	409600
last layer bytes/task	404391
BB biases bytes/task	201248
Mahalanobis bytes/task	66165680.1
SUPSUP bytes/task	3000000

Figure S3: Additional details for how we compute MACs and speedup. Different assumptions (e.g., higher or lower MACs/byte transmitted) can be used, which would update the results in the main paper Figs. 9 and 10.

E Summary of our new SKILL-102 for image classification

Fig. S5 shows a summary of 102 datasets we are using along with the accuracy of all our methods. Note that TM stands for Task Mapper. The red text indicates datasets with large domain gap which were mentioned in Sec. 6, the blue text indicates datasets with poor GMMC accuracy which are further examined below. Fig. S6 shows the baselines performance on SKILL-102.

F Cases of low accuracy in GMMC

In this section, we analyze in details the failures of GMMC on three datasets: Office Home Art, Dragon Ball, and Malacca Historical Buildings.

In certain cases, several datasets may share a common characteristic, such as all of them are anime pictures (e.g. Dragon Ball, Pokemon, and One-Piece). GMMC may capture the tasks' characteristics as animation but fail to further distinguish different tasks. On the other side, Mahalanobis focus on the characteristics

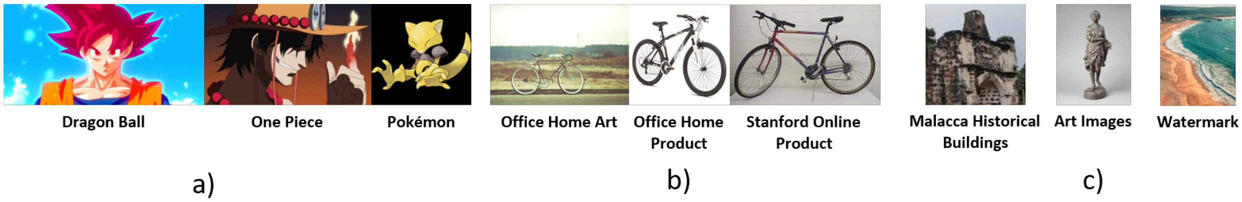

Figure S4: Here we analyze the top 3 tasks into which images may be misclassified by GMMC. a) Out of 18 test images from the (very small) Dragon Ball Dataset, 4 are correctly classified as belonging to Dragon Ball Dataset, 11 are misclassified as belonging to the One Piece dataset, and 2 are misclassified as belonging to the Pokemon dataset. Since all three datasets contain cartoon images, GMMC was confused to classify some images into an incorrect dataset. b) Out of 252 test images from Office Home Art, 86 are correctly classified, 33 are classified as belonging to the Stanford Online Product dataset, and 20 are classified as Office Home Product dataset. These three datasets have many objects in common such as bicycles, chairs, and tables. Hence, it is easy for GMMC to get confused. c) Out of 18 test images from Malacca Historical Buildings, 7 were correctly classified, 5 are classified as Art Images, and 5 are classified as belonging to the Watermark dataset. The Art Image and Watermark datasets contain a large variety of images which may confuse the GMMC to make wrong predictions.

classwise, witch captures the difference among classes (e.g., Wukong vs. Abra characters in Fig. S4-a) and hence is able to distinguish them.

Another case is that two tasks may share similar objects (e.g., Office Home Art, Office Home, and Stanford Online Products; Fig. S4-b). Although represented in different tasks, these are the same types of objects in the real life. We address these GMMC confusions with our proposed "corrective approach" that would declare correct classification for equivalent labels belonging to different tasks.

Other cases may include that one task is too general; for example, Watermark non Watermark includes a large variety of images with or without a watermark which may also confuse GMMC as many similar images are present in other datasets (Fig. S4-c).

G Amount of data shared by LLL

The analysis below includes 2 options not exercised in the main text of this paper:

- **Head2Toe:** If the input domain encountered by an agent is very different than what the frozen backbone was trained on, sharing only the last layer(s) + BPN biases may not always work well, because the features in the backbone are not able to well represent the new domain. Our backbone is pretrained on ImageNet, which is appropriate for many image classification and visually-guided RL tasks in the natural world. However, the latent features may not be well suited for highly artificial worlds. This was recently addressed by (Evci et al., 2022), who showed that this problem can be alleviated using a last layer that connects to several intermediary layers, or even to every layer in the network, as opposed to only the penultimate layer. Hence, instead of sharing the last layer, we may share a so-called *Head2toe* layer when a large domain shift is encountered. Note that AR will also be used in this case as it is another way to counter large domain shifts: the AR pattern essentially recasts an input from a very different domain back into the ImageNet domain, then allowing the frozen backbone to extract rich and meaningful features in that domain. Also see Parisi et al. (2022) for ideas similar to Head2toe, with applications in RL.
- **Adversarial reprogramming (AR) (Elsayed et al., 2018):** Adversarial reprogramming is quite similar in spirit to BB, with the main difference being that it operates in the input (image) space as opposed to BB operating in the activation space. In adversarial reprogramming, one computes a single noise pattern for each task. This pattern is then added to inputs for a new task and fed through the original network. The original network processes the combined input + noise and generates an

Note: TM = Task Mapper

Dataset Name	Sample Dataset: # of classes	Sample Dataset: Train # of images	Sample Dataset: Validation # of images	Sample Dataset: Test # of images	Full Dataset: # of classes	Full Dataset: Train # of images	Full Dataset: Validation # of images	Full Dataset: Test # of images	Perfect Task Mapper: no BB Accuracy	Perfect Task Mapper: BB Accuracy	GMMC: TM Accuracy	GMMC: no BB Accuracy	GMMC: BB Accuracy	MAHA: TM Accuracy	MAHA: no BB Accuracy	MAHA: BB Accuracy
102_Category_Flower_Dataset	102	6455	865	865	102	6455	865	865	93.64%	95.95%	92.95%	87.63%	89.36%	95.38%	90.29%	92.02%
MIT_Indoor_Scenes	67	12350	1587	1587	67	12350	1587	1587	76.69%	77.88%	71.90%	56.96%	57.72%	85.63%	67.49%	68.30%
Caltech-UCSD_Birds_200	200	9405	1191	1191	200	9405	1191	1191	74.22%	74.06%	96.47%	71.70%	71.70%	99.50%	74.06%	73.97%
Stanford_Cars	196	12723	1693	1693	196	12723	1693	1693	53.93%	63.85%	97.22%	53.16%	62.67%	99.65%	53.69%	63.67%
Fine-Grained_Visual_Classification_of_Aircraft	70	8000	1000	1000	70	8000	1000	1000	46.10%	70.60%	99.20%	45.80%	70.30%	99.20%	45.90%	70.50%
VOC_2012_Human_Action_Subset	11	3424	435	435	11	3424	435	435	71.95%	73.79%	70.80%	56.78%	56.09%	67.36%	52.87%	52.64%
Chars74k	62	8902	1150	1150	62	8902	1150	1150	59.57%	70.26%	84.70%	52.61%	61.91%	86.78%	54.43%	64.09%
Stanford_Street_View_House_Numbers	10	54000	6000	6000	10	504323	63047	63047	74.35%	95.10%	93.18%	69.28%	88.65%	89.55%	66.58%	85.20%
iNaturalist_Reptilia	313	53360	5930	5930	313	71751	9026	9026	39.83%	42.07%	84.01%	35.97%	37.77%	88.72%	37.98%	40.03%
iNaturalist_Fungi	341	53161	6121	6121	341	74541	9409	9409	57.34%	61.36%	88.76%	51.23%	55.14%	93.89%	54.88%	58.98%
iNaturalist_Amphibia	170	38220	4815	4815	170	38220	4815	4815	37.20%	39.50%	82.78%	33.25%	35.31%	81.77%	33.79%	35.76%
iNaturalist_Arachnida	153	33674	4245	4245	153	33674	4245	4245	52.49%	56.09%	84.66%	46.15%	48.95%	85.82%	48.20%	50.91%
iNaturalist_Mollusca	169	36921	4659	4659	169	36921	4659	4659	65.06%	67.91%	84.67%	56.62%	59.33%	87.55%	59.61%	61.79%
iNaturalist_Actinopterygii	183	37195	4723	4723	183	37195	4723	4723	61.00%	63.10%	89.86%	56.55%	58.48%	92.25%	58.27%	60.41%
iNaturalist_Insecta	500	54000	6000	6000	2526	549151	69336	69336	66.23%	71.30%	71.87%	50.20%	54.17%	88.22%	62.27%	66.85%
CORe50	10	54000	6000	6000	10	130390	16308	16308	97.85%	99.27%	98.52%	96.42%	97.80%	96.65%	94.68%	95.93%
Sketches	250	15999	2000	2000	250	15999	2000	2000	59.70%	65.80%	98.55%	58.50%	64.55%	99.65%	59.45%	65.55%
WikiArt_Dataset	27	34522	4037	4037	27	62917	7880	7880	50.61%	54.47%	68.91%	34.33%	37.60%	69.93%	36.64%	38.92%
Describable_Textures_Dataset	47	4497	563	563	47	4497	563	563	72.47%	74.07%	64.83%	50.09%	51.69%	77.80%	59.86%	60.57%
GTSRB	43	41463	5184	5184	43	41463	5184	5184	96.66%	99.69%	97.78%	94.54%	97.51%	96.01%	92.92%	95.79%
CelebA	5	40566	4622	4622	5	99837	12483	12483	82.04%	91.19%	97.34%	80.46%	88.92%	95.61%	79.27%	87.43%
Office-Home_Clipart	65	3300	445	445	65	3300	445	445	77.53%	79.55%	61.12%	47.64%	50.11%	68.31%	55.51%	57.30%
Office-Home_Product	65	3361	454	454	65	3361	454	454	90.97%	89.87%	45.59%	42.51%	42.29%	71.81%	68.94%	69.16%
Office-Home_Art	65	1793	252	252	65	1793	252	252	80.95%	79.76%	34.13%	29.76%	29.76%	48.02%	42.86%	43.65%
Food-101	101	54035	5959	5959	101	80738	10100	10100	66.64%	68.43%	78.54%	52.69%	53.83%	81.66%	56.37%	57.61%
EuroSAT	10	21600	2700	2700	10	21600	2700	2700	94.37%	96.15%	95.81%	90.70%	92.26%	94.48%	89.48%	91.04%
PatchCamelyon	2	54000	6000	6000	2	221983	27750	27750	87.63%	88.65%	98.63%	86.38%	87.35%	77.02%	67.30%	67.80%
Diabetic_Retinopathy_Detection	5	28096	3515	3515	5	28096	3515	3515	75.07%	76.75%	93.89%	71.01%	72.48%	72.34%	54.43%	55.80%
RVL-CDIP	16	54000	6000	6000	16	318696	39846	39846	69.47%	76.33%	97.15%	67.47%	74.37%	96.73%	67.18%	74.00%
HistAerial	7	50637	5763	5763	7	109933	13747	13747	86.73%	87.65%	97.78%	84.87%	85.82%	94.88%	82.58%	83.65%
OrigamiSet1.0	3	1193	151	151	3	1193	151	151	82.12%	83.44%	86.75%	70.86%	70.86%	83.44%	66.89%	67.55%
Brazilian_Coins	5	2443	308	308	5	2443	308	308	85.06%	93.18%	97.73%	83.12%	91.56%	99.35%	85.06%	92.86%
iMaterialist_Fashion_2019	46	36117	4537	4537	46	36117	4537	4537	28.72%	29.42%	79.15%	22.94%	23.21%	73.68%	20.08%	20.50%
Rice_Image_Dataset	5	54000	6000	6000	5	59761	7471	7471	99.38%	100.00%	99.87%	99.25%	99.87%	99.95%	99.33%	99.95%
Vegetable_images_Dataset	15	16796	2100	2100	15	16796	2100	2100	100.00%	100.00%	95.14%	95.14%	95.14%	93.57%	93.57%	93.57%
garbage_classification	12	12383	1555	1555	12	12383	1555	1555	94.98%	95.63%	62.19%	59.42%	60.39%	48.87%	46.69%	46.82%
Facial_Expression_Recognition_2013	7	27175	3401	3401	7	27175	3401	3401	53.87%	60.01%	98.06%	52.78%	58.78%	96.30%	52.04%	58.13%
7000_Labeled_Pokemon	150	5305	749	749	150	5305	749	749	69.83%	83.04%	81.04%	58.74%	68.09%	84.78%	63.95%	73.16%
Manga_Facial_Expressions	7	357	49	49	7	357	49	49	53.06%	55.10%	65.31%	34.69%	36.73%	95.92%	48.98%	51.02%
10_Monkey_Species	10	1007	130	130	10	1007	130	130	98.46%	99.23%	86.15%	85.38%	85.38%	100.00%	98.46%	99.23%
Oregon_Wildlife	20	5648	714	714	20	5648	714	714	88.38%	87.68%	91.60%	82.77%	82.07%	89.50%	81.65%	81.23%
Blood_Cell_Images_Dataset	4	10005	1254	1254	4	10005	1254	1254	83.33%	87.00%	100.00%	83.33%	87.00%	99.84%	83.17%	86.84%
Retinal_OCT_2017	4	42158	4896	4896	4	61339	7669	7669	85.11%	89.44%	99.88%	85.01%	89.32%	99.29%	84.50%	88.79%
APTOS_2019_Blindness_Detection	5	2798	353	353	5	2798	353	353	79.60%	79.32%	88.95%	70.54%	71.10%	92.35%	72.52%	71.67%
Cataract_Dataset	4	479	61	61	4	479	61	61	60.66%	60.66%	65.57%	45.90%	47.54%	95.08%	59.02%	60.66%
Freiburg_Groceries_Dataset	25	3921	506	506	25	3921	506	506	78.66%	79.05%	94.66%	74.90%	75.30%	95.06%	75.69%	75.69%
Fashion_Product_Images_Dataset	43	34885	4384	4384	43	34885	4384	4384	95.87%	96.44%	71.92%	69.41%	69.80%	63.16%	60.22%	60.61%
Apparel_Images_Dataset	24	9080	1146	1146	24	9080	1146	1146	83.33%	94.42%	70.77%	60.82%	67.02%	73.47%	63.44%	70.51%
Zalando_Clothing_and_Models	6	8527	1070	1070	6	8527	1070	1070	88.88%	90.75%	89.44%	79.07%	81.12%	80.47%	71.50%	72.62%
PlantDoc-Dataset	27	2017	265	265	27	2017	265	265	59.62%	62.64%	74.72%	45.66%	47.55%	79.25%	47.17%	51.70%
Images_LEGO_Bricks	50	32000	4000	4000	50	32000	4000	4000	81.88%	90.05%	99.35%	81.28%	89.43%	99.53%	81.43%	89.58%
Art_Images_Type_Classification	5	5684	713	713	5	5684	713	713	88.64%	90.88%	67.32%	61.85%	62.69%	65.92%	61.85%	62.41%
Multi-Class_Weather_Dataset	4	875	112	112	4	875	112	112	92.86%	98.21%	65.18%	64.29%	64.29%	72.32%	69.64%	72.32%
Simpsons_Characters_Data	42	17464	2209	2209	42	17464	2209	2209	72.88%	87.37%	96.51%	70.67%	84.97%	94.66%	70.35%	84.07%
Intel_Image_Classification	6	13597	1703	1703	6	13597	1703	1703	92.48%	92.60%	83.21%	77.33%	77.75%	84.20%	77.80%	78.21%
House_Room_Image_Dataset	5	4136	519	519	5	4136	519	519	88.63%	88.63%	86.51%	76.88%	76.88%	68.02%	60.50%	60.50%
UIUC_Sports_Event_Dataset	8	1258	160	160	8	1258	160	160	100.00%	99.37%	69.38%	69.38%	68.75%	73.13%	73.13%	72.50%
Land-Use_Scene_Classification	21	8399	1050	1050	21	8399	1050	1050	96.00%	96.29%	86.86%	83.43%	83.52%	88.48%	85.05%	85.33%
ASL_Alphabets_Dataset	29	53998	6003	6003	29	69600	8700	8700	99.78%	99.95%	99.60%	99.38%	99.55%	99.78%	99.57%	99.73%
Yoga-82	82	15486	1982	1982	82	15486	1982	1982	52.72%	64.78%	86.63%	47.68%	57.82%	83.40%	46.97%	56.66%
Russian_Letter_Dataset	33	30097	3785	3785	33	30097	3785	3785	81.51%	93.34%	99.05%	80.74%	92.47%	99.10%	80.98%	92.52%
UMIST_Face_Database	20	788	112	112	20	788	112	112	100.00%	100.00%	91.96%	91.96%	91.96%	100.00%	100.00%	100.00%
iFood2019	251	53815	6009	6009	251	104150	13159	13159	42.69%	44.55%	72.14%	30.82%	32.15%	85.52%	37.28%	38.92%
Oxford_Buildings	11	665	90	90	11	665	90	90	85.56%	83.33%	74.44%	68.89%	65.56%	90.00%	78.89%	76.67%
Texture_Dataset	64	6808	861	861	64	6808	861	861	99.65%	99.88%	90.59%	90.36%	90.48%	98.03%	97.68%	97.91%
electronic-components	36	8087	1033	1033	36	8087	1033	1033	47.82%	46.08%	88.38%	42.01%	41.14%	88.67%	42.21%	41.24%
Hurricane_Damage_Dataset	2	16838	2106	2106	2	16838	2106	2106	95.82%	97.91%	96.44%	92.64%	94.54%	88.79%	85.61%	87.32%
chest_xray	2	4658	583	583	2	4658	583	583	96.74%	97.60%	99.49%	96.23%	97.08%	97.08%	93.83%	94.68%
PAD-UFES-20	6	1807	231	231	6	1807	231	231	64.50%	67.10%	74.89%	49.35%	50.22%	85.71%	54.98%	57.14%
Brain_Tumor_Dataset	4	2292	289	289	4	2292	289	289	87.89%	92.39%	97.92%	86.16%	90.66%	99.65%	87.54%	92.04%
Kannada-MNIST	10	48000	6000	6000	10	48000	6000	6000	98.12%	99.42%	99.60%	97.73%	99.05%	99.90%	98.02%	99.32%
Breast_Ultrasound	3	620	79	79	3	620	79	79	82.28%	83.54%	84.81%	69.62%	72.15%	96.20%	79.75%	81.01%
BookCover30	30	45575	5700	5700	30	45575	5700	5700	26.46%	27.89%	83.44%	21.21%	22.49%	82.89%	21.47%	22.68%
boat-types-recognition	9	1160	150	150	9	1160	150	150	93.33%	93.33%	74.67%	72.00%	71.33%	76.67%	73.33%	72.67%
rock-classification	7	1620	206	206	7	1620	206	206	75.73%	76.21%	57.77%	45.15%	44.17%	63.11%	49.03%	48.54%
dermnet	23	14238	1794	1794	23	14238	1794	1794	44.20%	45.82%	95.21%	41.64%	43.42%	91.36%	40.25%	41.14%
dragon-ball-super-saiyan-dataset	6	112	18	18	6	112	18	18	50.00%	55.56%	22.22%	22.22%	16.67%	94.44%	44.44%	50.00%
concrete-crack	2	30720	3841	3841	2	30720	3841	3841	99.84%	99.92%	98.54%	98.44%	98.49%	95.81%	95.68%	95.76%
Malacca_Historical_Buildings	3	126	18	18	3	126	18	18	100.00%	100.00%	38.89%	38.89%	38.89%	100.00%	100.00%	100.00%
Satellite_Images_to_Predict_African_Poverty	4	20455	2558	2558	4	20455	2558	2558	77.01%	81.04%	90.34%	69.62%	73.53%	79.75%	62.16%	65.01%
Skin_Cancer_MNIST_HAM10000	7	8003	1005	1005	7	8003	1005	1005	77.91%	77.31%	96.42%	75.82%	75.02%	89.95%	70.95%	70.15%
watermarked-not-watermarked-images	2	24917	3115	3115	2	24917	3115	3115	70.88%	82.02%	54.70%	39.07%	45.04%	29.82%	21.25%	24.69%
Large-Scale_Fish_Dataset	9	7526	941	941	9	7526	941	941	99.89%	100.00%	98.51%	98.41%	98.51%	97.56%	97.45%	97.56%
DeepWeedsX	9	13996	1756	1756	9	13996	1756	1756	81.66%	84.17%	94.99%	77.68%	79.84%	91.80%	74.83%	76.88%
IP102_Dataset	102	36302	4303	4303	102	60089	7564	7564	59.08%	59.49%	69.77%	41.90%	42.57%	66.77%	41.78%	42.41%
planets-and-moons-dataset	11	1304	165	165	11	1304	165	165	99.39%	100.00%	96.97%	96.97%	96.97%	99.39%	98.79%	99.39%
polish-craft-beer-labels	100	6275	848	848	100	6275	848	848	99.41%	100.00%	99.29%	98.70%	99.29%	100.00%	99.41%	100.00%
Kvasir-Capsule_Dataset	14	30208	3784	3784	14	30208	3784	3784	93.97%	96.51%	99.84%	93.90%	96.43%	98.84%	92.97%	95.40%
NEU-Surface-Defect-Database	6	1439	180	180	6	1439	180	180	100.00%	100.00%	87.78%	87.78%	87.78%	99.44%	99.44%	99.44%
colorectal-histology-mnist	8	3992	504	504	8	3992	504	504	88.69%	87.50%	94.25%	84.13%	82.74%	97.02%	85.71%	84.72%
100_Sports	100	11558	1500	1500	100	11558	1500	1500	92.27%	92.73%	78.07%	72.20%	72.13%	92.40%	86.33%	86.20%
Labeled_Surgical_Tools_and_Images	5	2400	302	302	5	2400	302	302	85.76%	94.70%	97.02%	83.11%	92.05%	92.05%	78.48%	87.75%
Mechanical_Tools_Classification_Dataset	8	5861	737	737	8	5861	737	737	91.18%	90.64%	79.10%	73.81%	73.00%	62.14%	58.21%	57.26%
Galaxy10	10	14083	1766	1766	10	14083	1766	1766	63.53%	77.01%	99.55%	63.08%	76.61%	99.60%	63.25%	76.73%
Stanford_Online_Products	12	54000	6000	6000	12	94375	11804	11804	81.57%	80.82%	72.95%	62.50%	62.10%	53.45%	45.30%	44.78%
NWPU-RESISC45	45	25192	3150	3150	45	25192	3150	3150	86.57%	89.78%	84.89%	72.83%	75.87%	93.40%	80.79%	83.81%
FaceMask_Dataset	3	11001	1377	1377	3	11001	1377	1377	98.55%	98.77%	88.96%	87.65%	88.02%	83.22%	82.21%	82.57%
OnePiece_Dataset	18	9191	1156	1156	18	9191	1156	1156	78.11%	84.69%	86.07%	67.30%	73.10%	84.95%	67.65%	73.18%
ilab_80m	15	12000	1500	1500	15	12000	1500	1500	90.33%	90.40%	96.53%	87.20%	87.13%	95.60%	86.33%	86.33%
CLEVR_v1.0	8	54000	6000	6000	8	67994	8503	8503	48.73%	73.65%	99.90%	48.68%	73.55%	99.97%	48.73%	73.63%
ilab_atari	67	52735	5921	5921	67	295048	36911	36911	99.97%	100.00%	99.09%	99.07%	99.09%	99.27%	99.24%	99.27%
ilab_deepvp	9	53911	6003	6003	9	59422	7433	7433	58.09%	90.07%	97.63%	57.45%	88.77%	90.35%	54.67%	83.47%
Total # / Average Accuracy	5033	2041225	243464	243464	7059	4126063	518043	518043	77.92%	81.57%	84.94%	67.43%	70.58%	87.10%	68.87%	72.10%

Figure S5: Basic stats of dataset and accuracy. TM=task mapper.

Dataset Name	EWC	PSP	ER	MAS	SI	LwF	Online-EWC	SUPSUP
102_Category_Flower_Dataset	0.58%	1.04%	57.23%	1.62%	0.92%	0.35%	0.46%	50.64%
MIT_Indoor_Scenes	3.21%	3.02%	9.01%	4.91%	2.71%	2.02%	1.89%	28.86%
Caltech-UCSD_Birds_200	0.25%	0.42%	21.24%	0.67%	0.50%	0.59%	0.34%	12.59%
Stanford_Cars	0.41%	0.59%	13.59%	0.65%	0.47%	0.41%	0.35%	6.44%
Fine-Grained_Visual_Classification_of_Aircraft	1.70%	1.90%	19.70%	5.80%	2.90%	2.40%	2.90%	24.10%
VOC_2012_Human_Action_Subset	9.20%	15.86%	2.76%	11.72%	15.86%	7.59%	7.59%	20.92%
Chars74k	0.96%	3.22%	22.00%	3.13%	1.48%	0.17%	0.70%	73.22%
Stanford_Street_View_House_Numbers	11.05%	9.98%	55.85%	9.52%	10.43%	9.32%	8.80%	91.93%
iNaturalist_Reptilia	0.44%	0.20%	2.56%	0.32%	0.34%	0.30%	0.32%	1.70%
iNaturalist_Fungi	0.31%	0.34%	13.67%	0.28%	0.29%	0.20%	0.42%	8.45%
iNaturalist_Amphibia	0.56%	0.69%	3.26%	0.66%	0.46%	0.69%	0.62%	4.11%
iNaturalist_Arachnida	0.64%	0.73%	6.12%	0.97%	0.64%	0.90%	0.49%	5.91%
iNaturalist_Mollusca	0.75%	0.64%	11.05%	0.69%	0.56%	0.67%	0.71%	8.97%
iNaturalist_Actinopterygii	0.70%	0.51%	9.65%	0.83%	0.38%	0.42%	0.57%	5.27%
iNaturalist_Insecta	0.27%	0.17%	14.73%	0.25%	0.10%	0.33%	0.17%	3.77%
CORe50	10.35%	10.37%	19.03%	10.22%	11.77%	9.25%	10.17%	81.62%
Sketches	0.25%	0.65%	43.60%	0.40%	0.45%	0.05%	0.45%	19.65%
WikiArt_Dataset	3.12%	6.69%	4.58%	6.47%	1.41%	2.65%	4.24%	26.13%
Describable_Textures_Dataset	1.95%	1.95%	12.08%	4.26%	1.42%	1.60%	1.95%	20.60%
GTSRB	3.67%	4.26%	74.46%	5.34%	3.03%	4.57%	2.91%	86.67%
CelebA	12.12%	25.98%	50.58%	25.94%	25.96%	20.45%	10.80%	80.44%
Office-Home_Clipart	1.35%	1.57%	21.12%	3.15%	1.57%	1.57%	1.80%	39.55%
Office-Home_Product	2.20%	1.10%	26.21%	1.32%	2.20%	1.76%	0.44%	44.05%
Office-Home_Art	4.76%	1.59%	3.97%	4.37%	1.19%	3.97%	3.17%	9.92%
Food-101	1.12%	1.04%	8.00%	1.16%	1.26%	1.16%	1.19%	11.90%
EuroSAT	12.22%	10.41%	50.37%	20.96%	11.11%	7.85%	8.26%	84.11%
PatchCamelyon	12.60%	49.75%	41.08%	51.27%	52.48%	3.48%	7.28%	83.58%
Diabetic_Retinopathy_Detection	52.96%	32.09%	15.99%	72.40%	70.15%	26.15%	5.88%	72.57%
RVL-CDIP	6.33%	7.72%	29.40%	5.97%	6.33%	4.15%	8.07%	65.43%
HistAerial	8.47%	24.08%	49.44%	23.10%	15.50%	10.93%	1.94%	63.28%
OrigamiSet1.0	36.42%	34.44%	6.62%	50.33%	23.18%	16.56%	13.25%	43.05%
Brazilian_Coins	16.56%	4.55%	51.62%	21.43%	19.81%	1.62%	3.57%	93.18%
iMaterialist_Fashion_2019	1.43%	2.07%	1.68%	17.57%	0.42%	1.17%	0.82%	24.71%
Rice_Image_Dataset	14.48%	51.83%	82.18%	33.10%	20.00%	19.65%	8.72%	99.15%
Vegetable_images_Dataset	9.90%	11.14%	51.71%	6.95%	7.90%	5.86%	9.29%	95.90%
garbage_classification	4.50%	33.57%	5.72%	37.49%	5.40%	6.69%	5.98%	68.75%
Facial_Expression_Recognition_2013	12.88%	24.26%	26.64%	17.82%	17.17%	14.64%	11.76%	44.46%
7000_Labeled_Pokemon	0.40%	0.93%	53.81%	0.80%	0.67%	0.40%	0.67%	43.93%
Manga_Facial_Expressions	12.24%	18.37%	40.82%	22.45%	14.29%	16.33%	14.29%	30.61%
10_Monkey_Species	9.23%	1.54%	43.85%	11.54%	9.23%	7.69%	9.23%	43.08%
Oregon_Wildlife	6.02%	6.58%	12.89%	5.04%	5.88%	6.30%	4.76%	31.65%
Blood_Cell_Images_Dataset	21.93%	17.86%	35.41%	24.96%	25.20%	14.27%	15.95%	93.70%
Retinal_OCT_2017	11.76%	31.41%	57.56%	30.92%	29.78%	15.77%	29.13%	86.66%
APTOS_2019_Blindness_Detection	42.78%	51.56%	43.63%	51.84%	17.85%	20.96%	5.95%	67.99%
Cataract_Dataset	16.39%	47.54%	50.82%	39.34%	16.39%	9.84%	19.67%	49.18%
Freiburg_Groceries_Dataset	7.31%	3.75%	22.13%	8.50%	5.14%	4.55%	4.94%	44.07%
Fashion_Product_Images_Dataset	0.25%	34.08%	40.08%	36.95%	4.88%	0.57%	0.59%	43.25%
Apparel_Images_Dataset	4.28%	8.99%	30.37%	6.11%	8.55%	4.89%	5.15%	82.02%
Zalando_Clothing_and_Models	8.97%	41.40%	32.34%	46.07%	16.26%	24.11%	19.72%	72.43%
PlantDoc-Dataset	5.28%	4.15%	13.96%	7.17%	3.77%	3.77%	3.77%	18.11%
Images_LEGO_Bricks	1.68%	3.33%	28.90%	1.73%	2.18%	2.70%	2.60%	4.45%
Art_Images_Type_Classification	13.46%	33.94%	8.42%	37.59%	16.83%	21.46%	10.10%	76.72%
Multi-Class_Weather_Dataset	17.86%	43.75%	25.00%	45.54%	32.14%	24.11%	6.25%	83.93%
Simpsons_Characters_Data	2.04%	3.98%	22.45%	10.82%	5.93%	2.22%	2.90%	77.14%
Intel_Image_Classification	12.16%	31.24%	21.96%	20.26%	15.68%	9.92%	6.28%	80.50%
House_Room_Image_Dataset	21.19%	26.40%	10.79%	19.27%	17.73%	26.20%	18.69%	47.78%
UIUC_Sports_Event_Dataset	10.00%	20.00%	20.63%	28.75%	13.13%	12.50%	10.00%	61.25%
Land-Use_Scene_Classification	7.52%	7.05%	34.38%	14.00%	4.76%	3.81%	3.52%	67.90%
ASL_Alphabets_Dataset	4.45%	32.13%	58.35%	5.43%	3.28%	3.28%	4.11%	96.39%
Yoga-82	1.21%	1.06%	12.46%	2.57%	1.46%	0.91%	1.21%	27.70%
Russian_Letter_Dataset	2.99%	30.12%	44.76%	3.57%	1.59%	6.58%	4.83%	86.92%
UMIST_Face_Database	2.68%	24.11%	89.29%	5.36%	6.25%	4.46%	3.57%	99.11%
iFood2019	0.37%	0.78%	6.57%	0.63%	0.45%	0.22%	0.28%	4.58%
Oxford_Buildings	10.00%	25.56%	50.00%	38.89%	38.89%	4.44%	6.67%	52.22%
Texture_Dataset	1.74%	3.48%	80.60%	6.27%	1.05%	1.51%	1.28%	91.99%
electronic-components	2.81%	2.71%	9.68%	4.74%	3.19%	4.07%	3.78%	24.10%
Hurricane_Damage_Dataset	4.51%	85.57%	48.81%	65.34%	34.24%	20.47%	45.77%	94.82%
chest_xray	7.38%	71.36%	63.81%	72.90%	27.10%	6.86%	2.23%	95.71%
PAD-UFES-20	12.99%	32.47%	23.38%	31.60%	17.75%	11.69%	1.73%	50.65%
Brain_Tumor_Dataset	13.15%	37.37%	59.52%	30.10%	26.64%	8.30%	12.46%	79.93%
Kannada-MNIST	11.73%	98.97%	93.65%	10.02%	10.00%	12.10%	9.57%	98.85%
Breast_Ultrasound	0.00%	59.49%	45.57%	17.72%	54.43%	29.11%	13.92%	70.89%
BookCover30	3.67%	4.33%	3.56%	4.46%	2.81%	3.02%	3.77%	12.56%
boat-types-recognition	18.00%	22.00%	8.00%	34.00%	17.33%	13.33%	9.33%	34.00%
rock-classification	11.65%	26.21%	6.80%	24.27%	17.48%	9.71%	10.19%	31.55%
dermnet	5.02%	5.57%	13.10%	5.07%	2.79%	3.85%	3.07%	26.42%
dragon-ball-super-saiyan-dataset	22.22%	27.78%	33.33%	0.00%	16.67%	27.78%	16.67%	33.33%
concrete-crack	8.10%	98.54%	81.31%	49.00%	45.77%	1.17%	0.83%	99.11%
Malacca_Historical_Buildings	5.56%	77.78%	88.89%	33.33%	33.33%	22.22%	27.78%	66.67%
Satellite_Images_to_Predict_African_Poverty	7.66%	28.46%	29.36%	29.44%	21.23%	8.80%	1.64%	64.19%
Skin_Cancer_MNIST_HAM10000	1.49%	61.49%	31.74%	66.77%	66.37%	2.69%	15.32%	71.94%
watermarked-not-watermarked-images	13.32%	57.37%	1.06%	54.64%	50.75%	9.47%	4.01%	79.78%
Large-Scale_Fish_Dataset	11.37%	51.33%	74.18%	13.28%	10.95%	9.56%	10.41%	96.28%
DeepWeedsX	17.65%	53.02%	29.33%	51.88%	51.20%	10.59%	35.88%	65.95%
IP102_Dataset	0.49%	3.25%	8.58%	1.86%	1.46%	0.65%	0.91%	14.36%
planets-and-moons-dataset	7.88%	43.64%	96.36%	11.52%	7.88%	9.70%	6.67%	100.00%
polish-craft-beer-labels	1.53%	11.91%	95.87%	2.83%	1.53%	0.59%	0.12%	95.40%
Kvasir-Capsule_Dataset	0.63%	77.80%	36.73%	72.60%	25.08%	0.37%	1.14%	90.12%
NEU-Surface-Defect-Database	10.00%	73.89%	81.11%	32.78%	16.67%	6.67%	7.22%	93.33%
colorectal-histology-mnist	6.35%	28.37%	64.68%	31.75%	12.50%	8.73%	4.96%	89.68%
100_Sports	1.93%	4.73%	28.27%	3.80%	1.47%	0.93%	1.67%	40.20%
Labeled_Surgical_Tools_and_Images	12.25%	32.78%	53.31%	45.03%	24.83%	11.26%	13.58%	90.40%
Mechanical_Tools_Classification_Dataset	5.29%	26.73%	6.65%	32.16%	25.64%	9.77%	7.33%	43.83%
Galaxy10	11.61%	14.84%	39.30%	12.29%	14.95%	5.95%	11.49%	51.02%
Stanford_Online_Products	9.12%	26.77%	8.83%	20.92%	8.47%	8.47%	7.92%	39.28%
NWPU-RESISC45	1.97%	23.30%	41.40%	11.81%	2.25%	2.25%	2.03%	53.08%
FaceMask_Dataset	13.94%	80.46%	75.16%	62.31%	29.99%	8.50%	9.44%	94.34%
OnePiece_Dataset	8.22%	25.17%	22.15%	9.86%	5.71%	7.27%	6.83%	55.10%
ilab_80m	8.87%	40.27%	46.00%	26.33%	7.60%	7.73%	7.93%	67.00%
CLEVR_v1.0	10.12%	60.70%	29.35%	18.32%	20.47%	11.03%	13.17%	65.72%
ilab_atari	6.23%	99.93%	99.71%	14.63%	1.45%	7.50%	5.57%	99.27%
ilab_deepvp	91.84%	88.59%	93.02%	90.60%	41.41%	89.91%	92.09%	84.41%
Average Accuracy	**8.86%**	**25.49%**	**35.32%**	**20.54%**	**13.89%**	**8.41%**	**7.77%**	**56.22%**

Figure S6: Accuracy of baselines after all 102 tasks have been learned in the order shown.

output, which is then remapped onto the desired output domain. Unfortunately, the CPU cost of this approach is prohibitive with respect to $0.5N$ speedup.

We denote the number of BB biases by $\mathcal{N}$ in what follows (for xception, $\mathcal{N} = 17,472$). If Head2toe connects to the same feature maps as BB, then the number of weights is $\mathcal{N} \times c$ for c output classes. We assume that each task is modeled with k GMMC clusters ($k = 25$ currently), and each is represented by a 2048D mean and 2048D diagonal covariance. We denote by 4 the number of bytes per floating point number.

For a classification task with c classes: An agent receives an image as input and produces a vector of c output values (on SKILL-102, c is 49.34 on average), where the highest output value is the most likely image class for the input image (Table S1).

Shared params and data	Size (bytes)	Implemented: $\mathcal{N} = 17,472, c = 49.34, k = 25$
Last layer weights	$2048 \times c \times 4$	404 KBytes
BB biases	$\mathcal{N} \times 4$	70 KBytes
GMMC clusters	$k \times (2048 + 2048) \times 4$	409 KBytes
Optional: Head2toe	adds $\mathcal{N} \times c \times 4$	adds 3.45 MBytes
Optional: AR pattern	adds $299 \times 299 \times 3$	adds 268 KBytes
Alternative: 5 images/task for MAHA	$5 \times 299 \times 299 \times 3$	1.34 MBytes

Table S1: *Total average sharing per task in our current implementation with GMMC+BB: 404+70+409 = 883 KBytes/task; for Mahalanobis+BB: 404+70+1341=1.81 MBytes/task.*

H GMMC visual explanation

A visual explanation of how GMMC works in LLL agents is shown in Fig. S7.

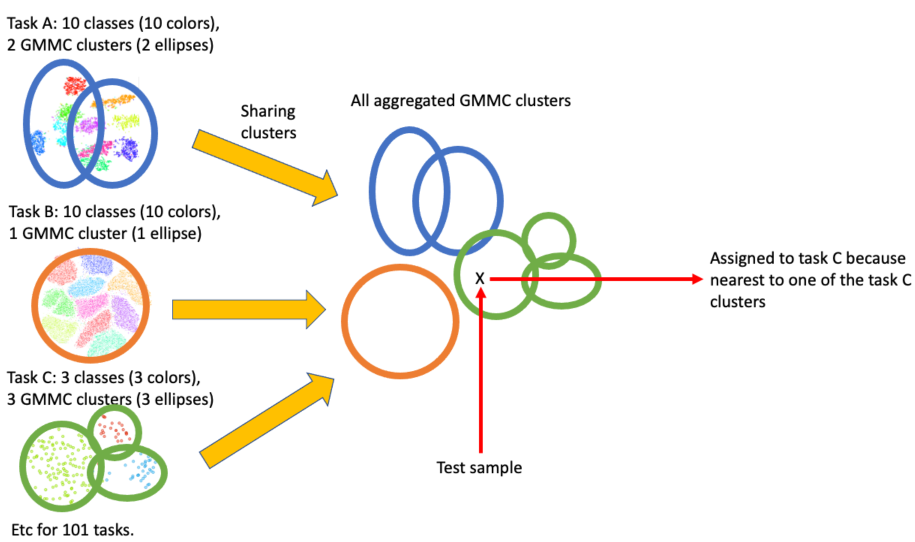

Figure S7: GMMC task mapper. (left) Each teacher clusters its entire training set into a number of Gaussian clusters. Here, a variable number of clusters is shown for each task, but in our results we use 25 clusters for every task. Each teacher then shares the mean and diagonal covariance of its clusters with all students. (right) Students just aggregate all received clusters in a bank, keeping track of which task any given cluster comes from. At test time, a sample is evaluated against all clusters received so far, and the task associated with the cluster closest to the test sample is chosen.

I Pairs of similar classes according to CLIP

Table S2, Table S3, Table S4, and Table S5 show examples of pairs of similar classes according to CLIP embedding. The first and the second column are the names of similar class pairs from two different tasks (i.e iFood2019 and Food-101). The third column is the cosine similarity score between the CLIP embeddings of the name of the class pairs.

Table S2: Matched Class for MIT_Indoor_Scene and House_Room_Images

	learned_class(weight source)	target_class	score
0	Dinning	dining_room	0.9106
1	Kitchen	kitchen	0.9995
2	Bathroom	bathroom	1.0
3	Bedroom	bedroom	1.0
4	Livingroom	livingroom	1.0

Table S3: Matched Class for Office-Home_Product and Standford_Online_Products

	learned_class(weight source)	target_class	score
0	stapler	Paper_Clip	0.757
1	toaster	Oven	0.838
2	coffee	Mug	0.882
3	cabinet	File_Cabinet	0.9126
4	lamp	Lamp_Shade	0.9453
5	sofa	Couch	0.9736
6	bicycle	Bike	0.978
7	mug	Mug	1
8	chair	Chair	1
9	fan	Fan	1
10	kettle	Kettle	1

Table S4: Matched Class for UIUC_Sports_Event_Dataset and 199_Sports
labellong

	learned_class(weight source)	target_class	score
0	bocce	bowling	0.7837
1	badminton	tennis	0.823
2	sailing	sailboat racing	0.9
3	snowboarding	snow boarding	0.9355
4	RockClimbing	rock climbing	0.989
5	polo	polo	0.9995
6	croque_madame	croque_madame	1
7	Rowing	rowing	1

Table S5: Food-101 vs iFood2019

	learned_class(weight source)	target_class	score
0	cheese_plate	grilled_cheese_sandwich	0.8574
1	cup_cakes	cupcake	0.904

2	steak	steak_au_poivre	0.9185
3	scallops	scallop	0.929
4	breakfast_burrito	burrito	0.939
5	nachos	nacho	0.9517
6	dumplings	dumpling	0.9536
7	mussels	mussel	0.955
8	churros	churro	0.9585
9	spring_rolls	spring_roll	0.9585
10	chicken_wings	chicken_wing	0.9604
11	escargots	escargot	0.962
12	waffles	waffle	0.966
13	baby_back_ribs	baby_back_rib	0.966
14	oysters	oyster	0.9663
15	beignets	beignet	0.97
16	tacos	taco	0.9727
17	donuts	donut	0.9756
18	crab_cakes	crab_cake	0.9756
19	deviled_eggs	deviled_egg	0.976
20	macarons	macaron	0.9775
21	pancakes	pancake	0.982
22	pad_thai	pad_thai	0.999
23	grilled_salmon	grilled_salmon	0.999
24	fried_calamari	fried_calamari	0.999
25	omelette	omelette	0.9995
26	beef_carpaccio	beef_carpaccio	0.9995
27	hamburger	hamburger	0.9995
28	clam_chowder	clam_chowder	0.9995
29	chocolate_cake	chocolate_cake	0.9995
30	lobster_roll_sandwich	lobster_roll_sandwich	0.9995
31	macaroni_and_cheese	macaroni_and_cheese	0.9995
32	seaweed_salad	seaweed_salad	0.9995
33	shrimp_and_grits	shrimp_and_grits	0.9995
34	sushi	sushi	0.9995
35	creme_brulee	creme_brulee	0.9995
36	sashimi	sashimi	0.9995
37	cheesecake	cheesecake	0.9995
38	chicken_curry	chicken_curry	0.9995
39	fried_rice	fried_rice	0.9995
40	pork_chop	pork_chop	0.9995
41	bruschetta	bruschetta	0.9995
42	edamame	edamame	0.9995
43	cannoli	cannoli	0.9995
44	caprese_salad	caprese_salad	0.9995
45	red_velvet_cake	red_velvet_cake	0.9995
46	spaghetti_bolognese	spaghetti_bolognese	1
47	spaghetti_carbonara	spaghetti_carbonara	1
48	takoyaki	takoyaki	1
49	tiramisu	tiramisu	1
50	tuna_tartare	tuna_tartare	1

J Performance on Visual Domain Decathlon

We also perform our methods on a well-known benchmark Visual Domain Decathlon (Ke et al., 2020) in Fig. S8. The baselines and our method implementations are the same as the experiments in SKILL-102 dataset.

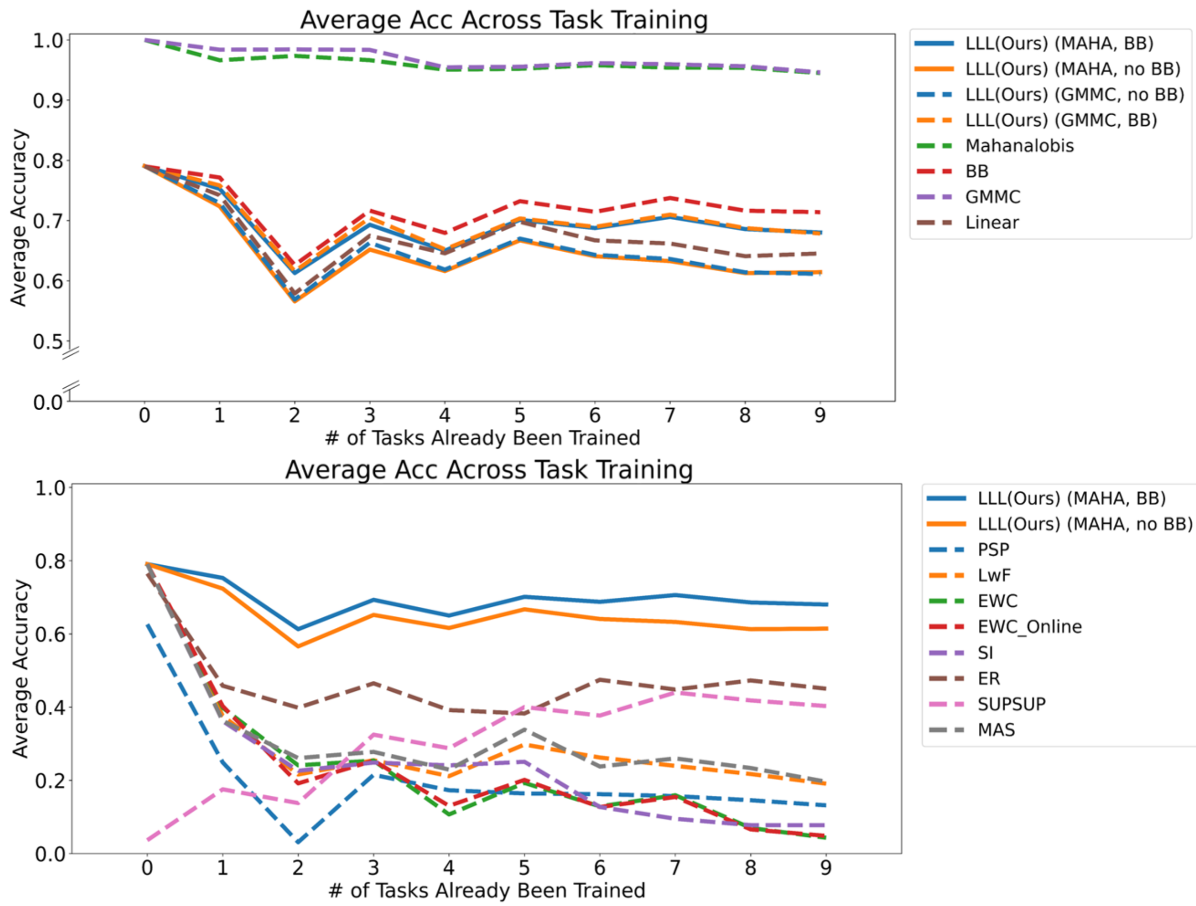

Figure S8: Average absolute accuracy on 10 Visual Domain Decathlon tasks learned so far, as a function of the number of tasks learned.

References

Gamaleldin F Elsayed, Ian Goodfellow, and Jascha Sohl-Dickstein. Adversarial reprogramming of neural networks. *arXiv preprint arXiv:1806.11146*, 2018.

Utku Evci, Vincent Dumoulin, Hugo Larochelle, and Michael C Mozer. Head2toe: Utilizing intermediate representations for better transfer learning. *arXiv preprint arXiv:2201.03529*, 2022.

Zixuan Ke, Bing Liu, and Xingchang Huang. Continual learning of a mixed sequence of similar and dissimilar tasks. *Advances in Neural Information Processing Systems*, 33:18493–18504, 2020.

Simone Parisi, Aravind Rajeswaran, Senthil Purushwalkam, and Abhinav Gupta. The unsurprising effectiveness of pre-trained vision models for control. *arXiv preprint arXiv:2203.03580*, 2022.