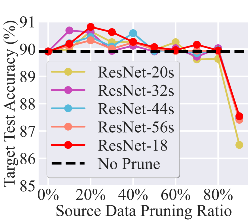
Table R1: **Experiments with MoCov3 on ViT** (@Reviewer RSvb). Other settings follow Tab. 2 of the original submission. The performance surpassing the unpruned setting (pruning ratio 0%) is highlighted in cyan. The best result in each setting is marked in **bold**. FM consistently outperforms other baselines and can find winning subsets with pruning ratios of more than 50%.

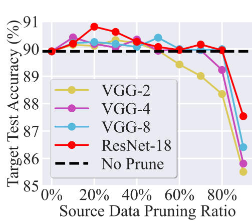
Dataset / Pruning Ratio	0%	50%	60%	70%	80%
OxfordPets					
RANDOM	87.34	82.13	80.27	75.42	68.34
MODERATE		86.17	85.29	84.01	81.32
GRAND		87.21	86.11	83.19	80.78
FM (ours)		**87.68**	**87.51**	**87.39**	**84.14**
SUN397					
RANDOM	60.36	59.22	58.45	56.73	54.29
MODERATE		60.13	59.61	58.21	56.44
GRAND		60.39	59.27	58.95	57.15
FM (ours)		**60.49**	**60.55**	**60.42**	**59.88**
Flowers102					
RANDOM	93.96	92.41	91.65	90.17	88.41
MODERATE		93.75	92.41	91.42	90.11
GRAND		93.88	93.21	91.77	90.45
FM (ours)		**94.11**	**94.28**	**93.97**	**91.42**

Table R2: **Sensitivity study on surrogate model size** (@Reviewer BXTc, @Reviewer vmFL). Experiments follow the setting of Fig. 4: RN-101 is first pretrained on the pruned source dataset (ImageNet) based on the surrogate model, and then finetuned on the downstream task OxfordPets. Under different surrogate models, the source class selection overlapping ratio with the used surrogate model RN-18 in the submission is reported under 50% pruning ratio.

Surrogate Model Architecture	RN-20s	VGG-2	RN-32s	VGG-4	RN-44s	RN-56s	VGG-8	RN-18 (Default)
Param. # (M)	0.236	0.417	0.516	0.698	0.706	0.896	5.53	11.69
Source Acc. (%)	36.25	22.56	40.77	29.44	43.74	45.72	58.45	68.73
Largest Pruning Ratio of Winning Subsets (%)	60	50	80	70	80	80	80	80
Source Class Selection Overlap (%)	89.3	84.4	90.7	87.2	93.5	94.8	97.7	100



(a) ResNets family



(b) VGG family

Figure R1: Source dataset pruning trajectory given the downstream task OxfordPets using different surrogate models associated with **Tab. R2**. Experiment presentation protocols are the same as Fig. 4 in the original submission. Pruning trajectory of **ResNet-18** (the default surrogate model used in the submission) is plotted as a reference for comparison.

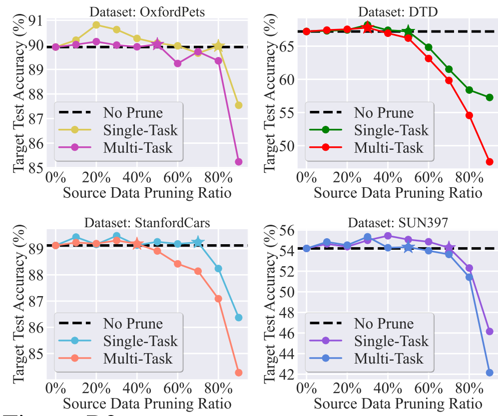


Figure R2: **DP achieved by LM in the multi-task setting given 4 downstream tasks** (@Reviewer BXTc, @Reviewer vmFL). This expands Fig. 4, *i.e.*, the single-task setting, where source data is pruned based on an individual task.

Table R3: Experiments on CIFAR-10C (@Reviewer omyc, @Reviewer vmFL). LM-based source dataset pruning on ImageNet (given CIFAR-10 as the downstream task) applies to transfer learning against CIFAR-10C. 5 out of the 19 corruption types are shown due to the space limit.

Dataset	Pruning Ratio 0%	20%	40%	60%	80%
CIFAR10	96.83	96.88	**97.03**	96.57	95.41
+ Gaussian Noise	82.13	82.67	**82.89**	82.60	81.19
+ Defocus Blur	84.73	85.22	**85.36**	84.92	82.75
+ Impulse Noise	84.62	85.21	84.78	**85.93**	85.11
+ Shot Noise	83.18	83.25	83.49	**83.76**	83.24
+ Speckle Noise	83.11	**83.59**	83.29	83.57	82.27

Table R4: **Experiments on the few-shot transfer learning benchmark VTAB** (@Reviewer omyc, @Reviewer vmFL). Seven tasks in the NATURAL set are studied following the setting of Fig. 4. Each task contains 800 training and 200 testing samples. LM can obtain winning subsets of the source dataset (ImageNet) with 40% pruning ratios for most tasks.

		● Caltech101	● CIFAR-100	● DTD	● Flowers102	● OxfordPets	● SUN397	● SVHN	● Mean
	No Prune	80.23	46.39	62.48	90.39	88.42	33.32	87.24	69.78
LM (Ours)	Pruning 20%	**80.39**	**46.65**	62.88	90.78	88.57	**34.22**	**87.45**	**70.13**
	Pruning 40%	80.34	46.49	**62.92**	**90.88**	**88.62**	33.81	87.21	70.04
	Pruning 60%	78.65	45.51	62.57	90.12	87.35	33.57	86.38	69.16
	Pruning 80%	73.31	41.33	61.18	89.42	85.11	31.49	84.22	66.58

Table R5: **Time consumption of dataset pruning of the SSL experiments in Tab. 2** (@Reviewer omyc). Each Experiments run on 8× Nvidia RTX A6000 GPUs. Time calculation follows Tab. 3. FM can save up to 168 hours without any downstream performance loss. Settings in cyan indicate settings where winning subsets are found for all the datasets in Tab. 2.

Pruning Ratio	0%	50%	60%	70%	80%
Time (h)	384	254	216	178	139