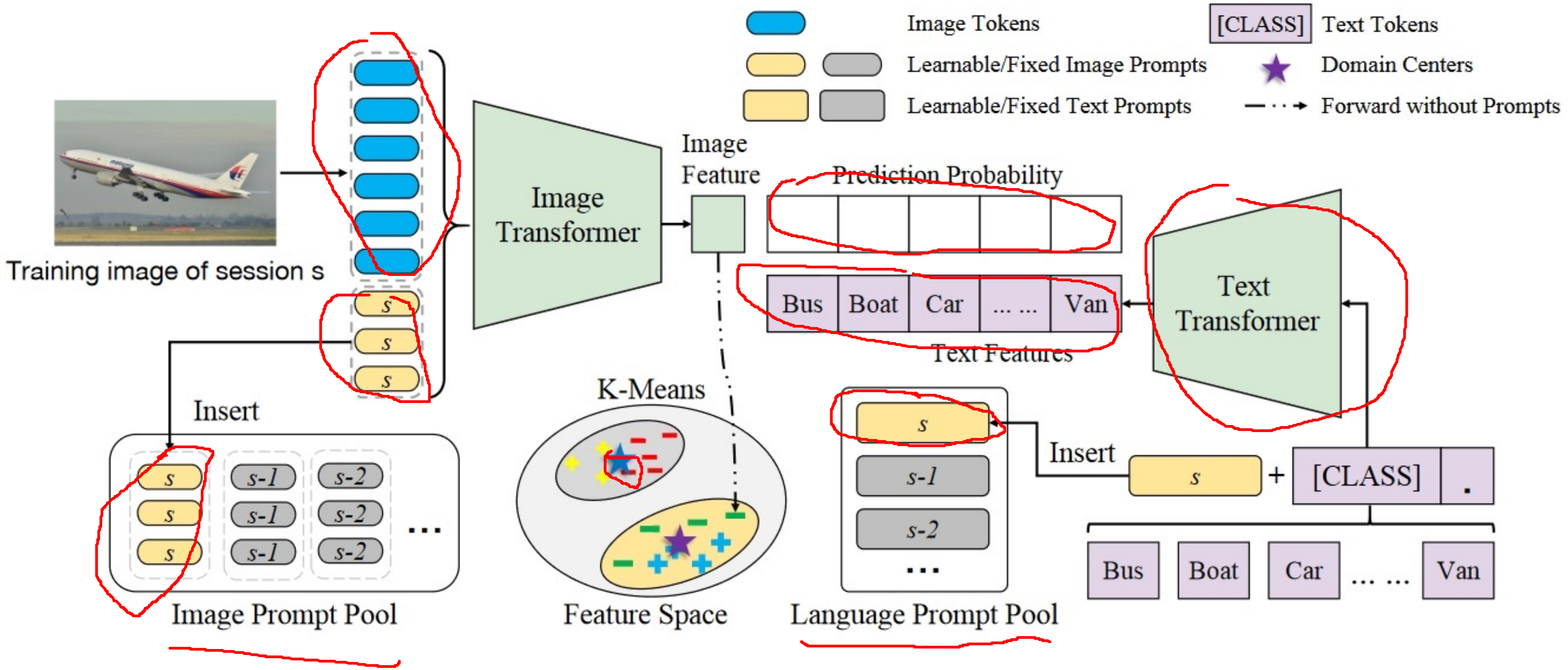
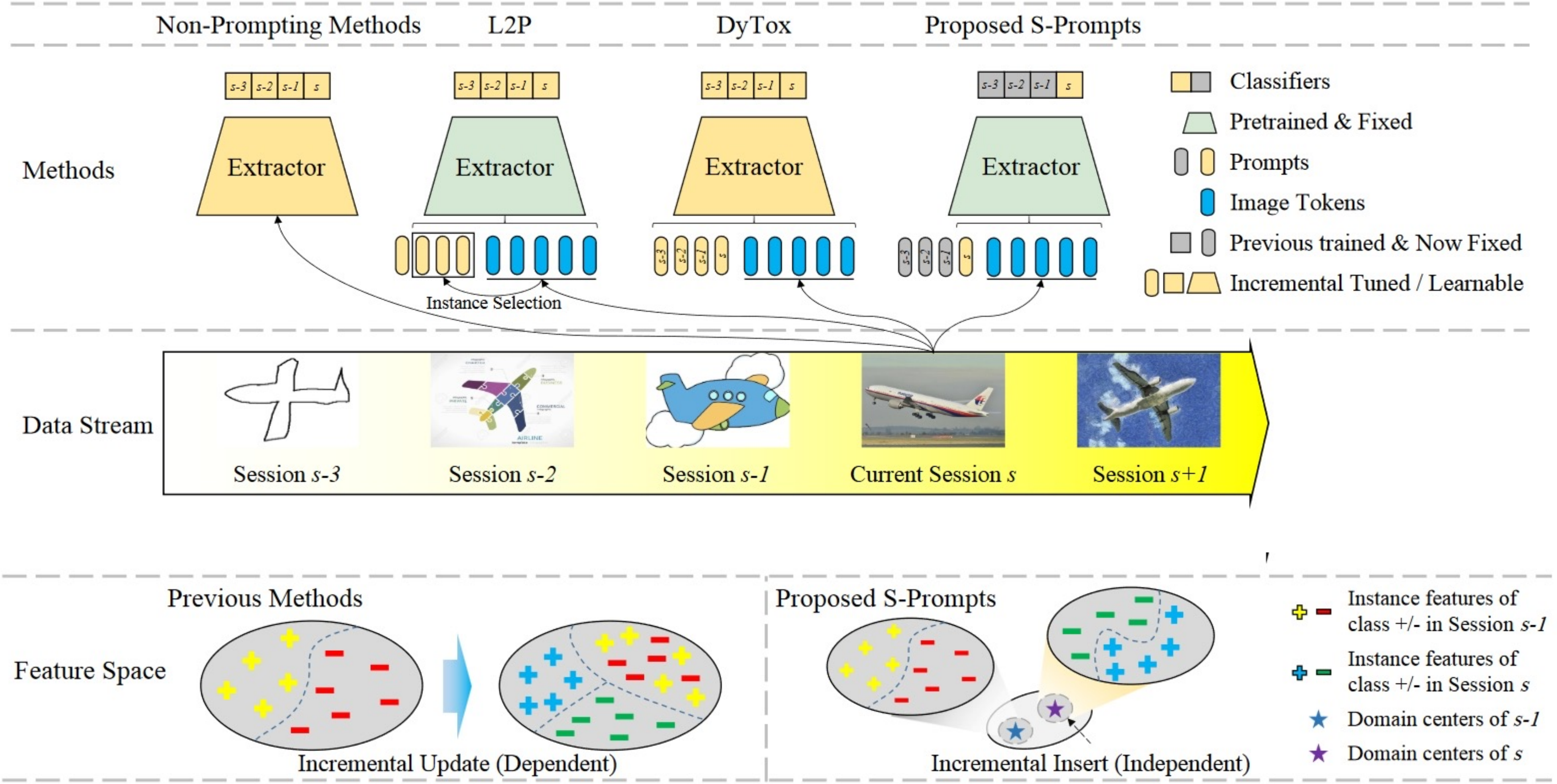


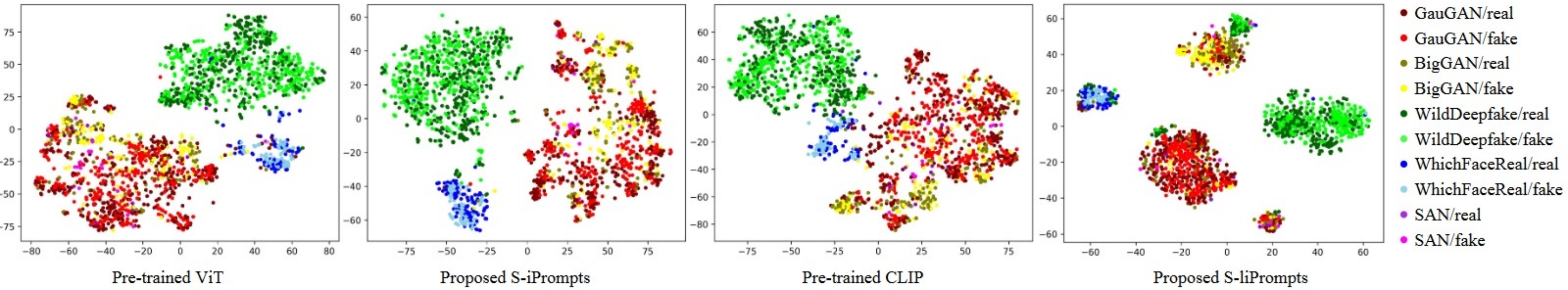
Image Tokens
Learnable/Fixed Image Prompts
Learnable/Fixed Text Prompts
[CLASS] Text Tokens
Domain Centers
Forward without Prompts
Training image of session s
Image Transformer
Image Feature
Prediction Probability
Bus
Boat
Car
... ...
Van
Text Features
Text Transformer
Insert
s
s-1
s-2
Image Prompt Pool
K-Means
Feature Space
Language Prompt Pool
[CLASS]


Non-Prompting Methods
L2P
DyTox
Proposed S-Prompts
Methods
Extractor
Extractor
Extractor
Extractor
s-3
s-2
s-1
s
Classifiers
Pretrained & Fixed
Prompts
Image Tokens
Previous trained & Now Fixed
Incremental Tuned / Learnable
Instance Selection
Data Stream
Session s-3
Session s-2
Session s-1
Current Session s
Session s+1
Previous Methods
Proposed S-Prompts
Feature Space
Incremental Update (Dependent)
Incremental Insert (Independent)
Instance features of class +/- in Session s-1
Instance features of class +/- in Session s
Domain centers of s-1
Domain centers of s

Method	Buffer size	Average Acc (↑)	Forgetting (↑)
LRCIL [42]	100/class	76.39*	-4.39*
iCaRL [40]		79.76*	-8.73*
LUCIR [20]		82.53*	-5.34*
LRCIL [42]	50/class	74.01*	-8.62*
iCaRL [40]		73.98*	-14.50*
LUCIR [20]		80.77*	-7.85*
DyTox [14]		86.21	-1.55
EWC [29]	0/class	50.59	-42.62
LwF [35]		60.94	-13.53
DyTox [14]		51.27	-45.85
L2P [55]		61.28	-9.23
S-iPrompts (ours)		<u>74.51</u>	<u>-1.30</u>
S-liPrompts (ours)		**88.65**	**-0.69**
Upper-bound (S-iPrompts)	-	85.50	-
Upper-bound (S-liPrompts)	-	91.91	-

Method	Buffer size	Average Acc (↑)
ER [8]	50/class	80.10±0.56*
GDumb [44]		74.92±0.25*
BiC [56]		79.28±0.30*
DER++ [4]		79.70±0.44*
Co^2L [6]		79.75±0.84*
DyTox [14]		79.21±0.10
L2P [55]		81.07±0.13*
EWC [29]	0/class	74.82±0.60*
LwF [35]		75.45±0.40*
L2P [55]		78.33±0.06*
S-iPrompts (ours)		<u>83.13</u>±0.51
S-liPrompts (ours)		**89.06**±0.86
Upper-bound (S-iPrompts)	-	84.01±0.53
Upper-bound (S-liPrompts)	-	93.19±0.21

Method	Buffer size	Average Acc (↑)
DyTox [14]	50/class†	62.94
EWC [29]		47.62
LwF [35]		49.19
SimCLR [10]-CaSSLe [17]		44.2*
BYOL [18]-CaSSLe [17]		49.7*
Barlow Twins[59]-CaSSLe [17]		48.9*
Supervised Contrastive [27]-CaSSLe [17]	0/class	<u>50.9*</u>
L2P [55]		40.15
S-iPrompts (ours)		50.62
S-liPrompts (ours)		**67.78**
Upper-bound (S-iPrompts)	-	63.22
Upper-bound (S-liPrompts)	-	64.08



CDDB
continual deepfake detection
5 domains GauGAN, BigGAN, WildDeepfake, WhichFaceReal, and SAN respectively
27000 images

CORe50
continual object recognition
It has 50 categories from 11 distinct domains. The continual learning setting uses 8 domains
120000

Domain Net
. Images from Domain Net is split into 6 domains
345 categories and roughly 600,000 images

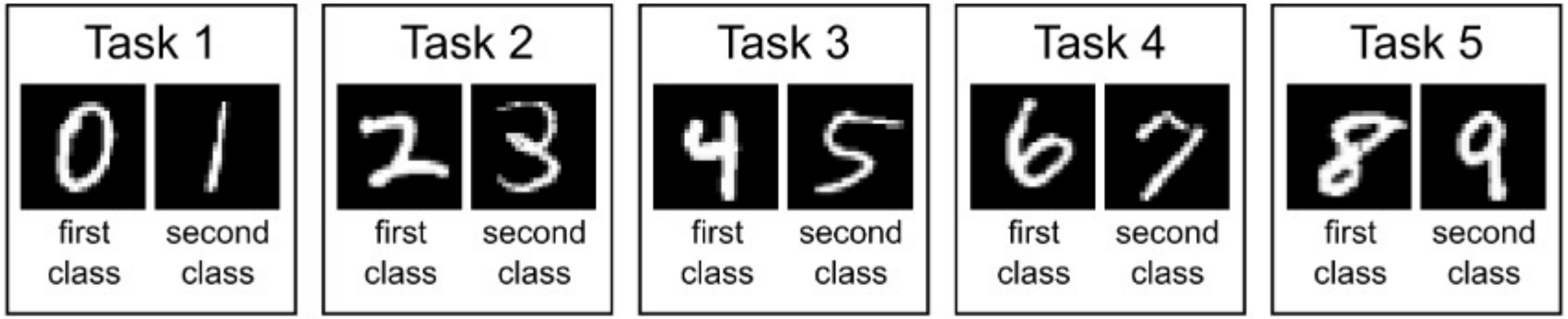
[1904.07734] Three scenarios for continual learning (arxiv.org)

Task IL

1.Each task has its own task-ID number.

2.The model can identify which task the current input data is from, making it the simplest scenario of continual learning.

Split MNIST 包含 5 个 task，task-ID 及 task 分别为 ID=1 {0,1}，ID=2 {2,3}，ID=3 {4,5}，ID=4 {6,7}，ID=5 {8,9}。给出 0 ~ 9 中的任意一个样本 x，首先需要给出 task-ID 向模型指明现在的预测任务是什么（比如给出 task-ID=3），那么模型会从 ID=3 {4,5} 这个任务中对样本 x 进行预测。针对 Task-IL 场景的连续学习模型能够为每个任务都训练一个专门的预测模块，典型架构就是 multi-headed 输出层。在 multi-headed 输出层每个任务都有专属自己的 output head，而网络较浅层（一般是特征提取层）通常是任务间共享的。



Domain IL

1.In Domain-IL, the model is not provided with a task-ID. Instead, it directly makes predictions on the incoming samples

2. In Domain-IL scenarios, tasks typically have the same structure but different input distributions (i.e., the task objectives remain constant, but the distribution of input data varies)

MNIST 和 Permuted MNIST 包含 0 ~ 9 所有数字。使用 MNIST 作为第一个任务，预测目标为 0 ~ 9；对 MNIST 进行 permutation，得到 Permuted MNIST 作为第二个任务，预测目标仍为 0 ~ 9；第三、第四 ⋯ 任务均基于 MNIST 进行不同的 permutation，这些任务的输入分布不同，但学习目标都是对样本 x 进行 0~9 的预测。

Scenario	Difference between D_{i-1} and D_i			Task-ID	Online
	$P(X_{i-1}) \neq P(X_t)$	$P(Y_{i-1}) \neq P(Y_i)$	$\{Y_{i-1}\} \neq \{Y_i\}$		
Task Incremental	✓	✓	✓	Train & Test	No
Class Incremental	✓	✓		No	Optional
Domain Incremental	✓			No	Optional

Class IL

1.Class-IL requires the model to infer which task the to-be-predicted sample x belongs to, and make predictions on x accordingly.

2. Class-IL models typically use a single-headed structure, meaning that predictions for all tasks utilize the same output layer.

Split MNIST 将数字 0 ~ 9 划分为 5 个任务，分别为 {0,1}，{2,3}，{4,5}，{6,7}，{8,9}。给出 0 ~ 9 中的任意一个样本 x，模型需要预测出这是“数字几”。预测“数字几”这个目标其实包含了两部分，模型要预测出现在是哪个任务（推断 task-ID）并在该任务中做出样本的正确分类。