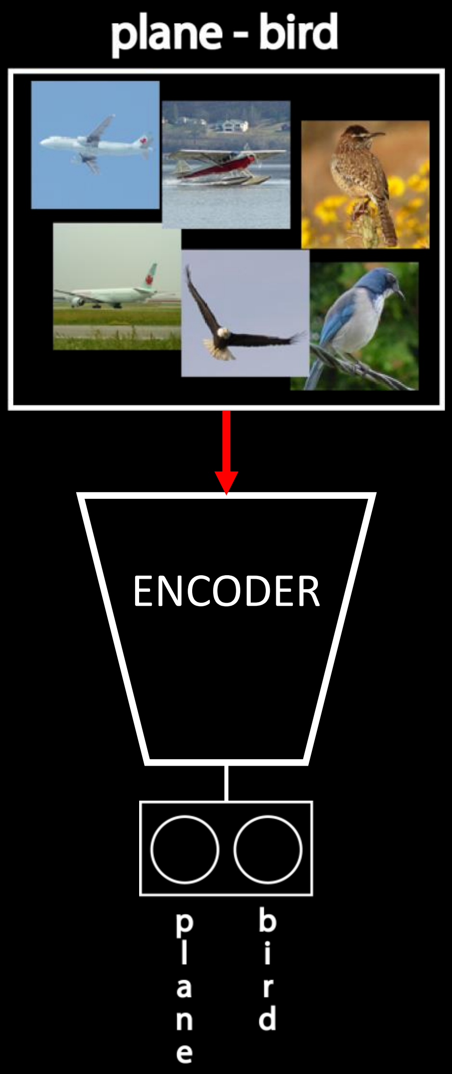
Semi-Supervised Class Incremental Learning

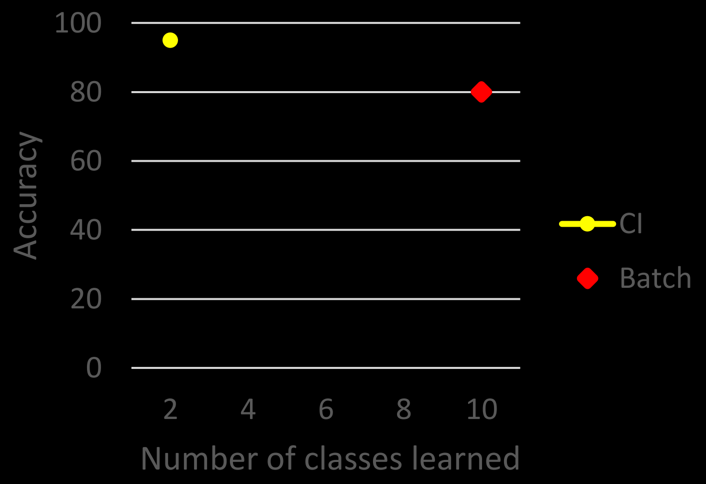
Alexis Lechat[1,2], Stéphane Herbin[1] and Frédéric Jurie[2]
[1]ONERA, [2]Normandie Université

ICPR 2020

Artificial Intelligence, Machine Learning for Pattern Analysis PS T1.16

Class Incremental (CI) Learning

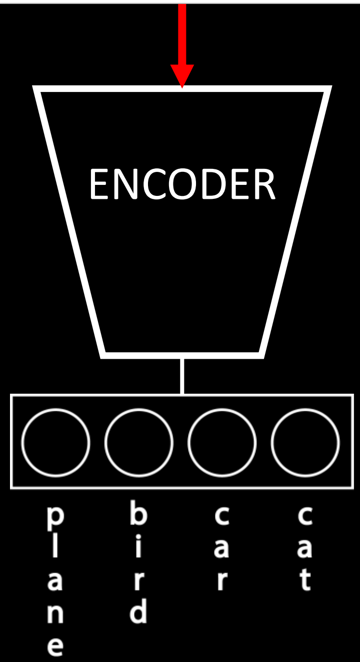




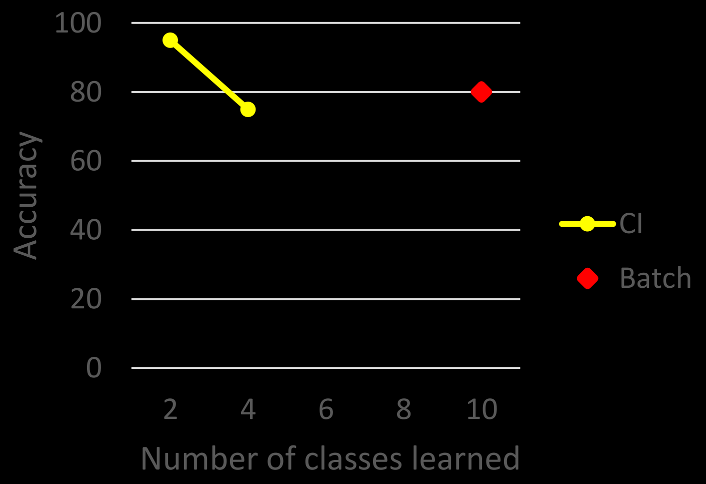
Class Incremental (CI) Learning

plane - bird

car - cat



deer - dog

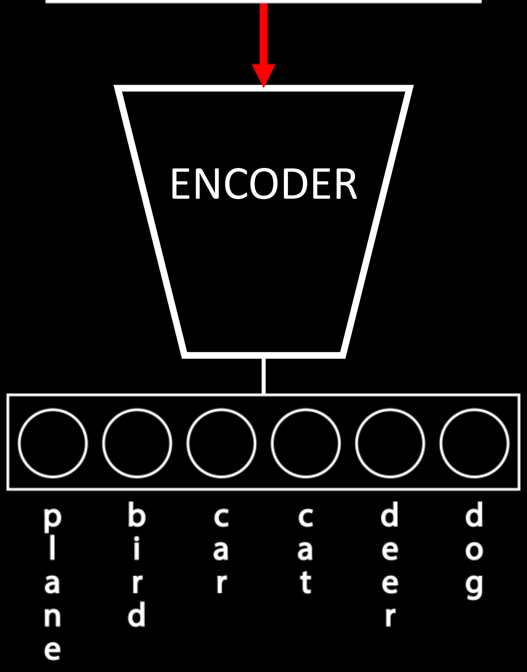


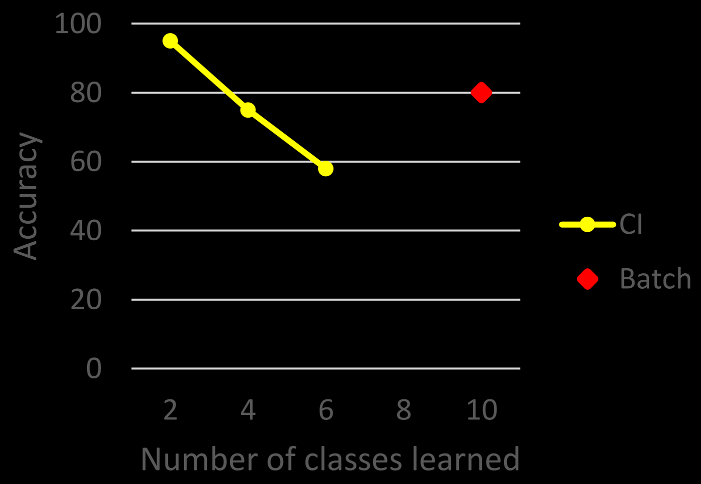
Class Incremental (CI) Learning

car - cat

deer - dog

horse - monkey



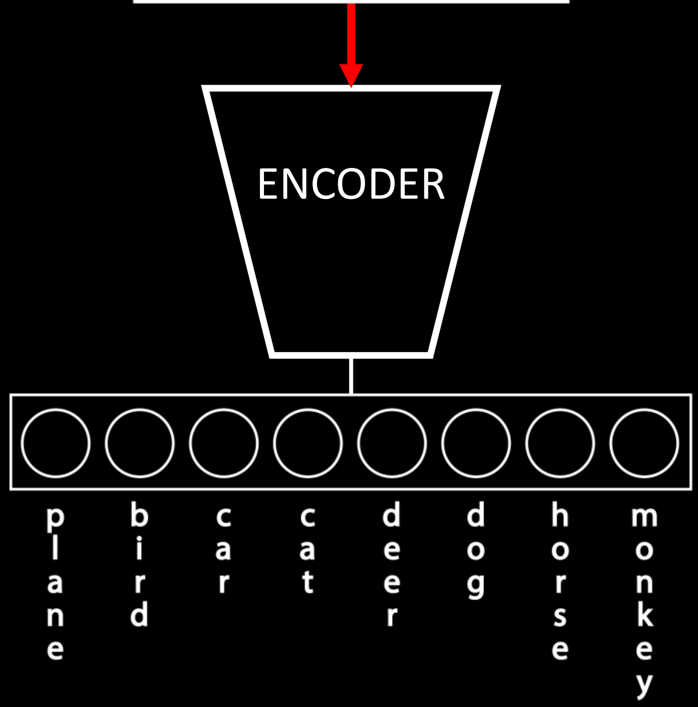


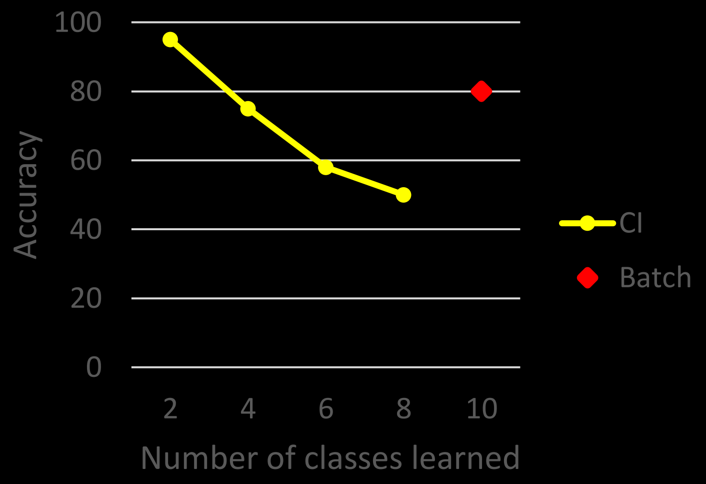
Class Incremental (CI) Learning

deer - dog

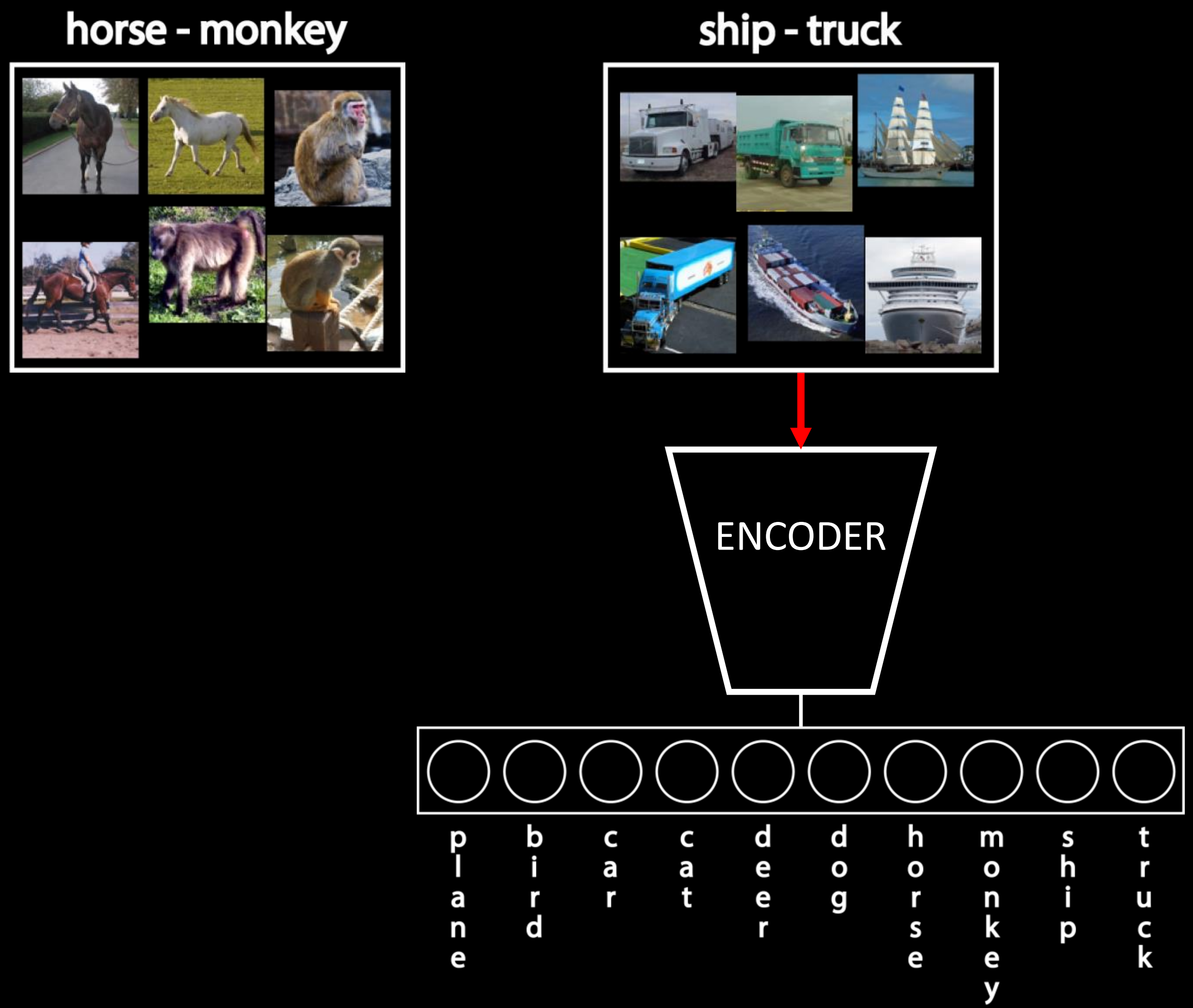
horse - monkey

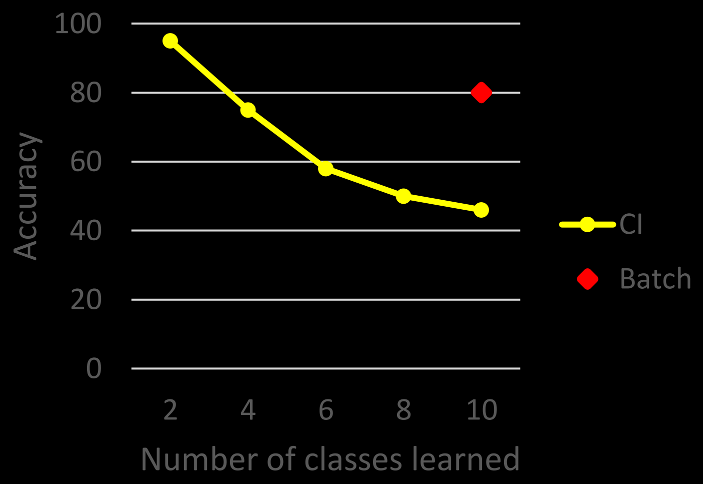
ship - truck



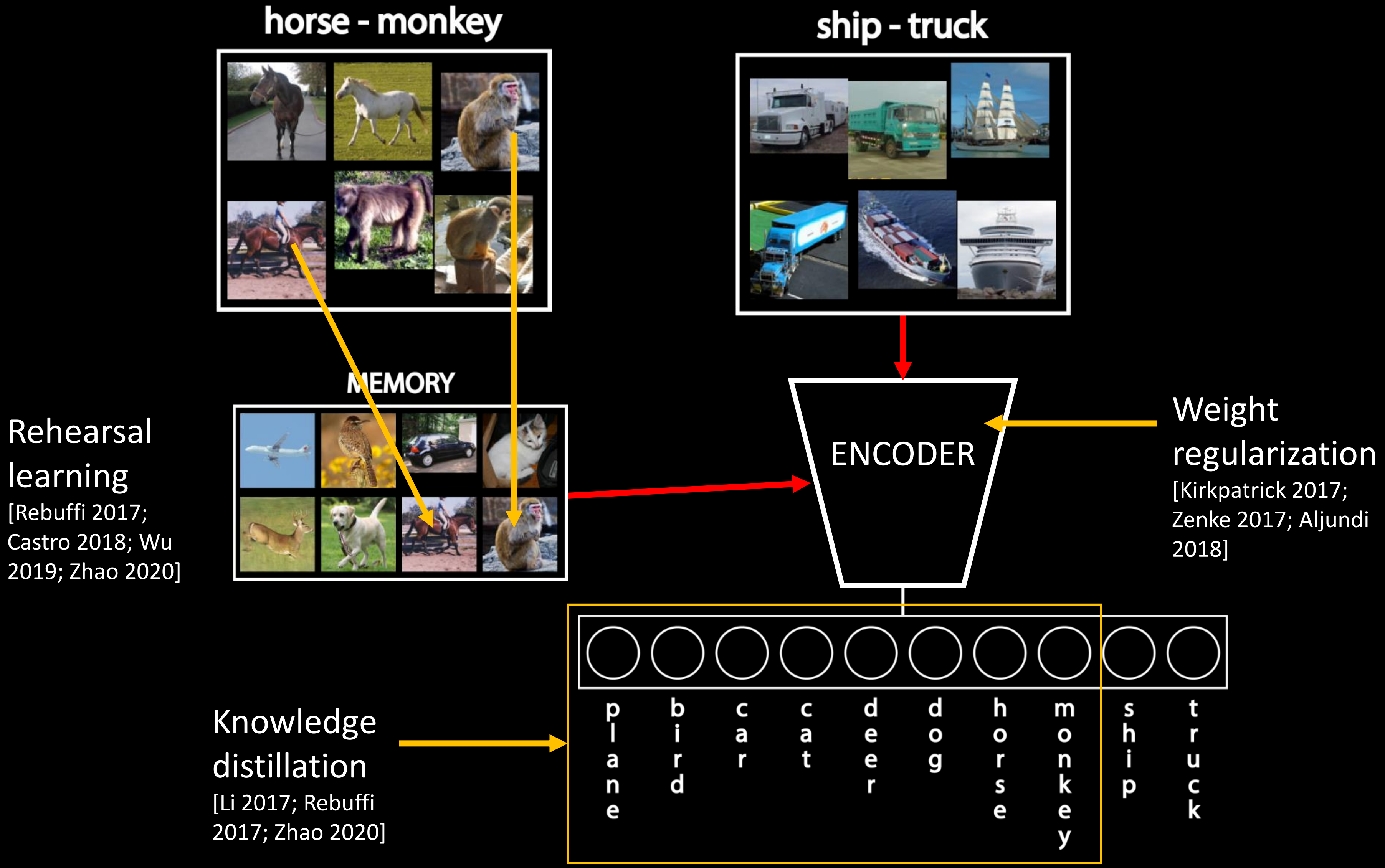


Class Incremental (CI) Learning



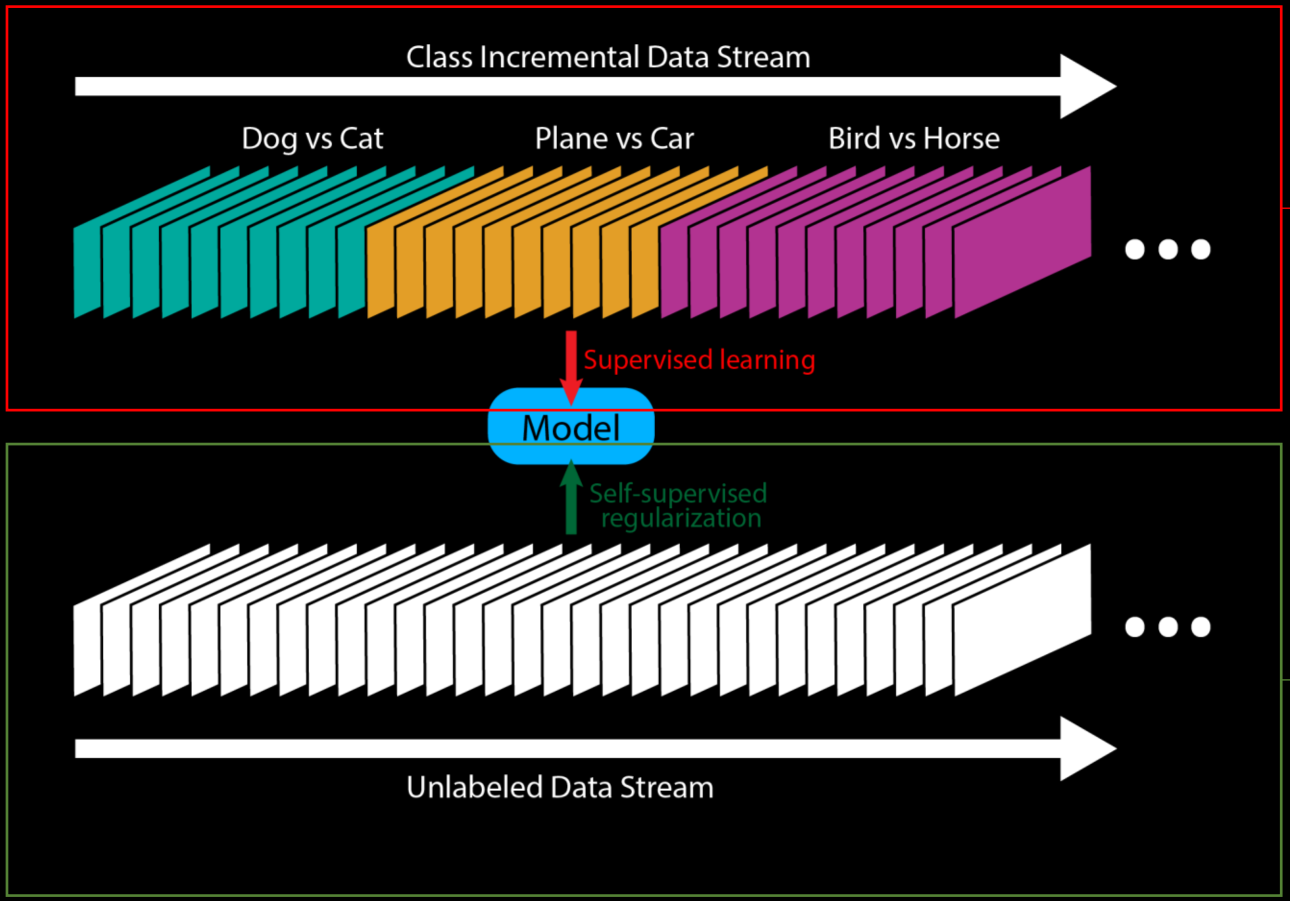


Prior works



horse - monkey
ship - truck
MEMORY
Rehearsal learning
[Rebuffi 2017; Castro 2018; Wu 2019; Zhao 2020]
ENCODER
Weight regularization
[Kirkpatrick 2017; Zenke 2017; Aljundi 2018]
Knowledge distillation
[Li 2017; Rebuffi 2017; Zhao 2020]
plane
bird
car
cat
deer
dog
horse
monkey
ship
truck

Our approach: Semi-Supervised Incremental Learning



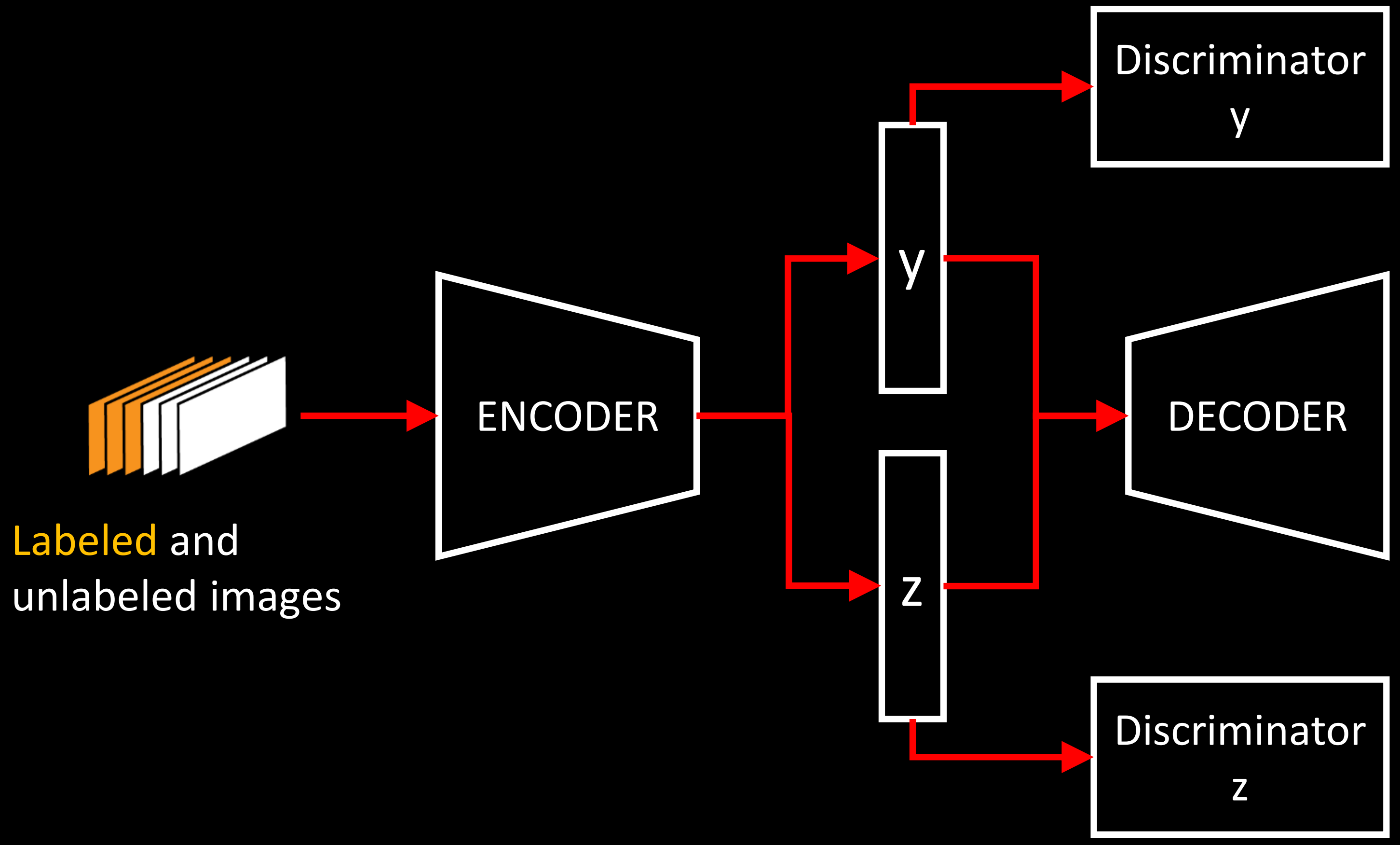
Standard CI process with rehearsal

Self-supervised training (pretext task) with cheap unlabeled data

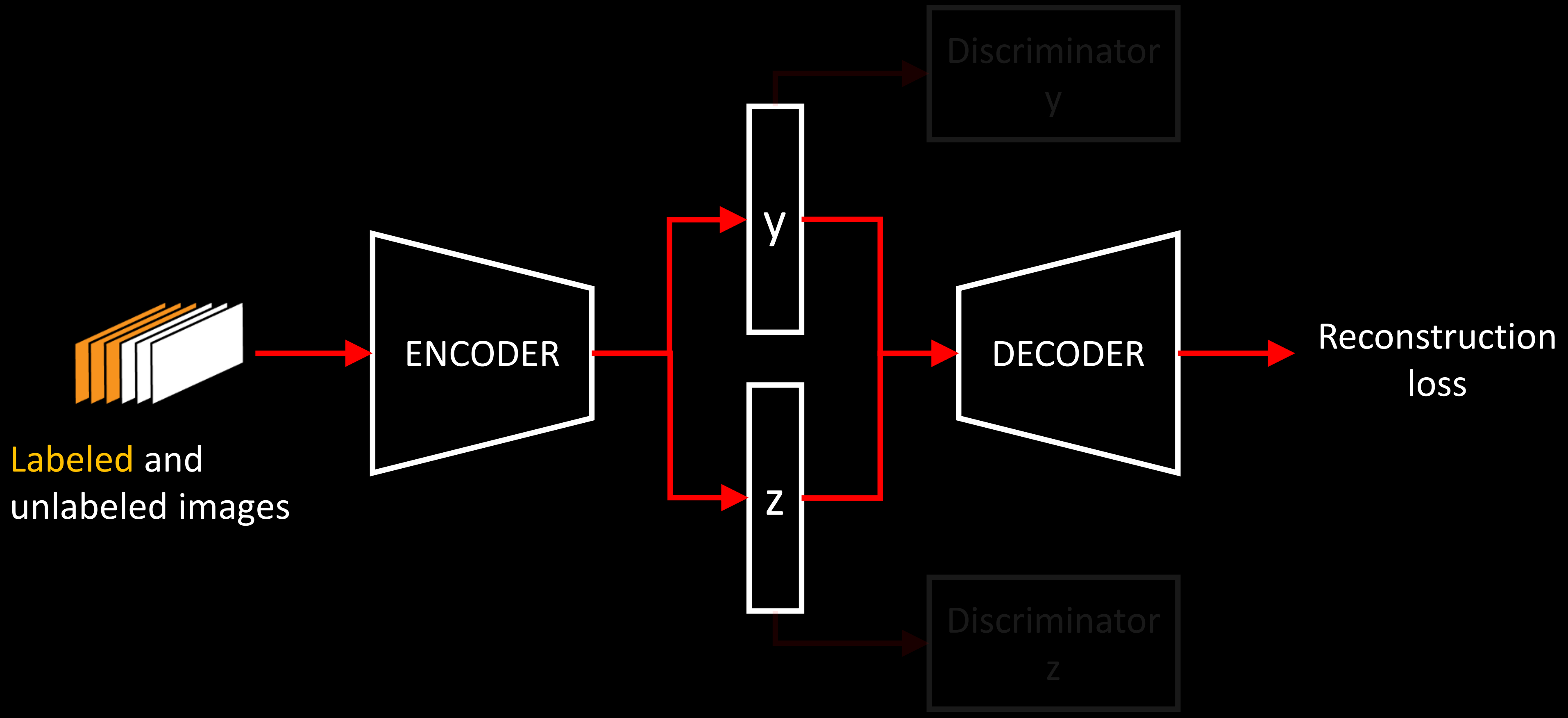
Objectives

- Profit from inexpensive unlabeled data to build a large self-supervised task
- Use the self-supervision as a regularization to alleviate the *Catastrophic Forgetting*
- Learn better representations for a more stable encoder/enhanced performances
- Further reduce the amount of labeled data needed

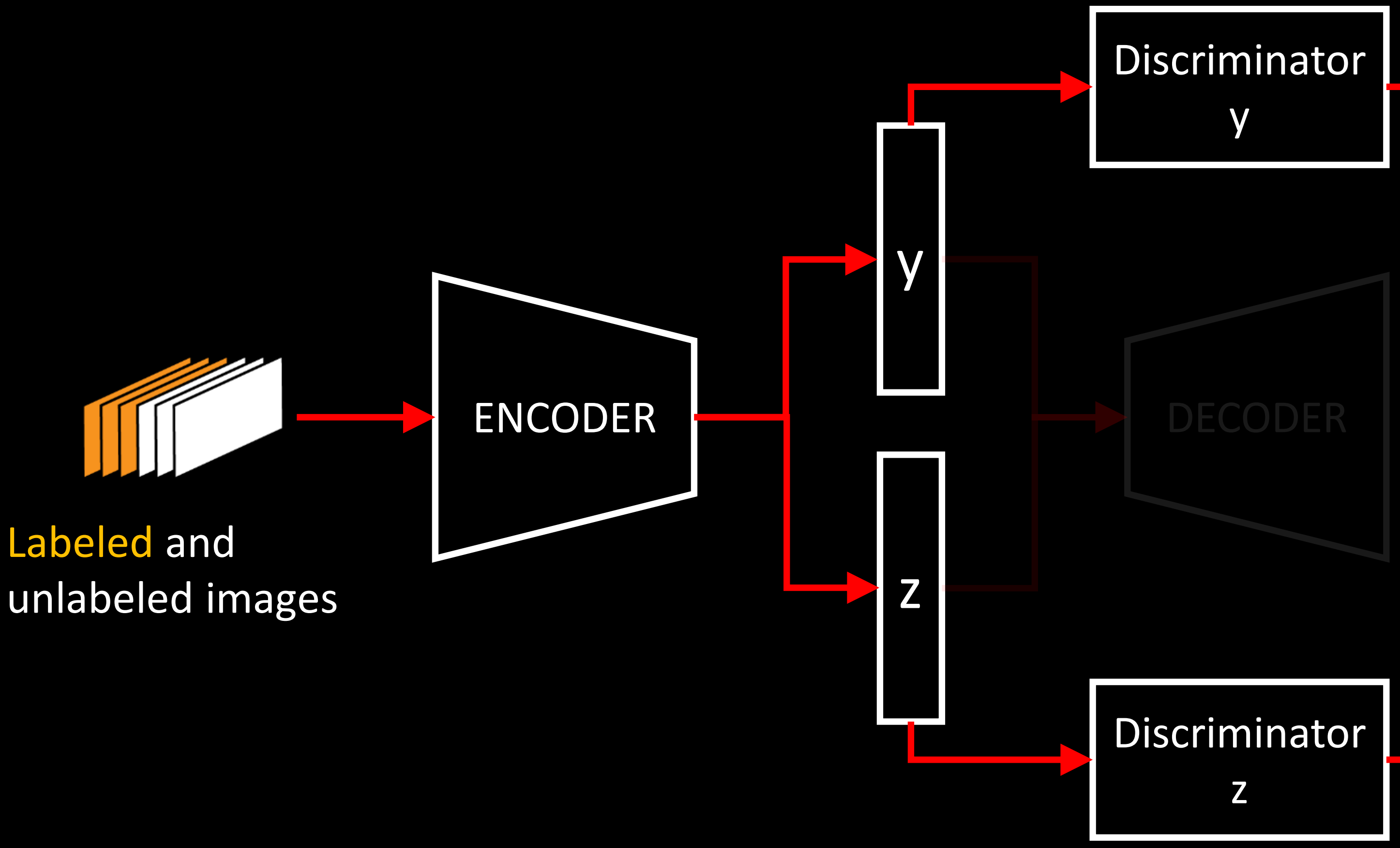
Our SSIL Framework



Step 1: reconstruction



Step 2: adversarial training

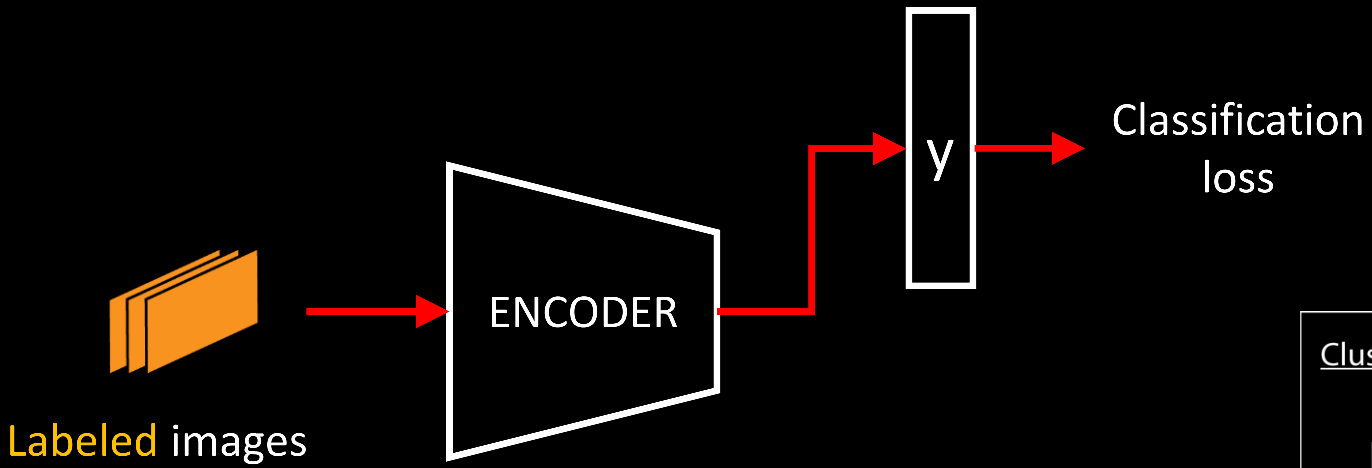


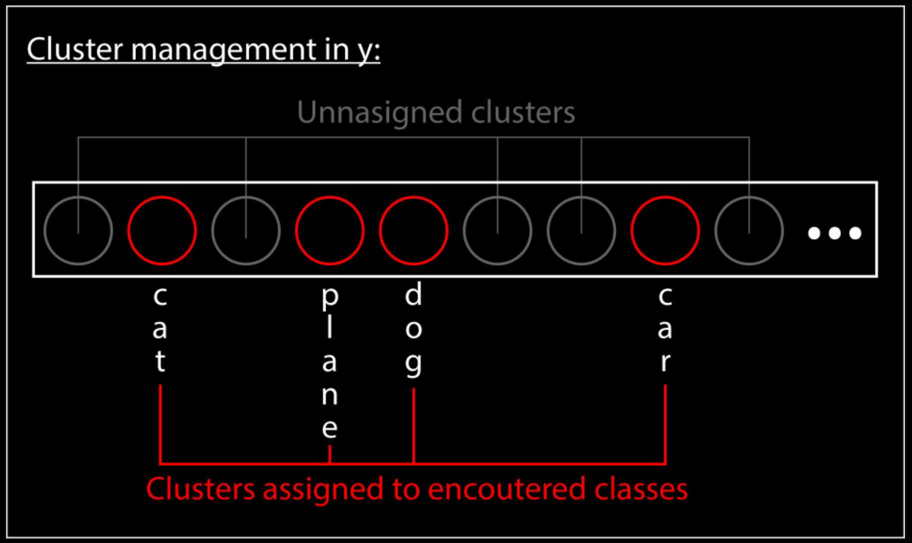
Adversarial loss
p(y)=Cat(y)

$y \in \mathbb{R}^C$ with C the number of clusters for self-supervised clustering

Adversarial loss
p(z)=$\mathcal{N}$(O, I)

Step 3: supervised classification





Results: class incremental

60,000 labeled samples

2,000 labeled samples

STL-10: 500 labeled samples

TABLE I
COMPARISON OF LATEST AND AVERAGE ACCURACY OF DIFFERENT CLASS-INCREMENTAL LEARNING METHODS ON MNIST AND STL-10

Method	MNIST		STL-10	
	Latest (%)	Average (%)	Latest (%)	Average (%)
Oracle	99.4	99.7	67.2	73.5
Fine-Tuning	19.8	44.9	16.2	38.3
LwF	71.3	85.2	17.9	42.5
DMC	81.1	87.4		
Naive Rehearsal	93.7	97.6	43.8	62.0
iCaRL	95.3	97.9	42.6	63.0
WA	96.0	98.3	47.3	63.5
Ours[a]	96.9	98.5	**57.3**	**72.0**
Ours[b] (EMNIST-digits)	**98.1**	**99.0**		
Ours[b] (EMNIST-letters)	95.9	98.5		

[a] Our standard baseline on MNIST uses EMNIST-full as unlabeled data stream.
[b] Additional results on MNIST benchmark when using EMNIST-digits and EMNIST-letters as unlabeled data stream instead of the whole EMNIST.

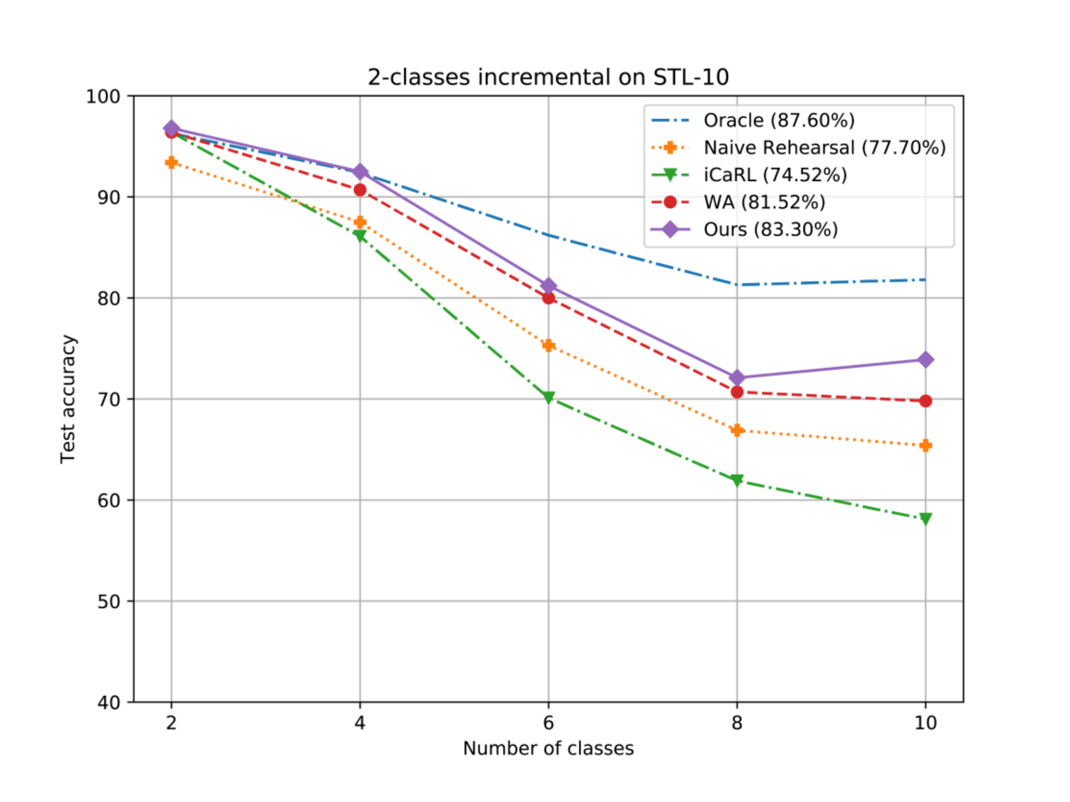
Memory size: K=400 for MNIST and K=500 for STL-10

Unlabeled dataset leveraged by our SSIL

MNIST: EMNIST (814,255 characters, digits and letters)

STL-10: 100,000 unlabeled images are provided in the dataset

Results: enhanced representations



Comparison of different rehearsal strategies initialized with a self-supervised encoder (pre-trained with RotNet)

Conclusion

- SSIL achieves better performance
- SSIL requires less labeled data
- Self-supervision is an efficient regularization for incremental learning

Thank you for watching

Poster Session T1.16

ID #2885

Semi-Supervised Class Incremental Learning

Alexis Lechat, Stéphane Herbin and Frédéric Jurie

ONERA

THE FRENCH AEROSPACE LAB

Normandie Université