Does Prior Data Matter? Exploring Joint Training in the Context of Few-Shot Class-Incremental Learning

Shiwon Kim[1]*, Dongjun Hwang[2]*†, Sungwon Woo[2]*, Rita Singh[3]†

[1]Yonsei University [2]Sogang University [3]Carnegie Mellon University (* Equal Contribution, † Corresponding Author)

Paper Github

6th CLVISION ICCV Workshop

ICCV OCT 19-23, 2025 HONOLULU HAWAII

Background

Class-Incremental Learning (CIL)

Adapt to new classes over time + Maintain strong performance on all previously observed classes

Few-Shot Class-Incremental Learning (FSCIL)

New classes emerge with only a few samples

base session (t = 0) → t = 1 → … → t = T

large label space + sufficient training samples → N-way K-shot → … → N-way K-shot

Fundamental assumption in FSCIL: ***"Previously seen data are no longer accessible in the following incremental sessions"***

However, in many real-world scenarios such as industrial deployments, **previously collected datasets often remain available.**

Motivation

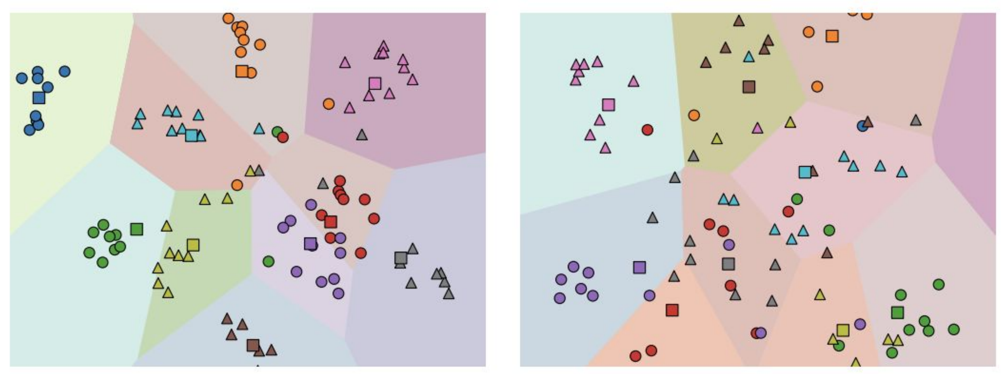
"If previous data is accessible, is it better to retrain a model using all data (i.e., joint training), or to update the model solely based on the newly introduced data (i.e., incremental learning)?"

CIL

"When previous data is available, joint training maximizes performance, while CIL is suitable under time or resource constraints."

FSCIL

Joint training is less effective in FSCIL due to class imbalance. *"It remains unclear whether retraining on the full dataset or incremental learning is preferable in FSCIL scenarios."*



(a) Joint training in CIL setting (b) Joint training in FSCIL setting

Feature space visualization.
base classes: **dots**, incremental classes: **triangles**

Contributions

I. **Develop a more reliable joint training benchmark** for comparison with FSCIL
- Explore 8 imbalance learning techniques and identify/evaluate the optimal combination
- Present this combination as a new imbalance-aware joint training benchmark for FSCIL

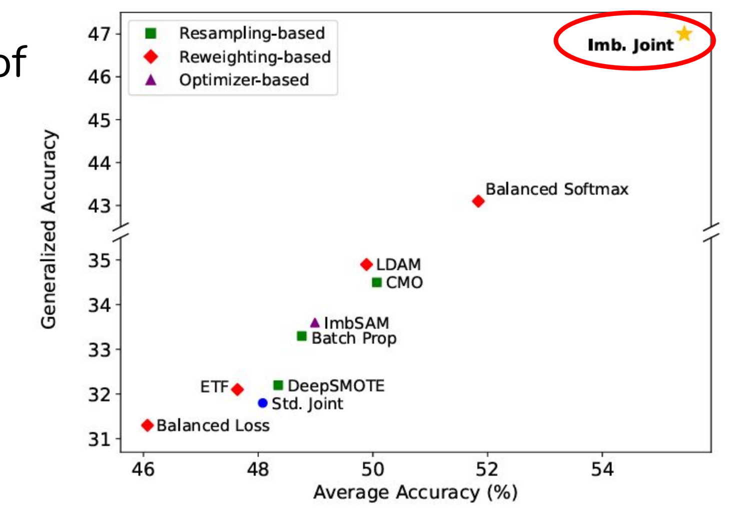
II. **Provide practical insights and guidelines** for selecting suitable training strategies in FSCIL
- Compare the new benchmark with FSCIL methods under varying resource constraints
- Reimplement and integrate all methods into a unified framework to ensure consistent comparision

Rethinking Joint Training in FSCIL

Imbalance-Aware Joint Training in FSCIL

1) Exploring *imbalanced learning* strategies

- We explore three independent categories of imbalanced learning:
  - 3 Resampling-based methods
  - 4 Reweighting-based methods
  - 1 Optimizer-based method
- Combining CMO, Balanced Softmax, and ImbSAM achieves the best performance.



2) Ablation Study

- The new imbalance-aware joint training benchmark <u>**improves**</u> *aAcc* by 7%p and *gAcc* by 15%p over standard joint training.

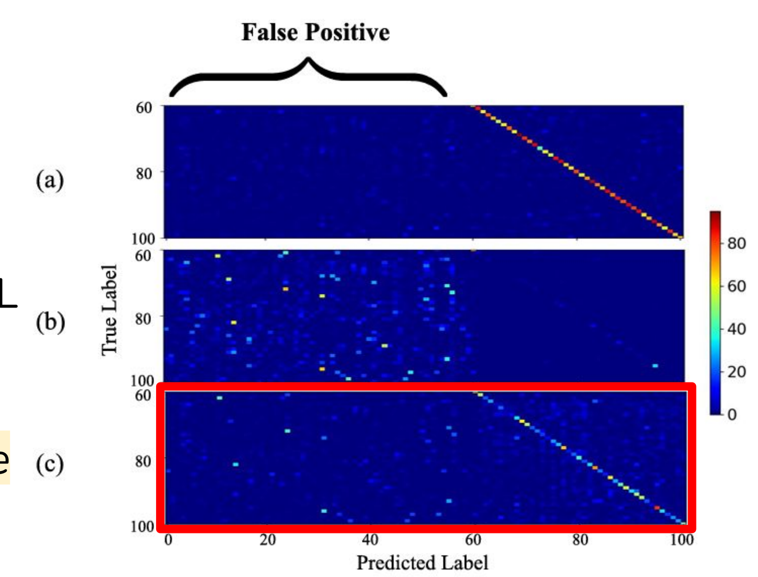
CMO	*BalancedSoftmax*	*ImbSAM*	*aAcc*	*gAcc*
			48.1	31.8
✓			50.1	34.5
✓	✓		55.5	45.9
✓	✓	✓	**55.8**	**46.8**

Analysis of Imbalance-Aware Joint Training

1) Resolving bias towards base classes

- Comparison of ***confusion matrices***:
  - (a) standard joint training in CIL
  - (b) standard joint training in FSCIL
  - (c) imbalance-aware joint training in FSCIL

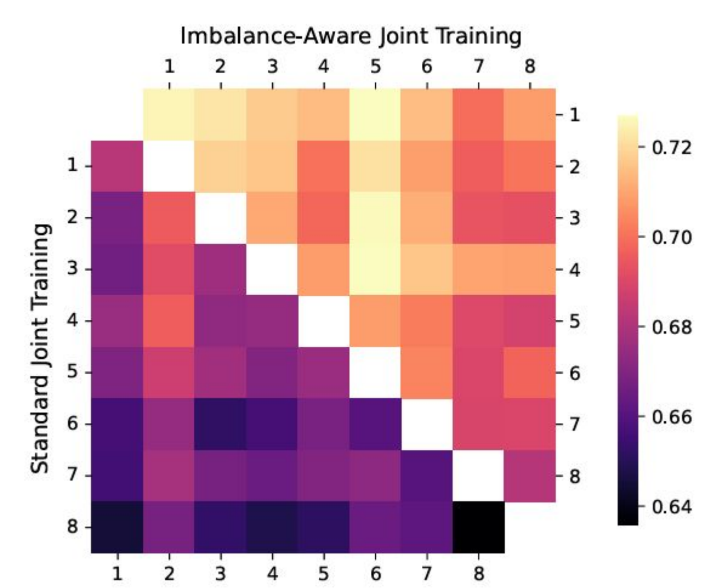
*(c) shows **fewer FPs for incremental classes** than (b), suggesting a more reliable benchmark for FSCIL.*



2) Resemblance to joint training in CIL

- ***Centered Kernel Alignment (CKA)*** feature similarity vs. standard joint training in CIL:
  (a) Upper: imbalance-aware joint training
  (b) Lower: standard joint training

*(a) exhibits brighter coloration than (b), indicating **stronger feature similarity** to standard joint training in CIL.*



Towards a Practical Guideline for FSCIL

Comparison of FSCIL and Joint Training

← **Joint training outperforms FSCIL**

Method	S8	*aAcc* S8 Base	*aAcc* S8 Inc.	*aAcc*	*gAcc*
Std. Joint	48.1	**78.8**	1.9	61.1	46.9
Imb. Joint	**55.3**	70.5	<u>32.5</u>	**65.9**	**58.0**
CEC [55]	39.7	50.5	23.5	50.5	45.9
FACT [60]	42.4	62.5	12.1	53.1	43.6
TEEN [48]	41.7	63.7	8.8	52.6	42.4
S3C [23]	41.3	47.4	32.1	48.4	45.5
WaRP [25]	47.1	64.2	21.6	57.0	49.2
SAVC [39]	<u>54.7</u>	<u>76.5</u>	22.2	<u>65.8</u>	<u>55.6</u>
LIMIT [61]	49.7	68.9	21.1	60.4	51.2
YourSelf [40]	48.5	56.0	**37.3**	58.7	54.6

Results on CIFAR-100

→ **FSCIL outperforms Joint training**

Method	S8	*aAcc* S8 Base	*aAcc* S8 Inc.	*aAcc*	*gAcc*
Std. Joint	44.2	<u>72.3</u>	2.1	56.0	43.0
Imb. Joint	<u>51.7</u>	66.7	<u>29.1</u>	<u>60.6</u>	<u>53.5</u>
CEC [55]	46.7	65.2	18.9	56.5	47.9
FACT [60]	44.1	66.8	9.9	54.9	44.4
TEEN [48]	43.8	58.0	22.4	52.7	45.8
S3C [23]	40.7	51.4	24.9	47.7	42.3
WaRP [25]	49.7	65.9	25.4	58.6	50.8
SAVC [39]	**54.1**	**76.2**	21.1	**65.3**	**55.6**
LIMIT [61]	47.1	64.2	21.6	57.3	49.2
YourSelf [40]	49.4	60.9	**32.2**	58.2	52.6

Results on *mini*ImageNet

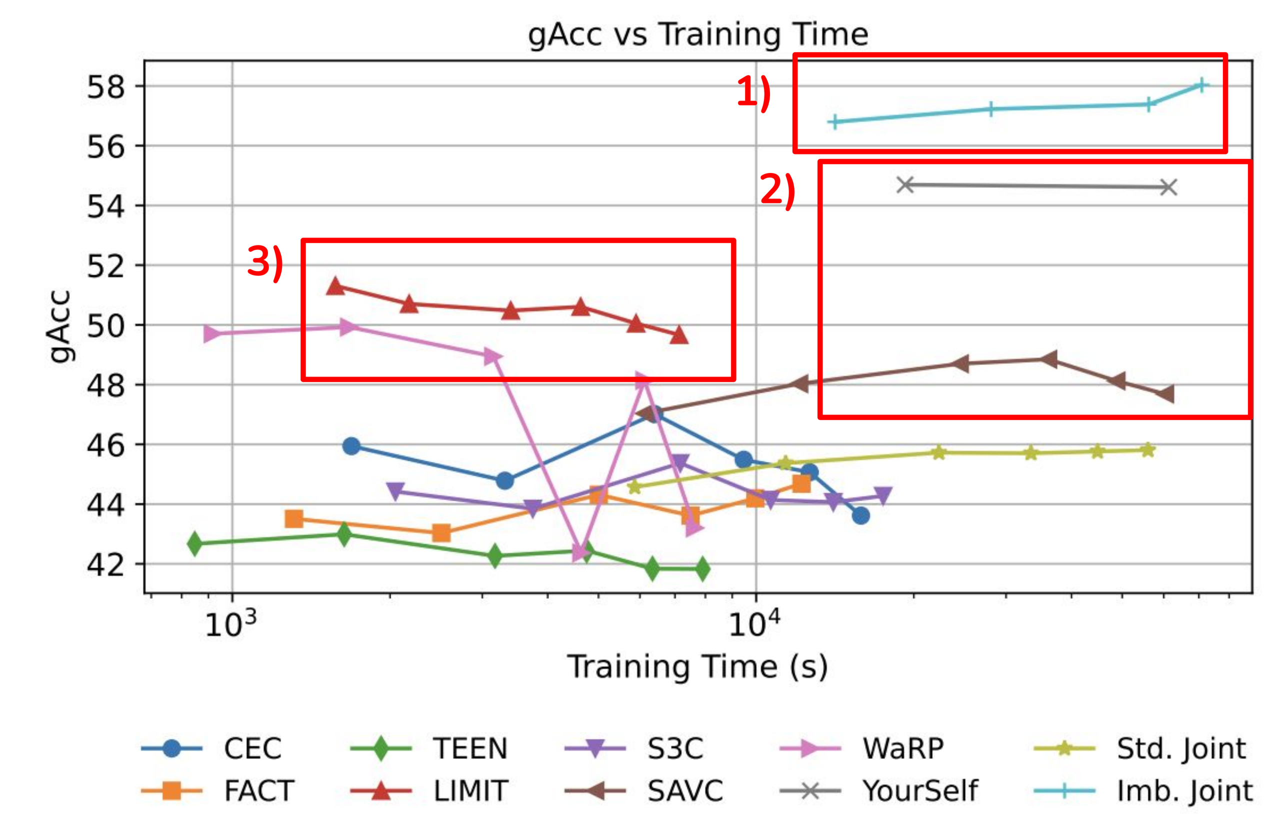
Method	S10	*aAcc* S10 Base	*aAcc* S10 Inc.	*aAcc*	*gAcc*
Std. Joint	58.3	**77.4**	40.1	62.6	57.0
Imb. Joint	<u>61.9</u>	73.2	<u>51.4</u>	<u>65.1</u>	<u>62.8</u>
CEC [55]	44.6	66.2	24.1	52.2	48.1
FACT [60]	55.2	73.0	38.1	61.0	57.1
TEEN [48]	54.9	70.5	40.0	60.7	57.8
S3C [23]	49.9	54.5	44.1	52.7	52.1
WaRP [25]	56.2	72.3	40.9	61.4	57.9
SAVC [39]	59.9	<u>75.1</u>	45.4	64.3	60.4
LIMIT [61]	40.3	58.1	23.3	48.2	44.2
YourSelf [40]	**62.5**	73.4	**52.3**	**66.4**	**64.3**

Results on CUB-200

"Contrary to expectations, imbalance-aware joint training does not always outperform FSCIL methods."

Resource-Aware Comparison

Under the standard training protocol, **performance trends over training time** suggest the following insights:



*1) **Imb.Joint** is effective with with sufficient resources and access to prior data*
*2) **SAVC/Yourself** performs well with sufficient resources but unavailable prior data*
*3) **LIMIT** provides the best trade-off between efficiency and performance when both resources and prior data are limited*