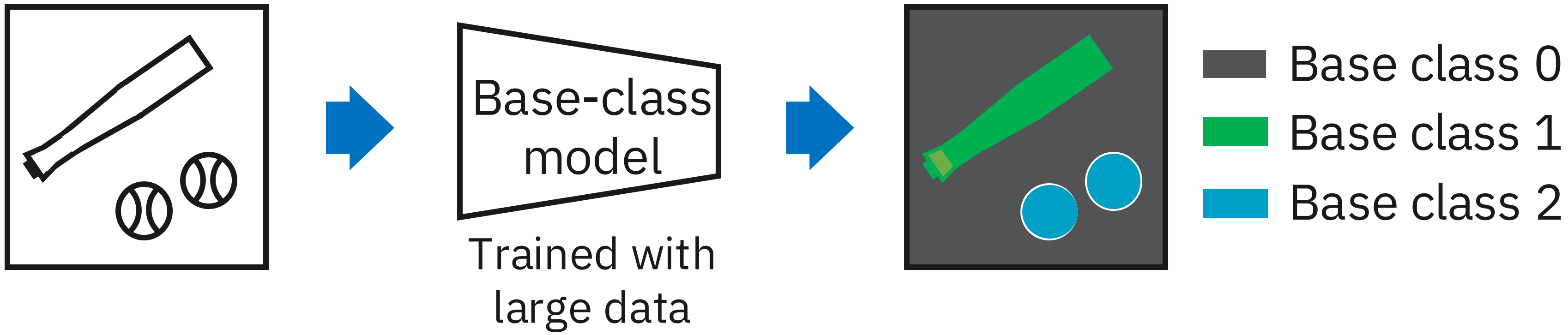
A Surprisingly Simple Approach to Generalized Few-Shot Semantic Segmentation

Tomoya Sakai　Haoxiang Qiu　Takayuki Katsuki
Daiki Kimura　Takayuki Osogami　Tadanobu Inoue

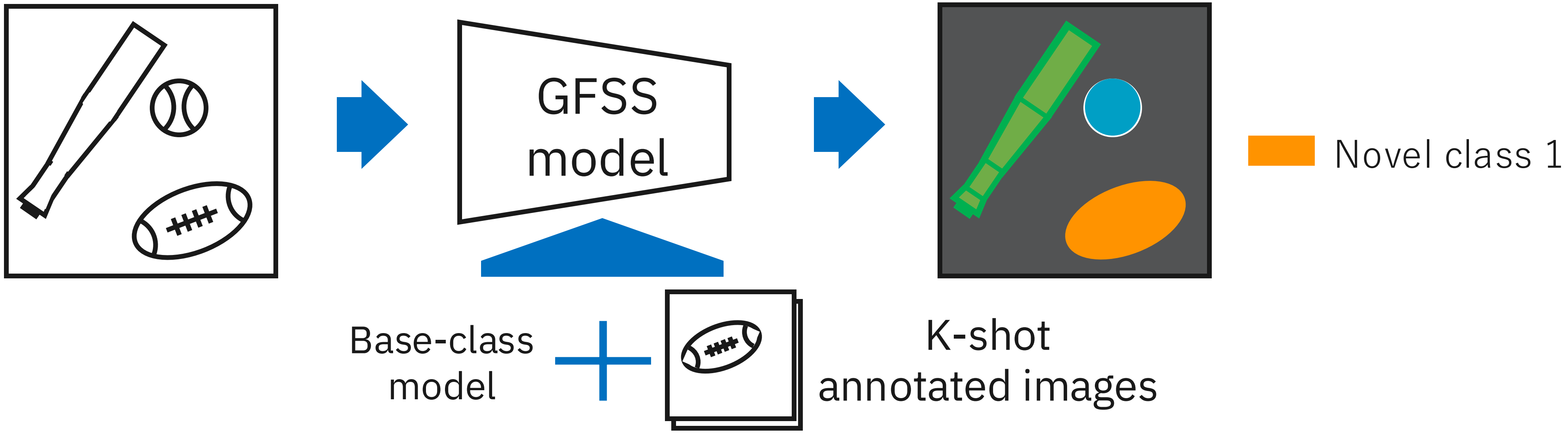
IBM Research

Setting | GFSS (Generalized Few-Shot Semantic Segmentation)

Assume **base-class model** is available



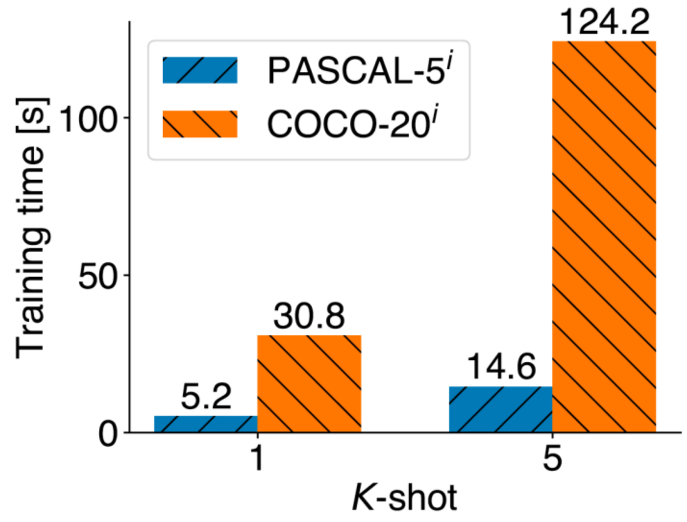
Goal: recognize novel objects with a few annotated images



Existing vs Our Approach

Existing approach

- Rely on several techniques:
  - Meta-learning
  - Information maximization principle
  - Knowledge distillation
  - Transductive learning



Our approach

- Simple rule + standard supervised learning
- Advantages
  - Efficient training
  - Theory: Perfectly maintain segmentation performance for most of the base classes

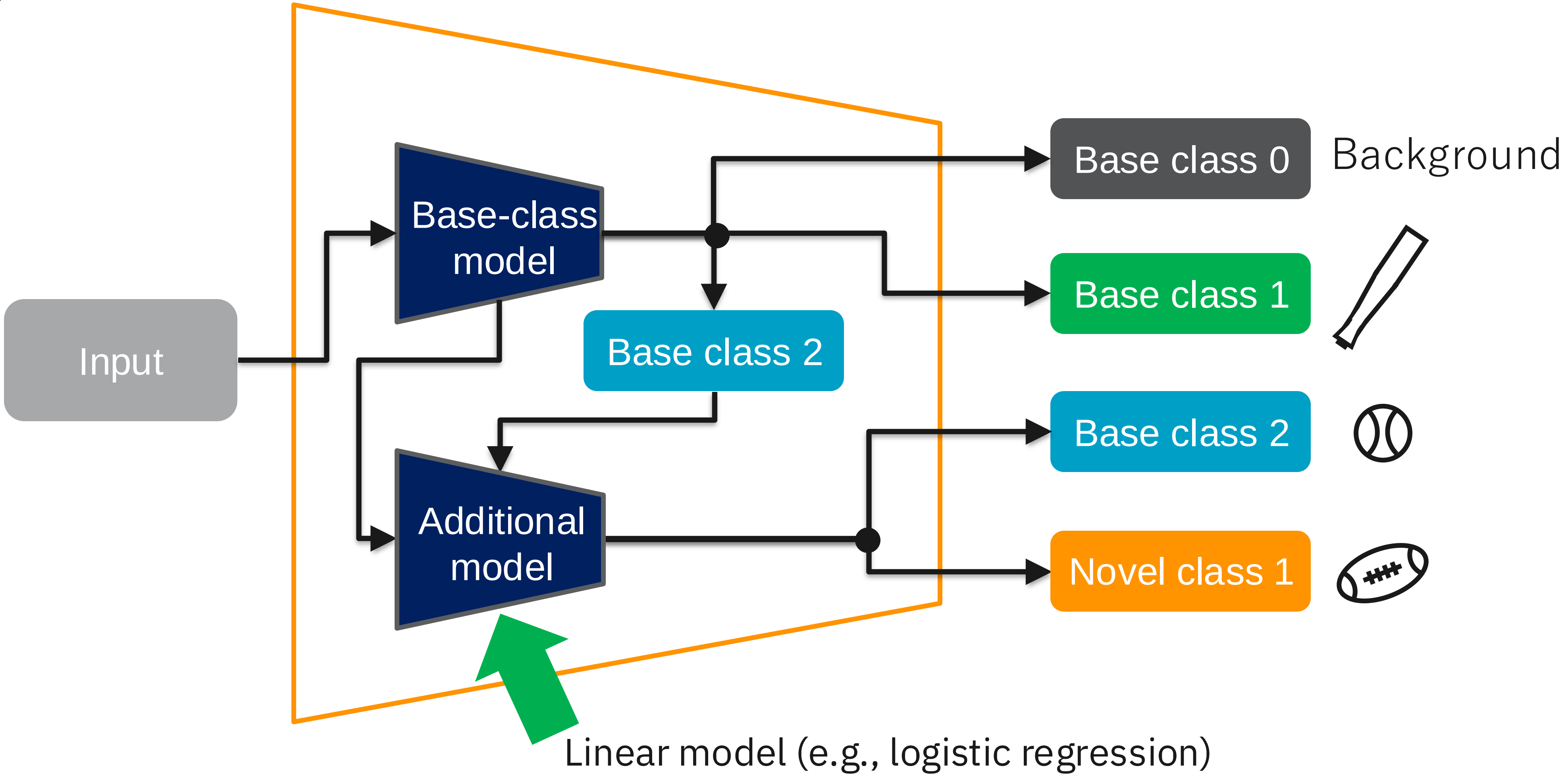
Proposition 4.1. *Let $\widehat{\boldsymbol{Y}}_{\mathrm{b}}$ and $\widehat{\boldsymbol{Y}}_{\mathrm{BCM}}$ be the predictions of the base-class model and BCM, respectively. The mIoUs of $\widehat{\boldsymbol{Y}}_{\mathrm{b}}$ and $\widehat{\boldsymbol{Y}}_{\mathrm{BCM}}$ over $\mathcal{Y}_{\mathrm{b}} \setminus \mathcal{B}$ are the same:*

$$\mathrm{mIoU}_{\mathcal{Y}_{\mathrm{b}} \setminus \mathcal{B}}(\boldsymbol{Y}, \widehat{\boldsymbol{Y}}_{\mathrm{b}}) = \mathrm{mIoU}_{\mathcal{Y}_{\mathrm{b}} \setminus \mathcal{B}}(\boldsymbol{Y}, \widehat{\boldsymbol{Y}}_{\mathrm{BCM}}). \quad (4)$$

If $|\mathcal{B}|$ is small, the segmentation performance of most of the base classes is perfectly maintained.

Overview of Proposed Method

Run additional prediction if predicted base-class is “similar” to novel class



Experiments

Setting

PASCAL-5^i dataset

15 base classes, 5 novel clases

Results (mIoU: higher is better)

Method		1-shot			5-shot		
		Base	Novel	Mean	Base	Novel	Mean
CAPL	(CVPR'22)	64.80	17.46	41.13	65.43	24.43	44.93
BAM	(CVPR'22)	71.60	27.49	49.55	71.60	28.96	50.28
DIaM	(CVPR'23)	70.89	35.11	53.00	70.85	55.31	63.08
BCM	(Ours)	71.15	41.24	56.20	71.23	55.36	63.29
			+6.1%				

Proposed method achieved superior performance without resorting to meta-learning, information maximization principle, and transductive learning