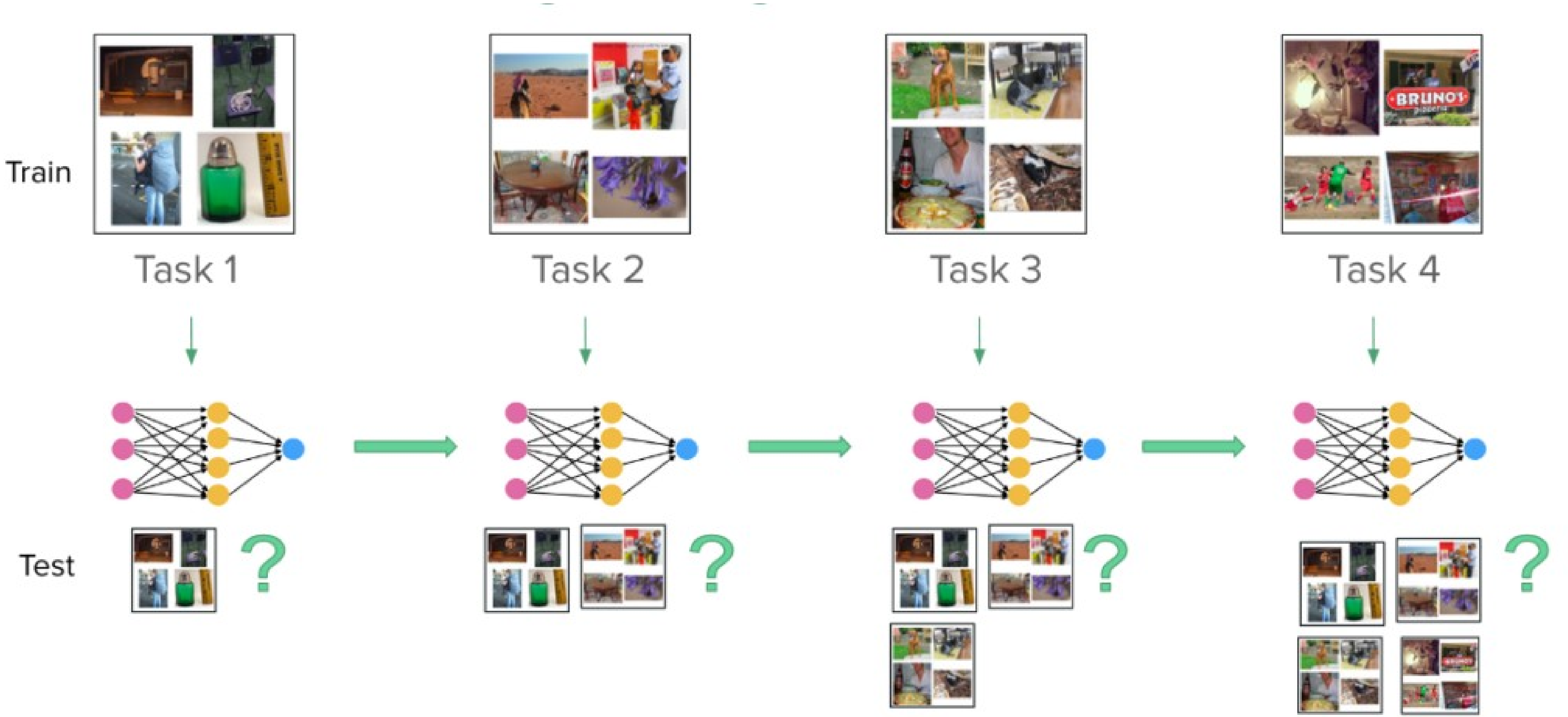
IDEAS NCBR

Divide and not forget: Ensemble of selectively trained experts in Continual Learning

Grzegorz Rypeść[1,2], Sebastian Cygert[1,3], Valeriya Khan[1,2], Tomasz Trzciński[1,2,4], Bartosz Zieliński[1,5], Bartłomiej Twardowski[1,6,7]

[1]IDEAS-NCBR, [2]Warsaw University of Technology, [3]Gdańsk University of Technology, [4]Tooploox, [5]Jagiellonian University, [6]Computer Vision Center, Barcelona, [7]Department of Computer Science, Universitat Autònoma de Barcelona

Exemplar-Free Class and Task Incremental Learning

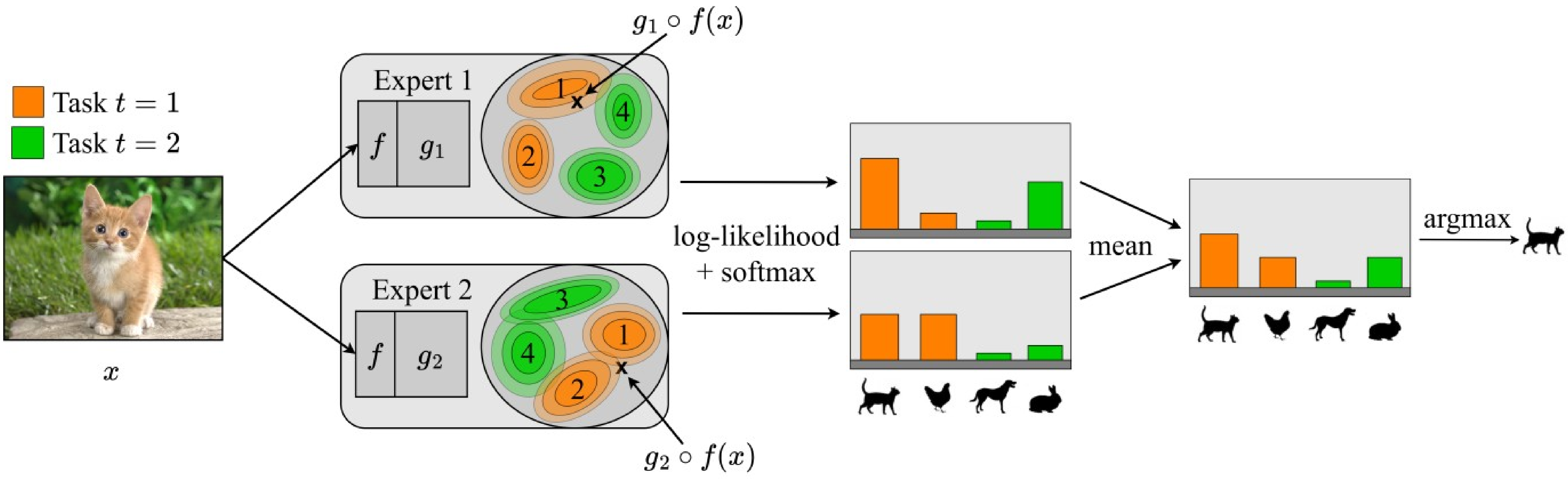


van de Ven, Gido M., and Andreas S. Tolias. "Three scenarios for continual learning." arXiv e-prints (2019): arXiv-1904.

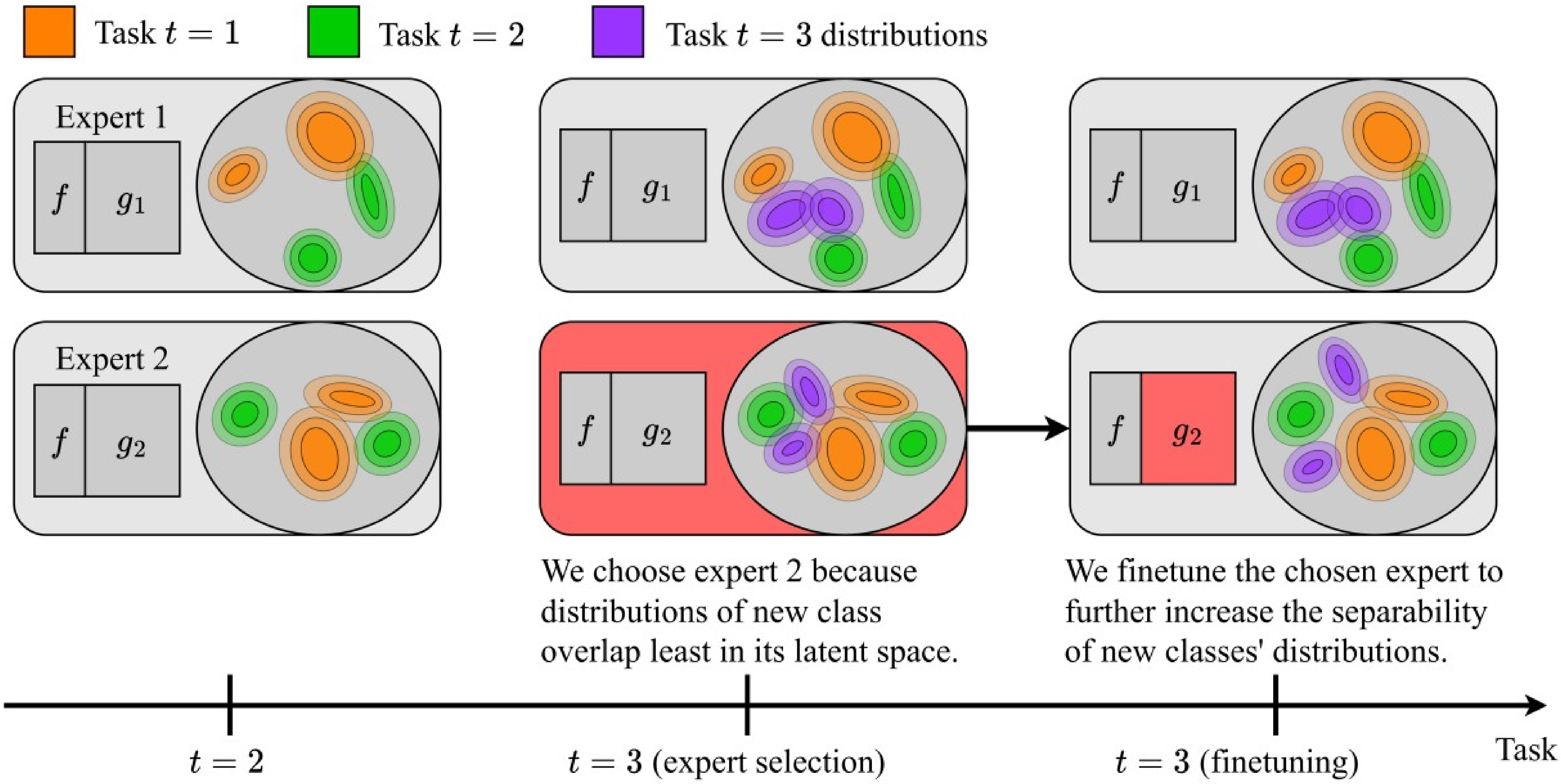
SEED - main aspects

- Unlike existing ensembling methods we train only a single expert on a task, but perform inference using all experts
- Each expert is a deep network and a set of Multivariate Gaussian distributions representing classes
- We utilize ensemble of Bayes classifiers for inference
- Expert diversification is obtained by training them on different data

SEED - Inference



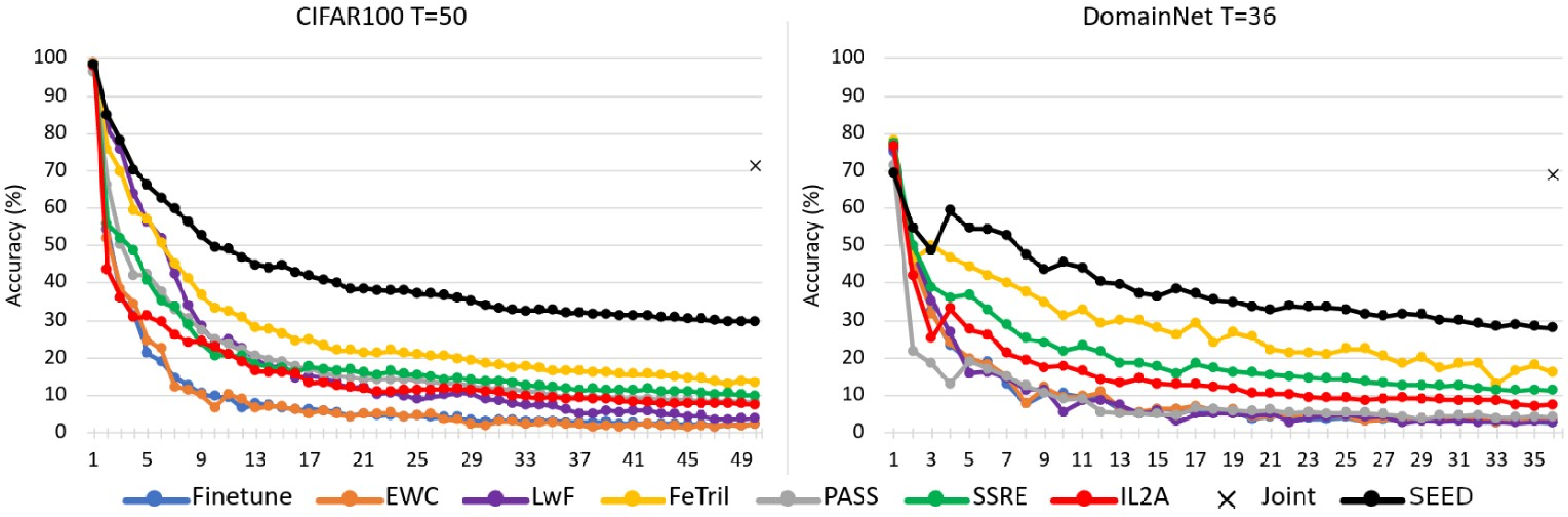
SEED - Training



SEED - Results

CIL Method	CIFAR-100 (ResNet32)			DomainNet			ImageNet-Subset
	T=10	T=20	T=50	T=12	T=24	T=36	T=10
Finetune	26.4±0.1	17.1±0.1	9.4±0.1	17.9±0.3	14.8±0.1	10.9±0.2	27.4±0.4
EWC (Kirkpatrick et al., 2017) (PNAS'17)	37.8±0.8	21.0±0.1	9.2±0.5	19.2±0.2	15.7±0.1	11.1±0.3	29.8±0.3
LwF* (Rebuffi et al., 2017) (CVPR'17)	47.0±0.2	38.5±0.2	18.9±1.2	20.9±0.2	15.1±0.6	10.3±0.7	32.3±0.4
PASS (Zhu et al., 2021b) (CVPR'21)	37.8±1.1	24.5±1.0	19.3±1.7	25.9±0.5	23.1±0.5	9.8±0.3	-
IL2A (Zhu et al., 2021a) (NeurIPS'21)	43.5±0.3	28.3±1.7	16.4±0.9	20.7±0.5	18.2±0.4	16.2±0.4	-
SSRE (Zhu et al., 2022) (CVPR'22)	44.2±0.6	32.1±0.9	21.5±1.8	33.2±0.7	24.0±1.0	22.1±0.7	45.0±0.5
FeTrIL (Petit et al., 2023) (WACV'23)	46.3±0.3	38.7±0.3	27.0±1.2	33.5±0.6	33.9±0.5	27.5±0.7	58.7±0.2
SEED	**61.7±0.4**	**56.2±0.3**	**42.6±1.4**	**45.0±0.2**	**44.9±0.2**	**39.2±0.3**	**67.8±0.3**
Joint		71.4±0.3		63.7±0.5	69.3±0.4	69.1±0.1	81.5±0.5

SEED - Results

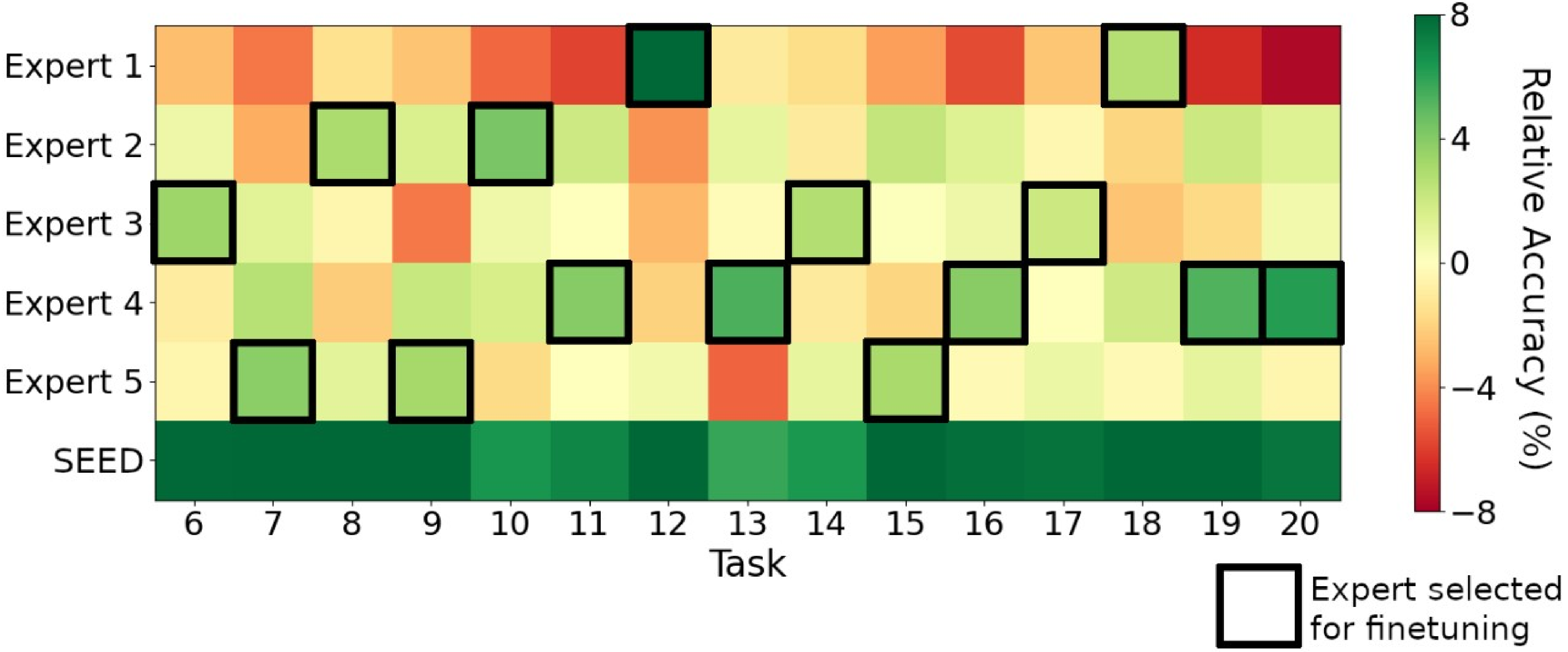


SEED - Task incremental setting + Abelation

Approach	#Params.	20-split	50-split
HAT	6.8M	77.0	80.5
MARK	4.7M	78.3	-
BNS	6.7M	-	82.4
CoSCL(EWC+LWF)	4.6M	79.4±1.0	87.9±1.1
SEED	3.2M	**86.8±0.3**	**91.2±0.4**
SEED(1 shared)	3.2M	86.7±0.6	91.2±0.5
SEED(11 shared)	3.1M	85.6±0.3	89.6±0.2
SEED(21 shared)	**2.7M**	82.4±0.4	88.1±0.5

Approach	Acc.(%)
SEED(5 experts)	**61.7 ±0.4**
standard ensemble	56.9±0.4
weighted ensemble	57.0±0.5
CoSCL ensemble	57.3±0.4
w/o multivariate Gauss.	53.5±0.5
w/o covariance	54.1±0.3
w/o temp. in softmax	59.2±0.5
w/ ReLU	57.8±0.6

SEED - Diversification of experts



IDEAS NCBR

Summary

- We introduce SEED method for CL, that achieves state-of-the-art results
- We develop a novel selection strategy that assigns expert to a task

Link to the paper and code:

Grzegorz Rypeść
LinkedIn
grzegorz.rypesc@ideas-ncbr.pl