NEURAL INFORMATION PROCESSING SYSTEMS

One-Time Initial Alignment Enables Resilient Learning without Weight Transport

NeurIPS 2026

Jeonghwan Cheon, Jaehyuk Bae, Se-Bum Paik

Department of Brain and Cognitive Sciences
Korea Advanced Institute of Science and Technology

KAIST

Introduction

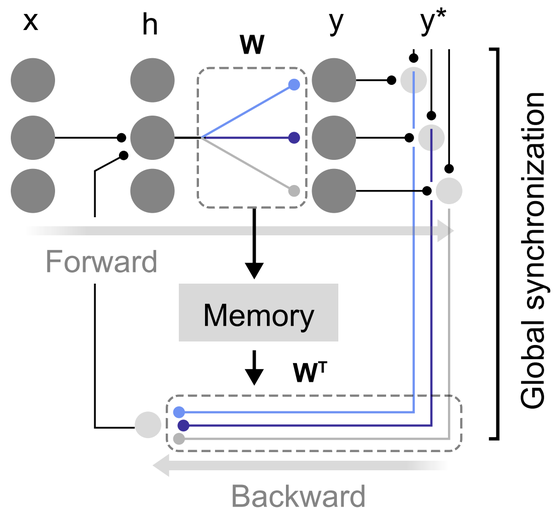
Feedback alignment

01

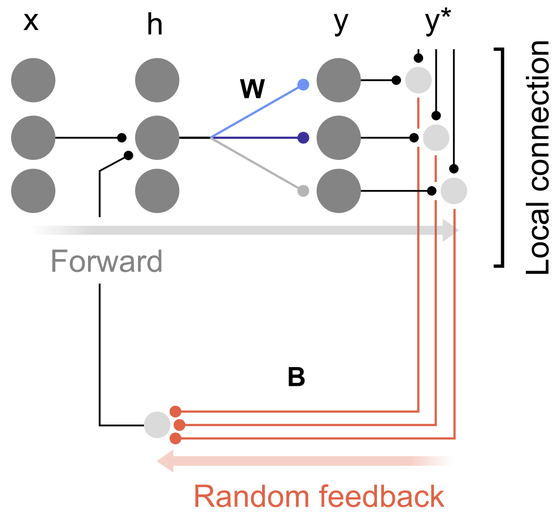
Feedback alignment replaces weight-symmetric backpropagation with fixed random feedback weights, but often suffers from poor learning performance.

Lillicrap 2016, Bartunov 2018

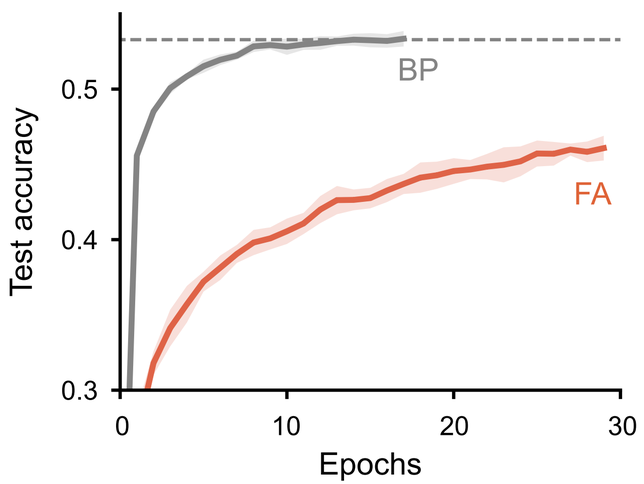
Backpropagation (BP)



Feedback alignment (FA)



Learning dynamics



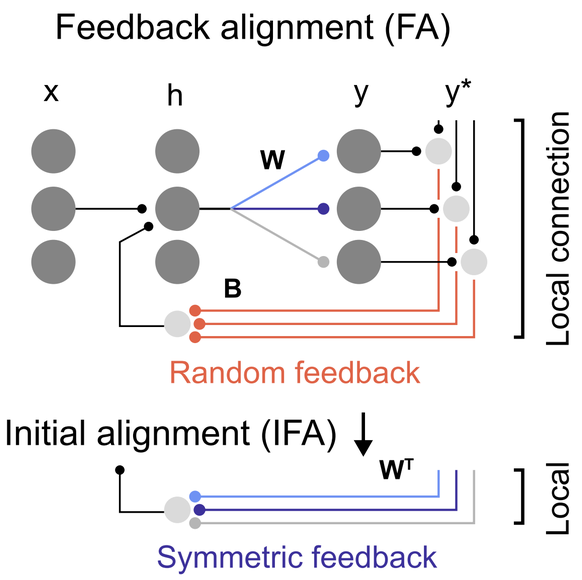
Method

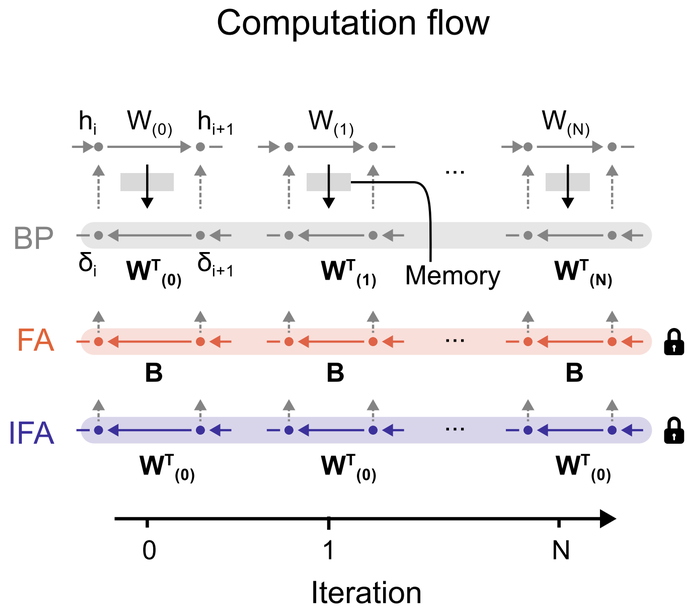
One-time initialization of fixed feedback weights

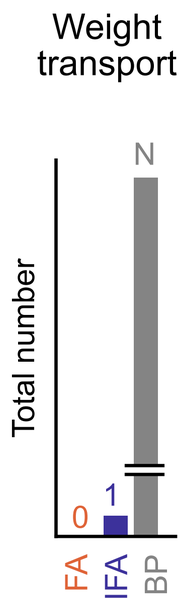
02

We align forward and feedback weights once at initialization and keep the feedback weights fixed throughout training.

Bordelon 2023, Cheon 2024





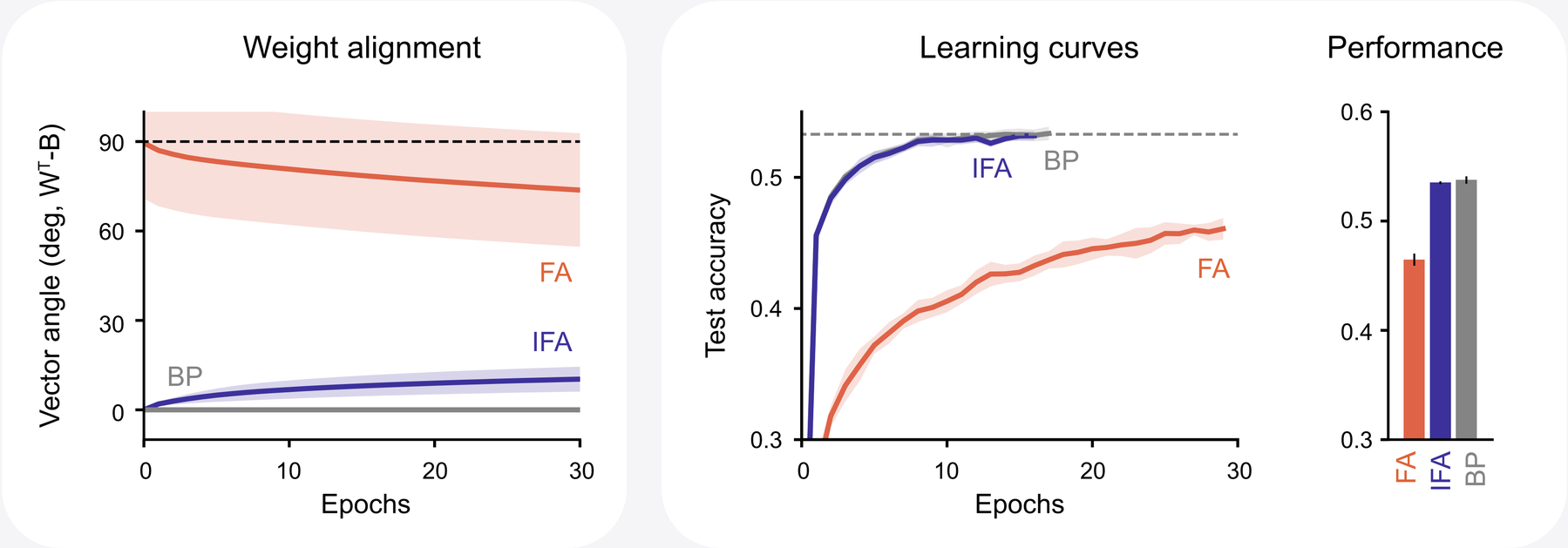


Results

Fixed feedback learning with initial alignment

03

One-time initial alignment improves convergence speed and final performance without requiring ongoing weight synchronization.

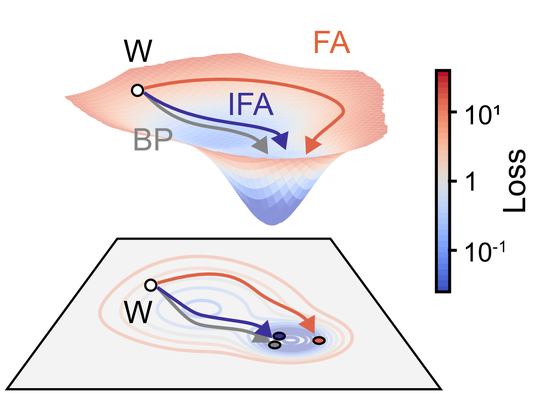


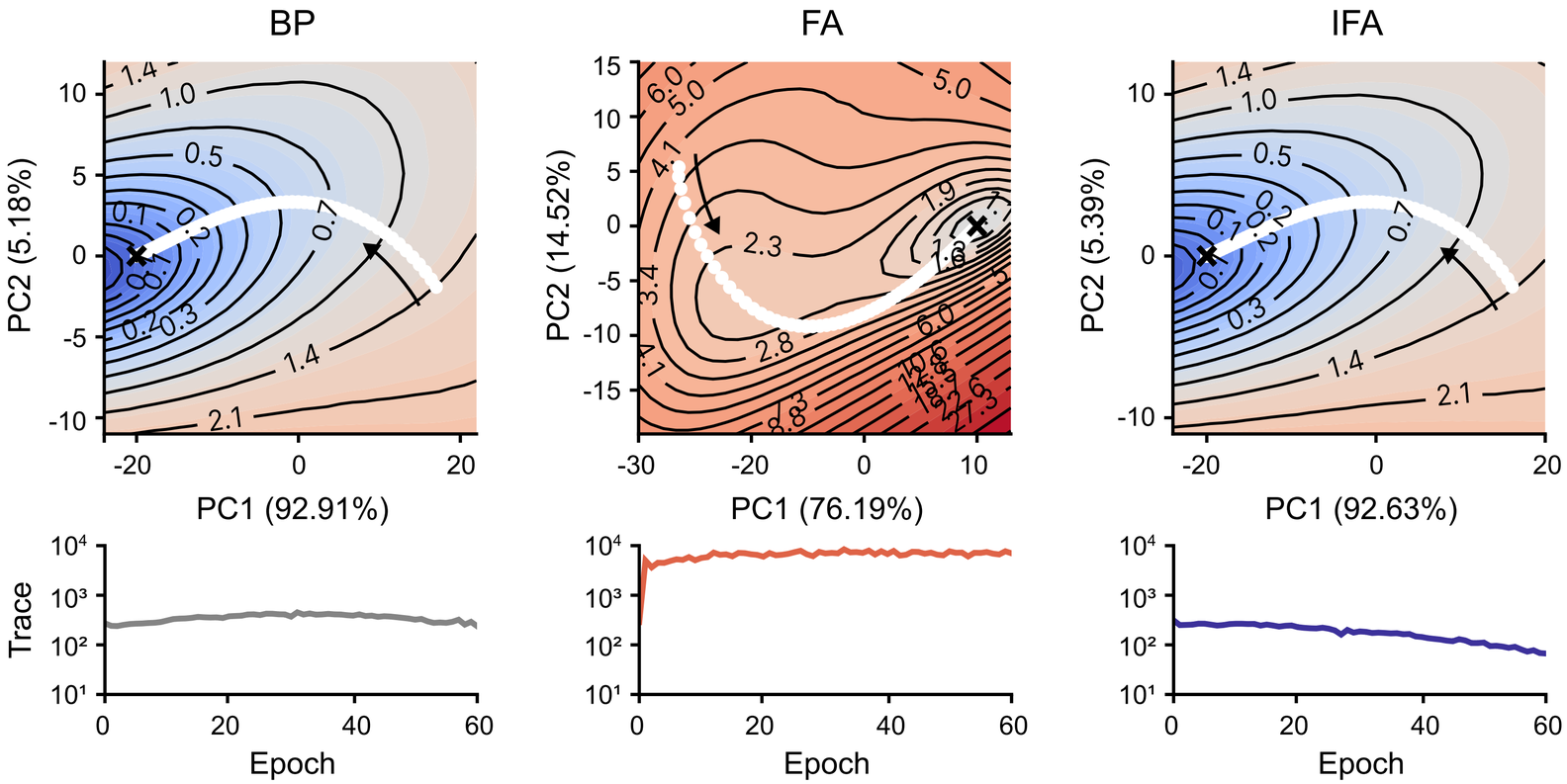
Results

Initial alignment stabilizes loss minimization

04

Initial alignment promotes stable optimization trajectories and smoother loss landscapes without repeated weight transport.



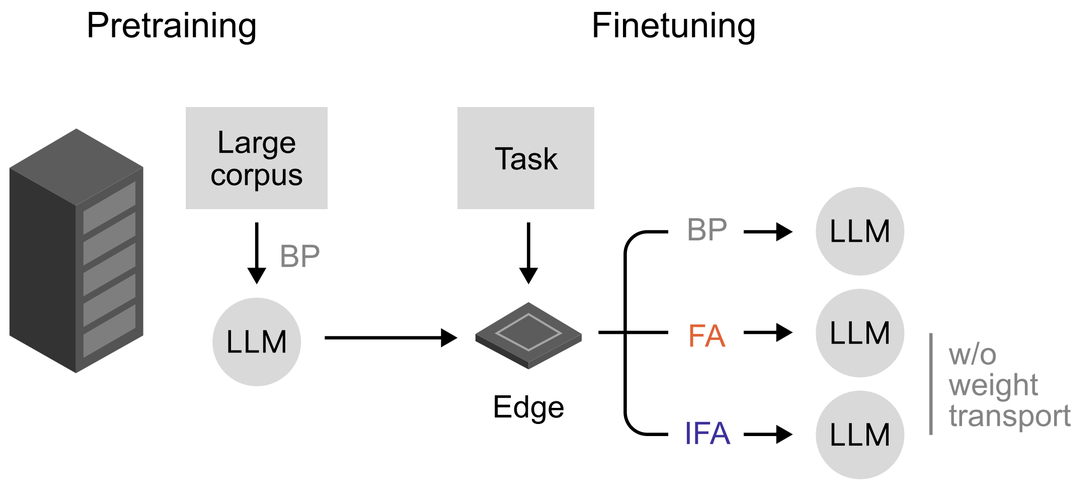


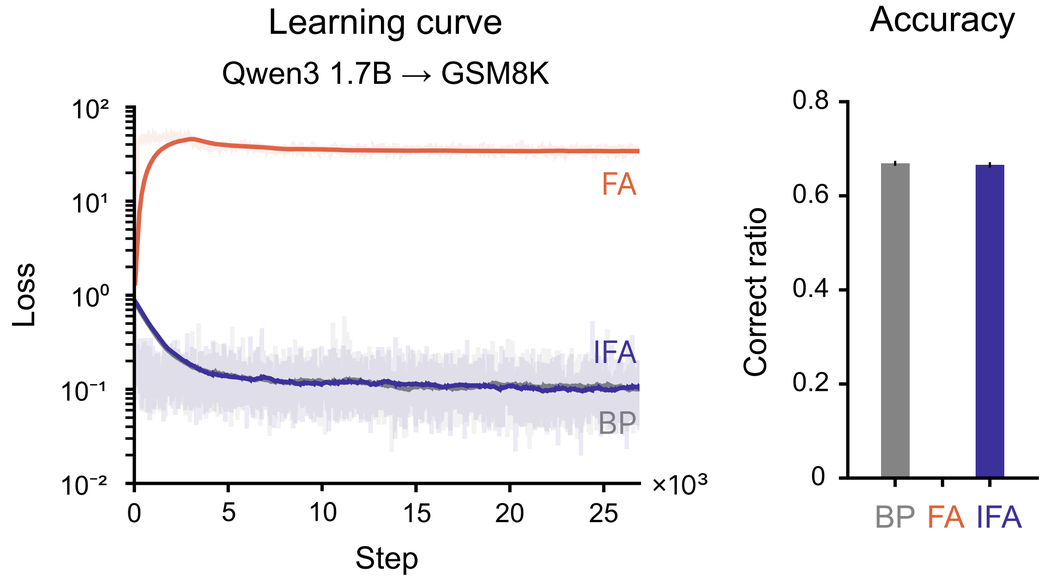
Results

Large language model finetuning with IFA

05

Standard FA diverged, whereas IFA enabled stable learning and performance comparable to backpropagation.





Results

IFA generalizes across language models and tasks

IFA achieves accuracy comparable to BP across multiple downstream tasks and model sizes, while standard FA fails to converge.

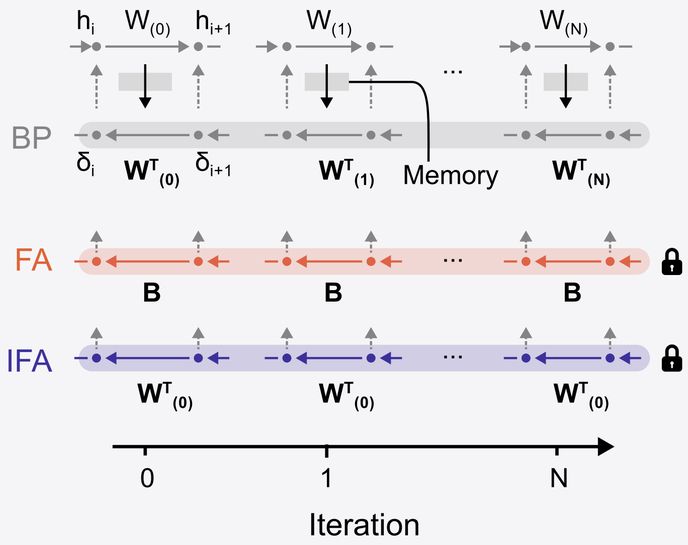
			Fine-tuned		
Model	Task	Base	BP	FA	IFA
Qwen3-1.7B	StrategyQA	19.07	40.27 ± 4.62	0.00 ± 0.00	38.72 ± 1.39
Qwen3-1.7B	BoolQ	77.58	85.47 ± 0.03	0.00 ± 0.00	85.19 ± 0.18
Qwen3-1.7B	ARC-Challenge	71.76	76.93 ± 0.30	0.37 ± 0.36	77.19 ± 0.44
Qwen3-1.7B	RTE	39.35	86.64 ± 1.30	0.00 ± 0.00	86.16 ± 0.75
Qwen3-1.7B	CommonsenseQA	60.52	73.16 ± 0.71	0.00 ± 0.00	73.22 ± 0.43
Qwen3-8B	ARC-Challenge	90.87	92.24 ± 0.31	0.03 ± 0.05	92.18 ± 0.18
Qwen3-8B	RTE	86.28	90.61 ± 0.00	0.00 ± 0.00	90.85 ± 0.21

Takeaway

07

Aligning forward and feedback weights
once at initialization improves fixed-feedback learning without repeated weight transport.

Computation flow



Weight transport



Learning curves

