How to Merge Your Multimodal Models Over Time?

Sebastian Dziadzio*, Vishaal Udandarao*, Karsten Roth*, Ameya Prabhu,
Zeynep Akata, Samuel Albanie, Matthias Bethge

27.04.2025 · ICLR MCDC · @sbdzdz

Model merging literature typically assumes a set of independent expert models that are merged **simultaneously.**

How to integrate the knowledge of trained model experts when new concepts and domains emerge **over time?**

About TIME

Initialization

all checkpoints (all) | moving average (ema) | last checkpoint (ft) | base model (zs)

Deployment

all checkpoints (all) | moving average (ema) | last checkpoint (ft)

Merge technique

averaging | task arithmetic | ties | dare | breadcrumbs | lines | magmax

Fomo-in-Flux

label: a satellite image of a bridge

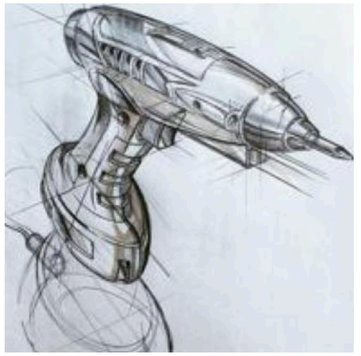
caption: the satellite image shows a bridge, an elevated highway built over water, with cars travelling over it

label: a painting by richard gerstl

caption: the painting shows a group of people on a large canvas and is attributed to richard gerstl

label: a photo of a greater praire chicken

caption: the greater praire chicken, a brown feathered creature, is captured in a photo as it sits on a rock surrounded by plants

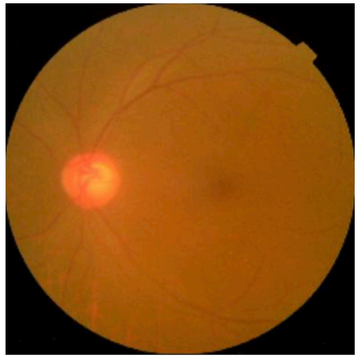


label: a pencil sketch of an electric screwdriver

caption: the pencil sketch of a screwdriver provides a detailed representation of this real-world tool

label: a photo of a bumpy texture

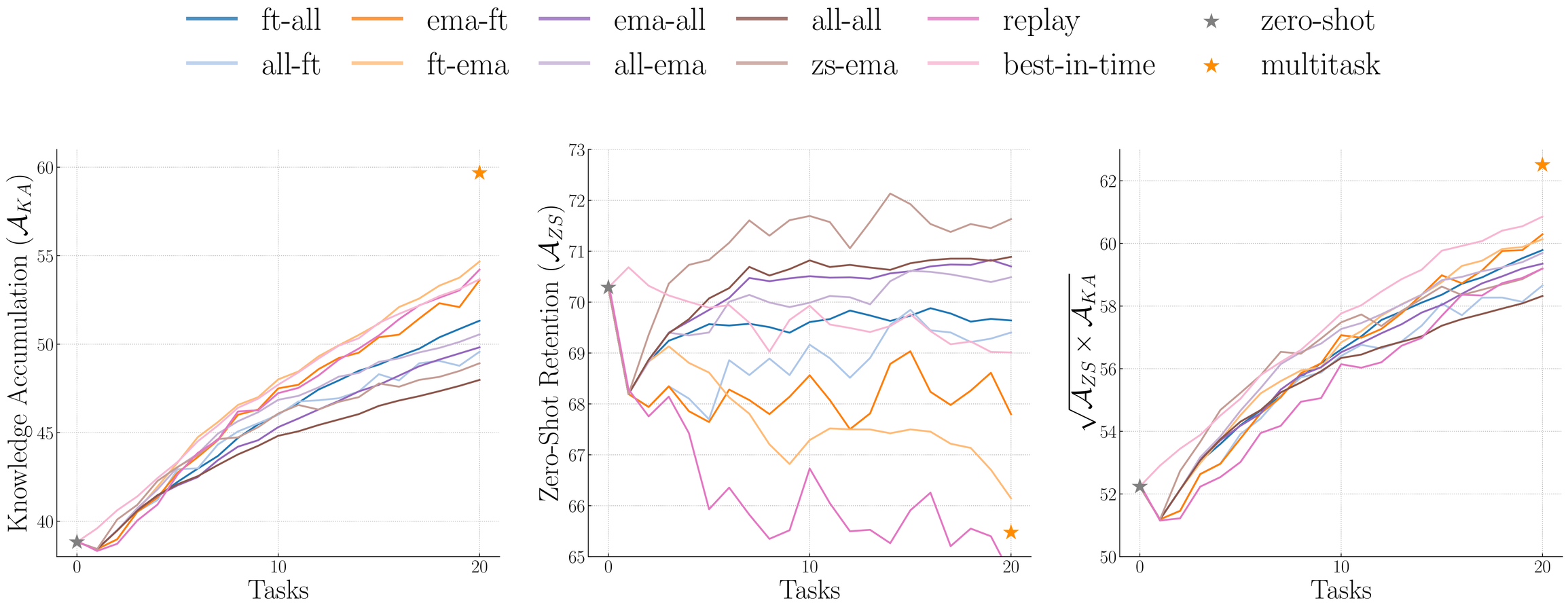
caption: there is a photo of a bumpy texture resembling a meringue



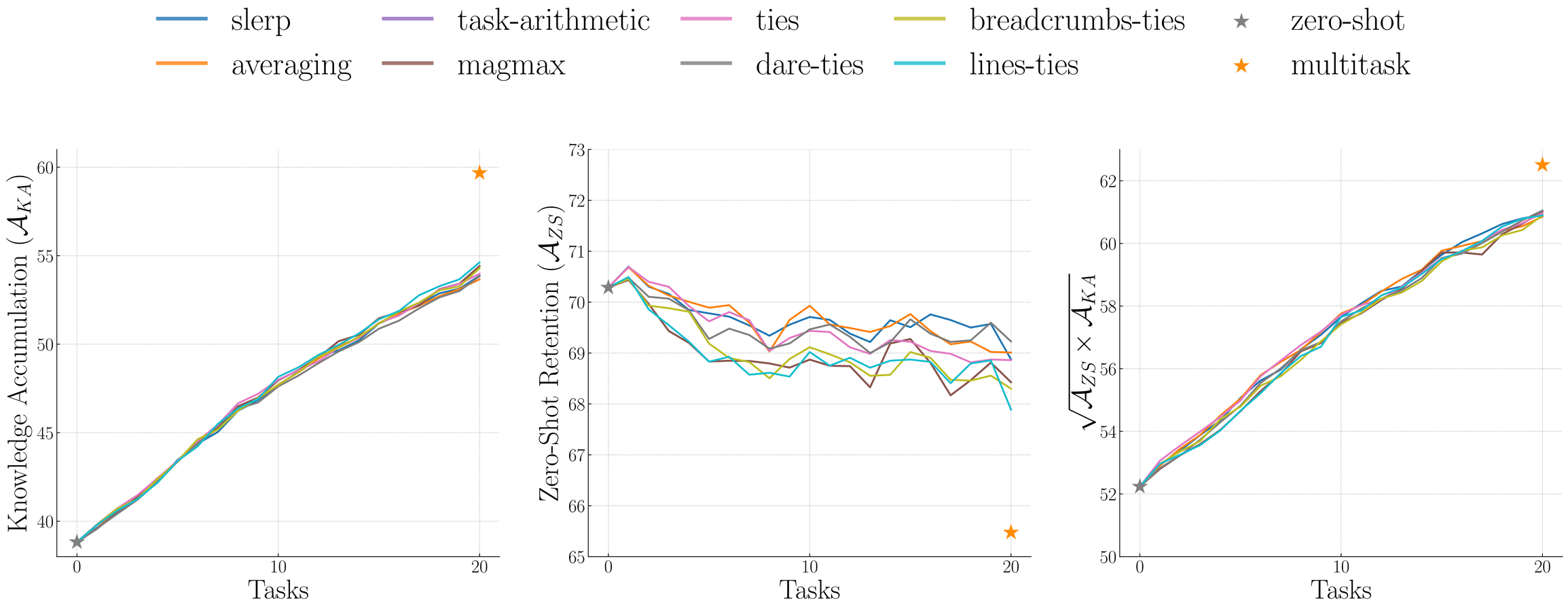
label: an image of an eye with no diabetic retinopathy

caption: a close-up light microscopy picture of a healthy eye reveals intricate details

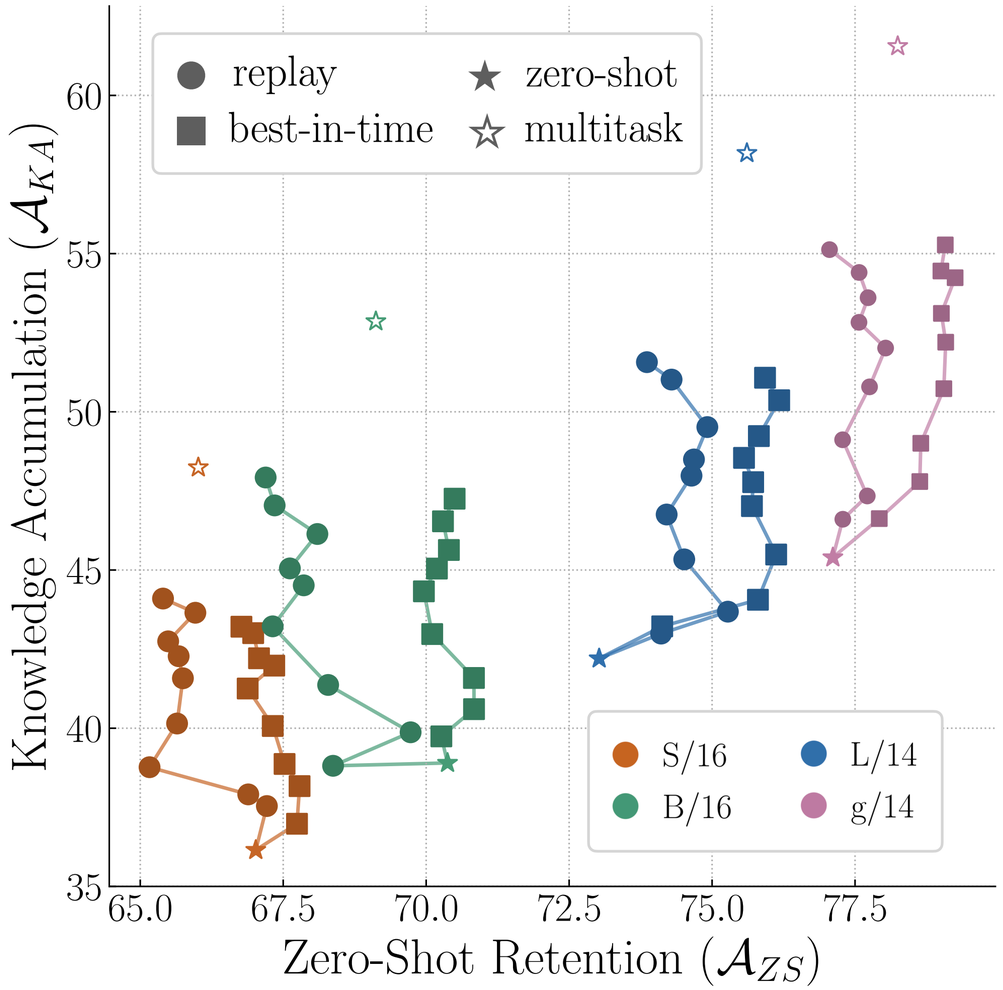
Initialization and deployment is important

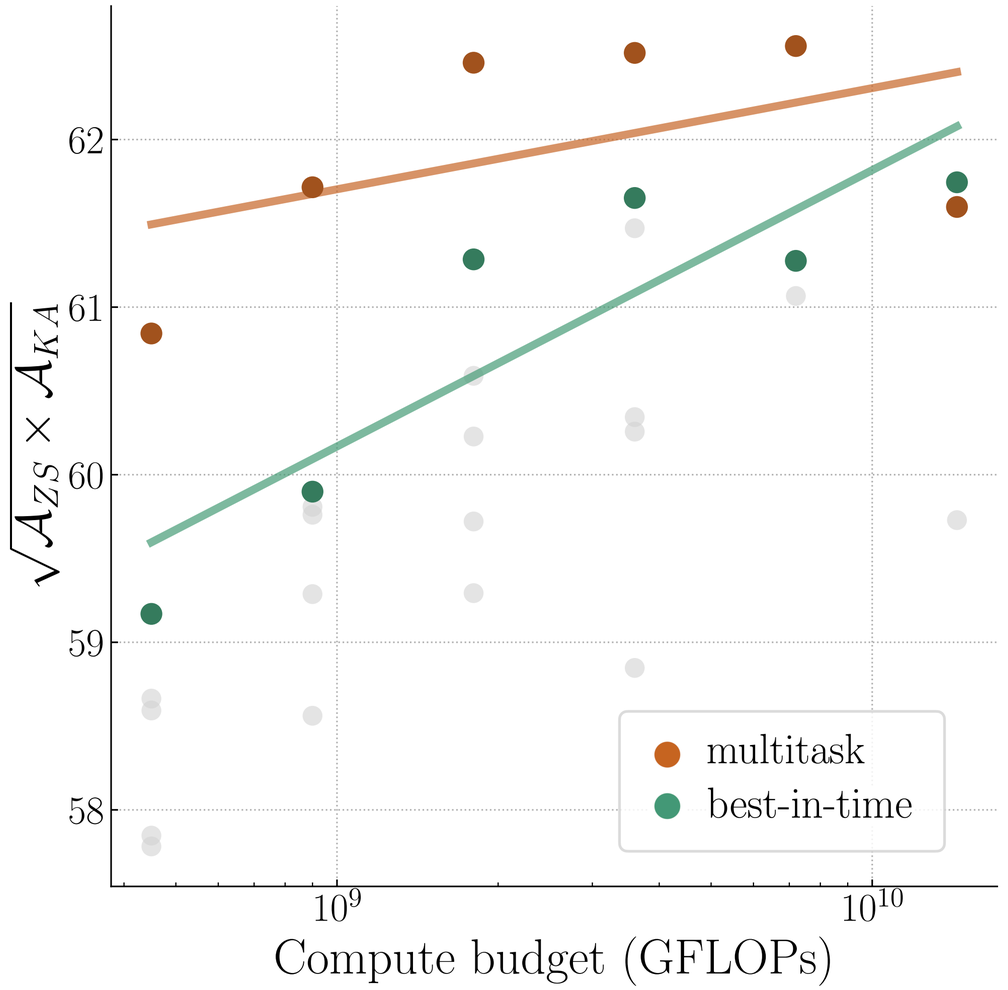


Merge technique matters less



Merging scales well with model size and compute





Thank you!

27.04.2025 · ICLR MCDC · @sbdzdz