Continual Learning Mini-Review

Scope. Survey of CL for task-incremental classification with known boundaries: (1) regularization; (2) replay; (3) distillation; (4) parameter isolation/adapters.

Regularization. EWC, SI, MAS penalize movement on important weights (diagonal approximations). Pros: tiny memory; Cons: weaker under large domain shift.

Replay. ER buffers prior samples; iCaRL prototypes; GEM/A-GEM gradient constraints. Pros: strong retention; Cons: memory + sometimes extra compute.

Distillation. LwF/POD keep teacher logits as soft targets. Pros: no raw exemplars; Cons: teacher quality critical; weaker for large shifts.

Isolation/expand. PNN, PackNet, DEN dedicate or grow parameters. Pros: near-zero forgetting; Cons: model size grows.

Adapters/LoRA. Task-specific adapters with shared backbone; low per-task memory; good for fast adaptation.

Metrics. Average accuracy, backward transfer, forgetting; validate on replay buffer and monitor old-task accuracy during new training.