Ảnh dụng Deep Reinforcement Learning (DRL) trong việc chọn giữa SAM và SGD để tối ưu chi phí và hiệu quả huấn luyện 1. Bối cảnh SAM (Sharpness-Aware Minimization) là một kỹ thuật hỗ trợ tối ưu loss giúp cải thiện độ tổng quát của mô hình nhưng tốn kém chi phí do cần tính gradient hai lần/mỗi bước. Do đó, ta đề xuất một cách tiếp cận sử dụng Deep Reinforcement Learning (DRL) để quyết định tại mỗi bước huấn luyện xem nên dùng: SAM (hiệu quả cao, tốn chi phí) hay SGD thông thường (nhanh, rẻ) Mục tiêu: Tối ưu độ chính xác trong khi giảm thiểu chi phí huấn luyện. 2. Cách định nghĩa bài toán DRL • State (Trạng thái): Vector đặc trưng tiền trình huấn luyện, bao gồm: Loss train Accuracy validation Gradient norm / sharpness Sự thay đổi loss giữa 2 bước Epoch hiện tại / tổng epoch • Action (Hành động): 0: Dùng SGD 1: Dùng SAM • Reward (Phần thưởng): Phụ thuộc vào: Độ cải thiện acc\_val Chi phí tính toán (phạt nặng khi dùng SAM không cần thiết) Công thức gợi ý: reward = α \* ΔAcc\_val - β \* Cost(SAM) 3. Triển khai agent • Mô hình DRL: Mạng nhỏ (MLP) dựa trên state vector DQN (nếu action rời rạc), hoặc PPO (nếu muốn continuous) • Huấn luyện agent: Chạy huấn luyện trên tập proxy (VD: CIFAR-10) Mỗi epoch, agent quyết định action Tính reward, update policy 4. Khả năng tái sử dụng Sau khi huấn luyện xong, agent có thể được lưu lại và dùng trong nhiều bài toán khác (nhiều task / dataset) Đặc biệt hiệu quả khi các task đó có bản chất tương đồng (VD: image classification) Trong thực tế, agent có thể chỉ cần vài ms để quyết định hãy suy nghĩ kĩ về cách thực hiện bạn có thể cải thiện cách làm không áp dụng cho object detection cho bộ cifa 100

* Cách DQN hoạt động trong bài (học từ đâu,…)