Final Project: Tiny and Efficient Model for Edge Detection Generalization

Team Members - Group 4

- Huda Aisyah Khoirunissa (胡艾莎)
- Steven Albert Yosephus Ndun (斯帝芬)
- Yu-Ti Tu (杜于緹)
- Novia Brilianty (諾菲雅)

Score: 94/100 ⭐ ⭐

Overview

This project enhances the TEED (Tiny and Efficient Edge Detector) model with two complementary approaches: targeted improvements for hard regions (foliage-dominant areas) and quantization-aware training for efficient deployment.

Problem Statement

TEED achieves good generalization on common edge detection benchmarks, but tends to **under-activate in foliage-dominant regions** where boundaries are subtle and texture is dense.

Proposed Improvements

Area (A): Hard Region Enhancement

Target: Improve edge detection in leaves and foliage-dominant scenes

Method: Multi-modal edge extraction combining three components:

1. **Color Prior (HSV Space)**
   - Identify green-colored regions to capture vegetation areas
2. **Texture Information**
   - High-texture regions using gradient magnitude: `sqrt(gx² + gy²)`
   - Computed from the V channel
3. **Early Deep Features**
   - Utilize TEED's early output (`out_1`) for fine-grained edge details
   - Early features: rich texture but noisy
   - Final predictions: stable but overly smooth

Formula:

```
edge_prob = final + γ × (out1 - final) × tex × gate
```

Benefits:

- Targeted enhancement without over-sharpening
- Significantly improves edge detection in foliage scenes
- Balances stability and detail

Area (B): Quantization-Aware Training (QAT)

Objective: Optimize TEED for efficient deployment on edge devices

Motivation:

- Original TEED (FP32): High accuracy but resource-intensive
- Target: Lightweight INT8 model for edge deployment
- Challenge: Maintain quality during precision reduction

Training Strategy:

1. FP32 warm-up
2. QAT training with fake quantization
3. Post-training quantization attempts
4. Qualitative evaluation on UDED dataset

Results:

- Validation loss: 0.0410
- MAE: 0.048
- Successfully maintains acceptable quality
- Model demonstrates improved robustness to quantization noise

Limitations:

- Full INT8 conversion not achievable due to architectural constraints
- Residual connections and tensor reduction operations unsupported by current PyTorch quantization
- Provides valuable insights for future quantization-compatible architecture design

Key Findings

Area (A) - Hard Region Improvement

Successfully combines color prior, texture information, and early deep features to avoid over-sharpening while enhancing foliage edge detection.

Area (B) - Quantization-Aware Training

QAT training converges successfully while preserving edge quality, though full INT8 deployment is limited by PyTorch quantization constraints.

Evaluation Metrics

For QAT Model:

- **Error Metrics** (Lower is better): MAE, MSE, RMSE
- **Similarity Metrics** (Higher is better): SSIM, PSNR, Correlation

Conclusion

This work demonstrates TEED's adaptability for both performance-oriented optimization (enhanced leaf edge extraction) and compression-oriented deployment (stable QAT training with preserved quality). The findings provide practical insights into task-specific optimization and the challenges of quantizing complex edge detection networks.

Future Work

1. Extend TEED to multi-scale edge detection for better fine and coarse structure capture
2. Integrate into higher-level vision tasks (object detection, segmentation)
3. Improve robustness to noise
4. Optimize for efficient deployment on edge and embedded devices
5. Design quantization-compatible architectures supporting full INT8 conversion

---

Final project completed as part of Deep Learning coursework