Overview

Structured predictions output correlated 2D masks, including semantic segmentation, depth estimation, surface normal prediction, etc.

IID (softmax/L2) is the most common approach, which ignores label correlations among pixels.

GAN is a prior based structural model, which encodes relationships in a one-to-one mapping.

Our ASM adversarially matches multiscale structures in the label space, featuring:

1. Adaptive structure prior
2. Instance specificity
3. Generalizability.

Adaptive Affinity Fields (2\textsuperscript{nd} order regularization): AAF only selects pixel relationships adaptively.

Adversarial Structure Matching for Structured Prediction Tasks
Jyh-Jing Hwang*, Tsung-Wei Ke*, Jianbo Shi, Stella X. Yu

Adversarial Structure Matching

What’s Learned in Analyzer A?

- Analyzer A is trained to maximize the multiscale structural mistakes of S. → Hard negative mining.
- Predictor S is trained to minimize the same error.
- Regularizer R ensures that A also forms a good basis for reconstructing the ground truth.
- Assuming infinite capacity for S and A, we proved $S^*(\cdot) = y$ and $V^*(A) = 0$. (Nash equilibrium)
- ASM retains critical assessments thru training:
  - A encodes multiscale pixel relationships, e.g., person riding bike, hand picking up bottle, etc.
  - A and R learn to complete shapes while A adaptively refines the focus of supervision.

Experimental Results