

Soft Target Training

Hard Targets

[0, 0, 1, 0]

One-hot vector

Only correct class is 1

✗ No inter-class relationship information

Soft Targets

[0.05, 0.15, 0.70, 0.10]

Probability distribution

Probabilities for all classes

✓ Learns inter-class similarity

Benefits of Soft Targets

- ✓ Richer information: Transfers inter-class relationships learned by Teacher
- ✓ Improved generalization: Prevents overfitting
- ✓ Medical imaging: Utilizes disease similarity information (e.g., borderline benign/malignant cases)