Figure 1a: Training Stability Comparison
Standard KD vs Decoupled KD

- v2: Standard KD (76.19%)
- v3: DKD (β=8.0) (66.85%)