World Models at Chicago Booth

ALTOS™

VISReg for World Model Training

Haiyu Wu[1] · Randall Balestriero[2] · Morgan Levine[1]

[1] Altos Labs [2] Brown University

VISReg stabilizes latent representations and improves planning across tasks and longer horizons.

01 VISReg

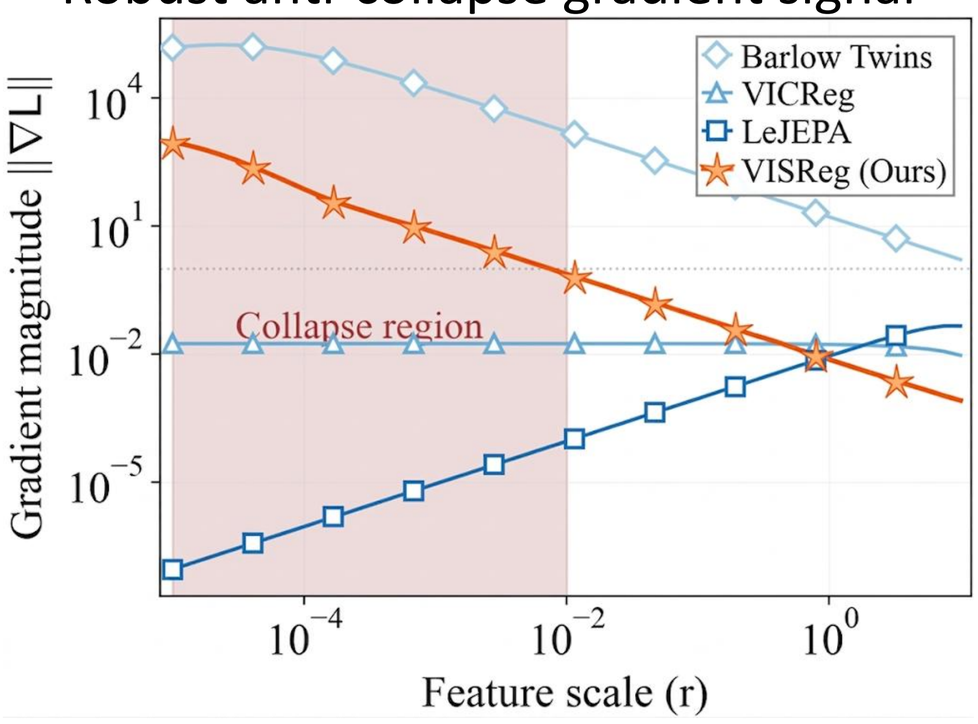
$$\mathcal{L}_{\text{scale}} = \frac{1}{D}\sum_{j=1}^{D}(1-\sigma_j(\hat{\mathbf{Z}}))^2$$

$$\mathcal{L}_{\text{shape}} = \frac{1}{K}\sum_{k=1}^{K}\left\|\text{sort}(\tilde{\mathbf{Z}}w_k) - \mathbf{q}_{\mathcal{N}}\right\|_2^2$$

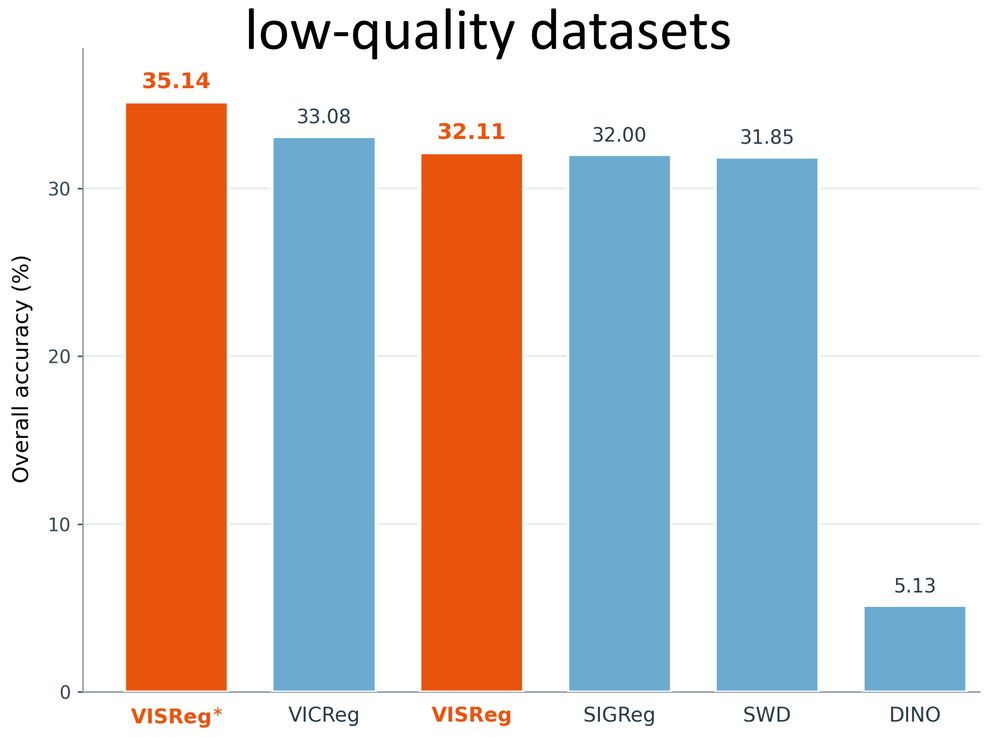
$$\mathcal{L}_{\text{center}} = \|\mu\|_2^2$$

What does it bring?

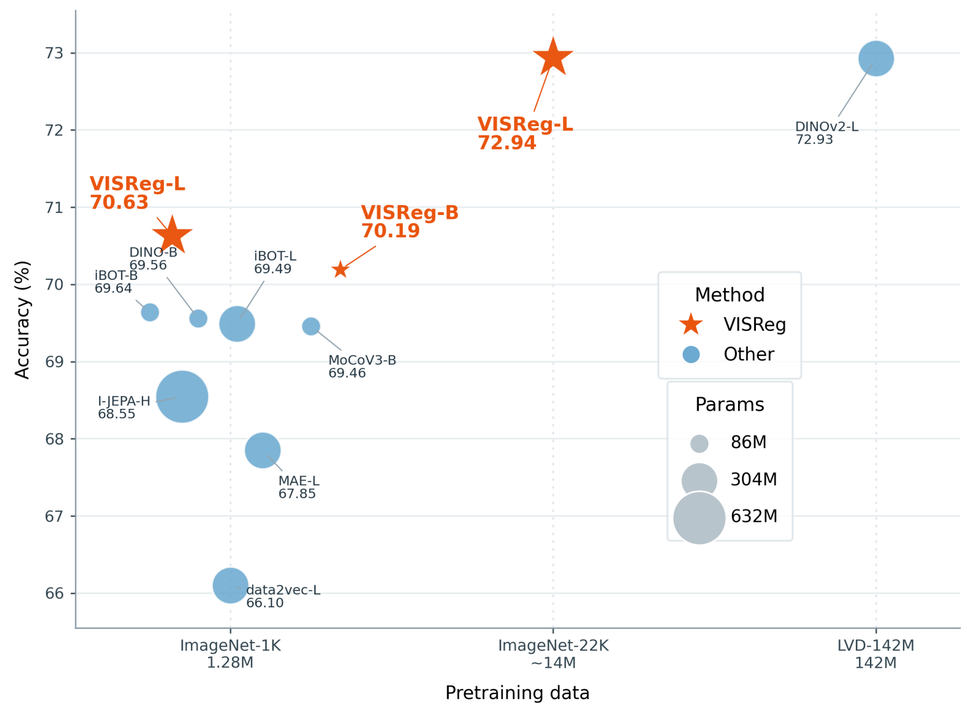
Robust anti-collapse gradient signal



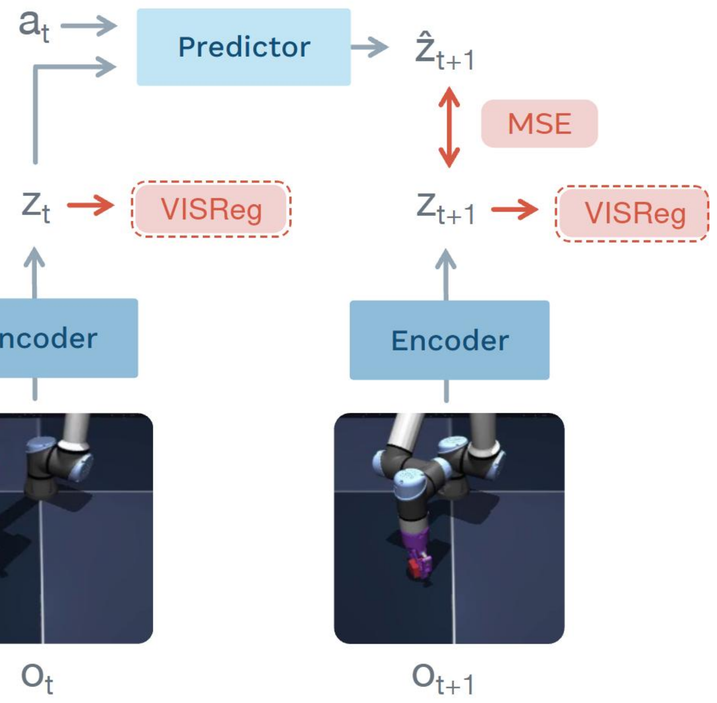
Reweighting brings a better result on low-quality datasets



Strong OOD performance



02 VIS-WM



03 Cross task performance

04 Long-horizon on PushT