

## Advanced Self-Supervised Methods

Architecture Comparison: MoCo vs BYOL



### MoCo

Momentum Contrast

Query Encoder



Query

Momentum  
Encoder



Key

Queue of Negative Samples



**Queue-based:** Maintains large dictionary of negatives



**Momentum update:** Consistent representations over time



**Contrastive:** Uses negative pairs



### BYOL

Bootstrap Your Own Latent

Online Network



Predictor

Target Network



Projection

Stop-Gradient



**No negatives:** Only positive pairs needed



**Predictor network:** Asymmetric architecture



**Stop-gradient:** Prevents collapse

✓ Excellent ImageNet Performance

✓ Excellent ImageNet Performance