

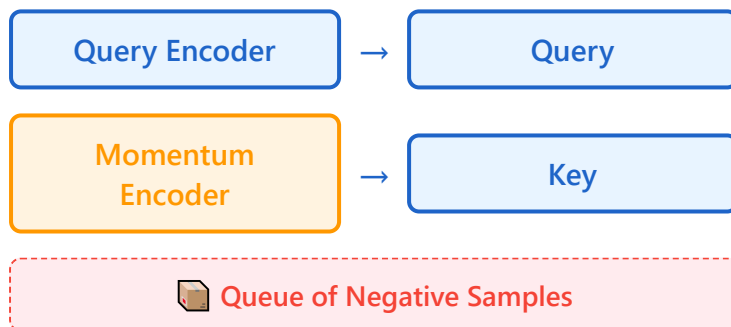
## Advanced Self-Supervised Methods

Architecture Comparison: MoCo vs BYOL



### MoCo

Momentum Contrast



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



**Momentum update:** Consistent representations over time

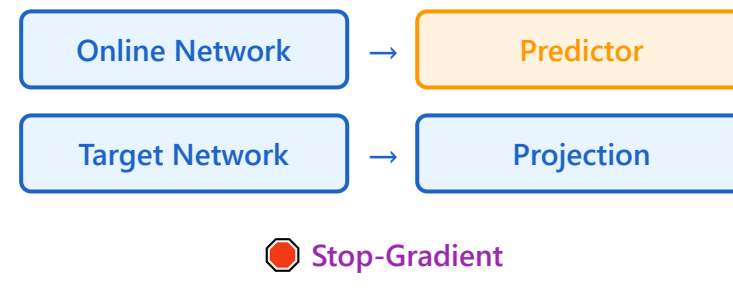


**Contrastive:** Uses negative pairs



### BYOL

Bootstrap Your Own Latent



**No negatives:** Only positive pairs needed



**Predictor network:** Asymmetric architecture



**Stop-gradient:** Prevents collapse

✓ Excellent ImageNet Performance

✓ Excellent ImageNet Performance