

## Comparison of Normalization Techniques

Technique	Best Use Cases	Key Characteristics
Batch Norm	CNNs Large batches	Normalizes across batch dimension
Layer Norm	RNNs Transformers Small batches	Independent of batch size
Instance Norm	Style Transfer GANs	Per instance, per channel
Group Norm	Small batch vision Detection Segmentation	Groups channels, batch-independent
Weight Norm	RNNs Reinforcement Learning	Normalizes weights, not activations
Spectral Norm	GAN discriminators BigGAN StyleGAN	Ensures Lipschitz continuity

**Key Insight:** Each normalization technique addresses different aspects of training stability. Choose based on your architecture, batch size constraints, and specific task requirements.