Solutions to Deep Learning Challenges

Better Activation Functions ⚡

Prevents vanishing gradients during backpropagation

ReLU, Leaky ReLU, ELU, GELU

→ Vanishing Gradient

Smart Initialization 🎯

Maintains gradient scale across layers

Xavier, He Initialization

→ Gradient Issues

Normalization Techniques 📊

Stabilizes training and speeds up convergence

Batch Norm, Layer Norm

→ Training Stability

Skip Connections 🔗

Creates gradient highways through the network

ResNet, Highway Networks

→ Vanishing Gradient

Careful Optimization 🚀

Adaptive learning rates for stable updates

Adam, RMSprop, AdaGrad

→ Training Efficiency

Regularization & Clipping 🛡️

Prevents overfitting and gradient explosion

Dropout, Weight Decay, Grad Clipping

→ Overfitting & Explosion

Modern Deep Learning

Robust training is achieved by combining multiple techniques together