Challenges of Deep Networks

Training Complexity

- Non-convex optimization
- Longer training time
- High computational cost

Gradient Problems

- Vanishing gradients in early layers
- Exploding gradients
- Unstable parameter updates

Generalization Issues

- Overfitting risk
- Memorizing training data
- Poor test performance

Configuration Sensitivity

- Critical initialization requirements
- Hyperparameter sensitivity
- Architecture design choices

⚠ HISTORICAL BARRIER (Pre-2006)

Deep networks often performed worse than shallow ones due to these challenges

These challenges motivated research into better training techniques and architectures