

## Practical Design Guidelines

### **START SIMPLE**

Begin with proven architectures (ResNet, EfficientNet), modify as needed

### **Architecture Choices**



- ▶ Depth vs. width: Deeper generally better
- ▶ Residual connections for > 10 layers

### **Key Components**



- ▶ Batch norm: After conv, before activation
- ▶ Activation: ReLU/GELU (avoid sigmoid/tanh)

### **Regularization Stack**



- ▶ Data augmentation
- ▶ Dropout
- ▶ Weight decay

### **Architecture Principles**



- ▶ Gradually reduce spatial size
- ▶ Increase channels through network



### **Validate on Real Hardware**

Consider **inference speed** and **memory**, not just accuracy