Fine-tuning Strategies Comparison

Full Fine-tuning

Update all parameters

- Task Head ✓
- Top Layers ✓
- Middle Layers ✓
- Bottom Layers ✓

✓ Best performance
⚠ Slower training
⚠ Risk of forgetting

Feature Extraction

Freeze base model

- Task Head ✓
- Top Layers ✗
- Middle Layers ✗
- Bottom Layers ✗

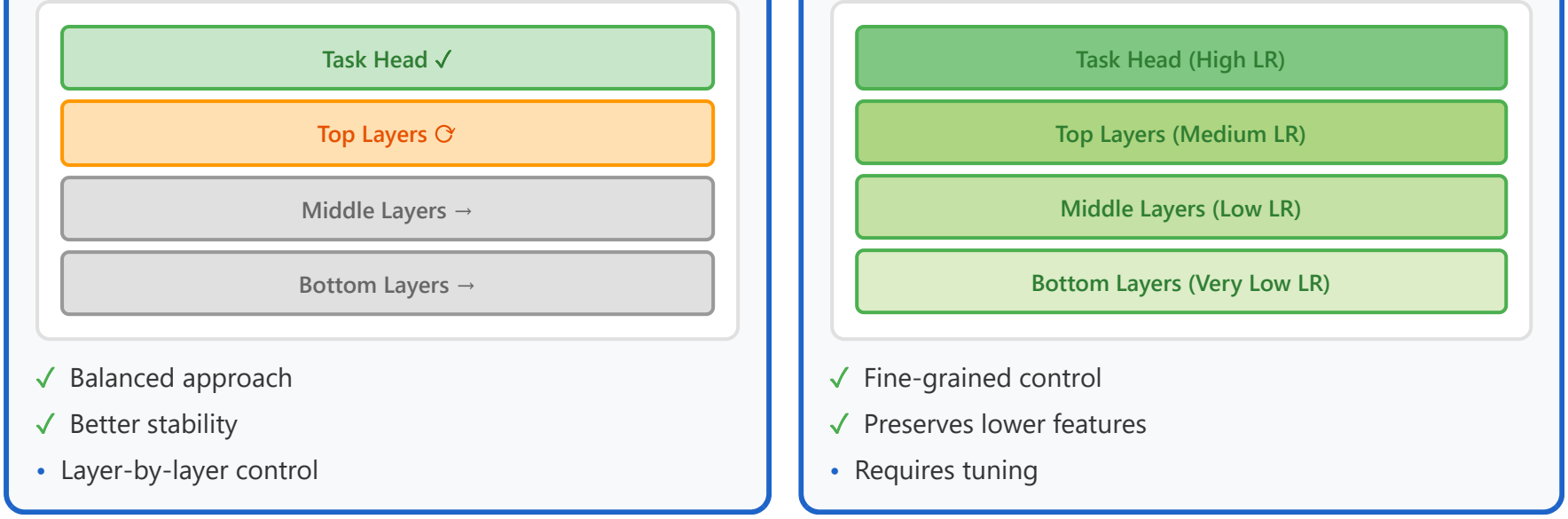
✓ Faster training
✓ No catastrophic forgetting
⚠ Lower performance

Gradual Unfreezing

Progressive unlocking

Discriminative LR

Different rates per layer



✓ Balanced approach
✓ Better stability
• Layer-by-layer control

✓ Fine-grained control
✓ Preserves lower features
• Requires tuning

🎯 Performance

Full FT: **Best**

Feature Ext: **Good**

⚡ Speed

Feature Ext: **Fastest**

Full FT: **Slowest**

💾 Memory

Feature Ext: **Low**

Full FT: **High**

💡 Decision Factors

Choice depends on **data size** and **computational budget** • Large data + resources → Full FT • Small data / limited resources → Feature Extraction