Multi-Teacher Distillation
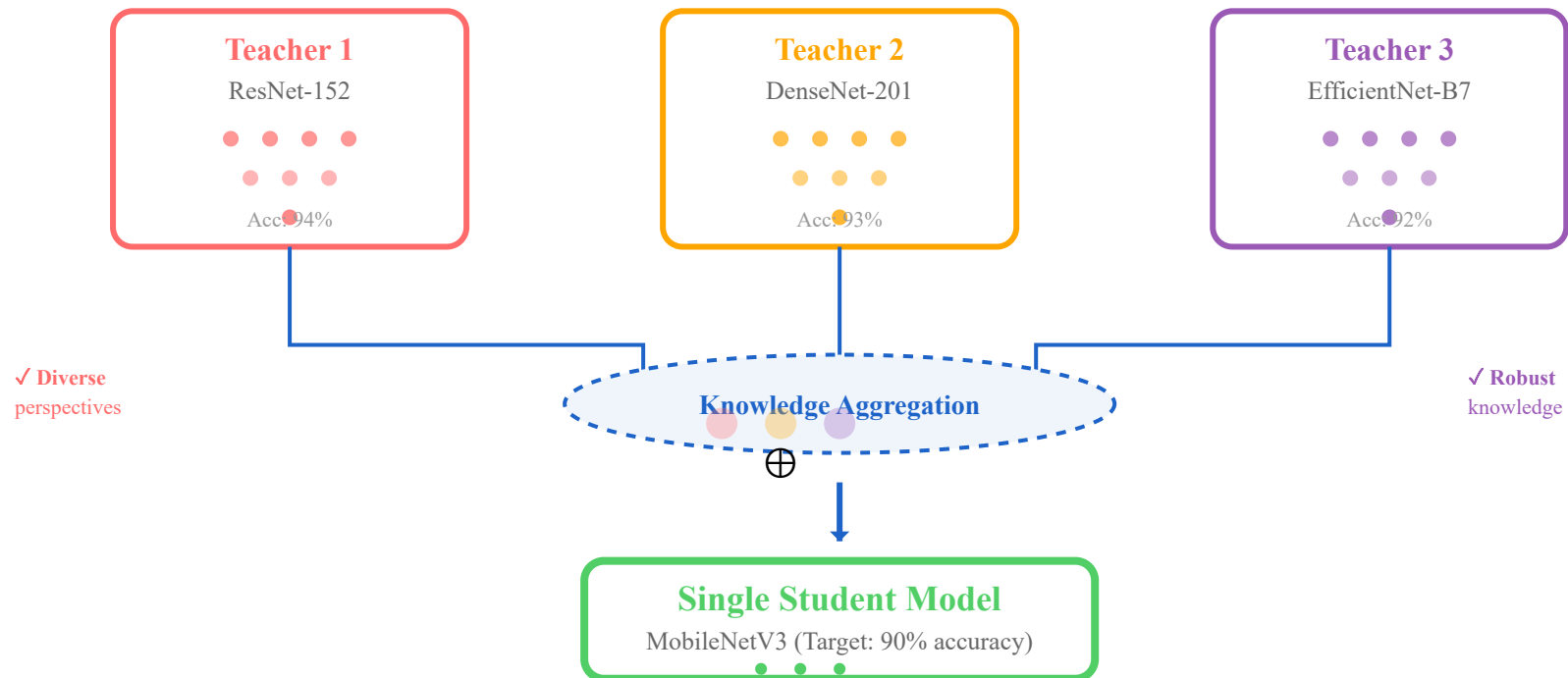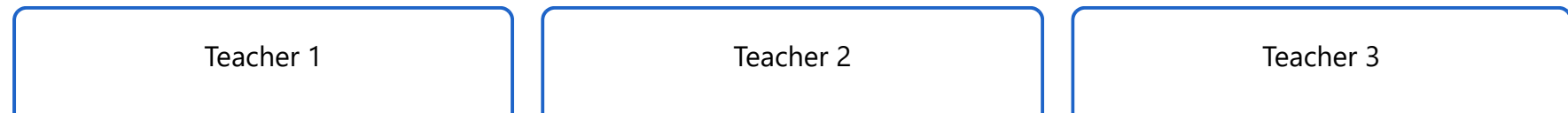
Multi-Teacher Distillation

Transferring ensemble knowledge from multiple Teacher models to Student

Teacher 1
ResNet-152

Acc:94%

Teacher 2
DenseNet-201

Acc:93%

Teacher 3
EfficientNet-B7

Acc:92%

✓ **Diverse** perspectives

Knowledge Aggregation

⊕

✓ **Robust** knowledge

Single Student Model
MobileNetV3 (Target: 90% accuracy)

Multiple Teachers

| Teacher 1 | Teacher 2 | Teacher 3 |

| ResNet-152 | DenseNet-201 | EfficientNet-B7 |

⬇️ Knowledge Aggregation ⬇️

Single Student Model

MobileNetV3

Knowledge Integration Methods

| Equal Weight | fferential Weig | ynamic Selecti |
| Average | Weighted Average | Attention-based |