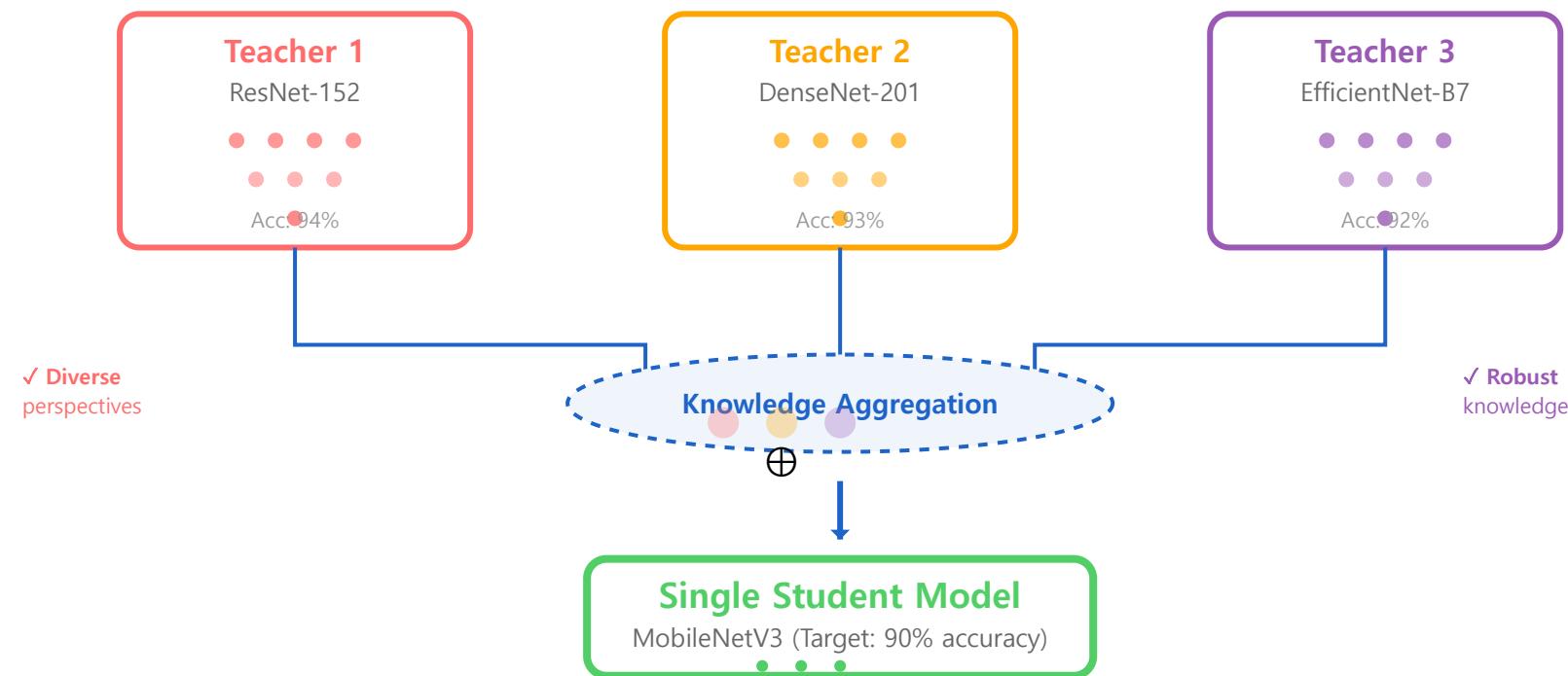


Multi-Teacher Distillation

Multi-Teacher Distillation

여러 Teacher 모델의 양상을 지식을 Student에 전달



Multiple Teachers

Teacher 1

Teacher 2

Teacher 3

ResNet-152

DenseNet-201

EfficientNet-B7

↓ Knowledge Aggregation ↓

Single Student Model

MobileNetV3

지식 통합 방법



평균 (Average)



가중 평균 (Weighted)



동적 선택
Attention 기반