

Case Study: Mobile Diagnostics

사례 연구: 모바일 피부암 진단 앱

딥러닝 모델 압축을 통한 실시간 모바일 진단 시스템 구현

문제 정의



목표: 피부 병변 사진으로 양성/악성 분류



플랫폼: iOS/Android 모바일 앱



요구사항: 3초 이내 결과, 오프라인 동작

압축 접근법

Step 1: Base Model

ResNet-50 (98MB, FP32)

정확도: 94.5%

추론: 2.5초



Step 2: Knowledge Distillation

MobileNetV3 Student

크기: 24MB

정확도: 93.2%




Step 3: INT8 Quantization

크기: 6MB (75% 감소)
정확도: 92.8%
추론: 0.8초



Final: TFLite Optimization

최종 크기: 5.2MB
정확도: 92.5%
추론: 0.6초 

최종 결과

94.7%

크기 감소

76%

속도 향상

-2.0%

정확도 변화