DermaLite

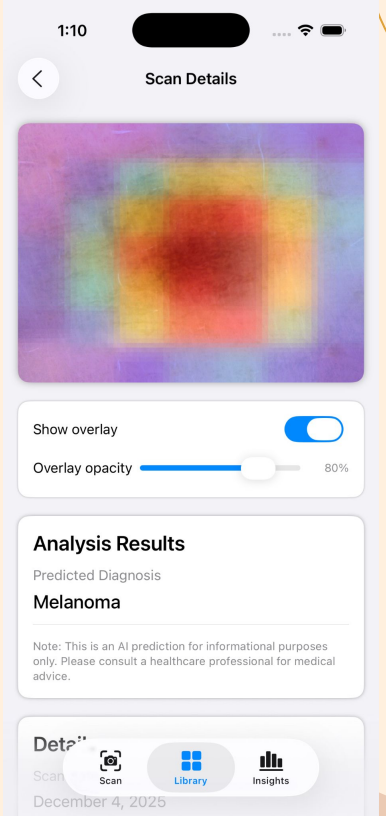
Ryan Caudill, Theo Villalva, & DJ Ravenell

DermaLite

Create an mobile app where a user can take a picture of a mole and be told:

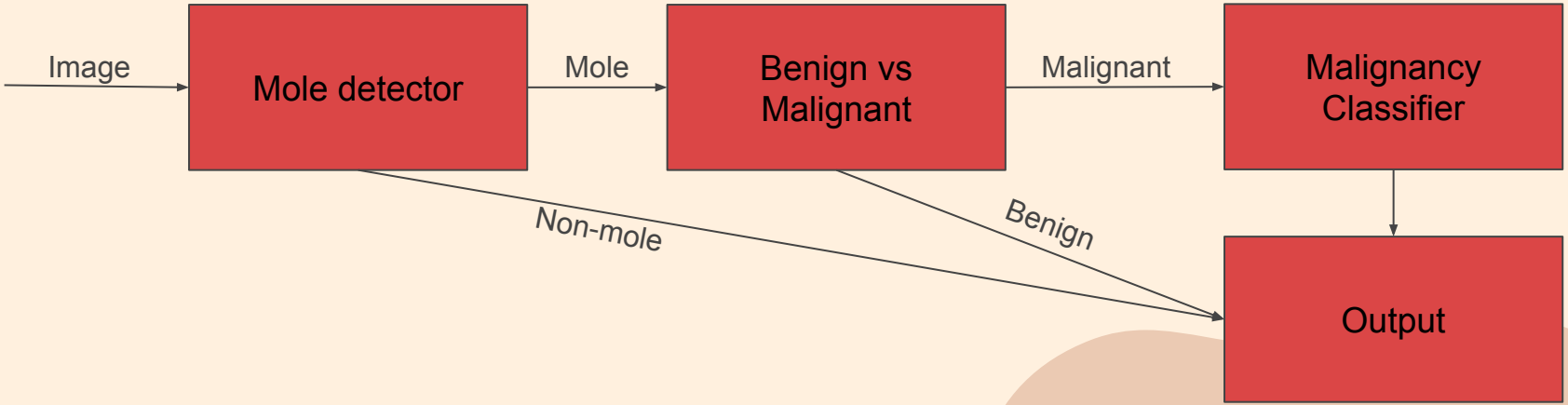
- If it is a mole
- Benign or Malignant
- Type of malignancy

All done on-device for privacy and efficiency



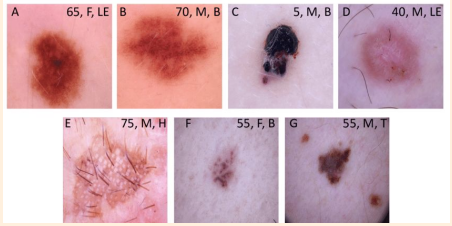
Method

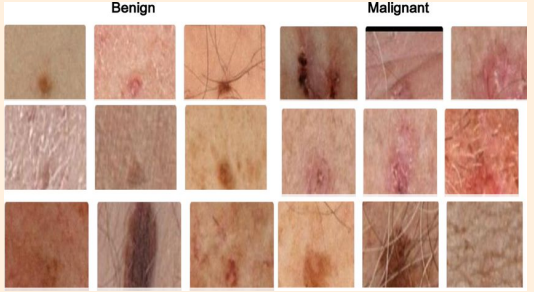
- Original plan
  - Single model, show output to user
  - Inaccurate, unreliable
- New plan



Mole Detector

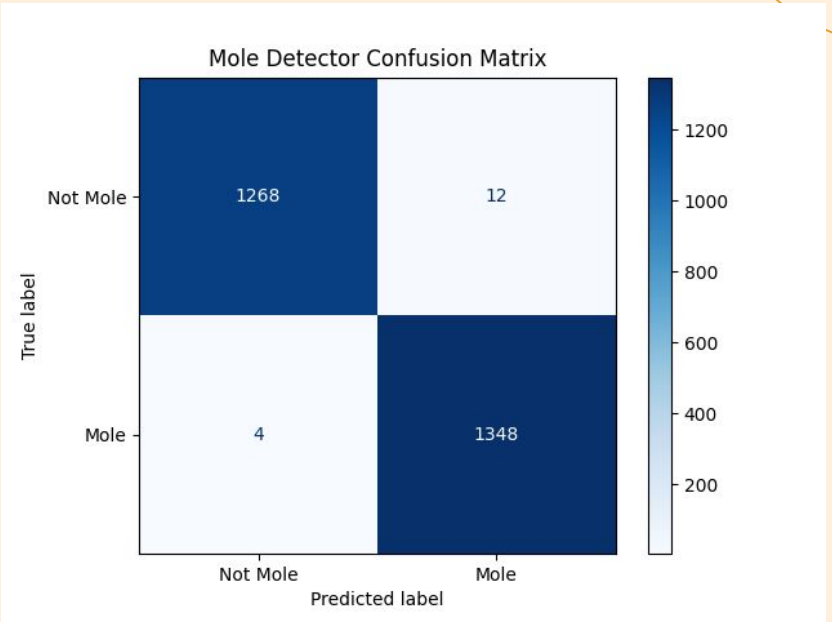
- Filter out pictures of non-moles
- Challenges
  - Initial Attempt
    - Would not work with any non-dataset image
    - Determined on dermoscopic imaging side-effects
      - Lens distortion, high-quality imaging, scales
  - Second Attempt
    - Heavy preprocessing
      - Cropping, Gaussian blur and noise, grayscale





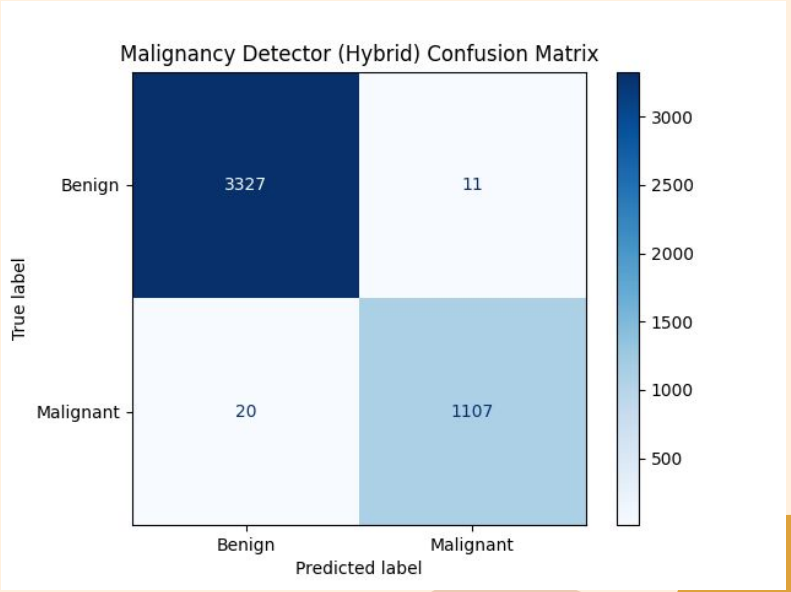
Mole Detector

- Results seemed quite effective
- Limitations
  - Skin color
  - Size of mole
  - Type of lesion
    - Warts, birthmarks, burns



Benign vs Malignant

- ResNet-50
- Challenges
  - Naturally more difficult and specific
  - Limited to HAM10000 and ISIC 2024
    - 3x benign vs malignant case
    - Malignant is more important
  - Training would plateau
- Solution
  - Preprocessing (same as previous model)
  - Replaced FC layer with random forest



Cancer Classifier

- MobileNetv2 (7 classes)
    - Default Batch normalization
    - 80% train / 20% validation
- Preprocessing
- Class Balancing
    - Oversampling of malignant class and class weights
- 15 Epochs
- Validation accuracy = 89%

Core ML 2-Stage Pipeline Evaluation

- Mole detector -> malignant classifier
- Critical oversight in malignancy detection (1k size sample)
    - Binary Accuracy: 84.10%
    - Cancer Recall: 11%
    - 89% of cancer cases were missed
    - Cancer Precision: 64%
    - 36% of cancer predictions were false alarms

Core ML 3-Stage Pipeline Evaluation

- Still critical, but better than before
- Confusion metrics (1k size sample):
    - Binary Accuracy: 46.70%
    - Cancer Recall: 30%
        - 70% of cancer cases were missed
    - Cancer Precision: 11%
        - 89% of cancer predictions were false alarms
    - Sensitivity over specificity for medical detection products

Future Improvements

- Improved accuracy
- Guide on taking/submitting ideal images
- Tell user next steps

Demo

Discussion

- Results
  - The task was more complex than we initially thought
  - Could possibly be viable for clinical use
    - Not prepared for real users
- Lessons Learned
  - Overfitting
  - Data augmentation
  - Data availability
  - Training techniques
- Limitations
  - Skin color
  - Malignancy type
  - On device computational limitations