Mobile Deep Learning Digital Narratives

Priya Pai B.A. Columbia University

We created an app that captures live video and outputs the highest emotion frames.

Pressing Problem: Mobile video capture rises by 100% every year. What frames should we focus on when dealing with so much video?

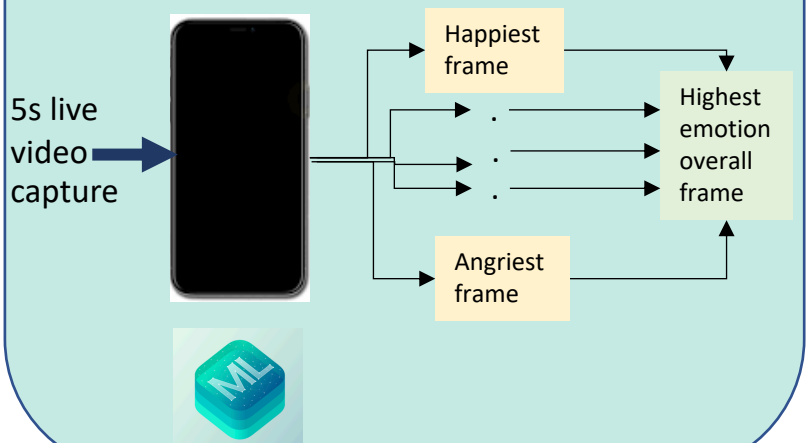
Goals:

- Optimize mobile user experience with video using intelligent video understanding models
- Extract “significant” frames given a selfie video.

Key Internship Question:

How can we combine facial detection & emotion recognition on-device for video editing purposes?

High-Level Control Flow:



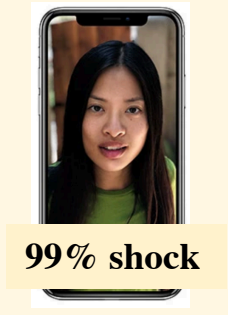
Technical Control Flow:

AVVideo Stream created

Face(s) detected (OpenCV haarcascade)

Emotion(s) detected. (CNNEmotions ML model – EmotionNet) **Add frame to array.**

AVVideoStream closed (after 5 seconds)



Loop over array of frames. Find highest emotion frame. Output image.

Research Mentors: Fabian Caba Heilbron, Joon-Young Lee

Adobe