

# Mobile Vision API

Shweta Garg



# Agenda

---

- Introduction
- Detectors
- Use cases
- Hands-On
- References
- Q&A

# Introduction

---

- A framework by Google for Mobile without cloud
- Capability to find objects in photos and videos
- Works on real time on device as well as on already clicked and recorded videos
- What all you can detect from this API
  - Face
  - Barcode
  - Text

# Android & iOS

---

- Android Framework
  - Face
  - Barcode
  - Text
- iOS Framework
  - Face
  - Barcode

# Detectors

---

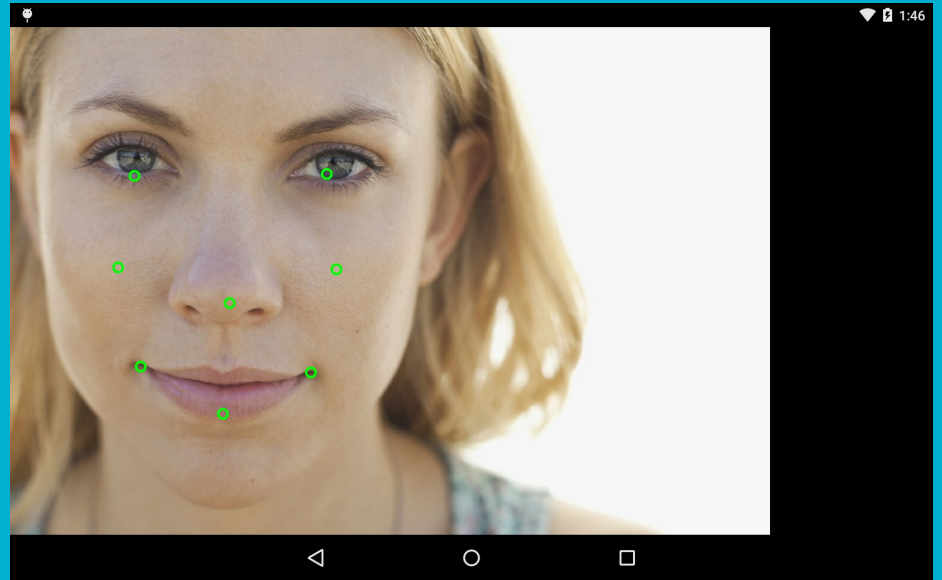
- It uses detectors to locate visual objects in images and video frames.
- Detectors return position of the object.
- Single detector or multiple detectors
- 3 detectors
  - Face Detectors
  - Barcode Detectors
  - Text Detectors

# Face Detector

---

- Face tracking
- Detects Face landmarks e.g. eyes, nose and mouth
- Face classification e.g. eyes open, smiling

Image source: Google Mobile Vision doc



# Barcode Detector

---

- Detect barcode in any orientation
- Multiple Barcodes at once
- Supports following Barcodes formats
  - 1D Barcodes (EAN-13, EAN-8, UPC-A, UPC-E, Code-39, Code-93, Code-128, ITF, Codabar)
  - 2D Barcodes (QR Code, Data Matrix, PDF-417, AZTEC)

# Text Detector

---

- Any Latin based language (e.g. English, Spanish, German, French, etc.)
- It segments text into blocks, lines, and words
- Block - set of lines, such as a paragraph or column,
- Line - set of words in a block
- Word - set of characters in a line



# Text Detector

Image source: Google Mobile Vision doc

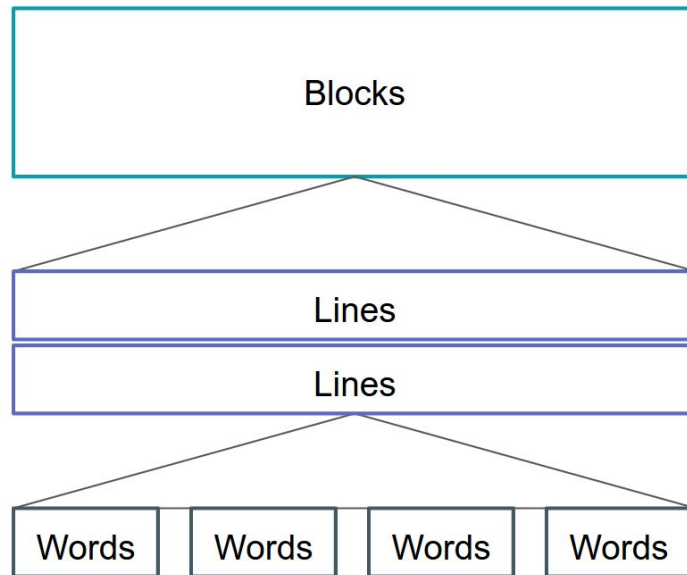
Dear Ms. Parker,

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed diam nonummy nibh euismod tincidunt ut laoreet dolore magna aliquam erat volutpat. Ut wisi enim ad minim veniam, quis nostrud exerci tation ullamcorper suscipit lobortis nisl ut aliquip ex ea commodo consequat.

Duis autem vel eum iriure dolor in hendrerit in vulputate velit esse molestie consequat, vel illum dolore eu feugiat nulla facilisis at vero eros et accumsan.

Nam liber tempor, cum soluta nobis eleifend option congue nihil imperdiet doming id quod mazim placerat facer possim assum. Typi non habent claritatem insitam; est usus legentis in iis qui facit eorum claritatem. Investigationes demonstraverunt lectores legere me lius quod ii legunt saepius.

Sincerely,



# Use cases

---

- Image sentiment analysis
- Organise photos based on the content
- Convert books into Text for your phone
- Scan multiple codes on live camera
- Read business cards
- React if people winks and smile
- Data entry jobs can be easier
- Gaming

# Text Detector Hands-On

---

- Clone or download the repo

<https://github.com/shweta-techjini/Start-TextRecognizer.git>

- Add mobile vision api in gradle

compile 'com.google.android.gms:play-services-vision:10.0.1'

- Sync gradle

# Limitations

---

- Does not support face recognition - it does not determine if two faces are likely to correspond to the same person
- Face detector supports only 2 classifications
  - Eyes open
  - Smiling
- Text detector reads randomly so you won't get the text in the same sequence in which they are in frame
- Text detector is not 100% accurate

# References

---

- [Google Mobile Vision](#)
- [Face Detection Code Labs](#)
- [Barcode Detection Code Labs](#)
- [Text Detection Code Labs](#)

Contact

shweta.grg@gmail.com

Thank you!!  
Q&A