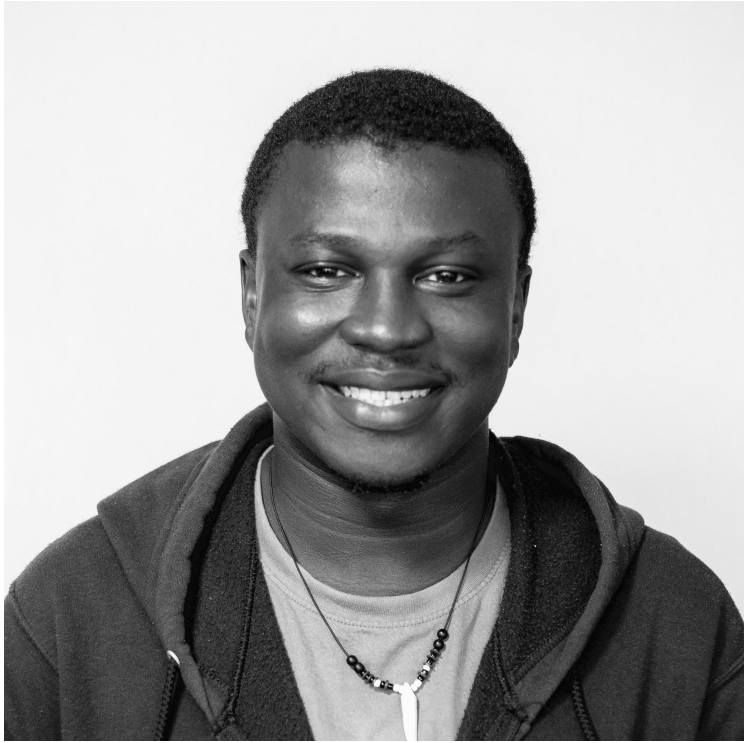
Innovating with Google Mobile Vision APIs

A case for **adKandi**, the greatest thing since sliced bread.

Well, sup!!

Akapo Damilola F.

- Android, iOS, IoT && Anything that solves the problem.
- Senior Consultant, Andela Inc.
- Can sing sometimes. Scared to death of HTML and CSS.
- Twitter: **@_fdamilola**
- **Ignore the fake smile ->**



“Every attempt at innovation starts with a problem.” - **Some apparently wise dude.**

The world is changing

- We’re getting access to cool tools used by big tech. Machine learning tools, AI models etc.
- These “tools” are opening up opportunities to explore and solve unique problems.
- Devices are evolving, becoming more powerful thus making edge computing much more easier

TOOLS

- **Microsoft**: Project Oxford (https://github.com/Microsoft/ProjectOxford-ClientSDK)
- **Google**: Cloud & Mobile vision, Tensorflow (https://cloud.google.com/vision/, https://developers.google.com/vision/, https://www.tensorflow.org)
- **IBM**: Watson (https://www.ibm.com/watson/)

Google CLOUD vision APIs

- Image processing on Google Cloud
- Easily **detect broad sets of objects** in your images, from flowers, animals, or transportation to thousands of other object categories commonly found within images
- Vision API enables you to detect different types of inappropriate content from adult to violent content.
- Image sentiment analysis and OCR
- Has a version for mobile

ENTERS GOOGLE MOBILE VISION

Google Mobile Vision APIs had what we needed to get started.

[P.S] We tried a number of other mobile focused Vision related SDKs but none stuck like the Vision API.

Why google mobile vision api

- Able to detect faces in video and pictures
- Works offline (After initial setup)
- Works locally
- Easy to integrate
- Free free free

HOW DO I GET STARTED WITH MOBILE VISION

- Introducing Mobile vision APIs: https://developers.google.com/vision/introduction
- Introduction to Mobile Vision APIs: https://hackernoon.com/machine-learning-for-android-developers-with-the-mobile-vision-api-part-1-face-detection-e7e24a3e472f
- Mobile Vision Documentation: https://developers.google.com/vision/android/detect-faces-tutorial

Cont: How do I get started with Mobile Vision

- “Seeing is Believing” talk by Moyin
- Totally Watch: https://www.youtube.com/watch?v=704qQzEnT0I

POINT(X, Y)

The point of this talk is to give you a glimpse into how we used Google Vision APIs in solving a couple of interesting problems in the digital advertising domain.

Let’s talk about **adKandi** and the problems therein.

adKandi: Coolest thing since sliced bread

adKandi is digital-out-of-home advertising platform that enables distribution of digital content to remote locations.

Adverts and Candies.

adKandi: Coolest thing since sliced bread

This simply means **adKandi** is a platform that provides advertisers with the means to distribute, track and measure the conversion rates of their advert content as efficiently as possible.

ADKANDI: COOLEST THING SINCE SLICED BREAD

Pfft! Why is this a big deal?

To answer that question, I'll need to explain what digital-out-of-home advertising platforms are and why they are "kewl".

Advert distribution channels you're familiar with

- Internet Advertising (Google, Facebook, Twitter, Snapchat)
- TV/Radio Advertising
- Newspapers
- Fill in the gaps maybe? :)

"DIGITAL OUT OF HOME" - YES PLEASE

I think one of the greatest advances we've had in tech over the past couple of years is the Internet of Things.

MORE DEFINITIONS

[NOTE TO SELF: SPEED READ THROUGH BORING DEFINITIONS]

The Internet of Things (IoT): is a system of interconnected computing devices with unique identifiers and the ability to transfer data over a network without requiring human-to-human or human-to-computer interaction.

"DIGITAL OUT OF HOME" - YES PLEASE!!!

Digital out of home advertising is advertising that reaches users outside their homes.

Traditionally, this implies advertising on digital billboards, electronic signage, transit advertising, etc etc.

What makes this type of advertising exciting?

Let’s look at a couple of pictures first before we get into that.

P.S: I personally prefer this kind of advertising because it takes the “cookies” off my computer.


From just £6 one-way*
WHY WAIT UNTIL YOU ARRIVE TO FEEL AWESOME?
Mainline Nice&fast
Chiltern Railways
If you think our way, travel our way.
dutyfree
More than you imagined...
EYE LOVE
1-20


OPENING JANUARY 20TH
00 : 11 : 37
DAYS HRS MINS
DUNKIN' DONUTS
LAMAR

EXHIBIT 1: WOMEN'S AID AD CAMPAIGN

In 2015, A U.K. advocacy group created an ad campaign using a digital billboard that shows a woman with cuts and bruises across her face — as passersby look at the woman, the wounds on her face start to heal and disappear.

The billboard, which was produced by ad agency WCRS and media company Ocean Outdoor, uses facial recognition technology to sense how many people are looking its way.

[Play Video 1]

Question

How cool was that?

Before you pass any judgements, let’s watch the next ad

EXHIBIT 2

2014...

Apotek, a pharmacy brand, outfitted subway platform ads in Stockholm with ultrasonic sensors that discerned when a train was coming.

The ad featured a model with a lush mane, and when the train came, her hair flapped in the wind and she struggled to keep it in place.

[Play Video 2]

ABOUT ENGAGEMENT

Can you think of **ads or campaigns** that have over the years caught your attention and think about **why** they may have caught your attention?

Now that we're all reminiscing, let me take your through the thought process that birthed **adKandi**.

WOW!! This is all cool...

DOOH with the level of technology we just watched can be quite expensive and usually takes a lot of time to build.

What if you wanted to build your own DOOH platform while making it extensible and eventually able to perform voodoo?

What would you need to make that possible?


ONE DOES NOT SIMPLY
BUILD A DOOH PLATFORM
imgflip.com

We want nice things

Once you get a glimpse at the potential of DOOH systems. You'll likely want to build one.

There are three core components needed in this context.

1. An extensible OS platform to run your DOOH system on
2. Connectivity support
3. Support for media in varying formats

We can have nice things

Luckily for us the Android platform....

- Is extensible (Supports Mobile vision duh!!) and quite pervasive.
- Has support for connectivity (Bluetooth, Wifi, Ethernet, Sim cards)
- Can play media in various formats

adKandi runs on high tablets, TVs and anything that runs Android.


YES! YES!!
I'M SO "ENTROPICAL"
imgflip.com

We have a platform now...

Let’s define a problem statement.

I think this should work...

“Build a DOOH platform with configurable advert formats and super awesomeeeeeeeee”

Let’s start small ehn, no vess. Skynet wasn’t built in a day.

Our DOOH System aka adkandi

Using the defined problem statement, we built our first MVP.

adKandi could...

- Display ads with different formats, on different android devices
- Display ads on TV setups at different locations
- Display ads based on a schedule
- Provide an amount of user feedback on adverts

But this didn’t exactly cut it.

Adverts are nice but feedback is platinum

Providing feedback led us to ask more questions like

1. What good ways do we measure conversion and engagement rate of a particular advert?
2. How can we categorize these feedback and make it useful to advertising agencies?
3. How do we make the process such that we leverage on the devices themselves to process this information?

Devices are getting more powerful by the day so edge computing for the win.

Solving the feedback problem

We figured out the best way to go about this would be to use the Vision API in

- Estimating the number of people who see these ads.
- Perform mood(sentiment analysis) of each individual watching the ads
- Measure viewer stay time and mood over the duration of watching the advert

ADKANDI II

Questions like

- How do we measure how long a face stares at an advert?
- How do we infer emotions these ads trigger with the viewers?
- What demographic does the viewer fall into?

The working theory here was, while not all encompassing, the Vision API should be able to pick up faces of people definitely looking at the screen.

What we did

Using Mobile Vision face detection APIs, we were able to

- Track faces watching an ad, we called them “eye-ball” count.
- Provide QR code access to exclusive content provided by the adverts.

This way we could give advertising agencies fairly accurate estimates of people interacted at their ads.

One of many problems solved.

Now, we needed to be able to tell how those faces reacted while watching the ads to help us answer our initial questions concerning feedback....

Pertinent questions

- Did they like or dislike your ads?
- Did they smile or frown while watching it?
- Are they likely to buy your product based on inference from their response?

Important and vital information to advertising agencies.

Obvious Next Steps

We needed a way to measure viewer sentiments while watching adverts.

Using M.Vision face tracking APIs, we were able to measure viewer sentiment over the duration which an advert played.

How?

The mobile vision API lets you track individual faces.

With it, we were able to...

- Detect unique faces (Eyeball count)
- Detect unique face landmarks
- Classify sentiment (Estimate if the viewer is smiling or not)

All over the duration of an advert being displayed.

Statistics will save you

Now, we have enough data to crunch, make inference from, make suggestions and a bunch of other interesting things.

All of these without compromising your privacy, serving you cookies you don’t want or shoving banners in your face.

Advertising agencies can have their badly needed data.

Internet of things

By leveraging on user devices and the connectivity advantage of the adKandi platform we were able to explore the IOT domain.

Let’s say you walk into a store and have any of our partner apps installed, you can get pushed vouchers, coupons, trivias, discount cards etc.

Things Considered

First, security was very important. We had to make sure that data collected was secure and the system wasn’t easy to compromise.

Things Considered

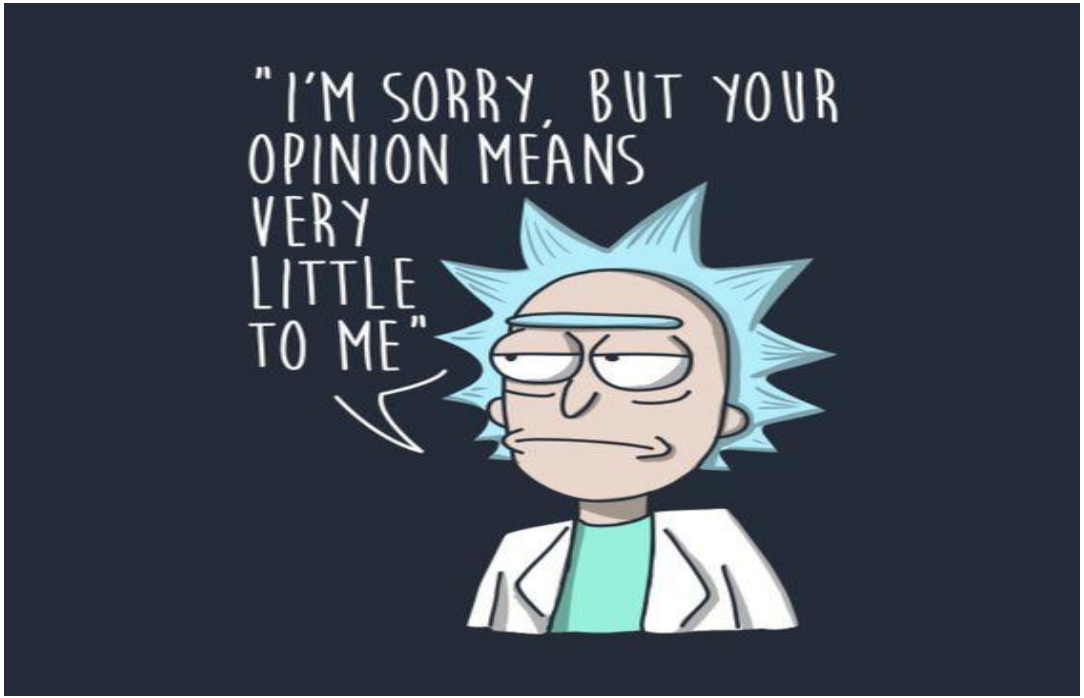
Building for the African context first brought in the need for a lot of redundancy, app auto-restarts, caching, fault tolerance etc.

More goodies

Building adKandi with an OS in mind gave us time to build a much more interactive system using IOT.

This is the point where I go vague as there’s only so much company detail I can share.

BE THIS GUY BUT NOT LIKE THIS



Fun fact

We only measure this when you watch an advert.

Not before or after, only during.

And it’s totally anonymous.

Ethics: don't be evil

Building a solution like adKandi touches on the subject of ethics a lot.

Data was secured, deleted appropriately and we didn't go out of our way to be assholes.

Ethics: because we don't discuss it enough

"That you can doesn't mean you should" - Me

I think it's important to know when to draw the line when software engineering is involved.

With the progress being made in AI and machine learning, we have to make sure we do not end up building a skynet.

I sure as hell don't want to be that guy. :)

EXPLORE THESE ARTICLES

- **Introduction to Mobile Vision APIs:** https://hackernoon.com/machine-learning-for-android-developers-with-the-mobile-vision-api-part-1-face-detection-e7e24a3e472f
- **Mobile Vision Documentation:** https://developers.google.com/vision/android/detect-faces-tutorial

Thank you for listening