

Overall:

“Articulating visual information into speech.”

Sponsors: Google, ABN AMRO, ImpactCity, The Hague

October 27 Webinar (2 years of the glasses)

https://us02web.zoom.us/webinar/register/WN_hbK-UWTFsveWYxdwlaN5Ig

Will collaborate with business to provide their tech and their platform

Based on my testing of the app, it seems like their strongest technology is definitely the text detection. (Was really good at recognizing different sections of things like a website). Also three personalized features (find people, find objects, teach envision faces). The object detection works pretty well for common objects (provides a similar spatialized sense as IMAGE). Scene detection isn't too strong I think, but it is also able to detect colors, it is less specific than IMAGE on average.

Products:

Envision Glasses: (Regular price \$3,476.00 CAD)

- “speak out text and environmental information, recognizes faces, light, and colors, and lets you share that information”
- Text recognition in 60+ languages
- Has Batch Scan (multiple pages), Smart Guidance and Layout Detection (heading, columns, etc)
- Built on Google Glass Enterprise Edition 2
- Detects color, light, languages, etc
- Integrated with other apps (Cash Reader, AIRA, etc)

Envision App:

- Free smart phone app, articulates visuals into text. Object and text detection (found a typo on their website here)
- “Our results constantly out perform benchmarks of visual recognition.”
- Read
 - Takes photos, and scans text, or upload files of text (pdf etc) and it will read it based on sections
- Identify
 - Scene identifier
 - Detect colors
 - Scan barcode
- Find
 - People and objects (common objects from a list). Phone vibrates when located
 - Teach envision faces

Object detection and scene detection are closest to IMAGE

Examples: (first bullet is output from Envision's scene detection, second is IMAGE)

Testing on the data set,

- 0007 (Treadmill)



- Probably a close up of a cage
 - This photo contains the following outlines of regions: wall, floor, stage, windowpane, and curtain. It also contains the following objects or people: a person.
- 0010 (ocean and sky photo)



- Might be a close up of a blue wall
 - This photo contains the following outlines of regions: sky, and sea.
- 0018 (bowl of fruit on paper)



- Looks like a bowl of fruit sitting on a table
- This photo contains the following outlines of regions: floor, wall, water, seat, and rock. It also contains the following objects or people: a dining table, and a bowl.

- 0023 (dog and lady in bed)



- It might be a cat lying on a bed
 - This photo contains the following outlines of regions: seat, wall, sofa, and rock. It also contains the following objects or people: 2 dogs, a bed, and a person.
- 0025 (city scene)



- Probably a blurry photo of a city
- This photo contains the following outlines of regions: sky, floor, building, skyscraper, and earth. It also contains the following objects or people: 5 cars.

Envision Ally:

- Can video call friends, family through the glasses

AI Used:

Have found very little information about the technical side. Found the reddit of some company executives but no GitHubs.

“Envision’s software provides the fastest and most accurate OCR (Optical Character Recognition) available”

“Working with Google Glass’ technical specifications and build quality has been ideal for our development needs,”

<https://www.mdpi.com/2079-9292/10/22/2756/htm>