

Additional Visual Recognition technologies and extended considerations

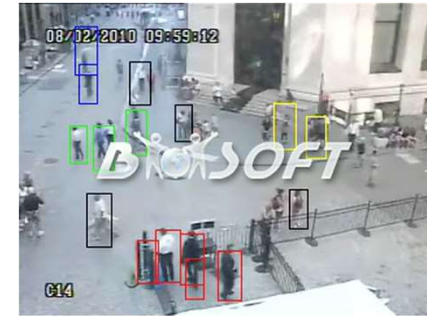


Visual Recognition : Industry Solutions

Intelligent Video Analytics

Challenges

- Inability to monitor large infrastructure
- Limited staffing and budget for security operations
- Silo solutions, infrastructure difficult to leverage



Capabilities

- Metadata enrichment of video
- Proactive searches, pattern identification, pre-configured conditions
- Rich search capabilities
- Critical area monitoring
- Geo fencing, Face matching



Use cases

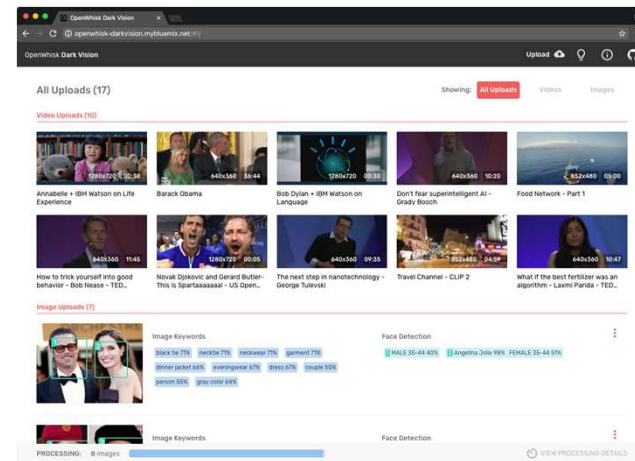
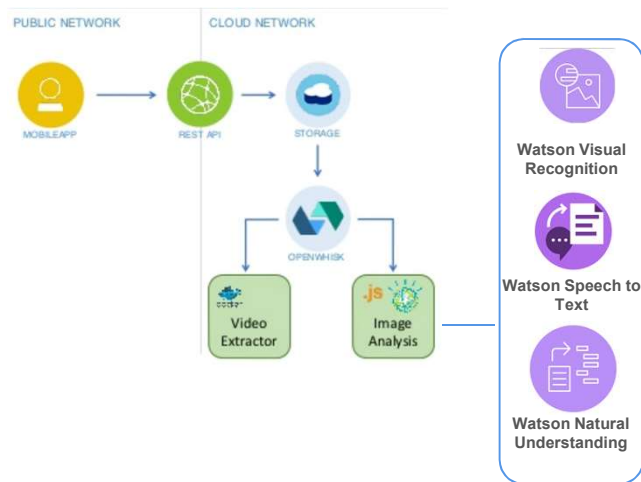
- Public Safety in critical infrastructures
- Public Transportation
- City Intelligent Operations Center



Dark Vision

Capabilities

- Metadata enrichment of video (images and sound)
- Improve search results
- Better navigation, make recommendations
- Detect brands



<https://github.com/IBM-Cloud/openwhisk-darkvisionapp>

Multimedia Enrichment Pipeline

IBM Watson Accelerator

ABOUT

Videos

Loaded Media

GridBreakers_Ep_01-Rapid_...

GridBreakers_Ep_02-Terror_...

GridBreakers_Ep_03-Saved_...

GridBreakers_Ep_04-Defying...

GridBreakers_Ep_06-Buried_i...

Watson Multimedia

Scene 00:02:33 - Moment 160

Show Moment

we're talking and joking around and

Classes[1] Faces[0] Words[]

gray color[95%]
nature[93%]
ice mass[89%]
ash grey color[83%]
glacier[77%]
/nature/ice mass/glacier
icecap[66%]
/nature/ice mass/icecap

02:41 05:09

00:00:01 00:00:23 00:00:59 00:01:21 00:02:00 00:02:33 00:03:04 00:03:37 00:03:59 00:04:34 00:05:01 00:05:33 00:06:01 00:06:17

Transcript

you know nobody around . I . we're talking and joking around and . happened . a ...

55 more words...

you know nobody around . I . we're talking and joking around and . happened . a quick . nicekino . my feet . tip of the richest . broke . fast free fell naked . all I could think about was gonna land water at that temperature you know five minutes . don

Categories

/food and drink[15%]
/art and entertainment/movies and tv/movies[12%]
/art and entertainment/music/singing[12%]



Visual Recognition : Product usages

Watson Content Hub

IBM Watson Content Hub (“WCH”) is a cloud-based content management solutions. It is a headless CMS that powers digital experiences. Watson services help transform disorderly assets into an easily searchable library, so everyone can always find just what they need.

- **Content Management System**
 - Manage unstructured contents
 - Image, Video, Documents
 - Manage structured contents
 - Define content template
 - Article, news, product, ...
- **Includes Watson Visual Recognition**
 - Automatically tag images
 - Use default classifier
 - Custom tagging in plan
- **Watson Tone Analyser**
 - Used for text and documents (pdf, word...)
- **Website Management**
 - Create and manage web pages and web site
 - Template based creation
- **Assisted content creation for Business Users**
 - Step by step content and website wizard
- **SaaS Solution**
 - Watson Visual Recognition access included in the license at no extra charge
 - Only for use in WCH

Watson Content Hub : <https://www.digitalexperience.ibm.com>

IBM Watson Content Hub

Hub setup FRANCK DESCOLLONGES Help

Content > All content and assets > Upload assets

My items

Watson analysis completed.
3 assets were uploaded

Upload assets Done

> Applies to all digital assets Tags (0) Add tag

Digital assets 3

- lab3_1.jpeg
- lab3_2.png
- lab3_3.jpeg

lab3_1.jpeg Open asset

Tags

Tags (0) Add tag

Watson tags (16) Reject all

Concepts

- athletic game
- sport
- ball carrier (football)
- soccer (football) player
- player
- contestant
- person
- professional football
- contact sport
- outdoor game
- basketball
- claret red color

Faces

- Male
- 25-28
- 26-29
- 23-26

IBM Watson Media

IBM Watson Media solutions enable you to infuse AI throughout your media workflow or video library - unearthing opportunities to improve monetization, viewer engagement, content performance, ad revenues, and more.

- **Watson Captioning**

- <https://www.ibm.com/watson/media/watson-captioning/>
- After training Watson, caption accuracy rate is up to 95%
- Custom corpora and custom vocabulary to further enhance the accuracy of first-run caption scripts
 - Language model customization creates domain-specific language models to increase recognition accuracy
- Leverages automated speech-recognition capabilities (using the Watson Speech to Text API) to ingest spoken and audial elements of video assets
 - Applies a range of cognitive functionality to assess and act on the interpreted data

- **Watson Media Enrichment**

- <https://www.ibm.com/watson/media/watson-video-enrichment/>
- Cognitive abilities to analyze audio, textual, and visual data within multimedia content
 - Watson analyzes multimedia content and builds deep, easily searchable metadata packages for every asset in your library
 - Watson can even determine the dominant emotion of a video like joy, sadness, or anger
- Target viewers based on what really motivates them
 - Recommend only highly relevant content personalized to each user's preferences, resulting in more viewing time and reduced churn

IBM Watson Media

IBM Watson Media solutions enable you to infuse AI throughout your media workflow or video library - unearthing opportunities to improve monetization, viewer engagement, content performance, ad revenues, and more.



Content search and discovery

Content enriched by Watson is inherently more searchable - and improved content discovery leads to increased usage.



Recommendation uplift

Enriched metadata gives you deeper insights into what really motivates viewers, so you can connect them with relevant content in completely new ways.



Highlight clipping

Sports broadcasters need to create video highlights as quickly as possible. Watson automatically watches, identifies, and clips those must-see moments.



Closed captioning

Watson automates real-time, accurate, and easily editable captions, saving you time and money.



Compliance monitoring

Watson understands when to flag video assets for potential violations, helping you stay compliant while eliminating the cost of manual search and identification.

IBM Cloud Video – Formerly known as UStream

When streaming video means business : Power your business with proven, scalable cloud video solutions and reveal insights hidden in your video data

<https://www.ibm.com/cloud-computing/solutions/video/>

- **IBM Cloud Video**
 - **Fully managed streaming solution on the cloud**
 - Private and public channels
 - Live broadcast and recorded video
 - **Watson Generated Captions for Video**
 - Closed captions have grown to be an important part of the video experience.
 - make video more accessible to a diverse viewer base like the deaf or hard of hearing or non-native speakers
 - IBM Watson Speech to Text technology to extract searchable contents from audio out-of-the-box



Corporate communications

Transform employee communications by producing secure, internal video assets that provide individual-viewer tracking.

→ [Learn more](#)



Live events

Put the potential of live content to work at a global scale with time-tested tools optimized for streaming video content online and integrating social media.

→ [Learn more](#)



Marketing

Introduce campaigns, generate leads, launch products, train users and share news and information with your prospects, customers and the public, while leveraging the power of immersive, persuasive video.



Visual Recognition : On premise technical preview

Power AI Vision

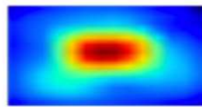
On-premise Visual Recognition Technology Preview

Technology Preview Capabilities

IBM PowerAI Vision's Technology Preview Release 3.0 currently includes three main use cases to demonstrate its capabilities:

Static Image Classification: Determine whether an image belongs to one or more classes of images based on overall image contents (e.g. "Determine the species of dog in this image").

Result:

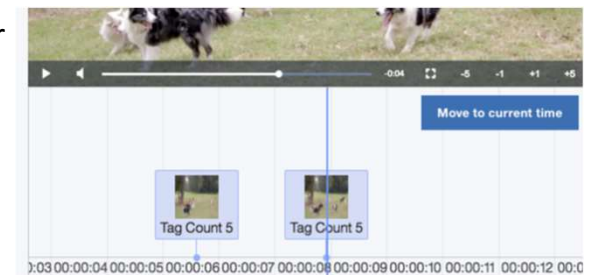


japanese_spaniel: 1.00000



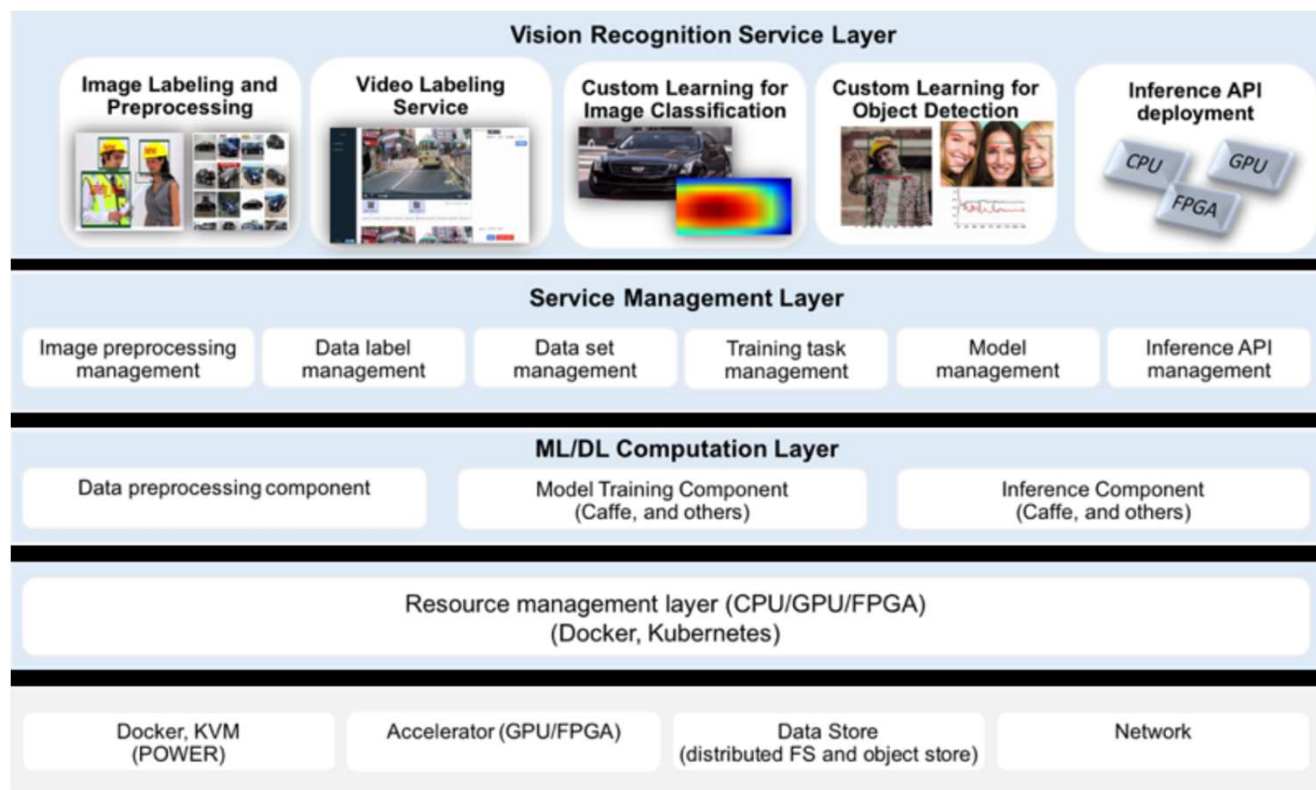
Static Image Object Detection: Determine and label the contents of an image based on userdefined data labels (e.g. "Locate and label all dogs in this image").

Video Object Detection: Determine and label the contents of an uploaded video or live video stream based on user-defined data labels (e.g. "Locate and label all dogs in this video").



Power AI Vision

On-premise Visual Recognition Technology Preview





Visual Recognition : Competitive landscape

Competitive Landscape - Major Players

x = GA

	IBM Watson	Microsoft	Clarifai	Google	Amazon
General Model	x	x	x	x	x
Custom Model	x	Beta	x	x (via AutoML)	
Food Model	Public Beta		x		
Multiple Object Detection	Research		x		
Face Detection	x	x	Beta	x	x
Face Recognition	Research	x			x
Face Sentiment/Emotion	Research	Beta		x	x
Logo Model	Awaiting prioritization		Beta	x	
Text Model	Private Beta	x		x	x
Video Recognition	Watson Media now owns	x			x
Fashion Model	Research		Beta		
Embeddings	Private Beta Q1 2018		Beta		
Celebrity Detection	Refactoring	x	Beta	x	x
Visual Search/Similarity Search	Refactoring		x	x (no custom libraries)	x
Explicit Model	Public Beta	x	x	x	x
Color Detection	Public Beta	x	x	x	
Landmark Detection		x		x	
Handwritten Text		Beta			
Deployment Options	Edge via CoreML Q1 On Prem via PowerAI	Core ML, On prem via Azure Stack	Edge beta (iOS today, Android/IoT coming soon)	Core ML via TensorFlow, Hybrid-Ready (Kubernetes)	Edge via DeepLens
Tooling	Beta (Q1 GA)	x			
Languages	English, Arabic, German, Spanish, Italian, Japanese, Korean	English for most features, 25 languages for OCR (full list)	23 languages (full list)	56 languages for OCR (full list)	English
Other		Document OCR	Travel, Wedding, Focus, Landscape Quality, Portrait Quality, Textures & Patterns	Document OCR	Face ID