

What is Visual Recognition?

Video and Image data is Growing exponentially

\$105B

2019 cloud-based video market¹

2.5 Quin-tillion

bytes (10^{18}) of data is created everyday.

1.3 Trillion

photos were taken in 2017.⁵

Video will represent **80%** of global Internet Traffic by 2019³

Social & Video



150 million daily active users
301 million monthly active users

Video is increasingly a communal experience, with social media connecting audiences viewing live and on-demand content.⁴

\$80B expected
image recognition market size by 2025⁶

MINUTE OF INTERNET

266K

Hours watched



4.3M

Videos viewed



2.4M

Snaps created



176K

Scrolling Instagram



MEDIA DATA

97% growth in 2017

82% unstructured

Media data comes from sources such as:



Video and Film



Images



Audio

Market Challenges

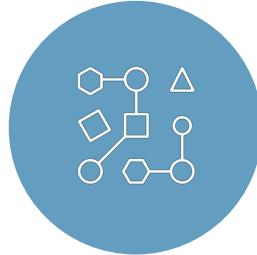
In 2017, more than **1.3 trillion photos** were taken, with 4.7 trillion photos supposedly to be stored.



Analysis of visual data requires advanced machine learning and computer vision algorithms that require significant training.



Creating custom machine learning systems can require specialized and detailed knowledge and have a significant learning curve for new users.



Different Flavors of Computer Vision

Image recognition

- Image recognition finds images shared online/in a platform's archive containing certain things – usually brand logos.

Image analysis

- Image analysis identifies what's within an image you already have and subsequently tags what's in the image, sometimes even “captions” an image.



Different Flavors of Computer Vision

Image Tagging

Similarity Search

People Counting

OCR

Visual Learning

Geo fencing

Color Detection

Face Detection

Abandoned bag

Content Classification

Facial Recognition

License Plate
recognition

Text from Images

Facial Sentiment

Object Tracking

Media Enrichment

Image Captioning

Group forming

Use Computer Vision to answer these questions

- What is this a picture/video/image of?
- Which ones and how many pictures/videos/images do I have which are similar to this one?
- What is the relevance to my bottom line of using one image vs. another image?
- What are the common themes in this batch of 1MM photos?
- Is this ship a known entity?
- What is my audience interested in?
- What user-generated content matches which products I sell?
- Which images will resonate best in a marketing campaign?
- Which image should I feature on my front page of my app?

Use Cases & Case Studies

Visual Recognition Use Cases

Visual Inspections

An **Insurance** company builds an image recognition solution to automate visual inspections for damage, defects, and quality assurance.

Resource Identification

A **Mining & Minerals** company uses image recognition to automatically identify assets and sites in satellite imagery.

Autonomous Vehicles

A **Car** equipped with computer vision to safely drive in our environment, capable of reading road signs, detecting pedestrians, other cars...

Social Media Listening

An **Advertising** agency analyzes visual content in social media posts to understand content, sentiment, and trends.

Content Enrichment

A **Media** company uses image recognition to automatically append metadata to visual content, turning dark data into searchable content.

Assisted Maintenance

An **Energy Producer** uses mobile application to empower technicians on the field to rapidly identify, report and solve maintenance issues with rapid access to contextual documentation and maintenance guidance.

Demographics

A **Retailer** uses face detection capabilities to gather age and gender estimates of its shoppers.

Calorie Counting

A **Health & Nutrition** company automatically identifies foods and meals in order to more quickly return nutrition facts.

Visual Assistance

A **Car Manufacturer** uses augmented reality and computer vision to educate new car owners on the devices and capabilities of their new car.

Fraud Detection

An **Insurance** company authenticate claims using image forensic search.

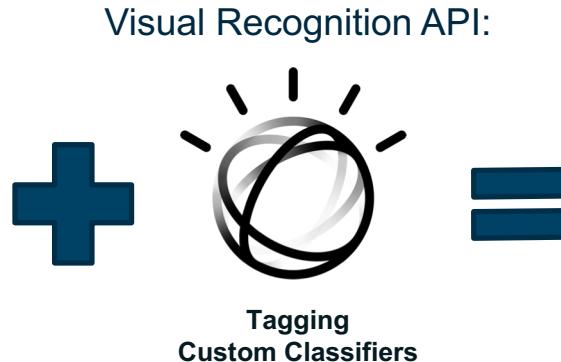
IBM Watson Advertising - Social Media Monitoring



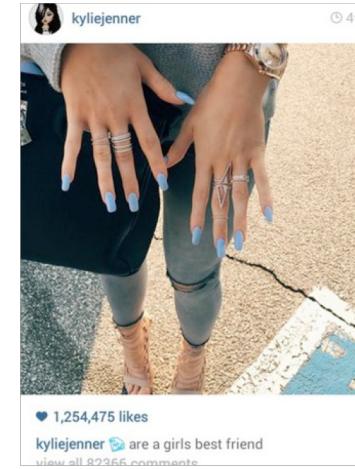
Companies are looking to understand what is trending on social media and how they can use that for advertising their own products:



photos often have more information than the associated captions



Can use trend insights in advertising campaigns to better target certain customer segments:



this photo with over a million likes shows a celebrity wearing skinny jeans

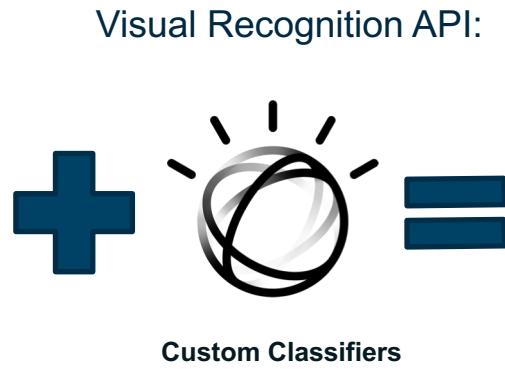
monitor the top trends on social media

get to know your customers' **interests**

IBM Watson Natural Resources - Resource Monitoring



A mining or oil company has a number of facilities it sends people to inspect on a regular basis:



must perform regular inspections of oil wells for leaks/spills to comply with regulations

Identify through satellite/security images when a leak has occurred at a facility:



instead of sending people out periodically, a resource manager can know in real time when a leak has occurred

quickly evaluate current state of facilities

improve margins by **reducing** resource **waste**

IBM Watson Natural Resources - Quality Control

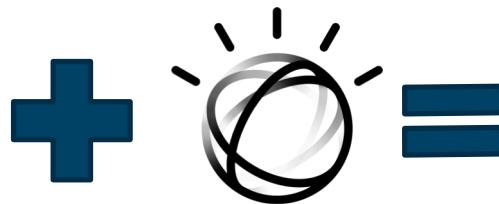


Equipment and infrastructure monitoring requires manual, time-intensive inspections:



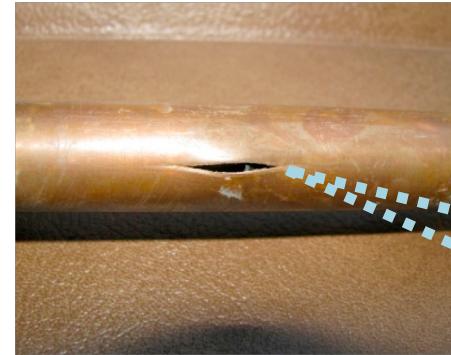
Lack of automation means monitoring takes time away from other tasks

Visual Recognition API:



Tagging
Custom Classifiers

Automatically inspect infrastructure and assets to identify areas that need maintenance:



this includes the ability to custom train on specific infrastructure

preventative measures help decision makers
address issues **before** they become problems

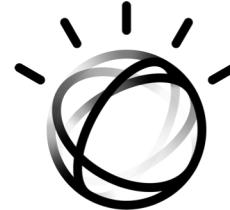
IBM Watson Insurance - Report Organization



A survey of 1,000 homes in the bay area has been ordered with ~20 images/property:



Visual Recognition API:



Custom Classifier



House
#145
House
#267
House
#325
etc...

list of houses with images that show roof damage

around 20,000 images not clearly labeled or organized

focus on the parts that matter

find what you need **faster**

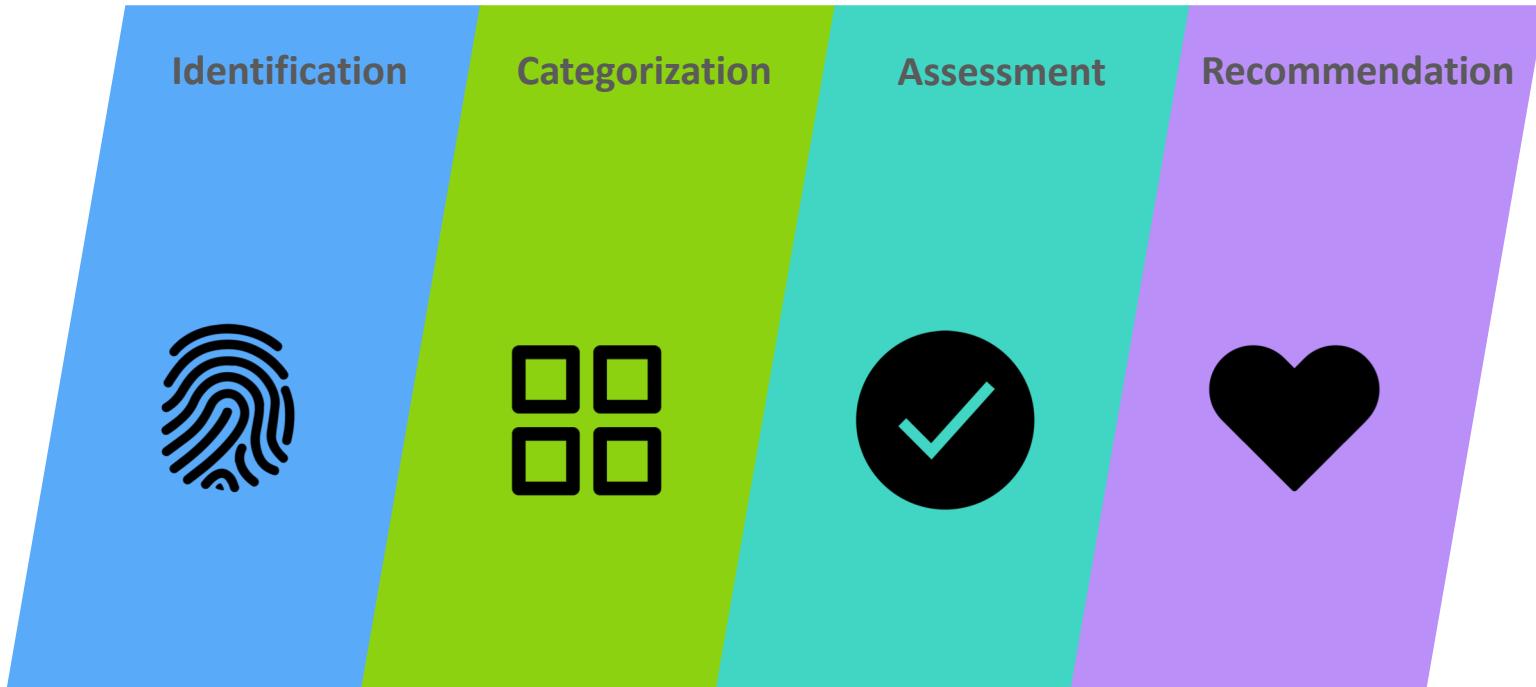
Watson Visual Recognition

Watson Visual Recognition



An **image recognition** service that enables users to quickly and accurately **tag**, **classify**, and **train** **visual content** using machine learning.

Watson Visual Recognition focuses on...



Watson Visual Recognition **identifies** objects and people.

hatchback
compact car
vehicle
claret red color



Watson Visual Recognition **categorizes** for easy organization.



vintage



modern

Watson Visual Recognition assesses for better problem-solving.



counterfeit, 78% confident

Watson Visual Recognition **recommends** for faster decision-making.



Historically, we've paid
\$7,500 for similar
types
of damage

Visual Recognition Features

General Model

Quickly understand the contents, scenes, and actions within an image.

English, Spanish, Arabic, Japanese, Korean, Italian, French, and German are supported today.

Face Model

Locate faces within an image and receive age and gender estimates.

Text Model (private beta)

Locate and read natural scene text within images. [Apply for access.](#)

Custom Model

Train Watson to understand and classify your own custom content.

Food Model (beta)

Recognize foods and meals with enhanced accuracy.

Color Model (beta)

Identify up to two most prominent colors that appear within an image.

Explicit Model (beta)

Determine if an image contains inappropriate content.

Watson Visual Recognition Service

The screenshot shows the Google Cloud Platform interface for Visual Recognition. At the top, there's a navigation bar with a profile icon, the title "Visual Recognition : visual-recognition-wws", and a sub-project name "Associated project : WasteSorter". Below the navigation is a tabs bar with "Overview" (which is active) and "Credentials".

The main content area is divided into several sections:

- Custom**: A section for creating custom visual classifiers. It includes a "Create Model" button.
- General**: A section for generating class keywords from images. It includes a "Copy model ID" button and a "Test" button.
- Faces**: A section for detecting human faces in images. It includes a "Copy model ID" button and a "Test" button.
- Food**: A section for detecting food items. It includes a "Copy model ID" button and a "Test" button.
- Explicit**: A section for detecting explicit content. It includes a "Copy model ID" button and a "Test" button.

Below these sections is a "Text" section labeled "PRIVATE BETA", which automatically detects and extracts recognized words within natural scene images. It includes a "Request Access" button.

At the bottom left, there's a "Custom Models" section with a "Default Custom Model" button.

Out of the box tooling

General

Associated Service : VM4-Workshop_1

Overview Test Implementation

Filter

x Clear results

Threshold 0.01

0 1

Class

ash gray color

auto racing

auto vehicle

bodywork

bonnet (of vehicle)

bumper (of vehicle)

cab

cab (vehicle)

car

charcoal color

compact car

compartment

convertible car

coupe car

dark red color

door panel

elevator car

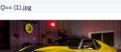
four-wheel drive

furniture

gear cluster

gear mechanism

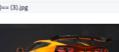
2Q== (1).jpg



2Q== (2).jpg



2Q== (3).jpg



2Q== (4).jpg



claret red color	0.50	ash gray color	0.44	wheeled vehicle	0.51
lavender yellow color	0.48	charcoal color	0.72	vehicle	0.42
vehicle	0.84	vehicular	0.72	motorcycle	0.92
wheeled vehicle	0.84	wheeled vehicle	0.72	racing	0.82
car	0.83	car	0.72	sport	0.82
sports car	0.66	motorcycle	0.40	indian red color	0.81
body (of vehicle)	0.61	sport utility	0.55	sports car	0.81
cab (taxi)	0.59	off-road vehicle	0.55	motorcycle	0.71
compartment	0.58	transmission system	0.55	motorcycle	0.71
coupe	0.58	gear mechanism	0.55	auto racing	0.79
dark red color	0.58	coupe car	0.53	race car	0.79
door panel	0.54	bodywork	0.51	roaster	0.56
elevator car	0.52	body (of vehicle)	0.50	motor vehicle	0.53
four-wheel drive				gas guzzler	0.53
furniture					
gear cluster					
gear mechanism					

☰ See all features

1/2

Pre-trained models

Watson Visual Recognition's category-specific models enable you to analyze images for scenes, objects, faces, colors, foods, and other content.

Paste an image URL here...



General Model

Quickly understand objects, actions, scenes, and colors within an image.

photographic film	0.98
photographic equipment	0.98
roll	0.94
bottle green color	0.89
electrical device	0.79
greenishness color	0.57
roll film	0.50

Face Model (no results)

Food Model (beta) +

Explicit Model (beta) +

Text Model (private beta) -

Extract text from natural scene images. To learn more, please request access at ibm.biz/request-text

← →

Visual Recognition Tools

The screenshot shows the IBM Watson Visual Recognition service interface. At the top, there's a navigation bar with 'IBM Watson' and 'Projects' selected. Below it, a breadcrumb trail shows 'Services / Watson Services / Visual Recognition-ps'. The main content area has a title 'Visual Recognition: Visual Recognition-ps' and a subtitle 'Associated Project - Food Classifier'. There are two tabs: 'Overview' (selected) and 'Credentials'. The 'Overview' tab displays several cards:

- Custom**: Create custom, unique visual classifiers. Use the service to recognize objects and concepts that are not available with general models.
[Create Model](#)
- General**:
 - Food**: Utilize a specialized vocabulary of over 2000 food-related items, food products, and dishes with enhanced accuracy.
[Test](#)
- Explicit**: Assess whether an image contains objectionable or adult content, helping you generate audiences.
[Test](#)

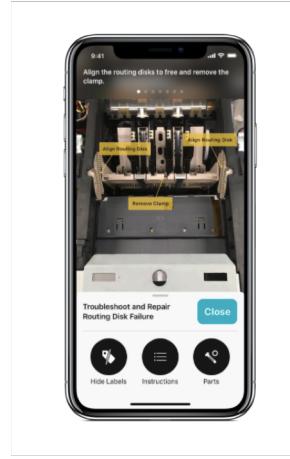
At the bottom left, there's a 'Text' section with 'Text' and 'Create model ID' buttons. On the far right, there's a blue circular icon with a white question mark.

Watson Studio

Easily provision a Visual Recognition instance and interact with the service through your browser.

Classify images using IBM's pre-trained models or train and manage your own Custom Models.

Try it now



IBM Watson Services for Core ML

Build apps that leverage Watson models on iOS devices, even when offline.

Export Custom Models for Core ML through Watson Studio. Core ML exports are free during our promotional period.

Get started

Key Capabilities of Watson Visual Recognition



Tooling

Not a developer? No problem! Our tooling allows anyone to get started immediately using your own API key.



Flexibility and Customizability

Whether you're looking to get started right away with IBM's models or custom train your own, we offer options for every user.



Privacy and Security

IBM ensures that images and data passed through our services remain secure. Your data is never used to train IBM models.



Simplified query language

Our API service and best practices guides make it easy to get started with your preferred language.

Watson Recognition Pricing Model

PLAN	FEATURES	PRICING
✓ Lite	1,000 Events per month towards: Pre-trained model classification (General, Face, Food, Explicit) (images) Custom Model classification (images) Custom Model training (images) 2 Custom Models 1 Lite Plan instance per IBM Cloud Organization Free Exports to Core ML	Free
<p>The Lite Plan gets you started with 1,000 events (images) per month and the ability to train two Custom Models. Users wishing to use more premium features or increase usage must upgrade to a Standard Plan or a Subscription Plan.</p> <p>Lite plan services are deleted after 30 days of inactivity.</p>		
Standard	Image Tagging Events Pay per Use Face Detection Events Pay per Use Training Events Pay per Use Custom Tagging Events Pay per Use Food Tagging Events Pay per Use Explicit Tagging Events Pay per Use	€0.001504 EUR/GeneralTagging €0.003009 EUR/FaceRecognition €0.0752 EUR/Training €0.001504 EUR/CustomTagging €0.001504 EUR/FoodTagging €0.001504 EUR/ExplicitTagging

Premium

Watson Premium plans offer a higher level of security and isolation to help customers with sensitive data requirements.

Thank You