Computer Vision

Can you identify the actor in this image?

Yes, he is none other than Tom Cruise who has acted in many movies like the Mission Impossible series (Does he ever age?!).

But how were you able to identify him from the image? And how can a computer identify him? While we identified him using our eyes and brain (and the fact that he doesn't seem to age), the computer uses computer vision to identify celebrity information from the image.

How? You will get the answer in this lesson!

Topic Covered in the Lesson

2. Computer vision and its applications
3. How self-driving car works

4. Artificial Intelligence extension in PictoBlox
5. Computer vision blocks in PictoBlox

Key Learning Outcomes

At the end of the lesson, you will be able to:

1. Use Artificial Intelligence blocks in PictoBlox to identify the following in images:
   0. Brand
   1. Celebrity
   2. Objects
   3. Image Tags
   4. Image Description

The following example shows a project that identifies the object in the image, i.e. a person, who that person is, i.e. a celebrity, his name, what he's wearing, and the brand his clothes belong to.
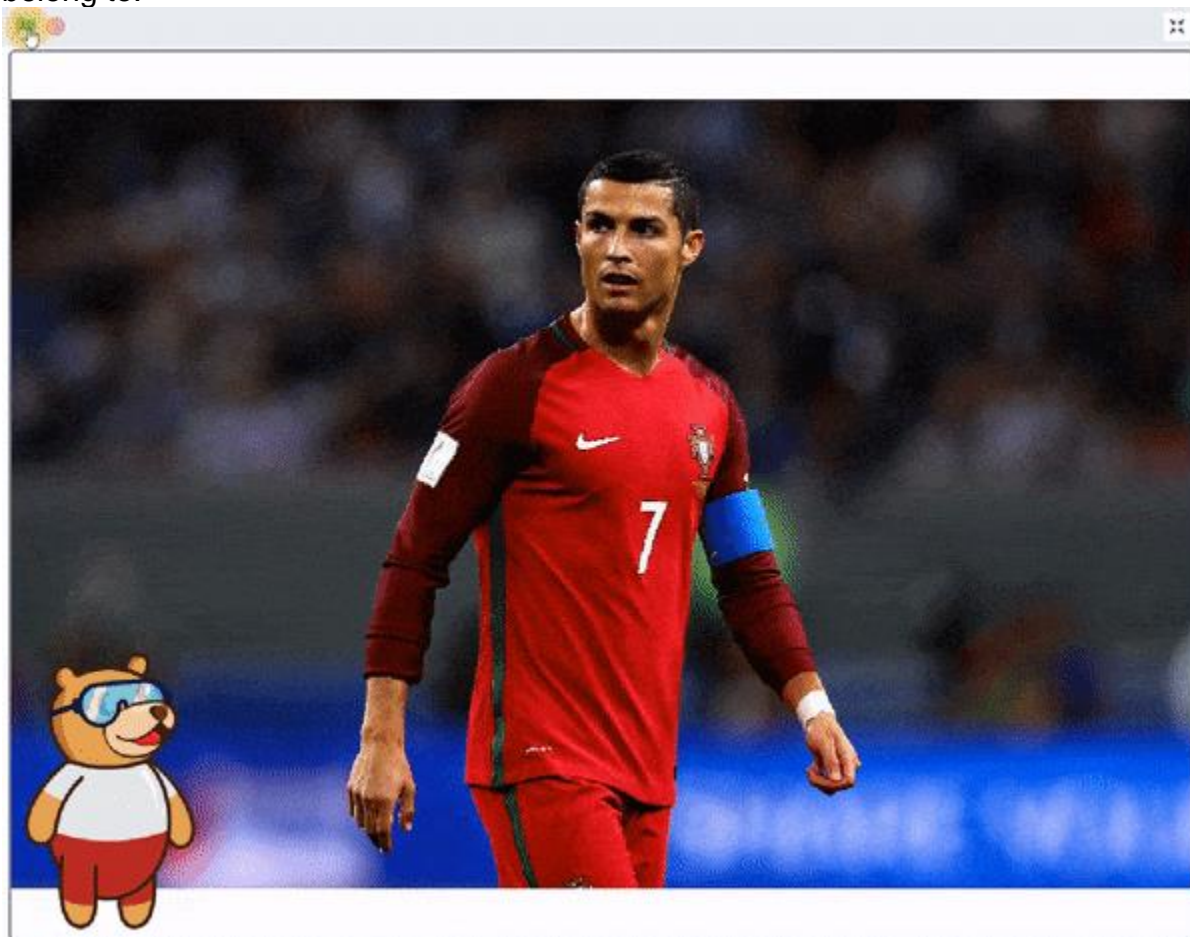
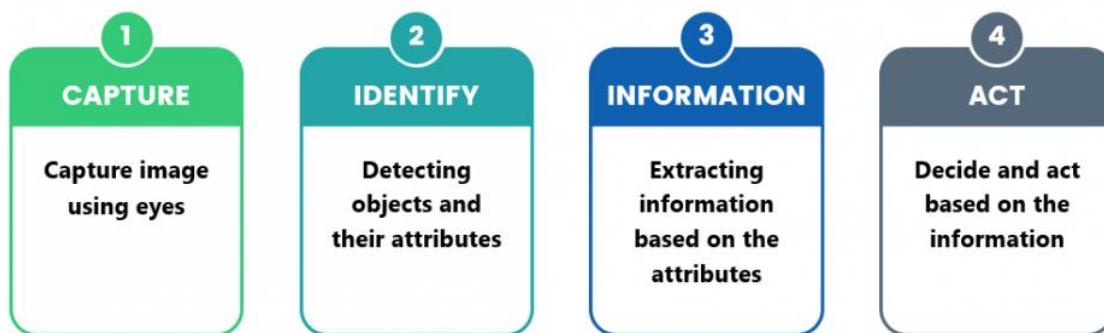1. Make AI projects in PictoBlox using computer vision.

2. Understand how computer vision works.
3. Understand the application of computer vision in various industries.
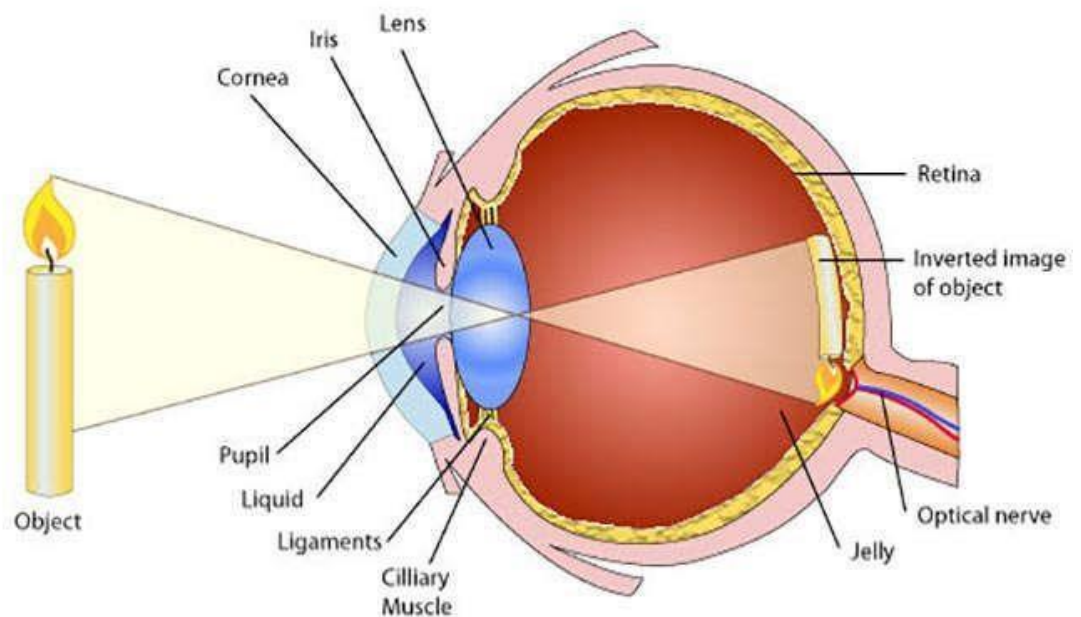
Let's begin!

Introduction to Computer Vision

How Human Vision Works?

To understand computer vision, first, we must look at how human vision works.

1. Capture image: Humans capture images using their eyes. The image captured is

2. Identify the objects and their features: The raw image is then transferred to the brain via the optical nerves for processing. The brain starts to identify different

objects like a candle, human, chair, and many others along with their features such as size, color, shape, and others.

3. Extract information: In this step, our brain compares the features of the object to its past knowledge to gather information. E.g., it can differentiate between your father and your mother because you can distinguish their visual features.

4. Act: Once you get the higher-level information, you can start acting on it. E.g., if you can identify that a ball is coming to your face, then you can move aside to avoid hitting the ball.

All these steps happen at a very fast pace due to the perfection of the human eye and brain coordination.
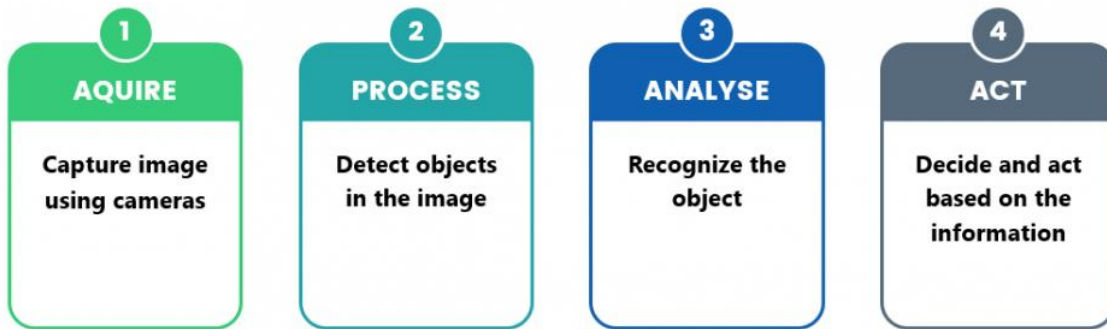
Computer Vision

Computer vision also follows a similar approach as human vision.

Computer vision deals with how computers can be made to gain high-level understanding from digital images or videos. From the perspective of engineering, it seeks to automate tasks that the human vision can do.

Computer vision tasks include methods for acquiring, processing, analyzing, and understanding digital images.

1	2	3	4
AQUIRE	**PROCESS**	**ANALYSE**	**ACT**
Capture image using cameras	Detect objects in the image	Recognize the object	Decide and act based on the information

Example: Self Driving Car Using Vision

A self-driving car is a vehicle that is capable of sensing its environment and moving safely with little or no human input.

For this example, let us consider that a self-driving car can go forward, turn left, right, or stop. Let's see how the car would react if a pedestrian comes in front of it.

1. Acquire: Self-driving cars use cameras to acquire images. They acquire and process images are at a very high rate. Let's consider that our camera has acquired this image:

2. Process: The computer starts to identify all the objects in the image and make a list of the objects with their position. In this case, there is something on the road.

The computer still has no information about what object it is.
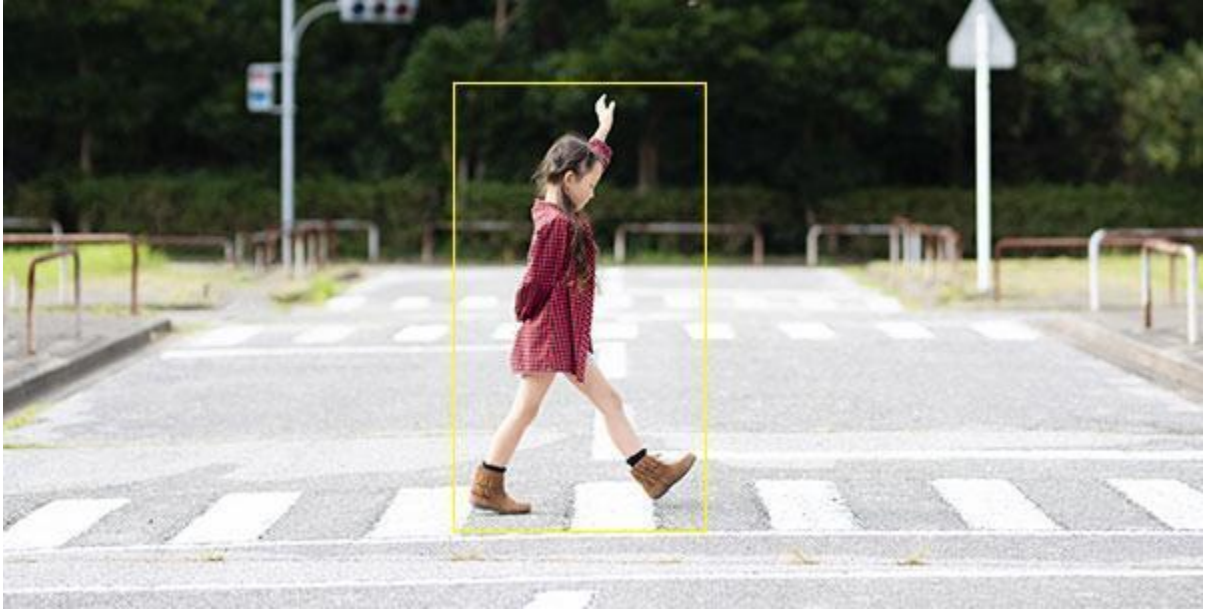
3. Analyze: The computer then classifies each object into different categories. In this case, it identifies the object as a girl. It also tags some information to the object as harmfulness, distance, and other parameters. These tags are the

higher-level information used to make a decision.

4. Act: Based on the higher-level information the computer can act. In this case, the car will stop.

Conclusion

In this topic, you have learned what computer vision is and its process using the example of a self-driving car. In the next topic, you will learn about the Artificial Intelligence extension in PictoBlox, using which you make various AI projects.

Artificial Intelligence Extension in PictoBlox

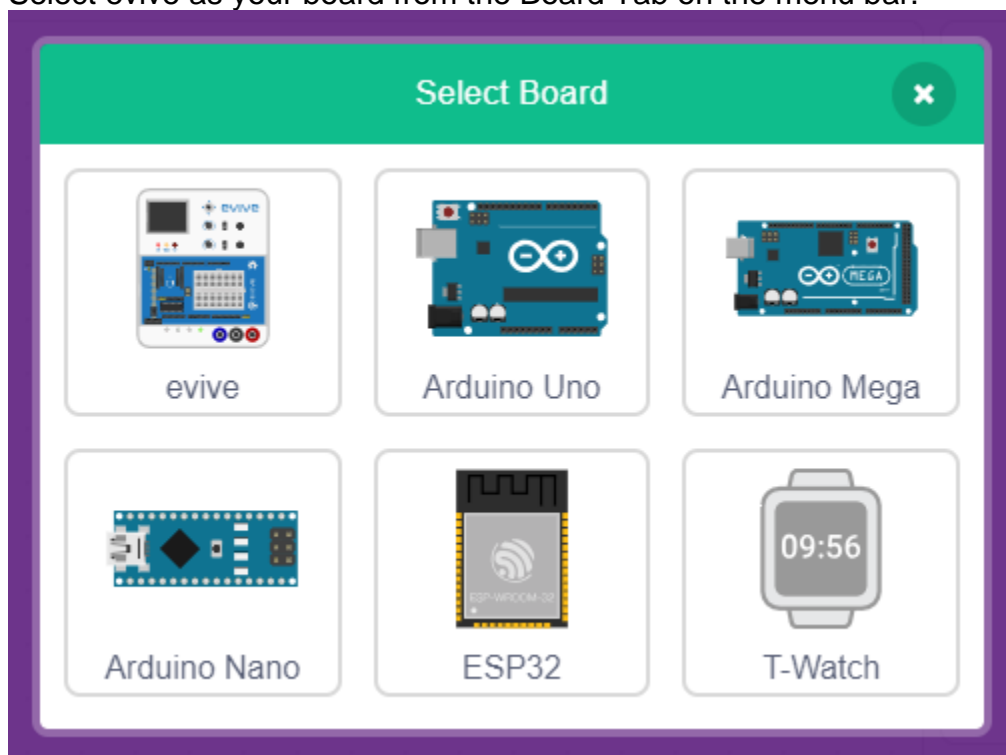## Artificial Intelligence Extension

PictoBlox's *Artificial Intelligence* extension allows you to make AI projects in PictoBlox. You can make projects in the following areas of AI:

2. Face Recognition: To identify age, gender, and emotion from the face recognized in the images.
3. Optical Character Recognition: To identify text in images.
4. Speech Recognition: To generate text from speech.

Loading the AI Extension

Follow these steps to load the AI extension in PictoBlox:

1. Open a new project in PictoBlox.
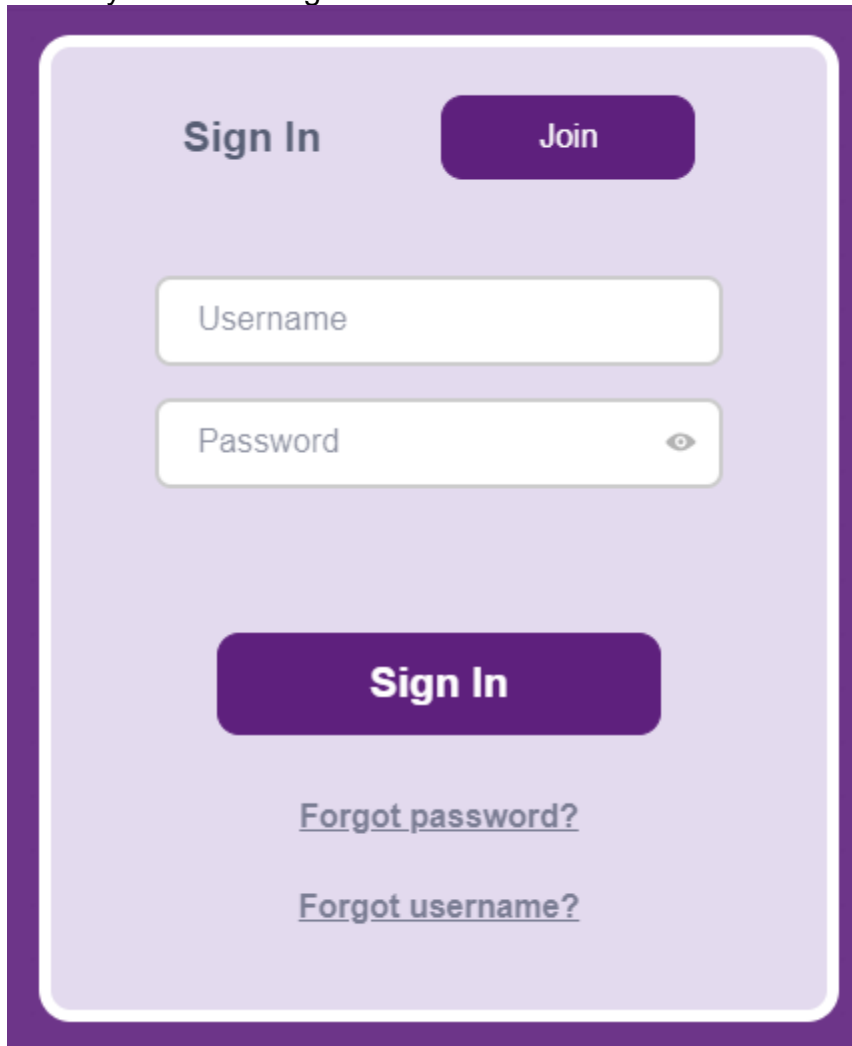2. Select evive as your board from the Board Tab on the menu bar.

3. Click the Add Extension button in the bottom left corner.

4. A window will open will all the available extensions. Select the *Artificial Intelligence* extension from the library.
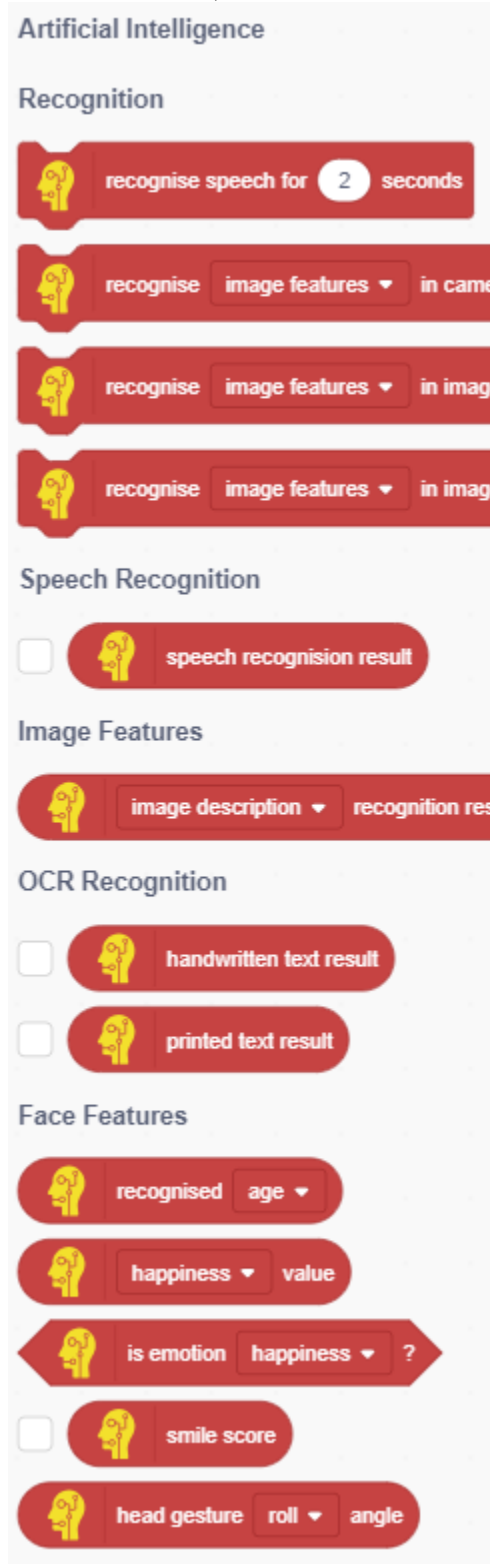

Artificial Intelligence
Speech, Text & Face Recognition

5. Once you have selected the AI extension, a sign in tab will open. To use the AI blocks you have to sign in PictoBlox.

6.

7. Alert: The login for PictoBlox is different from the website. So, you have to create a new account if you have not signed in PictoBlox before.

8. Once selected, all the blocks of the extension will load in the extension palette.

Artificial Intelligence

Recognition

recognise speech for (2) seconds

recognise image features ▼ in came

recognise image features ▼ in imag

recognise image features ▼ in imag

Speech Recognition

speech recognision result

Image Features

image description ▼ recognition res

OCR Recognition

handwritten text result

printed text result

Face Features

recognised age ▼

happiness ▼ value

is emotion happiness ▼ ?

smile score

head gesture roll ▼ angle

In the next topic, we will learn about the computer vision blocks in detail.

Computer Vision Blocks

In this topic, we will have a look at what blocks PctoBlox has in store for us for computer vision.
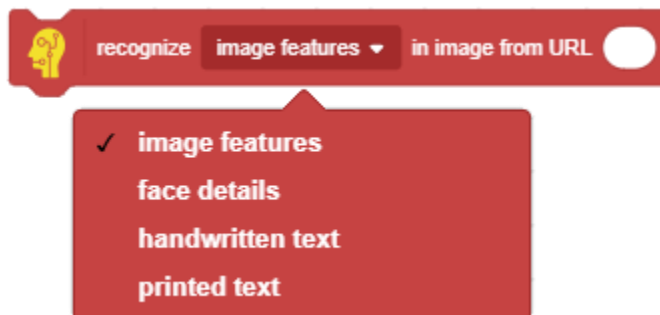
Computer Vision Blocks

The *Artificial Intelligence* extension has the following 3 blocks which take an image as the input and analyze for image features:

1. recognize () in camera image after () seconds: This block starts the camera and takes an image after the specified time and analyzes it. It then saves the image features in PictoBlox.
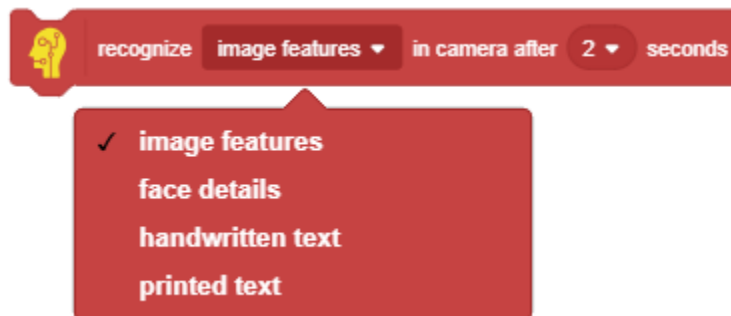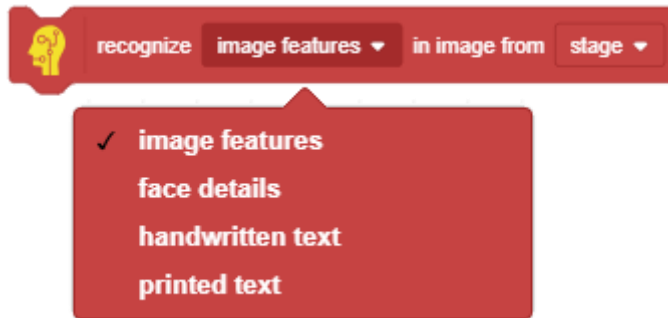

2. recognize () in image from URL (): This block extracts the image from the specified URL, analyzes it, and then saves information in PictoBlox.

3. recognize () in image from (): This block extracts the image from Stage, Costume, or Backdrop in PictoBlox, analyzes it, and then saves information in

PictoBlox.

You can recognize the following in images:

1. Image Features – Computer vision
2. Handwritten Text
3. Printed Text
4. Face Details

Note: Once you make a recognition request in PictoBlox, it identifies all the features of the image in local variables. Then, when you request a particular feature, it reports it from the local variable itself so that you do not have to make recognition requests again and again.

Image Features Results

By analyzing the images for image features, you can recognize the following things:

1. Brands: Brand detection uses a database of thousands of global logos to identify commercial brands in images. The Computer Vision service detects if there are brand logos in a given image; if so, it returns the brand name, else it returns NULL.

2. Celebrity: Celebrity detection uses a database to identify celebrities in images. The Computer Vision service detects if there is a celebrity in a given image; if so, it returns their name, else it returns NULL.

Robert Downey Jr., American actor born on April 4, 1965
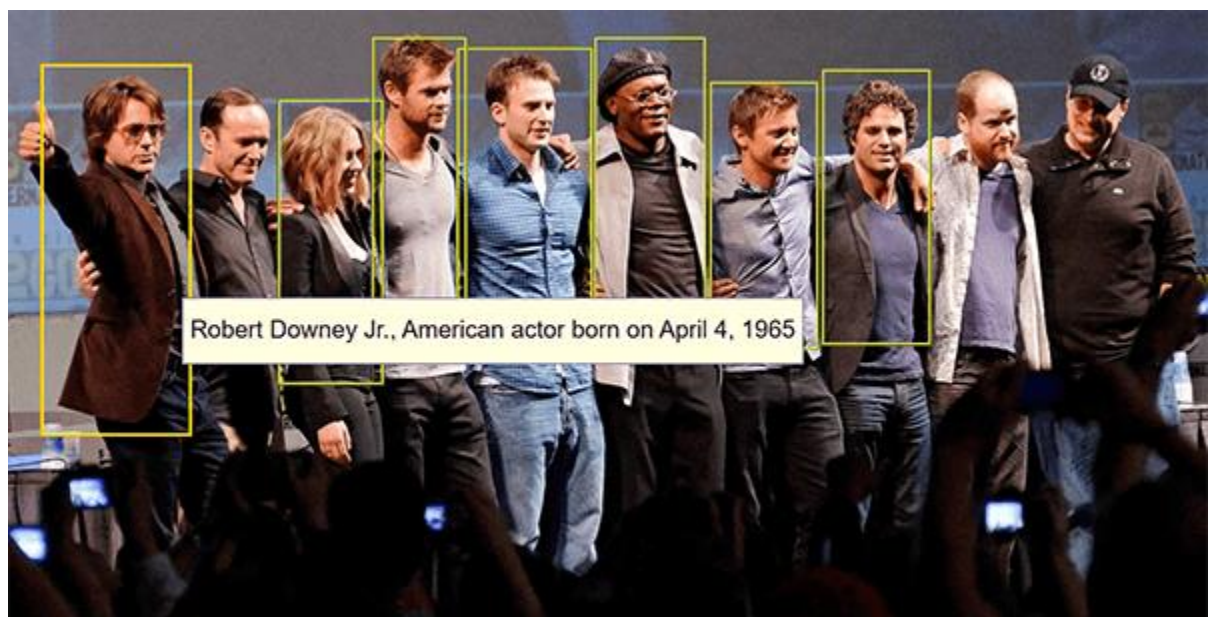
3. Objects: Computer vision detects if there are objects in a given image; if so, it returns their name, else it returns NULL.

4. Landmarks: Landmark detection uses a database of thousands of global landmarks to identify them in images, e.g. Taj Mahal.

5. Image Tags: Computer vision returns the taxonomy-based categories detected in an image. Computer Vision can categorize an image broadly or specifically

according to the 86 categories given in the following diagram:

abstract_	dark_	object_screen	people_	sky_cloud
abstract_net	dark_fire	object_sculpture	people_baby	sky_object
abstract_nonphoto	dark_fireworks	others_	people_crowd	sky_sun
abstract_rect	dark_light	outdoor_	people_group	text_
abstract_shape		outdoor_city	people_hand	text_mag
abstract_texture	drink_	outdoor_field	people_many	text_map
animal_	drink_can	outdoor_grass	people_portrait	text_menu
animal_bird	food_	outdoor_house	people_show	text_sign
animal_cat	food_bread	outdoor_mountain	people_swimming	trans_bicycle
animal_dog	food_fastfood	outdoor_oceanbeach	people_tattoo	trans_bus
animal_horse	food_grilled	outdoor_playground	people_young	trans_car
animal_panda	food_pizza	outdoor_pool	plant_	trans_trainstation
building_	indoor_	outdoor_railway	plant_branch	
building_arch	indoor_churchwindow	outdoor_road	plant_flower	
building_brickwall	indoor_court	outdoor_sportsfield	plant_leaves	
building_church	indoor_doorwindows	outdoor_stonerock	plant_tree	
building_corner	indoor_marketstore	outdoor_street		
building_doorwindows	indoor_room	outdoor_water		
building_pillar	indoor_venue	outdoor_waterside		
building_stair				
building_street				

6. Image Description: Human-readable sentence that describes the contents of the image.

Now, let's look at how we can use Computer vision in our PictoBlox projects. First, let see what blocks are available for computer vision.

recognized () count
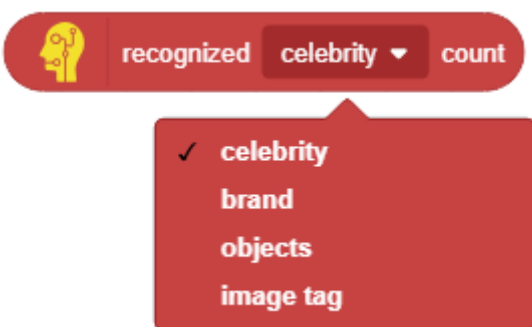
The recognized () count block reports the number of selected things to count in the image. There can be multiple celebrities, brands, objects, and image tags in a single image.
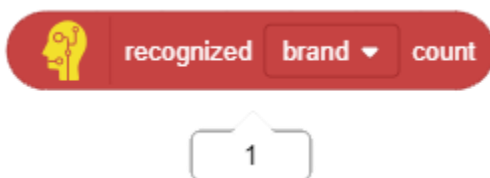
Let's recognize this image:

Copy the image from here (https://learn.thestempedia.com/wp-content/uploads/2020/04/Robert-and-Chris.jpg) and analyze the image using the URL.
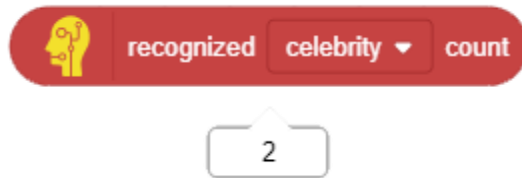

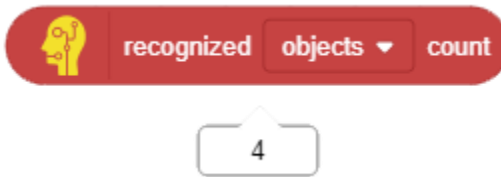The following counts are recognized in this image:
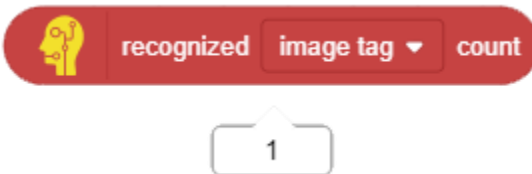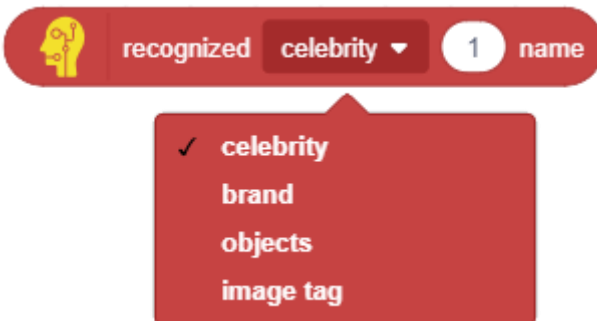
1. Brand: 1

2. Celebrity: 2

3. Object: 4

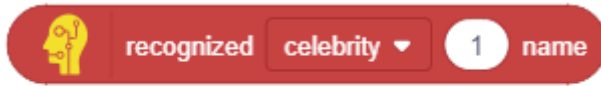4. Image Tag:1

recognized () () name

The recognized () () name block reports the name of the recognized celebrity/brand/object/image tag. There are two inputs to the block:
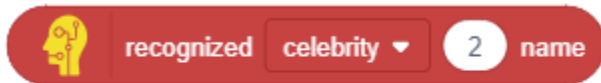
1. Type: Celebrity, Brand, Object or Image Tag
2. Number: # of the thing you want to analyze. E.g. if there are two celebrities, then you can get the celebrity 1 name and celebrity 2 name by changing this input.

If the number is out of range, then the block will return NULL.
E.g., for the image given above, you can get the following:

1. Celebrity Name: Chris Evans and Robert Downey Jr.

recognized | celebrity ▼ | 1 | name

Chris Evans

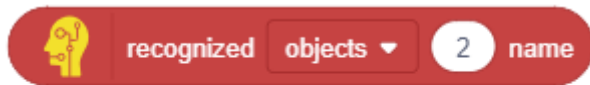recognized | celebrity ▼ | 2 | name

Robert Downey Jr.

2. Brand: Audi

recognized | brand ▼ | 1 | name
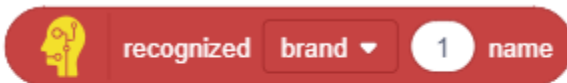
Audi

3. Objects: Tie, Tie, Person, Person

recognized | objects ▼ | 1 | name

tie

recognized | objects ▼ | 2 | name

tie

recognized | objects ▼ | 3 | name
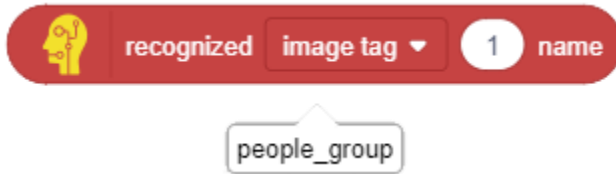
person

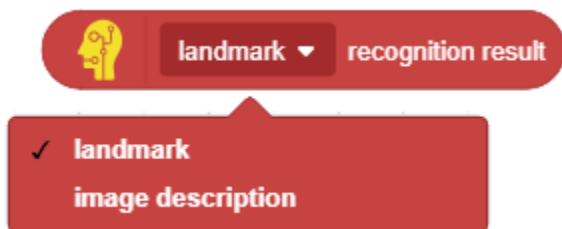recognized | objects ▼ | 4 | name
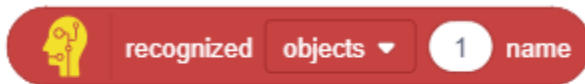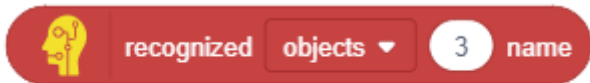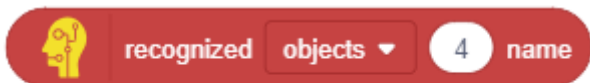
person

4. Image Tag: people_group

We will see how to use this block in our program in the next topic.

() recognition result

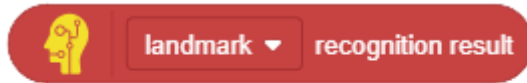The () recognition result block reports the landmark and the image description identified in the image.

Let's analyze this image:

Copy the image URL here: https://learn.thestempedia.com/wp-content/uploads/2020/04/Taj-Mahal-2.jpg

The following results come from the block:

1. Landmark: Taj Mahal

2. Image Description: a group of people standing in front of Taj Mahal

(END)