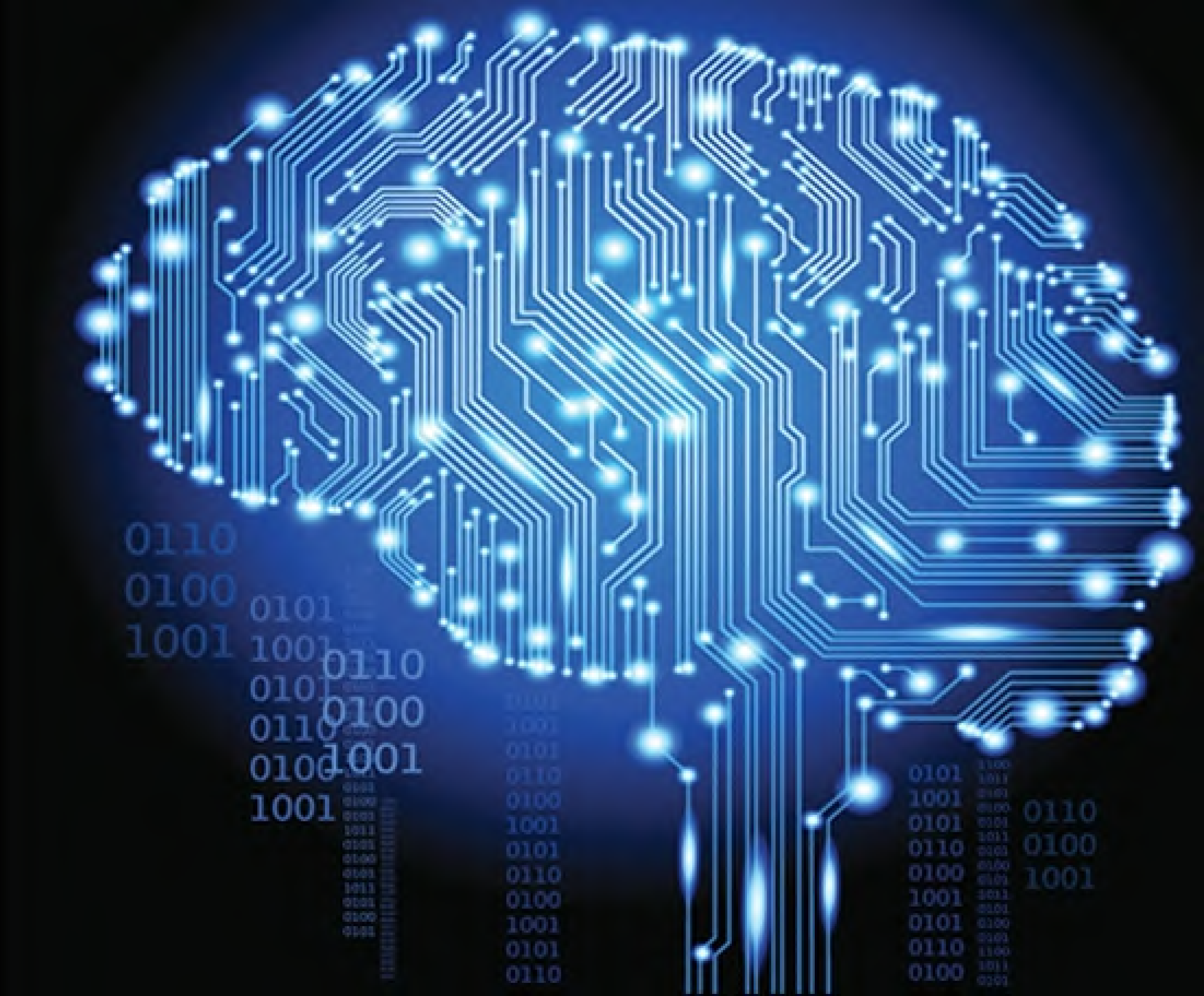
AI

Object Detection Project

Automated Object Detection using
Decision Trees & Random Forests



1) Project idea in details:

Object detection is a computer technology related to computer vision and image processing that deals with detecting instances of semantic objects of a certain class (such as humans, buildings, or cars) in digital images and videos.

Object detection is a computer vision technique for locating instances of objects in images or videos.

Object detection algorithms typically leverage machine learning or deep learning to produce meaningful results.

When humans look at images or video, we can recognize and locate objects of interest within a matter of moments. The goal of object detection is to replicate this intelligence using a computer.

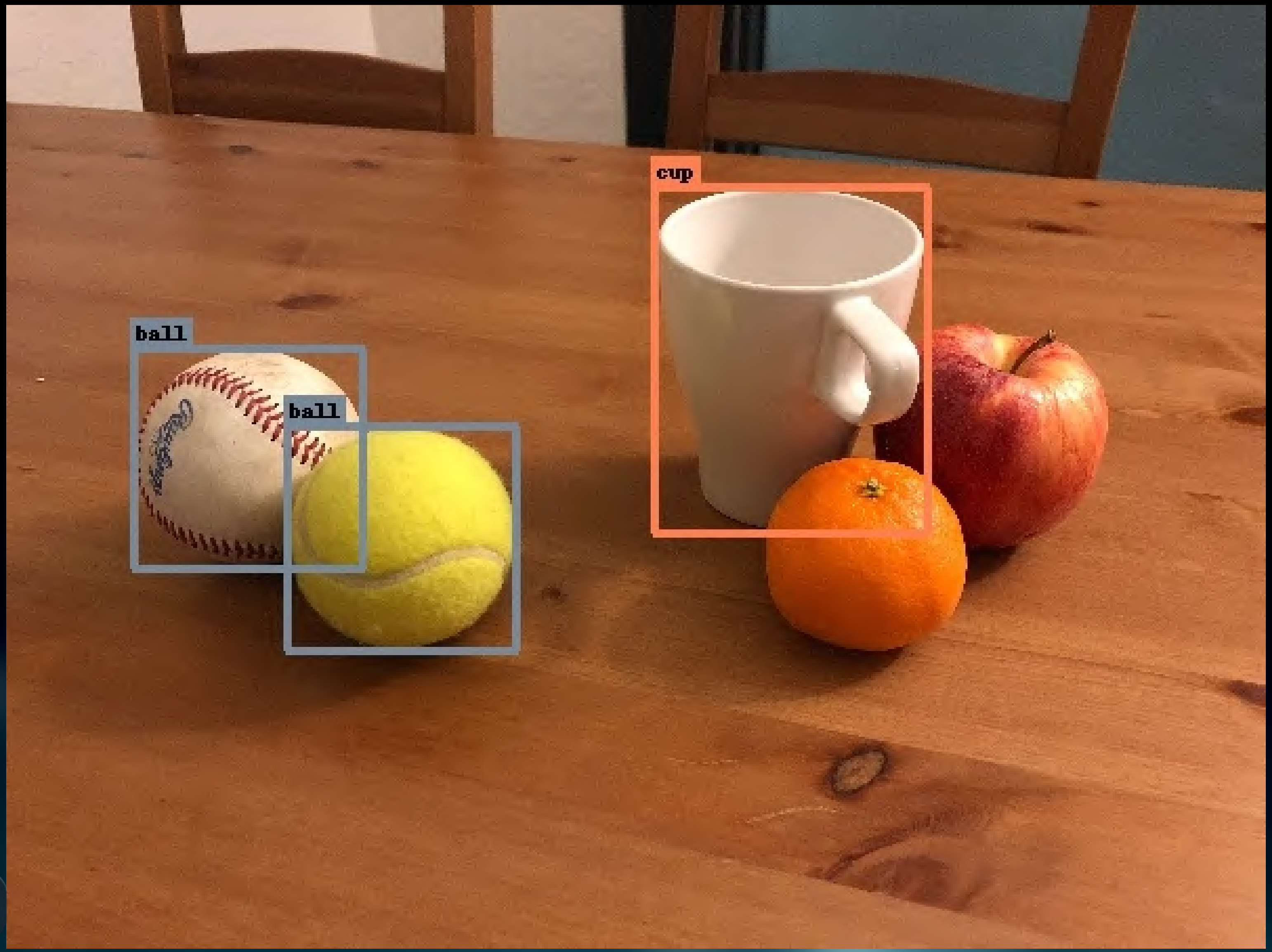
Object detection is a computer vision technique that allows us to identify and locate objects in an image or video. With this kind of identification and localization, object detection can be used to count objects in a scene and determine and track their precise locations, all while accurately labeling them

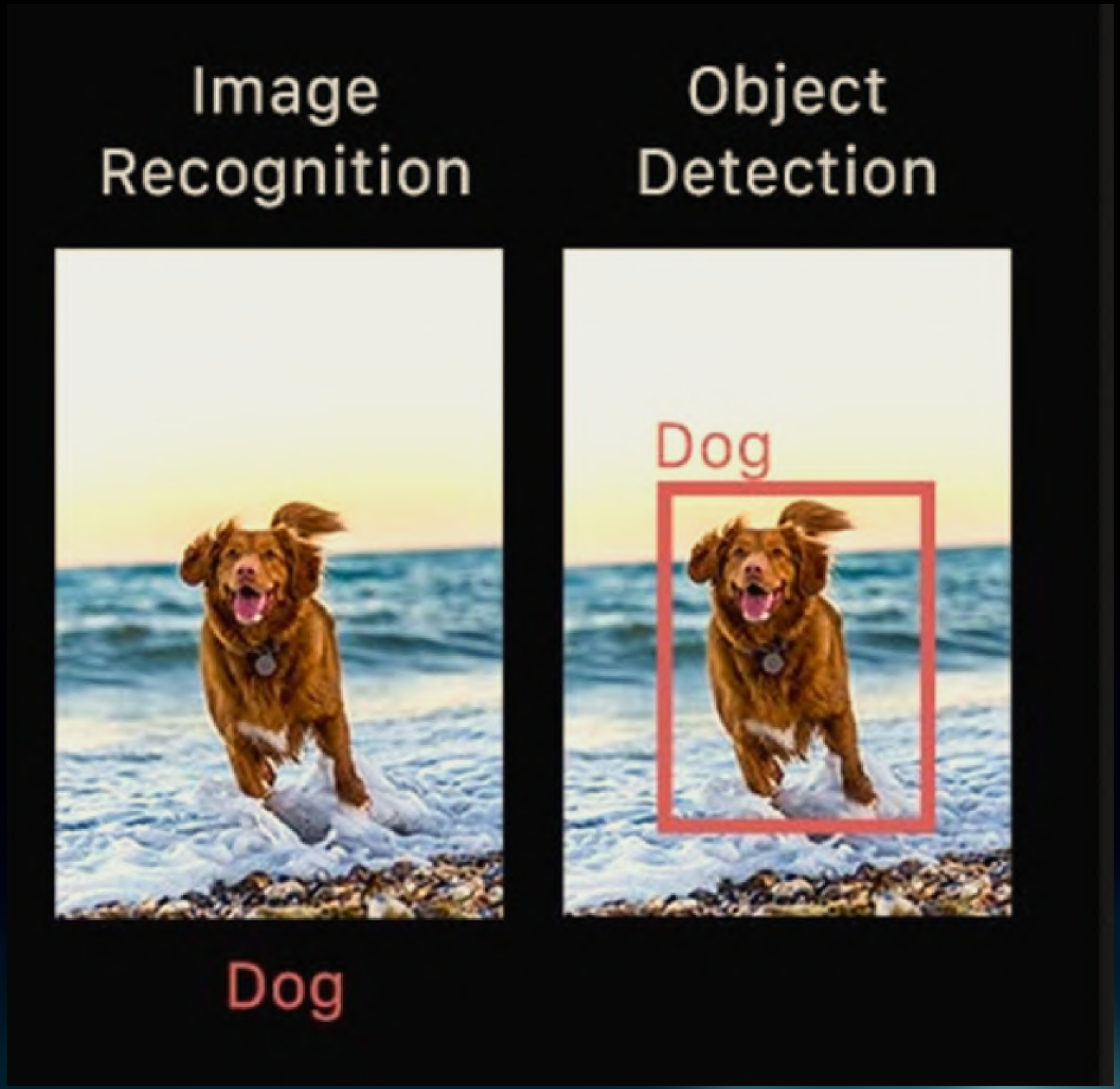
Imagine, for example, an image that contains cat . Object detection allows us to at once classify the type of thing found while also locating instance of it within the image.

Object detection is commonly confused with image recognition, so before we proceed, it’s important that we clarify the distinctions between them .

Image recognition assigns a label to an image. A picture of a dog receives the label “dog”. A picture of two dogs, still receives the label “dog”. Object detection, on the other hand, draws a box around each dog and labels the box “dog”. The model predicts where each object is and what label should be applied. In that way, object detection provides more information about an image than recognition.

Here’s an example of how this distinction looks in practice:





Modes and types of object detection

Broadly speaking, object detection can be broken down into:

1) machine learning-based approaches:

In more traditional ML-based approaches, computer vision techniques are used to look at various features of an image, such as the color histogram or edges, to identify groups of pixels that may belong to an object. These features are then fed into a regression model that predicts the location of the object along with its label.

2) deep learning-based approaches:

deep learning-based approaches employ convolutional neural networks (CNNs) to perform end-to-end, unsupervised object detection, in which features don't need to be defined and extracted separately. For a gentle introduction to CNNs, check out this overview

Because deep learning methods have become the state-of-the-art approaches to object detection, these are the techniques we'll be focusing on for the purposes of this guide.

hand, draws a box around each dog and labels the box "dog". The model predicts where each object is and what label should be applied. In that way, object detection provides more information about an image than recognition.

Why is object detection important?

Object detection is inextricably linked to other similar computer vision techniques like image recognition and image segmentation, in that it helps us understand and analyze scenes in images or video.
But there are important differences. Image recognition only outputs a class label for an identified object, and image segmentation creates a pixel-level understanding of a scene's elements. What separates object detection from these other tasks is its unique ability to *locate objects within an image or video*. This then allows us to count and then track those objects

Given these key distinctions and object detection's unique capabilities, we can see how it can be applied in a number of ways:

1) Crowd counting
2) Self-driving cars
3) Video surveillance
4) Face detection
5) Anomaly detection

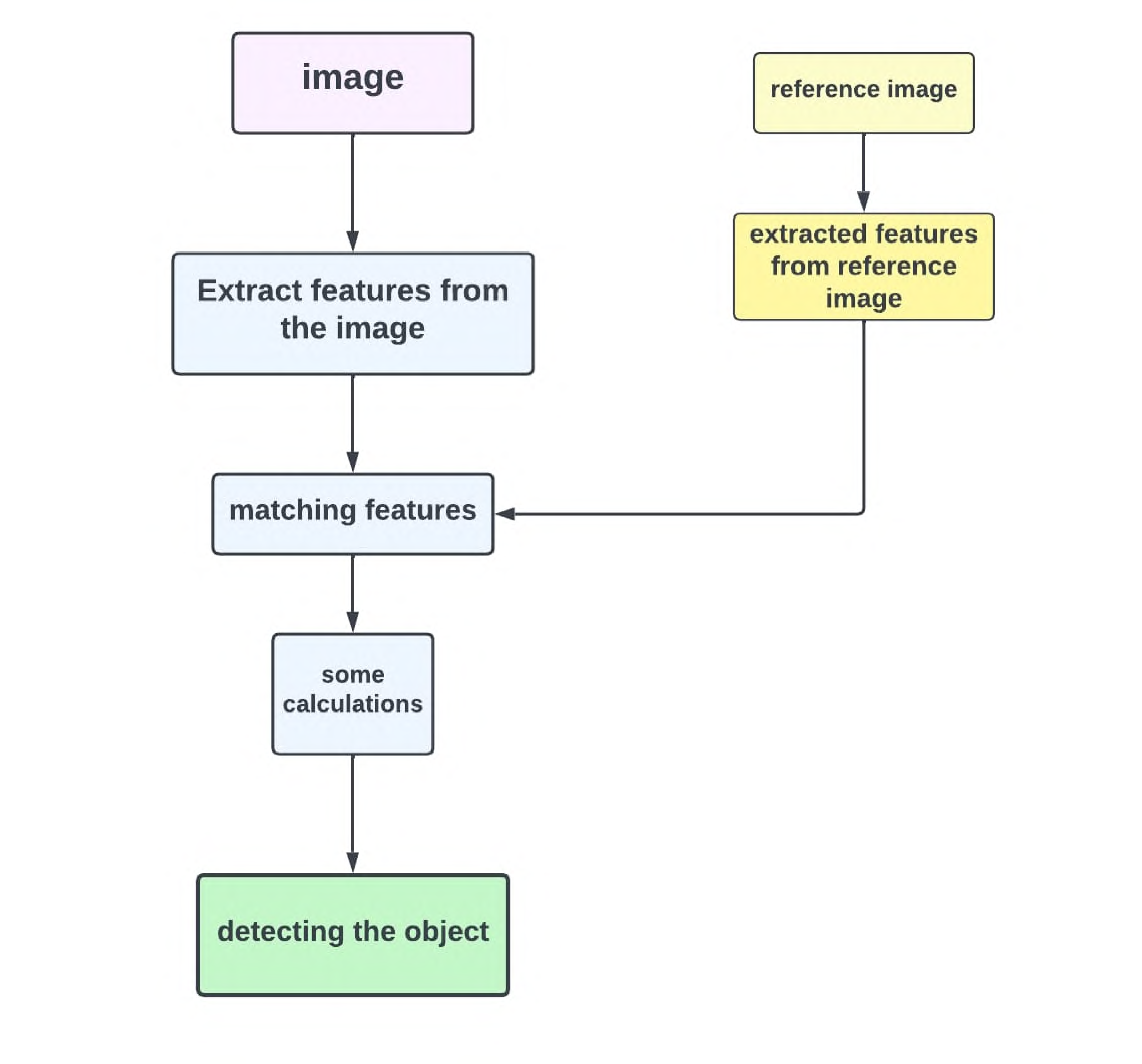
Pitch

2) Main functionalities:

Our main function is to identify and locate objects in an image, it allows us to classify the types of things while locating instances of them within the image So:

1) We firstly collect our related data
2) Preparing it (randomizing its order & preparing it to be used in the training)
3) Wrangling (preprocessing & cleansing).
4) Feature extraction (selection).
5) Data analysis (selection of machine learning techniques to be used [decision trees & random forest]).
6) Building the model.
7) Model training (so it can understand patterns & rules).
8) Testing (checking our model's accuracy).

block diagram to clarifies functionalities



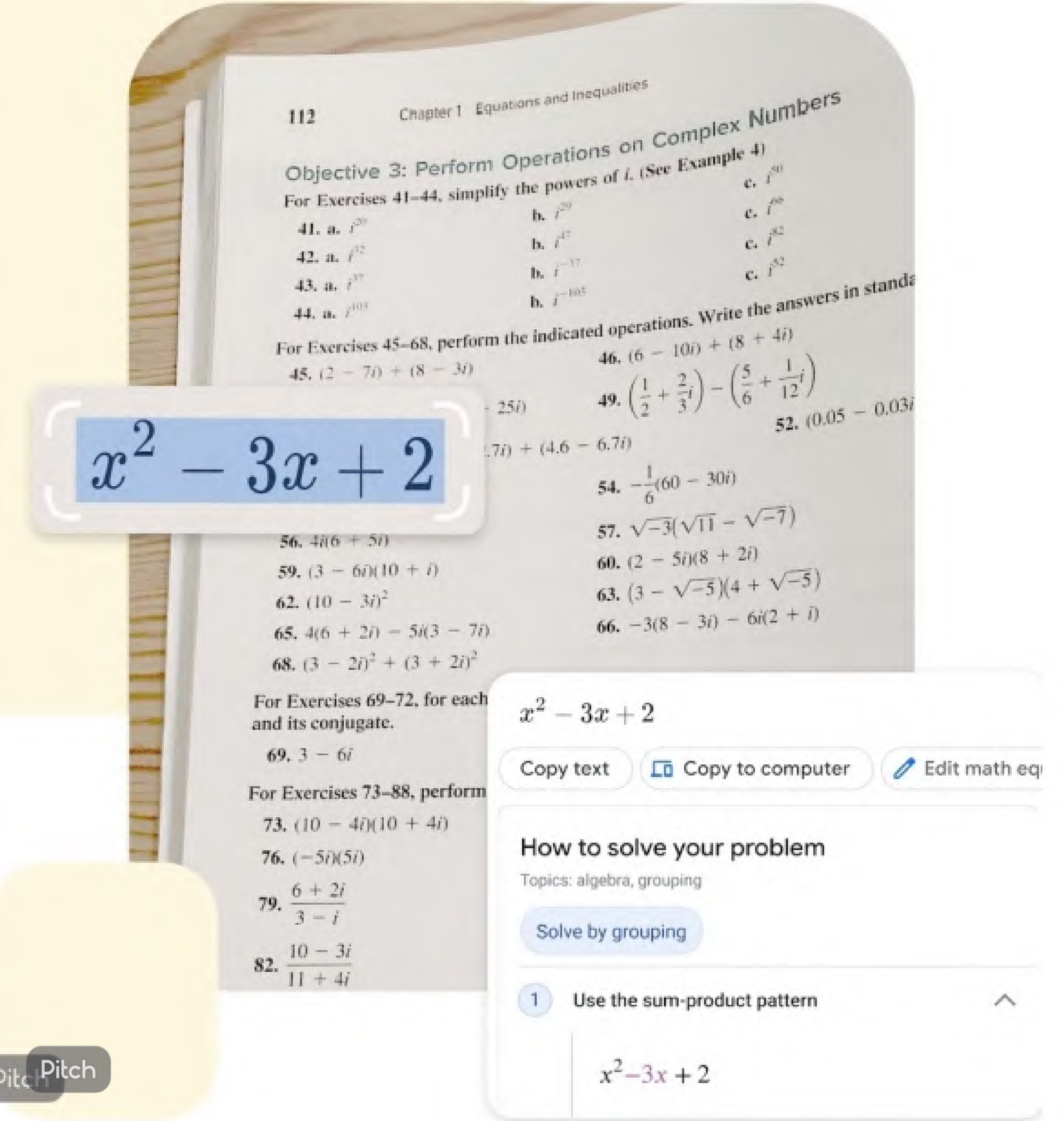
Pitch

3) Similar applications in the market

1) Google Lens https://lens.google/

Google Lens

Google Lens is an AI-powered technology that uses your smartphone camera and deep machine learning to not only detect an object in front of the camera lens, but understand it and offer actions such as scanning, translation, shopping, and more.



What can Google Lens do?

Google homework questions:

That's right, you can just scan the question and see what Google comes up with.

Shopping:

If you see a dress, you like while shopping, Google Lens can identify that piece and similar articles of clothing. This works for just about any item you can think of, accessing shopping or reviews

Search around you:

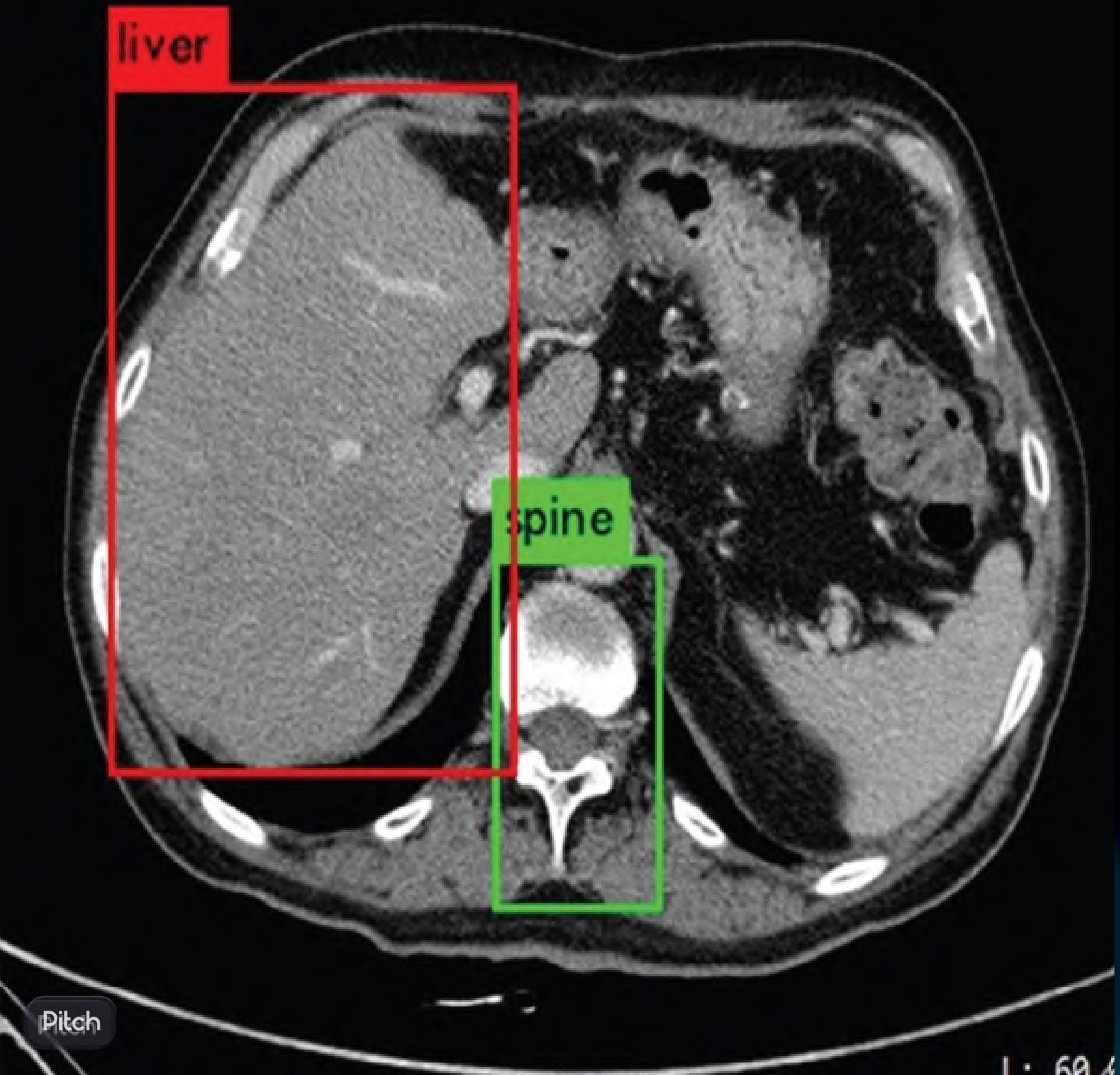
If you point your camera around you, Google Lens will detect and identify your surroundings. That might be details on a landmark or details about types of food - including recipes.

2) Managing SKUs in a retail store:

https://www.hindawi.com/journals/mpe/2022/4916818/

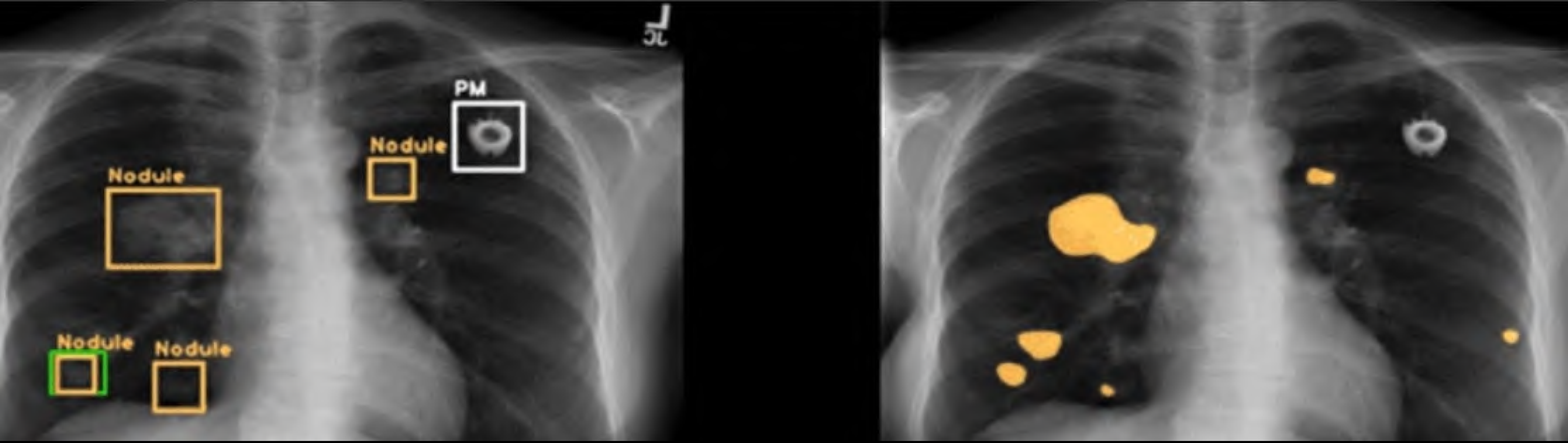
How does it work?

Object detection using machine learning detects SKUs (Stock Keeping Units) by analysing and comparing shelf images with the ideal state. Such neural networks are trained to flag gaps between reference planograms and the actual shelf images. It makes the job of auditor very easy, by providing them with real-time feedback on their handheld devices, so that they can take appropriate action immediately.



3) Medical Image Analysis :

https://paperswithcode.com/task/medical-object-detection



How does it work?

Source: IBM AI-based radiology tools don’t replace clinicians but support their decision-making. They flag acute abnormalities, identify high-risk patients or those needing urgent treatment, so that radiologists can prioritise their worklists. IBM’s research division in Haifa, Israel, is currently developing an AI/ML-based solution called Cognitive Radiology Assistant, which is a next-gen cognitive assistant for radiologists. The software solution provides support to the clinicians and radiologists, by analysing medical images and combining the insights with the patient’s medical records. The scientists also created a deep neural network that is specialised to identify potentially cancerous breast tissue.

4) An initial literature review of Academic publications (papers) relevant to the idea

1) Alternating Regression Forests for Object Detection and Pose Estimation by :

Samuel Schulter, Christian Leistner, Paul Wolhart, Peter M. Roth, Horst Bischof; Proceedings of the IEEE International Conference on Computer Vision

We presented Alternating Regression Forests, a novel Random Forest training procedure for regression tasks, which, in contrast to standard Random Regression Forests, optimizes any differentiable global loss function without sacrificing the computational benefits of Random Forests. ARFs are easy to implement and can be exchanged with standard Random Regression Forests without great efforts. This novel regressor gives better performance on machine learning benchmarks compared to Random Forests and Boosted Trees. Furthermore, we also integrated our ideas into two computer vision applications (object detection with Hough Forests and pose estimation from depth images). In both cases, ARFs could beat the baselines, illustrating the benefits of optimizing a global loss during training.

pdf:
https://openaccess.thecvf.com/content_iccv_2013/papers/Schulter_Alternating_Regression_Forests_2013_ICCV_paper.pdf

2) Context-Sensitive Decision Forests for Object Detection by :

Peter Kontschieder, Samuel Bulò, Antonio Criminisi, Pushmeet Kohli, Marcello Pelillo, Horst Bischof

In this work we have presented Context-Sensitive Decision Forests with application to the object detection problem. Our new forest has the ability to access intermediate prediction (classification and regression) information about all samples of the training set and can therefore learn from contextual information throughout the growing process. This is in contrast to existing random forest methods used for object detection which typically treat training samples in an independent manner. Moreover, we have introduced a novel splitting criterion together with a mode isolation technique, which allows us to (a) perform a priority-driven way of tree growing and (b) install novel context-based test functions to check for mutual object centroid agreements. In our experimental results on pedestrian detection we demonstrated superior performance with respect to state-of-the-art methods and additionally found that our new algorithm can significantly better exploit training data containing multiple training objects.

pdf: https://proceedings.neurips.cc/paper/2012/file/bcbe3365e6ac95ea2c0343a2395834dd-Paper.pdf

3) An Introduction to Random Forests for Multi-class Object Detection by :
Juergen Gall & Nima Razavi & Luc Van Gool

this paper might overcome the overfitting problem of random forests partially. Due to its relation to implicit shape models [23], the detection approach shares advantages and limitations of this type of models. While techniques like back- projection and feature sharing allow to reason about object hypotheses and the similarity of categories, which goes beyond black box classifiers, the independent assumption of the image patches is a weakness of these models that needs to be addressed in the future. Nevertheless, random forests have a strong potential for applications where many labeled examples are available. For instance, pose or body part estimation from depth data [13, 19] are examples where accurate results can be obtained in real-time. The work [19] also shows the benefit of engineering where a fine tuned version of [18] resulted in a speed-up by a factor of 3200

pdf: https://pages.iai.uni-bonn.de/gall_juergen/download/jgall_introforest.pdf

4) Hough Forests for Object Detection, Tracking, and Action Recognition by :
Juergen Gall & Angela Yao & Nima Razavi & Luc Van Gool & V. Lempitsky

In general, Hough forests provide an excellent balance between high detection accuracy and time efficiency both at training and test time. Similar to random forests, it is expected that an implementation of a Hough Forest on a GPU [34], [89] would give an extra significant speed-up.\

pdf: https://files.is.tue.mpg.de/jgall/download/jgall_houghforest_pami11.pdf

5) The Dataset employed

1) Balls

https://www.kaggle.com/datasets/gpiosenka/balls-image-classification

2) Cats & Dogs

https://www.kaggle.com/datasets/shaunthesheep/microsoft-catsvsdogs-dataset

3) Flowers

https://www.kaggle.com/datasets/imsparsh/flowers-dataset

4) Vegetable

https://www.kaggle.com/datasets/misrakahmed/vegetable-image-dataset

5) Cars

https://www.kaggle.com/datasets/jessicali9530/stanford-cars-dataset?select=cars_test

6) Details of the algorithm / approach that will be used

A decision tree is a supervised machine learning algorithm that can be used for both classification and regression problems. A decision tree is simply a series of sequential decisions made to reach a specific result.

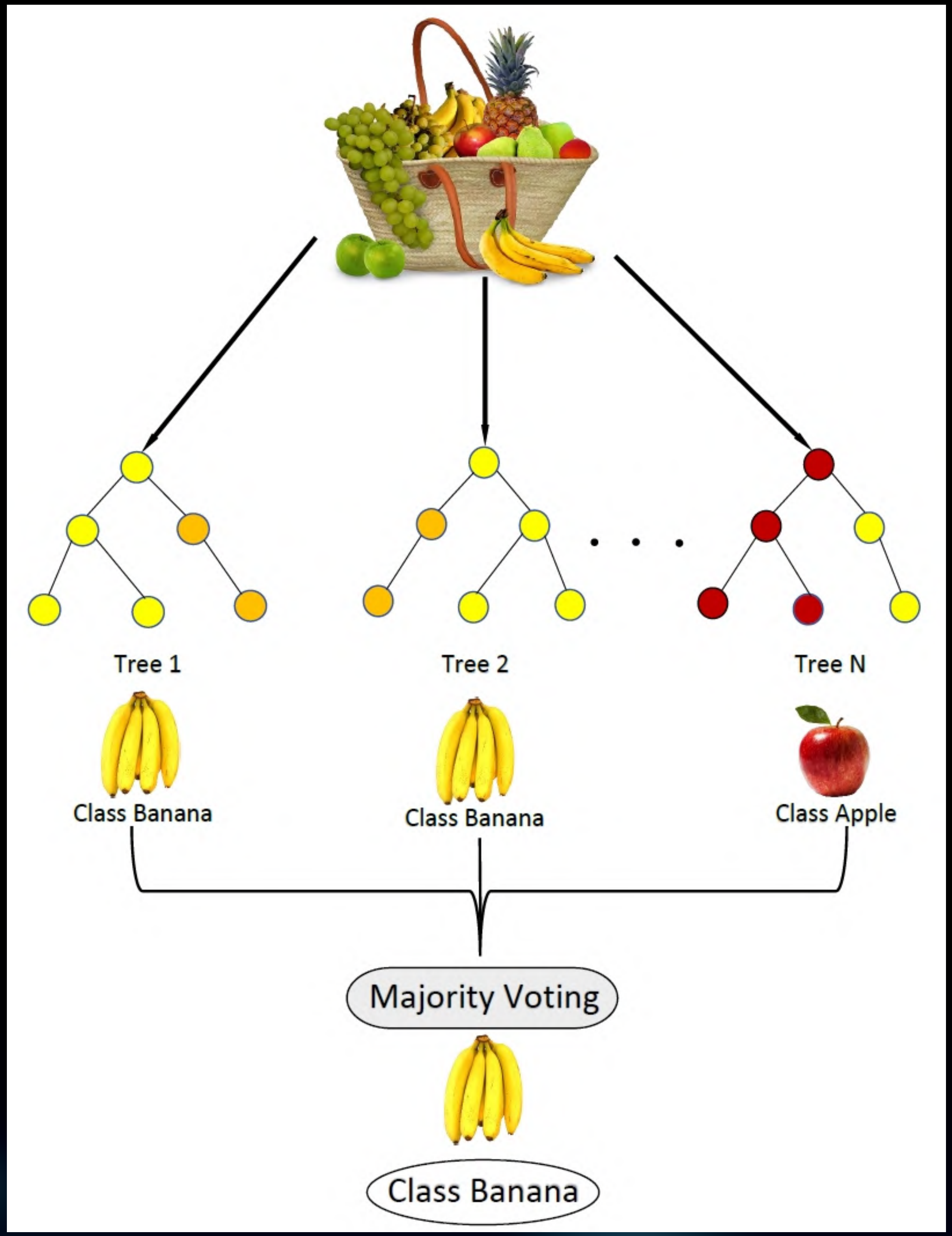
Random Forest is a tree-based machine learning algorithm that leverages the power of multiple decision trees for making decisions. As the name suggests, it is a “forest” of trees! It's random because it's forest of randomly created by decision tree.

In simple words:

The Random Forest Algorithm combines the output of multiple (randomly created) Decision Trees to generate the final output.

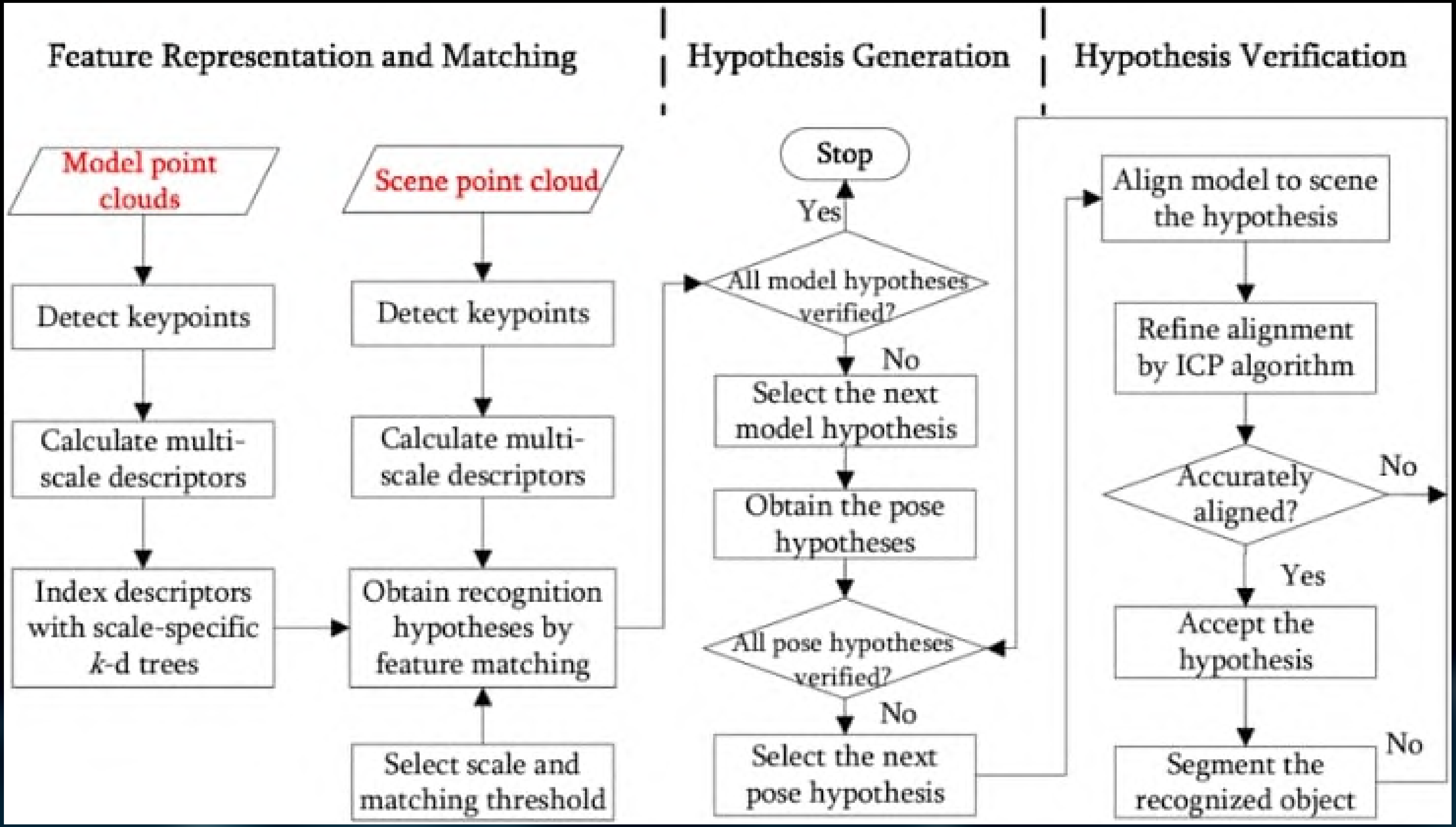
How does Random Forest work ?

Let's say we want to classify the different types of fruits in a bowl based on various features, but the bowl is cluttered with a lot of options. You would create a training dataset that contains information about the fruit, including colors, diameters, and specific labels (i.e., apple, grapes, etc.) You would then need to split the data by sorting out the smallest piece so that you can split it in the biggest way possible. You might want to start by splitting your fruits by diameter and then by color. You would want to keep splitting until that particular node no longer needs it, and you can predict a specific fruit with 100 percent accuracy.



Pitch

For more details we can read this flowchart :



Pitch

Advantages of Random Forest Algorithm :

- Can perform both Regression and classification tasks.
- Produces good predictions that can be understood easily.
- Can handle large data sets efficiently.
- Provides a higher level of accuracy in predicting outcomes over the decision algorithm

Disadvantages of Random Forest Algorithm :

- While using a Random Forest Algorithm, more resources are required for computation.
- Consumes more time compared to the decision tree algorithm.
- Less intuitive when we have an extensive collection of decision trees.
- Extremely complex and requires more computational resources.