



# CONTENTS

## Module # 1 – Theoretical Basis

### **Introduction**

- What is CV?
- Applications of CV across industries

### **Building Blocks of CV**

- Digital Images: Pixels and Channels
- Seeing Through Images: Image Processing
- Classifying vs. Detecting

### **The Power of Deep Learning in CV**

- Introducing Deep Neural Networks (DNNs)
- Convolutional Neural Networks (CNNs) – The Stars
- The Importance of Training Data

### **Conclusion**



# CONTENTS

## Module # 2 - Hands-on Projects

### **Introduction to environment**

- Familiarization with Python Libraries

### **Project # 1 Hand Gesture Recognition**

- Familiarization with the underlying principles
- Familiarization with Program and Algorithm
- Discussion and Demonstration

### **Project # 2 Gaze Detection**

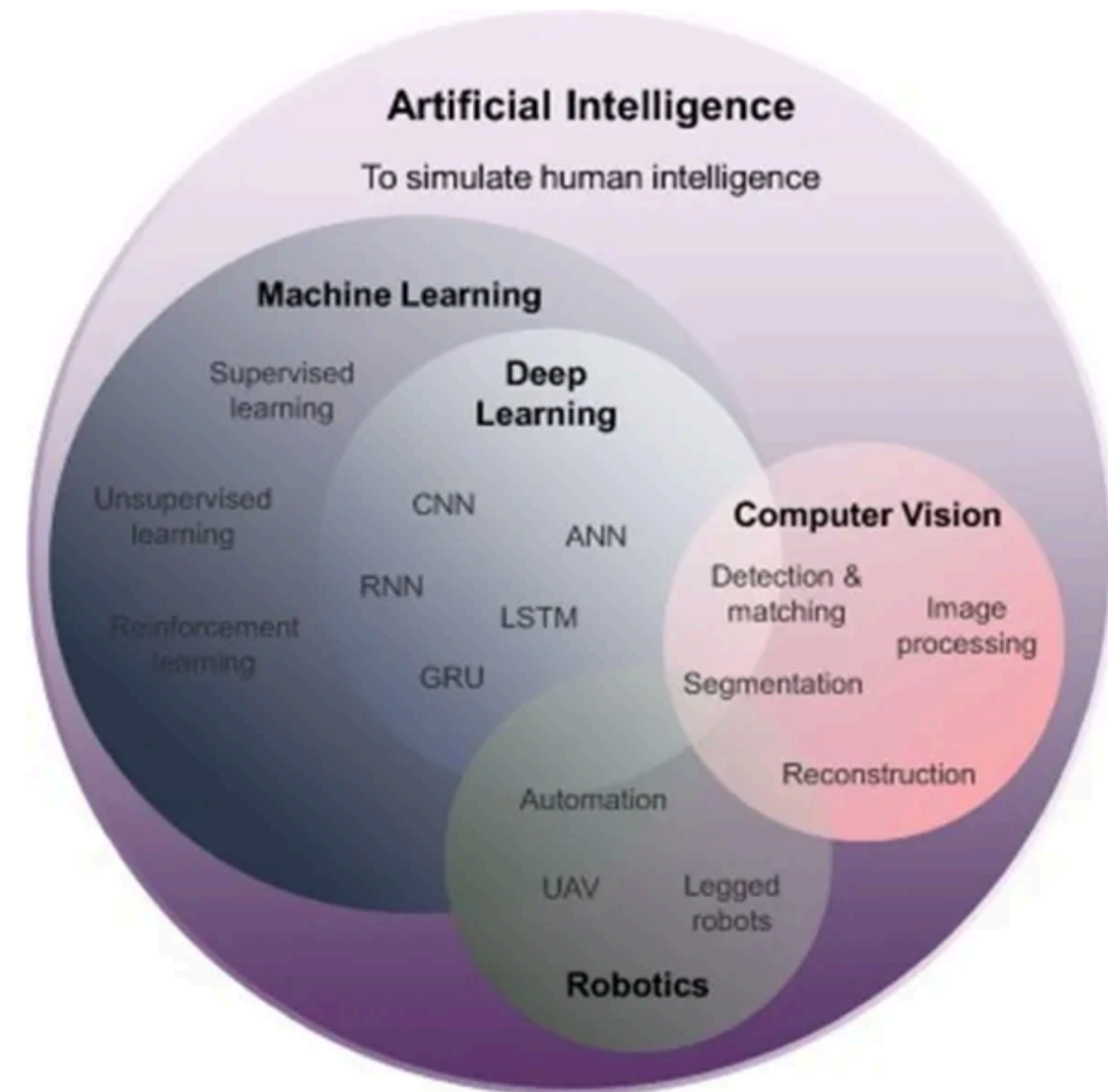
- Familiarization with the underlying principles
- Familiarization with Program and Algorithm
- Discussion and Demonstration



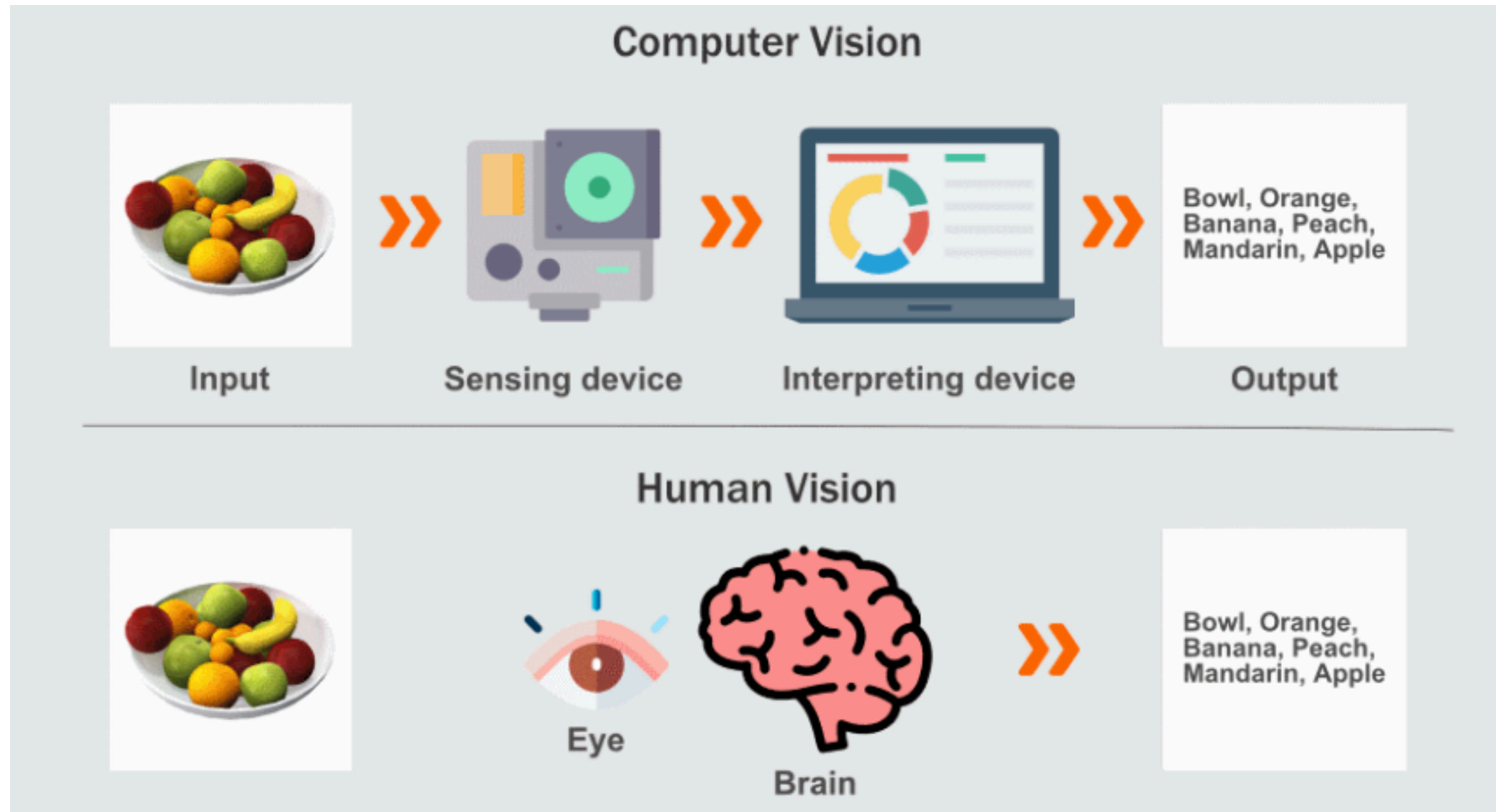
# INTRODUCTION

## What is Computer Vision?

- A field of computer science that focuses on enabling computers to identify and understand objects and people in images and videos.
- Seeks to perform and automate tasks that replicate human capabilities.
- Seeks to replicate both the way humans see, and the way humans make sense of what they see.



# INTRODUCTION

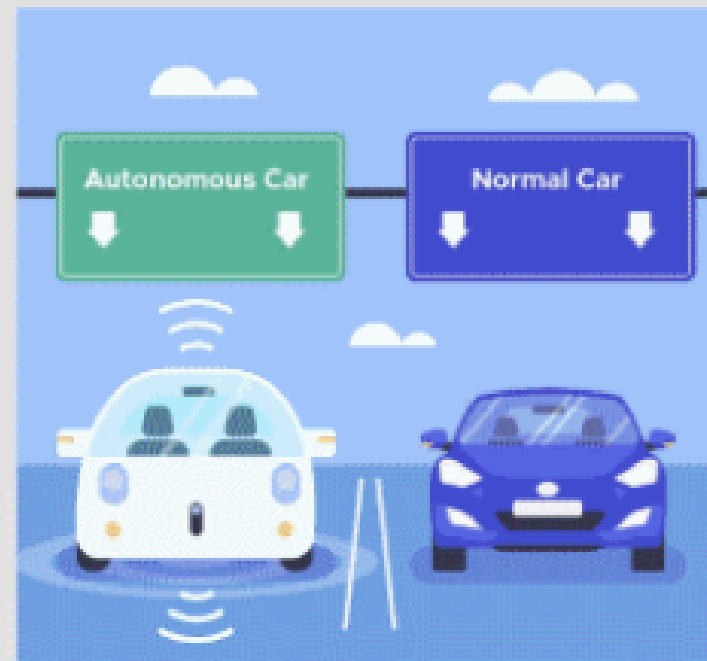


# INTRODUCTION

## World of CV



Facial recognition software



Self-driving cars



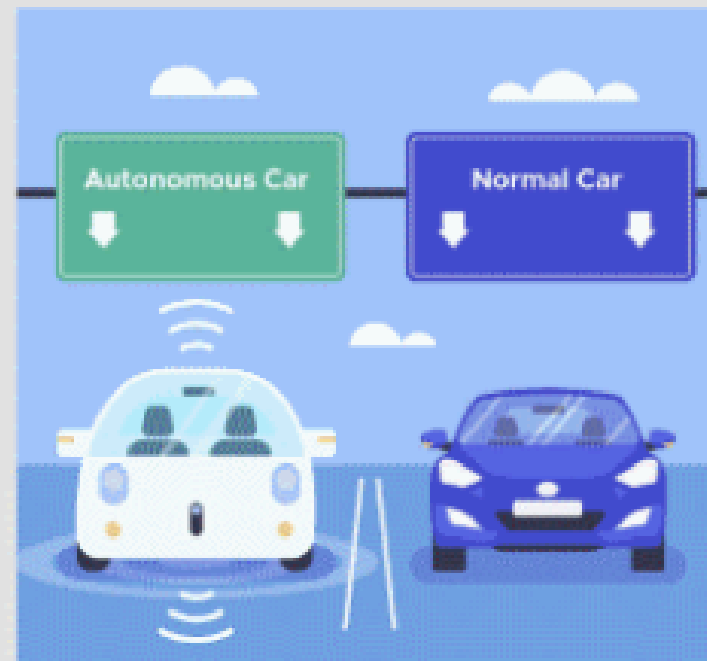
Manufacturing production line

# INTRODUCTION

## World of CV



Facial recognition software



Self-driving cars



Manufacturing production line

# INTRODUCTION

## World of CV



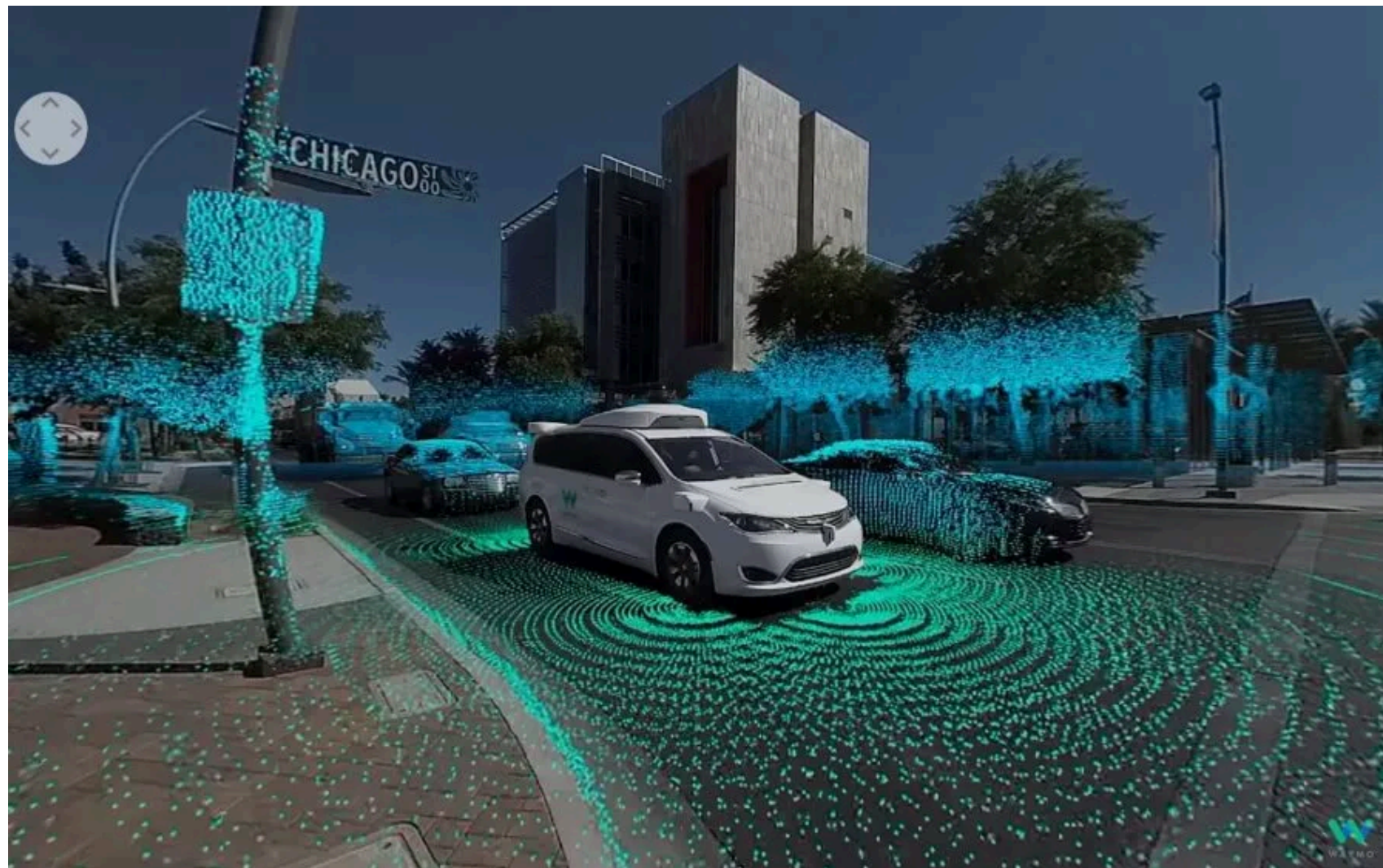
Mitek's computer vision-driven image recognition technology does away with the need for a customer to visit a branch to verify their identity to open an account.



# INTRODUCTION

---

## World of CV



*Waymo (part of Google) is equipping cars with cameras and computer vision systems, providing a 360-degree view around the car.*

*Waymo's computer vision system is one of the key technologies it uses to 'see' other cars and navigate traffic*

# INTRODUCTION

---

## World of CV



*Tesla Autopilot with computer vision technology.*

*The system also includes ultrasonic sensors, allowing the car to detect both hard and soft objects.*

*Tesla's forward-facing radar allows the car to see despite hazardous weather conditions such as heavy rain or fog*

# INTRODUCTION

---

## World of CV



Amazon is not the only tech giant launching checkout less store, the Chinese internet giant Lenovo has also jumped onto the bandwagon. Instead, CV powered facial recognition cameras monitor when items are taken from shelves and placed in carts.

# INTRODUCTION

---

## World of CV



*Shell is using predictive maintenance including computer vision tech to keep their equipment in prime condition.*

# INTRODUCTION

---

## World of CV



*Bossa Nova Robotics shelf-scanning robots*

These robots are able to identify products with missing labels as well as items that are out of stock or incorrectly priced.

# INTRODUCTION

## World of CV



*ModelingCafe created a fashion image with computer vision. **Guess which one is a fake?***

# INTRODUCTION

---

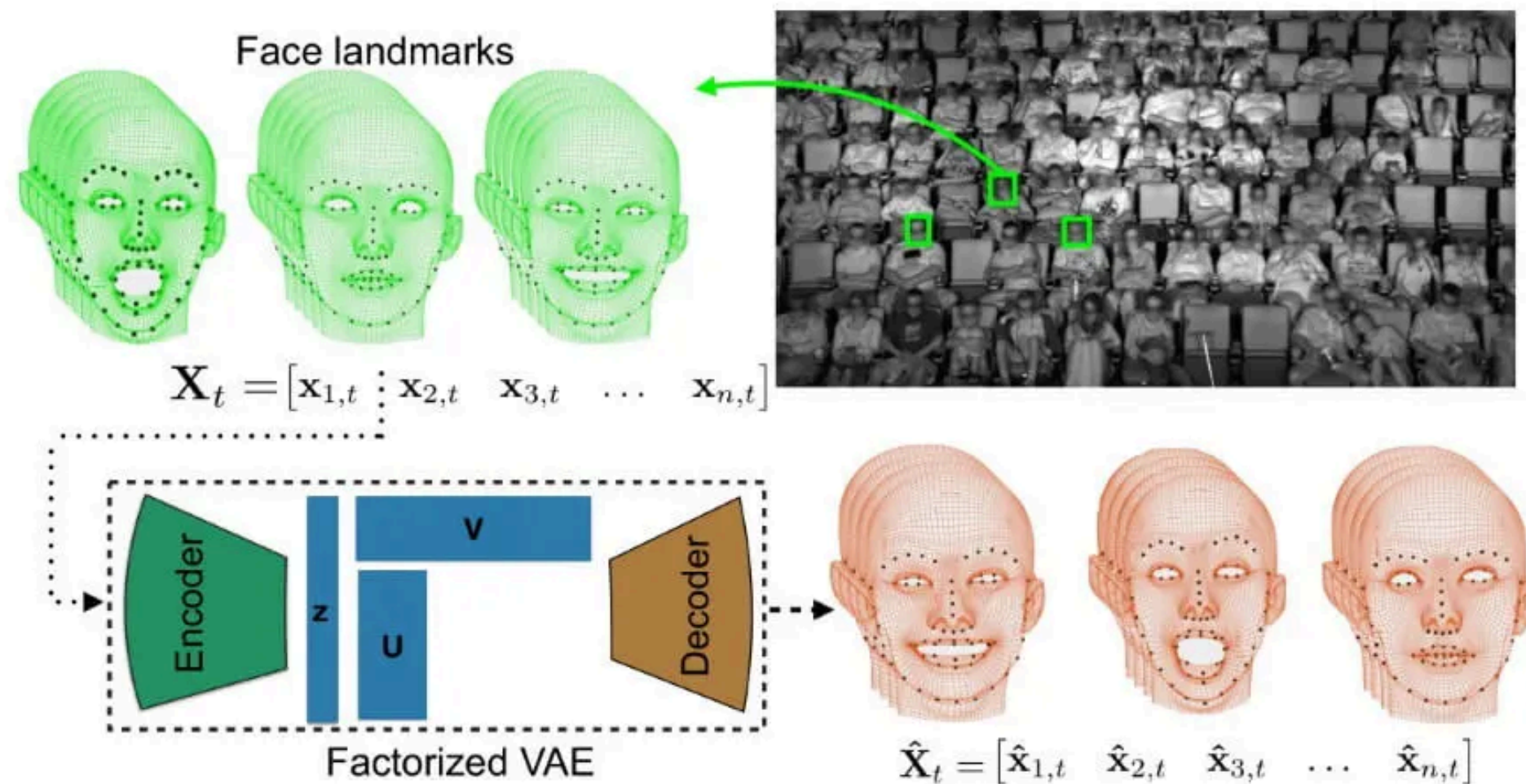
## World of CV



*Star Wars The Force Awakens was one of the films Disney used computer vision systems to track the focus and emotions of the customer.*

# INTRODUCTION

## World of CV



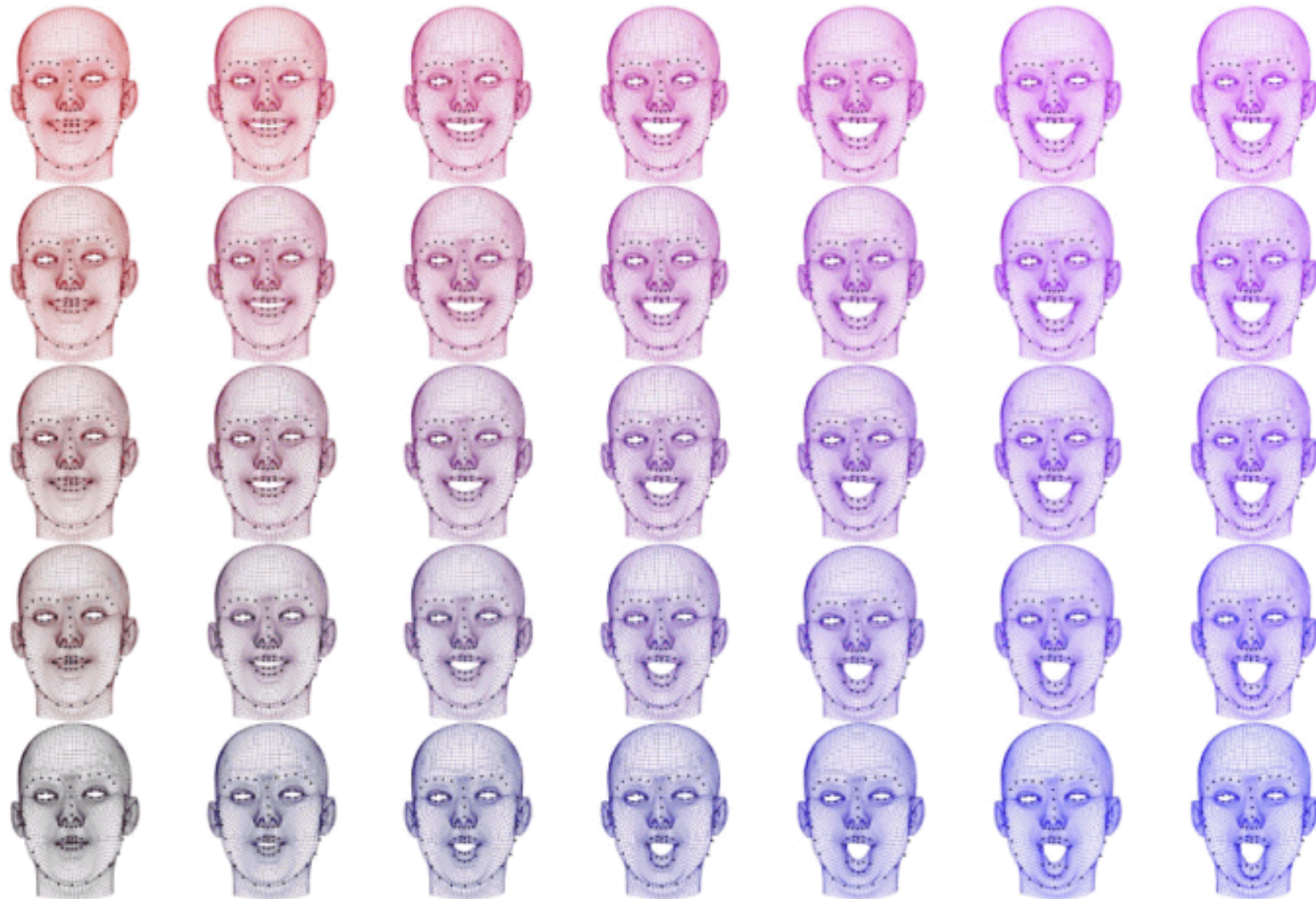
Disney utilized a new algorithm of facial recognition called factorized variational autoencoders or FVAEs. The FVAEs algorithm is searching from the audience who display similar facial expressions throughout the whole movie to determine what the typical expression to the film is, similar to how e-commerce sites recommend items for purchases.



# INTRODUCTION

---

## World of CV

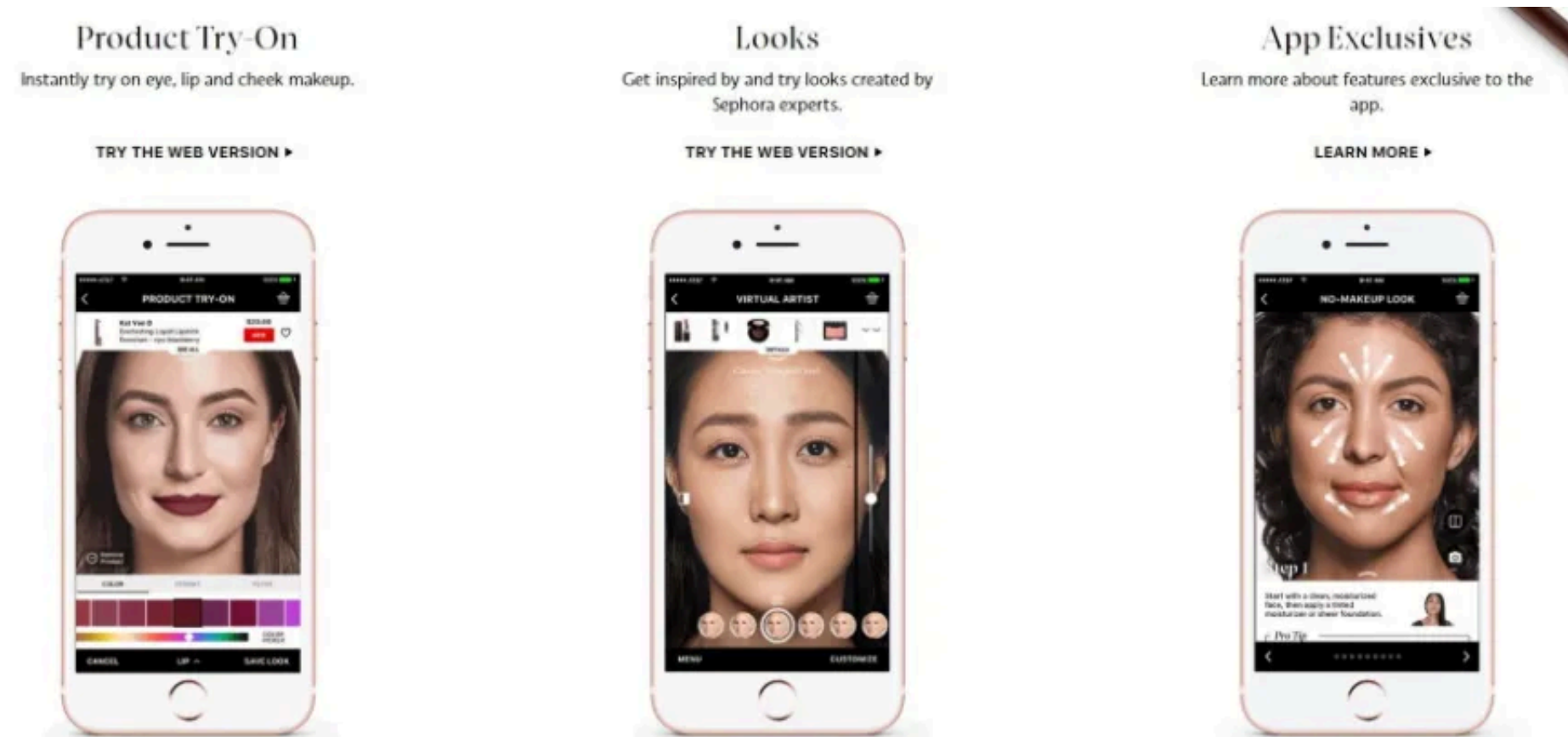


*Disney captured 16 million facial landmarks from 3,179 viewers. The FVAEs algorithm 'learnt' themselves how to recognize and understand facial expressions such as laughing smiling.*

# INTRODUCTION

---

## World of CV



*Sephora App makes use of computer vision to allow you to try before you buy makeup.*

Additionally, Sephora also uses AI and smart tools to help customers find the perfect shade or product for their skin tone.

# INTRODUCTION

---

## World of CV

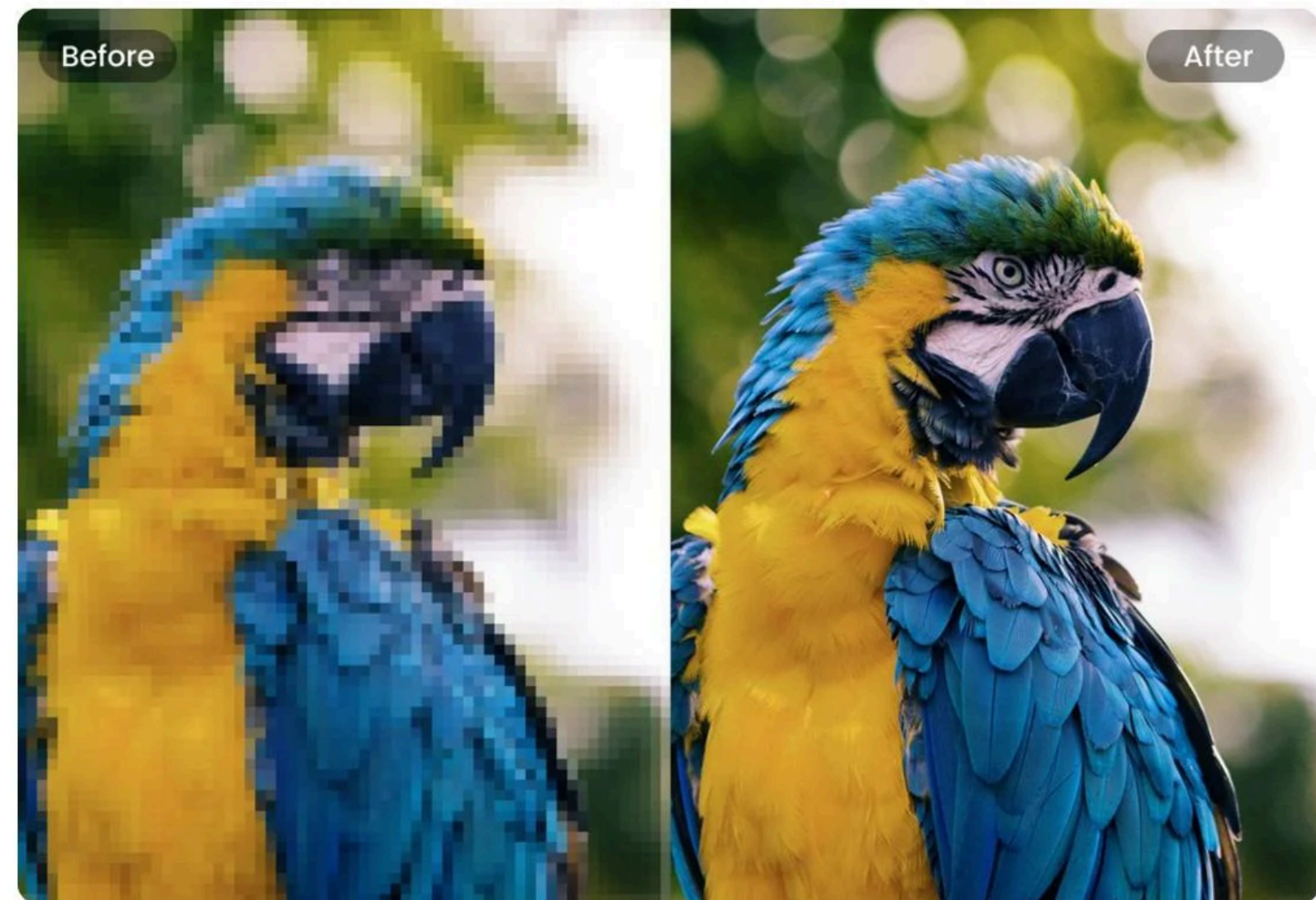


*Cainthus have developed a computer vision and predictive image analysis system that is able to identify an individual cow and its eating habits.*

# BUILDING BLOCKS OF CV

## Digital Images: Pixels and Channels

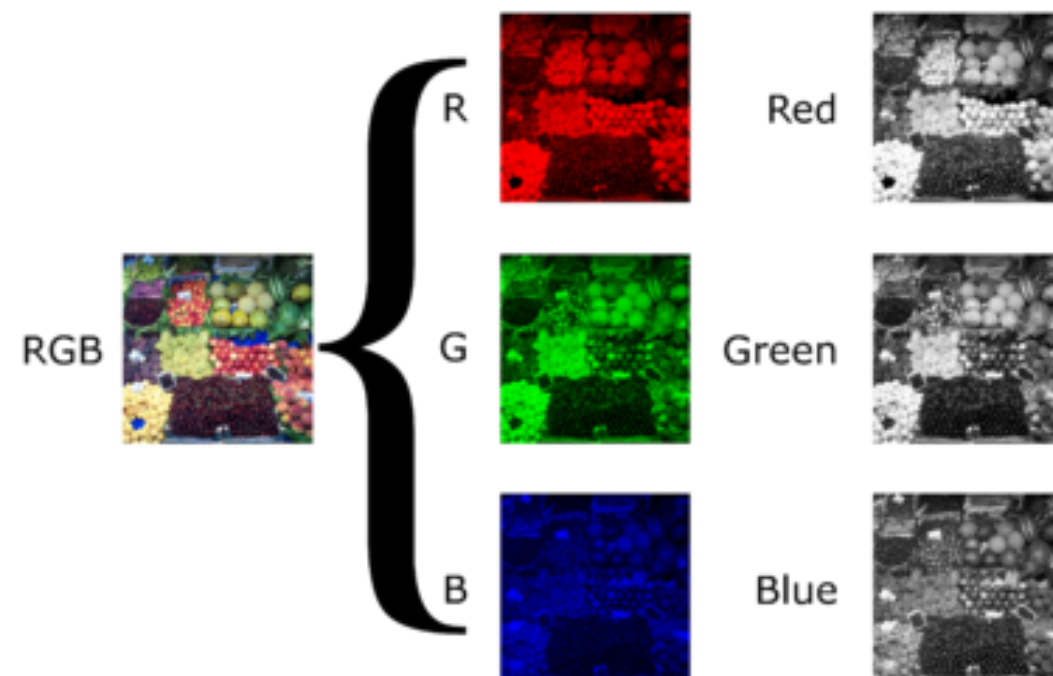
- Digital images are composed of tiny squares called pixels.
- Each pixel represents a single point in the image and contains color information.



# BUILDING BLOCKS OF CV

## Digital Images: Pixels and Channels

- RGB (Red, Green, Blue) channels combine to form color images.
- Grayscale images use a single channel to represent brightness variations.



# BUILDING BLOCKS OF CV

## Seeing through Images: Image Processing

- Feature extraction is a process in machine learning and data analysis that involves identifying and extracting relevant features from raw data.
- These features are later used to create a more informative dataset, which can be further utilized for various tasks such as: Classification. Prediction.



# BUILDING BLOCKS OF CV

## Capturing the World (Cameras & Sensors)

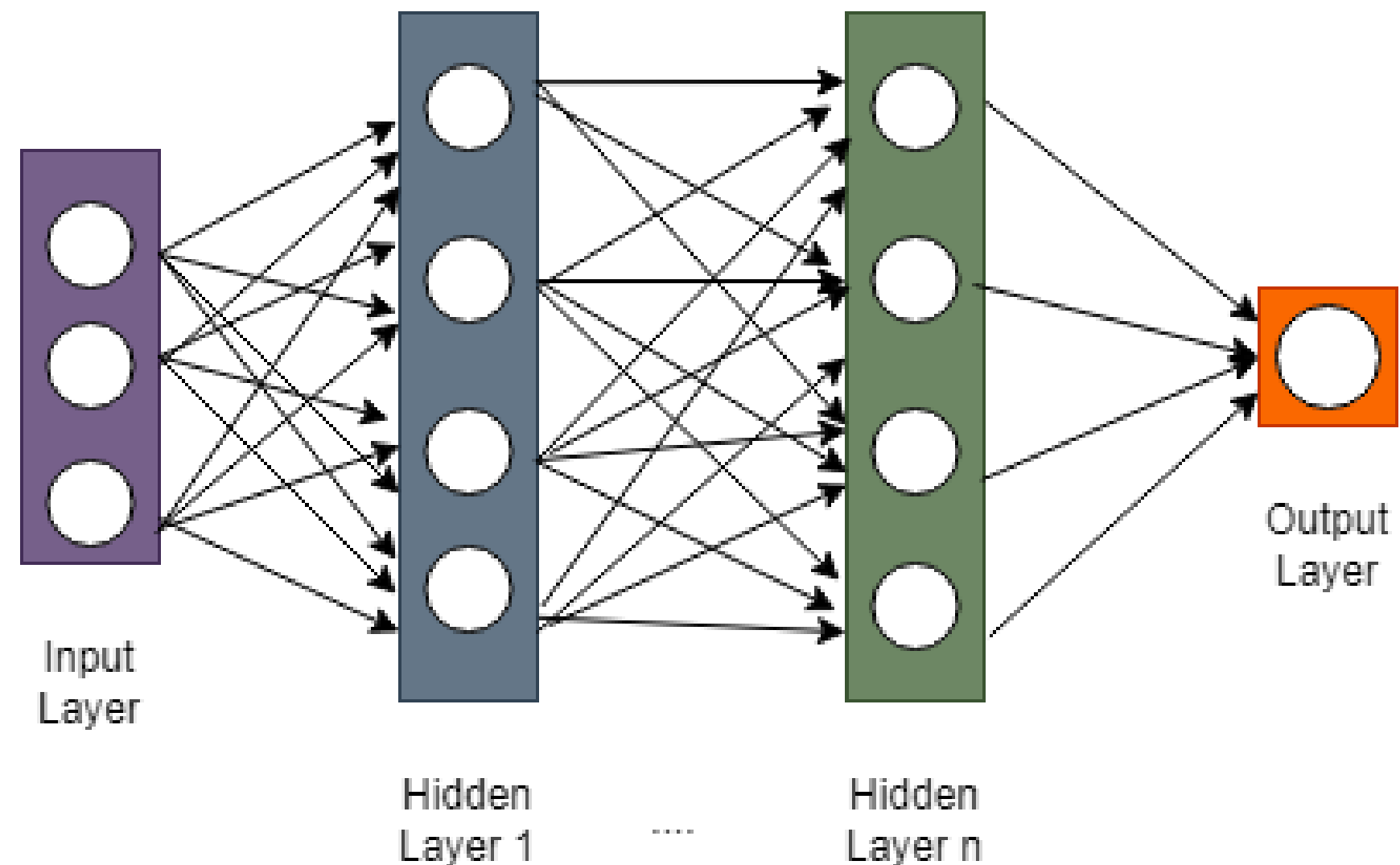
- Digital cameras and sensors capture visual data as digital images.
- Different types of cameras and sensors exist for various applications.
- For example:



# THE POWER OF DEEP LEARNING

## Introduction to DNN

- Deep Neural Networks (DNNs) are powerful tools for complex tasks.
- They learn from vast amounts of data to identify patterns and make predictions.
- DNNs excel in image recognition and other computer vision applications.

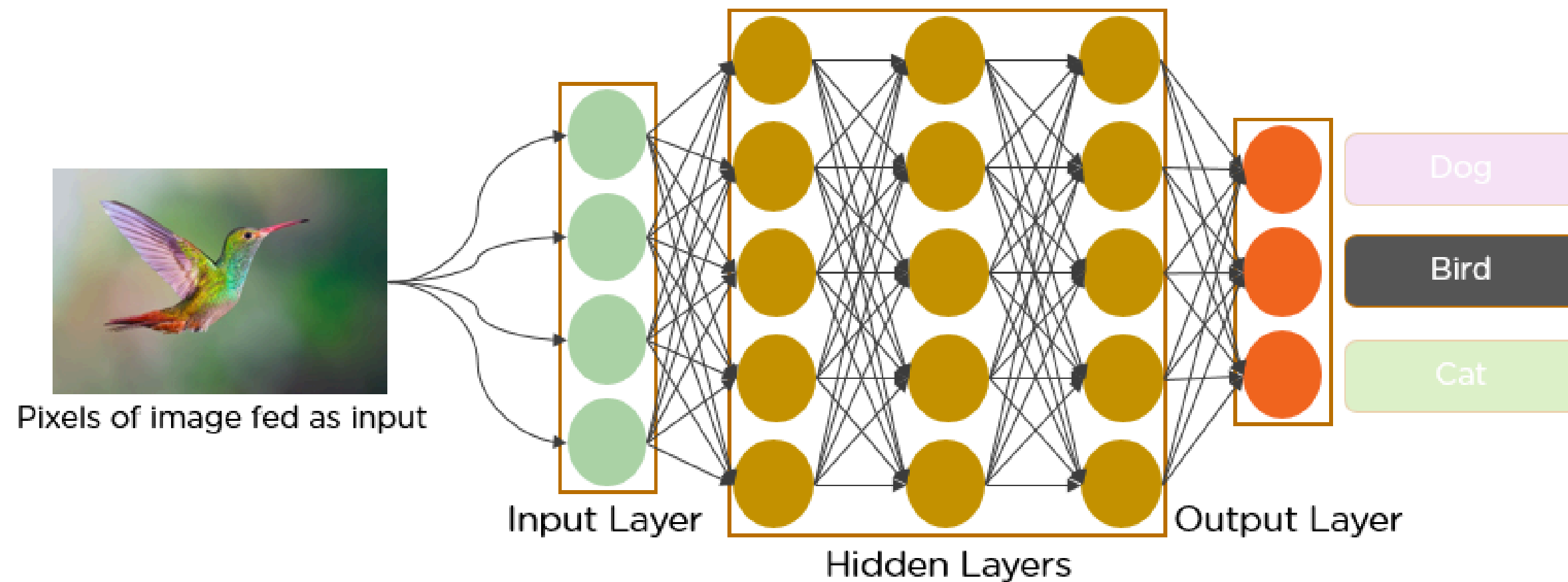




# THE POWER OF DEEP LEARNING

## Introduction to CNN

- Specialized DNNs designed for computer vision tasks.
- Excel at learning spatial features in images (shapes, arrangements).
- Convolutional layers extract features by applying filters, looking for patterns.



# THE POWER OF DEEP LEARNING

## Importance of Training Data

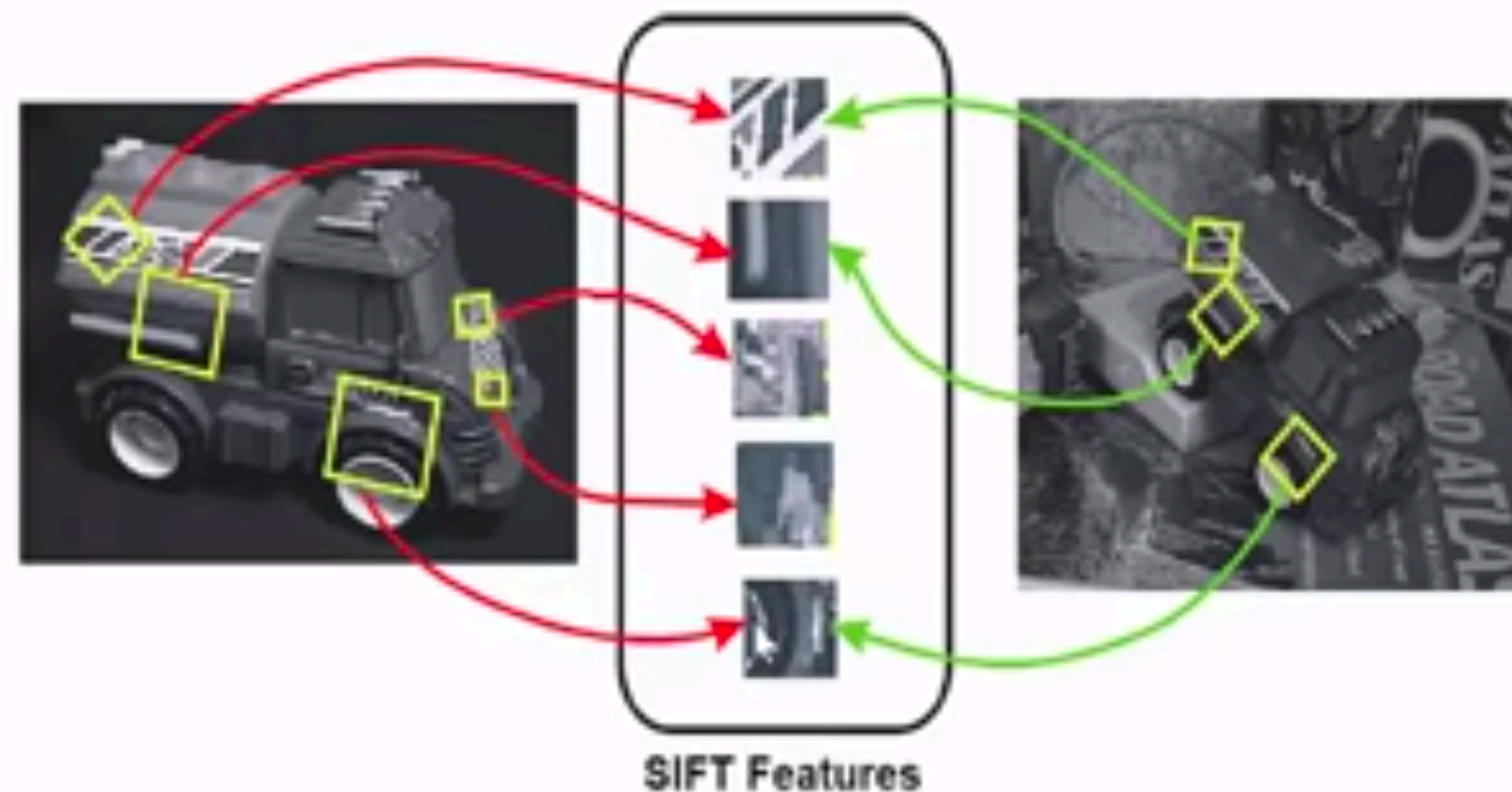
- The more data you use to train your model the better.



# CLASSIC TECHNIQUES

## Scale-Invariant Feature Transform (SIFT)

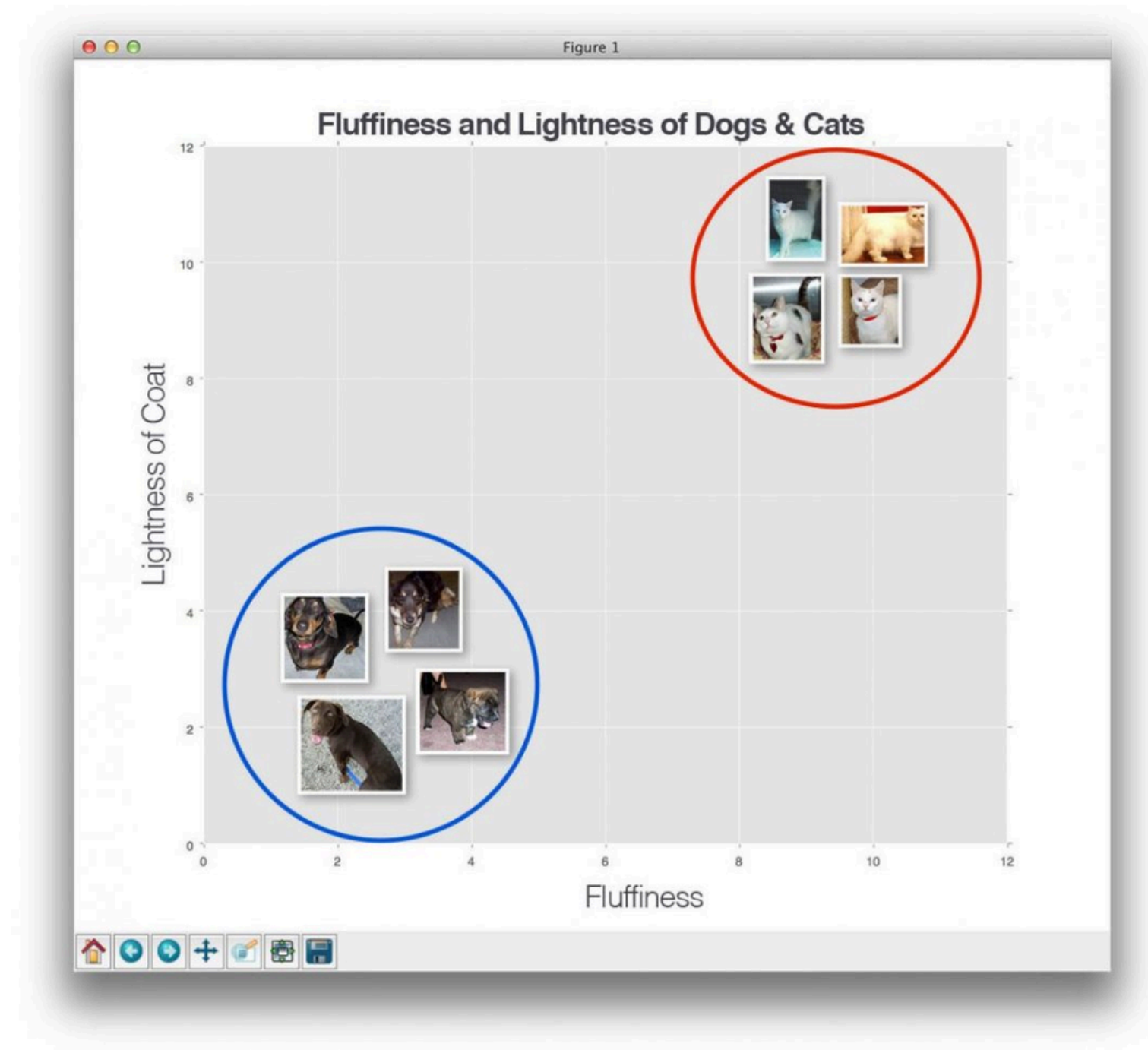
- SIFT is a traditional algorithm used for object detection and recognition.
- It identifies and matches keypoints within images, which are distinctive features.
- These keypoints are relatively stable even under changes in scale, rotation, or illumination



# CLASSIC TECHNIQUES

## K-Nearest Neighbors (k-NN)

- k-NN is a classification algorithm used to categorize new data points.
- It compares the features of a new image to existing labeled data points.
- The k-NN algorithm classifies the new image based on the majority vote of its k nearest neighbors.



*That is all folks!!*

**NOW BUCKLE UP FOR HANDS-ON MODULE**

534547657568  
675756756756  
7867876889  
7878678789789  
87798797  
7867886976  
78979878978



# COMMON LIBRARIES

## **mediapipe**

Framework by Google for building multimodal (video, audio, etc.) machine learning pipelines, used here for hand detection.

## **numpy**

Library for numerical operations on arrays, supporting large, multi-dimensional arrays and matrices.

## **tensorflow**

Open-source library for machine learning and artificial intelligence, used here for loading and using neural network models.

## **tensorflow.keras.models**

Submodule of TensorFlow for building and training deep learning models, used here for loading the pre-trained model.

# COMMON LIBRARIES

## **cv2**

OpenCV library used for computer vision tasks, including image and video processing.

## **dlib**

Library for machine learning and data analysis, used for face detection and landmark prediction.

## **imutils**

Convenience functions to make basic image processing functions such as translation, rotation, resizing, and skeletonization easier with OpenCV and Python.

## **math**

Python standard library for mathematical functions.

# CONTENTS

## Module # 2 - Hands-on Projects

### **Introduction to environment**

- Familiarization with Python Libraries

### **Project # 1 Hand Gesture Recognition**

- Familiarization with the underlying principles
- Familiarization with Program and Algorithm
- Discussion and Demonstration

### **Project # 2 Gaze Detection**

- Familiarization with the underlying principles
- Familiarization with Program and Algorithm
- Discussion and Demonstration



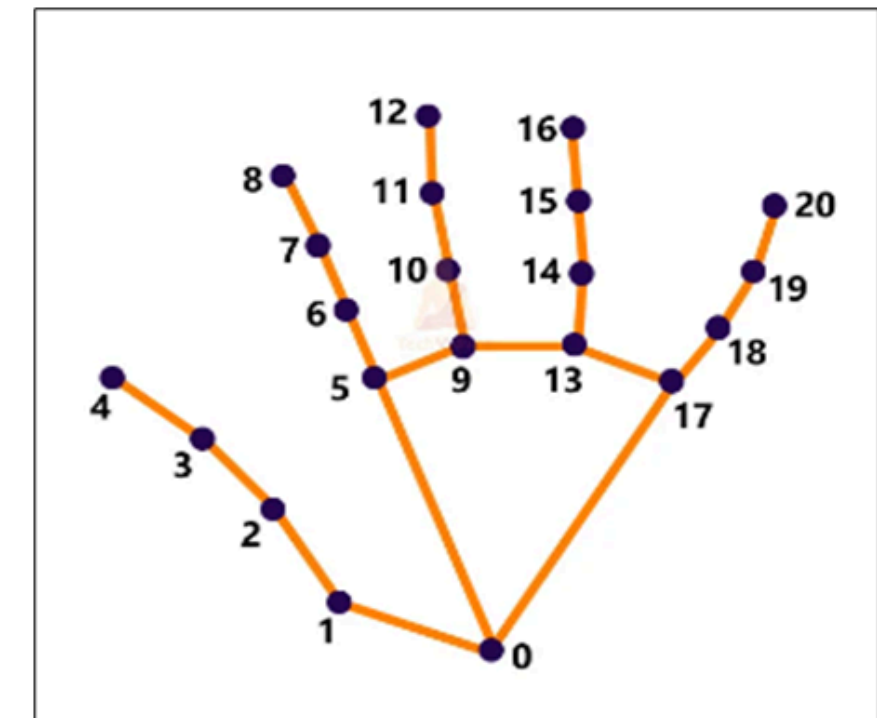


# HAND GESTURE RECOGNITION

## Summary

This code captures video from a webcam, detects hand landmarks using MediaPipe, and classifies hand gestures using a TensorFlow model.

- 1. Hand Detection:** Detects hand landmarks in each frame using MediaPipe.
- 2. Gesture Recognition:** Classifies the detected hand landmarks into gestures using a pre-trained neural network model.
- 3. Real-Time Display:** Displays the predicted gesture on the video frame in real-time and provides an interface to quit the application using the 'q' key.



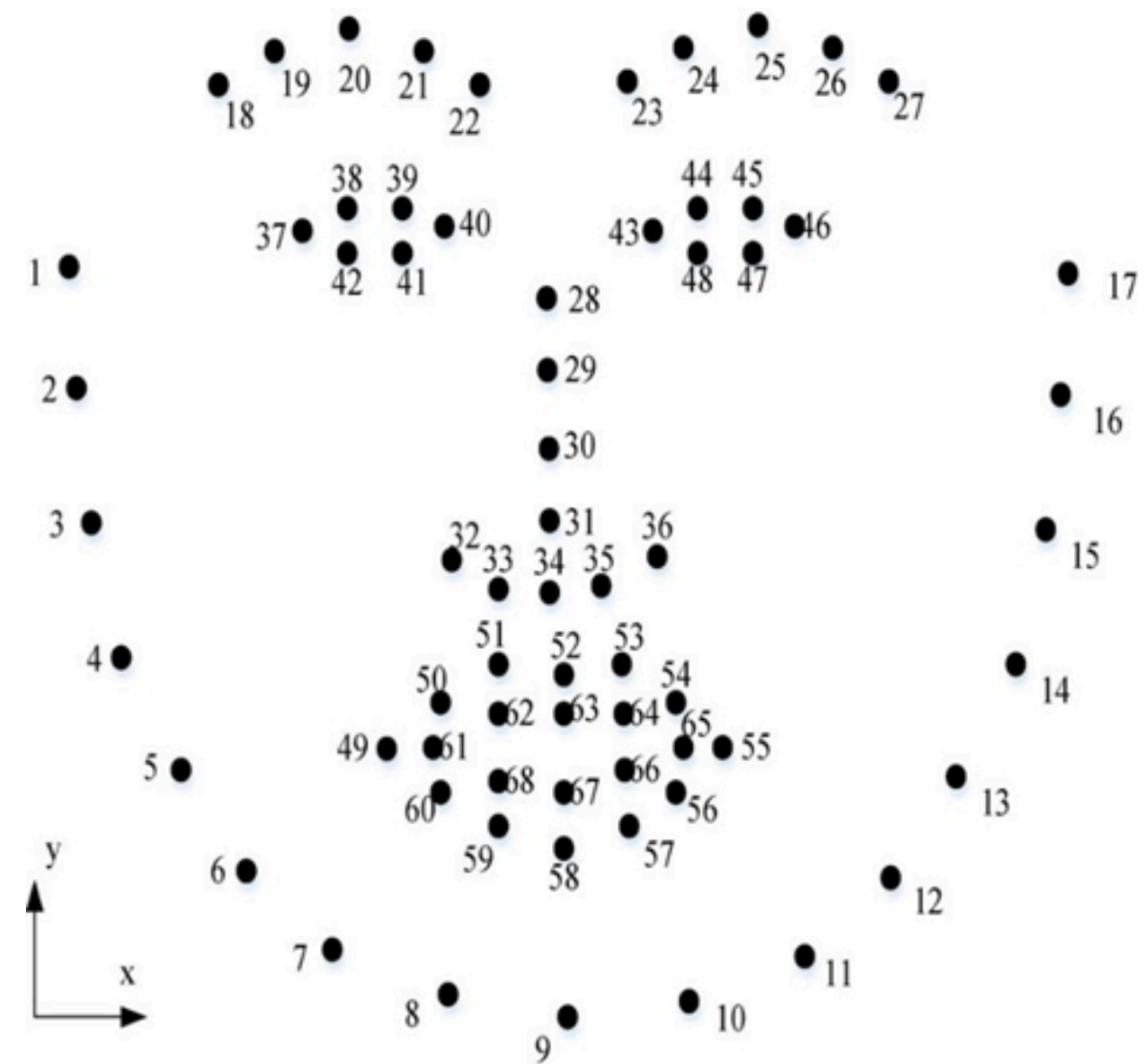
0. WRIST	11. MIDDLE_FINGER_DIP
1. THUMB_CMC	12. MIDDLE_FINGER_TIP
2. THUMB_MCP	13. RING_FINGER_MCP
3. THUMB_IP	14. RING_FINGER_PIP
4. THUMB_TIP	15. RING_FINGER_DIP
5. INDEX_FINGER_MCP	16. RING_FINGER_TIP
6. INDEX_FINGER_PIP	17. PINKY_MCP
7. INDEX_FINGER_DIP	18. PINKY_PIP
8. INDEX_FINGER_TIP	19. PINKY_DIP
9. MIDDLE_FINGER_MCP	20. PINKY_TIP
10. MIDDLE_FINGER_PIP	

# GAZE DETECTION

## Summary

This code captures video from a webcam, processes each frame to detect faces and their landmarks, and tracks the eyes' positions. It uses functions from the **Detector** module for:

1. **Face Detection:** Identifies faces in each frame.
2. **Landmark Detection:** Detects facial landmarks to identify key points like eyes.
3. **Eye Tracking:** Tracks the eyes and determines their position (right, center, left).
4. **Determine Eye Position:** Classify eye position based on pixel analysis.



**THANK YOU**

Any questions, feel free to reach out.

