Unit 2

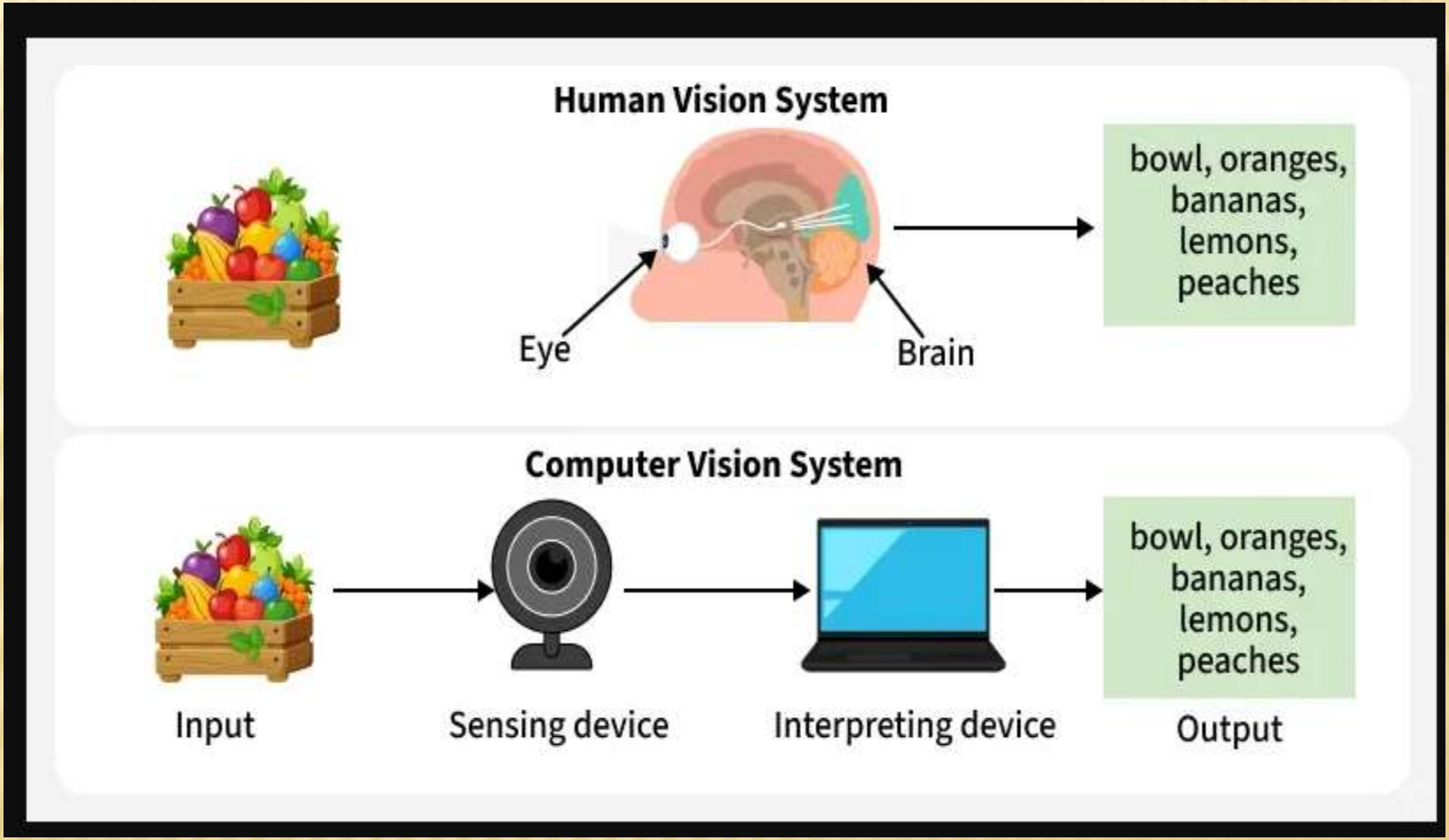
COMPUTER VISION

INDEX

- Introduction
- Basic techniques of Computer Vision
- Applications of Computer Vision
- Computer Vision Libraries and Tools
- Ethical Considerations in Computer Vision

INTRODUCTION

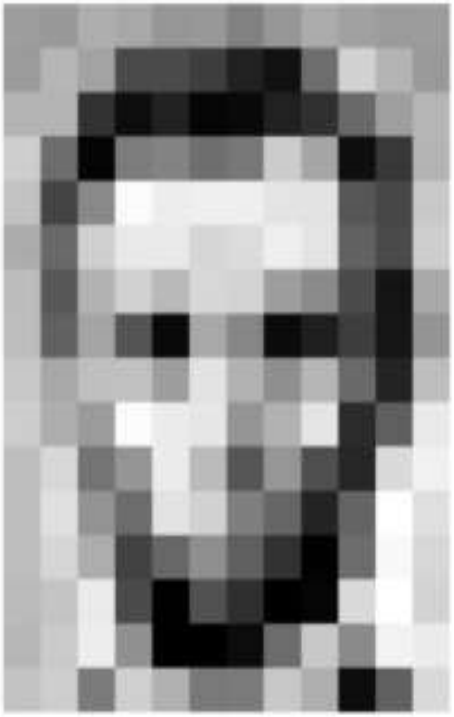
- Computer Vision (CV) is a branch of Artificial Intelligence (AI) that helps computers to interpret and understand visual information much like humans.
- beginners and experienced professionals and covers key concepts such as Image Processing, Feature Extraction, Object Detection, Image Segmentation and other core techniques in CV.

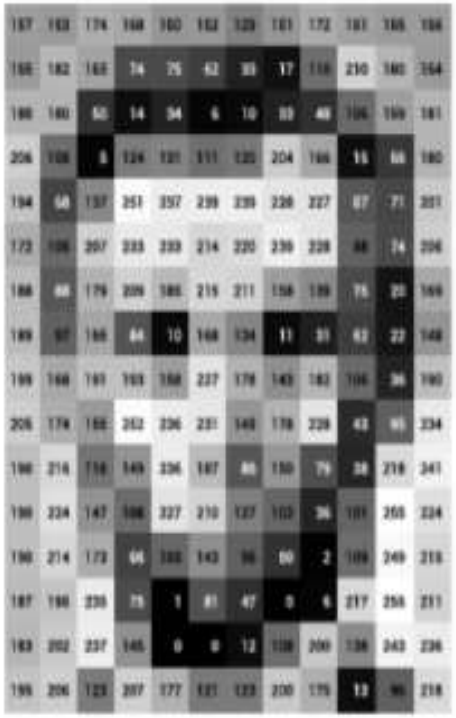

Human Vision System
bowl, oranges, bananas, lemons, peaches
Eye
Brain
Computer Vision System
bowl, oranges, bananas, lemons, peaches
Input
Sensing device
Interpreting device
Output

BASIC TECHNIQUES OF COMPUTER VISION

- the basics of computer vision seem easy, processing and understanding an image via machine vision are quite difficult. Here's why—
- An image consists of several pixels, with a pixel being the smallest quanta in which the image can be divided into.
- Computers process images in the form of an array of pixels, where each pixel has a set of values, representing the presence and intensity of the three primary colors: red, green, and blue.
- All pixels come together to form a digital image.

THIS IS HOW COMPUTER “SEES” IMAGE





157	153	174	168	150	152	129	151	172	161	155	156
155	182	163	74	75	62	33	17	110	210	180	154
180	180	50	14	34	6	10	33	48	106	159	181
206	109	5	124	131	111	120	204	166	15	56	180
194	68	137	251	237	239	239	228	227	87	71	201
172	105	207	233	233	214	220	239	228	98	74	206
188	88	179	209	185	215	211	158	139	75	20	169
189	97	165	84	10	168	134	11	31	62	22	148
199	168	191	193	158	227	178	143	182	106	36	190
205	174	155	252	236	231	149	178	228	43	95	234
190	216	116	149	236	187	86	150	79	38	218	241
190	224	147	108	227	210	127	102	36	101	255	224
190	214	173	66	153	143	96	50	2	109	249	215
187	196	235	75	1	81	47	0	6	217	255	211
183	202	237	145	0	0	12	108	200	138	243	236
195	206	123	207	177	121	123	200	175	13	96	218

BASIC TECHNIQUES OF COMPUTER VISION

- The values represent the pixel values at the particular coordinates in the image, with 255 representing a complete white point and 0 representing a complete dark point.

BASIC TECHNIQUES OF COMPUTER VISION

- Some operations commonly used in computer vision based on a Deep Learning perspective include:
- Convolution: Convolution in computer vision is an operation in which a learnable kernel is “convolved” with the image. In other words—the kernel is slided across the image pixel by pixel, and an element-wise multiplication is performed between the kernel and the image at every pixel group.
- Pooling: Pooling is an operation used to reduce the dimensions of an image by performing operations at a pixel level. A pooling kernel slides across the image, and only one pixel from the corresponding pixel group is selected for further processing, thus reducing the image size., eg., Max Pooling, Average Pooling.
- Non-Linear Activations: Non-Linear activations introduce non-linearity to the neural network, thereby allowing the stacking of multiple convolutions and pooling blocks to increase model depth.

FACE AND PERSON RECOGNITION

- Facial Recognition is a subpart of object detection where the primary object being detected is the human face.
- While similar to object detection as a task, where features are detected and localized, facial recognition performs not only detection, but also recognition of the detected face.

FACE AND PERSON RECOGNITION

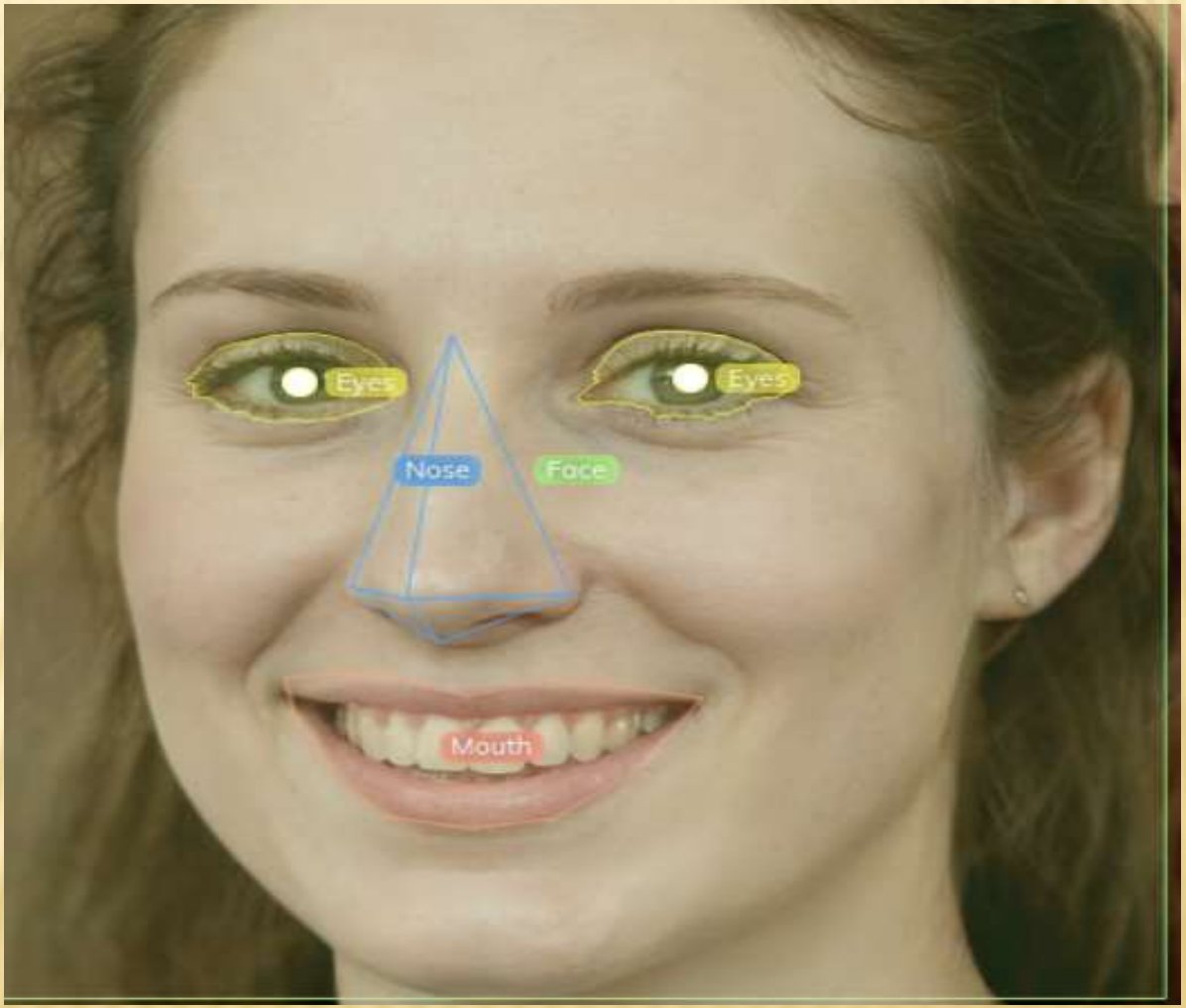


IMAGE RESTORATION

- image Restoration refers to the restoration or the reconstruction of faded and old image hard copies that have been captured and stored in an improper manner, leading to loss of quality of the image.

IMAGE RESTORATION

FEATURE MATCHING

- The applications of feature matching are found in computer vision tasks like object identification and camera calibration. The task of feature matching is generally performed in the following order:
- Detection of features: Detection of regions of interest is generally performed by Image Processing algorithms like Harris Corner Detection
- Formation of local descriptors: After features are detected, the region surrounding each keypoint is captured and the local descriptors of these regions of interest are obtained. A local descriptor is the representation of a point's local neighborhood and thus can be helpful for feature matching.
- Feature matching: The features and their local descriptors are matched in the corresponding images to complete the feature matching step.


inliers: 687/734

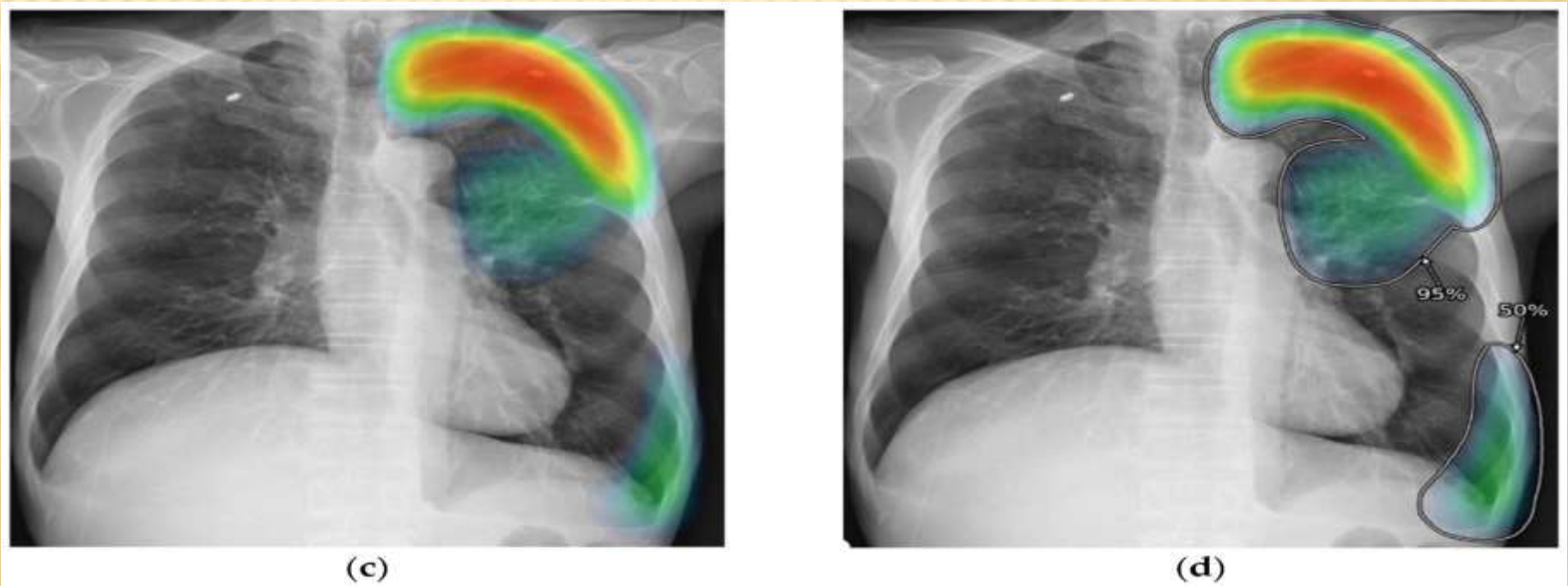
Application of computer vision

1. Healthcare

- **Medical Imaging Analysis**: Detecting diseases in X-rays, MRIs, and CT scans (e.g., tumors, fractures).
- **Surgical Assistance**: Real-time guidance during surgery using visual data.
- **Skin Cancer Detection**: Using image classification to identify malignancies from skin images.

Application of computer vision

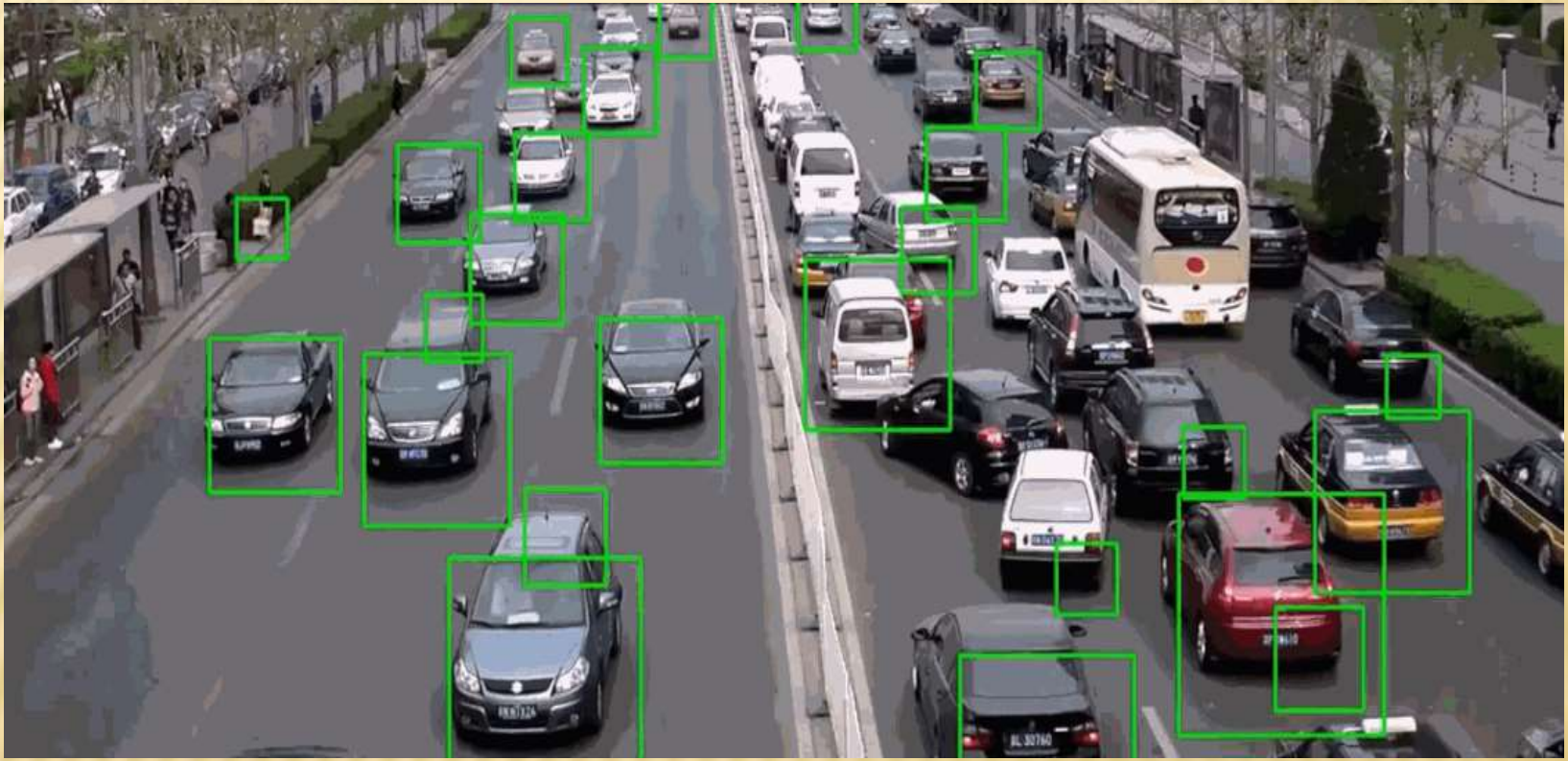
A chest X-ray of a pneumothorax case—AI overlays a heatmap (red-yellow) identifying air pocket region that corresponds with physician-confirmed abnormality



Application of computer vision

- **Autonomous Vehicles**
- **Self-Driving Cars:** Computer vision is used for object detection, lane detection, traffic sign recognition, pedestrian tracking, and obstacle avoidance.
- **Drone Navigation:** Drones use CV to detect and avoid obstacles in real-time while navigating.

Image cars count on road



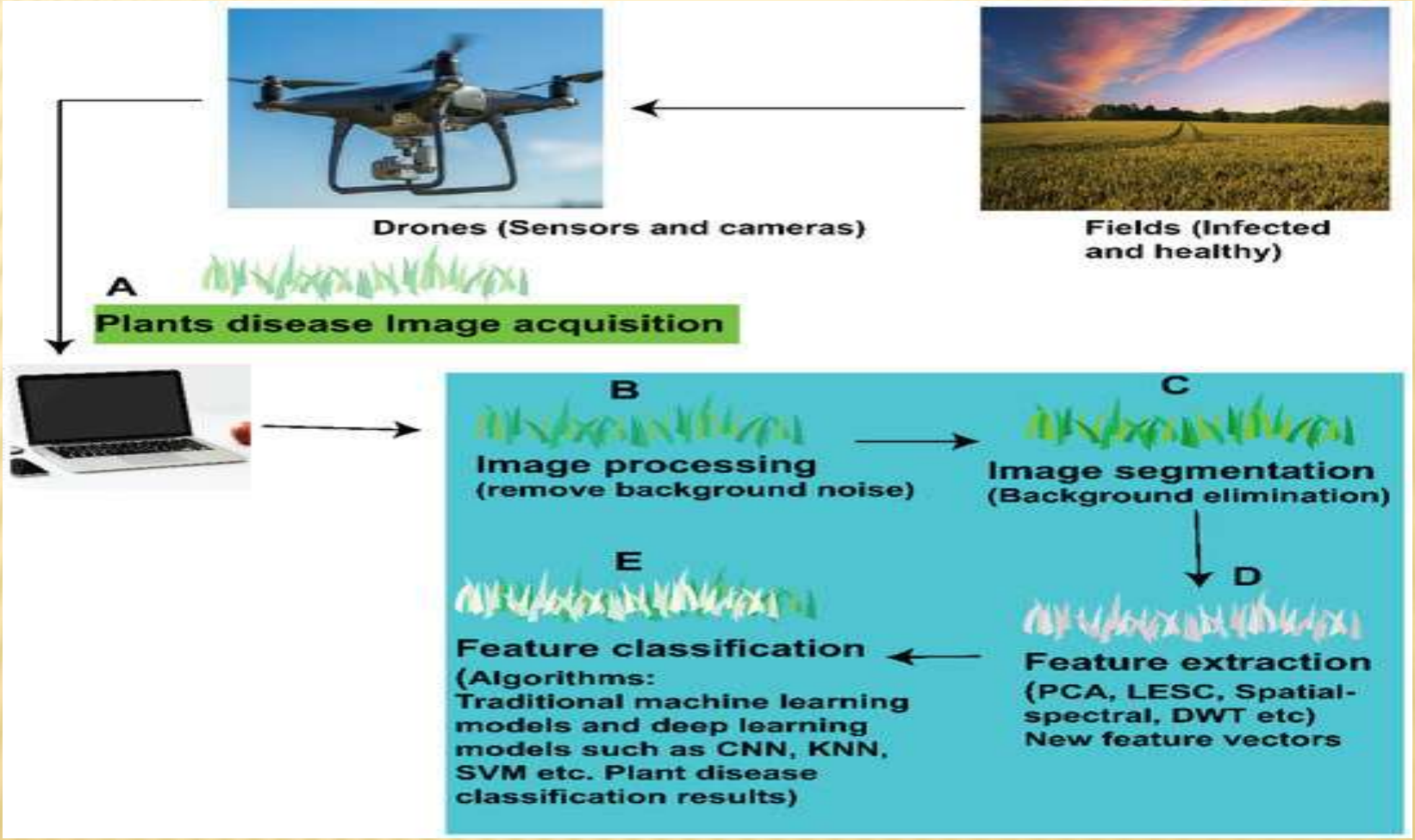
Application of computer vision

- **Retail and E-commerce**
- **Application:** Enhances the shopping experience through image-based search, recommendation systems, and even checkout-less stores.
- **Example:** Amazon Go stores use computer vision to track what customers pick up, allowing them to leave without manually checking out.


COLD DRINKS
ENTRÉES & BOWLS
SALADS
WRAPS
SANDWICHES
BREAKFAST

Application of computer vision

- **Retail and E-commerce**
- **Application:** Enhances the shopping experience through image-based search, recommendation systems, and even checkout-less stores.
- **Example:** Amazon Go stores use computer vision to track what customers pick up, allowing them to leave without manually checking out.


Drones (Sensors and cameras)
Fields (Infected and healthy)
A
Plants disease Image acquisition
B
Image processing
(remove background noise)
C
Image segmentation
(Background elimination)
D
Feature extraction
(PCA, LESC, Spatial-spectral, DWT etc)
New feature vectors
E
Feature classification
(Algorithms:
Traditional machine learning models and deep learning models such as CNN, KNN, SVM etc. Plant disease classification results)

Computer Vision Libraries and Tools

- **1. OpenCV (Open Source Computer Vision Library)**
- **Description:** OpenCV is one of the most popular and comprehensive open-source libraries for computer vision tasks. It provides tools for image processing, object detection, face recognition, and real-time video processing.
- **Languages Supported:** C++, Python, Java, and others.
- **Key Features:** Image filtering, feature detection, image transformations, machine learning integration, real-time video analysis.

Computer Vision Libraries and Tools

- **2. TensorFlow & TensorFlow.js**
- **Description:** TensorFlow, developed by Google, is a popular machine learning framework that also has strong support for computer vision tasks. TensorFlow.js brings machine learning to JavaScript for real-time computer vision in the browser.
- **Languages Supported:** Python, JavaScript.
- **Key Features:** Object detection, image segmentation, neural networks for visual tasks, support for deep learning.

Computer Vision Libraries and Tools

- **3. PyTorch**
- **Description:** PyTorch is a deep learning library that is widely used for computer vision tasks. It's known for its flexibility, ease of use, and support for dynamic computation graphs.
- **Languages Supported:** Python.
- **Key Features:** Deep learning for vision tasks like image classification, segmentation, and object detection. Popular models include ResNet, etc.

Computer Vision Libraries and Tools

- **Keras**
- **Description:** Keras is a high-level neural networks API that runs on top of TensorFlow, making it easier to develop deep learning models for computer vision tasks.
- **Languages Supported:** Python.
- **Key Features:** Simplified implementation of deep learning models for image classification, object detection, and segmentation.

Ethical consideration in CV

Ethical Concern	Description	Example
1. Privacy Invasion	CV often captures images/video in public/private spaces without consent.	CCTV systems in public spaces or facial recognition in retail stores.
2. Bias and Discrimination	Training datasets may lack diversity, leading to biased outputs.	Face recognition works better on lighter skin tones than darker ones.
3. Consent and Data Use	Individuals are often unaware their images are being used or analyzed.	Using social media photos to train facial recognition algorithms.
4. Misuse and Dual-Use	CV can be used for unethical or harmful purposes.	Military drones using CV for autonomous targeting.
5. Lack of Transparency	CV systems (especially deep learning) are often “black boxes.”	Inability to explain why an algorithm flagged a person as suspicious.
6. Accountability	It’s unclear who is liable when CV systems fail or cause harm.	Who is responsible if a self-driving car hits a pedestrian?
7. Deepfakes & Misinformation	CV enables creation of fake videos/images that can deceive and manipulate.	Political deepfakes spreading misinformation during elections.