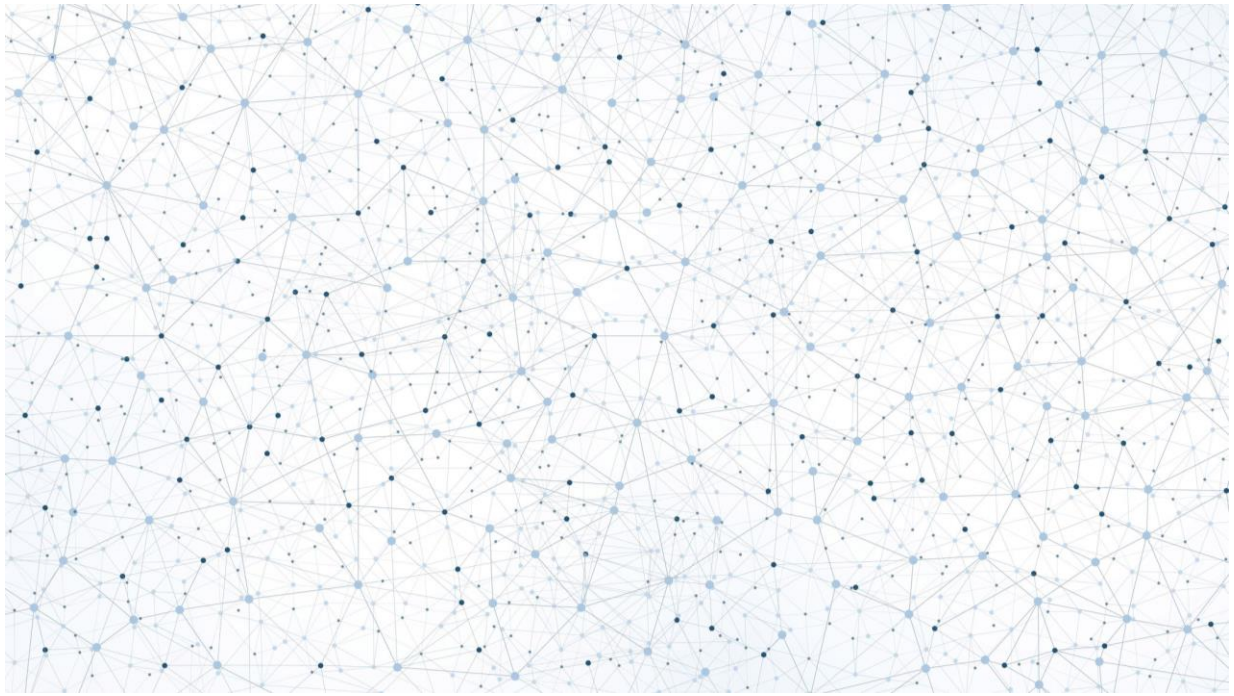
Mākslīgā intelekta risinājumi ražošanas procesu uzlabošanai - datorredze

2024

Valdis Saulespurens



RBS
RIGA BUSINESS SCHOOL
Riga Technical University

About Me

Valdis Saulespurens

30+ Years programming

Run retail and e-commerce store in 90s California,assisted with manufacturing startup

Lector RTU RBS

Researcher National Library of Latvia – OCR model improvement

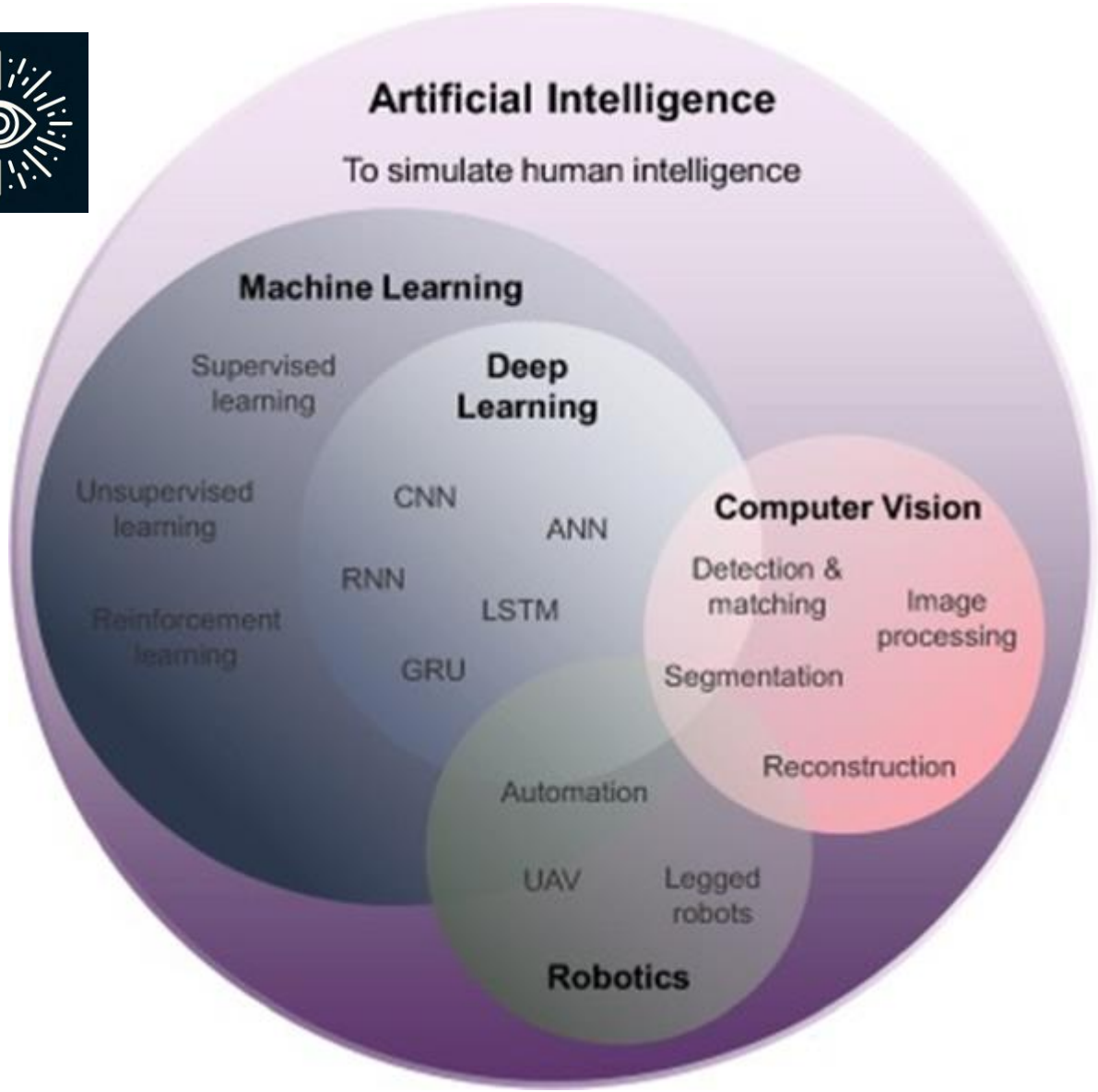
valdis.saulespurens@rtu.lv

Datorredze – Computer Vision

- Definition: A field of artificial intelligence(AI) that enables computers to interpret, analyze, and make decisions based on visual data from the real world, such as images and videos.
- Involves techniques for processing, understanding, and extracting meaningful information from visual inputs to automate tasks that typically require human vision.

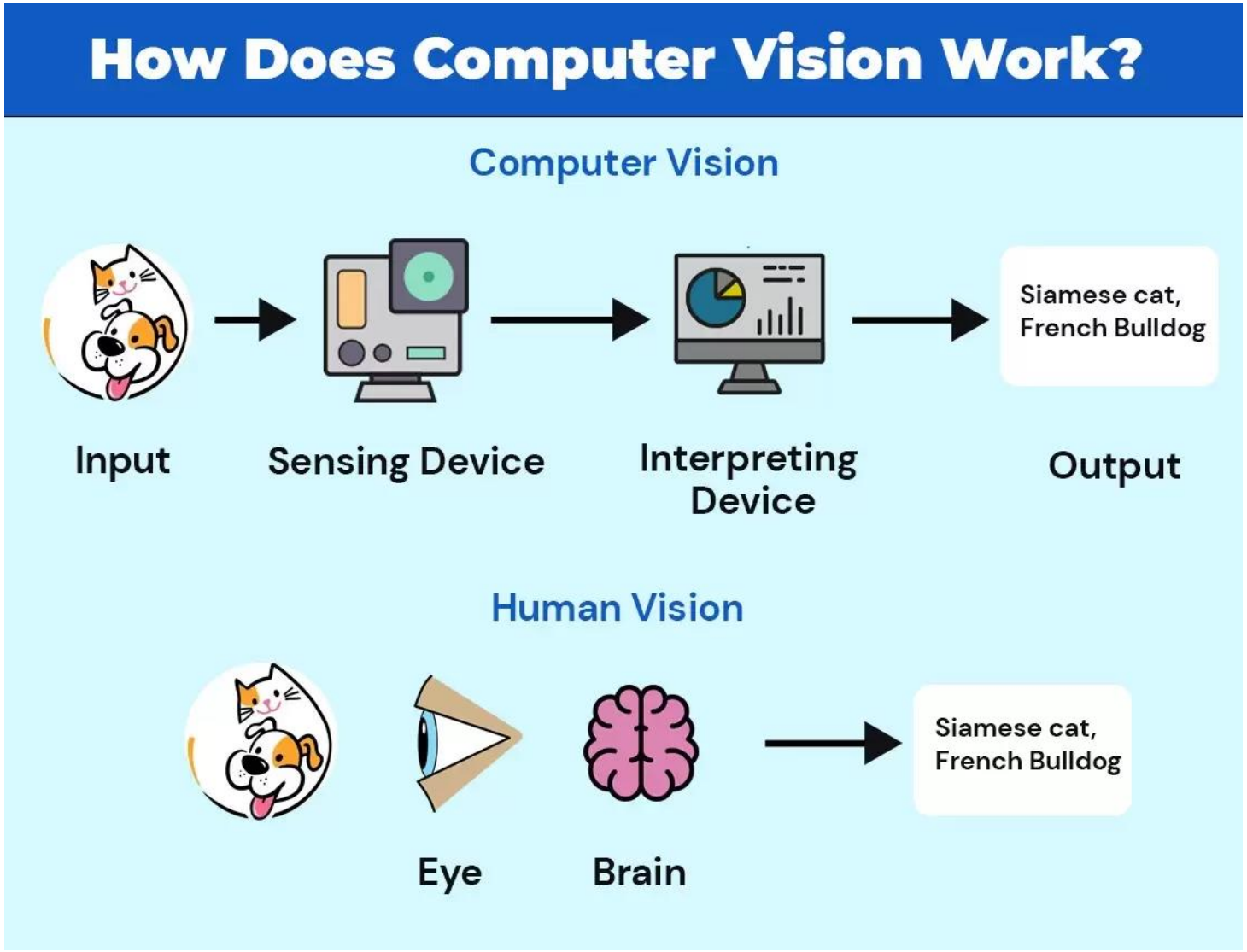
Computer Vision within AI field

- **Artificial Intelligence (AI)** - Simulation of human intelligence in machines.
- **Machine Learning (ML)** - Subfield of AI focused on algorithms that learn from data.
- **Deep Learning (DL)** - Subfield of ML using neural networks with multiple layers.
- **Robotics** - Design, construction, and operation of robots.
- **Machine Vision** - Technology enabling machines to interpret and process visual information. Primarily used in industrial and manufacturing settings.



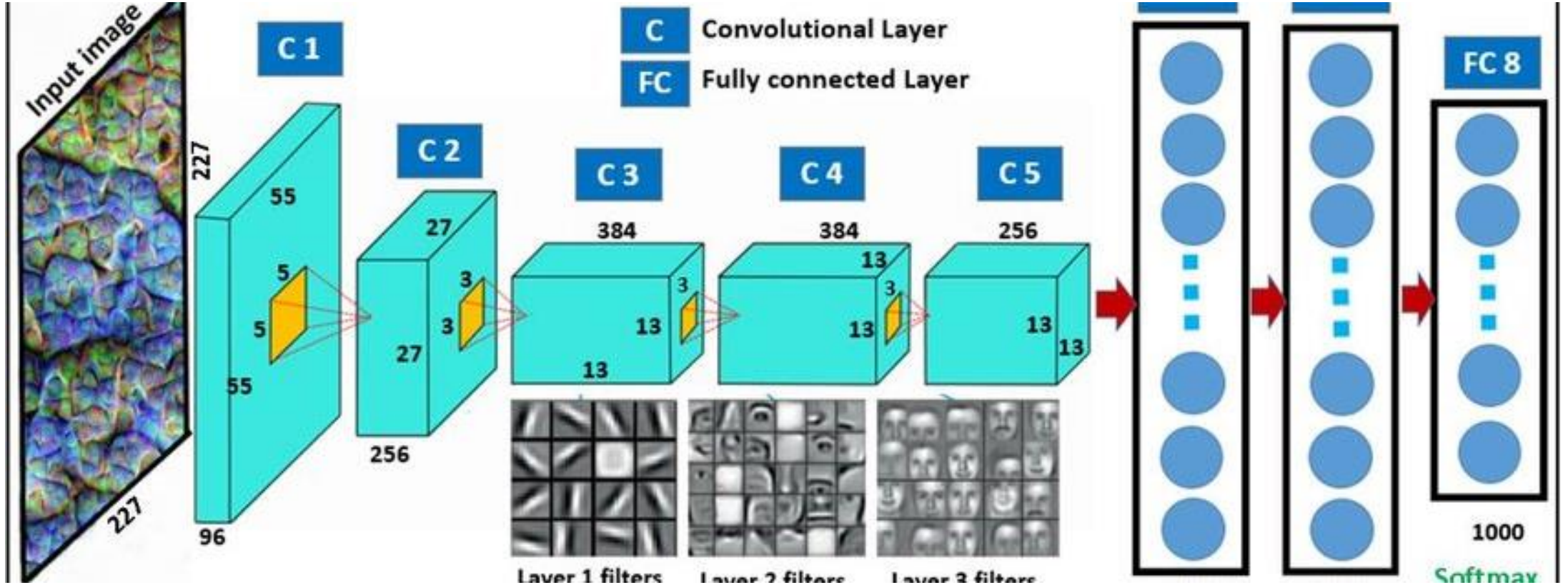
Img: https://medium.com/@su3230/computer-vision-cv-d204f81a9063

- **Input**
- Visual data such as images or videos.
- Acquired from real-world environments.
- Examples: photographs, live camera feeds, medical scans.
- **Sensing Device**
- Hardware capturing visual data.
- Converts real-world scenes into digital images.
- Examples: cameras, sensors, scanners.
- **Interpreting Device**
- Software or algorithms processing visual data.
- Extracts meaningful information from images.
- Examples: feature extraction, object detection, image classification.
- **Output**
- Processed results or actionable insights.
- Can be in various forms such as visualizations, decisions, or commands.
- Examples: labeled images, detected objects, autonomous vehicle navigation.



Early History of Computer Vision (1950s to 1970s)

- **First Scanned Image(NIST 1957) – Russell A. Kirsch , others**
- 1957 – D.H. Hubel and Wiesel – how do cats see? Neural basis of vision, simple and complex cells. (Nobel 1981)
- 1966 – Marvin Minsky (MIT) – Summer of Computer Vision – Block World
- 1970s – Early OCR (Optical Character Recognition) systems. – special fonts, still used in some fields such as bank checks
- David Marr's Theory of Edge Detection (1976) – importance of edges

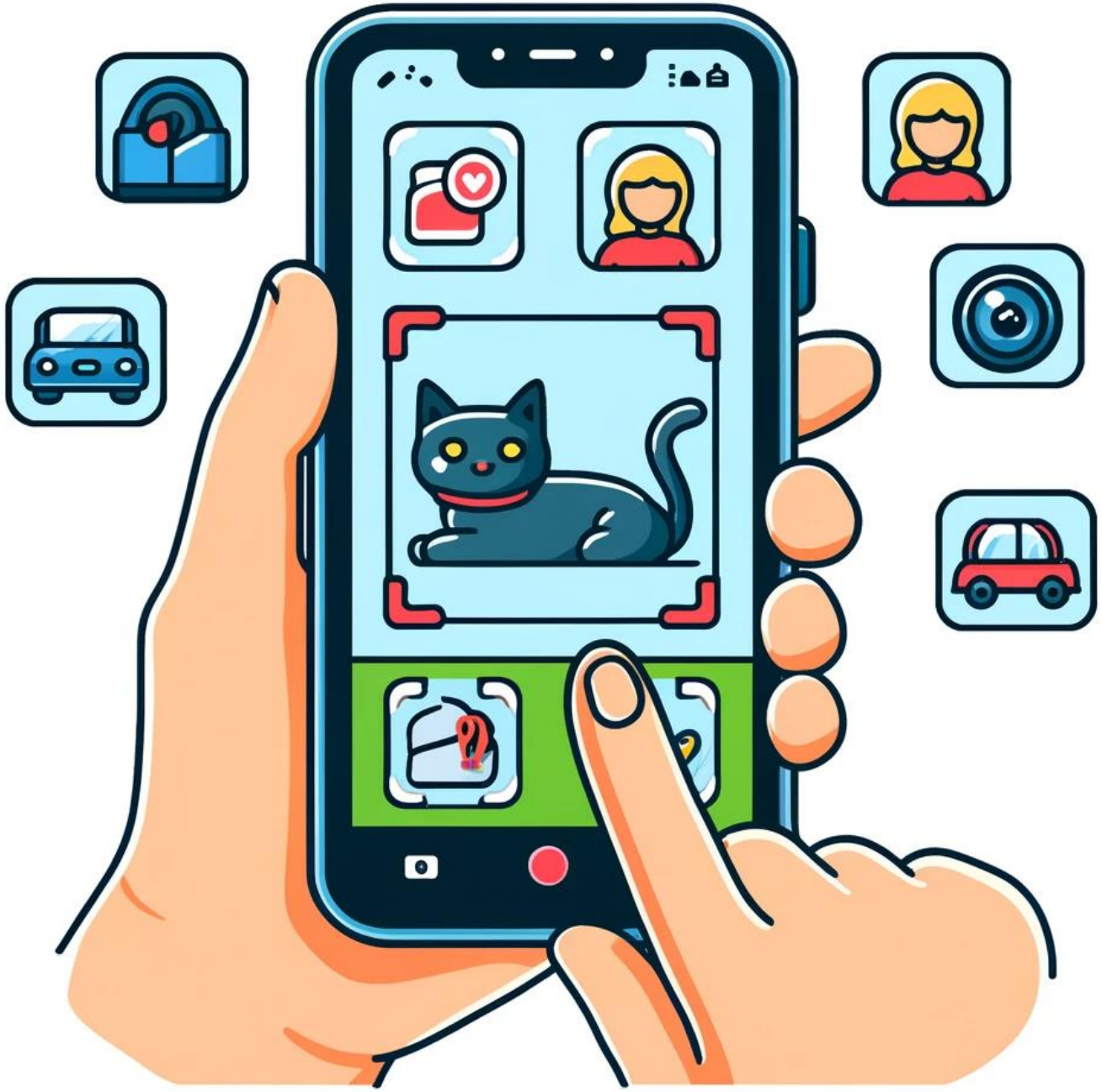


1980s to 2012
From Heuristics to Deep Learning

- **1980s –** Feature-Based Models (edges,corners, textures)
- 1990s – Early Experiments with Neural Networks and Deep Learning, backpropagation, 3D vision, advances in Face Recognition
- 2000s – Boosting Algorithms, Large Scale Image Datasets
- 2012 – ImageNet dataset (big labeled dataset), GPU acceleration, **AlexNet** top accuracy in the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) 2012, bringing Convolutional Neural Network (CNN) Architecture to forefront

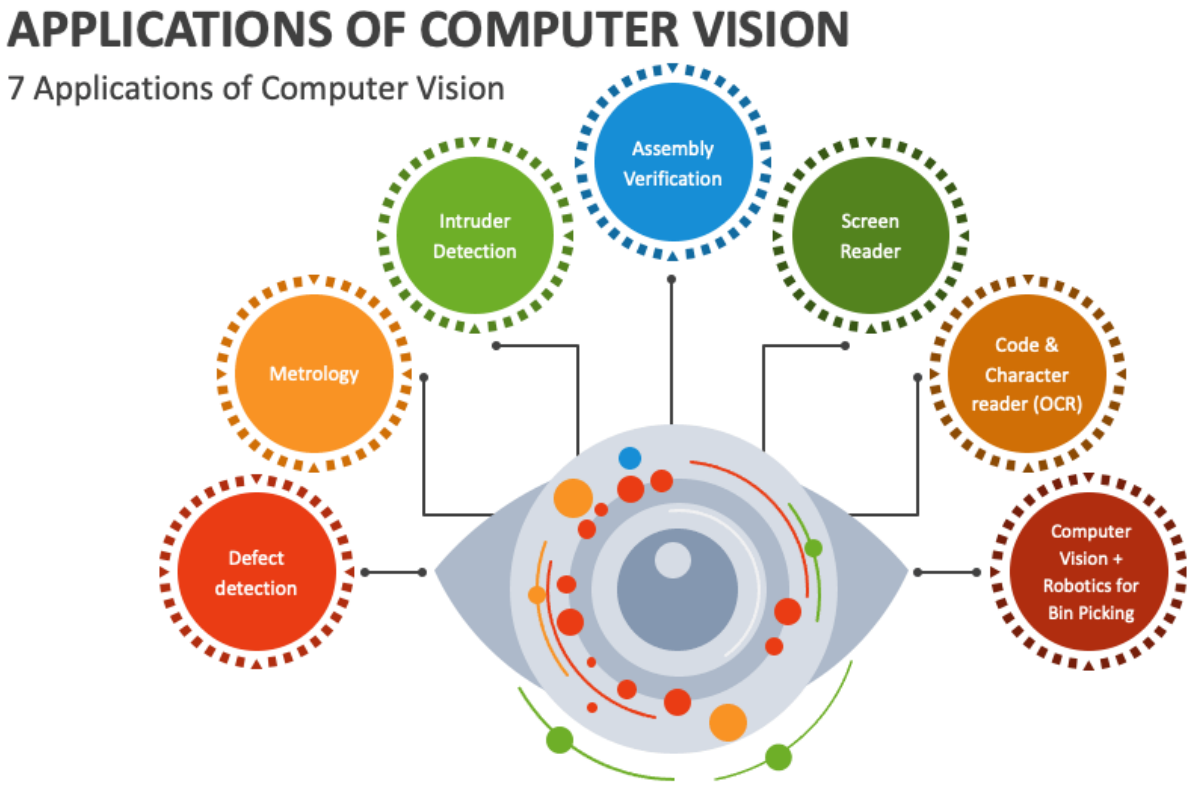
2012-Onwards

- Deeper Networks VGG(2013) Oxford, GoogleNet/Inception(2014), ResNet(2015)-Microsoft
- YOLO (You Only Look Once) Series (2016 onwards)
- Transformers in Vision (2017 onwards) – attention based models
- DALL-E and Generative Models (2021), Stable Diffussion, Midjourney, advances in GAN(Generative Adversarial Networks)
- Edge AI and On-Device Learning - complex vision models on edge devices like smartphones and IoT devices.



Subfields of Computer Vision

- Image Classification
- Face Recognition
- Optical Character Recognition (OCR)
- Object Detection
- Image Segmentation
- Activity Recognition
- 3D Vision
- Scene Understanding



Manufacturing use cases

- Automation and Robotics
- **Supply Chain Optimization**
- **Traceability**
- **Environmental Monitoring**

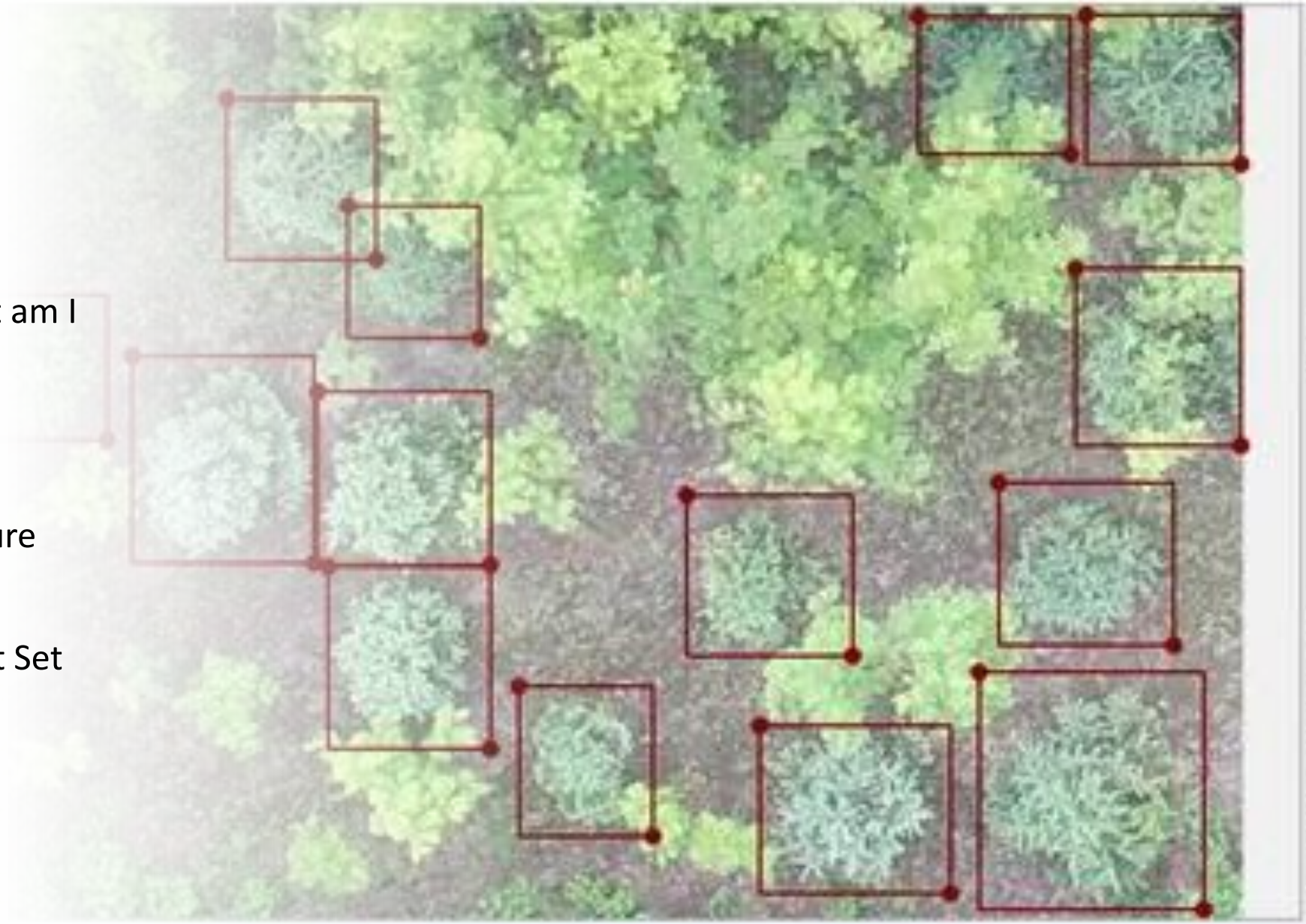
Src: viact.ai

Image Segmentation

- Segmentation – Object Detection
- Localization – Important Points
- Define Your Goals
- Gather and Prepare Data
- Preprocess Your Images
- Select a Segmentation Method (U-Net, Mask R-CNN others)
- Build and Train Your Segmentation Model
- Evaluate the Model
- Refine, Deploy, Monitor and Improve

Image Classification

- Define the Problem – what am I trying to classify
- Collect and Prepare Data
- Preprocess the Data
- Choose a Model Architecture
- Build and Train the Model
- Evaluate the Model on Test Set
- Fine-tune and Optimize
- Deploy the Model
- Monitor and Maintain



Src: https://www.mdpi.com/2076-3417/13/2/895

Optical Character Recognition

- Similar steps to General Classifier
- Well studied
- Rules based OCR overtaken by LTSM and other neural net architectures
- Business case for digitalization
- Open Source models such as Tesseract can achieve excellent results by performing supplemental training to existing models

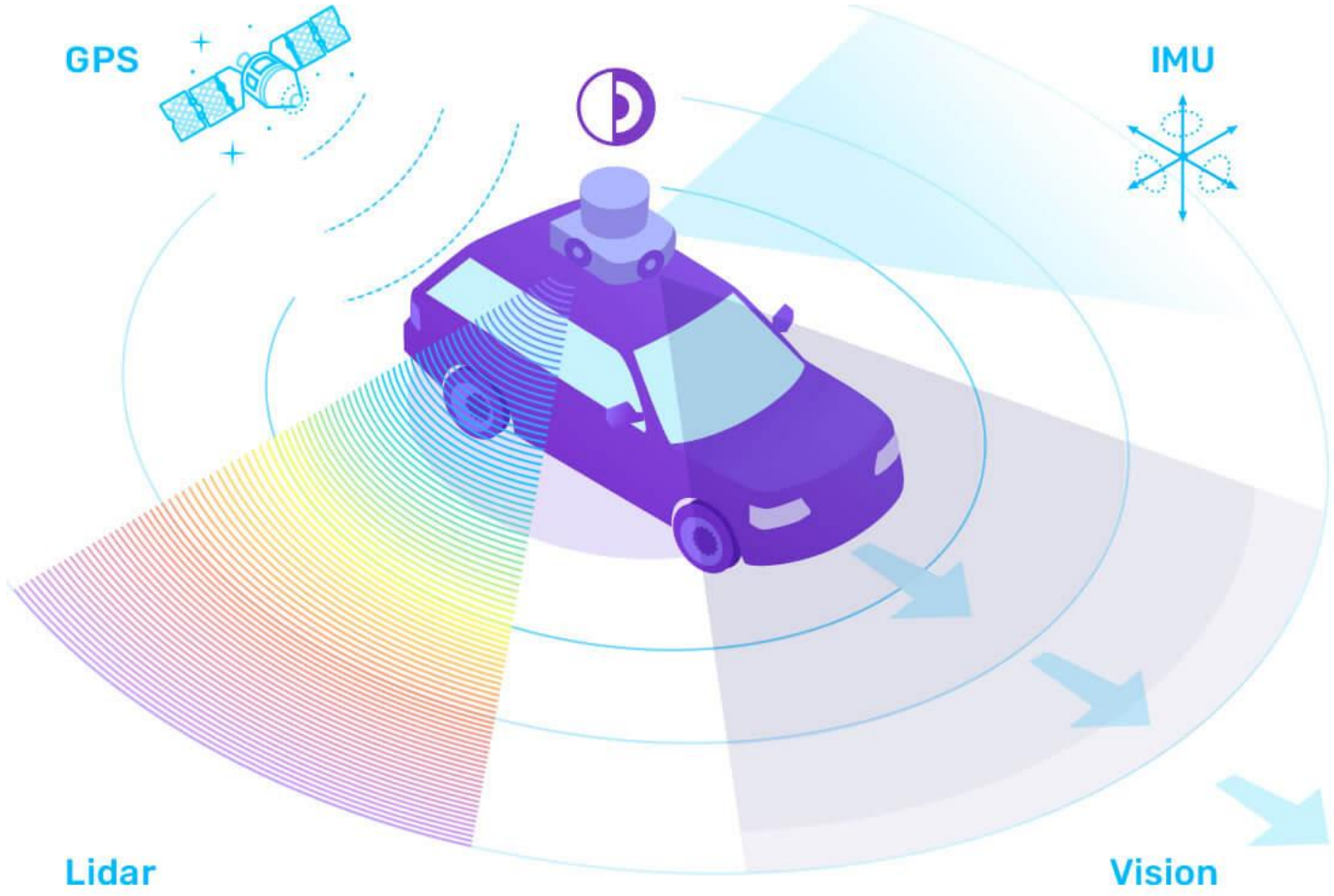
Face Recognition

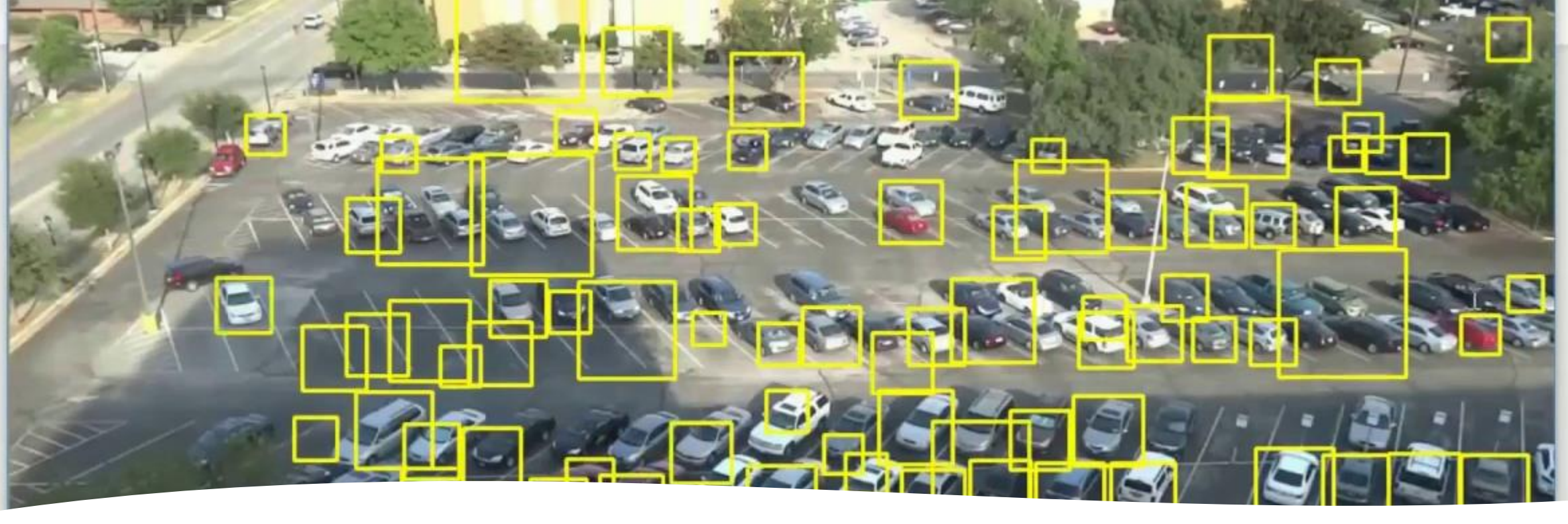
- Similar steps to General Classifier
- Use cases: Security, Health Monitoring
- Well studied
- Privacy and Ethics Concerns
- EU - GDPR

Simultaneous Localization and Mapping(SLAM)

- **Localization**: Determining the precise position and orientation of the observer within the environment.
- **Mapping**: Creating a representation (map) of the environment based on sensor data.

- Computer Vision part of SLAM
- Use in Robotics





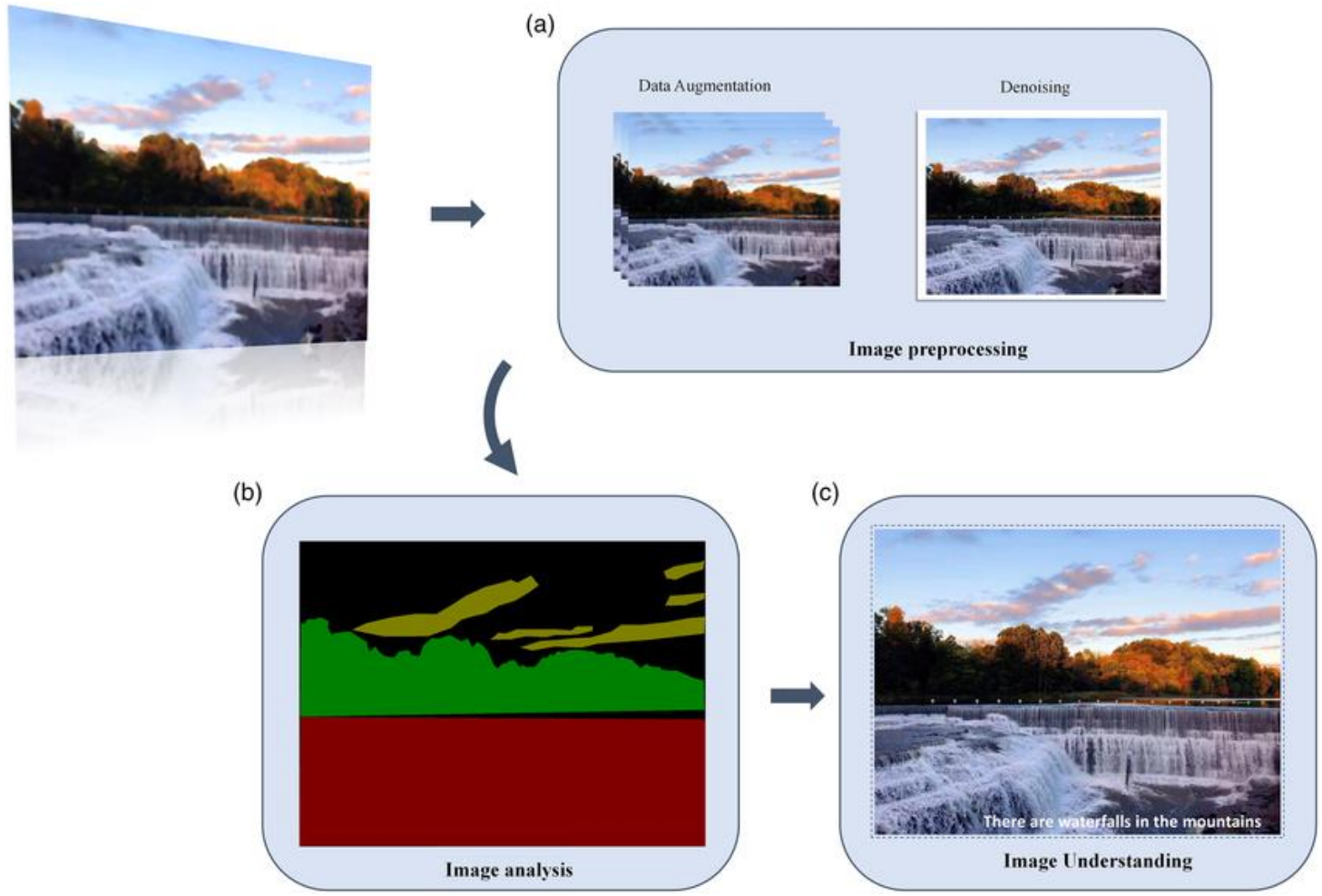
Real-Time Video Processing in Computer Vision

- **Real-time video processing involves analyzing and interpreting video data on-the-fly, enabling immediate decision-making and actions based on the visual input.**
- Frame-by-Frame Analysis
- Temporal Information Utilization – using multiple frames
- Applications in Security, Analytics, Autonomous Vehicles

Src: https://www.youtube.com/watch?v=y1M5dNkvCJc

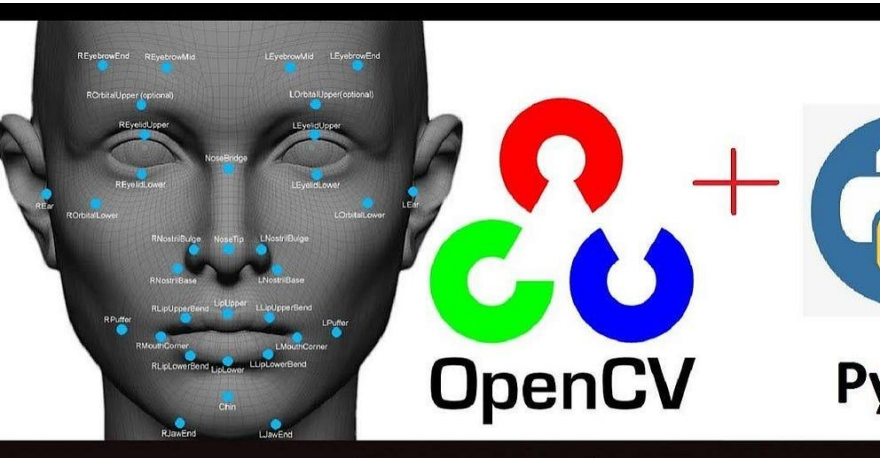
Preprocessing Pipeline Clean and Normalized Data is of Essence!

- Normalize: color, size
- Localize & crop
- Convert to intensity, normalize
- For Video might need to think about framerate



When starting CV project

Category	Question
Project Goals and Objectives	What are the primary goals of implementing a computer vision solution in our business or manufacturing process?
Use Case Definition	What specific use cases or problems are we aiming to address with computer vision?
Supervised vs. Unsupervised	Should we use supervised or unsupervised learning for our application?
Data Collection	What types of data do we need to collect, and how much data is required?
Data Cleaning and Preparation	What data cleaning and preprocessing steps are necessary?
Model Selection	What type of models should we use (e.g., CNN, RNN, GAN)?
Pretrained Models	Can we use pretrained models, or do we need to train our models from scratch?
Additional Training	If using pretrained models, do we need additional training or fine-tuning for our specific application?
Integration with AI Technologies	How will the computer vision solution integrate with other AI technologies (e.g., LLMs, IoT, RPA)?
Infrastructure Requirements	What infrastructure do we need to support computer vision implementation (e.g., hardware, software, cloud services)?
Scalability	How scalable is the proposed solution?
Cost and Budget	What is the estimated cost of implementing and maintaining the computer vision solution?
ROI and Benefits	What is the expected return on investment (ROI) and tangible benefits of the computer vision solution?
Data Privacy and Security	How will we ensure data privacy and security, especially if dealing with sensitive information?
Regulatory Compliance	Are there any industry-specific regulations or standards we need to comply with?
Team and Expertise	Do we have the necessary expertise in-house, or do we need to hire or partner with external experts?
Change Management	How will we manage the transition to a computer vision-based system?
Performance Metrics	What metrics will we use to measure the success of the computer vision implementation?
Maintenance and Updates	What is our plan for maintaining and updating the computer vision system post-implementation?
Pilot Testing	How will we conduct pilot tests to validate the computer vision solution before full-scale deployment?



Resources to play around:

- https://chatgpt.com/ - Now supports integration with DallE, also does object detection with OCR support
- https://gemini.google.com/ - playing catchup
- https://segment-anything.com/ - State of Art segmentation model developed by Meta
- https://roboflow.com/ - For developing your own models and supplemental training of existing ones
- https://docs.ultralytics.com/ - YOLO – You Only Look Once models
- https://opencv.org/ - Open Computer Vision library, integrates with various languages, tooling

Conclusions

CV works best along with other AI disciplines and human judgment

Questions

Thank You!

https://github.com/ValRCS/AI_Studio_24