Introduction to Computer Vision using Deep Learning

Vijayan K. Asari

Electrical and Computer Engineering

University of Dayton

Dayton, Ohio, USA

05 December 2022

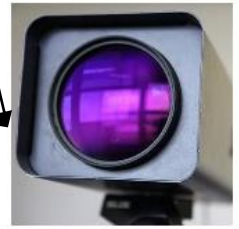
Visual Perception for Autonomy

Sensor Data Acquisition → Sensor Data Exploitation → Decision and Feedback

Canon Cameras

Hyperspectral Camera

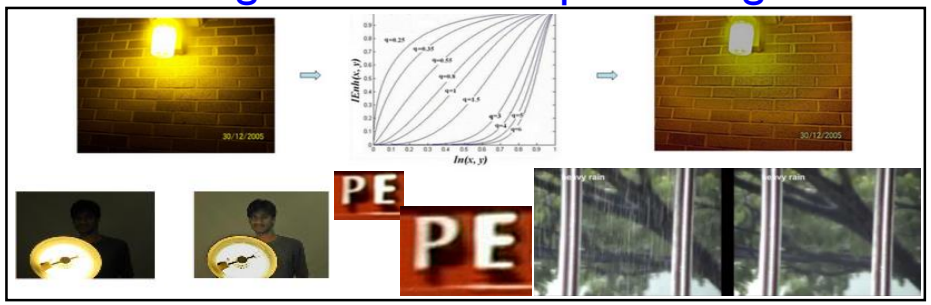
LiDAR



Long-range Cameras

Computer Vision and Scene Understanding

Image and Video Preprocessing

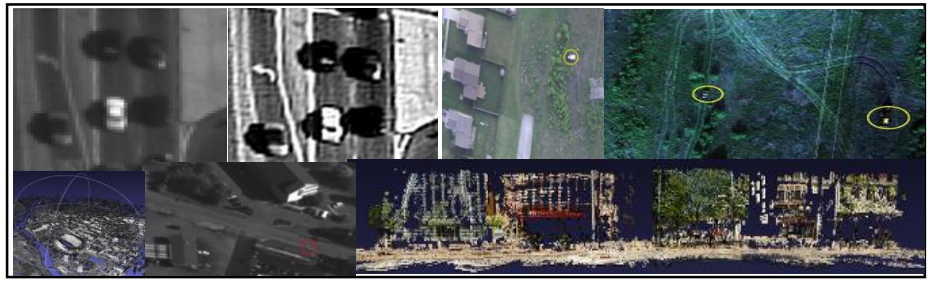


Enhancement
Super-resolution
Haze removal
Rain removal
Stabilization

Vision-Guided Robotics

Robotic navigation
Path planning
Object following
Behavior analysis
Threat analysis

Wide Area Surveillance

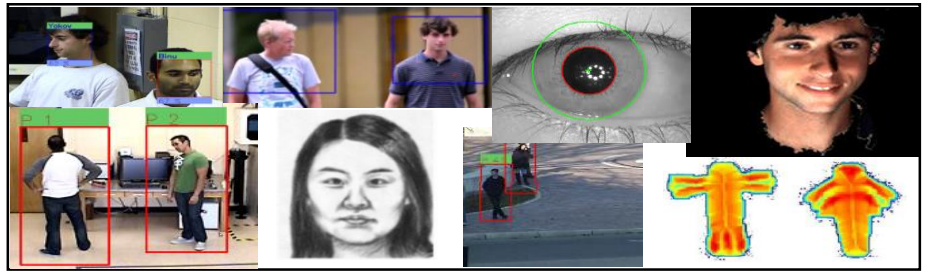


Object detection
Object recognition
Object tracking
3D reconstruction
Change detection

Perception Beyond Visible Spectrum

LiDAR data analysis
Hyperspectral data
IR/thermal data
Satellite imagery
EEG data analysis

Biometrics



Face recognition
Human action and activity recognition
Expression analysis
Emotion recognition

Brain Activity Analysis

Emotion recognition
Brain machine interface
Source localization
Neurofeedback

Computer Vision: Data Analysis and Processing

Distortion correction, enhancement, haze/fog removal, rain removal, stabilization.

Pre-processing

Spatial domain features
Spectral domain features
Statistical features
Phase features

Feature Extraction

Classification

Classical and machine learning based approaches.

Decision

Scene Analysis and Understanding

Wide Area Surveillance

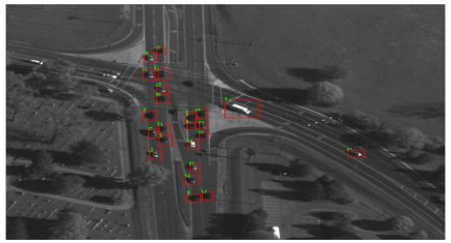
Biometric Identification

Human Activity Recognition

Brain Activity Analysis

Hardware Acceleration

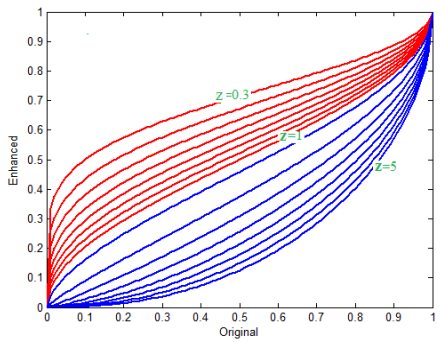
Data Acquisition



Face recognition.
Human action and activity recognition.
Object detection/tracking on WAMI data.
Emotion recognition by EEG analysis.
Brain machine interface.

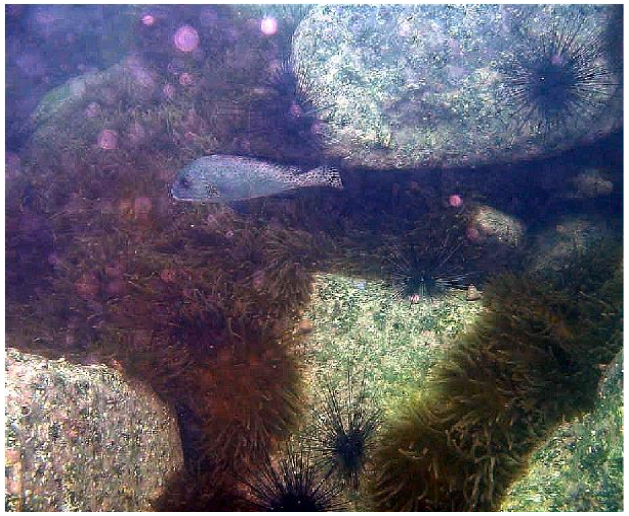
Enhancement of Low Lighting and Over Exposed Images

Input image

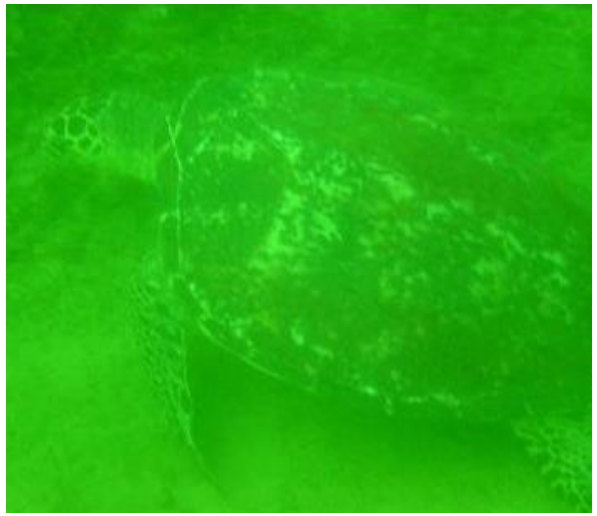


Enhanced image

Enhancement of Hazy/Foggy Images

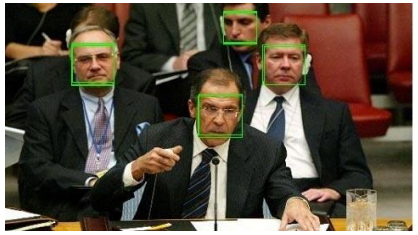





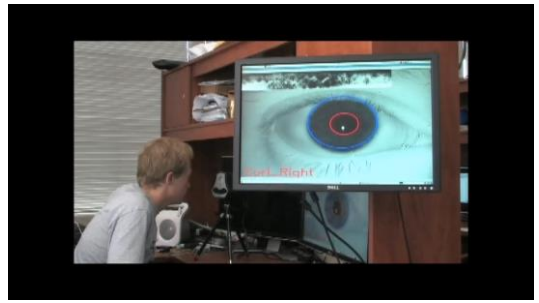

Biometric Applications

Face Detection

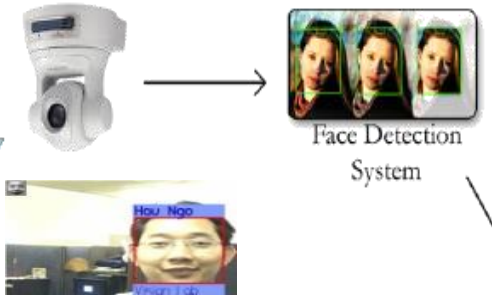


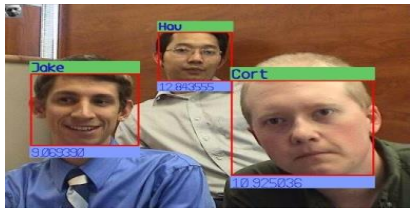
Action Recognition

Iris Recognition

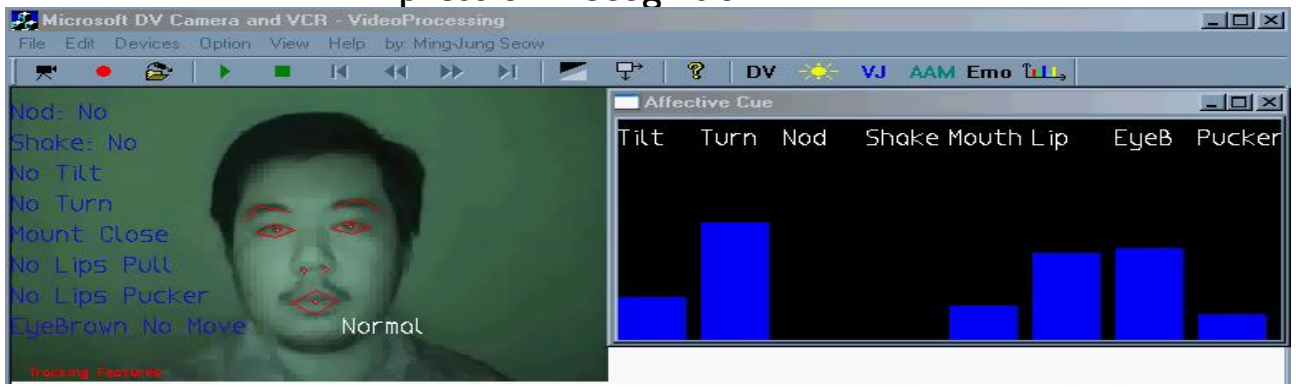


Face Recognition System

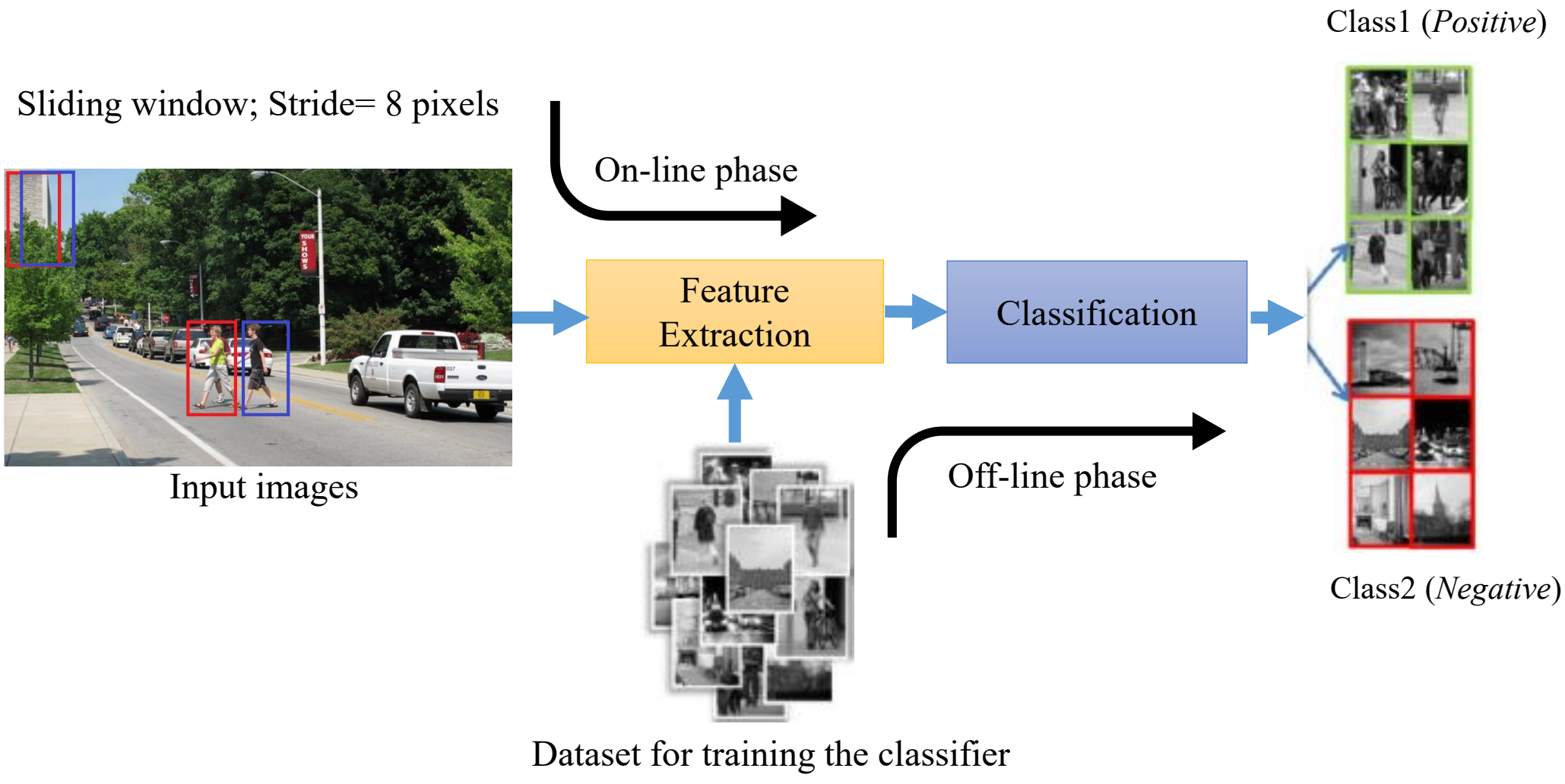




Expression Recognition

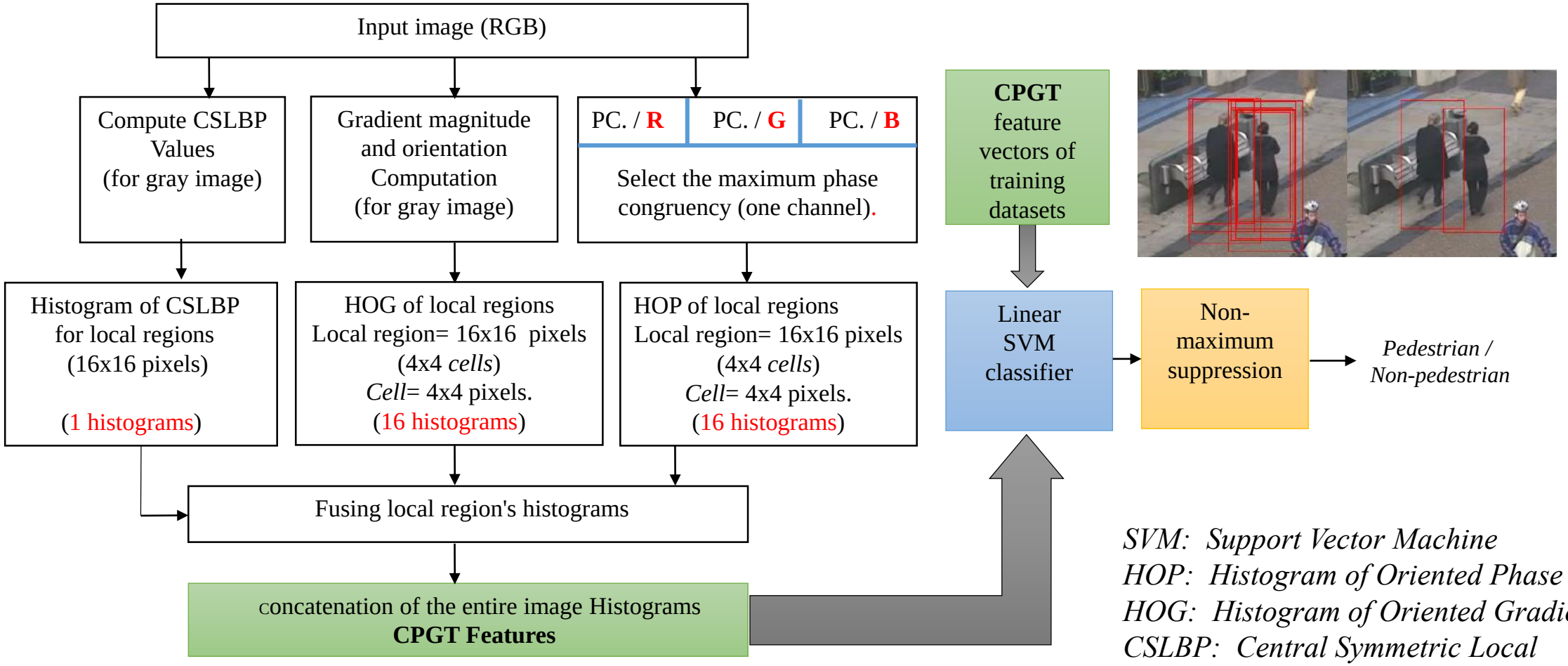


Human Detection: System Perspective



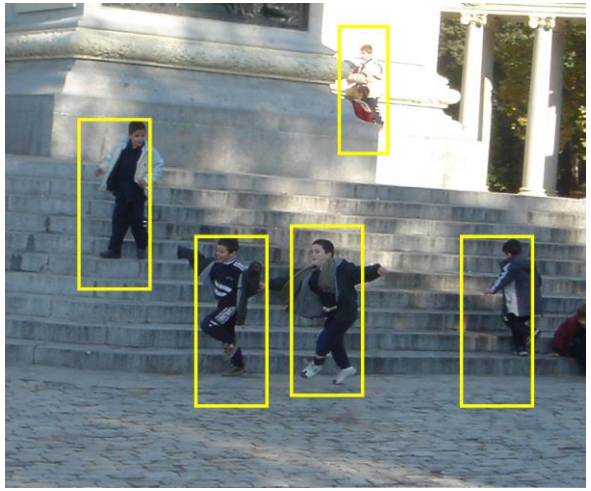
Framework of the Human Detection System

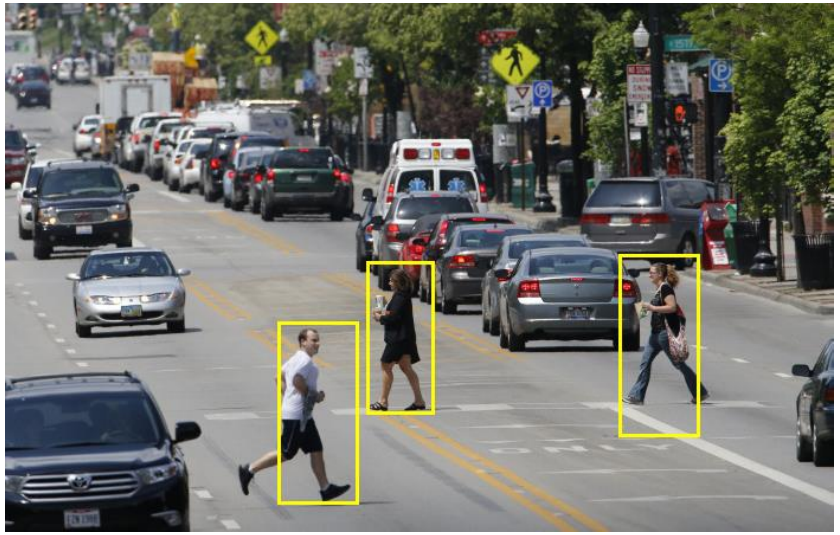
Chromatic domain phase features with gradient and texture (CPGT)

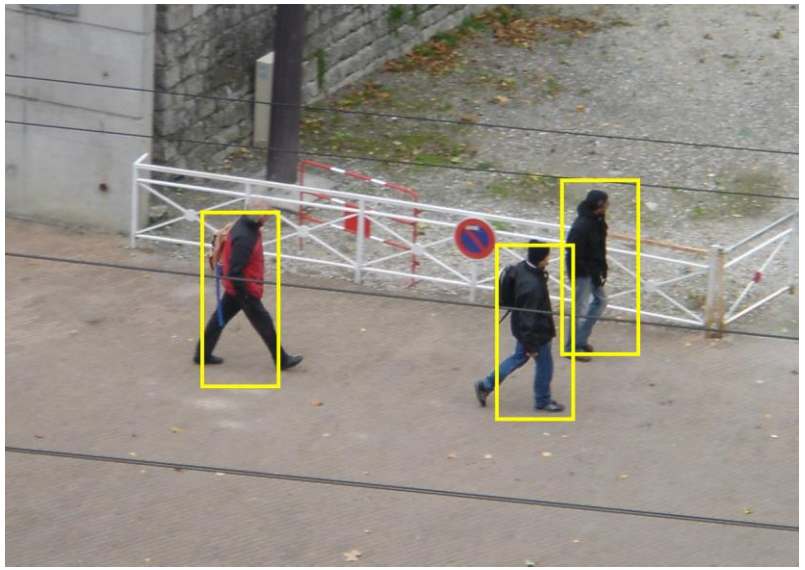


SVM: Support Vector Machine
HOP: Histogram of Oriented Phase
HOG: Histogram of Oriented Gradient
CSLBP: Central Symmetric Local Binary Pattern

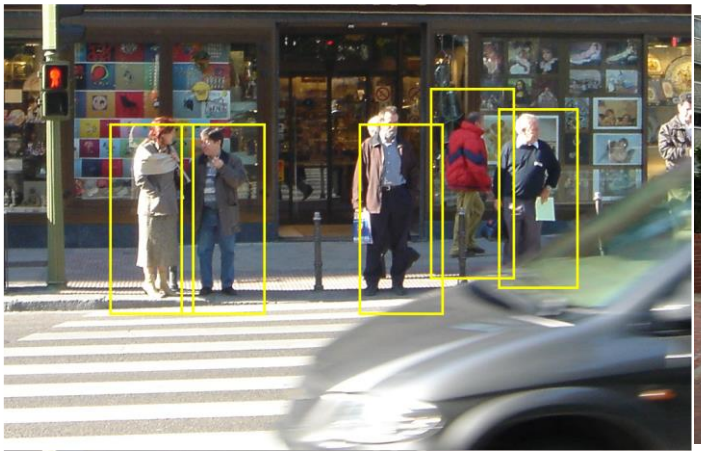
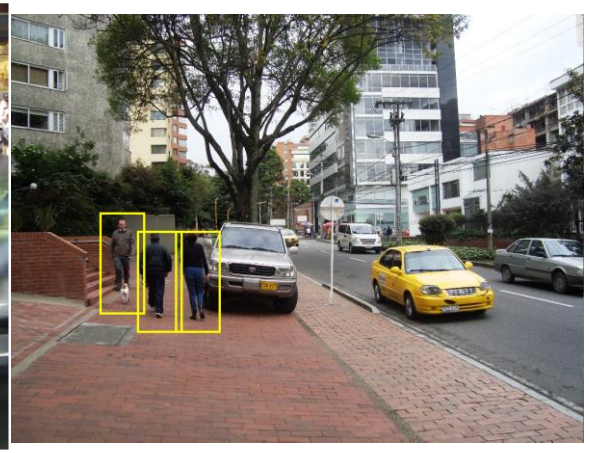
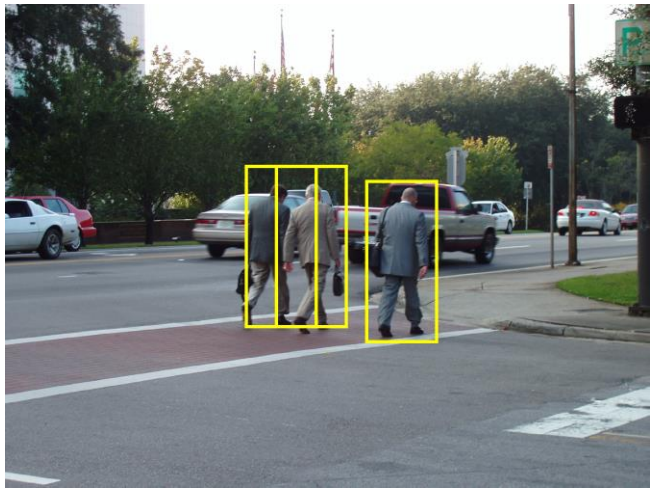
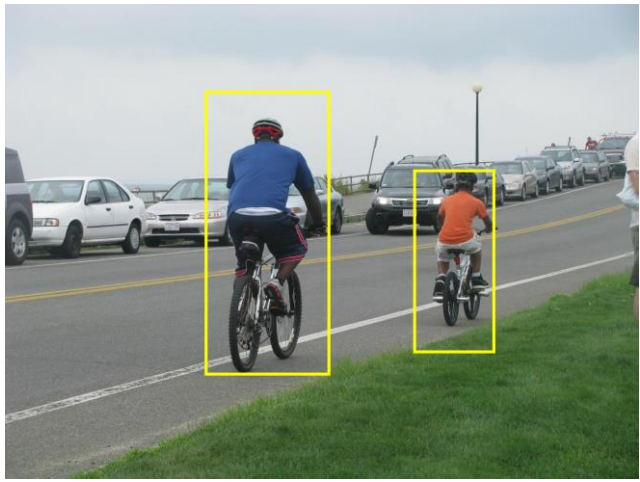
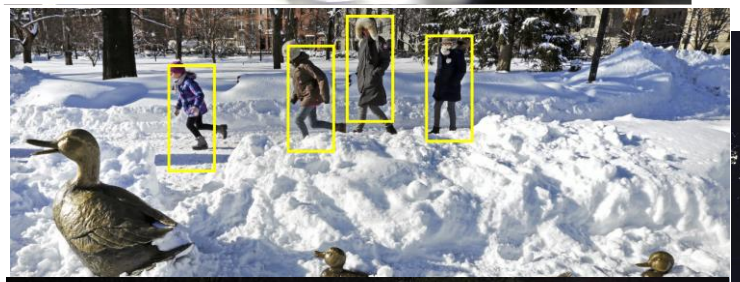
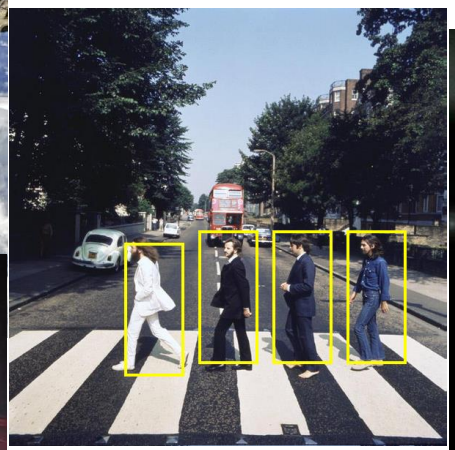
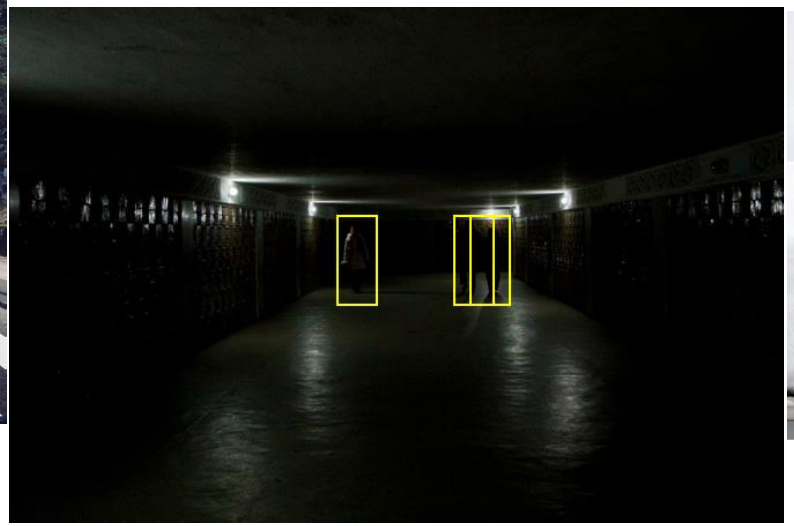
Human Detection







Human Detection

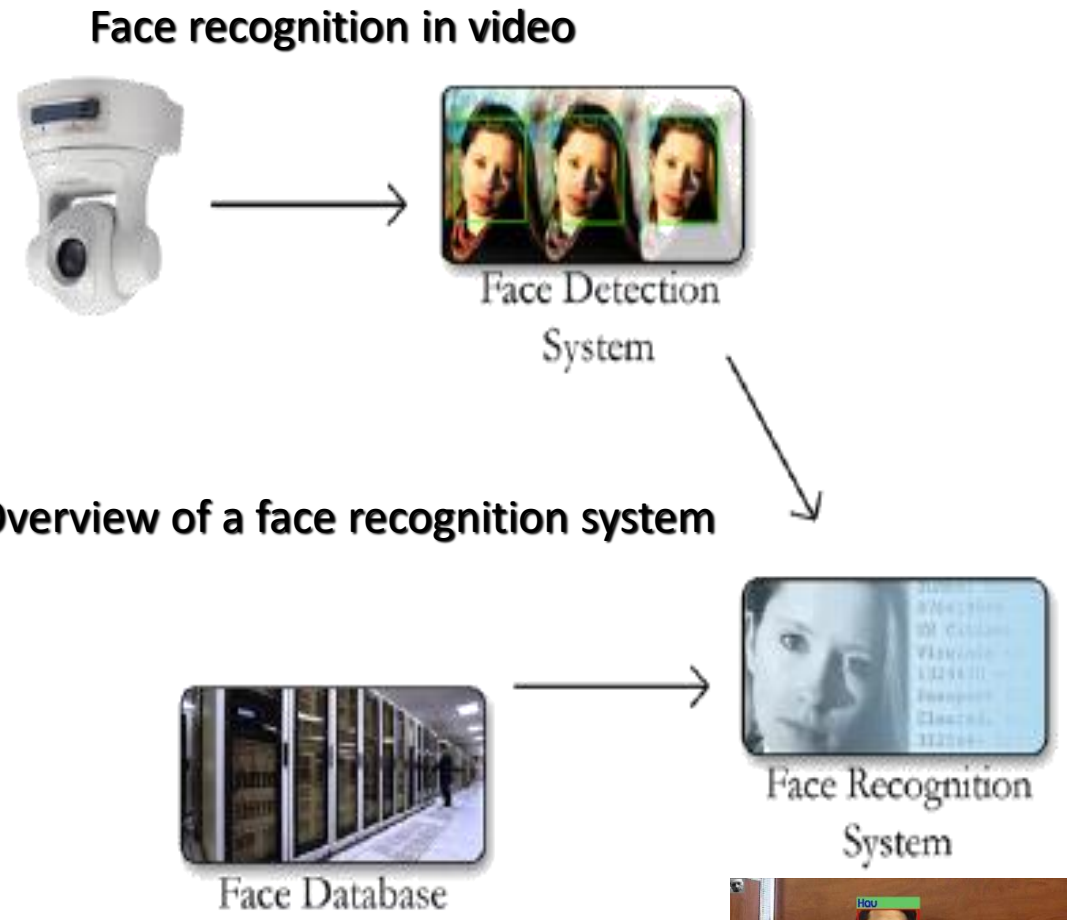








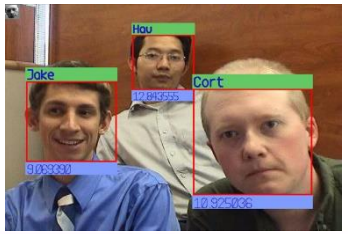
Face Recognition System

Face Detection – quickly and efficiently locates all faces in a given image region.

Face Features– calculates unique features of each person in the face database that can be used for accurate classification.

Feature Classification – compares features of face regions obtained from the detection process with face feature data computed from the training stage to determine the identity of individuals.

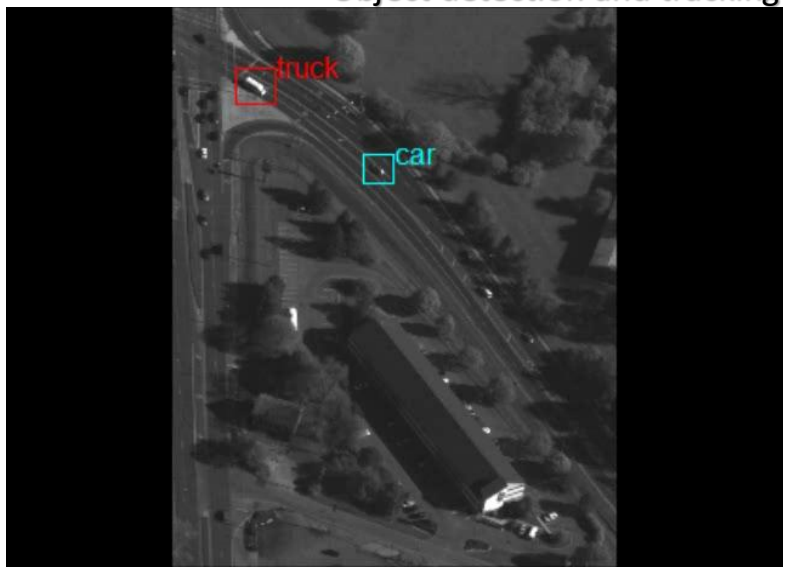


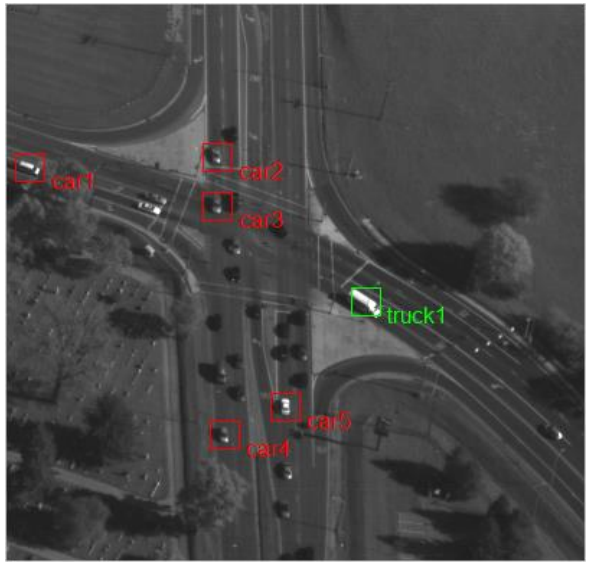


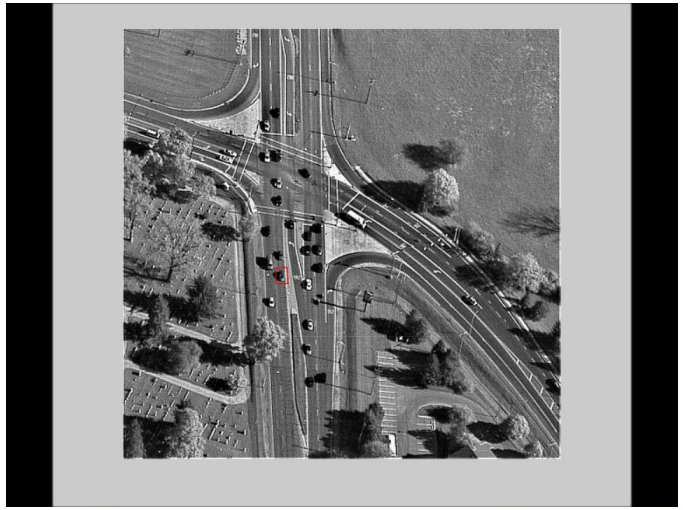
Face Recognition in Low Light

Object Detection, Tracking, and Identification

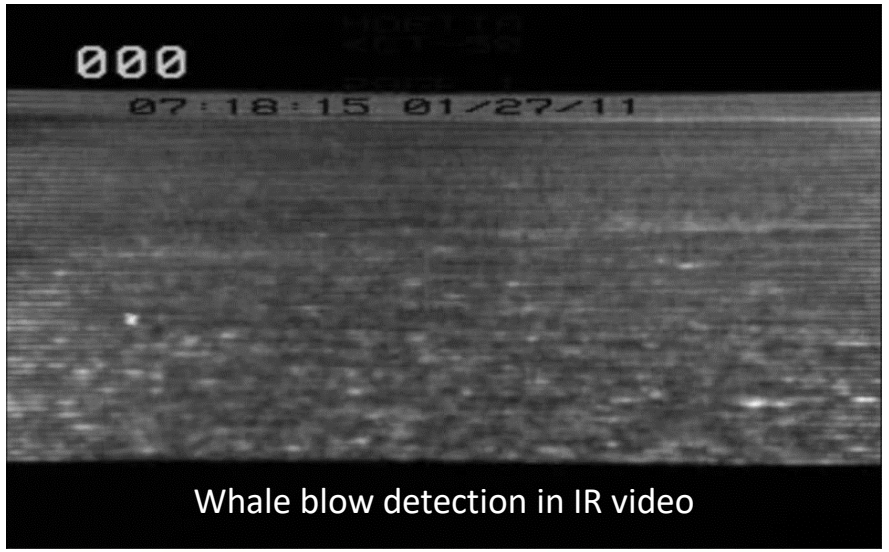
Object detection and tracking on WAMI data







Small boat detection

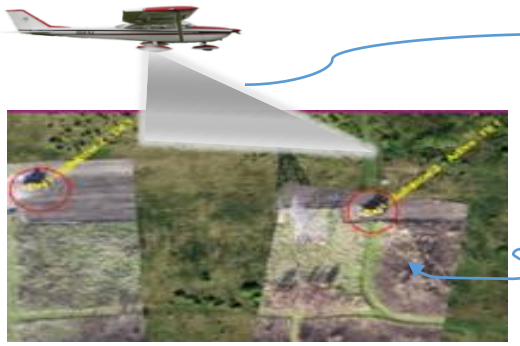


Whale blow detection in IR video

Object Tracking

- Description:
  - From data collected at Eglin Air Base during DARPA VIVID program. Tracking a pickup truck along a dirt road in a wooded area. Occlusion by trees. Illumination changes as truck passes in and out of tree shadows.
- Testing Scenario Challenges
  - Harshest test on features, due to partial occlusions and loss of features
  - Occlusions and shadows

Oil/Gas Pipeline Right of Way Intrusion Detection



Aerial Imagery

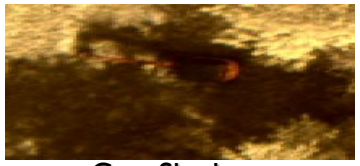
Image Analysis

Automated System

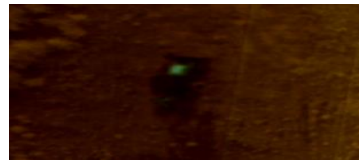
Object Classification

Threat Localization

KML File Mapping



Cast Shadows



Low illumination

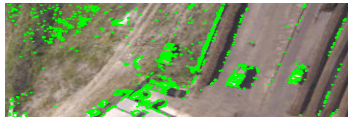
Orientation

Partial Occlusion

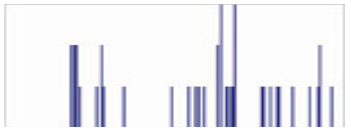
Small Scale

Color Transformation and Enhancement

Key Frame Selection



Key Region Localization



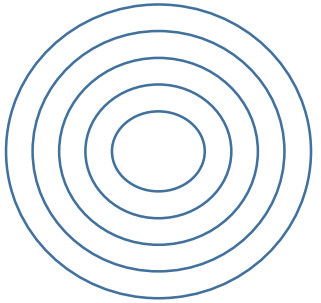
Local Feature Extraction

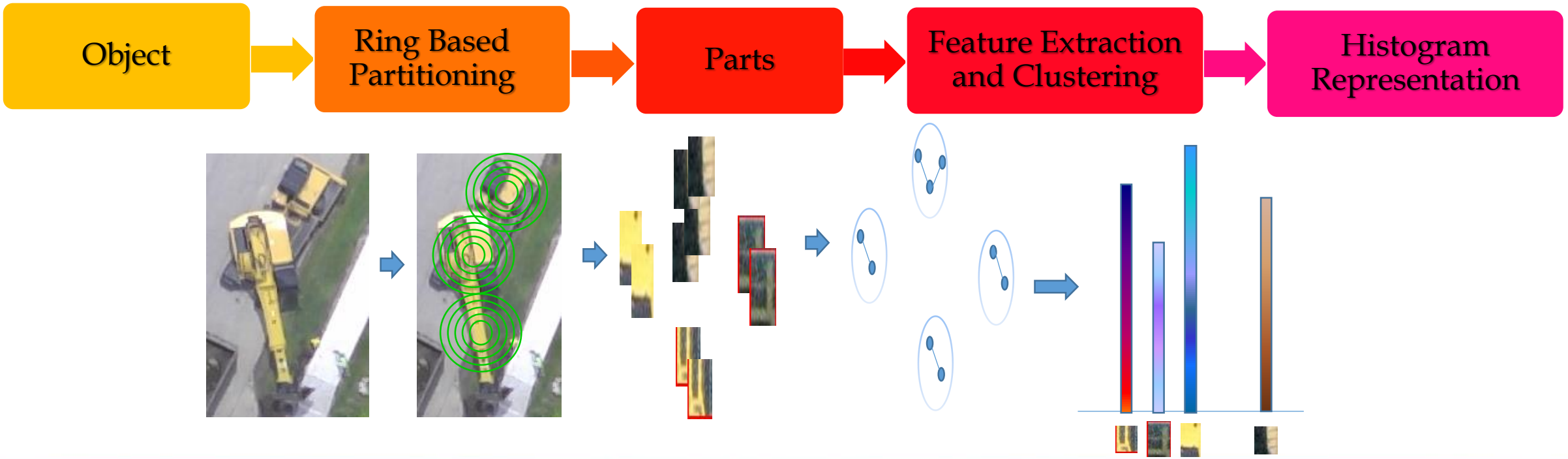
Object Detection and Identification

Part-Based Ringlet Model

Method: Using Ring Histogram for each part of objects

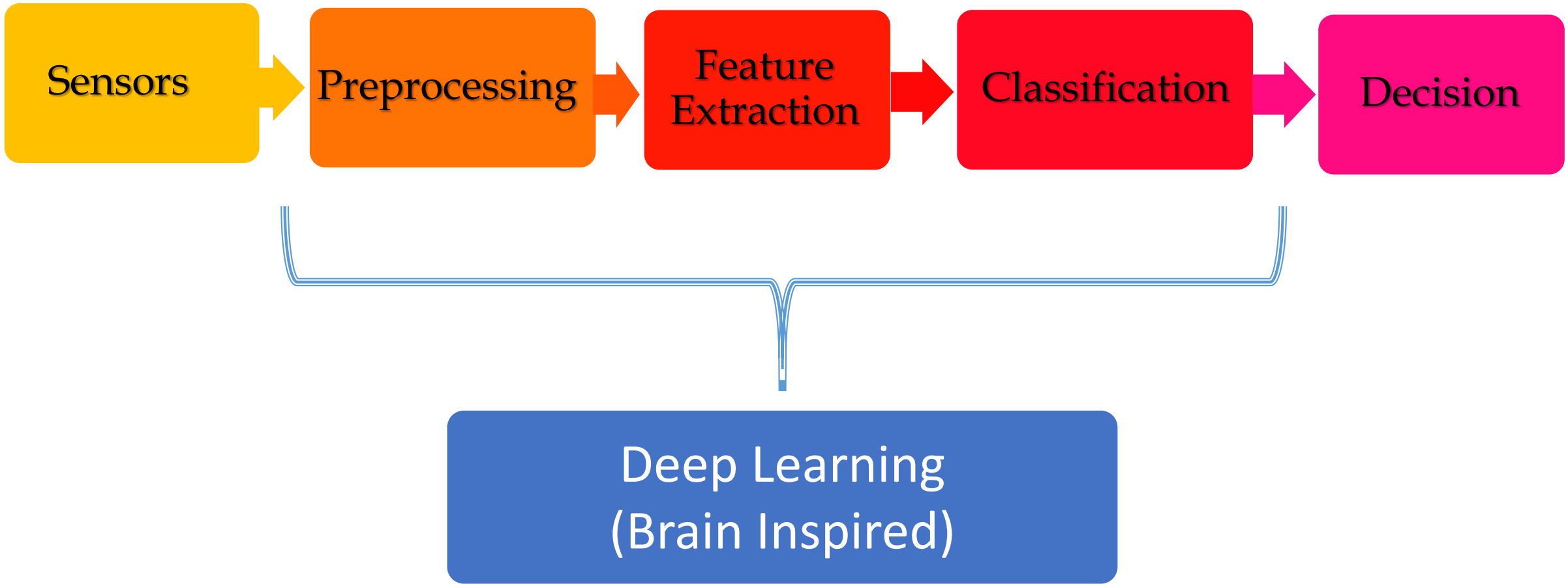
- Invariant to rotation
- Still contains spatial information
- Still contains partial occlusion ability



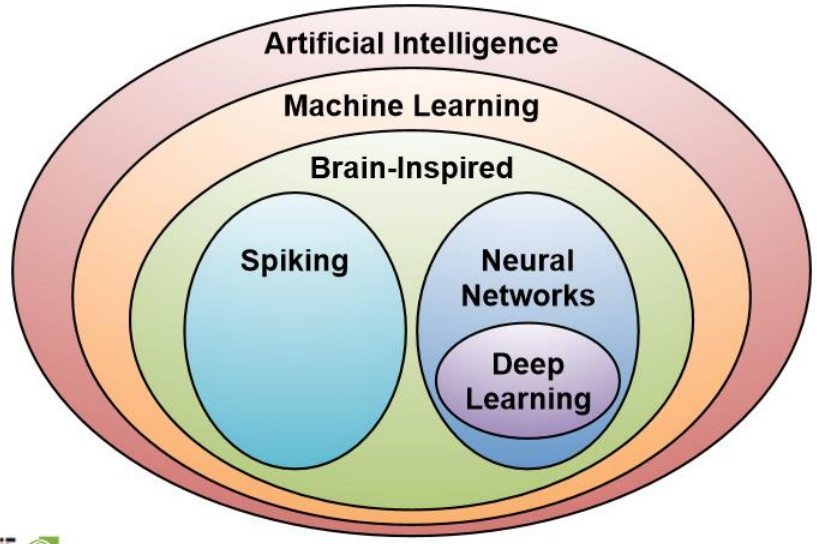


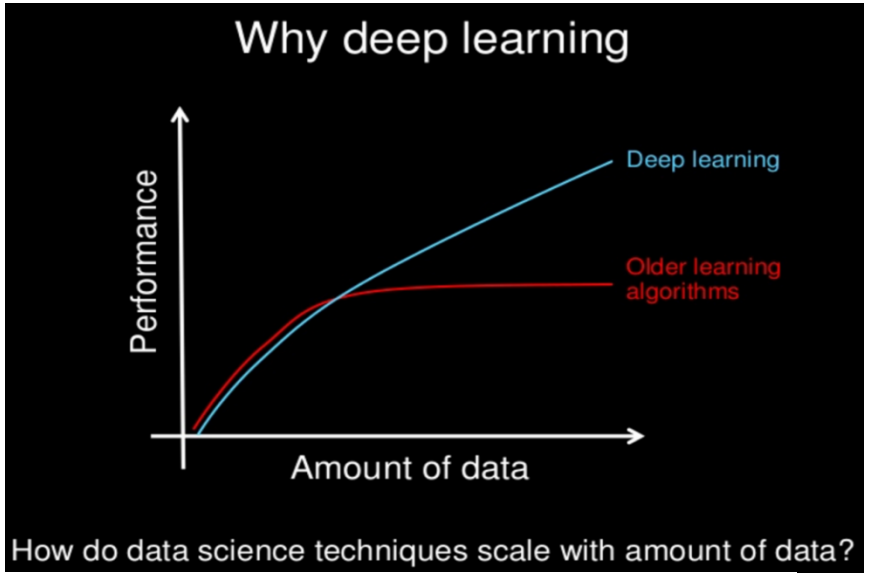
Object Detection

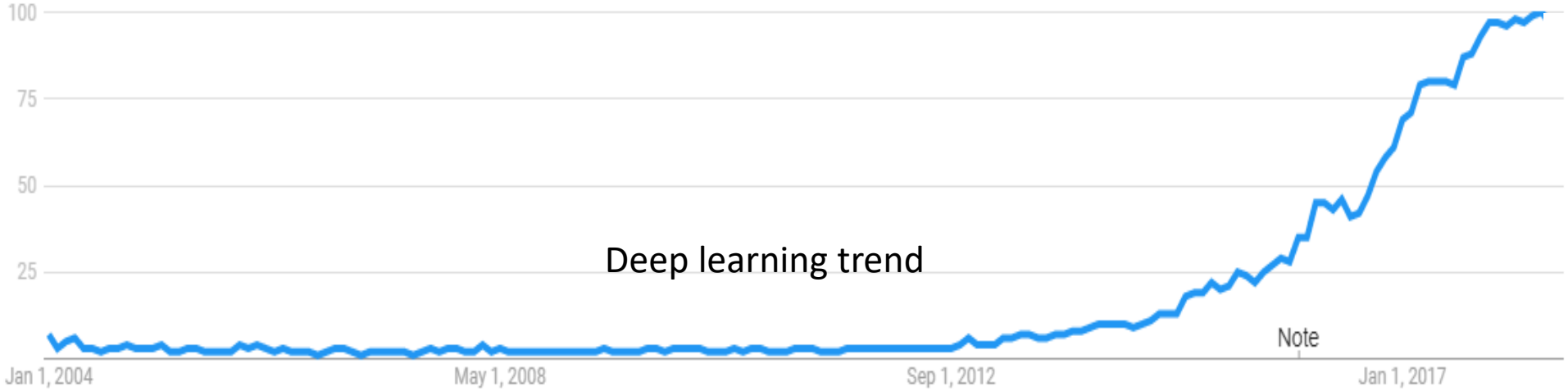
Human Perception Modeling: Deep Learning



Human Perception Modeling: Deep Learning

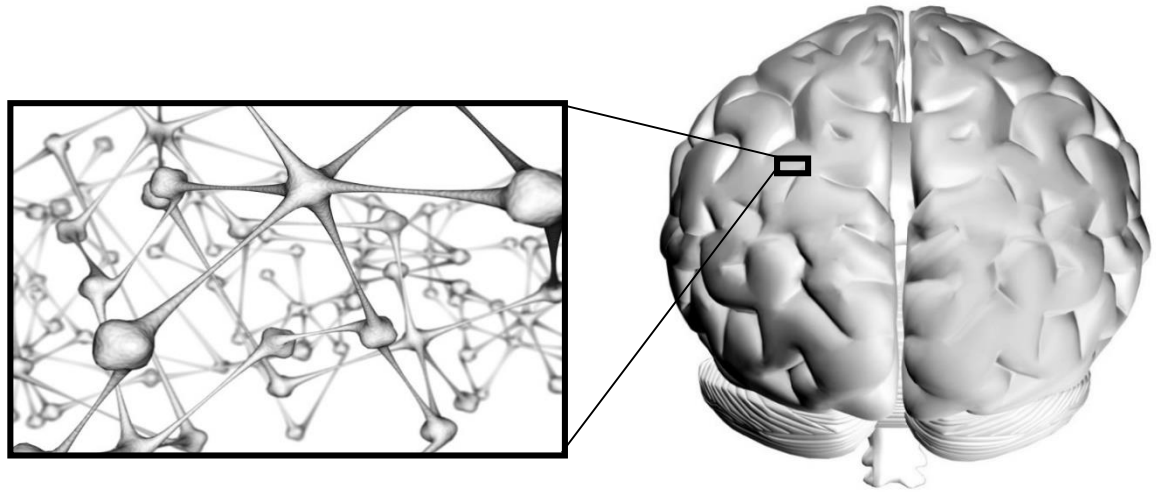


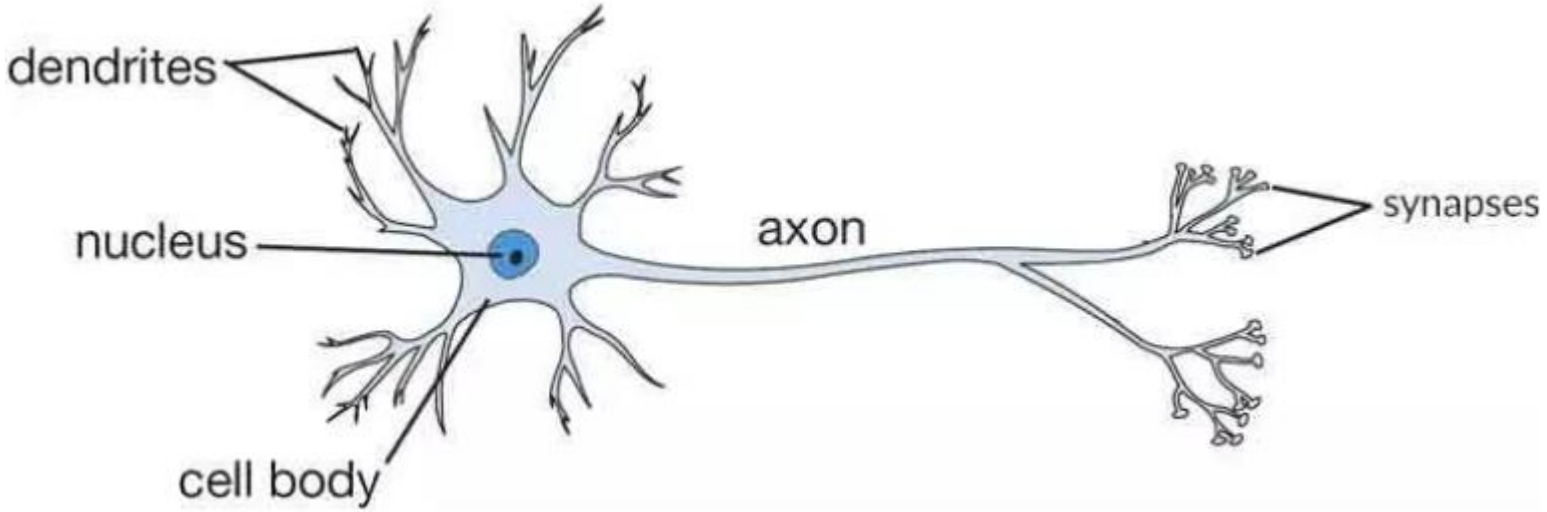




The Human Brain

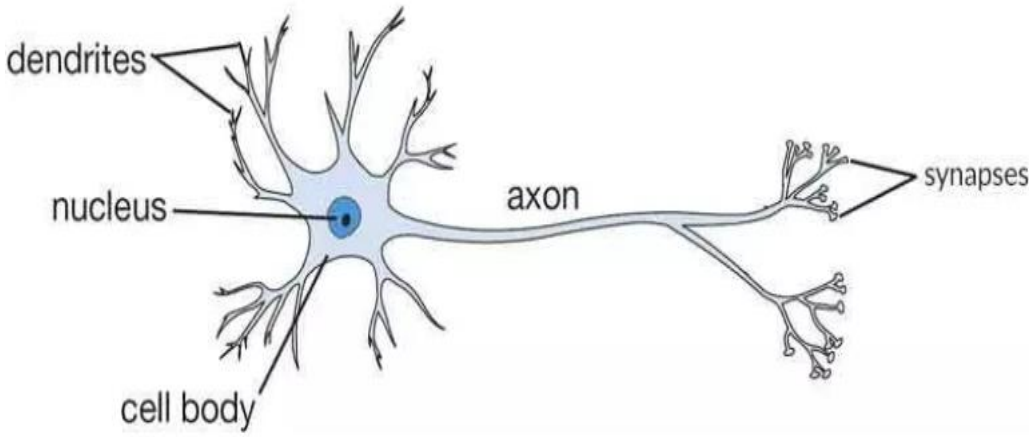
- Center of the Nervous System
  - Controls and monitors actions and reactions
    - Muscular activity
    - Organ function
    - Memory retention
- Several Neurons
  - 50-100 billion neurons
  - Each with 10,000 synaptic connections





The Neuron

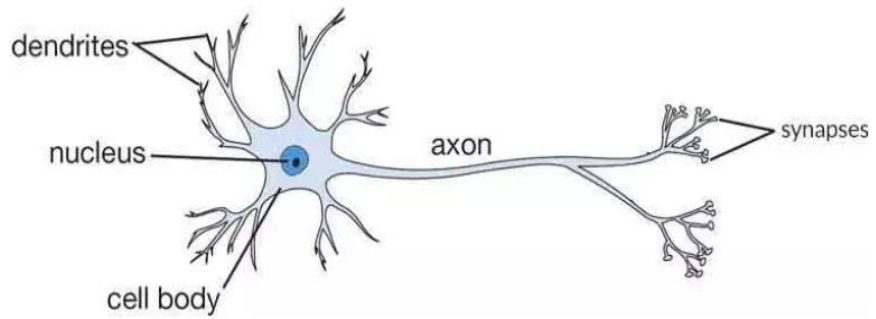
- Cell Body
  - Center of the neuron, contains the soma and nucleus, controls protein synthesis
- Dendrites
  - Cellular extensions from other neurons, input to the nucleus
- Axons
  - Only one axon per neuron, branches out through terminal buttons, creating synaptic junctions with other neurons

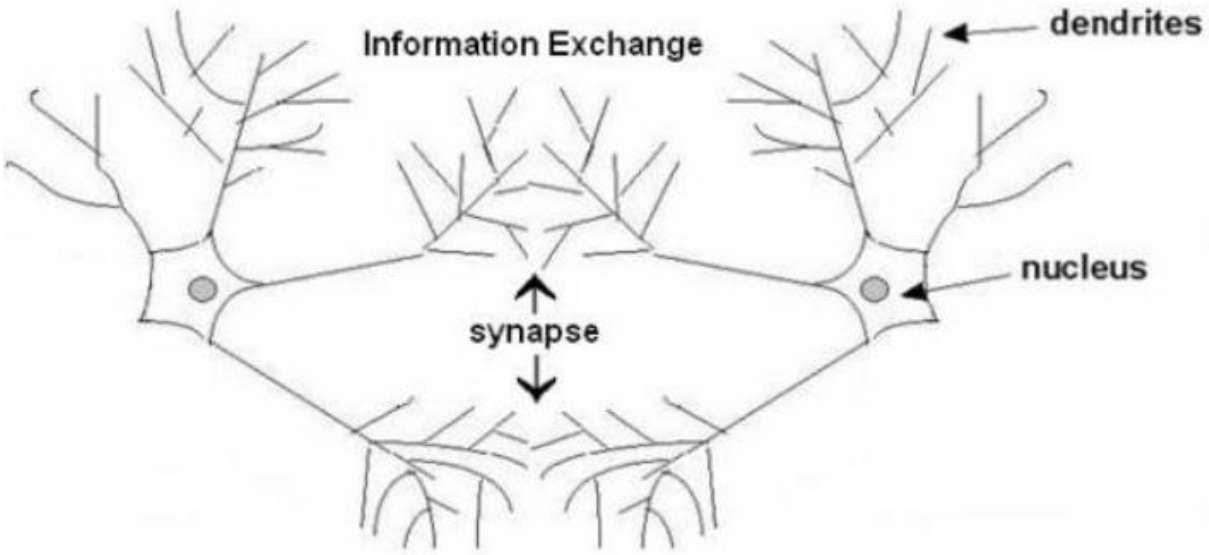


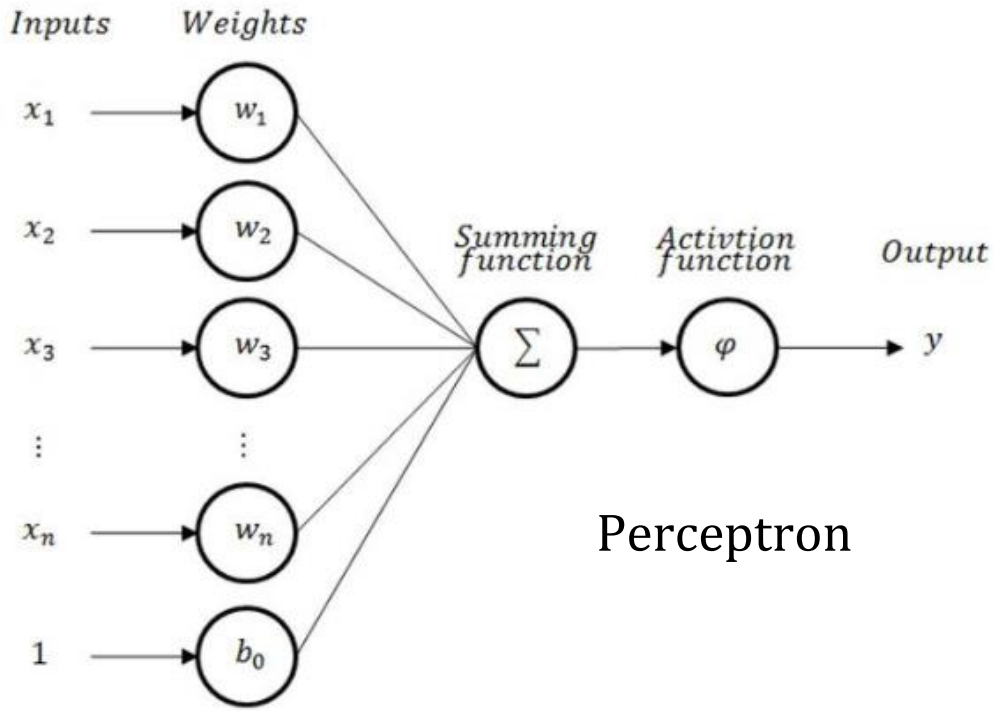
- Synaptic Junctions (Neurons communicate through synapses)
  - Connects terminal branches from the output (axon: pre-synaptic) of a neuron to the inputs (dendrites: post-synaptic) of other neurons
- Neuron Firing
  - Ions flow to branch ends releasing a neurotransmitter and binds to the post-synaptic neuron

Artificial Neuron

- Cell Body: Center of the neuron
- Dendrites: Cellular extensions from other neurons
- Axons: Only one axon per neuron, branches out through terminal buttons
- Synaptic Junctions: Connects terminal branches
- Neuron Firing: Ions flow to branch ends releasing a neurotransmitter and binds to the post-synaptic neuron

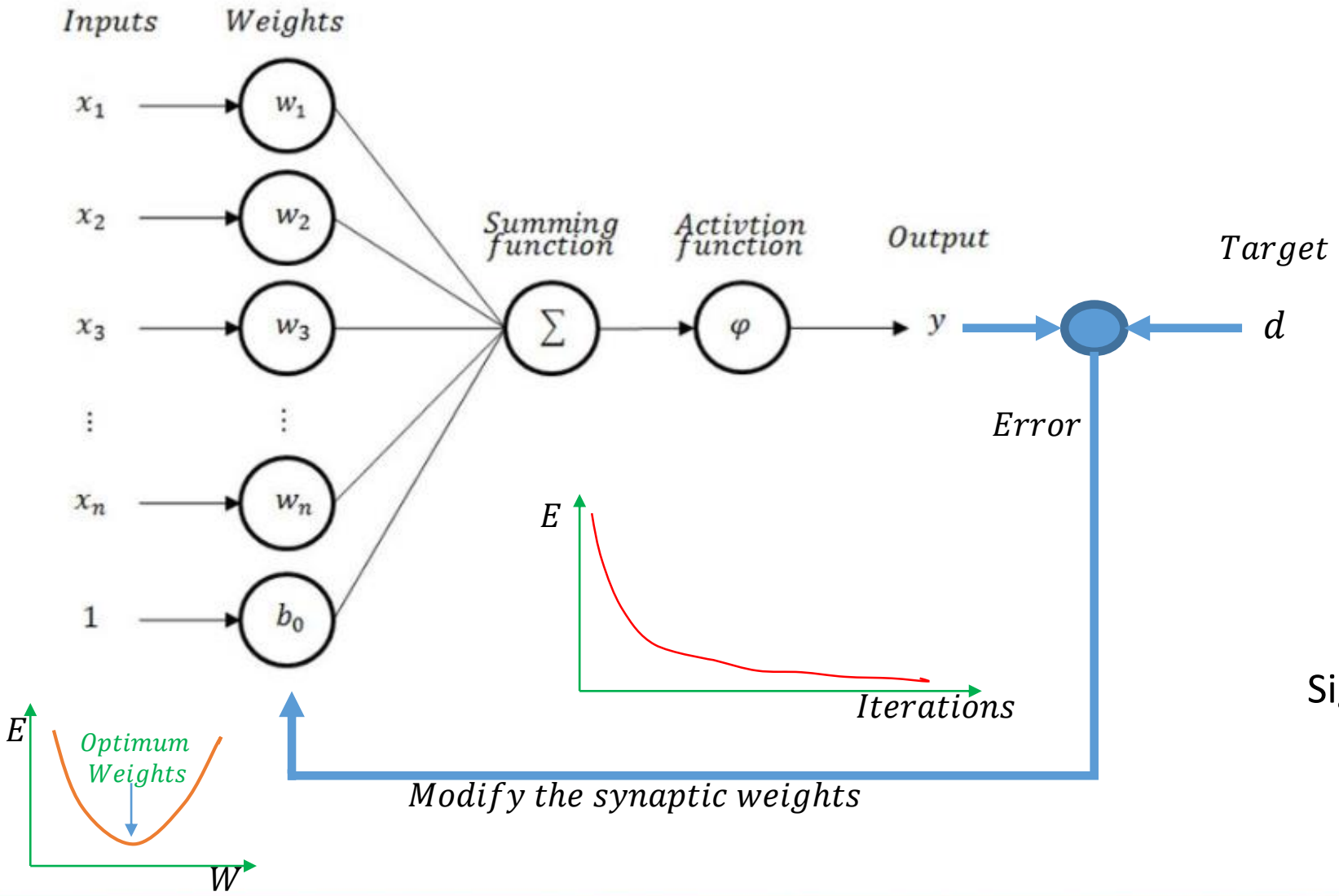






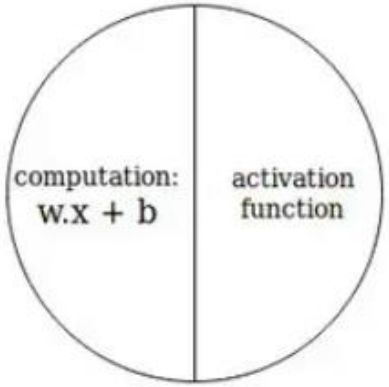
Perceptron

Perceptron Learning



$$X = \{x_1, \ x_2, \ \ldots, \ x_n\}$$

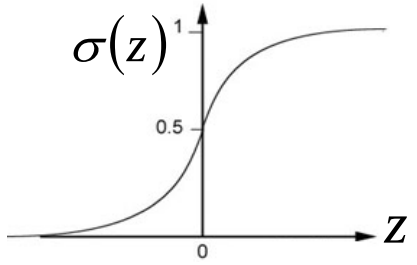
$$W = \{w_0, \ w_1, \ \ldots, \ w_n\}$$



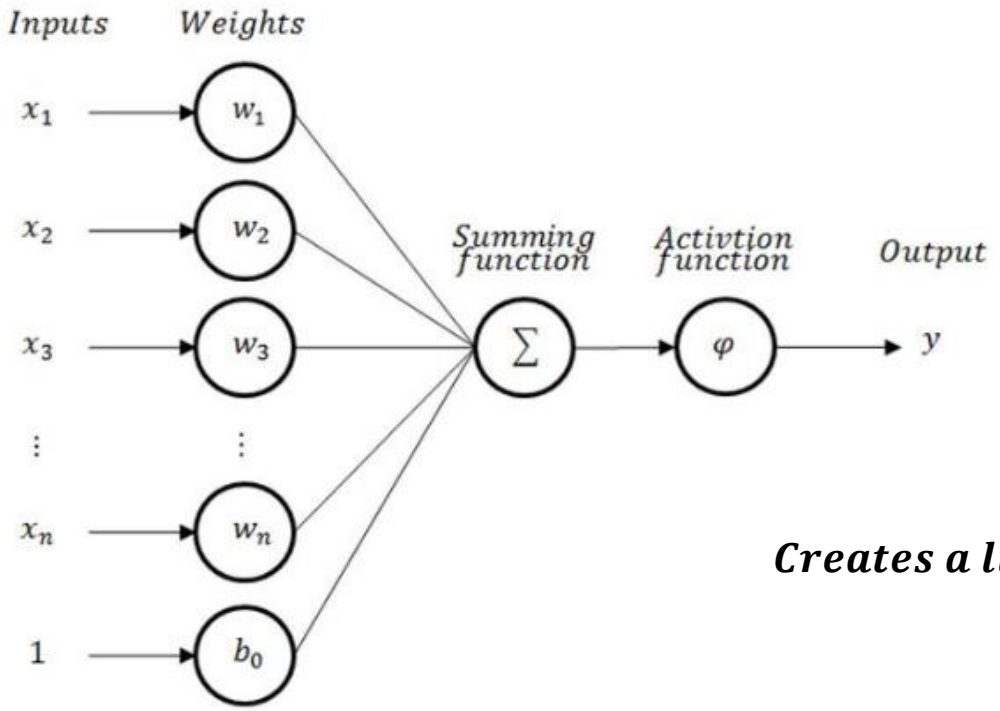
$$y = Fn\,[W.X + b]$$

Sigmoid Function

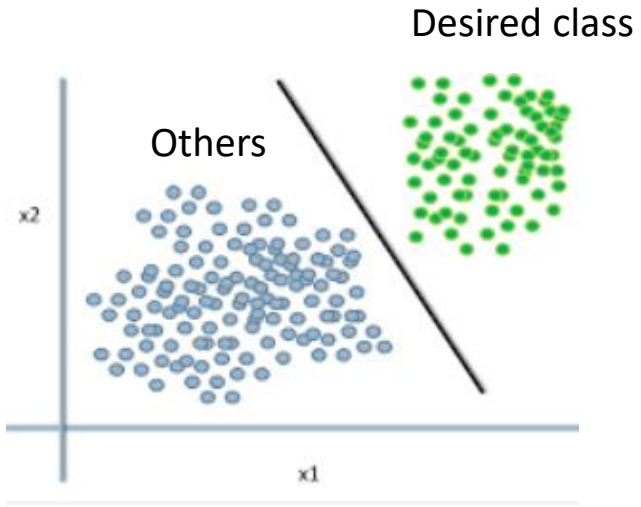
$$\sigma(z) = \frac{1}{1+e^{-z}}$$



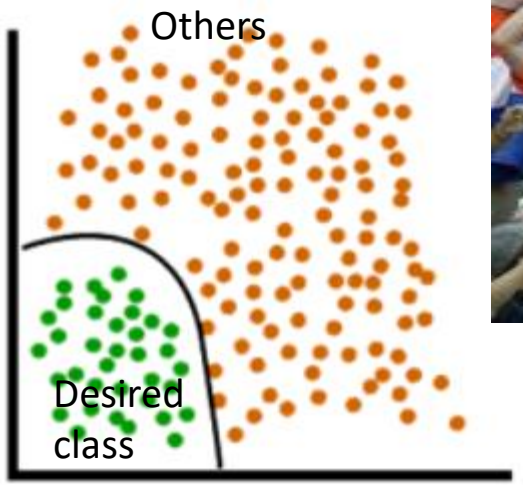
Perceptron Learning



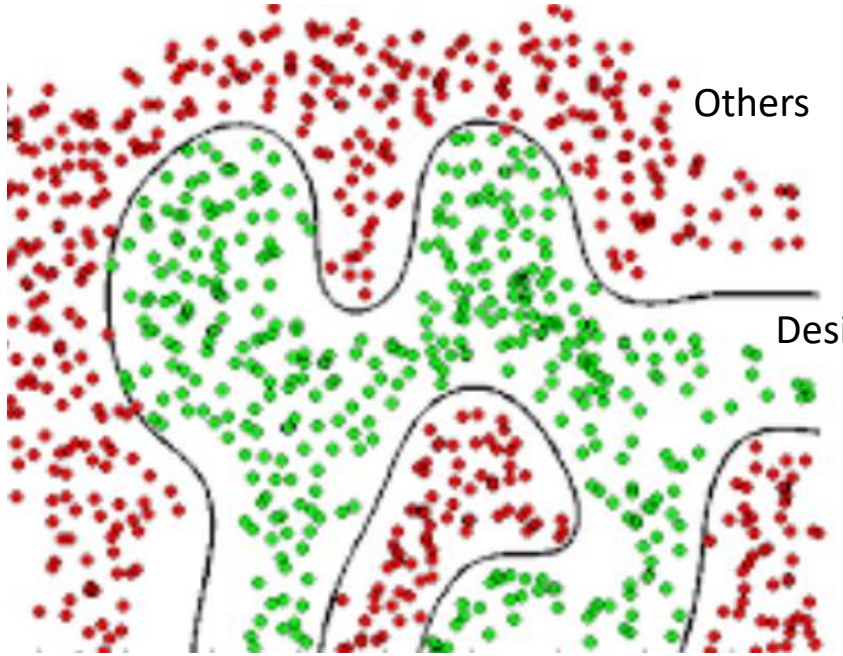
Creates a linear decision boundary



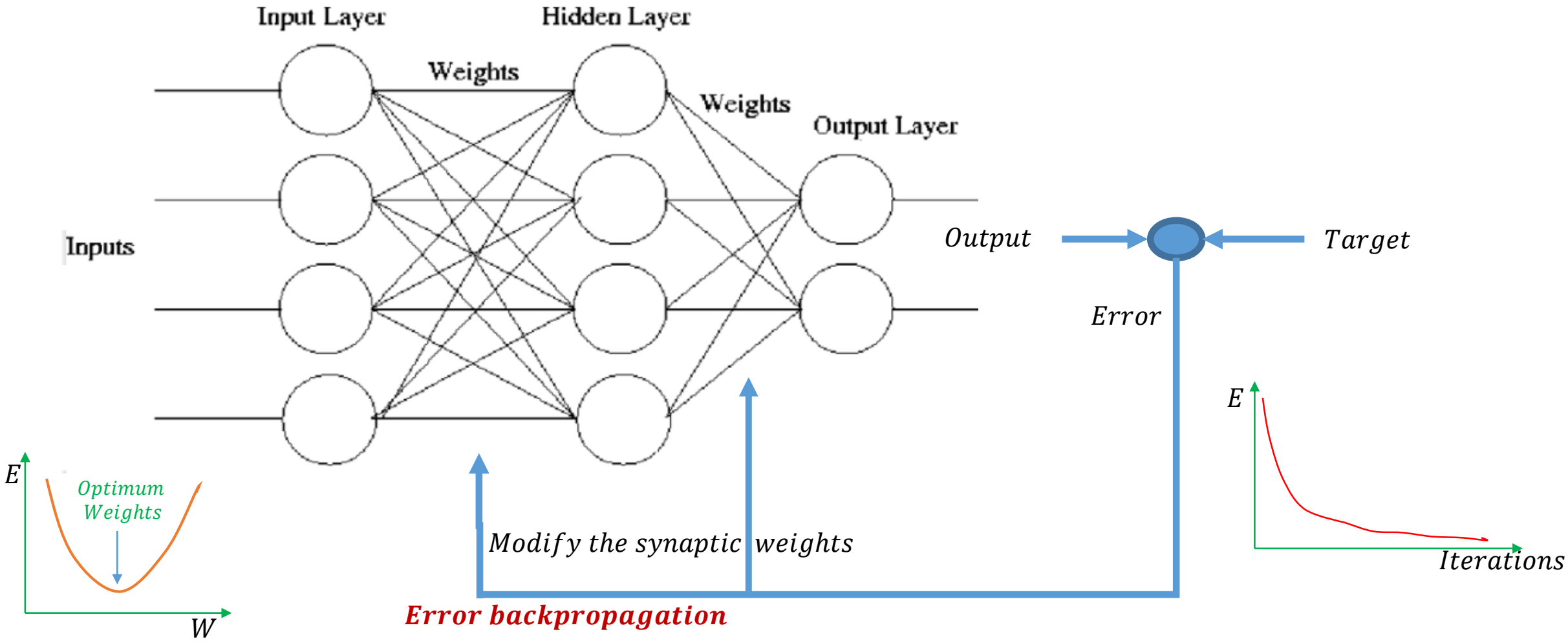
Real Life Situation



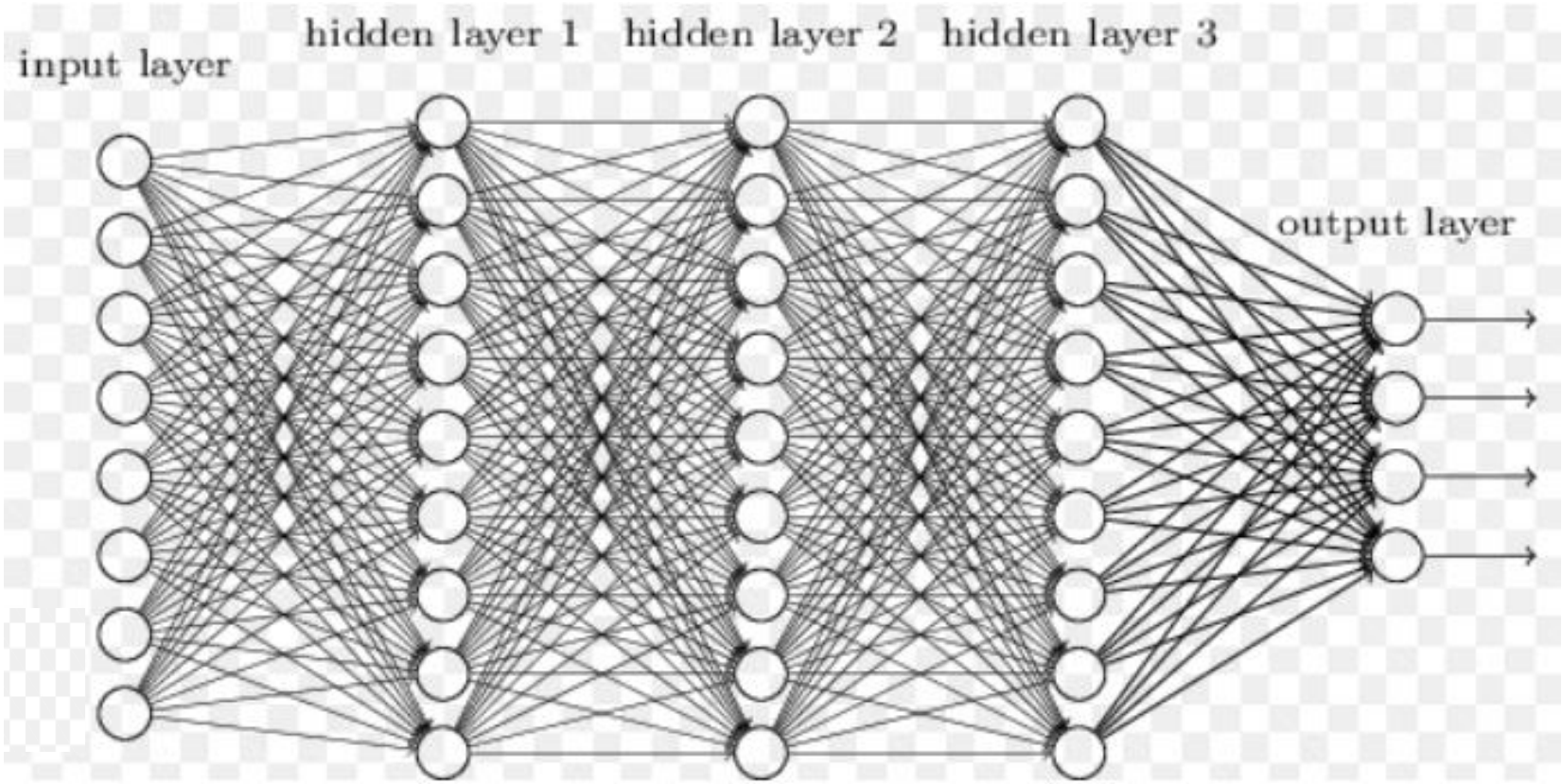
Need for nonlinear decision boundary

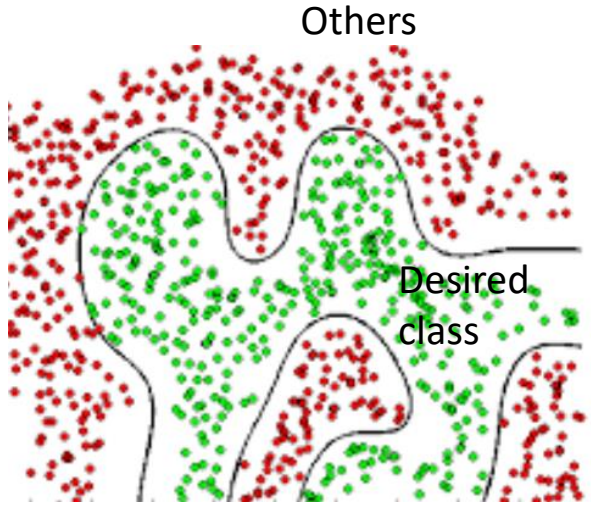


Multilayer Neural Networks

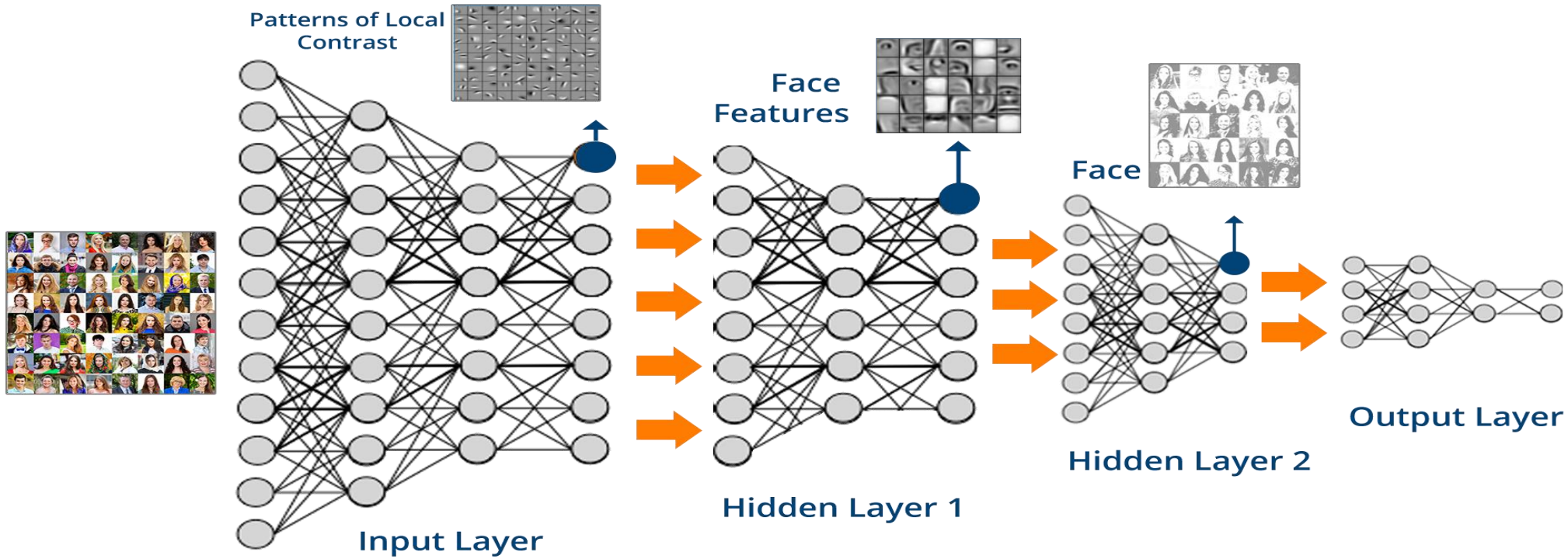


Multilayer Neural Networks





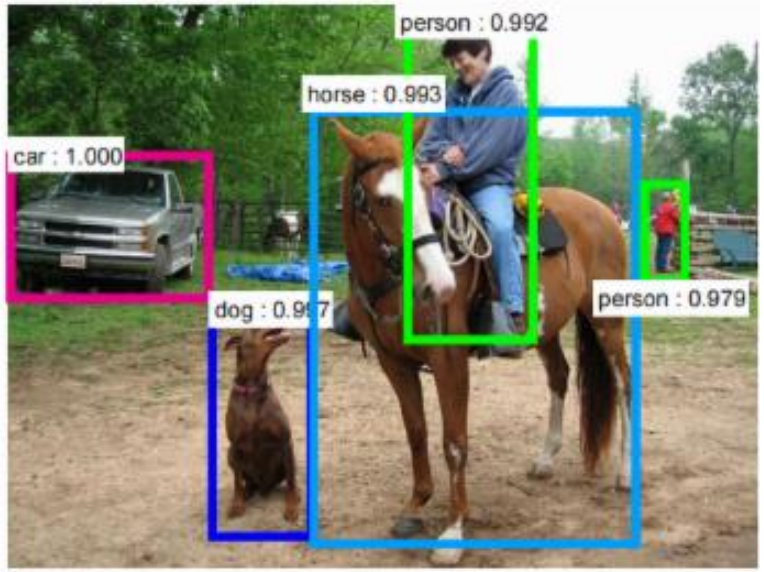
Convolutional Neural Networks (CNN)

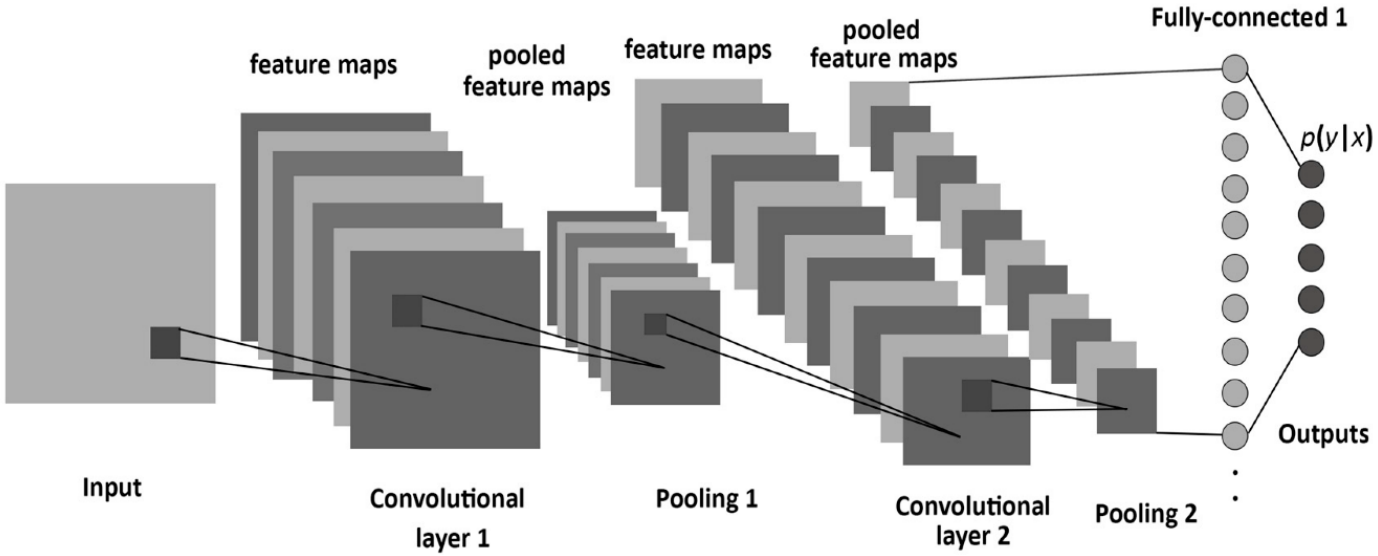


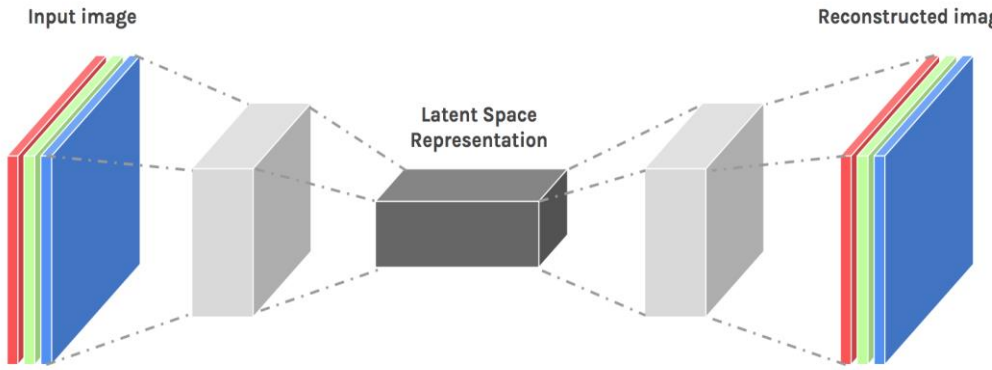
- A feed-forward network that can extract topological properties from an image.
- Trained with error back-propagation algorithm.
- Designed to recognize visual patterns directly from pixel images with minimal preprocessing.
- Can recognize patterns with extreme variability.

Deep Convolutional Neural Network

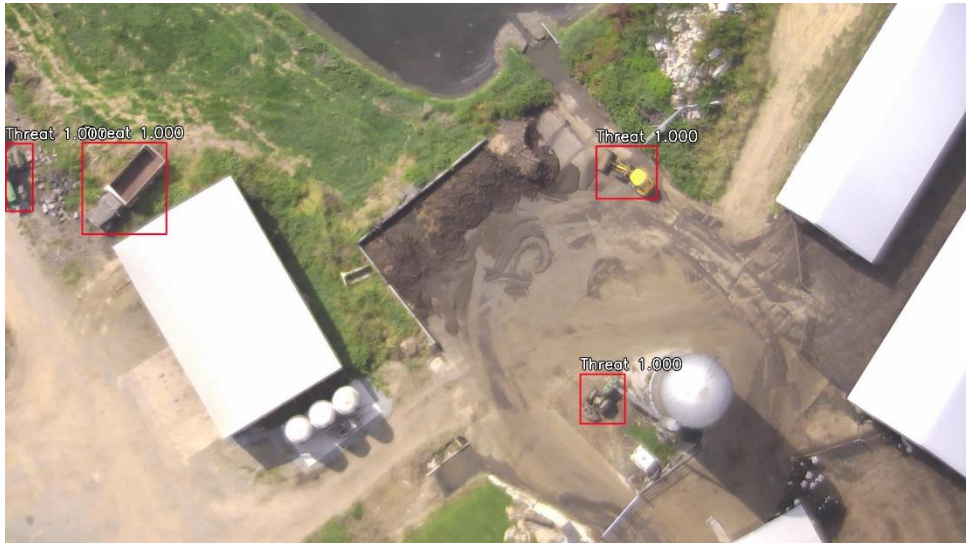
- Learning through Association
  - Higher level representations are created through various filtrations from multiple layers in a deep learning architecture



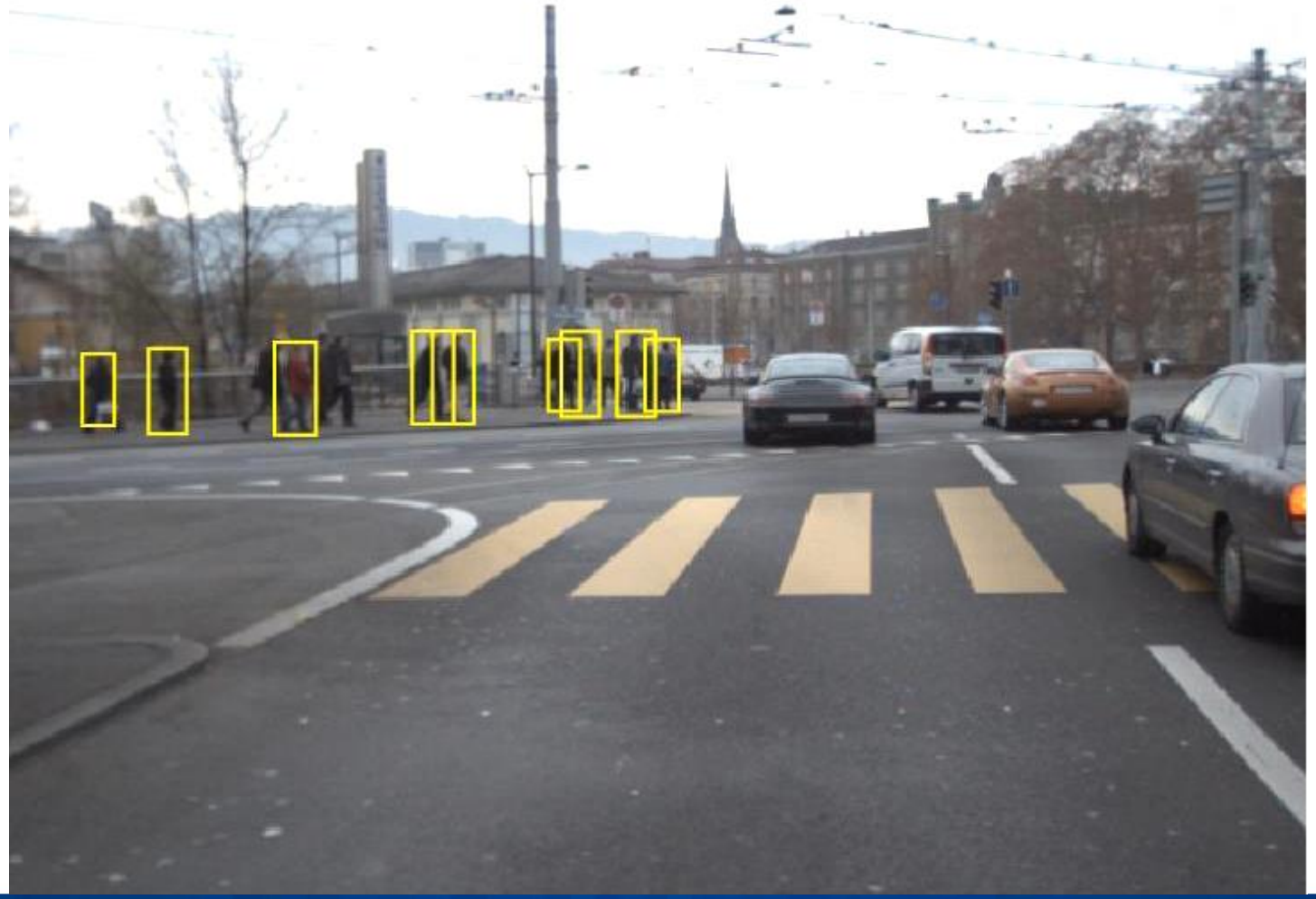




Detection Results

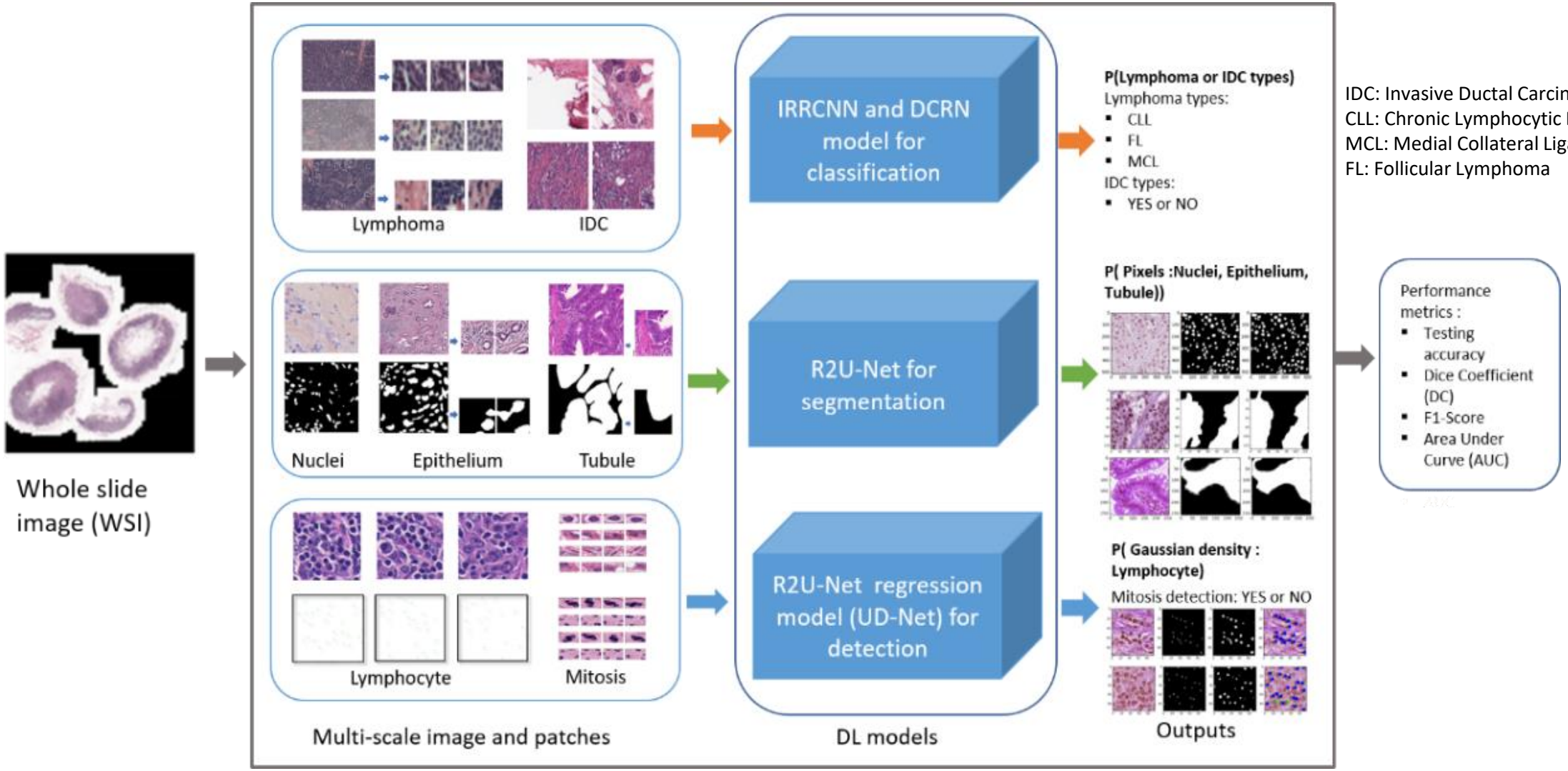


Human Detection



Human Detection in Infrared Imagery

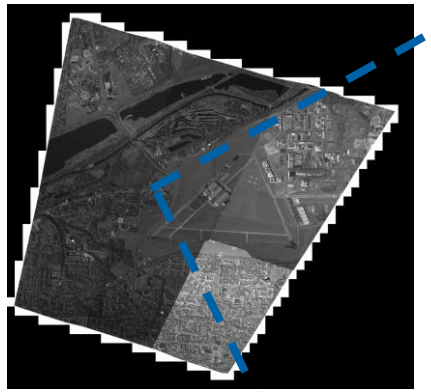
Computational Pathology

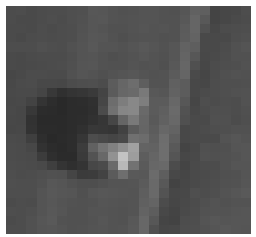


IDC: Invasive Ductal Carcinoma
CLL: Chronic Lymphocytic Leukemia
MCL: Medial Collateral Ligament
FL: Follicular Lymphoma

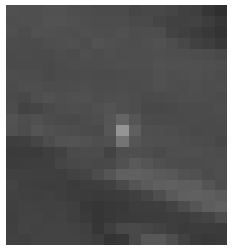
Multi-Object Detection and Tracking

- Utilizes the current frame, previous frame, and background model to obtain detection and tracking of object
  - Includes both feature and motion models
  - Includes image registration and stitching





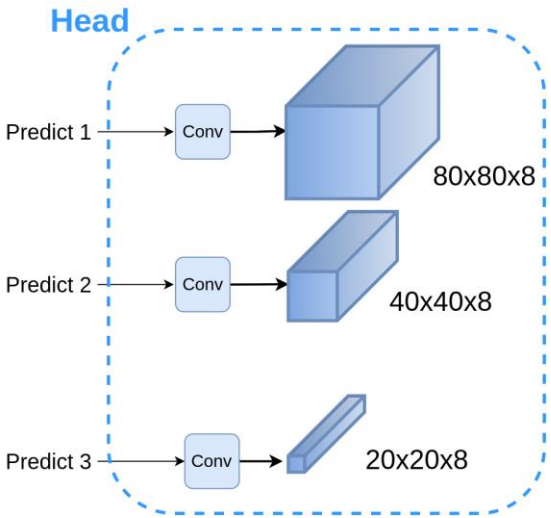
Objects of interest – cars, vans, trucks
15 x 15 pixels to 18 x 24 pixels

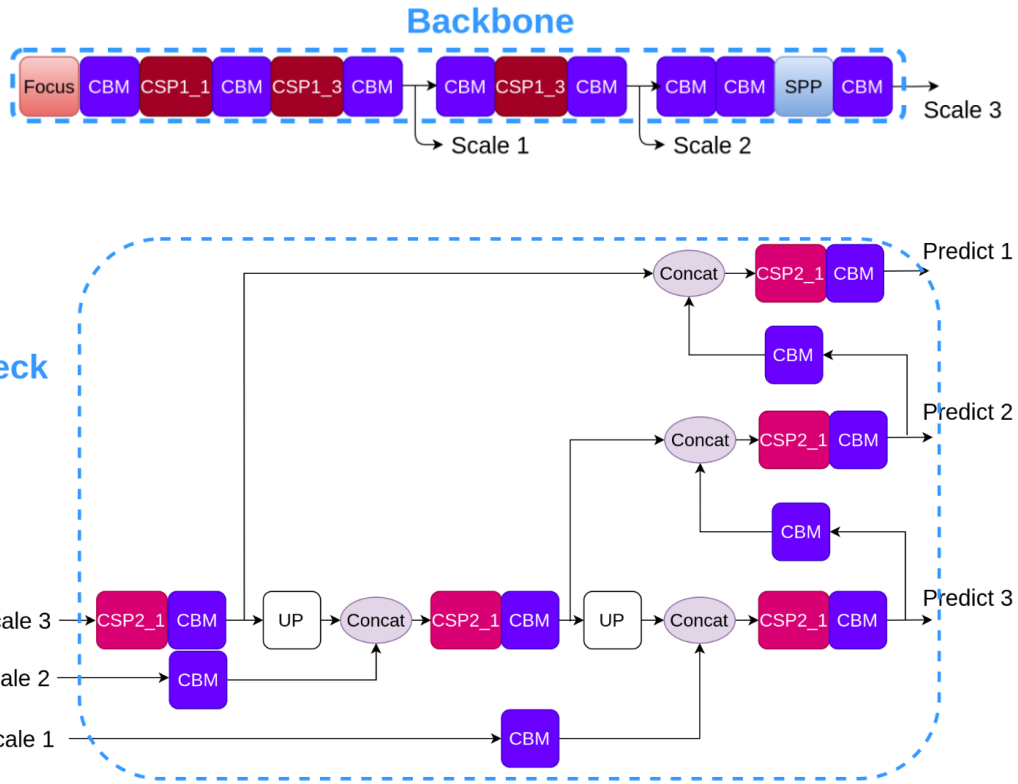


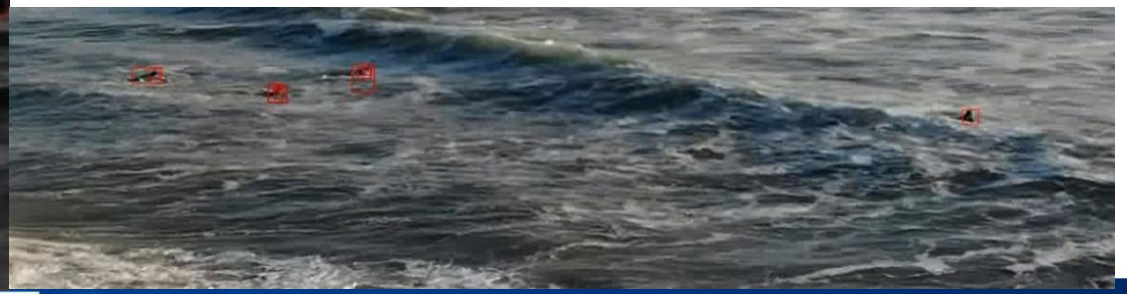
Person
2 x 7 pixels

Search and Rescue in Ocean

- Scale and Rotation Invariant Deep Neural Machine for Human Rescue in Ocean







University of Dayton

VISION LAB

Center of Excellence

Thanks

www.visionlab.udayton.edu