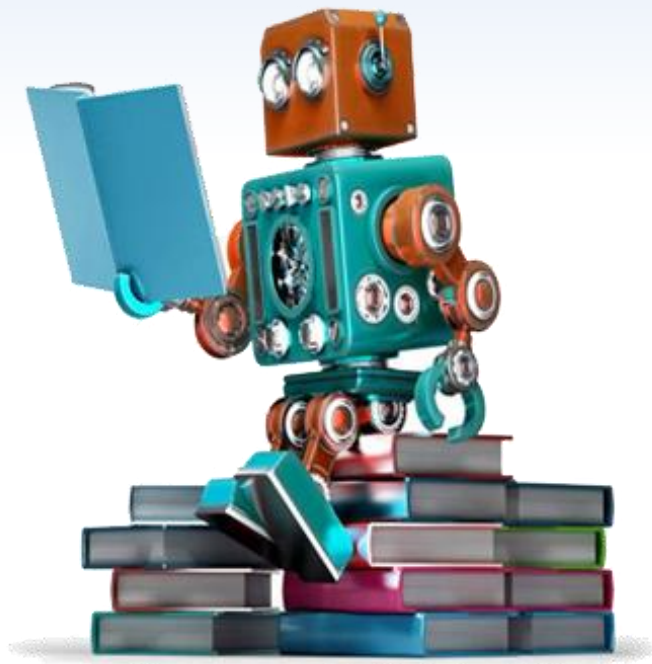


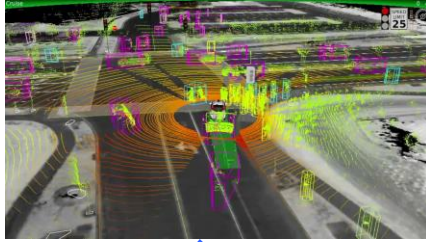
西北工业大学
NORTHWESTERN POLYTECHNICAL UNIVERSITY

机器学习 Machine Learning

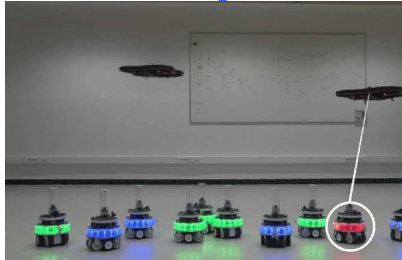
布树辉



Future?



AI



UAV

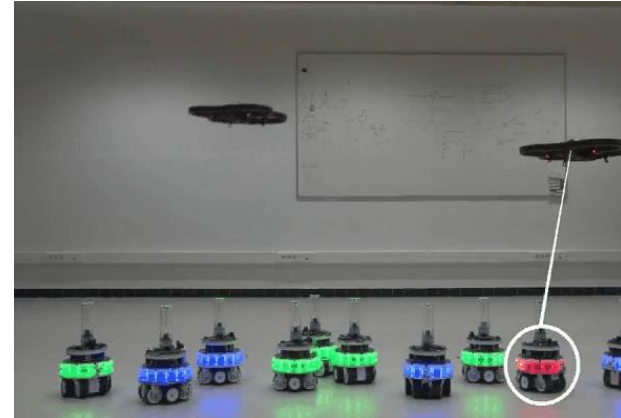
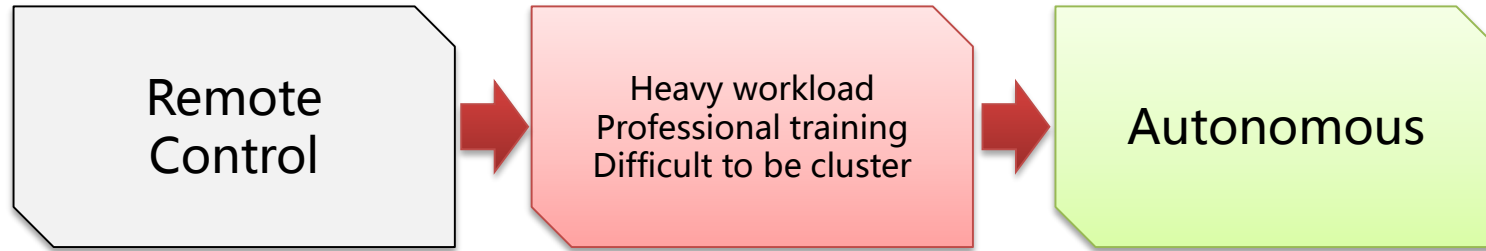


- Small, cheap
- No pilot
- Convenient
- Strong survivability

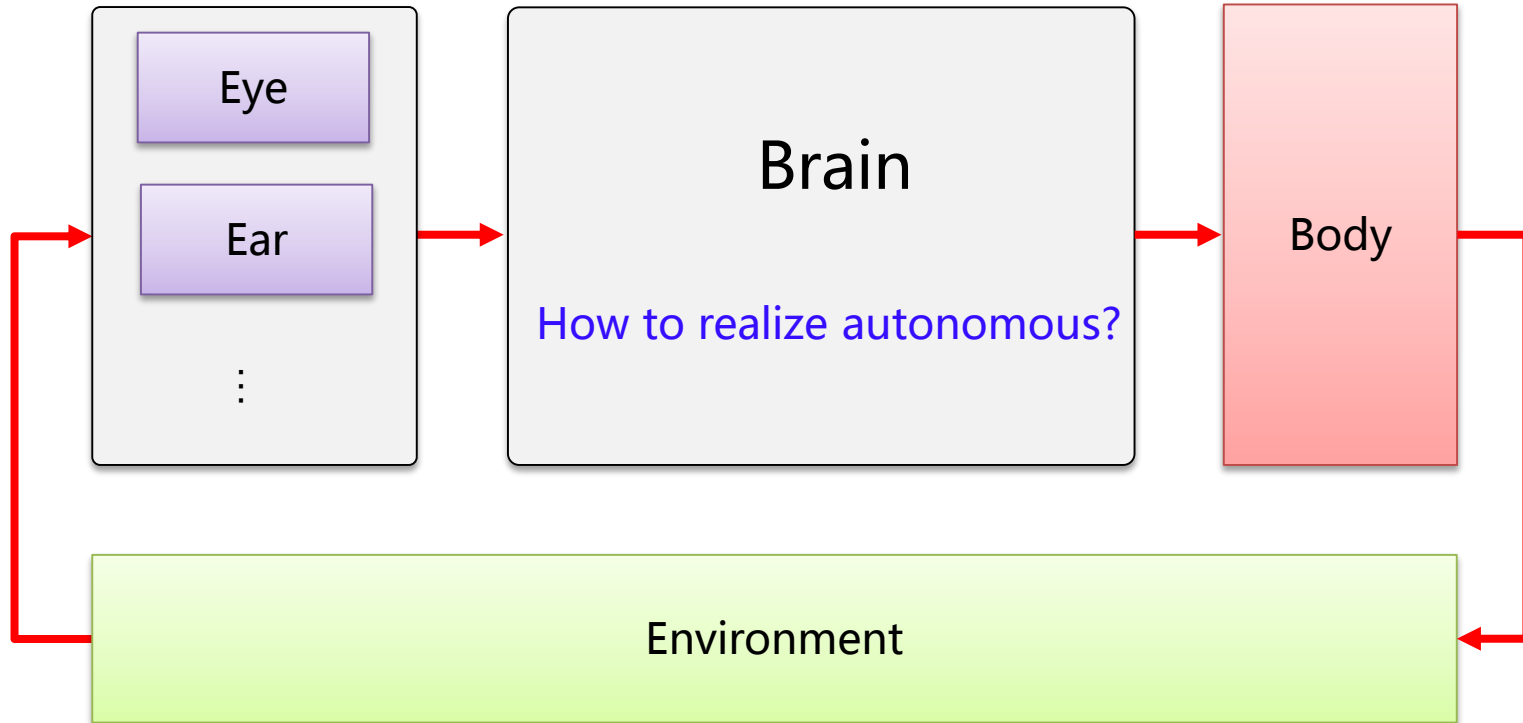
- Aerial photograph
- Attack
- Air platform
- General aviation



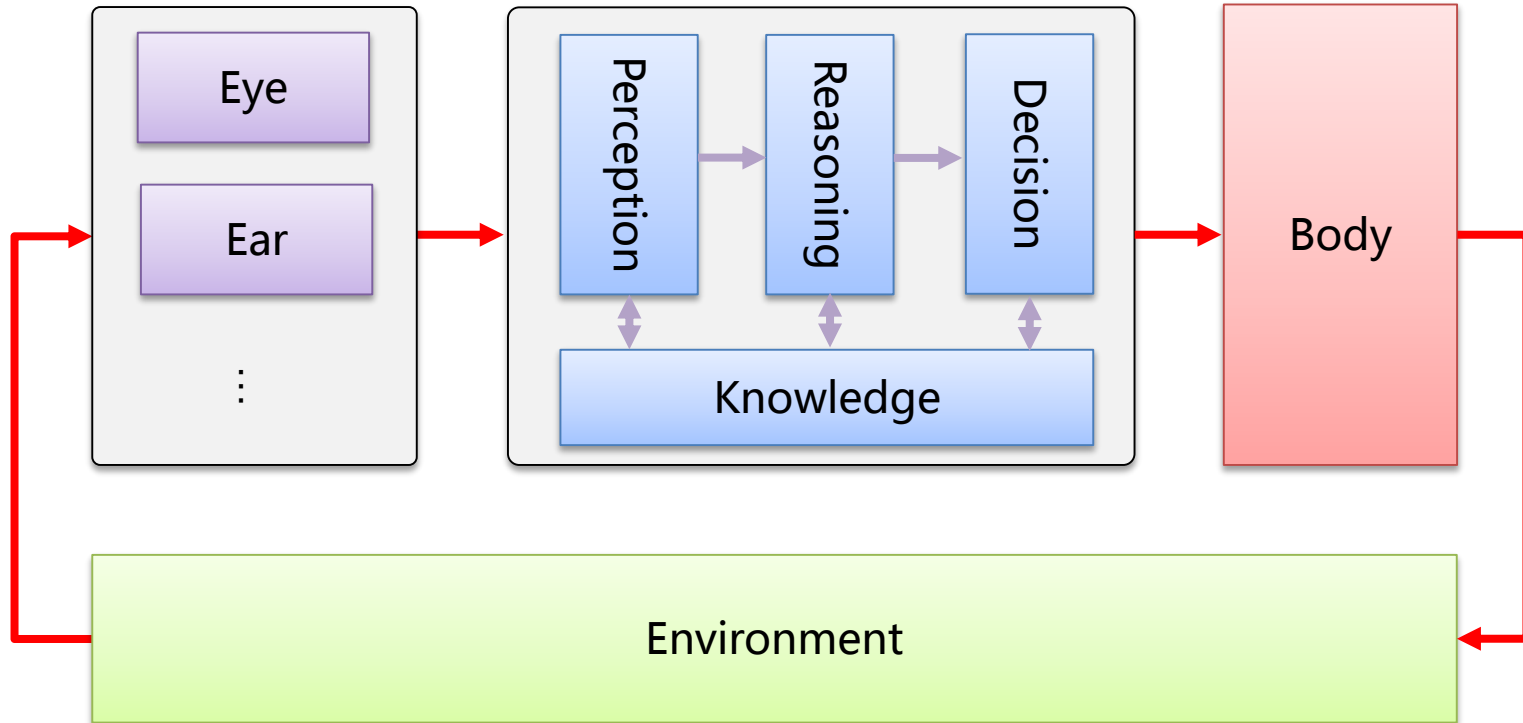
UAV - Autonomous



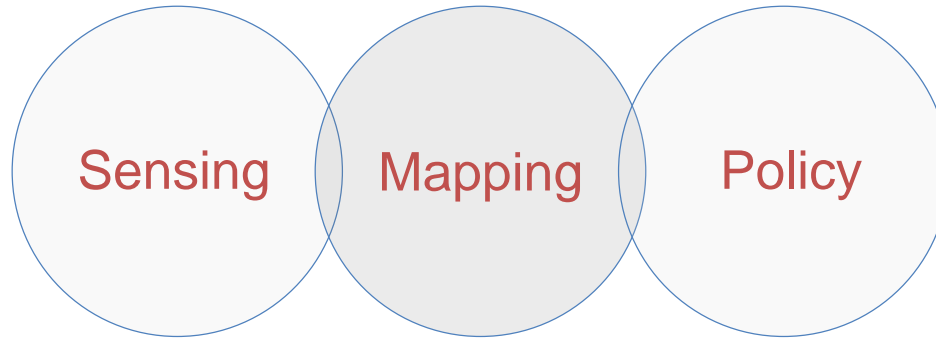
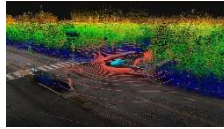
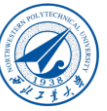
How to Realize Autonomous?



How to Realize Autonomous?

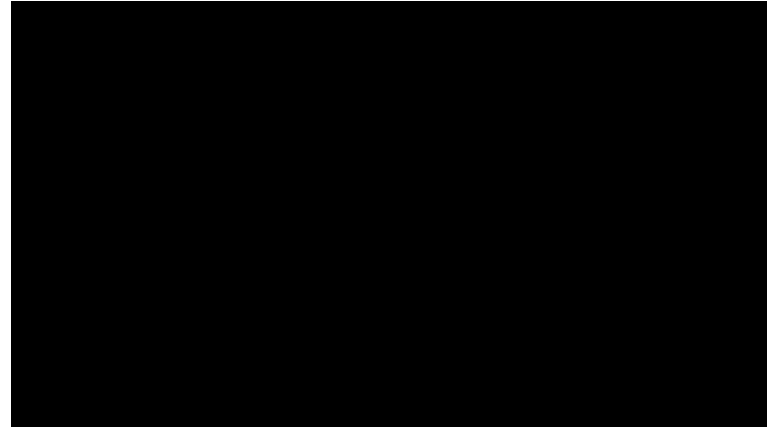


New Challenges

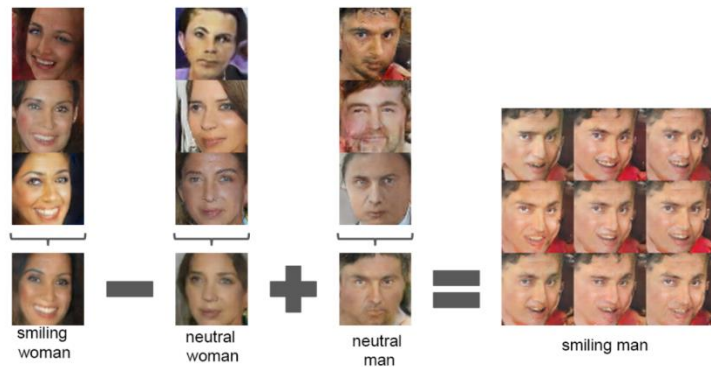
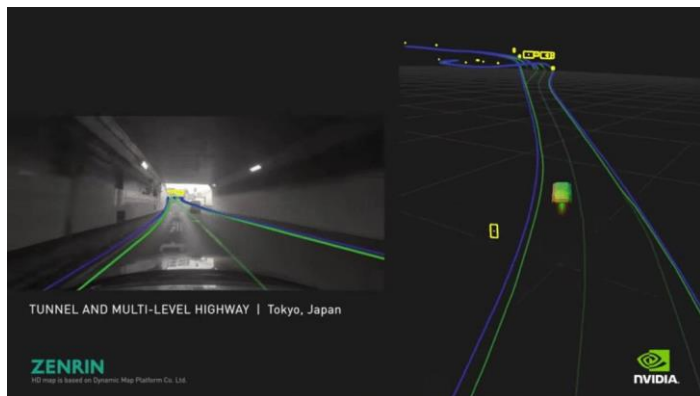


- Multi-type sensors: IMU, GPS, Image, LiDAR, RADAR ...
- High quality and real-time speed required
- Reasoning and knowledge are important for realizing strong AI

Applications



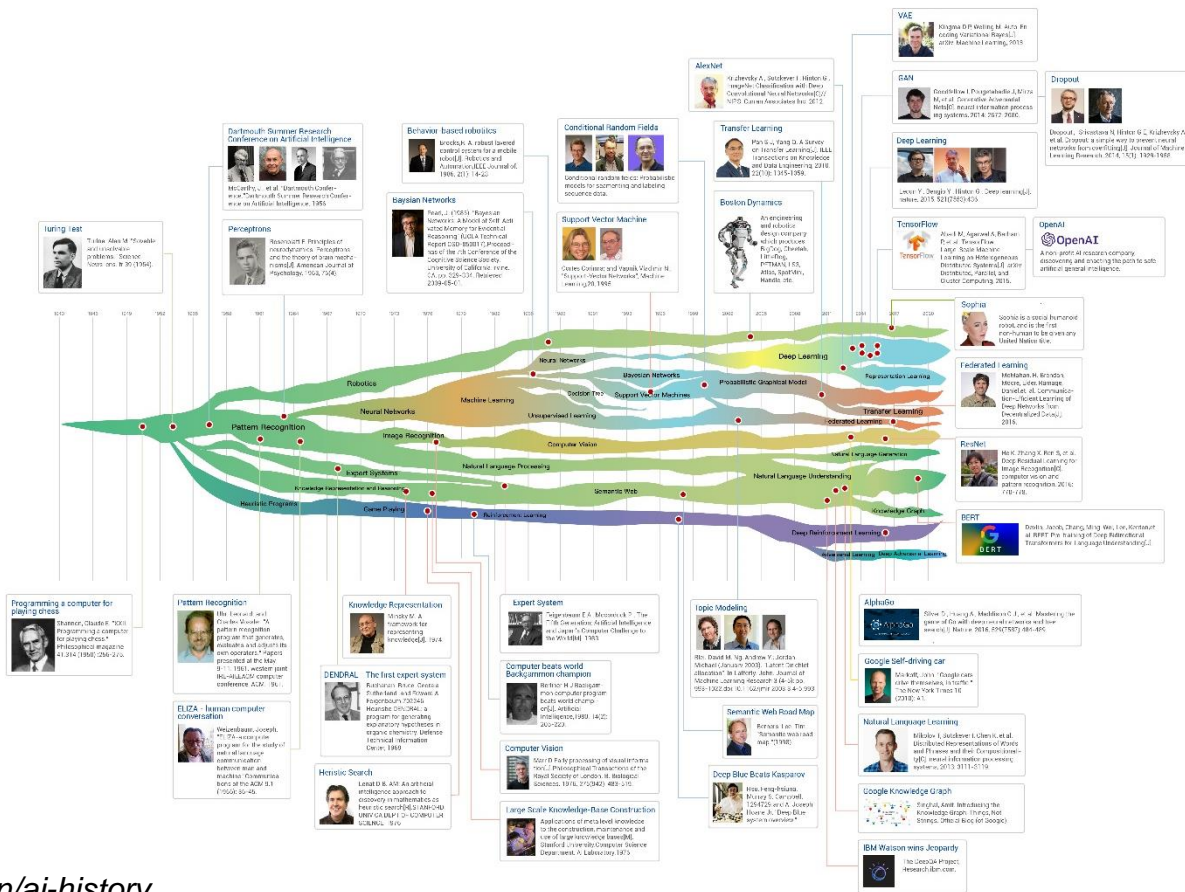
Applications



Applications



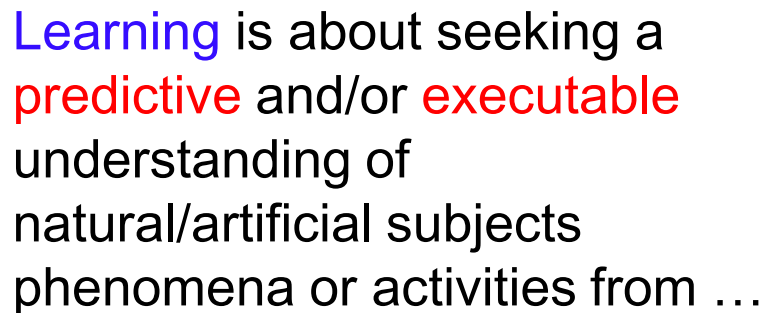
History of AI





西北工业大学
NORTHWESTERN POLYTECHNICAL UNIVERSITY

How to Achieve Intelligence?



Learning is about seeking a predictive and/or executable understanding of natural/artificial subjects phenomena or activities from ...

What is Machine Learning?

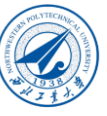
Machine learning seeks to develop theories and computer systems for

- Representing
- Classifying, clustering, recognizing
- Reasoning under uncertainty
- Predicting
- And reacting to
- ...

Complex, real world data, based on the **system's own experience with data**, and (hopefully) under a **unified model or mathematical framework**, that

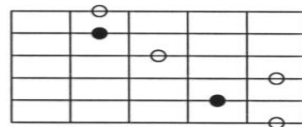
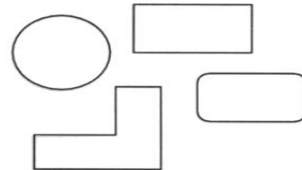
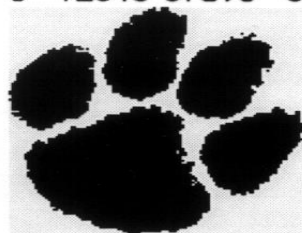
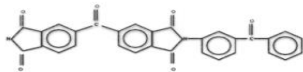
- Can be formally characterized and analyzed
- Can take into account human prior knowledge
- Can generalize and adapt across data and domains
- Can operate automatically and autonomously
- And can be interpreted and perceived by human

Pattern?



“A pattern is the opposite of a chaos; it is an entity vaguely defined, that could be given a name.”

- Watanabe

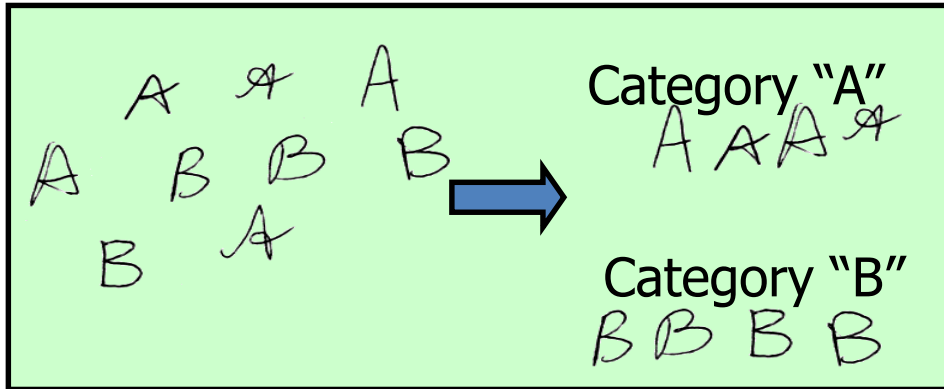


Recognition

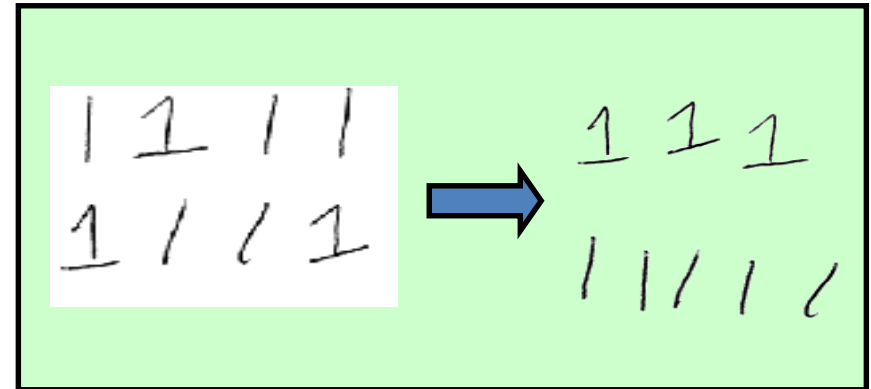


Identification of a pattern as a member of a category we already know, or we are familiar with

- **Classification** (known categories)
- **Clustering** (learning categories)



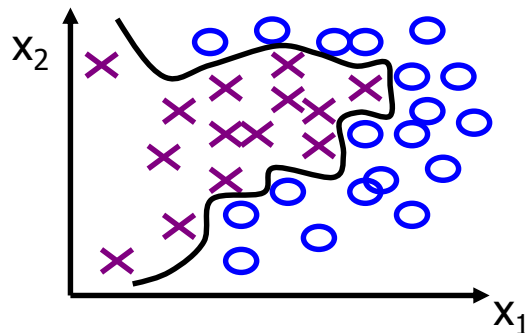
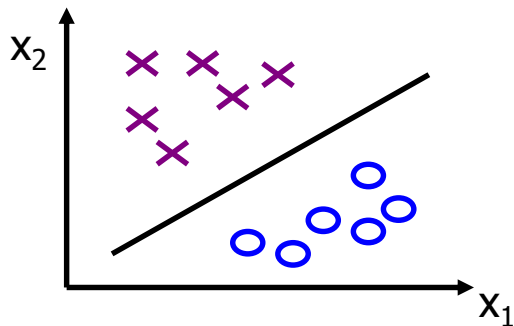
Classification



Clustering

Representation

- Each pattern is represented as a **point in d -dimensional feature space**
- Choice of features and their desired invariance properties are domain-specific



- Good representation implies (i) small intra-class variation, (ii) large inter-class separation and (iii) simple decision boundary

Pattern Class

- A collection of similar (not necessarily identical) objects
- A class is defined by class samples (paradigms, exemplars, prototypes, training/learning samples)
- Intra-class variability
- Inter-class similarity
- How do we define similarity?

Intra-class Variability



The letter “T” in different typefaces

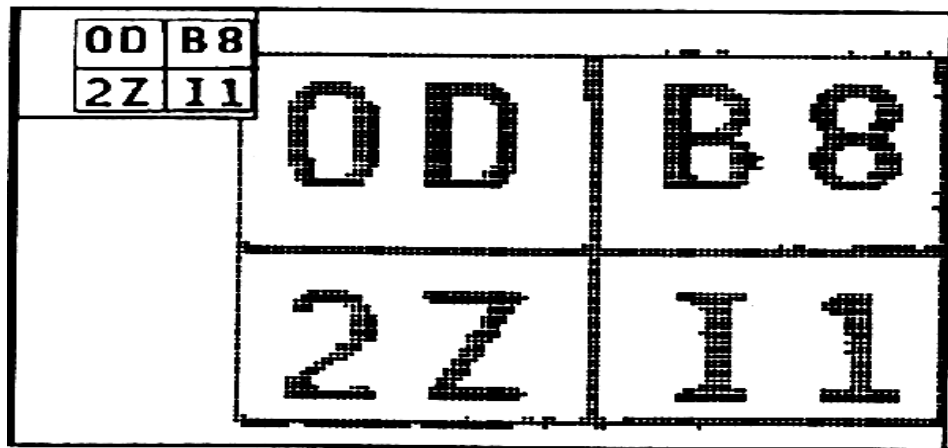


Same face under different expression, pose, illumination

Inter-class Similarity

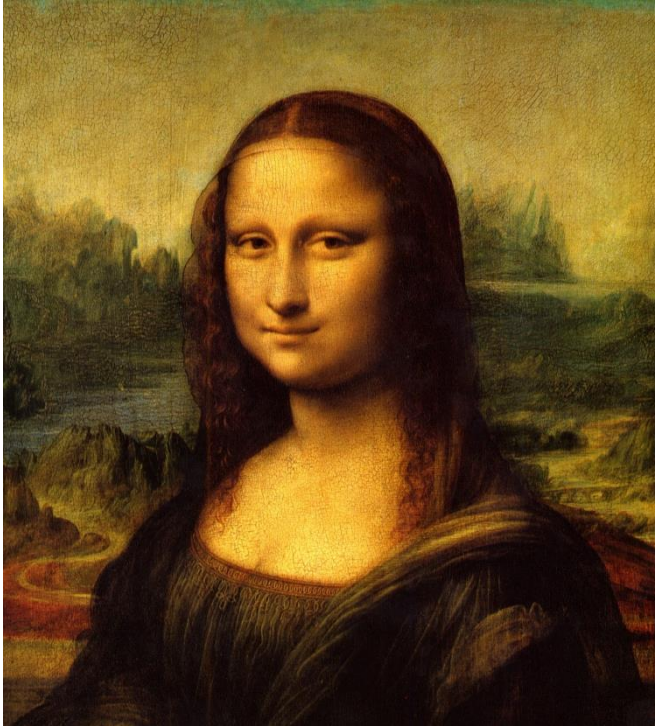


Identical twins



Characters that look similar

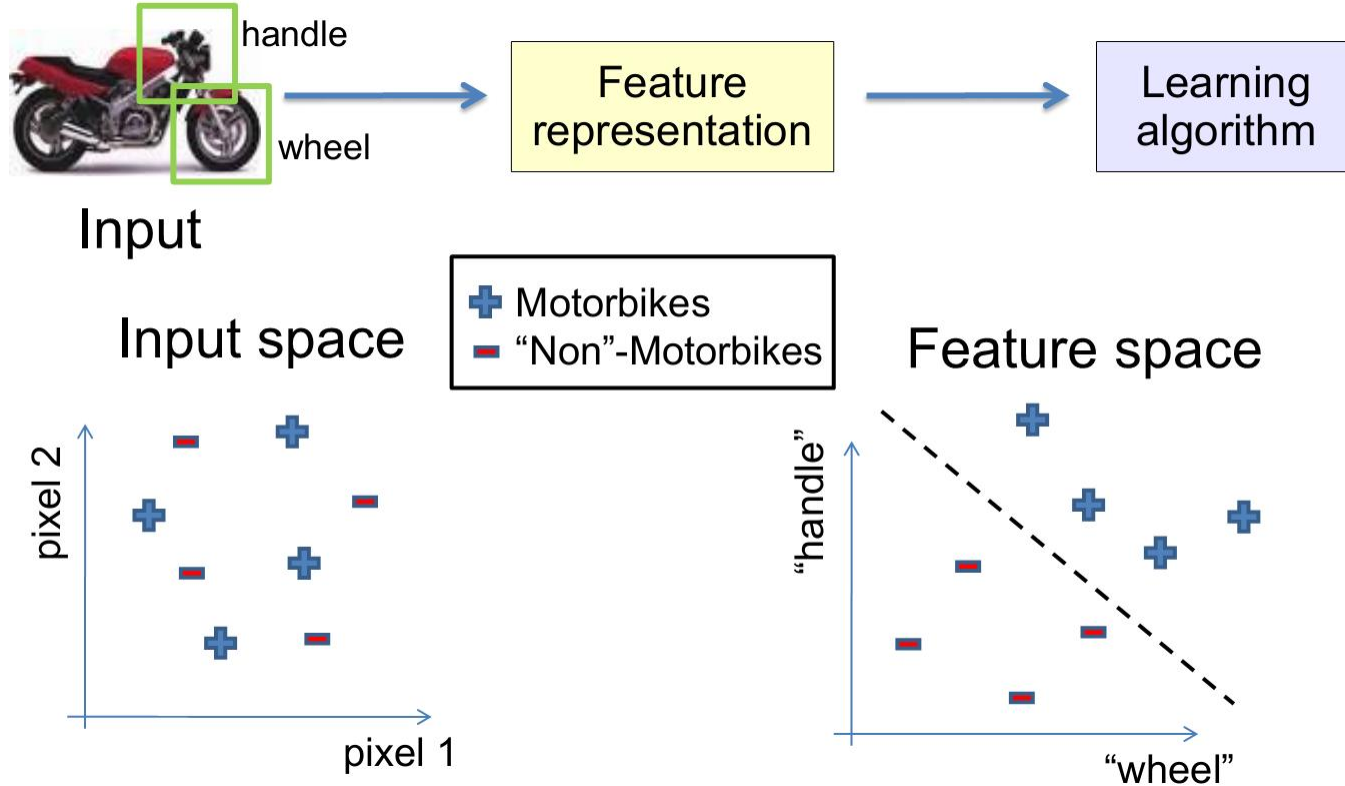
Inter-class or Intra-class?



Good Representation

- Should have some **invariant** properties (e.g., rotation, translation, scale, ...)
- Account for intra-class variations
- Ability to discriminate pattern classes of interest
- Robustness to noise, occlusion,...
- Lead to simple decision making strategies (e.g., linear decision boundary)
- Low measurement cost; real-time

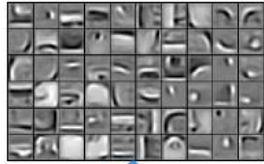
Good Representation



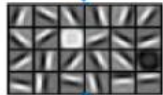
Good Representation



3rd layer
“Objects”



2nd layer
“Object parts”



1st layer
“edges”



Input

- Represent objects from low-level to high-level structure
- Can share the low-level representation for multiple tasks

Deep Learning = Learning Hierarchical Representation

Traditional Pattern Recognition: Fixed/Handcrafted Feature Extractor



Mainstream Modern Pattern Recognition: Unsupervised mid-level features

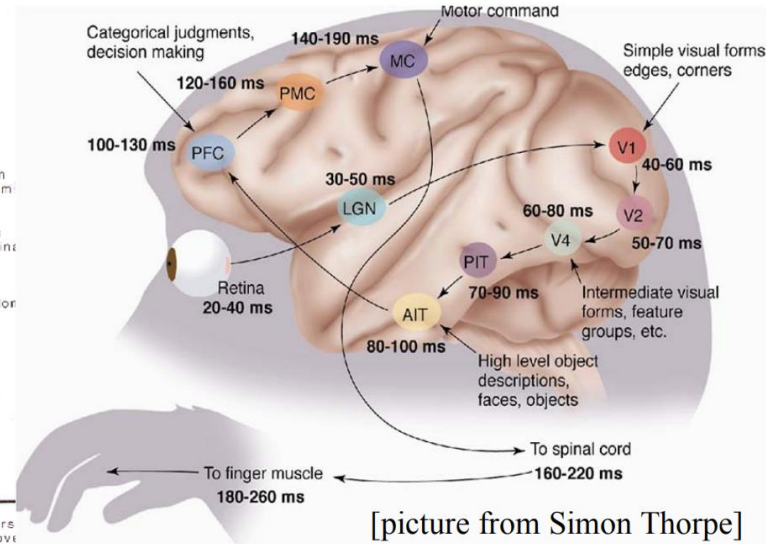
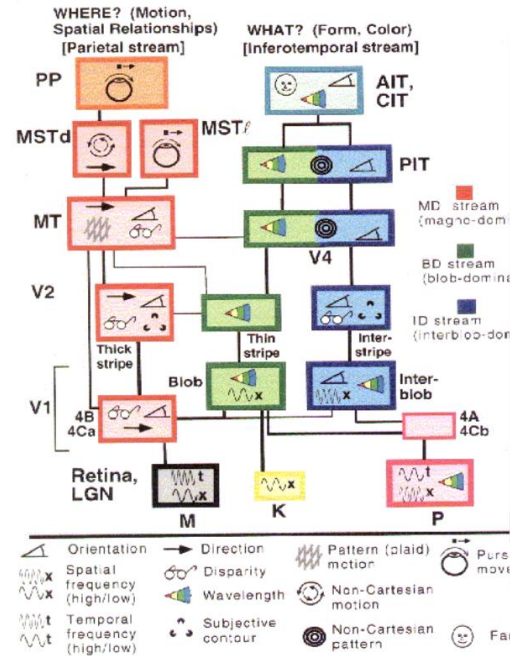


Deep Learning: Representations are hierarchical and trained



The Mammalian Visual Cortex is Hierarchical

- The ventral (recognition) pathway in the visual cortex has multiple stages
- Retina - LGN - V1 - V2 - V4 - PIT - AIT
- Lots of intermediate representations

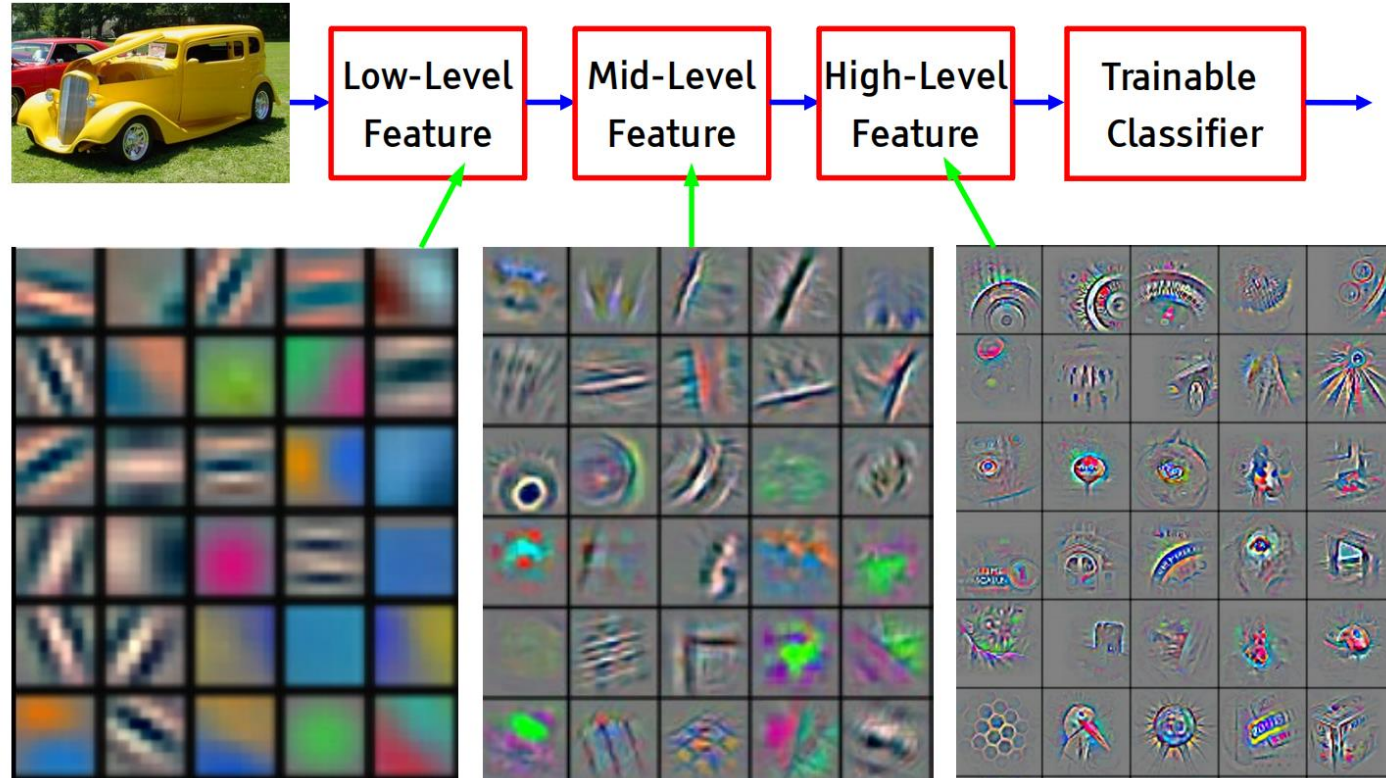


[picture from Simon Thorpe]

[Gallant & Van Essen]

Deep Learning = Learning Hierarchical Representation

It's deep if it has more than one stage of non-linear feature transformation



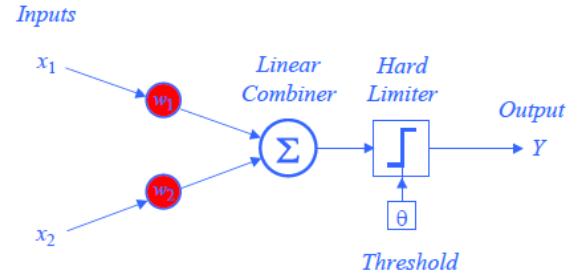
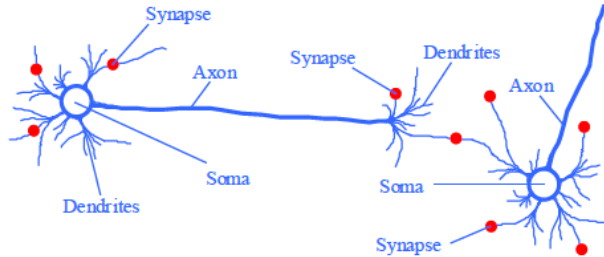


西北工业大学
NORTHWESTERN POLYTECHNICAL UNIVERSITY

Neural Networks and Deep Learning

Perceptron and Neural Networks

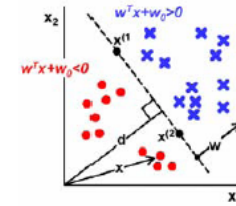
- From biological neuron to artificial neuron (perceptron)



- Activation function

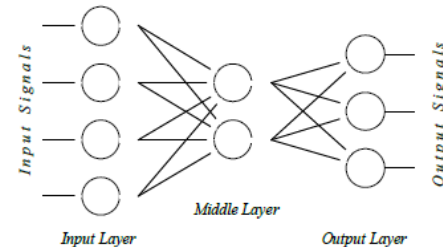
$$X = \sum_{i=1}^n x_i w_i$$

$$Y = \begin{cases} +1, & \text{if } X \geq \omega_0 \\ -1, & \text{if } X < \omega_0 \end{cases}$$

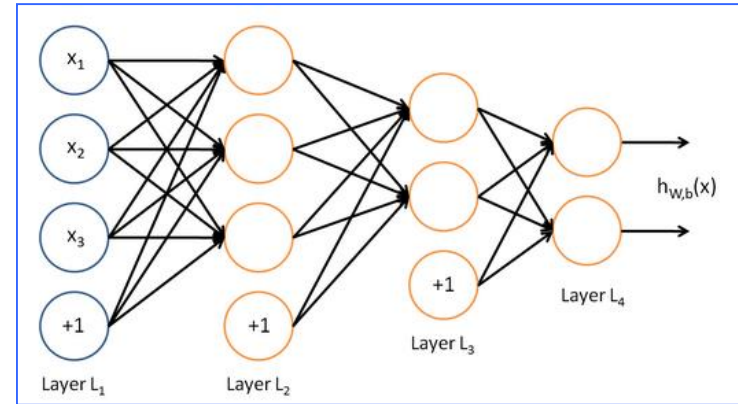
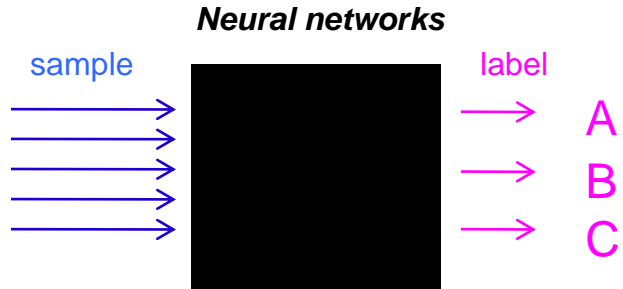


- Artificial neuron networks

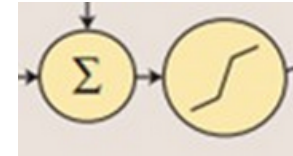
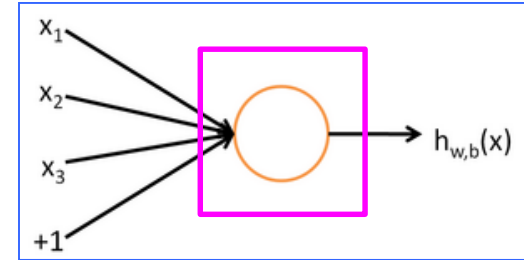
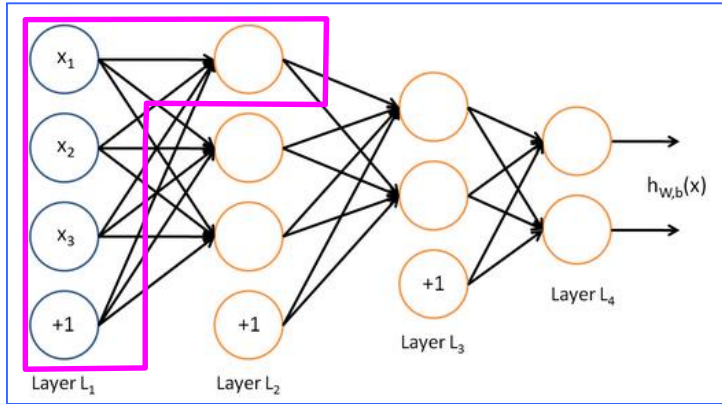
- supervised learning
- gradient descent



Neural Networks



Neural Networks - Feedforward Networks



$$h_{W,b}(x) = f(W^T x) = f(\sum_{i=1}^3 W_i x_i + b)$$

$$f(z) = \frac{1}{1 + \exp(-z)}$$

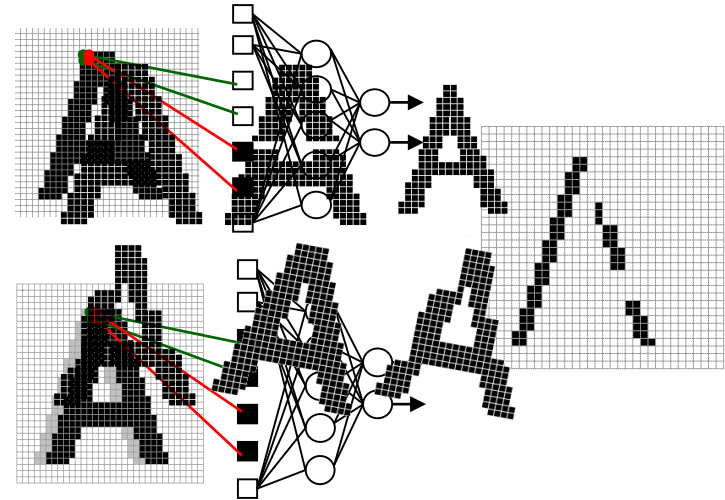
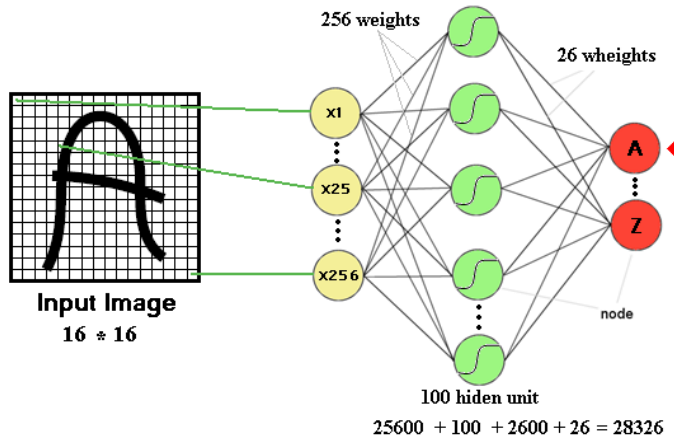
Neural Networks - Feedforward Networks



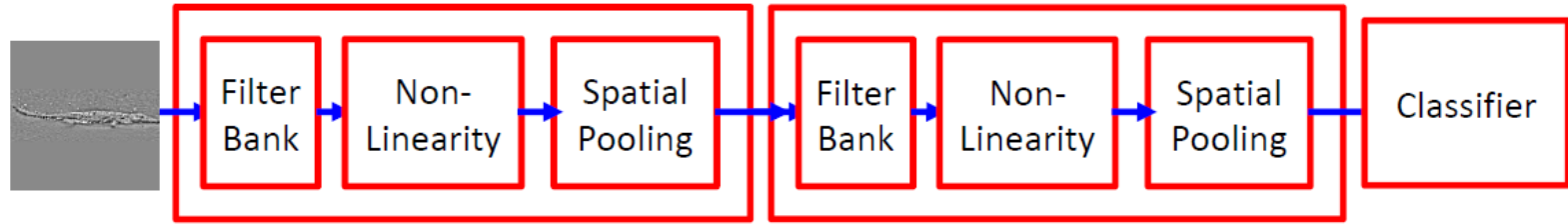
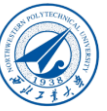
www.cybercontrols.org

Neural Networks - Disadvantages

- The number of **trainable parameters** becomes extremely **large**
- Little **or no invariance** to shifting, scaling, and other forms of distortion

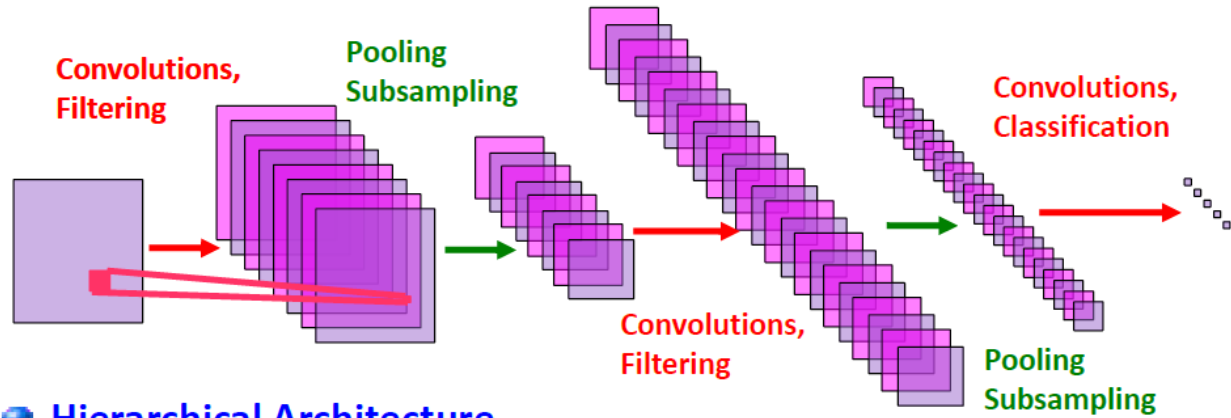
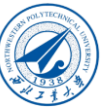


Hierarchical / Deep Architectures for Vision



- Multiple Stages
- Each Stage is composed of
 - A bank of local filters (convolutions)
 - A non-linear layer (may include harsh non-linearities, such as rectification, contrast normalization, etc...).
 - A feature pooling layer
- Multiple stages can be stacked to produce high-level representations
 - Each stage makes the representation more global, and more invariant
- The systems can be trained with a combination of unsupervised and supervised methods

Convolutional Neural Network: Multi-stage Trainable Architecture



● Hierarchical Architecture

- ▶ Representations are more global, more invariant, and more abstract as we go up the layers

● Alternated Layers of Filtering and Spatial Pooling

- ▶ Filtering detects conjunctions of features
- ▶ Pooling computes local disjunctions of features

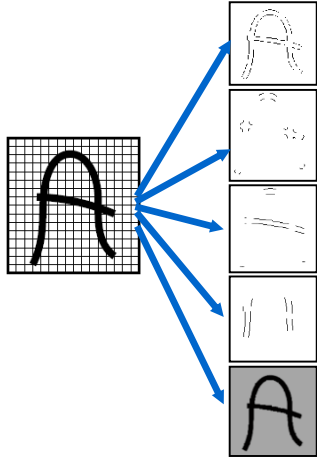
● Fully Trainable

- ▶ All the layers are trainable

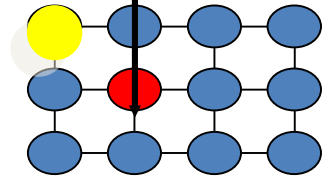
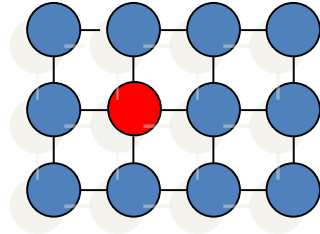
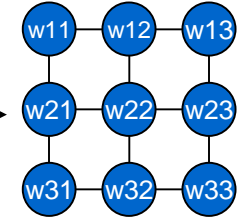
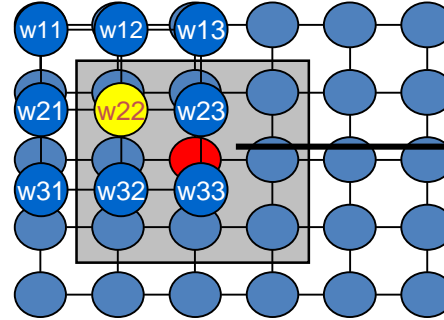
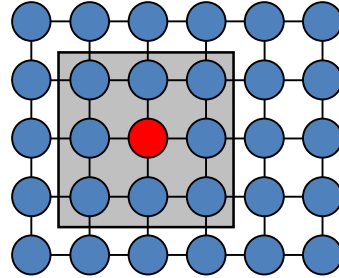
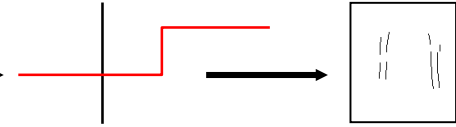
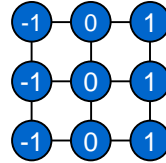
Convolutional Layer or Feature Extraction Layer



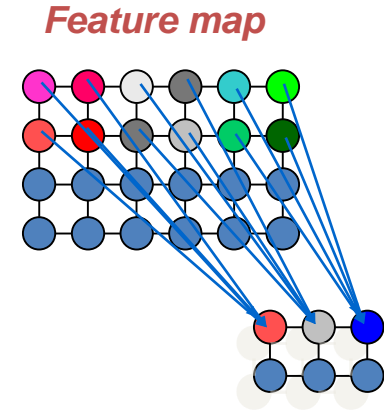
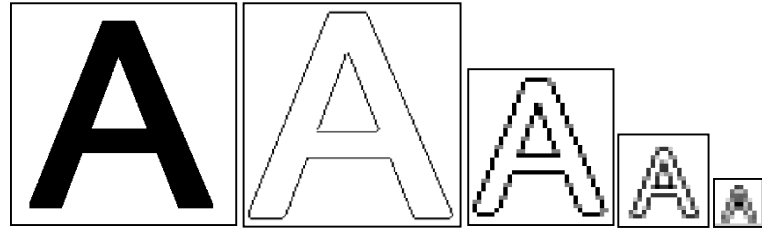
features



Convolve with



Subsampling Layer

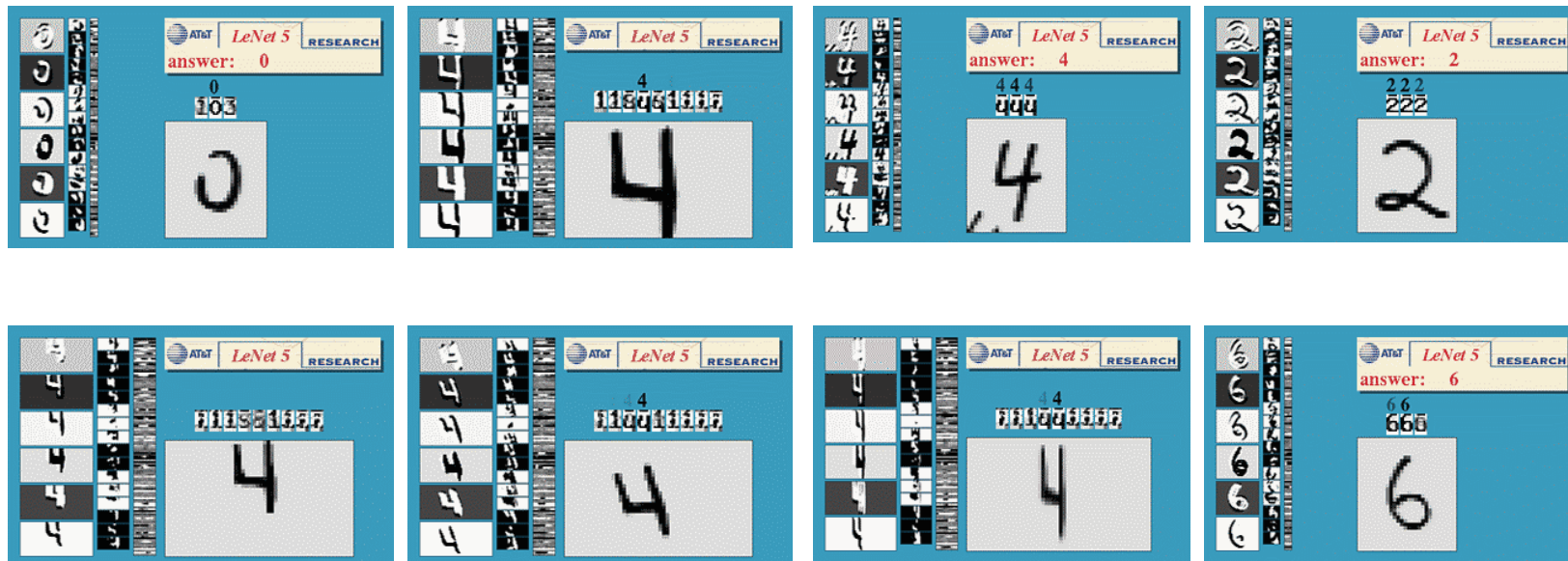


- By reducing the **spatial resolution** of the feature map, reduce the effect of **noises** and **shift or distortion**.
- The **weight sharing** is also applied in subsampling layers

Convolutional Neural Networks - Demo



Convolutional Neural Networks - Hand-Writing Recognition





西北工业大学
NORTHWESTERN POLYTECHNICAL UNIVERSITY

Applications

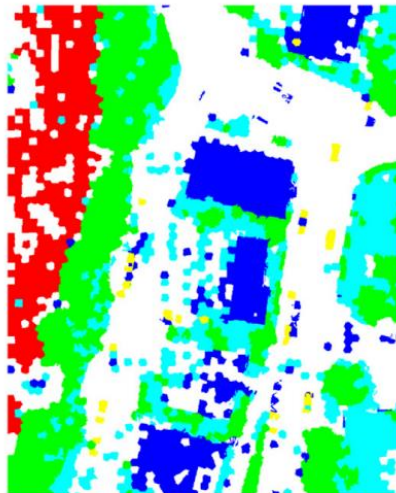
Application (1) Image Annotation



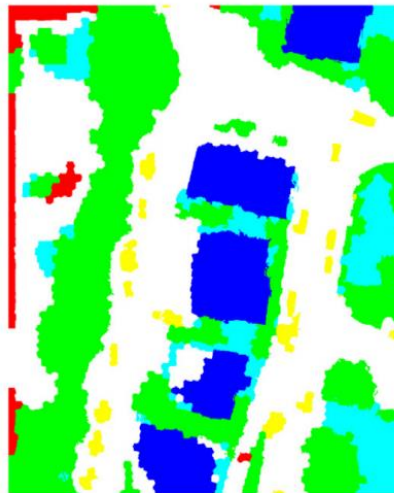
imp_surf Building Low vegetation Tree Car Clutter



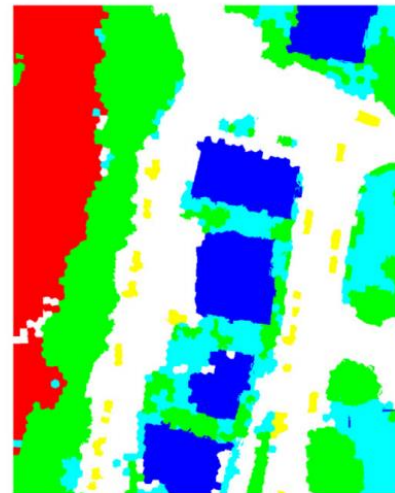
(a) original image



(b) shallow features

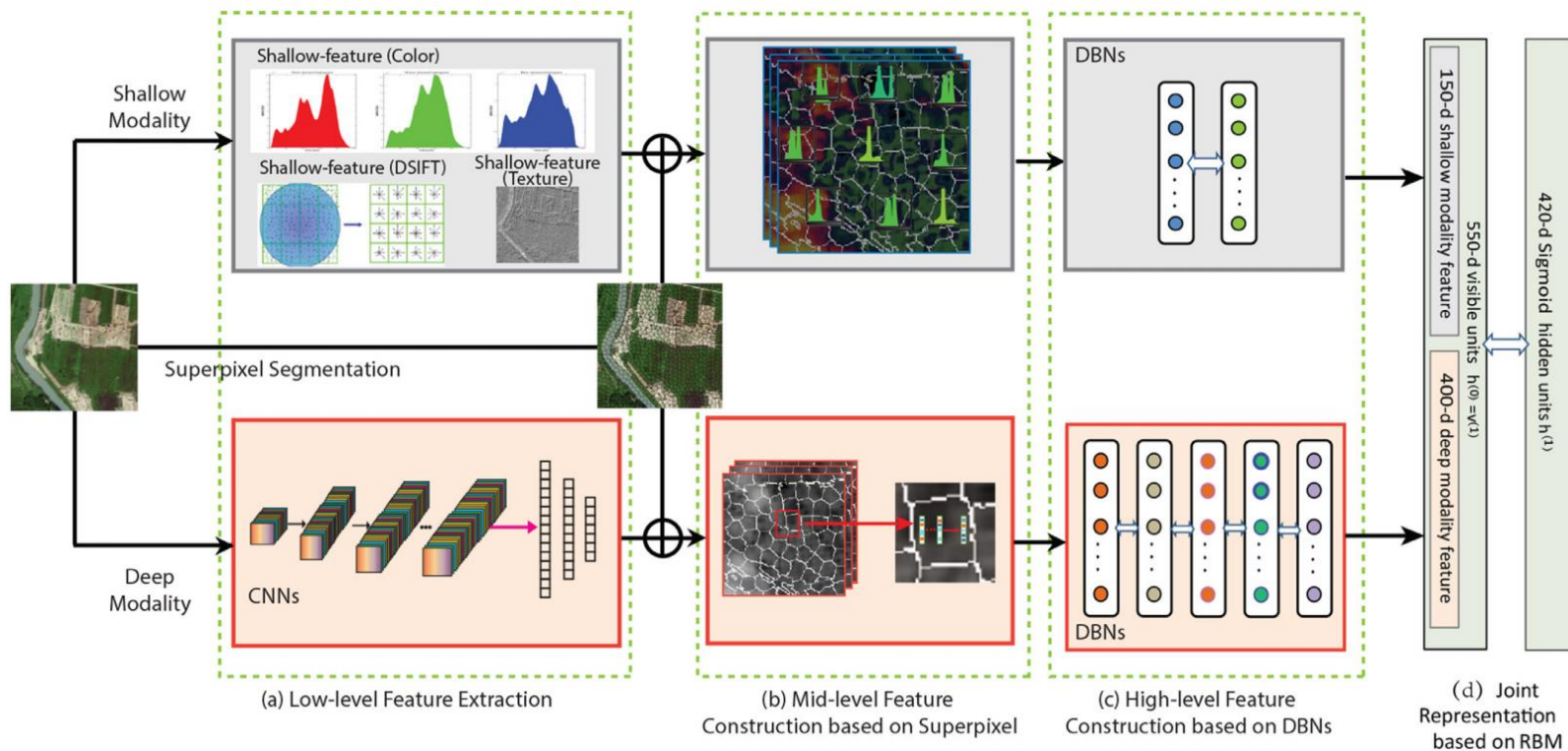


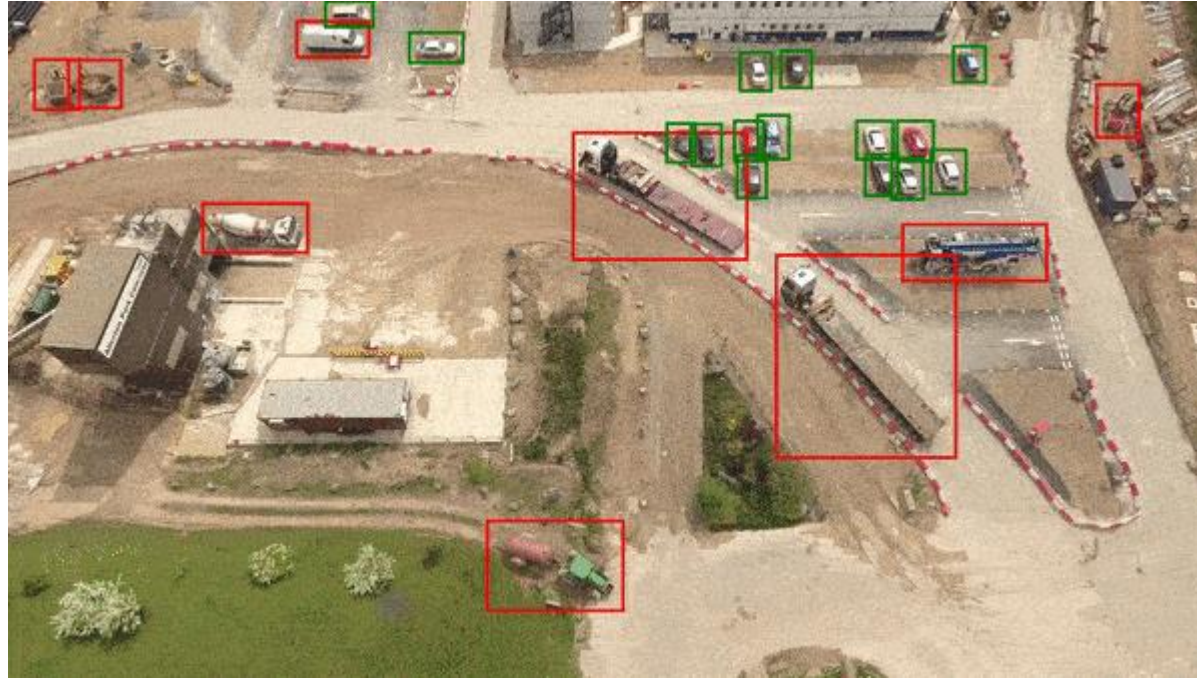
(c) deep features



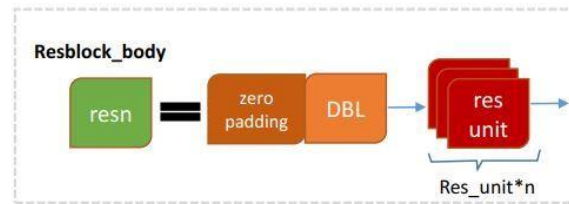
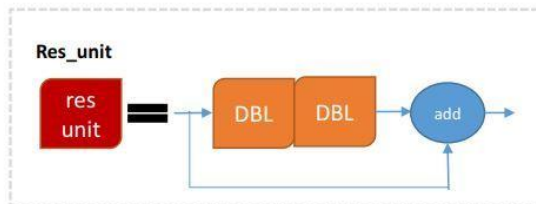
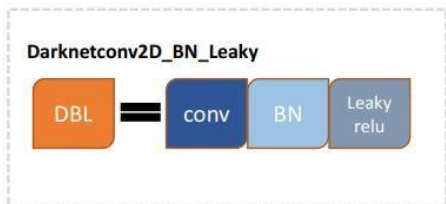
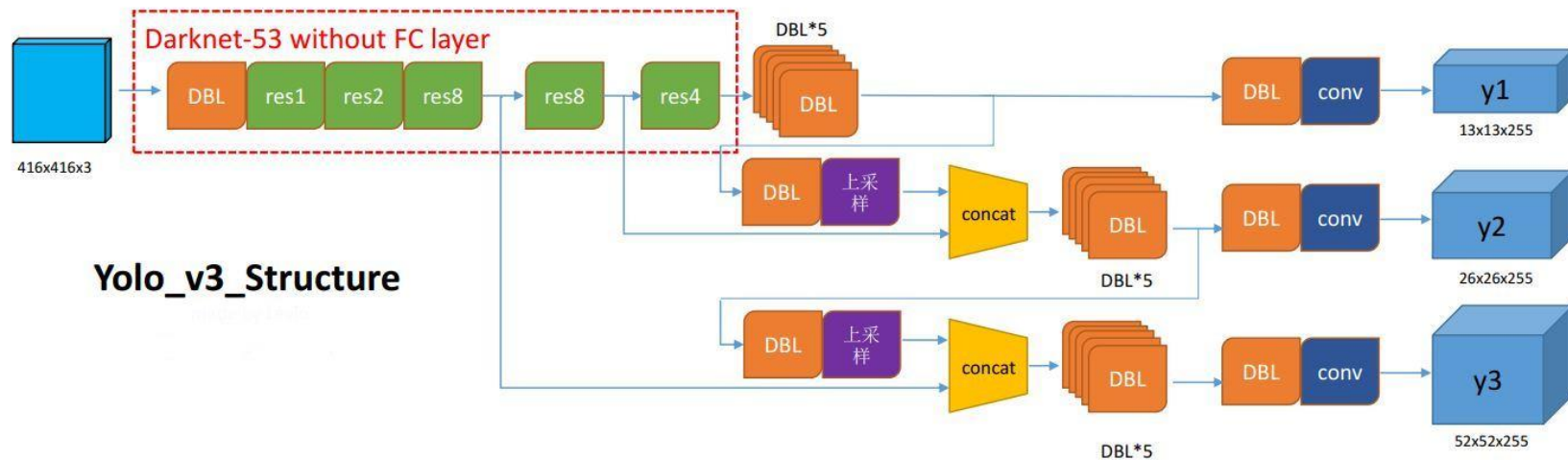
(d) multi-modal features

Application (1) Image Annotation





Application (2) Object Detection

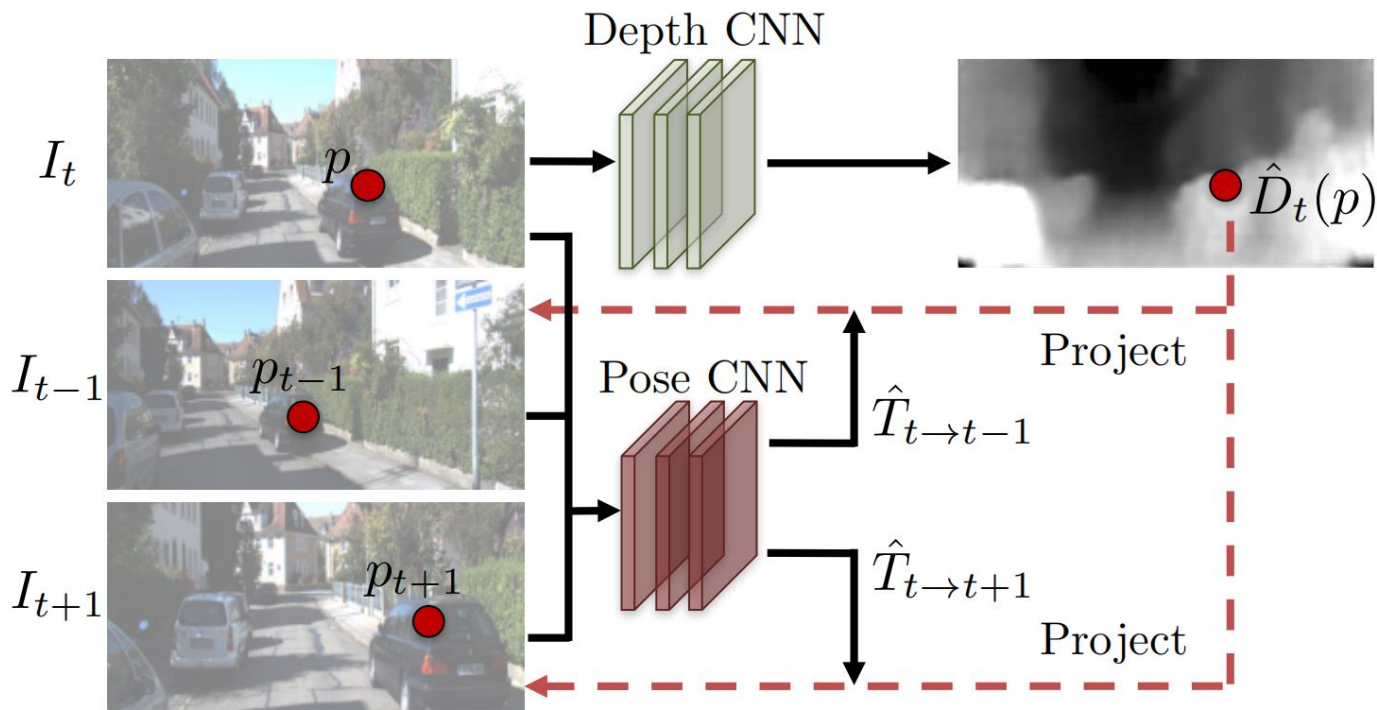


Application (3) Deep Learning based SLAM

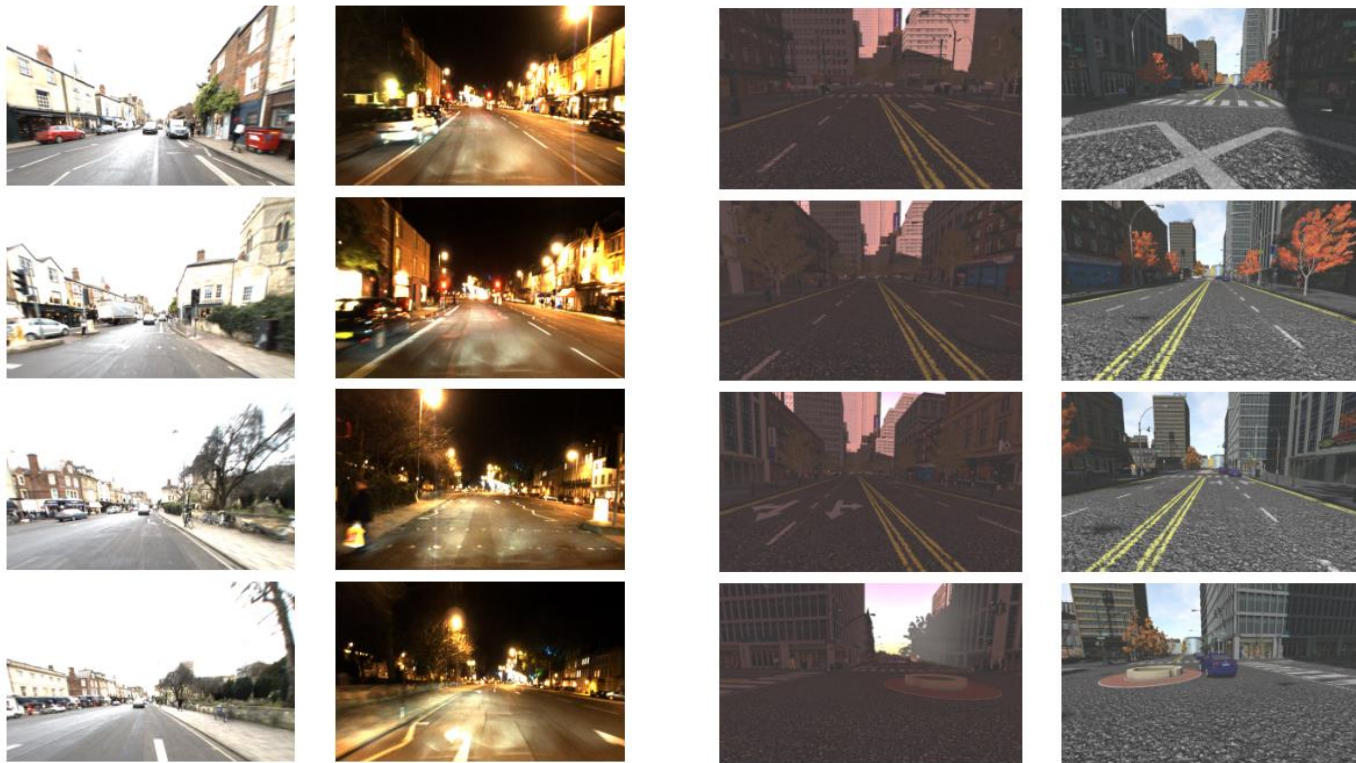


Estimate position and depth image simultaneously

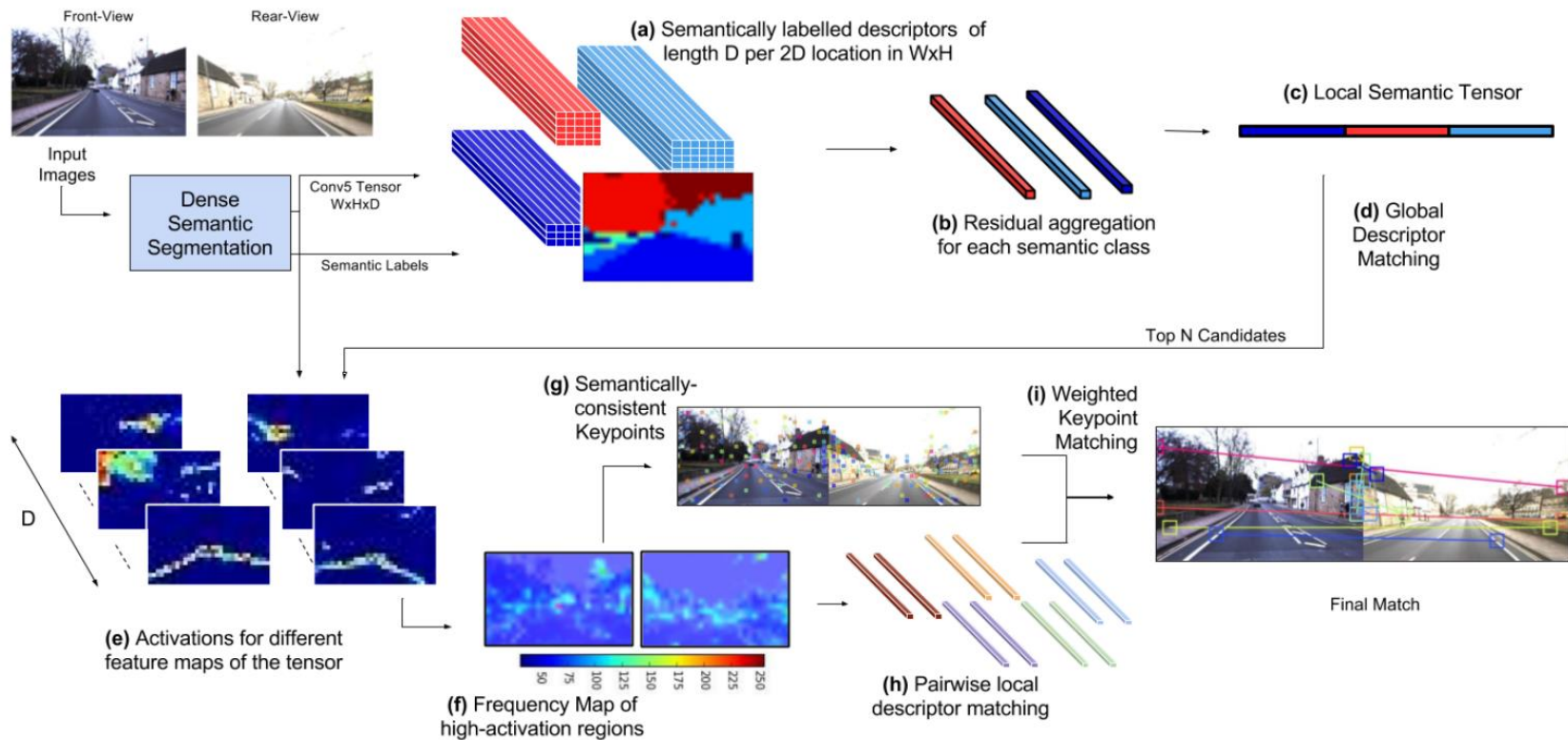
Application (3) Deep Learning based SLAM



Application (4) Place Recognition



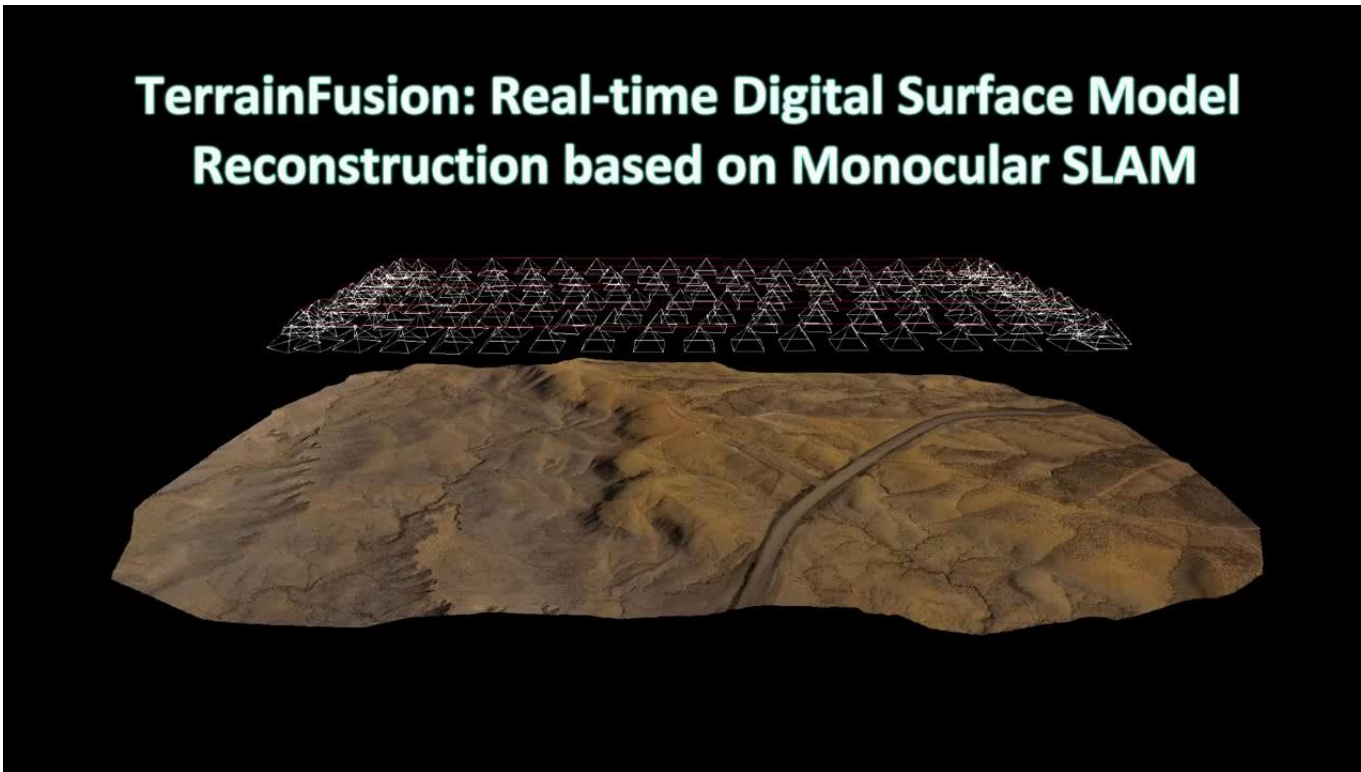
Application (4) Place Recognition



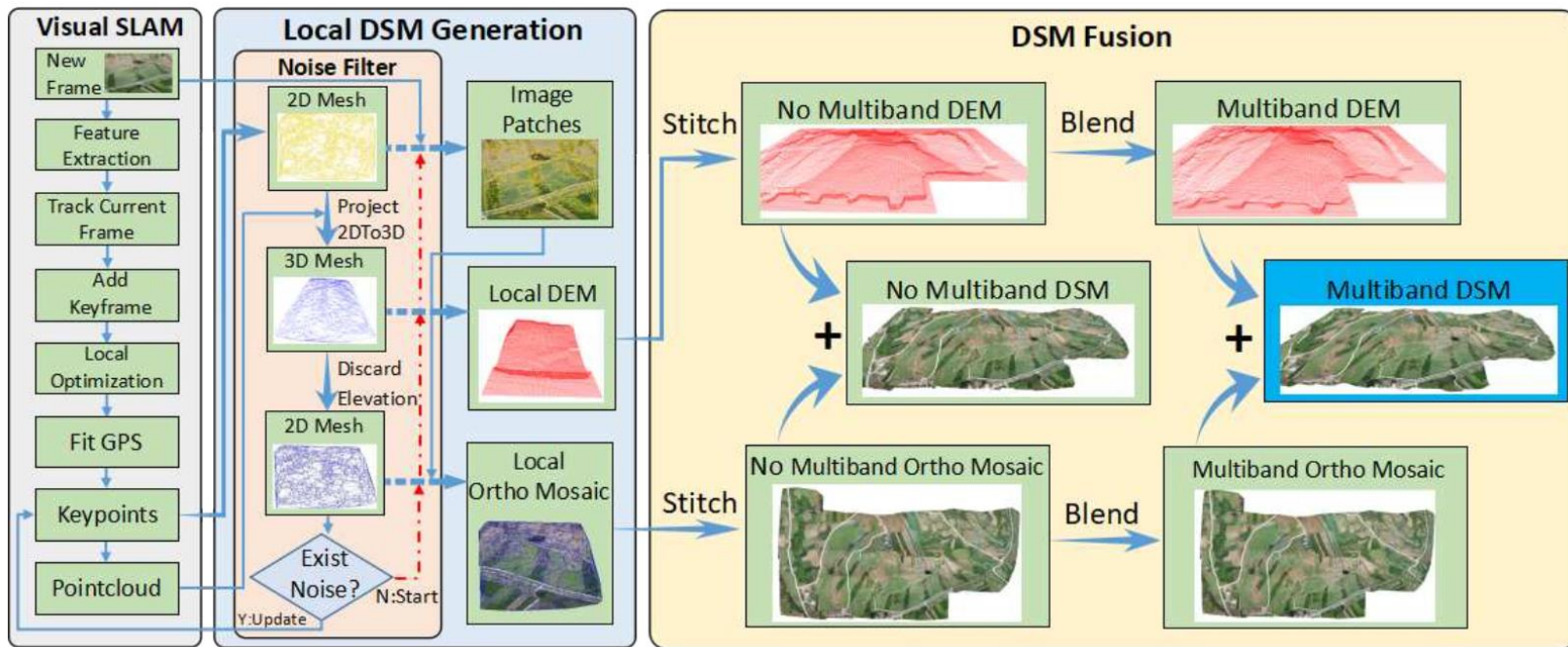
Application (5) Realtime 3D Map



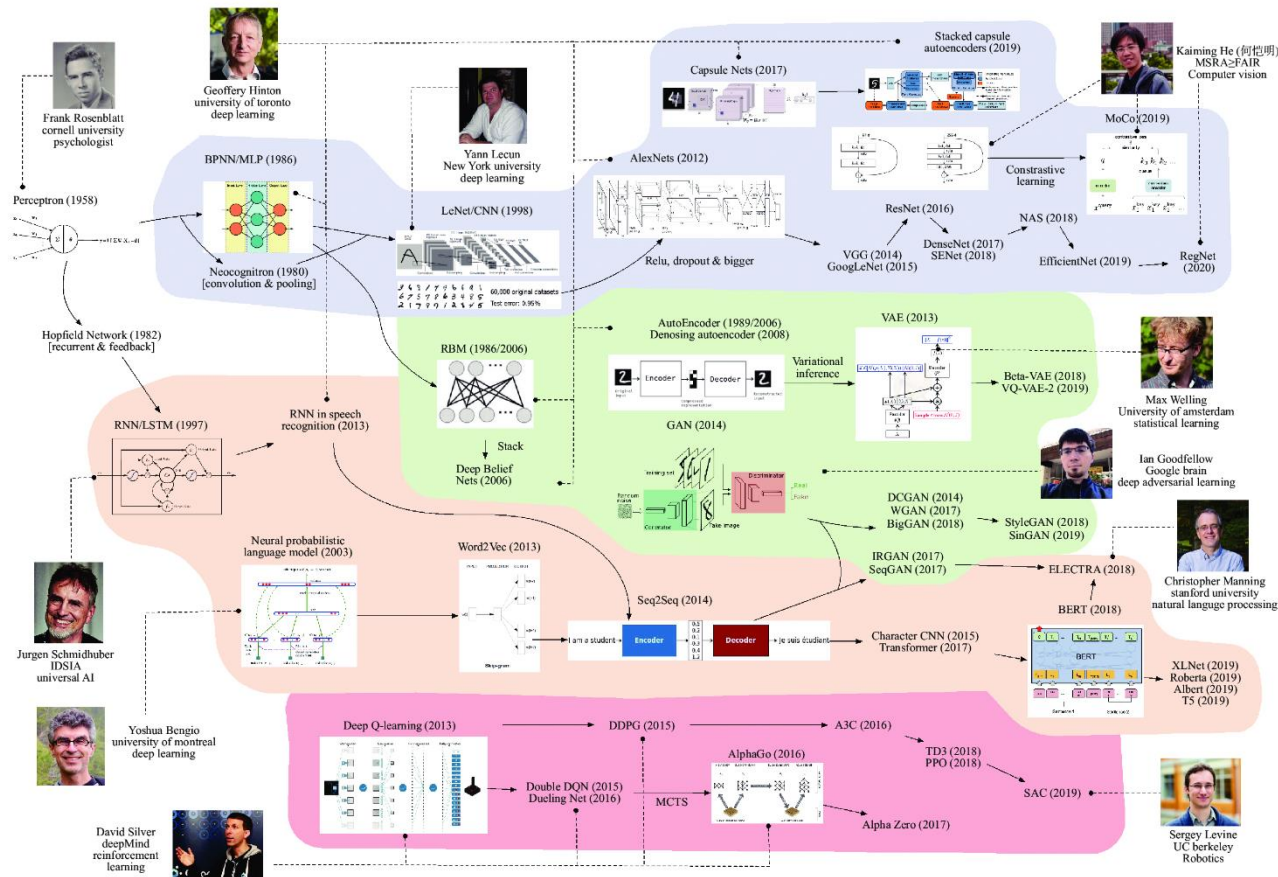
TerrainFusion: Real-time Digital Surface Model Reconstruction based on Monocular SLAM



Application (5) Realtime 3D Map



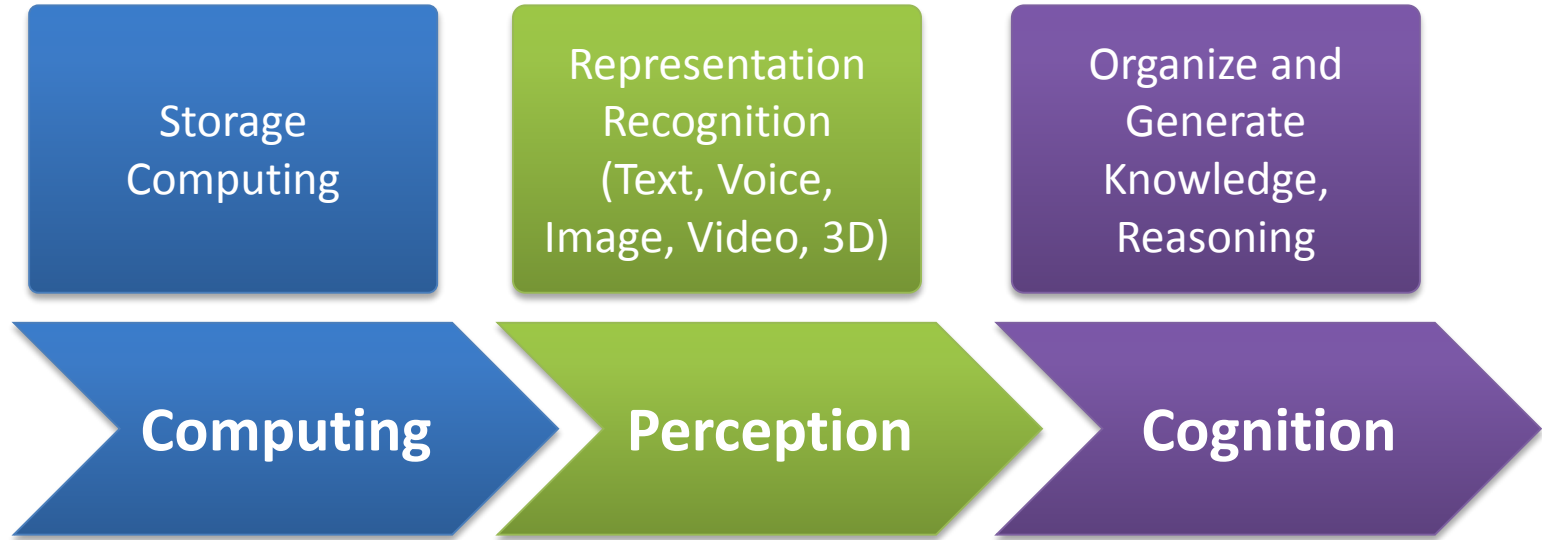
Summary - Methods



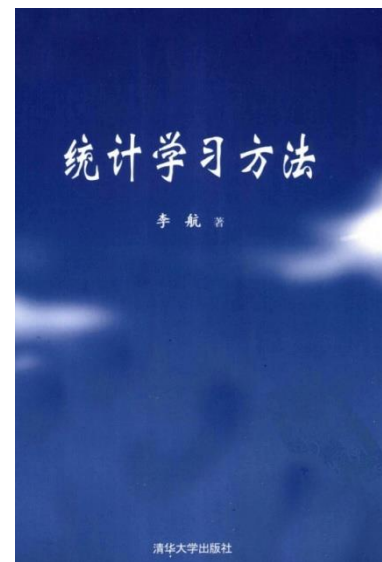
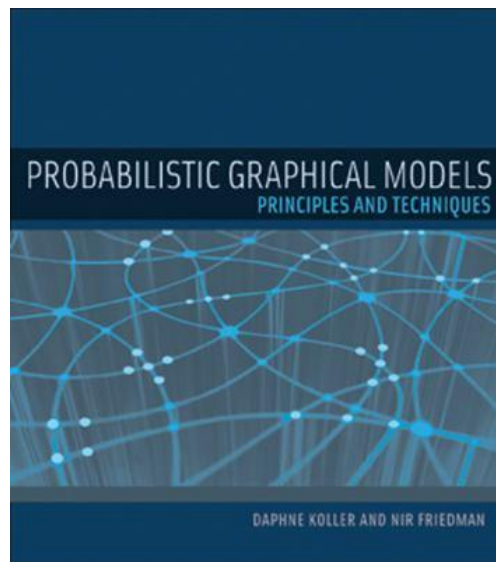
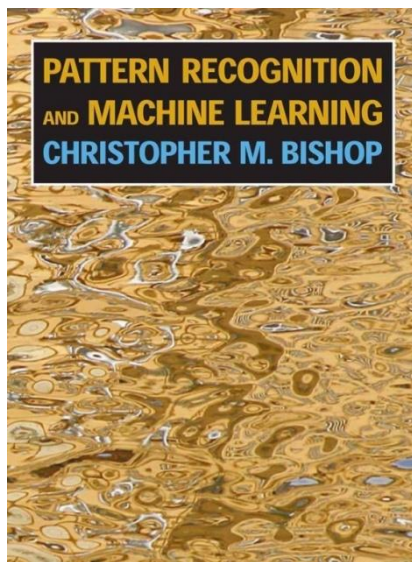
Summary - Problems

- General-purpose intelligent system is a very difficult problem
- Successful systems available in well-constrained domains
- All components are coupled
- No single approach has been found to be optimal for all problems
- Use of object models, constraints and context is necessary for identifying complex patterns
- Careful sensor design and feature extraction often lead to simple classifiers

Summary - Future



Books



Materials



Notebook: https://gitee.com/pi-lab/machinelearning_notebook

Homework: https://gitee.com/pi-lab/machinelearning_homework



Notebook



Homework

THANK YOU



www.adv-ci.com