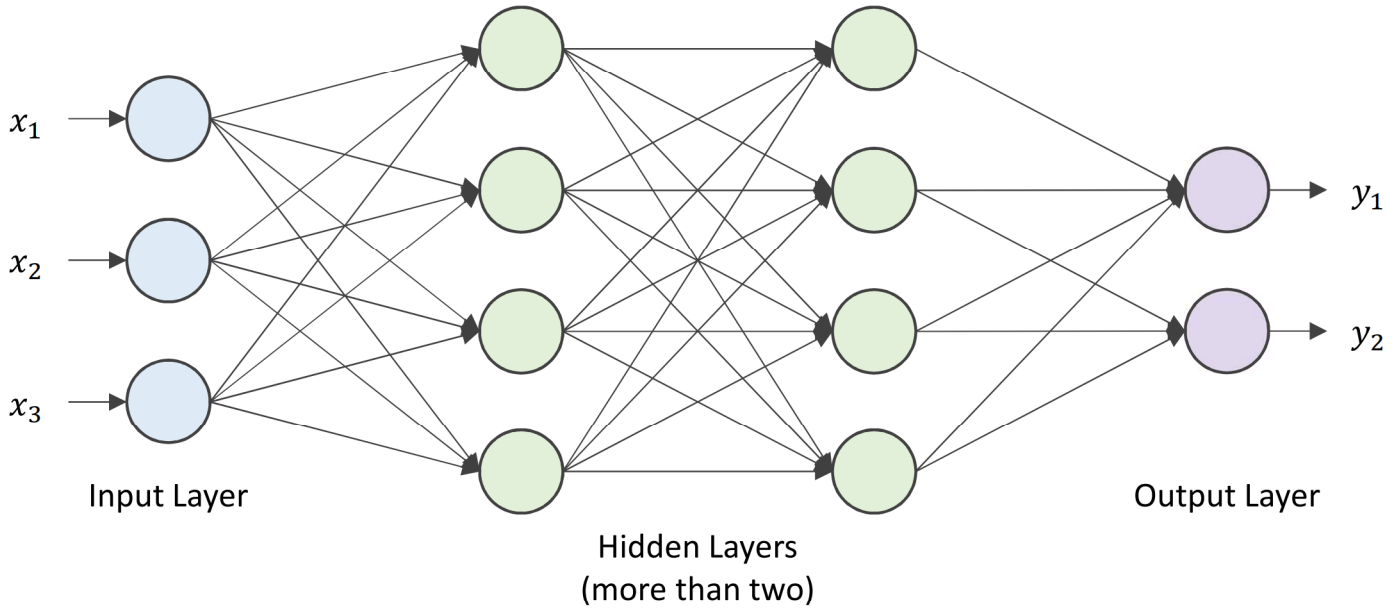
What is Deep Learning?

Definition

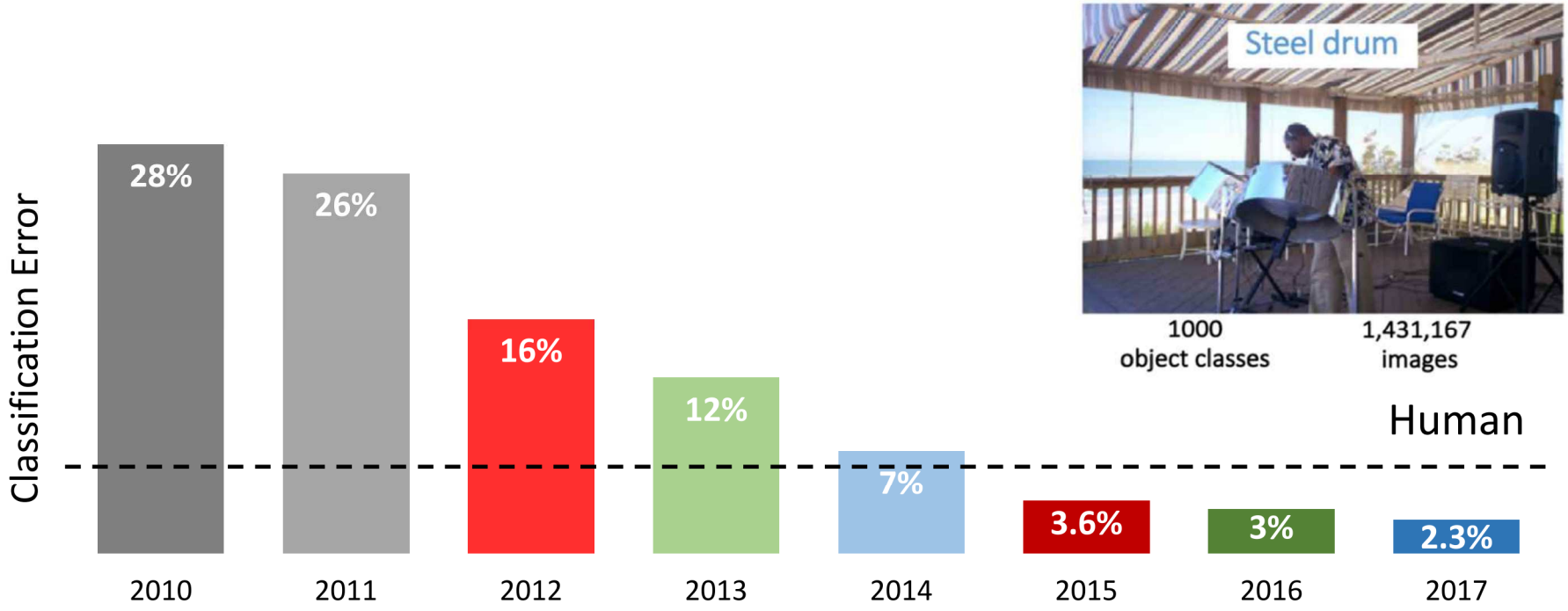
- A subfield of machine learning that is based on *deep neural networks* consisting of multiple hidden layers to mimic the structure and function of the human brain

Basic Architecture



Why Deep Learning?

- **Breakthrough in Image Classification Challenge** IMAGENET



(Source: Fei-Fei Li 2020)

The Deeper, The Better

#.#. Deep Learning

Deep Learning Revolution

- **Three Key Enablers**

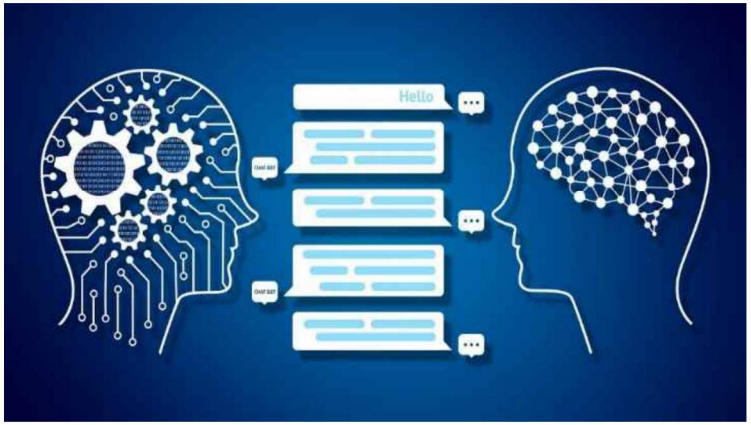
Computation

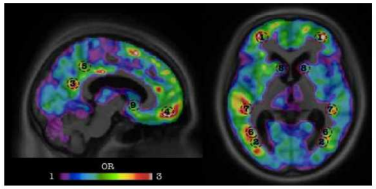
Big Data

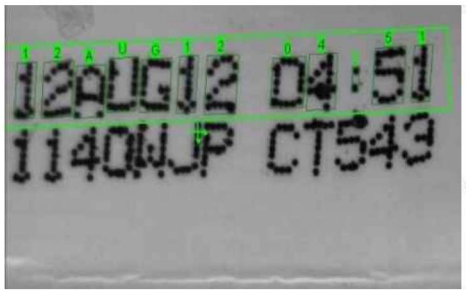
Algorithms

Deep Learning Revolution

- Applications in Daily Lives

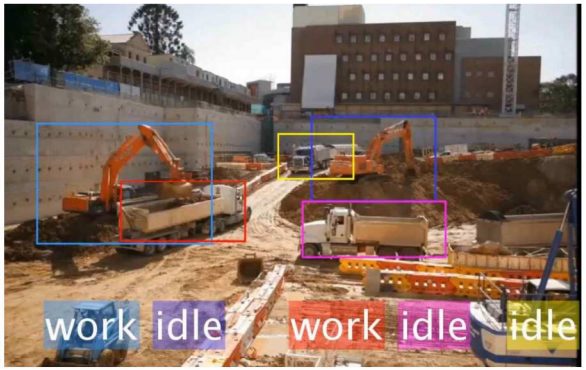






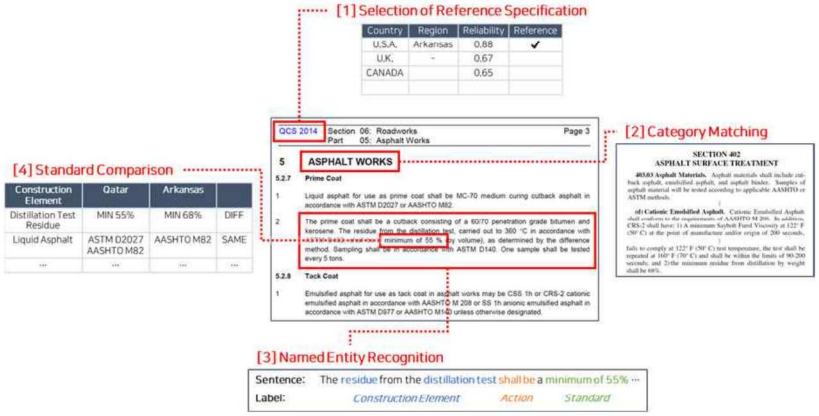
Deep Learning Revolution

- **Applications in Construction (C!Lab)**



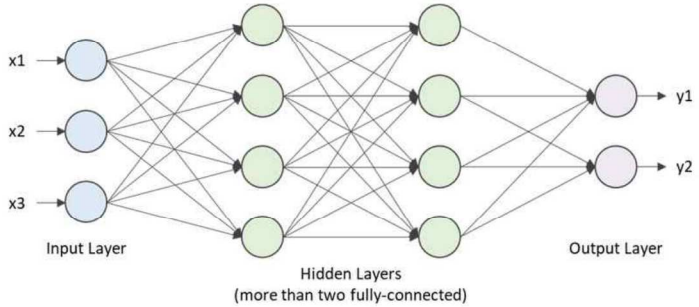
Automated Construction Site Monitoring

Infrastructure Damage Prediction

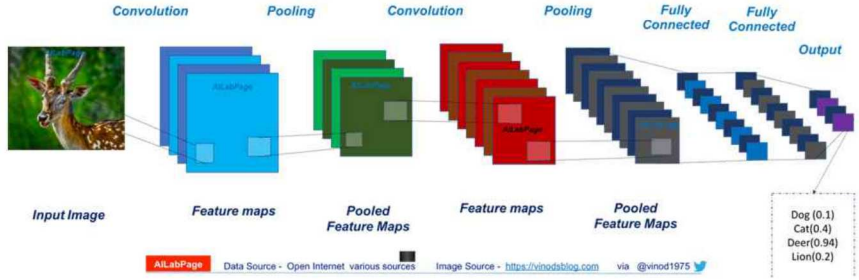


Automated Construction Document Analysis

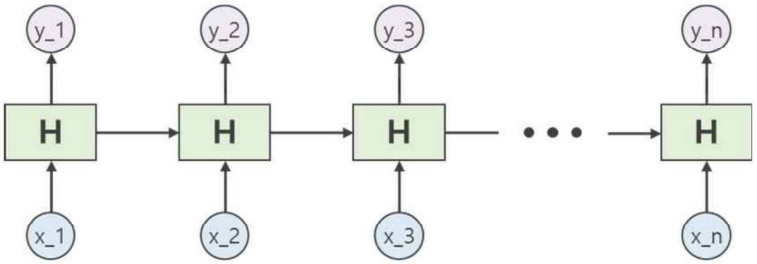
Major Architectures in Deep Learning



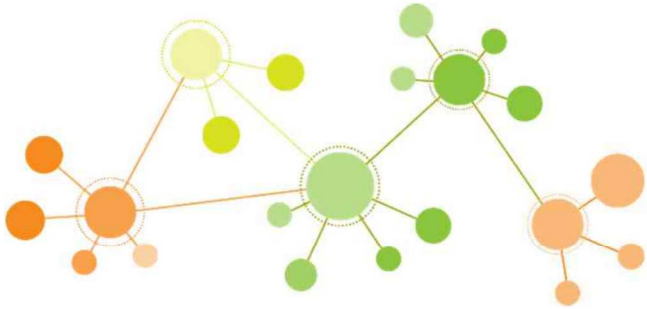
Deep Neural Network (DNN)



Convolutional Neural Network (CNN)



Recurrent Neural Network (RNN)

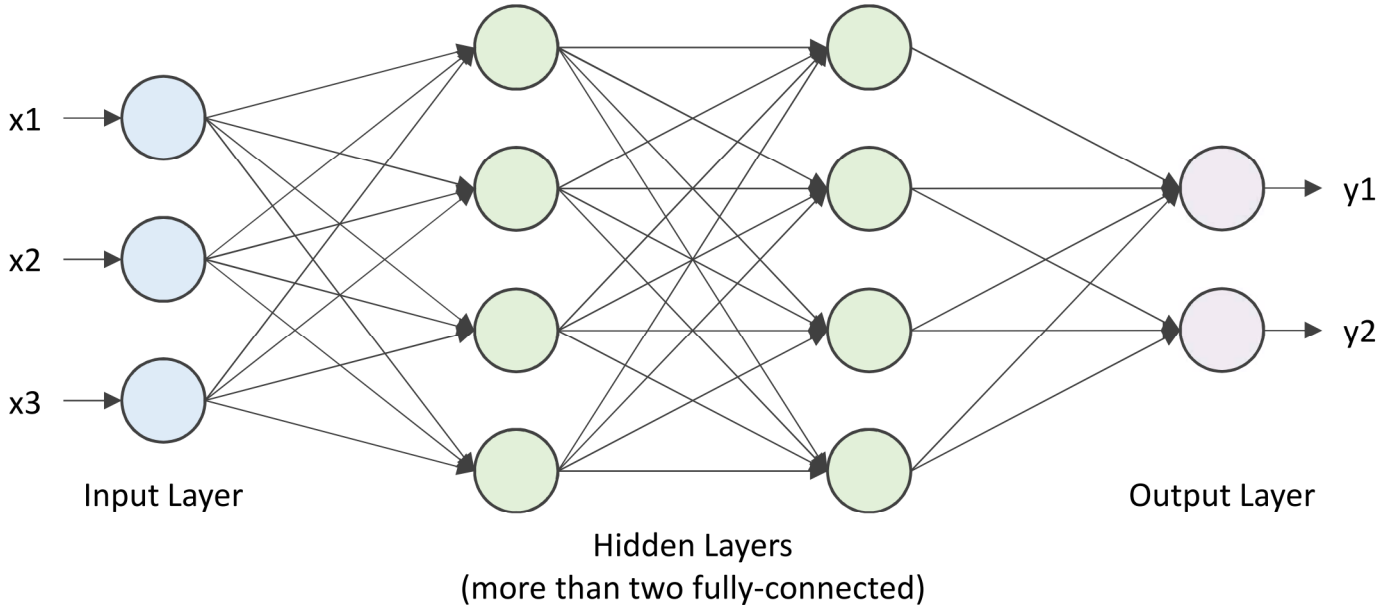


Graph Neural Network (GNN)

Major Architectures in Deep Learning

1. Deep Neural Network

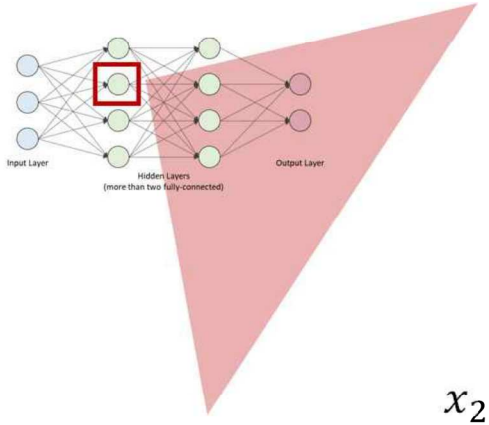
- A type of artificial neural networks that contain more than two hidden fully-connected layers

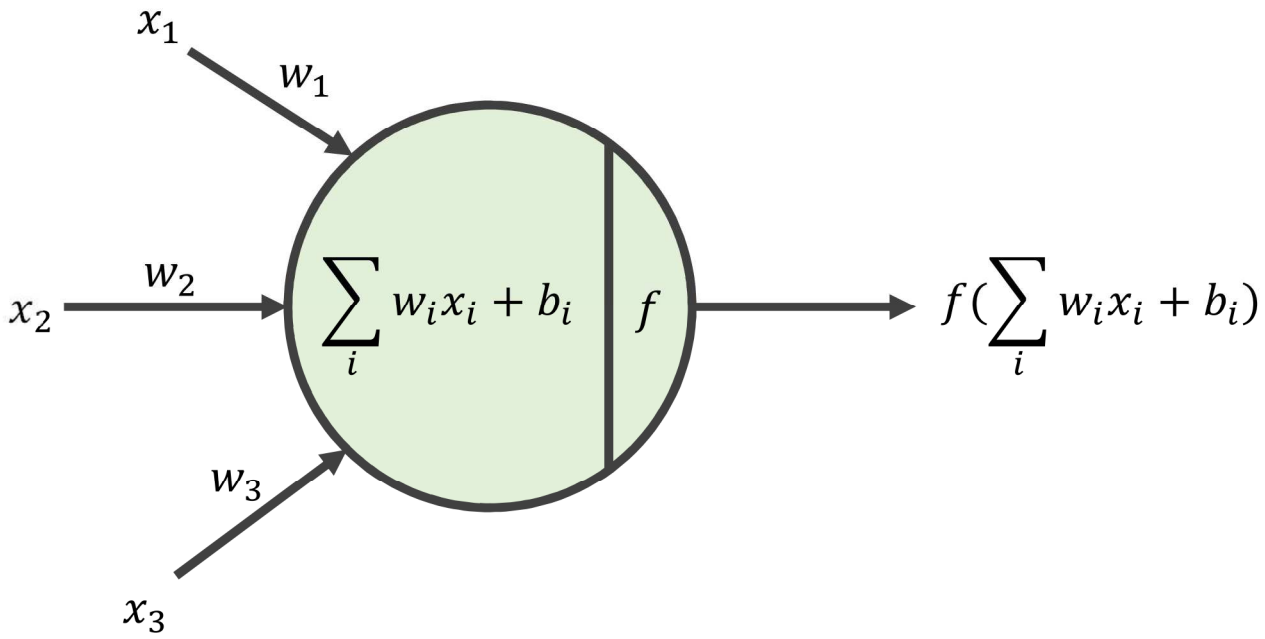


Major Architectures in Deep Learning

1. Deep Neural Network

- Model structure: input layer, (more than two) hidden layers, and output layer → A set of neurons
- Computational process: (1) weighting, (2) sum, and (3) activation





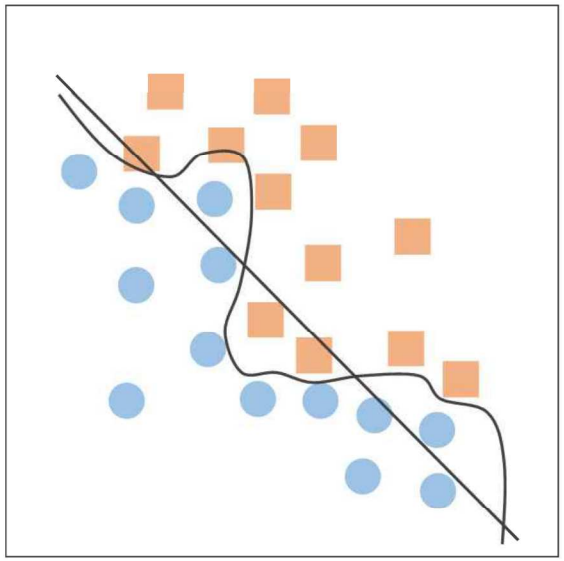
(1) weighting (2) sum (3) activation

Major Architectures in Deep Learning

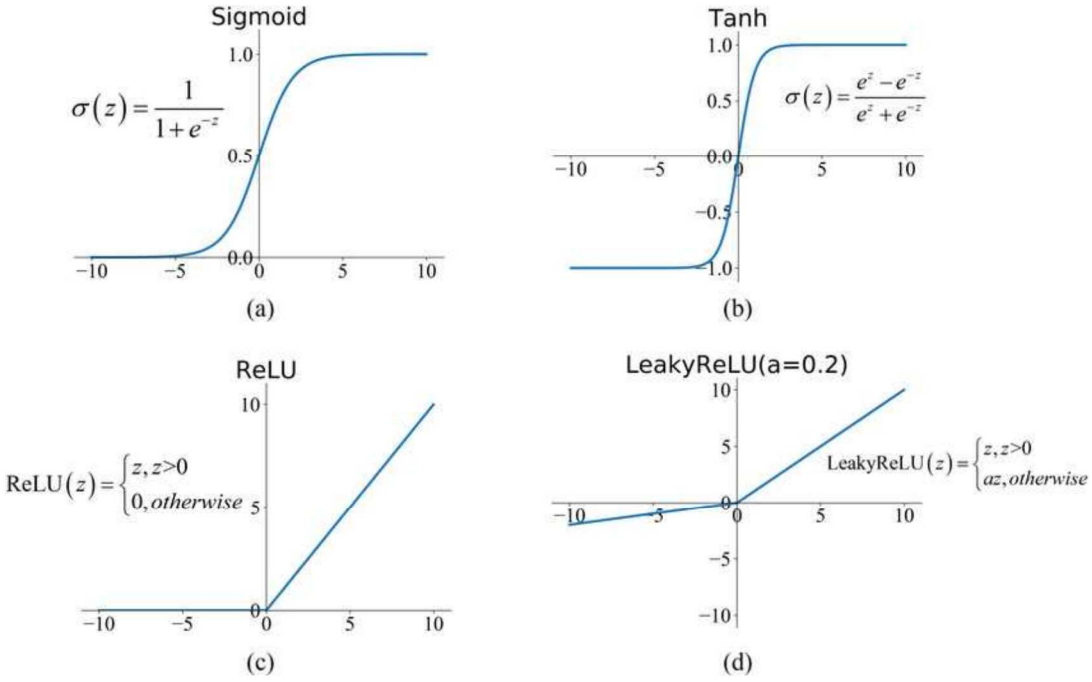
1. Deep Neural Network

- Computational process: (1) weighting, (2) sum, and **(3) activation**

  - An activation process is one of the most significant contributors to the performance of deep learning models. Specifically, *non-linear* activation functions allow deep neural networks to learn complex data and decision boundaries more effectively.



Linear vs. Non-linear
(Decision Boundaries)



(Source: Feng et al. 2019)

Widely Used Non-linear Activation Functions

Major Architectures in Deep Learning

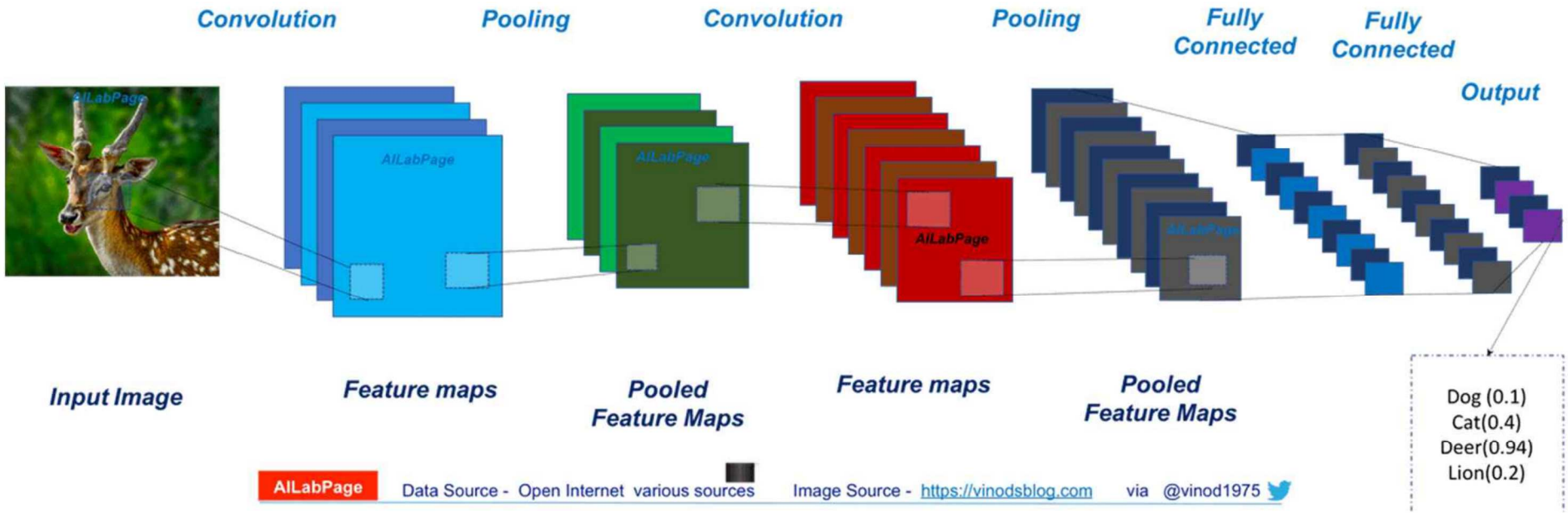
1. Deep Neural Network

- Advantages
  - Great ability to learn and represent complex relationships between input data (e.g., image) and output inferences (e.g., class), compared to traditional machine learning models (e.g., k-NN, support vector machine, etc.)
- Disadvantages
  - Black-box algorithms
  - Too dense and complex (← fully-connected layers)
  - A large number of parameters to be trained → possibility of overfitting
  - Need for a large amount of training data

Major Architectures in Deep Learning

2. Convolutional Neural Network

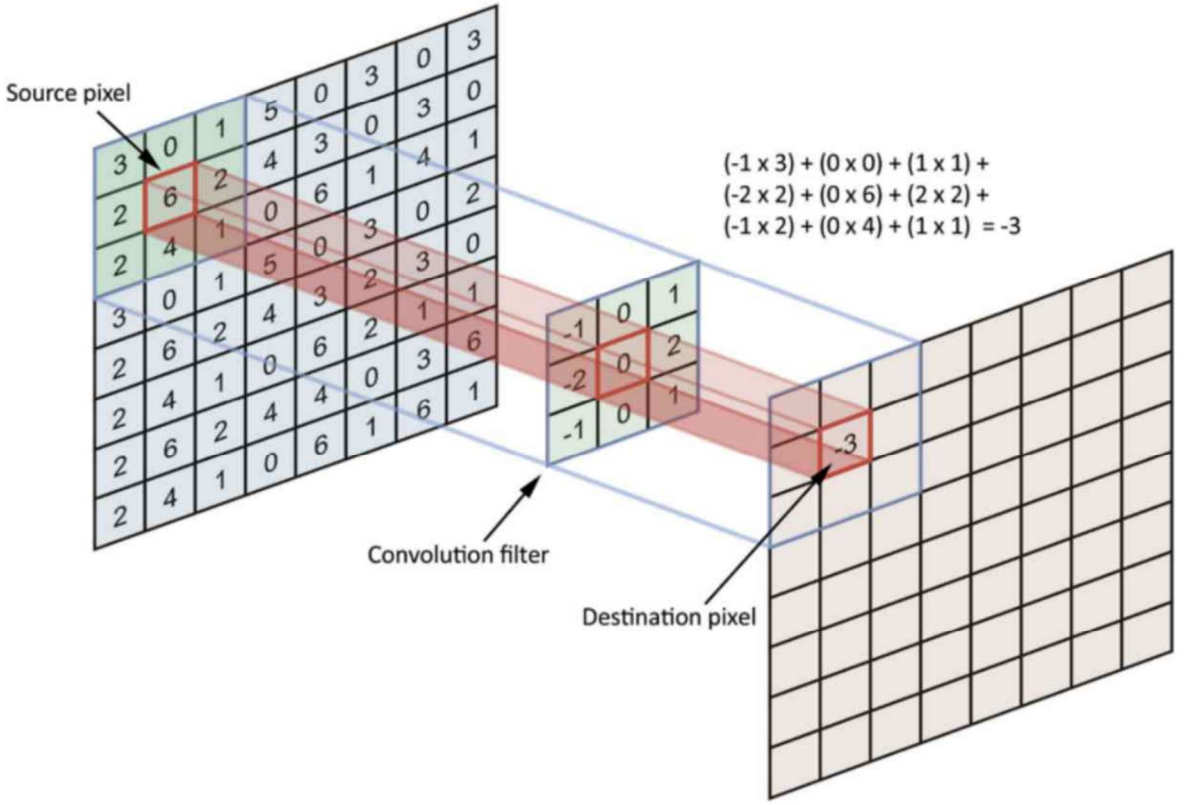
- A type of deep neural networks that split input data into small units, share weighting parameters (i.e., convolution), and extract various features of each unit for classification (e.g., image recognition)



Major Architectures in Deep Learning

2. Convolutional Neural Network

- Model structure: **convolution layer**, pooling layer, and fully-connected layer



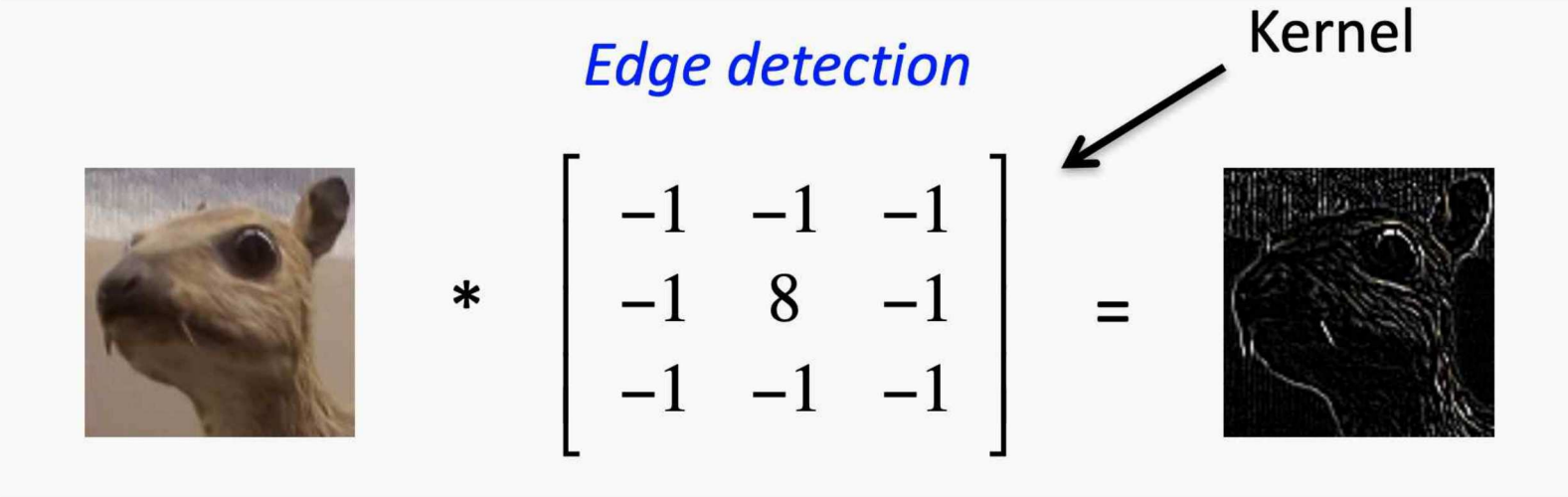
(Source: https://towardsdatascience.com/simple-introduction-to-convolutional-neural-networks-cdf8d3077bac)

Major Architectures in Deep Learning

2. Convolutional Neural Network

- Model structure: **convolution layer**, pooling layer, and fully-connected layer

  - Example of image convolution



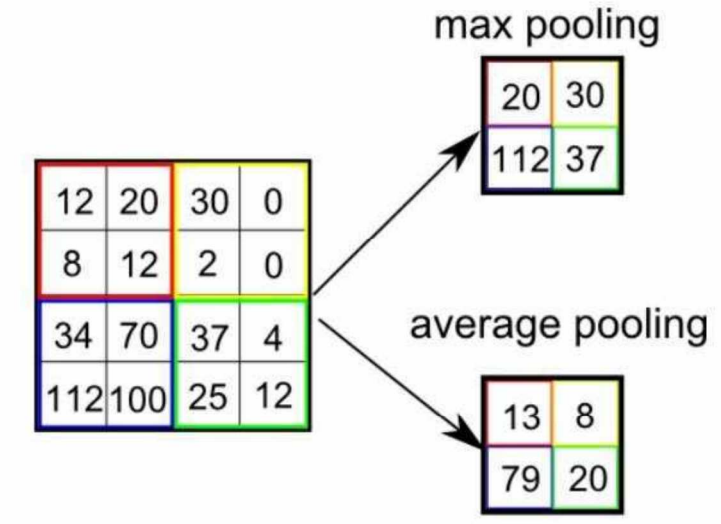
(Source: https://towardsdatascience.com/simple-introduction-to-convolutional-neural-networks-cdf8d3077bac)

Extract important visual features

Major Architectures in Deep Learning

2. Convolutional Neural Network

- Model structure: convolution layer, **pooling layer**, and fully-connected layer

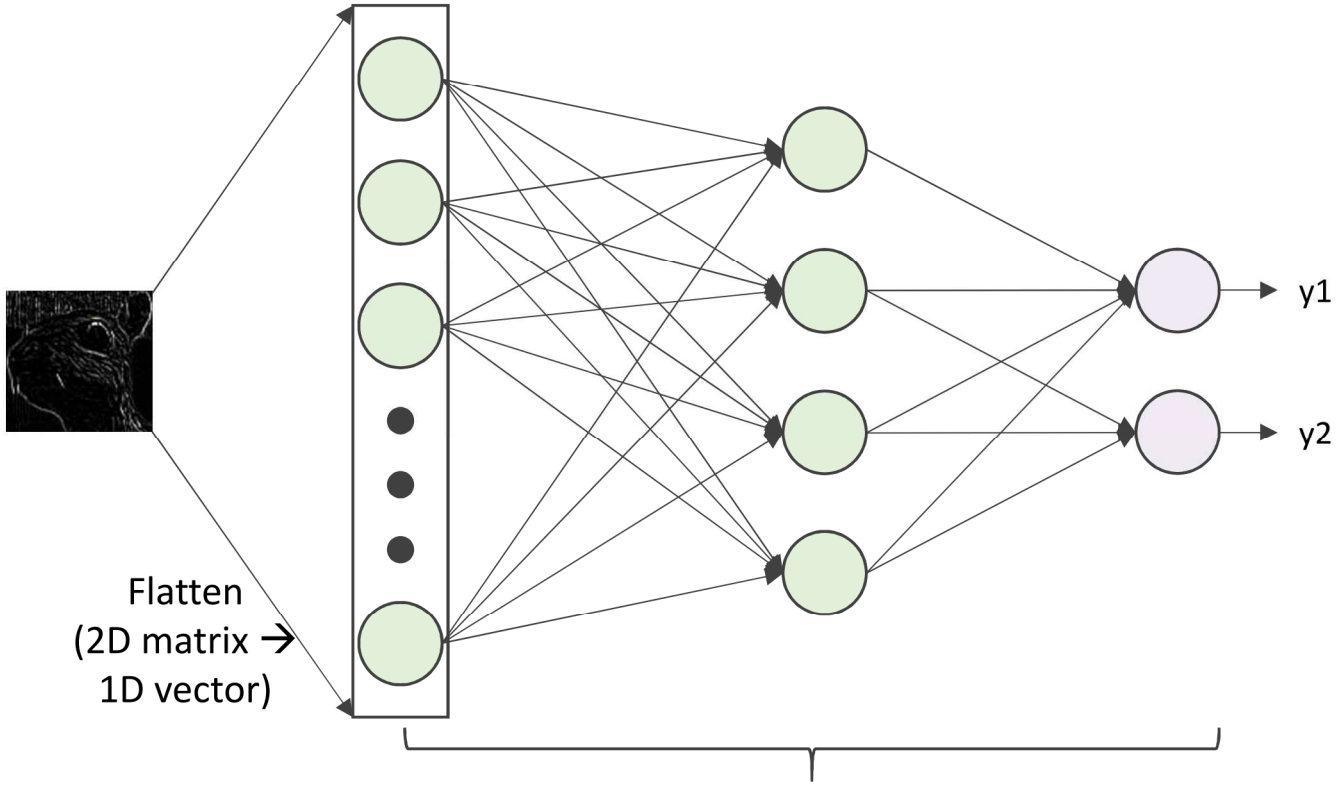


Reduce the spatial dimension while maintaining information details

Major Architectures in Deep Learning

2. Convolutional Neural Network

- Model structure: convolution layer, pooling layer, and **fully-connected layer**



Major Architectures in Deep Learning

2. Convolutional Neural Network

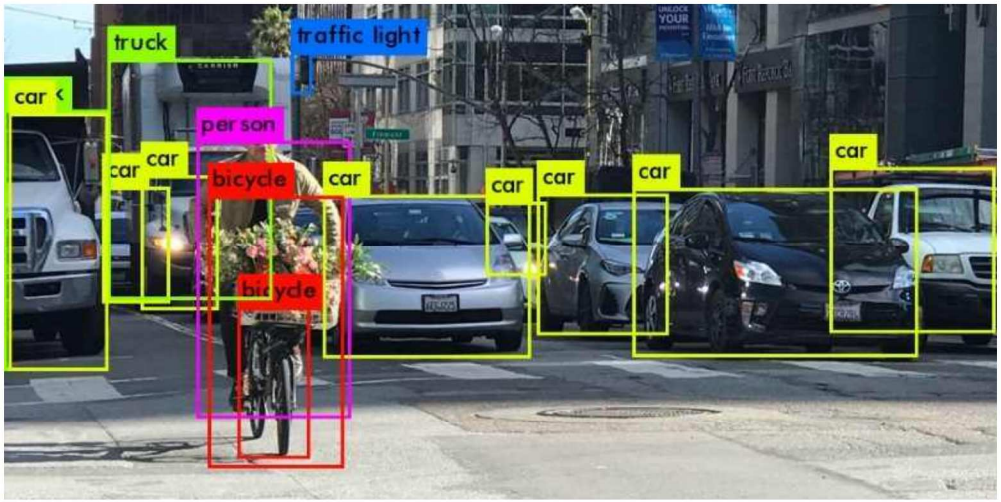
- Applications

  - Widely used for classification tasks

Image Classification

dog

Object Detection



(Source: https://medium.com/@jonathan_hui/real-time-object-detection-with-yolo-yolov2-28b1b93e2088)

Major Architectures in Deep Learning

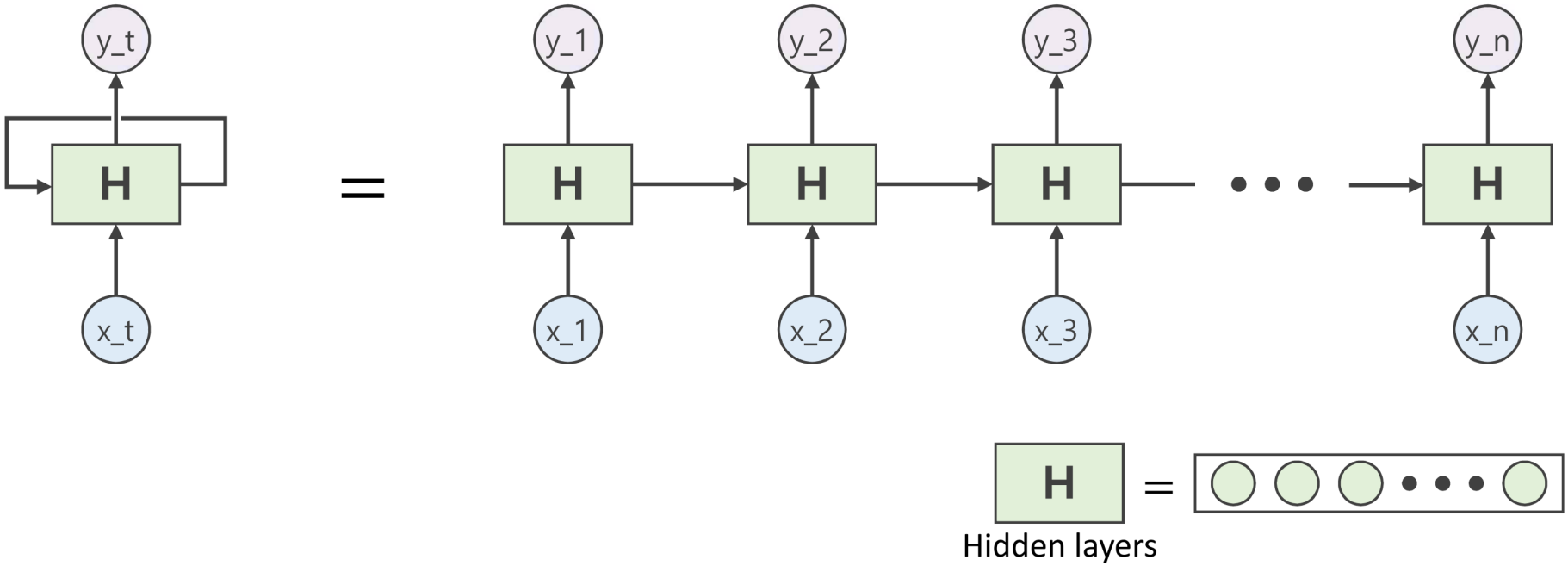
2. Convolutional Neural Network

- Advantages
  - Excellent ability to learn spatial features and local connectivity (through the convolution among neighbor pixels)
  - Comparatively sparse structure compared to deep neural networks
  - Reduce the number of parameters → less possibility of overfitting
- Disadvantages
  - Unable to learn temporal and time-series data (e.g., natural language processing, signal processing, video analysis)

Major Architectures in Deep Learning

3. Recurrent Neural Network

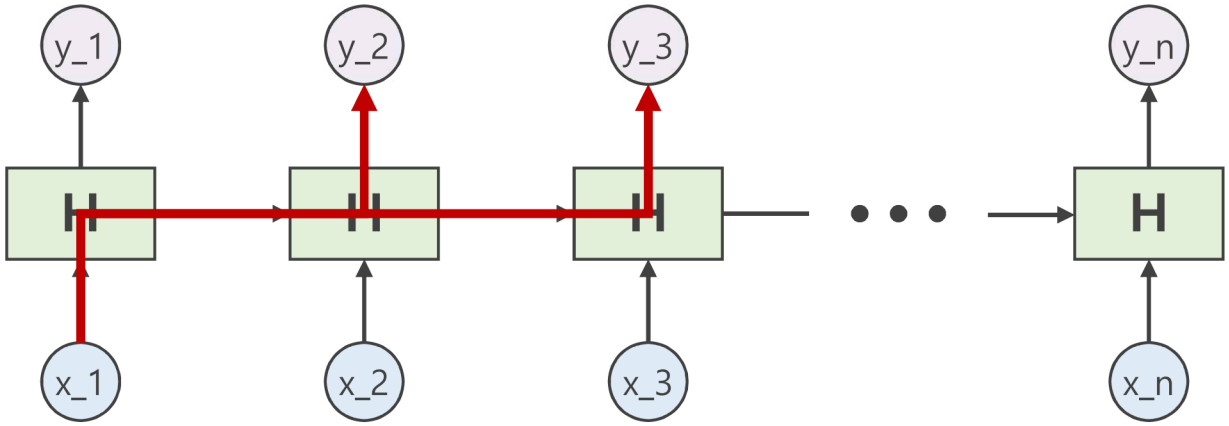
- A type of deep neural networks that consider temporal patterns of input sequence data



Major Architectures in Deep Learning

3. Recurrent Neural Network: Simple RNN

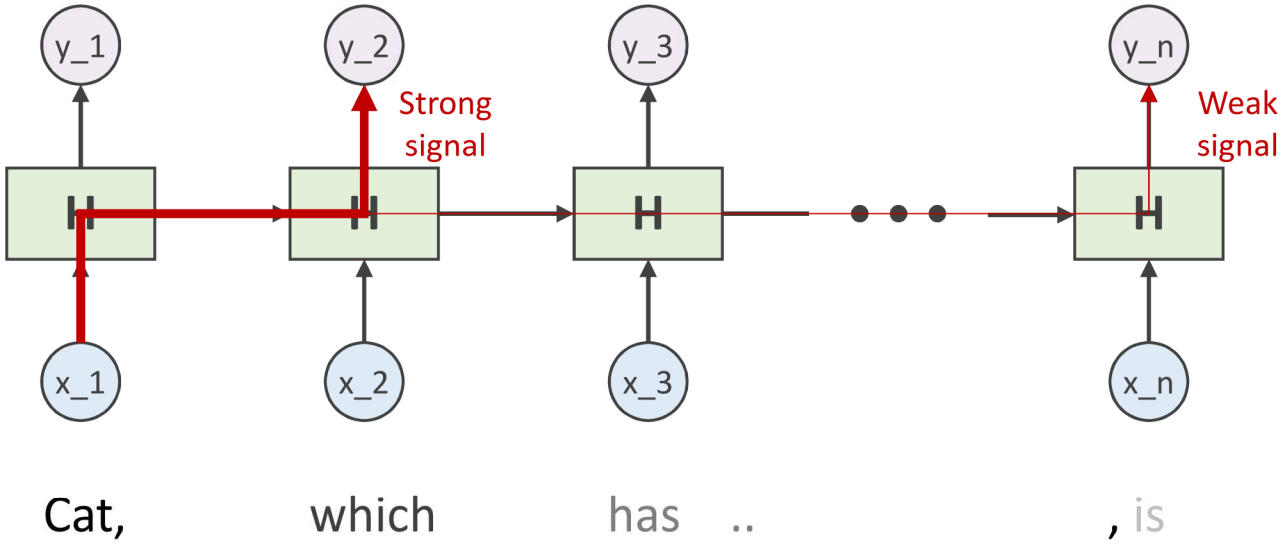
- Model structure: a sequence of neural networks
- Computational process: same with deep neural networks except that
  - (1) it receives the previous hidden state as input
  - (2) it transfers the current hidden state to the next hidden state



Major Architectures in Deep Learning

3. Recurrent Neural Network: Simple RNN

- Advantages
  - Can consider temporal dynamics of input data
  - Robust to the length of input data
- Disadvantages
  - Difficulty in accessing information from a long time ago, called *Gradient Vanishing problems*

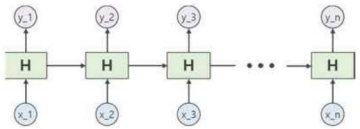


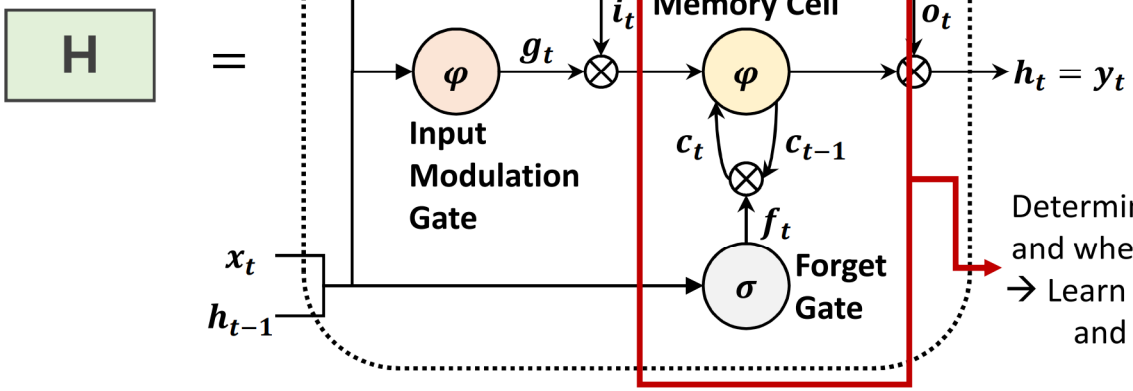
Major Architectures in Deep Learning

3. Recurrent Neural Network: Long Short Term Memory

- Model structure and computational process: Same with simple RNN except that

  - The forget gate and memory cell are additionally included in hidden layers





Determine when to forget previous hidden states and when to memorize and update hidden states
→ Learn critical sequential information selectively and long-range dependencies effectively

Major Architectures in Deep Learning

3. Recurrent Neural Network: Long Short Term Memory

- Advantages
  - Can learn important sequential information selectively
  - Can address gradient vanishing problems
- Disadvantages
  - Unable to consider backward sequential patterns

Cat, which has.. , is

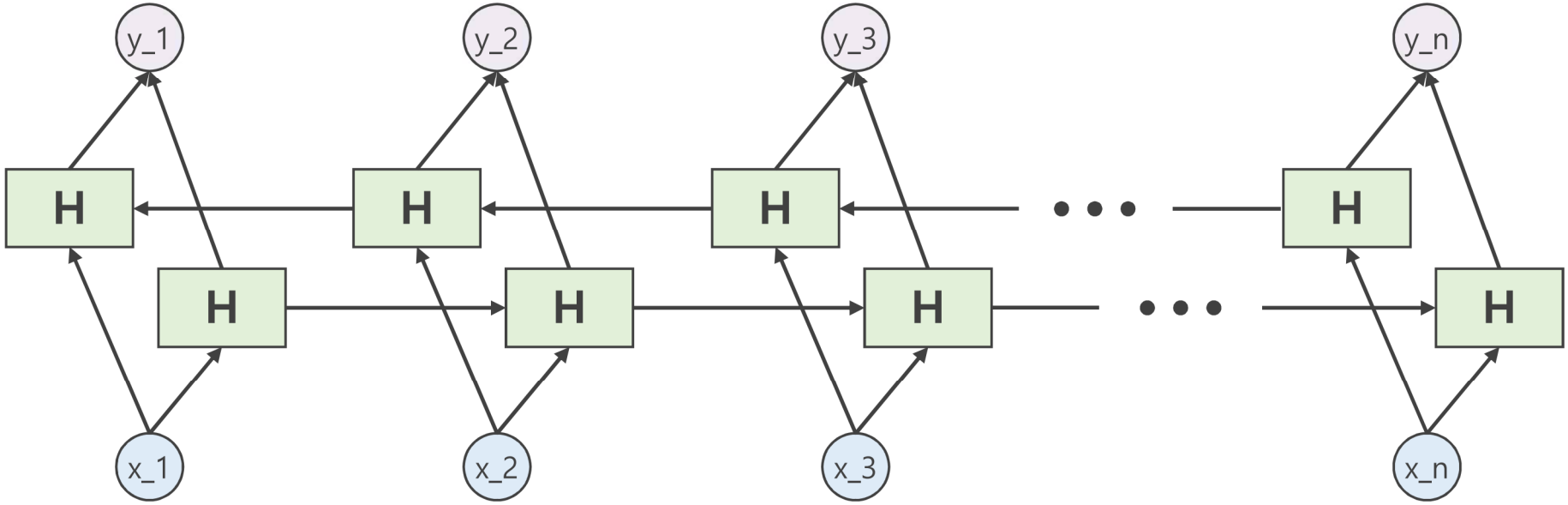
Apple
↓
사과

Apple iPhone
↓
사과 아이폰

Major Architectures in Deep Learning

3. Recurrent Neural Network: Bidirectional RNN and LSTM

- Model structure and computational process: Similar to simple RNN and LSTM except that
  - (1) it receives both previous and next hidden states as input
  - (2) it transfers the current hidden state to both previous and next hidden state



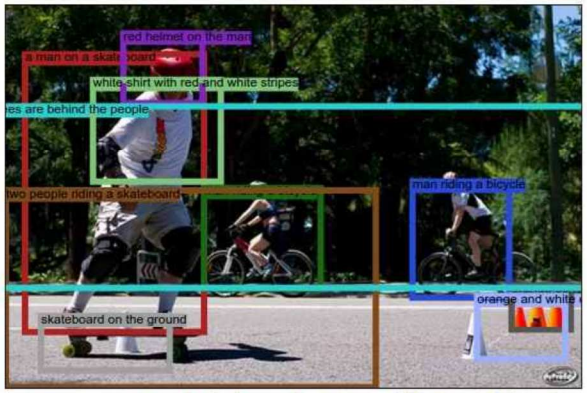
Major Architectures in Deep Learning

3. Recurrent Neural Network

- Applications

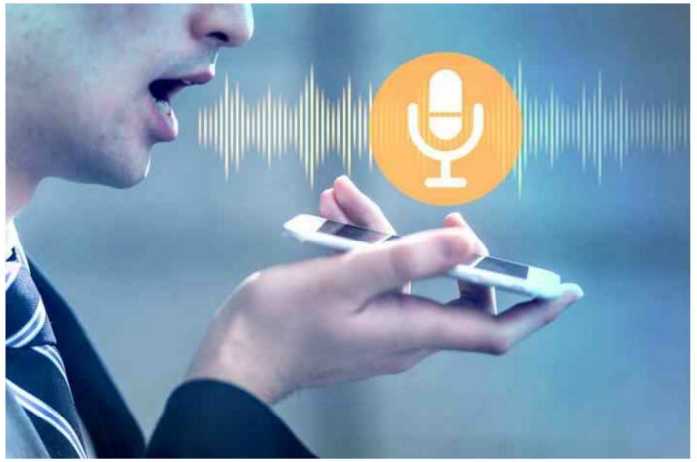
Machine Translation

Image Captioning



(Source: https://cs.stanford.edu/people/karpathy/densecap/)

Speech Recognition



(Source: https://becominghuman.ai/voice-recognition-beyond-smart-speakers-6b6c61c7b9e8)

Major Architectures in Deep Learning

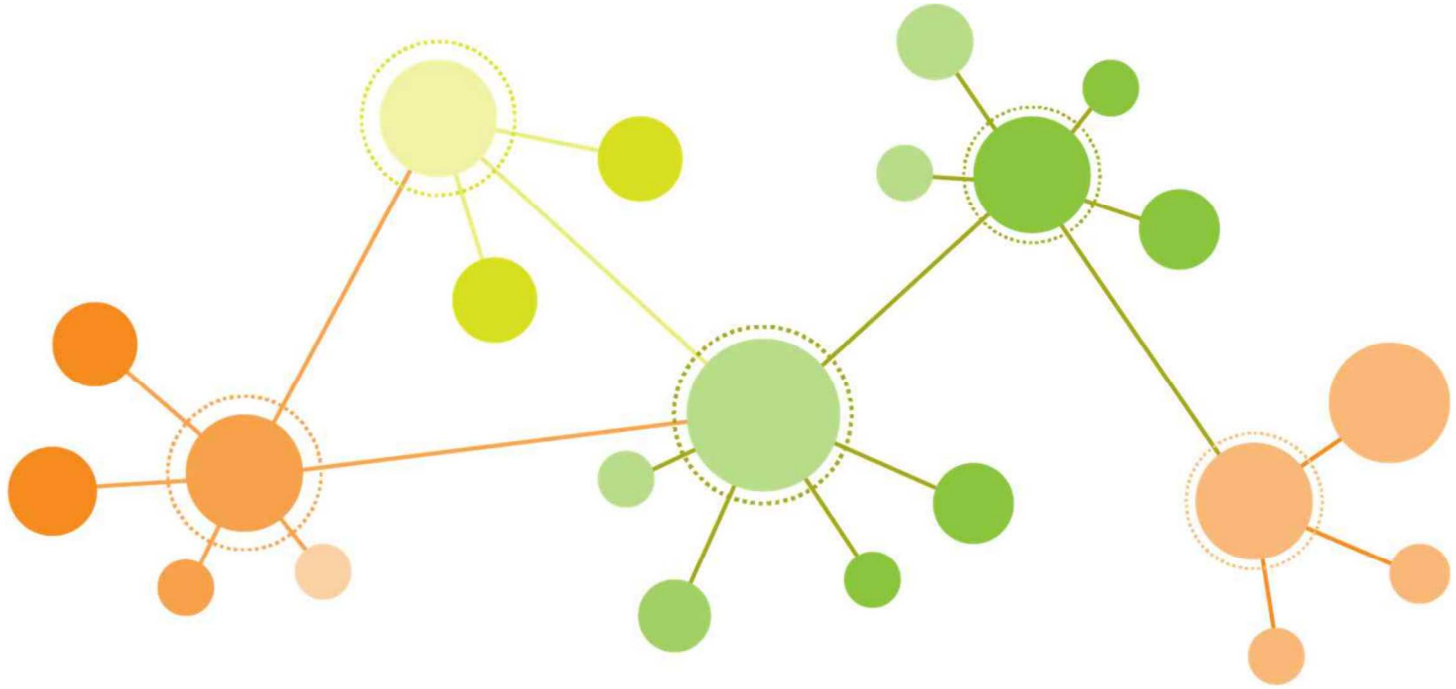
3. Recurrent Neural Network: Bidirectional RNN and LSTM

- Advantages
  - Can consider both forward and backward sequential patterns
- Disadvantages
  - More complex models than simple RNN and LSTM → Bidirectional models are not always best
  - Unable to learn complex multi-dimensional networks

Major Architectures in Deep Learning

4. Graph Neural Network (as an emerging architecture)

- A type of deep neural networks that learn a set of objects (i.e., nodes) and their relationships (i.e., edges)

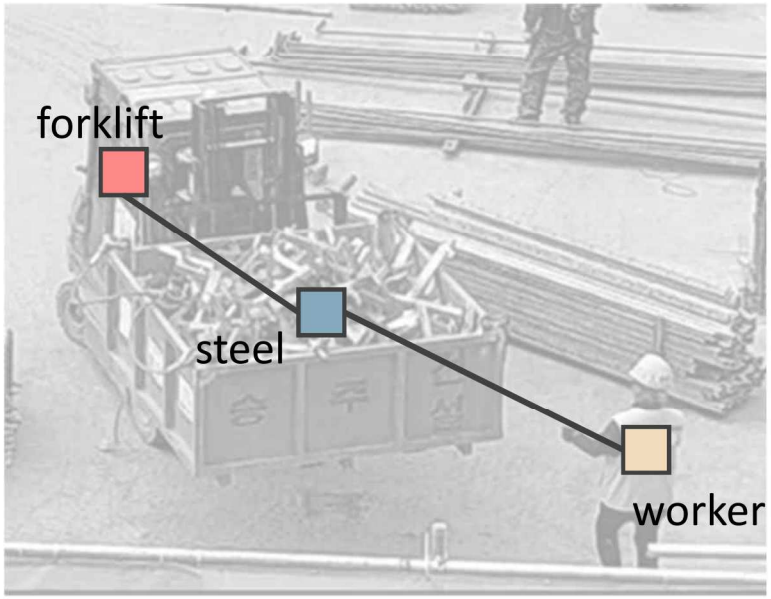


Major Architectures in Deep Learning

4. Graph Neural Network

- Model structure: a set of nodes and their linked edges
- Computational process: there are diverse processing methods, but one of the most typical processes is

  - Neighbor aggregation (i.e., message passing): node-to-edge embedding, edge-to-node embedding, and concatenation



Problem: What types are their relationships? (EDGE)
Given: Object features (NODE)

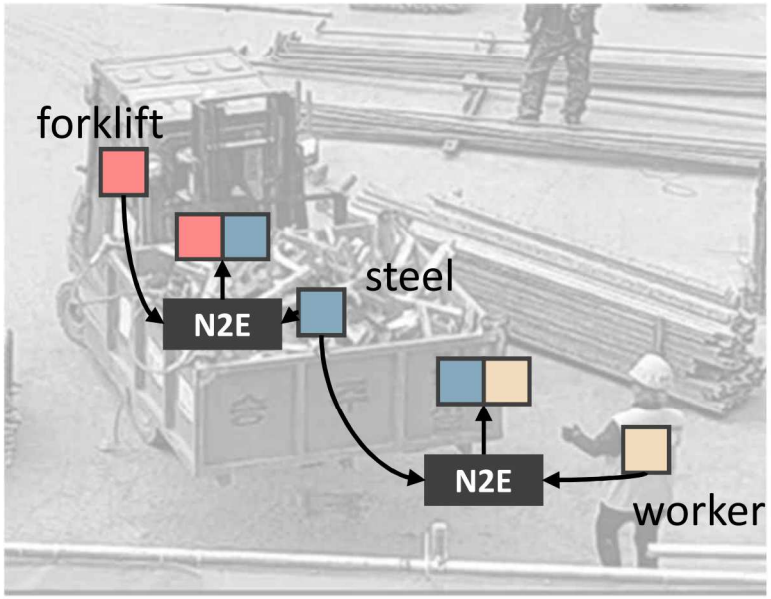
Object features (e.g., CNN visual features, spatial features, ..)

Major Architectures in Deep Learning

4. Graph Neural Network

- Model structure: a set of nodes and their linked edges
- Computational process: there are diverse processing methods, but one of the most typical processes is

  - Neighbor aggregation (i.e., message passing): **node-to-edge embedding**, edge-to-node embedding, and concatenation



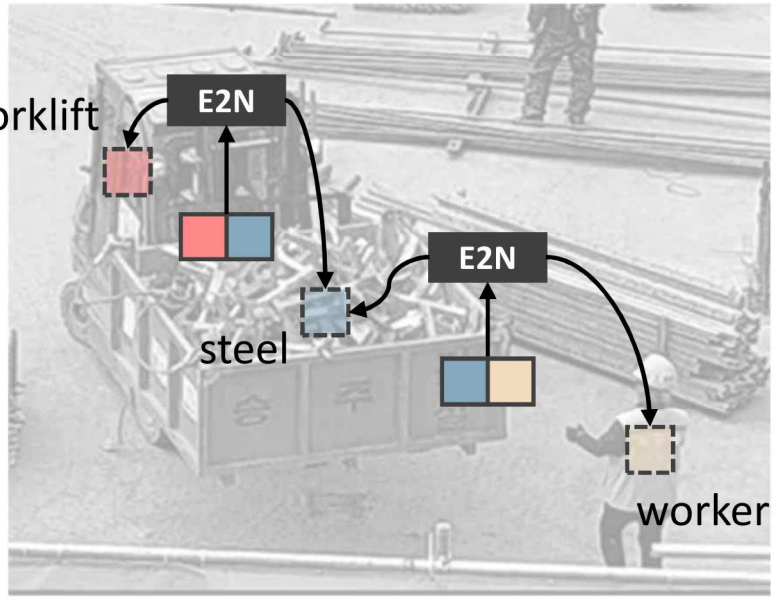
N2E

Embedding can be performed by various types of neural networks
(e.g., fully-connected, convolution) or
other unsupervised techniques (e.g., concatenation)

Major Architectures in Deep Learning

4. Graph Neural Network

- Model structure: a set of nodes and their linked edges
- Computational process: there are diverse processing methods, but one of the most typical processes is
  - Neighbor aggregation (i.e., message passing): node-to-edge embedding, **edge-to-node embedding**, and concatenation



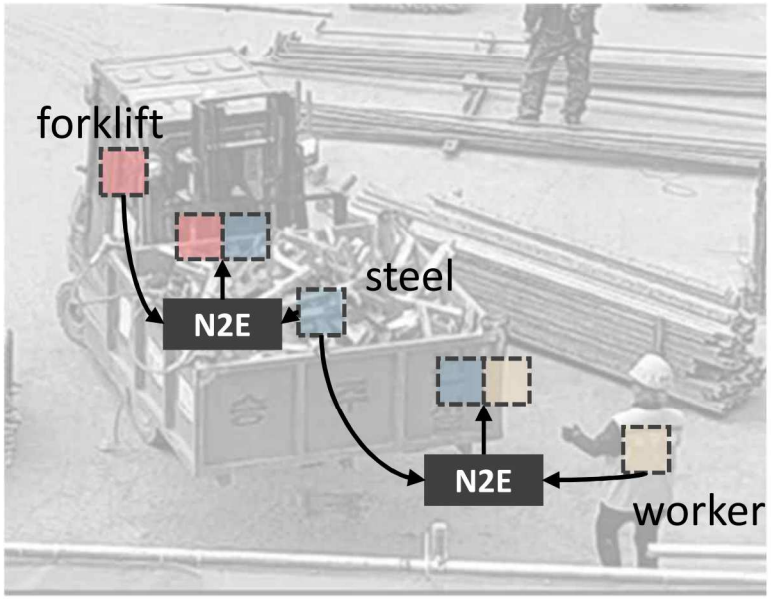
E2N

Embedding can be performed by various types of neural networks
(e.g., fully-connected, convolution) or
other unsupervised techniques (e.g., concatenation)

Major Architectures in Deep Learning

4. Graph Neural Network

- Model structure: a set of nodes and their linked edges
- Computational process: there are diverse processing methods, but one of the most typical processes is
  - Neighbor aggregation (i.e., message passing): **node-to-edge embedding**, edge-to-node embedding, and concatenation



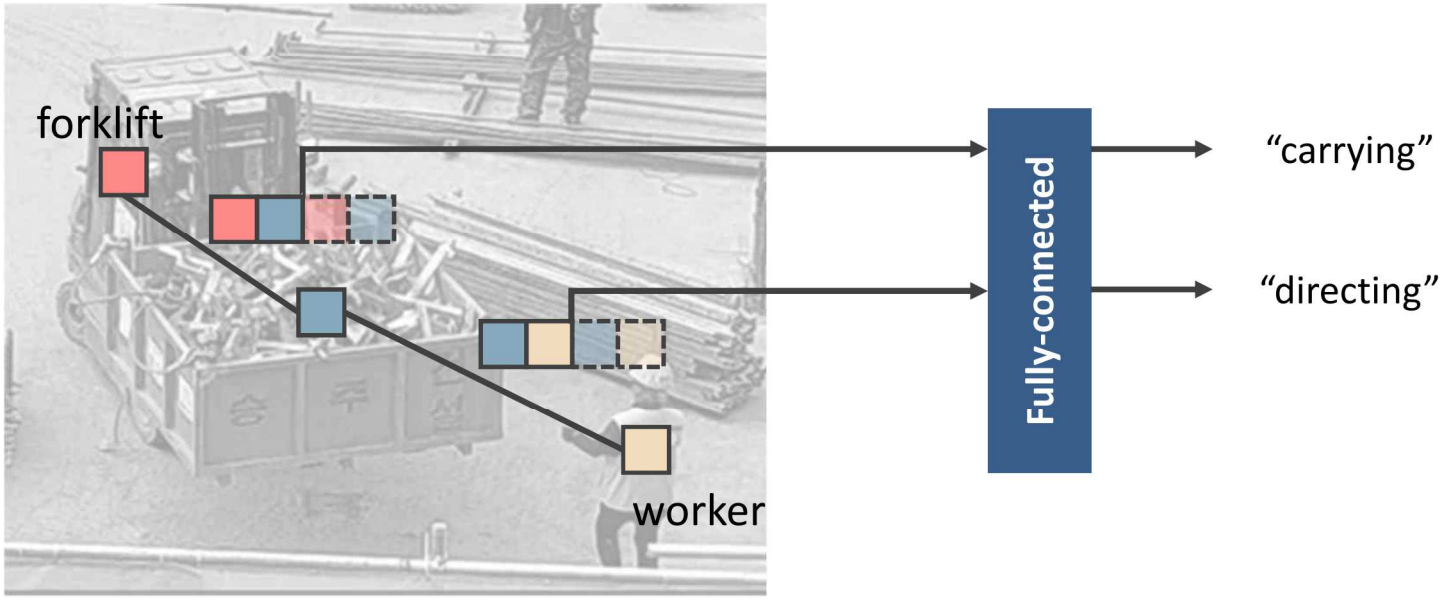
N2E

Embedding can be performed by various types of neural networks
(e.g., fully-connected, convolution) or
other unsupervised techniques (e.g., concatenation)

Major Architectures in Deep Learning

4. Graph Neural Network

- Model structure: a set of nodes and their linked edges
- Computational process: there are diverse processing methods, but one of the most typical processes is
  - Neighbor aggregation (i.e., message passing): node-to-edge embedding, edge-to-node embedding, and **concatenation**

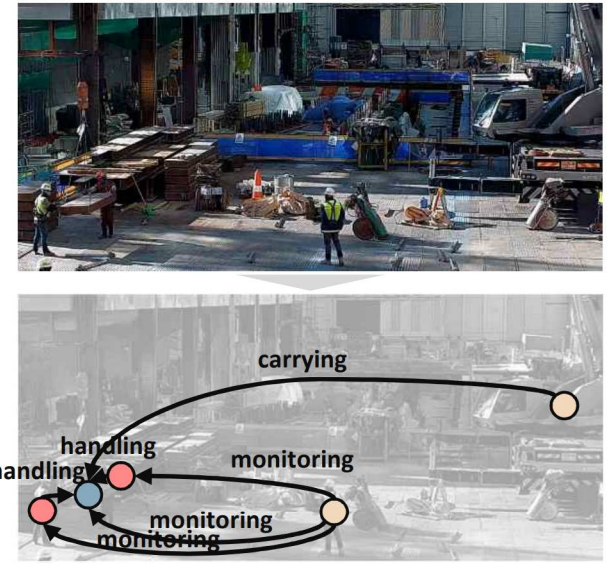


Major Architectures in Deep Learning

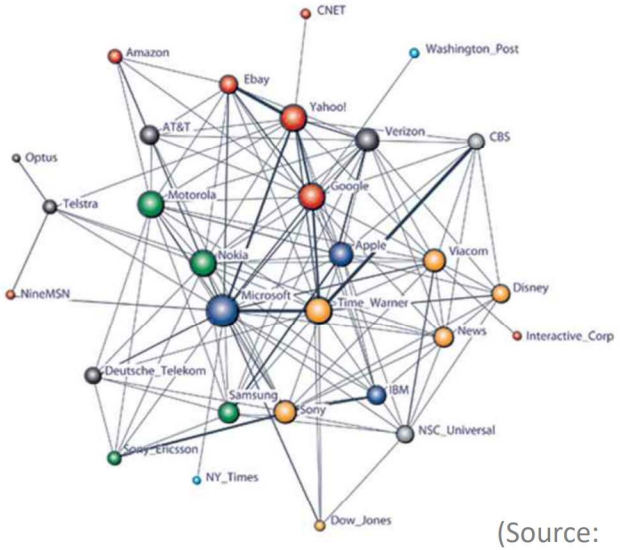
4. Graph Neural Network

- Applications

Visual Relationship Detection



Social Network Analysis



(Source: https://rossdawson.com/blog/analyzing_media/)

Major Architectures in Deep Learning

4. Graph Neural Network

- Advantages
  - Ability to learn and interpret complex multi-dimensional relationships
  - Can connect data points that exist far from each other but have relationships (i.e., Non-Euclidean space)
  - Can minimize the model complexity (i.e., the number of parameters to be trained)
- Disadvantages
  - Unable to learn temporal and time-series data (e.g., natural language processing, signal processing, video analysis)

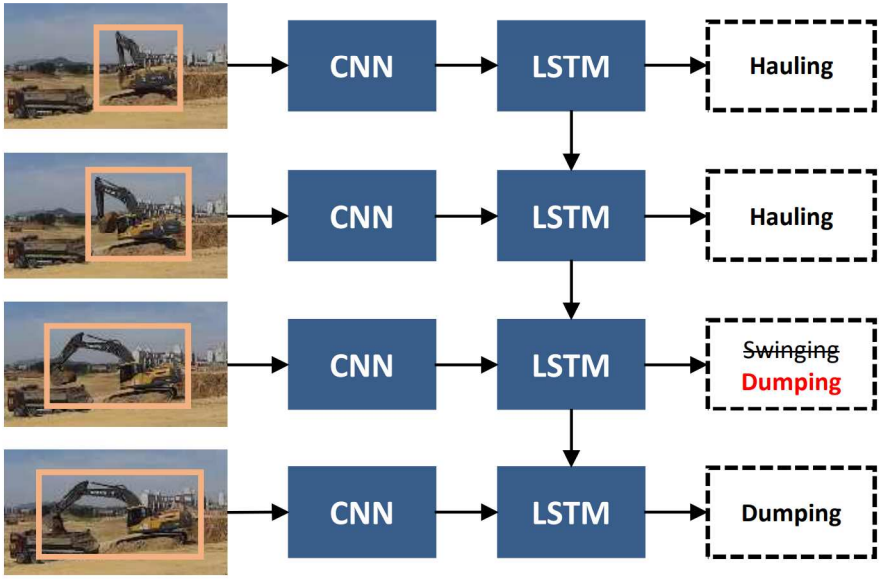
Major Architectures in Deep Learning

Hybrid Architecture

- Example 1: CNN-RNN, CNN-LSTM

  - CNN: extract important visual features from each image frame

  - RNN/LSTM: analyze temporal patterns of visual features extracted



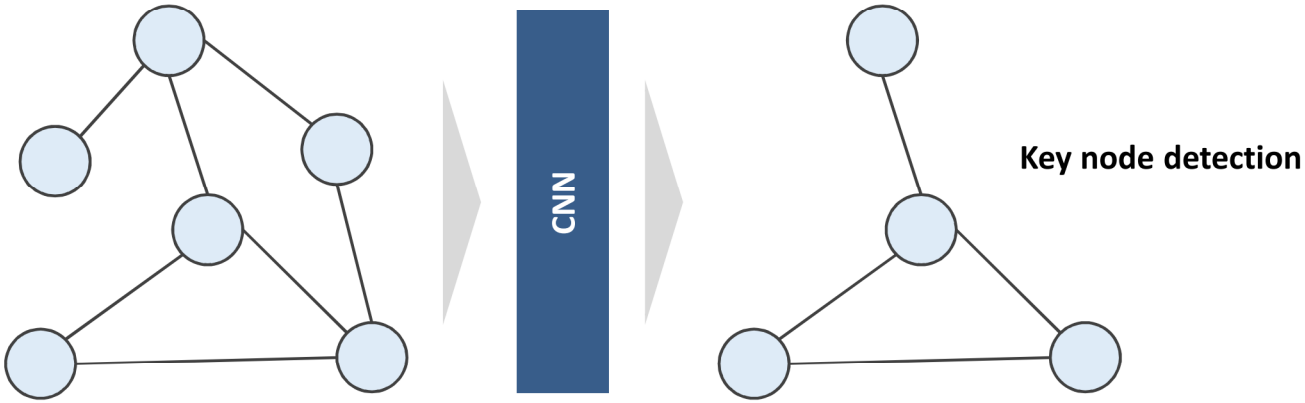
Major Architectures in Deep Learning

Hybrid Architecture

- Example 2: Graph Convolutional Neural Networks

  - GNN: process a graph-structured data

  - CNN: extract important nodes, edges, and sub-graphs from a graph-structured input data



Challenges in Deep Learning: (Annotated) Training Data

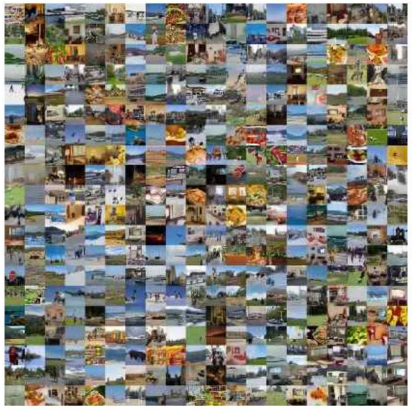
Traditional Approaches: The More Data, The Better Performance

- Challenges

  - A large amount of human efforts and time

  - A lack of training data in the real-world (e.g., medical data)

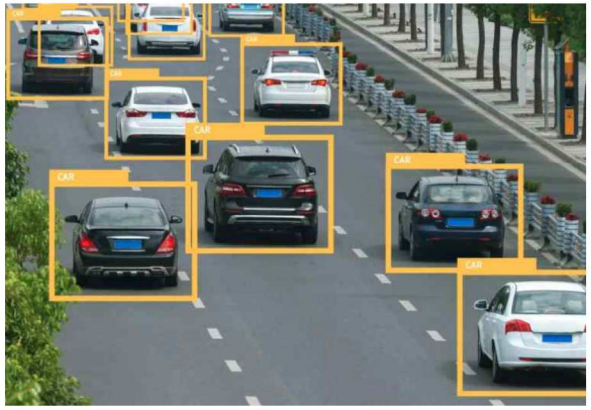
MS COCO Dataset



More than 330,000 images

X

Image Labeling



About 30 seconds per image

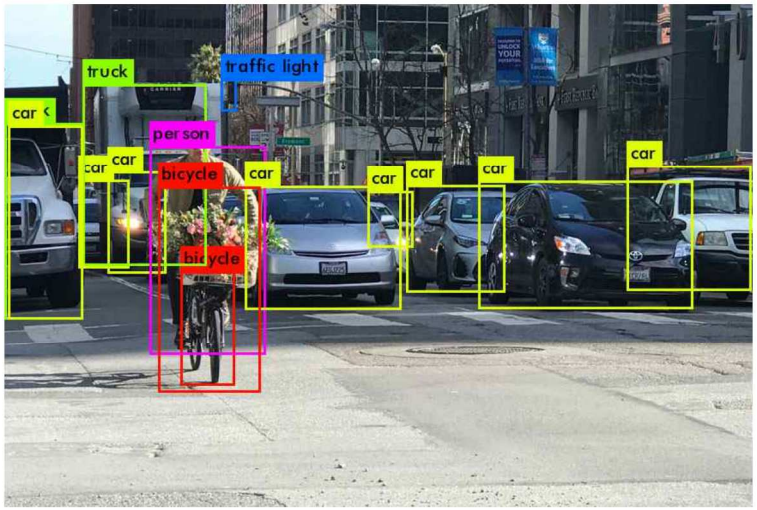
= More than 115 days

Challenges in Deep Learning: (Annotated) Training Data

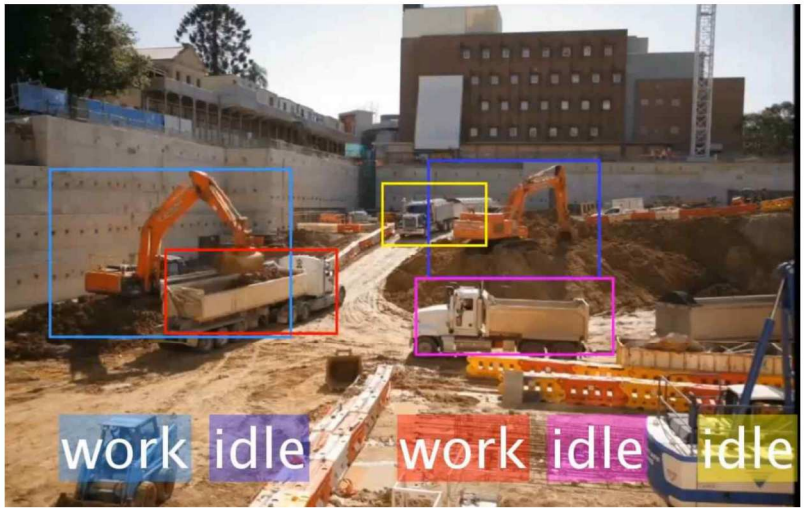
Novel Approaches: The Less Data, But Comparable Performance

- 1. Transfer learning

  - A machine learning method where a model developed for a task is reused as the starting point for a model on a second task.



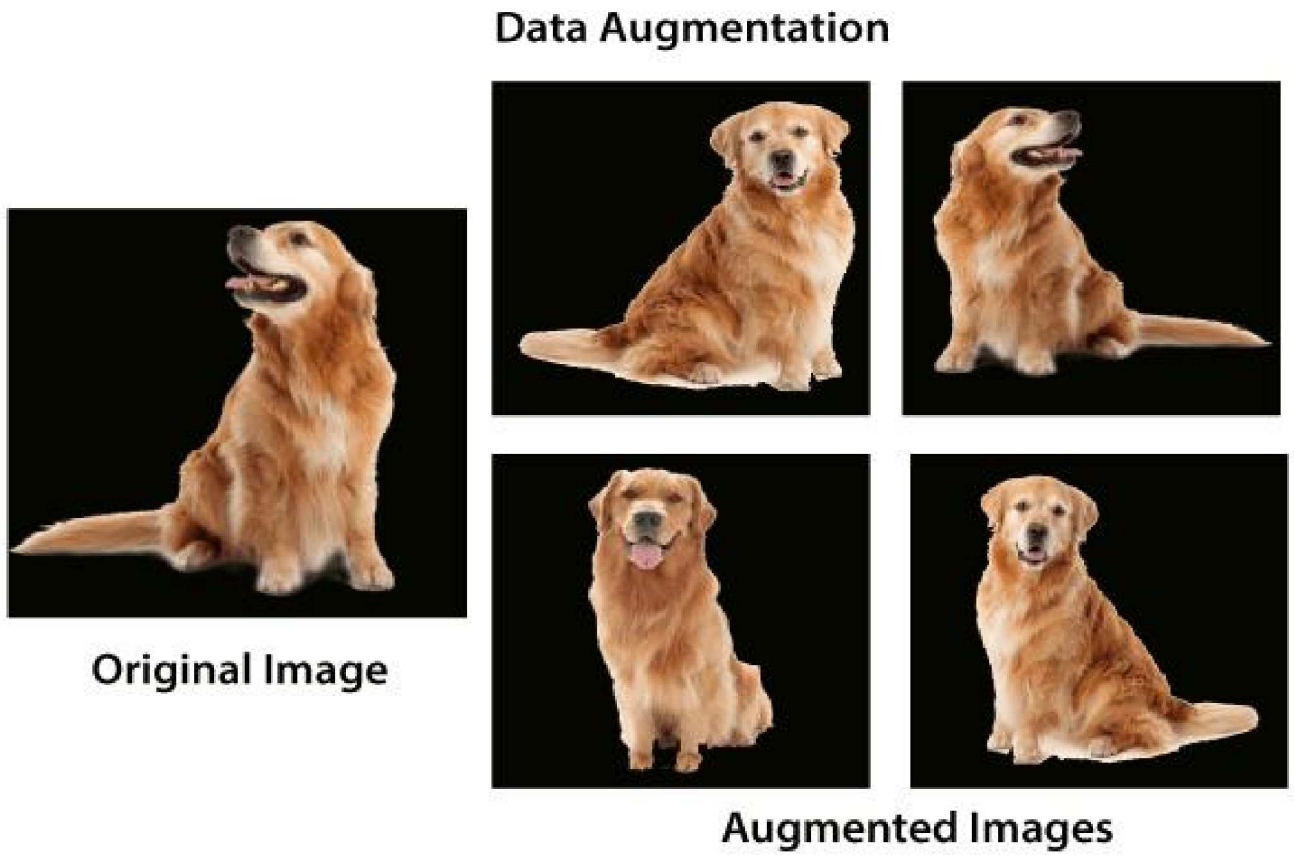
330,000 images



10,000 images

Challenges in Deep Learning: (Annotated) Training Data

- **Novel Approaches: The Less Data, But Comparable Performance**
  - 2. Data augmentation



Challenges in Deep Learning: (Annotated) Training Data

Novel Approaches: The Less Data, But Comparable Performance

- 5. Active learning

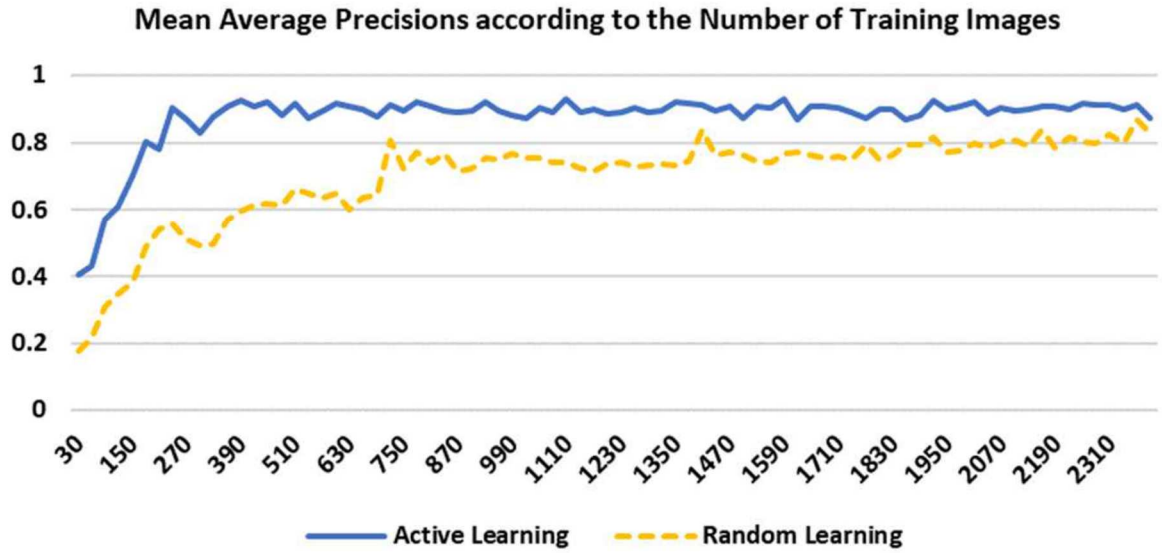
  - Active learning selects the most meaningful-to-learn instances from abundant training data and trains a deep learning model with the selected data first.

Low Uncertainty

No information to learn

High Uncertainty

New information to learn



Challenges in Deep Learning: (Annotated) Training Data

Novel Approaches: The Less Data, But Comparable Performance

- 6. Meta learning (i.e., “learning to learn”)

  - Human can learn new concepts and skills only with few examples.

    (e.g., if human knows how to ride a bicycle, he or she can easily learn how to ride a motorcycle.)