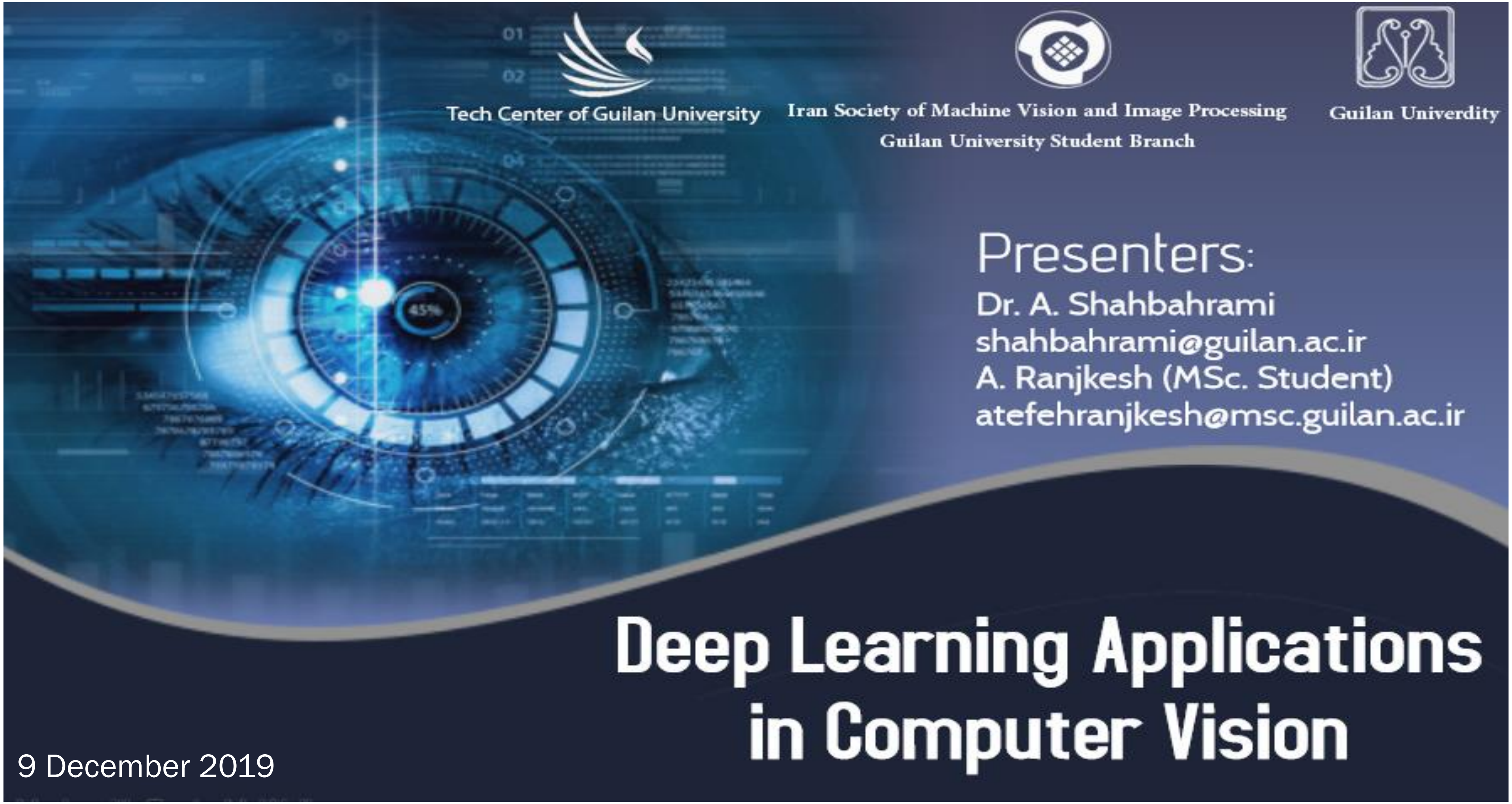
انجمن ماشین بینایی
و پردازش تصویر ایران
شاخه دانشجویی دانشگاه گیلان
دانشگاه گیلان
گروه مهندسی کامپیوتر
کارگاه آموزشی
کاربردهای یادگیری عمیق
در بینایی ماشین
ارائه دهندگان:
- دکتر اسدالله شاه بهرامی
- مهندس عاطفه رنجکش
زمان و محل برگزاری:
- دوشنبه ۱۸ آذر، ساعت: ۱۵:۳۰-۱۳:۳۰
- سالن شهید نورانی دانشکده فنی دانشگاه گیلان
برگزار کنندگان:
انجمن ماشین بینایی و پردازش تصویر ایران- شاخه دانشجویی دانشگاه گیلان و مرکز رشد واحدهای فناور دانشگاه گیلان
ismvip.guilan.ac.ir


Tech Center of Guilan University
Iran Society of Machine Vision and Image Processing
Guilan University Student Branch
Guilan Univerdity
Presenters:
Dr. A. Shahbahrami
shahbahrami@guilan.ac.ir
A. Ranjkesh (MSc. Student)
atefehranjkesh@msc.guilan.ac.ir
Deep Learning Applications
in Computer Vision
9 December 2019

Scopes

- Computer Vision (CV)
- Machine Learning (ML)
- Neural Network (NN)
- Deep Learning (DL)
- Conclusion

Computer Vision

➢ **Why image?**

- Characters of good shoes:
  - ✓ Beauty
  - ✓ Suitable Heel
  - ✓ Replaceable Insole
  - ✓ Fit

Computer Vision

What is computer vision?

1982

“the construction of explicit, meaningful descriptions of physical objects from images”

1998

“computing properties of the 3D world from one or more digital images”

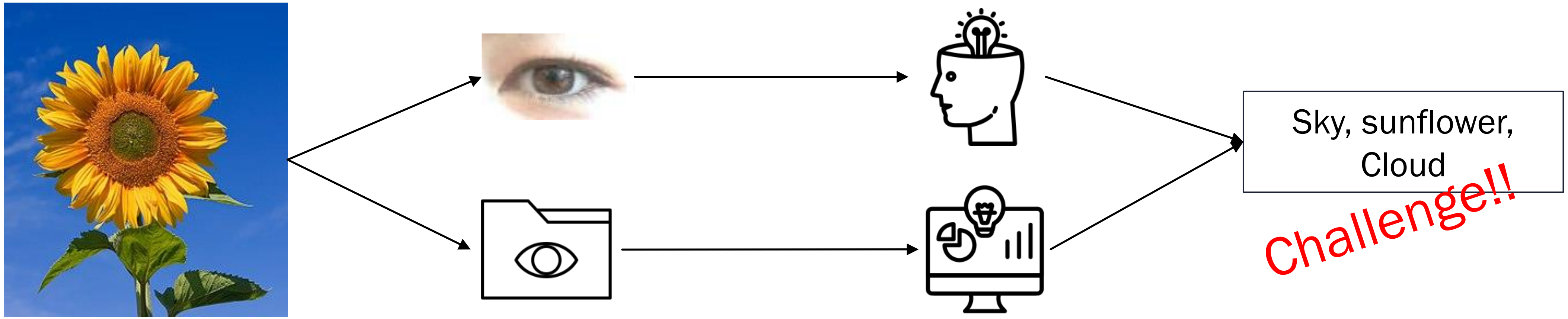
2001

“to make useful decisions about real physical objects and scenes based on sensed images”

Computer Vision

➢ What is computer vision?

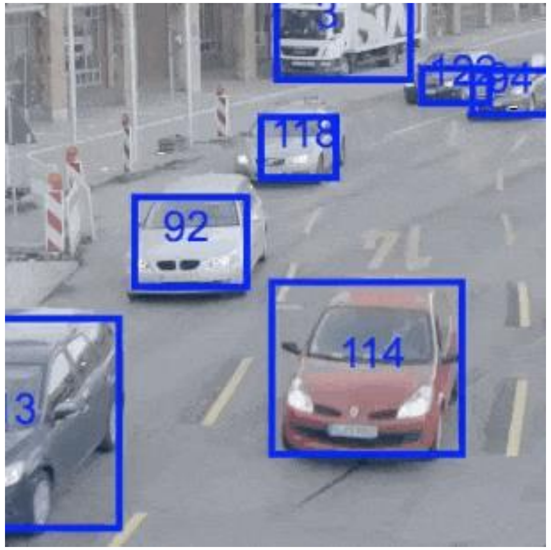
Image | Sensing device | Interpreting device | Interpretations



Computer Vision

➢ What is it related to?

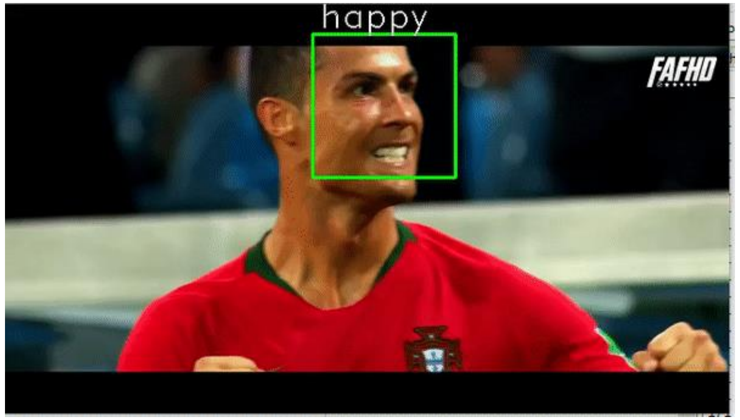
- Intelligent Transportation System (ITS)
- Face Recognition
- Object Detection
- Robotics
- Biometrics
- Image Classification
- Question & Answering
- Agriculture



Computer Vision

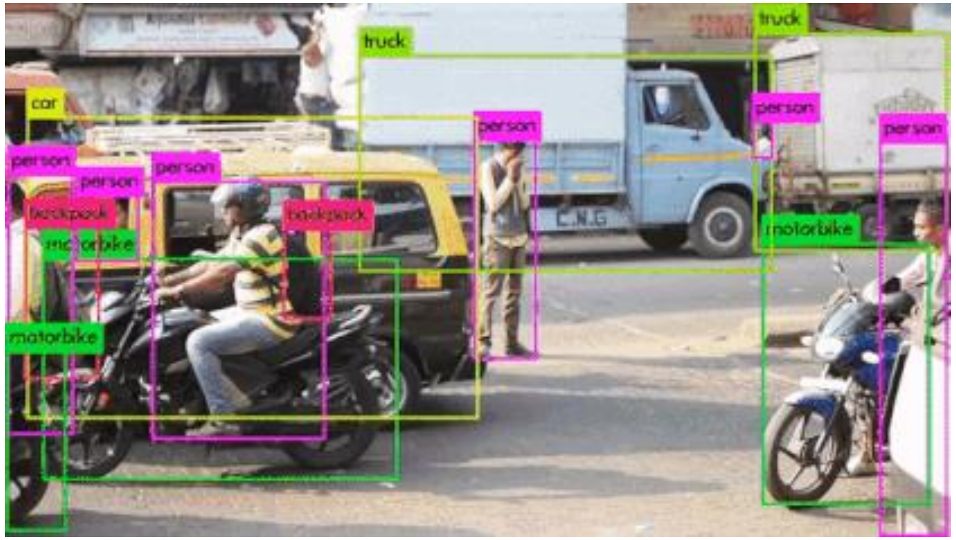
➢ What is it related to?

- Intelligent Transportation System (ITS)
- Face Recognition
- Object Detection
- Robotics
- Biometrics
- Image Classification
- Question & Answering
- Agriculture



Computer Vision

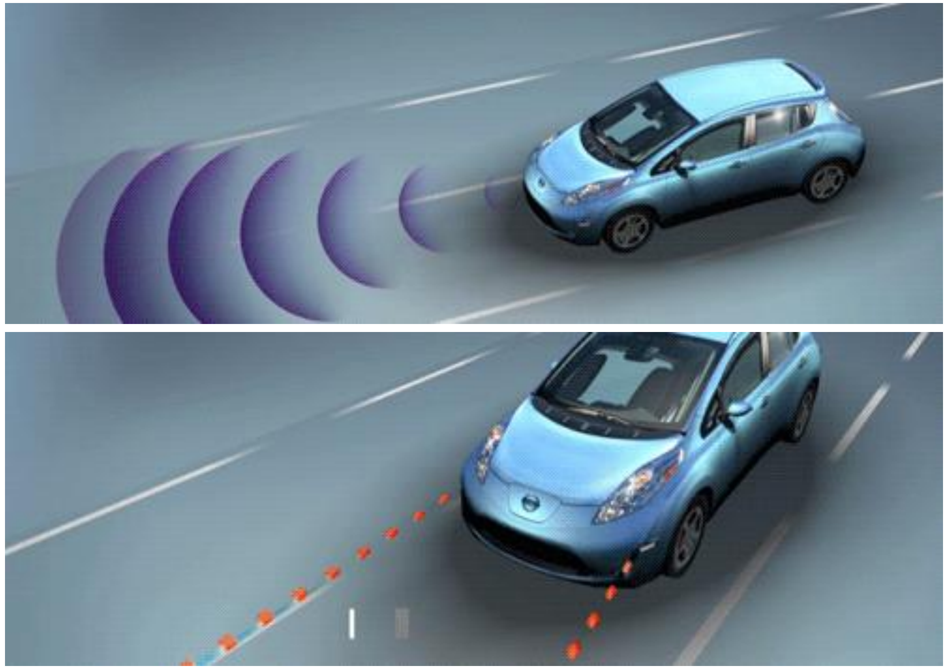
- What is it related to?
  - Intelligent Transportation System (ITS)
  - Face Recognition
  - Object Detection
  - Robotics
  - Biometrics
  - Image Classification
  - Question & Answering
  - Agriculture



Computer Vision

➢ What is it related to?

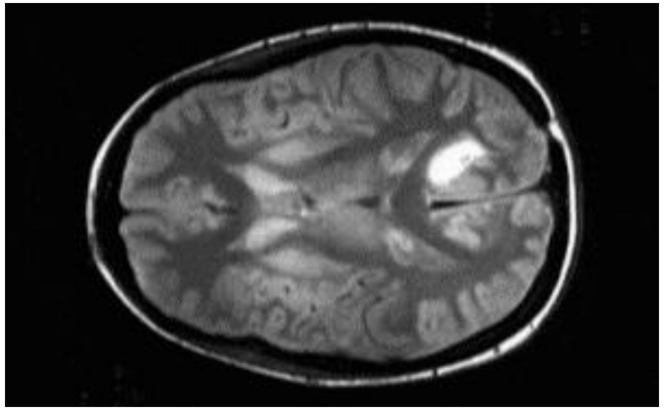
- Intelligent Transportation System (ITS)
- Face Recognition
- Object Detection
- Robotics
- Biometrics
- Image Classification
- Question & Answering
- Agriculture



Computer Vision

➢ What is it related to?

- Intelligent Transportation System (ITS)
- Face Recognition
- Object Detection
- Robotics
- Biometrics
- Image Classification
- Question & Answering
- Agriculture



Computer Vision

➢ What is it related to?

- Intelligent Transportation System (ITS)
- Face Recognition
- Object Detection
- Robotics
- Biometrics
- Image Classification
- Question & Answering
- Agriculture

Computer Vision

➢ What is it related to?

- Intelligent Transportation System (ITS)
- Face Recognition
- Object Detection
- Robotics
- Biometrics
- Image Classification
- Question & Answering
- Agriculture

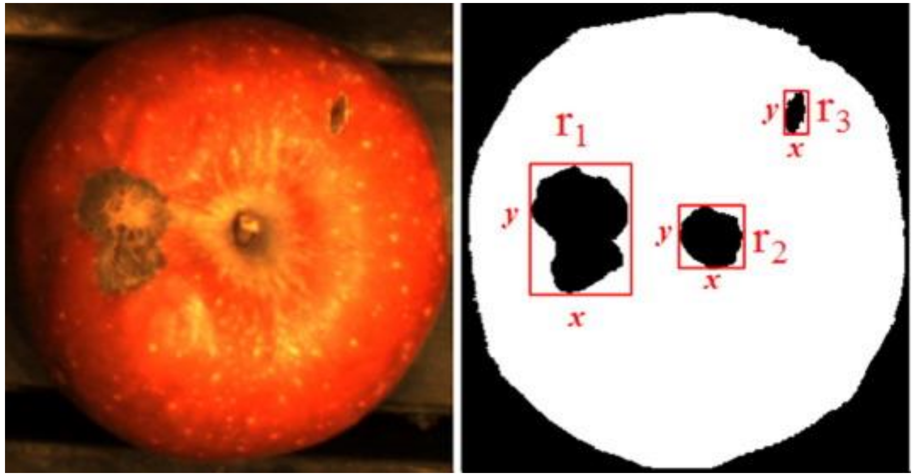
Where is the child sitting?!!

fridge OR chair

Computer Vision

➢ What is it related to?

- Intelligent Transportation System (ITS)
- Face Recognition
- Object Detection
- Robotics
- Biometrics
- Image Classification
- Question & Answering
- Agriculture



Computer Vision

➢ What is computer vision challenges?

- View point Variation
- Difference in Illumination
- Occlusion
- Scale
- Background Clutter

Computer Vision

➢ What is computer vision challenges?

- View point Variation
- Difference in Illumination
- Occlusion
- Scale
- Background Clutter

Computer Vision

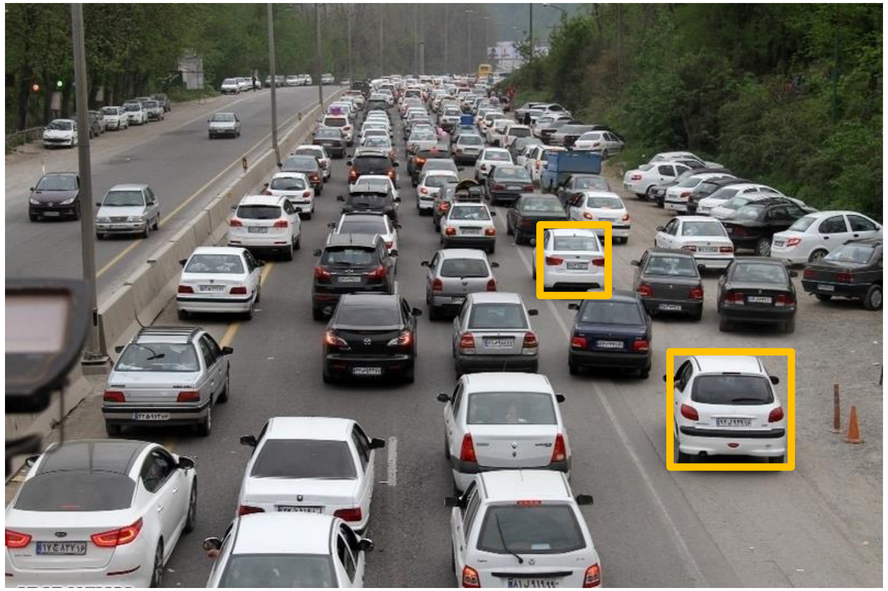
➢ What is computer vision challenges?

- View point Variation
- Difference in Illumination
- Occlusion
- Scale
- Background Clutter

Computer Vision

➢ What is computer vision challenges?

- View point Variation
- Difference in Illumination
- Occlusion
- Scale
- Background Clutter



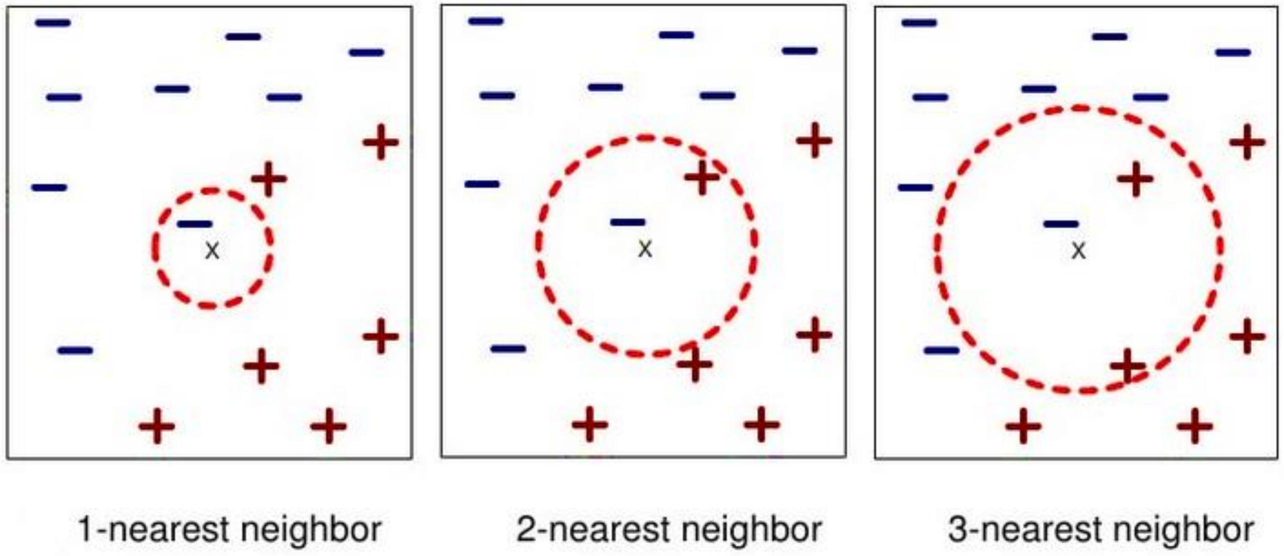
Computer Vision

➢ What is computer vision challenges?

- View point Variation
- Difference in Illumination
- Occlusion
- Scale
- Background Clutter

Common methods in ML

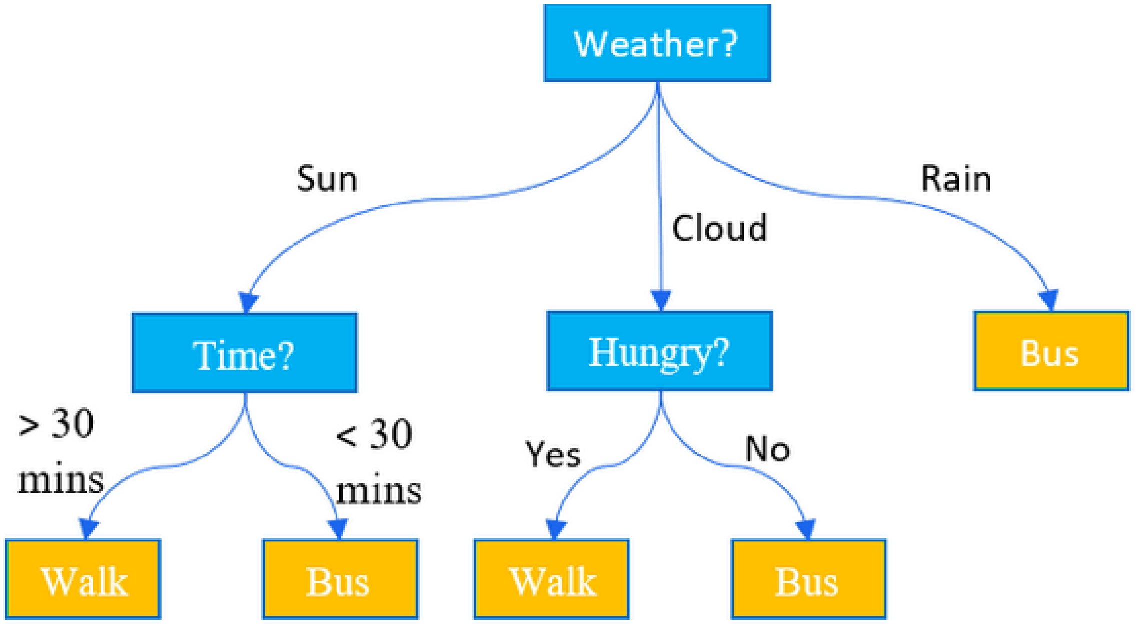
- K-nearest neighbor (KNN)
- Decision trees
- Genetic algorithm



Machine Learning (ML)

➢ Common methods in ML

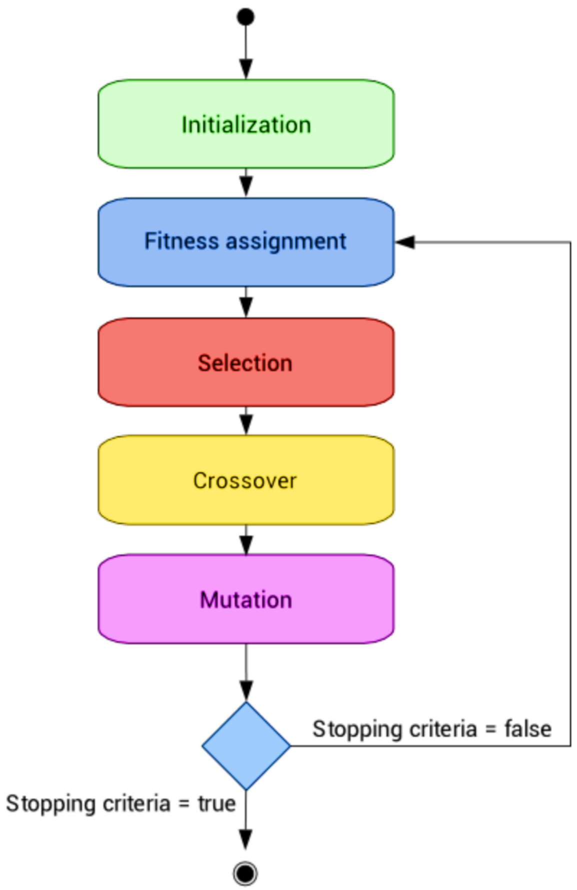
- K-nearest neighbor (KNN)
- Decision trees
- Genetic algorithm



Machine Learning (ML)

➢ Common methods in ML

- K-nearest neighbor (KNN)
- Decision trees
- Genetic algorithm



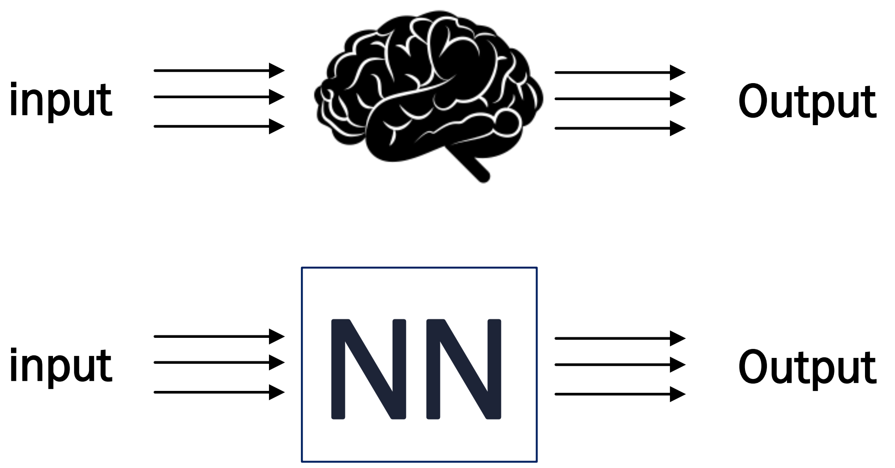
Machine Learning (ML)

- Advantage & disadvantage
  - ✓ Easily identifies trends and patterns
  - ✓ No human intervention needed (automation)
  - ✓ Handling multi-dimensional and multi-variety data
  - ✓ Wide Applications
  - × Data Acquisition
  - × Time and Resources
  - × Interpretation of Results
  - × High error-susceptibility

Neural Network (NN)

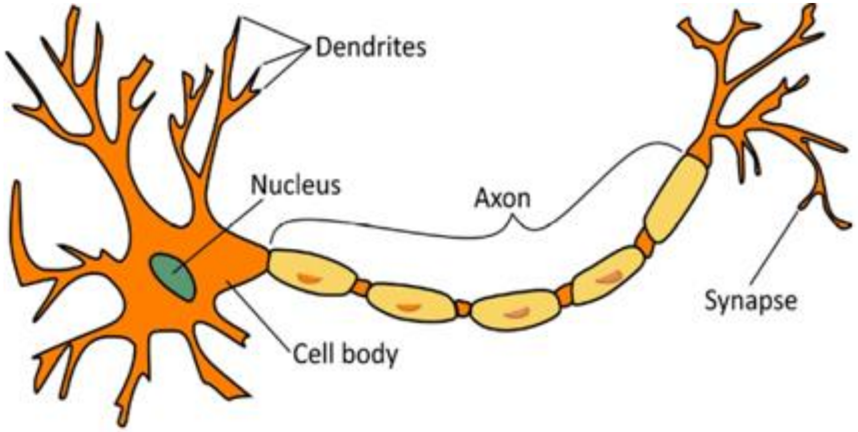
➢ What are neural networks?

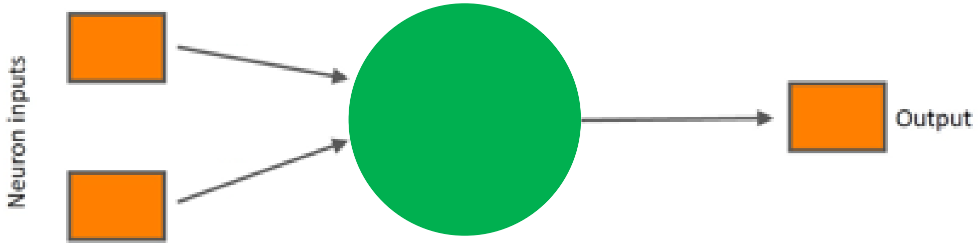
- A pool of simple processing units which communicate by sending signals to each other.
- Simplified model of the brain
- Function approximator



Neural Network (NN)

- What are neural networks?





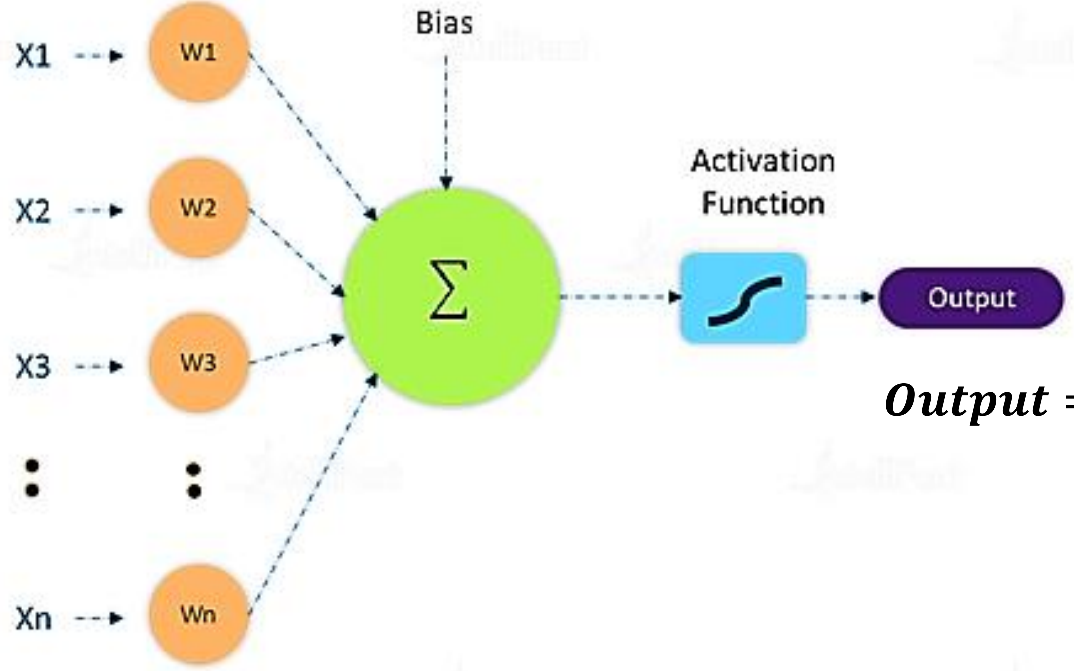
Neural Network (NN)

➢ What are they used for?

- ✓ Classification
  - Pattern recognition, feature extraction, image matching
- ✓ Noise Reduction
  - Produce noiseless outputs from input
- ✓ Prediction
  - Extrapolation based on historical data
- ✓ Learn & Generalize
  - Output based on input samples
  - produce reasonable outputs for inputs it has not learned

Neural Network (NN)

➢ How does neural network work?

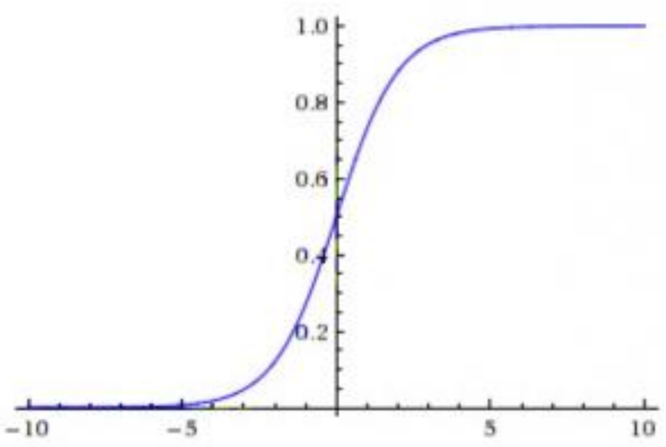


$$\boldsymbol{Output} = \boldsymbol{f}(\boldsymbol{x_1 w_1} + \dots + \boldsymbol{x_n w_n} + \boldsymbol{bias})$$

Neural Network (NN)

➢ What are activation function?

- ✓ Sigmoid Function
  - Majority of NN's use sigmoid functions
- ✓ tanh activation
- ✓ ReLU (Rectified Linear Unit)
- ✓ Sofmax

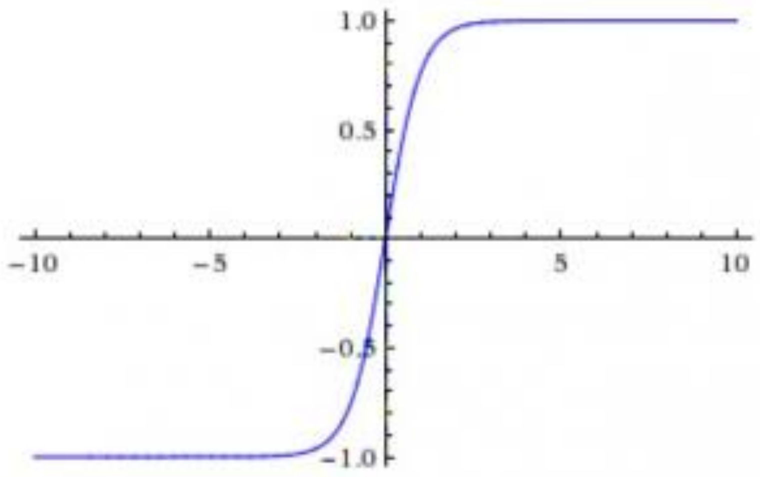


logistic function: $\boldsymbol{\sigma(x) = 1/(1 + e^{-x})}$

Neural Network (NN)

➢ What are activation function?

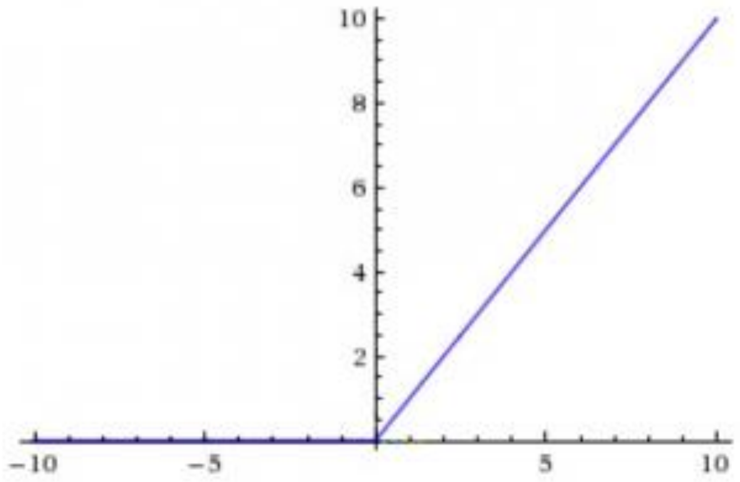
- ✓ Sigmoid Function
- ✓ tanh activation
- ✓ ReLU (Rectified Linear Unit)
- ✓ Softmax



Neural Network (NN)

➢ What are activation function?

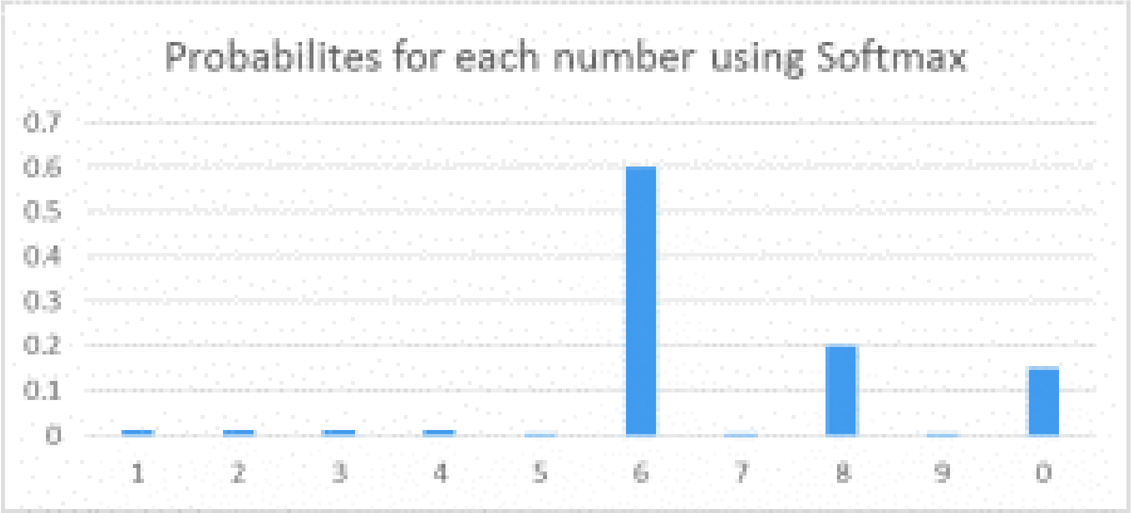
- ✓ Sigmoid Function
- ✓ tanh activation
- ✓ ReLU (Rectified Linear Unit)
- ✓ Sofmax



$$f(x) = max(0, x)$$

Neural Network (NN)

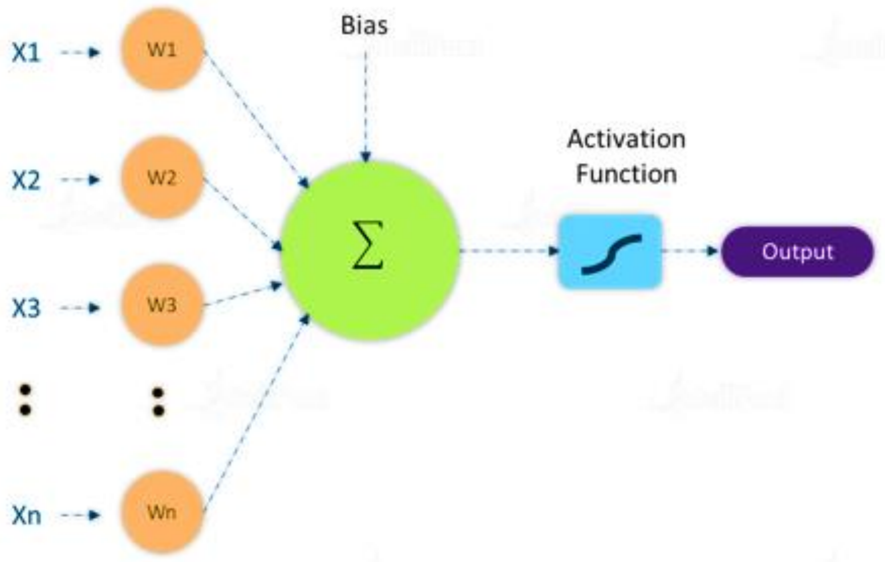
- What are activation function?
  - ✓ Sigmoid Function
  - ✓ tanh activation
  - ✓ ReLU (Rectified Linear Unit)
  - ✓ Sofmax



Neural Network (NN)

➢ Where do the weights come from?

- The weights are the most important factor in determining its function.
- Training methods
- Epoch
- Loss function



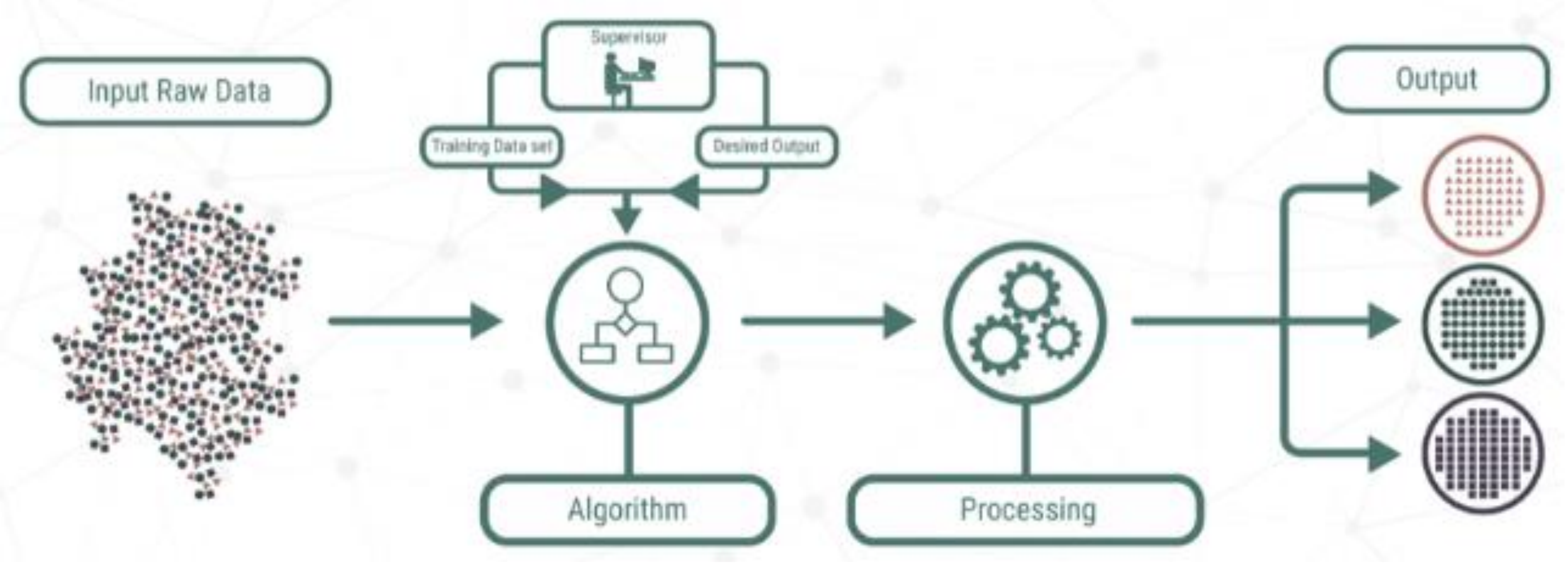
Neural Network (NN)

➢ Where do the weights come from?

❑ Two main types of training:

✓ Supervised

✓ Unsupervised



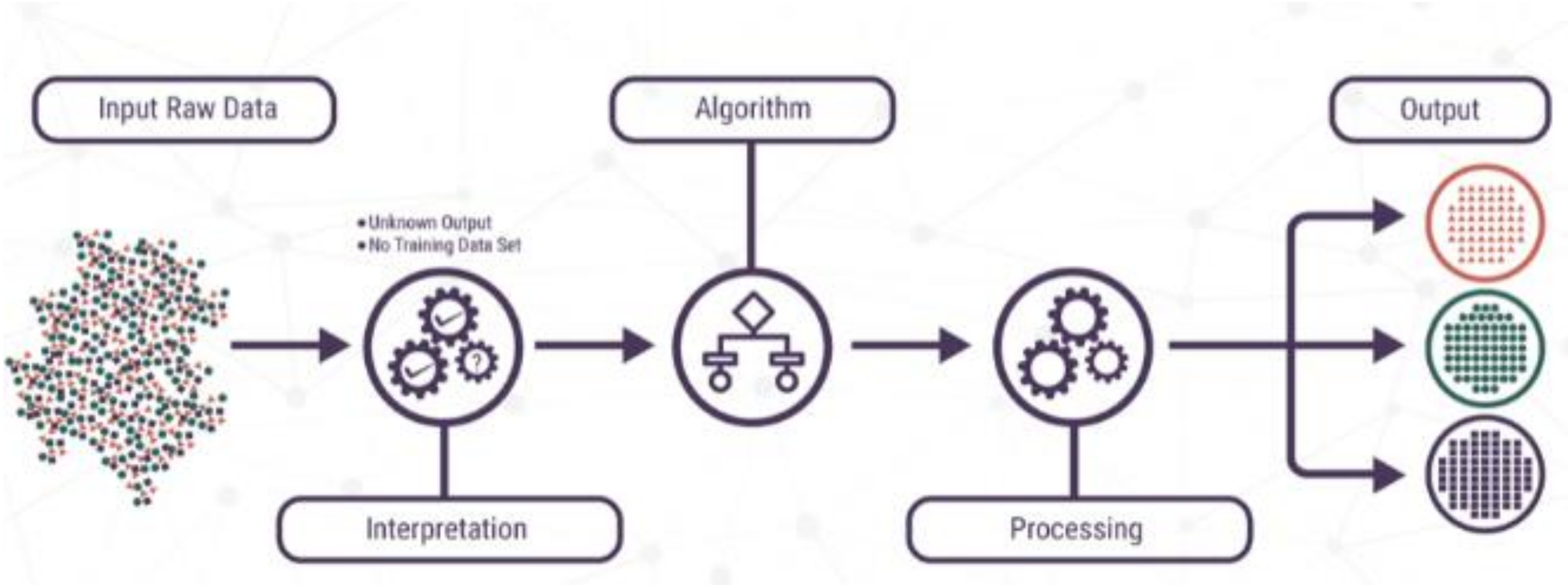
Neural Network (NN)

➢ Where do the weights come from?

❑ Two main types of training:

✓ Supervised

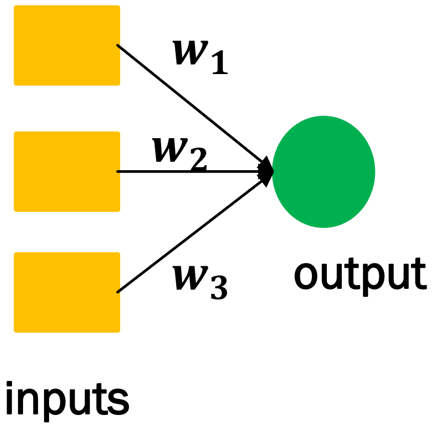
✓ Unsupervised



Neural Network (NN)

➢ What is perceptron?

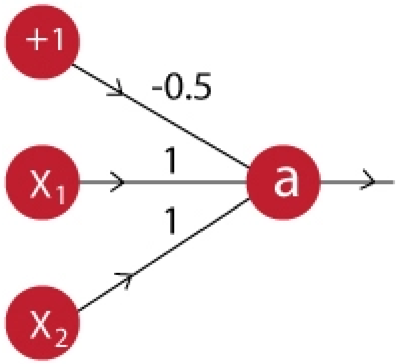
- First neural network with the ability to learn
- Made up of only input neurons and output neurons
- Output neurons use a simple threshold activation function
- In basic form, can only solve linear problems



Neural Network (NN)

- What is perceptron?
  - ✓ OR function

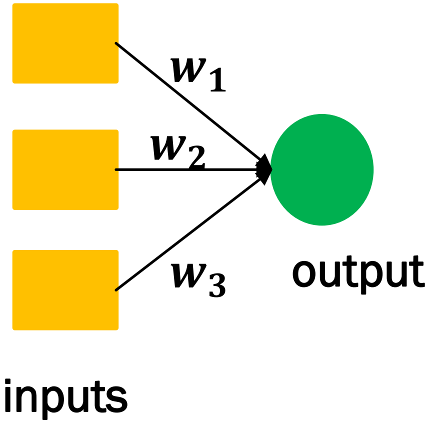
X1	X2	X1 OR X2	(-0.5+X1+X2)	a
0	0	0	-0.5	0
0	1	1	0.5	1
1	0	1	0.5	1
1	1	1	1.5	1



Neural Network (NN)

➢ How do perceptrons learn?

- Uses supervised training.
- If the output is not correct, the weights are adjusted.

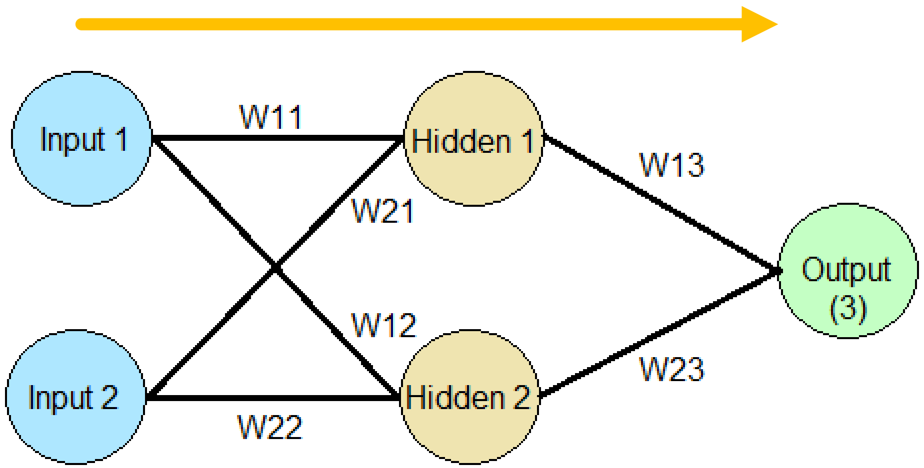


$$w_{new} = w_{old} + \alpha(desired - output) * input$$

Neural Network (NN)

➢ What is Multilayer Feedforward Networks?

- ✓ An extension of the perceptron
- ✓ Information flows in one direction



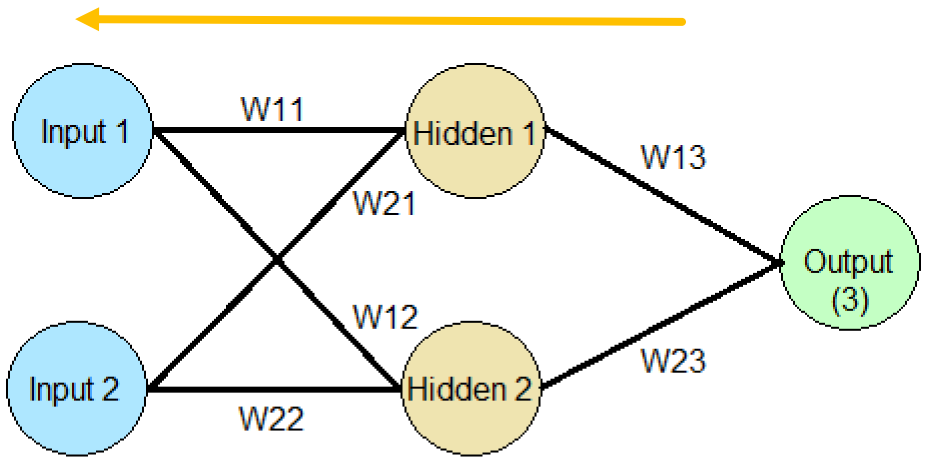
Neural Network (NN)

➢ What is Backpropagation?

- Uses supervised training.
- Based on minimizing the error of the network using the derivatives of the error function.
- Common measure of error is the mean square error:

$$\boldsymbol{E = (target - output)^2}$$

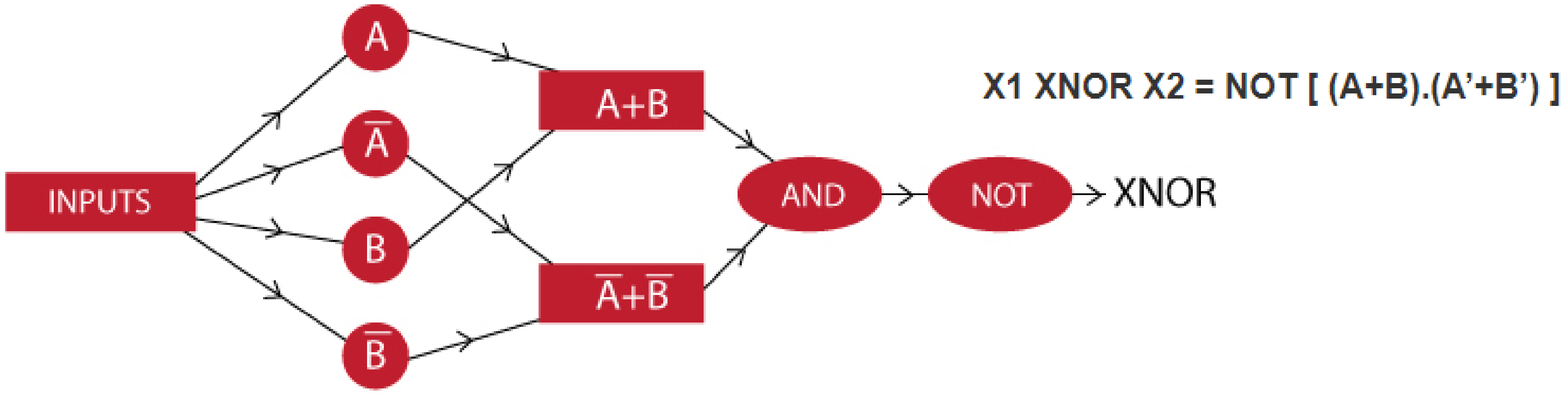
- learning rate: small or large



Neural Network (NN)

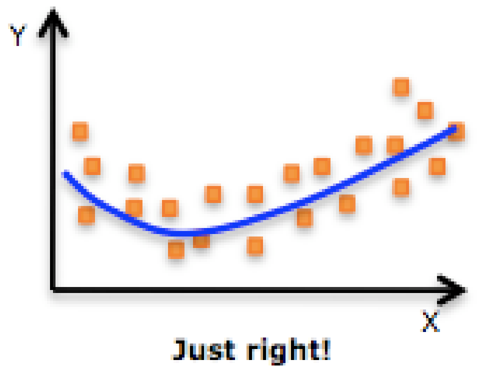
➢ Hidden layers

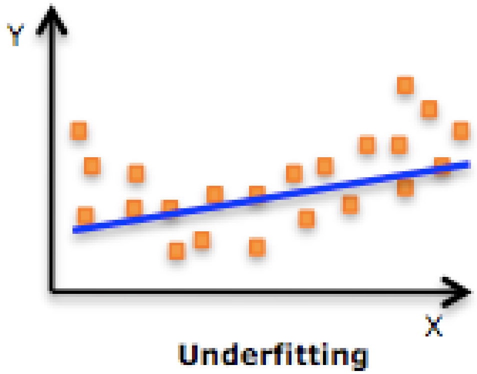
- For most problems, one layer is sufficient.
- Two layers are required when the function is discontinuous.



Neural Network (NN)

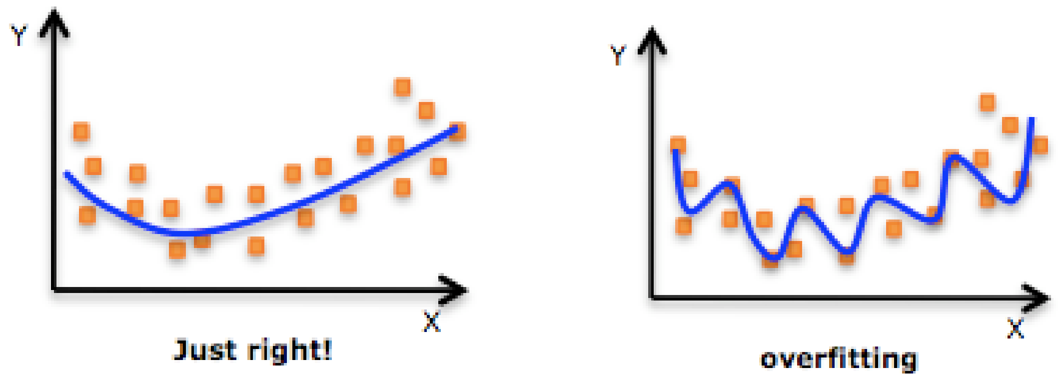
- Neurons
  - ✓ Too few
    - Underfitting
  - ✓ Too many





Neural Network (NN)

- Neurons
  - ✓ Too few
  - ✓ Too many
    - Overfitting



Neural Network (NN)

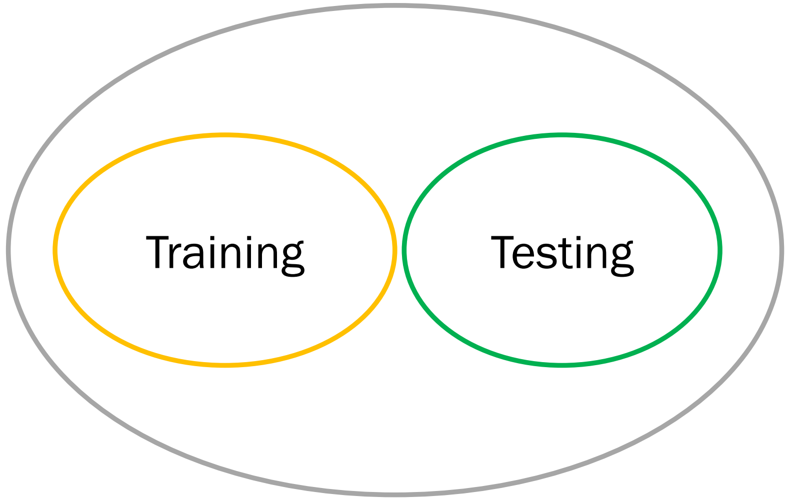
➢ How is the Training Set Chosen?

- Overfitting can also occur if a "good" training set is not chosen.
- What constitutes a "good" training set?
  - ✓ Samples must represent the general population.
  - ✓ Samples must contain members of each class.
  - ✓ Samples in each class must contain a wide range of variations or noise effect.

Neural Network (NN)

➢ Training & Testing

- ✓ Training set
  - A group of samples used to train the neural network
- ✓ Testing set
  - A group of samples used to test the performance of the neural network
  - Used to estimate the error rate



Neural Network (NN)

➢ Verification

- ✓ Provides an unbiased test of the quality of the network.
- ✓ Common error is to "test" the neural network using the same samples that were used to train the neural network.
  - The network was optimized on these samples.
  - Doesn't give any indication as to how well the network will be able to classify inputs that weren't in the training set.

Neural Network (NN)

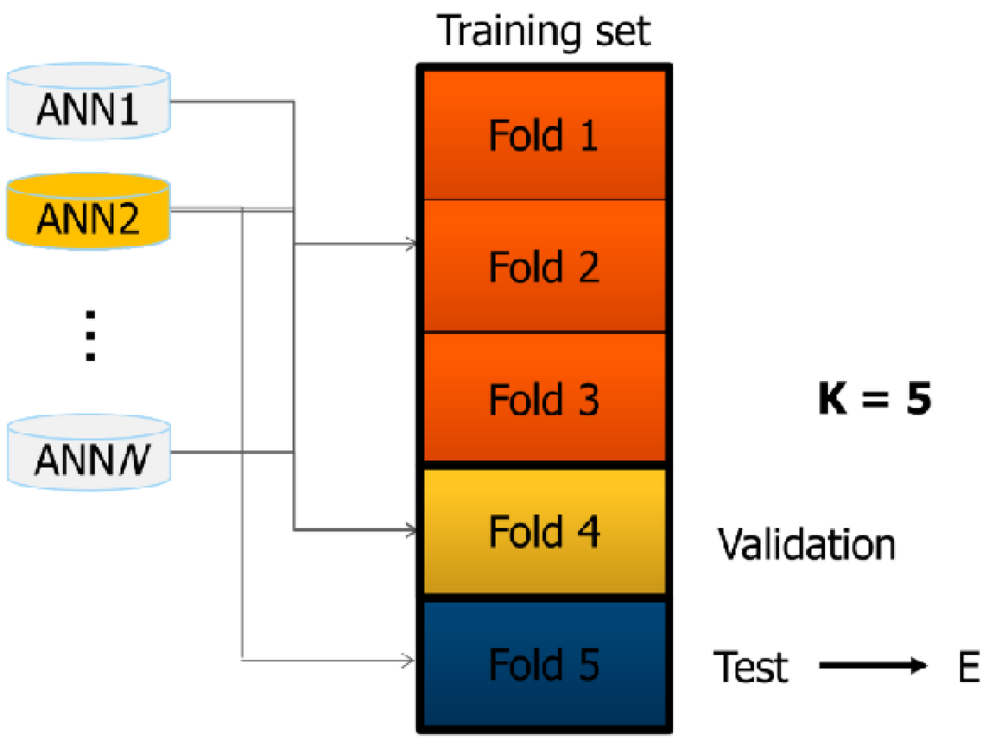
➢ Verification

- ✓ Various metrics can be used to grade the performance of the neural network.
  - Mean square error, SNR, etc.
- ✓ Resampling is an alternative method of estimating error rate of the neural network.
  - Iterate the training and testing procedures multiple times.
  - Two main techniques are used:
    - Cross-Validation
    - Bootstrapping

Neural Network (NN)

➢ Cross-Validation

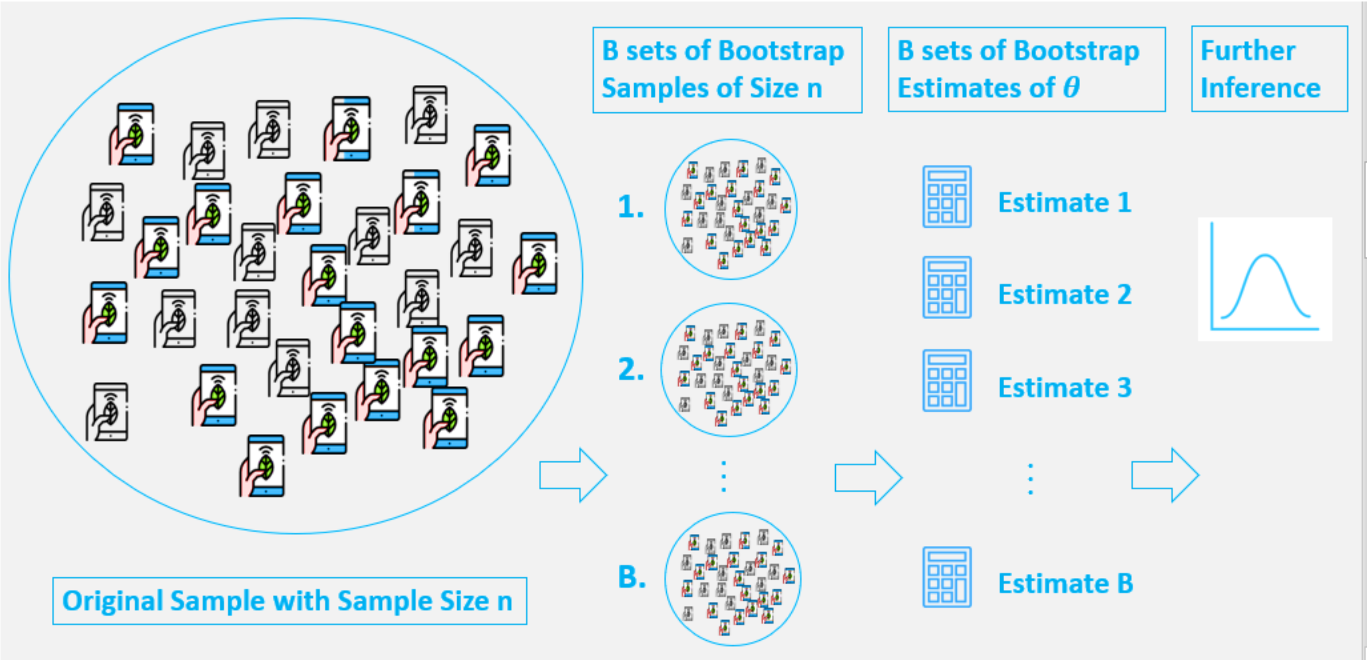
- ✓ Divide all samples into K folds.
- ✓ Overall performance comes from all folds performance.
- ✓ The network prevents overfitting by cross-validation.



Neural Network (NN)

➢ Bootstrapping

- ✓ Resampling method that sample from population with sample size n.
- ✓ containing so many statistic topics



Neural Network (NN)

- ✓ High cost
- ✓ Difficult to train as the number of hidden layers increases.
- ✓ Stuck in local optima.
- ✓ The random initialization does not guarantee starting from the proximity of global optima.

Solution

Deep Learning

Neural Network (NN)

- Generating Image Descriptions

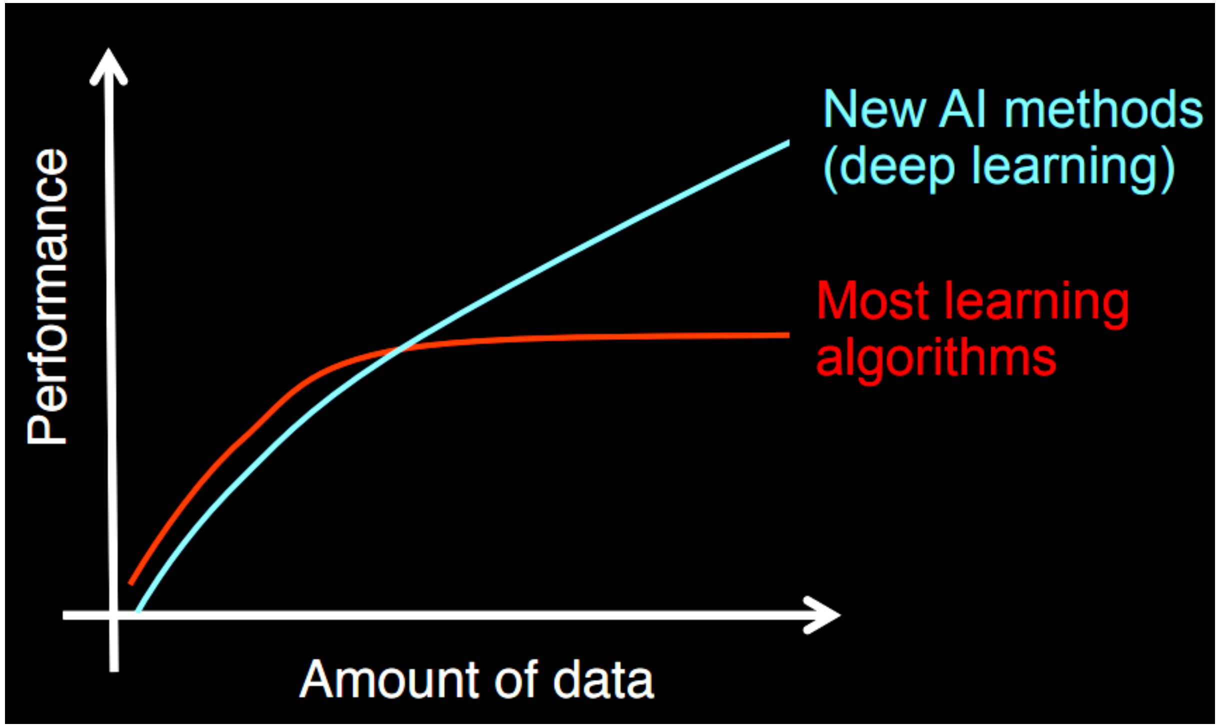
Neural Network (NN)

➢ Modelling Complex Behaviors

- ✓ Too many concepts to learn
  - Too many object categories
  - Too many ways of interaction between objects categories
- ✓ Behavior is a highly varying function underlying factors

Deep Learning (DL)

➢ Performance & Amount of data



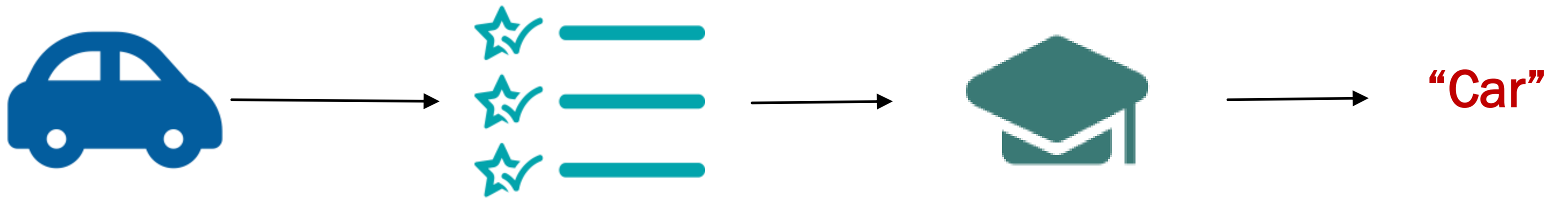
Deep Learning (DL)

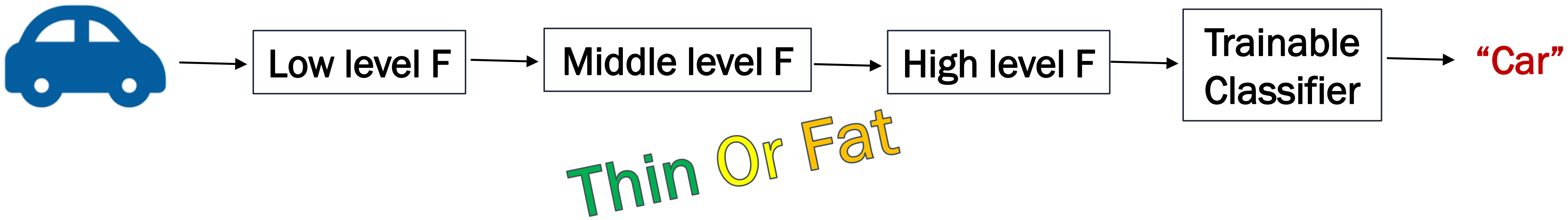
- Three key ideas
  - ✓ Hierarchical
  - ✓ End-to-End Learning
  - ✓ Distributed Representations

Deep Learning (DL)

➢ Hierarchical

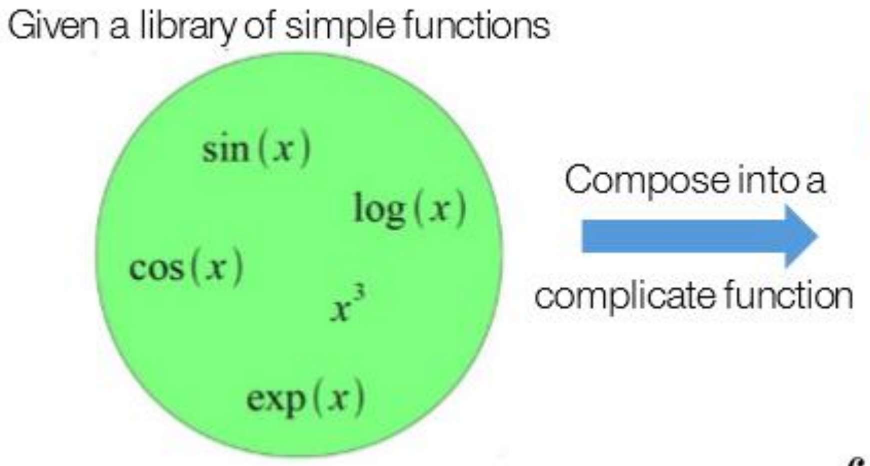
- ✓ Cascade of non-linear transformations
- ✓ Multiple layers of representations





Deep Learning (DL)

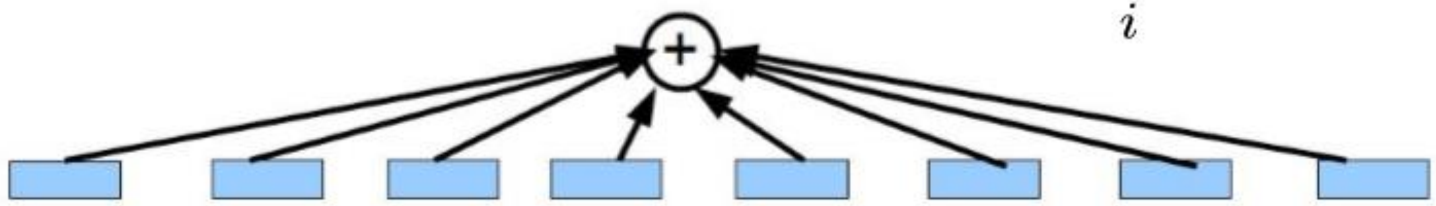
➢ Building a Complicated Function



Idea 1: Linear Combinations

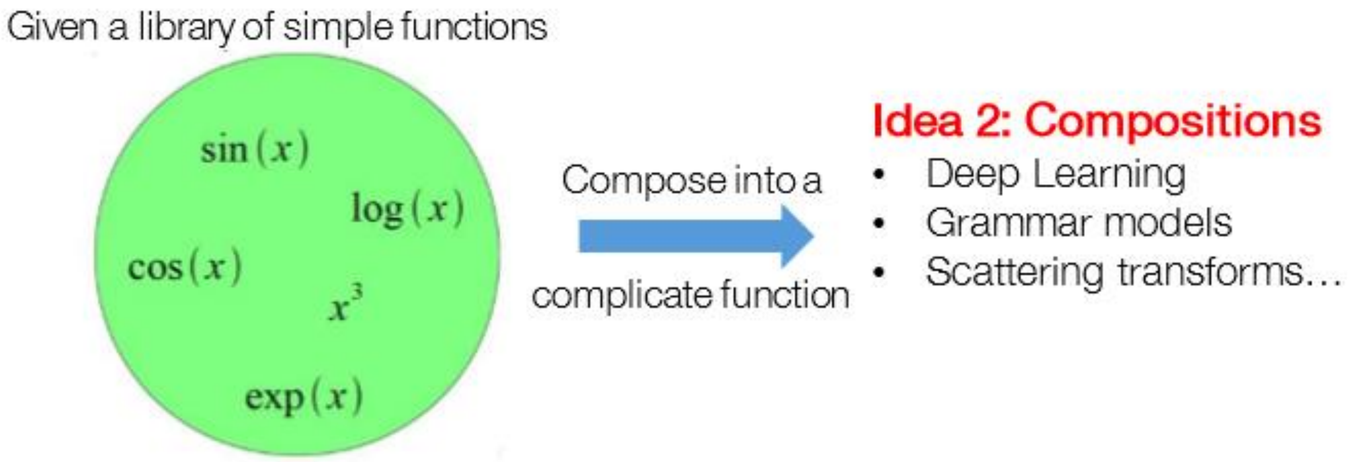
- Boosting
- Kernels
- …

$$f(x) = \sum_i \alpha_i g_i(x)$$

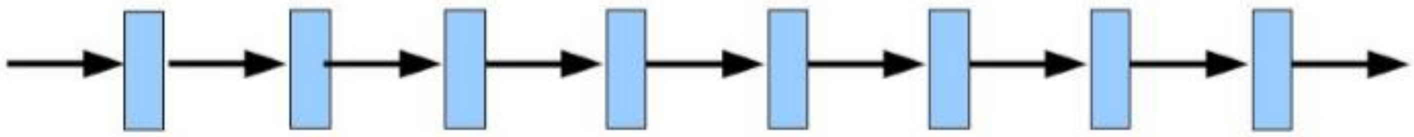


Deep Learning (DL)

➢ Building a Complicated Function

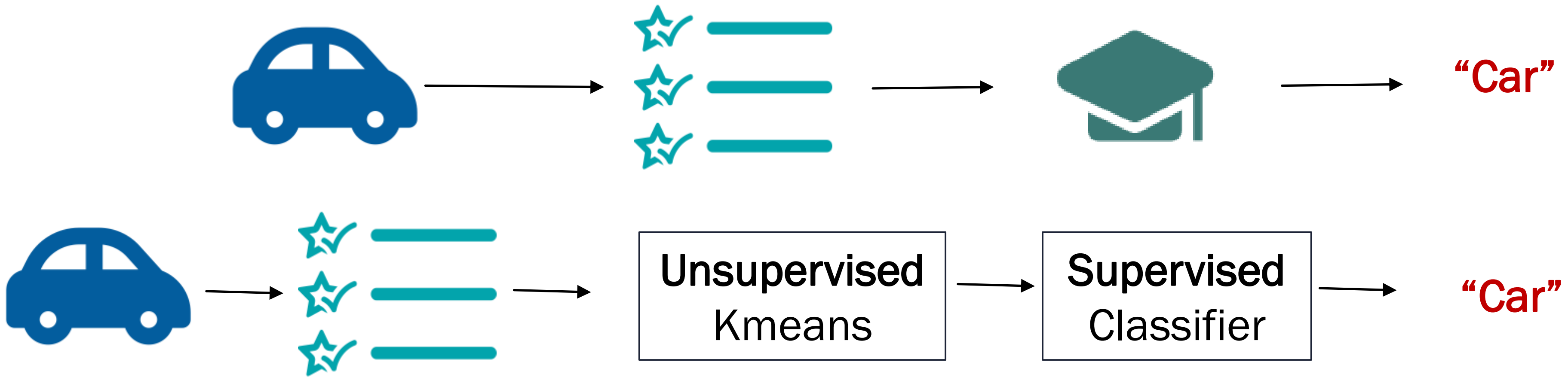


$$f(x) = g_1(g_2(\ldots(g_n(x)\ldots))$$



Deep Learning (DL)

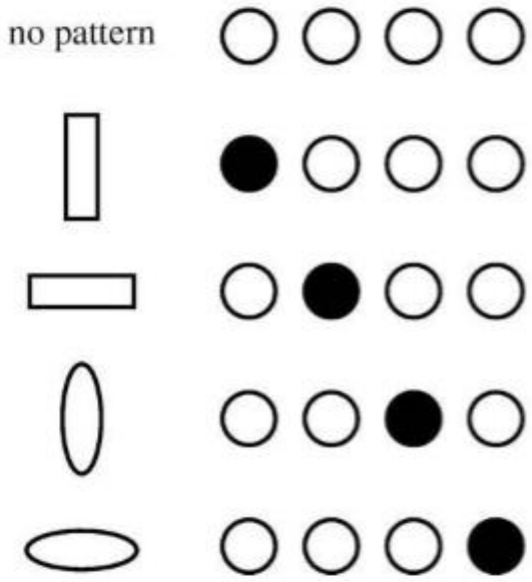
- End-to-End Learning
  - ✓ Learning (goal-driven) representations
  - ✓ Learning to feature extract



Deep Learning (DL)

➢ Representation

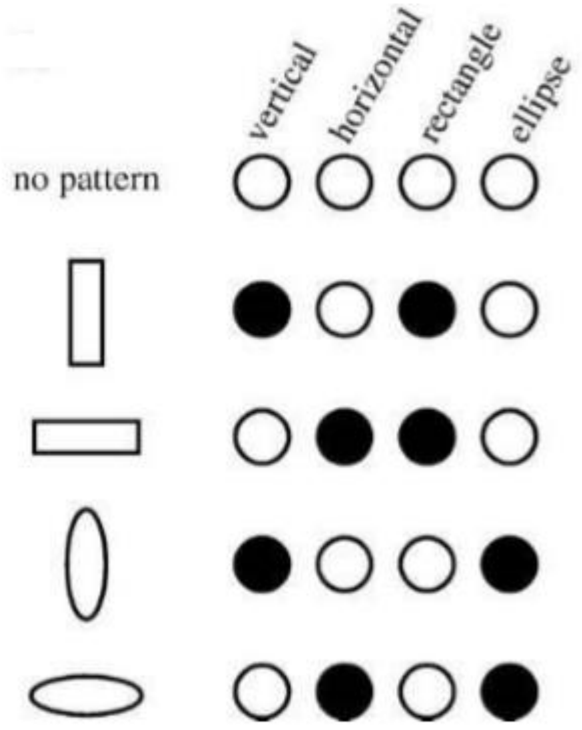
- Local
  - ✓ Each cluster corresponds to one neuron.
  - ✓ Easy to understand/learn/code.
  - ✓ Easy to associate with other representations or responses.
  - × But, very inefficient whenever the data has componential structure
- Distributed



Deep Learning (DL)

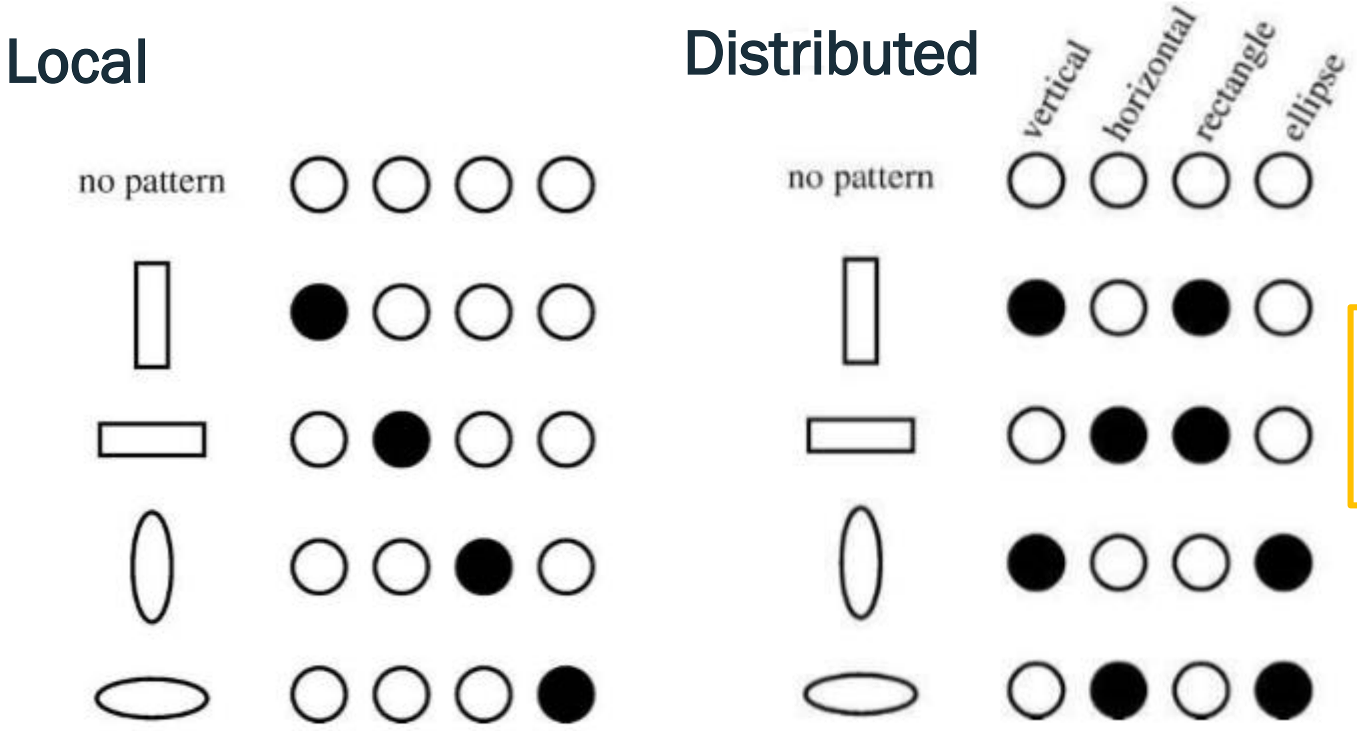
➢ Representation

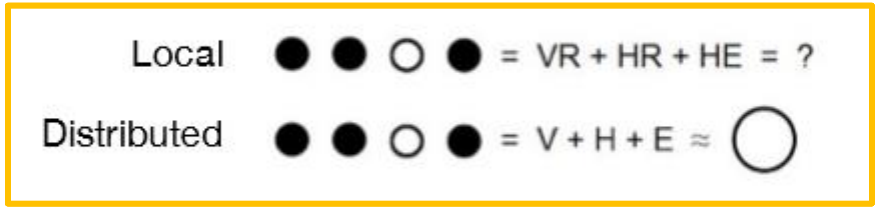
- Local
- Distributed
  - Each concept is represented by many neurons
  - Each neuron participates in the representation of many concepts



Deep Learning (DL)

➢ Representation





Deep Learning (DL)

- Deep learning methods
  - Unsupervised methods
    - Restricted Boltzmann Machines
    - Deep Belief Networks
    - Auto encoders
  - Supervised methods

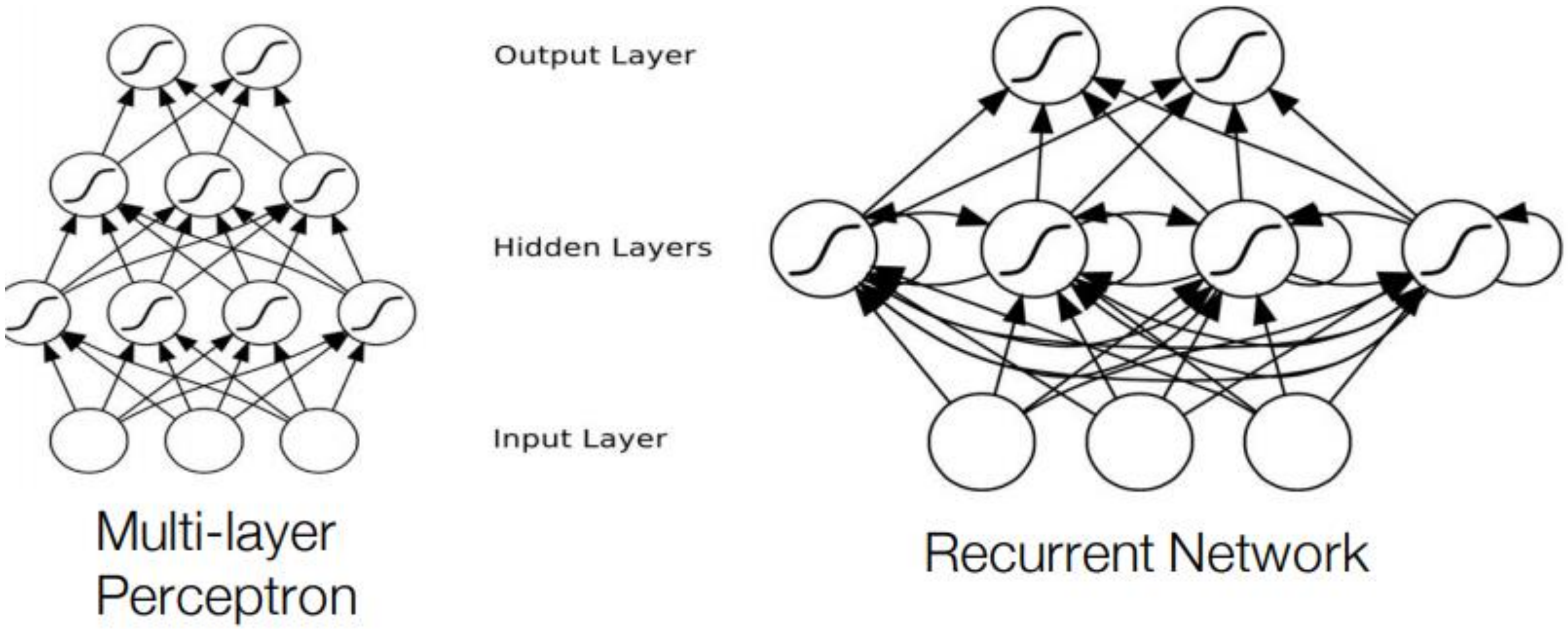
Deep Learning (DL)

- Deep learning methods
  - ✓ Unsupervised methods
  - ✓ Supervised methods
    - Deep Neural Networks
    - Recurrent Neural Networks
    - Convolutional Neural Networks

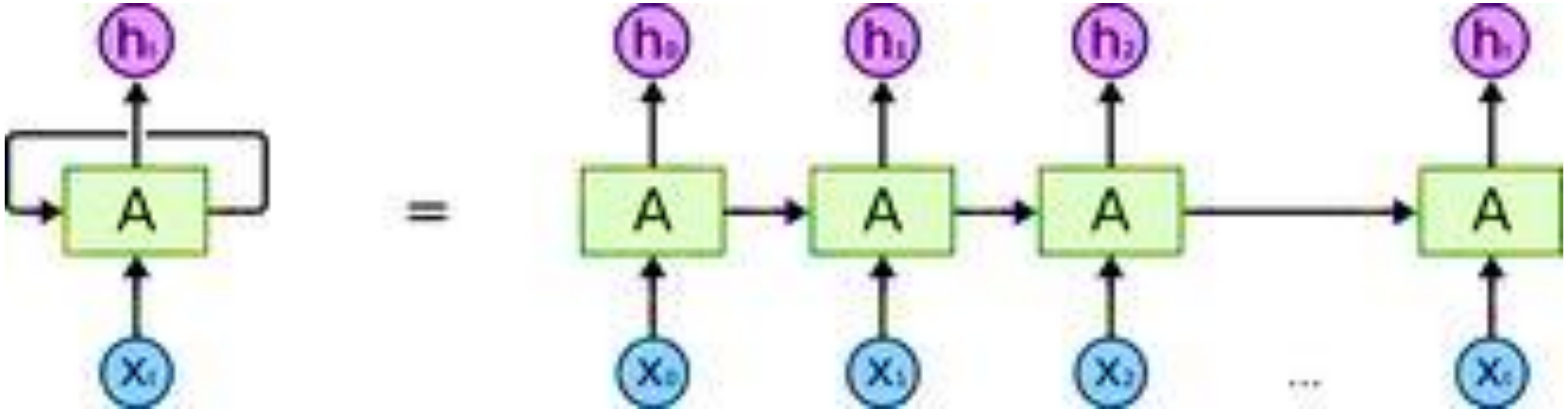
Deep Learning (DL)

➢ Recurrent neural network (RNN)

- ✓ Map from the entire history of previous inputs to each output.



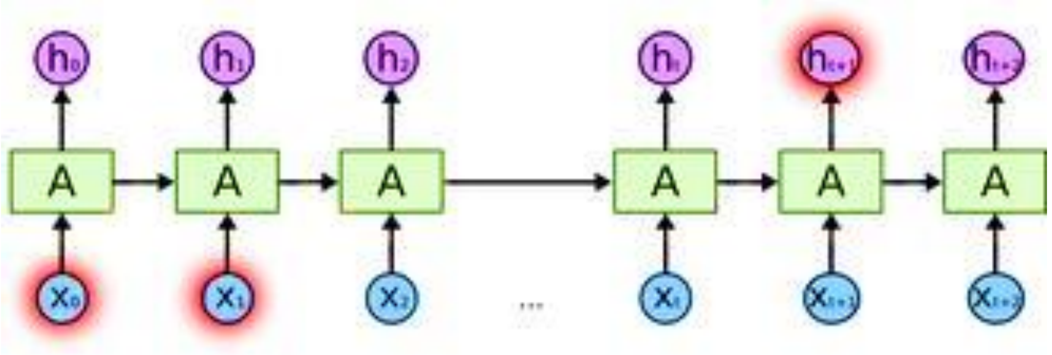
- Unrolled RNN



Deep Learning (DL)

➢ The Problem of Long-Term Dependencies

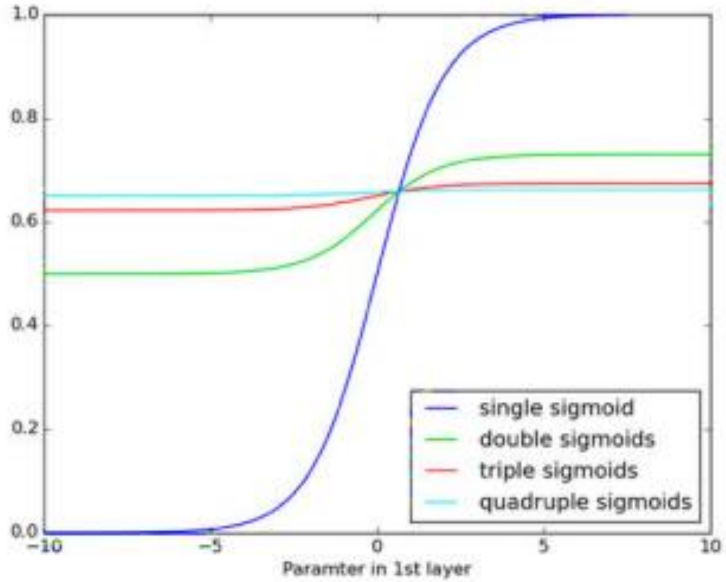
- ✓ Map appeals of RNNs is the idea that they might be able to connect previous information to the present task.
- ✓ The gap between the relevant information and the place that it's needed is small/large.
- ✓ When that gap grows, RNNs become unable to learn to connect the information.



Deep Learning (DL)

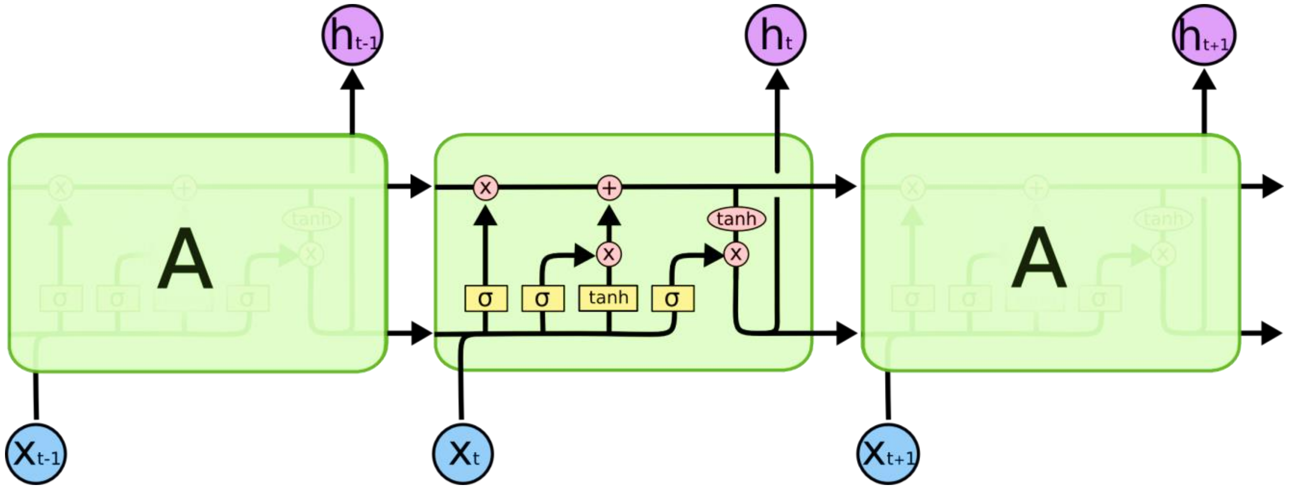
➢ The Problem of Vanishing and Exploding Gradient

- If the gradients are large
  - ❑ Exploding gradients, learning diverges
  - ✓ Solution: Clip the gradients to a certain max value.
- If the gradients are small
  - ❑ Vanishing gradients, learning very slow or stops
  - ✓ Solution: introducing memory via LSTM, GRU, etc.



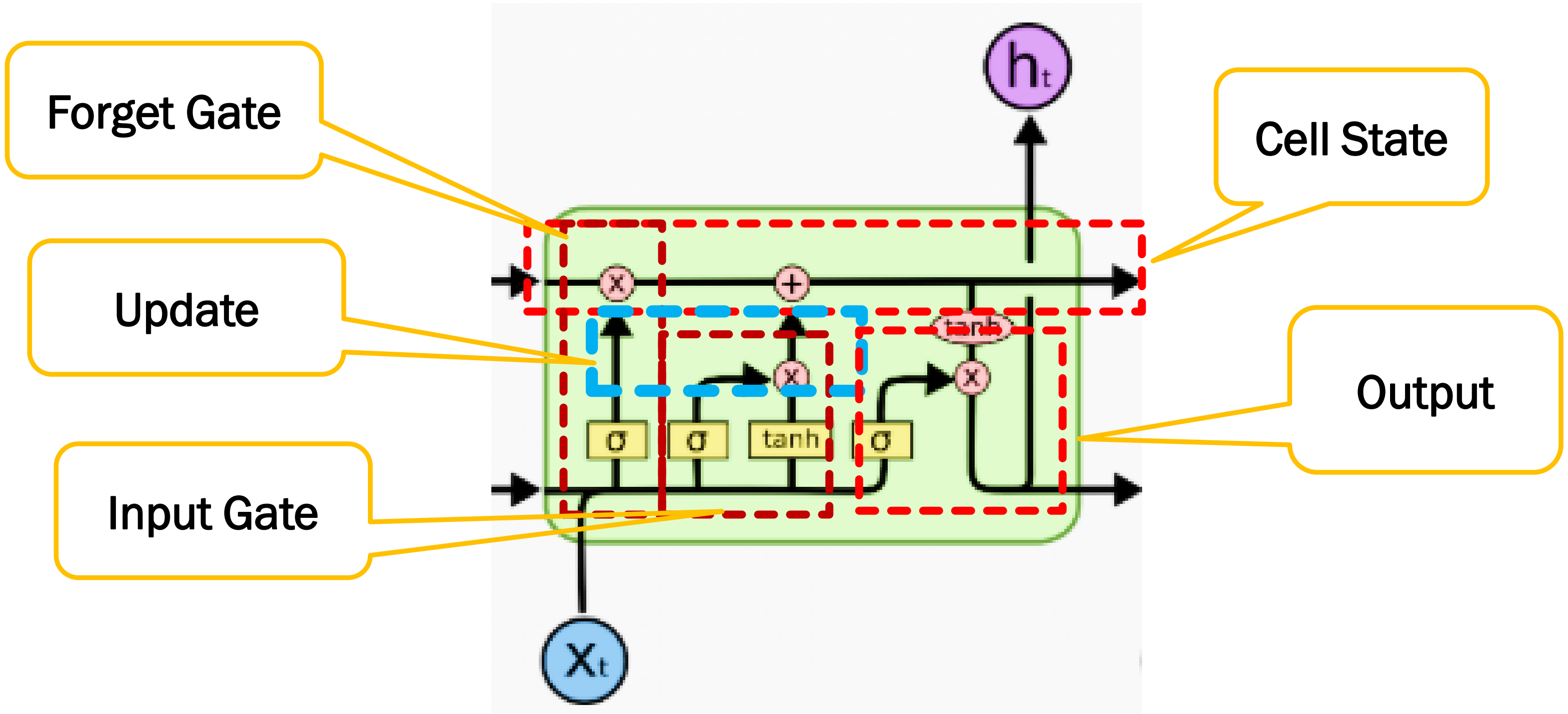
➢ Long-Short Term Memory (LSTM) networks

- ✓ LSTMs are explicitly designed to avoid the long-term dependency problem.
- ✓ Remembering information for long periods of time is practically their default behavior, not something they struggle to learn!



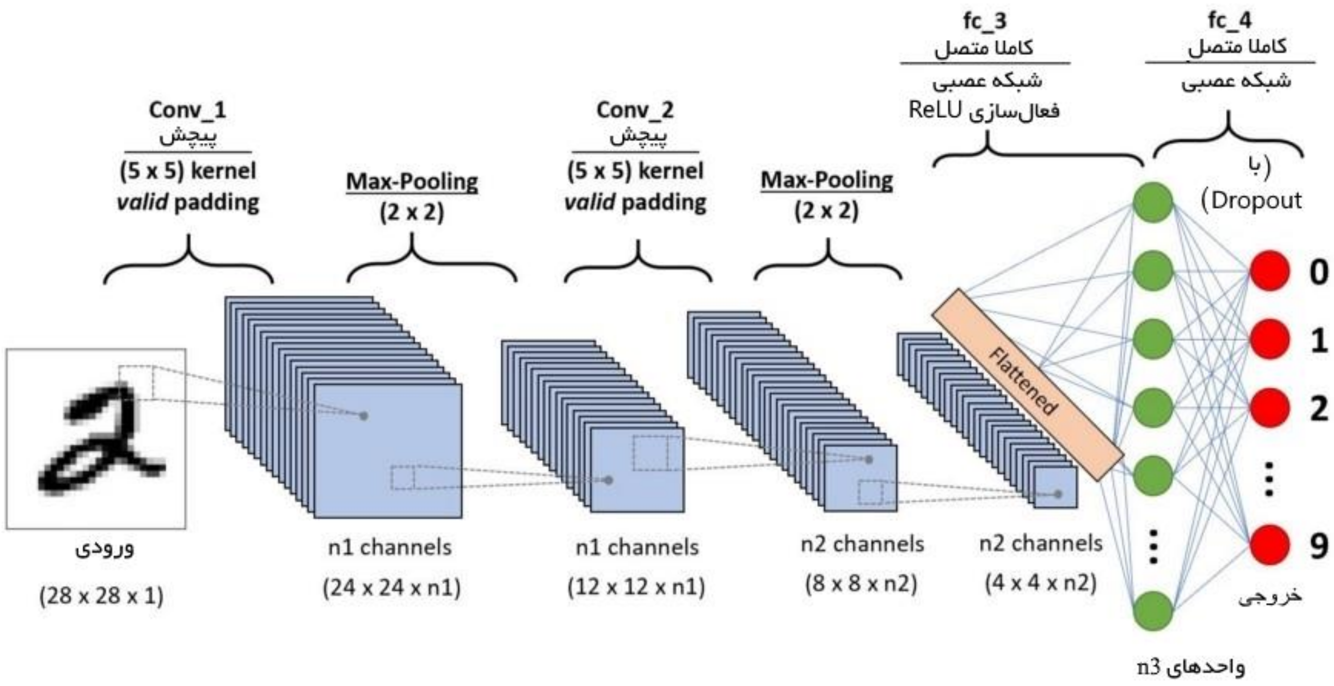
Deep Learning (DL)

➢ Step-by-Step LSTM Walk Through



➢ Convolutional Neural Network (CNN)

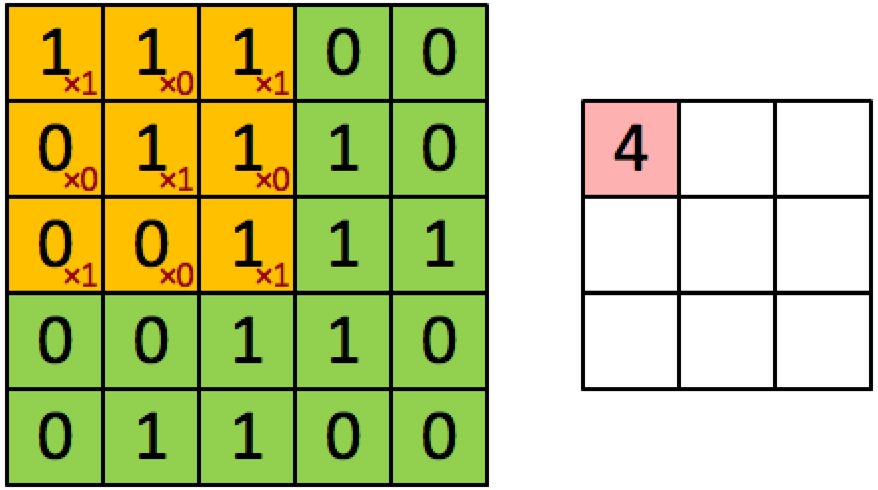
- ✓ CNN algorithm requires less pre-processing than other clustering algorithms.



Deep Learning (DL)

➢ Convolve Layer

- Kernel (3×3)
- Image size (5×5)
- Reduce feature space



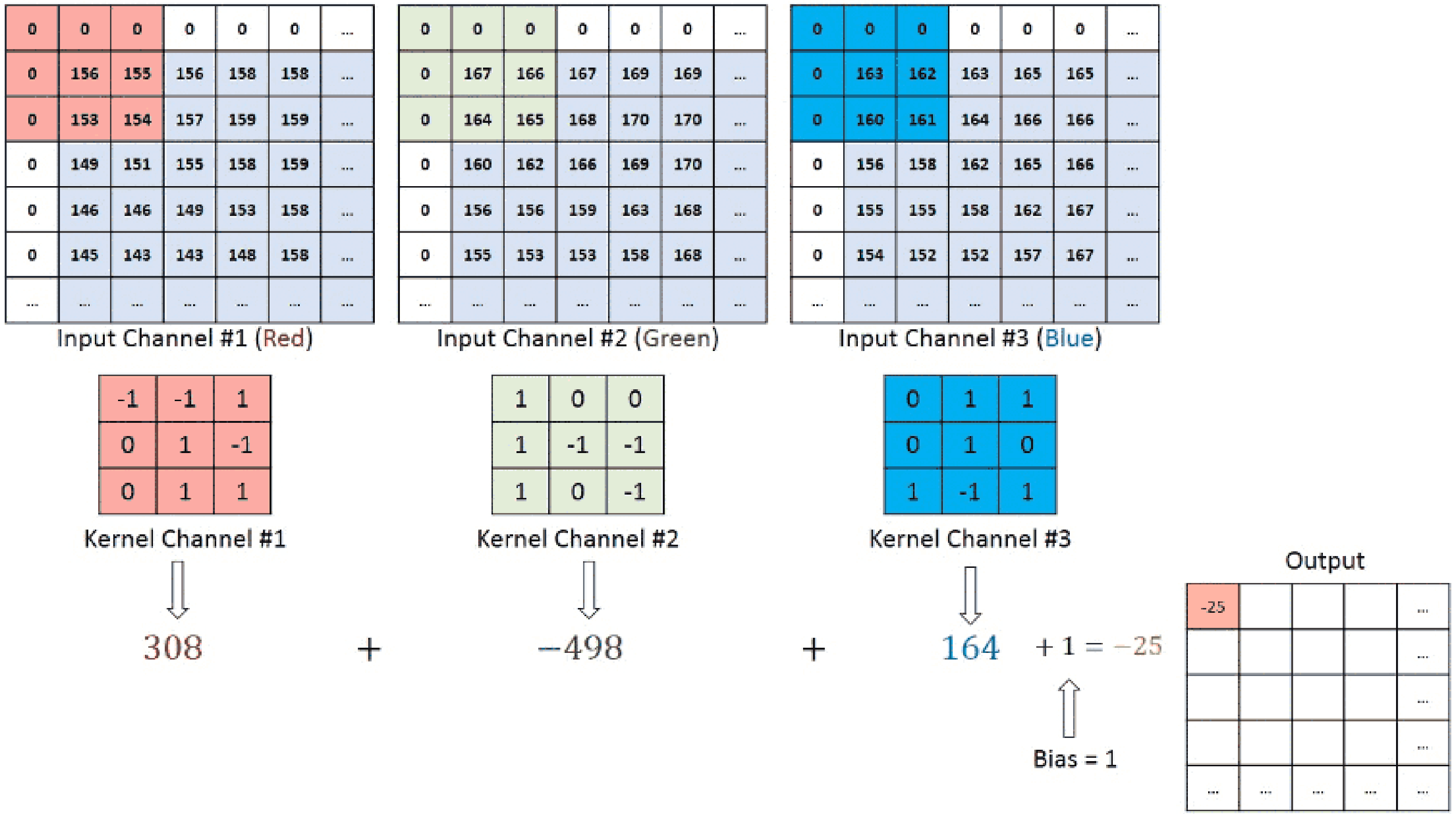
Image

Convolved Feature

Deep Learning (DL)

➢ Convolve Layer

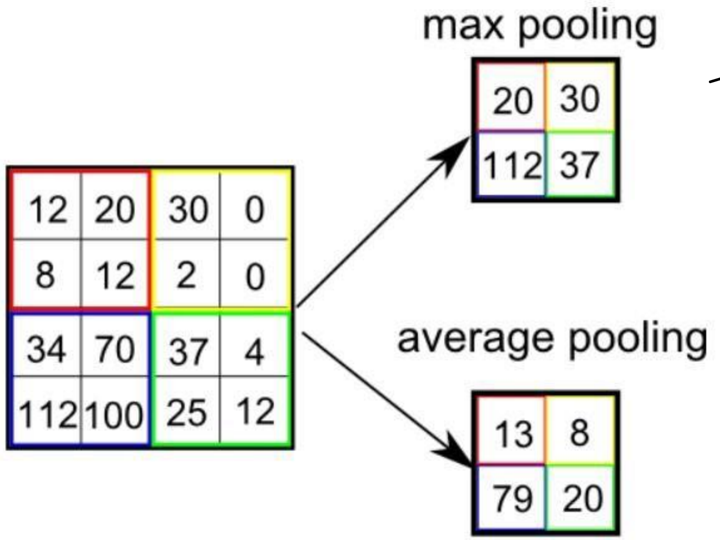
- Kernel (3×3×3)
- Matrix size (M×N×3)
- Extract features

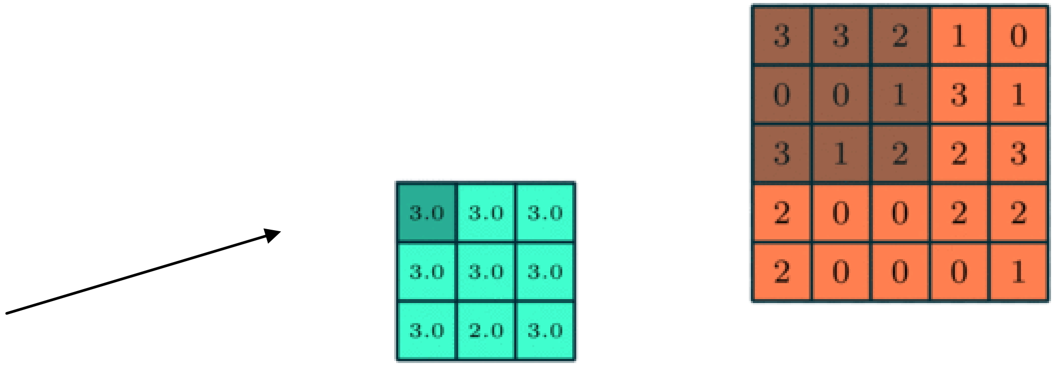


Deep Learning (DL)

➢ Pooling

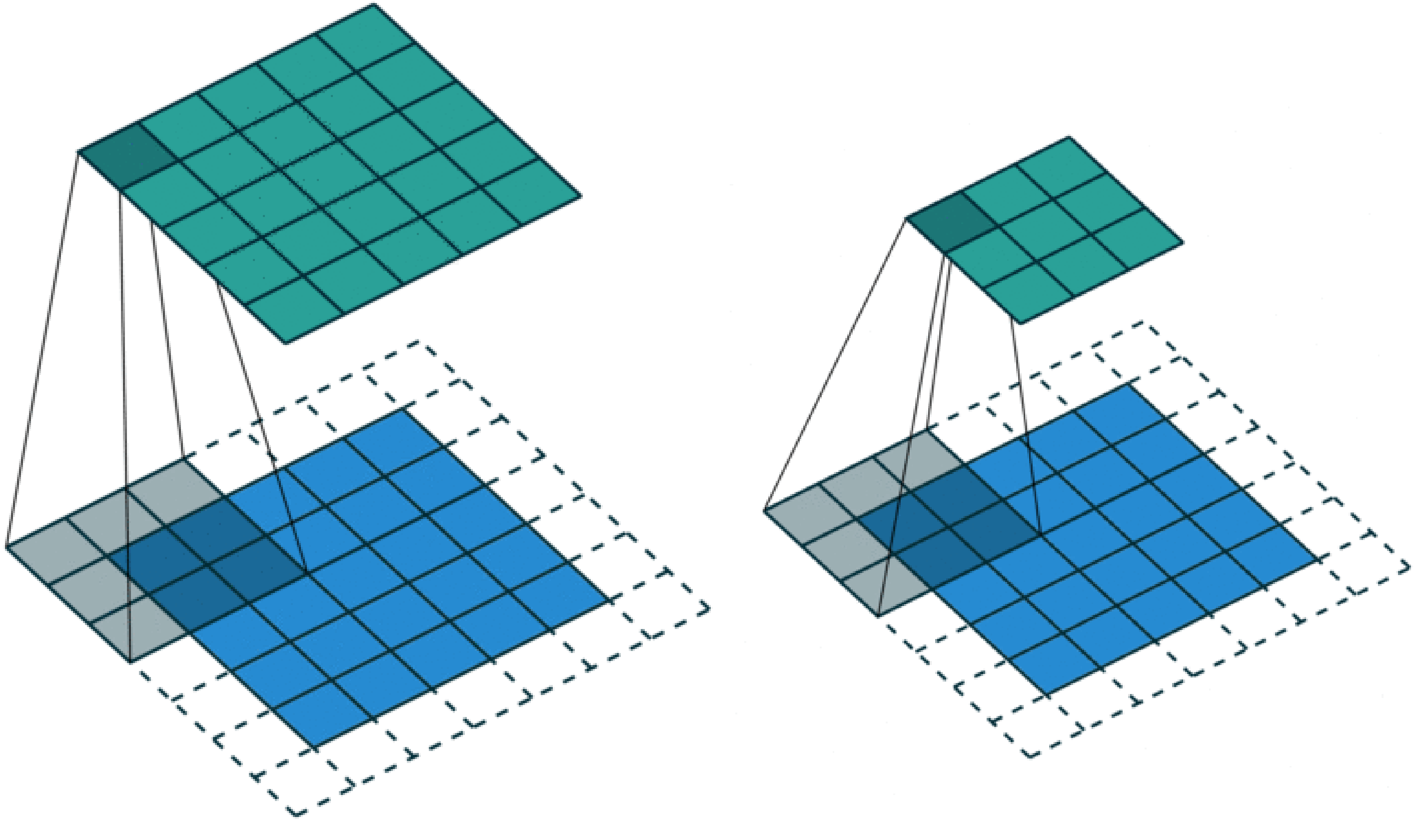
- Feature space Reduction
- Denoising





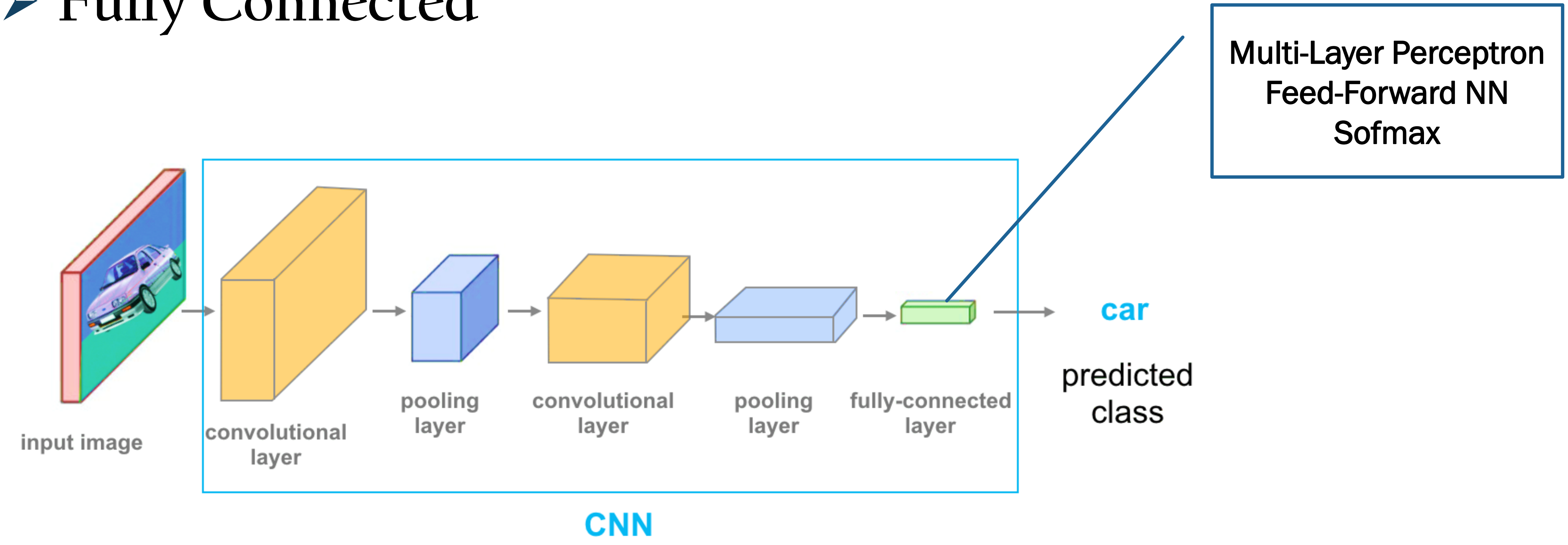
Deep Learning (DL)

- Padding
  - ✓ Same padding
  - ✓ Padding, strides



Deep Learning (DL)

- Fully Connected



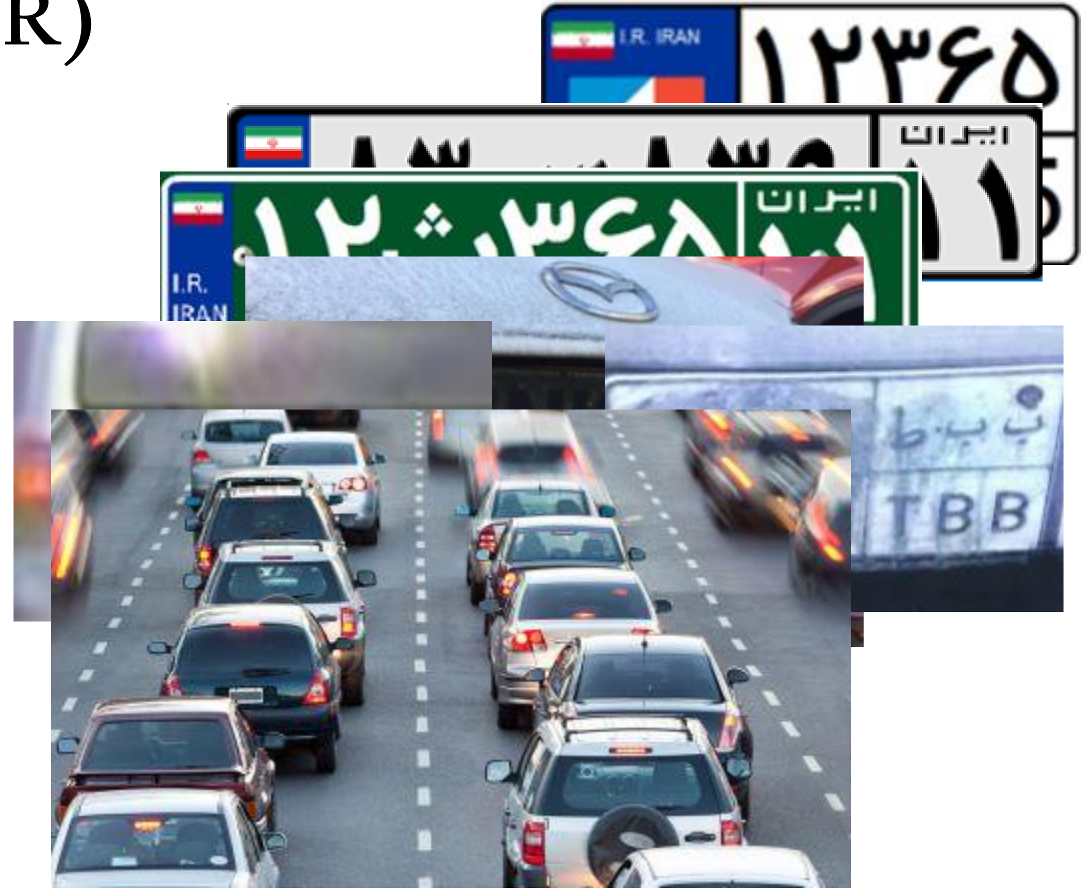
Discussion & Result

- License Plate Recognition (LPR)

Discussion & Result

➢ License Plate Recognition (LPR)

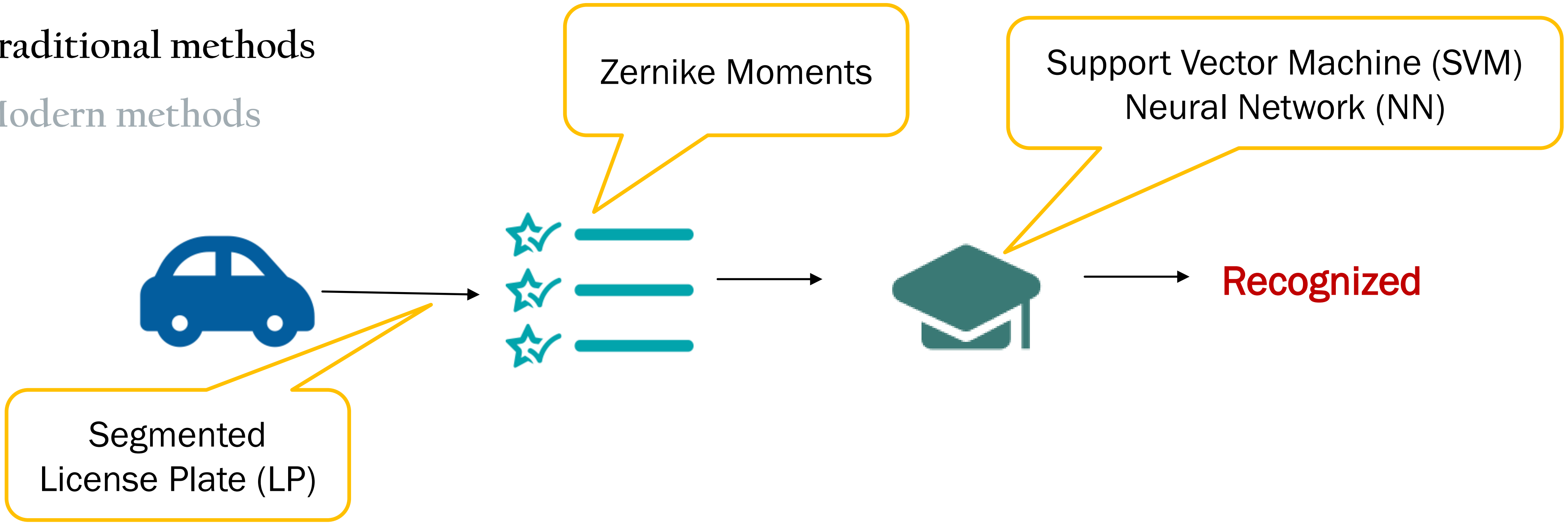
- ✓ Challenges
  - Style & font
  - Image quality
  - Occlusion
- ✓ Traditional methods
- ✓ Modern methods



Discussion & Result

➢ License Plate Recognition (LPR)

- ✓ Challenges
- ✓ Traditional methods
- ✓ Modern methods

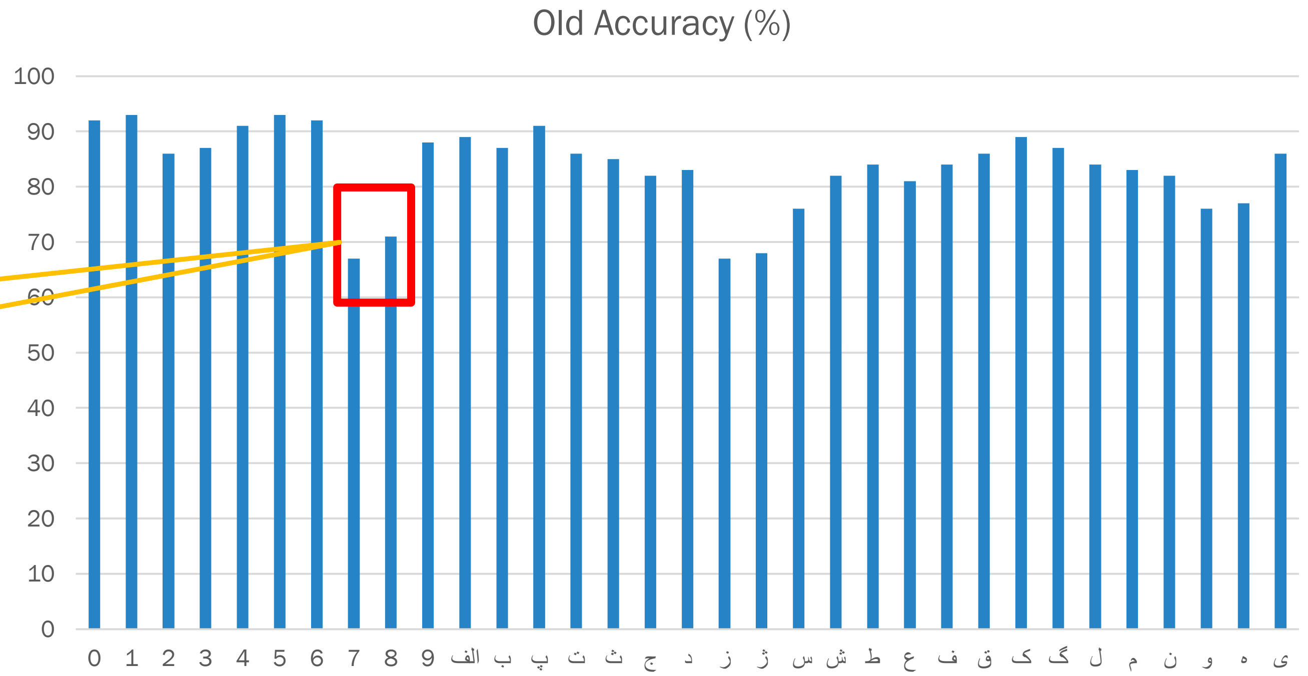


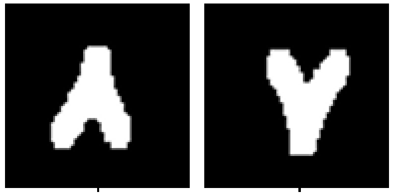
Discussion & Result

➢ License Plate Recognition (LPR)

- ✓ Challenges
- ✓ Traditional methods
- ✓ Modern methods

Rotation Invariant



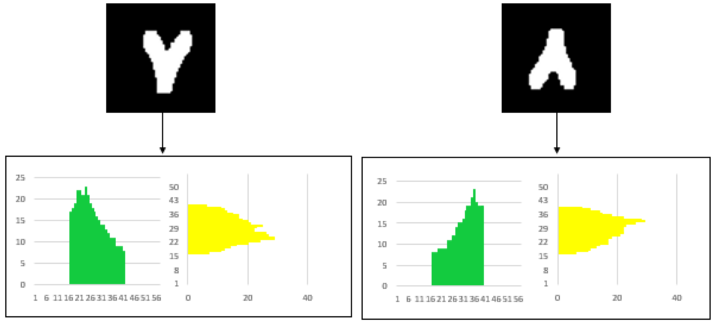


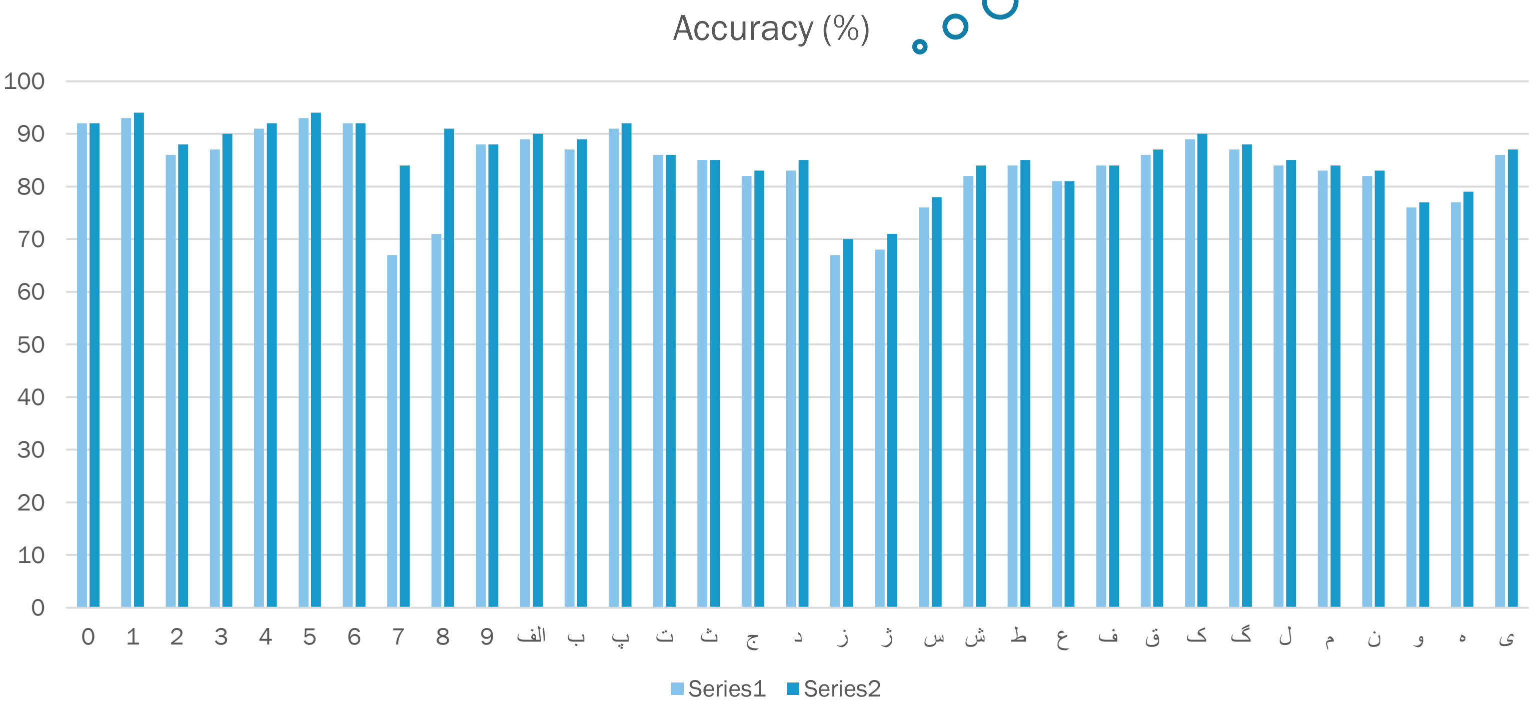
Discussion & Result

➢ License Plate Recognition (LPR)

- ✓ Challenges
- ✓ Traditional methods
- ✓ Modern methods

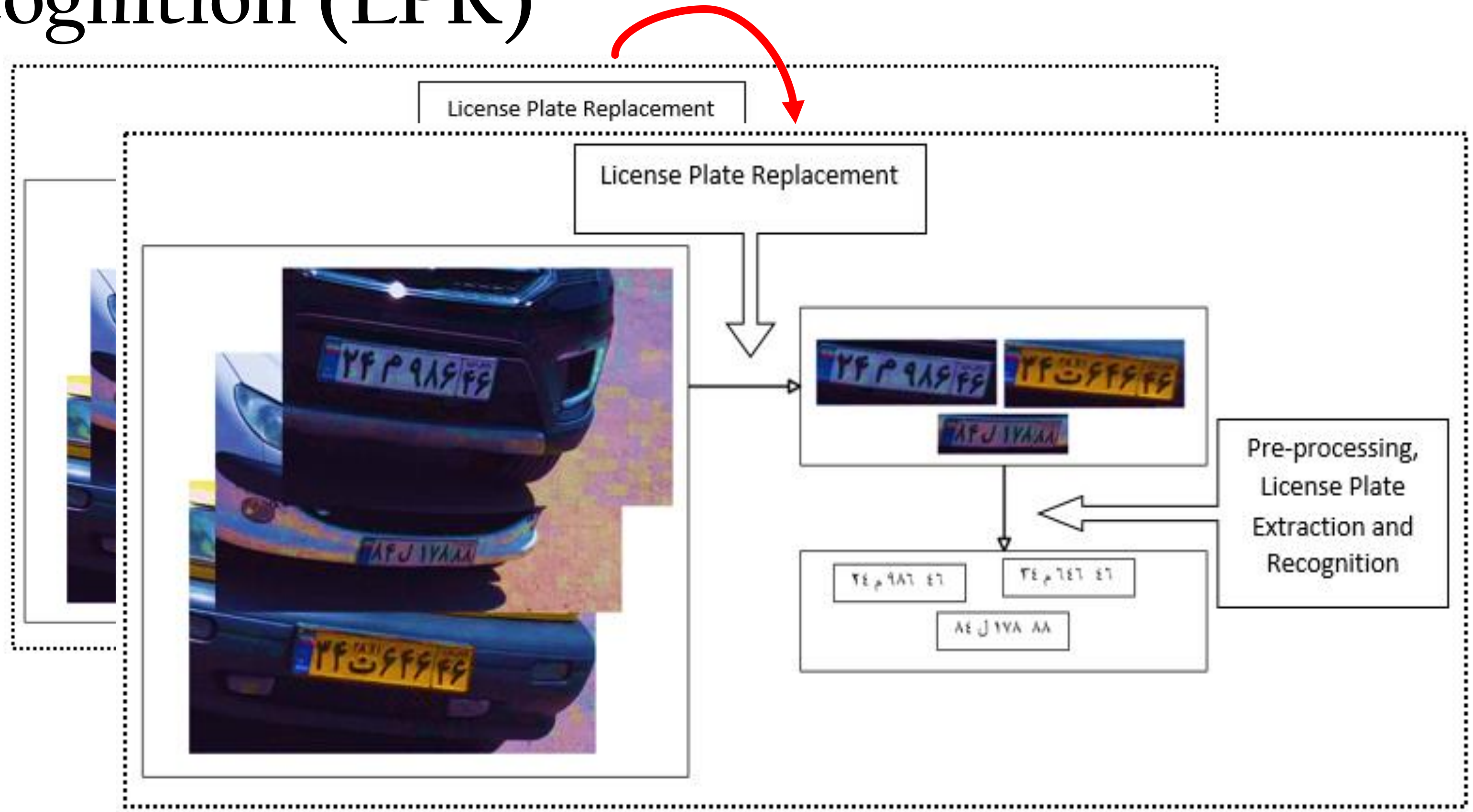
Adding Crossing Count feature





Discussion & Result

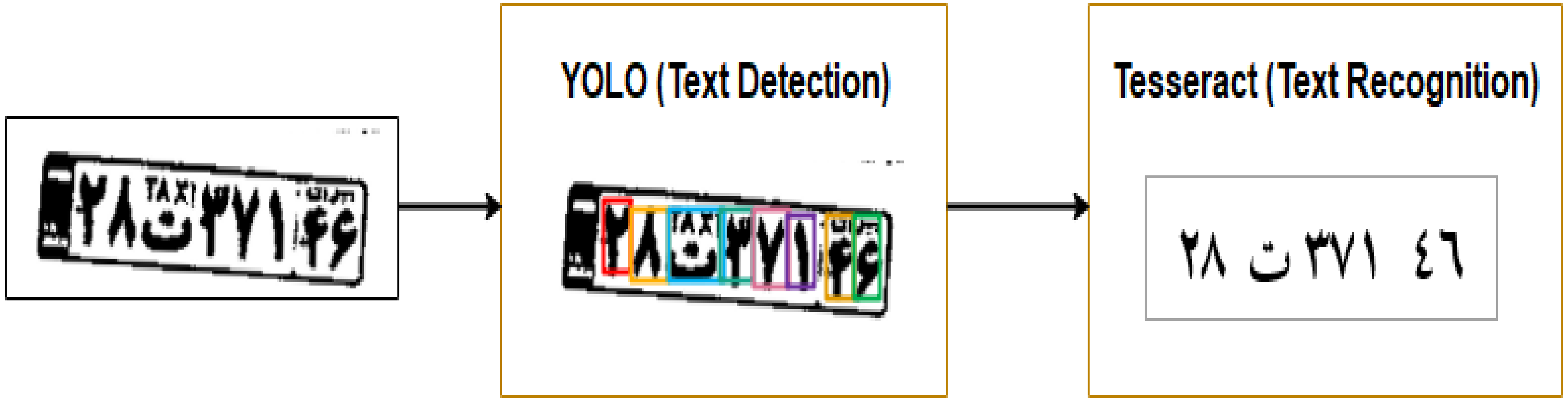
- License Plate Recognition (LPR)
  - ✓ Challenges
  - ✓ Traditional methods
  - ✓ Modern methods



Discussion & Result

➤ License Plate Recognition (LPR)

- ✓ Challenges
- ✓ Traditional methods
- ✓ Modern methods

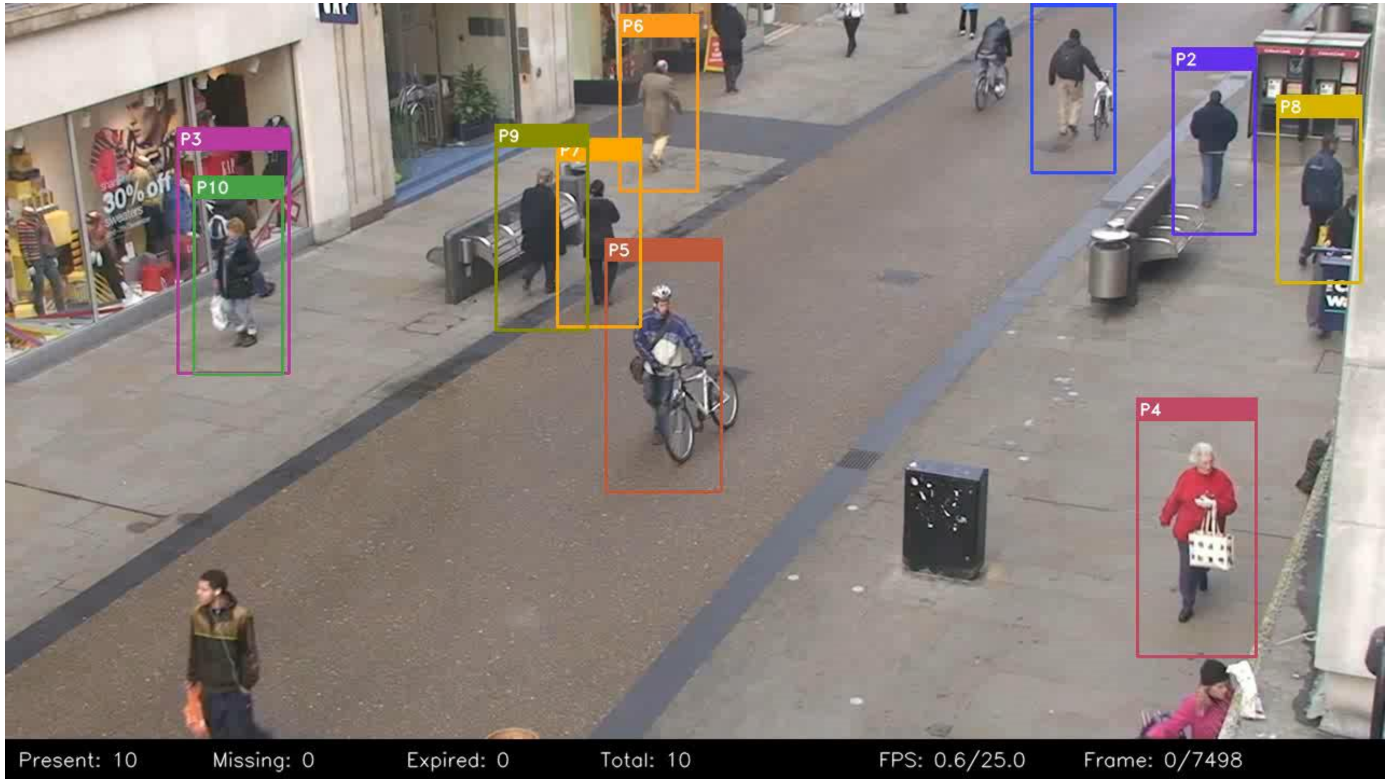


Discussion & Result

➢ License Plate Recognition (LPR)

- ✓ Challenges
- ✓ Traditional methods
- ✓ Modern methods

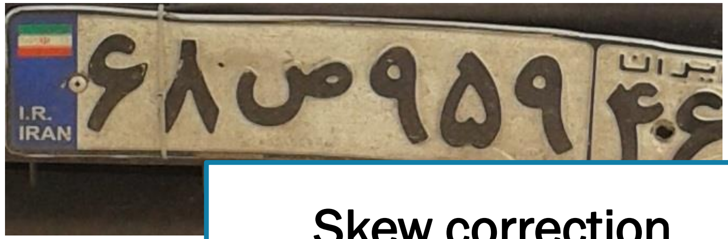
You Only Look Once (YOLO)



Discussion & Result

➢ License Plate Recognition (LPR)

- ✓ Challenges
- ✓ Traditional methods
- ✓ Modern methods
  - Pre-Processing
  - Augmentation
  - Annotation
  - Text Detection
  - Text Recognition

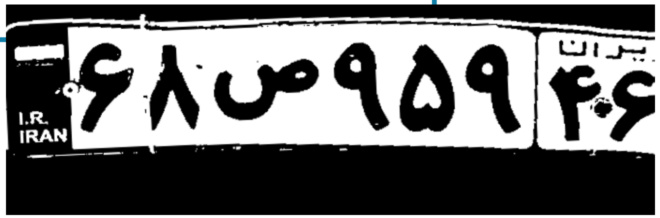


Skew correction

Shadow reduction

Noise reduction

Scale correction

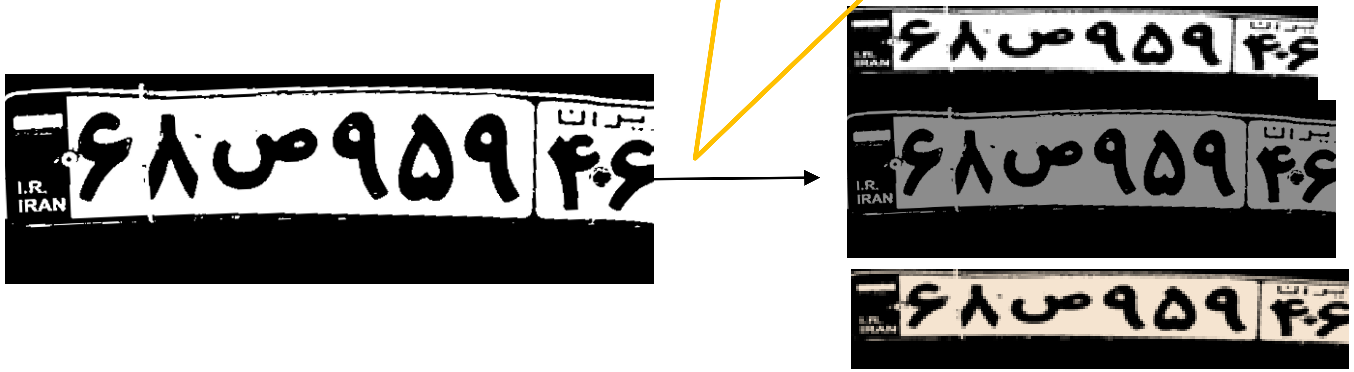


Discussion & Result

➢ License Plate Recognition (LPR)

- ✓ Challenges
- ✓ Traditional methods
- ✓ Modern methods
  - Pre-Processing
  - Augmentation
  - Annotation
  - Text Detection
  - Text Recognition

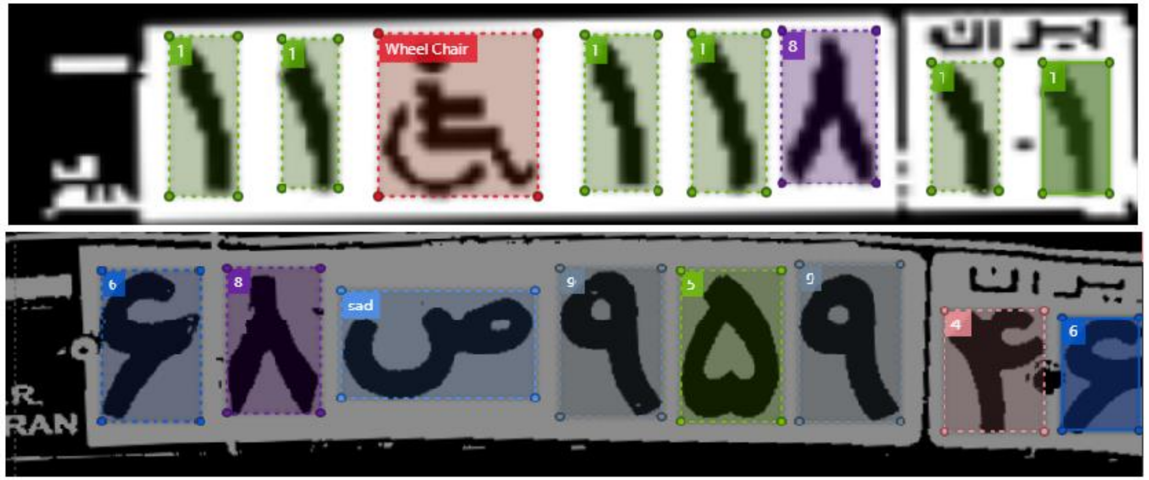
Size, Resolution, Rotation , etc



Discussion & Result

➢ License Plate Recognition (LPR)

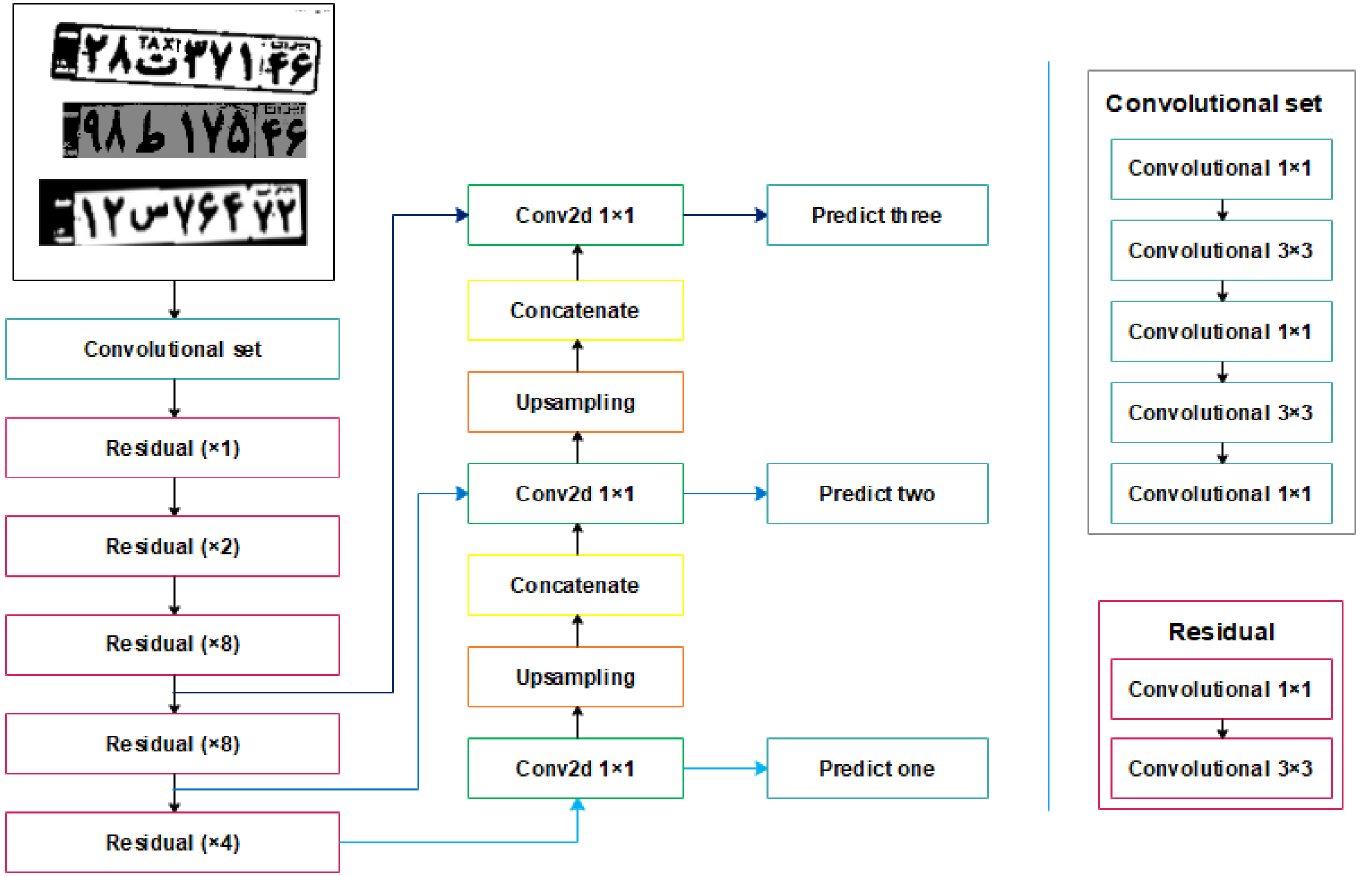
- ✓ Challenges
- ✓ Traditional methods
- ✓ Modern methods
  - Pre-Processing
  - Augmentation
  - Annotation
  - Text Detection
  - Text Recognition



Discussion & Result

➢ License Plate Recognition (LPR)

- ✓ Challenges
- ✓ Traditional methods
- ✓ Modern methods
  - Pre-Processing
  - Augmentation
  - Annotation
  - Text Detection
  - Text Recognition



Discussion & Result

- License Plate Recognition (LPR)

Segmented LP

Pre-processing

Feature extraction

Classification

Output

LP

Pre-processing

Augmentation

Annotation

Text Detection

Text Recognition

Output

Conclusions

- Neural network is an important issue in the field of computer vision.
- We used deep learning for getting better performance.
- An example of License Plate Detection (LPD) was used to understand all the concepts.

Questions

?

References

- H. Abdi, D. Valentin, B. Edelman, Neural Networks, Thousand Oaks, CA: SAGE Publication Inc., 1999.
- S. Haykin, Neural Networks, New York, NY: Nacmillan College Publishing Company, Inc., 1994.
- T. Masters, Practial Neural Network Recipes in C++, Toronto, ON: Academic Press, Inc., 1993.
- R. Schalkoff, Artificial Neural Networks, Toronto, ON: the McGraw-Hill Companies, Inc., 1997.
- S. M. Weiss and C. A. Kulikowski, Computer Systems That Learn, San Mateo, CA: Morgan Kaufmann Publishers, Inc., 1991.
- P. D. Wasserman, Neural Computing: Theory and Practice, New York, NY: Van Nostrand Reinhold, 1989.

References

- Learning to recommend from positive evidence. In: H. Lieberman (ed.) Proceedings of 2000 International Conference on Intelligent User Interfaces, New Orleans, LA, January 9-12, 2000, ACM Press, pp. 241-247 .
- J. Wang, X. Liu, A. Liu, J. Xiao., "A deep learning-based method for vehicle license plate recognition in natural scene," APSIPA Transactions on Signal and Information Processing, vol. 8, 2019.
- I. J. Goodfellow, Y. Bulatov, J. Ibarz, S. Arnoud, and V. Shet., "Multi-digit number recognition from street view imagery using deep convolutional neural networks," arXiv preprint arXiv:1312.6082, 2013.
- J. Redmon, S. Divvala, R. Girshick, A. Farhadi., "You only look once: Unified, real-time object detection.," Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 779-788, 2016.

References

- L. Jun, J. Su, and X. Zeng, "A solution method for image distortion correction model based on bilinear interpolation," ***Компьютерная оптика***, vol. 43, no. 1, 2019.
- D. Bradley and G. Roth, "Adaptive thresholding using the integral image," Journal of Graphics Tools, vol. 12, no. 2, pp. 13–21, 2007.
- J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," IEEE conference on computer vision and pattern recognition, pp. 779–788, 2016.
- J. Redmon, and A. Farhadi, "YOLOv3: An incremental improvement," 2018, arXiv: 1804.02767. [Online]. Available: https://arxiv.org/abs/1804.02767.