



EGYPT

IndabaX

3RD - 4TH MAY, 2019

The American University in Cairo



THE AMERICAN
UNIVERSITY IN CAIRO



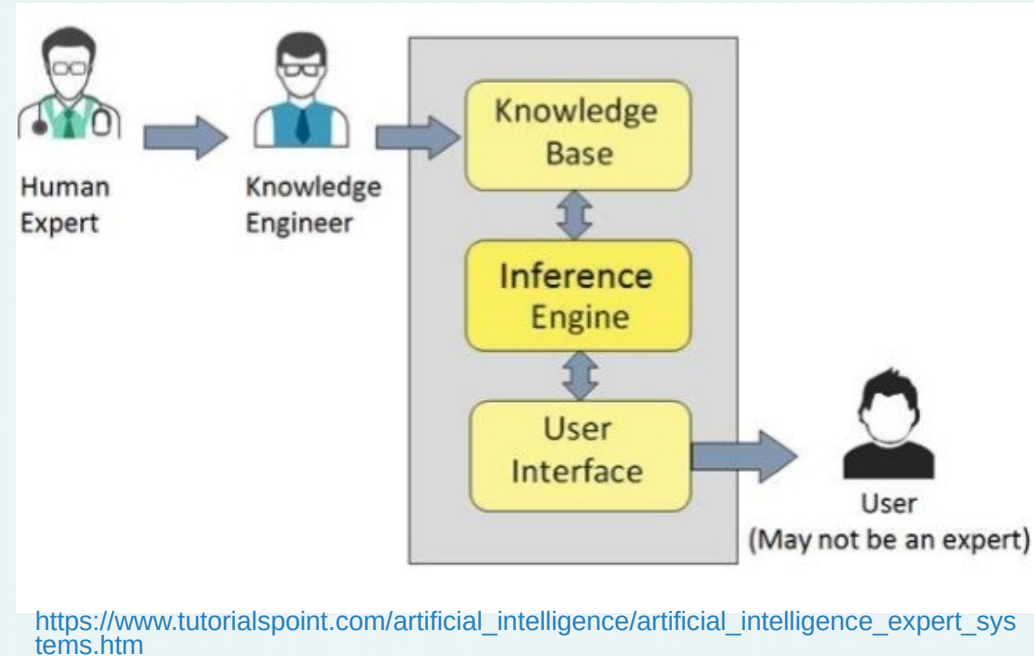
The Landscape of AI

Mohamed Moustafa, PhD
m.moustafa@aucegypt.edu
moustafa@ieee.org



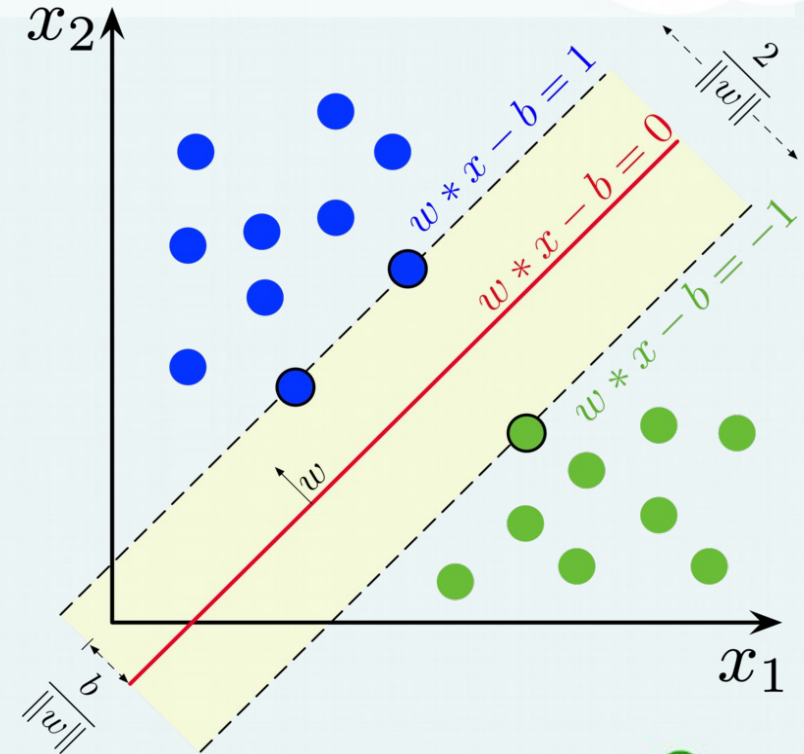
AI = Expert Systems

- 1980s
- Rule-based systems
- Simple implementation
- Mostly toward natural language processing
- Knowledge acquisition issue



AI = Statistical ML

- 1990s-2000s
- Increased computation power
- Availability of digital images
- Extract statistical features from the training data
- Vision, speech, bioinformatics
- Still needs an expert scientist to design features and classifier combination

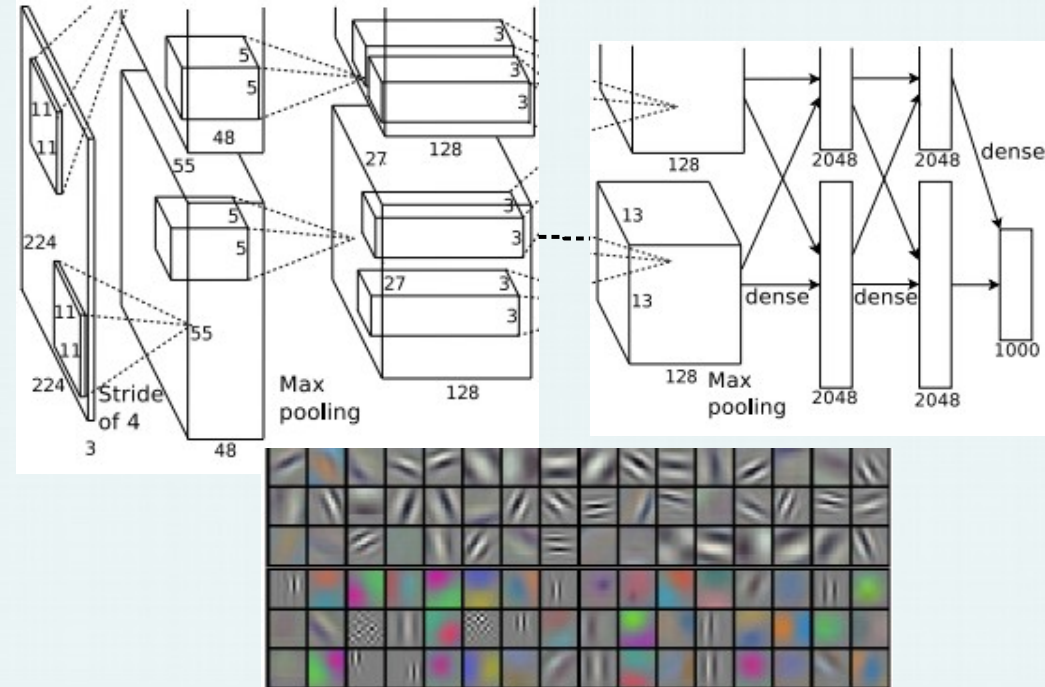


https://en.wikipedia.org/wiki/Support-vector_machine#/media/File:SVM_margin.png



AI = Deep Learning

- 2010s
- Parallel computing power
- Availability of data
- Try to mimic biological layered neural networks
- Feature extractor + Decision maker
- Overfitting and computation time issues



https://www.cs.toronto.edu/~kriz/imagenet_classification_with_deep_convolutional.pdf



Image Classification – CNN



motor scooter

leopard

motor scooter	leopard
go-kart	jaguar
moped	cheetah
bumper car	snow leopard
golfcart	Egyptian cat

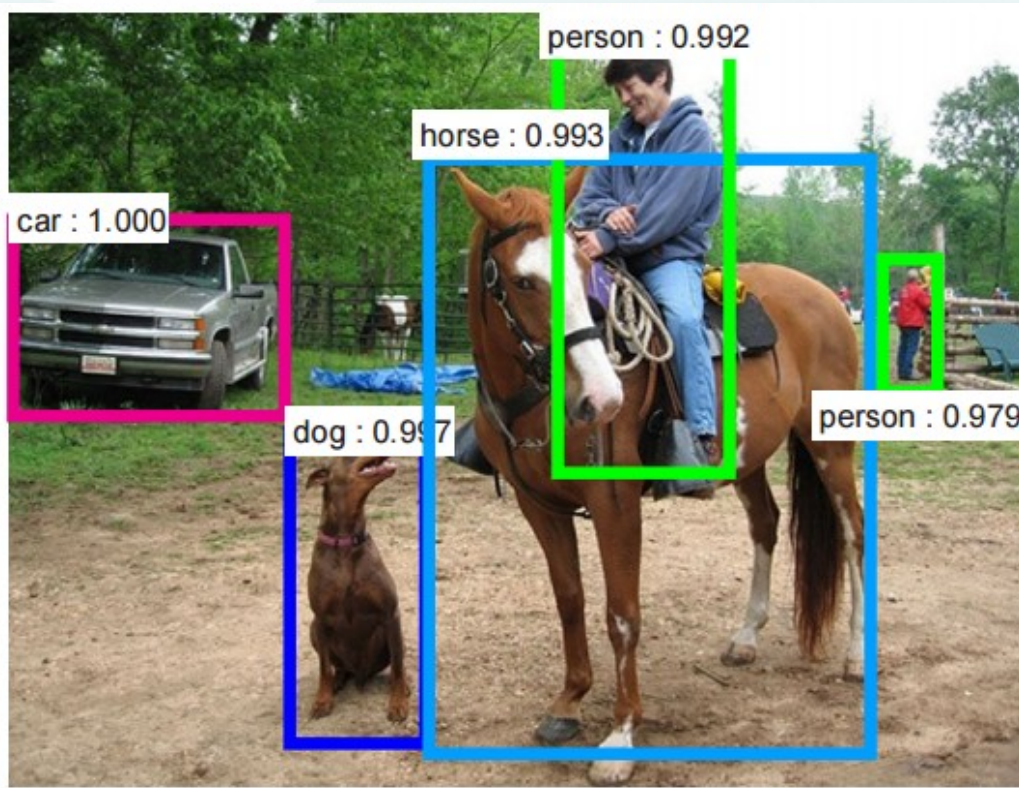
https://www.cs.toronto.edu/~kriz/imagenet_classification_with_deep_convolutional.pdf

```
[(0.17842512, 'n03478589 half track'),
 (0.17090011,
  'n04389033 tank, army tank, armored combat vehicle, armoured combat vehicle'),
 (0.11375312, 'n03445924 golfcart, golf cart'),
 (0.057274561, 'n03255030 dumbbell'),
 (0.03960849, 'n04461696 tow truck, tow car, wrecker')]
```



<http://english.ahram.org.eg/Media/News/2012/10/24/2012-634867103303985336-398.jpg>

Object Detection and Segmentation – CNN



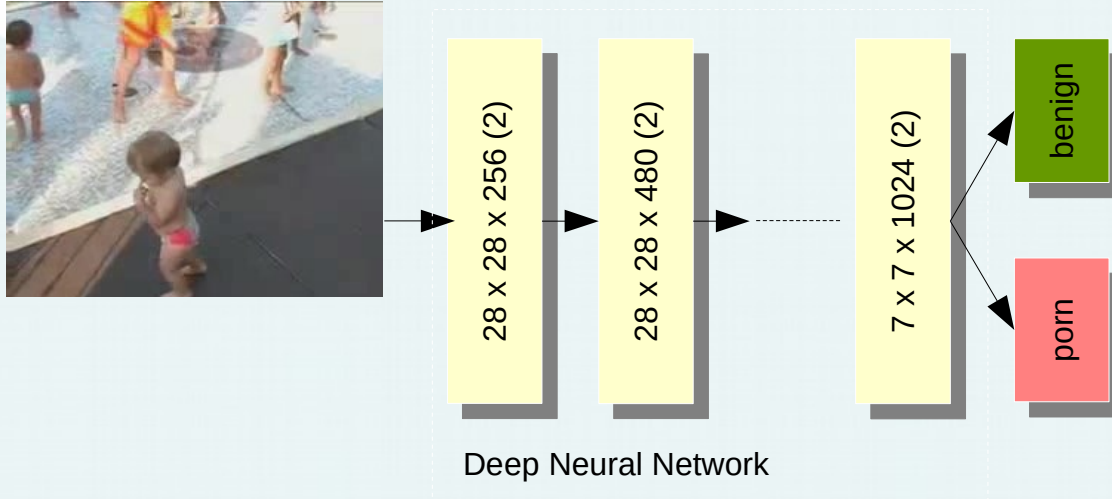
(from Shaoqing Ren, Kaiming He, Ross Girshick, Jian Sun, Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks, arXiv:1506.01497.)



(from Jifeng Dai, Kaiming He, Jian Sun, BoxSup: Exploiting Bounding Boxes to Supervise Convolutional Networks for Semantic Segmentation, arXiv:1503.01640.)



Image and Video Filtering – CNN



What is new?
Not a mere skin detector

Approach	Accuracy %
HueSIFT [2]	89.5%
BinBoost [3]	90.9%
Our DeepNN	94.1%

500 Hours of Video Uploaded To YouTube Every Minute

34% of *unwanted* exposure to porn. (*pop up ads, misdirected links*)

20% of men admit to watching *at work*.

Egypt ranks 2nd worldwide in search for 'porn' (2015).

https://www.researchgate.net/publication/284727434_Applying_deep_learning_to_classify_pornographic_images_and_videos

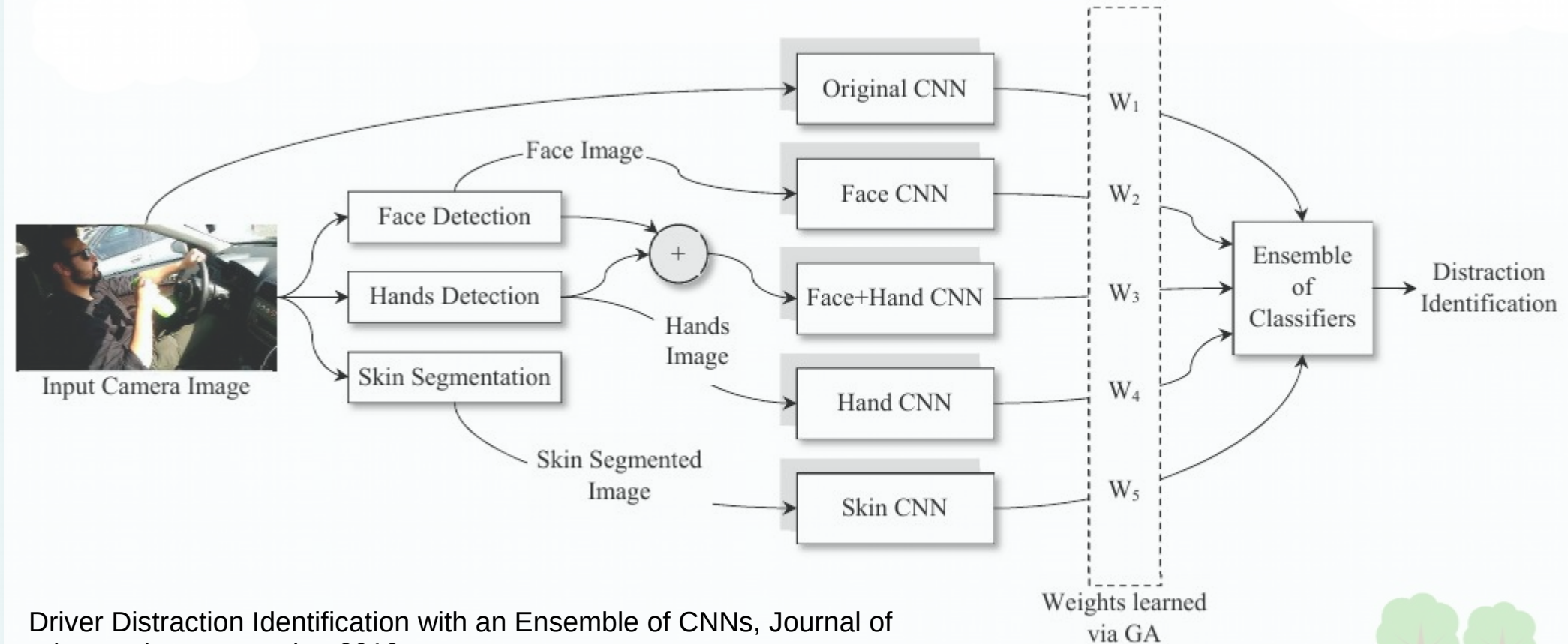
<http://arxiv.org/abs/1511.08899>



**7th Pacific Rim Symposium on
Image and Video Technology**

23-27 November, 2015, Auckland, New Zealand

Ensemble of CNNs



Driver Distraction Identification with an Ensemble of CNNs, Journal of advanced transportation 2019

Saliency Maps



Talking to Passenger



Phone Left



Text Right



Reaching Behind



Phone Right



Adjusting Radio



Hair or Makeup



Drinking



Text Left



Safe Driving

Saliency maps demonstrate that our network makes its decision based on relevant features/regions in the input images.

Driver Distraction Identification with an Ensemble of CNNs, Journal of advanced transportation 2019

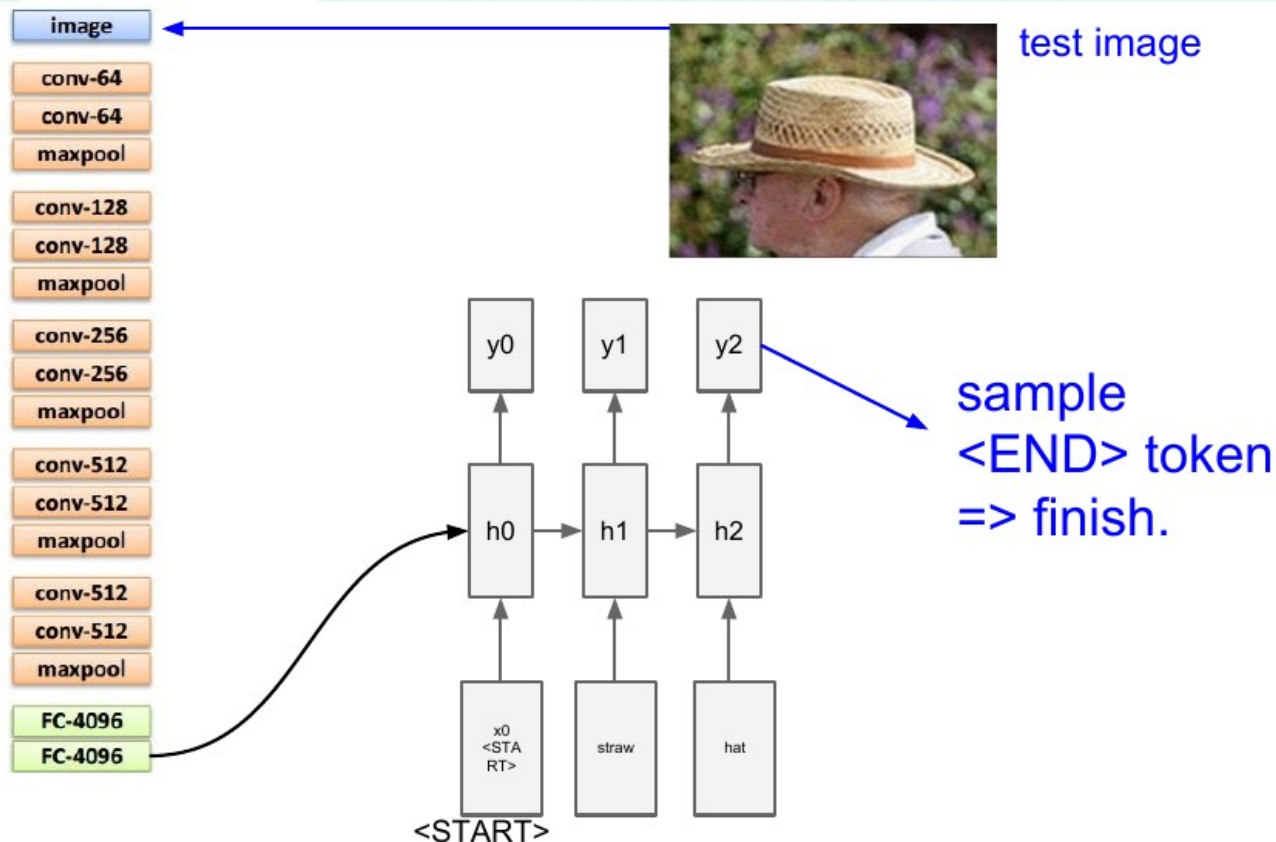


Spatio-Temporal Modalities

- Do not treat a video sequence like an unordered set of images
- There is information in the **sequence** itself, and recurrent nets use it to perform tasks that feedforward networks cannot
- The rise of language processing and video analytics



Image Caption : CNN+LSTM

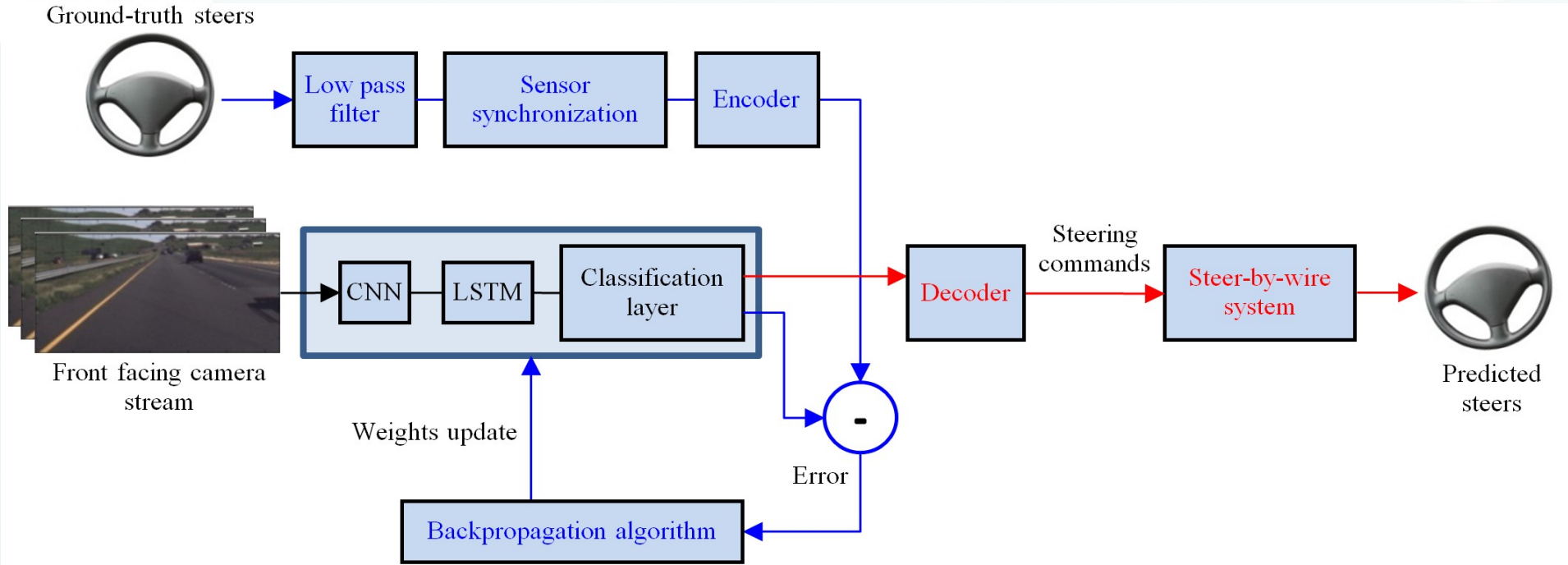


a man riding a bike on a beach
with a dog in the water

<https://github.com/karpathy/neuraltalk2>



Auto-Steering : CNN+LSTM



Eraqi, Moustafa, and Honer, "End-to-End Deep Learning for Steering Autonomous Vehicles Considering Temporal Dependencies," (NIPS 2017) <https://arxiv.org/abs/1710.03804>

Solid Conclusions ... so far

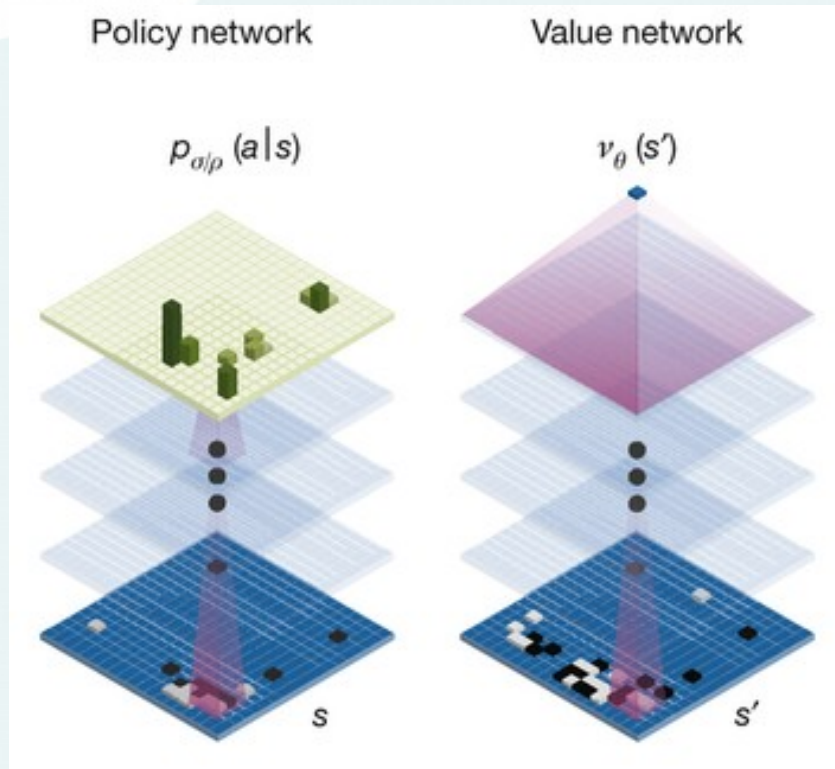
- Transfer learning is a must in case of data scarcity
- Ensemble of Networks is always better than individuals
- CNN+LSTM is a popular choice in video analytics



Deep Reinforcement Learning



AlphaGo : CNN+RL

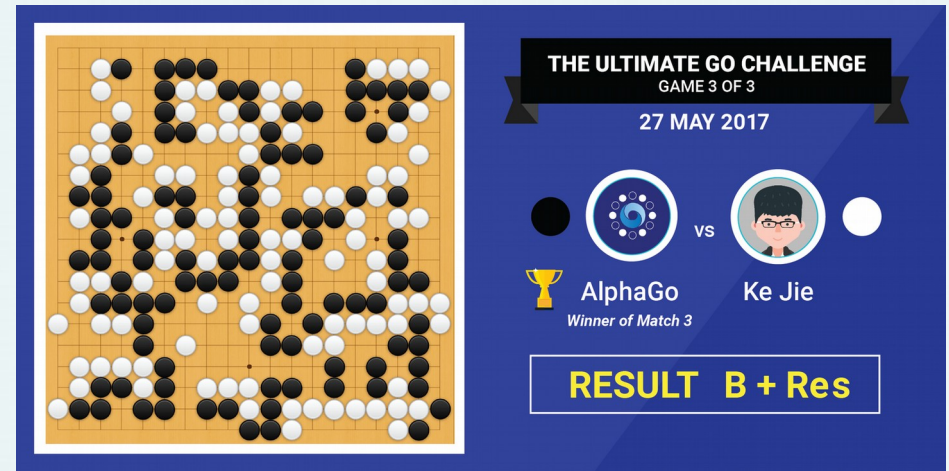


- Value network provides an estimate of: what is the probability of the black player to win the game, given the current state?
- Policy network provides a probability value for each possible legal move.

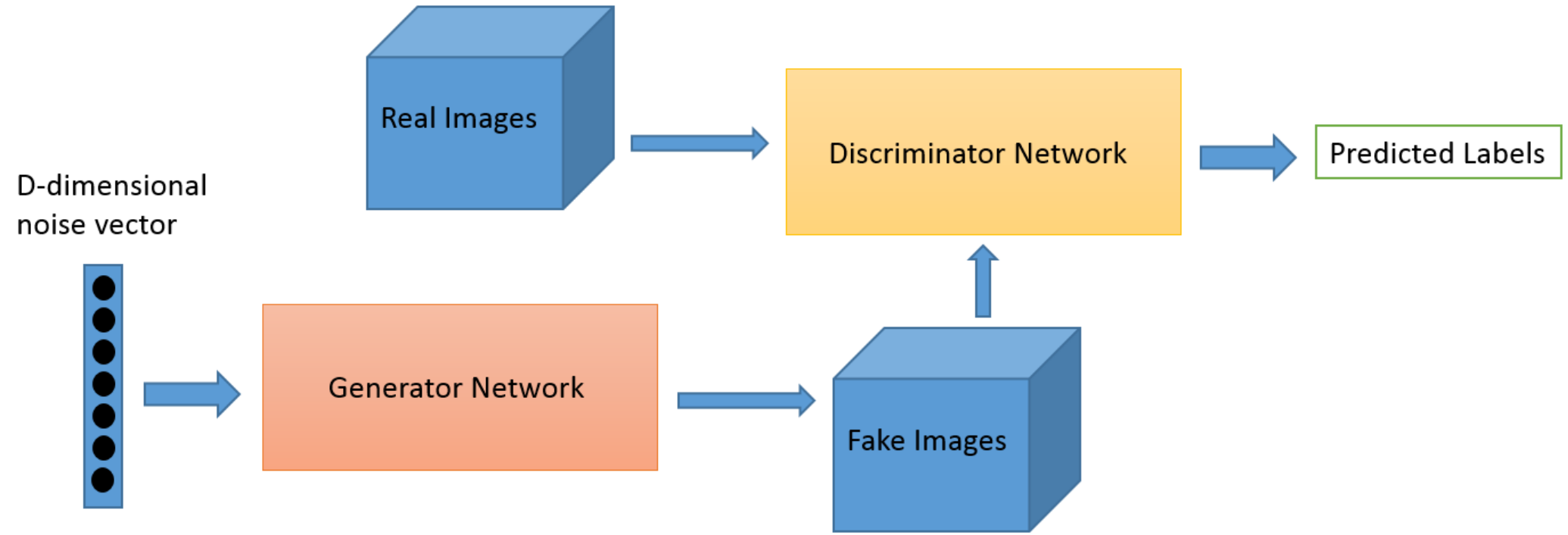


AlphaGo Zero: Single Network

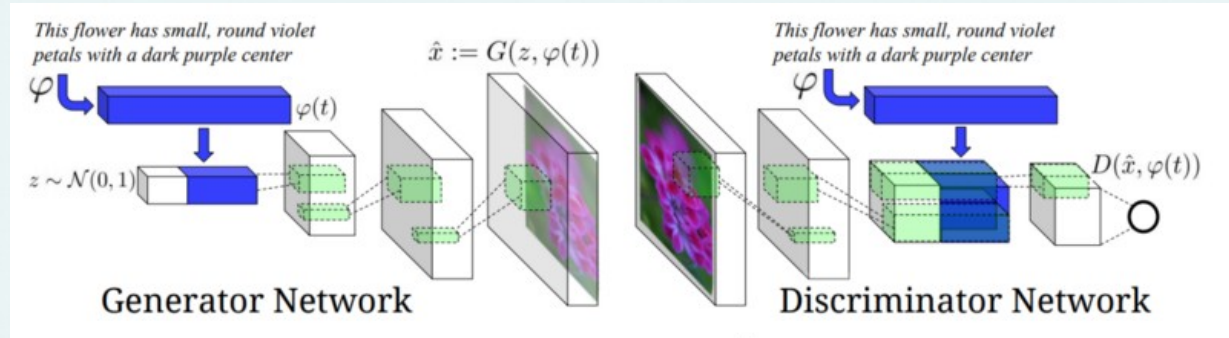
- Trained solely by self-play reinforcement learning, starting from random play
- Defeated Ke Jie (world number 1), May 2017



Generative Adversarial Networks



GAN applications



Mohsen and Moustafa, "Generating Large Scale Images Using GANs", The 11th International Conference on Digital Image Processing, icdip.org, May 2019

this white and yellow flower have thin white petals and a round yellow stamen



Reed et al., "Generative Adversarial Text to Image Synthesis", <https://arxiv.org/pdf/1605.05396.pdf>



Trending ...

- Weak supervising methods
- Edge computing
- Code of Ethics

