

NEURAL NETWORKS:

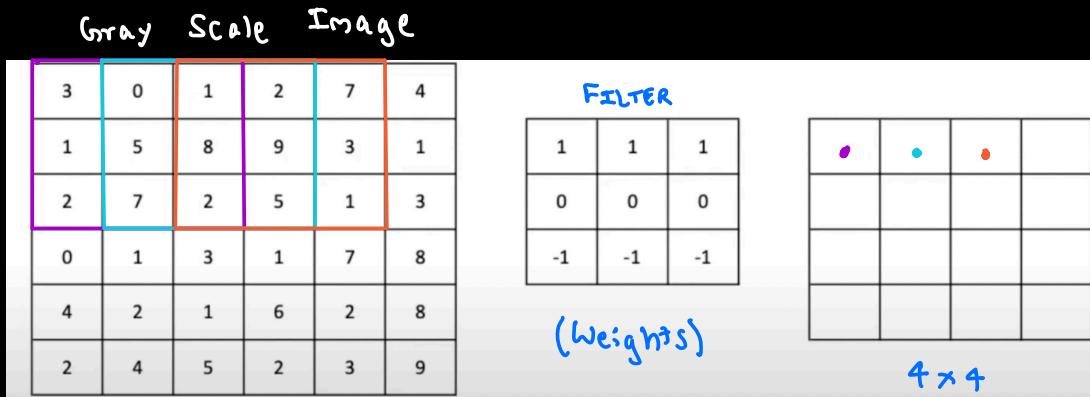
- Computer vision (Image Analysis) CNN
- Sequence-to-sequence translation RNN
(Speech, Text)
- Perspective on "model-based" vs "model-free" approaches to AI.

FUNDAMENTAL CONCEPTS:

1. Local Detection of features / Patterns
2. Aggregation / Abstraction

CONVOLUTION:

b

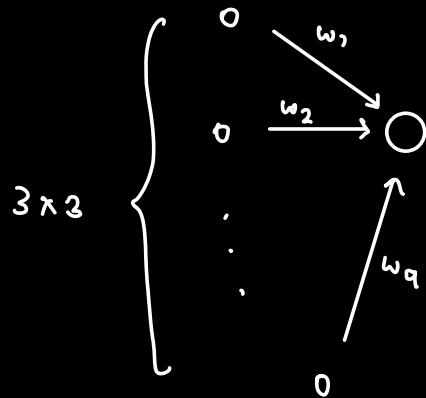


6x6

$$3 \times 1 + 0 \times 1 + 1 \times 1 +$$

$$1 \times 0 + 5 \times 0 + 8 \times 0 +$$

$$2 \times -1 + 7 \times -1 + 2 \times -1 = 0$$



HYPERPARAMETERS:

• Filter Size $f = 3$

• Padding $p = 1$

• Stride $s = 1$

| | | | | | |
|---|---|---|---|---|---|
| 3 | 0 | 1 | 2 | 7 | 4 |
| 1 | 5 | 8 | 9 | 3 | 1 |
| 2 | 7 | 2 | 5 | 1 | 3 |
| 0 | 1 | 3 | 1 | 7 | 8 |
| 4 | 2 | 1 | 6 | 2 | 8 |
| 2 | 4 | 5 | 2 | 3 | 9 |

8×8

Padding $p = 1$

* FILTER

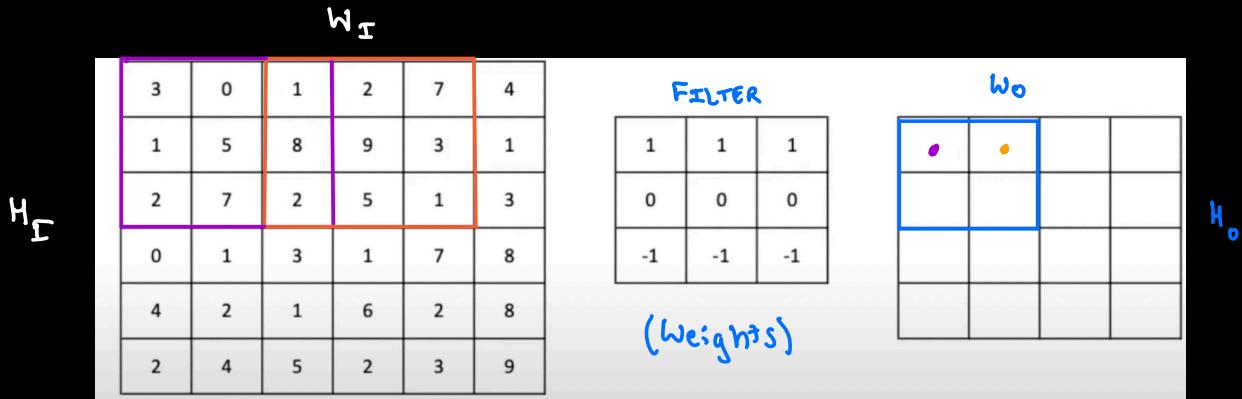
| | | |
|----|----|----|
| 1 | 1 | 1 |
| 0 | 0 | 0 |
| -1 | -1 | -1 |

| | | | |
|---|---|---|--|
| * | . | . | |
| | | | |
| | | | |
| | | | |
| | | | |

4×4

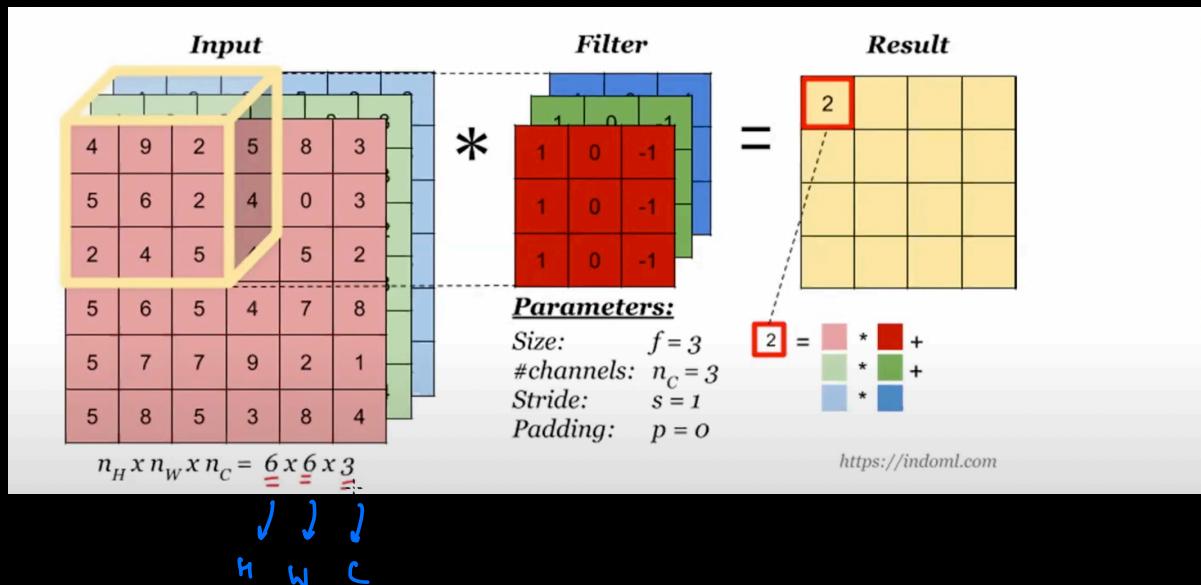
6×6

Same Convolution



Stride = 2

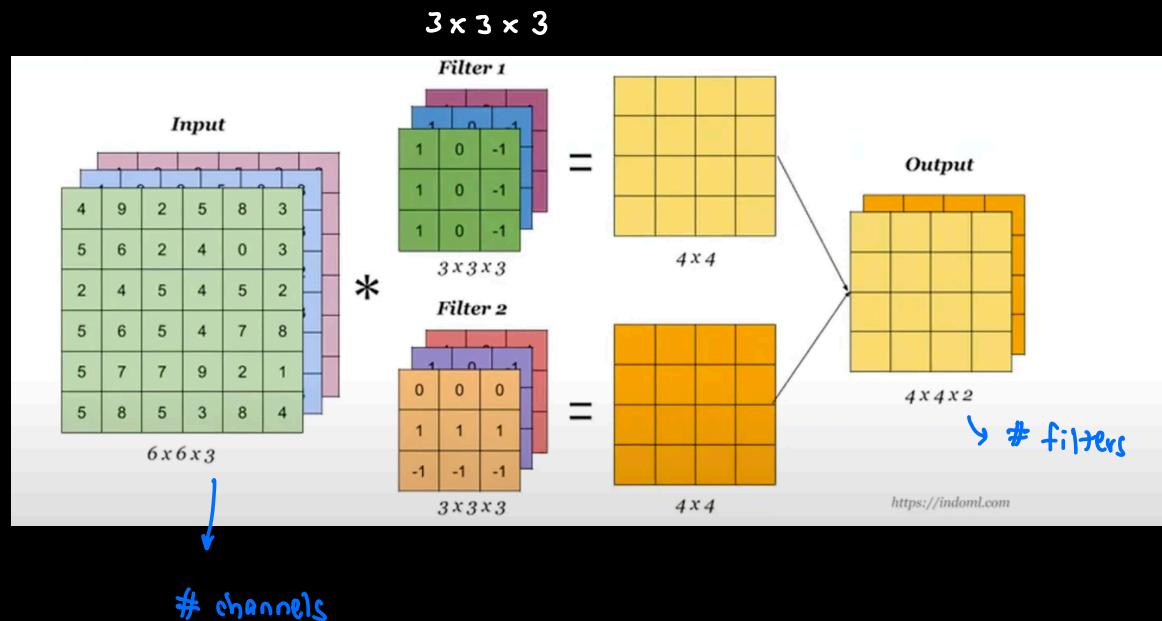
$3 \times 3 \times 3$



Channels of Input, Filter must

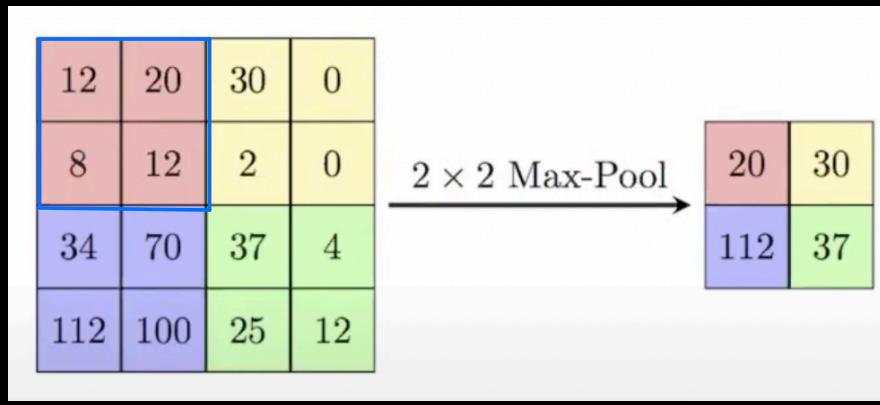
be same.

MULTIPLE FILTERS:



MAX POOLING:

AGGREGATION



4×4

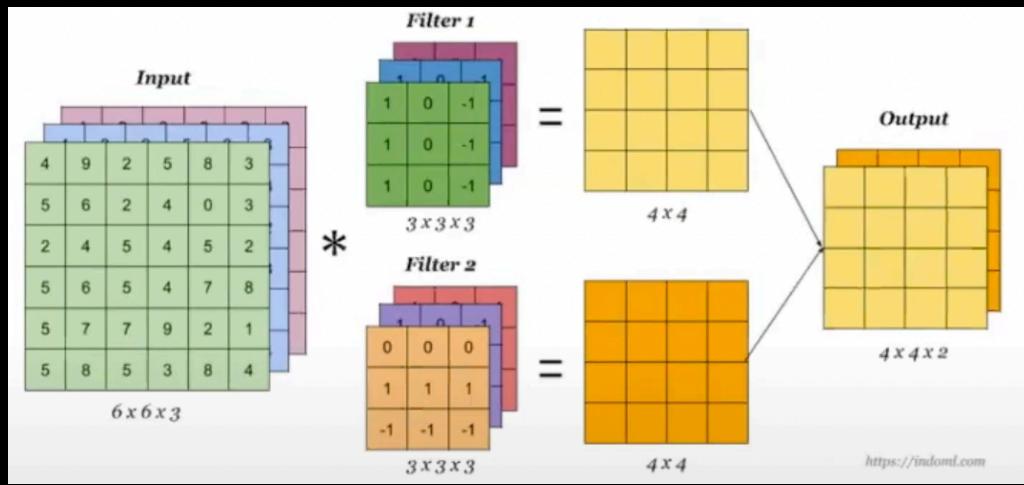
2×2

$S = 2$

$P = 0$

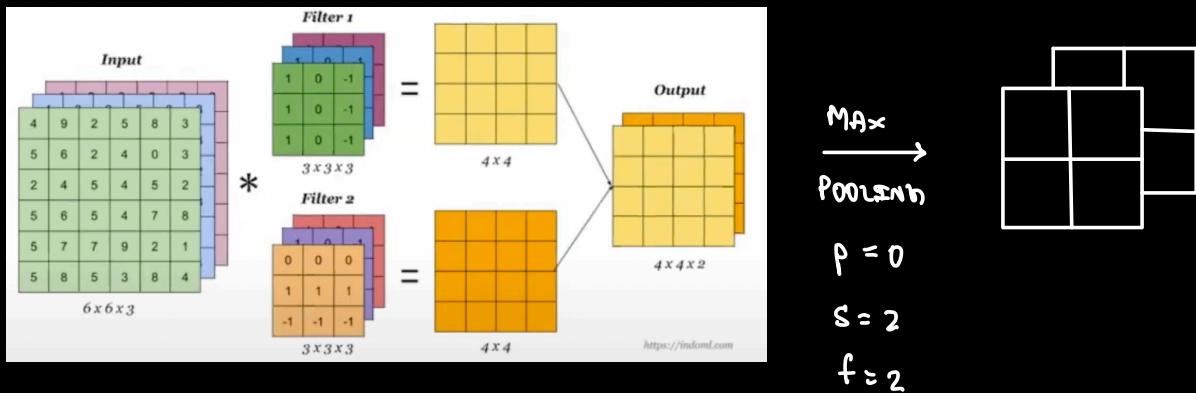
$f = 2$

* So filters have no weight to be learnt. Just take max of elements.

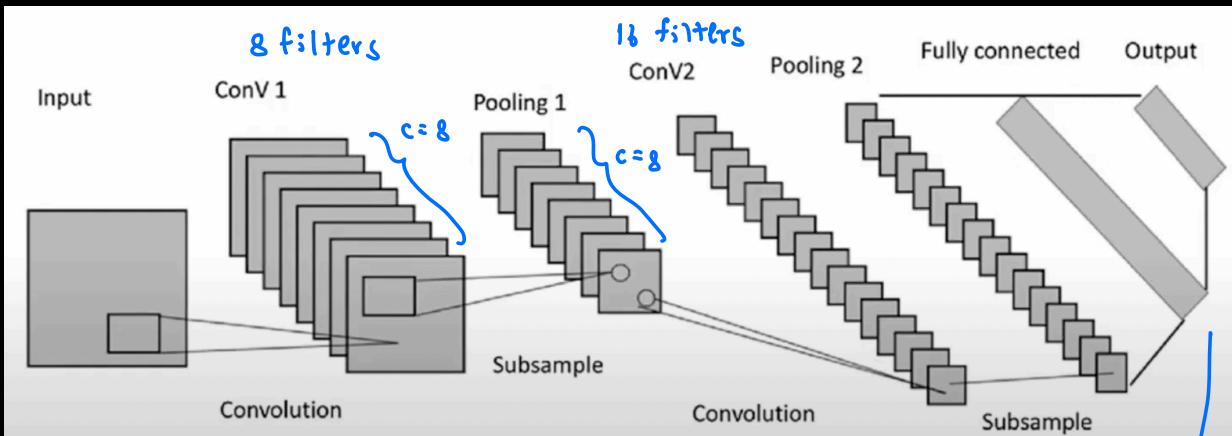


So , convolution collapsed the number of channels.

Max-pooling maintains it → applied to channels independently



CNN ARCHITECTURE:



number of channels increases

height , width decreases.

Most of
the weights
are here.

Tensors: Multi-directional arrays that can execute
parallelly.

PROMISE AND LIMITATIONS OF NN:

- (1) Hungry for data
- (2) Brittle / Lack of robustness
- (3) Not as easy to explain

} Explainable AI.

↓ Why?

Like Loan Rejected - But why?



Model-free

Function fitting

Curve fitting

Model-based - Used knowledge / domain

BRITTLE / LACK OF ROBUSTNESS:

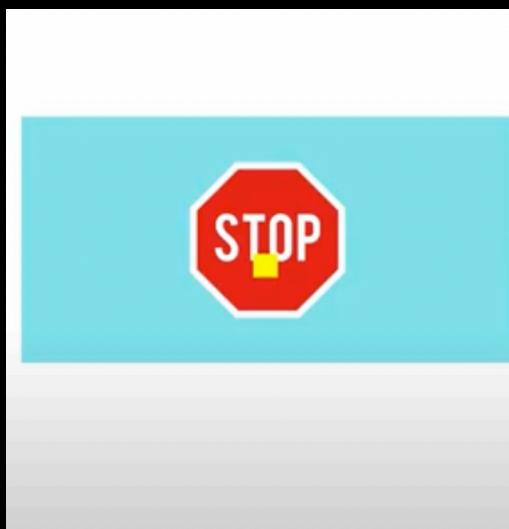


Image Transformations that do not change the image

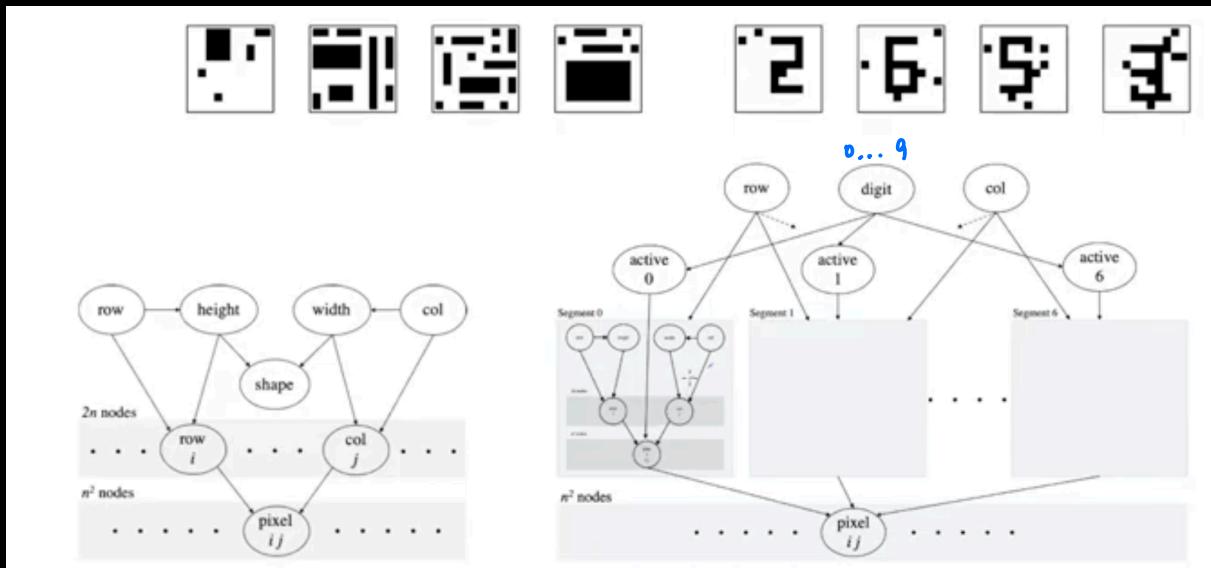
much - Affects image recognition.

→ Adversarial Attacks.

MODEL-BASED SUPERVISED LEARNING:

Recognize the largest box

7 line with noise



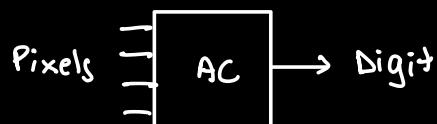
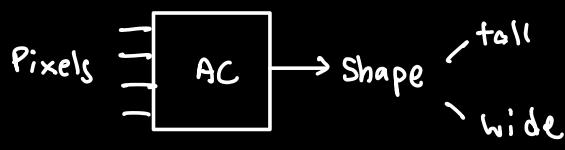
7 Line :



on some of

it to generate

numbers.



CHALLENGES:

- * Compiling into circuit - computationally intensive.
- * In some scenarios, hard to build the model.

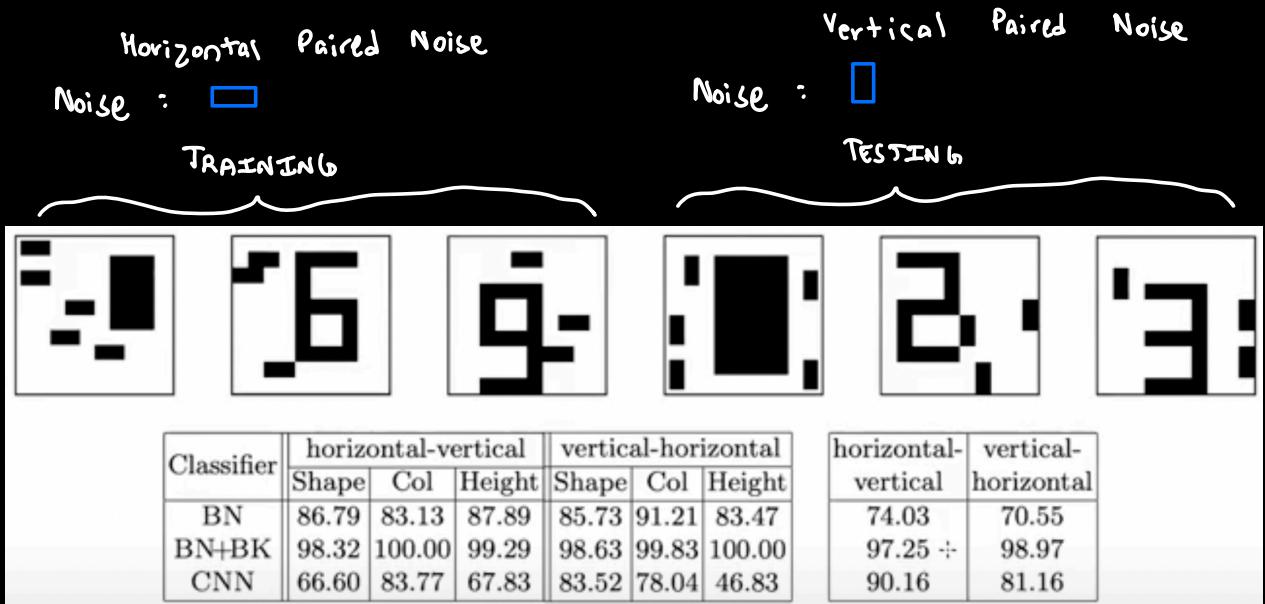


Table 2. Accuracy of detecting rectangle properties while varying training data size.

| Data | BN Classifier | | | BN+BK Classifier | | | CNN Classifier | | |
|------|---------------|-------|--------|------------------|-------|--------|----------------|-------|--------|
| | Shape | Col | Height | Shape | Col | Height | Shape | Col | Height |
| 25 | 50.64 | 29.53 | 20.35 | 96.71 | 87.23 | 81.88 | 72.02 | 32.26 | 20.72 |
| 50 | 53.17 | 34.77 | 19.95 | 93.12 | 91.88 | 86.79 | 81.73 | 49.74 | 24.71 |
| 100 | 56.32 | 49.67 | 24.08 | 94.44 | 97.37 | 93.04 | 83.73 | 75.21 | 31.27 |
| 250 | 67.54 | 63.51 | 26.49 | 98.97 | 98.42 | 97.26 | 86.13 | 84.74 | 39.09 |
| 500 | 77.92 | 73.02 | 31.45 | 99.82 | 98.61 | 97.38 | 90.05 | 91.67 | 61.54 |
| 1000 | 81.20 | 91.93 | 69.20 | 99.98 | 98.86 | 97.56 | 97.12 | 96.56 | 84.54 |
| 2000 | 83.40 | 98.11 | 91.63 | 99.99 | 98.90 | 97.66 | 98.37 | 98.77 | 92.52 |
| 4000 | 88.99 | 98.98 | 98.61 | 100.00 | 98.95 | 97.85 | 99.18 | 99.39 | 96.35 |
| 8000 | 95.44 | 99.79 | 99.15 | 99.99 | 99.18 | 98.01 | 99.64 | 99.77 | 97.51 |

Table 3. Accuracy of classifying digits while varying the size of training data.

| Training Data | 25 | 50 | 100 | 250 | 500 | 1000 | 2000 | 4000 | 8000 |
|---------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| BN | 12.68 | 11.66 | 17.42 | 23.57 | 32.19 | 42.49 | 56.35 | 84.33 | 91.49 |
| BN+BK | 72.07 | 96.39 | 98.52 | 98.56 | 98.64 | 98.83 | 99.10 | 99.07 | 99.14 |
| CNN | 30.34 | 29.22 | 43.45 | 78.79 | 95.13 | 97.29 | 97.97 | 98.48 | 98.48 |



So MODEL-BASED learns the concept!

MODEL-FREE only fits the data.

Note: BK : Background Knowledge.