Introduction to Machine Learning

Convolution Neural Networks

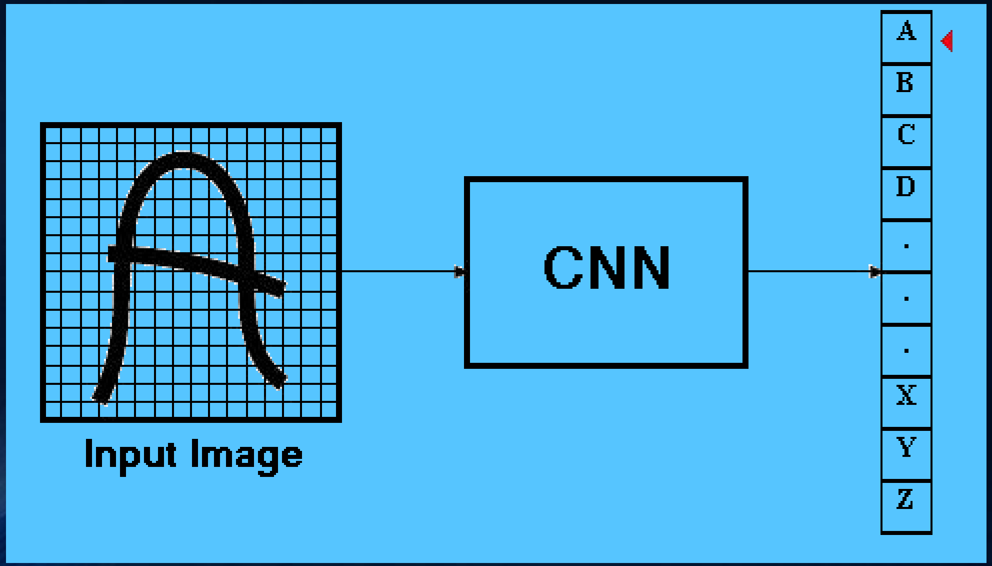
Dr Raymond Lee
Associate Professor
Division of Science and Technology
BNU-HKBU United International College

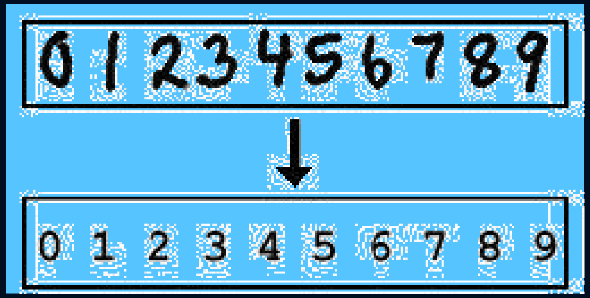
RoadMap

- Introduction
- Drawbacks of previous neural networks
- Convolutional neural networks
- LeNet 5
- Comparison
- Disadvantage
- Application

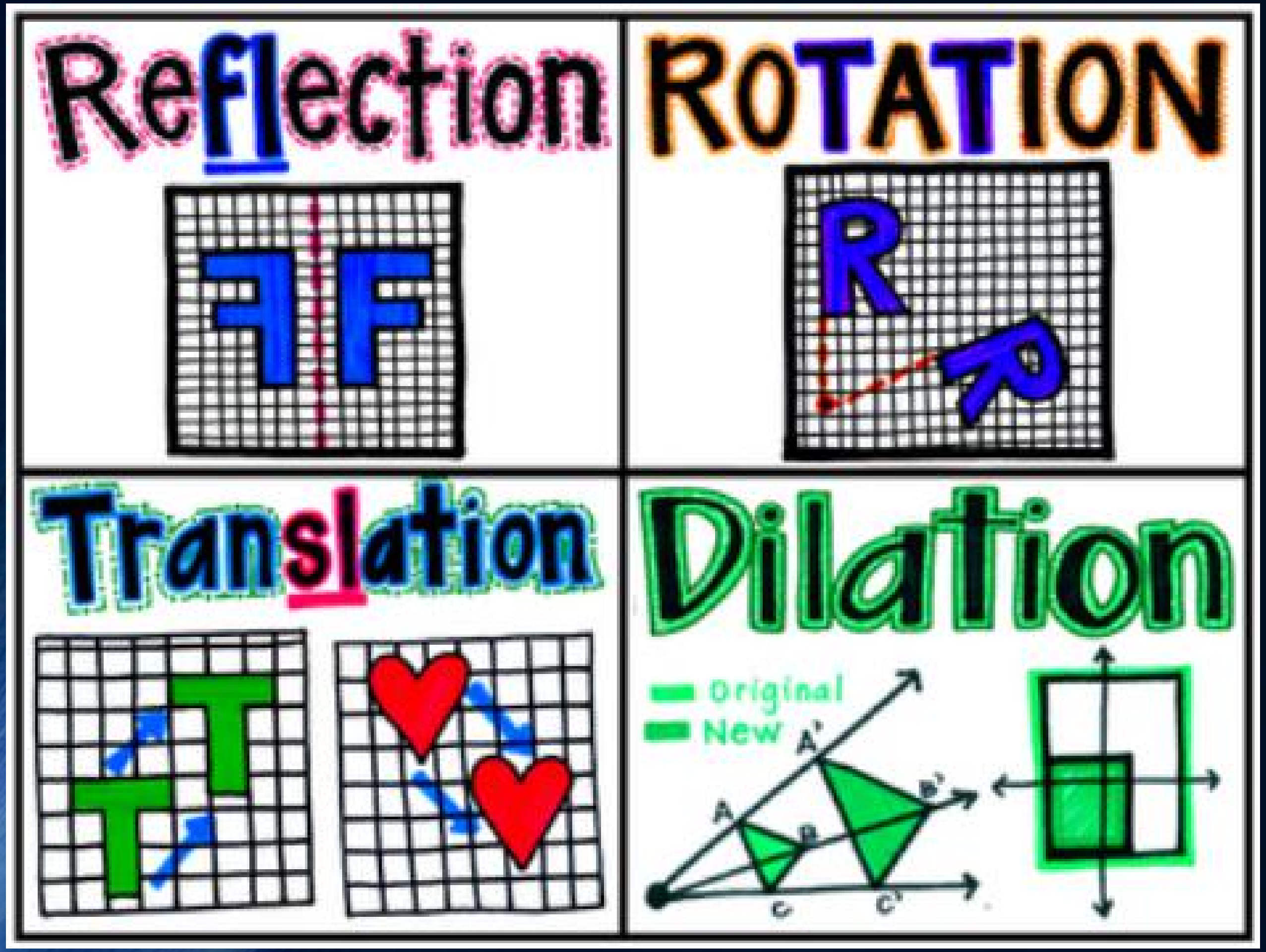
Introduction

- **Convolutional neural networks**
  - Signal processing, Image processing
- **improvement over the multilayer perceptron**
  - performance, accuracy and some degree of invariance to distortions in the input images





5 Different Types of Invariant Properties

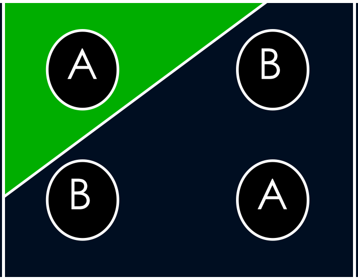
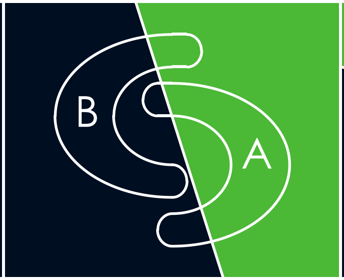
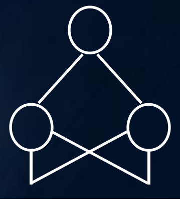
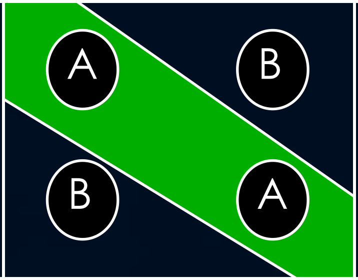
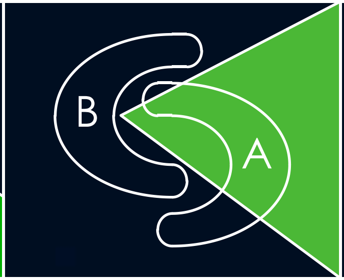
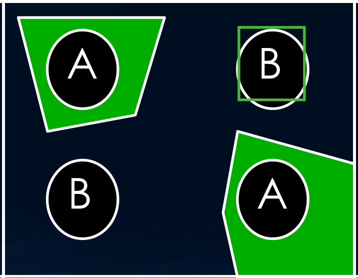
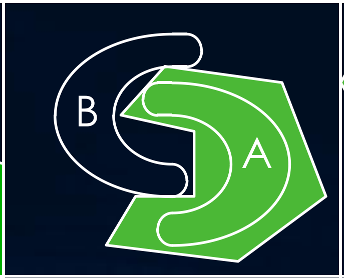
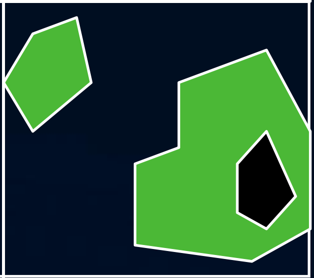


Distortion

RoadMap

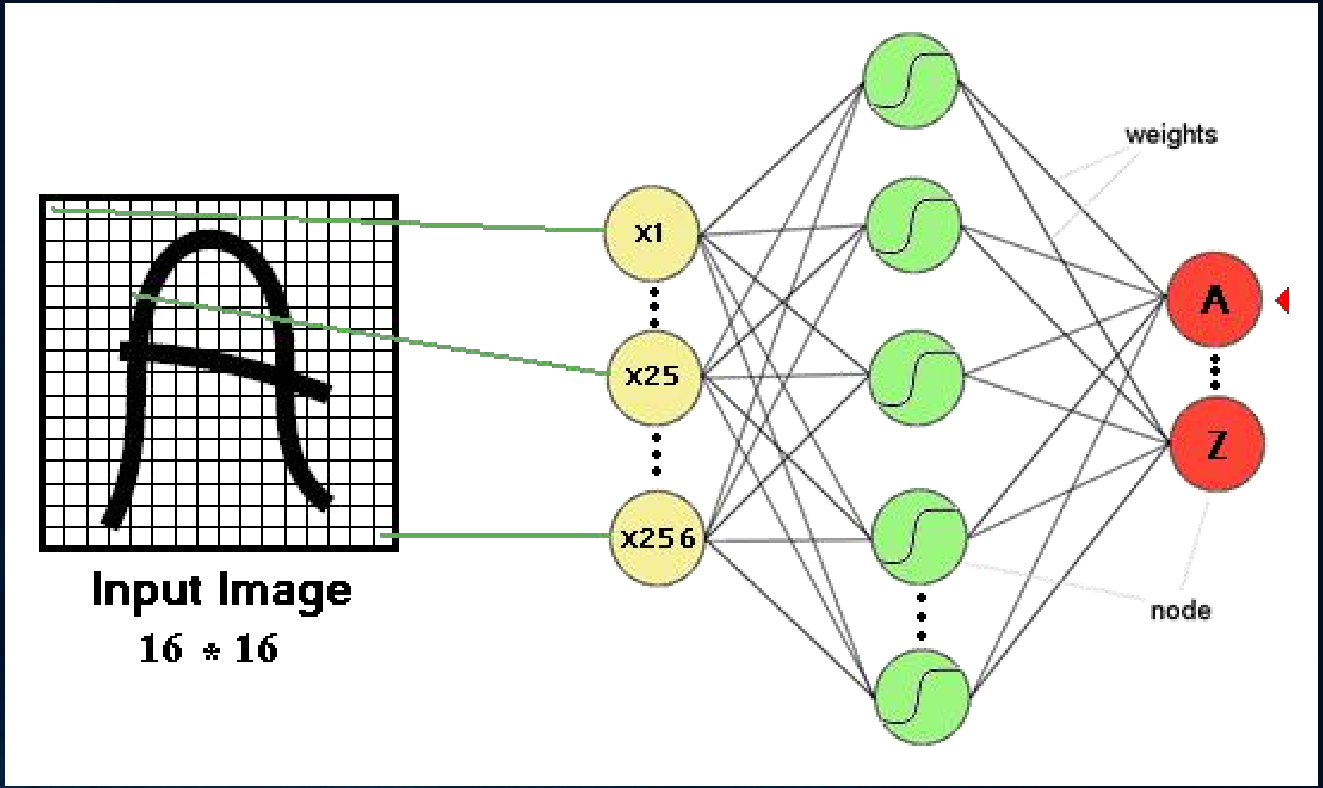
- Introduction
- Drawbacks of previous neural networks
- Convolutional neural networks
- LeNet 5
- Comparison
- Disadvantage
- Application

Behavior of multilayer neural networks

Structure	*Types of Decision Regions*	*Exclusive-OR Problem*	*Classes with Meshed regions*	*Most General Region Shapes*
Single-Layer 	*Half Plane Bounded By Hyper plane*			
Two-Layer 	*Convex Open Or Closed Regions*			
Three-Layer 	Arbitrary (Complexity Limited by No. of Nodes)			

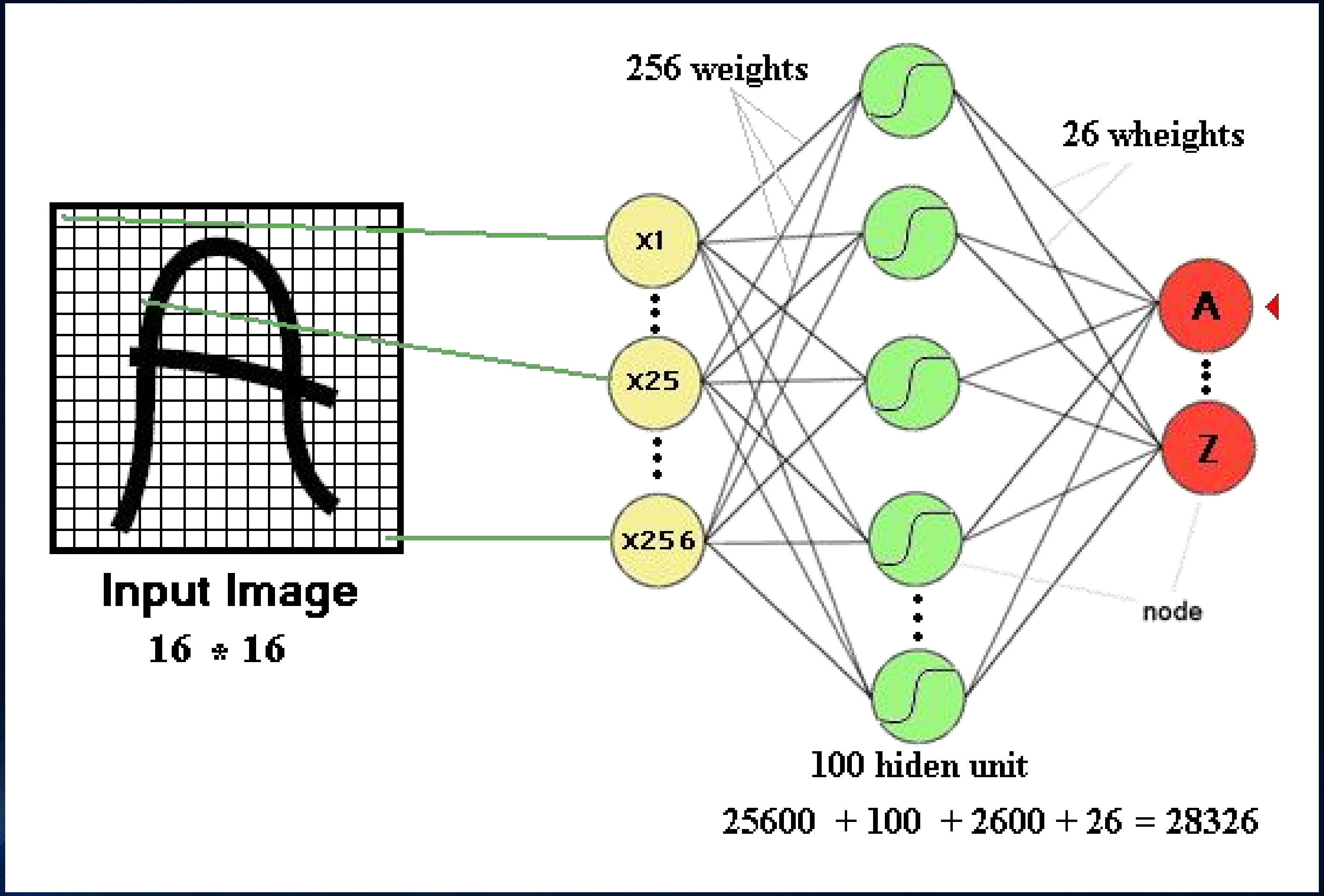
Multi-layer perceptron and image processing

- One or more hidden layers
- Sigmoid activations functions



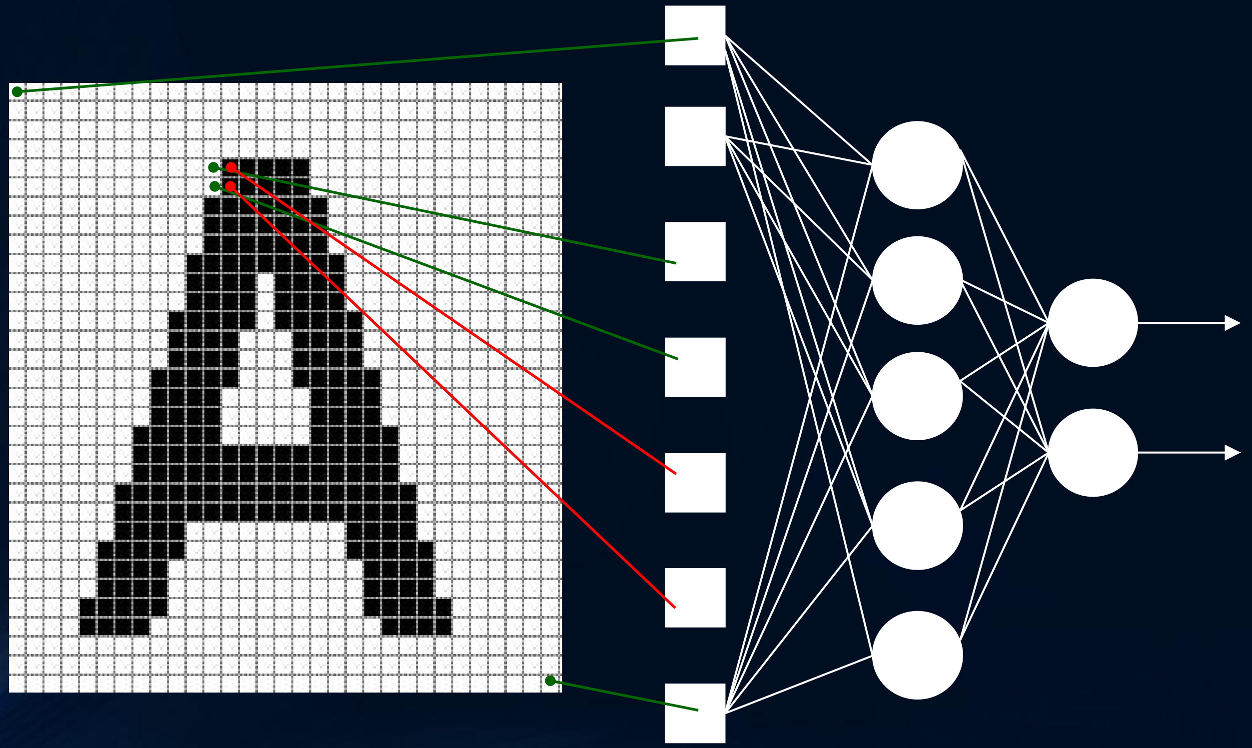
Drawbacks of previous neural networks

- the number of trainable parameters becomes extremely large



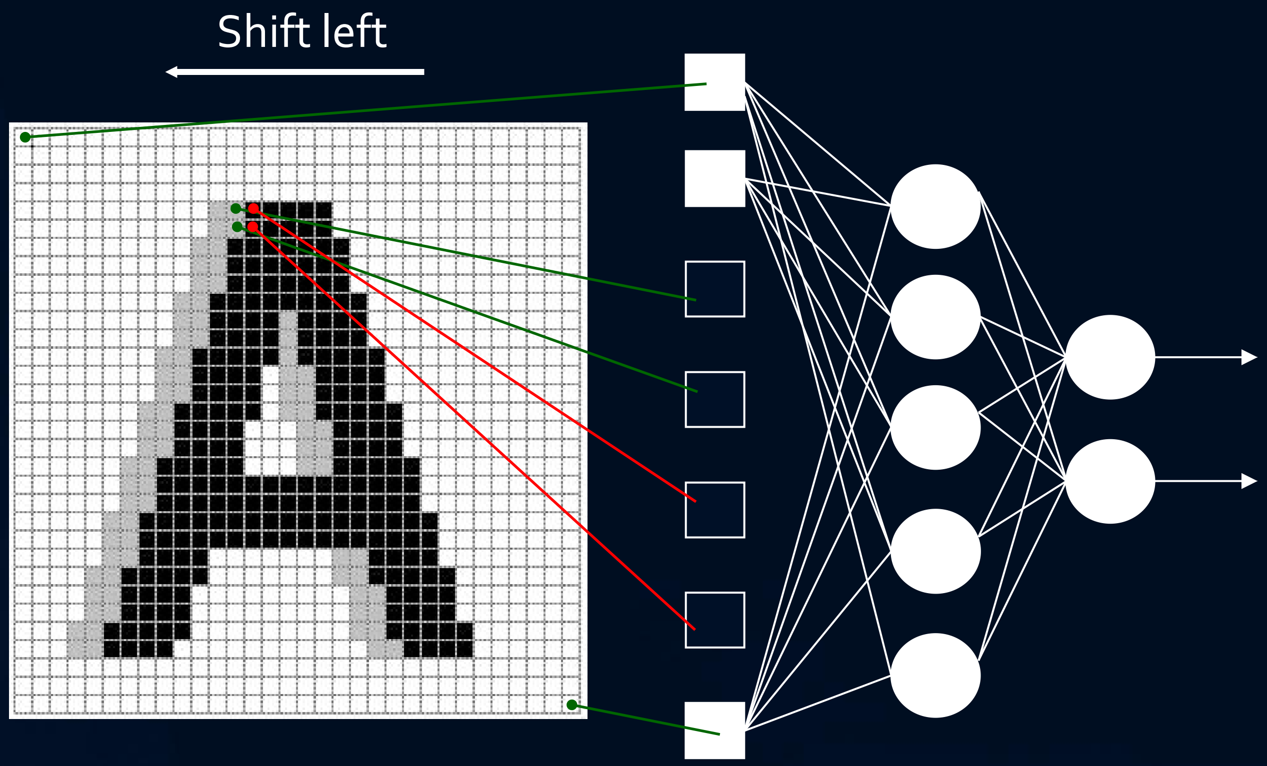
Drawbacks of previous neural networks

- Little or no invariance to shifting, scaling, and other forms of distortion

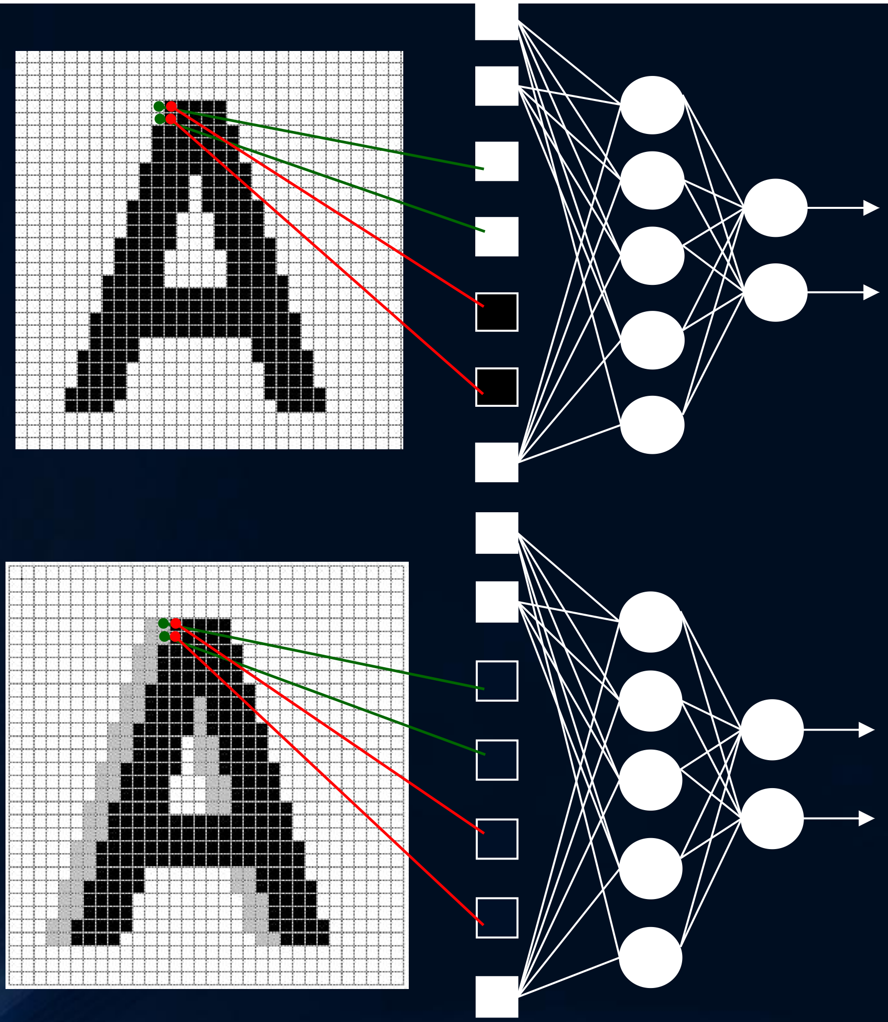


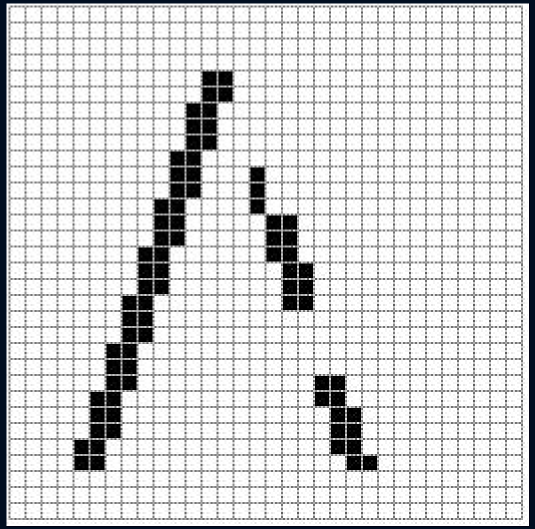
Drawbacks of previous neural networks

- Little or no invariance to shifting, scaling, and other forms of distortion



Drawbacks of previous neural networks

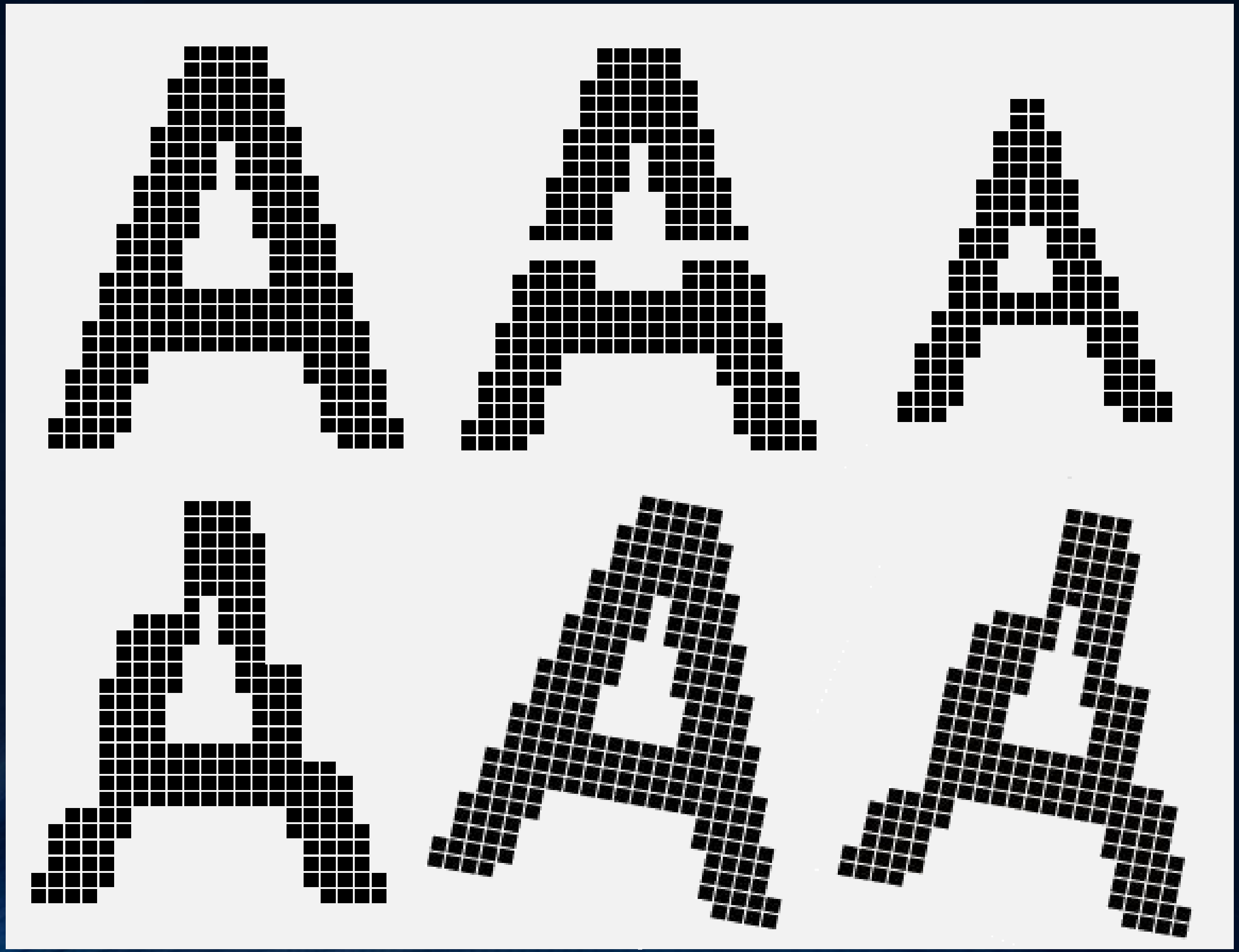




154 input change
from 2 shift left
77 : black to white
77 : white to black

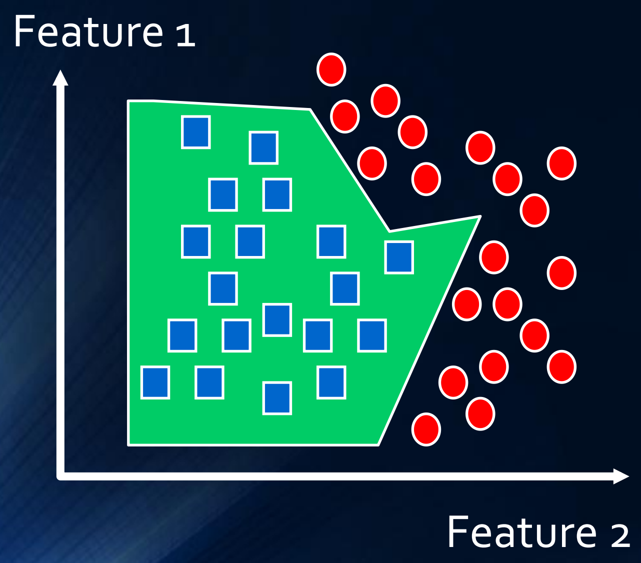
Drawbacks of previous neural networks

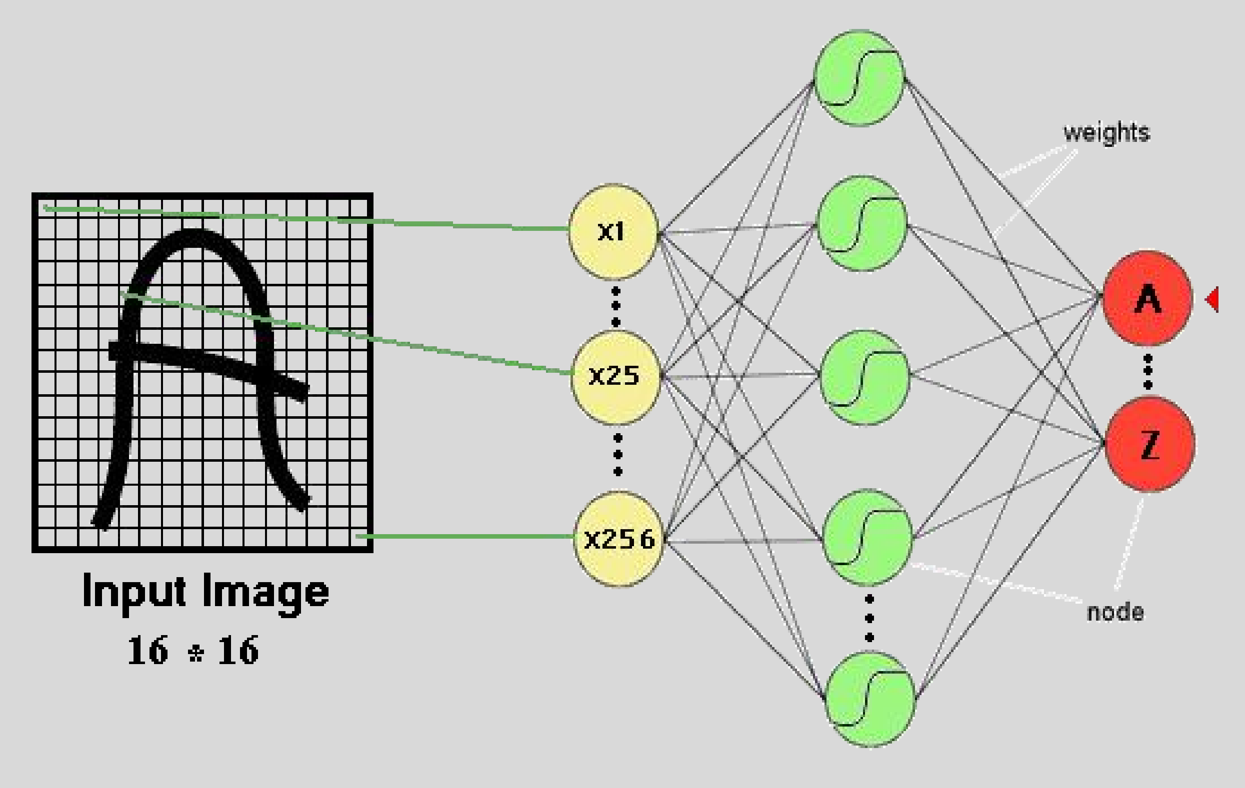
- scaling, and other forms of distortion



Drawbacks of previous neural networks

- the topology of the input data is completely ignored
- work with raw data.

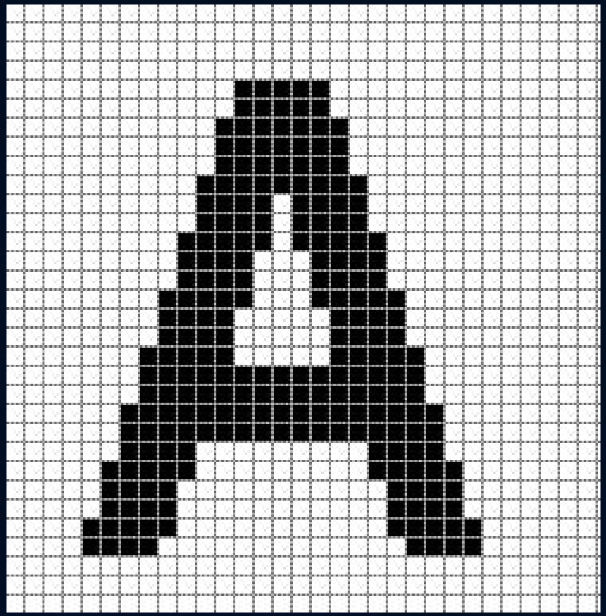




Drawbacks of previous neural networks

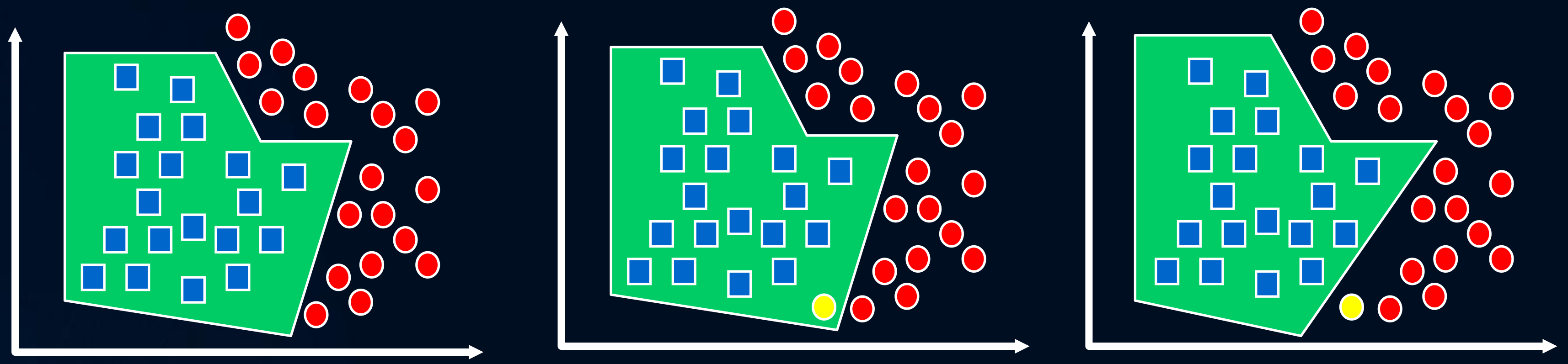
Black and white patterns: $2^{32*32} = 2^{1024}$

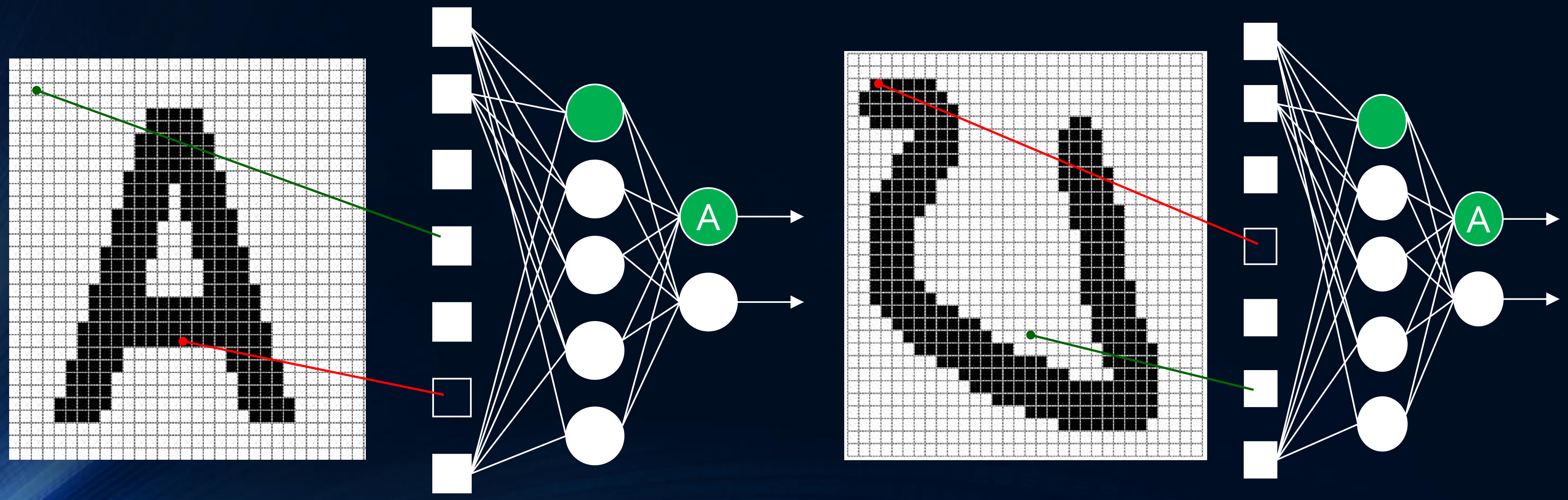
Gray scale patterns : $256^{32*32} = 256^{1024}$



32 * 32 input image

Drawbacks of previous neural networks





Improvement

- Fully connected network of sufficient size can produce outputs that are invariant with respect to such variations.
  - **Training time**
  - **Network size**
  - **Free parameters**

RoadMap

- Introduction
- Drawbacks of previous neural networks
- Convolutional neural networks
- LeNet 5
- Comparison
- Disadvantage
- Application

History

Yann LeCun, Professor of Computer Science
The Courant Institute of Mathematical Sciences
New York University
Room 1220, 715 Broadway, New York, NY 10003, USA.
(212)998-3283   yann@cs.nyu.edu

- In 1995, Yann LeCun and Yoshua Bengio introduced the concept of convolutional neural networks.

About CNN’s

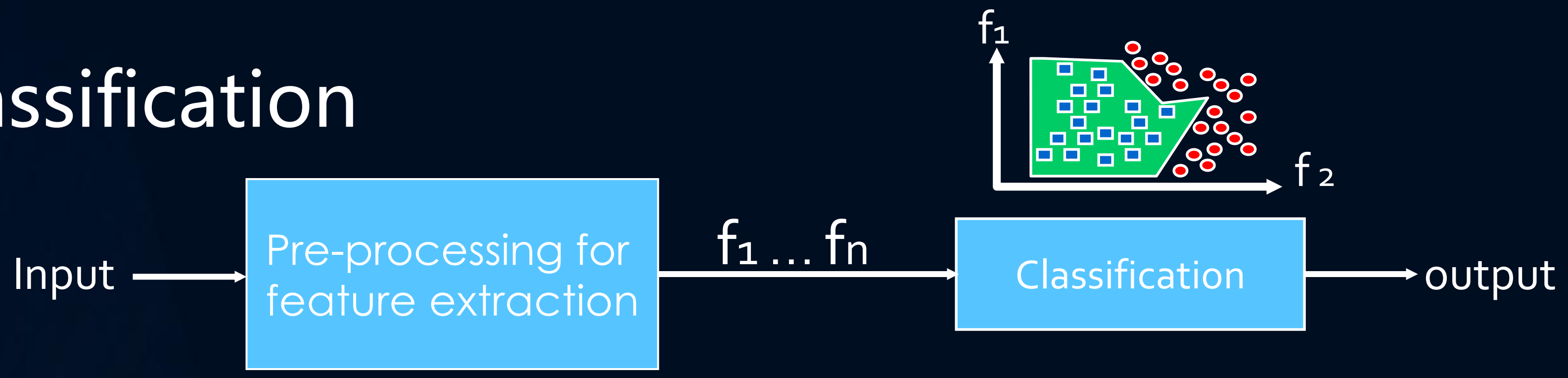
- CNN’s Were neurobiologically motivated by the findings of locally sensitive and orientation-selective nerve cells in the visual cortex.
- They designed a network structure that implicitly extracts relevant features.
- Convolutional Neural Networks are a special kind of multi-layer neural networks.

About CNN’s

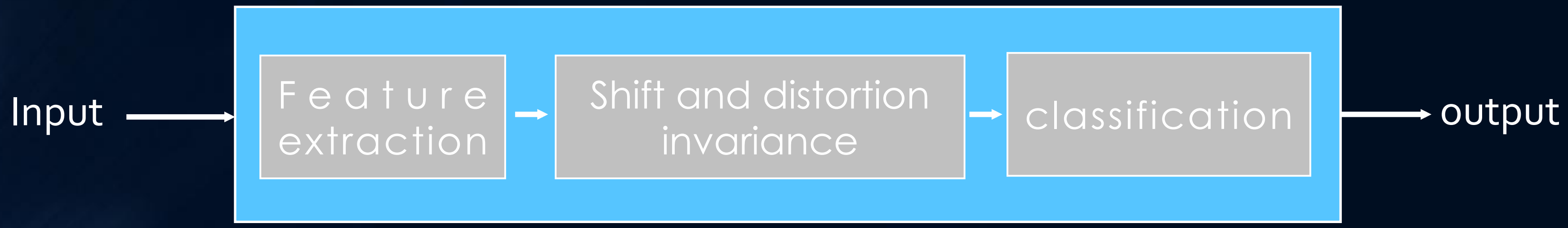
- CNN is a feed-forward network that can extract topological properties from an image.
- Like almost every other neural networks they are trained with a version of the back-propagation algorithm.
- Convolutional Neural Networks are designed to recognize visual patterns directly from pixel images with minimal preprocessing.
- They can recognize patterns with extreme variability (such as handwritten characters).

Classification

- Classification

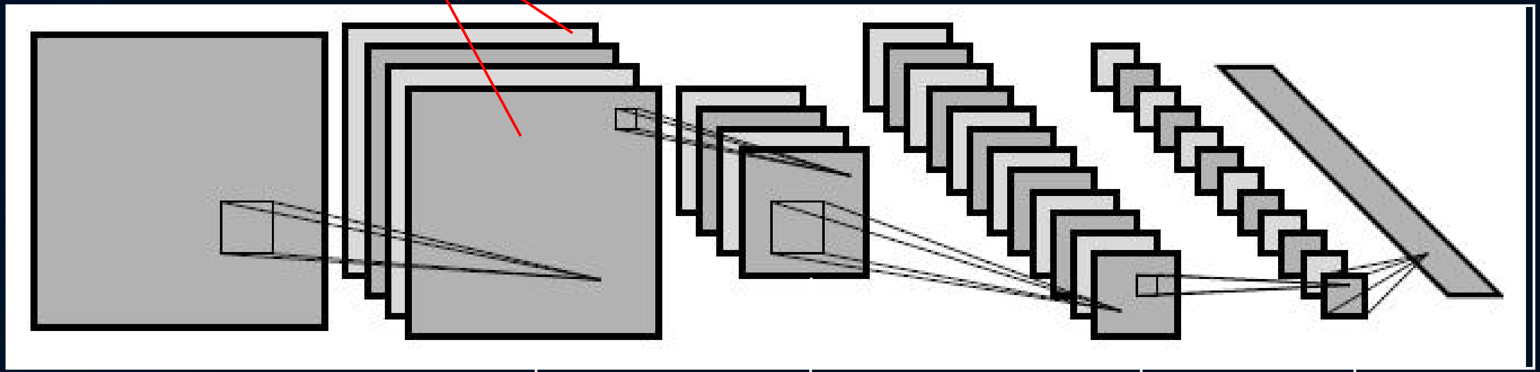


- Convolutional neural network



CNN’ s Topology

Feature maps



Feature extraction layer
Convolution layer

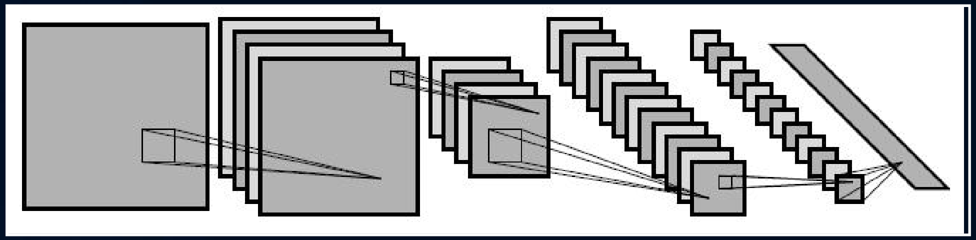
C

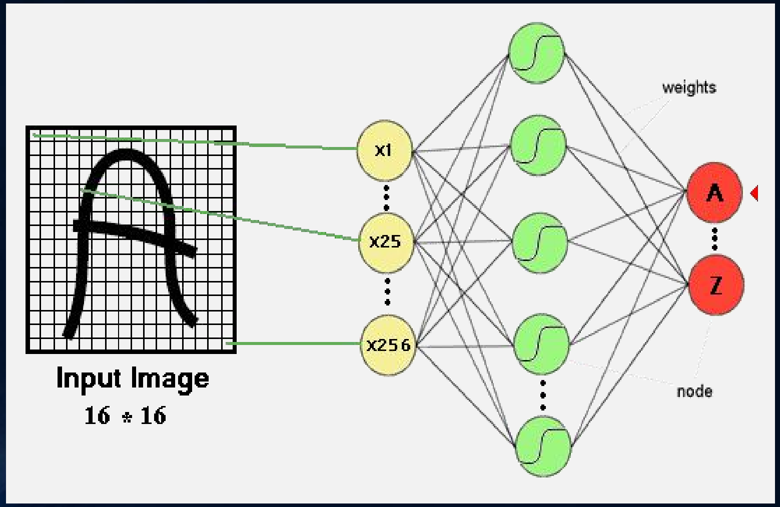
Shift and distortion invariance or
Subsampling layer

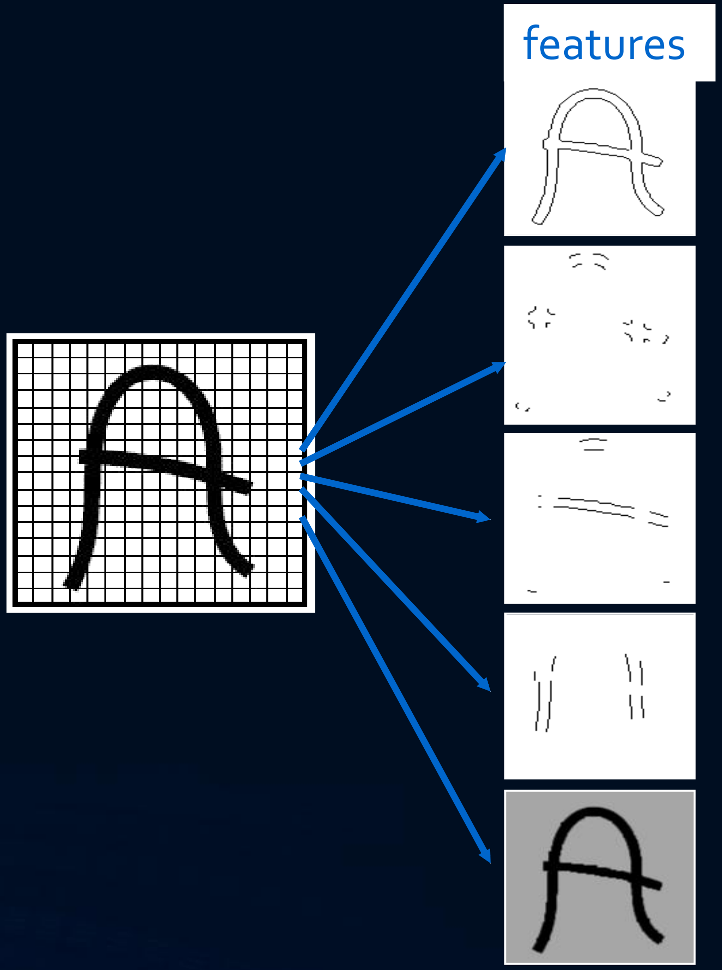
S

Feature extraction layer or Convolution layer

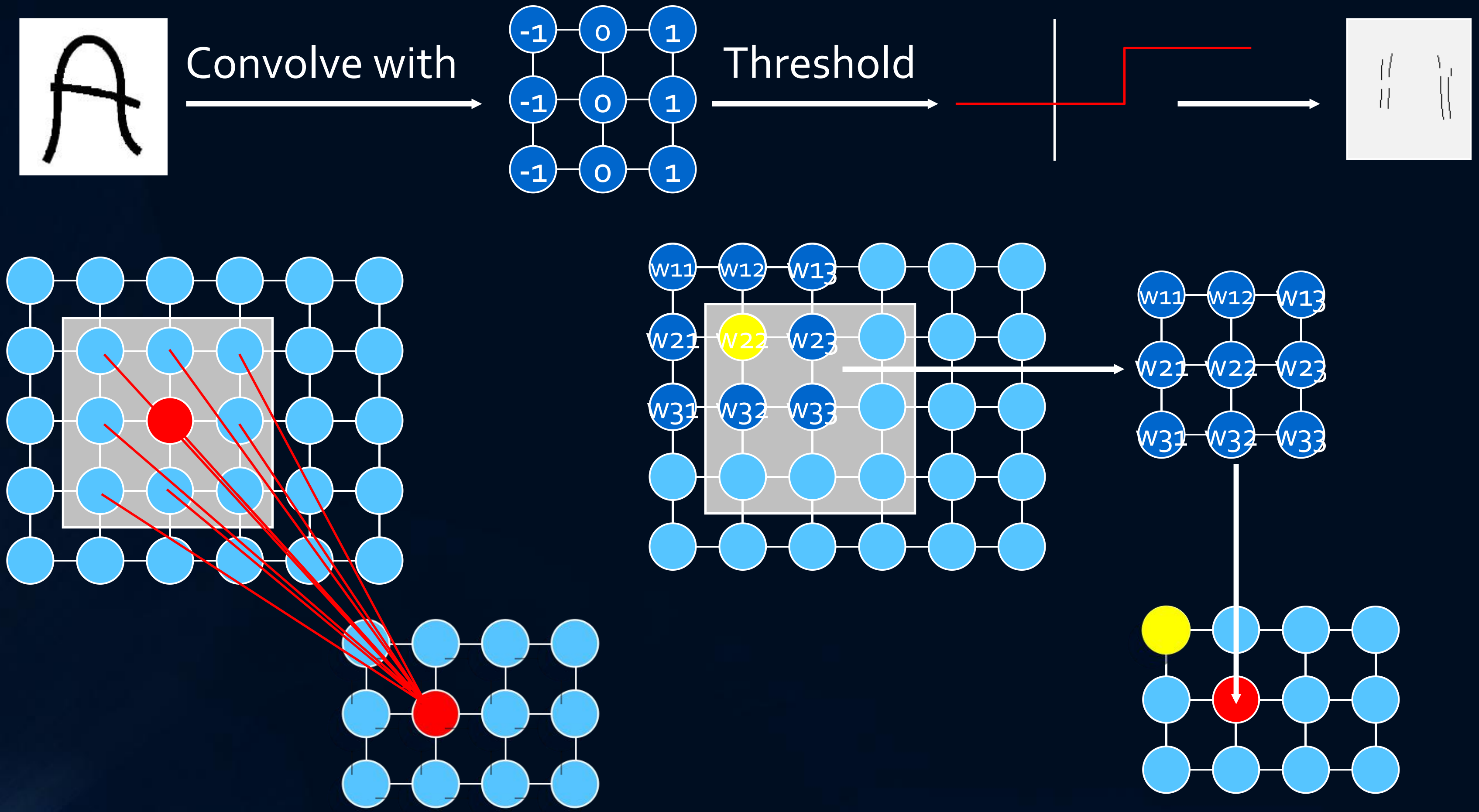
- detect the same feature at different positions in the input image.





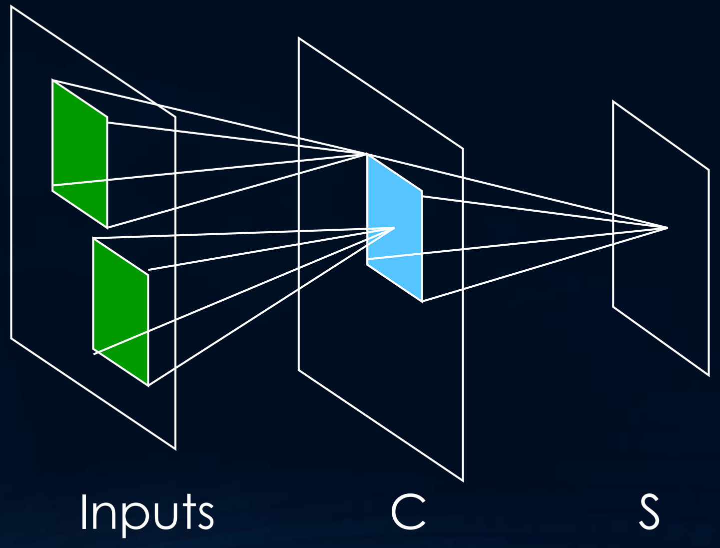


Feature extraction



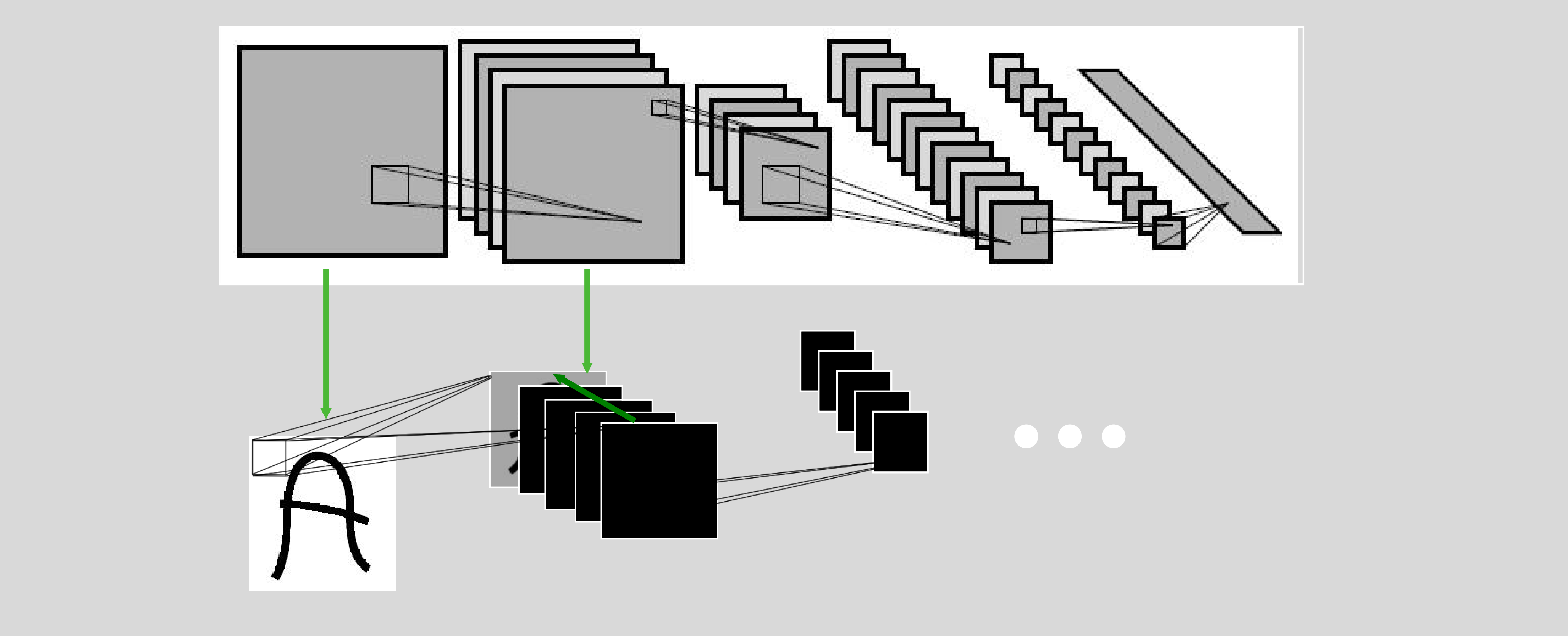
Feature extraction

- Shared weights: all neurons in a feature share the same weights (but not the biases).
- In this way all neurons detect the same feature at different positions in the input image.
- Reduce the number of free parameters.



Feature extraction

- If a neuron in the feature map fires, this corresponds to a match with the template.

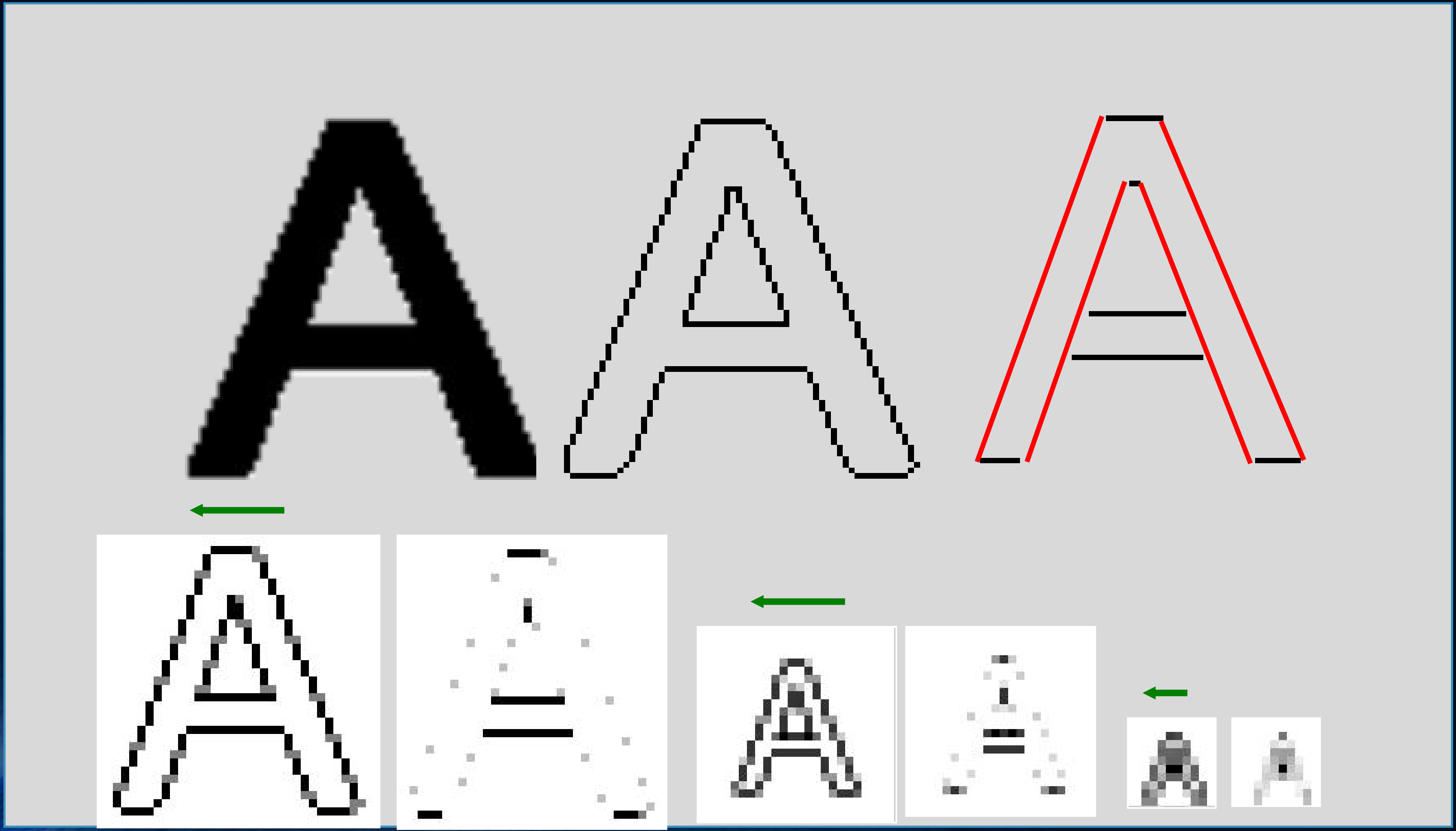


Subsampling layer

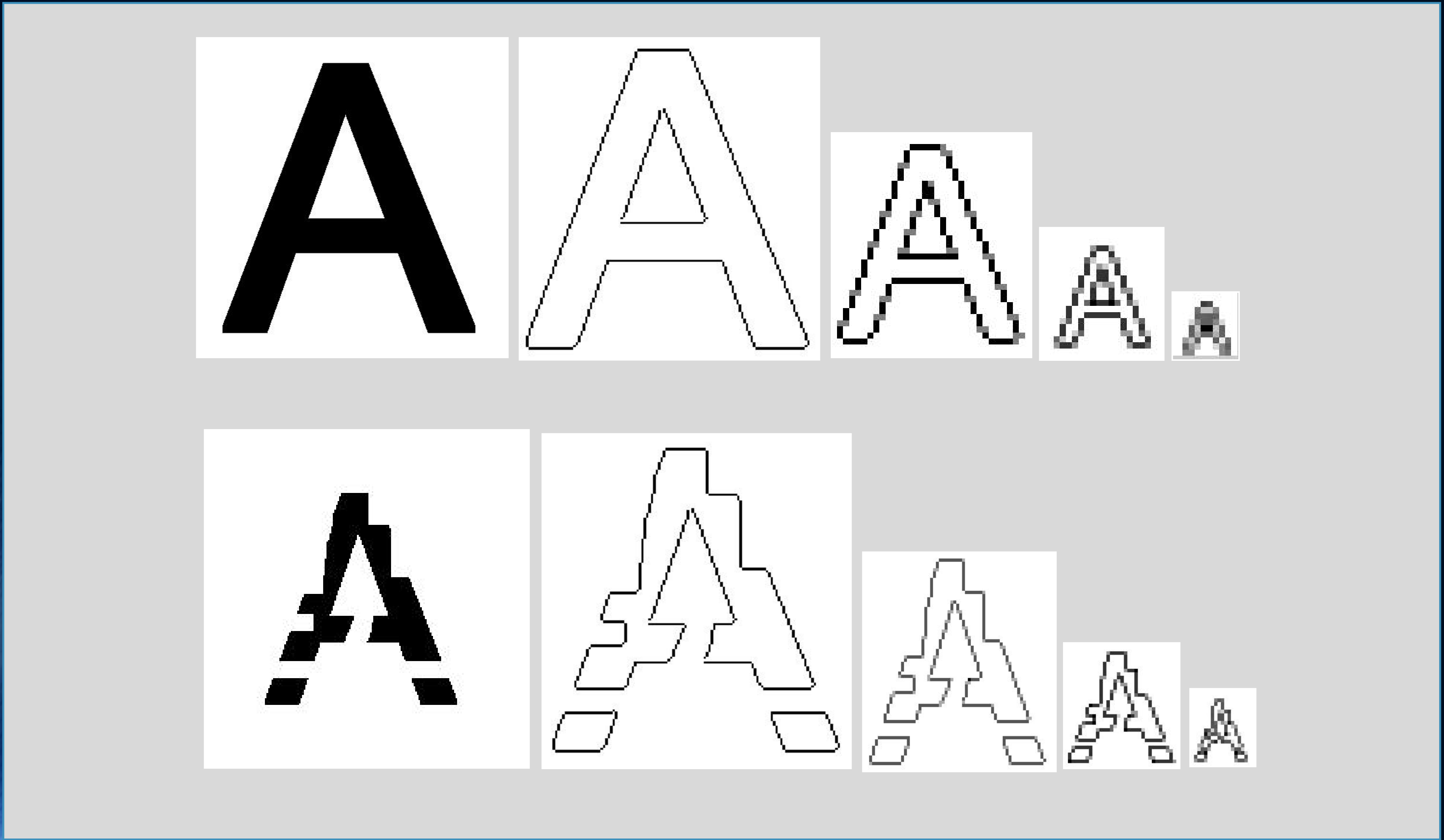
- the subsampling layers reduce the spatial resolution of each feature map
- By reducing the spatial resolution of the feature map, a certain degree of shift and distortion invariance is achieved.

Subsampling layer

- the subsampling layers reduce the spatial resolution of each feature map

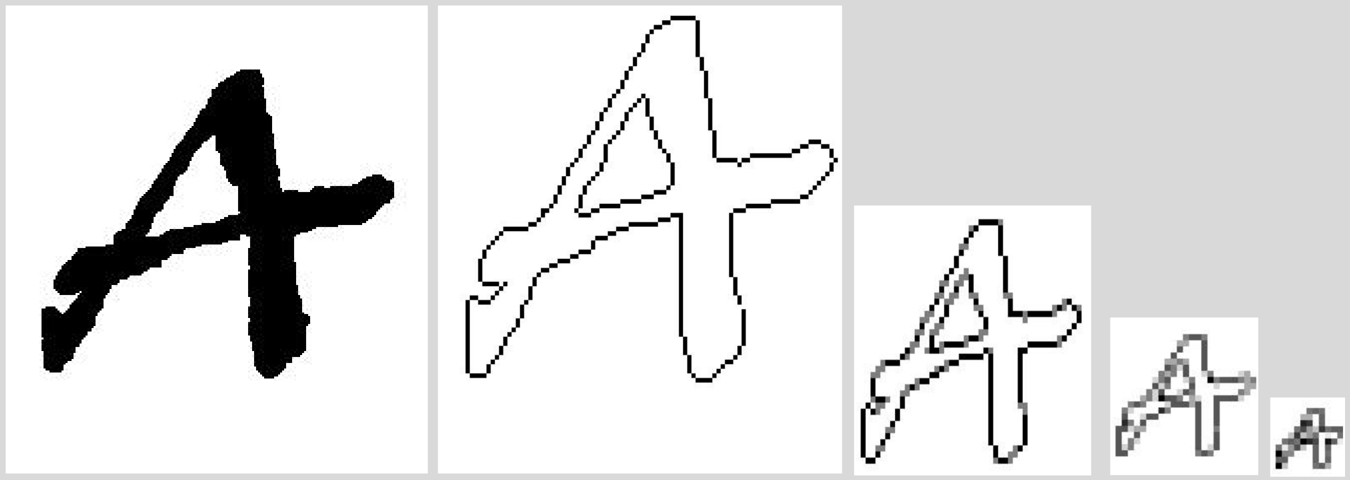


Subsampling layer



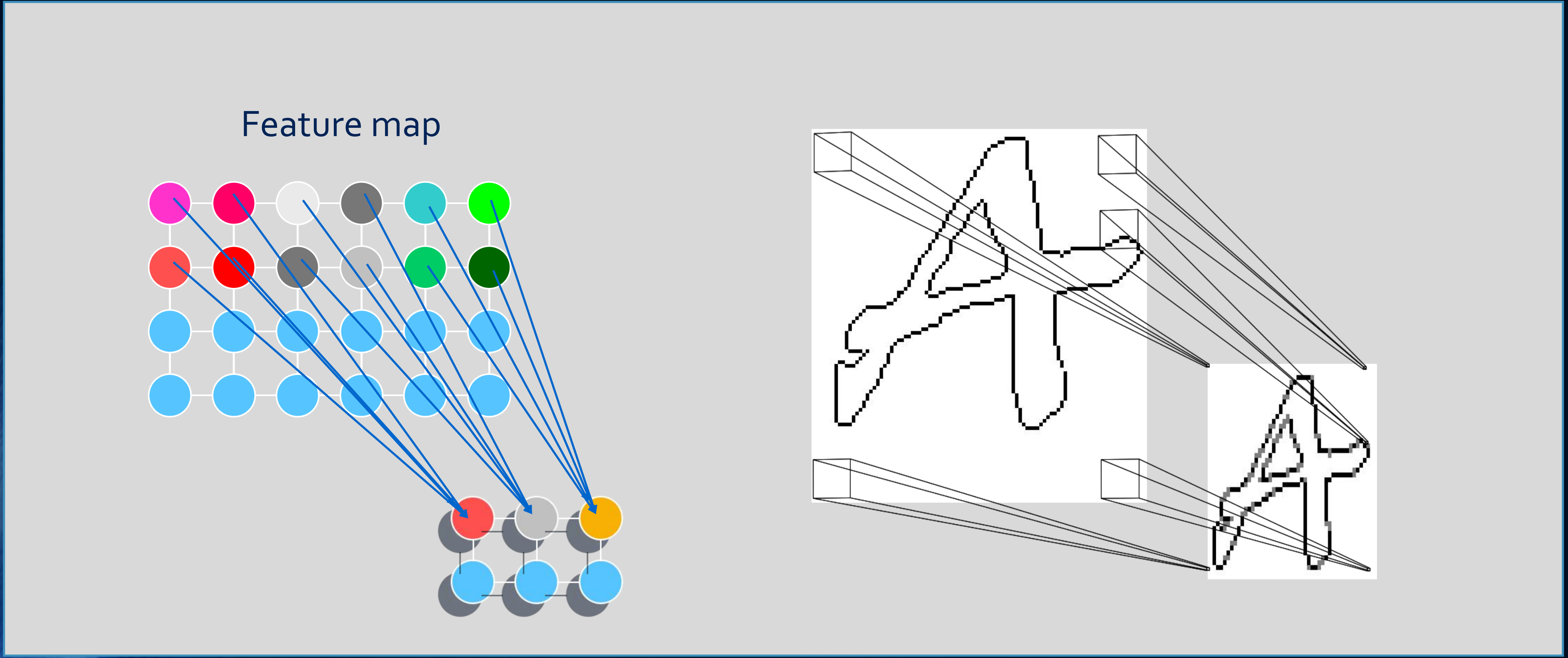
Subsampling layer

- The weight sharing is also applied in subsampling layers.

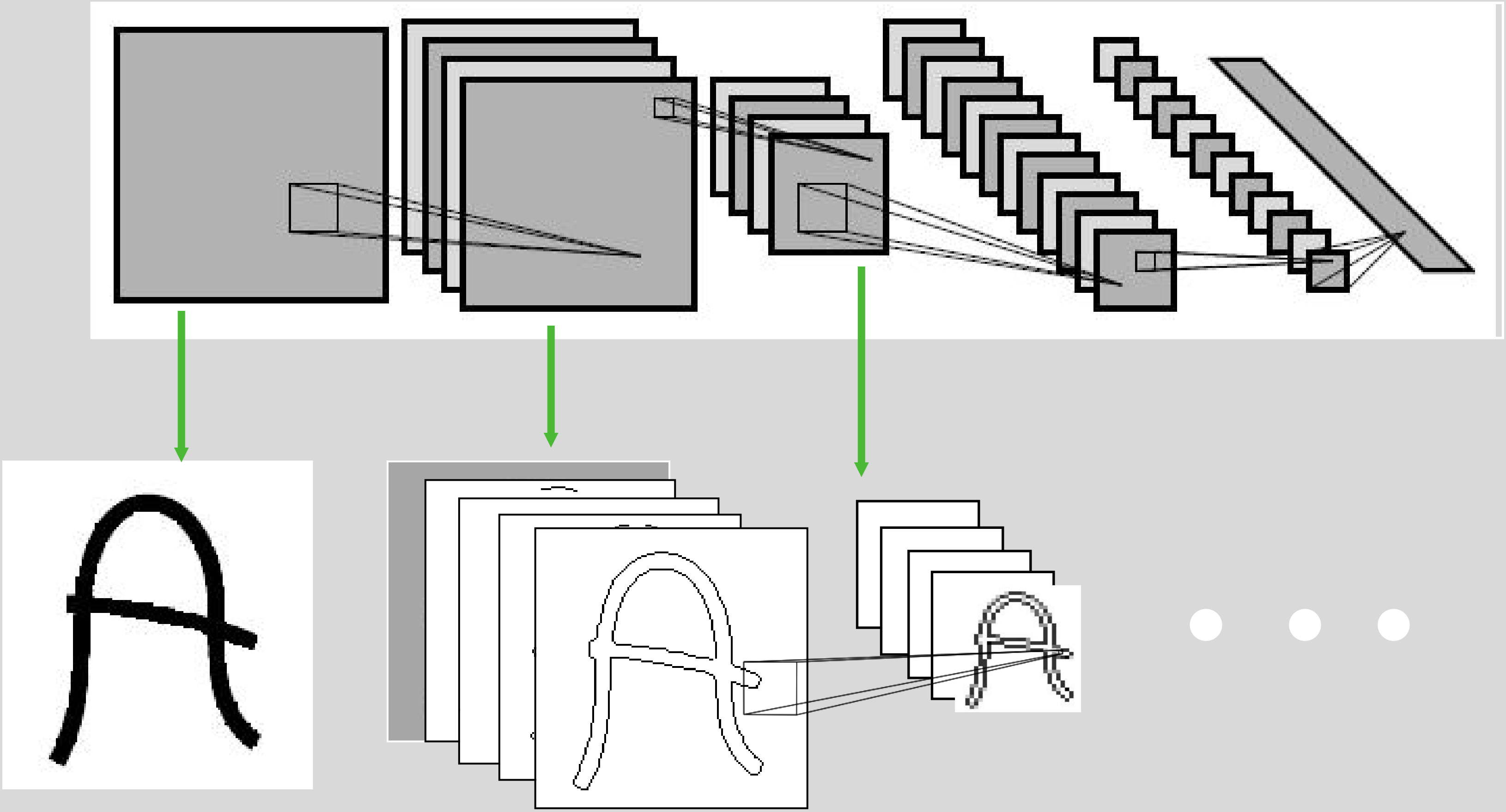


Subsampling layer

- the weight sharing is also applied in subsampling layers
- reduce the effect of noises and shift or distortion



Up to now …

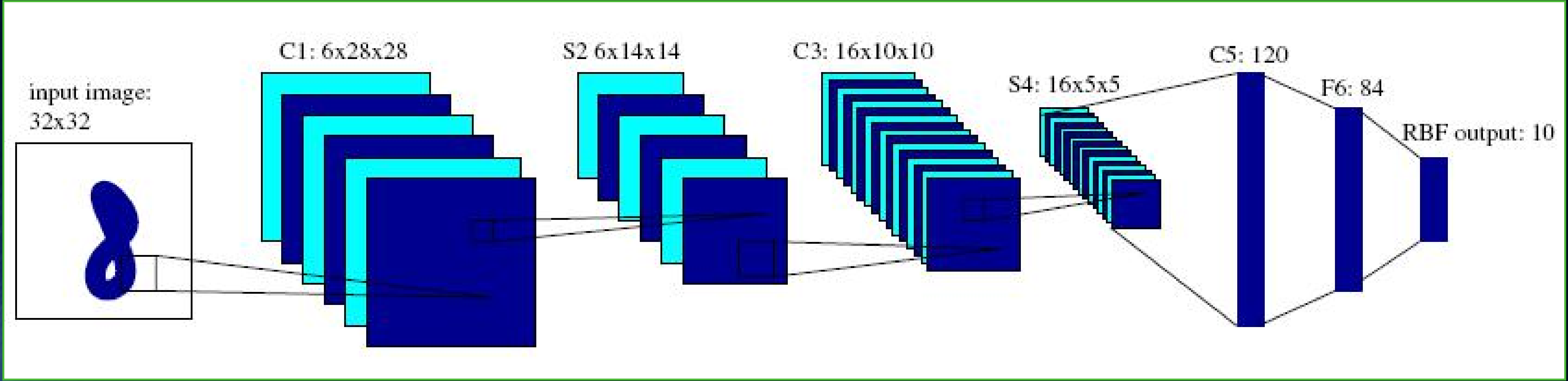


RoadMap

- Introduction
- Drawbacks of previous neural networks
- Convolutional neural networks
- LeNet 5
- Comparison
- Disadvantage
- Application

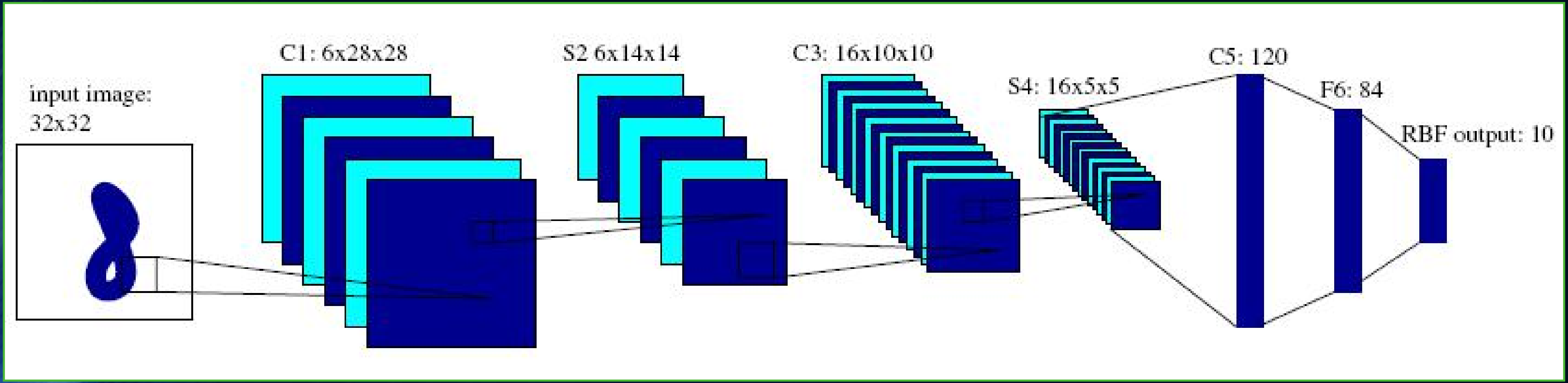
LeNet5

- ☯ Introduced by LeCun.
- ☯ raw image of 32 × 32 pixels as input.



LeNet5

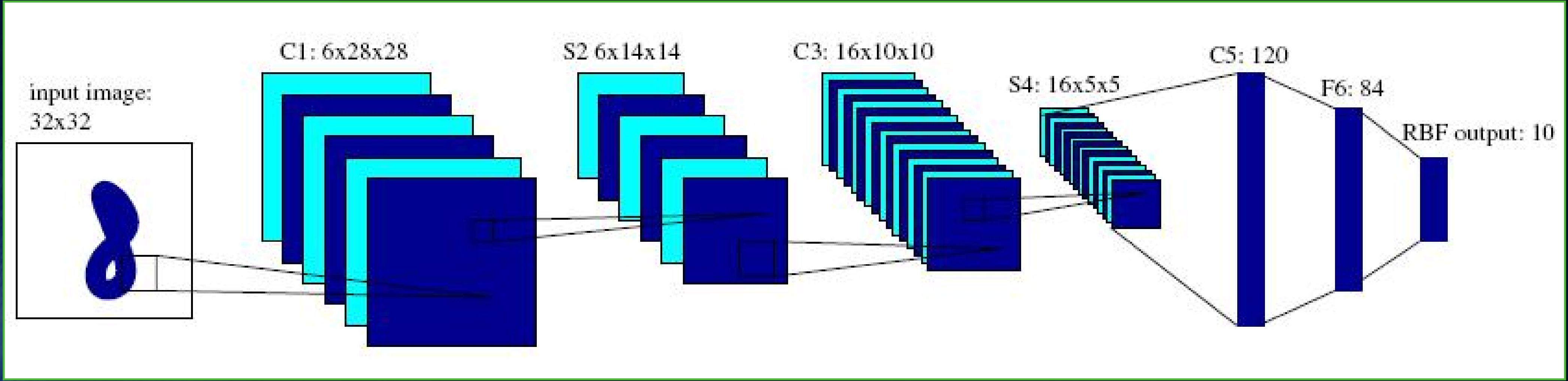
- C1,C3,C5 : Convolutional layer.
- 5 × 5 Convolution matrix.
- S2 , S4 : Subsampling layer.
- Subsampling by factor 2.
- F6 : Fully connected layer.



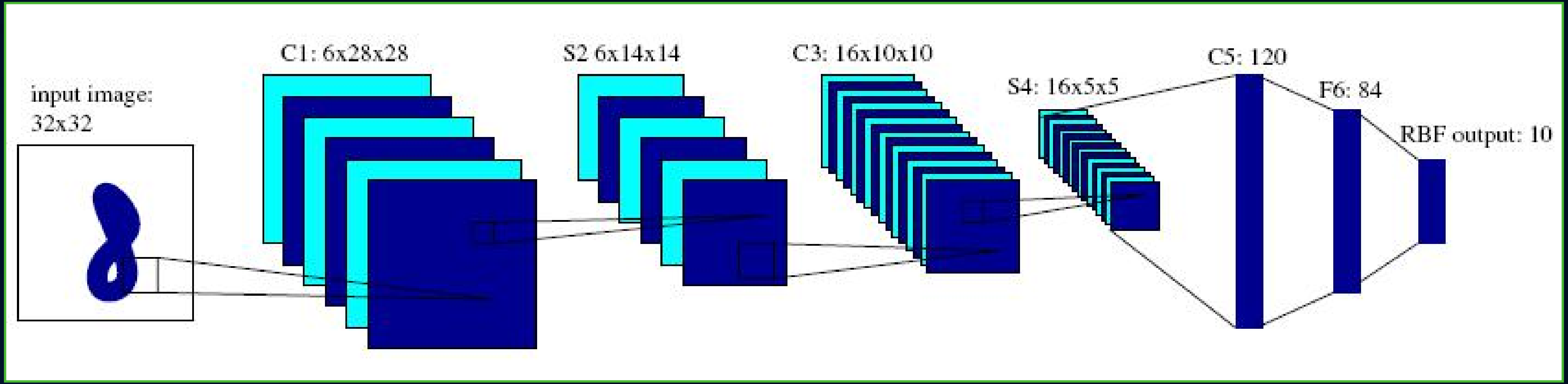
LeNet5

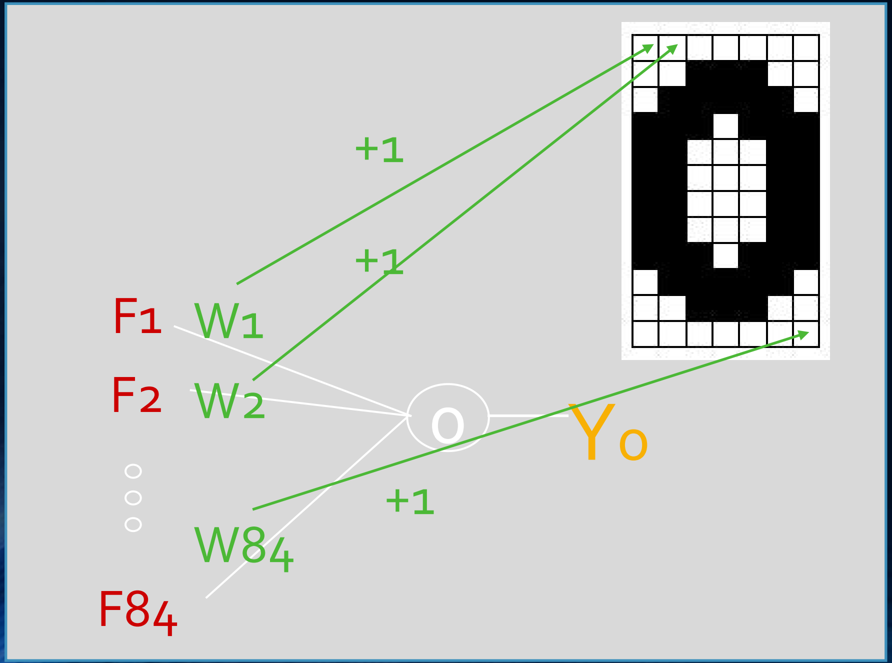
- All the units of the layers up to F6 have a sigmoidal activation function of the type:

$$y_j = \phi(v_j) = A\tanh\,(Sv_j)$$



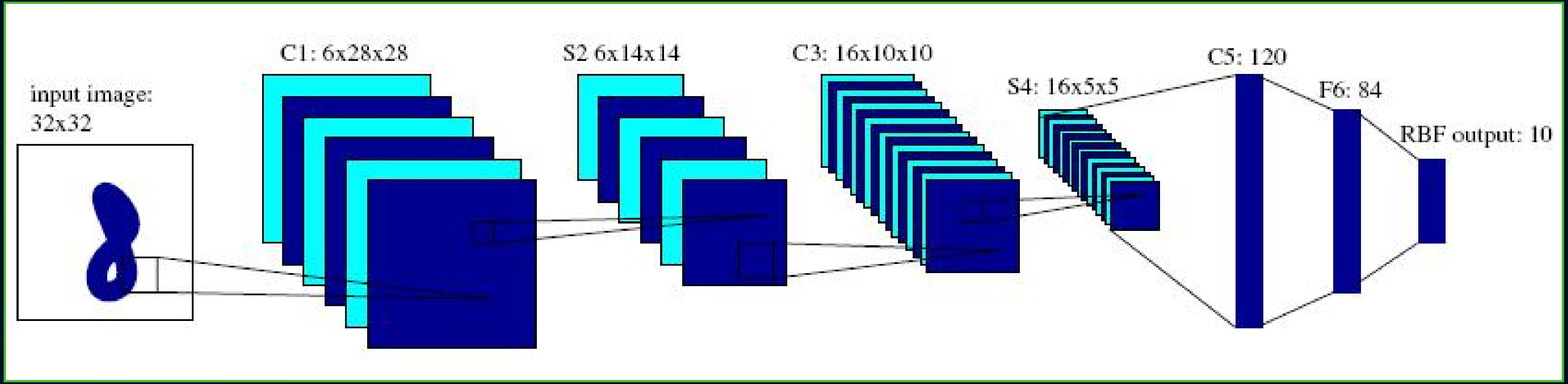
LeNet5



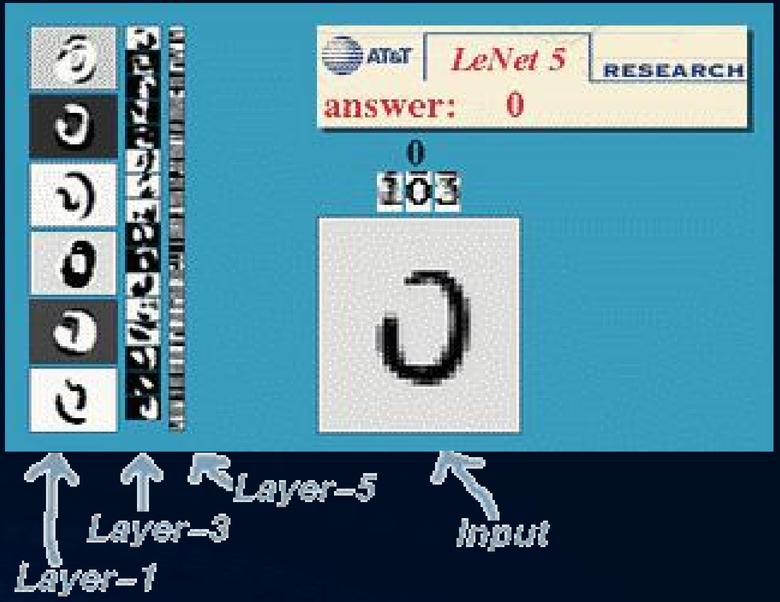


$$Y_j = \sum_{i=1}^{84} (F_i - W_{ij})^2, j = 0, \ldots, 9$$

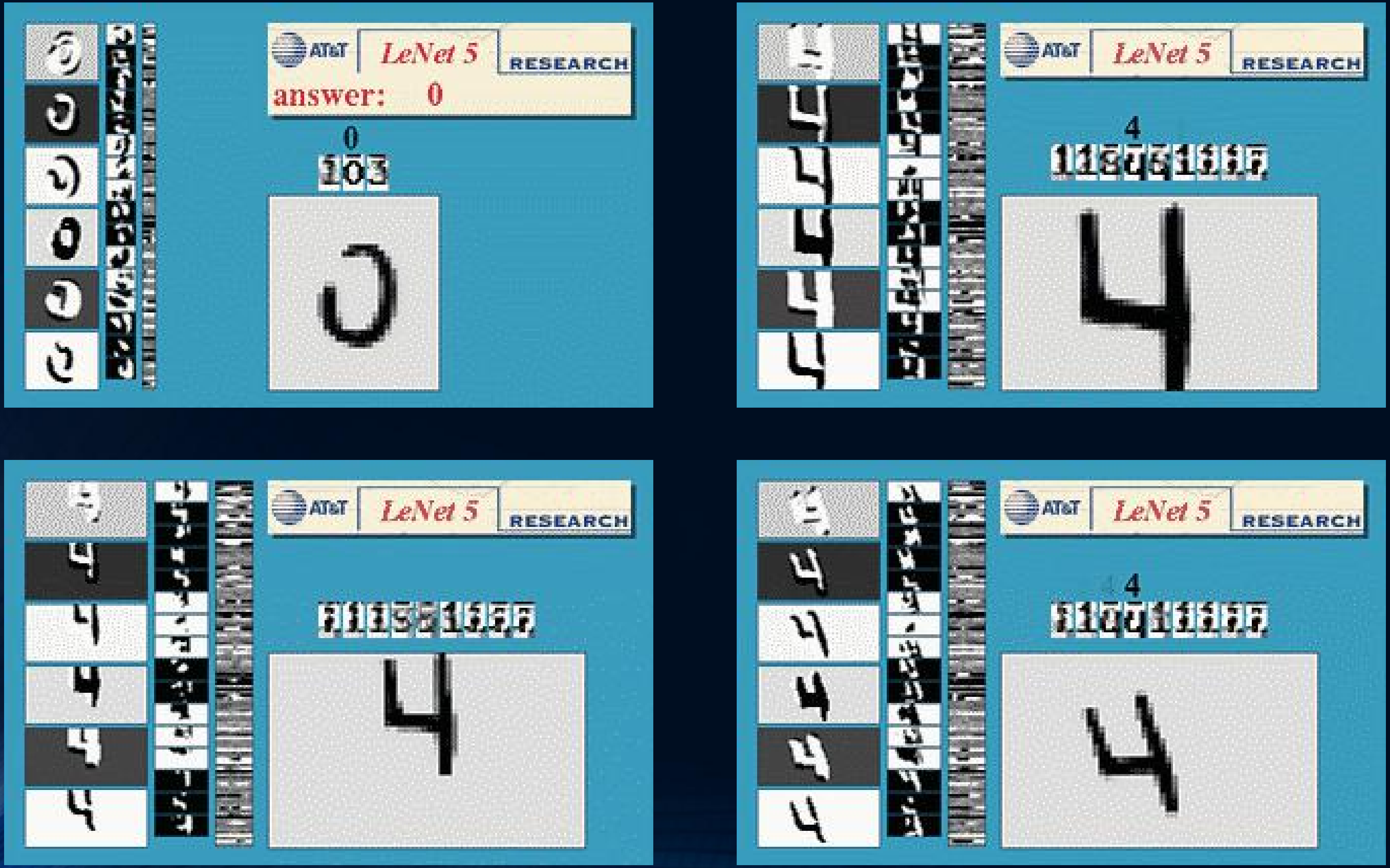
LeNet5



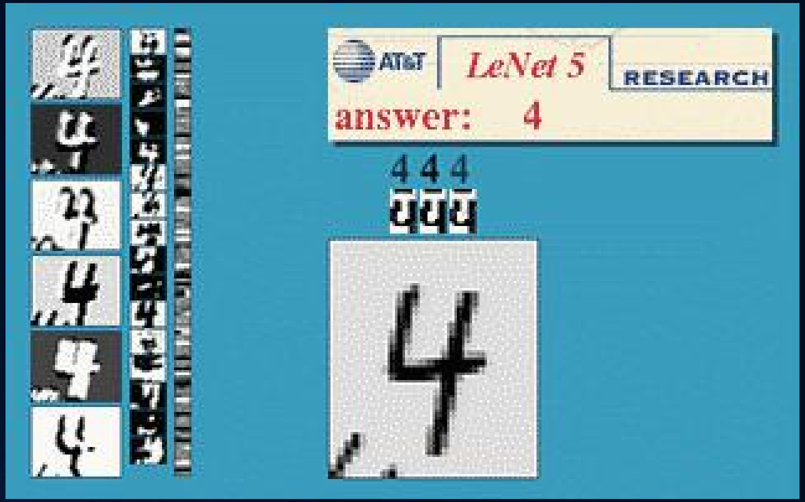
- About 187,000 connection.
- About 14,000 trainable weight.



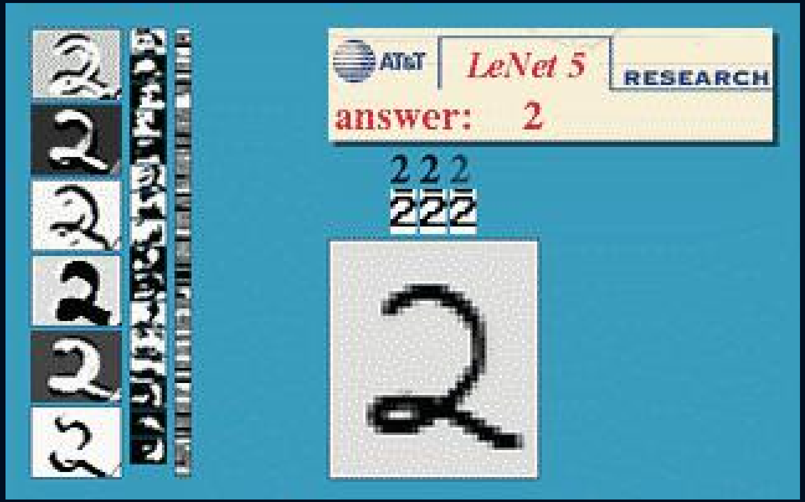
LeNet5


AT&T
LeNet 5
RESEARCH
answer: 0
AT&T
LeNet 5
RESEARCH
4
AT&T
LeNet 5
RESEARCH
AT&T
LeNet 5
RESEARCH
4

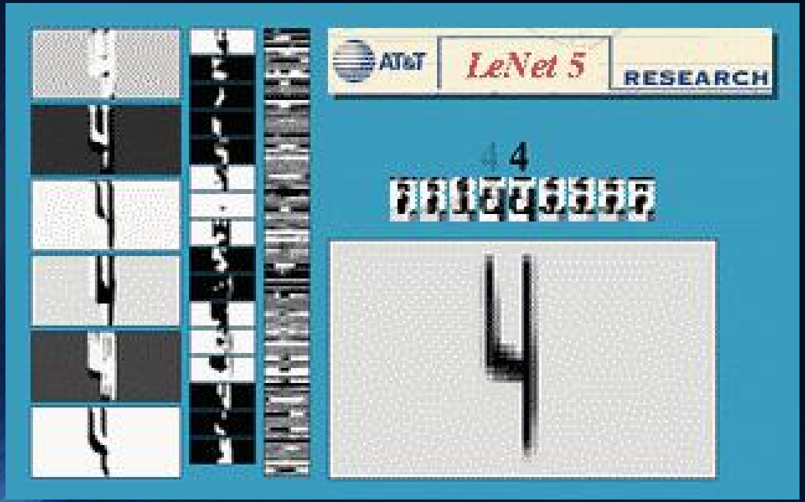
LeNet5



AT&T
LeNet 5
RESEARCH
answer: 4
4 4 4



AT&T
LeNet 5
RESEARCH
answer: 2
2 2 2



AT&T
LeNet 5
RESEARCH
4 4

AT&T
LeNet 5
RESEARCH
answer: 6
6 6

RoadMap

- Introduction
- Drawbacks of previous neural networks
- Convolutional neural networks
- LeNet 5
- Comparison
- Disadvantage
- Application

Comparison

- Database : MNIST (60,000 handwritten digits)
- Affine distortion : translation, rotation.
- elastic deformations : corresponding to uncontrolled oscillations of the hand muscles.
- MLP (this paper) : has 800 hidden unit.

Comparison

Algorithm	Distortion	Error	Ref.
2 layer MLP (MSE)	affine	1.6%	[3]
SVM	affine	1.4%	[9]
Tangent dist.	affine+thick	1.1%	[3]
Lenet5 (MSE)	affine	0.8%	[3]
Boost. Lenet4 MSE	affine	0.7%	[3]
Virtual SVM	affine	0.6%	[9]
2 layer MLP (CE)	none	1.6%	this paper
2 layer MLP (CE)	affine	1.1%	this paper
2 layer MLP (MSE)	elastic	0.9%	this paper
2 layer MLP (CE)	elastic	0.7%	this paper
Simple conv (CE)	affine	0.6%	this paper
Simple conv (CE)	elastic	**0.4%**	this paper

- "This paper" refer to reference[3] on references slide.

RoadMap

- Introduction
- Drawbacks of previous neural networks
- Convolutional neural networks
- LeNet 5
- Comparison
- Disadvantage
- Application

Disadvantages

- From a memory and capacity standpoint the CNN is not much bigger than a regular two layer network.
- At runtime the convolution operations are computationally expensive and take up about 67% of the time.
- CNN’s are about 3X slower than their fully connected equivalents (size-wise).

Disadvantages

- **Convolution operation**
  - 4 nested loops (2 loops on input image & 2 loops on kernel)
- **Small kernel size**
  - make the inner loops very inefficient as they frequently JMP.
- **Cash unfriendly memory access**
  - Back-propagation require both row-wise and column-wise access to the input and kernel image.
  - 2-D Images represented in a row-wise-serialized order.
  - Column-wise access to data can result in a high rate of cash misses in memory subsystem.

RoadMap

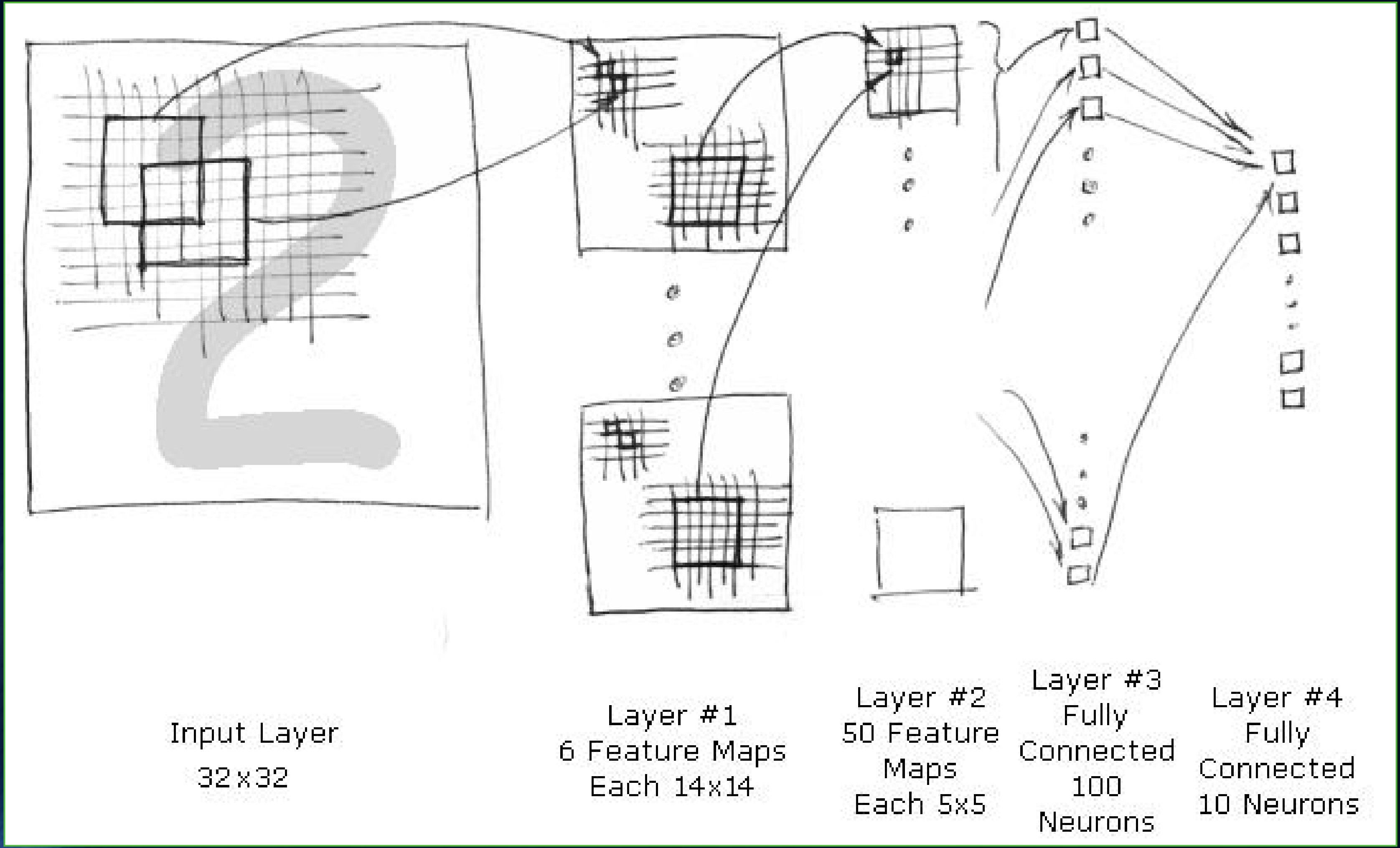
- Introduction
- Drawbacks of previous neural networks
- Convolutional neural networks
- LeNet 5
- Comparison
- Disadvantage
- Application

Application

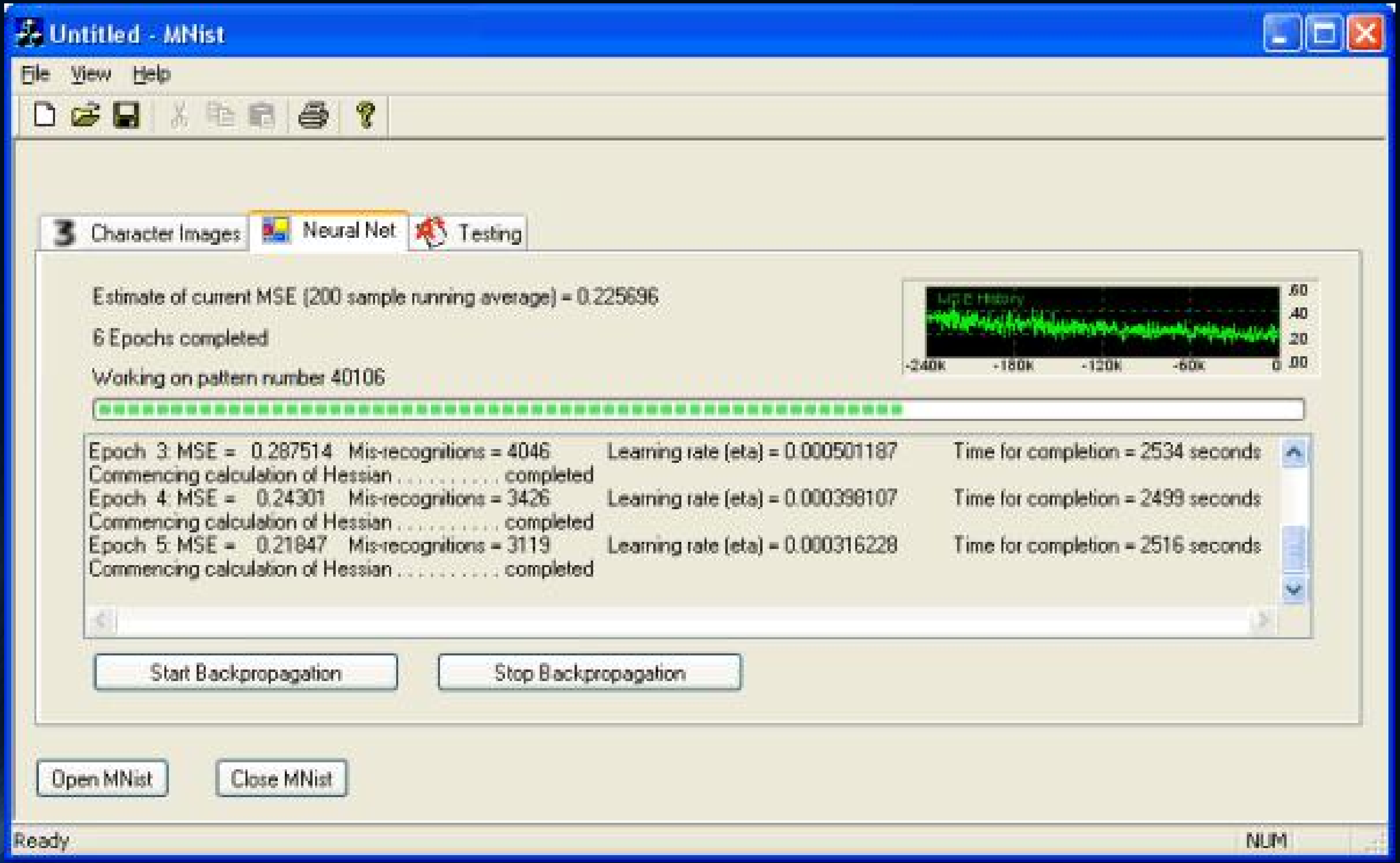
- A Convolutional neural network achieves 99.26% accuracy on a modified NIST database of
  hand-written digits.

- MNIST database : Consist of 60,000 hand written digits uniformly distributed over 0-9.

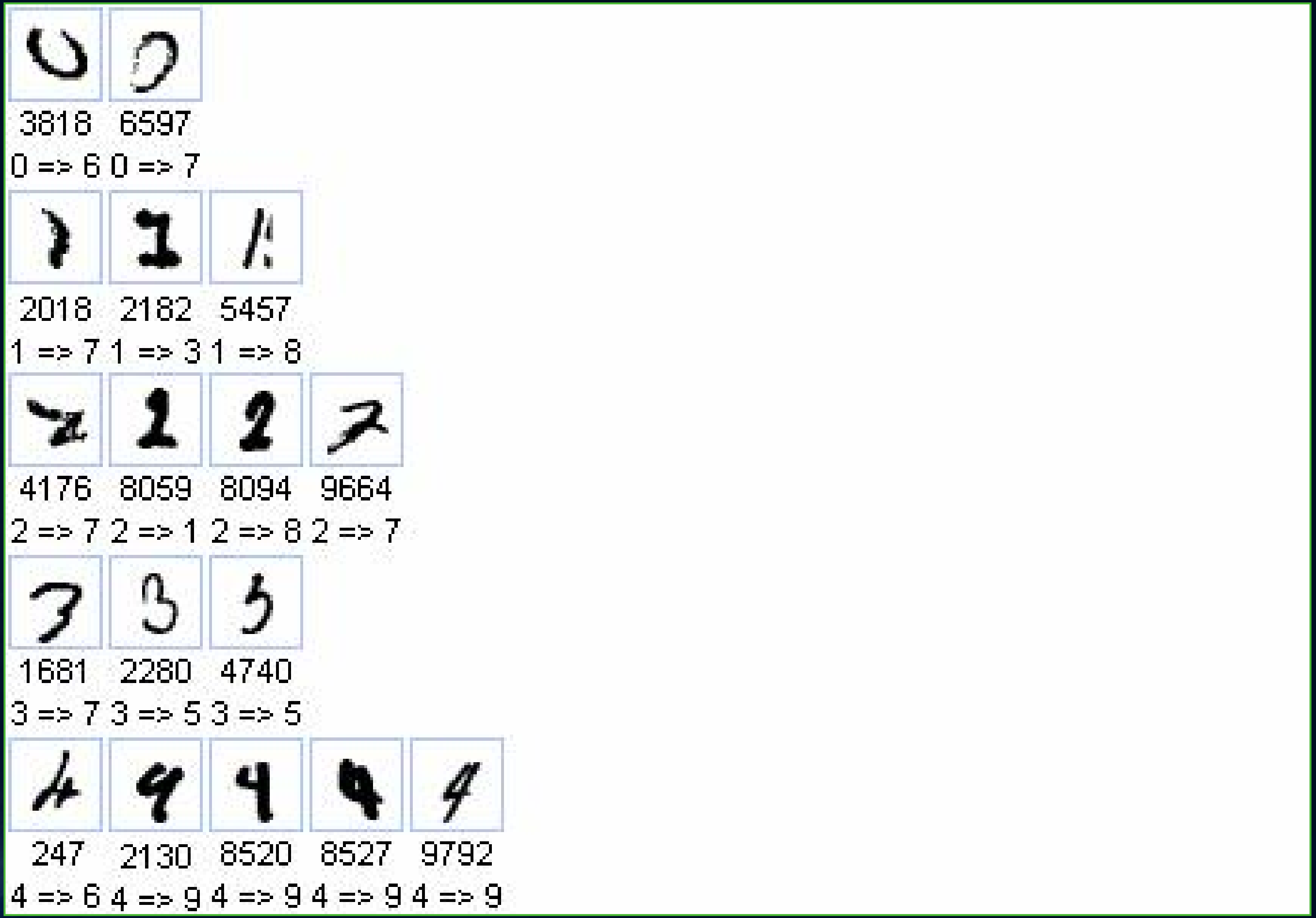
Application



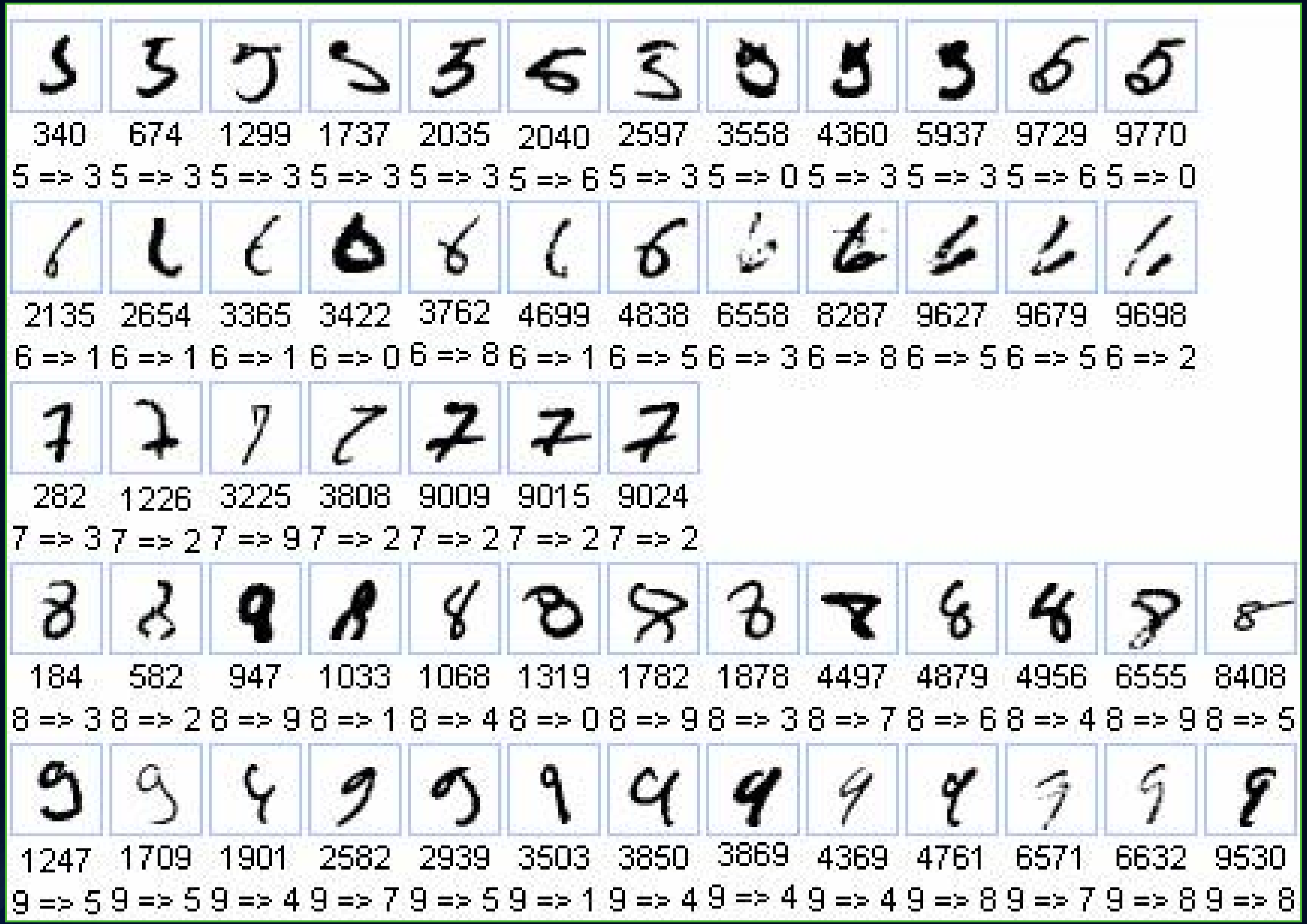
Application



Application



Application



References

[1].Y. LeCun and Y. Bengio."Convolutional networks for images, speech, and time-series." In M. A. Arbib, editor, *The Handbook of Brain Theory and Neural Networks*. MIT Press, 1995.

[2].Fabien Lauer, Ching Y. Suen, Gérard Bloch,"A trainable feature extractor for handwritten digit recognition",Elsevier, october 2006.

[3].Patrice Y. Simard, Dave Steinkraus, John Platt, "Best Practices for Convolutional Neural Networks Applied to Visual Document Analysis," International Conference onDocument Analysis and Recognition (ICDAR), IEEE Computer Society, Los Alamitos, pp. 958-962, 2003.

Next

Recurrent Neural Networks