

Machine Learning, Spring 2020

Deep Learning with Auto-Encoding

Instructor: Prof. Yi Fang
yfang@nyu.edu

Python tutorial: <http://learnpython.org/>

TensorFlow tutorial: <https://www.tensorflow.org/tutorials/>

PyTorch tutorial: <https://pytorch.org/tutorials/>

Acknowledge: The slides are partially referred to the online materials by Prof. Hung-yi Lee from NTU, Auto-encoder session, and also other online materials

Note that some figures and contents used in the presentation are from public available resources, the original authors own the copyright of the figures and content

Unsupervised Learning

“We expect unsupervised learning to become far more important in the longer term. Human and animal learning is largely unsupervised: we discover the structure of the world by observing it, not by being told the name of every object.”
– LeCun, Bengio, Hinton, Nature 2015

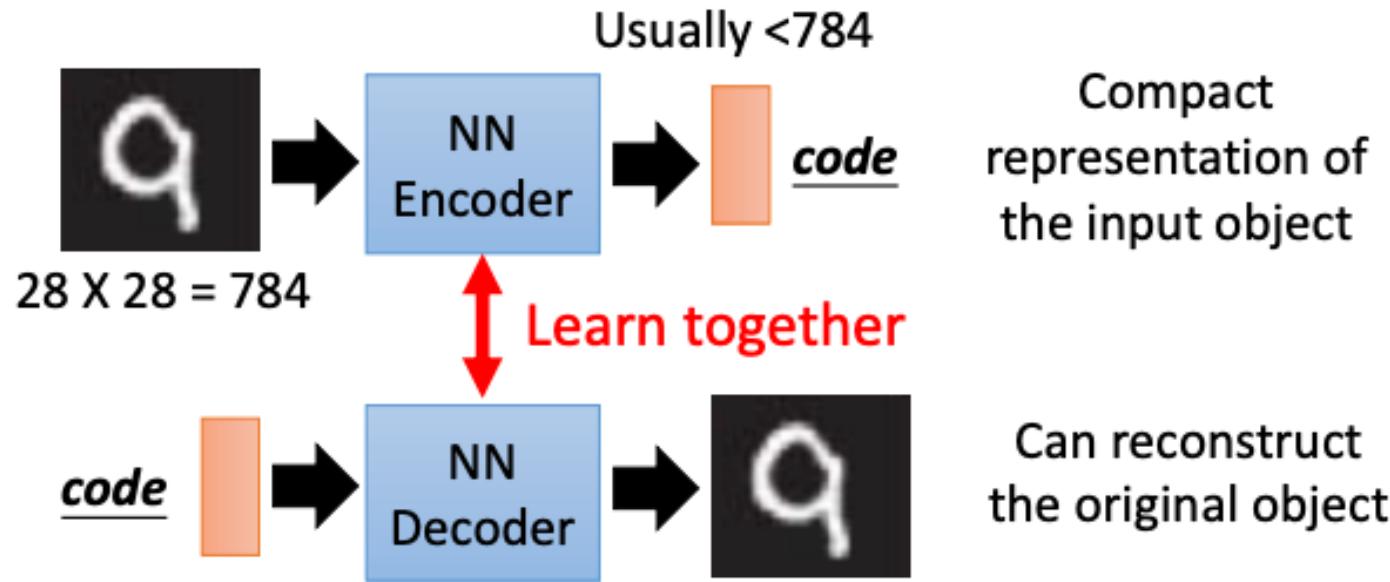
As I've said in previous statements: most of human and animal learning is unsupervised learning. If intelligence was a cake, unsupervised learning would be the cake, supervised learning would be the icing on the cake, and reinforcement learning would be the cherry on the cake. We know how to make the icing and the cherry, but we don't know how to make the cake.
- Yann LeCun, March 14, 2016 (Facebook)

Unsupervised Learning

What we have learnt:

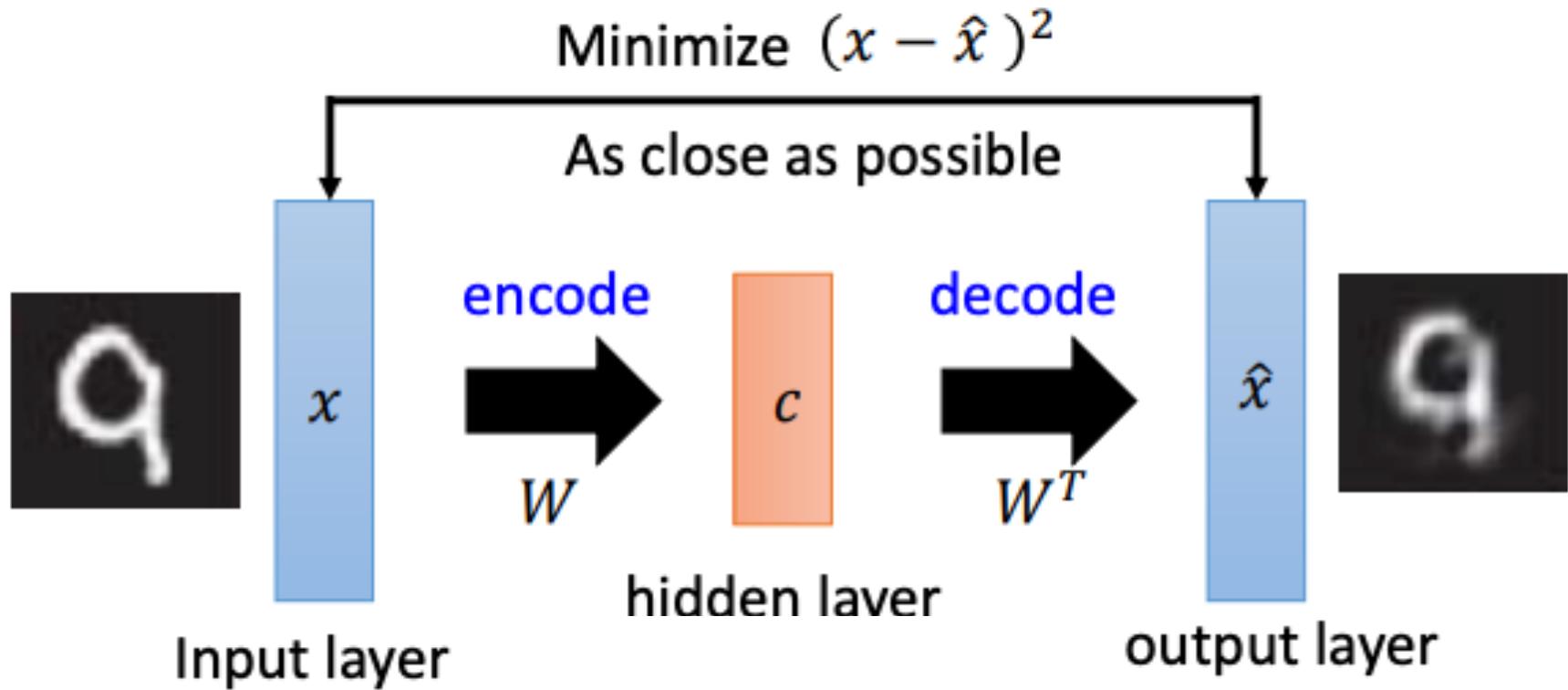
- Manifold Learning: PCA, LLE, MDS, etc.

How Deep Learning deals with unsupervised learning?



Acknowledge: The slides are partially referred to the online materials by Taegyun Joen, <https://www.slideshare.net/TaegyunJeon1/pr12-you-only-look-once-yolo-unified-realtime-object-detection> and online YOLO paper and other materials (from ECS289g by Prof. Lee)

Deep Auto-Encoder

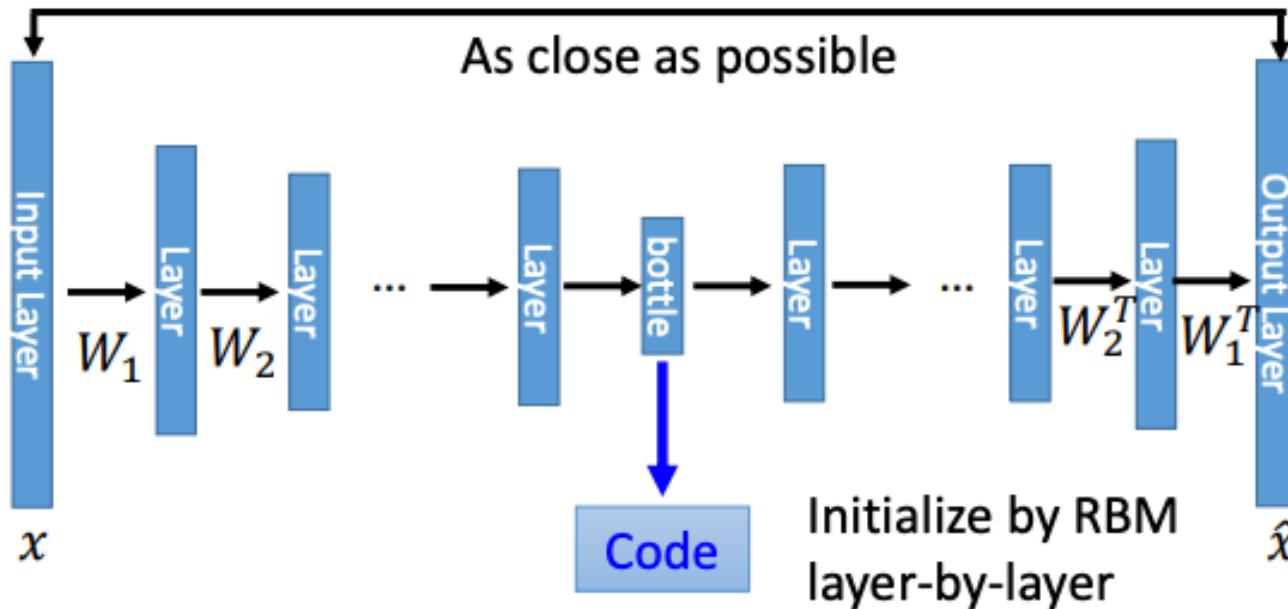


Acknowledge: The slides are partially referred to the online materials by Taegyun Joen, <https://www.slideshare.net/TaegyunJeon1/pr12-you-only-look-once-yolo-unified-realtime-object-detection> and online YOLO paper and other materials (from ECS289g by Prof. Lee)

Deep Auto-encoder

Symmetric is not necessary.

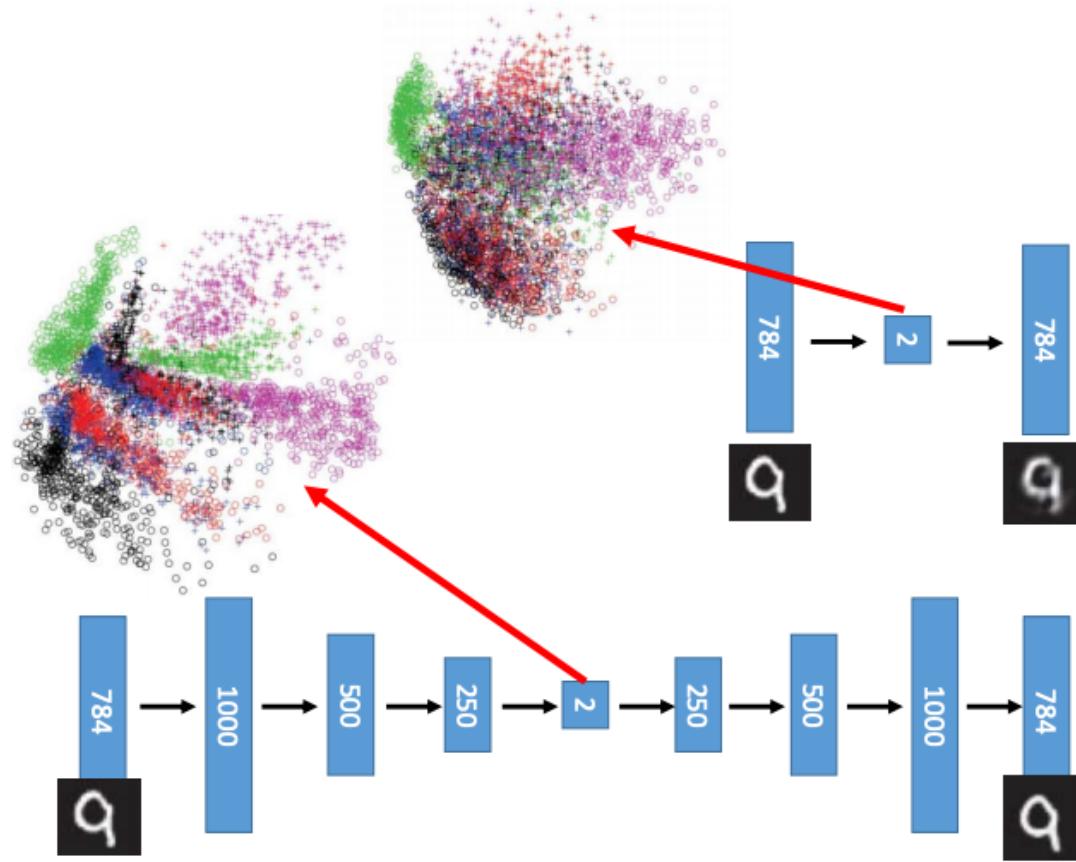
- Of course, the auto-encoder can be deep



Reference: Hinton, Geoffrey E., and Ruslan R. Salakhutdinov. "Reducing the dimensionality of data with neural networks." *Science* 313.5786 (2006): 504-507

Acknowledge: The slides are partially referred to the online materials by Taegyun Joen, <https://www.slideshare.net/TaegyunJoen1/pr12-you-only-look-once-yolo-unified-realtime-object-detection> and online YOLO paper and other materials (from ECS289g by Prof. Lee)

Manifold Embedding Visualization



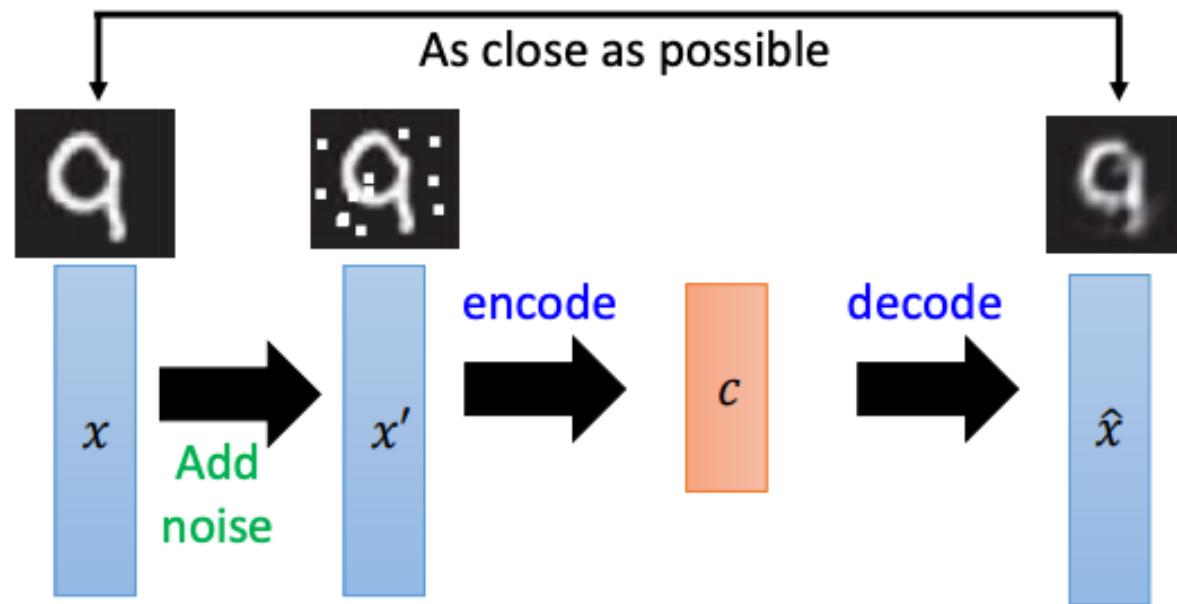
Acknowledge: The slides are partially referred to the online materials by Taegyun Joen, <https://www.slideshare.net/TaegyunJeon1/pr12-you-only-look-once-yolo-unified-realtime-object-detection> and online YOLO paper and other materials (from ECS289g by Prof. Lee)

Auto-encoder

- De-noising auto-encoder

More: Contractive auto-encoder

Ref: Rifai, Salah, et al. "Contractive auto-encoders: Explicit invariance during feature extraction." *Proceedings of the 28th International Conference on Machine Learning (ICML-11)*. 2011.



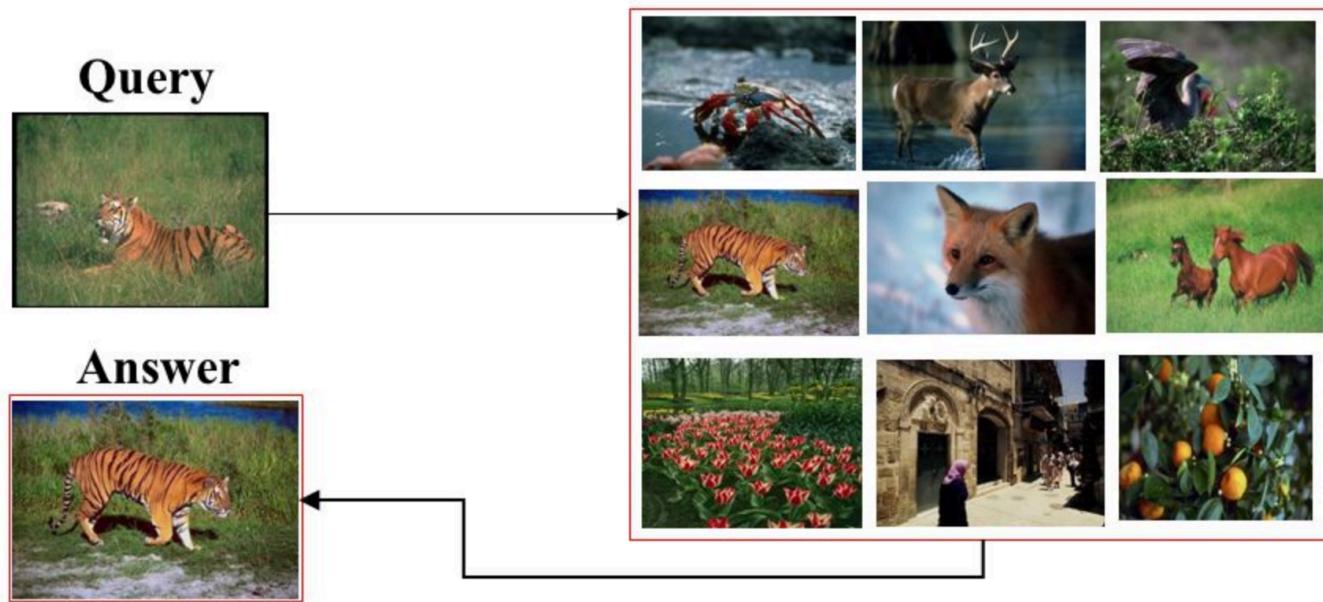
Vincent, Pascal, et al. "Extracting and composing robust features with denoising autoencoders." *ICML*, 2008.

Acknowledge: The slides are partially referred to the online materials by Taegyun Joen, <https://www.slideshare.net/TaegyunJoen1/pr12-you-only-look-once-yolo-unified-realtime-object-detection> and online YOLO paper and other materials (from ECS289g by Prof. Lee)

Auto-encoder for Image Retrieval

Content-based Image Retrieval

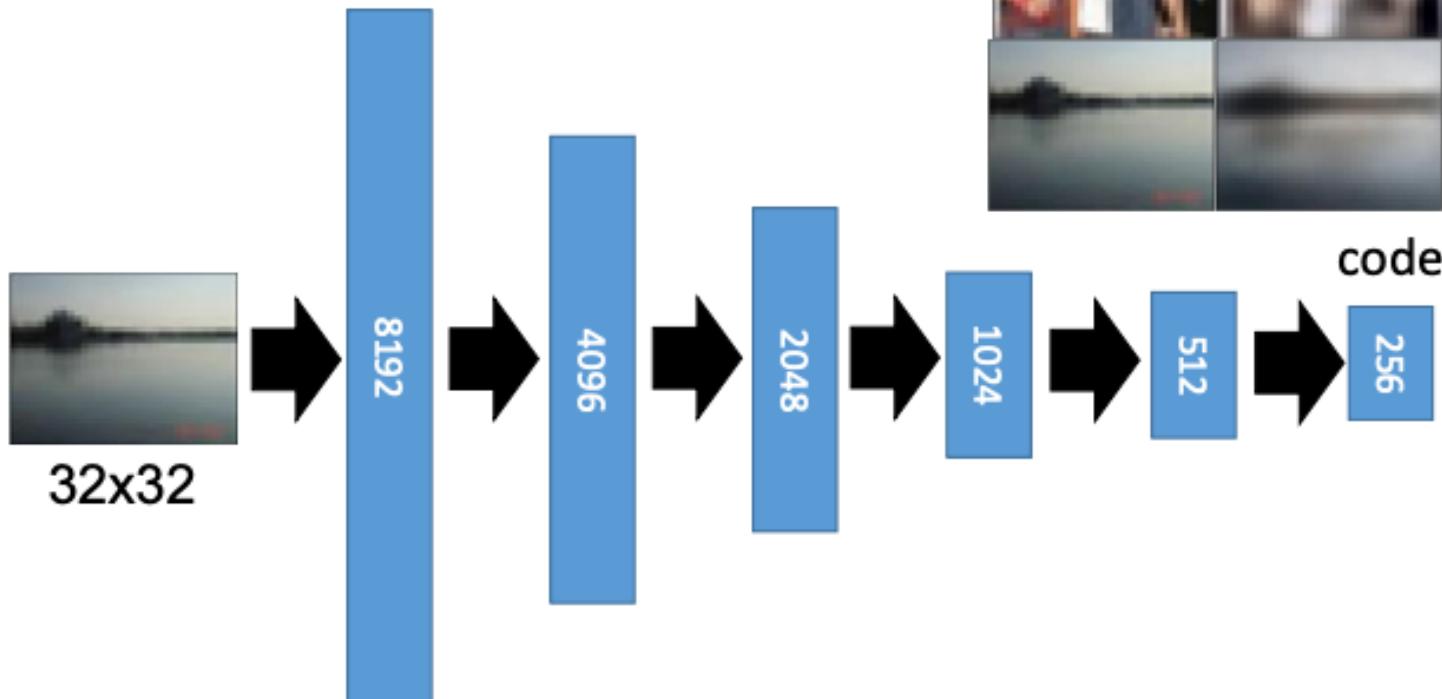
Given a query image, try to find visually similar images from an image database



Acknowledge: The slides are partially referred to the online materials by Taegyun Joen, <https://www.slideshare.net/TaegyunJeon1/pr12-you-only-look-once-yolo-unified-realtime-object-detection> and online YOLO paper and other materials (from ECS289g by Prof. Lee)

Image Search Engine Training

Auto-encoder –
Similar Image Search



Acknowledge:
[yolo-unified-realtime-object-detection](#) and online YOLO paper and other materials (from ECS289g by Prof. Lee)

[1ly-look-once-](#)

Retrieved using Euclidean distance in pixel intensity space

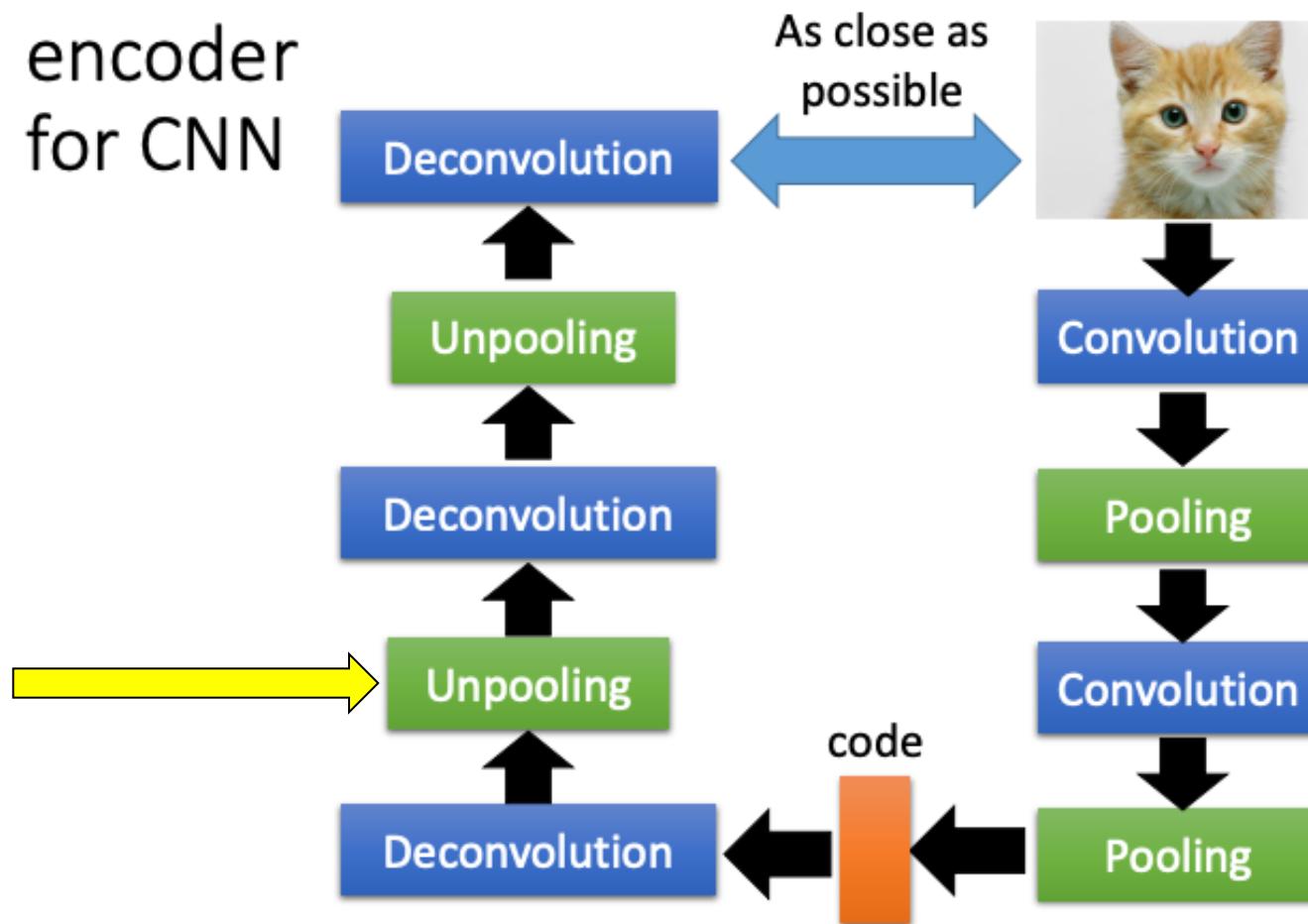


retrieved using 256 codes



CNN For Auto-Encoder

Auto-encoder for CNN



Unpooling

Nearest Neighbor

1	2
3	4



1	1	2	2
1	1	2	2
3	3	4	4
3	3	4	4

Output: 4 x 4

Input: 2 x 2

“Bed of Nails”

1	2
3	4

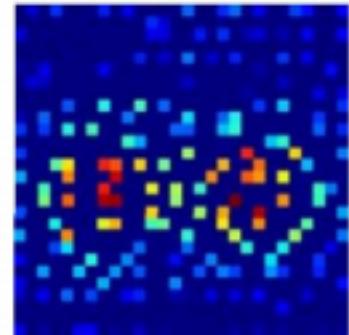
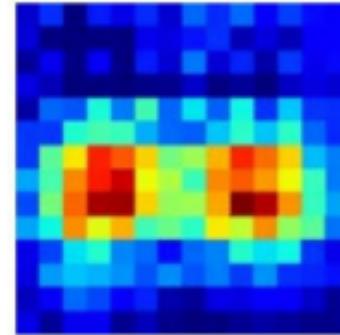


1	0	2	0
0	0	0	0
3	0	4	0
0	0	0	0

Output: 4 x 4

Input: 2 x 2

Unpooling Result



Max Unpooling

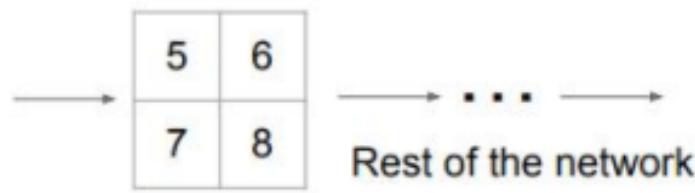
Max Pooling

Remember which element was max!

1	2	6	3
3	5	2	1
1	2	2	1
7	3	4	8

Input: 4 x 4

Output: 2 x 2



Max Unpooling

Use positions from pooling layer

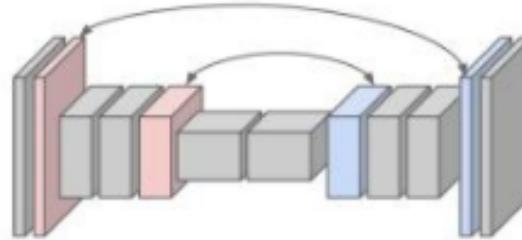
1	2
3	4

Input: 2 x 2

0	0	2	0
0	1	0	0
0	0	0	0
3	0	0	4

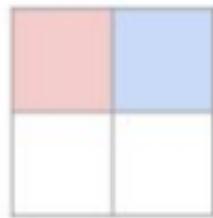
Output: 4 x 4

Corresponding pairs of
downsampling and
upsampling layers

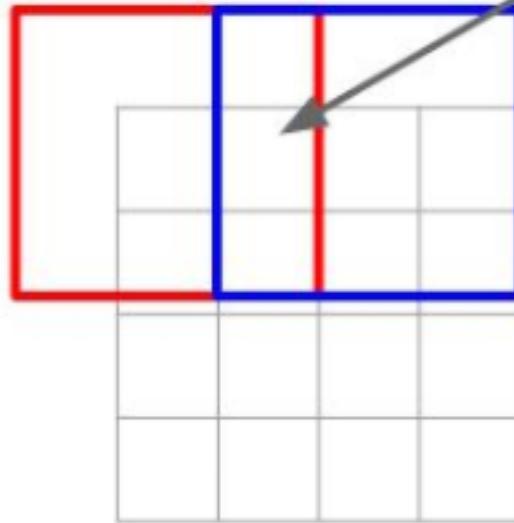


Transpose Convolution

3 x 3 **transpose** convolution, stride 2 pad 1



Input gives
weight for
filter



Sum where
output overlaps

Filter moves 2 pixels in
the output for every one
pixel in the input

Input: 2 x 2

Stride gives ratio between
movement in output and
input

Can we use Decoder to Generate Something?

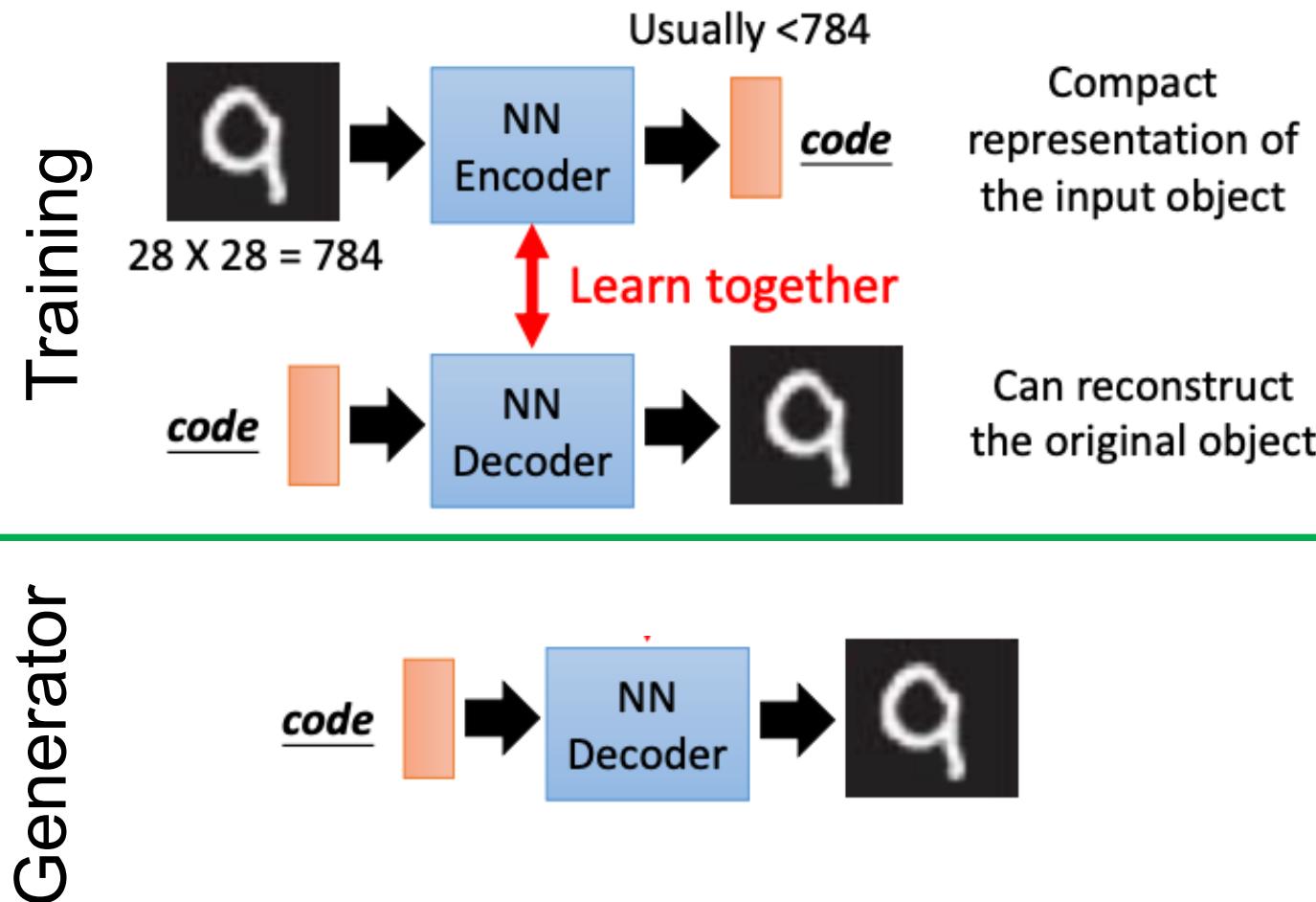


Image Generator

- Can we use decoder to generate something?

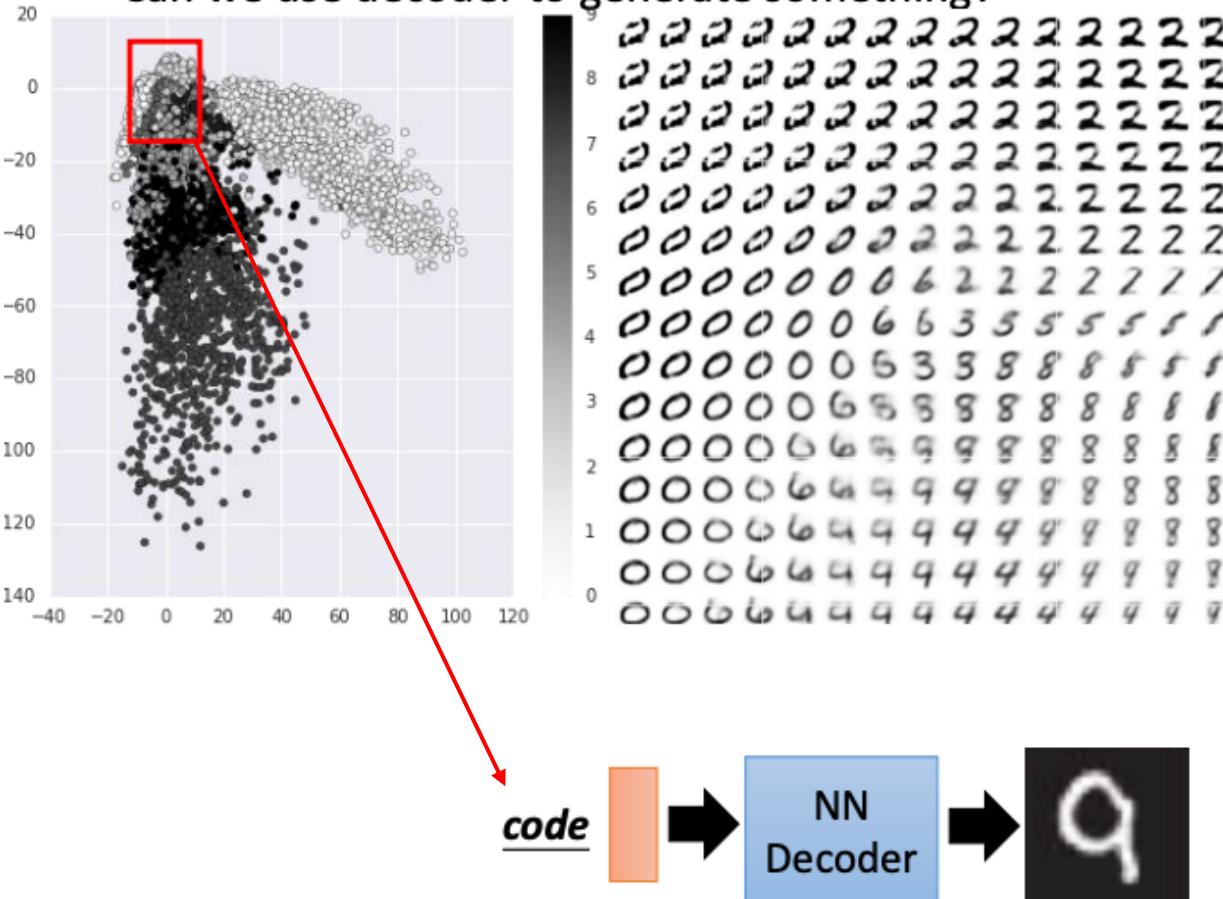


Image source: