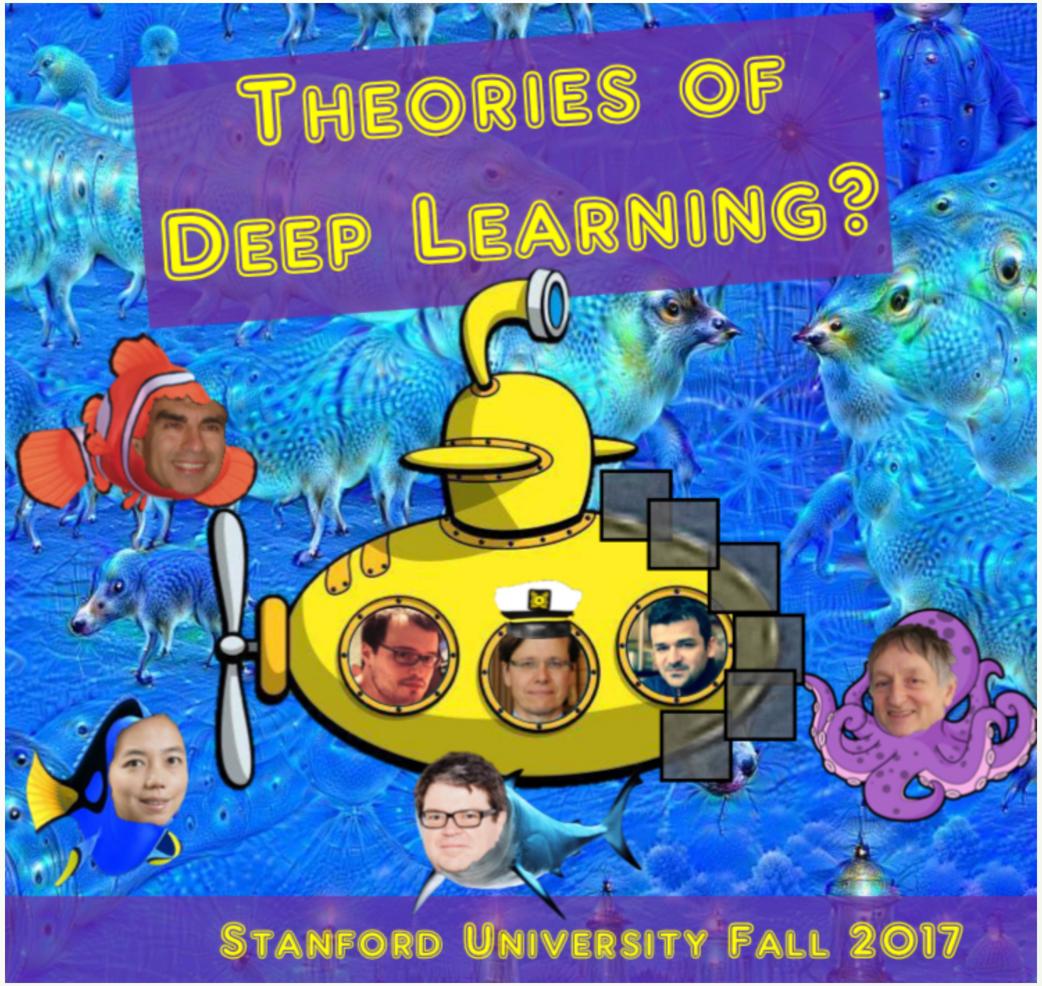
Deep Learning on Graphs and Manifolds

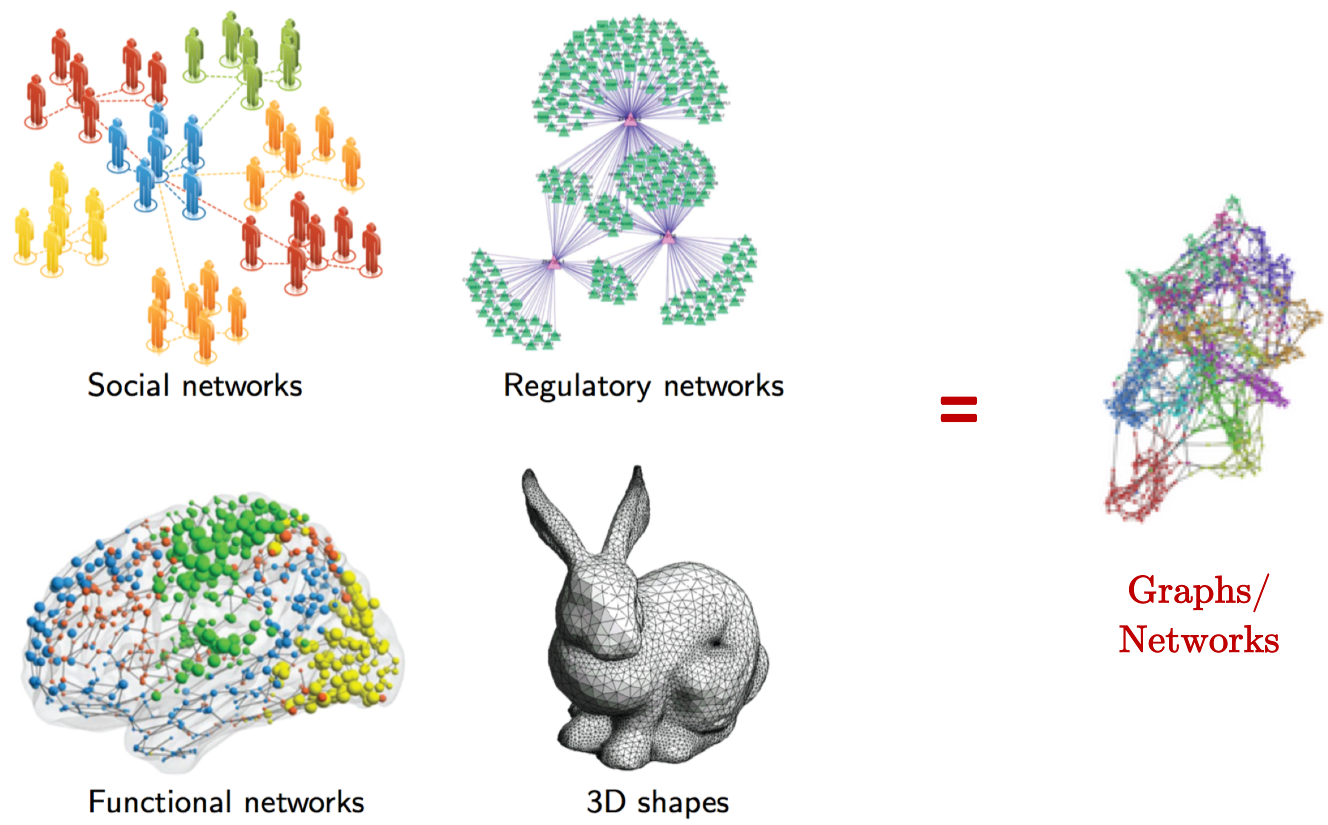
Yuan YAO
HKUST
Based on Xavier Bresson et al.

Acknowledgement



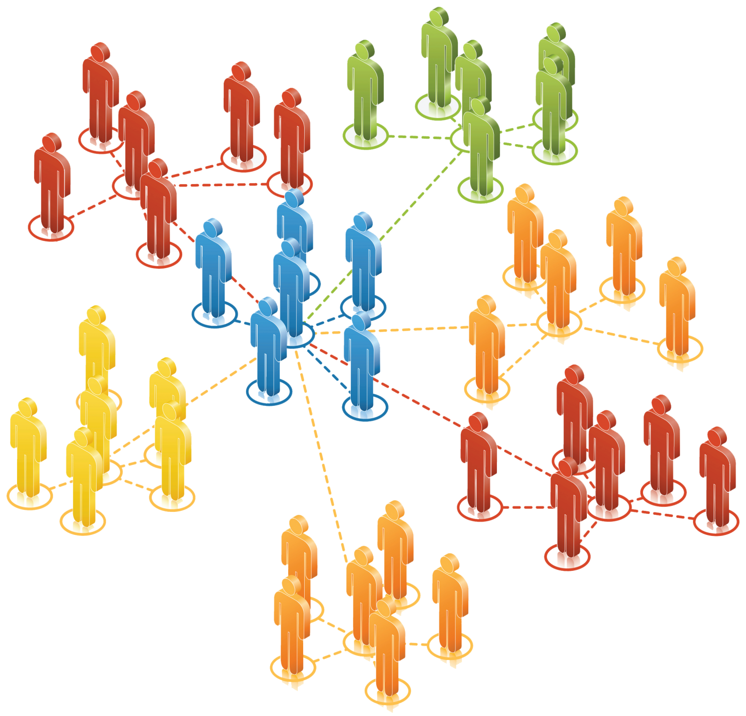
A following-up course at HKUST: https://deeplearning-math.github.io/

Non-Euclidean Data?



- Also chemistry, physics, social science, communication networks, etc.

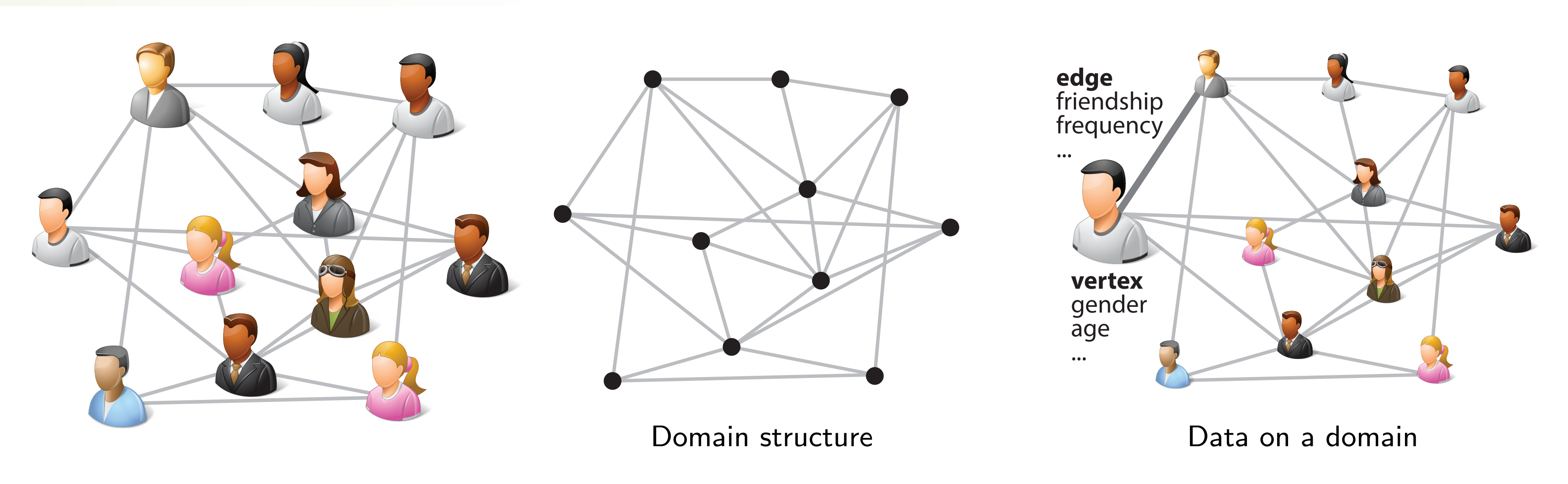
Graphs and Manifolds



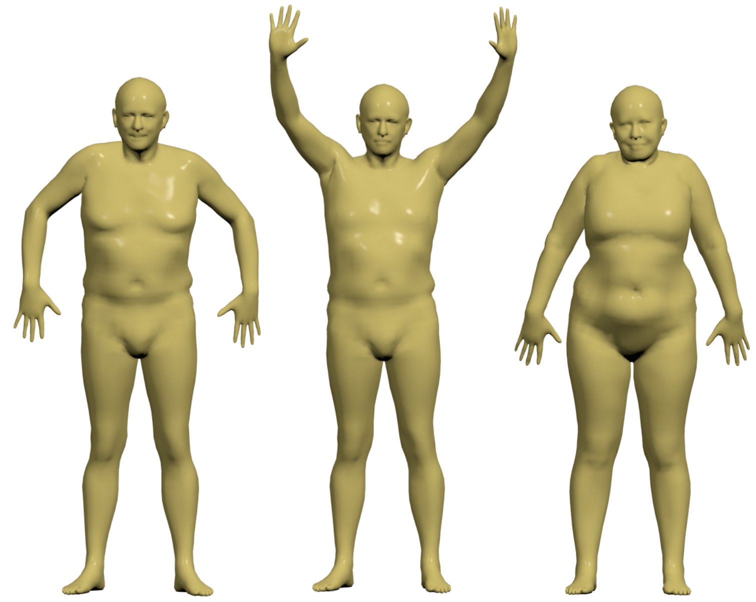
Graphs

Manifolds

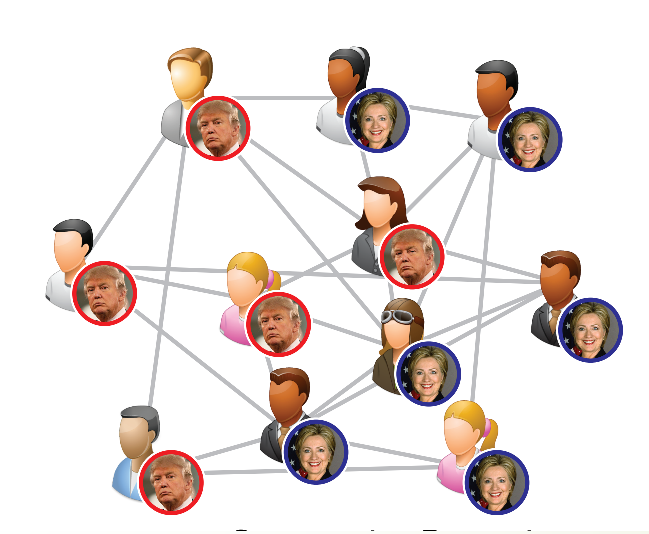
Social Networks as Graphs and Features on Edges and Vertices

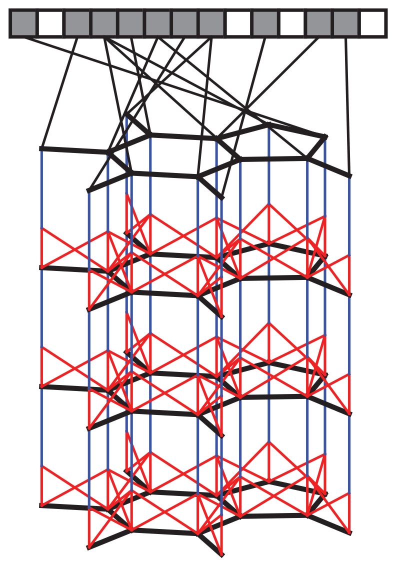


Graphs and Manifolds may vary



3D shapes
(different manifolds)





Molecule graph

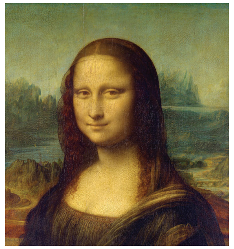
Challenges

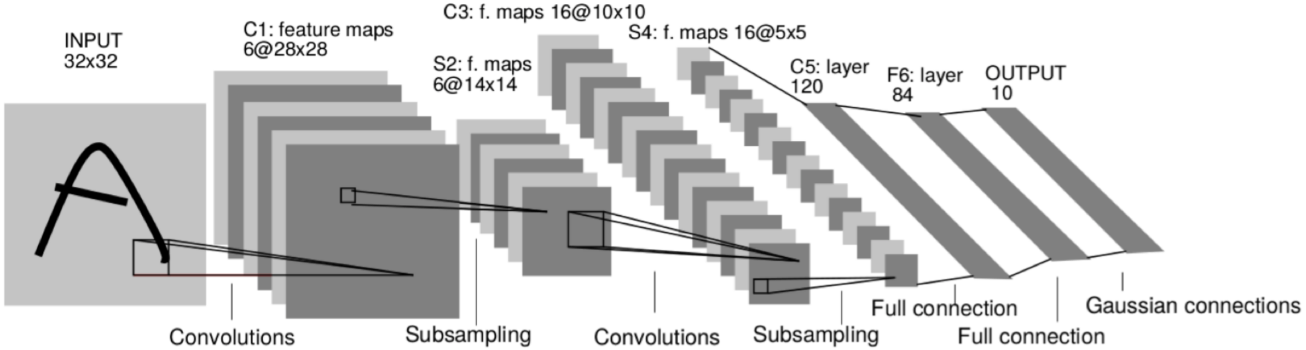
- ***What geometric structure in images, speech, video, text, is exploited by CNNs***?

- ***How to leverage such structure on non-Euclidean domains***?

Convolutional Networks on Euclidean Domain (e.g. LeNet for Images)

- An architecture for high-dimensional learning :



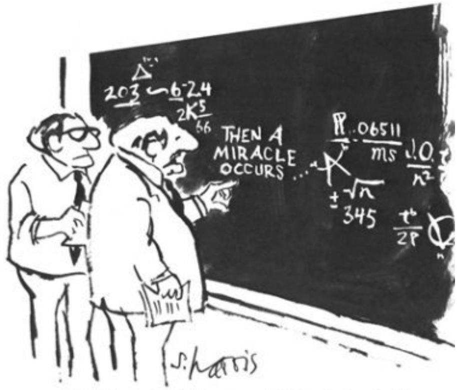


- Curse of dimensionality :

  dim(image) = 1024 x 1024 ≈ 10^6

  For N=10 samples/dim ⇒ $10^{1,000,000}$ points

- ConvNets are powerful to solve high-dimensional learning problems.

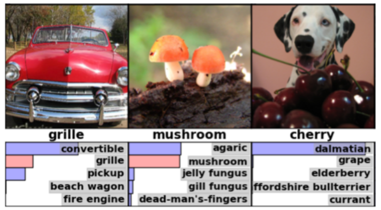


"I think you should be more explicit here in step two."

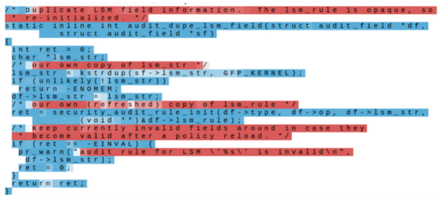
ConvNets on Euclidean Domains

- Main assumption :
  - Data (image, video, sound) is compositional, it is formed of patterns that are:
    - Local
    - Stationary
    - Multi-scale (hierarchical)

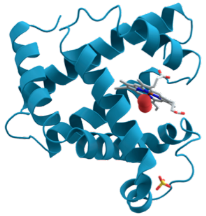
- ConvNets leverage the compositionality structure :
  - They extract compositional features and feed them to classifier, recommender, etc (end-to-end).



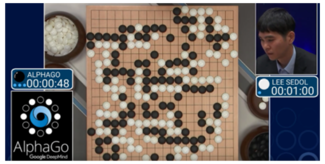
Computer Vision



NLP



Drug discovery

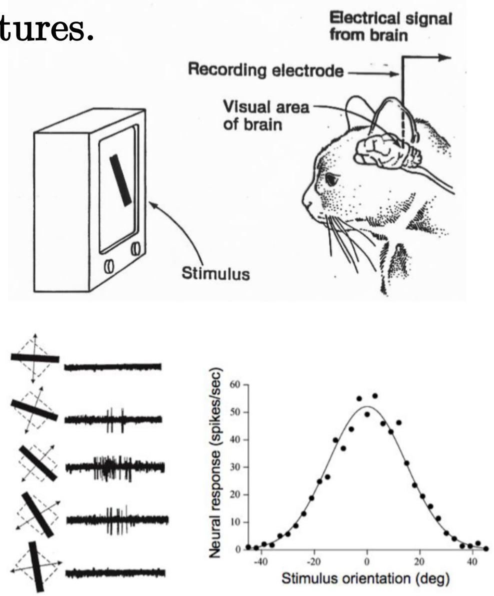


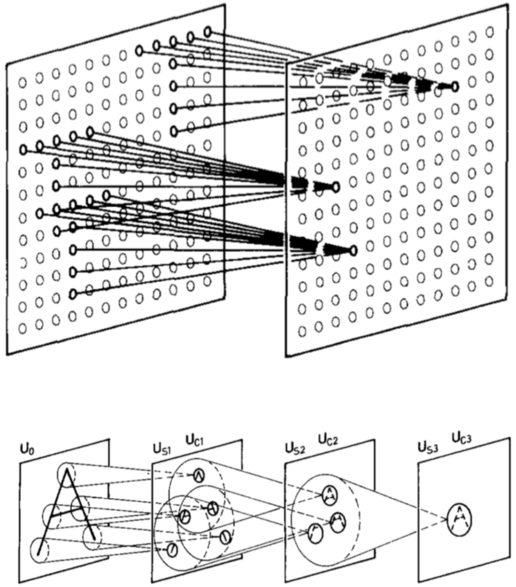
Games

Key Property: Locality

- Locality :
  - Property inspired by the human visual cortex system.

- Local receptive fields (Hubel, Wiesel 1962) :
  - Activate in the presence of local features.

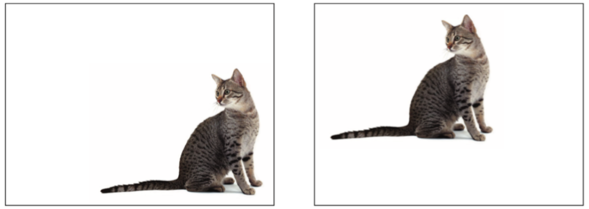




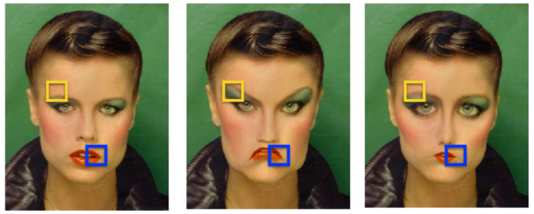
Neocognitron
Fukushima 1980

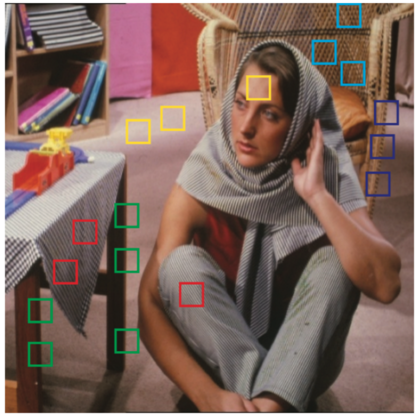
Key Property: Stationarity (Invariance)

- Stationarity ⇔ Translation invariance
  - Global invariance



- Local stationarity ⇔ Similar patches are shared across the data domain
  - Local invariance, essential for intra-class variations

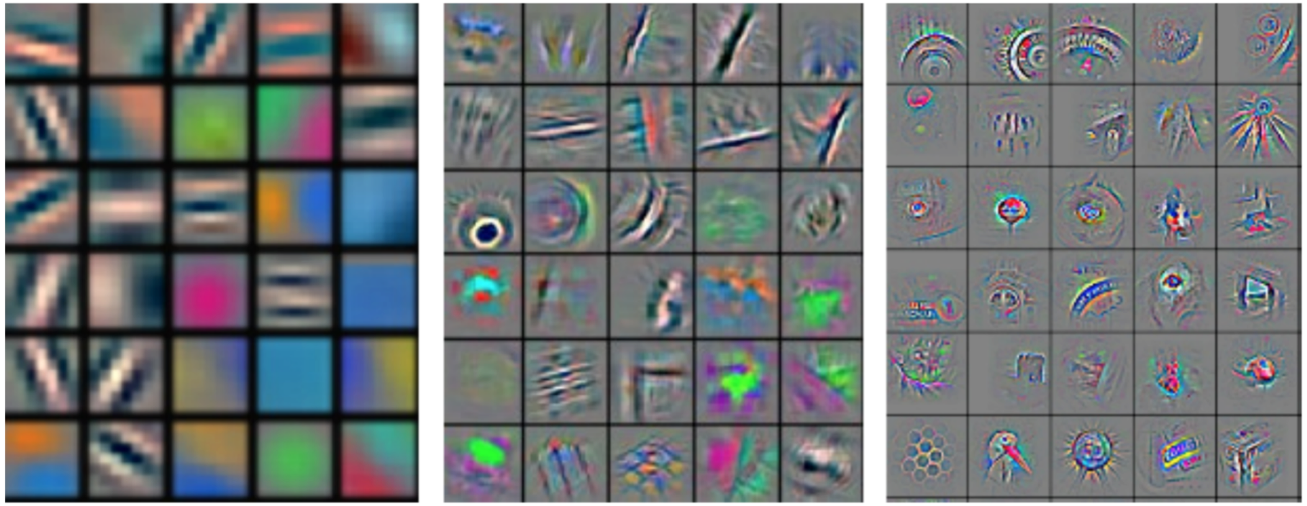




Key Property: Multiscale Representation

- Multi-scale :
  - Simple structures combine to compose slightly more abstract structures, and so on, in a hierarchical way.

- Inspired by brain visual primary cortex (V1 and V2 neurons).

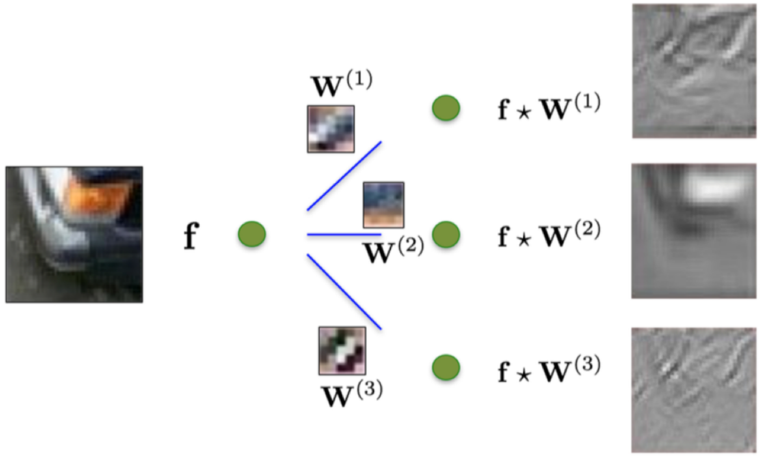


Features learned by ConvNet become increasingly more complex at deeper layers
(Zeiler, Fergus 2013)

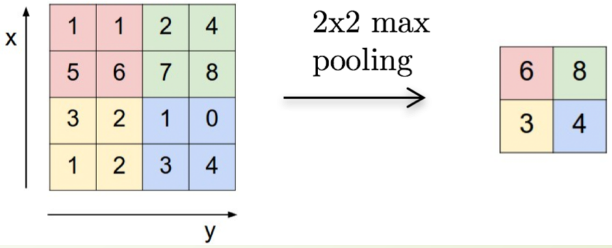
How to avoid the curse of dimensionality?

- Locality :
  - Compact support kernels $\Rightarrow O(1)$ parameters per filter.

- Stationarity :
  - Convolutional operators $\Rightarrow O(n\log n)$ in general (FFT) and $O(n)$ for compact kernels.

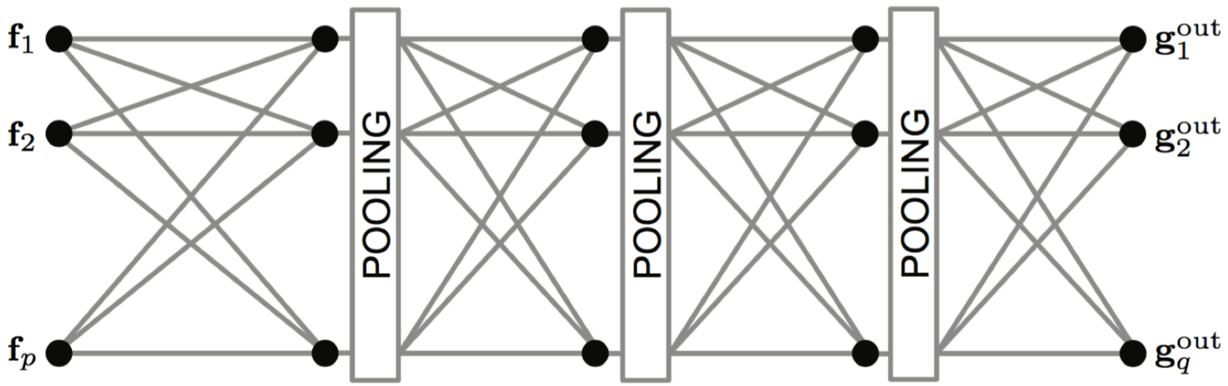


- Multi-scale :
  - Downsampling + pooling $\Rightarrow O(n)$



Implementation: Compositional Maps

$$\mathbf{f}_l = l\text{-th image feature (R,G,B channels)}, \dim(\mathbf{f}_l) = n \times 1$$
$$\mathbf{g}_l^{(k)} = l\text{-th feature map}, \dim(\mathbf{g}_l^{(k)}) = n_l^{(k)} \times 1$$



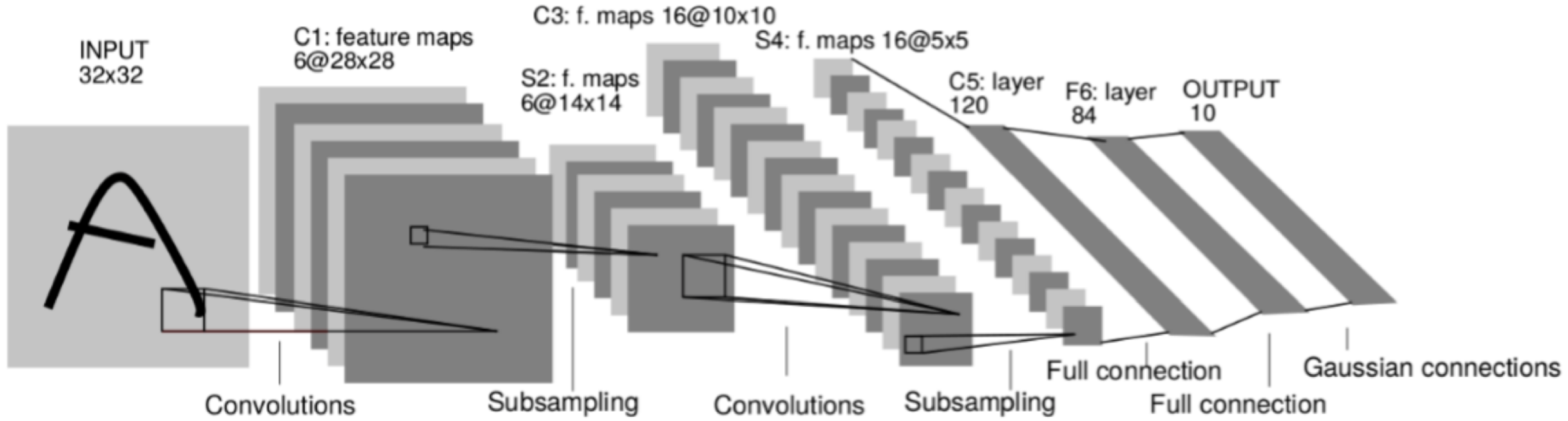
Compositional features consist of multiple convolutional + pooling layers.

Convolutional layer $\mathbf{g}_l^{(k)} = \xi\left(\sum_{l'=1}^{q_{k-1}} \mathbf{W}_{l,l'}^{(k)} \star \xi\left(\sum_{l'=1}^{q_{k-2}} \mathbf{W}_{l,l'}^{(k-1)} \star \xi\left(\cdots \mathbf{f}_{l'}\right)\right)\right)$

Activation, e.g. $\xi(x) = \max\{x, 0\}$ rectified linear unit (ReLU)

Pooling $\mathbf{g}_l^{(k)}(x) = \|\mathbf{g}_l^{(k-1)}(x') : x' \in \mathcal{N}(x)\|_p$ $p = 1, 2, \text{ or } \infty$

Summary of ConvNets



- ☺ Filters localized in space (locality)
- ☺ Convolutional filters (stationarity)
- ☺ Multiple layers (multi-scale)
- ☺ $O(1)$ parameters per filter (independent of input image size n)
- ☺ $O(n)$ complexity per layer (filtering done in the spatial domain)

Generalization to ConvNets on Graphs?

- How to extend ConvNets to graph-structured data?
- Assumption :
  - Non-Euclidean data is locally stationary and manifest hierarchical structures.
- How to define compositionality on graphs? (convolution and pooling on graphs)
- How to make them fast? (linear complexity)

Next:

- Prof. **Xavier Bresson**, NTU
  - IPAM talk on Convolutional Neural Networks on Graphs
  - https://www.youtube.com/watch?v=v3jZRkvIOIM
- Prof. **Zhizhen ZHAO**, UIUC
  - Seminar: Multi-Scale and Multi-Representation Learning on Graphs and Manifolds

Thank you!