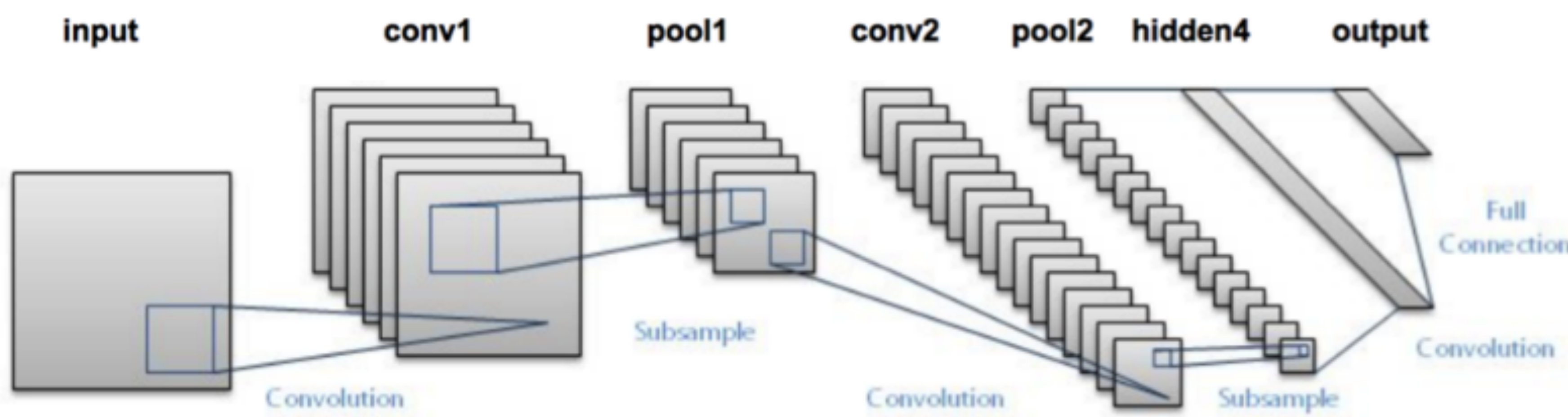
Intrinsic dimension and density profile of neural representations (2)

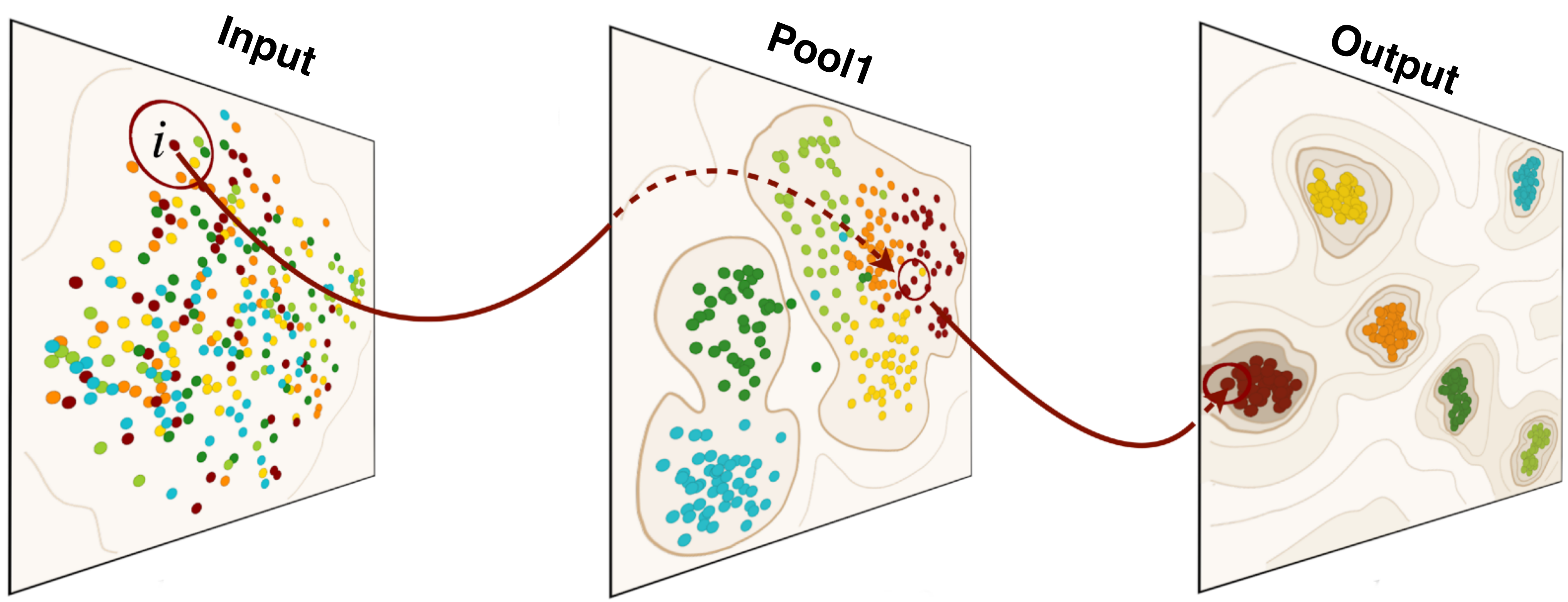
Diego Doimo and Aldo Glielmo
International School for Advanced Studies (SISSA), Trieste

SISSA
Scuola Internazionale Superiore di Studi Avanzati

Uni TS, 27 April 2021

Understanding deep networks 2)





1) **Intrinsic dimension** [Ansuini et al., NeurIPS 2019]

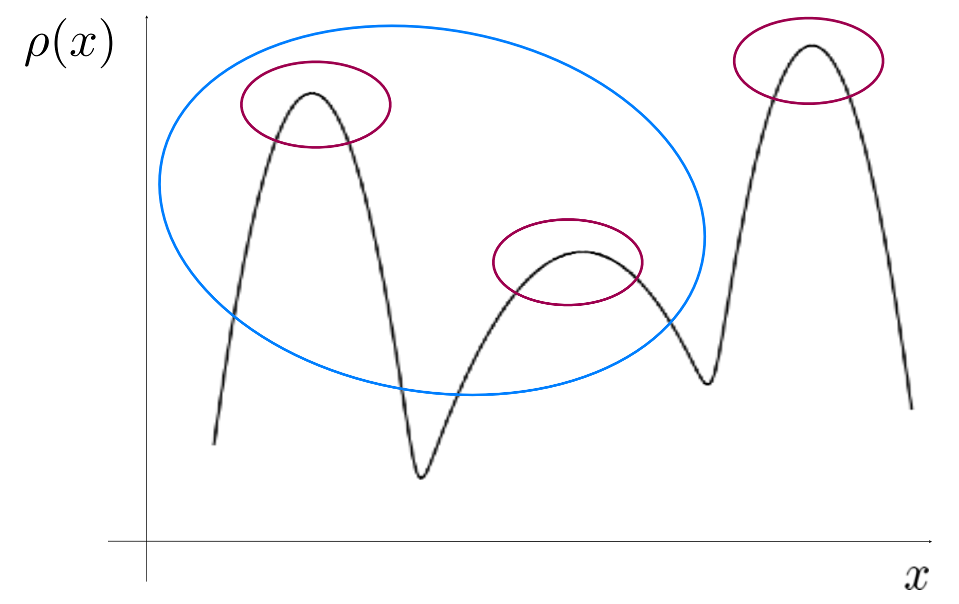
2) **Probability density** [Doimo et al., NeurIPS 2020]

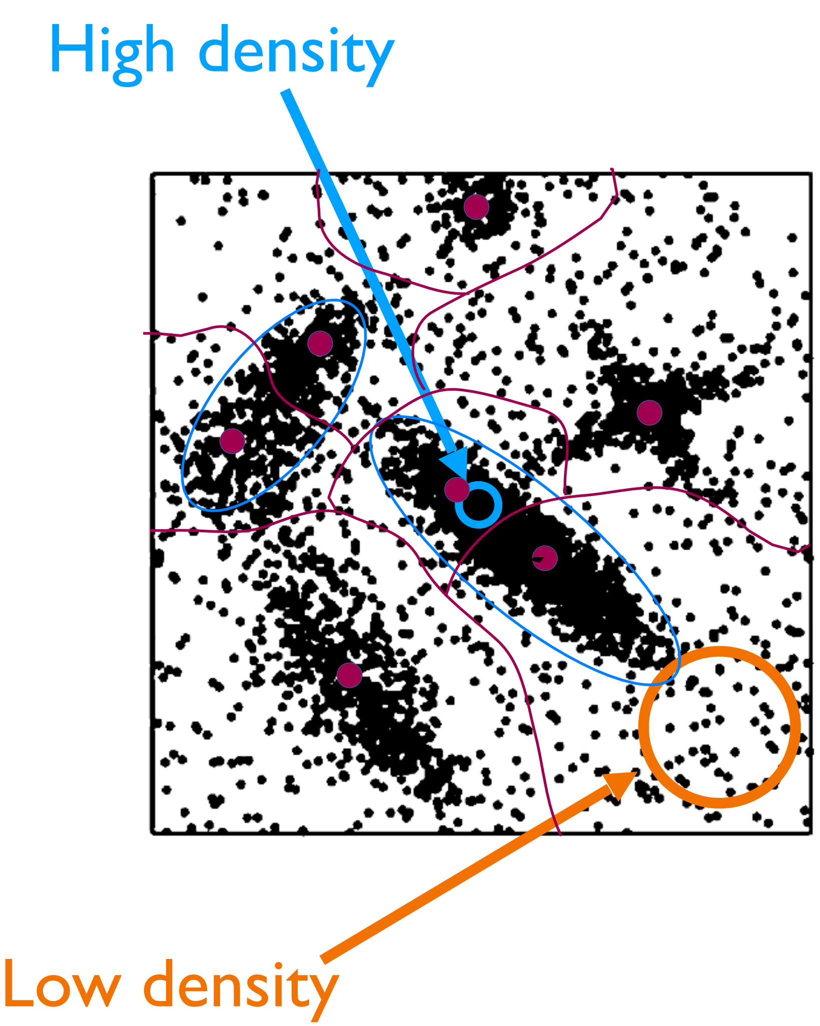
Density estimation: overview

Density estimation: kNN estimator

Density peaks estimation

- **Compute density** of each point (after computing ID)
- Find the **peaks** of the density
- Find the peak borders and **saddle points**
- **Merge peaks** according to a single hyper parameter Z

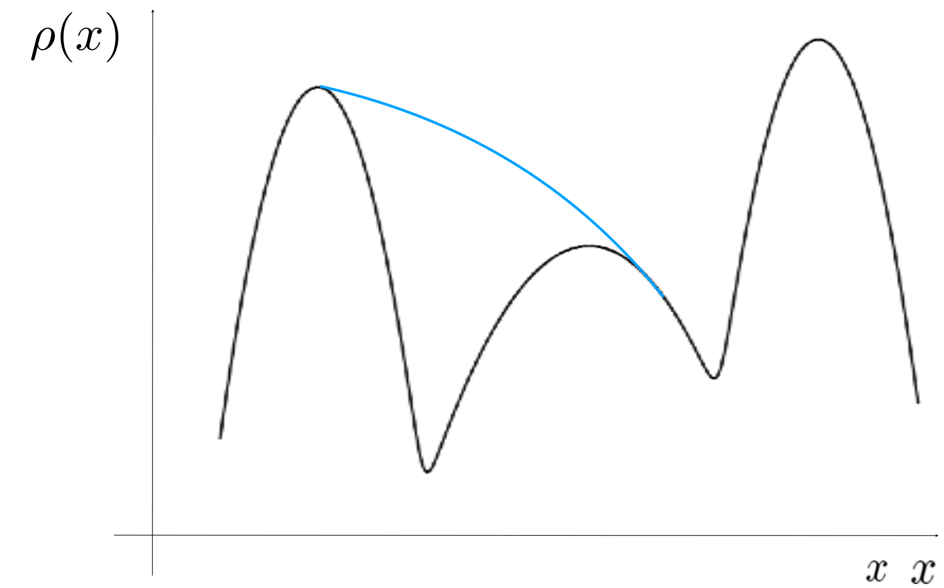


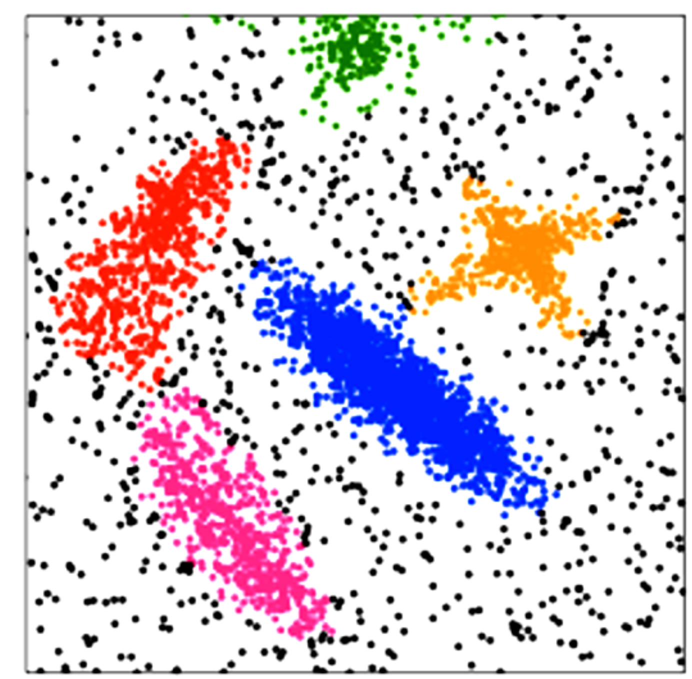


[Rodriguez and Laio, Science 2014]
[D'Errico et al., Information Sciences 2021]

Density peaks estimation

- **Compute density** of each point (after computing ID)
- Find the **peaks** of the density
- Find the peak borders and **saddle points**
- **Merge peaks** according to a single hyper parameter Z
- Assign points to the different peaks



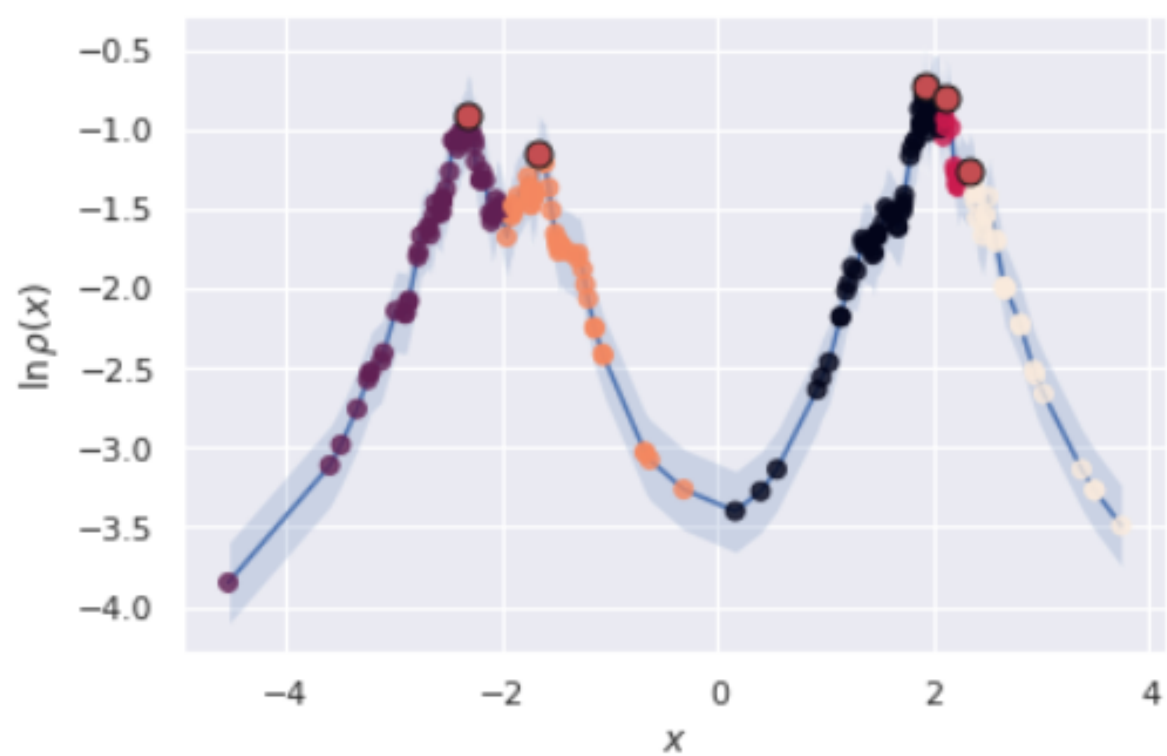


[Rodriguez and Laio, Science 2014]
[D'Errico et al., Information Sciences 2021]

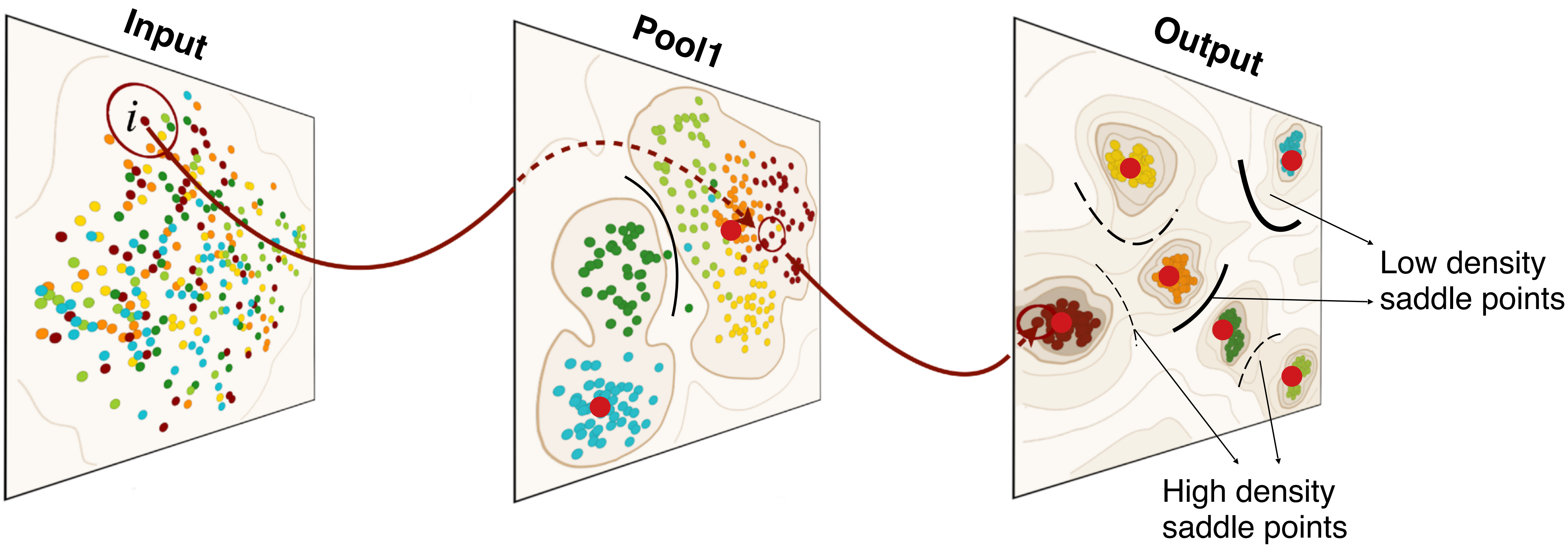
Tutorial1: Density peaks estimation

- Concepts that we will cover:
  - Estimation of the density profile
    → **kNN estimator**
  - Estimation of statistically significant peaks of the density
    → **Clustering**

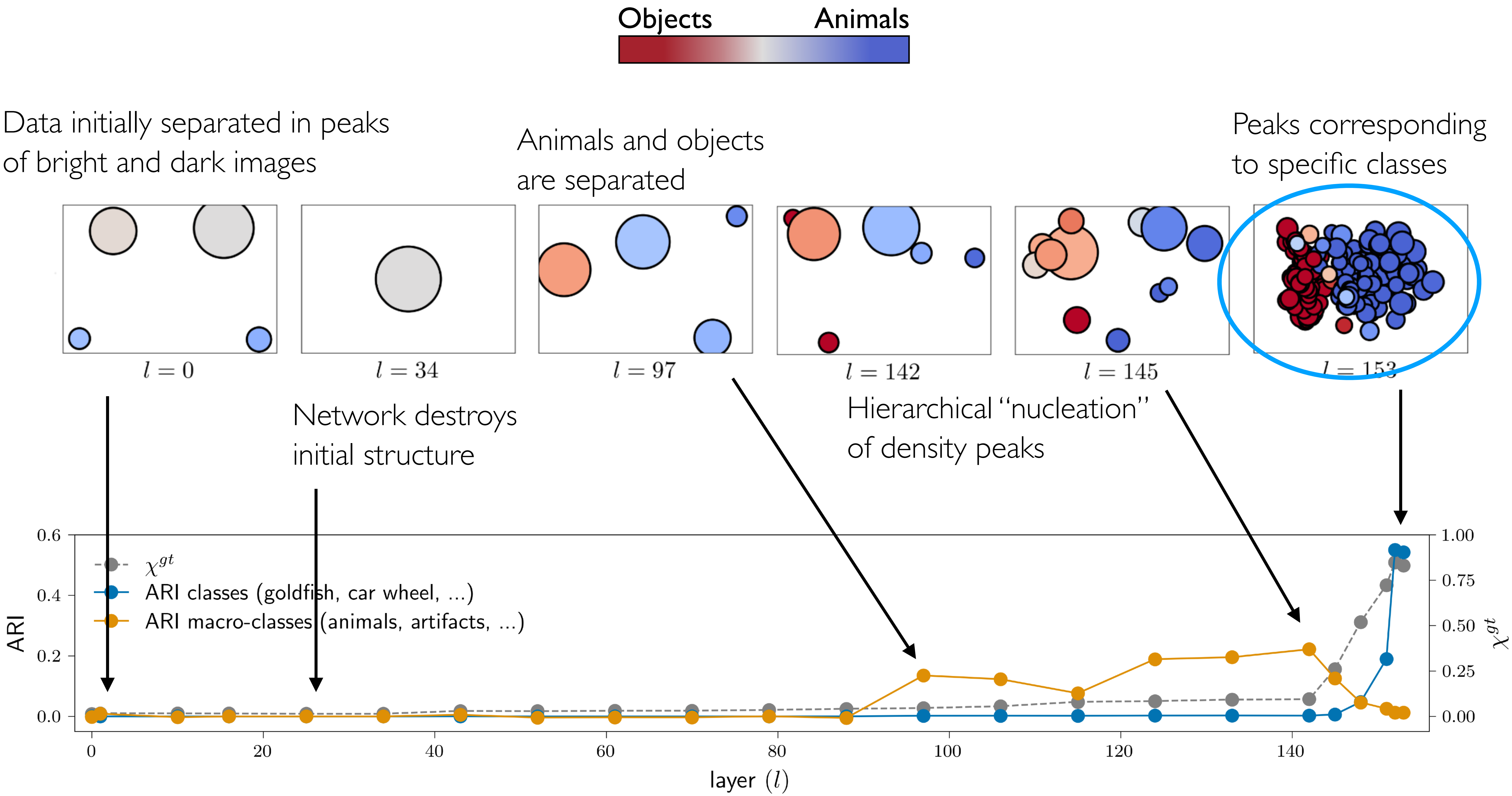
https://colab.research.google.com/drive/1fTxE0GWb5BobZhL3j6G6Ra5hBj__c9X-?usp=sharing



Density peaks estimation in deep networks



Evolution of density peaks

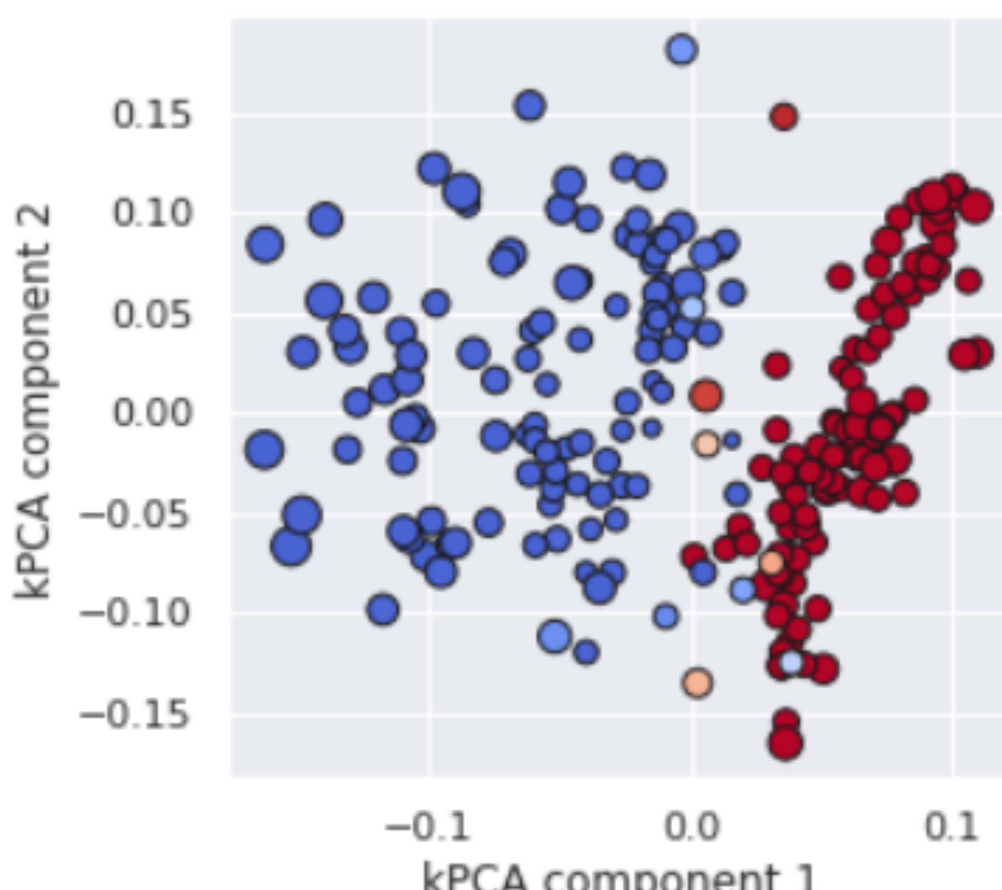


- Concepts are learned in **hierarchical order** through at least two transitions

Tutorial2: Density peaks in deep CNNs

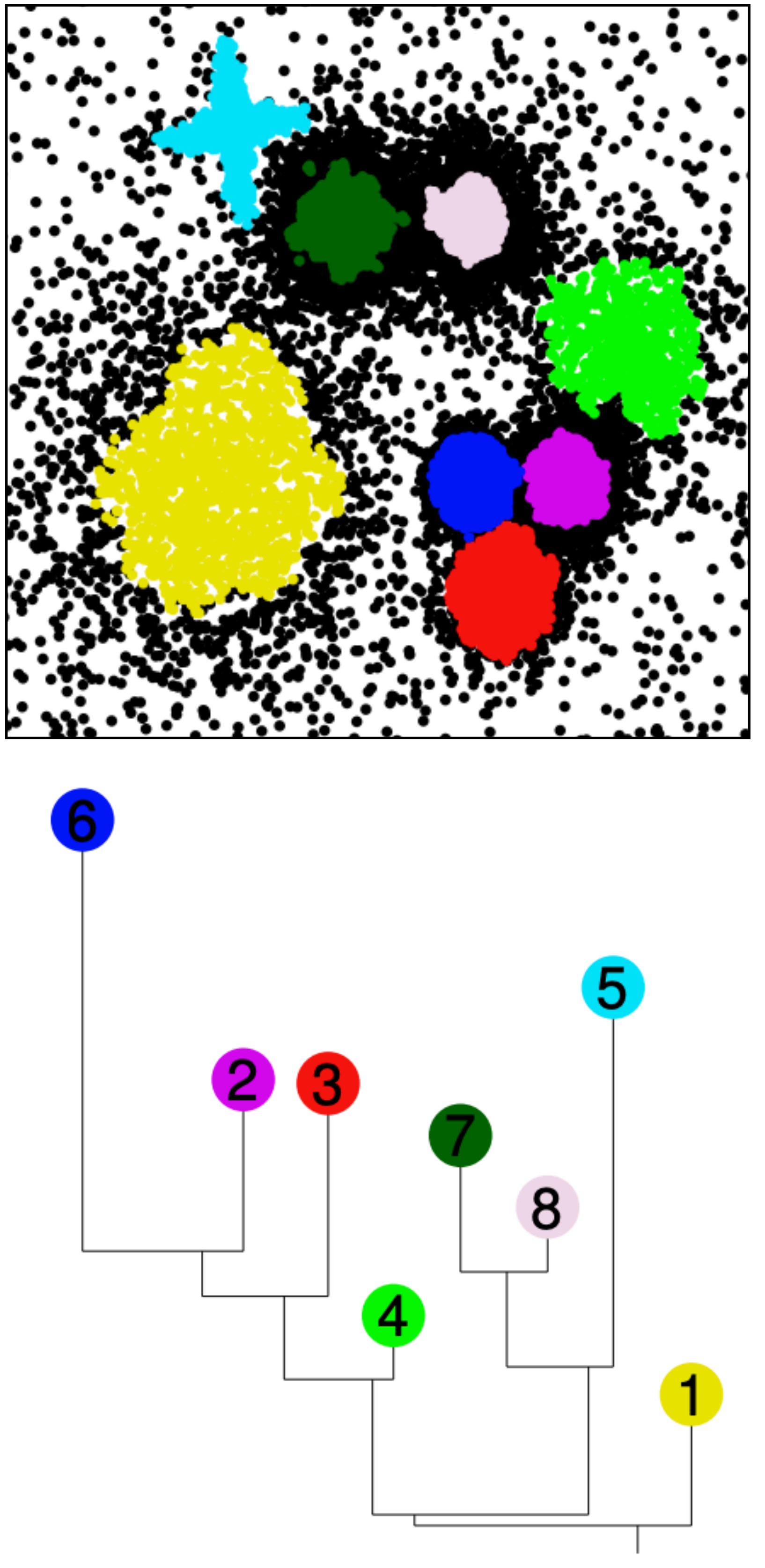
- What we will cover:
  - **"forgetting" and "learning"** as separate phases in deep CNNs
  - **High level concepts** are learned in early layers
  - **Low level concepts** are learned in a sharp "nucleation" transition near the end of the network

https://colab.research.google.com/drive/1fTxE0GWb5BobZhL3j6G6Ra5hBj__c9X-?usp=sharing



Hierarchical agglomeration

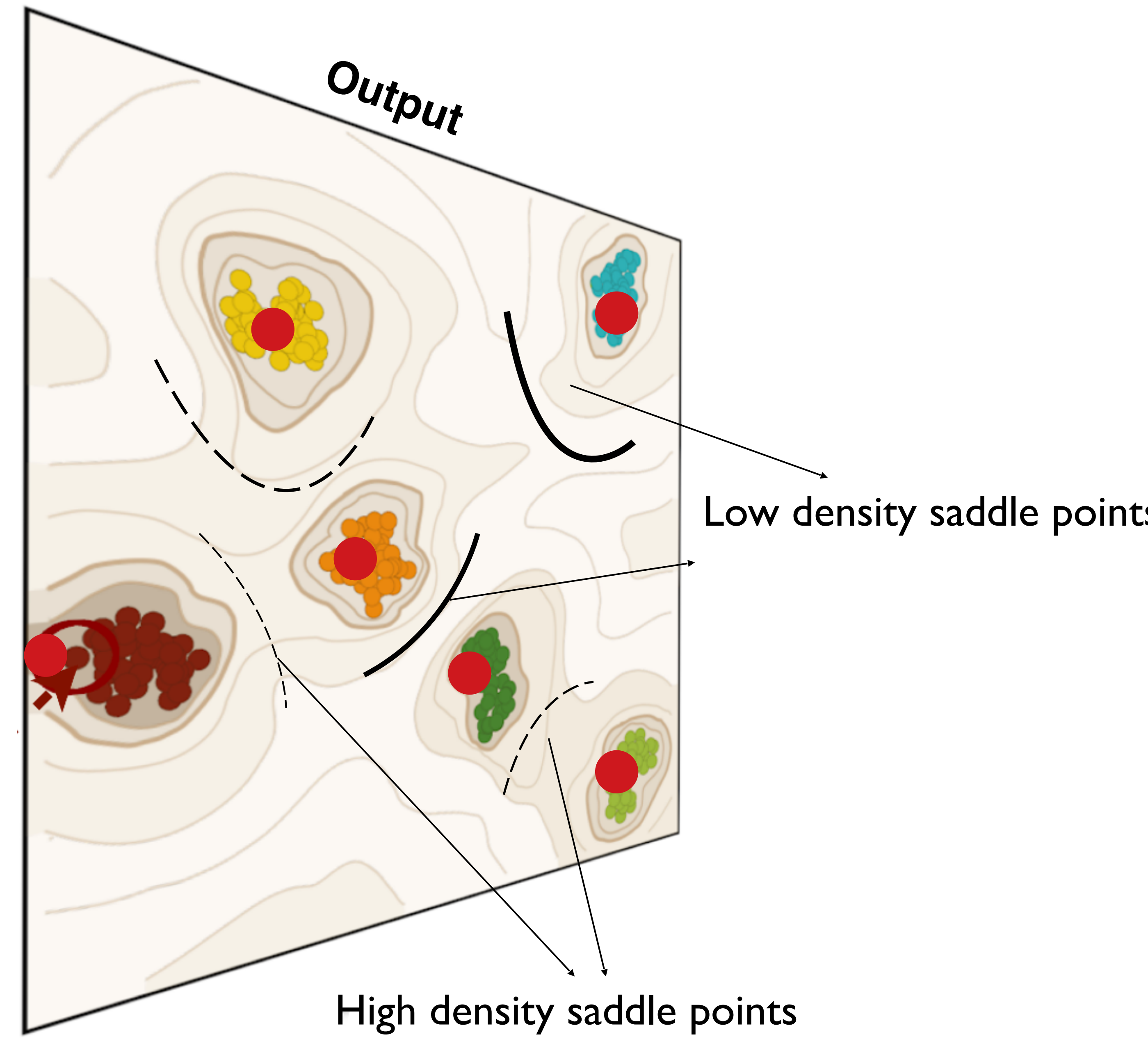
- Take highest density peak
- Merge it with the peak separated by the highest saddle point
- Repeat until there is only one peak left



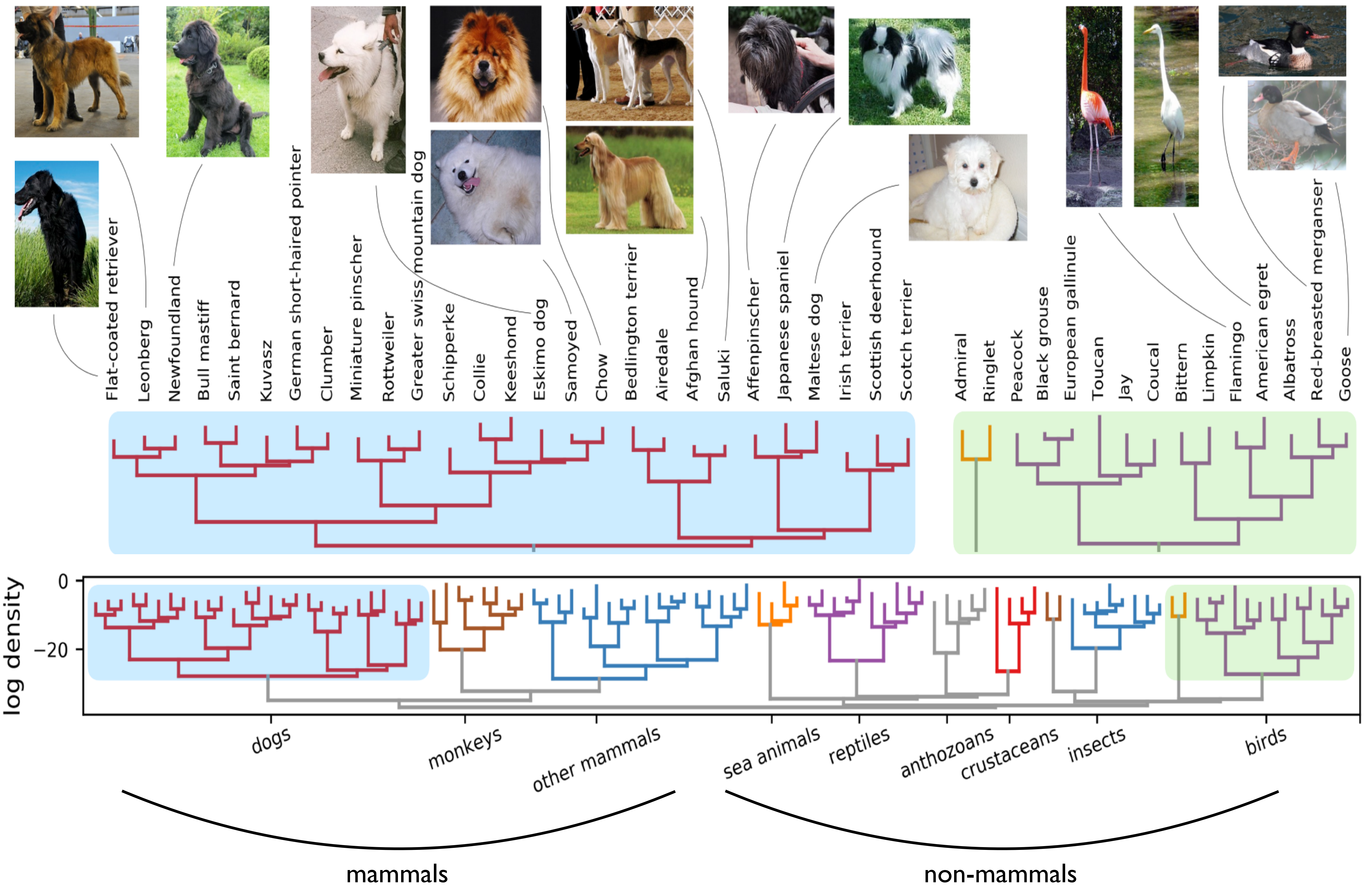
[D'Errico et al., Information Sciences 2021]

Hierarchical structure in the output layer

- Take highest density peak
- Merge it with the peak separated by the highest saddle point
- Repeat until there is only one peak left

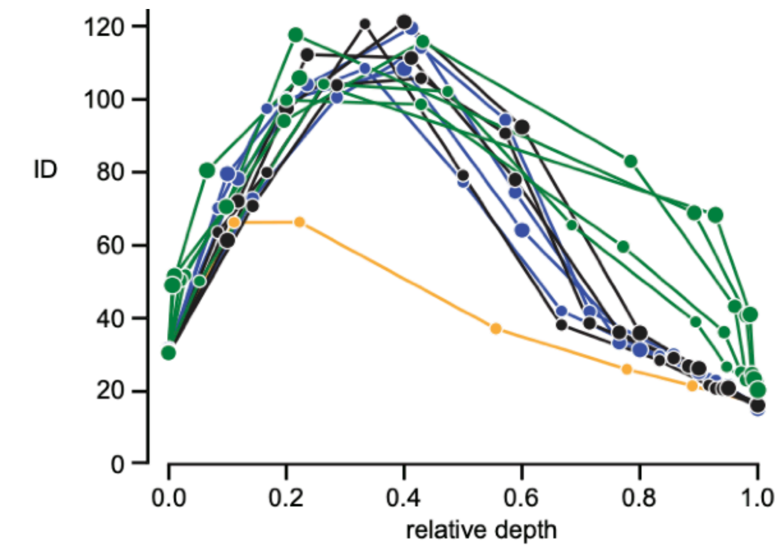


Semantic hierarchy in the last layer

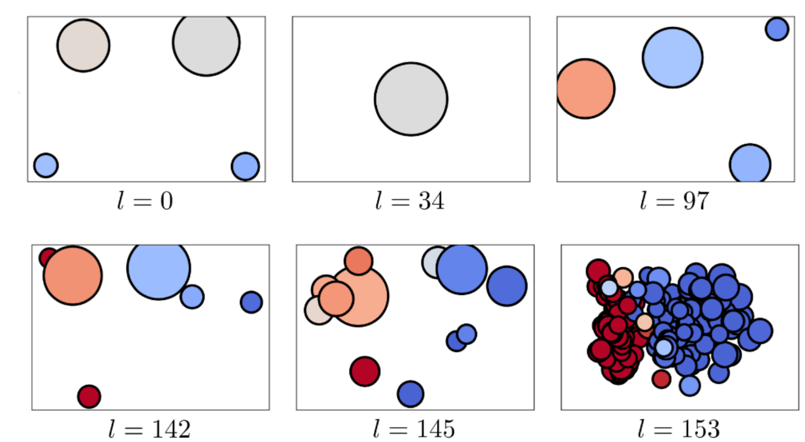


Take home messages

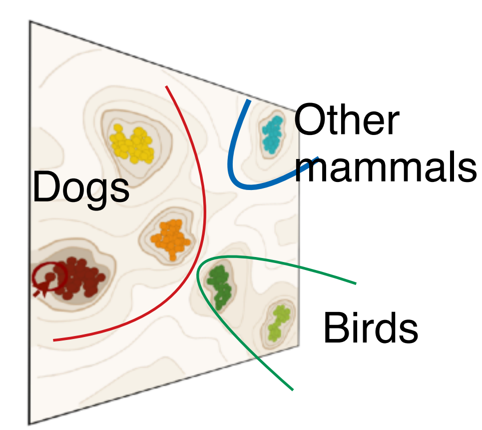
Intrinsic dimension of hidden representations is always small and shows a characteristic **hunchback profile** which can be interpreted as characteristic of a "forgetting" and a "learning" phase



Representations in neural networks trained for complex classification tasks evolve through **nucleation-like events**, with probability peaks appear in an order that mirrors the semantic **hierarchy** of the dataset



The peaks of the last layer are organised in **complex "mountain chains"** resembling the semantic kinship of the categories



Interesting research directions (extra)

Practical

- Improve triplet loss schemes via a principled choice of "anchor" images (e.g., the saddle points)
- Improve transfer learning schemes by educated guesses on network cuts
- Inform pruning schemes through the knowledge of "less important" layers

Theoretical/Exploratory

- Analyse the evolution of density profile during training
- Analyse the representations learned by different types of networks (not CNNs), for instance those used for Natural Language Processing

Neuroscience

- Use the analysis of the density profile as a tool to connect artificial networks with biological networks

Essential references:

- **ID**: *Estimating the intrinsic dimension of datasets by a minimal neighborhood information*
- **ID in CNNs**: *Intrinsic dimension of data representations in deep neural networks*
- **Clustering:** *Automatic topography of high-dimensional data sets by non-parametric Density Peak clustering*
- **Clustering in CNNs**: *Hierarchical nucleation in deep neural networks*

Code references:

- **GitHub** https://github.com/diegodoimo/hierarchical_nucleation GitHub
- **Supplementary Materials** of "*Hierarchical nucleation in deep neural networks*", NeurIPS 2020
- **Colab notebook** https://colab.research.google.com/drive/1fTxE0GWb5BobZhL3j6G6Ra5hBj__c9X-?usp=sharing
- **DULY Package** https://duly.readthedocs.io/en/latest/index.html

If you have other questions or ideas:

- aglielmo@sissa.it
- ddoimo@sissa.it