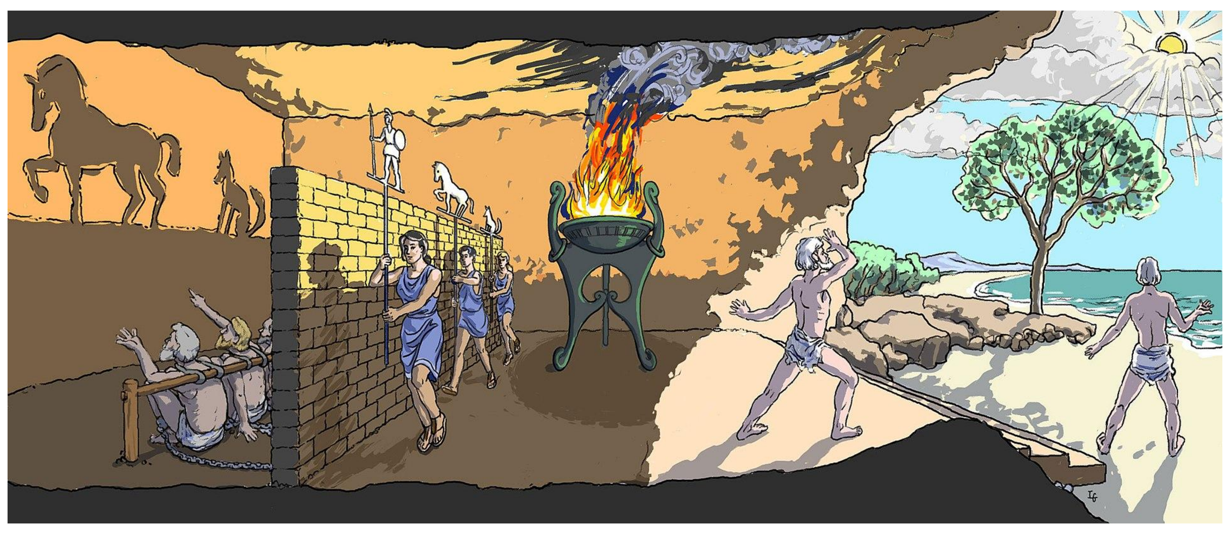
ScuDo
Scuola di Dottorato – Doctoral School
WHAT YOU ARE, TAKES YOU FAR
XXXVI Cycle

Representation Learning: from inverse problems to classification

Antonio Montanaro
Supervisor: Prof. Enrico Magli

Research context and motivation

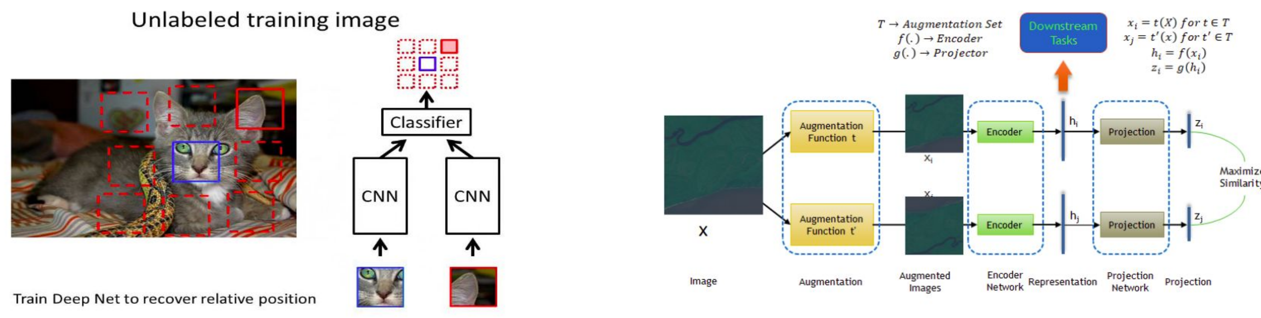
- The Plato's Cave is reminiscent of what machine learning systems, and specifically representation learning, aim to discover, that is understanding the higher levels of reality.
- There are prisoners that watch shadows on a wall projected by more complex objects in front of a fire. For them this is the unique reality, and only the reason can let them to go beyond this. The abstraction of the observations was the primary aim of philosophers in the Ancient Greek.



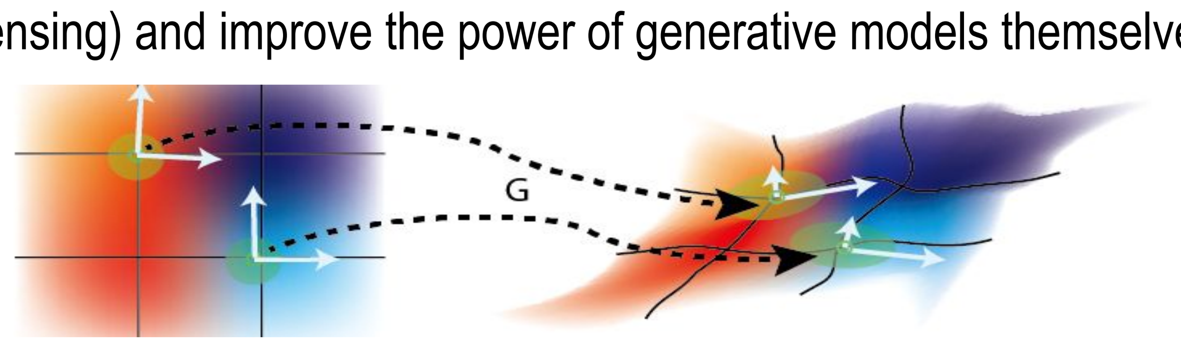
- Nowaday, machine learning systems are addressing the same kind of problems, but they are inspired by math, probability, geometry and biology. Representation learning allows a system to extract abstract features from observed data like images. Knowing high-level features is fundamental to make them able to perform accurate classification, generate new data and solve inverse problems.

Addressed research questions/problems

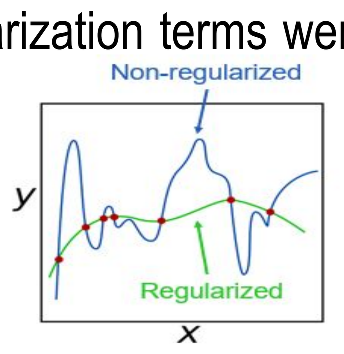
- There are different ways in which representation learning can help deep neural networks to better represent data. Possible solutions include:
  - Self-Supervised Learning
  - Exploring the geometry of the latent space
  - Regularization of the feature space
- **Self-Supervised Learning** is a new paradigm where the system is able to learn the features of the data by itself without label annotations. This is important in the majority of computer science tasks, where there are million of data often without a label and we want to make the system performing as in the supervised learning, where all the labels are available and the training procedure is the simplest to learn.



- **Exploring the geometry of the latent space** has become a interesting topic especially in generative models, where understanding how the latent space behaves is crucial for many applications like image editing, solving inverse problems (e.g. inpainting or compressive sensing) and improve the power of generative models themselves.



- **Regularization** is an old technique yet applied in the simplest regression analysis, i.e. the least squares. It induces a prior distribution on model parameters enabling to prevent overfitting and find more consistent solutions. If in the past the regularization terms were very simple, like minimizing the L_2 and L_1 weights' norm, today in deep learning more sophisticated regularizations has been proposed, both explicit or implicit, like early stopping and contrastive learning.



Submitted and published works

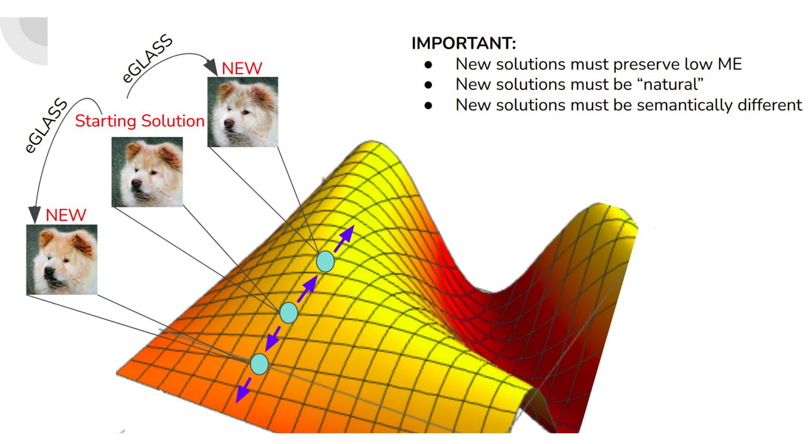
- A. Montanaro, D. Valsesia, G. Fracastoro and E. Magli, "Semi-Supervised Learning for Joint SAR and Multispectral Land Cover Classification," in *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1-5, 2022.
- A. Montanaro, D. Valsesia, E. Magli, "Exploring the solution space of linear inverse problems with GAN latent geometry", ICIP 2022
- A. Montanaro, D. Valsesia, E. Magli, "Exploring the solution space of linear inverse problems with GAN latent geometry", NIPS 2022

Novel contributions

- **SSCL** (Spatial-Spectral Context Learning) promotes bridging the gaps between sensing modalities and exploiting the spectral characteristics of the remote sensing images. Using the proposed self-supervised pretraining, followed by finetuning for land cover classification with SAR and multispectral data, it outperforms conventional approaches.
- **eGLASS** (explore Gan Latent Space Solution) is a new method able to find multiple solutions of linear inverse problems starting from an initial solution, where all the solutions are in the latent space of a GAN. We use GAN inversion to find one solution, then we use eGLASS to find directions corresponding to new solutions that are semantically different but that respects the measurement error.
- **HyCoRe** (Hyperbolic Compositional Regularization) is a new regularization aiming to represent the part-whole hierarchy in 3D objects. This is achieved only if the space can embed the tree-likeness structure, and it turns out that the hyperbolic space is the most adaptive (lowest distortion error). Reasoning about part-whole hierarchy allows different models to better classify the whole objects, other than having a more explainable feature space.

Adopted methodologies

- **eGLASS**



Algorithm 1 e-GLASS: exploring GAN LAtent Space Solutions

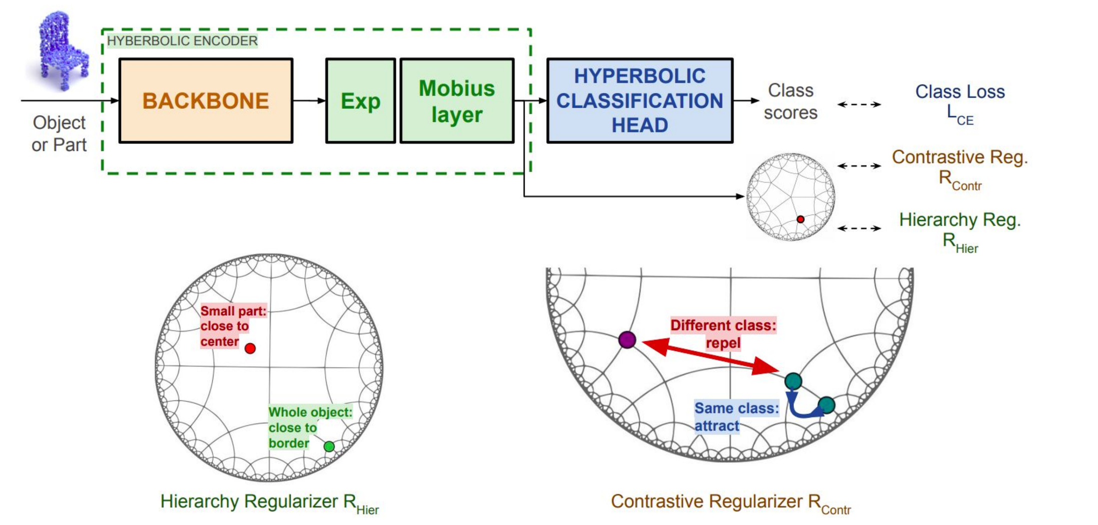
Input: $\mathbf{z}_0, K$
Output: New solution $\hat{\mathbf{x}}$
Compute Hessian $\mathbf{H}_{\mathcal{Y}}(\mathbf{z}_0) = \frac{1}{2}\frac{\partial^2}{\partial \mathbf{z}^2}\|\mathbf{A}G(\mathbf{z}) - \mathbf{A}G(\mathbf{z}_0)\|_2^2$
Compute Hessian $\mathbf{H}_{\mathcal{X}}(\mathbf{z}_0) = \frac{1}{2}\frac{\partial^2}{\partial \mathbf{z}^2}\mathrm{LPIPS}(G(\mathbf{z}), G(\mathbf{z}_0))$
Compute eigenvectors $\mathbf{H}_{\mathcal{Y}}(\mathbf{z}_0) = \mathbf{U}\Lambda\mathbf{U}^T$
Compute eigenvectors $\mathbf{H}_{\mathcal{X}}(\mathbf{z}_0) = \mathbf{V}\Omega\mathbf{V}^T$
$\mathbf{d} \leftarrow \mathbf{v}^K$
$\mathcal{J} = \{$top-k eigenvectors of $\mathbf{H}_{\mathcal{Y}}\}$
for $j \in \mathcal{J}$ **do**
  $\mathbf{d} \leftarrow \mathbf{d} - (\mathbf{d}^T\mathbf{u}^j)\mathbf{u}^j$
  $\mathbf{d} \leftarrow \mathbf{d}/\|\mathbf{d}\|$
end for
$\hat{\mathbf{x}} = G(\mathbf{z}_0 + \eta\mathbf{d})$

- **HyCoRe**

$$R_{\text{hier}}(\mathbf{z}^+_{\text{whole}}, \mathbf{z}^+_{\text{part}}) = \max(0, -\|\mathbf{z}^+_{\text{whole}}\|_{\mathbb{D}} + \|\mathbf{z}^+_{\text{part}}\|_{\mathbb{D}} + \gamma/N')$$

$$R_{\text{contr}}(\mathbf{z}^+_{\text{whole}}, \mathbf{z}^+_{\text{part}}, \mathbf{z}^-_{\text{part}}) = \max\left(0, d_{\mathbb{D}}(\mathbf{z}^+_{\text{whole}}, \mathbf{z}^+_{\text{part}}) - d_{\mathbb{D}}(\mathbf{z}^+_{\text{whole}}, \mathbf{z}^-_{\text{part}}) + \delta\right)$$

$$L = L_{\text{CE}} + \alpha R_{\text{contr}} + \beta R_{\text{hier}}$$



Future work

- **HyCoRe** for part segmentation. The generalization of this regularization to part and instance segmentation should be addressed, although is not obvious, since parts should be pushed on the Poincarè edge leading to a reversed hierarchy.
- New Representation Learning paradigm: **Diffusion Models.** These models aim to learn representations at higher level than the real data (like in the Plato's allegory). A diffusion process noise the data up to a simple distribution, then the process is reversed and from this simple distribution we can generate new data.

List of attended classes

- 01SCTIU Text mining and analytics (3,30/09/2021)
- 01SCVIU Data analytics for science and society (3,30/09/2021)
- 01TRERS Open geospatial data (3,28/06/2021)
- 01TSBRV Scienza dei dati applicata alle reti complesse (4,23/07/2021)
- 01UJBRV Adversarial training of neural networks (3,03/06/2021)
- 01UJCRVSpectral and machine learning methods for uncertainty quantification (4,31/05/2021)
- 01UMNRV Advanced deep Learning (didattica di eccellenza) (6,15/06/2021)
- 01UNRRV Entrepreneurship and start-up creation (8,31/05/2021)
- 03QTIIU Mimetic learning (4,15/01/2021)

Politecnico di Torino
1859

PhD program in
Electrical, Electronics and Communications Engineering