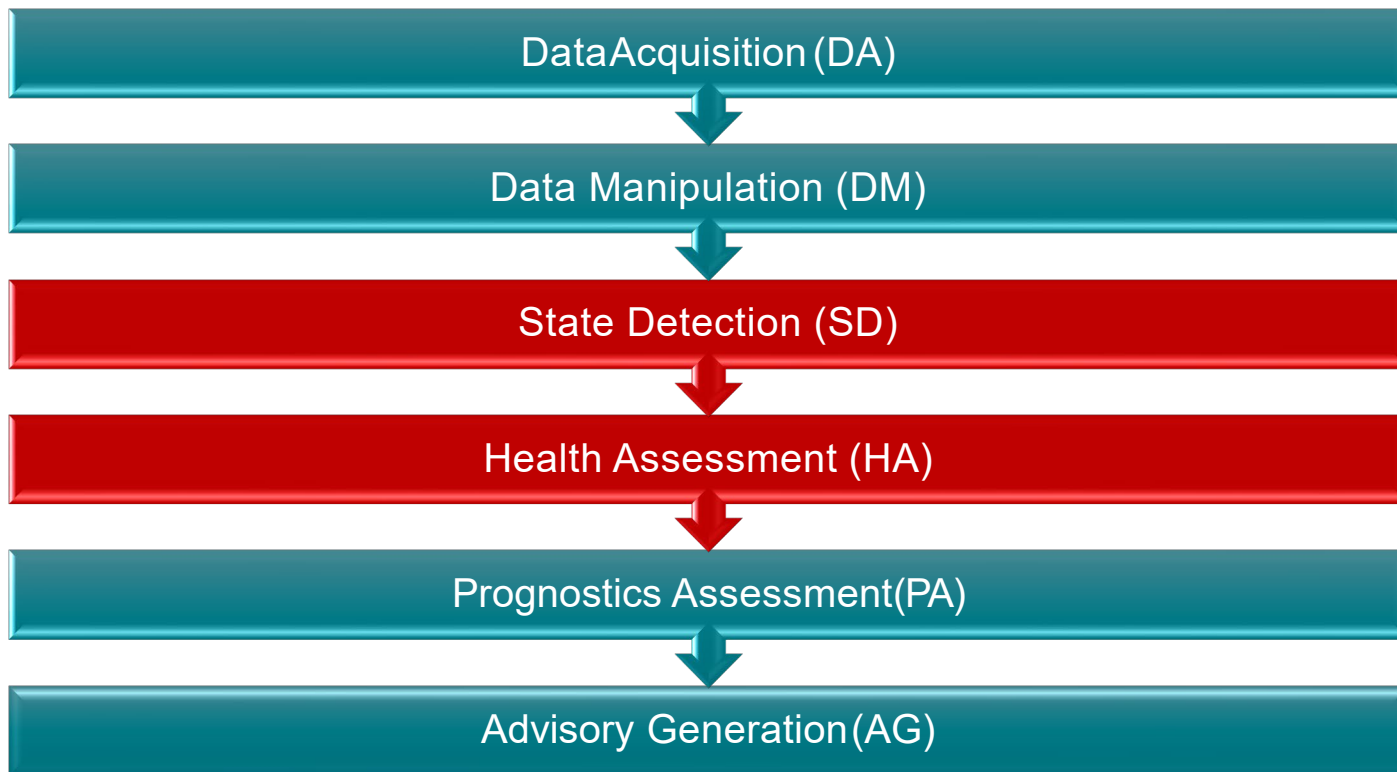
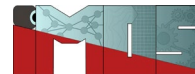
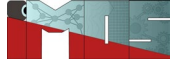
The background of the slide is an aerial photograph of the EPFL campus in Lausanne, Switzerland. The image shows various university buildings, green spaces, and a large lake in the distance under a cloudy sky. A large red rectangular area is overlaid on the right side of the image, containing the title text.

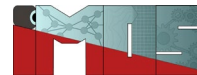
Machine Learning for Predictive Maintenance Applications: Self-Supervised learning

Prof. Dr. Olga Fink

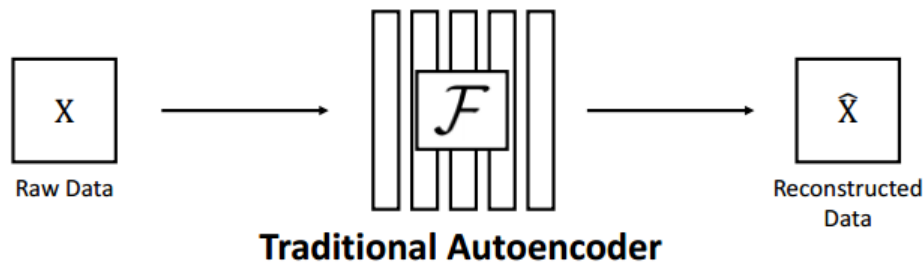




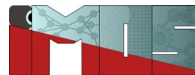
Self-Supervision



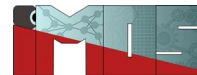
- Why may autoencoders not be sufficient?
 - Use pixel-wise loss, no structural loss incorporated
 - Reconstruction can hardly represent semantic information



What is Self-Supervision?

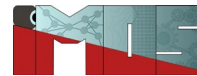


- A form of unsupervised learning where the data provides the supervision
- Use naturally existing supervision signals for training.
- (Almost) no human intervention
- In general, define an auxiliary (supervised) learning task with the labels derived from the data
- The task defines a proxy loss, and the network is forced to learn what we really care about, e.g. a semantic representation, in order to solve it
- Many self-supervised tasks for images
- Often complementary, and combining improves performance



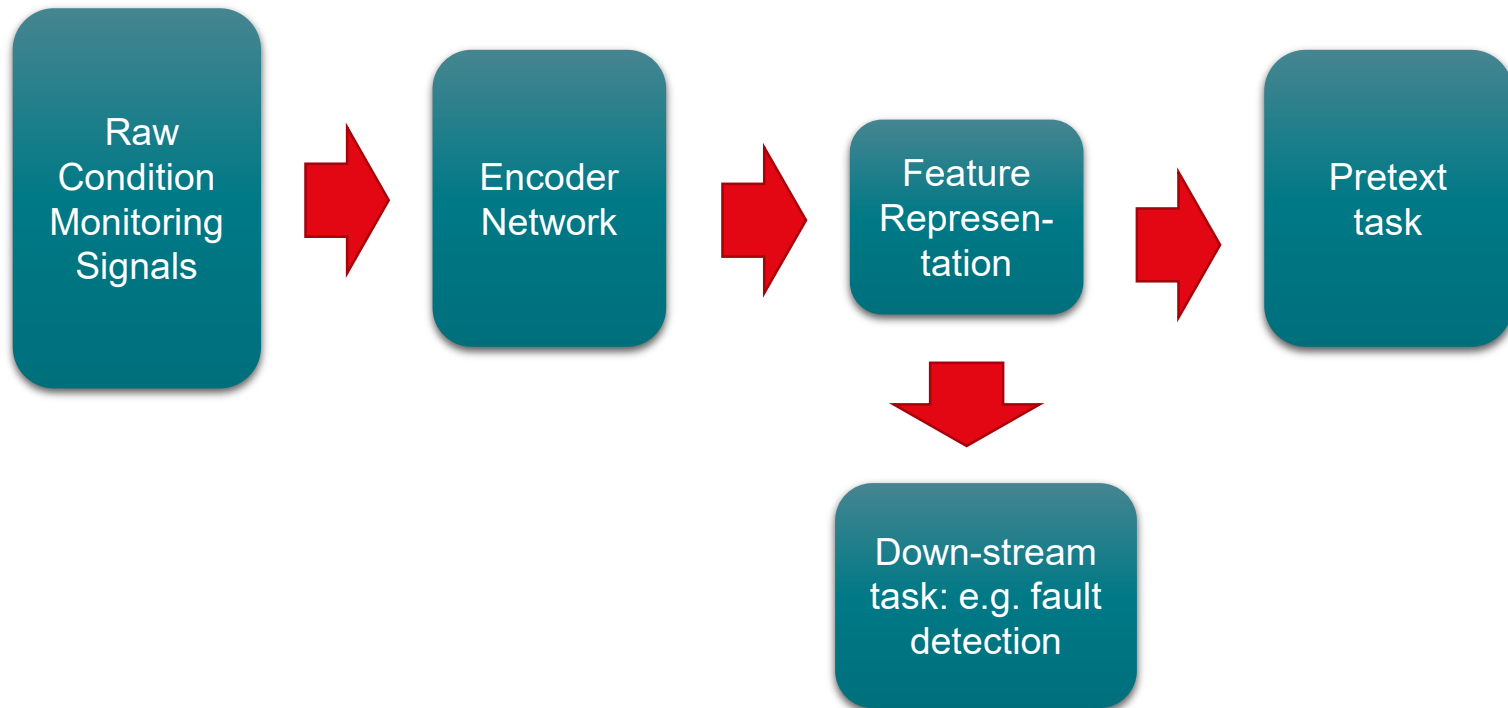
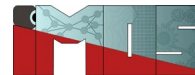
- Pretext task → important strategy for learning data representations under self-supervised mode
- Self-defined pseudo-labels
- Pseudo-labels automatically generated based on the attributes found in the unlabeled data

Important pretext tasks (for computer vision)

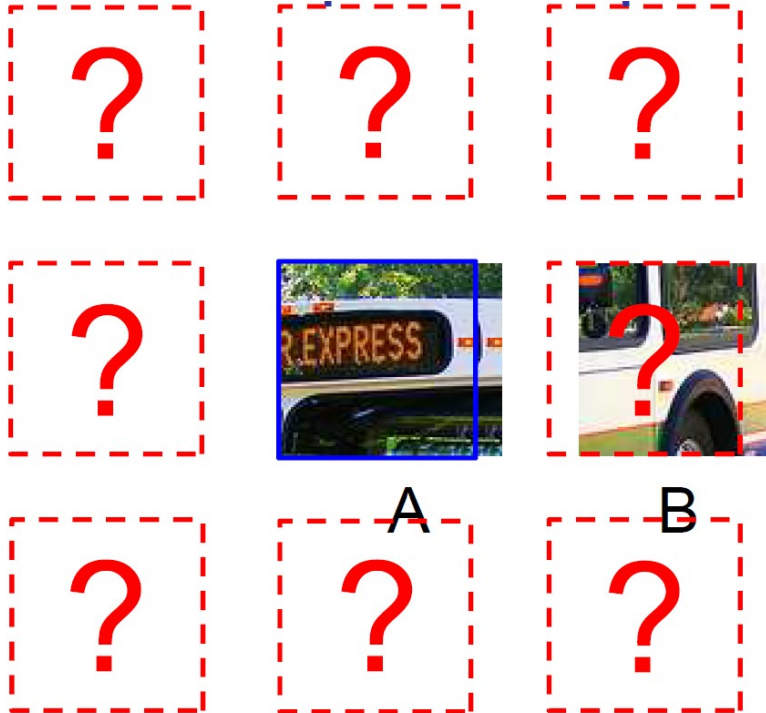
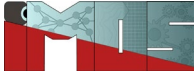


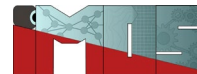
- color transformations
- geometric transformations
- context-based tasks
- cross-modal-based tasks

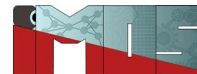
Basic idea of self-supervised learning



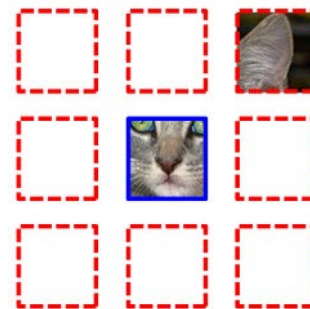
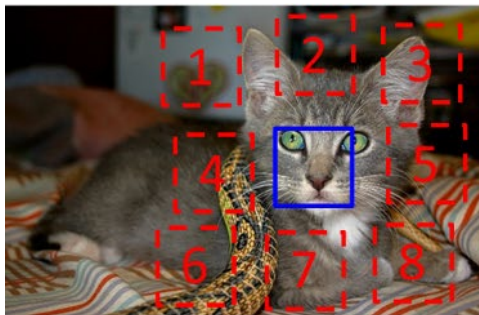
Relative position?



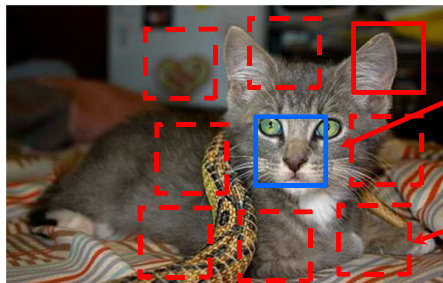
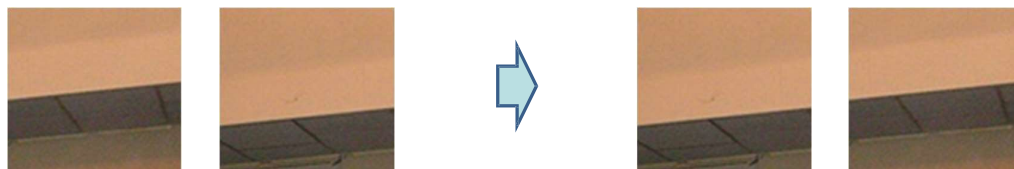
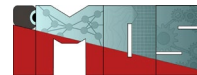




- Solving the Jigsaw
 - Predict relative positions of patches
 - You have to understand the object to solve this problem!
 - Be aware of trivial solution! CNN is especially good at it

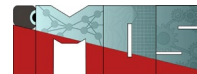


Avoiding Trivial Shortcuts



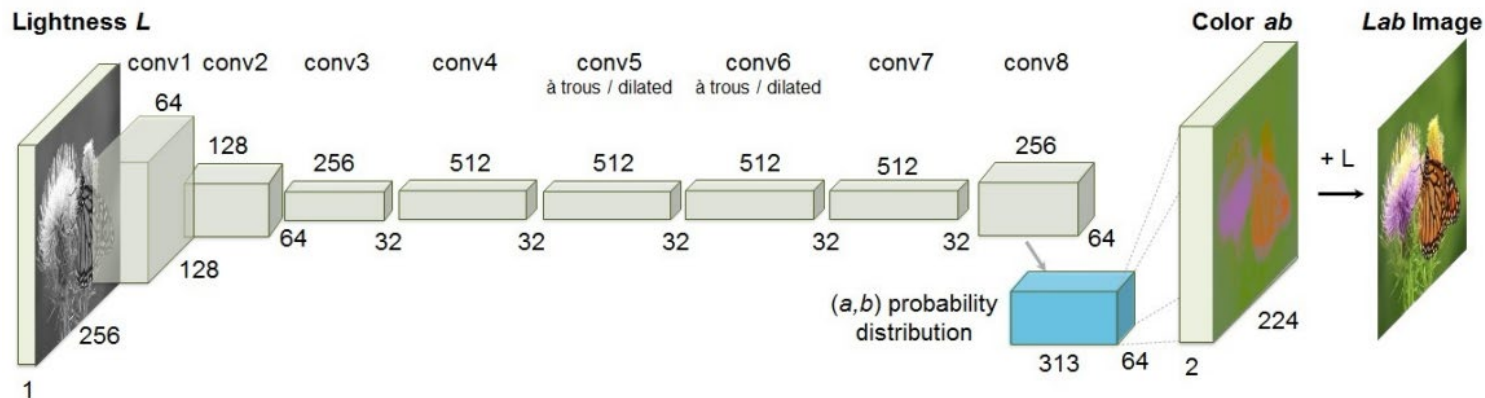
Include a
gap

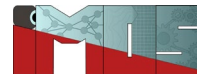
Jitter the patch
locations



■ Colorization

- You have to know what the object is before you predict its color
- E.g. Apple is red/green, sky is blue, etc.

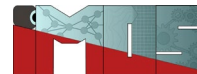




- Colorization

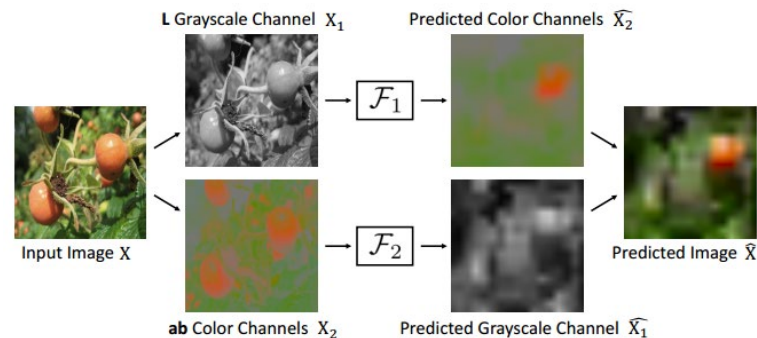
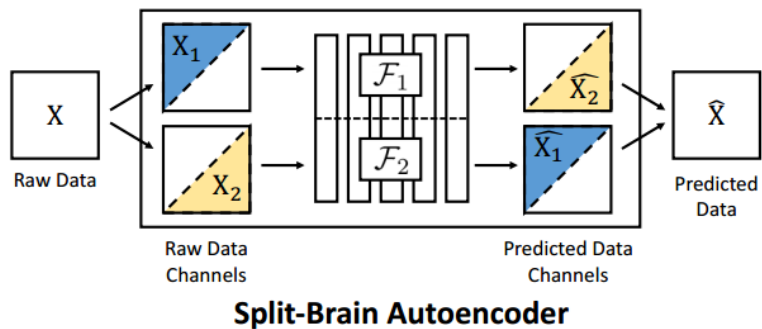
- It is important how to interpret your work!
- Example colorization of Ansel Adams's B&W photos



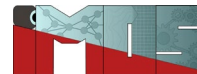


- Colorization

- Stronger supervision, cross-supervision of different parts of data



Zhang, R., Isola, P., & Efros, A. A. Split-Brain Autoencoders: Unsupervised Learning by Cross-Channel Prediction. *In CVPR 2017*



Which image has the correct rotation?



Unsupervised representation learning by predicting image rotations,
Spyros Gidaris, Praveer Singh, Nikos Komodakis, ICLR 2018