Popular science summary of the PhD thesis

PhD student	Andrea Dittadi
Title of the PhD thesis	On the Generalization of Learned Structured Representations
PhD school/Department	DTU Compute

Science summary

Despite tremendous progress in deep learning over the last decade, state-of-the-art methods still typically fail to generalize in a systematic and predictable manner beyond their training data. Humans, on the other hand, naturally generalize to unseen situations.

It has been argued that incorporating structure in deep neural networks may be beneficial in this respect. Humans perceive the world as highly structured. For example, we interpret visual scenes in terms of a few high-level explanatory factors, such as the shapes and colors of objects. This, together with an intuitive understanding of the interactions between objects, allows us to robustly predict future outcomes even in unseen situations, thus enabling effective planning and decision-making.

In this thesis, we focus on representation learning—a subfield of machine learning that uses deep neural networks to learn compact representations of complex data. For example, while a photo consists of millions of pixels, it can roughly be described in a significantly more concise and meaningful manner in terms of objects and their properties. More specifically, we consider *disentangled representations*—where the underlying explanatory factors of the data are represented separately—and *object-centric representations*—where a visual scene is interpreted in terms of objects which are then represented in distinct representation slots. Such structured representations are believed to be beneficial, for example, for learning downstream tasks and for robust generalization beyond the training data distribution. This thesis presents large-scale empirical studies to thoroughly investigate the performance of state-of-the-art structured representation learning methods, with a particular focus on their generalization.