Learning Equivariance in Deep CNNs

TL;DR Can we learn equivariance in a CNN for unknown input transformations?

Supervisors: Attila Lengyel, Jan van Gemert

Description:

CNNs are translation equivariant: first translating an image $\mathbf{x}$ and then convolving with a filter is the same as first convolving $\mathbf{x}$ with a filter and then translating. Translation equivariance makes it possible to share the same filter parameters over all spatial locations in the input, thereby reducing the parameter count, computation and the need for large annotated datasets and extensive data augmentation (i.e. we don't need to show examples of an object in all spatial locations in an image). Group Equivariant Convolutions [1] (GConvs) enable equivariance to a larger group of transformations G, including translations, rotations of multiples of 90 degrees ($p4$ group), and horizontal and vertical flips ($p4m$ group). This further improves data efficiency.

Given input $\mathbf{x}$, neural network $\phi(\mathbf{x})$ and image transformation $g(\mathbf{x}) \in G$, equivariance is formally defined as

$$g'(\phi(\mathbf{x})) = \phi(g(\mathbf{x})),$$

i.e. the transformation of the feature representation outputted by the network (g') is predictable from the input transformation g. Now assume that the transformation g is unknown or not parameterizable (e.g. style transfer). Can we still preserve the equivariance property in the network by making the feature transformation g' part of the network and learning it from data?

This research topic combines multiple active research fields within computer vision, including equivariant deep learning, contrastive learning and incorporating prior knowledge into deep neural networks. See the papers under references and pointers for further reading.

References and pointers:

- [1] Group Equivariant Convolutional Networks - https://arxiv.org/abs/1602.07576, blog post
- [2] Spatial Transformer Networks - https://arxiv.org/abs/1506.02025
- [3] Equivariant Transformer Networks - https://arxiv.org/abs/1901.11399
- [4] A Simple Framework for Contrastive Learning of Visual Representations - https://arxiv.org/abs/2002.05709
- [5] A Survey on Contrastive Self-supervised Learning - https://arxiv.org/abs/2011.00362