Deep Predictive Coding Models of Sensory Information Processing in the Brain

Shirin Dora (s.dora@uva.nl)
University of Amsterdam
Science Park 904, Amsterdam 1098 XH, Netherlands

Cyriel Pennartz (C.M.A.Pennartz@uva.nl)
University of Amsterdam
Science Park 904, Amsterdam 1098 XH, Netherlands

Sander Bohte (S.M.Bohte@cwi.nl)
Centrum Wiskunde & Informatica
Science Park 123, Amsterdam 1098 XG, Netherlands

Abstract:

Predictive coding describes how feedforward and feedback connections in the brain enable efficient processing of sensory information in the brain. In this framework, the bottom-up information represents the information received from the external environment and top-down information influences the processing of bottom-up information based upon context, experience, etc. Here, we used predictive coding to construct a deep neural network model of visual information processing in the sensory areas. The network uses an architecture in which each neuron responds only to information in its receptive field. The trained model is used to infer sets of hierarchical causes for real-world images. Here, we show that the model can capture the statistical regularities of real-world images by using the trained model to infer causes behind natural images that are never before presented to the network.

Keywords: Predictive coding; deep neural network

Introduction

The sensory areas in the brain receive feedforward connections from upstream regions (like thalamus) as well as feedback connections from downstream regions in the brain. Predictive coding describes the hypothesized role of these feedforward and feedback connections in the processing of sensory information in the brain. The sensory information flows through feedforward connections to higher regions in the brain. The feedback connections from higher regions affect how sensory information is processed in a particular region of the brain.

Rao and Ballard (Rao & Ballard, 1999) proposed one of the first computational models of predictive coding. The model used a three layered fully connected neural network with symmetric recurrent connections between all the layers. In (Spratling, 2012), a model of predictive coding that employs non-symmetric weights was developed. Most of the existing computational models of predictive coding have focused on training small neural networks in order to develop a model for information processing in area V1. Here, we describe the procedure to train a deep neural network model that is based on the model of predictive coding developed by Rao and Ballard. The network employs an architecture that is based upon the retinotopic organization of receptive fields in the sensory areas and is neural network is trained on real-world images to build a model of the visual information processing in the brain. We use the trained model to infer causes for sensory stimuli (natural images) never before presented to the network. We find that the causes inferred using the trained model are sufficient to reconstruct the original stimuli. This shows that the trained model can capture the statistical regularities that are characteristic of real-world images.

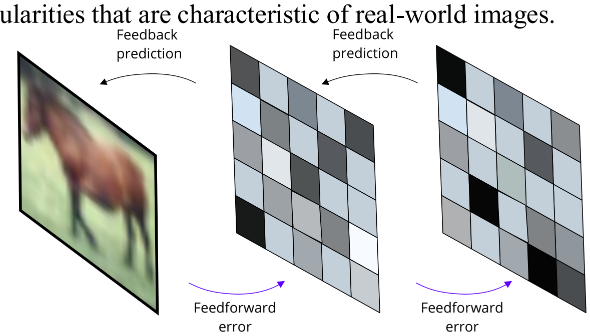

Figure 1 Direction of information propagation in the deep predictive coding network.

Methods

The architecture of the neural network used is similar to that of a Convolutional Neural Network (CNNs). But there are two important differences with respect to CNNs. First, the neural network uses locally connected neurons i.e. there is no weight sharing. This allows us to respect the constraints imposed by the retinotopic organization of receptive fields. Second, in a CNN, information serially propagates in the forward direction while computing the network output and backwards during learning. In the deep predictive coding network, however, the information propagation in both directions happens in parallel in each layer of the network (see Figure 1).

Consider a neural network with $(N+1)$ layers where 0^{th} layer denotes the input layer and subsequent layers in the network are numbered from 1 to N. Predictive coding is used to infer causes and learn the synaptic weights in the network. For a given stimulus, the model alternately adapts the causes and weights in the network. The activations of neurons in a given layer of the network represent the causes inferred by the model behind the neuronal activations at the layer below. For example, the neuronal activations in layer 1 represent the causes that generated the actual sensory stimulus presented to the network at layer 0.

It may be observed that the causes inferred in a given layer of the network can be used to reconstruct the causes inferred by the model in the layer below. Using this mechanism, it is possible to reconstruct the original stimulus starting from the inferred causes in any layer of the network. This allows us to use the model as a generative model while starting with the causes inferred in the topmost layer (see Figure 2).

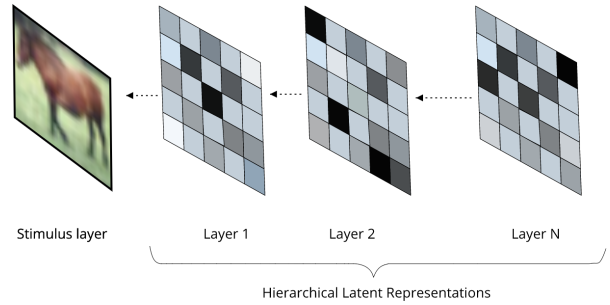

Figure 2 Using the trained model and inferred causes to reconstruct the sensory stimulus.

Results

We trained a deep neural network with six layers using predictive coding on 1000 images of horses and ships from the CIFAR-10 data set. Layer 0 has same shape as the real-world images presented to the network i.e. 32x32x3. Each layer in the model uses filters of size 7x7 except the last layer which uses filters of size 8x8. Layer 1 uses 32 filters and the number of filters are doubled in each subsequent layer resulting in 512 filters in the last layer. The network has close to 90000 neurons in total.

Figure 3(a) shows the reconstructions obtained using the inferred causes for the natural images used to train the model. Figure 3(b) shows the reconstructions obtained using the causes inferred by the model for the images that were never before presented to the model. These images are of objects like dogs, birds, airplanes, etc. which are not presented to the network during training. Figure 3(c) shows the causes inferred by the model for translated versions of the original images present in the CIFAR-10 data set. Note, for images shown in Figure 3(b) and 3(c) the model was only used to infer causes. There was no learning of weights.

Conclusions

We described the procedure to train a deep predictive coding network that uses an architecture inspired by the retinotopic organization of receptive fields in the brain. The network was trained on 1000 natural images of horses and ships from the CIFAR-10 data set. We showed that the trained model can be

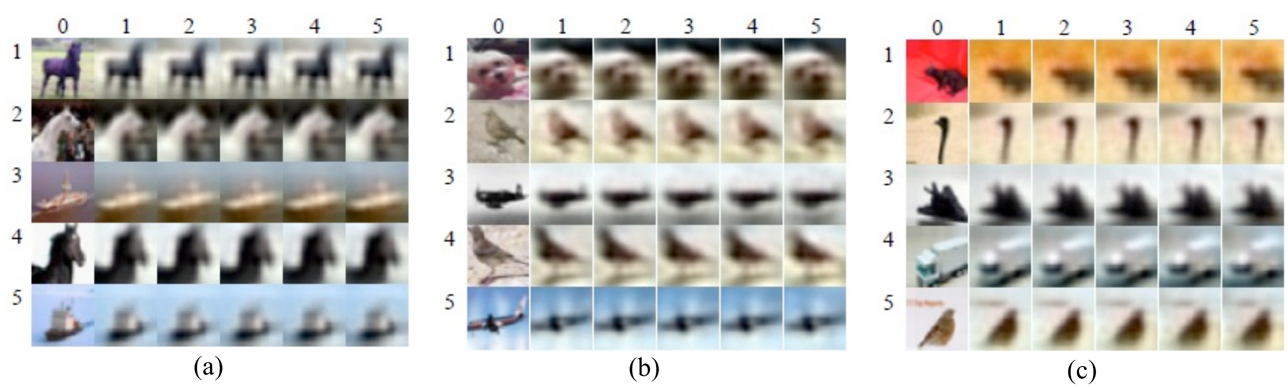

Figure 3 The images are reconstructions obtained using the causes inferred by the model. (a) Reconstructions for the images used to train the model. The numbers on the left side represent the index of the sample and the numbers at the top denote the layer in which inferred causes are used to reconstruct the sensory stimulus. The number ‘0’ denotes the original stimulus. (b) Reconstructions for images never before presented to the network. (c) Reconstruction for translated versions of the images in the CIFAR-10 data set.

used to infer causes for natural images of objects that are never before shown to the network during training.

References

Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*(1), 79–87.

Spratling, M. W. (2012). Unsupervised Learning of Generative and Discriminative Weights Encoding Elementary Image Components in a Predictive Coding Model of Cortical Function. *Neural Computation*, *24*(1), 60–103.