From biological vision to unsupervised hierarchical sparse coding

V. Boutin[1,2,*], A. Franciosini[1], F. Ruffier[2], L. Perrinet[1]

[1]Aix Marseille Univ, CNRS, INT, Marseille, France, [2]Aix Marseille Univ, CNRS, ISM, Marseille, France

*Contact: victor.boutin@univ-amu.fr

Introduction

The brain has to solve inverse problems to correctly interpret sensory data and infer the set of causes that generated the sensory inputs. Such a problem is typically ill-posed, and thus requires constraint the narrow down the number of solutions.

Predictive coding (PC) [1] is a computational neuroscience framework that finds the most likely causes for the sensory input by minimizing the mismatch between the sensory data and the predicted input. Such a framework could be used to build sparse hierarchical internal representations of a given input.

Method

In the proposed Sparse Deep Predictive Coding (SDPC) model, each layer attempts to predict the activity of the lower layer via feedback connections. The error between predicted and actual response is then sent back to the next higher level via feed-forward connections to correct the estimation of the representation.

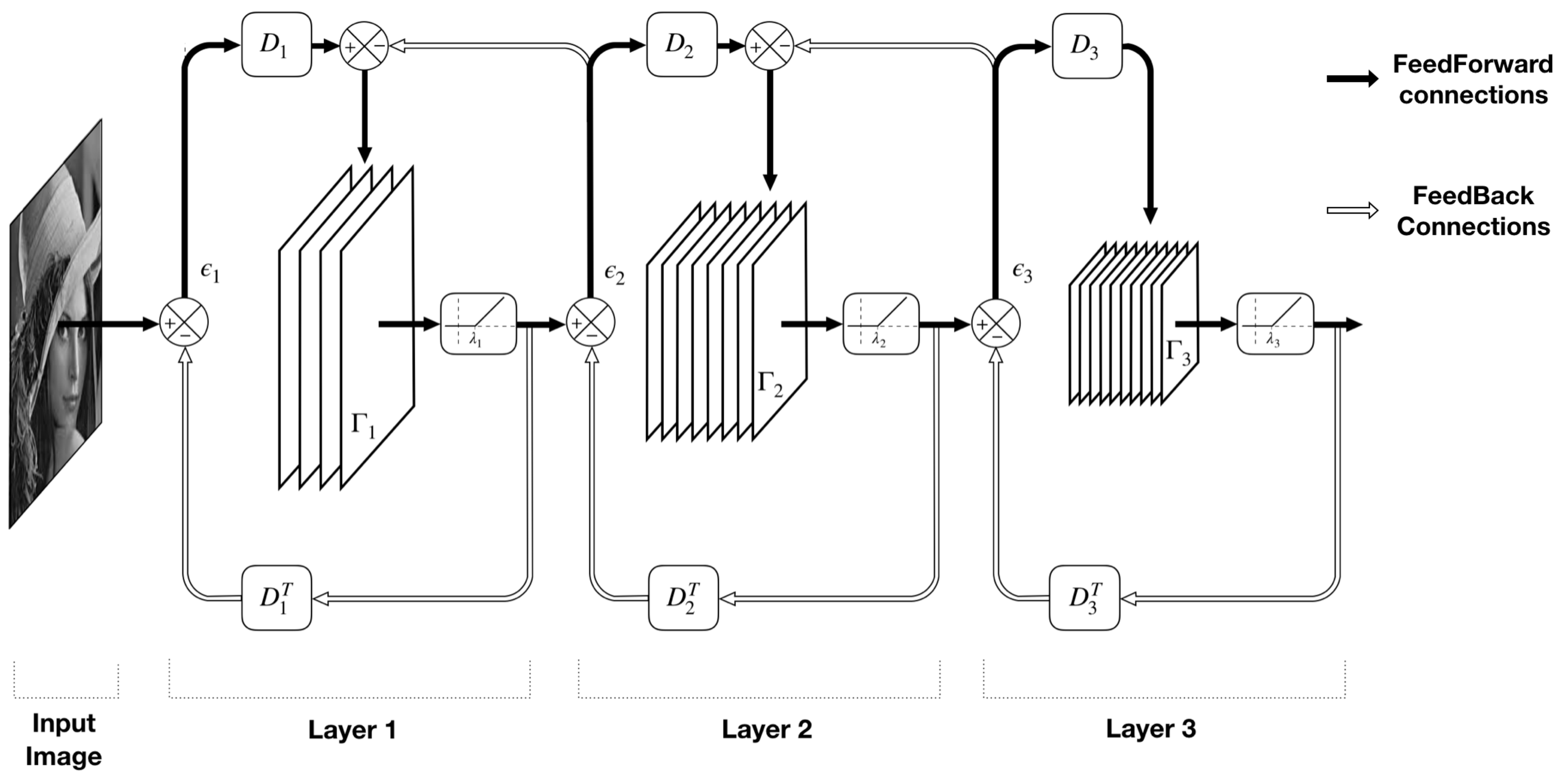


Figure 1: A 3-layered SDPC Model. In this model, Γ_i is the internal sparse representation of the image at layer i, ϵ_i is the representation error at layer i. D_i are the synaptic weights and λ_i control the level of sparsity.

Mathematical formulation:

$$\mathcal{L} = \frac{1}{2}\|I - \Gamma_1 D_1^T\|_2^2 + \frac{1}{2}\sum_{i=2}^{L}\|\Gamma_{i-1} - \Gamma_i D_i^T\|_2^2 + \sum_{i=1}^{L}\lambda_i\|\Gamma_i\|_1 \quad (1)$$

Inference algorithm:

Algorithm 1: SDPC inference algorithm inspired by [2]

input : I, $\{D_i\}_{i=1}^L$, $\{\lambda_i\}_{i=1}^L$, th
#Initialization of the representation
$\Gamma_0 = I$, $\{\Gamma_i\}_{i=1}^L = 0$
while (change in $\mathcal{L} >$ th) **do**
  $\mathcal{L} = 0$
  for $i = 1$ **to** L **do**
    #Update feedforward error
    $\epsilon_{FF} = \Gamma_{i-1} - \text{FBConv}(\Gamma_i, D_i)$
    if i=L **then**
      $\epsilon_{FB} = 0$
    else
      #Update feedback error
      $\epsilon_{FB} = \Gamma_i - \text{FBConv}(\Gamma_{i+1}, D_{i+1})$
    # Update internal representations
    $\Gamma_i = \mathcal{T}^{+}_{\lambda_i}(\Gamma_i + \eta\ \text{FFConv}(\epsilon_{FF}, D_i) - \eta\ \epsilon_{FB})$
    $\mathcal{L} \mathrel{+}= \epsilon_{FF}$
return $\{\Gamma_i\}_{i=1}^L$

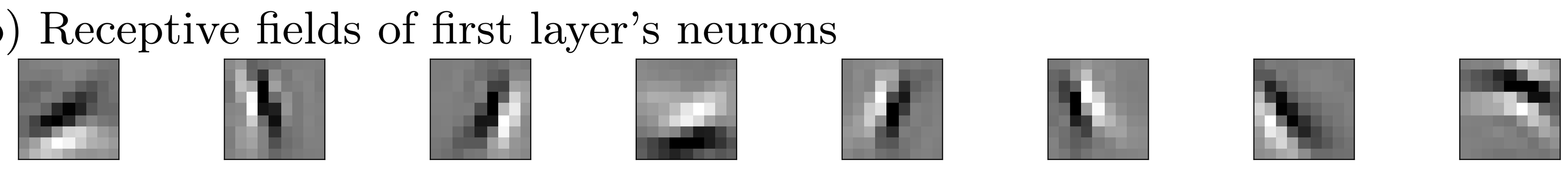
Note: FFConv and FBConv represent respectively the feed-forward and feed-back convolutions. $\mathcal{T}^{+}_{\lambda_i}(\cdot)$ denotes the positive-only soft thresholding operator. Comments are denoted by the # sign.

Results

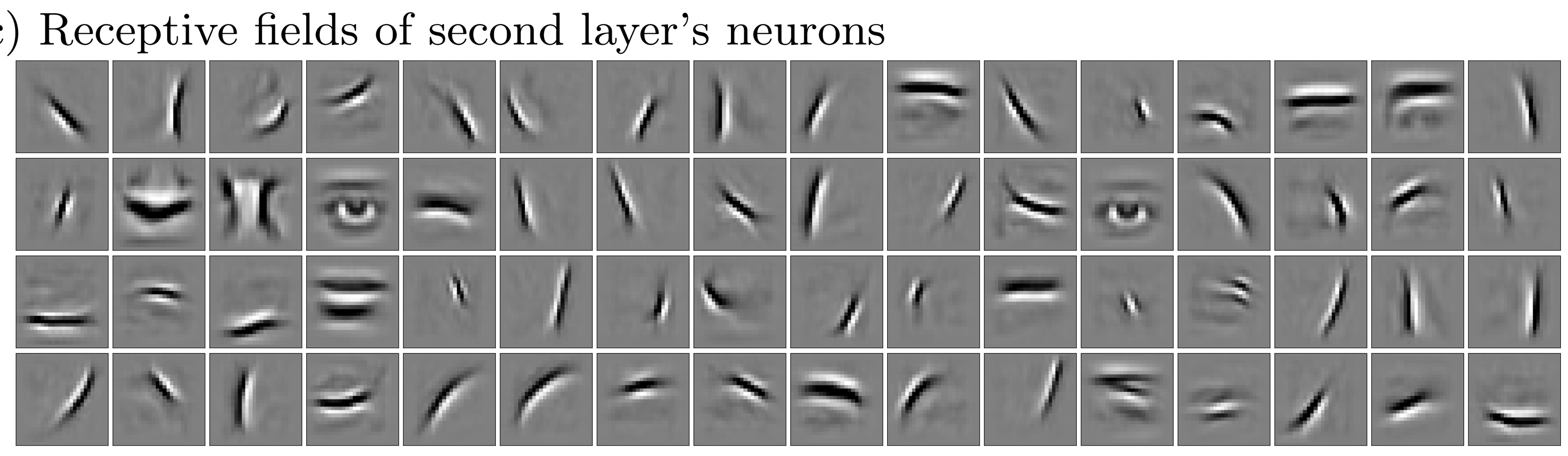
Receptive Fields (RF) on face database :

(a) Examples of AT&T images after pre-processing

(b) Receptive fields of first layer's neurons



(c) Receptive fields of second layer's neurons



(d) Example of Reconstructed Images

Figure 2: first and second layers' RF (b,c) and predicted images (d) obtain with a 2-layer SDPC model applied on AT&T database (a) after the learning.

Classification on MNIST database

Algorithms	Classification Error
Stacked Denoising Autoencoder	1.28 %
k-Sparse Autoencoder	1.35 %
Shallow WTA Autoencoder	1.20 %
Stacked WTA Autoencoder	1.11 %
ML-CSC (3 layers) [3]	1.15 %
SDPC	**2.8 %**

Table 1: Classification results on MNIST using unsupervisedly learnt features.

Conclusion

- First layer's RF are Gabor-like filters, like those observed in V1. Second layer's features are sensitive to curvatures and other complex features. By increasing the scale and the specificity of receptive fields along the network, the model is able to combine simple and low level representation to build a more abstract and meaningful representation of the presented image.
- SDPC accounts for reasonable recognition accuracy on MNIST

References

[1] Rajesh PN Rao and Dana H Ballard. "Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects". In: *Nature neuroscience* 2.1 (1999), p. 79.

[2] William Lotter, Gabriel Kreiman, and David Cox. "Deep predictive coding networks for video prediction and unsupervised learning". In: *arXiv preprint arXiv:1605.08104* (2016).

[3] Jeremias Sulam et al. "Multi-Layer Convolutional Sparse Modeling: Pursuit and Dictionary Learning". In: *CoRR* abs/1708.08705 (2017). arXiv: 1708.08705. URL: http://arxiv.org/abs/1708.08705.

Acknowledgment

This research received funding from the European Union's H2020 programme under the Marie Skłodowska-Curie grant agreement n°713750 and by the Regional Council of Provence-Alpes-Côte d'Azur, A*MIDEX (n°ANR-11-IDEX-0001-02).