Predictive Coding for Deep Neural Networks*

Maxime Chourré, Guilhem Fouilhé, Lison Kardassevitch,
Esteban Marco, and Chloé Michel
Advisor : Andrea Alamia (CerCo-CNRS)

Université Paul Sabatier (UPS), Toulouse, France

Abstract. Predictive Coding is a popular framework in neurosciences for explaining cortical function. In this model, higher-level cortical areas try to predict lower-level neural activity and prediction errors are passed back to higher layers. Deep Neural Networks (DNN) , which use brain-inspired architecture, could be augmented with such a model, providing robustness and a better understanding of spatio-temporal dependencies. We investigate research in this direction and give a quick review on tasks in which Predictive Coding (PC) for DNN has demonstrated its interest, with a strong emphasis on vision-related tasks.

Keywords: Deep Neural Networks · Predictive Coding · Robustness · Feedback connections · Computer Vision.

1 Introduction

Deep Neural Networks (DNN), and more specifically Deep Convolutional Neural Networks (CNN), demonstrated their large potential on the last decade by becoming the state-of-the-art models in a lot of high-level Artificial Intelligence (AI) tasks, including image classification, video prediction, image segmentation among many more. Their architecture takes inspiration from human vision system by using a hierarchical architecture. However, classical approach of DNN only includes feedforward (bottom-up) connections, whereas there is much evidence that the brain also features lateral [10] and feedback (top-down) [1] connections, which results in recurrent neuronal dynamics.

Predictive Coding (PC) designates the theory on how the brain performs probabilistic inference. Although several algorithms are called PC [20], we will most consider the model in which higher-layers tries to predict lower-layer activity trough feedback connections, whereas feedforward connections carry residuals errors (errors between predictions and actual lower-layer activity). This definition comes from the pioneering article of Rao & Ballard [16], which provides a model that is used in almost every articles discussed here.

* This work was produced for TIR (*Travaux d'Initiation à la Recherche*) course in M1 Computer Science of UPS during year 2021/2022.

1.1 Rao & Ballard's model

The model of Rao & Ballard (see Figure 1) can be written as :

$$y^i = \alpha y^i + \beta W e^{i-1} - \gamma e^i \tag{1}$$

where $\alpha, \beta, \gamma > 0$ and

$$e^i = y^i - V^{i+1} y^{i+1} \tag{2}$$

Where y^i is the activity of neurons in layer i, e^i is the residual error of layer i, W^i and V^i respectively are the weights of feedforward and feedback connections between layer i and $i+1$, which verify $V^T = W$ in this model.
The term $\beta W e^{i-1}$ represents the feedforward transmission of previous-layer residual error whereas the term γe^i can be interpreted as a feedback action of prediction by the upper layer, and therefore carries prediction errors.

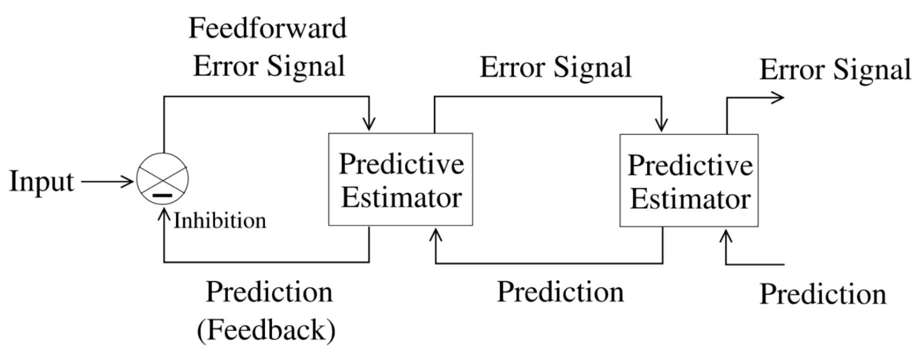


Fig. 1: Rao & Ballard's Hierarchical model for predictive coding. Picture from [16].

1.2 Biological motivations

Such a framework could explain several cognitive observations, like alpha oscillations [2] or filling-in at the blind spot [15].

Alpha oscillations are historically the first observed oscillations by Electroencephalography (EEG). The article [2] has shown that a simple 2-level of PC model explains the emergence of alpha-band rhythms (neural oscillations, 8-12 Hz) in the visual cortex.

As a consequence of the PC dynamics, the authors observed in their models traveling waves propagating forward or backward depending on the cognitive state of the system. Remarkably, the same dynamics was observed in experimental EEG data, with oscillatory waves propagating accordingly to the model's predictions

2 Predictive Coding for Machine Learning

PC turns out to be relevant when considering biological plausibility of DNN. This fact is not only interesting from a neurosciences perspective but also for machine learning applications. Indeed, it seems that PC could be an important element in bridging the gap between these two disciplines. Below we briefly introduce two areas in which ML algorithms can benefit from PC dynamics: training methods and latent representations.

Training methods The algorithm of error back-propagation is widely used to update parameters of a network. However, this update rule is considered as biologically implausible because of the non-locality of these updates. Using a variant of PC on the computational graph turned out to approximate exactly back-propagation [17]. In a similar idea, PredProp [13] proposes a PC approach for bidirectional stochastic optimization. This is interesting as it may provide a framework to justify current ML algorithm -i.e., back-propagation- from a biological point of view.

Latent representations From a data representation perspective, Deep Predictive Coding Networks (PCN) implement a way to contextually alter priors on the latent representations [4] [7], whereas most of the time fixed priors are used in ML, such as sparsity, which struggles to adapt to a new context. This allows models to learn complex general features of the world like spacio-temporal dependencies [8].

In the following section we will give an overview of tasks in which PC integration to a DNN architecture appears to be relevant on a machine learning perspective and will try to gather some important articles for each tasks.

2.1 Visual perception

CNN have reached good performances for tasks related to object recognition, classification, and segmentation, but they also have disclosed their limits and lack of robustness. Neurosciences lead us to think that PC can be a way to improve CNN's robustness.

Image classification is a classic neural network problem, which can be seen as a basic capability of picture's analysis. Most results in using PC for this task suggest that it provides robustness. For instance, [18] (Santana et al. 2015) has shown an improvement of performance against input noise.

In [9] (Han et al. 2018), PCN has shown competitive to state-of-the-art performance on benchmark datasets (CIFAR-10, CIFAR-100, SVHN and ImageNet), despite having fewer layers and parameters.

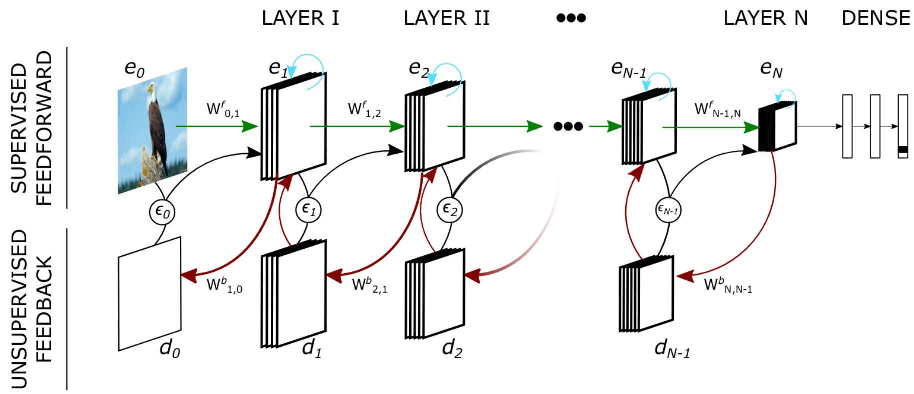


Fig. 2: Predify's model for augmenting a CNN. Picture from [5]. The connections in green are the feedforward connections already existing in the CNN while blue and red ones are respectively recurrent and feedback connections.

More recently, [5] (Choski et al. 2021) provided an open-source framework, *Predify* (see Figure 2), which can easily implement PC in any *Pytorch* CNN and shown that augmenting a few popular CNN (VGG16 and EfficientNetB0) improved their robustness against noise and adversarial attacks.

The paper [23] even proposes a way to increase PCN efficiency by increasing the recursive cycles of computation and by not letting feedforward and feedback connections share weights. To understand this performance's rise, the authors studied the case of testing images that were well-classified by PCN but not by CNN. They discovered that PCN and CNN were not computed the same representation of an image, which yielded a different probability distribution across different categories at each step.

Illusory contours Observations suggests that perceiving an optical illusion involves feedback connections of the brain. The article [14] showed that a PCN can perceive basic optical illusions and "misclassify" a shape because it saw an illusory contour. More specifically, they showed that the feedback connections are mostly responsible for the illusion perception whereas the feedforward connections goes against it, trying to ground the perception to the actual input.

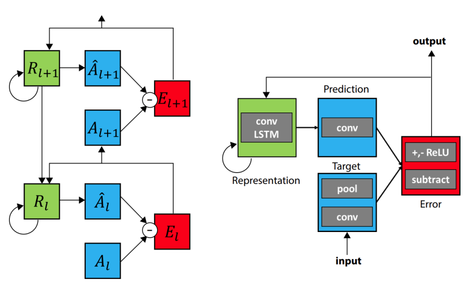
Image reconstruction consists of imitating bio-mechanism that allows our brain to rebuild a picture that we have seen. The issue is to not lose important information of the image while it is piped from the entry layer to the last layer. As explained in [6] and [7], PC is here proposed as a theory where the brain infers causes that generated a sensory stimulus : the inferred causes are related to the top-down flow and to the reconstruction of the image features at each level.

Those papers show that PCN has good results on inferring causes : even with little datasets (1000 images) with few classes (only two in the previous papers), PCN captures the important statistical regularities of photos, which means that it is able to infer representative causes of an image, even for unknown objects. However, those papers also raise an important issue : the images reconstructed by PCN are blurry. They propose to use L1 norm instead of MSE in order to get visually better images, but it stills an open problem. A really recent study [3] proposes an alternative model to have less blurry results by mixing PC with Sparse Coding (SC), they called it the Sparse Deep Predictive Coding (SDPC).

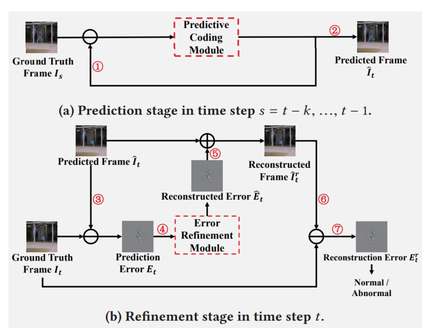
We noticed that in all those studies, they do not try to rebuild an image with several important objects on it, the presented models only shows their results for real-world images with one important object on the picture.

2.2 Perception in Video

Models An intuitive application of knowing spatio-temporal dependencies of the world, which is a motivation of PC, is Video Prediction. This task consists in trying to predict the next frame of a video, given all (or a part of) the previous ones. An approach of this problem using multi-scale convolution has been treated in [12].



(a) Predictive Coding Network (PredNet). Picture from [11]



(b) AnoPCN Module. Picture from [24]

Fig. 3: Two models for Predictive Coding in video

Lotter et al. approaches this task using predictive coding [11]. The model visible in Figure 3a is a succession of layers, and each layer predicts the next frame before sending the difference to the other layer.

They used an unsupervised learning approach, with unlabelled video data collected by themselves, which are videos of moving objects, and camera movements. Also, they were inspired by the fact that the human eye does not need millions of labelled data to learn.

Their model, called **PredNet**, can be used to predict the future frames of a video. The model works because it knows the next frame, so it can learn itself by calculating the error between the prediction and the actual image. This allows using temporal dependencies as supervision for training, even though the dataset is not labelled.

Ye et al. try to upgrade some already existing technologies[24], like anomaly detection in videos. Their model, called **AnoPCN**, takes inspiration from biological PC and achieves state-of-the-art performance in three different benchmarks. The model features a Predictive Coding Module (PCM) and an Error Refinement Module (ERM). The PCM generates frame prediction, and ERM enlarges the regularity score gaps. This model is distinct from **PredNet**[11] because it uses an encoder-decoder approach.

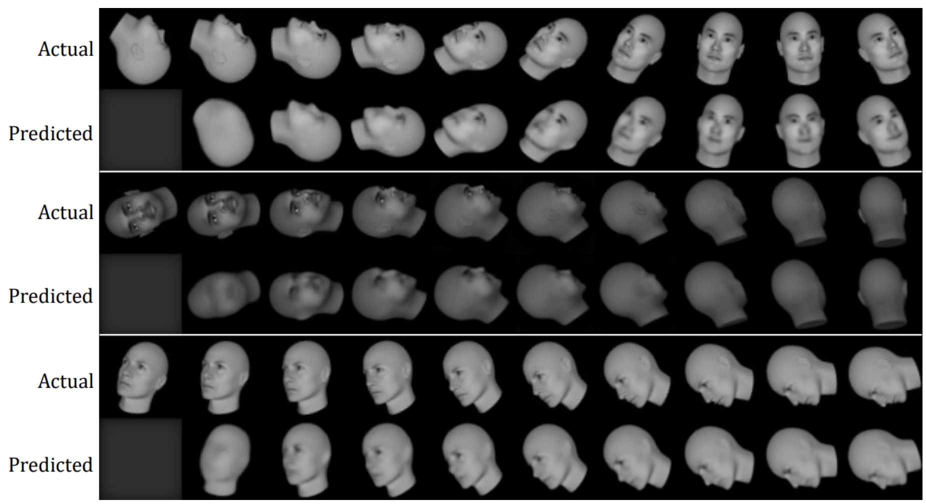


Fig. 4: Next-frame predictions for sequences of rendered faces rotating with two degrees of freedom. Picture from [11]

Examples and tests The video used to train the **PredNet** model includes ten frames of a face rotating with two degrees of freedom and all the faces used are different. They also chose to evaluate their model with some natural image sequences. For that, they used car-mounted camera videos, because the camera and the elements on the video both move. The result is similar to the previous test.

In **Figure 4**, there are three examples of the prediction by **PredNet**:

- The first row represents the actual video (or frame sequences) and the second row is the generated row by the **PredNet** model. We can see that the first image is greyed out in the generated row. It is a uniform representation set by default because the model can not predict the first frame as there is no previous one.

– Then, the model will learn and try calculating the second frame. Improvements can be noted in the second frame, but it's blurry.
– The last frame seems identical to the original one, thus the model learned successfully.

Table 1: Evaluation of next-frame predictions on Rotating Faces Dataset. Table from [11].

	Mean Square Error	Structural similarity index
PredNet L_0	**0.0152**	**0.937**
PredNet L_{all}	0.0157	0.921
CNN-LSTM Enc.-Dec.	0.0180	0.907
Copy Last Frame	0.125	0.631

Finally, the benchmark proved their model successful: as shown in Table 1, the best model on this dataset is their **PredNet** L_0.

3 Similar Ideas

Feedback connections The fact that the brains features not only feed-forward (bottom-up) but also lateral (recurrent) and feedback (top-down) connections is well known and massively supported by neurosciences observations, as we've seen in 1. Therefore, some articles investigated the implementation of these connections in CNN without necessarily adopting a PC perspective. In [19] for instance, feedback connections have shown beneficial for image classification with heavy occlusions or Gaussian noise, suggesting they play an important role in robustness.

Another neuro-inspired motivation of implementing feedback connections is for internal selective attention modelling [21]. Feedback connections allow modifying the weights of the previous convolutional filter, which leads to focusing on specifics parts of the image.

Auto-encoders The concept of PC described here and some of the architectures considered (e.g. Predify [5]) are similar to stacked denoising auto-encoders [22]. The convolution and prediction layers of a CNN augmented with PC can be seen respectively as an encoder and a decoder.

Sparse Coding As stated in 2, predictive coding implies imposing contextual prior knowledge on the representation of data. Sparse coding is also a biological inspired constraint on the representation of data (encouraging the proportion of activated neurons to be small, which is widely observed in cortical activity). However, sparse coding is a fixed prior that cannot adapt to a context while in a predictive coding framework these prior knowledge constraints are created contextually. Moreover, predictive coding is a much larger framework from a neurosciences point of view, whereas sparse coding is mainly this constraint.

4 Conclusion

The huge progression of DNN performances over the past years is mostly a consequence of the large quantity of computer science research in this domain that is allowed by its important potential in engineering. This led to more liberty regarding biological plausibility of architectures, which can be both beneficial to neurosciences -which may find in machine learning architectures interesting models- and of course to computer science because machine learning models end-up being more efficient. The exchange between these two disciplines also occurs on the other way around. The human vision system remains way more efficient for general perception than computers while only requiring around 20 watts to work. A better understanding of the brain can probably lead to improvements in machine learning performances.

This state-of-the-art emphasises this symbiosis between PC (from a biological aspect) and DNN (machine learning) by reviewing a significant part of the studies that have been led in this field. We have started to provide a biological framework with Rao & Ballard model. Then, we have discussed about the interest of applying PC to machine learning. In this context, we introduced the main scope of applications that have been studied in research : the perception in images and videos.

Because PC operates on biological perception, it is not surprising that researches are mainly led in computer vision. However, the improvements of DNN observed in this state-of-the-art suggests to try extending PCN to others machine learning tasks that doesn't necessarily deals with perception and computer vision.

We have highlighted in this work that Predictive Coding framework is interesting for both neurosciences and computer science. This suggests that further research on this topic by neurosciences or machine learning scientists can lead to a mutual profitable work.

References

1. Alamia, A., Mozafari, M., Choksi, B., VanRullen, R.: On the role of feedback in visual processing: a predictive coding perspective. arXiv:2106.04225 [cs, q-bio] (Jun 2021), http://arxiv.org/abs/2106.04225
2. Alamia, A., VanRullen, R.: Alpha oscillations and traveling waves: Signatures of predictive coding? PLOS Biology **17**(10), e3000487 (Oct 2019). https://doi.org/10.1371/journal.pbio.3000487, publisher: Public Library of Science
3. Boutin, V., Franciosini, A., Chavane, F., Ruffier, F., Perrinet, L.: Sparse deep predictive coding captures contour integration capabilities of the early visual system. PLOS Computational Biology **17**(1),

e1008629 (Jan 2021). https://doi.org/10.1371/journal.pcbi.1008629, https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1008629
4. Chalasani, R., Principe, J.C.: Deep Predictive Coding Networks. arXiv:1301.3541 [cs, stat] (Mar 2013), http://arxiv.org/abs/1301.3541, arXiv: 1301.3541
5. Choksi, B., Mozafari, M., O'May, C.B., Ador, B., Alamia, A., VanRullen, R.: Predify: Augmenting deep neural networks with brain-inspired predictive coding dynamics. arXiv:2106.02749 [cs, q-bio] (Nov 2021), http://arxiv.org/abs/2106.02749, arXiv: 2106.02749
6. Dora, S., Pennartz, C., Bohte, S.: A Deep Predictive Coding Network for Inferring Hierarchical Causes Underlying Sensory Inputs. In: Kůrková, V., Manolopoulos, Y., Hammer, B., Iliadis, L., Maglogiannis, I. (eds.) Artificial Neural Networks and Machine Learning – ICANN 2018. pp. 457–467. Lecture Notes in Computer Science, Springer International Publishing, Cham (2018). https://doi.org/10.1007/978-3-030-01424-7˙45
7. Dora, S., Pennartz, C., Bohte, S.: A Deep Predictive Coding Network for Learning Latent Representations (Mar 2018). https://doi.org/10.1101/278218
8. Fonseca, M.: Deep Predictive Coding for Spatiotemporal Representation Learning p. 63 (2018)
9. Han, K., Wen, H., Zhang, Y., Fu, D., Culurciello, E., Liu, Z.: Deep Predictive Coding Network with Local Recurrent Processing for Object Recognition. 32nd Conference on Neural Information Processing Systems p. 13 (2018)
10. Kietzmann, T.C., Spoerer, C.J., Sörensen, L.K.A., Cichy, R.M., Hauk, O., Kriegeskorte, N.: Recurrence is required to capture the representational dynamics of the human visual system. Proceedings of the National Academy of Sciences **116**(43), 21854–21863 (Oct 2019). https://doi.org/10.1073/pnas.1905544116
11. Lotter, W., Kreiman, G., Cox, D.: Deep Predictive Coding Networks for Video Prediction and Unsupervised Learning. arXiv:1605.08104 [cs, q-bio] (Feb 2017), http://arxiv.org/abs/1605.08104, arXiv: 1605.08104
12. Mathieu, M., Couprie, C., LeCun, Y.: Deep multi-scale video prediction beyond mean square error. arXiv:1511.05440 [cs, stat] (Feb 2016), http://arxiv.org/abs/1511.05440, arXiv: 1511.05440
13. Ofner, A., Stober, S.: PredProp: Bidirectional Stochastic Optimization with Precision Weighted Predictive Coding. arXiv:2111.08792 [cs] (Nov 2021), http://arxiv.org/abs/2111.08792, arXiv: 2111.08792
14. Pang, Z., O'May, C.B., Choksi, B., VanRullen, R.: Predictive coding feedback results in perceived illusory contours in a recurrent neural network. Neural Networks **144**, 164–175 (Dec 2021). https://doi.org/10.1016/j.neunet.2021.08.024, https://www.sciencedirect.com/science/article/pii/S0893608021003373
15. Raman, R., Sarkar, S.: Predictive Coding: A Possible Explanation of Filling-In at the Blind Spot. PLOS ONE **11**(3), e0151194 (Mar 2016). https://doi.org/10.1371/journal.pone.0151194, https://dx.plos.org/10.1371/journal.pone.0151194
16. Rao, R.P.N., Ballard, D.H.: Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. Nature Neuroscience **2**(1) (Jan 1999). https://doi.org/10.1038/4580
17. Salvatori, T., Song, Y., Xu, Z., Lukasiewicz, T., Bogacz, R.: Reverse Differentiation via Predictive Coding p. 9 (2022)
18. Santana, E., Principe, J.C.: Mixed generative and supervised learning modes in Deep Predictive Coding Networks. In: 2015 International Joint Conference on Neural Networks (IJCNN). pp. 1–4 (Jul 2015). https://doi.org/10.1109/IJCNN.2015.7280746, iSSN: 2161-4407

19. Spoerer, C.J., McClure, P., Kriegeskorte, N.: Recurrent Convolutional Neural Networks: A Better Model of Biological Object Recognition. Frontiers in Psychology **8** (2017), https://www.frontiersin.org/article/10.3389/fpsyg.2017.01551
20. Spratling, M.W.: A review of predictive coding algorithms. Brain and Cognition **112**, 92–97 (Mar 2017). https://doi.org/10.1016/j.bandc.2015.11.003
21. Stollenga, M., Masci, J., Gomez, F., Schmidhuber, J.: Deep Networks with Internal Selective Attention through Feedback Connections. Advances in Neural Information Processing Systems **4** (Jul 2014)
22. Vincent, P., Larochelle, H., Lajoie, I., Bengio, Y., Manzagol, P.A.: Stacked Denoising Autoencoders: Learning Useful Representations in a Deep Network with a Local Denoising Criterion. Journal of Machine Learning Research p. 38 (2010)
23. Wen, H., Han, K., Shi, J., Zhang, Y., Culurciello, E., Liu, Z.: Deep Predictive Coding Network for Object Recognition. arXiv:1802.04762 [cs] (Jul 2018), http://arxiv.org/abs/1802.04762, arXiv: 1802.04762
24. Ye, M., Peng, X., Gan, W., Wu, W., Qiao, Y.: AnoPCN: Video Anomaly Detection via Deep Predictive Coding Network. In: Proceedings of the 27th ACM International Conference on Multimedia. pp. 1805–1813. MM '19, Association for Computing Machinery, New York, NY, USA (Oct 2019). https://doi.org/10.1145/3343031.3350899