ESANN 2026 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 22-24 April 2026, i6doc.com publ., ISBN 9782875870964. Available from http://www.i6doc.com/en/.

# Predictive Coding inspired convolutional networks can capture the neural dynamics of recurrent processing in human image recognition

Manshan Guo[1,2], Bhavin Choksi[1], Sari Saba-Sadiya[1], Pablo Oyarzo[2], Radoslaw M. Cichy[2], and Gemma Roig[1]

1- Goethe Universität, Frankfurt 2- Freie Universität, Berlin

**Abstract**.

Inspired by the robustness of human vision, various attempts have been made to incorporate brain-inspired mechanisms into artificial neural networks. A popular candidate has been predictive coding, a prominent theory in neuroscience, that theorizes that feedback connections communicate top-down predictions to earlier regions. While recurrence in models has been demonstrated to be useful when processing noisy and difficult stimuli, a direct evidence of its utility for explaining brain data under such situation was yet to be shown. Here, we investigated whether such brain-inspired mechanism *actually* helps to capture neural dynamics. Specifically, we measured the brain alignment between representations of a predictive version of a popular feedforward CNN often used as a computational model of the visual cortex–VGG16–and human EEG collected when viewing images that were relatively easy (Control) or difficult to classify (Challenge). We demonstrate that the recurrent dynamics significantly enhanced the model's alignment with EEG responses, underscoring the importance of recurrent connectivity in computational models of human vision, an effect distinctly visible for challenging stimuli.

## 1 Introduction

In neuroscience, predictive coding theory posits that the brain functions as a sophisticated hypothesis-testing system continuously generating predictions about sensory inputs and updating its internal world models to minimize prediction errors [1]. Within this theoretical framework, feedback connections from higher to lower level brain regions transmit the predictions of incoming stimuli [1].

Motivated by the goal of achieving human-like visual capabilities, researchers have developed frameworks that integrate predictive coding principles into computational models [2, 3, 4], demonstrating remarkable capacities for reproducing human-like perceptual behaviors. Moreover, previous investigations with predictive models like PredNets—LSTM-based models trained in an unsupervised fashion to anticipate video frame sequences—revealed that it captures key aspects in various forms of brain data (fMRI and MEG)[3, 5]. In a similar vein, Guo et al. [6] found that predictive coding-inspired dynamics improved the alignment to human fMRI responses during scene recognition.

While these, and various other studies suggest that integrating predictive top-down information leads to brain-like representations for natural images, their

ESANN 2026 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 22-24 April 2026, i6doc.com publ., ISBN 9782875870964. Available from http://www.i6doc.com/en/.

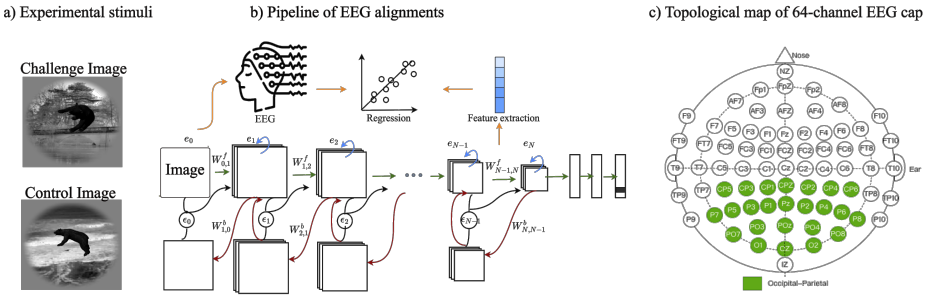


Fig. 1: **Experimental Design.** a) Experimental stimuli. The 'Challenge' set comprises images depicting objects with substantial degradation (occlusion, deformation, or missing parts). The 'Control' set consists of corresponding images with high visibility and minimal distortion. b) EEG-feature encoding pipeline. We integrated predictive coding units into a VGG16 backbone to construct a recurrent neural network (PVGG16). Three image sets-'Challenge', 'Control', and their combination ('Challenge & Control')-were processed through PVGG16 to extract feature representations from five hierarchical predictive coding stages (PCoders 1-5). A ridge regression model was trained to map these features onto EEG data, and model performance was evaluated using the Pearson correlation between predicted and actual EEG. c) Occipital-parietal electrodes (the green 24 channels) provide particularly strong visual signals.

ability to explain brain dynamics under challenging stimuli–where recurrence is hypothesized to play a prominent role–lacks a direct demonstration. In this study we directly examined whether the recurrent dynamics implemented in predictive models align with neural dynamics observed during challenging visual stimuli processing. Specifically, using ridge regression, we measure the alignment of the model representations with the EEG responses collected while the participants viewed images that are relatively easy (henceforth 'Control') and difficult (henceforth 'Challenge') for the model [7].

## 2 Methods

***Predictive coding model (PVGG16)*** Predify [2] integrates predictive coding dynamics into models by augmenting a pretrained classification model (here VGG16), with feedback connections trained in an unsupervised fashion. As illustrated in Figure 1, these feedback connections divide the model into a sequence of consecutive predictive coding (PCoder) modules (5 PCoders in PVGG16). The output of each PCoder, $e_N(t)$, over timesteps, incorporates input from feed forward, backward, and recurrent connections using the following equation:

$$e_N(t+1) = \underbrace{\beta_N [W^f_{N-1,N} e_{N-1}(t+1)]}_{\text{feedforward}} + \underbrace{(1-\beta_N-\lambda_N) e_N(t)}_{\text{memory}} - \underbrace{\alpha_N \nabla \epsilon_{N-1}(t)}_{\text{error}} + \underbrace{\lambda_N [W^b_{N+1,N} e_{N+1}(t)]}_{\text{feedback predictions}} \quad (1)$$

ESANN 2026 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 22-24 April 2026, i6doc.com publ., ISBN 9782875870964.
Available from http://www.i6doc.com/en/.

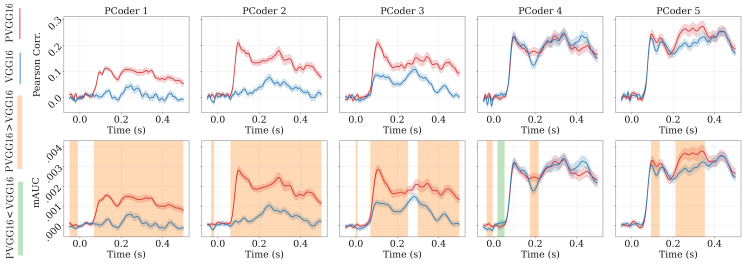


Fig. 2: **Whole-scalp EEG alignments, $\lambda_N = 0.5$, model iteration $n = 10$.** Upper panel shows pearson correlation between predicted and actual EEG signals while the lower panel demonstrates the Mean Area Under the Curve (mAUC). Red and blue lines represent values for PVGG16 and VGG16, respectively. The shaded regions denote the standard error for 34 subjects. Orange shaded regions represent time points where correlations with the predictive model are significantly higher than the corresponding feedforward (baseline) values.

where $\beta_N$, $\lambda_N$ $(0 \leq \beta_N + \lambda_N \leq 1)$, $\alpha_N$ are layer-specific coefficients controlling the weight of the feedforward, feedback and error-correction signals respectively. $\epsilon_{N-1}$ is the Mean Square Error between $e_{N-1}(t)$ and $W^b_{N,N-1}e_N(t)$ (the top-down prediction) at timestep $t$ and is used to train the feedback connections in-line with the predictive coding. We refer the readers to the original paper for additional details. For the sake of clarity, we restrict the use of the term 'timestep' for EEG data and use 'iteration' for model's recurrence instead.

***Experimental stimuli and human EEG recordings:*** To investigate recurrence in PVGG16, we employed the 'Challenge' and 'Control' image sets from Oyarzo et al. [7], derived from a 10-category stimulus set [8]. These sets were defined by comparing human (34 participants) and feedforward AlexNet classification performance: 'Challenge' images (N=121) depicted occluded, deformed, or partial objects where human accuracy exceeded the model's by $\geq 1.5d'$, whereas 'Control' images (N=121) showed minimal performance differences ($\leq 0.4d'$). More importantly, the authors reported that humans showed delayed information processing and specific recruitment of frontal brain regions for Challenge images. For our analysis, we used three conditions: 'Challenge', 'Control', and their combination ('Challenge + Control'). We aligned model representations with whole-scalp (64 channels) and occipital-parietal (24 channels) providing strong visual signals [9].

***Measuring EEG alignment:*** We extracted activations from the five PCoders in PVGG16 and their corresponding layers in the feedforward version of VGG16. To estimate the alignment between model activations and EEG signals, across iterations, we trained ridge regressions using nested cross-validation, and calculated the EEG encoding accuracy by measuring the pearson correlation between predicted and actual signals. The final scores were averaged across channels and subjects. We measured the mean Area Under the Curve (mAUC) using a 15ms

ESANN 2026 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 22-24 April 2026, i6doc.com publ., ISBN 9782875870964. Available from http://www.i6doc.com/en/.

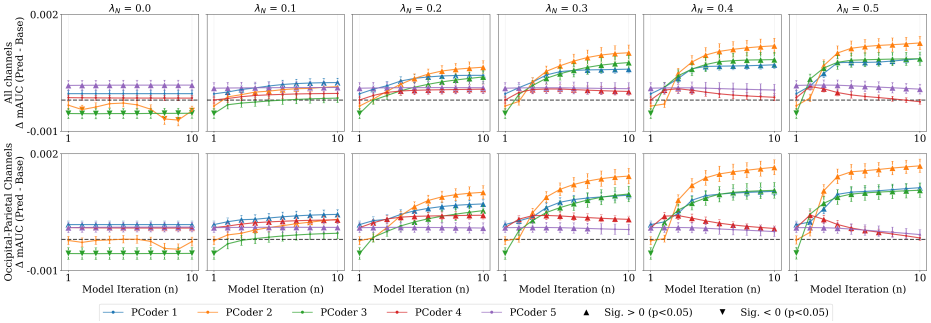


Fig. 3: $\Delta mAUC(predictive - baseline)$ **across model iterations n** under varying amounts of top-down ($\lambda_{\mathbf{N}}$) information. The top panel shows the alignment with whole-scalp and the bottom for occipital-parietal channels.

sliding window, inline with the 10-15 ms processing in the ventral stream[8], maintaining the fine-grained temporal resolution of neural dynamics.

## 3 Results

***'Challenge + Control'**: Increasing model recurrence and top-down influence ($\lambda_N$) progressively enhances whole-scalp EEG alignments.* Figure 2 presents the EEG encoding performance comparison between predictive VGG16 (PVGG16) and standard VGG16 architectures under $\lambda_N = 0.5$ (see Eq. 1) and model iteration n = 10. The upper panel displays EEG encoding accuracy across time, while the lower panel illustrates the corresponding mAUC values. Orange shading in the lower panel denotes temporal intervals where PVGG16 demonstrates statistically significant improvement in EEG alignment performance over VGG16 ($p < 0.05$), specifically at PCoder1–3.

Figure 3 examines the difference in mAUC between PVGG16 and VGG16 across whole-scalp and occipital-parietal configurations, revealing that increasing $\lambda_N$ from 0.0 to 0.5 produces larger improvements at early and middle PCoders (1–3). Conversely, late PCoders (4 and 5) show minimal benefits or performance degradation. While the lack of feedback to the highest layer (PCoder5) can explain its trendline, the alignment of PCoder4 indicates that the alignment is driven mostly by early visual cortex–it is possible that the features of a deeper layer like PCoder4 might have already peaked in their alignment, with additional recurrence only deteriorating or stagnating their performance. We validated our intuition by re-performing our analysis only on the occipital-parietal channels (Fig 3 bottom) and observing similar values for alignment across all layers.

***'Challenge' vs. 'Control':** with higher top-down influence, PVGG16 more comprehensively captures the divergent processing of challenge and control images inside brains between around 0.17s and 0.22s after stimuli onset.* In Figure 4, the upper panel shows occipital and parietal alignments with $\lambda_N = 0.5$ at model iteration $n = 10$, $\text{mAUC}_{\text{Challenge}}$ significantly exceeds $\text{mAUC}_{\text{Control}}$

ESANN 2026 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 22-24 April 2026, i6doc.com publ., ISBN 9782875870964. Available from http://www.i6doc.com/en/.

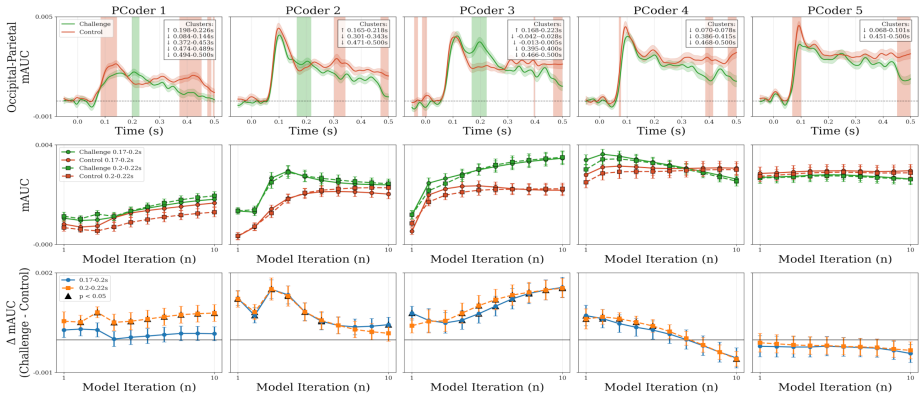


Fig. 4: **Influence of reccurrent dynamics on enhancing occipital-parietal alignments, $\lambda_N = 0.5$.** In the upper panel, green and orange lines represent mAUC values across EEG time points for the 'Challenge' and 'Control' conditions, respectively. Shaded regions indicate the standard error across subjects. Vertical green shading highlights time intervals where mAUC for the 'Challenge' condition significantly exceeds the 'Control' condition ($p < 0.05$). The bottom panel displays $\Delta\text{mAUC} = \text{mAUC}_{\text{Challenge}} - \text{mAUC}_{\text{Control}}$, where ▲ denotes statistically significant positive differences ($p < 0.05$), indicating model iteration n where recurrence improves EEG dynamic predictions for 'Challenge' stimuli.

between approximately 0.17–0.22 s at PCoder 1-3 (vertical green shading). This temporal window closely aligns with Oyarzo et al. [7] who identified divergent neural processing for challenge and control images in the human brain from 0.14–0.22 s. The middle panel illustrates that with $\lambda_N = 0.5$, mAUC generally increases with model iterations at PCoder 1–3 within both [0.17 s, 0.2 s] and [0.2 s, 0.22 s], indicating recurrent engagement during recognition of both conditions. Conversely, late PCoders exhibit declining mAUC with excessive recurrence, indicating stage-dependent effects.

We further asked whether additional model iterations increasingly contribute to the alignment with the brain data on challenge stimuli. We thus looked at the difference in mAUC scores and found significant differences within [0.2 s, 0.22 s] for PCoders 1–3 (Fig4, bottom), suggesting that recurrence enhances neural dynamic prediction specifically for challenging stimuli.

## 4 Discussion and Conclusion

This study provides direct evidence that predictive coding-inspired recurrent dynamics enhance the modeling of EEG activity, particularly for challenging visual stimuli, with gains scaling directly with top-down influence. These dynamics successfully capture the differential neural processing observed by [7].

Previous arguments for recurrent models in predicting brain data have relied on indirect evidence, such as late-onset decoding from natural stimuli or the alignment of deeper network layers. Our work offers the first direct demonstra-

ESANN 2026 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 22-24 April 2026, i6doc.com publ., ISBN 9782875870964. Available from http://www.i6doc.com/en/.

tion of their utility for explaining brain activity elicited by degraded stimuli, while also proposing a viable candidate architecture for implementing these dynamics.

Although the moderate effect sizes may reflect constraints of our EEG dataset (121 images), and our focus on PVGG16 architectures leaves open questions about generalizability across other predictive coding implementations, our work provides empirical support for predictive coding theory, demonstrating that recurrent mechanisms are engaged during core object recognition and become increasingly critical for capturing neural dynamics of demanding perceptual tasks. This research contributes to bridging computational models of predictive coding with the brain's own predictive processing mechanisms.

***Acknowledgements*** This project was funded by the German Research Foundation (DFG) - DFG Research Unit FOR 5368 (GR) awarded to GR, Deutsche Forschungsgemeinschaft (DFG; CI241/1-1, CI241/3-1, and CI241/7-1) awarded to RMC, and a European Research Council (ERC) starting grant (ERC-2018-STG 803370) awarded to RMC. We are grateful for access to the computing facilities of the Center for Scientific Computing at Goethe University and Freie universität Berlin. M. Guo is supported by a PhD stipend from the China Scholarship Council (CSC).

## References

[1] Rajesh Rao and Dana Ballard. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature neuroscience*, 2(1), 1999.

[2] Bhavin Choksi, Milad Mozafari, Callum Biggs O'May, Benjamin Ador, Andrea Alamia, and Rufin VanRullen. Predify: Augmenting deep neural networks with brain-inspired predictive coding dynamics. *Advances in Neural Information Processing Systems*, 2021.

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

[4] Beren Millidge, Mufeng Tang, Mahyar Osanlouy, Nicol Harper, and Rafal Bogacz. Predictive coding networks for temporal prediction. *PLOS Computational Biology*, 2024.

[5] William Lotter, Gabriel Kreiman, and David Cox. A neural network trained for prediction mimics diverse features of biological neurons and perception. *Nature machine intelligence*, 2(4):210–219, 2020.

[6] Manshan Guo, Michael Samjatin, Bhavin Choksi, Sari Sadiya, Radoslaw Cichy, and Gemma Roig. Predictive coding dynamics enhance model-brain similarity. In *Proceedings of the European Symposium on Artificial Neural Networks (ESANN)*, 2025.

[7] Pablo Oyarzo, Johannes JD Singer, Kohitij Kar, Diego Vidaurre, and Radoslaw M Cichy. Adaptive recruitment of cortex-wide recurrence for visual object recognition. *bioRxiv*, pages 2025–10, 2025.

[8] Kohitij Kar, Jonas Kubilius, Kailyn Schmidt, Elias B Issa, and James J DiCarlo. Evidence that recurrent circuits are critical to the ventral stream's execution of core object recognition behavior. *Nature neuroscience*, 22(6):974–983, 2019.

[9] Manshan Guo, Bhavin Choksi, Sari Sadiya, Alessandro Gifford, Martina Vilas, Radoslaw Cichy, and Gemma Roig. Limited but consistent gains in adversarial robustness by co-training object recognition models with human eeg. In *European Conference on Computer Vision*, pages 245–255, 2024.