Improving your vision model’s robustness with Predify

https://github.com/miladmozafari/predify

Enables you to easily extend your deep neural network with predictive coding dynamics.

Link to the preprint: https://arxiv.org/abs/2106.02749

1. Some theoretical background – Rufin VanRullen
2. Hands-on tutorial – Milad Mozafari (& Bhavin Choksi)
Deep predictive coding for more robust, human-like vision

**Predify: Augmenting deep neural networks with brain-inspired predictive coding dynamics**

Bhavin Choksi, Milad Mozafari, Callum Biggs O'May, Benjamin Ador, Andrea Alamia, Rufin VanRullen

**On the role of feedback in visual processing: a predictive coding perspective**

Andrea Alamia, Milad Mozafari, Bhavin Choksi, Rufin VanRullen

**Predictive coding feedback results in perceived illusory contours in a recurrent neural network**

Zhaoyang Pang, Callum Biggs O'May, Bhavin Choksi, Rufin VanRullen
1. Deep Predictive Coding Architecture

Deep predictive coding for more robust, human-like vision

\[ e_n(t + 1) = \beta \left[ W_{n-1,n}^f e_{n-1}(t + 1) \right]_+ + \lambda d_n(t) + (1 - \beta - \lambda) e_n(t) - \alpha \nabla e_{n-1}(t) \]

\[ d_n(t) = \left[ W_{n+1,n}^b e_{n+1}(t) \right]_+ \]

\[ e_{n-1}(t) = \| e_{n-1}(t) - d_{n-1}(t) \|_2^2 \]
Deep predictive coding for more robust, human-like vision

1. Deep Predictive Coding Architecture

**Predify** software (for PyTorch):
https://github.com/miladmozafari/predify

We *predified*: VGG16, EfficientNetB0, ResNet18…

\[
e_n(t + 1) = \beta [W_{n-1,n}^f e_{n-1}(t + 1)]_+ + \lambda d_n(t) + (1 - \beta - \lambda) e_n(t) - \alpha \nabla e_{n-1}(t)
\]
2. Robustness to noise

ImageNet accuracy:

(a) PVGG16

(b) PEfficientNetB0

Higher is better

---

input

$t=10$
3. Robustness to adversarial attacks

Predictive coding iterations decrease the efficacy of attacks:

- **L∞ BIM**
  - PVGG16
  - PEfficientNetB0

- **L2 RPGD**
  - PEfficientNetB0

- **L∞ HopSkipJump**
  - PEfficientNetB0

The graphs show the success rate of attacks against different models and perturbation sizes. Lower values indicate better robustness against adversarial attacks.
Deep predictive coding for more robust, human-like vision

4. Illusory contour perception

Kanizsa shapes:

- Humans see a white shape on black disks
- Computer vision models (DCNNs) typically don’t

→ Can predictive coding help?
Deep predictive coding for more robust, human-like vision

4. Illusory contour perception

3-layer PC model:

Pre-trained for (unsupervised) image reconstruction on CIFAR-100 natural image dataset
Deep predictive coding for more robust, human-like vision

4. Illusory contour perception

The model tends to classify the “illusory square” configuration as an actual square!

- $\sigma = 0$
- $\sigma = 0.1$
- $\sigma = 0.2$
- $\sigma = 0.3$
Deep predictive coding for more robust, human-like vision

Just like humans, the model “thinks” the square has a different luminance than the background!
Deep predictive coding for more robust, human-like vision

1. Brain-inspired Architecture
2. Robustness to image corruptions
3. Robustness to adversarial attacks
4. Predictive coding networks “perceive” illusory contours like we (humans) do
Deep predictive coding for more robust, human-like vision

Predify

- PyTorch-based Python package
- Automatically and simply adds predictive coding dynamics to any network
- Provides a super simple text-based user interface (TOML files)
- Code Generation

https://github.com/miladmozafari/predify