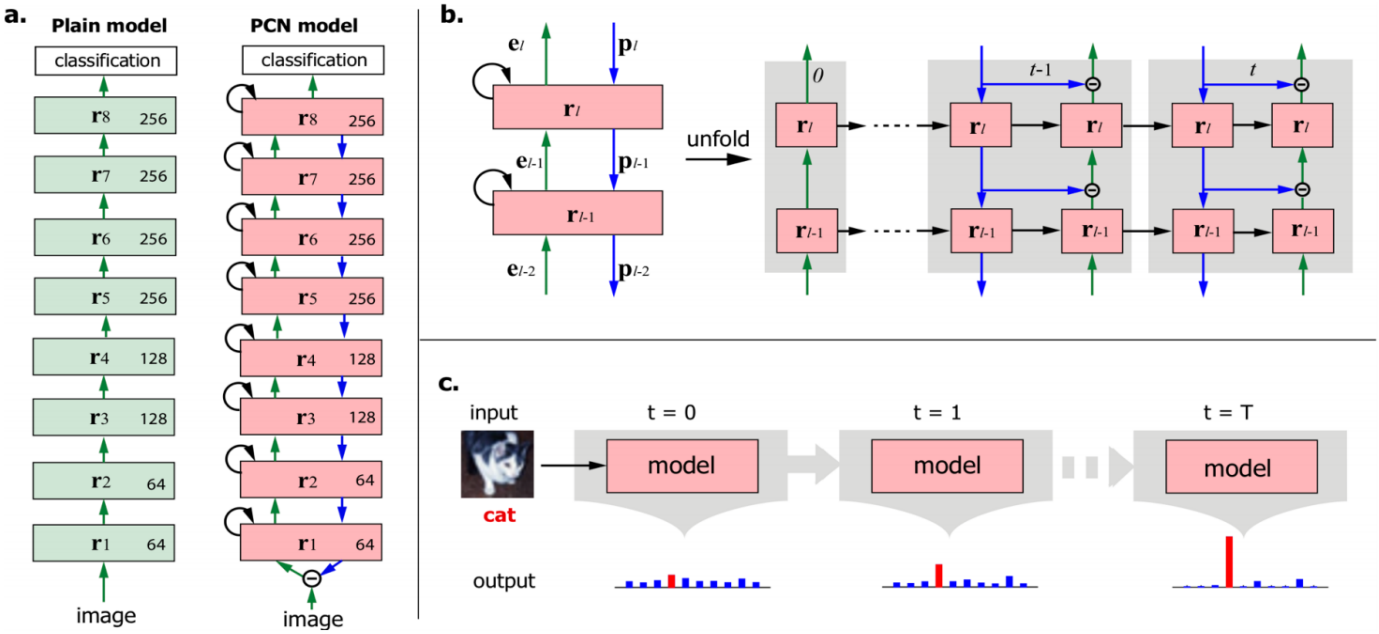
Review briefly predictive encoding and create a predictive encoding model on CIFAR10 and attack it to verify its robustness.

This task said to create a predictive encoding model on CIFAR10, and attack it to verify its robustness.

So first, i saw this url, a-new-kind-of-deep-neural-networks.
Then i know that, PredNet is a deep convolutional recurrent neural network inspired by the principles of predictive coding.
But it is trained for next-frame video prediction not the image.
And then i find a paper named [1], they propose a predictive coding networks(PCN) for Object Recognition.
The PCN is used for object recogntion, so i have a implement with it on CIFAR10.

The network structure is like the next figure.



So this network have 9 layers, like a.
b is Two-layer substructure of PCN.Feedback (blue), feedforward (green), and recurrent (black) connections convey the top-down prediction, the bottom-up prediction error, and the past information, respectively.

The next is the Algorithm for the PCN network.

Algorithm 1 Deep Predictive Coding Network

Input static image: $\mathbf{x}$
$\mathbf{r}_0(t) \leftarrow \mathbf{x}$
% initialize representations
for $l = 0$ **to** L-1 **do**
  $\mathbf{r}_{l+1}(0) \leftarrow \mathrm{ReLU}\left(\mathrm{FFConv}(\mathbf{r}_l(0))\right)$
% recurrent computation with T cycles
for t = 1 to T **do**
  % nonlinear feedback process
  for l = L **to** 1 **do**
    $\mathbf{p}_{l-1}(t\text{-}1) \leftarrow \mathrm{FBConv}\left(\mathbf{r}_l(t\text{-}1)\right)$
    if $l > 1$ **do**
      $\mathbf{r}_{l-1}(t\text{-}1) \leftarrow \mathrm{ReLU}\left((1\text{-}b)\mathbf{r}_{l\text{-}1}(t\text{-}1) + b\mathbf{p}_{l\text{-}1}(t\text{-}1)\right)$
  % nonlinear feedforward process
  for $l = 0$ **to** L-1 **do**
    $\mathbf{e}_l(t) \leftarrow \mathbf{r}_l(t) - \mathbf{p}_l(t\text{-}1)$
    $\mathbf{r}_{l+1}(t) \leftarrow \mathrm{ReLU}\left(\mathbf{r}_{l+1}(t\text{-}1) + a\ \mathrm{FFConv}(\mathbf{e}_l(t))\right)$
% classification
Output $\mathbf{r}_{\mathrm{L}}(\mathrm{T})$ for classification

These figure from [1].

And then, like task B3, i use the adversarial attack to verify its robustness.
The training log is saved in ./log folder, and the adversarial attack result is saved in the same folder.

I train the PCN in 250 epochs, and adversarial attack epsilons in [0, .05, .1, .15, .2, .25, .3, .5, 1.]

Reference

1. Wen H, Han K, Shi J, et al. Deep predictive coding network for object recognition[C]//International Conference on Machine Learning. PMLR, 2018: 5266-5275.