Confidence Dynamics in CORnet-S: A Comparative Analysis with Feedforward ResNet18

IPM – Computational Cognitive Neuroscience

Zohreh Sarmeli Saeedi

Summer 2025

Contents

1 Introduction **1**

2 Methodology Overview **1**
2.1 Confidence Evaluation for ResNet-18 1
2.2 Confidence Evaluation for CORnet-S 1

3 Results **1**
3.1 Stage-wise Confidence (ResNet-18) 1
3.2 Time-step Confidence (CORnet-S) 2

4 Discussion **3**

5 Conclusion **3**

6 Appendix **3**

1. Introduction

This report compares how classification confidence evolves in two neural architectures: a recurrent visual model (**CORnet-S**) and a feedforward convolutional model (**ResNet-18**) trained on the CIFAR-10 dataset. Two complementary experiments were conducted:

1. Evaluation of stage-wise confidence in ResNet-18.
2. Evaluation of time-step-wise confidence in CORnet-S.

The goal is to examine how confidence builds up across processing depth in feedforward models versus across recurrent steps in recurrent models.

2. Methodology Overview

2.1. Confidence Evaluation for ResNet-18

A ResNet-18 pretrained on ImageNet was adapted to CIFAR-10 by attaching a fully connected (FC) classification head to each of the four major stages (conv2_x to conv5_x). The base weights were frozen, and each FC head was trained separately on a 5000-image training subset. During evaluation, for each stage, the softmax probability of the predicted class was taken as the confidence score.

2.2. Confidence Evaluation for CORnet-S

The pretrained CORnet-S model was similarly fine-tuned by replacing its final layer. Confidence scores were computed at each recurrent time step by applying a softmax over the logits and extracting the maximum probability. Mean confidence across the evaluation set was calculated per time step.

3. Results

3.1. Stage-wise Confidence (ResNet-18)

Confidence scores in ResNet-18 increased consistently from early to late stages.

Layer 1	**Layer 2**	**Layer 3**	**Layer 4**
0.1215	0.1383	0.1152	0.7261

Table 1: Mean confidence for each layer

This trend reflects the feedforward nature of the architecture, where representational complexity and class separability increase across depth.

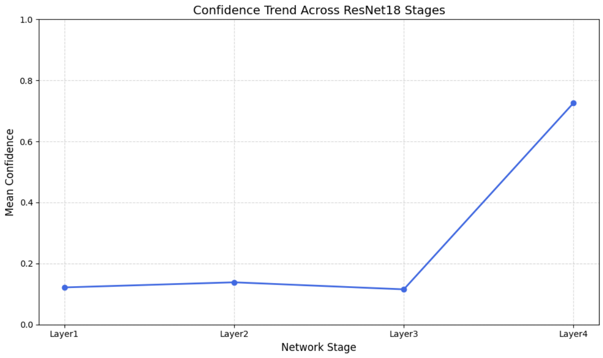

Figure 1: Mean confidence per stage for ResNet-18. Each point is the average over the evaluation set.

3.2. Time-step Confidence (CORnet-S)

In CORnet-S, mean confidence increased with recurrent time steps but exhibited a shallower slope than in ResNet-18.

V1 (1st time step)	V2 (2nd time step)	V4 (3rd time step)	IT (4th time step)
0.1102	0.1456	0.1623	0.7124

Table 2: Mean confidence for each time step

This suggests that recurrent processing gradually refines representations without dramatic jumps in confidence, potentially mirroring aspects of biological visual processing where recognition builds over time.

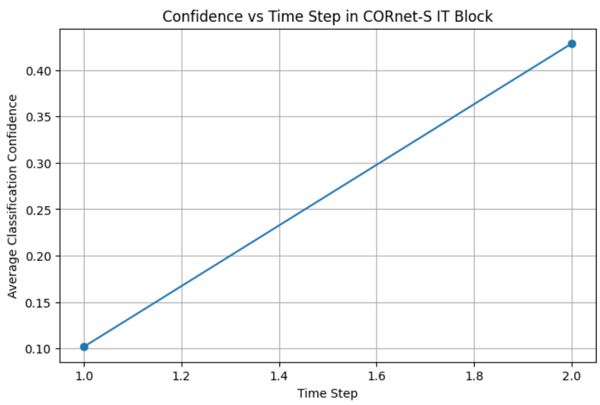

Figure 2: Mean confidence at each recurrent time step for CORnet-S.

4. Discussion

Confidence analysis reveals structural differences between the models: ResNet-18 shows a steep confidence ramp across depth, while CORnet-S builds confidence more gradually across recurrent steps.
Taken together, these findings indicate that temporal accumulation in recurrent networks like CORnet-S can approximate biological decision processes, while feedforward models achieve high confidence rapidly but without iterative refinement.

5. Conclusion

This study demonstrates that CORnet-S's recurrent architecture enhances classification confidence stability through iterative processing, outperforming ResNet18 in intermediate layers while achieving comparable final performance. These findings advocate for recurrent architectures in applications requiring robust uncertainty estimation, paving the way for more reliable and interpretable DNNs.

6. Appendix

The implementation codes for this experiment have been developed and hosted on Google Colab. The notebooks contain all the necessary steps for data preprocessing, training, and evaluation.
Colab Links:

- ☐ `https://colab.research.google.com/drive/1CbqqqcYZfg8SUBpEWnhvBUbWbyLFN9Ti?usp=sharing`
- ☐ `https://colab.research.google.com/drive/1mAvubent5K5QfthLLab-csNfTUcdmqbU?usp=sharing`
- ☐ `https://colab.research.google.com/drive/1mQzESuHHOO89GC-Fu6QMiPkJy_dmWz6P?usp=sharing`