Internal Neural Noise Progression for Emergent Classification Robustness

Suayb S. Arslan (sarslan@mit.edu) Dept. of Brain and Cog. Sciences, MIT Cambridge, MA, 02139, USA

Hojin Jang (hojin4671@korea.ac.kr) Dept. of Brain and Cog. Eng. Korea University, Seoul, South Korea

Pawan Sinha (psinha@mit.edu) Dept. of Brain and Cog. Sciences, MIT Cambridge, MA, 02139, USA

Abstract

Mammalian perceptual development follows a largely consistent progression along several dimensions. Acuity and chromatic sensitivity improve over the first several months of life after birth. Recent evidence suggests that the progressions may have adaptive value by inducing the formation of receptive field structures that enable later resilience to spatial or chromatic degradations. Here we examine whether the developmental changes in neural noise may follow a similar logic, i.e. does the temporal progression in neural noise lead to benefits in classification performance, especially under challenging conditions? Our preliminary results in two distinct experimental settings indicate that progressions from high network noise to lower levels lead to phenomena similar to those associated with stochastic resonance. These findings not only provide a potential teleological account of noise progressions in biological systems, they also suggest useful training regimens for artificial vision systems to improve their robustness.

Keywords: Noise training, biomimetic training, brain noise progression, stochastic resonance

Introduction

The brain experiences two primary types of noise. The first type is internal noise, which arises from the inherent variability in cellular functions including fluctuations in membrane potentials and variations in ion transmissions. The second type is external noise, which originates from the variability in external conditions that influence sensory input, such as changes in the environment or imperfections in the sensory apparatus. Our hypothesis is that internally developed noise could be a source of emergence of mechanisms of the kind that might potentially help neural systems to combat against the external noise and help generalize over unseen stimuli.

Unlike linear systems, which are limited in their capacity to model multi-faceted behavioral dynamics and are subject to reduced information loss in presence of noise, non-linear systems such as brain has the ability to express multi-stability and explore different functional network configurations provided that enough internal noise is generated (McIntosh et al., 2010). Consequently, the concurrent existence of nonlinearity and internal noise become essential for the spontaneous emergence of exploratory dynamics even in the absence of external stimulation/noise.

Despite internal noise being an intrinsic component of neural dynamics, in numerous practical scenarios, it is influenced by external stimulation and noise. This interaction is particularly noteworthy due to the emergent behaviors that arise, also known as stochastic resonance (SR) in biological systems. SR occurs when a nonlinear system's signal-to-noise ratio improves at moderate noise intensities. Moderate noise helps the signal reach the threshold without dominating the original signal, eventually optimizing signal transmission (Gammaitoni, Hänggi, Jung, & Marchesoni, 1998). In our study, we have found evidence to observe similar behavior in neural networks as long as an appropriate noise level and progression is employed.

Biological Noise Progression

It is known that the brain noise changes with maturation and aging, correlating with the stable behavior. However, the literature has conflicting arguments about the direction of noise progression in biological systems. (Wang et al., 2019) has found that spiking noise is much lower in the spike trains of infant V2 neurons compared with those of adults, despite the reduced information density for V2 neurons in infants. However, (Skoczenski & Norcia, 1998) argued that intrinsic neural noise in neonates is approximately nine times higher than in adults, suggesting high-to-low internal noise progression. Similar observations has been reported in ASD patients as well (Davis & Plaisted-Grant, 2015; Dinstein, Heeger, & Behrmann, 2015). Although these studies compare aggregate relative noise levels, there seems to be no consensus on the way the noise is quantified at the neuron-level.

Objective

In this work, we present some of the preliminary results regarding the time course of internal noise consolidation in neural networks. Objective is to explore two distinct ways of noise injection in a progressive way to demonstrate the emergence of biologically plausible mechanisms such as stochastic resonance and its positive effects on the generalization performance under degraded and challenging stimuli.

Experiments and Discussions

We have assumed two kinds of internal noise: First one is independent of the input stimuli, governed solely by intrinsic mechanisms. The latter is dependent on the input through the signal power available at the output of each neural unit. This way we have quantified the amount of noise based on signal-to-noise ratio (SNR) variations for each of the processed batch

Exp.	Network	Noise Type	Biomimetic	Datasets
1	AlexNet	Nueron-lvl	Resolution, Noise	CIFAR10, Fashion MNIST
2	AlexNet	Batch-SNR	Noise	ImageNet

Table 1: A summary of the two experiments.

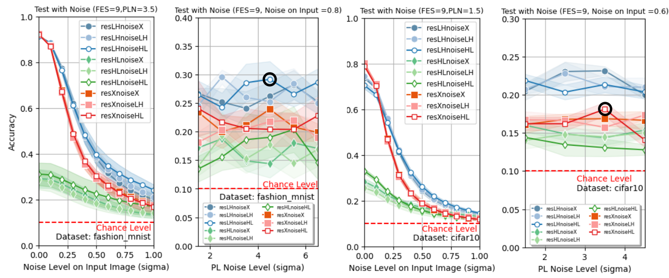

Figure 1: Test accuracy performance of joint bio-mimetic progressions for a specific set of parameter selections.

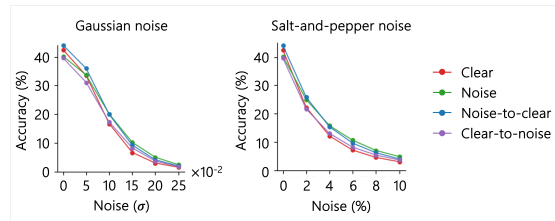

Figure 2: Test accuracy performance as a function of Gaussian noise (left) and salt-and-pepper noise (right). Four training regimes were compared.

of input images. The architectures, noise types, biomimetic progression details as well as datasets used are summarized in Table 1. The input layer of the networks are adjusted to fit the available image resolution.

Input-Independent Internal Noise

We have considered nine different combinations resulting from the product space of resolution (α) and noise (β) having three distinct progression directions, namely $\alpha, \beta \in \{\mathrm{LH}, \mathrm{HL}, \mathrm{X}\}$ where L is low, H is high and X represents no change, encoded in legends as "resαnoiseβ". In our simulations we have simply treated high resolution to be the clear images and low noise and no noise cases to be equivalent. The parameter FES denotes the standard deviation σ_R of the Gaussian kernel used to generate low resolution images whereas PLN is the σ_{pl} of the independent additive Gaussian noise applied to penultimate layer of the network. Two phases are defined covering 15 epochs each, where the learning rate is fixed in the former and reduced in the latter.

We observed in our experiment for a selection of variables ($\sigma_R = 9, \sigma_{lp} = 3.5$ for Fashion MNIST, $\sigma_{lp} = 1.5$ for CIFAR10) that although low-to-high resolution training provides robustness, training jointly with high-to-low noise progression provided even better generalization performance. Moreover, there seems to be a window of PL noise that helps networks achieve the best accuracy (e.g. $\sigma_{pl} = 4.5$ for resLHnoiseHL on Fashion MNIST and $\sigma_{pl} = 3.5$ for resXnoiseHL on CIFAR10) with/out noise consolidation. This seems to indicate a similar mechanism of stochastic resonance phenomenon for weakened (noisy) inputs. We also observe that the optimal parameter selections depend on the input despite the internal noise is independent, suggesting an input-dependent internal noise model (Mišic, Mills, Taylor, & McIntosh, 2010) which has been the motivation for the next experiment.

Input-Dependent Internal Noise

In this experiment, we introduced noise at every layer of the model, as internal noise in the biological brain would be present at every stage of visual processing. To ensure an equal amount of noise injection across layers, the variance of signals was dynamically measured within each batch, and random noise sampled from a Gaussian distribution with 25% of the variance was injected during the training. Four scenarios were compared: 1) "clear," involving training without any noise across all epochs; 2) "noise," involving training with noise across all epochs; 3) "noise-to-clear," where the first phase was conducted with noise, and the second phase without noise; and 4) "clear-to-noise," where the first phase was conducted without noise, and the second with noise. Each network was tested on inputs degraded by either Gaussian noise or salt-and-pepper noise. We hypothesized that injecting internal noise would enhance the networks' ability to generalize to conditions where input images were degraded, and further, by comparing the two training regimes, "clear-to-noise" and "noise-to-clear," we might gain insight into the benefit of noise progression direction regarding robustness.

Overall, our findings indicate that, although the effect was not substantial, when the network was initially trained with noise followed by clear conditions ("noise-to-clear"), it exhibited better generalization performance than the opposite direction (Figure 2). This directionality suggests that initially starting with noisy neural conditions may afford a better opportunity for developing more robust networks, potentially hinting at the early developmental progression of visual systems that might occur in the biological brain.

Discussions

Our initial investigations have led to conclude that the utilization of low spatial frequencies and internal noise, when applied progressively (high to low noise) during training, has the potential to enhance classification robustness under degraded input. It is important to note that our training protocol does not incorporate degraded (e.g. noisy) images, thus preventing the direct learning of degradation type and characteristics. While the progression from high to low noise may appear contrary to some of the existing studies that suggest noise levels and physiological variability increasing with age (McIntosh et al., 2010), our approach focuses on noise manipulation at the neuronal level. Furthermore, we have not considered the dynamic network changes, widely recognized within the neuroscience community as related to proliferation and pruning processes, which are integral to architectural optimizations. These dynamic alterations could potentially be modeled through techniques such as regularization and network adaptation methods involving growth and reduction.

Acknowledgments

This research is based upon work supported in part by the Office of the Director of National Intelligence (ODNI), Intelligence Advanced Research Projects Activity (IARPA), via [2022-21102100009]. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of ODNI, IARPA, or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright annotation therein.

References

Davis, G., & Plaisted-Grant, K. (2015). Low endogenous neural noise in autism. *Autism*, *19*(3), 351–362.

Dinstein, I., Heeger, D. J., & Behrmann, M. (2015). Neural variability: friend or foe? *Trends in cognitive sciences*, *19*(6), 322–328.

Gammaitoni, L., Hänggi, P., Jung, P., & Marchesoni, F. (1998). Stochastic resonance. *Reviews of modern physics*, *70*(1), 223.

McIntosh, A. R., Kovacevic, N., Lippe, S., Garrett, D., Grady, C., & Jirsa, V. (2010). The development of a noisy brain. *Archives italiennes de biologie*, *148*(3), 323–337.

Mišic, B., Mills, T., Taylor, M. J., & McIntosh, A. R. (2010). Brain noise is task dependent and region specific. *Journal of Neurophysiology*, *104*(5), 2667–2676.

Skoczenski, A. M., & Norcia, A. M. (1998). Neural noise limitations on infant visual sensitivity. *Nature*, *391*(6668), 697–700.

Wang, Y., Zhang, B., Tao, X., Shen, G., Smith, E. L., & Chino, Y. M. (2019). Spiking noise and information density of neurons in visual area v2 of infant monkeys. *Journal of Neuroscience*, *39*(29), 5673–5684.