Improving image classification robustness with stochastic resonance in a recurrent layer

Siegfried M. Ludwig[1]

[1]Theoretical Foundations for Cognitive Agents, Radboud University

July 2019

Abstract

Stochastic resonance describes the utility of noise in improving the detectability of weak signals in certain types of systems. It has been observed widely in natural and engineered settings, but its utility in image classification with rate-based neural networks has not been studied extensively. In this analysis a simple LSTM recurrent neural network is trained for digit recognition and classification. During the test phase, image contrast is reduced to a point where the model fails to recognize the presence of a stimulus. Controlled noise is added to partially recover classification performance. The results indicate the presence of stochastic resonance in rate-based recurrent neural networks.

1 Introduction

This paper investigates the occurrence of stochastic resonance (SR) in rate-based artificial neural networks (ANN), with SR describing the improved detectability of weak signals with the introduction of noise in certain types of systems. A recurrent neural network (RNN) is trained to recognize the presence of a stimulus and classify it into digits. At test time, the input intensity is reduced to a sub-threshold level and controlled noise is added to induce SR. The Python implementation is publicly available on GitHub[1].

Considerable research has been done on the occurrence of SR in various types of natural and engineered systems (see [6] for a recent review) and in spiking-based neural networks (SNN). SNNs in their current state have many drawbacks however, which is the reason why rate-based models are much more popular in practice.

While benefits of general noise introduction have been studied in rate-based models, including deep learning models, those studies focused on the broader benefit of noise, particularly to make models robust against noise itself and to combat overfitting [9]. Literature based on narrower definitions of SR relating to signal detectability is scarce for rate-based ANN and its applications in deep learning. In the following, the basic concept of more narrowly defined SR will be presented and its occurrence in natural neural systems will be introduced. Building on this base, the literature on SR in ANN will be introduced.

1.1 Stochastic resonance

SR is the benefit of optimal levels of noise for the detectability of weak input signals in nonlinear systems with some sort of threshold [4]. Multiple experiments have shown the benefit of optimal noise levels for detection of weak input signals in animals and humans, including the detection of sub-threshold visual stimuli (see [1] for a recent review).

[1] https://github.com/SMLudwig/lstm-stochastic-resonance

A typical curve of the output performance at different levels of noise of a system capable of exhibiting SR is given in figure 1. Performance follows a bell-shaped curve with an optimal level of noise, below and above which performance diminishes. This behavior justifies the term stochastic resonance.

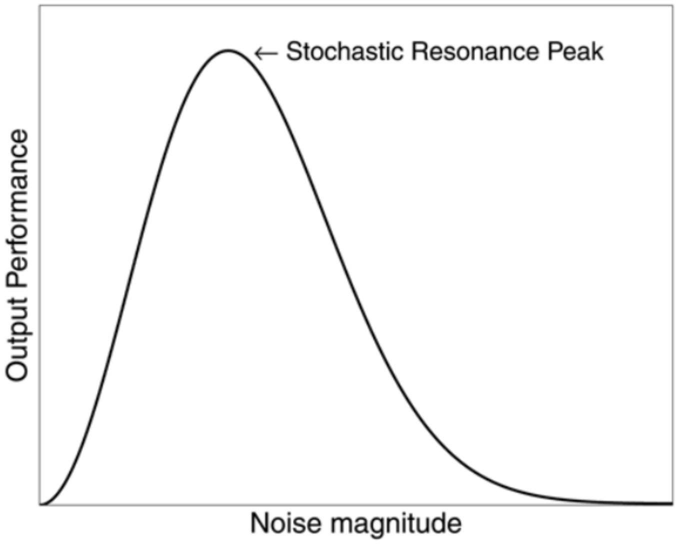

Figure 1: Typical output performance versus noise level in systems capable of stochastic resonance. Adapted from [6].

Since ANN are at least loosely based on natural neural networks, the aforementioned findings suggest that introducing noise to the input of ANN could increase signal detection performance.

1.2 SR in artificial neural networks

Reviewing the above mentioned definition of SR reveals the two conditions of nonlinearity and the presence of some threshold. ANN are highly nonlinear systems, a feature built in purposefully to give them universal function approximation power [2]. Non-linearity is usually introduced with activation functions like the sigmoid, tanh or ReLu, which are applied to the weighted and summed inputs to each neuron. While spiking neural networks clearly exhibit thresholding behavior, the situation is less clear for rate-based ANN, which by definition do not act on thresholds as straightforwardly. Nevertheless, both by introducing certain activation functions such as rectified linear units (ReLU), which is zero for negative inputs and linear for positive inputs, and by introducing a no-signal output class, threshold-like behavior can be implemented. SR could therefore be observed in certain rate-based ANN. Indeed there is theoretical mathematical work on SR in ANN, both spiking and continuous [8].

Reviewing the literature reveals the occurrence of SR in Spiking neural networks (SNN) [3], which are much more closely modelled on the human brain than rate-based ANN, with the latter including the deep learning models which recently have enjoyed much success. SR has also been observed in certain kinds of recurrent neural networks (RNN) [5, 7]. However, literature on more narrowly defined SR as noise-benefit in signal detection in the widely used long short-term memory (LSTM) RNN model is scarce. SR benefits for common rate-based models on common application tasks like digit classification is also scarce. Literature on SR in feed-forward ANN without recurrent connections is scarce, which is likely due to weak theoretical arguments for the occurrence of SR in those models. Since they only run inputs once through the network and lack recurrent connections, no temporal effects can emerge and the emergence of SR in the general form is therefore less conceivable.

1.3 Research aim and relevance

The aim of the project is to answer the question of whether and under which conditions the introduction of controlled noise can lead to SR and thereby improved signal detection in LSTM RNN. Due to the nonlinear and temporal nature of RNN and building on successes in previous literature on other kinds of ANN and RNN specifically (see introduction), SR is hypothesized to occur in LSTM RNN. Furthermore, the effect of noise on classification performance beyond mere signal detection on weak inputs will be studied as measured in the digit classification accuracy.

SR has not been studied satisfactorily in rate-based neural networks, even though they are by far the dominant network architecture, underlying deep learning systems that enjoy massive popularity. Since natural stimuli rarely resemble clear images with high contrast, robustness against weak input signals has practical utility. Potential performance enhancements and increased robustness resulting from the introduction of SR phenomena could therefore lead to many practical benefits.

2 Method

The model consists of a single LSTM layer with 20 hidden units and ReLU activation, followed by a dense layer with softmax activation. The ReLU activation theoretically serves as an additional thresholding function, since only values above zero do not get clipped to zero by the ReLU function. This setup roughly corresponds to a simple multilayer perceptron in the case of a sequence length of one.

The model will be trained on the MNIST digit dataset with a manually added empty stimulus class. For the empty class, black images and an additional label are added. During training, the model gets visual inputs with normal signal intensity and has to classify whether an input digit is present and if so, which digit it is. This results in a classification problem with eleven classes, ten for the input digits and an eleventh for no input. During test time, the input will be reduced to sub-threshold intensity by multiplying with a number (thresholding factor), leading the model to wrongly classify inputs as containing no signal (see figure 6).

The model is not trained on any thresholding or noise influence, since this might skew results and for practical applications it is more useful if noise can just be added during test phase, allowing the use of pre-trained models.

Following this state, controlled noise of different types and levels is introduced to the input and changes in the signal detection rate and the digit classification accuracy are observed. Uniform noise is sampled from a range between zero and a maximum noise level determined via grid search. Gaussian noise is sampled from a normal distribution with zero mean and standard deviation set by the noise level parameter during grid search. For Gaussian noise, noise values are clipped to be equal or greater to zero to avoid negative noise, which would remove parts of the stimulus even further.

The effect of different sequence lengths is studied. Theoretically longer sequences should provide the model with more information since at every step of the sequence, new noise is added to the original image. This potentially reveals different parts of the stimulus at each step.

3 Results

Results are reported as mean and standard deviations of the model accuracy for each setup across five independently trained models. Figures 2 and 3 show a clear positive effect of noise on performance under certain conditions, resembling the typical curve of stochastic resonance (see figure 1).

It can be observed that the classification of lower contrast stimuli (higher thresholding factor) benefit from higher noise levels, while higher contrast stimuli benefit from lower noise levels. Uniform (figure 2) and Gaussian noise (figure 3) show very similar results, with the latter showing best performance for slightly higher noise levels. This difference is expected since Gaussian noise with mean at zero as in this analysis is on average smaller than uniform noise.

Running the experiment without the added no-signal class destroys the SR effect (figure 4). This means that the threshold-like behavior of the no-signal class is essential to the occurrence of the SR

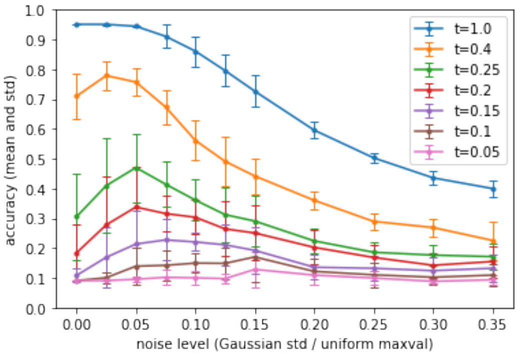

Figure 2: Uniform noise: mean and standard deviation of validation accuracy with different thresholding factors and uniform noise levels. t=1.0 corresponds to the original image.

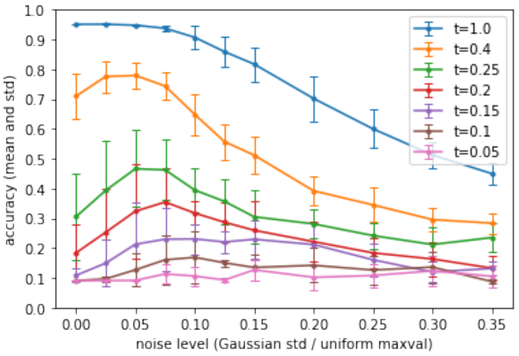

Figure 3: Gaussian noise: mean and standard deviation of validation accuracy with different thresholding factors and Gaussian noise levels. t=1.0 corresponds to the original image.

phenomenon. ReLU activation functions do not seem to be sufficient to introduce the desired effect.

An analysis of different sequence lengths reveals almost no effect, with longer sequences even having somewhat lower performance (figure 5). This is likely due to the fact that the model is not trained on any noise and therefore gets identical inputs at each sequence step during training. It can not adapt to making use of the small variations across the sequence that the noise introduces. Further research could determine the utility of training a sequence model with some level of noise and applying it to sub-threshold stimuli with added noise. Reported in the figure are mean and standard deviations of the validation accuracy over five independent iterations for each sequence length.

Figure 6 shows examples of the effect of noise on model performance on validation data. The first image of each row is the original image resembling the images the model is trained on. The model performance is very high on these images with accuracies around 95%. The second image of each row is the watered down version, achieved by multiplying the image with a thresholding factor. The model usually classifies these images as containing no stimulus, which means the stimulus is below the threshold internal to the model. The last image of each row shows the watered down version with added noise. It can be seen that the model usually recognizes that some stimulus is present and to some degree accuracy also recovers, although not to the full level.

4 Discussion

The LSTM model trained on digit recognition with an added class of empty stimuli exhibited clear SR behavior. Adding noise to sub-threshold stimuli partially recovered classification performance, with the optimal noise level being dependent on the stimulus intensity. This shows that rate-based recurrent neural networks do exhibit SR and can be made more robust against weak stimuli by making use of the SR phenomenon.

Interestingly, a sequence length of one exhibits at least equal performance to longer sequences (figure 5). This means that the recurrent nature of the model studied here is not a prerequisite for the occurrence of SR. Other models, like convolutional neural networks, could therefore also exhibit SR. This is a promising direction for future research.

The dependence of the optimal noise level on stimulus intensity presents a problem in practice, since the stimulus intensity can not be controlled in the way it is in this experiment. Potential ways do deal with this problem would be to preprocess stimuli or to add noise adaptively depending on the stimulus intensity. An intersting way could be to make use of the recurrent architecture by introducing varying noise levels at each step of the sequence, which could allow the model to pick the optimal noise level itself.

The results also show that running the experiment without the added no-signal class gives better

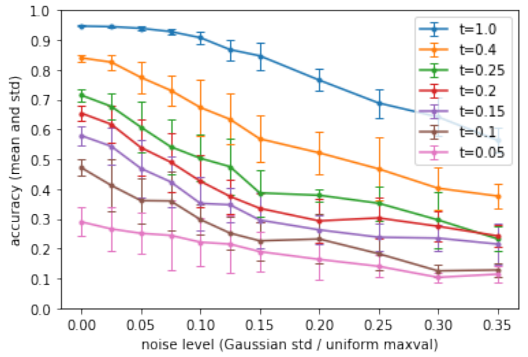

Figure 4: Without no-signal class: mean and standard deviation of validation accuracy with different thresholding factors and uniform noise levels. t=1.0 corresponds to the original image.

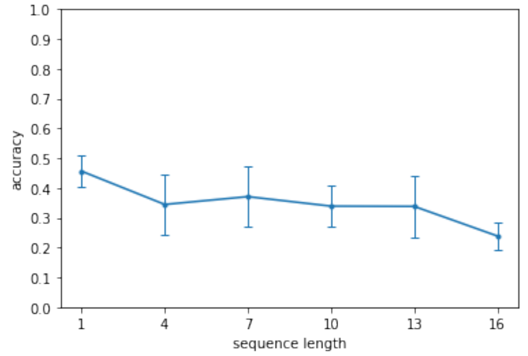

Figure 5: Mean and standard deviation of validation accuracy for different LSTM sequence lengths (t-factor=0.15, uniform noise level=0.075).

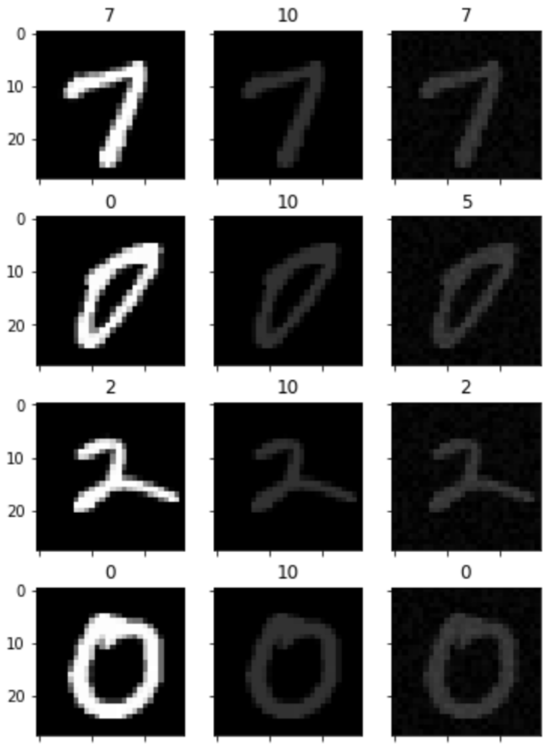

Figure 6: Examples of the effect of noise on model validation performance (t-factor=0.2, uniform noise level=0.05). Model predictions are given as numbers above each image, with 10 representing the empty class (no stimulus).

performance even though no SR effect is exhibited. On the type of controlled task the MNIST dataset represents, this is an argument against making use of noise to introduce SR. On object recognition tasks and in practical applications however, instances with absence of stimuli are to be expected and the no-signal class is a natural part of the problem formulation. Under those more practical conditions, SR effects could improve performance.

References

[1] Takatsugu Aihara et al. “How does stochastic resonance work within the human brain?–Psychophysics of internal and external noise”. In: *Chemical Physics* 375.2-3 (2010), pp. 616–624.

[2] Balázs Csanád Csáji. “Approximation with artificial neural networks”. In: *Faculty of Sciences, Etvs Lornd University, Hungary* 24 (2001), p. 48.

[3] Daqing Guo and Chunguang Li. “Stochastic and coherence resonance in feed-forward-loop neuronal network motifs”. In: *Physical Review E* 79.5 (2009), p. 051921.

[4] Peter Hänggi. “Stochastic resonance in biology how noise can enhance detection of weak signals and help improve biological information processing”. In: *ChemPhysChem* 3.3 (2002), pp. 285–290.

[5] Patrick Krauss et al. “Stochastic resonance in three-neuron motifs”. In: *arXiv preprint arXiv:1811.12091* (2018).

[6] Mark D McDonnell and Derek Abbott. “What is stochastic resonance? Definitions, misconceptions, debates, and its relevance to biology”. In: *PLoS computational biology* 5.5 (2009), e1000348.

[7] Christopher Monterola and Martin Zapotocky. “Noise-enhanced categorization in a recurrently reconnected neural network”. In: *Physical Review E* 71.3 (2005), p. 036134.

[8] Ashok Patel and Bart Kosko. “Stochastic resonance in continuous and spiking neuron models with Levy noise”. In: *IEEE Transactions on Neural Networks* 19.12 (2008), pp. 1993–2008.

[9] George Saon et al. “Sequence Noise Injected Training for End-to-end Speech Recognition”. In: *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE. 2019, pp. 6261–6265.