Feedback neural networks best explain human object recognition on degraded images

Vincent Roest (vincent.roest@student.auc.nl)
Amsterdam University College
Amsterdam, The Netherlands

Kandan Ramakrishnan (K.Ramakrishnan@uva.nl)
University of Amsterdam
Amsterdam, The Netherlands

Abstract

Feedforward neural networks are currently the dominant model to understand human object recognition. The visual cortex has more feedback connections than feedforward, which enable humans to robustly recognize degraded objects. A computational model of feedback has not yet been evaluated as a model for human object recognition. In this work, we test the hypothesis that a feedback neural network correlates closely to human behaviour and outperforms feedforward networks in recognition of degraded objects. A number of architectural extensions to the standard feedforward network is proposed with the introduction of feedback loops at different hierarchical levels. The computational models are evaluated on a noisy MNIST dataset and further compared against human behaviour to demonstrate its biological plausibility. Our results show that feedback models outperform feedforward model at different noise levels. Testing for human object recognition performance, increased reaction time is observed with increase in image noise level. Furthermore, feedback models have higher correlation to human performance compared to feedforward networks. Our results suggest that feedback networks are essential for object recognition and match human object recognition better than the standard feedforward models.

Keywords: feedback model;object recognition;behaviour

Introduction

Feedforward models such as Deep Neural Networks (DNNs) are state-of-the-art on automated object recognition (Krizhevsky, Sutskever, & Hinton, 2012). DNNs have achieved human level performance (He, Zhang, Ren, & Sun, 2015) and correlates to the hierarchy of visual representations in the human brain (Güçlü & van Gerven, 2015).

Visual processing in the human brain goes beyond feedforward connections, and feedback plays an important role in object recognition (Lamme & Roelfsema, 2000). One of the functions for feedback connections is to recognize objects degraded by noise or occlusion. This suggests that, the incorporation of feedback connections might improve object recognition performance compared to standard feedforward models in computer vision. Additionally, these models should match more closely to human performance.

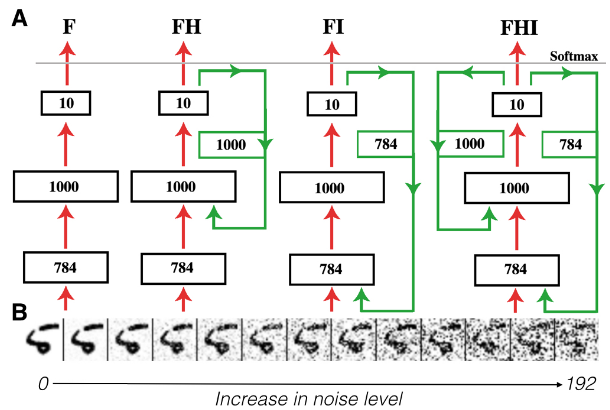

Figure 1: Computational models : A) We test a feedforward (F) and three variations of feedback (FH, FI and FHI) model on MNIST dataset. B) Example digit with different levels of gaussian noise added (varying from a standard deviation of 16 to 192).

In this article, we hypothesize that feedback models outperform feedforward models on noisy images and correlate more closely to human behaviour. A computational model is proposed by integrating feedback connections into a standard feedforward neural network. We test the proposed architecture on MNIST dataset with different levels of gaussian noise and compare it with a feedforward network. A behavioural experiment with human participants is used to test if feedback or feedforward models match human object recognition.

Materials and Methods

Data

The MNIST dataset is used to test the models for recognition performance. To test the effect of noise on recognition performance, different levels of gaussian noise ($\sigma = 16to192$) are introduced both in the training and testing sets.

Computational models

We test four different models as shown in Figure 1; i) the standard feedforward neural network (F), ii) F+ feedback connection to hidden layer (FH), iii) F + feedback connection to input layer (FI) and iv) F + feedback connection to both hidden and input layer (FIH). The layers in the network are fully connected and dimensions of the layers are as shown in Figure 1. The models are trained and tested on each level of gaussian noise added to the MNIST digits.

Human behaviour

The human behavioural experiment is conducted using 5 participants on 40 digits randomly selected from each of the noisy levels. The task is to recognize if a target digit is present or not in the presented image. If human participants indeed use feedback, it would be reflected in higher response time with increase in noise levels.

Results

Table 1: Recognition performance at different noise levels

	σ 16	σ 32	σ 48	σ 64	σ 80	σ 96
F	**0.940**	**0.937**	**0.929**	0.911	0.89	0.859
FH	0.938	0.934	0.925	0.906	0.886	0.857
FI	0.928	0.927	0.924	**0.917**	**0.908**	**0.889**
FHI	0.922	0.92	0.917	0.909	0.898	0.875
	σ 112	σ 128	σ 144	σ 160	σ 176	σ 192
F	0.824	0.782	0.742	0.694	0.651	0.614
FH	0.825	0.786	0.746	0.699	0.656	0.618
FI	**0.867**	**0.833**	**0.798**	**0.753**	**0.711**	**0.675**
FHI	0.849	0.814	0.776	0.733	0.691	0.656

Computational models

We evaluate the different models on the MNIST dataset at different noise levels as shown in Table 1. For extremely low noise levels the purely feedforward network (F) performs the best while, at higher noise levels the feedback network (FI) outperforms F. It is also the best performing network compared to FH and FHI.

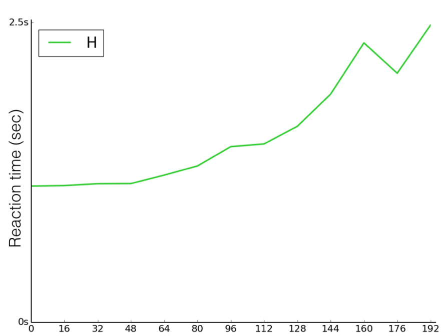

Figure 2: Reaction time of human object recognition with increase in gaussian noise added to the MNIST images.

Human behaviour

To test if feedback is used by humans in challenging conditions, we show the reaction time of participants in Figure 2. We observe that reaction time for human recognition increases with increase in noise level of the image. To determine if feedback networks match human behaviour we compared computational models (trained on clean images, tested on all noise levels) to human behaviour in Figure 3. The feedback model has a higher correlation (r = 0.95, p ¡ 0.005) to human object recognition performance as compared to feedforward model (r = 0.81, p ¡ 0.005).

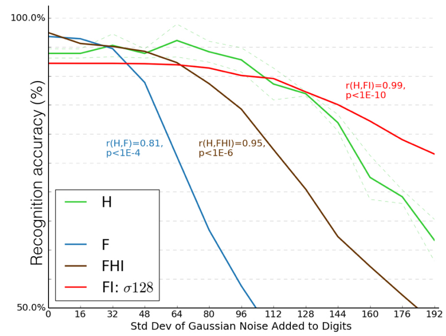

Figure 3: Feedforward (F:trained on clean images) and feedback model (FHI: trained on clean images, FI:trained on noise level 128) compared against human object recognition behaviour (H).

Conclusion

Our results show that feedback neural networks outperform feedforward network object recognition in presence of noise. Furthermore, feedback models best describe human object recognition performance. This suggests that feedback neural networks are a biologically plausible model to understand human object recognition.

Acknowledgments

KR is supported by COMMIT.

References

Güçlü, U., & van Gerven, M. A. (2015). Deep neural networks reveal a gradient in the complexity of neural representations across the ventral stream. *The Journal of Neuroscience*, *35*(27), 10005–10014.

He, K., Zhang, X., Ren, S., & Sun, J. (2015). Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. *arXiv preprint arXiv:1502.01852*.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Nips* (p. 1106-1114).

Lamme, V. A., & Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in neurosciences*, *23*(11), 571–579.