Oscillations in an ANN Convert Competing Inputs into a Temporal Code

Katharina Duecker (katharina.duecker@gmail.com)
Centre for Human Brain Health, School of Psychology, University of Birmingham,
52 Pritchatts Road, Birmingham, B15 2TT, UK

Marco Idiart (marco.idiart@gmail.com)
Institute of Physics, Federal University of Rio Grande do Sul
Avenida Bento Gonçalves, Porto Alegre, 9500, Brazil

Marcel van Gerven (marcel.vangerven@donders.ru.nl)
Donders Institute for Brain, Cognition and Behaviour
Thomas van Aquinostraat 4, Nijmegen, 6525, The Netherlands

Ole Jensen (ole.jensen@bham.ac.uk)
Centre for Human Brain Health, School of Psychology, University of Birmingham
52 Pritchatts Road, Birmingham, B15 2TT, UK

Abstract:

Deep convolutional neural networks (DNNs) show strong similarities to the primate ventral stream in terms of the hierarchical organization of representations. However, DNNs typically disregard the temporal dynamics experimentally observed in these areas. For instance, alpha oscillations dominate the dynamics of the human visual cortex, yet oscillations are rarely considered in neural networks. We propose a neural network that embraces oscillatory dynamics with the computational purpose of converting spatial information into a temporal code. The network was trained to classify three letters in four quadrants. Post-training, we added semi-realistic temporal dynamics to the hidden layer, introducing relaxation and pulsed inhibition mimicking neuronal alpha oscillations. The relaxation ensured non-sustained activation. Without the rhythmic inhibition, the trained network correctly classified individual letters but showed high uncertainty when presented with two stimuli, elucidating a bottleneck problem. When introducing pulsed oscillatory inhibition, the output nodes corresponding to the two stimuli activated sequentially, generating a temporal code. Our model provides a novel approach for implementing multiplexing in ANNs embracing physiologically plausible temporal dynamics. Future work will expand to larger networks and to constraining the dynamics based on neural recordings.

Keywords: Neural networks; Neuronal oscillations; Temporal dynamics; Inhibition; Multiplexing

Deep convolutional neural networks have had enormous success in solving a wide range of computer vision problems (Krizhevsky et al., 2012; LeCun and Bengio, 1995; Voulodimos et al., 2018). Originally inspired by the receptive fields of neurons in visual cortex, the hierarchically organized representations emerging in these networks have been repeatedly shown to correspond strongly to those identified from human MEG and fMRI recordings of the visual ventral stream (Cichy et al., 2017, 2016; Güçlü, and van Gerven, 2015) and intracranial recordings from the non-human primate brain (Kriegeskorte, 2015; Marques et al., 2021; Schrimpf et al., 2020; Yamins et al., 2014; Yamins and DiCarlo, 2016). While alpha oscillations (8-12 Hz) are strongly modulated in visual tasks, the temporal dynamics observed in electrophysiological recordings are typically not used to inform DNNs (Kuzovkin et al., 2018; Reddy et al., 2021).

Here, we propose a neural network that embraces oscillatory dynamics. Using relaxation dynamics and pulsed inhibition, the networks segments simultaneously presented inputs in time, organizing competing inputs in a temporal code.

Methods

Network architecture

We present a neural network with one hidden layer (68 units), fully connected to the three units in the output layer (Fig. 1). The input (56x56 pixels) could be one of three letters (“A”, “E”, “T”) presented in one of the 28x28 quadrants. To implement competition between the quadrants, the same 28x28 weight matrix was applied to each quadrant as illustrated by green square in Fig. 1. As such, each hidden unit received 4x28x28 inputs. The hidden activations were calculated as:

$$H_j = \sigma(Z_j) = \frac{1}{1 + e^{-2(Z_j - 2.5)}} \quad (1)$$

With H_j being the activation at hidden unit j, and Z_j being the sum over the quadrants multiplied with the 28x28 weight matrix.

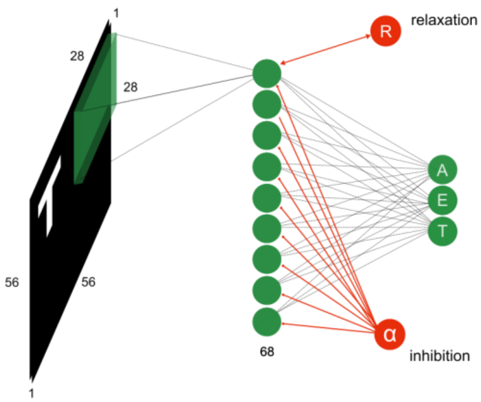

Fig. 1 The architecture of the one-layer fully connected network. Competition between quadrants was implemented by applying the same 28x28 weight matrix to each quadrant and summing the weighted inputs at each hidden unit. The dynamics were added post-training to the hidden layer units, using a relaxation term R and pulsed inhibition α.

Note that the sigmoid activation was steepened and shifted to ensure that the activations were somewhat sparse and outside the linear part of the sigmoid, i.e., close to 0 or 1. The adjustable parameters, i.e. the size of the hidden layer, slope of the sigmoid $a = -2$ and shift in x direction $b = -2.5$, where chosen manually, to achieve somewhat binary activations. This was relevant for the imposed dynamics described below. Output activations were calculated using the (arg)-softmax function. The weights were learned using stochastic gradient descent.

Oscillatory dynamics

After the training, we added semi-realistic temporal dynamics to each node in the hidden layer, based on differential equations, solved using the Euler integration method. Dynamics in each hidden unit j were defined as:

$$\tau_H \frac{dH_j}{dt} = -H_j + \sigma\left(\frac{Z_j - R_j - \alpha(t) + H_j}{S}\right) \tag{2}$$

Here, $\tau_H = 0.01$ determines the timescale by which H_j approaches the activation $\sigma(Z_j)$.The relaxation term R_j was added to ensure a non-sustained activation, and traced the activation in H_j with a time delay of $\tau_R = 0.05$:

$$\tau_R \frac{dR_j}{dt} = -R_j + c \cdot H_j \tag{3}$$

The pulsed inhibition α was implemented as $\alpha(t) = a\sin(2\pi t \cdot 10) + a$ with amplitude a being adjustable to modulate the strength of the inhibition. S was set to a small number to drive the sigmoid activations towards 1 ($S = 0.01$). The increase of R_j at each time step was defined by $c = 4 \cdot \max(Z)$, i.e. a multiple of the maximum pre-activation in the hidden layer.

Results

Pulsed inhibition generates a temporal code

Error! Reference source not found.a shows the activation of the output node to a single letter T, presented for 300ms at luminance 1. The activation in the output unit corresponding to the T builds up according to $\tau_h = 0.01\ s$. The activations in the output units corresponding to “E” and “A” approach 0 (**Error! Reference source not found.**a, top right).

When presenting two inputs simultaneously, we modulated their gain such that all values corresponding to the attended letter (here “T”) were set to 1.1 and pixel values in the unattended letter (here “E”) were set to 0.9 (**Error! Reference source not found.**b, top left). As expected, the softmax function divided the activation between the output units of the two letters but favored the “T” (**Error! Reference source not found.**b, top right).

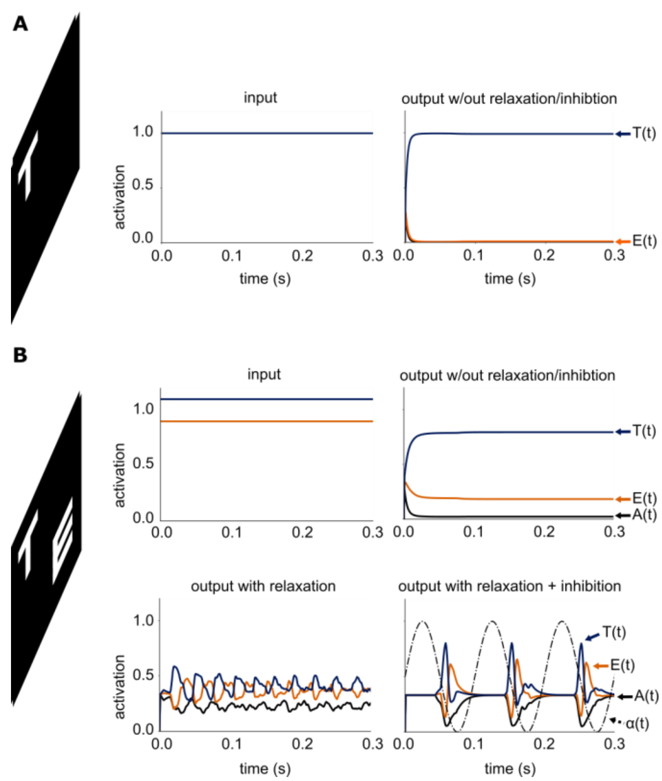

Fig. 2 Dynamics in the output units. **A** Activations in the output node corresponding to T increase following τ_h. **B** Presentation of two letters indicates the bottleneck problem. The weaker input (here “E”) loses the competition against the stronger “T” (top right). When imposing relaxation dynamics, the output nodes corresponding to the “E” and “T” oscillate in antiphase

(bottom left). When additionally imposing pulsed inhibition, the network generates a temporal code, whereby the “T” is read out before the “E” (bottom right).

When adding relaxation dynamics with $c > 0$ in equation (3), and the amplitude of the alpha oscillations set to $a = 0$, the activations in the output nodes oscillate in antiphase (Fig. 2B, bottom left). However, "E" eventually loses the competition against the stronger input "T", and never reaches activations above 0.5.

Adding the 10Hz pulsed inhibition ($a = 1$) to the network results in a temporal code, with "T" being read out at an earlier phase of the alpha oscillation than "E". As the activation of the "T" reduces due to the relaxation dynamics, the unit corresponding to the letter "E" starts to activate and then decays. This repeats in every alpha cycle. Unlike the dynamics without inhibition, this code is stable over time, and the weaker input "E" wins the competition against "T" in every cycle (**Error! Reference source not found.**b, bottom right).

Discussion

We here present a dynamical neural network that resolves the competition between simultaneously presented inputs using relaxation dynamics and pulsed inhibition. Without having been explicitly trained to disentangle simultaneously presented inputs, the network dynamics separate the hidden activations of the individual inputs in time. The strongest input overcomes the inhibition first and is read out at an earlier phase of the rhythmic inhibition than the weaker input. These results are in line with a pipelining mechanism suggested to be implemented by alpha oscillations in visual cortex (Jensen et al., 2021, 2014). It is also akin to hippocampal phase coding implemented by phase-coupling between theta and gamma oscillations (Lisman and Idiart, 1995). Our work provides a proof-of-principle for the use of dynamical deep neural networks as models of the visual system. Future work will focus on incorporating more layers and using a larger stimulus set to eventually constrain the network by electrophysiological recordings from the human and non-human primate visual cortex.

Acknowledgments

This research was supported by a Wellcome Trust Investigator Award in Science 207550 awarded to O.J. M.I. acknowledges a Brazilian Government CAPES/PRINT grant (88887.583995).

References

Cichy, R.M., Khosla, A., Pantazis, D., Oliva, A., 2017. Dynamics of scene representations in the human brain revealed by magnetoencephalography and deep neural networks. NeuroImage 153, 346–358. https://doi.org/10.1016/j.neuroimage.2016.03.063

Cichy, R.M., Khosla, A., Pantazis, D., Torralba, A., Oliva, A., 2016. Comparison of deep neural networks to spatio-temporal cortical dynamics of human visual object recognition reveals hierarchical correspondence. Sci. Rep. 6, 27755. https://doi.org/10.1038/srep27755

Güçlü, U., van Gerven, M.A.J., 2015. Deep Neural Networks Reveal a Gradient in the Complexity of Neural Representations across the Ventral Stream. J. Neurosci. 35, 10005–10014. https://doi.org/10.1523/JNEUROSCI.5023-14.2015

Jensen, O., Gips, B., Bergmann, T.O., Bonnefond, M., 2014. Temporal coding organized by coupled alpha and gamma oscillations prioritize visual processing. Trends Neurosci. 37, 357–369. https://doi.org/10.1016/j.tins.2014.04.001

Jensen, O., Pan, Y., Frisson, S., Wang, L., 2021. An oscillatory pipelining mechanism supporting previewing during visual exploration and reading. Trends Cogn. Sci. 25, 1033–1044. https://doi.org/10.1016/j.tics.2021.08.008

Kriegeskorte, N., 2015. Deep Neural Networks: A New Framework for Modeling Biological Vision and Brain Information Processing. Annu. Rev. Vis. Sci. 1, 417–446. https://doi.org/10.1146/annurev-vision-082114-035447

Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. ImageNet Classification with Deep Convolutional Neural Networks, in: Advances in Neural Information Processing Systems. Curran Associates, Inc.

Kuzovkin, I., Vicente, R., Petton, M., Lachaux, J.-P., Baciu, M., Kahane, P., Rheims, S., Vidal, J.R., Aru, J., 2018. Activations of deep convolutional neural networks are aligned with gamma band activity of human visual cortex. Commun. Biol. 1, 1–12. https://doi.org/10.1038/s42003-018-0110-y

LeCun, Y., Bengio, Y., 1995. Convolutional networks for images, speech, and time-series, in: Arbib, M.A. (Ed.), The Handbook of Brain Theory and Neural Networks. MIT Press.

Lisman, J.E., Idiart, M.A., 1995. Storage of 7 +/- 2 short-term memories in oscillatory subcycles. Science 267, 1512–1515. https://doi.org/10.1126/science.7878473

Marques, T., Schrimpf, M., DiCarlo, J.J., 2021. Multi-scale hierarchical neural network models that bridge from single neurons in the primate primary visual cortex to object recognition behavior (preprint). Neuroscience. https://doi.org/10.1101/2021.03.01.433495

Reddy, L., Cichy, R.M., VanRullen, R., 2021. Representational Content of Oscillatory Brain Activity during Object Recognition: Contrasting Cortical and Deep Neural Network Hierarchies. eNeuro 8. https://doi.org/10.1523/ENEURO.0362-20.2021
Schrimpf, M., Kubilius, J., Hong, H., Majaj, N.J., Rajalingham, R., Issa, E.B., Kar, K., Bashivan, P., Prescott-Roy, J., Geiger, F., Schmidt, K., Yamins, D.L.K., DiCarlo, J.J., 2020. Brain-Score: Which Artificial Neural Network for Object Recognition is most Brain-Like? bioRxiv 407007. https://doi.org/10.1101/407007
Voulodimos, A., Doulamis, N., Doulamis, A., Protopapadakis, E., 2018. Deep Learning for Computer Vision: A Brief Review. Comput. Intell. Neurosci. 2018, e7068349. https://doi.org/10.1155/2018/7068349
Yamins, D.L.K., DiCarlo, J.J., 2016. Using goal-driven deep learning models to understand sensory cortex. Nat. Neurosci. 19, 356–365. https://doi.org/10.1038/nn.4244
Yamins, D.L.K., Hong, H., Cadieu, C.F., Solomon, E.A., Seibert, D., DiCarlo, J.J., 2014. Performance-optimized hierarchical models predict neural responses in higher visual cortex. Proc. Natl. Acad. Sci. 111, 8619–8624. https://doi.org/10.1073/pnas.1403112111