Exploring More Biologically-Inspired Recurrent Neural Networks

Bhavana Jonnalagadda

Abstract

We examine features in the brain that have potential applications in artificial neural networks, and then review areas within the machine learning field that would benefit from strategies taken from observing the brain. Using the discussed intersection of properties between these neural networks, we will later attempt to construct an ANN to represent a biological circuit in the mouse visual system.

1 Introduction

2 The Biological Neural Network

The human brain is the most complex piece of machinery that nature has produced, and yet remains one of the most enigmatic. The processes behind how we reason and decide and think are unknown and highly debated, and all current models of the brain have yet to really describe the entirety of consciousness. A vast amount of research has been done in cognitive neuroscience, and from this we can at least see aspects of the brain that can be described as high-level processing [45]. Plasticity, the ability of our neurons to change their behavior, can lead to nonlinear computation even on the single neuron level [25]; it is how our brain adapts, learns, and manages to not only process vast amounts of information but also create increasingly-abstracted representations of the world in our head [40]. Changes in our neural circuits happen within the individual neuron but are reflective of a larger complex, dynamical system that manages to stay stable [43]. And the ability to update our neural network is crucial to learning; creating and integrating new neurons is how we learn as a child and form long term memories [6]. Learning and reacting to our environment are aspects of our brain that we aim to imitate with artificial neural networks (ANN); its compact and efficient information processing, combined with its ability for infinite abstraction would constitute a veritable revolution in machine learning.

2.1 Inspirations for Artificial Neural Networks

Processes in the brain have already been taken and applied to ANNs, with huge success. A classic example is the visual system. Retinal cells process the world not as an image of pixels, but as a combination of light-dark edges, movement, brightness, and distance. Some cells manage to compress the information they receive [14] by firing only when a specific orientation of light passes over them, and then pass it on to neurons (citations needed). This ability to extract features from an image is used in convolutional neural networks (CNN), and has been used in machine learning to identify and categorize images to staggering accuracy [22, 42]. Networks now can even generate natural sentences that accurately describe the contents of a picture [44].

Our brain's auditory system has also inspired better speech-to-text processing; circuitry for auditory processing and sound pattern recognition uses timing and relational signals to decide [36]. From this, deep CNNs have been created that try to process acoustic input more naturally, and have even managed to independently extract syllables from audio with no prompting [24].

Many other aspects of neural systems have been inspirations for important computational uses, from designing efficient distribution networks [30] to robot locomotion (citation needed) to user recommendations (citation needed).

2.2 Time-Dependent Decisions and Oscillations

A crucial aspect to the brain's processing is timing. Hebbian plasticity is based off of coincident firing in neuronal circuits [43], and individual pairwise correlations can be scaled to describe the behavior of the larger local neural network [35]. Learning in the brain is based on delayed/staggered signals that arrive to neurons in a specific order [15]. Even our brain's representation of time is relational: time is encoded within the brain as a function of its previous state, from time-dependent changes in synaptic and cellular propertie of local neural networks [20]. Timing is needed to precisely regulate everything from body movement, heart rate [17], and understanding the environment. Auditory processing for sound pattern recognition is based on a coincidence detector mechanism [36]; in fact, the brain's oscillations, and how tightly they match a music piece, predicts a person's musical experience and even their performance ability [5]. Thus the order (and delay) in which neurons grow, evolve, change [40], and fire signals, is everything to a neural network's performance (citation needed).

2.3 Fractal Properties

Identifying patterns in neural firing is important for understanding the chaotic system, and for increasing the complexity and functionality of our ANNs. Fractal can refer to power law scaling of output, or to the Hausdorff dimension, or to repeating patterns at all time and space scaling, but the base idea is that a simple unit or law can be applied repeatedly at different sizes to produce very intricate behavior/graphs that otherwise seem to have no simple function/explanation. Fractal dimensional patterns arise everywhere in nature, especially in complex dynamical systems [33] like weather, flora, and even surprising areas like coastlines, clouds, and lightning. It shows that a linear self-similarity is key for a simple unit to produce complicated behaviors and topologies. In the brain, the very wavelets of its electrical signals fire in a fractal pattern [29]. The brain's local networks have oscillating structures that scale in a fractal manner, and their signals can be modeled with recurrent neural networks with the fractal power law applied to take apart the complicated output [2, 32]. These neural fractals are at all levels in the nervous system: they are present from individual ion channels to peripheral nerves to macroscopic brain organization, to symbolic processing to motor behavior [46]. They are used to regulate circadian rhythm, and to organize regular body activity like heart rate [17]. They are also used in brain-wide information compression and scaling [2], giving potential applications for ANNs to more efficiently process data.

For a direct application to recurrent neural networks, they can potentially be used as a form of plasticity: currently, most models use back-propagation and gradient descent to learn their networks, but this method potentially erases or blows up errant input and does not account for the complex relations between neurons/nodes. Introducing a method to increase the weight of connections that fire "together" in a fractal manner (i.e. time delays between signal inputs to a neuron occur in power-law relations) could potentially increase the functionality and stability of an ANN, especially since this plasticity is infinitely scalable [40].

2.4 Nested Levels of Computation

This scaling ability in the brain is not just limited to fractal patterns. Computation in the nervous systems happens at all levels, from the molecular/chemical all the way to the abstraction of thought we possess [12]. Even epigenetic modifications (changes to the DNA in the neuron) are important for the storage and creation of memory, an integral part in decision making [4]. This ability to compute, from the smallest biological unit all the way up to the most superficial observation of the brain's activity, is crucial for its ability to model dynamical systems. Without the ability to adapt the decision-making at all levels, learning would not be nearly as flexible or intuitive.

3 Artificial Neural Networks

Machine learning has caused a veritable revolution in data analysis, to say the least. Artificial neural networks (ANN) are the peak of this work, and are our closest representation of an artificial brain that we have. They were loosely based on spiking neurons in the brain, but now are entrenched in graph theory [13]. ANNs are used for everything from designing more efficient distribution routes [30], to identifying images and videos [22, 42, 31, 3], to generating user recommendations, to robot control [41], to automatically playing video games [39, 28], to understanding spoken language [24]. The most popular form today are Deep Neural Networks (DNN), which are simply ANNs with more layers and many more neurons in each layer, along with more advanced processing at each layer [34]. Yet with all their popularity in cutting edge computational research today, ANNs still remain a black box when it comes to understanding how and why they work so well. This opaqueness is just one of the many ways ANNs naturally imitate the brain.

3.1 Recurrent Neural Networks

Recurrent neural networks (RNN) are a form of ANNs that can have the weights of the nodes connect in either direction, instead of each layer just feeding forward. This means that time then becomes a factor in this network; at each time step, there could be signals propagating forwards or backwards. This property has made RNNs very adept at handling sequential data such as text or audio [21], but has also made RNNs hard to train and keep stable [16, 47]. RNNs are some of the harder to interpret networks, since their output - from the network as well as from the nodes - could mean different things depending on what they outputted before and afterwards. But of course, this is how the brain works as well: no understood restriction on which neurons connect, and with multiple factors affecting the timing of these signals.

Recurrent neural networks are the first step in making ANNs more organic and flexible [2], to data over time but also to making more "educated" decisions. With signals depending on future input at delayed timesteps, decisions are made less erratically and more naturally; long-term potentiation properties give RNNs the ability to potentially handle complex spatio-temporal stimuli like the brain currently does [23]. This would be crucial in future models that parse text/speech full of noise like grammatical errors, or models that attempt to make decisions across large timescales - for example, taking second-by-second data on stock fluctuations in an attempt to predict their value at the end of a season.

3.2 Reinforcement Learning

Reinforcement learning (RL) is another popular form of an ANN where the output must change over time, but instead through trial-and-error interactions with the environment [19]. The model learns using reward and punishment values to update itself after each output, and then receiving a reaction from the environment. Currently, even reinforcement learning has been improved by using deep convolutional networks, using brain-inspired "experience replay" and periodic updating to automatically beat games with large success [28]. This method has become very popular, with ever-increasing accuracy and technique [7, 26]. Reinforcement learning can be used for areas other than just gaming, but also in robotics such as a mechanical task like pole-balancing [41], or even computing and performing sorting algorithms [48].

Reinforcement learning is a growing and important field, as its applications are potentially huge if the networks were easier to train; we would be able to deploy a network live in a situation and it would learn to perform the required tasks with only minimal information an feedback. Reinforcement learning also happens to be a big field of study in psychology - Pavlov's dog and bell experiment is a prime example of how important the research can be, and attempting to get as dramatic of results with humans is a prime topic. RL is an active intersection between machine learning and neuroscience, and developments in either field have the potential to greatly influence the other.

3.3 Specific Models and Their Properties

[11]: NTM; Describes Neural Turing Machines, a recurrent neural network enriched with extensible external memory source

[48]: Reinforcement Learning NTM: turing complete computation with efficient learning

[31]: Using CNN -¿ LSTM for video recognition/classification

[44]: Image classification with multiple words, generating a real sentence; CNN -¿ LSTM

[8]: DeViSE; a model that can identify visual objects using both labeled image data as well as semantic information gleaned from unannotated text, i.e. state-of-the-art image classification combined with word labels

[38]: CPPN; Describes compositional neural networks, a class of activation functions that produce patterns and scale independently of locality/relatively close nodes

[18]: Evolutionary algorithm that involves merging nodes

[41]: NEAT evolutionary alg

[39]: HyperNEAT; evolve connections with respect to space output; evolves composition of functions

[9]: Essentially describes evolving convolutional layers; An evolutionary algorithm for generating ANNs, using HyperNEAT, that have geometric regularities

[23]: SORN; combines three distinct forms of local plasticity to learn spatio-temporal patterns in its input: spike-timing-dependent plasticity (STDP) (weights b/w neurons), intrinsic plasticity (IP) (activation function), and synaptic scaling of the excitatory–excitatory connections (i.e. averaging/limiting incoming total signals)

3.4 Existing Limitations and Solutions

[37]: Compares regulation methods (to prevent overfitting) for convolutional NNs performing image classification; emphasizes dropout success

[10]: Adversarial examples; NN vulnerability, can be trained on to reduce errors

[1]: shallow feed-forward nets can learn the complex functions previously learned by deep nets, with similar accuracies

[27]: Get near performance with shallow, non-convolutional NN

[40]: Describes key features in artificial embryogeny that correlate to evolutionary mechanisms: Cell fate, targeting, heterochrony, canalization, and complexification

4 Future Work

4.1 Existing Models and Algorithms to Simulate the Biological Network

[1]: shallow feed-forward nets can learn the complex functions previously learned by deep nets, with similar accuracies

[27]: Get near performance with shallow, non-convolutional NN

[30]: large network with pruning afterwards

4.2 Designing an RNN for a Biological Network

4.3 Combining and Improving Existing Models

4.4 Proposed New Model

Drawing from ideas of the discussed models, a model based on RNNs is proposed: it has the structure of an LSTM for the temporal order, with four main additional properties:

- Neural layers may serve as activation functions.
- The timing of signal can also be changed in addition to the weight of a signal.
- Updating of weight/timing uses various rules of ”plasticity” in addition to backpropogation, similar to [23].
- The topology of the network is evolved using HyperNEAT [39].

These properties introduce some large deviations from standard RNNs:

- The output of a single neuron can be different for each weight, since a neuron can represent a layer of neurons.
- Learning the network needs to account for updating timing as well as strength; (for simplification, weight strength and timing can be equivalent).
- There needs to be a phase of learning the structure of a network, separate from learning the weights.
- Multiple types of cost functions are needed to evaluate the output of a network.

- Stability of the network is prioritized over correctness/ minimizing the error function.
- Patterns/sections of the RNN must be able to be added/removed in a logical manner, that changes the functionality in a semi-understandable manner.

The model with these additions will then be evolved, trained, regularized, and updated with the methods from discussed papers (Gauci 2010, Stanley 2009, Angeline 1994, Williams 1989).

Applications/ Data Sets: Data set(s) will be chosen that already have working NN models, and are in areas that could potentially benefit from the added memory and pattern identification of the proposed project, such as language parsing and other sequential events. Additionally, a speech recognition data set provided by Yisong Yue's group will be used to test the project model.

Results: The model will be compared to existing structures on its efficiency in training time as well as runtime. Features within the model that may have contributed to its function will be discussed, along with features that differ from existing structures .Additionally, a system for identifying and evolving patterned connections will be proposed.

References

[1] Lei Jimmy Ba and Rich Caruana. "Do Deep Nets Really Need to be Deep?" In: (Dec. 2013). arXiv: 1312.6184. URL: http://arxiv.org/abs/1312.6184.

[2] Erhard Bieberich. "Recurrent fractal neural networks: A strategy for the exchange of local and global information processing in the brain". In: *BioSystems* 66.3 (2002), pp. 145–164. ISSN: 03032647. DOI: 10.1016/S0303-2647(02)00040-0.

[3] Adam Coates et al. "Text Detection and Character Recognition in Scene Images with Unsupervised Feature Learning". In: *2011 International Conference on Document Analysis and Recognition.* IEEE, Sept. 2011, pp. 440–445. ISBN: 978-1-4577-1350-7. DOI: 10.1109/ICDAR.2011.95. arXiv: fa. URL: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6065350.

[4] Jeremy J Day and J David Sweatt. "Epigenetic Modifications in Neurons are Essential for Formation and Storage of Behavioral Memory". In: *Neuropsychopharmacology* 36.1 (Jan. 2011), pp. 357–358. ISSN: 0893-133X. DOI: 10.1038/npp.2010.125. URL: http://www.nature.com/doifinder/10.1038/npp.2010.125.

[5] Keith B. Doelling and David Poeppel. "Cortical entrainment to music and its modulation by expertise". In: *PNAS* (Oct. 2015), pp. 1508431112–. DOI: 10.1073/pnas.1508431112. URL: http://www.pnas.org/content/early/2015/10/21/1508431112.abstract.

[6] Dan Ehninger and Gerd Kempermann. "Neurogenesis in the adult hippocampus". en. In: *Cell and tissue research* 331.1 (Jan. 2008), pp. 243–50. DOI: 10.1007/s00441-007-0478-3. URL: http://link.springer.com/article/10.1007/s00441-007-0478-3/fulltext.html.

[7] Vincent François-Lavet, Raphael Fonteneau, and Damien Ernst. "How to Discount Deep Reinforcement Learning: Towards New Dynamic Strategies". In: (Dec. 2015). arXiv: 1512.02011. URL: http://arxiv.org/abs/1512.02011.

[8] Andrea Frome, GS Corrado, and Jonathon Shlens. "Devise: A deep visual-semantic embedding model". In: *Advances in Neural ...* (2013), pp. 1–11. URL: http://papers.nips.cc/paper/5204-devise-a-deep-visual-semantic-embedding-model.

[9] Jason Gauci and Kenneth O. Stanley. "Autonomous evolution of topographic regularities in artificial neural networks." In: *Neural computation* 22.7 (2010), pp. 1860–1898. ISSN: 0899-7667. DOI: 10.1162/neco.2010.06-09-1042.

[10] Ian J. Goodfellow, Jonathon Shlens, and Christian Szegedy. "Explaining and Harnessing Adversarial Examples". In: (Dec. 2014). arXiv: 1412.6572. URL: http://arxiv.org/abs/1412.6572.

[11] Alex Graves, Greg Wayne, and Ivo Danihelka. "Neural Turing Machines". In: (Oct. 2014). arXiv: 1410.5401. URL: http://arxiv.org/abs/1410.5401.

[12] Sten Grillner, Alexander Kozlov, and Jeanette Hellgren Kotaleski. "Integrative neuroscience: Linking levels of analyses". In: *Current Opinion in Neurobiology* 15.5 (2005), pp. 614–621. ISSN: 09594388. DOI: 10.1016/j.conb.2005.08.017.

[13] a Grüning and Sm Bohte. "Spiking Neural Networks: Principles and Challenges". In: *Elen.Ucl.Ac.Be* April (2014), pp. 23–25. URL: https://www.elen.ucl.ac.be/Proceedings/esann/esannpdf/es2014-13.pdf.

[14] Sarah Hammer et al. "Multiple Retinal Axons Converge onto Relay Cells in the Adult Mouse Thalamus." English. In: *Cell reports* 12.10 (Aug. 2015), pp. 1575–1583. ISSN: 2211-1247. DOI: 10.1016/j.celrep.2015.08.003. URL: http://www.cell.com/article/S2211124715008591/fulltext.

[15] Kaiwen He et al. "Distinct Eligibility Traces for LTP and LTD in Cortical Synapses". English. In: *Neuron* 88.3 (Nov. 2015), pp. 528–538. ISSN: 08966273. DOI: 10.1016/j.neuron.2015.09.037. URL: http://www.cell.com/article/S0896627315008260/fulltext%20http://linkinghub.elsevier.com/retrieve/pii/S0896627315008260.

[16] Sepp Hochreiter and Jürgen Schmidhuber. *Long Short-Term Memory.* 1997. DOI: 10.1162/neco.1997.9.8.1735.

[17] Kun Hu et al. "Fractal Patterns of Neural Activity Exist within the Suprachiasmatic Nucleus and Require Extrinsic Network Interactions". In: *PLoS ONE* 7.11 (2012).

[18] M. Islam et al. "A New Adaptive Merging and Growing Algorithm for Designing Artificial Neural Networks". In: *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 39.3 (June 2009), pp. 705–722. ISSN: 1083-4419. DOI: 10.1109/TSMCB.2008.2008724. URL: http://www.ncbi.nlm.nih.gov/pubmed/19203888%20http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4776509.

[19] Leslie Pack Kaelbling, Michael L Littman, and Andrew W Moore. "Reinforcement learning: A survey". In: *Journal of artificial intelligence research* 4 (1996), pp. 237–285. DOI: 10.1613/jair.301.

[20] Uma R. Karmarkar and Dean V. Buonomano. "Timing in the Absence of Clocks: Encoding Time in Neural Network States". In: *Neuron* 53.3 (2007), pp. 427–438. ISSN: 08966273. DOI: 10.1016/j.neuron.2007.01.006.

[21] Andrej Karpathy, Justin Johnson, and Fei-Fei Li. "Visualizing and Understanding Recurrent Networks". In: (June 2015). arXiv: 1506.02078. URL: http://arxiv.org/abs/1506.02078.

[22] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. "ImageNet Classification with Deep Convolutional Neural Networks". In: *Advances In Neural Information Processing Systems* (2012), pp. 1–9. ISSN: 10495258. arXiv: 1102.0183.

[23] Andreea Lazar, Gordon Pipa, and Jochen Triesch. "SORN: a self-organizing recurrent neural network." In: *Frontiers in computational neuroscience* 3 (2009), p. 23.

[24] Chia-ying Lee, Timothy J O 'donnell, and James Glass. "Unsupervised Lexicon Discovery from Acoustic Input". en. In: *Transactions of the Association for Computational Linguistics* 3 (July 2015), pp. 389–403. ISSN: 2307-387X. URL: https://tacl2013.cs.columbia.edu/ojs/index.php/tacl/article/view/520.

[25] Robert Legenstein and Wolfgang Maass. "Branch-specific plasticity enables self-organization of nonlinear computation in single neurons." In: *The Journal of neuroscience : the official journal of the Society for Neuroscience* 31.30 (2011), pp. 10787–10802. ISSN: 0270-6474. DOI: 10.1523/JNEUROSCI.5684-10.2011.

[26] Yitao Liang et al. "State of the Art Control of Atari Games Using Shallow Reinforcement Learning". In: (Dec. 2015). arXiv: 1512.01563. URL: http://arxiv.org/abs/1512.01563.

[27] Mark D McDonnell et al. "Fast, Simple and Accurate Handwritten Digit Classification by Training Shallow Neural Network Classifiers with the 'Extreme Learning Machine' Algorithm." In: *PloS one* 10.8 (Jan. 2015), e0134254. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0134254. URL: http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0134254.

[28] Volodymyr Mnih et al. "Human-level control through deep reinforcement learning". In: *Nature* 518 (2015). DOI: 10.1038/nature14236.

[29] Radu Mutihac. "Wavelets , Fourier Transform , and Fractals". In: *Analysis* i (2006), pp. 49–58. URL: http://www.fizica.unibuc.ro/mutihac.

[30] Saket Navlakha, Alison L. Barth, and Ziv Bar-Joseph. "Decreasing-Rate Pruning Optimizes the Construction of Efficient and Robust Distributed Networks". In: *PLOS Computational Biology* 11.7 (2015), e1004347. ISSN: 1553-7358. DOI: 10.1371/journal.pcbi.1004347. URL: http://dx.plos.org/10.1371/journal.pcbi.1004347.

[31] Joe Yue-Hei Ng et al. "Beyond Short Snippets: Deep Networks for Video Classification". In: *Cvpr2015* (Mar. 2015). arXiv: 1503.08909. URL: http://arxiv.org/abs/1503.08909.

[32] Tyler M. Reese et al. "Analyzing Self-Similar and Fractal Properties of the C. elegans Neural Network". In: *PLoS ONE* 7.10 (2012).

[33] Dierk Schleicher. "Hausdorff dimension, its properties, and its surprises". In: *American Mathematical Monthly* (2007). ISSN: 00029890. arXiv: 0505099 [math].

[34] Jürgen Schmidhuber. "Deep learning in neural networks: An overview". In: *Neural Networks* 61 (Oct. 2014), pp. 85–117. ISSN: 08936080. DOI: 10.1016/j.neunet.2014.09.003. URL: http://www.sciencedirect.com/science/article/pii/S0893608014002135.

[35] Elad Schneidman et al. "Weak pairwise correlations imply strongly correlated network states in a neural population". In: *Changes* 29.6 (2012), pp. 997–1003. DOI: 10.1016/j.biotechadv.2011.08.021.Secreted. arXiv: NIHMS150003.

[36] S. Schoneich, Konstantinos Kostarakos, and Berthold Hedwig. "An auditory feature detection circuit for sound pattern recognition". In: *Science Advances* 1.8 (Sept. 2015), e1500325–e1500325. ISSN: 2375-2548. DOI: 10.1126/sciadv.1500325. URL: http://advances.sciencemag.org/cgi/doi/10.1126/sciadv.1500325.

[37] Evgeny A. Smirnov, Denis M. Timoshenko, and Serge N. Andrianov. "Comparison of Regularization Methods for ImageNet Classification with Deep Convolutional Neural Networks". In: *AASRI Procedia* (2014). ISSN: 22126716. DOI: 10.1016/j.aasri.2014.05.013.

[38] Kenneth O. Stanley. "Compositional pattern producing networks: A novel abstraction of development". In: *Genetic Programming and Evolvable Machines* 8.2 (2007), pp. 131–162. ISSN: 13892576. DOI: 10.1007/s10710-007-9028-8.

[39] Kenneth O. Stanley, D. D'Ambrosio, and Jason Gauci. "A hypercube-based encoding for evolving large-scale neural networks". In: *Artificial life* 15.2 (2009), pp. 1–39. ISSN: 1064-5462. DOI: 10.1162/artl.2009.15.2.15202. URL: http://www.mitpressjournals.org/doi/abs/10.1162/artl.2009.15.2.15202.

[40] Kenneth O. Stanley and Risto Miikkulainen. "A Taxonomy for Artificial Embryogeny". In: *International Security* 6.2 (2003), pp. 209–217. URL: http://www.mitpressjournals.org/doi/abs/10.1162/089892903321208141.

[41] Kenneth O. Stanley and Risto Miikkulainen. "Evolving Neural Networks through Augmenting Topologies". In: *The MIT Press* 6.2 (2003), pp. 209–217. URL: http://www.mitpressjournals.org/doi/abs/10.1162/089892903321208141.

[42] Christian Szegedy et al. "Going Deeper with Convolutions". In: (Sept. 2014). arXiv: 1409.4842. URL: http://arxiv.org/abs/1409.4842.

[43] G G Turrigiano and S B Nelson. "Hebb and homeostasis in neuronal plasticity." In: *Current opinion in neurobiology* 10.3 (June 2000), pp. 358–64. ISSN: 0959-4388. URL: `http://www.ncbi.nlm.nih.gov/pubmed/10851171`.

[44] Oriol Vinyals et al. "Show and Tell: A Neural Image Caption Generator". In: (2015). arXiv: `arXiv:1411.4555v2`. URL: `http://arxiv.org/pdf/1411.4555v2.pdf`.

[45] Tim P Vogels, Kanaka Rajan, and L.F. Abbott. "Neural Network Dynamics". In: *Annual Review of Neuroscience* 28.1 (July 2005), pp. 357–376. ISSN: 0147-006X. DOI: `10.1146/annurev.neuro.28.061604.135637`. URL: `http://www.annualreviews.org/doi/abs/10.1146/annurev.neuro.28.061604.135637`.

[46] Gerhard Werner. "Fractals in the nervous system: conceptual implications for theoretical neuroscience." In: *Frontiers in physiology* 1 (Jan. 2010), p. 15. DOI: `10.3389/fphys.2010.00015`. URL: `http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3059969%7B%5C&%7Dtool=pmcentrez%7B%5C&%7Drendertype=abstract`.

[47] Ronald J. Williams and David Zipser. *A Learning Algorithm for Continually Running Fully Recurrent Neural Networks*. 1989. DOI: `10.1162/neco.1989.1.2.270`.

[48] Wojciech Zaremba and Ilya Sutskever. "Reinforcement Learning Neural Turing Machines". In: (2015). arXiv: `arXiv:1505.00521v1`.