The Compass

Volume 1 | Issue 7 Article 2

November 2020

Computational Cognition and Deep Learning

Andy Malinsky

Computational Cognition and Deep Learning

By: Andy Malinsky, *Arcadia University*

Introduction

One of the goals of both neuroscience and artificial intelligence (AI) is to understand the composition of intelligent cognition, otherwise simply referred to as intelligence. As commonly defined for empirical neuroscience research, intelligence is a general mental ability.[1] It consists of being able to learn new things quickly through logic and abstract reasoning, allowing us to understand an environment and solve problems. In order to make informed decisions and actions, the intelligent system must process some input from which to learn and make inferences. This is the main idea and inspiration behind deep learning, a subset of algorithms within AI. Decisions regarding surrounding environments are made based on inputs received from the senses (mainly sight, touch, and sound). Analogically, connections can be made between these inputs and areas of deep learning, such as vision to sight, reinforcement learning to touch, and natural language processing to sound.

How the brain works is still highly regarded as one of the biggest fundamental mysteries of our universe. Current understanding is challenged by the underlying principles behind the brain's processing of sensory input information up to thoughts and then up to actions. "Intelligence" can be viewed as human-level reasoning from these principles. Both cognitive neuroscientists and AI researchers seek to solve the problem of intelligence: "How can a system composed of relatively unintelligent parts (say, neurons or transistors) behave intelligently?"[2] A major component of this problem involves the process of learning.

To understand what we know so far about learning in the brain, it is important to have a general understanding of what neurons are and how they operate. The region of the brain where most of our cognitive thought takes place is the neocortex, which houses around 20 billion neurons.[3] Each neuron is interconnected with roughly 10,000 others. Communication between neurons happens through synapses, or connection points, by means of electrical signals. Incoming signals carry synaptic weights, which determine whether or not the receiving neuron will activate and send an output signal. Our fundamental understanding of how learning works in the brain is based on the local neural activity patterns between sending and receiving neurons. This activity results in changes in the connections between neurons, a process known as synaptic plasticity. Everything that we learn is essentially different patterns of synaptic weights.

Updating synaptic weights is the key inspiration for deep learning, a subset of machine learning algorithms within the field of AI. It mainly involves the use of more biologically inspired methods, as opposed to traditional machine learning approaches. These methods consist of architectures known as artificial neural networks. Artificial neural networks were designed to operate or "learn" similarly to the biological neural networks found in the brain by continuously updating connections between neuron units. Industry breakthroughs ranging from self-driving cars, self-play game mastering, image classification, and language translation demonstrate deep learning's powerful potential. Deep learning tools can also be beneficial to neuroscientists as they employ cognitive modeling to test their theories of how the brain performs certain tasks and computations. Essentially, we can attempt to use artificial neural networks to mimic the neuronal activity we see in biological systems, which also allows us to bypass ethical concerns over tweaking a human brain to identify contributing factors toward accomplishing a task.[4]

An artificial neural network is said to be "deep" if it consists of multiple layers. A typical model has an input layer, many hidden layers, and an output layer. Each layer is made up of "neurons," or nodes, and each node in a given layer is connected to each node in an adjacent layer with a given weight. These weights can

1 Richard J. Haier, *The Neuroscience of Intelligence* (Cambridge University Press, 2016), 4-5.

2 N.L. Cassimatis, "Artificial intelligence and cognitive modeling have the same problem," *in Theoretical Foundations of Artificial General Intelligence* (Paris: Atlantis Press, 2012), 11-24.

3 R.C. O'Reilly et al., *Computational Cognitive Neuroscience,* 1st ed (Wiki Book, 2012, 3), 12-13.

4 Neil Savage, "How AI and neuroscience drive each other forwards," *Nature* 571 (2019), 10.1038/d41586-019-02212-4.

then be adjusted based on a variation of the stochastic gradient descent model in order to minimize error in generated output values. Neuron activations are forward-and-back propagated through the layers of the network, and the weights are updated accordingly. The process of learning in this system is known as "backpropagation." After the network goes through this training process, it should be able to take new input and provide expected output with an ideally high accuracy. This is the fundamental way deep neural networks function, and Figure 1 provides a visual representation of the underlying mathematics involved.

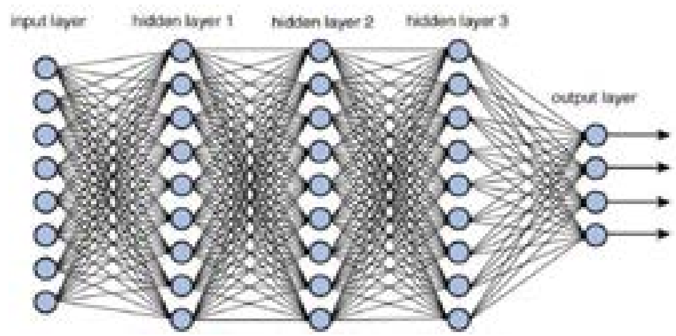


Figure 1: This is a common graphical representation of a deep neural network model.[5]

Neural networks and the backpropagation algorithm were originally designed from the simplest understanding of changes in synaptic weights between neurons. However, many AI researchers have contended that the human brain does not exactly learn the way standard deep learning demonstrates. This concern is about the theme of biological plausibility, the notion that algorithms more closely resembling the processes found in the human brain would improve performance. For example, spiking neural networks[6] have been proven to show promise in this direction. They consist of neuron units that relay information via spikes, or time-based signals, and model simplified versions of biological neuron thresholds and network dynamics. They have also been shown to perform unsupervised learning successfully, using spike-timing-dependent plasticity,[7] a highly biologically plausible approach not commonly utilized in practical applications. A path toward more unsupervised approaches seems to be in order, as humans are capable of learning without a "teacher" when it comes to general problem-solving and critical thinking, while the standard backpropagation model detailed above still requires a "teacher" to feed it labelled data.

Further integration of neuroscience with deep learning may have the potential to solve the biggest challenge in AI: artificial general intelligence, or AGI. The goal of AGI is to create software or hardware systems with general intelligence comparable to or greater than that of a human.[8] AI company DeepMind developed an algorithm involving neural networks called AlphaZero that was able to master the 2,000-year-old game of Go in the space of a few days.[9] Computer programs have the ability to perform amounts of computations that would take humans hundreds or even thousands of years. The end goal is to apply this same level of computational analysis to optimize problems or find solutions in various domains such as healthcare, the environment, finance, and society resource management. AI could uncover patterns in overwhelmingly complex datasets and suggest promising ideas and strategies.[10] To move the field forward, a unifying theory of intelligence would be most ideal, as it would apply to both neuroscience and AI research.

This paper will discuss three active areas of research pertaining to both cognitive neuroscientists and deep learning researchers, including vision, reinforcement learning, and natural language processing. Each of these areas requires some form of human-level reasoning, which allows for specific avenues of investigation in intelligence research. Introduced in the fol-

5 Ravindra Parmar, "Training Deep Neural Networks," *Towards Data Science*, September 11, 2018, https://towardsdatascience.com/training-deep-neural-networks-9fdb1964b964.

6 Guy Dove, "On the need for embodied and dis-embodied cognition," *Frontiers in Psychology* vol. 1 (2011): 242.

7 Peter U. Diehl and Matthew Cook, "Unsupervised learning of digit recognition using spike-timing-dependent plasticity," *Frontiers in Computational Neuroscience* vol. 9 (2015): 99, https://doi.org/10.3389/fncom2015.0099

8 B. Goertzel, "Artificial general intelligence: concept, state of the art, and future prospects," *Journal of Artificial General Intelligence* 5, no. 1 (2014): 1-48, https://doi.org/10.2478/jagi-2014-0001.

9 D. Silver et al., "Mastering chess and shogi by self- play with a general reinforcement learning algorithm." arXiv (2017). https://arxiv.org/abs/1712.01815.

10 "AI and the world's complex challenges," DeepMind Technologies Limited, 2019, https://deepmind.com/applied/deepmind-ethics- society/research/AI-worlds-complex-challenges/.

lowing sections for each research area are the aspects of human cognition involved, along with examples and limitations of the artificial counterparts realized by deep learning.

Vision

Vision plays a crucial part in how the world is perceived and understood. According to Jamie Ward, an acclaimed professor of neuroscience, vision is when "the brain divides a continuous pattern of light into discrete objects and surfaces to translate the two-dimensional retinal image into a three-dimensional interactive model of the environment."[11] The main visual pathway in the brain is known as the primary visual cortex. Neurons in this brain region encode sensory input in terms of edges, or transitions of illumination, to begin developing a mental representation of the perceived natural physical image.[12] This cognitive process involves three stages: perception, recognition, and action.[13] Perception is the sensory experience obtained from consciously translating an input image into a form of acknowledgement signaling that an object has been perceived. Recognition is the brain's way of categorically identifying what it sees. These categories may be developed through prior knowledge or experiences as they help shape individual understanding. Then, action takes place in terms of activity in our motor systems. Depending on the stimulus being perceived, a resultant behavior is expected. For example, if one perceives and recognizes a ball coming towards them, they may take action by putting their hands up to catch it. The cognitive process of vision continuously runs during average daily activity and contributes to how the brain learns new information by constantly providing inputs from its surroundings. The way the human brain is able to interpret and understand what it sees demonstrates a level of intelligent comprehension that is still not fully understood. The eye to brain process, however, still serves as an inspiration for the common way AI "sees."

Common cognitive models used for object recognition tasks are known as convolutional neural networks. Inspired by the biological eye, the neural net scans the input image pixel by pixel, extracting features of the given objects. For example, if the task is to distinguish between a cat and a dog, it can look for features such as "floppy" or "pointy" ears, "dog snout," or "fur."[14] Alternative techniques such as feature visualization allow for a semantic dictionary of how the network understands the input image as a whole. The network would need to go through a training process first in order to identify each feature representation.

Figure 2

Figure 3

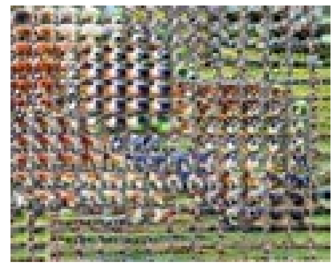


Figure 2 depicts a standard input image of two dominant objects, each with representative features. Figure 3 depicts the resulting feature visualization map, or a semantic dictionary of detected features from the input image.[15] In order for a computer to understand the difference between two given objects, this approach utilizes the idea that the biological brain stores memory of specific sets of features that help to identify similar objects in the future. By way of extracting and identifying features, we are attempting to mimic computationally how the human brain would understand a given visual input.

Convolutional neural networks have a great impact in emerging technologies like self-driving cars and facial recognition, but deep learning for computer vision still has limitations. Training neural networks requires a large amount of labelled data. A network has to learn what "floppy ears" are before it can recognize them in an image. Bias in the lack of realistic examples, such as varying viewpoints, results in worse scene understanding and lower accuracies for object recognition. If the majority of training examples are similar to a particular angle, the network may have a harder time recognizing an object in a new image that is from a completely different viewpoint.[16] For instance, the human brain can understand that a dog

11 Jamie Ward, *The Student's Guide to Cognitive Neuroscience*, 3rd ed. (London: Psychology Press, 2015), 107.
12 R.C. O'Reilly et al., *Computational Cognitive Neuroscience*, 1st ed. (Wiki Book, 2012), 76.
13 Bruce E. Goldstein, *Sensation and Perception*, 8th ed. (Boston, MA: Cengage Learning, 2010), 8-9.
14 C. Olah et al., "The building blocks of interpretability," *Distill 3* (2018) no. 10. https://doi.org/10.23915/distill.00010.
15 Ibid.
16 A.L. Yuille and C. Liu, "Deep Nets: What have they ever done for Vision?" *arXiv* (2018): 3.

is still a dog whether looking from in front or behind. The development of better methods for overall scene understanding and general contextualization is a current challenge. A self-driving car would become safer if it understood its surroundings rather than simply observing and reacting to them. Therefore, current models of computer vision still have room for improvement.

Reinforcement Learning

Responses from external stimuli affect human behavior. The brain region responsible for learning from reward and punishment signals is known as the basal ganglia.[17] This specialized area directly influences the actions humans take, which results in a form of learning called reinforcement learning. Dopamine neurons activate when a reward is delivered, conditioning the brain to expect a reward upon taking a given action.

This topic of conditioning is one that arises under procedural learning through associations.[18] This type of learning deals with procedural memories, in that a given sensory input generates a given motor or sensory output. These reactions can be learned, or conditioned, along reflex pathways in the brain. Classical conditioning is a type of associative learning that involves learning an association from two different stimuli to the same individual response.[19] As seen in the famous experiment from Russian physiologist Ivan Pavlov, depicted by Figure 4, dogs were conditioned to salivate upon hearing a certain sound. The figure below illustrates the conditioning process.

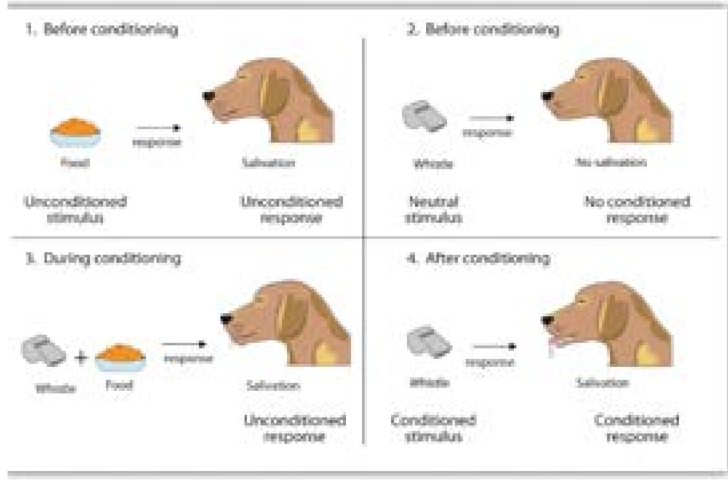


Figure 4: An example of classical conditioning.[20]

The dog in the experiment learns a new association between the presence of the food and an auditory stimulus. This reaction is essentially a form of reinforcement learning: the dog will predict food to arrive upon hearing the same sound, thus resulting in the expected response. Conditioning and reinforcement learning processes are evolutionarily beneficial[21] because they allow for the development of new associations between particular stimuli and good or bad events. Humans are constantly making decisions based on associations created to benefit survival. Learning through trial and error is one of the main methods of learning about the world from a young age. This type of trial-and-error cognitive reinforcement learning provides the inspiration for cutting-edge computational models employed for deep learning. DeepMind's program AlphaZero learned to master the game of chess through the process of reinforcement learning. Using a deep neural network architecture of many layers, it learns by self-play. Based on results from wins, losses, and draws, it adjusts the network's parameters to make it more likely to choose better moves in the future.[22]

However, AlphaZero is unique because it is general-purpose. It is able to learn any two-player perfect information game given just the set rules to follow. This is a big step forward in AGI, but a potential route for further research would be to implement transfer learning. For example, a person who knows how to

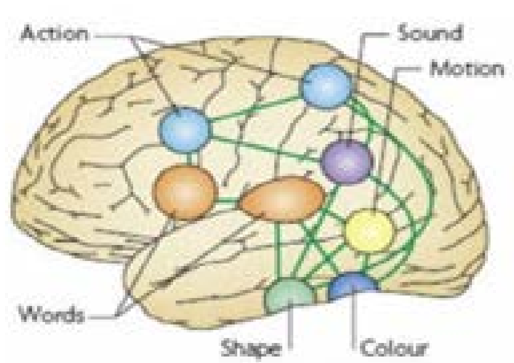
17 R.C. O'Reilly et al., *Computational Cognitive Neuroscience*, 1st ed. (Wiki Book, 2012), 78.
18 Mark F. Bear et al., *Neuroscience: exploring the brain* 4th ed., (Philadelphia: Wolters Kluwer, 2016, 827).
19 Ibid; Charles Stagnor, "Learning by Association: Classical Conditioning" in *Introduction to Psychology* 1st ed. (Boston, MA: Flatword, 2018), https://open.lib.umn.edu/intropsyc/chapter/7-1-learning-by-association-classical-conditioning/.
20 Stagnor, *Introduction to Psychology.*
21 Ibid.
22 Silver et al, "Mastering chess and shogi by self- play with a general reinforcement learning algorithm."

play the guitar may be able to pick up the piano faster than someone with no musical background.[23] Learning from reinforcement and then transferring gained knowledge to new tasks is a skill in which the human brain excels, so perhaps AI research could benefit from further inspiration from neuroscience. Reinforcement learning methods need to be more generalized, which would benefit robotics as well. Autonomous agents would have a better understanding of their environment and would be able to make even better decisions about which actions to take. Logical decision making is the fundamental goal of artificial general intelligence, so further research into reinforcement learning would help in this endeavor.

Natural Language Processing

In addition to translating knowledge to a variety of skills, the human brain also excels in processing one of the most complex forms of communication: language. People understand each other based on individual systems of semantic representations, or relationships between words and symbols to their meaning in reality. This ability to understand complex connections is the idea behind a dominant notion of amodal, or abstract memory.[24] Figure 5 depicts this cognitive process displayed in the hub-and-spoke model.

Figure 5: The hub-and-spoke model is a hybrid model of semantic memory, containing both amodal and grounded representations



in sensory and motor systems.[25] Processes similar to that of mental imagery take place, which can make sense intuitively. People essentially translate generated mental images and associations into representative words. These words are ingrained through a training or learning process, as people develop those connections, or associations.

For example, consider the word "elephant" and one may recall associated properties such as long trunk, animal, big ears, or gray skin. Because of these varied associations, a linguistic symbol grounding problem arises in that there must be a way to define words without the use of other words.

The idea of embodied cognition offers grounded associations, a solution to the symbol grounding problem. Theories of embodied cognition suggest that "neural systems involved in understanding real objects, actions, and events in the world are used to internally simulate those objects, actions, and events at later points in time."[26] As such, words or concepts can additionally be associated with simulations of sensorimotor activity. For example, the concept of "greeting a person" may be associated with the motor activity of shaking a hand. This additional sensorimotor simulation adds to the brain's understanding of a given concept. Also, it has been shown that perceptual representations can substantially increase the computational power of artificial intelligence systems.[27] Current language modelling networks in the industry may take slightly different approaches.

Similar to computer vision tasks, language reading may also take the form of object recognition. Brains go through a form of a training process when first learning how to read. Connections are made between visually interpreted symbols to their corresponding sound and ultimately to their corresponding objective meaning. Studying this phenomenon gives us insight into how the brain processes visual symbolic representations of words and ideas.

Computational algorithms that automatically analyze and represent human language fall under the category of Natural Language Processing (NLP).

23 K. Weiss, T. Khoshgoftaar, and D. Wang, "A survey of Transfer Learning." *Journal of Big Data* (2016), https://doi.org/10.1186/s40537-016-0043-6.

24 Jamie Ward, *The Student's Guide to Cognitive Neuroscience,* 3rd ed. (London: Psychology Press, 2015), 31.

25 Ibid.; Greg Hickok, Adapted from Patterson et al., 2007. "Semantics and Brain," *Talking Brains,* January 15, 2008, http://www.talkingbrains.org/2008/01/semantics-and- brain-more-on-atl-as-hub.html.

26 Guy Dove, "On the need for embodied and dis-embodied cognition," *Frontiers in Psychology* 1, (2011): 242.

27 J.I.Glasgow, "The Imagery Debate Revisted: A Computational Perspective," *Computational Intelligence* 9 (1993): 310-333, doi:10.1111/j.1467-8640.1993.tb00224.x.

One deep learning approach to NLP is the use of recurrent neural networks. This type of model uses a computational structure known as "Long Short-Term Memory" which is able to process through sequences of data with a temporal internal state memory.[28] An example that utilizes this technology is Google Translate.[29] Google Translate demonstrates Neural Machine Translation, in which it encodes the source language and decodes to the target language. It uses sequence to sequence modeling to summarize the source sentence in the translated version. Language modeling has seen great breakthroughs such as Google Translate, but there are still limitations. Semantic understanding is the direction for further research. Exactly how people acquire languages and develop higher-level semantic representations may provide insights into ways to improve state-of-the-art natural language processing applications.

Conclusion

The current technological revolution is rapidly developing. Computational cognitive neuroscience has set the foundation for deep learning, seen in the power of artificial neural networks. Inspired by areas of human cognition, including vision, reinforcement learning, and natural language processing, computational models have been developed in an effort to mimic the complexity of the human brain. However, there is still room for improvement as the function of intelligence has yet to be fully solved. Many say that AGI is a long way away, but with a stronger coordinated effort between neuroscientists and deep learning researchers, general intelligence could be fully understood sooner rather than later.

28 S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation* 9, no. 8 (November 1997): 1735-1780, https://doi.org/10.1162/neco.1997.9.8.1735.

29 Yonghui Wu et al., "Google's neural machine translation system: Bridging the gap between human and machine translation." arXiv (2016), https://arxiv.org/abs/1609.08144.

Bibliography

"AI and the world's complex challenges." DeepMind Technologies Limited, 2019. https://deepmind.com/applied/deepmind-ethics-society/research/AI-worlds-complex-challenges/.

Bear, Mark F., Barry W. Connors, and Michael A. Paradiso. *Neuroscience: Exploring the Brain* 4th ed. Philadelphia: Wolters Kluwer, 2016,

Cassimatis, Nicholas L. "Artificial intelligence and cognitive modeling have the same problem." In *Theoretical Foundations of Artificial General Intelligence,* 11-24. Paris: Atlantis Press, 2012.

Diehl, Peter U. and Matthew Cook. "Unsupervised learning of digit recognition using spike-timing-depen dent plasticity." *Frontiers in Computational Neuroscience* 9, (2015): 99.

Dove, Guy. "On the need for embodied and dis-embodied cognition." *Frontiers in Psychology* 1 (2011): 242.

Glasgow, J.I. "The Imagery Debate Revisted: A Computational Perspective." *Computational Intelligence* 9 (1993): 310-333, doi:10.1111/j.1467-8640.1993.tb00224.x.

Goertzel, Ben. "Artificial general intelligence: concept, state of the art, and future prospects." *Journal of Artificial General Intelligence* 5, no. 1 (2014): 1-48.

Goldstein, E. Bruce. *Sensation and Perception*. 8th ed. Boston, MA: Cengage Learning, 2010, 8-9.

Haier, Richard J. *The Neuroscience of Intelligence*. Cambridge, UK: Cambridge University Press, 2016, 4-5.

Hickok, Greg. Adapted from Patterson et al., 2007. "Semantics and Brain." *Talking Brains*. January 15, 2008. http://www.talkingbrains.org/2008/01/semantics-and-%20brain-more-on-atl-as-hub.html.

Hochreiter, Sepp and Jürgen Schmidhuber. "Long short-term memory." *Neural Computation* 9, no. 8 (1997): 1735-1780.

Olah C., Satyanarayan A., Johnson I., Carter S., Schubert L., Ye K., and Mordvintsev A. "The building blocks of interpretability," *Distill 3* no. 10. (2018) https://doi.org/10.23915/distill.00010.

O'Reilly, R.C., Munakata Y., Frank, M.J., Hazy, T.E., and contributors. *Computational Cognitive Neuro science,* 1st ed. (Wiki Book, 2012), 3-78.

Parmar, Ravindra. "Training Deep Neural Networks," Towards Data Science. September 11, 2018, https://towardsdatascience.com/training-deep-neural-networks-9fdb1964b964.

Savage, Neil. "How AI and neuroscience drive each other forwards." *Nature* 571 (2019), 10.1038/d41586-019-02212-4.

Siegfried, Tom. "There's a long way to go in understanding the brain." July 25, 2017, https://www.sciencenews.org/blog/context/neuroscience-understanding-brain.

Silver, D. et al., "Mastering chess and shogi by self- play with a general reinforcement learning algorithm." arXiv (2017). https://arxiv.org/abs/1712.01815

Stagnor, Charles. “Learning by Association: Classical Conditioning” in *Introduction to Psychology* 1st ed (Boston, MA: Flatword, 2018), https://open.lib.umn.edu/intropsyc/chapter/7-1-learning-by-association-classical-conditioning/.

Tavanaei, Amirhossein, Masoud Ghodrati, Saeed Reza Kheradpisheh, Timothée Masquelier, and Anthony Maida. “Deep learning in spiking neural networks.” *Neural Networks* 111, (2019): 47-63.

Ward, Jamie. *The Student's Guide to Cognitive Neuroscience* 3rd ed., London: Psychology Press, 2015.

Weiss, Karl, Taghi M. Khoshgoftaar, and DingDing Wang. “A survey of transfer learning.” *Journal of Big Data* 3, no. 1 (2016).

Wu, Yonghui, Schuster, Mike, Zhifeng, Chen, Le, Quoc V., Norouzi, Wolfgang, Mohammad, Macherey, Krikun, Maxim et al. “Google's neural machine translation system: Bridging the gap between hu man and machine translation.” arXiv. arXiv:1609.08144 (2016).

Yuille, Alan L., and Chenxi, Liu. “Deep Nets: What have they ever done for Vision?.” arXiv arX iv:1805.04025 (2018): 3.