Cognition-Based Deep Learning: Progresses and Perspectives

Kai Yi, Shitao Chen, Yu Chen, Chao Xia, Nanning Zheng

Cognition-based Deep Learning: Progresses and Perspectives

Kai Yi[1,2], Shitao Chen[1,2], Yu Chen[1,2], Chao Xia[1,2], Nanning Zheng[1,2,3]

[1] Institute of Artificial Intelligence and Robotics,
Xian Jiaotong University, Xian, Shaanxi, P.R.China
[2] National Engineering Laboratory for Visual Information
Processing and Applications,
Xi'an Jiaotong University, Xi'an, Shaanxi, P.R.China
{yikai2015, chenshitao, alan19960212, xc06210417}@stu.xjtu.edu.cn
[3] Correspondence: nnzheng@mail.xjtu.edu.cn

Abstract. The human brain is composed of multiple modular subsystems, with a unique way interacting among. These subsystems have their own unique characteristics and interact to support cognitive functions such as memory, attention and cognitive control. Nowadays, deep learning methods based on the above-mentioned functions accompanied with knowledge are widely used to design more dynamic, robust and powerful systems. We first review and summarize the progresses of cognition-based deep neural networks, and how cognitive mechanisms can be applied to more brain-like neural networks. Then we propose a general framework for the design of cognition-based deep learning system. Although great efforts have been made in this field, cognition-based deep learning is still in its early age. We put forward the potential directions towards this field, such as associative memory in deep learning, interpretable network with cognitive mechanisms, and deep reinforcement learning based on cognitive science.

Keywords: Cognitive mechanisms · Deep learning · Attention · Memory · Knowledge · Decision-making

1 Introduction

In recent years, deep learning, characterized by learning large neural-network-style models with multiple layers of representations, has received lots of attention. Those models based on deep learning have achieved remarkable gains in many domains, including image classification [1], [2], [3], control and decision-making [4], [5]. [1] trained a deep convolutional neural network that nearly halved the error rate of the previous state-of-the-art methods on the typical image classification dataset. In recent years, [2] even surpass the human-level performance on several challenging classification datasets. In the field of decision-making, deep learning in combination with reinforcement learning, have been widely used to play games. [5] achieved several human even superhuman level performance on

several games, illustrating that computer can master go learning from scratch by trial-and-error strategies. These accomplishments have helped Deep Neural Networks (DNNs), the core of deep learning, to regain their status as a leading paradigm in machine learning.

Nowadays, deep learning has shown more and more interests on how our brain works and how an artificial intelligence system can be inspired by cognitive science. The human brain doesn't learn through a unified undifferentiated neural network. It is composed of multiple modular subsystems, with a unique way interacting among. These subsystems have their own unique characteristics, and interact to support cognitive functions, such as memory, attention, language and cognitive control. What's more, the brain can combine knowledge (including internal knowledge from self experience, environment knowledge by interacting with objects around, global knowledge extracted from the universe), with different cognitive functions to conduct complicated tasks with a few data.

In this article, we review the latest progresses and future perspectives of deep learning systems based on the cognitive core elements, especially memory, attention and knowledge. In section two, we will review the fundamental concepts of deep learning, cognitive mechanism, and what deep neural networks can benefit from cognitive science. In section three to section five, we review and summarize the latest progresses of deep learning methods based on memory, attention and knowledge respectively. In section six, we propose a general framework of cognition-based deep learning and make assumptions of the essential future directions towards this field. In the last section, we'll make conclusions about all above-mentioned issues concerning cognition-based deep learning.

2 Fundamental Concepts

In this section, we will review the basic concepts of deep learning, cognitive mechanism, as well as the reason and the way that the former two are combined. We first introduce the core concepts of deep learning and it's foundations (i.e., two kind of neural works). Then, the major mechanisms of cognitive science, especially those that have been applied to more powerful deep learning systems will be presented. Further, we give some directions on how to build deep learning systems inspired or based on important elements of cognitive science.

2.1 Deep Learning

Deep learning is a kind of computational methods. It is composed of multiple processing layers which learn and represent the feature and the distribution of input data with multiple levels of abstraction, (i.e., different depth of feature map). Nowadays, the success of deep learning can own to two branches of well-designed neural networks: Convolutional Neural Networks (CNNs or ConvNets) and Recurrent Neural Networks (RNNs).

ConvNets are feature extractor indeed. They are very excellent at dealing with structured 2D arrays in areas like image processing. ConvNets achieved a

state-of-the-art performance on image classification [1], [2], [3], image segmentation [6], and object detection [7], [8]. RNNs are excellent at dealing with sequence inputs, such as speech and language [9], [10]. RNNs process an input sequence one element at a time, maintaining in their hidden units a "state vector" that implicitly contains information about the history of all the past elements of the sequence. For more detailed explanations, please refer to [11].

2.2 Cognitive Mechanisms

Cognitive science is a discipline and a recognition of a fundamental set of common concerns shared by psychology and artificial intelligence [12]. The key point of cognitive science is the way we reflect to our environment and the effect of our brain activities. Our brain is composed of several subsystems, which interact with each other in a very complicated way. And they communicate to support cognitive functions, including attention, memory, language and cognitive control. The combination of the above-mentioned functions with knowledge extracted from self experience, environments, intuitive psychology and physical worlds are among the key characteristics of human.

2.3 Combination

The most important elements of cognitive science are attention, memory and knowledge. The last one can be classified to internal knowledge, environment knowledge and global knowledge. Deep learning system can benefit from every element separately or together by increasing dynamics and target-oriented accuracy with fewer training data. Further, the cognitive science mechanisms are derived from human brains. Deep learning, which is regarded as the "black box", can be interpreted at the aspect of cognitive science, such as the view of decision tree [13] and the view of shape bias [14].

3 Deep Learning Inspired by Memory Mechanism

From the perspective of cognitive science and our intuitions, humans can learn in a continuous spatial sequence and memorize the pattern and characteristics. Therefore applying this mechanism to deep learning systems is of vital importance. How to guide those systems to memorize sequences of input and how to memorize according to the relative importance are two topics that will be reviewed next.

3.1 RNNs-based Memory Model

RNNs is a type of artificial neural network where connections between units form a directed cycle. This creates an internal state of the network which allows it to exhibit dynamic behavior. Unlike feed forward neural networks, RNNs can use their internal memory to process and work on arbitrary sequences of inputs. This

makes deep neural networks applicable to handle the data with rich structures, especially the sequences. It has been widely used in speech recognition [15], [16], natural language processing [10], [17] and object detection [18], [19].

Traditional RNNs structures have been used to deal with sequences of input mentioned above, but [20] found it is difficult to train RNNs to capture long-term dependencies because the gradient tends to either vanish or explode, which can cause severe effects. Long Short-Term Memory (LSTM) [21] and Gated Recurrent Units (GRU) [22] are two of the well-designed recurrent neural networks which can elevate this problem. The idea of those two RNNs is to design a more sophisticated activation function than an usual one, consisting of affine transformation followed by a simple element-wise nonlinearly part by using gating units.

3.2 Memory Model with Importance

In the process of human learning, the old information we stored will be overwritten by new incoming information[23]. However, what we memorized are rarely with equal importance: frequently used and important knowledge is often prevented from being erased. Inspired by the memory mechanism, we think evaluating what is important and what is not behind the structure of a deep neural network is really meaningful. Elastic Weight Consolidation [24] used an approximation of the diagonal term of the Fisher information matrix to identify the important parameters for the task. While training a new task, a regularizer is used to prevent those important weights from being overwritten by the new task. The Fisher information matrix needs to be computed in a separate phase after each task and also needs to be stored for each task for later use when learning a new task. Thus, this model stores a large number of parameters that grows with the number of executed tasks. To avoid this, Improved multi-task learning through synaptic intelligence [25] adopts an online way of computing the importance of the network parameters. [26] formulates the importance of memorized information as the absolute gradient of the parameters in deep neural network based on the sensitivity of the predicted output to a change in this parameter. When learning a new task, changes to important parameters are penalized. The memory-based deep learning method shows the ability to adapt the importance of the parameters towards what the network needs (not) to forget.

4 Attention Mechanism Applied to Deep Learning

Human attention is a built-in mechanism for deciding how to apply their brainpower from moment to moment, (e.g., decide where to see in saliency visual object detection [27]). Attention mechanism is a reasonably well studied subject within the field of cognitive psychology and is known to be a key feature of human artificial intelligence [28]. Nowadays, attention-based deep learning methods are active especially in dealing problems concerning sequence prediction or control, including object detection, natural language processing, and deep reinforcement learning.

4.1 Natural Language Processing

The seminal work of natural language processing with attention was proposed by [29] for English-to-French translation. They used a novel neural machine translation model that implements an attention mechanism in the decoder, which achieved much better performance than traditional phrase-based models. In order to allow parallelisation, [30] proposed a highly parallelisable multi-hop attention module with convolutional neural network, which takes multiple glimpses at the sentence to determine what will be translated next, and a separate attention module in each decoder layer. Moreover, the attention mechanisms have been widely used in other language processing fields, like text classification [31], [32] and text understanding [33].

4.2 Object Detection

Attention mechanism in object detection decides which object or which field we need to see. Recurrent attention models are widely used to deal with the problem. [34] represented how an attention mechanism can be applied to ignore irrelevant objects in a scene and how an object can be "seen" by the system with the help of the attention mechanism. [35] proposed a deep recurrent neural network trained with reinforcement learning to detect multiple objects automatically. [36] found the attention models with deep neural networks are the insights gained by approximately visualizing where and what the attention focused on (i.e., what the model "sees") after feeding a sequence of data. Besides, attention models with deep learning are hot methods in several topics related to object detection, including saliency detection [37], [38] (detecting the most salient object and segmenting the accurate region of that object), eye fixation [39], [40], [41] (maintaining the visual gaze on a single location).

4.3 Deep Reinforcement Learning

Deep reinforcement learning is widely used in decision-making and control. Deep Q-Network (DQN) proposed by [4] showed a single algorithm that can outperform human or even superhuman performance on Atari 2600 games. By combining the attention mechanisms into DQN, [42] proposed the Deep Attention Recurrent Q-Network (DARQN). By combining what they called "soft" and "hard" attention mechanisms, the model greatly outperformed the traditional DQN. The attention network takes the current game state as input and generates a context vector based on the features observed. Then an typical LSTM network takes the context vector with a previous hidden state and the memory state to evaluate the action that an agent can take. Further, [43] improved on DARQN by implementing a multi-focus attention network where the agent has the ability to attend to multiple important elements. They proposed a novel model by using multiple parallel attention to attend to entities concerning the problem instead of just one attention layer in DARQN.

5 Deep Neural Networks with Knowledge

Humans can combine different kinds of knowledge in a complicated manner to solve very difficult problems without being trained with plenty of data. On the one hand, our brains use knowledge accompanied with other elements (e.g. attention, reasoning) to realize associative memory and build high-level concepts. On the other hand, the logical and physical constraints derived from our knowledge can be used to build a more robust model, especially towards solving natural problems as it can be influenced by too many factors. Beriefly speaking, our knowledge originated from three different parts: self experience (i.e., internal knowledge), surrounding objects (i.e., environment knowledge) and universe (i.e., global knowledge). Human brains process and summarize that knowledge into three categories: intuitive originates psychology knowledge, intuitive derives from physical knowledge and domain specific knowledge.

5.1 Intuitive Psychology

Humans gain plenty of psychology knowledge by interacting with the environment. Infants can understand mental states of other people like beliefs and goals, and this understanding strongly guides and constraints decisions that they make [44]. Besides, humans tend to assign the same name to similarly shaped items rather than to items with similar color, texture, or size by psychological experiments [45]. Those psychology intuitions can help create more interpretable neural networks and create a new area of one-shot learning. [14] found that several well-performed one-shot learning models trained on ImageNet exhibit a similar bias to that observed in humans: they prefer to categorize objects according to shape rather than color. Inspired by cognitive psychology, [46] proposed shape Matching Network (MN) with inception network, which has the state-of-the-art one-shot learning performance on ImageNet.

5.2 Intuitive Physics

Deep learning can learn features and patterns not only from plenty of labeled data, but also from physical laws. The limitation and constraints can help neural networks learn from fewer labeled data, even without any labeled data (i.e., unsupervised learning). Further, deep learning methods with physical constraints can help build high-level structural models and solve complicated scientific problems. In many fields, labeled data and long-time training is scarce and obtaining more labels is expensive. Constraint learning with physical knowledge is another active field of machine learning, which is aimed at uncover the hidden structure of models. By using physical knowledge, [47] trained a convolutional neural network to detect and track objects without any labeled examples.

5.3 Domain Knowledge

Regulating deep neural networks (DNNs) with human structured domain knowledge has been confirmed to be of great benefit for improved accuracy and in-

terpretability with fewer training data. Recently, [48] proposed a general distillation framework that transfers knowledge into neural networks by combining first-order logic (FOL), where FOL constraints are integrated via posterior regularization [49]. Further, [50] used a generalized framework that makes it possible to learn procedure for knowledge representations and adapt their weights jointly with the help of the regulated DNN models. [50] proposed to transfer logical knowledge information into neural networks with diverse architectures such as recurrent networks and convolutional networks.

6 Perspectives

Cognition-based deep learning has become one hot research topic, and some of the most important functions of our human brains like memory and attention associated with knowledge extracted from experience and the universe, have been widely used in the design of a more human-like deep learning system. Meanwhile, the brain does not learn through a unified undifferentiated neural network. The brain is composed of multiple modular subsystems, with an unique and complicated way interacting among. Although deep neural network can process structural data well, it can't deal with dynamic clouds of data. What's more, data is really scarce in some fields. Deep learning systems can get a lot of inspirations from cognitive science, to alleviate and even eliminate those problems.

In this section, we will discuss the essential trend to apply more elements of cognitive science to build more dynamic, robust and intelligent deep learning systems. We are going to give a general framework of cognition-based learning firstly. Then we will discuss the key problems of fusing deep neural network with cognitive mechanisms and essential solutions.

6.1 General Framework of Cognition-based Deep Learning

We suggest the general framework of designing cognition-based deep learning systems. This framework use cognitive mechanisms in a particular way. It can help build more dynamic, robust and intelligent systems. More accurately speaking, it can process unstructured data as constructed one with the help of our memory with concepts, especially associative memory. Also the system based on the proposed framework can reasoning and infer based on the knowledge by gaining structural feature map with hierarchical knowledge sets in the top-down manner. Every layer in hierarchical knowledge sets is corresponding to each layer in the structural feature maps. As feedback is also very essential in our human brains, we can monitor this mechanism by designing two feedback loops. One is knowledge feedback loop, to update our knowledge based on attention select network, which is aimed at deciding what we need to see. Another is memory feedback loop, to update our memory (especially the experience), and gains high-level concepts after measuring actions/decisions the system make. The general framework is shown in Fig 1. The model based on this framework can be suitable for plenty domains such as image processing and nature language processing.

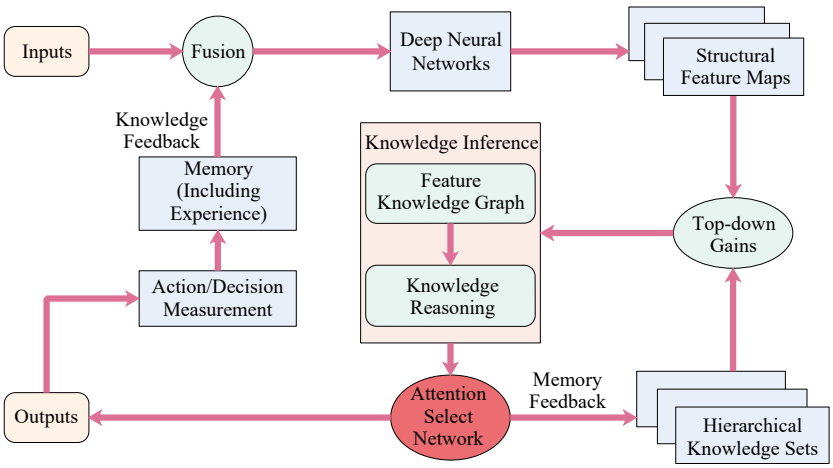

Fig. 1: General framework of cognition-based deep learning

6.2 Key Problems and Potential Solutions

This part will discuss the future directions of cognition-based deep learning. It is organized by current problems and essential solutions.

6.3 Associative Memory

Human brains can associate patterns similar to the input patterns when being stimulated. Associative memory model was once prevailing in 1980s and 1990s, accompanied with the popularity of Hopfield Neural Network [51], a typical network that can store patterns and realize associative memory. Due to the potential chaos state of network evolution, HNNs alone is difficult to handle natural real-world problems well. However, it has the potential as it is an important kind of brain-like neural network. Besides, synesthesia is a typical perceptual phenomenon in cognitive science. That is, a person can activate a sensory when stimulated by another sensory (e.g., grapheme-color synesthesia means a person can directly associate a colorful image when listening to music). The proposition of an effective associative memory model by combining human-like neural network and synesthesia with deep learning is a promising direction. A recent successful attempt was Dense Associative Memory [52], which combined associative memory with deep learning and achieved a good result on MNIST dataset.

6.4 Interpretable Network with Cognitive Mechanisms

For we human, it's difficult to understand how deep neural networks work and how they react towards a task. However, interpretable systems in many applications are of vital importance. For example, suppose that there is a person who

may be in the early stage of cancer, the system based on deep neural networks needs to infer whether he is suffering from cancer. We can gather all features of the person as the input of DNNs, such as age, history of disease. The question is why we can trust the output of this system as we can not check the correctness. What if the process of inference can be understood or monitored (e.g., the decision tree) by an expert? Interpretability is important in these fields.

[13] proposed a tree regularization to interpret the neural network in the perspective of decision tree. This method can not train towards the typical backpropagation learning rule as the tree is undifferentiated. They suggested replacing trees with multi-layer perceptrons in the training phase, but this solution is not very elegent and does not create a really interpretable network indeed. According to psychological experiments [45], humans tend to assign the same name to similarly shaped items instead of items with similar color, texture or size. [46] proposed shape MN with inception network, which achieved better performance than several state-of-the-art methods in the field of one-shot learning on ImageNet. And [14] found that this kind of networks that exhibits a similar shape bias to that observed in humans. Cognitive mechanisms like shape bias, decision and inference can help design more interpretable neural networks.

6.5 Cognition-based Deep Reinforcement Learning

Deep reinforcement learning has raised a lot of interests nowadays. However, due to the uncertainty of the state space and the complexity of the reward function, it is difficult for the traditional trial-and-error strategies to associate continuous actions with reward. Imagination is utilized to make use of the knowledge embedded in the model. However, deep reinforcement learning is still in its early stage.

As decision making and feedback mechanism are very similar to that of humans, there is a trend to apply cognitive mechanisms to reinforcement learning. As for attention mechanism, [42] proposed the Deep Attention Recurrent Q-Network (DARQN), which greatly outperformed the traditional Deep Q-Network (DQN) on Atari 2600 games by combining what they called "soft" and "hard" attention mechanisms. Besides, [43] used a multi-focus attention network where the agent can give attention to multiple important elements. This model achieved better performance than the traditional DARQN. Further, [53] extended the typical LSTM-based memory network to choose more sophisticated addressing schemes over the past k frames by using memory mechanism. [54] proposed a spatially structured 2D memory image that is capable of learning to store arbitrary information about the environment over long time lags. As for our human knowledge mechanism, in which field we call usually transfer learning. What's more, a novel policy distillation (i.e., knowledge-based reinforcement learning policy) architecture was proposed by [55] for deep reinforcement learning. This architecture was well organized by implementing task-specific high-level convolutional features as the inputs to the multi-task policy network. However, how to hierarchically reconstruct the knowledge and uncover the hidden characteristics, how to abstract our knowledge and experience for the feasibility to deal

with unstructured data by fusion, how to design a generalized attention selection network, may remain issues that lead the future research direction in this field.

7 Conclusion

Cognition-based deep learning has gained widely interests recent years. Several core functions of cognitive science (i.e. attention, memory) and knowledge, are used to design more dynamic and robust systems based on deep neural network. We reviewed the recent progress related to this field. Meanwhile, deep neural networks are not interpretable to our human brains. We can design a more interpretable neural network in the perspective of cognitive science. Finally, we proposed a general framework of cognition-based deep learning and made assumptions of the essential future directions towards this field.

Acknowledgment

This research was partially supported by the Programme of Introducing Talents of Discipline to University (No. B13043), the National Natural Science Foundation of China (No. 61773312, 61790563).

References

1. Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.
2. Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.
3. Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, Andrew Rabinovich, et al. Going deeper with convolutions. Cvpr, 2015.
4. Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Alex Graves, Ioannis Antonoglou, Daan Wierstra, and Martin Riedmiller. Playing atari with deep reinforcement learning. *arXiv preprint arXiv:1312.5602*, 2013.
5. Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A Rusu, Joel Veness, Marc G Bellemare, Alex Graves, Martin Riedmiller, Andreas K Fidjeland, Georg Ostrovski, et al. Human-level control through deep reinforcement learning. *Nature*, 518(7540):529, 2015.
6. Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3431–3440, 2015.
7. Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster r-cnn: Towards real-time object detection with region proposal networks. In *Advances in neural information processing systems*, pages 91–99, 2015.
8. Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, and Alexander C Berg. Ssd: Single shot multibox detector. In *European conference on computer vision*, pages 21–37. Springer, 2016.

9. Alex Graves, Abdel-rahman Mohamed, and Geoffrey Hinton. Speech recognition with deep recurrent neural networks. In *Acoustics, speech and signal processing (icassp), 2013 ieee international conference on*, pages 6645–6649. IEEE, 2013.
10. Wenpeng Yin, Katharina Kann, Mo Yu, and Hinrich Schütze. Comparative study of cnn and rnn for natural language processing. *arXiv preprint arXiv:1702.01923*, 2017.
11. Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *nature*, 521(7553):436, 2015.
12. Herbert A Simon. Cognitive science: The newest science of the artificial. *Cognitive science*, 4(1):33–46, 1980.
13. Mike Wu, Michael C Hughes, Sonali Parbhoo, Maurizio Zazzi, Volker Roth, and Finale Doshi-Velez. Beyond sparsity: Tree regularization of deep models for interpretability. *arXiv preprint arXiv:1711.06178*, 2017.
14. Samuel Ritter, David G. T Barrett, Adam Santoro, and Matt M Botvinick. Cognitive psychology for deep neural networks: A shape bias case study. 2017.
15. Haşim Sak, Andrew Senior, and Françoise Beaufays. Long short-term memory recurrent neural network architectures for large scale acoustic modeling. In *Fifteenth annual conference of the international speech communication association*, 2014.
16. Yajie Miao, Jinyu Li, Yongqiang Wang, Shi-Xiong Zhang, and Yifan Gong. Simplifying long short-term memory acoustic models for fast training and decoding. In *Acoustics, Speech and Signal Processing (ICASSP), 2016 IEEE International Conference on*, pages 2284–2288. IEEE, 2016.
17. Prakash M Nadkarni, Lucila Ohno-Machado, and Wendy W Chapman. Natural language processing: an introduction. *Journal of the American Medical Informatics Association*, 18(5):544–551, 2011.
18. Ming Liang and Xiaolin Hu. Recurrent convolutional neural network for object recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3367–3375, 2015.
19. Subarna Tripathi, Zachary C Lipton, Serge Belongie, and Truong Nguyen. Context matters: Refining object detection in video with recurrent neural networks. *arXiv preprint arXiv:1607.04648*, 2016.
20. Yoshua Bengio, Patrice Simard, and Paolo Frasconi. Learning long-term dependencies with gradient descent is difficult. *IEEE transactions on neural networks*, 5(2):157–166, 1994.
21. Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.
22. Kyunghyun Cho, Bart Van Merriënboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. Learning phrase representations using rnn encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*, 2014.
23. John Jonides, Richard L Lewis, Derek Evan Nee, Cindy A Lustig, Marc G Berman, and Katherine Sledge Moore. The mind and brain of short-term memory. *Annu. Rev. Psychol.*, 59:193–224, 2008.
24. James Kirkpatrick, Razvan Pascanu, Neil Rabinowitz, Joel Veness, Guillaume Desjardins, Andrei A Rusu, Kieran Milan, John Quan, Tiago Ramalho, Agnieszka Grabska-Barwinska, et al. Overcoming catastrophic forgetting in neural networks. *Proceedings of the National Academy of Sciences*, 114(13):3521–3526, 2017.
25. Friedemann Zenke, Ben Poole, and Surya Ganguli. Improved multitask learning through synaptic intelligence. *arXiv preprint arXiv:1703.04200*, 2017.

26. Rahaf Aljundi, Francesca Babiloni, Mohamed Elhoseiny, Marcus Rohrbach, and Tinne Tuytelaars. Memory aware synapses: Learning what (not) to forget. *arXiv preprint arXiv:1711.09601*, 2017.
27. Tie Liu, Zejian Yuan, Jian Sun, Jingdong Wang, Nanning Zheng, Xiaoou Tang, and Heung-Yeung Shum. Learning to detect a salient object. *IEEE Transactions on Pattern analysis and machine intelligence*, 33(2):353–367, 2011.
28. Helgi Pll Helgason. General attention mechanism for artificial intelligence systems. 2013.
29. Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. Neural machine translation by jointly learning to align and translate. *arXiv preprint arXiv:1409.0473*, 2014.
30. Jonas Gehring, Michael Auli, David Grangier, Denis Yarats, and Yann N Dauphin. Convolutional sequence to sequence learning. *arXiv preprint arXiv:1705.03122*, 2017.
31. Zichao Yang, Diyi Yang, Chris Dyer, Xiaodong He, Alexander J. Smola, and Eduard H. Hovy. Hierarchical attention networks for document classification. In *HLT-NAACL*, 2016.
32. Devendra Singh Chaplot, Kanthashree Mysore Sathyendra, Rama Kumar Pasumarthi, Dheeraj Rajagopal, and Ruslan Salakhutdinov. Gated-attention architectures for task-oriented language grounding. *arXiv preprint arXiv:1706.07230*, 2017.
33. Bhuwan Dhingra, Hanxiao Liu, Zhilin Yang, William W Cohen, and Ruslan Salakhutdinov. Gated-attention readers for text comprehension. *arXiv preprint arXiv:1606.01549*, 2016.
34. Volodymyr Mnih, Nicolas Heess, Alex Graves, et al. Recurrent models of visual attention. In *Advances in neural information processing systems*, pages 2204–2212, 2014.
35. Jimmy Ba, Volodymyr Mnih, and Koray Kavukcuoglu. Multiple object recognition with visual attention. *Computer Science*, 2014.
36. Kelvin Xu, Jimmy Ba, Ryan Kiros, Kyunghyun Cho, Aaron Courville, Ruslan Salakhudinov, Rich Zemel, and Yoshua Bengio. Show, attend and tell: Neural image caption generation with visual attention. In *International Conference on Machine Learning*, pages 2048–2057, 2015.
37. Jason Kuen, Zhenhua Wang, and Gang Wang. Recurrent attentional networks for saliency detection. *arXiv preprint arXiv:1604.03227*, 2016.
38. Guanbin Li and Yizhou Yu. Visual saliency based on multiscale deep features. *arXiv preprint arXiv:1503.08663*, 2015.
39. Nian Liu, Junwei Han, Dingwen Zhang, Shifeng Wen, and Tianming Liu. Predicting eye fixations using convolutional neural networks. In *Computer Vision and Pattern Recognition (CVPR), 2015 IEEE Conference on*, pages 362–370. IEEE, 2015.
40. Srinivas SS Kruthiventi, Kumar Ayush, and R Venkatesh Babu. Deepfix: A fully convolutional neural network for predicting human eye fixations. *IEEE Transactions on Image Processing*, 26(9):4446–4456, 2017.
41. Marcella Cornia, Lorenzo Baraldi, Giuseppe Serra, and Rita Cucchiara. Predicting human eye fixations via an lstm-based saliency attentive model. *arXiv preprint arXiv:1611.09571*, 2016.
42. Ivan Sorokin, Alexey Seleznev, Mikhail Pavlov, Aleksandr Fedorov, and Anastasiia Ignateva. Deep attention recurrent q-network. *arXiv preprint arXiv:1512.01693*, 2015.

43. Jinyoung Choi, Beom-Jin Lee, and Byoung-Tak Zhang. Multi-focus attention network for efficient deep reinforcement learning. *arXiv preprint arXiv:1712.04603*, 2017.
44. Brenden M. Lake, Tomer D. Ullman, Joshua B. Tenenbaum, and Samuel J. Gershman. Building machines that learn and think like people. *CoRR*, abs/1604.00289, 2016.
45. Barbara Landau, Linda B Smith, and Susan S Jones. The importance of shape in early lexical learning. *Cognitive Development*, 3(3):299–321, 1988.
46. Oriol Vinyals, Charles Blundell, Tim Lillicrap, Daan Wierstra, et al. Matching networks for one shot learning. In *Advances in Neural Information Processing Systems*, pages 3630–3638, 2016.
47. Russell Stewart and Stefano Ermon. Label-free supervision of neural networks with physics and domain knowledge. In *AAAI*, pages 2576–2582, 2017.
48. Zhiting Hu, Xuezhe Ma, Zhengzhong Liu, Eduard Hovy, and Eric Xing. Harnessing deep neural networks with logic rules. *arXiv preprint arXiv:1603.06318*, 2016.
49. Kuzman Ganchev, Jennifer Gillenwater, Ben Taskar, et al. Posterior regularization for structured latent variable models. *Journal of Machine Learning Research*, 11(Jul):2001–2049, 2010.
50. Zhiting Hu, Zichao Yang, Ruslan Salakhutdinov, and Eric Xing. Deep neural networks with massive learned knowledge. In *Conference on Empirical Methods in Natural Language Processing*, pages 1670–1679, 2016.
51. John J Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the national academy of sciences*, 79(8):2554–2558, 1982.
52. Dmitry Krotov and John J. Hopfield. Dense associative memory for pattern recognition. *CoRR*, abs/1606.01164, 2016.
53. Junhyuk Oh, Valliappa Chockalingam, Satinder Singh, and Honglak Lee. Control of memory, active perception, and action in minecraft. *arXiv preprint arXiv:1605.09128*, 2016.
54. Emilio Parisotto and Ruslan Salakhutdinov. Neural map: Structured memory for deep reinforcement learning. *arXiv preprint arXiv:1702.08360*, 2017.
55. Haiyan Yin and Sinno Jialin Pan. Knowledge transfer for deep reinforcement learning with hierarchical experience replay. In *AAAI*, pages 1640–1646, 2017.