Deep Learning Technologies and Their Applications in the Advancement of Artificial Intelligence

Jianzhi Wang
Xiangzhang Avenue, Shushan District, Hefei City, Anhui Province, China
jianzhi6655@outlook.com

ABSTRACT

The study briefly summarizes and analyzes the use of deep learning technology in the development of artificial intelligence. In response to the actual needs of AI's evolution from perceptual intelligence to cognitive and decision-making intelligence, the algorithm of multi-layer representation learning model for deep learning has changed from the initial convolutional recurrent networks to residual structures and generative adversarial networks. In order to carry out large-scale training in computing, it is particularly noteworthy that real-time object detection, semantic segmentation, semantic understanding of natural language and machine translation in computer vision are being analyzed, and new usage methods in important fields such as judging complex feedback learning and generative model are being analyzed. The research shows that deep learning uses end-to-end feature abstraction and data-driven mechanism, which greatly improves the scope and practical application advantages of AI system, effectively promoting the implementation of strategic applications such as autonomous driving, precision medicine, intelligent interaction and autonomous decision-making. However, there are still some unsolved problems such as model interpretability, energy consumption optimization, data bias, etc. Future research needs to be more deeply integrated with symbolic reasoning, causal learning, edge computing and other technologies to establish a safer and more efficient general-purpose AI system. This study provides an organized reference framework for the theoretical discovery and industrial practice of deep learning-enabled AI.

Keywords: Deep Learning; Artificial Intelligence Development; Computer Vision Applications; Natural Language Processing; Reinforcement Learning Decision Making

1. INTRODUCTION

1.1 Research Background

Since the emergence of artificial intelligence, it has evolved in various ways, such as symbolism and connectionism. After AI winter in 1980s and 1990s, deep learning, an important way forward for ism, was full of energy after Mr. Hinton introduced the fast learning algorithm of deep belief networks in 2006 [1]. This method used complementary priors. Make the deep belife network practice layer by layer, which solves the gradient vanishing problem well and opens up a way for the practical use of deep networks. Due to the explosive increase of data on the Internet, hardware acceleration technologies such as graphics processing units have also become solid, so there will be others where deep learning automatically remembers the characteristics of multi-stages from the original data.

2012, the AlexNet convolutional neural network designed by Krizhevsky et al. significantly outperformed traditional methods in the ImageNet image classification challenge [2], reducing the top-5 error rate to 15.3%, which was regarded as a landmark event in the era of deep learning. Subsequently, LeCun, Bengio and Hinton published an article in Nature that systematically described the mechanism by which deep learning achieves data abstraction representation through multi-layer nonlinear processing [3]. Meanwhile, the Transformer architecture based on the attention mechanism proposed by Vaswani et al. completely changed the way sequence data is modeled [4], directly giving rise to modern artificial intelligence systems represented by large language models. These technological breakthroughs together constitute the core driving force for the leap of artificial intelligence from perceptual intelligence to cognitive and decision-making intelligence.

1.2 Significance of the Research

Research on the application of deep learning technology in the development of artificial intelligence has important theoretical value and practical significance. At the theoretical level, deep learning breaks through the dependence of traditional machine learning on artificial feature design and significantly improves the ability of models to process complex high-dimensional data through end-to-end multi-layer representation learning, injecting new vitality into machine learning theory [3]. At the same time, the connection between the interpretability research of deep learning models and the modeling of biological neural systems also provides a new research perspective for the intersection of cognitive science and artificial intelligence [1].

At the application level, deep learning has triggered a performance revolution in core branches of artificial intelligence such as computer vision, natural language processing, and reinforcement learning. For example, the success of convolutional networks has enabled tasks such as real-time object detection and medical image analysis to reach a practical level [2], while the Transformer architecture has supported the rapid development of generative artificial intelligence and expanded the application boundaries of artificial intelligence in scenarios such as content creation and human-computer interaction [4]. In addition, the application of deep learning technology in industry has brought great economic and social benefits. It has helped the development of strategic areas such as intelligent manufacturing and autonomous driving. In addition, issues such as energy consumption, data privacy and fairness of algorithm are clarified, which provides an opportunity for academia and industry to find a responsible development path of artificial intelligence. This study can organize and review the changes and application effects of related technologies, and provide valuable reference for future academic research and technical judgment.

2. THE THEORETICAL FOUNDATION AND EVOLUTION OF DEEP LEARNING TECHNOLOGY

2.1 Conceptual connotation and basic principles of deep learning

“Deep learning” is a machine learning method, which automatically learns the representation of multi-stage abstract features from the original data by creating a multi-layer nonlinear network. Its core significance is to use hierarchical structure to gradually summarize low-level features into a higher-level semantic concept, thus effectively modeling complex data [5]. Compared with the simple model, deep learning pays more attention to end-to-end learning, and the work of making features does not need manual work. Therefore, it has obvious advantages in processing high-dimensional unstructured data such as images, voices and texts. The basic principles include calculating network output through forward propagation, determining loss function and quantifying prediction error, and optimizing network parameters based on gradient descent in error back propagation [5]. By using nonlinear activation function, the model can further enhance its expressive strength, and the deep structure can capture the complex nonlinear relationship in the data.

In addition, deep learning is built on the basis of representation learning theory. General approximation theorem proves that a neural network large enough can approximate any connected function. Deep structure has advantages in parameter usage efficiency and generalization ability. The combination of these rules is the core that makes deep learning different from traditional machine learning. The design of the model, the stability of training and the popularization and application in the future provide solid theoretical support.

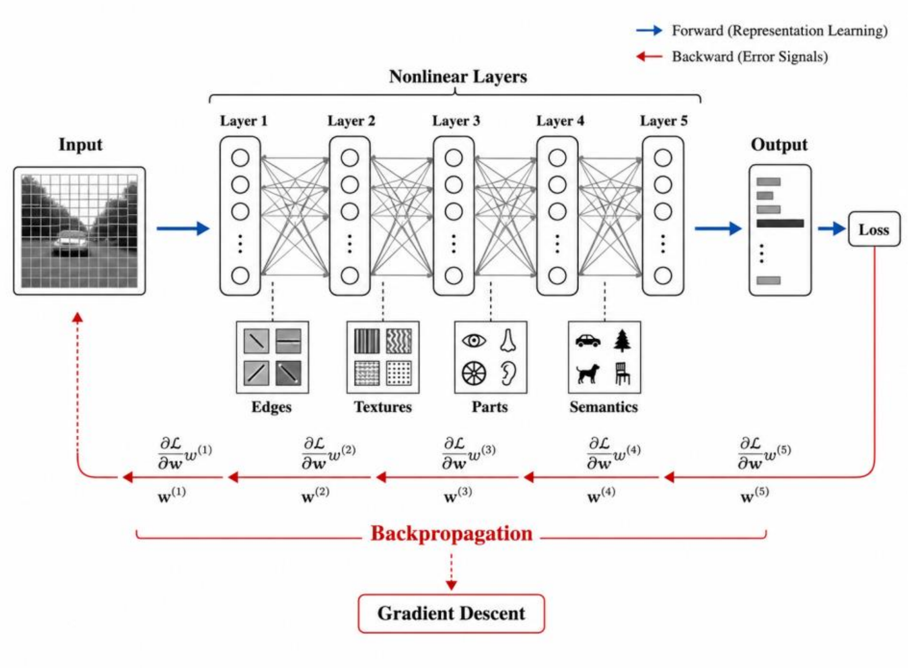

Figure 1. Deep Learning Multilayer Representation Learning and Backpropagation Principle.

As shown in Figure 1, deep learning conducts hierarchical representation learning of data through multi-layer nonlinear transformation. The input data is abstracted layer by layer from low-level edge and pattern features to high-level semantic concept. Backpropagation algorithm uses chain rule to send the output loss gradient back to the weight of each layer. Then it is combined with gradient descent to achieve end-to-end parameter optimization.

2.2 Evolution of Deep Learning Models and Algorithms

The evolution of models and algorithms of deep learning began with the resurrection of neural network. In 1986, Rumelhart and others put forward the backpropagation algorithm, which laid the foundation for effectively training multi-layer networks [5]. Subsequently, in 1998, LeCun et al. successfully applied the LeNet convolutional neural network to document recognition tasks, marking the practical application of convolutional structures in image processing [6]. In the 21st century, Hochreiter and Schmidhuber proposed the Long Short-Term Memory (LSTM) network in 1997, which effectively solved the long-range dependency problem of recurrent neural networks and promoted the progress of sequence modeling [7].

In the mid-2010s, the improvement of model depth and performance became the mainstream. In 2014, Generative Adversarial Networks (GANs) proposed by Goodfellow et al. pioneered a new paradigm of generative learning [9]. In 2016, Residual Networks (ResNet) proposed by He et al. solved the degradation problem in deep network training through residual connections, enabling the network depth to reach hundreds of layers and achieving breakthroughs in tasks such as ImageNet [8]. These algorithmic evolutions have expanded from perception tasks to the fields of generation and decision-making, continuously driving the performance leap of deep learning in various branches of artificial intelligence.

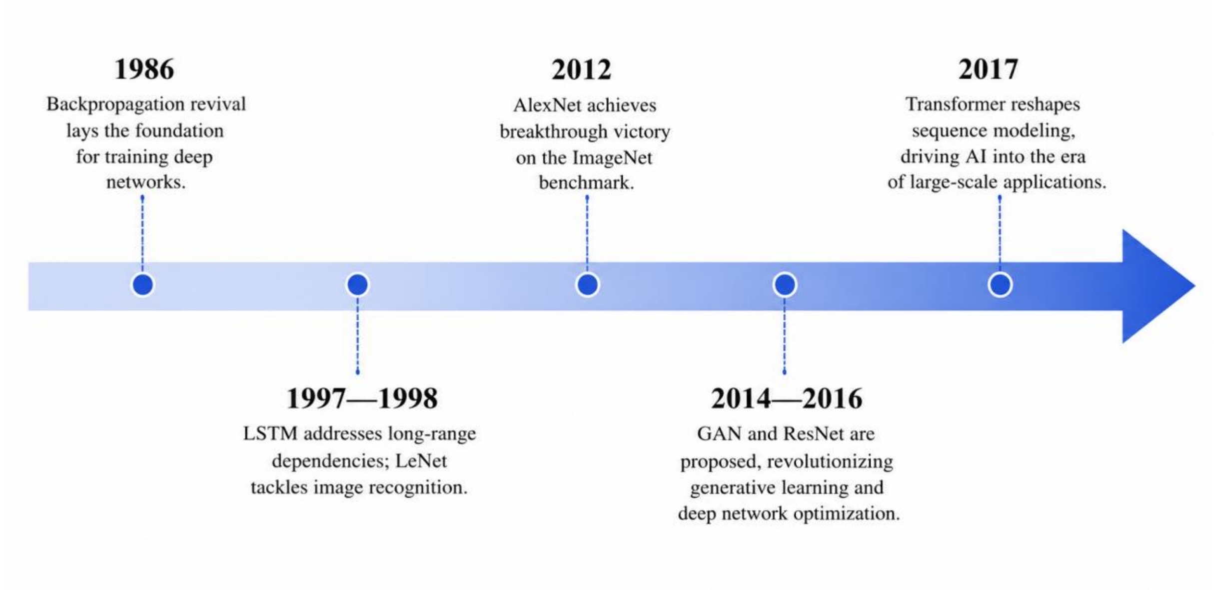

Figure 2. Timeline of the evolution of major deep learning models and algorithms.

2.3 Development of Deep Learning Supporting Technologies and Platforms

The coordinated development of computing hardware, software frameworks and data infrastructure is very important for the successful operation of deep learning. The parallel computing capabilities of graphics processing units (GPUs) is a force that can be calculated in parallel, which makes the calculation of large matrix and the update of gradient very fast. This is an important hardware force for training deep network. In terms of software, the system named TensorFlow introduced by Abadi et al. in 2016 uses data flow graphs to represent the calculation process. Apps that support distributed training and run on different device are more likely to increase deep learning used in factories [10]. The framework integrates automatic differentiation, model optimization and mobile tools, and has become a standard platform widely used by universities and enterprises.

In addition, by combining the distributed training framework with cloud computing resources, we further reduced the wall needed to start deep learning experiments. As a result, the training of a machine has evolved to the training of a large cluster. The establishment of these supporting technologies and platforms not only ensures the rapid and stable model training, but also lays a solid foundation for the rapid update and practical application of deep learning.

3. APPLICATIONS AND IMPACTS OF DEEP LEARNING IN CORE AREAS OF ARTIFICIAL INTELLIGENCE

3.1 Computer Vision and Image Intelligence Applications

Computer vision is the field where deep learning has made the greatest progress. The center is a convolutional structure, which can automatically extract image features and make a high degree of semantic understanding. YOLO model proposed by Redmon completes the target location and classification simultaneously through a forward propagation network. Therefore, the efficiency and accuracy of real-time detection have been greatly improved, and it has become a practical solution for autonomous driving perception, video surveillance, robot navigation and other image intelligent systems [11]. This complete detection framework avoids the lag problem caused by traditional multi-stage flow, which shows that deep learning has matured in the field of understanding moving visual scenes. The fully convolutional network (FCN) proposed by Long transforms the topic of semantic segmentation into end-to-end pixel-level prediction. In this way, image intelligence can expand the scope from target recognition to detailed scene analysis [12]. This directly promotes the development of high-value applications such as medical image aided diagnosis and industrial defect detection.

The progress of these technologies has had a great impact on the leap of AI from perception layer to judgment layer. The image intelligent system working in Deep learning can not only do as well as experts in the analysis of static images, but also work well in the real-time working environment. This provides key technical support for key areas such as intelligent transportation, smart security and precision medicine, and accelerates the overall evolution of AI from special tools to multi-purpose intelligence.

3.2 Applications of Natural Language Processing and Semantic Understanding

In the field of natural language processing, due to deep learning, the machine has changed from a simple statistical model to a deep semantic model, and its ability to understand human language intentions has become very good. If the attention mechanism proposed by Bahdanau et al. is combined with the sequence-to-sequence framework, the relevant parts of the original word and the target word will be matched at the same time [13], thus losing meaning in the translation of long sentences, which is the basis for creating cross-language information retrieval, intelligent customer service and multilingual content. This method goes beyond the fixed context limitation of the old recurrent networks, which enables the model to focus on the important parts of semantics and greatly improves the accuracy and fluency of natural language understanding.

Using deep learning in semantic understanding is better than cognitive capabilities of artificial intelligence. From word vector representation to context-aware encoding, models can find difficult semantic relationships in the whole word, sentence and text. Now, you can perform tasks such as question-answering systems, sentiment analysis and Knowledge Graph Construction. These advances not only improve the communication between people and computers, but also help people from semantic when using artificial intelligence in professional fields such as law, medicine and education. The progress from perceptual intelligence to cognitive intelligence will also be accelerated.

3.3 Application of Reinforcement Learning and Generative Models in Decision Intelligence

The combination of reinforcement learning and generative models opens up a new path for AI's decision intelligence. In this way, the system can learn the best practices through trial and error and simulation in an uncertain environment. The Deep Q-Network (DQN) proposed by Mnih, together with deep convolutional networks and Q-learning, has become as strong as human experts in Atari games. This proves that deep reinforcement learning can be used in high-dimensional state judgment [14] This framework can be stably learned through experience replay and target network, so it can be greatly popularized for complex judgment work such as robot control, resource scheduling, game AI, etc. The AlphaGo system developed by Silver further combines deep neural networks with Monte Carlo tree search and wins the world champion in the field of Go. This shows that the judgment system run by deep learning may be stronger than human beings in very complicated games [15].

Generative model, such as Generative Adversarial Networks (GANS), use adversarial training to make real data. In order to decision intelligence, it will give us a lot of training environment and improvement methods. It is found that these technologies can be flexibly used in autonomous driving road planning, factory work optimization, financial risk control and so on. Artificial intelligence is developing from passive execution of commands to active decision-making. In addition, in order to cooperate with the easy-to-understand decision and safety research, an important road is also given.

4. CONCLUSION

This study sorts out and introduces the core approaches and application results of deep learning technology in the development of artificial intelligence. Starting with the basic rules of multi-layer expression learning and backpropagation, this paper analyzes the theoretical significance of deep learning in breaking through the limitations of traditional models in end-to-end feature extraction. The evolution from the early convolutional and recurrent networks to the residual structures and generative adversarial networks, as well as the computational frameworks and distributed platforms used for large-scale training. On this basis, the breakthrough application of deep learning is emphatically discussed. These advances in computer vision, such as real-time object detection, semantic segmentation, natural language semantic understanding and machine translation, and complex decision-making driven by reinforcement learning and generative models, jointly promote the development of intelligent intelligence from a special perceptual system to general intelligence with cognitive and independent decision-making ability. It greatly improves the theoretical depth, engineering usability and industrial value of related technologies.

Looking ahead, there is still a problem in deep learning. For example, it is difficult to explain the idea of the model, make good use of computer energy, data deviation and so on. It needs to be more deeply integrated with technologies such as symbolic reasoning, causal learning and edge intelligence to establish a more efficient, safe and reliable AI system. This study will provide a systematic reference framework for future theoretical discoveries and technical work. This will help accelerate the use and sustainable development of deep learning in key areas such as intelligent manufacturing, precision medicine and autonomous systems.

REFERENCES

[1] Hinton, GE, Osindero, S., Teh, Y.-W.: A fast learning algorithm for deep belief nets. Neural Computation 18(7), 1527–1554 (2006). https://doi.org/10.1162/neco.2006.18.7.1527.

[2] Krizhevsky, A., Sutskever, I., Hinton, GE: ImageNet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems 25, pp. 1097–1105 (2012).

[3] LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. Nature 521(7553), 436–444 (2015). https://doi.org/10.1038/nature14539.

[4] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, AN, Kaiser, Ł., Polosukhin, I.: Attention Is All You Need. In: Advances in Neural Information Processing Systems 30, pp. 5998–6008 (2017).

[5] Rumelhart, DE, Hinton, GE, Williams, RJ: Learning representations by back-propagating errors. Nature 323(6088), 533–536 (1986). https://doi.org/10.1038/323533a0.

[6] LeCun, Y., Bottou, L., Bengio, Y., Haffner, P.: Gradient-based learning applied to document recognition. Proceedings of the IEEE 86(11), 2278–2324 (1998). https://doi.org/10.1109/5.726791.

[7] Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural Computation 9(8), 1735–1780 (1997). https://doi.org/10.1162/neco.1997.9.8.1735.

[8] He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770–778 (2016). https://doi.org/10.1109/CVPR.2016.90.

[9] Goodfellow, IJ, Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y.: Generative adversarial nets. In: Advances in Neural Information Processing Systems 27, pp. 2672–2680 (2014).

[10] Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., Kudlur , M., Levenberg, J., Monga, R., Moore, S., Murray, DG, Steiner, B., Tucker, P., Vasudevan, V., Warden, P., Wicke, M., Yu, Y., Zheng, X.: TensorFlow: A system for large-scale machine learning. In: 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16), pp. 265–283 (2016).

[11] Redmon, J., Divvala, S., Girshick, R., Farhadi, A.: You Only Look Once: Unified, real-time object detection. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 779–788 (2016). https://doi.org/10.1109/CVPR.2016.91.

[12] Long, J., Shelhamer, E., Darrell, T.: Fully convolutional networks for semantic segmentation. In: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3431–3440 (2015). https://doi.org/10.1109/CVPR.2015.7298965.

[13] Bahdanau, D., Cho, K., Bengio, Y.: Neural machine translation by jointly learning to align and translate. In: 3rd International Conference on Learning Representations (ICLR 2015), Conference Track Proceedings (2015).

[14] Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, AA, Veness, J., Bellemare, MG, Graves, A., Riedmiller, M., Fidjeland, AK, Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S., Hassabis, D.: Human-level control through deep reinforcement learning. Nature 518(7540), 529–533 (2015). https://doi.org/10.1038/nature14236.

[15] Silver, D., Huang, A., Maddison, CJ, Guez, A., Sifre, L., van den Driessche, G., Schrittwieser, J., Antonoglou , I., Panneershelvam , V., Lanctot, M., Dieleman, S., Grewe, D., Nham, J., Kalchbrenner, N., Sutskever , I., Lillicrap, T., Leach, M., Kavukcuoglu , K., Graepel, T., Hassabis, D.: Mastering the game of Go with deep neural networks and tree search. Nature 529(7587), 484–489 (2016). https://doi.org/10.1038/nature16961.