Proceedings of CONF-CDS 2025 Symposium: Data Visualization Methods for Evaluation
DOI: 10.54254/2755-2721/2025.PO25828

Deep Learning Techniques and Applications

Tangyi Huang

Computer Science Department, Durham University, Durham, The United Kingdom
xnds74@durham.ac.uk

Abstract. Deep learning has brought about a significant transformation in artificial intelligence, facilitating progress in various domains, including computer vision, natural language processing, and biomedical imaging. This paper undertakes a review of the fundamental breakthroughs, approaches, and uses of deep learning. The emphasis is placed on key architectures, namely convolutional neural networks (CNNs), recurrent neural networks (RNNs), transformers, and generative models. By conducting an analysis of the existing body of literature and real - world case studies, this research delves into the advantages and drawbacks of these models. It underscores the far - reaching and revolutionary influence of these models while also tackling issues such as scalability, interpretability, and the high computational requirements. Key findings include the potential of emerging techniques like diffusion models and hybrid generative approaches to overcome current limitations. However, the study acknowledges limitations in its reliance on secondary sources and the scope of reviewed literature. Subsequent research stands to gain significantly from the incorporation of empirical investigations and more extensive data sets. This approach would offer a more holistic comprehension of the prospects and effects of deep learning. The objective of this paper is to offer guidance to researchers and professionals in the progress of both the theoretical and practical uses of deep learning.

Keywords: Deep learning, convolutional neural networks(CNNs), recurrent neural networks (RNNs), Artificial Neural Networks (ANNs)

1. Introduction

Deep learning, a specialized area within machine learning, has revolutionized artificial intelligence. It allows systems to directly learn intricate patterns from unprocessed data. Based on neural network framework and hierarchical feature description, deep learning models have achieved first-class performance in many fields, such as computer vision, natural language processing, and biomedical applications [1-2].

AlexNet's breakthrough in the ImageNet competition marked a key milestone, as mentioned in the literature [3], and demonstrated the full potential of convolutional neural networks, and since then innovations such as residual networks, transformers, and generative adversarial networks have been developed. Both have broadened the horizons of deep learning. These advances have been driven by increasing computing power, the availability of large-scale data sets, and innovative algorithm design [1].

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2025.PO25828

Although deep learning has achieved widespread success, it still faces some challenges, including relatively high computing requirements, data dependence, and limited interpretability [1-2]. These limitations are addressed while improving the robustness and efficiency of deep learning systems. It's still an ongoing pursuit. This paper makes a comprehensive review of the basic concepts, model categories and applications of deep learning. It also discusses the limiting factors and challenges that hinder the wide application of deep learning technology, and also mentions the promising direction of future research and development. By clarifying the current research direction, This paper provides valuable insights for future researchers, as well as practical perspectives for professionals in related fields, which can drive the development and implementation of innovative deep learning solutions.

2. Fundamental concepts

2.1. Components of neural networks

Artificial neural networks are the foundation of deep learning. They mimic the structure and operation of biological neural networks. Artificial neural networks are made up of multiple layers of interconnected nodes, also known as neurons, each of which computes a weighted sum of its inputs and then applies a nonlinear activation function. Like rectifiers or S-type functions [1]. These layers can be divided into three main types, the input layer is responsible for receiving the raw data, the hidden layer is responsible for extracting features from the input data, and the output layer generates predictions based on the information obtained after processing.

As the depth of the network, or the number of layers, increases, the ability of the network to learn layers and abstract representations becomes stronger, which is critical when dealing with complex tasks such as image classification and language modeling [1,3].

2.2. Principles of backpropagation

Backpropagation is the basic algorithm for training neural networks, which allows the neural network to slowly reduce the prediction error during the iteration process. The purpose of this algorithm is to make the loss propagate in the opposite direction in the network, which refers to the difference between the predicted value and the actual value, and it uses the chain rule in calculus to achieve this [1]. At each layer of the network, the gradient of the loss function relative to the weight will be calculated, and then the gradient descent method will be used to adjust the weights corresponding to these gradients [1,4]. The whole process consists of three key steps: the first step is forward transmission, during which the network will calculate its own predicted results; the second step is loss calculation. This step is used to measure the size of the gap between the predicted value and the actual value. In this step, the gradient is calculated and the weights are updated according to the calculated results. Backpropagation can ensure that the network gradually reduces errors by adjusting its own parameters. Although this algorithm is very effective, it may encounter challenges such as disappearing gradient in deep networks. These challenges have been mitigated by innovative approaches such as residual connectivity and activation functions such as ReLU [5].

3. Deep learning model categories

3.1. Convolutional Neural Networks (CNNs)

Convolutional neural network is a kind of neural network with a unique feature, which is specially designed to process data arranged into a grid structure, just like the common data type of image,

DOI: 10.54254/2755-2721/2025.PO25828

originally proposed by LeCun and his colleagues, cnn will use convolutional layer to extract hierarchical features from the input data [1]. These convolutional layers apply learnable filters to identify patterns such as edges, textures, and objects. CNNS are particularly well suited for activities related to image classification, object discovery, and segmentation [3]. There have been some notable advances, such as AlexNet, which demonstrates the full capabilities of deep CNNS for handling large data sets. In the ImageNet competition achieved very outstanding results [3]. Later advances, such as ResNet, which introduced residual connections to train deeper networks in a way that mitigated the disappearing gradient problem [5], cemented cnn's status as a cornerstone of computer vision.

3.2. Recurrent Neural Networks (RNNs) and transformers

The point of this article is that recurrent neural networks are specially designed to deal with sequential data. In this case, the result of each stage depends on the input of the previous stage. RNN uses hidden layers to preserve the information content between sequences, which makes it suitable for applications such as speech recognition and machine translation [1]. Theoretically, it can model arbitrary sequence length, but in the actual implementation process, it will appear gradient dissipation when dealing with long-term dependence, long short-term memory and gated cycle unit and other technologies have made progress, these technologies rely on integrated gating mechanism to solve the problem of gradient dissipation [1].

As mentioned in this article, transformer, introduced by Vaswani et al., replaces recursion with a attention mechanism, which enables parallel computation and makes training more efficient [2] , and a self-attention mechanism allows the model to focus on the relevant parts of the input sequence, making it easy to capture long-term dependencies. Transformers have become an excellent architecture for natural language processing, and are the basis for models such as BERT and GPT, which are much better at scalability and efficiency than RNNS for large-scale pre-training tasks [2].

3.3. Generative models (VAEs, GANs, and normalizing flows)

As mentioned in this paper, the generation model contains a variety of methods to learn the distribution of the underlying data, which are used to synthesise new samples, and the variational autoencoder does some fine-tuning of the evidence lower limit in order to achieve a balance between the reconstruction accuracy and the underlying regularization [2]. Despite their theoretical grounding in Bayesian principles, VAEs often exhibit blurry outputs due to their reliance on Gaussian assumptions,which constrains their application in tasks requiring high-resolution outputs [6].

Generative Adversarial Networks (GANs) engage a generator and a discriminator in a minimax competition. In this setup, the generator produces data specimens to deceive the discriminator. Meanwhile, the discriminator differentiates between authentic and synthetic data [7]. GANs excel in producing realistic images but are prone to instability during training, with issues like mode collapse. Despite the advantages of GANs in generating high-fidelity samples [6], thet have certain defects including susceptibility to instability and the computational cost of fine-tuning hyperparameters.

Normalizing Flows enable exact likelihood computation by transforming simple probability distributions through a series of invertible functions [4]. They provide a powerful framework for density estimation and sample generation but require significant computational resources for high-

DOI: 10.54254/2755-2721/2025.PO25828

dimensional data. Recent advancements, such as Glow, have demonstrated their capability for high-quality image generation [4].

4. Applications of deep learning

4.1. Computer vision

Deep learning has made great progress in computer vision by promoting automatic analysis of visual data. Convolutional neural networks, also known as CNNS, have become a key part of many tasks, such as image classification, object recognition and semantic division [3]. The emergence of AlexNet shows the ability of deep CNNS to deal with large-scale data sets. It achieved innovative results in the ImageNet Competition [3], and subsequent innovations, such as ResNet, solved challenges such as disappearing gradients, enabled deeper network training, and improved image recognition [5]. Advanced frameworks such as Faster R-CNN integrate regional proposal networks to achieve real-time target detection and have an impact on areas such as autonomous driving and surveillance [5,7].

In addition to these advances mentioned above, visualization techniques have played a key role in understanding and optimizing convolutional neural networks. Zeiler and Fergus present a method for visualizing the middle layers of convolutional neural networks, which shows how different layers extract layered features such as edges and textures. And it has also helped researchers make improvements to network design [8]. These understandings of convolutional neural networks have played a key role in creating more efficient and understandable deep learning models, thus pushing the limits of computer vision.

4.2. Natural language processing

In the field of natural language processing, deep learning has revolutionized the way machines understand and produce human language. Recurrent neural networks, long and short term memory networks, play a key role in tasks such as machine translation and speech recognition, and they rely on modeling sequential relationships to achieve this role [1]. However, the emergence of Transformer architecture has given a big change, which uses a self-attention mechanism to enable parallel processing and effectively capture remote dependencies [2,7], and Transformer is now the basis for pre-trained language models such as BERT and GPT. These models have achieved top-notch results for many NLP tasks, including text enrichment, question answering, emotion assessment, etc [2,7]. and their success highlights the profound impact of deep learning on natural language understanding and generation.

4.3. Biomedical image processing

Deep learning has revolutionized biomedical image processing, playing a key role in helping to diagnose diseases, develop treatment plans, and promote medical research. The U-Net framework is designed specifically for biomedical image segmentation, integrating a downlink sampling path that extracts contextual information. At the same time, an upsampling path is also integrated, which can accurately locate the target [9]. In addition to model segmentation, generative adversarial networks are also used to solve the problem of shortage of labeled medical data, which is achieved by creating synthetic images and then adding these synthetic images to the training data set. This augmentation process can improve the robustness and generalization of the model [7]. Diffusion models have improved medical imaging by reconstructing high-quality images from noisy inputs, enabling

DOI: 10.54254/2755-2721/2025.PO25828

improved MRI and CT scan results [4], and these advances have provided healthcare professionals with more accurate and effective diagnostic tools, marking a critical step toward personalized medicine.

5. Limitations and challenges

5.1. Data and computational resource requirements

Deep learning models need large data sets and extensive computing resources to be effective, which can pose some obstacles to research and model deployment. Training on state-of-the-art architectures like Transformers or Gans generally requires access to specialized hardware like Gpus or Tpus. At the same time, a large amount of labeling data is required [1-2]. However, in areas such as healthcare, it is challenging to obtain annotated data sets, which limits model generalization and performance [9], and the carbon footprint issue associated with training large models has raised concerns about sustainability. To address these challenges, advances in efficient training algorithms, self-supervised learning, and low-lens learning are necessary prerequisites to reduce reliance on data and computing power [1].

5.2. Model interpretability and controllability

One of the most prominent shortcomings of deep learning is its lack of interpretability, which is because deep learning models are mostly treated as "black boxes," and this problem is particularly difficult to solve in areas such as health care, finance, and autonomous vehicles, where, Understanding the decision-making mechanism is critical to building trust in the model and ensuring accountability is in place [3]. While techniques such as maximality graphs and attention mechanisms can provide interpretability to some extent [2], they cannot fully explain the behavior of the model, and controlling the output of generative models such as Gans is still a challenge, because these models may produce undesirable or biased results [7]. Developing models that are inherently interpretable, and incorporating interpretability into the whole process of deep learning, is critical to addressing these issues [9].

5.3. Training stability and optimization challenges

What this paper will say is that the training process of deep learning models, generating adversarial networks and recurrent neural networks, generally lacks stability, which requires very careful adjustment of hyperparameters. Take gan for example, it is prone to modal collapse, in which case the output generated by the generator has limited variation. In the training process, oscillations occur most of the time, which will make the model unable to reach the convergence state [7]. The same situation also happens to rnn, which also encounters difficulties in learning long-term dependencies because the gradient disappears, even though LSTMs has made some progress [1]. Large-scale models like Transformers bring more challenges, including high memory requirements and low computational efficiency [2]. To solve these problems, it is necessary to improve optimization technology, adopt adaptive learning rate plan and regularization strategy, so as to improve the stability and efficiency of training [4].

DOI: 10.54254/2755-2721/2025.PO25828

6. Future directions

6.1. Emerging models and algorithms

The evolution of deep learning is marked by continuous innovation in model architectures and algorithms, and recent breakthroughs in techniques like diffusion models have demonstrated their ability to produce best-in-class data by representing the noise cancellation process as a step-by-step operation. For details, please refer to relevant literature [4]. These models keep the training process stable and achieve competitive results in image synthesis tasks. Similarly, hybrid generation models combine the advantages of gan, VAEs, and normalized flow to address the shortcomings of individual models, such as instability in GANs or fuzzy output in VAEs. Relevant content can be found in literature [7,10]. Future research can focus on improving the scalability, efficiency, and elasticity of these models so that they can be more widely used in a wide range of different fields.

6.2. Cross-domain integration and efficient deep learning

The future direction of deep learning is to integrate across disciplines, such as combining biological knowledge and machine learning to apply it to drug discovery, or to improve real-time decision making in autonomous systems [9], and cross-domain applications rely on transfer learning and domain adaptation methods to promote different data sets. As the demand for computing resources continues to increase, deep learning needs to prioritize efficiency improvements, such as model pruning, quantification, and knowledge transfer, to reduce the complexity of models without affecting their performance. The rise of green artificial intelligence can reduce energy consumption in training and reasoning. Promote sustainable practices [1,4]. These developments highlight the need for deep learning to evolve, only in terms of capabilities, but also in terms of accessibility and environmental impact.

7. Conclusion

Deep learning has made great strides in recent years, providing powerful tools for solving complex problems in different fields. There are some noteworthy developments, such as convolutional neural networks for computer vision, recurrent neural networks and deformers for natural language processing, and generative models such as generative adversarial networks. Has sparked a revolution in fields as diverse as health care and entertainment. These achievements have not only increased the capabilities of AI systems, but also expanded their capabilities, driving deeper and more detailed understanding of data.

While deep learning models have made some impressive achievements, as discussed in detail in this article, they still face several critical challenges, such as the need for large amounts of labeled data, high computational requirements, and limited interpretability of models. These continue to be prominent obstacles. In addition, the problems of stability and efficiency in training large-scale models are still areas of ongoing research.

Although this review provides an overview of the key advances in deep learning and the challenges it faces, it is not without limitations. First, the review does not include empirical research or experimental verification, but relies on secondary sources to explain it. Second, the number of references cited is relatively small. This may exclude other works that might provide more insight.

Looking ahead, the future of deep learning will likely be shaped by ongoing innovations in model architecture, optimization techniques, and cross-domain integration. Emerging models like diffusion

DOI: 10.54254/2755-2721/2025.PO25828

models and hybrid generative models show great promise, while efforts to reduce computational costs and enhance sustainability will play a critical role in making deep learning more accessible. With future studies adressing these challenges, deep learning is poised to drive further advancements in AI, ultimately transforming industries and improving lives in profound ways.

References

[1] LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521: 436–444
[2] Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser Ł, Polosukhin I (2017) Attention is all you need. Adv Neural Inf Process Syst 30:
[3] Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. Adv Neural Inf Process Syst 25:
[4] Ho J, Jain A, Abbeel P (2020) Denoising diffusion probabilistic models. Adv Neural Inf Process Syst 33: 6840–6851
[5] He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp 770–778
[6] Bond-Taylor S, Leach A, Long Y, Willcocks CG (2021) Deep generative modelling: A comparative review of vaes, gans, normalizing flows, energy-based and autoregressive models. IEEE Trans Pattern Anal Mach Intell 44: 7327–7347
[7] Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y (2014) Generative adversarial nets. Adv Neural Inf Process Syst 27:
[8] Zeiler MD, Fergus R (2014) Visualizing and understanding convolutional networks. In: Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part I 13. pp 818–833
[9] Ronneberger O, Fischer P, Brox T (2015) U-net: Convolutional networks for biomedical image segmentation. In: Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18. pp 234–241
[10] Kingma DP, Dhariwal P (2018) Glow: Generative flow with invertible 1x1 convolutions. Adv Neural Inf Process Syst 31