A Comparative Overview of Artificial Neural Network Architectures: Theoretical Foundations and Applications

Brandon Marquez Salazar

***Abstract*—Artificial Neural Network (ANN) architectures have rapidly diversified, each designed to address specific computational challenges and application domains. These architectures range from enhancements of foundational algorithms to specialized designs for particular problem-solving tasks. This overview examines prominent ANN architectures—including Feedforward Neural Networks (FFNs), Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Transformers, and Generative Models—and their applications in fields such as geotechnical engineering, energy forecasting, and computer vision. Based on a meta-analysis of recent literature, this review highlights key architectural innovations, provides performance comparisons, and discusses emerging trends within the field.**

I. Introduction

The field of Artificial Intelligence (AI) was mainly fueled by the development of Artificial Neural Networks (ANNs). Inspired by biological neural research in attempts of understanding the neuronal behaviour. ANNs are computational models capable of learning complex, non-linear patterns from data [7]. Over time, a diverse ecosystem of architectures has emerged, each designed to address specific computational challenges.

From the foundational Feedforward Neural Networks (FFNs) [9] to the specialized designs of Convolutional Neural Networks (CNNs) [4] for spatial data and Recurrent Neural Networks (RNNs) [5] for time series, the evolution of ANNs has been driven by the need of reaching enhanced performance, better efficiency and extended applicability. More recently, the Transformer architecture [10] has revolutionized natural language processing and is now challenging CNNs in computer vision [2]. Simultaneously, Generative Models like GANs [3] and VAEs [6] have opened new frontiers in creating novel, realistic data.

This review provides a structured overview of these key ANN architectures, focusing on their theoretical underpinnings, architectural innovations, and performance characteristics. Based on a meta-analysis of recent literature, this paper aims to synthesize the strengths and limitations of each paradigm and highlight emerging trends that are shaping the future of deep learning. The review is structured according to the IMRaD format, covering the core methodologies (architectures), results (performance comparisons), and discussion (trends and conclusions).

II. Architectural Overview

This section delineates the fundamental architectures under review, detailing their core structural principles and theoretical innovations.

A. Feedforward Neural Networks (FFNs)

FFNs, or Multi-Layer Perceptrons (MLPs), represent the simplest class of ANNs. They consist of an input layer, one or more hidden layers, and an output layer, with information flowing strictly forward without cycles. Their training is enabled by the backpropagation algorithm [9], which calculates the gradient of the loss function with respect to each weight. Theoretically, FFNs are universal function approximators but lack internal memory, making them suitable for static pattern recognition but ineffective for sequential data.

B. Convolutional Neural Networks (CNNs)

CNNs are the dominant architecture for processing grid-like data such as images. Their design incorporates three key ideas to exploit spatial locality: (1) *sparse connectivity* (neurons connect only to local regions), (2) *parameter sharing* (using the same weights across spatial locations), and (3) *hierarchical feature learning* [7]. Architectures like ResNet [4] introduced skip connections to mitigate the vanishing gradient problem in very deep networks, while designs like Xception [1] further improved parameter efficiency through depthwise separable convolutions.

subsectionRecurrent Neural Networks (RNNs)

RNNs are designed for sequential data by introducing cyclic connections, allowing them to maintain an internal state or ”memory” of previous inputs. The Long Short-Term Memory (LSTM) unit [5] was a pivotal innovation, introducing gating mechanisms (input, forget, and output gates) to regulate information flow and effectively learn long-range dependencies, overcoming the vanishing gradient problem of vanilla RNNs.

C. Transformer Networks

The Transformer architecture [10] marked a paradigm shift by abandoning recurrence and convolution entirely. Its core mechanism is *self-attention*, which weighs the significance of all elements in a sequence simultaneously when processing any single element. This allows for direct modeling of long-range dependencies and massive parallelization during training. The Vision Transformer (ViT) [2] demonstrated that this architecture could be applied directly to sequences of image patches, achieving state-of-the-art results in classification.

D. Generative Models

This class of models learns the underlying probability distribution of data to generate novel samples.

- **Generative Adversarial Networks (GANs)** [3] frame learning as an adversarial game between a generator (that creates fake data) and a discriminator (that distinguishes real from fake).
- **Variational Autoencoders (VAEs)** [6] take a probabilistic approach, learning a latent representation of the data and generating new data by sampling from this latent space.

III. Performance and Applications

The efficacy of each architecture is contingent on its alignment with the problem domain. This section synthesizes their performance across key applications.

A. Computer Vision

In image classification, CNNs like ResNet [4] have long been the benchmark due to their efficient spatial feature extraction. However, Vision Transformers (ViTs) [2] have demonstrated comparable or superior performance on large datasets, leveraging their ability to model global context. For dense prediction tasks like image segmentation, Fully Convolutional Networks (FCNs) and specifically the U-Net architecture [8]—with its encoder-decoder structure and skip connections—remain the gold standard, effectively combining high-level semantics with low-level spatial precision.

B. Sequential and Temporal Data

RNNs and their LSTM variants [5] have been widely successful in time-series forecasting and natural language processing due to their inherent temporal dynamics. The Transformer [10] has largely superseded them in NLP by offering superior training speed and performance on large datasets, thanks to parallelization and global context awareness. Hybrid models, such as the Meta-ANN for Short-Term Load Forecasting [11], showcase how dynamic architectures refined by meta-learning can outperform static models by adapting to non-stationary patterns.

C. Generative Tasks

GANs [3] are renowned for generating high-fidelity, sharp images, though their training can be unstable and prone to mode collapse. VAEs [6], in contrast, provide a more stable training process and a structured latent space but often produce blurrier samples. The choice between them involves a trade-off between sample quality and training stability.

IV. Discussion

The evolution of ANN architectures reveals a clear trend: from general-purpose function approximators to highly specialized designs, and now towards a new synthesis of these specializations. The comparative analysis indicates that there is no single "best" architecture; rather, the optimal choice is dictated by the inductive biases required for the task—spatial invariance for images, temporal consistency for sequences, and the ability to model data distributions for generation.

A key emerging trend is the *hybridization* of architectural paradigms. The Vision Transformer is a prime example, applying a sequence-modeling architecture to computer vision [2]. Similarly, models like Meta-ANN [11] integrate meta-learning to create dynamic networks that adapt to changing data distributions, pointing towards more robust and generalizable systems.

Future research directions are likely to focus on several areas: (1) improving the computational and energy efficiency of large models like Transformers, (2) enhancing the stability and controllability of generative models, and (3) automating architecture design through Neural Architecture Search (NAS). Furthermore, the development of architectures that can seamlessly learn from multiple data modalities (vision, language, sound) within a single model represents a significant frontier in AI research.

References

[1] François Chollet. “Xception: Deep learning with depthwise separable convolutions”. In: *In Proceedings of the IEEE conference on computer vision and pattern recognition* (2017). An influential paper on efficient CNN architectures, relevant for understanding modern design principles., pp. 1251–1258.

[2] Alexey Dosovitskiy et al. “An image is worth 16x16 words: Transformers for image recognition at scale”. In: *arXiv preprint arXiv:2010.11929* (2020). Introduces the Vision Transformer (ViT), applying the Transformer architecture to image classification.

[3] Ian Goodfellow et al. “Generative adversarial nets”. In: *Advances in neural information processing systems* 27 (2014). The seminal paper introducing Generative Adversarial Networks (GANs).

[4] Kaiming He et al. “Deep residual learning for image recognition”. In: *In Proceedings of the IEEE conference on computer vision and pattern recognition* (2016). Introduces the ResNet architecture with skip connections, enabling the training of very deep CNNs., pp. 770–778.

[5] Geoffrey Hinton et al. “Long short-term memory”. In: *Neural computation* 9.8 (1997). The original paper introducing the LSTM architecture., pp. 1735–1780.

[6] Diederik P Kingma and Max Welling. “Auto-encoding variational bayes”. In: *arXiv preprint arXiv:1312.6114* (2013). The paper introducing the Variational Autoencoder (VAE).

[7] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. “Deep learning”. In: *nature* 521.7553 (2015). A foundational review of deep learning, covering CNNs, RNNs, and the core concepts., pp. 436–444.

[8] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. “U-net: Convolutional networks for biomedical image segmentation”. In: *In Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18* (2015). Presents the U-Net architecture, a cornerstone for image segmentation tasks., pp. 234–241.

[9] David E Rumelhart, Geoffrey E Hinton, and Ronald J Williams. “Learning representations by back-propagating errors”. In: *nature* 323.6088 (1986). The classic paper on the backpropagation algorithm for training Feedforward Neural Networks., pp. 533–536.

[10] Ashish Vaswani et al. “Attention is all you need”. In: *Advances in neural information processing systems* 30 (2017). The seminal paper introducing the Transformer architecture.

[11] Xun Xiao et al. “Meta-ANN – A dynamic artificial neural network refined by meta-learning for Short-Term Load Forecasting”. In: *Energy* 246 (May 2022), p. 123418. issn: 0360-5442. doi: 10.1016/j.energy.2022.123418. url: http://dx.doi.org/10.1016/j.energy.2022.123418.