Advances in Artificial Intelligence and Machine Learning; Review 6 (4) 6122-6155 Received 15-04-2026; Accepted 23-08-2026; Published online 30-08-2026

Deep Learning: A Survey of Recent Advances And Applications in Machine Learning

Nahla Flayyih Hasani n.hassani@uos.edu.iq
Department of science, College of Basic Education,
University of Sumer,
Rifai/Thi-Qar, IRAQ

Mohammed Salih hussein muhammadsalihhussain@gmail.com
Computer center, university of Sumer,
Thi-Qar, IRAQ

Ibrahim Haitham Abdulelah ibr3111990@gmail.com
College of Basic Education, University of Sumer,
Rifai/Thi-Qar, IRAQ

Corresponding Author: Nahla Flayyih Hasani

Abstract

Deep Learning has had a huge impact on a wide variety of computational tasks and it has become one of the major pillars of modern Artificial Intelligence. This paper presents a comprehensive and consistent organizational structure on recent developments in the area of deep learning, including algorithm enhancements and proof-of-concept implementations. The survey was conducted through a methodological literature review of scientific databases such as ACM Digital Library, Scopus, Web of Science, IEEE Xplore, and arXiv using a keyword menu designed to identify the most common core applications, architectures, and cutting-edge research fronts. We synthesized the development of leading popular deep learning libraries linking to models that are well known such as CNN, RNN, to the present day of more advanced and sophisticated models such as transformers, graph neural networks, and diffusion-based generating models. The use of various applications in different domains has been reviewed and their performance, limitations, and interdisciplinary impacts have been detailed. There are many different persistent challenges persistent (such as lack of data, explainability, too much computational power required to run AI models, ethics, and adversarial robustness) that are pointing to key research gaps in the field that still exist. At the conclusion of this study there will be multiple examples of future work in this field (explainable AI, resource-efficient modelling, learning from multiple types of data, collaboration between humans & AI, and creating general-purpose intelligent systems). By reviewing all current literature thoroughly, this study aims to serve as a comprehensive, accurate, up-to-date reference for researchers, practitioners and policy-makers who either study the current state or predict future states of deep learning.

Keywords: Deep learning, Artificial intelligence trends, Machine learning applications, Computer vision, Generative models

Citation: Nahla Flayyih Hasani, et al. Deep Learning: A Survey of Recent Advances And Applications in Machine Learning. **Advances in Artificial Intelligence and Machine Learning**. 2026;6(4):339. https://dx.doi.org/10.54364/AAIML.2026.64339

1. INTRODUCTION

1.1 Background and Evolution of Machine Learning

Over time Machine Learning (ML) has made significant advances from its beginnings with basic statistical classification and classification in general to the more sophisticated ML Techniques that are now used primarily for processing very large amounts of complex data. The traditional approaches to ML have relied heavily on manually engineered features (e.g. SVM's, DT, and Naive Bayes Classifiers) to calculate parameters w which require considerable experience and domain knowledge along with extensive preparation steps to achieve success in applying the techniques that were initially used to construct classifiers. These were useful on low-dimensional or structured problems but were less flexible and have certainly not scaled to accommodate the new marketplace of heterogeneous, unstructured and/or high-dimensional data with which we now have to work. As the volume of digital data, the development of more powerful computational infrastructure and the creation of gradient descent-based learning algorithms have all contributed to the growth of automated representation learning, this has eventually led to the development of Deep Learning which is based on creating hierarchical neurons or synapses that are used to learn and to create multiple levels of representation of information directly from unedited data [1].

1.2 Emergence and Impact of Deep Learning

Last year's advances in deep learning (DL) have established the field as an emergent technology that has the capability to exceed results obtained by conventional machine learning methods. The demonstrated excellence of deep convolutional neural networks applied to large-scale computer vision (a.k.a. visual recognition) problems has confirmed the viability of very deep networks performing at an extremely high level, taking into account multiple neural architecture definition variants, as well as GPU-based training and utilization methods. The introduction of sequence modelling using Recurrent Neural Networks (RNNs), particularly LSTMs [2], and GRUs, expands the amount of language, speech, and temporal data that can be learned from neural networks, resulting in significant improvements in relative rankings of models used to represent these types of data over some time. The creation of the transformer architecture, which has allowed parallel processing on sequential data, has produced a transformational breakthrough in providing significant improvements in the ability of neural networks to model long-term dependencies. More recent developments involving generative models, such as diffusion models and large language models, have enhanced the power of DL to the generation of synthetic images and temporal data, as well as in the areas of chemical and molecular dynamics (simulated – computational) modelling and storage of naturalistic information (e.g., images, voice, etc.) and human-like reasoning. These advancements continue to demonstrate how DL can be utilized as an overall catalyst for advancing innovation in the fields of science, engineering, medicine, and the global economy[3].

1.3 Why Survey Studies Are Essential in Rapidly Evolving Fields

Research into deep learning has seen a tremendous boost in recent times, resulting in an exponentially increasing volume of literature. With so many new models, theories, optimization techniques and applications changing almost daily, researchers are finding it increasingly difficult to keep track of everything. Understanding how to navigate this rapidly changing landscape will be achieved through research, by identifying emerging themes and collecting diverse pieces of evidence into large, coherent themes from different disciplines. This study aims to identify existing linkages between novel and foundational conceptual models. The body of work also highlights the impact of gaps in the literature, including the lack of clear methodological approaches. Due to the speed of such rapid change in technology such as deep learning, conducting broad based studies of this topic is essential for the research community to have a knowledge basis for advancing research work [4].

1.4 In Latest AI, Deep Learning Like a Transformative Model

In artificial intelligence (AI), deep learning has emerged as a vital area of expertise that allows the development of sophisticated new algorithms for encoding complex semantic representations of both spatial and temporal structures. This capability allows deep learning architectures to outperform classical machine learning (ML) methodologies in areas relying upon generative modeling, inference, predictions, and perception. Due to deep needs from industrywide adoption and increasing evidence from success in creating systems performing at a greater than human-level of accuracy, AI-driven technologies are used to assist people in accomplishing more efficiently the missions they are on through automating laborious tasks or building smart devices; therefore, deep learning has been involved in revolutionary contributions to several industry sectors: natural language processing (NLP), computer Vision (CV), cyber-Security robotics; etc and has produced numerous scientific discoveries by creating systems capable of solving problems faster, cheaper, and with greater accuracy than their human counterparts. Using the scalability and flexibility characteristics of deep learning has also enabled the development of smart technology in many diverse real-world applications like environmental improvements, medical fields, and automating vehicles; therefore being at the forefront of developing smart technology[5, 6].

1.5 Key Milestones and Paradigm Shifts

Changes in deep learning are a result of two major developments which have caused the development of the technology to progress. These include; successful implementations of the backpropagation algorithm (the efficient backpropagation algorithm) and use of GPU's for training neural networks. The introduction of convolutional networks has set new performance standards for computer vision and the use of recurrent networks will allow the ability of deepening networks for speech and language processing. Transformer [7], a newer class of neural network models, have developed a completely different family of models based on providing 'attention' to more than one input sequence model as well as providing a method of scaling such models in terms of their training and production performance. Finally, a new model, variational diffusion, has provided additional opportunities for producing high quality images with multimodal datasets and scientific simulations. The rapid and continuing disruptions due to advances in these technology will continue to advance the evolving use and types of architectures used for constructing deep learning systems[8].

1.6 Objectives, Scope, and Contributions of the Present Survey

The aim of this report is to show different emerging technologies for increasing the efficiency of DL Models and provide examples of applications and architectures of DL Networks as a repository of the ever-evolving hardware and neural networks landscape. The required information includes (a) a working knowledge of DL Fundamentals, (b) comprehensive discussion of DL architectures including Transformer [7], RNNs, CNNs, GNNs, and Generative Models; (c) examples of developing applications in the technological field and the social/natural sciences; (d) examples of technical barriers to the development of this field. In addition to ethical issues related to deployment and considerations; and (e) A forecast of the evolution of DL from this study [9].

2. METHODOLOGY

2.1 Literature Search Strategy

The thorough literature search strategy was employed to ensure the most up-to-date and valid sources of deep learning were captured in the survey. The strategy targeted breadth and depth of the included studies; therefore, including theoretical works on the foundations of deep learning, upgrades to architectures, benchmark studies and application-specific research. Controlled vocabulary terms and free text keywords were used to maximise search sensitivity e.g. "deep learning OR neural networks OR convolutional networks OR transformers OR graph neural network AND generative model AND machine learning applications" [10]. Boolean operators and parameters were used in the search to assist with removing irrelevant studies; for example forward and backward tracing used in the source search process looked at references within the study to find key original references that were not obtained using database searches. The iterative search process was used to identify an unbiased, representative sample of the current literature on deep learning.

2.2 Databases Used (Scopus, IEEE Xplore, ACM, Web of Science, arXiv)

To create an extensive listing of existing research in relation to machine learning, we searched through nine of the highest quality academic databases that collectively represented all types of academic literature (journals, conference presentations, technical articles and preprints). Both Scopus and Web of Science were chosen because they contain a significant volume of peer-reviewed literature in the field of engineering, computer science and other interdisciplinary areas. IEEE Xplore and the ACM Digital Library were selected as they are both leading publishers of cutting-edge work in areas such as artificial intelligence, machine learning and computing algorithms, among others. Finally, the database arXiv was also included so as to provide access to preprint literature, which frequently represents the first published account of novel progress made in relation to deep learning technologies. Both of these databases thus serve as a comprehensive and authoritative base of contemporary evidence upon which to draw conclusions [11].

2.3 Inclusion and Exclusion Criteria

Inclusion and exclusion criteria were transparent to ensure the selection of only relevant high quality publications. I- I.C Inclusion Criteria Study Inclusion Criteria: All studies must examine at least one of the following: deep learning architectures, training paradigms for deep learning, performance evaluations of applications of deep learning. Preference will be given to those articles which introduce new models, have state-of-the-art evaluation results, or provide significant advancement(s) of theory of deep learning and its application to real world problems. Articles will not be included in this review if they lack methodological details, evaluation results or in-depth discussions of deep learning, or for any reason, do not focus on some aspect of deep learning (for instance, if the article is about some classical machine learning approach that does not involve a neural network). Non English papers and/or those without full text availability will also be excluded. These criteria together provide for the methodological and conceptual soundness of the studies included in the review[12].

2.4 Data Extraction and Thematic Categorization

Structured data extraction provided a method for collecting the most central information from the identified articles that included essential details about the model architecture, training strategy, experimental setup and equipment used, data set and evaluation metrics, as well as the reported results. Each text data set was formed from information on several themes that can be identified as significant topics of study within the fields of deep learning as well as other major activity areas made up of convolutional structures, recurrent networks, attention mechanisms, graph structures and generative models. Additionally, several different categories for application of computer vision, natural language processing, speech recognition, health care, Internet of Things (IoT), agriculture, and cybersecurity also have been created for use with these types of research studies. The thematic classification will enable researchers to generate more consistent syntheses of evidence from systematic meta-analysis studies; compare and validate findings across multiple studies; and make all methodological and convergence issues relative to performance among each discipline explicit. The structured data extraction and thematic category creation establishes a degree of transparency and reproducibility of the analytical framework utilized within the study for the purposes of designing and conducting survey research [13].

2.5 Evaluation Parameters and Synthesis Approach

We synthesized a comparison of the literature based on a set of evaluation criteria that were developed with a focus on the methodological quality of studies, the scalability of the methodology based on existing practice, whether the methodology could be computationally efficient for implementation, and if the methodology had practical meaning. Evaluation criteria included the novelty of the architecture used, diversity of data sets used in conducting experiments, performance relative to baseline studies, human interpretability/understanding, robustness to noise or adverse effects, and potential for real-world implementation. We utilized comparative analysis methods to assess trends, variations in relative strengths and weaknesses between deep learning models (we characterize as MOLM), and similarities and differences between agreements and disagreements within the literature. The synthesis also includes in-depth analysis of how specific innovative architectures address foundational problems associated with using machine learning and affect the performance of machine learning applications. This review format permits more balanced, less biased, and richer analyses of literature[14].

2.6 Limitations of Survey Methodology

Moreover, although the survey approach is thorough and methodical (as you would expect with any research), it suffers from many shortcomings that arise because of the speed at which research in this area is evolving. For example, you can be certain that new deep-learning paradigms and discoveries will be created well after your literature review is over. In addition, when you rely only on peer-reviewed literature, there is the possibility of a publication bias; published works usually tend not to include any negative results from studies that have been conducted, nor are there many instances of negative findings that are published. Third, while preprint repositories vary in terms of how each is peer-reviewed, the ability to use preprint articles helps document how quickly advances in deep learning have been able to progress to publication. Finally, thematic categorisation for hybrid/multi-domain models is typically unusable. Defining your research boundaries will assist you in defining the extent of your survey, as well as how you intend to interpret it through the use of your chosen survey methodology [4].

3. FOUNDATIONS OF DEEP LEARNING

3.1 Neural Network Fundamentals

Deep learning uses artificial neural networks (ANN) as models based on the connected networks of neurons in the human brain. In a neural network each node (or neuron) has its own input(s) and computes a weighted sum of those inputs, then applies an activation function. The more hidden layers there are, the more inputs can be combined into higher order abstractions of the input, which makes them better able to approximate complex functions and decision boundaries (i.e., single functions or boundaries that can not be accomplished with less than five inputs). The training of the neural network is actually tuning the weight values of the various parameters in the network using backpropagation, where the gradient of the loss function (with respect to the input/output relationship of the network) is computed with respect to the weights of all the connections between the nodes (i.e., the parameters of the network model). There are many neural networks that represent networks of varying numbers of layers (tens to hundreds) that have enough capacity to represent the complex inter-relationships or hierarchies in the data. The underlying characteristic of the artificial neural networks is the depth, which differentiates modern deep learning algorithms from previous models of the same type (i.e., models to assist with perceptions, predictions or generative synthesis); the depth is what gives neural networks the means to provide performance beyond any previous models [15].

3.2 Learning Paradigms: Supervised, Unsupervised, Semi-Supervised, and Reinforcement

Different types of supervision and data accessibility are explored through the application of deep learning techniques in various learning paradigms. Supervised learning employs a model trained by utilizing labeled data that provides a direct mapping from the input to the corresponding desired output. Supervised Learning is the most used application for the tasks of image classification, speech recognition, and language translation. An example of an unsupervised learning paradigm would be learning patterns from the underlying structure of large datasets that do not contain any labeled data by utilizing either clusters or procedures involving auto-encoders to learn significant lower-level patterns. In the case of semi-supervised learning, there are two types of models that utilize both a small set of labeled data and a much larger set of unlabeled data for the purpose of developing greater generalization potential. Finally, reinforcement learning is another well-established type of learning paradigm in which agents perform actions in an environment to receive feedback based on a trial and error approach in order to maximize the amount of reward they have received over time. The two paradigms are therefore complementary to one another in that they each have the ability to leverage the unique strengths of deep learning when working with varied data types, as well as meeting the needs of different applications[16].

3.3 Optimization Algorithms and Training Strategies

At the core of deep learning is optimization since the efficiency of a neural network's output is determined by how accurately a loss function is minimized (e.g., measuring how far off predictions were from being accurate or how many different types were used to manage larger data sets when utilizing multiple computers on a distributed system). Stochastic Gradient Descent (SGD) continues to be the main optimization algorithm due to its ability to provide good results while using small amounts of data (i.e., ”mini-batch”) instead of requiring complete access to a full dataset (i.e., having to process the entire dataset). Adaptive optimizers such as Adam [17], RMSProp [18] and Adagrad allow for more flexible learning rate adaptation by using metrics such as historical gradient data. These methods allow for both increased stability in convergence and faster convergence when compared to other adaptive optimizers such as SGD. Other important improvements to the training process are realized through advanced training techniques such as momentum-based acceleration of training, adaptive scheduling of the learning rate, and proper initialization of weights. Additionally, PSA methods for inductive bias on circuit-based architectures provide a hardware solution to avoid exploding/vanishing gradients through the training of deep and complex neural networks. These combined aspects of optimizations all play an important role in the speed at which current leading-edge neural networks can perform effectively, performance-related improvements continue! [19].

3.4 Regularization, Normalization, and Generalization Techniques

In order to manage generalizability for models, avoiding overfitting, and producing consistent results across unseen data is critical, therefore, regularization techniques are essential. Traditional techniques such as the L1 and L2 regularization constrain a model's parameters through various methods in order to control its complexity, while techniques such as dropout deactivate neurons randomly throughout the training period support distributed representations. Normalization techniques such as batch normalization [20], layer normalization [21, 22], and group normalization [23], stabilize the training dynamics of intermediary activations and reduce the time required to converge. Data augmentation helps to generalize models by artificially expanding the training distribution through various processes such as noise injection, rotation, and scaling. More recent techniques such as adversarial training, mixup, and manifold-based regularization build upon these fundamentals, especially when seeking to ensure robustness against distributions' shifts and perturbations. The use of these various techniques may provide a wealth of possibilities to improve the robustness and transferability capabilities of deep learning models [24].

3.5 Performance Metrics in Deep Learning

Evaluation of deep learning systems requires measurement of performance according to specific properties of the task being performed. Typically, with any type of classification task, the metrics of accuracy and the metrics of precision, recall and F1 score will give you a meaningful way to determine how well the correct solution matches up to the incorrect solutions. Often in regression, mean squared error, mean absolute error and correlation coefficients will be the key evaluation standards. In general, when looking at evaluating models for ranking or retrieval, you will often come across metrics such as AUC, mAP and NDCG that are considered common evaluation metrics. In application of computer vision to real-world tasks, many objects will be identified through object detection or segmentation. In this case, object detection identification is completed using metrics such as IoU, AP and Dice and used to compare performance against other object detection applications. In natural language processing, different metrics will be used such as BLEU, ROUGE and perplexity that vary based on the type of task being studied. Choosing an appropriate metric to evaluate your models will become critical to evaluate the comparison between models across architectures, and also to ultimately choose an appropriate model for your production systems[25].

4. RECENT ADVANCES IN DEEP LEARNING ARCHITECTURES

4.1 Convolutional Neural Networks (CNNs)

Convolutional Neural Network (CNN) is one of the most widely used architectures in deep learning for Most primarily with respect to computer vision. CNN architecture represents the application of pooling and local receptive fields as representations of the underlying hierarchical structure seen in images. While the development of CNNs can be traced back to earlier network designs such as LeNet, there was a resurgence of CNN architectures into popularity when the CNN architecture of AlexNet [26] demonstrated that networks with many layers, can learn from numerous labelled training examples with processing power from GPUs, can achieve significant improvements over previously existing methods. Later architectural designs such as VGGNet and ResNet further developed deeper and more robust architectures through the adjustment of convolutional blocks/filters and the addition of skip connections between layers in order to enable learning from more extensive networks. More recent advancements in architectural design, including ConvNeXt and EfficientNet have improved normalization of architectural design, the compound scaling method, and the parameter efficiency of architectures and have demonstrated that increased efficiency does provide for improved performance with fewer computations. This series of architectural advances has expanded the number of different application areas where CNNs are able to be applied, to include remote sensing, medical imaging, industrial inspection, and others, and has helped to validate the role of CNNs as fundamental methods for solving high dimensional spatial learning tasks [27]. TABLE 1 shows a summary of Deep Learning Architectures.

Table 1: Summary of Deep Learning Architectures

Architecture	**Key Innovation**	**Primary Applications**	**Representative Models**
CNN	Local receptive fields, weight sharing	Image classification, object detection	LeNet [28], AlexNet [26], ResNet [29]
RNN/LSTM/GRU	Temporal dependencies, gating mechanisms	Sequence modeling, time series	LSTM [2], GRU [30]
Transformer	Self-attention, parallel processing	NLP, vision, multimodal tasks	Transformer [7], BERT [31], GPT [32]
GNN	Graph structure learning	Molecular modeling, social networks	GCN [33], GAT [34]
GAN	Adversarial training	Image generation, style transfer	DCGAN [35], StyleGAN [36]
VAE	Probabilistic latent space	Data generation, anomaly detection	VAE [37], β-VAE [38]
Diffusion Models	Denoising process	High-quality image generation	DDPM [39], Latent Diffusion [40]

4.1.1 Historical evolution

CNNs have gone from simple features to advanced networks that can handle complex inputs. LeNet was the first architecture defined with convolution-pooling. AlexNet was the first to really show off effective use of ReLU activations and GPU training for deep-learning models. VGGNet was also a very successful recognition architecture, and it provided evidence that depth of layers is important; it also costs a lot and provides no new solutions. ResNet introduced multiple updates by using a residual learning paradigm, which addressed the vanishing gradient problem and allowed for the stable training of ultradeep networks. These innovations have led to the implementation of modern networks that are appropriately scaled for resolution, width, and depth [41].

4.1.2 Advanced CNN architectures (EfficientNet, ConvNeXt)

CNN architectures today are designed for high performance in terms of scalability, accuracy, and efficiency. The new scalable approach of EfficientNet [42], uses a compound scaling method, enabling model scaling (e.g., resolutions, depth & width) to occur consistently with less overall degradation than conventional scaling methods.

By implementing design principles from transformer architectures, ConvNeXt [43], has also helped advance CNNs into the era of deep learning, allowing them to compete more closely with attention-based networks. Ultimately, these CNN enhancements have demonstrated the significant contribution of CNNs to the current growth of transformer-based architectures, and, more importantly, they have exhibited comparable performance with image classification accuracy (IF-classification) to that of an attention-based model, and thus are expected to improve data propagation and use at the edge and in practical applications [44].

4.2 Recurrent Neural Networks (RNNs) and LSTM [2]/GRU

The main reason why RNNs were crucial is that they allowed deep learning methods to process sequential data because an RNN's use of recurrent hidden states allows it to take into account time-based relationships between the hidden states of each time step. Unfortunately, RNNs have been affected by both vanishing gradients and exploding gradients which limits their ability to model long range dependencies. The introduction of GRUs and LSTMs solved the problems of vanishing and exploding gradients through the use of gating mechanisms that control the flow of information between the outputs of each time step. As such, GRUs and LSTMs transformed how algorithms can be applied to tasks such as time-series forecasting, machine translation, and speech recognition. Currently, despite the superior performance of transformer models relative to RNNs, LSTMs and GRUs are still the best choice when you need a lightweight, real-time, or low-computational-load solution. Additionally, they are comparatively simple, stable, and interpretable for use with signal processing, biomedicine, or embedded systems [45].

4.2.1 Sequential modeling improvements

Sequential modeling has improved much more than just how we use classic RNNs to model sequential data. Bidirectional processing, the use of an attention mechanism, and hierarchical recurrent networks have all been proposed as methods for capturing forward and backward dependencies of sequentially dependent phenomena. Dilated kernels in temporal convolutional networks (TCNs) further provide a design for parallelizable networks while still maintaining the temporal order necessary to model these sequentially dependent phenomena. However, LSTM and GRU models remain much more robust for use within compact networks on smaller data sets. The continued success of LSTM and GRU models indicates the continuing relevance of some of the more traditional forms of recurrent modeling for domains where the temporal aspect of data has significant importance and computationally limited resources are abundant [46].

4.2.2 Transition toward attention-Based models

Recently, attention-based architectures have changed the traditional way we use recurrence in sequence modeling. By allowing for easy representation of long-distance dependencies without incurring a delay due to sequential calculations, attention enables us to create scalable and more effective models than have been produced by RNNs to this point in time. As a result, RNNs have essentially been replaced by transformer networks in the field of Natural Language Processing (NLP) and are beginning to pervade other areas of multimodal processing (e.g., Vision, Audio). A fascinating aspect of this transition from RNNs to transformer networks is to investigate how these models bridge the gap between these two types of approaches to enable a greater degree of flexibility and generality for modeling sequences [47].

4.3 Transformer Models [7]

4.3.1 Architecture fundamentals

One of the most transformational technologies of deep learning has been the Transformer (Trans) [7], as it has transitioned from using a recurrence-based architecture to a self-attention based architecture, allowing the model to process all the elements in parallel. The major building blocks of this new architecture consist of: an encoder-decoder architecture; multi-head attention; a location (positional) encoding for each input sequence; and a feed-forward network. The Trans can also take into consideration long-range dependencies among the various elements of an input sequence, allowing the network to build out its architecture into more scalable and detailed representations. The properties mentioned above indicate that Transformers will likely remain the primary technology for NLP and many other applications beyond natural language (such as in vision, audio, reinforcement learning, and computational biology), and so forth [48].

4.3.2 Vision Transformer (ViT)

Implementing self-attention in vision transformers (ViTs) leads to improved performance in image classification over convolutional networks. ViTs can divide images into ”patches” which become the ”tokens” for attention modeling at their global spatial coordinates (positions). Since there is no longer a requirement for the inductive bias that is built into convolutional networks, transformer-based models have been able to grow rapidly in visual tasks. Variants on the original ViT design like Swin Transformer [7], and Data Efficient Image Transformer (DeiT) have developed new hierarchical structures, used motion windows, and come up with new efficient data training methods that allow ViTs to be competitive across the range of tasks in vision [49].

4.3.3 Large Language Models (LLMs)

With the advent of Large Language Models (LLMs)2,3 that incorporate transformer architectures and trained on vast amounts of data, natural language processing (NLP) has experienced dramatic change. The performance demonstrated by language models on zero- or few-shot tasks, generating fluent and contextually relevant text and performing reasoning across many domains has made them very impactful both in academia and in industry. LLMs are representative of the scalable hypothesis in that they exhibit predictable model capacity growth through increasing parameters, compute resources, and data set size. To date, LLMs have been utilized for a variety of purposes including multi-modal AI, discovering new science, providing decision support for healthcare, generating code, etc., and LLMs will be among today's most significant advances in terms of machine learning systems [50].

4.4 Graph Neural Networks (GNNs)

Graph Neural Networks extend conventional neural networking into the modeling of relationships and the representation of graph-structured data by using complex algorithms that integrate an entire network's nodes into single network nodes. Initial architectures like Graph Convolutional Networks (GCNs) use both the topological and the relational features of the original graph to develop new representations of the original graphs and graph structures. More flexible and powerful models were then created based on GNNs, and created to meet increasing demand for their effectiveness in solving relational data modeling problems. Different architectures of GNN, just like GCN, GMLs, have either created or improved the speed of processing through Attention mechanisms (GATs), created or improved their capabilities through Graph Transformers [7], and while no single model of GNN is perfect for all problems, many of these models excel at solving the relational data modeling problems of molecular discovery, recommendation systems, social network analysis, and knowledge or semantic graphs where determining the relationship between different graph nodes is critical to their successful development and use [51].

4.4.1 Message passing and embedding strategies

Message-Passing Neural Networks (MPNNs) are a general form of Graph Neural Networks (GNNs). In MPNN's, each node updates it's representation/model based on combining features from the nodes/oother vertex on the graph that are neighbouring to it; this update it iteratively performed so that the resulting embedded representation (or model) of the object at each vertex contains is a function of both the features at the object (node) level and the local topology of the graph. The MPNN framework has produced a number of key schema/techniques for creating powerful representation/model learning and scalability approaches such as sampling from a neighbourhood, using spectral convolution, and creating multi-hop aggregates of feature representations for the nodes in a neighbourhood [52, 53]. MPNNs have produced a number of state of the art GNN models that could be applied to problems in the area of physical/chemical science, and also for a variety of tasks involving structured prediction [54] (e.g., classification or regression).

4.4.2 Recent innovations: GAT, graph transformer [7]

In addition to GATs, practices based on graph transformers utilize message passing containment; this allows the modeler to gather all of the neighbor nodes in each receiver node. This is extremely efficient in constructing complex topology structures, particularly those containing wide reaching delegate dependencies. The advancement of these types of models has greatly aided in the outcome of predictive molecular property modeling, the modeling of protein structure, drug product development, and in the building of recommendation engines on a massive scale, all of which have transformed how graph-based machine learning is being applied [55, 56].

4.5 Generative Deep Learning Models

4.5.1 Generative Adversarial Networks (GANs)

A GAN (generative adversarial network) consists of a generator and a discriminator that are trained together in an adversarial manner to produce realistic synthetic data. GANs have driven many of the recent advancements in image generation, super-resolution, style transfer, as well as many other creative uses for images and video. The advancements made in GAN architectures, particularly with the development of Wasserstein GANs, StyleGAN and Progressive Growing GANs, have greatly enhanced the training stability, quality of generated samples, and variety of generated samples for all tasks and input sources within the GAN family, making it one of the most commonly used classes of generative models in deep learning [57].

4.5.2 Variational Autoencoders (VAEs)

Variational Autoencoders (VAE [37]) that are configured to have a 'bottleneck' structure provide a probabilistic way to model data via generating a dataset and learning the latent distribution of that dataset. VAEs are scalable via being able to create latent space that is both smooth as well as interpretable; thus, they can assist in tasks related to reconstructing data, detecting anomalies within a data set and helping create representations of the data set. There have been a number of improvements made to VAEs; including, but not limited to: β-VAE, hierarchical VAEs, and flow based decoders which have resulted in improved disentanglement and expressiveness, thus providing VAEs with potential applications for vision, speech and scientific modeling [58].

4.5.3 Diffusion models [39], and their breakthroughs

Diffusion models represent a different class of prominent types of recently developed generative AIs through the use of learned process for denoising random noise to create realistic images. Compared to GAN [59], diffusion models produce more stable, more controllable and more believable results and are therefore being employed in a wide variety of creative, scientific, and industrial applications. The ability of diffusion models to generate authentic images has been greatly enhanced by recent development of DDPMs, Latent Diffusion Models, and complicated text-to-image architectures; in fact, diffusion is now playing a pivotal role in generating finding new methods for creating generative work within the last few years [60].

4.6 Emerging and Hybrid Models

4.6.1 Neuro-symbolic deep learning

Neuro-symbolic systems combine both sets of benefits; logically defined rules and deep learning together means systems can receive input through the application of symbolic links with the ability to apply logic as a result of a new experience. The approach can address many limitations with data driven models such as providing the user with interpretability, robustness, and the ability to reason. Other areas of interest include automated theorem proving, program synthesis, knowledge representation and explainable artificial intelligence [61], and these areas will continue to have new types of neuro-symbolic systems that will fill these areas.

4.6.2 Federated deep learning

Decentralized datasets can be used to train deep models without having to share any of that sensitive information with the central server with the help of federated learning. This concept is necessary for creating AI systems which respect privacy in industries such as healthcare, finance, mobile systems and collaborative systems across industries. By employing techniques such as secure aggregation, differential privacy, and personalized federated optimizations, these two requirements (model utility and confidentiality) can coexist.

4.6.3 TinyML and edge-based deep learning

The intention of TinyML is wide-ranging, from deploying deep learning models on low-powered, resource-constrained devices such as microcontrollers and internet-based sensors, through to producing rapid real time, cross-platform controls in order to drive large scale autonomous sensor development as well as embed AI applications into the physical world [62]; by leveraging model compression, quantization, pruning and hardware-aware design one can run a neural network at the edge with extremely low energy use.

5. APPLICATIONS OF DEEP LEARNING

5.1 Computer Vision

Computer vision has witnessed many advancements due to deep learning as it enables the development of models to perform high-level visual tasks with an accuracy level comparable to human accuracy. Convolutional neural networks (CNNs) in combination with transformer-like architectures provide a means to build spatial scaling hierarchies, which have been critical to advancements in segmentation, detection and classification of objects, and reasoning based on visual input. In industrial applications, deep learning is often used for quality control, defect detection and automated inspection, increasing efficiency and reliability. In automated driving applications, perception modules utilize computing resources from deep vision systems to process data from multiple (radar, LiDAR, and cameras) to assist with safe navigation through the identification of road conditions and obstacles. Deep learning applications in medical imaging can be used to interpret histopathological slides, MRIs, radiographs, and CT scans, and have demonstrated a high degree of accuracy, which helps practitioners make more accurate decisions when creating treatment plans and diagnosing patients earlier. These examples illustrate how significant deep learning is for image analysis in industrial and scientific applications [63].

5.1.1 Image classification, detection, segmentation

Images classified using deep learning have made significant advancements in the basic tasks associated with visual perception. State-of-the-art deep learning architectures (e.g., ViT, DenseNet, ResNet [29]) provide high accuracy for image classification-based object detection. Also, object detection models (e.g. Faster R-CNN, DETR, YOLO) are able to perform real-time object recognition and localization. Models used for pixel-level segmentation, such as semantic segmentation models, are effective at performing vision-related tasks like tracking objects, environmental monitoring, self-driving vehicles and medical imaging. These tasks provide the foundation for visual perception intelligence systems to be used in a variety of industries including consumer electronics, robotics, health care and security [64, 65].

5.1.2 Industrial inspection, biometrics, autonomous driving

Deep Learning is used in the manufacturing sector to automate quality control processes and defect detection, thereby improving the accuracy of production and decreasing human error. In the area of biometric security (gait recognition, iris recognition and facial recognition), deep embedding is used for identity verification. Solutions to autonomous mobility applications rely on deep models to interpret complex visual environments and road hazards to make real time driving decisions. These examples illustrate the real-world advantages of deep learning applications in diversely challenging environments with a significant need for rapid visual evaluation, robustness and accuracy [66], as shown in Table 2.

Table 2: Deep Learning Applications by Domain

Domain	Key Tasks	Architectures Used	Impact
Computer Vision	Classification, detection, segmentation	CNN, ViT, YOLO, ResNet	Human-level accuracy in image recognition
Natural Language Processing	Translation, sentiment analysis, QA	Transformer, BERT, GPT, LLMs	Near-human language understanding
Speech & Audio	ASR, TTS, voice generation	RNN, Transformer, WaveNet	Real-time transcription and synthesis
Healthcare	Medical imaging, diagnosis, drug discovery	CNN, GNN, Transformer	Clinical-level diagnostic accuracy
IoT & Smart Systems	Anomaly detection, predictive maintenance	LSTM, CNN, TinyML	Edge AI and intelligent automation
Agriculture	Crop monitoring, yield prediction	CNN, Transformer, Satellite imagery	Precision agriculture and resource optimization
Cybersecurity	Intrusion detection, malware classification	CNN, RNN, GNN, Autoencoders	Automated threat detection

5.2 Natural Language Processing (NLP)

The impact of deep learning in natural language understanding is creating machines capable of understanding and generating human languages. Numerous large-scale, transformer-based language models (LLMs) utilize contextual semantics, thus enabling accurate representations of long-distance relationships between elements of text in a way that had never been achieved with earlier technologies. Applications of LLMs include question and answer systems, sentiment analysis, machine translation, summarization of multiple content sources, information retrieval, and facilitating conversations between humans or between a human and a machine. Transfer learning and multilingual models have significantly benefitted low resource language technology, and the implementation of domain adaptation for LLMs allows businesses to further customize their models to perform very specific tasks, such as analysing legal/trade/scientific documents, or conducting clinical document analysis. Continuing advancements in deep learning systems will push computational systems beyond their current capabilities relative to processing of natural languages and the communication between human beings and computers [67, 68].

5.2.1 Machine translation

Using deep learning, or neural networks as we call them, transforms how we translate between multiple languages. With the advent of new attention-based approaches to translation (e.g. the transformer model), these neural machine translations (NMTs) have widely replaced traditional statistical methods by producing more natural (fluent) and contextualized translations across multiple language pairs. The ability for NMTs to account for idioms, long-distance relations, and specific terminologies enables language services that require NMTs to support global usage of low-resourced languages, thus diminishing the cost and effort associated with translating from one language to another [69].

5.2.2 Sentiment analysis and document classification

Methods that utilize deep learning are helping to improve sentiment analysis through greater precision, by using context-based polarity, irony, and other nuances of emotion associated with textual materials. Once we apply our pre-trained model representations, we can also do document classification, topic modeling, and opinion mining across a variety of segments including; marketing, political analysis and social network research. These types of analytic software help organizations make better choices by extracting valuable insights from large sets of textual material [70].

5.2.3 Large language models

According to LLM's ability to reason, produce content, match meaning and infer context at a higher level than any prior designed architecture has been demonstrated through training with large amounts of text data. The LLM architecture can adapt its ability to generalize (via zero-shot, few-shot, or learning via fine-tuning methods) to many large-scale downstream applications, including logical reasoning on multiple pieces of media simultaneously (text, images, sound, and structured information) [71].

5.3 Speech and Audio Processing

Speech and audio processing have benefited significantly from deep learning. By using this same technology, deep learning models have allowed computers to process, synthesize, and manipulate sound in ways that closely resemble those found in actual life or the world around us. There are several different styles of network architectures (e.g., CNNs, RNNs, and Transformers [7]) [47], that provide deep learning models with powerful methods to learn spectral patterns as well as temporal patterns. The models developed with deep learning can be used in numerous applications, including but not limited to, speech recognition, speaker verification (voice), emotion detection (speech), noise removal (speech and audiobooks), audio generation/music production. Many of these models are capable of real-time processing using on-device inference. This includes applications such as virtual assistants (e.g., Siri), intelligent automation of call centers, and assistive communication technologies. Therefore, deep learning has become the enabling technology upon which to make sound-based human-computer interaction more natural, more convenient, and more widely used [72].

5.3.1 Automatic Speech Recognition (ASR)

Deep Learning-based ASR systems with complete end-to-end functionality, including but not limited to RNN-T, CTC and Attention-Only Transformer [7], architectures, have been shown to perform accurately enough to be comparable to humans' transcribed output. There are several real-world applications that such systems can be used for, including: smart devices, transcription service providers, computer accessibility solutions for people with disabilities [73].

5.3.2 Speech synthesis and voice generation

Generative models like WaveNet [74], Tacotron, and diffusion audio synthesizers allow us to create realistic and expressive speech. Thus, they represent new opportunities to build high-quality text-to-speech systems. As a result, these models will help improve the naturalness of virtual assistant voices, the dubbing of movies, cloning voices, and other forms of interactive entertainment. The realism of these models will keep improving even further by continuing to advance the models of prosody and emotion [75].

5.3.3 Audio event detection

Generative Audio Synthesis Available Today WaveNet [74] (by Google) and diffusing-based synthesizers provide natural sounding, expressive speech with high quality text-to-speech (TTS) synthesis systems. This technology powers virtual assistant applications, video dubbing or cloning and interactive games. To enhance these systems further will require developing models for up-to-date prosody and emotion expression [76].

5.4 Healthcare and Biomedical Engineering

Deep learning technology has led to significant advances in diagnostics, prognostics, and personalized treatment options in the fields of healthcare and biomedical engineering. By utilizing deep models to analyze multiple sources of complex clinical data (for example, medical imaging, electronic health records (EHRs), genomic data, and biosignals), deep learning support clinical decision-making activities. In the area of radiology, for example, deep learning helps clinicians identify, classify, and estimate the risk of lesions. In genomics, deep models built from the sequence of DNA are used to identify variations in the genome that may be associated with susceptibility to certain diseases. Deep learning will also enable researchers to better monitor cardiovascular and/or blood-related changes (for example, vascular age) over time from continuous wearable sensor data (for example, sore wrists). These advances in medical science will improve the accuracy, availability, and efficacy of the healthcare system on a global level [77].

5.4.1 Medical image analysis

Deep Learning can perform powerful tasks such as: tumor detection, segmenting organs, and characterizing lesions utilizing CT, MRI, Ultrasound, and Digital Pathology images. Models such as U-Net [78], and transformer type architectures produce clinical-level accuracy that can be utilized by clinicians in the screening and diagnostic workflow as well as to detect diseases at the earliest stages [66, 67].

5.4.2 Clinical prediction models

Employing deep learning models to evaluate EHR, lab, and time-series patient data, can offer predict future risks and progress of conditions, and treatment options based on patients' previous clinical outcomes. Migrating attention means that the emphasis has shifted away from static changes in time to the dynamic changes that occur within time, thus improving the interpretation of clinical diagnosis at time of care [79].

5.4.3 Drug discovery and genomics

Deep learning has sped up drug development with its focus on generating new molecules, predicting molecular properties, and modeling the interactions between proteins and ligands. Molecular structure representation has been helped by GNNs and transformers, which have achieved impressive success. In addition, many genomics sequence-based models have produced interpretable functional variants and regulatory patterns for complex diseases, supporting research in precision medicine as well as bioinformatics [80].

5.5 Smart Systems and the Internet of Things (IoT)

Deep learning lays the groundwork for smart appliances and other smart devices through on-demand analysis of data in real-time, learning to adapt to new situations and generate predictive thoughts. Edge-optimized models or lightweight models are used on sensor data input devices in a variety of examples of smart homes, connected vehicles, industrial IoT platforms, and environmental monitoring systems. These applications of deep learning will improve intelligent management of energy resources, predictive maintenance, activity detection, and situational awareness in IoT systems. With the increasing prevalence of TinyML, there will be increasing opportunities for AI-based IoT ecosystems to have an intelligent capability at individual devices with extremely low power usage [81].

5.5.1 Smart homes and devices

Deep Learning is revolutionizing smart homes by allowing for activity recognition, intrusion detection, environmental control and automated personalization of interaction with appliances as a result of the use of artificial intelligence (AI) → AI built into appliances will improve user comfort and provide potential energy savings [82].

5.5.2 Traffic and mobility prediction

Deep learning models that utilize multimodal transport data to make predictions (eg GPS data from phones, traffic sensor data, public transportation logs) can be used to optimize routes, reduce congestion, etc. for Intelligent Transportation Systems → the interactions between dependence of mobility are modelled via spatio-temporal and graph based models [64, 65].

5.5.3 Environmental and sensor data analytics

The Internet of Things (IoT) can collect information through sensors from many different industries such as weather (connectivity), air pollution, agricultural monitoring, industrial process monitoring, etc. → Deep learning Systems provide improved online monitoring, locating abnormal wells and forecasting maintenance opportunities, providing improved safety and support to protect the Environment in these industries [83].

5.6 Agriculture and Remote Sensing

Precision Agriculture, along with Environmental Monitoring, uses deep learning techniques to analyse satellite imagery, drone flying data and/or multi-spectral sensor feeds. Developers create models thatperform Crop Health Monitoring, Yield Estimating, Soil Properties Inference, Vegetation Identification and Water Management. On the other hand, convolutional models provide an opportunity for utilising RS imagery with deep learning-based solutions to create semantic segmentation and land-use classifications, large area change detections via remote sensing. Finally, these technology deliverables further sustainble agriculture by providing early warning systems to hazards imposed by the environment and significantly enhance resource management [84].

5.6.1 Crop monitoring

Aerial imagery is inspected for pest invasion, nutrient loss, and disease using Transformers [7], and CNNs. By using these models, wasted resources are reduced as a result of increased agricultural production [55].

5.6.2 Soil and vegetation analysis

Using deep learning, soil moisture, biomass, and vegetation indices are produced from hyperspectral and multi-spectral data. This can be utilized for rangeland management and in situ conservation [85].

5.6.3 Satellite image interpretation

The use of models such as DeepLab, Swin, UNet++, and Transformer [7], will provide disaster management, monitoring of deforestation, and detailed land cover mapping based on high resolution satellite images. These applications support disaster response, weather research, and ecological monitoring [86].

5.7 Cybersecurity and Anomaly Detection

Cybersecurity has been using deep learning to automate how we detect and reduce the changing nature of threats, as there are no two attacks are alike; systems need solutions that will continuously adapt as these threats evolve. Using an array of data sources—network traffic, user activity, and system logs—to identify anomalies and detect intrusions/malware with a high level of accuracy, some techniques include: autoencoders, graph networks, and transformers are often utilized for detecting anomalous activity patterns. In addition, techniques associated with deep learning have been implemented for anti-phishing techniques, fraud detection, biometric authentication, and secure access control. The ability of deep learning to capture complex temporal signatures and relational signatures also makes it a very attractive option for the ever-evolving cyber-attacks [87].

5.7.1 Intrusion detection

Using Deep Networks, suspicious patterns(in this case,i.e. Distributed Denial of Service Attack, Unauthorised Access and Covert Communication Channels) are searched within High Dimensional Network Flow Data. The use of Hybrid CNN-RNN architectures in conjunction with GNNs, increases accuracy of detection [88].

5.7.2 Fraud detection

As far as financial systems go, abnormal transaction detection/fraud detection can be aided via the use of deep learning to recognize patterns of fraudulent activity. When it comes to detecting these types of challenges, the sequence and graph models realize there are complicated dependency structures between different accounts that will assist in the early detection of fraud [89].

5.7.3 Malware classification

The use of CNN and transformer neural networks has improved malware detection using deep learning by allowing the analysis of executable files based not only on signatures but also on their respective system behaviours and characteristics while taking into account effective database boundary conditions. These models provide a means for identifying malware that cannot easily be detected by traditional methods, such as obfuscated and polymorphic malware [90].

6. TECHNICAL, ETHICAL, AND PRACTICAL CHALLENGES

6.1 Data Availability and Annotation Issues

Due to the exponential growth of deep learning, it continues to be challenging to obtain high-quality, diverse, large-scale datasets across domains, with the lack of vast quantities of labelled training datasets to allow for developing generalisable features in real-world applications. Furthermore, some domains such as medical and security have limitations on the number or sensitivity of data available that result in increased annotation costs due to regulatory requirements for the proper storage and transmission of sensitive data. Additionally, the difficulty in annotating these datasets provides limited opportunity for most persons to conduct the annotation as it is often less relevant or valuable than with most other applications (e.g., scientific research). The challenge associated with obtaining labelled datasets is also compounded by the time and resources required to complete the annotation process manually, particularly in challenging applications such as the diagnosis of patients within the healthcare domain or the monitoring of environmental data. Differences in the data collection process, particularly in terms of obtaining a representative sample of the population, can lead to poor performance of the resulting algorithms at predicting the intended outcomes for members of the population that is under-represented in the training datasets. These challenges highlight the need for better capabilities within the data-sharing community, advancement of synthetic data generation efforts, and use of semi-supervised or self-supervised learning approaches that are less dependent upon the successful completion of manually annotated datasets [91].

6.2 Computational Resource Constraints

Most state-of-the-art deep learning models require significant computing power, typically in the form of large GPU clusters or special hardware like TPUs, to train them. Accomplishing this is a daunting task for only the most well-resourced researchers/institutions. The costs associated with training large-scale models create barriers for many researchers accessing state-of-the-art models; additionally, reproducibility is limited due to the inability of many researchers to utilize training results produced on such a large-scale. Furthermore, there are also practical limitations associated with inference models for incorporating deep learning into real-time applications/mobile/plugged-in devices: they typically require both low power and low-latency. Quantization, pruning, compression, and knowledge distillation are all potential methods for easing constraints on the above challenges but developing viable solutions for determining acceptable trade-offs between processing accuracy and model efficiency remain a problem [92].

6.3 Model Explainability and Interpretability

Black-box characteristics associated with algorithms utilized within deep learning architectures present unique challenges to fulfilling the criteria of accountability, transparency, and trust, particularly in high risk applications such as finance, health care, and autonomous systems. Many deep learning architectures use “dark” layers which are extremely difficult to interpret and consequently difficult to clarify how models utilize certain inputs to produce specific predictions. Users utilizing non-interpretable models lack trust in these types of models, do not achieve compliance with appropriate regulatory requirements, and have difficulty debugging models in use due to lack of model transparency. Many approaches have been developed to address these issues, including Explainable AI (XAI) methodologies such as saliency maps, feature attribution methodologies, surrogate models, and intrinsically interpretable architectures. However, technical interpretability methodologies do not create a complete linkage between interpretability tools available to users of deep learning models and the users' true ability to interpret results. Achieving consistently interpretable results for black box deep learning methods that deliver performance competitive with existing state of the art methods remains a primary focus area for developing reliable and trustworthy AI [93].

6.4 Bias, Fairness, and Ethical Considerations

The potential for deep learning systems to replicate, reinforce or amplify existing societal biases, such as gender, race, ethnicity, age, and class, is largely a function of the structural inequality of the training data sets used to develop the systems. As a result, these systems can lead to socially unacceptable outcomes in a variety of real-world applications (i.e., hiring algorithms, facial recognition systems, loan approvals). The ethical issues associated with how an AI model can be misapplied (e.g., deceptive content) also play a critical role in the difficulty of developing ethical AI systems. Ethical considerations for AI systems must include (1) developing approaches to mitigate algorithmic bias (e.g., publishing best practices), (2) establishing ethical oversight, and (3) having responsible auditing practices in place across the entire lifecycle of the AI system [94].

6.5 Robustness, Security Threats, and Adversarial Attacks

Deep learning models can be attacked with many types of adversarial threats on their dependability and security. An Example: Attackers can use adversarial examples to lead a model into producing a harmful and/or incorrect output, creating significant security threats for some applications such as automated monitoring, biometric verification, and driverless cars. Likewise, models could also be vulnerable to Backdoor attacks, Data poisoning, Shifts in distribution, and over-sensitivity to unseen or noise conditions. Robustness includes all that requires designing a system that has been built to be inherently safe from being exploited by adversarial means through defensive techniques, i.e., input sanitization, certified robustness, or adversarial training, to real-time adaptive monitoring of everyday systems. However, the problem of adversarial robustness continues to be an unsolved problem that still needs to be addressed in order to develop reliable and safe implementations of deep learning in real-world settings [95].

6.6 Deployment Challenges in Real-World Systems

Deep learning systems are being applied to many real-world situations, but there are many operational challenges related to their implementation including regulatory compliance, and challenges related to scalability, maintainability and integration of deep learning models, as well as issues with operational performance degradation when the actual data observed in a real-world application deviates significantly from the dataset used to train the deep learning model due to shifts in distribution or a domain between the training and actual datasets. These deep learning models need to be monitored continually following their deployment to ensure reliability and hence it may be necessary to make adjustments or retrain the model based on dynamic principles. Furthermore, there may also be challenges related to network connectivity at the edge or with distributed applications and constraints related to latency and computer resources for the deployment of the deep learning technologies. Additionally, deep learning models must adhere to regulations related to domain-specific requirements, safety standards, and privacy laws. A successful implementation of a deep learning model requires extensive planning of the pipeline involved with monitoring and lifecycle management, as well as effective documentation and validation of the model in order to establish that deep learning technologies can be expected to perform consistently under true operational conditions [96].

7. FUTURE TRENDS AND RESEARCH DIRECTIONS

7.1 Explainable and Transparent Deep Learning

Due to their essential role in deep learning systems, increasingly focused on high-stakes applications of deep learning, such as public safety, healthcare, governance, and finance, there is increased demand for explainable AI systems. Future work should aim to develop models that are competitive in performance but also exhibit natural interpretable and trustworthy patterns along their reasoning pathways in a way that makes sense to human users.

Hybrid models that combine deep neural representations with some form of symbolic reasoning are one promising direction in terms of interpretability. For instance, [97] contains some preliminary evidence that this could be an area of good potential for developing more interpretable models.

7.2 Resource-Efficient and Green AI

The computational burden associated with many of the large deep learning models creates concerns around equity of access to AI resources, environmental impact, and energy consumption. Studies to follow will research techniques for developing "green AI" which both reduce energy consumption while achieving high performance; these techniques include more efficient training algorithm designs, compact design of CNN architectures, and creating sparse computation for reducing redundancy. The following are two examples of model compression techniques that have been developed to reduce computational costs: quantization, pruning [98].

7.3 Multimodal Learning Architectures

The study of multimodal learning is a rapidly growing and advancing area of research in the deep learning field, providing opportunities for models to interpret various forms of data including text, imagery, audio, video, and sensor information, as well as to reason about multiple forms of data simultaneously. Current research efforts are focusing on developing unified systems that enable the processing of multiple forms of data through a shared embedding space, while utilizing cross-modal attention mechanisms. The overall aim of building these types of models is to create systems that can process their environment holistically, such as when building autonomous vehicles, having an AI system assist with medical diagnosis or helping to create intelligent robots or Interactive AI. Ultimately, the ongoing development of multimodal fusion, as well as alignment and generative modelling within the AI space, will be vital components for creating artificial intelligence systems that perceive and reason in an integrated manner across all aspects of the human-like environment [99].

7.4 Continual, Lifelong, and Meta-Learning

Classic deep learning models typically train for their entirety on unchanging datasets and are thus not well suited for working within frequently changing domains because they will require eventual and periodic self-execution of modifications. A major cognitive obstacle with creating lifelong learning models continues to be catastrophic forgetting (i.e., the ability for newer information to overwrite previously learned knowledge). As we look into the future, the goal for creating new models will be to develop systems to learn continuously over the course of the lifetime of the model so that when new information is incorporated into a model, it does not lose access to the old knowledge it has learnt from the past. The goal with meta-learning or learning to learn is to develop a system that can rapidly extend existing knowledge based upon a few experiences (i.e., instances of training), have access to relevant knowledge, and have access to all training opportunities in order to produce optimal results. Such functionality will be critical in order to provide personalised AI systems, independent robotic systems, and real-time decision making in unstable environments [100].

7.5 Human–AI Collaboration Frameworks

In the new evolution of AI, technologies will be able to work in conjunction with people rather than being independently intelligent. One area of research to help create a new evolution of AI is to help develop machines that can support human and AI working together by having the machines learn how to incorporate human knowledge and context-based norms into their reasoning. Collaborative AI has many aspects where humans are involved in both learning by collaborating with machines and collaborative machine teaching through a more adaptive interface based on the user's intent. Emphasis on collaboration also ensures that advanced learning will increase human function and safety; will produce effective, and socially responsible deployment, in all domains such as education, health care, engineering, and creativity [101].

7.6 Evolution of Generative Models

Diffusion models [39], large scale generative transformers and multimodal generative architectures will fuel generative modeling advances. Generative Systems will also focus on increasing controllability, interpretability, and safety that can allow high-level semantic instructions to enable fine-grained manipulation of generated content. The use of generative AI with reinforcement learning and generally probabilistic modelling have the potential to further catalyse creativity, accuracy, and generalisability capabilities of generative models that tackle similar general tasks. Research into

Providing ethical security within generative modelling will continue to be a primary research focus [102].

7.7 Toward General-Purpose AI Systems

Long-term research in deep learning has included research into creating general-purpose AI/AGI systems that can perform many different tasks and exhibit some human-level flexibility when reasoning. The foundation models that scientists have been building in recent years are an example of this type of system: they are composed of large, heterogeneous datasets on which they have been created; they also have the ability to adapt quickly to a variety of downstream tasks (with relatively light fine-tuning). In order for scientists to build real AGI/AGIs, it is necessary that these systems have properties such as scalability, generalizability, robustness, and alignment with human values. Increasing numbers of researchers will be investigating architecture models that integrate perception/reasoning/planning/action within one system. These models have the potential to completely transform how we do scientific discovery, automate business processes, and make decisions; however, they also pose serious challenges for governance mechanisms designed to ensure those systems are safe to operate, fair and align with the will of society [103].

8. CONCLUSION

The development of deep learning has completely transformed the entire area of artificial intelligence and even though most people do not know how much or rapidly this technology is evolving and being used, this paper provides a detailed overview of the backgrounds, architecture modifications, real-life examples of usages, issues with technical implementation and potential research areas for both current and future deep learning ecosystem developments. This paper will look at how different types of deep neural networks have changed from the early studies using recurrent and convolutional networks to the newer implementations of transformers, graph neural networks and generative networks by showing all of these things can adapt to any type of architectural modification; therefore, it proves how flexible/convenient all modern day deep learning systems are. The new architectures that have developed using deep learning continue to be successful in all areas including: computer vision, natural language, speech/audio processing, health care, IoT, agriculture and cyber security demonstrate the vast documentation of real-world examples where deep learning is being applied.

Ever since deep learning had its transformational breakthrough, it has been faced with challenges of restricted access to data, high computational demands, lack of robustness, low levels of interpretability, and formal ethical considerations. Developing solutions for these problems is essential to creating deep learning-based systems that are reliable, transparent, fair, and safe across all use cases where they are deployed. The ongoing development of Explainable Artificial Intelligence (XAI), resource-bounded modelling, multimodal architectures, LIFE member support (long term evolving multi-task learning), human-AI collaborative systems, and generative models are creating a foundation for a new level of research into AI. Therefore, these emerging frontiers provide insights into future AIs that will be more general-purpose systems that adapt to human needs, are collaborative with humans as partners, and think like real people.

Deep Learning (DL) is at an important milestone in its evolution. It is moving from separate models designed for specific tasks to unified models that can perform multiple functions including perception, reasoning, creativity and decision-making. Future research must find a way to balance both innovation and accountability, which will ensure that the development of the best possible state-of-the-art DL systems will have a genuine affect on both science and society. This survey presents an overall, holistic viewpoint of the entire process associated with this evolution and enables researchers and developers of DL systems, policy-makers and teachers of DL systems to have a thorough comprehension of the most recent results related to DL in all areas of application.

References

[1] Trigka M, Dritsas E. A Comprehensive Survey of Deep Learning Approaches in Image Processing. Sensors. 2025;25:531.

[2] Hochreiter S, Schmidhuber J. Long Short-Term Memory. Neural Comput. 1997;9:1735-1780.

[3] LeCun Y, Bengio Y, Hinton GE. Deep Learning. Nature. 2015;521:436-444.

[4] Xu R, Peng J. A Comprehensive Survey of Deep Research: Systems, Methodologies, and Applications. Cornell University. 2025. ArXiv preprint: https://arxiv.org/pdf/2506.12594.

[5] Chen C, Liu J, Tan H, Li X, Wang KI, et al. Trustworthy Federated Learning: Privacy, Security, and Beyond. Knowl Inf Syst. 2024;67:2321-2356.

[6] Chen L, Xia C, Zhao Z, Fu H, Chen Y. Ai-Driven Sensing Technology: Review. Sensors. 2024;24:2958.

[7] Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, et al. Attention Is All You Need. Advances in neural information processing systems. 2017:30.

[8] Ali MA, Dornaika F. Agentic AI: A Comprehensive Survey of Architectures, Applications, and Future Directions. 2025. ArXiv preprint: https://arxiv.org/pdf/2510.25445

[9] Bengio Y. Deep Learning of Representations: Looking Forward. In International Conference on Statistical Language and Speech Processing. Berlin, Heidelberg: Springer Berlin Heidelberg. 2013:1-37.

[10] Hadjinikolis C, Siantos Y, Modgil S, Black E, McBurney P. Opponent Modelling in Persuasion Dialogues. In: Proceedings of the 23rd international joint conference on artificial intelligence (IJCAI). 2013:164-170.

[11] Azizov D, Manzoor MA, Bojković V, Wang Y, Wang Z, et al. A Decade of Deep Learning: A Survey on the Magnificent Seven. 2024. ArXiv preprint: https://arxiv.org/pdf/2412.16188

[12] Sharma SD, Leung SH, Viatte S. Genetics of Rheumatoid Arthritis. Best Pract Res Clin Rheumatol. 2024;38:101968.

[13] Stupple A, Singerman D, Celi LA. The Reproducibility Crisis in the Age of Digital Medicine. npj Digit Med. 2019;2.

[14] Indykov V, Strüber D, Wohlrab R. Architectural Tactics to Achieve Quality Attributes of Machine-Learning-Enabled Systems: A Systematic Literature Review. J Syst Softw. 2025;223:112373.

[15] Bengio Y, LeCun Y, Hinton GE. Deep Learning for AI. Commun ACM. 2021;64:58-65.

[16] François-Lavet V, Henderson P, Islam R, Bellemare MG, Joelle P. An Introduction to Deep Reinforcement Learning. Found Trends Mach Learn. 2018;11:219-354.

[17] Kingma DP, Ba J. Adam: A Method for Stochastic Optimization. 2014. ArXiv preprint: https://arxiv.org/pdf/1412.6980

[18] Hinton G, Krizhevsky A, Jaitly N, Tieleman T, Tang Y. Does the Brain Do Inverse Graphics. In Brain and Cognitive Sciences Fall Colloquium. 2012:2.

[19] Kirtas M, Passalis N, Tefas A. Multiplicative Update Rules for Accelerating Deep Learning Training and Increasing Robustness. Neurocomputing. 2024;576:127352.

[20] Santurkar S, Tsipras D, Ilyas A, Madry A. How Does Batch Normalization Help Optimization? NIPS'18: Proceedings of the 32nd International Conference on Neural Information Processing Systems. 2018;2488-2498.

[21] Ba J, Zhao J, Carcione JM, Huang X. Compressional Wave Dispersion Due to Rock Matrix Stiffening by Clay Squirt Flow. Geophys Res Lett. 2016;43:6186-6195.

[22] Ghosal S, Sengupta S, Majumder M, Sinha B. Linear Regression Analysis to Predict the Number of Deaths in India Due to Sars-Cov-2 at 6 Weeks From Day 0 (100 Cases—March 14th 2020). Diabetes Metab Syndr Clin Res Rev. 2020;14:311-315.

[23] Wu H, Zhang Z, Yue K, Zhang B, He J, et al. Dual-Regularized Matrix Factorization With Deep Neural Networks for Recommender Systems. Knowl-Based Syst. 2018;145:46-58.

[24] Mądry A, Makelov A, Schmidt L, Tsipras D, Vladu A. Towards Deep Learning Models Resistant to Adversarial Attacks. 2017. arXiv preprint: https://arxiv.org/pdf/1706.06083v1

[25] Reiter E. A Structured Review of the Validity Of BLEU. Comput Linguist. 2018;44:393-401.

[26] Krizhevsky A, Sutskever I, Hinton GE. Imagenet Classification With Deep Convolutional Neural Networks. Advances in neural information processing systems. 2012:25.

[27] Li Z, Liu F, Yang W, Peng S, Zhou J. A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects. IEEE Trans Neural Netw Learn Syst. 2022;33:6999-7019.

[28] LeCun Y, Bottou L, Bengio Y, Haffner P. Gradient-Based Learning Applied to Document Recognition. Proc IEEE. 1998;86:2278-2324. doi: 10.1109/5.726791.

[29] He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition. 2016:770-778.

[30] Cho K, Van Merriënboer B, Bahdanau D, Bengio Y. On the Properties of Neural Machine Translation: Encoder–Decoder Approaches. InProceedings of SSST-8, eighth workshop on syntax, semantics and structure in statistical translation. 2014:103-111. doi: 10.3115/v1/W14-4012.

[31] Devlin J, Chang MW, Lee K, Toutanova K. Bert: Pre-Training of Deep Bidirectional Transformers for Language Understanding. InProceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers). 2019:4171-4186.

[32] https://www.semanticscholar.org/paper/Improving-Language-Understanding-by-Generative-Radford-Narasimhan/cd18800a0fe0b668a1cc19f2ec95b5003d0a5035

[33] Berg RV, Kipf TN, Welling M. Graph Convolutional Matrix Completion. 2017. ArXiv preprint: https://arxiv.org/pdf/1706.02263.

[34] Veličković P, Fedus W, Hamilton WL, Liò P, Bengio Y, et al. Deep Graph Infomax. 2018. ArXiv preprint: https://arxiv.org/pdf/1809.10341.

[35] Salimans T, Goodfellow I, Zaremba W, Cheung V, Radford A, et al. Improved Techniques for Training GANs. Advances in Neural Information Processing Systems. 2016;29.

[36] Karras T, Laine S, Aila T. A Style-Based Generator Architecture for Generative Adversarial Networks. In2019 IEEE/CVF conference on computer vision and pattern recognition (CVPR). IEEE. 2019:4401-4410.

[37] Kingma DP, Welling M. Stochastic Gradient VB and the Variational Auto-Encoder. InSecond international conference on learning representations. ICLR. 2014;19:121.

[38] Higgins I, Matthey L, Pal A, Burgess C, Glorot X, et al. beta-VAE: Learning Basic Visual Concepts with a Constrained Variational Framework. In International conference on learning representations. ICLR. 2017.

[39] Ho J, Jain A, Abbeel P. Denoising Diffusion Probabilistic Models. Advances in Neural Information Processing Systems. 2020;33:6840-6851.

[40] Rombach R, Blattmann A, Lorenz D, Esser P, Ommer B. High-Resolution Image Synthesis With Latent Diffusion Models. In 2022 IEEE/CVF conference on computer vision and pattern recognition (CVPR). IEEE/CVF. 2022:10674-10685.

[41] Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, et al. Going Deeper With Convolutions. In Proceedings of the IEEE conference on computer vision and pattern recognition. 2015:1-9.

[42] Tan M, Le QV. Mixconv: Mixed Depthwise Convolutional Kernels. 2019. ArXiv preprint: https://arxiv.org/pdf/1907.09595

[43] Liu B, Zheng D, Zhou S, Chen L, Yang J. Vfdb 2022: A General Classification Scheme for Bacterial Virulence Factors. Nucleic Acids Res. 2022;50:D912-D917.

[44] Nauen TC, Palacio S, Dengel A. Which Transformer to Favor: A Comparative Analysis of Efficiency in Vision Transformers. 2023. ArXiv preprint: https://arxiv.org/pdf/2308.09372v1

[45] Siebra CA, Kurpicz-Briki M, Wac K. Transformer in Health: A Systematic Review on Architectures for Longitudinal Data Analysis. Artif Intell Rev. 2024;57:32.

[46] Tiezzi M, Casoni M, Betti A, Gori M, Melacci S. State-Space Modeling in Long Sequence Processing: A Survey on Recurrence in the Transformer Era. Neural Netw. 2026;193:108039.

[47] Alomar K, Aysel HI, Cai X. RNNs, Cnns and Transformer in Human Action Recognition: A Survey and a Hybrid Model. 2024. ArXiv preprint: https://arxiv.org/pdf/2407.06162

[48] Wolf T, Debut L, Sanh V, Chaumond J, Delangue C, et al. Transformer: State-Of-The-Art Natural Language Processing. InProceedings of the 2020 conference on empirical methods in natural language processing: system demonstrations. 2020:38-45.

[49] Liu Z, Lin Y, Cao Y, Hu H, Wei Y, et al. Swin Transformer: Hierarchical Vision Transformer Using Shifted Windows. In 2021 IEEE/CVF international conference on computer vision (ICCV). IEEE. 2021:9992-10002.

[50] Yang X, Chen A, PourNejatian N, Shin HC, Smith KE, et al. A Large Language Model for Electronic Health Records. npj Digit Med. 2022;5:194.

[51] Dwivedi VP, Kanatsoulis CI, Huang S, Leskovec J. Relational Deep Learning: Challenges, Foundations and Next-Generation Architectures. In Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V. 2. 2025:5999-6009.

[52] Niepert M. Discriminative Gaifman Models. Advances in Neural Information Processing Systems. NIPS. 2016;29.

[53] Defferrard M, Bresson X, Vandergheynst P. Convolutional Neural Networks on Graphs With Fast Localized Spectral Filtering. Advances in Neural Information Processing Systems. NIPS. 2016;29.

[54] Veličković P. Everything Is Connected: Graph Neural Networks. Curr Opin Struct Biol. 2023;79:102538.

[55] Zhang O, Lin H, Zhang X, Wang X, Wu Z, et al. Graph Neural Networks in Modern Ai-Aided Drug Discovery. Chem Rev. 2025;125:10001-10103.

[56] Zhang R, Wu X, Li L, Zhao P, Zhang Q, Wuri L et al. A Bibliometric Review of Deep Learning in Crop Monitoring: Trends, Challenges, and Future Perspectives. Front Artif Intell. 2025;8:1636898.

[57] Gulrajani I, Ahmed F, Arjovsky M, Dumoulin V, Courville A. Improved Training of Wasserstein Gan. 2022. ArXiv preprint: https://arxiv.org/pdf/1704.00028

[58] Zietlow D, Rolínek M, Martius G. Demystifying Inductive Biases for (Beta-)Vae Based Architectures. International Conference on Machine Learning. In International Conference on Machine Learning. PMLR. 2021:12945-12954.

[59] Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, et al. Generative Adversarial Nets. Adv Neural Inf Process Syst. Advances in Neural Information Processing Systems. 2014;27.

[60] Hadiyya K, Dina R, Mokhtar B. Diffusion Models in Generative AI: Principles, Applications, and Future Directions. 2025. Available at: https://www.preprints.org/manuscript/202502.0524

[61] Garcez AS, Lamb LC. Neurosymbolic AI: The 3rd Wave. Artif Intell Rev. 2023;56:12387-12406.

[62] Meuser T, Lovén L, Bhuyan M, Patil SG, Dustdar S, et al. Revisiting Edge AI: Opportunities and Challenges. IEEE Internet Comput. 2024;28:49-59.

[63] Kumar RR, Shankar SV, Jaiswal R, Ray M, Budhlakoti N, et al. Advances in Deep Learning for Medical Image Analysis: A Comprehensive Investigation. J Stat Theory Pract. 2025;19.

[64] Li H, Zhao Y, Mao Z, Qin Y, Xiao Z, et al. Graph Neural Networks in Intelligent Transportation Systems: Advances, Applications and Trends. 2024. ArXiv preprint: https://arxiv.org/pdf/2401.00713

[65] Li X, Ding H, Yuan H, Zhang W, Pang J, et al. Transformer-Based Visual Segmentation: A Survey. IEEE Trans Pattern Anal Mach Intell. 2024;46:10138-10163.

[66] Liu A. Anticipating Driving Behavior Through Deep Learning-Based Policy Prediction. 2023. arXiv preprint: https://arxiv.org/pdf/2307.11058

[67] Liu Z, Lv Q, Yang Z, Li Y, Lee CH, et al. Recent Progress in Transformer-Based Medical Image Analysis. Comput Biol Med. 2023;164:107268.

[68] Deng L, Liu Y. Deep Learning in Natural Language Processing. Springer. 2018.

[69] Han L, Gladkoff S, Erofeev G, Sorokina I, Galiano B, et al. Neural Machine Translation of Clinical Text: An Empirical Investigation Into Multilingual Pre-Trained Language Models and Transfer-Learning. Front Digit Health. 2024;6:1211564.

[70] Karamouzas D, Mademlis I, Pitas I. Public Opinion Monitoring Through Collective Semantic Analysis of Tweets. Soc Netw Anal Min. 2022;12:91.

[71] Qiu J, Li L, Sun J, Peng J, Shi P, et al. Large AI Models in Health Informatics: Applications, Challenges, and the Future. 2023. arXiv preprint: https://arxiv.org/pdf/2303.11568

[72] Diehl PU, Zilly H, Sattler F, Singer Y, Kepp K, et al. Deep Learning-Based Denoising Streamed From Mobile Phones Improves Speech-In-Noise Understanding for Hearing Aid Users. Front Med Eng. 2023;1:1281904.

[73] Prabhavalkar R, Hori T, Sainath TN, Schlüter R, Watanabe S. End-To-End Speech Recognition: A Survey. IEEE/ACM Trans Audio Speech Lang Process. 2024;32:325-351.

[74] Van Den Oord A, Dieleman S, Zen H, Simonyan K, Vinyals O, et al. Wavenet: A Generative Model for Raw Audio. 2016. arXiv preprint: https://arxiv.org/pdf/1609.03499

[75] Anastassiou P, Chen J, Chen J, Chen Y, Chen Z, et al. Seed-Tts: A Family of High-Quality Versatile Speech Generation Models. 2024. arXiv preprint: https://arxiv.org/pdf/2406.02430

[76] Yang G, Yang C, Chen Q, Ma Z, Chen W, et al. Emovoice: Llm-Based Emotional Text-To-Speech Model With Freestyle Text Prompting. 2025. arXiv preprint: https://arxiv.org/pdf/2504.12867

[77] Nie G, Zhao Q, Tang G, Li Y, Hong S. Artificial Intelligence-Derived Photoplethysmography Age as a Digital Biomarker for Cardiovascular Health. Commun Med. 2025;5:481.

[78] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In International Conference on Medical image computing and computer-assisted intervention. Cham: Springer international publishing. 2015:234-241.

[79] Choi E, Bahadori MT, Kulas JA, Schuetz A, Stewart WF, et al. Retain: An Interpretable Predictive Model for Healthcare Using Reverse Time Attention Mechanism. 2016. arXiv preprint: https://arxiv.org/pdf/1608.05745v1

[80] Generalized AI Models for Genomics Applications. Nat Methods. 2025;22:231-232.

[81] Pazmiño Ortiz LA, Maldonado Soliz IF, Guevara Balarezo VK. Advancing Tinyml in Iot: A Holistic System-Level Perspective for Resource-Constrained AI. Preprints. org. 2025;17:257.

[82] Mugalakhod MS, Miss Nirmanik MS. AI Driven Smart Homes Energy Efficiency and Model. Int J Res Appl Sci Eng Technol. 2022;10:461-465.

[83] Yan P, Abdulkadir A, Luley PP, Rosenthal M, Schatte GA, et al. A Comprehensive Survey of Deep Transfer Learning for Anomaly Detection in Industrial Time Series: Methods, Applications, and Directions. IEEE Access. 2024;12:3768-3789.

[84] Dua R, Saxena N, Agarwal A, Wilson A, Singh G, et al. Agricultural Landscape Understanding at Country-Scale. 2024. ArXiv preprint: https://arxiv.org/pdf/2411.05359v1.

[85] Barnetson J, Phinn S, Scarth P. Climate-Resilient Grazing in the Pastures of Queensland: An Integrated Remotely Piloted Aircraft System and Satellite-Based Deep-Learning Method for Estimating Pasture Yield. AgriEngineering. 2021;3:681-702.

[86] Shi Y, Zhu J, Jiang C, Kuang L, Letaief KB. Satellite Edge Artificial Intelligence With Large Models: Architectures and Technologies. Sci China Inf Sci. 2025;68:170302.

[87] R V, Alazab M, Kp S, Srinivasan S, Venkatraman S, et al. Deep Learning for Cyber Security Applications: A Comprehensive Survey. TechRxiv. 2021;10.

[88] Vaisman Y, Katz G, Elovici Y, Shabtai A. Detecting Anomalous Network Communication Patterns Using Graph Convolutional Network. 2023. ArXiv preprint: https://arxiv.org/pdf/2311.18525.

[89] Lou C, Wang Y, Li J, Qian Y, Li X. Graph Neural Network for Fraud Detection via Context Encoding and Adaptive Aggregation. Expert Syst Appl. 2025;261:125473.

[90] Quertier T, Marais B, Barrué G, Morucci S, Azé S, et al. A Lean Transformer Model for Dynamic Malware Analysis and Detection. 2024. ArXiv preprint: https://arxiv.org/pdf/2408.02313.

[91] Huang SC, Pareek A, Jensen M, Lungren MP, Yeung S, et al. Self-Supervised Learning for Medical Image Classification: A Systematic Review and Implementation Guidelines. npj Digit Med. 2023;6:74.

[92] Schwartz R, Dodge J, Smith NA, Etzioni O. Green AI. Commun ACM. 2020;63:54-63. doi: 10.1145/3381831.

[93] Li X, Xiong H, Li X, Wu X, Zhang X, et al. Interpretable Deep Learning: Interpretation, Interpretability, Trustworthiness, and Beyond. Knowl Inf Syst. 2022;64:3197-3234.

[94] Horvitz E, Conitzer V, McIlraith SA, Stone P. Now, Later, and Lasting: Ten Priorities for AI Research, Policy, and Practice. Cornell University. 2024. ArXiv preprint: https://arxiv.org/pdf/2404.04750.

[95] Chen PY, Liu S. Holistic Adversarial Robustness of Deep Learning Models. InProceedings of the AAAI Conference on Artificial Intelligence. 2023;37:15411-15420.

[96] Baltensperger J, Salza P, Gall HC. Continuous Deep Learning: A Workflow to Bring Models Into Production. 2022. ArXiv preprint: https://arxiv.org/pdf/2208.12308.

[97] Zschech P, Weinzierl S, Kraus M. Inherently Interpretable Machine Learning: A Contrasting Paradigm to Post-Hoc Explainable AI. Bus Inf Syst Eng. 2025;68:445-463.

[98] Geng X, Wang Z, Chen C, Xu Q, Xu K, et al. From Algorithm to Hardware: A Survey on Efficient and Safe Deployment of Deep Neural Networks. IEEE Trans Neural Netw Learn Syst. 2025;36:5837-5857.

[99] Liang PP, Zadeh A, Morency L. Foundations and Trends in Multimodal Machine Learning: Principles, Challenges, and Open Questions. 2022. ArXiv preprint: https://arxiv.org/pdf/2209.03430.

[100] Huisman M, van Rijn JN van, Plaat A. A Survey of Deep Meta-Learning. Artif Intell Rev. 2021;54:4483-4541.

[101] Spera C, Agrawal G. Reversing the Paradigm: Building Ai-First Systems With Human Guidance. 2025. ArXiv preprint: https://arxiv.org/pdf/2506.12245.

[102] Barrett C, Boyd B, Bursztein E, Carlini N, Chen B, et al. Identifying and Mitigating the Security Risks of Generative AI. Foundations Trends Priv Secur. 2023;6:1-52. doi: 10.1561/3300000041.

[103] Duraisamy K. Active Inference AI Systems for Scientific Discovery. 2025. arXiv preprint: https://arxiv.org/pdf/2506.21329.

9. List of Acronyms

Acronym	Full Form
AI	Artificial Intelligence
ML	Machine Learning
DL	Deep Learning
ANN	Artificial Neural Network
CNN	Convolutional Neural Network
RNN	Recurrent Neural Network
LSTM	Long Short-Term Memory
GRU	Gated Recurrent Unit
GNN	Graph Neural Network
GAN	Generative Adversarial Network
VAE	Variational Autoencoder
NLP	Natural Language Processing
ASR	Automatic Speech Recognition
TTS	Text-To-Speech
IoT	Internet of Things
XAI	Explainable AI
AGI	Artificial General Intelligence
LLM	Large Language Model
ViT	Vision Transformer
GAT	Graph Attention Network
GCN	Graph Convolutional Network
MPNN	Message-Passing Neural Network
DDPM	Denoising Diffusion Probabilistic Model
SGD	Stochastic Gradient Descent
GPU	Graphics Processing Unit
TPU	Tensor Processing Unit
EHR	Electronic Health Record
CV	Computer Vision
mAP	mean Average Precision
IoU	Intersection over Union
AP	Average Precision
BLEU	Bilingual Evaluation Understudy
ROUGE	Recall-Oriented Understudy for Gisting Evaluation
AUC	Area Under the Curve
NDCG	Normalized Discounted Cumulative Gain
TCN	Temporal Convolutional Network
TinyML	Tiny Machine Learning
SVM	Support Vector Machine
DT	Decision Tree