Exploring the Depths: A Comprehensive Survey of Neural Networks and their Applications

Harshit Gupta*, Manya Verma, Anushka Arora, Dhriti, Gyanender Kumar

Department of Computer Science, HMR Institute of Technology & Management, Hamidpur, 110033, Delhi, India

* Corresponding author

doi: https://doi.org/10.21467/proceedings.178.6

ABSTRACT

This paper reviews the history and use of neural networks technology in great details. It starts by demonstrating the range of tasks neural networks can tackle, which are not limited to structural engineering or medical imaging. Then, it proceeds to outline the principles of artificial neural network models such as feedforward, recurrent and convolutional neural networks. The paper focuses on the training of neural networks and provides an account of the development of the backpropagation algorithm, optimization strategies, and regularization approaches. Moreover, a variety of case studies of neural networks applications are presented, for example, in climate modelling, natural language and speech recognition, computer vision and defect diagnosis techniques. In addition, the paper highlights the impact of the latest transformer models on self-supervised learning in natural language processing. The last part tackles ethical issues that are related to the use of neural networks such as fairness, bias and interpretability. All in all, this survey witnesses the growing importance of neural networks as a tool for enhancing the computational capabilities and addressing difficult problems.

Keywords: Neural Networks, Machine Learning, Computer Vision, Natural Language Processing, Artificial Intelligence

1 Introduction

Neural networks are advanced models of computation that imitate the biological functioning of the human which are used in a number of activities including building construction, character recognition, and in the field of medicine. The models resemble artificial ones connected with layers, but in this case, the model is a unit of a neuron, so these units are interlinked, allowing the processing of various types of conversion on the network various nodes. Each module, or neuron, makes a quick computation and passes the output to the next tier. Neural networks enable the identification of intricate relationships in the data. They have been effectively applied to complex problems including images, videos, speech recognition and synthesis, language translation, voice generation as well as predictive analysis and decision making. The neural network approach was first provided in the framework of the structural engineering applications, indicating also the possibilities of its application [1]. The applications of neural networks in the field of pattern classification were elucidated, particularly in relation to the design as having shadowing capabilities [2]. It was also possible to test other aspects of neural networks on the problems of statistical dependence analysis, decision making modelling, and forecasting where nonlinear relationships were discerned [3]. The application of neural network in building services engineering provided an appreciation of its contribution as a design tool to other disciplines [4]. The use of neural networks was further enhanced to include mathematical modelling, thus increasing their scope even further. Both feedforward and recurrent networks were studied, explaining the backpropagation algorithm integrational through the prospects of deep neural networks in brain computation modelling [5]. To summarize, the studies of neural networks represent varied disciplines, which add to their range of application. Neural networks have from structural engineering to medical diagnostics proven useful for dealing with complex relationships and challenging problems. Perhaps an

Proceedings DOI: 10.21467/proceedings.178; Series: AIJR Proceedings; ISSN: 2582-3922; ISBN: 978-81-984081-8-1

appropriate image of the structure of the feedforward neural network would make these potentials clearer [6] .

2 Artificial Neural Networks

Artificial neural networks artificial neural networks are automated systems that consist of layers of interconnected nodes or neurons, which receive an input, then there is one or more hidden layers, which carry some operations on the input and finally, there is an output neuron which provides the system response [7]. The successive advances achieved by artificial neural networks in numerous domains are attributed to their capacity to learn and develop based on the data. Learning techniques which are applied in the training of ANNs are statistical based, therefore, there is significance of statistical theory to the different learning approaches of the networks [8]. It has been established that Neural networks have been used in the field of industrial control systems, particularly those that combined neural networks fuzzy logic artificial intelligence hybrid systems [9]. With respect to econometrics and a particular lean towards ANN application, artificial neural networks have been applied for control, prediction and identification which underscores the significance of the applications [10].In clinical settings, the potential of ANN to solve more complicated issues has been a work in progress. In this context, several of its advancements and specifications have been introduced and supervised feed forward ANNs training processes have been discussed in great detail [11]. Studies and explanations of fundamental ANN constructs have been conducted in order to build cognition about the models [12]. Recurrent neural networks are introduced, and their features of recursive gravitation and activation delay are emphasized [13]. In the field of building services engineering, ANNs have also found usefulness in conducting energy assessments in buildings [14]. To address the scalability issue in the training of ANNs, sparse evolutionary training strategies have been introduced in which sparse topologies are utilized during learner training, thereby enabling the use of deep learning in different fields [15].

Table 1: *Key Findings on Artificial Neural Networks.*

S.no	Reference	Year	Key Findings
1	White, 1989	1989	Statistical techniques for training ANNs
2	Fukuda et al., 1992	1992	ANNs in industrial control systems with fuzzy sets and AI
3	Kuan et al., 1994	1994	Econometric perspective on ANNs
4	Haykin, 1994	1994	ANNs as computational models with interconnected nodes
5	Xing et al., 1995	1995	ANNs for identification, prediction, and control of weather
6	Porto et al., 2001	2001	Clinical applications of evolving ANNs
7	Hammer, 2001	2001	Supervised learning in feedforward ANNs
8	Krogh, 2008	2008	Fundamental questions about ANNs
9	Marhon et al., 2013	2013	Recurrent neural networks with recursive dynamics
10	Kumar et al., 2013	2013	ANNs in building energy analysis
11	Mocanu et al., 2018	2018	Sparse evolutionary training of ANNs for deep learning

2.1 Feedforward Neural Networks

Feedforward neural networks have been extensively studied and applied in various fields [15]. The bias/variance dilemma in neural networks trained by error backpropagation has been analysed, emphasizing the statistical strengths and weaknesses of these models [16]. A discrete affine wavelet transformation representation for feedforward neural networks was developed, enabling both analysis and synthesis [17].

Neural networks have been explored for tasks such as identification, prediction, and control [18]. Methodologies for feedforward neural networks have been proposed, with considerations on learning capabilities and storage capacity, including the reduction of required hidden neurons in two hidden layer networks [19].The extreme learning machine (ELM) was introduced as a new learning scheme to improve learning speed by randomly choosing input weights and analytically determining output weights [20].Comparisons between the generalization characteristics of complex valued and real valued feedforward neural networks have been conducted, particularly for applications like radar and coherent imaging systems [21]. Feedforward and recurrent neural network language models were evaluated for speech recognition, highlighting differences in context dependencies [22]. Research has demonstrated that increasing depth in feedforward neural networks can be exponentially more effective than increasing width for approximating specific functions. Neural networks have shown promise as substitutes for conventional mesh based approaches, approximating solutions to partial differential equations in complex geometries [23].

Table 2: *Feedforward Neural Network Research Summary.*

S. No	Research Paper Title	Author(s)	Year	Key Findings
1	Bias/Variance Dilemma in Neural Networks Trained by Error Backpropagation	Geman et al.	1992	Statistical viewpoint on bias/variance trade off in neural networks
2	Representation of Feedforward Neural Networks Using Discrete Affine Wavelet Transformations	Pati et al.	1993	Method for analyzing and synthesizing feedforward neural networks using wavelets
3	Use of Neural Networks for Identification, Prediction, and Control Tasks	Xing et al.	1995	Exploration of neural networks for identification, prediction, and control tasks
4	Methodology for Feedforward Neural Networks	Fine	1999	Methodology for working with feedforward neural networks
5	Learning Capability and Storage Capacity of Two Hidden Layer Feedforward Networks	Huang	2003	Reduced hidden neuron requirement for two hidden layer feedforward networks
6	Introduction of Extreme Learning Machine for Feedforward Neural Networks	Huang et al.	2004	Extreme learning machine (ELM) for faster learning in feedforward neural networks
7	Generalization Characteristics of Complex Valued and Real Valued Feedforward Neural Networks	Hirose et al.	2012	Comparison of complex valued and real valued neural networks for signal coherence
8	Comparison of Feedforward and Recurrent Neural Network Language Models for Speech Recognition	Sundermeyer et al.	2013	Differences in context dependencies between feedforward and recurrent neural networks for speech recognition
9	Power of Depth for Feedforward Neural Networks	Eldan et al.	2016	The value of depth in feedforward neural networks for approximating certain functions
10	Utilization of Deep Feedforward Artificial Neural Networks for Approximating Solutions to PDEs	Berg et al.	2018	Deep feedforward neural networks for approximating solutions to partial differential equations

2.2 Recurrent Neural Networks

Recurrent Neural Networks (RNN) are made to remember past observations, which is particularly used for sequential data, such as natural language or time series, as it has feedback connections. Long Short Term

Memory (LSTM) and Gated recurrent unit (GRU) are commonly used variants of RNNs that deal with the vanishing gradient problem [24]. Since RNNs are able to identify temporal relations in sequential data, they have been researched and used in a number of domains. The robust learning algorithm of RNNs turning into NARMA (p,q) networks, has consistently outperformed feed forward neural networks in time series forecasting applications [25].The learning algorithms for recurrent neural networks with hidden units have been investigated including the advantages and disadvantages of temporal and non temporal based neural networks [26]. The prediction tasks, through the application of RNN networks, confirm the usefulness of this type of architecture for sequential data [27]. The Long Short Term Memory (LSTM) architecture was proposed in order to overcome the issues of training the traditional RNNs, providing better performance for language modelling tasks. Due to the existence of internal loops in RNNs some recursive dynamics are present, which are important for learning the sequential patterns. Stability analysis in the continuous time RNNs has been studied comprehensively concentrating on models including Hopfield and Cohen Grossberg neural networks [28]. In a similar modelling effort, deep RNNs have been used to comprehend tasks such as video translation to natural language and this shows their practicality for sequential data processing across different modalities [29]. Another approach, hierarchical multiscale RNNs, emerged as a solution to two well known problems in RNNs studies: hierarchical and temporal encoding of the sequence [30]. The functionality of attentive RNNs has also been tested in other tasks such as abstractive sentence summarization and joint event extraction [31]. Such works extend the idea of establishing long range dependencies in sequential data and fills in further reasons for employing RNNs for various NLP tasks.

Table 3: *Research on Recurrent Neural Networks.*

S. No	Research Paper	Author(s)	Year	Key Findings
1	A Robust Learning Algorithm for RNNs	Connor et al.	1994	Introduced robust learning for RNNs, leading to NARMA networks for time series prediction.
2	Survey of Learning Algorithms for RNNs	Pearlmutter	1995	Surveyed learning algorithms for RNNs with hidden units.
3	Application of RNNs for Prediction Tasks	Mandic et al.	2001	Highlighted RNN potential in modeling sequential data for prediction.
4	Introduction to LSTM Architecture in RNNs	Sundermeyer et al.	2012	Introduced LSTMs to address training challenges in RNNs for language modeling.
5	Recursive Dynamics in RNNs	Marhon et al.	2013	Introduction to RNNs for those familiar with ANNs, emphasizing internal loops.
6	Stability Analysis in Continuous time RNNs	Zhang et al.	2014	Reviewed stability analysis in continuous time RNNs (Hopfield, Cohen Grossberg networks).
7	Application of Deep RNNs in Translating Videos	Venugopalan et al.	2015	Explored deep RNNs for translating videos to natural language.
8	Hierarchical Multiscale RNNs for Temporal Representations	Chung et al.	2016	Proposed hierarchical multiscale RNNs to capture hierarchical and temporal representations.
9	Effectiveness of Attentive RNNs in Natural Language Processing	Chopra et al.	2016	Demonstrated effectiveness of attentive RNNs in abstractive summarization and event extraction.

2.3 Convolutional Neural Networks

Due to the presence of pooling and convolutional layers that utilize local spatial patterns, Convolutional neural networks are suitable for images and videos as grid like structures [32]. CNNs are said to perform state of the art results in notable computer vision tasks including image analysis, object recognition, and semantic image segmentation. Many studies regarding the use of CNNs have been conducted recently with a lot of studies reporting on their multifunctional nature. CNNs training has also been reported to be parallelized across different GPUs to enhance the training process [33]. Sentence classification tasks have also been successfully tackled by performing CNNs [34]. 3D CNNs have been employed in studies for the prediction of Alzheimer s disease and have produced impressive outcomes in neuroimaging research [35].It has been established through extensive literature that CNN Inception V3 architecture is used efficiently across deep learning networks [36]. Orders of words have been cleverly captured through the use of CNNs for text classification tasks [37]. Different forms of CNNs have been employed for the relation perspective extraction tasks of the more extensive natural language processing systems[38].Dynamic multi pooling CNNs have reportedly been effective in event extraction thus demonstrating that CNN architecture is suited for a variety of tasks [39].The training of deep CNN architectures on large datasets has been a computationally intensive task, yet it is such endeavours that have led to the creation of accelerated training optimization algorithms [40]. In addition, feeding a 4D spatio temporal convolutional neural network has been reported to be more resistant to noise and was more efficient than the traditional 3D CNN when utilized for spatio temporal perception tasks [41]. Thus, it is apparent that CNNs possess a wide range of uses and offer robust solutions including in the fields of image and natural language processing, as well as spatio temporal perception. Research on CNN s capabilities still is expanding the boundaries of their performance.

3 Training Neural Networks

3.1 Backpropagation

Backpropagation stands out as potentially the most popular approach used for training neural networks whereby weights and biases in inter neuron connections are altered. Achieving this is based upon the difference between what is expected and what is intended [42]. Deeply staged networks are trained efficiently through this process which has also sparked the interest of many researchers for decades. A novel momentum step along with a dynamic scheme for momentum rate selection was developed to speed up the training of backward propagation networks [43]. However, as operational procedures show, backpropagation is often stuck in local minima as it is too sensitive to changes in network parameters prompting the search for alternative approaches including genetic algorithms [44]. A combination of genetic algorithms and backpropagation was synthesized, which was found to result in less training time than using either backpropagation or SOFT algorithm alone [45].Training algorithms based on backpropagation and intended for specific network architectures have also been improved, mentioning as well their evolution in the past [46].The problem of generalization for advanced spiking neural networks was tackled by developing a new spatio temporal backpropagation technique allowing supervised training of such networks [47]. The issue of backpropagation in through the optical neural networks nonlinear units was solved by using saturable absorption as the means of implements force all optical training [48]. A Spike Train level RSNNs Backpropagation algorithm was created for the purposes of training deep recurrent spiking neural networks which in turn enhance the computational efficiency for RSNNs [49]. Another algorithm for physics aware training has been proposed where in situ in silico methods were conjugated with backpropagation so that manipulable physical entities can be trained thus showing the benefits of backpropagation as a teaching tool for large neural networks [50].

Table 4: *Backpropagation and its Advancements in Neural Network Training.*

Research Paper Title	Author(s)	Year	Description
Backpropagation Algorithm	Rumelhart, Hinton, & Williams	1986	Fundamental method for training deep neural networks.
Momentum Step	Qiu et al.	1992	Introduced to accelerate backpropagation training.
Genetic Algorithms for Training	Siddique et al.	2001	Explored genetic algorithms as alternatives to backpropagation.
Hybrid Technique with Genetic Algorithms	Adawy et al.	2002	Combined genetic algorithms with backpropagation for faster training.
Advancements in Backpropagation	Huang	2009	Discussed advancements in backpropagation for various architectures.
Spatio temporal Backpropagation	Wu et al.	2018	Introduced for training high performance spiking neural networks.
All Optical Training Solution	Guo et al.	2019	Proposed a solution for all optical training of optical neural networks.
Spike Train Level RSNNs Backpropagation	Zhang et al.	2019	Presented for training deep recurrent spiking neural networks.
Physics Aware Training Algorithm	Wright et al.	2022	Introduced for training controllable physical systems.

3.2 Optimization Algorithms

To reduce the loss function and adjust the parameters of the neural networks in the course of training, a number of minimization optimization, such as SGD together with such derivatives as Adam and RMSProp, are put into practice. Learning a neural network is one of the most important steps in supervised learning as it is the process of establishing the most effective weight arrangement across training. This process has also been emphasized by various studies which have proposed optimization algorithms to improve the process and performance of neural networks. It has been found that the application of a new global minimization algorithm NOVEL allows constructing networks more compact and less prone to errors, achieving best results in training of neural networks [51].The Artificial Bee Colony Optimization Algorithm has been used in a training process of feed forward neural networks due to its explorative ability for optimal set of weights during training [52]. An Ant Colony Optimization Algorithm was realized for a number of continuous optimization problems including the training of feed forward neural networks [53]. The outcome showed that the proposed algorithms are robust when compared to gradient based strategies and perform better than standard genetic algorithms. There was a comparison of five algorithms namely BA, GA, PSO, BP and LM which were applied for the training of feed forward neural networks in an e learning domain based on a benchmark dataset in order to offer results that would help in the future comparisons of these algorithms [54]. A method of quantum learning of neural networks that uses low dimensional circuits was described, which used the quantum approximation optimization method to sample low energy distributions of Ising Hamiltonians to train neural networks [55]. There have also been attempts to use adaptive schemes to optimize neural networks. It was modified the Adam optimizer in order to have non-existent generalization gap in some cases, thus, increasing the usefulness of deep neural networks with it [56]. Also, a training algorithm based on the Whale Optimization Algorithm was reported for the optimization of connection weights in the neural network [57]. This algorithm worked to solve the concerns of local optimum stagnation and low convergence speed which are associated with standard training algorithms. Neural networks have also been studied to assist in the quantum learning stage which is referred to as meta learning to search for possible optimal values in the sphere of quantum variational algorithms [58].In general, the optimization of neural networks comprises various algorithms and techniques; all with different advantages and limitations. More investigations in this field are also necessary in order to improve the training process and performance of neural networks.

Table 5: *Alternative Optimization Methods for Training Neural Networks.*

S. No	**Research Paper Title**	**Author(s)**	**Year**	**Description**
1	Global Minimization Method (NOVEL)	Shang et al.	1996	Introduced NOVEL for training neural networks, achieving smaller errors.
2	Artificial Bee Colony Optimization	Karaboga et al.	2007	Utilized an algorithm for optimal weight training in feed forward networks.
3	Ant Colony Optimization Algorithm	Socha et al.	2007	Applied an algorithm for training feed forward networks, comparable to gradient based methods.
4	Comparison of Training Algorithms	Khan et al.	2012	Compared five training algorithms for neural networks in an e learning context.
5	Quantum Algorithm for Training	Verdon et al.	2017	Introduced a quantum method for training neural networks with low dimensional circuits.
6	Improved Adam Optimizer	Zhang	2018	Improved the Adam optimizer to address the generalization gap in deep neural networks.
7	Whale Optimization Algorithm	Aljarah et al.	2018	Presented a training algorithm to optimize connection weights in neural networks.
8	Meta Learning with Neural Networks	Verdon et al.	2019	Investigated using traditional neural networks to aid in quantum meta learning.

3.3 Regularization Techniques

Overfitting of neural networks can be lessened and learning improved through a number of techniques, such as dropout, L1 by L2 regularization, and early stopping. Regularization methods substantially improve performance and mitigate overfitting risk when training neural networks [59]. It appears that the application of RL DNNs administered with tied scalar regularization was advantageous in LVCSR applications [60]. The importance of different regularization techniques in addressing the overfitting problem when CNNs are used to model facial emotions has been proven [61].To speed up computations using deep neural networks, Force Regularization was proposed as a means of synchronizing filters in DNNs [62]. A Generative Adversarial Trainer was designed to secure neural networks against adversarial perturbations, and to alleviate overfitting [63]. In particular, deep neural networks were trained with dropout regularization while accounting for label noise to increase robustness [64,65]. The distance transform regression method was used in developing a new segmentation focused regularization technique [66].The acquisition of effective image super resolution networks was achieved using structured pruning to enhance performance and efficiency. Deep convolutional neural networks with L2 regularization were proposed for fire detection, highlighting the diverse applications and benefits of regularization techniques in training neural networks.

Table 6: *Regularization Techniques for Neural Networks.*

S. No	Research Paper Title	Author(s)	Year	Description
1	Combining NN with PCA for Concrete Properties	Boukhatem et al.	2012	Combined Neural Networks (NN) with PCA to effectively predict concrete properties.
2	RL DNNs for LVCSR	Zhang et al.	2015	Highlighted benefits of rectified linear deep neural networks (RL DNNs) with regularization for speech recognition.
3	Regularization in CNNs for Facial Expressions	Hinz et al.	2016	Emphasized the importance of regularization in CNNs to reduce overfitting when learning facial expressions.
4	Force Regularization for Faster DNN Computation	Wen et al.	2017	Introduced Force Regularization for faster computation in Deep Neural Networks (DNNs).
5	Generative Adversarial Trainer for Overfitting Reduction	Lee et al.	2020	Proposed a method to reduce overfitting in neural networks using a Generative Adversarial Trainer.
6	Dropout Regularization for Label Noise in DNNs	Jindal et al.	2016	Discussed using dropout regularization to handle label noise during DNN training.
7	d SNE for Domain Adaptation in Neural Networks	Xu et al.	2019	Introduced d SNE for domain adaptation using embedding techniques in neural networks.
8	Semantic Segmentation Regularization	Audebert et al.	2019	Presented a novel distance transform based regularization method for semantic segmentation.

3.4 Applications and Future Directions

Different domains including computer vision, speech production, and recognition, natural language processing, and robotic control have effectively applied neural networks [67].Some structural approaches, such as generative adversarial networks (GANs) [68], attention mechanisms [69], and transformer models [70], are becoming popular considering the rise in computing power as well as the availability of data. Computer processing tasks for training and deploying large scale neural networks have been made easier through the adoption of hardware accelerators like graphics processing units and tensor processing units

[71,72].Also, targeted efforts are dealing with the issues of explainable, robust, and less resource intensive neural networks and their impact on ethical and social issues [73,74]. With regard to several works, it is worth noting the interesting applications of CNNs in different domains. For example, one study demonstrated the capability of CNNs to predict where a cell would move to in the future given its morphological shape, thus pointing out the possibilities of automated morphological assessment automation of factors responsible for movement of cells [75]. Individual patient data was combined with DANN using the inter net of things and artificial intelligence to forecast health care of the patients in an improved manner in the oyer studied networks [76]. Besides, in healthcare as well, the employment of artificial neural networks was studied, as exemplified in the modelling design of a watch with media networks community, Numeric Modelling of Community Network Created by Technology Design Exploratorium. The potential of neural stem cell based therapies and regeneration after stroke with functional recovery that would in turn facilitate stroke recovery research and its clinical application have been systematically reviewed [77]. CNNs are also being used in the classification of plants diseases using deep learning technologies [78]. There was a proposal of an enhanced deep convolutional neural network using SSD and multi level modular converters to design a fault diagnostic method for high voltage direct current transmission systems [79]. There have been discussions regarding issues and possible solutions for further studying deep learning and other elementary learning models for remote sensing of soil organic carbon measurement [80]. The vital role of the deep learning approaches, quite a number, has been stated in the prediction and design of protein structures [81]. A brief survey of deep learning using graph convolutional networks has been provided which discusses its relevance in computational intelligences and the communication networks as a field of future research [82]. In broad strokes, it appears that the literature demonstrates a rather broader possibilities of the application of the neural networks in general and CNN in particular in such areas as healthcare, plant disease recognition, fault detection, remote sensing, protein structure prediction, computational intelligence etc. Such studies enable one to focus on cutting edge applications and also future prospects of neural networks.

Table 7: *Recent Advancements in AI*

S. No	Research Paper Title	Author(s)	Year	Description
1	CNNs for Cell Movement Prediction	Nishimoto et al.	2019	CNNs predict cell movement for morphology analysis
2	AI for Patient Health Analysis	Fouad et al.	2020	AI analyzes patient data (IoT sensors) for improved future assistance.
3	Neural Stem Cell Treatment for Recovery from Strokes	Jiao et al.	2021	Examines developments in neural stem cell treatment for stroke rehabilitation
4	Deep Learning in Plant Disease Classification	Lu et al.	2021	Deep learning (CNNs) gaining traction in plant disease classification
5	High Voltage Transmission Engineering: Fault Diagnosis	Ke et al.	2021	Neural networks for high voltage transmission system diagnostics
6	Utilizing Remote Sensing to Estimate Soil Organic Carbon	Odebiri et al.	2021	Review on using machine learning for soil organic carbon estimation

4 Recent Advances in Neural Network Architectures

4.1 Advancements in Transformer Models

As of 2017, the Transformer became the most common architecture with respect to natural language processing (NLP) tasks. These models have evidently surpassed the performance of previous models to a large extent because of the attention mechanism that has been inserted into the architectures that enable these models to learn long range textual dependencies. BERT (Bidirectional Encoder Representations from Transformers), presented by Devlin et al. in 2018, represented a major advancement in NLP. As it is BERT s case, its tasks can contribute to a range of further applications including sentiment analysis, question answering, and text classification among others. The success made by BERT also inspired the creation of other large models such as GPT 3. Released in 2020 by Brown et al., GPT 3 (Generative Pre trained Transformer 3) is a gigantic model comprising 175 billion parameters. It exhibits great capabilities such as writing essays that resemble the writing of humans, translating languages, writing all kinds of creative works and answering questions reasonably well [83]. But the capabilities of such models also invoke possible biases and ethical issues.

4.2 Self Supervised Learning Techniques

Self supervised learning is a method that is helpful in constructing relevant representations from data sources that haven t been labelled, and this method has found much applicability in the field of natural language processing (NLP) given the lots of unlabelled text data sources available at no cost. One such method in this self supervised learning approach, named as contrastive learning, has performed quite well in the domain of representation learning. Here the focus is on constructing models which are able to identify similar and dissimilar data points, hence being able to construct appropriate representations which bear some essential semantic meaning. Thanks to the ingenuity of diabetes modifier emergence of transformer architecture and self supervised learning approaches, a lot of progress has been recorded in developing sophisticated and flexible language models in NLP. As this domain of study advances, more discoveries in NLP and representation learning are anticipated and this will extend the frontiers of human computer interactions.

4.3 Other developments

An approach that effectively speeds up the machine learning based chip design was demonstrated which has the potential to facilitate enhancement in chip design processes [84].Also, the spatio temporal graph neural networks were explored in the context of traffic forecasting and their potential in enhancing traffic forecasting was pointed out [85]. Further, a wavelet transformation based Convolutional Neural Network technique for vessel monitoring using SAR data was presented and enhanced object detection under high naval traffic [86]. A study on human computer interaction neural networks enabled by HRI with a focus on natural language understanding and interaction engines was presented, demonstrating progress in human machine communication systems [87]. In context of medicine and surgery, merging these principles with Artificial neural networks were shared, stressing the importance of Artificial neural networks in data analysis and the fast emerging technologies [88]. A neural network based representation for the Pauli potential was formulated in pursuit of the orbital free density functional theory showing promise in enhancing theoretical models [89]. At the same time, the concept of inter system interaction networks was proposed as a self organizing mechanism for the development of autonomous systems, which led to the understanding of neural networks as systems that are not just composed of wires between modules, but organized into larger hierarchies, which the authors referred to as meta lateness structures. The problem of integrating uncertainty estimation in predictive capabilities with Bayesian neural networks (BNNs) was also addressed as a problem of conventional networks that have restricted capabilities, thus enhancing the behaviour and performance of the models [90]. These articles explained the use of advanced approaches through neural networks using structural feature information as well as convolutional neural networks for structural health monitoring systems [91]. The combination of these two components resulted in successful achievements in the advancement of artificial intelligence by optimizing analogue AI through software hardware co design with edge pruning [92]. Subspace clustering useful for nonlinear problems in FNNs was accomplished with high computational efficiency that resulted in clustering performance that was quite desirable [93]. The specificity of graph neural networks demonstrated using graph Lambda to predict the protein ligand binding affinity showed an enhancement in predictive accuracy [94]. Clinical diagnostic support using deep learning networks for convolutional neural networks to propose PS5 Net, which is based on medical image segmentation neurons for deep neural networks, was also illustrated [95].

5 **Cutting Edge Applications and Ethical** Considerations

As seen with the Deep Drive approach which indeed enhances autonomous driving systems, deep learned driving models can be trained end to end, that is CNNs are able to perform steering commands only from raw video sensor pixels [96]. It has been demonstrated that an inception model algorithm trained for skin lesions over a dataset of more than 100000 can classify skin cancer with dermatologists level of expertise, thus affirming the power of neural networks to complement clinical practitioners and improve diagnostic [97]. Leibovich and colleagues developed a neural network dedicated for representing the planetary boundary layer to use within a deeper learning framework for global high resolution modeling of climate, indicating the capability of deep learning to derive and reproduce dynamics of highly turbulent climate systems. Such research demonstrates the capability of neural networks in climate modeling and general areas of climate behavior to advance global climate understanding [98]. The application of generative models, graph neural networks, and reinforcement learning presented new molecular designs and optimized chemical synthesis routes which indicate the ability of neural networks to speed up the processes of modern drug design and development further helping in the global battle against various diseases [99].

There are a number of ethical issues regarding the effectiveness of neural networks that the scientists are addressing. However, fairness continues to be a challenge, as the models may replicate pre existent societal biases that are found in the data. Anything such as fairness aware machine learning tackles the issues more mildly, where de biasing pre training techniques managed to eliminate gender stereotypes in word embeddings [100]. Besides, the complicated structure of neural networks tends to render them as black boxes characteristic and thus making them less interpretable and usable for high risk applications. Explainable AI tools like SHAP (SHapley Additive exPlanations) allow interpreting machine learning models, which makes it possible to explain their output based on the contribution of individual features to the predicted value improving reliability, accountability and trust [101]. The combination of neural networks with deep learning algorithms can produce astonishing results, but this also puts an important responsibility into the hands of users since the outcome must be just, impartial and clear to a person. Seeing to reasonably solve these issues is key to build artificial intelligence systems that integrate tenets of fairness, openness and privacy and at the same time serve the community. In order to achieve this goal, however, more studies need to be done, and also ethical principles need to be embedded into the machine learning lifecycle from the beginning.

6 Conclusions

To summarize, a range of complex challenges have been resolved by the provision of neural networks in a number of applications. In their various forms including feedforward and recurrent networks alongside newer models like CNNs or transformers, these systems have found considerable applicability in computer vision, natural language processing and healthcare diagnostics. Problems like climate change modelling, drug design or the quest for driverless cars clearly indicate the tremendous scope of these systems to support and improve forecasting and decision making tasks. But at the same time with these capabilities the sophisticated design aspects of neural networks bring challenges of interpretability, fairness and bias out of the box which underline the essentials for the future of ethical AI research. Further, the field of regularization techniques and therefore optimization techniques is also progressing leading to better models in terms of computational efficiency and scalability. Advancing features of neural network models evidently promotes advancement in AI technologies, however, this calls for accountability and operational transparency for the building of public confidence and trust on such technologies. It is therefore promising to believe that better foresight will come within a responsible development framework speedily improving the standing of neural network applications in the future.

7 Declarations

7.1 Competing Interests

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

7.2 Publisher's Note

AIJR remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

How to Cite

Harshit Gupta, Manya Verma, Anushka Arora, Dhriti, Gyanender Kumar (2025). Exploring the Depths: A Comprehensive Survey of Neural Networks and their Applications. *AIJR Proceedings*, 42-57. https://doi.org/10.21467/proceedings.178.6

References

[1] R. D. Vanluchene and R. SUN, "Neural Networks in Structural Engineering," *Computer-Aided Civil and Infrastructure Engineering*, vol. 5, no. 3, pp. 207–215, 1990, doi: 10.1111/j.1467-8667.1990.tb00377.x.

[2] I. GUYON, "Applications Of Neural Networks To Character Recognition," *Intern J Pattern Recognit Artif Intell*, vol. 05, no. 01n02, pp. 353–382, 1991, doi: 10.1142/s021800149100020x.

[3] V. L. Wiggins, L. T. Looper, and S. K. Engquist, "Neural Networks: A Primer," Defense Technical Information Center, 1991. doi: 10.21236/ada235920.

[4] J. Nannariello and F. R. Fricke, "Introduction to neural network analysis and its application to building services engineering," *Building Services Engineering Research and Technology*, vol. 22, no. 1, pp. 58–68, 2001, doi: 10.1191/014362401701524127.

[5] N. Yadav, A. Yadav, and M. Kumar, "Conclusion," *SpringerBriefs in Applied Sciences and Technology*, pp. 101–109, 2015, doi: 10.1007/978-94-017-9816-7.

[6] T. Golan, P. C. Raju, and N. Kriegeskorte, "Controversial stimuli: adjudicating between deep neural network models of biological vision with synthetic images," *J Vis*, vol. 20, no. 11, p. 947, 2020, doi: 10.1167/jov.20.11.947.

[7] R. Lippmann, "Book Review: 'Neural Networks, A Comprehensive Foundation', by Simon Haykin," *Int J Neural Syst*, vol. 05, no. 04, pp. 363–364, 1994, doi: 10.1142/s0129065794000372.

[8] H. White, "Learning in Artificial Neural Networks: A Statistical Perspective," *Neural Comput*, vol. 1, no. 4, pp. 425–464, 1989, doi: 10.1162/neco.1989.1.4.425.

[9] T. Fukuda and T. Shibata, "Theory and applications of neural networks for industrial control systems," *IEEE Transactions on Industrial Electronics*, vol. 39, no. 6, pp. 472–489, 1992, doi: 10.1109/41.170966.

[10] V. W. Porto and D. B. Fogel, "Evolving artificial neural networks," *Clinical Applications of Artificial Neural Networks*, pp. 223–234, Dec. 2009, doi: 10.1017/CBO9780511543494.010.

[11] B. Hammer, "Neural Smithing – Supervised Learning in Feedforward Artificial Neural Networks," *Pattern Analysis & Applications*, vol. 4, no. 1, pp. 73–74, 2001, doi: 10.1007/s100440170029.

[12] A. Krogh, "What are artificial neural networks?," *Nat Biotechnol*, vol. 26, no. 2, pp. 195–197, 2008, doi: 10.1038/nbt1386.

[13] S. A. Marhon, C. J. F. Cameron, and S. C. Kremer, "Recurrent Neural Networks," *Intelligent Systems Reference Library*, vol. 49, pp. 29–65, 2013, doi: 10.1007/978-3-642-36657-4_2.

[14] R. Kumar, R. K. Aggarwal, and J. D. Sharma, "Energy analysis of a building using artificial neural network: A review," *Energy Build*, vol. 65, pp. 352–358, 2013, doi: 10.1016/j.enbuild.2013.06.007.

[15] S. Geman, E. Bienenstock, and R. Doursat, "Neural Networks and the Bias/Variance Dilemma," *Neural Comput*, vol. 4, no. 1, pp. 1–58, Jan. 1992, doi: 10.1162/neco.1992.4.1.1.

[16] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Networks*, vol. 61, pp. 85–117, 2015, doi: 10.1016/j.neunet.2014.09.003.

[17] Y. C. Pati and P. S. Krishnaprasad, "Analysis and synthesis of feedforward neural networks using discrete affine wavelet transformations," *IEEE Trans Neural Netw*, vol. 4, no. 1, pp. 73–85, 1993, doi: 10.1109/72.182697.

[18] D. T. Pham and X. Liu, *Neural Networks for Identification, Prediction and Control*. Springer London, 1995. doi: 10.1007/978-1-4471-3244-8.

[19] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: a new learning scheme of feedforward neural networks," in *2004 IEEE International Joint Conference on Neural Networks (IEEE Cat. No.04CH37541)*, in IJCNN-04. IEEE. doi: 10.1109/ijcnn.2004.1380068.

[20] A. Hirose and S. Yoshida, "Generalization Characteristics of Complex-Valued Feedforward Neural Networks in Relation to Signal Coherence," *IEEE Trans Neural Netw Learn Syst*, vol. 23, no. 4, pp. 541–551, Apr. 2012, doi: 10.1109/tnnls.2012.2183613.

[21] M. Sundermeyer, I. Oparin, J.-L. Gauvain, B. Freiberg, R. Schluter, and H. Ney, "Comparison of feedforward and recurrent neural network language models," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, IEEE, May 2013. doi: 10.1109/icassp.2013.6639310.

[22] "The Power Of Depth For Feedforward Neural Networks | Paper Digest." Accessed: Apr. 17, 2024. [Online]. Available: https://www.paperdigest.org/paper/?paper_id=colt-eldan16.html-2016-12-30

[23] J. Berg and K. Nyström, "A unified deep artificial neural network approach to partial differential equations in complex geometries," *Neurocomputing*, vol. 317, pp. 28–41, Nov. 2018, doi: 10.1016/j.neucom.2018.06.056.

[24] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Comput*, vol. 9, no. 8, pp. 1735–1780, 1997, doi: 10.1162/neco.1997.9.8.1735.

[25] J. T. Connor, R. D. Martin, and L. E. Atlas, "Recurrent neural networks and robust time series prediction," *IEEE Trans Neural Netw*, vol. 5, no. 2, pp. 240–254, 1994, doi: 10.1109/72.279188.

[26] B. A. Pearlmutter, "Gradient calculations for dynamic recurrent neural networks: a survey," *IEEE Trans Neural Netw*, vol. 6, no. 5, pp. 1212–1228, 1995, doi: 10.1109/72.410363.

[27] D. P. Mandic and J. A. Chambers, *Recurrent Neural Networks for Prediction*. Wiley, 2001. doi: 10.1002/047084535x.

[28] H. Zhang, Z. Wang, and D. Liu, "A Comprehensive Review of Stability Analysis of Continuous-Time Recurrent Neural Networks," *IEEE Trans Neural Netw Learn Syst*, vol. 25, no. 7, pp. 1229–1262, 2014, doi: 10.1109/tnnls.2014.2317880.

[29] "Translating Videos To Natural Language Using Deep Recurrent Neural Networks | Paper Digest." Accessed: Apr. 17, 2024. [Online]. Available: https://www.paperdigest.org/paper/?paper_id=naacl-N15-1173-2015-12-30

[30] C.-M. Kuan and H. White, "Artificial neural networks: an econometric perspective*," *Econom Rev*, vol. 13, no. 1, pp. 1–91, 1994, doi: 10.1080/07474939408800273.

[31] "Joint Event Extraction Via Recurrent Neural Networks | Paper Digest." Accessed: Apr. 17, 2024. [Online]. Available: https://www.paperdigest.org/paper/?paper_id=naacl-N16-1034-2016-12-30

[32] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998, doi: 10.1109/5.726791.

[33] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun ACM*, vol. 60, no. 6, pp. 84–90, 2017, doi: 10.1145/3065386.

[34] "Convolutional Neural Networks For Sentence Classification | Paper Digest." Accessed: Apr. 17, 2024. [Online]. Available: https://www.paperdigest.org/paper/?paper_id=emnlp-D14-1181-2014-12-30

[35] "Predicting Alzheimer's Disease: A Neuroimaging Study With 3D Convolutional Neural Networks | Paper Digest." Accessed: Apr. 17, 2024. [Online]. Available: https://www.paperdigest.org/paper/?paper_id=arxiv-1502.02506

[36] J. Gu *et al.*, "Recent advances in convolutional neural networks," *Pattern Recognit*, vol. 77, pp. 354–377, 2018, doi: 10.1016/j.patcog.2017.10.013.

[37] "Effective Use Of Word Order For Text Categorization With Convolutional Neural Networks | Paper Digest." Accessed: Apr. 17, 2024. [Online]. Available: https://www.paperdigest.org/paper/?paper_id=naacl-N15-1011-2015-12-30

[38] "Classifying Relations By Ranking With Convolutional Neural Networks | Paper Digest." Accessed: Apr. 17, 2024. [Online]. Available: https://www.paperdigest.org/paper/?paper_id=acl-P15-1061-2015-12-30

[39] "Recent Advances In Convolutional Neural Networks | Paper Digest." Accessed: Apr. 17, 2024. [Online]. Available: https://www.paperdigest.org/paper/?paper_id=arxiv-1512.07108

[40] "Fast Algorithms For Convolutional Neural Networks | Paper Digest." Accessed: Apr. 17, 2024. [Online]. Available: https://www.paperdigest.org/paper/?paper_id=cvpr-Lavin_Fast_Algorithms_for_CVPR_2016_paper.html-2016-06-23

[41] "4D Spatio-Temporal ConvNets: Minkowski Convolutional Neural Networks | Paper Digest." Accessed: Apr. 17, 2024. [Online]. Available: https://www.paperdigest.org/paper/?paper_id=cvpr-Choy_4D_Spatio-Temporal_ConvNets_Minkowski_Convolutional_Neural_Networks_CVPR_2019_paper.html-2019-06-14

[42] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, no. 6088, pp. 533–536, 1986, doi: 10.1038/323533a0.

[43] G. Qiu, M. R. Varley, and T. J. Terrell, "Accelerated training of backpropagation networks by using adaptive momentum step," *Electron Lett*, vol. 28, no. 4, p. 377, 1992, doi: 10.1049/el:19920236.

[44] M. N. H. Siddique and M. O. Tokhi, "Training neural networks: Backpropagation vs genetic algorithms," *Proceedings of the International Joint Conference on Neural Networks*, vol. 4, pp. 2673–2678, 2001, doi: 10.1109/IJCNN.2001.938792.

[45] M. I. El Adawy, M. E. Aboul-Wafa, H. A. Keshk, and M. M. El Tayeb, "A SOFT-backpropagation algorithm for training neural networks," *National Radio Science Conference, NRSC, Proceedings*, vol. 2002-January, pp. 397–404, 2002, doi: 10.1109/NRSC.2002.1022647.

[46] Y. Huang, "Advances in Artificial Neural Networks – Methodological Development and Application," *Algorithms*, vol. 2, no. 3, pp. 973–1007, 2009, doi: 10.3390/algor2030973.

[47] Y. Wu, L. Deng, G. Li, J. Zhu, and L. Shi, "Spatio-Temporal Backpropagation for Training High-Performance Spiking Neural Networks," *Front Neurosci*, vol. 12, p. 331, May 2018, doi: 10.3389/fnins.2018.00331.

[48] X. Guo, T. D. Barrett, Z. M. Wang, and A. I. Lvovsky, "Backpropagation through nonlinear units for the all-optical training of neural networks," *Photonics Res*, vol. 9, no. 3, p. B71, 2021, doi: 10.1364/prj.411104.

[49] "Spike-Train Level Backpropagation for Training Deep Recurrent Spiking Neural Networks | Paper Digest." Accessed: Apr. 17, 2024. [Online]. Available: https://www.paperdigest.org/paper/?paper_id=nips-8995-spike-train-level-backpropagation-for-training-deep-recurrent-spiking-neural-networks-2019-11-15

[50] L. G. Wright *et al.*, "Deep physical neural networks trained with backpropagation," *Nature*, vol. 601, no. 7894, pp. 549–555, Jan. 2022, doi: 10.1038/s41586-021-04223-6.

[51] Y. Shang and B. W. Wah, "Global optimization for neural network training," *Computer (Long Beach Calif)*, vol. 29, no. 3, pp. 45–54, Mar. 1996, doi: 10.1109/2.485892.

[52] D. Karaboga and B. Akay, "Artificial Bee Colony (ABC) Algorithm on Training Artificial Neural Networks," in *2007 IEEE 15th Signal Processing and Communications Applications*, IEEE, Jun. 2007. doi: 10.1109/siu.2007.4298679.

[53] K. Socha and C. Blum, "An ant colony optimization algorithm for continuous optimization: application to feed-forward neural network training," *Neural Comput Appl*, vol. 16, no. 3, pp. 235–247, Mar. 2007, doi: 10.1007/s00521-007-0084-z.

[54] K. Khan and A. Sahai, "A Comparison of BA, GA, PSO, BP and LM for Training Feed forward Neural Networks in e-Learning Context," *International Journal of Intelligent Systems and Applications*, vol. 4, no. 7, pp. 23–29, Jun. 2012, doi: 10.5815/ijisa.2012.07.03.

[55] "A Quantum Algorithm To Train Neural Networks Using Low-depth Circuits | Paper Digest." Accessed: Apr. 17, 2024. [Online]. Available: https://www.paperdigest.org/paper/?paper_id=arxiv-1712.05304

[56] Z. Zhang, "Improved Adam Optimizer for Deep Neural Networks," in *2018 IEEE/ACM 26th International Symposium on Quality of Service (IWQoS)*, IEEE, Jun. 2018. doi: 10.1109/iwqos.2018.8624183.

[57] I. Aljarah, H. Faris, and S. Mirjalili, "Optimizing connection weights in neural networks using the whale optimization algorithm," *Soft comput*, vol. 22, no. 1, pp. 1–15, Nov. 2016, doi: 10.1007/s00500-016-2442-1.

[58] "Hyper-Parameter Optimization: A Review Of Algorithms And Applications | Paper Digest." Accessed: Apr. 17, 2024. [Online]. Available: https://www.paperdigest.org/paper/?paper_id=arxiv-2003.05689

[59] B. Boukhatem, S. Kenai, A. T. Hamou, Dj. Ziou, and M. Ghrici, "Predicting concrete properties using neural networks (NN) with principal component analysis (PCA) technique," *Computers & concrete*, vol. 10, no. 6, pp. 557–573, Dec. 2012, doi: 10.12989/cac.2012.10.6.557.

[60] S. Zhang, H. Jiang, S. Wei, and L.-R. Dai, "Rectified linear neural networks with tied-scalar regularization for LVCSR," in *Interspeech 2015*, in interspeech_2015. ISCA, Sep. 2015. doi: 10.21437/interspeech.2015-562.

[61] T. Hinz, P. Barros, and S. Wermter, “The Effects of Regularization on Learning Facial Expressions with Convolutional Neural Networks,” in *Lecture Notes in Computer Science*, Springer International Publishing, 2016, pp. 80–87. doi: 10.1007/978-3-319-44781-0_10.

[62] W. Wen, C. Xu, C. Wu, Y. Wang, Y. Chen, and H. Li, “Coordinating Filters for Faster Deep Neural Networks,” in *2017 IEEE International Conference on Computer Vision (ICCV)*, IEEE, Oct. 2017. doi: 10.1109/iccv.2017.78.

[63] J. Lee, R. Zhang, W. Zhang, Y. Liu, and P. Li, “Spike-Train Level Direct Feedback Alignment: Sidestepping Backpropagation for On-Chip Training of Spiking Neural Nets,” *Front Neurosci*, vol. 14, p. 143, Mar. 2020, doi: 10.3389/fnins.2020.00143.

[64] I. Jindal, M. Nokleby, and X. Chen, “Learning Deep Networks from Noisy Labels with Dropout Regularization,” in *2016 IEEE 16th International Conference on Data Mining (ICDM)*, IEEE, Dec. 2016. doi: 10.1109/icdm.2016.0121.

[65] X. Xu, X. Zhou, R. Venkatesan, G. Swaminathan, and O. Majumder, “d-SNE: Domain Adaptation Using Stochastic Neighborhood Embedding,” in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Jun. 2019. doi: 10.1109/cvpr.2019.00260.

[66] N. Audebert, A. Boulch, B. Le Saux, and S. Lefèvre, “Distance transform regression for spatially-aware deep semantic segmentation,” *Computer Vision and Image Understanding*, vol. 189, p. 102809, Dec. 2019, doi: 10.1016/j.cviu.2019.102809.

[67] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553, pp. 436–444, 2015, doi: 10.1038/nature14539.

[68] J. Heaton, “Ian Goodfellow, Yoshua Bengio, and Aaron Courville: Deep learning,” *Genet Program Evolvable Mach*, vol. 19, no. 1–2, pp. 305–307, 2017, doi: 10.1007/s10710-017-9314-z.

[69] A. Vaswani *et al.*, “Attention is All you Need,” *Adv Neural Inf Process Syst*, vol. 30, 2017.

[70] J. Devlin, M.-W. Chang, K. Lee, K. T. Google, and A. I. Language, “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding,” *Proceedings of the 2019 Conference of the North*, pp. 4171–4186, 2019, doi: 10.18653/V1/N19-1423.

[71] N. P. Jouppi *et al.*, “In-datacenter performance analysis of a tensor processing unit,” *Proc Int Symp Comput Archit*, vol. Part F128643, pp. 1–12, Jun. 2017, doi: 10.1145/3079856.3080246.

[72] “Interspeech 2017, 18th Annual Conference of the International Speech Communication Association, Stockholm, Sweden, August 20-24, 2017 - researchr publication.” Accessed: Apr. 17, 2024. [Online]. Available: https://researchr.org/publication/interspeech-2017

[73] F. Doshi-Velez and B. Kim, “Towards A Rigorous Science of Interpretable Machine Learning,” Feb. 2017, Accessed: Apr. 17, 2024. [Online]. Available: https://arxiv.org/abs/1702.08608v2

[74] I. J. Goodfellow, J. Shlens, and C. Szegedy, “Explaining and Harnessing Adversarial Examples,” *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, Dec. 2014, Accessed: Apr. 17, 2024. [Online]. Available: https://arxiv.org/abs/1412.6572v3

[75] S. Nishimoto, Y. Tokuoka, T. G. Yamada, N. F. Hiroi, and A. Funahashi, “Predicting the future direction of cell movement with convolutional neural networks,” *PLoS One*, vol. 14, no. 9, pp. e0221245–e0221245, Sep. 2019, doi: 10.1371/journal.pone.0221245.

[76] H. Fouad, A. S. Hassanein, A. M. Soliman, and H. Al-Feel, “Analyzing patient health information based on IoT sensor with AI for improving patient assistance in the future direction,” *Measurement*, vol. 159, p. 107757, 2020, doi: 10.1016/j.measurement.2020.107757.

[77] Y. Jiao, Y.-W. Liu, W.-G. Chen, and J. Liu, “Neuroregeneration and functional recovery after stroke: advancing neural stem cell therapy toward clinical application,” *Neural Regen Res*, vol. 16, no. 1, pp. 80–92, Jan. 2021, doi: 10.4103/1673-5374.286955.

[78] J. Lu, L. Tan, and H. Jiang, “Review on Convolutional Neural Network (CNN) Applied to Plant Leaf Disease Classification,” *Agriculture*, vol. 11, no. 8, p. 707, 2021, doi: 10.3390/agriculture11080707.

[79] L. Ke, Y. Zhang, B. Yang, Z. Luo, and Z. Liu, “Fault diagnosis with synchrosqueezing transform and optimized deep convolutional neural network: An application in modular multilevel converters,” *Neurocomputing*, vol. 430, pp. 24–33, Mar. 2021, doi: 10.1016/j.neucom.2020.11.037.

[80] O. Odebiri, J. Odindi, and O. Mutanga, “Basic and deep learning models in remote sensing of soil organic carbon estimation: A brief review,” *International Journal of Applied Earth Observation and Geoinformation*, vol. 102, p. 102389, Oct. 2021, doi: 10.1016/j.jag.2021.102389.

[81] R. Pearce and Y. Zhang, “Deep learning techniques have significantly impacted protein structure prediction and protein design,” *Curr Opin Struct Biol*, vol. 68, pp. 194–207, Jun. 2021, doi: 10.1016/j.sbi.2021.01.007.

[82] U. A. Bhatti, H. Tang, G. Wu, S. Marjan, and A. Hussain, “Deep Learning with Graph Convolutional Networks: An Overview and Latest Applications in Computational Intelligence,” *International Journal of Intelligent Systems*, vol. 2023, pp. 1–28, 2023, doi: 10.1155/2023/8342104.

[83] T. B. Brown *et al.*, “Language Models are Few-Shot Learners,” *Adv Neural Inf Process Syst*, vol. 2020-December, May 2020, Accessed: Apr. 17, 2024. [Online]. Available: https://arxiv.org/abs/2005.14165v4

[84] Y. Li, “WITHDRAWN: Design of cultural creative products based on FPGA and convolutional neural network,” *Microprocess Microsyst*, p. 103483, 2020, doi: 10.1016/j.micpro.2020.103483.

[85] B. Khailany *et al.*, "Accelerating Chip Design with Machine Learning," *IEEE Micro*, vol. 40, no. 6, pp. 23–32, Nov. 2020, doi: 10.1109/MM.2020.3026231.

[86] B. Yu, H. Yin, and Z. Zhu, "Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting," *IJCAI International Joint Conference on Artificial Intelligence*, vol. 2018-July, pp. 3634–3640, 2018, doi: 10.24963/ijcai.2018/505.

[87] S. P. Tiwari, S. K. Chaturvedi, S. Adhikary, S. Banerjee, and S. Basu, "Automatized marine vessel monitoring from sentinel-1 data using convolution neural network," *International Geoscience and Remote Sensing Symposium (IGARSS)*, vol. 2021-July, pp. 1311–1314, Apr. 2023, doi: 10.1109/igarss47720.2021.9555149.

[88] Y. Ma, L. Zhang, and X. Wang, "Natural language understanding and interaction engine oriented to human-computer interaction based on neural network," *https://doi.org/10.1117/12.2660383*, vol. 12511, pp. 781–786, Feb. 2023, doi: 10.1117/12.2660383.

[89] M. Jiřík, V. Moulisová, M. Hlaváč, M. Železný, and V. Liška, "Artificial neural networks and computer vision in medicine and surgery," *Rozhl Chir*, vol. 101, no. 12, pp. 564–570, Dec. 2022, doi: 10.33699/PIS.2022.101.12.564-570.

[90] J. Arbel, K. Pitas, M. Vladimirova, and V. Fortuin, "A Primer on Bayesian Neural Networks: Review and Debates".

[91] Y. Lu, D. Wang, D. Liu, and X. Yang, "A Lightweight and Efficient Method of Structural Damage Detection Using Stochastic Configuration Network," *Sensors (Basel)*, vol. 23, no. 22, Nov. 2023, doi: 10.3390/s23229146.

[92] Y. Li *et al.*, "Pruning random resistive memory for optimizing analogue AI," Nov. 2023, Accessed: Apr. 18, 2024. [Online]. Available: https://arxiv.org/abs/2311.07164v1

[93] L. Shi, L. Cao, Z. Chen, B. Chen, and Y. Zhao, "Nonlinear subspace clustering by functional link neural networks", Accessed: Apr. 18, 2024. [Online]. Available: https://lshi91.github.io/

[94] G. Mqawass and P. Popov, "graphLambda: Fusion Graph Neural Networks for Binding Affinity Prediction," *J Chem Inf Model*, vol. 64, no. 7, pp. 2323–2330, 2024, doi: 10.1021/acs.jcim.3c00771.

[95] "PS5-Net: A Medical Image Segmentation Network with Multiscale Resolution | Paper Digest." Accessed: Apr. 18, 2024. [Online]. Available: https://www.paperdigest.org/paper/?paper_id=pubmed-38379775

[96] M. Bojarski *et al.*, "End to End Learning for Self-Driving Cars," Apr. 2016, Accessed: Apr. 17, 2024. [Online]. Available: https://arxiv.org/abs/1604.07316v1

[97] A. Esteva *et al.*, "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, no. 7639, pp. 115–118, Feb. 2017, doi: 10.1038/nature21056.

[98] S. Rasp, M. S. Pritchard, and P. Gentine, "Deep learning to represent subgrid processes in climate models," *Proc Natl Acad Sci U S A*, vol. 115, no. 39, pp. 9684–9689, Sep. 2018, doi: 10.1073/pnas.1810286115.

[99] J. M. Stokes *et al.*, "A Deep Learning Approach to Antibiotic Discovery," *Cell*, vol. 180, no. 4, pp. 688-702.e13, Feb. 2020, doi: 10.1016/j.cell.2020.01.021.

[100] T. Bolukbasi, K. W. Chang, J. Zou, V. Saligrama, and A. Kalai, "Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings," *Adv Neural Inf Process Syst*, pp. 4356–4364, Jul. 2016, Accessed: Apr. 17, 2024. [Online]. Available: https://arxiv.org/abs/1607.06520v1

[101] S. M. Lundberg and S. I. Lee, "A Unified Approach to Interpreting Model Predictions," *Adv Neural Inf Process Syst*, vol. 2017-December, pp. 4766–4775, May 2017, Accessed: Apr. 17, 2024. [Online]. Available: https://arxiv.org/abs/1705.07874v2