DOI: http://dx.doi.org/10.26483/ijarcs.v15i2.7052

ISSN No. 0976-5697

Volume 15, No. 2, March-April 2024

International Journal of Advanced Research in Computer Science

SURVEY REPORT

Available Online at www.ijarcs.info

A Comprehensive Survey of Deep Learning and Its Applications in Advancing Artificial Intelligence

A.Jasmine Sugil
Research Scholar (Reg.No:23111242282011)
Department of Computer Applications & Research Centre
Sarah Tucker College (Autonomous), affiliated to
Manonmaniam Sundaranar University, Tirunelveli-627007
Tamilnadu, India

Dr.K.Merriliance
Associate Professor
Department of Computer Applications & Research Centre
Sarah Tucker College (Autonomous), Tirunelveli-627007
Tamilnadu,India,

Dr. Mary Immaculate Sheela Lourdusamy
Professor, Department of Informatics
Heritage Christian University College
Accra, Ghana

Abstract: Deep learning, a subset of machine learning, stands at the forefront of artificial intelligence, striving to bridge the gap to its ultimate goal. This paper employs summary and induction methodologies to research into the area of deep learning. It begins by surveying the global development and current landscape of deep learning. Next, it elucidates the structural principles, characteristics, and key models, including stacked auto encoders, deep belief networks, deep Boltzmann machines, and convolutional neural networks. Furthermore, it examines the latest advancements and applications of deep learning across diverse domains such as speech processing, computer vision, natural language processing, and medical diagnostics. Finally, the paper outlines the challenges and future research directions within the realm of deep learning.

Keywords: Deep learning, Structural Principles, Neural Networks, Deep Applications, Challenges

1. INTRODUCTION

A powerful branch of artificial intelligence (AI) and machine learning, deep learning (Figure 1)has garnered a lot of interest and popularity recently because of its capacity to solve challenging issues and produce outstanding outcomes in a variety of fields. At its core, deep learning seeks to mimic the workings of the human brain by using artificial neural networks composed of multiple layers of interconnected nodes, or neurons, to process and learn from vast amounts of data. One of the key advantages of deep learning is its scalability, as it can handle large datasets with high dimensionality, such as images, audio, text, and more. This has led to groundbreaking advancements in various fields, including computer vision, natural language processing, speech recognition, medical diagnosis, autonomous vehicles, and many others.

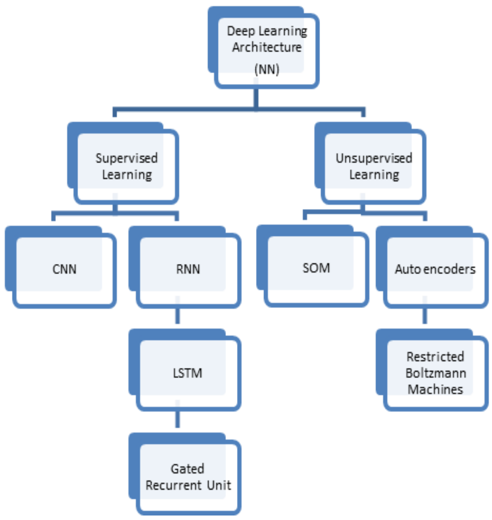


Figure 1. Arichitecture of Deep learning

1.1. Important of Deep Learning

There is only one reason why deep learning is significant: Capable of obtaining significant, practical precision on important assignments. For many years, machine learning has been used to classify text and images, but it has had trouble breaking through since commercial settings require a minimum level of accuracy for algorithms to function. We can now finally cross that boundary in areas where we previously couldn't thanks to deep learning.

One excellent example of a task that Deep Learning has made feasible for corporate applications is computer vision.

Not only is Deep Learning superior to other conventional algorithms in image classification and labelling, but it is beginning to surpass human performance as well.

- Image, text, and sound analysis are examples of complicated problems that are solved with deep learning algorithms. With a lot of data, deep learning algorithms can get a high level of accuracy.

1.2 The primary factors contributing to the widespread popularity of deep learning.

- The deep learning networks can be efficiently implemented on massively parallel graphics processing units (GPUs).
- They are easy to implement.
- Deep learning networks can handle huge amounts of data.
- Deep learning networks can perform feature extraction and classification in one model.
- Deep learning mimics the way organic brain sort the information, which provides favorable results for the complex problems.

2. NEURAL NETWORK

In deep learning, a pc version learns to carry out classification duties without delay from images, text, or sound. Deep learning fashions can attain cutting-edge accuracy, on occasion exceeding human-stage performance. Models are educated with the aid of using the use of a big set of categorized facts and Neural Network architectures that comprise many layers. It is a set of diverse algorithms to discover a hidden sample of reputation of facts. The NN may be used to understand the sample and classifies in case of numeric facts found in vectors and different real-time facts like text, image, and sound. The unique tasks—classification, regression, and clustering—are accomplished with NN.

Convolutional Neural Networks (CNNs)(Figure 2), Recurrent Neural Networks (RNNs)(Figure3), and Generative Adversarial Networks (GANs) are among the popular architectures within deep learning, each designed to address specific types of tasks and data structures. These networks have revolutionized industries by enabling tasks that were once considered difficult or even impossible, such as image recognition, language translation, and drug discovery.

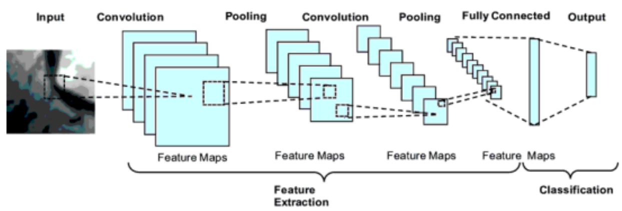


Figure 2. Convolution Neural Network Frame Work

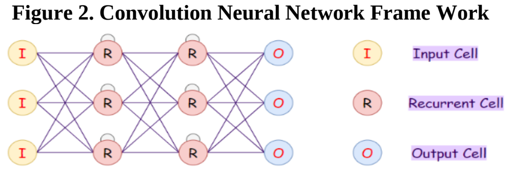


Figure 3. Recurrent Neural Network

Some of the most popular basic structures are the logistic regression [1–5], Support Vector Machines [6–7], Gaussian Mixture Models [8- 10], and so on. In 2019, OpenAI Five [11] made headlines by defeating professional e-sports DOTA2 human teams, making use of a single-layer LSTM in their NN design (Figure 4). More recently in 2021, Efficient Zero [12] shows for the first time a DRL algorithm to achieve super-human performance on Atari games with very limited data. Efficient Zero is able to match DQN's performance at 200 million frames while using 500 times less data using LSTM-based NN architecture. Particular to physics, in 2022, the application of DRL for magnetic control of tokamak plasmas in nuclear fusion [13] utilized an LSTM core in their network.

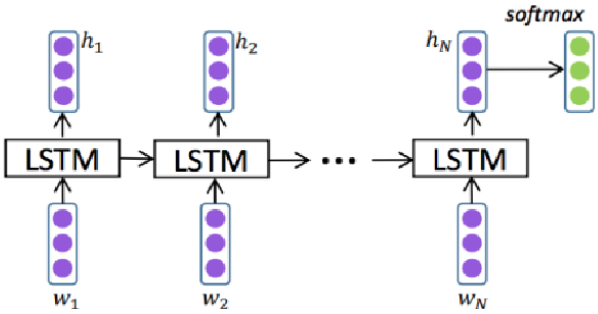
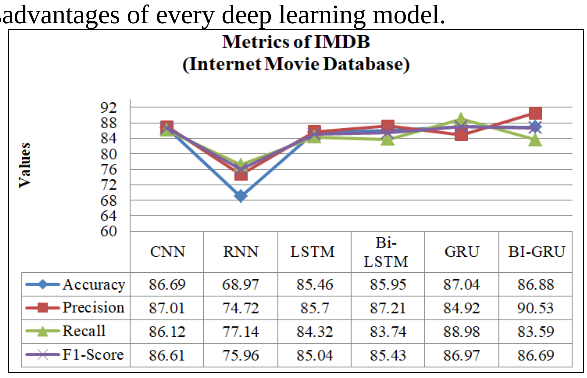


Figure 4. Long Short-Term Memory Frame Work[14]

Shiri et al., [14] provides an extensive examination and comparison of several popular deep learning architectures: Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Long Short-Term Memory networks (LSTMs), and Gated Recurrent Units (GRUs). The paper likely covers various aspects of these models, including their architectures, strengths, weaknesses, and applications in different domains, particularly in natural language processing (NLP), image recognition, time series analysis and used three datasets [20-22] for their investigation, In order to compare the performance of various deep learning models with six well-known models, including the Auto-Encoder [15], Generative Adversarial Network (GAN) [16], Restricted Boltzmann Machine (RBM) [17], Deep Belief Network (DBN) [18], and Self-Organizing Map (SOM) [19]. Using measures for accuracy, precision, F1 -Score and recall, they carried out a thorough investigation and discovered the advantages and disadvantages of every deep learning model.



	CNN	RNN	LSTM	Bi-LSTM	GRU	BI-GRU
Accuracy	86.69	68.97	85.46	85.95	87.04	86.88
Precision	87.01	74.72	85.7	87.21	84.92	90.53
Recall	86.12	77.14	84.32	83.74	88.98	83.59
F1-Score	86.61	75.96	85.04	85.43	86.97	86.69

Figure. 5a Metrics of IMDB (Internet Movie Database) [14]

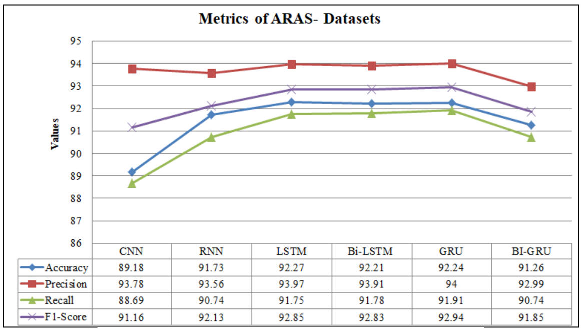


Figure. 5b Metrics of ARAS- Datasets [14]

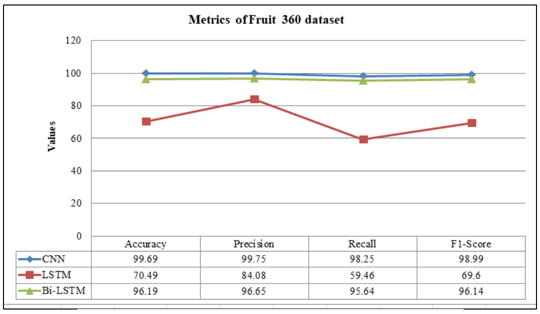


Figure.5c Metrics of Fruit 360 dataset [14]

Based on the IMDB dataset, the GRU, CNN, and Bi-GRU models have been found to be effective deep learning models for sentiment analysis (Fig 5a). In Fig 5b, GRU outperformed the other recurrent models in terms of accuracy or other evaluation metrics for Activity Recognition with Ambient Sensing dataset and Fig 5c Metrics show that CNN models are typically more effective and efficient than recurrent models such as LSTM and Bi-LSTM. The outcomes suggest that the RNN models, like LSTM and GRU, perform well in time series analysis, while the CNN model does remarkably well in image classification tasks.

The above study showcases the versatility and efficiency of deep learning models across multiple fields. It highlights how crucial it is to choose the best deep learning model for the job at hand and the type of data involved. The knowledge gathered from the tests helps to clarify the advantages and disadvantages of various deep learning models, which makes it easier to make wise decisions in real-world situations.

2.1 Comparison for various DL Techniques

Deep Learning Technique	Strengths	Weaknesses	Applications
Feedforward Neural Networks[23]	- Simple architecture	- Limited capability for sequential data	Tabular data analysis, regression, classification
Convolutional Neural Networks[24]	- Effective for image processing tasks	- Requires large amounts of labeled data	Image recognition, object detection, segmentation
Recurrent Neural Networks[25]	- Suitable for sequential data processing	- Prone to vanishing/exploding gradients	Time series prediction, natural language processing
Long Short-Term Memory (LSTM)[26]	- Captures long-term dependencies	- expensive, complex architecture	Language modeling, speech recognition
Gated Recurrent Unit (GRU)[27]	- Simpler architecture than LSTM	- May not capture long-term dependencies well	Language modeling, sentiment analysis
Autoencoders[28]	- Unsupervised feature learning	- Requires careful tuning, may suffer from overfitting	Dimensionality reduction, anomaly detection
Generative Adversarial Networks (GAN)[29]	- Generates realistic synthetic data	- Training instability, mode collapse	Image generation, data augmentation
Transformer Networks[30]	- Highly parallelizable, state-of-the-art NLP performance	- Large memory requirements, computationally intensive	Machine translation, text generation
Capsule Networks [31,32]	- Better handling of spatial hierarchies	- Limited empirical evidence, computationally expensive	Image recognition, object pose estimation
Deep Reinforcement Learning[33]	- Learns policies for sequential decision-making tasks	- High sample complexity, training instability	Game playing, robotics, recommendation systems
DeepBelief Networks (DBNs)[34]	- Effective for unsupervised learning - Suitable for dimensionality reduction -pre-training for other neural networks.	-expensive, and may face vanishing or explosion gradient problems.	Image and Speech recognition, Feature learning
Deep Boltzmann Machines (DBMs)[35]	-Effective in modeling complex probability distributions Suitable for both supervised and unsupervised learning tasks	-Training is computationally demanding, -particularly for large models, and can lead to getting stuck in local minima during training and scaling to deep architectures.	Collaborative filtering, Feature learning, hierarchical representation

3. Applications and Challenges

Deep learning techniques have revolutionized various fields, including computer vision, natural language processing, speech recognition, healthcare, finance, autonomous vehicles, recommendation systems, and robotics. They have improved image recognition, object detection, segmentation, machine translation, speech recognition, healthcare, finance, autonomous vehicles, recommendation engines, and robotics. These advancements have led to improved accuracy, personalized treatment planning, and improved user experiences in various industries.

Deep learning techniques face challenges such as data quality, computational resources, interpretability, overfitting, ethical and legal concerns, adversarial attacks, transfer learning, domain adaptation, and robustness to variability. High-quality labeled datasets are crucial for training, while computational resources are expensive and time-consuming. Interpretability is difficult, and models are prone to overfitting. Ethical and legal concerns, such as dataset biases and privacy concerns, also pose challenges.

4. Conclusion

In conclusion, this paper provides a comprehensive overview of deep learning technology in artificial intelligence, focusing on models like CNN, RNN, generative models, DRL, and transfer learning. It highlights their specific applications, such as temporal dependencies in RNN models, spatial features in CNN models, and faster training times in GRU models. Overall, this paper highlights the diverse applications and effectiveness of deep learning models in various domains. The survey's findings help to clarify the advantages and disadvantages of various deep learning models, which make it easier to make wise decisions in real-world applications.

5. References

[1] Rajkomar, A., Oren, E., Chen, K., Dai, A. M., Hajaj, N., Hardt, M., ... & Dean, J. (2018). Scalable and accurate deep learning with electronic health records. NPJ digital medicine, 1(1), 18.

[2] Marcus, G. (2018). Deep learning: A critical appraisal. arXiv preprint arXiv:1801.00631.

[3] D. Ravì et al., "Deep Learning for Health Informatics," in IEEE Journal of Biomedical and Health Informatics, vol. 21, no. 1, pp. 4-21, Jan. 2017, doi: 10.1109/JBHI.2016.2636665.

[4] Min, S., Lee, B., & Yoon, S. (2017). Deep learning in bioinformatics. Briefings in bioinformatics, 18(5), 851-869.

[5] Sequeira, S. S., Sneha, N., Pallavi, M. O., & Kumar, P. (2022, July). Deep Learning Medical Solutions: A short review on COVID-19 Detection Using Chest X-ray Image with CNN Multiple Architectures. In 2022 Fifth International Conference on Computational Intelligence and Communication Technologies (CCICT) (pp. 198-203). IEEE.

[6] J. Li, A. Sun, J. Han and C. Li, "A Survey on Deep Learning for Named Entity Recognition," in IEEE Transactions on Knowledge and Data Engineering, vol. 34, no. 1, pp. 50-70, 1 Jan. 2022, doi: 10.1109/TKDE.2020.2981314.

[7] Piero Fariselli, Pier Luigi Martelli, Castrense Savojardo, Rita Casadio, INPS: predicting the impact of non-synonymous variations on protein stability from sequence, Bioinformatics, Volume 31, Issue 17, September 2015, Pages 2816–2821, https://doi.org/10.1093/bioinformatics/btv291.

[8] Cao, Y., Geddes, T. A., Yang, J. Y. H., & Yang, P. (2020). Ensemble deep learning in bioinformatics. Nature Machine Intelligence, 2(9), 500-508.

[9] Sekaran, K., Chandana, P., Krishna, N. M., & Kadry, S. (2020). Deep learning convolutional neural network (CNN) With Gaussian mixture model for predicting pancreatic cancer. Multimedia Tools and Applications, 79(15-16), 10233-10247.

[10] Viroli, C., & McLachlan, G. J. (2019). Deep Gaussian mixture models. Statistics and Computing, 29, 43-51.

[11] C. Berner, G. Brockman, B. Chan, V. Cheung, P. Debiak, C. Dennison, D. Farhi, Q. Fischer, S. Hashme, C. Hesse, et al., Dota 2 with large scale deep reinforcement learning, arXiv preprint arXiv:1912.06680 (2019).

[12] W. Ye, S. Liu, T. Kurutach, P. Abbeel, Y. Gao, Mastering atari games with limited data, Advances in Neural Information Processing Systems 34 (2021).

[13] J. Degrave, F. Felici, J. Buchli, M. Neunert, B. Tracey, F. Carpanese, T. Ewalds, R. Hafner, A. Abdolmaleki, D. de Las Casas, et al., Magnetic control of tokamak plasmas through deep reinforcement learning, Nature 602 (7897) (2022) 414–419

[14] Shiri, F., Perumal, T., Mustapha, N., & Mohamed, R. (2023). A Comprehensive Overview and Comparative Analysis on Deep Learning Models: CNN, RNN, LSTM, GRU. ArXiv, abs/2305.17473.

[15] D. Bank, N. Koenigstein, and R. Giryes, "Autoencoders," arXiv preprint arXiv:2003.05991, 2020.

[16] I. Goodfellow et al., "Generative adversarial networks," Communications of the ACM, vol. 63, no. 11, pp. 139-144, 2020.

[17] N. Zhang, S. Ding, J. Zhang, and Y. Xue, "An overview on restricted Boltzmann machines," Neurocomputing, vol. 275, pp. 1186-1199, 2018.

[18] G. E. Hinton, "Deep belief networks," Scholarpedia, vol. 4, no. 5, p. 5947, 2009.

[19] T. Kohonen, Self-organizing maps. Springer Science & Business Media, 2012.

[20] A. Maas, R. E. Daly, P. T. Pham, D. Huang, A. Y. Ng, and C. Potts, "Learning word vectors for sentiment analysis," in Proceedings of the 49th annual meeting of the association for computational linguistics: Human language technologies, 2011, pp. 142-150.

[21] H. Alemdar, H. Ertan, O. D. Incel, and C. Ersoy, "ARAS human activity datasets in multiple homes with multiple residents," in 2013 7th International Conference on Pervasive Computing Technologies for Healthcare and Workshops, 2013: IEEE, pp. 232-235.

[22] H. Mureşan and M. Oltean, "Fruit recognition from images using deep learning," arXiv preprint arXiv:1712.00580, 2017.

[23] Arulampalam, G., & Bouzerdoum, A. (2003). A generalized feedforward neural network architecture for classification and regression. Neural networks, 16(5-6), 561-568.

[24] Yang, R., & Yu, Y. (2021). Artificial convolutional neural network in object detection and semantic segmentation for medical imaging analysis. Frontiers in oncology, 11, 638182.

[25] Hewamalage, H., Bergmeir, C., & Bandara, K. (2021). Recurrent neural networks for time series forecasting: Current status and future directions. International Journal of Forecasting, 37(1), 388-427.

[26] Medennikov, I., & Bulusheva, A. (2016). LSTM-based language models for spontaneous speech recognition. In Speech and Computer: 18th International Conference, SPECOM 2016, Budapest, Hungary, August 23-27, 2016, Proceedings 18 (pp. 469-475). Springer International Publishing.

[27] Al Wazrah, A., & Alhumoud, S. (2021). Sentiment analysis using stacked gated recurrent unit for arabic tweets. IEEE Access, 9, 137176-137187.

[28] Chen, Z., Yeo, C. K., Lee, B. S., & Lau, C. T. (2018, April). Autoencoder-based network anomaly detection.

In 2018 Wireless telecommunications symposium (WTS) (pp. 1-5). IEEE.

[29] Antoniou, A., Storkey, A., & Edwards, H. (2018). Augmenting image classifiers using data augmentation generative adversarial networks. In Artificial Neural Networks and Machine Learning–ICANN 2018: 27th International Conference on Artificial Neural Networks, Rhodes, Greece, October 4-7, 2018, Proceedings, Part III 27 (pp. 594-603). Springer International Publishing.

[30] Li, Y., Feng, R., Rehg, I., & Zhang, C. (2020). Transformer-based neural text generation with syntactic guidance. arXiv preprint arXiv:2010.01737.

[31] Sun, K., Zhang, J., Xu, S., Zhao, Z., Zhang, C., Liu, J., & Hu, J. (2023). CACNN: Capsule Attention Convolutional Neural Networks for 3D Object Recognition. IEEE Transactions on Neural Networks and Learning Systems.

[32] Ha, M. H., & Chen, O. T. C. (2021). Deep neural networks using capsule networks and skeleton-based attentions for action recognition. IEEE Access, 9, 6164-6178.

[33] Zhang, W., Zhao, X., Zhao, L., Yin, D., Yang, G. H., & Beutel, A. (2020, July). Deep reinforcement learning for information retrieval: Fundamentals and advances. In Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval (pp. 2468-2471).

[34] Prasetio, M. D., Hayashida, T., Nishizaki, I., & Sekizaki, S. (2017, November). Deep belief network optimization in speech recognition. In 2017 International Conference on Sustainable Information Engineering and Technology (SIET) (pp. 138-143). IEEE.

[35] Hu, G., Li, H., Xia, Y., & Luo, L. (2018). A deep Boltzmann machine and multi-grained scanning forest ensemble collaborative method and its application to industrial fault diagnosis. Computers in Industry, 100, 287-296.