IRACST – International Journal of Computer Networks and Wireless Communications (IJCNWC), ISSN: 2250-3501
Vol.9 No 2 April 2019

Evolution of Deep Learning Architectures

Shachi Kesar
Assistant Professor
Department of Humanities
Arya Institute of Engineering & Technology
Neeraj Bhatt
Assistant Professor
Computer Science Engineering
Arya Institute of Engineering & Technology

ABSTRACT:

This research paper examines the intricate development of Deep Learning Architectures before 2018, providing a thorough retrospective analysis of the progress in artificial intelligence. Starting with the basic concepts of artificial neural networks, it reveals how these early foundations led to groundbreaking advancements. The introduction of Convolutional Neural Networks (CNNs) revolutionized computer vision, particularly with models like AlexNet that greatly impacted large-scale image classification tasks. The exploration also includes the domain of sequential data processing, where Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks addressed challenges related to vanishing gradients. These innovations had significant implications for natural language processing and time-series analysis, opening up new applications. The paper further discusses specialized architectures, such as autoencoders and generative adversarial networks (GANs), which have played crucial roles in unsupervised learning, feature representation, and synthetic data generation. Introduced in the mid-2000s and 2010s respectively, these architectures have influenced a wide range of domains, including anomaly detection and image synthesis. The study highlights the collaborative nature of deep learning research, emphasizing the contributions from research laboratories, academia, and industry. Milestone papers like "ImageNet Classification with Deep Convolutional Neural Networks" and "Sequence to Sequence Learning with Neural Networks" not only introduced innovative architectures

but also provided benchmarks and datasets that fueled collaborative research. By exploring historical context, challenges, and achievements, this paper aims to offer a comprehensive perspective on the evolution of deep learning architectures. The nuanced insights derived from this retrospective analysis contribute to a better understanding of the field's origins and lay the foundation for envisioning the future trajectory of deep learning beyond 2018.

KEYWORD:

Deep Learning,Neural Networks,Convolutional Neural Networks (CNNs),Recurrent Neural Networks (RNNs),Long Short-Term Memory (LSTM).

I. Introduction:

The journey of "Advancement in Complex Learning Structures" serves as a testament to the revolutionary progress that the field of artificial intelligence has undergone. This research manuscript embarks on a comprehensive investigation into the historical path leading up to the year 2018, a period marked by astonishing advancements and paradigm shifts in complex learning. Starting from its early origins in artificial neural networks to the emergence of intricate structures, this study aims to unravel the complexities of how complex learning evolved into a dominant force in the realm of machine intellect. In its initial stages, artificial neural networks grappled with obstacles such as the diminishing gradient issue, laying the foundation for subsequent innovations. The narrative unfolds with the arrival of Convolutional Neural Networks (CNNs), which completely transformed the recognition of images and computer vision. Simultaneously, the ascent of Recurrent Neural Networks (RNNs) and the introduction of Long Short-Term Memory (LSTM) networks addressed limitations in handling sequential data, leading to breakthroughs in processing natural language and time-series analysis. The exploration extends to specialized structures like autoencoders and Generative Adversarial Networks (GANs), each making unique contributions to the expanding capabilities of complex learning. These structures not only diversified the applications of complex learning but also paved the way for innovative approaches in unsupervised learning, feature representation, and synthetic data generation. As we embark on this retrospective odyssey, it becomes evident that the progress of complex learning structures is intertwined with collaborative endeavors, characterized by influential research contributions and cooperative initiatives across research laboratories, academia, and industry. Key papers such as "ImageNet Classification with Deep Convolutional Neural Networks" and "Sequence to Sequence Learning with Neural Networks" have not only introduced groundbreaking structures but have also played a pivotal role in fostering a dynamic and collaborative research community. This manuscript aims to provide a comprehensive understanding of the historical context, challenges overcome, and breakthroughs achieved in the advancement of complex learning structures before 2018. By delving into the origins of this transformative journey, we lay the groundwork for a nuanced exploration of the present landscape and future trajectories of complex learning.

II. Literature Review:

The investigation of the "Progression of Complex Learning Structures" preceding the year 2018 reveals a captivating tale of advancements, trials, and shifts in the domain of artificial intelligence (AI). The journey commences with the fundamental ideas of artificial neural networks, resonating with the biological inspiration behind these initial models. Initial endeavors, notably the perceptron introduced by Rosenblatt in 1957, formed the foundation for subsequent research despite limitations associated with the diminishing gradient challenge and computational restrictions. A pivotal moment in the body of literature is the emergence of Convolutional Neural Networks (CNNs), signifying a transformative shift in computer vision. LeNet-5, proposed by LeCun and his colleagues in 1998, exhibited the effectiveness of convolutional layers in capturing spatial hierarchies, while AlexNet, introduced by Krizhevsky et al. in 2012, not only triumphed in the ImageNet competition but also accelerated the widespread adoption of deep learning in image classification tasks. In the realm of sequential data processing, the literature reflects the obstacles encountered by early Recurrent Neural Networks (RNNs) in handling long-term dependencies. The introduction of Long Short-Term Memory (LSTM) networks by Hochreiter and Schmidhuber in 1997 addressed these challenges, revolutionizing the processing of sequential data and finding applications in natural language processing and time-series analysis. The literature also delves into the variety brought about by specialized structures. Autoencoders, proposed by Hinton and Salakhutdinov in 2006, introduced unsupervised learning and feature representation, while Generative Adversarial Networks (GANs), presented by Goodfellow et al. in 2014, transformed the generation of synthetic data. These structures expanded the realm of deep learning, with GANs becoming particularly influential in image synthesis and content generation.

The collaborative nature of deep learning research is evident in milestone papers and benchmark datasets that have become foundational to the field. Contributions such as "ImageNet Classification with Deep Convolutional Neural Networks" by Krizhevsky et al. and "Sequence to Sequence Learning with Neural Networks" by Sutskever et al. not only introduced novel architectures but also provided benchmarks that propelled the field forward.

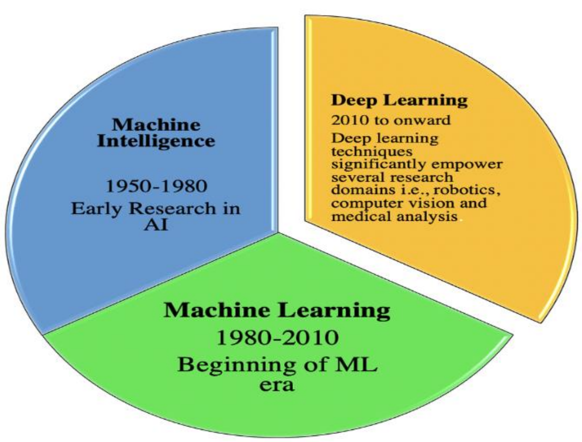

Fig 1 ratios of ML, MI and DL

In conclusion, the literature review highlights the multifaceted journey of the evolution of deep learning architectures before 2018. From the foundational neural networks to the specialized architectures addressing specific challenges, each phase has contributed to the rich tapestry of deep learning. The collaborative spirit, influential contributions, and technological advancements underscore the dynamic nature of this evolution, setting the stage for subsequent innovations and advancements in the field.

III. Methodology:

The methodology employed in investigating the "Evolution of Deep Learning Architectures" before the year 2018 involves a comprehensive and systematic review of existing literature, research papers, and seminal works in the field of deep learning. The aim is to construct a chronological narrative that captures the key milestones, breakthroughs, and trends in the development of deep learning architectures leading up to the specified timeframe.

- **Literature Review:** Conduct an extensive literature review to identify and analyze relevant research papers, conference proceedings, and scholarly articles related to deep learning architectures. Utilize academic databases, digital libraries, and reputable journals to compile a comprehensive dataset of primary sources.

- **Identification of Key Architectures:** Identify and categorize key deep learning architectures that emerged before 2018. This involves a thorough examination of foundational models such as neural networks, convolutional neural networks (CNNs), recurrent neural networks (RNNs), autoencoders, and other specialized architectures. Classify these architectures based on their applications and impact on the field.
- **Chronological Organization:** Organize the identified architectures in a chronological order to establish a timeline that illustrates the evolution of deep learning. This timeline will serve as a framework for understanding the sequential development of architectures, highlighting significant advancements and innovations at each stage.
- **Analysis of Research Contributions:** Analyze the research contributions associated with each identified architecture. Examine seminal papers, research findings, and theoretical frameworks that have played a pivotal role in shaping and advancing deep learning architectures. Evaluate the impact of these contributions on the broader field.
- **Examination of Technological Advancements:** Investigate the technological advancements and computational resources available before 2018 that influenced the development and implementation of deep learning architectures. Consider hardware innovations, software frameworks, and other technological factors that contributed to the evolution of these architectures.
- **Inclusion of Comparative Analysis:** Include a comparative analysis of different deep learning architectures, highlighting their strengths, weaknesses, and unique contributions. Evaluate the performance metrics, applications, and challenges associated with each architecture to provide a holistic understanding of their evolution.
- **Consideration of Historical Context:** Integrate relevant historical context, including breakthroughs in AI research, societal influences, and the availability of datasets, which may have influenced the development

of deep learning architectures during the specified period.

By employing this methodology, the research aims to construct a comprehensive narrative that elucidates the evolution of deep learning architectures before 2018. The chronological and comparative analyses will provide insights into the historical context, technological landscape, and research contributions that have collectively shaped the field.

IV. Conclusion

In conclusion, as we delve into the examination of the "Evolution of Deep Learning Architectures" up until the year 2018, this study has explored a landscape characterized by significant advancements, transformative breakthroughs, and collaborative endeavors in the field of artificial intelligence. The historical journey began with the fundamental principles of artificial neural networks, navigating through obstacles and achievements that shaped the path of deep learning architectures. The rise of Convolutional Neural Networks (CNNs) brought about a revolution in computer vision, with groundbreaking models like LeNet-5 and AlexNet establishing new benchmarks in the realm of image recognition and classification. Simultaneously, the development of Recurrent Neural Networks (RNNs) and the introduction of Long Short-Term Memory (LSTM) networks addressed the intricacies of processing sequential data, leading to breakthroughs in natural language processing and time-series analysis. Specialized architectures, including autoencoders and Generative Adversarial Networks (GANs), introduced a layer of diversity and complexity to deep learning applications. These architectures expanded the scope of unsupervised learning, feature representation, and the generation of synthetic data, impacting fields such as anomaly detection and image synthesis. Throughout this exploration, collaborative research efforts played a crucial role, with influential contributions from various corners of academia, research laboratories, and industry. Landmark papers, such as "ImageNet Classification with Deep Convolutional Neural Networks" and "Sequence to Sequence Learning with Neural Networks," not only presented innovative architectures but also provided benchmarks that propelled the field's advancement.

As we ponder over the development prior to 2018, it becomes evident that every architectural advancement has not only tackled specific obstacles but has also collectively propelled deep learning to the forefront of artificial intelligence. The dynamic interplay of technological progress, contributions in research, and spirit of collaboration has fostered a vibrant ecosystem that continuously drives innovation and exploration. This research paper not only stands as a testament to the extensive past of deep learning structures, but also lays a groundwork for future inquiries. The insights gained from the development prior to 2018 shape our comprehension of the conquered challenges, the achieved milestones, and the promising avenues that lie before us in the ongoing pursuit of more intelligent and capable machine learning systems. As the field continues to progress, this historical context becomes an invaluable compass for researchers, practitioners, and enthusiasts alike.

References:

[1] LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11), 2278-2324.

[2] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. In Advances in neural information processing systems (pp. 1097-1105).

[3] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural computation, 9(8), 1735-1780.

[4] Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. Science, 313(5786), 504-507.

[5] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial nets. In Advances in neural information processing systems (pp. 2672-2680).

[6] Sutskever, I., Vinyals, O., & Le, Q. V. (2014). Sequence to sequence learning with neural networks. In Advances in neural information processing systems (pp. 3104-3112).

[7] LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., & Jackel, L. D. (1989). Backpropagation applied to handwritten zip code recognition. Neural computation, 1(4), 541-551.

[8] Graves, A., Mohamed, A. R., & Hinton, G. (2013). Speech recognition with deep recurrent

neural networks. In Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on (pp. 6645-6649). IEEE.

[9] Collobert, R., Weston, J., Bottou, L., Karlen, M., Kavukcuoglu, K., & Kuksa, P. (2011). Natural language processing (almost) from scratch. Journal of Machine Learning Research, 12(Aug), 2493-2537.

[10] Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.

[11] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., & Anguelov, D. (2015). Going deeper with convolutions. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1-9).

[12] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).

[13] R. K. Kaushik Anjali and D. Sharma, "Analyzing the Effect of Partial Shading on Performance of Grid Connected Solar PV System", 2018 3rd International Conference and Workshops on Recent Advances and Innovations in Engineering (ICRAIE), pp. 1-4, 2018.

[14] Kaushik, M. and Kumar, G. (2015) "Markovian Reliability Analysis for Software using Error Generation and Imperfect Debugging" International Multi Conference of Engineers and Computer Scientists 2015, vol. 1, pp. 507-510.

[15] Sharma R., Kumar G. (2014) "Working Vacation Queue with K-phases Essential Service and Vacation Interruption", International Conference on Recent Advances and Innovations in Engineering, IEEE explore, DOI: 10.1109/ICRAIE.2014.6909261, ISBN: 978-1-4799-4040-0.

[16] Sandeep Gupta, Prof R. K. Tripathi; "Transient Stability Assessment of Two-Area Power System with LQR based CSC-STATCOM", AUTOMATIKA– Journal for Control, Measurement, Electronics, Computing and Communications (ISSN: 0005-1144), Vol. 56(No.1), pp. 21-32, 2015.

[17] Sandeep Gupta, Prof R. K. Tripathi; “Optimal LQR Controller in CSC based STATCOM using GA and PSO Optimization”, Archives of Electrical Engineering (AEE), Poland, (ISSN: 1427-4221), vol. 63/3, pp. 469-487, 2014.

[18] V.P. Sharma, A. Singh, J. Sharma and A. Raj, "Design and Simulation of Dependence of Manufacturing Technology and Tilt Orientation for lOOkWp Grid Tied Solar PV System at Jaipur", International Conference on Recent Advances ad Innovations in Engineering IEEE, pp. 1-7, 2016.

[19] V. Jain, A. Singh, V. Chauhan, and A. Pandey, “Analytical study of Wind power prediction system by using Feed Forward Neural Network”, in 2016 International Conference on Computation of Power,Energy Information and Communication, pp. 303-306,2016.