IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Recent Trends In Kolmogorov-Arnold Networks

[1]Seema Sabharwal*, [2]Varun Chand H., [3]Pardeep Kumar, [4]Reema Gupta

[1,3,4]Assistant Professor, Department of Computer Science, Govt. P.G College for Women, Panchkula, Haryana, India

[2]Assistant Professor, Department of Computer Science and Engineering, College of Engineering, Perumon, Kollam, Kerala, India

Abstract: As we know, the machine learning field has been continuously evolving and Kolmogorov-Arnold Networks (KANs) have caught the attention of many researchers recently due to their performance and profound applications in mathematical and physics problems. These networks have been originated from the theorem of Kolmogorov Arnold and dominated traditional machine learning models. These networks have been equipped with dynamic adaptable and learnable univariate parameters and replacing traditional networks having fixed learnable weights. A lot of variants of KAN are already in use such as Graph KAN, Temporal KAN, Convolutional KAN, and Rational KAN, etc. With the objective of assisting the keen researchers to navigate this complex field of KAN, this article aims to provides in-depth study of various KAN architectures and discusses future potential applications of KAN in computer vision and other domains.

***Index Terms* - KAN, Kolmogorov Arnold Network, Deep Learning, Computer Vision.**

I. INTRODUCTION

The development of neural networks has revolutionized the field of machine learning, providing powerful tools for a wide range of applications, from image recognition to natural language processing. Among the various theoretical foundations that support the architecture of neural networks, the Kolmogorov–Arnold representation theorem stands out as a pivotal mathematical result. This theorem forms the basis for Kolmogorov–Arnold Networks (KANs), a class of neural networks designed to leverage the theorem's insights into function approximation [1]. The Kolmogorov-Arnold representation theorem has long been a cornerstone of mathematical analysis, offering a powerful framework for understanding the structure of multivariate functions. In recent years, this theorem has inspired the development of Kolmogorov-Arnold Networks (KANs), a class of neural networks designed to leverage the theorem's decomposition principles for improved efficiency and interpretability. Unlike traditional deep learning models, which often operate as black boxes, KANs aim to break down complex functions into simpler, univariate components, making them more transparent and computationally tractable. This article explores the recent trends in KAN research, highlighting key advancements, applications, and challenges in this rapidly evolving field.

II. LITERATURE REVIEW

In this section, we will highlight recent research works of KAN in various domains and how it has evolved over the years. [2] [3] proposed wavelet-based multilayered KAN for classing of high dimensionality hyperspectral images on datasets such as Indian Pines[4], Pavia[5], Salinas[6] and MNIST etc. [7] proposed DeepOKAN in the field of mechanics using Gaussian radial basis function along with KAN architecture to optimise the performance of neural operators. It has been observed that proposed model performed better that baseline DeepONet employing multilayer perceptron architecture and Fourier neural operator.
[8] proposed KAN with Graph learning to effectively approximate rather than traditional graph neural network (GNN). Two graph-based KAN architectures based on Graph Convolution Network (KAGCN) and Graph

Isomorphism Network (KAGIN) alongside B-splines and radial basis neural operators have been used with variations of Graphical KAN to improve the performance of proposed model.

Another GNN based Kan have been proposed by [9] Graph Kolmogorov–Arnold Network (GKAN) using spline based activation functions in deal with graphical structured information. The architecture comprises of KAN-Convolution layer, linear and output layer apart from spline activation function. Five benchmark datasets CORA, PUBMED, CITESEER, MUTAG, PROTEINS have been employed to train and test the proposed system using accuracy parameter.

[10] demonstrated the importance of KAN in computer vision domain by proposing KAN-Mixer on MNIST, CIFAR10, CIFAR100 datasets. The input image is divided into patches and linear transformation is applied on each patch followed by alternating token and channel mixing layer. The final output is calculated after applying various operations such as layer normalisation, mean and linear layer in output KAN layer. The model achieved highest accuracy of 98.16% on MNIST dataset with 64 channels with 132.12sec epoch time, 11.12 sec test time, 32 batch size.

[11] demonstrated the use of KAN in satellite image classification. A ConvNeXT [12] (a modernized CNN architecture based on the principles of Vision Transformer) has been used along with KAN architecture for classification of satellite images in remote sensing.

We can conclude from above studies than KAN architecture has found its applications in various domains. It has been analyzed

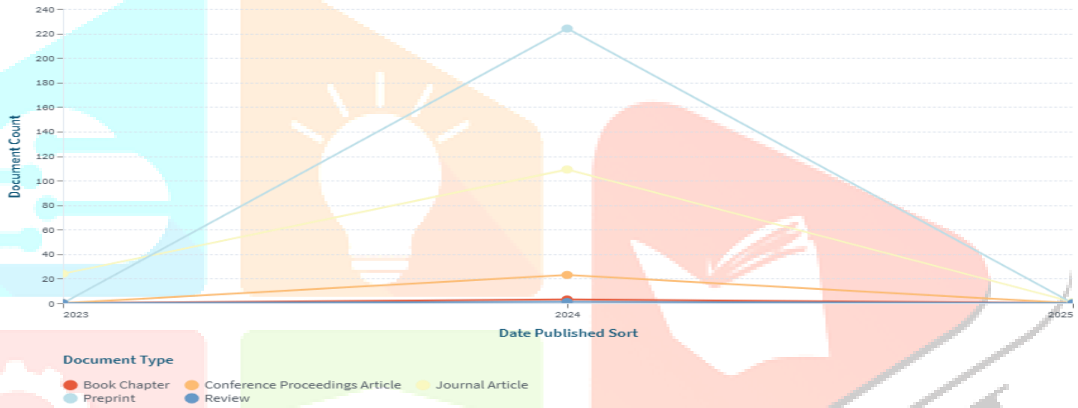

Figure 1 Various research papers published as document type

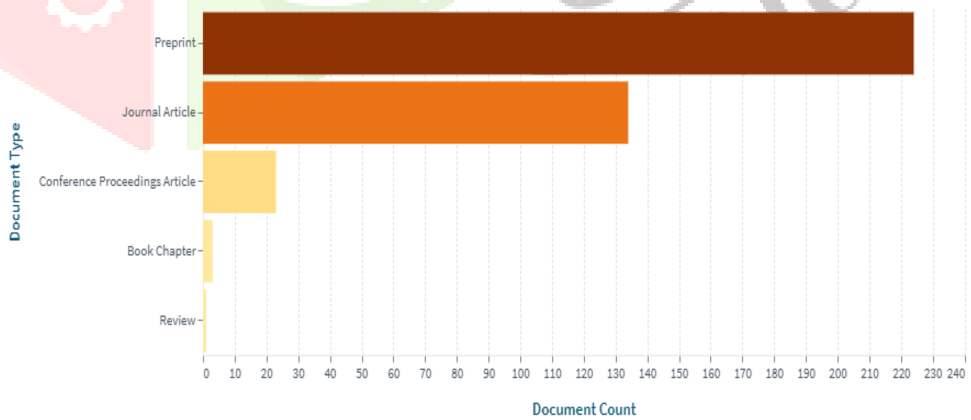

Figure 2. Document count of various research papers published

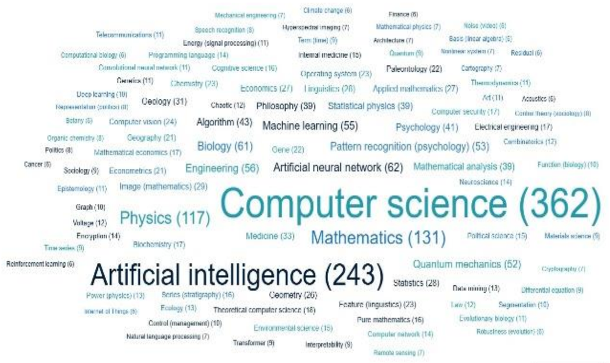

Figure 3. Published papers field of study

from Figure 1 and 2, that there has been profound increase in publications in the year 2024 than previous year. Figure 3 demonstrates the applications of KAN in various domains on the basis of various research articles. Further, it has been interpreted from Figure 3 that KAN has been used mostly in the field of computer science, artificial intelligence, physics, mathematics etc. [13].

III. BACKGROUND

In Kolmogorov-Arnold theorem, any continuous multivariate function can be represented as a finite superposition of continuous univariate functions and addition. Mathematically, this can be expressed as equation (1):

$$f(x) = \sum_{q=1}^{2n+1} \emptyset_q \left(\sum_{p=1}^{n} \varphi_{q,p}\left(x_p\right)\right) \quad (1)$$

where $\Phi_q\Phi_q$ are outer function comprising of single variable continuous function and mapping sum of transformed input thereby moulding the end result. And, ϕ_{pq} $\phi_{p,q}$ are continuous univariate functions mapping individual input variable x_p. Finite superposition

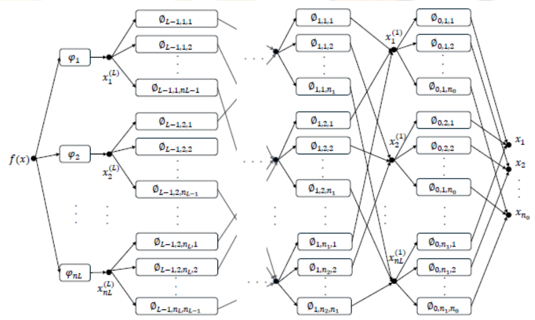

Figure 4- Architecture of KAN [14]

of continuous univariate functions has been represented with f(x), which transforms given input into results by applying summation operation on functions. This decomposition provides a theoretical basis for designing neural networks that can approximate high-dimensional functions with fewer parameters and greater interpretability. Due to continuous learnable univariate functions, KANs are now replacing traditional fixed learnable weights. The architecture of KAN has been demonstrated with the help of figure 4 with layers L. In case of multilayer implementation of KAN, Recent research has focused on translating this theoretical framework into practical machine learning models, leading to the development of KANs[15].

IV. RECENT ADVANCEMENTS

There is lot of work going on in this field such as

4.1 Hybrid Architectures

One of the most significant trends in KAN research is the development of hybrid architectures that combine KAN principles with existing deep learning frameworks. These hybrid models aim to leverage the strengths of both approaches, such as the expressive power of deep neural networks and the interpretability of KANs. Researchers have proposed integrating KAN-inspired layers into convolutional neural networks (CNNs) and recurrent neural networks (RNNs) to improve performance in tasks such as high-dimensional regression and time-series forecasting[16]. These hybrid models have shown promise in applications ranging from image processing to natural language processing, offering a balance between accuracy and transparency.

4.2 Applications in Scientific Machine Learning

KANs have also found its major applications in scientific machine learning, where interpretability and computational efficiency are critical. As KAN-inspired models have been used to solve partial differential equations (PDEs) and model physical systems, providing insights into the underlying structure of complex datasets [17]. In fields such as physics and engineering, KANs have been employed to approximate high-dimensional functions with fewer parameters, reducing computational costs while maintaining accuracy[18].

4.3 Interpretability and Explainability

A key advantage of KANs is their potential to improve the interpretability of machine learning models. By decomposing complex functions into simpler components, KANs offer a more transparent alternative to traditional black-box models. Recent studies have explored how KANs can provide insights into the decision-making process of neural networks, particularly in domains such as healthcare and finance [19]. KANs have been used to analyze medical data, enabling clinicians to understand the factors influencing a model's predictions and make more informed decisions.

V. CHALLENGES AND OPEN QUESTIONS

Despite the success the KAN, there are various challenges faces by researchers due to its complexity, some of them are discussed below:-

5.1 Computational Complexity

Despite their potential, KANs face several challenges, particularly in terms of computational complexity. Learning the univariate functions in practice can be computationally expensive, especially for high-dimensional data. Recent work has focused on developing efficient algorithms for training KAN-inspired models, such as gradient-based optimization techniques and parallel computing strategies.

5.2 Generalization to discontinuous functions

Another open question is how to generalize the Kolmogorov-Arnold theorem to non-smooth or discontinuous functions, which are common in real-world applications. While the theorem provides a powerful framework for continuous functions, extending it to more general cases remain an active area of research.

5.3 Scalability and Robustness

Scalability and robustness are also critical concerns for KANs. As the size and complexity of datasets continue to grow, researchers must develop scalable and robust KAN architectures that can handle large-scale problems without sacrificing accuracy or interpretability.

VI. CONCLUSION

Kolmogorov-Arnold Networks represent a promising direction in the quest for more interpretable and efficient machine learning models. By leveraging the Kolmogorov-Arnold representation theorem, KANs offer a powerful framework for approximating high-dimensional functions with fewer parameters and greater transparency. Recent advancements in hybrid architectures, scientific machine learning, and interpretability have demonstrated the potential of KANs to revolutionize fields such as healthcare, finance, and physics. However, significant challenges remain, including computational complexity, generalization to non-smooth functions, and scalability. Addressing these challenges will require continued innovation and collaboration across disciplines. As research in this area progresses, KANs are likely to play an increasingly important role in the development of next-generation machine learning models.

VII. FUTURE SCOPE

Future research on KANs is likely to focus on integrating these models with emerging technologies such as quantum computing and edge computing. Quantum-inspired algorithms could be used to accelerate the training of KANs, while edge computing could enable the deployment of KANs in resource-constrained environments. The potential applications of KANs are vast, ranging from healthcare and finance to climate modeling and robotics. Future work could explore how KANs can be applied to new domains, such as autonomous systems and generative modeling to address complex real-world problems. Finally, enhancing the interpretability of KANs will remain a key focus for researchers. By developing new techniques for visualizing and explaining the components of KANs, researchers can make these models more accessible and useful for practitioners in various fields.

REFERENCES

[1] Z. Liu *et al.*, "KAN: Kolmogorov-Arnold Networks," Feb. 09, 2025, *arXiv*: arXiv:2404.19756. doi: 10.48550/arXiv.2404.19756.

[2] S. T. Seydi, Z. Bozorgasl, and H. Chen, "Unveiling the Power of Wavelets: A Wavelet-based Kolmogorov-Arnold Network for Hyperspectral Image Classification," Oct. 08, 2024, *arXiv*: arXiv:2406.07869. doi: 10.48550/arXiv.2406.07869.

[3] Z. Bozorgasl and H. Chen, "Wav-KAN: Wavelet Kolmogorov-Arnold Networks," 2024, *arXiv*. doi: 10.48550/ARXIV.2405.12832.

[4] M. Baumgardner, L. Biehl, and D. Landgrebe, "220 Band AVIRIS Hyperspectral Image Data Set: June 12, 1992 Indian Pine Test Site 3." Purdue University Research Repository, 2015. doi: 10.4231/R7RX991C.

[5] G. Licciardi, P. R. Marpu, J. Chanussot, and J. A. Benediktsson, "Linear Versus Nonlinear PCA for the Classification of Hyperspectral Data Based on the Extended Morphological Profiles," *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 3, pp. 447–451, May 2012, doi: 10.1109/LGRS.2011.2172185.

[6] F. Melgani and L. Bruzzone, "Classification of hyperspectral remote sensing images with support vector machines," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 8, pp. 1778–1790, Aug. 2004, doi: 10.1109/TGRS.2004.831865.

[7] D. W. Abueidda, P. Pantidis, and M. E. Mobasher, "DeepOKAN: Deep operator network based on Kolmogorov Arnold networks for mechanics problems," *Comput. Methods Appl. Mech. Eng.*, vol. 436, p. 117699, Mar. 2025, doi: 10.1016/j.cma.2024.117699.

[8] R. Bresson, G. Nikolentzos, G. Panagopoulos, M. Chatzianastasis, J. Pang, and M. Vazirgiannis, "KAGNNs: Kolmogorov-Arnold Networks meet Graph Learning," Dec. 13, 2024, *arXiv*: arXiv:2406.18380. doi: 10.48550/arXiv.2406.18380.

[9] G. D. Carlo, A. Mastropietro, and A. Anagnostopoulos, "Kolmogorov-Arnold Graph Neural Networks," Jun. 26, 2024, *arXiv*: arXiv:2406.18354. doi: 10.48550/arXiv.2406.18354.

[10] M. Cheon, "Demonstrating the Efficacy of Kolmogorov-Arnold Networks in Vision Tasks," Jun. 21, 2024, *arXiv*: arXiv:2406.14916. doi: 10.48550/arXiv.2406.14916.

[11] M. Cheon, "Kolmogorov-Arnold Network for Satellite Image Classification in Remote Sensing," Jun. 02, 2024, *arXiv*: arXiv:2406.00600. doi: 10.48550/arXiv.2406.00600.

[12] Z. Liu, H. Mao, C.-Y. Wu, C. Feichtenhofer, T. Darrell, and S. Xie, "A ConvNet for the 2020s," Mar. 02, 2022, *arXiv*: arXiv:2201.03545. doi: 10.48550/arXiv.2201.03545.

[13] "Results The Lens - Free & Open Patent and Scholarly Search," The Lens - Free & Open Patent and Scholarly Search. Accessed: Sep. 18, 2023. [Online]. Available: https://www.lens.org/lens

[14] Y. Nagai and M. Okumura, "Kolmogorov--Arnold networks in molecular dynamics," Jul. 25, 2024, *arXiv*: arXiv:2407.17774. doi: 10.48550/arXiv.2407.17774.

[15] T. Ji, Y. Hou, and D. Zhang, "A Comprehensive Survey on Kolmogorov Arnold Networks (KAN)," Jan. 28, 2025, *arXiv*: arXiv:2407.11075. doi: 10.48550/arXiv.2407.11075.

[16] C. J. Vaca-Rubio, L. Blanco, R. Pereira, and M. Caus, "Kolmogorov-Arnold Networks (KANs) for Time Series Analysis," Sep. 25, 2024, *arXiv*: arXiv:2405.08790. doi: 10.48550/arXiv.2405.08790.

[17] A. Pal and S. Nagarajaiah, "KAN/MultKAN with Physics-Informed Spline fitting (KAN-PISF) for ordinary/partial differential equation discovery of nonlinear dynamic systems," Nov. 18, 2024, *arXiv*: arXiv:2411.11801. doi: 10.48550/arXiv.2411.11801.

[18] B. C. Koenig, S. Kim, and S. Deng, "KAN-ODEs: Kolmogorov–Arnold network ordinary differential equations for learning dynamical systems and hidden physics," *Comput. Methods Appl. Mech. Eng.*, vol. 432, p. 117397, Dec. 2024, doi: 10.1016/j.cma.2024.117397.

[19] S. Azimi, L. Spekking, and K. Staňková, “Kolmogorov-Arnold Networks and Evolutionary Game Theory for More Personalized Cancer Treatment,” Jan. 12, 2025, *arXiv*: arXiv:2501.07611. doi: 10.48550/arXiv.2501.07611.