Kolmogorov-Arnold Networks: Key Developments and Uses

Bochra Hadj Kilani[1]

1 University of Carthage

Funding: No specific funding was received for this work.

Potential competing interests: No potential competing interests to declare.

Abstract

Kolmogorov-Arnold Networks (KANs) have emerged as a promising advancement in the field of neural networks, offering enhanced interpretability, efficiency, and adaptability. This literature review explores various applications and developments of KANs. Temporal KANs (TKANs) combine the strengths of Recurrent Neural Networks (RNNs) and KANs for improved multi-step time series forecasting. DeepOKAN employs Gaussian radial basis functions (RBFs) in lieu of B-splines for computational mechanics, resulting in a notable acceleration and enhancement in efficiency. Wav-KANs utilize wavelet functions for data analysis, offering a balance between detail and overview. KANs have also been utilized in image classification, integrating with pre-trained Convolutional Neural Network (CNN) models for remote sensing scene classification tasks. The Variational Quantum Kolmogorov-Arnold Network (VQKAN) implements a quantum version of KAN on a quantum circuit, optimizing synaptic connection weights for improved efficiency. KANs have also been applied for Explainable Natural Language Processing (NLP), forming continuous word embeddings based on the meaning profiles of words. The review also mentions UKAN and IKAN, These advances underscore the versatility and potential of KANs in enhancing various machine learning applications. Future research will likely concentrate on further refinement of these architectures and exploring their suitability to diverse datasets and tasks.

Bochra Hadj KILANI

University of Carthage

[a] Email: hadjki@gmail.com

[b] ORCID iD: 0009-0000-8648-7670

Keywords: Kolmogorov-Arnold Networks (KANs), Time series forcasting, Image Classification, Natural Language Processing (NLP), Variational Quantum Kolmogorov-Arnold Network (VQKAN).

1. Introduction

The field of artificial intelligence has been undergoing rapid evolution, with new advancements and techniques being introduced on a regular basis. One such recent development is the introduction of Kolmogorov-Arnold Networks (KANs). Inspired by the Kolmogorov-Arnold representation theorem, KANs have emerged as a promising alternative to the traditional Multi-Layer Perceptrons (MLPs).

In contrast to MLPs, which have fixed activation functions on nodes (often referred to as "neurons"), KANs have learnable activation functions on edges (often referred to as "weights"). This implies that KANs lack any linear weights. Instead, every weight parameter is replaced by a univariate function parametrized as a spline.

The principal advantages of KANs include their interpretability, enhanced accuracy, efficiency, and adaptability. KANs are designed to be more expressive, meaning they can capture more complex patterns. Additionally, they are less likely to overfit, which means they are better at generalizing from the training data to unseen data. Furthermore, they are more interpretable, making it easier for us to understand how they are making their predictions. One of the key applications of KANs is in time series forecasting. The paper, titled "Kolmogorov-Arnold Networks (KANs) for Time Series Analysis," employs the adaptive activation functions of KANs to enhance predictive modeling. The paper demonstrates that KANs outperform conventional Multi-Layer Perceptrons (MLPs) in a real-world satellite traffic forecasting task. They provide more accurate results with considerably fewer number of learnable parameters.

Additionally, KANs have been utilized in image classification, integrating with pre-trained Convolutional Neural Network (CNN) models for remote sensing scene classification tasks. The Variational Quantum Kolmogorov-Arnold Network (VQKAN) implements a quantum version of KAN on a quantum circuit, optimizing synaptic connection weights for improved efficiency.

KANs have also been applied to the field of Explainable Natural Language Processing (NLP), with the objective of forming continuous word embeddings based on the meaning profiles of words. The research proposes a method for the formation of a continuous word embedding, whereby a basis is established for the expression of the meaning of words.

Structure of the paper: After providing an overview of the theoretical foundations of KANs, we delve into their unique advantages over traditional neural networks, including their interpretability, enhanced accuracy, efficiency, and adaptability. We then present a comprehensive review of the applications of KANs in various domains, including time series forecasting, image classification, natural language processing, and quantum computing. Furthermore, we examine the challenges and limitations of KANs and suggest future research avenues to enhance KAN architectures and expand their applications and the potential use of Kolmogorov-Arnold Networks (KANs) in urban research. The objective of this study is to provide a comprehensive overview of KANs and illustrate their potential to transform machine learning and artificial intelligence.

2. Background and Theoretical Foundations

The theoretical foundations of Kolmogorov-Arnold Networks (KANs) are anchored in the Kolmogorov-Arnold representation theorem. This theorem, initially proposed by Andrey Kolmogorov and subsequently refined by Vladimir Arnold, posits that any multivariate function can be represented as a superposition of univariate functions. This theorem serves as the foundation for the distinctive architectural design of KANs.

The theorem states that any multivariate continuous function can be represented as a superposition of continuous functions of one variable[1]. In simpler terms, it suggests that any complex function involving multiple variables can be broken down into simpler functions, each involving only one variable, and then combined using the binary operation of addition.

$$f\left(x_1, \ldots, x_n\right) = \sum_{q=1}^{2n+1} \Phi_q\left(\sum_{p=1}^{n} \phi_{q,p}\left(x_p\right)\right)$$

In this context, the multivariate function (complex recipe) (1,…,) f(x1,…, xn) is defined as , . The univariate functions (simple, one-ingredient recipes) ( )ϕq, p(xp) are combined by the function Φ Φq, which takes the univariate functions and combines them.

In simple words, this implies that any function, regardless of its complexity or the number of variables involved, can be decomposed into simpler functions that operate on a single variable at a time. This is the fundamental concept underlying the Kolmogorov-Arnold Networks (KANs).

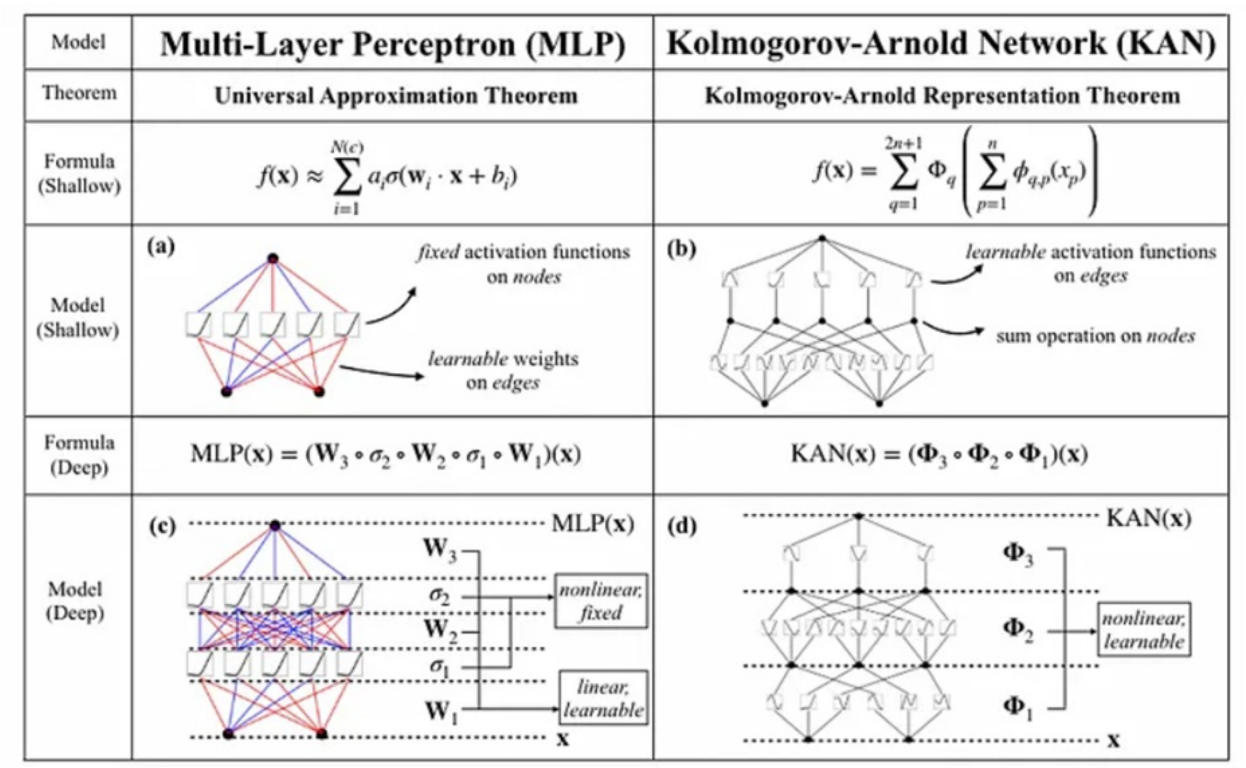

Figure 1. Comparison of the MLP and the KAN (Liu et al.,2024)

In deep learning, KANs is a class of neural network architectures that focus on the Kolmogorov-Arnold representation theorem (Bethell, 2024). In contrast to MLPs, which employ fixed activation functions at the node level (commonly referred to as "neurons"), KANs utilize learnable activation functions at the edge level (commonly referred to as "weights"). This implies that KANs lack any linear weights. In contrast, each weight parameter is substituted with a univariate function that is parameterized as a spline.

3. Advantages of KANs

Interpretability: One of the key advantages of Kolmogorov-Arnold Networks (KANs) is their interpretability. Unlike traditional neural networks, KANs can be intuitively visualized, making it easier to understand how they make predictions.

For example, let's consider a scenario where a KAN is used to fit a mathematical function or solve a partial differential equation[2]. In this case, the learned univariate functions in the KAN can be visualized and analyzed individually. This allows us to see how each function contributes to the overall output, providing a clear picture of the network's internal workings[3]

Moreover, KANs can interact with human users in a meaningful way For instance, in the field of symbolic regression, KANs can help scientists discover new mathematical laws. By visualizing and interpreting the learned functions in the KAN, scientists can gain insights into the underlying mathematical relationships in their data.

Enhanced Accuracy and Efficiency: KANs have demonstrated the capacity to provide enhanced accuracy in a variety of tasks, outperforming traditional neural networks in numerous instances. This is largely attributable to their distinctive architectural configuration, which enables them to capture intricate patterns and relationships in the data with greater efficacy. Additionally, KANs are also more efficient, as they are capable of achieving comparable or superior performance with fewer parameters compared to traditional neural networks.

Adaptability: KANs are highly adaptable. KANs can be applied to a wide range of tasks, including time series forecasting, image classification, natural language processing, and quantum computing. This adaptability makes KANs a versatile tool in the field of machine learning.

Comparison with Traditional Neural Networks: Compared to traditional neural networks like Multi-Layer Perceptrons (MLPs), KANs offer several advantages. As previously mentioned, they are more interpretable, accurate, efficient, and adaptable. Moreover, the use of learnable activation functions on edges in KANs represents a significant departure from traditional neural networks, which typically use fixed activation functions.

KANs represent a significant advancement in neural network architectures, offering several advantages over traditional models. Their interpretability, enhanced accuracy and efficiency, and adaptability make them a promising tool for a wide range of applications in machine learning and artificial intelligence.

4. Methodology

The Litmaps Knowledge Graph was employed in this study as a foundational tool for conducting a comprehensive literature review. Litmaps is a sophisticated platform that visualizes the interconnected landscape of academic research through a dynamic knowledge graph, thereby enabling researchers to explore the relationships between various scholarly articles.

Explanation of the Knowledge Graph:

The knowledge graph in Litmaps functions as a network where each node represents an academic article, and the edges between nodes symbolize the relationships between these articles. These relationships are typically based on citation data, where one article cites another, indicating a flow of knowledge and influence. This graphical representation provides a clear visualization of how research ideas evolve and interlink over time, offering a macroscopic view of the academic discourse within a particular field.

In the context of Knowledge-Aware Networks (KANs), the knowledge graph enables the identification of seminal papers, influential studies, and emerging trends. By examining the citation patterns between articles, it is possible to trace the development of key concepts and methodologies, as well as to gain insight into the hierarchical structure of knowledge within the domain.

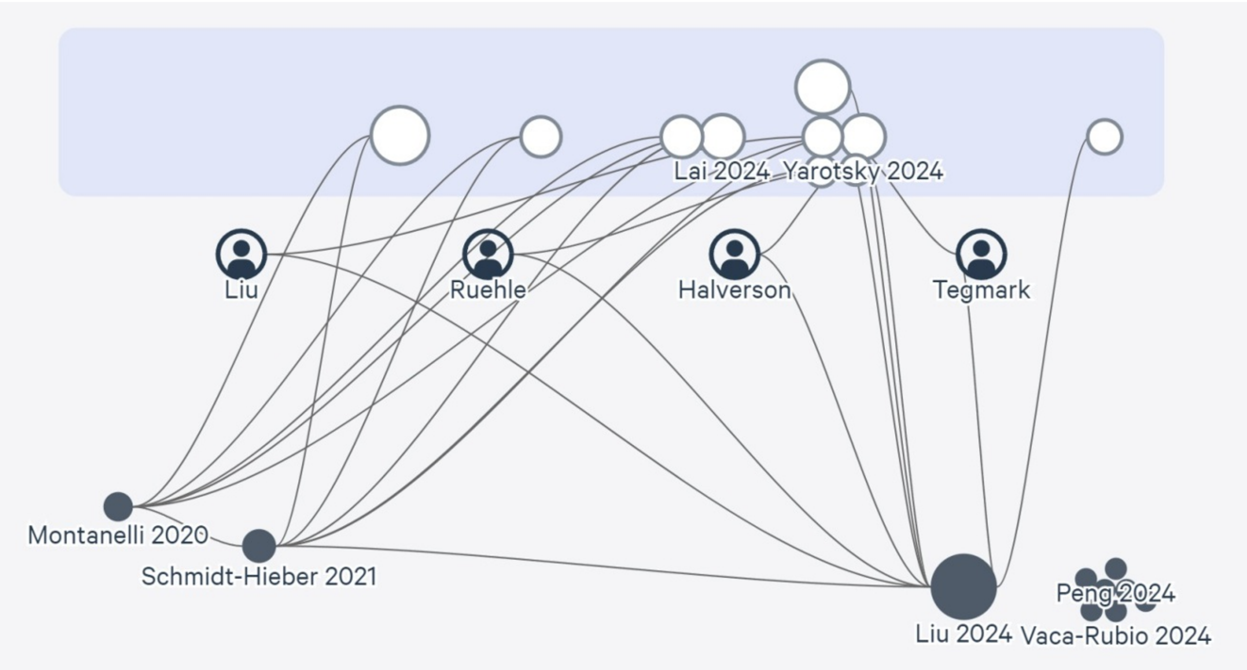

Figure 2. Knowledge graphs used for selection (keyword:KANs). Source: the author (2024)

In the context of Litmaps, the knowledge graph is constructed based on the citation relationships between academic articles. The knowledge graph enables users to ascertain which articles cite one another, the nature of their interconnectivity, and the clusters of articles that emerge around specific topics.

In selecting articles for the literature review, a total of 12 articles were chosen. The knowledge graph was then employed to identify key articles on Kolmogorov-Arnold Networks (KANs) and their various applications. By examining the citation relationships, it was possible to identify the most influential articles in this field, as well as more recent articles that build upon this foundational work.The knowledge graph also facilitated an understanding of the broader context of research on KANs. By examining the clusters of articles in the graph, it was possible to identify the key themes and trends in the research, as well as the areas where further research might be needed.

5. Applications of KANs

The Kolmogorov-Arnold Network (KAN) represents a novel approach to the construction of neural networks. The KAN is designed to offer enhanced performance compared to the Multi-Layer Perceptron (MLP), which is a prevalent component of deep learning models.

The KAN is designed to be more expressive, thereby enabling it to capture more complex patterns. Additionally, the KAN is less susceptible to overfitting, which enables it to generalize from the training data to unseen data more effectively. Moreover, it is more interpretable, facilitating a more comprehensive understanding of the underlying mechanisms by which it generates its predictions.

Multi-Layer Perceptrons (MLPs) are pervasive in the field of deep learning. For example, they are employed in transformer models such as GPT-2 and GPT-3, and it is likely that they will be used in GPT-4 as well. These models have had a profound impact on the field of machine learning. Therefore, enhancing the MLP could result in significant advancements within the field.

5.1. Time Series Forecasting

The authors (Cristian J. et al.,2024) leverage the adaptive activation functions of KANs for enhanced predictive modeling. Inspired by the Kolmogorov-Arnold representation theorem, KANs replace traditional linear weights with spline-parametrized univariate functions This allows KANs to learn activation patterns dynamically.

The paper demonstrates that KANs outperform conventional Multi-Layer Perceptrons (MLPs) in a real-world satellite traffic forecasting task

They provide more accurate results with considerably fewer number of learnable parameters[2]. The authors also provide an ablation study of KAN-specific parameters impact on performance.This research opens new avenues for adaptive forecasting models, emphasizing the potential of KANs as a powerful tool in predictive analytics. It’s a significant contribution to the field of time series analysis and forecasting, showcasing the potential of KANs to improve upon traditional methods.[4]

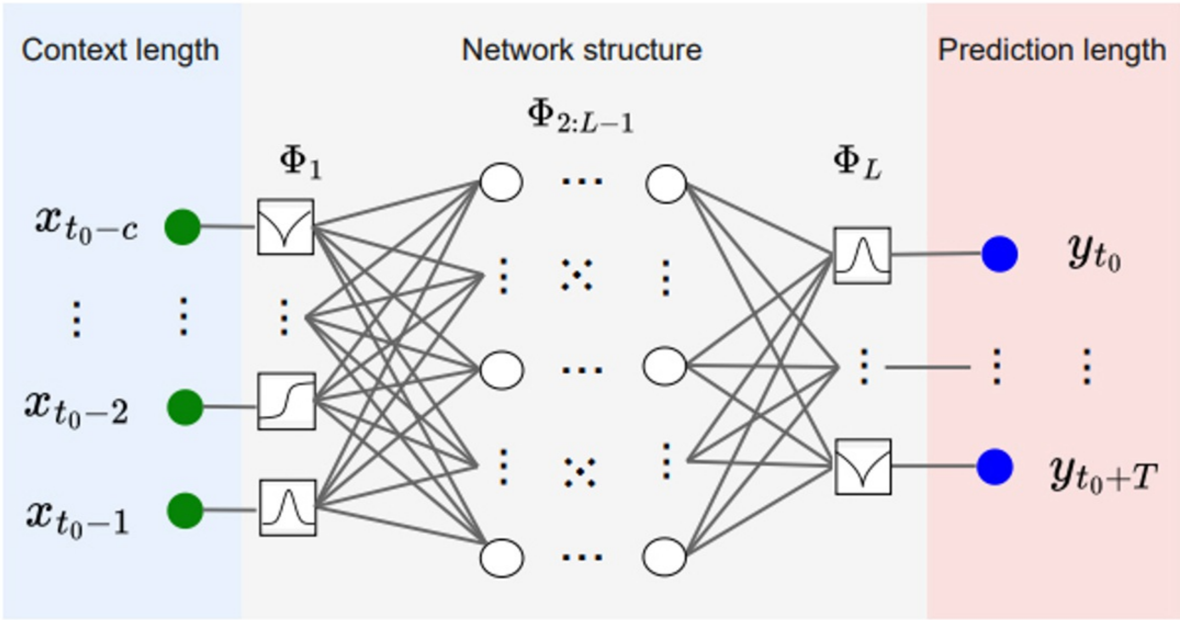

Figure 3. Example of the flow of information in the KAN network architecture for our traffic forecasting task. Learnable activations are represented inside a square box. (Cristian J.et al.,2024)

5.2. FastKAN

Li, Z. (2024). A preliminary investigation is required to ascertain whether the three-order B-splines utilized in Kolmogorov-Arnold Networks (KANs) can be adequately approximated by Gaussian radial basis functions. In essence, KANs represent a specific type of neural network, and they typically employ three-order B-splines. However, in this paper, Li demonstrates that Gaussian radial basis functions can be employed in lieu of the 3-order B-splines, with the result that the KANs will function identically.

This results in the development of a novel variant of KANs, designated as FastKAN. As the name implies, FastKAN is considerably more expeditious than the original KANs. Moreover, FastKAN is a type of radial basis function (RBF) network, which is another kind of neural network. This paper thus concerns the acceleration and enhancement of KANs through the utilisation of a distinct functional approach.

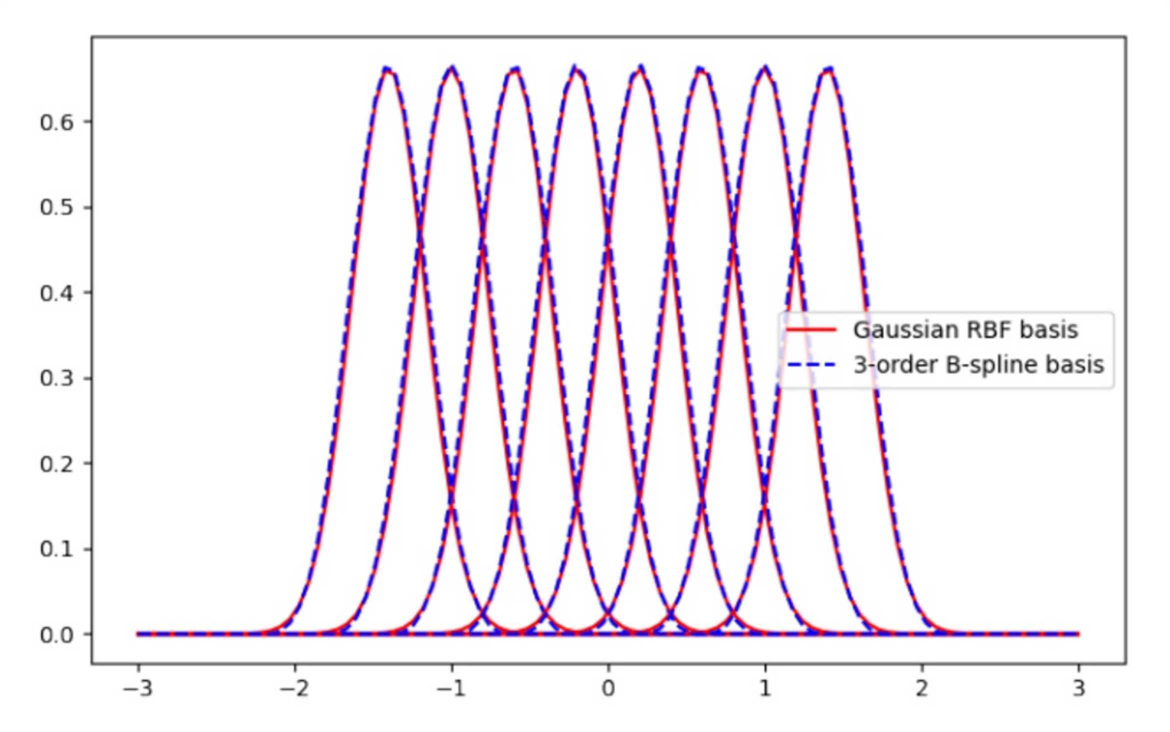

Figure 4. Under proper linear transformations, Gaussian RBFs well approximate 3-order B-spline bases with good precision (Li, Z.,2024).

5.3. Wav-KANs

Wav-KAN, a new type of neural network (Bozorgasl, Z., & Chen, H. 2024) that improves how machines learn and understand data. Traditional learning models, like MLPs, have some issues with being clear, fast, and efficient. Wav-KAN solves these by using wavelet functions, which are mathematical tools that break down data into different frequency components. This allows the network to focus on both the important details and the overall trends in the data.

- Wavelets: They work like a zoom lens, capturing both the big picture and the fine details.
- Efficiency: Wav-KAN is faster and smarter in learning from data, avoiding unnecessary complexity.
- Adaptability: It adjusts to the data's structure, much like water takes the shape of its container.
- Potential: This approach could lead to better and easier-to-understand AI systems across various industries.

The Wavelet Kolmogorov-Arnold Networks (Wav-KAN) signifies a substantial progression in the construction of comprehensible neural networks. Its proficiency in managing high-dimensional data and offering lucid insights into the model's behavior positions it as a promising instrument for an extensive array of applications, spanning from scientific investigation to industrial implementation.

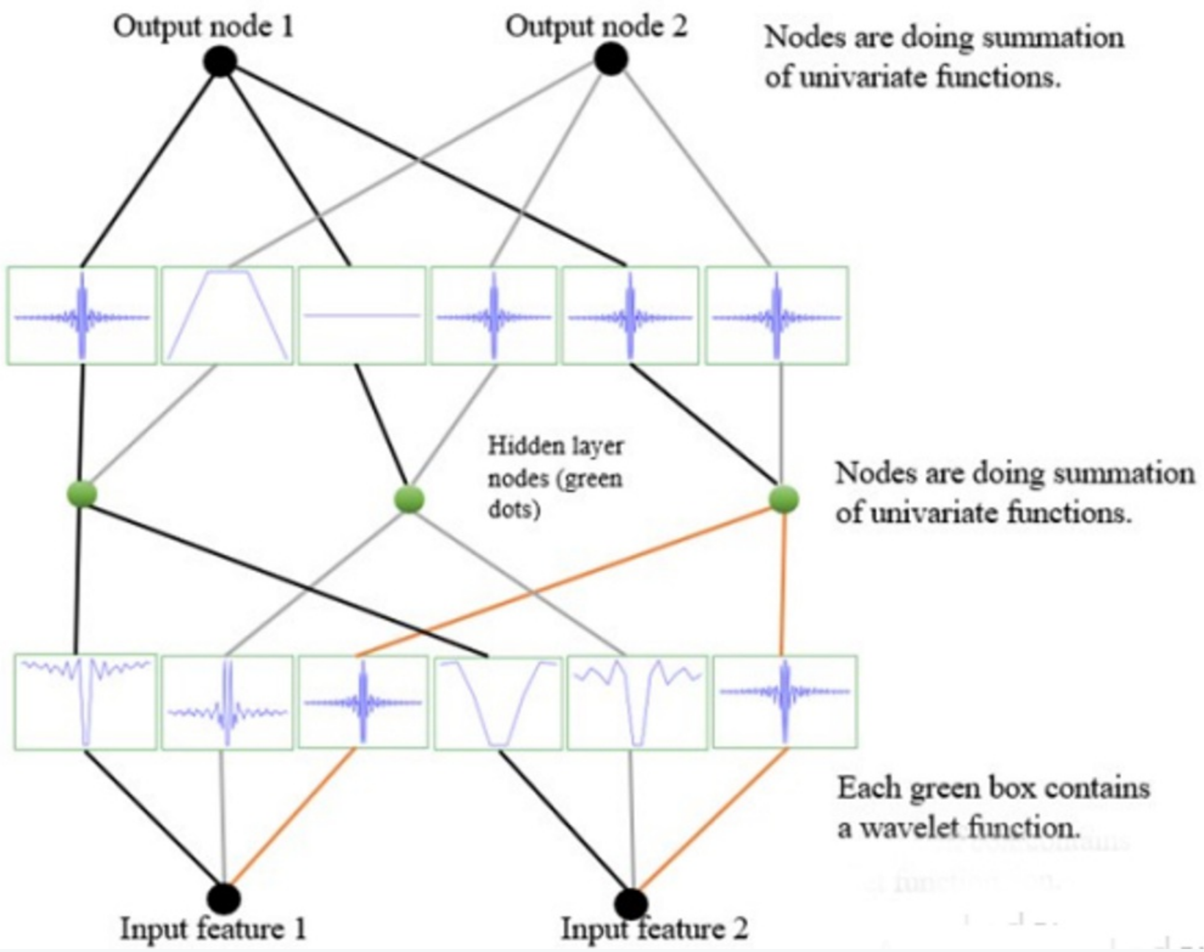

Figure 5. Wav-KAN with arbitrary number of layers (Bozorgasl, Z., & Chen, H.,2024)

A neural network architecture that can have an arbitrary number of layers. This flexibility allows Wav-KAN to adapt to various data structures and complexities, enhancing its accuracy, training speed, and robustness. The graph in the paper likely visualizes the structure of a Wav-KAN with a specific number of input features, hidden nodes, and output nodes. This illustration may demonstrate how data flows through the network and how the wavelet functions in the network help capture both high-frequency and low-frequency components of the input data efficiently. Future endeavors will concentrate on further refinement of the Wav-KAN architecture, probing its suitability to diverse datasets and tasks, and incorporating the framework into prevalent machine learning libraries such as PyTorch and TensorFlow.

In essence, Wav-KAN emerges as a potent and adaptable model, laying the groundwork for the evolution of more transparent and efficient neural network architectures. Its capacity to amalgamate high performance with interpretability signifies a pivotal advancement in the realm of artificial intelligence. This amalgamation underscores the potential of Wav-KAN to revolutionize the landscape of neural network design, thereby marking a critical stride in the field of artificial intelligence.

5.4. TKAN: Temporal Kolmogorov-Arnold Networks

Genet, R., & Inzirillo, H. (2024), presents a paper on a novel neural network architecture called Temporal Kolmogorov-

Arnold Networks (TKANs), which combines the strengths of Recurrent Neural Networks (RNNs) and Kolmogorov-Arnold Networks (KANs) to enhance multi-step time series forecasting.

Temporal Kolmogorov-Arnold Networks (TKANs) are a novel neural network architecture that combines the strengths of Recurrent Neural Networks (RNNs) and Kolmogorov-Arnold Networks (KANs).

Compared to traditional RNNs, TKANs offer several advantages:

Table 1. Table summarizing the key features of Temporal Kolmogorov-Arnold Networks (TKANs). (Author,2024)

Feature	Description
Enhanced Accuracy and Efficiency	TKANs are designed for multi-step time series forecasting, offering improved accuracy and efficiency. They overcome the limitations of traditional models in handling complex sequential patterns.
Memory Management	TKANs incorporate memory management into their architecture, which is crucial for processing sequential data. This feature enables TKANs to capture long-term dependencies in the data.
Adaptability	TKANs adapt to the structure of the data, much like how water conforms to the shape of its container. This results in faster training speeds and increased robustness compared to traditional RNNs.
Interpretability	TKANs, inspired by KANs, apply activation functions on the connections between nodes. These functions can learn and adapt during training, making TKANs more interpretable than traditional RNNs.

5.5. DEEPOKAN

DeepOKAN represents a novel iteration of neural operators designed for computational mechanics. In contrast to conventional neural network architectures, DeepOKAN employs Kolmogorov-Arnold networks (KANs).

The distinctive feature of DeepOKAN is its utilization of Gaussian radial basis functions (RBFs) in lieu of B-splines. This modification has led to a notable acceleration in computational speed, while maintaining the model's efficacy and precision.

DeepOKAN is employed to construct surrogates for diverse mechanics problems. It has been observed that DeepOKANs require a smaller number of learnable parameters than current MLP-based DeepONets to achieve comparable accuracy.

This approach should facilitate further enhancements to the performance of neural operators. This represents a promising development in the field of computational mechanics and machine learning.

This approach is particularly useful in modern digital engineering design, which often requires costly repeated simulations for different scenarios. The capacity of neural networks to make predictions renders them suitable surrogates for providing design insights. Nevertheless, only a select few neural networks are capable of efficiently handling complex engineering scenario predictions. DeepOKAN is one such network (Abueidda, D. W. et al., 2024).

5.6. KANs in image classification

Cheon (2024) proposes a novel approach for integrating the Kolmogorov-Arnold Network (KAN) with various pre-trained Convolutional Neural Network (CNN) models for remote sensing (RS) scene classification tasks using the EuroSAT dataset. The methodology, designated as KCN, is intended to supplant traditional Multi-Layer Perceptrons (MLPs) with KAN in order to enhance classification performance.

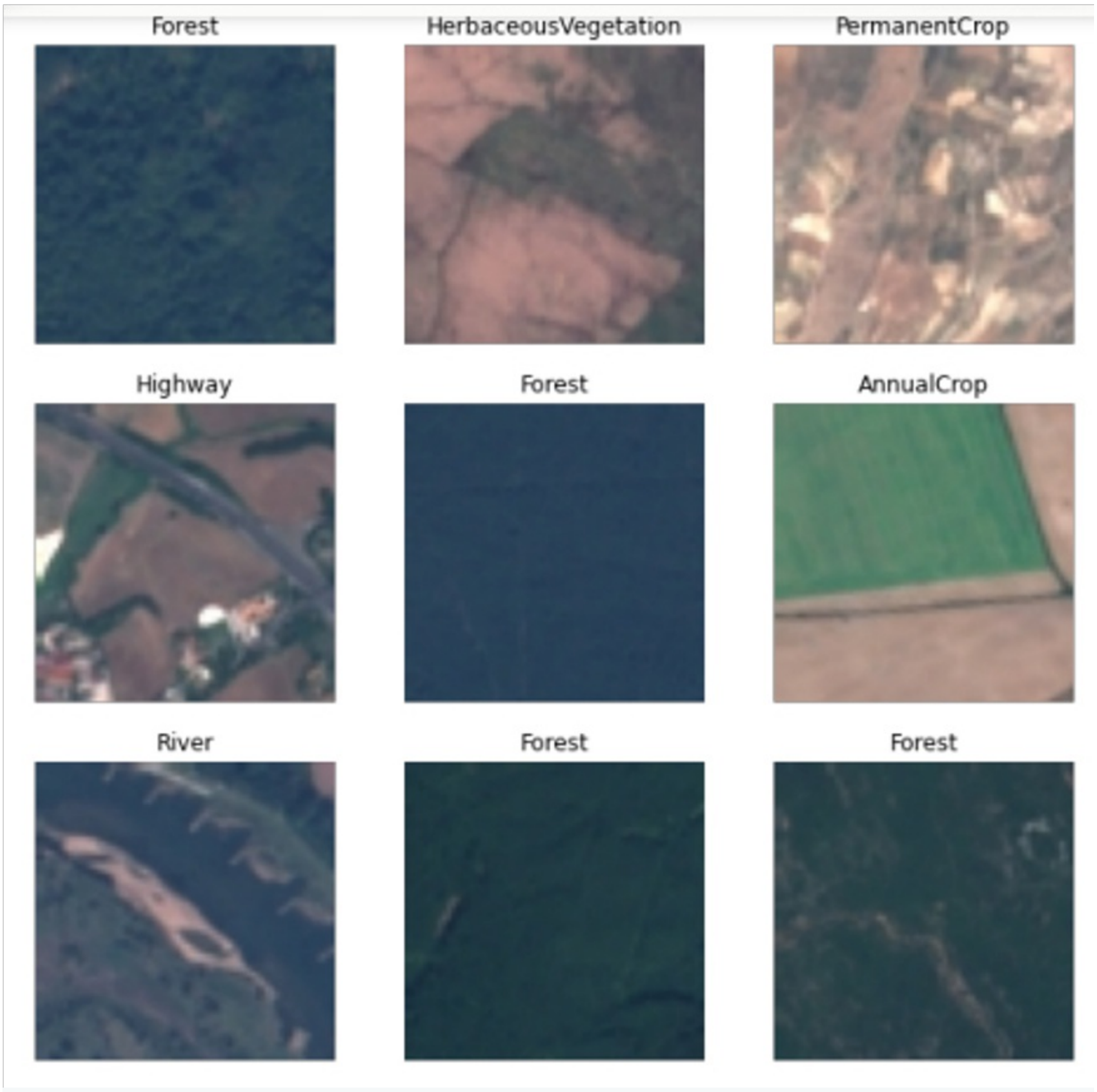

Figure 6. Example datasets from the EuroSAT used in the experiment: Cheon, M. (2024)

A variety of CNN-based models, including VGG16, MobileNetV2, EfficientNet, ConvNeXt, ResNet101, and Vision Transformer (ViT), were employed and their performance was evaluated when paired with KAN. The experiments demonstrated that KAN achieved high accuracy with fewer training epochs and parameters than other models. In particular, the ConvNeXt model paired with KAN demonstrated the most favorable performance, achieving 94% accuracy in the initial epoch, which increased to 96% and remained consistent across subsequent epochs. (Cheon,2024)

The results indicated that both KAN and MLP achieved similar levels of accuracy, with KAN demonstrating a slight advantage in later epochs. The utilization of the EuroSAT dataset provided a robust testbed for the investigation of the suitability of KAN for remote sensing classification tasks. Given that KAN is a novel algorithm, there is substantial potential for further development and optimization. This indicates that KCN represents a promising alternative for efficient image analysis in the field of remote sensing.

This research is particularly pertinent in the context of the rapid development of satellite technology and the proliferation of high-resolution remote sensing (RS) images. The integration of data from multiple satellites into time series has facilitated the detection of specific events and trends, thereby enhancing our understanding of global climate patterns and ecological changes. Deep learning models have demonstrated remarkable success in a variety of applications, including land cover classification, object detection, and environmental monitoring.

The integration of advanced deep learning techniques, including convolutional neural networks and transformers, is anticipated to further enhance the analysis and interpretation of remote sensing data. It is anticipated that these advancements will result in a notable enhancement in the accuracy and scope of remote sensing applications. (Cheon,2024)

5.7. Variational Quantum Kolmogorov-Arnold Network

Wakaura, H., & Suksmono, A. B. (2024) proposes a quantum version of the Kolmogorov-Arnold Network (KAN) implemented on a quantum circuit, named the Variational Quantum Kolmogorov-Arnold Network (VQKAN).

The KAN is a multi-layer network, analogous to a traditional neural network, with each layer represented by a matrix containing parameters and feedback. This structure is linear and mirrors the form of variational quantum algorithms.

The VQKAN optimizes the weight of each synaptic connection, thereby enhancing the overall performance of the neurons. This approach offers a significant enhancement in efficiency over traditional neural networks, as it can be optimized with greater precision and requires only a few neurons to solve complex problems.

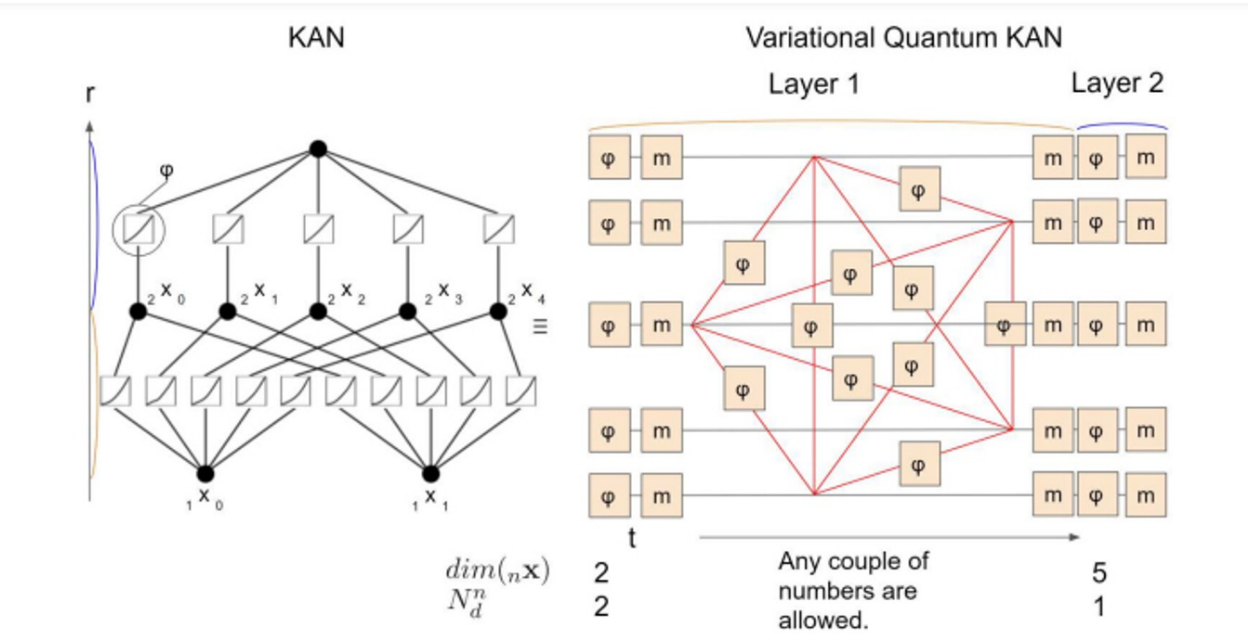

Figure 7. The abstract picture of VQKAN; Left; classical KAN

The research demonstrated the optimization of VQKAN for the fitting of the given function in various ways, including the use of ansatzes and comparisons. This suggests that VQKAN could be a promising approach for efficient image analysis in the field of quantum computing and machine learning.

5.8. KAN for Explainable NLP

Galitsky, B. A. (2024) discusses the application of the Kolmogorov-Arnold Network (KAN) for Explainable Natural Language Processing (NLP), with a particular focus on the context of continuous word embeddings.

Word embeddings are dense vector representations of words in a vector space, where each dimension of the vector captures a distinct aspect of the word's meaning. The generation of word embeddings is a common practice in the field of natural language processing. Popular methods include Word2Vec, GloVe (Global Vectors for Word Representation), and transformer-based models such as BERT (Bidirectional Encoder Representations from Transformers) and GPT (Generative Pre-trained Transformer).

The research proposes a method for forming a continuous word embedding by establishing a basis for expressing the meaning of words. In order to express the meaning of a noun, a profile is constructed based on the verbs that are applicable to this noun. Similarly, to express the meaning of a verb, its profile is constructed on the basis of the nouns to which it is applicable.

The profile for a word is a sequence of distances to the elements of the basis. For words that are representative of a wide range of contexts, such as "human," the profile can be constructed using a set of representative verbs, such as "move," "walk," "sit," "hike," and "jump." A typical profile is a monotonic decreasing function. This function is continuous, as the elements of the basis possess "intermediate" meanings situated between the meanings of the words of the basis. These intermediate meanings correspond to "fictitious" words positioned between the meanings of the words of the basis.

The research also discusses a verb-noun word2vec distance matrix, which is a list of verbs and corresponding nouns. An example of this is the list of verbs and corresponding nouns for the verb "run" and the noun "runner." The verbs are ordered by a distance to the head verb, with the lowest distance representing the most proximal verbs and the highest distance representing the most distal verbs within a given semantic family of verbs for movement types.

This approach to forming continuous word embeddings using KAN provides a novel method for capturing the semantic relationships between words in a continuous vector space, thereby enabling NLP models to effectively comprehend and process natural language text.

5.9. UKAN: for medical image segmentation and generation

The U-KAN approach to medical image segmentation and generation represents a novel methodology.

U-KAN represents a modification of the U-Net architecture, which has been established as a cornerstone in various visual applications, including image segmentation and diffusion probability models (Li, C. et al., 2024). U-KAN integrates

Kolmogorov-Arnold Networks (KANs) into the U-Net pipeline. KANs are inspired by the Kolmogorov-Arnold representation theorem and reshape neural network learning through the stack of non-linear learnable activation functions. In contrast to Multi-Layer Perceptrons (MLPs), which employ fixed activation functions at the node level, KANs utilize learnable activation functions at the edge level. This distinction renders KANs more accurate and interpretable than MLPs.

In the U-KAN model, the established U-Net pipeline is modified and redesigned by integrating dedicated KAN layers on the tokenized intermediate representation. The superiority of U-KAN is validated by rigorous medical image segmentation benchmarks, which demonstrate that it achieves higher accuracy with less computation cost.

Furthermore, U-KAN has been investigated as an alternative U-Net noise predictor in diffusion models, thereby demonstrating its applicability in generating task-oriented model architectures.

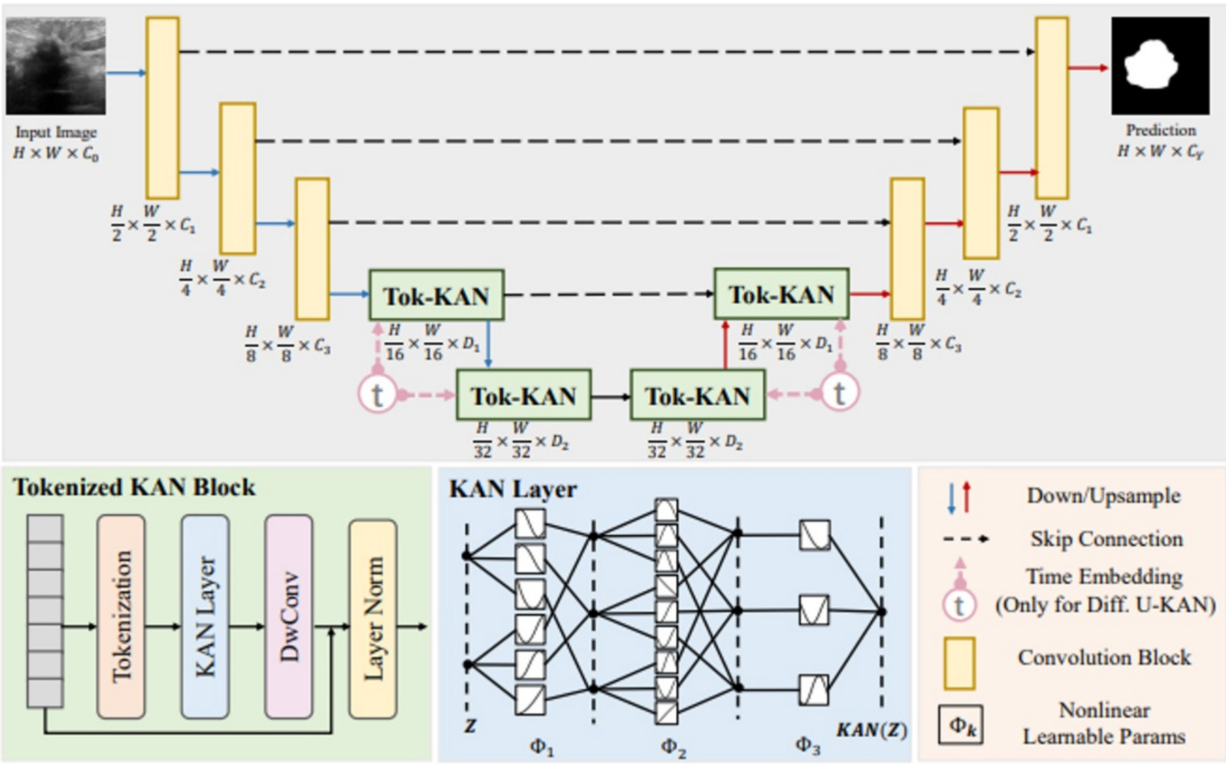

Figure 8. Overview of U-KAN pipeline. After feature extraction by several convolution blocks in Convolution Phrase, the intermediate maps are tokenized and processed by stacked Tok-KAN blocks in Tokenized KAN Phrase. The time embedding is only injected into the KAN blocks when applied for Diffusion U-KAN. (Li, C et al., 2024)

In the context of diffusion models, these are sophisticated machine learning algorithms that generate high-quality data by gradually introducing noise into a dataset and then learning to reverse this process. They are employed to generate novel, distinctive images, sounds, or other data types. They represent a highly effective tool within the generative model toolkit, offering a versatile approach to the generation of novel, high-quality data samples derived from learned patterns within existing data.

U-KAN represents a promising alternative to MLPs, offering the potential for further enhancement of contemporary deep learning models that rely heavily on MLPs. It serves as a robust foundation for medical image segmentation and generation. Furthermore, it is a promising alternative to the U-Net noise predictor in diffusion models. The adaptability and effectiveness of U-KAN also highlight its potential as a superior alternative to U-Net for noise prediction in diffusion models. These findings underscore the importance of exploring non-traditional network structures like KANs for advancing a broader range of vision applications.

5.10. IKAN

iKAN (Liu et al., 2024) is a novel incremental learning (IL) framework for wearable sensor human activity recognition (HAR) that addresses two challenges simultaneously: catastrophic forgetting and non-uniform inputs. iKAN pioneers IL with Kolmogorov-Arnold Networks (KANs) to replace multi-layer perceptrons as the classifier. The framework leverages the local plasticity and global stability of splines. To adapt KAN for HAR, iKAN employs task-specific feature branches and a feature redistribution layer. In contrast to existing IL methods, which primarily adjust the output dimension or the number of classifier nodes to adapt to new tasks, iKAN focuses on expanding the feature extraction branches to accommodate new inputs from different sensor modalities while maintaining consistent dimensions and the number of classifier outputs.

The iKAN framework demonstrated incremental learning performance across six public HAR datasets. The last performance was 84.9% (weighted F1 score), with an average incremental performance of 81.34%. This significantly outperforms the two existing incremental learning methods, namely EWC (51.42%) and experience replay (59.92%).The iKAN approach represents a promising avenue for incremental learning in HAR scenarios, offering a solution to the challenges of catastrophic forgetting and non-uniform inputs. This is the inaugural study to utilise KAN for IL. Through empirical investigation, the optimal KAN configurations for the selected HAR tasks were identified.

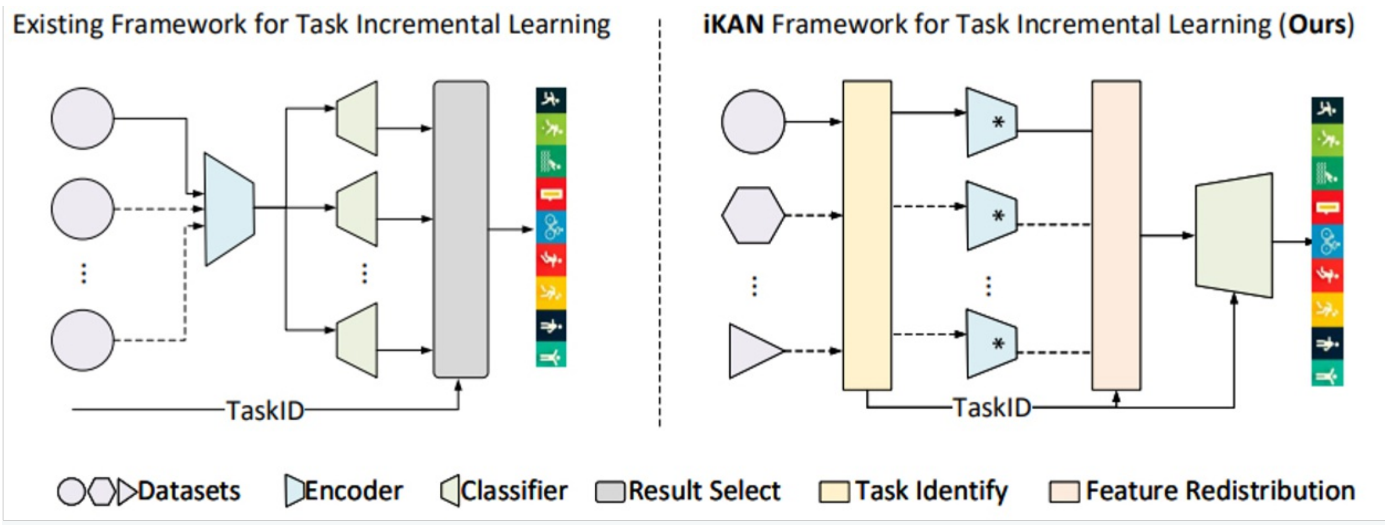

Figure 9. Comparison between the existing framework for task incremental learning and the proposed iKAN framework.

The existing framework has one encoder and increased classifiers to incrementally learn the new tasks, which can only process the task with the same input, while the iKAN framework has multiple encoders aiming to receive non-uniform inputs and one KAN-based classifier. Thus iKAN can incrementally learn the tasks across heterogeneous datasets. The shape of the datasets(tasks) represents the different sensor modalities in the datasets(tasks)) (Liu et al., 2024)

6. Challenges and Limitations

The literature review presents a comprehensive exploration of Kolmogorov-Arnold Networks (KANs), a novel neural network architecture that has emerged as a promising alternative to traditional Multi-Layer Perceptrons (MLPs). The Kolmogorov-Arnold representation theorem inspired the development of KANs, which replace linear weights with spline-parametrized univariate functions. This allows the networks to learn activation patterns dynamically.

The review discusses various applications and advancements of KANs, emphasizing their versatility and potential to enhance machine learning applications. The literature review demonstrates the innovative nature of KANs and their potential to revolutionize various domains of machine learning. Their interpretability and efficiency render them a valuable addition to the toolkit of machine learning algorithms, particularly for tasks involving large datasets and complex nonlinear relationships.

Table 2. comparison table of different uses of Kolmogorov-Arnold Networks (KANs), their domains, advantages, and disadvantages: the author (2024)

Use of KANs	Domain	Advantages	Disadvantages
Time Series Forecasting	Machine Learning, Predictive Analytics	Enhanced predictive modeling, dynamic learning of activation patterns, outperforms MLPs in terms of accuracy and interpretability	Requires careful optimization of node counts and grid sizes
Data Fitting and PDE Solving	Mathematics, Physics	Can achieve comparable or better accuracy than larger MLPs, useful for helping scientists (re) discover mathematical and physical laws	As a new model, further research is needed to explore their robustness across diverse datasets
Image Classification	Computer Vision, Remote Sensing	Integrates with pre-trained CNN models for remote sensing scene classification tasks, achieves high accuracy with fewer training epochs and parameters	As a novel algorithm, there is substantial capacity for further development and optimization
Medical Image Segmentation	Medical Imaging, Computer Vision	Enhances the established U-Net pipeline by integrating dedicated KAN layers, achieves higher accuracy with less computation cost	Further investigation is needed to develop more complex solutions that can compete with advanced architectures
Non linear fonction approximation	Mathematics, Machine learning	Accurate approximation of complex nonlinear functions, outperforms MLPs in terms of accuracy and interpretability	As a new model, further research is needed to explore their robustness across diverse datasets

Despite these encouraging outcomes, KANs remain a nascent model. Further research is necessary to investigate their resilience across a range of datasets and their compatibility with other deep learning architectures. As we continue to investigate their resilience and compatibility with other deep learning architectures, KANs are positioned to facilitate the development of novel adaptive forecasting models.

Implementing Kolmogorov-Arnold Networks (KANs) can present several challenges:

Table 3. Different challenges in Implementing Kolmogorov-Arnold Networks (KANs): the author (2024)

Challenges in Implementing KANs	Description
Complexity	KANs replace linear weights with spline-parametrized univariate functions, increasing the complexity of the model and making the training process more computationally intensive and time-consuming.
Parameter Tuning	The performance of KANs can be highly sensitive to the choice of parameters, such as the number of nodes and the grid size. Finding the optimal set of parameters can be challenging.
Overfitting	While KANs are less prone to overfitting compared to traditional neural networks, they can still suffer from this problem if not properly regularized, especially when dealing with high-dimensional data or complex tasks.
Interpretability	Although KANs are more interpretable than traditional neural networks, understanding their internal workings can still be challenging, especially for complex models.
Lack of Pretrained Models	Unlike popular architectures like CNNs or RNNs, there are not many pretrained KAN models available. This means that researchers often have to train their models from scratch, which can be time-consuming and resource-intensive.
Limited Research	As a relatively new field, there is still much to learn about KANs. There are not as many resources or studies available compared to more established neural network architectures.

7. Future Research Directions

Kolmogorov-Arnold Networks (KANs) are particularly adept at analyzing complex, nonlinear data relationships, and thus have a wide range of applicability across various domains. From the field of financial forecasting to the realm of personalized medicine, KANs provide predictive insights and optimization solutions.

In the field of healthcare, KANs have the potential to revolutionize genomics analysis and drug discovery. With their versatility, KANs redefine data analytics across diverse fields, offering invaluable tools for informed decision-making and optimized processes.

Potential Use of Kolmogorov-Arnold Networks (KANs) in Urban Research:

Kolmogorov-Arnold Networks (KANs) offer a versatile and powerful tool for urban research, which involves analyzing complex, multifaceted data across various domains such as traffic management, environmental monitoring, infrastructure maintenance, and public health.

Smart city the Internet of Things (IoT):

In the context of rapid urbanization, the concept of smart cities has emerged as a promising solution to address the complex challenges facing modern urban environments. At the core of smart city development lies the integration of advanced technologies, notably the Internet of Things (IoT), with the objective of enhancing the efficiency, sustainability, and livability of cities. Among these technologies, Kolmogorov-Arnold Networks (KANs) offer a powerful toolset for analyzing data from IoT sensors and optimizing urban services.

KANs facilitate the ability of smart cities to address a multitude of urban challenges through their capacity to process and interpret voluminous and intricate datasets generated by IoT sensors.

By employing KANs, city planners and policymakers are able to gain insights into various aspects of urban life, including transportation patterns, environmental conditions, energy usage, and public health. This data-driven approach enables informed decision-making and proactive management of city resources.

One of the primary applications of KANs in smart cities is predictive analytics. By employing historical data, KANs are able to forecast future trends and events. This enables predictive maintenance of critical infrastructure, optimization of resource allocation, and proactive measures to address emerging issues such as traffic congestion and energy demand fluctuations. This predictive capability empowers city authorities to implement proactive strategies to mitigate potential disruptions and enhance urban resilience.

Another area where KANs demonstrate significant potential is in the optimization of urban services. By analyzing data on transportation flows, energy consumption patterns, waste generation rates, and service usage metrics, KANs facilitate the optimization of urban service delivery. From the optimization of public transportation routes and scheduling of waste collection to the dynamic adjustment of energy distribution networks, KANs contribute to the efficient and effective delivery of essential services to city residents.

Table 4. Potential Use of Kolmogorov-Arnold Networks (KANs) in Urban Research: the author (2024)

Domain	Application	Description	Advantages
Traffic and Mobility	Multi-step traffic forecasting	TKANs improve the accuracy and efficiency of predicting traffic patterns and mobility trends.	Enhanced accuracy, fewer learnable parameters, improved congestion management.
Environmental Monitoring	Real-time environmental analysis	Wav-KANs use wavelet functions to analyze remote sensing data, improving the detection and classification of environmental changes.	Efficient data analysis, balances detail and overview, adaptable to various environmental datasets.
Infrastructure Management	Predictive maintenance	KANs model complex relationships in infrastructure data to predict wear and tear, optimizing maintenance and resource allocation.	Accurate predictions, enhanced efficiency, reduced maintenance costs.
Urban Health Studies	Medical image segmentation and analysis	U-KAN enhances U-Net architecture for improved segmentation and analysis of medical images, supporting urban health initiatives.	Higher accuracy, lower computational cost, effective health data analysis.
Sociological Research	Analysis of public sentiment and social trends	Explainable NLP models using KANs analyze text data to provide insights into public sentiment, social trends, and policy impacts.	Improved interpretability, effective analysis of large text datasets, enhanced policy insights.

The table presents a comprehensive overview of the diverse applications of Kolmogorov-Arnold Networks (KANs) in urban research, spanning various domains including traffic and mobility, environmental monitoring, infrastructure management, urban health studies, and sociological research. Each application is accompanied by a brief description of the specific use case and the advantages offered by KANs in addressing the associated challenges.

8. Conclusion

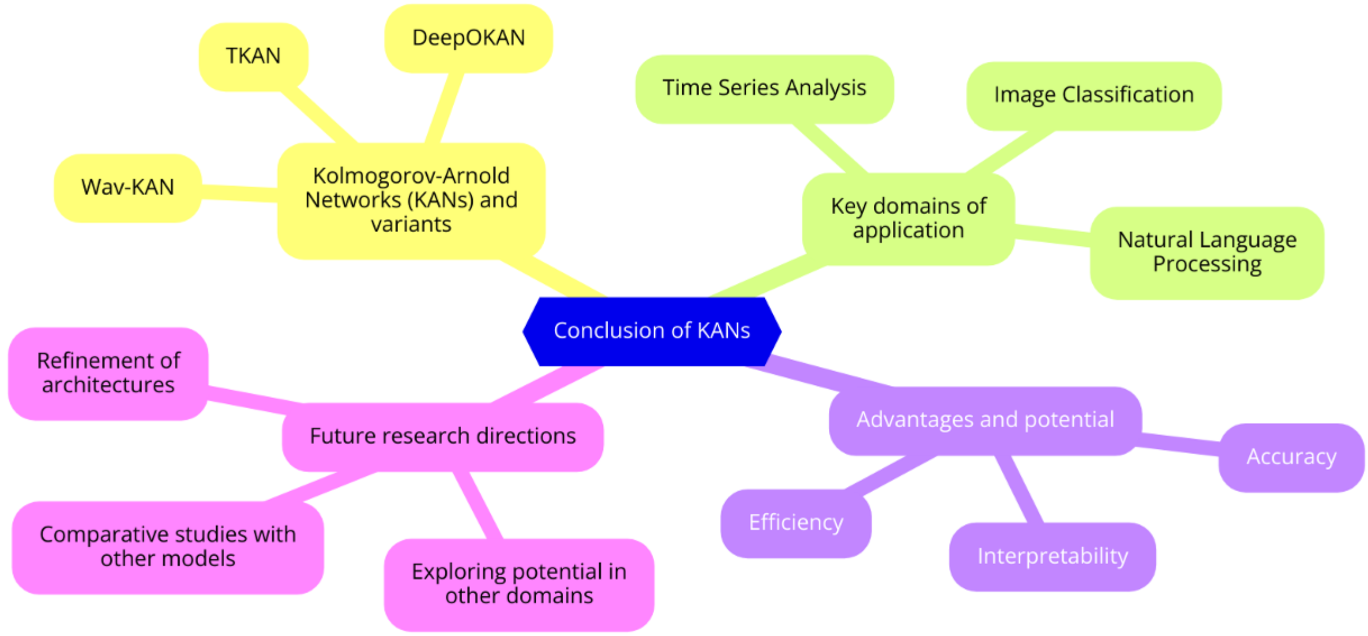

Figure 10. KAN's domains, variants, and future research possible direction:: the author (2024)

Kolmogorov-Arnold Networks (KANs) and their variants, such as Wav-KAN, DeepOKAN, and TKAN, have demonstrated remarkable potential across a spectrum of domains, including time series analysis, image classification, natural language processing, and emerging applications in smart city IoT and urban research. These networks are rooted in the Kolmogorov-Arnold representation theorem, offering a distinctive approach to function approximation. This approach is underpinned by a theoretical guarantee of universal approximation capability.

Despite the promising results demonstrated thus far, further research is imperative to comprehensively comprehend the strengths and limitations of these KAN variants and explore their untapped potential in the burgeoning domains of smart city development and urban research. The integration of KANs into these realms holds significant promise, given their unique properties and ability to analyze complex, nonlinear relationships in data generated by IoT sensors and urban systems.

As research in the field progresses, it is becoming increasingly evident that KANs are poised to assume an increasingly pivotal role in the landscape of machine learning and artificial intelligence. Their transformative potential in reshaping how we perceive and construct neural networks makes them a captivating area of exploration. The journey of unlocking and harnessing the full potential of KANs is in its nascent stages, brimming with anticipation and promise.

Moreover, comparative studies between KANs and other machine learning models in the context of smart city IoT and urban research could yield invaluable insights into their relative performance and suitability for diverse tasks. As the field of KANs continues to evolve and mature, it is foreseeable that these networks will emerge as indispensable tools in driving the advancement of machine learning and artificial intelligence, particularly in the dynamic and complex urban environment.

Footnotes

[1] Federico Girosi, Tomaso Poggio; Representation Properties of Networks: Kolmogorov's Theorem Is Irrelevant*Neural Comput* 1989; 1 (4): 465–469. doi: https://doi.org/10.1162/neco.1989.1.4.465

[2] GitHub - KindXiaoming/pykan: Kolmogorov Arnold Networks

[3] GitHub - mintisan/awesome-kan: A comprehensive collection of KAN(Kolmogorov-Arnold Network)-related resources, including libraries, projects, tutorials, papers, and more, for researchers and developers in the Kolmogorov-Arnold Network field.

[4] This paper introduces a novel application of KolmogorovArnold Networks (KANs) to time series forecasting, leveraging their adaptive activation functions for enhanced predictive modeling. Inspired by the Kolmogorov-Arnold representation theorem, KANs replace traditional linear weights with spline-parametrized univariate functions, allowing them to learn activation patterns dynamically. The proposed approach opens new avenues for adaptive forecasting models, emphasizing the potential of KANs as a powerful tool in predictive analytics.

References

- Bozorgasl, Z., & Chen, H. (2024). Wav-KAN: Wavelet kolmogorov-arnold networks. arXiv preprint arXiv:2405.12832. Retrieved from https://arxiv.org/abs/2405.12832
- Cheon, M. (2024). Kolmogorov-Arnold Network for Satellite Image Classification in Remote Sensing. arXiv preprint arXiv:2406.00600. Retrieved from https://arxiv.org/abs/2406.00600
- Dhiman, V. (2024). KAN: Kolmogorov–Arnold Networks: A review. Retrieved from https://vikasdhiman.info/reviews/KAN_a_review.pdf
- Federico Girosi, Tomaso Poggio; Representation Properties of Networks: Kolmogorov's Theorem Is Irrelevant. Neural Comput 1989; 1 (4): 465–469. doi: https://doi.org/10.1162/neco.1989.1.4.465
- Galitsky, B. A. (2024). Kolmogorov-Arnold Network for Word-Level Explainable Meaning Representation. Preprints. Retrieved from https://www.preprints.org/manuscript/202405.1981
- Genet, R., & Inzirillo, H. (2024). TKAN: Temporal Kolmogorov-Arnold Networks. arXiv preprint arXiv:2405.07344. Retrieved from https://arxiv.org/abs/2405.07344
- Hao, H., Zhang, X., Li, B., & Zhou, A. (2024). A First Look at Kolmogorov-Arnold Networks in Surrogate-assisted Evolutionary Algorithms. arXiv preprint arXiv:2405.16494. Retrieved from https://arxiv.org/abs/2405.16494
- Huang, Z., Cui, J., Yu, L., & Herbozo Contreras, L. F. (2024). Abnormality Detection in Time-Series Bio-Signals using Kolmogorov-Arnold Networks for Resource-Constrained Devices. medRxiv. Retrieved from https://www.medrxiv.org/content/10.1101/2024.06.04.24308428.abstract

- Liu, Z., Wang, Y., Vaidya, S., Ruehle, F., & Halverson, J. (2024). KAN: Kolmogorov-Arnold Networks. arXiv preprint arXiv:2404.19756. Retrieved from https://arxiv.org/abs/2404.19756
- Peng, Y., He, M., Hu, F., Mao, Z., & Ding, J. (2024). Predictive Modeling of Flexible EHD Pumps using Kolmogorov-Arnold Networks. arXiv preprint arXiv:2405.07488. Retrieved from https://arxiv.org/abs/2405.07488
- Qiu, Q., Zhu, T., Gong, H., Chen, L., & Ning, H. (2024). ReLU-KAN: New Kolmogorov-Arnold Networks that Only Need Matrix Addition, Dot Multiplication, and ReLU. arXiv preprint arXiv:2406.02075. Retrieved from https://arxiv.org/abs/2406.02075
- Samadi, M. E., Müller, Y., & Schuppert, A. (2024). Smooth Kolmogorov Arnold networks enabling structural knowledge representation. arXiv preprint arXiv:2405.11318. Retrieved from https://arxiv.org/abs/2405.11318
- Schmidt-Hieber, J. (2021). The Kolmogorov–Arnold representation theorem revisited. Neural networks, 137, 119-126.
- Vaca-Rubio, C. J., Blanco, L., & Pereira, R. (2024). Kolmogorov-Arnold Networks (KANs) for Time Series Analysis. arXiv preprint arXiv:2405.08790. Retrieved from https://arxiv.org/abs/2405.08790
- Wakaura, H., & Suksmono, A. B. (2024). The Quantum Version of Kolmogorov-Arnold Network (KAN): Variational Quantum Kolmogorov-Arnold Network. ResearchSquare. Retrieved from https://www.researchsquare.com/article/rs-4504342/latest
- Xu, K., Chen, L., & Wang, S. (2024). Kolmogorov-Arnold Networks for Time Series: Bridging Predictive Power and Interpretability. arXiv preprint arXiv:2406.02496. Retrieved from https://arxiv.org/abs