Iranian Society of Acoustics and Vibration
The 13th ISAV2023 International Conference on Acoustics and Vibration
20, 21 Dec 2023 Tehran - Iran

Leveraging Koopman operator and Deep Neural Networks for Parameter Estimation and Future Prediction of Duffing oscillators

Yassin Riyazi[a], NavidReza Ghanbari[a], Arash Bahrami[b*]

[a] *M.Sc. Student, School of Mechanical Engineering, College of Engineering, University of Tehran, Tehran, Iran.*

[b] *Assistant Professor, School of Mechanical Engineering, College of Engineering, University of Tehran, Tehran, Iran.*

** Corresponding author e-mail: arash.bahrami@ut.ac.ir*

Abstract

The exploration of nonlinear dynamical systems holds significant importance across scientific and engineering disciplines, primarily for its applications in modeling real-world phenomena. Traditional methods employed for the analysis and prediction of the behavior of these systems typically involve intricate mathematical techniques and numerical simulations. This paper presents an innovative approach that combine the capabilities of the Koopman operator and deep neural networks to establish a linear representation of the Duffing oscillator. This newly developed methodology facilitates effective parameter estimation and the accurate prediction of the oscillator's future behavior. Moreover, the paper proposes a modified training procedure aimed at confining the Koopman operator to a linear layer within the neural network, as opposed to its application across the entire network. This synergy between the Koopman operator and deep neural networks not only simplifies the analysis of nonlinear systems but also paves the way for significant advancements in predictive modeling across diverse fields.

Keywords: Koopman operator; Parameter estimation; Duffing oscillator; Nonlinear dynamical systems

---

1. Introduction

Nonlinear dynamical systems, recognized for their intricate and sometimes chaotic behavior, permeate the realms of natural phenomena and technological applications. They transcend the simplicity of linear systems, giving rise to phenomena such as bifurcations, limit cycles, and chaotic attractors. These systems have long captivated the interest of scientists and engineers, presenting substantial challenges in understanding, characterizing, and predicting their trajectories. Across diverse fields, from physics and biology to economics and engineering, nonlinear systems underscore the fundamental complexity of our world.

At the heart of this intricate landscape lies the Duffing oscillator [1], an iconic archetype of nonlinear dynamical systems. Its versatility enables it to emulate a wide spectrum of behaviors, making it a pertinent model for various physical phenomena. From capturing the subtle interplay of mechanical vibrations in structures subjected to external forces to mirroring the rhythmic patterns of biological oscillations, the Duffing oscillator encapsulates the essence of nonlinear dynamics.

Traditionally, dissecting and forecasting the behavior of Duffing oscillators have relied on a combination of analytical techniques and numerical simulations. While these methods provide valuable insights, they often encounter limitations in handling nonlinear intricacies with precision. Analytical solutions may prove elusive and algebraically complex, especially for higher-dimensional or strongly nonlinear systems. On the contrary, numerical simulations, though powerful, demand extensive computational resources and face challenges in long-term predictions due to inherent numerical errors and uncertainties.

The Koopman operator [2], with its inherent structure involving a mapping to a higher dimension, a linear transformation, and an inverse mapping, bears resemblance to the structure of autoencoders [3]. Bethany Lusch, J. Nathan Kutz, and Steven L. Brunton in [4] have developed a neural network deliberately tailored for conciseness and interpretability, with the specific aim of encapsulating system dynamics within a low-dimensional manifold. They introduce nonlinear coordinates, which are identified through a modified auto-encoder, and under these coordinates, the dynamics of the system exhibit a global linearity property. Furthermore, they expand the concept of Koopman representations to encompass systems with continuous spectra, introducing an auxiliary network to efficiently parameterize continuous frequencies. This innovative approach establishes a bridge between deep learning models and decades of asymptotic research, offering a fusion of the advantages of deep learning with the capacity for physical interpretability provided by Koopman embedding.

In a recent study, Kathleen Champion, Bethany Lusch, J. Nathan Kutz, and Steven L. Brunton [5] addressed the challenge of discovering governing equations from scientific data in data-rich fields lacking well-characterized quantitative descriptions. They utilized advances in sparse regression to identify both the structure and parameters of nonlinear dynamical systems from data, resulting in models that balance simplicity with descriptive power. To achieve this, they introduced a custom deep autoencoder network designed to uncover a coordinate transformation into a reduced space, simplifying the representation of dynamics. This innovative approach allows for simultaneous learning of governing equations and the associated coordinate system. They applied this technique to various examples of high-dimensional systems exhibiting low-dimensional behavior, creating a modeling framework that combines the flexibility of deep neural networks with the parsimonious nature of sparse identification of nonlinear dynamics (SINDY) [6].

While the sole use of a neural network proves accurate and meets our requirements, it does not guarantee the exclusive confinement of the Koopman operator to a designated linear layer. Consequently, the entire network structure incorporates elements of mapping, linear transformation, and inverse mapping simultaneously. This challenges the utility of using the linear layer weight as a representation of the system, as they only encapsulate a portion of the Koopman operator. To overcome these challenges, we introduce a novel approach that capitalizes on the synergy between the Koopman operator and Deep Neural Networks [7]–[9]. This groundbreaking fusion is aimed at converting the Duffing oscillator into a linearized representation, offering promising solutions to the intricacies encountered in traditional methods. By harnessing the computational power of deep learning and the Koopman operator's capability to provide a linear representation of nonlinear systems [10], our approach enables a more accurate Koopman linearized representation of system behavior.

In the following sections, we first dig into the foundational principles of Koopman operator theory and the adaptability of deep neural networks. We show how their fusion forms a compelling framework for analyzing and predicting the behavior of Duffing oscillators. We outline the process of transforming Duffing oscillator dynamics into a linear representation and introduce a modified loss function designed to enhance the generality of the Koopman linear representation of the dynamical system within this context. Through numerical simulations and comparisons with traditional methods, we demonstrate the efficacy of our approach in providing accurate predictions for the future behavior of Duffing oscillators. Finally, this work enriches our understanding of nonlinear dynamics and offers a powerful tool with transformative potential across scientific, engineering, and practical applications.

For access to the project's code and resources, please refer to the GitHub repository available at: github.com/yriyazi/Koopman-Operator-and-Deep-Neural-Networks-ISAV2023/

2. Koopman operator and Its Application

In recent years, the Koopman operator has emerged as a powerful mathematical tool that provides a fresh vantage point for studying dynamical systems. Rooted in functional analysis, the Koopman operator introduces a paradigm shift by transitioning the focus from the state space to the space of observable functions. Doing so allows us to view the system's evolution in a linear framework, even when dealing with inherently nonlinear systems. This perspective offers a new lens through which we can gain insights into the dynamics of complex systems.

2.1 Dynamical System Representation

We consider a dynamical system described by a set of state variables $x(t)$, which evolve over time t. Mathematically, we can represent this as:

$$\frac{dx}{dt} = f(x,t) \tag{1}$$

where

- n is the state vector dimension.
- $x \in \mathbb{R}^n$ is the state vector representing the system's state variables.
- $t \in \mathbb{R}^+$ is time.
- $f(x): \mathbb{R}^n \to \mathbb{R}^n$ is a vector-valued function describing how the state variables change over time.

2.2 Koopman operator Transformation

The Koopman operator, denoted as $\mathcal{K}$, is an infinite-dimensional linear operator that acts on observables or functions of the state variables. Let $g(x)$ be such an observable. The Koopman operator maps this observable from the state space to a higher-dimensional space:

$$\mathcal{K}g(x) = g\big(f(x)\big) \tag{2}$$

where:

- m supposed to be infinite-dimensional but in numerical approximation a value will be assigned.
- $\mathcal{K}(.): \mathbb{C}^m \to \mathbb{C}^m$ is the Koopman operator Generator.
- $g(x): \mathbb{R}^n \to \mathbb{C}^m$ is an observable or function defined on the state space.
- $g\big(f(x)\big)$ represents the observable after the system evolves according to $f(x)$.

2.3 Koopman operator in Discrete Time

In discrete-time dynamical systems, the Koopman operator is applied at discrete time steps. For these systems Equation (2) may be represented as:

$$\mathcal{K}g(x_k) = g(x_{k+1}) \tag{3}$$

where:

- x_k represents the state of the system at time k.
- x_{k+1} represents the state of the system at the next time step $k+1$.

In Figure 1, an important relationship is apparent: the transformation of data x_t from its low-dimensional representation in Euclidean space to an infinite-dimensional Hilbert space is facilitated through the utilization of Koopman observables $g(x_t)$. Leveraging this Koopman transformation allows for the mapping of y_t to y_{t+1} via a linear matrix transformation. Furthermore, employing the inverse Koopman observable mapping $g^{-1}(y_{t+1})$ enables the derivation of x_{t+1}. With the help of the Koopman operator $f(x_t) = g^{-1}(\mathcal{K} \times g(x_t)) = x_{t+1}$.

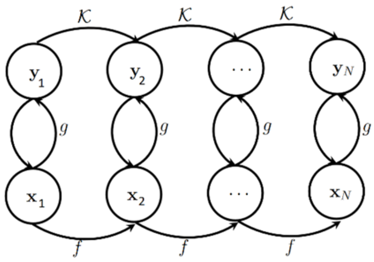

Figure 1. Koopman operator Evolution and a discrete dynamical system.

3. Coupling Koopman operators with Deep Neural Networks

Deep neural networks have showcased remarkable abilities in approximating intricate functions and mastering complex patterns from data. One of the key challenges encountered in the realm of Koopman operators is the identification of suitable observable functions. In methods such as DMD [11], the observable function is typically the identity function, and in extended DMD (EDMD), observable functions take the form of polynomials or trigonometric functions. While these approaches are straightforward and accurate, they exhibit resilience to noise and initial conditions.

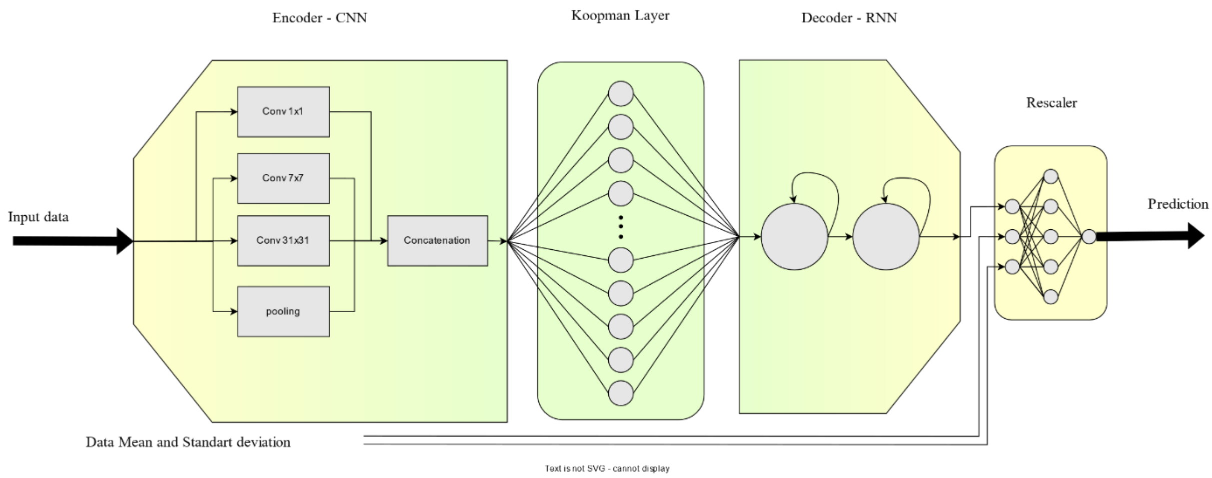

Figure 2. Neural Network Diagram.

In this study, we combine a deep neural network with the Koopman operator, thereby generating a linearized representation of the Duffing oscillator. This neural network effectively learns the intricate relationship between system parameters and observed behaviors, facilitating efficient parameter estimation. Moreover, the neural network undergoes training to predict the future trajectory of the Duffing oscillator, thereby equipping us with a valuable tool for forecasting system behavior.

neural networks are not necessarily always better than feature crosses, but neural networks do offer a flexible alternative that works well in many cases.

3.1 Data acquisition

The Duffing oscillator is a dynamical system described by the following second-order differential equation:

$$\frac{d^2x}{dt^2} + \delta\frac{dx}{dt} + \alpha x + \beta x^3 = \gamma\, cos(\omega t) \quad (5)$$

where:

- x represents the displacement of the oscillator from its equilibrium position.
- δ denotes the damping coefficient.
- α is the linear stiffness coefficient.
- β characterizes the nonlinearity in the system.
- γ is the amplitude of the external driving force.
- ω is the angular frequency of the driving force.

Duffing oscillator solution is generated using the Runge-Kuta method [12] and the initial condition for Solving the equation is $x_0 = 1.5$ and $v_0 = -1.5$.

A normal distribution with noise in the range of $[-0.5,0.5]$ is added to the data to simulate real-world data.

3.2 Data Normalization

Given the nature of regression, it is advisable to normalize the data before training the neural network. This normalization is crucial because even small variations in the input data can lead to significant changes in the output, potentially reducing the model's robustness against changes in input conditions. In our approach, we employ data normalization before feeding it into the convolutional neural network (CNN) architecture. During this process, we pass the statistical properties of the data, such as its mean and variance, through the network.

After the recurrent neural network (RNN) block, the data is remapped to its original statistical properties before being passed to the Rescaler Block. The purpose of the Rescaler Block is to further reduce the variation in the output, ultimately leading to a more stable and controlled model response.

3.3 Structure

The architecture of the model is shown in Figure 2. Neural Network is structured around an Encoder-Decoder paradigm, which effectively captures the essence of complex dynamics. Specifically, the Encoder component is meticulously designed, featuring a sequence of Inception Blocks [13] in a convolutional neural network (CNN) [14]. These Inception Blocks serve as robust feature extractors, enabling the model to discern intricate patterns and relevant features from the input data.

Following the Inception-based Encoder, a pivotal transformation takes place through a linear layer. This linear layer assumes a distinct role within the architecture, embodying the essence of the Koopman operator evolution function $\mathcal{K}$. It is important to note that this linear layer operates without an activation function and bias, preserving the linear nature of the Koopman operator's transformation.

Table 1. CNN parameters. Out hyperparameter is 20.

	Sequential Blocks	In Channels	Out Channels	kernel size	padding
branch1x1	Conv 1×1	1	out	1	0
	ReLU	-	-	-	-
branch7x7	Conv 1×1	1	out	1	0
	ReLU	-	-	-	-
	Conv 7×7	out	out	7	3
	ReLU	-	-	-	-
branch31x31	Conv 1×1	1	out	1	0
	ReLU	-	-	-	-
	Conv 31×31	out	out		15
	ReLU	-	-	-	-
branch pool	MaxPool1d	1	1	3	1
	-	-	-	-	-
	Conv 1×1	1	out	1	0
	ReLU	-	-	-	-

Transitioning from the Koopman operator layer, the architecture takes an intriguing turn with the integration of a two-layer Long Short-Term Memory (LSTM) [15] network. This LSTM component acts as the Decoder, expertly leveraging its sequential memory to unravel the transformed linearized representation. This sequence-to-sequence modeling approach facilitates the reconstruction of the system's temporal evolution, a crucial aspect in capturing its intricate behaviors.

3.4 Training

The network is trained end-to-end, without the need for custom loss functions or specialized training algorithms. However, it is important to note that the evolution function of the Koopman operator does not remain confined solely to the Koopman part; instead, it spreads throughout the network. In a sense, the network operates as a black box, handling this evolution internally.

To address this issue and restrict the Koopman operator's influence exclusively to the Koopman linear layer, a two-stage training algorithm has been proposed. In this algorithm, after each optimization step:

1. The weights of all layers except the Koopman Linear Layer are frozen.
2. The output of the Koopman Linear Layer is calculated for time steps n_0 to n_{KPH} (KPH is the Hyperparameters and due to the cost of calculating matrix power 20 was selected).
3. The weights of the Koopman Linear Layer are updated based on the linearity property. This update aims to minimize the prediction error of the nth output.

$$\sum_{n}^{KPH} \mathscr{L}\left(g\left(x_{n_0}\right) \times \left(W_{Koopman}^{n}\right)^{T}, g\left(x_n\right) \times \left(W_{Koopman}\right)^{T}\right) \quad (4)$$

By implementing this two-stage training process, we ensure that the Koopman operator's influence is confined and utilized specifically within the Koopman linear layer, enhancing the network's predictive accuracy and control.

Table 2. Optimizer and loss parameters.

		Stage 1	Stage 2
Optimizer	Type	SGD	SGD
	Learning Rate	5.00E-02	5.00E-04
	momentum	0.9	-
	weight decay	1.00E-04	-
loss	Type	MSE	MSE

4. Results and Discussion

Numerical results demonstrate the effectiveness of the proposed approach. The combination of Koopman operator-based linearization and deep neural networks yields impressive results in terms of parameter estimation accuracy and future prediction. During the Networks training input horizon for extracting features was 200 previous samples.

Throughout this paper, the numerical values are used: $\alpha = -1$, $\beta = +1$, $\delta = 0.3$, and $\omega = 1.2$. Additionally, the initial conditions for the training dataset are set to $x(0) = 1.5\ [m]$ and $\dot{x}(0) = -1.5\ [m/s]$. Also, we study periodic and quasi-periodic responses of the Duffing equation by changing value of γ.

4.1 Periodic $\gamma = 0.2\ [N]$

For periodic oscillations, neural networks demonstrate an ability to effectively capture the underlying oscillatory structure, yielding accurate predictions even in the presence of substantial noise.

Figure 3 depicts various scenarios: a) illustrates the neural network's performance under normal training conditions, b) presents a similar scenario with an increased noise level, and c) represents a worst-case situation where noise completely overwhelms the available data, resulting in the network's inability to provide accurate predictions.

In general, neural networks exhibit robustness against noise within a range spanning $(-1, +1)$.

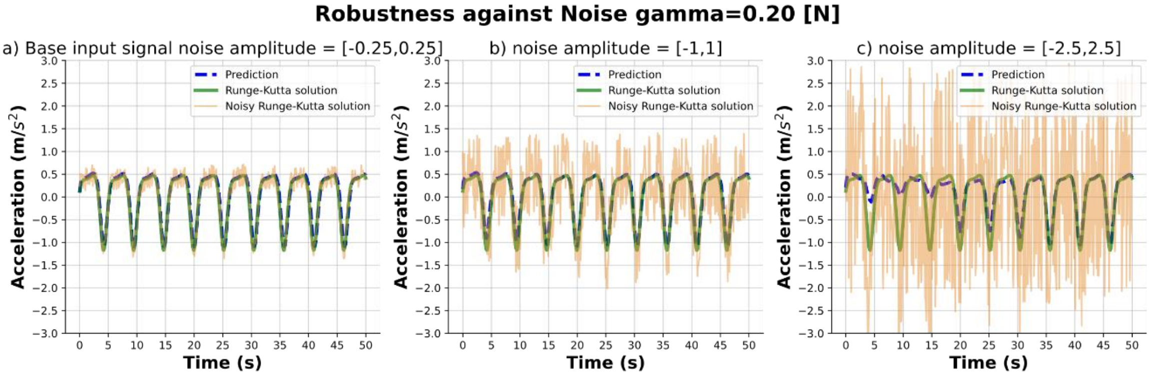

Figure 3. Robustness against noise $\gamma = 0.2$ $[N]$.

4.2 Quasi-periodic $\gamma = 0.37$ $[N]$

For quasi-periodic oscillations, akin to simple periodic ones, neural networks exhibit the ability to effectively capture the underlying oscillatory structure, leading to accurate predictions even in the presence of significant noise.

Figure 4 delineates various scenarios: a) shows the neural network's performance under typical training conditions, b) illustrates a similar scenario but with an elevated noise level, and c) portrays a worst-case scenario where noise completely dominates the available data, resulting in the network's inability to provide precise predictions.

Broadly speaking, neural networks demonstrate robustness against noise within a range of $(-1, +1)$.

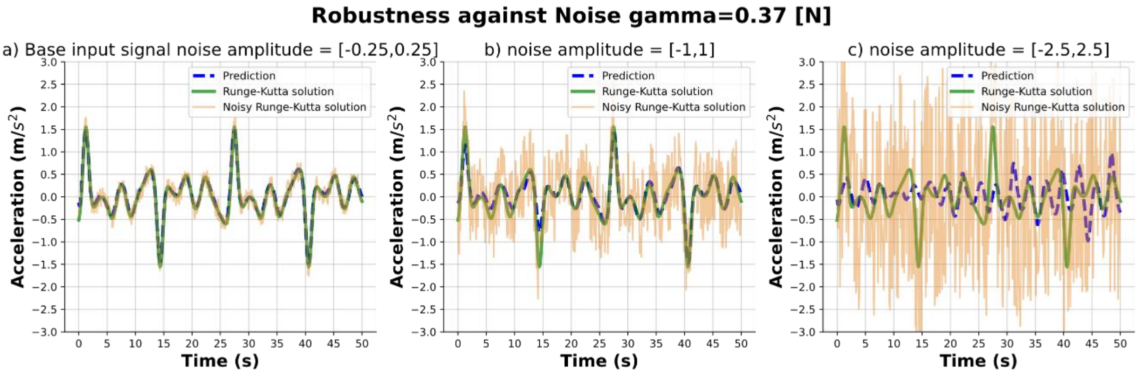

Figure 4. Robustness against noise $\gamma = 0.37$ $[N]$.

4.3 Koopman Eigenvalues

One of the drawbacks associated with Dynamic Mode Decomposition (DMD) [11] and Extended Dynamic Mode Decomposition (EDMD) methods [16] is the presence small number of Koopman operator eigenvalues. To address this issue, a radial basis function has been proposed in a previous study [17] to approximate the Koopman operator eigenvalues.

In Figure 5, we present a plot of 1600 Koopman operator eigenvalues, demonstrating their diversity and their ability to capture various system behaviors. It is noteworthy that, in the study mentioned [16], these results were obtained for a free Duffing oscillator. Nevertheless, the results achieved through Deep Neural Networks surpass those obtained by other means.

5. Conclusion

In this research, our objective was to implement an Autoregressive self-supervising autoencoder utilizing a two-stage training process. The primary goal was to confine the Koopman operator to a specific layer within the network and assess its resilience to noise. In addition, we aimed to predict the steady-state behavior of the system while ensuring that the network remains robust across different initial conditions, indeed with the same Duffing coefficient.

The utilization of an autoregressive framework played a crucial role in enabling the network to generate predictions, akin to the capabilities demonstrated by large language models.

Furthermore, in comparison to traditional methods such as DMD and EDMD, our approach yielded an increased diversity in the Koopman eigenvalues, even surpassing radial basis function (RBF) methods in this regard.

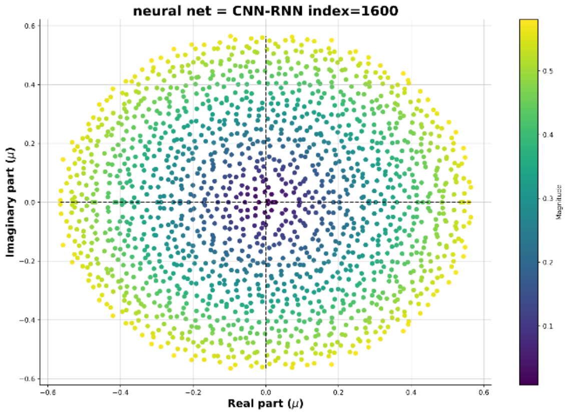

Figure 5. Koopman Layer Eigen values.

REFERENCES

[1] Hamel, "Georg Duffing, Ingenieur: Erzwungene Schwingungen bei veränderlicher Eigenfrequenz und ihre technische Bedeutung. Sammlung Vieweg. Heft 41/42, Braunschweig 1918. VI+134 S," *ZAMM - J. Appl. Math. Mech. / Zeitschrift für Angew. Math. und Mech.*, vol. 1, no. 1, 1921.

[2] B. O. Koopman, "Hamiltonian Systems and Transformation in Hilbert Space," *Proc. Natl. Acad. Sci.*, vol. 17, no. 5, 1931.

[3] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science (80-.).*, vol. 313, no. 5786, 2006.

[4] B. Lusch, J. N. Kutz, and S. L. Brunton, "Deep learning for universal linear embeddings of nonlinear dynamics," *Nat. Commun.*, vol. 9, no. 1, 2018.

[5] K. Champion, B. Lusch, J. Nathan Kutz, and S. L. Brunton, "Data-driven discovery of coordinates and governing equations," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 116, no. 45, 2019.

[6] E. Kaiser, J. N. Kutz, and S. L. Brunton, "Sparse identification of nonlinear dynamics for model predictive control in the low-data limit," *Proc. R. Soc. A Math. Phys. Eng. Sci.*, vol. 474, no. 2219, 2018.

[7] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "2012 AlexNet," *Adv. Neural Inf. Process. Syst.*, 2012.

[8] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, 2015.

[9] M. Shafiq and Z. Gu, "Deep Residual Learning for Image Recognition: A Survey," *Applied Sciences (Switzerland)*, vol. 12, no. 18. 2022.

[10] S. L. Brunton, M. Budišić, E. Kaiser, and J. N. Kutz, "Modern Koopman Theory for Dynamical Systems," *SIAM Rev.*, vol. 64, no. 2, 2022.

[11] P. J. Schmid, "Dynamic mode decomposition of numerical and experimental data," *J. Fluid Mech.*, vol. 656, 2010.

[12] C. Runge, "Ueber die numerische Auflösung von Differentialgleichungen," *Math. Ann.*, vol. 46, no. 2, 1895.

[13] C. Szegedy *et al.*, "Going deeper with convolutions," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2015.

[14] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, 1998.

[15] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Comput.*, vol. 9, no. 8, 1997.

[16] M. O. Williams, I. G. Kevrekidis, and C. W. Rowley, "A Data–Driven Approximation of the Koopman operator: Extending Dynamic Mode Decomposition," *J. Nonlinear Sci.*, vol. 25, no. 6, 2015.

[17] Q. Li, F. Dietrich, E. M. Bollt, and I. G. Kevrekidis, "Extended dynamic mode decomposition with dictionary learning: A data-driven adaptive spectral decomposition of the Koopman operator," *Chaos*, vol. 27, no. 10, 2017.