Proceedings of CONF-MLA 2025 Symposium: Intelligent Systems and Automation: AI Models, IoT, and Robotic Algorithms
DOI: 10.54254/2755-2721/2025.LD28244

Application of Ordinary Differential Equations and Partial Differential Equations in Neural Networks

Jialin Zhao

Nanjing Jinling Hexi School, Nanjing, China
verity00716z@outlook.com

Abstract. As an important research direction in the field of artificial intelligence, neural networks have made significant progress in many fields such as image recognition and natural language processing in recent years. The mathematical principles behind it, particularly ordinary and partial differential equations, play a pivotal role in describing the dynamic behavior and learning processes of neural networks. Therefore, the study of the mathematical principles of these two and their application in neural networks has important theoretical significance and practical value. This paper first introduces the mathematical principles of ordinary differential equations and their applications in neural networks, then expounds the mathematical principles of partial differential equations and their applications in neural networks, and finally analyzes the limitations of existing research and looks forward to future development. By systematically reviewing the application of ordinary differentiation and partial differentiation in neural networks, this paper provides a new perspective and idea for related research. At the same time, the research in this paper is helpful to deepen the understanding of the dynamic behavior of neural networks, promote the development of neural network theory, and provide a theoretical basis for model design and optimization in practical applications.

Keywords: Ordinary differential equation, Partial differential equation, Neural Networks.

1. Introduction

In today's era of technological development, artificial intelligence has entered various industries. It has changed people's way of life and work. Neural network is one of the core technologies of artificial intelligence. It can handle a large amount of data. It is good at recognizing patterns. It has achieved significant results in image recognition, natural language processing, and speech recognition. The application scenarios are becoming increasingly complex. Traditional neural networks are not powerful enough. They perform poorly over time. They cannot accurately simulate complex systems.

Researchers have proposed a new idea. They introduced differential equations into neural networks. God often combines network structure and ordinary differential equations in differential equations. It can model processes that change over time. This method enhances the ability of neural networks. It can handle continuous time data. It can simulate complex systems more effectively. It provides a new path to solve existing problems.

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2025.LD28244

Neural partial differential equations are another approach. It embeds partial differential equations into neural networks. It enhances the modeling ability of the network in space and time. Researchers use this method to improve the accuracy of complex system models. For example, how fluids flow. For example, how climate changes. This indicates that neural networks have made progress in scientific computing and engineering simulation.

The application of differential equations in neural networks has achieved certain results. Practical applications still face challenges. The calculation speed of neural differential equations is slow. High dimensional systems are difficult to effectively handle. Nonlinear systems are difficult to accurately model. Model performance decreases when there is insufficient data. These issues are of great importance in research. There are also difficulties in solving these problems.

Studying differential equations in neural networks is meaningful. People need to understand the mathematical mechanisms behind it. It is necessary to verify its usability in real-world scenarios. People need to analyze the shortcomings in current research. People need to consider the future development direction. This work has theoretical value. It also has potential for application in practice.

This article starts from the basic concepts of ordinary differential equations. Introduced its specific application in neural networks. It explains the use of ordinary differential equations in time modeling. It elucidates the fundamental theory of partial differential equations. It analyzed its advantages in spatiotemporal modeling. It summarizes the current research on integrating ordinary differential equations and partial differential equations into neural networks. It points out the difficulties and shortcomings that exist within it. It proposes possible future research directions and application scenarios.

By exploring these contents, this article hopes that readers can understand the application of ordinary and partial differentials in neural networks. It hopes readers can recognize the limitations of current research. It hopes to provide useful information for subsequent research and practical applications.

2. Ordinary differential equations and their applications in neural networks

2.1. Mathematical principles of ordinary differential equations

Ordinary differential equations rely on one independent variable. Their form is usually

$$(\frac{dy}{dx} = f(x, y)) \tag{1}$$

.(y) is the dependent variable. (x) is the independent variable. $(f(x, y))$ is a known function. It is defined on an open set [1]. Ordinary differential equations assist in the study of dynamic systems. They are used in physics, engineering, economics, and computer science. Solving them can display the system's changes over time. The pattern of change can be observed. The trend becomes easy to understand [2].

2.2. Application of ordinary differential equations in neural networks

In the study of neural networks, ordinary differential equations are models of dynamic behavior. These equations help establish complex time series models. They can be used for sequence prediction and time series analysis. ODE allows neural networks to process data related to time. This

DOI: 10.54254/2755-2721/2025.LD28244

improves generalization and prediction accuracy. This integration helps the network handle the constantly changing flow of information. It extends the use beyond static input.

ODE in neural networks is an important research direction in the field of deep learning. The purpose of doing so is to improve the expressiveness and learning efficiency of the model. ODE provides a solid mathematical framework. They capture complex nonlinear relationships. They describe how the system state changes over time. For example, long short-term memory networks work like solving ode. The change in internal state simulates temporal correlation. This makes effective sequence data modeling possible. This method captures long-term dependencies. It also improves the performance of time series prediction. It is achieved by establishing parametric differential equations [3].

On this basis, subsequent work focused on the stability and robustness of ode based neural networks. Ilya et al. studied the stability of solutions to nonlinear differential equations with right-handed discontinuities. They applied the results to the Hopfield artificial neural network. Their analysis shows the crucial role of stable conditions. When using ode, these conditions ensure reliable and predictable network behavior [4]. In addition, adding ode to neural architecture helps to incorporate physical laws. This improves interpretability and generalization. This method has been extended to fractional order differential equations. These equations provide powerful tools for capturing complex dynamics in neural networks. Imran et al. proved this [5]. These studies indicate that ODE provides a solid mathematical framework. This framework is helpful for designing neural networks. These networks simulate continuous dynamics. They have improved their understanding of the theory. Compared to traditional discrete layer models, they provide practical benefits. These benefits include better adaptability and efficiency.

In medical image analysis, Hu and Jin proposed a diffusion based anomaly detection method for medical images. This method utilizes diffusion modes to estimate the likelihood of image features. It uses neural networks to fit these estimates. It calculates anomaly scores by diffusion ODE and weighted average estimated by neural networks. They also introduced an anomaly localization method based on image reconstruction. This method identifies anomalous regions by calculating reconstruction errors. The experiments on the BraTS2021 brain tumor dataset and X-ray chest X-ray dataset show that this method outperforms existing methods in terms of anomaly detection performance and robustness [6].

Rong studied a cell image segmentation method based on ODEs and watershed algorithms. This work aims to improve the accuracy and efficiency of cell image segmentation. It applies ode to image processing. It combines watershed algorithm for segmentation. Effective cell image segmentation has been achieved [7]. These applications show ODEs can model complex dynamics in neural networks across different fields.

3. Partial differential equations and their applications in neural networks

3.1. Mathematical principles of partial differential equations

Partial differential equations are an important mathematical tool used in many fields of science and engineering. Their basic forms include multivariate functions and their partial derivatives. Partial differential equations can be expressed as

$$(F(x_1, x_2, \ldots, x_n, u, u_{x_1}, u_{x_2}, \ldots, u_{x_n}) = 0) \tag{2}$$

DOI: 10.54254/2755-2721/2025.LD28244

, where $(u = u(x_1, x_2, \ldots, x_n))$ is the pending function, and (u_{x_i}) represents the partial derivative of the independent variable (x_i)[8]. Its basic types include elliptical, parabolic, and hyperbolic. Each type has different mathematical characteristics and applicable scenarios.

Elliptical partial differential equations describe steady-state distribution behavior. They are applicable to the heat conduction equation. The parabolic equation simulates dynamic processes. They are applicable to changes in velocity fields in fluid dynamics. Hyperbolic equation analysis of wave phenomena. They are suitable for sound propagation.

In the mathematical theory of partial differential equations, constants and variables are defined to establish models. These elements form the theoretical basis for understanding equations. They affect the behavior of the system to a certain extent. In mathematical modeling, constants represent parameters that remain fixed over time or space. A variable is a quantity that changes under given conditions. Understanding their connections is important for exploring partial differential equations in neural networks [9].

From a mathematical perspective, adding constants helps the model maintain stability in complex systems. In the Navier Stokes equations of fluid mechanics, fluid viscosity is set as a constant. This simplifies the calculation of velocity and pressure fields. In many practical problems, fluid characteristics can change due to factors such as temperature and pressure. Setting these features as variables makes the model more realistic. In the heat conduction equation, temperature is regarded as a function of time and space. Its changes reflect the heat transfer process in the system [10].

3.2. Mathematical principles of partial differential equations

In modern computer vision, CNN has been widely applied due to its outstanding performance in image tasks. Using partial differential equations in CNN helps support how to discover and display image features. These equations help the network find edges and textures in the image. They can also help the network better process new data and more effectively handle noise.

Convolution operation is a type of linear partial differential equation. Their working principle is to sum weighted input images through convolution kernels. This method helps the network find key spatial structures in the data. Improved image processing performance. In edge detection, using the Laplacian operator to construct convolution kernels can enhance the edges of images. This makes future feature learning more accurate. The use of PDE in the form of time or frequency can help CNN better handle changing or blurry images.

In complex applications, image denoising using partial differential equations is integrated into the training of convolutional neural networks. This method adopts bilateral filtering or anisotropic diffusion theory. It increases local smoothness and global consistency to optimize training data. This greatly improves the quality of denoising and reconstruction. Compared with traditional algorithms, convolutional neural networks based on PDE have better local adaptability and global resolution in processing complex images.

In the context of deep learning, the effectiveness and accuracy of spatial feature extraction are considered to be one of the key factors in improving model performance. Partial differential equations, as a mathematical tool capable of describing spatial phenomena and complex dynamic systems, have been gradually applied to the field of deep learning, especially in the process of feature extraction. Partial differential equations provide neural networks with rich spatial properties by modeling continuous changes in images and other high-dimensional data, thereby enhancing their recognition and classification capabilities.

DOI: 10.54254/2755-2721/2025.LD28244

Boundary conditions and initial conditions are key to solving partial differential equations. They help capture local and global features in complex systems. In image processing, using partial differential equations for edge detection can identify important contours. This is achieved by analyzing the gradient information of the image. The Canny edge detection algorithm uses the principle of partial differential equations. It improves the way image edges are extracted. Experiments have shown that incorporating partial differential equation models into convolutional neural networks can improve the accuracy of region segmentation. This confirms that PDE is helpful for feature extraction [11].

In neural network training, ordinary differential equations and partial differential equations can guide the learning process. This method is called "PDE controlled neural network training". It uses PDE features to shape network design or improve training. The main task is to link PDE solutions with network parameters. This allows the network to learn how to approximate the function required to solve the target problem. People can define a network outputu (x,t) and impose constraints by introducing a partial differential equation, such as:

$$[\frac{\partial u}{\partial t} + \nabla \cdot (D\nabla u) = f(x,t)] \tag{3}$$

In the above formula, $u(x,t)$ the output of the network is the diffusion coefficient, the gradient operator, and the source term is represented. This control strategy helps the neural network maintain features that match the PDE solution during training. Improved the predictive and generalization abilities of the model. By using appropriate loss functions, equation based learning can improve network performance in complex dynamic systems $D\nabla f(x,t)$[12].

In neural network research, partial differential equations are used for image processing and denoising. Cheng Shuhong proposed a Poisson noise removal method for low light images based on partial differential equations. This method utilizes the PDE model to remove Poisson noise and improve image quality [13]. Wang Haobai and Zhang Zhongbo studied the denoising and deblurring model based on PDE and its experiments. They proposed a new PDE model that simultaneously processes image denoising and deblurring. The experimental results confirmed the effectiveness of the model. These studies provide a theoretical foundation and methodological support for the application of neural networks in image processing [14].

4. Existing limitations and future prospects

4.1. Existing limitations

In neural network applications, model complexity has become a key bottleneck in current research. It affects training efficiency and generalization. Model complexity occurs in many parts of the network, such as the number of layers, nodes, and parameter settings. As the depth of the network increases, the training time is also rapidly increasing. This will result in higher usage of computing resources. For example, training deep convolutional neural networks with millions of parameters for image classification typically takes several days or weeks [15].

Model complexity leads to overfitting. This situation occurs when neural networks are too complex in terms of the size and type of training data. The model fits well with the training data. Its ability to process new data has sharply declined. This has been seen in practical applications. It often occurs in small sample tasks. Medical image analysis is an example. Overfitting limits the application of neural networks in such tasks [16]. Researchers use regularization. They use early stop. These help reduce this problem.

DOI: 10.54254/2755-2721/2025.LD28244

In today's field of deep learning, computing resources limit the scale and performance of block neural networks. Training large neural networks requires a significant amount of computational resources. Requires high-performance graphics processing units. Tensor processing units are necessary. The memory requirement is very high. Resource bottlenecks lead to training delays. These delays affect the convergence speed of the model. They will affect the accuracy of the model.

4.2. Future prospects

Neural networks are developing rapidly. Improving computational efficiency is a key issue. The complex network structure poses a challenge to computing resources. Massive datasets pose a challenge to computing resources. Deep learning faces these issues. Traditional algorithms lack computational power. The demand continues to grow. Optimizing efficiency will be helpful. It improves the real-time performance of the model.

Advanced numerical algorithms are helpful. Adaptive time stepping technology is an example. It adjusts the solution of ordinary differential equations. It adjusts the solution of partial differential equations. It is dynamic. It reduces unnecessary calculations. This method checks the stability of the model. It checks the accuracy of the model. It is real-time. It reduces the number of calculation steps. It improves the calculation speed. It makes the results reliable. Research has shown that this method reduces computation time. Under specific conditions, it can be reduced by 30% to 50% [17].

Interdisciplinary integration frameworks have strong potential. This is evident when applying ode and pde in neural networks. Cross disciplinary collaboration is active. Researchers combine mathematical modeling with other fields. A new idea has emerged. ODE and PDE are standard tools. They model dynamic systems. They transform the complexity of the real world into mathematics. Current research emphasizes the cross use of ODE and PDE. Dynamic modeling has become better and control methods have advanced.

Biomedical applications ode. They track the drug levels in their bodies. PDE depicts the growth of cells in space. These models deepen insights into biology. They provide data for new drugs. Neural networks work together with ode. Improve accuracy in parameter adjustment. Improved prediction of dynamic systems. The use of interdisciplinary approaches is growing.

The research on neural networks is developing rapidly. New algorithms have driven progress. The old restrictions need a solution. Most models rely on classical deep learning. They have achieved success in many cases. They are unable to handle complex data. SGD is widely used. Its variants are very common. They are trapped in a local optimum state. Model generalization decreases. People need a new algorithm. Breaking restrictions is crucial.

5. Conclusion

This article introduces the basic ideas of ordinary differential equations and partial differential equations. It focuses on how to use them in neural networks and their limitations. It also discussed where they may be used in the future. This article explores how these equations work in neural networks. It demonstrates how mathematical models can help neural networks perform better and run more efficiently.

Ordinary differential equations describe systems of change. They help neural networks predict sequences more accurately. They established a better time series model. These equations can show how neurons change over time. They simulate activation functions. This helps researchers understand how information moves in neural networks. Partial differential equations deal with problems involving many variables. They can handle data related to space very well. They are very

DOI: 10.54254/2755-2721/2025.LD28244

useful in image processing and time-based prediction. These equations are applicable in many situations. However, combining partial differential equations with modern neural networks still has limitations. These issues require further research in the future.

The study of ordinary differential equations and partial differential equations provides a new perspective for neural network theory. It also supports research and practice in related fields. With the passage of time, deep learning theory will be developed. Mathematical tools will become better. The application of these equations in neural networks will be expanded. A new method of combining them will emerge. The mixture of these fields will help solve the challenges in the fields of science and engineering. It will provide solid theoretical and practical methods.

References

[1] Chang, Y., Sun, T., Zhang, X., & Chen, X. (2021). Output feedback stabilization for a class of cascade nonlinear ODE-PDE systems. International Journal of Control and Automation, 19, 2519–2528.
[2] Qian, G., Yan, R., Wang, Z., & Zheng, S. (2021). Green's function and its applications in ordinary differential equations. 40, 81–85.
[3] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9, 1735–1780.
[4] Belykh, I., Rachinskii, V., & Belykh, A. (2022). Stability of solutions to systems of nonlinear differential equations with discontinuous right-hand sides: Applications to Hopfield artificial neural networks. Mathematics, 10(9), 1524.
[5] Imran, M. A., Rizwan, K., Abid, H., et al. (2023). Applications of fractional partial differential equations for MHD Casson fluid flow with innovative ternary nanoparticles. Processes, 11(1), 218.
[6] Hu, X., & Jin, C. (2024). Abnormality detection in medical images based on diffusion ordinary differential equations. Journal of Zhejiang Sci-Tech University (Natural Science Edition), 51(6), 851–860.
[7] Rong, Y. (2022). Cell image segmentation based on ordinary differential equations and labeled watershed algorithm.
[8] Wu, P. (2021). Truncation functions and their applications in partial differential equations. Journal of Applied Mathematics, 44(6), 801–806.
[9] Sun, W., Yang, Z., & Yan, X. (2024). Acoustic neural networks based on partial differential equations. Reviews of Geophysics and Planetary Physics, 55(2), 257–266.
[10] Antonelli, M., et al. (2002). Statistics of temperature fluctuations in a buoyancy-dominated boundary layer flow simulated by a large eddy simulation model. Journal of the Atmospheric Sciences, 60, 215–224.
[11] Ruthotto, L., & Haber, E. (2020). Deep neural networks motivated by partial differential equations. Journal of Mathematical Imaging and Vision, 62(3), 352–364.
[12] Mowlavi, S., & Nabi, S. (2023). Optimal control of PDEs using physics-informed neural networks. Journal of Computational Physics, 473, 111731.
[13] Cheng, S. (2023). Low-light image Poisson noise removal based on partial differential equations. Journal of Chongqing University of Science and Technology (Natural Science Edition), 25(4), 60–64.
[14] Wang, H., & Zhang, Z. (2025). PDE-based denoising and deblurring model and experiments. Experimental Technology and Management, 42(7), 151–155.
[15] You, Y., Zhang, Z., Hsieh, C. J., Demmel, J., & Keutzer, K. (2018). ImageNet training in minutes. In Proceedings of the 47th International Conference on Parallel Processing (pp. 1–10).
[16] Chen, Y., Lin, H., Zhang, W., Feng, L., Zheng, C., Zhou, T., ... & Liu, L. (2023). Research progress on medical image dataset expansion methods. Journal of Biomedical Engineering, 40(1), 185–192.
[17] Kelly, C., & Lord, G. J. (2018). Adaptive time-stepping strategies for nonlinear stochastic systems. IMA Journal of Numerical Analysis, 38(3), 1523–1549.