CNM-UNet: Continuous Ordinary Differential Equations for Medical Image Segmentation

Tianqi Xu, Yashi Zhu, Quansong He, Yue Cao, Kaishen Wang, Zhang Yi, Tao He*

College of Computer Science, Sichuan University, Chengdu, China
{xutianqi, zhuyashi, hequansong, caoyue1, wangks}@stu.scu.edu.cn
{zhangyi, tao_he}@scu.edu.cn

Abstract

Integrating Ordinary Differential Equations (ODEs) with U-shaped neural networks has emerged as a novel direction in medical image segmentation. Current networks predominantly employ discretization methods incorporating ODEs. However, these methods face inherent trade-offs between model compactness, computational accuracy, and efficiency. Continuous ODE solutions were rarely studied because they face three limitations: high computational costs, long training time, and poor generalization ability. To address these limitations, we propose an innovative Continuous Neural Memory ODE UNet (CNM-UNet), which replaces all hierarchical decoder layers in vanilla UNet with a single Continuous Neural Memory ODEs Block (CNM-Block) decoder, significantly reducing computation costs and improving training efficiency. CNM-UNet leverages ODEs' dynamic properties to establish continuous temporal feature extraction. For alleviating the generalization problem, a **DU**al **SE**lf-updated (DUSE) strategy based on test-time adaptation principles is introduced to enhance cross-domain generalization. Experimental results demonstrate CNM-UNet's comprehensive advantages in computational capacity, convergence speed, and cross-domain adaptability, offering new insights for practical deployment of continuous ODE methodologies for medical image segmentation.

Code — https://github.com/Tianqi04/CNM-UNet

Introduction

Medical image segmentation plays a pivotal role in healthcare by precisely delineating anatomical structures and pathological regions, providing critical decision support for clinical diagnosis, treatment planning, and pathological research (Xia et al. 2025). In recent years, deep learning has driven revolutionary progress in this domain. UNet (Ronneberger, Fischer, and Brox 2015), with its distinctive encoder-decoder architecture and skip connections, has emerged as the benchmark model for medical image segmentation through multi-level feature extraction and fusion. To enhance segmentation performance, advanced architectures such as ResUNet (Zhang, Liu, and Wang 2018), Attention-UNet (Oktay et al. 2018), and VM-UNet (Ruan, Li, and Xiang 2024) have been developed. However, these improved methods often incur high computational complexity and parameter redundancy, prompting ongoing efforts to balance model performance with lightweight design.

Notably, the integration of Neural Memory ODEs (nmODEs) (Yi 2023) has introduced novel breakthroughs in lightweight medical image segmentation models. Current methods (He et al. 2024, 2025; He, He, and Yi 2023; Wang et al. 2025a,b) constructed parameter-sharing discrete nmODE decoders to replace the conventional UNet's independently designed decoders. These Discretized nmODE UNets (DiODE-UNets) significantly reduce parameter counts and computational complexity, offering a novel paradigm for efficient lightweight segmentation. Although existing DiODE-UNets achieve model lightweighting, their reliance on discretization and iterative approximation of continuous solutions inherently introduces multiple limitations through this approximate estimation of data trends. Numerical solvers employing polynomial approximations via Taylor expansion incur local truncation errors from neglected higher-order terms, which accumulate across iterations into substantial global deviations. Furthermore, strict time-step constraints and computational instability risk triggering error explosions and phase shift phenomena, undermining the temporal smoothness of hidden state evolution (Wanner and Hairer 1996).

Continuous methods based on Neural ODEs (NODEs) (Chen et al. 2018) are widely applied in image recognition (Paoletti et al. 2019; Zakwan, Xu, and Ferrari-Trecate 2022). Representative continuous methods in medical image segmentation include NODEs-UNet (Li et al. 2021), NmODE-KD (Hu et al. 2023), 2nd-order Continuous UNet (Cheng et al. 2023), and Implicit U-KAN 2.0 (Cheng et al. 2025). These continuous methods innovatively integrate continuous modules into medical image segmentation networks, thus enhancing segmentation accuracy. However, these methods face three main limitations: Firstly, the inherent symmetric encoder-decoder architecture requires feature inference across multiple channels and resolutions. The resulting network depth and width significantly increase the parameter scale and computational complexity. When continuous methods replace discrete modules with high-precision ODE solvers, the iterative process further exacerbates the computational load and memory consumption, reducing training

*Corresponding Author.

Properties	CNM-UNet	Continuous UNet	DiODE-UNet	Vanilla UNet
Single Decoder	✓	✗	✗	✗
Continuous	✓	✓	✗	✗
GFlops	**22.53**	394.15	**22.53**	47.07
Params (M)	**9.45**	13.39	**9.45**	13.40
Memory (GB)	5.66	14.24	**4.34**	6.19
Training time (h)	1.80	14.90	**1.12**	1.38
DSC	**0.861**	0.855	0.846	0.841
DSC (Generalization)	**0.322**	0.242	0.288	0.219

Table 1: Overview of CNM-UNet, Continuous UNet, DiODE-UNet, and vanilla UNet (Ronneberger, Fischer, and Brox 2015). "Continuous UNet" refers to a variant where all encoder and decoder layers within the UNet are replaced with continuous ODE modules. "DiODE-UNet" refers to the Explicit Euler Discretized (EED) UNet method (He et al. 2024). All models were trained under identical experimental conditions: 300 epochs, batch size of 8, and using an RTX 4090 GPU. Training speed is reported by total training time (hours), DSC represents the Dice Similarity Coefficient (DSC) metric on the training set (BKAI-IGH-NeoPolyp), and DSC (Generalization) represents the DSC scores on the test set (ETIS-LaribPolypDB). Bold values denote optimal metrics.

efficiency (Table 1). Secondly, high-precision solvers necessitate numerous small-step iterative calculations. These repeated iterations drastically increase the time for each forward and backward pass, significantly prolonging the training period. Finally, repeated iterations of high-precision solvers on the training set can lead to the model becoming overly sensitive to training noise and local details (e.g., specific textures, background environments). This weakens the model's ability to extract critical discriminative features and, in the presence of distribution shifts between training and test sets, tends to impair its generalization on the test sets (Guan and Liu 2022). As shown in Table 1, these drawbacks severely restrict the practical application efficiency of continuous methods in medical image segmentation.

To mitigate the challenges of high computational costs and long training time, we propose Continuous Neural Memory ODE UNet (CNM-UNet). Our innovation lies in replacing all hierarchical decoder layers of vanilla UNet with a single Continuous Neural Memory ODEs Block (CNM-Block), which achieves continuous hidden-state extraction by employing a continuous nmODE solver and enables precise modeling of data distribution characteristics. A multi-scale feature fusion mechanism is implemented, unifying cross-scale features through weighted fusion, upsampling, and convolutional transformations of encoder outputs at each stage. Analytical results demonstrate that this single-decoder architecture maintains a compact parameter scale while delivering superior training efficiency and segmentation accuracy. This method establishes a practical pathway for applying continuous ODE-based methods in real-world medical image segmentation tasks.

To mitigate the limitation in generalization, we integrate a **DU**al **SE**lf-updating (DUSE) strategy based on Test-time Adaptation (TTA) principles. This strategy comprises two core components: Firstly, we design an adaptive parameter update strategy for the CNM-Block based on the characteristics of the test data distribution. Secondly, we develop a prompt-learning-driven Batch Normalization (BN) layers statistic calibration method, which enhances the domain alignment capability of the CNM-Block's BN layers through a differentiated weight amplification strategy. Experiments demonstrate that this method significantly enhances performance on dynamically shifting target test sets while maintaining model lightweightness with minimal parameter overhead. As shown in Table 1, the proposed method exhibits comprehensive performance advantages over Continuous UNet, DiODE-UNet, and vanilla UNet methods.

Related Work

UNet with ODEs

ODEs, which model continuous-time dynamical systems, are essential for analyzing neural network dynamics. The development of Neural ODEs (NODEs) (Chen et al. 2018) has introduced novel perspectives for medical image segmentation. Inspired by these advances, NODEs-UNet (Li et al. 2021) replaced the convolutional layers of UNet with NODE modules. NmODE-KD (Hu et al. 2023) incorporated continuous nmODE modules. 2nd-order Continuous UNet (Cheng et al. 2023) replaced traditional discrete modules with second-order ODEs. Implicit U-KAN 2.0 (Cheng et al. 2025) introduced second-order NODE modules. These continuous methods boost segmentation accuracy by integrating continuous modules. However, their dependence on high-precision ODE solvers significantly increases computational cost, lengthens training time, and reduces domain generalization, thereby limiting practical application.

Yi (2023) proposed a novel Neural Memory ODEs (nmODEs), which decoupled learning neurons from memory neurons to leverage the memory properties of the dynamical system. NmODEs established dynamic mappings where learning units update parameters and memory units project inputs to global attractors, enabling input-to-memory space transformations. Recent studies have extended this framework through ODE discretization methods for medical image segmentation. He et al. (2024) incorporated nmODEs decoders based on explicit Euler (Euler 1845), Heun (Heun et al. 1900), and linear multistep (Bashforth and Adams 1883) methods into diverse UNet variants. Wang et al. (2025a) constructed nmSSM-UNet by fusing nmODEs

with state-space models (SSMs). Wang et al. (2025b) proposed nmS4M-UNet, a lightweight self-supervised variant using discrete ODE-based decoders. He et al. (2025) developed FuseUNet, which integrated a multiscale feature fusion method incorporating linear multistep (Bashforth and Adams 1883), predictor-corrector, and nmODE method. While existing discretization methods exhibit versatility and parameter efficiency in medical image segmentation, their hidden state approximations compromise temporal continuity during dynamical evolution.

Test-time Adaptation

Test-time Adaptation (TTA) aims to adapt models trained in the source domain during inference in an online and source domain-agnostic way to improve prediction performance (Liang, He, and Tan 2024). Mainstream TTA methods leverage model updates by constructing self-supervised auxiliary tasks to guide learning on test data. Sun et al. (2020) predicted image rotations as an auxiliary task to facilitate adaptation. Wang et al. (2021) proposed a test-time entropy minimization scheme that reduces generalization errors by minimizing prediction entropy. Wang et al. (2022) designed weight averaging and prediction enhancement with stochastic source weight restoration. Another TTA paradigm addresses domain shifts by adjusting Batch Normalization (BN) statistics. AdaBN (Li et al. 2019) replaced the source BN statistics with the globally estimated target domain statistics. Chen et al. (2024) generated low-frequency prompts from test domains, dynamically adjusting BN statistics via memory bank and warm-up mechanisms. These methods effectively enhanced model generalization capability.

Methods

The Architecture of CNM-UNet

The architecture of Continuous Neural Memory ODE UNet (CNM-UNet) is illustrated in Figure (1b). The model comprises three main components: an encoder module, a feature fusion module, and a decoder module (CNM-Block). Assuming the input image for the segmentation task is $x \in \mathcal{R}^{C\times W\times H}$, the CNM-UNet's objective is to learn a function $f(x,\theta) \rightarrow y$, where the output segmentation result is $y \in \mathcal{R}^{N\times W\times H}$. Firstly, the encoder module of CNM-UNet shares the same architecture as the vanilla UNet (Ronneberger, Fischer, and Brox 2015). Each stage consists of a double convolution block, which contains two convolutional layers, Batch Normalization (BN), and ReLU activation, followed by a max-pooling layer.

Secondly, in the feature fusion module, CNM-UNet employs a simple yet efficient multi-scale feature fusion strategy: The output of the i-th encoder stage, x_i, is multiplied by a learnable weight parameter w_i, i.e., $w_i \odot x_i$, where $\odot$ denotes element-wise multiplication. These weight parameters are dynamically optimized during backpropagation based on data distribution, effectively adjusting the contribution of each feature level. Subsequently, heterogeneous features are uniformly mapped to the same channel space via 3×3 convolutions. The decoding stage innovatively decouples parameterized and non-parameterized operations: firstly, the fused features are upsampled to the target resolution using bilinear interpolation, yielding $\hat{x}_i \in \mathcal{R}^{C\times W\times H}$; then, feature aggregation is performed by $X = \sum_i \hat{x}_i + b$, where b is a learnable parameter with the same shape as $\hat{x}_i$, enabling dynamic adjustment of feature fusion to construct the input for CNM-Block. This weighted summation effectively merges multi-scale features, enhancing feature representation capability.

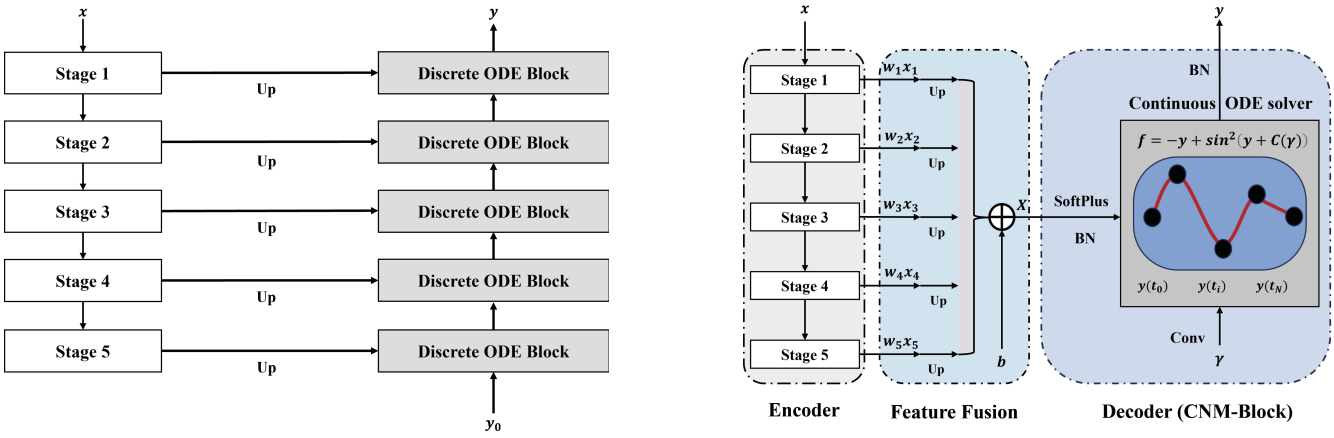
In DiODE-UNet (Figure 1a), each encoder stage corresponds to a discrete ODE decoder module. These decoder modules are limited by low-precision numerical solvers, where the iterative process introduces local truncation errors because of the neglect of higher-order terms. These errors accumulate with each iteration, leading to significant global deviations. Simultaneously, strict time-step constraints can lead to computational instability, severely restricting the accuracy of data trend fitting (Table 1). To migrate these issues, we introduce a CNM-Block, the third component of CNM-UNet, replacing the structure of multiple discrete ODEs.

The CNM-Block employ a continuous ODE solver to achieve a precise fitting of data distribution characteristics through dynamic system modeling. To ensure smooth hidden-state updates aligned with data trends, the Softplus activation function is adopted as the activation function at the CNM-Block input, which theoretically guarantees unique adjoint gradients. After activation and BN, the processed input is fed into the ODE solver as the initial input y. The system's dynamical characteristics are governed by Neural Memory ODEs (nmODEs) (Yi 2023):

$$f = -y + \sin^2\left(y + \mathcal{C}(\gamma)\right), \tag{1}$$

where $\mathcal{C}(\cdot)$ represents the convolution operation. The global learnable parameter $\gamma \in \mathcal{R}^{C\times W\times H}$ is initialized to zero. After a 3×3 convolutional transformation, it is fed into the continuous ODE solver as a phase parameter. This parameter is then automatically optimized via backpropagation to match the frequency characteristics of the target data.

Difference Between CNM-Block and Discrete nmODE Decoder Conventional discrete methods treat y in Equation (1) as a randomly initialized memory parameter, continuously fitting data trends during training, while γ is regarded as an external input parameter from the encoder's skip connection. In contrast, CNM-UNet uses y in Equation (1) as the actual network input, and γ is set as a global memory parameter. This innovative utilization of nmODEs effectively achieves both rapid convergence and high segmentation accuracy. The key innovation lies in using the target variable y to directly represent the input data, providing explicit directional learning signals. This design directly couples loss function dynamics with y optimization, dramatically enhancing gradient update efficiency. The immediate feedback mechanism accelerates feature learning and promotes rapid convergence. Simultaneously, parameter γ is repurposed as an auxiliary global memory parameter that fine-tunes the network's adaptability to data trends. Optimizing γ enables precise approximation of training data patterns, yielding superior output accuracy.



(a) The general architecture of DiODE-UNets

(b) The architecture of CNM-UNet

Figure 1: The comparison between the architecture of DiODE-UNets and the proposed CNM-UNet.

During forward propagation, the system receives the initial state $y(t_0)$, the state derivative f, and the time step Δt. The continuous evolution is approximated by the Dormand-Prince-Shampine 5th-order Runge-Kutta (dopri5) method (Dormand and Prince 1986), ensuring high-precision numerical integration while balancing computational efficiency:

$$y(t_N) = y(t_0) + \int_{t_0}^{t_N} f(y(t), t)dt. \tag{2}$$

Specifically, the continuous-time domain $[t_0, t_N]$ is discretized into N dense time points $t_0 < t_1 < t_2 < ... < t_N$. The interpolation formulas are used to enable high-precision continuous-state approximations between discrete intervals. This continuous feature extraction ensures smooth transitions in learned representations, yielding near-continuous model outputs. The dopri5 solver provides fifth-order accuracy, reducing local truncation errors to $\mathcal{O}(h^6)$, thereby stabilizing feature extraction. Furthermore, the method adaptively adjusts step sizes, dynamically balancing precision and computational efficiency. This strategy significantly enhances the model's capability to fit complex data trends.

We mitigate backpropagation bottlenecks in ODE solvers by using an augmented adjoint method (Chen et al. 2018). Gradients are computed through vector-Jacobian products via adjoint states $\mathbf{a}(t) = \frac{\partial L}{\partial y(t)}$, eliminating intermediate-state storage. This reduces memory complexity to $\mathcal{O}(1)$, significantly enhancing training efficiency for large-scale data. The output of the continuous ODE solver undergoes BN and convolutional operations to produce the final segmentation result $y \in \mathcal{R}^{N \times W \times H}$.

Innovations of CNM-UNet CNM-UNet introduces innovations at two levels: Firstly, the CNM-Block deploys a continuous ODE solver that leverages nmODEs to model feature extraction continuously, enabling accelerated convergence while simultaneously achieving superior segmentation accuracy. Secondly, to migrate the high computational and temporal complexity of the dopri5-based numerical solver, the architecture replaces all multi-stage decoders in vanilla UNet with a single lightweight CNM-Block decoder, achieving significant compression in parameter size, computational overhead, and training time. Through these coordinated optimization strategies, the architecture achieves an optimal balance between model performance and computational efficiency.

Test-time Adaption Method for CNM-UNet

In continuous ODE methods, repeated iterations of high-precision solvers within the continuous modules on the training set can lead to the model becoming overly sensitive to training noise and irrelevant local details (e.g., specific textures, background). This weakens its ability to extract critical discriminative features. Consequently, when distribution shifts exist between the training and test sets, the model's generalization on the test sets is limited (Guan and Liu 2022; He et al. 2024). Specifically, for CNM-UNet, this phenomenon manifests in two parts: (1) parameter bias. The global learnable parameter γ of Equation (1) in the continuous ODE solver predominantly fits the feature of the training set, failing to adapt to test sets distribution shifts. (2) misalignment of the BN statistics. Inspired by prior works (Hu, Liao, and Xia 2022; Zhang et al. 2023; Chen et al. 2024), we observe significant discrepancies between the BN statistics calculated on the test features (mean μ_t, variance σ_t) and stored in the source model (mean μ_s, variance σ_s), particularly within the BN layer of the CNM-Block. To mitigate these issues, we design a **DU**al **SE**lf-updating (DUSE) strategy, as shown in Algorithm 1, to alleviate the impacts of distribution shifts on CNM-Block. The adapted image undergoes forward propagation through the weight-optimized CNM-Block to generate the optimized segmentation result.

Firstly, to dynamically adjust the parameter γ in CNM-Block during testing, we develop a gradient statistics-based adaptation mechanism inspired by BN's moving average. During training, for each time step i $(0 \leq i \leq N)$ in the CNM-Block, we feed its output y_i into Equation (1) to obtain the corresponding state gradient derivative f_i. We then record the mean $(\mu_s^{f_i})$ and variance $(\sigma_s^{f_i})$ of f_i and update

global statistics (mean $\mathcal{M}_{f_i}$, variance $\mathcal{V}_{f_i}$) via:

$$\begin{cases} \mathcal{M}_{f_i} = (1-m)\mathcal{M}_{f_i} + m\mu_s^{f_i} \\ \mathcal{V}_{f_i} = (1-m)\mathcal{V}_{f_i} + m\sigma_s^{f_i} \end{cases}, \tag{3}$$

where m is a momentum coefficient to balance historical information with current batch statistics.

During testing, we compute real-time statistics ($\mu_t^{f_i}$, $\sigma_t^{f_i}$) of state gradient derivatives f_i across all time steps in CNM-Block for test set data. Significant deviations emerge between the globally tracked statistics ($\mathcal{M}_{f_i}$, $\mathcal{V}_{f_i}$) from the training set and test set statistics ($\mu_t^{f_i}$, $\sigma_t^{f_i}$) under substantial distribution shifts. We construct an absolute-distance-based loss:

$$\mathcal{L}_n = \frac{1}{i}\sum_i \left(|\mu_t^{f_i} - \mathcal{M}_{f_i}| + |\sigma_t^{f_i} - \mathcal{V}_{f_i}| \right). \tag{4}$$

This loss dynamically adjusts the weights of the convolutional layer θ_γ within the CNM-Block that govern the parameters γ in Equation 1. To prevent error propagation, we implement single-sample parameter updating, where pretrained weights are immediately restored after each test image's adaptation.

Secondly, inspired by Chen et al. (2024), we generate a low-frequency prompt $\mathcal{P}_i \in R^{C\times(W\times\alpha)\times(H\times\alpha)}$ per test image $x_i \in \mathcal{R}^{C\times W\times H}$. The adapted image $\hat{x}_i$ is obtained through:

$$\hat{x}_i = \mathcal{F}^{-1}\left(\left[\text{OnePad}(\mathcal{P}_i) \odot \mathcal{F}^A(x_i), \mathcal{F}^P(x_i) \right] \right), \tag{5}$$

where $\mathcal{F}^A(\cdot)$, $\mathcal{F}^P(\cdot)$ and $\mathcal{F}^{-1}(\cdot)$ denote the amplitude component, phase component of the Fast Fourier Transform (FFT), and Inverse Fast Fourier Transform (IFFT) operations (Frigo and Johnson 1998), respectively. $\odot$ represents element-wise multiplication, and OnePad performs unitary padding to expand $\mathcal{P}_i$ to dimensions of $\mathcal{R}^{C\times W\times H}$.

To optimize $\mathcal{P}_i$ for adaptive BN statistic calibration, we develop a dynamically weighted BN loss function:

$$\mathcal{L}_p = \frac{H}{i}\sum_i \left(|\mu_s^i - \mu_t^i| + |\sigma_s^i - \sigma_t^i| \right), \tag{6}$$

where a differentiated weight allocation strategy is applied. The standard BN layers are assigned H_{base}, while CNM-Block BN layer adopts a strengthened coefficient $H_{CNM} = k \cdot H_{base}$, where k is a scaling factor. This strategy enforces enhanced alignment of BN statistics within CNM-Block, significantly improves its cross-domain generalization.

Innovations of DUSE Firstly, a test-distribution-aware adaptive parameter update mechanism is tailored for CNM-Block, enabling dynamic adjustments based on target-set data patterns. Secondly, a prompt learning-driven BN statistic calibration method is proposed, which significantly enhances domain alignment in CNM-Block BN layers through a differentiated weight allocation strategy.

Experiments

The experimental design comprises: (1) CNM-UNet segmentation validation on the source set. (2) Generalization performance assessment of the **DU**al **SE**lf-updating (DUSE) strategy. (3) Ablation study.

Algorithm 1: DUSE algorithm

Input: Current test image x_i, initialized prompt $\mathcal{P}_i$, CNM-Block weights θ_γ

Compute adapted image $\hat{x}_i$ by Equation (5)
First Forward Pass:
  Compute $\mathcal{L}_n$ by Equation (4)
  Compute $\mathcal{L}_p$ by Equation (6)
First Backward Pass:
  Update θ_γ: $\theta_\gamma \leftarrow \theta_\gamma - \eta\nabla_{\theta_\gamma}\mathcal{L}_n$
  Update $\mathcal{P}_i$: $\mathcal{P}_i \leftarrow \mathcal{P}_i - \eta\nabla_{\mathcal{P}_i}\mathcal{L}_p$
Recompute $\hat{x}_i$ by Equation (5)
Second Forward Pass:
  Compute $\mathcal{L}_n$ by Equation (4)
Second Backward Pass:
  Update θ_γ: $\theta_\gamma \leftarrow \theta_\gamma - \eta\nabla_{\theta_\gamma}\mathcal{L}_n$

Output: Adapted image $\hat{x}_i$, optimized weights θ_γ

Datasets

ISIC2018 Dataset contains approximately 2,700 dermoscopic images. Images were normalized and resized to 256×256 resolution. **OD/OC Dataset** integrates five public ophthalmic imaging resources: A (RIM-ONE-r3 (Fumero et al. 2011)), B (REFUGE (Orlando et al. 2020)), C (ORIGA (Zhang et al. 2010)), D (Drishti-GS (Sivaswamy et al. 2014)) and E (REFUGE validation/test set (Orlando et al. 2020)), comprising a total of over 2,000 cases. Images were processed to 512×512 resolution. **Polyp Dataset** combines four public medical imaging collections: A (CVC-ClinicDB (Bernal et al. 2015)), B (ETIS-LaribPolypDB (Silva et al. 2014)), C (Kvasir-Seg (Jha et al. 2019)), and D (BKAI-IGH-NeoPolyp (Ngoc Lan et al. 2021)), totaling over 2,800 cases. Images were resized to 352×352 resolution.

Implementation Details

During training, the continuous ODE solver timestep Δt was set to 0.05, with data augmentation via horizontal flipping, vertical flipping, and random rotation. The AdamW optimizer was used with a CosineAnnealingLR scheduler, with a maximum of 50 iterations and a minimum learning rate of 1×10^{-5}. Training spanned 300 epochs with a batch size of 8. All experiments were conducted on a single RTX 4090 GPU using PyTorch. Experiments were repeated five times and the results were reported as mean and standard deviation for each dataset. For DUSE evaluation, single-source training with multi-target testing was implemented. Key parameters include a prompt size $\alpha = 0.01$, momentum coefficient $m = 0.2$, $H_{base} = 0.5$, with the scaling factor k set to 500 for OD/OC dataset and 1000 for Polyp dataset. The CNM-Block employed the Adam optimizer for parameter updates, with task-specific initial learning rates: 0.005 (OD/OC dataset) and 0.01 (Polyp dataset). Evaluation metrics included Mean Intersection over Union (mIoU) and Dice Similarity Score (DSC).

Comparative Results

CNM-UNet vs. Other Methods CNM-UNet is validated on the ISIC2018, Polyp-A, Polyp-C, and Polyp-D datasets,

Dataset	Model	Params(M)	GFLOPs	mIoU	DSC
ISIC18	UNet (Ronneberger, Fischer, and Brox 2015)∗	13.40	31.12	0.773	0.855
	ResUNet (Diakogiannis et al. 2020)∗	35.99	58.56	0.772	0.854
	Att-UNet (Oktay et al. 2018)∗	26.16	53.31	0.776	0.857
	UNet (LMD) (He et al. 2024)	9.45	11.91	0.775±0.002	0.873±0.001
	nmS4M-UNet (Wang et al. 2025b)∗	9.42	11.77	0.780	0.878
	nmSSM-UNet (Wang et al. 2025a)∗	9.41	11.77	0.779	0.876
	Continuous UNet	13.39	209.10	**0.793±0.002**	**0.885±0.001**
	CNM-UNet (Ours)	9.45	11.92	0.784±0.001	0.879±0.001
Polyp-A	UNet (Ronneberger, Fischer, and Brox 2015)	13.40	47.07	0.577±0.010	0.732±0.008
	UNet (EED) (He et al. 2024)	9.45	22.53	0.588±0.007	0.741±0.007
	MALUNet (Ruan et al. 2022)	0.18	0.13	0.392±0.005	0.563±0.005
	EGE-UNet (Ruan et al. 2023)	0.05	0.11	0.479±0.006	0.648±0.006
	nmS4M-UNet (Wang et al. 2025b)	9.90	33.03	0.600±0.009	0.750±0.007
	Continuous UNet	13.39	394.15	0.600±0.008	0.750±0.006
	CNM-UNet (Ours)	9.45	22.53	**0.606±0.009**	**0.754±0.007**
Polyp-C	UNet (Ronneberger, Fischer, and Brox 2015)	13.40	47.07	0.662±0.006	0.797±0.004
	UNet (EED) (He et al. 2024)	9.45	22.53	0.676±0.006	0.807±0.004
	MALUNet (Ruan et al. 2022)	0.18	0.13	0.508±0.008	0.674±0.007
	EGE-UNet (Ruan et al. 2023)	0.05	0.11	0.509±0.006	0.675±0.005
	nmS4M-UNet (Wang et al. 2025b)	9.90	33.03	0.665±0.009	0.799±0.007
	Continuous UNet	13.39	394.15	0.668±0.010	0.801±0.007
	CNM-UNet (Ours)	9.45	22.53	**0.687±0.008**	**0.814±0.006**
Polyp-D	UNet (Ronneberger, Fischer, and Brox 2015)	13.40	47.07	0.726±0.004	0.841±0.003
	Ulite (Dinh et al. 2023)	0.88	1.14	0.667±0.001	0.800±0.001
	UNet (EED) (He et al. 2024)	9.45	22.53	0.732±0.006	0.846±0.004
	Att-UNet (Oktay et al. 2018)	34.88	100.78	0.684±0.008	0.812±0.005
	MALUNet (Ruan et al. 2022)	0.18	0.13	0.540±0.008	0.701±0.007
	EGE-UNet (Ruan et al. 2023)	0.05	0.11	0.575±0.004	0.730±0.004
	nmS4M-UNet (Wang et al. 2025b)	9.90	33.03	0.737±0.006	0.848±0.004
	Continuous UNet	13.39	394.15	0.743±0.011	0.852±0.007
	CNM-UNet (Ours)	9.45	22.53	**0.756±0.005**	**0.861±0.004**

Table 2: Comparison of mIoU and DSC on ISIC2018, Polyp-A, Polyp-C and Polyp-D datasets. ∗ means the experimental results from Wang et al. (2025b). "Continuous UNet" refers to a variant where all encoder and decoder layers within the UNet are replaced with continuous ODE modules. Bold values denote optimal metrics.

with comparative results against mainstream models summarized in Table 2. To ensure fair comparison, the experimental design excluded the use of the DUSE strategy. In terms of model performance, CNM-UNet achieves the best results on the Polyp-A, Polyp-C, and Polyp-D datasets and ranks second on the ISIC2018 dataset. Compared to Continuous UNet, CNM-UNet exhibits slight performance improvements on Polyp-A, C, and D datasets, particularly achieving gains of 0.019 mIoU and 0.013 DSC on Polyp-C dataset. Compared to lightweight models like MALUNet and EGE-UNet, CNM-UNet demonstrates more significant improvements, with increases exceeding 0.127 in mIoU and 0.106 in DSC. Regarding model complexity, CNM-UNet reduces parameters by 3.95M compared to vanilla UNet, while reducing GFLOPs by 19.20 and 24.54 on ISIC2018 and Polyp datasets, respectively. The parameter and computational reductions are more pronounced for the ResUNet and Att-UNet variants. Compared to Continuous UNet, CNM-UNet shows a 3.94M reduction in parameters and a remarkable 197.18 and 371.62 decrease in GFLOPs on the ISIC2018 and Polyp datasets, respectively. Notably, CNM-UNet exhibits comparable model complexity to several Discretized nmODE UNets (DiODE-UNets), including UNet (LMD), UNet (EED), nmSSM-UNet and nmS4M-UNet. The experimental results indicate that CNM-UNet achieves high segmentation accuracy with low computational costs, effectively balancing precision and efficiency.

DUSE vs. Other TTA Methods Using CNM-UNet as the benchmark, as shown in Table 3, we integrated the DUSE strategy and compared its generalization performance with mainstream TTA methods. The "Source Only" baseline indicates the original model without any TTA methods applied. On the OD/OC dataset, DUSE improved DSC values over "Source Only" baseline and comparative methods for test sets A, D and continuous scenarios, achieving improvements of 0.374 on set A and 0.214 on set D. On the Polyp dataset, DUSE outperforms the "Source Only" baseline and comparative methods for test sets A, B, and continuous scenarios, with DSC improvements of 0.064 on set A and 0.084 on set B. It is worth noting that parameter-updating methods like AdaBN, Tent, and CoTTA suffer from error accumula-

Method	OD/OC Dataset					Polyp Dataset			
	A	B	C	D	Con	A	B	C	Con
Source Only	0.335	0.789	**0.732**	0.498	0.683	0.495	0.237	**0.656**	0.550
AdaBN (Li et al. 2019)	0.175	0.150	0.096	0.433	0.388	0.225	0.142	0.234	0.417
Tent (Wang et al. 2021)	0.445	0.445	0.344	0.678	0.237	0.488	0.320	0.595	0.525
CoTTA (Wang et al. 2022)	0.688	0.561	0.430	0.625	0.455	0.488	0.320	0.595	0.525
EATA (Niu et al. 2022)	0.704	**0.851**	0.663	0.662	0.737	0.488	0.320	0.596	0.526
SAR (Niu et al. 2023)	0.705	0.841	0.667	0.663	0.740	0.488	0.320	0.596	0.526
VPTTA (Chen et al. 2024)	0.684	0.818	0.677	0.692	0.732	0.554	0.317	0.624	0.568
DUSE (Ours)	**0.709**	0.827	0.709	**0.712**	**0.766**	**0.559**	**0.321**	0.626	**0.573**

Table 3: Generalization performance comparison between DUSE and other TTA methods on OD/OC and Polyp dataset. Models are trained on OD/OC-E and Polyp-D for their respective datasets. "Con" represents sequential testing across test sets $A \rightarrow B \rightarrow C \rightarrow D$ for the OD/OC dataset and $A \rightarrow B \rightarrow C$ for the Polyp dataset. All reported values represent DSC. Bold values denote optimal metrics. Variances are omitted due to minimal values.

tion and catastrophic forgetting due to pretrained parameter modifications (Chen et al. 2024), which can lead to performance degradation. In contrast, prompt-learning methods including VPTTA and our DUSE strategy effectively mitigate these issues by generating learnable prompts without updating model parameters, outperforming the "Source Only" baseline across most test sets.

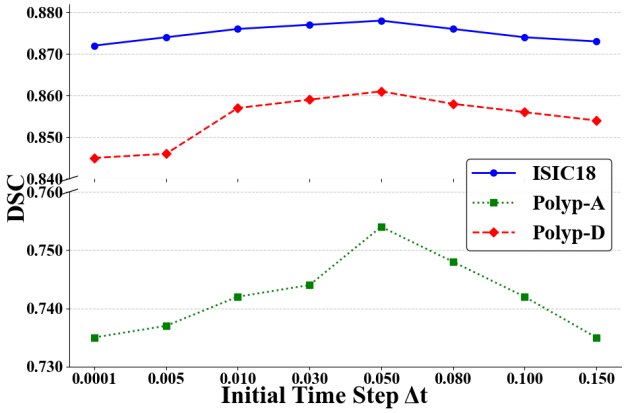


Figure 2: Performance of CNM-UNet with various initial time step Δt on ISIC2018, Polyp-A, and Polyp-D dataset.

Ablation Study

Analysis of Initial Time Step Δt in Continuous ODE Solver We conduct a parameter sensitivity analysis on the initial time step Δt of the continuous ODE solver in CNM-UNet on the ISIC2018, Polyp-A, and Polyp-D dataset. As shown in Figure 2, optimal convergence is achieved at $\Delta t = 0.05$. Excessively large Δt values induce significant accumulation of numerical errors in the ODE solver, while overly small Δt values exacerbate gradient instability.

Ablation Analysis Under Image Corruption Following Zhang et al. (2022), corruptions at severity level 5 were implemented on the ISIC18 dataset, including six types: Brightness, Contrast, Defocus Blur, Gaussian Noise, Shot Noise, and Saturate. As shown in Table 4, the DUSE strategy consistently outperformed the "Source Only" baseline in mIoU and DSC across all corruption conditions. Notably, substantial improvements in mIoU and DSC were observed under Contrast, Gaussian Noise, and Shot Noise. The results demonstrate that the DUSE strategy effectively enhances model robustness to image corruption.

Corruption	Source Only		DUSE (Ours)	
	mIoU	DSC	mIoU	DSC
Brightness	0.661	0.796	**0.728**	**0.842**
Contrast	0.134	0.236	**0.554**	**0.713**
Defocus Blur	0.743	0.853	**0.752**	**0.858**
Gaussian Noise	0.000	0.000	**0.434**	**0.606**
Shot Noise	0.000	0.000	**0.504**	**0.670**
Saturate	0.438	0.609	**0.461**	**0.631**

Table 4: Robustness evaluation under severity level 5 corruptions on ISIC2018 dataset. Models are trained on the original ISIC2018 dataset without corruptions. Bold values denote optimal metrics. Variances are omitted due to minimal values.

Conclusion

In the paper, we propose the Continuous Neural Memory ODE UNet (CNM-UNet) for medical image segmentation, which replaces all hierarchical decoder layers in vanilla UNet with a single Continuous Neural Memory ODEs Block (CNM-Block) decoder, significantly reducing computation costs and improving training efficiency. CNM-Block leverages Neural Memory ODEs (nmODEs) to establish continuous temporal feature extraction. To alleviate the generalization problem, we innovatively introduce a **DU**al **SE**lf-updated (DUSE) strategy based on test-time adaptation principles, which significantly enhances model cross-domain generalization. Experimental results demonstrate that CNM-UNet achieves improved segmentation accuracy across multiple datasets while maintaining competitive parameter efficiency and training speed. Furthermore, the DUSE strategy notably boosts the model's adaptability in test scenarios involving substantial distribution shifts. We hope that our work can establish a novel technical pathway for continuous ODE methodologies in medical image segmentation.

Acknowledgments

This work was supported by the Major Program of the National Natural Science Foundation of China under Grant No. 62495064, the Youth Program of the National Natural Science Foundation of China under Grant No. 62206189, and the Innovation Research Groups Project of the Natural Science Foundation of Sichuan Province under Grant No. 2024NSFTD0051.

References

Bashforth, F.; and Adams, J. C. 1883. *An Attempt to Test the Theories of Capillary Action: By Comparing the Theoretical and Measured Forms of Drops of Fluid*. University Press.

Bernal, J.; Sánchez, F. J.; Fernández-Esparrach, G.; Gil, D.; Rodríguez, C.; and Vilariño, F. 2015. WM-DOVA Maps for Accurate Polyp Highlighting in Colonoscopy: Validation vs. Saliency Maps From Physicians. *Computerized Medical Imaging and Graphics*, 43: 99–111.

Chen, R. T. Q.; Rubanova, Y.; Bettencourt, J.; and Duvenaud, D. K. 2018. Neural Ordinary Differential Equations. In *Advances in Neural Information Processing Systems (NeurIPS)*, 6572—-6583.

Chen, Z.; Pan, Y.; Ye, Y.; Lu, M.; and Xia, Y. 2024. Each Test Image Deserves A Specific Prompt: Continual Test-Time Adaptation for 2D Medical Image Segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 11184–11193.

Cheng, C.-W.; Runkel, C.; Liu, L.; Chan, R. H.; Schönlieb, C.-B.; and Aviles-Rivero, A. I. 2023. Continuous U-Net: Faster, Greater and Noiseless.

Cheng, C.-W.; Zhao, Y.; Cheng, Y.; Montoya, J.; Schönlieb, C.-B.; and Aviles-Rivero, A. I. 2025. Implicit U-KAN2. 0: Dynamic, Efficient and Interpretable Medical Image Segmentation. *arXiv preprint arXiv:2503.03141*.

Diakogiannis, F. I.; Waldner, F.; Caccetta, P.; and Wu, C. 2020. ResUNet-a: A Deep Learning Framework for Semantic Segmentation of Remotely Sensed Data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 162: 94–114.

Dinh, B.-D.; Nguyen, T.-T.; Tran, T.-T.; and Pham, V.-T. 2023. 1M Parameters Are Enough? A Lightweight CNN-Based Model for Medical Image Segmentation. In *Asia Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, 1279–1284.

Dormand, J.; and Prince, P. 1986. A Reconsideration of Some Embedded Runge—Kutta Formulae. *Journal of computational and applied mathematics*, 15(2): 203–211.

Euler, L. 1845. *Institutionum Calculi Integralis*. Number 4 in Institutionum calculi integralis. Impensis Academiae Imperialis Scientiarum.

Frigo, M.; and Johnson, S. G. 1998. FFTW: An Adaptive Software Architecture for the FFT. In *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 1381–1384.

Fumero, F.; Alayon, S.; Sanchez, J. L.; Sigut, J.; and Gonzalez-Hernandez, M. 2011. RIM-ONE: An Open Retinal Image Database for Optic Nerve Evaluation. In *International Symposium on Computer-Based Medical Systems (CBMS)*, 1–6.

Guan, H.; and Liu, M. 2022. Domain Adaptation for Medical Image Analysis: A Survey. *IEEE Transactions on Biomedical Engineering*, 69(3): 1173–1185.

He, Q.; He, T.; and Yi, Z. 2023. Medical Image Segmentation Using Discretized nmODE. In *International Annual Conference on Complex Systems and Intelligent Science (CSIS-IAC)*, 497–03.

He, Q.; Min, X.; Wang, K.; and He, T. 2025. FuseUNet: A Multi-Scale Feature Fusion Method for U-like Networks. *arXiv preprint arXiv:2506.05821*.

He, Q.; Yao, X.; Wu, J.; Yi, Z.; and He, T. 2024. A Lightweight U-like Network Utilizing Neural Memory Ordinary Differential Equations for Slimming the Decoder. In *International Joint Conferences on Artificial Intelligence Organization (IJCAI)*, 821–829.

Heun, K.; et al. 1900. Neue Methoden zur approximativen Integration der Differentialgleichungen einer unabhängigen Veränderlichen. *Zeitschrift fur Angewandte Mathematik und Physik (Z. Math. Phys)*, 45: 23–38.

Hu, J.; Yu, C.; Yi, Z.; and Zhang, H. 2023. Enhancing Robustness of Medical Image Segmentation Model With Neural Memory Ordinary Differential Equation. *International Journal of Neural Systems*, 33(12): 2350060.

Hu, S.; Liao, Z.; and Xia, Y. 2022. ProSFDA: Prompt Learning based Source-Free Domain Adaptation for Medical Image Segmentation. arXiv:2211.11514.

Jha, D.; Smedsrud, P. H.; Riegler, M. A.; Halvorsen, P.; De Lange, T.; Johansen, D.; and Johansen, H. D. 2019. Kvasir-SEG: A Segmented Polyp Dataset. In *International Conference on Multimedia Modeling (MMM)*, 451–462.

Li, D.; Tang, P.; Zhang, R.; Sun, C.; Li, Y.; Qian, J.; Liang, Y.; Yang, J.; and Zhang, L. 2021. Robust Blood Cell Image Segmentation Method Based on Neural Ordinary Differential Equations. *Computational and Mathematical Methods in Medicine*, 2021(1): 5590180.

Li, Y.; Wang, N.; Shi, J.; Liu, J.; and Hou, X. 2019. Revisiting Batch Normalization For Practical Domain Adaptation. *Pattern Recognition*, 96: 106996.

Liang, J.; He, R.; and Tan, T. 2024. A Comprehensive Survey on Test-Time Adaptation Under Distribution Shifts. *International Journal of Computer Vision (IJCV)*, 133(1): 31–64.

Ngoc Lan, P.; An, N. S.; Hang, D. V.; Long, D. V.; Trung, T. Q.; Thuy, N. T.; and Sang, D. V. 2021. NeoUNet: Towards Accurate Colon Polyp Segmentation and Neoplasm Detection. In *Advances in visual computing (ISVC)*, 15–28.

Niu, S.; Wu, J.; Zhang, Y.; Chen, Y.; Zheng, S.; Zhao, P.; and Tan, M. 2022. Efficient Test-Time Model Adaptation without Forgetting. In *International Conference on Machine Learning (ICML)*, 16888–16905.

Niu, S.; Wu, J.; Zhang, Y.; Wen, Z.; Chen, Y.; Zhao, P.; and Tan, M. 2023. Towards Stable Test-Time Adaptation in Dynamic Wild World. In *Internetional Conference on Learning Representations (ICLR)*.

Oktay, O.; Schlemper, J.; Folgoc, L. L.; Lee, M.; Heinrich, M.; Misawa, K.; Mori, K.; McDonagh, S.; Hammerla, N. Y.; Kainz, B.; Glocker, B.; and Rueckert, D. 2018. Attention U-Net: Learning Where to Look for the Pancreas. In *Conference on Medical Imaging with Deep Learning (MIDL)*.

Orlando, J. I.; Fu, H.; Breda, J. B.; Van Keer, K.; Bathula, D. R.; Diaz-Pinto, A.; Fang, R.; Heng, P.-A.; Kim, J.; Lee, J.; et al. 2020. REFUGE Challenge: A Unified Framework for Evaluating Automated Methods for Glaucoma Assessment from Fundus Photographs. *Medical Image Analysis (MEDIA)*, 59: 101570.

Paoletti, M. E.; Haut, J. M.; Plaza, J.; and Plaza, A. 2019. Neural Ordinary Differential Equations for Hyperspectral Image Classification. *IEEE Transactions on Geoscience and Remote Sensing*, 58(3): 1718–1734.

Ronneberger, O.; Fischer, P.; and Brox, T. 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 234–241.

Ruan, J.; Li, J.; and Xiang, S. 2024. VM-UNet: Vision Mamba UNet for Medical Image Segmentation. arXiv:2402.02491.

Ruan, J.; Xiang, S.; Xie, M.; Liu, T.; and Fu, Y. 2022. MALUNet: A Multi-Attention and Light-Weight UNet for Skin Lesion Segmentation. In *IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 1150–1156.

Ruan, J.; Xie, M.; Gao, J.; Liu, T.; and Fu, Y. 2023. EGE-UNet: An Efficient Group Enhanced UNet for Skin Lesion Segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 481—-490.

Silva, J.; Histace, A.; Romain, O.; Dray, X.; and Granado, B. 2014. Toward Embedded Detection of Polyps in WCE Images for Early Diagnosis of Colorectal Cancer. *International Journal of Computer Assisted Radiology and Surgery*, 9(2): 283–293.

Sivaswamy, J.; Krishnadas, S. R.; Datt Joshi, G.; Jain, M.; and Syed Tabish, A. U. 2014. Drishti-GS: Retinal Image Dataset for Optic Nerve Head (ONH) Segmentation. In *IEEE International Symposium on Biomedical Imaging (ISBI)*, 53–56.

Sun, Y.; Wang, X.; Liu, Z.; Miller, J.; Efros, A.; and Hardt, M. 2020. Test-Time Training with Self-Supervision for Generalization under Distribution Shifts. In *Proceedings of the 37th International Conference on Machine Learning (ICMI)*, 9229–9248.

Wang, D.; Shelhamer, E.; Liu, S.; Olshausen, B.; and Darrell, T. 2021. Tent: Fully Test-Time Adaptation by Entropy Minimization. In *International Conference on Learning Representations (ICLR)*.

Wang, Q.; Fink, O.; Van Gool, L.; and Dai, D. 2022. Continual Test-Time Domain Adaptation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 7201–7211.

Wang, Z.; Gu, J.; Zhou, W.; He, Q.; Zhao, T.; Guo, J.; Lu, L.; He, T.; and Bu, J. 2025a. Neural Memory State Space Models for Medical Image Segmentation. *International Journal of Neural Systems (IJNS)*, 35(01): 2450068.

Wang, Z.; He, Y.; Yi, Z.; He, T.; and Bu, J. 2025b. Neural Memory Self-Supervised State Space Models With Learnable Gates. *IEEE Signal Processing Letters (SPL)*, 926–930.

Wanner, G.; and Hairer, E. 1996. *Solving Ordinary Differential Equations II*, volume 375. Springer.

Xia, Q.; Zheng, H.; Zou, H.; Luo, D.; Tang, H.; Li, L.; and Jiang, B. 2025. A Comprehensive Review of Deep Learning for Medical Image Segmentation. *Neurocomputing*, 613: 128740.

Yi, Z. 2023. NmODE: Neural Memory Ordinary Differential Equation. *The Artificial intelligence review*, 56(12): 14403–14438.

Zakwan, M.; Xu, L.; and Ferrari-Trecate, G. 2022. Robust Classification Using Contractive Hamiltonian Neural ODEs. *IEEE Control Systems Letters*, 7: 145–150.

Zhang, J.; Qi, L.; Shi, Y.; and Gao, Y. 2023. DomainAdaptor: A Novel Approach to Test-Time Adaptation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 18971–18981.

Zhang, Y.; Sun, Y.; Li, H.; Zheng, S.; Zhu, C.; and Yang, L. 2022. Benchmarking the Robustness of Deep Neural Networks to Common Corruptions in Digital Pathology. In *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 242—252.

Zhang, Z.; Liu, Q.; and Wang, Y. 2018. Road Extraction by Deep Residual U-Net. *IEEE Geoscience and Remote Sensing Letters (GRSL)*, 749—-753.

Zhang, Z.; Yin, F. S.; Liu, J.; Wong, W. K.; Tan, N. M.; Lee, B. H.; Cheng, J.; and Wong, T. Y. 2010. ORIGA-Light: An Online Retinal Fundus Image Database for Glaucoma Analysis and Research. In *Annual International Conference of the IEEE Engineering in Medicine and Biology (EMBC)*, 3065–3068.