

LaTiM: Longitudinal representation learning in continuous-time models to predict disease progression

Rachid Zeghlache^{1,2}, Pierre-Henri Conze^{1,3}, Mostafa El Habib Daho^{1,2}, Yihao Li^{1,2}, Hugo Le Boité⁵, Ramin Tadayoni⁵, Pascale Massin⁵, Béatrice Cochener^{1,2,4}, Alireza Rezaei^{1,2}, Ikram Brahim^{1,2}, Gwenolé Quéllec¹, and Mathieu Lamard^{1,2}

¹ LaTIM UMR 1101, Inserm, Brest, France

² University of Western Brittany, Brest, France

³ IMT Atlantique, Brest, France

⁴ Ophthalmology Department, CHRU Brest, Brest, France

⁵ Lariboisière Hospital, AP-HP, Paris, France

Abstract. This work proposes a novel framework for analyzing disease progression using time-aware neural ordinary differential equations (NODE). We introduce a "time-aware head" in a framework trained through self-supervised learning (SSL) to leverage temporal information in latent space for data augmentation. This approach effectively integrates NODEs with SSL, offering significant performance improvements compared to traditional methods that lack explicit temporal integration. We demonstrate the effectiveness of our strategy for diabetic retinopathy progression prediction using the OPHDIAT database. Compared to the baseline, all NODE architectures achieve statistically significant improvements in area under the ROC curve (AUC) and Kappa metrics, highlighting the efficacy of pre-training with SSL-inspired approaches. Additionally, our framework promotes stable training for NODEs, a commonly encountered challenge in time-aware modeling.

1 Introduction

Deep learning has embraced self-supervised learning (SSL) for representation learning in downstream tasks. Existing SSL methods often rely on contrastive learning [5] or hand-crafted pretext tasks [26]. The latter leverages inherent data properties to automatically generate supervisory signals without manual annotations. While effective, hand-crafted tasks require domain-specific knowledge. Recently, longitudinal SSL approaches have emerged for disease progression analysis, aiming to capture disease evolution at patient or population levels [23, 19, 27, 6, 20, 17, 11]. Longitudinal self-supervised learning (LSSL) was initially introduced in the context of disease progression as a pretext task [21] involving a Siamese-like model. The model takes as input a consecutive pair of images and predicts the difference in time between the two examinations. Since then, more

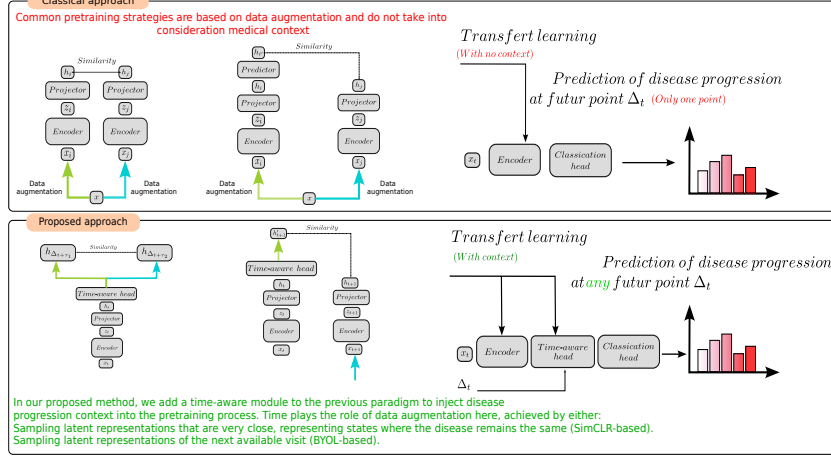


Fig. 1: a), c) refer to SimCLR [5] and BYOL [9], and e) refer to the standard paradigm for predicting disease progression at a certain point in the future, using a single image. b), d), and f) adapt the common SSL paradigm and supervised classification to disease progression with the introduction of a time-aware head.

sophisticated LSSL algorithms have been proposed. The framework in [27] attempted to theorize the notion of longitudinal pretext task with the purpose of learning the disease progression. LSSL was embedded in an auto-encoder (AE), taking two consecutive longitudinal scans as inputs. A cosine alignment term was added to the classic reconstruction loss to force the topology of the latent space to change in the direction of longitudinal changes. A Siamese-like architecture was developed in [21, 12] to compare longitudinal imaging data with deep learning. The strengths of this approach were to avoid any registration requirements, leverage population-level data to capture time-irreversible changes that are shared across individuals and offer the ability to visualize individual-level changes. The previously mentioned methods primarily focus on learning an adequate feature extractor for disease progression analysis. While these approaches offer valuable contributions, we claim that incorporating a time-aware component during the learning process can lead to more accurate disease progression models.

Neural ordinary differential equations (NODEs) [4] offer a promising approach for modeling time-dependent processes like disease progression. NODEs define the relationship between input and output through the solution of an ordinary differential equation (ODE). Moreover, this allows to effectively handle irregular time series data, a common feature in disease progression analysis [22]. However, current disease progression models often lack a dedicated time-aware component. We propose to integrate a time-aware head based on a NODE architecture between the encoding and classification stages. This approach aims to condition the predictions with time, towards more accurate models [14, 24, 25].

Recent works support the use of NODEs for disease progression. Studies have employed NODEs to model COVID-19 progression [18] and predict Alzheimer’s disease progression [14]. This work investigates the hypothesis that incorporating a time-aware head based on NODE can outperform traditional classification heads for disease progression tasks [1]. We claim that this approach can lead to a more informative latent space representation, facilitating more advanced disease progression analysis. To establish our proposed framework, we define and train a set of similarity criteria specifically tailored to the learning paradigm. These criteria, referred to as *temporal evolution* (TE), *temporal consistency* (TC), and *disease progression alignment* (DPA), are critical to guide the learning process.

This paper presents the following key contributions:

1. We propose a novel framework for pre-training time-aware models to address disease progression downstream tasks, specifically focusing on diabetic retinopathy progression.
2. We explore the application of SSL to NODEs for disease progression analysis.
3. We bridge the gap between SSL and time-aware models by introducing novel temporal augmentations.

To the best of our knowledge, this is the first attempt to leverage SSL to learn improved weights specifically for NODE-based models in the context of disease progression analysis using longitudinal medical images.

2 Methods

This section describes the methodology developed for disease progression analysis. We start by briefly describing two important frameworks commonly used for pre-training. Our different hypotheses will be used on top of these frameworks. To harmonize the notation, let f , g , u be encoder, projection head, and time-aware head, respectively. The core novelty lies in employing NODEs within the BYOL/SimCLR framework. With the proposition of two straightforward augmentation schemes to pre-train a time-aware head using the most popular SSL paradigms: SimCLR [5] and BYOL [9]. Let $\mathcal{V}$ be the set of consecutive patient-specific image pairs from the collection of all CFP images. $\mathcal{V}$ contains all $(\mathbf{x}_{t_i}, \mathbf{x}_{t_{i+1}})$ that are from the same patient where $\mathbf{x}_{t_i}$ is scanned before $\mathbf{x}_{t_{i+1}}$ and $i \in [0, m - 2]$ with m the number of visits for a given eye. For a given patient, the scans are irregularly sampled due to clinical constraints ($\Delta_{t_i} \neq \Delta_{t_{i+1}}$) with $\Delta_{t_i} = t_{i+1} - t_i$ the time difference between two consecutive exams and we denote S_{t_i} the severity grade at time t_i . Time-aware models are deep neural networks that take as inputs both time and embedding representation. We note $h'_{t_{i+1}} = u(h(t_i), t_i, t_{i+1}, \theta)$ where u denotes a given deep neural network, and θ its trainable parameters and h_{t_i} some latent vector at time t_i and $h'_{t_{i+1}}$ the latent representation produced by u at time t_{i+1} . In our experiments, Neural ODE (NODE) based methods will be followed.

Neural Ordinary Differential Equations (NODEs) approximate unknown ordinary differential equations by a neural network [4] that parameterizes the

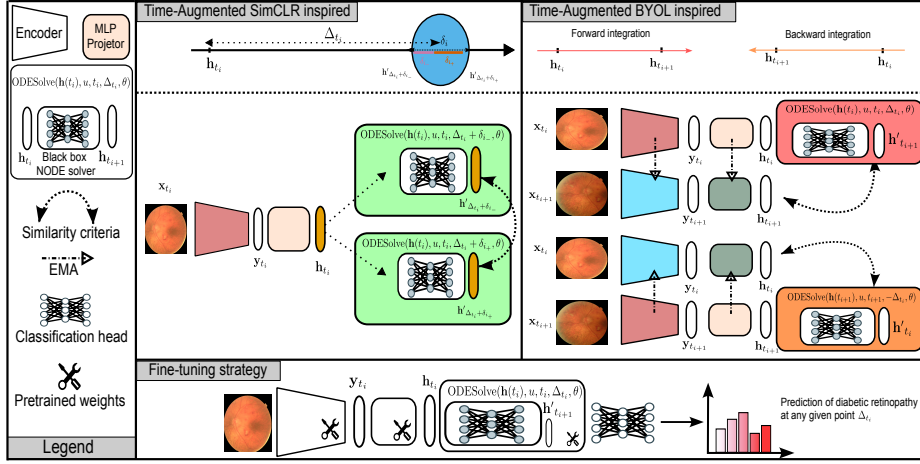


Fig. 2: Proposed augmentation techniques that we employed to mimic the popular SSL paradigms (SimCLR [5] on the left, BYOL [9] on the right) with neural ODE.

continuous dynamics of hidden units $h \in \mathbb{R}^n$ over time with $t \in \mathbb{R}$. NODEs are able to model the instantaneous rate of change of h with respect to t using a neural network u with parameters θ .

$$\lim_{\epsilon \rightarrow 0} \frac{h_{t+\epsilon} - h_t}{\epsilon} = \frac{dh}{dt} = u(t, h, \theta) \quad (1)$$

The analytical solution of Eq.1 is given by:

$$h_{t_{i+1}} = h_{t_i} + \int_{t_i}^{t_{i+1}} u(t, h, \theta) dt = \text{ODESolve}(h(t_i), u, t_i, t_{i+1}, \theta) \quad (2)$$

where $[t_i, t_{i+1}]$ represents the time horizon for solving the ODE, u being a neural network, and θ is the trainable parameters of u . By using a black-box ODE solver introduced in [4], we can solve the initial value problem (IVP) and calculate the hidden state at any desired time using Eq.2. We can differentiate the solutions of the ODE solver with respect to the parameters θ , the initial state h_{t_i} at initial time t_i , and the solution at time t . This can be achieved by using the adjoint sensitivity method [4]. Through the latent representation of a given image, we define an IVP that aims to solve the ODE from t_i to a terminal time t_{i+1} :

$$\dot{h}(t) = u(h(t), t, \theta), \text{ with the initial value } h(t_i) = h_{t_i} \quad (3)$$

Following the same way, we can define a final value problem (FVP). This is possible due to the fact that neural ODE are invertible and learn homeomorphic transformation [4]. Through the latent representation of a given image, we define a FVP that aims to solve the ODE from t_{i+1} to a reverse time t_i :

$$\dot{h}(t) = u(h(t), -t, \theta), \text{ with the final value } h(t_{i+1}) = h_{t_{i+1}} \quad (4)$$

2.1 SimCLR-based approach: *disease progression alignment*

During training, we sample pairs from $\mathcal{V}$ and encode the first element of the pair $\mathbf{y}_{t_i} = f(\mathbf{x}_{t_i})$ using the encoder f . Then, using the projector g , we obtain the representation of $h_{t_{i+1}} = g(\mathbf{y}_{t_i})$. To transpose the SimCLR (Fig.1.a) learning paradigm by defining close time in the patient disease trajectory such that the transformation plays the role of augmentation for the time-aware model: $h'_{\Delta_{t_i} + \delta_i} = \text{ODESolve}(h(t_i), u, t_i, \Delta_{t_i} + \delta_i, \theta)$. The value of δ_i is chosen to align with the known progression of the disease. For a given pair of image $(\mathbf{x}_{t_i}, \mathbf{x}_{t_{i+1}})$, we first calculate the rate of variation such that $r_i = \frac{S_{t_{i+1}} - S_{t_i}}{t_{i+1} - t_i}$. By using r_i , we define an interval δ_i where the disease rate of progression is the smallest on a period of one month, with $r_i \cdot \frac{12}{365}$. Using δ_i we now define δ_{i+} and δ_{i-} two sub-intermediate time, which will play the role of time augmentation for the neural NODE when solving the ODE. We choose $\delta_{i+} \sim \text{Unif}(0, \delta_i)$ (Unif denote the uniform distribution) and $\delta_{i-} = \delta_i - \delta_{i+}$ so that the maximum time distance between the two sampled times is at most δ_i . A schematic representation of the operations is illustrated in Fig.2 (left). We follow the standard SimCLR loss where τ and *sim* stand for the temperature and the cosine similarity, respectively. We can rewrite the objective loss for one pair in a batch:

$$h'_{\Delta_{t_i} + \delta_{i+}} = \text{ODESolve}(h(t_i), u, t_i, \Delta_{t_i} + \delta_{i+}, \theta), \quad (5)$$

$$h'_{\Delta_{t_i} - \delta_{i-}} = \text{ODESolve}(h(t_i), u, t_i, \Delta_{t_i} - \delta_{i-}, \theta), \quad (6)$$

$$\mathcal{L} = -\log \frac{\exp(\text{sim}(h'_{\Delta_{t_i} + \delta_{i+}}, h'_{\Delta_{t_i} + \delta_{i-}})/\tau)}{\sum_{k=1}^{2N} \mathbb{1}_{[k \neq i]} \exp(\text{sim}(h'_{\Delta_{t_i} + \delta_{i+}}, h'_{\Delta_{t_i} + \delta_{i+}}^k)/\tau)} \quad (7)$$

2.2 BYOL-based approach: *temporal evolution and temporal consistency*

Similar to BYOL, we leverage two neural networks, denoted to as online (parameterized by μ) and target (parameterized by ξ), that interact and guide each other's learning. The target network shares the same architecture as the online network but uses exponentially moving average weights $\xi \leftarrow \alpha \xi + (1 - \alpha) \mu$ with α controlling the moving average strength. During training, a pair of data $(\mathbf{x}_{t_i}, \mathbf{x}_{t_{i+1}})$ are sampled from the data distribution $\mathcal{V}$. These are encoded by the online and target networks' shared encoder f into representations $\mathbf{y}_{t_i}$ and $\mathbf{y}_{t_{i+1}}$, respectively. Subsequently, projectors g_μ and g_ξ project them into latent representations h_{t_i} and $h_{t_{i+1}}$. The online network leverages the initial latent code h_{t_i} to predict the future latent code $h'_{t_{i+1}}$ using the IVP of the neural ODE (Eq.3). Both predicted and target representations are L2-normalized. The BYOL loss

$\mathcal{L}_\mu^{\text{BYOL NODE}}$ is the mean squared error between the normalized prediction $h'_{t_{i+1}}$ and $h_{t_{i+1}}$. A complementary symmetric loss $\tilde{\mathcal{L}}_\mu^{\text{BYOL NODE}}$ is formulated using the FVP of the neural ODE, predicting h'_{t_i} from $h_{t_{i+1}}$ (Eq.4). Fig.2 (right) visually depicts this process. The final loss function combines both BYOL losses:

$$\mathcal{L} = \mathcal{L}_\mu^{\text{BYOL NODE}} + \tilde{\mathcal{L}}_\mu^{\text{BYOL NODE}}$$

3 Experiments and results

3.1 Dataset and implementation details.

We trained and evaluated our models on OPHDIAT [16], a vast dataset of fundus photographs (CFPs) collected from the Ophthalmology Diabetes Telemedicine network. This dataset encompasses examinations from over 101,000 patients between 2004 and 2017. Among the 763,848 interpreted CFP images, nearly 673,000 received diabetic retinopathy (DR) severity grade, while the remainder were ungradable. The patient’s age ranges from 9 to 91 years. We trained the networks for 400 epochs using AdamW optimizer, OneCycleLR scheduler, learning rate of 1e-3, weight decay of 1e-4, and a batch size of 128 on a single Nvidia A6000 GPU with PyTorch. For NODEs, we utilized the Torchdiffeq library [3] for solving ODEs, back-propagation, and the adjoint method. Our NODE architecture comprised dense layers with tanh activation and dopri5 solver. We monitored validation performance and saved the best model. We used ResNet50 [10] as a backbone and a two-layer MLP projector. By leveraging consecutive image pairs, we achieved diversity without additional augmentations, fostering a robust time-aware predictor for longitudinal tasks. We evaluated the effectiveness of our framework in enhancing time-continuous models (TMCs) by comparing our approach with popular TMCs like NODE [4], NODE-RNN [22], NODE-GRU [2], and NODE-LSTM [15]. We evaluated our pre-trained model by fine-tuning the time-aware and feature extractor component for 100 epochs on two tasks: predicting DR progression at 1-3 years (task 1), and next visit timing (task 2). The area under the receiver operating characteristic curve (AUC) with three binary tasks: predicting at least non-proliferative diabetic retinopathy (NPDR), (**AUC1**), at least moderate NPDR (**AUC2**) and finally at least severe NPDR (**AUC3**) and the quadratic-weighted Kappa were used metrics. When δ_i is null a default disease rate of progression equivalent of 3 month is given. Further training dataset details and hyper-parameters are given in supplementary materials.

3.2 Results

Evaluation of time aware model on a fixed time interval (Task 1): Prediction of diabetic retinopathy progression for 1,2,3 years. Tab.1 compares methods for predicting diabetic retinopathy (DR) progression at 1, 2, and 3 years using AUC. A baseline model (ResNet50) achieved moderate performance (AUC 0.57-0.66). Adding a NODE layer generally decreased performance, and further

incorporating recurrent layers (RNN, LSTM, GRU) yielded mixed results. However, pre-training the NODE layer with either disease progression alignment or temporal consistency and evolution significantly improved performance (AUC > 0.6 for all predictions), with disease progression alignment achieving the highest AUCs. These results suggest that pre-training the NODE layer with task-specific information is crucial for accurate DR progression prediction.

Method	Weights	Prediction 1 year			Prediction 2 year			Prediction 3 year		
		AUC1	AUC2	AUC3	AUC1	AUC2	AUC3	AUC1	AUC2	AUC3
ResNet50	-	0.609	0.658	0.663	0.5747	0.620	0.624	0.603	0.649	0.636
ResNet50+NODE	-	0.533	0.624	0.568	0.538	0.572	0.599	0.501	0.503	0.512
ResNet50+NODE-RNN	-	0.498	0.465	0.549	0.606	0.655	0.653	0.499	0.467	0.505
ResNet50+NODE-LSTM	-	0.524	0.619	0.564	0.504	0.459	0.575	0.542	0.616	0.615
ResNet50+NODE-GRU	-	0.510	0.607	0.603	0.501	0.503	0.513	0.504	0.574	0.543
ResNet50+NODE	SimCLR-based	0.645	0.694	0.674	0.606	0.651	0.664	0.615	0.679	0.678
ResNet50+NODE-RNN	SimCLR-based	0.627	0.675	0.677	0.627	0.675	0.677	0.592	0.644	0.647
ResNet50+NODE-LSTM	SimCLR-based	0.617	0.669	0.672	0.622	0.669	0.665	0.603	0.653	0.630
ResNet50+NODE-GRU	SimCLR-based	0.615	0.669	0.679	0.614	0.663	0.667	0.581	0.634	0.657
ResNet50+NODE	BYOL-based	0.654	0.708	0.669	0.654	0.708	0.669	0.640	0.685	0.673
ResNet50+NODE-RNN	BYOL-based	0.650	0.696	0.680	0.650	0.696	0.680	0.608	0.659	0.661
ResNet50+NODE-LSTM	BYOL-based	0.622	0.669	0.665	0.640	0.679	0.625	0.609	0.654	0.667
ResNet50+NODE-GRU	BYOL-based	0.654	0.697	0.698	0.605	0.667	0.680	0.612	0.654	0.673

Table 1: Comparison of method and pre-trained weights in terms of AUC for the prediction of diabetic retinopathy at 1, 2, and 3 years. Best results are in Bold.

Evaluation for DR progression of time aware model on varying time interval (Task 2): Results shown in Tab.2 demonstrate the effectiveness of incorporating recurrent architectures (NODE-RNN, NODE-LSTM, NODE-GRU) and inspired pre-training strategies (SimCLR, BYOL) for predicting diabetic retinopathy progression. Compared to the baseline ResNet50+NODE model, all recurrent architectures achieved statistically significant improvements ($p < 0.05$, compared to the baseline without pretraining, DeLong test) in all three AUC metrics (AUC1, AUC2, AUC3) and in Kappa across all pre-training settings. Notably, BYOL pre-training consistently yielded the highest performance across all NODE-recurrent models, with the NODE-GRU variant achieving the best overall results (Kappa: 0.472, AUC1: 0.773, AUC2: 0.836, AUC3: 0.848).

The findings reported in Tab.1, 2 and 3 suggest that NODE-based recurrent architectures effectively capture temporal information in retinal disease progression. Pre-training with self-supervised learning methods inspired by SimCLR or BYOL further enhances model performance. One significant advantage of the proposed paradigm lies in its ability to train models with modular time steps. This flexibility contrasts with current methods, which often require imposing strict criteria on datasets to fit a specific prediction task. Such constraints can limit the model’s ability to handle variable durations and sample numbers within the training data. While the tasks in our experiments share similar characteristics, we observed a greater performance gain for the task presented in Tab.2 compared to Tab.1. We hypothesize that this difference stems from the varying number of training samples between the two tasks. The results in Tab.2 showcase

Method	Weights	Task 2			
		Kappa	AUC1	AUC2	AUC3
ResNet50+NODE	-	0.007	0.516	0.568	0.514
ResNet50+NODE-RNN	-	0.042	0.561	0.623	0.628
ResNet50+NODE-LSTM	-	0.127	0.571	0.618	0.643
ResNet50+NODE-GRU	-	0.105	0.571	0.639	0.672
ResNet50+NODE	SimCLR-based	0.090	0.560	0.631	0.680
ResNet50+NODE-RNN	SimCLR-based	0.390	0.764	0.802	0.815
ResNet50+NODE-LSTM	SimCLR-based	0.459	0.763	0.825	0.800
ResNet50+NODE-GRU	SimCLR-based	0.472	0.769	0.836	0.820
ResNet50+NODE	BYOL-based	0.177	0.589	0.646	0.714
ResNet50+NODE-RNN	BYOL-based	0.421	0.760	0.831	0.847
ResNet50+NODE-LSTM	BYOL-based	0.511	0.764	0.830	0.841
ResNet50+NODE-GRU	BYOL-based	0.459	0.773	0.835	0.848

Table 2: Comparaison of method and pre-trained weights in term of AUC for the prediction of diabetic retinopathy for time varying interval. Best results are in Bold.

Method	TC	Kappa	AUC1	AUC2	AUC3
ResNet50+NODE (BYOL-based)	No	0.096	0.572	0.633	0.641
ResNet50+NODE (BYOL-based)	Yes	0.177	0.589	0.646	0.714
δ_i					
ResNet50+NODE (SimCLR-based) 1	1	0.033	0.548	0.577	0.573
ResNet50+NODE (SimCLR-based) 2	2	0.049	0.546	0.607	0.628
ResNet50+NODE (SimCLR-based) 3	3	0.090	0.560	0.631	0.680

Table 3: Comparaison of the importance of configuration for BYOL/SimCLR-based pre-training technique on Task 2. Best results are in Bold.

the ability of our framework to fully leverage the strengths of continuous-time models, allowing us to train models with samples featuring diverse progression times (as depicted in Fig.1.c) compared to the limitations of classical paradigms.

3.3 Ablation study

Impact of value of the chosen δ_i and inverse constraint. To examine the impact of temporal enhancement on model performance, we conducted an ablation study focusing on the values of δ_i : (1) fixed: a constant value was assigned to δ_i for all data points; (2) variable (unaligned): δ_i was varied randomly for each data point, but not aligned with the progression of the disease; (3) variable (aligned): δ_i was varied based on the disease stage of each data point, aligned with the temporal dynamics of the disease as explained in Sec.2.2. Tab.3 highlights the crucial role of both the inverse constraint in the BYOL-inspired approach, and the alignment with disease progression and diverse values of δ_i in the SimCLR-inspired training strategy.

Convergence challenges and benefits of pre-training. During experiments, when training our time-aware classifier using the "dopri5" solver, we encountered convergence issues, specifically when NODE was not pre-trained. These issues manifested as unstable learning curves and gradient instability during training. Notably, pre-trained models exhibited significantly faster convergence and consistently stable gradients throughout the training process. This finding suggests that pre-training can effectively mitigate a prominent challenge associated with NODE training stability [7], aligning with previous observations reported in [25, 24, 8] about NODE pre-training and embedding capacities [13].

4 Conclusion

This study pre-trained NODEs, a time-aware model, to enhance representation learning for longitudinal medical image analysis, significantly improving performance in predicting DR progression. Our novel time augmentation paradigm

bridges the gap between time-aware methods and popular SSL approaches like BYOL/SimCLR, with potential generalizability to other techniques. This paves the way for further exploration of re-training TAM, especially NODEs, for improved temporal modeling in medical imaging. Future research should optimize hyperparameters and extend our framework beyond classification to various longitudinal tasks.

Acknowledgments. This work was conducted within the framework of the ANR RHU project Evired (ANR-18-RHUS-0008).

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Bora, A., Balasubramanian, S., Babenko, B., Virmani, S., Venugopalan, S., Mitani, A., de Oliveira Marinho, G., Cuadros, J., Ruamviboonsuk, P., Corrado, G.S., Peng, L., Webster, D.R., Varadarajan, A.V., Hammel, N., Liu, Y., Bavishi, P.: Predicting the risk of developing diabetic retinopathy using deep learning. *The Lancet Digital Health* **3**, e10–e19 (1 2021). [https://doi.org/10.1016/S2589-7500\(20\)30250-8](https://doi.org/10.1016/S2589-7500(20)30250-8)
2. Brouwer, E.D., Simm, J., Arany, A., Moreau, Y.: Gru-ode-bayes: Continuous modeling of sporadically-observed time series. In: *Neural Information Processing Systems* (2019), <https://api.semanticscholar.org/CorpusID:168169891>
3. Chen, R.T.Q.: torchdiffEq (2018), <https://github.com/rtqichen/torchdiffEq>
4. Chen, R.T.Q., Rubanova, Y., Bettencourt, J., Duvenaud, D.: Neural ordinary differential equations (2018). <https://doi.org/10.48550/ARXIV.1806.07366>, <https://arxiv.org/abs/1806.07366>
5. Chen, T., Kornblith, S., Norouzi, M., Hinton, G.: A simple framework for contrastive learning of visual representations (2020). <https://doi.org/10.48550/ARXIV.2002.05709>, <https://arxiv.org/abs/2002.05709>
6. Couronné, R., Vernhet, P., Durrleman, S.: Longitudinal self-supervision to disentangle inter-patient variability from disease progression (09 2021). https://doi.org/10.1007/978-3-030-87196-3_22
7. Ghosh, A., Behl, H.S., Dupont, E., Torr, P.H.S., Nambodiri, V.: Steer: Simple temporal regularization for neural odes (2020). <https://doi.org/10.48550/ARXIV.2006.10711>, <https://arxiv.org/abs/2006.10711>
8. Gong, S., Meng, Q., Wang, Y., Wu, L., Chen, W., Ma, Z.M., Liu, T.Y.: Incorporating node with pre-trained neural differential operator for learning dynamics (2021)
9. Grill, J.B., Strub, F., Altché, F., Tallec, C., Richemond, P.H., Buchatskaya, E., Doersch, C., Pires, B.A., Guo, Z.D., Azar, M.G., Piot, B., Kavukcuoglu, K., Munos, R., Valko, M.: Bootstrap your own latent: A new approach to self-supervised learning (2020)
10. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition (2015)
11. Holland, R., Leingang, O., Bogunovi'c, H., Riedl, S., Fritsche, L.G., Prevost, T., Scholl, H.P.N., Schmidt-Erfurth, U.M., Sivaprasad, S., Lotery, A.J., Rueckert, D., Menten, M.J.: Metadata-enhanced contrastive learning from retinal optical coherence tomography images. *ArXiv abs/2208.02529* (2022), <https://api.semanticscholar.org/CorpusID:251320568>

12. Kim, H., Sabuncu, M.R.: Learning to compare longitudinal images (2023)
13. Kuehn, C., Kuntz, S.V.: Embedding capabilities of neural odes (2023)
14. Lachinov, D., Chakravarty, A., Grechenig, C., Schmidt-Erfurth, U., Bogunović, H.: Learning spatio-temporal model of disease progression with neuralodes from longitudinal volumetric data. *IEEE Transactions on Medical Imaging* pp. 1–1 (2023). <https://doi.org/10.1109/TMI.2023.3330576>
15. Lechner, M., Hasani, R.: Learning long-term dependencies in irregularly-sampled time series (2020)
16. Massin, P., Chabouis, A., Erginay, A., Viens-Bitker, C., Lecleire-Collet, A., Meas, T., Guillausseau, P., Choupot, G., André, B., Denormandie, P.: Ophdiat©: A telemedical network screening system for diabetic retinopathy in the ile-de-france. *Diabetes & metabolism* **34**, 227–34 (07 2008). <https://doi.org/10.1016/j.diabet.2007.12.006>
17. Ouyang, J., Zhao, Q., Adeli, E., Peng, W., Zaharchuk, G., Pohl, K.M.: Lsor: Longitudinally-consistent self-organized representation learning (2023)
18. Qian, Z., Zame, W.R., Fleuren, L.M., Elbers, P., van der Schaar, M.: Integrating expert odes into neural odes: Pharmacology and disease progression. In: *Advances in Neural Information Processing Systems*. vol. 34 (2021), <https://papers.neurips.cc/paper/2021/file/5ea1649a31336092c05438df996a3e59-Paper.pdf>
19. Ren, M., Dey, N., Styner, M.A., Botteron, K., Gerig, G.: Local spatiotemporal representation learning for longitudinally-consistent neuroimage analysis (2022)
20. Ren, M., Dey, N., Styner, M.A., Botteron, K., Gerig, G.: Local spatiotemporal representation learning for longitudinally-consistent neuroimage analysis (2022). <https://doi.org/10.48550/ARXIV.2206.04281>, <https://arxiv.org/abs/2206.04281>
21. Rivail, A., Schmidt-Erfurth, U., Vogel, W.D., Waldstein, S.M., Riedl, S., Grechenig, C., Wu, Z., Bogunovic, H.: Modeling disease progression in retinal octs with longitudinal self-supervised learning. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* **11843 LNCS**, 44–52 (2019). https://doi.org/10.1007/978-3-030-32281-6_5
22. Rubanova, Y., Chen, R.T.Q., Duvenaud, D.: Latent odes for irregularly-sampled time series (2019). <https://doi.org/10.48550/ARXIV.1907.03907>, <https://arxiv.org/abs/1907.03907>
23. Vernhet, P., Durrleman, S.: Longitudinal self-supervision to disentangle inter-patient variability pp. 231–241 (2021). <https://doi.org/10.1007/978-3-030-87196-3>
24. Zeghlache, R., Conze, P.H., Daho, M.E.H., Li, Y., Le Boité, H., Tadayoni, R., Massin, P., Cochener, B., Brahim, I., Quéllec, G., et al.: Lmt: Longitudinal mixing training, a framework to predict disease progression from a single image. In: *International Workshop on Machine Learning in Medical Imaging*. pp. 22–32. Springer (2023)
25. Zeghlache, R., Conze, P.H., Daho, M.E.H., Li, Y., Le Boité, H., Tadayoni, R., Massin, P., Cochener, B., Brahim, I., Quéllec, G., Lamard, M.: Longitudinal self-supervised learning using neural ordinary differential equation. In: *International Workshop on Predictive Intelligence in Medicine*. Vancouver (Canada), Canada (Oct 2023), <https://imt-atlantique.hal.science/hal-04171357>
26. Zhang, R., Isola, P., Efros, A.A.: Colorful image colorization. In: *Computer Vision—ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part III* 14. pp. 649–666. Springer (2016)
27. Zhao, Q., Liu, Z., Adeli, E., Pohl, K.M.: Longitudinal self-supervised learning. *Medical Image Analysis* **71** (2021). <https://doi.org/10.1016/j.media.2021.102051>