Progression Models for Imaging Data with Longitudinal Variational Auto Encoders

Benoît Sauty[(✉)] and Stanley Durrleman

Inria, Sorbonne Université, Institut du Cerveau - Paris Brain Institute - ICM,
Inserm, CNRS, AP-HP, Paris, France
`benoit.sautydechalan@icm.institute.org`

Abstract. Disease progression models are crucial to understanding degenerative diseases. Mixed-effects models have been consistently used to model clinical assessments or biomarkers extracted from medical images, allowing missing data imputation and prediction at any time-point. However, such progression models have seldom been used for entire medical images. In this work, a Variational Auto Encoder is coupled with a temporal linear mixed-effect model to learn a latent representation of the data such that individual trajectories follow straight lines over time and are characterised by a few interpretable parameters. A Monte Carlo estimator is devised to iteratively optimize the networks and the statistical model. We apply this method on a synthetic data set to illustrate the disentanglement between time dependant changes and inter-subjects variability, as well as the predictive capabilities of the method. We then apply it to 3D MRI and FDG-PET data from the Alzheimer's Disease Neuroimaging Initiative (ADNI) to recover well documented patterns of structural and metabolic alterations of the brain.

Keywords: Variational auto encoders · Mixed-effects models · Disease progression models · Alzheimer's Disease

1 Introduction

Estimating progression models from the analysis of time dependent data is a challenging task that helps to uncover latent dynamics. For the study of neurodegenerative diseases, longitudinal databases have been assembled where a set of biomarkers (medical images, cognitive scores and covariates) are gathered for individuals across time. Understanding their temporal evolution is of crucial importance for early diagnosis and drug trials design, especially the imaging biomarkers that can reveal a silent prodromal phase.

In this context, several approaches have been proposed for the progression of scalar measurements such as clinical scores or volumes of brain structures [12,18,29] or series of measurements across brain regions forming a network [4,21]. These approaches require the prior segmentation and extraction of the

measurements from the images. Providing progression models for high dimensional structured data without prior processing is still a challenging task. The difficulty is to provide a low dimensional representation of the data, where each patient's trajectory admits a continuous parametrization over time. It should allow sampling at any time point, be resilient to irregularly spaced instances and disentangle temporal alterations from the inter-patients variability. 0

1.1 Related Work

When dealing with high dimensional data, it is often assumed that the data can be encoded into a low dimensional manifold where the distribution of the data is simple. Deep Generative Neural Networks such as Variational Auto Encoders (VAE) [19] allow finding such embeddings. Several approaches have explored longitudinal modeling for images within the context of dimensionality reduction.

Recurrent Neural Networks (RNN) provide a straightforward way to extract information from sequential data. Convolutional networks with a recurrent structure have been used for diagnosis prediction using MRI [11] or PET [23] scans in Alzheimer's Disease (AD). The main caveat of these approaches is that the recurrent structure is highly sensible to the temporal spacing between instances which is troublesome in the context of disease modeling, where visits are often missing of irregularly spaced.

Mixed-effects models provide an explicit description of the progression of each patient, allowing to sample at any timepoint. Through a time reparametrization, all patients are aligned on a common pathological timeline, and individual trajectories are parametrised as small variations (random effects) around a reference trajectory (fixed effects) that can be seen as the average scenario. Now considered a standard tool in longitudinal modeling [21,28,29], mixed-effects models have yet been scarcely used for images within the context of dimension reduction. In [24], a RNN outputs the parameters of a mixed-effect model that describes patients' trajectories as straight lines in the latent space of a VAE across time.

Self supervised methods have proposed to alleviate the need for labels, in our case the age of the patients at each visit. In [9], the encoder of a VAE learns a latent time variable and a latent spatial variable to disentangle the temporal progression from the patient's intrinsic characteristics. Similarly, in [30], the encoder is penalized with a cosine loss that imposes one direction in the latent space that corresponds to an equivalent of time. Both these methods allow the model to learn a temporal progression that does not rely on the clinical age of the patients, offering potential for unlabeled data, at the cost of interpretability of the abstract timeline and the ability to sample at any given timepoint.

Longitudinal VAEs architectures have been proposed in order to endow the latent space with a temporal structure. Namely, Gaussian Process VAEs (GPVAE) [7] introduced a more general prior for the posterior distribution in the latent space, in the form of a Gaussian Process (GP) that depends on the age of the patients [13] as well as a series of covariates [2,26]. This approach

poses challenges as to the choice of parametrization for the Gaussian Process, and does not provide an expected trajectory for each patient.

Diffeomorphic methods provide progression models for images. The main approaches are based off of the geodesic regression framework [3,25] and allow learning a deformation map that models the effect of time on the images for a given subject. While providing high resolution predictions, these methods show limited predictive abilities further in time when compared to mixed-effects models, that aggregate information from all the subjects at different stages of the disease [6].

1.2 Contributions

In this context, we propose to endow the latent space of a VAE with a linear mixed-effect longitudinal model. While in [24], the networks predict the random effects from visits grouped by patients, we propose to enrich a regular VAE that maps each individual visits to a latent representation, with an additional longitudinal latent model that describes the progression of said representations over time. A novel Monte Carlo Markov Chain (MCMC) procedure to jointly estimate the VAE and the structure of its representation manifold is proposed. To sum up the contributions, we:

1. use the **entire 3D scan** without segmentation or parcellation to study relations across brain regions in an unsupervised manner,
2. proceed to **dimension reduction** using a convolutional VAE with the added constraint that latent representations must comply with the structure of a generative statistical model of the trajectories,
3. provide a **progression model** that disentangles temporal changes from changes due to inter-patients variability, and allows sampling patients' trajectories at any timepoint, to infer missing data or predict future progression,
4. demonstrate this method on a synthetic data set and on both **MRI** and **PET scans** from the Alzheimer's Disease Neuroimaging Initiative (ADNI), recovering known patterns in normal or pathological brain aging.

2 Methodology

2.1 Representation Learning with VAEs

Auto Encoders are a standard tool for non-linear dimensionality reduction, comprised of an *encoder* network q_ϕ that maps high dimensional data $x \in \mathcal{X}$ to $z \in \mathcal{Z}$, in a smaller space refered to as the *latent space*, and a *decoder* network $p_\theta : z \in \mathcal{Z} \mapsto \hat{x} \in \mathcal{X}$. VAEs [19] offer a more regularized way to populate the latent space. Both encoder and decoder networks output variational distributions $q_\phi(z|x)$ and $p_\theta(x|z)$, chosen to be multivariate Gaussian distributions. Adding a prior $q(z)$, usually the unit Gaussian distribution $\mathcal{N}(0, \mathrm{I})$, on $\mathcal{Z}$ allows to derive a tractable Evidence Lower BOund for the log-likelihood $ELBO = \mathcal{L}_{recon} + \beta \mathcal{L}_{KL}$ where $\mathcal{L}_{recon}$ is the ℓ_2 reconstruction error, $\mathcal{L}_{KL}$ is the Kullback-Leibler (KL) divergence between the approximate posterior and the prior on the latent space and β balances reconstruction error and latent space regularity [16].

2.2   Longitudinal Statistical Model

In this section, we propose a temporal latent variables model that encodes disease progression in the low-dimensional space $\mathcal{Z}$. Given a family of observations from N patients $\{x_{i,j}\}_{1 \leqslant i \leqslant N}$, each observed at ages $t_{i,j}$ for $1 \leqslant j \leqslant n_i$ visits, and their latent representations $\{z_{i,j}\}$, we define a statistical generative model with

$$z_{i,j} = p_0 + \left[e^{\xi_i}(t_{i,j} - \tau_i) \right] v_0 + w_i + \varepsilon_{i,j}$$

where e^{ξ_i} and τ_i, respectively the *acceleration factor* and the *onset age* of patient i, allow an affine time warp aligning all patients on a common pathological timeline, and $w_i \in \mathcal{Z}$ is the *space shift* that encodes inter-subjects variability, such as morphological variations across regions that are independent from the progression. These parameters position the individual trajectory with respect to the typical progression that is estimated at the population level. These three parameters form the **random effects** of the model ψ_r. Vectors w_i and v_0 need to be orthogonal in order to uniquely identify temporal and spatial variability. We choose the Gaussian priors for the noise $\varepsilon_{i,j} \sim \mathcal{N}(0, \sigma_\varepsilon^2)$ and random effects $\tau_i \sim \mathcal{N}(t_0, \sigma_\tau^2)$, $\xi_i \sim \mathcal{N}(0, \sigma_\xi^2)$ and $w_i \sim \mathcal{N}(0, \mathrm{I})$. The parameters $p_0 \in \mathcal{Z}$, $v_0 \in \mathcal{Z}$, $t_0 \in \mathbb{R}$ are respectively a reference *position*, *velocity* and *time* and describe the average trajectory. Together with the variances $\sigma_\varepsilon, \sigma_\tau, \sigma_\xi$, they form the **fixed-effects** of the model ψ_f. We note $\psi = (\psi_r, \psi_f)$.

2.3   Longitudinal VAE

We combine dimension reduction using a regular β-VAE and the aforementioned statistical model to add a temporal structure to the latent space. To do so, we consider a composite loss that accounts for both the VAE loss and the goodness-of-fit of the mixed-effect model:

$$\mathcal{L} = \mathcal{L}_{recon} + \beta\mathcal{L}_{KL} + \gamma\mathcal{L}_{align} \ \text{ where } \begin{cases} \mathcal{L}_{recon} = \sum_{i,j} ||x_{i,j} - \hat{x}_{i,j}||^2 \\ \mathcal{L}_{KL} \ \ = \sum_{i,j} KL(q_\phi(z|x_{i,j})||\mathcal{N}(0, \mathrm{I})) \\ \mathcal{L}_{align} = \sum_{i,j} ||z_{i,j} - \eta_\psi^i(t_{i,j})||^2 \end{cases}$$

where $z_{i,j}$ and $\hat{x}_{i,j}$ are the modes of $q_\phi(x_{i,j})$ and $p_\theta(z_{i,j})$, and $\eta_\psi^i(t_{i,j}) = p_0 + \left[e^{\xi_i}(t_{i,j} - \tau_i) \right] v_0 + w_i$ is the expected position of the latent representation according to the longitudinal model and γ balances the penalty for not aligning latent representations with the linear model. Since the loss is invariant to rotation in $\mathcal{Z}$, we set $p_0 = 0$ and $v_0 = (1, 0, \cdots, 0)$.

Since $\mathcal{L}_{align}$ is a ℓ_2 loss in the latent space, it can be seen as the log-likelihood of a Gaussian prior $z_{i,j} \sim \mathcal{N}(\eta_\psi^i(t_{i;j}), \mathrm{I})$ in the latent space, which defines an elementary Gaussian Process, and which supports the addition of GP priors in the latent space of VAEs to model longitudinal data [13,26]. Besides, $\mathcal{X}$ can be seen as a Riemannian manifold, the metric of which is given by the pushforward of the Euclidean metric of $\mathcal{Z}$ through the decoder, such that trajectories in $\mathcal{X}$ are geodesics, in accordance with the Riemannian modeling of longitudinal data

[5,14,24,28,29]. The metric on $\mathcal{X}$ thus allows to recover non linear dynamics, as is often the case for biomarkers. Our approach thus bridges the gap between the deep learning approach to longitudinal data and a natural generalization of well studied disease progression models to images.

Network Implementation and Estimation. Both the encoder and decoder are chosen to be vanilla convolutional Neural Networks (4 layers of Convolution with stride/BatchNorm/ReLU and transposition for decoder) with a dense layer towards the latent space, as described in Fig. 1. The implementation is in PyTorch and available at https://github.com/bsauty/longitudinal-VAEs.

Algorithm 1: Monte Carlo estimation of the Longitudinal VAE

Input   : Longitudinal visits $\{(x_{i,j}, t_{i,j})\}$ and hyperparameters β and γ.
Output: Estimation of (ϕ, θ) for the VAE and ψ for the temporal model.
Init   : Initialize (ϕ, θ) as a regular β-VAE.
        Set $k = 0$ and $z^0 = q_\phi(x)$;
while *not converged* **do**
> **Simulation**
> > Draw candidates $\psi_r^c \sim p(.|\psi_f^k)$;   `// Sampling with prior` $p(.|\psi_f^k)$
> > $\forall i, j$ compute $\eta_{\psi_k}^i(t_{i,j})$;   `// Expected latent trajectories`
> > Compute likelihood ratio $\omega = \min\left(1, \frac{q(\psi_r^c | z^k, \psi_f^k)}{q(\psi_r^k | z^k, \psi_f^k)}\right)$
> > **if** $u \sim \mathcal{U}(0, 1) > \omega$ **then** $\psi_r^{k+1} \leftarrow \psi_r^c$ **else** $\psi_r^{k+1} \leftarrow \psi_r^k$
> **Approximation** Compute sufficient statistics S^k for ψ_r^k
> **Maximisation** $\psi_f^{k+1} \leftarrow \psi_f^*(S^k)$
> **VAE optimization** Run one epoch using $\mathcal{L}$ with the target latent representation $\eta_{\psi^k}^i$ for $\mathcal{L}_{align}$ and update $z^{k+1} \leftarrow q_\phi(x)$

end

During training : $z_{i,j} = \mu_{i,j} + \varepsilon \sigma_{i,j}$
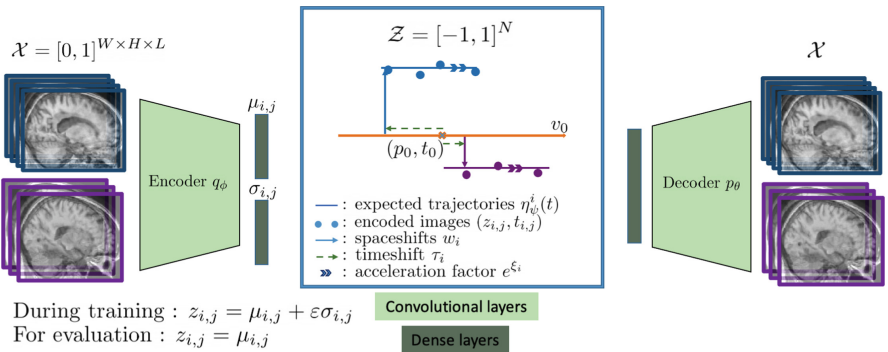For evaluation : $z_{i,j} = \mu_{i,j}$

Fig. 1. Images $\{x_{i,j}\}$ are encoded into $\mathcal{Z}$ such that the $\{z_{i,j}\}$ are close to the estimated latent trajectories. Individual trajectories (straight lines) are parametrized with w_i, τ_i and e^{ξ_i} as variations around the reference trajectory (orange arrow). (Color figure online)

The difficulty lies in the joint estimation of (θ, ϕ) and ψ, which are co-dependant since $\mathcal{L}_{align}$ depends on η_ψ^i and ψ depends on the encoded representation $z = q_\phi(x)$. The longitudinal statistical model is part of a family of geometric models that have been studied in [20,29]. Given $\{z_{i,j}\}$, we can proceed to a Maximum a Posteriori estimation of ψ with the MCMC-SAEM procedure in which the estimation step of an EM algorithm is replaced by a stochastic approximation. See [1,22] for details. Given the target trajectories $\{\eta_\psi^i\}$, the weights from both networks of the VAE is optimized through backpropagation using $\mathcal{L}$ with an optimizer with randomized batches. Both estimation schemes are iterative so we designed a Monte Carlo estimator for (ϕ, θ, ψ), presented in Algorithm 1, alternating between both schemes.

Once calibrated with a training set, we freeze the VAE parameters (ϕ, θ) and fixed effects ψ_f, and learn the individual parameters ψ_r, via gradient descent of the likelihood, to personalize the model for new subjects.

Hyperparameters are $\dim(\mathcal{Z})$, which should be small enough to allow the mixed-effect model to be interpretable but big enough to reach good reconstruction; β, which should minimize overfitting while not impairing reconstruction quality; and γ, which should also not be too big to avoid loosing contextual information in $\mathcal{Z}$. These parameters were set using grid search. Lastly, the MCMC-SAEM is computationaly inexpensive compared to backpropagation so memory footprint and runtime are similar to training a regular VAE. All training was performed with a Quadro RTX4000 8Go GPU.

3 Experiments and Results

3.1 Results on Synthetic Experiments

We first validated our approach on a synthetic data set of images of silhouettes of dimension 64×64 [10] . Over time, the silhouette raises its left arm. Different silhouettes are generated by varying the relative position of the three other limbs, all of them raising their left arm in time. The motion is modulated by varying the time stamp at which the motion starts and the pace of motion. This is done using an affine reparametrization of the time stamp $t_{i,j}$ of the silhouette with Gaussian log-acceleration factor ξ_i and onset age τ_i. This data set contains $N = 1,000$ subjects with $n = 10$ visits each, sampled at random time-points.

We choose $\dim(\mathcal{Z}) = 4$ to evaluate the ability of our model to isolate temporal changes (motion of the left arm) from the independent spatial generative factors (the position of the other 3 limbs). Results are displayed in Fig. 2. The 5-fold reconstruction mean squared errors (MSE) (times 10^{-3}) for train/test images are $7.88 \pm .22/7.93 \pm .29$, showing little over-fitting. Prediction error for missing data, when trained on half-pruned data set, is $8.1 \pm .78$ which shows great extrapolation capabilities. A thorough benchmark of six former approaches on this data set was provided in [9] displaying similar MSE to ours. Although a couple of approaches [9,30] also disentangle time from space, ours is the only one to yield the true generative factors, with the direction of progression encoding the motion of the

left arm and the 3 spatial directions orthogonal to it encoding legs spreading, legs translation, and right arm position respectively.

3.2   Results on 3D MRI and PET Scans

We then applied the method to 3D T1w MRI and FDG-PET scans from the public ADNI database (http://adni.loni.usc.edu). For MRI, we selected two cohorts: patients with a confirmed AD diagnosis at one visit at least ($N = 783$ patients for a total of $N_{tot} = 3{,}685$ images) and Cognitively Normal (CN) patients at all visits for modeling normal aging ($N = 886$ and $N_{tot} = 3{,}205$). We considered PET data for AD patients only ($N = 570$ and $N_{tot} = 1{,}463$). Images are registered using the T1-linear and PET-linear pipelines of the Clinica software [27] and resampled to $80 \times 96 \times 80$ resolution.

We set $\dim(\mathcal{Z}) = 16$ for both modalities, as it is the smallest dimension that captured the reported dynamics with satisfying resolution. For MRI, error (10^{-3}) for train/test reconstruction and imputation on half-pruned data set for the AD model are $14.15 \pm .12/15.33 \pm .23$ and $18.65 \pm .76$, again showing little over-fit and good prediction abilities. In Fig. 3, the reference trajectory for AD patients reveals the structural alterations that are typical of AD progression. The control trajectory displays alterations more in line with normal aging.

We tested differences in the mean of individual parameters between subgroups using Mann-Whitney U test within a 5-fold cross-validation. AD average onset age occured earlier for women than for men: $72.2 \pm .4$ vs. $73.7 \pm .6$ years, $p < 3.10^{-7} \pm 5.10^{-8}$). APOE-$\varepsilon 4$ mutation carriers experience also earlier onset
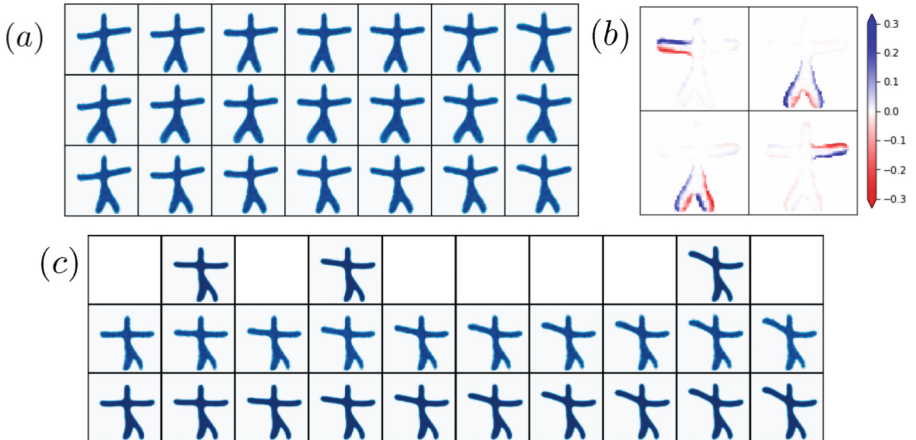

Fig. 2. Synthetic experiment. (a) the average trajectory over time (left to right) on first row, followed by its translation in the directions $w_1 = (0,1,0,0)$ and $w_2 = (0,0,1,0)$ in the latent space (second and third row). (b) the gradient of the image at $p_0 = 0$ in the 4 directions of the latent space v_0, w_1, w_2, w_3. (c) data of a test subject (first row), its reconstructed trajectory (second row) and the ground truth (third row)

than non-carriers ($71.8\pm.2$ vs $73.1\pm.4$, $p < 3.3.10^{-2}{\pm6.10^{-3}}$) and greater pace of progression ($.1\pm3.10^{-2}$ vs $-.08\pm2.10^{-2}$, $p < 1.4.10^{-4}{\pm6.10^{-3}}$). The normal aging model shows an earlier onset for men ($71.2\pm.4$ vs $73.7\pm.6$, $p < 3.10^{-10}{\pm6.10^{-11}}$). These results are in line with the current knowledge in AD progression [15,17] and normal aging [8]. For PET scans, the 5-fold train/test reconstruction MSE (10^{-2}) are $4.71 \pm .32/5.10 \pm .23$ (Fig. 4).
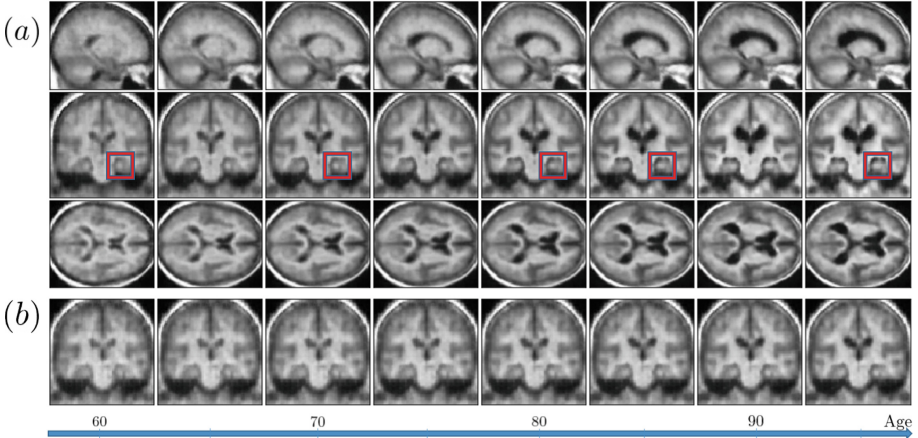

Fig. 3. (a) Sagittal, coronal and axial views of the population trajectory over pathological time (left to right) for the AD cohort. Enlargment of the ventricles and atrophy of the cortex and the hippocampus are visible. Red squares around the hippocampus are positioned at the 5 stages of the Schelten' scale used in the clinics to evaluate AD progression. (b) Coronal view of the estimated normal aging scenario, with matched reparameterized age distribution. Atrophy is also visible but to a smaller extent. As is common in atlasing methods, these images average anatomical details from different subjects to provide a population trajectory, and are thus not as sharp as true images. (Color figure online)
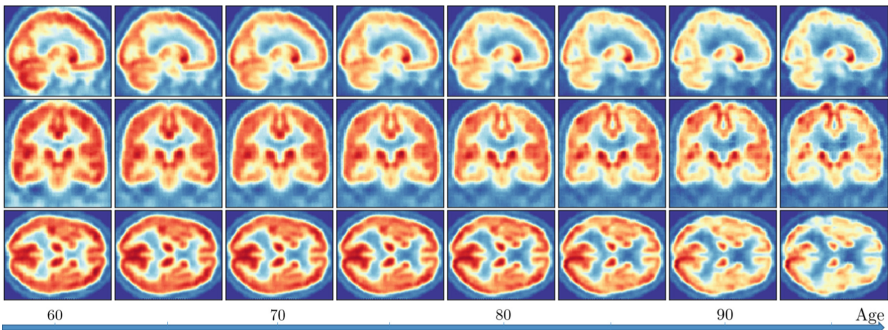

Fig. 4. Sagittal, coronal and axial views of the average trajectory for FDG-PET scans, showing decreased level of metabolism across brain regions.

4   Conclusion

We proposed a generative Variational Autoencoder architecture that maps longitudinal data to a low dimensional Euclidean space, in which a linear spatiotemporal structure is learned to accurately disentangle the effects of time and inter-patient variability, while providing interpretable individual parameters (onset age and acceleration factor). This is the first approach to provide a progression model for 3D MRI or PET scans and it relies on vanilla deep learning architectures that only require the tuning of the loss balance. We showed that it bridges the gap between former approaches to handle longitudinal images, namely GP-VAEs, and Riemannian disease progression models. The method applied to MRI and PET data retrieves known patterns of normal and pathological brain aging but without the need to extract specific biomarkers. It does not only save time but also makes the approach independent of prior choice of biomarkers.

Current work focuses on linking this progression model of brain alterations with cognitive decline, and exploring disease sub-types in the latent space.

References

1. Allassonnière, S., Kuhn, E., Trouvé, A.: Construction of Bayesian deformable models via a stochastic approximation algorithm: a convergence study. Bernoulli **16**(3), 641–678 (2010)
2. Ashman, M., So, J., Tebbutt, W., Fortuin, V., Pearce, M., Turner, R.E.: Sparse gaussian process variational autoencoders. arXiv preprint arXiv:2010.10177 (2020)
3. Banerjee, M., Chakraborty, R., Ofori, E., Okun, M.S., Viallancourt, D.E., Vemuri, B.C.: A nonlinear regression technique for manifold valued data with applications to medical image analysis. In: Proceedings of the IEEE Conference On Computer Vision and Pattern Recognition, pp. 4424–4432 (2016)
4. Bauer, S., May, C., Dionysiou, D., Stamatakos, G., Buchler, P., Reyes, M.: Multiscale modeling for image analysis of brain tumor studies. IEEE Trans. Biomed. Eng. **59**(1), 25–29 (2011)
5. Bône, A., Colliot, O., Durrleman, S.: Learning distributions of shape trajectories from longitudinal datasets: a hierarchical model on a manifold of diffeomorphisms. In: CVPR 2018 - Computer Vision and Pattern Recognition 2018, Salt Lake City, United States, June 2018. https://hal.archives-ouvertes.fr/hal-01744538
6. Bône, A., et al.: Prediction of the progression of subcortical brain structures in Alzheimer's disease from baseline. In: Cardoso, M.J., et al. (eds.) GRAIL/MFCA/MICGen -2017. LNCS, vol. 10551, pp. 101–113. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-67675-3_10
7. Casale, F.P., Dalca, A.V., Saglietti, L., Listgarten, J., Fusi, N.: Gaussian process prior variational autoencoders. arXiv preprint arXiv:1810.11738 (2018)
8. Coffey, C.E., et al.: Sex differences in brain aging: a quantitative magnetic resonance imaging study. Arch. Neurol. **55**(2), 169–179 (1998)

9. Couronné, R., Vernhet, P., Durrleman, S.: Longitudinal self-supervision to disentangle inter-patient variability from disease progression. In: de Bruijne, M., et al. (eds.) MICCAI 2021. LNCS, vol. 12902, pp. 231–241. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-87196-3_22

10. Couronné, R., Vernhet, P.: Starmen longitudinal (2021). https://doi.org/10.5281/zenodo.5081988

11. Cui, R., Liu, M., et al.: RNN-based longitudinal analysis for diagnosis of Alzheimer's disease. Comput. Med. Imaging Graph. **73**, 1–10 (2019)

12. Fonteijn, H.M., et al.: An event-based model for disease progression and its application in familial Alzheimer's disease and Huntington's disease. NeuroImage **60**(3), 1880–1889 (2012)

13. Fortuin, V., Baranchuk, D., Rätsch, G., Mandt, S.: GP-VAE: deep probabilistic time series imputation. In: International Conference On Artificial Intelligence and Statistics, pp. 1651–1661. PMLR (2020)

14. Gruffaz, S., Poulet, P.E., Maheux, E., Jedynak, B., Durrleman, S.: Learning Riemannian metric for disease progression modeling. Adv. Neural Inf. Process. Syst. **34**, 23780–23792 (2021)

15. Gurvich, C., Hoy, K., Thomas, N., Kulkarni, J.: Sex differences and the influence of sex hormones on cognition through adulthood and the aging process. Brain Sci. **8**(9), 163 (2018)

16. Higgins, I., et al.: BETA-VAE: Learning basic visual concepts with a constrained variational framework (2016)

17. Jack, C.R., et al.: Age, sex, and apoe $\varepsilon4$ effects on memory, brain structure, and β-amyloid across the adult life span. JAMA Neurol. **72**(5), 511–519 (2015)

18. Jedynak, B.M., et al.: A computational neurodegenerative disease progression score: method and results with the Alzheimer's disease neuroimaging initiative cohort. Neuroimage **63**(3), 1478–1486 (2012)

19. Kingma, D.P., Welling, M.: Auto-encoding variational bayes. arXiv preprint arXiv:1312.6114 (2013)

20. Koval, I., et al.: Statistical learning of spatiotemporal patterns from longitudinal manifold-valued networks. In: Descoteaux, M., Maier-Hein, L., Franz, A., Jannin, P., Collins, D.L., Duchesne, S. (eds.) MICCAI 2017. LNCS, vol. 10433, pp. 451–459. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-66182-7_52

21. Koval, I., et al.: AD Course Map charts Alzheimer's disease progression. Sc. Rep. **11**(1), -1-6 (2021). https://doi.org/10.1038/s41598-021-87434-1, https://hal.inria.fr/hal-01964821

22. Kuhn, E., Lavielle, M.: Coupling a stochastic approximation version of EM with an MCMC procedure. ESAIM: Probabil. Statist **8**, 115–131 (2004)

23. Liu, M., Cheng, D., Yan, W., Initiative, A.D.N., et al.: Classification of Alzheimer's disease by combination of convolutional and recurrent neural networks using FDG-pet images. Front. Neuroinform. **12**, 35 (2018)

24. Louis, M., Couronné, R., Koval, I., Charlier, B., Durrleman, S.: Riemannian geometry learning for disease progression modelling. In: Chung, A., Gee, J., Yushkevich, P., Bao, S. (eds.) Information Processing in Medical Imaging. IPMI 2019. LNCS, vol. 11492, pp. 542–553. Springer, Cham (2019)

25. Niethammer, M., Huang, Y., Vialard, F.-X.: Geodesic regression for image time-series. In: Fichtinger, G., Martel, A., Peters, T. (eds.) MICCAI 2011. LNCS, vol. 6892, pp. 655–662. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-23629-7_80

26. Ramchandran, S., Tikhonov, G., Kujanpää, K., Koskinen, M., Lähdesmäki, H.: Longitudinal variational autoencoder. In: International Conference on Artificial Intelligence and Statistics, pp. 3898–3906. PMLR (2021)
27. Routier, A., et al.: Clinica: an open-source software platform for reproducible clinical neuroscience studies. Front. Neuroinform. **15**, 689675 (2021)
28. Sauty, B., Durrleman, S.: Riemannian metric learning for progression modeling of longitudinal datasets. In: ISBI 2022-International Symposium on Biomedical Imaging (2022)
29. Schiratti, J.B., Allassonniere, S., Colliot, O., Durrleman, S.: Learning spatiotemporal trajectories from manifold-valued longitudinal data. In: Neural Information Processing Systems, vol. 28, Advances in Neural Information Processing Systems, Montréal, Canada, December 2015. https://hal.archives-ouvertes.fr/hal-01163373
30. Zhao, Q., Liu, Z., Adeli, E., Pohl, K.M.: Longitudinal self-supervised learning. Med. Image Anal. **71**, 102051 (2021)