Heteroscedastic Temporal VAE for Irregular Time Series

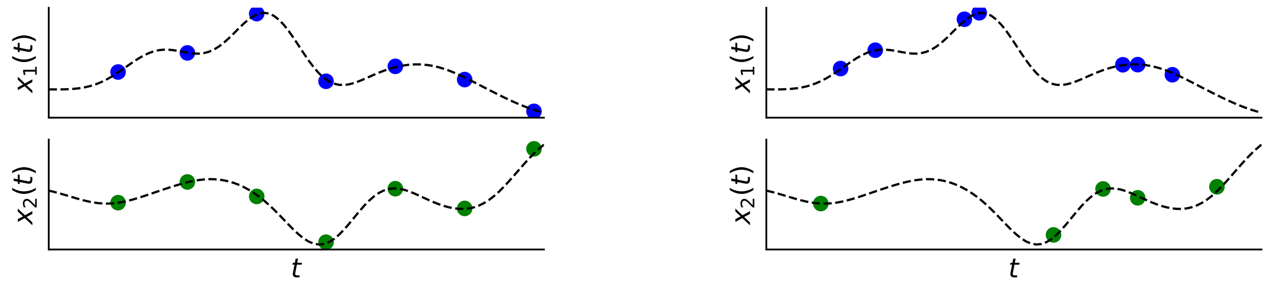
UMASS AMHERST

Satya Narayan Shukla and Benjamin M. Marlin

{snshukla,marlin}@cs.umass.edu

Introduction

Problem: Probabilistic interpolation of irregularly sampled time series.

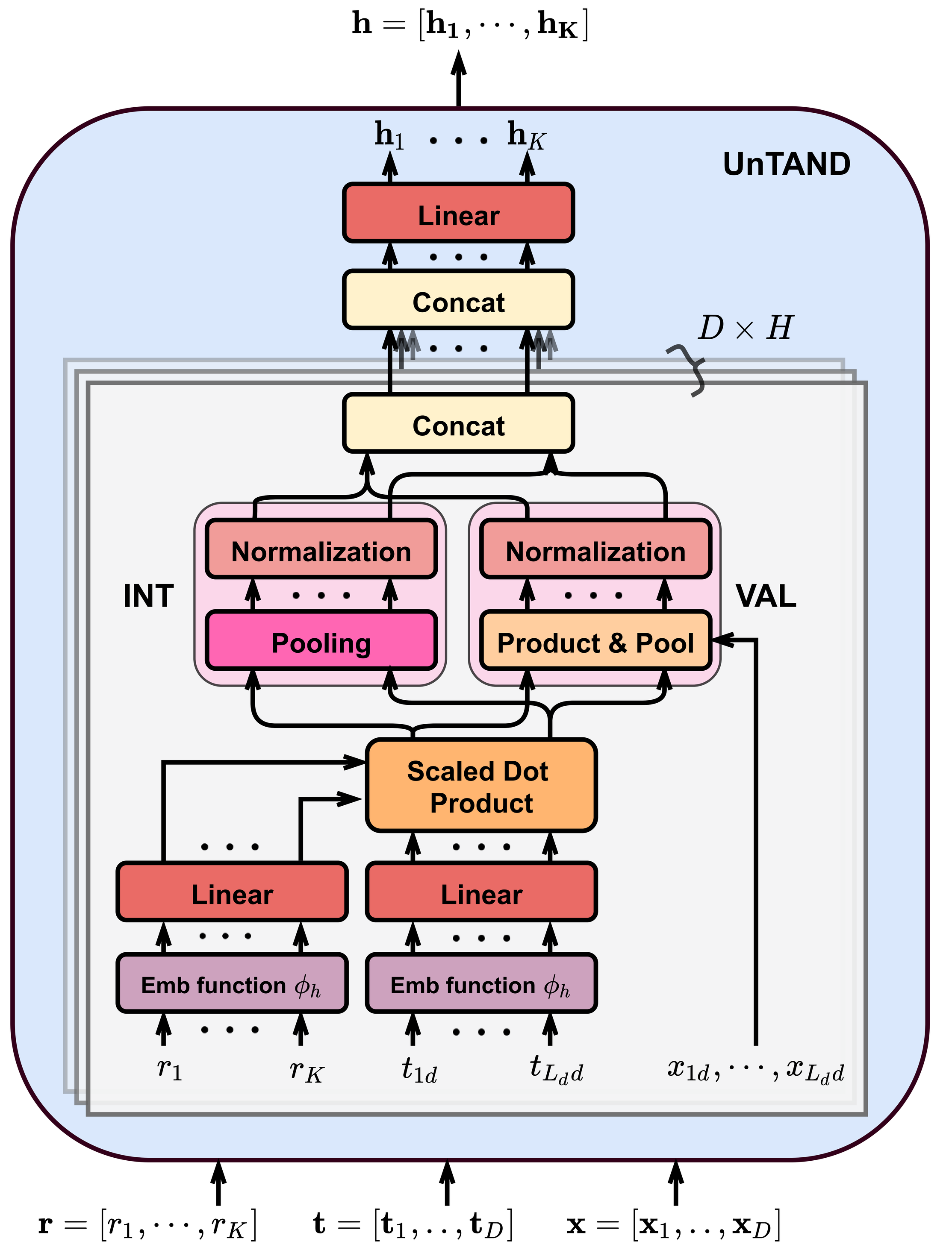


Multivariate regularly (left) and irregularly (right) sampled time series

Contributions

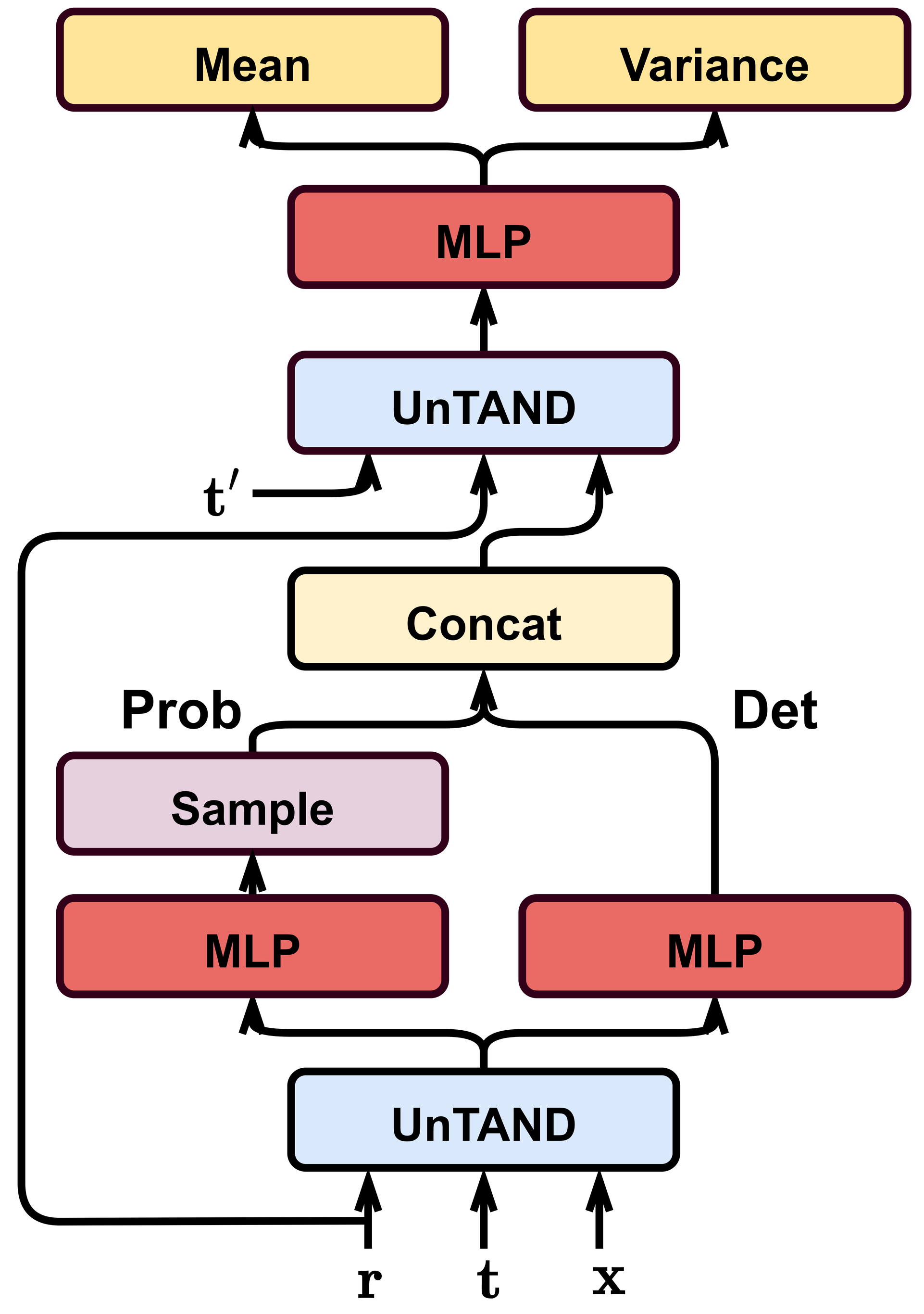
- Proposed Uncertainty-Aware Multi-Time Attention Network layer to encode information about input uncertainty due to variable sparsity.
- Proposed augmented learning objective for training the HeTVAE architecture.
- Our results show that HeTVAE significantly improves uncertainty quantification in the output interpolations compared to several baselines and SOTA methods.

UnTAN: Uncertainty Aware Multi-Time Attention Network



- **UnTANs** contain two encoding pathways to encode information about input uncertainty due to variable sparsity.
- Intensity pathway (INT) focuses on representing information about the sparsity of observations while value pathway (VAL) focuses on representing information about values of observations.

HeTVAE Architecture



- HeTVAE produces a heteroscedastic output distribution to directly represent variable output uncertainty.
- Parallel deterministic and probabilistic latent pathways for propagating input information to the output distribution, including information about input sparsity.
- Use of heteroscedastic output layer can lead to local optima where the mean is essentially flat and all of the structure in the data is explained as noise.

Augmented Learning Objective

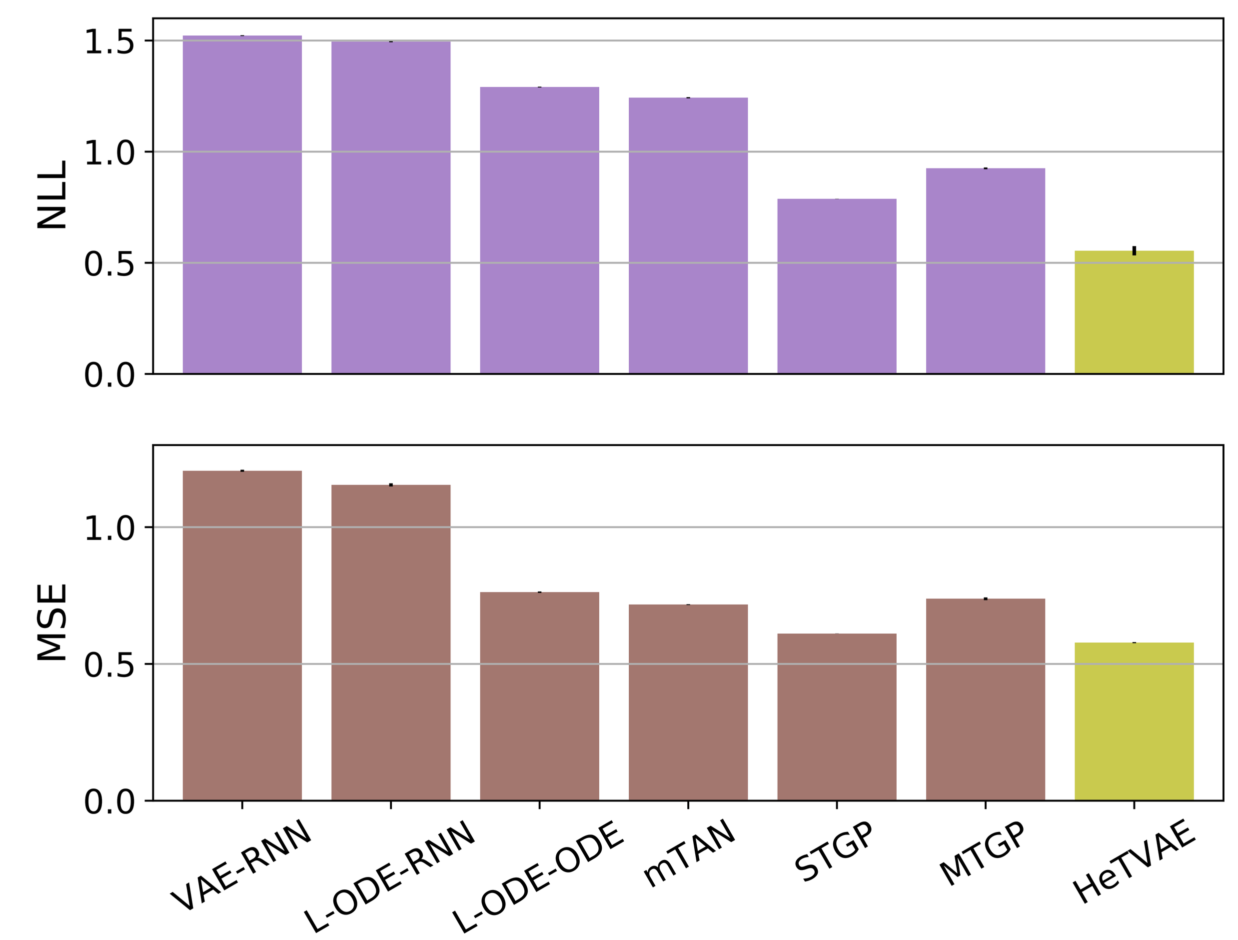
$$\mathcal{L}_{\mathrm{NVAE}}(\theta,\gamma) = \sum_{n=1}^{N} \frac{1}{\sum_d L_{dn}} \Big(\mathbb{E}_{q_\gamma(\mathbf{z}|\mathbf{r},\mathbf{s}_n)}[\log p_\theta^{het}(\mathbf{x}_n|\mathbf{z}_n^{\mathrm{cat}}, \mathbf{t}_n)] - D_{\mathrm{KL}}(q_\gamma(\mathbf{z}|\mathbf{r},\mathbf{s}_n)||p(\mathbf{z})) - \lambda\, \mathbb{E}_{q_\gamma(\mathbf{z}|\mathbf{r},\mathbf{s}_n)} ||\mathbf{x}_n - \boldsymbol{\mu}_n||_2^2 \Big)$$

- The uncertainty agnostic component helps find more informative parameters by introducing a fixed penalty for the mean deviating from the data.
- HeTVAE is trained by maximizing the augmented learning objective on the interpolated time points.

Experimental Protocols

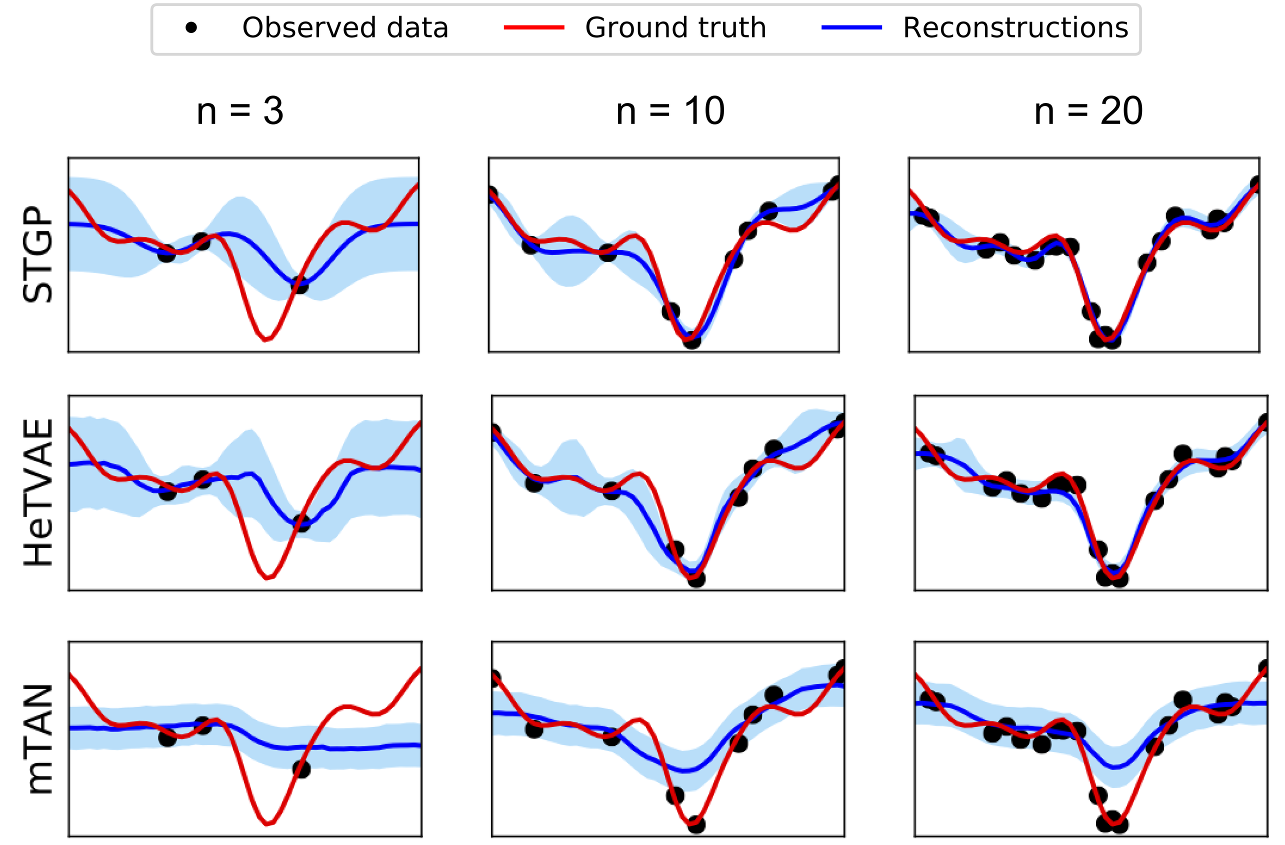
- We randomly divide the data set into a training set (80%) and a test set (20%). We use 20% of the training data for validation.
- During training, we condition on 50% of the available points and compute loss on the rest of the time points.
- At test time, each model is used to infer single time point marginal distributions at the rest of the available time points.

Results on PhysioNet

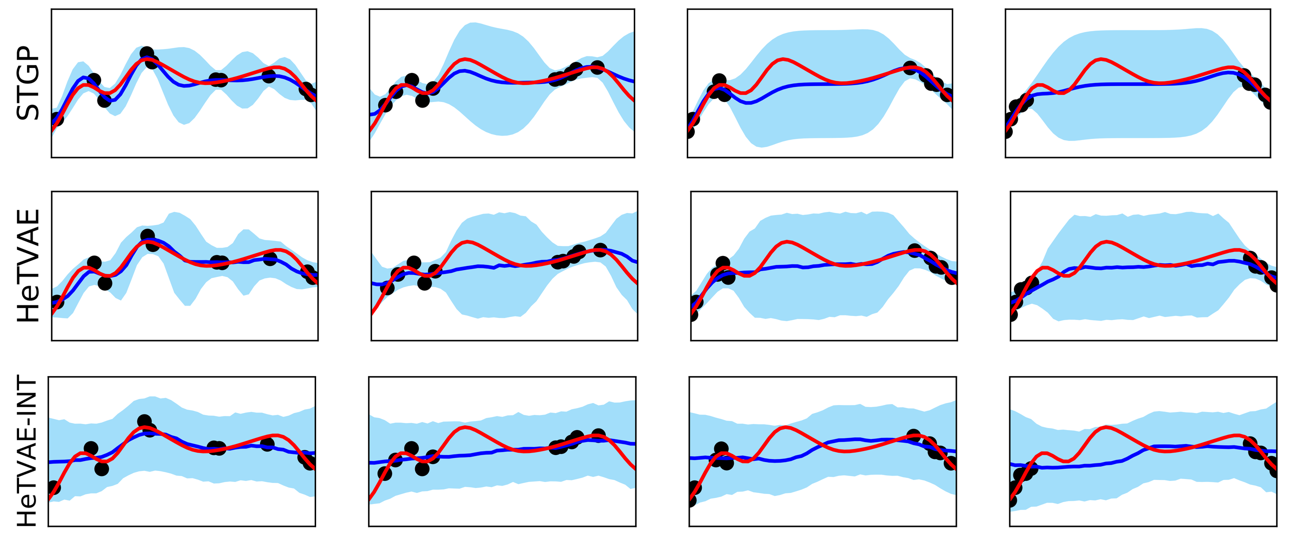


Qualitative Evaluation

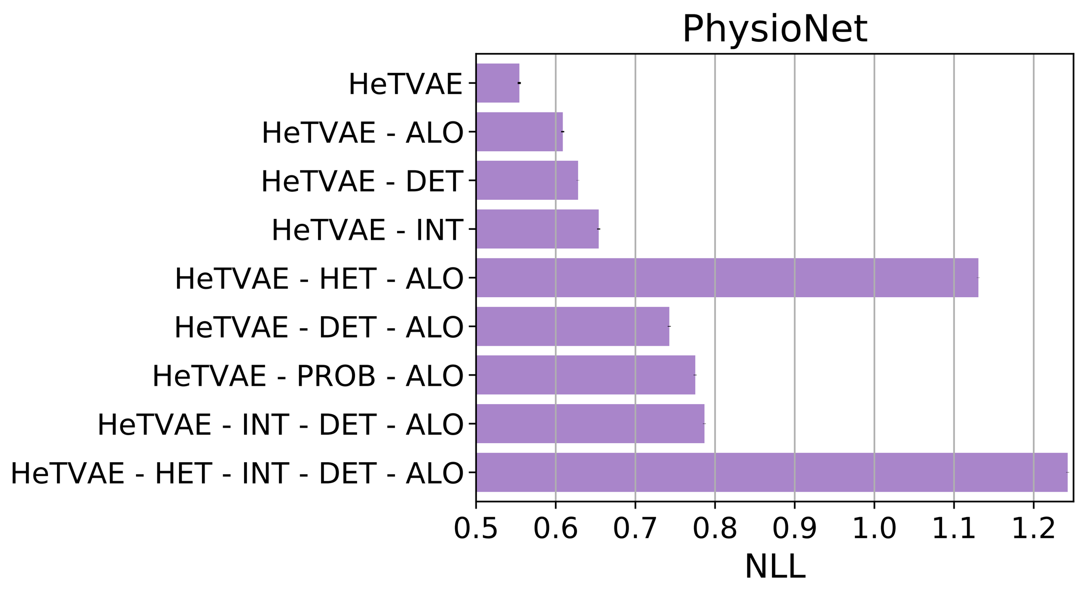
Effect of Sparsity



Effect of Inter-observation Gap



Ablation Study



HET: heteroscedastic layer, ALO: augmented learning objective, INT: intensity encoding, DET: deterministic pathway