

Simple Video Generation using Neural ODEs



David Kanaa^{1*}

Vikram Voleti^{1,2*}

Samira Kahou^{1,3}

Chris Pal^{1,4}

¹Quebec Artificial Intelligence Institute (Mila)

²Université de Montréal, Canada

³McGill University, Canada

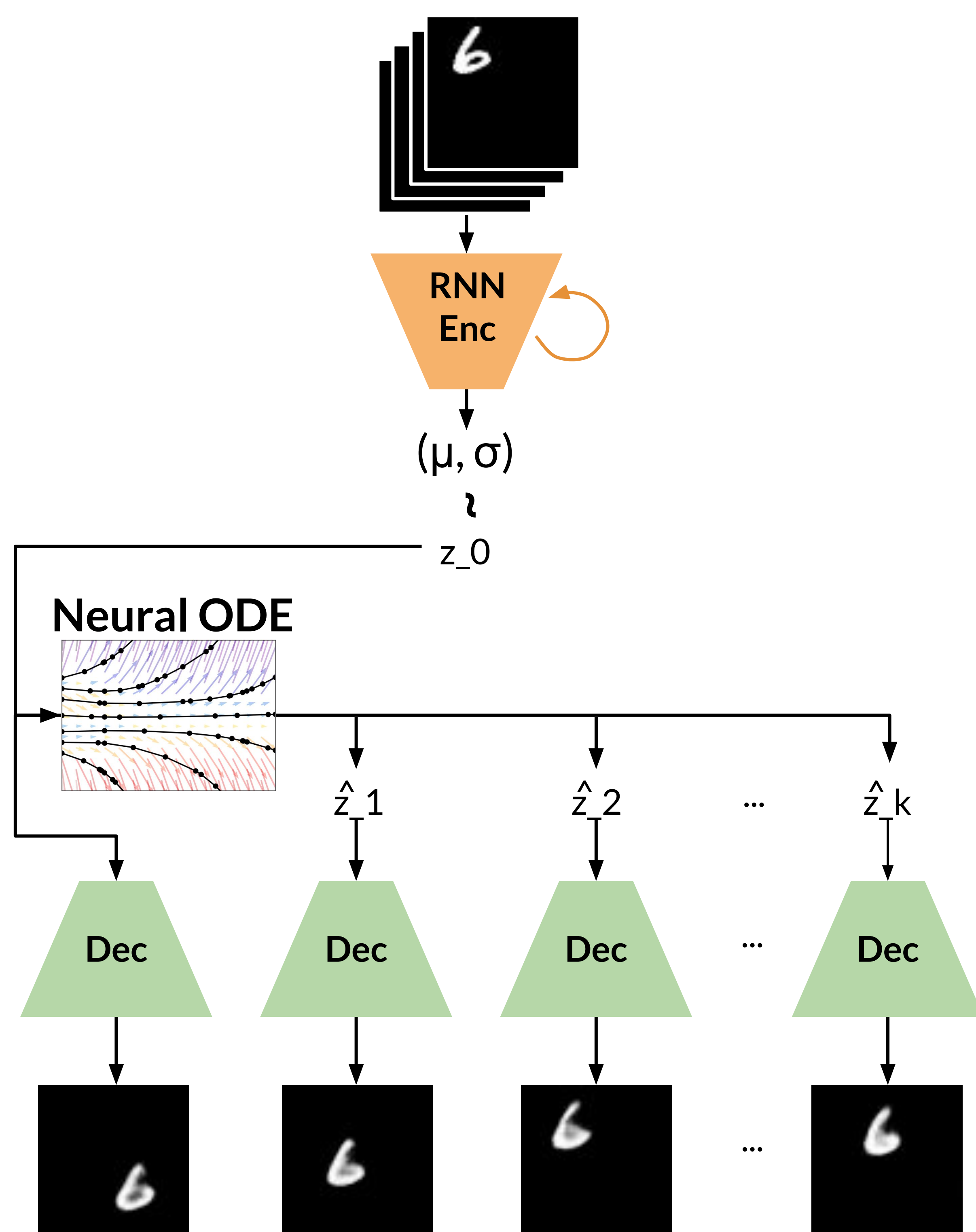
⁴Polytechnique Montreal, Canada

- We study the use of Neural ODEs in video generation.
- We train an encoder-decoder architecture with a Neural ODE in latent space to generate video frames in future time steps.

**Neural ODEs
can be trained
to model
latent dynamics
in video.**

- We show that Neural ODEs can model the latent dynamics in video for the 1-digit and 2-digit Moving MNIST dataset.
- We list the potential use cases and future steps in this direction.

Architecture

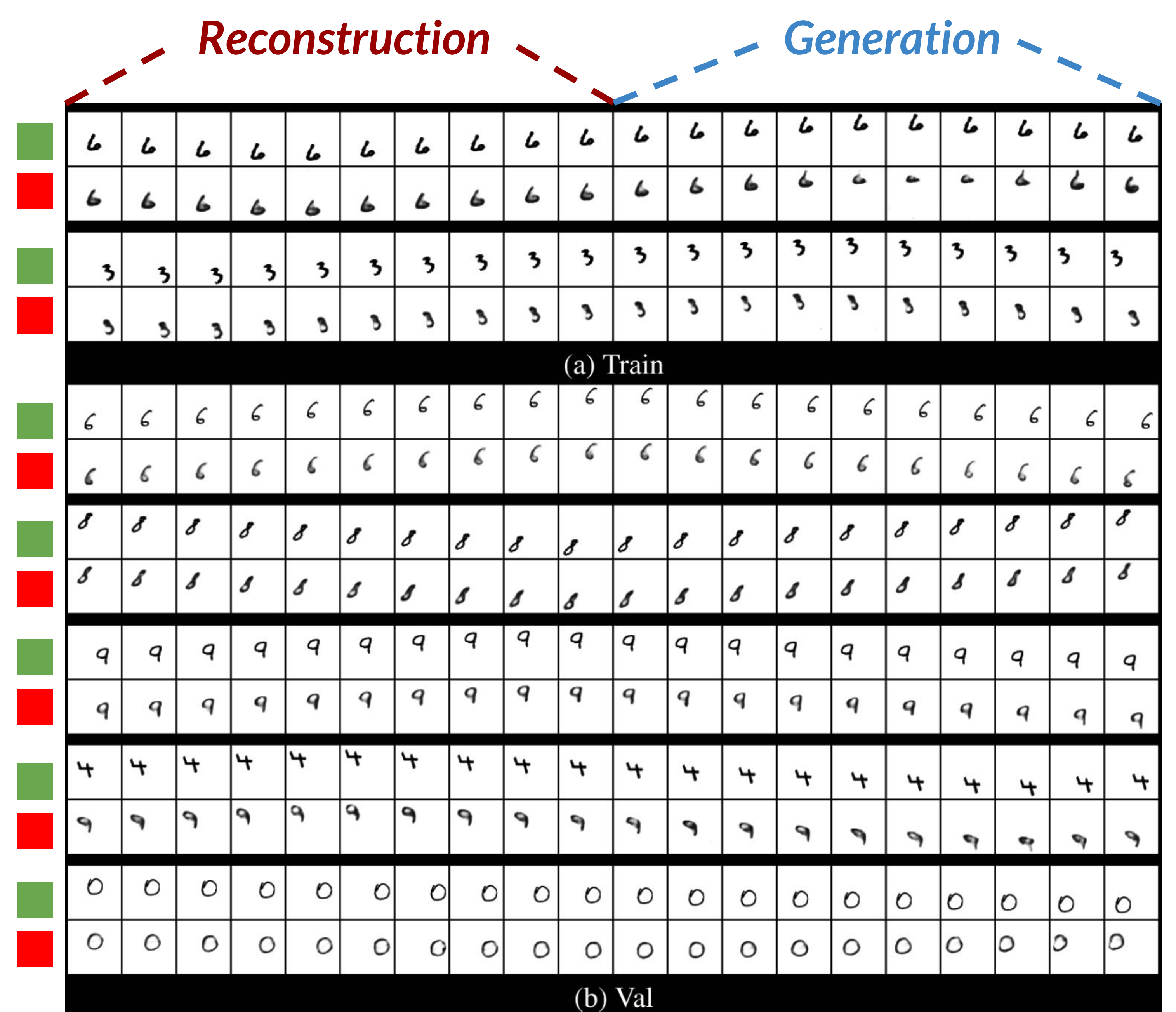


Training & Inference

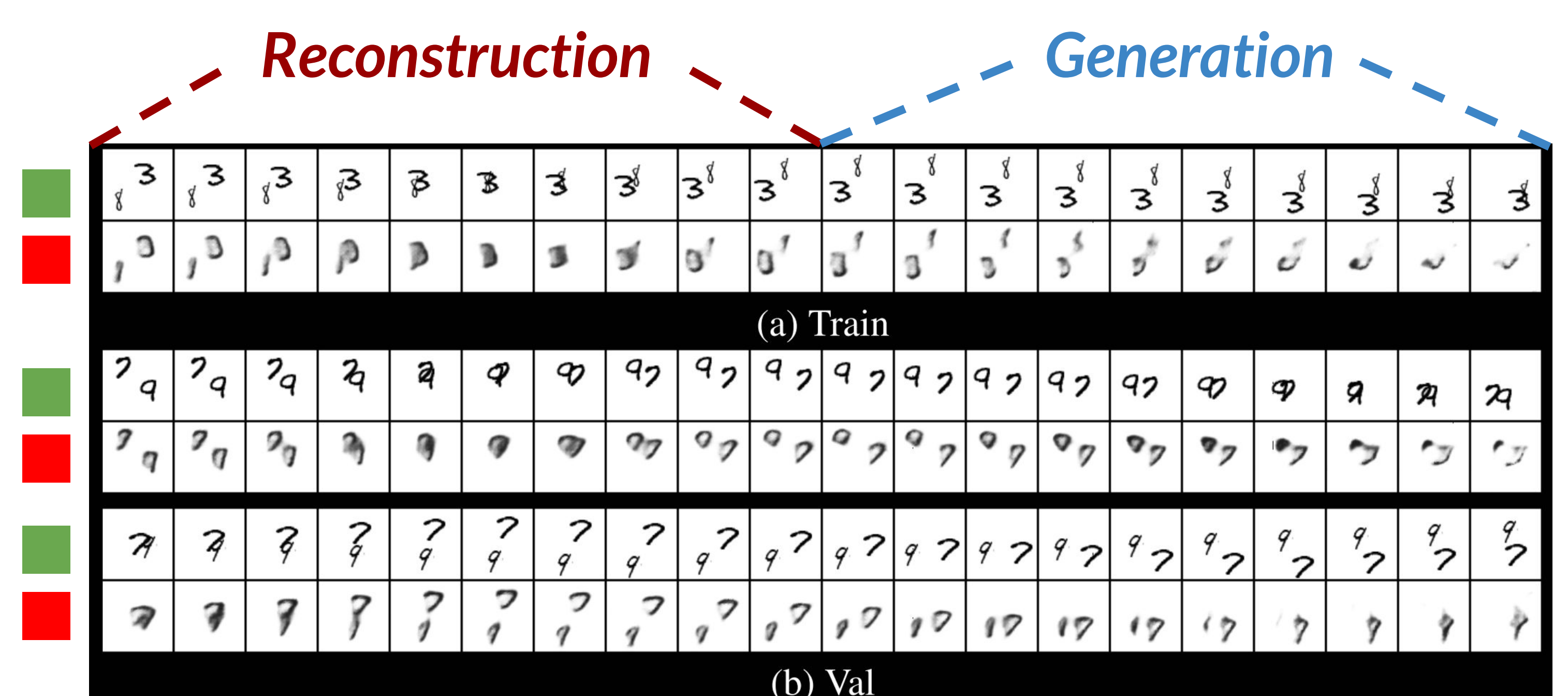
- We trained our model to reconstruct the first 10 frames in the video.
- We use the Neural ODE to extrapolate 10 time steps in the future.
- We decode the future latent points to show video frame generation.

Results

Ground truth
Reconstructed / Generated



Moving MNIST 1-digit



Moving MNIST 2-digit