Physics-aware Spatiotemporal Modules with Auxiliary Tasks for Meta-Learning
Sungyong Seo, Chuizheng Meng, Sirisha Rambhatla, Yan Liu
Department of Computer Science, University of Southern California

Motivation

Limited real-world observations
- Modeling natural phenomena with deep neural networks when only a limited number of observations are available is challenging.
- The sparsely available sensor-based data cause substantial numerical error when we utilize existing differential methods.
- Temporally short events and inevitable missing values from sensors further shorten the length of fully-observed sequences.

Challenges in existing few-shot learning methods
- It is not easy to find a set of similar meta-tasks which provide sharable latent representations needed to understand observations.
- Unlike computer vision or NLP tasks where a common object (images or words) is clearly defined, it is not straightforward to find analogous objects in the spatiotemporal data.
- Exact equations behind natural phenomena are usually unknown, leading to the difficulty in reproducing the similar dataset via simulation.

Proposed Architecture

Schematic overview of the physics-aware meta-learning (PMeta).

Contributions

Modularized PDEs and auxiliary tasks:
- Inspired by forms of PDEs in physics, we decompose PDEs into sharable (spatial) and adaptation (temporal) parts. The sharable one is PDE-independent and specified by auxiliary tasks: supervision of spatial derivatives.

Physics-aware meta-learning:
- We provide a framework for physics-aware meta-learning, which consists of PDE-independent/specific modules. The framework is flexible to be applied to the modeling of different or unknown dynamics.

Synthetic data for sharable modules:
- We extract sharable parameters in the spatial modules from synthetic data, which can be generated from different dynamics easily.

Spatial Derivative Modules: Reusable Modules

- We have claimed that SDM provide reusable features associated with spatial derivatives such as \( \nabla u, \nabla^2 u \) and \( \nabla^2 u \) across different dynamics or PDEs.
- We explore if the proposed SDM based on graph networks can be used as a feature provider for different spatial functions and discretization.

Presentation of Auxiliary Tasks

- We extract shareable parameters in the spatial modules from synthetic data, which can be generated from different dynamics easily.

Modularized PDEs

Decomposability of Variants of a Continuity Equation
- One of fundamental equations in physics describing the transport of physical quantity over space and time is a continuity equation:

\[
\frac{\partial u}{\partial t} + \nabla \cdot u = \sigma,
\]

where \( \rho \) is the amount of the target quantity (u) per unit volume, \( J \) is the flux of the quantity, and \( \sigma \) is a source or sink, respectively.
- Based on the form of \( \sigma \) and \( J \) with respect to a particular quantity \( u \), Eq. 1 can be generalized as:

\[
\frac{\partial u}{\partial t} + \nabla \cdot (F(u, \nabla u, \ldots)) = \sigma,
\]

where the function \( F(\cdot) \) describes how the target \( u \) is changed over time from its spatial derivatives. This equation underlies many specific equations such as the convection-diffusion equation and Navier-Stokes equations:

\[
\begin{align*}
\frac{\partial u}{\partial t} + \nabla \cdot (F(u, \nabla u, \ldots)) &= \sigma, \\
\text{(Convection-Diffusion eqn.)} \\
\frac{\partial u}{\partial t} + \nabla \cdot (\nabla u) &= 0.
\end{align*}
\]

Inspired by the form of Eq. 2, we propose two modules: spatial derivative modules (SDM) and time derivative modules (TDM).

Spatial Derivative Modules (SDM): PDE-independent Modules

Finite difference method (FDM) is used to compute a \( d \)-order derivative as a linear combination of \( n \)-point stencil values.

\[
\frac{\partial^d u}{\partial x^d} \approx \sum_{i=1}^{\alpha} \alpha_i u(x_i),
\]

where \( n > d \). Since Eq. 3 is independent for a form of \( F(\cdot) \) in Eq. 2, we can modularize spatial derivatives as PDE-independent modules and use them as input of \( F(\cdot) \).

Time Derivative Module (TDM): PDE-specific Module

- Once up to \( d \)-order derivatives are modularized by the SDM, the approximated spatial derivatives are fed into an additional module to learn the function \( F(\cdot) \) in Eq. 2.
- This module is PDE-specific as the function \( F(\cdot) \) describes how the spatiotemporal observations change.

Table: Prediction error (MAE) of the first (top) and second (bottom) order spatial derivatives.

<table>
<thead>
<tr>
<th>Region</th>
<th>1-shot (USA)</th>
<th>5-shot (USA)</th>
<th>1-shot (Europe)</th>
<th>5-shot (Europe)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SDM (from scratch)</td>
<td>1.337</td>
<td>2.395</td>
<td>0.768</td>
<td>1.190</td>
</tr>
<tr>
<td>PA-DGN (scratch)</td>
<td>0.022</td>
<td>0.066</td>
<td>0.005</td>
<td>0.009</td>
</tr>
</tbody>
</table>

Experimental Results

Graph Signal Generation

We adopt a set of multi-step spatiotemporal sequence generation tasks to evaluate our proposed framework on two real-world dataset (AQI-CO [2]: air quality index, ExtremeWeather [3]: the extreme weather dataset).

Graph Signal Regression [4] conducted a graph signal regression task: predict the temperature \( x_t \) from the temperature on the previous 5 days \( (x_{t-1} \ldots x_{t-5}) \). We split the GHCN dataset (Global Historical Climatology Network (GHCN) provided by National Oceanic and Atmospheric Administration (NOAA) spatially into two regions: 1) the USA (1,705 stations) and 2) Europe (703 stations) where there are many weather stations full functioning.

Table: Graph signal regression results (MAE) on the two regions of weather stations.

<table>
<thead>
<tr>
<th>Region</th>
<th>1-shot (USA)</th>
<th>5-shot (USA)</th>
<th>1-shot (Europe)</th>
<th>5-shot (Europe)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SDM (from scratch)</td>
<td>1.337</td>
<td>2.395</td>
<td>0.768</td>
<td>1.190</td>
</tr>
<tr>
<td>PA-DGN (scratch)</td>
<td>0.022</td>
<td>0.066</td>
<td>0.005</td>
<td>0.009</td>
</tr>
</tbody>
</table>

Conclusion

In this paper, we propose a framework for physics-aware meta-learning with auxiliary tasks. By incorporating PDE-independent knowledge (spatial derivatives) from simulated data, the framework provide reusable features and the features help improve the meta-test tasks with a limited amount of data.

References