Scaling Up AI-driven Scientific Discovery via Embedding Physics Modeling into End-to-end Learning and Harnessing Random Projection

Yexiang Xue[1], Md Nasim[1], Xinghang Zhang[2], Anter El-Azab[2]

[1]Department of Computer Science, Purdue University. [2]School of Materials Engineering, Purdue University. *{yexiang, mnasim, xzhang98, aelazab}@purdue.edu*

PURDUE UNIVERSITY

Motivation

- Simulation-based methods to learn physics models requires perfectly annotated data and computationally expensive simulation
- Human error during annotation leads to error in physics models
- Expensive high dimensional simulation prohibits learning of physics models from fine temporal and spatial resolution data

Data

- Examples:
  - Sensor readings
  - Videos
  - Remote sensing (satellite images)
  - ...
- High dimensional
- Noisy

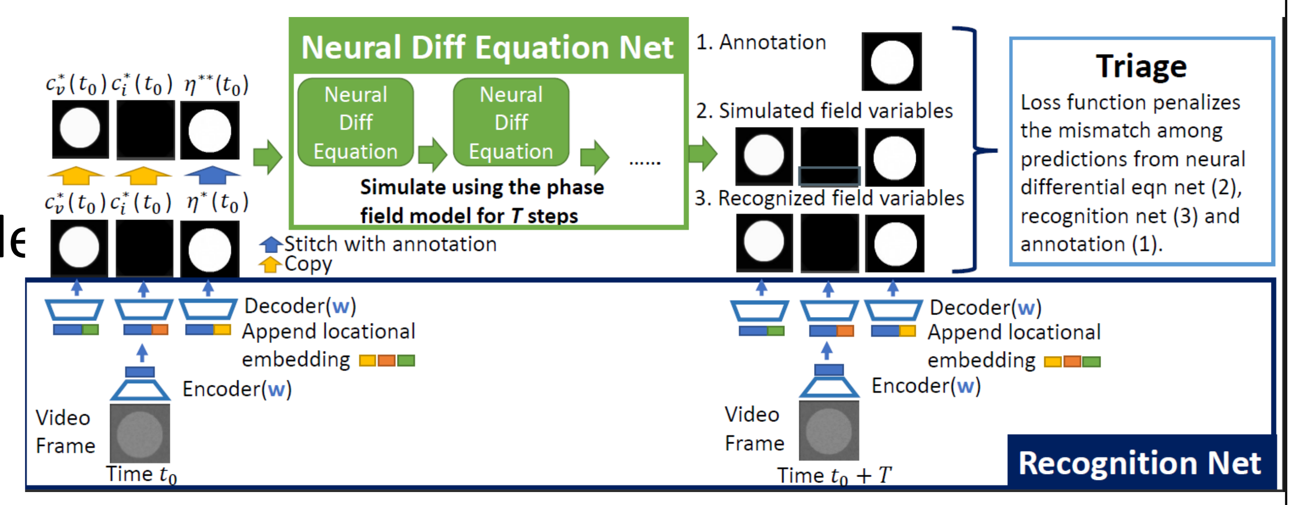
Need efficient methods to accelerate scientific discovery

Physics Models

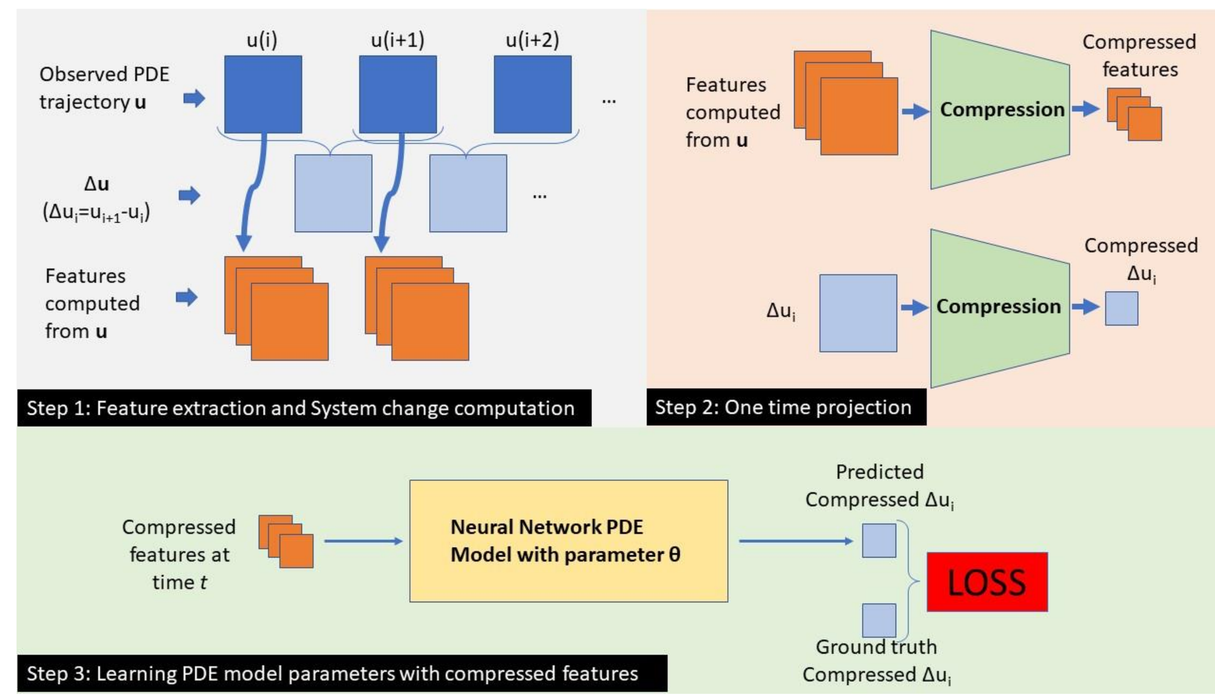
- Examples:
  - Conservation laws
  - Phase field models
  - ...
- Explains the hidden dynamics of the physical world
- Leads to innovation of new tech

Technical Approach

- **Embedding physics modeling into end-to-end learning**
  - Uses a combination of 2 neural networks
  - Embeds differential equations as fully differentiable neural network layers
  - Exhibits superior performance in learning physics model of void evolution dynamics from data,

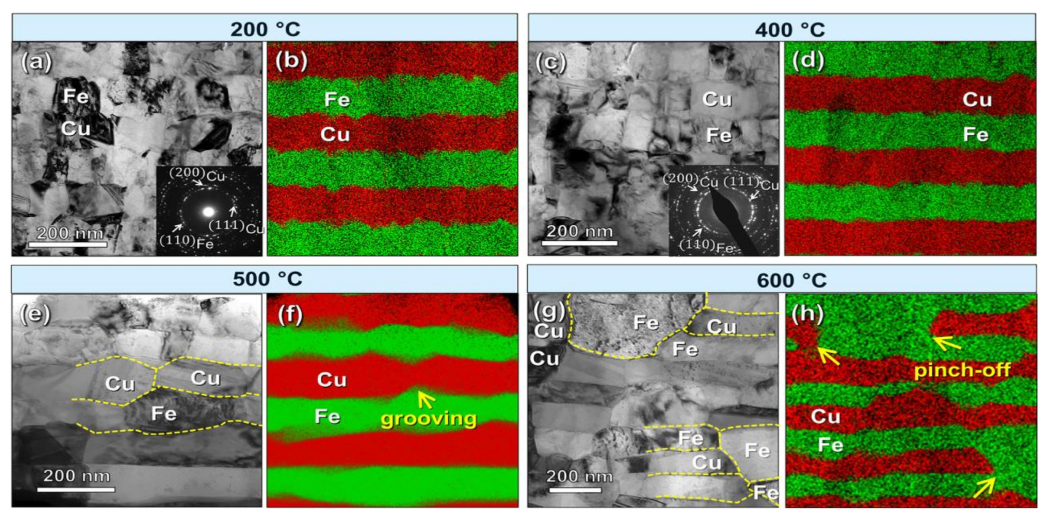


- **Efficient learning of sparse and decomposable PDE models using random projection**
  - Identifies a class of PDE models, having sparse temporal updates and decomposable spatial features
  - Uses random projection to compress sparse features and temporal updates into low dimensional space, and conduct learning in the low dimensional space
  - In experiment, can learn correct physics models from data compressed at 0.05% of original size

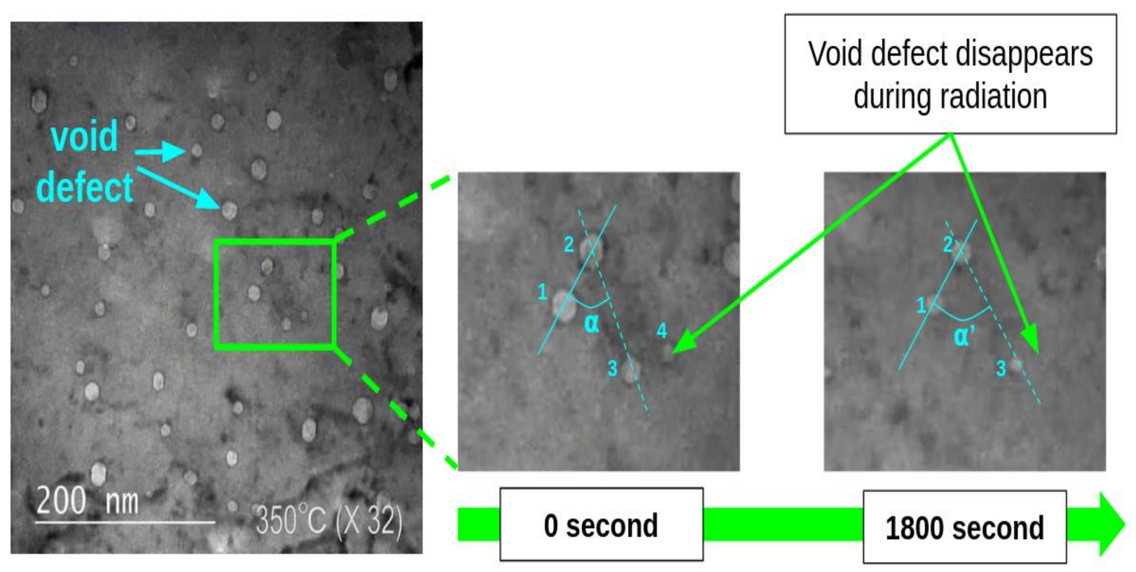


Real World Impact of Our Learning Methods

Many real-world systems (e.g., phase field model of grain growth in materials, material defect evolution in irradiated materials) are modeled with partial differential equations, and our method can greatly accelerate learning these models from data efficiently.



(b) Grain growth in Cu/Fe 100 nm multilayer upon annealing at different temperatures. XTEM micrographs and EDS maps show the microstructure evolution at different temperatures. [1]



(a) Void shaped defects in *Cu* specimen at 350 C. These defects are dynamic and can change size and position and can also disappear as shown in figure.

Broader Impact

- Designing sustainable materials for extreme environment
  - Sustainable material design for extreme environments e., inside nuclear reactors, requires analysis of experimental data, which are noisy and high resolution. Our methods can efficiently analyze these data, and in essence accelerate the design process.
- Understanding properties of materials
  - Broad understanding of the physical world can lead to fundamentally transformative technologies (e.g., semiconductors). Our method has the potential to accelerate scientific discovery in material science domain by streamlining the analysis process and reducing manual effort.

Summary

We propose efficient methods to learn partial differential equation-based physics models from data. Our methods are robust against human annotation error and can greatly reduce the cost of high dimensional simulation associated with learning. Our proposed techniques can greatly accelerate scientific discovery in the domain of material science, and lead to the design of more sustainable materials for high temperature high irradiation environments such as that inside of a nuclear reactor.