

A comparison between transformers and recurrent neural networks for non-linear elastoplastic behavior in 3D micro-mechanical modeling

Petter Uvdal^{1,*}, Mohsen Mirkhalaf¹

¹Department of Physics
University of Gothenburg
Origovägen 6B, 41296 Gothenburg, Sweden

Abstract

During the last decades, Short Fiber Reinforced Composites (SFRCs) have gained attention in various industries for their favorable strength-to-density ratio, manufacturability, and cost-effectiveness. They can be injection-molded into complex 3D shapes, giving rise to a wide variety of fiber orientation distributions, which necessitates the generation and homogenization of 3D Representative Volume Elements (RVEs), as illustrated in Figure 1. However, high-fidelity micro-mechanical simulations of SFRCs remain computationally expensive - especially for non-linear, path-dependent elasto-plastic analyses, where each RVE homogenization involves a considerable numerical cost. Recurrent Neural Network (RNN) approaches have shown promise as surrogate models for history-dependent plasticity [1–2], yet challenges remain in accurately capturing long-term dependencies within a loading history. Recently, pre-trained (PT) transformer models using self-attention mechanisms [2] have been proposed for surrogating micromechanical homogenization of 2D composites [4]. By leveraging self-attention mechanisms, the transformer architecture can capture sequence-to-sequence dependencies and provides great accuracy for complex modeling cases. Yet, questions remain regarding potential specific advantages of PT transformers over RNNs for non-linear plasticity models, including handling data scarcity during training.

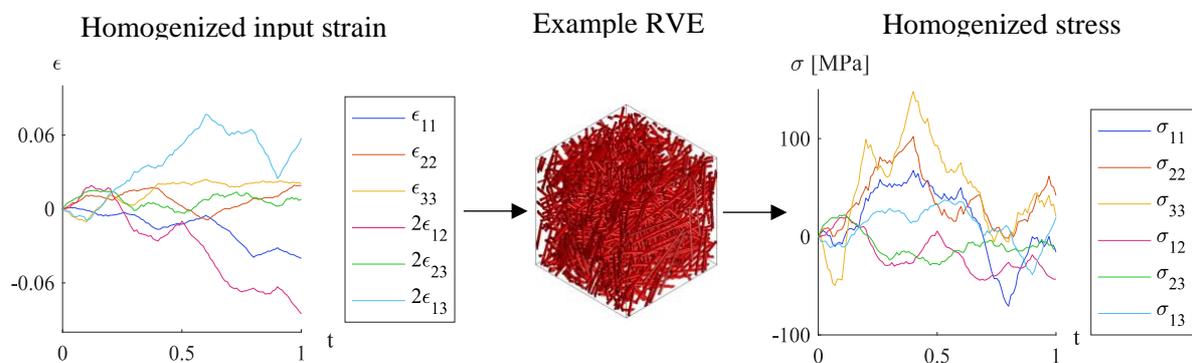


Figure 1: Homogenized strain inputs and stress outputs using DIGMAT-FE (re-illustrated from [5]).

To address these questions, we propose a direct comparison of a classical RNN model and a transformer-based surrogate model for capturing the non-linear behavior in 3D RVEs of a glass-fiber composite material. Our approach involves generating synthetic data via high-fidelity FE or FFT-based micro-mechanical analyses, and subsequently training both the RNN and the transformer. The transformer architecture can capture both local and global strain-stress dependencies across sequential data, often providing an improved predictive accuracy for longer time sequences. Moreover, by enabling parallelized computation, it provides relatively fast training and inferences time. However, increasing the number of attention layers introduces more parameters, leading to potentially difficult training with limited data. While multiple

*Corresponding author: petter.uvdal@physics.gu.se

approaches exist for transfer learning and pre-training in both RNNs and transformers, we mitigate data scarcity by incorporating a previously developed augmentation strategy [1]. This expands the smaller datasets, thus improving model robustness without additional high-fidelity simulations.

References

- [1] H.L. Cheung, P. Uvdal, M. Mirkhalaf, (2024). Augmentation of scarce data - a new approach for deep-learning modelling of composites. *Composites Science and Technology* 249, 110491.
- [2] M. Mozaffar, R. Bostanabad, W. Chen, K. Ehmann, J. Cao, and M. A. Bessa, (2019). Deep learning predicts path-dependent plasticity. *Proceedings of the National Academy of Sciences*, 116(52):26414–26420.
- [3] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, I. Polosukhin (2017). Attention is all you need. *Advances in Neural Information Processing Systems*.
- [4] Y. Zhongbo and P. L. Hien (2024). Pre-trained transformer model as a surrogate in multiscale computational homogenization framework for elastoplastic composite materials subjected to generic loading paths. *Computer Methods in Applied Mechanics and Engineering*, 421:116745.
- [5] P. Uvdal and M. Mirkhalaf, (2024). Test-time data augmentation: improving predictions of recurrent neural network models of composites. *arXiv preprint*, 2409.02478.