State Space Models Are All You Need

Cai, Junxiang (School: National Junior College)
Ong, Aidan (School: Hwa Chong Institution)

Conventional architectures such as the Transformer struggle to scale to very long sequences of over 10 000 steps. Promisingly, recent advances in sequence models based on the Structured State Space Sequence (S4) model like Liquid-S4, S5 and S6 have shown remarkable performance in handling sequences with long-term dependencies such as image, text, audio, and medical time-series data. In this paper, we propose two state-of-the-art (SOTA) architectures based on the State Space Model (SSM). Our first architecture, Liquid-S5, is a multi-input, multi-output (MIMO) SSM that can dynamically modify its state based on incoming inputs during inference. Liquid-S5 achieves SOTA on the Long Range Arena benchmark, demonstrating its ability to handle intricate long-range dependencies on continuous sequences. Our second architecture, LiquidMamba, builds upon selective SSMs (S6) by introducing the notion of data-dependent state updates to the S6 architecture. LiquidMamba also introduces two modifications. Firstly, we use a gated convolution over the normal convolution in vanilla Mamba to improve parameter efficiency. Secondly, we add the Feed-Forward Network that was removed in the original Mamba architecture back into LiquidMamba to improve its generalization capabilities. Notably, LiquidMamba exceeds the performance of Mamba even with the Feed-Forward Network added, as well as Transformer architectures for a given parameter-count on the WikiText-103 benchmark. We posit that both Liquid-S5 and LiquidMamba can serve as highly efficient and accurate architectures for learning representations from sequential data.

Awards Won:
Third Award of $1,000
Association for the Advancement of Artificial Intelligence: AAAI Student Memberships for each finalist that is part of the 1st, 2nd, and 3rd Prize Winning projects and 5 Honorable Mention winning projects (up to 3 students per project) (in-kind award / part of the 1st-3rd prize)
Association for the Advancement of Artificial Intelligence: AAAI Membership for the School Libraries of All 8 Winners (in-kind award / part of 1st-3rd prize and honorable mentions' prize)
Association for the Advancement of Artificial Intelligence: First Award of $1,500
National Security Agency Research Directorate : Third Place Award “Cybersecurity”