We trained a model based on Qwen-32B-SFT-R1 [1] using the CodeContests+ dataset. The RL algorithm used was GRPO [2], and the evaluation dataset was LiveCodeBench [3]. The experimental results are as follows. The experimental results show that conducting RL using CodeContests+ can significantly enhance the model's coding and reasoning abilities, and enable the model to reach the same upper-bound of capabilities as the current SOTA models.

Model	LiveCodeBench pass@1
GPT-4o	32.9
OpenAI o1-mini	53.8
OpenAI o1	63.4
OpenAI o3-mini (Low)	60.9
Qwen2.5-32B-SFT-R1	57.2
Qwen2.5-32B-SFT-R1 + CodeContests+ (Ours)	**64.0**

References

[1] Guo, Daya, et al. "Deepseek-r1: Incentivizing reasoning capability in llms via reinforcement learning." *arXiv preprint arXiv:2501.12948* (2025).
[2] Shao, Zhihong, et al. "Deepseekmath: Pushing the limits of mathematical reasoning in open language models." *arXiv preprint arXiv:2402.03300* (2024).
[3] Jain, Naman, et al. "Livecodebench: Holistic and contamination free evaluation of large language models for code." arXiv preprint arXiv:2403.07974 (2024).