

Transferring Foundation Models for Generalizable Robotic Manipulation Supplementary Material

Jiange Yang¹, Wenhui Tan², Chuhao Jin², Keling Yao³, Bei Liu⁴,
Jianlong Fu⁴, Ruihua Song², Gangshan Wu¹, Limin Wang^{1,5*}

¹State Key Laboratory for Novel Software Technology, Nanjing University, China

²Renmin University of China

³The Chinese University of Hong Kong, Shenzhen

⁴Microsoft Research ⁵Shanghai AI Lab

jiangeyang.jgy@gmail.com, {tanwenhui404, jinchuhao, rsong}@ruc.edu.cn,
120090220@link.cuhk.edu.cn, {jianf, Bei.Liu}@microsoft.com, {gswu, lmwang}@nju.edu.cn

1. Low-cost dual-arm robot experiments details.

In this section, we report the experimental setup details of our TPM extended to new low-cost dual-arm robot capabilities for folding and stacking skills. Specifically, we co-fine-tune the previous TPM model weight using 50 new demonstrations and deploy it using Mixformer-small [1] and SAM-tiny [2]. The demos can be seen Fig 1. It is worth mentioning that our low-cost dual-arm does not require additional motion planning time due to without movelt calls.

References

- [1] Yutao Cui, Tianhui Song, Gangshan Wu, and Limin Wang. Mixformerv2: Efficient fully transformer tracking. *arXiv preprint arXiv:2305.15896*, 2023. 1
- [2] Chaoning Zhang, Dongshen Han, Yu Qiao, Jung Uk Kim, Sung-Ho Bae, Seungkyu Lee, and Choong Seon Hong. Faster segment anything: Towards lightweight sam for mobile applications. *arXiv preprint arXiv:2306.14289*, 2023. 1

*Corresponding author.

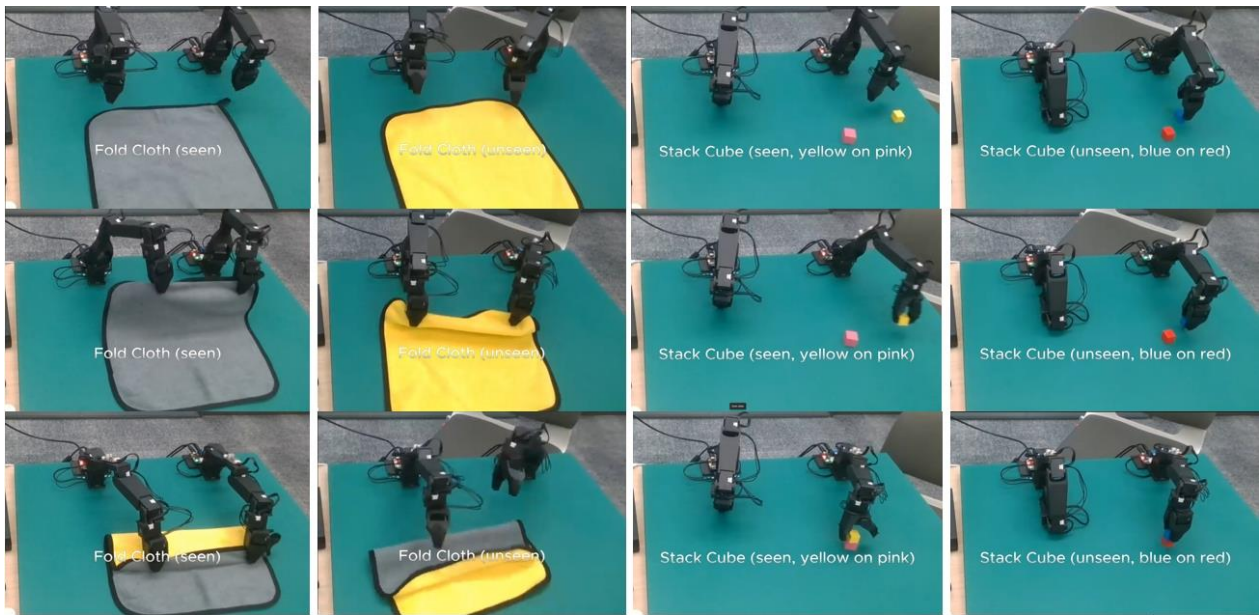


Figure 1. The demos of folding cloth and stacking cube skills.