TDRI: Two-Phase Dialogue Refinement and Co-Adaptation for Interactive Image Generation

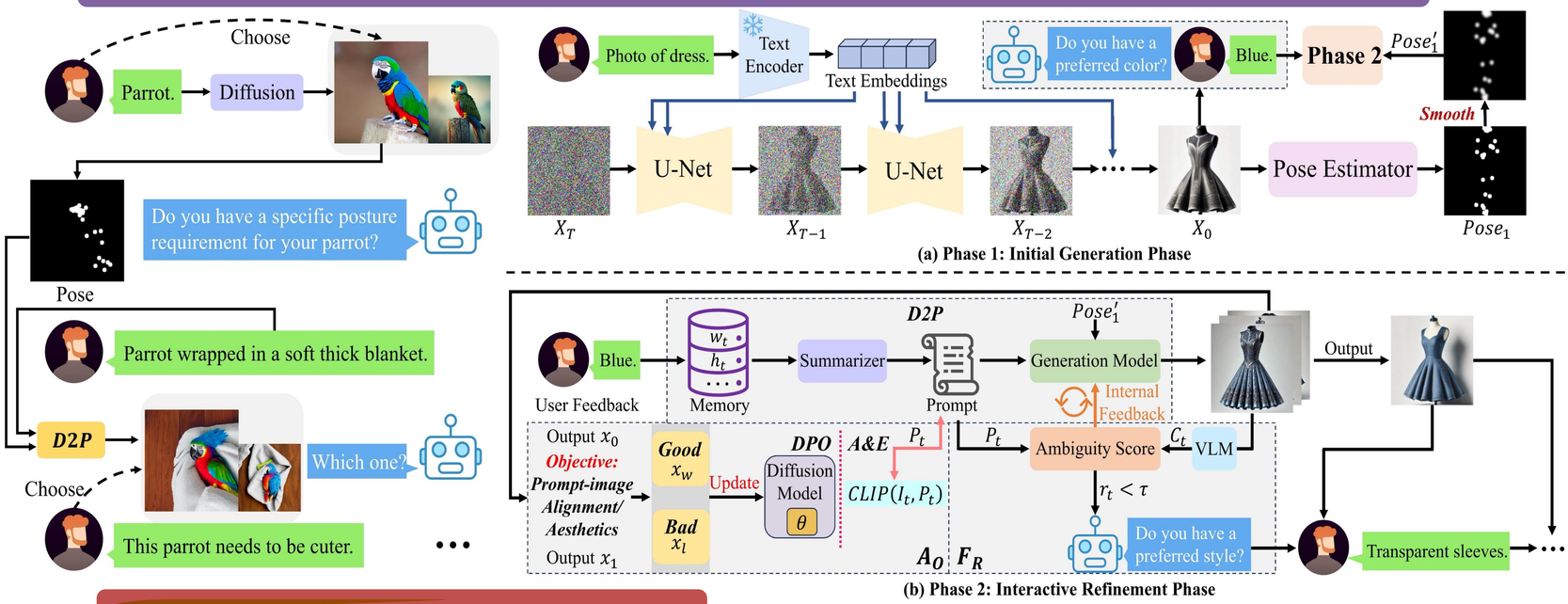
Yangfan He, Yuheng Feng, Jianhui Wang, Kun Li, Yijin Wang, Haoyuan Li, Sida Li, Yinghui Xia, TIANYU SHI, Miao Zhang
22061300002@stu.xidian.edu.cn ty.shi@mail.utoronto.ca edward.yang@autoagents.ai

AutoAgents 未来式智能

Introduction

- **Motivation and Goal**
  - Text-to-image models struggle with ambiguitiesin user prompts and understanding user intent
  - Non-expert users without prompt engineeringtraining face challenges in getting desired results
  - Multiple iterations are typically needed toachieve satisfactory outputs
- **Our Approach: TDRI Framework**
  - A two-phase dialogue system that refines image outputs through iterative feedback

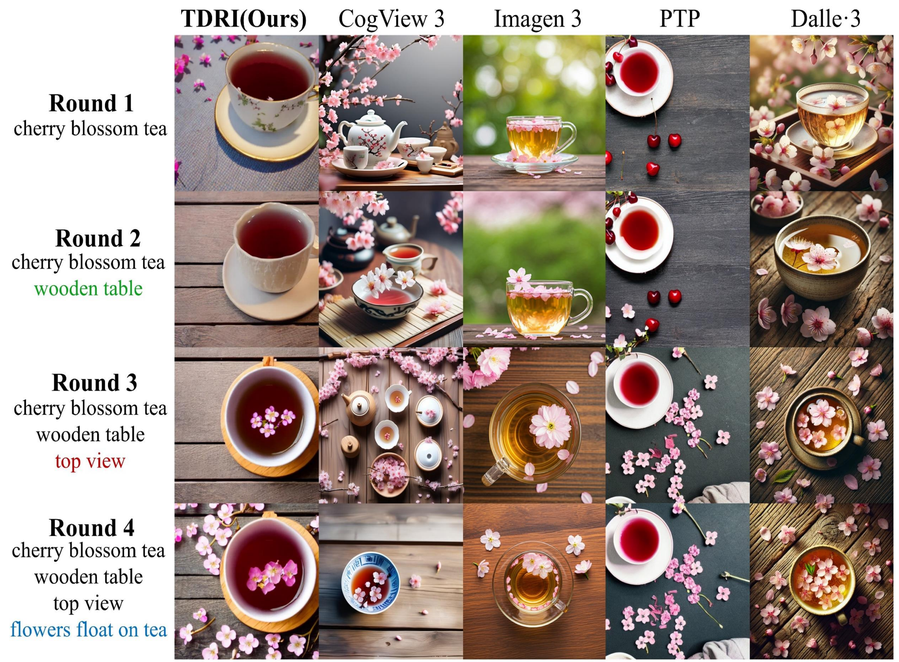
Proposed Framework for TDRI



Experiment Results

Experiment	Key Metrics	Main Results
Text-to-Image Alignment Comparison	CLIP/BLIP scores, Human Voting	TPR (0.93/0.98/0.385) outperformed GPT-4 representation (0.92/0.151)
Multi-dialog Performance	CLIP/BLIP similarity scores across rounds	TDR reached 94.7% compared to 84% for GPT-4
Attend-and-Excite Performance	Usage frequency, T2I similarity improvement	SD-1.4 improved from 0.728 to 0.804 (CLIP)
User Perception	Survey results	Most users found ideal results by Round 5 (21.1%)
Image Editing vs. From Scratch	Consistency, Satisfaction,	Image editing (0.88, 90%, 9 min) outperformed generation from scratch
Simple vs. Complex Prompts	Success rate, CLIP score, Human voting	Simple prompts (92%, 0.95, 87%) significantly outperformed complex
Model Size Comparison	Satisfaction, CLIP score, Computation time	7B model (90%, 0.83, 15 min) showed best quality 3B model (78%, 0.77, 8 min) was faster

Comparison Effect



Key Contributions

- Specialized human-machine interactiontechniques for interactive image generation
- A two-phase dialogue methodology combiningexternal user interactions with internaloptimization processes
- Demonstrated applicability across variousimage generation tasks

If you have any questions, please feel free to contact us