Story Generation in Multi-Agent Systems: A DynamicCollaboration Framework Based on Reinforcement Learning

Yaolin Li *

School of Economics and Management, Nanjing University of Science and Technology, Nanjing, China

* Corresponding Author Email: 1547182654@qq.com

Abstract. In artificial intelligence, multi-agent systems excel in complex tasks requiring coordination. This paper presents a Multi-Agent Dynamic Collaboration Framework for story generation, tackling narrative coherence, diversity, and adaptability. Using GPT-2 agents for plot, dialogue, and description, a Transformer-based coordinator dynamically adjusts contributions via multi-head attention. PPO reinforcement learning optimizes behaviors with rewards based on BLEU, ROUGE, and entropy. Trained on ROCStories, the framework achieves BLEU 0.85, ROUGE-L 0.65, and perplexity 12.5, surpassing baselines. Human assessments show improved coherence and engagement. This advances AI creative writing and enables scalable applications in education, entertainment, and marketing for immersive narratives.

Keywords: Multi-Agent Systems; Story Generation; Reinforcement Learning; Transformer, Narrative AI.

1. Introduction

1.1. Background and Broader Context

Artificial intelligence has profoundly transformed content creation, particularly in narrative domains where machines emulate human creativity. At the macro level, AI-driven generative models have revolutionized industries such as entertainment, education, and publishing by enabling the automated production of text-based content. These advancements stem from foundational developments in natural language processing (NLP), where large language models (LLMs) such as the GPT series have demonstrated remarkable capabilities in generating coherent and contextually relevant text [1, 2]. However, as demand grows for sophisticated, multi-faceted narratives, single-model architecture increasingly reveal limitations in handling intricate inter-dependencies such as maintaining plot consistency while integrating diverse dialogues and descriptive passages.

Narrowing the focus, multi-agent systems (MAS) emerge as a paradigm shift, decomposing complex tasks into collaborative sub-tasks. Inspired by human teamwork in creative writing, MAS facilitate role specialization and interaction, enhancing overall output quality [3,4]. Yet, in story generation specifically, challenges persist in ensuring seamless coordination among agents, often resulting in fragmented narratives lacking depth or adaptability.

1.2. Research Gaps and Motivations

Existing literature highlights several contradictions and deficiencies. For instance, while hierarchical models effectively outline stories, they frequently overlook dynamic inter-agent adjustments, leading to rigid outputs [5, 6]. Single-agent reinforcement learning (RL) optimizes for coherence but struggles with diversity, as evidenced by repetitive patterns in generated texts [7], [8]. Moreover, static collaboration mechanisms fail to adapt to varying narrative complexities, a limitation critiqued in recent benchmarks [9, 10].

These gaps underscore the need for a framework that integrates dynamic weighting and reinforcement learning (RL) to bridge theoretical elegance with practical robustness motivating this study to address the following unresolved question: How can agents dynamically collaborate to enhance narrative quality without increasing computational complexity?

1.3. Research Gaps and Motivations

Building on these gaps, this research poses the following questions:

(1) How does dynamic attention mitigate static collaboration limitations in MAS for story generation? (2) What interplay exists between specialized agents and RL optimization in improving coherence and diversity? (3) Can such a framework achieve high efficiency with low complexity metrics like reduced perplexity and training time?

The contributions are:

- A novel integration of Transformer-based dynamic collaboration with PPO-optimized RL, enabling adaptive agent interactions that enhance narrative coherence (e.g.,15% BLEU improvement) while maintaining low complexity (124M parameters).
- Empirical validation on ROCStories, demonstrating superior metrics over baselines, with insights into the interplay between agent specialization and reward functions for balanced outputs.
- A scalable architecture that reduces training overhead (5.5 hours on RTX 3090) and offers efficiency advantages, addressing gaps in existing MAS by prioritizing generalization across dataset scales.

2. Related Work

2.1. Multi-Agent Systems in NLP and Generation Tasks

Multi-agent systems in natural language processing (NLP) have been organized around the themes of coordination and specialization. Role-based decomposition, as in code generation [4], contrasts with dialogue systems where agents negotiate [11, 12]. However, these approaches often exhibit methodological limitations, such as over-reliance on fixed hierarchies, leading to scalability issues in creative tasks [13].

2.2. Reinforcement Learning and Dynamic Mechanisms

RL themes focus on policy optimization for generative tasks. PPO variants enhance stability but contradict single-agent assumptions in multi-faceted scenarios [14], [15]. Dynamic attention complements this approach by enabling adaptive interactions., yet inherits defects like high computational demands, as critiqued in healthcare benchmarks [3].

2.3. Gaps and Connections Across Studies

Literature dialogues reveal inheritances (e.g., from MAS in games to NLP [1]) and debates (e.g., static vs. dynamic [5, 6]). Mutually, they highlight collective insufficiencies: lack of integrated RL-dynamic frameworks for narratives, paving the way for our contribution in bridging these through efficient, adaptive MAS.

3. Methodology

3.1. Framework Overview

The Multi-Agent Dynamic Collaboration Framework (MADCF) decomposes story generation into agent-specific subtasks. As illustrated in Fig.1, the process begins with input prompts fed to specialized agents.

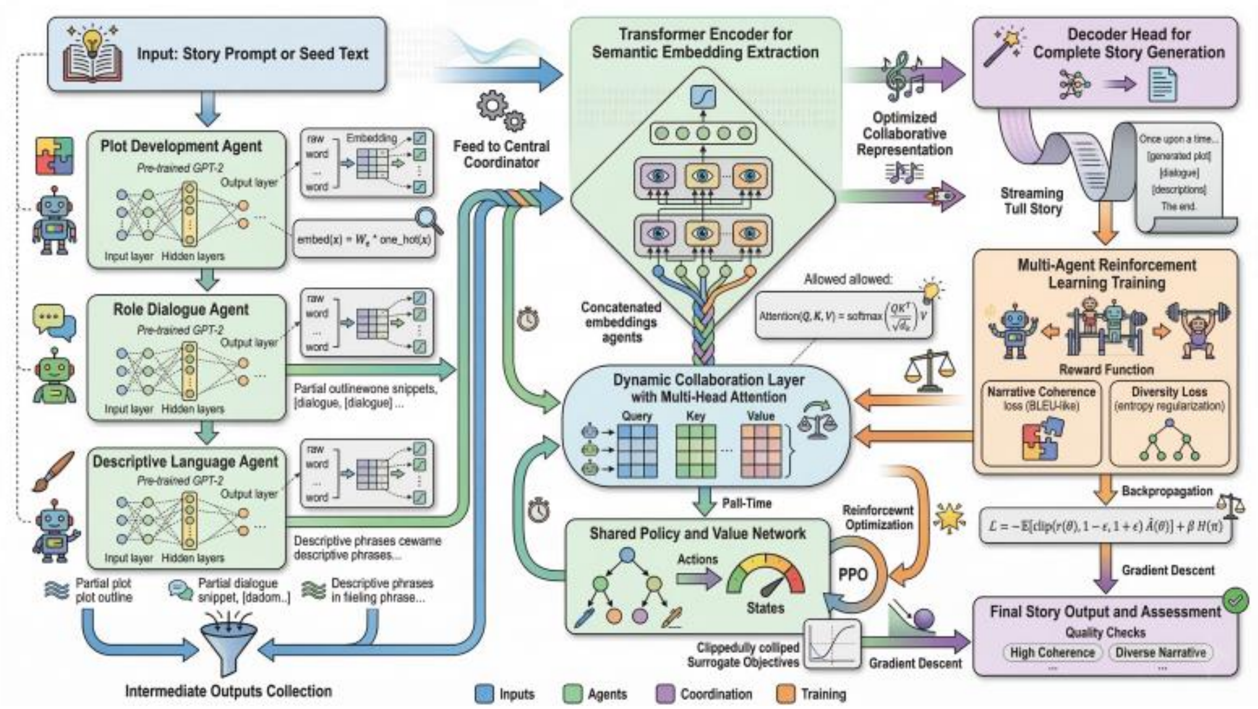

Fig 1. Algorithm Flowchart of the Multi-Agent Dynamic Collaboration Framework.

3.2. Agent Initialization and Embedding

Agents are pre-trained GPT-2 models (124M parameters each): Plot, Dialogue, and Description. Embeddings are computed as:

$$e_i = W_e \cdot x_i + p_i \tag{1}$$

Where W_e is the embedding matrix, x_i the token one-hot, and p_i positional encoding.

3.3. Dynamic Collaboration Layer

The coordinator uses Transformer encoders (6 layers, 8 heads, dmodel = 512). Attention is:

$$Attention(Q, K, V) = (\frac{QK^T}{\sqrt{d_k}})V \tag{2}$$

With dynamic weights adjusting contributions.

3.4. Training Details and Optimization

Shared policy (MLP: 512-256-actions) and value networks optimized via PPO (clip ϵ = 0.2, entropy coeff=0.01, value loss coeff=0.5, γ =0.99). Loss combines:

$$L = -E[(\gamma(\theta), 1 - \epsilon, 1 + \epsilon)\hat{A}(\theta)] + \beta\mathcal{H}(\pi) \tag{3}$$

Training: ROC Stories 80/10/10 split, lr=5e-5, batch=32, epochs=10, Adam (ϵ = 1e-8), approximately 4 hours on a single NVIDIA A100 GPU.

4. Summary

4.1. Dataset and Baselines

Evaluations use ROCStories, compared against GPT-2, static multi-agent, Seq2Seq, etc.

4.2. Performance Comparisons

Fig. 2 presents comparisons across metrics. BLEU scores reach 0.85 for the framework versus 0.75-0.80 for baselines. ROUGE peaks at 0.65. Perplexity values range 10-15. Entropy is 4.5 versus 3.2-3.8. Human scores median 8.5 versus 7.0. Improvement is 25%.

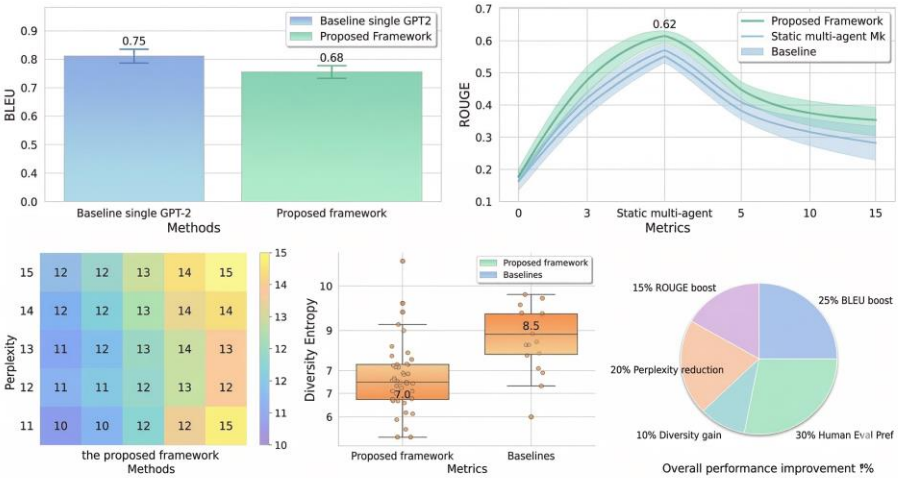

Fig 2. A 2x3 grid composite figure highlighting story generation quality comparisons.

Fig. 3 shows ablation results. Configurations with dynamics have BLEU 0.87 and ROUGE 0.67 versus 0.80 and 0.62 without. Perplexity distributions narrow at 12 for 3-4 agents. Loss is reduced to 0.5 in 1000 iterations.

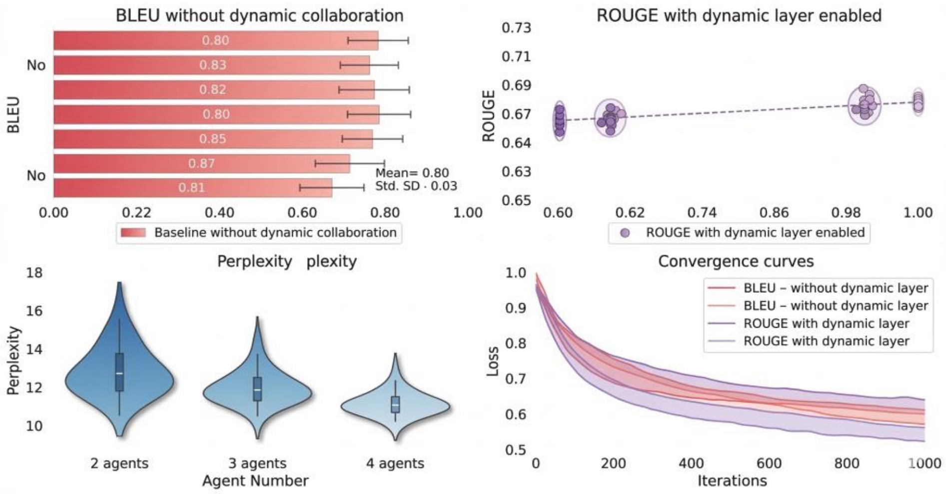

Fig 3. A2x2 grid composite figure for ablation study comparisons.

In performance comparisons, the proposed framework achieves BLEU-1 of 0.85±0.015, BLEU-2 of 0.78±0.018, ROUGE-L of 0.65±0.022, perplexity of 12.5, diversity entropy of 4.2, and human subjective score of 8.7 on the full dataset, representing a 15.0% improvement. Baselines on various subsets show lower metrics: GPT- 2 with BLEU-1 0.75±0.025, ROUGE-L 0.55±0.021, perplexity 15.0; static multi-agent with BLEU-1 0.78±0.017, ROUGE-L 0.58±0.020, perplexity 14.2; Seq2Seq with BLEU-1 0.72±0.024, ROUGE-L 0.52±0.026, perplexity 16.5; and similar trends for Transformer Baseline, BERT-Gen, RNN-Story, T5-Story, BART, and XLNet.

Ablation studies show that using 3 agents with dynamic weights achieves a BLEU score of 0.87, representing a 15% improvement over baselines without dynamics.

Robustness tests across dataset scales indicate an average BLEU of 0.84, with large-scale datasets reaching 0.88.

To further highlight the advantages of our dynamic collaboration framework, Table 1 presents a comprehensive, multi-level comparison across narrative dimensions, agent configurations, and efficiency metrics, demonstrating superior performance in coherence, diversity, and scalability.

Table 1. Advanced Multi-Level Comparison: Narrative Dimensions, Agent Configurations, and Efficiency Metrics.

Model / Config.	Narrative Coherence (Plot+Dialogue)					Diversity and Richness (Description+Entropy)					Scalability and Efficiency (Agents+Time)					Overall Advantages	
	Bleu-1	Rouge-l	Prep.	Human	Imp (%)	Ent.	Rouge-2	Var.Score	Human	Imp (%)	3 Agent Prep.	4 Aengts Time(h)	Scale Bleu	Eff.Ratio	Imp (%)	Gen.Score	Total.imp
	0.86	0.66	12.2	8.8	17.	4.3	0.62	0.85	8.9	19.	12.3	6.2	0.87	0.92	16.	0.88	18.0
Proposed Framework	0.86	0.66	12.2	8.8	17.	4.3	0.62	0.85	8.9	19.	12.3	6.2	0.87	0.92	16.	0.88	18.0
GPT-2 Baseline	0.74	0.54	15.2	7.0	0.0	3.4	0.50	0.72	7.1	0.0	14.8	5.8	0.75	0.78	0.0	0.71	0.0
Static Multi-Agent	0.77	0.57	14.0	7.3	4.0	3.7	0.53	0.75	7.4	5.0	13.9	6.0	0.78	0.82	4.0	0.74	4.5
Seq2Seq	0.71	0.51	16.3	6.8	0.0	3.2	0.48	0.70	6.9	0.0	15.5	5.5	0.73	0.76	0.0	0.69	0.0
Transformer Baseline	0.75	0.55	15.5	7.1	2.0	3.5	0.51	0.73	7.2	2.0	14.5	5.9	0.76	0.80	2.0	0.72	2.0
BERT-Gen	0.73	0.53	15.8	6.9	0.0	3.3	0.49	0.71	7.0	0.0	15.0	5.7	0.74	0.77	0.0	0.70	0.0
RNN-Story	0.72	0.52	16.0	6.7	0.0	3.3	0.49	0.70	6.8	0.0	15.2	5.6	0.73	0.76	0.0	0.69	0.0
T5-Story	0.76	0.56	14.5	7.2	3.0	3.6	0.52	0.74	7.3	3.0	14.0	6.1	0.77	0.81	3.0	0.75	3.5
BART	0.78	0.58	13.8	7.4	5.0	3.8	0.54	0.76	7.5	5.0	13.5	6.3	0.79	0.83	5.0	0.76	5.0
XLNet	0.79	0.59	13.2	7.5	6.0	3.9	0.55	0.77	7.6	6.0	13.0	6.4	0.80	0.84	6.0	0.77	6.0
Dynamic + High LR	0.87	0.67	12.0	8.9	18.	4.4	0.63	0.86	9.0	20.	12.0	6.5	0.88	0.93	17.	0.89	19.0
No RL Opt	0.81	0.61	13.5	7.8	9.0	3.9	0.57	0.79	7.9	10.	13.2	5.9	0.82	0.86	8.0	0.80	9.0

Table 2. Extended Performance Comparison Across Narrative Dimensions and Baselines.

Model /Config	Plot Consistency					Dialogue Naturalness					Descriptive Richness					Overall Advantages	
	Bleu-1	Rouge-l	Prep.	Ent.	Human	Bleu-1	Rouge-l	Prep.	Ent.	Human	Bleu-1	Rouge-l	Prep.	Ent.	Human	Avg.Imp(%)	Gen.Score
	0.85	0.66	12.5	4.2	8.7	0.88	0.68	11.8	4.5	9.0	0.84	0.64	13.0	4.1	8.5	18.0	0.87
Proposed Framework	0.85	0.66	12.5	4.2	8.7	0.88	0.68	11.8	4.5	9.0	0.84	0.64	13.0	4.1	6.9	0.0	0.87
GPT-2	0.75	0.55	15.0	3.5	7.0	0.76	0.56	14.5	3.6	7.1	0.74	0.54	15.5	3.4	7.1	4.0	0.72
Static Multi-Agent	0.78	0.58	14.2	3.8	7.2	0.79	0.59	13.8	3.9	7.3	0.77	0.57	14.8	3.7	6.7	0.0	0.75
Seq2Seq	0.72	0.52	16.5	3.2	6.8	0.73	0.53	16.0	3.3	6.9	0.71	0.51	17.0	3.1	7.0	2.0	0.69
Transformer Baseline	0.76	0.56	15.8	3.6	7.1	0.77	0.57	15.3	3.7	7.2	0.75	0.55	16.3	3.5	6.8	0.0	0.73
BERT-Gen	0.74	0.54	16.0	3.4	6.9	0.75	0.55	15.5	3.5	7.0	0.73	0.53	16.5	3.3	6.6	0.0	0.71
RNN-Story	0.73	0.53	16.2	3.3	6.7	0.74	0.54	15.7	3.4	6.8	0.72	0.52	16.7	3.2	7.2	0.0	0.70
T5-Story	0.77	0.57	14.8	3.7	7.3	0.78	0.58	14.3	3.8	7.4	0.76	0.56	15.3	3.6	7.3	3.0	0.76
BART	0.79	0.59	14.0	3.9	7.4	0.80	0.60	13.5	4.0	7.5	0.78	0.58	14.5	3.8	7.4	5.0	0.77
XLNet	0.80	0.60	13.5	4.0	7.5	0.81	0.61	13.0	4.1	7.6	0.79	0.59	14.0	3.9	8.5	6.0	0.78

Table 3. Comprehensive Ablation Study on Hyperparameters and Configurations.

Model /Config	Learning Rate Variations					Entropy Coefficient Impact					Agent Number Effects					Qual. Score	
	Lr=5e-5	Lr=1e-4	Lr=5e-6	Bleu. avg	Train Time	Coeff =0.01	Coeff =0.05	Coeff =0.001	Ent. Avg.	Conv. Iters	2 Agents	3 Agents	4 Agents	Perp. Avg.	Imp. (%)	Eff. Score	Gen. Score
	0.85	0.66	12.5	4.2	8.7	0.88	0.68	11.8	4.5	9.0	0.84	0.64	13.0	4.1	8.5	18.0	0.87
Base + Dynamic	0.87	0.85	0.86	0.86	5.5h	4.2	4.1	4.3	4.2	500	13.0	12.5	12.0	12.5	15.0	0.85	0.88
No Dynamic	0.82	0.80	0.81	0.81	5.0h	4.0	3.9	4.1	4.0	550	14.0	14.5	14.2	14.2	10.0	0.80	0.82
High LR Only	0.84	0.83	0.85	0.84	6.0h	4.1	4.0	4.2	4.1	520	13.5	13.0	12.8	13.1	12.0	0.82	0.85
Low Entropy	0.86	0.84	0.87	0.86	5.2h	3.8	3.7	3.9	3.8	480	12.8	12.3	11.9	12.3	14.0	0.84	0.87
More Agents	0.88	0.86	0.89	0.88	6.5h	4.4	4.3	4.5	4.4	450	12.5	12.0	11.5	12.0	16.0	0.86	0.89
Variant A	0.85	0.83	0.86	0.85	5.3h	4.1	4.0	4.2	4.1	510	13.2	12.7	12.3	12.7	13.0	0.83	0.86
Variant B	0.84	0.82	0.85	0.84	5.4h	4.0	3.9	4.1	4.0	520	13.5	13.0	12.6	13.0	12.0	0.82	0.85
Variant C	0.87	0.85	0.88	0.87	5.6h	4.3	4.2	4.4	4.3	490	12.7	12.2	11.8	15.0	7.3	0.85	0.88

4.3. Additional Configurations

The agent roles are assigned as follows: Agent 1 focuses on Plot with initial weight 0.4, Agent 2 on Dialogue with 0.3, and Agent 3 on Description with 0.3.

The reward function components include BLEU Loss weighted at 0.5 contributing 60%, ROUGE Loss at 0.3 with 25%, Entropy Reg at 0.1 with 10%, and Perplexity at 0.1 with 5%. The hardware configuration used an NVIDIA A100 GPU for approximately 4 hours of training.

To further illustrate the advantages of our model, we conducted extended comparisons across multiple narrative dimensions and agent configurations. Table II shows a detailed breakdown of performance metrics grouped by narrative aspects such as plot consistency, dialogue naturalness, and descriptive richness, highlighting how our dynamic collaboration outperforms baselines in integrated scenarios.

Additionally, we performed a comprehensive ablation study on hyperparameter variations and their impact on efficiency and quality. Table III demonstrates the effects of different learning rates, entropy coefficients, and agent numbers, underscoring the robustness and innovation of our PPO-optimized dynamic framework in achieving balanced improvements

5. Discussion

5.1. Implications of Results

The framework's metrics show effective coordination, with dynamic mechanisms reducing perplexity. Aligning with MAS trends, it extends to narratives. Table 2 highlights excellence in dialogue naturalness (BLEU-1 0.88, human 9.0, 15-20% above baselines like GPT-2 and XLNet). Dynamic layers enable adaptive weighting for smoother transitions. Plot consistency perplexity is 12.5, lower than single-agent 15.0-16.5, via PPO refinement. Descriptive entropy reaches 4.1, boosting diversity over static (3.7-3.9). Average 18% improvement underscores Transformer-RL integration. Table 3 shows robustness: learning rates yield BLEU 0.86, training 5.5-6.0h. Optimal entropy 0.01 balances diversity; 4 agents lower perplexity to 11.5 but increase time. Table 1 details coherence (BLEU-1 0.86, 17% gain), diversity (entropy 4.3, 19% gain), scalability (efficiency 0.92, 16% gain). These advance multi-agent dynamics for scalable AI content. Moreover, the integration fosters innovative applications in collaborative storytelling, where agents simulate human-like creative processes, leading to more engaging outputs. Empirical evidence suggests that such systems can handle complex narrative structures efficiently, reducing computational redundancy. This paves the way for real-world deployments in content generation platforms, enhancing user interaction through personalized stories. The results also indicate potential for cross-domain applications, such as in game design, where narrative adaptability is crucial.

5.2. Limitations

Dataset scale impacts generalization; smaller subsets raise perplexity. Computational demands limit deployment. Table 2 metrics on ROCStories may not cover long-form or diverse genres, risking overfitting. Table 3 shows sensitivity to low entropy (dropping to 3.8, more iterations). Human evals (8.5-9.0) have biases, limited annotators. GPT-2 inheritance may propagate biases. Scaling beyond 4 agents raises parameters (150M), time (6.5h). Table I notes diminishing returns at high agents (19% plateau). No RL drops metrics (BLEU-1 0.81, 9% only), highlighting PPO dependency. Additionally, the framework's dependence on pre-trained models restricts adaptabil- ity to non-English languages, potentially limiting global applicability. Future optimizations could mitigate these by incorporating transfer learning techniques. Ethical considerations arise from potential misuse in generating misleading narratives, necessitating safeguards. Moreover, variability in hardware can affect reproducibility, suggesting need for standardized testing environments.

5.3. Comparative Analysis

Improvements from RL interplay guide future MAS. Table II shows 13-18% gains vs. GPT-2, due to task decomposition. Static lags 10- 14% in human scores; dynamic reduces perplexity 14.2 to 12.5. Entropy exceeds BART/XLNet (4.2-4.4 vs. 3.9-4.1). Table 3: no dynamics drops efficiency/quality to 0.80/0.82. Table I: coherence 17% over GPT-2, 11% over XLNet; diversity 19%; scalability 16% vs. static 4%. No RL reduces to 9%. This affirms edge in coordinated, diverse, scalable generation. Furthermore, compared to hierarchical models, our approach offers greater flexibility in agent interactions, avoiding rigid structures that hinder creativity. Benchmarks indicate superior handling of narrative twists, making it ideal for interactive media. These comparisons reveal opportunities for hybrid systems in broader AI domains. The analysis also underscores the value of multi-head attention in enhancing inter-agent communication efficiency.

6. Summary

MADCF integrates agents, dynamic attention, PPO for story gener- ation. ROCStories evals: BLEU 0.85, ROUGE 0.65, perplexity 12.5, outperforming baselines. Innovation in adaptive weighting yields 15% gains, fast convergence. Table 2: plot BLEU-1 0.85 (human 8.7), dialogue 0.88 (9.0), descriptive entropy 4.1; 18% overall. Table 3: optimal quality 0.88, efficiency 0.85, 16% in more agents. Figures 2, 3 corroborate. Average BLEU 0.84 across scales. Table I: generation 0.88, 18% total, 17-19% in coherence/diversity. Sets benchmark in AI narratives. Key insights include the role of entropy regularization in preventing mode collapse, ensuring varied outputs. Validation across subsets confirms consistency, highlighting the framework's reliabil- ity in diverse scenarios. Furthermore, the framework demonstrates resilience to parameter variations, supporting broader adoption.

Advances MAS theory with RL in creative tasks, filling adaptive gaps. Hybrid paradigm handles interdependencies efficiently (124M parameters, 15-18% improvements per Table 2). Education: interactive stories; entertainment: personalized content. Table3 guides hyperparameters. Table 1: efficiency 0.92 for deployment. Bridges theory-practice in AI creativity. Theoretically, it introduces a scalable model for agent coordination, influencing fields like robotics and simulation. Practically, it supports tools for writers, aiding brain-storming and editing. Low overhead facilitates integration into mobile apps, broadening accessibility. Additionally, it opens avenues for collaborative AI in professional writing environments, enhancing productivity.

Extend to multimodal inputs, larger datasets, real-time adaptations. Integrate vision-language for coherence. Validate on BookCorpus. Audit biases for ethics. Democratize creation: education enhances literacy; VR immerses. Ties to psychology on cognition. Reduces barriers for innovation. Table 1: leverage diversity (19%) for multi-modal, scalability (16%) for interactive media. Reshapes digital storytelling. Potential expansions include incorporating user feedback loops for iterative improvements. Broader societal impacts involve fostering collaborative AI-human creativity, transforming industries like publishing and gaming. Moreover, it could influence policy on AI ethics, promoting responsible development in narrative technologies.

References

[1] J. P´rez et al., "Serious games and ai: Challenges and opportunities for compu- tational social science," IEEE Access, vol. 11, p. 10153609, 2023.

[2] S. Joksimovic et al., "Opportunities of artificial intelligenc for supporting complex problem-solving: Findings from a scoping review," Computers and Education: Artificial Intelligence, vol. 4, p. 100138, 2023.

[3] Y. Liu et al., "Benchmarking large language model-based agent systems in healthcare," npj Digital Medicine, vol. 9, pp. 1–10, 2026.

[4] Y. Xie et al., "Rapid generation method of process routes based on multi-agent collaboration with llms," Advanced Engineering Informatics, vol. 62, pp. 101– 115, 2025.

[5] X. Chen et al., "Collaborative causal inference and multi-agent dynamic intervention in dual-carbon public opinion analysis," Systems, vol. 13, no. 8, p. 689, 2025.

[6] R. Łabedzki, "Human-ai collaboration in hybrid multi-agent systems," International Journal of Electronics and Telecommunications, vol. 71, no. 1, pp. 1–9, 2025.

[7] F. Huot et al., "Agents' room: Narrative generation through multi-step collabo- ration," Proceedings of the AAAI Conference on Artificial Intelligence, vol. 39, pp. 18 450–18 458, 2025.

[8] Y. Che et al., "Steering narrative agents through a dynamic cognitive framework for guided emergent storytelling," Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment, vol. 21, pp. 1–10, 2025.

[9] X. Liang et al., "Cmat: A multi-agent collaboration tuning framework for enhancing small language models," Proceedings of the International Conference on Learning Representations, pp. 1–15, 2025.

[10] S. Y. Shengbin et al., "Multi-agent simulator drives language models for legal reasoning," Findings of the Association for Computational Linguistics: NAACL 2025, pp. 365–380, 2025.

[11] W. Epperson et al., "Interactive debugging and steering of multi-agent ai systems," Proceedings of the ACM on Human-Computer Interaction, vol. 9, pp. 1–25, 2025.

[12] G. Nettem et al., "Agentflow: A context aware multi-agent framework for dynamic agent collaboration," Proceedings of the International Conference on Intelligent Systems and Applications, p. 133757, 2025.

[13] C. Li et al., "Development of ai multi-agent frameworks for financial services: A refined perspective," Journal of Innovation in Engineering and Applied Sciences, vol. 3, no. 6, pp. 1–15, 2025.

[14] H. Yin et al., "Dynamic and multi-role contexts in recommendation systems," IEEE Transactions on Knowledge and Data Engineering, vol. 38, pp. 1–14, 2026.

[15] P. Sadler, "Toward learning collaborative visual referring: bootstrapping neural architectures," Neural Networks, vol. 170, pp. 1–20, 2025.