

NARRATIVE AI • RESEARCH NOTE • DECEMBER 2025

Maintaining Visual and Character Consistency in AI-Generated Multi-Chapter Storybooks

Kites Narrative AI Group • narrative@kites.dev

Abstract: Automated generation of long-form illustrated storybooks frequently suffers from character identity drift, pacing degradation, and plot incoherence across chapters. We present a hybrid attention architecture that extracts invariant visual features from initial artwork into persistent prompt vectors. Evaluated on 10,000+ illustrated story generations in Pictales, our method improves multi-chapter character retention by 68% compared to unconditioned autoregressive baselines.

1. Background & Motivation

While text-to-image and large language models excel at generating single isolated scenes, creating multi-chapter children's books demands strict visual and emotional continuity. When a young reader turns a page, the protagonist's clothing, eye color, companion animals, and expressive traits must remain constant despite changing environments and poses.

Pictales addresses this by combining localized visual embeddings with memory-augmented language prompting.

2. Visual Memory Embeddings

Our system processes input character drawings through an invariant visual feature extractor. Key anatomical and stylistic attributes are encoded into persistent cross-attention key-value vectors that condition all subsequent chapter narrative prompts and image generation seeds.

3. Empirical Evaluation

In blind tests across 500 parents, educators, and children, stories generated with Pictales visual memory scored 4.7/5.0 on character recognition consistency, compared to 2.1/5.0 for standard prompt-chaining workflows.

4. Conclusion

By decoupling visual identity from situational context, Pictales enables creators, parents, and authors to produce cohesive, publication-ready storybooks from single concept sketches.