Proceedings of the 3rd International Conference on Machine Learning and Automation
DOI: 10.54254/2755-2721/2025.28119

Algorithmic Storytelling and Cinematic Narrative: a Comparative Study of Ai-generated Screenplays and Contemporary Auteur Cinema

Danqing Liu[1], Yuanchen Li[2*]

[1]The Hang Seng University of Hong Kong, Hong Kong, China
[2]Sapienza University of Roma, Rome, Italy
**Corresponding Author. Email: rara481846778@gmail.com*

Abstract. With the advancement of generative artificial intelligence, AI-based text generation has been increasingly applied to the domain of screenplay writing, raising critical questions about whether algorithmic storytelling can embody literariness, cultural expression, and philosophical depth. This study uses Life of Pi as a case and constructs two AI-generated screenplay samples (theme-driven and adaptation-driven) to compare systematically with Ang Lee's directorial version. Methods include narrative structure modeling, thematic weight analysis, symbolic language density computation, and philosophical abstraction measurement. A multidimensional comparison across narrative coherence, thematic focus, linguistic tension, and cultural depth is conducted, complemented by blind expert interviews involving five specialists to evaluate literary expressiveness from a humanistic perspective. The results show that while AI scripts perform well in structural control and thematic identification, they lag behind auteur-driven screenplays in philosophical abstraction, symbolic system construction, and aesthetic articulation. The study concludes that current AI systems are not yet capable of independently producing screenplays with humanistic depth but can function as effective tools in generating genre-oriented drafts.

Keywords: AI screenplay generation, Auteur cinema, Literariness, Textual comparison, Narrative construction

1. Introduction

As generative AI technologies are progressing at a fast-paced rate, algorithmic text generation has increasingly been incorporated into the screenwriting practice. Recent models such as GPT-4 and ScriptWriter API are shown to create coherent narratives and genre-centered plot lines, presenting screenwriters with new means of creating drafts and conceptualizing story structure [1]. There has, though, been a persistent challenge, AI-generated screenplays tend to lag behind in expressing literary quality, intellectual depth, and cultural subtlety—essential dimensions traditionally maintained by author-centric cinema. In spite of past research having focused predominantly on the logic of structure and generative parsimony, scant consideration has accrued to the degree to which

DOI: 10.54254/2755-2721/2025.28119

AI may mimic aesthetic and symbolical value embedded in human-created film narratives [2]. In this study, Life of Pi serves as a case to explore these tensions, creating two AI-driven screenwriting versions (thematically-oriented and adaptation-oriented) and comparing them systematically with Ang Lee's directing vision via semantic modeling, narrative cartography, and symbolical language detection. The ambition involves understanding the limitations and potential of literary expressiveness embedded in the prevailing frameworks of algorithmic storytelling.

2. Literature review

2.1. Evolution of AI narrative systems

More recent innovations in AI narrative systems employing large language models have greatly improved machine competency in creating plots, building characters, and linguistic flow. Programs such as GPT-4 now effectively mimic the three-act structure and generate dialogue sequences consonant with genre conventions [3]. Most of these systems, however, are fine-tuned to optimize structural competency and probabilistic fluency more than thematic resonance or symbolic articulation. Most of the literature emphasizes enhancing narrative coherence and genre conformity but neglects whether AI may create narratives projecting complex human experience or cultural specificity [4]. These systems then end up recreating familiar narrative shapes instead of creating original aesthetic configurations and thus suffer semantic flattening when confronted with abstract, philosophical, or metaphorical materials. As such, literary assessment of AI storytelling demands analytical instruments more sophisticated even than structural verification, and these are particularly needed in measuring symbolic compactness and hermeneutic penetration.

2.2. Auteur cinema: signature, philosophy, and cultural mediation

Auteur theory places the director at the forefront of the primary creative mind whose stylistic and thematic consistency characterizes the narrative structure of the film. Directors like Ang Lee epitomize this model, infusing works with personal philosophy and trans-cultural identity. In Life of Pi, Lee interweaves existential inquiry with mythic narrative visualization, generating a story narrative that goes beyond functional plot devices [5]. Auteur films tend to resist genre expectations by celebrating nonlinear narratives, symbolical indeterminacy, and open-readability qualities hard to replicate with AI. Comparative studies demonstrate that though AI replicates logic coherence and pattern structure, it cannot replicate the symbolic denseness and cultural specificity of auteur-driven narratives [6]. What AI cannot provide is the "creative deviation" mimicking lived reality, emotive complexity, and ideological positioning. This shortcoming highlights an abiding limitation in using AI at the high level of literary or philosophical narrative storytelling.

2.3. Literariness in screen narrative

Literariness of screenwriting extends aesthetic beauty to engage abstraction, symbolic depth, and multiple-reading complexity. Informed by Jakobson's poetic function, Barthes's writerly texts, and semiotic narrative theory, literariness as a register combines linguistic form, cultural context, and reader response [7]. In film, it manifests in the form of metaphor, symbolic interlayering, structural indeterminacy, and reflexivity. State-of-the-art AI systems, overwhelmingly oriented toward statistical pattern recognition, produce surface-level coherence but are unlikely to generate texts with semantic flexibility and symbolic tension. Unable to register cultural abstraction or contextual complexity, AI narratives invariably rely on literal representation and anticipated meaning and close

DOI: 10.54254/2755-2721/2025.28119

their multiple-reading openness [8]. Consequently, evaluating AI's literary potential entails more than syntactic or structural analysis; indeed, semiotic and cultural-pragmatic analysis capable of judging if AI texts register multivalent reading and affective resonance comparable to human-generated literary narratives is needed.

3. Methodology

3.1. Sample and script generation

In this study, Life of Pi by Ang Lee is selected as the core case for comparative screenplay generation. A dual-path AI generation framework is designed. On one hand, a theme-driven script is created using GPT-4, prompted with semantic anchors such as "faith," "survival," "drift," and "illusion." Structural markers and emotion tags guide the model to produce a three-act format screenplay, approximately 6000 words in length [9]. On the other hand, an adaptation-driven script is generated by segmenting the source novel and inputting it into GPT-4, assisted by LangChain-based summarization and a custom event extraction module to identify key narrative elements. The system is fine-tuned using screenplay corpora from IMDb and Academy Award-winning films to ensure stylistic fidelity to cinematic language. The output includes three screenplay types, AI-original, AI-adapted, and the original director's version, forming the dataset for subsequent textual analysis, semantic comparison, and expert interviews.

3.2. Analytical dimensions and evaluation metrics

To evaluate the narrative and literary quality of AI-generated versus director-authored screenplays, five quantifiable dimensions are modeled using computational linguistic techniques [10].

First, Narrative Coherence (NC) is measured by event alignment ratio:

$$NC=\frac{N_{aligned}}{N_{total}} \tag{1}$$

Where $N_{aligned}$ refers to the number of events in the text that successfully match standard narrative structure elements (such as conflict, turning point, and climax), and N_{total} represents the total number of narrative events in the text.

Second, Character Depth (CD) evaluates psychological complexity via internal conflict density:

$$CD=\frac{\sum_{i=1}^{n} C_i}{n} \tag{2}$$

Where n is the number of main characters in the text and C_i denotes the frequency of internal conflict, emotional fluctuation, or self-contradiction presented by character i in the screenplay.

Third, Thematic Density (TD) employs a weighted TF-IDF formula to assess thematic focus:

$$TD_k=\frac{tf_k \cdot \log\left(\frac{N}{df_k+1}\right)}{L} \tag{3}$$

Where tf_k indicates the term frequency of the kth thematic keyword such as faith or survival in the text, N is the total number of documents in the corpus, df_k is the number of documents containing that keyword, and L is the total number of words in the text.

DOI: 10.54254/2755-2721/2025.28119

Fourth, Symbolic Language Density (SLD) is calculated as symbolic phrase frequency per sentence:

$$SLD=\frac{N_{symbolic}}{N_{sentences}} \tag{4}$$

Where $N_{symbolic}$ refers to the number of symbolic expressions identified through semantic annotation such as ocean symbolizing faith or tiger symbolizing inner fear and $N_{sentences}$ is the total number of sentences in the text.

Fifth, Philosophical Abstraction Tension (PAT) utilizes cosine distance in BERT embeddings:

$$PAT=\frac{1}{|P|}\sum_{p\in P}\text{cosine\_distance}\,(v_p,v_c) \tag{5}$$

Where P is the set of philosophical keywords such as truth existence and faith v_p is the embedding vector of the keyword in the BERT semantic space and v_c is the average contextual vector in which the word appears.

These formulas enable the structured comparison of symbolic density, semantic abstraction, and narrative architecture across the three screenplay types.

3.3. Expert interviews and qualitative assessment

This study invited five experts including screenwriters, literary critics, and narrative scholars to participate in expert interviews. A combination of semi-structured interviews and blind review was adopted, conducted through remote video conferencing. Each interview lasted 60 minutes. The purpose was to conduct qualitative validation from the perspectives of textual meaning, symbolic construction, and cultural depth. Each expert was randomly assigned one AI-original screenplay, one AI-adapted screenplay, and one director version for review. The interview content is shown in Table 1.

Table 1. Interview question table

Analytical Dimension	Guiding Question
Authorial Perspective	Can you perceive any distinct creative standpoint or value expression in the text?
Symbolism and Metaphor	Do you observe any symbolic or non-literal expressions in the text? Can you give examples?
Cultural or Philosophical Position	Does the text reflect any cultural identity, religious reflection, or philosophical theme?
Emotional and Aesthetic Expression	Does the text evoke any emotional reaction or contain a distinct linguistic aesthetic?

All expert data were transcribed and thematically coded with informed consent. Using NVivo 14 for semantic clustering, the study extracted shared and divergent views to build a humanistic evaluation framework for AI-generated screenplays. Results confirm that while AI scripts show technical coherence in narrative structure, they lack depth in symbolism, cultural positioning, and emotional aesthetics [11]. Experts noted a tendency toward functional rather than authorial language, with symbolic elements often template-based rather than contextually emergent. This qualitative analysis complements the quantitative findings and supports later discussion from a human-centered perspective.

DOI: 10.54254/2755-2721/2025.28119

4. Results

4.1. Narrative structure and thematic consistency

Through quantitative analysis of the three types of screenplays, the study finds that AI-generated scripts exhibit clear structural characteristics in terms of narrative coherence. As shown in Figure 1, the narrative coherence index of the theme-driven AI screenplay reaches 0.78, that of the adaptation-driven AI screenplay is 0.82, while the director's original version reaches 0.91. In the measurement of thematic density, AI-generated scripts demonstrate a high concentration of keyword frequencies. The theme-driven script shows TF-IDF weighted values of 0.156 and 0.143 for the core concepts of faith and survival respectively, significantly higher than the director's version values of 0.089 and 0.092. The data indicate that AI systems possess strong capabilities in maintaining narrative logical consistency and can effectively identify and reinforce core thematic elements. However, this high degree of thematic concentration also reflects a certain mechanical tendency in AI semantic expression, lacking the natural evolution of thematic variation and layered meaning commonly found in human-authored works. The adaptation-driven script performs slightly better in structural integrity than the theme-driven version, which suggests that AI generation based on original texts is more stable in maintaining narrative coherence.

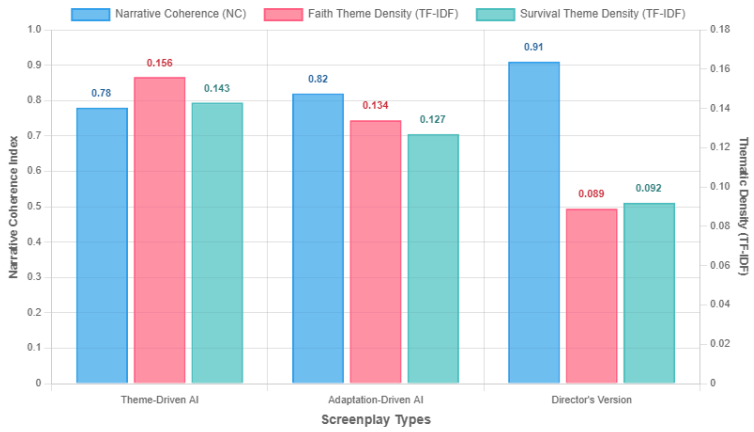

Figure 1. Comparative analysis of narrative coherence and thematic density

4.2. Philosophical and literary expression

In the measurement of symbolic language density and philosophical abstraction tension, the three types of screenplays present significant differences as shown in Table 2. The director's version reaches 0.31 symbolic expressions per sentence in symbolic language density, while the AI-generated scripts contain only 0.18 in the theme-driven version and 0.22 in the adaptation-driven version. More critically, in the BERT embedding distance analysis of philosophical abstraction tension, AI screenplays show lower semantic complexity, with an average cosine distance of 0.67, which is substantially lower than the director's version at 0.89.

DOI: 10.54254/2755-2721/2025.28119

Table 2.

Screenplay Type	Symbolic Language Analysis		Philosophical Abstraction		Expert Assessment	
	SLD (per sentence)	Metaphor Frequency	PAT (BERT distance)	Abstract Concepts	Cultural Depth (1-5 scale)	Aesthetic Quality (1-5 scale)
Theme-Driven AI	0.18	23	0.67	42	2.1	2.3
Adaptation-Driven AI	0.22	31	0.71	38	2.4	2.6
Director's Version (Ang Lee)	0.31	67	0.89	78	4.2	4.5

The expert interview results further validate the findings of the quantitative analysis. Four out of five experts believe that AI screenplays exhibit clear limitations in cultural positioning and emotional aesthetic expression. Although AI can identify and apply basic symbolic elements, these symbols often appear templated and lack organically generated contextual relevance. In terms of philosophical depth, AI screenplays tend to use direct statements rather than metaphorical expressions. While this mode of expression ensures the accessibility of content, it diminishes the interpretive openness and aesthetic tension of the text. The research results indicate that current AI systems perform well in replicating surface structures but still face fundamental challenges in generating literary texts with profound cultural connotation and philosophical reflection.

5. Discussion

By means of comparative examination of three types of screenplays in structural logic, thematic articulation, symbolic regimes, and philosophical abstractions, this research unveils AI-generated script's technical potential in narrative compactness and structure management, and critical weakness in literariness and cultural articulation. AI programs display great consistency and templating strengths in recognizing themes and narrative coherence, best utilized in replicating genre-oriented patterns. In doing so, however, comes semantic uniformity and symbolic compression, making the texts less qualified to accommodate ambiguity, cultural resonance, or deep emotion. Expert interview findings further highlight the reality that though AI script may simulate human scripting in appearance, they are passive and disjointed in creative purpose, linguistic tension, and worldview building. In summary, AI at this moment serves best as a structural aid rather than an independent literary creator.

6. Conclusion

Using Life of Pi as a case study, this paper constructed both theme-driven and adaptation-driven AI screenplays and compared them with the director's original version to systematically examine the boundaries and potential of generative AI in narrative construction and literary expression. The study finds that AI screenplays demonstrate a certain degree of technical maturity in structural clarity, pacing, and thematic focus, making them suitable for preliminary genre-based storytelling. However, in terms of symbolic system construction, cultural embedding, and philosophical semantic expression, AI-generated texts still exhibit noticeable flatness and mechanization, lacking aesthetic and intellectual depth. Qualitative interviews further confirm AI's limitations in authorial presence, emotional engagement, and linguistic aesthetics. Future research should focus on multimodal inputs, value-based modeling, and cross-cultural training to explore hybrid narrative frameworks that

DOI: 10.54254/2755-2721/2025.28119

integrate algorithmic generation with humanistic design, thereby enhancing AI's expressive capacity and ethical awareness in literary creation.

Contribution

Danqing Liu and Yuanchen Li contributed equally to this paper.

References

[1] Mirowski, Piotr, et al. "Co-writing screenplays and theatre scripts with language models: Evaluation by industry professionals." Proceedings of the 2023 CHI conference on human factors in computing systems. 2023.
[2] Lin, Zefeng, et al. "R $^ 2$: A LLM Based Novel-to-Screenplay Generation Framework with Causal Plot Graphs." arXiv preprint arXiv: 2503.15655 (2025).
[3] Tang, Yuying, et al. "Understanding Screenwriters' Practices, Attitudes, and Future Expectations in Human-AI Co-Creation." Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems. 2025.
[4] Xu, Yiren. "Balancing Creativity and Automation: The Influence of AI on Modern Film Production and Dissemination." arXiv preprint arXiv: 2504.19275 (2025).
[5] Elias, Silvia, et al. "Rethinking literary creativity in the digital age: a comparative study of human versus AI playwriting." Humanities and Social Sciences Communications 12.1 (2025): 1-11.
[6] Serena, Andrea. The Impact of Generative AI on the Film Industry: Opportunities, Challenges, and Ethical Implications. Diss. Politecnico di Torino, 2025.
[7] Prabowo, Bayu Ade, and Rahmanti Asmarani. "Generative Literature: The Role of Artificial Intelligence in the Creative Writing Process." Allure Journal 5.1 (2025): 1-9.
[8] Holopainen, Jussi, et al. "Infinity Book: Speculating Literary Expressions in the Age of Generative AI." Proceedings of the 2025 ACM Designing Interactive Systems Conference. 2025.
[9] Dayo, Fatima, Ahmed Ali Memon, and Nasrullah Dharejo. "Scriptwriting in the age of AI: Revolutionizing storytelling with artificial intelligence." Journal of Media & Communication 4.1 (2023): 24-38.
[10] Vainikka, Eliisa, Anne Soronen, and Saara-Maija Kallio. "Is prompt engineering the future of screenwriting? Views of professional screenwriters and commissioners about the impact of AI technologies on their profession." Critical Studies in Television 20.1 (2025): 54-72.
[11] Anantrasirichai, Nantheera, and David Bull. "Artificial intelligence in the creative industries: a review." Artificial intelligence review 55.1 (2022): 589-656.