AI-Assisted Automated Micro-film Script Generation And Analysis System

Dongxue Tan

Lecturer, School of Journalism and CommunicationShenyang Normal University, Shenyang 110000, China

Corresponding author. E-mail: dongxue01tan@outlook.com, cold000000@163.com

Received: February 26, 2026; Accepted: April 17, 2026

With the rise of short stories in the digital medium, there has been a growth of interest in developing intelligent, scalable systems that generate micro-film scripts. This paper introduces an Artificial Intelligence-assisted automated framework to generate micro-film scripts, and to analyze their narrative structure and content using a Fine-tuned GPT-2 Transformer Model. Movie scripts are pre-processed to normalize, tokenize by sub-words, and segment the text into fixed-length sequences to develop "screenplay" specific narrative patterns. The model is then further trained using the AdamW Optimiser with a learning rate of 5E-5, a batch size of 8 and training over 3 epochs; additionally, micro-film scripts will be generated using Nucleus Sampling (top-p = 0.9), and Temperature Scaling (temperature = 0.8). An extensive evaluation framework has been used to evaluate the outputs of the proposed framework via BLEU, ROUGE, Distinct-N, Semantic Similarity, Narrative Coherence, and Sentiment Consistency metrics. The experimental evaluations showed a proposed framework generated high semantic similarity (0.93), strong lexical diversity (Dist-2 = 0.83), and stable sentiment consistency (0.555) indicating a proposed framework generated fluent, creative scripts. However, global narrative coherence was evidenced to be only moderate (0.264); therefore, the proposed framework could provide a reproducible and scalable solution to allow for automated micro-film script creation and use in creative, educational and commercial industries.

Keywords: Micro-films; Short dramas; documentaries; And other film and television genres; AI-Assisted Script Generation

http://dx.doi.org/10.6180/jase.202610_33.020

1. Introduction

Short-form digital storytelling has transformed how narratives are created and consumed across entertainment, education, and social media [1]. Micro-films, typically only a few minutes long, require concise yet impactful scripts to preserve narrative flow and emotional engagement within limited time [2]. Typically ranging between 1 to 10 minutes, this duration imposes structural constraints requiring compact storytelling and efficient scene progression. This structural compression strengthens emotional resonance by concentrating plot development, character intent, and thematic cues into a highly focused cinematic span. It emphasizes focused character development and minimal dialogue redundancy. In AI-assisted frameworks, these constraints guide structurally optimized script generation. Effective micro-film writing demands careful handling of dialogue, scene transitions, and thematic consistency to maintain coherence [3]. With the rapid growth of micro-film content, there is an increasing need for intelligent systems that can efficiently generate high-quality scripts.

Automated micro-film script generation systems are gaining importance due to the rising demand for fast and scalable content production across digital platforms [4, 5]. Traditional scriptwriting is time-intensive, requires domain expertise, and often results in stylistic inconsistencies [6].

Additionally, the lack of structured screenplay datasets and the complexity of modeling character interactions, emotional progression, and thematic depth highlight the necessity for AI-driven narrative generation approaches [7, 8].

Earlier script-generation methods relied on rule-based or template-driven techniques, which limited creativity and flexibility [9, 10]. Although modern approaches such as recurrent neural networks improve fluency, they struggle with maintaining structured screenplay formats. Transformer-based models further enhance generation capabilities but often produce incoherent plots or inconsistent character dialogues without proper guidance [11, 12].

To address these challenges, an AI-assisted framework for automated micro-film script generation is introduced by integrating data preprocessing, fine-tuned transformer models, and prompt engineering. The framework enables scalable and efficient script creation across diverse narrative contexts with reduced manual effort and faster turnaround. By extracting scenes, dialogues, and character actions, the fine-tuned GPT-2 model generates coherent screenplay-style content, while the evaluation module ensures narrative coherence, dialogue quality, and thematic consistency through AI-driven structural analysis.

Contributions of the study,

- Unified AI Framework: AI-based system for automated micro-film script generation and integrated narrative analysis.
- Pre-Processing: Structured preprocessing with GPT-2 tokenization for screenplay-specific data preparation.
- Transformer-Based Generation: Fine-tuned GPT-2 with prompt design for coherent and structured script generation.
- Script Evaluation: Analysis module assessing coherence, dialogue quality, semantics, themes, and sentiment consistency.

Organization of the Paper

The paper is organized as follows: Section 2 reviews related work on micro-films, AI script generation, and narrative analysis. Section 3 presents the proposed methodology, including preprocessing, transformer-based script generation, and prompt design. Section 4 describes the script analysis system and evaluation metrics. Section 5 details the experimental setup, results, and comparisons. Section 6 concludes the study and outlines future research directions.

2. Literature review

Ge [13] highlights that optimizing AI algorithms improves narrative coherence, visual rhythm, and audience engagement in micro-films, supporting automated script generation. Fan [14] emphasizes the importance of concise storytelling, emotional impact, and strong thematic grounding in micro-film creation. Honglei Zhan [15] focuses on visual symbolism, narrative compression, and consistent aesthetic structure, essential for AI-based script modeling.

Shi and Li [16] identify the need for semantic and thematic analysis due to the shift toward concise, realistic, and audience-centric storytelling. Kong [17] demonstrates the role of cultural adaptability and global narrative structures in modern cinema, relevant for AI-generated scripts. Yuan [18] shows that emotional storytelling enhances brand engagement in micro-films, while Li [19] highlights the rise of fast-paced, episodic micro-narratives shaped by AI and platform analytics. Roshan et al. [20] evaluated cement stabilized residual soil at macro and micro scales, demonstrating improved strength and durability for road construction applications.

Jia and Mohamed [21] examine media convergence and digital transformation, highlighting the role of AI, data analytics, and automated systems in modern media production, including micro-film generation. Zschocke [22] proposes a multimethod approach combining spatial and narrative film analysis, which supports extending such concepts into AI-based computational frameworks.

2.1. Problems Statement

The rapid growth of short video platforms [23] has increased demand for high-quality, engaging, and well-structured storytelling [24]. However, existing studies lack systematic approaches for micro-film script development and fail to address challenges such as limited duration, episodic structure, and emotional continuity [25]. Traditional scriptwriting methods further limit scalability and consistency in modern digital production environments.

Recent research highlights the role of AI in content optimization, yet most works remain conceptual or focus on isolated aspects. This study proposes an integrated AI-based framework for automated micro-film script generation and evaluation, combining structured preprocessing, advanced language models, and multi-level analysis to ensure scalable and high-quality script production.

3. Proposed methodology

The process begins with a movie script dataset, which undergoes preprocessing to enable tokenization, sequence management, and encoding for transformer compatibility. The processed data is then used to fine-tune a GPT-2 model with optimized hyperparameters such as learning rate, batch size, and epochs to capture screenplay structure. The fine-tuned model generates micro-film scripts based on input prompts, preserving narrative attributes. Generated scripts are evaluated using metrics such as BLEU, ROUGE, cosine similarity, distinct-n scores, dialogue structure, and sentiment and thematic coherence to ensure both linguistic and narrative quality.

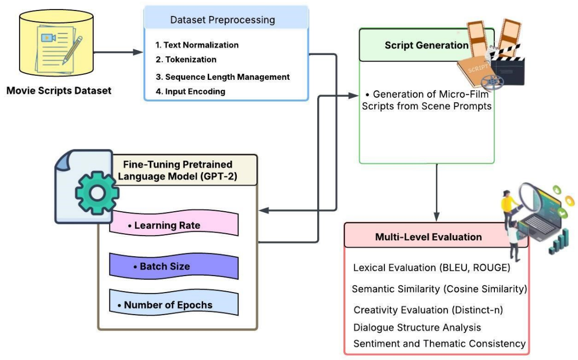

Fig. 1. Proposed Workflow for GPT-2–Based Micro-Film Script Generation.

The Fig. 1 workflow starts with preprocessing movie scripts through normalization, tokenization, sequence management, and encoding, which are converted into embeddings and fed into a fine-tuned GPT-2 model for narrative learning. The model generates micro-film scripts from structured prompts, and the outputs are evaluated using lexical, semantic, creativity, dialogue, and sentiment-based metrics to ensure overall quality.

3.1. Dataset Overview

This study utilizes a dataset of 1,172 English movie scripts sourced from Hugging Face, covering diverse genres, narrative styles, and dialogue structures. The corpus spans genres including drama, comedy, action, and thriller, with substantial variability in script length from concise dialogue-driven narratives to scene-rich full-length screenplays. The dataset is used in accordance with its licensing terms for research and non-commercial purposes only. No personal or sensitive information is included, ensuring ethical compliance and data privacy. These measures support transparent and responsible use of publicly available screenplay data. The dataset supports modeling complex storytelling patterns and long-range textual dependencies [26].

- Data Source and Accessibility
- The dataset is publicly available on Hugging Face under the Apache-2.0 open-source license, enabling unrestricted academic use and reproducibility. All scripts are sourced from publicly available screenplay repositories, ensuring transparency and ethical data collection.

Dataset Structure

Each record in the dataset contains two primary attributes:

- Movie Name: The official title of the movie.
- Script: The full screenplay text, including scene headings, dialogue, narrative descriptions, and formatting conventions commonly used in professional scripts.

This structured format enables effective extraction of scene prompts, dialogue blocks, and narrative transitions for downstream modeling tasks.

Dataset Splitting

The dataset is split into training (80%), validation (10%), and test (10%) subsets. Training uses only the training data, validation supports monitoring and hyperparameter tuning, and testing evaluates final performance and qualitative results. To ensure independence, scripts from the same movie are assigned to a single subset, improving generalization.

3.2. Data Preprocessing

Data preprocessing standardizes raw movie scripts into a uniform format using a script normalizer, removing unnecessary spaces while preserving punctuation. The processed data is tokenized using the GPT-2 subword tokenizer to handle diverse dialogue structures. Token sequences are truncated or split into overlapping windows to fit GPT-2 length limits, ensuring coherence. These sequences are converted into embeddings that capture semantic and contextual relationships, enabling stable training and high-quality script generation.

1. Text Normalization

   Text normalization ensures lexical consistency across scripts by removing noise such as inconsistent capitalization, special characters, and extra whitespaces. Specifically, screenplay text was converted into a uniform case format, redundant punctuation symbols were filtered, and excessive whitespace between scene descriptions and dialogue turns was normalized. Important punctuation is retained to preserve screenplay structure, reducing vocabulary sparsity and enabling stable GPT-2 training. This is expressed in Eq. (1) as:

$$S' = \mathcal{N}(S) \tag{1}$$

where S represents the raw script text and $\mathcal{N}(\cdot)$ denotes normalization operations.

2. Tokenization

A GPT-2 based subword tokenization procedure is applied to normalized scripts, enabling efficient handling of frequent and rare tokens. This compact segmentation further enhances dialogue structure representation by preserving speaker turns, contextual phrasing, and screenplay-specific lexical patterns. The GPT-2 subword vocabulary constructs token sequences by preserving frequent lexical units while decomposing rare words, character names, and uncommon screenplay expressions into reusable subcomponents. It improves representation of character names and diverse dialogue while preserving contextual continuity across the screenplay. It is applied as shown in Eq. (2):

$$T = \mathcal{T}\left(S'\right) = \{t_1, t_2, \ldots, t_M\} \tag{2}$$

where T denotes the generated token sequence and M is the total number of tokens.

3. Sequence Length Management

Lengthy scripts are split into overlapping or cut off token sequences so that they fit into the fixed input length of GPT-2. A 512-token sliding window with a stride of 256 tokens was employed, creating a 50% overlap between adjacent segments to preserve local narrative continuity. This is done to minimize data loss while keeping the training data contextually relevant is defined as Eq. (3):

$$T_k = \{t_k, t_k, \ldots, t_{k+L}\} \tag{3}$$

where L is the maximum sequence length supported by the model.

4. Input Encoding

Mapping tokenized sequences to dense vector embeddings captures semantic and contextual relationships. Dense vector embeddings are the input representation of the GPT-2 model when being fine-tuned is expressed as Eq. (4),

$$X = \text{Embed}(T), X \in \mathbb{R}^{M \times d} \tag{4}$$

where d is the embedding dimension.

3.3. Script Generation Model Training

Fine-tuned GPT-2 transformers are used to generate coherent and context-rich micro-film scripts from preprocessed and tokenized screenplay data. The model adapts to screenplay structures such as scene transitions, character interactions, and dialogue flow. Nucleus Sampling (top-p = 0.9) is used during generation to select tokens from a probability distribution, filtering low-probability choices. This balances randomness and control in text generation. As a result, it improves creative diversity while maintaining narrative coherence. A Fig. 2 depicts a fine-tuning method where a pre-trained GPT-2 model is used along with domain-specific screenplay data to generate scripted micro-films automatically. The token and positional embedding vectors go through the stacked transformer blocks using a multi-head attention and feed-forward network to generate well-structured sentences comprising an entire script, including dialogue, scenes, and a narrative flow.

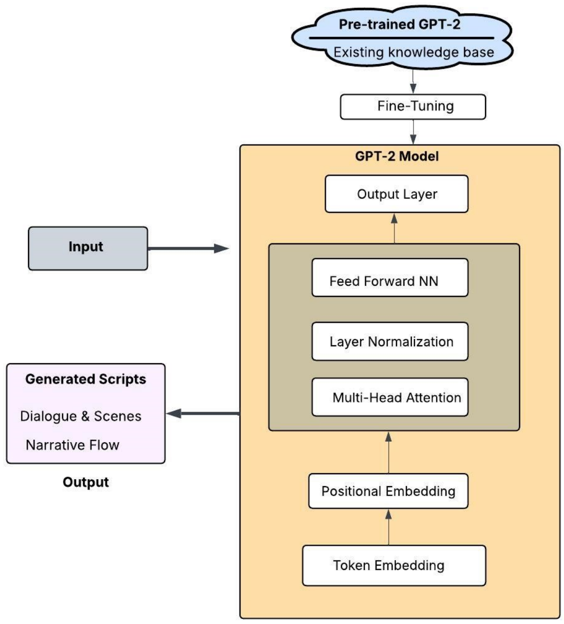

Fig. 2. Fine-Tuned GPT-2 Architecture for Automated Micro-Film Script Generation.

3.3.1. *Fine-Tuning Pretrained Language Model (GPT-2)*

The GPT-2 transformer-based model for microfilm script generation is based on an autoregressive model with strong abilities for modelling long-range contextual dependencies. The selection of GPT-2 is justified by its efficient fine-tuning capability, strong generative performance for structured text, and its suitability for screenplay-style narrative synthesis with moderate computational cost. The adaptation process was initialized from the pretrained GPT-2 checkpoint, preserving baseline transformer weights while updating all parameters through supervised screenplay-specific gradient optimization. The self-attention mechanism within the model can effectively track the progression within a narrative, the flow of dialogue between characters, and the sequences of events in time within the development of a coherent and logically-developed short cinematic script.

This is mathematically defined Eq. (5) as:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^\top}{\sqrt{d_k}}\right) V \tag{5}$$

where Q, K, V denote query, key, and value matrices, respectively, and d_k is the dimensionality of the key vectors.

1. Input Representation

   The scripts in the data set will be processed through a tokenizer that is based on sub words of GPT-2, which removes all explicit identifiers of scenes or characters so that the model can tell apart narrative text and dialogue easily. Each script sequence as depicted in Eq. (6):

   $$X = \{x_1, x_2, \ldots, x_T\} \tag{6}$$

   where x_t is the token at position t in the sequence.

2. Fine-Tuning Strategy

   Domain-specific micro-film scripts are used to fine-tune GPT-2, adapting its general language understanding to screenplay conventions such as dialogue, narration, and scene transitions. This enables coherent event progression, realistic dialogue, and consistent character continuity.

   Formally, given a tokenized script sequence $X = \{x_1, x_2, \ldots, x_T\}$, the model learns to predict each token conditioned on all previous tokens. This autoregressive learning objective can be defined as Eq. (7),

   $$P(X) = \prod_{t=1}^{T} P\left(x_t \mid x_1, x_2, \ldots, x_{t-1}\right) \tag{7}$$

   The model parameters θ are optimized by minimizing the cross-entropy loss, which measures the difference between the predicted probability distribution and the actual tokens in the script is expressed as Eq. (8):

   $$\mathcal{L}_{CE} = -\frac{1}{T} \sum_{t=1}^{T} \log P\left(x_t \mid x_{<t}\right) \tag{8}$$

   Gradient-based optimization updates the model parameters iteratively is expressed. The AdamW optimization strategy is employed to decouple weight decay from gradient updates, enabling improved generalization and stable convergence during fine-tuning. Its adaptive learning rate mechanism effectively manages parameter updates across screenplay-style textual features, ensuring consistent optimization behavior as Eq. (9):

   $$\theta_{t+1} = \theta_t - \eta \nabla_\theta \mathcal{L}_{CE} \tag{9}$$

   where η is the learning rate controlling the step size of each update.

   Learning rate and batch size control enable GPT-2 to adapt to screenplay patterns while preserving pretrained language knowledge. This supports coherent narrative flow, consistent pacing, and structured dialogue in micro-film scripts.

3. Narrative Learning and Temporal Modeling

   Through the fine-tuning of GPT-2, we are capable of modelling short-form story arcs, integrating the dynamic interaction between characters, and also creating a fluid transition between scenes to ensure that all tokens within a script can be interpreted contextually, enabling a consistent and coherent narrative across the entire script. The given Eq. (10) as,

   $$P\left(x_t \mid x_{<t}\right) = \text{softmax}\left(f_\theta\left(x_{<t}\right)\right) \tag{10}$$

   where f_θ represents the GPT-2 transformer network.

3.4. Prompt Engineering

Prompt engineering is used to guide GPT-2 in generating micro-film scripts by providing constraints such as theme, characters, and length (e.g., "Generate a two-minute micro-film scene"). It further supports more coherent chronicle generation by preserving temporal progression and character consistency across successive scenes. Few-shot prompting further improves output by supplying example scripts, helping the model match desired style and structure. The prompt templates were further organized to include scene context, thematic boundaries, target duration, and character-role specifications for controlled screenplay generation. The prompt design incorporates structured narrative attributes such as scene context, character roles, and emotional tone to guide script generation. The prompt design incorporates structured narrative attributes such as scene context, character roles, and emotional tone to guide script generation. These attributes explicitly control character interactions and scene progression, ensuring consistency in storytelling. This structured mechanism enhances narrative coherence and context-awareness in generated scripts. Given a prompt P, the generated script G is computed as Eq. (11):

$$G = \arg\max P(G \mid P) \tag{11}$$

where $P(G \mid P)$ represents the conditional probability of generating the script given the prompt.

Table 1 The fine-tuned GPT-2 model demonstrates the ability to generate coherent and contextually relevant micro-film scripts from structured prompts. The outputs follow screenplay conventions, including scene headers, narrative descriptions, temporal continuity, and genre-specific styles.

Table 1. Scene Prompt–Based Micro-Film Script Generation Using Fine-Tuned GPT-2.

Scene Prompt	Generated Micro-Film Script	Source Movie
INT. HALLWAY LATER	The generated script opens with a formal screenplay structure, including authorship metadata, temporal markers, and a cinematic narration describing riders approaching along railroad tracks. The narrative voice contextualizes historical themes and character introduction using descriptive prose.	I Am Sam
EXT. LOS ANGELES STREET DAY	The output emphasizes strong visual symbolism, focusing on painted eyes, emotional expressions, and montage-style transitions. The script employs cinematic imagery and rapid scene progression to establish mood and thematic depth.	The Book of Eli
CUT TO:	The generated content adopts a commercial-style screenplay format, including voice-over narration, scene wipes, and character introduction. It reflects non-linear storytelling and realistic dialogue structure.	Sex, Lies and Videotape
INT. SHOE STORE CONTINUOUS	The script follows crime-drama conventions with concise scene transitions, action-driven narration, and character-focused exposition. Temporal and spatial continuity is maintained across indoor and outdoor scenes.	The Passion of Joan of Arc
EXT. FOREST - DAWN	The generated script highlights atmospheric setting and genre-specific cues, using minimal dialogue and descriptive language to convey tension and action-oriented storytelling.	Wall Street

4. Script analysis module

The Script Analysis Module systematically evaluates narrative structure, dialogue, and themes in both original and generated scripts. The module is organized into dialogue flow analysis, emotional progression modelling, and thematic representation evaluation, where dialogue consistency, sentiment transitions, and topic distributions are assessed. Distinct-n metrics measure lexical diversity by computing unique n-gram ratios, indicating originality and reduced repetition. Cosine similarity evaluates semantic alignment by comparing embedding representations of generated and reference scripts, reflecting thematic consistency. It ensures storytelling coherence, emotional connection, and contextual relevance comparable to professionally written micro-film scripts.

4.1. Narrative Coherence and Structure

Sequence-based analysis models how scenes and dialogues progress over time, enabling visualization of narrative development in micro-films. Narrative coherence is quantified by computing a coherence score based on thematic continuity and logical transitions between consecutive scenes. The coherence score is computed using embedding-based cosine similarity between consecutive scene representations, followed by weighted aggregation of semantic similarity and topic transition stability to obtain a normalized coherence value.

4.2. Dialogue Quality Evaluation

Dialogue quality is evaluated based on natural character interactions, emotional flow, and contextual relevance. Sentiment analysis measures emotional consistency across scenes, while contextual continuity ensures dialogue aligns with prior events. A structured evaluation methodology is further employed, where speaker attribution detection is performed using contextual embedding similarity to accurately assign dialogue to characters.

Algorithm 1: Automated Micro-Film Script Generation and Evaluation Framework

Input:
  $D = \{S_1, S_2, \ldots, S_n\}$ // Raw movie script dataset
  P // User-defined scene prompt
  Lmax // Maximum token sequence length
  θ_0 // Pre-trained GPT-2 parameters
Output:
  G // Generated micro-film script
  M // Evaluation metrics set
Begin
1. Data Preprocessing
2. For each script $S \in D$ do
3. Normalize text (case folding, whitespace removal, punctuation retention)
4. Tokenize script using GPT-2 subword tokenizer
5. Split tokens into overlapping sequences of length $\leq$ Lmax
6. End For
7. Dataset Preparation
8. Split tokenized data into Train (80%), Validation (10%), Test (10%)
9. Model Fine-Tuning
10. Initialize GPT-2 model with parameters θ_0
11. F or epoch = 1 to E do
12. For each batch B in Training Set do
13. Predict next-token probabilities using autoregressive objective
14. Compute cross-entropy loss
15. Update model parameters using AdamW optimizer
16. End For
17. End For
18. Obtain fine-tuned parameters θ^*
19. Script Generation
20. Encode prompt P using GPT-2 tokenizer
21. Generate script G using θ^* with nucleus sampling (top-p) and temperature scaling
22. Script Analysis Module
23. Compute BLEU and ROUGE scores for lexical quality
24. Compute Distinct-n scores for diversity
25. Compute semantic similarity using embedding cosine similarity
26. Evaluate narrative coherence and topic continuity
27. Perform sentiment consistency analysis
28. Store evaluation results as metrics set M
29. Return generated script G and evaluation metrics M
End

4.3. Semantic and Thematic Analysis

Embedding-based language models create dense semantic representations of scripts, enabling deeper understanding of underlying themes beyond surface text. These embeddings help identify clusters of storytelling styles across genres such as action, comedy, and drama. Within this embedding space, cosine similarity quantifies the conceptual alignment between generated scripts and reference screenplays by measuring the angular closeness of their semantic vectors.

5. Results and discussion

The proposed framework is implemented in Python 3.9 using PyTorch and Hugging Face Transformers, with GPT-2 fine-tuned using AdamW (learning rate 5×10 batch size 8,3 epochs). The training environment is configured on a CUDA-enabled system with an NVIDIA GPU, 16 GB RAM, and an Intel i.7. processor, ensuring efficient training, reproducibility, and stable execution of experiments. Tokenization uses GPT-2 subword methods with 512-token segments and overlapping windows to preserve context. Nucleus sampling (topp $= 0.9$) and temperature (0.8) control generation, with outputs limited to 300 tokens. The top-p value of 0.9 was selected empirically to retain only the most probable cumulative token distribution while filtering low-probability noisy continuations during inference. Evaluation metrics include BLEU, ROUGE, Distinct-n, semantic similarity, and sentiment consistency using NLTK, SacréBLEU, Scikit-Learn, and Sentence-Transformers. Results show that the system generates linguistically coherent, semantically accurate, and lexically diverse scripts. The learning rate of 5×10^{-5} was selected to ensure stable convergence while preventing overfitting, as commonly recommended for fine-tuning transformer-based language models. The chosen hyperparameters ensure stable training and reduce repetition, reflected in strong Distinct-n and semantic similarity scores. While scene-level coherence is strong, long-term thematic consistency is limited due to GPT-2's fixed context window. This limitation highlights the need for hierarchical scene-level planning mechanisms to better preserve long-range thematic continuity across generated micro-film narratives. Overall, the framework provides a reproducible and scalable solution for automated micro-film script generation and analysis. Fig. 3 presents

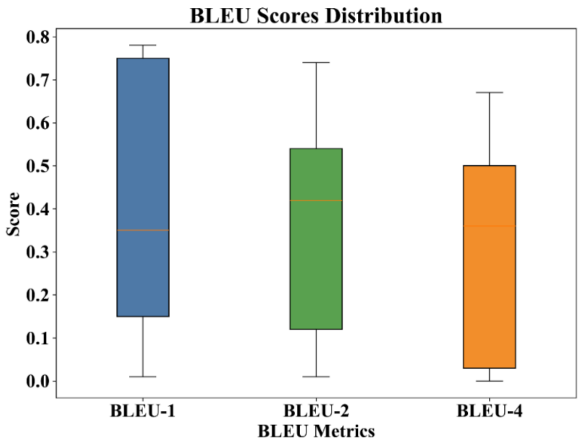

Fig. 3. Distribution of BLEU Scores for Generated Micro-Film Scripts.

mean and median BLEU-1, BLEU-2, and BLEU-4 scores for generated micro-film scripts, showing n-gram overlap with reference scripts. Higher BLEU-1 median scores indicate stronger unigram similarity, while lower BLEU-2 and BLEU-4 reflect reduced alignment in longer phrases. Fig. 4 illustrates the distribution of ROUGE-1, ROUGE-2, and ROUGE-L scores, measuring lexical and structural overlap between generated and reference scripts. The higher ROUGE-1 median indicates strong unigram consistency and appropriate vocabulary usage. Fig. 5 presents Distinct-1, Distinct-2, and Novel N-gram scores, highlighting the creativity and diversity of generated scripts. The Distinct-1 (0.64) and Distinct-2 (0.83) values indicate high lexical variation with minimal repetition. Fig. 6 evaluates dialogue quality across Turn Consistency, Speaker Attribution, Context Retention, and Emotional Coherence. High Speaker Attribution (0.80) and good Turn Consistency (0.645) indicate logical flow and correct character identification. Fig. 7 shows the distribution of narrative coherence scores, with a mean value of approximately 0.264. Most scripts fall within the 0.25–0.35 range, indicating generally consistent

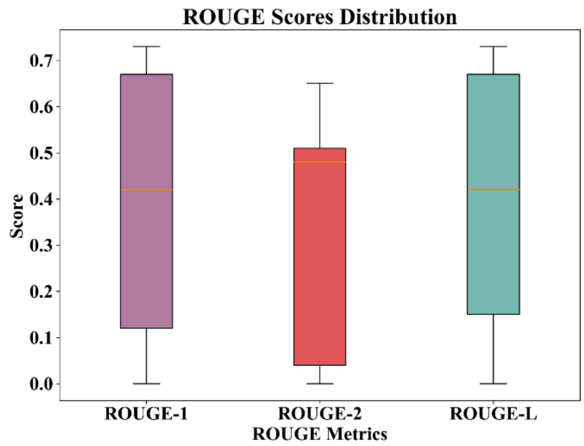

Fig. 4. Distribution of ROUGE Scores for Generated Micro-Film Scripts.

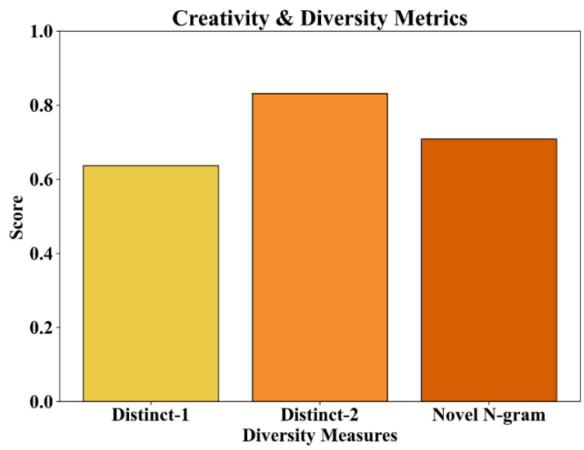

Fig. 5. Creativity and Diversity Evaluation of Generated Scripts.

coherence and smooth scene transitions. Fig. 8 illustrates variation in topic coherence across thematic indices in generated scripts. High coherence (≈0.74 at Topic Index 2) indicates strong thematic consistency, while very low coherence (≈0.03 at Topic Index 3) reflects difficulty in maintaining less-represented themes. Fig. 9 compares Entity Grid, Topic Continuity, and Scene Progression for narrative clarity. Scene Progression scores highest (0.80), indicating strong linear development across scenes. Entity Grid (0.68) reflects moderate consistency in tracking characters, while Topic Continuity (0.55) shows limited thematic consistency. Fig. 10 evaluates script quality using narrative coherence, sentiment consistency, semantic similarity, and dialogue density. High semantic similarity (0.935) indicates strong alignment with reference narratives, while dialogue density

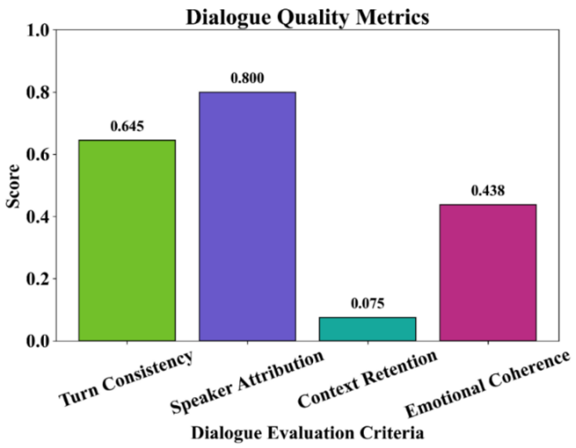

Fig. 6. Dialogue Quality Evaluation Metrics.

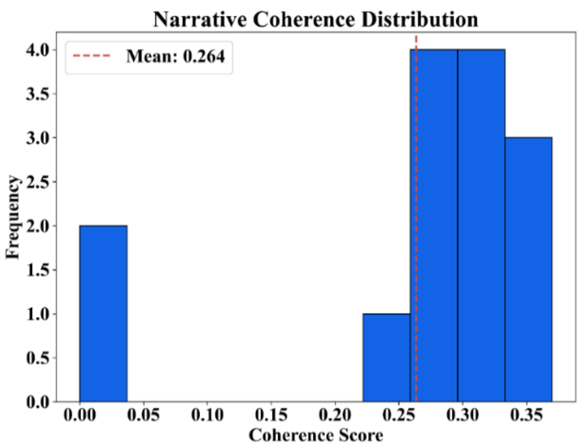

Fig. 7. Distribution of Narrative Coherence Scores.

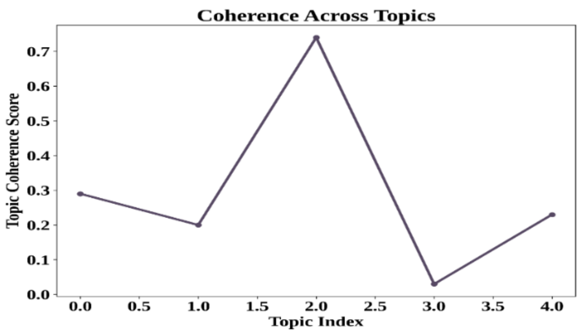

Fig. 8. Topic-Level Coherence Variation Across Generated Scripts.

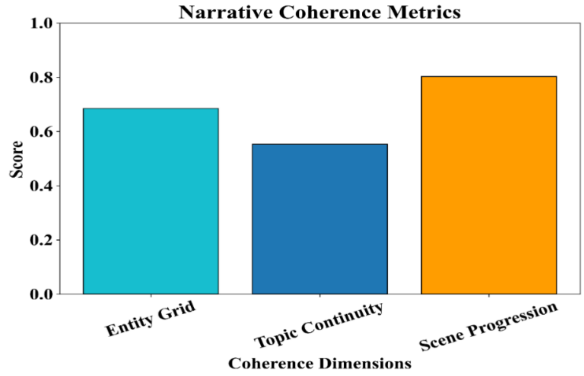

Fig. 9. Narrative Coherence Metrics Across Structural Dimensions.

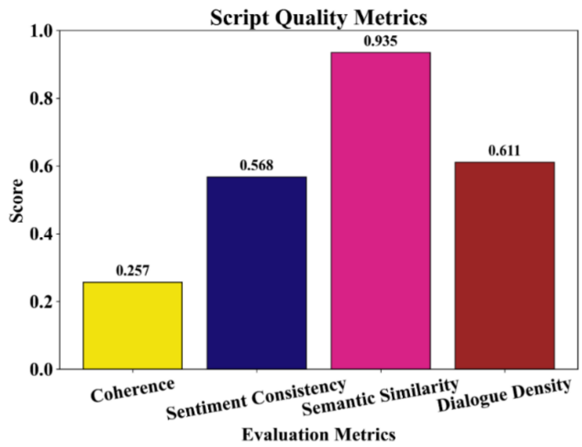

Fig. 10. Overall Script Quality Evaluation Based on Multiple Metrics.

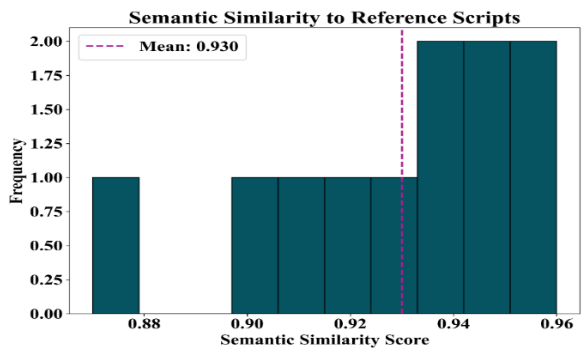

Fig. 11. Semantic Similarity Distribution Between Generated and Reference Scripts.

(0.611) and sentiment consistency (0.568) reflect balanced dialogue and stable emotional progression. Fig. 11 shows a high average semantic similarity (0.93) between generated and reference micro-film scripts, with values concentrated in the upper range. This indicates that the model effectively preserves author intent, context, and thematic relevance.

Fig. 12 evaluates semantic alignment using Cosine Sim-

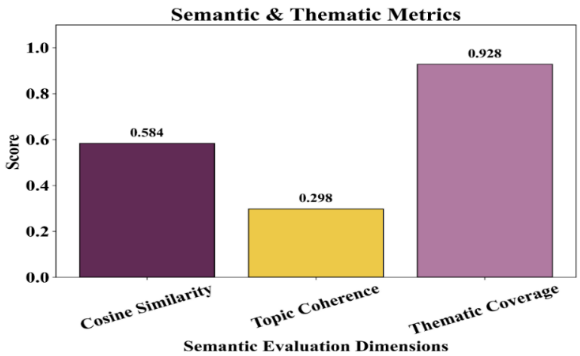

Fig. 12. Evaluation of Semantic and Thematic Consistency Metrics.

ilarity, Topic Coherence, and Thematic Coverage. High Thematic Coverage (0.928) indicates strong inclusion of key narrative themes. Fig. 13 shows the distribution of

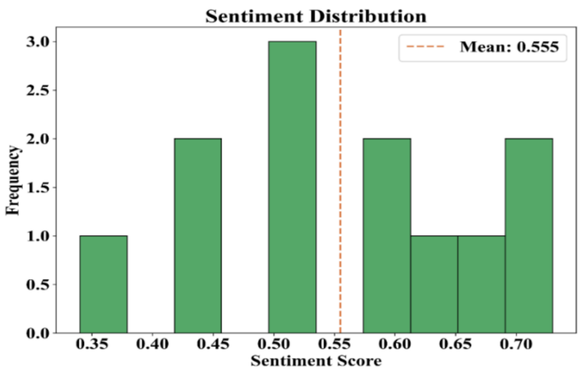

Fig. 13. Sentiment Score Distribution Across Generated Scripts.

sentiment scores across generated micro-film scripts, with an average of 0.555, indicating balanced emotional tone. Most scripts exhibit neutral to slightly positive sentiment, reflecting stable emotional progression. Fig. 14 evaluates semantic similarity across generated scripts using cosine similarity. Most script pairs show similarity above 0.6, with averages near 0.8, indicating stable thematic consistency and effective context retention. Fig. 15 shows wide variation in script length (mean ≈ 21,862 words), indicating diverse narrative scope and the need for robust preprocessing and segmentation. This token variation reflects differences in story complexity, where longer scripts capture richer events and character development. Shorter scripts, in contrast, require denser narrative compression within limited token spans. Vocabulary richness is moderately

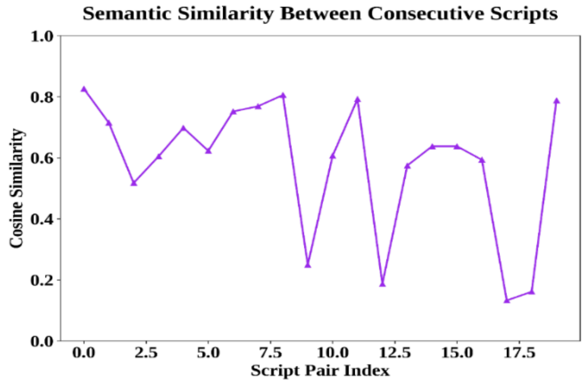

Fig. 14. Semantic Similarity Between Consecutive Scripts.

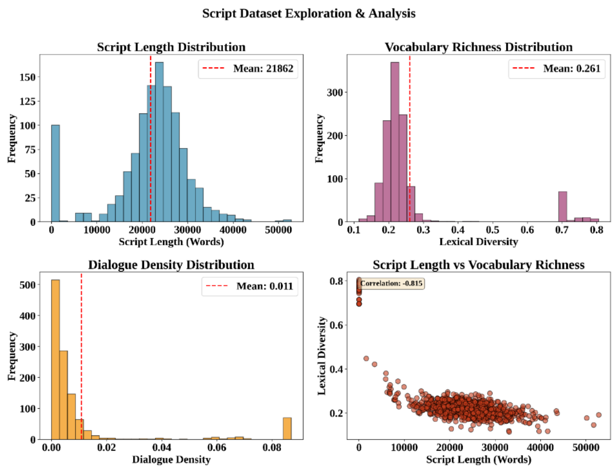

Fig. 15. Script Length Distribution and Linguistic Diversity Analysis.

low (≈0.26), reflecting a mix of unique expressions and repeated patterns. Fig. 16 shows significant structural variation in screenplays, with an average of 146 scenes, indicating high narrative segmentation. Dialogue distribution varies, with some scripts being dialogue-rich while others are less so.

6. Conclusion and future works

The study presents an AI-assisted framework for generating and analyzing micro-film scripts using structured preprocessing, fine-tuned GPT-2 models, and multi-dimensional evaluation metrics. Results demonstrate the ability to produce grammatically correct, fluent scripts with coherent dialogue flow and scene progression, validating the effectiveness of transformer-based approaches for short-form narrative generation. However, limitations remain in maintaining long-term coherence, topic continuity, and emotional progression due to fixed context constraints. Fu-

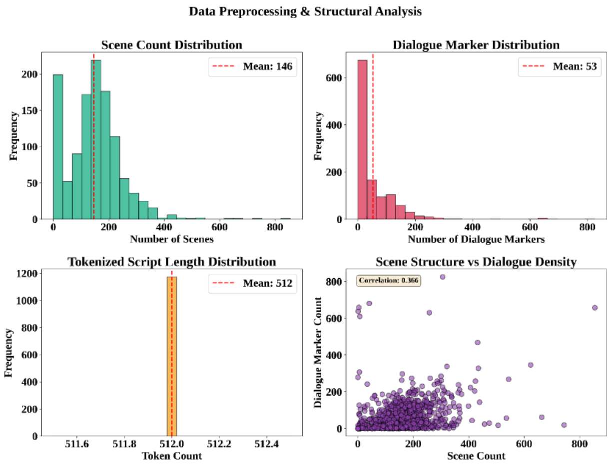

Fig. 16. Structural Properties of Scripts and Preprocessing Effects.

ture work focuses on hierarchical narrative planning, topic-aware conditioning, extended context modeling, and advanced architectures such as larger transformers, memory-enhanced networks, and reinforcement learning. Incorporating multimodal inputs (visual, audio, and genre constraints) is expected to further enhance realism, scalability, and practical applicability across creative, educational, and commercial domains.

References

[1] R. Luo, R. Huang, Y. Wang, and L. Ru, (2024) *"Digital Storytelling as a New Narrative Technique in Short Videos Telling China Stories"* **Advances in Journalism and Communication** *12*(2): 246–257. DOI: 10.4236/ajc.2024.122013.

[2] A. Ocal, (2025) *"Cinematic Narratives as Socio-Technical Systems: Emotion Mining and Script–Audience Emotional Fidelity"* **Systems** *13*(11): 994. DOI: 10.3390/systems13110994.

[3] A. Vázquez, A. López Zorrilla, J. M. Olaso, and M. I. Torres, (2023) *"Dialogue Management and Language Generation for a Robust Conversational Virtual Coach: Validation and User Study"* **Sensors 23**(3): 1423. DOI: 10.3390/s23031423.

[4] Y. Zhu, R. Song, J.-Y. Nie, P. Du, Z. Dou, and J. Zhou, (2022) *"Leveraging Narrative to Generate Movie Script"* **ACM Transactions on Information Systems** *40*(4): 86:1–86:32. DOI: 10.1145/3507356.

[5] Z. Zhang and Z. Hua, (2025) *"Balancing Revenue Streams in Online Video Platforms: The Impact of Original Content Provision on Business Model Selection"* **Journal of Theoretical and Applied Electronic Commerce Research** *20*(2): 98. DOI: 10.3390/jtaer20020098.

[6] S. Indra, M. A. I. Yassin, S. A. Shekh Alsagoff, and R. Bidin, (2023) *"New Journalism in Documentary Script Writing"* **International Journal of Academic Research in Business and Social Sciences** *13*(3): 767–779. DOI: 10.6007/IJARBSS/v13-i3/16335.

[7] W. Hanmei, (2024) *"Deep Learning-Based Analysis of Emotional and Content Relevance Between Bullet Screens and Subtitles as Movie Narrative Medium"* **SAGE Open** *14*(3): 21582440241280840. DOI: 10.1177/21582440241280840.

[8] D. T. K. Ng, W. Luo, H. M. Y. Chan, and S. K. W. Chu, (2022) *"Using Digital Story Writing as a Pedagogy to Develop AI Literacy Among Primary Students"* **Computers and Education: Artificial Intelligence** *3*: 100054. DOI: 10.1016/j.caeai.2022.100054.

[9] A. Bucaioni, F. Di Silvestro, I. Singh, M. Saadatmand, and H. Muccini, (2022) *"Model-Based Generation of Test Scripts Across Product Variants: An Experience Report from the Railway Industry"* **Journal of Software: Evolution and Process** *34*(11): e2498. DOI: 10.1002/smr.2498.

[10] S. Cake, (2025) *"Artificial Intelligence as a Collaborative Tool for Script Development"* **Media Practice and Education** *26*(3): 302–317. DOI: 10.1080/25741136.2025.2454074.

[11] N. A. Arjomand, (2024) *"Empirical Fiction: Composite Character Narratives in Analytical Sociology"* **The American Sociologist** *55*(4): 436–472. DOI: 10.1007/s12108-022-09546-z.

[12] M. Skowron, M. Trapp, S. Payr, and R. Trappl, (2016) *"Automatic Identification of Character Types from Film Dialogs"* **Applied Artificial Intelligence** *30*(10): 942–973. DOI: 10.1080/08839514.2017.1289311.

[13] S. Ge, (2025) *"The Impact of AI Algorithm Optimization on Content Quality Enhancement in Short Video and Micro-Film Production"* **International Journal of Advanced Applied Science Research** *4*(8): 113–118.

[14] H. Fan, (2024) *"The Art and Challenge of Micro Film Creation"* **Studies in Art and Architecture** *3*(2): 51–56.

[15] H. Zhan, (2024) *"Aesthetic Characteristics and Creative Practice of Micro-Films from the Perspective of Directors"* **Frontiers in Art Research** *6*(11): DOI: 10.25236/FAR.2024.061103.

[16] Z. Shi and L. Li, (2024) *"A Study on the Aesthetic Characteristics of Documentary Films in the Internet Communication Environment"* **Advances in Education, Humanities and Social Sciences Research** ***9***(1): 64. DOI: 10.56028/aehssr.9.1.64.2024.

[17] L. Kong, (2023) *"When State-Endorsed Cinema Meets Marvel: A Study of Wolf Warrior II's Global Superhero Vernacular"* **Arts** ***12***(6): 252. DOI: 10.3390/arts12060252.

[18] J. Yuan, (2023) *"Application of Micro Film Advertising Communication Under New Media Form"* **Frontiers in Social Sciences and Technology** ***5***(15): DOI: 10.25236/FSST.2023.051513.

[19] G. Li, (2025) *"Analysis of New Trends in Microdramas: On the Example of Social Media Drama Shorts in China and US"* **Perspectives on Communication** ***1***(3): 1–11. DOI: 10.63808/pcm.v1i3.251.

[20] M. J. Roshan, A. S. B. A. Rashid, M. A. B. Hezmi, M. N. Nejabi, S. N. B. Jusoh, S. Tamassoki, and R. Razali, (2022) *"Evaluation of Cement Stabilised Residual Soil on Macro-and Micro-Scale for Road Construction"* **Journal of Engineering and Applied Science** ***69***(1): 109. DOI: 10.1186/s44147-022-00165-6.

[21] Z. Jia and S. B. Mohamed, (2023) *"Navigating the Tides of Media Convergence: Shanxi Radio and Television Station's Transformation Amidst the Digital Tsunami"* **International Journal of Academic Research in Business and Social Sciences** ***13***(4): 1289–1298. DOI: 10.6007/IJARBSS/v13-i4/16757.

[22] J. Zschocke, (2022) *"Towards a Multimethod Approach for Film Geography: A Case Study of Los Angeles' Nate Starkman Building"* **GeoJournal** ***87***(1): 85–93. DOI: 10.1007/s10708-022-10660-1.

[23] C. Luo, N. A. M. Hasan, A. M. Z. Ahmad, and G. Lei, (2025) *"Influence of Short Video Content on Consumers Purchase Intentions on Social Media Platforms with Trust as a Mediator"* **Scientific Reports** ***15***(1): 16605. DOI: 10.1038/s41598-025-94994-z.

[24] R. Zainal and M. A. M. Desa, (2025) *"The Benefit of Using Emotion in 3D Animation Film: Pre-Production Stage"* **PaperAsia** ***41***(1b): 255–270. DOI: 10.59953/paperasia.v41i1b.162.

[25] G. B. Takala, (2023) *"The Interactive Creativity of the Digital Era: Exploring How Media Art Redefines the Relationship Between Audience and Artwork"* **Studies in Art and Architecture** **2**(3): 28–44.

[26] I. Mousa. *README.md · IsmaelMousa/movies at main*. Accessed: 2026-01-09. 2026.