A Study on Audience Discourse and Emotional Identification of YOLO from the Perspective of Digital Humanities—Based on Corpus Text Mining of Douban Reviews

Wanjing Zhang

University of Electronic Science and Technology of China, Qingshuihe Campus, Chengdu, China
1303195785@qq.com

Abstract. The social influence of films is closely related to audience discourse. In the context of film cultural communication, there is growing attention to exploring how audience discourse reflects film reception and social responses. This study employs a corpus linguistics approach, grounded in discourse analysis, identity theory, narrative theory, and feminist theory, to construct a corpus of audience reviews for the film YOLO (You Only Live Once) on the Douban platform from February 10 to April 10, 2024. It deeply examines the cognition, attitudes, and underlying emotional identification and value orientations expressed in audience discourse. The findings can provide valuable reference for filmmakers to understand audience needs, optimize creative practices, and promote the healthy development of the film industry.

Keywords: YOLO, corpus linguistics, discourse analysis, female perspective

1. Introduction

Since its official release in mainland China on February 10, 2024, the film YOLO has rapidly ignited enthusiasm on social media and in public discourse, sparking extensive and lively discussions. By the end of its theatrical run on April 10, 2024, the film had achieved a cumulative box office of 3.46 billion RMB, with 72.034 million viewers, earning the dual championship in box office revenue and audience attendance for the 2024 Spring Festival period, and breaking the box office record for narrative films during the Spring Festival period in Chinese film history. Films serve as an important window for understanding popular aesthetic preferences, social psychology, and cultural values. In today's highly competitive arts market, audience feedback and evaluations of films are increasingly evolving into a form of socially influential discourse practice. At the micro level, these reviews reflect individuals' understanding of film content and emotional projection, while at the macro level, they constitute the collective perception of a film's artistic value, social significance, and cultural expression. Online audience reviews, due to their high engagement and authenticity, have become an important discursive resource for studying film reception and social impact.

However, the success of YOLO has not been without controversy. From the film's initial release, a variety of voices have emerged regarding its content, thematic expression, and directorial

techniques. Douban, as one of the most representative book, film, and music review communities in China, gathers perspectives from audiences of diverse ages, professional backgrounds, and cultural levels, forming a pluralistic and representative commentary field. From artistic aspects to ideological values, reviews on the platform encompass reflections on plot development, character construction, directorial methods, social issues, and gender expression. Therefore, selecting audience reviews on Douban as the source corpus for systematic discourse analysis enables a more comprehensive understanding of audiences' cognition and attitudes toward YOLO.

This study takes corpus linguistics as its methodological foundation and selects a large number of film review texts on Douban concerning YOLO as the research objects. Leveraging natural language processing techniques, the analysis is conducted along three dimensions: high-frequency vocabulary, sentiment orientation, and thematic focus. This research aims to provide a new analytical path and methodological perspective for examining the artistic value and social impact of YOLO from a linguistic perspective. By systematically analyzing audience discourse, the study seeks to offer valuable references for filmmakers, critics, academic researchers, and related practitioners, helping them more accurately grasp audience needs and aesthetic preferences, thereby promoting a positive interaction between film creation and industry development and further advancing the diversified development and continuous innovation of Chinese film culture.

2. Research content and methods

2.1. Theoretical foundations

This study takes corpus linguistics as its methodological foundation and integrates multiple theoretical perspectives to construct the analytical framework. Discourse analysis theory [1] posits that language serves as a carrier of social practice, and the aggregation of public discourse reflects social cognition and collective consensus. By analyzing lexical choices, syntactic structures, and narrative strategies in public discourse, it is possible to uncover specific social perceptions and power relations. This approach helps to understand how readers or audiences express their understanding of a topic and emotional projection through language, thereby revealing their deeper value orientations and social psychology. Kenneth Burke's [2] identification theory emphasizes that rhetoric facilitates cooperation through emotional resonance and symbolic association. Its core concept, Identification by Sympathy, refers to the resonance individuals feel due to shared experiences. This theory has been widely applied to analyze audience emotional identification with a subject, especially when the creator's real-life experiences resonate with the content of the text, enhancing the audience's perception of genuine personal growth or transformation [3]. Narrative theory [4,5] focuses on the intertextuality between stories and reality. The theory suggests that when storylines reflect real-life situations, audiences are more likely to resonate with them. Therefore, in analyzing audience feedback, narrative theory can help explore audiences' recognition of the protagonist's growth process and their acceptance of non-traditional endings. Feminist theory, particularly perspectives on bodily politics and subjectivity [6,7], reveals that female characters achieve self-mastery through embodied practices, refusing to define their value through dependence on others, thereby challenging patriarchal discipline. This theoretical lens helps to understand that, in cultural products, the growth of female characters is not merely a reconstruction of outward appearance but also symbolizes the enactment of their own agency.

2.2. Research content

Based on discourse analysis, identification theory, narrative theory, and feminist theory, this study aims to explore audience preferences and criticisms expressed in responses to the film YOLO, as well as the underlying emotional identification and value orientations. The corpus data for this study are drawn from Douban, one of China's most influential film and media review communities, which combines professional and popular perspectives. Since its establishment in 2005, Douban has become an important platform for reflecting audience evaluations of audiovisual works, owing to the authenticity and diversity of user-generated content.

Specifically, this study collected audience reviews of YOLO from its release on February 10, 2024, to the end of its theatrical run on April 10, 2024, covering the film's complete screening period and capturing audience feedback and emotional changes at different stages. Using keyword searches on Douban (e.g., "YOLO," "Ling Jia," "weight loss"), a large volume of discussion content related to the film was retrieved. The search results were then sorted by relevance, and each review was carefully read and screened, excluding images, videos, and comments unrelated to the film's content. Ultimately, a small thematic corpus containing 3,000 valid audience reviews, totaling approximately 80,000 words, was constructed. This corpus is representative and can comprehensively reflect the diverse perspectives of audiences toward YOLO, providing a solid data foundation for subsequent analyses of high-frequency words, sentiment orientation, and thematic focus.

2.3. Research questions and methods

This study primarily seeks to address the following three questions:

(1) In Douban reviews of YOLO, what core concerns of the audience are reflected through high-frequency words and core semantic clusters?

(2) What is the overall sentiment orientation of the reviews? Which dimensions of the film are most associated with different sentiment orientations?

(3) What thematic foci have audiences constructed through their reviews, and what forms of consensus and identification do these themes reflect?

Specifically, this study employs a mixed-methods approach combining quantitative analysis and qualitative interpretation to achieve a multidimensional examination of audience evaluations of YOLO.

At the quantitative level, the raw corpus was first preprocessed using the "jieba" word segmentation tool in the Python programming language, including tokenization and the removal of stop words (such as "的," "了," "是," which carry little semantic content). Subsequently, the cleaned corpus was analyzed using the Pandas data processing library to calculate word frequencies and generate a list of high-frequency words, identifying the most representative audience concerns in the reviews. In addition, based on the "emotional resonance" framework in Burke's identification theory, combined with manual annotation and the SnowNLP sentiment analysis tool, the sentiment orientation of each review was classified into three categories: positive, negative, or neutral. By examining the proportion of these three sentiment categories and their distribution across different topical dimensions (e.g., actor performance, plot structure, thematic expression), the study further reveals audiences' emotional reception and its dynamic trends.

At the qualitative level, grounded theory's three-level coding method—open coding, axial coding, and selective coding—was applied based on high-frequency semantic clusters to gradually extract the core themes embedded in audience reviews. Centered on high-frequency words such as

"weight loss," "self," and "growth," themes such as "weight loss and self-identification" and "female growth narratives" were synthesized. These themes were then further interpreted through the lens of Burke's identification theory to uncover the underlying emotional and identification mechanisms.

To protect user privacy, identifiable information in the corpus, including usernames, avatars, and IP addresses, was anonymized. When citing review content, only the core expressions were extracted, without presenting the full original text.

3. Research results

3.1. Core connotations of high-frequency words in YOLO reviews

Using word frequency statistical tools, 3,000 audience reviews of YOLO were analyzed, and the top 20 high-frequency content words (excluding function words and semantically meaningless terms) were identified, as shown in Table 1:

Table 1. Top 20 high-frequency content words in 3,000 reviews of YOLO

No.	Keyword	Frequency
1	Weight loss	1049
2	Ying Le	560
3	Director	533
4	Story	474
5	Women	444
6	Like	436
7	Boxing	357
8	Feeling	345
9	Inspirational	326
10	Plot	312
11	Spring Festival	306
12	Life	304
13	Marketing	264
14	Success	260
15	Friend	253
16	Life journey	238
17	Huanying Li	238
18	Female lead	231
19	Remake	219
20	Character	218

By observing the keywords in the table, the core focus of audience reviews can be summarized into four main aspects:

(1) Core narrative and protagonist image. The term "weight loss" (1,049 occurrences) ranks first, followed by "Ying Le" (560), "female lead" (231), and "boxing" (357), clearly pointing to the film's central narrative centered on the protagonist Ying Le's weight loss and boxing journey. This

combination reflects the audience's high attention to the protagonist's growth trajectory—"weight loss" is not only a plot symbol but also a vehicle for the character's self-breakthrough, while "boxing" may serve as a key behavior facilitating her transformation. Together, they constitute the audience's fundamental understanding of the film's story.

(2) Discussion related to artistic expression and creation. Terms such as "story" (474), "plot" (312), and "director" (533) reflect audience attention to the film's artistic expression. The high frequency of "director" suggests that viewers relate the film to the director's creative style (similar to the director's previous work Hello, Huanying Li, which also explains why "Huanying Li" appears 238 times in the high-frequency list). The co-occurrence of "story" and "plot" indicates that audiences value narrative completeness and logical coherence, which are important dimensions in evaluating film quality.

(3) Emotional orientation and value resonance. Words such as "like" (436), "inspirational" (326), and "success" (260) convey an overall positive emotional tendency. "Like," as a direct expression of attitude, confirms that the film received substantial positive recognition from viewers. "Inspirational" and "success" point to the film's conveyed value core—through the protagonist's struggle, it elicits emotional resonance with themes of "growth" and "breakthrough," which is a central source of positive evaluation.

(4) Controversial topics. Terms such as "remake" (219), "marketing" (264), and "Spring Festival" (306) involve external associations and points of debate. "Spring Festival" reflects the influence of the film's release period (Spring Festival) on audience discussion. "Remake" indicates that the film is adapted from the Japanese movie Centennial Love, triggering comparisons and discussions of adaptation quality. The appearance of "marketing" is related to audience attention to promotional strategies, such as the alignment between publicity and actual content.

In summary, the combination of high-frequency words clearly outlines the focal concerns of audiences regarding YOLO: narrative content centered on the protagonist's growth, the quality of the film's artistic expression, emotional resonance evoked by inspirational values, and external topics associated with the film. Together, these constitute the core connotations of audience discourse in film reviews.

3.2. Sentiment analysis results

Using the SnowNLP sentiment analysis toolkit, audience reviews of YOLO on Douban were analyzed, yielding the following results:
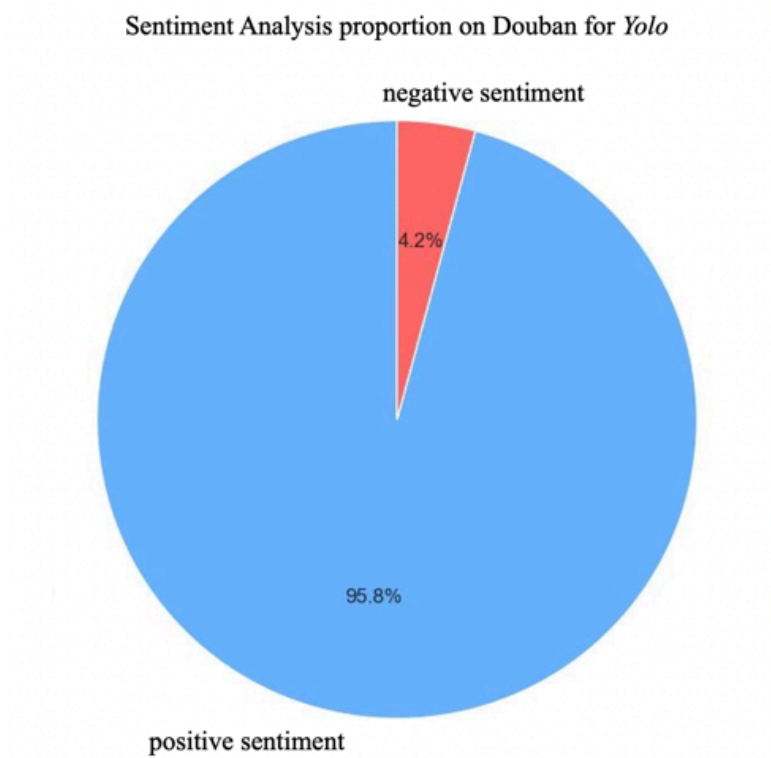
Figure 1. Pie chart of sentiment analysis for Douban reviews of YOLO
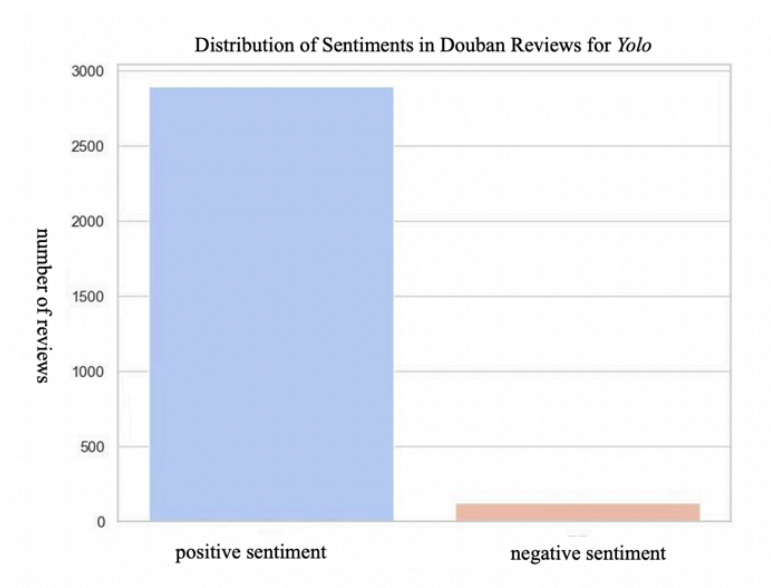

Figure 2. Bar chart of sentiment analysis for Douban reviews of YOLO

The sentiment analysis of YOLO reviews shows that positive emotions account for 95.8% of all comments, while negative emotions account for only 4.2%. This distribution aligns with the film's high box office performance, demonstrating that the movie has been widely loved and recognized by audiences, successfully attracting a large number of viewers to theaters.

Audience appreciation can be primarily summarized across three dimensions:

(1) Resonance with the theme: Many viewers were moved by the inspirational spirit conveyed in the film. As reflected in the previous high-frequency word analysis, terms such as "inspirational" and "success" indicate that the motivational theme is a key reason for eliciting positive emotions. The protagonist Ying Le's resilience and pursuit of success through weight loss and boxing inspired audiences. One viewer commented: "At the end, the film reminds anyone with a people-pleasing personality to 'live for yourself once.' This statement, delivered by a female comedian like Ling Jia, who has long used her weight as a source of self-deprecating humor, is particularly persuasive." Another noted: "So many training details resonated with me—the jump rope warm-up for coordination, the roars during strength training, the screams in stretching and massage, the bodyweight cardio that made me want to cry. But more importantly, the resonance comes from being the same gender as the protagonist and facing similar situations." Many viewers saw themselves or people around them reflected in Ying Le's story, experiencing the courage to persevere in the face of difficulties, which generated strong emotional resonance and positive expressions of appreciation.

(2) Authentic and vivid character portrayal: The character of Ying Le won the audience's hearts. "Ying Le has flaws, feels lost at times, but she keeps growing, which is very authentic. She is like a friend around us, allowing us to see ourselves in her and cheer for her." With her distinct personality and charm, Ying Le's growth and transformation elicited audience affection, and liking the character contributed to an overall positive sentiment toward the film. As one comment explained: "Ying Le likely has an INFP personality—highly sensitive, empathetic, and low-energy. She empathizes strongly, often sacrificing her own needs to satisfy others, and every grievance accumulates internally. Being low-energy and fearing harm, she gradually becomes extremely repressed and lacking in confidence." Such realistic character development allows audiences to immerse themselves fully in her story and care deeply about her fate.

(3) Plot arrangement: The film's innovative narrative structure also attracted viewers. "I loved the ending. It's lonely, but what a light and graceful loneliness it is—walking alone under the pure moonlight, having done one's utmost, facing the unknown path ahead calmly." The film breaks the conventional "happy-ending" model of motivational movies, redefining "growth" and "success" through a "lonely yet graceful" conclusion. As noted in comments, phrases like "having done one's utmost" indicate that the film shifts the standard of success from external recognition to internal self-consistency. This redefinition of the endpoint of growth moves beyond reliance on conventional closure, making self-acceptance a more important value than external validation.

Negative evaluations are comparatively few and focus mainly on "film marketing" and "remake issues." Some viewers felt a gap between the film's promotion and its actual content. One comment stated: "Just looking at Ling Jia's self-indulgent marketing is off-putting. Years ago, actors losing or gaining weight, training, or practicing martial arts for an art film would move audiences. Now, emphasizing 'Ling Jia lost 100 jin' repeatedly without linking it to character development or plot depth makes 'weight loss' a marketing gimmick, reducing professionalism to a spectacle." Such marketing, detached from the film itself, led audiences to perceive dedication as being instrumentalized. Weight loss ceased to serve Ying Le's character growth and became merely an attention-grabbing tactic, naturally provoking resistance.

Additionally, as a remake, comparisons with the original work disappointed some viewers. "The protagonist's environment changed from the original 'underclass mud' to 'everyone is malicious.' The original depicted a realistic social microcosm; characters' behaviors followed logic, producing stronger realism. YOLO abstracted the story, heightened conflicts, and added comedy, but drifted away from reality, making the setting vague." The original's realistic narrative allowed audiences to see real-life reflections in the characters, creating emotional impact. The adaptation of YOLO

deliberately intensified antagonism (relatives' exploitation, friends' betrayal) and diluted the original's depth with humor. This treatment, while superficially heightening conflict, detached the story from reality, causing viewers familiar with the original to feel a loss of authenticity. "Honestly, when I first heard Ling Jia made a movie called YOLO, I had slight expectations—the title wasn't off-putting. But learning about the plot—loser, boxing, training, weight loss, motivational—immediately reminded me of a Japanese movie, 100 Yen Love." From the comments, the adaptation failed to achieve effective localization; the core storyline "loser transforms through boxing" closely mirrors the original, yet lacks local contextual integration. High expectations of the original led audiences to critique the remake.

3.3. Analysis of core themes

Audience reviews of YOLO revolve around the film's textual content and its multiple connections to real-life experiences, forming several distinct themes and achieving broad consensus on core values. These themes reflect the film's narrative strategies and emotional intentions, while also revealing viewers' deep projections regarding social issues, individual growth experiences, and cultural value recognition. The focus of audience commentary extends beyond simple plot evaluation to encompass broader layers, such as "intertextuality between text and reality" and "individual growth and social resonance." Specifically, discussions center on three core themes:

First, self-growth: This theme represents a dual breakthrough from physical transformation to psychological independence. The transformation of Du Ying Le from a people-pleasing personality to gaining agency over her own life became a focal point of audience discussion. Her journey—from unlimited compromise with family and excessive accommodation of others, to finally responding to Hao Kun with a "depending on my mood" attitude—illustrates a dual realization of self-worth. Physically, she achieved health through weight loss and boxing; psychologically, she shifted from defining success through others' recognition to self-defined achievement. Audiences widely acknowledged the film's redefinition of "growth," which no longer relies solely on conventional notions of victory, but emphasizes persistence and self-transcendence. Common review expressions such as "she didn't win the competition, but she won herself" exemplify emotional identification with this non-utilitarian view of growth.

Second, interweaving of film and real life: The extreme physical transformation undergone by Ling Jia for role preparation—gaining 40 jin and then losing 100 jin—creates strong intertextuality with Du Ying Le's growth trajectory, enhancing viewers' emotional resonance. Audiences generally perceive Ling Jia's real-life actions as embodied symbols of inspiration, believing that consistency between on-screen and off-screen experiences strengthens the film's emotional authenticity and narrative persuasiveness. Ling Jia's embodiment of the character in reality creates a vivid blend of fact and fiction, allowing audiences to empathize strongly and develop value recognition, with a sentiment of "if she can do it, we can believe we can do it too," reflecting high appreciation for tangible effort and replicable growth paths.

Third, localized narrative: The film incorporates socio-cultural issues specific to contemporary China—such as disputes over school district housing, workplace sexual harassment, and malicious media editing—stimulating audience attention and discussion of local social problems. Plot elements such as family conflicts, betrayal by friends, and online public pressure are interpreted by viewers as accurate depictions of pressing social issues in China. These narrative choices allow audiences to feel seen and understood, facilitating emotional engagement and reflection on real-world social experiences.

Through thematic analysis of audience reviews, it is evident that viewers generally reject "suspended motivational" narrative templates and instead value ordinary characters persisting in adversity [8]. They no longer seek sudden heroic reversals, but rather empathize with relatable figures who continue striving in difficult circumstances. The emotional impact of this narrative lies in audiences seeing themselves reflected in the characters—they may not experience boxing competitions, but have all been knocked down and gotten back up; they may not achieve complete turnaround, but continually strive to become better versions of themselves. This grounded resonance represents the film's core strength beyond conventional "motivational" tropes.

Furthermore, data analysis indicates that contemporary audiences are inclined to trust narratives verifiable in reality. The film establishes multiple "anchors" between the fictional story and the real world: Ling Jia's real-life experience serves as a behavioral anchor, localized details as contextual anchors, and emotional experience as a psychological anchor [9]. These anchors convince viewers that the story is not an unreachable legend, but a narrative that could happen around them and has even been validated in reality. This trust surpasses the requirement for plot plausibility, essentially reflecting audience recognition of "the power of authenticity." They believe in narratives of self-transformation through sustained effort because such change is "verifiable, tangible, and replicable" in real life.

Finally, the film conveys a concept of female self-subjectivity: "self-worth should be self-defined," which has been widely accepted by viewers. Du Ying Le's growth trajectory—from dependent existence to awakening of subjectivity—demonstrates how women gradually break free from external value standards and establish a value system centered on personal perception. On the level of bodily narrative, the film challenges traditional gendered representations of women as objects of the gaze, detaching weight from societal aesthetic standards and transforming it into symbols of "self-control" and "agency." [10] Du Ying Le's physical training and Ling Jia's real-life weight fluctuation jointly construct a narrative logic in which the body serves the self, embedding the concept that bodily autonomy equates to value autonomy. By downplaying male dominance, the film further strengthens the proactive nature of female self-redemption. Du Ying Le's critical growth moments stem from her own choices rather than male guidance or rescue. This narrative strategy breaks the conventional logic in film and television that female value must be validated by male recognition, enabling audiences to recognize that women's growth and value establishment do not require dependence on others.

4. Conclusion

This study utilized a corpus of 3,000 audience reviews of YOLO from the Douban platform, employing high-frequency word analysis, sentiment orientation interpretation, and thematic summarization to examine viewers' perceptions and attitudes toward the film. The findings indicate that audience attention extends beyond surface-level symbols such as "weight loss" and "boxing," centering instead on the protagonist's personal growth, embodied practices, and the integration of localized social issues, which together evoke emotional resonance and value recognition among viewers. High-frequency words emphasize themes of "self-growth" and "female value," and, combined with a positive sentiment proportion as high as 95.8% and thematic analysis, reflect collective audience identification with concepts such as authentic growth and self-defined value. The study also captured negative feedback related to the disconnect between marketing and content, highlighting the need for creators to be cautious of prioritizing form over substance. Overall, audience discourse surrounding YOLO mirrors contemporary social psychology and provides practical insights for film production: grounding narratives in reality, conveying sincere values, and

building emotional connections are key factors for a work to gain audience recognition. Future research may explore the integration of authentic storytelling with emotional resonance to further promote constructive interactions between Chinese cinema and broader social culture.

References

[1] Fairclough, N. (2013). Critical discourse analysis. Taylor and Francis.
[2] Burke, K. (1969). A rhetoric of motives. Berkeley: University of California Press.
[3] Liu, Y. (2008). History of Western rhetoric [M]. Beijing: Foreign Language Teaching and Research Press.
[4] Prince, G. (2003). A dictionary of narratology. Lincoln: University of Nebraska Press.
[5] White, H. (1975). Metahistory: The historical imagination in nineteenth-century Europe. Baltimore: Johns Hopkins University Press.
[6] Beauvoir, S. (2009). The second sex (S. Shu, Trans.). Beijing: Xiyuan Press.
[7] Butler, J. (2006). Gender trouble: Feminism and the subversion of identity. New York: Routledge.
[8] Wan, C. (2024). Narrative overflow, boundary resetting, and strategies of localizing genres: A review of the film Hot Hot Heat [J]. Contemporary Film, (3), 39–43.
[9] Yuan, Y. (2020). The consumption of imagination in film aesthetics: Motivation, production, and purpose [J]. Journal of Shanghai University: Social Sciences Edition, (1), 48–56.
[10] Wang, L. (2022). The artistic discourse logic of contemporary Chinese female film directors [J]. Film Literature, (3), 90–93.