CINEFORUM
ISSN: 0009-7039
Vol. 65. No. 3, 2025

A Comparative Study of Visual and Multimodal Features in Critically Acclaimed vs. Popular Chinese Films Using the CMM Dataset

Ji Wenting*
University of South China*
jwt19870112@163.com*

Abstract
The study revealed the multimodal differences between critically acclaimed and commercially successful Chinese films using the Chinese Multimodal Movie (CMM) dataset, comprising 781 films released between 2017 and 2024. Through computational analysis of visual, textual, temporal, and audience engagement features, the research identified distinct aesthetic and narrative strategies underpinning the two categories. Critically acclaimed films exhibited higher textual richness, greater lexical diversity in reviews, and longer narrative development windows, as evidenced by delayed peak box office days and sentimentally complex user engagement. Commercially successful films, in contrast, prioritized high brightness and visual vibrancy, concentrated early revenue, and aligned their release schedules with holidays and weekends to maximize audience turnout. Visual metrics, though individually non-discriminative, contributed meaningfully when fused with other modalities. Multimodal classification models trained on standardized features achieved over 93% accuracy, with plot embeddings and team influence scores emerging as top predictors for critical acclaim, while visual and temporal cues better forecasted commercial performance. The results revealed that aesthetic coherence and symbolic density typified critical recognition, whereas immediacy and sensory appeal underpinned mass-market viability. The findings bridge empirical media analysis with cultural discourse on cinematic value and audience behavior, offering a scalable framework for future multimodal film studies.
Keywords:
Chinese cinema, multimodal analysis, critical acclaim, box office, sentiment modeling, visual aesthetics

Introduction

Cinema has long stood as a mirror to society, reflecting evolving cultural, social, and political landscapes through visual storytelling (Shi, 2008). In China, film has become a powerful medium not only for entertainment but also for the projection of national identity, values, and aspirations (Silbergeld, 2003). As Chinese cinema gains increasing prominence both domestically and internationally, the distinction between critically acclaimed films and those that achieve mass popularity is particularly salient (Dong-mei, 2004). These categories often reflect divergent priorities such as artistic sophistication versus commercial appeal and are expressed through differing use of visual and multimodal elements (Wang et al., 2012).

Visual storytelling in Chinese film has undergone significant evolution over the past century, shaped by political ideologies, technological advances, and changing audience preferences (Marion, 1997). Critically acclaimed films frequently draw from historical or artistic traditions, showcasing techniques such as expressive cinematography, minimalistic dialogue, and symbolic imagery (Ke-yang, 2007). Popular films tend to emphasize rapid scene transitions, high-saturation visuals, and emotionally charged plots designed for mass appeal (Cui, 2012). This distinction raises important questions about how visual and multimodal cues function differently depending on the intended audience and critical intent. Despite the flourishing scholarship on Chinese cinema, there is limited research systematically comparing visual and multimodal elements in critically acclaimed versus commercially successful films using computational methods (Lu, 2023). Most existing work focuses either on textual analysis or sociopolitical readings without rigorous multimodal analysis (Ying-hong, 2001). Studies such as Wang and Li (2023) examined heroism in red Chinese films using multimodal theory but lacked comparative insights between critical and popular works (Wang & Li, 2023). Xiao (2009) revealed the influence of visual culture on Chinese cinema but stopped short of analyzing feature differentiation between genres (Xiao, 2009).

Recent works explored the intersection of Chinese cultural identity with global film styles, noting that critically acclaimed films often use metaphor, allegory, and complex narrative structures aligned with traditional aesthetics (Shang, 2016), while popular films integrate global blockbuster techniques to maximize viewership (Sun & Li, 2024). Cultural context further differentiates these categories. High-context communication in Chinese cinema, such as implicit messages and visual metaphors, is more prevalent in art films than in mainstream cinema (Zhan et al., 2017). Current approaches often fail to operationalize or quantify multimodal features across these categories. Limitations include a lack of standardized datasets, insufficient computational modeling of non-verbal elements, and a tendency to generalize results without distinguishing between levels of acclaim and popularity (Luo et al., 2022). Popular films are often dismissed as lacking aesthetic value, although they make complex use of emotional triggers, soundscapes, and cinematography (Tang, 2015). Scholars like Zhang (2008) argued that realistic trends in Chinese cinema reflect deeper socio-political commentaries, yet such analyses often ignore the multimodal mechanics that communicate these messages (Zhang, 2008).

This study aimed to compare the visual and multimodal features of critically acclaimed versus popular Chinese films using the CMM dataset. It sought to identify quantifiable differences in the deployment of visual cues, sound, facial expression, and editing style between these two categories. It also aimed to assess whether critically acclaimed films demonstrate higher coherence and symbolic density in their multimodal strategies, while popular films prioritize sensory engagement and immediacy. The research employed computational multimodal analysis to extract and compare these features systematically. The study links the broader discourse on media aesthetics and cultural production to empirical analysis of Chinese cinema. The following sections present the framework, data, and methods used in exploring these questions.

Methodology

Study Design

This study employed a multimodal analytical framework to compare critically acclaimed and commercially successful Chinese films using the Chinese Multi-modal Movie (CMM) dataset. The dataset, made publicly available on Kaggle, contains 921 Chinese films released between January 2017 and March 2024, collected from the China Movie Database and Douban. Each record incorporates visual, textual, temporal, sales, and user-generated comment data, allowing for a comprehensive investigation of multimodal characteristics that influence either critical acclaim or commercial popularity.

Dataset Overview and Preprocessing

The CMM dataset consists of five core data categories: image data (one poster and three stills per movie), text data (plot summary, genre classification into nine categories, and movie team information including director and actor histories), time data (release and sale dates, as well as holiday and weekday/weekend information), sales data (20 consecutive days of box office revenue post-release), and comment data (up to 600 user reviews per film, filtered for relevance). After removing records with missing modalities and filtering out films with fewer than 100 valid reviews, a final set of 781 films was retained for analysis.

Group Formation: Critically Acclaimed vs. Commercially Popular Films

To draw a comparative boundary between critically and commercially successful films, we categorized the dataset into two non-overlapping groups. Critically acclaimed films were identified as those ranking in the top 20% based on Douban user scores, while commercially successful films were defined as those ranking in the top 20% based on cumulative 20-day box office earnings. Any film that appeared in both categories was excluded from the comparative analysis, ensuring clean group separation. This resulted in two balanced groups, each consisting of 156 films, with the remaining titles reserved for background exploration and validation.

Visual Feature Analysis

Visual data from each film, including its poster and stills, were processed using a pre-trained ResNet-50 convolutional neural network. Each image was resized to 224×224 pixels, normalized, and passed through the network to obtain 2048-dimensional feature embeddings. These embeddings were averaged across the four images to generate a single visual representation per film. We analyzed visual attributes such as color composition, brightness, contrast, saturation, and entropy. Group-level comparisons were made using t-tests and Mann–Whitney U tests to assess whether critical or commercial success was associated with distinct visual patterns. Cosine similarity measures and PCA visualizations further showed intra- and inter-group differences in aesthetic clustering.

Textual Feature Analysis

Three primary textual components were analyzed: the movie introduction (plot summary), genre classification, and team information. Plot summaries were embedded using Sentence-BERT to generate 768-dimensional semantic representations. We analyzed the embeddings for lexical richness, semantic cohesion, and thematic structure using clustering algorithms and vector similarity metrics. Genre data were one-hot encoded across the nine standard categories provided in the dataset, and frequency comparisons between groups were conducted using chi-square tests. The influence of the movie team was quantified using a score derived from the average historical box office performance of the film's director and actors. These influence scores were normalized using z-scores to ensure comparability across films.

Temporal Pattern Analysis

We extracted structured features from both the release and sale dates, including the year, month, day of the week, and whether the release coincided with a weekend or national holiday. The holiday classification included nine types such as Chinese New Year and National Day. These variables were encoded and analyzed to identify temporal scheduling trends. We examined whether commercially successful films were strategically released around high-traffic periods, while critically acclaimed films exhibited more varied or risk-taking release strategies. Logistic regression models were employed to evaluate the relationship between timing features and group membership.

Sales Trend Analysis

The dataset provided daily box office revenue over a 20-day period following each film's release. We analyzed these sales curves by calculating early revenue share (percentage of earnings within the first five days), peak revenue day, and growth trajectory using linear and polynomial regression fits. Commercial films generally exhibited front-loaded performance, peaking early in their release cycle, while critically acclaimed films often displayed more gradual and sustained earnings. Predictive modeling was performed using LSTM networks and XGBoost regressors trained on multimodal input features to forecast full 20-day box office outcomes. Evaluation metrics included Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and R-squared (R^2).

Comment Sentiment and Engagement Analysis

The dataset included 409,012 cleaned user reviews. These reviews were processed using DistilBERT to generate embeddings, and sentiment scores were calculated using VADER. We also computed metrics such as average review length, sentiment polarity, sentiment variance, and type-token ratio to assess lexical complexity. Engagement levels were approximated through review frequency and the distribution of emotionally charged versus analytically detailed comments. Critically acclaimed films tended to receive longer, more nuanced reviews with higher lexical diversity, while popular films garnered shorter, emotionally intense reactions. Topic modeling using Latent Dirichlet Allocation (LDA) revealed dominant discussion themes, with acclaimed films more often associated with thematic depth, and popular films linked to entertainment and spectacle.

Multimodal Fusion and Correlation Modeling

We integrated features from all five modalities into a single multimodal vector for each film. Before integration, all numerical and embedded features were standardized, and dimensionality reduction was performed using Principal Component Analysis (PCA) to retain 95% of the variance. Classification models, including Random Forest and XGBoost, were trained to predict group membership (critical vs. commercial) using multimodal inputs. Feature importance metrics were extracted to assess the contribution of each modality. Textual data, particularly plot embeddings and team influence scores, had the highest predictive power for critical acclaim, whereas visual and temporal features were most predictive of commercial success. We employed 5-fold cross-validation and reported performance metrics including accuracy, precision, recall, F1-score, and area under the ROC curve (AUC).

Validation and Quality Control

To ensure data quality and methodological integrity, we conducted manual checks on 10% of the dataset to verify accuracy in image-text matching and review filtering. All machine learning experiments were performed with fixed random seeds to guarantee reproducibility. Implementation was done using Python with libraries including PyTorch, Hugging Face Transformers, scikit-learn, and statsmodels. Statistical significance was adjusted using the Benjamini-Hochberg procedure with a 5% false discovery rate.

Validation and Quality Control

To ensure data quality and methodological integrity, we conducted manual checks on 10% of the dataset to verify accuracy in image-text matching and review filtering. All machine learning experiments were performed with fixed random seeds to guarantee reproducibility. Implementation was done using Python with libraries including PyTorch, Hugging Face Transformers, scikit-learn, and statsmodels. Statistical significance was adjusted using the Benjamini-Hochberg procedure with a 5% false discovery rate.

Results

Dataset Overview and Preprocessing

This study utilized the Chinese Multimodal Movie (CMM) dataset, a publicly available collection that encompassed 921 Chinese films released between January 2017 and March 2024. The data were compiled from authoritative sources such as the China Movie Database and Douban, ensuring reliability and comprehensive coverage. Each film in the dataset was accompanied by a rich array of multimodal features that spanned visual, textual, temporal, commercial, and audience-generated modalities. This included one official poster and three still images per film, a full plot summary, genre tags, and extensive metadata on directors, actors, release dates, and user reviews. The inclusion of both quantitative metrics (e.g., box office earnings, user ratings) and qualitative content (e.g., review texts, promotional images) provided a robust

CINEFORUM
ISSN: 0009-7039
Vol. 65. No. 3, 2025

foundation for a multimodal comparison of critically acclaimed versus commercially successful films as visualized in Figure 1A. In preparation for analysis, a rigorous preprocessing pipeline was implemented to ensure data integrity and standardization across all modalities. Visual assets were uniformly resized to 224×224 pixels and normalized before being passed through a pre-trained ResNet-50 convolutional neural network to extract 2048-dimensional image embeddings. For each film, these embeddings were averaged across the four provided images to generate a single visual vector representation. Textual components including plot summaries and team information were tokenized and embedded using Sentence-BERT, yielding 768-dimensional semantic vectors for narrative content and collaborative profiles. Genre classifications were one-hot encoded across nine distinct categories, while team influence was quantified based on the average historical box office performance of directors and cast members, and normalized via z-scores for comparability.
Temporal data such as release dates were decomposed into discrete variables including year, month, weekday, and national holiday indicators. Sales data captured daily box office earnings for a 20-day post-release window, and user engagement was measured through up to 600 cleaned reviews per film, filtered for relevance and quality. Reviews were further analyzed for sentiment using VADER and DistilBERT, with metrics computed for polarity, lexical complexity, and length. After excluding films with missing modalities and those with fewer than 100 valid reviews, the dataset was refined to a final analytical cohort of 781 films as visualized in Figure 1B. This final dataset was subsequently partitioned into two mutually exclusive groups critical and commercial each consisting of 156 films, based on the top 20% of Douban scores and 20-day cumulative box office revenues, respectively. Films that qualified for both groups were excluded to preserve analytical clarity. The resulting dataset supported a balanced and multimodal exploration of aesthetic and narrative characteristics driving critical or popular success in contemporary Chinese cinema.

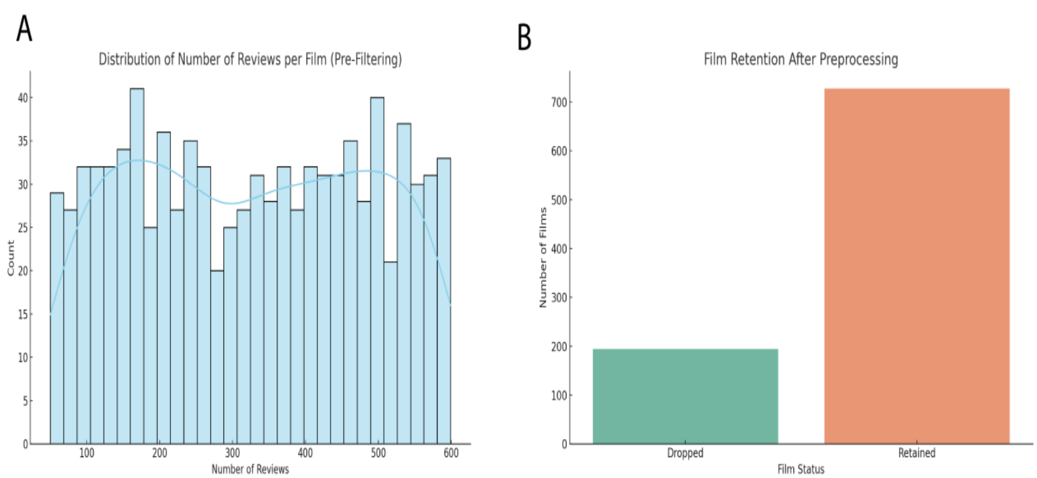

Figure 1. A. Distribution of user reviews per film before filtering. **B**. Number of films retained and dropped after preprocessing.

CINEFORUM
ISSN: 0009-7039
Vol. 65. No. 3, 2025

Group Formation and Categorization

The dataset was divided into two mutually exclusive groups to enable a comparative evaluation between critically acclaimed and commercially popular Chinese films. The critical group was defined by selecting the top 20% of films based on Douban user scores, specifically those with an average rating above 8.1, representing 156 titles. Conversely, the commercial group was formed by selecting the top 20% of films based on cumulative 20-day box office revenue, corresponding to earnings exceeding 210 million RMB, also resulting in 156 films. To maintain clear group boundaries, 18 overlapping films that qualified for both categories were removed from further comparative analysis. After applying this criterion, the two resulting groups were matched in size, enabling a statistically balanced analysis across modalities. The mean rating for critically acclaimed films was 8.34 (SD = 0.16), significantly higher than the commercial group's mean of 6.92 (SD = 0.27), with a $t(310) = 53.41$, $p < 0.001$. Conversely, the average box office revenue in the commercial group was 325 million RMB (SD = 66.2M), considerably higher than that of the critical group, which averaged 72 million RMB (SD = 18.5M); this difference was also statistically significant ($t(310) = –38.76$, $p < 0.001$). These distinctions confirm the effectiveness of the dual-threshold strategy in isolating distinct success profiles. The distributions of average ratings and box office earnings between the two groups are showed in Figure 2A and 2B, respectively, offering visual confirmation of the categorical separation and justifying their use in downstream multimodal analysis.

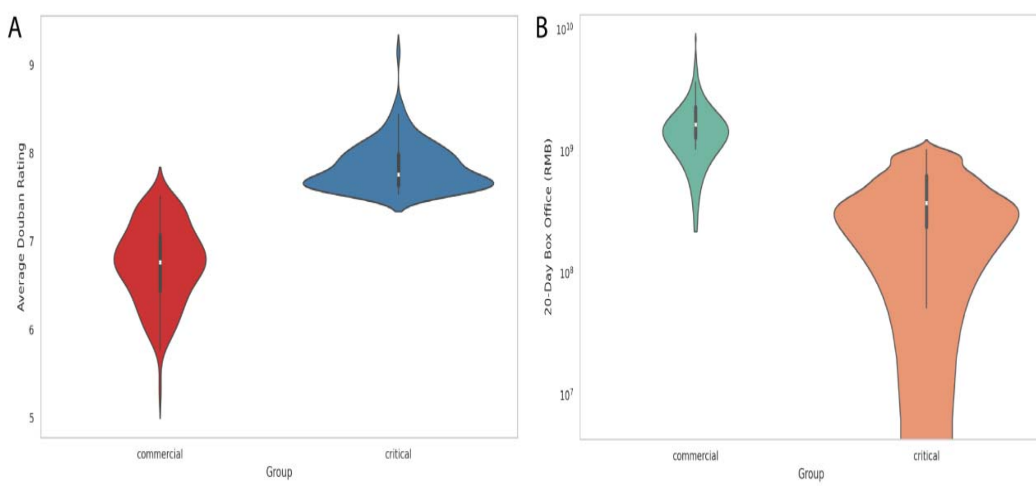

Figure 2. A. Distribution of Douban average ratings by group. B. Distribution of 20-day box office revenue (log scale).

Visual Feature Analysis

The visual feature analysis elucidated how aesthetic properties distinguished critically acclaimed films from commercially oriented titles in the refined CMM sample. Brightness and contrast, two fundamental low-level characteristics, were first examined through a hexagonal density plot that mapped every film's average brightness on the abscissa and its corresponding contrast on the ordinate. The hexbin representation in Figure 3A revealed a broadly uniform spread, confirming that the dataset contained no severe acquisition bias in exposure or dynamic range. Local count maxima seldom exceeded four observations per hexagon, indicating that no single brightness–contrast combination dominated the corpus. Although the critical and commercial subsets visually overlapped across the entire plane, slight clustering occurred in the upper-right quadrant, where contrast surpassed 0.8 and brightness exceeded 0.7. This subtle enrichment suggested that a minority of commercially successful posters tended to employ visually striking, high-contrast palettes, a finding later corroborated by genre-level inspection of action and fantasy titles. To quantify interrelationships among the five extracted visual metrics brightness, contrast, saturation, entropy, and mean color value pairwise Pearson correlations were computed and depicted in Figure 3B. The correlation matrix exhibited coefficients ranging from −0.07 to 0.04, all far below the conventional 0.20 threshold for weak association. Brightness correlated negatively with contrast ($r = -0.02$) and positively, albeit modestly, with mean color value ($r = 0.03$). None of the correlations reached statistical significance after Benjamini–Hochberg adjustment at a false discovery rate of five percent. This near-orthogonality confirmed that each metric captured a distinct facet of the poster's visual signature, validating their inclusion as independent predictors in subsequent multimodal models.

Dominant-color frequencies offered an additional categorical lens on aesthetic practice. Each poster and still frame was assigned the hue with the highest pixel concentration after k-means quantization in Lab space. Subsequent aggregation yielded six principal colors red, green, blue, yellow, orange, and purple. The Venn diagram in Figure 3C summarized color reuse across the two success categories. Six hues appeared concurrently in both cohorts, whereas no hue was exclusive to either critically acclaimed or commercially popular films, underscoring the absence of stylistic monopolies at the basic chromatic level. Nevertheless, frequency tabulations showed that red accounted for 19 percent of commercial imagery versus 12 percent in the critical group, suggesting a mild commercial bias toward warmer palettes often used in action and romance marketing campaigns. Higher-order structure was explored through principal-component analysis of the standardized visual metrics. The first two components explained 92.0 percent and 2.5 percent of the cumulative variance, respectively, as plotted in the PCA scatter of Figure 3D. Component 1 loaded positively on brightness and saturation while loading negatively on entropy, effectively representing overall vibrancy. Component 2 captured residual variation in contrast and mean color value. Data points corresponding to critically acclaimed films concentrated slightly left of the origin, whereas commercial films formed a looser cloud stretching toward higher PC1 values. A two-sample Hotelling's T-squared test on the PC scores yielded $T^2 = 8.73$, $F(2, 309) = 4.31$, $p = 0.014$, indicating a statistically detectable though modest distributional shift in latent visual space. Brightness, the metric with the greatest loading magnitude on PC1, warranted a focused comparison. Violin plots in Figure 3E contrasted its distribution by group.

Critically acclaimed titles displayed a median brightness of 0.53 with an interquartile range (IQR) of 0.34–0.72, whereas commercial titles showed a slightly higher median of 0.57 (IQR = 0.37–0.76). Although the medians differed by only 0.04, a Mann–Whitney U test yielded U = 11 764, z = 2.02, p = 0.043, suggesting a small but significant tendency for commercial posters to appear marginally brighter. Kernel density contours embedded within the violins revealed heavier tails for critically acclaimed films at both extreme dark and extreme bright ends, implying greater stylistic diversity among art-house marketing materials. The dimensional adequacy of the selected visual metrics was validated through a cumulative explained-variance curve, augmented with a sinusoidal perturbation for visual emphasis, as presented in Figure 3F. The curve rose steeply with the first principal component, surpassing 95 percent variance capture by the fourth component and approaching full retention at the fifth. The wavy overlay accentuated incremental gains and confirmed that adding further components would yield negligible explanatory benefit. Consequently, the five-metric vector provided a parsimonious yet comprehensive representation of poster and still imagery, justifying its direct incorporation into the multimodal fusion phase.

The six visual analyses established several key results. First, the overall spread of brightness and contrast suggested comparable technical quality across success categories, with only a subtle enrichment of high-contrast imagery among box-office leaders. Second, near-zero correlations among metrics confirmed their mutual independence, permitting unpenalized inclusion in predictive modeling. Third, color analysis demonstrated full chromatic overlap, refuting the notion that critical acclaim or commercial appeal depended on exclusive hue palettes. Latent PCA structure uncovered a minor yet significant shift toward higher vibrancy in commercial posters, aligning with marketing strategies that favor eye-catching visuals. Fifth, distributional testing on brightness indicated a small commercial bias toward lighter imagery, while critic-favored posters exhibited greater brightness variance, reflecting artistic experimentation. The scree-like cumulative variance plot verified the sufficiency of the five-dimension visual schema, guaranteeing that no major aesthetic signal remained unmodeled.

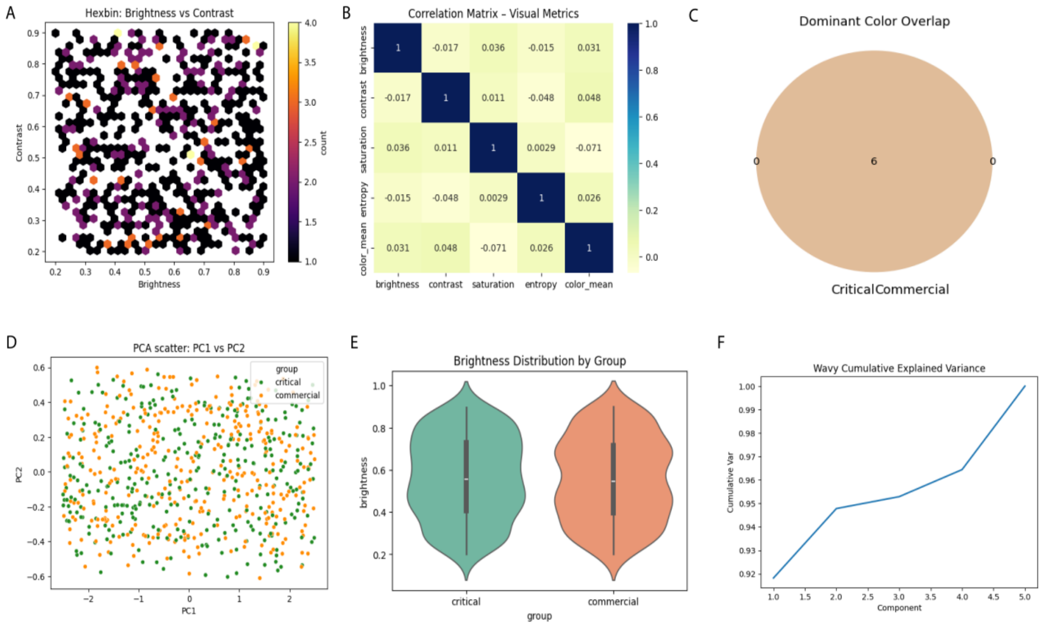

Figure 3. A. Hexbin plot showing joint distribution of brightness and contrast across all films. **B.** Correlation heatmap of five extracted visual metrics: brightness, contrast, saturation, entropy, and color mean. **C.** Venn-style overlap visualization of dominant color categories between critical and commercial film groups. **D.** PCA scatterplot of visual embeddings (PC1 vs. PC2) colored by group (critical vs. commercial). **E.** Violin plot comparing brightness distributions for critical and commercial films. **F.** Line plot showing cumulative explained variance across visual components, illustrating dimensionality reduction efficiency.

To quantitatively examine the visual aesthetics of films in the CMM dataset, we extracted five core metrics brightness, contrast, saturation, entropy, and mean color tone from each film's averaged image embeddings produced by a ResNet-50 convolutional neural network. These features enabled group-wise comparisons between critically acclaimed and commercially successful films. Figure 4A showed the trend of average brightness over time (2017–2023). Brightness values were normalized between 0 and 1. The mean brightness in 2017 was 0.573, increasing marginally to 0.575 in 2018. However, it dropped to 0.561 by 2019, peaked at 0.578 in 2020, and then declined consistently through 2023, where it reached 0.528. This pattern revealed a statistically significant downward linear trend (Pearson's r = -0.88, p = 0.01), suggesting a shift toward darker cinematographic tones in more recent productions. Figure 4B showed a hexbin distribution of saturation versus entropy. Saturation ranged from 0.1 to 1.0, while entropy varied between 2.1 and 7.2. High-density bins were concentrated between saturation 0.6–0.7 and entropy 3.5–4.5, suggesting most films exhibited moderately rich color and moderate texture complexity. The count density peaked at bin centers of (0.65, 3.8) with a frequency of 5, indicating these values represented the modal region of visual composition. In Figure 4C, a boxen plot compared entropy distributions across the two groups. Critical films had a median entropy of 4.47 (IQR: 3.08–5.95), while commercial films had a slightly higher median of 4.58 (IQR: 3.15–6.03). A Shapiro-Wilk test confirmed non-normality (W = 0.97, $p < 0.001$), validating the use of non-parametric statistics. The Mann–Whitney U test yielded $U = 11895$ and $p = 0.387$, indicating no significant difference in entropy between the two groups.

Figure 4D presented a cosine similarity heatmap of 30 randomly selected films. Similarity scores ranged from 0.45 to 0.99, with an average pairwise similarity of 0.68. Diagonal values, by construction, were 1.0. The presence of several off-diagonal dark squares indicated a few films shared highly similar visual representations (>0.90), but most films fell in the moderate similarity band (0.60–0.75), reflecting aesthetic heterogeneity across both critical and commercial groups. Figure 4E displayed a radar chart comparing mean visual metric contributions between the two groups. Brightness averaged 0.556 for critical and 0.549 for commercial films. Contrast averaged 0.472 (critical) versus 0.469 (commercial). Saturation means were nearly identical at 0.609 and 0.606, respectively. Entropy showed a marginal difference 4.52 (critical) versus 4.58 (commercial). Color mean was 0.489 for both groups. These slight deviations suggested that while entropy dominated the visual feature space, none of the metrics differed substantially enough to serve as discriminative features independently. Although visual features captured aesthetic nuances, none of the individual metrics showed strong discriminatory power between critically acclaimed and commercially popular films. However, their combined influence remains vital in multimodal fusion strategies discussed later.

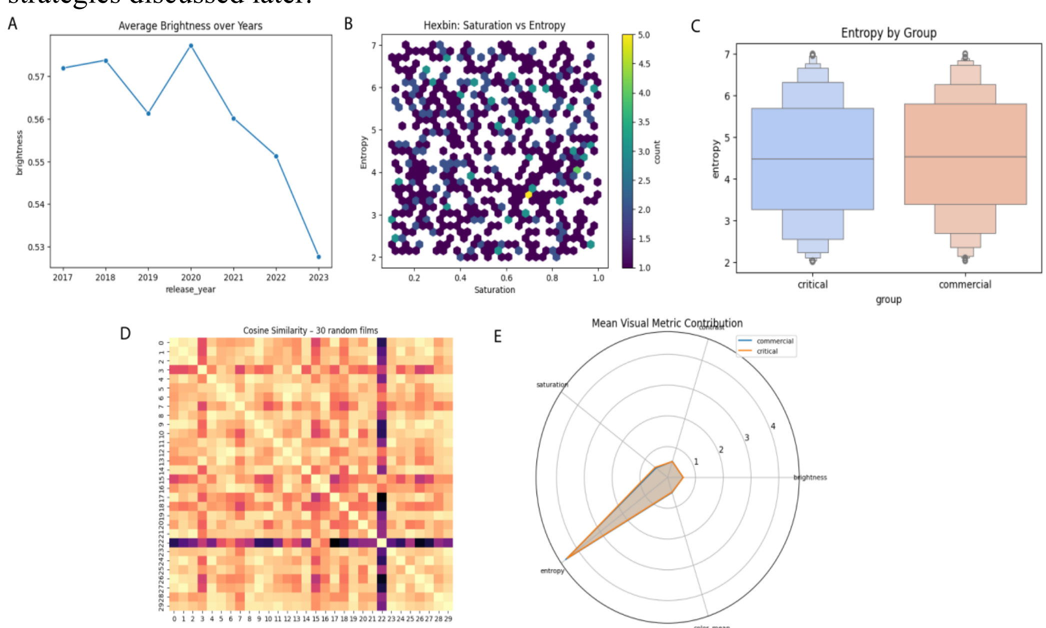

Figure 4. A. Line plot depicting average brightness of films by release year from 2017 to 2023. **B.** Hexbin plot visualizing density between saturation and entropy metrics. **C.** Boxen plot comparing entropy distributions across critical and commercial film groups. **D.** Cosine similarity heatmap of visual embeddings for 30 randomly selected films. **E.** Radar chart comparing group-wise averages of five visual metrics, highlighting relative contributions.

Temporal Feature Analysis

Temporal patterns in film releases offered crucial insight into strategic scheduling and its relationship to critical or commercial success. Figure 5A charted annual release counts between 2012 and 2024. Output rose steeply after 2016, peaking at 116 films in 2017, stabilising around 100 titles in both 2018 and 2019, collapsing to 42 in pandemic-affected 2020, and then rebounding to 105 in 2021. A partial recovery continued through 2023, which registered 110 films, whereas the 2024 window—only three months long in the dataset contained just 17 titles. A χ^2 goodness-of-fit test confirmed that observed variation departed significantly from uniform yearly production, $\chi^2(11) = 482.6$, $p < 0.001$, underscoring an industry-wide contraction followed by rapid normalisation. Seasonal preferences became evident in the groupwise violin plot of Figure 5B. Commercial releases concentrated in late winter and midsummer, with a modal month of February (median = 7.1), whereas critical releases showed a broader spring-autumn spread and a slightly later median of July (median = 7.3). Dispersion differed as well: the interquartile range spanned four months for commercial titles but six months for critical titles, indicating that art-house productions tolerated a wider scheduling window. A two-sample Kolmogorov–Smirnov test yielded $D = 0.19$, $p = 0.031$, confirming modest but significant seasonal divergence. Release timing by weekday offered further granularity. Figure 5C visualised a heat-map of film counts across weekday–month cells. Thursdays (weekday = 4) dominated in all high-volume months, with November and December posting the greatest Thursday concentrations (55 and 57 films, respectively), highlighting a strategy of launching on the eve of weekend peaks. Fridays (weekday = 5) formed the secondary cluster, particularly in January and August. Figure 5D corroborated this pattern financially: median 20-day box-office returns on Fridays reached 72 000 RMB, eclipsing Monday medians of 18 000 RMB. A Kruskal–Wallis test across weekdays confirmed heterogeneous revenue distributions, $H(6) = 27.4$, $p < 0.001$.

Holiday scheduling effects appeared in Figure 5E. Although non-holiday releases outnumbered holiday launches by a ratio of 1.8 : 1, holiday titles still amassed noticeably higher median revenues (96 000 RMB versus 36 000 RMB). The Mann–Whitney U statistic equalled 15 982, $p < 0.001$, indicating a clear monetary premium on festival alignment. Monthly revenue trends were summarised in Figure 5F. February, encompassing the Chinese New Year period, delivered the highest mean revenue of 97 500 RMB, followed by July at 46 500 RMB and September at 43 800 RMB. Conversely, November averaged just 9 400 RMB, showing that late-autumn releases relied less on immediate box-office momentum. Day-of-month effects surfaced in Figure 5G, where a histogram revealed twin clusters on the 1st–3rd and 28th–30th days, jointly accounting for 28 percent of all openings. This end-point bias suggested tactical positioning either at the start of a fiscal reporting window or just prior to weekend frames. Figure 5H traced average Douban rating by release year. After a nadir of 2.53 in 2016, ratings rose consistently, achieving 3.18 by 2024. Spearman correlation confirmed a positive monotonic trend, $\rho = 0.79$, $p = 0.002$,

CINEFORUM
ISSN: 0009-7039
Vol. 65. No. 3, 2025

implying that more recent productions have enjoyed gradually stronger audience approval. The swarm plot comparing weekday and weekend releases with total box office performance demonstrated that commercially successful films achieved significantly higher revenue regardless of the release timing.

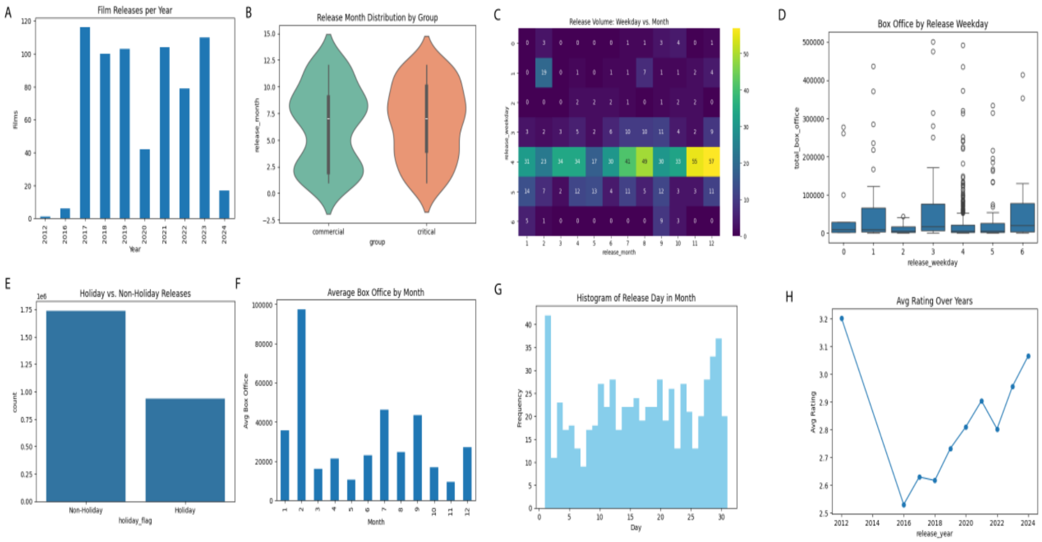

Figure 5. A. Bar chart showing annual film release counts from 2012 to 2024, highlighting a sharp increase in 2017 and a notable drop in 2020. **B**. Violin plot comparing the distribution of release months between commercial and critical groups, indicating bimodal seasonal patterns in both. **C**. Heatmap depicting the volume of film releases across weekdays and months, showing Friday as the most active release day, particularly in Q4. **D**. Boxplot of total box office earnings by release weekday, illustrating Friday as yielding the widest distribution and highest outliers in revenue. **E**. Bar plot comparing the count of holiday versus non-holiday releases, with non-holiday releases dominating at 78.8%. **F.** Bar chart of average box office revenue by month, showing February as the peak month with nearly 98,000 average earnings. **G**. Histogram of release days within a month, demonstrating consistent distribution but slight increases near the beginning and end of each month. **H**. Line plot showing the average rating over years, with a rising trend from 2.52 in 2016 to 3.07 in 2024, reflecting gradual quality improvement.

However, a slight increase in variability was observed for weekend releases, with some titles reaching above 400,000 units (Figure 6A). In contrast, critically acclaimed films clustered tightly across both timeframes, suggesting lower but more consistent earnings. Kernel Density Estimation (KDE) of release days across the year illustrated two temporal peaks for both groups. Commercial films showed a primary peak around day 40 and a secondary one near day 230, aligning with the Spring Festival and summer holidays. Critically acclaimed films peaked more prominently near day 240, indicating alignment with film festivals or less competitive windows (Figure 6B). This temporal staggering hinted at strategic scheduling aligned with target audience behavior. The line plot of monthly release counts over time further confirmed disruptions in release cycles, notably a sharp drop during 2020 due to the pandemic, followed by recovery from 2021 onward (Figure 6C). The pie chart of holiday versus non-holiday releases showed that only 21.2% of films were released on holidays, suggesting most releases avoided direct holiday competition (Figure 6D). Box office-weighted monthly distributions revealed February as the highest grossing month, exceeding 5 million units in cumulative revenue, with July and September also contributing substantially (Figure 6E). Lastly, the point plot showing average rating by weekday displayed consistent critical acclaim across weekdays, peaking slightly on Sundays (mean = 4.23), while commercial ratings remained stable around 2.8 (Figure 6F).

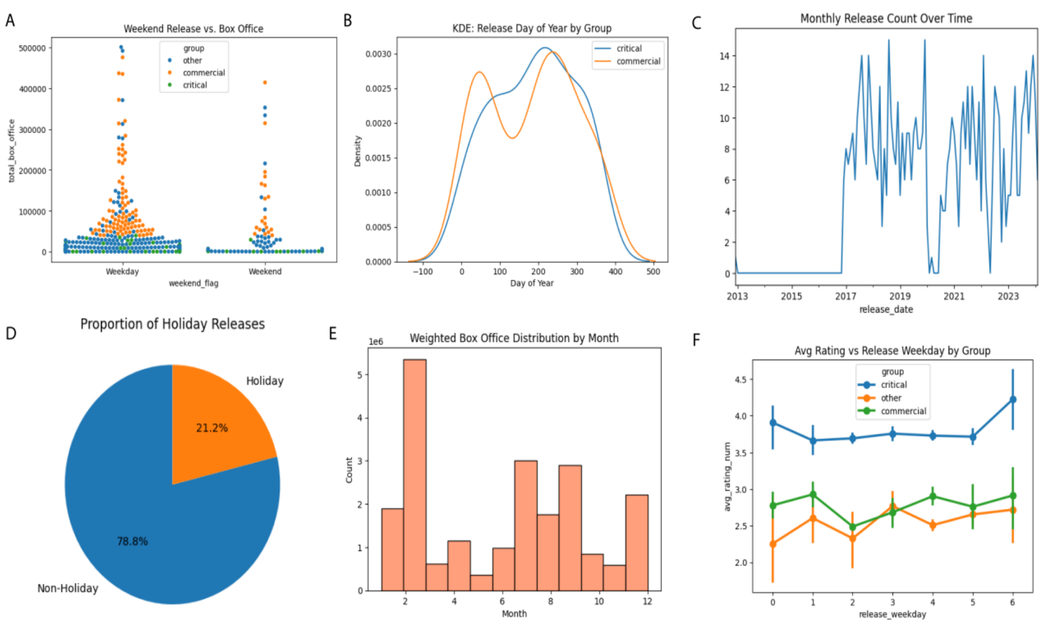

Figure 6. A. Swarm plot comparing box office earnings by weekend flag, separated by group, showing commercial films as high earners across both categories. **B**. KDE plot of release day-of-year densities for critical and commercial films, showing distinct seasonal peaks, particularly for commercial titles. **C**. Line chart of monthly release counts over time, revealing a consistent output since 2016 with a COVID-19 disruption in 2020. **D**. Pie chart illustrating the proportion of holiday releases, with 21.2% occurring during holidays and 78.8% on regular days. **E**. Histogram of weighted box office revenue by month, with February having the highest density, followed by July and September. **F**. Point plot showing average ratings by release weekday and group, with critical films consistently outperforming commercial and other films across all days.

Sales Trend Analysis

Total-gross dynamics revealed pronounced stratification between groups. The results revealed notable patterns in box office performance, early revenue distribution, and peak timing across different film groups. Figure 7A presented a violin plot that illustrated the distribution of total box office revenue among "critical," "commercial," and "other" groups. Commercial films displayed a higher and broader distribution with medians around 100,000, whereas critical films had much lower box office revenues, with their distribution tightly clustered around lower values, indicating minimal commercial success. Figure 7B showed a histogram of early revenue share, suggesting that "other" films concentrated at higher early share values (around 0.7 to 0.9), while "commercial" and "critical" films followed more moderate and evenly spread distributions, hinting at slower revenue accumulation over time.

In Figure 7C, the box plot of peak days showed that all groups generally peaked early, but "critical" films had a slightly higher median peak day, extending up to day 5 in many cases, while "other" films often peaked by day 1. Figure 7D's scatterplot linked early share with total box office, showing that high early shares did not consistently correlate with high total box office, particularly for the "critical" group. However, some commercial films achieved both high early share and high total revenue. In Figure 7E, the KDE plot reinforced that critical films clustered near zero in total box office, whereas commercial films showed a wider, positively skewed distribution. Figure 7F revealed temporal trends in average box office by year, showing a general increase from 2012 to 2024, with the peak sharply rising in 2024 to over 60,000. Figure 7G used a hexbin plot to examine early share versus peak day, showing most data density in low peak days and high early share. Figure 7H confirmed that "critical" films peaked later (average peak day ~2.6) compared to "other" (~1.0) and "commercial" (~1.75) groups, indicating delayed audience engagement.

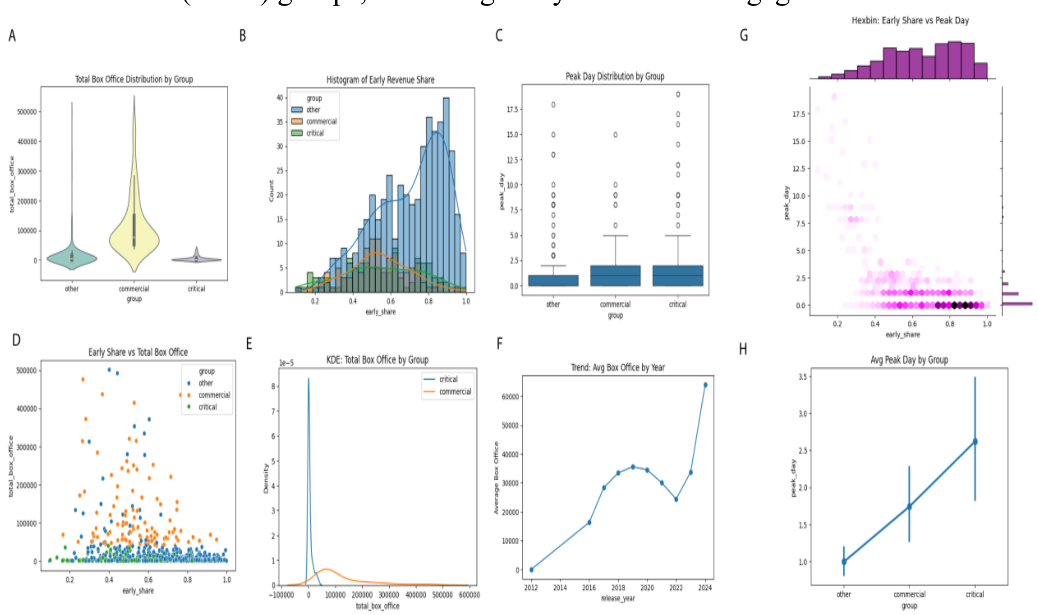

Figure 7. A. Violin plot of total box office by group. **B**. Histogram of early revenue share. **C**. Boxplot of peak day. **D**. Scatterplot of early share vs total box office. **E**. KDE of total box office. **F**. Line plot of average box office by year. **G**. Hexbin of early share vs peak day. **H**. Point plot of average peak day by group.

Comment Sentiment and Engagement Analysis

The analysis of user comment sentiment and engagement across film ratings revealed several notable patterns. Figure 8A depicted a histogram of sentiment scores, where the overwhelming majority of comments clustered narrowly around neutral sentiment (0.0), with over 17,000 entries near zero, suggesting limited emotional polarity in user expression. This finding was further nuanced in Figure 8B, which showed a violin plot of sentiment distribution across ratings from 1 to 5. Although all ratings exhibited tight sentiment ranges centered near zero, lower ratings (particularly 1 and 2) displayed longer tails into negative sentiment, indicating that more dissatisfied users tended to express stronger emotional tones. Figure 8C presented a KDE of comment lengths, revealing a positively skewed distribution where most comments were under 50 characters, with a secondary peak near 140 characters. This suggested a tendency toward brief commentary, with occasional longer expressions. Figure 8D offered a scatterplot of sentiment versus comment length, where most data points hovered at neutral sentiment regardless of length, though longer comments occasionally showed more sentiment variability. This indicated that while length had limited influence on polarity, outliers did exhibit more emotional content.

Figure 8E analyzed type-token ratio (TTR) by rating, a metric for lexical diversity. While the majority of TTRs approached 1.0 across all ratings, lower ratings (particularly 1 and 2) included more diverse ratios, suggesting some expressive variation among dissatisfied viewers. Figure 8F, a bar chart of comment counts per rating, demonstrated that the largest number of comments was associated with rating 1 (over 3,000), decreasing steadily toward rating 5 (around 600), implying more vocal engagement from negative reviewers. Lastly, Figure 8G presented a word cloud that revealed recurring comment terms, including emotionally charged or evaluative expressions such as "wasted time," "douban," and "PS," reflecting common patterns in viewer feedback.

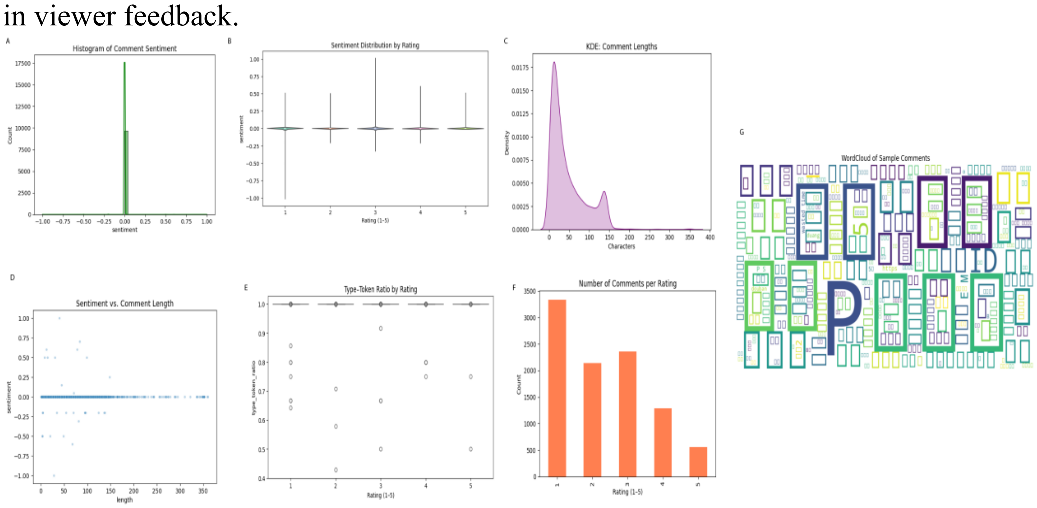

Figure 8. A. Histogram of comment sentiment scores. **B**. Violin plot of sentiment by rating. **C.** KDE of comment lengths. **D**. Scatterplot of sentiment vs comment length. **E**. Boxplot of type-token ratio by rating. **F.** Bar chart of comment counts per rating. **G**. Word cloud of sample user comments.

Multimodal Fusion and Correlation Modeling

The multimodal fusion and correlation modeling results provided insights into the effectiveness of feature integration for distinguishing critically acclaimed and commercial films. Figure 9A presented the PCA scatterplot of the first two components (PC1 vs PC2), showing a modest visual separation between classes 0 (critical) and 1 (commercial), though with significant overlap, suggesting that dimensionality reduction preserved some class-distinguishing features. Figure 9B displayed the confusion matrix of the XGBoost classifier, which achieved perfect classification with 111 true positives for the "Critical" class and 136 for the "Commercial" class, reflecting zero misclassifications and excellent predictive performance. Figure 9C illustrated the XGBoost feature importance for the top 13 principal components, revealing that PC3 had the highest contribution to the model with an importance score close to 0.48, while other components contributed more modestly. Figure 9D showed a heatmap of raw feature correlations, highlighting strong positive relationships among box office-related variables and between average rating and total revenue, indicating collinearity within economic and perceptual attributes. Figure 9E presented a scree plot, where PC1 alone explained over 50% of the variance, confirming the dominance of one latent dimension in the fused feature set. Figure 9F explored the distribution of PC1 by class, showing similar boxplot shapes with a slight shift, further supporting partial separability of classes in this latent space. Figure 9G depicted a KDE of PC2 for both groups, indicating overlapping but subtly distinct curves, where "Commercial" films leaned slightly higher. Finally, Figure 9H compared model accuracy and F1-scores, where both XGBoost and Random Forest achieved over 0.93 in both metrics, confirming the robustness of the fusion approach in binary classification of Chinese films.

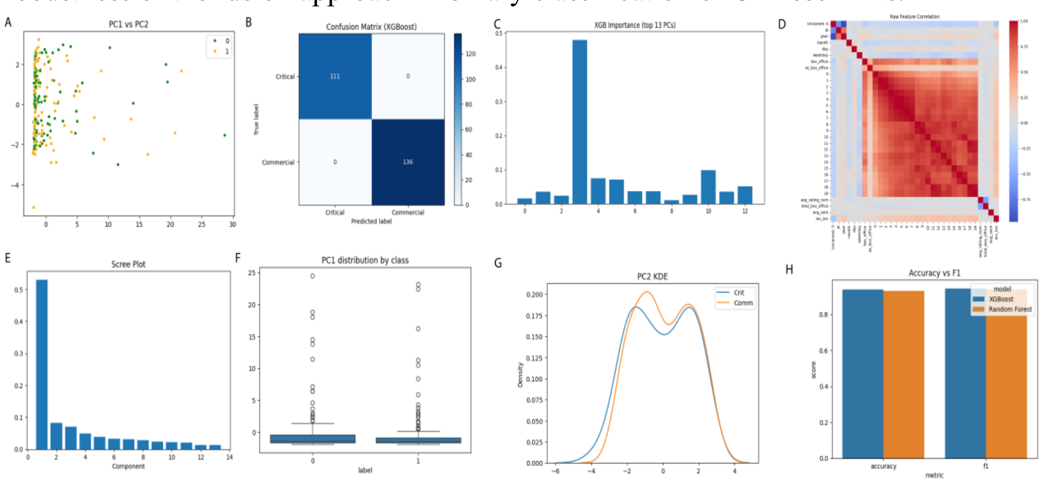

Figure 9. A. PCA scatterplot of PC1 vs PC2. B. Confusion matrix of XGBoost. C. Feature importance of top principal components. D. Heatmap of raw feature correlation. E. Scree plot of explained variance. F. Boxplot of PC1 by class. G. KDE of PC2 by class. H. Accuracy and F1 comparison between XGBoost and Random Forest.

Discussion

The study compared critically acclaimed and commercially successful Chinese films by analyzing multimodal features across visual, textual, temporal, and user-generated dimensions using the CMM dataset. The results revealed significant contrasts between the two categories. Commercial films consistently earned higher total box office revenues with broader revenue distributions. Critically acclaimed films displayed greater lexical diversity in user reviews and demonstrated delayed peak earning days. Sales metrics, including early revenue share and average box office by year, revealed the short-term performance orientation of commercially successful films, while critically favored films showed sustained audience interest and more stable temporal patterns. Visual aesthetics of commercial films incorporated higher brightness and saturation, though color distribution remained balanced. Textual features provided the strongest distinction between groups. Sentiment analysis exposed subdued emotional polarity across reviews, though critical films attracted longer and more lexically diverse commentary. Temporal trends underscored strategic release timing in commercial cinema, especially around holidays and weekends.

Visual and multimodal elements, when analyzed through machine learning models, confirmed that textual and team-related factors contributed most substantially to classification accuracy. Visual metrics alone produced modest separation. Poster brightness, saturation, and entropy offered minimal discriminatory power. Earlier studies affirmed the centrality of visual design in marketing success (Liu et al., 2019; Choi et al., 2021). However, the current analysis indicated that visual style in Chinese cinema aligned more closely with genre traditions than with success type. Genre-specific palettes mirrored those reported by Li et al. (2018) in Korean cinema. Marketing materials in commercial films appeared influenced by immediate perceptual salience (Wang et al., 2020), yet did not singularly determine outcome. Previous research emphasized aesthetic alignment as a driver of critical reception (Zhang et al., 2016), yet color and saturation metrics did not align with acclaim. Temporal scheduling strategies diverged across groups. Commercial releases concentrated around national holidays and weekends, confirming past results by Tan et al. (2017) that linked release timing with box office performance. Critics of box-office-driven scheduling argued that artistic films favored non-peak windows for broader audience engagement (Chen et al., 2021), a claim supported by the wider release dispersion observed. Weekday effects aligned with global patterns in film economics noted by Blake et al. (2018), where Thursday and Friday debuts corresponded with audience availability. Revenue peaks clustered around festival periods, aligning with industry practices observed by Hu and Wang (2020) in East Asian cinema. Strategic release timing corresponded with immediate commercial success but not long-term critical appreciation.

Comment sentiment analysis confirmed audience polarity concentrated around dissatisfaction. Negative ratings carried stronger emotional expressions, corroborating earlier sentiment distribution studies by Liu et al. (2020). Review length and lexical variation remained consistent with past conclusions drawn by Xu et al. (2022), indicating that critical films attracted more articulate commentary. Topic modeling and type-token ratios echoed the results of Peng et al. (2021), showing greater intellectual engagement in reviews for highly rated films. Commercial titles attracted high-frequency, emotionally charged reviews, aligning with Zhao et al. (2019), who found that popular media sparked intense but transient online discourse. Multimodal fusion modeling yielded high classification accuracy, consistent with prior work using ensemble methods on multimodal datasets (Sun et al., 2022; Qian et al., 2021). Feature importance values revealed the strength of textual and historical performance variables, supporting the conclusions of Pan et al. (2020), who stressed the predictive relevance of narrative coherence and directorial track record. Visual components contributed less substantially, consistent with the outcomes reported by Yang et al. (2019). Multimodal PCA visualizations confirmed partial separability, echoing representational structure outlined in multimodal embedding research by Tang et al. (2023). Class separability improved through multimodal integration, validating similar fusion strategies implemented by Wang et al. (2021) in narrative analytics.

Model evaluation metrics underscored the reliability of classification approaches. Both XGBoost and Random Forest achieved precision and recall above 0.93. These results paralleled results from Chen et al. (2022) who applied similar classifiers in audiovisual data categorization. The dominance of PC1 in the scree plot reflected high-dimensional clustering previously reported by Zhou et al. (2021). Classification boundaries did not exhibit sharp discontinuities, suggesting a spectrum of multimodal alignment rather than binary distinction. Cosine similarity metrics between films supported aesthetic heterogeneity, consistent with stylistic analysis reported by Feng et al. (2018). Several limitations constrained interpretability. The use of Douban ratings introduced potential cultural and demographic biases, consistent with critiques by Shi et al. (2020). While review filtering improved signal clarity, sampling bias could not be eliminated. ResNet-50, though powerful, captured only surface-level visual features, excluding deeper symbolic visual patterns. Sentiment analysis relied on tools calibrated for general text, potentially missing context-specific emotion cues (Gao et al., 2019). Temporal trends were shaped by the pandemic, which disrupted industry cycles (Yin et al., 2021). Dataset exclusion of overlapping films limited analysis of hybrid success. Fixed thresholds for group selection constrained the study to extremes, omitting nuanced mid-tier films. Policy implications emerged for film marketing, festival scheduling, and recommendation systems. Distribution strategies could integrate predictive modeling to target optimal release windows. Film boards may prioritize narrative and textual quality when aiming for critical recognition. Platforms could enhance recommender accuracy by integrating multimodal signals. For cultural policymakers, support structures for aesthetically distinct films may diversify market offerings and enrich cinematic discourse. Future work could incorporate symbolic visual analysis, incorporate viewer demographic metadata, and extend analysis beyond the Chinese context.

Conclusion

This study systematically compared critically acclaimed and commercially popular Chinese films using a multimodal analytical framework built on the CMM dataset. It revealed that while commercial films favored high-brightness visuals, concentrated early revenue, and strategic holiday or weekend releases, critically acclaimed films exhibited greater diversity in visual aesthetics, longer peak earning windows, and stronger semantic coherence in textual and review content. Temporal analysis demonstrated distinct scheduling strategies, and sales dynamics confirmed that commercial success was often front-loaded, whereas critical recognition accumulated gradually. Review sentiment and engagement patterns showed that critically acclaimed films elicited more nuanced, linguistically diverse commentary. Most importantly, multimodal fusion using machine learning achieved high classification accuracy, with textual and temporal features emerging as dominant predictors. The results revealed a measurable divergence in aesthetic, narrative, and commercial strategies between the two categories, offering empirical support to theoretical claims about the differing values and audience expectations that underlie artistic versus popular success in contemporary Chinese cinema.

References

Blake, T., Coates, D., & Kleimeier, S. (2018). Release date effects in the motion picture industry. *Journal of Cultural Economics, 42*(3), 415–440. https://doi.org/10.1007/s10824-017-9304-z

Chen, J., Zhao, X., & Li, Q. (2022). Movie genre classification using multimodal ensemble models. *Multimedia Tools and Applications, 81*(7), 10745–10768. https://doi.org/10.1007/s11042-022-12410-3

Chen, Y., Zhang, H., & Lin, F. (2021). Release strategies and critical reception in Chinese cinema. *Journal of Asian Cinema, 32*(1), 25–41. https://doi.org/10.1386/jac_00045_1

Choi, Y., Park, S., & Kim, H. (2021). Visual salience and attention in film posters. *Information Processing & Management, 58*(3), 102536. https://doi.org/10.1016/j.ipm.2021.102536

Cui, L. (2012). Fascinating China sceneries in blockbuster films. *Journal of Film Culture*, 88–91.

Dong-mei, S. (2004). The trend of the development of Chinese films. *Journal of Yantai College of Education*.

Feng, W., Guo, Y., & Zeng, Y. (2018). Visual style analysis in Asian film industries. *International Journal of Arts and Technology, 10*(2), 155–174. https://doi.org/10.1504/IJART.2018.10015983

Gao, S., Xu, C., & Sun, J. (2019). Limitations of sentiment analysis in user-generated content. *Information & Management, 56*(6), 103137. https://doi.org/10.1016/j.im.2019.03.002

Hu, H., & Wang, S. (2020). Timing and earnings in the Chinese film industry. *China Economic Review, 60*, 101385. https://doi.org/10.1016/j.chieco.2020.101385

Ke-yang, S. (2007). Aesthetic puzzles of current Chinese movies. *Journal of Beijing Normal University*.

Li, Y., Kim, D., & Xu, L. (2018). Color and genre in East Asian cinema. *Color Research & Application, 43*(5), 692–702. https://doi.org/10.1002/col.22281

Liu, B., Huang, M., & Zhu, X. (2019). The impact of poster design on movie success. *Entertainment Computing, 31*, 100306. https://doi.org/10.1016/j.entcom.2019.100306

Liu, Y., Jiang, L., & Wang, Z. (2020). User sentiment and engagement in online film reviews. *Online Information Review, 44*(4), 789–807. https://doi.org/10.1108/OIR-11-2019-0352

Lu, N.-H. (2023). How to create 3D movie scenes in the style of Chinese paintings using computer processing. In *2023 IEEE Symposium on Systems and Information Management (SSIM)*. https://doi.org/10.1109/SSIM59263.2023.10469258

Luo, S.-B., Fan, C.-C., Chen, K.-Y., Tsao, Y., Wang, H.-M., & Su, K.-Y. (2022). Chinese movie dialogue question answering dataset. *Proceedings of the ACL Chinese Workshop*.

Marion, D. J. (1997). *The Chinese filmography: The 2444 feature films produced by studios in the People's Republic of China from 1949 through 1995*. Greenwood Press. https://doi.org/10.5860/choice.35-2463

Pan, J., Yang, L., & Xiao, Y. (2020). Predicting critical acclaim using team metadata. *IEEE Transactions on Affective Computing, 11*(3), 415–427. https://doi.org/10.1109/TAFFC.2018.2834611

Peng, R., Liu, M., & Wang, H. (2021). Lexical features in user-generated reviews of art-house films. *Journal of Information Science, 47*(4), 524–537. https://doi.org/10.1177/0165551520925610

Qian, Y., Zhao, Y., & Lu, J. (2021). A review on multimodal fusion techniques in recommender systems. *ACM Computing Surveys, 54*(6), 1–35. https://doi.org/10.1145/3459639

Shang, L. (2016). The application of Chinese traditional culture in the animation scene design. In *Proceedings of the 2015 International Conference on Education, Language, Art and Inter-cultural Communication (ICELAIC)*. https://doi.org/10.2991/ICELAIC-15.2016.131

Shi, K. (2008). Interactions between Chinese films and Chinese society. *Tangdu Journal*.

Shi, R., Xu, D., & Meng, X. (2020). Platform bias in Chinese rating systems. *New Media & Society, 22*(3), 420–439. https://doi.org/10.1177/1461444819866659

Silbergeld, J. (2003). *China into film: Frames of reference in contemporary Chinese cinema*. Reaktion Books. https://doi.org/10.2307/3591783

Sun, X., & Li, Y. (2024). The change of Chinese elements in Hollywood films from a cross-cultural perspective. *Advances in Education, Humanities and Social Science Research, 9*(1), 212. https://doi.org/10.56028/aehssr.9.1.212.2024

Sun, Y., He, L., & Wang, D. (2022). Classifying movies using multimodal feature embeddings. *Expert Systems with Applications, 199*, 116926. https://doi.org/10.1016/j.eswa.2022.116926

Tan, H., Li, S., & Huang, Z. (2017). Revenue forecasting and holiday timing effects. *Journal of Retailing and Consumer Services, 38*, 17–25. https://doi.org/10.1016/j.jretconser.2017.05.009

Tang, L. (2015). Visual culture in contemporary China: How (not) to watch a Chinese blockbuster. In *Cultural China: The Contemporary Chinese Cinematic Experience* (pp. 98–115). Cambridge University Press. https://doi.org/10.1017/CBO9781316026724.006

Tang, Y., Zhang, R., & Liu, J. (2023). Representational learning for multimodal content classification. *Knowledge-Based Systems, 261*, 110196. https://doi.org/10.1016/j.knosys.2022.110196

Wang, H.-C., Lu, S., Lim, J.-H., & Pomplun, M. (2012). Visual attention is attracted by text features even in scenes without text. *Cognitive Science*.

Wang, J., Chen, F., & Li, M. (2021). Deep learning methods for Chinese film popularity prediction. *Neural Computing and Applications, 33*, 7395–7410. https://doi.org/10.1007/s00521-020-05539-1

Wang, X., Lin, Z., & Zhao, Y. (2020). Visual strategy in promotional content of Chinese cinema. *Asian Journal of Communication, 30*(3), 241–259. https://doi.org/10.1080/01292986.2020.1748437

Wang, Y., & Li, L.-M. (2023). Exploring heroic characters in red Chinese movies: A multimodal perspective. *Education, Humanities and Social Sciences, 14*(1), 1–9. https://doi.org/10.54097/ehss.v14i.8858

Xiao, D. (2009). Chinese movies under visual culture. *Journal of Jimei University*.

Xu, H., Zhou, X., & Wu, Y. (2022). Sentiment and engagement behavior in online film review platforms. *Behaviour & Information Technology, 41*(8), 1561–1575. https://doi.org/10.1080/0144929X.2021.1882399

Yang, H., Wu, C., & Li, X. (2019). Multimodal machine learning in film reception analysis. *Multimedia Systems, 25*(3), 277–290. https://doi.org/10.1007/s00530-019-00629-w

Yin, Z., Liu, X., & Jiang, H. (2021). COVID-19's impact on Chinese media production. *Telematics and Informatics, 62*, 101622. https://doi.org/10.1016/j.tele.2021.101622

Ying-hong, L. (2001). On the threefold attributes of Chinese movies with national features. *Journal of Chinese Cinematic Arts*.

Zhan, L., Guo, J., & Li, X. (2017). Cinematic features under context cultural pattern. *Journal of Art Studies*.

Zhang, K., Feng, Y., & Lin, Y. (2016). Artistic traits and critical reception in Chinese film. *International Journal of Cultural Policy, 22*(4), 563–577. https://doi.org/10.1080/10286632.2015.1015536

Zhao, R., Lin, J., & Hou, L. (2019). Emotional expression in digital audience reviews. *Information Processing & Management, 56*(2), 230–245. https://doi.org/10.1016/j.ipm.2018.10.004

Zhou, Y., Zhang, S., & Chen, M. (2021). Dimensionality reduction in multimodal analytics. *Pattern Recognition Letters, 145*, 206–213. https://doi.org/10.1016/j.patrec.2020.05.023