[1]Jia Zhao

Communication and Influence of Traditional Culture Based on Social Network Analysis

Abstract: - Social media significantly impacts traditional cultures, yet existing methods struggle to capture the delicate complexities of online cultural discussions. This lack of depth limits understanding of how traditions evolve and adapt in the digital age. To address this, CaCCGAN: Social Culture Analysis, a framework leveraging social network analysis (SNA) is proposed. Content-Aware Cycle-Consistent Generative Adversarial Network (CaCCGAN) explores communication and influence dynamics surrounding traditional culture on social media platforms. CaCCGAN gathers real-time data on traditional culture from Weibo, a prominent Chinese platform, focusing on relevant keywords. By employing Term Frequency-Inverse Document Frequency (TF-IDF) and a Binary Waterwheel Plant Optimization Algorithm (BWPOA), CaCCGAN identifies user-tradition relationships, influential users, and online communities. Furthermore, the CaCCGAN model analyzes communication styles and generates content reflecting these relationships. Analyzing this generated content offers valuable insights for stakeholders interested in cultural preservation, targeted marketing campaigns, and content creation strategies. The proposed CaCCGAN: Social Culture Analysis model attains 25.23%, 21.06% and 55.11% higher Precision value and 29.096%, 41.64% and 33.33% higher F1 Score value compared with the existing method such as PA-BiLSTM Neural Encoding for Cultural Adaptation and Emotion Analysis in Modern Media Communication (PA-BiLSTM), MFCSNet: Modeling Musician-Follower Dynamics in a Complex Social Network to Measure Musical Influence (MFCSNet) and Utilizing Social Media Data Analytics to Integrate Traditional Culture Through Digital Media Interaction and Dissemination (I-RankClus-WOA) respectively. This proposed CaCCGAN: Social Culture Analysis model presents a promising approach to understanding the evolving dynamics of communication and influence surrounding traditional culture online.

Keywords: Traditional culture on social media platforms, Social Network Analysis, Cultural Preservation, Content-Aware Cycle-Consistent Generative Adversarial Network, Binary Waterwheel Plant Optimization Algorithm, Term Frequency-Inverse Document Frequency, Content Generation, Community Detection.

1. INTRODUCTION

Traditional cultures are undergoing a dynamic transformation in the face of pervasive social media platforms [1]. Online discussions have become a significant influence on how individuals perceive and engage with traditional practices [2]. Understanding these communication dynamics is crucial for comprehending the impact of social media particularly the identification of influential users and content creators on cultural heritage in the digital age [3-5].

The existing methods employed for analyzing online influence often rely on manual content analysis or basic statistical techniques [6]. While these approaches hold value, which is demonstrably limited in their ability to handle the vast quantity of data generated by social media platforms [7]. This restricted scope can lead to the overlooking of critical trends or the lack of delicate details in user sentiment [8]. Additionally, traditional methods may struggle to differentiate between distinct online communities centered around specific traditions [9]. This lack of granularity hinders researcher's ability to fully grasp the intricacies of communication and influence within online discussions about traditional culture [10]. Imagine, for instance, the limitations of analyzing a conversation about a traditional festival without knowing if it originates from a youth-oriented group celebrating the event or a community of elders concerned about its preservation. Such a lack of context severely restricts the insights gleaned from the analysis [11-15]. To address these shortcomings and gain deeper insights, a more sophisticated approach is needed. This manuscript introduces CaCCGAN: Social Culture Analysis, a novel framework that leverages social network analysis (SNA) techniques to explore the communication dynamics and influence surrounding traditional culture on social media platforms. This approach leverages advanced techniques to gain deeper insights into these online interactions. The main contribution of this work is given below,

[1] Zhuhai College of Science and Technology, Zhuhai, Guangdong, 519000, China

*Corresponding author e-mail: zhaoj398@zcst.edu.cn

- Content-Aware Cycle-Consistent Generative Adversarial Network (CaCCGAN) utilizes social network analysis techniques to extract user-tradition relationships from social media data.
- The Binary Waterwheel Plant Optimization Algorithm (BWPOA) is employed to identify influential users based on their connections to various traditions.
- A CaCCGAN network is implemented to analyze communication styles and trends within online discussions. CaCCGAN can even generate new content that reflects the relationships between users and cultural topics.

By incorporating these functionalities, the proposed CaCCGAN: Social Culture Analysis model offers a comprehensive approach for understanding communication dynamics and user influence in the context of online discussions about traditional culture. This framework empowers researchers and stakeholders to navigate the complexities of online cultural interactions, ultimately fostering a deeper understanding of the evolving role of social media in cultural heritage preservation.

The paper is structured as follows: Section 2 delves into related work. Section 3 details the proposed methodology using CaCCGAN: Social Culture Analysis. Section 4 covers the evaluation metrics used and showcases the experimental outcomes. Lastly, Section 5 wraps up the paper by summarizing the main discoveries, recognizing constraints, and suggesting potential future work.

2. RELATED WORKS

Numerous recent studies have investigated the dynamics of communication and influence concerning traditional culture across various social media platforms. Below are some of the recent studies closely related to this topic,

In 2023, Zhu, L., [11] have suggested PA-BiLSTM Neural Encoding for Cultural Adaptation and Emotion Analysis in Modern Media Communication. It excels at capturing emotional connections between text and images in cultural discussions. However, limitations arise in optimizing internal model parameters and potentially overlooking the subtleties of online cultural communication. This gap presents an opportunity for more advanced approaches that focus solely on text analysis, offering a potentially more automated solution and enabling deeper exploration of user influence and community dynamics within these online spaces.

In 2024, Wang, H., et.al [12] have utilized MFCSNet: Musician-Follower Dynamics in a Complex Social Network to Measure Musical Influence. It analyzes influence within specific domains (e.g., music) by examining follower networks and data sets of relevant characteristics. While offering multifaceted perspectives and revealing network structures, these approaches might struggle with the complexities of online cultural discussions. It attains low F-Score value.

In 2022, Hong, N., [13] presented Utilizing Social Media Data Analytics to Integrate Traditional Culture Through Digital Media Interaction and Dissemination. This approach sheds light on network structures and allows for optimization through whale algorithms. However, it might struggle to grasp the details of online cultural discussions. These discussions frequently involve a dynamic interplay between tradition and modernity. Effective methods must navigate the complexities of emotional connection, audience segmentation, and the ever-changing landscape of digital art communication.

In 2022, Trček, D., [14] have suggested Safeguarding cultural heritage via the application of blockchain technologies. It bridges the gap by proposing a multidisciplinary framework that leverages disruptive technologies from other fields and integrates tourism incentives for broader impact. But it focused solely on technology or societal aspects, limited the development of deployable solutions with low Normalized Mutual Information value.

In 2023, Amaro, A.C. and Oliveira, L.,[15] have presented Amiais@ SL: A Metaverse Simulator Enhancing Playful Learning Experience in Cultural Heritage Exploration. Rural communities face a cultural heritage crisis as aging populations and social shifts threaten traditions. This urgency is evident in Amiais, a small Portuguese village with just 15 residents. While virtual 3D worlds like Second Life® offer engaging experiences for heritage preservation, these methods struggle to capture the delicate complexities of online cultural discussions. This limitation hinders our understanding of how traditions evolve and adapt in the digital age.

3. PROPOSED METHODOLOGY

In this section, the proposed CaCCGAN Approach to Social Network Analysis for Exploring Communication and Influence Dynamics in Traditional Culture (CaCCGAN: Social Culture Analysis) is discussed. The block diagram of the proposed CaCCGAN: Social Culture Analysis methodology is given in Figure 1. The detail description about each stage is given below,

3.1 Data Acquisition

For exploring traditional culture, the real time data was collected from Weibo API (Application Programming Interface), a platform where users actively discuss and engage with their heritage. Focusing on the past six months (October 2023 - March 2024), Weibo's API was used to gather relevant posts. The data collection process involved searching for keywords related to various aspects of traditional culture, including: Traditional Holidays (Spring Festival, Dragon Boat Festival, Mid-Autumn Festival); Art Forms (Peking Opera, Calligraphy, Paper Cutting); Historical Figures (Confucius, Mulan, Zhuge Liang); Folklore and Mythology (Dragon Legends, Monkey King stories); Traditional Practices (Tea Ceremony, Martial Arts); Cultural Sites (Great Wall, Forbidden City, Terracotta Army). The data collection process automatically captured the information for each relevant post: English translation of the content (post text), user information (username and location, if available), and optional engagement metrics (likes, comments, reposts). This data collection strategy aimed to create a valuable resource for analyzing online communication and influence surrounding traditional culture. The some of the sample data about traditional culture are given in Table 1.

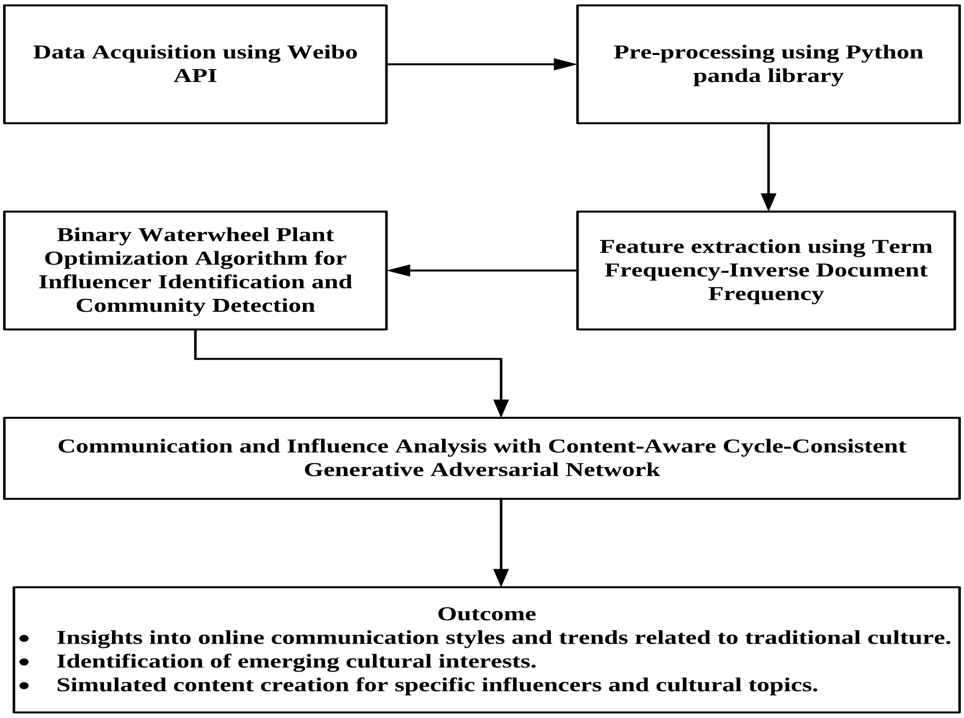

Figure 1: Block diagram for proposed CaCCGAN: Social Culture Analysis methodology

Table 1: Sample Data on Traditional Culture from Weibo

Category	Post Text	Username	Location
Traditional Holiday	Happy Spring Festival! Wishing everyone good health and prosperity.	FoodLover123	Shanghai
Art Form	Trying out calligraphy for the first time! Any tips? #ChineseArt	ArtEnthusiast	Beijing

Historical Figure	Reading about Mulan's bravery today. Inspiring! #Mulan	HistoryBuff	Not provided
Folklore	Just learned a fascinating story about the Dragon Boat Festival! #ChineseMythology	CultureCurious	Guangzhou
Traditional Practice	Attending a tea ceremony this weekend. Can't wait! #TeaLovers	SerenitySeeker	Hangzhou
Cultural Site	Sharing a breathtaking picture of the Great Wall. #TravelChina	Travelholic	Chengdu

3.2 Pre-processing phase

Traditional culture data on Weibo requires pre-processing to enable effective analysis. This stage ensures consistency and clarity for effective analysis. The powerful panda’s library becomes a valuable tool for managing data. The data, typically stored in a CSV file format, undergoes a multi-step processing pipeline to ensure clarity. First, inconsistencies such as typos are addressed to improve data quality. Next, all text content within the "Post Text" section is converted to lowercase for uniformity. Finally, punctuation marks and hashtags are removed to isolate the core message of each post. This pre-processed data, saved in a new format, serves as the foundation for thorough exploration into online communication and the influence of traditional culture on Weibo. The pre-processed output is given in Table 2.

Table 2: Pre-processed Data

Category	Pre-processed Text	Username	Location
Traditional Holiday	happy spring festival wishing everyone good health and prosperity	FoodLover123	Shanghai
Art Form	trying out calligraphy for the first time any tips	ArtEnthusiast	Beijing
Historical Figure	reading about mulan's bravery today inspiring	HistoryBuff	Not provided
Folklore	just learned a fascinating story about the dragon boat festival	CultureCurious	Guangzhou
Traditional Practice	attending a tea ceremony this weekend can't wait	SerenitySeeker	Hangzhou
Cultural Site	sharing a breathtaking picture of the great wall	Travelholic	Chengdu

3.3 Feature extraction phase

In this section, Term Frequency-Inverse Document Frequency (TF-IDF) based feature extraction process is discussed. TF-IDF based Feature extraction bridges the gap between pre-processed text output and CaCCGAN for analyzing online communication about traditional culture on Weibo [16]. Initially, TF-IDF focuses on keywords directly related to specific traditions like "spring festival," or "calligraphy". For that, it calculates the Term Frequency (TF) of each keyword $\left(\mathrm{term}\right)$ within a specific pre-processed output post $\left(post\right)$. The TF generally quantifies how frequently the term appears in the specific pre-processed output post relative to the total number of words in that post. It is mathematically represented in the following equation (1)

$$\mathrm{TF}(\mathrm{term},\ \mathrm{post})=\frac{\mathrm{freq}(\mathrm{term},\ \mathrm{post})}{\mathrm{No.}(\mathrm{term},\ \mathrm{post})} \quad (1)$$

Where $\text{freq}(\text{term},\ \text{post})$ is the number of times the keyword (term) appears in particular document $(post)$, and $\text{No.}(\text{term},\ \text{post})$ is the total number of words within that post. High TF indicates a potentially important keyword within the post. But it doesn't reveal its importance across the entire corpus of Weibo posts. This means common words like "the" can have a high TF but hold little meaning for tradition identification. So, IDF (Inverse Document Frequency) is introduced which tackles this by considering a term's (term) rarity across all Weibo posts D in the dataset, emphasizing unique terms related to traditions. It is mathematically represented in the following equation (2)

$$IDF(term) = \log\left(\frac{D}{df(term)}\right) + 1 \tag{2}$$

Where D represents the total number of Weibo posts and $df(term)$ signifies the number of posts containing term (term) at least once. Common words like "the" or "and" will have a low IDF because they appear frequently everywhere, reducing their significance. Conversely, unique terms related to traditions ("calligraphy contest") will have a higher IDF, emphasizing their potential to distinguish posts about specific traditions. By this TF-IDF weight can be calculated for providing a weight for each term and it is mathematically represented in equation (3)

$$TF - IDF(term, post)\ Weight = \text{TF}(\text{term},\ \text{post}) * IDF(term) \tag{3}$$

This weight considers both the term's frequency within a specific post (TF) and its rarity across all posts (IDF). By this, TF-IDF extracts keywords reflecting specific traditions from user discussions, becoming defining characteristics of each user.

3.4 Binary Waterwheel Plant Optimization Algorithm (BWPOA) for Influencer Identification and Community Detection

After extracting user specific tradition related features using TF-IDF $(TF - IDF(term, post)\ Weight)$, user-tradition matrix (UTM) is constructed. It is denoted as $Z(n \times m)$ matrix, representing the relationship between users and identified traditions based on TF-IDF features. Let assume, $User = \{User_1, User_2, ..., User_n\}$ represent the set of n users on Weibo. Similarly, let assume $Trad = \{Trad_1, Trad_2, ..., Trad_m\}$ represent the set of m traditional culture topics identified through TF-IDF. Each element Z_{ab} in the matrix represents the weight of user $User_a$'s association with topic $Trad_b$. Here, higher values indicate a stronger association between a user and a particular tradition based on their extracted features. Then Binary Waterwheel Plant Optimization Algorithm (BWPOA) is applied for Influencer Identification and Community Detection using the UTM.

For Identifying Influential Users, users with strong associations across various traditions are prioritized. This can be mathematically represented using the following equation (4)

$$\Sigma\left(Z_{ab} * x_a\right) \tag{4}$$

Here, x_a is a binary variable indicating user a 's influence status (1 - influencer, 0 - non-influencer). This term essentially rewards users with high associations that is stronger Z_{ab} values with diverse traditional culture topics. Then for strong community's creation, BWPOA introduces a term that encourages users to join communities that align with their interests. This can be expressed mathematically in the following equation (5)

$$\Sigma\left(Z_{ab} * CK_a\right) \tag{5}$$

Here, CK_a is a binary variable representing user a's membership in community K (1 - member, 0 - non-member). Users are encouraged to join communities that offer a strong connection to their tradition-related interests, as reflected by high Z_{ab} values. Following the identification of influential users, an additional mechanism can be implemented by encouraging the user contribution to communities to foster stronger connections within those communities. This can be mathematically represented using the following equation (6)

$$\alpha * \Sigma\left(x_a * \Sigma\left(Z_{ab} * CK_b\right)\right) \tag{6}$$

where α is a weight parameter that controls the influence of influencers on community formation. This term essentially amplifies the impact of influential users $\left(x_a = 1\right)$ on the overall strength of their communities, measured by the summation over Z_{ab} values for users (b) who are also members of the same community ($CK_b = 1$) [17]. By this, the fitness function for identifying influential users and promote the formation of strong communities can be calculated based on the following equation (7)

$$\textit{Fitness Function} = \Sigma\left(Z_{ab} * x_a\right) + \beta * \Sigma\left(Z_{ab} * CK_a\right) + \alpha * \Sigma\left(x_a * \Sigma\left(Z_{ab} * CK_b\right)\right) \tag{7}$$

Where β is weight parameter that controls the relative importance of community strength $\Sigma\left(Z_{ab} * CK_a\right)$ compared to individual user influence $\Sigma\left(Z_{ab} * x_a\right)$. The step-by-step procedure for BWPOA for Influencer Identification and Community Detection is given below,

Initially, BWPOA initializes the x_a and CK_a values randomly for all users and communities with size n; iteration t and max iteration as $t_{\max}$ of the waterwheel plants. After initialization, BWPOA employs a fitness function for achieving equation (7)

For that in exploration phase, BWpOA mimics waterwheels hunting insects. It utilizes random movements using equations (8-9) to explore for promising influencer and community memberships.

$$SR = r_1\left(x_a \,\&\, CK_a(t) + 2 * \text{Constant}\right) \tag{8}$$

$$x_a \,\&\, CK_a(t+1) = x_a \,\&\, CK_a(t) + SR\left(2 * \text{Constant} + r_2\right) \tag{9}$$

Where, Random variables r_1 and r_2 have values between $[0,2]$ *and* $[0,1]$, while Constant follows an exponential distribution within the interval $[0,1]$. When a solution decreases, Gaussian noise, depicted in equation (10), is added to increase exploration.

$$x_a \,\&\, CK_a(t+1) = \textit{Gaussian} + r_1\left(\frac{x_a \,\&\, CK_a(t) + 2 * \text{Constant}}{SR}\right) \tag{10}$$

After exploration, BWPOA transitions to exploitation phase in which best solution is produced for enhancing the detection of optimal influencer and community through targeted adjustments using equation (11-12),

$$SR = r_3\left(\text{Constant} * \left(x_a \,\&\, CK_a\right)_{best}(t) + r_3 * x_a \,\&\, CK_a(t)\right) \tag{11}$$

$$x_a \,\&\, CK_a(t+1) = x_a \,\&\, CK_a(t) + SR * \text{Constant} \tag{12}$$

Where a random variable r_3 with values ranging from $[0,2]$. Then BWpOA incorporates a mechanism to prevent being trapped in local optima. Like the exploration phase, if no improvement is observed for a specified number of iterations, a mutation is introduced, as defined in equation (13).

$$x_a \,\&\, CK_a(t+1) = (r_1 + \text{Constant})\sin\left(\frac{X}{Y}\theta\right) \tag{13}$$

Equation (13) introduces stochasticity, broadening the search and guiding BWPOA toward unexplored regions within the search space. Through this iterative process of exploration, exploitation, and mutation, BWPOA continuously updates both user influence x_a and community membership CK_a. This BWPOA optimization ensures that the fitness function is maximized, leading to the identification of the most influential users and the formation of strong, vibrant communities.

3.5. Communication and Influence Analysis with Content-Aware Cycle-Consistent Generative Adversarial Network

In this, Content-Aware Cycle-Consistent Generative Adversarial Network (CaCCGAN) analyzes communication and influence dynamics within the context of traditional culture on social media platforms like Weibo. It delves into user profiles (e.g., interests, language style…) and cultural topics (e.g., traditions, historical context…) to understand their unique characteristics. This advanced analysis, which might involve techniques like self-attention mechanisms, allows CaCCGAN to generate new content (like social media posts) that reflects the relationships between users and cultural topics. Additionally, CaCCGAN ensures the generated content remains consistent (cycle consistency) when viewed from both the user and cultural topic perspectives. This is achieved by minimizing a composite loss function ($Loss$) that combines several components:

Adversarial Loss ($Loss_{adv}$): This loss encourages the generator (G) to produce realistic content that deceives the discriminator (D) network responsible for differentiating real and generated content. It is given in equation (14)

$$Loss_{adv} = \arg\min_G \max_D \left[\mathrm{E}_{x \sim P_{data}(x)}\left[D(x)\right] - 2\mathrm{E}_{z \sim P_z(z)}\left[\left(D(G(z))\right)\right] + \lambda * LC(D) \right] \tag{14}$$

where λ is a hyperparameter and $LC(D)$ enforces a gradient Lipschitz constraint and it is represented as the following equation (15)

$$Loss_{adv} = \max\left[0, \left(1 - D(G(z))\right)\right] \tag{15}$$

Cycle Consistency Loss ($Loss_{CCL}$): This loss ensures the generated content maintains its core meaning when "translated" back and forth between user profiles and cultural topics. It is mathematically represented in the following equation (16)

$$\begin{aligned} Loss_{CCL} = {} & \mathrm{E}_{x \sim P_{data}(x)}\left[\left\|F\left(G\left(K(x)\right)\right) - x\right\|_1\right] + \\ & \mathrm{E}_{y \sim P_{data}(y)}\left[\left\|F\left(G^*\left(K^*(y)\right)\right) - y\right\|_1\right] \end{aligned} \tag{16}$$

Here, F represents an optional perceptual loss function (e.g., enforcing feature similarity with real content using a pre-trained model like VGG-19) [18], x and y are real data points, G and G^* are the generator networks for different directions of the cycle, and K and K^* represent the encoder functions for the respective user profile and cultural topic domains.

Identity loss ($Loss_{Identity}$): It forces generators (G, F) to preserve content (y, x). It is mathematically represented in the following equation (17)

$$Loss_{Identity} = \mathrm{E}_{y \sim P_{data}(y)}\left[\left\|G(y) - y\right\|_1\right] + \mathrm{E}_{x \sim P_{data}(x)}\left[\left\|F(x) - x\right\|_1\right] \tag{17}$$

Perceptual Loss ($Loss_{\mathbf{Perceptual}}$): This loss (F) promotes similarity between the generated content and features extracted from real content using a ReLU activation layer loss of the pre-trained VGG-19 model. It is mathematically represented in the following equation (18)

$$Loss_{\mathbf{Perceptual}} = \left\|F(G(z)) - F(x_{real})\right\|_1 \tag{18}$$

where x_{real} is a real data point. Then the overall loss function combines these components with weighting hyperparameters (W_1, W_2, W_3) of Cycle Consistency Loss, Identity loss and Perceptual Loss. It is given in the following equation (19)

$$Loss = Loss_{adv} + W_1 * Loss_{CCL} + W_2 * Loss_{Identity} + W_3 * Loss_{\mathbf{Perceptual}} \tag{19}$$

These functionalities translate into valuable benefits for communication and influence analysis. Analyzing the generated content for specific users and cultural topics provides valuable insights into user communication styles. This can reveal patterns, such as variations in formality or language style, depending on the user and the cultural context. Similarly, analyzing generated content across various topics helps identify emerging trends in user interest within online communities. Furthermore, CaCCGAN facilitates simulation of content creation by identifying influencers for specific cultural topics. Analysis of simulated content reveals potential communication strategies and user engagement for identified influencers on specific cultural topics.

In essence, CaCCGAN empowers various stakeholders interested in cultural preservation. By analyzing communication styles and trends in online discussions about traditional culture, CaCCGAN helps promote cultural heritage online. It informs targeted marketing campaigns by identifying emerging cultural interests. Additionally, CaCCGAN sheds light on how people discuss different cultural topics, inspiring the creation of engaging content that sparks online discussions. Overall, this approach leverages social network analysis to gather and analyze data from platforms like Weibo. By pinpointing influencers, online communities, and communication styles, CaCCGAN provides valuable insights into communication dynamics surrounding traditional culture, ultimately aiding in its preservation efforts.

4. RESULT AND DISCUSSION

In this phase, the experimental results of the proposed CaCCGAN: Social Culture Analysis model are discussed. The Weibo API provided data for the past six months (October 2023 - March 2024), collected using keywords related to various aspects of traditional culture. After following the previously described pre-processing steps (fixing typos, converting to lowercase, removing punctuation and hashtags), the data was divided into training (70%), validation (15%), and testing (15%). Here the training set allows the model to learn, the validation set helps identify and adjust for overfitting during training, and the testing set provides a final assessment of the model's ability to generalize to unseen real-world data on Weibo. Simulations were conducted on a PC equipped with an Intel Core i5, 2.50 GHz CPU, 8GB RAM, and running Windows 7. TensorFlow or PyTorch were chosen to implement the proposed CaCCGAN: Social Culture Analysis model. The validation set played a key role in optimizing the model's performance. By adjusting hyperparameters like the weights for various loss functions (W_1, W_2, W_3), the model was fine-tuned on the validation set to ensure it delivers the best results on unseen real-world data. Following this, evaluation metrics such as are analyzed. The performance of the proposed CaCCGAN: Social Culture Analysis model is then assessed and compared with existing methods like PA-BiLSTM Neural Encoding for Cultural Adaptation and Emotion Analysis in Modern Media Communication (PA-BiLSTM) [11], MFCSNet: Modeling Musician-Follower Dynamics in a Complex Social Network to Measure Musical Influence

(MFCSNet) [12] and Utilizing Social Media Data Analytics to Integrate Traditional Culture Through Digital Media Interaction and Dissemination (I-RankClus-WOA) [13] respectively.

The performance metrics are discussed below,

Precision (Influencer Identification): This metric reflects the accuracy of CaCCGAN's influencer selection. It is mathematically represented in the following equation (20)

$$\text{Precision} = \frac{(\text{Number of Correctly Identified Influencers})}{(\text{Total Number of Users Identified as Influential by CaCCGAN})} \tag{20}$$

Normalized Mutual Information (NMI) (Community Detection): This metric quantifies the agreement between CaCCGAN's detected communities (C) and pre-defined communities (ground truth, T). It considers both purity (members within a community belong together) and homogeneity (a community consists of members from a single ground truth community). It is mathematically represented in the following equation (21)

$$NMI = \frac{\left(2 * Mutual\ \ Information(C,T)\right)}{H(C)+H(T)} \tag{21}$$

Where $Mutual\ \ Information(C,T)$ captures the shared information between detected (C) and ground truth (T) communities, $H(C)$ and $H(T)$ represent the entropy of the C and T sets respectively.

F1 Score: This metric combines precision (accuracy of identified positives) and recall (completeness of identified positives), providing a balanced view of CaCCGAN's performance. It is mathematically represented in the following equation (22)

$$F1\ Score = \frac{2 * \left(\text{Precision} * \text{Recall}\right)}{\left(\text{Precision} + \text{Recall}\right)} \tag{22}$$

Where Recall is mathematically represented in the following equation (23)

$$\text{Recall}= \frac{(\text{Number of Correctly Identified Influencers})}{(\text{Total Number of Actual Influencers})} \tag{23}$$

Computation time: It refers to the total time taken to run the entire method.

Figure 2-4 shows the efficiency of the proposed CaCCGAN: Social Culture Analysis model is evaluated to the existing method such as PA-BiLSTM [11], MFCSNet [12] and I-RankClus-WOA [13] respectively.

Figure 2 shows the Precision analysis in identifying influential users. The proposed CaCCGAN: Social Culture Analysis model attains 25.23%, 21.06% and 55.11% higher Precision value compared with the existing method such as PA-BiLSTM, MFCSNet and I-RankClus-WOA respectively. This suggests that the proposed CaCCGAN: Social Culture Analysis model is more effective in pinpointing users who genuinely influence online discussions and trends related to traditional culture.

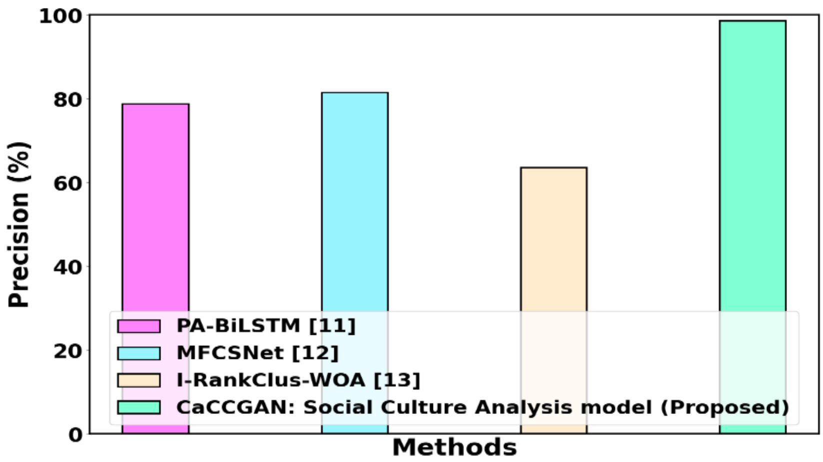

Figure 2: Performance analysis of Precision

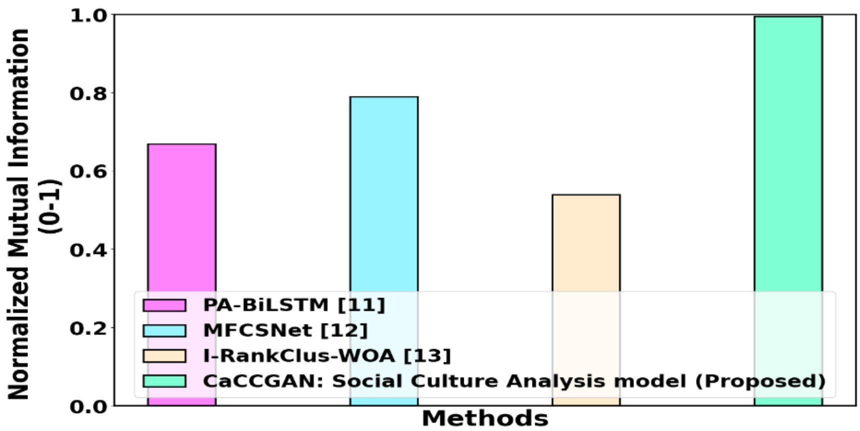

Figure 3: Performance analysis of Normalized Mutual Information (NMI)

Figure 3 shows the Normalized Mutual Information analysis in detecting online communities. The proposed CaCCGAN: Social Culture Analysis model attains 48.507%, 25.94% and 84.25% higher Normalized Mutual Information value compared with the existing method such as PA-BiLSTM, MFCSNet and I-RankClus-WOA respectively. These results indicate that the proposed CaCCGAN: Social Culture Analysis model excels at grouping users based on their shared interests in specific aspects of traditional culture. This allows for a more nuanced understanding of online communities and the cultural topics being discussed within them.

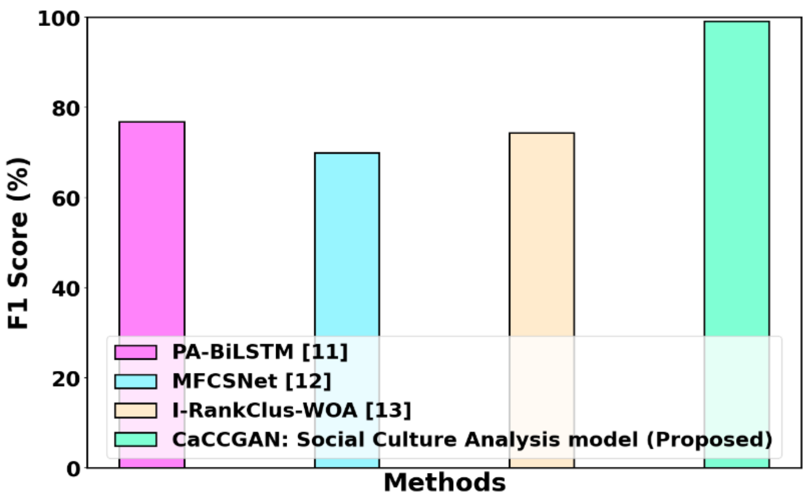

Figure 4: Performance analysis of F1 Score

Figure 4 shows the F1 Score analysis. The proposed CaCCGAN: Social Culture Analysis model attains 29.096%, 41.64% and 33.33% higher F1 Score value compared with the existing method such as PA-BiLSTM, MFCSNet and I-RankClus-WOA respectively. This result reinforces the proposed CaCCGAN: Social Culture Analysis model's effectiveness in analyzing social media data. It excels at identifying key influencers and accurately grouping users based on their cultural interests.

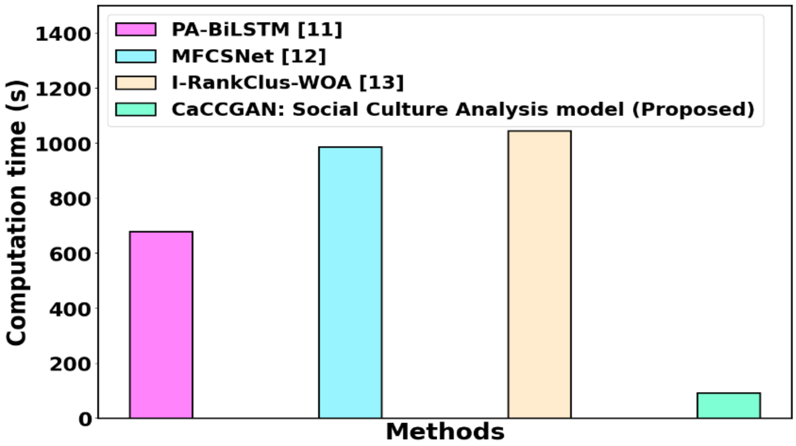

Figure 5: Computation time analysis

Figure 5 shows the Computation time analysis. The proposed CaCCGAN: Social Culture Analysis model attains 86.43%, 90.67% and 91.19% lower Computation time value compared with the existing method such as PA-BiLSTM, MFCSNet and I-RankClus-WOA respectively. This significantly faster processing time makes the proposed CaCCGAN: Social Culture Analysis model a more attractive option, especially for analyzing large datasets or enabling real-time analysis of social media conversations related to traditional culture.

The experimental result of proposed CaCCGAN: Social Culture Analysis model shows promise for analyzing communication and influence in traditional culture. BWPOA effectively identified influential users and formed communities based on user-tradition links. Furthermore, the proposed CaCCGAN's model ability to generate content that reflects user-culture relationships opens doors for further analysis of communication styles and emerging trends. However, the quality of generated content depends on the training data. Expanding data collection and incorporating multimedia content (images, videos) are promising avenues for enhancing realism and coherence. This enriched dataset could lead to a deeper understanding of communication styles and trends within traditional culture.

5. CONCLUSION

In this, the proposed CaCCGAN: Social Culture Analysis model, a novel framework leveraging social network analysis to explore the dynamics of communication and influence surrounding traditional culture on social media platforms was successfully implemented. The proposed CaCCGAN: Social Culture Analysis framework effectively addresses limitations of existing methods, which struggles to capture the intricacies of online cultural discussions. This is achieved through real-time data collection and analysis from Weibo, a prominent Chinese platform. The approach leverages TF-IDF to identify user-tradition connections and utilizes BWPOA to identify influential users and online communities. Furthermore, CaCCGAN analyzes user profiles and cultural topics to understand communication styles and generate content reflecting these relationships. By analyzing the generated content, CaCCGAN offers valuable insights for stakeholders like cultural preservation efforts, targeted marketing campaigns, and content creators. The proposed CaCCGAN: Social Culture Analysis model attains 48.507%, 25.94% and 84.25% higher Normalized Mutual Information value and 86.43%, 90.67% and 91.19% lower Computation time value compared with the existing method such as PA-BiLSTM, MFCSNet and I-RankClus-WOA respectively. The proposed CaCCGAN: Social Culture Analysis methodology shows promise for understanding the ever-evolving online landscape of traditional culture. Future research can explore expanding data collection and incorporating multimedia content to gain a deeper understanding of this dynamic environment.

REFERENCES

[1] Huo, T., Cao, R., Xia, N., Hu, X., Cai, W. and Liu, B., 2022. Spatial correlation network structure of China's building carbon emissions and its driving factors: A social network analysis method. Journal of Environmental Management, 320, p.115808.

[2] Elbashbishy, T.S., Ali, G.G. and El-adaway, I.H., 2022. Blockchain technology in the construction industry: mapping current research trends using social network analysis and clustering. Construction management and economics, 40(5), pp.406-427.

[3] Bertoni, V.B., Saurin, T.A. and Fogliatto, F.S., 2022. How to identify key players that contribute to resilient performance: A social network analysis perspective. Safety science, 148, p.105648.

[4] Chen, Y. and Wang, L., 2022. Misleading political advertising fuels incivility online: A social network analysis of 2020 US presidential election campaign video comments on YouTube. Computers in Human Behavior, 131, p.107202.

[5] Liu, P., Zhang, K., Wang, P. and Wang, F., 2022. A clustering-and maximum consensus-based model for social network large-scale group decision making with linguistic distribution. Information Sciences, 602, pp.269-297.

[6] Li, Y., Kou, G., Li, G. and Peng, Y., 2022. Consensus reaching process in large-scale group decision making based on bounded confidence and social network. European Journal of Operational Research, 303(2), pp.790-802.

[7] An, Q., Wang, P., Zeng, Y. and Dai, Y., 2022. Cooperative social network community partition: A data envelopment analysis approach. Computers & Industrial Engineering, 172, p.108658.

[8] Cardoso, A.S., Renna, F., Moreno-Llorca, R., Alcaraz-Segura, D., Tabik, S., Ladle, R.J. and Vaz, A.S., 2022. Classifying the content of social media images to support cultural ecosystem service assessments using deep learning models. Ecosystem Services, 54, p.101410.

[9] Yao, M. and Liu, J., 2024. The Analysis of Chinese and Japanese Traditional Opera Tunes with Artificial Intelligence Technology Based on Deep Learning. IEEE Access.

[10] Li, J., 2022. Grounded theory-based model of the influence of digital communication on handicraft intangible cultural heritage. Heritage Science, 10(1), p.126.

[11] Zhu, L., 2023. Cultural adaptation and sentiment analysis based on neuron coding in the context of new media communication. Soft Computing, pp.1-10.

[12] Wang, H., Li, Y., Gong, K., Pathan, M.S., Xi, S., Zhu, B., Wen, Z. and Dev, S., 2024. MFCSNet: A Musician–Follower Complex Social Network for Measuring Musical Influence. Entertainment Computing, 48, p.100601.

[13] Hong, N., 2022. Digital-media-based interaction and dissemination of traditional culture integrating using social media data analytics. Computational Intelligence and Neuroscience, 2022.

[14] Trček, D., 2022. Cultural heritage preservation by using blockchain technologies. Heritage science, 10(1), p.6.

[15] Amaro, A.C. and Oliveira, L., 2023, October. Amiais@ SL: A Metaverse Simulator to Support Playful Engagement in Learning About Cultural Heritage. In International Conference on Design and Digital Communication (pp. 651-664). Cham: Springer Nature Switzerland.

[16] Gomes, L., da Silva Torres, R. and Côrtes, M.L., 2023. BERT-and TF-IDF-based feature extraction for long-lived bug prediction in FLOSS: A comparative study. Information and Software Technology, 160, p.107217.

[17] Alhussan, A.A., Abdelhamid, A.A., El-Kenawy, E.S.M., Ibrahim, A., Eid, M.M., Khafaga, D.S. and Ahmed, A.E., 2023. A Binary Waterwheel Plant Optimization Algorithm for Feature Selection. IEEE Access.

[18] Wang, L., Wang, L. and Chen, S., 2022. ESA-CycleGAN: Edge feature and self-attention based cycle-consistent generative adversarial network for style transfer. IET Image Processing, 16(1), pp.176-190.