IIAI Open Conference Publication Series
IIAI Letters on Informatics and Interdisciplinary Research
Vol.006, LIIR374
DOI: https://doi.org/10.52731/liir.v006.374

Examining the Connection Between the Color Scheme of Event Announcement Images and View Counts on Social Media Through Machine Learning Models

Kayu Morishige *, Daichi Inoue †,
Shimpei Matsumoto ‡

Abstract

In this study, we examined whether there is a relationship between the colors used in event announcement images and the number of views. We compared two machine learning models: a conventional model using event images, latitude/longitude, and titles as explanatory variables, and a proposed model incorporating the number of colors, representative HSV colors, latitude/longitude, and titles. The results showed that the proposed model performed similarly to the conventional model. Further analysis revealed that the number of colors and HSV values influenced the number of views. Therefore, it was revealed that the colors in event announcement images are related to the number of views.

Keywords: CNN, SHAP, Machine Learning, Social Media

1 Introduction

In recent years, various social media platforms have emerged with the goal of revitalizing local communities. In recent years, social media platforms that primarily handle visual media, such as images and videos, have become mainstream. To enhance the usability and convenience of these platforms—for instance, through recommendation of similar content or filtering of noise—it is effective to leverage user activity logs. On traditional text-based social media, such activity log analysis has been actively conducted. For example, by analyzing user activity logs and the content of posts, it is now possible to deliver personalized content tailored to users' interests, or to recommend other users with similar preferences [1]. Additionally, techniques for filtering out irrelevant or potentially illegal content [2] have already been implemented, and their effectiveness has been widely acknowledged through real-world application.

In recent years, the adoption rate of smartphones has increased significantly, making them an indispensable part of daily life. With advances in network infrastructure, the development of sensors, and the realization of the Internet of Things (IoT), a wide range of real-world phenomena have been digitized, resulting in a rapid growth in the volume of data transmitted over networks. For instance, between October 2004 and November 2016, total upload traffic increased from 209 Gbps to 1,464 Gbps, while total download traffic grew from 241 Gbps to 8,254 Gbps—representing approximately a 7-fold increase in upload traffic and a 34-fold increase in

* Hiroshima Institute of Technology Graduate School, Hiroshima, Japan
† D2C Inc., Hiroshima, Japan
‡ Hiroshima Institute of Technology, Hiroshima, Japan

download traffic over a 12-year period [3]. As a result, the amount of information accessible to individuals through smartphones has grown tremendously. However, in order to efficiently identify desired information within this vast data landscape, information aggregation has become increasingly important.

One application that incorporates both of these aspects is a service called "Tamemap" [4]. Tamemap aims to make it easy for anyone to post, browse, and share information about small-scale local events, which often do not circulate widely on conventional social media (see Figure 1). Currently, many individuals in local communities face isolation and struggle to find accessible information about local activities in which they can participate. Tamemap addresses this issue by providing an online space where users can easily connect with one another and access relevant local event information. Through this approach, the aim is to create opportunities for communication among local residents, address information gaps, and alleviate social isolation. By supporting connections among community members, the initiative seeks to contribute to the revitalization and regeneration of local communities.

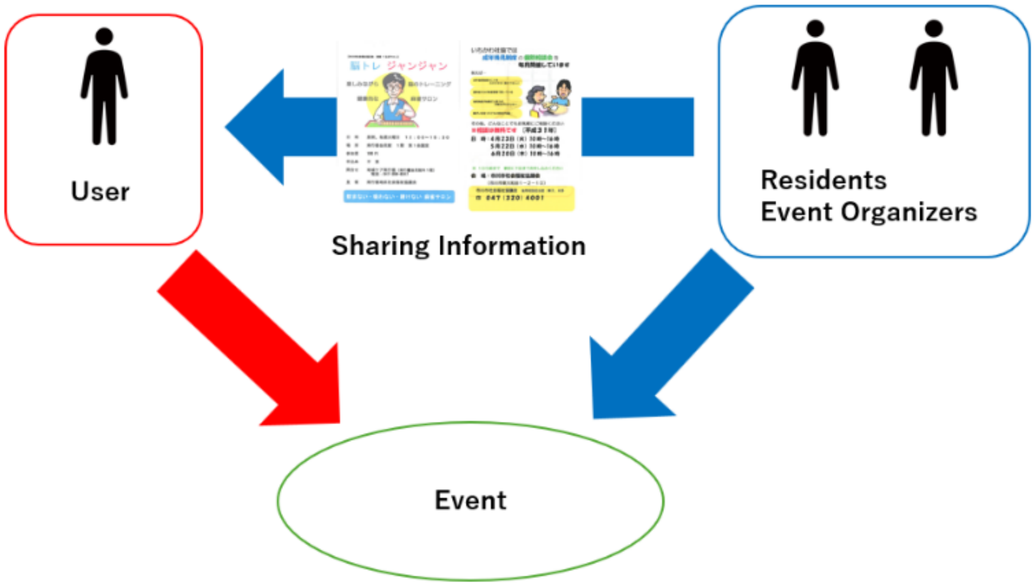

Figure 1: Connection between local events and people through "Tamemap"

Specifically, while event information has traditionally been shared through neighborhood bulletin boards or word of mouth, the use of an application enables users to easily disseminate a wide range of event information and allows others to access that information more efficiently. As a result, compared to conventional methods, smartphones empower both event organizers and users to share local event information, enabling access to events of various scales—from small, individually organized gatherings to large-scale events such as concerts and festivals. This ensures that even individuals who previously had no access to such information can now stay informed. In this way, the use of Tamemap provides opportunities for discovering diverse events and contributes significantly to the revitalization of local communities.

Local communities have traditionally played a crucial role in areas such as public safety, disaster preparedness, childcare, support for the elderly, and environmental preservation. However, due to factors such as the rise of nuclear families, geographic dispersion of relatives, and the separation of workplace and residence, the strength of local communities has been steadily declining in recent years. As community ties weaken, residents increasingly rely on non-profit organizations, municipal services, or commercial services for many aspects of daily life. However, there are regions and households that cannot fully depend on such services, leading to the emergence of various issues. One of the key problems stemming from the decline of local communities is information scarcity. This lack of access to information can lead to social isolation, preventing individuals from receiving necessary support and services, as well as limiting opportunities for interaction with others, thereby contributing to communication breakdowns. Therefore, the revitalization of local communities is not merely desirable, but essential.

The format of event promotional images used on Tamemap differs from those commonly found on major social media platforms such as Facebook. On mainstream SNS, promotional images often feature a brief, attention-grabbing phrase, with the primary goal of redirecting users to a product or event web page. As a result, detailed event information is typically not included in the image itself. In contrast, Tamemap adopts a flyer-style format, where essential event details—such as date and time, location, and whether a reservation is required—are directly embedded within the image. Therefore, the amount of information contained in event promotional images differs significantly between general SNS platforms and Tamemap.

However, Tamemap has several limitations. One of the key issues lies in the posts made by organizers of small-scale events who use the platform. First, due to financial constraints and other limitations, these organizers often create promotional images for their events by themselves, rather than hiring professional designers. In general, it is difficult for individuals with little or no experience in graphic design to produce event promotional images of the same quality as those created by professionals. As a result, the designs for small-scale events tend to be of lower quality compared to those for large-scale events, making them less visually appealing and less likely to attract viewers. Second, many small-scale event organizers are unable or unwilling to invest the time and effort required to revise or improve their promotional images. Consequently, some continue to reuse poorly designed templates that fail to capture attention, further diminishing the visibility of their events.

When aiming to improve usability or optimize conversion rates (CRO) in the design of websites or advertising media, it is common to use A/B testing. A/B testing is one of the standard CRO techniques in web marketing, involving the preparation of two design variants to experimentally determine which performs better in achieving a specific goal. However, A/B testing does not account for external factors beyond the creative elements, making it difficult to interpret why one version outperforms the other. Moreover, Tamemap hosts both small-scale and large-scale event promotional images. While the latter are often designed by professional designers, the former are typically created by event organizers, such as community center staff or volunteer group members, as part of their regular duties. For these individuals, creating promotional content is a time-consuming and unfamiliar task. Given this context, it is unrealistic to repeatedly conduct A/B tests to determine what constitutes an effective design for highly viewable event promotional images, particularly for small-scale event organizers.

However, if a system could be developed that provides both predicted view counts and specific design issues for event promotional images posted on Tamemap, content creators would be able to efficiently improve their designs. This, in turn, could help increase visibility and engagement, even for small-scale events.

Research on such a system has already been conducted. Inoue et al. [5] developed a mock system modeled after Tamemap to investigate the usefulness of a view count prediction feature and whether the predicted view count could influence changes in event promotional image designs. They compared two versions of the system—one with the prediction feature and one without—and conducted user surveys. The results revealed that users were more inclined to improve their designs when the predicted view count was low. This suggests that predicted view counts can serve as an indicator for prompting design changes, highlighting the utility of a prediction-based system. However, their system only presented the predicted view count and did not provide feedback on specific design issues. While this may help users decide whether to revise their designs, it does not assist them in identifying which aspects need improvement. As a result, users might struggle to take concrete action, potentially leading to no design changes at all. In fact, Inoue et al. explicitly acknowledged this limitation in the abstract of their paper.

To provide feedback on specific design issues in event promotional images posted on Tamemap, it is essential to define the visual features that influence user attention and to clarify the relationship between these features and the number of views.

We identify three key visual features that attract user attention: color, large text, and title. These features were selected based on the mechanisms by which view counts are recorded on Tamemap. One method of view count acquisition involves displaying thumbnails of event promotional images organized by date. When a user clicks on a thumbnail of interest, they are directed to the event's detail page, and a view is recorded. Given this mechanism, it is likely that users focus on visually prominent elements such as color and large text when browsing thumbnails. Another method involves displaying map pins within a specified distance from the user's current location. When a user taps a pin and subsequently clicks on a promotional image of interest, the system similarly records a view upon navigating to the event detail page. In this case, it is reasonable to assume that users pay attention to the title when interacting with map pins. Based on these two mechanisms, we conclude that color, large text, and title are critical features influencing user attention and, consequently, view counts.

The relationship between event titles and view counts has already been studied. Inoue et al. [6] investigated view count prediction on Tamemap using event promotional images, geographic coordinates (latitude and longitude), and title information as explanatory variables, with view count as the target variable. They compared the prediction accuracy with a model that used only promotional images and the day of the week as explanatory variables. The results showed that including geographic and title information improved prediction accuracy. Furthermore, Inoue et al. [7] proposed a quantitative classification method for the design of Tamemap's promotional images and examined how design complexity patterns relate to both aesthetics and view counts. Their findings indicated that the readability of titles—such as their size and placement—had a significant impact on the number of views.

However, no prior studies have investigated the relationship between color and view counts. Therefore, this study focuses on analyzing color and its association with the number of views. It is worth noting that research on large text is being conducted as a separate collaborative study with a researcher who is not directly involved in the present work.

Research on the relationship between color and view counts in event promotional images posted on Tamemap has been progressively refined over time [8][9][10]. A recent study [10], for example, involved the extraction of representative colors used in event promotional images. Specifically, RGB values were obtained from each pixel of an image, and K-means clustering was applied to these values. By setting the number of clusters to five and extracting the centroid of each cluster, the representative colors of the image were determined. K-means is a clustering method that partitions a given dataset into k clusters based on the proximity to k centroids. An example is shown in Figure 2. The left side of Figure 2 shows an actual event promotional image used on Tamemap. The graph on the right depicts the RGB components of the representative colors on the vertical axis, and the number of pixels (i.e., the proportion of the image each color occupies) on the horizontal axis.

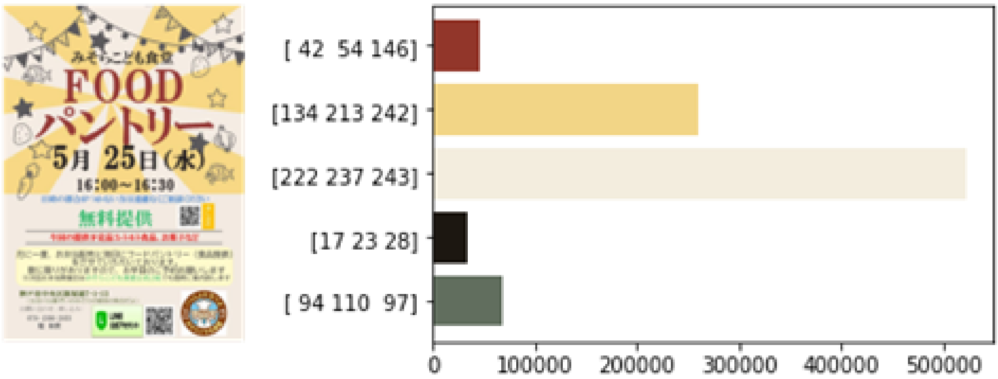

Figure 2: Example of representative color extraction

Second, the RGB components of the five representative colors extracted from each promotional image were treated as separate features. The data was then processed to form four attributes: the RGB values of each representative color and the corresponding view count, as illustrated in Figure 3. These attributes were normalized prior to clustering, and the optimal number of clusters was determined. To perform clustering, we used both K-means and the Gaussian Mixture Model (GMM). For defining the number of clusters, we applied the elbow method and the silhouette method for K-means, and Akaike's Information Criterion (AIC) and Bayesian Information Criterion (BIC) for GMM. GMM is a method that approximates a distribution by combining multiple Gaussian distributions. The elbow method identifies the point at which the total within-cluster sum of squares begins to decrease at a slower rate ("elbow point") and uses it as the optimal number of clusters. The silhouette method calculates how closely data points are clustered within their assigned groups and recommends the number of clusters that maximizes the silhouette score. AIC selects the number of clusters by balancing model fit with the number of parameters, aiming for the lowest score. BIC is similar to AIC but tends to prefer simpler models. By comparing these four methods, we selected a single optimal number of clusters, which was then used for the visualization step described in the third phase.

	hit	col1_R	col1_G	col1_B	…	col5_R	col5_G	col5_B
Image1	29	211	211	211	…	150	154	158

	hit	col1_R	col1_G	col1_B
Image1-1	29	211	211	211

⋮

	hit	col5_R	col5_G	col5_B
Image1-5	29	150	154	158

Figure 3: Data processing method for five RGB elements

As a result, it was observed that when the range of possible cluster numbers was increased, the values computed using AIC and BIC continued to decrease. This suggests that using AIC and BIC tends to favor a much larger number of clusters. Therefore, we concluded that these criteria are not suitable for defining an appropriate number of clusters in this context, and clustering based on AIC and BIC was avoided. In contrast, the elbow method and silhouette method were able to suggest a specific number of clusters. The elbow method indicated four clusters, while the silhouette method suggested two clusters. Since this study aims to examine distinctive characteristics of each cluster, having a larger number of clusters allows for more granular and meaningful analysis. For this reason, we selected the cluster number defined by the elbow method—four clusters—for the final clustering process.

Third, clustering and visualization were performed based on the representative colors used in the event promotional images and their corresponding view counts. Specifically, clustering was conducted using K-means with four clusters, based on four attributes: view count and the RGB components of the representative colors. The resulting high-dimensional data was then reduced from four dimensions to two using UMAP (Uniform Manifold Approximation and Projection). UMAP is a dimensionality reduction technique that preserves both the local and global structure of high-dimensional data in a low-dimensional representation. Finally, a scatter plot was used to visualize the reduced data. However, the results did not reveal a clear relationship between the colors used in the promotional images and their view counts. This suggests that color may not be as strong an attention-driving feature for users as initially assumed.

To further investigate this issue, the present study aims to verify whether color is related to view counts on Tamemap by comparing the performance of two machine learning models. The first model, referred to as the baseline model, uses event promotional images, geographic coordinates (latitude and longitude), and title information as explanatory variables, with view count as the target variable. The second model, referred to as the proposed model, uses the number of colors in the promotional image, the HSV color space values of the representative colors, as well as geographic and title information as explanatory variables. The HSV color space represents colors using three components: Hue, Saturation, and Value (brightness), offering a more perceptually meaningful representation of color compared to the RGB color space.

Color has been widely studied across various fields, including product packaging and face detection. In particular, interest in the role of color in consumer behavior has grown significantly since the early 21st century [11], with numerous studies focusing on the diverse effects that color can have. This trend is closely tied to the rise of sensory marketing, which aims to influence consumer behavior by appealing to multiple senses. Among the five senses, vision is considered to be the most dominant [12], which likely explains the abundance of research on color—an attribute closely associated with visual perception. Therefore, focusing on color in the context of event promotional images is not only relevant, but also highly meaningful.

In addition, the usefulness of employing event promotional images as explanatory variables has been empirically validated by Shimoyama et al. [13]. Specifically, they examined the machine learning models used in the studies by Inoue et al. [6][14], which incorporated event promotional images as explanatory variables. Shimoyama et al. developed a new model by excluding the promotional images from the input features and compared its performance with the models proposed by Inoue et al. The results showed that the performance of Shimoyama et al.'s model was inferior, thereby demonstrating the effectiveness of including event promotional images as input features in view count prediction tasks.

2 Data Preprocessing

In this study, we analyzed a dataset consisting of 88,270 records from Tamemap. The target variable, view count, is a numerical value and therefore did not require any preprocessing. Preprocessing of the explanatory variables—geographic coordinates and title information—was conducted based on the methods used by Inoue et al. [6]. Specifically, the latitude and longitude values were stored as separate numerical entries in the database, and were vectorized accordingly. The title information consisted of textual data, often including multiple titles, and was processed using the Bag-of-Words (BoW) method to convert them into numerical vectors. In this process, morphological analysis was first applied to each title to generate a list of individual words. A vocabulary was then constructed from all title information in the dataset, and finally, for each title, the frequency of each word in the vocabulary was counted and used to generate a numerical vector representation.

Regarding the event promotional images as explanatory variables, we applied a preprocessing method different from that used by Inoue et al. [6]. In their study, all images were resized to 197×197 pixels to standardize image size and aspect ratio, followed by normalization of pixel values. However, such resizing without preserving the original aspect ratio may distort images that are not square, potentially affecting the accuracy of color-based statistical analysis. To address this issue, we employed a technique known as padding to resize the images to 197×197 pixels. Padding is a method in which extra pixels are added around the outside of an image and is commonly used to standardize image sizes or to reduce boundary effects in image processing tasks. In our approach, we first replicated the pixels at the image boundaries, then padded the image by adding these pixels to the outer edges, and finally resized the result to 197×197 pixels. After resizing, we applied pixel value normalization to the padded images.

The number of colors and the HSV color space of representative colors used in the event promotional images were extracted based on our previous studies [8][9][10]. Specifically, to determine the number of colors, we obtained the HSV values for each pixel in an image and calculated the number of unique HSV values, thereby constructing a dataset representing color diversity. For the HSV color space of representative colors, we performed K-means clustering on the HSV values extracted from each image. By setting the number of clusters to five, we extracted the centroid of each cluster, resulting in a dataset representing the HSV color space of the image's representative colors. The decision to use five clusters was based on two key considerations. First, certain colors, such as black and white, are commonly used in text elements—particularly for non-emphasized information like contact details or participation fees. These colors tend to dominate when fewer clusters are used, which may obscure the presence of more visually salient colors. Second, the color distribution within event promotional images varies; some images feature one dominant color, while others display multiple high-usage colors. Using too few clusters may fail to capture the full range of prominent colors in such cases. Therefore, five clusters were chosen to ensure a more comprehensive extraction of meaningful representative colors. Since both the number of colors and the HSV values of representative colors are numerical features, they were vectorized without additional preprocessing.

3 Model Creation

In this study, we investigate the relationship between color and view count by comparing the baseline model and the proposed model. For the baseline model, we followed the approach of Inoue et al. [6] and first extracted feature vectors for each modality. Specifically, for event promotional images, we input the preprocessed images into a convolutional neural network (CNN), in which the final softmax layer of a Residual Network pre-trained on ImageNet [15] was removed. From this model, we extracted a 128-dimensional feature vector. A CNN is a deep learning model primarily used for image recognition and feature extraction. For the title information, we passed the preprocessed text through a three-layer fully connected network, resulting in a 128-dimensional feature vector. A fully connected layer (or dense layer) in a neural network is used to integrate all extracted features. The latitude and longitude values were used as-is from the preprocessing step and treated as a 2-dimensional feature vector. Next, the three feature vectors were concatenated to form a 258-dimensional unified feature vector. Finally, this unified vector was passed through an additional fully connected layer to reduce its dimensionality to one, enabling the prediction of view count. An overview of the entire baseline model architecture is shown in Figure 4.

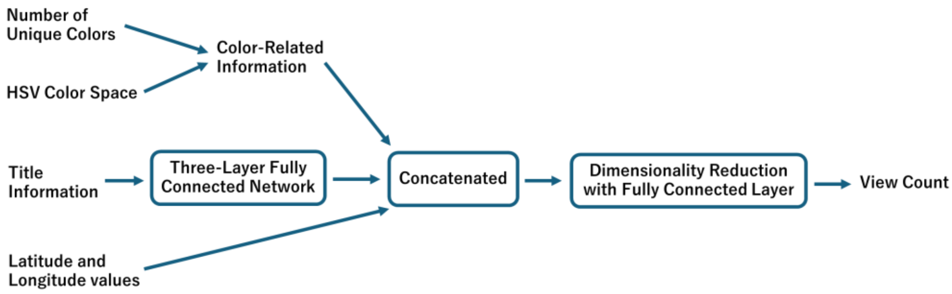

Figure 4: Baseline Model Architecture

For the proposed model, the overall structure follows the same pipeline as the baseline model. First, feature vectors were extracted. The title information and geographic coordinates (latitude and longitude) were processed in the same manner as in the baseline model to obtain their respective feature vectors. For the color-related information, we extracted a 16-dimensional feature vector by combining two components: the number of unique colors (1 dimension), and the HSV color space representation of five representative colors, each consisting of three elements (Hue, Saturation, and Value), resulting in 15 dimensions. Next, the three feature vectors were concatenated to form a 146-dimensional unified feature vector. Finally, this unified vector was passed through a fully connected layer to reduce it to a single dimension, enabling the prediction of view count. An overview of the entire architecture of the proposed model is illustrated in Figure 5.

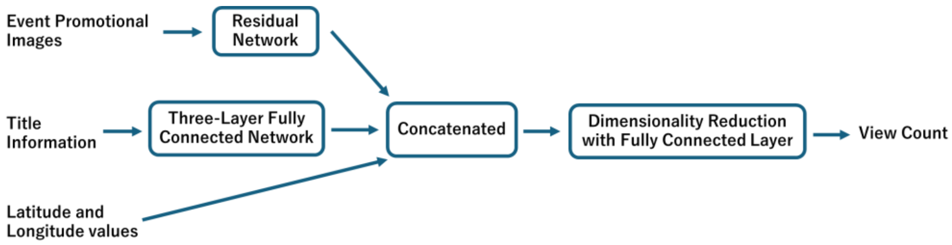

Figure 5: Proposed Model Architecture

For both the baseline and proposed models, the dataset was split into 70% training data and 30% test data. Model performance was evaluated using Mean Squared Error (MSE), following the same approach as Inoue et al. [6]. MSE is calculated by taking the squared difference between the predicted and actual values for each data point, summing these squared differences, and then dividing by the total number of data points. Model optimization was performed using Stochastic Gradient Descent (SGD), a widely used optimization method in machine learning that enables efficient training, particularly for large-scale datasets. Due to GPU memory limitations, it was not feasible to train the model on all images simultaneously. Therefore, we set a batch size of 5000 and a mini-batch size of 32, effectively dividing the training data into 32 blocks of approximately 5000 images for iterative learning. Other hyperparameters followed those used in Inoue et al. [6]: the learning rate for SGD was set to 0.01, and momentum was set to 0.9. To prevent overfitting in the fully connected layers, L2 regularization was applied with a value of 0.0005. The number of training epochs was set to 200.

4 Research Results

The MSE result for the baseline model was 80.38, while the proposed model achieved an MSE of 80.62. Since a lower MSE indicates a better model, the baseline model—which incorporates image data—outperformed the proposed model based on color information. However, this result is not surprising, as event promotional images contain various types of information beyond color, such as text and layout. Therefore, it is expected that a model using the full image would perform better. Nonetheless, the proposed model, which relies only on limited information such as color, achieved an MSE that is comparable to that of the baseline model. This suggests that color information is indeed related to view count and plays a meaningful role in user engagement.

However, relying solely on MSE for evaluation may lead to subjective interpretations. Therefore, we applied SHAP (SHapley Additive exPlanations) to the proposed model to visualize the contribution of each explanatory variable to the prediction of view counts. SHAP is a method based on Shapley values from cooperative game theory, which quantifies how much each feature contributes to a particular prediction, thereby providing a detailed interpretation of the model's decision-making process. The SHAP analysis revealed that geographic coordinates had a significant influence on view count prediction, while the number of colors and the HSV color space of representative colors also had a measurable impact, although to a lesser extent. In contrast, title information contributed minimally to the prediction. From these results, we conclude that the colors used in event promotional images are indeed related to view count, as both the number of colors and their HSV characteristics showed meaningful influence in the model.

5 Future Tasks

5.1 Relationship between Color and View Counts

In this study, we focused solely on comparing the evaluation metrics of the baseline and proposed models and on assessing the feature contributions of the proposed model using SHAP. However, we did not conduct a detailed investigation into the specific relationships between view counts and the number of colors or the HSV color space of the representative colors in the promotional images. For example, potential relationships such as "a greater number of colors leads to higher view counts" or "higher color saturation results in more views" were not examined. Investigating these aspects could enable us to provide concrete design feedback for event promotional images posted on Tamemap. One possible approach for uncovering such relationships is correlation analysis. Specifically, we could compute the correlation coefficients between view count and four variables: the number of colors and the three HSV components (hue, saturation, and value) of the representative colors. The results could then be visualized using a heatmap, providing insights into how each color-related feature correlates with user engagement.

In the present study, the color-related explanatory variables used in the proposed model were limited to the number of colors and the HSV color space of representative colors. However, we believe that there are additional color features that could be incorporated as explanatory variables. For example, the proportion of the most dominant color and the HSV color space of the background color are two promising candidates. The proportion of the most dominant color can be obtained by performing K-means clustering on the HSV values of all pixels in the event promotional image (as done for the representative colors), identifying the cluster with the largest number of pixels, and dividing that number by the total number of pixels in the image. The background color can be estimated by extracting the HSV values from a 10-pixel-wide margin along the top, bottom, left, and right edges of the image, and computing the mode (most frequent value) within this region. Using these features, we can construct a dataset and build a machine learning model following the same approach as in the proposed method. By applying SHAP, we can analyze the contribution of these new features to view count prediction. If the analysis reveals that these features are indeed related to view count, we can then further investigate their specific relationships—just as we proposed for the number of colors and the HSV values of representative colors.

5.2 Improvement of a Function to Predict View Counts

In Chapter 1, we discussed a limitation of the view count prediction feature proposed by Inoue et al. [5], namely that it does not provide specific feedback on design issues within the system. As a result, content creators may find it difficult to identify which aspects of their promotional images need improvement, and thus may choose not to revise the design at all. We believe this issue can be addressed by enhancing the view count prediction feature. Specifically, we propose developing an improved model—an extension of our current method—that incorporates three types of explanatory variables: color features, title features, and large-text features, with view count as the target variable. This model would be integrated into the mock Tamemap system used by Inoue et al. [5]. Based on the relationships between these three visual elements and view counts, the system would then identify which components are lacking or insufficient in a given event promotional image and provide concrete design suggestions accordingly. Finally, as in the study by Inoue et al. [5], we would compare two versions of the system—one with and one without the design suggestion feature—and conduct a user survey to evaluate their effectiveness.

The results of this study revealed that title information, in its textual form, does not significantly contribute to view count prediction. A similar observation was made by Inoue et al. [6], who reached the same conclusion when focusing specifically on small-scale events. Combining this finding with the results of Inoue et al. [7], we can infer that while the textual content of the title may not influence view count, the visual attributes of the title—such as its size and placement within the promotional image—do have an impact. Therefore, when incorporating title-related features into a view count prediction model, it is essential to consider not just the text itself, but also how it is visually presented within the image.

6 Summary

This study aimed to clarify whether the colors used in real-world event promotional images are related to their view counts. To do so, we compared the performance of two machine learning models: a baseline model that used event promotional images, geographic coordinates, and title information as explanatory variables; and a proposed model that used the number of colors, the HSV color space of representative colors, geographic coordinates, and title information. First, in order to use the data as explanatory variables, we applied the following preprocessing procedures. The geographic coordinates were used as-is and vectorized, while the title information was converted into numerical vectors using the Bag-of-Words (BoW) method. The event promotional images were padded and resized to 197×197 pixels, and then passed through a CNN to extract feature vectors. The number of colors was calculated by extracting HSV values from each pixel and counting the number of unique HSV values. Lastly, the HSV color space of representative colors was obtained by applying K-means clustering (with five clusters) to the HSV values of all pixels in each image, and extracting the centroids of these clusters as feature vectors. Second, using the above feature vectors, we constructed neural network models for both the baseline and proposed methods. When evaluating the models using Mean Squared Error (MSE), the proposed model achieved a performance comparable to that of the baseline model. Furthermore, we applied SHAP to the proposed model to assess the contribution of each explanatory variable to the predicted view count. The results showed that both the number of colors and the HSV color space of representative colors had a measurable impact on view count prediction. These findings indicate that the colors used in event promotional images are indeed related to view counts.

Future work includes the incorporation of additional color-related explanatory variables beyond the number of colors and the HSV color space of representative colors, the establishment of concrete relationships between color features and view counts, and the enhancement of the view count prediction feature. Addressing these challenges is expected to enable content creators to efficiently improve the design of their event promotional images, ultimately contributing to increased visibility and view counts even for small-scale events.

Acknowledgement

This study was conducted with the generous support of Mr. Koki Okimoto and Mr. Yoshihiro Shimizu of Tamemap Corporation. Additionally, this research was partially supported by the Japan Society for the Promotion of Science (JSPS) Grants-in-Aid for Scientific Research (Grant Numbers 21H01567 and 22K02147). We would like to express our deepest gratitude for their contributions.

References

[1] Z. Li, X. Fang, O. Sheng, “A Survey of Link Recommendation for Social Networks: Methods, Theoretical Foundations, and Future Research Directions,” ACM Transactions on Management Information Systems, vol. 9, no. 1, 2018, pp. 1-26.

[2] U. Shardanand, P. Maes, “Social information filtering: Algorithms for automating ‘word of mouth’,” In Proceedings of the SIGCHI conference on Human factors in computing systems, pp. 210-217, 1995.

[3] Part 1: Special Feature: The Data-driven Economy and Social Transformation, Ministry of I nternal Affairs and Communications, http://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h29/html/nc121210.html.

[4] D. Inoue and S. Matsumoto, “A View Count Prediction Method for Micro-Event Promotional

Images Using CNN," IEEJ Transactions on Electronics, Information and Systems, vol. 144, no. 9, 2024, pp. 942-954.

[5] D. Inoue and S. Matsumoto, "Evaluating the Usefulness of View Count Prediction in a Local Information Sharing Service," Proceedings of the 24th IEEE Hiroshima Section Student Symposium, TP-D-15, 2022, pp. 225-228.

[6] D. Inoue and S. Matsumoto, "Click Count Prediction of Event Promotional Images Focusing on Title Information," Proceedings of the 2023 IEEJ Conference on Electronics, Information and Systems, GS4-5, 2023, pp. 1285-1290.

[7] D. Inoue and S. Matsumoto, "A Design Classification Method for Event Promotional Images Using GMM," Proceedings of the 90th Information Systems Research Meeting, IS-22-041, 2022, pp. 13-18.

[8] K. Morishige, S. Matsumoto, "Relationship Between Color Schemes of Event Announcement Images and View Counts on Social Media Using GMM", 16th International Congress on Advanced Applied Informatics, DOI: 10.1109/IIAI-AAI 63651.2024.00107, 2024, pp. 562-567.

[9] K. Morishige, D. Inoue, S. Matsumoto, "A Study on Increasing View Counts of Micro-Events on Social Media: Detecting the Relationship Between Color Schemes and View Counts," Proceedings of the 98th Information Systems Research Meeting, IS-24-029, 2024, pp. 5-8.

[10] K. Morishige, D. Inoue, S. Matsumoto, "A Study on Improving View Counts of Micro-Events on Social Media - Detecting the Relationship Between Color Schemes and View Counts -", 30th International Symposium on Artificial Life and Robotics, OS16-1, 2025, pp. 1429-1432.

[11] L. I. Labrecque, "Color research in marketing: Theoretical and technical considerations for conducting rigorous and impactful color research", Psychology & Marketing, vol. 37, no. 7, 2020, pp. 855–863.

[12] A. Krishna, "An integrative review of sensory marketing: Engaging the senses to affect perception, judgment and behavior", Journal of Consumer Psychology, vol. 22, no. 3, 2012, pp. 332–351.

[13] E. Shimoyama, K. Morishige, S. Matsumoto, "Effectiveness of Image Data as Explanatory Variables in Click Count Prediction Models for Local Information Sharing Services," Proceedings of the 75th Joint Conference of the Chugoku Chapters of Institutes of Electrical and Information Engineers, R24-26-02, 2024.

[14] D. Inoue and S. Matsumoto, "Click Count Prediction in a Local Information Sharing Service Using CNN Considering Micro-Event Information," Proceedings of the 25th IEEE Hiroshima Section Student Symposium, TP-B-27, 2023, pp. 257–260.

[15] K. He, X. Zhang, S. Ren, J. Sun. "Deep Residual Learning for Image Recognition," In Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770-778.