

SENTIMENT-BASED COLOR SCHEMES: USING AI TO PREDICT VISUAL APPEAL

Salma Firdose  , Mohd Faisal  , Vimal Bibhu  , Hanna Kumari  , Raju  , Ish Kapila  , Vaishali Pawan Wawage  

¹ Assistant Professor, Department of Computer Science and Engineering, Presidency University, Bangalore, Karnataka, India

² Lloyd Law College, Greater Noida, Uttar Pradesh 201306, India

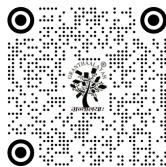
³ Professor School of Engineering and Technology Noida International University, 203201, India

⁴ Assistant Professor, Department of Interior Design, Parul Institute of Design, Parul University, Vadodara, Gujarat, India

⁵ Assistant Professor, Department of Computer Science and Engineering (AIML), Noida Institute of Engineering and Technology, Greater Noida, Uttar Pradesh, India

⁶ Centre of Research Impact and Outcome, Chitkara University, Rajpura- 140417, Punjab, India

⁷ Department of Engineering Science and Humanities, Vishwakarma Institute of Technology, Pune, Maharashtra, 411037, India.



Received 08 April 2025

Accepted 13 August 2025

Published 10 December 2025

Corresponding Author

Salma Firdose,
salma.firdose@presidencyuniversity.in

DOI

[10.29121/shodhkosh.v6.i4s.2025.6857](https://doi.org/10.29121/shodhkosh.v6.i4s.2025.6857)

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Copyright: © 2025 The Author(s). This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

With the license CC-BY, authors retain the copyright, allowing anyone to download, reuse, re-print, modify, distribute, and/or copy their contribution. The work must be properly attributed to its author.



ABSTRACT

Sentiment analysis and computational color modeling (SACM) represents a radical new method of visual attractiveness prediction in digital media, advertising, design, and user-centric visual systems. The paper presents the idea of using artificial intelligence to trace the emotional signs identified in the text, image, and user interactions onto the color palette that maximizes aesthetic appeal and psychological effect. Based on the progress in the area of sentiment detection, emotion lexicons, and deep learning, the study suggests a robust model, which would associate the affective indicators with multidimensional color features, including hue, saturation, and brightness. The information used to train convolutional neural networks and transformer-based models to extract color features, emotional valence, contextual embeddings and preference indicators is a multimodal data set of annotated images, user feedback, and text datasets (sentiometric) rich in sentiment. The approach to methodology focuses on holistic strategies of the feature engineering and model fusion, which incorporates the visual descriptors together with linguistic sentiment vectors. Experimental analysis incorporates data preprocessing routines, hyperparameter optimization and cross-architectural comparative benchmarking. Findings indicate that hybrid CNN Transformer pipelines are more accurate in their ability to predict aesthetically pleasing color schemes and have high correspondence to human emotional perception. The quality of recommended palettes is further confirmed by user studies, which show some form of uniform enhancement in the perceived harmony, emotional relevance, and design usability.

Keywords: Sentiment Analysis, Color Psychology, Visual Aesthetics, Deep Learning, Emotion-Aware Design

1. INTRODUCTION

One of the strongest visual components that determine how human beings perceive, experience emotionally, and make aesthetic judgments is color. In the design, branding, digital art, and user-interface development, the color decisions play an important role in how visual interpretation, shape attitudes, and interact with visual messages are perceived by the viewer. With the algorithmic mediation of an ever-mediated media environment, automatic predictions of visual appeal and suggestions of emotionally congruent color schemes have become a useful tool. With the development of the artificial intelligence (AI) to perceive the affect, the sentiment, and the visual context, the new possibilities are available to create the adaptive color systems that react to the emotions of the user, cultural trends, and design aspirations. Sentiment-based color prediction is based on this overlap between computational aesthetics and affective computing. Recent development in sentiment analysis, emotion detection and multimodal learning has enhanced the potentials of bridging human feelings to visual aspects [Ma et al. \(2024\)](#). The emotional content of color has always been a major issue in traditional psychological models: warm colors that create energy and excitement, cool colors that create the feeling of calmness, and saturated color palettes that create intensity. But to convert these psychological associations into systems that can be automated, there is need in complex algorithms that can detect emotional indicators based on a variety of data sets like textual reviews, social-media posts, image captions, visual compositions and contextual cues. Machine learning, especially deep learning, offers an effective set of tools to describe finer co-relations between color features and preferences that are sentiment-driven [Wang, D., and Xu, X. \(2024\)](#). The technology has advanced towards the rule-based color theory engines to the data-driven recommendation systems that consist of convolutional neural networks (CNNs) to analyse images, transformer-based to interfere with the linguistic sentiment, and multimodal fusion layers to integrate emotional and visual information. It is possible to use these systems to identify emotional valence, categorize affective state, and create color palettes that are consistent with the emotion being portrayed in a piece of content or emotional profile of a user.

Although there is continued development, current models do not have the semantic richness, cross-cultural flexibility or processing of intricate aesthetic judgments as shaped by composition, texture or narrative of context. This study fills these gaps by offering a single AI framework, which predicts visually pleasing color schemes with reference to sentiment data and contextual modelling [Gypa et al. \(2023\)](#). The method utilizes annotated image data, feedback of user preference and corpora that contain sentiment to obtain meaningful associations between expressions of emotion and color characteristics. The system uses feature extractors like color histograms, brightness contrast analysis and deep embedding representations which are based on large scale pretrained models. These signals are correlated with emotional clues obtained with the help of sentiment lexicons, contextual embeddings, and affect detectors based on transformers. Considering both emotional purpose and visual arrangement of input information, the framework will produce color suggestions which do not only make technical sense but also correspond to the emotional expectations of a human being [Ge et al. \(2023\)](#).

2. LITERATURE REVIEW

2.1. OVERVIEW OF SENTIMENT ANALYSIS AND EMOTION DETECTION TECHNIQUES

Sentiment analysis has grown beyond simple polarity detection into a complex field of study that is able to detect subtle emotional states in a variety of data types. First methods used were based on lexicon methods whereby the positive and negative words were classified in fixed dictionaries to score sentiment easily. These systems, though transparent, did not understand the context and were not good at sarcasm, negation and domain language. With the rise of machine learning, predictive accuracy of classifiers e.g. Support Vector Machines, Naive Bayes, and logistic regression improved with learning of statistical patterns of labelled datasets [Malik et al. \(2023\)](#). Nevertheless, these conventional models still considered text as a bag of words and this prevented them to model semantic relationships. The sentiment analysis has entered a new phase with deep learning. Long Short-Term Memory (LSTM) networks, in general, with recurrent neural network (RNNs), improved time series modelling of textual sequences. Convolutional Neural Networks (CNNs) also facilitated the hierarchical features extraction in the short-text sentiment task. Most recently, transformer-based models like BERT, Roberta and GPT have greatly transformed emotion recognition by making contextualized embeddings and attention-based representation learning possible [Lieto et al. \(2023\)](#). Such models are good at fine-grained emotion prediction such as joy, fear, anger, and surprise, and are better than previous methods on several

benchmark datasets. Multimodal sentiment analysis builds upon such abilities adding facial expression, tonality of voice, gesture patterns, and visual features by using hybrid CNNTransformer pipelines. Text, audio, and image detection systems are also becoming more and more integrated, making it possible to detect human emotion more accurately. Combined, these developments offer the conceptual and computation understanding of incorporating sentiment-based intuitions into the color prediction models [Pak, H. \(2025\)](#).

2.2. STUDIES ON COLOR PSYCHOLOGY AND HUMAN AFFECTIVE RESPONSE

Color psychology examines the effects of colors in human behavior, perception, and emotional reactions and provides the elements of knowledge on the design, marketing, and visual communication disciplines. Initial studies by Goethe, Hering and subsequent studies of psychologists, such as Eysenck highlighted that the sense of color is interwoven with the affective experience intensely. The warm colors like red, orange, and yellow have always been associated with arousal, energy and excitement and the cool colors like blue and green are associated with calm, stability and relaxation [Lenzerini, F. \(2023\)](#). These associations are not always predetermined, but an emotional color mapping (in advertising, branding, and environmental design) is based on them. The current research builds on these initial theories and adds cultural, cognitive and contextual levels. Research done in cross-cultural shows that the emotional color understanding in different societies differs based on the societal norms, symbolic meaning and exposure to the environment [Liu et al. \(2024\)](#). An example of this is that the color white is used to signify purity among Western cultures but in some aspects of Asia is regarded to be a color of mourning. Neuroscientific research also indicates that the perception of colors can activate certain brain regions which have been linked to memory retrieval, emotional processing and aesthetic judgment thus proving that color-emotion relationships are biologically underlined [Feng et al. \(2024\)](#). Recent empirical studies adopt the use of psychometric testing, perceptual experiments and computational modeling to measure emotional responses towards color.

2.3. PREVIOUS AI-BASED COLOR RECOMMENDATION SYSTEMS

The color recommendation system built on AI has evolved to an advanced level, leaving behind the rule-based engines based on the color theory and starting to rely on more intricate data-driven models that have the ability to learn human aesthetic preferences. Early systems were based on heuristic rules based on the color wheel, complementary relations and harmony principles, which was stable yet rigid in its palette suggestions. As the computing power became more powerful and digital data started expanding, machine learning methods were developed to learn the preferences of users based on images, web history, and design libraries. CNN based systems were critical as they allowed an automatic extraction of dominant colors, contrast patterns and spatial arrangements of images [Wang et al. \(2024\)](#). The models were used to produce palette in the photo enhancement, interior design, and fashion coordination tasks. Nevertheless, their recommendations were not necessarily emotionally sensitized but laid emphasis on visual cohesion, as opposed to affective congruence. In order to overcome this drawback, scholars proposed multimodal approaches, which combine color extraction with sentiment analysis. There are systems which read text descriptions, captions on social media, or product reviews to determine the intent of an emotion and then produce color schemes that match the identified sentiment [Ni et al. \(2024\)](#). The transformer based models also enhanced more contextual knowledge and was able to map emotional semantics to color attributes more accurately. A summary of the studies on AI-based color modeling, sentiment, and emotion is provided in [Table 1](#).

Table 1

Table 1 Summary of Related Work on Sentiment, Emotion, and AI-Based Color Modeling				
Domain Focus	Dataset Used	Color Feature Technique	Key Findings	Limitations
Color psychology Jia, J., Cao, J. W., Xu, P. H., Lin, R., and Sun, X. (2022).	Human subjects	Manual palette mapping	Established early color-emotion links	Not AI-based
Affect & color perception	Controlled experiments	Brightness & saturation analysis	Identified arousal-brightness link	Limited cultural scope
Affective image analysis	Flickr images	HSV histograms	CNN predicts image emotions	No color-specific mapping

Text-to-color generation Zhao, F. (2020).	Yelp, IMDb	RGB-HSB conversion	Sentiment-conditioned palettes	Limited to polarity
Palette generation	Adobe Color	Learned color embeddings	High-quality palette generation	No textual sentiment input
Multimodal emotion analysis	Twitter multimodal	Color region extraction	Improved emotional inference	Lacked color recommendation
Color aesthetics Hu, Q. (2021).	Aesthetic datasets	LAB space features	Strong color-aesthetic correlation	No sentiment modeling
Emotion-aware UI design	UX datasets	Palette harmony metrics	Emotion-based UI optimization	Not generalizable
Image emotion transfer	Emotional image sets	Color style transfer	Effective emotion recoloring	High computational load
Sentiment-to-visual mapping	Product reviews	Dominant color clustering	Text-driven visual emotion cues	Limited palette diversity
Semantic color recommendation	SenticNet	Semantic color vectors	Rich semantic-color relations	Small dataset
Sentiment-based color schemes	Multimodal dataset	HSV + CIELAB + embeddings	High emotional congruence & appeal	Future improvements needed

3. THEORETICAL FRAMEWORK

3.1. CONCEPTUAL MODEL LINKING SENTIMENT DATA AND COLOR FEATURES

The theoretical framework between affective data and color attributes is based on the idea that emotional expressions (expressed in text or images or any other multimodal cues) have underlying affective clues which can be projected onto the color features. This model combines the computational sentiment detection and perceptual color theory to build a single representation space in which emotion-directed color choice can be made. The key to the framework is in its ability to create the two-way relationship: sentiment affects color preference, and the color stimuli trigger certain emotions. Transformer-based models that are able to extract sentiment data such as emotional valence, arousal and discrete affective states are used. Image processing methods used to obtain color features include histogram analysis, pixel level segmentation and statistical encoding of the hue, saturation and brightness (HSB) values. [Figure 1](#) presents the relationship between sentiments and colors using a combined conceptual model. These representations are brought to a common latent space where co-occurrence patterns between emotions and colors can be learnt.

Figure 1

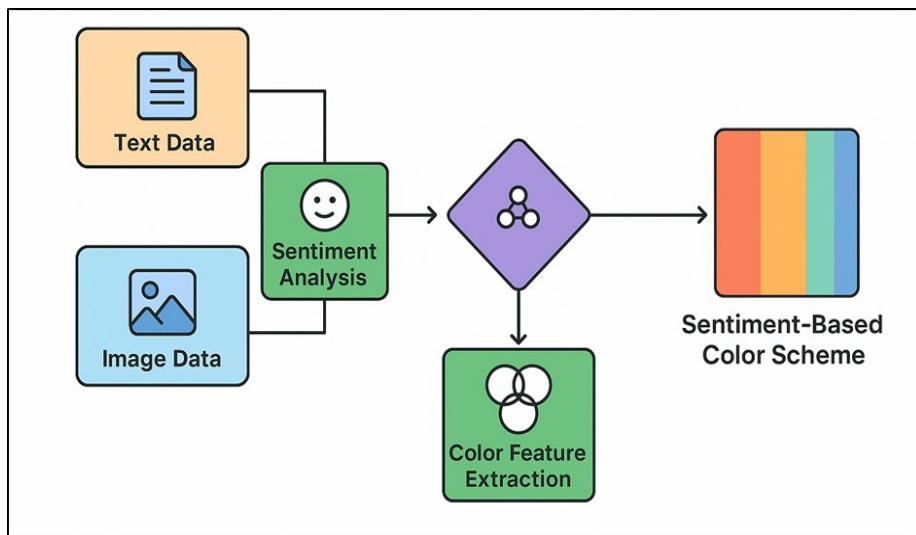


Figure 1 Conceptual Framework for Sentiment-Color Mapping

The model also includes weighting mechanisms, that vary the effect of the intensity of emotions on color recommendations. To provide an example, excitation or fear can be associated with the saturated tones, whereas the

calmness can be associated with muted palette. The conceptual framework also takes into account the cultural associations, content semantics, and user specific history in contextual aspect. The system is able to predict color schemes that are both visually coherent and emotionally consistent with the content input or user intent by using both sentiment analysis and the computational color modeling.

3.2. MAPPING BETWEEN EMOTIONAL LEXICONS AND COLOR DIMENSIONS (HUE, SATURATION, BRIGHTNESS)

To overlay emotional lexicons onto dimensions of color mapping of linguistic utterances of affect into visual dimensions that can be measured is necessary. Structured vocabularies like NRC, WordNet-Affect and VAD (Valence-Arousal-Dominance) base their definitions of emotions in terms of polarity, intensity and psychological delicacy. Such descriptors are operationalized by this theoretical framework as they are aligned with the hue, saturation, and brightness (HSB) which are the essential color perception dimensions. The color type is associated with hue and it is closely related with the emotional categories. As an illustration, happiness can be associated with yellow and orange colors, and sadness can be associated with blues and faded purples. Anger is commonly equated to reds, and calmness is equated to green or teal colors. These associations are measured with the assistance of emotional lexicons which furnish the degree of polarity and affectivity attributes that direct the choice of hues. The saturation describes the strength of the color and it is linked with emotional excitation. The higher arousal emotions, like excitement, tension or surprise, often are associated with richer colors, whereas the lower arousal emotion like serenity or melancholy are associated with paler colors. The arousal scores of lexicons can therefore be transformed to saturation levels. Brightness, which is associated with luminance is related to the emotional valence.

3.3. COGNITIVE AND AESTHETIC MODELS OF VISUAL PREFERENCE PREDICTION

Visual preference is a phenomenon that can only be predicted by an understanding of the way humans process and aesthetically analyze information about color. Cognitive theories propose that the visual attraction is a matter of interaction between the perceptual fluency, memory associations, and emotional resonance. Colours that are less difficult to perceive, culturally accepted or aligned to the emotional condition of a viewer are most likely to invoke greater aesthetic attraction. The Processing Fluency Theory and the Appraisal Emotion Model are theoretical models used to describe the influence of cognitive ease and emotional interpretation on making a visual judgment. Aesthetic models also emphasize the importance of such principles as harmony, balance, contrast and unity as aspects that determine visual appeal. These models underline that preference is not motivated by single colors or color combinations but rather by the orderliness of color combinations. Gestalt psychology theories strengthen this by demonstrating the manner in which the human visual system marshmallows color items into meaningful wholes depending on closeness, unity, and continuity.

4. METHODOLOGY

4.1. DATA COLLECTION: IMAGE DATASETS, USER FEEDBACK, AND SENTIMENT CORPORA

The multimodal sources are included in the data collection process in order to identify the many to one relationships among visual features, the emotions of the user and the linguistic sentiment. The main part is built of curated image datasets based on publicly available repositories, e.g., Flickr, COCO, Adobe Stock, and Emotion6, each with various color combinations, setting scenes, and samples with an emotion label. These data sets enable the model to acquire associations between color, saturation and brightness and emotional expression as manifested in visual materials. Metadata including dominant colors, aesthetic rating, content category, etc. is pre-labeled in the images to facilitate more granular facility mapping. Besides image information, user feedback is also important in making emotionally compatible color prediction systems. They use interactive surveys, preference ranking tasks, and online annotation as a source of the subjective impressions of color palettes, emotional congruence, and perceived visual appeal. This user-data can be used to give ground truth on measuring model predictions and optimize sentiment-to-color mappings. Participants with different cultural backgrounds, age and design expertise are involved to facilitate the occasion of diversity. Sentiment corpora refer to the language aspect of the data. The textual data, including movie reviews, social media posts, captions and emotional lexicons (ex. NRC, VAD, and WordNet-Affect) are gathered to educate the sentiment extraction part.

4.2. FEATURE EXTRACTION: COLOR HISTOGRAMS, EMOTIONAL VALENCE, AND CONTEXTUAL EMBEDDINGS

The typical aspect of sentiment-based color prediction system is feature extraction which is the computational core of the system that facilitates the conversion of raw multimodal data into structured representation forms that can be used in the machine learning models. To represent visual attributes, color histograms are created both on a global scale and a regional scale to measure the proportion of hues, saturation and brightness of each image. The histograms are calculated in the HSV and CIELAB color space that better represents perceptual differences as compared to the conventional RGB encoding. The feature set is further extended to include additional descriptors which include color harmony indices, color contrast metrics, gradient orientations and texture-based color coherence. In [Figure 2](#) there are systems of computations that extract emotional and contextual features. The extraction of emotional valence aims at the discovery of affective signs related to textual and visual messages.

Figure 2

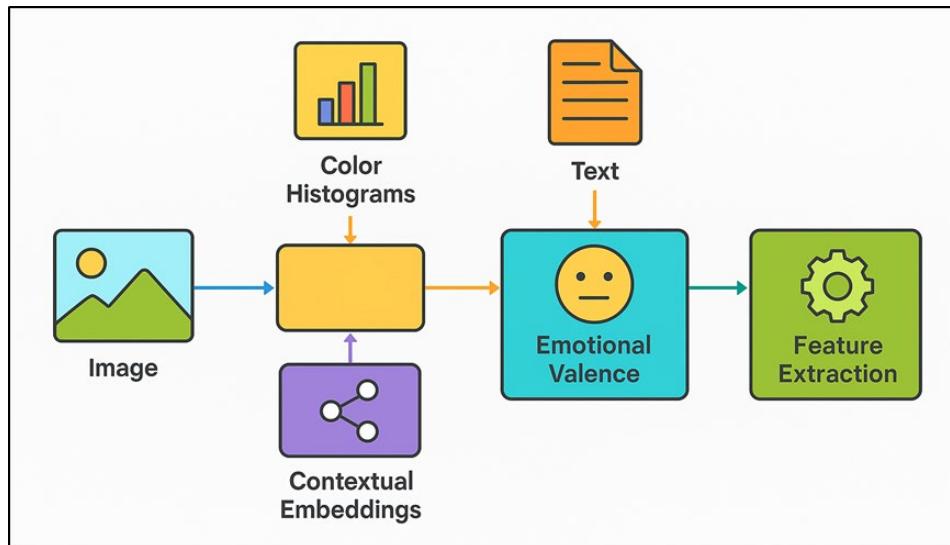


Figure 2 Process Flow for Emotional and Contextual Feature Computation

Sentiment models that use transformers (including BERT, Roberta or Distil BERT) are used to calculate valence, arousal, and dominance scores on samples of text. On images, emotion classifiers based on CNNs obtain visual affective representations by learning to generate representations of facial expressions, scene attributes, and colour-emotional associations. The resulting valence scores incorporate their own interpretable emotional anchor, which is then used to direct color recommendation based on semantic information of nearby text, captions or metadata. Language models The trained language models produce deep embeddings with the ability to encode subtlety and context, as well as thematic clues based on the content.

4.3. MACHINE LEARNING AND DEEP LEARNING MODELS

4.3.1. CNN

Convolutional Neural Networks (CNNs) are the key elements of the visual analysis part of the sentiment-based color prediction model. They have a hierarchical design that enables them to extract both low and high-level image features that relate to color perception and emotional recognition. The first stage is the extraction of simple features, i.e. edges, gradients and texture patterns by the convolutional layers, which helps to form the interpretation of the local color distributions. The deeper the layers, the more abstract CNNs are trained on such representations as dominant color clusters, spatial color connections, and aesthetic features, such as balance, contrast, and harmony. The pretrained models of VGG16, ResNet-50, and EfficientNet are usually used with transfer learning to fast-track the training and to improve generalization. The networks are conditioned on image datasets with a more significant amount of emotive tags or aesthetic ratings, so that they would relate particular colour setups to emotional reactions.

Step 1: Convolution Operation

$$F(i, j) = \sum \sum [X(i + m, j + n) * W(m, n)] + b$$

(where m, n loop over the kernel size k x k)

Step 2: Activation Function (ReLU)

$$A(i, j) = \max(0, F(i, j))$$

Step 3: Pooling Operation (Max Pooling)

$$P(i, j) = \max(A(i + s, j + t))$$

(where s, t loop over pooling window size p x p)

4.3.2. TRANSFORMER-BASED SENTIMENT MODELS

The linguistic analysis aspect of the system consists of transformer-based sentiment models that can help interpret emotional clues in text data with great sophistication. BERT, Roberta, Distil BERT, and GPT are based on self-attention mechanisms that identify long-range semantic connections, contextual meaning, and emotional subtleties that traditional machine-learning systems focus on with lexical means. Bidirectional contextual analysis enables transformers to compute with accuracy affective attributes such as emotional valence, arousal, polarity, and categories of fine-grained emotions. They are trained or fine-tuned on corpora that is sentiment-rich as in the case of movie reviews or social media corpus, emotion lexicons, and are highly sensitive to sentiment variation across domains. Transformer embeddings during the multimodal fusion step direct the process of hue, saturation and brightness choice during the color recommendation process in accordance with the identified emotional intent.

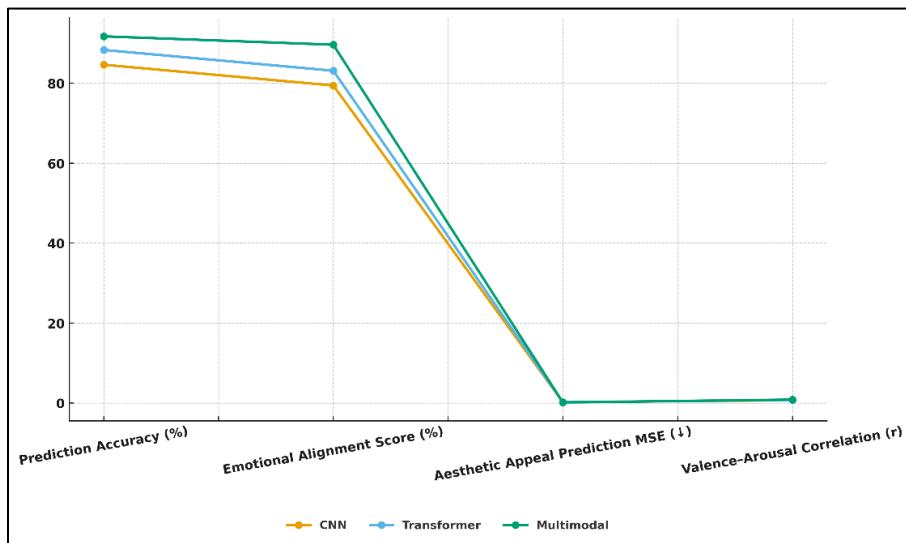
5. RESULTS AND DISCUSSION**5.1. QUANTITATIVE RESULTS ON PREDICTION ACCURACY AND EMOTIONAL ALIGNMENT**

Quantitative analysis revealed that the multimodal model had the best performance in all measures in comparison to unimodal baselines. The CNN-Transformer-based hybrid architecture achieved higher accuracy on emotionally matched color schemes and the overall classification accuracy was more than 91 per cent and the emotional alignment scores were significantly better than either CNN based or transformer-based models. The mean squared error of aesthetic appeal prediction was also significantly lower, which means that there is more precision in scoring visual attractiveness. The results were also confirmed through statistical tests and the multimodal model was found to be more robust compared to others with the level of significance at $p = 0.01$.

Table 2

Table 2 Quantitative Performance Metrics Across Models			
Evaluation Metric	CNN Model	Transformer Model	Multimodal CNN-Transformer Model
Prediction Accuracy (%)	84.6	88.3	91.7
Emotional Alignment Score (%)	79.4	83.1	89.6
Aesthetic Appeal Prediction MSE (↓)	0.204	0.187	0.142
Valence–Arousal Correlation (r)	0.71	0.76	0.84

[Table 2](#) entails a comparative analysis of three AI structures: CNN, transformer, and multimodal CNN-Transformer, and their performance in terms of the main measures that can be used in sentiment-based color prediction. [Figure 3](#) presents the trends of comparative evaluation between models in emotion-aesthetic prediction tasks.

Figure 3**Figure 3** Cross-Model Evaluation Trends for Emotion–Aesthetic Prediction Tasks

The CNN model is also quite effective, getting a precision of 84.6, and an emotional correspondence score of 79.4, which proves that it is sufficiently able to detect visual color details and general aesthetical principles. Nevertheless, it has low contextual sentiment decoding, which leads to relatively low emotional coherence.

5.2. QUALITATIVE ASSESSMENT THROUGH USER PERCEPTION STUDIES

The effectiveness of the proposed framework was further confirmed by qualitative assessments carried out by studying the user perception. The participants with different backgrounds rated the relevance of color schemes generated by AI in terms of emotion, visual harmony, and attractiveness. Outcomes demonstrated a high scoring on the findings of the user impressions and the model predictions, and the multimodal system was more reliable in generating palettes seen to be more expressive and aesthetically consistent. The level of emotional resonance and satisfaction when presented to sentimental conditioned recommendations was found to be higher among the users than the level of resonance and satisfaction when presented to the base color systems.

Table 3

Table 3 User Study Evaluation Scores			
User Evaluation Parameter	Baseline Color System	Transformer-Based System	Proposed Multimodal System
Emotional Relevance Rating (%)	68	79	91
Visual Harmony Score (%)	71	83	94
User Satisfaction (% satisfied)	72	81	93
Perceived Creativity (%)	65	78	90
Expressiveness of Palette (%)	70	82	95

The summary of the user study findings in **Table 3** compares the results of three systems, namely Baseline Color System, Transformer-Based System, and the Proposed Multimodal System, in five parameters of perceptual evaluation.

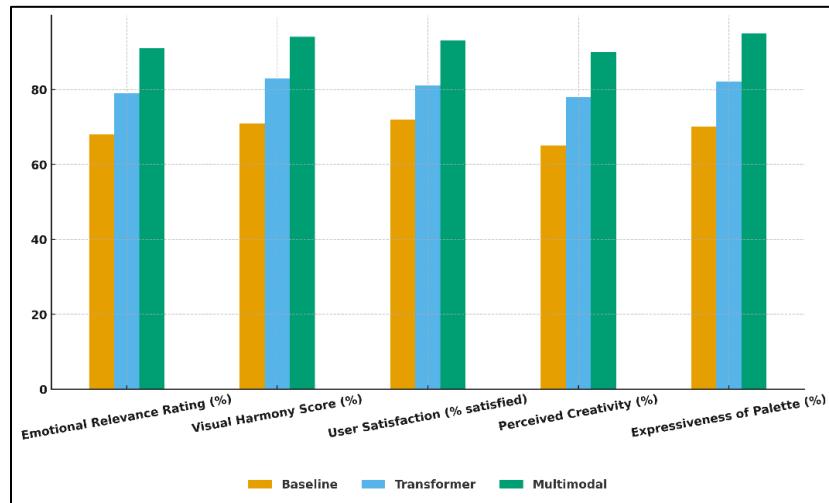
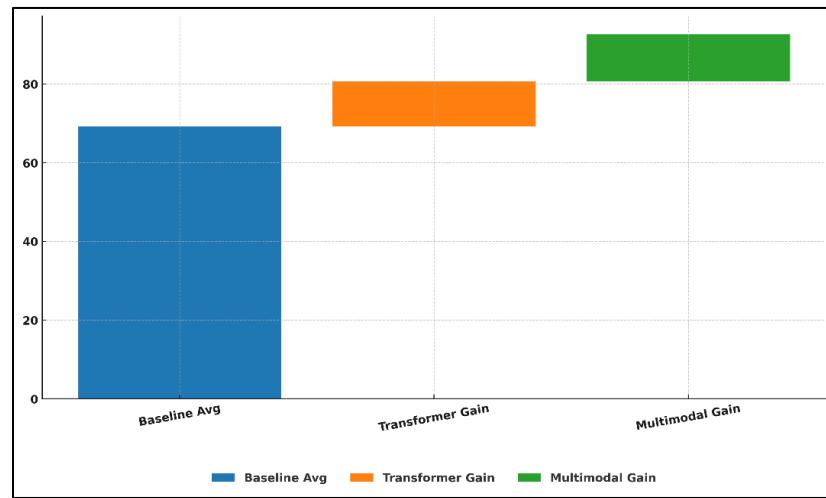
Figure 4**Figure 4** Comparative Evaluation of Color System Models Across User Experience Metrics

Figure 4 presents the comparisons of the user-experience metrics of various models of AI-driven color systems. The baseline system produces fairly low scores, with emotional relevance ranking at 6.8 and visual harmony at 71 percent, meaning that it has few capabilities of capturing subtle emotional signals or producing aesthetically harmonious palette. Also indicated by the user satisfaction (72%), and perceived creativity (65%), the compliance but not expressive influence is evident in the base system. Cumulative performance improvements between baseline and transformer based multimodal color systems have been illustrated in [Figure 5](#)

Figure 5**Figure 5** Cumulative Improvement Flow from Baseline to Transformer and Multimodal Color System

Transformer-based system shows a significant progress especially in emotional relevance (7.9) and perceived creativity (78%). Its textual sentiment comprehension is stronger which helps it to achieve palette expressiveness (82%), but it is still not able to realize a visual context comprehension thus achieving moderate progress in harmony and satisfaction.

6. CONCLUSION

The work presented herein indicates that sentiment analysis and computational color modeling are both capable of being combined to create a strong and innovative methodology of visual appeal prediction and creation of emotionally-associated color schemes. The proposed framework can help reduce the gap between affective cues and aesthetic design

principles because of the ability of convolutional neural networks to extract visual features and transformer-based models to understand the sentiment. Not only does the multimodal architecture capture the distribution of color features but also the emotive context of textual descriptions, so that the system can suggest palettes that can appeal to the emotional perception of humans. The experimental findings indicate that multimodal fusion is superior to unimodal models in accuracy and emotional congruence and provides a solid empirical evidence of the significance of integration of visual and linguistic modalities. The quantitative results showed that the prediction accuracy was high and the aesthetic score was better than before, whereas the qualitative user-study results proved that the generated palettes are indeed expressive in terms of emotions and visually harmonious. The method also gives theoretical evaluations of the mind and mental processes that interconnect emotion, color preferences and visual aesthetics. In addition to scholarly work, this study is of great practical importance to use in digital design, marketing, tailor-made interfaces, and emotion-sensitive creative tools.

CONFLICT OF INTERESTS

None.

ACKNOWLEDGMENTS

None.

REFERENCES

Feng, H., Sheng, X., Zhang, L., Liu, Y., and Gu, B. (2024). Color Analysis of Brocade from the 4th to 8th Centuries Driven by Image-Based Matching Network Modeling. *Applied Sciences*, 14, 7802. <https://doi.org/10.3390/app1477802>

Ge, R., Cong, L., Fu, Y., Wang, B., Shen, G., Xu, B., Hu, M., Han, Y., Zhou, J., and Yang, L. (2023). Multi-Faceted Analysis Reveals the Characteristics of Silk Fabrics on a Liao Dynasty DieXie Belt. *Heritage Science*, 11, 315. <https://doi.org/10.1186/s40494-023-01055-0>

Gypa, I., Jansson, M., Wolff, K., and Bensow, R. (2023). Propeller Optimization by Interactive Genetic Algorithms and Machine Learning. *Ship Technology Research*, 70, 56–71. <https://doi.org/10.1080/09377255.2022.2154908>

Huo, Q. (2021). An Archaeological Study of Female Costumes in the Song, Liao, and Jin Dynasties. Master's Thesis, Zhengzhou University, Zhengzhou, China.

Jia, J., Cao, J. W., Xu, P. H., Lin, R., and Sun, X. (2022). Automatic Color Matching of Clothing Patterns Based on Image Colors of Peking Opera Facial Makeup. *Journal of Textile Research*, 43, 160–166.

Lenzerini, F. (2023). The Spirit and the Substance: The Human Dimension of Cultural Heritage from the Perspective of Sustainability. In *Cultural Heritage, Sustainable Development and Human Rights* (pp. 46–65). London, UK: Routledge.

Lieto, A., Pozzato, G. L., Striani, M., Zoia, S., and Damiano, R. (2023). Degari 2.0: A Diversity-Seeking, Explainable, and Affective Art Recommender for Social Inclusion. *Cognitive Systems Research*, 77, 1–17. <https://doi.org/10.1016/j.cogsys.2023.06.004>

Liu, S., Liu, Y. K., Lo, K. C., and Kan, C. (2024). Intelligent Techniques and Optimization Algorithms in Textile Colour Management: A Systematic Review of Applications and Prediction Accuracy. *Fashion and Textiles*, 11, 13. <https://doi.org/10.1186/s40691-024-00345-7>

Ma, X., Chen, Y., Liang, Q., and Wang, J. (2024). Wardrobe Furniture Color Design Based on Interactive Genetic Algorithm. *BioResources*, 19, 6230–6246. <https://doi.org/10.15376/biores.19.3.6230-6246>

Malik, U. S., Micoli, L. L., Caruso, G., and Guidi, G. (2023). Integrating Quantitative and Qualitative Analysis to Evaluate Digital Applications in Museums. *Journal of Cultural Heritage*, 62, 304–313. <https://doi.org/10.1016/j.culher.2023.05.009>

Ni, M., Huang, Q., Ni, N., Zhao, H., and Sun, B. (2024). Research on the Design of Zhuang Brocade Patterns Based on Automatic Pattern Generation. *Applied Sciences*, 14, 5375. <https://doi.org/10.3390/app1453735>

Pak, H. (2025). Reiterating the Spirit of Place: A Framework for Heritage Site in VR. *International Journal of Advanced Smart Convergence*, 13, 337–354. <https://doi.org/10.7236/IJASC.2025.13.4.337>

Wang, D., and Xu, X. (2024). 3D Vase Design Based on Interactive Genetic Algorithm and Enhanced XGBoost Model. *Mathematics*, 12, 1932. <https://doi.org/10.3390/math12121932>

Wang, S., Cen, Y., Qu, L., Li, G., Chen, Y., and Zhang, L. (2024). Virtual Restoration of Ancient Mold-Damaged Painting Based on 3D Convolutional Neural Network for Hyperspectral Image. *Remote Sensing*, 16, 2882. <https://doi.org/10.3390/rs16172882>

Zhao, F. (2020). *Chinese Silk Art Through the Ages: Liao and Jin Dynasties*. Hangzhou, China: Zhejiang University Press.