https://doi.org/10.2352/ISSN.2470-1173.2021.11.HVEI-163

Prediction of individual preference for movie poster designs based on graphic elements using machine learning classification

Hyeon-Jeong Suk, Juhee Kim; Department of Industrial Design, KAIST; Daejeon, Korea
Chul Min Kim; KAIST; Daejeon, Korea

Abstract

We attempted to predict an individual's preferences for movie posters using a machine-learning algorithm based on the posters' graphic elements. We transformed perceptually essential graphic elements into features for machine learning computation. Fifteen university students participated in a survey designed to assess their movie poster designs ($N_{poster} = 619$). Based on the movie posters' feature information and participants' judgments, we modeled individual algorithms using an XGBoost classifier. We achieved prediction accuracies for these individual models that ranged between 44.70 and 71.70%, while the repeated human judgments ranged between 61.90 and 87.50%. We discussed technical challenges to advance prediction algorithm and summarized reflections on using machine learning-driven algorithms in creative work.

Introduction

The topic of individual preferences concerning graphics has received considerable attention for many decades. Traditional theories have been utilized in the graphic designs used in educational and practical fields. For instance, the Gestalt theory has been used to explain human perceptions of visual elements [1]. The Law of Prägnanz is a well-known principle that states that individuals commonly prefer simple forms that can be perceived more easily. Also, Lauer considered layouts by classifying visual elements in two-dimensional areas and described how certain feelings could be evoked by using compositions of basic graphic elements [2]. Based on these well-structured theories, graphic designs have played a significant role in marketing from past to present. Well-designed visuals have largely influenced product desirability [3], corporate image formation [4, 5] and advertising [6, 7]. Graphic designs can capture consumers' attention and promote products' sales when they are based on an understanding of what people commonly prefer.

Online media and various display devices expose people to graphic images regularly. Traditionally, graphic designs have been mainly expressed and provided in printed materials in that the graphic designs were propagated uniformly, i.e., in images of the same quality, to the public. However, as images are now sent to individual digital devices in real-time, graphic images have diversified and been given flexible characteristics that can change at any time. For instance, online media, such as Instagram and Pinterest, recommend images similar to those previously selected by a user [8, 9]. Facebook and Google can send advertising images to target groups via the use of the A/B testing method, which transmits multiple ad banner images simultaneously and monitors the ads' performance in real-time [10, 11].

Despite the importance of graphics, graphical properties have not been investigated much. Although some systems have put some effort into measuring the quality of marketing visuals, most studies have been focused on predicting marketing performance data, such as the click-through-rate [12, 13], rather than discovering the relationship between graphical characteristics and individual preferences. However, Artificial Intelligence (AI) technologies can support the predictions of preferences induced by images by analyzing large-scale data. These computational methods enable us to examine complex combinations of various types of visual elements. For instance, systems can predict users' preferred images [14, 15, 16] or to estimate feelings induced while users are viewing images [17, 18, 19].

A significant number of studies have been carried out in the computer science field regarding the computational approaches used to reveal aesthetic qualities and preferences. However, the experiments have usually been based on photographic images, which are insufficient when analyzing the various formative characteristics of graphic images and applying them to actual services. Besides, most studies have been conducted to reveal average preferences, while the latest AI studies have focused on optimizing individuals' tastes [20, 21, 22]. In this circumstance, this study aims to model individual's preferences in graphic images using machine learning-driven estimation.

In AI-driven attempts, the image features can be considered to predict human's preferences for images. By convention, the estimation considers color, texture, shape, and the corresponding semantic connotation thereof. In particular, the factors that can increase the aesthetic quality of, and preferences towards, images have been investigated to improve images' quality. For example, in a study conducted by Bianco[23], the Alignment-Free Facial Attribute Classification Technique (AFFACT) [24] was utilized to classify facial characteristics, such as hairstyle, accessories, and appearance, in a meaningful way. The results of these kinds of studies are actively applied in practical fields. Samsung Galaxy's single-take model recommends the best photo within 10 seconds of recorded videos based on AI algorithms [25]. Google Photos' style function auto-enhances the quality of a photo [26].

In preference studies, Pre-constructed image datasets, such as the Aesthetic Visual Analysis (AVA) dataset [27] and Aesthetic, and Attributes DataBase (AADB) [28], have been used frequently. By utilizing these datasets, studies have predicted average preferences by analyzing image features [29, 16, 30, 23]. Prediction accuracy has reached $r = 0.21 \sim 0.77$, a relatively lower range than those found in conventional machine learning studies. Real advertising banners were recently analyzed to predict the CTR (click-through rate) with deep learning approaches [12, 13]. These results have practical applications. For instance, Alibaba's LuBan generates high-quality banner images by learning many existing banner images [31].

Up to this point, most studies have been carried out on photographic images and paintings. Predicting the quality of an image is not straightforwardly applicable for design practice. Moreover, preference prediction has aimed at understanding public tendencies rather than individual differences. Because preferences vary from person to person, this study is centered on individual preferences.

Therefore this study intends to adopt a machine-learning algorithm to predict an individual's preference for graphic designs. In particular, we pay attention to involving graphic elements that designers consider in creating poster designs. Through this computational approach, we anticipate identifying which graphic elements will have driven the preference individually. Ultimately, we pursue actionable lessons and guides for an AI system to generate appealing visual designs in a customized manner and for designers to have early feedback about user preference.

Methods

We facilitated the machine-learning computation to related graphic elements within each poster design with individuals' subjective preference judgments selected among three categories, Like, Neutral, and Dislike. The participants' subjective judgments were considered as labels in the computation. The prediction modeling was individually performed, considering individually different preferences for graphic style. Furthermore, the prediction accuracy was compared with the participants' repeated assessment after a time interval, two weeks in this study. We admit that the human preference for a graphic design varies even within a person depending on the context the person is engaged in. The method details are described in the following.

Data collection

Data Set: Movie posters

We focused on movie posters as stimuli because they implicitly illustrate various feelings using combinations of graphics. Also, most of the movies are released with many poster alternatives to express the stories and characters in diverse view points. We collected 619 movie posters released from 1950 to 2019 from the NAVER Movies and Metacritic website.

Label Assignment: Preferences Survey

We recruited 15 participants and conducted a survey regarding their preferences for the 619 movie posters. The participants had diverse professional backgrounds, such as design, electronic engineering, and computer science. Their average age was 25.93 years old, with a standard deviation of 5.97 years. Participants were recruited via an online university community. All subjects were paid volunteers, and all of them had a normal or corrected-to-normal vision.

We printed poster images on 5 x 7-inch glossy photo cards and then asked participants to divide the poster cards into three groups: Like, Neutral, and Dislike. Also, We allowed them to spend as much time on the assessment as they needed, and each spent around 30 minutes to judge 619 movie posters.

Also, we conveyed individual interviews with all 15 participants about why they liked or disliked the posters. The moderator provided detailed questions when the answers provided were too general or not descriptive enough. The surveys and interviews were done individually. The poster classifications and interview contents were recorded on video, and the interviews were transcribed for the content analysis.

Prediction baseline

Next, we set the repeatability of human judgments as a baseline of prediction accuracy, with a focus on the variability involved, as a previous study noted concerns in this area [32]. Some recent machine-learning studies have indicated that the prediction accuracy regarding the emotional reactions to images is around 68% [33] or 67% [34], suggesting that human emotions keep changing. Among the 15 participants, four participants volunteered to re-judge the 619 posters after a two-week interval. We did not force all participants to judge the posters due to cognitive loads. Based on four volunteers' repeated assessments for the preference, we figured out the repeatability ration between 61.9% and 87.5% as presented in Table 1.

Table 1. Human repeatability rates for the four volunteer participants

Participant	P1	P2	P3	P4	Avg.
Repeatability	70.0%	87.5%	61.9%	65.0%	71.1%

Feature construction

To analyze graphic design characteristics in a formative way, we began with basic elements such as points, lines, faces, colors, and typography. We also considered the principles concerning the interactions among the elements, thereby focusing on uniformity, balance, and rhythm [2, 35, 36]. They are the principles of graphic design that communication design education is still following. To transform the graphic elements and their interactions, we adopted the graphic elements used in a previous study [37] and had experienced designers sort the movie posters based on these characteristics. The study claimed that color- and layout-related graphic elements were essential and that the human presence had a high impact. In the study, the research team attempted to discover what kinds of graphic elements were considered when distinguishing between posters. We adopted the graphic elements as well as the numerical transformations of these elements. We also included information on the movies themselves, such as genre and release year. Table 2 presents the list of features. Hence, the features consisted of both categorical and continuous scales.

Color-related features

The color-related features concerned hue, brightness, and saturation. When articulating the color-related features, we used the CIELab system to match human color perceptions more straightforwardly compared to other digital color systems [38]. Also, we extracted key perceptual colors of poster images while weighting the vivid colors. We mainly utilized the mean values of all pixels and key colors and their standard deviations to examine contrasts in images.

Overall, we adopted the following steps while observing a designer's key-color extraction process. Firstly, we converted the RGB values of image pixels into the L, a, b values used in the CIELab system. We then weighted the vivid colors ($Chroma >$

Table 1. The features used in preference-prediction machine-learning modeling

Category		Features
Color	All-pixels	Brightness, Hue(red-greenishness, yellow-bluishness), Chroma, Contrast in Brightness, Contrast in hue (red-greenishness, yellow-bluishness), Contrast of chroma
	Key-colors	Brightness, Hue(red-greenishness, yellow-bluishness), Chroma, Contrast in Brightness, Contrast in hue (red-greenishness, yellow-bluishness), Contrast in chroma, The number of key colors
Layout	Size	The number of humans, Size of humans, Size of title, Size of other text, Sizes of the main objects
	Position	Positions of humans, Average position of humans (horizontal, vertical), Average position of humans and the title (horizontal, vertical)
	Balance	Weighted positions of humans (horizontal, vertical), Weighted position of the title (horizontal, vertical)
	Dispersion	Position variability for major objects (horizontal, vertical), Weighted position variability for major objects (horizontal, vertical)
Style		Simplicity
Movie profile		Release year, Audience review score, Expert review score, Genre of movie

70). Next, we created a pool of color candidates using the K-means clustering of all pixel colors and weighted vivid pixel colors. Thirdly, we ordered the extracted colors according to their cluster sizes and computed the color differences as "delta E's," i.e., the Euclidean distances between the first and second colors. In each instance, we accepted the second color only when the color difference between it and the first color exceeded the threshold ($\Delta E = 27$). Only those colors passing the threshold test remained in the end.

Layout-related features: human and text

In the poster designs, the human presence and text components play major roles. Therefore, we searched for deep learning techniques that could identify the human and text properties automatically. In the end, we used Dual Shot Face Detector (DSFD) [39], Character Region Awareness for Text Detection (CRAFT) [40], Mask R-CNN [41], and Torchvision in Python. However, we inevitably carried out manual corrections because the human shapes and text arrangements in the posters were often expressed in such a manner that they were beyond the abilities of deep learning-based recognition. As shown in Figure 1, the current techniques can well-detect the photographic human faces and non-calligraphy typefaces. However, it had some trouble to detect stylized carton humans, humanized animal characters, and calligraphy texts. In such cases, we manually corrected the results by defining the areas with red rectangle boxes in the Shiny library in R. All posters were screened to avoid any misdetection.

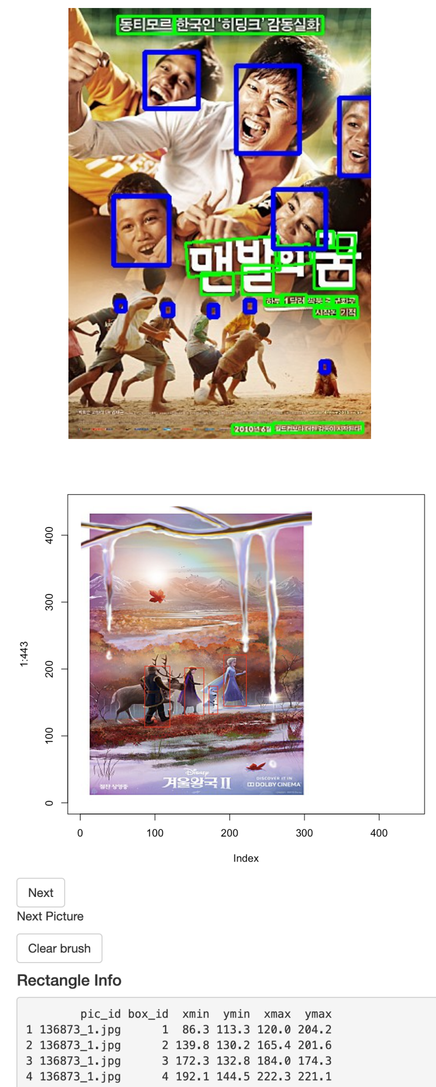

Figure 1. (Top) Face and text detection via deep learning, (Bottom) Manual detection is performed for supplementary feature information. Shiny library in R is used to facilitate this step

Concentrating on the human (humanized characters) and text areas, we calculated object positions, sizes, balance, and dispersion, all of which are related to the basic layout theories. The object positions were estimated by averaging the position values of the targeted objects. The object sizes were calculated by the summing the sizes of the targeted objects. The object balance was calculated by weighting sizes by position. Lastly, object dispersion was estimated with the deviation in the positions of the objects.

Graphic style: simplicity

In a study conducted by Kim and Suk [37], perceptions of complexity and simplicity were estimated via edge proportions using the Canny edge method [42]. The edge proportion is proportional to the outline amount implying the complexity of a poster.

Movie profiles

Apart from the perception-based graphic elements, we scraped movie profiles and used features such as the release year, review score, and genre to account for the movies' influences on poster preferences. Release year and score were measured using continuous scales, but the genre information was categorical. Also, a movie could belong to more than one genre. We collected the genre-related features in the form of TRUE or FALSE.

In summary, we collected 59 features, of which 17 were color-related, 20 were layout-related, and 21 were movie-related properties, while one was a graphical style-related property.

Classification

We chose a decision tree classifier, XGboost in the Scikit-learn, and using Python for the classification. As the decision tree approach, the XGboost helps understand where the decisions were made. Often called "white box," in contrast with "black box" methods, the interpretable machine learning methods have received increasing attention, especially in the creative domains. Consequently, despite over-fitting and bias issues, the decision trees highlight consequences, including chance event outcomes, resource costs, and utility scores.

Also, the XGBoost is the latest and popular ensemble method developed by [43]. It uses an extreme gradient boosting framework. The gradient boosting introduced by [44] was invented to solve a prediction model problem in the form of an ensemble of weak prediction models, typically decision trees. XGBoost usually yields superior results among diverse methods using fewer computing resources in the shortest amount of time.

In modeling the algorithm, we validated the trained individual models with ten times cross-validation, considering the random computational procedures, as shown in Table 4.

Results

Feature importance

When the XGboost modeled the individual algorithm individually, we observed that the importance of features varied remarkably from person to person. In Table 3, the top five important features are presented, and individuals considered different aspects when they assessed the preference for the movie poster designs. For example, P2 considered the edge proportion the most relevant, indicating that the poster design's simplicity or complexity is a critical aspect of the preference. P2 did not consider the genre of the movie. On the contrary, P8 showed a particular concern regarding the animation genre, probably apart from the poster aesthetics. Overall, P8 highly depended on the genre types rather than graphic elements of the poster design compared with P2. Interestingly, P4, P9, P10, P12, P13, and P15 were highly influenced by the "release year" than other features, implying that their preferred styles can group users or consumers.

Table 5. The top five important features in the individual movie poster preference models

	Important features
P1	the number of humans, sizes of title and other text, contrast in brightness, position variability of major objects
P2	simplicity, expert review score, size of humans, other text, and title
P3	weighted position variability of major objects, release year, contrast in hue, musical genre, contrast in brightness
P4	release year, position variability of major objects, thriller genre, size of other text, and major objects
P5	size of other text, contrast in hue, contrasts in key colors' hue and Chroma, thriller genre
P6	weighted position variability of major objects, sizes of major objects, contrast in key color's hue, expert review score, positions of major objects
P7	hue contrast, release year, animation genre, contrasts in key colors' hues, adventure genre
P8	animation genre, contrast in key color's hue, release year, thriller genre, family genre
P9	release year, contrast in hue, size of the title, simplicity, comedy genre
P10	release year, size of other text, contrasts in key colors' Chroma, position and size of the title
P11	size of other text, simplicity, adventure genre, SciFi genre, romance genre
P12	release year, musical genre, size of other text, position of the title, SciFi genre
P13	release year, sizes of major objects, Chroma, brightness, Chroma of key colors
P14	size of other text, contrast in Chroma, release year, weighted positions of humans and the title
P15	release year, sizes of humans, sizes of major objects, weighted position variability of major objects, expert review score

Accuracy

We admit that human's preference for graphic artwork can vary depending on one's mood and change over time. In this circumstance, we consult the human repeatability as the baseline, as shown in Table 1. Then, we compare the accuracy obtained from the machine-learning algorithm with the human repeatability. The average prediction accuracy across the 15 individual models from the XGBoost algorithm is 57.29%. In particular, the P2 model has the highest accuracy score at 71.7%, as presented in Table **??**). When the accuracy performance is compared with the human repeatability, the performance of 57.29% can be regarded as 80.58% (57.29% / 71.1%).

Table 4. Individual preference prediction classified with XG-Boost algorithm. Ten times of computation were conveyed, and the ranges of scores are in the parentheses. The Baseline refers to the result shown in Table 1.

Participant	Accuracy	Baseline
P1	67.2% (62.0-70.5)	70.0 %
P2	71.7% (60.6-73.4)	87.5 %
P3	59.2% (52.0-77.5)	61.9 %
P4	53.7% (43.1-61.2)	65.0 %
P5	51.9% (43.7-61.2)	-
P6	44.7% (36.7-52.9)	-
P7	54.7% (32.0-67.0)	-
P8	63.8% (54.9-68.0)	-
P9	51.2% (33.3-62.0)	-
P10	52.7% (44.0-67.3)	-
P11	53.1% (44.8-65.3)	-
P12	64.0% (60.7-74.5)	-
P13	50.3% (42.0-65.3)	-
P14	53.2% (48.9-62.0)	-
P15	55.1% (51.0-62.0)	-

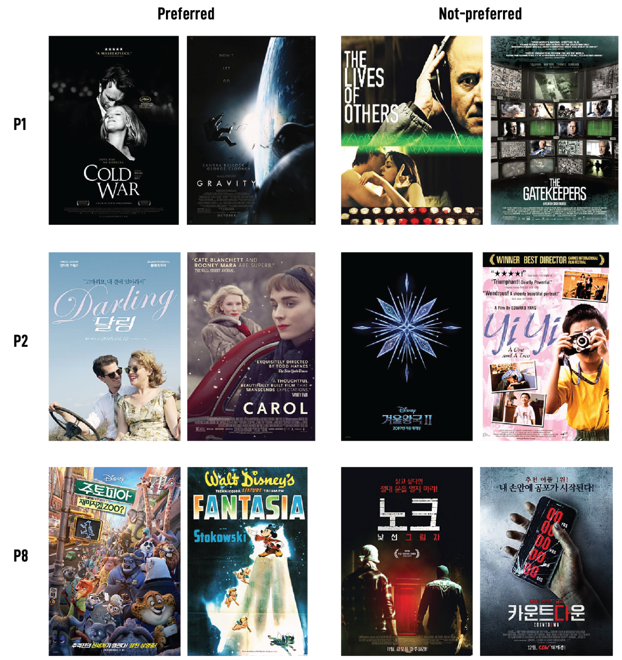

Figure 2. The posters that were preferred and not preferred by both humans- P1, P2, and P8- and individual machine learning models

Interview with participants

When participants completed the assessments, we made individual interviews additionally. We mainly asked about any influential graphical or informational features they had considered in making preference judgments.

In the case of P1(67.20%), she said that the graphical characteristics were more important than the movies themselves. Arbitrary, though, she felt that the combination of poster elements was meaningful. When we figured out the feature importance(see Table 3), the movie-related features such as genre or release year were not highly relevant. Alternatively, P2 (71.70%) said that she disliked complex posters, and indeed the simplicity feature was highly influential in the model. Besides, in P8's model (63.80%), the movie genre affected the results dominantly. During the interview with P8, he often mentioned the movie genre, making such statements as "I like animations," "I do not like horror movies," or "I like movies that address social issues." In Figure 2, two pairs of posters are shown for P1, P2, and P8. The selected posters are examples that three participants assessed as preferred or not-preferred. At the same time, the machine-learning algorithm could predict the preference concerning those posters accurately. Resuming both quantitative accuracy and qualitative interview analysis, we found a potential of AI-driven algorithm modeling that predicts individual preference. Currently, not yet at a satisfactory level of accuracy, but there seems a potential that an algorithm can predict customers' preferences through keen observation of the individual's appraisal for style taste.

Discussion

Contribution to imaging research

Quantitative approach for considering many graphical combinations

Communicating image-related issues has occasionally been a gray area for designers, decision-makers, and consumers because graphic images are difficult to analyze quantitatively and describe verbally. Therefore, design decisions have relied on individual subjective judgments based on knowledge and experience. Although emerging image analysis technology can computationally interpret images, it is largely used for object detection purposes or image enhancements.

However, machine learning techniques make it possible to deal with complex combinations of design elements. Design studies have adopted traditional analysis methods, such as conjoint and regression analyses. However, the results have been limited to the small number of independent variables included in the stimuli. The machine learning approaches can provide more options for improving designs while avoiding limiting designers' creativity since the models can consider various graphical combinations. When a study employs algorithms that transparently offer how the decision has been made, i.e., white-box algorithms, designers may better understand weighted graphic elements included in the modeled algorithms. Simultaneously, what algorithm accounts for human'shuman decision-making, such as preference, can inspire the human to generate creative outputs.

The adoption of user research methodologies in machine learning

The latest AI trend is to adopt deep learning features and methods to increase prediction performances. This method has shown remarkable results compared to the previous feature engineering method. Nevertheless, the process of deep learning is hidden behind layers, and humans cannot understand or evaluate the machine decision processes. Furthermore, researchers who often utilize deep learning approaches do not understand the processes but focus on the results only. However, when it comes to image creation, the process matters. Imaging research does strive for one ultimate answer. We do not only investigate whether the animal in an image is a cat or dog but pursue exploring why this animal is adorable to people. As confirmed in our research, there could be many reasons why certain movie posters are preferred. We are studying people, and human beings are complex. Thus, investigating ”why” and ”how” is meaningful.

In this study, we utilized the features developed from user studies involving workshops and in-depth interviews. We attempted to build a linkage between features and results, while previous machine learning studies have focused on finding unexplainable patterns to achieve higher performances. This study showed that interpretable machine learning models could be improved by meticulously following people’s mental processes.

Capturing individual preferences

The relationship between visual attributes and human preference has been studied mainly in terms of average tendencies until recently. Products had to satisfy the tastes of the majority of people for mass production. In recent years, small-batch production has become a part of the mainstream in various industries [45]; thus, understanding individuals’ emotions and preferences has become a new business tactic nowadays. Online media can present content based on the user’s mood and preferences. Graphic images are not an exception. By investigating individuals’ emotions and preferences, graphic images can be adjusted to deliver brands or products effectively. According to our previous research [46], graphical tastes should be examined individually.

Limitations and future research

Developing a visual guideline generation system

Against the fuzzy backdrop of AI, some attempts at explanation have received attention. For example, machine learning methods based on decision trees, although somewhat classical, can point to where decisions occurred explicitly. As white-box models, they provide information that might inspire humans to create new methods or items. In particular, they might help designers generate new concepts. Human creativity can benefit from how machines make decisions based on complex data. The human search for causation is partially satisfied by some new approaches. For example, SHAP (SHapley Additive exPlanations) [47] can explain what elements in a poster’s design influenced judgments and how these elements were quantified in a specific case. Hence, design guidelines can be developed based on an understanding of the decision-making processes of the machine.

This study produced some evidence on how we can make links between prediction results and graphical features using the feature exploration process. Thus, we can conclude that specific graphical attributes affect people’s preferences. Nevertheless, in future studies, we should explore methods that can address how the machine maintains high prediction quality. Although interpretable machine learning has been investigated recently [48], it has yet to be described as user-friendly. At this point, a designer can contribute to human-centered machine learning.

Limited data collection for personalized modeling

By convention, machine learning models are developed with large-scale data, usually provided by public databases or systems. However, no open datasets contained the individual data that we wanted to work with. This study required individual datasets to predict the individuals’ preferences regarding movie posters. Consequently, we surveyed to collect the necessary data: preference responses regarding a set of movie posters. However, we could only collect a limited number of responses from individuals due to cognitive fatigue. In particular, participants refused to judge the posters repeatedly. Therefore we could not force them to finish all processes. We managed to acquire the human repeatability from four participants and used that as the ”baseline” to evaluate models’ performances. Unfortunately, such data collection issues have arisen widely in artificial intelligence research [49]. Although some researchers have utilized collaborative data acquisition processes or transformed data acquired via web scraping, it is still difficult to acquire a rich dataset without collaborating with large companies or governments.

Furthermore, we did not correctly consider the pre-exposed movie posters, although participants’ responses could be biased. Due to technical limitations, the manual correction process of movie posters was very time-consuming. Current face and text detection algorithms are not appropriate for movie posters yet, as previously mentioned. In the future study, we may adopt the AI models that are trained for the graphic domain.

Developing secondary and semantic features

We examined features that reflected design theories, mainly focusing on color- and layout-related features. Compared to the color features, layout features might be treated as features on a different level since these features are found through human and text processes, while color features are extracted directly from pixel values. Nevertheless, we defined them as primary features because color and layout are fundamental elements in graphic design. Based on our observations, features can be transformed into numerical formats. However, we have confirmed from the interviews that secondary features, such as creativity, aesthetic quality, sophistication, or level of devotion, should be considered. These features do not have proper calculation methods yet, but they can be defined with combinations of primary features in future research. Human-centered approaches, such as observing humans’ decision processes, can be utilized to discover essential secondary features.

Also, since the stimuli were movie posters, the movies themselves bias the subjects’ responses. Some participants may have already watched some of the movies; therefore, background information affected their decisions. The reputation of a film and the popularity of an actor can also be significant factors. Additionally, the objects and texts in the images influence preferences. For example, if one actor is holding a red rose and another actor is holding a bloody knife in two different images, the messages conveyed will vary regardless of their similar color and layout combinations. Facial expressions and body postures of actors can also be influential.

Although we adopted some emotions and movie profiles, additional semantic and secondary features should be considered to increase models’ performances. As for the secondary features, machine recognition of the objects in images is yet in an early stage, especially for graphic designs. They are too premature to use in a real system, and the algorithms can produce much misleading information. Hence, we could manually input such information or scrape semantic data corresponding to images to generate well-founded features. We can develop secondary features by investigating the impact of primary graphical features.

Developing a graphic design auto-generation system

In this research, we explored the important aspects of machine learning analysis. Based on the results, designers can improve graphics by utilizing preferences. However, it is still difficult to understand the machine learning interpretation process. Therefore, we may consider developing a graphic auto-generation system in which the machine can capture complex graphic elements. In the online advertising area, banner auto-generation methods designed to reduce routine jobs have been researched. Alibaba has an AI banner design system called LuBan [50]. The system can make banners by training on numerous human-created banners. Google provides a similar auto-generated banner function in Google Ads [51]. Also, [52] developed a banner image editing method that can be utilized under various banner ratios. Finally, [53] presented an interactive layout suggestion system.

However, these systems remain in their early stages. They are for adjusting graphical inputs within prepared design templates or limited formats. The systems cannot creatively utilize various graphical skills but can only create designs based on understanding the relationships among design elements' colors and position values. Although these systems cannot replace human creativity at this point, they can still free people from iterative and routine tasks. Also, these systems can support people without the financial resources to hire design and marketing professionals. By combining our approach with the current system, marketing efficiency can be increased with reasonable quality graphics.

Conclusion

We attempted to predict movie posters' preferences through individually- optimized machine learning models that used basic graphical features. To develop the algorithm, we collected preference responses from 15 participants regarding 619 movie posters. The XGBoost decision tree classifier was applied. The average prediction accuracies of the individual models ranged from 44.70 to 71.70%. The results showed improvement when meaningful features were adopted.

This study was limited to small data set and did not consider every semantic feature. Nevertheless, this research demonstrated how design research could utilize AI-driven techniques to gather meaningful information relating to design practice when dealing with many graphic elements at once. This approach shows that it is possible to compute many combinations of design attributes. Thus, design research can be conducted more rigorously by utilizing machine learning methods.

By adopting the machine learning-driven approach, design research can overcome the qualitative analysis limitations of previous studies. The study demonstrated the potential of data-driven design, thus assisting designers.

References

[1] W. Köhler, "Gestalt psychology," *Psychologische Forschung*, vol. 31, no. 1, pp. XVIII–XXX, 1967.

[2] D. A. Lauer and S. Pentak, *Design basics*. Cengage Learning, 1979.

[3] B. Rundh, "Packaging design: creating competitive advantage with product packaging," *British Food Journal*, 2009.

[4] P. Foroudi, T. Melewar, and S. Gupta, "Linking corporate logo, corporate image, and reputation: An examination of consumer perceptions in the financial setting," *Journal of Business Research*, vol. 67, no. 11, pp. 2269–2281, 2014.

[5] G. Ad, V. Adr, N. E. Pascu, *et al.*, "Logo design and the corporate identity," *Procedia-Social and Behavioral Sciences*, vol. 51, pp. 650–654, 2012.

[6] T. Altstiel, J. Grow, and M. Jennings, *Advertising creative: Strategy, copy, and design*. Sage Publications, 2018.

[7] R. Lohtia, N. Donthu, and E. K. Hershberger, "The impact of content and design elements on banner advertising click-through rates," *Journal of advertising Research*, vol. 43, no. 4, pp. 410–418, 2003.

[8] S. Wang, Y. Wang, J. Tang, K. Shu, S. Ranganath, and H. Liu, "What your images reveal: Exploiting visual contents for point-of-interest recommendation," in *Proceedings of the 26th International Conference on World Wide Web*, pp. 391–400, 2017.

[9] Y. Jing, D. Liu, D. Kislyuk, A. Zhai, J. Xu, J. Donahue, and S. Tavel, "Visual search at pinterest," in *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1889–1898, 2015.

[10] Facebook, Inc., "About a/b testing." `https://www.facebook.com/business/help/1738164643098669?id=445653312788501`. Accessed: 2020-03-02.

[11] Google LLC, "Run experiments in adsense." `https://support.google.com/adsense/answer/6321879?hl=en/`. Accessed: 2020-03-02.

[12] K.-W. Park, J. Lee, S. Kwon, J.-W. Ha, K.-M. Kim, and B.-T. Zhang, "Which ads to show? advertisement image assessment with auxiliary information via multi-step modality fusion," *arXiv preprint arXiv:1910.02358*, 2019.

[13] J. Chen, B. Sun, H. Li, H. Lu, and X.-S. Hua, "Deep ctr prediction in display advertising," in *Proceedings of the 24th ACM international conference on Multimedia*, pp. 811–820, 2016.

[14] N. Ma, A. Volkov, A. Livshits, P. Pietrusinski, H. Hu, and M. Bolin, "An universal image attractiveness ranking framework," in *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp. 657–665, IEEE, 2019.

[15] D. Kuzovkin, T. Pouli, R. Cozot, O. Le Meur, J. Kervec, and K. Bouatouch, "Image selection in photo albums," in *Proceedings of the 2018 ACM on International Conference on Multimedia Retrieval*, pp. 397–404, 2018.

[16] J. Ren, X. Shen, Z. Lin, R. Mech, and D. J. Foran, "Personalized image aesthetics," in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 638–647, 2017.

[17] S. Corchs, E. Fersini, and F. Gasparini, "Ensemble learning on visual and textual data for social image emotion classification," *International Journal of Machine Learning and Cybernetics*, vol. 10, no. 8, pp. 2057–2070, 2019.

[18] S. Zhao, X. Zhao, G. Ding, and K. Keutzer, "Emotiongan: Unsupervised domain adaptation for learning discrete probability distributions of image emotions," in *Proceedings of the 26th ACM international conference on Multimedia*, pp. 1319–1327, 2018.

[19] S. Zhao, H. Yao, Y. Gao, G. Ding, and T.-S. Chua, "Predicting personalized image emotion perceptions in social networks," *IEEE transactions on affective computing*, vol. 9, no. 4, pp. 526–540, 2016.

[20] H. Liu, R. Wang, S. Shan, and X. Chen, "Learning multifunctional binary codes for personalized image retrieval," *International Journal of Computer Vision*, pp. 1–20, 2020.

[21] W. Zhang, Z. Wang, and T. Chen, "Personalized image recommendation with photo importance and user-item interactive attention," in *2019 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*, pp. 501–506, IEEE, 2019.

[22] A. Kovashka and K. Grauman, "Attribute adaptation for personalized image search," in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 3432–3439, 2013.

[23] S. Bianco, L. Celona, and R. Schettini, "Aesthetics assessment of images containing faces," in *2018 25th IEEE International Conference on Image Processing (ICIP)*, pp. 2820–2824, IEEE, 2018.

[24] M. Günther, A. Rozsa, and T. E. Boult, "Affact: Alignment-free facial attribute classification technique," in *2017 IEEE International Joint Conference on Biometrics (IJCB)*, pp. 90–99, IEEE, 2017.

[25] SAMSUNG ELECTRONICS CO., LTD., "Galaxy s20 camera." `https://www.samsung.com/global/galaxy/galaxy-s20/camera/`. Accessed: 2020-04-07.

[26] Google LLC, "Google photos." `https://www.google.com/photos/about/`. Accessed: 2020-04-07.

[27] N. Murray, L. Marchesotti, and F. Perronnin, "Ava: A large-scale database for aesthetic visual analysis," in *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2408–2415, IEEE, 2012.

[28] S. Kong, X. Shen, Z. Lin, R. Mech, and C. Fowlkes, "Photo aesthetics ranking network with attributes and content adaptation," in *European Conference on Computer Vision*, pp. 662–679, Springer, 2016.

[29] A. Khosla, A. Das Sarma, and R. Hamid, "What makes an image popular?," in *Proceedings of the 23rd international conference on World wide web*, pp. 867–876, 2014.

[30] Y. Deng, C. C. Loy, and X. Tang, "Image aesthetic assessment: An experimental survey," *IEEE Signal Processing Magazine*, vol. 34, no. 4, pp. 80–106, 2017.

[31] Alibaba Cloud, "Luban." `https://luban.aliyun.com/`. Accessed: 2020-04-07.

[32] R. Datta, D. Joshi, J. Li, and J. Z. Wang, "Studying aesthetics in photographic images using a computational approach," in *European conference on computer vision*, pp. 288–301, Springer, 2006.

[33] H. Zhang, E. Augilius, T. Honkela, J. Laaksonen, H. Gamper, and H. Alene, "Analyzing emotional semantics of abstract art using low-level image features," in *International Symposium on Intelligent Data Analysis*, pp. 413–423, Springer, 2011.

[34] R. Swanson, D. Escoffery, and A. Jhala, "Learning visual composition preferences from an annotated corpus generated through gameplay," in *2012 IEEE Conference on Computational Intelligence and Games (CIG)*, pp. 363–370, IEEE, 2012.

[35] C. Leborg, *Visual Grammar: A Design Handbook (Visual Design Book for Designers, Book on Visual Communication)*. Princeton Architectural Press, 2006.

[36] A. W. White, *The elements of graphic design: space, unity, page architecture, and type*. Skyhorse Publishing, Inc., 2011.

[37] J. Kim and H.-J. Suk, "Prediction of the emotion responses to poster designs based on graphical features: A machine learning-driven approach," *Archives of Design Research*, vol. 33, no. 2, 2020.

[38] Commission internationale de l'eclairage, "Commission internationale de l'eclairage proceedings, 1976, colorimetry–part 4: Cie 1976 l* a* b* colour space," 1976.

[39] N. Li and Z. Xiong, "Automated brain tumor segmentation from multi-modality mri data based on tamura texture feature and svm model," in *Journal of Physics: Conference Series*, vol. 1168, p. 032068, IOP Publishing, 2019.

[40] Y. Baek, B. Lee, D. Han, S. Yun, and H. Lee, "Character region awareness for text detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 9365–9374, 2019.

[41] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask r-cnn," in *Proceedings of the IEEE international conference on computer vision*, pp. 2961–2969, 2017.

[42] J. Canny, "A computational approach to edge detection," *IEEE Transactions on pattern analysis and machine intelligence*, no. 6, pp. 679–698, 1986.

[43] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pp. 785–794, 2016.

[44] J. H. Friedman, "Greedy function approximation: a gradient boosting machine," *Annals of statistics*, pp. 1189–1232, 2001.

[45] K. Muita, M. Westerlund, and R. Rajala, "The evolution of rapid production: How to adopt novel manufacturing technology," *IFAC-PapersOnLine*, vol. 48, no. 3, pp. 32–37, 2015.

[46] J. Kim, J. Kim, and H.-J. Suk, "Layout preference for korean movie posters: Contextual background or character dominance," *International Association of Societies of Design Research Conference 2019: DESIGN REVOLUTIONS*, 2019.

[47] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in *Advances in neural information processing systems*, pp. 4765–4774, 2017.

[48] C. Molnar, *Interpretable machine learning*. Lulu. com, 2019.

[49] Y. Roh, G. Heo, and S. E. Whang, "A survey on data collection for machine learning: a big data-ai integration perspective," *IEEE Transactions on Knowledge and Data Engineering*, 2019.

[50] X.-S. Hua, "Challenges and practices of large scale visual intelligence in the real-world," in *Proceedings of the 26th ACM international conference on Multimedia*, pp. 364–364, 2018.

[51] Google LLC, "Google ads." `https://ads.google.com/home/`. Accessed: 2020-06-04.

[52] Y. Zhang, K. Hu, P. Ren, C. Yang, W. Xu, and X.-S. Hua, "Layout style modeling for automating banner design," in *Proceedings of the on Thematic Workshops of ACM Multimedia 2017*, pp. 451–459, 2017.

[53] P. O'Donovan, A. Agarwala, and A. Hertzmann, "Designscape: Design with interactive layout suggestions," in *Pro-*

ceedings of the 33rd annual ACM conference on human factors in computing systems, pp. 1221–1224, 2015.

Author Biography

Hyeon-Jeong Suk received her B.S. and M.S. Industrial Design degrees from KAIST in 1998 and 2000, respectively. Having practiced in the industry as a designer, she received a doctoral degree in social science major in psychology from the University of Mannheim in Germany in 2006. She is currently an associate professor at KAIST, leading a color laboratory. She acts as editor-in-chief of the Korea Society for Color Studies and an executive board member of the Korea Society for Emotion and Sensibility.

Juhee Kim received a B.A.(2010) degree in visual communication design from Kyunghee University and M.S.(2020) in Industrial design from Korea Advanced Institute of Science and Technology (KAIST). She has worked in the IT industry for over ten years. Her research interest lies in emotional user experiences regarding various visual stimuli.

Chul Min Kim received B.S. and M.S. Nuclear and Quantum Engineering degrees from KAIST in 2014 and 2016, respectively. His research interest lies in machine learning and data science based on graphic and bio-signal elements, including electroencephalography(EEG) and facial expression.