IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received February 16, 2021, accepted March 6, 2021, date of publication March 10, 2021, date of current version March 23, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3065407

Smart Culture Lens: An Application That Analyzes the Visual Elements of Ceramics

JI HYUN YI[1], WOOJIN KANG[1], SONG-EI KIM[1], DOYUN PARK[1], AND JIN-HYUK HONG[1,2]
[1]School of Integrated Technology, Gwangju Institute of Science and Technology (GIST), Gwangju 61005, Republic of Korea
[2]AI Graduate School, Gwangju Institute of Science and Technology (GIST), Gwangju 61005, Republic of Korea

Corresponding author: Jin-Hyuk Hong (jh7.hong@gist.ac.kr)

This research is supported by Ministry of Culture, Sports and Tourism (MCST) and Korea Creative Content Agency (KOCCA) in the Culture Technology (CT) Research & Development Program 2021 under Grant R2020040174.

ABSTRACT The Smart Culture Lens developed in this study is an application developed by utilizing the visual element classification system of ceramic and AI technology. The user can analyze the visual elements of the ceramic photo taken with a smartphone and search for similar ceramics related to each visual element. For this development, as a first step, visual elements such as color, form, material, and pattern were defined as criteria for classifying ceramic appearance, and the visual element classification system of ceramics was organized. In the second step, 19,610 images of 7,346 ceramics were collected through museum visit photography and web, and a database was built by annotating these images with a visual element classification system. In the third step, representative object detection models, Faster R-CNN and Mask R-CNN were trained based on a ceramic classification system. Through those trained object detection models, visual elements and masks of the input image were recognized, and representative colors of the area were extracted using the k-means algorithm through the recognized masks. The performance of the trained object detection models (Average precision of Form / Material / Pattern 1st-level category = 0.87 / 0.89 / 0.72) shows that the amount of collected data and the established classification system are useful. Finally, by applying the above development results, a mobile application called 'Smart Culture Lens' was developed, and the usefulness of this application was confirmed through a user experience test. This study combines AI technology into cultural heritage so that people can intuitively explore artifacts from a new perspective, which differs from traditional artifact exploring methods. All the detailed processes of this development will be a guide to how to apply AI technology to the cultural heritage.

INDEX TERMS Classification scheme of ceramics, visual elements analysis, visual search, ceramic dataset.

I. INTRODUCTION

The appearance analysis of artifacts refers to the information from the perspective of purposeful thinking of design. Artifacts are products related to food, clothing, and shelter in human life. Therefore, the appearance of the artifacts should be analyzed in the context of the result of purposeful design beyond judgment about the aesthetic value [1], [2]. People design and make things with color, form, and material that are needed to realize a function according to their purposes. They understand the purposes and functions of an object by looking at the appearance elements of the object [3]–[5]. Therefore, the analysis of objects from the perspective of purposeful thinking of design is the starting point of the fundamental analysis from the viewpoint of visual elements. That is why education about visual elements is the first step in fine arts and design education [6]–[8]. Therefore, the analysis of the visual elements of artifacts is the most fundamental and crucial element.

When we visited museums, we can see real artifacts and read or hear their information about manufacturing time, formative features, and historical background. In the field of cultural heritage, the information on artifacts is generally based on archeological studies about the periods and geographical features or other scientific studies that analyze the materials or the manufacturing methods of the artifacts. However, these details can be easily obtained from the internet search and online museum by using search keywords, even without viewing the real objects. In this study, we focus on analyzing the appearance information of artifacts viewed by the users

The associate editor coordinating the review of this manuscript and approving it for publication was Gulistan Raja.

TABLE 1. Comparison of visual information providing services based on image search technologies.

	Art Palette (Google)	Runway Palette (Google)	Beyond Scrolls & Screens (Google)	Artimage, Saatchi Art (Visii)	ArtPI (Artendex)	Smart Culture Lens
Target	Artwork	Fashion show	Japanese painting	Modern artwork	Artwork	Korean ceramic
Database features	Refined single image	Refined single image	Refined single image	Multi-angle images	Refined single image	Multi-angle images
Personal photos can be applied	Available	Available	N/A	N/A	Available	Available
Classification	Color	Color	Shape:34 classes, Pattern:1 class	Color, Composition	Color, Composition, Light, Space, Line	Color, Form:57 classes, Material:33 classes, Pattern:53 classes
Background exclusion	N/A	Available	N/A	N/A	N/A	Avilable

from the viewpoint of intuitive exploration and providing systematic and fundamental information about the observed appearance. The purpose of the analysis of the artifacts in this study is not to provide information for archaeological and historical judgment but to intuitively analyze the appearance information of the artifact from the user's point of view and provide systematic and fundamental information about the appearance. For this, we established data about the visual elements of artifacts, which have not been sufficiently investigated to date. The data may be used for further studies on the details of artistic expression and provide a basis for the materials that can be employed by the present artists and designers for their creative works.

II. RELATED WORKS

In order to make it easier to obtain the information of artifacts, it has been actively investigated to recognize objects intuitively with smartphone cameras and provide the necessary information through analysis and search by using machine learning technology [9]. Representative examples of the technology include "Google Lens" (https://lens.google.com). These services employ visual search technology to recognize objects from a photographic image, just as the human eyes can intuitively identify objects, and provide relevant shopping information [10], [11]. In addition, although not using a smartphone camera, similar analysis and search services have been developed to identify the characteristics of query photographic images and search similar images by using visual search technologies [12]–[14].

As shown in Table 1, there are several attempts on incorporating machine learning and computer vision technology to provide interesting and useful services on cultural heritage with image search technology in the field of art and culture. Art traders, such as "Artimage" and "Saatchi Art", as well as the "Art Palette" project by Google (https://artsexperiments.withgoogle.com/artpalette/) employ visual search technology to intuitively find information about artworks that have colors and compositions similar to the artwork selected by the users. The "Beyond Scrolls & Screens" project (https://experiments.withgoogle.com/beyondscrolls) by Google provides an art appreciation method where the objects in Japanese paintings are classified according to hierarchical standards and Japanese paintings in a similar category may be viewed. Since most of the currently available artifact data are structured according the classification system based on the periods and archaeological features, the services suggest a novel method of appreciating artworks, widen the human understanding of the artworks, and increase the satisfaction of viewing them. However, few attempts have been made to analyze and interpret the appearance characteristics of artifacts by using a detailed system.

The Vissi's API-based (https://www.visii.com/) "Artimage" and the "Saatchi Art" search artworks having similar colors and compositions and provide the relevant information. The "Art Palette" and "Runway Palette" projects by Google (https://experiments.withgoogle.com/business-of-fashion) extract the representative colors from the color palettes and search the images having similar colors. These projects use only the color information among the visual elements of images. Other elements are also used in other services to analyze objects. For example, the "Beyond Scrolls & Screens" project by Google classifies articles that often appear in Japanese paintings into 35 kinds to search for artworks that have similar objects. However, the project classifies objects according to large categories, and the detailed forms of the individual objects are not specifically analyzed. "ArtPI" by Artendex (https://www.artpi.co/) analyzes more detailed sculptural components based on not only the colors but also the light, space, line, and composition, providing a more diverse visual search function.

The Smart Culture Lens developed in this study employs visual search technology to analyze the on-site pictures of artifacts taken by using a smartphone to provide sculptural information of the artifacts and has a function to search the established ceramic database based on the visual elements. Unlike the conventional services, the Smart Culture Lens includes a multi-angle image database of artifacts (see Table 1). It deals with various visual elements of ceramic appearance such as the color, form, and material, and the service has been developed as a mobile phone app so that users can easily and intuitively use the function. The service separates an artifact from the background in the photo of the artifact taken by a camera. The color, form, and material of the separated artifact are analyzed, and the relevant information is provided. The pattern of the artifact, which is a visual element that gives aesthetic value to ceramics, is also analyzed to provide the relevant information.

III. SMART CULTURE LENS

In this study, an automated image analysis system (named Smart Culture Lens) that can provide information about the appearance of artifacts intuitively and systematically when exploring actual artifacts was established by using deep learning technology. Among the various kinds of Korean artifacts, the target artifacts of the Smart Culture Lens are ceramics, because ceramics are commonly excavated relics in archaeology [15] and there are a number of ceramics currently displayed in museums. It was developed as a smartphone app so that common users can easily use it. A user may take a picture of a ceramic by using the Smart Culture Lens app, which automatically analyzes the characteristics of its appearance. The app then searches the ceramics that have the same appearance characteristics from the ceramics database server built based on museums throughout the Republic of Korea and lets the user view the detailed information. Figure 1 shows the key elements in the development of the Smart Culture Lens. First, the pictures of the ceramics were collected from many museums in the Republic of Korea to systemize the visual elements that are necessary to classify the appearances of ceramics. In this study, the form, color, material, and pattern were determined as the key visual elements to design the detailed classification structure. Based on the classification structure, annotation was performed to tag the meta-information with the collected pictures of the ceramics, establishing the ceramic visual element database. A deep learning technology suitable for image segmentation and classification was employed for the model to learn the automated analysis of the visual elements from the pictures of the ceramics. The model was coupled with the Smart Culture Lens mobile app so that the visual elements can be searched through simple photo-taking.

A. VISUAL ELEMENTS CLASSIFICATION SYSTEM

The visual elements that are commonly used in the field of visual arts and visual analysis include color, form (shape), and material (texture) [16]–[19]. In order to systemize the visual elements of Korean ceramics, the three basic visual element categories were included, and the category of pattern, which gives a symbolical meaning to ceramics, was added [20], [21]. Therefore, the Korea ceramic visual element classification system defined in this study includes the highest-level categories of color (as seen in relic), form (depending on function), material (depending on surface color and production technique), and pattern (depending on symbolic meaning). For color, a classification system was not designed but the representative color was extracted by using an image analysis method. For the other three visual elements, detailed classification systems were designed on the basis of the pictures and a number of studies of the ceramics.

The form of a ceramic is generally determined by the use (function). As presented in Figure 2(a), the highest level category of form was divided into six first-level categories: bottle, jar, bowl, dish, cup, lidded bowl, and ware. Each of the first-level categories was further divided into second-level categories by collecting the ceramics having the same use and form, as classified by various ceramic studies [22], [23], and sorting the ceramics according to the characteristic differences of the appearance. For example, the bottle form ceramics were classified by the 10 basic body forms of JangGyeong-Byeong, MaeByeong, OkHoChoonByeong, TongHyung-Byeong, PyoHyungByeong, SaKakByeong, YouByeong, JaRaByeong, JeongByeong, and JangKoon. According to the additional features applied to the body, PyunByeong, DaGakByeong, and KwaHyungByeong were added. KwangGooHyungByeong was also added for the ceramics having a largemouth. In this way, the first-level categories were further divided into second-level categories.

As a result, the highest category for the ceramic form was divided into a total of 57 second-level categories. For each second-level category, the representative image and the ceramic name used in Republic of Korea were provided with a statement about the characteristics of the appearance.

One of the most commonly used criteria in the ceramic classification system is material. The reason is that the ceramic material used varies depending on the periods, so the material is used as a classification criterion. The highest-level category of material was divided into five first-level categories of earthenware, celadon, buncheong, white porcelain, and others, depending on the manufacturing process, the used glaze, and the baking methods [24]–[26]. The second-level categories were defined as the detailed materials determined by the additional techniques (inlaid, sgraffito, glazed, underglaze, neriage, paste-on-paste, brushstroke, splashing, and openwork) applied to the materials of the individual first-level categories. As presented in Figure 2(b), the category of material was divided into five first-level categories and 31 second-level categories.

The patterns of ceramics (see Figure 2(c)), which increase the aesthetic values, are often derived from common plants, animals, and other natural elements [20], [21]. Research articles about ceramic patterns and the actual ceramic patterns were analyzed to prepare the pattern classification system. The first-level categories of the pattern were the ontological categories such as plant, animal, human, landscape, objects, letters, geometry, and technique [27]–[30]. The first-level categories of plant, animal, and technique were further divided, depending on the details, into second-level categories, which were then further divided into 53 third-level categories for the individual species. However, an actual ceramic may include a single pattern or a plurality of patterns. Therefore, the Smart Culture Lens was designed to analyze

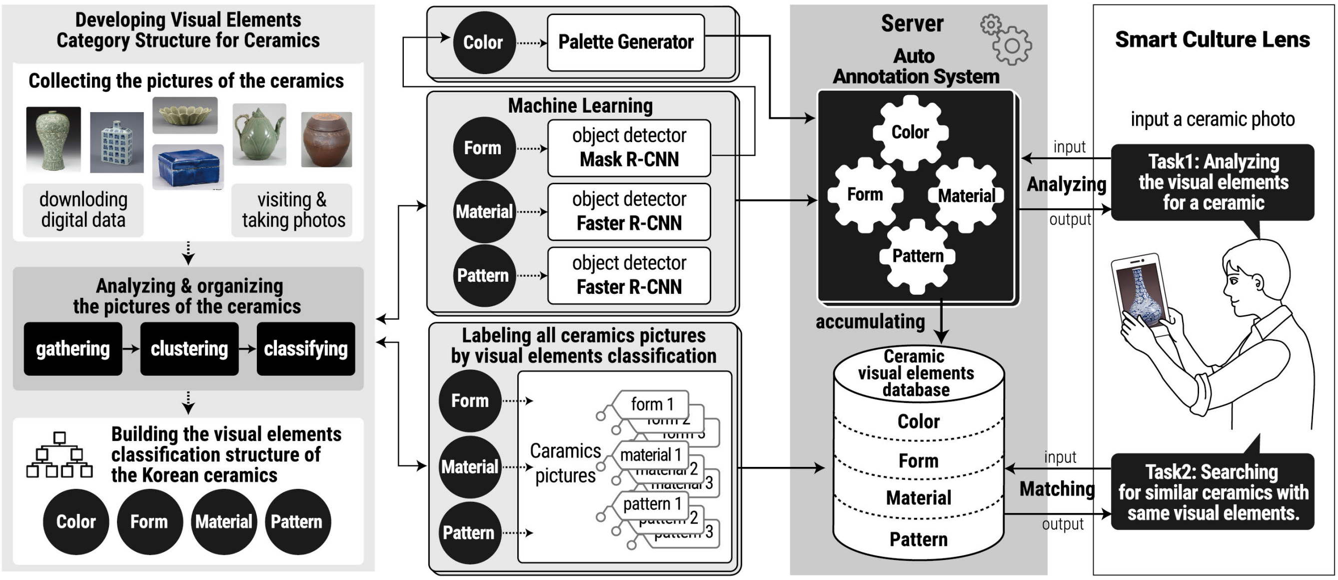


FIGURE 1. **Developmental overview of the smart culture lens.**

only the representative pattern found on the picture taken by users.

Unlike the categories of form, material, and pattern, an explicit classification system was not designed for the color of ceramics but an image analysis was applied. The accurate ceramic color usually refers to the color of the ceramic material, which can be known only by taking the ceramic out of the cabinet and analyzing the material by using a material analyzer. Because that is impossible for the users, the ceramic color analysis in this paper is based on the picture of the ceramic taken by users. However, the color in the picture taken by a user is highly dependent upon the brand of her smartphone, the exhibit environment, and the lighting condition. Different colors may be identified from a single ceramic, but in this paper, the background of a ceramic picture was removed, and five representative colors were automatically extracted from the ceramic part, from the brightest one to the darkest one, and expressed in digital color codes. This method enables common users to analyze color by using their own smartphones.

Figure 3 shows an overview of the visual elements of a ceramic analyzed by the Smart Culture Lens, including two steps for the form, two steps for the material, and two or three steps for the pattern as well as the five representative colors extracted from the ceramic image.

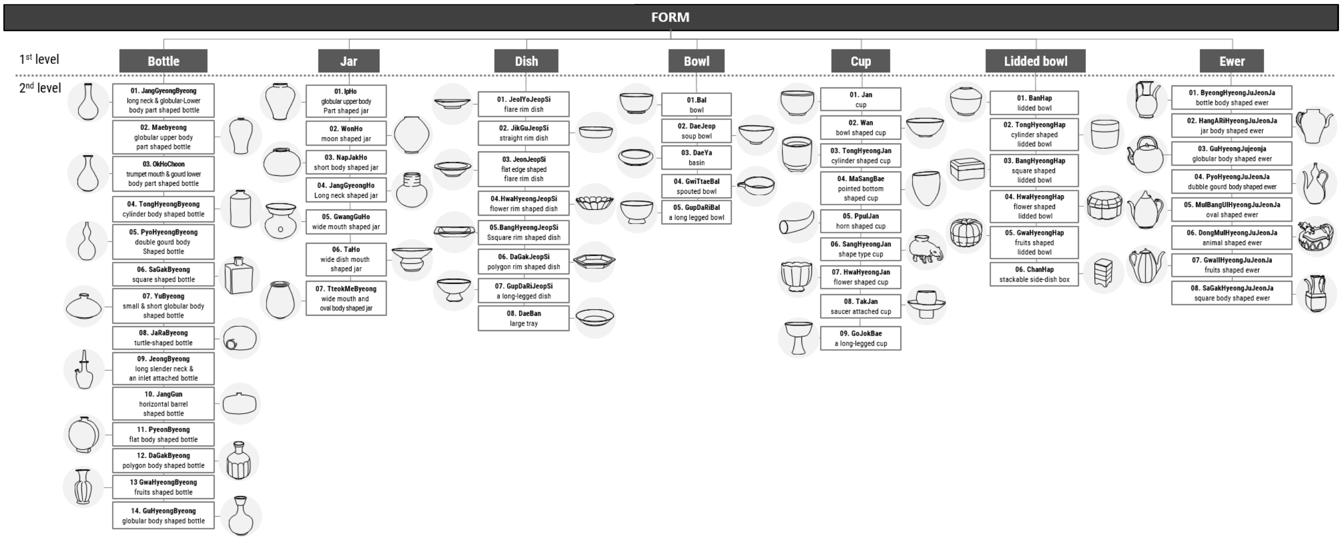
B. CERAMIC DATASET

The dataset needed for the automated visual element analysis and ceramic search in this study was established by two approaches. First, a number of the ceramic images provided by eMuseum (https://emuseum.go.kr/), the website of the National Museum of Korea that provides artifacts information of the museums in the entire Republic of Korea, were downloaded. A total of 5,684 images of 5,211 ceramics were collected, excluding the black-and-white images, scanned images of print-out, the images of incomplete broken ceramics, and the images out of the analytical categories of the periods and nationalities. The eMuseum basically provides a single front image of each ceramic in a well-arranged environment, giving only the basic information of the ceramic. In contrast to the well-refined images provided by the eMuseum, pictures taken by users with their smartphones are not typified because the pictures are affected by various shooting angles and distances as well as various lighting circumstances of museums.

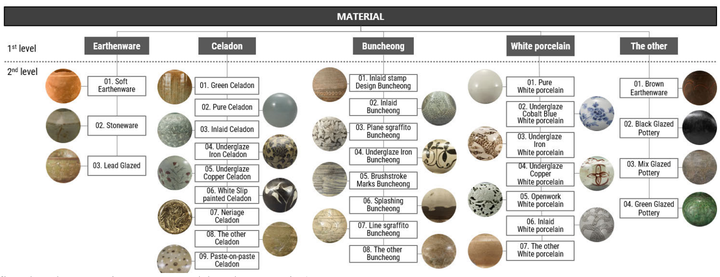
Therefore, for the automated analysis of ceramic pictures taken by using smartphones, on-site pictures of the ceramics were taken at the museum with smartphones from various angles and distances to establish the ceramic database (including pictures and description/information). Pictures of 2,135 ceramics displayed at 19 institutions in the Republic of Korea, including the National Museum of Korea, which is the largest museum in the Republic of Korea, were taken, and 13,926 pictures were finally selected, excluding blurred and noisy images.

An average of 6.5 pictures were taken from each ceramic. The established database was examined by the experts to label all the pictures and classify them into the first and second-level categories of the form, the first and second-level categories of the material, and the first, second, and third-level categories of the pattern. A single ceramic may have several patterns but the labeling was performed with reference to the most outstanding pattern found in the picture.

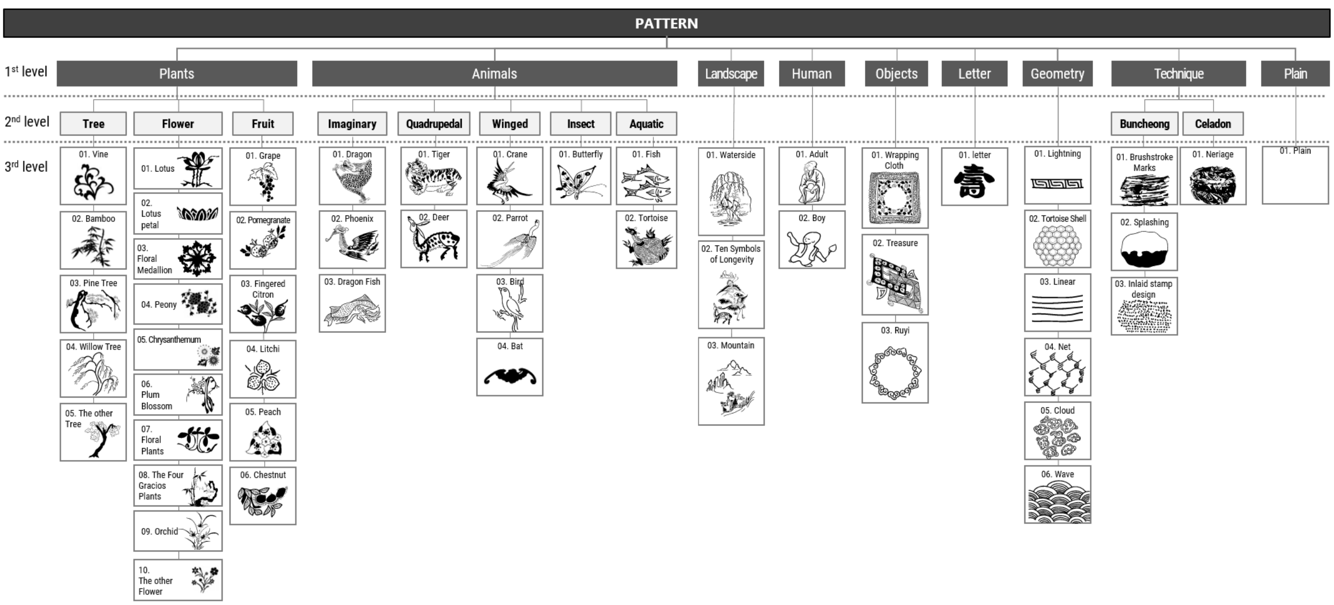
The color information was acquired by extracting the representative colors of the ceramics by using a palette generator. The palette generator reads a ceramic image based on the RGB colors, clusters the color of each pixel, and then extracts the median value as the representative color of each



(a) The Korean ceramics form categories (7 first-level categories, 57 second-level categories).



(b) The Korean ceramics material categories (5 first-level categories, 31 second-level categories).



(c) The Korean ceramics pattern categories
(9 first-level categories, 10 second-level categories, 53 third-level categories).

FIGURE 2. The classification scheme of Korean ceramics visual elements.

cluster. The "Art Palette" by Google extracts five pieces of color information from an artifact, including the background color. In this study, a mask was prepared by using a single front image of each ceramic and the pixel annotation tool (https://github.com/abreheret/PixelAnnotationTool). The mask was used to extract five colors from each ceramic, excluding the background, to express the representative colors of the ceramic more precisely.

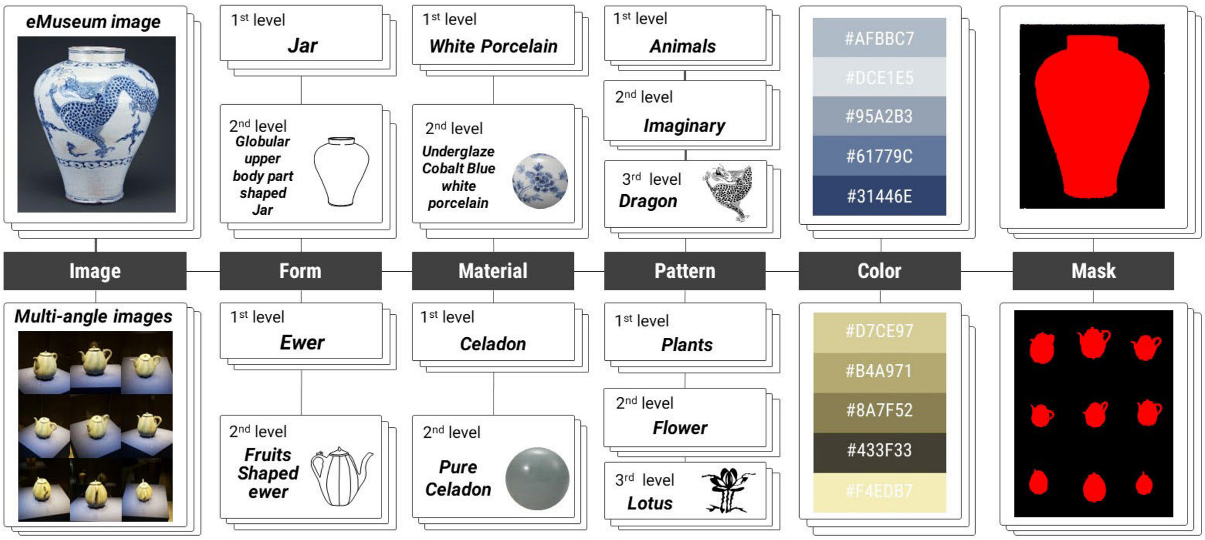


FIGURE 3. **Classification structure and meta-labels used in the smart culture lens.**

As shown in Figure 3, the database consists of the ceramic picture and the annotated visual elements (form, material, pattern, and color) as well as the mask for extracting the colors only from the ceramic, and the bounding box information defining the maximum length and breadth and the position of the mask. Besides, the database incorporates other information collected from the museums, including the name in Chinese characters, period, size, notation, collection institution, and collection number.

C. EXPLORING MODES OF SMART CULTURE LENS

Based on the established ceramic database, the Smart Culture Lens was developed as a smartphone app so that users may search the ceramic information intuitively and simply according to their preferences. The app provides two search modes. In the mode ''Analyzing Visual Elements'' where a user can take a picture of a ceramic to analyze the visual elements, the picture is transmitted to the server where the deep learning model performs automated and real-time search of the ceramic information and the analytical results are sent back to the smartphone, as shown in Figure 4. In the mode ''Searching for Ceramic by Visual Elements,'' the ceramics having the visual elements selected by a user are searched by a directory-based search method.

1) ANALYSIS

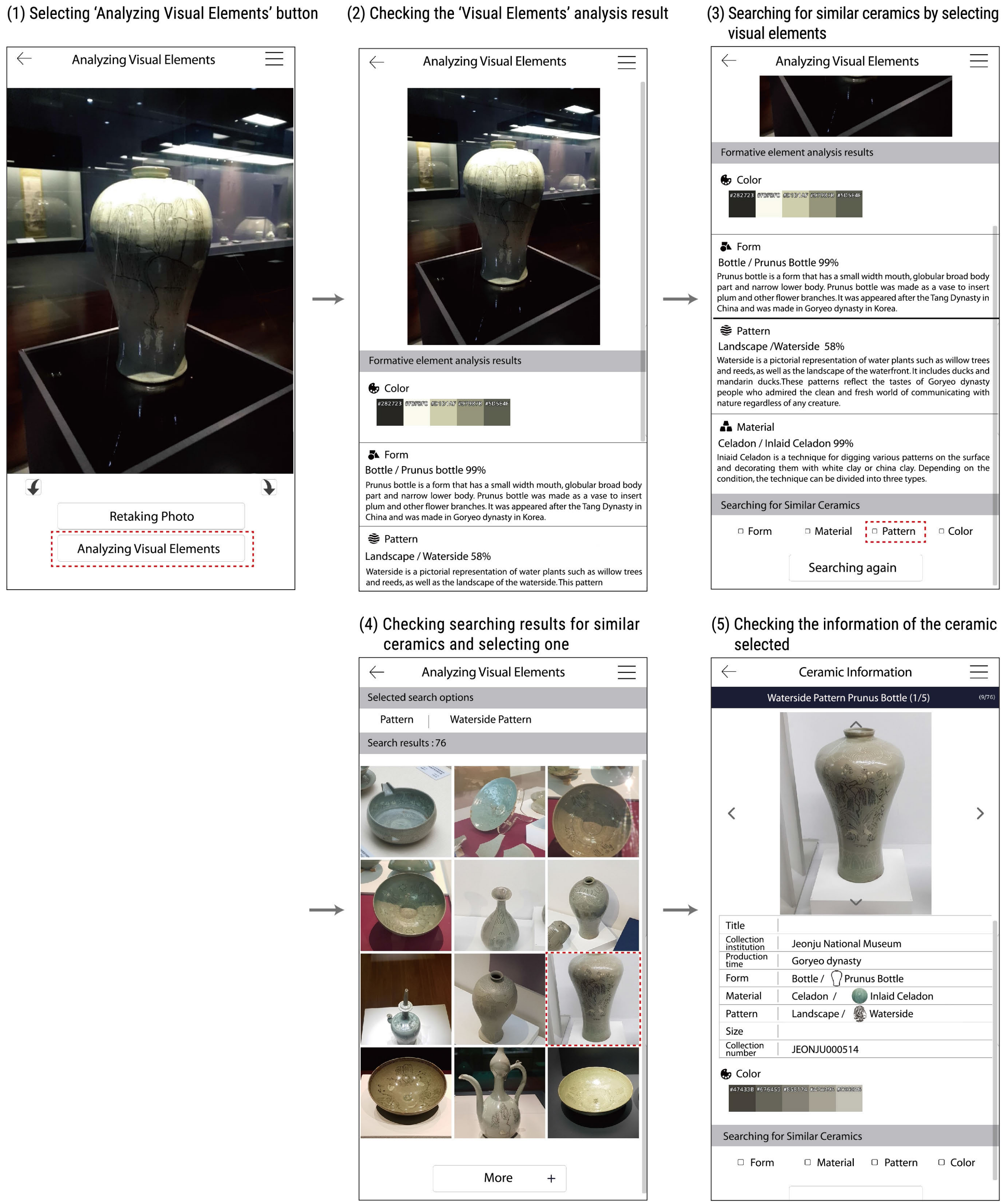
As shown in Figure 4(a1), for 'Analyzing Visual Elements', the input image (picture taken by the user or one stored in the gallery) is transmitted to the server, which then operates the deep learning model and the palette generator to analyze the form, material, pattern, and color. The categories of the visual elements and the description for each of the categories are provided, as shown in Figure 4(a2). When the user selects one or more of the analyzed visual elements as the search keywords (Figure 4(a3)), the ceramics having the categories of the selected visual elements are searched from the database and displayed (Figure 4(a4)). Especially for color, the similarity of the five representative colors is measured, and the ceramics having the colors are sorted and displayed in the order of the highest similarity. When the user clicks one of the pictures of the search ceramics, the detailed information is shown, as in Figure 4(a5).

In this study, a Mask R-CNN model was used to classify the ceramic form and generate the mask of the ceramic, with which only the colors of the ceramic are analyzed as shown in Figure 5(a). Mask R-CNN is a deep learning model that can simultaneously perform the classification through the Faster R-CNN structure, the bounding-box regression, and the instance segmentation of each region of interest (RoI) through the added parallel structure of the fully convolutional network (FCN) [31], [32]. Faster R-CNN does not perform pixel-to-pixel alignment between network inputs/outputs while Mask R-CNN employs the RoIAlign technique for more precise segmentation. For the learning of classification, bounding-box regression, and mask regression, the following multi-task loss is defined:

$$L = L_{cls} + L_{box} + L_{mask} \tag{1}$$

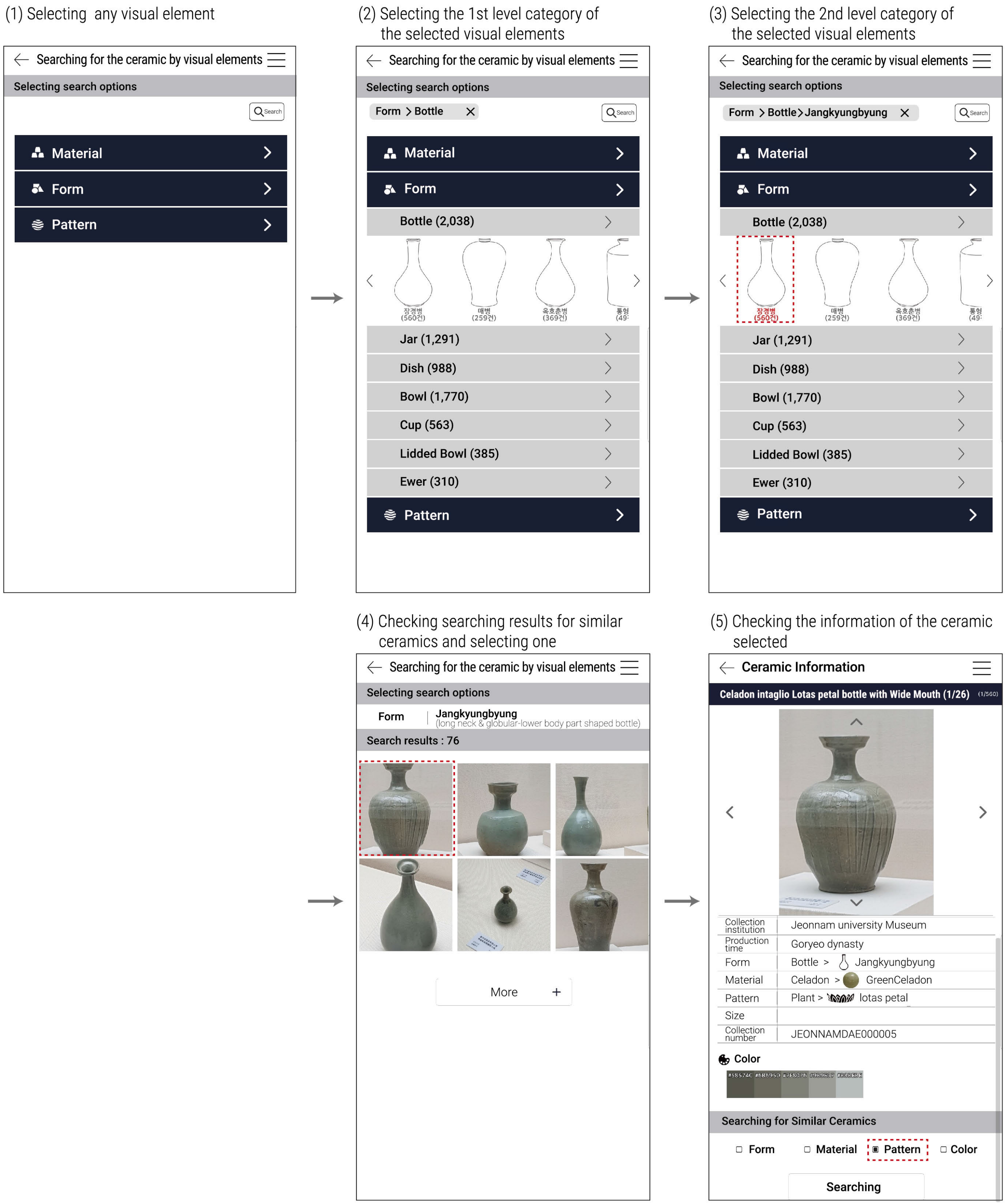
In this study, the Mask R-CNN model was fine-tuned for the 57 low-level categories from the pre-trained inception v2 network. After extracting the part of the ceramic only from the original picture, the palette generator extracts the respective mean values of the five clusters over the colors of the ceramic as 8-bit RGB values as shown in Figure 5(b). While the Google Art Palette randomly displays the five colors, the Smart Culture Lens sorts the colors in the order of the highest color dominance in the ceramic.

For the analysis of material and pattern, the Faster R-CNN was applied, which outputs the object proposal and objectness score of an image in the form of a bounding box repression



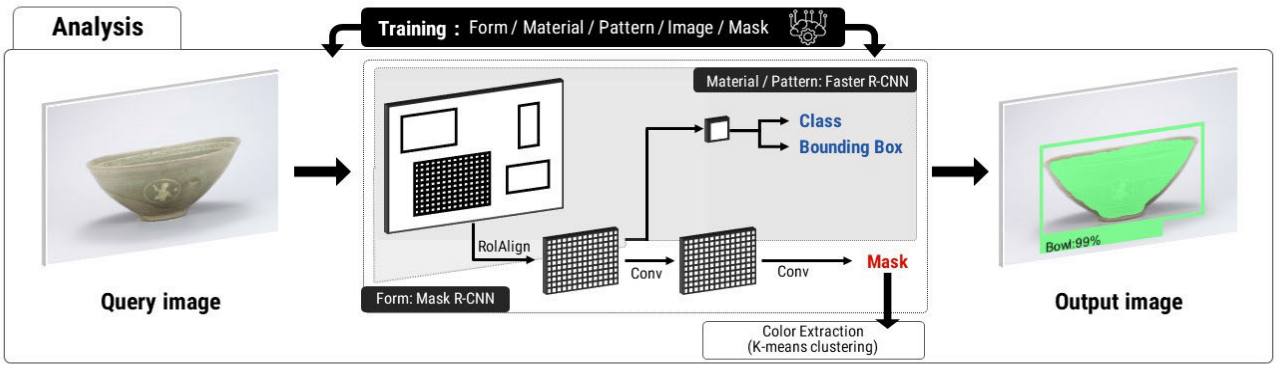
(a) An example of the mode "Analyzing visual elements"

FIGURE 4. Working examples of the smart culture lens.

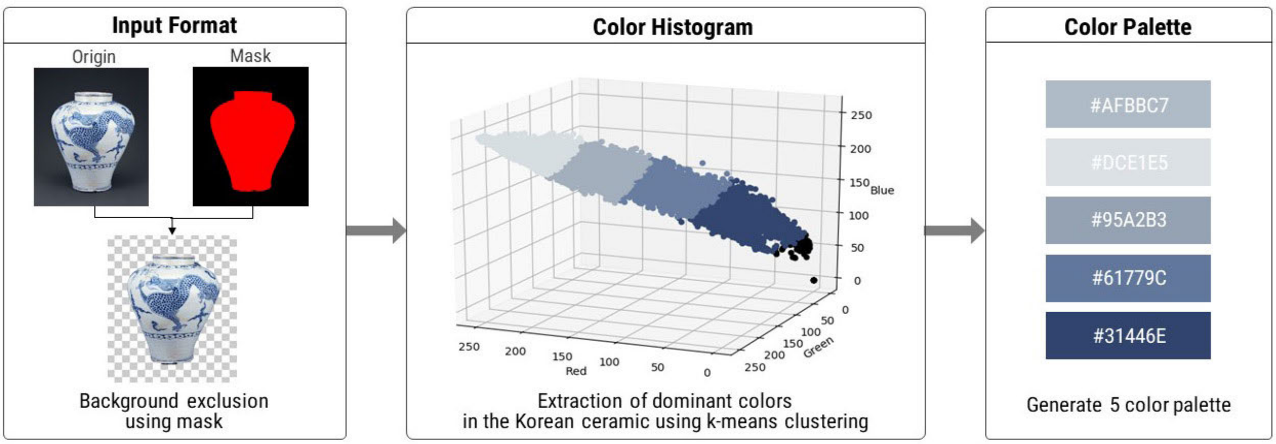


(b) An example of the mode "Searching for the ceramic by visual elements"

FIGURE 4. ***(Continued.)*** **Working examples of the smart culture lens.**



(a) Image recognition models



(b) Color palette generator

FIGURE 5. **Analysis methods for visual elements of ceramics.**

by using the region proposal networks (RPN) [32], [33]. The loss function is defined as in Equation (2):

$$L(\{p_i\},\{t_i\}) = \frac{1}{N_{cls}}\sum_i L_{cls}(p_i, p_i^*) + \lambda\frac{1}{N_{reg}}\sum_i p_i^* L_{reg}(t_i, t_i^*) \quad (2)$$

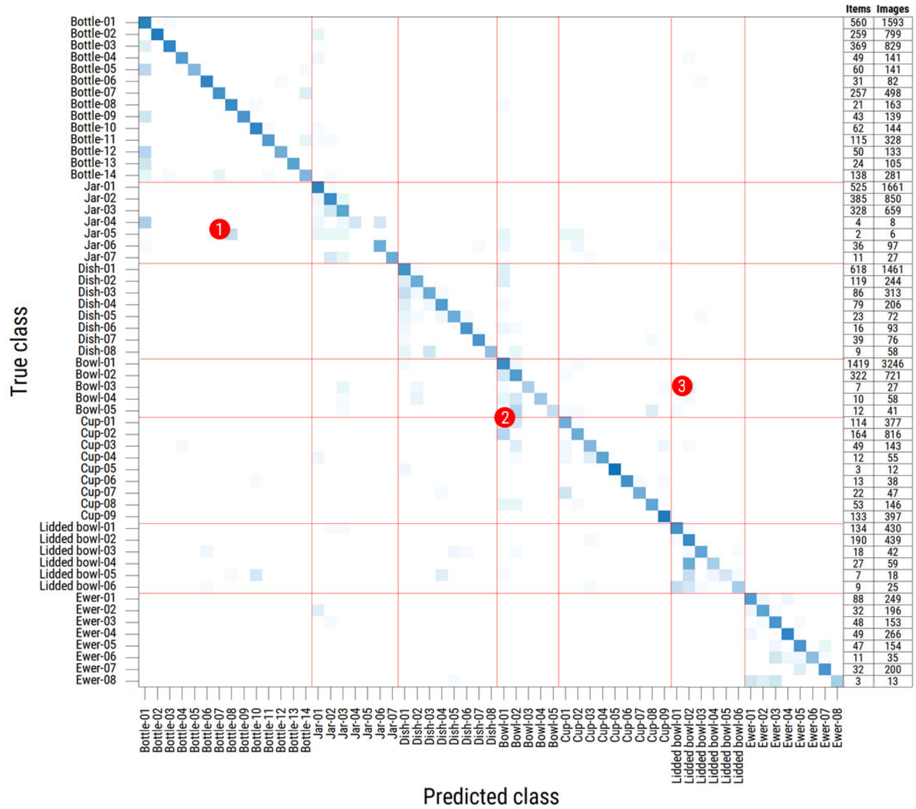
where p_i is the predicted probability of anchor i; in the presence of an object, p_i is 1; otherwise, 0. Equation (2), $t_i = \{t_x, t_y, t_\omega, t_h\}_i$, representing the correlation between the ground truth box and the positive anchor. The classification L_{cls}, representing the classification performance, was the log loss, and the box regression, L_{reg}, representing the object proposal performance, was the smooth $L1$ loss. The Faster R-CNN in this study basically consists of the Inception v2 and performs the image analysis rapidly at a rate of 58 ms per image with reference to the COCO dataset [34]. The material and pattern models were prepared by fine-tuning the pre-trained model according to the number of lowest-level categories (33 for material and 55 for pattern).

2) SEARCHING

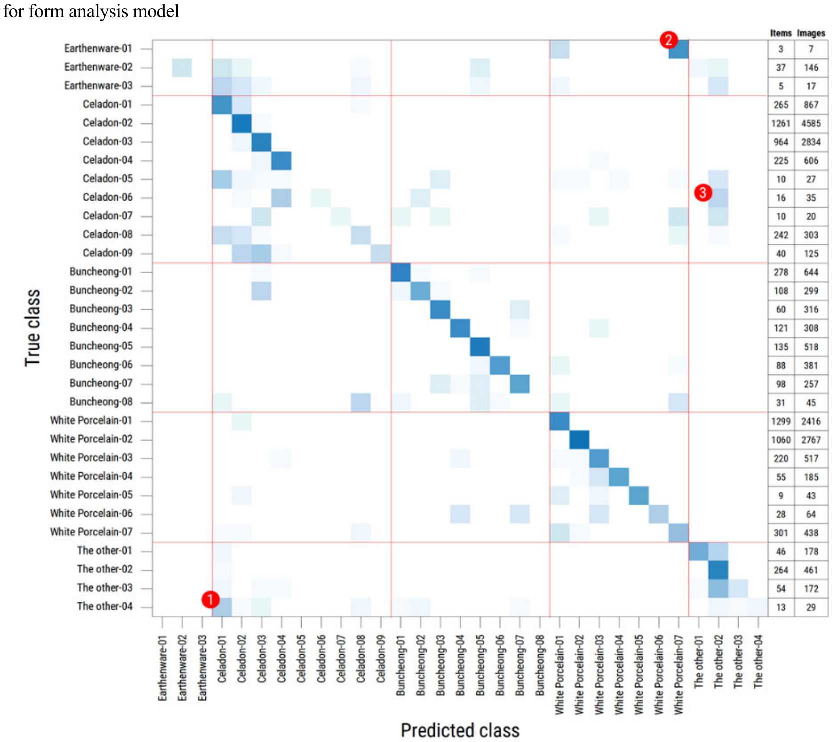
The second mode of the app is to explore the ceramics in a directory search method according to the three visual elements (form, material, and pattern), except the color, as shown in Figure 4(b). For the intuitive search by users, the representative image icons and names are displayed together. The users can also know how many ceramics of a certain category are registered to the ceramic database. They can sequentially select the individual visual elements from the first-level category to the third-level category in the classification system. As shown in Figure 4(b4), the ceramics included in the detailed categories selected by a user are searched from the database and displayed. On the screen showing the search results, the user can select a specific ceramic to view the detailed information about the ceramic (see Figure 4(b5)) and choose the visual elements shown at the bottom to further search the similar ceramics as well.

TABLE 2. **Performance analysis by models.**

Model	Form		Material		Pattern		
Category level	1st	2nd	1st	2nd	1st	2nd	3rd
Average precision (AP)	0.87	0.72	0.89	0.68	0.72	0.58	0.5
Average recall (AR)	0.86	0.62	0.75	0.5	0.58	0.5	0.38
F1 score	0.86	0.67	0.82	0.59	0.65	0.54	0.44
Unweighted accuracy (UA)	0.87	0.73	0.92	0.8	0.82	0.77	0.71
Weighted accuracy (WA)	0.86	0.62	0.75	0.5	0.58	0.5	0.38

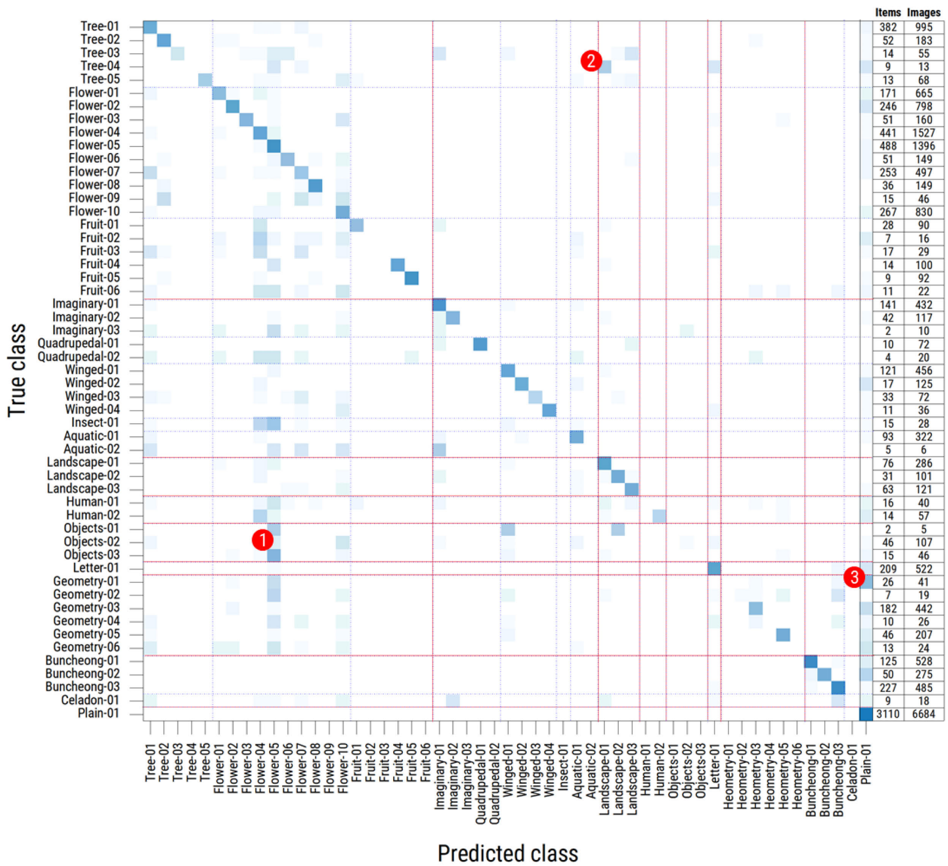


(a) The confusion matrix for form analysis model



(b) The confusion matrix for material analysis model

FIGURE 6. **Confusion matrices for the models learned with reference to the lowest categories. The red solid line represents the boundary of the first-level categories, the blue dash line represents the boundary of the second-level categories, and the red color numbers represent some incorrect cases. Two numbers ("Items" and Images") on the right represent the number of ceramics and corresponding pictures, respectively.**



(c) The confusion matrix for pattern analysis model

FIGURE 6. *(Continued.)* **Confusion matrices for the models learned with reference to the lowest categories. The red solid line represents the boundary of the first-level categories, the blue dash line represents the boundary of the second-level categories, and the red color numbers represent some incorrect cases. Two numbers ("Items" and Images") on the right represent the number of ceramics and corresponding pictures, respectively.**

IV. EXPERIMENTAL RESULTS

A. AUTOMATED VISUAL ELEMENT ANALYSIS

We first evaluated the performance of three deep learning models that classify the form, material, and pattern of an input ceramic image. A 10-fold cross-validation was performed with reference to the lowest categories. The data of 19,610 images were randomly allocated for training, validation, and testing at a ratio of 8:1:1. Based on the errors in each epoch recorded in each fold, the optimal epoch (epoch#: 39) was found based on the accuracy of the validation dataset. The accuracy about the highest categories was indirectly calculated from a confusion matrix on the lowest categories. Table 2 and Figure 6 show results on the classification with respect to form, material, and pattern. Models with higher (fewer) categories naturally show a higher accuracy than those considering lower (more) categories. Due to class imbalance, there are some differences between unweighted and weighted accuracies.

The ceramic dataset established in this study is an imbalanced dataset where the number of images included in the individual classes varies from less than 10 to thousands. Such an imbalanced dataset is evaluated by the F1 score, which is the weighted harmonic average of the precision and the recall calculated by Equation (3):

$$F1 = \frac{Precision \times Recall}{Precision + Recall} \times 2 \tag{3}$$

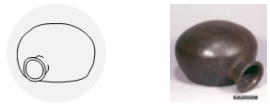
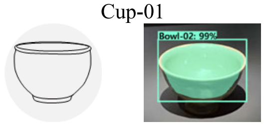
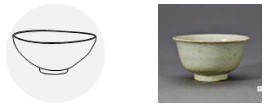
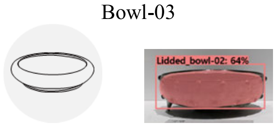
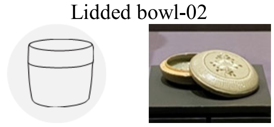
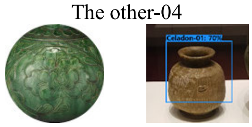
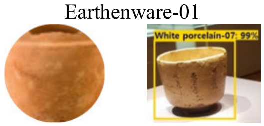
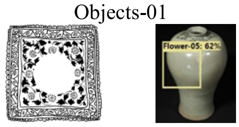
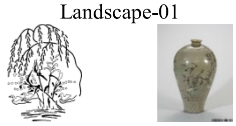
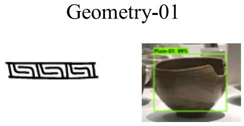
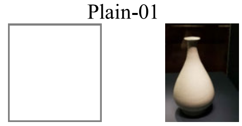
AP and AR were calculated for the F1 score. The form model showed the highest F1 score, indicating that the analytical performance of the form model was higher than that of the material and pattern models.

The number of classes included in the pattern model was similar to the form model. Therefore, UA, one of the main measurement criteria for balanced datasets calculated by Equation (4), was similar to the form model. However, the F1 score of the pattern model was lower than that of the form model, because the data imbalance was higher due to the lower AP and AR values (the data included in the Plain-01 class accounted for even 34% of the entire data.).

$$UA = \frac{1}{N_{tot}} \sum_{c} TP_c \tag{4}$$

Compared with the pattern model, the material model showed a relatively high WA, one of the main measurement

TABLE 3. Representative incorrectly-classified cases analysis.

Model	No.	True class / Incorrectly-classified case	Predicted class / Sample image	Description
Form	1	Jar-05 	Bottle-08 	Abnormal placement of the object
	2	Cup-01 	Bowl-02 	Cup-01 and Bowl-02 are similar in shape, but different in size.
	3	Bowl-03 	Lidded bowl-02 	The small and imbalanced number of data between Bowl-03 (7 items / 27 images) and Lidded bowl-02 (190 items / 439 images)
Material	1	The other-04 	Celadon-01 	A certain light condition distorts the color of the object.
	2	Celadon-06 	The other-02 	The small and imbalanced number of data between Celadon-06 (16 items / 35 images) and The other-02 (264 items / 461 images), and between Earthenware-01 (3 items / 7 images), and White porcelain-07 (301 items, 438 images)
	3	Earthenware-01 	White porcelain-07 	
Pattern	1	Objects-01 	Flower-05 	The small and imbalanced number of data between Objects-01 (2 items / 5 images) and Flower-05 (488 items / 1396 images)
	2	Three-04 	Landscape-01 	'Landscape' means a collection of patterns such 'Tree' and 'Animal', but AI sometimes recognizes individual patterns only.
	3	Geometry-01 	Plain-01 	Limited resolution and light condition

criteria for imbalanced datasets calculated by Equation (5), although the data imbalance of the material model was also high (the data included in the Celadon-01 class accounted for even 23% of the entire data). This was because the total number of the classes was as small as 33.

$$WA = \frac{1}{N_{class}} \sum_{c} (\frac{TP_c}{N_c}) \tag{5}$$

For the further analysis of the classification result, we present the confusion matrix for the three models (form, material, and pattern). As shown in Figure 6, they could achieve good performance by 1st categories of form and material, but show more errors for the classification by lower categories. For pattern, however, it shows poor performance by 1st category due to the imbalance and complexity of data (see Figure 6(c), Flower and Plain categories occupy 32%, 34% of the total data). Some categories include more than hundreds, but some other categories have a few samples, e.g., Jar-05, Earthenware-01, and Objects-01. This imbalance over classes in the data is commonly appeared from the three models, leading to decreased accuracies. Further studies need to reduce such errors and increase the accuracy, and

TABLE 4. Survey items and evaluation results.

(*n*=57) (*Cra*=.93)

Factors		Items	SD	MEAN	
Focused on Experiences (Subjective/Qualitative)	meaningful	Smart Culture Lens App has cultural and social value.	0.63	4.32	4.31
		It is meaningful to analyze ceramics as visual elements from a visual design point of view, not from a historical point of view.	0.66	4.09	
		The smart culture lens app should be extended to analyze other cultural artifacts as well as ceramics analysis.	0.57	4.44	
		This is a technique that made possible work from impossible work in the past.	0.66	4.23	
		These products, which combine culture and AI technology, must continue to be developed.	0.54	4.49	
	pleasurable	The overall flow is well connected.	0.63	4.00	4.00
		This shows the information I want to know from ceramics	0.89	3.77	
		The information presented as a result of each function is interesting.	0.71	4.12	
		This is worth recommending to others.	0.68	4.04	
		I have fun when using this.	0.85	4.00	
		I want to continue exploring with smart cultural lens app.	0.74	4.05	
	convenient	It conveys simple, organized and clear information.	0.80	4.12	4.24
		It requires only the minimum necessary steps when looking up information.	0.65	4.39	
		It easily conveys information in a visual way.	0.78	4.25	
		The information shown is easy to understand.	0.62	4.21	
		Terms and expressions used in menus and functions are easy to understand.	0.77	4.21	
Focused on task (Objective/Quantifiable)	usable	This can be used without difficulty.	0.66	4.49	4.34
		This can be used without looking at the written instruction.	0.75	4.28	
		Everyone will learn how to use this app very quickly.	0.60	4.56	
		The layout of the information and content of the app is useful.	0.74	4.09	
		Each function is useful.	0.70	4.26	
	reliable	When I do the task what I want, I can do it exactly.	0.65	4.07	4.11
		Overall, the design elements used are reliable.	0.69	4.09	
		Overall, the functional elements used are reliable.	0.79	4.02	
		Overall, the results of "Analyzing Visual Elements" are reliable.	0.68	4.23	
		Overall, the results of "Exploring Visual Elements" are reliable.	0.68	4.16	
	functional	Each function worked as I expected.	0.70	4.26	4.21
		The function of "Analyzing Visual Elements" is sufficient.	0.63	4.25	
		The function of " Exploring Visual Elements" is sufficient.	0.76	4.23	
		Information and layout are appropriate.	0.77	3.95	
		It's fast until you get the result.	0.69	4.35	
		The pages are designed intuitively.	0.69	4.25	

data augmentation, zero-shot or few-shot learning technology could be a good option [35], [36]. In practical situations, some other causes make classification poorer, e.g., form or color distortion in the picture due to different camera angles, the placement of an artifact, and various lighting conditions in the environments. Pre-processing or post-processing to resolve the distortion could improve the performance of the analysis. Several examples of incorrectly classified cases marked in the confusion matrixes are presented in Table 3.

B. USER EXPERIENCE EVALUATION

A survey was conducted with 57 undergraduate and graduate students to examine whether the Smart Culture Lens reflects the original purposes and to assess the user satisfaction with the app. Various real ceramic artifacts were displayed, and the participants were asked to use the Smart Culture Lens freely to get the information of the ceramic artifacts (see Figure 7). The survey items used in the evaluation of the usability were those included in 'creating pleasurable interfaces' model developed by Anderson(2009). Anderson's model incorporates Peter Morville's Honeycomb model(2004) and Maslow's Hierarchy of Needs theory(1954). The low level of this model comprises the factors focused on tasks, such as usable, reliable, and functional, and the high level comprises the factors focused on experience, such as meaningful, pleasurable, and convenient [37]–[40]. The participants were

FIGURE 7. Smart Culture Lens app usability evaluation.

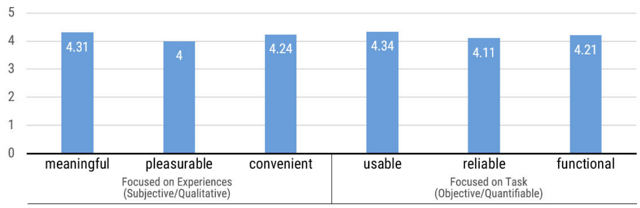


FIGURE 8. A graph comparing the results of the six evaluation factors.

asked to respond on a 5-point Likert scale (1-Strongly disagree, 2-Disagree, 3-Neutral, 4-Agree, 5-Strongly agree).

As shown in Figure 8 and Table 4, the overall scores for 6 measures were four points or higher. With regard to the factors focused on experience, relatively higher scores were obtained from the meaningful factor, including the survey items such as ''These products, which combine culture and AI technology, must continue to be developed (4.49, SD = 0.54),'' ''The smart culture lens app should be extended to analyze other cultural artifacts as well as ceramics analysis (4.44, SD = 0.57).'' In the convenient factor, subjects expressed easiness and intuitiveness in their experience of the Smart Culture Lens, such as ''It requires only the minimum necessary steps when looking up information (4.39, SD = 0.65)'' and ''It easily conveys information in a visual way (4.25, SD = 0.78).'' Only for one question of the pleasurable factor out of 15 experience-related questions, they scored lower than four, which is ''This shows the information I want to know from ceramic (3.77, SD =0.89).'' The reason for the low score of pleasurable factor was that users wanted to obtain additional ceramic information such as the date of manufacture as well as visual elements.

And there were user comments on the focused on experience such as ''It's good to see that science and technology are used to analyze the cultural heritage,'' ''This is exactly what I needed when visiting a museum,'' ''I could intuitively understand artifacts without the background knowledge about the history,'' and ''I could learn more about the artifacts.''.

With regard to the factors focused on tasks, the highest score was found in the usable factor from the survey items ''This can be used without difficulty (4.5, SD = 0.68),'' and ''Everyone will learn how to use this app very quickly (4.58, SD = 0.61).'' These results show that the users can easily view the automatic analytical results without complicated steps, thanks to the application of the AI technologies to the app and that the users are satisfied with the performance including the processing speed and functions. The survey results clearly showed that the app can analyze the detailed visual elements of the artifacts by using AI technologies and can be easily and intuitively used by any user.

In addition, many participants expressed their impression on the Smart Culture Lens, such as ''It's good to see that science and technology are used to analyze the cultural heritage,'' ''This is exactly what I needed when visiting a museum,'' ''I could intuitively understand artifacts without the background knowledge about the history,'' and ''I could learn more about the artifacts.'' It signifies that the app provides information that helps the intuitive understanding during the experience of exploring real artifacts at museums. In addition to viewing the artifacts and obtaining the historical facts from the museums, the visitors can extend their experiences in the museums by acquiring specific visual information about artifacts as well as searching visually similar artifacts by using the Smart Culture Lens.

V. CONCLUSION

Recent advances in computer vision and machine learning technology enable to acquisition of information on an object by simply taking its picture with people's own smartphones. In this paper, we developed the Smart Culture Lens to support an easy and intuitive way to explore the cultural heritage, through the entire process of converging the AI technology into the analysis of photographic images of artifacts. In order to get some information on a cultural artifact, users are simply required to take a picture of artifacts in a natural environment.

The Smart Culture Lens app is one product from the process but many other studies on humanities and design can be conducted by the same process. If the method employed in this study is applied to museums, they can not only provide the typical information about the artifacts but also discover artifacts information services from a novel point of view. Although the main purpose of the Smart Culture Lens was to provide intuitive exploration information, interesting research topics were obtained from the huge amount of data prepared in this study and the classification of the visual elements. What are the visual elements that were most frequently used in each period? What is the people's emotional aspect connected with each of the visual elements? What are the visual elements that were used in each region? What are the visual elements used in the neighboring countries? Objective and quantitative answers to these questions may be provided by analyzing the accumulated data by various methods.

There are some limitations to our work. First, in order to analyze the entire ceramics, e.g., Chinese or Japanese ceramics, we need to expand the classification scheme further. We have designed a flexible classification scheme, so we will do this with data augmentation in future studies. Second, the number of data per category is somewhat small and imbalanced. As this adversely affects analytical performance, it will be solved through state-of-the-art technologies such as data augmentation and few-shot learning later. Finally, our system is influenced by light conditions of the environment, where data augmentation through light simulation could be a solution.

REFERENCES

[1] P. Kroes, "Design methodology and the nature of technical artefacts," *Des. Stud.*, vol. 23, no. 3, pp. 287–302, May 2002, doi: 10.1016/S0142-694X(01)00039-4.
[2] M. A. Rosenman and J. S. Gero, "Purpose and function in design: From the socio-cultural to the techno-physical," *Des. Stud.*, vol. 19, no. 2, pp. 161–186, Apr. 1998, doi: 10.1016/S0142-694X(97)00033-1.
[3] B. Chandrasekaran and J. R. Josephson, "Function in device representation," *Eng. with Comput.*, vol. 16, nos. 3–4, pp. 162–177, Dec. 2000, doi: 10.1007/s003660070003.
[4] N. Crilly, "The roles that artefacts play: Technical, social and aesthetic functions," *Des. Stud.*, vol. 31, no. 4, pp. 311–344, Jul. 2010, doi: 10.1016/j.destud.2010.04.002.
[5] N. F. Roozenburg and J. Eekels, *Product Design: Fundamentals and Methods*, 1st ed. Hoboken, NJ, USA: Wiley, 1995.
[6] C. G. Boucharenc, "Research on basic design education: An international survey," *Int. J. Technol. Des. Educ.*, vol. 16, no. 1, pp. 1–30, Jan. 2006, doi: 10.1007/s10798-005-2110-8.
[7] O. Gude, "Postmodern principles: In search of a 21st century art education," *Art Educ.*, vol. 57, no. 1, pp. 6–14, Jan. 2004, doi: 10.1080/00043125.2004.11653528.
[8] C. Wallschlaeger, C. Busic-Snyder, and M. Morgan, *Basic Visual Concepts and Principles for Artists, Architects, and Designers*. Dubuque, IA, USA: Wm. C. Brown Publishers, 1992.
[9] Y.-G. Jiang, M. Li, X. Wang, W. Liu, and X.-S. Hua, "DeepProduct: Mobile product search with portable deep features," *ACM Trans. Multimedia Comput., Commun., Appl.*, vol. 14, no. 2, pp. 1–18, May 2018, doi: 10.1145/3184745.
[10] F. Yang, A. Kale, Y. Bubnov, L. Stein, Q. Wang, H. Kiapour, and R. Piramuthu, "Visual search at eBay," in *Proc. 23rd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2017, pp. 2101–2110, doi: 10.1145/3097983.3098162.
[11] Y. Zhang, P. Pan, Y. Zheng, K. Zhao, Y. Zhang, X. Ren, and R. Jin, "Visual search at Alibaba," in *Proc. 24th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Jul. 2018, pp. 993–1001, doi: 10.1145/3219819.3219820.
[12] S. Bell and K. Bala, "Learning visual similarity for product design with convolutional neural networks," *ACM Trans. Graph.*, vol. 34, no. 4, pp. 1–10, Jul. 2015, doi: 10.1145/2766959.
[13] W. Zhou, H. Li, and Q. Tian, "Recent advance in content-based image retrieval: A literature survey," 2017, *arXiv:1706.06064*. [Online]. Available: http://arxiv.org/abs/1706.06064
[14] J. Xiang, N. Zhang, R. Pan, and W. Gao, "Fabric image retrieval system using hierarchical search based on deep convolutional neural network," *IEEE Access*, vol. 7, pp. 35405–35417, 2019, doi: 10.1109/ACCESS.2019.2898906.
[15] E. Roman-Rangel, D. Jimenez-Badillo, and S. Marchand-Maillet, "Classification and retrieval of archaeological potsherds using histograms of spherical orientations," *J. Comput. Cultural Heritage*, vol. 9, no. 3, pp. 1–23, Nov. 2016, doi: 10.1145/2948069.
[16] P. Aigrain, H. Zhang, and D. Petkovic, "Content-based representation and retrieval of visual media: A state-of-the-art review," *Multimedia Tools Appl.*, vol. 3, no. 3, pp. 179–202, Nov. 1996, doi: 10.1007/BF00393937.
[17] S. Banerji, A. Sinha, and C. Liu, "New image descriptors based on color, texture, shape, and wavelets for object and scene image classification," *Neurocomputing*, vol. 117, pp. 173–185, Oct. 2013, doi: 10.1016/j.neucom.2013.02.014.
[18] C. W. Niblack, R. Barber, W. Equitz, M. D. Flickner, E. H. Glasman, D. Petkovic, P. Yanker, C. Faloutsos, and G. Taubin, "QBIC project: Querying images by content, using color, texture, and shape," *Proc. SPIE*, vol. 1908, pp. 173–187, Apr. 1993, doi: 10.1117/12.143648.
[19] R. C. Veltkamp and M. Hagedoorn, "State of the art in shape matching," in *Principles of Visual Information Retrieval*. London, U.K.: Springer, 2001, pp. 87–119, doi: 10.1007/978-1-4471-3702-3_4.
[20] Y. J. Im, *Korean Traditional Patterns*, 8th ed. Seoul, South Korea: Daewonsa Publishing Co., 2004.
[21] Y.-I. Yun, Y. J. Lee, and H. Sim, *Korean Traditional Pattern2-Ceramics*. Daejeon, South Korea: Nat. Res. Inst. Cultural Heritage, 2008.
[22] Y. J. Kim, K. J. Park, B. S. Bang, H. J. Seo, S. H. Eum, S. J. Lee, J. M. Lee, S. C. Jeon, and J. T. Choi, *Korean Ceramic Dictionary*, 1st ed. Seoul, South Korea: Kyungin Publishing, 2015.
[23] Y. W. Kim, S. K. Im, H. J. Yun, and H. B. Jeon, *Terminology of Korean Archaeology*. Daejeon, South Korea: Nat. Res. Inst. Cultural Heritage, 2012.
[24] D. G. Kang, U. R. Kim, H. J. Kim, H. S. Hwang, S. H. Kwon, W. S. Jang, and J. A. Im, *In Blue and White: Porcelains of the Joseon Dynasty*. Seoul, South Korea: Nat. Museum of Korea, 2014.
[25] K. S. Kang, *Korean Ceramics*, 2nd ed. Seoul, South Korea: Yekyong Publishing Co., 2012.
[26] J. I. Lee, *Mt.Mudeung Buncheong Wares: The First Epoch of Joseon Dynasty Ceramics*. Gwangju, South Korea: Gwangju Nat. Museum, 2013.
[27] S. Carey, *Conceptual Change in Childhood*. Cambridge, MA, USA: MIT Press, 1985.
[28] M. T. H. Chi, J. D. Slotta, and N. De Leeuw, "From things to processes: A theory of conceptual change for learning science concepts," *Learn. Instruct.*, vol. 4, no. 1, pp. 27–43, 1994, doi: 10.1016/0959-4752(94)90017-5.
[29] F. Keil, *The MIT Press Series in Learning, Development, and Conceptual Change*. Cambridge, MA, USA: MIT Press, 1989.
[30] E. J. Wisniewski, "Construal and similarity in conceptual combination," *J. Memory Lang.*, vol. 35, no. 3, pp. 434–453, Jun. 1996, doi: 10.1006/jmla.1996.0024.
[31] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," in *Proc. IEEE Int. Conf. Comput. Vis.*, Oct. 2017, pp. 2961–2969. [Online]. Available: https://openaccess.thecvf.com/content_iccv_2017/html/He_Mask_R-CNN_ICCV_2017_paper.html
[32] L. Jiao and J. Zhao, "A survey on the new generation of deep learning in image processing," *IEEE Access*, vol. 7, pp. 172231–172263, 2019, doi: 10.1109/ACCESS.2019.2956508.
[33] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137–1149, Jun. 2017, doi: 10.1109/TPAMI.2016.2577031.
[34] A. Sharma, V. Singh, and A. Rani, "Implementation of CNN on Zynq based FPGA for real-time object detection," in *Proc. 10th Int. Conf. Comput., Commun. Netw. Technol. (ICCCNT)*, Jul. 2019, pp. 1–7, doi: 10.1109/ICCCNT45670.2019.8944792.
[35] Y. Wang, Q. Yao, J. Kwok, and L. M. Ni, "Generalizing from a few examples: A survey on few-shot learning," 2019, *arXiv:1904.05046*. [Online]. Available: http://arxiv.org/abs/1904.05046
[36] Y. Xian, B. Schiele, and Z. Akata, "Zero-shot learning—The good, the bad and the ugly," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 4582–4591. [Online]. Available: https://openaccess.thecvf.com/content_cvpr_2017/html/Xian_Zero-Shot_Learning_-_CVPR_2017_paper.html
[37] J.-H. Ahn and S.-I. Kim, "Comparison the difference of user experience for mobile Facebook and Instagram using nonparametric statistics methods–focused on emotional interface Model," *J. Digit. Converg.*, vol. 14, no. 11, pp. 481–488, Nov. 2016, doi: 10.14400/JDC.2016.14.11.481.
[38] S. P. Anderson. (2006). *Creating pleasurable interfaces: Getting from tasks to experiences*. [Online]. Available: https://www.slideshare.net/stephenpa/creating-pleasurable-interfaces-getting-from-tasks-to-experiences
[39] A. H. Maslow, *Motivation and Personality*, 1st ed. New York, NY, USA: Harper & Row Publishers, 1954.
[40] P. Morville and P. Sullenger, "Ambient findability: Libraries, serials, and the Internet of Things," *Serials Librarian*, vol. 58, nos. 1–4, pp. 33–38, Apr. 2010, doi: 10.1080/03615261003622999.

JI HYUN YI received the B.F.A. degree in occidental painting from Ewha Woman's University, the M.F.A. degree in fine arts from the University of Pennsylvania, the M.I.D. degree from the Industrial Design of University of the Arts, and the Ph.D. degree from the Culture and Technology of Korea Advanced Institute of Science and Technology (KAIST). She is currently an Assistant Professor with the School of Integrated Technology, Gwangju Institute of Science and Technology (GIST). Her research interests include culture data classification structure for AI technology applications, interactive content design, UI/UX design, AR/MR interaction, and creative image creation method.

WOOJIN KANG received the B.E. degree in bioelectronics from Yonsei University, in 2014, and the M.S. degree in robotics from the Daegu Gyeongbuk Institute of Science and Technology (DGIST), in 2016. He is currently pursuing the Ph.D. degree with the School of Integrated Technology, Gwangju Institute of Science and Technology (GIST). In 2017, he worked as a Researcher with DGIST. His research interests include computer vision and human–computer interaction.

SONG-EI KIM received the bachelor's degree in art and the master's degrees in aesthetics and art history from Chosun University. From 2016 to 2018, she worked as a Registrar (Cultural heritage management) with the Gwangju National Museum, South Korea. Since 2018, she has been working as a Researcher with the Gwangju Institute of Science and Technology (GIST). Her research interests include content research on museums, cultural heritage and art, and design of content classification systems.

DOYUN PARK received the B.F.A. degree in digital design from Daegu Catholic University, in 2018. He is currently a Researcher with the School of Integrated Technology, Gwangju Institute of Science and Technology (GIST).

JIN-HYUK HONG received the B.S., M.S., and Ph.D. degrees in computer science from Yonsei University, Seoul, South Korea. He is currently an Assistant Professor with the School of Integrated Technology and the AI Graduate School, Gwangju Institute of Science and Technology (GIST). From 2009 to 2014, he was a Postdoctoral Researcher and a Systems Scientist with the Human–Computer Interaction Institute, Carnegie Mellon University. From 2014 to 2018, he worked with Naver and Samsung Electronics, South Korea, as a Research Engineer. His research interests include context awareness, pattern recognition, and user interaction design, particularly focusing on the understanding of human behaviors.