Journal of Network Intelligence
Taiwan Ubiquitous Information
 ISSN 2414-8105
Volume **9**, Number **4**, November **2024**

Extraction of Sustainable Environmental Elements and Application in Product Design Based on Visual Analysis Techniques

Xiaoran Duan*, Hui Ren

College of New Media
Sichuan Film and Television University, Chengdu 610036, China
Theqfish@163.com, renhui2407@126.com

*Corresponding author: Xiaoran Duan

Received November 25, 2023, revised February 28, 2024, accepted May 26, 2024.

Abstract. *The quality of the industrial milieu plays a pivotal role, exerting a direct influence on the trajectory of development across various spheres. As an integral component of our cultural heritage, sustainable environmental aesthetics are systematically distilled from pottery styles, deftly preserved through the medium of digital imagery, or meticulously crafted models. The existing methodology for element extraction is constrained by its capacity to segregate only singular objects, thus revealing a deficiency in precise cataloging and evolving tracking. To address this limitation and enhance the overall standard of the industrial environment, this study employs visual analysis and fusion management technology, with a specific emphasis on harnessing image processing and convolutional neural networks for element extraction.Initially, binocular vision analysis technology is employed to acquire a differential image, followed by meticulous pixel screening. Subsequently, an intricate image segmentation model is meticulously constructed, comprising two PCA convolution layers and an output layer. The outcomes from the two convolution filter layers are then normalized. Ultimately, the fusion management technique applies the channel dimension merging operation to each pixel block of the extracted differential image. According to the annotation diagram, the amalgamated image pixels are categorized into change points and constant points, facilitating feature extraction to obtain the desired element segmentation result. The experimental findings underscore the notable efficacy of our method, achieving accuracy rates of 99.01%, 99.43%, 98.3%, and 98.3% for four distinct types of porcelain patterns (geometric, human, animal, and landscape). Post-analysis reveals that the exemplary segmentation effect significantly contributes to the extraction and utilization of visual art elements, thereby demonstrating its practical significance in safeguarding and innovating excellent traditional culture to enhance the industrial environment and elevate the industrial standard.*
Keywords: visual analysis techniques; convolutional neural networks; industrial environment; image segmentation; fusion management

---

1. **Introduction.** In recent years, a heightened societal awareness concerning ecological conservation and sustainable development has emerged. Particularly within the realm of artistic and cultural product design, designers are increasingly acknowledging their inherent responsibilities and actively exploring practices and procedures aimed at mitigating their ecological impact. The extraction of sustainable environmental elements represents a noteworthy trend within the design domain, providing companies with a means to embody sustainability principles in their product design methodologies [1]. By responding to the growing consumer affinity for eco-friendly products and services, this approach aligns

with the heightened performance expectations of discerning individuals. Concurrently, designers are gradually emancipating themselves from the conventional notion that design is solely driven by product function and appearance. They are now contemplating ways in which design can better serve society and the natural environment. Through the extraction and segmentation of environmentally relevant elements, such as product patterns, shapes, colors, and other sustainable environmental attributes, designers strive to achieve a harmonious synthesis of cultural heritage and ecological preservation [2]. In the realm of art and design, the incorporation of sustainable environmental elements signifies not only a commitment to environmental responsibility but also a quest for innovation and aesthetics. The utilization of sustainable environmental elements offers artists and designers a wellspring of fresh creative inspiration and material. Simultaneously, the integration of these elements facilitates the creation of ecologically sound, sustainable, and reusable products. Furthermore, numerous sustainable environmental elements bear rich cultural connotations and historical values. Leveraging these elements in design serves not only to promote traditional culture but also to actualize the inheritance and evolution of cultural heritage.

As the scope of environmental assessment broadens and resources continue to advance, the manual extraction and segmentation of environmental elements have proven to be inefficient and susceptible to classification inaccuracies. However, recent years have witnessed notable advancements in technologies such as computer vision [3], image processing [4], and machine learning [5], which have provided robust support for the extraction of sustainable environmental elements. Leveraging high-performance computing power, abundant image data, and sophisticated algorithms, we can now achieve more precise extraction of environment-related information from images or videos. In the realm of sustainable environmental element extraction, numerous research scholars have conducted studies employing visual analytics. For instance, in literature [6], the environmental impact of various materials is assessed and compared through the application of visual analysis techniques. By scrutinizing the raw material sources, energy consumption, and emissions during the production process, as well as the environmental impact during usage and disposal stages, designers can make informed choices regarding more sustainable materials, thereby reducing the environmental footprint of their products. Furthermore, the literature [7] employs image recognition algorithms to analyze energy usage in different scenarios, facilitating the monitoring and evaluation of product energy consumption patterns.

Drawing from the aforementioned studies, it is apparent that visual analysis techniques facilitate prompt and accurate identification and extraction of content from image or video data [8]. However, while existing research predominantly concentrates on extracting individual elements from water and electricity environmental factors, there exists a scarcity of investigations into the extraction of diverse environmental elements in art product design. Notably, in the extraction of environmental elements from porcelain, a culturally significant category, patterns display complexity and overlapping characteristics. Furthermore, the observed patterns often exhibit variations from different angles. Consequently, achieving precise extraction of sustainable environmental elements from porcelain patterns becomes imperative. Such extraction can aid designers in promptly discerning people's preferences for porcelain patterns and drawing inspiration for artistic creation from the distinctive elements inherent in these patterns. Ultimately, this can facilitate the creation of design solutions and artworks that align more expeditiously with sustainable design and production principles.

To tackle the inherent challenges in extracting comparable porcelain pattern elements, this study employs binocular visual analysis techniques to extract image blocks from porcelain representations at each pixel point. Subsequently, a novel visual detection algorithm is formulated, utilizing convolutional neural network image segmentation. The algorithm takes the neighborhood image block of pixel points as input to adapt to the image detection model. Finally, through the network composed of two PCA convolution layers and one output layer, the algorithm can effectively predict the porcelain pattern under different observation angles.

By astutely utilizing the extracted sustainable environmental elements, the image can be finely segmented into distinct areas, enhancing our understanding and application of sustainable artistic elements such as lines, shapes, and colors. This approach significantly improves the harmony between the pattern and the form of the product. These distinctive pattern elements serve as a wellspring of rich creative inspiration for artists and designers. This comprehensive analysis and application not only broaden the horizons for artistic creation but also offer designers a fount of inspiration for creating products aligned with the principles of sustainable development and humanistic care.

2. **Related Work.** Visual analysis constitutes an analytical approach grounded in the observational capabilities and habits of the human eye [9,10]. Existing research on visual analysis encompasses various facets, including viewpoints, dynamic perspectives, visual interfaces, visual frequencies, and visual aesthetics. It is commonly applied in fields such as art, design, landscape gardening, and geography research. However, as the complexity of observed scenes escalates, relying solely on human visual analysis proves inefficient, resulting in incomplete analyses. To surmount these limitations, the integration of external cameras and computer-based vision techniques has proven invaluable.

Early studies identified that relying solely on a single camera for image capture can lead to image quality issues and subsequent inaccuracies in analyses [11,12]. To address this, literature [13] compared linear model camera calibration methods based on the direct linear transformation method with those based on the perspective transformation matrix method. They introduced the concept of using at least two cameras for image acquisition from the same source, coupled with binocular vision analysis techniques. Additionally, literature [14] utilized the parallax principle of left and right images in binocular vision. This involved incorporating camera models, internal and external parameters of camera lenses, and platform structures to devise a straightforward camera calibration method, meeting the requisites for measuring long-range targets. In 2022, literature [15] harnessed binocular vision technology, leveraging high-resolution GPU capabilities, to achieve real-time tracking and localization of multiple objects within intricate backgrounds. These advancements in employing binocular vision techniques and utilizing multiple cameras have substantially heightened the accuracy and efficiency of visual analysis, paving the way for more comprehensive and reliable analyses across various domains.

Advancements in binocular vision technology have provided essential procedural foundations for studying images within natural environments. To enable computers to comprehend and interpret images, researchers have introduced image segmentation techniques for the further analysis of visual data. These techniques encompass methods such as histogram thresholding [16], feature clustering and region growth [17], or artificial neural networks [18], aimed at distinguishing regions with similar characteristics from those lacking such attributes. With the evolution of computer vision analysis techniques, researchers turned to convolutional neural networks (CNNs) [19], coupled with standard backpropagation algorithms, for image recognition and extraction [20].

Subsequent efforts focused on the application of deep learning for unsupervised training of images, leading to rapid progress across various domains. For instance, in literature [21], classification extraction was accomplished using deep convolutional neural networks

(DCNNs) [22], leveraging the ImageNet LSVRC-2010 dataset. In literature [23], a convolutional neural network named AlexNet, comprising five convolutional layers and three fully connected layers, was fine-tuned to address foreground and background segmentation of characters. Moreover, in the field of computer science, literature [24] integrated hyperpixel analysis with convolutional neural networks to conduct image segmentation on millimeter-wave cloud radar maps for meteorological observations. This approach substantially reduced the number of pixel points requiring classification for foreground and background attributes.

These advancements represent significant progress in harnessing deep learning techniques, notably convolutional neural networks, for image analysis and segmentation. They have played a pivotal role in enhancing the accuracy and efficiency of data interpretation, enabling researchers to derive valuable insights from visual data across diverse domains.

In 2020, literature [25] introduced an accelerated Full Convolution Network (FCN) as an extension of CNNs for pixel-level classification, showcasing substantial improvements in accuracy. FCNs offer greater flexibility and efficiency compared to CNNs [26]. Building upon the FCN, literature [27] enhanced its structure by integrating both convolutional neural network and Reflux neural network [28] into a unified framework, enabling simultaneous training. By incorporating superpixel analysis, literature [29] seamlessly merged a fully convolutional neural network with a convolutional neural network, leading to enhanced algorithm speed and more effective segmentation and extraction outcomes for millimeter-wave cloud radar images.

In the field of art product design, where artworks and designs often possess subjective and complex characteristics, the extraction of sustainable environmental elements necessitates an understanding and interpretation of the artwork's theme, intention, style, and form. Environmental elements within artworks often encompass diverse attributes, such as color, shape, texture, and line, which intertwine and interact, forming multidimensional elements. However, given the intricate and unstructured nature of artworks, hyper-pixel analysis and full convolutional neural networks may prove unsuitable for extracting elements in art product design. To tackle challenges posed by lighting conditions, image quality, and algorithm performance affecting the extraction of color and shape, this paper focuses on image segmentation using superpixel analysis and convolutional neural networks. This approach aims to extract the shape and texture of porcelain patterns from various viewing angles, considering the intricacy and variability inherent in art product designs.

3. Model design.

The image collection process in this study is based on the utilization of binocular vision technology, employing a camera for observation. Figure 1 illustrates the comprehensive process of scene detection. The collected images undergo pre-processing, facilitating the subsequent extraction of environmental element features from the porcelain images using binocular vision analysis techniques. This methodology enables the extraction of sustainable environmental elements in the form of patterns on porcelain within the camera's observation area.

3.1. Image pixel filtering.

The extraction of visual elements relies on the collection of a substantial amount of visual information. The collected porcelain images encompass various abstract and figurative elements, including points, lines, surfaces, colors, and shapes. These images contain extractable sustainable environmental elements, which can be categorized into two groups: human visual elements and natural visual elements. However, for the objectives of this paper's product design, the focus is placed on the extraction of natural visual elements, specifically porcelain pattern shapes and texture

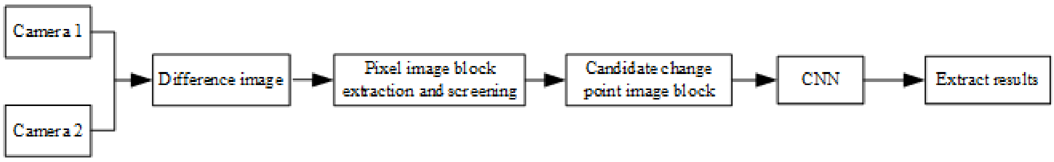

Figure 1. Frame diagram of image segmentation and extraction based on binocular vision technology

elements. To initiate the process, data collection is conducted. Neighboring images of pixel points within the designated model are selected as input to train the image detection model. Furthermore, a coarse filtering step is applied to the pixels of the differential image. This filtering procedure aims to eliminate a majority of the invariant points, thereby enhancing the efficiency of the visual detection algorithm. The implementation flow for extracting sustainable environmental elements is as follows:

(1) The pre-processed image and post-change image undergo image alignment and reduction operations. Afterward, the two images are differentiated by the channel dimension, resulting in a differential image that represents the changes in the infrared and visible light channels.

(2) For the differential image containing the infrared and visible channels, an image block of size $K \times K$ is created around each pixel. If the size of the image block is less than $K \times K$, it is zeroed out.

(3) Simultaneously, the differential image is subjected to a coarse filtering process to remove most of the invariant points. The differential image is summed across the channel dimension, generating a matrix A that represents the total change values for each pixel. This matrix A is then binarized based on the magnitude of the total change values, yielding the following Equation (1):

$$A(x, y) = \begin{cases} 0, & \delta \geq A(x, y) \\ 1, & \delta < A(x, y) \end{cases} \tag{1}$$

Where δ is the change threshold, the pixel point (x, y) corresponds to the matrix $A(x, y)$.

0 is the invariant point, and 1 is the change point. In general, invariant and change points have different data distributions, and the optimal change threshold δ is obtained by solving for the maximum interclass variance.

Assuming a change threshold of $\delta = t$, the number of change points is N_c, the set of change points is V_c, the number of invariant points is N_{uc}, the set of invariant points is V_{uc}, and $V = V_c \cup V_{uc}$, at which point the index of the elements in V can be obtained as follows:

$$V = [v_1, v_2, ..., v_p] \tag{2}$$

The total number of pixels is $N_p = N_c + N_{uc}$, then the inter-class variance between changing and unchanging pixels δ is:

$$\delta(t_i) = \frac{N_c}{N_p} \sum_{i=1}^{N_c} (v_i - \overline{v}_c)^2 + \frac{N_{uc}}{N_p} \sum_{i=1}^{N_{uc}} (v_i - \overline{v}_{uc})^2 \quad (i = 1, 2, ..., N_s) \tag{3}$$

Where $\overline{v}_c$ and $\overline{v}_{uc}$ are the mean values of V_c and V_{uc}, respectively. The pixel points are arranged in descending order of total variation value, and N_s points are sampled in equal proportion, and the total variation value of the sampled pixel points is traversed as the threshold T to solve for the class small variance, and the threshold $\hat{\delta}$ with the largest class small variance is.

$$\hat{\delta} = \arg\max(\delta(t_i))(i = 1, 2, ..., N_s) \tag{4}$$

(4) Finally, to avoid inaccurate detection rates due to broken detection regions or sporadic noise outside the change region, we fill the change region and remove the noise to correct the undetected points. The first fill is performed using the closed operation method, where the structural element B is introduced on the basis of the image matrix A. The expansion operation $\oplus$, the erosion operation $\ominus$, and the closed operation $\cdot$ are performed.

$$A = A \cdot B = (A \oplus B) \ominus B \tag{5}$$

Building upon the previous steps, the following actions are performed: The pixel values of the current pixel point and its neighboring pixels in the detection image are sorted. The value at the middle position in the ranking is selected as the value for the current pixel point. This step utilizes median filtering to eliminate noise in the area of change.

(5) The matrix A, where a value of 1 represents a candidate change point, is utilized to identify the locations of change. These candidate change points serve as representatives of the areas where the change has occurred. Finally, the image block corresponding to each candidate change point is output, completing the image pre-processing phase.

3.2. **CNN-based extraction of sustainable environmental elements.** In the context of image pre-processing, we established a model dedicated to feature extraction, as elucidated in Figure 2.

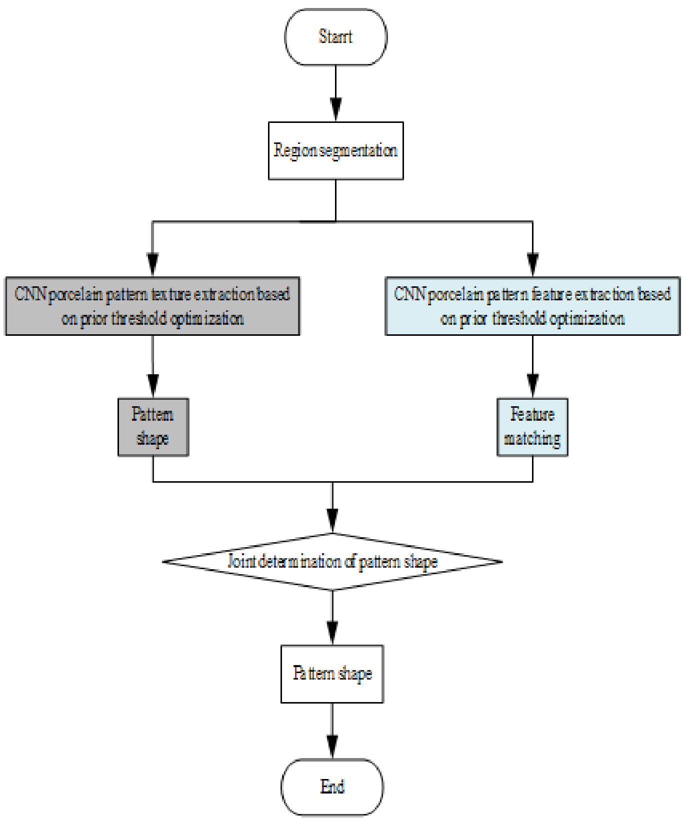

Figure 2. Algorithm flow chart

Initially, the image shall undergo segmentation into discrete regions, and the pixel parallax of the botanical specimen observed through two detection instruments will be quantified via the application of feature matching principles. Subsequently, the inference of the porcelain pattern's configuration will be conducted through the amalgamation of data from the observation devices, hardware parameters of the observation system, and the utilization of the principle of similar triangles. Concurrently, a convolutional neural network, utilizing a priori threshold optimization, will be deployed to extract the texture and characteristics inherent in the porcelain pattern. Ultimately, through a comparative analysis of information acquired from these two components, the segmentation and extraction of the pattern's configuration will be proficiently accomplished.

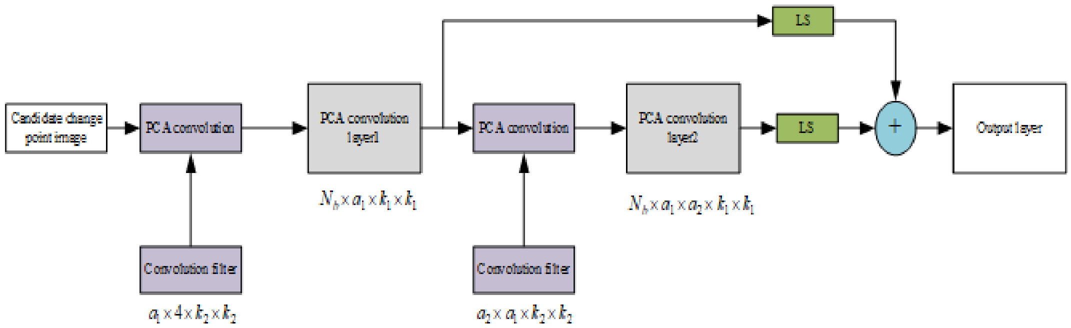

Figure 3. Model network construction

The model network consists of two PCA convolutional layers and one output layer, as shown in Figure 3. In PCA convolution layer 1, a PCA convolution filter of size $4 \times k_2 \times k_2$ is used to convolve the input data to obtain $N_b \times a_1$ a feature map of size $k_1 \times k_1$. Before convolution, the input needs to be zeroed to ensure that the image size is constant before and after convolution. Then, the feature map of PCA convolution layer one is used as the input of PCA convolution layer 2, and the PCA convolution filter of size two and $a_1 \times k_2 \times k_2$ is used to convolve the feature map of PCA convolution layer 1 to obtain a feature map of size and $N_b \times a_1 \times a_2$ $k_1 \times k_1$. Then, the feature maps of PCA convolution layer one and PCA convolution layer two are combined in the a_2 dimension after Layer Normalisation (LN) processing.

Finally, the feature maps of each image block are flattened to one dimension and fed into a support vector machine to determine whether the candidate pixels have changed.

The amalgamation of feature graphs from PCA convolution layers 1 and 2 in dimension is undertaken, facilitating the fusion of distinct levels of feature information. This integration enhances the training stability and convergence speed of the model during LN processing. The algorithm, adept at effectively extracting the shape, texture, and other features inherent in porcelain patterns, employs the PCA convolution layer to concentrate on key image features through reduction and principal component analysis. Subsequently, the support vector machine classifies pixels based on these features, ensuring precise segmentation and successful extraction of pattern elements. This algorithm demonstrates commendable performance and accuracy in visual inspection tasks, particularly for objects exhibiting intricate texture and shape variations, as exemplified by porcelain.

Acknowledging the susceptibility of porcelain pattern shape and texture to varying light angles, coupled with potential segmentation errors arising from color blending, our detection model necessitates a training sample incorporating both pre- and post-change images alongside annotated maps. In the domain of porcelain change detection, solely the regions manifesting alterations are annotated, while the remaining areas are designated as invariant. Consequently, image blocks corresponding to the change points and invariant points are transformed into positive and negative samples, respectively.

In the broader context, the number of change points is notably smaller than that of invariant points in change detection. This engenders a significant imbalance between positive and negative samples, predisposing the model to overfitting and potential oversights in change detection. Furthermore, the data distribution of invariant points is uneven, marked by a scarcity of high-value invariant data.

To address these challenges, we adopt a discerning approach in generating positive and negative samples for each training image within this model. Specifically, for change points with limited data, all samples are utilized as positive samples. In contrast, the invariant points undergo selective filtering to attain a balanced ratio of positive and negative samples. Thus, the process of generating positive and negative samples for each training image within this model unfolds as follows:

(1) Extract the image block of each pixel point of the difference image before and after the change, merge the channel dimension of the difference image, and divide the pixel points of the merged image according to the change point and the unchanged point according to the annotation map, and sort each by the value size from largest to smallest.

(2) Assuming that the error in the manual annotation is 5

(3) The image blocks corresponding to all points in the change point set V_c are designated as positive samples for model training. Pixel points sorted in the invariant point set V_{uc} are then sampled using a second-order equivariance series, ensuring that invariant points are captured within each value region. The general formula for the second-order equivariance series is

$$f_n = 1 \tag{6}$$

where N_{ss} is the number of invariant points to be sampled, f_n is the pixel index of the nth sampled point, and setting $f_n = 1$, then,

$$f_{N_u} = \mathrm{len}(V_{uc}) \tag{7}$$

The equivalence d in Equation (6) at this point is calculated as shown in Equation (8):

$$d = \frac{2(f_{N_u} - N_{ss} f_1)}{N_{ss}(N_{ss} - 1)} \tag{8}$$

After the training samples are obtained, all the positive and negative sample image blocks are input into the model, and the feature vectors of each image block are created by cascading the image blocks in sliding windows of size $k_2 \times k_2$. The optimal standard orthogonal matrices for PCA convolutional Layer 1 and PCA convolutional Layer 2 are computed as the PCA convolution filter parameters for each respective layer. Subsequently, the feature vectors generated by the network model serve as the training samples for the SVM. The SVM is then trained using these feature vectors. Finally, the PCA filter parameters for PCA convolutional Layer 1 and PCA convolutional Layer 2, along with the model parameters of the SVM, are saved.

4. **Experiment and analysis.** The presented methodology prioritizes the extraction of porcelain pattern shapes through the utilization of a CNN optimized by a priori threshold-based techniques. This section evaluates the performance of the proposed scheme concerning image segmentation and subsequently contrasts the accuracy of extracting individual elements of the porcelain pattern shape. Finally, the implications for the design of artistic and creative products will be deliberated based on the outcomes derived from the extraction of elements constituting the porcelain pattern shape.

4.1. **Image segmentation results.** The initial training dataset for our model comprised 1645 images, encapsulating four distinct pattern types: geometric, human, animal, and landscape patterns. Concurrently, a dedicated validation set consisting of 600 images was employed for the fine-tuning of the network. The training process encompassed 5000 iterations, utilizing a learning rate of 0.00001. Notably, the images within the validation set were independent of the training set and remained uninvolved in the training procedure. Given the disparate viewing angles of the two cameras, disparities in images necessitate registration to align the two. To address this requirement in the experiment, the image registration algorithm ICP is employed for the alignment of the two images. Figure 4 illustrates a comparative analysis of the classification performance between our

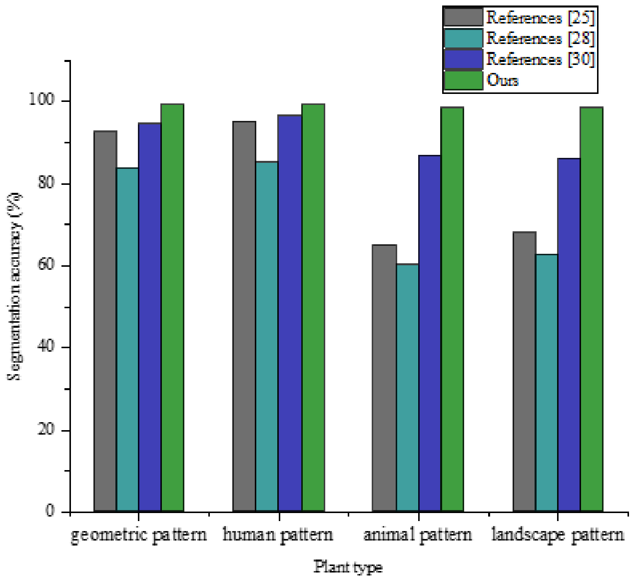

Figure 4. Comparison of segmentation accuracy

proposed method and referenced literature sources. As evident in Figure 4, our model, as presented in this paper, adopts the neighborhood image block of pixel points as input for adaptation to the image detection model. Through the preliminary screening of pixels in the differential image, a substantial portion of invariant points is effectively filtered out, enhancing the characteristics of the pixel itself and its adjacent pixel regions such as color, shape, and gradient.

Ultimately, our method demonstrates a notable proficiency in pattern-type detection, achieving a detection rate exceeding 98.3%. Only two pattern types, specifically "animal pattern" and "landscape pattern," encountered a singular segmentation error, while the segmentation and detection outcomes for the remaining pattern types aligned precisely with the actual patterns.

In [25], the model's detection rate varied between 65% and 95%, with the "character pattern" achieving a detection rate of 95%. [28] reported a detection rate ranging from 60% to 85%, with the lowest accuracy observed for the "animal" and "landscape" pattern types, reaching 60% and 62.59%, respectively. Only in the study presented in [30], a relatively good segmentation performance of 86.79% was achieved for the "animal pattern" and "landscape pattern" types. However, a substantial disparity exists compared to the 98.3% segmentation accuracy rate attained in this paper.

The primary source of this disparity lies in the fact that does not account for the potential impact of lighting conditions on the extraction of color and shape elements. Consequently, the segmentation edges in [30] may be prone to misidentification. In contrast, our method accurately segments the pixel points themselves and their neighboring regions by leveraging color, shape, and gradient features, ensuring precise segmentation. Notably, our method demonstrates exceptional performance in accurately extracting the intricate edges and colors of "animal prints," successfully isolating them from complex background information. The segmentation results exhibit minimal background information while effectively preserving the details of the animal prints.

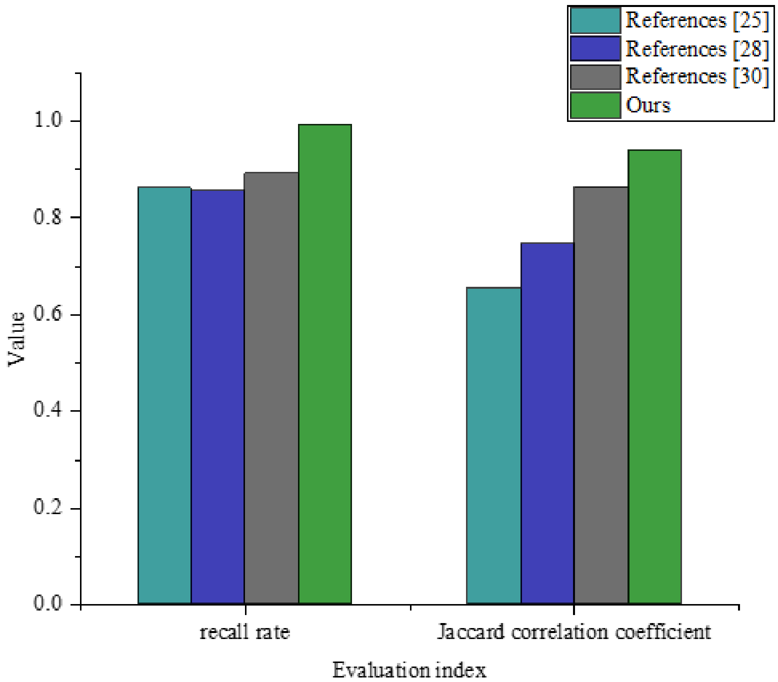

Figure 5. Objective evaluation index comparison

To underscore the advantages of our new model, we employed evaluation metrics such as recall and the Jaccard similarity coefficient to assess the segmentation results, as depicted in Figure 5. As both metrics approach a value of 1, it signifies a higher level of segmentation accuracy. Figure 5 illustrates that the evaluation metrics of the models presented in [25] and [28] yield poor numerical performance, while the metrics in [30] demonstrate better results due to the inclusion of superpixel analysis.

In this paper, our model incorporates superpixel analysis while introducing binocular visual analysis to mitigate the impact of different lighting conditions and observation angles on false segmentation. Consequently, our model achieves superior numerical performance. When compared to other models, the average values of recall and Jaccard's similarity coefficient on the test images are at least 1.31% and 2.68% higher, respectively. Furthermore, after conducting an analysis, the accuracy and recall variation curves presented in Figure 6 were plotted. The model attains a recall of 99.1% at a segmentation accuracy of 98.4%, aligning with the theoretical analysis and visually intuitive results achieved by our model in this paper.

The above experimental results validate the potential impact on artists' design thinking and methods under the methodology proposed in this paper. Firstly, traditional art

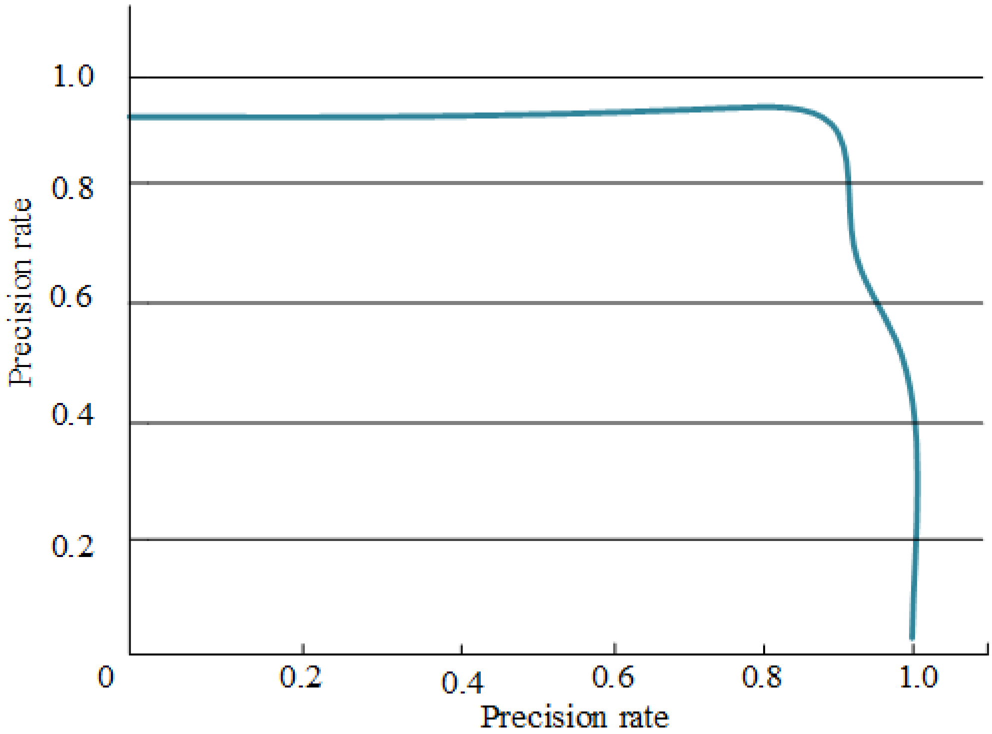

Figure 6. The curve of precision and recall rate

creation and design often rely on the artist's intuition and experience, whereas the image elements extracted by binocular vision analysis technology necessitate artists to focus more on the analysis and understanding of images. This shift enables artists to objectively and accurately grasp the composition and beauty of the image.

Secondly, traditional artistic creation and design typically commence from a holistic perspective, achieving the final effect through continuous adjustment and refinement. However, the image elements extracted by binocular vision analysis require artists to pay greater attention to the processing of the parts and details of the image. This adjustment allows artists to delve deeper into the inner beauty and expressiveness of images, albeit demanding heightened attention to detail and the coordination and balance between the part and the whole.

Moreover, traditional art creation and design are often solitary endeavors by a single artist, whereas the image elements extracted by binocular vision analysis technology provide additional opportunities and possibilities for collaboration and communication among artists. This transformation fosters artists' openness and inclusivity to different perspectives and styles, thereby promoting innovation and development in artistic styles and techniques. Simultaneously, this necessitates artists to adapt to a new collaborative working approach and workflow, fully harnessing the strengths and creativity of the team.

4.2. **Ablation experiment.** In this section, ablation experiment is conducted to validate the effectiveness of the modules incorporated into the algorithm proposed in this paper for the extraction of sustainable environmental elements. Unlike the approach adopted in [30], which combines FCN and CNN using Superpixel Analysis (SPA), our method utilizes two PCA convolutional layers and an output layer to construct a CNN. This configuration provides the network with auxiliary supervisory information and refines the segmentation results.

By comparing the proposed improved SPA1+CNN model with the SPA+FCN+CNN model, it is evident that our method outperforms all evaluation metrics, as depicted in Figure 7. In this paper, the SPA model is an enhancement of the superpixel analysis module. It involves filtering the neighboring image blocks of the pixel points and employing the filtered blocks as inputs to train the image detection model. This approach proves significantly more effective than the original SPA module.

Additionally, the CNN model in this paper takes into account the influence of irrelevant elements in complex environments. To prevent the detection target and the background from being merged into a single block, the target change region is kept adjacent to the pseudo-change region. This design demonstrates the algorithm's ability to distinguish between actual change points and pseudo-change points. In contrast, the FCN+CNN model may mistakenly detect nearby pseudo-change points, leading to a decrease in detection accuracy.

Overall, the algorithm presented in this paper effectively detects target changes while eliminating pseudo-change points, resulting in improved accuracy when compared to the SPA+FCN+CNN model.

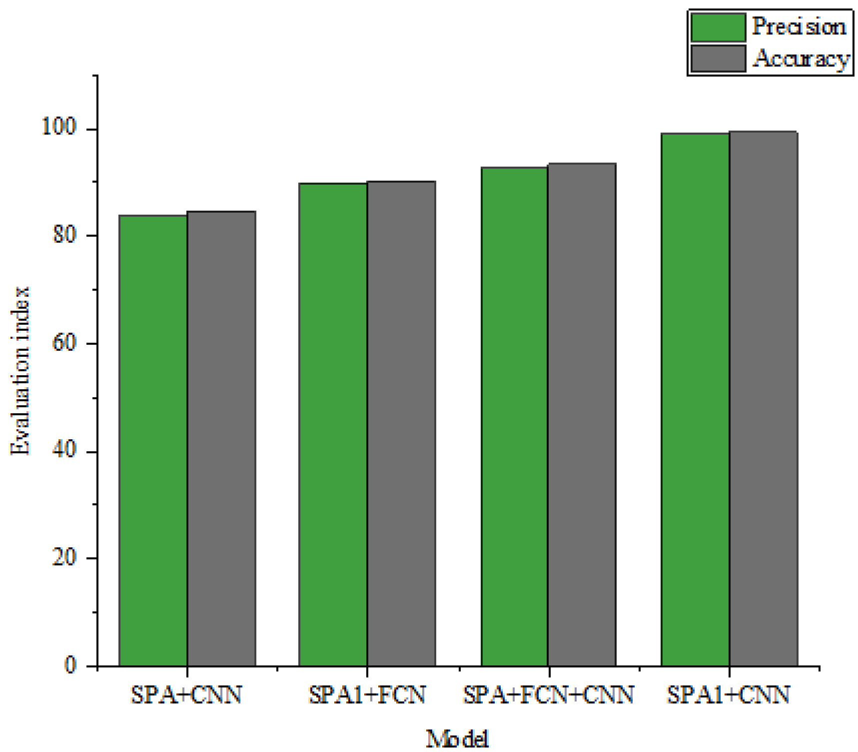

Figure 7. The results of ablation experiments

Based on the analysis above, it is evident that the method developed in this paper excels at accurately segmenting different shape elements within an image, enabling further analysis and processing. Particularly in dynamic environments that change, the method effectively detects and categorizes unchanging points and changing points by extracting the boundaries of each element in the image and assigning them to their respective categories.

This change-vision detection algorithm is not limited to the extraction of sustainable environmental elements alone. It can also be applied to other categories, such as the sky, roads, buildings, and vegetation in natural environments. By leveraging the shape and texture characteristics of the detection target, the algorithm facilitates the precise segmentation of target areas within complex environments. Consequently, it enables more advanced image processing and enhances the visual understanding capabilities for various complex tasks.

4.3. **Discussion.** The empirical analysis discloses the usefulness of the proposed methodology in capably segmenting porcelain images into varied area through the synergistic application of binocular visual analysis, image segmentation, and convolutional neural network technologies. This enables expeditious extraction of the floral art effect, color, and texture from each delineated area. The accuracy accomplished in image segmentation and element extraction markedly strengthens the visual impact and artistic expression embodied in art products.

In addition, binocular vision analysis technology aligns closely with the human-centered design concept. Firstly, it assists designers in gaining a deeper understanding of the interaction between users and products. By extracting and analyzing the three-dimensional shape, texture, color, and other intricate pattern elements of the product, designers can more accurately comprehend the user's perception and experience, thus better addressing the needs and expectations of users in the design process. Secondly, binocular vision analysis technology contributes to the sustainability and humanized design of products. Designers, by incorporating environmentally friendly materials and energy-saving technologies, can create products that are both aesthetically pleasing and environmentally conscious, thereby reducing their impact on the environment. Simultaneously, through the examination of human behavior and habits, designers can craft more ergonomic, comfortable, and convenient products to enhance the overall user experience. Lastly, products infused with a sustainable art component hold the potential for a positive impact on end users and consumers. These products not only meet practical user needs but also convey the principles of environmental protection and sustainable development, encouraging users to prioritize environmental concerns. Additionally, through artistic design, these products can evoke aesthetic appreciation and cultural identity, ultimately enhancing users' quality of life and fostering a sense of social responsibility.

By fine-tuning and designing each section, the level of detail, contrast, and expressiveness of the image can be enhanced, resulting in artwork that is more vibrant, distinctive, and compelling. This approach also empowers artists and designers to better understand and utilize artistic elements, opening up new creative possibilities and inspiring them throughout the art-making process. It aids in the creation of individual and unique artistic products that showcase a distinct artistic perspective and style. In conclusion, the merger of visual analysis techniques, represented by image segmentation, within the realm of artistic product design, covers the imaginative vistas and opens novel possibilities for artists and designers. The assessment and analysis of porcelain patterns not only offer valuable insights but also serve as a base for analysing and improving the methodologies employed in the design and texture production of porcelain. The geometric shapes and textures found in porcelain patterns can also be applied in sustainable building design to improve energy efficiency and enhance the environmental friendliness of building materials. This fosters the harmonious integration of sustainable environments and cultural preservation, elevates the personalization, visual impact, and user experience of products, and introduces new avenues of expression in artistic product design.

5. **Conclusion.** This paper presents a novel approach to constructing a sustainable environmental element extraction model by leveraging image segmentation and convolutional

neural networks. The proposed method utilises binocular visual analysis technology to extract and filter image blocks associated with each pixel point in a landscape image. A new visual detection algorithm is implemented through the training of two PCA convolutional layers within the convolutional neural network. This allows for the extraction of shape, texture, and other features of porcelain patterns, enabling accurate segmentation of each pattern's shape and successful extraction of pattern elements. The experimental analysis of image segmentation and extraction performance demonstrates an impressive segmentation accuracy of 98.3% across various pattern types. This discovery corroborates the uniqueness and usefulness of the proposed resolution. The extraction of intricate pattern elements emerges as a practical reservoir for artists and designers, affording them encouragement and contributory means for their creative quests. Through the thorough dissection of images into discrete zones, artificers and designers gain a deeper understanding and harnessing of sustainable artistic components, incorporating lines, shapes, and hues.

By meticulously dissecting images into discrete areas, craftsmen and designers can employ various methods to deconstruct and reconstruct images. Analyzing the elements comprising the image, such as line thickness, shape changes, and tonal warmth and coolness, allows them to identify potential innovations. These elements can be individually extracted and then recombined to generate entirely new visual compositions. Through an in-depth exploration of different image elements, coupled with sustainable artistic components, the artistic creation process becomes a source of creative possibilities and inspiration. This approach enables the design of products that align more closely with sustainable and human culture. Such endeavors not only propel creative pursuits forward but also exert a positive impact on society and the environment. In human-centered design, user feedback and interaction assume paramount importance. Future research could focus on incorporating artist feedback mechanisms into the binocular vision analysis process, enabling real-time adjustments and optimizations of design schemes.

Acknowledgment. The authors would like to thank those who contributed to this article, as well as anonymous reviewers who provided valuable comments.

REFERENCES

[1] Y. F. Wang, S. P. Chen, Y. C. Lee and C. T. S. Tsai, "Developing green management standards for restaurants: An application of green supply chain management," *International Journal of Hospitality Management*, vol. 34, pp. 263-273, 2013.

[2] Z. Liu, M. Adams, R. P. Cote, Y. Geng and Y. Li, "Comparative study on the pathways of industrial parks towards sustainable development between China and Canada," *Resources, Conservation and Recycling*, vol. 128, pp. 417-425, 2018.

[3] A. Parvaiz, M. A. Khalid, R. Zafar, H. Ameer and M. Ali, "Vision Transformers in medical computer vision-A contemplative retrospection," *Engineering Applications of Artificial Intelligence*, vol. 122, pp. 106126, 2023.

[4] D. C. Prakash, R. C. Narayanan and N. Ganesh, "A study on image processing with data analysis," *AIP Conference Proceedings*, AIP Publishing LLC, vol. 2393, no. 1, pp. 020225, 2022.

[5] G. Carleo, I. Cirac, K. Cranmer, L. Daudet and M. Schuld, "Machine learning and the physical sciences," *Reviews of Modern Physics*, vol. 91, no. 4, pp. 045002, 2019.

[6] M. Tawalbeh, A. Al-Othman, F. Kafiah, E. Abdelsalam, and F. Almomani, "Environmental impacts of solar photovoltaic systems: A critical review of recent progress and future outlook," *Science of The Total Environment*, vol. 759, pp. 143528, 2021.

[7] Y. Peng, A. Rysanek, Z. Nagy and A. Schluter, "Using machine learning techniques for occupancy-prediction-based cooling control in office buildings," *Applied Energy*, vol. 211, pp. 1343-1358, 2018.

[8] H. Hao, E. R. Bartusiak and D. Güera, "Deepfake detection using multiple data modalities," *Handbook of Digital Face Manipulation and Detection: From DeepFakes to Morphing Attacks*, Cham: Springer International Publishing, pp. 235-254, 2022.

[9] D. L. Gast and A. D. Spriggs, "Visual analysis of graphic data," *Routledge*, pp. 199-233, 2014.

[10] Z. Gao, H. Liu, X. Xu, X. Ahou, P. Cui and P. Mao, "Research Progress on Carbon Emissions of Public Buildings: A Visual Analysis and Review," *Buildings*, vol. 13, no. 3, pp. 677, 2023.

[11] S. M. Shah, R. A. Khan, S. Arif and U. Sajid, "Artificial intelligence for breast cancer analysis: Trends & directions," *Computers in Biology and Medicine*, pp. 105221, 2022.

[12] C. K. P Vallabh and Zhao X, "Melt pool temperature measurement and monitoring during laser powder bed fusion based additive manufacturing via single-camera two-wavelength imaging pyrometry (STWIP)," *Journal of Manufacturing Processes*, vol. 79, pp. 486-500, 2022.

[13] R. Xia, R. Su, J. Zhao, Y. Chen and S. Fu, "An accurate and robust method for the measurement of circular holes based on binocular vision," *Measurement Science and Technology*, vol. 31, no. 2, pp. 025006, 2019.

[14] D. Jiang, Z. Zheng, G. Li, Y. Sun, J. Kong and G. Jiang, "Gesture recognition based on binocular vision," *Cluster Computing*, vol. 22, pp. 13261-13271, 2019.

[15] X. Yuan, D. Li, P. Sun, G. Wang and Y. Ma, "Real-Time Counting and Height Measurement of Nursery Seedlings Based on Ghostnet-YoloV4 Network and Binocular Vision Technology," *Forests*, vol. 13, no. 9, pp. 1459, 2022.

[16] X. Zhou, Z. Li, H. Xie, T. Feng, Y. Lu and C. Wang, "Leukocyte image segmentation based on adaptive histogram thresholding and contour detection," *Current Bioinformatics*, vol. 15, no. 3, pp. 187-195, 2020.

[17] Y. Song and H. Yan, "Image segmentation techniques overview," *2017 Asia Modelling Symposium (AMS)*, IEEE, pp. 103-107, 2017.

[18] E. Song, J. Long, G. Ma, H. Liu and C.C. Hung, "Prostate lesion segmentation based on a 3D end-to-end convolution neural network with deep multi-scale attention," *Magnetic Resonance Imaging*, vol. 99, pp. 98-109, 2023.

[19] M. Xin and Wang Y, "Research on image classification model based on deep convolution neural network," *EURASIP Journal on Image and Video Processing*, vol. 2019, no. 1, pp. 1-11, 2019.

[20] Z. Li, F. Liu, W. Yang, S. Peng and J. Zhou, "A survey of convolutional neural networks: analysis, applications, and prospects," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 12, pp. 6999-7019, 2021.

[21] S. A. P. R Sekaran, C. P. Lee A and K. M. Lim, "Facial emotion recognition using transfer learning of AlexNet," *2021 9th International Conference on Information and Communication Technology (ICoICT)*, IEEE, pp. 170-174, 2011.

[22] P. Oza, P. Sharma, S. Patel and P. Kumar, "Deep convolutional neural networks for computer-aided breast cancer diagnostic: a survey," *Neural Computing and Applications*, vol. 34, no. 3, pp. 1815-1836, 2022.

[23] K. J. Praveen and S. Domnic, "Rosette plant segmentation with leaf count using orthogonal transform and deep convolutional neural network," *Machine Vision and Applications*, vol. 31, no. 1, pp. 6, 2020.

[24] S. Shangguan, X. Qiu and K. Fu, "Research on A Special Hyper-Pixel for SAR Radiometric Monitoring," *Remote Sensing*, vol. 15, no. 8, pp. 2175, 2023.

[25] Y. Ren, J. Huang, Z. Hong, W. Lu, J. Lin and L. Zou, "Image-based concrete crack detection in tunnels using deep fully convolutional networks," *Construction and Building Materials*, vol. 234, pp. 117367, 2020.

[26] X. Chang, H. Pan, D. Zhang and Q. Sun, "A memory-optimized and energy-efficient CNN acceleration architecture based on FPGA," *2019 IEEE 28th International Symposium on Industrial Electronics (ISIE)*, IEEE, pp. 2137-2141, 2019.

[27] R. Iqbal, A. Behjat, R. Adlakha and J. Callanan, "Auto-differentiable Transfer Mapping Architecture for Physics-infused Learning of Acoustic Field," *IEEE Transactions on Artificial Intelligence*, pp. 1-15, 2023.

[28] M. Y. Wong, C. W. Chen, Y. H. Tseng, S. K. Zhou, Y. H. Lin, "Noncontrast MRI in assessing venous reflux of legs using QFlow analysis and radial basis function neural network technique," *Scientific Reports*, vol. 13, no. 1, pp. 3263, 2023.

[29] B. Su and M. Yuan, "Object Recognition for Millimeter Wave MIMO-SAR Images Based on High-resolution Feature Recursive Alignment Fusion Network," *IEEE Sensors Journal*, vol. 23, no. 14, pp. 16413-16427, 2023.

[30] S. Niwattanakul, J. Singthongchai and E. Naenudorn, "Using of Jaccard coefficient for keywords similarity," *Proceedings of the International Multiconference of Engineers and Computer Scientists*, vol. 1, no. 6, pp. 380-384, 2013.