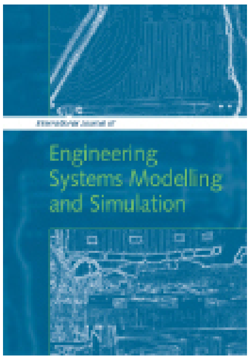


International Journal of Engineering Systems Modelling and Simulation

ISSN online: 1755-9766 - ISSN print: 1755-9758
https://www.inderscience.com/ijesms

Forms and innovative applications of fine arts factors in the design of literary and artistic products

Xin Yu

DOI: 10.1504/IJESMS.2026.10079183

Article History:

Received:	12 April 2025
Last revised:	07 February 2026
Accepted:	03 May 2026
Published online:	25 June 2026

Forms and innovative applications of fine arts factors in the design of literary and artistic products

Xin Yu

Department of Art,
Shangluo University,
Shangluo, Shaanxi Province
726000, China
Email: forwards226@163.com

Abstract: As material living standards improve, consumers now seek literary and artistic products that offer both function and visual appeal. Attention is increasingly focused on design, form, and aesthetics, making purely functional items less attractive. However, current research on image aesthetics often overlooks the role of human visual saliency and attention in beauty perception, limiting evaluation performance. To address this, the study proposes a novel method that mirrors the brain's processing of visual stimuli. It introduces two algorithms: one based on compositional edge features and another using weakly supervised learning for aesthetic classification through attention mechanisms. By combining visual saliency with aesthetic evaluation, the approach better aligns with human perception and improves assessment accuracy. This method enhances understanding of aesthetic judgment and supports the development of more appealing designs, offering a valuable framework for evaluating the visual quality of literary and artistic products in today's design-conscious market.

Keywords: fine arts; literary design; product design; artistic products; saliency-based aesthetic analysis; visual saliency; consumer preferences; human visual perception.

Reference to this paper should be made as follows: Yu, X. (2026) 'Forms and innovative applications of fine arts factors in the design of literary and artistic products', *Int. J. Engineering Systems Modelling and Simulation*, Vol. 17, No. 7, pp.1–11.

Biographical notes: Xin Yu graduated from the School of Art of Anhui University in 2015, with a Master's in Art. She is currently a Lecturer at the School of Art of Shangluo University, mainly focusing on research of art creation and art design. She has hosted various scientific research projects, has published more than 20 papers.

1 Introduction

As an effective means of product upgrading and design innovation, literary product design plays a vital role in improving product quality and sales (Gao et al., 2020). It has established a close relationship with China's economy and manufacturing industry (Talei and Pearlman, 2022). As former British Prime Minister Margaret Thatcher said, "one point of investment in the design of literary and artistic products can generate a thousand points of return" (Aslan et al., 2020). The development of literary and artistic product design in China only has a short history of 30 years (Bessghaier et al., 2022). While its vigorous development is undeniable, this emerging industry's shortcomings, such as the confusion of design concepts and the ambiguous design style, must also be acknowledged. The design expression is rendered exceedingly faint by these elements, and numerous contemporary literary product designs now depend on the principle of "one imitation, two changes, and three creations" to generate innovative products (Carroll et al., 2022). Most of these designs are imitations of European and American designs, as well as plagiarisms of Japanese and Korean design styles. However, the influence of Chinese national culture on design cannot be found (Rahman and Koszewska, 2020).

All this makes Chinese literary product design appear weak and powerless in market competition (Papadimitriou et al., 2021). To develop the design of Chinese literary and artistic products and to stand out in the world of design, it must possess a unique design style and aesthetic principles characteristic of the country. This depends on fine arts. Although the design of Chinese literary and artistic products has experienced rapid development over the past 30 years, there are still significant shortcomings in terms of aesthetic expression and cultural inheritance. European and American design emphasises creating visual impact through strong colour contrast and exaggerated styling, while traditional Chinese aesthetics focuses more on implicit expressions such as 'leaving blank space' and 'hiding dew', guiding viewers' attention flow through the virtual and real changes of edge lines. The current imitative design forcibly superimposes the visual salience mechanisms of different cultures, leading to cognitive conflicts among consumers during visual decoding – for example, in cultural and creative products, using both Western geometric colour blocks and traditional patterns weakens the aesthetic coherence of the design. More importantly, the lack of ethnic cultural elements has caused the design to lose its unique aesthetic anchor point. The principles of 'vivid charm' and 'bone technique and pen use' in traditional Chinese art are essentially a summary of the organisational law of visual elements: the line rhythm of calligraphy and the hollowed out rhythm of paper cuttings all contain the composition logic that conforms to the visual perception habits of Chinese people. To construct a design style with Chinese characteristics, we need to start from the visual essence of artistic factors: on the one hand, we need to explore the elements in traditional aesthetics that are compatible with the commonality of human visual perception, such as the satisfaction of visual balance through symmetrical composition. On the other hand, transforming ethnic cultural symbols into distinctive features that conform to modern aesthetics, such as extracting the texture of seal carved gold and stone into the texture treatment of product edges. By organically integrating visual elements, design can not only carry cultural connotations, but also be quickly perceived and recognised by consumers.

The content of product design must first be understood to analyse its factors (Keshari et al., 2021). In the traditional sense, product design is only regarded as the design work of beautifying the appearance of the product (Seifert and Chattaraman, 2020). There are many factors involved in product design that determine the entire process of the product, from the initial design concept to production, sales, use, and disposal. From this, it can be seen that product design is a series of activities that solve problems throughout a product's existence (Grein et al., 2021). Therefore, product designers should not only consider the appearance of the product when designing but also ensure that the function and structure of the product are reasonable, easy to operate and use, safe, easy to maintain, and coordinated with the environment. The relationships among these three aspects must be well-balanced (Kam, 2021). This requires designers to collect and accumulate a lot of knowledge and materials (Hamid and Celik-Butler, 2021). Design activity is not a closed, self-contained process but rather a symbol encoded by designers who integrate people, market competition, fashion, social culture, and other factors into market competition, and consumers decode it through market sales (Vieira, 2021). This requires that the design department cooperate closely with the sales and production departments in the process of product design to obtain excellent products with good performance, suitable for the market, and easy to manufacture and sell (Landy et al., 2020).

Designing a product is a creative activity in itself. Therefore, the application of creative thinking in product design plays a crucial role in determining the design's outcome (Naderi et al., 2020). According to the creation process of product design, the types of product design are initially divided into two categories: realistic product design and conceptual product design. The general characteristics of creative thinking in design, as well as the advanced nature of creative thinking in conceptual design, are then summarised (Alcaide-Marzal et al., 2020). Then, the relationship between design methods and creative thinking is analysed. The relationship between creative thinking and macro-level systematic design methods is examined, as is the creative thinking reflected in micro-level application techniques within the program (Petrusch et al., 2020). Finally, two typical exceptional cases of creative thinking in product design are analysed: the divergence-convergence process of thinking and the inspiration behind design (Schneider et al., 2020). Therefore, this work suggests an image aesthetic evaluation algorithm based on the fusion of visual saliency and compositional edge information (Gatchalian et al., 2020). Inspired by the saliency mechanism of the human visual system, the saliency regions of images are used in aesthetic evaluation research (Afrianti and Zainul, 2021). However, unlike common recognition tasks, when evaluating image aesthetics, the human brain also focuses on factors such as the image background and composition, which can affect its evaluation.

This article explores the construction of an aesthetic evaluation method for art product images that is more in line with human perception, in order to accurately reflect consumers' aesthetic needs for art products that combine functionality and visual appeal. The research gap lies in the fact that existing studies on image aesthetics have overlooked the role of human visual salience and attention in aesthetic perception, and have not fully considered the impact of factors such as image background and composition on aesthetic evaluation, resulting in limited evaluation effectiveness and difficulty in meeting the practical needs of visual quality evaluation in art product design. The novelty lies in proposing an image aesthetic evaluation method that combines visual saliency and compositional edge information. On the one hand, it utilises the saliency mechanism of the human visual system to focus on salient regions, and on the other hand, it uses edge features to describe background regions and composition rules. At the same time, by introducing attention mechanisms based on combined edge features and weakly supervised learning aesthetic classification algorithms, effective integration of multidimensional information is achieved. The research has constructed an aesthetic evaluation model that is more in line with human perception laws, improving the accuracy of aesthetic evaluation of artistic product images. Enhance the understanding of aesthetic judgment and provide scientific basis for designing more attractive artistic products. This provides a valuable framework for the visual quality assessment of art products in today's design awareness market, which helps promote the innovative application of artistic factors in art product design.

The research and innovation contributions are as follows:

1 This study innovatively combines visual saliency mechanism with compositional edge features. By focusing on key areas of human natural concern through visual salience, capturing the core appeal in

aesthetic perception. By utilising edge features to analyse background structure and composition rules, it fills the gap in traditional methods' consideration of "overall visual context".

2 Innovatively combining attention mechanisms with weakly supervised learning algorithms: attention mechanisms can dynamically allocate weights and highlight the interactive effects between salient regions and composition elements. Weakly supervised learning automatically mines hidden aesthetic patterns in limited annotated data.

3 By accurately reflecting consumers' comprehensive needs for 'functionality + visual appeal', provide designers with quantifiable optimisation directions. Adjust the composition edges to enhance product feature recognition, or enhance visual appeal by strengthening prominent areas.

Section 1 of the study elaborates on the research background of literary product design in improving product quality and sales. This section analyses a method proposed in the work that combines saliency objectives with aesthetic evaluation methods. Section 2 elaborates on materials and methods, analyses an overview of aesthetic evaluation, and identifies three levels of image quality assessment: fidelity, comprehensibility, and aesthetics. Section 3: image aesthetics evaluation model based on visual highlighting and composition edge information fusion. Section 4 verifies whether each module in the proposed algorithm contributes to its overall performance to a certain extent, and tests the performance contributed by using each module separately. Section 5 summarises the entire text.

2 Product aesthetic design

2.1 Design aesthetics

Beauty is a quality that can stir people's hearts and spirits or bring pleasure to their senses. 'Beauty' reflects the psychological connection between the aesthetic subject and the aesthetic object, where the aesthetic object has an impact on the aesthetic subject. Understanding the form created by industrial design from this perspective involves two levels of meaning. The term 'shape' typically refers to the physical characteristics, such as colour and texture, that are perceived by people's senses. Normally, it aims to convey a functional beauty to consumers, encompassing aspects such as the appropriateness of the human-machine scale, the comfort of texture, and the suitability of shape, among others. The 'state' refers to the spiritual situation, or 'Shen Yun', contained within the object, which often has specific meanings and symbols or evokes people's experiences of a particular emotion.

Aesthetics in industrial design refers to the combination of a product's shape and its overall character. Here, the complementary combination of 'shape' and 'spirit' is not a simple patchwork but the essence of a particular spirit integrated with the external shape of the product. In essence, beauty is the unity of the feeling of beauty and the moving of beauty, the unity of the pleasure of the senses and the satisfaction of reason. Regarding industrial design, the purpose of product modelling is to meet the physiological and psychological needs of human beings, to make the product user feel that the product is easy to use, to make the viewer of the product feel that the product is pleasing to the eye, and to make the holder of the product feel the taste and style that the product represents.

The evaluation standard of product aesthetics comprises formal aesthetics derived from identification, cognition, integrity, and anthropomorphism, as well as empirical rules derived from practical operations. The following common characteristics are generally possessed by them:

1 A harmonious relationship between the overall shape of the product and its environment, as well as the value of the product expressed through its shape, colour, and material.

2 Whether the overall shape conveys the product's function, whether it is safe, and whether it meets its operational requirements.

3 Whether the product's shape is deliberate, expressing clear structure and modelling principles.

4 Whether the shape can evoke spiritual resonance and whether the overall performance can capture the user's interest and curiosity, evoking a pleasant feeling.

5 When selecting modelling materials, the impact on the ecological environment should be considered during production and in future scrap recycling.

The concept of design must first be clarified before design aesthetics can be addressed. What is 'design'? In the broadest sense, all the biological and social activities of human beings can be considered a form of design. Design is 'the process of thinking formation', and design functions as a distinct art discipline. The study's content and service targets diverge from traditional art groups, resulting in design aesthetics that are distinct from conventional painting and decoration; furthermore, the research content is grounded in nature. Traditional aesthetic theories cannot be copied entirely. Design itself has three attributes: utilitarian, aesthetic, and ethical. Aesthetic features play a crucial role in design.

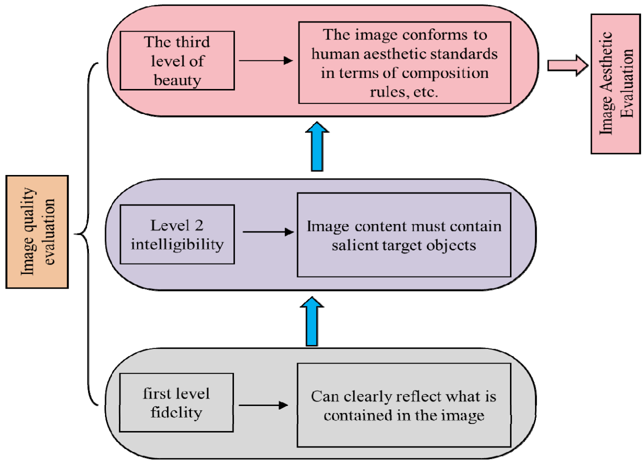
Design is an all-encompassing field that encompasses various elements, including society, culture, economy, market, technology, and ethics. The aesthetic standards in design evolve in tandem with these elements. The main content of the basic principles of design aesthetics is a new discipline developed based on modern design theory and application, combined with traditional theories of aesthetics and art research. Its basic tenets mainly include formal beauty, social beauty, technical beauty, object beauty, style, design subject, design language, and aesthetic taste.

Artistic aesthetics and general aesthetics diverge primarily in their focus on aesthetic objects, with works of art serving as the central objects of aesthetic consideration in artistic aesthetics. In contrast, general aesthetics centres on the appreciation of the beauty found in reality. In reality, tangible entities or occurrences do not function as the subjects of people's practical activities or cognitive processes. Within social practice, a will-practice relationship and an intellectual-cognitive relationship are first formed with people. On this basis, when society develops to the point where people no longer treat the object with a direct utilitarian and practical attitude, a more mature and refined aesthetic relationship will emerge. The concept of artistic beauty differs from that of realistic beauty, as it is specifically created and exists as an aesthetic entity. Artists dedicate themselves to creating works of art, aiming to fulfil people's aesthetic requirements in art. Art provides individuals with a form of aesthetic pleasure, enjoyment, and evaluation, playing a subtle yet significant role in shaping the human spirit.

2.2 Overview of aesthetic evaluation

The theoretical exploration of product design aesthetics mentioned earlier in this article lays the core theoretical foundation for the subsequent technical modelling of aesthetic evaluation. The core of product aesthetics should revolve around user sensory pleasure, rational satisfaction, and other aesthetic needs to construct evaluation criteria, which is the logical starting point of aesthetic evaluation technology modelling. In product design aesthetics, the core dimensions of psychological connection between aesthetic subjects and objects, as well as aesthetic evaluation, clarify the core elements of human aesthetic judgment that aesthetic evaluation technology modelling needs to simulate. Based on this, the following text naturally transitions to image aesthetic evaluation, focusing on users and using technologies such as deep learning (DL) to model human aesthetic perception and cognitive processes in the aforementioned theory. The organic connection between theoretical concepts and computational frameworks is achieved, filling the transitional gap between the two and enhancing the logical coherence of the paper. The three levels of image quality evaluation, as shown in Figure 1, include fidelity, intelligibility, and aesthetics. Image aesthetic quality evaluation relies on image quality evaluation and is the third level in image quality evaluation. Traditional image quality evaluation primarily assesses the impact of image distortion type and degree on image quality. However, with the update of image acquisition equipment and the improvement of image processing technology, it is no longer challenging to obtain more explicit images. Therefore, the research focuses on image quality evaluation, shifting to the third level, namely, image aesthetic evaluation. Image aesthetic evaluation aims to put users at the core, guided by computer vision and other theories, utilising DL and other technologies to model the process of human perception and cognitive aesthetics, thereby achieving the goal of simulating human aesthetic judgment.

Figure 1 Three levels of image quality evaluation (see online version for colours)



In the era of big data, helping users access the necessary resources more efficiently and effectively is a crucial and pressing issue. Users often reveal their emotions and preferences in the social process. Aesthetics help users present themselves more clearly in the social process. In the face of massive data, subjective evaluation of image aesthetics requires not only a high cost but also a large workload. Currently, it is necessary to develop an aesthetic evaluation model that can automatically assess the aesthetic quality of images. The key point in designing an image aesthetic evaluation model with high accuracy and performance is understanding how to utilise the mechanisms of human aesthetics. Since human aesthetics is highly subjective and abstract, it encompasses various fields, including psychology, photography, and visual art, making it more challenging than traditional image quality evaluation. In the face of complex aesthetic evaluation tasks, it is necessary to analyse human perception of aesthetics more accurately from the user's perspective so that the designed algorithm can better align with human aesthetics and serve users more effectively.

2.3 The core design principles of literary and artistic products

Culture is the soul of a product, and it is the key to creating culturally and creatively rich products. Literary and artistic products should fully reflect the cultural connotation and heritage of the product; that is also a crucial element in the core competitiveness of cultural and creative products. Delivering product value through culture meets people's material and spiritual needs. Unique and charming literary and artistic products effortlessly capture consumers' attention and evoke emotional resonance.

Innovation is the key to product design. The characteristic of literary and artistic products is to meet people's spiritual needs. Innovative design can increase the added value of products. The higher the creativity of the product, the greater the value will be. In the face of dazzling products and fierce market competition, innovative designs can quickly capture the market, gain consumer recognition, and meet the unconventional and distinctive psychological needs of today's people, thereby increasing product variety.

The current economy is the experience economy. People attach great importance to their sense of experience, and whether it resonates with the audience serves as a critical metric for assessing the effectiveness of a product's development. Experience design can enable the audience to engage with the experience product actively, allowing the culture to be disseminated imperceptibly and increasing people's sense of identity with the product. Through the experience of using the product, people's recognition and acceptance of it are strengthened. This is conducive to the occurrence of purchasing behaviour, making it a product recognised by the market.

Design is inseparable from aesthetics. Product design adheres to the principles of design beauty. The design of a product requires an in-depth study of the product's shape, colour, and expression, which echoes the theme and design elements of cultural and creative products. People cannot resist aesthetically rich products. Products not only have practical value but also require ornamental value and must be multifunctional. In an era of a highly developed economy, where technology and product supply often exceed demand, the aesthetics of products are critical. Designers can only stand out from many products of the same type by creating products that combine practicality and appreciation, thereby enhancing the cultural value of the products.

3 Fusion operation method

3.1 *Image aesthetic evaluation model based on the fusion of visual saliency and compositional edge information*

In this chapter, a block diagram illustrating the proposed image aesthetic evaluation algorithm is presented in Figure 2. The algorithm, designed to combine visual saliency and compositional edge information, consists of three primary components: a visual saliency detection module, a compositional edge information feature module, and a deep feature learning module. To begin with, the visual saliency detection algorithm is employed to achieve the saliency area map corresponding to the image in the aesthetic image database. The edge features of composition are the direct reflection of object contours, region boundaries, and spatial layout in an image. They not only define the form and distribution of visual elements, but also carry core aesthetic elements such as balance, rhythm, and hierarchy in the picture. In aesthetic evaluation, the human brain's perception of images often starts from the edges, grasping the correspondence and overall structure between objects through edge features, and judging whether the composition is coordinated and whether the primary and secondary are clear. Subsequently, the edge feature network model is applied to acquire the composition edge feature map of the image. Lastly, a deep neural network (DNN) is employed to extract the depth feature of the image in the aesthetic image database. Once the saliency features, edge features, and DL features are obtained, a feature fusion operation is performed. This operation combines the features above, resulting in fused features that are then utilised to predict the final image aesthetic score.

Figure 2 The block diagram of the image aesthetic evaluation algorithm based on the fusion of visual saliency features and composition edge information (see online version for colours)

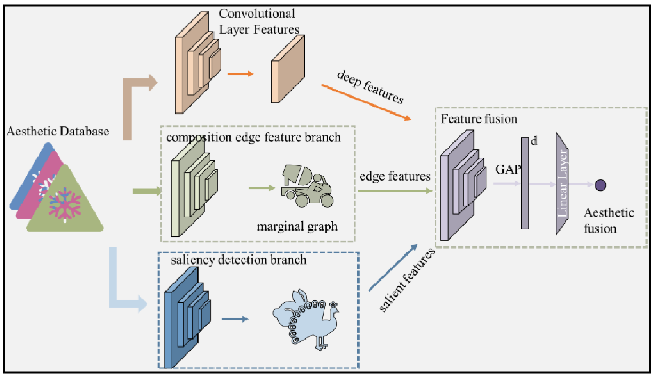


Figure 2 shows the process of an image aesthetic evaluation model based on the fusion of visual saliency and compositional edge information. The image data comes from an aesthetic database, and the images are first processed through convolutional layers to extract deep features. These deep features capture high-level abstract features of the image. The image is processed through an edge feature extraction network to generate edge features. These edge features are used to construct a marginal graph, capturing the structure and contour information of the image. The features from three branches (depth features, edge features, saliency features) are input into the feature fusion module. The feature fusion module processes these features through a series of conv layers and global average pooling (GAP). Finally, the fused features are mapped to an aesthetic score through a linear layer. This model extracts depth features, edge features, and saliency features of an image through three parallel branches, and then fuses these features to predict the aesthetic rating of the image. This multi feature fusion method can comprehensively capture the aesthetic attributes of images, improving the accuracy and robustness of evaluation. Early fusion refers to the fusion of multi-layer features first, and then the predictor is trained based on the fused features. *Add* and *concatenate* are two commonly used methods for early fusion. Inside-outside net and hypernet are classic early fusion methods. The proposed model can be regarded as a converged network, as it integrates deep features, edge features, and saliency features within a unified framework for aesthetic assessment.

Assuming that the input is two-way, then the outputs after *add* and *concatenate* are:

$$Z_{add} = \sum_{i=1}^{c} X_i * K_i + \sum_{i=1}^{c} Y_i * K_i \tag{1}$$

$$Z_{concatenate} = \sum_{i=1}^{c} X_i * K_i + \sum_{i=1}^{c} Y_i * K_{i+c} \quad (2)$$

Among them, * represents the convolution operation. The *add* method can be understood as the superposition of values, and the number of channels remains unchanged. The add operation can increase the amount of information in each dimension of the feature describing the image, but the dimension of the feature remains unchanged. The *concatenate* method combines features in the channel domain, which enhances the description of image features without increasing the amount of information per feature dimension.

Two loss functions, namely Euclidean loss and mean absolute error loss (MAE), are employed concurrently to facilitate network training, thereby enhancing the stability of the training process. The formula outlined below is used for performing calculations:

$$Loss = \frac{1}{N}\sum_{i=1}^{N} \underbrace{s_i - p_i{}_2^2}_{MSE} + \frac{1}{N}\sum_{i=1}^{N} \underbrace{|s_i - p_i|}_{MAE} \quad (3)$$

In the given context, s denotes the score value associated with the label of the i^{th} image, while p represents the score value forecasted by the network for the i^{th} image.

3.2 Converged network

The DFs and AAFs are used to generate weights for each other, and the activation unit generates the corresponding weights. The Sigmoid function is used as the activation function. The calculation formula is:

$$S(x) = \frac{1}{1+e^{-x}} \quad (4)$$

The Sigmoid function offers several advantages, including simple implementation, convenient derivation, reduced sensitivity to noise data, and a pronounced effect when the difference between features is not significant. Therefore, the Sigmoid function is more suitable as an activation function for weight generation during fusion. After the weight is calculated, it is multiplied by the feature and then added to the original feature as the final fused feature; that is, the fused feature is obtained by the following calculation method:

$$F_{new-aes}(x) = f_{aes}(x) * \left(1+\sigma\left(f_d(x)\right)\right) \quad (5)$$

$$F_{new-d}(x) = f_d(x) * \left(1+\sigma\left(f_{aes}(x)\right)\right) \quad (6)$$

Among them, f_{aes} represents the aesthetic attention feature, It refers to the shallow aesthetic features with visual significance that are first captured by the human eye and visual cognitive system in the process of visual aesthetics. It is the fundamental visual feature that triggers aesthetic cognition, and also the prerequisite for subsequent deep aesthetic feature extraction and fusion. f_d represents the deep aesthetic feature, σ represents the activation unit, which is the Sigmoid activation function. It refers to the deep features obtained through visual cognitive processing, combined with aesthetic laws and scene aesthetic needs, based on aesthetic attention characteristics. $F_{new-aes}$ represents the new aesthetic attention feature fused with the deep aesthetic feature information after the interactive operation, and F_{new-d} represents the new deep aesthetic features after the interaction.

In the training stage of the image aesthetic classification model, this study mainly uses two major databases, aesthetic visual analysis (AVA) and aesthetic attributes database (AADB), as the core training data. Among them, AVA database is currently a large-scale and widely used image aesthetic evaluation dataset, containing over 250,000 images from professional photography websites. Each image is accompanied by a 1–10 point aesthetic rating and detailed semantic labels (such as composition, lighting, colour, etc.) annotated by thousands of reviewers, which can comprehensively reflect human subjective perception of image aesthetics in different scenarios. The AADB database focuses on fine-grained annotation of image aesthetic attributes, covering approximately 5,000 images. In addition to aesthetic ratings, each image also contains labels for 12 specific aesthetic attributes such as symmetry, balance, and rhythm, providing precise annotation support for the model to learn the association between composition edges and aesthetic evaluations. The combination of two databases not only ensures the scale and diversity of training data, but also enhances the model's ability to capture aesthetic features through fine-grained labelling, laying a solid data foundation for the subsequent training of aesthetic evaluation models that integrate visual saliency and composition edge information. In the first stage, the image aesthetic classification model is first trained. This chapter utilises the images in the AVA database and their corresponding aesthetic classification labels as the data for training a deep network. The backbone network selects EfficientNet and utilises the EfficientNet model, pre-trained on ImageNet, for initialisation. The cross-entropy loss function is employed to optimise the parameters of the classification network during training. The calculation formula for the cross-entropy loss function is as follows:

$$CE = -\frac{1}{N}\sum_{i=1}^{N}\left[s_i \log p_i + (1-s_i)\log(1-p_i)\right] \quad (7)$$

Within this framework, s represents the score value associated with the label of the i^{th} image, whereas p denotes the score value predicted by the DNN for the i^{th} image.

In the second stage of training, which primarily focuses on integrating aesthetic attention features and deep aesthetic features, the ResNet and EfficientNet models pre-trained on ImageNet are utilised for initialisation. Euclidean loss and MAE are employed. The MAE is an error loss that, when used in conjunction with another loss function, can effectively direct the training of a network. This combined approach enhances the stability of the network training procedure. The computation formula for the loss function is given by:

$$Loss = \frac{1}{N}\sum_{i=1}^{N} s_i - p_{i2}^{2} + \frac{1}{N}\sum_{i=1}^{N}\left|s_i - p_i\right| \tag{8}$$

In the given context, s denotes the score value associated with the label of the i^{th} image, while p represents the score value forecasted by the network for the i^{th} image.

Therefore, it is necessary to clarify the various levels of the elements, so it is required to use multi-factor fuzzy comprehensive evaluation (FCE), that is, to start the comprehensive review from the lowest-level factors, use the obtained results to continue to carry out a thorough assessment based on the factors of the upper level, and move up layer by layer. Evaluation: until the review reaches the highest level so far, to get the total result.

Let the factor set be:

$$U = \{u_1, u_2, u_m\} \tag{9}$$

u_i is the i^{th} factor in the first level (the highest level) and is also determined by n factors in the second level, namely:

$$u_i = \{u_{i1}, u_{i2}, \ldots, u_{in}\}(i = 1, 2, \ldots, m) \tag{10}$$

The weight of each factor is determined by its importance at the specified level. The weight set of each factor level is as follows:

The weight vector of the first level:

$$\underline{A} = (a_1, a_2, \ldots, a_m) \tag{11}$$

That is, the weight of the first factor U in the first level is a ($i = 1, 2, \ldots, m$).

The weight vector of the second level:

$$\underline{A_i} = (a_{i1}, a_{i2}, \ldots, a_{in})(i = 1, 2, \ldots, m) \tag{12}$$

That is, the weight of the j^{th} factor u_{ij} of the determinant u in the second level is a_{ij} ($i = 1, 2, \ldots, m$; $j = 1, 2, \ldots, n$). Factors at different levels have their corresponding weight vectors and weights.

Since the lower-level factors all determine the factors at the upper level, the single-factor evaluation (SFE) of any factor is essentially a comprehensive multi-factor evaluation of its lower level. Therefore, an FCE should start at the lowest level. The first-level FCE involves a thorough examination of the lowest-level factors. In the case of considering only two-factor levels, the first-level FCE should be conducted following the second-level factors. Assuming that the judgment object is judged by the factor u_{ij} in the second level, the degree of membership of the k elements in the factor set are:

$$r_{ijk}\,(i = 1, 2, \ldots, m;\ j = 1, 2, \ldots, n;\ k = 1, 2, \ldots, p) \tag{13}$$

Therefore, the SFE matrix of the second level is:

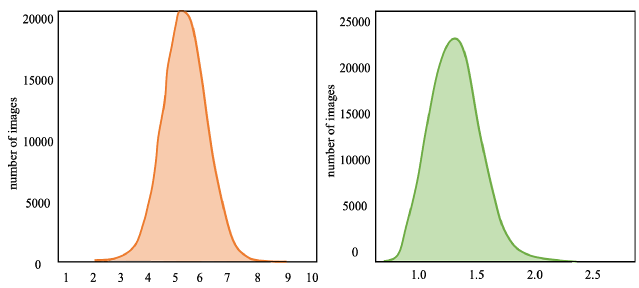
$$R_i = \begin{bmatrix} r_{i11} & \cdots & r_{i12} & \cdots & r_{i1p} \\ r_{i21} & \cdots & r_{i22} & \cdots & r_{i1p} \\ r_{in1} & \cdots & r_{in2} & \cdots & r_{inp} \end{bmatrix} \tag{14}$$

The j^{th} row in the matrix represents the judgment result by the j^{th} factor, u_{ij}, in the second level. The number of rows in the R_i matrix is determined by the number of u_{ij} factors that determine it. The number of columns in the R_i matrix is determined by the number of elements in the judgment set. The second-level FCE set is:

$$\underline{B_i} = \underline{A_i} \circ R_i = (a_{i1}, a_{i2}, \ldots, a_{in}) \tag{15}$$

Existing studies often use AVA or AADB datasets alone for image aesthetic evaluation, or simply integrate their basic data, without fully tapping into the subjective perceptual value of AVA's large-scale semantic labels and the compositional feature advantages of AADB's fine-grained aesthetic attribute labels. This study innovatively combines two major databases, relying on AVA to ensure the scale and diversity of training data, and utilising the 12 aesthetic attribute labels of AADB to enable the model to accurately learn the intrinsic relationship between composition edges and aesthetic evaluation. At the same time, based on the EfficientNet backbone network, an aesthetic classification model is constructed by integrating visual saliency and composition edge information, filling the research gap in the existing fusion of dual datasets that have not targeted fine-grained attribute mining and lack multi-dimensional visual feature integration, and improving the model's ability to capture aesthetic features. As shown in Figure 3, although the AVA database contains a large amount of data, the images in the AVA database comprise a significant number of professionally processed images, and some images also contain text information. In contrast, the images in the AADB database are more closely aligned with real-life scenarios. For images in the scene, AADB is more suitable than AVA in terms of realism. However, examining the distributions of the aesthetic score labels in the two databases reveals that the score distribution of the AVA database is closer to a normal distribution, making the normal distribution scores easier for the network to learn and predict.

Figure 3 Comparison of AADB database and AVA database, (a) average of aesthetic scores (b) standard deviation of aesthetic scores (see online version for colours)



(a) (b)

Figure 3 shows a comparison of aesthetic scores between the AADB database and the AVA database. The horizontal axis represents the average aesthetic score, ranging from 1 to 10. The vertical axis represents the number of images. The orange curve and shaded area show the distribution of the average aesthetic score in the AADB database. It can be seen that the average score is concentrated between 5 and 6, showing a clear peak, indicating that the average aesthetic score of most images is within this range. This figure shows the distribution of the average and standard deviation of aesthetic scores in the AADB database through two subgraphs. From the graph, it can be seen that the average aesthetic score of most images is concentrated between 5 and 6, while the standard deviation is concentrated between 1.2 and 1.6. This indicates that the images in the AADB database have a certain degree of concentration and dispersion in aesthetic ratings.

4 Experimental results and analysis

4.1 Experimental setup

This chapter employs two image aesthetic databases, namely AADB and AVA, to assess the effectiveness of the proposed aesthetic evaluation algorithm. The backbone network utilised in both the AADB and AVA databases consists of two widely recognised deep network models, ResNet and EfficientNet. Thus, ResNet and EfficientNet serve as the backbone networks in both AADB and AVA image aesthetic databases.

Figure 4 Performance comparison of EfficientNet and other CNN networks (see online version for colours)

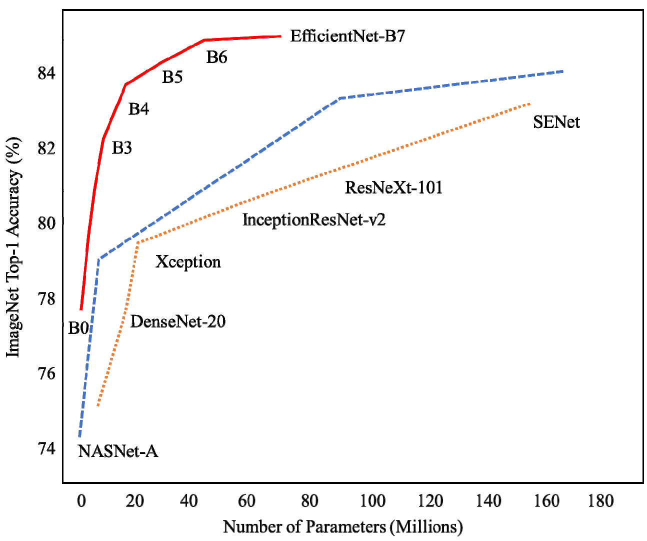


Figure 4 compares the performance of EfficientNet and other classic convolutional neural networks in terms of model size and computational effort. Combining the results in Figure 4, it can be concluded that the TOP1 classification accuracy index of the classification task on ImageNet and the EfficientNet network achieves the best performance compared to existing CNN networks, including the popular ResNet network model. At the same time, the number of parameters in the EfficientNet network is significantly lower than that of other types of CNN networks, and it also maintains absolute superiority in the calculation volume (FLOPS) indicator. With the continuous improvement of model efficiency in computer vision, EfficientNet strikes a balance between speed and accuracy, maintaining fewer parameters, which significantly facilitates the training of network models, reduces the consumption of computing resources during training, and lowers the cost of training DNN models. Table 1 compares the performance of the model with other state-of-the-art methods.

Table 1 Performance comparison table

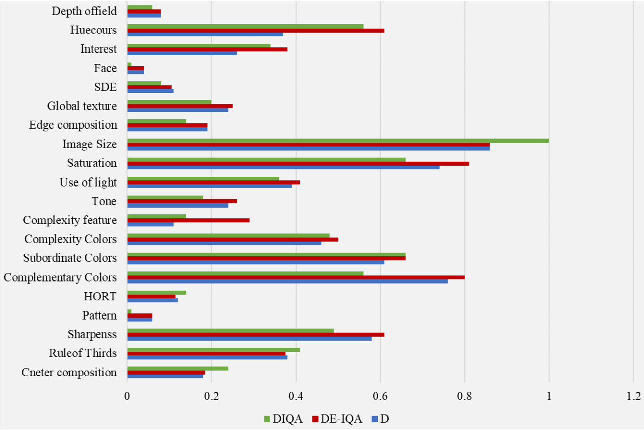
Model name	*ImageNet Top-1 accuracy*	*Number of parameters (in millions)*	*Core strengths*
EfficientNet-B7	84.3%	66	The highest accuracy and optimal parameter efficiency
EfficientNet-B0	77.3%	5.3	Lightweight and efficient, with low training costs
SENet	84.0%	145	High accuracy, but large number of parameters
ResNeXt-101	83.1%	83	Good accuracy, high computational cost
InceptionResNet-v2	80.4%	55	Balance accuracy and speed, but efficiency is average
NASNet-A	74.0%	5.3	Lightweight but lacks precision

4.2 Experimental analysis

The purpose of this experiment is to verify whether each module in the proposed algorithm makes a certain degree of contribution to its overall performance and to test the performance contributed by using each module individually. It can be seen from the table that the test results using visual saliency or aesthetic attention alone are better than the baseline performance, further demonstrating that both the visual saliency described above and the aesthetic attention extracted by the aesthetic classification model have contributed to image aesthetic evaluation tasks. In addition, it can be concluded from the table that the contribution of the aesthetic attention module to the algorithm's performance is higher than that of the visual saliency module to the overall algorithm performance in both the AADB database and the AVA database. It can be inferred that the aesthetic attention module provides information closer to the image aesthetics that the network needs to learn than the visual saliency.

From the results of the KL divergence, it can be seen that the aesthetic quality model, integrated with personalised image enhancement, optimises the results of personalised aesthetic evaluation and better represents the user's aesthetic preferences during the aesthetic process. For optimisation, this section selects a user's personalised aesthetic distribution under different experimental settings, as shown in Figure 5.

Figure 5 Aesthetic distribution models under different experimental settings (see online version for colours)



This horizontal bar chart compares the weight distribution of DIQA, DE-IQA, and D models on 19 image aesthetic features, clearly presenting the optimisation effect of personalised aesthetic models under different experimental settings. From the figure, it can be seen that in the feature of ‘image size’, the DIQA model has the highest weight, close to 1.2, far higher than the DE-IQA and D models, indicating that this model has the strongest aesthetic sensitivity to image size. In the ‘saturation’ and ‘complementary colours’ features, the weight of the DE-IQA model exceeds 0.8, significantly highlighting the core role of colour dimension in personalised aesthetics. Overall, the weight distribution of most aesthetic features in the DE-IQA model combined with image enhancement is more in line with users’ real aesthetic preferences. Compared to the other two models, its weight allocation in complex colours and composition features is more discriminative, which also confirms the advantage of this model in optimising personalised aesthetic evaluation.

By examining Figure 6, it becomes apparent that the BLG-PIAA model exhibits enhanced predictive capabilities as the values of k and S are raised. When the value of S is increased from 1 to 5, the prediction performance of the BLG-PIAA model improves significantly. When k exceeds 4, the BLG-PIAA model demonstrates a tendency towards stable prediction performance. The variable k denotes the number of tasks that are to be executed concurrently for gradient optimisation. The larger the value of k, the more significant the amount of training data in a batch. When the value of k is too large, achieving this is difficult in practice. Therefore, in the meta-training stage of the BLG-PIAA method proposed in this chapter, the four key parameters are set to $ms = 80$, $mq = 20$, $k = 4$, and $S = 5$, respectively.

The predicted performance (SROCC) is illustrated in Figure 7. It is evident that the prediction performance of both the personalised aesthetic evaluation models, namely BA-PIAA and BLG-PIAA, surpasses that of the aesthetic prior model. In particular, the mean prediction results (SROCC) of the BA-PIAA model and the BLG-PIAA model for 37 users increased by about 0.079 (from 0.504 to 0.583) and 0.129 (from 0.540 to 0.669), respectively, compared to the aesthetic prior model. In comparison to the BA-PIAA method, the BLG-PIAA method showcases enhanced efficacy in establishing an aesthetic prior model (0.540 versus 0.504). It exhibits superior capabilities in quickly adapting to individual aesthetic preferences (0.129 versus 0.079). This demonstrates that the meta-learning method, based on double-layer gradient optimisation proposed in this paper, can quickly and effectively learn personalised aesthetics tailored to specific users.

Figure 6 Performance comparison of the proposed method in different previous scenarios (see online version for colours)

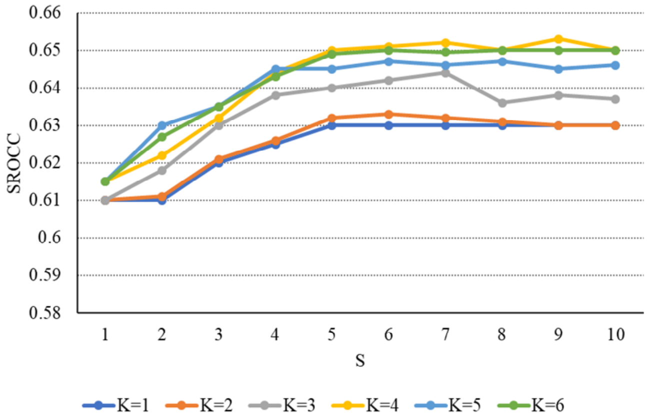
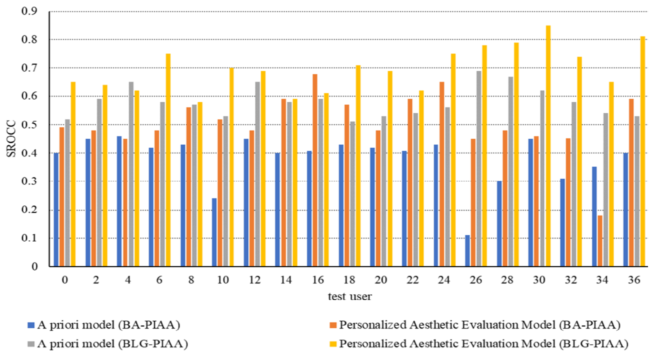


Figure 7 Direct use of aesthetic prior model and personalised aesthetic evaluation model (see online version for colours)



Compared with the BA-PIAA method, the BLG-PIAA method also shows higher efficiency in establishing aesthetic prior models, with an SROCC value of 0.540 for its aesthetic prior model, which is higher than the 0.504 of the BA-PIAA model. Moreover, in terms of quickly adapting to personal aesthetic preferences, the improvement of the BLG-PIAA model (0.129) is also greater than that of the BA-PIAA model (0.079), which further proves the superiority of the meta learning method based on double-layer gradient optimisation. This method can quickly and effectively learn personalised aesthetics tailored to specific users, which has important guiding significance for accurately grasping consumers’ aesthetic needs and designing art products that better meet market demands. Compared with Western art aesthetic evaluation methods that focus on quantitative indicators such as geometric proportions and colour contrast (such as composition analysis based on the golden ratio), this research framework emphasises the combination of visual salience and subjective aesthetic experience, making it more adaptable when dealing with Eastern literary products with implicit artistic conception; compared to Japan's qualitative evaluation system based on abstract concepts such as 'wabi sabi' and 'youxuan', this study achieves subjective aesthetic quantification expression by generating personalised aesthetic distributions, balancing the objectivity and cultural specificity of evaluation. Meanwhile, compared to some countries that rely on a single cultural sample evaluation model, this study relies on AVA (containing 250,000 images and 1–10 aesthetic scores) and AADB (focusing on Asian art images, including multicultural scenes) databases to demonstrate a wider applicability in cross-cultural art aesthetic evaluation.

5 Conclusions

This article validates the effectiveness of the proposed personalised image aesthetic evaluation algorithm using two image aesthetic databases, AADB and AVA, with ResNet and EfficientNet as the backbone networks. Experiments have shown that EfficientNet outperforms existing CNN networks (including ResNet) in TOP1 classification accuracy, and has more advantages in parameter quantity and computational complexity (FLOPS), achieving a balance between speed and accuracy, and reducing the resource consumption and cost of model training. The use of visual saliency or aesthetic attention modules alone outperforms baseline performance, and the contribution of aesthetic attention modules is more significant. The BLG-PIAA model has the best predictive performance when $k = 4$ and $S = 5$. Compared with the BA-PIAA model, its average SROCC is improved by 0.129 and 0.079 respectively, and BLG-PIAA is more efficient in establishing aesthetic priors and adapting to personal preferences. This algorithm combines visual saliency and subjective aesthetic experience to quantify subjective aesthetics. It relies on multiple databases and has cross-cultural applicability, which has important guiding significance for grasping consumer aesthetics and designing art products that meet market demand.

Although the BLG-PIAA model has shown good performance in experiments, it is built based on a specific algorithm framework and parameter settings. When faced with significant changes in new types of art products or aesthetic trends, models may need to undergo extensive retraining and parameter adjustments to maintain good performance. If the dataset contains more images of art products from specific regions or styles, the evaluation and enhancement effects of the model may decrease when facing art products from other regions or styles. In the future, the model can be improved in three aspects: firstly, expanding the cross-cultural aesthetic feature library to include more cultural and artistic product image data from countries and regions, optimising the adaptability of the model to multicultural aesthetic differences, and enabling it to accurately capture aesthetic preferences in different cultural backgrounds. The second is to introduce dynamic visual feature analysis, combined with the temporal changes of composition edges in video sequences and the dynamic transfer of visual saliency, to enhance the aesthetic evaluation ability of dynamic literary works (such as digital interactive design). The third is to integrate user physiological feedback data (such as eye movement trajectories and EEG signals), deeply associate objective features with subjective physiological perceptions, make the model more in line with human real aesthetic experience, and further enhance the accuracy and personalised expression of evaluation.

Authorship contribution statement

Xin Yu: supervision, conceptualisation, writing-original draft preparation, and project administration.

Declarations

All authors declare that they have no conflicts of interest.

References

Afrianti, T. and Zainul, R. (2021) 'E-learning development on basic chemical law materials in senior high school (SMA/MA) to improve high order thinking skill ability', *Journal of Physics: Conference Series*, Vol. 1783, No. 1, p.012128.

Alcaide-Marzal, J., Diego-Mas, J.A. and Acosta-Zazueta, G. (2020) 'A 3D shape generative method for aesthetic product design', *Design Studies*, Vol. 66, No. 1, pp.144–176.

Aslan, I., Seiderer, A., Dang, C.T., Rädler, S. and André, E. (2020) 'PiHearts: resonating experiences of self and others enabled by a tangible somaesthetic design', *Proceedings of the 2020 International Conference on Multimodal Interaction*, pp.433–441.

Bessghaier, N., Soui, M. and Ghaibi, N. (2022) 'Towards the automatic restructuring of structural aesthetic design of Android user interfaces', *Computer Standards & Interfaces*, Vol. 81, No. 1, p.103598.

Carroll, M.K., Anderson, A.M., Mangu, S.T., Hajjaj, Z. and Capron, M. (2022) 'Aesthetic aerogel window design for sustainable buildings', *Sustainability*, Vol. 14, No. 5, p.2887.

Gao, D., Sun, Y., Zhou, C., Bu, Y., Bao, Y. and Chai, J. (2020) 'Numerical and experimental study on a double-layered coating design using supplemental property particles for achieving user-desired thermal and aesthetic performance', *Energy*, Vol. 211, No. 1, p.118683.

Gatchalian, S.R., Agathis, N.T., Castillo-Carandang, N.T., Gunter, S.M., Murray, K.O. and Mandalakas, A.M. (2020) 'Design and evaluation of risk assessment tools to identify pediatric tuberculosis infection in Bohol, the Philippines, a low-HIV and high-TB-burden setting', *The American Journal of Tropical Medicine and Hygiene*, Vol. 103, No. 5, p.1818.

Grein, M., Wiecek, A. and Wentzel, D. (2021) 'It's a question of talent! The interplay of design complexity and talent information on consumers' product design responses', *European Journal of Marketing*, Vol. 55, No. 5, pp.1338–1358.

Hamid, H.M.A. and Celik-Butler, Z. (2021) 'Self-packaged, flexible, bendable MEMS sensors and energy harvesters', *IEEE Sensors Journal*, Vol. 21, No. 11, pp.12606–12617.

Kam, S. (2021) 'Three-dimensional printing fashion product design with emotional durability based on Korean aesthetics', *Sustainability*, Vol. 14, No. 1, p.240.

Keshari, A., Simeone, A. and Tiwari, M.K. (2021) 'Assembly design of additive manufacturing products: a computational framework for part separation', *Procedia CIRP*, Vol. 96, pp.121–126.

Landy, M.F.B., Sousa, S. and Romero, F. (2020) 'Service quality factors in the construction sector: a literature review', *IOP Conference Series: Materials Science and Engineering*, Vol. 800, No. 1, p.012035.

Naderi, E., Naderi, I. and Balakrishnan, B. (2020) 'Product design matters, but is it enough? Consumers' responses to product design and environment congruence', *Journal of Product & Brand Management*, Vol. 29, No. 7, pp.939–954.

Papadimitriou, D., Apostolopoulou, A., Branvold, S. and Gargalianos, D. (2021) 'Product meanings as drivers of sport consumer behavior: evidence from the Greek sport industry', *International Journal of Sports Marketing and Sponsorship*, Vol. 22, No. 3, pp.529–549.

Petrusch, N., Schliephack, W. and Kohl, H. (2020) 'Evaluation model for mobility design of learning factories', *Procedia CIRP*, Vol. 91, pp.659–664.

Rahman, O. and Koszewska, M. (2020) 'A study of consumer choice between sustainable and non-sustainable apparel cues in Poland', *Journal of Fashion Marketing and Management: An International Journal*, Vol. 24, No. 2, pp.213–234.

Schneider, D., Fröhlich, T., Huth, T. and Vietor, T. (2020) 'Design for flexibility – evaluation interactions between product properties and production processes', *Procedia CIRP*, Vol. 91, pp.814–818.

Seifert, C. and Chattaraman, V. (2020) 'A picture is worth a thousand words! How visual storytelling transforms the aesthetic experience of novel designs', *Journal of Product & Brand Management*, Vol. 29, No. 7, pp.913–926.

Talei, B. and Pearlman, S.J. (2022) 'CUPID lip lift: advanced lip design using the deep plane upper lip lift and simplified corner lift', *Aesthetic Surgery Journal*, Vol. 42, No. 12, pp.1357–1373.

Vieira, M.C. (2021) 'Design and tactile perception: a case study on the finishes of the ceramic tile textures', *Gestão & Produção*, Vol. 28, No. 4, p.e5013.