Emotionally Intelligent Fashion Design Using CNN and GAN

Cheng Yang[1], Yuliang Zhou[1,2], Bin Zhu[1,2], Chunyang Yu[1] and Lingang Wu[1,2]

[1]Department of Industrial Design, Zhejiang University City College
[2]Institute of Modern Industrial Design, Zhejiang University

Corresponding author: Yuliang Zhou, yuliang.zhou@hotmail.com

Abstract. In order to match consumers' emotional needs with the design of fashion goods, this paper proposes an emotionally intelligent design method to discover consumers' feelings about fashion goods, and based on that, modified CNN and GAN models are used to classify and generate product images satisfying consumers' emotional needs. The method consists of two parts: (1) a product image recognition model, which takes product image score as the weight to calculate the loss function. (2) an intelligent design generation model, which combines network architectures of DCGAN and Conditional GAN. The experimental results of shoe design demonstrate the feasibility and validity of the proposed method.

Keywords: Fashion Design, Product Image, Convolutional Neural Network, Generative Adversarial Network
DOI: https://doi.org/10.14733/cadaps.2021.900-913

1 INTRODUCTION

With the rapid development of science and technology, goods can be produced on a large scale. People's life is flooded with standardized industrial products, which leads to senses of loss, though the quality of these products is good. These highly standardized products bring us aesthetic experience of simplicity and order, but lack in nature and personalization. Therefore, in the life surrounded by modular products, people began to have an urgent desire for emotions, and more and more users paid more attention to the spiritual pleasure and emotional satisfaction of products, rather than just considering the basic functions and utilities. These abstract feelings are closely related to the product image, i.e., how people perceive and expect about the product. In other words, people can sense product stimulation through sensory organs and produce perceptions and recognition after processing information in brain, so as to receive the image of products [14-15]. Hence, product image is a scientific theory that concerns about the relationship between consumers and products, which helps to understand and explore consumers' emotional needs for ensuring a satisfactory product design. It is the basis of emotional product design. However, the product image in designers' minds usually differs from the expectation of consumers, which often misleads the final solution.

In order to solve such cognitive difference between designers and consumers in emotional product design, plenty of research has been carried out to establish a correct evaluation of product image to facilitate the communication between designers and users at an emotional level. Desmet et al. [2] added emotional values to the design of products on the basis of a theoretical framework and a nonverbal instrument to measure emotional responses. Guo et al. [7] proposed a systematic emotional design method based on the theory of Kansei Engineering. It applied the consumer-oriented technology into a Kansei engineering model, which reflects the relationship between Kansei image words and multi-dimensional key design variables. By expending the definition of value, Wang et al. added emotional response to the assessment of the value of visualization [20]. These are traditional methods that normally needs the statistical experiment to evaluate user emotions in specific scenarios. Thus, the process is usually time-consuming, and the results in most cases are not repeatable.

At the same time, various machine learning methods, such as genetic algorithm, support vector regression, clustering, and artificial neural network, are widely used in the field of product design to promote design efficiency. For instance, Dou et al. [4] applied multi-stage interactive genetic algorithm (MS-IGA) into the conceptual design system for car consoles to meet the personalized user needs. Using grey system theory (GST) and support vector regression (SVR), Wang [19] achieved to transform customers' needs into design elements. Jaber et al. [8] clustered actors according to the closeness of their relationships to facilitate the collaborative decision-making process. Bell and Bala [1] integrated visual search based on the technology of convolutional neural networks into interior design to identify products in scenes or find stylistically similar products. These are traditional machine learning algorithms requiring a complexly manual procedure for feature engineering (e.g. feature extraction and selection).

Evolved from artificial neural network, deep learning can automatically extract features from original data for modeling, avoiding the artificial intervention and data loss [12]. Kaiyrbekov and Sezgin proposed a new generative Variational Auto-Encoder (VAE) model to automatically segment hand-drawn objects into stroke-level components based on their semantics [9]. Dibia and Demiralp trained a multilayered attention-based encoder-decoder network with long short-term memory (LSTM) units to automatically generate data visualizations [3]. Fogel et al. suggested a Clustering-driven deep embedding for nonparametric clustering using a neural network [5].

This paper takes shoe as an example to explore the procedure of emotional intelligent design. Shoes, as a daily necessity worn by people, reflect the aesthetics and taste of the wearer. Although shoes have many attributes such as comfort and stability, the shape of the shoes often gives people the first impression. The shapes of different shoes contain tremendous emotional information, bringing people of different psychological feelings and emotional experience. Therefore, how to match the shoe shape with the user's emotional preference is the key to emotional product design.

To sum up, in the field of product design, so far there has been few applications using the technologies of deep learning to promote the intelligent design in accordance with the specific product image of users. For this reason, aiming at integrating consumer emotional needs into the practical product design, this paper attempts to propose a method for emotionally intelligent design that includes a couple of specific deep learning models for product image recognition and intelligent design generation respectively. The product image recognition model is set up based on the convolutional neural network to achieve an efficient and effective product image evaluation. The intelligent design generation model is established on the basis of GAN. By this means, either the designer or the user only needs to define the product image to automatically acquire a number of satisfactory product designs, which can significantly save the design effort. Finally, considering the familiarity and diversity of products, shoes as a necessity of daily life, were used to verify the effectiveness of the proposed method.

2 PRODUCT IMAGE RECOGNITION MODEL

Figure 1 shows the framework of the product image recognition model based on convolutional neural network. First of all, to build a dataset for product image evaluation, two major steps are required. One is to collect and select the representative product samples, while the other is to create the product image space and carry out the experiment of product image evaluation. After that, the features can be extracted by convolutional neural network, and the corresponding mapping relationship between product pictures and product images can be constructed. Finally, based on this relationship, the product images can be evaluated automatically.

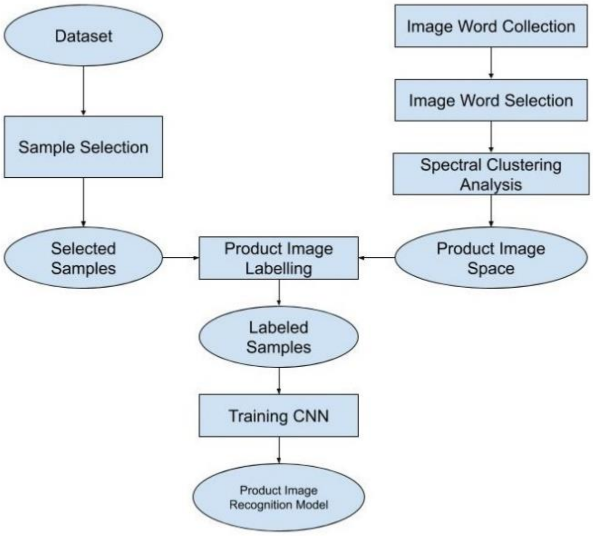

Figure 1: Framework of product image recognition model.

2.1 Construction of Product Image Space

2.1.1 Selection of product samples

The quality of sample selection will directly affect the subsequent experiment stages and final results. Hence, to reduce the noise of experimental samples and obtain relatively pure data, several principles should be satisfied. First, the product picture must be clear enough. Second, the background color should be standardized, like white. Third, samples should be displayed from the same perspective.

UT Zappos 50K [21] was chosen as the dataset, which contains 50,025 shoe pictures of various brands in the market. 1200 shoe pictures were randomly selected from the dataset as the experimental samples for product image recognition. Some of them are shown in Figure 2.

2.1.2 Screening product image words

Product images can be described intuitively using image word pairs. Referring to relevant literature, fashion magazines, e-commerce websites, and advertising proposals, 86 pairs of image words were selected. After preliminary screening (e.g. removing the pairs of words with repetitive meanings, such as "publicized-introverted" and "publicized-understated"), 63 pairs got left. These 63 pairs of words were further filtered by 30 university teachers and graduate students who majored in industrial design through questionnaire. Based on their daily experience, the 30 subjects were required to select the most suitable pairs of image words after observing the representative samples.

Those selected more than 15 times (more than half) were chosen as the final word pairs, in a total of 45 pairs.

Figure 2: Samples from UT Zappos 50K.

2.1.3 Establishing product image space

In order to further determine the relationship between different pairs of image words and thereby establish the product image space, 22 graduate students and teachers majoring in industrial design (subjects) were invited to participate in a clustering experiment. The subjects were asked to cluster the selected 45 pairs of image words based on their own experiences and judgments. There was no restriction on the quantities of groups and pairs of image words in each group. After that, the times of every two pairs of words that appeared in a same group were counted, thus obtaining a 45*45 similarity matrix. Obviously, in this matrix, the larger the value, the higher the similarity between the two pairs of image words.

Then, based on the similarity matrix, the 45 image word pairs were grouped into several clusters using spectral clustering [16]. Words in the same cluster have similar meaning. Furthermore, to determine the optimal number of clusters, the Sum of Squared Errors (SSE) was applied to evaluate the clustering performance. There is an obvious inflection point at 4 clusters, after which the decline trend becomes flat. Therefore, we considered 4 as the optimal number of clusters and divided the image words into four dimensions. The one nearest to the center of each cluster was selected as the representative word pair of each dimension, i.e., gorgeous-plain, modern-retro, casual-formal, and male-female. The whole procedure is shown in Figure 3.

2.2 Product Image Labeling

2.2.1 Subject selection

38 university teachers and graduate students majoring in industrial design were invited to participate in the image labeling experiment, including 17 males and 21 females aged from 23 to 40. Compared to the ordinary consumers without design background, professional designers tend to have a deeper understanding in product images and design styles, and thus can provide more accurate experimental data.

2.2.2 Questionnaire Design

As it is difficult for subjects to concentrate their focuses in a long time, the samples (product pictures) were divided into small groups to make sure each subject can finish the questionnaire in a short time, which ensures the data accuracy and effectiveness. For this reason, 1200 shoe samples were randomly classified into 12 groups corresponding to 12 questionnaires. Except for the experimental samples, other parts of the questionnaire were the same. As shown in Figure 4, the questionnaire applied semantic differential scale to evaluate the product image of each sample, where the left-most image word corresponded to -2, the right-most image word corresponded to 2, and the neutral was 0.

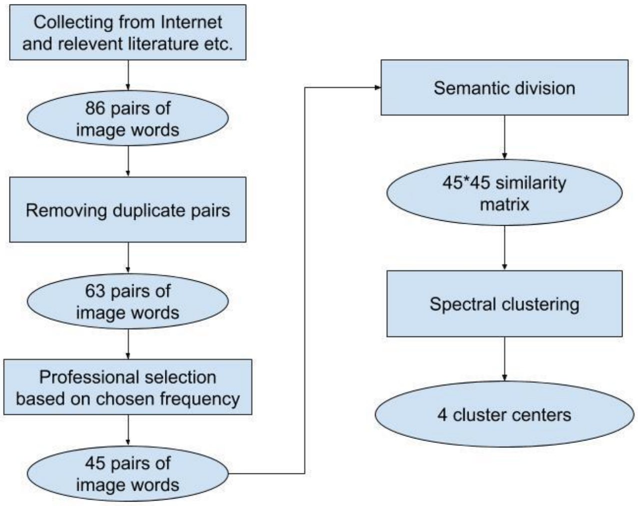

Figure 3: Construction of product image space.

Figure 4: Semantic Differential Scale of a Shoe Sample.

2.2.3 *Results of product image labeling*

Subjects can complete the product image evaluation of all samples within 30 minutes. In each dimension, the average score of each sample was computed to define its corresponding product image. Taking "casual-formal" as an example, if the average score of the sample is less than -1, then the sample would be labeled as "casual", while if it is greater than 1, then the sample would be "formal". Otherwise, the sample was neutral and would be removed from this dimension because of the lack of image characteristics. Table 1 shows the number of samples under each product image, in which positive samples represent the samples corresponding to the right word of the image word pairs, and negative samples represent the samples related to the left word. As can be seen from Table 1, except for the dimension of "casual-formal", the distributions of positive and negative samples in others are basically balanced.

Product Image	***Positive***	***Negative***	***Total***
Male-Female	426	503	929

Modern-Retro	525	398	923
Casual-Formal	778	200	978
Gorgeous-Plain	417	509	926

Table 1: Positive and negative sample distributions of product images.

2.3 Data Preprocessing

The samples in UT Zappos 50K dataset required to be preprocessed for the convenience of operation in convolutional neural network, i.e., the sample pictures should be turned into square. To realize this, the original sample picture, in a size of 136*102 and with a white background, was added with white paddings to avoid the distortion.

The preprocessed samples were randomly divided into a training set (including 70% samples) and a testing set (including the rest 30% samples). Furthermore, the mean values of the R, G, B channels of the training samples were calculated and then subtracted to obtain a set of zero-mean training samples. In addition, data enhancement was used to prevent overfitting. Pictures were horizontally flipped and added to the training set to improve the robustness of the model. Finally, the training sample was enlarged to a size of 224*224 as the input to the convolutional neural network.

2.4 Implementation

2.4.1 Network architecture

VGGNet [18] is good at extracting image features, as it can replace large convolution kernels (5×5 or 7×7) with small ones (3×3) by the way of stacking and combination, which reduces the parameters and improves the training effect. For this reason, this paper adjusts the VGGNet-16 [18] to fulfill the image recognition of shoe products. The network structure is shown in Figure 5.

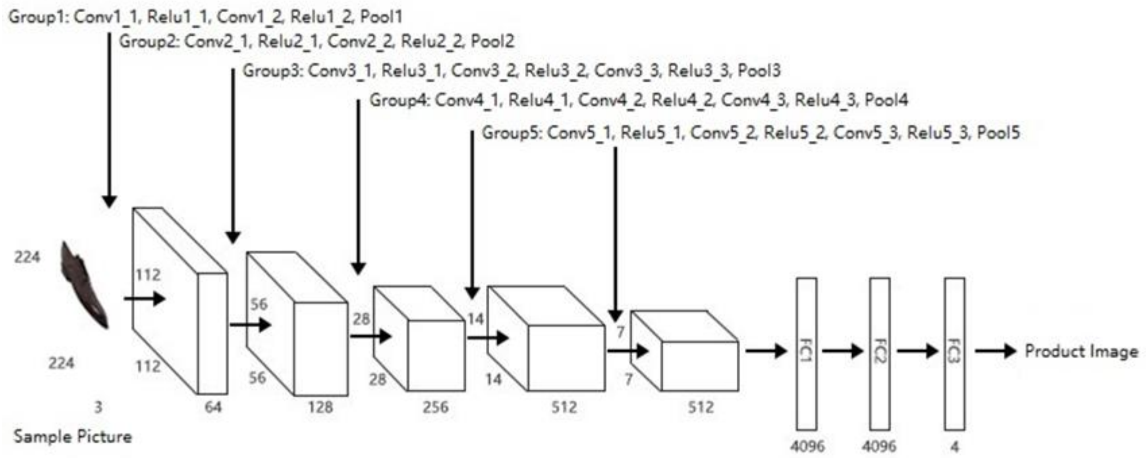

Figure 5: Network structure of product image recognition model.

The network consists of 13 convolutional layers and 3 fully connected layers. The convolutional layers are divided into five groups, and each group is followed by a pooling layer which compresses image features and simplifies the computational complexity of the network. Image features are extracted through these convolutional layers.

The last three layers are fully connected layers. The first two share the identical settings with that of VGG-16, while the neurons of the last layer is changed to 4 corresponding to the 4 dimensions of product images. In the end, a softmax function is adopted to map the output to a probability distribution over predicted product images, ranging from 0 to 1. Rather than standard cross entropy, we use weighted cross entropy in the loss function:

$$loss = \frac{\sum_{i=1}^{n} \alpha_i CE_i}{\sum_{i=1}^{n} \alpha_i} \tag{1}$$

where n is the quantity of training samples, α_i is the average image score given by the professionals, CE_i is the cross entropy of the ith sample. Introducing average image score to the loss function is to help the model converge faster and improve the accuracy, because training samples with higher product image score have more obvious characteristics.

2.4.2 Training details

We apply two techniques to stable the training process. First, we use weight decay to prevent the model from overfitting. L2 regularization $\lambda \sum w_i^2$ is added to the loss function, where the decay rate λ is set as 0.0005. Second, parameters in the network are updated using SGD with momentum. The learning rate is set as 0.0001, which gradually declines as training progresses (i.e., multiplied by 0.1 after every 100 epochs) and momentum value is set as 0.9.

In the experiment, we did not classify the category of "neutral" for product image, although many products can be classified in this category. This is because that the image style of "neutral" is not obvious and "neutral" often represents mediocre, however the product design requires innovation. We hope to teach computers to design brightly, doing something creative instead of repetitive boring work.

In the product image recognition model, we visualized the hidden convolutional layers to trace the machine learning patterns by feature maps. It shows the feature extraction process of the recognition model, which is the way computers understanding the design styles (seen in Figure 6). There is a trend can be traced that the features are transformed from concrete to abstract. For instance, in conv2_1, conv3_1 and conv4_1, the shapes and edges of the shoe are clearer, while in conv5_1 and conv5_3, there is almost none concrete features but the abstract one. This is similar to the hierarchical structure of the human visual system — humans use eyes to observe into the high-level semantic information through the transmission of neurons. Therefore, by learning representations from original dataset through convolutional layers, the product image recognition model is able to identify the product image and understand the design style.

2.5 Results

In order to explore whether introducing product image score as the weight into the loss function helps improve recognition accuracy and training efficiency, a comparative experiment was conducted between the standard loss and weighted loss. Table 2 shows test accuracies of two loss functions in the experiment of product image recognition after 50 epochs of training. The overall accuracies of weighted loss are better, which may be due to the weighted loss function converges faster than the standard loss function. To verify it, we conducted a longer comparative experiment with 100 epochs as shown in Table 3. The results of weighted loss and standard loss are close, which means after enough training, both loss functions perform equally well. Affected by overfitting, the accuracies of weight loss after 100 epochs are slightly lower than those after 50 epochs.

Product Image	***weighted loss***	***standard loss***
Male-Female	85.7%	72.6%
Modern-Retro	79.4%	76.2%
Casual-Formal	82.6%	75.4%
Gorgeous-Plain	75.9%	70.4%

Table 2: Test accuracies of two loss functions after 50 epochs of training.

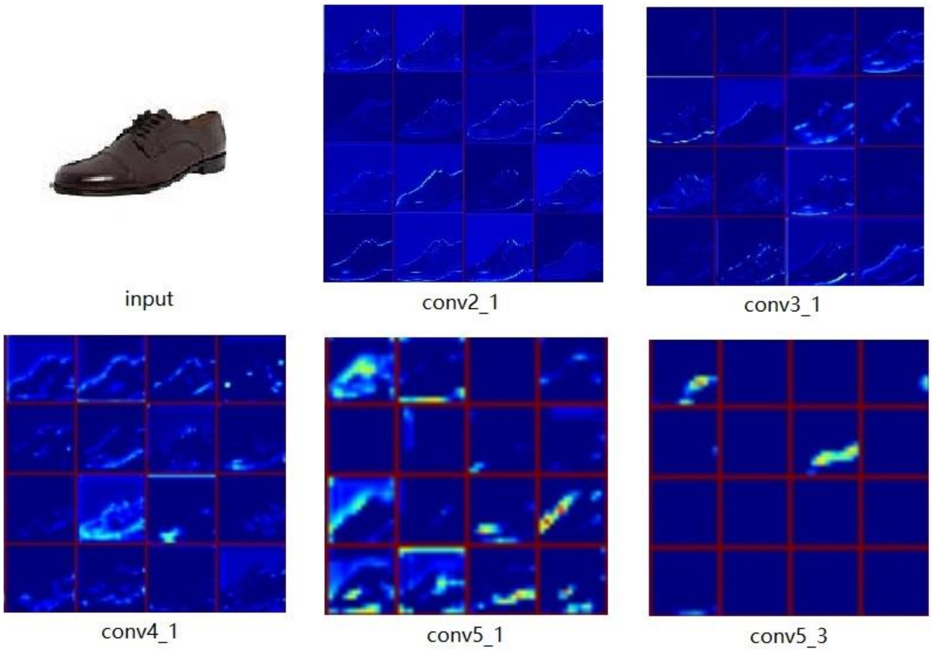

Figure 6: Visualization of the hidden convolutional layers.

Product Image	***weighted loss***	***standard loss***
Male-Female	80.2%	78.5%
Modern-Retro	75.4%	76.1%
Casual-Formal	79.6%	77.4%
Gorgeous-Plain	72.7%	73.1%

Table 3: Test accuracies of two loss functions after 100 epochs of training.

3 INTELLIGENT DESIGN GENERATION MODEL

Generative Adversarial Network (GAN) [6] has not received enough attention in the area of product design. Based on the product image recognition model (in Section 2), this paper attempts to make a preliminary exploration of using GAN to design products, and the technology of Conditional GAN [13] is applied into the intelligent design generation model which can intelligently generate innovative product designs according to the given product image, so as to meet consumer emotional requirements and save costs for enterprises.

As deep learning is applied into the intelligent design generation model, it needs a large number of training samples attached with product image labels. Thus, labeling samples in manual is inadequate to this environment. However, the product image recognition model trained in Section 2 is able to automatically mark the unlabeled samples meets the requirement.

3.1 Data Preprocessing

The small size of dataset in the experiment of design generation may easily lead to the lack of diversity in the generated designs. The labeled pictures in the dataset built in Section 2 was far from enough. Therefore, we needed to extend the dataset. Hence, 10240 shoes were randomly selected from UT Zappos 50K, and then were automatically labeled by the product image recognition model trained in section 2 to produce a much larger dataset. The quantities of samples with different product images are shown in Table 4. Positive samples represent samples on the right side of semantic differential scale and the negative for the left.

Product Image	***Positive***	***Negative***	***Total***
Male-Female	*4807*	*5433*	*10240*
Modern-Retro	*5844*	*4396*	*10240*
Casual-Formal	*7689*	*2551*	*10240*
Gorgeous-Plain	*4995*	*5245*	*10240*

Table 4: Distributions of different product image samples.

3.2 Implementation

3.2.1 Network architecture

Intelligent design generation model combines the network structures of DCGAN and conditional GAN, consisting of two parts: a generator and a discriminator. For the generator, the input is a 100-dimensional random noise concatenated with a 4-dimensional one-hot encoding of the product image, and the output is a generated product picture. For the discriminator, the input is a real product picture or a generated product picture, and the output is a probability value between 0 and 1, judging whether the input is real or fake. Product images were used as guiding conditions to generate product designs that meet users' emotional requirements. The generator and the discriminator were trained alternately until the model converges.

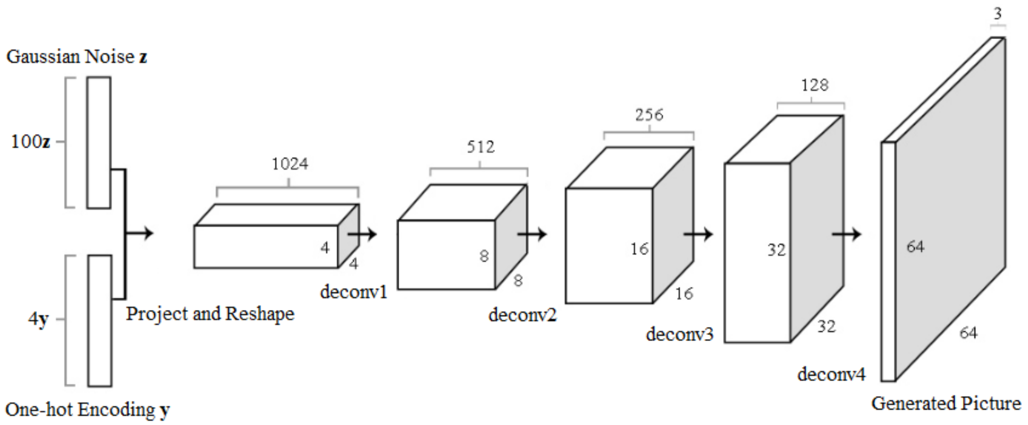

Figure 7: Network structure of the generator.

Figure 7 shows the network structure of the generator, which consists of one input layer, one project and reshape layer and four deconvolutional layers. The input layer consists of a 100-dimensional

Gaussian noise z concatenated with a one-hot encoding y of the product image. The dimension of y depends on how many product images are there. For convenience, we selected "male-female" and "casual-formal" for the design generation experiment. Hence, the dimension of y is 4 and the one-hot encoding is shown in Table 3. There are two reasons for only selecting two image dimensions. On one hand, a product rarely has obvious features across 3 or 4 dimensions, and image words usually have redundancy and repetition. On the other hand, most users are accustomed to picking one or two adjectives to describe the product style, seldom using as many as three words, e.g., "male casual gorgeous retro shoes". The 104-dimensional input layer is projected and reshaped to the dimension of 4*4*1024, and then is followed by four deconvolutional layers whose dimensions are 8*8*512, 16*16*256, 32*32*128, 64*64*3 respectively. The output layer of the generative model is the product picture.

Male Casual	***Male Formal***	***Female Casual***	***Female Formal***
0001	0010	0100	1000

Table 3: One-hot encoding of 'Male-Female' and 'Casual-Formal.'

The structure of the discriminator is shown in Figure 8, which consists of an input layer, four convolutional layers and a fully connected layer. The input layer contains a generated picture or a real picture, followed by four convolutional layers whose dimensions are 32*32*64, 16*16*128, 8*8*256, 4*4*512 respectively. The output layer generates the probability of truth.

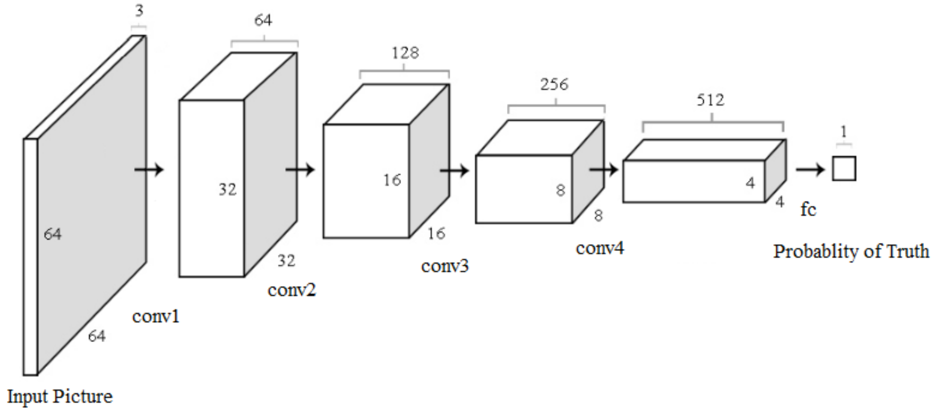

Figure 8: Network structure of the discriminator.

3.2.2 Training details

In the training process, the generator and discriminator were updated alternately. During each iteration, the generator was fixed at first and the discriminator updated twice, and then the discriminator was fixed and the generator updated once. The batch size was set as 32. The training iterations in an epoch was 320, and the model was trained 100 epochs. All parameters in the generator and discriminator were initialized with a zero-mean Gaussian distribution whose standard deviation was 0.02.

In order to accelerate the training of the model, parameters were updated by the Adam optimizer [10] with a learning rate of 0.002. The parameters β_1 and β_2 in the Adam optimizer were set to 0.5 and 0.999. The slope of Leaky ReLU in the discriminator was set to 0.2.

We extracted some generated pictures during training, as shown in Figure 9. Figure 9(a), 9(b), 9(c) are the generated pictures after 1 epoch, 10 epochs and 100 epochs respectively. As it shows, in Figure 9(a), the shoes are blur without clear outlines. In Figure 9(b), the outlines become even clearer while still few colors, mainly consisting of black ones. In Figure 9(c), the shoes have more details and colors. Obviously, with the increasing of training time, the generated pictures tend to have higher qualities. This is similar to the process of drawing — people sketch out the outlines first, and then draw the details, and finally color the paintings. In contrast to the recognition model, the generation model can create concrete designs from the abstract product image, as it applies deconvolutional layers.

Figure 9: Generated pictures during training.

3.3 Results

3.3.1 Quality evaluation

The model was trained for 100 epochs, and the experimental results are shown in Figure 10, illustrating some generated samples under the product image of "male-female" and "casual-formal", also including some bad samples.

The experiment shows that the Conditional GAN can well generate shoe designs under the restriction of product images, and the generated designs have a diversity of styles. Whether the generated design is an effective one was also evaluated by professional designers based on their knowledge, and there were a small number of bad results (some of them shown in Figure 10). According to the statistics of the last two training epochs (2000 randomly selected samples), the proportion of bad results was about 4.1%. There are mainly three types of bad results: failure to form the shape of shoes, some parts of the shoe disengaged from the main body, and too much noise in the picture. We conclude two reasons for the failure: first, the color of shoes is too complex which changes rapidly in local part; second, some sandals have very slim laces, and the connection between the lace and the main body is not clear. Although these two points lead to some difficulties, most design samples were generated well, and the texture and the shape can be sensed intuitively.

3.3.2 Accuracy evaluation

To verify whether the generated design samples are consistent with the target product images, the results were evaluated through a questionnaire survey. Twenty graduate students and university teachers majoring in industrial design, aged from 23 to 40, were invited as subjects, consisting of 7 males and 13 females.

The product images of "male casual", "male formal", "female casual" and "female formal" were used as input to generate 320 samples. After removing the bad results (4.1%), 25 samples were selected for each product image.

Figure 10: Generated design samples of "male-female" and "casual-formal".

Based on their design experience, the subjects chose the most suitable product image (including non-conformance) for each generated sample. To avoid the personal tendencies of subjects, if more than 80% of subjects (i.e. more than 16 people) made the right choice, the generated sample is considered to be consistent to the specified product image. The results are shown in Table 5.

Furthermore, subjects were interviewed about the invalid generated samples (inconsistent with the subjects' product image) after the experiment. Taking the generated "female casual" samples as an example, there were two invalid experimental samples. For these two samples, only half of people (8 and 11 respectively) considered that they were 'female casual'. The others, (9 and 12 respectively) thought them as non-conformance. In the interviews, the subjects who selected non-conformance considered the two design samples are "neutral casual". In this case, the product image of the generated sample is ambiguous and controversial, thus considered invalid.

Product Image	***Male Casual***	***Male Formal***	***Female Casual***	***Female Formal***	***Total***
Consistent Number	*23*	*24*	*23*	*21*	*91*
Effective Proportion	*92%*	*96%*	*92%*	*84%*	*91%*

Table 5: Validation results of 'Male-Female' and 'Casual-Formal'

4 DISCUSSION

4.1 Applicability

Emotionally intelligent design model is not limited to the fashion field, but applies equally well to other product fields with different product image dimensions. Taking the cases of cars, first a large number of car pictures need to be collected, and then users are invited to evaluate the product images to build the emotionally intelligent design model of cars. The product dimensions for cars may be divided into "modern-retro", "fashion-traditional" and so on. For the case that products in the same category differ greatly, such as airplanes are divided into gliders, helicopters, airliners, etc, emotionally intelligent design can be carried out for sub-categories.

Emotionally intelligent design model can create sufficient design solutions according to the target product image and style. It can maintain the brand style and save design costs for enterprises. Designers can get inspirations from the generated designs and continue to improve them, which reduces design effort and increases working efficiency. At the same time, it also broadens the horizon of designers, and alleviates the negative influence of cognitive difference between designers and users in an intelligent manner.

Emotionally intelligent design model is also better for users to engage in design processes. For example, users can actively define the appearance style of the generated solution according to their own preferences by choosing different combinations of product images. Based on the numerous generated solutions, users can choose their favorite ones for further processing (by designers). The user engagement in design incorporates user preferences into products and promotes the efficiency in customized design.

4.2 Limitations

Although intelligent design generation model can quickly generate a large number of different types of images under human control, amounts of high-quality labeled pictures are required for supervised learning. All those training images should be taken from a similar perspective and labeled manually, which would cost a lot of effort. Product image recognition model can alleviate this problem, but it has not been fundamentally solved. The recognition model still needs labeled images for training. Hence, unsupervised learning tends to be applied in product image recognition in further research.

5 CONCLUSION

This paper proposes an emotionally intelligent design method based on deep learning, to meet consumers' emotional needs and enterprises' demands for rapid product iteration. Product image recognition model in support of convolutional neural network bridges the designers and consumers to share the identical product images. The intelligent design generation model on the basis of generative adversarial network provides a smart method to generating innovative product designs. Taking shoe products as an example, the feasibility and validity of the model are demonstrated.

In the future work, we are going to increase the creativity of generated solutions and bring more versatile designs to inspire designers, instead of being confined to the original training data mode in this paper. In addition, as vivid pictures can provide better visual effects, it is also important to enhance the resolution of generated pictures.

ACKNOWLEDGEMENTS

This research was supported by the National Natural Science Foundation of China (No. 61772468), and Zhejiang Provincial Natural Science Foundation of China (No. Y18E050014).

Cheng Yang, https://orcid.org/0000-0002-1783-4760
Yuliang Zhou, https://orcid.org/0000-0002-7390-6597
Bin Zhu, https://orcid.org/0000-0002-9213-2611
Chunyang Yu, https://orcid.org/0000-0001-8864-0106
Lingang Wu, https://orcid.org/0000-0002-2301-9199

REFERENCES

[1] Bell, S.; Bala, K.: Learning visual similarity for product design with convolutional neural networks, ACM Transactions on Graphics (TOG), 34(4), 2015, 98. https://doi.org/10.1145/2766959

[2] Desmet, P.; Overbeeke, K.; Tax, S.: Designing products with added emotional value: Development and appllcation of an approach for research through design, The design journal, 4(1), 2001, 32-47. https://doi.org/10.2752/146069201789378496

[3] Dibia, V.; Demiralp, Ç.: Data2vis: Automatic generation of data visualizations using sequence-to-sequence recurrent neural networks, IEEE computer graphics and applications, 39(5), 2019, 33-46. https://doi.org/10.1109/MCG.2019.2924636

[4] Dou, R.; Zong, C.; Nan, G.: Multi-stage interactive genetic algorithm for collaborative product customization, Knowledge-Based Systems, 92, 2016, 43-54. https://doi.org/10.1016/j.knosys.2015.10.013

[5] Fogel, S.; Averbuch-Elor, H.; Cohen-Or, D.; Goldberger, J: Clustering-driven deep embedding with pairwise constraint, IEEE computer graphics and applications, 39(4), 2019, 16-27.

[6] Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y.: Generative adversarial nets, Advances in neural information processing systems, 2014, 2672-2680.

[7] Guo, F.; Liu, W. L.; Liu, F. T.; Wang, H.; Wang, T. B.: Emotional design method of product presented in multi-dimensional variables based on Kansei Engineering, Journal of Engineering Design, 25(4-6), 2014, 194-212. https://doi.org/10.1080/09544828.2014.944488

[8] Jaber, H., Marle, F., & Jankovic, M.: Improving collaborative decision making in new product development projects using clustering algorithms, IEEE Transactions on Engineering Management, 62(4), 2015, 475-483. https://doi.org/10.1109/TEM.2015.2458332

[9] Kaiyrbekov, K.; Sezgin, M.: Stroke-based sketched symbol reconstruction and segmentation, arXiv preprint arXiv:1901.03427, 2019.

[10] Kingma, D. P.; Ba, J.: Adam: A method for stochastic optimization, arXiv preprint arXiv:1412.6980, 2014.

[11] Krizhevsky, A.; Sutskever, I.; Hinton, G. E.: Imagenet classification with deep convolutional neural networks, Paper presented at the Advances in neural information processing systems, 2012.

[12] LeCun, Y.; Bengio, Y.; Hinton: Deep learning, nature, 521(7553), 2015, 436-444. https://doi.org/10.1038/nature14539

[13] Mirza, M.; Osindero, S.: Conditional generative adversarial nets, arXiv preprint arXiv:1411.1784, 2014.

[14] Nagamachi, M.: Kansei engineering as a powerful consumer-oriented technology for product development, Applied ergonomics, 33(3), 2002, 289-294. https://doi.org/10.1016/S0003-6870(02)00019-4

[15] Nagamachi, M.: Kansei engineering: a new ergonomic consumer-oriented technology for product development, International Journal of industrial ergonomics, 15(1), 1995, 3-11. https://doi.org/10.1016/0169-8141(94)00052-5

[16] Ng, A. Y.; Jordan, M. I.; Weiss, Y.: On spectral clustering: Analysis and an algorithm, Paper presented at the Advances in neural information processing systems, 2002.

[17] Radford, A.; Metz, L.; Chintala, S: Unsupervised representation learning with deep convolutional generative adversarial networks, arXiv preprint arXiv:1511.06434, 2015.

[18] Simonyan, K.; Zisserman, A: Very deep convolutional networks for large-scale image recognition, arXiv preprint arXiv:1409.1556, 2014.

[19] Wang, K.-C.: A hybrid Kansei engineering design expert syst0em based on grey system theory and support vector regression, Expert Systems with Applications, 38(7), 2011, 8738-8750. https://doi.org/10.1016/j.eswa.2011.01.083

[20] Wang, Y.; Segal, A.; Klatzky, R.; Keefe, D. F.; Isenberg, P.; Hurtienne, J.; Hornecker, e.; Dwyer, T.; Barrass, S.: An emotional response to the value of visualization, IEEE computer graphics and applications, 39(5), 2019, 8-17. https://doi.org/10.1109/MCG.2019.2923483

[21] Yu, A.; Grauman, K.: Fine-grained visual comparisons with local learning, Paper presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014. https://doi.org/10.1109/CVPR.2014.32