学校代号 10731　　　　学　　号 171080201043

分 类 号 TB472　　　　密　　级 公开

LANZHOU UNIVERSITY OF TECHNOLOGY

博士科研成果

学位申请人姓名 李　雄

培　养　单　位 机电工程学院

导师姓名及职称 苏建宁　教授

学　科　专　业 机械制造及其自动化

研　究　方　向 智能设计理论与方法

科研成果目录

1、SCI 论文：Product innovation concept generation based on deep learning and Kansei engineering

2、EI 论文：基于深度学习的产品概念草图生成设计研究

3、EI 论文：特征迁移的细粒度产品形态智能决策方法

4、EI 论文：基于深度学习的产品风格精细识别研究

5、中文核心论文：产品意象造型设计应用研究进展

6、EI 国际会议论文：Research on parametric form design based on natural patterns

7、发明专利：基于自然模式的 Logo 生成方法、智能 Logo 生成器

8、专著：工业产品设计草图

兰州理工大学博士研究生科研情况一览表

学院：机电工程学院　　姓名：李 雄　　学号：171080201043　　2023年8月13日

发表学术论文情况				
发表日期	论文题目	刊物名称/刊号	本人排名	刊物级别
2021	Product innovation concept generation based on deep learning and Kansei engineering	Journal of Engineering Design ISSN：0954-4828	1/4	SCI
2023	基于深度学习的产品概念草图生成设计研究	机械工程学报 ISSN：0577-6686	1/3	EI（网络版）
2023	特征迁移的细粒度产品形态智能决策方法	计算机辅助设计与图形学学报 ISSN：1003-9775	1/4	EI（网络版）
2023	基于深度学习的产品风格精细识别研究	计算机集成制造系统 ISSN：1006-5911	1/5	EI（网络版）
2019	产品意象造型设计应用研究进展	包装工程 ISSN：1001-3563	1/6	中文核心期刊
2018	Research on parametric form design based on natural patterns	2018th International Form on Industrial Design (IFID 2018)	1/2	EI（国际会议）
合计：6 篇　本人签名：李雄　导师签名：[signature]　主管院长签名：[signature]				

申请发明专利情况				
授权日期	专利名称	专利号	本人排名	专利级别
2021	基于自然模式的 Logo 生成方法、智能 Logo 生成器	ZL201810343073.9	2/3	发明专利
合计：1 项　本人签名：李雄　导师签名：[signature]　主管院长签名：[signature]				

注：1、仅列出本人第一或者导师第一、本人第二的学术论文和科研成果。

2、刊物级别分为：SCI、SSCI、A&HCI、EI（注明光盘版或网络版）、CPCIE、CSCD（注明核心版或扩展版）、CSSCI、中文核心期刊、普通 CN 刊物。

3、奖励级别分为：国家级、省部级、厅局和省级及以上学会级、校级（需写明获奖的具体等级，如省部级三等奖）。

4、专利分为：发明专利、实用新型专利和外观设计专利。

5、双面打印，仅限一页。

6、一式 25 份，一份原件，其他可复印。

报告编号：2023-YZ-5197

论 著 检 索 报 告

委托内容： 李雄的论文收录与引用情况

委托单位： 兰州理工大学

检索单位： 中国科学院兰州查新咨询中心

完成日期： 2023 年 8 月 16 日

中 国 科 学 院 兰 州 查 新 咨 询 中 心

二〇二二年制

一、检索要求

1、被检作者：李雄（Li Xiong；Li X）

2、委托单位：兰州理工大学机电工程学院

3、论文发表年限：2018-2023 年

5、提供待检索论文篇数：5 篇

二、检索范围

Science Citation Index Expanded (SCI-EXPANDED)	1900 - present	网络版
Engineering Village 2-Compendex (EI)	1884 - present	网络版
科睿唯安 Journal Citation Reports(JCR)	2022 年	网络版
中国科学院文献情报中心期刊分区表（升级版）	2022 年	网络版
中国学术期刊网络出版总库(CNKI)	1979 - present	网络版

三、检索结果

1）收录：

- 有 1 篇被 SCI 收录（记录见附件一）；
- 有 1 篇被 EI 收录（记录见附件二）；
- 有 2 篇被 CNKI 收录（记录见附件三）；
- 有 1 篇专利已被中国知识产权局公开并授权（记录见附件四）。

检索人：白丽　　检索单位：中国科学院兰州查新咨询中心

审核人：任珩　　完成时间：2023 年 8 月 16 日

附件：

一、SCI 收录情况

1. 收录论文题录

Record 1 of 1
Title: Product innovation concept generation based on deep learning and Kansei engineering
Author(s): Li, X (Li, Xiong); Su, JN (Su, Jianning); Zhang, ZP (Zhang, Zhipeng); Bai, RS (Bai, Ruisheng)
Source: JOURNAL OF ENGINEERING DESIGN Volume: 32 Issue: 10 Pages: 559-589 DOI: 10.1080/09544828.2021.1928023 Early Access Date: MAY 2021 Published: OCT 3 2021
Times Cited in Web of Science Core Collection: 15
Total Times Cited: 16
Accession Number: WOS:000657539500001
Language: English
Document Type: Article
Addresses: [Li, Xiong; Bai, Ruisheng] Lanzhou Univ Technol, Sch Mech & Elect Engn, Lanzhou, Peoples R China.
[Li, Xiong] Lanzhou City Univ, Sch Bailie Mech Engn, Lanzhou, Peoples R China.
[Su, Jianning; Zhang, Zhipeng] Lanzhou Univ Technol, Sch Design Art, Lanzhou, Peoples R China.
Corresponding Address: Su, JN (corresponding author), Lanzhou Univ Technol, Sch Design Art, Lanzhou, Peoples R China.
E-mail Addresses: sujn@lut.cn
Web of Science Index: Science Citation Index Expanded (SCI-EXPANDED); Social Science Citation Index (SSCI)
IDS Number: UX4XE
ISSN: 0954-4828
eISSN: 1466-1837
Output Date: 2023-08-15
End of File
影响因子（2022 年）：2.7
科睿唯安 JCR 分区（2022 年）：

JOURNAL OF ENGINEERING DESIGN

Publisher name: TAYLOR & FRANCIS LTD

Journal Impact Factor ™

2.7	3.1
2022	Five Year

JCR Category	Category Rank	Category Quartile
ENGINEERING, MULTIDISCIPLINARY *in SCIE edition*	42/90	Q2

Source: Journal Citation Reports 2022. Learn more

中国科学院文献情报中心期刊分区（升级版）（2022 年）：

JOURNAL OF ENGINEERING DESIGN

刊名	JOURNAL OF ENGINEERING DESIGN
年份	2022
ISSN	0954-4828
Review	否
Open Access	否
Web of Science	SCIE

	学科	分区	Top期刊
大类	工程技术	3	否
小类	ENGINEERING, MANUFACTURING 工程：制造	3	-
	ENGINEERING, INDUSTRIAL 工程：工业	4	

二、EI 收录情况

<RECORD 1>

Accession number:20184706119642
Title:Research on Parametric Form Design Based on Natural Patterns (Open Access)
Authors:Li, X (1); Su, Jn (2)
Author affiliation:(1) School of Mechanical and Electronical Engineering, Lanzhou University of Technology, Lanzhou Gansu; 730050, China; (2) School of Design Art, Lanzhou University of Technology, Lanzhou Gansu; 730050, China
Corresponding author:Su, Jn(sujn@lut.cn)
Source title:MATEC Web of Conferences
Abbreviated source title:MATEC Web Conf.
Volume:176
Part number:1 of 1
Issue title:2018 6th International Forum on Industrial Design, IFID 2018
Issue date:July 2, 2018
Publication year:2018
Article number:01012
Language:English
ISSN:22747214
E-ISSN:2261236X
Document type:Conference article (CA)
Conference name:6th International Forum on Industrial Design, IFID 2018
Conference date:May 18, 2018 - May 20, 2018
Conference location:Luoyang, China
Conference code:141841
Publisher:EDP Sciences
Abstract:Parametric form design method based on natural patterns is proposed for the design schema of the traditional form bionics. Firstly, the concepts of natural patterns are analyzed, concluded and summarized. Secondly, the parametric design thinking is elaborated from three aspects: thinking mode, design flow, tools and scripts. It is also proposes a parametric logo design method. This paper takes logarithmic spiral pattern as an example to describe the process from the law of natural pattern to logo design, which includes nature pattern analysis, control rules of from and color, algorithm research and generating design.
 © The Authors, published by EDP Sciences, 2018.
Number of references:20
Main heading:Design
Controlled terms:Biomimetics
Uncontrolled terms:Algorithm researches - Control rules - Design flows - Logarithmic spiral - Parametric design - Parametric forms - Pattern analysis - Thinking modes
Classification code:461.8 Biotechnology - 461.9 Biology
DOI:10.1051/matecconf/201817601012
Funding details: Number: 51465037, Acronym: NSFC, Sponsor: National Natural Science Foundation of China;Number: 201406, Acronym: LUT, Sponsor: Lanzhou University of Technology;
Funding text:The project is sponsored by National Natural Science Foundation of China (51465037) and

Hongliu Outstanding Talent Development Program of Lanzhou University of Technology (201406). In addition, Zbloom Culture Media also provided support for this project.

Database:Compendex

Open Access type(s): All Open Access, Gold, Green

三、CNKI 收录情况

DataType: 1
Title-题名: 基于深度学习的产品概念草图生成设计研究
Author-作者: 李雄;苏建宁;张志鹏;
Source-刊名: 机械工程学报
Year-年: 2023
PubTime-出版时间: 2023-02-01 11:39
Keyword-关键词: 视觉认知;深度学习;生成设计;草图设计;生成对抗网络
Summary-摘要: 概念草图设计作为人类的高阶视觉认知活动，是辅助设计师记录、构思、创造和评估想法的重要手段，对生成创新概念具有积极影响。为模拟设计师的这种高阶视觉认知行为，实现智能化辅助创意草图设计，提出一种基于深度学习的产品概念草图智能生成设计集成方法框架，包括端到端的草图设计GAN(Sketch2Render-GAN)和草图神经风格迁移网络(Sketch-NST)两个核心模块。前者实现概念草图生成与渲染，后者执行草图风格特征变换。分别以手电钻和自行车头盔为实验对象进行了验证，结果表明该方法框架可快速获得大量具有创新概念的草图，并实现草图自动渲染及风格变换。有助于辅助设计师在视觉认知层面突破设计固化，提高设计效率。此外，为改善工业设计师与 AI 模型间的人机设计协作模式，还开发了智能草图设计生成器(S-SDG__v0.1)，从而有效降低设计师应用智能算法辅助设计的门槛。
Period-期: 11
Roll-卷: 59
PageCount-页数: 15
Page-页码: 16-30
SrcDatabase-来源库: 期刊
Organ-机构: 兰州理工大学机电工程学院;兰州城市学院培黎机械工程学院;
Link-链接: https://kns.cnki.net/kcms/detail/detail.aspx?FileName=JXXB202311002&DbName=CJFQTEMP

DataType: 1
Title-题名: 产品意象造型设计应用研究进展
Author-作者: 李雄;苏建宁;陈彦蒿;张秦玮;张新新;杨文瑾;
Source-刊名: 包装工程
Year-年: 2019
PubTime-出版时间: 2019-04-20
Keyword-关键词: 产品设计;意象造型;研究进展
Summary-摘要: 目的对产品意象造型设计应用研究进行综述,分析其发展现状、热点和趋势等问题。方法通过对国内外相关文献的研究,分析产品意象造型设计的体系结构和应用过程。结果产品意象造型设计应用过程主要包括产品意象挖掘定位、产品造型要素分析、意象造型设计等方面,其中意象造型设计可从单目标意象、多维意象、意象形态仿生、意象形态融合等角度展开。结论产品意象造型设计是现代工业设计的重要发展方向,其应用产品种类广泛,核心建模思想和方法在不断地完善和更新,其中准确挖掘产品意象,深层次解析产品要素、多维意象与智能化设计等,将是未来应用研究的难点和热点。
Period-期: 08
Roll-卷: 40
PageCount-页数: 9
Page-页码: 1-9
SrcDatabase-来源库: 期刊
Organ-机构: 兰州理工大学;西北工业大学;北方民族大学;华东理工大学;
Link-链接: https://kns.cnki.net/kcms/detail/detail.aspx?FileName=BZGC201908005&DbName=CJFQ2019
DOI-DOI: 10.19554/j.cnki.1001-3563.2019.08.001

四、专利授权情况

http://pss-system.cnipa.gov.cn/sipopublicsearch/portal/uilogin-forwardLogin.shtml

基于自然模式的 Logo 生成方法、智能 Logo 生成器
申请号 CN201810343073.9
申请日 2018.04.17
申请人所在国家/地区/组织 CN
代理人汤东凤
代理机构北京科亿知识产权代理事务所(普通合伙) 11350
公开公开号（公开）CN108549530A
公开日期（公开）2018.09.18
申请人（公开）兰州理工大学
发明人（公开）苏建宁; 李雄; 唐钊山
IPC 分类（公开）G06T11/40;G06F8/20;G06T11/60
CPC 发明（公开）G06F8/20;G06T11/40;G06T11/60

申请号	法律状态公告日	中文含义	英文含义
CN201810343073	20210312	发明专利权授予	granted
CN201810343073	20181016	实质审查的生效	initiative for examination as to substance
CN201810343073	20180918	发明专利申请公布	publication

（End）

Journal of Engineering Design

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/cjen20

Product innovation concept generation based on deep learning and Kansei engineering

Xiong Li, Jianning Su, Zhipeng Zhang & Ruisheng Bai

To cite this article: Xiong Li, Jianning Su, Zhipeng Zhang & Ruisheng Bai (2021): Product innovation concept generation based on deep learning and Kansei engineering, Journal of Engineering Design, DOI: 10.1080/09544828.2021.1928023

To link to this article: https://doi.org/10.1080/09544828.2021.1928023

Published online: 03 Jun 2021.

JOURNAL OF ENGINEERING DESIGN
https://doi.org/10.1080/09544828.2021.1928023

Product innovation concept generation based on deep learning and Kansei engineering

Xiong Li[a,b], Jianning Su[c], Zhipeng Zhang[c] and Ruisheng Bai[a]

[a]School of Mechanical & Electronical Engineering, Lanzhou University of Technology, Lanzhou, People's Republic of China; [b]School of Bailie Mechanical Engineering, Lanzhou City University, Lanzhou, People's Republic of China; [c]School of Design Art, Lanzhou University of Technology, Lanzhou, People's Republic of China

ABSTRACT

Industrial designers often present their initial concepts as design sketches. Rapid creation of new product conceptual images that meet users' affective preferences remains challenging in real design environments. However, few published works in affective design directly assist industrial designers in creating product conceptual images. Thus, we propose a product concept generation approach framework based on deep learning and Kansei engineering (PCGA-DLKE) to assist industrial designers. Our work focuses on dataset collection, pre-processing, affective preferences recognition, conceptual image generation model and product style transfer networks. To mark users' affective preferences, we established an affective recognition model by Kansei engineering and deep convolutional neural networks. To address the product conceptual image generation problem, we proposed a product design GAN model (PD-GAN), generating product conceptual images with affective preferences. An improved fast neural style transfer network was successfully trained to meet users' style preferences. This study aims to assist industrial designers in finding innovative concepts with affective preference. The Kansei evaluation shows that the innovation of the new product concept has been enhanced, indicating that the approach can better assist industrial designers in creating designs that meet users' emotional needs. Hand drill design and bicycle helmet design are taken as a case study.

ARTICLE HISTORY
Received 25 May 2020
Accepted 6 May 2021

KEYWORDS
deep learning; PCGA-DLKE; Kansei engineering; PD-GAN; product concept generation

1. Introduction

Nowadays, consumers are not only concerned with the functionality and reliability of a product, but they are also concerned with product emotions related to the feelings and impressions of the product, such as form, texture, colour and style (Yanagisawa and Fukuda 2005). People's demands are have become diversified, especially with the rapid progress of technology, emotional needs have become more prominent. Simultaneously, for enterprises, the era of gaining a competitive advantage by focusing only on product functions has passed (Khalid 2006). Enterprises need to consider the consumption preferences of

CONTACT Jianning Su sujn@lut.cn

users and consumers, especially emotional needs (Kwong et al. 2013). Products need to appeal to users and customers on an affective level to succeed in the highly competitive market (Chan et al. 2020). Therefore, affective design becomes the key point in the product upgrade process.

However, most of the initial concepts of product design are designed by industrial designers to explore the design direction through the hand-painted conceptual image pattern, which is a design pattern of personal introspection (Self 2019). This design process operates entirely in a black box by the designer's brain (Sutera, Yang, and Elsen 2014). This traditional design pattern is difficult to capture the emotional preference needs of users and consumers quickly. Therefore, how to quickly and effectively create conceptual images of products that meet consumer's emotional needs is a challenge for industrial designers in a competitive design environment. In this study, we propose a product concept generation approach framework based on deep learning and Kansei engineering (PCGA-DLKE).

Kansei engineering is widely used as a quantitative analysis method of new product affective design and development. It has three core tasks: affective modelling, determination of design elements settings, and design execution, that is, conceptual generation (Nagamachi 1995). Many previous researches mainly focused on the first two. Reviewing related literatures, we find that there are three types of concept generation methods. The first type concerns algorithm or data-driven design generation. A small number of investigations developed a specialized design tool in the case study, but they cannot be used as general-purpose tools. For example, Chen et al. developed a 3D knife design tool based on Kansei engineering and Visual Basic software (Chen and Chang 2014). Big-data mining brings new challenges and opportunities for designer's idea generation. For example, Liu et al. proposed data-driven concept network-assisted design concept generation (Liu et al. 2020). Second, generate a new design scheme using 3D software. Guo et al. (2014), for example, used Rhino to create a digital camera scheme after the Kansei engineering mole was determined. Third, design sketches show the results of Kansei engineering. For example, the case in (Chen, Yeh, and Lin 2010) shows chair design sketches drawn by a product designer base on design elements determined by Kansei engineering. These three methods lack continuity with the first two core tasks in Kansei engineering. Because the design schemes are not directly generated by Kansei engineering. This shows that concept generation is a shortcoming of Kansei engineering. Such a short board can easily lead to the separation of design research and design practice in the implementation process.

Over the past few years, deep learning technology, such as deep convolutional neural networks (DCNN), have made breakthroughs in many fields, for instance, computer vision, autopilot, Medical Science and games (Hadji and Wildes 2018). These efforts have resulted in new state-of-the-art performance on a wide range of classification (Krizhevsky, Sutskever, and Hinton 2017) and regression tasks (Eigen and Fergus 2015). In industrial design, Pedro et al. (2018) proposed using CNN to evaluate the usability of a product, and a case study using a thermostat as an example for rapid product evaluation and development. Pan et al. (2017) used a scalable deep learning approach to predict and interpret customer aesthetic perceptions of heterogeneous market design attributes, and used automotive aesthetic perceptions as a case study. However, the model constructed lacks explanations for the aesthetic perception of the heterogeneous market. Wang, Mo, and Tseng (2018) proposed an affective modelling approach based on deep learning techniques that automatically correlate customer needs with product design parameters. Besides, the invention of the

Generative Adversarial Networks (GAN) (Goodfellow et al. 2014) has breathed new life into the innovative design of the product. Kim et al. (Kim et al. 2017) used GAN to implement the cross-domain generative design, which improves the associative innovation capability of an aided design system. Chai et al. (2018) proposed a GAN-based automatic colouring model for footwear design sketches, but their model lacks consideration of consumer affective preferences. Chen et al. (2019) proposed a visual concepts combination model of the GAN model consisting of a double discriminator for product conceptual innovation, but their model almost ignores the functionality of the product. Quan, Li, and Hu (2018) proposed a combined deep learning and Kansei engineering method for product style transfer that can automatically generate a new scheme with a specific style, but this approach simply changes the style of the product and does not create a new product form. In addition, the recognition and prediction of product affective preferences is a key issue in Kansei engineering research (Nagamachi 1995; Quan, Li, and Hu 2018; Wang, Mo, and Tseng 2018). Traditional machine learning methods, such as SVM and Multi-Layer Perceptron, are unable to recognise the affective preferences of the products at the level of overall visual perception as humans do.

To overcome the above problem, we proposed a product concept generation approach framework based on deep learning and Kansei engineering (PCGA-DLKE). This framework benefits for applying computer-aided design and creation and realizes the integration of creativity and inspiration.

The success of deep convolutional neural networks in the field of image recognition such as Alex-Net (Krizhevsky, Sutskever, and Hinton 2017), VGG-Net (Simonyan and Zisserman 2015) reminds us of the recognition of product affective preferences in affective design. While the invention of GAN (Goodfellow et al. 2014) shocked us, we saw its great potential in design. The style transfer network (Gatys, Ecker, and Bethge 2015) reborn Van Gogh's style, which aroused widespread concern in academia and industry. This also prompted us to imagine the space of product style. In short, these works inspired our passion for developing research into product design applications.

This paper proposes a framework model that produces product concept design for an industrial designer to inspire innovation. We propose a relatively complete product concept image generation framework based on deep learning technology and Kansei engineering.

Our contribution can be described as three points of innovation: (1) An integrated approach (PCGA-DLKE), which combines Kansei engineering and deep learning techniques, is proposed for innovation concept generation, to improve the product conception design process efficiently. It has three core modules. A product affective preferences recognition module, based on the modified deep residual networks and Kansei engineering. A valid PD-GAN algorithm framework for product conceptual image generation and a suitable PD-GAN model based on DCGAN and Residual Networks. We propose using the fast-neural style transfer technique for a new generative product by reconstructing and merging the style image's pattern features. We also improved the structure of the original fast-neural style transfer networks with more robust feature transfer capability. (2) Previous studies generate a limited space for conceptual designs. In this paper, our approach framework generates many innovative conceptual designs with affective preferences and rapidly changes the style features of the generated solutions. (3) We applied the proposed PCGA-DLKE framework to the hand drill and bicycle helmet creative concept generation problem to demonstrate the value of our method. Experiments demonstrate the innovative design potential of the methodological framework to improve affective design efficiency through an end-to-end learning approach.

The rest of this paper is organised as follows. In Section 2, we firstly present the overall research framework. Next, we describe the product image data preparation, affective preferences recognition and label estimation, product design generative adversarial networks (PD-GAN), and product style fast-neural style transfer networks (PS-FNSTN). In Section 3, an empirical case is given to verify the proposed framework for an industrial designer, and the related experimental results are shown. Finally, discussion, conclusions, and future work are given in Sections 4 and 5.

2. Methods

2.1. Research framework

In order to exploit artificial intelligence (AI) for innovative conceptual design, we propose a generative approach to the product concept image, the PCGA-DLKE based on deep convolutional neural networks, deep convolutional generative adversarial networks and fast neural style transfer networks model. As shown in Figure 1, the PCGA-DLKE framework contains four parts. In part 1, product image data crawling and a series of image data pre-processing. Step A of data preparation aims to obtain the original data of the target product. In this step, crawler technology crawls the product pictures from the target websites, and the design team draws the concept image. Step B involves the pre-processing of the original data, aiming to obtain a clean image data set. Based on this module, we can obtain a clean and uniform product image dataset. In part 2, Kansei engineering (KE) and deep residual networks are used for label estimation of product images. In part 3, we use a variant of vanilla generative adversarial networks (named PD-GAN) to create a newer product design that meets user and customer's affective needs. In the last part, the style transfer technique is applied to extend the affective domain of the product. Because we applied the fast-neural style networks method to product images, we named it PS-FNSTN. By trained the VGG-19 (Simonyan and Zisserman 2015) model, we can turn the content image and style image into a stylised image, which is generated entirely new product.

2.2. Product image data preparation

Data is one of the three critical components (big data, computing power and algorithms) of artificial intelligence. High-quality image dataset is indispensable for successful image generation. Currently, there are few dedicated open-source product images dataset for training image generation models. Previous works (Jin et al. 2017; Mattya 2015) have shown that images crawled directly from the web have high inter-image variance and noise. To obtain high quality, clean product images dataset, we proposed an efficient operation that includes two parts. In the first part, we obtain the original product images data through the web crawler and the work of the design team. The design team can only draw a limited number of product images. Therefore, the original images data are mainly captured from the object web pages through web crawlers. We found the Bing search engine to be a practical tool for collecting product images. To collect more object images, it is important to

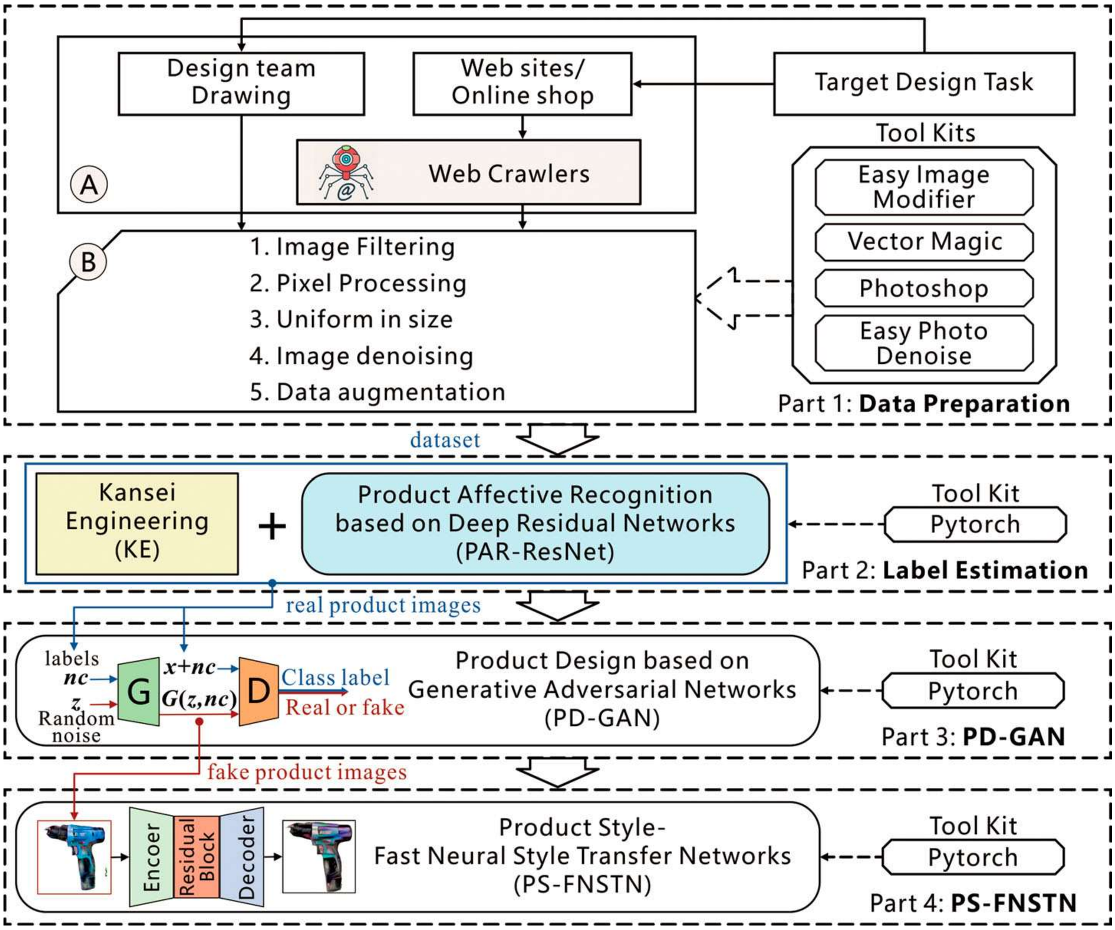


Figure 1. The product concept generation approaches framework based on deep learning and Kansei engineering (PCGA-DLKE).

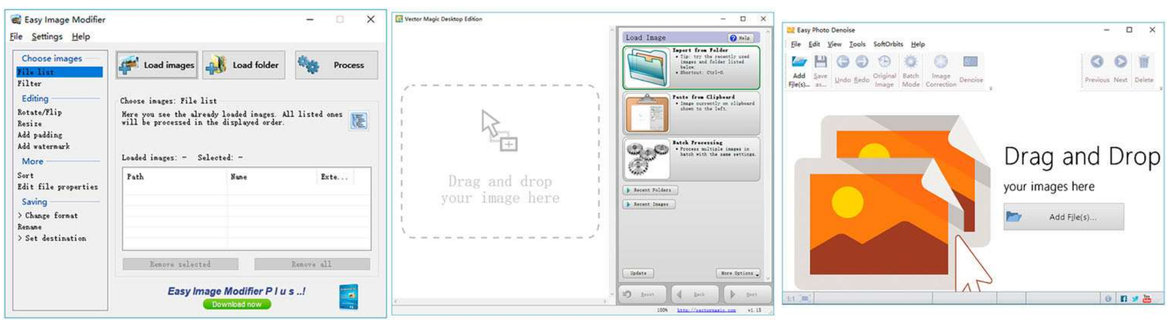


Figure 2. Left: Easy Image Modifier. Middle: Vector Magic. Right: Easy Photo Denoise.

constantly change search keywords for web crawlers and continuously update the landing page. In the second part, we need to perform some series of pre-processing on the original images, including image filter for selecting good images, pixel pre-treatment, uniform in size, data augmentation, and image noise reduction. During image data pre-processing, we found that three tools can provide effective auxiliary functions. They are Easy Image Modifier, Vector Magic, Photoshop and Easy Photo Denoise, as shown in Figure 2. Vector Magic and Photoshop can be used together to enhance image quality. The combination of Easy Image Modifier and Photoshop can be used to achieve data augmentation. We did not use any data augmentation techniques provided by deep learning frameworks (e.g.

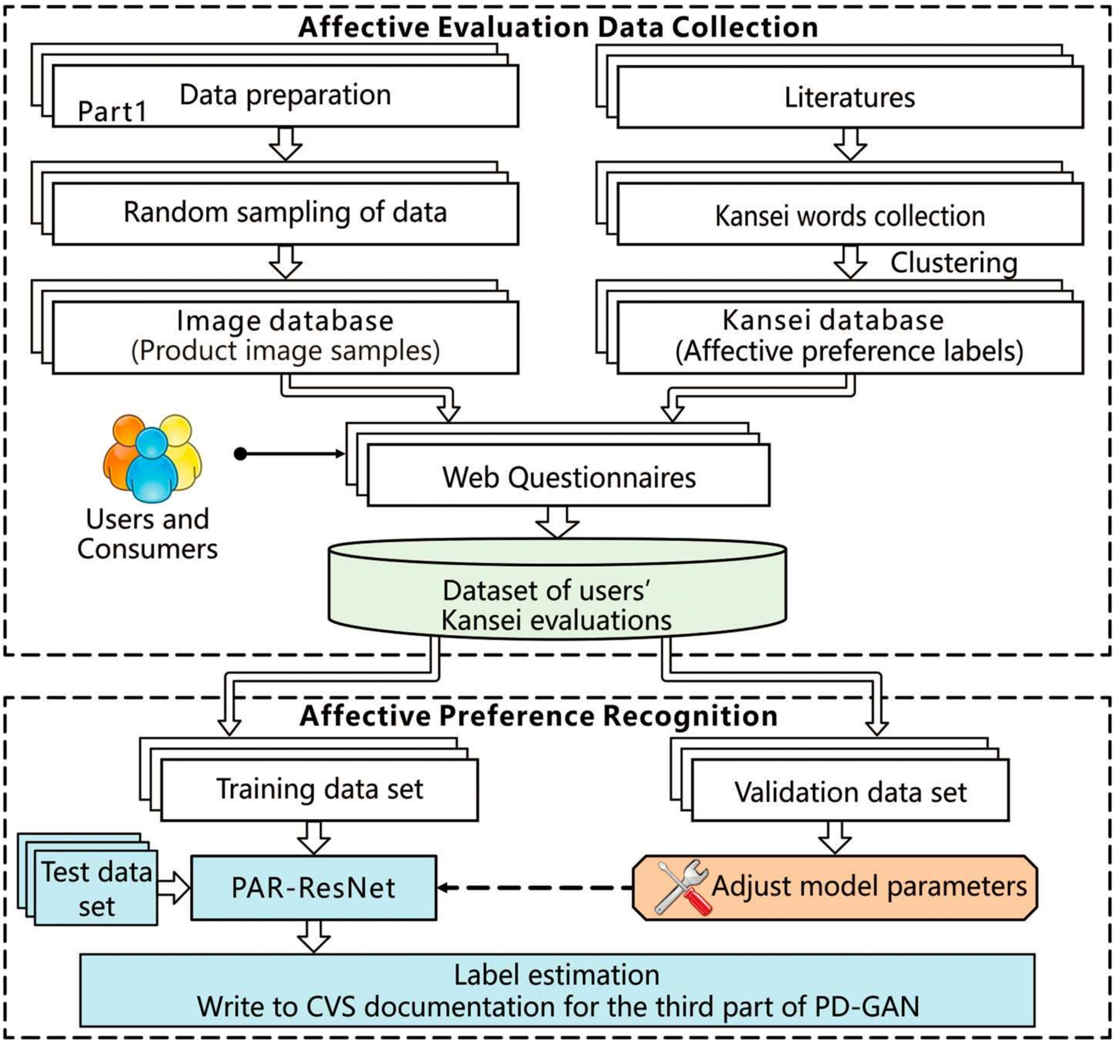


Figure 3. Overall flow of labels estimation approach based on Kansei engineering (KE) and deep residual networks (ResNet).

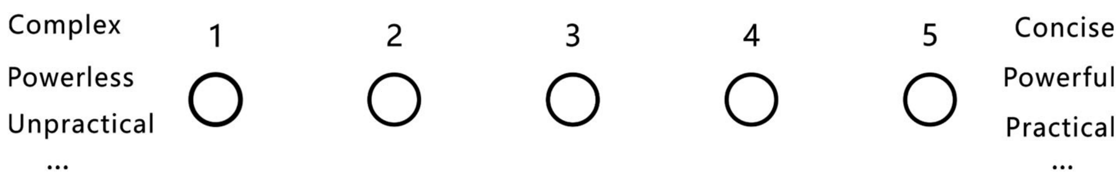


Figure 4. 5-point semantic differential scale for the online questionnaire.

PyTorch and TensorFlow), as we found that they often lead to incomplete product images. This is not good for subsequent affective preference recognition and product image generation. To obtain a clean image, image noise reduction is necessary. In this study, Easy Photo Denoise is used for image noise reduction.

2.3. *Affective preference recognition and label estimation*

According to Figure 3, the proposed approach to label estimation consists of two modules: separately affective evaluation data collection and product affective preference recognition. this section gives a detailed introduction to the method of affective preference labels estimation.

Affective evaluation data collection of the product includes collecting product images from various websites, which is taken from the first part of the approach framework and Kansei words collection and web questionnaires. This affective evaluation data collection process aims to build the data onto users and customers' Kansei evaluations. The product affective preference recognition includes data splitting and a training process of the deep convolutional neural networks. It uses the dataset collected from the last step to train the deep learning models of the product features and affective preference labels (that is Kansei labels). The network is continuously updated with the growth of data. Lastly, the affective preference recognition application involves applying the PAR-ResNet to help designers estimate the types of users' affective preference and label automatically for the product image.

2.3.1. Affective evaluation data collection

The data used for affective evaluation includes product image sample data (image database) and corresponding affective preference labels (Kansei database). Product image sample data can be randomly extracted from the image data collected in the previous step. The affective preference labels in the second part can be obtained from the Kansei words. Because Kansei words are usually adjectives that describe people's emotions and feelings, and are widely used in design research and practice.

Affective preferences collection. Generally, the acquisition of affective preferences in design research and practice requires two steps. First, the extensive collection of Kansei vocabulary. Kansei words usually use adjectives, an expression channel that intuitively reflects people's emotions and feelings about the product. They can be collected through various channels, such as magazines, academic paper, product test reports, product manuals, expert reviews, user opinions, web reviews and customer interview (Chou 2016). In this study, we collect Kansei words from related academic literatures and the internet. Second, clustering all collected Kansei words result in affective preference attributes. This process is called affective clustering. Currently, there are three Kansei clustering methods that are frequently used in an affective design, which are the clustering method based on fuzzy equivalence relation (Chou 2016), the Kansei clustering method based on a design structure matrix (DSM) (Huang, Chen, and Khoo 2012), the rough set-based clustering method (Zhai, Khoo, and Zhong 2009). In this study, we used the methods and steps proposed by Li et al. (2018), and hand drills are selected to conduct the case study. There was a similar study about battery drill used 25 adjectives as Kansei words, and cluster analysis (Grimsaeth et al. 2010).

Questionnaire construction. Because the internet and smartphones have become an integral part of people's lives, such as online office, online learning and online shopping, it is natural to save time and labour by performing surveys online traditional surveys. Besides, what excites us most is that the speed of data update online surveys is unmatched by traditional surveys. Consumers' and user's affective responses obtained through online questionnaires are more authentic than conventional surveys. We used the determined affective preference tags and product images to construct an online questionnaire based on the semantic differential. Semantic differential, proposed by Osgood, Suci, and Tannenbaum (1957), an American psychologist, is a user-centric design technique commonly used to obtain consumer affective scores for products. This study uses the 5-point semantic differential scale for quantifying consumers and user's affective preferences. As shown in Figure 4, from 1 to 5, each point represents a preference level of the customers and users. For instance, 1 and 5 represent a pair of bipolar adjectives, while 3 represents a medium level. Finally, the user Kansei evaluation data set was obtained through online surveys.

2.3.2. Affective preferences recognition

Essentially, the product affective preferences recognition is a classification problem. Obtaining the user evaluation Kansei evaluation data set, we use a deep learning algorithm to model the relation between product (images) and consumers' and users' affective preferences tags.

Deep learning, especially deep convolutional neural networks (CNNs), has been widely used in image recognition (Krizhevsky, Sutskever, and Hinton 2017; Simonyan and Zisserman 2015). Because of their ability to hierarchically abstract representations with local operations (Hadji and Wildes 2018). Many experiments have proved that CNNs can extract the identifiable overall features of the images through local receptive fields, weight sharing and local down sampling (Dong et al. 2014). More and more network architectures with high representation capabilities have been proposed and refreshed the image recognition records again and again, for example, LeNet (Lecun et al. 1998), Alex-Net (Krizhevsky, Sutskever, and Hinton 2017), VGG-Net (Simonyan and Zisserman 2015), GooleLeNet (Szegedy et al. 2015) and ResNet (He et al. 2016). In this study, we used a modified ResNet18 as the affective preferences recognizer, namely PAR-ResNet. Therefore, more new product images can be automatically recognised by the trained model and labelled with the affective preference tags. Since it is a classification problem, we use the cross-entropy loss function to evaluate PAR-ResNet.

$$\mathcal{L}(\boldsymbol{\theta})_{\text{PAR-ResNet}} = -\left(\frac{1}{m}\right)\sum_{j=1}^{m}[\log h_{\boldsymbol{\theta}}(\boldsymbol{x}^{(j)}) + (1-\boldsymbol{y}^{(j)})\log(1-h_{\boldsymbol{\theta}}(\boldsymbol{x}^{(j)}))] \tag{1}$$

The architecture of par-resnet will be presented in Section 3.2, the recognition accuracy is measured by the confusion matrix, which is calculated by Equation (2):

$$\text{accuracy}_{\text{Dataset}} = \frac{TP + TN}{TP + TN + FP + FN} \tag{2}$$

where, TP is True Positive, TN is True Negative, FP is False Positive and FN is False Negative.

2.4. Product design generative adversarial network

Generative Adversarial Networks (GANs) is a hybrid generative model whose core idea comes from the two-person zero-sum game in game theory (Goodfellow et al. 2014). The basic GAN model consists of two networks: A Generator and a Discriminator. Sampling random variable $\boldsymbol{z}$ from a probability distribution P_z (for example, Gaussian) as input to the Generator $\boldsymbol{G}$, through a nonlinear mapping of $\boldsymbol{G}$, output signal $\boldsymbol{G}(\boldsymbol{z})$. The discriminator $\boldsymbol{D}$ takes $\boldsymbol{G}(\boldsymbol{z})$ or $\boldsymbol{x}$ as input, and determines whether the input data comes from real data or generated data by calculating the probability that it belongs to real data. Specifically, the original GAN adopts the adversarial learning strategy to train $\boldsymbol{G}$ and $\boldsymbol{D}$, so that the training objectives of the two are opposite. Mathematically, the goal of GAN can be expressed as

$$\min_{\boldsymbol{G}} \max_{\boldsymbol{D}} V(\boldsymbol{D}, \boldsymbol{G}) = \mathbb{E}_{\boldsymbol{x} \sim \boldsymbol{P_{data}}(\boldsymbol{x})}[\log \boldsymbol{D}(\boldsymbol{x})] + \mathbb{E}_{\boldsymbol{z} \sim \boldsymbol{P_z}(\boldsymbol{z})}[\log(1 - \boldsymbol{D}(\boldsymbol{G}(\boldsymbol{z})))] \quad (3)$$

In Equation (3), the discriminator $\boldsymbol{D}$ tries to maximise the log likelihood function from real data. Meanwhile, it also minimises the log likelihood function from fake data generated by generator $\boldsymbol{G}$. In contrast, the goal of G is to minimize the log-likelihood function and make the distribution $\boldsymbol{P_G}$ of $\boldsymbol{G}(\boldsymbol{z})$ approach the distribution $\boldsymbol{P}_{data}$ of real data. The discriminator $\boldsymbol{D}$ is the object deceived by generator $\boldsymbol{G}$.

Although Goodfellow et al. (2014) have theoretically proved the convergence of the GAN model. However, in practice, GANs still has problems such as training instability, mode collapse, convergence difficulties, difficulty inaccurate control of generated content (Arjovsky and Bottou 2017). Therefore, different architectures, loss functions, conditional techniques and constrain methods were introduced, easing the convergence of GAN models. In improving the architecture of networks, such as Deep Convolutional GAN (2016) (Radford, Metz, and Chintala 2016), Deep Regret Analytic GAN (2017) (Kodali et al. 2017), PGAN (2018) (Karras et al. 2018) and Big-GAN (2019) (Brock, Donahue, and Simonyan 2019). In terms of improving the loss function, such as Least Squares GAN (2016) (Mao et al. 2017), Wasserstein GAN (2016) (Arjovsky, Chintala, and Bottou 2017) and Boundary-Seeking GAN (2018) (Hjelm et al. 2017). In terms of improving the conditional techniques, such as Conditional GAN (2014) (Mirza and Osindero 2014), Auxiliary Classifier GAN (2016) (Odena, Olah, and Shlens 2017), Triple GAN (2017) (Li et al. 2017) and Style GAN (2019) (Karras, Laine, and Aila 2019). Unfortunately, there is currently no comprehensive comparative study of GAN models. However, we can always make the right choice since no free lunch theorem in machine learning and considering the reality of industrial product design.

In this paper, we choose DCGAN (Radford, Metz, and Chintala 2016) as the basic GAN model. We have three reasons: First, the number of images of the same product in real world is limited, for example, the number of hair dryers is far less than the number of faces. Second, the network architecture should not be too complex because of the limited amount of data. The parameters that need to be calculated for a complex network structure will also increase dramatically. Third, the model must be easy to train. Besides, the work of Mattya (2017) and Elgammal et al. (2017) also shows that DCGAN's generation ability is excellent. In our experiments, we successfully train the PD-GAN model based on ResNet (He et al. 2016) and DCGAN (Radford, Metz, and Chintala 2016). Figure 5 shows an overview of our PD-GAN algorithm framework. The architecture of the generative network will be presented in Section 3.3.

The discriminator model needs to determine whether each sample is true or false, and it also needs to complete to a categorisation task to predict affective preference type $\boldsymbol{c}$ by adding an auxiliary classifier. Thus, the loss function of the discriminator is described as following:

$$\mathcal{L}_{adv}(\boldsymbol{D}) = -E_{\boldsymbol{x} \sim \boldsymbol{P}_{data}(\boldsymbol{x})}[\log \boldsymbol{D}(\boldsymbol{x})] - E_{\boldsymbol{z} \sim \boldsymbol{P_z}(\boldsymbol{z}), \boldsymbol{nc} \sim \boldsymbol{P_{nc}}(\boldsymbol{nc})}[\log(1 - \boldsymbol{D}(G(\boldsymbol{z}, \boldsymbol{nc})))] \quad (4)$$

$$\mathcal{L}_{cls}(\boldsymbol{D}) = \mathbb{E}_{\boldsymbol{x} \sim \boldsymbol{P_{data}}(\boldsymbol{x})}[log \boldsymbol{P_D}(\boldsymbol{c_x}|\boldsymbol{x})] + \mathbb{E}_{\boldsymbol{z} \sim \boldsymbol{P_z}(\boldsymbol{z}), \boldsymbol{nc} \sim \boldsymbol{P_{nc}}(\boldsymbol{nc})}[log \boldsymbol{P_D}(\boldsymbol{nc}|\boldsymbol{G}(\boldsymbol{z}, \boldsymbol{nc}))] \quad (5)$$

$$\mathcal{L}_{gp}(\boldsymbol{D}) = \mathbb{E}_{\tilde{\boldsymbol{x}} \sim \boldsymbol{P_{data}}(\tilde{\boldsymbol{x}})}[(\|\nabla_{\tilde{\boldsymbol{x}}} \boldsymbol{D}(\tilde{\boldsymbol{x}})\|_2 - 1)^2] \quad (6)$$

where the term $\mathcal{L}_{gp}$ from DRAGAN (Kodali et al. 2017).

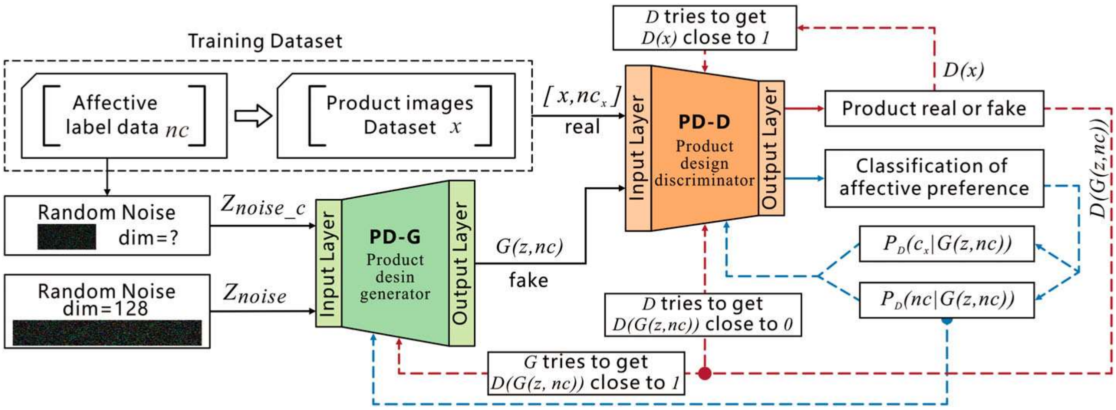


Figure 5. Product Design-GAN algorithm framework.

For the generator model, the loss function is as follows:

$$\mathcal{L}_{adv}(\boldsymbol{G}) = \mathbb{E}_{\boldsymbol{z}\sim \boldsymbol{P_z}(\boldsymbol{z}),\boldsymbol{nc}\sim \boldsymbol{P_{nc}}(\boldsymbol{nc})}[\log(\boldsymbol{D}(\boldsymbol{G}(\boldsymbol{z},\boldsymbol{nc})))] \tag{7}$$

$$\mathcal{L}_{cls}(\boldsymbol{G}) = \mathbb{E}_{\boldsymbol{z}\sim \boldsymbol{P_z}(\boldsymbol{z}),\boldsymbol{nc}\sim \boldsymbol{P_{nc}}(\boldsymbol{nc})}[\boldsymbol{P_D}(\boldsymbol{nc}|\boldsymbol{G}(\boldsymbol{z},\boldsymbol{nc}))] \tag{8}$$

Summarily, the loss function of PD-GAN can be expressed as the following two simple expressions:

$$\mathcal{L}(\boldsymbol{D}) = \mathcal{L}_{adv}(\boldsymbol{D}) + \beta\mathcal{L}_{cls}(\boldsymbol{D}) + \lambda\mathcal{L}_{gp}(\boldsymbol{D}) \tag{9}$$

$$\mathcal{L}(\boldsymbol{G}) = \mathcal{L}_{adv}(\boldsymbol{G}) + \beta\mathcal{L}_{cls}(\boldsymbol{G}) \tag{10}$$

where β and λ are balance factors for the adversarial loss and gradient penalty, respectively.

2.5. Fast neural style transfer network

Style transfer is also called style conversion. The intuitive analogy adds a filter to the input image, but it is different from traditional filters. Generally, style transfer uses a convolutional neural network to automatically apply styles from one image to another. A content image and a style image are used to create an output image whose 'content' mirrors the content image and whose style resembles that of the style image (Gatys, Ecker, and Bethge 2015). In contrast to Neural Style, Fast Neural Style has designed a network specifically for style transfer. When inputting a picture, the network automatically generates the target picture in real time (Johnson, Alahi, and Fei-Fei 2016). This network needs to train a corresponding style network for each style image, but once the training is completed, it only takes 20 s less to complete a style transfer. This is well suited to early product concept design needs and helps designers quickly experiment with different styles in order to make the right judgment. The network structure of a fast-neural style consists of two parts: an image transformation network ***fw*** (that is generative network) and a loss network $\boldsymbol{\phi}$. Figure 6 shows a structural overview of our fast-neural style transfer network, in other words, generative network on the left and loss network on the right. The former is the network we need to train, the latter is a trained network. We utilised the 'Encoder-IN-Decoder' architecture. The architecture of the generative network is presented in section 3.4.

The transform network ***fw*** takes a content image *c* and a style image *s* as inputs and synthesises an output image that recombines the content of the former and style of the

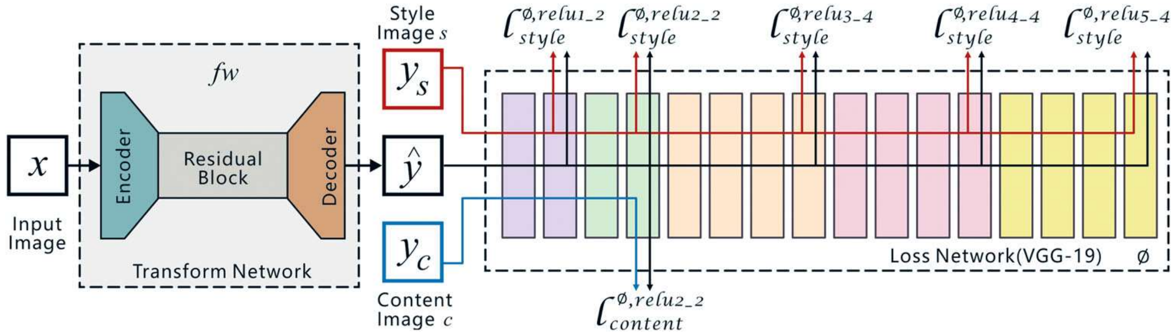


Figure 6. Fast style transfer model for product design. Left: style transfer network. Right: loss network.

later. In the task of style transform, $\boldsymbol{x}$ is the input image, $\boldsymbol{y}_c = \boldsymbol{x}$, $\boldsymbol{y}_s$ is the style image. For the input image $\boldsymbol{x}$, $\boldsymbol{fw}$ can return a new image $\hat{\boldsymbol{y}}$, so $\boldsymbol{fw}$ is naturally the style transform network we want to design. $\hat{\boldsymbol{y}}$ is similar to $\boldsymbol{y}_c$ in content, but is similar to $\boldsymbol{y}_s$ in style. Generally, the loss network can be used to calculate visual features and style features without training. In this study, we used VGG-19 (Simonyan and Zisserman 2015), which has been trained on the ImageNet dataset. The choice of content presentation layer and style presentation layer is derived from the literature (Luan et al. 2017). Therefore, all three ($\boldsymbol{y}_s$, $\hat{\boldsymbol{y}}$ and $\boldsymbol{y}_c$) are input into the loss network $\boldsymbol{\phi}$, and corresponding losses (including content loss and style loss) are generated. In this way, we can get ideal style images by end-to-end training.

The content loss is the (squared and normalised) Euclidean distance between the stylised image and the original image:

$$\mathcal{L}_{content}^{\phi,j}(\hat{\boldsymbol{y}}, \boldsymbol{y}_c) = \frac{1}{\boldsymbol{C}_j \boldsymbol{H}_j \boldsymbol{W}_j} ||\boldsymbol{\phi}_j(\hat{\boldsymbol{y}}) - \boldsymbol{\phi}_j(\boldsymbol{y}_c)||_2^2 \tag{11}$$

where $\boldsymbol{C}_j \boldsymbol{H}_j \boldsymbol{W}_j$ is the shape of the feature map. According to our definition of style perception, the style loss can be simply understood as removing the spatial information and fusing the feature response of each channel. The style loss is the squared Frobenius norm of the difference between the Gram matrices of the output and target images:

$$\mathcal{L}_{style}^{\phi,j}(\hat{\boldsymbol{y}}, \boldsymbol{y}_s) = \frac{1}{\boldsymbol{C}_j \boldsymbol{H}_j \boldsymbol{W}_j} ||\boldsymbol{G}_j^{\phi}(\hat{\boldsymbol{y}}) - \boldsymbol{G}_j^{\phi}(\boldsymbol{y}_s)||_F^2 \tag{12}$$

The loss function that is ultimately used for training is weighted sum of content loss and style loss.

$$\mathcal{L}_{total} = \alpha \mathcal{L}_{content}^{\phi,j}(\hat{\boldsymbol{y}}, \boldsymbol{y}_c) + \beta \mathcal{L}_{style}^{\phi,\boldsymbol{J}}(\hat{\boldsymbol{y}}, \boldsymbol{y}_s) + \gamma \mathcal{L}_{tv}(\hat{\boldsymbol{y}}) \tag{13}$$

where $\mathcal{L}_{total}$ is the loss function, α is the weight of the content loss, β is the weight of the style loss, $\mathcal{L}_{tv}(\hat{\boldsymbol{y}})$ total variation regularizer, and γ is total variation weight.

3. Empirical study

In this paper, two case studies of generating hand drills and bicycle helmets were conducted to verify the proposed framework's practicality and effectiveness. It had the following steps: (1) Data collection and data pre-processing. To obtain a clean dataset, we pre-processed the messy raw data by web crawler using a series of the combined steps and methods. (2) A modified ResNet18 (PAR-ResNet) was trained to identify the type of

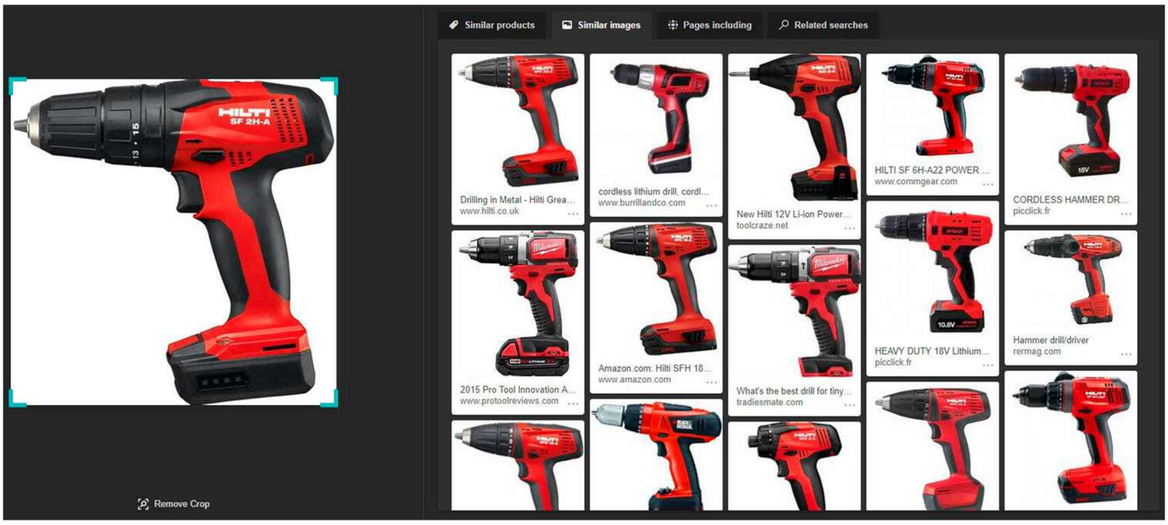


Figure 7. Example of a Bing Search Web page search hand drills.

affective preference of the target product. (3) The PD-GAN model based on DCGAN and residual blocks was trained to generate new product images and evaluate them. (4) A modified fast neural style transfer networks (PS-FNSTN) was trained to transfer the style of the new product concept image.

Our PCGA-DLKE framework was developed using Python. Comprehensively, PAR-ResNet, PD-GAN, PS-FNSTN were built using the Python language along with Python modules such as PyTorch, TorchVision, TorchNet, Torchsummary, Visdom and NumPy. All experiments were run on a YunXuan workstation (YunXuan Inc., Shanghai, China) with Intel i7-9700K and RTX2080 8G (double) and Ubuntu 18.04. operating system.

3.1. Image data preparation

It is well known that an image dataset in high quality is essential, if not most important, to the success of product concept image generation. However, it is not easy to get enough product pictures. Our collection of images is based on the approach mentioned in Section 2.2. First, some of the product images are crawled from a web page provided by the Bing search engine using a web crawler tool, and others are crawled from Amazon.com, JD.com and Taobao.com. For example, the results of the Bing search are shown in Figure 7. Table 1 lists the keywords and samples we used in the process of searching for hand drill images. Second, we manually check all product images and remove undesired images. It should be noted that we do not use a hammer drill as a search keyword. Thus, there are no hammer drill images in our dataset. Third, we used Easy Image Modifier for image filtering, format conversion (∗.jpg), standard size (resolution: 256×256, 128×128). The combination of Vector Magic and Photoshop was used to improve pixel quality. Easy Photo Denoise was used for image noise reduction. Easy Image Modifier and Photoshop were combined for data augmentation technology (random colour and image flipping). Finally, we harvested 18,285 images of hand drills (i.e. hand drill dataset, named HD dataset) and 15,456 images of bicycle helmets (i.e. bicycle helmet dataset, named BH dataset).

Table 1. The list of search keyword for hand-held electric tools.

Sample picture	Search keyword	Sample picture	Search keyword	Sample picture	Search keyword
	Hand Drill		Power Drill		Electric Drill
	Cordless Drill		Battery Drill		Electrical Drill
	Cord Drill		Automatic Screwdriver		Electric Screwdriver

3.2. Product affective recognition and labelling

3.2.1. Affective evaluation data

The procedure for obtaining affective evaluation data uses the approach proposed in Section 2.3. The data consists of two parts: image database, that is hand drill images and Kansei database, that is affective preference labels data. In this study, 840 hand drills are randomly selected from the previous hand drill image dataset.

Table 2 shows the collected Kansei words and Kansei labels (i.e. affective preference labels). We collected 168 Kansei words from the existing literatures and online shop reviews. We referred to the method mentioned in (Li et al. 2018) to cluster and optimise the Kansei words. First, the collected Kansei words with opposite meanings are clustered into a cluster by semantic cluster analysis. Sixteen clusters were obtained in total. Second, all the clusters were ranked by the value of the sum of the amount and total frequency of the Kansei words. Third, the first six high-scoring clusters were naturally selected, and the affective attribute of each cluster is determined based on the semantic matching degree and common degree. Finally, six representative affective attributes were identified. They are Ergonomic–Uncomfortable, concise–complex, powerful–powerless, Practical–Unpractical, Handy–Bulky and Appealing–Unaesthetic. This means that we have 6 Kansei preference dimensions, that is, 12 Kansei labels. These Kansei words relate to four essential aspects of the hand drill: function, material, operation control and appearance, which are of interest to users (Grimsaeth et al. 2010).

The online questionnaire consists of three parts: a hand drill image, 6 Kansei dimensions (i.e. 12 Kansei labels) and a 5-point semantic difference scale. The first example is shown in Figure 8. About 840 different hand drill samples we selected to be investigated in total.

There are two types of questionnaires, mobile and computer versions, distributed through social networking sites. To ensure the accuracy of the evaluation and considering the respondent's visual tolerance, we randomly divided 840 hand drill samples into 20 parts, each part containing 42 samples. Each participant was given an incentive during the implementation. The minimum number of participants for each online questionnaire was set at 32, and each participant answered the question up to 2 times. In the end, a total of 655 people participated in the online survey on affective preferences for the hand drill. Their age range: 28–45 years. They have diverse careers, including engineers, construction workers, designers, college students, renovators, etc. Therefore, only 42 sample images were scored for each participant, and 640 (32×20) complete questionnaires were obtained. Each of the six Kansei dimensions is evaluated independently of each other. Finally, the number

Table 2. Kansei labels and Kansei words.

	Kansei labels	Kansei words (examples)	References (examples)
K1	Uncomfortable–Ergonomic	comfortable, handling comfort, cozy, restrained, Ergonomic, round	(Kim et al. 2019), (Grimsaeth et al. 2010), (Chang and Chen 2016), Online shop reviews
K2	Complex–Concise	minimal, simple, plain, complex, complicated, dazzling, simplistic, compact	(Grimsaeth et al. 2010), (Chang and Chen 2016), (Razza and Paschoarelli 2015), Online shop reviews
K3	Powerless–Powerful	powerful, strong, forceful, vigorous, energetic, energetical, weak, flaccid, powerless	(Vieira et al. 2017), (Wang et al. 2016), online shop reviews
K4	Unpractical–Practical	quality, reliable, high-quality, sturdy, safe, accurate, robust, solid, durable, unreliable,	(Kim et al. 2019), (Guo et al. 2016), (Hsiao, Chen, and Liao 2017), Online shop reviews
K5	Bulky–Handy	handy, portable, heavy, ingenious, bulky, flexible, lightweight	(Kim et al. 2019), (Chou 2016), Online shop reviews
K6	Unaesthetic–Appealing	artistic, aesthetic, appealing, cute, elegant, good-looking, exquisite, eye-catching, unaesthetic, artless, attractive	(Chou 2016), (Fung et al. 2014), (Jiao, Zhang, and Helander 2006), Online shop reviews

Table 3. The Kansei evaluation data of hand drill.

Label	Ergonomic	Uncomfortable	Concise	Complex
Number	447	393	452	388
Label	Powerful	Powerless	Practical	Unpractical
Number	445	395	446	394
Label	Handy	Bulky	Appealing	Unaesthetic
Number	464	376	416	424

of samples corresponding to each Kansei label is shown in Table 3. Figure A1 shows the distribution of hand drill images in our dataset.

3.2.2. PAR-ResNet architecture

Figure 9 shows PAR-ResNet's architecture, a minor modification from ResNet18 (He et al. 2016). We only added a data deformation layer before the last Linear layer, which is named Flatter layer. Of course, the number of output features of the Linear layer must modify to what we need. The Model contains four modules made up of ResBlocks, each of which uses two ResBlocks with the same number of output channels. There are four convolutional layers in each module except for 1×1 convolutional layer. The number of all channels in the first module is 64. Because the previous layer is a maximum pooling layer with a stride of 2, there is no need to reduce the height and width. A residual block with a 1×1 convolution layer can double the number of channels of the previous residual block, and the height and width are also halved. Furthermore, Gaussian error linear units (GELUs) (Hendrycks and Gimpel 2016) is introduced as activation function in our model.

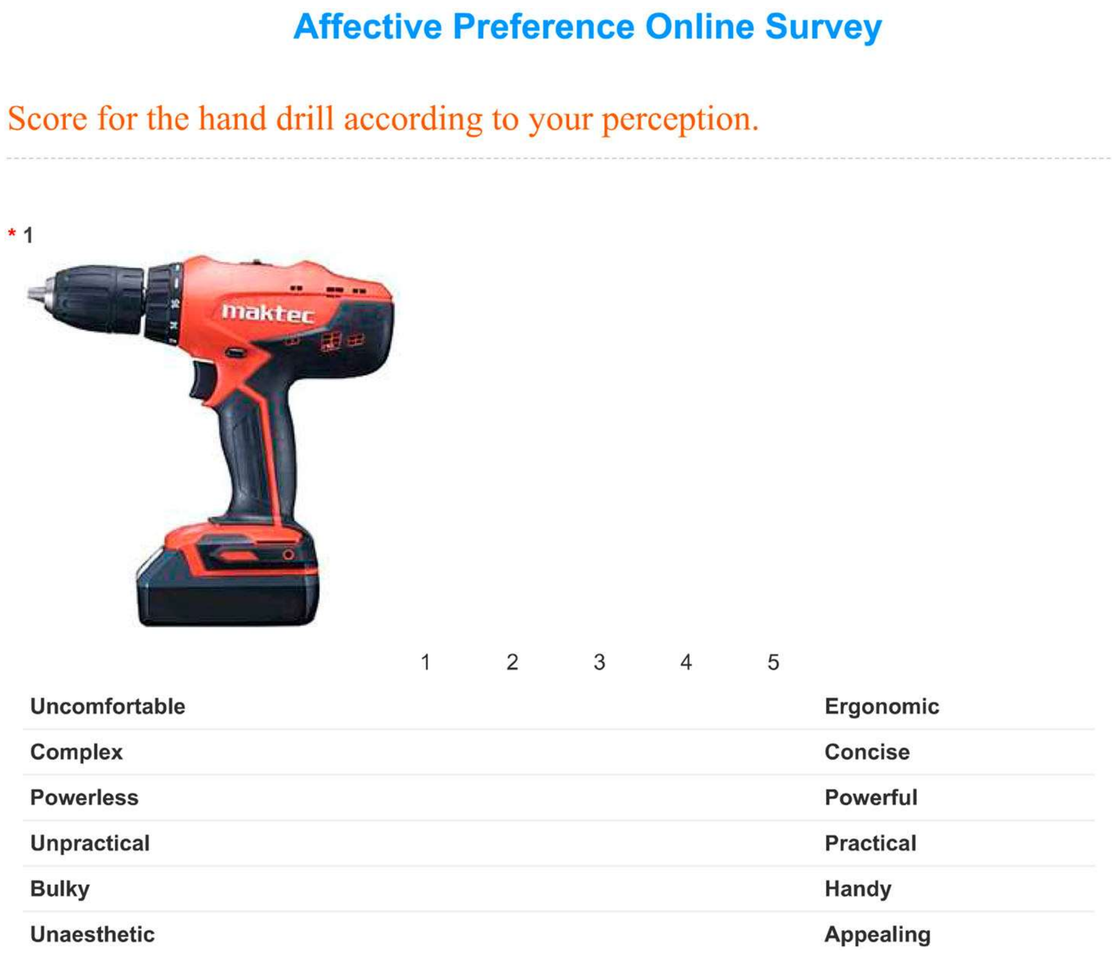


Figure 8. The question corresponding to the first-hand drill.

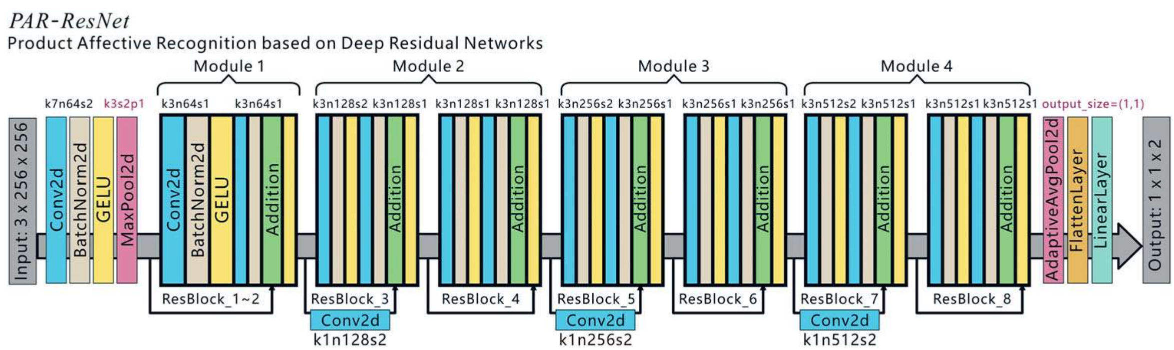


Figure 9. Product affective recognition resnet architecture.

3.2.3. Training and results

Obviously, 840 data samples are not enough. During training, therefore, we used data augmentation techniques to augment the dataset. All 840 images were flipped, including mirror flipping (left and right) and five random rotations (−25 degrees to 25 degrees), but the labels were not changed. To ensure the consistent visual perception of humans and machines, we did not change the colour and cropping on all sample images.

Since there are 6-dimensioanl emotion categories, the affective recognition task in this paper is a 6-dimensional binary classification problem, so the model needs to be trained six times individually.

Table 4. Evaluation results of hand drill affective preferences recognition.

Method	K1 Uncomfortable-Ergonomic	K2 Complex–Concise	K3 Powerless–Powerful	K4 Unpractical–Practical	K5 Bulky–Handy	K6 Unaesthetic–Appealing	Mean (%)
	Recognition accuracy (%)						
Alex-Net	56.34	75.59	74.20	71.82	69.98	70.83	64.83
VGG-11	58.88	73.21	81.52	77.97	72.26	80.95	54.13
VGG-16	63.96	78.42	75.99	83.53	78.51	82.93	77.22
PAR-ResNet	78.57	82.93	92.31	88.49	88.86	91.46	87.10

Table 5. Number of hand drill images for each label.

Tag	Ergonomic	Uncomfortable	Concise	Complex
Number	9457	8828	7881	10,404
Tag	Powerful	Powerless	Practical	Unpractical
Number	8652	9633	8835	9450
Tag	Handy	Bulky	Appealing	Unaesthetic
Number	11,880	6405	8401	9884

PAR-ResNet model is optimized using AdamW optimizer (Loshchilov and Hutter 2017) with $\beta 1 = 0.9$, $\beta 2 = 0.999$. We use a batch size of 32 in the training procedure. The learning rate is initialised to $5\boldsymbol{e} - 2$, and the learning rate decay is initialised to $5\boldsymbol{e} - 1$. The number of max training epoch is set to 50. Figure A2 shows the loss of PARA-ResNet in our dataset.

Given a hand drill image, PAR-ResNet can predict probabilities of belonging to 12 kinds of Kansei labels such as 'concise', 'appealing' and 'powerful'. The average accuracy rate of six Kansei dimensions is 87.10% in validation set (Table 4). Besides, we used PyTorch to construct the Alex-Net, VGG-11 and VGG-16 models for comparison with PAR-ResNet. The recognition accuracy is calculated by the formula (2). The evaluation results are shown in Table 4. Table 5 shows the Kansei labels and the number of HD dataset images corresponding to each estimated label.

3.3. Product design generative adversarial network

3.3.1. PD-GAN architecture

The generator's architecture is shown in Figure 10, which is a modification from ResNet (He et al. 2016) and DCGAN (Radford, Metz, and Chintala 2016). The network contains five Residual Blocks and six Transpose convolutional layer for feature map extraction. We use Gaussian error linear units (GELUs) (Hendrycks and Gimpel 2016) as the activation function layer. Figure 11 shows the discriminator architecture, which includes five Residual Blocks in all. The GELUs (Hendrycks and Gimpel 2016) is also used as the activation function layer in the discriminator. The batch norm layer is replaced by the Instance norm layer in the discriminator. The batch normalisation layer is not used in all Residual Blocks, since it would bring correlations with the mini-batch, which is undesired for the computation of the gradient norm. Additionally, the output includes two branches that perform discriminatory and classification tasks, respectively.

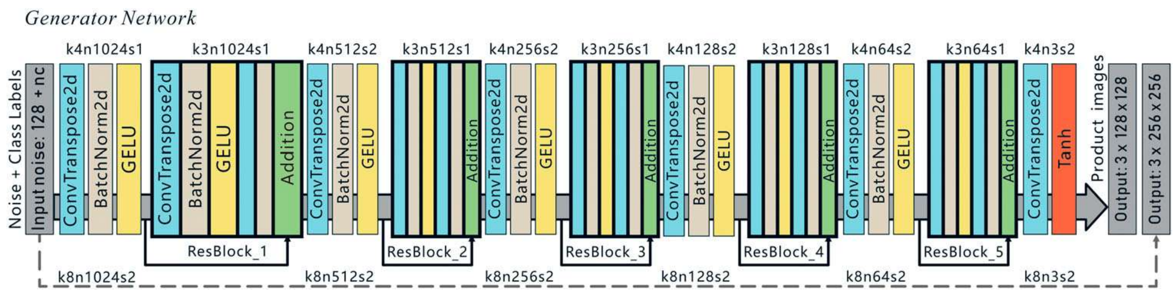


Figure 10. Generator architecture.

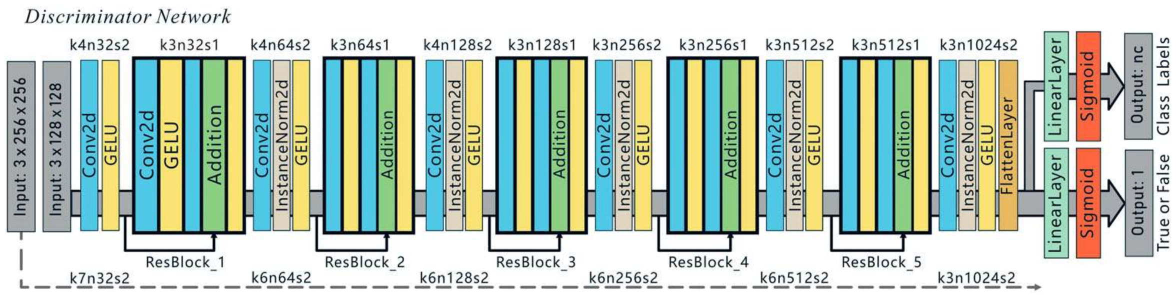


Figure 11. Discriminator architecture.

3.3.2. Network training and results

Noise data is designed to 128 dimensions. The learning rate is initialised to $2\mathbf{e}-4$. Two models (PD-G and PD-D) are optimised using Adam optimiser (Kingma and Ba 2015) with $\beta 1 = 0.5$, $\beta 2 = 0.999$ and learning rate $\eta = 2\boldsymbol{e} - 4$. We use a batch size of 64 in the training procedure. Regarding the super parameters β and λ of the loss function, we refer to the literature (Zhou et al. 2018). Here we set $\beta = 3$ and $\lambda = 0.05$.

Two types of PD-GAN models were successfully trained during our actual experiments. The first type is a hand drill generator with Kansei labels trained based on the hand drill data set with a resolution of 256×256 pixels. The second type is a bicycle helmet generator without labels, it is trained based on the bike helmet dataset with a resolution of 128×128 pixels.

Figure 12 shows an example with Kansei labels. By fixing the random noise part and Kansei labels, the model can generate hand drill images that have similar morphological features. This is a valid proof of the learning ability of the PD-GAN under labelled conditions, showing that our generator can avoid memory training samples.

Figure 13 shows hand drill images generated from the PD-GAN model. It is worth noting that these hand drill images were generated without any labelling conditions. From the perspective of visual layer perception, we found that some of the newly generated hand drill images result from PD-GAN innovative synthesis after learning real product samples. Figure 14 shows an example of bicycle helmet designed by PD-GAN without any Kansei labels. However, there are also new samples that memorise training samples. The main reason for this result is the small number of some samples in the training sample. It is worth emphasising that when training without labels, β is set to 0 and λ is set to 10.

We constructed a questionnaire (Figure 15) to verify whether the hand drills (Figure 12) produced by PD-GAN matches the relevant Kansei label and the level of form innovation. Eight professional designers, eight experts and eight graduate students from industrial

Figure 12. Generated hand drills with random noise and Kansei labels. First column: Unaesthetic. Second column: Practical. Third column: Powerful. Fourth column: Ergonomic.

design were invited to finish the questionnaires. By sorting the data from questionnaires, and calculating the averages, we obtained the results shown in Figure 16.

For result 1, the first column in Figure 12 had the lowest average score (3.7) on the 'Unaesthetic-Appealing' dimension, which is consistent with the Kansei label (Unaesthetic) of the first column in Figure 12. The second column has the highest average score (5.5) on 'Unpractical-Practical' dimension, which is consistent with the Kansei label (Practical) of the second column in Figure 12. The same consistency also happened in the third and fourth columns of Figure 12.

For result 2, the total average of form innovation is 4.4. The average values of the four Kansei categories are 4.2, 4.1, 4.6 and 4.5, which shows that the overall form innovation difference was not obvious. Although the differences between the means for each category were small, this fluctuation indicates that the PD-GAN has a different form of innovation capabilities in different Kansei preference categories.

Figure 13. Generated hand drills with random noise. Left: 256 × 256 pixels. Right: 128 × 128 pixels.

Figure 14. Generated Bicycle Helmets without affective preference labels.

Comparing Result 1 and Result 2, we found that the second conceptual design in Figure 12 had the lowest score for the six Kansei preferences. However, the form innovation score of 4.1 was close to the middle. This apparent difference indicates that the second

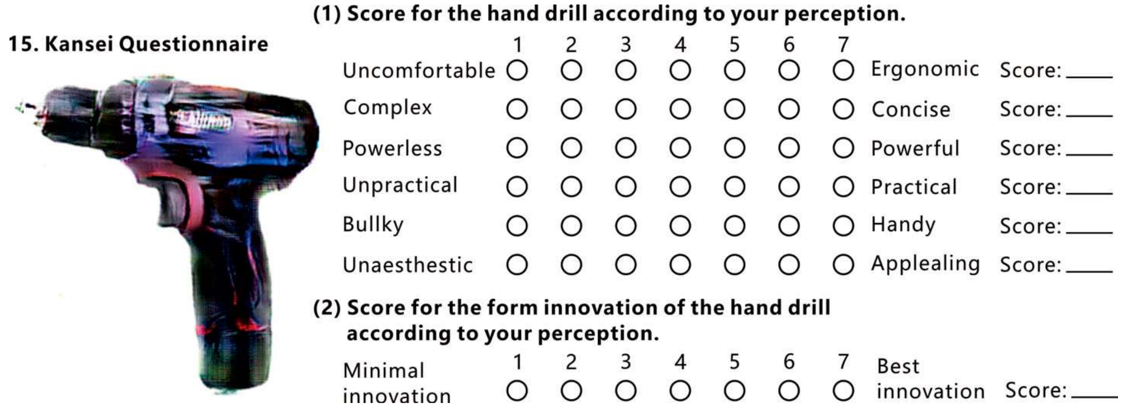


Figure 15. The question corresponding to the concept image of the fifteenth hand drill.

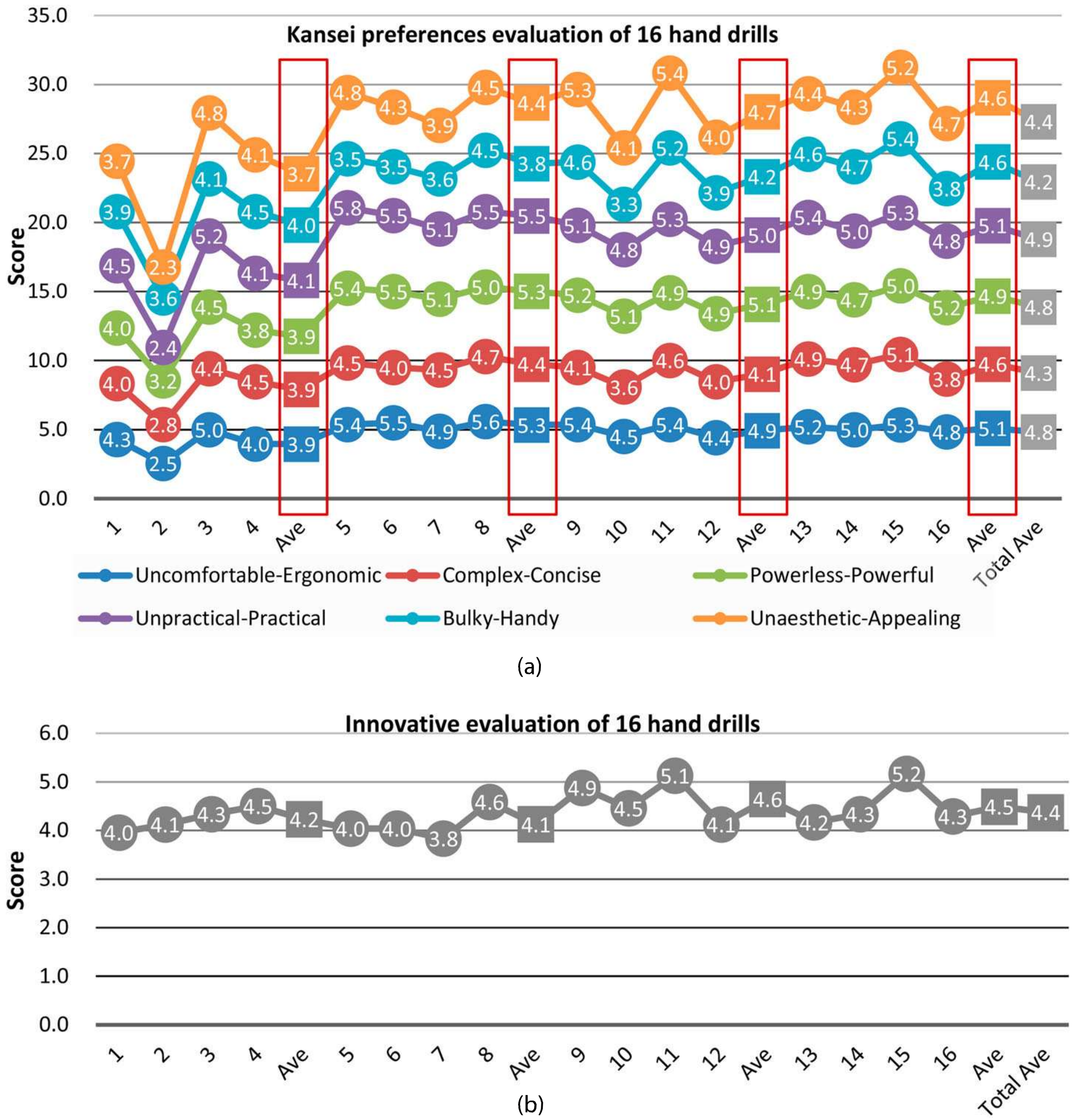


Figure 16. New concept image evaluation results. (a) Result 1: Score results of Kansei preferences (b) Result 2: Score results of form innovation.

conceptual design is the result of the PD-GAN model's imbalance in the game between Kansei preferences and innovation. However, this is not evident in our model.

We also did a simple test just to check whether the product images generated by PD-GAN are instructive for industrial designers. First, two industrial designers were invited to view the product images generated by PD-GAN. Then, they were asked to draw some quick sketches, as shown in Figure 17. Finally, we learned their feelings through interviews. The core points as follows.

> Although these images still look a little fuzzy, and the details are not clear, the forms created by artificial intelligence have created more imagination for designers. Because design usually starts with fuzzy concepts, it is wonderful that these concepts suddenly appear before you.

Therefore, we believe that the combination of deep learning and Kansei engineering can generate product designs and stimulate the innovation of industrial designers.

3.4. *Product style fast transfer*

3.4.1. *PS-FNSTN architecture*

The fast-neural style model's architecture is shown in Figure 18, which is a modification from (Johnson, Alahi, and Fei-Fei 2016). The model contains three parts: encoder (composed of down sampling convolutional layers), deep residual module (12 ResBlocks and a skip connection ResBlock) and decoder (composed of up sampling convolutional layers). We use GELUs (Hendrycks and Gimpel 2016) as the activation function layer.

3.4.2. *Network training and results*

We used the HD dataset and the MS COCO (2014) (Lin et al. 2014) dataset for content images to train our transform network ***fw***. The former was used to train a style transfer model for hand drills, and the latter was used to train a style transfer model for bicycle helmets. There are 80,000 samples in the training data of MS COCO. Although MS COCO has fewer categories than ImageNet and SUN, it has more instances of each category, more like life photos. We think this is more suitable for transfer learning. We used the AdamW optimiser (Loshchilov and Hutter 2017) with default parameters setting ($\beta 1 = 0.9$ and $\beta 2 = 0.999$). We used a batch size of eight content-style image pairs. The learning rate is initialised to $1e - 3$. Content weight and style weight are $1e5$ and $1e10$, respectively. The total variation weight $\gamma = 0$. In the pre-processing, we used TorchVision's Scale and CenterCrop tools to standardise the image to 256×256 or 128×128. However, our transform network ***fw*** does not have any restrictions on the size of the image during the test because it is a fully convolutional neural network (Figure 18). For all product style transfer experiments, we compute content loss at layer *relu*2_2 and style loss at layers *relu*1_2, *relu*2_2, *relu*3_4, *relu*4_4 and *relu*5_4 of the VGG-19 loss network $\boldsymbol{\phi}$ (Figure 6).

Partial results are shown in Figure 19. We inputted the content images (Figure 19(a)) and style images (Figure 19(b)) into our style transform network model to obtain new product style images (Figure 19(c)). For the bike helmet results, we removed the background colour, while the hand drills have no background colour. We found no style image suitable for product style transfer design in our style transfer experiments. We needed to choose the right style image according to the shape and function of the product. Otherwise, the result is not satisfactory, as shown in Figure 20. Because the product image and style image have

(a)

(b)

Figure 17. Quick sketches drawn by two industrial designers. (a): hands drill sketches were drawn by designer A. (b): bicycle helmet sketches were drawn by designer B.

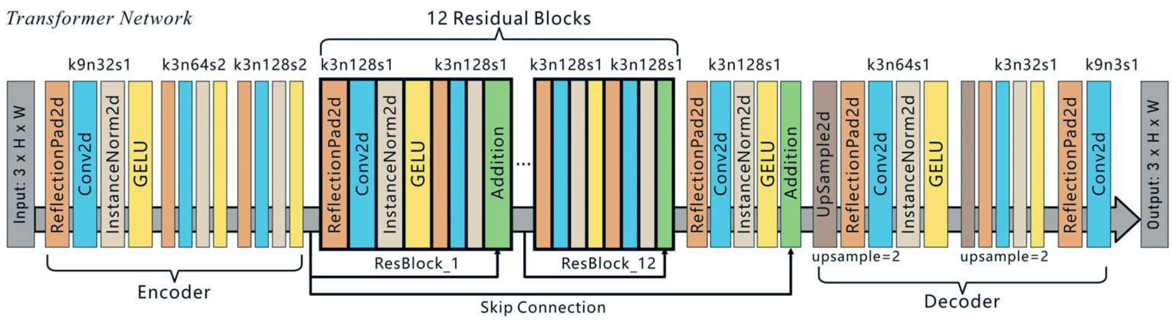


Figure 18. Fast neural style transfer architecture.

Figure 19. The results of product style transfer based on Fast Neural Style Architecture. (a) The content images. (b) The style images. (c) The generated results.

too much loss in content. Besides, fabric and leather products are more suitable for style transfer, which explains why literature (Quan, Li, and Hu 2018) chose women's coats as a research case.

In addition, the training time, and the number of parameters for the above three deep neural network models (i.e. PAR-ResNet, PD-GAN and PS-FNSTN) are shown in Table A1.

4. Discussion

According to the literature (Nagamachi 1995), (Quan, Li, and Hu 2018) and (Wang et al. 2016), there is always a fatal flaw (cannot directly generate products) in product concept

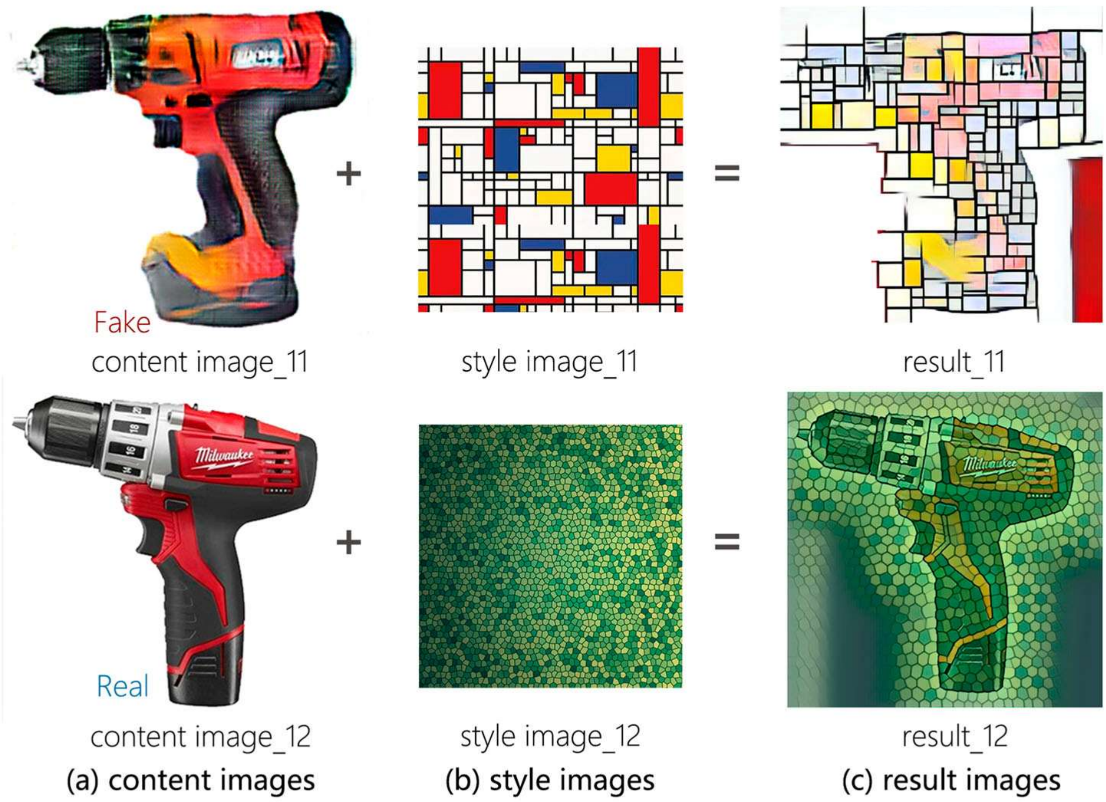


Figure 20. The results of product style transfer based on Fast Neural Style Architecture. (a) The content images. (b) The style images. (c) The generated results.

design based on Kansei engineering. Although the genetic algorithm has been used for product design generation (Hsiao and Tsai 2005; Lo, Ko, and Hsiao 2015), its design space is limited. It is challenging to deal with complicated products like cars and hand drills. Industrial designers still need to deal with specific design implementations based on experience in early design development. Therefore, it is necessary to develop an effective design method to directly generate product conceptual images based on the users' and customers' affective preferences. A product concept generation approach framework based on deep learning and Kansei engineering (PCGA-DLKE) is proposed in this article to automatically recognise affective preferences, to generate designs and to quickly transfer product styles directly.

Comparing the generated results in Figures 12 and 13 from the visual perception level, the results with Kansei preferences in Figure 12 are more beneficial to inspire designers to manoeuver the affective design direction of the product and help improve design efficiency compared to the randomly generated results in Figure 13. The evaluation results in Figure 16(a) show that there is still Kansei error in the hand drill generated under the same Kansei label, but the average of the six Kansei dimensions indicates that PD-GAN has a strong design generation capability. The results in Figure 19 show that style transfer can modify the results generated by PD-GAN and give new style preferences to the results randomly generated by PD-GAN to meet the emotional preferences of consumers.

The advantages of the PCGA-DLKE are described below. Most of the previous works use traditional approaches for product innovation design. Still, it can only provide theoretical guidance for industrial designers and cannot directly generate visible product forms, which is a fatal flaw in industrial product form design. Although the parametric design based on script technology can directly create product forms, this is only the result of the creation of algorithms, and it has no 'learning' process. For the proposed PCGA-DLKE, we build a product form generation system based on deep learning and Kansei engineering techniques, including product images data acquisition methods, product affective preference recognition, automatically generate product images, and product style transfer. As a result, this method constructs a relatively complete generation design system, which makes up for the lack of Kansei engineering and cannot directly generate product conceptual images, while providing a more effective design tool for industrial designers. Therefore, comparing to traditional approaches and digital approaches, the proposed approach not only can maximise user satisfaction, but more importantly, the ability to create product concepts is far superior to previous approaches, greatly improving design efficiency. The case study result shows that PCGA-DLKE is feasible and reasonable for product design generation. However, there is one thing to be reminded of in product style transfer: the results shown in Figure 20 prove that no style image can be used for product style transfer design. At percent, we do not think this is directly related to the architecture and performance of the transform network, mainly due to the functions and attributes of the product itself.

Besides, comparing Figures 13 and 14 at the level of visual perception, we found that most of the hand drill concept images generated by PD-GAN are better than the bicycle helmet images it generates. This is because the morphology of bicycle helmets is inherently more complex. It is more difficult for PD-GAN to learn the data distribution patterns of bicycle helmet images than the hand drill dataset. Furthermore, we found that when training large pixel datasets in our experiments, the residual blocks in the PD-GAN can be removed, which helps to save memory and shorten the training time. Style transfer is the deconstruction and reconstruction of two image features by a deep convolutional neural network. Still, for product design sketches, it requires the industrial designer to give more consideration to the gains and losses of the product's morphological details.

5. Conclusions

Kansei appeal is of critical importance to customer-centric product designs in a competitive market environment. This creates an ongoing challenge for industrial designers as they need to understand the affective preferences that influence users and consumers and comp up with innovative design concepts as quickly as possible.

In this research, we propose PCGA-DLKE, a deep learning-based product concept generation design framework for an industrial designer. Firstly, we use a web crawler to obtain product image data and propose a simple and effective image preprocessing method. Secondly, we construct a tag estimation model to recognise the Kansei label of the product image based on Kansei engineering and deep residual networks. Thirdly, we train a product design GAN model to generate new designs. Finally, we construct a fast-neural style transfer model to create a new style for the previous step. Taking the hand drill and bicycle helmet as examples, we demonstrate the effectiveness and feasibility of PCGA-DLKE. While deep learning-based design method has been used to create anime characters before (Jin et al. 2017; Mattya 2015), our study is the first to combine deep learning with Kansei engineering to assist industrial designers in creating product concept images, which lays a solid foundation for deep learning-based product design.

Deep learning provides a new interface for human–machine co-design, and our proposed approach is an exploration of the human–machine co-design approach. This approach makes up for the shortcomings that Kansei engineering cannot directly generate designs and provides a new innovative design paradigm for industrial designers. The essence of PD-GAN is to use high-dimensional random noise data and Kansei labels to approximate and simulate distribution patterns that approximate a large amount of real product image pixel data, which is why PD-GAN creates new concepts as well as new colour schemes. Our results show that this research framework is indeed able to predict users' affective preferences, and generate innovative product conceptual images and new style, which can effectively help industrial designers break for design fixation. Besides, we show a visual interpretation of the Kansei attributes of PD-GAN-generated conceptual images.

This paper integrates emotion recognition, concept generation and style transformation into an effective intelligent design methodology, which is of high value in a real design environment, not only to improve affective design efficiency, but also to stimulate the potential of designers. The combination of deep learning and Kansei engineering provides an end-to-end intelligent design approach that meets the emotional preferences of the user. This also indicates that designers will have to play the design game with intelligent design systems in the future. That is human–machine design game. This game is reflected in the fact that AI is driven by data and algorithms, while human designers are driven by design knowledge and experience. Artificial intelligence, such as GAN, will help us discover the underlying patterns in human design outcomes. Conversely, the knowledge and experience of professional designers can compensate for the random creation of artificial intelligence. Soon, AI will change the way industrial designers work, their habits of mind, and their design processes.

Although our framework can automatically generate brand-new products with new styles, there are still some drawbacks. For example, colour labels are overlooked in affective recognition, our data preprocessing capabilities are not fully automated, Kansei labels are obtained in a single way, and there is room for improvement in the generative adversarial networks for product design. Besides, the size of the dataset also affects the quality of the generated image. Therefore, in the future, we will focus on improving the data preprocessing capabilities, the accuracy in capturing the user's affective preference, the structure and loss function of generator to generate conceptual images with high affective preferences, and the scale of the extended dataset to improve our framework. This is some very challenging and interesting work for us.

In the research of this project, we also found that there is still much more to explore in industrial design assisted by artificial intelligence. In the cyber-physical network computing environment of the Internet of Everything, Big Data, Artificial Intelligence, Cloud Computing, and Cloud Services, AI will open another window for designers. In the future, therefore, artificial intelligence will be applied to various fields as a general technology, and it is naturally no exception in industrial design.

Acknowledgements

The author first thank own long-term interest in the design innovation and design drawing. Secondly, the author thank to the open source of GAN, CGAN and DCGAN authors on GitHub, our ideas can be implemented efficiently.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by National Natural Science Foundation of China [grant number 51465037].

References

Arjovsky, M., and L. Bottou. 2017, April 24–26. *Towards Principled Methods for Training Generative Adversarial Networks*. Paper presented at the International Conference on Learning Representations (ICLR).

Arjovsky, M., S. Chintala, and L. Bottou. 2017. “Wasserstein Generative Adversarial Networks.” In *Proceedings of the 34th International Conference on Machine Learning*, 214-223 PMLR.

Brock, A., J. Donahue, and K. Simonyan. 2019. “Large Scale GAN Training for High Fidelity Natural Image Synthesis.” In. https://arxiv.org/abs/1809.11096v1: ArXiv e-Print.

Chai, C., J. Liao, N. Zou, and L. Sun. 2018. “A One-to-Many Conditional Generative Adversarial Network Framework for Multiple Image-to-Image Translations.” *Multimedia Tools and Applications* 77 (17): 22339–22366. doi:10.1007/s11042-018-5968-7.

Chan, K. Y., C. K. Kwong, P. Wongthongtham, H. Jiang, Cky Fung, B. Abu-Salih, Z. Liu, T. C. Wong, and P. Jain. 2020. “Affective Design Using Machine Learning: A Survey and its Prospect of Conjoining Big Data.” *International Journal of Computer Integrated Manufacturing*, 33 (7) 645–669.

Chang, Y. M., and C. W. Chen. 2016. “Kansei Assessment of the Constituent Elements and the Overall Interrelations in Car Steering Wheel Design.” *International Journal of Industrial Ergonomics* 56: 97–105. doi:10.1016/j.ergon.2016.09.010.

Chen, H. Y., and Y. M. Chang. 2014. “Development of A Computer Aided Product-Form Design Tool Based on Numerical Definition Scheme and Neural Network.” *Journal of Advanced Mechanical Design, Systems, and Manufacturing* 8 (3): JAMDSM0033–JJAMDSM033..

Chen, L., P. Wang, H. Dong, F. Shi, J. Han, Y. Guo, P. R. N. Childs, J. Xiao, and C. Wu. 2019. “An Artificial Intelligence Based Data-Driven Approach for Design Ideation.” *Journal of Visual Communication and Image Representation* 61: 10–22. doi:10.1016/j.jvcir.2019.02.009.

Chen, C. F., C. H. Yeh, and Y. C. Lin. 2010. A Neural Network Approach to Eco-Product Form Design. Paper presented at the 2010 5th IEEE Conference on Industrial Electronics and Applications, 15-17 June 2010.

Chou, J. R. 2016. “A Kansei Evaluation Approach Based on the Technique of Computing with Words.” *Advanced Engineering Informatics* 30 (1): 1–15. doi:10.1016/j.aei.2015.11.001.

Dong, C., C. C. Loy, K. M. He, and X. Tang. 2014. *Learning a Deep Convolutional Network for Image Super-Resolution. Paper presented at the European Conference on Computer Vision (ECCV)*.

Eigen, D., and R. Fergus. 2015. Predicting Depth, Surface Normals and Semantic Labels with a Common Multi-scale Convolutional Architecture. Paper presented at the 2015 IEEE International Conference on Computer Vision (ICCV), 7-13 Dec. 2015.

Elgammal, A., B. Liu, M. Elhoseiny, and M. Mazzone. 2017. “CAN: Creative Adversarial Networks, Generating “Art” by Learning About Styles and Deviating from Style Norms.” In. https://arxiv.org/abs/1706.07068v1: ArXiv e-Print.

Fung, C. K. Y., C. K. Kwong, K. Y. Chan, and H. Jiang. 2014. “A Guided Search Genetic Algorithm Using Mined Rules for Optimal Affective Product Design.” *Engineering Optimization* 46 (8): 1094–1108. doi:10.1080/0305215X.2013.823196.

Gatys, L. A., A. S. Ecker, and M. Bethge. 2015. “A Neural Algorithm of Artistic Style.” In. https://arxiv.org/abs/1508.06576: ArXiv e-Print.

Goodfellow, I., J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. C. Courville, and Y. Bengio. 2014. “Generative Adversarial Networks.” In. https://arxiv.org/abs/1406.2661: ArXiv e-Print.

Grimsaeth, K., M. Gangwar, B. Baggerud, and R. Narain. 2010. “Kansei Engineering Based Emotional Product Design: An Illustration by a Battery Drill Case-Study.” *Kansei Engineering International Journal* 9 (2): 67–79. doi:10.5057/kei.9.67.

Guo, F., W. L. Liu, Y. Cao, F. T. Liu, and M. L. Li. 2016. “Optimization Design of a Webpage Based on Kansei Engineering.” *Human Factors and Ergonomics in Manufacturing & Service Industries* 26 (1): 110–126. doi:10.1002/hfm.20617.

Guo, F., W. L. Liu, F. T. Liu, H. Wang, and T. B. Wang. 2014. “Emotional Design Method of Product Presented in Multi-Dimensional Variables Based on Kansei Engineering.” *Journal of Engineering Design* 25 (4-6): 194–212. doi:10.1080/09544828.2014.944488.

Hadji, I., and R. P. Wildes. 2018. “What Do We Understand About Convolutional Networks?” In. https://arxiv.org/abs/1803.08834: ArXiv e-Print.

He, K., X. Zhang, S. Ren, and J. Sun. 2016. Deep Residual Learning for Image Recognition. Paper presented at the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 27-30 June 2016.

Hendrycks, D., and K. Gimpel. 2016. “Gaussian Error Linear Units (GELUs).” In. https://arxiv.org/abs/1606.08415v4: ArXiv e-Print.

Hjelm, R. D., A. P. Jacob, T. Che, A. Trischler, K. Cho, and Y. Bengio. 2017. “Boundary-Seeking Generative Adversarial Networks.” In. https://arxiv.org/abs/1702.08431: ArXiv e-Print.

Hsiao, Y. H., M. C. Chen, and W. C. Liao. 2017. “Logistics Service Design for Cross-Border E-Commerce Using Kansei Engineering with Text-Mining-Based Online Content Analysis.” *Telematics and Informatics* 34 (4): 284–302. doi:10.1016/j.tele.2016.08.002.

Hsiao, S. W., and H. C. Tsai. 2005. “Applying A Hybrid Approach Based on Fuzzy Neural Network and Genetic Algorithm to Product Form Design.” *International Journal of Industrial Ergonomics* 35 (5): 411–428. doi:10.1016/j.ergon.2004.10.007.

Huang, Y., C. H. Chen, and L. P. Khoo. 2012. “Kansei Clustering for Emotional Design Using A Combined Design Structure Matrix.” *International Journal of Industrial Ergonomics* 42 (5): 416–427. doi:10.1016/j.ergon.2012.05.003.

Jiao, J., Y. Zhang, and M. Helander. 2006. “A Kansei Mining System for Affective Design.” *Expert Systems with Applications* 30 (4): 658–673. doi:10.1016/j.eswa.2005.07.020.

Jin, Y., J. Zhang, M. Li, Y. Tian, H. Zhu, and Z. Fang. 2017. “Towards the Automatic Anime Characters Creation with Generative Adversarial Networks.” In. https://arxiv.org/abs/1708.05509: ArXiv e-Print.

Johnson, J., A. Alahi, and L. Fei-Fei. 2016. “Perceptual Losses for Real-Time Style Transfer and Super-Resolution.” *European Conference on Computer Vision (ECCV)*, 694-711.

Karras, T., T. Aila, S. Laine, and J. Lehtinen. 2018. “Progressive Growing of GANs for Improved Quality, Stability, and Variation.” In. https://arxiv.org/abs/1710.10196: ArXiv e-Print.

Karras, T., S. Laine, and T. Aila. 2019. A Style-Based Generator Architecture for Generative Adversarial Networks. Paper presented at the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 15-20 June 2019.

Khalid, H. M. 2006. “Embracing Diversity in User Needs for Affective Design.” *Applied Ergonomics* 37 (4): 409–418. doi:10.1016/j.apergo.2006.04.005.

Kim, T., M. Cha, H. Kim, J. Lee, and J. Kim. 2017. Learning to Discover Cross-Domain Relations with Generative Adversarial Networks. Paper presented at the Proceedings of the 34th International Conference on Machine Learning (ICML).

Kim, W., T. Ko, I. Rhiu, and M. H. Yun. 2019. “Mining Affective Experience for A Kansei Design Study on A Recliner.” *Applied Ergonomics* 74: 145–153. doi:10.1016/j.apergo.2018.08.014.

Kingma, D. P., and J. L. Ba. 2015. “Adam: A Method for Stochastic Optimization.” In. https://arxiv.org/abs/1412.6980: ArXiv Preprint.

Kodali, N., J. Abernethy, J. Hays, and Z. Kira. 2017. “On Convergence and Stability of GANs.” In. https://arxiv.org/abs/1705.07215v5: ArXiv e-Print.

Krizhevsky, A., I. Sutskever, and G. E. Hinton. 2017. “ImageNet Classification with Deep Convolutional Neural Networks.” *Communications of The ACM* 60 (6): 84–90. doi:10.1145/3065386.

Kwong, C. K., K. Y. Fung, H. Jiang, K. Y. Chan, and K. W. M. Siu. 2013. “A Modified Dynamic Evolving Neural-Fuzzy Approach to Modeling Customer Satisfaction for Affective Design.” *The Scientific World Journal* 2013: 1–11. doi:10.1155/2013/636948.

Lecun, Y., L. Bottou, Y. Bengio, and P. Haffner. 1998. "Gradient-Based Learning Applied to Document Recognition." *Proceedings of the IEEE* 86 (11): 2278–2324. doi:10.1109/5.726791.
Li, Z., Z. G. Tian, J. W. Wang, W. M. Wang, and G. Q. Huang. 2018. "Dynamic Mapping of Design Elements and Affective Responses: A Machine Learning Based Method for Affective Design." *Journal of Engineering Design* 29 (7): 358–380. doi:10.1080/09544828.2018.1471671.
Li, C., K. Xu, J. Zhu, and B. Zhang. 2017. Triple Generative Adversarial Nets." In. https://arxiv.org/abs/1703.02291: ArXiv e-Print.
Lin, T. Y., M. Maire, S. Belongie, L. Bourdev, R. Girshick, J. Hays, P. Perona, D. Ramanan, C. L. Zitnick, and P. Dollár. 2014. Microsoft COCO: Common Objects in Context. Paper presented at the European Conference on Computer Vision (ECCV), September 2014.
Liu, Q., K. Wang, Y. Li, and Y. Liu. 2020. "Data-Driven Concept Network for Inspiring Designers' Idea Generation." *Journal of Computing and Information Science in Engineering* 20 (3): 1–12. doi:10.1115/1.4046207.
Lo, C. H., Y. C. Ko, and S. W. Hsiao. 2015. "A Study That Applies Aesthetic Theory and Genetic Algorithms to Product Form Optimization." *Advanced Engineering Informatics* 29 (3): 662–679. doi:10.1016/j.aei.2015.06.004.
Loshchilov, I., and F. Hutter. 2017. "Fixing Weight Decay Regularization in Adam." In. https://arxiv.org/abs/1711.05101v3: ArXiv e-Print.
Luan, F., S. Paris, E. Shechtman, and K. Bala. 2017. Deep Photo Style Transfer. Paper presented at the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 21-26 July 2017.
Mao, X., Q. Li, H. Xie, R. Y. K. Lau, Z. Wang, and S. P. Smolley. 2017. Least Squares Generative Adversarial Networks. Paper presented at the 2017 IEEE International Conference on Computer Vision (ICCV), 22-29 Oct. 2017.
Mattya. 2015. "Chainer-DCGAN." In. https://github.com/mattya/chainer-DCGAN.
Mattya. 2017. "Chainer-Gan-Lib." In. https://github.com/pfnet-research/chainer-gan-lib.
Mirza, M., and S. Osindero. 2014. "Conditional Generative Adversarial Nets." In. https://arxiv.org/abs/1411.1784: ArXiv e-Print.
Nagamachi, M. 1995. "Kansei Engineering: A New Ergonomic Consumer-Oriented Technology for Product Development." *International Journal of Industrial Ergonomics* 15 (1): 3–11. doi:10.1016/0169-8141(94)00052-5.
Odena, A., C. Olah, and J. Shlens. 2017. Conditional Image Synthesis with Auxiliary Classifier GANs. Paper presented at the International Conference on Machine Learning (ICML), August 2017.
Osgood, E. C., G. J. Suci, and P. H. Tannenbaum. 1957. *The Measurement of Meaning*. Urbana, IL: University of Illinois Press.
Pan, Y., A. Burnap, J. Hartley, R. Gonzalez, and P. Y. Papalambros. 2017. "Deep Design: Product Aesthetics for Heterogeneous Markets." In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1961–1970. Halifax, NS, Canada: Association for Computing Machinery.
Ponce, P., D. Balderas, T. Peffer, and A. Molina. 2018. "Deep Learning for Automatic Usability Evaluations Based on Images: A Case Study of the Usability Heuristics of Thermostats." *Energy and Buildings* 163: 111–120. doi:10.1016/j.enbuild.2017.12.043.
Quan, H., S. Li, and J. Hu. 2018. "Product Innovation Design Based on Deep Learning and Kansei Engineering." *Applied Sciences* 8 (12): 2397–2317. doi:10.3390/app8122397.
Radford, A., L. Metz, and S. Chintala. 2016. "Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks." In. https://arxiv.org/abs/1511.06434: ArXiv e-Print.
Razza, B., and L. C. Paschoarelli. 2015. "Affective Perception of Disposable Razors: A Kansei Engineering Approach." *Procedia Manufacturing* 3: 6228–6236. doi:10.1016/j.promfg.2015.07.750.
Self, J. A. 2019. "Communication Through Design Sketches: Implications for Stakeholder Interpretation During Concept Design." *Design Studies* 63: 1–36. doi:10.1016/j.destud.2019.02.003.
Simonyan, K., and A. Zisserman. 2015. "Very Deep Convolutional Networks for Large-Scale Image Recognition." In. https://arxiv.org/abs/1409.1556: ArXiv e-Print.

Sutera, J., M. C. Yang, and C. Elsen. 2014. "The Impact of Expertise on the Capture of Sketched Intentions: Perspectives for Remote Cooperative Design." In *International Conference on Cooperative Design, Visualization and Engineering*, 245-252.

Szegedy, C., W. Liu, Y. Q. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich. 2015. Going Deeper with Convolutions. Paper presented at the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 7-12 June 2015.

Vieira, J., J. M. A. Osório, S. Mouta, P. Delgado, A. Portinha, J. F. Meireles, and J. A. Santos. 2017. "Kansei Engineering as A Tool for the Design of In-Vehicle Rubber Keypads." *Applied Ergonomics* 61: 1–11. doi:10.1016/j.apergo.2016.12.019.

Wang, Y., D. Y. Mo, and M. M. Tseng. 2018. "Mapping Customer Needs to Design Parameters in the Front End of Product Design by Applying Deep Learning." *CIRP Annals* 67 (1): 145–148. doi:10.1016/j.cirp.2018.04.018.

Wang, C. C., C. H. Yang, C. S. Wang, T. R. Chang, and K. J. Yang. 2016. "Feature Recognition and Shape Design in Sneakers." *Computers & Industrial Engineering* 102: 408–422. doi:10.1016/j.cie.2016.05.003.

Yanagisawa, H., and S. Fukuda. 2005. "Interactive Reduct Evolutional Computation for Aesthetic Design." *Journal of Computing and Information Science in Engineering* 5 (1): 1–7. doi:10.1115/1.1846055.

Zhai, L. Y., L. P. Khoo, and Z. W. Zhong. 2009. "A Rough Set Based Decision Support Approach to Improving Consumer Affective Satisfaction in Product Design." *International Journal of Industrial Ergonomics* 39 (2): 295–302. doi:10.1016/j.ergon.2008.11.003.

Zhou, Z., H. Cai, S. L. Rong, Y. Song, K. Ren, W. Zhang, J. Wang, and Y. Yu. 2018. "Activation Maximization Generative Adversarial Nets." In. https://arxiv.org/abs/1703.02000v9: ArXiv e-Print.

Appendices

Appendix 1. Data distribution of survey sample images.

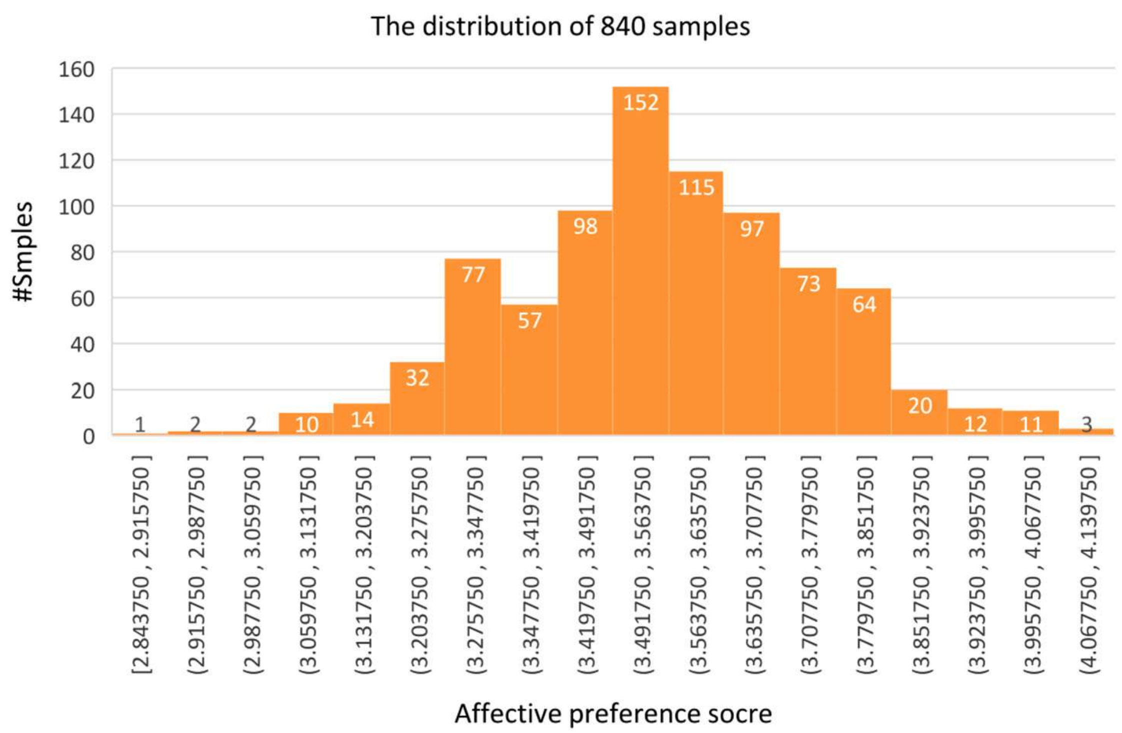


Figure A1. Data distribution of affective preferences (840 samples)

Appendix 2. PAR-ResNet losses on six Kansei preference dimensions.

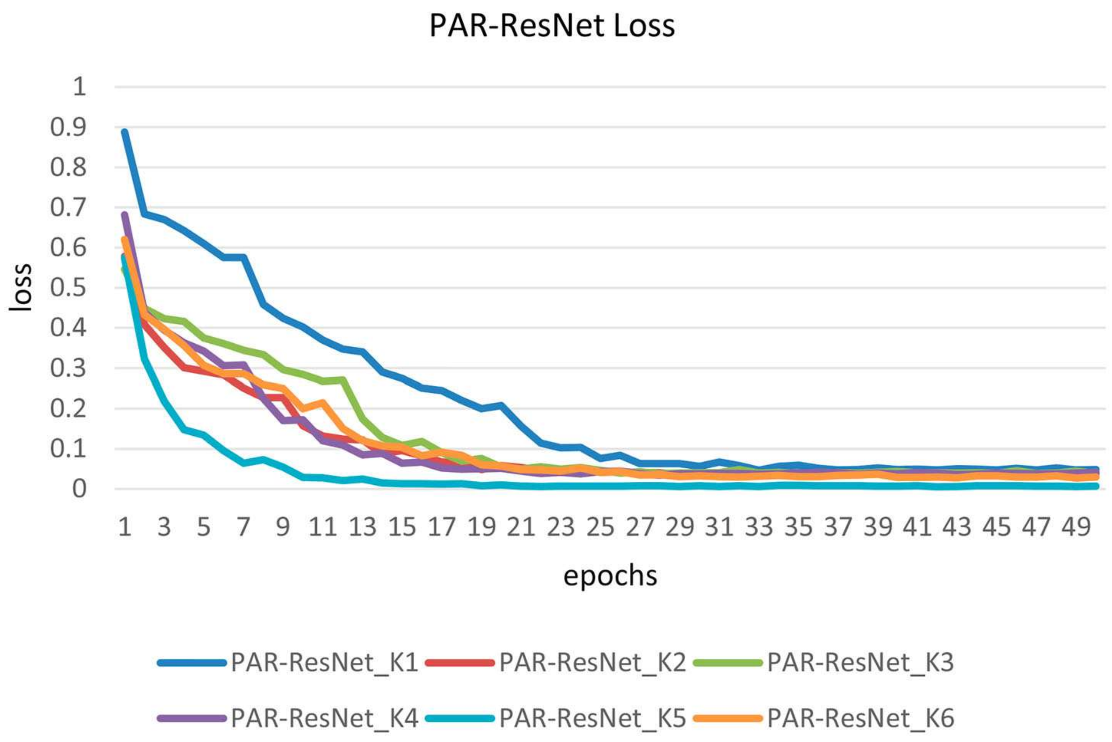


Figure A2. The losses of PAR-ResNet for six Kansei dimensions.

Appendix 3. The training time and the number of parameters of three models (PAR-ResNet, PD-GAN and PS-FNSTN).

Table A1. The training time and the number of parameters of three models.

	Model		Dataset	Max epoch	Number of GPU	Training time (s)	Trainable params	Total params
1	PAR-ResNet		HD dataset	50	Single	avg. 245.46	11,180,546	11,180,546
2	PD-GAN	Generator (No ResBlocks) Discriminator (No ResBlocks)	HD dataset (256 × 256 pixels)	500	Double	96,511.18	52,969,344 11,019,744	52,969,344 11,019,744
		Generator Discriminator	HD dataset (128 × 128 pixels)	500	Double	47232.16	38,394,496 1,26,60,128	38,394,496 12,660,128
		Generator Discriminator	BH dataset (128 × 128 pixels)	500	Double	41278.95	38,394,496 12,660,128	38,394,496 12,660,128
3	PS-FNSTN		MS COO 2014	3	Double	avg. 14546.37	2,714,115	2,714,115
			HD dataset	6		avg. 3948.42		

机械工程学报
Journal of Mechanical Engineering

版面费收取通知

李雄　　　作者（老师）您好：

贵文　基于深度学习的产品概念草图生成设计研究

（稿号 2021-2019 ；录排号 3 ），拟刊登在《机械工程学报》, 须缴纳版面费 8500 元，　请尽快办理。

感谢您的支持与厚爱，欢迎您以后继续投稿，愿我们合作良好！

汇款方式

*提示：烦请您在汇款前，登录系统，在作者工作区填写版面费开票信息。同时，请您登陆系统核对通讯作者（系统中的通讯作者实为联系人）的地址、身份证等信息。如有疑问，请致电010-88379881

银行汇款

行号：102100000144
户名：机械工业信息研究院
开户银行：工行北京百万庄支行
帐号：0200001409014473834
汇款附言：JME 第一作者姓名 稿号 录排号

《机械工程学报》编辑部

机械工业信息研究院
《机械工程学报》编辑部
2023年1月30日

注: 1.随同本通知还将排版后的作者校样发给了您，请在5个工作日内将校对结果提交系统
2. 汇款务必填写附言！如有字数限制，'JME和第一作者姓名'为必填项！发票如未特殊说明发票抬头按照付款人户名开具，询问发票等事宜请致电010-88379881或jme@cjmenet.com

Tel：010-88379907/9909　Fax：010-88379504
http：//www.cjmenet.com.cn；www.cjmenet.com
E-mail：jme@cmes.org
通信地址：北京百万庄大街22号《机械工程学报》编辑部 (100037)

第59卷第11期 机 械 工 程 学 报 Vol.59 No.11
2023 年 6 月 JOURNAL OF MECHANICAL ENGINEERING Jun. 2023

DOI：10.3901/JME.2023.11.016

基于深度学习的产品概念草图生成设计研究*

李 雄[1,2] 苏建宁[1,3] 张志鹏[1]
(1. 兰州理工大学机电工程学院 兰州 730050；
2. 兰州城市学院培黎机械工程学院 兰州 730070)
3. 兰州理工大学设计艺术学院 兰州 730050；

摘要：概念草图设计作为人类的高阶视觉认知活动，是辅助设计师记录、构思、创造和评估想法的重要手段，对生成创新概念具有积极影响。为模拟设计师的这种高阶视觉认知行为，实现智能化辅助创意草图设计，提出一种基于深度学习的产品概念草图智能生成设计集成方法框架，包括端到端的草图设计 GAN(Sketch2Render-GAN)和草图神经风格迁移网络(Sketch-NST)两个核心模块。前者实现概念草图生成与渲染，后者执行草图风格特征变换。分别以手电钻和自行车头盔为实验对象进行了验证，结果表明该方法框架可快速获得大量具有创新概念的草图，并实现草图自动渲染及风格变换。有助于辅助设计师在视觉认知层面突破设计固化，提高设计效率。此外，为改善工业设计师与 AI 模型间的人机设计协作模式，还开发了智能草图设计生成器(S-SDG_v0.1)，从而有效降低设计师应用智能算法辅助设计的门槛。

关键词：视觉认知；深度学习；生成设计；草图设计；生成对抗网络
中图分类号：TG156

Product Conceptual Sketch Generation Design Using Deep Learning

LI Xiong[1, 2] SU Jianning[1, 3] ZHANG Zhipeng[1]
(1. School of Mechanical & Electronical Engineering, Lanzhou University of Technology, Lanzhou 730050;
2. School of Bailie Mechanical Engineering, Lanzhou City University, Lanzhou 730070;
3. School of Design Art, Lanzhou University of Technology, Lanzhou 730050)

Abstract：Concept sketching, as a higher order human visual cognitive activity, is an important tool to assist designers in recording, ideating, creating, and evaluating ideas and has a positive impact on the generation of innovative concepts. In order to simulate this higher order visual cognitive behaviors of designers and to achieve intelligent assistance in creative sketching, a deep learning-based design integrated framework for intelligent generation of product concept sketches is proposed, which includes two core modules: an end-to-end sketch design GAN (Sketch2Render-GAN) and a sketch-neural style transfer network (Sketch-NST). The first module implements sketch generation and rendering, while the second performs sketch style features transformation. The hand drill and bicycle helmet were used as design objects respectively, and experimental results show that the proposed approach framework can quickly obtain many innovative concept sketches and implement automatic sketch rendering and style transformation. The findings also show that the approach framework helps designers to break through design solidification at the visual perception level and increase design efficiency. Furthermore, a smart-sketch design generator (S-SDG_v0.1) was developed to facilitate human-machine design collaboration between designers and AI models, which effectively reduces the threshold of designers to apply intelligent algorithms to assist design.

Key words：visual cognition；deep learning；generative design；sketch design；GAN

* 国家自然科学基金资助项目(52165033)。20211119 收到初稿，20221226 收到修改稿

0　前言

工业设计师最初的设计概念常以视觉化的概念草图呈现，但创意概念的生成往往包含了一个思维循环的所有要素，包括产生若干个构思、各种概念以及有创意、有价值的想法[1]。草图作为人类视觉化推理的高阶思维认知活动，对设计师具有重要作用。同时，也是与团队及客户高效沟通的重要手段之一。草图设计过程集中体现着设计师左右脑互动构思的认知过程，并通过视觉化的输出进行演绎与推理。尽管计算机辅助工业设计软件(如 Rhino、Alias、Creo 等)已在设计实践中得到广泛应用，但长期以来设计之初的构思和概念设计依旧由草图(传统纸笔或数字草图)的形式主导[2]。特别是在早期概念设计阶段，设计师会频繁通过手绘草图的方式表达想法，以期自我沟通。然而，在人工智能(AI)时代，这种高阶的视觉认知活动能否用现有的人工智能技术模拟，并有效提升设计效果及概念收敛的速度和质量是近期探讨的热门问题[3-5]。

随着人工智能、机器学习的迅猛发展，深度神经网络算法的不断演进加速了机器的智能，图像生成问题的解决方案不断涌现。特别是 GOODFELLOW[6]于 2014 年发明生成对抗网络(Generative adversarial networks，GANs)后，深度生成算法研究及应用成为一个热点话题，被用于图像生成(如手写数字、人脸、室内场景等[7])、3D 建模[8]、视频预测[9]、文本到图像合成[10]等智能生成任务。与此同时，GAN 的改进常从网络结构、损失函数、条件技术和约束方法四个方面进行，可有效提升 GAN 模型的稳定性和收敛性[11]。通过对大量图像数据的学习，GAN 可生成新的图像。因此，GAN 的发明不仅激发了设计研究者的兴趣，而且加速了计算创意的研究进程。通过学习众多已有的设计方案，进而生成具有创意的概念设计草图是一项有价值的研究工作，也是设计智能领域的一个潜在应用。

近年来，一些研究者已将 GAN 应用于产品、平面及建筑等可视化概念设计中，并取得了一定的效果。LIU 等[12]使用两个 GAN 模块进行椅子图像生成设计，第一个 GAN 用于椅子图像生成，第二个 GAN 用于提高图像分辨率，但由于所收集的椅子图像视角多变导致生成的概念图像缺乏细节。DAI 等[13]以智能手表为设计对象，同时使用三个 GAN 的变体(DCGAN[7]，LSGAN[14]、WGAN[15])进行智能手表概念图像生成设计。SAGE 等[16]提出基于 GAN 的 Logo 智能设计方法。JIN 等[17]利用改进的条件生成对抗网络(C-GAN[18])进行动漫头像生成设计。KIM 等[19]使用 GAN 的变体实现关联生成设计，如将鞋子图像作为输入，模型可生成相关联的鞋子图像，具有关联创造性的能力。CHAI 等[20]利用 GAN 实现了鞋类产品草图的自动着色。PAN 等[21]使用可扩展的深度学习方法预测和解释客户对异构市场设计属性的客户美学感知，并利用 DCGAN 生成汽车概念图，但形态美学预测精度有待提升。刘跃中等[22]提出运用 C-GAN 处理跨学科陌生数据的方法，以支持城市设计过程。这些研究验证了 GAN 在概念设计图像生成方面的潜力，进一步拓展了应用领域。

上述研究是直接采用 GAN 及其变体进行图像到图像的创意生成，图像一次融合了产品的形态、功能、色彩等内容，但这种生成模式与设计师的视觉思维过程不够一致。通常，设计师在设计思维活动中借助草图进行推理和可视化表达，从产品形态、功能、结构、比例、位置关系等入手，进行产品草图绘制，然后考虑色彩、材质、纹理等信息。图像到图像的直接创意过程显然与设计师草图推理过程不符。在视觉化思维活动中，人的思维是由简单到复杂、局部到整体的过程，一次性引入多种复杂信息不仅会增加设计师的认知负荷，而且还会增加网络负荷及网络训练难度。另外，有学者在创意生成设计中为获取更多视觉刺激，往往预先制作风格各异的情绪看板(Mood Board)，以帮助他们跨域引入其他风格特征。例如，CHEN 等[1]提出使用基于双判别器的 GAN 模型进行产品概念融合创新设计，并以勺子数据集和树叶数据集为判断对象进行概念融合创新生成。在计算创意策略上采用语义关联，并融合刺激图像产生了创意方案，但在该创作模式下模型需要学习多种复杂内容，导致生成的结果缺乏设计细节。

针对以上问题，本文从计算创意的角度探索产品概念草图智能设计过程。通过解读设计师概念草图设计认知过程，并基于深度学习方法构建了一个新的产品概念草图智能设计集成方法框架，实现产品创意草图生成设计和风格特征迁移。

1　基本理论

1.1　生成对抗网络

GAN 由 2 个交替训练的深度神经网络构成，属于一种混合模型。一个为生成模型 G(Generative model)，用于捕获数据分布；另一个是判别模型 D(Discriminative model)，用于判别数据的真伪。其发明者 GOODFELLOW[6]以造伪钞为例解释 GAN 的原理：生成模型类似于一个假币伪造团伙，试图制造伪钞并在不被发现的情况下使用；判别模型则相当于警察，他们试图检测假币。这种对抗思想源自博弈论(Game theory)中两个玩家的零和博弈(Two-player game，两人的利益之和为零，一方所失正是对方所得)，对抗中的双方根据对方的策略不断变换自己的策略集直到双方达到“纳什均衡”，这是一种动态“博弈”过程。因此，GAN 被视为一种生成逼近真实世界数据的工具，已应用到处理图像、文本、视频、音频等数据类型[7, 9-10]。其目标函数为

$$\min_G \max_D V(D,G) = \boldsymbol{E}_{x\sim P_{data}(x)}[\log D(x)] + \boldsymbol{E}_{z\sim P_z(z)}[\log(1-D(G(z)))] \tag{1}$$

GAN 的算法框架如图 1 所示，$G(z)$ 是生成模型创造的设计方案。其中生成模型 G 依据输入的随机噪声 z 生成样本 $G(z)$；判别模型 D 判别的数据来自真实世界的样本，而不是生成模型 G 产生的伪数据。G 和 D 一般采用高度非线性的深度神经网络结构。

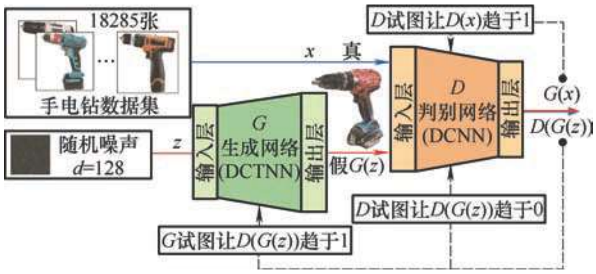


图 1　GAN 算法框架(以手电钻生成为例)

1.2　风格迁移

风格迁移是一种将目标图像的风格转移到某个指定图像上，并最大限度的保持指定图像在内容上保持不变的图像处理技术。早期的风格迁移主要针对图像的局部特征，应用数学模型或统计模型分析对象。一般包括两种类型，其一为纹理生成，有纹理映射、过程纹理合成和基于样图的纹理合成[23]三种实现方法，其缺点是基础工作过于复杂；另一种为单一风格转换，如油画风格迁移[24]、人物肖像风格迁移[25]，其缺点是一种算法只能单独作用于某一种特定风格，很难推广到多变的应用场景。

基于深度卷积神经网络的风格转换技术始于 2015 年，由 GATYS 等[26-27]提出，并证明了图像内容和风格的表示是可被分离操作，可以生成高感知的艺术风格图像和高感知的自然纹理，称为神经风格迁移(Neural style transfer，NST)。该方法突破了之前无法变换应用场景的缺陷，此后风格迁移便与卷积神经网络密不可分。尽管此方法可以生成高感知的风格图像，然而他们的方法实现过程过于复杂且以耗时耗内存为代价，从而限制了其实际应用。因此，2016 年 JOHNSON 等[28]提出了一种快速风格迁移算法，即快速风格迁移网络(Fast neural style transfer network，F-NST)，由风格迁移网络和损失网络两部分组成，如图 2 所示。损失网络通常用来计算视觉特征和风格特征，采用预训练的 VGG-16，此架构影响了后续诸多神经风格迁移研究。本文采用此架构的变体 AdaIN 进行草图风格迁移任务。

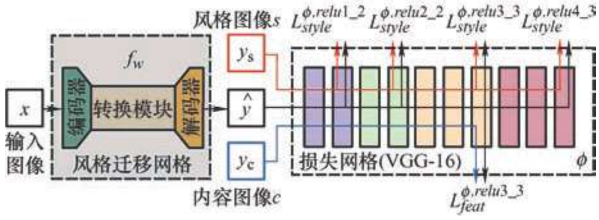


图 2　快速风格迁移(F-NST)网络架构

2　方法

2.1　方法框架

对于设计师来说，在概念未明确之前绘制草图是一种自然、直觉的可视化推理方式。为实现让 AI 模型模拟工业设计师的草图创作过程，构建了一个混合集成方法框架，共有三个关键步骤：草图数据准备、端到端草图生成设计、草图神经风格迁移，如图 3 所示。首先是创建草图数据集；然后执行草图生成与渲染，由 Sketch2Render-GAN 来实现，是由 SketchGAN 和 RenderGAN 组成的两个平行训练的 GAN 联合体，前者用于草图生成设计，后者则是对前者设计输出的草图进行着色渲染，其中还包括草图聚类分析；最后执行草图风格特征变换，由草图神经风格迁移网络(Sketch-NST)来实现，可弥补 Sketch2Render-GAN 无法像设计师那样跨域映射设计特征要素(如纹理、形态、色彩)的缺陷。

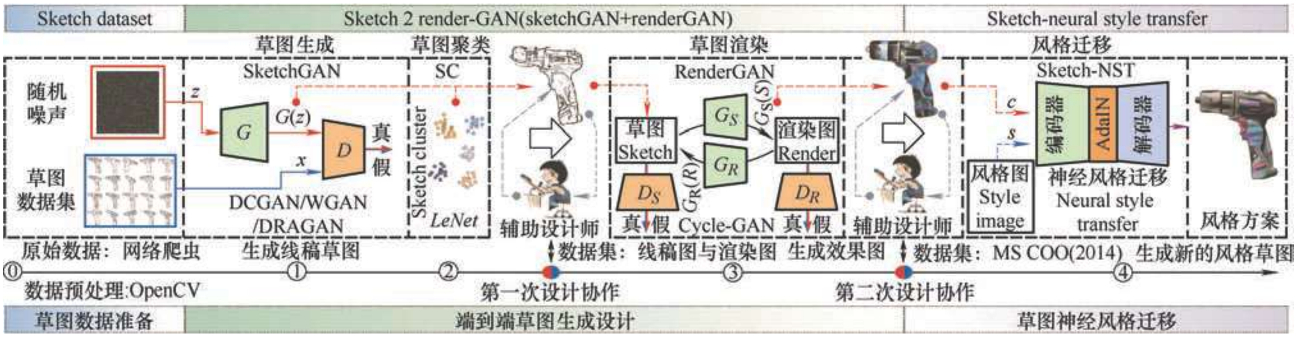


图 3　基于深度学习的产品概念草图智能设计集成方法框架

2.2　数据准备

如同人类设计师通过临摹学习草图设计一样，AI 模型同样需要学习用的草图样本，即草图数据集。因此，如何获取训练用的草图数据集成为第一个关键点。通常包括两个核心内容，即获取原始的产品图像数据和数据预处理。原始数据通过网络爬虫和人工方式获取在线网店和各种搜索引擎上的产品图像数据。产品图像数据预处理包括人工筛选(删除低像素图片)、清洁数据(背景处理成白色只保留产品图像)、裁剪、调整大小(大于 256×256 像素)和数据增广，最后应用开源计算机视觉库 OpenCV 将产品图像自动转换成线稿草图，并进行降噪处理。

2.3　端到端草图设计

设计之初为探索解决问题的新概念，工业设计师会绘制一系列的创意草图并有选择性的着色渲染以过滤概念，直到有新的创新概念产生，这是一种连续性的产品草图设计思维模式。所过滤出的创意草图是大脑逻辑与非逻辑即时非线性思维互动的结果，既包括抽象的逻辑思维过程，也包含了直觉、形象等非逻辑思考。所构建的 Sketch2Render-GAN 是应用 AI 算法模拟设计师的连续性互动视觉认知模式，从而实现从概念草图生成到着色渲染的一站式产品草图生成设计。

2.3.1　草图生成

长期以来，制图与绘画是人们交流的重要方式。然而设计草图是有别于绘画与工程制图的视觉推理语言，是一种设计师自我及与他人探讨、推理和交流想法的高阶视觉认知活动，也是设计师快速可视化所见、所思、所想的直观方式[29]。它需要在左脑逻辑思维与右脑形象思维的互动下，做到心手合一，并强调由前者引导后者。SketchGAN 由两个神经网络构成，判别网络和生成网络。判别网络接收真实数据负责逻辑判断，类似人的左脑；生成网络接收高维随机噪声，负责捕获概念形象的数据分布与重构，类似人的右脑。先训练判别网络后训练生成网络，即判别网络引导生成网络，如此互动博弈，这与设计师左脑逻辑引导右脑形象的草图设计过程类似。

SketchGAN 将从现有草图数据集中学习，并将高维随机噪声数据映射成新的概念草图。为此，SketchGAN 可选择的 GAN 模型有如 DCGAN[7]、LSGAN[14]、WGAN[15]、WGAN-gp[30]、C-GAN[18]等。它们是研究者从网络结构、损失函数、约束条件等方面对原始 GAN 的不同改进。为更适应草图样本生成，基于 DCGAN 设计了新的网络结构及参数，具体见 3.2 节。损失函数由 WAGN[15]与 DRAGAN[31]构成。

SketchGAN 的判别器 D 执行的是一个二分类任务，用于判断每个草图样本的真或假。它的对抗损失函数为

$$L_{adv}(D) = -\boldsymbol{E}_{s\sim P_{data}(s)}[\log D(s)] - \boldsymbol{E}_{z\sim P_z(z)}[\log(1-D(G(z)))] \tag{2}$$

式中，z 是高维随机噪声，s 是草图样本数据。

此外，为防止梯度消失，稳定判别器，需对其进行梯度惩罚，惩罚项 $L_{gp}(D)$ 来自 DRAGAN[31]

$$L_{gp}(D) = \boldsymbol{E}_{\hat{s}\sim P_{perturbed\_data}(\hat{s})}[(\|\nabla_{\hat{s}} D(\hat{s})\|_2 - 1)^2] \tag{3}$$

式中，$\hat{s}$ 代表被扰动的草图数据。

综合式(2)、(3)，判别器 D 的总损失为

$$L(D) = \lambda_{adv} L_{adv}(D) + \lambda_{gp} L_{gp}(D) \tag{4}$$

式中，λ_{adv} 是对抗损失平衡因子，λ_{gp} 是梯度惩罚平衡因子。

SketchGAN 生成器 G 的输入为 128 维的随机噪声 z，其损失函数为

$$L(G) = \boldsymbol{E}_{z\sim P_z(z)}[\log(D(G(z)))] \tag{5}$$

所构建的 SketchGAN 与基本 GAN[6]类似，具有双网络结构和循环交替训练模式，可用于模拟设计师的“读”-“写”草图认知过程[32-33]，判别器 D

负责读取草图数据，生成器 G 输出概念草图；此外，最新研究发现设计师绘制设计草图过程中存在认知组块的经验性证据[2]，可通过深度神经网络聚类草图进行可视化模拟与分析，建构的认知模型如图 4 所示，详细结构见 3.3 节。

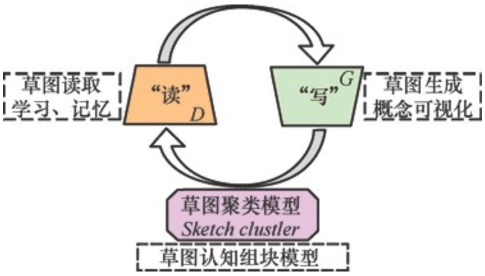


图 4 AI 草图视觉认知模型

另外，模型学习到真实的产品草图数据特征分布后，用训练的 SketchGAN 设计出新的产品概念方案草图，进行第一次人机概念草图设计协作，以便设计师在视觉认知层面突破设计固化。

2.3.2 草图渲染

RenderGAN 的目标是实现从草图到效果图的跨域映射，即将 SketchGAN 设计输出的草图翻译成效果图方案。RenderGAN 可选择的模型有 Pix2Pix[34]，Pix2PixHD[35]，Cycle-GAN[36]等，它们是 GAN 在图像翻译(image-to-image)方面的代表性成果。其中，Cycle-GAN 是一个由两个 GAN 组成的循环结构的神经网络，一个 GAN 负责将草图转化成渲染图，另一个 GAN 则与之相反，如图 5 所示。手绘设计草图作为视觉推理的一种手段，是一个非静态的过程[2, 29]，设计师将大脑中意象通过手绘草图的方式呈现在纸面上，但为解决问题，设计师不断迭代之前的方案并探索新的可能，因此手绘草图是一个循环互动过程，即草图设计迭代循环模式[37]。Cycle-GAN 的循环机制与设计师绘制草图的循环互动具有类似的特点。

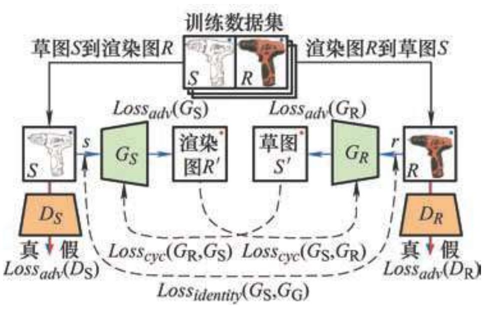


图 5 RenderGAN(Cycle-GAN)算法框架

RenderGAN 选择 Cycle-GAN，由两个相同的 GAN 构成。第一个 GAN 用于从草图 S 到渲染图 R 的映射，与之相反，第二个 GAN 实现由渲染图 R 到草图 S 的映射，训练过程属于双对抗训练模式，具体网络结构见 3.4 节。因此，数据集也由两部分构成，草图数据集和渲染图数据集。训练完成后，效果图由草图渲染器 G_S 实现。损失函数包括三部分，循环一致损失 L_{cyc}、对抗损失 L_{adv} 和恒等损失 $L_{identity}$，如图 5 所示。

两个生成器 G_S 和 G_R 的损失函数为

$$L_{cyc}(G_S,G_R)=\boldsymbol{E}_{s\sim P_{data}(s)}[\|G_R(G_S(s))-s\|_1]+\boldsymbol{E}_{r\sim P_{data}(r)}[\|G_S(G_R(r))-r\|_1] \tag{6}$$

$$L_{adv}(G_S)=\boldsymbol{E}_{s\sim P_{data}(s)}[(D_R(G_S(s))-1)^2] \tag{7}$$

$$L_{adv}(G_R)=\boldsymbol{E}_{r\sim P_{data}(r)}[(D_S(G_R(r))-1)^2] \tag{8}$$

$$L_{identity}(G_S,G_R)=\boldsymbol{E}_{r\sim P_{data}(r)}[\|G_S(r)-s\|_1]+\boldsymbol{E}_{s\sim P_{data}(s)}[\|G_R(s)-s\|_1] \tag{9}$$

综合式(6)~(9)，两个生成器的总损失 $L(G)$ 为

$$L(G)=\alpha L_{cyc}(G_S,G_R)+L_{adv}(G_S)+L_{adv}(G_R)+\beta L_{identity}(G_S,G_R) \tag{10}$$

式中，α 是循环平衡因子，β 是恒等平衡因子。

由于判别器负责判断真伪，只有对抗损失，所以 D_S 和 D_R 的损失函数分别为

$$L_{adv}(D_S)=\boldsymbol{E}_{s\sim P_{data}(s)}[(D_S(s)-1)^2]+\boldsymbol{E}_{r\sim P_{data}(r)}[(D_S(G_R(r)))^2] \tag{11}$$

$$L_{adv}(D_R)=\boldsymbol{E}_{r\sim P_{data}(r)}[(D_R(r)-1)^2]+\boldsymbol{E}_{s\sim P_{data}(s)}[(D_R(G_S(s)))^2] \tag{12}$$

综合式(11)和(12)，两个判别器的总损失 $L(D)$ 为

$$L(D)=\gamma L_{adv}(D_S)+\delta L_{adv}(D_R) \tag{13}$$

式中，γ，δ 是判别器的平衡因子。

卷积神经网络是受视觉感受野机制启发而提出。基于深度卷积神经网络的 Sketch2Render-GAN 框架构建出一个端到端的草图视觉推理模型。其中 SketchGAN 集中体现在产品草图特征线的生成博弈上，RenderGAN 更多的是特征线与色彩方案的视觉循环互动，因此需进行第二次概念草图人机设计协作。

2.4 草图神经风格迁移

神经风格迁移是一种使用卷积神经网络自动将某个图像的风格变换到另一图像上的技术[38]。因此，草图风格迁移需要两张图像重构输出新图像，一张是内容草图，另一张是风格图像。深度卷积神

经网络修改内容草图使其在风格上接近风格图像。此外，归一化技术已经成为深度卷积神经网络的重要组成部分，对优化神经网络参数、提高泛化性能有着重要作用[39]，现已发展成独立神经网络层。例如批归一化层(Batch normalization layer，BN)[40]、实例归一化层(Instance normalization layer，IN)[41]、条件实例归一化层(Conditional instance normalization layer，CIN)[42]和自适应实例归一化层(Adaptive instance normalization layer，AdaIN)[43]常常被用于神经风格迁移网络。BN 层批量计算图像样本每个通道的均值和方差，造成单个样本独特细节的丢失，而 IN 层单独计算每个样本和通道的均值和方差，不依赖批量。CIN 是对 IN 超参数的探索，使用不同的超参数组合可生成不同风格的图像。与前三种归一化技术相比，由于训练过程不需要学习任何参数，AdaIN 层则更简单有效，可将内容特征的均值和方差与风格特征的均值和方差对齐。

因此，草图风格迁移网络 T 采用“编码器-AdaIN-解码器”结构[43]，框架如图 6 所示。编码器 f 采用预训练的 VGG19 的前 9 层，AdaIN 层没有需要学习的参数，只需要训练解码器 g，具体结构见 3.5 节。

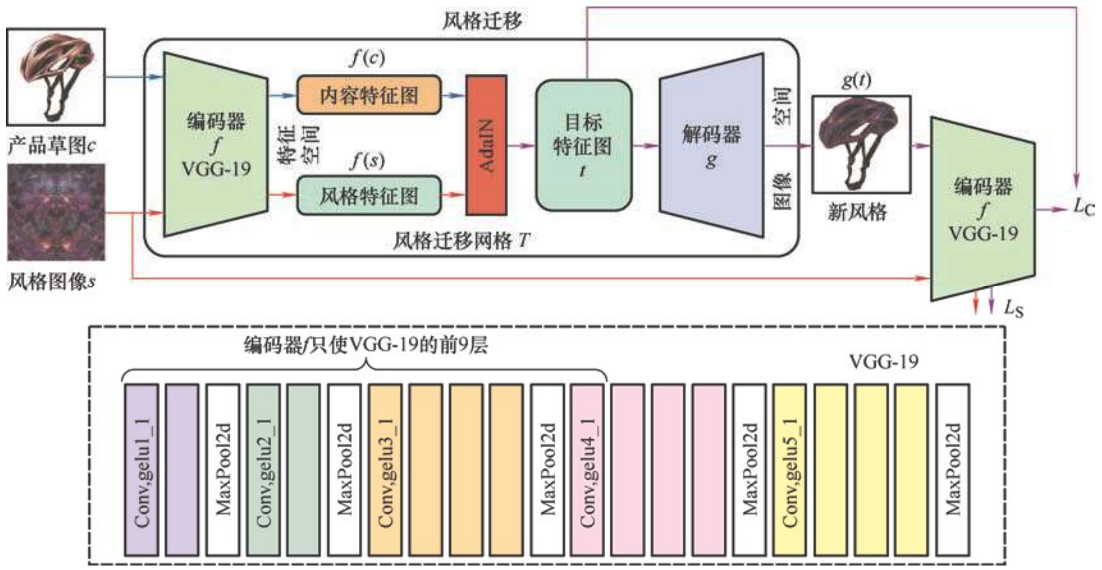


图 6 神经风格迁移网络算法框架

产品图像 c 和风格图像 s 作为风格迁移网络 T 的输入，经过“编码器-AdaIN-解码器”结构计算出一个具有风格 s 的产品创意草图。首先，预训练模型 VGG19(即编码器 f)对产品图像 c 和风格图像 s 进行编码后得到 $f(c)$ 和 $f(s)$。接着，AdaIN 层在特征空间执行风格迁移，将设计草图特征图的均值和方差与风格特征图的均值和方差在通道上对齐，从而生成目标特征 t

$$t = AdaIN(f(c), f(s)) = \sigma(f(s))\left(\frac{f(c)-\mu(f(c))}{\sigma(f(c))}\right)+\mu(f(s)) \qquad (14)$$

式中，μ 是特征图的均值，σ 是特征图的方差。

最后，训练一个随机初始化解码器 g 将目标特征 t 映射到图像空间，从而生成一个具有风格 s 的产品创意草图方案 $T(c,s)$

$$T(c,s) = g(t) \qquad (15)$$

使用预训练的 VGG-19 计算损失函数来训练解码器 g，损失函数为

$$L_{total} = \alpha L_C + \beta L_S \qquad (16)$$

式中，L_{total} 是 T 的总损失，L_C 是草图的内容损失，L_S 是风格损失，α 和 β 分别是两种损失的平衡权重。

目标特征 t 与输出图像的特征 $f(g(t))$ 之间的欧式距离是概念草图的内容损失

$$L_C = \| f(g(t)) - t \|_2 \qquad (17)$$

由于 AdaIN 层仅传递风格特征的均值和方差，因此风格损失可用下式计算

$$L_S = \sum_{i=1}^{L} \| \mu(\phi_i(g(t))) - \mu(\phi_i(s)) \|_2 + \sum_{i=1}^{L} \| \sigma(\phi_i(g(t))) - \sigma(\phi_i(s)) \|_2 \qquad (18)$$

式中，ϕ_i 表示 VGG-19 中计算风格损失的层。下文实验中的风格损失从 $GELU1\_1,2$、$GELU2\_1,2$、$GELU3\_1,2,3$ 和 $GELU4\_1$ 八个层中提取。

3　实验及结果分析

以手电钻为设计对象，基于上述方法框架进行创意草图生成设计实验。分三阶段进行，首先是数据准备；其次是手电钻概念草图设计，包括概念草图生成与渲染；最后是草图风格迁移。

所提方法框架使用 Python 语言开发。开发平台为：双 GPU(RTX2080/8G)，Ubuntu18.04，PyTorch1.6，OpenCV3.4.2，Visdom。同时，使用 Dear PyGui 开发了一款智能草图设计生成器(S-SDG_v0.1)。

3.1　数据集及实验安排

依据前文 2.2 节所述方法创建了一个手电钻草图数据集，包括三种类型的手电钻共有 13 027 张草图(见表 1)。

表 1　草图数据集示例

样本类型	样本数量/张	草图样例	产品样例
直流 带电池底座	6402 (256×256像素)		
直流 无电池底座	5587 (256×256像素)		
直流 电动螺丝刀	1038 (256×256像素)		

首先训练 Sketch2Render-GAN，即同时训练 SketchGAN 和 RenderGAN。SketchGAN 使用手电钻草图数据集进行训练，而 RenderGAN 的训练数据由两部分构成，即草图部分与产品图部分，它们是从总数据中随机抽取。然后，用训练好的 RenderGAN 为已训练的 SketchGAN 生成的草图方案进行着色渲染。最后，在 MS-COO[44]数据上训练一个草图神经风格迁移网络，可依据风格偏好对设计草图进行风格迁移。此外，还在自行车头盔数据集上进行了扩展验证实验。

3.2　草图生成实验

3.2.1　SketchGAN 网络结构

SketchGAN 的网络结构基于 DCGAN[7]而设计。在卷积核尺度设计方面，考虑到线稿草图中像素点的稀疏特性，采用较大尺寸的卷积核。图 7 为生成器的网络结构，包括 6 个用于特征映射的转置卷积层。用高斯误差线性单元(GELUs)[45]代替修正线性单元(ReLU)作为激活函数层。卷积核尺寸均设计成 8×8，网络输入为 128 维的随机噪声，输出为一批 256×256 像素的手电钻概念草图。图 8 为判别器的网络结构，包括 6 个用于草图特征识别的卷积层。第 1 个卷积层和最后一个卷积层的卷积核尺寸设计为 7×7，其余卷积层均为 6×6。用 Instance Norm 层代替了 Batch Norm 层，网络输入为一批 256×256 像素的真草图数据，输出为判断草图是真或假的概率值。

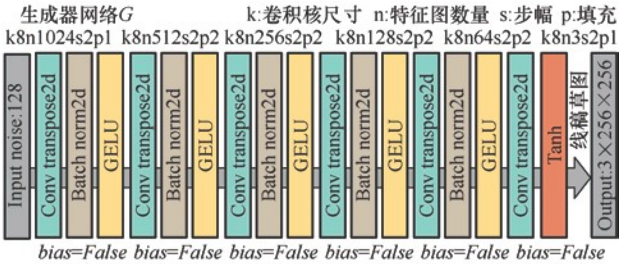


图 7　SketchGAN 的生成器网络结构及参数设置

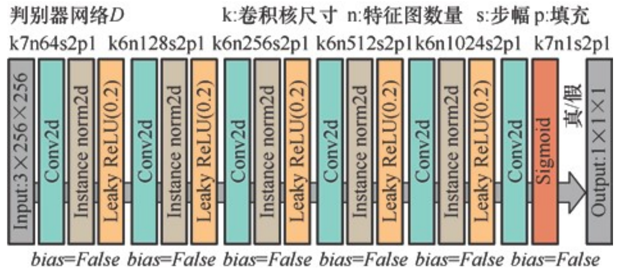


图 8　SketchGAN 的判别器网络结构及参数设置

3.2.2　网络训练与结果

SketchGAN 的最大训练轮数(max epoch)设置为 500，输入生成器的随机噪声为 128 维，批尺寸 batch size 设为 64。使用 Adam[46]优化器训练生成器和判别器，初始学习率均为 0.000 2，$\beta 1 = 0.5$，$\beta 2 = 0.999$，无衰减策略。对抗损失平衡因子 $\lambda_{adv} = 1$，判别器的梯度惩罚平衡因子 $\lambda_{gp} = 10$。经实验测试 λ_{gp} 的取值在[6,10]区间均可成功训练。先训练判别器，后训练生成器，每 1 个 batch 训练 1 次判别器，每 7 个 batch 训练 1 次生成器。如此交替循环，训练过程的损失变化，如图 9 和图 10 所示。

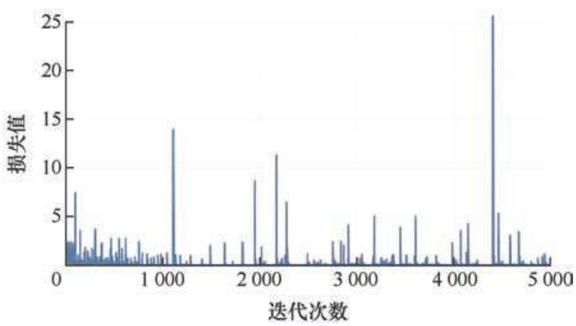


图 9　判别器的损失变化

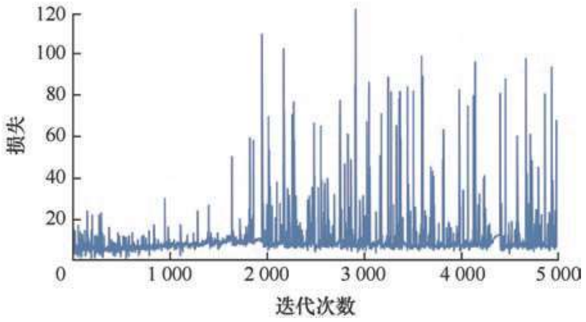


图 10　生成器的损失变化

从图 9 中可知判别器的损失呈现间断跳跃式变化，表明其训练过程没有出现梯度消失，从而为训练有效的生成器提供保障。

从图 10 中生成器的损失函数变化中发现，约 2000 次迭代(200 个 epoch)之后的模型变得不稳定，误差增大，出现了模式崩塌现象，模型设计输出的草图缺乏多样性，可停止训练。但在第 179 个 epoch 左右训练的模型误差基本在 10%附近，模型稳定性良好。

图 11 为训练好的 SketchGAN(第 179 个 epoch 的生成器和判别器)在输入随机噪声后一次性由生成器 G 输出 512 张候选方案，并经过判别器 D 打分后优选出 64 张草图方案。该计算创意操作耗时约为 4 s，视觉结果显示草图的多样性基本得到了保障。

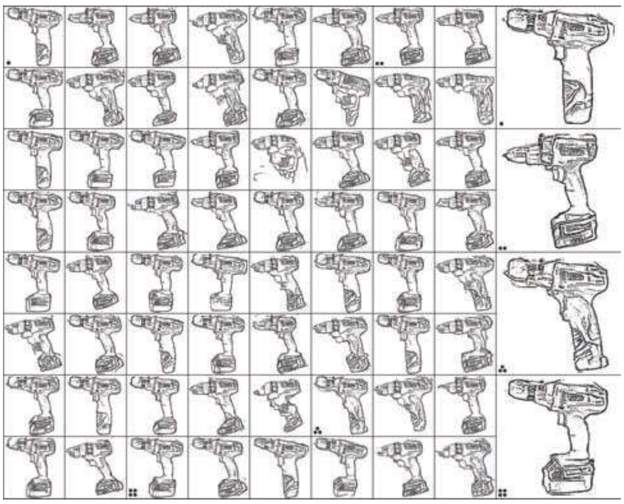


图 11　SketchGAN 设计输出的手电钻概念草图

3.3　草图聚类实验

3.3.1　草图分类器网络结构

草图聚类模型通过草图识别分类器实现，使用 t-NSE[47]降维可视化分析。其网络结构是基于 LeNet-5[48]改进，每个卷积层之后增加批归一化层，并使用最大池化层代替平均池化层；前两个线性层也做批归一化处理，并使用丢弃法舍弃 50%的神经单元防止过拟合；所有激活函数使用 GELUs[45]替换 Sigmoid，如图 12 所示，输出为草图类别。

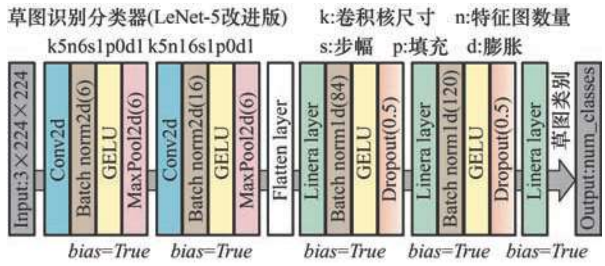


图 12　草图识别分类器网络结构及参数设置

3.3.2　训练与结果

网络最大训练轮数设为 50 个 epoch，批尺寸 batch size 为 16。使用交叉熵损失函数进行性能度量。采用 AdamW[49]优化器训练，初始学习率为 0.001，$\beta 1=0.9$，$\beta 2=0.999$，权重衰减为 0.001。训练集和验证集的比例为 8:2，训练准确率为 99.1%，验证准确率为 97.2%。选择 SketchGAN 生成的 64 张草图和设计师绘制的 68 张草图进行测试，并通过 t-SNE[47]聚类可视化，如图 13 所示。发现相似的草图被聚集到一起，实现草图认知组块模拟。既可以帮助设计师快速筛选方案，又能对草图进行智能标注，以便形成闭环设计系统。

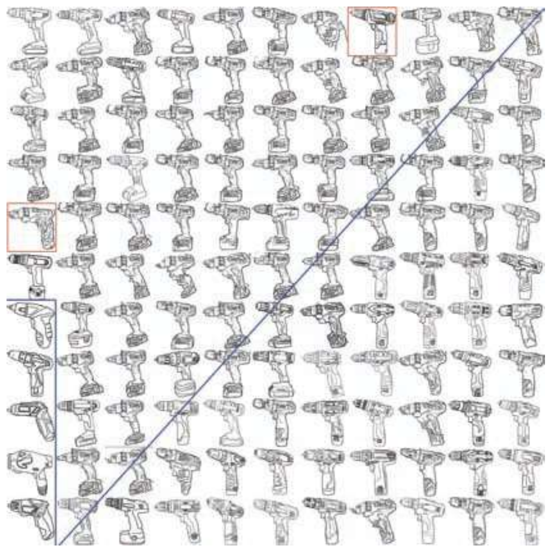


图 13　手电钻草图聚类结果

3.4　草图渲染实验

3.4.1　RenderGAN 网络结构

为更好地适应产品设计草图渲染任务，在 Cycle-GAN[36]的基础上构建 RenderGAN 网络结构。图 14 为判别器的网络结构，沿用 DCGAN 的结构及参数，在第 2、第 3、第 4 卷积层后使用 Instance Norm 层，并使用 GELUs[45]代替 LeakyReLU 作为激活函数层，在输出层使用了平均池化层和数据展平层，输入尺寸为 256×256 像素，输出为单个判断值。

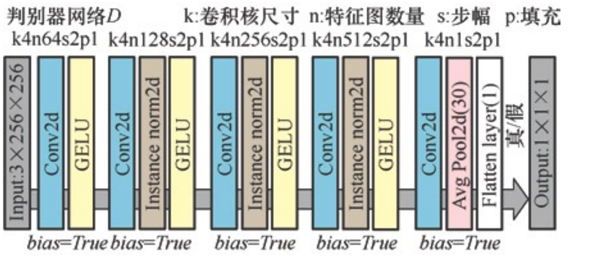


图 14 RenderGAN 的判别网络结构及参数设置

图 15 为生成器的网络结构，由编码器、转换器和解码器构成。编码器是三个卷积层，转换器由是 16 个残差块和一个跳转连接的残差块组成，解码器由两个转置卷积层和一个卷积层组成。同时，使用 GELUs[45]层替换 ReLU 层作为生成器的激活函数层。输入和输出尺寸是相同的，均为 256×256 像素。

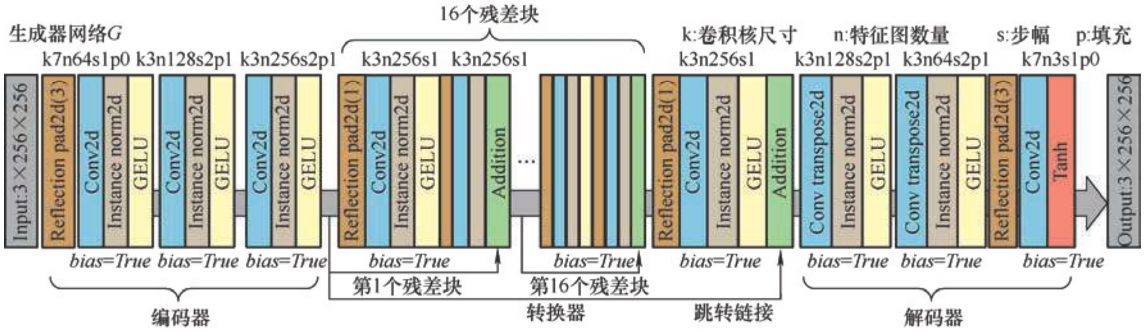


图 15 RenderGAN 的生成网络结构及参数设置

3.4.2 训练与结果

模型最大训练轮数设为 200 个 epoch，批尺寸 batch size 为 1。两个 GAN 模块均使用 Adam[46]优化器进行训练，学习率为 0.000 2，$\beta 1=0.5$，$\beta 2=0.999$。循环平衡因子$\alpha=10$，恒等平衡因子$\beta=5$。判别器的平衡因子$\gamma=\delta=0.5$。所用数据集是从手电钻草图数据集中随机抽取，无需配对。训练集由 120 张草图和 120 张产品图组成，测试集由 30 张草图(其中 15 张是 SketchGAN 设计的草图)和 30 张产品图组成，像素尺寸均为 256×256 像素。

训练过程的损失变化如图 16 所示，其中 $Loss\_G$ 是生成器的总损失，$Loss\_D$ 是判别器的总损失，二者均逐步下降趋于稳定，表明两个模型性能良好。

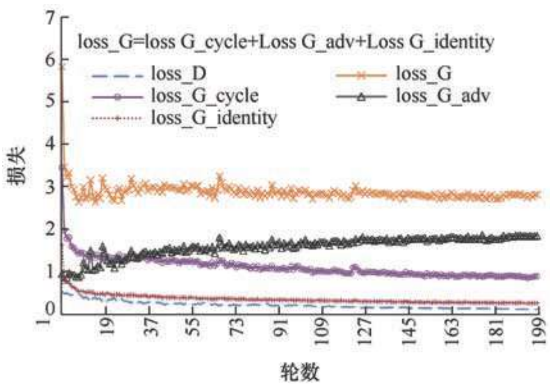


图 16 RenderGAN 的损失变化

图 17 是由训练好的生成网络 G_S 对草图渲染输出的结果，其中第 2 行是对测试集草图渲染的结果，第 4 行是对 SketchGAN 生成的草图渲染的结果。与渲染测试集草图(真草图)相比，从视觉上不难发现 RenderGAN 在 SketchGAN 生成的草图(假草图)上的表现略显逊色，但仍有益于设计之初辅助设计师快速获得可视化概念方案。另外，这也表明 SketchGAN 具有较强的草图创新设计能力，而使得 RenderGAN 在其生成的草图上表现出了低泛化能力。此外，草图中的噪声点也影响了渲染效果。虽然渲染结果在视觉细节上与人类设计师还存在明显差距，但这表明了 AI 模型在产品草图渲染方面的潜在应用。

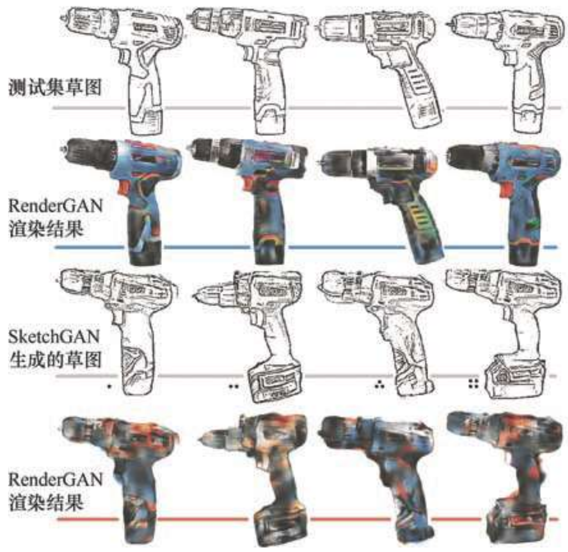


图 17 RenderGAN 渲染的手电钻效果图对比

Sketch2RenderGAN 设计的概念草图也激发工业设计师的灵感和创作欲望，在“看-动-看”(seeing-moving-seeing)[50]的循环互动中进行人机草图设计协作，结果如图 18 所示。从视觉感知层发现，

图 18　两次人机草图设计协作结果

SketchGAN 生成的草图引导设计师突破思维固化，创造新造型，RenderGAN 渲染的效果图为其提供了更丰富的配色方案。

3.5 草图风格迁移实验

3.5.1　网络结构

风格迁移网络结构由编码器、AdaIN 和解码器三部分组成，激活函数层使用 GELUs[45]，如图 19 所示。编码器使用预训练 VGG-19 的前 9 个 GELU 层，AdaIN 层负责风格迁移，解码器通常是编码器的镜像。

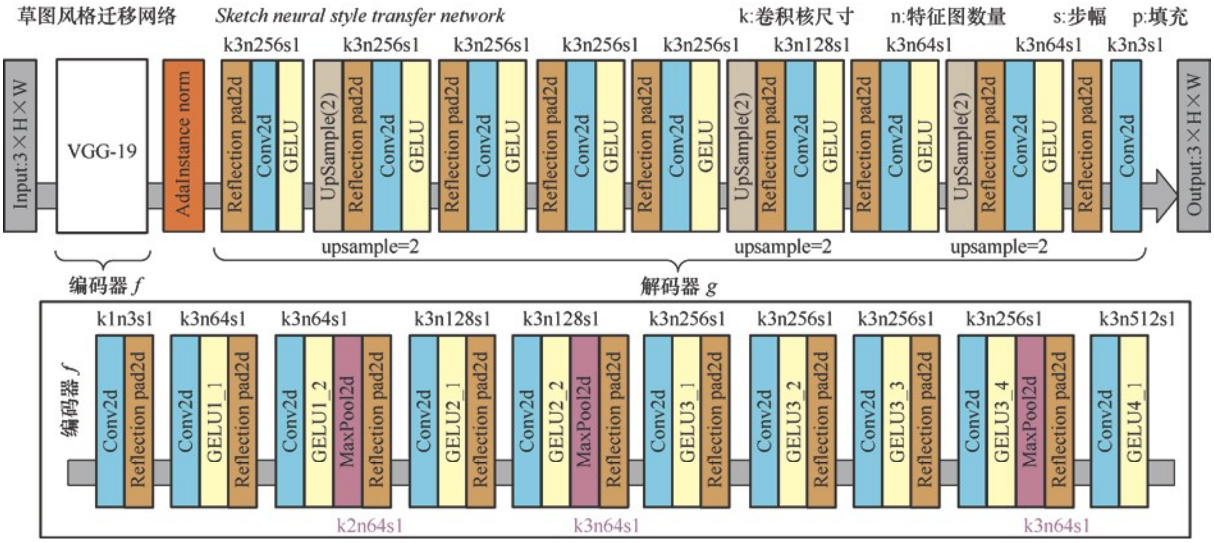


图 19　Sketch-NST(AdaInNorm)网络结构及参数设置

3.5.2　训练与结果

模型最大训练轮数设为 50 个 epoch，批尺寸 batch size 为 8。使用 Adam[46]优化器训练，初始学习率为 0.000 1，$\beta 1 = 0.9$。内容权重 $\alpha = 1$，风格权重 $\beta = 20$。内容数据集使用 MS-COO[44]，风格数据集使用 Wikiart。与 ImageNet 相比，MS-COO 数据集的种类虽少，但每个类别的数量多且丰富，更适合草图风格迁移学习。

选择与手电钻草图形态大致相似的两款风格不同的吹风机和两张风格迥异的纹理图像，对 RenderGAN 渲染的手电钻效果图进行风格迁移实验，结果如图 20 所示。

由于所训练的模型具有较好的自适应能力，具体形态上的差异并没有导致原始手电钻草图细节丢失，同时还引入了新的纹理特征。这表明风格迁移的思想和技术将在产品概念个性化创新方面会产生积极的推动作用。这为草图方案自动引入了新的造型特征、纹理和色彩方案，从而有效弥补了 Sketch2Render-GAN 无法像人类设计师跨域映射风格造型特征的想象能力。相比形态内容差异较大的吹风机，如果选择形态内容更为接近的真实手电钻作为风格图像，结果发现 AI 模型引入了更多的设计细节，如图 21 所示。

图 20　手电钻草图风格迁移(1)

图 21　手电钻草图风格迁移(2)

3.6　人机交互界面

为便于设计师与 AI 模型交互，使用 Dear PyGui 开发了一个流程完整的智能草图设计生成器(S-SDG_v0.1)，如图 22 所示。该交互界面包括四个主要模块，从左至右依次为用户指南(User guide)、草图生成模块(SketchGAN)、草图渲染模块 (Render-GAN) 和草图风格迁移模块(Sketch-NST)。用户指南提供基本使用说明和文档，以图和简要的文字帮助设计师理解该系统，其他三个模块都提供了加载模型、生成和保存操作的功能性接口。

图 22　智能草图设计生成器交互界面(S-SDG_v0.1)

3.7　自行车头盔案例

为进一步证明所提方法框架的有效性，避免单一数据集的局限性，文章还验证了自行车头盔概念草图的生成、渲染及风格迁移。

依据所述方法搜集数据及预处理，共获头盔草图 4 864 张。图 23 为 SketchGAN 生成的自行车头盔概念草图。

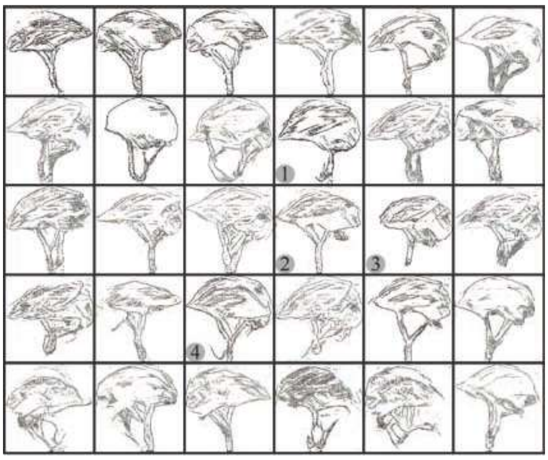


图 23　SketchGAN 设计输出的自行车头盔概念草图

对于头盔草图的渲染，尝试训练能够渲染不同颜色的 RenderGAN。为此，创建了一个具有不同颜色的头盔数据集。这里以渲染的不同效果头盔为例进行展示，如图 24 第 2，3 行所示，第 4 行为风格迁移后的效果。

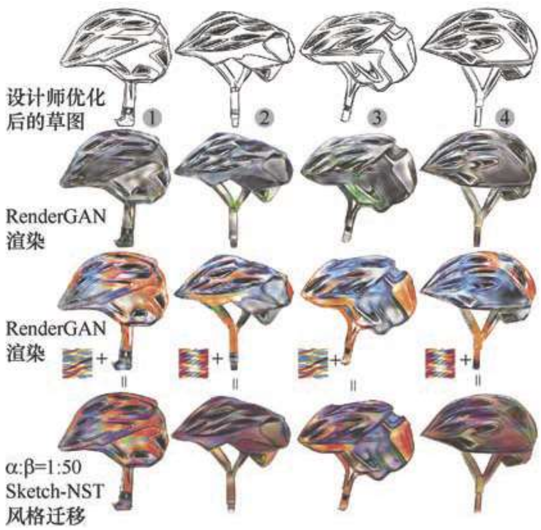


图 24　头盔渲染效果图和风格迁移结果

4　讨论

人工智能与创新设计正在从科学与艺术的角度

不断地融合，已经实现了学习和创新的跨越，形成了类似人类智能的创作闭环[51]。文章初步尝试了用人工智能方法模拟设计师的草图思维过程，提出了一个相对完整的产品草图智能设计集成方法框架，包括草图生成、渲染及风格迁移，探索了人工智能辅助工业产品概念草图设计的实现模式。

在概念生成方面，SketchGAN 的本质是用高维随机噪声数据逼近，同时模拟出与大量真实草图像素数据近似的分布规律，正是这个原因 SketchGAN 在创造新概念的同时创造了新的视觉逻辑模式。此外，与直接应用 GAN 生成产品图像[12-13]相比，SketchGAN 能够创造更多的细节特征，如图 25 所示。然而，仍有三个无法回避的问题。首先，人脸[7]、动漫头像[17]等数据集的整体结构相对统一，即都是两个眼睛一个鼻子一张嘴，但手电钻和自行车头盔草图数据集的稀疏性相对较高，并且自行车头盔的数据分布比手电钻要更复杂一些。因此，对于输入不同随机噪声实现生成功能和结构可控的创意草图仍是一个瓶颈。其次，如何高效、智能化的创建干净、高质量、数量可观、种类多样的产品草图数据集仍是一个问题。最后，SketchGAN 模型的创新设计能力仍有提升空间，例如附加设计约束空间，开拓基于设计师文本描述智能生成相应的概念草图[52-53]，从而让 SketchGAN 成为设计师草图思维的直接拓展者。

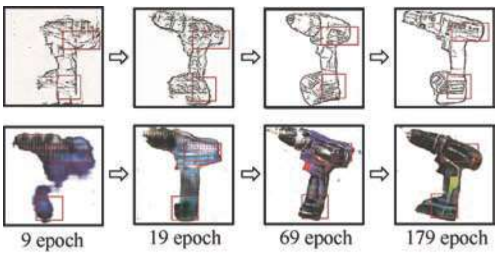


图 25　手电钻草图生成与图像生成对比

草图渲染模型 RenderGAN 是基于 Cycle-GAN 的循环一致性实现，这与人类设计师草图行为的循环互动特性相似。然而，目前的 RenderGAN 只能针对某一种产品实现草图渲染，例如针对手电钻、头盔、吹风机等单一类型数据集而建立有效的 AI 模型是可行的，若想泛化到多个品类仍需深入研究。同时，针对不同的草图数据集，RenderGAN 的表现也是不同的。对比图 18 和图 24 发现手电钻的渲染效果更好，这是由于自行车头盔的形态结构相对复杂，数据分布更稀疏。相比智能算法，设计师只要掌握基本的光影逻辑和色彩关系就可以渲染各种不同品类的产品设计草图。从这个角度来讲，人类设计师的高阶视觉认知活动还是远超 AI 模型的。因此，将设计师的经验和专业知识引入到生成模型中以提高模型的认知推理能力仍是一项长期的工作。

风格迁移是深度卷积神经网络对两张图像特征的解构与重建，但对产品设计草图则需要更多考虑形态细节上的得与失。对比图 20 和图 21 可知，风格产品的选择对草图风格迁移效果存在不同程度的影响。另外，内容权重与风格权重的配比也是影响设计细节的关键，不同的产品需提供合适的参数配比。未来希望探索动态草图风格迁移，进一步拓展模型的泛化能力。

从设计师和 AI 模型的两次草图设计协作中不难发现，设计师和人工智能之间的人机设计协作模式有待进一步探索。AI 模型设计擅长处理大量错综复杂的、形式多样的数据，设计师则擅长处理设计细节，二者该如何高效、流畅的沟通是人机设计协作的重点。尽管该项研究为工业设计师研发了简洁的交互辅助设计界面，然而现阶段深度学习模型的可解释性和设计师对 AI 技术的理解深度都对人机设计协作模式提出挑战。因此，探索人机设计协作新模式仍是现阶段及未来设计智能的核心问题[5]。

另外，从学习所需要的数据量来看，人工智能设计所需要的数据量远大于人类设计师。人类设计师绘制草图体现的是一种创造性的视觉思维。然而现阶段的人工智能、深度神经网络尽管能模拟草图设计过程，所创造的结果在视觉认知层面对设计师具有积极的影响，但它更多的是一种数据思维。设计师通过绘制一系列的概念草图，直到探索出新的概念，这是人类的高阶视觉认知活动的表现；人工智能、GAN 能在极短的时间内设计输出数量远超人类设计师，但还不完全具有高阶认知活动的能力。即便是它创造了比工业设计师更多、更好的设计方案草图，现阶段其实并没有真正“理解”和“解释”方案的能力，这也正是人们继续探索的动力。

5　结论

设计师用草图进行概念探索的过程，需要绘制一系列的草图方案，从中进行观察、比较、思考和过滤，这是他们与草图方案之间的互动协作过程。Sketch2RenderGAN 和 Sketch-NST 的联合使得 AI 模型模拟了设计师“看-动-看”[50]的互动协作过程。

设计的 SketchGAN 模型能够适应不同的草图数据集，可在一定程度上避免模式崩塌，设计输出的概念草图具有多样性和细节特征。RenderGAN 渲

染器在特征提取方面性能良好，小样本训练渲染效果明显，可有效帮助设计师过滤概念。分别以手电钻和自行车头盔为设计对象进行验证性实验。结果表明所构建的方法框架可有效产出大量创新概念草图，激发设计师的创新潜力，提高设计效率。

所提出的 AI 草图视觉认知模型，是将草图认知研究成果与深度学习方法融合，有助于促进 AI 辅助草图设计形成闭环，并为进一步深入研究 AI 辅助工业产品草图设计提供参考。

所开发的 S-SDG_v0.1 为工业设计师创造了一个人机协创交互窗口，有效降低他们应用智能算法辅助设计的门槛，从而提升概念构思的高效性、互动性和便捷性。

人工智能、机器学习、深度学习引入工业设计中，为计算创意方法注入了新动力，特别是设计之初的概念草图设计阶段能引发一系列的重要改进。有助于推动草图思维的革新，优化概念草图设计过程的互动性，提升概念构思过程的效率，直至改变设计师的思维方式和工作模式。

参 考 文 献

[1] CHEN L，WANG P，DONG H，et al. An artificial intelligence based data-driven approach for design ideation[J]. Journal of Visual Communication and Image Representation，2019，61：10-22.

[2] MAO X，GALIL O，PARRISH Q，et al. Evidence of cognitive chunking in freehand sketching during design ideation[J]. Design Studies，2019，67：1-26.

[3] HA D，ECK D. A neural representation of sketch drawings[C]// 6th International Conference on Learning Representations (ICLR)，Vancouver，BC，Canada，April 30-May 3，2018.

[4] CAO N，YAN X，SHI Y，et al. AI-Sketcher：A deep generative model for producing high-quality sketches[J]. Proceedings of the AAAI Conference on Artificial Intelligence，2019，33(1)：2564-2571.

[5] TANG Y C，HUANG J J，YAO M T，et al. A review of design intelligence：progress，problems，and challenges[J]. Frontiers of Information Technology & Electronic Engineering，2019，20(12):1595-1617.

[6] GOODFELLOW I，POUGET-ABADIE J，MIRZA M，et al. Generative adversarial nets[C]// Proceedings of the 27th International Conference on Neural Information Processing Systems. Cambridge：MIT Press，2014：2672-2680.

[7] RADFORD A，METZ L，CHINTAL S. Unsupervised representation learning with deep convolutional generative adversarial networks[C]// 4th International Conference on Learning Representations (ICLR)，San Juan，Puerto Rico，May 2-4，2016.

[8] NGUYEN-PHUOC T，LI C，THEIS L，et al. HoloGAN：unsupervised learning of 3D representations from natural images[C]// 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCV)，Seoul，Korea，Oct. 27-Nov. 2，2019：7588-7597.

[9] KWON Y H，PARK M G. Predicting future frames using retrospective cycle GAN[C]// IEEE Computer Vision and Pattern Recognition (CVPR)，Long Beach，CA，June 16-20，2019：1811-1820.

[10] XU T，ZHANG P，HUANG Q，et al. AttnGA：Fine-grained text to image generation with attentional generative adversarial networks[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)，Salt Lake City，Utah，June 18-23，2018：1316-1324.

[11] CRESWELL A，WHITE T，DUMOULIN V，et al. Generative adversarial networks：an overview[J]. IEEE Signal Processing Magazine，2017，35(1)：53-65. https://arxiv.org/pdf/1710.07035.pdf.

[12] LIU Z，GAO F，WANG Y. A Generative adversarial network for AI-aided chair design[C]// 2019 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR)，San Jose，California，March 28-30，2019：486-490.

[13] DAI Y，LI Y，LIU L J. New product design with automatic scheme generation[J]. Sensing and Imaging an International Journal，2019，20(1)：1-16.

[14] MAO X，LI Q，XIE H，et al. Least squares generative adversarial networks[C]// 2017 IEEE International Conference on Computer Vision (ICCV)，Venice，Italy，October 22-29，2017：2813-2821.

[15] ARJOVSKY M，CHINTALA S，BOTTOU L. Wasserstein generative adversarial networks[C] //Proceedings of the 34th International Conference on Machine Learning (ICML). Sydney，Australia，August 6-11，2017：214–223.

[16] SAGE A，AGUSTSSON E，TIMOFTE R，et al. Logo synthesis and manipulation with clustered generative adversarial networks[C]// 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)，Salt Lake City，UT，USA，June 18-23，2018：5879-5888.

[17] JIN Y H，ZHANG J K，LI M J，et al. Towards the automatic anime characters creation with generative adversarial networks[J]. arXiv preprint arXiv：1708.05509，2017. https://arxiv.org/pdf/1708.05509.pdf.

[18] MIRZA M，OSINDERO S. Conditional generative adversarial nets[J]. arXiv preprint arXiv：1411.1784，2014. https://arxiv.org/pdf/1411.1784.pdf

[19] KIM T，CHA M，KIM H，et al. Learning to discover cross-domain relations with generative adversarial networks[C]// Proceedings of the 34th International Conference on Machine Learning (ICML)，Sydney NSW，Australia，August 6-11，2017：1857-1865.

[20] CHAI C，LIAO J，ZOU N，et al. A one-to-many conditional generative adversarial network framework for multiple image-to-image translations[J]. Multimedia Tools and Applications，2018，77(17)：22339-22366.

[21] PAN Y，BURNAP A，HARTLEY J，et al. Deep design：product aesthetics for heterogeneous markets[C]// Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)，Halifax，August 14-17，2017：1961-1970.

[22] 刘跃中，斯托夫斯·卢迪，杨阳. 基于条件生成对抗网络的城市设计研究[J]. 建筑学报，2018(9)：108-113.
LIU Yuezhong，STOUFFS R，YANG Yang. Urban design process with conditional generative adversarial networks[J] Architectural Journal，2018(9)：108-113.

[23] PORTILLA J，SIMONCELLI E P. A parametric texture model based on joint statistics of complex wavelet coefficients[J]. International Journal of Computer Vision，2000，40(1)：49-70.

[24] HAYS J，ESSA I. Image and video based painterly animation[C]// Proceedings of the 3rd International Symposium on Non-photorealistic Animation and Rendering (NPAR)，Annecy France，June 7-9，2004，113-120.

[25] SHIH Y C，PARIS S，BARNES C，et al. Style transfer for headshot portraits[J]. ACM Transactions on Graphics，2014，33(4):1-14

[26] GATYS L A，ECKER A S，BETHGE M，et al. A neural algorithm of artistic style[J]. Journal of Vision，2016，16(12)：1-16.

[27] GATYS L A，ECKER A S，BETHGE M，et al. Texture synthesis using convolutional neural networks[C]// Proceedings of the 28th International Conference on Neural Information Processing Systems，Montreal，Canada，December 7-12，2015：262–270.

[28] JOHNSON J，ALAHI A，FEI-FEI L. Perceptual losses for real-time style transfer and super-resolution[C]// 14th European Conference on Computer Vision (ECCV). Amsterdam，Netherlands，October 11-14，2016：694-711.

[29] FORBUS，KENNETH D. Vision and sketching[J]. Perception，2012，41(9)：1031-1039.

[30] GULRAJANI I. AHMED F，ARJOVSKY M，et al. Improved training of Wasserstein GANs[C]// Proceedings of the 31st International Conference on Neural Information Processing Systems，California，USA，December 4-9，2017：5769–5779.

[31] KODALI N，ABERNETHY J，HAYS J，et al. On convergence and stability of GANs[C]// 6th International Conference on Learning Representations (ICLR)，Vancouver，BC，Canada，April 30-May 3，2018.

[32] EREZ P. Sketching in design and caid：a theoretical exploration[D]. Calgary：The University of Calgary，1999：50-66.

[33] 刘征，鲁娜，吴剑锋. 基于设计认知的草图研究综述[J]. 浙江大学学报(工学版)，2010，44(12)：2376-2382.
LIU Zheng，LU Na，WU Jianfeng. Review of sketch based on design cognition[J]. Journal of Zhejiang University (Engineering Science)，2010，44(12)：2376-2382.

[34] ISOLA P，ZHU J，ZHOU T，et al. Image-to-image translation with conditional adversarial networks[C]// Proceedings of the IEEE Computer Vision and Pattern Recognition (CVPR)，Honolulu，Hi，USA，July 21-26，2017：5967-5976.

[35] WANG T C，LIU M Y，ZHU J Y，et al. High-resolution image synthesis and semantic manipulation with conditional GANs[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)，Salt Lake City，Utah，June 18-23，2018：8798-8807.

[36] ZHU J Y，PARK T，ISOLA P，et al. Unpaired image-to-image translation using cycle-consistent adversarial networks[C]// IEEE International Conference on Computer Vision (ICCV)，Venice，Italy，October 22-29，2017：2242-2251.

[37] JIN Y，CHUSILP P. Study of mental iteration in different design situations[J]. Design Studies，2006，27(1)：25-55.

[38] GATYS L A，ECKER A. S，BETHGE M. Image style transfer using convolutional neural networks[C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas，June 26-July 1，2016：2414-2423.

[39] HADJI I，WILDES R P. What do we understand about convolutional networks?[J]. arXiv preprint arXiv：1803.08834，2018. https://arxiv.org/pdf/1803.08834.pdf.
[40] IOFFE S，SZEGEDY C. Batch normalization：accelerating deep network training by reducing internal covariate shift[C]// Proceedings of the 32nd International Conference on International Conference on Machine Learning (ICML). Lille France，June 6-11，2015：448-456.
[41] ULYANOV D，VEDALDI A，LEMPITSKY V. Instance normalization：The missing ingredient for fast stylization[J]. arXiv preprint arXiv：1607.08022v3，2018. https：//arxiv.org/pdf/1607.08022v3.pdf.
[42] DUMOULIN V，SHLENS J，KUDLUR M，et al. A learned representation for artistic style[C]// 5th International Conference on Learning Representations (ICLR)，Toulon，France，April 24-26，2017.
[43] HUANG X，BELONGIE S. Arbitrary style transfer in real-time with adaptive instance normalization[C]// IEEE International Conference on Computer Vision (ICCV)，Venice，Italy，October 22-29，2017：1510-1519.
[44] LIN T，MAIRE M，BELONGIE S，et al. Microsoft COCO：common objects in context[C]// European Conference on Computer Vision (ECCV)，Zürich，Switzerland，September 6-12，2014：740-755.
[45] HENDRYCKS D，GIMPEL K. Gaussian error linear units (GELUs)[J]. arXiv preprint arXiv：1606. 08415，2016. https://arxiv.org/pdf/1606. 08415v4.pdf.
[46] KINGMA D，BA J. Adam：a method for stochastic optimization[C]// 3rd International Conference on Learning Representations (ICLR)，San Diego，CA，USA，May 7-9，2015.
[47] MAATEN L V D，HINTON G. Visualizing Data using t-SNE[J]. Journal of Machine Learning Research，2008，9(2605)：2579-2605.
[48] LECUN Y，BOTTOU L，BENGIO Y，et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE，1998，86(11)：2278-2324.
[49] LOSHCHILOV I，HUTTER F. Fixing Weight Decay Regularization in Adam[C]// 6th International Conference on Learning Representations (ICLR)，Vancouver，BC，Canada，April 30-May 3，2018.
[50] SCHON D A，WIGGINS G. Kinds of seeing and their functions in designing[J]. Design Studies，1992，13(2)：135-156.
[51] 胡洁. 人工智能驱动的艺术创新[J]. 装饰，2019(11)：12-17.
HU Jie. Artificial intelligence driven art innovation[J]. Zhuang Shi，2019(11)：12-17.
[52] HUANG F，CANNY J F. Sketchforme：composing sketched scenes from text descriptions for interactive applications[C]// Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology (UIST)，New Orleans LA USA，October 20-23，2019：209-220.
[53] FORREST H，ELDON S，DAVID H，et al. Scones：towards conversational authoring of sketches[C]// Proceedings of the 25th International Conference of Intelligent User Interfaces (IUI)，Cagliari Italy，March 17-20，2020：313-323.

作者简介：李雄，男，1984 年出生，博士研究生。主要研究方向为智能设计、参数化设计、草图思维。
E-mail：leaondesign@163.com
苏建宁(通信作者)，男，1974 年出生，博士，教授，博士研究生导师。主要研究方向为感性工学、智能设计等。
E-mail：sujn@lut.edu.cn
张志鹏，男，1994 年出生，博士研究生。主要研究方向为智能设计、参数化设计。
E-mail：Japadesign@163.com

《计算机辅助设计与图形学学报》录用通知

李雄 苏建宁 张志鹏 李晓晓　　同志：您们好！

非常高兴地通知您，您的修改稿已于 2022-11-02 收到。

论文题目：　特征迁移的细粒度产品形态智能决策方法

论文编号：19830

经专家评审，编委会决定予以录用。具体发表日期另行通知。感谢您对本刊的支持，期望您继续为本刊投来有关您最新科研成果的高水平的学术论文。

此致

敬礼！

《计算机辅助设计与图形学学报》编辑部

2022-11-04

计算机辅助设计与图形学学报 编辑部

特征迁移的细粒度产品形态智能决策方法

李雄 [1,3], 苏建宁 [1,2]*, 张志鹏 [1], 李晓晓 [4]

[1] (兰州理工大学机电工程学院 兰州 730050)
[2] (兰州理工大学设计艺术学院 兰州 730050)
[3] (兰州城市学院培黎机械工程学院 兰州 730070)
[4] (新疆艺术学院设计学院 乌鲁木齐 830049)
(sujn@lut.edu.cn)

摘 要: 针对产品形态智能决策框架系统性不强、模型决策机制单一且历史样本数据量少等问题, 提出一种基于混合迁移学习的细粒度产品形态智能决策方法. 该方法将 Swin Transformer 和 ResNets 作为骨干网络设计了 3 个并行混合迁移学习子网络, 即产品形态识别网络(Form-CN)、产品形态深度回归评价网络(Form-REN)和产品形态分布拟合评估网络(Form-DFEN). 首先应用 Form-CN 对产品进行细粒度形态分类判别, 实现产品形态设计定位识别任务; 其次应用 Form-REN 对产品整体形态语义进行预测评价; 然后通过 Form-DFEN 对产品形态进行分布拟合评估; 最后由 Form-REN 和 Form-DFEN 完成综合决策. 以创建的手电钻数据集进行实验, 并与其他经典模型进行比较, 结果表明所设计的 3 个网络分别取得了 99.0%的准确率、0.4058 的均方误差和 84.3%的准确率; 所提方法能够精细、高效地辅助设计师进行综合智能决策, 为产品形态智能决策提供了一个更为系统的参考框架.

关键词**:** 产品形态智能决策; 细粒度识别; 迁移学习; 并行网络
中图法分类号**:** TP391.41 **DOI:** 10.3724/SP.J.1089.2023.19830

Intelligent Decision-Making of Fine-grained Product Form with Feature Transfer

Li Xiong [1,3], Su Jianning [1,2]*, Zhang Zhipeng [1], Li Xiaoxiao [4]

[1] (*School of Mechanical & Electrical Engineering, Lanzhou University of Technology, Lanzhou* 730050)
[2] (*School of Design Art, Lanzhou University of Technology, Lanzhou* 730050)
[3] (*School of Bailie Mechanical Engineering, Lanzhou City University, Lanzhou* 730070)
[4] (*School of Design, Xinjiang Arts University, Urumqi* 830049)

Abstract: A fine-grained product form intelligent decision-making method based on hybrid transfer learning is proposed. The aim is to solve the problems of weak systematicity of the product form intelligent decision-making framework, single decision mechanism of the model, and small amount of historical sample data. The method uses Swin Transformer and ResNets as the backbone network to design three parallel hybrid transfer learning sub-networks, including product form classification network (Form-CN), product form deep regression evaluation network (Form-REN) and product form distribution fitting evaluation network (Form-DFEN). Firstly, Form-CN is applied to classify the products with fine-grained form to achieve the product form design location identification. Secondly, Form-REN is applied to predict and evaluate the overall product form semantics. Then, Form-DFEN is used to evaluate the product form by distribution fitting. Finally, Form-REN and Form-DFEN are used to complete the integrated decision making. Experiments were performed on the created hand drill dataset and compared with other classical models. The results show that the three designed networks achieved 99.0% accuracy, 0.4058 mean square error and 84.3% accuracy, respectively. The proposed method can finely and efficiently assist designers to make comprehensive intelligent decisions, which provides a more systematic reference framework for intelligent decision-making of product forms.

收稿日期: 2022-07-06; 修回日期: 2022-11-02. **基金项目:** 国家自然科学基金(52165033). **李雄**(1984 年—), 男, 博士研究生, 讲师, 主要研究方向为智能设计、参数化设计、草图思维; **苏建宁**(1974 年—), 男, 博士, 教授, 博士生导师, 论文通信作者, 主要研究方向为感性工学、智能设计等; **张志鹏**(1994 年—), 男, 博士研究生, 主要研究方向为智能设计、参数化设计等; 李晓晓(1991 年—), 男, 讲师, 主要研究方向为感性工学、可持续设计等.

Key words: intelligent decision-making of product forms; fine-grained recognition; transfer learning; parallel networks

设计是创新的过程, 评价与决策是其中的重要环节, 而高效决策的关键是决策群体快速达成共识性的评价标准. 传统的工业设计决策过程常由用户代表、营销人员、工程师、设计师、管理者等协作实施. 由于业务偏好、知识背景和经验等不同导致设计评价标准形式多样, 并且真实环境中的设计决策过程呈现模糊性、感性与理性并存, 以及多阶段性、非线性与动态性等特点[1-2]. 产品形态往往是工业设计决策过程中的焦点任务, 既是视觉感知与物理功用的统一, 也是理性与感性的交融, 更是物质层面与精神层面的互通. 设计过程不仅考虑了视觉感知层面的心理因素, 也考虑了结构、人机、环境等物理因素, 复杂的设计信息容易导致决策群体的认知冲突. 此外, 传统人为导向的评价决策模式已无法适应大数据环境下的数据量级和维度要求[2]. 因此, 面对复杂的决策任务,高效准确可信的决策方法和技术一直是研究者与实践者积极探索的问题[3].

近年来, 由于人工智能(artificial intelligence, AI)、机器学习(machine learning, ML)的迅速发展, 深度学习作为 ML 的重要支撑与分支, 其方法及技术的更迭速度超过以往任何历史时期, 并向各个领域渗透, 已在图像分类与生成、目标检测、语义分割等方面取得了重要突破[4], 研究内容和应用场景也不断被扩展, 其中智能决策已成为研究与应用热点之一. 如医疗诊断中 AI 辅助医生进行新冠状病的诊断决策[5]、自动驾驶中的智能决策系统[4]等. 同样地, 在工业设计领域设计研究者们开始应用深度学习方法及技术探索智能设计决策模式, 以加速设计过程中多方参与者的决策流通速度及准确度[6], 从而改变以数据挖掘[7]、经典 ML[3,8]、优化传统决策流程与方法[9]等为主无法实现端到端智能决策的现状.

神经网络是深度学习主要采用的模型, 早期设计评价决策模型研究以多层感知机(multilayer perceptron, MLP)神经网络为主, 在感性设计元素与用户需求之间的非线性决策关系中发挥了重要作用, 探究对象如办公座椅[10]、摩托车头盔和乒乓球拍[11]等. 与 MLP 相比, 卷积神经网络(convolutional neural networks, CNNs)受生物神经科学启发, 其窗口扫视计算操作特性更接近人类的视觉识别方式. 因此, 近几年的探索则聚焦到深度卷积神经网络(deep CNNs, DCNNs). 例如, 朱斌等[12]提出一种基于深度学习的产品意象识别方法,并训练 VGG16 对椅子图像进行智能识别, 对比实验证明其准确率超过以支持向量机(support vector machine, SVM)为代表的经典 ML 方法; GONG 等[13]提出一种像素级的感性分析与识别(pixel-level image Kansei analysis and recognition, PIKAR)系统, 并用改进的 AlexNet 对产品包装进行感性评价和用户体验预测; LI 等[14]采用改进的 ResNet18 对手电钻图像进行情感偏好识别, 实验结果表明不同感性标签的手电钻准确率存在较明显差异, 证明 DCNNs 存有潜在识别偏好特性. SU 等[15]对汽车的细粒度感性分类识别实验结果同样存在该现象. 此类工作多从整体造型感性语义分类角度考虑, 缺乏对产品形态特征的深度回归分析, 因此决策方法单一、决策精度仍有提升空间. 王亚辉等[6]为消除设计决策者偏好对产品开发的影响, 提出了基于 ResNets 的 AI 设计决策模型; 裴卉宁等[16]考虑有限数据资源条件问题, 提出基于胶囊网络的产品形态设计智能决策模型. 这两项工作虽然从产品部件造型语义分割角度考虑, 但所给定的评价标签仍是方案分类标签, 因此使用单一神经网络模型很难在整体形态语义与部件造型特征之间进行均衡性判断. 此外, 考虑到文本描述在设计决策中的作用, 循环神经网络(recurrent neural networks, RNNs), Transformer 常用于文本信息的提取与解读[17-19], 可辅助设计师在设计元素和用户需求匹配方面进行决策[20], 但仍以分类决策为主.

总结上述工作发现产品形态智能决策框架在系统性设计、模型决策机制、小数据训练等方面尚有不足. (1) 模型缺乏系统性. 传统设计评价决策时, 参与者们往往先整体评价后考虑局部造型, 在整体与局部之间给出一个综合决策, 现有智能评价方法未将二者系统性结合, 仅从单方面进行探究; 同时, 现有基于深度学习的评价、决策方法仅以分类任务执行评价过程[6,12-14,16], 然而真实评价过程不仅有解读设计定位的分类识别问题, 还有连续性的回归预测任务. 因此, 现有方法尚未构建出相对系统的智能决策框架. (2) 决策机制单一. 现有设计智能评价、决策方法中所使用的 MLP, CNNs, RNNs 模型尚未考虑注意力机制(attention mechanism, AM)[6,10,12,15-16], 该机制是人工神经网络借鉴认知神经科学中的注意力, 具有选择重点信息的能力. 同时, 现有工作鲜有将不同类型的网络混合使用, 多为单一类型的神经网络, 如 MLP 或 CNNs[6,10,12,16], 特征决策过程存有潜在偏好[14-15]. 因此, 模型提取特征缺乏重点, 决策机制相对单一, 决策精度仍有提升空间. (3) 产品历史数据有限. 迁移学习有利于平衡小数据集与模型泛化能力的矛盾, 可有效缓解产品历史数据量少的问题. 由于单一骨干模型无法全面表征图像细节[21], 本文采用融合不同类型的骨干模型构建混合迁移学习方法, 以便抽取更多有效特征.

目前生成对抗网络(generation adversarial networks, GAN)[22]及其各种变体已被应用到创意概念的生成设计中[23]. 如椅子[24]、智能手表[25]、手电钻[14]等产品的概念图像生成; 城市及建筑立面设计[26]、房屋平面智能布局[27-28]; 海报[29]、Logo[30]的智能生成. 面对不同的设计对象应用 GAN 能生成大量候选方案, 然而面对众多备选方案如何缩小范围, 通常是由 GAN 的判别器进行打分决策, 该方式主要从图像真伪的角度进行评判, 缺乏从设计角度的评价. 因此, 面对此种情况探索具有系统性、精细化的设计评价决策模型, 不仅能够弥补 GAN 判别器的非专业性评价, 而且有助于构建具有计算连续数据流的智能设计系统.

综上所述, 本文从视觉认知的角度出发, 将计算机视觉(computer vision, CV)新成果 Swin Transformer 与经典卷积神经网络 ResNets 共同融入设计视觉评估任务当中, 提出基于混合迁移学习的产品形态智能决策方法框架, 包括细粒度产品形态识别、整体形态语义评价和细粒度产品形态分布拟合评估, 并应用预训练迁移学习方法解决小样本数据训练问题. 相比 MLP, RNNs 和 CNNs, Swin Transformer

具有层级化的视觉 AM, 对人为设计视觉评估具有较高模拟性, 并与 ResNets 联合更具优势. 以手电钻图像数据集[14]为基础构建了手电钻细粒度设计评价数据集, 对所提方法进行了实验验证, 结果表明所构建的细粒度评价体系可有效辅助设计团队进行智能决策, 同时也进一步深化了 CV 的应用研究内容.

1 基础理论

1.1 细粒度识别与智能决策

与粗粒度大类识别任务相比, 细粒度识别是探究某一大类别中的子类识别问题. 通常, 子类别之间具有较高的相似度, 形态差异细小, 从而使得细粒度识别比粗粒度识别更具有挑战性, 例如鸟类图像细粒度识别问题[31]. 与此相似, 在真实产品设计评价决策环节中, 参与者们往往面对众多类似设计方案效果图的细粒度评价决策. 该问题可抽象成一个 ML 任务, 即同类产品间的细粒度识别, 涉及各个方案间整体形态差异识别和单个方案中局部形态间的语义解读.

考虑上述实际问题和场景, 本文提出细粒度产品形态智能决策框架, 并使用具有 AM 的深度神经网络(deep neural networks, DNNs)进行模型训练, 如图 1 所示. 首先, 从细粒度形态差异和功能语义层面构建数据集; 其次, 对方案形态的细粒度形态对比性识别, 实现细粒度产品形态分类; 然后, 考虑整体形态语义评价作为真实设计评价的重要内容, 构建回归评价模型, 从而获得整体形态语义得分; 最后从产品形态设计优劣分布的角度出发, 对其进行更加细粒度的形态分布识别决策, 获得分布拟合评价分.

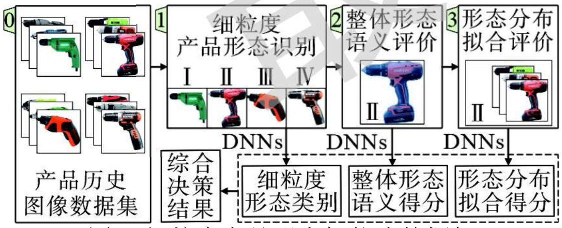


图 1　细粒度产品形态智能决策框架

1.2 迁移学习

从小数据中学习的能力似乎是人类智能的一个特别强大的方面, 这归功于人类智能具备强大的知识和经验迁移能力. 通常, 迁移学习被认为是 ML 的一个重要方面, 特别是面对小数据学习问题时, 可以使 ML 系统更加可靠和稳健. 在真实的设计环境中研究和应用设计智能方法, 多是各种小数据学习任务, 如用户需求分析[11]、概念方案生成[14,25]、设计评价与决策[6,12,14,16]. 与常规方法训练一个完整的神经网络模型相比, 迁移学习试图在小数据集上从已训练好的模型实施相关知识的迁移训练[32], 这些训练好的模型往往是在较大的数据集上完成, 被称为预训练模型. 因此, 迁移学习在一定程度上可以解决学习算法受训练数据量影响泛化能力的问题, 以实现设计决策特征(知识)迁移, 也是研究设计智能[23]可借鉴的思想.

1.3 注意力机制

众所周知, 人脑每时每刻都在通过五感接收来自外界的大量信息. 其中, 眼睛向大脑提供的视觉信息比所有其他器官提供的总和还要多, 一条视神经含有 100 万根神经纤维, 研究表明半数以上的意识信息通过眼睛传入大脑[33]. 然而人脑在有限的资源下, 保持高效、准确的信息处理能力. 这是由于人脑神经系统的两个重要协调机制可以有效处理信息过载问题, 即 AM 和记忆机制(memory mechanism, MM). 前者已经在 CV 的各种任务中取得了重要突破和应用, 如图像识别、语义分割、图像生成等. 在视觉系统中, AM 可以被视为一个动态的选择过程, 它是通过根据输入的重要性对特征进行适应性加权来实现的, 通常 AM 可以抽象化为[34]

$$A_{\text{Attention}} = f\left[g(x), x\right] \tag{1}$$

其中, x 为输入特征; $f(\cdot)$ 表示注意力模型, 通常是某种特别设计的神经网络结构及其计算方式; $g(x)$ 表示对 x 进行处理并产生注意力的过程; $f\left[g(x), x\right]$ 则表示基于注意力 $g(x)$ 对输入特征 x 进行处理的过程.

2 本文方法

2.1 方法框架

本文方法整体框架如图 2 所示, 包括历史数据训练网络和测试数据推理模型 2 个部分. 针对第 1 部分, 首先构建可用的产品图像, 并通过人工标注获取细粒度产品图像数据集; 然后根据数据特征、任务复杂度和决策目标, 基于 Swin Transformer[35]和 ResNets[36]搭建 3 个并行混合迁移学习网络, 构成细粒度智能决策系统, 即产品形态识别网络(product form classification network, Form-CN)、产品形态深度回归评价网络(product form deep regression evaluation network, Form-REN)和产品形态分布拟合评估网络(product form distribution fitting evaluation network, Form-DFEN); 最后训练这 3 个网络, 实现产品形态的细粒度识别、整体形态语义评价和细粒度形态分布解码. 针对第 2 部分,首先通过设计师定义或 GAN 生成获得待评估的方案效果图, 接着从形态类型识别、整体造型语义和形态分布拟合 3 个层面对给定的方案进行细粒度综合评估, 以便帮助设计师和决策团队等进行高效智能决策.

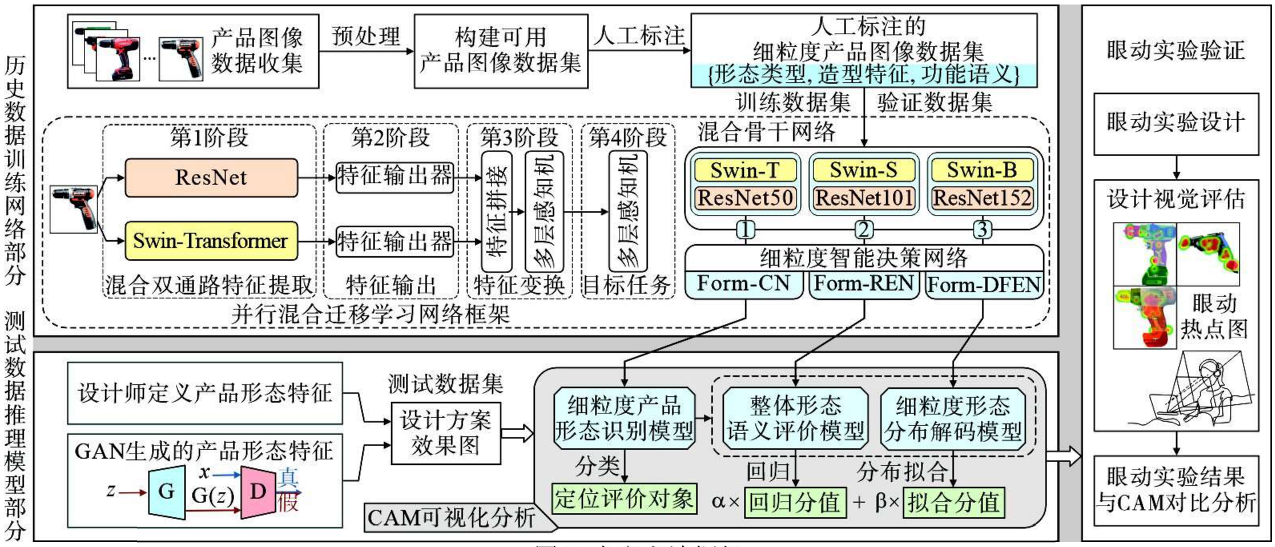


图2 本文方法框架

此外, 为了更进一步对比和探讨智能决策与人工决策在视觉经验上的差异, 引入类别激活映射(class activation maps, CAM)[37]可视化技术对Form-CN, Form-REN和Form-DFEN的决策结果进行可视化分析; 同时, 通过设计眼动实验对神经网络决策进行对比验证, 具体将对比分析眼动热点分布图与CAM热力图, 有助解释神经网络决策,提高决策可信度.

2.2 构建产品图像数据集

如同设计评价团队通过设计知识和视觉经验评估设计方案一样, 深度神经网络同样需要学习高质量的历史数据以获得智能评价模型参数集. 因此, 依据细粒度产品形态智能决策方法框架所提的3种任务, 需构建对应的3类数据集, 即细粒度形态识别数据集、整体形态语义评价数据集和细粒度形态分布拟合数据集. 首先, 借助爬虫技术搜集在线商店和各种搜索引擎上的产品原始图像数据; 然后, 对原始图像进行预处理以获得较高质量的产品图像数据集; 最后, 依据3种具体任务依次制作相关数据集. 具体实施采用文献[14]所提方法流程.

针对形态细粒度分类识别任务, 依据产品造型、结构和功能等细粒度差异, 采用人工分类产品图像并标注. 综合考虑文献[6,15]中的数据集样本量, 本文将每个类别的产品图像数量设置在800~1000个样本.

针对产品整体形态语义评价任务, 需依据相关的设计评价指标$\boldsymbol{x}$组织专业人员对每个样本进行评分, 并依据指标权重$\boldsymbol{\omega}$计算综合基准评分

$$R_{\text{baseline}}(\boldsymbol{x};\boldsymbol{\omega})=\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{x} \tag{2}$$

其中, $R_{\text{baseline}}(\boldsymbol{x};\boldsymbol{\omega})$为人工基准评分函数, $\boldsymbol{x}=[x_1,x_2,\cdots,x_M]^{\mathrm{T}}$为$M$维的评价分数向量, $\boldsymbol{\omega}=[\omega_1,\omega_2,\cdots,\omega_M]^{\mathrm{T}}$为$M$维的权重向量.

依据文献[38]在感知图像质量回归评价任务中所使用的LIVE和CSIQ数据集样本量, 可将产品形态语义评价任务的样本数量控制在1500~2000.

针对分布拟合评价任务, 首先依据形态美学确定产品形态分布级别, 然后采用人工方式对每幅产品图像进行细粒度标注, 参考文献[39]可知每个级别需标注至少1000个样本.

2.3 基于ResNets和Swin Transformer的细粒度智能决策模型

2.3.1 并行骨干模型

在DCNNs中, 不同深度的卷积层会得到不同级别的特征, 较浅的层可抽取低级视觉特征, 较深得层可获得高级语义特征. ResNets的提出使得DCNNs的层数可达1202层[36], 其性能表现得更加优异, 使得模型能学到更加细腻的特征. 基于卷积操作的ResNets是一种静态特征学习模式, 而基于Transformer的Swin Transformer则具有空间动态特征抽取能力[35], 二者可学习到的特征层次各有差异. 综合考虑这两种特征学习模式, 并结合孪生网络(Siamese Network)[40]和GoogLeNet[41]并行网络结构设计思想, 设计了一种具有双通路并行特征提取的混合迁移学习网络框架, 如图2(左上)所示. 下面具体讨论ResNets和Swin Transformer网络结构.

理论上网络越深可获得更高级的语义特征, 识别效果也会得到提升. 然而这种理论上的深层网络并没有取得更好的效果, 实验表明深度网络出现退化现象: 网络深度增加时, 网络准确度出现饱和, 网络难以训练、训练误差增大、准确率下降[36]. 为此, HE等[36]提出了卷积残差恒等映射的思想来解决上述问题, 具体网络结构通过快速跳转补偿连接的方式实现. 依据网络层数的不同, ResNets又分为ResNet18, ResNet34, ResNet50, ResNet101和ResNet152. 已有工作将其用于产品形态智能评估与决策, 如文献[14]将改进的ResNet18用于手电钻的感性偏好评估, 文献[6]把修改的ResNet50应用到起重机部件造型的决策模型.

针对图像分辨率高、像素点多, Transformer[42]的全局自注意力导致模型计算量大; 视觉实体变化大, 不同视觉场景下ViT[43]的性能不稳定两个问题. Swin Transformer[35]提出一种不重叠滑动窗口操作代替固定重叠窗口的卷积操作, 并具有全局感受野. 首先, 通过对图像进行分块操作, 将图片切成一个个小图块, 并进行线性嵌入; 然后利用提出的Swin-transformer模块(Swin-transformer block, STB)进行特征学习, 如图3所示. 1个STB包含1个W-MSA(或SW-MSA)和1个两层的MLP组成.该模块与常规Transformer块(Transformer block, TB)[42]的不同之处在于使用了窗口多头自注意力(window multi-head self attention, W-MSA)和移位窗口多头

自注意力(shifted window multi-head self attention, SW-MSA), 取代了标准的多头自注意力(multi-head self attention, MSA), 使得连续 STB 具有动态空间信息抽取能力, 其性能超过以静态特征抽取的 MLP, RNNs 和 CNNs. 同时, STB 吸收了残差网络恒等映射的思想, 引入前馈跳转补偿后, 当网络结构更深时可以有效防止模型过拟合、梯度消失和爆炸等问题.

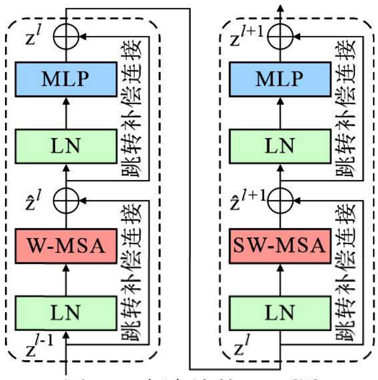


图 3 2 个连续的 STB[35]

Swin Transformer 由特征提取器和分类器构成, 特征提取分为 4 个阶段, 每个阶段执行分片合并操作和移动窗口转换操作, 具体通过 PatchMerging 层和不同数量的 STB 组合实现[35]. 依据第 4 阶段 STB 数量的不同和注意头数的数量可分为 4 种变体网络, 即 Swin-T, Swin-S, Swin-B 和 Swin-L. 网络性能也随模块数量和头数的增加而增强. 本文网络结构基于前 3 个而设计, 依据任务的难度不同采用不同的 Swin Transformer 结构. 具体地说, 细粒度产品形态识别网络基于 Swin-T 和 ResNet50;产品整体形态语义评价网络基于 Swin-S 和 ResNet101; 产品形态分布拟合评估网络基于 Swin-B 和 ResNet152.

2.3.2 细粒度形态识别

智能决策的首要任务是对象的精准识别. 就某一种产品设计而言, 同种产品由于应用场合需求、人机、功能、结构等差异, 在形态设计上也相应存在不同程度的差异. 例如手电钻, 可分为工地场景下的交流冲击钻、家用环境下的直流螺丝刀等.在真实的设计评价环境中, 同样会考虑这些关键因素进行差异化评价. 因此, 细粒度识别是产品形态智能决策的基础工作, 其本质是对设计定位的识别.

本节将细粒度产品形态识别任务表示为预测离散标签的分类建模问题. 将输入网络的产品图像表示为 $\boldsymbol{x}_n$, 基于 ResNet50 和 Swin-T 构造一个真实条件概率分布 $f_{\text{Form-CN}}(y\,|\,\boldsymbol{x}_n)$, 从而预测产品的形态类别 $f_{\text{Form-CN}}(\hat{y}\,|\,\boldsymbol{x}_n)$, 给定产品图像训练数据集 $D=\{(\boldsymbol{x}_n, y_n)\}_{n=1}^{N}$, 其中, $y_n\in\{1,2,\cdots,C\}$ 为类别标签, N 为训练集大小. 通过交叉熵损失函数来评价整个网络, 数学表达式为

$$L_{\text{Form-CN}}\left(y_{n,c}, f_{\text{Form-CN}}(\hat{y}_{n,c}\,|\,\boldsymbol{x}_n;\boldsymbol{w})\right)= -\frac{1}{2N}\sum_{n=1}^{N}\sum_{c=1}^{C} y_{n,c}\log \hat{y}_{n,c} \tag{3}$$

其中, $\hat{y}_{n,c}$ 是细粒度形态识别网络 Form-CN 的预测概率, $y_{n,c}$ 是产品图像真实类别.

具体网络结构设计见第 3.1 节.

2.3.3 整体形态语义评价

与文献[6,12-14,16]的离散分类评价决策不同, 本节将产品整体形态评价任务抽象为预测连续分数的回归评估建模问题. 分类评估结果是从属某个类别的概率值, 最终的得分还需进一步转化, 对设计方案的评价相对粗糙. 回归评估获得的分数直接且连续, 具有较强的针对性和自适应性, 适用于不同设计阶段. 因此, 回归评估模型对方案的评分相对精细, 也是真实环境中方案评估的首选.

将产品图像表示为 $\boldsymbol{x}_n$, 基于 ResNet101 和 Swin-S 创建一个深度回归评估映射函数 $f_{\text{Form-REN}}:\mathbf{R}^2\to\mathbf{R}$, 从而对产品进行预测评分 $\hat{y}_n=f_{\text{Form-REN}}(\boldsymbol{x}_n;\boldsymbol{w})$, 其中 $\boldsymbol{w}$ 是网络可学习的参数集. 给定产品图像训练数据集 D 由 N 个样本组成, 其中每个样本都是独立同分布的, 则训练集 D 表示为

$$D=\{(\boldsymbol{x}_n, y_n)\}, n\in[1,N] \tag{4}$$

其中, N 为训练集的大小, y_n 为人工依据产品评价标准给出的基准评分, 由式(2)计算得到.

通过实验对比了均方绝对误差(L_1)、均方误差(L_2)和平滑 L_1 (Smooth L_1)3 种回归损失函数, 发现 Smooth L_1[44]能够更好的量化所构建的 Form-REN 网络, 数学表达式为

$$L_{\text{Smooth}L1}\left(y_n, f_{\text{Form-REN}}(\boldsymbol{x}_n;\boldsymbol{w})\right)=\frac{1}{2N}\sum_{n=1}^{N} z_n \tag{5}$$

其中, $z_n=\begin{cases}\frac{1}{2}(y_n-\hat{y}_n)^2, & |y_n-\hat{y}_n|<\gamma \\ |y_n-\hat{y}_n|-\frac{1}{2}\gamma, & |y_n-\hat{y}_n|\geq\gamma\end{cases}$, 且 γ 是一个可变参数, 通常 $\gamma=1$; $\boldsymbol{w}$ 代表产品形态语义评价网络 Form-REN 的参数集; $\hat{y}_n$ 代表网络的预测值; N 为训练样本总数.

具体网络结构设计见第 3.2 节.

2.3.4 形态分布拟合评估

细粒度形态识别使得智能决策框架可聚焦评价对象. 回归评价模型可直接给出连续型的评分值, 但偏向主观性; 分布拟合评价模型是对产品形态差异分布区间上的评分, 本质上是产品部件造型特征的精细分类识别, 具有较好的客观性. 因此,二者的联合使得智能决策框架更完善, 决策结果相对客观合理. 此外, 设计评价贯穿设计活动的全部过程, 设计前期团队成员重点关注概念的创新;中期创新概念的有效转化成为主角; 后期则聚焦到细节和可实现性上. 与之对应的设计评价内容同样发生变化, 如果整体形态评价有助于前期概念的推进, 那么对局部细节差异的评估则是中后期推动概念深入的关键. 因此, 针对成型方案部件造型的细粒度形态语义评估是智能决策的核心工作之一.

综上所述, 为了实现产品造型特征级别的细粒度形态评估, 同时又能够比较方便地获取用于训练的数据及标签. 本节将这一任务归纳为一个分布拟合问题来解决[39]. 具体来说, 将任务先分解为分类识别任务, 然后对结果进行分布拟合评分计算. 因此, 目标函数使用交叉熵损失函数, 形式同公式(3), 用于计算 Form-DFEN 网络输出层对产品形态预测

分布与标注真实样本分布之间的差距, 数学表达式为

$$L_{\text{Form-DFEN}}\left(y_{n,c}, f_{\text{Form-DFEN}}\left(\hat{y}_{n,c} \mid \boldsymbol{x}_n; \boldsymbol{w}\right)\right) = -\frac{1}{2N}\sum_{n=1}^{N}\sum_{c=1}^{C} y_{n,c} \log \hat{y}_{n,c} \tag{6}$$

其中, $\hat{y}_{n,c}$ 是形态分布拟合网络 Form-DFEN 的预测概率, 即 SoftMax 输出的预测值, $y_{n,c}$ 是产品形态分布级别, 且 $y_{n,c} \in \{1,2,3,\cdots\}$.

至此实现产品形态分布评估任务, 最终的形态设计质量分布拟合评分(design quality distribution fitting, DF-Score)为

$$D_{\text{DF\_Score}} = \sum_{y_{n,c}\in\{1,2,3,\cdots\}} y_{n,c} \cdot f_{\text{Form-DFEN}}(\boldsymbol{x};\boldsymbol{w}) \tag{7}$$

其中, $y_{n,c}$ 是形态分布级别, $f_{\text{Form-DFEN}}(\boldsymbol{x};\boldsymbol{w})$ 代表 Form-DFEN 输出的预测概率分布值.

具体网络结构设计见第 3.3 节.

2.4 综合决策

对于给定的一批手电钻设计方案效果图 $\boldsymbol{x}$, 首先由 $f_{\text{Form-CN}}(\boldsymbol{x};\boldsymbol{w})$ 进行细粒度识别, 预测设计所属类别; 然后通过回归模型 $f_{\text{Form-REN}}(\boldsymbol{x};\boldsymbol{w})$ 和分布拟合模型 $f_{\text{Form-DFEN}}(\boldsymbol{x};\boldsymbol{w})$ 进行综合评估, 最终的评分为

$$E_{\text{OvScore}} = \frac{\alpha \cdot f_{\text{Form-REN}}(\boldsymbol{x};\boldsymbol{w}) + \beta \cdot D_{\text{DF\_Score}}}{2} \tag{8}$$

其中, α 和 β 分别表示 From-REN 和 Form-DFEN 的平衡因子; $D_{\text{DF\_Score}}$ 由公式(7)计算.

3 实验及结果分析

本文实验平台为 RTX2080/8 GB, CUDA11.0, Ubuntu18.04, PyTorch1.8.1.

3.1 细粒度形态识别实验

3.1.1 形态识别数据集

应用前文所述的数据处理方法, 通过人工标注的方式创建了一个具有细粒度标签的手电钻形态识别数据集(HD256CLA), 包括 4 种类型的手电钻, 共有 4 000 幅图像, 用于形态识别网络的训练和验证, 像素为 256×256, 数据样例见表 1.

表 1 HD256CLA 数据样例

名称	样例	数量	标签
交流冲击钻		1 000	wired
直流电池座		1 000	battery
直流无电池座		1 000	cordless
直流螺丝刀		1 000	screwdriver

此外, 为了便于测试和比较模型的最终性能,还创建了一个测试数据集, 同样包含上述 4 种类型, 共 100 张手电钻组成, 其中 32 张是设计师绘制的效果图, 另外 32 张由 PD-GAN[14]生成, 其余 36 张为是历史数据.

3.1.2 Form-CN 网络结构

ResNet50 和 Swin-T 有类似的 FLOPs(4 G vs 4.5 G), 基于二者设计了细粒度形态识别网络 Form-CN, 由特征提取器、特征输出器、特征变换器和分类器组成, 如图 4 所示.

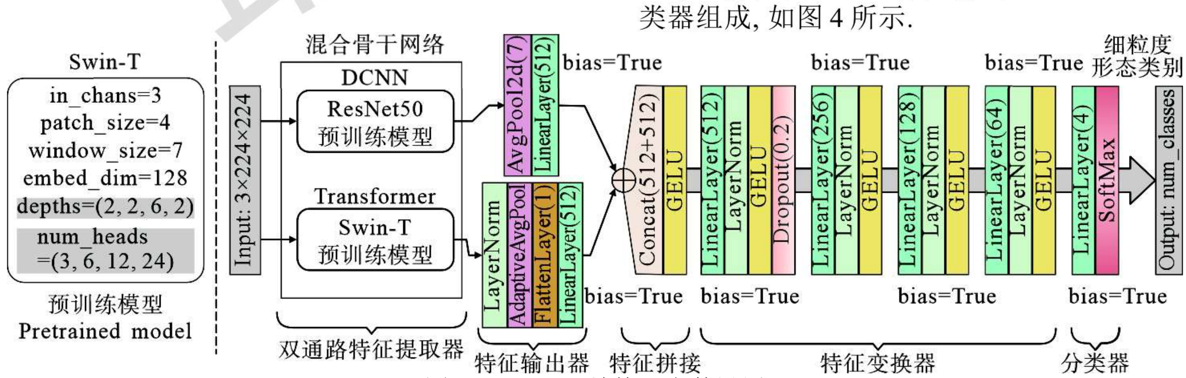


图 4 Form-CN 结构及参数设置

首先 ResNet50 和 Swin-T 同时对电钻图像进行特征提取, 其次特征输出器将提取的特征压缩输出, 然后进行特征变换, 最后由分类器完成图像分类识别. Swin-T 是 Swin Transformer 提出的最小的网络结构单元, 其网络深度为(2, 2, 6, 2), 共 12 个 STB, 分别对应自注意力头数为(3, 6, 12, 24); ResNet50 的特征输出器有平均池化层和 1 个全连接层(Linear Layer)组成; Swin-T 的特征输出器由层归一化(Layer Norm)、自适应平均汇聚层(AdaptiveAvg-Pool)、数据展平层(Flatten Layer)和 1 个全连接层组成; 特征变换器是 4 层的 MLP, 将特征从 1 024 维度变换到 64 维, 并对第 1 个全连接层使用了暂退法(Dropout); 分类器由 1 个全连接层和 1 个 SoftMax 层构成. 激活函数层均使用高斯误差线性单元 GELUs[45]. 网络输入为 3 通道的手电钻图像, 数据变换时将 256×256 像素处理成 224×224 像素, 网络输出以 4 个概率值判断所属类别, 并作 SoftMax 规范化处理.

3.1.3 网络训练与结果分析

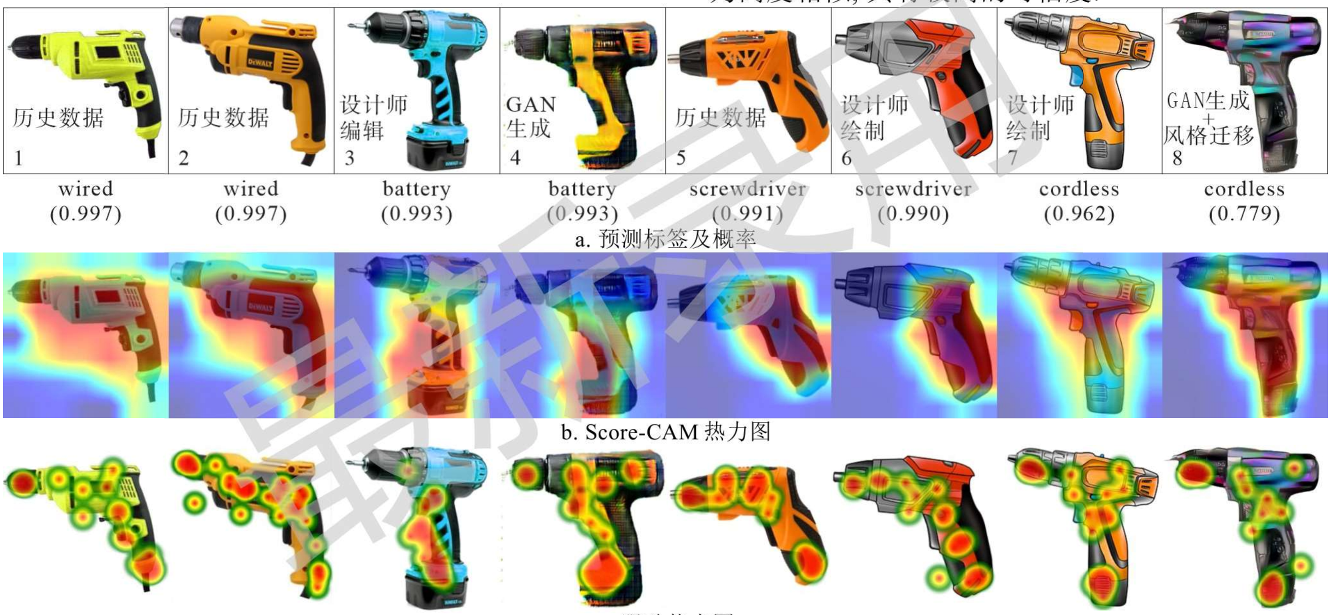
Form-CN 最大训练轮数 max_epoch 设为 60,批量大小 batch_size 为 32. 训练过程中将手电钻形态识别数据集 HD256CLA 划分为训练集和验证集, 比例为 8∶2. 使用 AdamW[46]优化器对 Form-CN 的参数进行优化, 初始学习率为 1E-4, β_1=0.9, β_2=0.999, 权重衰减为 5E-2.

在测试数据集上对训练好的 Form-CN 模型进行细粒度识别性能测试. 还训练了 7 个经典深度卷积神经网络 LeNet-5[47], AlexNet[48], VGG11[49], GoogLeNet[41], ResNet18[36], DenseNet[50] 和 ResNet50[36]与之对比, 结果见表 2. 将 LeNet-5 网络作为基准进行性能比较, Form-CN 在准确率、精确率、召回率和 F1 值 4 个评价指标均高于 Swin-T 以及其他 7 个 DCNNs 模型, 且 F1 值达 99.0%, 表明 Form-CN 模型最稳健. 另外, Form-CN 各项指标均高出 Swin-T 两个点, 而 Swin-T 和 ResNet50 的性能相当.

表 2 **Form-CN 与经典网络对比结果**

DCNNs	准确率	精确率	召回率	F_1 值	是否预训练
Form-CN	**0.990**	**0.990**	**0.990**	**0.990**	是
Swin-T	0.970	0.971	0.970	0.969	是
LeNet-5	0.950	0.945	0.940	0.940	否
AlexNet	0.952	0.949	0.950	0.950	否
VGG11	0.969	0.970	0.960	0.960	否
GoogLeNet	0.949	0.950	0.949	0.949	否
ResNet18	0.959	0.962	0.859	0.960	否
DenseNet	0.898	0.903	0.898	0.898	否
ResNet50	0.969	0.971	0.969	0.969	否

部分预测结果及可视化分析, 如图 5 所示. 由于 Score-CAM 的定位比 CAM 更精准[51], 因此使用 Score-CAM[51]可视化 Form-CN 的决策激活图. Score-CAM 热力图显示 Form-CN 在决策过程中所关注的不同类别手电钻造型特征区域不尽相同. 图 5 中的眼动实验热点图展示了人的视觉决策结果, 与 Score-CAM 热力图具有类似的结果, 表明 Form-CN 的神经细粒度决策行为与人的视觉经验决策行为高度相似, 具有较高的可信度.



a. 预测标签及概率

b. Score-CAM 热力图

c. 眼动热点图

图 5 Form-CN 决策结果及热力图分析(部分测试数据集)

3.2 整体形态语义评价实验

3.2.1 形态语义评价数据集

形态语义评价数据集(HD256REG)包含手电钻图像和相对应的人工基准评分. 具体以直流带电池底座的手电钻类型为对象, 从文献[14]筛选 1 606 幅较高质量的手电钻图像组成训练集和验证集, 部分示例见表 3.

表 3 **HD256REG 数据样例**

编号	样例	得分
0001		3.20
0002		3.50
0003		3.25
0004		3.00
…	…	…

对于人工基准评分, 首先确定手电钻设计评价标准, 然后组织人员对所有手电钻进行在线评分. 评价标准采取以用户为导向, 并结合相关文献的方式制订. 对网络爬虫抓取的 2 000 余条手电钻用户在线评论数据进行词频统计, 预选前 800 个分词进行筛选, 剔除噪声分词后共保留 507 个与设计相关的评论分词进行 K-means 聚类分析, 轮廓图显示聚为 5 类相对合理, 轮廓系数大于零且均值为 0.327. 并结合文献[52-53]为手电钻设置了 5 个维度的评价指标, 即{实用感, 造型美感, 动力感, 色彩方案, 操控舒适感}, 还设计了一个人工评估打分页面, 采用 5 级李克特量表形式, 如图 6 所示. 为降低人工评价强度, 提高人工基准评分质量, 将 1 606 幅手电钻图像随机分为 50 个组, 每组 30 个样本, 最后一组 16 个样本. 评价人员由工业设计师和用户组成. 采用公式(2)对每幅手电钻图像进行评价统计, 其中 5 个评价指标权重均为 1, 最终得分示例见表 3.

图 6 人工评估打分页面

另外, 为了测试和比较模型的最终性能, 还创建了一个回归评价测试数据集, 共有 120 幅手电钻图像组成. 其中设计师绘制效果图 20 幅, 历史图像数据 80 幅, PD-GAN[14]生成方案 20 个.

3.2.2 Form-REN 网络结构

ResNet101 与 Swin-S 具有相似的 FLOPs(7.6 G vs 8.7 G), 同时为更好的适应手电钻形态回归评估任务, 在二者的基础上设计了具有残差连接的 Form-REN, 如图 7 所示.

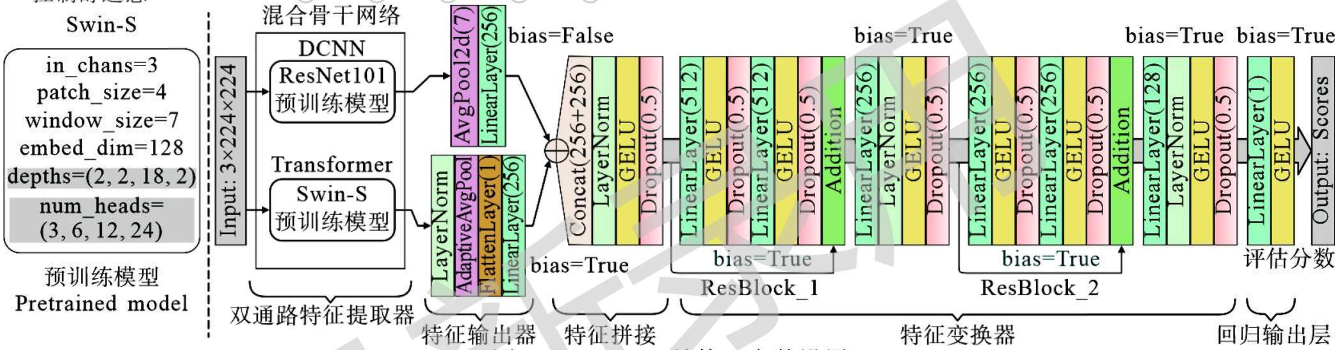


图 7 Form-REN 结构及参数设置

与 Swin-T 相比, Swin-S 在第 3 个特征提取阶段增加了 12 个 STB, 自注意力头数仍保持不变. 具体结构从左至右依次为, 双通路特征提取器、特征输出器、特征变换器, 回归输出层. 对两个特征输出器各抽取的 256 维特征进行拼接操作. 为了防止过拟合和网络退化, 特征变换器设计成具有两个残差块的 MLP 以提高网络性能, 并使用了 Dropout, 概率均为 0.5. 激活函数层均使用 GELUs[45]. 网络输入为 3 通道 224×224 像素的手电钻图像, 网络输出为单个回归评价分数.

3.2.3 网络训练与结果分析

Form-REN 最大训练轮数 max_epoch 设为 100, 批量大小 batch_size 为 8, 初始学习率为 1E-5. 使用 AdamW[46]优化器优化网络参数, 并使用权重衰减方法, 衰减系数为 1E-3. 训练过程中将手电钻数据集划分为训练集和验证集, 二者的比例为 7:3. 最好模型出现在第 86 轮, 其验证均方误差为 0.007 024 165.

在 120 张手电钻测试数据集上对 Form-REN 进行性能测试, 并训练了 9 个模型方法与之对比, 其中 Swin-S 和 ResNet101 为预训练模型, 结果见表 4. Form-REN 在 4 个回归评价指标平均绝对误差(mean absolute error, MAE)、均方误差(mean squared error, MSE)、均方根误差(root-mean-square error, RMSE)和平均绝对百分比误差(mean absolute percent error, MAPE)上的误差均最小, 其中 MSE 达 0.258 3. 部分测试集样例预测结果如图 8 所示.

表 4 Form-REN 与经典网络对比结果

DCNNs	MAE	MSE	RMSE	MAPE	是否预训练
Form-REN	**0.405 8**	**0.258 3**	**0.508 3**	**0.133 4**	是
Swin-S	0.530 4	0.479 8	0.692 7	0.157 9	是
ResNet101	0.537 5	0.495 3	0.703 8	0.159 7	是
LeNet-5	0.522 6	0.465 3	0.682 1	0.155 3	否
AlexNet	0.571 9	0.562 6	0.750 1	0.164 6	否
VGG11	0.532 1	0.482 1	0.694 3	0.158 0	否
GoogLeNet	0.530 7	0.478 1	0.691 5	0.157 9	否
ResNet18	0.524 7	0.468 7	0.684 6	0.157 0	否
DenseNet	0.537 5	0.499 0	0.706 4	0.160 0	否
ResNet50	0.533 8	0.483 5	0.695 4	0.160 2	否

图 8 Form-REN 回归评价结果(部分测试结果)

3.3 形态分布拟合评估实验

3.3.1 形态分布拟合评估数据集

应用 2.2 节所述的数据处理方法, 针对带电池底座的手电钻图像数据, 通过人工标注的方式创建了一个具有细粒度形态风格差异等级的手电钻分布拟合数据集(HD256DFE). 人工标注时的评价指标参考图 6 中的评价指标, 共设置了 5 个评估级别: 1—bad, 2—poor, 3—fair, 4—good, 5—excellent. 基础数据为 4 352 幅图像, 经过设计师编辑和数据增广将数据扩充至 7 602 幅图像, 用于形态分布拟合网络 Form-DFEN 的训练和验证, 像素为 256×256, 部分数据样例见表 5.

表 5 HD256DFE 数据样例

评估级别	样例	数量	标签级别
较差		1 344	1
一版		954	2
中等		1 532	3
好		1 778	4
优秀		1 994	5

另外, 为了测试模型的最终性能, 还创建了一个分布拟合测试数据集, 同样包含上述 5 个级别,共 153 张手电钻图片组成. 其中设计师绘制 50 张效果图, 利用历史数据合成 43 张, 其余 60 张由 PD-GAN[14]生成.

3.3.2 Form-DFEN 网络结构

与前两个网络执行分类和回归任务相比, 形态分布拟合评价是对同类型手电钻形态进行部件级细粒度识别及评估, 形态差异要求聚焦到部件造型细节, 同时模型应具有较强的稳健性和较好的泛化能力. 为适应这种相对复杂的识别及评价任务, 双通路特征提取器使用更高复杂度的网络结构, 即 ResNet152 和 Swin-B(FLOPs: 11.3 G vs 15.4 G)构成, 如图 9 所示.

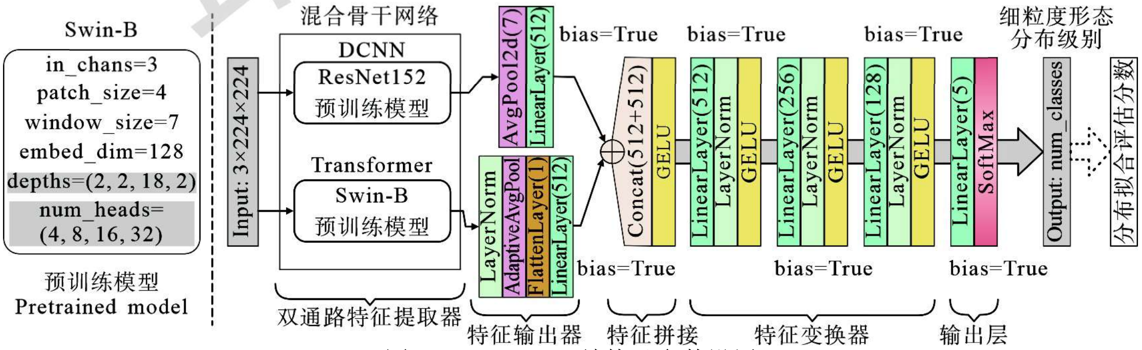


图 9 Form-DFEN 结构及参数设置

与 Swin-S 不同, Swin-B 在每个阶段的自注意力头数从(3, 6, 12, 24)变为(4, 8, 16, 32), 网络深度相同, 而 ResNet152 由 151 个卷积层和 1 个线性层构成. 手电钻图像同时输入到 ResNet152 和 Swin-B 两个通路, 每个通路经过相同结构的特征输出器分别输出 512 维的特征, 然后进行特征拼接操作获得 1 024 维的特征, 接着经 3 层全连接层组成的特征变换器, 最后由输出层输出. 特征变换器使用 Layer Norm 进行层归一化, 所有激活函数层使用 GELUs[45].

3.3.3 网络训练与结果分析

Form-DFEN 最大训练轮数 max_epoch 设为 80, 批量大小 batch_size 为 10. 使用 AdamW[46]优化器训练 Form-DFEN, 初始学习率为 1E-4, β_1 =0.9, β_2 =0.999, 使用权重衰减策略, 衰减值为 5E-2. 训练时将数据集 HD256DFE 按 8:2 划分为训练集和验证集. 网络输入为 224×224 像素手电钻图像, 输出为 5 个级别的形态分布概率, 最终的结果通过公式(7)计算出形态的分布拟合评估分数.

在测试集上对 Form-DFEN 模型进行性能测试, 并训练了 7 个经典 DCNNs 模型方法与之对比, 结果见表 6. 同时给出部分测试结果, 如图 10 所示.

表 6 Form-DFEN 与经典网络对比结果

DCNNs	准确率	精确率	召回率	F_1值	是否预训练
Form-	**0.843**	**0.859**	**0.843**	**0.84**	是

DFEN				3	
Swin-S & Swin-B	0.797	0.834	0.797	0.804	是
ReNet152	0.810	0.834	0.810	0.812	是
LeNet-5	0.728	0.793	0.728	0.733	否
AlexNet	0.756	0.756	0.708	0.718	否
VGG16	0.728	0.767	0.728	0.732	否
GoogLeNet	0.622	0.698	0.622	0.621	否
ResNet18	0.708	0.735	0.708	0.702	否
ResNet50	0.730	0.735	0.662	0.671	否

数据说明: 预测标签级别(概率值)/分布拟合得分

图 10 Form-DFEN 决策结果(部分样例)

通过对比表 6 中的所有模型结果可知, 具有混合双通路设计特点的 Form-DFEN 模型在 4 个评价指标上, 不仅高于 6 个非预训练 DCNNs 模型, 而且高出预训练的 ResNet152 和 Swin-S & Swin-B 双通路模型, 且 F1 值达 84.3%, 可知模型的稳健性相对更好, 这表明混合迁移学习模型的性能在复杂任务上表现更优异.

3.4 智能综合评价与眼动验证

为了进一步验证和解释神经网络决策的可信度, 本小结对比了神经网络综合评价结果和眼动实验评估决策结果. 神经网络综合决策结果依据公式(8)计算, 由于两个网络均采用 5 分制评分标准, 并且两种评估方式同等重要, 因此 $\alpha=\beta=1$.这里对设计师定义的 5 个方案和 GAN 生成的 3 个方案进行了综合评价, 如图 11 所示. 采用眼动实验验证的方式对图中的 8 个方案进行了对比, 同时给出了设计师的综合评分, 如图 11 中的第 6, 7 行.从眼动实验热点图和视觉兴趣域(Areas of Interest, AOI)分布图中可知, 人类评价决策对于一些设计细节给予更多时间的思考, 这是由于人具有设计知识和经验, 评价过程会显得相对细腻, 而神经网络决策不主动依赖设计知识和经验. 对比最终的评价分数发现, 神经网络综合评分和人工评分相当, 仅在方案 6(3.046 vs. 4.4)和方案 8(2.626 vs. 3.6)上的绝对误差较为明显, 分别为 1.354 和 0.974, 表明人对 GAN 生成的方案给予了更宽容的评价态度, 而神经网络则不然.

图 11 智能综合评分与眼动评估决策

4 讨 论

4.1 实验结果讨论

对于 Form-CN 的分类定位决策结果, 由图 5 可知, Form-CN 对交流冲击钻的决策准确率最高, Score-CAM 热力图显示神经网络几乎关注了该类型产品的全部形态, 而眼动热点图显示了类似的结果. 这表明神经网络在决策过程中对于该类型的产品给予了更多的视觉关注, 因此决策准确率最高. 排名次之的是直流带电池底座型手电钻, Score-CAM 热力图和眼动热点图显示神经网络和人类视觉都聚焦到了电池座. 排名第 3 的是形态小巧的直流螺丝刀, Score-CAM 热力图显示神经网络在决策依据转移到了手柄和变速按钮. 排名最后的是直流无电池座手电钻, Score-CAM 热力图显示神经网络将手柄与机身交接处作为决策焦点特征, 而眼动热点图表明 Form-CN 与人的视觉关注点相近. 由于使用了 AM, 使得神经网络决策关注到更重要的设计细节.

对于 Form-REN 的回归评价结果, 由图 8 可知, 模型对于形态和色彩具有较好的敏感性. 对第 1 行的中 6 个方案给出了较明显差异的回归评分, 其中比较特别的是对第 6 个方案给出了 2.381 的低分.其一是因为该方案所属细粒度类型与训练 Form-REN 的数据类型不同; 其二是该设计的适用场景为家庭环境, 而文中采集的数据集和标签主要面对工业环境中的手电钻; 另外, 该方案在设计风格和理念与其他方案存在较大差异. 第 2 行主要测试了神经网络对色彩方案的评价决策. 第 1 组的色彩方案注重高低速调节、变速按钮、正反转按钮和电池开启按钮 4 个操作按钮; 第 2 组色彩方案重点强调手柄和机身后部; 第 3 组色彩方案同时包括手柄和 4 个操作按钮. 评分结果表明, 在造型特征相同的条件下, Form-REN 对色彩方案具较好的敏感性, 3 组方案的回归评分依次降低.

对于 Form-DFEN 的分布拟合评价结果, 由图 10 可知, 预测级别与最终的分布拟合得分并非一致, 后者相对综合和客观, 这是分布拟合评价的优点, 即概率分布拟合具有自适应性. 方案 1 预测级别是 4(概率为 0.442), 而分布拟合得分为 3.293, 表现出不一致性. 方案 2~方案 5 预测级别均为 5, 概率依次降低, 其中方案 5 的概率最低, 可能是由于该方案是个合成方案(由设计师主观制作), 给 Form-DFEN 的决策带来了一些挑战, 但 4 个方案表现出了一致性. 方案 6 的预测级别是 3(概率为 0.475), 但最终的分布拟合得分为 3.894, 表现出不一致性. 方案 7 和 8 是 GAN 生成的方案, 预测级别均为 4, 但二者的概率差别显著, 后者的概率明显高于前者, 但分布拟合得分则相反, 在同级别中也表现出了不一致性. 其余 5 个方案同样表现出了分布拟合的这种特性, 层级化鲜明.

此外, Form-REN 的回归评价直接且精细, 但受人工标注基准回归评分数据的影响, 其评分存在一定的主观性; Form-DFEN 则是从客观数据分布的角度出发, 所获得的评分是分布区间上的某个值, 结果相对客观, 但缺乏一定的精准度. Form-REN 的回归评分与 Form-DFEN 的分布拟合评价联合可相互弥补各自的不足, 使得神经网络综合评分与人工评分之间的相对误差较小, 表明模型稳定、有效, 具有一定的可信度, 如图 11 所示.

设计决策团队在进行设计方案评价过程中, 需要借助所掌握的知识和视觉经验对众多方案给出综合决策评分, 基于 ResNets 和 Swin Transformer 构建的智能设计混合决策框架使得 AI 模型模拟了设计决策者们的细粒度视觉评价特性. 因此, 该项工作既是对设计应用型 AI(design applied AI, DA-AI)的扩展, 也是对 AI 工业设计评价决策系统的深化和补充.

现阶段人工决策以设计视觉为导向、以设计知识为前提, 而智能设计决策主要以 CV 为基础. 面对产品历史图像数据, 深度神经网络采用端到端的训练进行学习, 测试样本从训练好的模型直接产生决策结果, 但是这种黑箱操作会让设计决策群体缺乏信任感. AI 决策在真实环境应用时, 若作为辅助决策, 需要和人的决策进行耦合, 若为自主决策, 必须让人完全相信[54]. 为突破思维惯性和去同质化, 设计师在创造产品形态时融入诸如人机特性、品牌识别、结构工艺、材质色彩等造型要素和特征. 它们可能会被神经网络隐式的发现并在决策中使用, 但很难将其从外部直接提供给神经网络并显式地使用. 然而, 人的视觉感知和神经系统构成的决策闭环与深度学习决策方法不同, 对造型要素和特征的决策可出可入、动态显隐, 特别是对设计理念的评价, 人工神经网络还不能比拟人类在决策过程中表现出的心智能力[55]. 关于这一点, 神经网络在图 8 方案 6 上的表现得到了证实. 因此, 针对智能决策受益者, 深度神经评价网络模型的可解释性是一项保证性工作, 在此基础上的人机协作评价是一个重要的研究方向.

4.2 数据集与未来问题

目前工业设计研究领域鲜有开源标准数据集, 相关工作均使用自己创建的数据集[6,12,14-16,56], 且未公开, 所训练模型的性能评分相对孤立. 这给不同设计评价决策模型之间的性能比较带来一定障碍, 同时也不利于智能设计决策研究的发展. 因此,发展开源标准产品设计评价数据集是一项基础工作, 更是发展产品形态智能设计与评价预训练模型的保障. 另外, 分类标签和回归评分数据不可避免的具有主观性, 且获取方式相对单一. 因此, 为提高基准评价数据的质量, 通过实验评价的方式获取生理数据[1,57], 可进一步提升数据标签质量、扩展标签类型模式, 进而发展多模态深度神经决策模型.

未来有两个重要的问题值得长期关注. 一方面, 对于真实环境中的设计评价及决策, 参与者们主要依靠设计知识和经验, 并以视觉互动的方式做出综合评估决策, 但以特征学习为主的深度神经决策缺乏设计知识和经验的引导, 自适应性和动态性相对较弱, 因此也存在一定的局限性. 同时也为未来研究指出了一个有价值的方向, 即进一步深入探索人机协作设计决策系统. 另一方面, 当智能设计决策模型的效率和准确率达到一定的水准后, 还需要关注设计智能的闭环性, 如语义部件检测和定位有助于为设计智能闭环系统提供精准有效的反馈信号, 进而对设计方案进行智能编辑[56].

5 结 论

设计智能研究涉及用户需求分析、概念构思、方案生成和设计评价[23], 其中设计评价直接关系到方案的决策结果. 方案效果图评估是现阶段以人为主导的工业设计决策过程中的重要环节, 本文结合 AI, CV 研究动向, 提出一个相对完整的智能决策方法框架, 包括细粒度的形态分类、整体形态语义评价和精细形态分布拟合检测. Form-CN 从形态类内间差异出发, 经形态分类辨识出产品形态最初的设计定位; Form-REN 从用户的角度切入, 尝试连续性的感知回归评价; Form-DFEN 则从产品部件造型特征优劣分类的角度出发, 探索产品形态本身的等级化分布特性. 3 个网络的联合构成了一个集成智能决策系统.

提出的双通路并行混合网络框架为特征迁移学习提供了新思路, 融合 Swin Transformer 的动态信息与 ResNets 的静态信息提取能力, 形成并行复合决策机制, 使得模型具有更强的特征提取和识别能力. Form-CN 模型在手电钻测试数据集上的准确率达 99.0%, 可实现相对精准的细粒度设计定位识别. From-REN 可对方案直接给出形态语义评分, 在对比模型中的各项误差指标最小, 其中均方误差为 0.2583; 而 Form-DFEN 则是从数据分布角度切入, 在测试集上的准确率可达 84.3%. 二者的联合使得评价决策结果具有更强的可信度, 能够对设计师定义的方案和 GAN 生成的方案给出一个相对客观的评分, 可有效帮助设计团队进行快速决策, 并为优化设计方向提供参考依据.

本文提出的双通路并行混合网络框架也可以应用到其他细粒度识别和评价任务, 针对类内间差异及目标种类数量, 构建有针对性的特征变换器网络结构和输出层即可.

参考文献(References):

[1] Wang Y, Yu S, Ma N, *et al*. Prediction of product design decision making: An investigation of eye movements and EEG features[J]. Advanced Engineering Informatics, 2020, 45: 101095

[2] Yang Yanpu, Chen Dengkai, Cheng Peiyao. Decision-making problems of industrial design[J]. Packaging Engineering, 2020, 41(4): 1-6(in Chinese)
(杨延璞, 陈登凯, 成沛瑶. 工业设计决策问题研究[J]. 包装工程, 2020, 41(22): 1-6)

[3] Yang Yanpu, Gu Rong, Gu Yufeng. Method for reaching a consensus in industrial design evaluation based on PAM clustering[J]. Computer Integrated Manufacturing Systems, 2017, 23(7): 1386-1393(in Chinese)
(杨延璞, 顾蓉, 古玉锋. 基于 PAM 聚类的工业设计方案评价共识度达成方法[J]. 计算机集成制造系统, 2017, 23(7): 1386-1393)

[4] Minar M R, Naher J. Recent advances in deep learning: An overview[OL]. [2021-12-02]. https://arxiv.org/abs/1807.08169

[5] Zhao W, Jiang W, Qiu X. Deep learning for COVID-19 detection based on CT images[J]. Scientific Reports, 2021, 11: 14353

[6] Wang Yahui, Yu Suihuai, Chen Dengkai, *et al*. Artificial intelligence design decision making model based on deep learning[J]. Computer Integrated Manufacturing Systems, 2019, 25(10): 2467-2475(in Chinese)
(王亚辉, 余隋怀, 陈登凯, 等. 基于深度学习的人工智能设计决策模型[J]. 计算机集成制造系统, 2019, 25(10): 2467-2475)

[7] Yang Chunying. Research on the decision model of automobile conceptual design based on the theory of data mining[D]. Wuhan: Wuhan University of Technology, 2011(in Chinese)
(杨春莹. 基于数据挖掘的汽车整车方案设计决策模型的研究[D]. 武汉: 武汉理工大学, 2011)

[8] Su Jianning, Zhang Qinwei, Wu Jianghua, et al. Evolutionary design product multi-image styling[J]. Computer Integrated Manufacturing Systems, 2014, 20(11): 2675-2682(in Chinese)
(苏建宁, 张秦玮, 吴江华, 等. 产品多意象造型进化设计[J]. 计算机集成制造系统, 2014, 20(11): 2675-2682)

[9] D'Anjou P. An alternative model for ethical decision-making in design: A Sartrean approach[J]. Design Studies, 2011, 32(1): 45-59

[10] Hsiao S W, Huang H C. A neural network based approach for product form design[J]. Design Studies, 2002, 23(1): 67-84

[11] Diego-Mas J A, Alcaide-Marzal J. Single users' affective responses models for product form design[J]. International Journal of Industrial Ergonomics, 2016, 53(2016): 102-114

[12] Zhu Bin, Yang Cheng, Yu Chunyang, *et al*. Product image recognition based on deep learning[J]. Journal of Computer-Aided Design & Computer Graphics, 2018, 30(9): 1778-1784(in Chinese)
(朱斌, 杨程, 俞春阳, 等. 基于深度学习的产品意象识别[J]. 计算机辅助设计与图形学学报, 2018, 30(9): 1778-1784)

[13] Gong Y, Wang B, Rau P L P. PIKAR: A pixel-level image Kansei analysis and recognition system based on deep learning for user-centered product design[J]. Cross-Cultural Design User Experience of Products, Services, and Intelligent Environments, 2020: 60-71

[14] Li X, Su J, Zhang Z, *et al*. Product innovation concept generation based on deep learning and Kansei engineering[J]. Journal of Engineering Design, 2021, 32(10): 559-589

[15] Su Z, Yu S, Chu J, *et al*. A novel architecture: Using convolutional neural networks for Kansei attributes automatic evaluation and labeling[J]. Advanced Engineering Informatics, 2020, 44: 101055

[16] Pei Huining, Huang Xueqin, Li Haitao, *et al*. Decision model of product form design based on capsule network[J]. Computer Integrated Manufacturing Systems, 2022, 28(3): 853-863(in Chinese)
(裴卉宁, 黄雪芹, 李海涛, 等. 基于胶囊网络的产品形态设计决策模型[J]. 计算机集成制造系统, 2022, 28(3): 853-863)

[17] Wang W M, Wang J W, Li Z, *et al*. Multiple affective attribute classification of online customer product reviews: A heuristic deep learning method for supporting Kansei engineering[J]. Engineering Applications of Artificial Intelligence, 2019, 85: 33-45

[18] Wang Xiaohui, Qin Jingyan. Affective modeling of social networks based on deep recurrent neural networks[J]. Packaging Engineering, 2021, 42(4): 77-82(in Chinese)
(王晓慧, 覃京燕. 基于深度循环神经网络的社交网络用户情感研究[J]. 包装工程, 2021, 42(4): 77-82)

[19] Du Peng, Lu Yiqing, Han Changfeng. Sentiment analysis of commodity reviews based on Transformer model[J]. Journal of Chinese Information Processing, 2021, 35(2): 125-132(in Chinese)
(杜朋, 卢益清, 韩长风. 基于 Transformer 模型的商品评论情感分析[J]. 中文信息学报, 2021, 35(2): 125-132)

[20] Wang Y, Mo D Y, Tseng M M. Mapping customer needs to design

parameters in the front end of product design by applying deep learning[J]. CIRP Annals, 2018, 67(1): 145-148
[21] Wang Weining, Wang Li, Zhao Mingquan, *et al*. Image aesthetic classification using parallel deep convolutional neural networks[J]. Acta Automatica Sinica, 2016, 42(6): 904-9143(in Chinese)
(王伟凝, 王励, 赵明权, 等. 基于并行深度卷积神经网络的图像美感分类[J]. 自动化学报, 2016, 42(6): 904-914)
[22] Goodfellow I, Pouget-Abadie J, Mirza M, *et al*. Generative adversarial nets[C] //Proceedings of the 27th International Conference on Neural Information Processing Systems. Cambridge: MIT Press, 2014: 2672-2680
[23] Tang y C, Huang J J, Yao M T, *et al*. A review of design intelligence: progress, problems, and challenges[J]. Frontiers of Information Technology & Electronic Engineering, 2020, 20(12): 1595-1617
[24] Liu Z, Gao F, Wang Y. A generative adversarial network for AI-aided chair design[C] //Proceedings of 2019 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR), Los Alamitos: IEEE Computer Society Press, 2019: 486-490
[25] Dai Y, Li Y, Liu L J. New product design with automatic scheme generation[J]. Sensing and Imaging, 2019, 20(1): 1-16
[26] Liu Yuezhong, Stouffs R, Yang Yang. Urban design process with conditional generative adversarial networks[J]. Architectural Journal, 2018(9): 108-113(in Chinese)
(刘跃中, 斯托夫斯·卢迪, 杨阳. 基于条件生成对抗网络的城市设计研究[J]. 建筑学报, 2018(9): 108-113)
[27] Nauata N, Hosseini S, Chang K H, *et al*. House-GAN++: Generative adversarial layout refinement networks[OL]. [2021-12-02]. https://arxiv.org/abs/2103.02574
[28] Nauata N, Chang K H, Cheng C Y, *et al*. House-GAN: Relational generative adversarial networks for graph-constrained house layout generation[OL]. [2021-12-02]. https://arxiv.org/abs/2003.06988v1
[29] Yang X, Mei T, Xu Y-Q, *et al*. Automatic generation of visual-textual presentation layout[J]. ACM Transactions on Multimedia Computing, Communications, and Applications, 2016, 12(2): 1-22
[30] Sage A, Agustsson E, Timofte R, *et al*. Logo synthesis and manipulation with clustered generative adversarial networks[C] // Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Los Alamitos: IEEE Computer Society Press, 2018: 5879-5888
[31] Zhao Yili, Xu Dan. Joint semantic parts for fine-grained bird images recognition[J]. Journal of Computer-Aided Design & Computer Graphics, 2018, 30(8): 1522-1529(in Chinese)
(赵毅力, 徐丹. 联合语义部件的鸟类图像细粒度识别[J]. 计算机辅助设计与图形学学报, 2018, 30(8): 1522-1529)
[32] Cai Liwei, Hu Xiaofeng, Liu Mengxiang. Prediction method of tool remaining useful life based on transfer learning[J]. Computer Integrated Manufacturing Systems, 2021, 27(6): 1541-1549(in Chinese)
(蔡伟立, 胡小锋, 刘梦湘. 基于迁移学习的刀具剩余寿命预测方法[J]. 计算机集成制造系统, 2021, 27(6): 1541-1549)
[33] Parker S. The human body book[M]. Shanghai: Shanghai Scientific & Technical Publishers, 2014(in Chinese)
(帕克. 人体[M]. 上海: 上海科学技术出版社, 2014)
[34] Guo M H, Xu T X, Liu J J, *et al*. Attention mechanisms in computer vision: A survey[OL]. [2021-12-02]. https://arxiv.org/ abs/2111.07624
[35] Liu Z, Lin Y, Cao Y, *et al*. Swin transformer: Hierarchical vision transformer using shifted windows[C] //Proceedings of IEEE/CVF International Conference on Computer Vision (ICCV). Los Alamitos: IEEE Computer Society Press, 2021: 9992-1000
[36] He K, Zhang X, Ren S, *et al*. Deep residual learning for image recognition[C] //Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Los Alamitos: IEEE Computer Society Press, 2016: 770-778
[37] Zhou B, Khosla A, Lapedriza A, *et al*. Learning deep features for discriminative localization[C] // Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Los Alamitos: IEEE Computer Society Press, 2016: 2921-2929
[38] Cheon M, Yoon S J, Kang B, *et al*. Perceptual image quality assessment with Transformers[C] //Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Los Alamitos: IEEE Computer Society Press, 2021: 433-442
[39] Talebi H, Milanfar P. NIMA: Neural image assessment[J]. IEEE Transactions on Image Processing, 2018, 27(8): 3998-4011
[40] Chopra S, Hadsell R, Lecun Y. Learning a similarity metric discriminatively, with application to face verification[C] //Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR). Los Alamitos: IEEE Computer Society Press, 2005: 539-546
[41] Szegedy C, Liu W, Jia Y, *et al*. Going deeper with convolutions[C] //Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Los Alamitos: IEEE Computer Society Press, 2015: 1-9
[42] Vaswani A, Shazeer N, Parmar N, *et al*. Attention is all you need[C] //Proceedings of 31st Conference on Neural Information Processing Systems (NIPS). New York: ACM Press, 2017: 6000–6010
[43] Dosovitskiy A, Beyer L, Kolesnikov A, *et al*. An image is worth 16x16 words: Transformers for image recognition at scale[OL]. [2021-12-02]. https://arxiv.org/abs/2010.11929
[44] Girshick R. Fast R-CNN[C] //Proceedings of 2015 IEEE International Conference on Computer Vision (ICCV). Los Alamitos: IEEE Computer Society Press, 2015: 1440-1448
[45] Hendrycks D, Gimpel K. Gaussian error linear units (GELUs) [OL]. [2021-12-02]. https://arxiv.org/abs/1606.08415v4
[46] Loshchilov I, Hutter F. Decoupled weight decay regularization[OL]. [2021-12-02]. https://arxiv.org/abs/1711.05101v3
[47] Lecun Y, Bottou L. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324
[48] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[C] //Proceedings of the 25th International Conference on Neural Information Processing Systems (NIPS). New York: ACM Press, 2012:1097-1105
[49] Simonyan K, Zisserman A. Very Deep convolutional networks for large-scale image recognition[OL]. [2021-12-02]. https://arxiv.org/abs/1409.1556v6
[50] Huang G, Liu Z, Laurens V, *et al*. Densely connected convolutional networks[C] //Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Los Alamitos: IEEE Computer Society Press, 2017: 2261-2269

[51] Wang H, Wang Z, Du M, *et al*. Score-CAM: Score-weighted visual explanations for convolutional neural networks[C] //Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Los Alamitos: IEEE Computer Society Press, 2020: 111-119

[52] Grimsaeth K, Gangwar M, Baggerud B, *et al*. Kansei engineering based emotional product design: An illustration by a battery drill case-study[J]. Kansei Engineering International Journal, 2010, 9(2): 67-79

[53] Wei Yongxia. The design of electric drill based on Kansei engineering[J]. Packaging Engineering, 2016, 37(24): 108-113(in Chinese)
(魏永侠. 基于感性工学的手电钻设计[J]. 包装工程, 2016, 37(24): 108-113)

[54] Kong Xiangwei, Tang Xinze, Wang Ziming. A survey of explainable artificial intelligence decision[J]. Systems Engineering Theory & Practice, 2021, 41(2): 524-536(in Chinese)
(孔祥维, 唐鑫泽, 王子明. 人工智能决策可解释性的研究综述[J]. 系统工程理论与实践, 2021, 41(2): 524-536)

[55] Xu Zhilei, Dong Zhanxun. Cognitive science and design[J]. Packaging Engineering, 2021, 42(2): 1-4(in Chinese)
(徐志磊, 董占勋. 认知科学与设计研究[J]. 包装工程, 2021, 42(2): 1-4)

[56] Nobari H A, Rashad M F, Ahmed F. CreativeGAN: Editing generative adversarial networks for creative design synthesis[OL]. [2021-12-02]. https://arxiv.org/abs/2103.06242v1

[57] Feng Yixiong, Lou Shanhe, Wang Xupeng, *et al*. Research on performance-oriented perceptual image evaluation method for customized products[J]. Journal of Mechanical Engineering, 2020, 56(9): 181-190(in Chinese)
(冯毅雄, 娄山河, 王绪鹏, 等. 面向性能的定制产品感性意象评价方法研究[J]. 机械工程学报, 2020, 56(9): 181-190)

稿件录用通知

李雄，苏建宁，张志鹏，祝铎，鱼宝银同志：

您投来的题为**《基于深度学习的产品风格精细识别研究》**，（稿件编号：2022-0550），经专家评审后决定录用，特此通知。

版面费 5400 元。

您的稿件刊登以后，我们将给您寄去相应的稿费，并给所有作者寄去一份期刊。如您的通信地址有变动，请及时通知我们。

如产出论文的课题题目系基金资助项目，且获得了有关部门的科学技术等奖励，请将获奖证书复印件寄送编辑部，以作统计用。

欢迎您继续为本刊撰写稿件，谢谢您的支持与合作。

此致

敬礼！

北方科技信息研究所

《计算机集成制造系统》编辑部

2023 年 4 月 13 日

证书编号：2023042342259

中国知网学术期刊网络首发论文出版证书

《中国学术期刊（网络版）》是教育部主管、清华控股有限公司主办、国家新闻出版广电总局 2015 年 8 月 14 日批准的网络版连续出版物（国际标准连续出版物号 ISSN 2096-4188；国内统一连续出版物号 CN 11-6037/Z），在《中国学术期刊（光盘版）》电子杂志社有限公司的互联网出版网站中国知网（www.cnki.net，新出网证（京）字 008 号）上进行出版。

《计算机集成制造系统》与《中国学术期刊（光盘版）》电子杂志社有限公司合作，在中国知网创办了与《计算机集成制造系统》内容一致的网络版，以单篇或整期出版形式，在印刷版出版之前刊发纸质期刊已正式录用定稿的论文，著作权受法律保护。论文发表时间按中国知网的网络出版时间确认。

兹有，李雄，苏建宁，张志鹏，祝铎，鱼宝银同志的题为《基于深度学习的产品风格精细识别》的论文

链接地址：https://kns.cnki.net/kcms/detail/11.5946.TP.20230423.1050.002.html

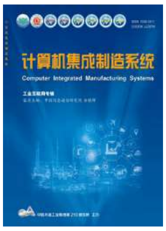
已于 2023 年 04 月 23 日在中国知网出版，出版证书验证地址 www.cnki.net，特此证明。

《计算机集成制造系统》编辑部（签章）

计算机集成制造系统 编辑部

《中国学术期刊（光盘版）》电子杂志社有限公司（签章）

2023 年 04 月 23 日



计算机集成制造系统
Computer Integrated Manufacturing Systems
ISSN 1006-5911,CN 11-5946/TP

《计算机集成制造系统》网络首发论文

题目：　　　　基于深度学习的产品风格精细识别
作者：　　　　李雄，苏建宁，张志鹏，祝铎，鱼宝银
收稿日期：　　2022-07-11
网络首发日期：2023-04-23
引用格式：　　李雄，苏建宁，张志鹏，祝铎，鱼宝银．基于深度学习的产品风格精细识别[J/OL]．计算机集成制造系统．
https://kns.cnki.net/kcms/detail/11.5946.TP.20230423.1050.002.html

基于深度学习的产品风格精细识别

李 雄[1,3]，苏建宁[1,2+]，张志鹏[1]，祝 铎[2]，鱼宝银[2]
(1.兰州理工大学 机电工程学院，甘肃 兰州 730050；2.兰州理工大学 设计艺术学院，甘肃 兰州 730050；3.兰州城市学院 培黎机械工程学院，甘肃 兰州 730070)

摘 要：为有效提取具有差异性的产品风格特征，提出一种基于复合学习通路的细粒度风格识别卷积神经网络(Fine-grained style recognition convolutional neural networks，FSR-CNN)。一是注意力学习通路，以残差结构为基础，采用串并结合的方式将坐标注意力、卷积块注意力和多头注意力嵌入其中，提出轻量化的混合注意力残差网络(Hybrid attention-based ResNet，HA-ResNet)，用于抽取“专用特征”。二是迁移学习通路，应用微调预先训练的 GoogLeNet 以扩充 HA-ResNet 模型容量，实现多感受野“通用特征”抽取。最后对二者输出的特征进行融合，并使用 MLP 分类器识别产品风格类型。在自行车头盔数据集上进行实验，并与其他的经典深度卷积神经网络模型进行比较，实验结果表明 FSR-CNN 模型表现出较高的准确率和良好的稳健性，为产品风格精细检索与知识重用提供一种新的模型算法架构。

关键词：产品造型；风格识别；混合注意力；迁移学习；复合学习机制
中图分类号：TB472；TP183　　　　**文献标识码**：A

Recognition method for fine-grained product styles based on deep learning

LI Xiong[1,3]，SU Jianning[1,2+]，ZHANG Zhipeng[1]，ZHU Duo[2]，YU Baoyin[2]
(1.School of Mechanical & Electrical Engineering, Lanzhou University of Technology, Lanzhou 730050, China；2.School of Design Art, Lanzhou University of Technology, Lanzhou 730050, China；3. School of Bailie Mechanical Engineering, Lanzhou City University, Lanzhou 730070, China)

Abstract：A fine-grained styles recognition convolutional neural network (FSR-CNN) based on composite learning pipelines is proposed to effectively extract product style features with differences. The first one is the attention learning pipeline, which is based on the residual structure and embeds coordinate attention, convolutional block attention and multi-head attention in a string-parallel combination to form a lightweight hybrid attention residual network (HA-ResNet) for extracting “specialized features”. Secondly, the transfer learning pipeline is used to fine-tune the pre-trained GoogLeNet to expand the capacity of HA-ResNet model for extracting multi-receptive field “generic features”. Finally, the output features of both are fused and the MLP classifier is used to identify the product style types. Experiments are performed on a self-built bicycle helmet dataset and compared with other classical deep convolutional neural network models. The experimental results show that the FSR-CNN model exhibits higher accuracy and stronger robustness, providing a new model algorithm architecture for product styles fine retrieval and reuse.

Keywords：product form; style recognition; hybrid attention; transfer learning; composite learning mechanism

1 引言

风格策略及设计是提升产品设计质量的重要手段之一[1]。良好的产品风格设计可以有效传达消费者的情感需求，也能充分体现设计师的理念，不仅是设计师与消费者沟通的重要方式，而且是设计师准确把握和理解消费者对产品情感认知的重要途经[2,3]。从广义角度来看，产品风格是地域、文

收稿日期：2022-07-11；修订日期：2023-04-03。Received 11 July 2022; accepted 03 Apr. 2023.
基金项目：国家自然科学基金资助项目(52165033)。Foundation item：Project supported by the National Natural Science Foundation，China(No. 52165033)。

化、经济、生产技术等环境因素综合折射的结果。但从狭义的角度看，产品风格是通过不同的造型方法将各种造型元素分析、组织后，构建出的一种具有相似造型特征的集合[2,4]。其内涵涉及造型特征和意象特征[2]，是物理层面与精神层面高度融合的结果。因此，如何快速、准确、有效地计算产品风格及生成方法构建是研究者和实践者们长期探索的方向[2,4,5]，风格识别作为产品风格设计计算的首要任务，在产品形态概念生成、设计评价和用户偏好推荐等方面扮演着重要角色[2,4-7]。

针对产品风格计算，文献[2]总结的四类方法相对全面合理，包括基于形状文法的产品风格描述与再现、基于感性工学的产品风格与造型要素映射、基于认知心理学的产品风格认知计算和基于模式识别理论的产品风格计算模型。其中第四类方法属于人工智能范畴，然而受限于当时的算力、算法和数据，研究并没有获得较好的效果，仅作为一种风格查询系统在使用[2]。近年来，深度学习作为人工智能的一个重要分支发展迅猛，特别是深度卷积神经网络(Deep convolutional neural networks，DCNNs)的发展使得计算机视觉(Computer vision，CV)在图像识别、目标监测、语义分割等方面取得了重要突破[8]。2012 年 AlexNet[9]横空出世，在 ImageNet(2010)上的图像特征学习能力首次超过人工设计的特征，从而改变了人们对 CV 的理解方式。AlexNet 不仅良好继承了它的前贤 LeNet-5(1998)[10]，同时对后续 DCNNs 研究影响深远。例如，NiN(2013)[11]，VGG(2015)[12]、GoogLeNet(2015)[13]和 ResNets(2015)[14]都不同程度的参考和学习 AlexNet。其中 ResNets 提出了具有残差块的网络结构，可将网络深度提升至千层，为深度神经网络的收敛做出了重要贡献。另外，研究者们提出的各种激活函数(如 ReLU[15]，GELUs[16])、暂退法(Dropout)[9,17]、层归一化(Layer normalization，LN)[18]和批归一化(Batch normalization，BN)[19]等方法与技术是避免深度神经网络过拟合或梯度消失的重要手段。这些神经网络结构组件的提出不仅促进了深度学习技术发展，也为产品风格智能识别与生成提供了技术基础和研究思路。

近几年，一些研究者开始应用 DCNNs 提取目标产品的特征信息。HU 等[20]针对家具风格视觉分类任务，通过实验对比了基于人工特征设计的支持向量机(Support vector machine，SVM)和具有端到端学习能力的 DCNNs(AlexNet、VGG16、GoogLeNet 等)，结果显示 DCNNs 具有明显优势，但人工设计的特征也不容忽视。朱斌等[21]应用 VGG16 模型对座椅进行感性偏好识别，对比实验结果表明 VGG16 感性识别准确率超过经典机器学习算法 SVM。GONG 等[22]基于 AlexNet 对产品包装进行像素级的感性分析。ZHOU 等[23]应用简化后的 VGG11 对汽车进行二分类美学评估。王亚辉等[24]结合感性评价提出基于 ResNets 的人工智能设计决策模型，以起重机造型语义分类为实例进行验证。SU 等[25]则应用 DFL-CNN 对汽车进行细粒度感性偏好分类识别。可见 DCNNs 在产品风格、感性意象识别、美学评价方面的研究已取得一些成果。然而，现有研究多侧重于验证深度学习方法在风格意象识别、美学评价等方面的可行性，所使用的算法模型多以通用经典深度学习算法为主，鲜有针对产品风格特征精细识别问题提出新的神经网络算法框架。尽管文献[20]发现将人工设计特征和神经网络自动提取特征相结合能提高准确率，但人工设计特征不仅耗时费力，而且不利于扩展风格对象。因

此，针对产品风格特征提取及识别的神经网络算法仍需进一步研究。

注意力机制(Attention mechanism，AM)使得神经网络具备专注于核心特征的能力，最初在自然语言处理中被证实具有突出的效果，现已经被广泛用于不同的 CV 任务，如图像分类、语义分割、目标检测[26]，可以有效降低网络结构和模型复杂度。在 DCNNs 方面多以视觉注意力为主，常采用通道注意力和空间注意力。迁移学习(Transfer learning，TL)是机器学习中解决训练数据不足问题的重要方法[27]。它试图通过放宽小样本数据集必需为独立同分布的假设，将知识和经验从源域迁移到目标域。迁移学习在 CV 领域发挥了重要作用，通过在大数据集上预先训练获得一套模型参数，针对新的任务模型参数不再随机初始化，从而实现网络模型保留了在大数据集上获得的经验和知识。

产品风格精细识别是图像识别的一种特殊形式，属于 CV 应用研究子领域。相比粗粒度大类识别，产品风格精细识别属于细粒度子类识别问题，要求神经网络提取更多具有差异性的细节特征。另外，不同风格产品图像的数量是有限的，可归为小样本学习范畴[28]。因此，产品风格精细识别是一项更有挑战性的图像识别任务。注意力使得神经网络具有自上而下的特征选择而忽略无关特征的机制；预先训练模型作为迁移学习的重要范式为新的图像识别任务提供“通用视觉特征”，如“线条”、“轮廓”等。

基于上述分析，为更专注、更高效地抽取具有差异性的风格特征，提出一种基于复合学习机制(通路)的细粒度风格识别卷积神经网络，同时融入注意力机制和迁移学习模式，以提升产品风格识别准确率。该算法模型有助于设计师在风格特征设计阶段更有效地检索预期风格产品，实现对已有风格特征的检索和重用，也为识别用户风格偏好提供新的技术支持。

2　产品风格精细识别算法模型

2.1　风格识别网络

产品风格精细识别任务属于类内间的细粒度识别问题。DCNNs 在图像特征学习方面的能力出众，特别是在数据量充足的条件下能够实现接近或超过人类分类识别水平[8]。然而大多数产品历史图像数据较为有限，具体到不同风格的样本数据量则更小。论文创建的自行车头盔风格数据集，平均每种风格约有 1200 幅图像。使用较浅的网络模型难学习到更多有效特征，使用较深的网络模型又易产生过拟合，二者均容易导致模型性能下降[8,28]。

研究表明人脑神经系统有两个相对重要的学习机制，即迁移学习和注意力机制，前者能将从前积累的经验用于解决新问题，后者可有效处理信息过载问题。为融合、模拟这两种学习机制，论文提出一种复合并行学习的神经网络算法框架，包括注意力学习通路和迁移学习通路，用于提高复杂产品造型风格的识别精度，如图 1 所示。注意力学习通路实现“专用特征”学习，迁移学习通路基于大数据预先训练模型实现“通用特征”学习。同时给出一种具体的网络结构实现形式，并命名为细粒度风格识别卷积神经网络(Fine-grained style recognition convolutional neural networks，FSR-CNN)，如图 2 所示。一方面注意力学习通路由一种混合注意力残差网络(Hybrid attention-based ResNet，HA-ResNet)结构实现，是在修改后的残差网络结构中嵌入多种注意力，优点是注意力强且结构简洁，

但模型容量较小；另一方面迁移学习通路采用预先训练的 GoogLeNet 实现，优点是能通过小量数据学习引入外部经验记忆特征，以扩充 HA-ResNet 的容量。两条并行的学习通路优势互补，“专用特征”与“通用特征”的数量比为 2:1。最终的风格识别输出通过特征融合层和多层感知机(Multilayer perceptron，MLP)分类器实现。

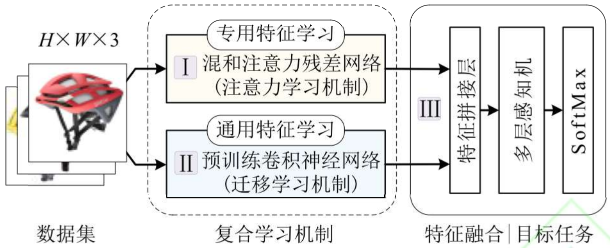


图 1　复合学习机制的神经网络算法框架

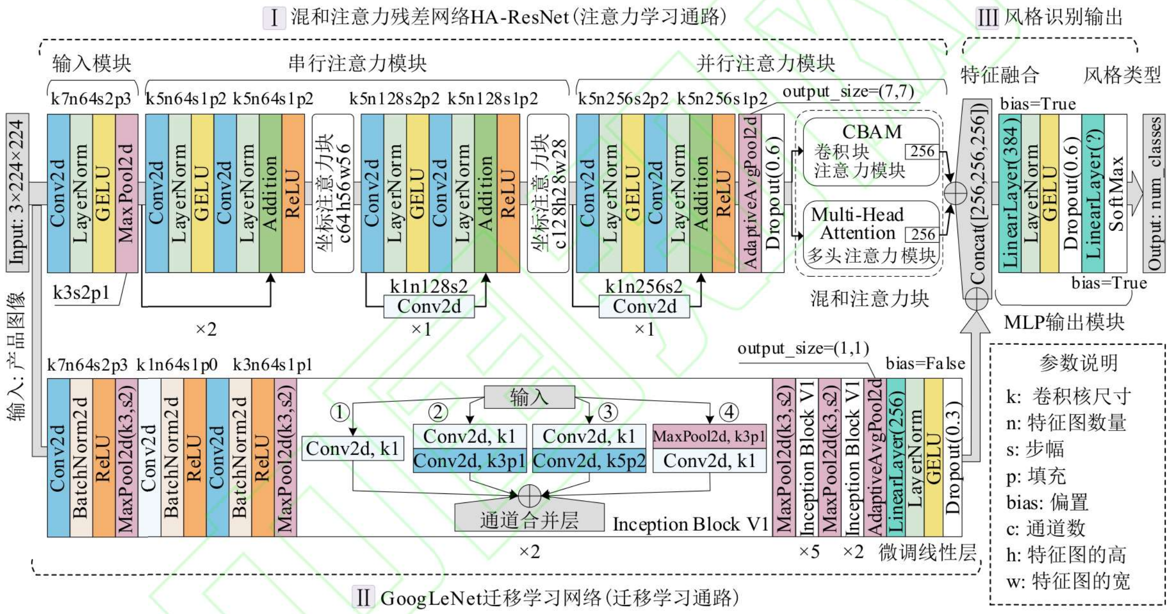


图 2　细粒度风格识别卷积神经网络结构

2.2　混合注意力残差网络

在卷积神经网络中，较浅层用于捕获低阶语义特征，较深层可学习到更高阶的语义特征，因此卷积神经网络一直向着更深的网络结构发展。然而深层网络容易出现退化现象，模型训练难度大，容易出现梯度消失和梯度爆炸。为此，HE 等[14]提出了残差学习框架，即通过快速跳跃连接的方式将信号前馈补偿进行残差表示学习解决了上述问题。这使得残差结构既可以增加网络层数，又能够防止网络退化现象。另外，注意力机制作为一种资源分配方案，能将有限的计算资源用来处理更关键的信息，可有效提高神经网络的计算效率[26,29]。基于此，论文提出混合注意力残差网络 HA-ResNet，如图 2 所示。以残差结构作为网络基础架构，采用串并结合的方式将三种不同的注意力分阶段嵌入其中，有助于注意力学习通路捕获产品风格关键特征信息。

从图 2 可知，HA-ResNet 输入的是一幅产品图像，输出是该产品造型风格注意力特征张量，共

有 3 个模块组成，分别是图像输入模块、串行注意力模块和并行注意力模块。输入模块采用 1 个卷积层对输入图像进行处理，卷积核的尺寸为 7×7，步长为 2，将输入的 3 通道特征图映射后输出 64 通道的特征图。使用 LN 代替 BN 对卷积层的输出进行层归一化，将高斯线性误差单元 GELUs 作为激活函数，并用最大汇聚层(MaxPool2d)对激活后的特征进行降维。接着是串行注意力模块，由卷积残差块和坐标注意力块顺序交替组成，将特征图从 64 维映射到 128 维。最后是并行注意力模块，包括 1 个卷积残差块和 1 个混合注意力块。先由卷积残差块将特征图从 128 维提升至 256 维；然后利用自适应平均汇聚(AdaptiveAvgPool2d)提取特征并输出尺寸为 7×7 的特征图，为防止过拟合在其后加入 Dropout 层(概率为 0.6)；最后分别由卷积块注意力模块(Convolutional block attention module，CBAM)和多头注意力(Multi-head attention，MHA)模块并行关注而成。另外，所有卷积残差块中卷积核的尺寸为 5×5，且激活函数混合使用 GELUs 和 ReLU。

2.2.1 串行注意力模块

深度神经网络靠前的浅层可识别物体的边、角和轮廓，靠后的深层可识别整体对象部分[30]，这是一种从局部到整体的识别模式。然而这与 WERTHEIMER[31]最先提出的格式塔视觉心理感知组织原则存在差异性。该视觉组织原则包括接近原则(相近的元素倾向于被组织成单元)、相似原则(看上去相像的物体常常被组合为一体)、连续性原则(除非有尖锐的拐点出现，不然视觉知觉认为是连续的)和闭合原则(倾向于完整地构建一个图形，而不是观察残缺的线条或形状)。通常人们对产品风格的视觉心理感知同样遵循某一条或某几条原则。这说明人的视觉感知具有先整体后局部的先验性特点，为人类快速识别对象提供了有效支持。为了模拟这种视觉感知特点，论文在卷积神经网络较浅层嵌入具有全局注意力机制的坐标注意力块(Coordinate attention block，CAB)[32]。如此可使较浅的卷积层提前关注到全局信息，有助于降低网络的深度。由图 2 可知，串行注意力模块是由残差块和 CAB 串联而成，其中残差块已在上文给出详述，在此重点讨论 CAB 的算法结构。

CAB 机制的实现如图 3 所示。首先，分别从水平方向和垂直方向进行平均汇聚操作得到两个特征向量；其次，在空间维度上先进行特征拼接操作(Concat operation)，后进行 1×1 卷积运算，从而实现压缩通道数；然后，通过 BN 处理和非线性变换来编码垂直方向和水平方向的空间信息，并对其进行分割操作(Split operation)；接着，再各自通过 1×1 卷积运算获得与输入通道数相同的特征图，并使用 Sigmoid 激活函数对特征数据进行归一化；最后，实现加权输出与输入相同维度的特征图。总结来说，CAB 机制首先是在水平方向和垂直方向上同时进行平均汇聚，然后通过一系列变换方法对空间信息进行编码，最后把空间信息在通道维度上以加权求和的方式进行融合，从而实现更大区域特征信息的关注。具体网络设计是在第 2 个卷积残差块后使用了 1 个通道数为 64、高宽均为 56 的 CAB，以及在第 3 个卷积残差块之后使用了 1 个通道数为 128、高宽均为 28 的 CAB，从而构建了一个残差坐标串行注意力模块，如图 2 所示。

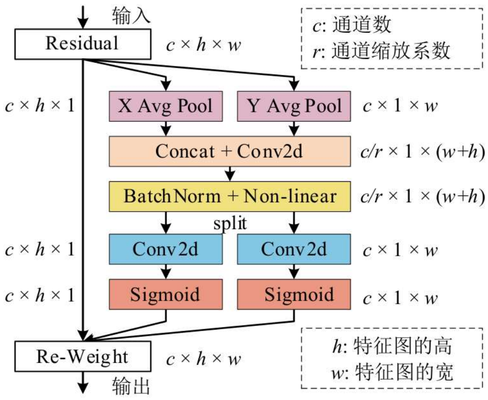


图 3　坐标注意力模块[32]

2.2.2　并行注意力模块

由图 2 可知，并行注意力模块由 1 个卷积残差块和两个并行通路的注意力块组成，前者执行产品造型特征的高层语义表示，后者实现对重要信息的提取。由于 HA-ResNet 使用较少的卷积层进行特征提取，但为了从网络深层获得更多关键特征表示，论文提出混合两种不同的注意力机制进行并行特征关注，并与残差块顺序连接构成并行注意力模块，如图 4 所示。其中，CBAM 是一种静态注意力计算模式；而 MHA 则是以动态生成注意力权重的方式捕捉交互信息，并且能以多头并行运算方式过滤信息。因此，以动静并行的注意力计算模式有利于关注到深层次的风格特征。

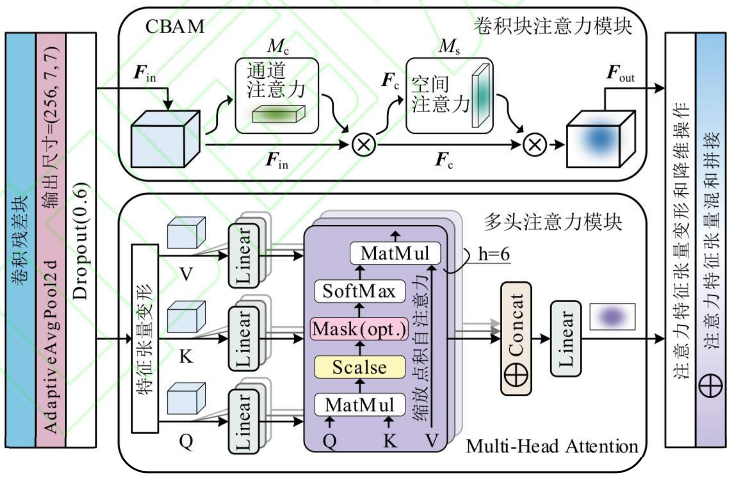


图 4　并行注意力模块

CBAM[33]是一种用于前馈卷积神经网络的注意力模块，由通道注意力和空间注意力混合而成，如图 4(上)所示。可对给定任何中间特征图在通道和空间依次推断出注意力图，然后将注意力图与给定的中间特征图相乘以进行自适应特征精炼，从而实现对关键特征提取。其连续的注意力计算过程为：

$$
\begin{aligned}
\boldsymbol{F}_{\mathrm{c}} &= \boldsymbol{F}_{\mathrm{in}} \otimes M_{\mathrm{c}}\left(\boldsymbol{F}_{\mathrm{in}}\right), \\
\boldsymbol{F}_{\mathrm{out}} &= \boldsymbol{F}_{\mathrm{c}} \otimes M_{\mathrm{s}}\left(\boldsymbol{F}_{\mathrm{c}}\right)
\end{aligned}
\tag{1}
$$

式中：⊗表示按元素相乘，$\boldsymbol{F}_{\mathrm{in}} \in \mathbb{R}^{256\times7\times7}$ 是前一层 AdaptiveAvgPool2d 输出的特征图，$M_{\mathrm{c}} \in \mathbb{R}^{256\times1\times1}$ 表示 1 维通道注意力映射，$M_{\mathrm{s}} \in \mathbb{R}^{1\times7\times7}$ 表示 2 维空间注意力映射。

MHA[34]是通过将 h 个自注意力头以并行独立学习的方式表示不同的关注行为，能让每个头都关注输入的不同部分，可以表示比简单加权平均值更复杂的函数，如图 4(下)所示。具体来说，当给定相同的查询($\boldsymbol{q} \in \mathbb{R}^{d_{\mathrm{q}}}$)、键($\boldsymbol{k} \in \mathbb{R}^{d_{\mathrm{k}}}$)和值($\boldsymbol{v} \in \mathbb{R}^{d_{\mathrm{v}}}$)信息集合时，可以用独立学习得到的 h 个不同的线性映射来变换它们，并对其进行注意力汇聚和拼接，注意力头 $\boldsymbol{h}_i\left(i=1,\ldots,h\right)$ 计算方法为：

$$\begin{aligned}\boldsymbol{h}_i &= f_{\text{self-att}}\left(\boldsymbol{W}_i^{(\mathrm{q})}\boldsymbol{q},\boldsymbol{W}_i^{(\mathrm{k})}\boldsymbol{k},\boldsymbol{W}_i^{(\mathrm{v})}\boldsymbol{v}\right)\\ &= f_{\text{self-att}}\left(\boldsymbol{Q}_i,\boldsymbol{K}_i,\boldsymbol{V}_i\right) \in \mathbb{R}^{p_{\mathrm{v}}}\end{aligned} \tag{2}$$

式中：可学习的参数为 $\boldsymbol{W}_i^{(\mathrm{q})} \in \mathbb{R}^{p_{\mathrm{q}}\times d_{\mathrm{q}}}$、$\boldsymbol{W}_i^{(\mathrm{k})} \in \mathbb{R}^{p_{\mathrm{k}}\times d_{\mathrm{k}}}$ 和 $\boldsymbol{W}_i^{(\mathrm{v})} \in \mathbb{R}^{p_{\mathrm{v}}\times d_{\mathrm{v}}}$，且 $\boldsymbol{Q}_i=\boldsymbol{W}_i^{(\mathrm{q})}\boldsymbol{q}$、$\boldsymbol{K}_i=\boldsymbol{W}_i^{(\mathrm{k})}\boldsymbol{k}$、$\boldsymbol{V}_i=\boldsymbol{W}_i^{(\mathrm{v})}\boldsymbol{v}$；$f_{\text{self-att}}(\cdot)$ 表示注意力汇聚函数，这里为缩放点积自注意力，即 $\boldsymbol{V}_i\,\mathrm{SoftMax}\left(\boldsymbol{K}_i^{\mathrm{T}}\boldsymbol{Q}_i\big/\sqrt{d_{\mathrm{k}}}\right)$。

最后的输出需经过一个线性变换得到：

$$\boldsymbol{W}_{\mathrm{o}}\begin{bmatrix}\boldsymbol{h}_1\\ \vdots\\ \boldsymbol{h}_h\end{bmatrix} \in \mathbb{R}^{p_{\mathrm{o}}} \tag{3}$$

式中：$\boldsymbol{W}_{\mathrm{o}} \in \mathbb{R}^{p_{\mathrm{o}}\times hp_{\mathrm{v}}}$ 为可学习的参数。

综合考虑 CBAM 和 MHA，混合注意力特征计算方法为：

$$\begin{aligned}\boldsymbol{F}_{\mathrm{CBA}} &= f_{\mathrm{RD}}\left(\boldsymbol{F}_{\mathrm{out}}\right),\\ \boldsymbol{F}_{\mathrm{MHA}} &= f_{\mathrm{RD}}\left(\boldsymbol{h}_{i\in\{1,2,3,\cdots,6\}}\right),\\ \boldsymbol{F}_{\mathrm{CON}} &= f_{\mathrm{CON}}\left(\boldsymbol{F}_{\mathrm{CBA}},\boldsymbol{F}_{\mathrm{MHA}}\right)\end{aligned} \tag{4}$$

式中：f_{RD} 表示特征张量变形及均值降维，$\boldsymbol{F}_{\mathrm{CBA}} \in \mathbb{R}^{256}$ 表示经 f_{RD} 映射后的卷积块注意力特征；$\boldsymbol{h}_{i\in\{1,2,3,\cdots,6\}}$ 表示使用了 6 个自注意力头，$\boldsymbol{F}_{\mathrm{MHA}} \in \mathbb{R}^{256}$ 表示经 f_{RD} 映射后的多头自注意力特征；f_{CON} 表示特征拼接操作，$\boldsymbol{F}_{\mathrm{CON}} \in \mathbb{R}^{512}$ 表示经 f_{CON} 映射后的混合注意力特征。

2.3 基于 GoogLeNet 的迁移学习网络

通常，提升网络性能最直接的办法就是增加网络深度，但一味地增加会导致网络参数激增、模型变大、容易过拟合、梯度消失、难以收敛等问题。为了解决这些问题，GoogLeNet 研究人员基于赫布原理(Hebbian principle)和多尺度处理方法，提出具有并行卷积计算的 Inception block 算法[13]，如图 2 所示。Inception block 由 4 条并行的卷积运算路径构成。前 3 条路径分别使用卷积核为 1×1、3×3 和 5×5 的卷积层抽取不同尺度的空间特征。第 2、3 条路径为减少通道数，均使用了 1×1 卷积层对输入进行处理，从而降低了模型的复杂度。第 4 条路径则首先使用核尺寸为 3×3 的最大汇聚层，然后应用 1×1 卷积层来改变通道数。为了使输入和输出的尺寸一致，4 条路径的卷积层均使用了相适应的填充。最后将每条路径上的输出特征在通道维度上进行拼接操作。该算法重点解决了适度卷

积核尺寸问题。通过组合不同大小卷积核的方式抽取不同尺度的细节特征，从网络宽度的角度提升模型性能，最终的 GoogLeNet 共串联了 9 个 Inception block。相比 LeNet-5[10]、AlexNet[9]、NiN[11]、VGG[12]等串行卷积运算方式，GoogLeNet 的卷积层并行运算方式更有利于细粒度产品风格特征抽取。因此，论文采用在 ImageNet 上预先训练的 GoogLeNet 作为迁移学习通路，以便抽取多尺度"通用特征"，有助于扩充 HA-ResNet 容量，同时也扩展了 FSR-CNN 的宽度。对模型最后一个线性层进行微调，先将其输出特征修改为 256，然后进行层归一化处理，并将 GELUs 作为激活函数层，最后使用 Dropout 层预防过拟合，概率为 0.3，如图 2 所示。

2.4 损失函数

论文将产品造型风格精细识别任务表示为预测离散标签的多分类建模问题。将输入到 FSR-CNN 的产品图像表示为 $\boldsymbol{x}_n$，且基于该网络构造一个真实条件概率分布 $f_{\text{FSR-CNN}}(y \mid \boldsymbol{x}_n)$，从而预测产品造型风格类型 $f_{\text{FSR-CNN}}(\hat{y} \mid \boldsymbol{x}_n)$，给定产品风格数据集 $SD=\{(\boldsymbol{x}_n, y_n)\}_{n=1}^{N}$，其中，$y_n \in \{1,2,3,\cdots,C\}$ 为风格类型标签，N 为训练集样本数量。通过交叉熵损失函数来量化 FSR-CNN 的误差：

$$Loss_{\text{FSR-CNN}}\left(y_{n,c}, f_{\text{FSR-CNN}}\left(\hat{y}_{n,c} \mid \boldsymbol{x}_n; \boldsymbol{w}\right)\right) = -\frac{1}{2N}\sum_{n=1}^{N}\sum_{c=1}^{C} y_{n,c}\log\hat{y}_{n,c} + \lambda\|\boldsymbol{w}\|_2^2 \tag{5}$$

式中：$y_{n,c}$ 是产品的真实风格类型；$\hat{y}_{n,c}$ 是 FSR-CNN 预测的风格概率，由 FSR-CNN 最后的 SoftMax 层输出；$\lambda\|\boldsymbol{w}\|_2^2$ 表示权重衰减正则化，λ 是权重衰减系数。

2.5 评价指标

为了验证比较模型的风格识别性能，论文采用与文献[21,23,25]相同的准确率(Accuracy)、精确率(Precision)、召回率(Recall)和 F_1 值(F_1Score)对训练好的模型进行评估比较，如公式(6)~(9)所示：

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{6}$$

$$Precision = \frac{TP}{TP+FP} \tag{7}$$

$$Recall = \frac{TP}{TP+FN} \tag{8}$$

$$F_1Score = 2\times\frac{Precision\times Recall}{Precision+Recall} \tag{9}$$

式中：TP 是指实际为正例，预测为正例的样本数量；TN 是指实际为负例，预测为负例的样本的数量；FP 代表实际为负例，但被预测为正例的样本数量；FN 代表实际为正例，但被预测为负例的样本的数量。

3 实验

实验平台为双 GPU(RTX2080/8GB 显存)，并使用 CUDA11.0 加速；操作系统为 Ubuntu18.04，深度学习框架为 PyTorch1.8.1；训练过程由 Visdom 可视化监测。以自行车头盔数据集对 FSR-CNN

及对比算法模型进行训练。

3.1 数据集

由于目前鲜有开源细粒度产品风格数据集，因此论文创建了 1 个自行车头盔风格数据集，共包含 6502 幅自行车头盔图像，训练/验证集为 6217 幅，测试集为 285 幅。一共有 5 类风格的自行车头盔图像，即波线型(线条交错)、科幻型(造型独特)、流线型(多线条并行)、硬朗型(型面交错规整)、包裹型(形态圆润)，每种风格的样例和数量见表 1。

表 1 自行车头盔风格数据集

风格类型	波线型	科幻型	流线型	硬朗型	包裹型
产品样例					
训练/验证集	1337	878	1047	1455	1500
测试集	58	52	45	62	68
标签	1Sleek	2Sci-fi	2Streamline	4Hale	5Wrap

3.2 网络验证实验

3.2.1 网络对比实验

FSR-CNN 最大训练轮数 max_epoch 设为 800，批量大小 batch_size 为 88。使用 AadmW[35]优化器优化网络参数，初始学习率为 2E-6，$\beta_1 = 0.9$，$\beta_2 = 0.999$，同时使用权重衰减策略，衰减系数 λ 为 5E-2。训练时将自行车头盔风格数据集划分为训练集与验证集，比例为 6:4。

图 5 为 FSR-CNN 训练损失与验证损失变化对比，误差均逐轮降低并趋于稳定，其中验证损失曲线下降地更快，表明网络具有较好地学习和抗过拟合能力。图 6 为训练准确率与验证准确率曲线对比，二者均逐轮上升，其中验证准曲率增长相对较快，并最终趋于稳定，表明具有较好地泛化性能。最好模型出现在第 725 轮，其验证准确率为 87.79%。

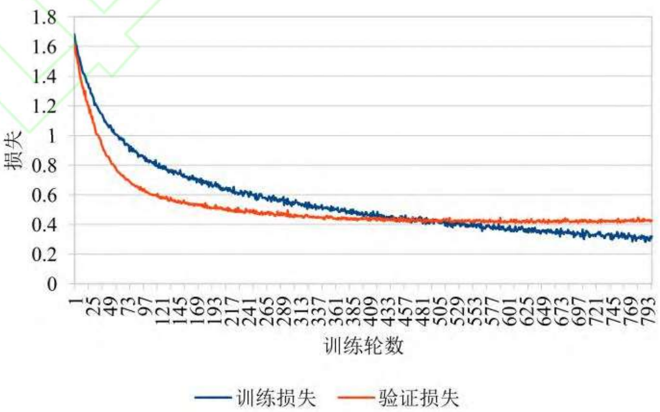


图 5 训练损失与验证损失对比

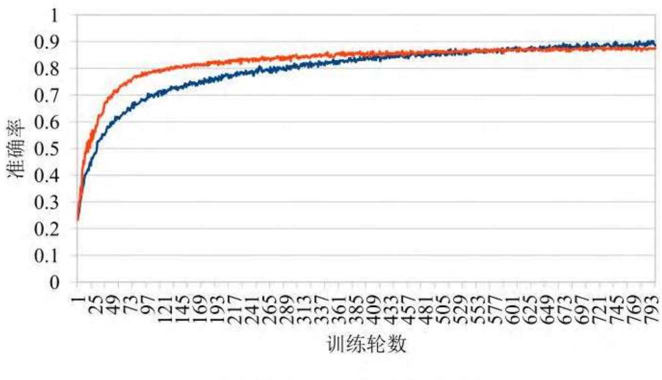


图 6 训练准确率与验证准确率对比

为对比验证，还训练了 7 个经典 DCNNs 算法模型，即 LeNet-5[10]、AlexNet[9]、VGG11[12]、VGG16[12]、GoogLeNet[13]、GoogLeNet(pretrained)[13]和 ResNet18[14]，网络深度依次增加。需要说明的是不同的超参数会对训练结果产生不同程度的影响，本着较小过拟合的原则经过多次实验，7 个对比模型选择了训练曲线与验证曲线振荡较小的超参数，具体见表 2。训练模型时，验证集的损失变化曲线与准确率变化曲线如图 7、图 8 所示。对比可知 FSR-CNN 的验证误差下降明显，小至 0.4 附近且振荡小；同时验证准确率曲线逐轮提升，且振荡最小。这表明 FSR-CNN 算法模型具有更好的泛化能力和稳健性。

表 2 经典算法模型的超参数

序号	算法模型	学习率	权重衰减系数	批量大小	优化器
1	LeNet-5	1E-4	1E-3	88	AdamW
2	AlexNet	3E-6	5E-2	88	SGD
3	VGG11/VGG16	1E-5	1E-3	88	Adam
4	GoogLeNet	5E-5	1E-3	88	AdamW
5	GoogLeNet(pretrained)	1E-6	1E-1	88	AdamW
6	ReNet18	6E-6	5E-2	88	AdamW

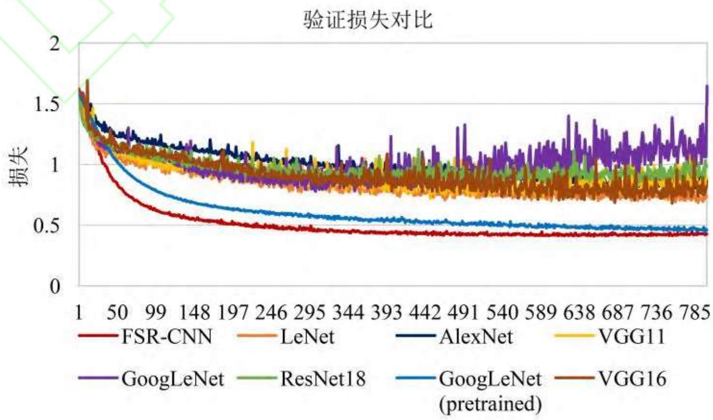


图 7 8 个算法模型的验证损失变化曲线对比

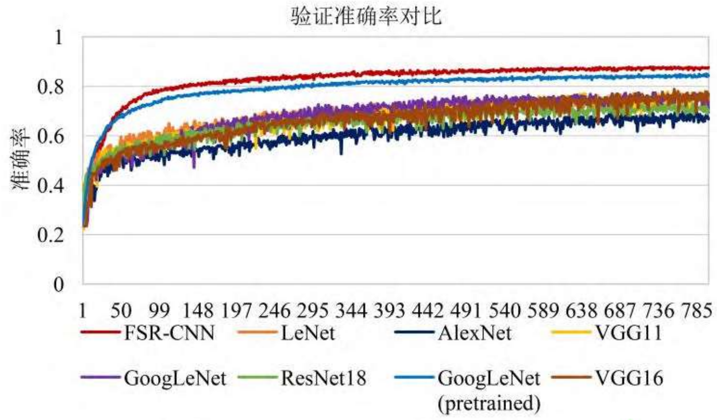


图 8 8 个算法模型的验证准确率变化曲线对比

在测试集上对训练好的 FSR-CNN 及上述 7 个算法模型进行风格识别性能测试，结果见表 3。可知，FSR-CNN 在准确率、精确率、召回率和 F1 值 4 项评价指标均高于其他 7 个 DCNNs 算法模型，说明该模型与自行车头盔风格数据集的复杂度最匹配。同时，还可看出 GoogLeNet 在 6 个非预训练算法模型中的表现相对更好，这也验证了选择 GoogLeNet 作为迁移学习通路具有一定的合理性。

表 3 FSR-CNN 与经典算法模型对比结果

序号	算法模型	准确率	精确率	召回率	F_1 值	是否预训练
0	**FSR-CNN**	**0.835**	**0.845**	**0.835**	**0.836**	**混和**
1	LeNet-5	0.670	0.685	0.670	0.669	否
2	AlexNet	0.677	0.677	0.677	0.668	否
3	VGG11	0.695	0.697	0.695	0.680	否
4	VGG16	0.737	0.754	0.737	0.734	否
5	GoogLeNet	0.712	0.713	0.712	0.696	否
6	GoogLeNet(pretrained)	0.821	0.826	0.821	0.822	是
7	ReNet18	0.688	0.699	0.689	0.678	否

另外，通常由于数据集中的每种风格类型的样本数量是不相等的，表 3 中的评价结果并不能直接反应出模型在每个风格类型中的识别性能。因此，必须使用更有效的指标来衡量 FSR-CNN 的性能。混淆矩阵(Confusion matrix)是评价分类识别模型优劣的更直观的工具。图 9 为 FSR-CNN 模型在测试集上的混淆矩阵，矩阵的每一列代表模型识别的结果，矩阵的每一行表示样本的真实风格标签。

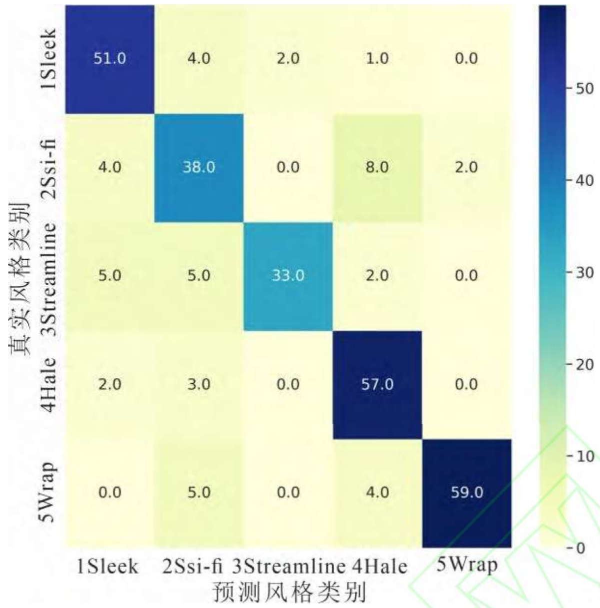


图 9　FSR-CNN 在测试集上的混淆矩阵

由图 9 可获得识别错误的风格类型和样本数量，波线型(1Sleek)、流线型(3Streamline)、硬朗型(4Hale)和包裹型(5Wrap)各有 4 个、5 个、3 个和 5 个的自行车头盔被识别为科幻型风格(2Sci-fi)，同时科幻型风格有 4 个被识别为波线型，8 个被识别为硬朗型，2 个被识别为包裹型。波线型中有 4 个被识别为科幻型、2 个被识别为流线型、1 个被识别为硬朗型。另外，前四种风格中除了科幻型有 2 个被识别为包裹型外，其余风格均识别正确，而包裹型风格中有 5 个被识别为科幻型，4 个被识别为硬朗型。由此表明，模型识别科幻型风格的准确率最低，而识别包裹型风格的准确率最高。进一步通过公式(7)~(9)可得到每种风格类型的具体精确率、召回率和 F_1 值，结果见表 4。比较 F_1 值可知模型在 5 种风格类型上的稳健性由高到低依次为包裹型、硬朗型、波线型、流线型和科幻型。部分风格样例识别结果，如图 10 所示。

表 4　FSR-CNN 风格类型评价结果

风格类型	标签	精确率	召回率	F_1 值	测试总量
波线型	1Sleek	0.82	0.88	0.85	58
科幻型	2Sci-fi	0.69	0.73	0.71	52
流线型	3Streamline	0.94	0.73	0.83	45
硬朗型	4Hale	0.79	0.92	0.85	62
包裹型	5Wrap	0.97	0.87	0.91	68

图 10　FSR-CNN 的风格预测结果(部分)

3.2.2　网络消融实验

对于较为复杂的神经网络，消融实验常用于分析各神经网络算法模块对整个网络的贡献性，可采用删除网络中的部分算法模块以验证其对网络整体性能的影响。为了验证 2 个学习通路对 FSR-CNN 的影响，以及不同的注意力机制对 FSR-CNN 的影响，论文以模块组别的方式设计了消融实验。其中，组别 0 是 FSR-CNN，组别 1 是 FSR-CNN 中的注意力学习通路 HA-ResNet，组别 2 是 FSR-CNN 中的迁移学习通路，组别 3 是删除了 FSR-CNN 中的坐标注意力机制(CAB)，组别 4 是删除了 FSR-CNN 中的混合注意力块(CBAM 和 MHA)。表 5 为各组别神经网络算法在测试数据集上的识别性能对比结果。可看出，FSR-CNN 在准确率、精确率、召回率和 F_1 值 4 个评价指标上的表现均最佳。因此，FSR-CNN 模型具有一定的合理性和先进性。

表 5　各组别神经网络消融实验对比

组别	模块组别	说明	准确率	精确率	召回率	F_1 值
0	**FSR-CNN**	**论文方法**	**0.835**	**0.845**	**0.835**	**0.836**
1	HA-ResNet	注意力学习通路	0.639	0.668	0.639	0.625
2	GoogLeNet	迁移学习通路	0.818	0.824	0.818	0.817
3	无串行注意力的 FSR-CNN	删除 FSR-CNN 中的 2 个 CAB	0.796	0.811	0.796	0.797
4	无并行注意力的 FSR-CNN	删除 FSR-CNN 中的 CBAM 和 MHA	0.814	0.823	0.814	0.814

4　讨论

面对 FSR-CNN 对自行车头盔风格的预测结果，可由图 10 中的一些结果作进一步分析和推断。在波线风格中，3 号方案被识别为科幻型，但识别概率(0.558)较低，这是由于该方案的线条交错相对规整。在科幻风格中，8 号方案被识别为包裹型，这是由于该方案在整体造型上与包裹型比较接近，

仅通过头盔后部翘起的渐消面体现科技感，容易导致识别错误；对比 9 号和 10 号方案发现，受视角影响 9 号方案被识别为包裹型。在流线型风格中，因风格特征不够强烈导致识别错误，如 12 号方案。在硬朗型风格中，因部分方案存在风格特征的交叉而容易出现识别偏差，如 17 号方案兼具波线型和硬朗型，但标签更偏向硬朗型风格。在包裹型风格中，同样出现因风格特征差异不明显导致的识别错误，如 22 号方案，其标签为包裹型，但模型识别为科幻型。由此可知产品风格分类与识别，不仅需要从总体上识别产品的造型结构，还需要考虑视角，以及判别具体形态细节，如线条、尺度、色彩等。

结合图 10 中的识别结果，考虑到神经网络算法缺乏人类视觉的空间构思能力，单一视角容易出现识别错误。为此，论文尝试了多视角综合识别任务，如图 11 所示。发现综合评估多个视角的风格特征概率能够提高预测准确率，同时也表明 FSR-CNN 具有较好的泛化能力。

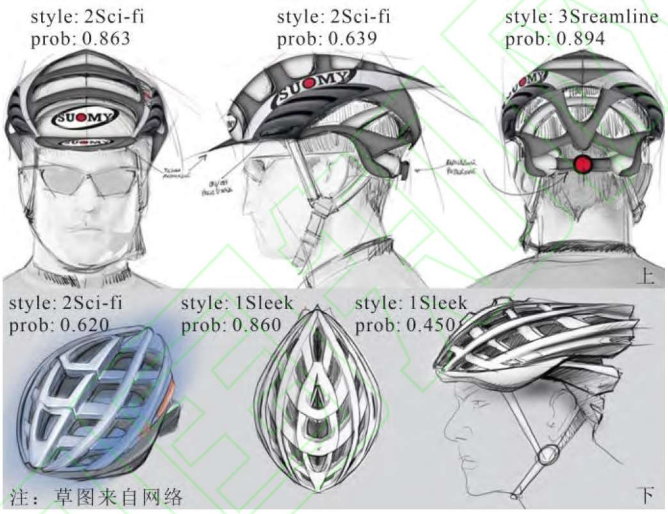


图 11　多视角综合识别结果

FSR-CNN 本质上输出的是设计方案的风格概率分布，因此能够对设计师绘制的概念草图进行风格概率分布预测，如图 12 所示。从该图中可以清晰地看出每种概念设计方案的造型特质和风格趋向，有助于高效引导设计师分析、探索和聚焦符合用户风格认知的设计概念。例如，图中 Sketch_6 以流线型(3Strealine)为主，兼具波线(1Sleek)和硬朗(4Hale)，且波线风格强于硬朗，而包裹和科幻风格特征则相对很小。与形状文法、感性工学等[2]传统风格计算方法相比，无论是在风格要素评价上，还是在产品风格继承性和竞争对手产品决策分析方面[1]，FSR-CNN 的风格计算识别方法为设计师提供了一种相对理性、高效的智能分析手段，有利于设计团队及时调整和把控产品风格策略。

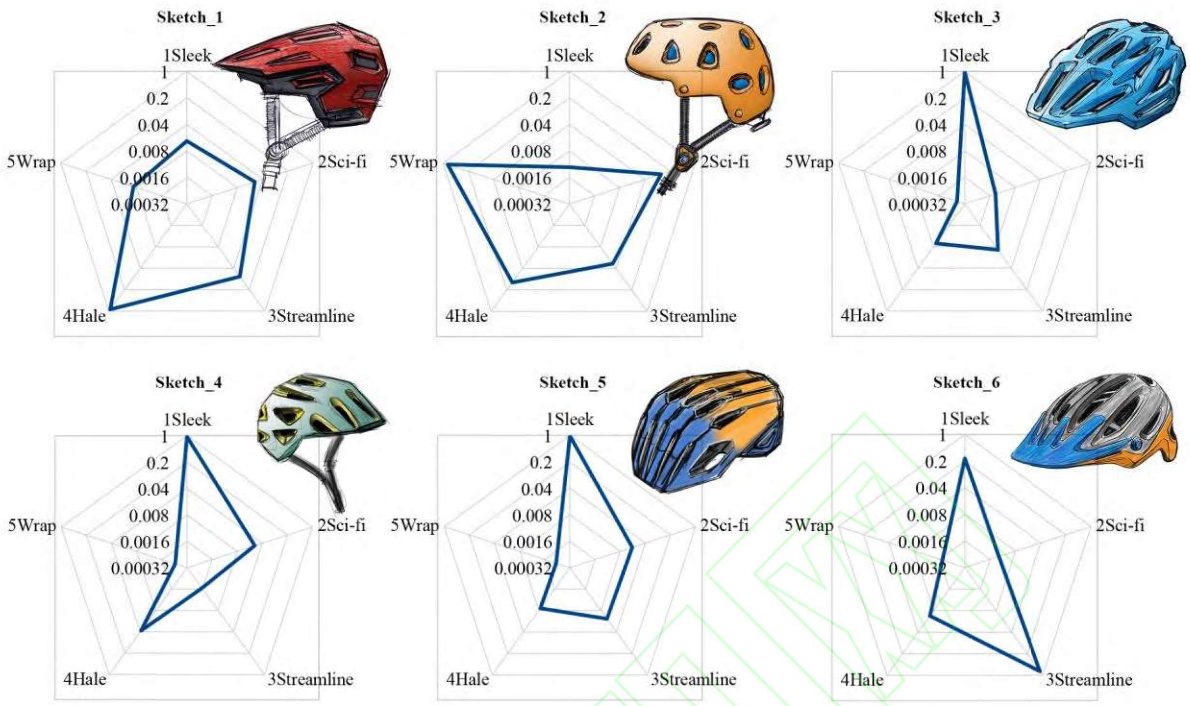


图 12　FSR-CNN 风格策略分析

FSR-CNN 为产品风格智能分类与识别提供了新方法、新思路，但该项工作仍有一个问题必须面对。产品风格传递的是一种综合体验，既有物理层面的造型特征，还有精神层面的意象特征。从上一节的实验过程及结果可知，FSR-CNN 更多的是在物理层面学习不同风格间的差异性特征，所提供的单一风格标签标是其进行学习的引导性准则，未融入多样化的风格情感属性。这也是细粒度风格识别出错的一个重要因素。因此，还需挖掘产品在线评论数据，进一步探索多模态多标签产品风格识别方法，以及多标签风格策略分析。

5 结束语

针对产品风格精细识别任务，论文提出一个细粒度风格识别深度卷积神经网络 FSR-CNN，以复合并行通路连接的方式融入了迁移学习和注意力机制。这两种学习机制的联合为产品风格精细识别提供了新思路，不仅有利于抽取更加细腻的特征，而且面对小数据集有抗过拟合特性。消融实验不仅验证了 FSR-CNN 在产品风格精细识别上的优良性能，也进一步地表明了复合学习机制的优势。

注意力学习通路使用论文提出的混合注意力残差网络 HA-ResNet 实现。该网络在残差映射结构中先以串行的方式嵌入了两个坐标注意力块，而后以并行的方式同时嵌入了卷积块注意力和多头注意力，不仅能够较早地关注自行车头盔风格特征的全局信息，还能重点关注空间位置信息和有效的特征差异信息，对提升模型的识别准确率起到了至关重要的作用。迁移学习通路采用预先训练的 GoogLeNet，其网络结构特点是多感受野并行计算，能提取更加细腻的风格特征细节。另外，面对样本数量有限的条件下，在网络中加入层归一化、自适应平局汇聚和暂退法，并混合使用 ReLU 和 GELUs 激活函数，有利于缓解过拟合现象，提升模型泛化能力。

通过实验与 7 个经典深度卷积神经网络对比，证明 FSR-CNN 能以较高的准确率和良好的稳健性对自行车头盔图像进行风格识别。与传统风格认知计算模式相比，FSR-CNN 无需手工提取特征，实现端到端的风格识别，且省时省力。不仅可以辅助设计师实施产品智能风格策略分析及设计定位，而且可为用户精准风格推荐提供支持，还为产品风格聚类奠定了基础。

论文所提出的具有复合学习机制的并行网络架构也可以应用到其他精细分类和识别任务，针对类内间差异大小、数据集规模等，设计出更有针对性的混合注意力网络通路，以及尝试更合适的迁移学习网络通路。

参考文献：

[1] PERSON O, SCHOORMANS J, SNELDERS D, et al. Should new products look similar or different? The influence of the market environment on strategic product styling [J]. Design Studies, 2008, 29(1): 30-48.

[2] HUANG Qi, SUN Shouqian. State-of-the-art of research on product style computation[J]. Journal of Computer-Aided Design & Computer Graphics, 2006, 18(11): 1629-1636 (in Chinese). [黄　琦, 孙守迁. 产品风格计算研究进展[J]. 计算机辅助设计与图形学学报, 2006, 18(11): 1629-1636.]

[3] HUANG Qi. Research on the technology of computer-aided conceptual design based on product style recognition model[D]. Hangzhou: Zhejiang University (in Chinese). [黄　琦. 基于产品风格认知模型的计算机辅助概念设计技术的研究[D]. 杭州: 浙江大学.]

[4] JING Nan, FANG Hai, ZHANG Qinwei. Research of product style innovative design based on typology[J]. Journal of Machine Design, 2015, 32(4): 121-125 (in Chinese). [景　楠, 方　海, 张秦玮. 基于类型学的产品风格创新设计研究[J]. 机械设计, 2015, 32(4): 121-125.]

[5] FU Yetao, LUO Shijian. Style perception-oriented product family shape gene design[J]. Computer Integrated Manufacturing Systems, 2012, 18(3): 449-457 (in Chinese). [傅业焘, 罗仕鉴. 面向风格意象的产品族外形基因设计[J]. 计算机集成制造系统, 2012, 18(3): 449-457.]

[6] WANG D, LI Z, DEY N, et al. Case-based reasoning for product style construction and fuzzy analytic hierarchy process evaluation modeling using consumers linguistic variables[J]. IEEE Access, 2017, 5: 4900-4912.

[7] XU Jiang, WANG Haixian, SUN Shouqian. Method of product generative design based on style evolution model[J]. Journal of Southeast University (Natural Science Edition), 2012, 42(4): 654-658 (in Chinese). [徐　江, 王海贤, 孙守迁. 基于风格进化模型的产品生成设计方法[J]. 东南大学学报(自然科学版), 2012, 42(4): 654-658.]

[8] HADJI I, WILDES R P. What do we understand about convolutional networks? [EB/OL]. (2018-3-23). [2022-4-9]. https://arxiv.org/ pdf/1803.08834.pdf.

[9] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[J]. Commun ACM, 2017, 60(6): 84-90.

[10] LECUN Y, BOTTOU L. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.

[11] LIN M, CHEN Q, YAN S. Network in network [EB/OL]. (2014-03-04) [2022-5-10]. https://arxiv.org/pdf/1312.4400v3.pdf.

[12] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [EB/OL]. (2015-04-10) [2022-5-10]. https://arxiv.org/pdf/1409.1556.pdf.

[13] SZEGEDY C, WEI L, YANGQING J, et al. Going deeper with convolutions[C] //proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). New York: IEEE, 2015: 1-9.
[14] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [EB/OL]. (2015-12-10) [2022-4-10]. https://arxiv.org/pdf/1512.03385.pdf.
[15] NAIR V, HINTON G E. Rectified linear units improve restricted boltzmann machines[C] //Proceedings of the 27th International Conference on International Conference on Machine Learning. Madison: Omnipress. 2010: 807–814.
[16] HENDRYCKS D, GIMPEL K. Gaussian error linear units (GELUs) [EB/OL]. (2020-07-08) [2022-4-10]. https://arxiv.org/pdf/1606.08415.pdf.
[17] SRIVASTAVA N, HINTON G, KRIZHEVSKY A, et al. Dropout: A simple way to prevent neural networks from overfitting[J]. 2014, 15(1): 1929-1958.
[18] BA L J, KIROS J R, HINTON G E. Layer normalization [EB/OL]. (2016-07-21) [2022-4-10]. https://arxiv.org/pdf/1607.06450v1.pdf.
[19] IOFFE S, SZEGEDY C. Batch normalization: Accelerating deep network training by reducing internal covariate shift [EB/OL]. (2015-03-02) [2022-4-10]. https://arxiv.org/pdf/1502.03167.pdf.
[20] HU Z, WEN Y, LIU L, et al. Visual classification of furniture styles[J]. ACM Transactions on Intelligent Systems and Technology, 2017, 8(5): 1-20.
[21] ZHU Bin, YANG Cheng, YU Chunyang, et al. Product image recognition based on deep learning[J]. Journal of Computer-Aided Design & Computer Graphics, 2018, 30(9): 1778-1784 (in Chinese). [朱 斌, 杨 程, 俞春阳, 等. 基于深度学习的产品意象识别[J]. 计算机辅助设计与图形学学报, 2018, 30(9): 1778-1784.]
[22] GONG Y, WANG B, RAU P-L P. PIKAR: A pixel-level image Kansei analysis and recognition system based on deep learning for user-centered product design[C] //International Conference on Human-Computer Interaction (HCII). Berlin, Heidelberg: Springer-Verlag, 2020: 60-71.
[23] ZHOU A, LIU H, ZHANG S, et al. Evaluation and design method for product form aesthetics based on deep learning[J]. IEEE Access, 2021, 9: 108992-109003.
[24] WANG Yahui, YU Suihuai, CHEN Dengkai, et al. Artificial intelligence design decision making model based on deep learning[J]. Computer Integrated Manufacturing Systems, 2019, 25(10): 2467-2475 (in Chinese). [王亚辉, 余隋怀, 陈登凯, 等. 基于深度学习的人工智能设计决策模型[J]. 计算机集成制造系统, 2019, 25(10): 2467-2475.]
[25] SU Z, YU S, CHU J, et al. A novel architecture: Using convolutional neural networks for Kansei attributes automatic evaluation and labeling[J]. Advanced Engineering Informatics, 2020, 44: 1-10.
[26] GUO M-H, XU T-X, LIU J-J, et al. Attention mechanisms in computer vision a survey [EB/OL]. (2021-11-15) [2022-3-9]. https://arxiv.org/pdf/2111.07624v1.pdf.
[27] PAN S J, YANG Q. A Survey on transfer learning[J]. IEEE Transactions on Knowledge and Data Engineering, 2010, 22(10): 1345-1359.
[28] Liu Ying, Lei Yanbo, Fan Jiulun, et al. Survey on image classification technology based on small sample learning[J]. Acta Automatica Sinica, 2021, 47(2): 297-315 (in Chinese). [刘 颖, 雷研博, 范九伦, 等. 基于小样本学习的图像分类技术综述[J]. 自动化学报, 2021, 47(2): 297-315.]
[29] QIU Xipeng. Neural networks and deep learning[M]. Beijing: China Machine Press, 2020: 197-200 (in Chinese). [邱锡鹏. 神经网络与深度学习[M]. 北京: 机械工业出版社, 2020: 197-200.]

[30] ZEILER M D, FERGUS R. Visualizing and understanding convolutional networks [EB/OL]. (2013-11-28) [2022-4-10]. https://arxiv.org/pdf/1311.2901v3.pdf.
[31] WERTHEIMER M. Laws of organization in perceptual forms[J]. 1923, 4: 301-350.
[32] HOU Q, ZHOU D, FENG J. Coordinate attention for efficient mobile network design [EB/OL]. (2021-03-04) [2022-4-10]. https://arxiv.org/pdf/2103.02907.pdf.
[33] WOO S, PARK J, LEE J-Y, et al. CBAM: Convolutional block attention module [EB/OL]. (2018-07-18) [2022-4-10]. https://arxiv.org/pdf/1807.06521.pdf.
[34] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [EB/OL]. (2017-12-06) [2022-4-10]. https://arxiv.org/pdf/1706.03762.pdf.
[35] LOSHCHILOV I, HUTTER F. Fixing weight decay regularization in Adam [EB/OL]. (2019-01-04) [2022-4-10]. https://arxiv.org/pdf/1711.05101v3.pdf.

作者简介：

李 雄（1984-），男，新疆石河子人，讲师，博士研究生，研究方向：智能设计、参数化设计、草图思维等，E-mail：leao_lx@lut.edu.cn；
+苏建宁（1974-），男，甘肃康乐人，教授，博士，博士生导师，研究方向：感性工学、智能设计等，通讯作者，E-mail：sujn@lut.edu.cn；
张志鹏（1994-），男，甘肃天水人，博士研究生，研究方向：智能设计、参数化设计，E-mail：zhangzhipeng@lut.edu.cn；
祝 铎（1997-），男，甘肃兰州人，硕士研究生，研究方向：智能设计、参数化设计等，E-mail：zhuduo@lut.edu.cn；
鱼宝银（1997-），男，陕西榆林人，硕士研究生，研究方向：感性工学、智能设计等，E-mail：yubaoyin@lut.edu.cn。

录　用　通　知

李雄，苏建宁，陈彦蒿 ，张秦玮，张新新，杨文瑾作者：

您（们）好!

您（们）的论文《**产品意象造型设计应用研究进展**》已收悉，经专家评审，同意发表。根据科技部和国家新闻出版署有关科技期刊收取版面发表费的规定，并从目前编辑部自筹全部办刊经费的实际情况出发，本刊拟征收论文发表费。按规定版面发表费可由您所在单位、课题组、项目组或本人支付。

您（们）的论文拟收 5000 元版面费，并在《包装工程》2019 年第 40 卷第 8 期《选题策划》栏目发表。

为保证您的文章能按时编辑、出版、发行，在收到本通知后，若文章已在其他刊物上发表，请来电、来信告知。若需要发表您的文章，请及时汇出版面费。不需要正式发票及有其他要求和意向，请及时来电、来函与编辑部有关人员联系。由于版面有限，原则上未收到版面费之前暂缓发表，请谅解。

邮 局 汇 款：
收款单位地址：重庆市九龙坡区渝州路 33 号　　　**邮编：**400039
收　　款　　人：《包装工程》编辑部

银 行 信 汇：
收款单位：重庆五九期刊社
开 户 行：中国银行重庆高新支行（行号：104653004623）
帐　　号：108802941586

联系地址：重庆市九龙坡区渝州路 33 号《包装工程》编辑部　**邮编：**400039
联系电话：023-68792836　　　**传真：**023-68793154
E-mail:bzgc59@126.com　　　**网址：**www.packjour.com

款到即寄正式发票或收据。感谢支持，愿合作愉快，欢迎继续来稿。

《包装工程》编辑部

2019 年 2 月 14 日

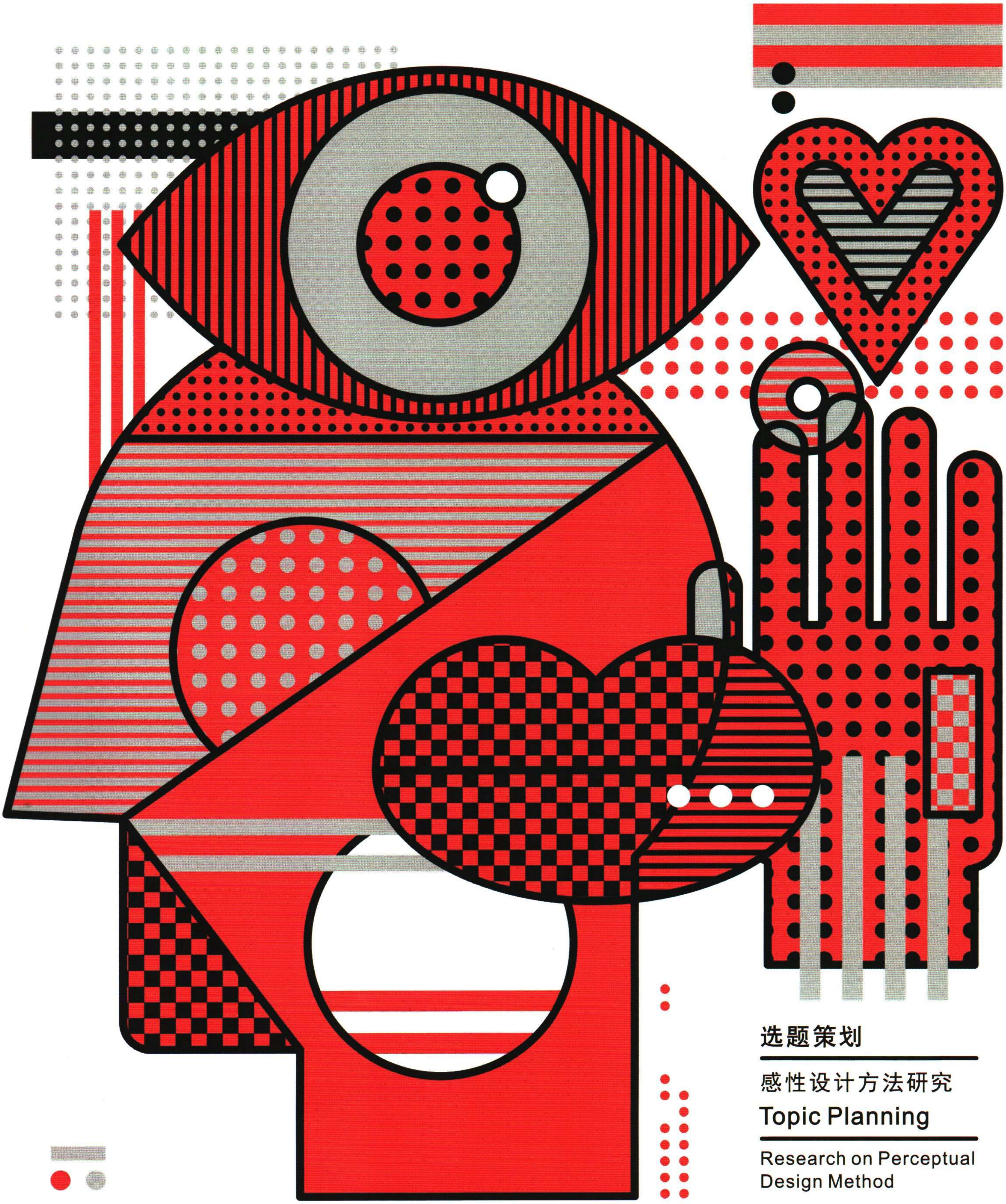
全 国 中 文 核 心 期 刊

2019年4月 第40卷 第8期

主管单位：中国兵器装备集团有限公司
主办单位：中国兵器工业第五九研究所

ISSN 1001-3563
CN 50-1094 / TB

包装工程
Packaging Engineering



选题策划

感性设计方法研究

Topic Planning

Research on Perceptual Design Method

包装工程
PACKAGING ENGINEERING

半月刊　1980 年创刊
Semimonthly　Initial Issue 1980

第 40 卷　第 8 期　2019 年 4 月 20 日　总第 398 期
Vol.40 No.8　Apr.20,2019　Total No.398

SUPERINTENDENT: China South Industries Group CO.,LTD.
SPONSOR: Southwest Technology and Engineering Research Institute
PRESIDENT: WU Hu-lin
EDITOR-IN-CHIEF: WU Hu-lin
VICE PRESIDENTS: ZHANG Lun-wu　LI Wen-tao　HU Jiang-hua
EXECUTIVE EDITOR: LI Wen-tao
EDITORIAL DEPARTMENT DIRECTOR: LIANG Ai-feng
ASSOCIATE EDITORIAL DEPARTMENT DIRECTOR: TANG Yao-yao
ASSOCIATE EDITORS-IN-CHIEF: TANG Yao-yao
ADDRESS: NO.33 Yuzhou Road,Jiulongpo District,Chongqing,China
POSTCODE: 400039
TEL: 023-68792836
FAX: 023-68792283
WEBSITE: www.designartj.com
EMAIL: designartj@126.com
PUBLISHER: Chongqing Wujiu Periodicals Press
PRINTER: Chongqing Gongchuang Printing Plant
RESPONSIBLE EDITORS: CHEN Ying　LI Hao-yu
COVER DESIGN: HE Lei
TYPESETTING: GAO Li-hong　XU Yang-yang
ACADEMIC EXCHANGE DEPARTMENT: FENG Li-jun　ZHONG Ke-ping　CHENG Zheng-chong　CHEN Zuo
TEL: 023-68792283
EMAIL: wjqk59@126.com
INTERNATIONAL UNITARY ISSUE NUMBER: ISSN 1001-3563
DOMESTIC UNITARY ISSUE NUMBER: CN 50-1094/TB
ADVERTISING LICENSE: Chongqing Industrial and Commercial Advertisements No.040297
OVERSEA DISTRIBUTOR: China International Book Trading Corporation
DOMESTIC DISTRIBUTOR: Chongqing Bureau for Distribution of Newspapers and Journals, China Post
ORDER OFFICE: All Local Post Offices in China（78-30）
OVERSEA DISTRIBUTION: BM1799
PRICE: RMB 45.00

主管单位：中国兵器装备集团有限公司
主办单位：中国兵器工业第五九研究所

社长：吴护林
主编：吴护林
副社长：张伦武　李文涛（常务）　胡江华
执行主编：李文涛
编辑部主任：梁爱锋
编辑部副主任：唐瑶瑶
副主编：唐瑶瑶

编辑部
地址：重庆市九龙坡区渝州路 33 号
邮编：400039
电话：023-68792836
传真：023-68792283
网址：www.designartj.com
邮箱：designartj@126.com

编辑出版：重庆五九期刊社
印刷单位：重庆共创印务有限公司

责任编辑：陈　英　李昊瑜

封面设计：何　蕾
排版：高丽红　徐扬扬

学术交流部：冯利军　钟克萍　程正冲　陈　作
电话：023-68792283
邮箱：wjqk59@126.com

国际标准刊号：ISSN 1001-3563
国内统一刊号：CN 50-1094/TB

广告经营许可证：渝工商广字 040297 号
国外发行：中国国际图书贸易集团有限公司
国内发行：中国邮政集团公司重庆市报刊发行局
邮发代号：78-30
国外发行代号：BM1799
定价：45.00 元
国内外公开发行

数据库收录

全国中文核心期刊
中国学术期刊综合评价数据库来源期刊
中国核心期刊（遴选）数据库收录期刊
《中国期刊网》全文数据库收录期刊
《中国学术期刊（光盘版）》入编期刊
中文科技期刊数据库入编期刊
中国科技期刊精品数据库收录期刊
美国《化学文摘》（CA）收录期刊
美国《剑桥科学文摘》（CSA：MI）收录期刊

包装工程
BAOZHUANG GONGCHENG
第40卷 第8期 2019年4月20日

目 次

选题策划：感性设计方法研究

设计前沿

视觉传达设计

工业设计

【选题策划：感性设计方法研究】

产品意象造型设计应用研究进展

李雄[1]，苏建宁[1]，陈彦蒿[2]，张秦玮[3]，张新新[4]，杨文瑾[1]
（1.兰州理工大学，兰州 730050；2.西北工业大学，西安 710072；
3.北方民族大学，银川 750021；4.华东理工大学，上海 200237）

摘要：目的 对产品意象造型设计应用研究进行综述，分析其发展现状、热点和趋势等问题。**方法** 通过对国内外相关文献的研究，分析产品意象造型设计的体系结构和应用过程。**结果** 产品意象造型设计应用过程主要包括产品意象挖掘定位、产品造型要素分析、意象造型设计等方面，其中意象造型设计可从单目标意象、多维意象、意象形态仿生、意象形态融合等角度展开。**结论** 产品意象造型设计是现代工业设计的重要发展方向，其应用产品种类广泛，核心建模思想和方法在不断地完善和更新，其中准确挖掘产品意象，深层次解析产品要素、多维意象与智能化设计等，将是未来应用研究的难点和热点。

关键词： 产品设计；意象造型；研究进展
中图分类号： TB472　　**文献标识码：** A　　**文章编号：** 1001-3563(2019)08-0001-09
DOI： 10.19554/j.cnki.1001-3563.2019.08.001

Review of Product Image Form Design and Its Application

LI Xiong[1], *SU Jian-ning*[1], *CHEN Yan-hao*[2], *ZHANG Qin-wei*[3], *ZHANG Xin-xin*[4], *YNAG Wen-jin*[1]
(1.Lanzhou University of Technology, Lanzhou 730050, China; 2.Northwestern Polytechnical University, Xi'an 710072, China; 3.North Minzu University, Yinchuan 750021, China; 4.East China University of Science and Technology, Shanghai 200237, China)

ABSTRACT: The paper aims to review the application research of product image form design and understand its development status, hot spots and trends. It analyzed the system structure and application process of product image form design through research on relevant literature at home and abroad. Product image form design and application process mainly included product image mining and positioning, analysis of product form elements, image form design and other aspects. Image form design can be carried out from the perspective of single-objective image, multi-dimensional image, image form bionics, image form fusion and so on. Product image form design is an important development direction of modern industrial design. It is widely used in various product designs. Its core modeling ideas and methods are constantly updated and evolved. The difficulties and hotpots of future application research are accurately mine of product image, implicit mechanism of cognitive products, high-dimensional image fusion and intelligent design.

KEY WORDS: product design; image form; research progress

随着云计算、物联网、大数据、区块链、人工智能等新一代信息技术与先进设计方法理论体系的融合，如今设计的定义，不仅仅是指创建实体产品的过程，已发展为创建包括服务、流程、战略和政策等泛化创新的一种核心手段。从强调对产品特质属性的提取，转向关注人与产品之间复杂的相互作用，以其本

收稿日期： 2019-02-24
基金项目： 国家自然科学基金资助项目（51465037）
作者简介： 李雄（1984—），男，新疆人，兰州理工大学博士生，兰州城市学院讲师，主要研究方向为智能设计、参数化设计、设计思维。
通信作者： 苏建宁（1974—），男，甘肃人，兰州理工大学教授、博士生导师，主要从事感性工学、智能设计等方面的研究。

能、行为及反思的基本维度交替运行[1]，情感特征已成为产品研究的关键模块[2]。人们对产品直接体验的静态表现、动态思考与姿态行动已满足浅层的需求动机，但其抽象层的认知情感通常是促发行为产生的关键驱动因素。其中，物理约束、相关人因尺度、颜色及美学参数等外显物质层，逐渐被表征为情感需求的感知反应。感性设计旨在将用户的期望和情感转化为产品属性，以此对产品的语义和情感特征进行有效整合，形成一种基于情感语境和个人关系整体性的认知机制。产品设计过程中的多物理特性与其唤起的用户主观感知之间的关系，是感性认知链接的重要环节，有效选择与产品设计要素相关的感性变量，是产品情感化发展的基础[3]。随着情感测量、认知计算等方法与技术在交叉学科背景下的发展和完善，个体及社群的情感意识被提取，认知主体对产品表象信息的传递逐渐明晰，并将其表达为一种高度凝结的意识产物，即产品的感性意象[4]。产品感性意象是透过产品的外显语义、内在含义与认知主体所形成的沟通语言的媒介，是感性属性的具象精炼。产品意象造型设计是探讨设计中蕴含的用户感性属性与引起响应的设计要素之间关联的设计方法[5]。感性意象造型设计的过程中，依据不同的意象和产品类型，以情感契合度为标准进行优选，产品的感性信息被传达，情感认知纽带呈现出多因素、多层次的高阶综合体状态。

通过运用感性工学等理论进行产品意象造型设计，提升消费者对产品的情感体验，已成为工业设计的重要方向[6]，在现代制造与服务业中体现出相当重要的竞争力[7]。目前在产品的形态设计、色彩设计、材质设计方面，做出了较多的应用探索，其在产业领域的应用研究逐步建立了一些实用的方法和体系。

1 产品意象造型设计

产品意象造型设计的研究涉及感性工学、符号学、美学、计算机、语言学、心理学、认知学、人机工程等多种领域[8]，目前主要的方法有语意差分法、形态分析法、多元回归法、灰色关联法、邓氏关联法、数量化Ⅰ类理论、神经网络、遗传算法等，面向家居产品、交通工具、装备产品、文创产品等展开应用和研究。

产品意象造型设计大致可分为如下关键步骤：首先，理解设计任务，明确设计目标；其次，通过调研分析等，挖掘定位用户的需求意象；然后，针对目标产品，进行相关产品造型要素分析；最后，将用户需求意象与产品设计要素关联，进而运用相关方法展开产品造型创新设计。从应用研究的角度出发，综合考虑设计方法、意象维度、设计对象等主要因素，将产品意象造型设计分为单目标意象造型设计、多维意象造型设计、意象形态仿生设计、意象形态融合造型设计等 4 种类型，并进行深入探讨。产品意象造型设计的类别见图 1，从方法和技术的角度来看，4 种类型也是不断演进与交叉的结果。

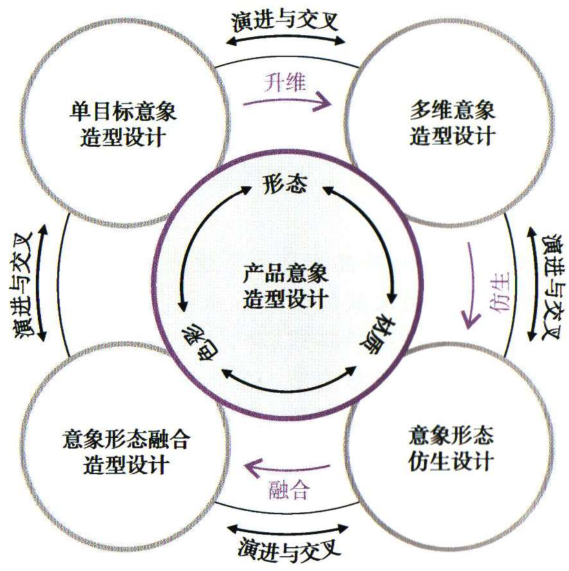


图 1 产品意象造型设计的类别
Fig.1 Category of product image form design

2 产品意象挖掘定位

产品意象的形成源自人们对产品的认知过程，人们透过产品本身的形态、色彩、质感和结构等因素，结合外在文化环境所赋予的内在含义，形成与产品沟通的语言[9]。产品造型意象一般由感性形容词表述，挖掘定位即确定产品意象造型设计的目标意象词汇，主要包括形态意象、色彩意象和材质意象等，见图 2。

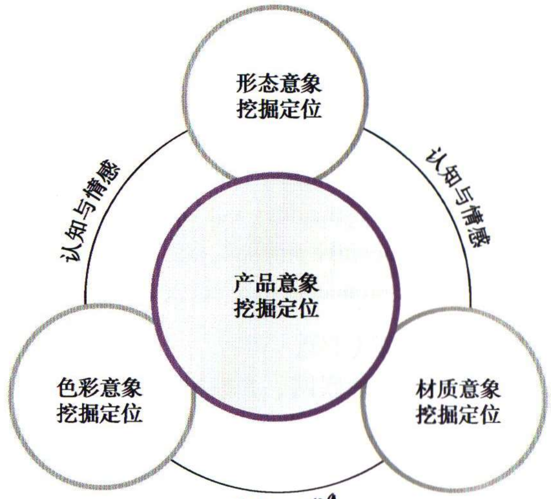


图 2 产品意象造型设计中的意象挖掘定位分类
Fig.2 Types of image mining positioning in product image form design

产品意象挖掘定位一般包括收集初始意象词汇、人工初步筛选和分析确定目标意象 3 个步骤。收集初

始意象词汇是广泛收集相关的意象词汇，主要方法有文献法、设计师描述[10]、用户描述、自然语言处理[11]和文本挖掘[12]等。人工初步筛选是设计师对收集的初始感性意象词汇进行分析和比较，删除明显不符合的意象词汇，以降低后期数据的处理数量。确定目标意象的方法有人工分类法、问卷调查法、数理统计分析法、粗糙集等。如曾栋等人[13]结合市场医用产品的现状，通过项目组讨论，确定了符合离心机造型语义的 4 个意象词语；陈黎等人[10]经过问卷调查，统计出了办公自动化设备的 4 个目标意象；苏畅等人[14]针对车身色彩设计，按照科学性等原则选出了 100 个感性词汇，并按照审美属性、描述属性、时间属性等划分成了 11 个类别，筛选出了 21 组词汇，最后通过认可度问卷调查筛选出了 6 组目标感性词汇；沈洁等人[15]整理出了市面上最常用的手机材质及描述手机的感性词汇，运用因子分析法得出了常用材质语义与用户感性词汇之间的映射关系，为后续手机的感性设计提供了指导。

感性形容词表达造型意象是相对主观的，其具有一定的局限性。目前，人们与网络的互动越发密切，产生了大量的用户数据，例如从在线产品描述与评论中，可有效提取和总结用户情感特征和感知偏好[16—17]。另外，基于生理信号的感知意象挖掘也是近年研究的一个方向[18—19]，同时应用机器学习、深度学习技术在产品意象识别和情感挖掘分析中亦可获得良好效果[20—21]，联合生理信号数据和互联网用户数据，并借助深度学习技术能够有效地逼近和预测用户的感性意象需求。由此可见，构建以生理信号数据、互联网用户数据为联合数据，并应用深度学习技术挖掘定位产品感知意象是未来的一个研究热点。

3 产品造型要素分析

产品造型要素分析的核心是产品造型特征的解构，是对原始设计的逆向思维过程。产品激发出的情感意象，是产品的形态、色彩、材质等多维造型要素联合作用的结果[22]。产品造型要素是用户感知中的关键作用点，包含显性特征和隐性特征，所有特征在人机交互过程中显现，与人们的综合感官密切相关。依据 Norman[23]的三层次理论，产品造型要素分析可从外观层、使用交互层和精神层方面展开[24—25]，见图 3。

外观层关注的是产品的本体特征，即形态、色彩、材质纹理，是在本能层的解构。外观要素可视为形态变量集、色彩变量集、材质变量集等 3 个主要本体设计变量的集合，解构常用形态分析、因素分析、聚类分析等方法。其中最常用的是形态分析法，由 Zwick[26]提出，最早用于设计巡航导弹，在工业设计领域得到了广泛应用，其关键步骤为要素分解、要素分析、构造设计矩阵等，如熊艳等人[27]通过产品形态特征线解析了产品外观形态上的特征要素。另外，也有学者从外观基因特征入手分析，如李雪瑞等人[28]采用 21 个形态基因参数解析并定义了汽车侧面轮廓特征参数；刘肖健等人[29]基于聚类提取特征颜色，构建出了色彩网络辅助设计配色。由于视网膜上的图像与物体反射之间并不是简单的映射关系，人们对色彩和材料的感知是复杂视觉计算的结果[30]。对产品材质的解析，主要从视觉感知和触觉感知方面进行，还需从材质本质特征出发探寻材质变量[31]，例如木材可考虑粗糙度、表面纹理、平滑度等，玻璃材质则包含透明度、折射率等。

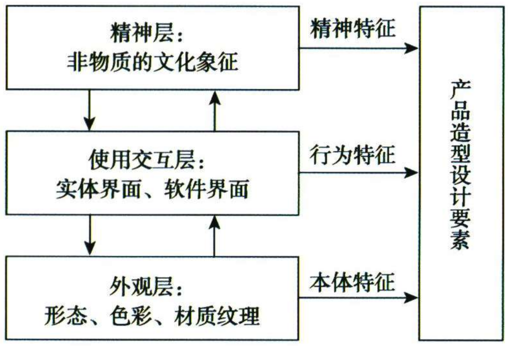


图 3　产品造型要素解构的三层次模型

Fig.3 Three hierarchical models of deconstruction of product form elements

使用交互层是在行为层的解构，重点在于分析信息的输入与反馈，以及形态构成的逻辑顺序、层级架构、界面布局等。产品造型通过使用和交互才能让用户获得真实的体验，包括能量与信息的交互，在三层次中属于中间过渡层，是精神层与外观层的桥梁，因此，从使用交互层面解构造型要素能获得更重要的特征参数，让信息交互得更有秩序，能量交换更合理。Han S H 等人[32]将设计变量定义为用户看、听、触摸或操作等综合感官的人机界面元素（HIEs）的集合，从软件和硬件两个方面对产品造型要素进行了解析；刘玲玲等人[24]从使用交互层对 Philips 空气净化器的造型要素解构出了 21 项人机特征参数，并进行了编码分析。

精神层是在反思层的解构，偏重于审美和情感的共鸣，具有沟通心灵的作用。在产品外观和使用功能解析的基础上，偏重对产品内涵的解析，关注产品造型要素传递的符号和象征，关注非物质层面的主体概念、故事意蕴、社会特质等。精神层面的解析是研究的难点，目前在产品造型要素分析中往往被忽略，如 Guo Fu 等人[33]在数码相机的特征解构中则忽略了精神层面的解析。

4 单目标意象造型设计

单目标意象造型设计，首先是通过研究多个设计

因子和一个目标意象之间的对应关系，常用的方法有类目层次法、数量化Ⅰ类、多元回归、人工神经网络、以及目前备受关注的深度学习等[4]；然后是从定性和定量层面上，分析并压缩设计师与用户需求间的认知距离，进而针对需求意象展开创新设计，见图4。从用户意象集中甄选一个目标意象，分析造型要素解构设计变量，依据具体条件选择合适的意象建模方法，逐步压缩认知主体间的认知距离，满足用户感性偏好。

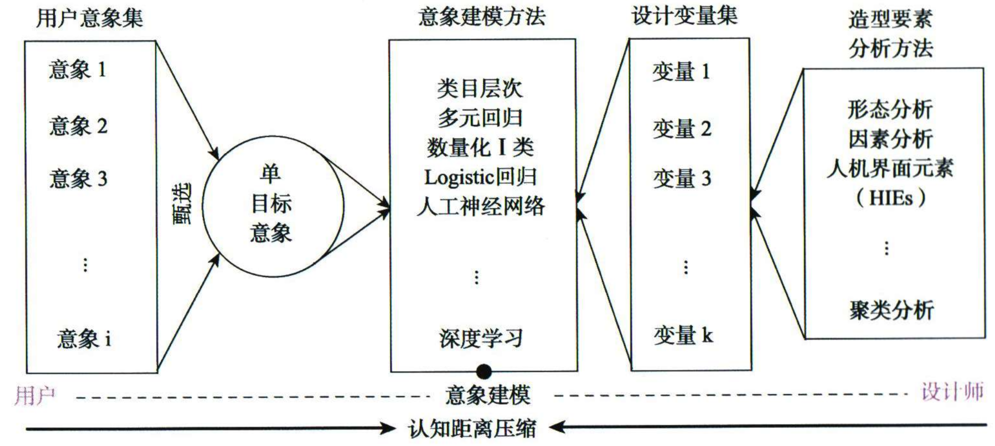


图4　单目标意象造型设计示意
Fig.4 Schematic diagram of single target image form design

苏建宁等人应用类目层次法逐级透析出与电热水壶设计相关的设计要素，从而对电热水壶进行了感性设计[34]；Hsiao S W[35]应用数量化Ⅰ类，并结合语意差分法进行了办公椅的设计；熊艳等人[27]通过多元回归探索单目标意象与手机特征线之间的关系，有助于手机形态的创新设计；李永锋等人[36]以办公椅为设计对象，利用次序Logistic回归构建了意象与设计要素间的相关性，发现了办公椅的"腿脚"对用户的感性意象影响最大，而"头枕"的影响最小；朱炜等人[37]从单个品牌意象出发，构建了以硬朗-圆润为目标品牌意象的多元线性模型，并设计了吉利SUV；李雪瑞等人[28]通过复杂网络理论构建了感性意象形态基因网络(K-FGN)，并以"优雅"为单目标意象对汽车侧轮廓进行了设计实践，取得了良好的设计效果；张硕等人[38]通过反向传播（BP）神经网络为壁挂式充电桩构建了感性意象与设计因子间的关联模型；Diego-Mas等人[39]提出了一种基于神经网络的产品形态设计情感反应建模方法，为单个用户的感知开发了一个理论框架，建立了预测单个用户对不同产品反应的数学模型；胡志刚等人[40]基于产品感性意象与配色之间的关联性，应用神经网络实现了豆浆机配色设计的智能化；Chen H Y等人[41]则开发了一种基于数字定义方案（NDS）和BP神经网络（BPNN）的计算机辅助产品形态设计（CAPD）工具。

单目标意象与设计因子间映射模型的构建，从最初的类目层次法到复杂网络的应用，从多元回归到神经网络，从线性映射机制到非线性映射机制，计算模型的能力逐步增强，智能设计的效果越来越接近人脑的形象思考，更好地拉近了设计师与用户感知之间的距离。

5　多维意象造型设计

现实中消费者对产品的情感意象需求具有多样性和复杂性[5]，要求产品设计同时满足多个情感复合意象需求。人类的情感受视觉、触觉、嗅觉、味觉、听觉5种感觉的影响，具有多维特性，用户的情感意象自然也与多种感觉器官捕获的能量和信息有关[42]。另外，多维意象也可能是用户、设计师、工程师等共同作用下的意象融合，因此，用户复合意象、多意象融合等都属于多维意象形态设计的范畴。由此可见，多维意象形态设计是指多个目标意象与众多设计因子间的关系，即多对多的映射关系，见图5。应用研究的基本思路是将多维意象降维或加权融合[43]，常用的方法有熵值法、模糊理论、灰度关联、基于精英保留的非支配排序遗传算法（Non-dominated Sorting Genetic Algorithm-II，NSGA-Ⅱ）[9]、复杂网络理论等，后续的过程与单目标意象造型设计基本一致。

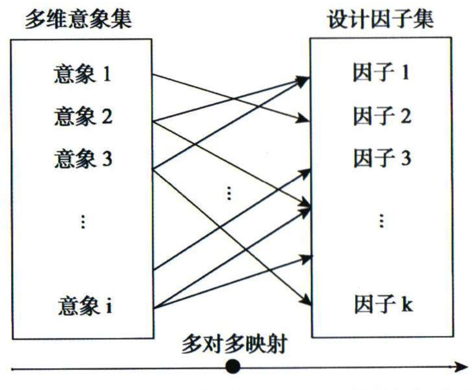


图5　多维意象造型设计的多对多映射关系示意
Fig.5 Schematic diagram for many-to-many mapping of multi-dimensional image form design

苏建宁等人[44]在融合用户、设计师、工程师认知的基础上，以“简洁”、“舒适”、“时尚”为复合目标意象，对矿泉水瓶形态设计进行了分析；陈国东等人[45]为了满足用户的多维度意象需求，构建了一种基于多目标优化的复合意象形态设计方法，并以豆浆机为例进行了设计验证；Ding Man 等人[46]以儿童药片为设计对象，通过灰色预测模型构建了设计要素与用户感知图像的关联模型，并采用粒子群算法对片剂形态设计方案进行了优化；柳禄等人[47]提出了一种多意象驱动的拖拉机产品族外形基因进化设计方法；李愚等人[48]以 20 对目标意象词为网络节点，构建出了汽车三维形态的意象复杂网络，并借助数据可视化技术分析了其复杂关系；Yumer M E 等人[49]将形态语义作为设计驱动因子构建了形态驱动模型，即意象语义与设计间的复杂网络模型，并借助数据可视化技术实现了对产品形态的意象编辑；LI Z 等人[50]提出了一种基于机器学习的情感设计动态映射方法，采用 4 种机器学习算法对设计元素与用户情感意象进行了关联建模，并以智能手表设计进行了验证。目前，随着降维技术和处理复杂模型能力的提升，以及多目标优化技术的应用[51]，多维意象造型设计的应用研究越来越深入，在高维意象融合、意象建模精度提升等方面取得了一定的成果。

6 意象形态仿生设计

模仿大自然是人类获得设计灵感的重要来源之一。意象形态仿生设计是将感性工学与仿生学相结合的成果，是意象造型设计研究及应用的一个重要方面。意象形态仿生重点在于生物神态特征及其象征意义的挖掘，用高度凝练的手法将生物的神态特征提炼出来，应用在产品形态设计中[52]，其目标是融情于物，情态共鸣。在意象形态仿生设计中，从挖掘用户意象到选择匹配的生物，再从生物形态向产品形态的融合过程，并不是某个单一的匹配过程，而是经过两次映射达到设计目标，即二重映射关系，见图 6。用户意象通过仿生对象与设计要素间的关联模型获得表达，其意象维度可以是单目标意象，也可以是多维意象，形态、色彩、材质都可通过此过程获得呈现。二次映射模型则是设计的关键，不仅建立了生物形态特征与产品形态间的映射模型，更重要的是实现了从用户意象到生物形态特征语义的映射过程。

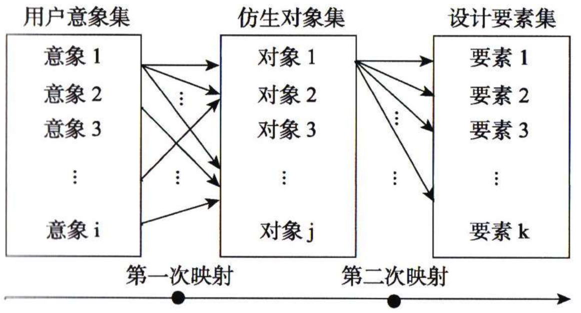


图 6 意象形态仿生二重映射关系示意
Fig.6 Schematic diagram for double mapping of image form bionics

Ding L 等人[53]应用语意差分法获取了用户对产品的感性意象和仿生对象，用聚类分析和模糊综合评价获取了电饭锅形态特征和仿生对象特征，并最终指导设计出了具有贝壳意象的电饭锅新形态；陆冀宁等人[54]提出了高速列车意象仿生思维映射设计方法，以大白鲨为仿生对象，考虑高速列车约束条件与大白鲨特征的匹配问题，成功设计了具有鲨鱼意象的高速列车，并对方案进行了气动性能评估；袁雪青等人[55]采用灰度关联法对用户意象词进行了意象评价实验，从而聚类出核心意象词汇匹配生物原型，并借助 CorelDraw 平台开发了意象形态仿生设计基因库，从而提高了意象造型设计的效率；高小针等人[56]以大象为形态仿生对象，借助眼动实验分析了高压电机的关键形态设计要素，在高压电机造型特征中融入了大象的意象特质，隐喻地传达了高压电机高强度的性能；朱赫[57]在认知耦合的基础上，以企鹅为意象仿生对象，开发了企鹅水壶意象仿生智能设计系统。以上研究，均体现出意象仿生的二重映射模式，折射出设计认知的过程，但还不够深入和完善，需要不断地去探索。

7 意象形态融合造型设计

借助计算的形态融合技术，能够自动生成大量与众不同的新形态。形态融合是指将两个或多个初始形态，光滑且连续地变换为继承初始形态特征的中间形态的技术过程，主要依靠形态融合算法[58]来实现。形态融合的同时意象也随之融合，由此展开基于形态融合技术的意象形态设计，意象形态融合造型设计示意，见图 7。目前，常利用形态混合技术，在甄选的目标形态和初始形态之间生成新的中间形态。在二维方面，通常应用插值法拟合计算产品形态特征线，如局部特征线的融合[59—60]。三维方面，则主要使用混合变换技术融合形态特征数据。无论是二维或是三维融合后所获得的新形态，往往具有序列

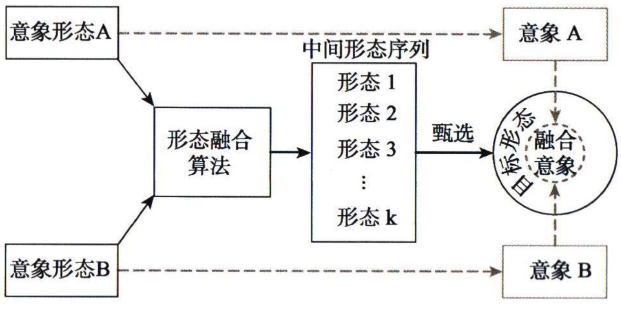


图 7 意象形态融合造型设计示意
Fig.7 Schematic diagram for image form fusion design

性和衍生性，形态间会有微妙的变化，从中筛选融合度较好的方案进行二次创作，如此，形态意象便可巧妙融入其中。

李明珠等人[61]以数码相机为研究对象，将形状混合技术引入了产品意象造型设计中，构建出了基于特征的形态混合设计方法，并对数码相机的轮廓线和表面装饰线进行了意象形态混合；韩超艳[62]通过差值算法构建了二维形态混合器，以"硬朗"为目标意象对SUV侧面轮廓进行了融合设计；Breen D E[63]提出了一种基于层级设置方法的变形技术，可以与多种扫描转换、模型处理技术相结合，创建出了一个通用的变形方法，用户可在一个动画中创造出一个变形序列；Lin C H等人[64]提出了一种通过动态添加或删除顶点，将三维多面体的连通性从源模型逐步转化为目标模型，同时生成中间形态的三维变形技术；苏建宁等人[65]提出了基于球面调和映射的方法，对两个具有形态差异的鼠标进行了融合，从而获得了更多的中间鼠标形态。形态融合的本质是数据的融合，近年来参数化脚本技术的发展对三维形态融合设计产生了良好的促进作用[66]，例如Grasshopper、Dynamo Studio等可视化编程语言的出现，降低了实现形态融合的算法门槛。

8 结语

意象造型设计面向的产品种类广泛，关键步骤大致相同，但核心建模思想和方法在不断完善和更新。意象挖掘定位已不满足于感性词汇的映射，开始向生理数据和网络大数据渗透，特别是基于生理信号数据和网络用户数据的感知意象挖掘得到了研究者的青睐，应用人工智能辅助挖掘分析可取得更好的效果。不同的目标意象确定方法各有特点，初始意象词汇的广泛性、人工分类法的客观性、问卷调查法的准确性以及数理统计分析中基础数据的有效性等，将是未来研究的重点。

产品造型要素分析可从3个层次分析，外观层属于底层，是造型的本体层，产品意象造型设计常在该层面展开；使用交互层处于中间衔接层，是信息与能量的置换渠道，产品的交互设计需在此层面展开；精神层位于顶层，是对产品内涵的解析，该方面的应用研究是未来重要的方向。

意象建模技术和方法从简单到复杂，已呈现多样化、非线性化、交叉化的特点，所构建的模型逐渐拉近了用户认知和设计师认知间的距离。从一对多到多对多映射模型的构建，反映了设计创新的真实情况和复杂程度，意象分析模型能够处理的设计变量逐步增多。产品意象造型设计研究朝着多维意象与智能化设计的方向发展，并呈现出向大自然获取意象灵感和形态融合的发展态势。

参考文献：

[1] RAZZA B, PASCHOARELLI L C. Affective Perception of Disposable Razors: a Kansei Engineering Approach[C]. Amsterdam: Procedia Manufacturing, 2015: 6228—6236.

[2] CHANG Yu-ming, CHEN Chun-wei. Kansei Assessment of the Constituent Elements and the Overall Interrelations in Car Steering Wheel Design[J]. International Journal of Industrial Ergonomics, 2016, 56: 97—105.

[3] VIEIRA J, OSORIO J M A, MOUTA S, et al. Kansei Engineering as a Tool for the Design of In-vehicle Rubber Keypads[J]. Applied Ergonomics, 2017, 61: 1—11.

[4] 苏建宁，王鹏，张书涛，等. 产品意象造型设计关键技术研究进展[J]. 机械设计, 2013, 30(1): 97—100.
SU Jian-ning, WANG Peng, ZHANG Shu-tao, et al. Review of Key Technologies of Product Image Styling Design[J]. Journal of Machine Design, 2013, 30(1): 97—100.

[5] 苏建宁，张秦玮，张书涛，等. 产品意象造型进化设计研究进展[J]. 机械设计, 2014, 31(2): 97—102.
SU Jian-ning, ZHANG Qin-wei, ZHANG Shu-tao, et al. Development of Product Image Modeling Evolutionary Design[J]. Journal of Machine Design, 2014, 31(2): 97—102.

[6] ZHOU Yan, YANG Ping, WANG Si-yu. Research of Kansei Image Based on Product Appearance Form Deconstruction[J]. Advanced Materials Research, 2014, 971—973: 1316—1320.

[7] WONJOON K, TAEHOON K, IISUM R, et al. Mining Affective Experience for a Kansei Design Study on A Recliner[J]. Applied Ergonomics, 2019, 74: 145—153.

[8] 罗仕鉴，潘云鹤. 产品设计中的感性意象理论、技术与应用研究进展[J]. 机械工程学报, 2007, 43(3): 8—13.
LUO Shi-jian, PAN Yun-he. Review of Theory, Key Technologies and Its Application of Perceptual Image in Product Design[J]. Journal of Mechanical Engineering, 2007, 43(3): 8—13.

[9] 苏建宁，张秦玮，吴江华，等. 产品多意象造型进化设计[J]. 计算机集成制造系统, 2014, 20(11): 2675—2682.
SU Jian-ning, ZHANG Qin-wei, WU Jiang-hua, et al. Research on Evolutionary Design of Product Multi-image Styling[J]. Computer Integrated Manufacturing Systems, 2014, 20(11): 2675—2682.

[10] 陈黎，朱如鹏，周海海. 新技术商品化中产品造型感性需求挖掘方法研究[J]. 中国机械工程, 2010, 21(23): 2803—2808.
CHEN Li, ZHU Ru-peng, ZHOU Hai-hai. Research on Mining Method of Perceptual Needs of Product Form in Commercialization of New Technology[J]. China Mechanical Engineering, 2010, 21(23): 2803—2808.

[11] 卢兆麟, 程若丹, 石清吟, 等. 基于自然语言处理的汽车造型风格推导与评价[J]. 汽车工程, 2016, 38(5): 553—560.
LU Zhao-lin, CHENG Ruo-dan, SHI Qing-yin, et al. Vehicle Styling Feature Derivation and Evaluation Based on Natural Language Processing[J]. Automotive Engineering, 2016, 38(5): 553—560.

[12] 蒲骄子, 李延来, 刘宗鑫. 基于文本挖掘与神经网络的高速列车意象造型设计[J]. 机械设计, 2017, 34(9): 101—105.
PU Jiao-zi, LI Yan-lai, LIU Zong-xin. Image Modeling Design of High-Speed Train Based on Text Mining and Neural Network[J]. Journal of Machine Design, 2017, 34(9): 101—105

[13] 曾栋, 巩敦卫, 李梅子, 等. 产品造型设计中的思维固化策略及应用[J]. 机械工程学报, 2017, 53(15): 58—65.
ZENG Dong, GONG Dun-wei, LI Mei-zi, et al. Thinking Fixation Strategy in Product Form Design and Its Application[J]. Journal of Mechanical Engineering, 2017, 53(15): 58—65.

[14] 苏畅, 付黎明, 魏君, 等. 基于感性工学和主成分分析的车身色彩设计[J]. 吉林大学学报(工学版), 2016, 46(5): 1414—1419.
SU Chang, FU Li-ming, WEI Jun, et al. Design Method in Exterior Color Based Kansei Engineering and Principal Component Analysis[J]. Journal of Jilin University(Engineering and Technology Edition), 2016, 46(5): 1414—1419.

[15] 沈洁, 薛澄岐, 朱嫣绯. 材质语义在手机设计中的应用[J]. 设计, 2018(18): 123—125.
SHEN Jie, XUE Cheng-qi, ZHU Yan-fei. Application of Material Semantics in Mobile Design[J]. Design, 2018(18): 123—125.

[16] WANG W M, LI Z, TIAN Z G, et al. Extracting and Summarizing Affective Features and Responses from Online Product Descriptions and Reviews: a Kansei Text Mining Approach[J]. Engineering Applications of Artificial Intelligence, 2018, 73: 149—162.

[17] IRELAND R, LIU A. Application of Data Analytics for Product Design: Sentiment Analysis of Online Product Reviews[J]. Cirp Journal of Manufacturing Science and Technology, 2018, 23: 128—144.

[18] 张艳河. 基于脑电的用户感知意象思维表征[J]. 机械设计, 2017, 34(6): 117—122.
ZHANG Yan-he. User Thinking Representation of Product Image Based on ERPs[J]. Journal of Machine Design, 2017, 34(6): 117—122.

[19] 张艳河, 杨颖, 罗仕鉴, 等. 产品设计中用户感知意象的思维结构[J]. 机械工程学报, 2010, 46(2): 178—184.
ZHANG Yan-he, YANG Ying, LUO Shi-jian, et al. Mental Construction of User Perception Image in Product Design[J]. Journal of Mechanical Engineering, 2010, 46(2): 178—184.

[20] 朱斌, 杨程, 俞春阳, 等. 基于深度学习的产品意象识别[J]. 计算机辅助设计与图形学学报, 2018, 30(9): 1778—1784.
ZHU Bin, YANG Cheng, YU Chun-yang, et al. Product Image Recognition Based on Deep Learning[J]. Journal of Computer-Aided Design & Computer Graphics, 2018, 30(9): 1778—1784.

[21] ZHANG Lei, WANG Shuai, LIU Bing. Deep Learning for Sentiment Analysis: a Survey[J]. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 2018, 8(4): 1—34.

[22] CHANG C C, WU J C. The Underlying Factors Dominating Categorical Perception of Product Form of Mobile Phones[J]. International Journal of Industrial Ergonomics, 2009, 39(5): 667—680.

[23] NORMAN D A. 设计心理学 3: 情感设计[M]. 北京: 中信出版社, 2015.
NORMAN D A. Emotional Design: Why We Love(or Hate) Everyday Things[M]. Beijing: China Citic Press, 2015.

[24] 刘玲玲, 薛澄岐. 面向多维 KE 模型构建的产品特征解构方法[J]. 工程设计学报, 2014, 21(4): 323—328.
LIU Ling-ling, XUE Cheng-qi. Product Feature Deconstruction Method by Constructing a Multi-dimensional Kansei Engineering Model[J]. Chinese Journal of Engineering Design, 2014, 21(4): 323—328.

[25] 苏建宁, 刘志君. 基于感性工学的传统器物形态创新设计方法研究[J]. 创意与设计, 2016, 47(6): 80—83.
SU Jian-ning, LIU Zhi-jun. A Research of Innovative Design Methodology for the Modality of Traditional Utensils Based on the Methods of Kansei Engineering[J]. Creation and Design, 2016, 47(6): 80—83.

[26] ZWICK W R, VELICER W F. Factors Influencing Four Rules for Determining the Number of Components to Retain[J]. Multivariate Behav Res, 1982, 17(2): 253—269.

[27] 熊艳, 李彦, 李文强, 等. 基于形态特征线意象量化的产品形态设计方法[J]. 工程科学与技术, 2011, 43(3): 233—238.
XIONG Yan, LI Yan, LI Wen-qiang, et al. An Image Quantization of Product Form Design Method Based on Form Feature Lines[J]. Advanced Engineering Sciences, 2011, 43(3): 233—238.

[28] 李雪瑞, 余隋怀, 初建杰, 等. 意象驱动的产品形态基因网络模型构建与应用[J]. 计算机集成制造系统, 2018, 24(2): 464—473.
LI Xue-rui, YU Sui-huai, CHU Jian-jie, et al. Construction and Application of Product Form Gene Network Model Driven by Kansei[J]. Computer Integrated Manufacturing Systems, 2018, 24(2): 464—473.

[29] 刘肖健, 曹愉静, 赵露唏. 传统纹样的色彩网络模型及配色设计辅助技术[J]. 计算机集成制造系统, 2016, 22(4): 897—905.
LIU Xiao-jian, CAO Yu-jing, ZHAO Lu-xi. Color Networks of Traditional Cultural Patterns and Color Design Aiding Technology[J]. Computer Integrated Manufacturing Systems, 2016, 22(4): 897—905.

[30] BRAINARD D H, COTTARIS N P, RADONJIC A, et al. The Perception of Colour and Material in Naturalistic Tasks[J]. Interface Focus, 2018, 8(4): 1—12.

[31] FLEMING R W. Visual Perception of Materials and Their Properties[J]. Vision Research, 2014, 94: 62—75.

[32] HAN S H, YUN M H, KIM K J, et al. Evaluation of Product Usability: Development and Validation of Usability Dimensions and Design Elements Based on Empirical Models[J]. International Journal of Industrial Ergonomics, 2000, 26(4): 477—488.

[33] GUO Fu, LIU Wei-lin, LIU Fan-tao, et al. Emotional Design Method of Product Presented in Multi-dimensional Variables Based on Kansei Engineering[J]. Journal of Engineering Design, 2014, 25(4—6): 194—212.

[34] 苏建宁, 王瑞红, 赵慧娟, 等. 基于感性意象的产品造型优化设计[J]. 工程设计学报, 2015, 22(1): 35—41.
SU Jian-ning, WANG Rui-hong, ZHAO Hui-juan, et al. Optimization Design of Product Modeling Based on Kansei Image[J]. Chinese Journal of Engineering Design, 2015, 22(1): 35—41.

[35] HSIAO S W, CHEN C H. A Semantic and Shape Grammar Based Approach for Product Design[J]. Design Studies, 1997, 18(3): 275—296.

[36] 李永锋, 朱丽萍. 基于序次 Logistic 回归的产品意象造型设计研究[J].机械设计, 2011, 28(7) : 8—12.
LI Yong-feng, ZHU Li-ping. Research on Product Image Form Design Based on Ordinal Logistic Regression[J]. Journal of Machine Design, 2011, 28(7): 8—12.

[37] 朱炜, 赵磊, 王小龙. 基于品牌意象的 SUV 造型设计方法[J]. 机械设计, 2017, 34(1): 122—125.
ZHU Wei, ZHAO Lei, WANG Xiao-long. SUV Modeling Design Method Based on Brand Image[J]. Journal of Machine Design, 2017, 34(1): 122—125.

[38] 张硕, 吴俭涛. 基于 BP 神经网络的壁挂式充电桩产品形态研究[J]. 图学学报, 2017, 38(6): 865—868.
ZHANG Shuo, WU Jian-tao. Research on Form of Wall Set Charging Pile Base on BP Neural Network[J]. Journal of Graphics, 2017, 38(6): 865—868.

[39] DIEGO-MAS J A, ALCAIDE-MARZAL J. Single Users' Affective Responses Models for Product Form Design[J]. International Journal of Industrial Ergonomics, 2016, 53: 102—114.

[40] 胡志刚, 吕娜, 乔现玲, 等. BP 神经网络在产品配色中的应用研究[J]. 包装工程, 2016, 37(10): 136—141.
HU Zhi-gang, LYU Na, QIAO Xian-ling, et al. Application of BP Neural Network in Product Color-matched Design[J]. Packaging Engineering, 2016, 37(10): 136—141.

[41] CHEN H Y, CHANG Y M. Development of a Computer Aided Product-form Design Tool Based on Numerical Definition Scheme and Neural Network[J]. Journal of Advanced Mechanical Design, Systems, and Manufacturing, 2014, 8(3): 1—15.

[42] HSIAO S W, TSAI H C. Applying a Hybrid Approach Based on Fuzzy Neural Network and Genetic Algorithm to Product Form Design[J]. International Journal of Industrial Ergonomics, 2005, 35(5): 411—428.

[43] 苏珂, 王硕. 产品多意象造型设计研究[J]. 机械设计, 2016, 33(3): 115—119.
SU Ke, WANG Shuo. Study of Product Modeling Design with Multiple Images[J]. Journal of Machine Design, 2016, 33(3): 115—119.

[44] 苏建宁, 张新新, 景楠, 等. 认知差异下的产品造型意象熵评价研究[J]. 机械设计, 2016, 33(3): 105—108.
SU Jian-ning, ZHANG Xin-xin, JING Nan, et al. Research on the Entropy Evaluation of Product Styling Image Under the Cognitive Difference[J]. Journal of Machine Design, 2016, 33(3): 105—108.

[45] 陈国东, 陈思宇, 王军, 等. 面向复合意象的产品形态多目标优化[J]. 中国机械工程, 2015, 26(20): 2763—2770.
CHEN Guo-dong, CHEN Si-yu, WANG Jun, et al. Multi-objective Optimization Method of Composite Imagery-oriented Product Form[J]. China Mechanical Engineering, 2015, 26(20): 2763—2770.

[46] DING Man, BAI Zhong-hang, LI Mei. Image-based Tablet Form Design[J]. Cluster Computing, 2018(6): 1—8.

[47] 柳禄, 章永年, 丁为民, 等. 多目标驱动的拖拉机产品族外形基因设计[J]. 农业工程学报, 2017, 33(17): 82—90.
LIU Lu, ZHANG Yong-nian, DING Wei-min, et al. Design of Multi-objective Driven Product Family Shape Gene for Tractor[J]. Transactions of the Chinese Society of Agricultural Engineering, 2017, 33(17): 82—90.

[48] 李愚, 卢纯福, 刘肖健, 等. 汽车外形设计的基因网络模型[J]. 计算机集成制造系统, 2018, 24(5): 189—200.
LI Yu, LU Chu-fu, LIU Xiao-jian, et al. Gen Network Model of Automobile Styling Design[J]. Computer Integrated Manufacturing System, 2018, 24(5): 189—200.

[49] YUMER M E, CHAUDHURI S, HODGINS J K, et al. Semantic Shape Editing Using Deformation Handles[J]. ACM Transactions on Graphics, 2015, 34(4): 1—12.

[50] LI Z, TIAN Z G, WANG J W, et al. Dynamic Mapping of Design Elements and Affective Responses: a Machine Learning Based Method for Affective Design[J]. Journal of Engineering Design, 2018, 29(7): 358—380.

[51] SHIEH M D, LI Y, YANG C C. Comparison of Multi-objective Evolutionary Algorithms in Hybrid Kansei Engineering System for Product Form Design[J]. Advanced Engineering Informatics, 2018, 36: 31—42.
[52] 祝莹, 曹建中, 韦艳丽. 汽车造型设计中的形态仿生研究[J]. 合肥工业大学学报（自然科学版）, 2010, 33(10): 1458—1460.
ZHU Ying, CAO Jian-zhong, WEI Yan-li. Research on Bionic Form Design in Automobile Modeling Design[J]. Journal of Hefei University of Technology(Natural Science), 2010, 33(10): 1458—1460.
[53] DING Lei, CHE Jian-ming, DONG Xiao-fei. Product Bionics Form Design Based on Users' Kansei Image[J]. Applied Mechanics and Materials, 2010, 37—38: 910—914.
[54] 陆冀宁, 徐伯初, 支锦亦, 等. 高速列车头型的意象仿生设计[J]. 机械设计, 2017, 34(9): 106—110.
LU Ji-ning, XU Bo-chu, ZHI Jin-yi, et al. Image Bionic Design of High Speed Stain Head Modeling[J]. Journal of Machine Design, 2017, 34(9): 106—110.
[55] 袁雪青, 陈登凯, 杨延璞, 等. 意象关联产品形态仿生设计方法[J]. 计算机工程与应用, 2014, 50(8): 178—182.
YUAN Xue-qing, CHEN Deng-kai, YANG Yan-pu, et al. Bionic Imagery Associated with Product Form Design[J]. Computer Engineering and Applications, 2014, 50(8): 178—182.
[56] 高小针, 张阿维. 基于眼动实验的高压电机意象仿生设计[J]. 西安工程大学学报, 2018, 32(3): 330—335.
GAO Xiao-zhen, ZHANG A-wei. Image Bionic Design of High-Voltage Motor Based on Eye-Tracking Experiment[J]. Journal of Xi'an Polytechnic University, 2018, 32(3): 330—335.
[57] 朱赫. 基于认知耦合的产品意象形态仿生进化设计方法[D]. 兰州: 兰州理工大学, 2018.
ZHU He. Cognitive Coupling Based Product Image Morphology Bionic Evolutionary Design Method[D]. Lanzhou: Lanzhou University of Technology, 2018.
[58] 沙春发, 卢章平, 何灿群. 三维形状混合技术在产品造型设计中的应用[J]. 包装工程, 2008, 29(8): 201—203.
SHA Chun-fa, LU Zhang-ping, HE Can-qun. Application of 3D Metamorphosis in Product Modeling Design[J]. Packaging Engineering, 2008, 29(8): 201—203.
[59] 袁浩, 薛澄歧, 卢章平, 等. 基于局部形态融合的产品创新设计研究[J]. 包装工程, 2011, 32(20): 30—33.
YUAN Hao, XUE Cheng-qi, LU Zhang-ping, et al. Product Innovative Design Based on Local Shape Integration[J]. Packaging Engineering, 2011, 32(20): 30—33.
[60] HUI K C, LI Y. A Feature-based Shape Blending Technique for Industrial Design[J]. Computer-Aided Design, 1998, 30: 823—824.
[61] 李明珠, 卢章平, 王春艳, 等. 基于形状混合的产品意象造型设计方法[J]. 机械设计, 2015, 32(7): 105—108.
LI Ming-zhu, LU Zhang-ping, WANG Chun-yan, et al. Method of Product Image Modeling Design Based on Shape Blending[J]. Journal of Machine Design, 2015, 32(7): 105—108.
[62] 韩超艳. 基于二维形状混合的产品形态设计方法[J]. 机械设计, 2015, 32(9): 116—119.
HAN Chao-yan. Product Form Design Method Based on 2D Shape Blending[J]. Journal of Machine Design, 2015, 32(9): 116—119.
[63] BREEN D E, MAUCH S, WHITAKER R T, et al. 3D Metamorphosis between Different Types of Geometric Models[J]. Computer Graphics Forum, 2001, 20(3): 36—48.
[64] LIN C H, LEE T Y. Metamorphosis of 3D Polyhedral Models Using Progressive Connectivity Transformations[J]. IEEE Transactions on Visualization & Computer Graphics, 2005, 11(1): 2—12.
[65] 苏建宁, 吴江华, 刘芸, 等. 基于球面调和映射的三维产品形态融合方法研究[J]. 图学学报, 2014, 35(3): 417—422.
SU Jian-ning, WU Jiang-hua, LIU Yun, et al. Research on 3D Product Shape Merging Technology Based on Spherical Harmonic Map[J]. Journal of Graphics, 2014, 35(3): 417—422.
[66] OXMAN R. Thinking Difference: Theories and Models of Parametric Design Thinking[J]. Design Studies, 2017, 52(9): 4—39.

第六届国际（洛阳）工业设计高峰论坛

2018年5月18-20日 中国洛阳

论 文 录 用 通 知

尊敬的作者，您（们）好！

经大会科学技术委员会的专家评审，您(们)的论文（见下表）已被本次会议录用，将被MATEC Web of Conferences(ISSN(Electronic Edition): 2261-236X)出版，并将由出版社提交给EI Compendex和ISTP检索。谨向您（们）表示衷心的祝贺！

稿件编号	DF6251
姓名	X Li， JN Su
论文名称	Research on Parametric Form Design Based on Natural Patterns
备注	

我们诚挚地邀请您前来中国洛阳参会，现将有关事项通知如下：

一、 提交电子文档

请按专家给出的审稿意见认真修改论文，填写和签署版权协议。并把修改后的论文（word和pdf版本）和版权协议一起发到组委会邮箱。

二、 会议注册

每篇被录用的论文必须注册，方可发表。会议注册费用为2900元人民币/篇；如果论文页数超出5页，则超出部分需另加300元/页。本次会议由洛阳市生产力促进中心代收中国大陆作者的会议注册费和出具发票。请按附件中给出的方法按时注册。

第六届国际（洛阳）工业设计高峰论坛

组委会

2018年4月27日

Research on Parametric Form Design Based on Natural Patterns

X Li[1], JN Su[2,a]

[1]School of Mechanical & Electronical Engineering, Lanzhou University of Technology,730050 Lanzhou Gansu, China
[2]School of Design Art, Lanzhou University of Technology, 730050 Lanzhou Gansu, China

Abstract. Parametric form design method based on natural patterns is proposed for the design schema of the traditional form bionics. Firstly, the concepts of natural patterns are analyzed, concluded and summarized. Secondly, the parametric design thinking is elaborated from three aspects: thinking mode, design flow, tools and scripts. It is also proposes a parametric logo design method. This paper takes logarithmic spiral pattern as an example to describe the process from the law of natural pattern to logo design, which includes nature pattern analysis, control rules of from and color, algorithm research and generating design.

1 Introduction

The aesthetic laws of statistical aesthetics, most of which are based on the structure and growth of plants and animals in the natural world[1]. Learning in nature exists at every stage of design development. Different stages of development have different understandings. From the initial appearance simulating to rational creation born out of the inside, humans gradually increase the integration of design and nature. Design absorbs nourishment from nature and helps to create beauty actively and regularly, including new forms and new structures. The design pursues the realm of natural harmony between man and nature.

Parametric technology has been widely used in architectural design, industrial product design, landscape design, fashion design, jewelry design and other design fields[2]. The goal of parameterization is to construct an automatic design system that can be edited at any time. Parametric technology provides infinite creativity for design[3]. It can obtain several design schemes at one time, thus improving the design efficiency.

For the design schema of the traditional bionic form, this paper focused on the mathematical logic behind natural patterns, and studied the parametric form design method based on natural patterns.

In the following Section II, we provide a brief introduction to the concept of natural pattern, including analysis and induction. Section III discusses the parametric design, including process, thinking model, tools and scripts. Section IVstudies the method of parametric logo design. Case study in Section V. The last part is the conclusion.

2 Natural Pattern

The term pattern is not unfamiliar and covers a wide range of fields, such as biology, economics, computer science and technology, physics.

We are in nature. Human exploration of natural patterns began as early as ancient Greek philosophers. Plato, Pythagoras, and Empedocles had tried to reveal the order of the natural world, and established the philosophical foundation for the study of natural patterns. In 1202, Leonardo Fibonacci published his famous book *Liber Abaci*, and raised the Fibonacci numbers on the issue of rabbit breeding[4]. In the 19th century, the Belgian physicist Joseph Plateau studied the soap film, prompting him to propose the concept of minimal surface [5]. The Scottish biologist and molecular mathematician D'Arcy Thompson took the lead in researching the growth patterns of plants and animals. In 1917, he published a book *On Growth and Form*. In his book, Thompson proposed to associate phyllotaxis with the Fibonacci sequence[6]. He showed that complex spiral growth can be explained by simple equations. In the 20th century, the British mathematician and the “father of artificial intelligence” Alan Turing pushed the study of natural patterns to a climax. In 1952, he wrote *The Chemical Basis of Morphogenesis*[7], and the seemingly complex and irregular pattern of spots and stripes could be described by the Turing equation. Then, in 1968, Hungarian biologist Aristid Lindenmayer developed the L- system to simulate the fractal of plant growth pattern[8]. The scientific research results of these natural patterns are the premise of ‘design research based on natural patterns’.

Patterns in nature are visible regularities of form found in the natural world[6]. Symmetry patterns are ubiquitous in nature. Humans and animals are mainly mirror-symmetrical and plants usually have radial or rotational symmetry. Non-living bodies also have symmetrical patterns, such as snowflakes, ice crystals, crystals and so on. Spirals is undoubtedly the most

*Corresponding author: [a] sujn@lut.cn

beautiful universal pattern, the initial pattern of life system. From the inanimate spiral galaxy to the genetic material of life the double helix of DNA molecule, petals, fingerprints, eddy currents, cochlea, the tendrils of plants and so on. Insects and animals show the extremes of the spots and stripes pattern. For example, leopards, cows and ladybugs are full of spots, zebras and angelfish have distinctive stripes, butterflies blend spots and stripes. Meanders are common in nature and life, such as the whereabouts of animals, brain gyrus, artificial mazes, and pedestrian trails. Branches and fractals are widespread in nature and have infinite self-similarity. Such as lightning, mountains, coastline, the body's nervous system and blood vessels branch, etc. Alongside fractals, chaos theory ranks as an essentially universal influence on patterns in nature. Mathematics tries to discover and explain various abstract patterns and rules. The visual model of nature can find explanations in logarithmic spirals, fractals, topology, chaos theory, and other mathematics, which lays a mathematical foundation for parametric design.

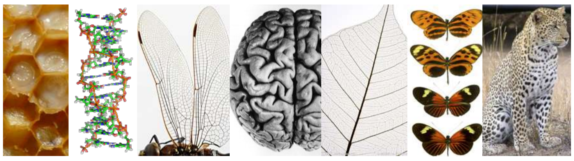


Figure 1. Examples of natural patterns (Wikipedia).

Patterns organizes and defines the relationship of nature, and can be applied to practical design to enhance and support visual communication[6]. This article attempted to develop the intersection between natural patterns and design sciences, and tried to combine natural patterns with the form design through computer programming techniques to explore a new form-aided design approach.

3 Parametric design

Parametric design originated from mechanical design[9]. The fundamental difference in parametric design is the application of computer programming techniques, compares to a typical sketch-manual modeling design schema. Its purpose is to construct an automated generative design system that can edit and modify parameters at any time. Therefore, the parametric design has changed in three aspects: Design Processes (DP), Design Thinking (DT) and Design Tools (DT). This also leads to the design cognitive model of designers to be changed.

3.1 Parametric design thinking(PDT)

Hugh Whitehead thinks that parameterization is more about a way of thinking[10]. Rivka Oxman[11] claims that this is mainly due to the intersection of three areas of knowledge: professional design knowledge, computer programming knowledge, and mathematical knowledge. Parametric design does not deal directly with the form, but rather studies the mathematical logic behind the form. Computer programs are used to calculate the various elements that affect the styling (parameter variables).

Robert Woodbury[10], in his book *Elements of Parametric Design* (2010), proposed that Parametric Design Thinking(PDT) has three main characteristics - abstract thinking, mathematical thinking, algorithmic thinking. Parametric design thinking processes shown in Figure 2 can also prove this view. Abstract thinking is the basis of parametric design, abstracting design ideas or concepts into symbols, data, variables, functions, etc. Mathematical thinking mainly involves how to translate mathematical theorems and data structures into useful algorithms [12]. The essence of algorithmic thinking is the step-by-step problem solving. In parametric design, algorithmic thinking means writing functions in scripting languages. Functions are part of the digital form, the editing function is to edit digital form.

In traditional paper-based design, designers mostly reduce the distance from the ideal design scheme by drawing sketches and making models (digital or physical models). At the beginning, however, parametric design schema need to transform concepts into data. Compared with traditional paper-based design, parametric schema emphasizes systematic design process and the geometry logic of form.

The organic management of data is the key for designers to form designs by programming technology. That is: data flow driven form, data structure organize form. All of this is accomplished by data transfer between the various functions written in a scripting language. In computers, form is represented by data. Designers achieve design goals through ‘design’ code while the code records the thought process of the designer. The design process of the human brain is no longer a black-box operation.

Therefore, the parametric thinking schema can be understood as a process of integrating abstract thinking, mathematical thinking and algorithm thinking in order to realize the design goal of organic automation under the guidance of design knowledge.

3.2 Parametric design process

Whether there is a specific thinking process in parametric design, there is no definite conclusion yet[12]. However, no matter what type of design has its own general design process. Figure 2 is a typical parametric design thinking process. First of all, consider concepts and ideas based on customer needs and design tasks, and present them in sketches, text, or data. Secondly, designers constantly abstract and digitizes ideas and concepts, explore the geometric logic rules, and include evaluation rules[13]. Gradually, designer establish an associative set of algorithmic rules and data structure. In the third step, the designer must standardize all parameters, including the definition, coding, and naming of the parameters, and further clarify the relationship between the design input and the design response (design output)[3]. Further, designers need to consider the internal data stream transfer parameters to form a control parameter of the organic system. Step Four: designer writes the script code in the script environment according to the algorithms, rules and data structure, runs and debugs the program, and checks the design response. Finally, the designer evaluates the output and communicates with the customer to provide timely feedback. However, the design is not a linear process, especially the process of parameterization. Modifying rules, algorithms, and code is a regular job.

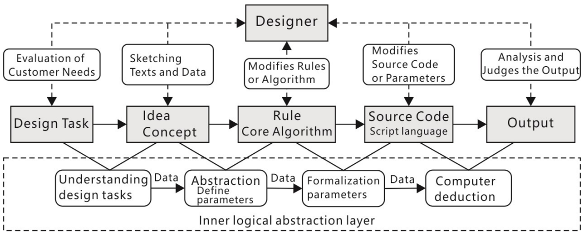


Figure 2. The process of parametric design.

3.3 Parametric design tool and scripts

The core work of parametric design is algorithm research and programming. The design of parametric digital form by programming technology requires graphic design platform and script environment. For example, CorelDraw is a graphic design platform, VBA can be used as a scripting language. However, this is not a secondary development in the general sense, which is a branch or direction of design method and design technology development. In script environment, form is driven by algorithm and code. Program statements and the definition of functions and classes can express the form logic construction process[14].

Grasshopper is the most popular parametric design tool, and is a visual scripting language. It has been built into Rhino6.0. Grasshopper itself supports multiple scripting languages such as Python, C#, and VB. It is often used in architecture design, industrial design, jewelry and so on. Illustrator supports multiple scripting environments (VB, JavaScript, AppleScript). Processing is an extension of the Java language, which can be used for visual communication design, interactive media art design, and information visualization[15], such as the logo designed by the Danish design studio NR2154 for United Nations Climate Change Conference in Copenhagen. Mathematica and Matlab can be used for the calculation of parametric form design[16]. The features of parametric tools and scripts are listed in Table 1.

Table 1. Tools and scripts for parametric design.

Tools and scripts	Language type/ Futures	Scope of application
Rhino / Grasshopper /Python[14]	Visual scripting code / Code and model in parallel	Architecture/ Product etc.
CorelDRAW / VBA[17]	Macro scripting language	graphic design
Illustrator/ JavaScrip[18]	scripting language	graphic design
Processing (JavaScrip)[15]	Graphic design language	Interactive/Visual Art/Information Visualization
Mathematica (wolfram)[19]	symbolic language	2D/3D design calculation

The parametric algorithmic schema has been completely different from the traditional paper-based design schema. However, this does not mean that the designer must become a programmer, or that the programmer can replace the work of designer. In the process of parametric design, therefore, designers need to constantly balance programming techniques and the understanding of design knowledge [11].

4 The method of parametric logo design

As a typical form, the logo is the core of enterprise brand image. A good logo is both attractive and meaningful. It is an important bridge between the public and the company. The core steps of the parametric logo design based on the natural mode are as follows.

Firstly, designers must have a deep understanding of the culture, market environment and core values of the enterprise, and extract core demand and key visual elements [20].

Secondly, we need to find natural patterns that match the design requirements. The natural patterns that match the enterprise image can be crossed and merged in several natural patterns. For instance, symmetry and rotation can be fused with spiral pattern. After the match is successful, the visual constituent laws and associated logic of the selected natural patterns can be studied, especially the geometric logic relations. At this point, designer can initially define the parameter set of the logo form: $\{s_1, s_2, s_3, \ldots, s_n\}$.

Then, according to the geometrical logic relation and the form association logic, designers can construct the transformation rule of the form, and transform all the rules into the transform algorithm of the logo form. On the other hand, the overall image of the logo is also affected by other factors, such as proportion, location, text, etc. This requires the construction of some additional rules for systematic coordination. In addition, designers should define color control parameter set: $\{c_1, c_2, c_3, \ldots, c_n\}$ and color transformation rules based on the color design scheme of the logo and color space. In order to facilitate the color transformation and algorithm implementation, it is necessary to map the value range of the color space to the interval of [0, 1]. Sometimes, in order to show better results, it is necessary to associate the color parameters of the logo with the form parameters of the logo.

Finally, according to design requirements, we need to define the font control parameter set: $\{f_1, f_2, f_3, \ldots, f_n\}$ and font control rules.

The process of studying natural patterns is also the process of studying algorithms. Parametric form design is a process of parametric experiment. It constantly adjusts rules and modifies parameters to satisfy the visual demand of the customer based on design responses.

5 Case study

The design task from Gansu ZBloom Culture Media Co. Ltd. and its core philosophy: the culture spread to the hearts of the people, the spirit of sustainable artisans. The spiral pattern is radially scattered from the center to the outside, implying the infinite transmission of the energy of cultural media, which coincided with the core philosophy of ZBloom Cultural Media. This case, therefore, applied a logarithmic spiral pattern to the parametric form design.

5.1 Logarithmic spiral pattern

Logarithmic spiral, also known as isometric spiral or growth spiral, is a self-similar spiral curve and it is a common pattern in nature. Logarithmic spiral was first described by Descartes and later extensively investigated by Jacob Bernoulli, who called it 'the marvelous spiral'[5]. The Fibonacci spiral also known as the 'golden spiral' in geometry. It is a special case of logarithmic spiral and is also the most perfect classic gold proportion in nature. The logo design of this case used a logarithmic spiral pattern which comes from Nautilus and seed head of sunflower (Figure 3).

The head of sunflower is spiral. Its seeds are Fibonacci spiral arrangement, and there are two groups of spiral in the opposite direction. This arrangement is described by any spiral, wherein the spiral of Archimedes easiest. The definition of Archimedes spiral line: r(θ)=aθ (polar equation). Generally, there are 34 to 55 spiral lines from inside to outside. 34 and 55 are two consecutive numbers in the Fibonacci sequence.

Figure 3. The classical logarithmic spiral in nature: Seed head of sunflower and Nautilus (Wikipedia)

The polar coordinate definition of logarithmic spiral is

$$r(\theta) = ae^{b}\theta \quad (1)$$

where $a, b \in R$, and $a \neq 0$. In order to facilitate the implementation of the algorithm, the formula (1) must be converted into a parametric equation as shown in formula (2).

$$\begin{cases} x(t) = r(t)cos(t) = ae^{bt}cos(t) \\ y(t) = r(t)sin(t) = ae^{bt}sin(t) \end{cases} \quad (2)$$

5.2 The control rules of form and color

The parametric design of the logo includes three aspects: parametric form, parametric color, and parametric text. In this case, we used disks instead of sunflower seeds for parametric logo design.

The sine function was selected as the radius transformation function of disks. The radius changes with the number of disks, but the radius of the disk cannot be negative. So we defined the disk radius transformation function as shown in formula (3).

$$\text{Disks\_Radius}(n) = \begin{cases} a\sin(bn) + c & \text{if } n \geq 0 \\ 0 & \text{if } n < 0 \end{cases} \quad (3)$$

where a, b and c are positive real numbers, and a ≤ c. The independent variable n is the number of disks, $n \in [0,+\infty)$. The function image is shown in Figure 4.

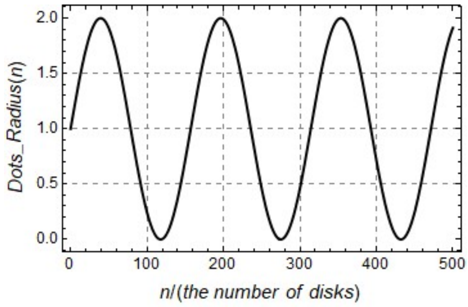


Figure 4. Function image of the disks radius transform.

RGB, HLV, CMKY and other color spaces can be parameterized by logo colour. The value of the color space was converted to [0, 1] according to the script environment and programming needs. Therefore, the color transformation function can be defined by the hyperbolic tangent function, which is the quotient of the hyperbolic sine and the hyperbolic cosine. Its definition process is as follows:

$$\sinh(x) = (e^{x}-e^{-x})/2 \quad (4)$$

$$\cosh(x) = (e^{x}+e^{-x})/2 \quad (5)$$

$$\tanh(x)=\sinh(x)/\cosh(x) = (e^{x}-e^{-x})/(e^{x}+e^{-x}) \quad (6)$$

From the formula (6), we have $x \in (-\infty,+\infty)$ and $\tanh(x) \in (-1,+1)$. But the color values cannot be negative. Therefore, we defined the following function:

$$\text{Color\_Transform}(n) = \begin{cases} \tanh(kn) & \text{if } n \geq 0 \\ 0 & \text{if } n < 0 \end{cases} \quad (7)$$

where $k \in R^{+}$, $n \in [0,+\infty)$, and the independent variable n is the number of disks. The value of constant k determines the quality of the color transformation. The process of the color change is shown in figure 5.

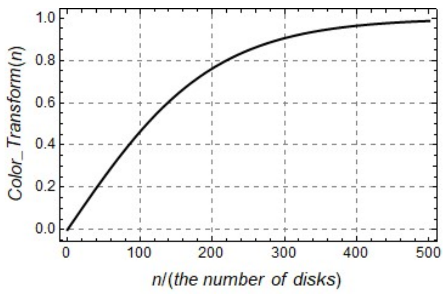


Figure 5. Function image of color transformation.

The combination of the blending colour algorithm and the color transformation function can present multicolour nonlinear gradient of the color design. In this case, the mixture of the three colors is defined as the color scheme of the logo.

5.3 Core algorithm

Firstly, the logarithmic spiral arrangement of disks is implemented according to formula (2), where the parameters are the radius and the angle. [0, a + c] is the domain of the radius changes of disks, where a and c are constants in equation (3). [0°, 360°] is the domain of angle changes. Secondly, according to the formula (3) to achieve the control algorithm of the change of disks radius, the range of the number of disks is [0, N], where N is a constant. The constant N is determined according to the design response and can be modified at any time. Thirdly, three color nonlinear gradient design is achieved by the blending colour algorithm and the formula (7). The color change of the logo with the number of disks n change. Finally, we can achieve the function of calling and selecting fonts by calling the standard font of the Window system and the Font library of the script environment.

5.4 The generator of parametric logo design

This article developed a generator of parametric logo design (Figure 6) with Wolfram language in mathematica for Gansu ZBloom Cultural Media. The left side of the logo generator interface is the parameter control bar of the logo, and the right side is the logo browser. The form parameters of the logo includes four parameters: the angle, the radius of disks, the number of disks, and the size of the logo image.

Figure 6. The generator of the parametric logo design.

Design results in Figure 7 selected out by the designers, which were selected in the first round. They were designed by the logo generator under different parameter values. The parameter values are shown in Table 2. Figure 7(c) was selected as the logo for the ZBloom Cultural Media judged by a evaluation group.

(a) **(b)** **(c)** **(d)**

Figure 7. The forms of logo in different parameters.

Table 2. Parameters of the logo form

Para meter s	angle	Min radius Max radius	Number of disks	Image size	Color value (RGB)
(a)	2.3	0.1 1.75	271	390^2	(255, 0, 50) (230, 255, 30) (20, 128, 155)
(b)	119	0.31 3.11	261	272^2	(15, 128, 253) (250, 255, 15) (255, 0, 50)
(c)	178.33	0.25 5.41	164	374^2	(162, 0, 0) (255, 255, 23) (10, 150, 210)
(d)	195	0.6 3.5	165	348^2	(0, 128, 255) (255, 255, 10) (240, 20, 45)

6 Conclusions

Natural patterns has provided an unlimited source of design and can be effectively applied to form design. Parametric design can effectively enrich the content of form design. Finding the mathematical logic behind natural patterns has been a key step. The combination of natural patterns and parametric design technology will open up new design schema and methods for form design. This has been a fusion of aesthetics and technology. Parametric languages and scripting techniques will have an unprecedented impact on design thinking. Designers has been required to constantly update their knowledge structure. Parametric design has required designers to respect the mathematical and logical laws of the scripting environment. It has been difficult for designers to anticipate design results in a scripting environment. The combination of natural patterns and parametric thinking has prompted the form design becoming an generative system of organic closed-loop from design requirements to design results.

The next work of this article will study the parametric design method of complex product form based on natural patterns.

Acknowledgments

The project is sponsored by National Natural Science Foundation of China (51465037) and Hongliu Outstanding Talent Development Program of Lanzhou University of Technology (201406). In addition, Zbloom Culture Media also provided support for this project.

References

1. WJ Yue. The natural theorem of art and design. Shanghai Art & Crafts, 1, 82-83 (2011)

2. J Chen. Research and practice of product design based on parameterization. Art Journal, 3, 111-115 (2015)
3. WG Xu. Parametric design and algorithm generating. World Architecture, 6, 110-111 (2011)
4. DL Ma. *A study on Fibonacci's Liber Abaci* (PhD thesis). (ShangHai Jiao Tong University, Shanghai, 2009)
5. S Hildebrandt, A.J Tromba, S Sheng. *The Parsimonius Universe*. (ShangHai Education Press, Shanghai, 2004)
6. M Macnab, WB Fan. *Design by nature: using universal forms and principles in design* (China Machine Press, Beijing, 2013)
7. AM Turing. The chemical basis of morphogenesis. 1953. Bull Math Biol, **52**,153-197(1990)
8. A Lindenmayer. Mathematical models for cellular interactions in development. Journal of Theoretical Biology, **18**, 3, 280-315 (1968)
9. CRB Hernandez. Thinking parametric design: introducing parametric Gaudi. Design Studies, **27**, 3, 309-324 (2006)
10. R Woodbury. *Elements of Parametric Design* (Routledge, New York, 2010)
11. R Oxman, N Gu. Theories and Models of Parametric Design Thinking. eCAADe 33, **2**, 477-482 (2015)
12. R Oxman. Thinking difference: Theories and models of parametric design thinking. Design Studies, **52**, 4-39 (2017)
13. YC Liu. Parametric Design: Method, Thinking Strategy and Framework. Architecture Technique, Z1, 34-37 (2011)
14. RQ Bao. *Parametric Logic Design Process* (Phoenix Science Press, Nanjing , 2015)
15. L Tan. Application Research of Code Art Based on Processing. Art & Design, **2**, 104-105 (2012)
16. JN Su, JH Wu, Y Liu, QW Zhang. Research on 3D Product Shape Merging Technology Based on Spherical Harmonic Map. Journal of Graphics, **35**, 3, 417-422 (2014)
17. XJ Liu, YY Wu. The Research on Rapid Design of Fan Products by Computer-Aided Intelligence. Packaging Engineering, 10, 42-45+56 (2013)
18. M O'Neill, A Brabazon. Evolving a logo design using Lindenmayer systems, Postscript & Grammatical Evolution. Evolutionary Computation. IEEE, 3788-3794 (2008)
19. S Wolfram. *An Elementary Introduction to the Wolfram Language* (Wolfram Media Inc, Champaign, 2016)
20. DQ Jin. *The Brand Design 100+1* (Peking University Press, Beijing, 2017)

证书号第4293403号

发明专利证书

发明名称：基于自然模式的Logo生成方法、智能Logo生成器

发　明　人：苏建宁;李雄;唐钊山

专　利　号：ZL 2018 1 0343073.9

专利申请日：2018年04月17日

专利权人：兰州理工大学

地　　址：730050 甘肃省兰州市七里河区兰工坪路287号

授权公告日：2021年03月12日　　　　授权公告号：CN 108549530 B

国家知识产权局依照中华人民共和国专利法进行审查，决定授予专利权，颁发发明专利证书并在专利登记簿上予以登记。专利权自授权公告之日起生效。专利权期限为二十年，自申请日起算。

专利证书记载专利权登记时的法律状况。专利权的转移、质押、无效、终止、恢复和专利权人的姓名或名称、国籍、地址变更等事项记载在专利登记簿上。

局长
申长雨

国家知识产权局

2021年03月12日

其他事项参见续页

证书号第4293403号

国家税务总局北京市海淀区税务局
印花税
代扣专用章
国家知识产权局

专利权人应当依照专利法及其实施细则规定缴纳年费。本专利的年费应当在每年 04 月 17 日前缴纳。未按照规定缴纳年费的，专利权自应当缴纳年费期满之日起终止。

申请日时本专利记载的申请人、发明人信息如下:

申请人:

兰州理工大学

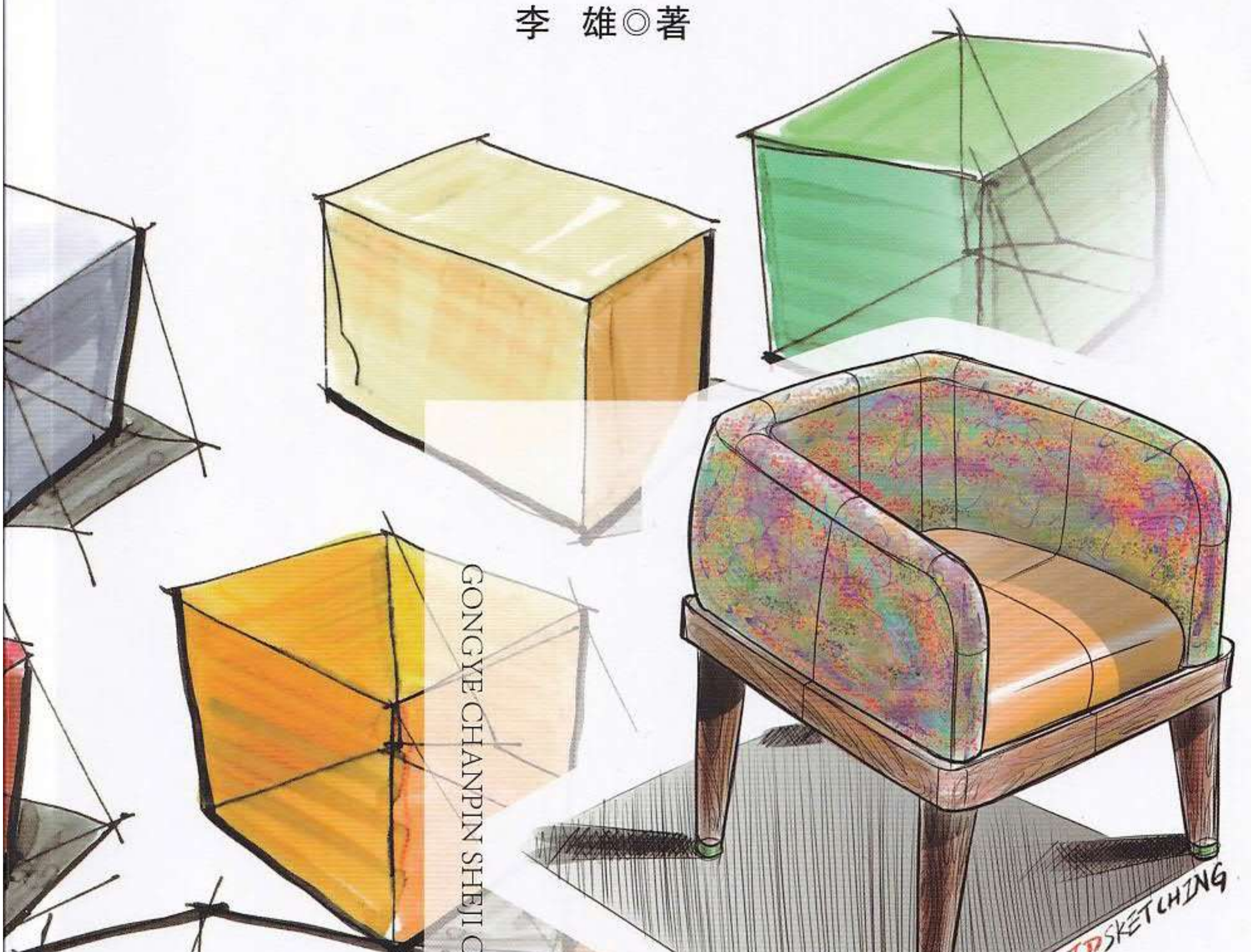
发明人:

苏建宁;李雄;唐钊山

工业产品设计草图

李 雄◎著

GONGYE CHANPIN SHEJI CAOTU



中国铁道出版社有限公司
CHINA RAILWAY PUBLISHING HOUSE CO., LTD.

内 容 简 介

本书讲述工业设计手绘，针对无任何绘画基础的理工科工业设计本科生，注重草图逻辑训练，轻画面效果训练。首先，帮助初学者建立设计手绘视觉基础（物与像、光与影）；接着，渐进到设计绘图的基本方法和产品形态构建技巧，以及数字草图的绘制技法（SketchBook、Krita），提倡在一定的基础上应加速向数字草图推进，初步探讨了人工智能辅助产品概念设计的思想、方法和技术；然后，通过基本色彩和材质的表现训练，延伸到效果图的绘制，包括传统效果图和数字效果图；最后，产品说明图则包括产品爆炸图、立体剖视图、流程板、场景图等绘制方法。另外，本书还特别探讨人体结构与产品设计，涵盖了基本的人体解剖学结构、绘制人体结构草图和基于人体结构的运动鞋草图设计。每个部分均给出具体的绘图方法和绘图步骤。

本书适合高等院校工业设计专业学生和工业产品设计师阅读参考，以及设计手绘爱好者交流与借鉴。另外，本书部分内容具有研究性质，可以作为工业设计研究生的参考资料。

图书在版编目（CIP）数据

工业产品设计草图 / 李雄著 . —北京：中国铁道出版社有限公司，2020. 12

ISBN 978-7-113-27473-3

Ⅰ. ①工… Ⅱ. ①李… Ⅲ. ①工业产品 - 产品设计 - 绘画技法 Ⅳ. ① TB472

中国版本图书馆 CIP 数据核字（2020）第 245103 号

书　　名：工业产品设计草图
作　　者：李　雄

策划编辑：潘晨曦　　　　**编辑部电话**：（010）63549458
责任编辑：祁　云　李学敏
封面设计：郑春鹏
责任校对：孙　玫
责任印制：樊启鹏

出版发行：中国铁道出版社有限公司（100054，北京市西城区右安门西街 8 号）
网　　址：http://www.tdpress.com/51eds/
印　　刷：三河市兴达印务有限公司
版　　次：2020 年 12 月第 1 版　2020 年 12 月第 1 次印刷
开　　本：787 mm×1 092 mm　1/16　印张：23　字数：540 千
书　　号：ISBN 978-7-113-27473-3
定　　价：82.00 元

关于本书

内容提要

本书讲述工业设计手绘，针对无任何绘画基础的理工科工业设计专业本科生，注重草图逻辑训练。首先，帮助初学者建立设计手绘视觉基础（物与像、光与影）；接着，渐进到设计绘图的基本方法和产品形态构建技巧，以及数字草图的绘制技法（SketchBook，Krita），提倡在一定的基础上应加速向数字草图推进，初步探讨了人工智能辅助产品概念设计的思想、方法和技术；然后，通过基本色彩和材质的表现训练，延伸到效果图的绘制，包括传统效果图和数字效果图；最后，产品说明图则包括产品爆炸图、立体剖视图、流程板、场景图等绘制方法。另外，本书还特别探讨了人体结构与产品设计，涵盖了基本的人体解剖学结构、绘制人体结构草图和基于人体结构的运动鞋草图设计。每个部分均给出具体的绘图方法和绘图步骤。

项目支持

本书受兰州城市学院青年教师项目（校级）基金的支持：项目编号LZCU–QN2017–28。

读者对象

本书适合作为高等院校工业设计专业学生和工业产品设计师的参考书，以及设计手绘爱好者交流与借鉴用书。另外，本书部分内容具有研究性质，可以作为工业设计研究生的参考资料。

本书结构

本书分为两大部分，即草图基础与绘图方法（第 1 章 ~ 第 6 章），效果图与说明图（第 7 章 ~ 第 10 章），具体安排如下：

第一部分　草图基础与绘图方法

第 1 章　设计草图——手 & 脑 & 心

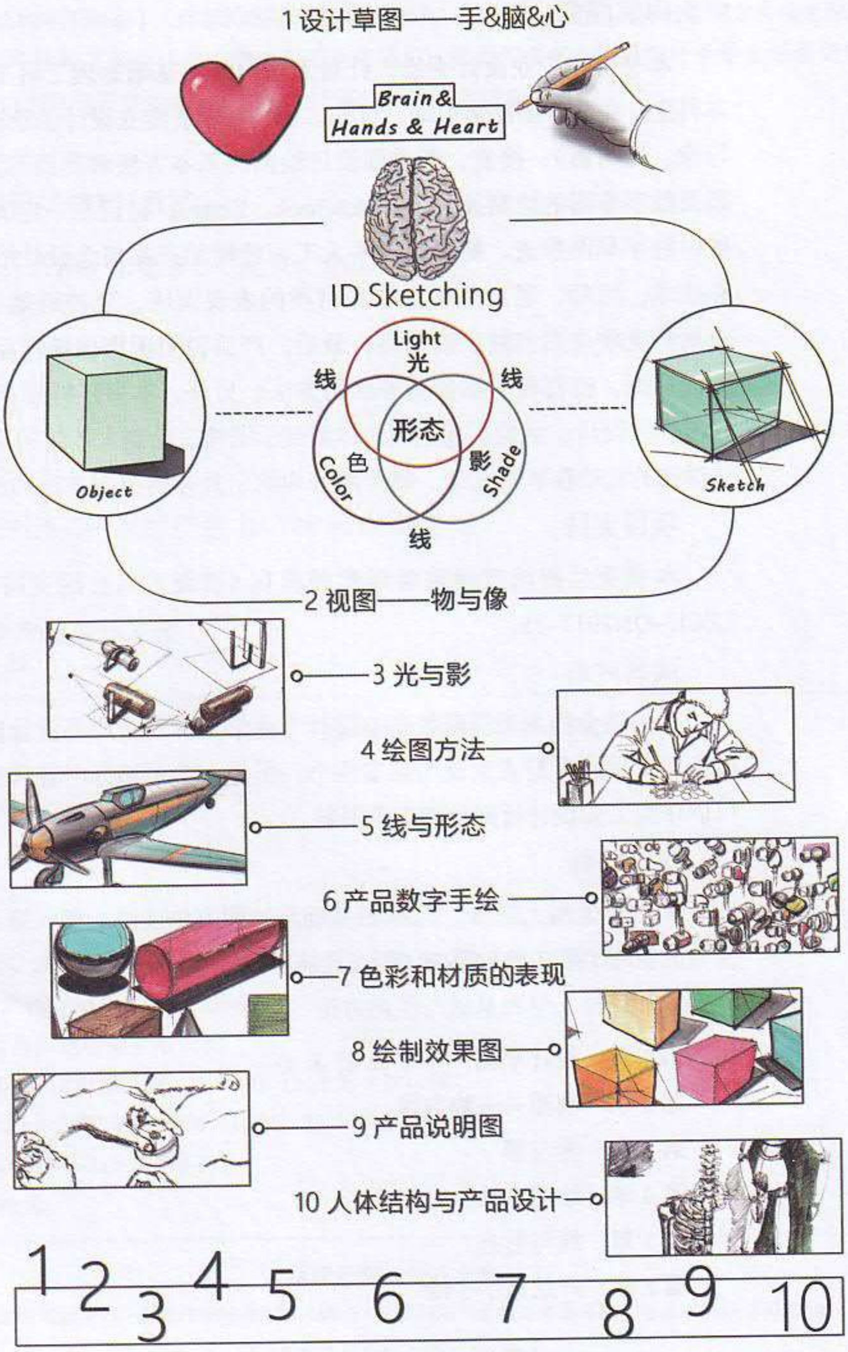
第 2 章　视图——物与像

第 3 章　光与影

第 4 章　绘图方法

第 5 章　线与形态

第 6 章　产品数字手绘


1 设计草图——手&脑&心
Brain &
Hands & Heart
ID Sketching
Light
光
线
线
形态
色
Color
影
Shade
线
Object
Sketch
2 视图——物与像
3 光与影
4 绘图方法
5 线与形态
6 产品数字手绘
7 色彩和材质的表现
8 绘制效果图
9 产品说明图
10 人体结构与产品设计
1 2 3 4 5 6 7 8 9 10

前 言

这本书并不是一本产品设计草图作品集。

我希望带给读者新的视角，特别是理工类工业设计专业本科生。设计草图并非绘画，也并非工程技术制图，它是设计师应该掌握的视觉语言。设计草图重在记录、构思、创造，它可以呈现对问题的观察、分析、归纳、联想、创造、评价六个维度连续性思考的成果。但请注意，设计草图不是第一步也不是最后一步。

书中不仅包含了近期绘制的设计方案草图，还包括专门为本书绘制的大量示范性的草图，也收录了以前参与过的一些项目中的设计草图，并邀请朋友或其他设计师提供一些实践案例草图。

在书中的设计草图最终定稿之前，我已经淘汰了许多设计草图。我始终希望把最能说明问题的草图呈现给读者，但做到这一点并非易事，只能尽力为之。

书中的每一张图都给出了简要解释和说明，这些图都凝聚着我对工业产品设计草图及草图思维的理解，但这仅仅是我个人的观点，难免出现偏差、疏漏及不足，敬请广大读者批评指正，不吝赐教。

接下来，就请大家翻开这本书，跟随我的视角来阅读它。当然你不一定要按顺序来阅读，但无论怎样，阅读是一方面，动手练习则必不可少，空闲时间的思考也是重要的。真心希望这本书能够帮助大家了解产品设计草图究竟是什么，并对产品设计草图产生兴趣，重要的是应用它帮助你思考问题、解决问题。

作　者

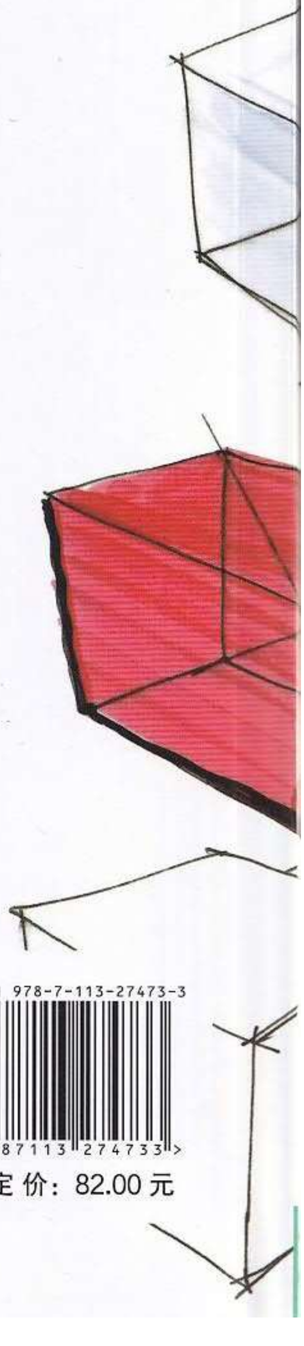
2020 年 6 月

目 录

责任编辑：祁　云　李学敏
封面设计：郑春鹏

设计草图不是第一步，也不是最后一步

Design Sketching is not the first or the last step



中国铁道出版社有限公司
CHINA RAILWAY PUBLISHING HOUSE CO., LTD.
地址：北京市西城区右安门西街8号
邮编：100054
网址：http://www.tdpress.com/51eds/

ISBN 978-7-113-27473-3

定价：82.00元