Construction of computer-aided product innovation design system based on AIGC

Menghao Yu*
School of Art and Design, Zhengzhou University of Light Industry, Zhengzhou, China

Abstract. With the development of AI technology, AIGC applications represented by ChatGPT are continually emerging, offering designers brand-new design tools. How designers should flexibly utilize the advantages of AIGC to expand creative boundaries, has become a crucial topic in exploring the future development path of design. To this end, this paper proposes a computer-aided design system (CADS) based on AIGC, and verifies the feasibility of CADS through the case of a center console in electric vehicles. In CADS, we regard the latest achievements of AIGC technology as beneficial tools to aid innovation. Simultaneously, to enhance the originality and depth of content generated by AIGC, we use sentiment analysis based on online review data to mine users' emotional tendencies, and use them as the emotional source for AIGC-generated content. Then, through training ChatGPT, we obtained Prompts in a professional format, which was input into Stable Diffusion for the text-to-image task. At this stage, we conducted three rounds of large-scale image generation. Following fine-tuning via the image-to-image function, we ultimately obtained high-quality product rendering images. For the first time, the research combines AIGC with product design from a practical application standpoint, laying a solid foundation for future product design practices based on AIGC.

Keywords: Industrial Design, Artificial Intelligence, Emotional Mining, ChatGPT, Stable Diffusion

1. Introduction

With the iterative development of artificial intelligence (AI) technology and deep learning models, the era of AI 2.0 is gradually unfolding, and one of the most significant hallmarks of this era is the breakthrough progress in Artificial Intelligence Generated Content (AIGC) [1]. As one of the most prominent cutting-edge technologies of the present time, AIGC can generate corresponding text and image content based on user preferences. Additionally, the continuous iterative innovation of pre-trained models and algorithms further enhances AIGC's generality and foundational capabilities. This holds milestone significance for the development of both human society and artificial intelligence [2].

Today, the application of AIGC technologies represented by ChatGPT is continually emerging, bringing disruptive impacts to human production and lifestyles [3,4]. Confronting the increasingly powerful AIGC, how to correctly view the creativity of AI and its potential societal roles has sparked intense discussions across various sectors of society [5]. Researchers like Fortuna [6], Modlinski [7] have explored the principles and efficacy of AI from aspects like intelligent industries, smart machines, and intelligent production. While their methods and themes may vary, the central question they repeatedly discuss remains the impact of AI's level of intelligence on humanity [5]. The results indicate that the integration of intelligent machines and humans has never ceased since the birth of AI. Concurrently, the presence and collaboration of AI in societal sectors and industries is also an unstoppable trend [8]. Hence, the current development of AIGC is akin to a "double-edged sword." How to better address the industry changes brought about by AIGC, responsibly utilize the achievements of technological advancement, bravely embrace the baptism of new technological waves, seize new opportunities in a timely manner, and ush-

*Corresponding author. Menghao Yu, School of Art and Design, Zhengzhou University of Light Industry, Zhengzhou, 450000, China. E-mail: 2665685823@qq.com.

er in a new round of industrial transformation and development, has become an urgent issue demanding attention.

In the field of innovative design, the emergence and development of AIGC have brought entirely new creative tools, thought paradigms, and design methodologies [9]. As a series of AIGC-generated artworks continue to fetch high prices at auctions, the incredible creative potential inherent in the content produced by AIGC becomes further evident [10]. Creativity is a vital element in the design industry and with the continuous development and application of AI in creative design, contemplation of AI's creativity has prompted a reconsideration of the essential characteristics of human creativity and the creative process. Undoubtedly, AI has successfully emulated certain aspects of this process [8]. Despite some debate surrounding the creative aspects between artificial art and AI-generated art, the majority of creative professional view AIGC as a tool to enhance their work rather than a threat [11]. According to a survey by Adobe [12], 48% of creative professionals have used ChatGPT in their work, and 25% have used Midjourney. They envision using AIGC to explore new design mediums, create higher-quality works, and distinguish their creations.

Similarly, in industrial design, traditional industrial designers translate innovative design ideas into conceptual product designs through creative thinking, while the emergence of AIGC provides a richer set of creative tools and inspirational resources for this transformation process [13]. On one hand, AIGC can serve as an extension and enhancement of designers' creativity, helping them break through traditional design boundaries and inspire more unique and innovative design solutions. On the other hand, AIGC can also serve as an imaginative creative partner for designers, and through its unlimited creative potential, it not only imparts new creative impetus to designers but also leads new waves of innovation in the design field. Therefore, in this era of rapid AI development, how designers should flexibly leverage the advantages of AIGC to expand creative boundaries and timely embrace the latest technological achievements has become an important issue in exploring the future path of design development [14]. This study utilizes the latest AIGC technological advancements, and treats them as beneficial tools for innovative design, to explore methods and pathways for AIGC's application in product innovation design. Through emotional mining, it integrates human emotions, concepts, and uniqueness into the AI design process, ensuring originality and depth of creativity. The goal is to construct a scientific computer-aided design system (CADS) that facilitates collaboration between AIGC and designers, thereby establishing a creative space filled with boundless possibilities.

The remainder of this paper is organized as follows: Section 2 reviews the concept of sentiment analysis the development of the AIGC. Section 3 shows the details and process of the proposed CADS. Case study is applied in Section 4 and the results and discussion of this study is summarized in Section 5. Section 6 gives the conclusions of this paper.

2. Literature review

2.1. Sentiment mining based on online review data

In order to ensure the originality and depth of AIGC-generated content, the research endeavors to establish designers as pivotal actors throughout the CADS process. Adopting the perspective of designers, human emotions are infused into the content generated by AIGC. In the realm of exploring and perceiving human emotions, Kansei engineering stands as one of the most crucial theoretical foundations [15]. In the domain of traditional Kansei engineering, the mining of user emotions usually involves collecting affective words through sources such as relevant magazines and literature, followed by evaluating user emotional attitudes through methods like surveys. However, such approaches result in collected emotional terms that lack specificity and exhibit a high degree of subjectivity in the evaluation process [16]. Today, with the advancement and ubiquity of online information technology, online review mining relying on online shopping platforms and big data has gradually become a focal point in relevant research [17, 18, 19]. A vast number of internet users contribute abundant information resources through online platforms daily, granting online comment data a characteristic of extensive coverage and significant quantity. This data can promptly reflect users' genuine needs. Consequently, utilizing online comments as a source of data for sentiment analysis can substantially aid businesses and designers in obtaining a more comprehensive understanding of users' intricate emotions.

Online review data, acting as a treasure trove of immense user knowledge, encompasses a wealth of precise and authentic user sentiment expressions. However, this data's vast quantity and inclusion of non-standardized and unstructured textual representations pose a complex challenge in extracting valuable information from it [20]. With the advancement of natural language processing techniques, methods and tools for processing online review data have become increasingly diverse and refined. Among them, sentiment analysis is a methodology that can identify and categorize users' emotional tendencies from text, thereby aiding decision-makers in gaining insight into emotional information [21]. Current sentiment analysis primarily falls into two categories. The first category is lexicon-based sentiment analysis [22], which involves matching words in the text with words from a pre-constructed sentiment lexicon. The polarity of these matched sentiment words is then used to determine the corresponding content's emotional tendency. The second category is machine learning-based sentiment analysis [23], which employs machine learning algorithms to automatically learn emotional information from text data and categorize it. These two methods of sentiment analysis have their own advantages and limitations in practical applications. In lexicon-based sentiment analysis, designers can create specialized sentiment lexicons according to specific domain requirements, thereby enhancing adaptability and accuracy [24]. Moreover, during the process of determining emotional tendencies based on a lexicon, designers can clearly identify factors that play a significant role in the judgment of emotional tendencies, making the results more comprehensible and interpretable [25]. On the other hand, in machine learning-based sentiment analysis, designers rely more on pre-trained machine models and lack some necessary human intervention [26]. Given that the primary purpose of using sentiment analysis in this research is to use the results as emotional input for designers regarding AIGC, choosing lexicon-based sentiment analysis aligns more closely with the research's task requirements.

2.2. Development and application of AIGC

AIGC primarily refers to a new content generation method that relies on artificial intelligence technologies such as Generative Adversarial Networks (GAN) and large-scale pretrained models. It can automatically generate text, images, audio, and video content based on specific input data [27]. Compared to traditional content generation methods like Professional Generated Content (PGC) and User Generated Content (UGC), AIGC has greatly improved efficiency and quality in content generation, gradually becoming one of the primary forms of content in the Web 3.0 era [28]. In fact, the history of AIGC can be traced back to earlier times when early AIGC involved AI tools assisting in generating fixed-template content. It was mainly applied in professional tasks such as film and entertainment, industrial modeling, and more [29]. Subsequently, with the advancement of AI technology, industry-driven AIGC experienced explosive growth [30]. During this process, there have been some representative technological revolutions. One of them was the introduction of GAN in 2014 by Goodfellow [31]. GAN employs a method of adversarial training between the generator and the discriminator, and generates new content through their adversarial game and continuous iterations. GAN models gradually became the mainstream models in generative machine learning and spawned modified models like Deep Convolutional GAN (DCGAN), Conditional GAN (cGAN), InfoGAN, etc., which accelerated convergence and enhanced model interpretability [32]. In addition, in the same year, the introduction of two major generative models, Variational Auto Encoder (VAE) proposed by Kingma et al. [33] and flow-based models proposed by OpenAI [34], also played a significant role in advancing generative machine learning. In 2021, OpenAI introduced the Contrastive Language-Image Pre-Training (CLIP) model [35]. CLIP employs text and image encoders to separately learn features from text and images. It utilizes multi-modal embedding space contrastive learning to transform image classification tasks into text-image matching tasks. By utilizing unsupervised text information as a supervisory signal, CLIP can effectively learn visual features, enabling efficient multi-modal recognition, fusion, and transformation [36]. In 2022, the popularity of diffusion models [37] once again propelled technological advancements and content innovation in the field of AIGC. These models efficiently achieve image-text generation through forward diffusion processes and reverse generation processes, gradually becoming a hot research focus in the current AIGC landscape. It can be foreseen that with the accelerated development of artificial intelligence algorithms, computing power, and data, AIGC technology applications represented by AI art creation will become a significant trend in the development of online information resources in the digital intelligent environment.

Text generation. In November 2022, OpenAI released ChatGPT, a natural language processing technology application developed based on a generative pre-training model [39]. Within a week of its release, ChatGPT had already gained one million users. The application of ChatGPT aims to generate naturally fluent text through extensive pretraining and fine-tuning, and it is capable of engaging in conversational interactions. The publicly released ChatGPT is a member of the GPT-3 series of models. GPT-3 is an immensely powerful pre-trained language model with 175 billion parameters [40]. As of March 2023, this model has been iteratively updated to GPT-4. Although the specific details of GPT-4's parameters have not been publicly disclosed, it is reasonable to anticipate significant increases in model capacity and pretrained data volume based on ChatGPT 4.0's strong performance in natural language processing [41]. The larger training data and parameters of ChatGPT 4.0 enable it to generate even more complex and higher-quality text. It can accept and recognize more intricate forms of input, including images and text, making it suitable for a broader range of application scenarios [42]. The emergence of ChatGPT has triggered explosive growth in the AIGC industry, bringing this specialized term into the public eye. It has swiftly led to a global rush among major technology companies to enter and explore applications in scenarios similar to ChatGPT. As a representative work of the latest wave of AIGC, ChatGPT 4.0 remains one of the most powerful natural language processing tools available today [43].

Image generation. Since the emergence of generative models like GAN and CLIP, text-to-image synthesis applications using deep learning have made significant progress. Currently, there are three mainstream text-to-image generation systems: Midjourney [44], Dall-E2[45], and Stable Diffusion [46]. Among them, Midjourney is created by an independent research lab with the same name and is primarily available to users through the Discord community. The operation is relatively simple – users only need to input the prompt they want to depict in the chat box, and AI will automatically generate images based on the text description. Midjourney tends to produce surrealistic images, making it quite popular among artists [47]. DALL-E and DALL-E 2 are text-image models developed by OpenAI. The technical principle of DALL-E is mainly based on GPT-3 and GAN models for image generation. DALL-E 2 is the successor to DALL-E and is primarily trained on approximately 650 million pairs of images and text collected from the internet. Its goal is to generate more realistic images at higher resolutions [48]. Stable Diffusion, released by Stability AI in 2022, utilizes a deep learning technique called Latent Diffusion to generate images based on textual descriptions [49]. Unlike some other text-to-image models such as Midjourney and Dall-E, Stable Diffusion's code and model weights are completely open-source and users can deploy it locally on hardware that meets the requirements. In addition to text prompts, Stable Diffusion's text-to-image generation script allows users to input various parameters like sampling type, output image size, and seed values. Users can also train their own personalized Lora models to shape the generated images according to their desired style. Moreover, the open-source community has contributed powerful plugins like ControlNet and Tagger, greatly expanding the functionality of Stable Diffusion. It's evident that compared to mainstream Text-to-Image models like Midjourney and Dall-E, Stable Diffusion offers greater controllability and tunability in generating corresponding image content [50]. Therefore, this research also considers SD as the preferred tool for the Text-to-Image process.

Audio and Video generation. The applications of AIGC in audio generation include text-to-speech synthesis and voice cloning [51]. In video generation, it encompasses video editing processes to create trailers and promotional videos [52]. However, since the relevant technological applications are not within the scope of this research, they will not be discussed in detail.

3. CADS for collaborative designers and AIGC

3.1. System framework

The research aims to construct a scientifically effective CADS, enabling designers and AIGC to collaborate intensively during the design process, creating a creative space full of infinite possibilities. At the same time, to effectively mitigate the randomness and unpredictability of AI-generated content, and enhance its originality and depth, The research attempts to inject human emotions into AIGC by designers, allowing designers to play a significant role throughout the CADS. To achieve this, the research uses sentiment analysis based on online comment data to mine users' emotional tendencies, using them as the emotional source for AI-generated content. At the same time, it selects ChatGPT and Stable Diffusion, which are currently mainstream in AIGC applications, as text and image generation tools, and then iterates repeatedly based on emotional input information using text-to-image and image-to-image processes until it finally outputs high-quality innovative content. The specific system construction process of this research is shown in Figure 1.

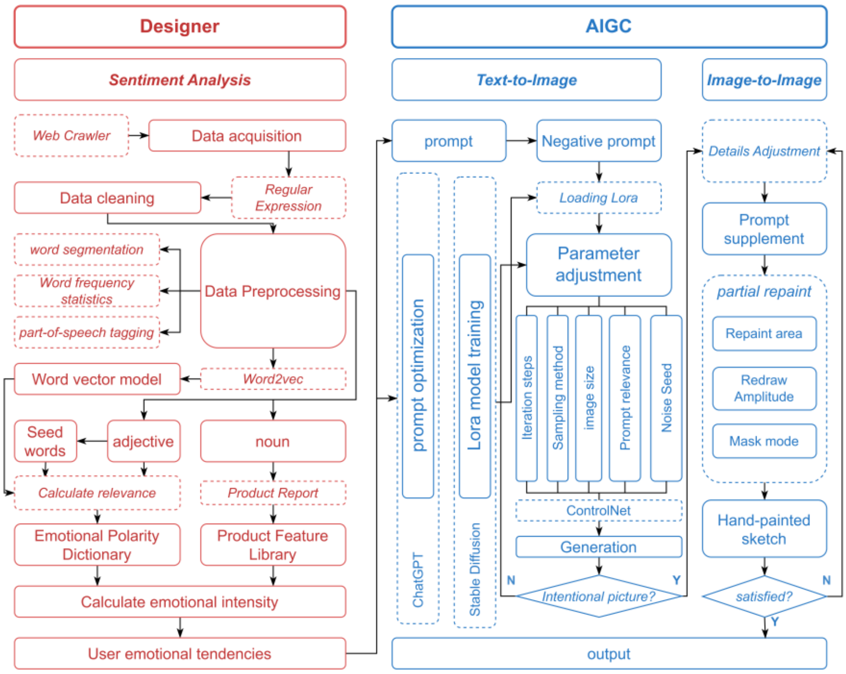

Fig.1. CADS construction process and framework

3.2. Sentiment Analysis

3.2.1. Collection and processing of online review data

To effectively mine users' sentiment tendencies from online comments on internet platforms, the first step is to collect and organize user comments on the relevant platforms. Faced with a massive amount of online comment data, manually collecting them one by one would consume a significant amount of time and effort. Web crawling technology provides an efficient tool for obtaining online data. It is a program or script that can automatically fetch information according to rules and is the primary method for collecting online comment data. It offers convenience, speed, and strong scalability [53].

Online comment data obtained through web crawling typically contain a large amount of erroneous or abnormal noise data. Therefore, statistical data analysis and other methods are needed to clean the collected data [54]. Additionally, taking into account the coherent nature of Chinese sentences in the comment data, and considering words as the basic units of natural language processing, it is necessary to employ word segmentation techniques to structure the text. [55]. Currently, there are numerous Chinese word segmentation technologies, including BosonNLP, LTP, Jieba, SCWS, and more. Different word segmentation technologies perform differently depending on the dataset's characteristics. Given that online comment text exhibits diverse vocabulary, misspellings, a wide range of internet slang, and a high degree of colloquialism, the study chose the Jieba word segmentation tool, which is more suitable for review data, to perform text segmentation, part-of-speech tagging, and word frequency analysis.

3.2.2. Word vector model

Online comments are unstructured data, and they need to be transformed into structured data for computer processing. Word vectors are a common representation method, and there are several models such as Glove, ELMo, Word2vec, and BERT [56]. The Word2vec model, proposed by the Google team in 2013, includes two training modes: CBOW and skip-gram. CBOW constructs the word vector space by predicting the target word based on context, while skip-gram primarily learns word vectors based on the context of the target word [57]. Word2vec uses low-dimensional, dense, real-valued vectors to represent the semantic information of words, making it easier to calculate semantic distances between words. Moreover, the skip-gram model is more advantageous in handling large datasets compared to the CBOW model [58]. Therefore, this study chooses to use the skip-gram model to generate word vector spaces for online comment data.

3.2.3. Sentiment polarity dictionary construction

Polarity words are words in text sentences that carry subjective emotions and are vocabulary used by users to express their attitudes and opinions. They are generally adjectives. Building a suitable emotional polarity dictionary can significantly improve the accuracy of sentiment analysis. The method and steps for mining polarity words from online review raw corpora and constructing an emotional polarity dictionary are as follows:

-Step 1: By text segmentation and part-of-speech tagging, obtain adjectives in the original online review corpus, filter and merge synonyms and near-synonyms to obtain adjective groups.

-Step 2: Use adjectives with high word frequency as the benchmark for sentiment evaluation words. Determine the emotional polarity of benchmark words, i.e., positive polarity words and negative polarity words, based on the HowNet dictionary, and define their emotional intensity.

-Step 3: Utilize the word vector space trained by the Word2vec model to calculate the relevance between non-benchmark words and benchmark words.

-Step 4: Use the emotional intensity of benchmark words as the standard to obtain the emotional intensity T_b of non-benchmark words as follows:

$$T_b = T_a \times L_{ab} \quad (1)$$

In the formula: T_a represents the emotional intensity of benchmark words, and L_{ab} represents the relevance between non-benchmark words and benchmark words.

3.2.4. Acquisition of product feature terms

Product features primarily refer to the design elements of a product, typically including its appearance, materials, functions, and other aspects. These elements serve as the objects of the user's emotional perceptions. Organizing product feature words effectively helps to better understand users' emotional attitudes and achieve fine-grained user requirement analysis based on product features [59]. Product feature words in online reviews mainly appear in the form of nouns. Therefore, in this paper, after text segmentation and noun frequency statistics, manual selection is applied to filter out nouns that clearly do not belong to the product feature dimension. Additionally, since user descriptions of product features in online reviews tend to be colloquial, it is necessary to combine the official professional descriptions of product features with product specifications. A manually defined product feature dimension table is created to enhance the organization of product feature words. Finally, product feature words are categorized based on the feature dimension table to obtain a product feature dictionary.

3.2.5. Feature-sentiment evaluation

Sentiment evaluation is primarily used to uncover users' emotional inclinations toward a specific feature dimension of a product, understanding the core product features that users are most concerned about. thereby providing a basis and guidance for product innovation and improvement. In this study, using product feature words as a reference, a semi-automated approach was employed to extract emotional words within a certain threshold distance from the feature words using the sliding window technique. The emotional evaluation of the product feature was then calculated based on the emotional intensity and frequency of emotional words appearing near the feature words. The specific calculation method is as follows:

$$Q = \frac{\sum_{i=1}^{m}(T_i \times X_i)}{\sum_{i=1}^{m} X_i} + \frac{\sum_{j=1}^{n}(T_j \times X_j)}{\sum_{j=1}^{n} Y_j} \quad (2)$$

In the formula: Q represents the sentiment evaluation value. m and n represent the quantities of positive and negative polarity words for the product feature, respectively. T_i and X_i represent the sentiment intensity and word frequency of the i positive polarity word for the product feature, while T_j and Y_j represent the sentiment intensity and word frequency of the j negative polarity word for the product feature.

Through sentiment evaluation, analyze the user's emotional inclination towards product features. When the sentiment evaluation value is positive, it indicates that the user has a positive attitude towards that product feature, which is a positive evaluation. When the sentiment evaluation value is negative, it indicates that the user has a negative attitude towards that product feature, which is a negative evaluation. The magnitude of the sentiment evaluation value reflects the user's satisfaction level and disappointment level with that product feature. Based on the results of sentiment evaluation, key elements for product innovation design can be extracted. After optimization, these elements are input in the form of a prompt into AIGC, serving as the emotional input source for AIGC-generated content.

3.3. AIGC as a partner in innovative design

3.3.1. Text-to-Text

Through sentiment analysis, we can extract product features that users are particularly interested in as keywords. However, because the product feature keywords at this stage are relatively scattered and lack strong logical connections. Directly inputting them into Stable Diffusion results in image outputs that do not fully meet expectations. In this scenario, ChatGPT, as a powerful natural language processing model, can effectively leverage its strengths. The research uses ChatGPT to optimize product keywords and conducts standardized training on ChatGPT through language description so that it can diverge and sort out based on the product core keywords we extracted earlier, add suitable scene descriptions, thematic style directions, and rendering image requirements to the product keywords. Ultimately, generate prompts that conform to the Stable Diffusion recognition logic.

3.3.2. Text-to-Image

Model

Stable Diffusion is primarily based on Latent Diffusion Models for text-to-image generation tasks. The large model implemented as the underlying architecture of Latent Diffusion Models is the basis for Stable Diffusion to perform all text-to-image tasks. In terms of large models, it has undergone iterative development, and recently, Stability AI has released the latest generation text-to-image model, SDXL 1.0. SDXL 1.0 is claimed to have the largest number of parameters among all open-source text-to-image models to date. It employs an innovative new architecture, consisting of a base model with 3.5 billion parameters and an optimization model with 6.6 billion parameters, capable of generating high-quality images in almost any artistic style [60]. Furthermore, thanks to the open-sourcing of Stable Diffusion parameters and models, a plethora of Checkpoint large models have emerged on the internet, in addition to the official large models. These large models are trained based on the official model and are primarily used to generate specific styles of images. Compared to the comprehensive style attributes of the official model, these Checkpoint large models tend to produce better image results when targeting certain specific styles.

The large model forms the foundation for Stable Diffusion to perform text-to-image tasks. However, relying solely on the capabilities of the large model is insufficient to achieve the desired image results. It is also necessary to use some small models that can fine-tune the image style. Some commonly used models for this purpose include:

1. VAE Model: This model can add filters to images, thereby adjusting their original color styles.
2. Embedding Model: It can generate features or styles specific to designated characters or elements.
3. Hypernetworks: Trained through artistic style, they can specify particular artistic styles for models.
4. Lora Model: The most widely used model is the Lora model, primarily trained based on large models. It fine-tunes the large model by acting on different parts of it. When the training direction is well-defined, it tends to produce better image results.

In addition, since the parameter volume of large models is relatively large, training them requires high-quality hardware facilities and computational power. Therefore, to effectively control image styles, most people currently choose to train specific small models for image generation.

Parameter

After configuring the basic style for generating images using the model, Stable Diffusion can further adjust the details of the generated images through various parameters. In addition to positive and negative prompts, these parameters include:

1. “Steps” which control the total number of iterations in the generation process.
2. Image size adjustments.
3. “Guidance Scale” which controls the balance between positive and negative prompts.
4. Setting the noise seed. Etc.

Plug-in (software component)

The open-source community provides various powerful plug-in support for Stable Diffusion, with ControlNet being one of the widely used plugins. ControlNet primarily controls the diffusion model by adding additional conditions, guiding Stable Diffusion to generate images in line with the creator's artistic vision [61]. This effectively enhances the controllability and precision of AI image generation. ControlNet effectively addresses the issue of limited control over image details by keywords in text-to-image models. Furthermore, the open-source community offers other powerful plugins, such as Deforum, which can convert images into videos, and Face-editor, which can be used for facial editing. These plugins significantly extend the functionality of Stable Diffusion and enhance the efficiency of image generation.

3.3.3. Image-to-Image

After setting all the parameters, a large batch of images is generated through text-to-image. After repeated iterations, preliminary concept images can be filtered and imported into Image-to-Image for fine-tuning of image details. In the image-to-image section, we can supplement and guide specific details of the image using new prompts or manually modify certain details through localized repainting. After multiple iterations of text-to-image and image-to-image adjustments, high-quality rendered images are produced.

4. Empirical study

This article selects the central console of pure electric vehicles as a case study for innovative design to validate the practicality and effectiveness of the system. The choice of this subject for empirical research is motivated by several factors.

Firstly, the new energy vehicle industry is currently experiencing vigorous growth, with a continuous influx of brand-new pure electric vehicle models entering the market. This not only injects fresh vitality into the automotive industry but also intensifies competition within the sector. To secure a significant market share in this increasingly competitive landscape, major automotive brands must continuously enhance their innovative capabilities and explore distinctive design languages to capture consumers' attention.

Secondly, in the field of automotive design, pure electric vehicles offer greater creative freedom, more extensive design space, and a multitude of possibilities in interior design compared to traditional vehicles. This is due to the absence of components like internal combustion engines, traditional dashboards, and fuel gauges. Consequently, designers have the flexibility to incorporate more innovative stylistic features. Additionally, a high-quality central console design directly influences consumers' driving experiences and emotional perceptions.

Considering these factors comprehensively, this research selects the central console of pure electric vehicles as the subject for investigation. It aims to verify the innovative design capabilities of CADS when applied to complex products and assess the effectiveness of infusing human emotions into product design during the design process.

4.1. Data collection

The study selected three of the most popular domestic car platforms in terms of user base: DCar, Autohome, and PCAuto, as data sources. Through web scraping technology, user comments were collected from these three platforms for a total of 60 best-selling models of compact, midsize, and large electric cars on the market. After merging, a total of 54,121 online comments were obtained.

4.2. Data pre-processing

Due to the initial comment data containing evaluations from users on various aspects of the car such as exterior, comfort, and performance, and the study's focus being on user feedback related to car interiors, the research used regular expressions in Python to clean the initial data. Sentences containing the keywords "interior" and "dashboard" were extracted, and the data was filtered to remove emoticons and short texts, resulting in 50,629 comments for analysis.

The study employed the Jieba segmentation tool to perform Chinese word segmentation and part-of-speech tagging on the comments selected for analysis. The segmented data was used as a corpus for training with the Skip-gram model. The word vector dimensions were set at 200, with 20 iterations, and a maximum context distance of 5, ultimately generating a word vector space for online comments.

4.3. Construction of sentiment polarity dictionary

The online comment data was subjected to part-of-speech tagging and word frequency statistics using the Jieba segmentation tool to extract all adjectives in the text. Redundant or similar adjectives were filtered and categorized based on their word frequency. They were then sorted by word frequency, with the top 20 adjectives and their frequencies displayed in Table 1.

From Table 1, six pairs of adjectives with higher frequencies (based on positive polarity words) were selected as seed words, as shown in Table 2. Simultaneously, the positive polarity words in the seed

Table 1

Top 20 adjectives and their frequency

Adjective	Word frequency	Adjective	Word frequency
Good	12465	Soft	1858
Large	5603	Rich	1604
Comfortable	4888	Clear	1449
Luxurious	3453	Smooth	1395
High	3268	Fine	1301
Small	3201	Hard	1166
Concise	3099	Young	1042
Exquisite	2660	Bad	949
Powerful	2411	Beautiful	873
General	2119	Meticulous	638

Table 2

Seed word

Positive polarity words	Negative polarity words
Good	Bad
Large	Small
Comfortable	Uncomfortable
Luxurious	Rough
High	Low
Concise	complex

Table 3

Dictionary of sentiment polarity (Partial)

Positive emotion word	Emotional intensity	Negative emotion word	Emotional intensity
Favorite	+0.85	Bad	-0.82
Breathtaking	+0.83	Insufficient	-0.78
Nice	+0.81	Regrettable	-0.75
Wonderful	+0.80	General	-0.75
Pretty good	+0.75	Mediocrity	-0.73
Prettier	+0.72	Doggedly	-0.62
Seamless	+0.68	Small	-0.62

words were assigned a sentiment strength of +1, while the negative polarity words were assigned a sentiment strength of -1. Using these six pairs of seed words as a basis, the word vector space obtained from training was used to determine the relevance between other non-seed words and seed words. The sentiment strength of each non-seed word was calculated according to Equation (1), resulting in the user's sentiment polarity dictionary, as shown in Table 3.

4.4. Building a Product Styling Feature Repository

Through part-of-speech tagging and word frequency analysis, we have obtained all the nouns in the comment data. We then excluded elements that were clearly not related to the design of the car's center console and interior. Finally, in conjunction with the professional descriptions of automotive design features found on the official websites of various car brands, we categorized and summarized the remaining nouns. After organizing the data, we primarily classified the design points of the interior and center console into two categories.

The first category is "Element Features," which mainly includes important characteristic elements directly involved in the design and structure of the

Table 4

Product Styling Feature Repository

Level 1 features	Level 2 features	Level 3 features	Level 4 features
Element Features	Material Used	Craft Material	Plastic Dermis Tumbled leather Artificial leather Lacquer Imitation leather Chrome Fine velvet Carbon fiber
		Textured	Embroider Decorative design Pinstripe Ripples Woodgrain Leather strap
		Tactility	Scrubby Smooth
	Color Scheme	Dominant Color Palette	Light tone Deep color Warm colors Cool colors
		Double Color Matching	Blue and White Blue and Gray Blue and Black Gray and White Black and Gray
	Function	Meter	Full LCD
		Central Control Screen	Suspended Embedded
		Size	⩾15.6 10.3-15.6
		Joint Screen	Dual screen Triple screen
		Button	Touch keys Physical buttons
		Detail	Seam Line Ambient lighting
Style Features	Overall Atmosphere	-	-
	Space Layout	-	-

Table 5

Emotional intensity of level 4 Features elements

Level 4 Features	Emotion Words	Intensity	Frequency	Emotional Intensity
Plastic	hard	-0.60	450	-0.69
	General	-0.75	330	
	Insufficient	-0.78	306	
Dermis	Wonderful	+0.80	276	+0.41
	Expensive	-0.44	254	
	Nice	+0.81	143	
	Good	+0.74	97	
	Luxurious	+0.70	54	
…		…		…
Ambient lighting	Exquisite	+0.61	317	+0.66
	Beautiful	+0.65	246	
	Good	+0.74	154	
	Cool	+0.71	139	

Table 6

Emotional description of secondary style features

Level 2 features	Emotional words
Overall Atmosphere	Bonzer Modern Sporty Younger Warm Chinoiserie
Space Layout	Three-Dimensional Feeling Structured Proportionate Coordination All-In-One Reasonable Enveloping Spacious Space Utilization Space Sense

car's center console. This includes aspects such as the materials and colors used in the center console, as well as the usage of various functional components.

The second category is “Style Features,” which focuses on the style or unique characteristics that can be reflected in the design of the center console and interior, such as the overall style of the center console and the characteristics of the space layout. Within each category, we further expanded on the noun features according to hierarchical levels. For detailed classification results, please refer to Table 4.

4.5. Feature-based user needs analysis

By utilizing online review data to construct a product’s design feature library, we have obtained the four-level elemental features and the secondary style features required for the design of the electric vehicle’s central control panel. In order to effectively explore user sentiment towards these product features, we used these feature nouns as the center and set a sliding window size of 5. We employed sliding window techniques to retrieve sentiment words near each feature word.

For the sentiment words obtained in relation to the four-level elemental features, we combined them with the sentiment polarity dictionary and used Equation (2) to calculate the sentiment intensity of each feature element, as shown in Table 5. The calculated results provide us with reference feature words for AI image generation prompts.

Regarding the secondary style features, the sliding window technique is mainly used to extract user sentiment descriptions of product style features, such as “spacious” space and a “luxurious” atmosphere, as shown in Table 5. These sentiment description words can be used in AIGC prompts to define the product’s style features.

4.6. Composing Prompts

Based on the results of sentiment mining, we composed prompts to guide the image generation via stable diffusion. Using the aforementioned product features with high user demand as the core, we divided the linguistic structure of the prompts into three main sections: prefix + subject + scene. The prefix portion mainly determines the image’s quality and artistic style, with words such as “best quality” and “masterpiece.” Using these terms in the prompts effectively enhances the image quality produced by stable diffusion. The subject is the central part of image generation, focusing on the core features of electric cars obtained through sentiment mining, such as genuine leather material, suede craftsmanship, dual-tone coloration, and full LCD instrument panels. The more detailed the description of product features, the more it aids in generating images that meet expectations. The scene content mainly describes the background and details of the product’s main body in the generated image, such as the environment in which it’s generated and the perspective of the main subject. Detailed content and commonly used expressions related to these three sections can be found in Table 7.

Based on the three-part structure of the prompt, we utilized ChatGPT to generate prompts in a professional format for us. We first trained ChatGPT with a natural language description, transforming it into an artistic assistant for stable diffusion prompts. We provided it with conceptual descriptions of the

Fig.1. Product Design (minimalism-eddiemauro)

Table 7

Prompt Keywords

	Keywords
Prefix	HDR, UHD,8K, Best quality, Masterpiece, highly detailed, Ultra-fine painting, Professional Futuristic, purism, minimalism, Chinese style···
Subject	Softcover, Tumbled Leather Craftsmanship, woodgrain veneer, embroidered floral embellishments, High-quality texture, pale tones dominate, Gray and White Duplex, Full Liquid Crystal Instrumentation···
Scene	virtualized scenario, first-person perspective, Sensible Perspective···

prompts as well as actual examples, allowing ChatGPT to learn the grammatical format of prompts. Ultimately, using natural language, we supplied ChatGPT with a series of keywords related to the generation theme. After optimization by ChatGPT, a detailed and professional prompt was outputted.

4.7. Preparing the base model and parameters

Large Model. To guide the generation of high-quality product images based on prompts, we sought to select high-quality Checkpoints suitable for product design from the website "Civitai". Ultimately, we chose the "Product Design Minimalism" variant model v2.0 trained by Eddiemauro based on SD 1.5, as detailed in Figure 1. This model's application in AI image generation within the product design domain is already well-established. The variant model v2.0 has better adaptability to objects, can produce minimalist-style images, and the object shapes are more realistic. Therefore, based on relevant parameters and generated image examples, we believe that this Checkpoint will fully meet our usage requirements.

Lora Model. We collected 20 interior images of purely electric vehicles, all from a front-facing angle, from major image websites like Pinterest, Behance, and PuxiangNet. Using the large model "product design Minimalism" as the base model, we trained

Fig.2. The first round of AI output

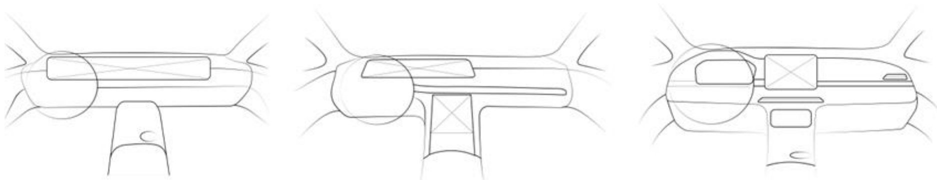

Fig.3. The second round of AI output Line art control drawing

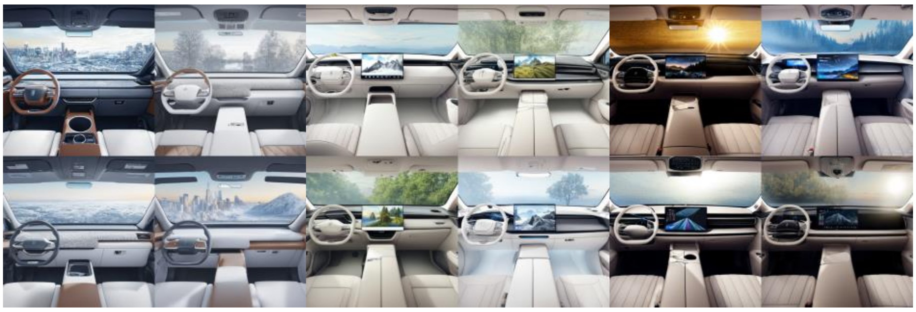

Fig.4. Line art control drawing

and developed our Lora model named “interior”.

Other Parameter. In order to produce high-quality image effects, the parameters set for stable diffusion must meet the requirements of large models. For the VAE model, we choose to use “vae-ft-mse-840000-ema-pruned.” For Clip skip, we use a value of “2”. Additionally, we use steps from “20-40” and CFG scale from “6-9”. For the sampler, we opt for DPM++SDE Karras. Regarding image dimensions, we mainly use 768×512. Some parameters may need adjustments within the range based on the actual imaging results.

5. Results and discussion

Through basic prompt guidance and the loading of related models and parameters, we utilized the text-to-image functionality of stable diffusion to conduct the first round of batch image generation. At this point, we did not introduce too many control conditions, allowing stable diffusion design outcomes to maintain a high degree of randomness, which can assist designers in brainstorming initial forms during the early stages of design. During this first round of batch image generation, we continuously optimized the structure and the weighting of keywords within the prompt based on the results of each generation, aligning this with adjustments to the related parameter ranges to incessantly refine the image output effects of stable diffusion until the final image quality reached a level of stability and preliminary satisfaction of our expectations. The results of this first round of image generation and their corresponding evolutionary processes can be seen in Figure 2. From the first round of image results, it is evident that without adding specific control conditions, the image quality produced by stable diffusion isn’t ideal. There are some perspective errors and structural deformations in the generated product images. To enhance the picture quality, further improvements and adjustments are necessary.

Next, we proceeded with the second round of image generation through the integration of the ControlNet’s Canny model. In this phase, we initially referred to some images generated in the first round by stable diffusion that substantially met our expecta

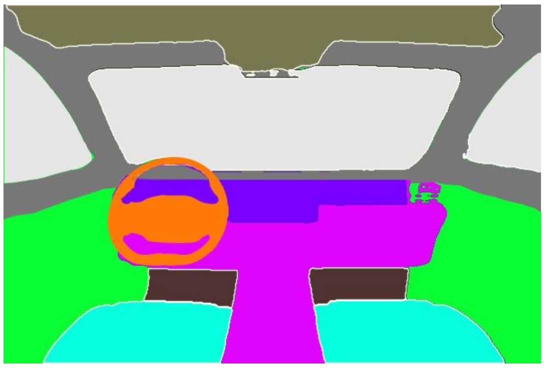

Fig.5. Semantic segmentation map

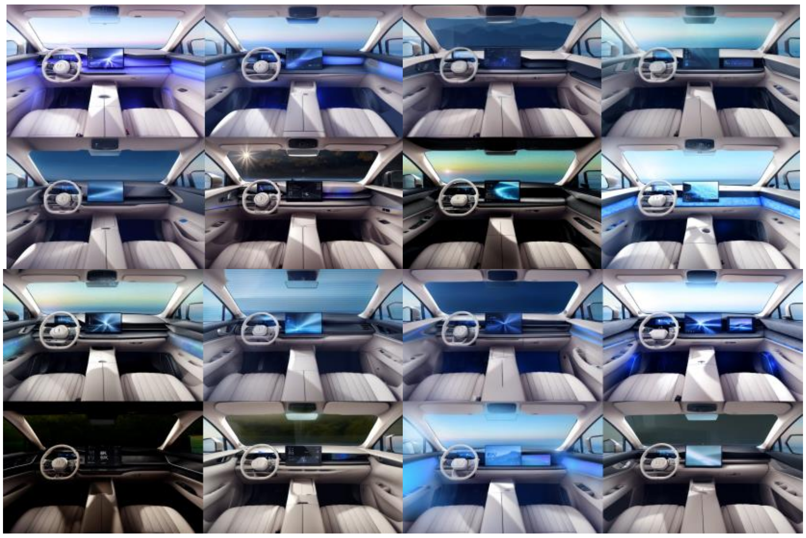

Fig.6. The fourth round of AI output

Fig.7. The fourth round of AI output

tions. We employed Photoshop to enhance and rectify the relationships between proportions and lines, and drafted the line-art sketches, as illustrated in Figure 3. Subsequently, these sketches were leveraged as control conditions for ControlNet, selecting the “Invert” pre-processing model and the “Canny” model for the second round of batch image generation. Throughout this process, we persistently fine-tuned the pertinent parameters to optimize the image output quality, with the results presented in Figure 4. From these outcomes, it can be discerned that the second round of image generation has improved upon the first, notably in the fundamental structure and spatial arrangement of various details. However, it is also vividly apparent that a portion of the images manifested chaotic color presentations. This implies that relying solely on the ControlNet’s Canny model is insufficient to govern the image generation quality. Utilizing only white-background black-line sketches might provoke errors in stable diffusion's recognition of color blocks between two components, consequently leading to some discordant color combinations, thereby affecting the output effects of the schemes.

To address the aforementioned issues, we incorporated the “Segmentation” feature from ControlNet for the third round of image generation. Referring to the visual results of the second round, we utilized the “OfADE20k” pre-processing model to produce semantic segmentation images, as seen in Figure 5. Similarly, these were employed as control conditions for ControlNet, with the Segmentation model being applied for the third round of batch image generation. The results are showcased in Figure 6. As can be observed from the figure, with the combined effects of ControlNet’s “Canny” and “Segmentation” functionalities, stable diffusion demonstrated significant improvements in design details and overall ambiance of the images.

After three rounds of adjustments and revisions, Stable Diffusion can now generate images in batches that align with the designer’s expectations. Next, the designer can select a design scheme and utilize the image-to-image functionality to tweak the details, thereby outputting high-quality product renderings. For instance, we initiated a fourth round of image generation using the image-to-image functionality on selected images, with the results displayed in Figure 7. Up to this point, the research has successfully set up CADS to collaborate between designers and AIGC, ultimately achieving high-quality product design outcomes.

6. Conclusion

In this study, we introduce CADS, a product concept generation design system for industrial designers based on AIGC. As one of the most attention-grabbing frontier technologies of the present day, AIGC has undoubtedly spurred lively discussions across various sectors of society. Numerous scholars have explored the developmental opportunities and risk challenges posed by AIGC from an ethical perspective. Building on this body of research, our study is the first to combine AIGC with product design from a practical application perspective. We explore the methods and pathways for applying AIGC in the product design process, striving to establish a scientific AIGC-assisted design system. This aims to assist designers in effectively creating product concept images, laying a solid foundation for future AIGC-based product design.

In our research on how AIGC can generate better content, we also present our own insights and methodologies. We aspire to infuse AIGC with human emotions from a designer's perspective, thereby enhancing the originality and depth of the content produced by AIGC. To achieve this objective, we opt to use an emotion analysis method based on online reviews. By constructing an emotion polarity dictionary and a product form feature library, we can effectively mine users' emotional preferences and extract core product feature keywords. Our findings confirm that using these as information and emotional inputs for AIGC significantly enhances the quality of the generated content. Moreover, this method of obtaining keywords also offers a novel approach for scientifically deploying AIGC in the future.

In our project, we selected ChatGPT and Stable Diffusion as the tools for text and image generation. ChatGPT was trained to act as a professional Prompt generation assistant to aid us in crafting appropriate prompts. Employing this prompt optimization method can effectively enhance the image quality produced by Stable Diffusion. Simultaneously, it explores the synergies between different forms of AIGC applications, increasing the efficiency of content generation. Subsequently, we carried out four large-scale image generation cycles using Stable Diffusion. During this process, by continuously adding control conditions, we optimized the final product renderings. Our practice confirmed that introducing an appropriate ControlNet control model can considerably improve the image outcomes of Stable Diffusion. The large models, Lora models, and parameter settings utilized in our study, as well as our choices and procedural methods related to the ControlNet control model during the image generation process, have pioneered a scientific design methodology for Stable Diffusion in product design. This serves as a reference for future industrial designers who wish to employ Stable Diffusion in product design.

It's essential to note that the development of AIGC has been particularly rapid in recent times. Among many technological domains, its associated application's update frequency and functional enhancements are remarkably evident. This technological iteration has significantly expanded the potential for design and application. Faced with such rapidly advancing AI technology, how to effectively keep pace with its evolution and fully tap into its inherent design value still requires ongoing exploration in future research.

References

[1] C.H. Zhang, C.S. Zhang, S. Zhang, Y. Qiao, C.H. Li, M.C. Zhang, S.K. Dam, C.M. Thwal, Y.L. Tun, L.L. Huy, D. Kim, S.H. Bae, L.H. Lee, Y. Yang and H.T. Shen, A complete survey on generative AI (AIGC): Is chatgpt from GPT-4 to GPT-5 all you need? *arXiv preprint arXiv:2303.11717* (2023).

[2] C. Zhang, C.S. Zhang, C.H. Li, S. Zheng, S.K. Dam, M.C. Zhang, J.U. Kim, S.T. Kim, J. Choi, G.M. Park, S.H. Bae, L.H. Lee and P. Hui, One small step for generative AI, one giant leap for AGI: A complete survey on ChatGPT in AIGC era, *arXiv preprint arXiv:2304.06488* (2023).

[3] Y. Zhang, L. Wang, B.J. Feng, A.L. Zhao and Y.J. Wu, Innovating Healthcare: The Role of ChatGPT in Streamlining Hospital Workflow in the Future. *Ann Biomed Eng* (2023).

[4] H. Lee, Using ChatGPT as a Learning Tool in Acupuncture Education: Comparative Study, *JMIR Med Educ* **9**(2023), e47427.

[5] W. Tao, S. Gao and Y.L. Yuan, Boundary crossing: an experimental study of individual perceptions toward AIGC, *Frontiers in Psychology* **14**(2023).

[6] P. Fortuna, A. Modlinski and M. McNeill, Creators Matter. Perception and Pricing of Art Made by Human, Cyborgs and

Humanoid Robots, *Empirical Studies of the Arts* **41**(2) (2023), 331-351.
[7] A. Modlinski and M.E. Gladden, An organizational metaphor for the 4th industrial revolution: the organization as Cyborg, *World Futures* **78**(6) (2022), 372-391.
[8] H.K. Lee, Rethinking creativity: creative industries, AI and everyday creativity, *Media, Culture & Society* **44**(3) (2022), 601-612.
[9] N. Anantrasirichai and D. Bull, Artificial intelligence in the creative industries: a review, *Artificial intelligence review* (2022), 1-68.
[10] J. Oppenlaender, The creativity of text-to-image generation, in: *Proceedings of the 25th International Academic Mindtrek Conference*, New York, United States, 2022, pp. 192-202.
[11] E. Cetinic and J. She, Understanding and creating art with AI: Review and outlook, *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)* **18**(2) (2022), 1-22.
[12] O. Stefan, Creative pros see generative AI as part of their future, Retrieved from https://blog.adobe.com/en/publish/2023/03/21/research-creative-pros-see-generative-ai-as-part-of-their-future.
[13] Y.M. Fang, The Role of Generative AI in Industrial Design: Enhancing, in: *2023 9th International Conference on Applied System Innovation*, Chiba, Japan, 2023.
[14] T. Knearem, M. Khwaia, Y. Gao, F. Bentley and C.K. Silver, Exploring the future of design tooling: The role of artificial intelligence in tools for user experience professionals, in: *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*, New York, United States, 2023, pp.1-6.
[15] X.H. Kang and S. Nagasawa, Integrating Kansei engineering and interactive genetic algorithm in Jiangxi red cultural and creative product design, *Journal of Intelligent & Fuzzy Systems* **44**(1) (2023), 647-660.
[16] H.C. Xue and D.P. Wu, Big Data-Driven Product Innovation Design Modeling and System Construction Method, *Mathematical Problems in Engineering* (2022).
[17] N.L. Li, X. Jin and Y.P. Li, Identification of key customer requirements based on online reviews, *Journal of Intelligent & Fuzzy Systems* **39**(3) (2020), 3957-3970.
[18] Y. L, Y. Dong, Y. Wang and N. Zhang, Product design opportunity identification through mining the critical minority of customer online reviews, *Electronic Commerce Research* (2023).
[19] Z.L. Wei, C.T. Wang and X.X. Yang, Weak tagging and imbalanced networks for online review sentiment classification, *Journal of Intelligent & Fuzzy Systems* **44**(1) (2023), 185-194.
[20] Y.W. Lai and M.Y. Chen, Review of Survey Research in Fuzzy Approach for Text Mining, *IEEE Access* **11**(2023), 39635-39649.
[21] A. Purohit and P.S. Patheja, Product review opinion based on sentiment analysis, *Journal of Intelligent & Fuzzy Systems* **44**(2) (2023), 3153-3169.
[22] L. Jiang, Z.W. Zou, J. Liao and Y. Li, A Study on the Application of Sentiment-Support Words on Aspect-Based Sentiment Analysis, *International Journal of Pattern Recognition and Artificial Intelligence* **37**(7) (2023).
[23] L.A. Zhu, Z.C. Zhu, C.W. Zhang, X.F. Xu and X.J. Kong, Multimodal sentiment analysis based on fusion methods: A survey, *Information Fusion* **95**(2023), 306-325.
[24] H. Cho, Comparison of Three Korean Sentiment/Emotion Word Dictionaries, in: *2023 IEEE International Conference on Big Data and Smart Computing, Bigcomp*, Jeju, South Korea, 2023, pp.334-335.
[25] M. Bordoloi and S.K. Biswas, Sentiment analysis: A survey on design framework, applications and future scopes, *Artificial Intelligence Review* (2023).
[26] R.K. Jana and S. Maity, An Accuracy Based Comparative Study on Different Techniques and Challenges for Sentiment Analysis, *Pervasive Computing and Social Networking, ICPCSN 2022* **475**(2023), 601-619.
[27] J.Y. Wu, W.S. Gan, Z.F. Chen, S.C. Wan and H. Lin, Ai-generated content (aigc): A survey, *arXiv preprint arXiv: 2304.06632* (2023).
[28] B.Y. Li, Y. Bai, X.N. Zhan and G. Li, The Technical Features and Aromorphosis of Artificial Intelligence Generated Content (AIGC), *Documentation, Information & Knowledge* **40**(1) (2023), 66-74.
[29] X.Y. Li, Y.Q. Fan and S.H. Cheng, AIGC In China: Current Developments and Future Outlook, *arXiv preprint arXiv: 2308.08451* (2023).
[30] Y.H. Cao, S.Y. Li, Y.X. Liu, Z.L. Yan, Y.T. Dai, P.S. Yu and L.C. Sun, A comprehensive survey of ai-generated content (AIGC): A history of generative AI from GAN to ChatGPT, *arXiv preprint arXiv:2303.04226* (2023).
[31] I. Goodfellow, J. Pouget-Abadie, M. Miraz, B. Xu, W.F. David, S. Ozair, A. Courville and Y, Bengio, Generative Adversarial Networks, *Communications of the ACM* **63**(11) (2020), 139-144.
[32] J.R. Cheng, Y. Yang, X.Y. Tang, N.X. Xiong, Y. Zhang and F.F. Lei, Generative Adversarial Networks: A Literature Review, *KSII Transactions on Internet and Information Systems* **14**(12) (2020), 4625-4647.
[33] D.P. Kingma and M. Welling, An introduction to variational autoencoders, *Foundations and Trends in Machine Learning* **12**(4) (2019), 307-392.
[34] D. Samarasinghe, M. Barlow and E. Lakshika, Flow-Based Reinforcement Learning, *IEEE Access* **10**(2022), 102247-102265.
[35] A. Radford, J.W. Kim, C. Hallacy, A. Ramesh, G. Goh, S. Agarwal, G. Sastry, A. Askell, P Mishkin, J. Clark, G. Krueger and I. Sutskever, Leaning Transfeble Visual Models from Natural Language Supervision, in: *International Conference on Machine Learning*, New York, United states, 2021, pp.8748-8763.
[36] Y.M. Pan, H. Cheng, Y.Q. Fang and Y.F. Liu, Region-Attention Prompt Learning for CLIP, *Journal of Intelligent & Fuzzy Systems* (2023), 1-15.
[37] F.A. Croitoru, V. Hondru, R.T. Lonescu and M. Shah, Diffusion Models in Vision: A Survey, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **45**(9) (2023), 10850-10869.
[38] B. Kawar, S. Zada, O. Lang, O. Tov, H.W. Chang, T. Dekel, I. Mosseri and M. Irani, Imagic: Text-Based Real Image Editing with Diffusion Models, in: *Proceeding of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Vancouver, Canada, 2023, pp.6007-6017.
[39] P.P Ray, ChatGPT: A comprehensive review on background applications, key challenges, bias, ethics, limitations and future scope, *Internet of Things and Cyber-Physical Systems* **3**(2023), 121-154.
[40] Y.T. Wang, Y.H. Pan, M. Yan, Z. Su and T.H. Luan, A Survey on ChatGPT: AI-Generated Contents, Challenges, and Solutions, *arXiv preprint arXiv:2305.18339* (2023).
[41] R. Mao, G.Y. Chen, X.L. Zhang, F. GuerY.T.in and E. Cambria, GPTEval: A Survey on Assessments of ChatGPT and GPT-4, *arXiv preprint arXiv:2308.12488* (2023).
[42] T.Y. Wu, S.Z. He, J.Q. Liu, S.Q. Sun, K. Liu, Q.L. Han and Y. Tang, A Brief Overview of ChatGPT: The History, Status

Quo and Potential Future Development, *IEEE/CAA Journal of Automatica Sinica* **10**(5) (2023), 1122-1136.
[43] K.M. Caramancion, News Verifiers Showdown: A Comparative Performance Evaluation of ChatGPT 3.5, ChatGPT 4.0, Bing AI and Bard in News Fact-Checking, *arXiv preprint arXiv: 2306.17176* (2023).
[44] A. Borji, Generate faces in the wild: Quantitative comparison of Stable Diffusion, Midjourney and Dall-E 2, *arXiv preprint arXiv: 2210.00586* (2022).
[45] A. Ramesh, P. Dhariwal, A. Nichol, C. Chu and M. Chen, Hierarchical text-conditional image generational with clip latents, *arXiv preprint arXiv:2204.06125* (2022).
[46] R. Rombach, A. Blattmann, D. Lorenz, P. Esser and B. Ommer, High-resolution image synthesis with latent diffusion models, in: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, New Orleans, Louisiana, 2022, pp.10684-10695.
[47] D.M. Hanna, The Use of Artificial intelligence Art Generator "Midjourney" in Artistic and Advertising Creativity, *Journal of Design Sciences and Applied Arts* **4**(2) (2023), 42-58.
[48] G. Marcus, E. Davis and S. Aaronson, A very preliminary analysis of DALL-E 2, *arXiv preprint arXiv:2204.13807* (2022).
[49] N. Dehouche and K. Dehouche, What's in a text-to-image prompt? The potential of stable diffusion in visual arts education, *Heliyon* **9**(6) (2023), e16757.
[50] P. Jie, X. Shan and J. Chung, Comparative Analysis of AI Painting Using [Midjourney] and [Stable diffusion]-A Case Study on Character Drawing, *International Journal of Advanced Culture Technology (IJACT)* **11**(2) (2023), 403-408.
[51] X.B. Liu, Z.K. Zhu, H.H. Liu, Y. Yuan, M. Cui, Q.S. Huang, J.H. Liang, Y. Cao, Q.Q. Kong, M.D. Plumbley and W.W. Wang, Wavjourney: Compositional Audio Creation with Large Language Models, *arXiv preprint arXiv: 2307.14335* (2023).
[52] S. Loeschcke, S. Belongie, S. Benaim, Text-driven stylization of video objects, *Computer Vision-ECCV Workshops* (2022), 594-609.
[53] Z. Chang, A Survey of Modern Crawler Methods, in: *Proceedings of the 6th International Conference on Control Engineering and Artificial Intelligence*, New York, United States, 2022, pp.21-28.
[54] K. Maharana, S. Mondal and B. Nemade, A review: Data pre-processing and data augmentation techniques, *Global Transitions Proceedings* **3**(1) (2022), 91-99.
[55] K.Q. Wang, Z.J. Gao, Y.N. Qiao, J. Li and S.R. Tong, User Requirements Identification and Trend of Evolution Mining in Online Reviews, *Mechanical Science and Technology for Aerospace Engineering* **42**(7) (2023), 1070-1080.
[56] N. Hernandez, I. Batyrshin and G. Sidorov, Evaluation of deep learning models for sentiment analysis, *Journal of Intelligent & Fuzzy Systems* **43**(6) (2022), 6953-6963.
[57] X.Q. Gao, Y.W. Jin, X.P. Wang and J. Hao, Research on product perceptual evaluation method based on online review mining, *Modern Manufacturing Engineering* **12**(2021), 13-20.
[58] D.P. Wei, N.S. Tang, T.L. Lei and S.W. Wen, The deep learning word vector model using part of speech and sentiment information, *Journal of Intelligent & Fuzzy Systems* **38**(1) (2020), 427-440.
[59] D.P. Jia, J. Jin, Q. Geng and S.Y. Deng, A Kansei Engineering Integrated Approach for Customer-needs Mining from Online Product Reviews, *Journal of the China Society for Scientific and Technical Information* **39**(3) (2020), 308-316.
[60] D. Podell, Z. English, K. Lacey, A. Blattmann, T. Dockhorn, J. Muller, J. Penna and R. Rombach, SDXL: Improving Latent Diffusion Models for High-Resolution Image Synthesis, *arXiv preprint arXiv:2307.01952* (2023).
[61] S.H. Zhao, D.D. Chen, Y.C. Chen, J.M. Bao, S.Z. Hao, L. Yuan and K.K. Wong, Uni-ControlNet: All-in-One Control to Text-to-Image Diffusion Models, *arXiv preprint arXiv: 2305. 16322* (2023).
[62] A. Gupta and H. Kumar, Multi-dimensional perspectives on electric vehicles design: A mind map approach, *Cleaner Engineering and Technology* **8**(2022), 100483.
[63] T. Ba, S. Li, Y. Gao and D. Y. Tan, Research on the interactive design of electric vehicle interior based on voice sensing and visual imagery, *International Journal of Vehicle Information and Communication Systems* **8**(1-2) (2023), 152-169.
[64] G. Huang and Y. Chen, Effects of Visual and Cognitive Load on User Interface of Electric Vehicle-Using Eye Tracking Predictive Technology, in: International Conference on Human-Computer Interaction, York, UK, 2023, pp.375-384.