Application and Problems of Artificial Intelligence Model in Photography Teaching

Zhiyan Li

School of Art and Design, Central Academy of Fine Arts, No. 8, Huajiadi South Street, 100000, Beijing
100224@yzpc.edu.cn

Abstract: With the continuous improvement of computing power and the explosive growth of data scale, the application of large models in the field of artificial intelligence (AI) has become more extensive and deeper. In particular, large-scale machine learning models, such as deep neural networks ((Generative Pre-trained Transformer-3 (GPT-3)), have demonstrated excellent performance in natural language processing, image recognition, and speech recognition. However, how to effectively integrate machine learning and large models to improve their application performance is still an important topic in current research and application. This paper describes the possibilities of two artificial intelligence models in the field of photography teaching, analyzes the application process of the two artificial intelligence models in different stages of photography teaching, and assumes specific practical scene cases to speculate the possibility of their application, explores whether the artificial intelligence model is feasible in the field of photography teaching, analyzes its potential advantages and disadvantages, and discusses the potential problems of artificial intelligence models in the application and how human beings should deal with the challenges and problems in the era of human-computer interaction.

Keywords: Application and Problems, Artificial Intelligence, Photography Teaching

1 Introduction

Artificial intelligence models (AI) have come to life from science fiction movies, and their definition is concise and profound: a technology that enables machines to mimic human intelligent behavior, including the ability to learn, reason, self-correct, and understand complex language [1]. At present, artificial intelligence is no longer an untouchable program code, it has penetrated all aspects of person lives, from big data collection to customized personalized information flow of complex algorithms to various targeted functions such as Chatgpt, Midjourney, Runway, Stable diffusion, and AI-generated websites that can train models. Therefore, in the field of photography, which has professional rational knowledge and artistic independent creation and innovation, it may be possible to expand the possibility of photography

Y. Wang (ed.), *Proceedings of the 2024 International Conference on Artificial Intelligence and Communication (ICAIC 2024)*, Advances in Intelligent Systems Research 185,
https://doi.org/10.2991/978-94-6463-512-6_71

teaching and promote the further development of photography methods and image language by using artificial intelligence technology to solve the professional knowledge of science popularization and explore the possibility of independent iteration and learning of artificial intelligence technology. There are two main types of artificial intelligence models used in photography teaching: large language models and image recognition models. These two types of models can automatically discover the correlation and characteristics between data through large-scale data learning, so as to achieve efficient processing and understanding of text, images and other data. In the field of natural language processing, AI large models have made remarkable achievements in tasks such as text generation and semantic understanding. In the field of computer vision, they have also demonstrated strong capabilities in tasks such as image classification and object detection. However, due to the fact that the accuracy of the artificial intelligence model for humanized and creative language response still has certain defects, and it cannot accurately understand the colloquial expression, the meaning beyond the words, and the artistic creative ideas, so the artificial intelligence model in the field of photography teaching has problems such as insufficient accuracy of language understanding, weak diversity of application scenarios, and low ability of independent creation.

The purpose of this paper is to explore the possibility of more applications of these two types of models in the field of visual and linguistic coexistence, such as photography teaching, to describe the applicable artificial intelligence models, and to put forward hypotheses and speculate on their potential problems in the application of real scenes.

2 AI models in Photography Teaching

The rapid development of artificial intelligence is not only convenient for us, but also poses a challenge for us to understand and adapt to this technological world [1]. Different types of AI models form the basis of this field, each with its own unique capabilities and use cases. From generative models, which can create images that are hard to distinguish between real and imaginary, to reinforcement learning models, which optimize the decision-making process through trial and error, each model represents an important aspect of AI technology. There are two main categories of AI models: supervised learning models and unsupervised learning models. A supervised learning model is a model that learns from labeled (labeled) data to predict the type or properties of new, unlabeled data. An unsupervised learning model is a model that discovers the rules, relationships, and structures of unlabeled data through learning. Common unsupervised learning models include clustering, association rules, principal component analysis, and more. In addition to supervised learning models and unsupervised learning models, there are some other types of AI models, such as semi-supervised learning models, reinforcement learning models, deep learning models, etc. The Table 1 is a summary of the classification and characteristics of artificial intelligence models used in the field of photography teaching.

Table 1. AI models in Photography Teaching

Type	Segment type	Characteristic
Generative Models Generative models, models that learn from the distribution of data and generate entirely new instances of data from it.	Generate adversarial networks (GANs)	GANs are a particularly interesting type of generative model that involves two competing networks: one is a generator, which can be seen as an artist trying to create a work of art that is enough to deceive the viewer; The other is the discriminator, a role that is more like that of an art critic whose task is to distinguish between what is a real work of art and what is created by an artist. Through this internal competition, generators learn to create more and more realistic compositions. GANs are already being used to create incredible new artwork, design fashion-forward clothing, and even create virtual gaming environments.
Discriminative Models Discriminant model, which maps input data to output labels.	Convolutional neural networks (CNNs)	CNNs are a type of discriminant model that is especially good at processing image data. Faced with thousands of image data, CNNs identify and classify objects in images by analyzing features such as shape, color, and texture. This capability allows CNNs to play a key role in facial recognition technology, such as unlocking smartphones or automatically tagging photos on social media.
Self-supervisedn Learning Models Self-supervised learning is a cutting-edge area of machine learning that allows models to learn from input data through automatically generated supervised signals without	GPT (Generative Pre-trained Transformer)	In the discussion of self-supervised learning, this paper has to mention the GPT model, which is very hot today. GPT, or generative pre-trained transformer, is a technology that uses self-supervised learning to pre-train language models. The GPT model is first pre-trained on a large-scale text dataset to self-learn language structure and knowledge by predicting the next word in the text. After completing the pre-training, GPT can be fine-tuned for a variety of specific language tasks, such as text generation, translation, Q&A, and summarization, showing amazing flexibility and ability.

relying on externally provided labels.	BERT (Bidirectional Encoder Representations from Transformers)	In addition to GPT, BERT is another important model in the field of self-supervised learning. It better understands the context of the language by processing text data bidirectionally, i.e., taking into account both the preceding and following words of each word in the text.

3 AI models in the Field of Photography Teaching

3.1 Overview of Large Language Models

Large language models (LLMs) are a class of foundational models that have been trained on large amounts of data to understand and generate natural language and other types of content to perform a variety of tasks [2, 3]

LLMs represent a major breakthrough in the field of NLP and artificial intelligence, and are easily accessible to the public through interfaces such as Open AI's Chat GPT-3 and GPT-4. Other examples include Meta's Llama model and Google's bidirectional encoder representations from Transformer (BERT/RoBERTa) and PaLM models. In short, LLMs are designed to be trained on large amounts of data to understand and generate text, as well as other forms of content, like humans. This model has the ability to extrapolate from the environment, generate coherent and contextually relevant responses, translate into languages other than English, summarize text, answer questions (general conversations and FAQs), and even assist with creative writing or code generation tasks. LLMs are revolutionizing applications in everything from chatbots and virtual assistants to content generation, research assistance and language translation, and teaching.

LLM in Photography Teaching In photography teaching, LLM is mainly used in the popularization of basic professional knowledge of photography in the early stage and the output of adjustment plan suggestions in the later stage of photography. In the form of AI assistants, communicate with users, provide the professional knowledge and science required by users, or provide solutions according to user needs.

For example, in a pre-shoot, the user provides a Japanese-style photo of the cherry blossoms he wants to take and his existing photographic equipment. For example, the user's existing equipment includes Sony T-100, Canon AE-1, and iPhone 15 Pro. And to provide users with these three types of equipment before the shooting example.

After the user selects the equipment, the LLM will provide the user with a shooting plan, that is, how to take Japanese-style cherry blossom photos, and popularize the photographic expertise, such as the need to adjust the sensitivity, white balance, aperture, etc. For example, in Shanghai, Jing'an Park is the best shooting location based on factors such as distance and fare, and the best shooting time is 9 a.m. and 5 p.m.

In the user's post-production adjustment of the photo, the LLM will also provide the adjustment plan and applicable software according to the user. For example, after

shooting cherry blossoms, the user feels that the photo effect does not meet the expected Japanese style, and the user wants to adjust the photo in post, then LLM will provide the user with a solution according to the user's needs, such as the need to crop the photo composition ratio, adjust the tone and color temperature, adjust the curve, etc., and recommend adjustment software for the user, such as Photoshop or Lightroom.

Fig. 1 enumerates the possible applications of LLM in pre- and post-photography.

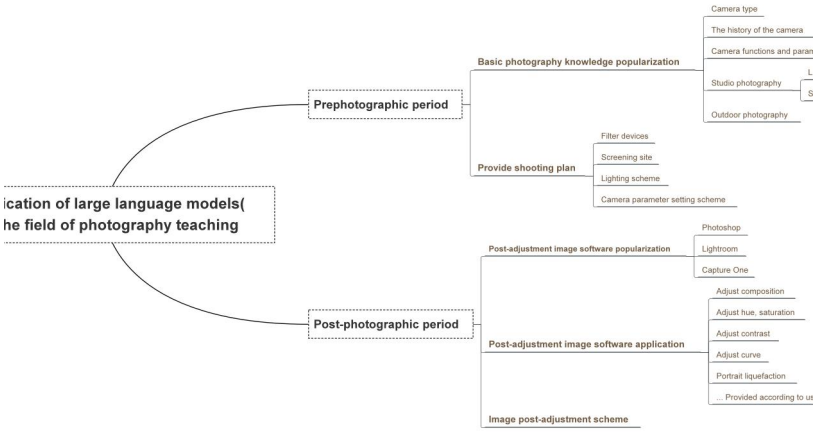

Fig. 1 The application of LLM in the field of photography(Photo/Picture credit :Original)

3.2 Discriminant Model—Overview of CNNS

Image recognition, also known as pattern recognition of images, is the specific application of pattern recognition technology in the field of images, which is a technology that establishes a graph image recognition model for the input image information, analyzes and extracts the features of the image, and then establishes a classifier to classify and recognize the features of the image. Common methods of image recognition include: Bayesian taxonomy, template matching method, etc. The image recognition method based on deep learning uses a large-scale deep convolutional neural network model (CNNS) to automatically learn image features from massive data sets and generalize them to the actual recognition scene. It simplifies complex work by simulating the human cerebral cortex, using multi-layer non-linear processing, step-by-step extraction, and layer-by-layer mapping [4,5]

CNNS in Photography Teaching In the field of photography teaching, CNNS is mainly used in the process of photography to perform image recognition, image classification and image segmentation on the phased photographic works uploaded by users, so as to point out the problems of current photographic works or put forward improvement plans for users. For example, according to the shooting plan and suggestions given by the LLM in the early stage, the user came to the designated place - Jing'an Park to shoot cherry blossoms, the user tried to take a few photos and

uploaded the photos, and then CNNS will analyze the uploaded images and compare the characteristics and parameters of the Japanese-style cherry blossom photos.

First, image recognition is performed on the image, and the cherry blossoms are identified in the image, and then the image is segmented, the proportion of cherry blossoms in the image and the location area are calculated, and then the image is classified according to the overall brightness and hue of the image. Subsequently, comparing the images taken by the user with the Japanese-style cherry blossom photos, according to the above analysis of the pictures, the user will systematically ask questions about the composition, the position of the main body of the image, the brightness and tone of the image, and then provide the user with adjustment solutions that are closer to the Japanese-style cherry blossom photos, such as adjusting the aperture, increasing the ISO, changing the focal length, etc. Fig. 2 enumerates the possible applications of CNNS in mid-shooting period.

Fig. 2 The application of CNN in the field of photography(Photo/Picture credit :Original)

4 AI Examples of the Application in Teaching

At present, there are few applications of photography in the field of artificial intelligence teaching, but there have been some cases of applying artificial intelligence models in secondary education, such as the model of "integrating artificial intelligence technology in teaching knowledge of geography"[6], which aims to integrate information technology with the curriculum.

4.1 AI-assisted Collection of Photographic Information Big Data

Information Processing of Photographic Knowledge In the era of information explosion in the 21st century, artificial intelligence technology has become an important tool for organizing and processing large amounts of information, especially in the field of photography teaching. AI technology makes information processing more efficient and accurate by reasonably classifying and intelligently matching the abundant photographic knowledge in the network. First, by massively collecting the basics of photography on the web, AI technology is able to learn and mimic well-established teaching resources such as "The School of Photography". These platforms cover everything from camera types and development history to specific parameter settings.

To make the collected knowledge more systematic, AI filters and classifies this information according to established classification criteria. For example, camera types can be subdivided into SLR, mirrorless, film camera, etc.; The historical stages of camera development can be divided by chronology or technological breakthroughs; Specific parameters such as aperture, shutter speed, ISO speed, etc., are also precisely categorized into the corresponding knowledge modules. Through this detailed

categorization and intelligent matching, photography beginners and enthusiasts can quickly find the knowledge they need and improve their learning efficiency. In addition, AI can also recommend relevant photography content based on the user's search habits and browsing history, making the learning process more personalized and convenient.

Information Collection of Basic Expertise in Photography In the field of photography, the accumulation and updating of professional knowledge is an important guarantee for the teaching effect [7]. AI can collect, organize, update, and supplement a large amount of photographic expertise through the powerful analysis capabilities of LLMs. The general information collection system consists of two parts: the collection of printed information and the collection of digital information. The collection of printed information mainly relies on the digitization of paper books, scanning and input, and then using AI technology to complete the data conversion. For example, classic photography books such as "Fundamentals of Photographic Composition" and "The Art of Light and Shadow" can be digitized into electronic versions by scanners and stored in a database for immediate reference.

Digital information collection combines four modules: site mirroring, intelligent information monitoring, intelligent resource discovery and intelligent information resource transformation. Site mirroring technology copies the entire site information to the system background, which not only speeds up the efficiency of information collection, but also ensures the integrity of the data. The intelligent information monitoring function is responsible for monitoring changes in network information in real time, ensuring that the stored data is always in sync with the latest information. For example, information about the latest changes in camera functions and the progress of AI applications in photography can be updated to the database in a timely manner through intelligent monitoring. The intelligent resource discovery module further optimizes the acquisition of information, and when the system detects new content related to photography, it will automatically collect and integrate it to provide users with the latest knowledge resources.

In practice, AI can also collect data on people's preferences for photographic devices and books from various e-commerce platforms and social media. For example, by analyzing the sales data and user reviews of photography books and photography equipment on platforms such as Taobao and Xiaohongshu, it is possible to understand the needs and preferences of different user groups. These data not only help teachers understand the learning context of students, but also provide data support for the design of teaching content. For example, if a particular model of camera is popular with a large number of amateur photographers, AI can use it to organize relevant educational materials to help novices get started with the device more quickly. In addition, through the analysis of user feedback, AI can continuously optimize the teaching content to ensure that the curriculum is always in line with the actual needs of users and technological development trends.

4.2 Assist in Learning the Basic Professional Knowledge of Photography

User Data Analysis and Personalized Teaching Path Planning In the teaching of photography knowledge, the artificial intelligence system collects and analyzes the user's knowledge base, including information such as photography practice behavior, photography application scenarios, content of photography works, and photography equipment, to understand the learning characteristics and needs of each user. Based on these data, the AI model can develop personalized teaching paths, and create appropriate learning content and learning progress plans for users according to different aspects such as the user's photography ability level, learning needs and interests in photography knowledge. Under the personalized learning path planning, users can learn according to their own situation, so as to improve learning efficiency and achieve results (Fig.3).

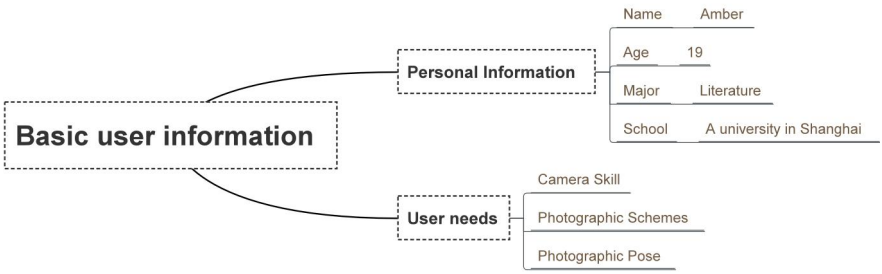

Fig 3 The personalized learning path planning(Photo/Picture credit :Original)

Diversification and Personalization of Multimedia Resources The artificial intelligence model provides a wealth of multimedia resources for the teaching of photography knowledge, such as text, images, audio, video and other forms of learning materials [8]. The AI model can provide personalized multimedia learning resources for users according to their learning characteristics and needs. With the help of multimedia teaching, users can access the learning content in a more intuitive and vivid way, thereby deepening their understanding and memory of knowledge.

In the early stage of knowledge popularization, it can provide users with a variety of resources to learn. For example, in the process of image recognition (CNNS), images can be used to show users the adjusted scheme, so that users can understand the scheme more accurately.

Real-time Feedback and Evaluation Mechanism The artificial intelligence model can give a real-time shooting plan in the process of user practice, according to the information provided by the user: existing equipment, shooting scene, shooting subject, shooting subject, shooting day, shooting day, light on the shooting day, indoor and outdoor, outdoor and outdoor studio, etc., give the user the closest to the needs of the program, and point out the application of photography knowledge for the

user in the program, and the user can feedback to the artificial intelligence model when there are problems or improper operation in the practice of shooting. The AI model gives real-time solutions to new problems and indicates to the user the applied photographic knowledge. After a period of teaching, the AI model no longer gives the user a shooting plan in advance, but requires the user to provide his final shooting work, and the AI model will correspond to the user's final shooting picture according to the user's initial shooting needs, and evaluate the user's application of photography knowledge.

Interactivity and Intelligent Dialogue System AI models are able to interact with users through LLMs. Users can have a conversation with the AI model by asking questions and interacting with each other, and get instant answers and solutions. The AI model is able to understand the user's needs based on their questions and give them answers and explanations accordingly. This interactive learning method not only assists users in learning knowledge, but also trains AI models to learn and train themselves to better understand users' needs and respond to diverse scenarios.

5 Problems of AI in Photography Teaching

Although the application of artificial intelligence technology in photography teaching has brought significant convenience and improvement, there are still many challenges in its development. These questions are not only about the technical aspects, but also about the uniqueness of the art of photography and the practical needs of teaching.

5.1 Low level of Language Comprehension

In the teaching of photography, the artificial intelligence model needs to understand the user's feedback, as well as the problems and needs raised by the user, for the popularization of knowledge and the provision of shooting and adjustment plans for the user [9]. From the perspective of media, there is no difference between machine natural language and human language, and it uses the linguistic symbol system created by humans for thousands of years. It is precisely because of this identity that machines and humans can communicate. But in essence, machine language has always been just a simulation of human language. Human language is native, while machine language is derived; Human language is active and growing, while machine language is completely passive. Human language is dynamic, family-similar, and somewhat ambiguous, while machine language is relatively fixed, clear, and completely logical. There are many reasons for these differences, such as the mechanism of behavior, the purpose of generation, and the historical-social dimension. From the perspective of behavioral mechanism, human language behavior includes expression, speech, communication and other parts, so human language is inseparable from emotion, cognition and communication. However, machine language behavior does not include the expression of emotions, the recognition of objects, the interaction between subjects, etc., but only a kind of instruction and information transmission [10].

Therefore, when it is applied in the field of photography teaching, when it comes to many more active or dynamic languages, it is often impossible to accurately understand the true meaning of language, that is, it is impossible to truly accurately understand the needs of users, and there will be repeated generation of similar answers, unable to answer from the position of the questioner, only satisfactory or dissatisfied judging criteria, some creative, non-general knowledge is likely to be ignored in the language model, and the artificial intelligence model still has a long way to go in simulating human language.

5.2 Model Data Quality and Authenticity

The effectiveness of AI in photography teaching relies heavily on access to a large amount of high-quality training data. However, obtaining authentic, reliable, and high-quality photographic data poses significant challenges. Photographic information collected from the web can be problematic in various ways, including issues of accuracy, relevance, and completeness. Inaccurate data can lead to the development of flawed AI models that may misinterpret photographic techniques or concepts, thus providing misleading guidance to students. Outdated content can fail to reflect current trends and technological advancements in photography, rendering the teaching process less effective. Furthermore, incomplete data sets can result in gaps in the AI's knowledge, preventing it from covering the full spectrum of photographic principles and practices. The process of curating and validating photographic data for AI training involves meticulous effort to ensure that the images and associated metadata are accurate, up-to-date, and comprehensive. This task often requires collaboration with professional photographers, educational institutions, and industry experts to source data that meets high standards of quality and authenticity. Additionally, ethical considerations must be addressed, such as obtaining proper usage rights and ensuring the privacy of subjects in the photographs.

5.3 Adaptability to Diverse Scenarios

Photography encompasses a wide range of scenes and styles, from portrait photography, landscape photography, to still life photography, each with its own unique technical requirements and artistic expression. Current AI models perform well in certain scenarios, but how to improve their adaptability in a variety of photographic scenarios is still a challenge. Photography is not only a technique, but also a form of artistic creation. AI is good at processing and analyzing existing data and giving recommendations based on historical patterns, but its capabilities are relatively limited in terms of creativity and artistic expression. For example, AI can analyze the composition and exposure of a photographic work, but it is difficult to output art direction that is truly creative and emotionally profound. Teaching photography requires not only the right technical guidance, but also the stimulation of students' creativity and personal expression. And this is exactly what AI is currently difficult to fully achieve.

6 Conclusion

This paper makes a preliminary exploration and speculation on the application of artificial intelligence model in the field of photography teaching, and explores the possibility of its application in various stages of photography practice. However, the accuracy of the response to diverse scenarios and creative needs is uncertain, and the understanding of creative subjectivity is not precise enough. At present, the creative landscape is richly adorned with masterpieces borne from the heart of artificial intelligence, even giving rise to AI as an artist in its own right. Anticipation mounts for an era where human-computer interaction evolves into profound dialogue, fostering a milieu conducive to robotic creativity and the pedagogical application of AI models. This is particularly resonant within the artistic domain, where the interplay between human and machine sparks endless possibilities and fresh societal narratives. As a result, novel avenues for artistic expression and media are poised to emerge, transforming the very fabric of creative endeavor. When knowledge and skills can be learned through human-computer dialogue, the main scene of teaching changes from knowledge teaching to personalized guidance and research activity organization, and teachers are upgraded from teaching craftsmen to tutors, so as to achieve higher quality and more creative teaching guidance in collaboration with intelligent machines. It must be pointed out that human-computer interaction will not replace interpersonal interaction, but the teacher-student relationship and student-student relationship will be more intimate due to the change of learning methods, thus creating a new type of learning knowledge and educational interpersonal communication environment in the intelligent era, especially the efficient cooperative relationship formed by learners in the research learning team, which is of great significance for cultivating core competencies in the intelligent era such as leadership and creativity.

Reference

1. Sun X., Che T., Yi J., Zhao B. The impact of AI big model paradigm and the change of scientific and technological innovation[J/OL]. Journal of Dalian University of Technology. Social Sciences, 1-3(2024)
2. What is a large language model [OL]. IBM (2024)
3. Wang Y., Li Q., Dai Z., Xu Y.. Research status and trends of large language models. Chinese Journal of Engineering, pp. 9, (2024).
4. A review of image recognition models based on deep learning[OL]. Artificial Intelligence Column, (2021).
5. Li Y., Zhao X., Wang R., et al. Image recognition of Chinese herbal slices based on improved convolution neural network. Science Technology and Engineering, . pp. 1-3 (2024)
6. Zhang W., Shan L., Yao M. A Knowledge Model for Geography Teaching Integrating Artificial Intelligence Technology: Connotation and Future Directions. Geography Teaching Reference for Secondary Schools, 2-3 (2024).

7. Feng Y. The application of artificial intelligence in network information retrieval in the era of big data. Scientific and technological innovation and application, pp. 2-3, (2023)
8. Li H., Application of Artificial Intelligence Technology in Computer-Aided Teaching. Shanghai Light Industry, 1-2, (2023)
9. Tao F., Liu X.. From human-machine dialogue to human-computer interaction-Philosophical Reflections on Large Language Models for Artificial Intelligence. Social Science Front, 3-4, (2024).
10. Wang Xiaolu, Yang Yunxuan, Xie Yangbin. Create a new form of human-machine dialogue learning-Current status and prospect of the educational application of large language models. Information technology education in primary and secondary schools, 1-2, (2024)