Proceedings of the 5th International Conference on Computing and Data Science
DOI: 10.54254/2755-2721/18/20230963

Visualizing the role of analytical diffusion models in AR--Focused analysis of the application of stable diffusion

Yizhi Zhang

The student of Nanjing University of Finance and Economics in Massey college

2025522608@qq.com

Abstract. The article focuses on the application and development of the Stable Diffusion module in the field of artificial intelligence image generation. The article presents a comprehensive description, analysis and discussion of the module's overview, operating environment, usage methods and its instructions, and points out the corresponding advantages and disadvantages.

Keywords: cite space, stable diffusion, augmented reality.

1. Introduction

1.1. Significance of the study

The development of Internet technology and globalization has led to the increasing importance of business and social networks. The emergence and application of diffusion model theory has provided new ideas and methods for the development of business. Diffusion model is a mathematical model describing the process of information dissemination, which can help us better understand the law of information diffusion in social networks and the influencing factors. Business development needs to continuously expand the market and attract more customers, which cannot be achieved without the dissemination and promotion of information. Therefore, the theory of diffusion model can provide a scientific and reasonable method for business development. Prediction and analysis of market trends; diffusion model can analyze the competition of different products in the market, predict market trends and provide important references for business development. Diffusion model can analyze market trends from different perspectives, such as product innovation, competition level, market capacity, etc., to help companies develop more scientific and reasonable development strategies.2. Optimize product marketing strategies: Diffusion model can help companies develop more reasonable product marketing strategies. By studying the needs, preferences and behaviors of potential customers, enterprises can tailor product promotion programs to improve product awareness and sales. At the same time, the diffusion model can also assist enterprises in choosing the most effective marketing channels and promotion methods to improve marketing efficiency and effectiveness.3. Control risks and costs: The diffusion model can help enterprises control risks and costs. By analyzing the market, enterprises can better understand the competitive situation and market changes in the industry, so that they can adjust their business strategies and reduce investment risks. In addition, diffusion modeling can help companies to position their markets more precisely and avoid over-expansion and unnecessary waste of resources.4. Improve corporate competitiveness: Diffusion modeling can help companies to improve their

DOI: 10.54254/2755-2721/18/20230963

competitiveness. By understanding information about the market, customers and competitors, enterprises can better identify their differentiation advantages and optimize their products and services to improve customer satisfaction and loyalty. At the same time, diffusion models can provide support for the development dynamics of enterprises, help them better respond to market changes and development trends, and improve their adaptability and competitiveness. To sum up, the purpose of combining diffusion modeling with business is to provide a scientific and rational approach to the development of business. By adopting diffusion model, enterprises can better understand the information of market, customers and competitors, become market leaders and improve their competitiveness.

1.2. Purpose of the study

With the development of society and the advancement of technology, the business sector has become more and more demanding for marketing. How to promote products and services efficiently has become an important issue and challenge for companies. Diffusion model, as a tool for predictive marketing, can help companies to achieve the spread of products and services. There are many benefits of combining diffusion model with business. One of them is that diffusion modeling can help companies predict marketing. Diffusion models are used to predict what is likely to happen when consumers buy a certain product or service. Diffusion models are based on two key concepts: innovation and diffusion. In marketing strategy, diffusion models can help companies choose the best marketing channel and time to promote new products and services quickly and with minimal cost. The diffusion model enables companies to develop different promotion strategies for different target markets, and it also enables them to predict market reactions and prepare response plans in advance through the model. Second, the diffusion model can help companies identify potential consumer groups. The diffusion model can predict the possible market response according to the characteristics and needs of potential consumer groups, helping enterprises to better understand the characteristics and needs of potential consumer groups. At the same time, diffusion models can also segment the market into different consumer groups, thus helping enterprises to make marketing strategies and decisions more accurately. Third, the diffusion model can promote cooperation and win-win among enterprises. Marketing needs resources and information not only from the company itself, but also from external partners and stakeholders. The application of the diffusion model requires the cooperation of multiple parties: diffusion not only relies on the company itself, but also can use various resources such as industry associations, news media, social media, etc. to carry out marketing activities. Therefore, the application of the diffusion model can promote cooperation and win-win among companies and create a better business environment. In short, the application of diffusion model in the business field can help companies improve the efficiency of marketing promotion, identify potential consumers, and also promote cooperation and win-win among companies. In order to successfully apply the diffusion model, companies need to understand its concepts and techniques, and conduct targeted analysis and trials to achieve better marketing results.

Especially for the AR field, augmented reality (AR) has gradually become an important technology under the continuous development of technology. However, AR still faces many challenges, one of which is how to expand its reach and enable AR to be more widely used in different fields. In this process, the diffusion model has become an important tool in the field of AR and has had a profound impact on the AR field. Diffusion model is a concept that has been widely used to study information, innovation and technology diffusion. In the AR field, diffusion models are often used to study the diffusion and adoption of AR technologies in different markets and users. Studies based on diffusion models can help understand the rate and path of AR technology adoption by users and find better ways to promote AR technology. Therefore, diffusion models have become an essential tool in the AR field. In the AR field, diffusion models are mainly related to market research and user research. For market research, diffusion model can help vendors understand the acceptance and popularity of AR technology in the market, as well as the number of potential users and user demand in the market. In addition, the diffusion model can help vendors understand the rate and path of AR technology adoption in order to maximize the promotion of AR technology. For user research, diffusion models can help understand the reasons and barriers for AR technology adoption by different users and provide better ways to meet user needs.

DOI: 10.54254/2755-2721/18/20230963

Therefore, diffusion models are of great value in the AR field. Currently, there have been many studies using diffusion models in the AR field. For example, some studies use diffusion models to study the acceptance and popularity of AR technologies in different markets; others use diffusion models to understand the adoption rate and pathways of AR technologies among different users. All of these studies provide valuable references and guidance for the development of the AR field. Diffusion models have become an important tool in the AR field and have had a profound impact on the development of the AR field. In the future, with the continuous development and popularity of AR technology, the application of diffusion model will be more extensive, providing better support for the innovation and application in AR field.

2. AR field research analysis

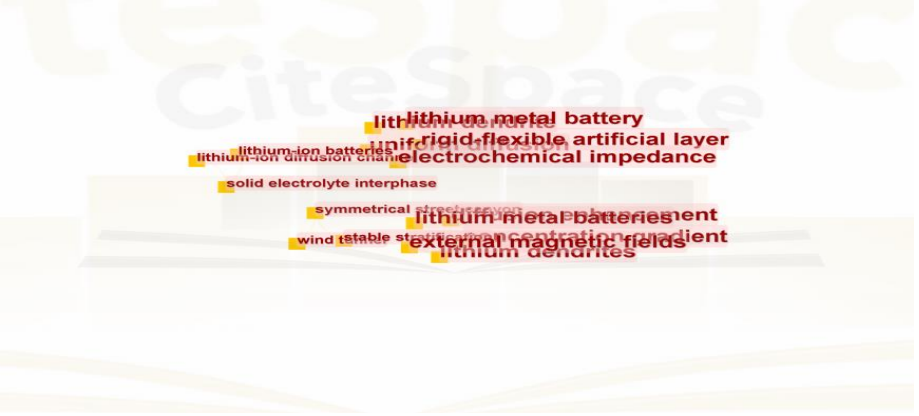

Figure 1. Keyword analysis based on cite space.

2.1. Hotspot analysis of postings

From the keyword analysis in Figure 1cite space, we can see that lithium ion is the focus of research in this field. In contemporary society, the demand for lithium-ion batteries is growing. However, the production process of lithium-ion batteries is accompanied by a large amount of waste and energy consumption. Therefore, it has become urgent to develop a sustainable model that can optimize the production of lithium-ion batteries. One of the solutions that can help optimize the efficiency of the lithium-ion production process is the diffusion model. A diffusion model is a mathematical model that predicts the performance of a material based on the physical properties of the diffusion process. In Li-ion batteries, diffusion models can be used to predict the ion transport between the cathode and anode materials and the electrolyte. Therefore, by studying the diffusion model, we are able to find some innovative points to optimize the production process of lithium-ion batteries. First, the production process of lithium-ion batteries involves a large amount of solvents and electrolytes. Therefore, by optimizing the solvent diffusion coefficient in the diffusion model, we are able to reduce the consumption of solvent in the battery production process. In addition, by increasing the concentration of electrolyte, the transport rate of solvent can be increased, thus improving the production efficiency. Second, the diffusion model can also be used to optimize the performance of the positive and negative electrode materials of the battery. By analyzing the diffusion coefficient and morphology of electrode materials, materials that achieve efficient transport can be found. In addition, by changing the particle size and morphology of the material, the efficiency of the electrode material can also be improved. Finally, diffusion models can also be used to optimize the cycle life of lithium-ion batteries. By studying the diffusion and solid solution processes of lithium ions in lithium-ion batteries, methods can be found to extend the life of lithium-ion batteries. For example, surface coating or alloying techniques can be used to enhance the cycle life of electrode materials. In summary, diffusion models play an important role in optimizing the production process of lithium-ion batteries. By studying the diffusion processes of solvent and electrolyte, the diffusion properties of electrode materials, and the transport and solid

DOI: 10.54254/2755-2721/18/20230963

solution processes of lithium ions in lithium-ion batteries, the performance and cycle life of lithium-ion batteries can be optimized. These research results will contribute to the advancement of lithium-ion battery production technology and environmental sustainability.

2.2. Analysis of research hotspots

From the perspective of commercial development, lithium-ion batteries are the new generation of batteries after traditional dry cell batteries, capacitors and electric motors in recent years and have become one of the most widely used batteries in today's commercial field. With high energy density, long life and light weight, lithium-ion batteries have been widely used in mobile communication, laptop computers, smart watches, smart homes, drones, electric vehicles and other fields, and stability is one of their important characteristics. Stable lithium-ion batteries have many benefits, one of which is to improve the efficiency of commercial development. Compared with other batteries, it can effectively improve the production efficiency and shorten the product development time because of its higher stability and shorter charging time and more times of charging and discharging. Whether it is for products in mobile communication, electric vehicles, etc., stable lithium-ion batteries can fully drive product innovation and make them more competitive and leading edge. Stable lithium-ion batteries can also reduce the production cost of business. Compared to other batteries, lithium-ion batteries have a significant advantage in terms of charge/discharge times and lifetime, which means that commercial enterprises can use fewer batteries to meet more demand, thus reducing their costs of buying and replacing batteries, which, of course, also reduces the energy consumption of commercial enterprises in the production process. Stable lithium-ion batteries offer better performance than other batteries and can reduce maintenance costs, giving commercial businesses an extra edge in the future. Stable lithium-ion batteries can also greatly improve the brand image of commercial enterprises. Because lithium-ion batteries have a high degree of stability, can very largely reduce the incidence of future failures, the video to ensure that the products produced by commercial enterprises can have high quality, stable performance, and then by the user's unanimous praise, such praise will also be beneficial to the promotion of the brand image of commercial enterprises, commercial enterprises in the highly competitive market to obtain more competitive advantage. To sum up, lithium-ion stability is crucial to business development. Stable lithium-ion battery can not only improve the development efficiency of business, reduce production costs, but also can greatly improve the commercial brand image, so that commercial enterprises in the competition in the market invincible.

Focusing on the AR field, with the increasing development of modern technology, people's lifestyles and communication patterns have changed dramatically. Among them, Augmented Reality (AR) technology has received more and more attention and importance. With the development of lithium-ion battery technology, the field of AR will usher in new opportunities and challenges. The development of lithium-ion battery technology has many benefits for the AR field. First, the development of lithium-ion battery technology can improve the endurance of AR devices. AR devices usually need a lot of energy to display the virtual reality world, and lithium-ion batteries can store more power in a small size to provide a longer endurance. With the development of lithium-ion battery technology, AR devices will have longer battery life, and users can enjoy the convenience and fun brought by AR technology for a longer time. Second, the development of lithium-ion battery technology can also improve the performance and efficiency of AR devices. As lithium-ion batteries have higher energy density and power density, the computing and image processing capabilities of AR devices will be improved, and users can get a more smooth and fine user experience. At the same time, lithium-ion battery technology can also provide higher stability and safety to ensure the safety of AR devices as well as users. In addition, the development of lithium-ion battery technology can also expand AR technology to a wider range of fields. As we all know, AR technology usually requires a large amount of energy to support its operation, and without sufficient power, the application of AR technology will be limited. However, due to the high energy density and lightweight design of lithium-ion batteries, AR technology can be more easily applied and promoted in various situations and applications. In summary, the development of lithium-ion battery technology will have a positive impact on the AR field. It can improve the range, performance

DOI: 10.54254/2755-2721/18/20230963

and stability of AR devices, and also enable the expansion of AR technology to a wider range of fields.

2.3. Advantages and limitations

The diffusion model is a commonly used tool for analyzing marketing strategies and predicting business trends. The model uses mathematical, statistical and behavioral science methods to study the process of spreading a product, service or idea, as well as the factors and patterns that influence consumer acceptance behavior. The diffusion model has both the ability and limitations to facilitate business development. Advantages: 1. Understanding market demand - Diffusion modeling can help companies understand the changing trend of consumer demand and predict the market trend so that they can adjust their marketing strategies to meet the market demand in time. In addition, the diffusion model helps to obtain consumers' evaluation of products or services to further improve the company's products or services. 2. Increase sales - By studying the process of product or service communication, the diffusion model can help companies to learn how to better communicate their products or services, so as to increase sales. Also, the model can help companies identify more effective promotional strategies to attract more potential customers.3. Achieve innovation - Diffusion models help companies understand how consumers react to and accept new products or services. This helps companies to better meet consumer needs when designing new products or services.

Limitations: 1. Data reliability - Diffusion models are based on a large amount of data. If the data is inaccurate or insufficient, it will affect the accuracy of the model. 2. Non-exact simulation of real life - Diffusion models are detached from the real environment, so there may be certain factors that are not taken into account. As a result, the results of a diffusion model do not fully predict real market conditions.3. Requires a lot of time and resources - Conducting a diffusion model requires a lot of time and resources. If companies do not make full use of this time and resources, it may lead to a lack of sufficient data, resulting in inaccurate analysis results. In conclusion, diffusion models are widely used as a tool for analyzing and predicting market trends. However, its results need to be interpreted with caution and should not simply be used as the sole criterion for business development. When using diffusion models for analysis, companies need to make decisions based on their own circumstances and a combination of factors.

Diffusion modeling is a commonly used prediction and analysis technique to effectively assess the diffusion of a new product, service or idea in the marketplace. In the digital era, augmented reality (AR) technology, a new and upgraded technology, has an increasingly significant impact on society. Through the diffusion model, we can evaluate the diffusion of AR technology in the market and understand the facilitation and limitations of AR technology. AR facilitation: First, AR technology, as a digital innovation, provides a new user experience. AR technology can interact and integrate virtual elements with real scenes, bringing users a richer and more vivid multi-dimensional experience of vision, hearing and touch. This new experience mode will provide a strong boost to the promotion and popularity of AR technology. Secondly, AR technology can achieve more accurate advertising push and customization; AR technology can identify and track users more accurately, and make advertising push more personalized and meet users' needs, thus improving the effect of advertising and user experience. Finally, AR technology has high technical barriers, which will be an advantage for AR technology to increase its competitiveness. AR technology involves knowledge and technology in virtual reality, computer vision, machine learning and other fields, which requires high professional skills and experience, which makes AR technology have a more solid position and secluded high value in the market. AR limitations: However, AR technology also has some limitations. First, the promotion of AR technology requires high cost and technical support. Because AR technology involves knowledge and technology in multiple fields, it requires high investment in R&D and team support. Secondly, the popularization of AR technology also faces the challenge of user understanding and acceptance. Because AR technology is an emerging technology field, the audience is currently limited to niche groups such as technology professionals and explorers, and the popularity is relatively narrow. Finally, there are certain restrictions on the scenarios and functions of AR technology applications, which will limit the application scope and landing realization of AR technology. Overall, AR technology, as an emerging digital technology,

DOI: 10.54254/2755-2721/18/20230963

has great potential for real-world problem solving. the promotion of AR technology needs to face a number of challenges and limitations, but the promotion of AR technology is also significant, and in the future, with the continuous progress and popularity of technology, AR technology is expected to become one of the important pillars of the digital economy era.

2.4. Outlook for the future

In recent years, with the development of Internet and mobile Internet, diffusion models have been more and more widely used in the business field. Diffusion model is a mathematical model to study the spread of information, products or services in a group, which can predict the promotion effect of products or services by modeling the audience's influence, spread speed, time and other factors. By combining with business, diffusion modeling becomes more practical and the future outlook is increasingly broad. First, diffusion models can be used in the development of commercial marketing strategies. By analyzing the target audience, brand characteristics and market competition, commercial companies can use diffusion models to predict the promotion effect of a product or service in the market. On this basis, commercial companies can develop more accurate and effective marketing strategies to improve the promotion effect and market share of their products or services. Second, the diffusion model can be used for the improvement and optimization of commercial products or services. According to the feedback mechanism in the diffusion model, commercial companies can learn consumers' feedback on products or services, and improve consumers' satisfaction, loyalty and experience through the improvement of products or services. At the same time, commercial companies can combine the prediction results of the diffusion model to continuously optimize the functions and features of their products or services to gain more market share by improving their differentiated competitive advantages. Finally, the diffusion model can be used for innovation and upgrading of business platforms. For Internet platforms such as e-commerce platforms and social media platforms, the application of diffusion models is more in-depth and extensive. Merchants can use the influence and user groups of the platform to predict the promotion effect and make more accurate marketing strategies through diffusion model. At the same time, platform owners can also use the ideas and practicality of the diffusion model to innovate and upgrade the business platform, improve the user experience and value of the platform, and further expand the user scale and market influence of the platform. Overall, the combination of diffusion model and business has become one of the most important and practical tools in the business field. In the future, as the Internet and mobile Internet become more popular and efficient, the application prospect of diffusion modeling will become more and more broad. Business companies can continuously improve their market competitiveness and market share by using the ideas and applications of diffusion model, especially in the Internet+ era, which is more practical and forward-looking.

The diffusion model is of good help in the research of AR field. In the future, AR technology diffusion model will become an important indicator for the development of AR field and can predict the trend and development direction of AR technology market. Diffusion model can analyze and predict the characteristics of the future market through data, such as market size, audience demand and consumer behavior, in order to assist AR companies to make timely decisions on the market demand and development direction of products and services. At the same time, the diffusion model can also enable AR enterprises to better understand the competitive environment and market trends, and improve the match between enterprises and the market to better meet the market demand and strengthen the market position. AR is applied in different fields, involving training, games, advertising, design, medical, education and many other fields. In the future, some of the central directions in AR will include: user experience, technological innovation, advertising integration, etc. The diffusion model helps companies achieve better development in these fields. Companies can draw on the analysis results of the diffusion model to determine the target audience of their products and optimize the user experience to further improve product quality and engagement. In the field of AR games, diffusion models can better predict and infer users' consumption habits and game demands, so as to develop better user strategies and marketing strategies to effectively improve the profitability and user retention rate of games. In the AR medical field, the diffusion model can be used to better predict patients' acceptance of AR technology

DOI: 10.54254/2755-2721/18/20230963

through sample studies and data analysis to provide better medical services for patients, and also encourage the application and development of the technology. To conclude, the future development of AR field is very extensive and high speed, and the application of diffusion model in this field will become one of the important development trends, AR companies can use diffusion model to understand the market trend and user demand, and make more optimized and accurate marketing strategy, which will help the better development and promotion of AR technology. At the same time, practitioners engaged in AR research and marketing can also learn from the prediction and analysis tools of AR technology diffusion model to promote the development of AR technology more widely and deeply.

3. In-depth on stable diffusion

3.1. Introduction to stable diffusion

In 2022, Artificial Intelligence (AI) has made great strides in the field of image generation. Various AI-generated image modules such as Midjourney, DiscoDiffusion, Dall-E 2, etc. are becoming available. The common features of these modules are: enhanced modal deep learning and generation of relevant learning models, the use of prompt (cue word) text for finding and associating in the model, and finally image generation based on the text. As the quality of AI-generated images has increased, so has the interest generated.

Since August, two major events have happened in the field of AI image generation: first, on August 22, "Stable Diffusion", an AI image generation module developed by Stability AI, a British AI company, was made fully open source on "Hugging Face", a community website for AI enthusiasts. About a week later, on August 30, "ERNIE-ViLG", an AI image generation module developed by Baidu, China, was also made public on the Hugging Face community.

Both of these AI image generation modules have attracted a lot of attention from the industry in this field. Stable diffusion is a text-to-image model based on Latent Diffusion Models (LDMs), which is introduced in this paper. Specifically, thanks to the computational resources support of Stability AI and the data resources support of LAION, Stable Diffusion trains a Latent Diffusion Models on a subset of LAION-5B, which is specifically used for text-to-image generation.

3.2. The description language of Stable-diffusion painting

The descriptive language of Stable-diffusion painting can be seen as a fractal-based language, which mainly describes the painting by describing the fractal shape, color, texture and other characteristics.

Specifically, this language encompasses the following:

Fractal shape description: Use mathematical formulas or other algorithms to describe the fractal shape formed by stable diffusion, such as the number of fractal branches, angles, lengths, branching relationships, etc.

Color Description: Create rich color effects by adjusting the hue, saturation, lightness and darkness properties of colors according to the needs of painting.

Texture description: Use complex texture patterns, such as texture transitions and color gradients, to enhance the three-dimensionality and sense of depth of the painting.

Motion description: Create dynamic visual effects, such as motion blur, distortion, etc., by controlling the speed and direction of movement of image elements.

Painting style description: stable diffusion painting style is unique, some special bracket shape use, etc.

In short, in the descriptive language of stable-diffusion painting, fractal shapes, colors, textures, movements and styles are all inseparable elements that intertwine and influence each other to create a unique work of art.

3.3. Overview of stable diffusion

As a leader in AI image generation since its launch, Stable Diffusion has been in the public domain for a long time, and its developer, Stability AI, conducted internal testing between May and August 2022.

DOI: 10.54254/2755-2721/18/20230963

The project was open-sourced and the code and documentation were made public at the same time.

While modules such as Midjourney, mentioned above, usually require many complex parameters to be set in practice, Stable Diffusion simplifies the parameter settings to a large extent and achieves certain improvements in image generation speed and details [1], thus gaining attention rapidly after its open sourceization.

The major difference between Stable Diffusion and other AI image generation modules is the addition of the "image generation" function on top of the "text-generated image" function. In other words, a new image can be generated by inputting a reference image and a text prompt. This feature is also one of its biggest competitive advantages.

Another noteworthy aspect is that Stability AI announced that for learning models made with Stable Diffusion and the images it generates, the makers are free to use them, even commercially, as long as they follow the relevant licensing rules adopted by the original project and are not applied to illegal and unethical scenarios. Although Stable Diffusion makes its code documentation publicly available, it is currently very difficult for users to produce or modify their deep learning AI learning models on their own. This is because the learning model is generated by running a cluster of 4000 A100 graphics cards for one month, which is very expensive for small and medium-sized enterprises and individual users to rebuild the learning model. The V1.4 learning model is about 4G in size.

3.4. Running environment

The Stable Diffusion runtime environment is divided into three main types: an official public demo version, a locally deployed version based on the Python environment, and a supported cloud deployment version based on the Jupyter Notebook environment.

-Official public demo version

Stable Diffusion currently publishes two application pages, one of which is a free public demo version with simple setup parameters and a network queue for each use. The other page is a public beta version for commercial services, where complex parameters can be set, and is available on a paid membership basis.

Since the code has been fully open source, we recommend users to use the following two methods to deploy it by themselves, which can realize the full functionality of this AI module for free.

-Python environment based local deployment version

To establish a local piggyback environment, the computer should have at least a graphics card with NVIDIA 4G or higher video memory. It is recommended to use a graphics card with NVIDIA 10G or higher video memory to generate images using the GPU, which will result in a better speed experience.

For local deployment, it is recommended to use Anaconda/Miniconda to configure the basic environment and NVIDIA's CUDA computing platform to make full use of the GPU. we can also use some user-integrated code packages that incorporate Web UI functionality. For example, AUTOMATIC version 1111, which has recently become popular on github. In addition, we need to configure the learning model accordingly.

If we encounter a situation where the local environment is not successful, we can use a cloud-based Jupyter Notebook environment for deployment. If possible, we can use Google Colab cloud service environment with GPU mode enabled, or other domestic deep learning platforms that provide free computing power.

Also, we need to write or obtain the relevant project configuration files based on this environment. It is recommended to use the AUTOMATIC version 1111 Web UI project file on the github code platform, which can be imported into the appropriate cloud or local environment, and execute the code step by step. The file will be downloaded and configured automatically when executed. we will need to fill in the user token once, and we will need to register with the Hugging Face community's Stable Diffusion project and open a free write token. A total of about 10G of configuration files and models will be downloaded from the beginning to the end of the code execution. A temporary URL will be generated after all executions are completed, and we can access it to open a dedicated Web UI interface to use.

Although Colab has a limited amount of free daily GPU computing power, the higher configuration

DOI: 10.54254/2755-2721/18/20230963

of Google servers will yield faster image generation results than the average local computer. With the same parameters set, I built a local environment with a NVIDIA GeForce RTX 3070 Laptop with 8GB of video memory, and generated images in about 2 minutes per image, while in the Colab cloud environment, I was able to achieve less than 30 seconds per image.

In order to use the GPU to its full potential, we will need to install NVIDIA's CUDA computing platform, which is available on the github code platform, as well as some user-integrated Web UI packages. For example, AUTOMATIC version 1111, which has recently become popular on github. we also need to configure the learning model accordingly.

3.5. Using method and related instructions

For example, in the AUTOMATIC version 1111 Web UI, we only need to enter the English prompt for beginners, upload an additional reference image if we want to use the graph-generate feature, and click the generate button to generate the image result after a certain period of time.

If we want to adjust the image generation process, we need to pay attention to the following parameter settings.

Prompt: The most important element. In Stable Diffusion, we need to enter the entire English language, and we can enter multiple descriptors at the same time. Too many prompts can affect the speed of image generation to some extent. If we want to highlight specific prompts within multiple prompts, wrap them in multiple parentheses.

Strength: The range of values is 0.0T.0. Setting this value in the Bunsen burner mode has no effect, while in the graph burner mode determines the degree of resemblance to the original image. The basic value is 0.75, and the measured value is generally above 0.7, which starts to differ significantly from the original image.

Seed: specify the corresponding seed. The default value of -1 is random. Each seed is considered to be a fixed token from the learning model, and different users can reproduce the same generated image by setting the same prompt and the same value of seed in the default environment using the same learning model.

Scale (scale): indicates how to follow the instructions of the prompt, although the larger this value will be the more in line with the content of the instructions, but the image itself will be more easily deformed. Generally in the range of 7 ~ 11 is more appropriate, usually we can keep the default value of 7.

Sampling Steps: The default value of 50 steps is recommended. Increasing the number of steps will cause the image generation time to increase further and may change the structure of the screen.

Width and Height: the width and height of the image, both recommended to use the default value of 512 pixels. While it is possible to set larger image sizes (but follow a multiple of 64), they also require larger video memory to run. For example, an image size of 512*768 pixels may require more than 12G of video memory, and it is very easy to make errors during image generation when the local video memory is insufficient.

Batch count: Set how many sets (times) of image results to generate. Note that if we set to generate multiple sets, we cannot specify a fixed seed, otherwise it will take a lot of time and generate the same results.

Batch size: the number of images contained in each group, it is recommended to set it to 1 by default, otherwise it will lead to multiplying the occupied video memory, and errors will occur when the video memory is not enough.

3.6. Advantages, disadvantages and partial solutions

The advantages of Stable Diffusion are more obvious: it runs faster than other AI modules in the same configuration. The necessary parameters are simple to set. The image display is optimized for landscape images and close-up images. The model also includes learning keywords of various artists' styles, and the faces of famous people are displayed better. In addition, NSFW safety device is set in the model, which forces the generated images to pure black when the AI recognition image generation results may

DOI: 10.54254/2755-2721/18/20230963

be inappropriate.

The disadvantage of Stable Diffusion is that it sometimes generates faces or limbs with obvious problems, where the facial problems can be fixed to some extent by combining the AI facial repair generation adversarial network "gfpgan". There is no automatic solution for limb problems (mainly multi-hand, multi-foot, multi-finger). If this type of image is not considered for post retouching, it is necessary to generate a batch of images and then select a small number of images that fit the normal human body characteristics, and finally determine whether they need to be further processed with other AI restoration methods. In addition, the limb issue is common to all AI-generated image modules and is related to the modeling method. This issue may become a new factor that restricts the development of AI image generation.

If it is difficult to generate images beyond the default pixel resolution due to insufficient video memory, the AI image overscaling adversarial network "realesrgan" can be used in combination to enable the output of images at the default size to be scaled up for a larger range of applications.

Perceptual Image Compression (PIC)

Perceptual compression is essentially a tradeoff, and many previous diffusion models can be performed without using this technique, but one big problem with the original non-perceptually compressed diffusion model is that, since the model is trained on the pixel space, if we want to generate a very high resolution image, this means that the space we train on is also a very high dimensional space. Introducing perceptual compression means that the original image is processed by a self-coding model like VAE, which ignores the high-frequency information in the image and keeps only the important, basic features. The benefit of this approach, as mentioned in the introduction section, is that it can significantly reduce the computational complexity of the training and sampling phases, allowing tasks such as text-image generation to generate images on consumer GPUs at the 10-second level, greatly reducing the threshold for implementation.

Perceptual compression mainly utilizes a pre-trained self-coding model that is able to learn a potential representation space that is perceptually equivalent to the image space. An advantage of this approach is that only a generic self-coding model needs to be trained, which can be used for the training of different diffusion models to be used on different tasks. In this way, the approach of perceptual compression can be very easily extended to various image-to-image (inpainting, super-resolution) and text-to-image (text-to-image) tasks in addition to its application to standard unconditional image generation.

It follows that the training of a diffusion model based on perceptual compression is essentially a two-stage training process, where a self-encoder needs to be trained in the first stage before the diffusion model itself is trained in the second stage. In the first stage of training the self-encoder, in order to avoid a high degree of dissimilarity in the potential representation space, the authors use two regularization methods, one is KL-reg and the other is VQ-reg, so that both KL and VQ implementations will be seen in the officially released first stage pre-trained model. This implementation, AutoencoderKL, is mainly used in Stable Diffusion.

Specifically, given an image $\chi \in R^{H*W*3}$, we can first use an encoder $\mathcal{E}$ to encode the image into the potential representation space $Z = \varepsilon(\chi)$, where $Z \in R^{h*w*c}$, and then a decoder to reconstruct the image from the potential representation space $\hat{x} = D(Z) = D(\varepsilon(\chi))$. In perceptual compression compression, the size of the downsampling factor is $f = H/h = W/\omega$, which is the power of 2, i.e., $f = 2^m$.

Latent Diffusion Models (LDM)

To begin with a brief introduction to the ordinary diffusion model (DM), the diffusion model can be interpreted as a temporally denoising autoencoderequally weighted sequence of denoising autoencoders) $\varepsilon\theta(\chi_t, t); t = 1...T$, whose objective is to predict a corresponding denoised variant, or predicted noise, based on the input χ_t, where χ_t is the noisy version of the input χ. The corresponding

DOI: 10.54254/2755-2721/18/20230963

objective function can be written in the following form:

$$LDM = E_{\chi,\varepsilon} \sim N(0,1)\mathrm{t}[\left\| \varepsilon\text{-}\varepsilon\theta(\chi_{\mathrm{t}},\ \mathrm{t}) \right\|_2^2]$$

where t is obtained from {1, . .T} is obtained by uniform sampling in

In contrast, a pre-trained perceptual compression model is introduced in the latent diffusion model, which consists of an encoder ε and a decoder D . This allows the model to learn in the potential representation space by using the encoder that can be obtained at training time Z_{t} , and the corresponding objective function can be written in the following form:

$$L_{LDM}:\ E_{\varepsilon(\chi)},\varepsilon \sim N(0,1)\mathrm{t}\left[\left\| \varepsilon\text{-}\varepsilon\theta(\mathrm{z_t},\ \mathrm{t}) \right\|_2^2 \right]$$

Conditioning Mechanisms (CMM)

In addition to unconditional image generation, we can also perform conditional image generation, which is mainly achieved by expanding to obtain a conditional denoising autoencoder $\varepsilon\theta(Z_{\mathrm{t}},\ \mathrm{t},\ \mathrm{y})$ so that we can control the process of image synthesis by y. Specifically, the paper achieves this by adding a cross-attention mechanism to the UNet backbone network. In order to preprocess y from several different modalities, the paper introduces a domain specific encoder $\mathrm{T}\theta$, which is used to map y to an intermediate representation $\mathrm{T}\theta(\mathrm{y}) \in R^{M*\mathrm{d_r}}$, so that we can easily introduce various morphological conditions (text, category, lawet, etc.). The final model can then incorporate control information into the intermediate layer of UNet through a cross-attention layer mapping. The implementation of the cross-attention layer is as follows:

$$\mathrm{Attention}(Q,K,V) = \mathrm{softmax}(\frac{QK^T}{\sqrt{\mathrm{d}}}).V, \mathrm{with}$$

$$Q = W^{(\xi)}_{Q}.\varphi\mathrm{i}(\mathrm{z_t}),\ K = W^{(\mathrm{i})}\mathrm{k}.\mathrm{T}\theta(\mathrm{y}), V = W^{(\mathrm{i})}V.\mathrm{T}\theta(\mathrm{y})$$

where $\varphi_{\mathrm{i}}(\mathrm{z_t}) \in R^{N*\mathrm{d}^{\mathrm{i}}_{\mathrm{e}}}$ is an intermediate representation of UNet. The corresponding objective function can be written in the following form:

$$L_{LDM} \coloneqq E\varepsilon(\chi),\mathrm{y},\in\sim N(0,1),\ \mathrm{t}\left[\left\| \in\text{-}\in\theta(\mathrm{z_t},\ \mathrm{t},\ \mathrm{T}\theta(\mathrm{y}) \right\|_2^2 \right]$$

4. Summary

As an open-source and promising AI image generation module, the main difficulty in using Stable Diffusion is the pre-construction of the environment. Once the environment is set up, care should be taken to combine it with appropriate prompt words, as a more detailed description is usually easier to generate better images. If the generated image shows the face or limbs of a person, additional attention should be paid to whether the generated result is normal.

References

[1] ZK. Lin, "A homogeneous and mechanically stable artificial diffusion layer using rigid-flexible hybrid polymer for high-performance lithium metal batteries," Journal of energy chemistry, (1):9, 2023.

[2] YX. Chen, "Magnetically enhancing diffusion for dendrite-free and long-term stable lithium metal anodes," Green Energy & Environment, 7(5):965-974, 2022.

[3] PC. Li, "β-cyclodextrin as Lithium-ion Diffusion Channel with Enhanced Kinetics for Stable Silicon Anode," Green Energy & Environment, (1):9, 2021.

[4] L. Zhou, "Highly Efficient and Stable Planar Perovskite Solar Cells with Modulated Diffusion Passivation Toward High Power Conversion Efficiency and Ultrahigh Fill Factor," Solar RRL,

DOI: 10.54254/2755-2721/18/20230963

3(11), 2019.

[5] L. Pickart, A. Margolina, “Regenerative and protective actions of the GHK-Cu peptide in the light of the new gene data,” International journal of molecular sciences, 19(7): 1987, 2018.