

MATN ORQALI RASM YARATISH, SUN'IY INTELLEKTNING VIZUAL INNOVATSIYALARI

Cho'liiyev Shoxrux Ibadullayevich

Muhammad al-Xorazmiy nomidagi

Toshkent Axborot Texnologiyalari Universiteti

Normo'minov Akbar Kamol o'g'li

Muhammad al-Xorazmiy nomidagi

Toshkent Axborot Texnologiyalari Universiteti

Annotatsiya Matn asosida rasm generatsiya qilish texnologiyasi, sun'iy intellekt yordamida matn tavsiflarini vizual tasvirlarga aylantirish jarayonini o'z ichiga oladi. Bu jarayon asosan, chuqur o'rganish modeli yordamida amalga oshiriladi, bu model matn va tasvir o'rtasidagi murakkab bog'liqliklarni o'rganishga qodir. Ushbu texnologiya ko'pincha generativ raqobatbardosh tarmoqlar (GANs) va transformer kabi ilg'or arxitekturalardan foydalanadi. Rasm generatsiya qilish jarayoni matn kirishlarini qabul qilib, mos rasmlarni yaratish orqali amalga oshiriladi, bu esa modelning matn kontekstini qanchalik yaxshi tushunishiga bog'liq. Texnologiya nafaqat san'at va dizayn sohasida, balki o'yin ishlab chiqarish, ta'lim va ilmiy tadqiqotlar kabi turli sohalarda ham keng qo'llaniladi. OpenAI DALL-E, Google Imagen va Stable Diffusion kabi platformalar bu jarayonni amalga oshirishda yetakchi hisoblanadi. Ushbu texnologiyalar yordamida foydalanuvchilar o'z matnlarini aniq va chiroyli rasmlarga aylantirish imkoniyatiga ega bo'ladi.

Kalit so'z: Text-to-Image Generation, Generative Adversarial Networks (GANs), Transformer Architectures, Loss Functions, Image Synthesis, Technology Applications, OpenAI DALL-E, Google Imagen

Matn orqali rasm generatsiya qilish jarayoni, ya'ni matn asosida rasm yaratish (Text-to-Image Generation) - bu sun'iy intellekt (SI) texnologiyalari yordamida matndagi tasvirlarni vizual tasvirlarga aylantirish jarayonidir. Bu jarayon, odatda, chuqur o'rganish modeli (deep learning models) yordamida amalga oshiriladi. Quyida ushbu texnologiya ishlaydigan asosiy bosqichlar haqida batafsil ma'lumot beraman.

- **Ma'lumotlarni tayyorlash**

- Matn kirishlari: Matn kirishlari, rasm tasvirlash uchun ishlatiladigan ma'lumotlar (masalan, "bir dengiz bo'yida quyosh botayotgan manzara"). Bu matnlar modelga nima yaratish kerakligi haqida ko'rsatma beradi.

- Rasm ma'lumotlari: Tayyorlangan rasmlar to'plami. Bu rasmlar modelni o'qitishda foydalaniladi, shunda u matn tasvirlarini qanday vizual tasvirlarga aylantirishni o'rganadi.

- **Modelni o'qitish**

- Arxitektura: Generativ raqobatbardosh tarmoqlar (GANs) yoki transformer kabi arxitekturalar ishlatiladi. Bu arxitekturalar matn va rasm o'rtasidagi bog'liqlikni tushunish uchun mo'ljallangan.

- Optimizatsiya va yo'qotish funksiyalari: Model, maqsadli rasimga o'xshash rasmlarni generatsiya qilish uchun yo'qotish funksiyasi (loss function) yordamida optimizatsiya qilinadi. Bu jarayon matn ma'lumotlari bilan birga berilgan rasmlarga qanchalik yaqin ekanligini baholash orqali amalga oshiriladi.

- **Generatsiya**

- Matn kirishi: Foydalanuvchi tomonidan berilgan yangi matn kirishi asosida model yangi rasm yaratadi.

- Rasm chiqishi: Model matn tasvirlariga mos keladigan rasmni generatsiya qiladi. Bu jarayon, modelning oldingi o'qitilishi asosida, matnning mohiyatini va kontekstini tushunish qobiliyatiga bog'liq.

- **Optimizatsiya va sozlash**

- Sinov va baholash: Generatsiya qilingan rasmlar foydalanuvchi va mutaxassislar tomonidan ko'rib chiqiladi va baholanadi. Bu orqali modelning samaradorligi o'lchanadi.

- Takomillashtirish: Modelning aniqligini oshirish uchun kerakli sozlamalar amalga oshiriladi. Masalan, yo'qotish funksiyasini yoki arxitekturani sozlash orqali.

Texnologiyalar va Platformalar

- OpenAI DALL-E, Google Imagen va Stable Diffusion kabi platformalar ushbu texnologiyani qo'llab, foydalanuvchilarga matn orqali rasm generatsiya qilish imkonini taqdim etadi. Ushbu platformalar, odatda, katta ma'lumotlar to'plamlari bilan o'qitilgan bo'lib, turli xil matn kirishlariga javob berishda yuqori samaradorlikka ega.

Bu jarayon, nafaqat san'at va dizaynda, balki o'yinlar ishlab chiqarish, ta'lim va hattoki ilmiy tadqiqotlar kabi sohalarda ham qo'llanilishi mumkin.

Adabiyotlar ro'yxati

1. Goodfellow, I., Bengio, Y., Courville, A. (2016). Deep Learning. MIT Press. [Online mavjud: <http://www.deeplearningbook.org>]

2. Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Amodei, D. (2020). Language Models are Few-Shot Learners. In NeurIPS.

3. Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sutskever, I. (2021). Learning Transferable Visual Models From Natural Language Supervision. In ICML.

4. Brock, A., Donahue, J., Simonyan, K. (2019). Large Scale GAN Training for High Fidelity Natural Image Synthesis. In ICLR.