Actividad 1 Modelos Generativos de video y 3D

Diego Iván Perea Montealegre (2185751) [diego.perea@uao.edu.co](mailto:diego.perea@uao.edu.co)

Carlos Iván Osorio Moreno [carlos\_ivan.osorio@uao.edu.co](mailto:carlos_ivan.osorio@uao.edu.co)

Samir Hassan Ordoñez [samir\_hassan@uao.edu.co](mailto:samir_hassan@uao.edu.co)

Luis Fernando Pareja Bernal [luis.pareja@uao.edu.co](mailto:luis.pareja@uao.edu.co)

Gabriel Angel Jeannot viaña [gabriel.jeannot@uao.edu.co](mailto:grabriel.jeannot@uao.edu.co)

Facultad de Ingeniería, Universidad Autónoma de Occidente

Cali, Valle del Cauca

Los modelos de inteligencia artificial para generación de video y 3D han experimentado una tendencia en aumento en los últimos años debido a su capacidad para producir imágenes y animaciones de alta calidad en tiempo real.

Los modelos IA son capaces de generar contenido visual de alta calidad que puede ser utilizado en una amplia gama de aplicaciones, desde entretenimiento hasta publicidad y cine. Sin embargo, aún hay desafíos técnicos y éticos en el uso de estos modelos, incluyendo la necesidad de una mayor diversidad en la generación de imágenes y la preocupación por la privacidad y la seguridad de los datos utilizados para entrenarlos. Además, algunos modelos pueden generar contenido inapropiado o ofensivo sin la supervisión adecuada.

Los modelos de Inteligencia Artificial (IA) para la generación de video y 3D ofrecen ventajas como procesos más eficientes, libres de errores y mayor control sobre las líneas de producción[1]. Por ejemplo, NVIDIA ha presentado un modelo de IA para crear objetos y personajes en 3D[2], mientras que Google ha creado una herramienta para generar vídeos en alta definición con comprensión de objetos[3]

**Impactos positivos**

Aspectos positivos de los modelos de IA de generación de video y 3D:

1. Mejora en la eficiencia y velocidad de producción de contenido visual.
2. Capacidad para generar imágenes y animaciones realistas y de alta calidad.
3. Aplicaciones en una amplia gama de industrias, incluyendo entretenimiento, publicidad, cine y juegos.
4. Reducción de costos y tiempo en la producción de contenido visual.
5. Mejora en la accesibilidad y disponibilidad de contenido visual de alta calidad.
6. Menor contaminación ambiental derivada de la huella de carbono producida por la industria cinematográfica.
7. Disminución de la brecha socioeconómica en la producción de películas y otros tipos de material audiovisual, permitiendo generar contenido de gran calidad a industrias y países con poca relevancia en el séptimo arte.
8. Mayor libertad y enfoque a la parte narrativa de una producción, debido a la reducción de dificultades en la realización de escenas.
9. Crecimiento exponencial de contenido VR/AR para mundos virtuales simulados y sus diferentes aplicaciones en el ámbito laboral, académico y ocio.

**Impactos negativos**

Aspectos negativos de los modelos de IA de generación de video y 3D:

1. Preocupaciones sobre la privacidad y seguridad de los datos utilizados para entrenar los modelos.
2. Desafíos éticos en el uso de modelos que generan contenido inapropiado u ofensivo.
3. Necesidad de supervisión humana para asegurar la calidad y la apropiabilidad del contenido generado.
4. Falta de diversidad en la generación de imágenes y animaciones.
5. La posibilidad de reemplazar trabajos humanos y afectar negativamente a la industria de la producción de contenido visual.

**Aplicaciones**

1. Modelos Generativos para descubrimiento molecular: El desarrollo de nuevos productos a menudo se basa en el descubrimiento de nuevas moléculas. Si bien el diseño molecular convencional implica el uso de la experiencia humana para proponer, sintetizar y probar nuevas moléculas, este proceso puede ser costoso y requiere mucho tiempo, lo que limita la cantidad de moléculas que se pueden probar razonablemente. El modelado generativo proporciona un enfoque alternativo al descubrimiento molecular al reformular el diseño molecular como un problema de diseño inverso[4].
2. Interpolación de fotogramas generados por IA: La predicción de video genera fotogramas futuros de un video en función de una secuencia dada de fotogramas pasados. Las GAN se han utilizado para desarrollar sistemas que pueden generar cuadros de video realistas y de alta calidad que predicen con precisión la evolución futura de la escena. [5]. En el portal de [Paperswithcode](https://paperswithcode.com/task/video-prediction), se presentan diferentes modelos de predicción de videos.
3. Los modelos de difusión de imágenes y videos : Los modelos de difusión de imágenes y videos basados en texto han logrado recientemente un realismo de generación sin precedentes. Si bien los modelos de difusión se han aplicado con éxito para la edición de imágenes, muy pocos trabajos lo han hecho para la edición de video [6]. Los modelos de difusión son una nueva clase de modelos generativos que generan diversas imágenes de alta resolución[7]. Estos modelos replican el proceso de añadir ruido a una imagen hasta tener una imagen deseada[8]. Las GAN, por otro lado, producen imágenes a partir del ruido[9].
4. El renderizado neuronal es un método de creación de nuevas imágenes y secuencias de vídeo a partir de escenas existentes, que combina técnicas de aprendizaje automático generativo con conocimientos físicos procedentes del ordenador[10][11]. Neural Radiance Fields (NeRF) es un tipo de modelo de renderizado neuronal que puede aprender una escena 3D de alta resolución en segundos y renderizar imágenes de ella en milisegundos[12][13]

* **Posibles impactos:**
  + Posible eliminación del estándar 24fps para cine.
  + Crecimiento de usuarios hacia tecnologías de VR/AR.
  + Aumento de costos en la producción de paneles (o bombillas) de equipos de imagen y video como televisores, proyectores de imagen, etc.
  + Disminución de costos en la producción de procesadores para equipos de imagen y video.
  + Menor capacidad de cómputo requerida involucrada en la proyección de contenido multimedia.
  + Mayor facilidad para estandarizar altas resoluciones de imagen.
  + Edición de video con superior simplicidad.
  + Incremento de creatividad.
  + Aumento de accesibilidad.
  + Aumento de creación de nuevos empleos

Referencias

[1] NexusAdmistraIntegra, “Inteligencia artificial (IA): Ventajas y desventajas de su uso,” Nexus Integra, 26-Jul-2022. [Online]. Available: https://nexusintegra.io/es/ventajas-y-desventajas-de-la-inteligencia-artificial/.

[2] J. Marquez, “Nvidia Acaba de presentar UN Modelo de ia Para Crear objetos y personajes en 3d: Una Nueva Herramienta Para construir El Metaverso,” Xataka, 23-Sep-2022. [Online]. Available: https://www.xataka.com/robotica-e-ia/nvidia-acaba-presentar-modelo-ia-para-crear-objetos-personajes-3d-otra-herramienta-para-construir-metaverso.

[3] B. Bécares, “La Nueva Ia de Google Crea vídeos en Alta Definición de Nuestras Sugerencias. similar a dall-e 2 pero en vídeo,” Genbeta, 06-Oct-2022. [Online]. Available: https://www.genbeta.com/actualidad/nueva-ia-google-genera-videos-alta-definicion-a-traves-descripcion-que-le-hagas-similar-a-dall-e-2-video.

[4] Bilodeau, C. et al. (2022) “Generative models for Molecular Discovery: Recent advances and challenges,” WIREs Computational Molecular Science, 12(5). Available at: <https://doi.org/10.1002/wcms.1608>.

[5] Simplilearn. "List Of Generative Adversarial Networks Applications | Simplilearn". Simplilearn.com. <https://www.simplilearn.com/generative-adversarial-networks-applications-article>

[6] “Video diffusion models are general video editors,” Dreamix. [Online]. Available:<https://dreamix-video-editing.github.io/?ref=futurepedia>. [Accessed: 07-Feb-2023]

[7] B. Bustos, “≫ Cómo funcionan los modelos de difusión: Las matemáticas desde cero,” La Ciencia de Jaun, 25-Oct-2022. [Online]. Available: <https://lacienciadejaun.com/como-funcionan-los-modelos-de-difusion-las-matematicas-desde-cero>. [Accessed: 07-Feb-2023]

[8] “Cómo funciona stable diffusion - y dalle-2 O midjourney -: Open source generative-ia que crea imágenes artísticas (1 de 2),” Cómo funciona Stable Diffusion - y Dalle-2 o Midjourney -: Open Source Generative-IA que crea imágenes artísticas (1 de 2). [Online]. Available: <https://www.elladodelmal.com/2022/10/como-funciona-stable-diffusion-y-dalle.html>. [Accessed: 07-Feb-2023]

[9] Jay, “Modelo de Difusión vs gan,” HashDork, 30-Aug-2022. [Online]. Available: <https://hashdork.com/es/diffusion-model-vs-gan/>. [Accessed: 07-Feb-2023]

[10] “Video diffusion models are general video editors,” Dreamix. [Online]. Available: <https://dreamix-video-editing.github.io/?ref=futurepedia>. [Accessed: 07-Feb-2023]

[11] A. Tewari, O. Fried, J. Thies, V. Sitzmann, S. Lombardi, K. Sunkavalli, R. Martin-Brualla, T. Simon, J. Saragih, M. Nießner, R. Pandey, S. Fanello, G. Wetzstein, J.-Y. Zhu, C. Theobalt, M. Agrawala, E. Shechtman, D. B. Goldman, and M. Zollhöfer, “State of the art on neural rendering,” arXiv.org, 08-Apr-2020. [Online]. Available: <https://arxiv.org/abs/2004.03805>. [Accessed: 07-Feb-2023]

[12] “Nerf,” Neural Radiance Fields. [Online]. Available: <https://www.matthewtancik.com/nerf>. [Accessed: 07-Feb-2023]

[13] I. Salian, “Nerf research turns 2D photos into 3D scenes,” NVIDIA Blog, 15-Nov-2022. [Online]. Available: <https://blogs.nvidia.com/blog/2022/03/25/instant-nerf-research-3d-ai/>. [Accessed: 07-Feb-2023]