Publicado en Ciudad de México el 24/06/2021

# [Más allá de la IA: Cómo la Investigación con las GAN está Remodelando las Videoconferencias](http://www.notasdeprensa.es)

## El SDK NVIDIA Maxine de streaming de video con IA en cloud se basa en la investigación de las GAN y se exhibirá en CVPR 2021

 Byline: Isha Salian. "Sales de la cama, enciendes la laptop, activas la cámara web y disfrutas de la perfección en cada videollamada, con la ayuda de la IA desarrollada por los investigadores de NVIDIA". Vid2Vid Cameo , uno de los modelos de deep learning detrás del SDK de NVIDIA Maxine para videoconferencias, utiliza redes generativas antagónicas, conocidas como GAN, para sintetizar videos realistas de cabezas parlantes utilizando una sola imagen 2D de una persona. Eso significa que un asistente a una reunión con el cabello desarreglado y la pijama puesta puede aparecer como si estuviera vestido con una camisa formal, mientras la IA mapea los movimientos faciales del usuario con una foto previamente cargada del individuo, con un atuendo apropiado para el trabajo. Si el sujeto está mirando hacia la izquierda, la tecnología puede ajustar el punto de vista para que el asistente parezca estar haciendo contacto visual con la cámara web. Además de ayudar a los asistentes a las reuniones a lucir lo mejor posible, esta técnica de IA también reduce el ancho de banda necesario para las videoconferencias hasta en 10 veces, lo que evita la fluctuación y el retraso en la recepción de la imagen. Pronto estará disponible en el SDK de NVIDIA Video Codec como el Códec AI Face. “Muchas personas tienen un ancho de banda de Internet limitado, pero aun así quieren tener una videollamada sin problemas con amigos y familiares", dijo el investigador de NVIDIA Ming-Yu Liu, coautor del proyecto. "Además de ayudarlos, esta tecnología también podría usarse para ayudar al trabajo de animadores, editores de fotos y desarrolladores de juegos". Vid2Vid Cameo se presentó esta semana en la prestigiosa Conferencia sobre Visión de Computación y Reconocimiento de Patrones, uno de los 28 documentos de NVIDIA en el evento virtual. La IA se Roba el ShowEn un guiño a las películas clásicas de atracos (y un exitoso programa de Netflix), los investigadores de NVIDIA pusieron su modelo GAN de cabeza parlante para que participe de una reunión virtual. La demostración destaca las características clave de Vid2Vid Cameo, incluida la redirección facial, los avatares animados y la compresión de datos. Liga de la grabación del Pre Briefing: https://drive.google.com/drive/folders/1p\_P-pes9unkav8R568iGSZhhgwU4P-Tp?usp=sharing Estas capacidades llegarán pronto al SDK de NVIDIA Maxine, que ofrece a los desarrolladores modelos optimizados previamente entrenados para efectos de video, audio y realidad aumentada en videoconferencias y transmisión en vivo. Los desarrolladores ya pueden adoptar los efectos de Maxine AI, incluida la eliminación inteligente del ruido, la ampliación de video y la estimación de la postura corporal. El SDK de descarga gratuita también se puede combinar con la plataforma NVIDIA Jarvis para aplicaciones de IA conversacionales, incluidas la transcripción y la traducción. Saludo desde el Lado de la IAVid2Vid Cameo requiere sólo dos elementos para crear una cabeza parlante de IA realista para videoconferencias: una sola toma de la apariencia de la persona y una transmisión de video que dicta cómo se debe animar esa imagen. El modelo se entrenó en un sistema NVIDIA DGX, utilizando un conjunto de datos de 180,000 videos de cabezas parlantes de alta calidad. La red aprendió a identificar 20 puntos clave que se pueden utilizar para modelar el movimiento facial sin anotaciones humanas. Los puntos codifican la ubicación de los rasgos como los ojos, la boca y la nariz. A continuación, se extrae estos puntos clave de una imagen de referencia de la persona que llama, que puede enviarse a otros participantes de la videoconferencia previamente o volver a utilizarse tomada de reuniones anteriores. De esta manera, en lugar de enviar voluminosos streamings de video en vivo de un participante a otro, las plataformas de videoconferencia pueden simplemente enviar datos sobre cómo se mueven los puntos faciales clave del orador. Del lado del receptor, el modelo GAN utiliza esta información para sintetizar un video que imita la apariencia de la imagen de referencia. Al comprimir y enviar solo la posición de la cabeza y los puntos clave hacia adelante y hacia atrás, en lugar de secuencias de video completas, esta técnica puede reducir las necesidades de ancho de banda para las videoconferencias hasta en 10 veces, lo que proporciona una experiencia de usuario más fluida. El modelo se puede ajustar para transmitir una cantidad diferente de puntos clave para adaptarse a diferentes entornos de ancho de banda sin comprometer la calidad visual. El punto de vista del video de la cabeza parlante resultante también se puede ajustar libremente para mostrar al usuario desde un perfil lateral o recto, así como desde un ángulo de cámara más bajo o más alto. Esta función también la pueden aplicar los editores de fotografías que trabajan con imágenes fijas. [embed GTC fall vid] Los investigadores de NVIDIA descubrieron que Vid2Vid Cameo supera a los modelos de vanguardia al producir resultados más realistas y nítidos, ya sea que la imagen de referencia y el video sean de la misma persona, o cuando la IA tiene la tarea de transferir el movimiento de una persona a una imagen de referencia de otra. Esta última característica se puede utilizar para aplicar los movimientos faciales de un orador a fin de animar un avatar digital, prestando expresión y movimiento realistas a una figura de dibujos animados. El documento detrás de Vid2Vid Cameo fue escrito por los investigadores de NVIDIA Ting-Chun Wang, Arun Mallya y Ming-Yu Liu. El equipo de investigación de NVIDIA está formado por más de 200 científicos de todo el mundo y se centra en áreas como la IA, la Visión de Computación, los Vehículos Autónomos, la Robótica y los Gráficos. NVIDIA desea agradecer y dar crédito al actor Edan Moses (voz en off en inglés del Profesor en “La Casa De Papel” / ”Money Heist” en Netflix) por su contribución en este video con su última investigación de IA.

**Datos de contacto:**

Carlos Valencia

MKQ PR Agency

55 39 64 96 00

Nota de prensa publicada en: [https://www.mexicopress.com.mx/mas-alla-de-la-ia-como-la-investigacion-con](http://www.notasdeprensa.es/educalivecom-disfruta-de-cursos-presenciales-y-clases-particulares-a-traves-de-internet)

Categorías: Imágen y sonido Telecomunicaciones Inteligencia Artificial y Robótica Cine Comunicación Programación Hardware E-Commerce Software Ciudad de México



[**http://www.mexicopress.com.mx**](http://www.notasdeprensa.es)