

Herramienta «Glaze» para protege imágenes de la IA

Investigadores de la Universidad de Chicago han desarrollado «Glaze», una herramienta diseñada para proteger imágenes de ser usadas por modelos de inteligencia artificial sin permiso. Glaze agrega capas casi invisibles sobre imágenes y fotografías, confundiendo a los algoritmos de IA para prevenir la replicación no autorizada del estilo de los artistas. Aunque la herramienta no es una solución definitiva contra el mimetismo de la IA, representa un paso importante hacia la protección centrada en el artista.

En un esfuerzo por combatir la creciente amenaza de la imitación de estilos mediante modelos de inteligencia artificial generativa, se ha lanzado un nuevo sistema llamado Glaze. Este innovador sistema utiliza algoritmos de aprendizaje automático para introducir cambios sutiles en las obras de arte, que aunque imperceptibles para el ojo humano, son detectados por los modelos de IA como estilos artísticos completamente diferentes.

Desarrollado por un equipo de investigadores en aprendizaje automático adversarial, Glaze se propone como una solución robusta frente a la proliferación de falsificaciones y copias de baja calidad que plagian el mercado y afectan negativamente la carrera y reputación de los artistas independientes. Los artistas han expresado su preocupación por cómo la imitación de estilos no solo reduce sus ingresos y oportunidades, sino que también desvirtúa su identidad creativa, comparando este fenómeno con un tipo de robo de identidad.

«Glaze actúa como una nueva dimensión del arte, visible solo para los modelos de IA, similar a cómo los humanos no perciben la luz ultravioleta,» explica Ben, líder del proyecto Glaze. «Esto asegura que cuando se solicita a un modelo de IA generar

arte en el estilo de un artista específico, el resultado sea significativamente diferente de lo esperado, protegiendo así la autenticidad del estilo original del artista.»

Sin embargo, Glaze enfrenta desafíos, ya que la tecnología de IA es rápidamente evolutiva y las estrategias actuales podrían ser superadas por algoritmos futuros. A pesar de estos riesgos, el equipo ha diseñado Glaze para ser lo más resistente posible y ha realizado extensas pruebas contra los sistemas y contramedidas conocidas.

En respuesta a la demanda de una solución más accesible, el equipo también lanzó WebGLaze en agosto de 2023. Esta plataforma web permite a los artistas aplicar Glaze desde cualquier dispositivo con navegador, como teléfonos móviles y tabletas, usando servidores GPU financiados por subvenciones de investigación y situados en la nube de Amazon AWS.

«WebGlaze nació de la retroalimentación directa de la comunidad artística, que trabaja cada vez más desde dispositivos móviles y no siempre tiene acceso a recursos informáticos avanzados», añadió Ben.

Glaze no se ofrece como una solución completa ni permanente contra la imitación por IA, sino como un paso importante hacia herramientas de protección centradas en el artista. Además, al no ser de código abierto, se aumenta la barrera contra ataques adaptativos, asegurando un nivel de protección elevado para las obras de arte en un mundo cada vez más dominado por la tecnología generativa.

Mientras tanto, la comunidad artística acoge con entusiasmo esta nueva herramienta, esperanzada en que Glaze y futuros proyectos proporcionen la protección necesaria para preservar la integridad y la pasión que define al arte humano en la era digital.

Más detalles en el sitio web <https://glaze.cs.uchicago.edu/download.html>