Vision AI

**Extrae información valiosa de imágenes, documentos y vídeos**

Accede a modelos de visión avanzados a través de APIs para automatizar tareas de visión, optimizar los análisis y obtener métricas útiles. También puedes crear aplicaciones personalizadas sin código de preparación de modelos y de bajo coste en un entorno gestionado.

Información general

¿Qué es la visión artificial?

La visión artificial es un campo de la inteligencia artificial (IA) que permite a los ordenadores y sistemas interpretar y analizar datos visuales y obtener información pertinente a partir de imágenes digitales, vídeos y otros datos visuales. Algunas de sus aplicaciones habituales en el mundo real son la detección de objetos, el procesamiento de contenido visual (imágenes, documentos, vídeos), la comprensión y el análisis, la búsqueda de productos, la clasificación y búsqueda de imágenes y la moderación de contenido.

IA generativa multimodal avanzada

Vertex AI de Google Cloud ofrece acceso a [Gemini](https://cloud.google.com/vertex-ai/docs/generative-ai/start/quickstarts/quickstart-multimodal), una familia de modelos multimodales y vanguardistas que pueden comprender prácticamente cualquier tipo de entrada, combinar diferentes tipos de información y generar casi cualquier salida. Aunque Gemini es más adecuado para tareas que combinan elementos visuales, texto y código, [Gemini Pro Vision](https://cloud.google.com/vertex-ai/docs/generative-ai/multimodal/overview#visual_understanding) destaca en una gran variedad de tareas relacionadas con la visión, como el reconocimiento de objetos, la comprensión de contenido digital, los subtítulos y las descripciones. Se puede acceder a él a través de una API.

IA generativa centrada en la visión

[Imagen en Vertex AI](https://cloud.google.com/vertex-ai/docs/generative-ai/image/overview) ofrece a los desarrolladores de aplicaciones las funciones vanguardistas de IA generativa de imágenes de Google a través de una API. Algunas de sus [funciones clave](https://cloud.google.com/vertex-ai/docs/generative-ai/image/overview#feature-launch-stage) son la [generación de imágenes](https://cloud.google.com/vertex-ai/docs/generative-ai/image/generate-images) (GA restringido)con peticiones de texto y la [edición de imágenes](https://cloud.google.com/vertex-ai/docs/generative-ai/image/edit-images) (GA restringido) con peticiones de texto, la [descripción de una imagen en texto](https://cloud.google.com/vertex-ai/docs/generative-ai/image/image-captioning) (también llamada "descripción visual", GA) y ajuste del modelo sujeto (GA restringido).

Vision AI lista para usar

La [API de Cloud Vision](https://cloud.google.com/vision/docs), que usa la tecnología de Google de modelos de aprendizaje automático de visión artificial entrenados previamente, es una API de disponibilidad inmediata (REST y RPC) que permite a los desarrolladores integrar fácilmente [funciones](https://cloud.google.com/vision/docs/features-list) habituales de detección de visión en sus aplicaciones, como etiquetado de imágenes, detección de caras y puntos de referencia, reconocimiento óptico de caracteres (OCR) y etiquetado de contenido explícito.

IA generativa para interpretar documentos

[Document AI](https://cloud.google.com/document-ai) es una plataforma de comprensión de documentos que combina la visión artificial y otras tecnologías, como el procesamiento del lenguaje natural, para extraer texto y datos de documentos escaneados y transformar los datos no estructurados en información estructurada y estadísticas empresariales.

Ofrece [una amplia gama de procesadores entrenados previamente](https://cloud.google.com/document-ai/docs/processors-list#processor_cde) optimizados para diferentes tipos de documentos. También facilita la creación de procesadores personalizados para clasificar, dividir y extraer datos estructurados de documentos mediante [Document AI Workbench](https://cloud.google.com/document-ai-workbench" \t "_blank).

Vision AI lista para usar en vídeos

La [API de Video Intelligence](https://cloud.google.com/video-intelligence), centrada en la tecnología de visión artificial, permite procesar, analizar y comprender fácilmente el contenido de vídeo.

Sus modelos de aprendizaje automático preentrenados reconocen automáticamente una gran cantidad de objetos, lugares y acciones en vídeos almacenados y en streaming con una calidad excepcional. Es muy eficaz en [casos prácticos habituales](https://cloud.google.com/video-intelligence#use-cases), como la moderación y las recomendaciones de contenido, los archivos multimedia y los anuncios contextuales. También puedes entrenar modelos de aprendizaje automático personalizados con Vertex AI Vision según tus necesidades.

Vision AI lista para usar en la búsqueda de productos en el comercio

[Product Search de la API de Vision](https://cloud.google.com/vision/product-search/docs) es un servicio especializado dentro del paquete de herramientas de Google Cloud Vision AI que permite a los usuarios buscar productos con sus propias imágenes. Se puede interpretar como un buscador de imágenes optimizado para productos. Admite las siguientes categorías de productos: artículos para el hogar, ropa, juguetes, productos de gran consumo y generales.

Visual Inspection AI

[Visual Inspection AI](https://cloud.google.com/solutions/visual-inspection-ai) automatiza las tareas de inspección visual en fábricas y otros entornos industriales. Utiliza técnicas avanzadas de visión artificial y aprendizaje profundo para analizar imágenes y vídeos, identificar anomalías, detectar y localizar defectos y comprobar piezas que falten o que tengan defectos en los productos ensamblados.

Puedes entrenar modelos personalizados sin conocimientos técnicos y usar un mínimo de imágenes etiquetadas, ejecutar inferencias de forma eficiente en las líneas de producción y actualizar continuamente los modelos con datos actualizados de la planta de producción.

Vision AI Platform unificada

[Vertex AI Vision](https://cloud.google.com/vision-ai/docs) es un entorno de desarrollo de aplicaciones totalmente gestionado que permite a los desarrolladores crear, desplegar y gestionar fácilmente aplicaciones de visión artificial para procesar [varias modalidades de datos](https://cloud.google.com/vertex-ai/docs/training-overview#types_of_models_you_can_build_using), como texto, imágenes, vídeo y datos tabulares. Además, reduce el tiempo de desarrollo de días a minutos por una décima parte del coste de los productos actuales.

Puedes crear y desplegar tus propios modelos personalizados, así como gestionarlos y escalarlos con flujos de procesamiento de CI/CD. Además, se integra con herramientas populares de software libre, como TensorFlow y PyTorch.

Privacidad y seguridad de los datos

Google Cloud cuenta con funciones punteras en el sector que permiten a nuestros clientes controlar sus datos y saber cuándo y cómo se accede a ellos.

Como cliente de Google Cloud, tú eres el propietario de tus datos de clientes Por tanto, implementamos unas medidas de seguridad estrictas para proteger tus datos de clientes y te proporcionamos las herramientas y funciones para que puedas controlar esos datos como desees. Tú eres el propietario de los datos de clientes, no Google. A la hora de tratar tus datos, seguimos estrictamente tus contratos.

Usos habituales:  
  
1) Detecta texto en archivos sin procesar y resúmelo automáticamente

Resume documentos extensos con IA generativa

La solución que se muestra en el diagrama de arquitectura de la derecha despliega un flujo de procesamiento que se activa cuando añades un nuevo documento PDF a tu segmento de Cloud Storage. El flujo de procesamiento extrae el texto del documento, crea un resumen a partir del texto extraído y almacena el resumen en una base de datos para que puedas verlo y buscarlo.

Puedes invocar la aplicación subiendo archivos a través de Jupyter Notebook o directamente a Cloud Storage en la consola de Google Cloud.

2) Búsqueda de productos basada en imágenes para comercio electrónico

Pónselo fácil a tus clientes a la hora de descubrir productos

[Product Search de la API de Vision](https://cloud.google.com/vision/product-search/docs?_gl=1*mv2n8h*_ga*MTEwNzAxMzI4MC4xNzAwMDAzMzAz*_ga_4LYFWVHBEB*MTcwNjIxNjczMS40My4xLjE3MDYyMTY4MDIuMC4wLjA.) permite a los comercios crear productos, cada uno con imágenes de referencia que describen visualmente los productos desde varios puntos de vista. Después, los comercios pueden añadir esos productos a sus conjuntos de productos.

Cuando los usuarios consultan el conjunto de productos con sus propias imágenes, Product Search de la API de Vision aplica el aprendizaje automático para comparar el producto que aparece en la imagen de la consulta del usuario con las imágenes del conjunto de productos del comercio y, a continuación, devuelve una lista ordenada de resultados similares visual y semánticamente.

3) Crea un flujo de procesamiento de imágenes

Procesamiento de imágenes escalable en una arquitectura sin servidor

La solución, que se muestra en el diagrama de la derecha, utiliza modelos de aprendizaje automático preentrenados para analizar imágenes proporcionadas por los usuarios y generar anotaciones de imágenes. Al desplegar esta solución, se crea un servicio de procesamiento de imágenes que puede ayudarte a gestionar contenido peligroso o dañino creado por los usuarios, digitalizar texto de documentos físicos, detectar y clasificar objetos en imágenes y mucho más.

Podrás revisar los ajustes de configuración y seguridad para saber cómo adaptar el servicio de procesamiento de imágenes a distintas necesidades.

4) Consigue descripciones de imágenes automatizadas con la IA generativa

La función de [descripción visual](https://cloud.google.com/vertex-ai/docs/generative-ai/image/image-captioning) de [Imagen](https://cloud.google.com/vertex-ai/docs/generative-ai/image/overview) te permite generar una descripción pertinente de una imagen. Puedes utilizarla para obtener metadatos más detallados sobre las imágenes que puedes almacenar y buscar, o para crear subtítulos automáticos. en casos prácticos de accesibilidad y recibir descripciones rápidas de productos y recursos visuales.

Esta función está disponible en alemán, español, francés, inglés e italiano. Puedes acceder a ella en la consola de Google Cloud o a través de una llamada a la API.

5) Vídeos procesados en streaming

Extrae información valiosa de vídeos en streaming con Vertex AI Vision

Antes de analizar los datos de vídeo con tu aplicación, crea un flujo de procesamiento para el flujo continuo de datos con el servicio Streams de Vertex AI Vision. Los datos ingeridos los analizan los modelos entrenados previamente de Google o tu modelo personalizado. Los resultados del análisis de los flujos se almacenan en Vertex AI Vision Warehouse, donde puedes usar funciones avanzadas de búsqueda basadas en IA para consultar contenido multimedia sin estructurar.

6) Extrae texto y estadísticas de documentos con IA generativa

Descubre información valiosa a partir de documentos específicos con Document AI

[El extractor personalizado de Document AI](https://cloud.google.com/document-ai/docs/processors-list#processor_cde) se basa en un modelo básico que extrae texto y datos de documentos genéricos y específicos de un dominio, de forma más rápida y precisa. Ajusta fácilmente entre 5 y 10 documentos para mejorar el rendimiento.

 Si quieres entrenar tu propio modelo, [etiqueta automáticamente](https://cloud.google.com/document-ai/docs/workbench/build-custom-processor#use_generative_ai_to_auto-label_documents) tus conjuntos de datos con el modelo básico para agilizar la producción.

También puedes utilizar procesadores especializados entrenados previamente. Consulta la [lista completa de procesadores](https://cloud.google.com/document-ai/docs/processors-list).

7) Inspección visual de alta precisión

Automatiza las inspecciones de calidad con Visual Inspection AI

[Visual Inspector AI](https://cloud.google.com/solutions/visual-inspection-ai) se optimiza en cada paso para que puedas configurar el dispositivo fácilmente y comprobar el retorno de la inversión en un abrir y cerrar de ojos. Gracias a que ofrece hasta 300 veces menos imágenes etiquetadas para empezar a entrenar modelos de inspección de alto rendimiento en comparación con las plataformas de aprendizaje automático de uso general, ha demostrado que ofrece una precisión hasta 10 veces mayor. Puedes entrenar modelos que se ejecutan on-premise sin conocimientos técnicos. Y lo mejor de todo es que los modelos se pueden actualizar continuamente a medida que fluyen datos desde la fábrica, lo que ofrece una mayor precisión a medida que descubres nuevos casos prácticos.