22/11/2023

Tema 2: Aplicaciones de la IA en tareas de visión por computador

Visión por computador

La visión por computador consisten en darle ojos a la máquina. Este campo aplica tanto imágenes como videos (un video es una serie de imágenes).

Para procesar videos necesitas modelos que funcionen en tiempo real (un ocche no se puede seperar un mintuo a procesar la imagen)

Con la vision por computador se puede dar información sobre el ambiente al modelo

La redes convolucionales que se utilizan para esta tarea se inventaron hacen mucho tiempo pero están en ague, gracias a la mejora del HW y mayor cantidad de datos.

Como resolver problemas de vision por computador

* Manual: hacer una serie de reglas
* Machine Learning tradicional: modelos que no son redes neuronales. Necestian una base de datos estructuradas para aprender. Sin embargo, los modelos son muy sensibles a variaciones de la imágenes.
* Deep learning: supone un gran avance. Utiliza redes neuronales.

Tipos de aprendizaje

* Supervisado: Utiliza datos etiquetados. Para cada registro de la DB sabemos a que clase pertenece
* No supervisado: datos no etiquetados. Permite detectar patrones, agrupándolos en clusters.

Aprendizaje superivsador

* Clasificación/Regresión
* Detección de objetos
* Segmentación
  + Segmentación de instancias: dentro de la misma clase se diferencias objetos
  + Segmentación semántica: los objetos de la mismsa clase son iguales

Aprendizaje no supervisado

* Reducción de ruido
* Detección de anomalías: se usa en banca para detectar fraudes
* Generación de imágenes sintéticas
* Transformación de imágenes

Aprendizaje semi-supervisado

Cierta parte del conjunto de datos esta etiquetado. Otro conjunto de datos sin etiquetar (aunque siempre pertenecerá a los de la misma clase)

1. Se hace un aprendizaje supervisado con los datos etiquetados
2. Se utiliza el modelo generado para etiquetar los datos no etiquetaods. Las etiquetas que generan se llaman etqiuetas débiles
3. Entrenar el modelo final utilizando las pseudo-etiquetas

Auto-aprendizaje

Es un tipo de aprendizaje no supervisado. Las etiquetas se extraen de los propios datos.

Ejimplo: ChatGPT

* Coger texto de internet, quitarles la ultima palabra y entrenar el modelo para predecir la ultima palabra.

Ejemplo: Impainting

* Eliminar elementos que obstaculizan una foto

Uno de los motivos por los que el auto-aprendizaje es tan bueno es porque es barato, porque no se necesitan inventir tanto en datos.

Clasificación

Dada una imagen de entrada se le asigna una etiqueta

* MNIST: bases de datos pública de dígitos escritos a mano
* ImageNet: Clasificación de imágenes de la vida cotidiana
  + Lanzaban retos abiertos anuales
  + En 2015 se supero la habilidad humana

Detección de objetos

* Venta al público (marketing para detectar si un escaparate se para mucha gente o no)
* Coches autónomos
* Monitorización de animales
* Detección de personas en peligro (persona en vía de tren)
* Detección del tráfico
* Detección de lesiones en imágenes médicas

La base de dato abiertas más utilizada es COCO. Tiene objetos ya segmentados

Segmentación

Identificar todos los píxeles que pertenecen a un objeto en concreto

A través del aprendizaje no supervisado puedes pasar de una zebra a un caballo o de invierno a verano.

Aplicaciones como DeOldify para convertir imágenes a color

IA generativa

* Redes generativas adversarias (GANs): una GAN tiene dos modelos: un generador y un discriminador. El generador hace las imágenes y el discriminador sabe si una imagen es real o no
* Modelos de difusión: transformar ruido a una imagen y al revés.

¿Para qué se pueden emplear estos modelos?

Reducir sesgos en las bases de datos. Ser capaces de replicar situaciones inusuales para que el modelo aprenda mucho mejor

HuggingFace

Comunidad donde se pueden encontrar modelos de código abierto. También te pueden encontrar modelos entrenados con muchos datos lo que se denomina modelos fundacionales.