

Cuestionario 3 VC.

Alberto Armijo Ruiz. GII.

1. ¿Cuáles son las propiedades esenciales que permiten que los modelos de recuperación de instancias de objetos de una gran base de datos a partir de descriptores sean útiles? Justificar la respuesta.

La característica esencial de estos modelos es que al contar una base de datos grande, podemos encontrar características que se aproximen bastante a las características del objeto que queremos reconocer.

2. ¿Justifique el uso del modelo de bolsa de palabras en el proceso de detección y reconocimiento de instancias de objetos? ¿Qué ganamos?, ¿Qué perdemos? Justificar la respuesta

El modelo de bolsa de palabras nos permite generalizar un conjunto de descriptores en uno solo, el cuál se le llama palabra visual. Una de las cosas que ganamos es que a la hora de obtener parecidos entre los descriptores obtenidos y los descriptores que tengamos en la base de datos, solo necesitamos comparar con la palabra visual y no con todos los descriptores que pueden encontrarse en el espacio de esa palabra visual. Además de lo anterior, los modelos de bolsa de palabras son más flexibles a las deformaciones, cambios en el punto de vista, etc...

Los inconvenientes de este modelo es que no está claro el tamaño del vocabulario que se debe utilizar, otro inconveniente es que este modelo ignora la geometría, por lo que hay que añadir más complejidad al modelo para que este no la ignore.

3. ¿Describa la diferencia esencial entre los problemas de reconocimiento de instancias y reconocimiento de categorías? ¿Qué deformaciones se presentan en uno y otro? Justificar la respuesta

La principal diferencia que hay entre el reconocimiento de instancias y el reconocimiento de categorías es que en el reconocimiento de instancias tratamos de reconocer un único objeto con unas características determinadas y en el reconocimiento de categorías tratamos de reconocer un conjunto de instancias diferentes que representan al mismo objeto, lo cual tiene mucha más variabilidad que el reconocimiento de una instancia. En el reconocimiento de instancias puede haber cambios de luz, el lugar donde se encuentra el objeto en la imagen, el punto de vista. En el reconocimiento de categorías además de las anteriormente mencionadas también puede haber cambios dentro de la geometría del objeto (no tiene la misma forma).

4. ¿Es posible usar el modelo de bolsa de palabras para el reconocimiento de categorías de objetos? Justificar la contestación

Sí que es posible utilizar el modelo de bolsa de palabras para el reconocimiento de objetos, para ello, podemos utilizar los histogramas de las imágenes para que en vez de que voten por una instancia sola, voten por una categoría.

5. Suponga que desea detectar, en una imagen, una instancia de un objeto a partir de una foto del mismo tomada desde el mismo punto de vista del que aparece en la imagen y en un entorno de iluminación similar. Analice la situación en el contexto de las técnicas de reconocimiento de objetos e identifique qué algoritmo concreto aplicaría que fuese útil para cualquier objeto. Argumente porqué funcionaría y especifique los detalles necesarios que permitan entender su funcionamiento.

Para esta situación, podríamos utilizar los descriptores del objeto para el reconocimiento de este. Para ello, lo primero que deberíamos hacer es obtener puntos de interés dentro de la zona donde se encuentra el objeto y normalizarlos, tras esto deberíamos de obtener descriptores sobre esas zonas de interés. Una vez hemos obtenido esto (hemos entrenado el modelo), debemos de obtener los descriptores de la imagen y buscar si dentro de la imagen existen descriptores que sean iguales o muy parecidos a los que tenemos dentro de nuestro modelo, si esto es así, reconoceríamos el objeto. Este modelo funcionaría debido a que para imágenes donde aparezca el objeto con situaciones parecidas (iluminación, punto de vista,...), los descriptores que se obtendrán en ambas imágenes del objeto serán muy parecidas.

6. Suponga de nuevo el problema del ejercicio anterior pero la foto que le dan está tomada con un punto de vista del objeto distinto respecto del objeto en la imagen. Analice qué repercusiones introduce esta modificación en su solución anterior y que cambios debería de hacer para volver a tener un nuevo algoritmo exitoso. Justificar la respuesta.

Esto supondría un problema, ya que podría ocurrir que obtengamos descriptores muy parecidos a los del objeto que queremos reconocer, pero siendo un objeto diferente. Para solucionar este problema, debemos analizar si las correspondencias entre las parejas de descriptores son realmente consistentes, para ello, podemos utilizar el algoritmo de RANSAC o la transformada generalizada de Hough. Si tras aplicar alguno de estos métodos, el número de correspondencias entre los descriptores del objeto que buscamos y los descriptores de la imagen sigue siendo lo suficientemente grande, diremos que en la imagen de entrada está el objeto que estamos reconociendo.

7. Suponga que una empresa de Granada le pide implementar un modelo de recuperación de información de edificios históricos de la ciudad a partir de fotos de los mismos. Explique de forma breve y clara que enfoque le daría al problema. Que solución les propondría. Y como puede garantizar que la solución podrá ser usada de forma eficiente a través de dispositivos móviles.

8. Suponga que desea detectar la presencia/ausencia de señales de tráfico en imágenes tomadas desde una cámara situada en la parte frontal de un coche que viaja por una carretera. Diga qué aproximación usaría y porqué. Identifique las principales dificultades y diga como las resolvería. Los argumentos deben ser sólidos y con fundamento en las técnicas estudiadas.

9. ¿Qué han aportado los modelos CNN respecto de los modelos de reconocimiento de objetos empleados hasta 2012? Enumerar las propiedades comunes entre ellos y aquellas claramente distintas que hayan permitido una mejora en la solución del problema por parte de las CNN. Dar una opinión razonada de por qué significan realmente una mejora.

Las CNN han conseguido una mejora muy grande con respecto a los modelos de reconocimiento de objetos empleados hasta 2012. En todos los modelos de reconocimiento de características existe una parte de extracción de características, que es diferente en cada modelo (descriptores SIFT, bolsa de palabras, etc...) y una parte de entrenamiento del modelo con las características que se han obtenido de las imágenes que utilizamos para entrenar el modelo.

La principal diferencia entre las redes neuronales convolucionales es la forma que tiene de extraer las características, éstas, a diferencia de los modelos convencionales, extraen las características de la imagen a partir de la aplicación de filtros sobre la imagen hasta conseguir características con las que la red neuronal pueda aprender; en cambio, los modelos utilizados hasta 2012 no utilizan una características lo suficientemente buenas; esto provoca que el modelo no pueda representar toda la variabilidad que los objetos presentan.

10. Razone y argumente a favor y en contra de usar modelos de redes CNN ya entrenados, y que se conocen han sido efectivos en otras tareas distintas de la que tiene que resolver, como modelos para aplicar directamente o como modelos a refinar para la tarea que tiene entre manos. Dar argumentos que no sean genéricos o triviales y que fundamenten su postura.

A favor:

-

En contra:

-