

Reconocimiento Visual con Deep Learning

Tarea: Evaluación de características aprendidas por redes convolucionales en el contexto de búsqueda por similitud

Prof: José M. Saavedra Rondo
Prof. Auxiliar: Cristóbal Loyola

Octubre 2021

1. Objetivo

Entender y evaluar la capacidad de las redes convolucionales como extractores de características visuales en el contexto de búsqueda por similitud.

2. Descripción

En esta tarea, los y las estudiantes deberán evaluar las características aprendidas por redes convolucionales entrenadas en un dataset de dibujos de animales, como los que se presentan en la Figura 1. En este caso, tendrán que entrenar dos modelos convolucionales, AlexNet y ResNet-34, que luego serán ocupadas como extractores de características en el contexto de búsqueda por similitud.

2.1. Detalle técnico

En esta tarea deberán desarrollar las siguientes actividades:

2.1.1. Entrenamiento del los modelos

- Entrenar una modelo basado en el arquitectura AlexNet.
- Entrenar una modelo basado en el arquitectura ResNet-34.
- Los datos de entrenamiento y validación (test) pueden se descargados de : <https://www.dropbox.com/sh/hsqjv0kd13xda3g/AABYkVk0ruG85s4aL4C1nDKaa>

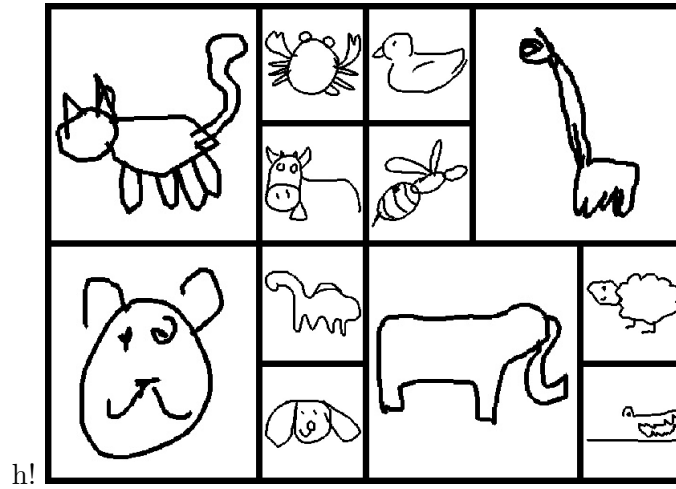


Figura 1: Ejemplo de sketches de animales. (Sería interesante que traten de adivinar qué animal representa cada uno de los dibujos)

2.1.2. Búsqueda por similitud

- Extraer el vector de características de AlexNet ocupando como salida la capa *fc7*.
- Extraer el vector de características de ResNet-34 ocupando como salida la capa *avg_pool*.
- Para las búsquedas, se debe utilizar como catálogo todas las imágenes de test, y como consulta utilizarán la estrategia *leave-one-out*, es decir todas las imágenes del catálogo son consultas, y se descarta esa imagen del catálogo.
- La relevancia se mide de acuerdo a la clase, es decir dos imágenes son similares si pertenecen a la misma clase.
- Ustedes deberán elegir entre usar una función de distancia (p.ej. Euclidiana), o una de similitud (p.ej. coseno).

2.1.3. Métricas a evaluar

- Mean Average Precision (total y separado por clases)
- Precisión@1 (total y separado por clases)
- Gráfico Recall-Precision (total)

2.1.4. Repositorios disponibles

Para realizar la tarea cuentan con los siguientes repositorios:

- Para entrenamiento: <https://github.com/jmsaavedrar/convnet2>
- Para búsquedas: https://github.com/jmsaavedrar/visual_attributes

2.1.5. Grupos

Esta tarea se puede realizar en grupos de 2 estudiantes.

3. Informe

1. **Abstract o Resumen:** es el resumen del trabajo.
2. **Introducción:** Aquí se describe el problema y el contexto (10 %)
3. **Desarrollo:** Aquí se describe el diseño e implementación del programa. Describa los pasos seguidos para resolver el problema. (40 %)
4. **Resultados Experimentales y Discusión:** Aquí se presentan los resultados, pero lo más importante es analizarlos. Observe y describa el comportamiento de los modelos en base a las métricas mencionadas anteriormente. ¿Puede genera algunas recomendaciones en base a sus observaciones?. Además presente ejemplos de resultados, bueno y malos casos. (40 %).
5. **Conclusiones** (10 %)

4. Entrega

La entrega del informe se realiza por u-cursos hasta el domingo 07 de noviembre, 2021, 23:50 hrs.