

CC7221: Reconocimiento Visual con Deep Learning

Tarea 1: Evaluación de redes convolucionales para clasificación de imágenes

Prof: José M. Saavedra Rondo
Ayudantes: Pablo Torres, Cristóbal Loyola

Abril 2021

1. Objetivo

Familiarizarse con la implementación, entrenamiento y evaluación de modelos convolucionales para la clasificación de imágenes.

2. Descripción

En esta tarea, los estudiantes deberán entrenar tres diferentes arquitecturas convolucionales para clasificación de prendas de vestir, según el dataset que se describe en Sección 2.2. Las arquitecturas a evaluar son AlexNet, ResNet-50¹ y ResNext² (compatible con ResNet-50).

Para implementar las redes se sugiere seguir el código descrito en <https://github.com/jmsaavedrar/convnet2>, que trae la implementación de AlexNet y ResNet. La arquitectura ResNext no viene incluida, por lo que es parte de la tarea extender la implementación existente para soportar ResNext.

Todas las redes descritas deben ser entrenadas sin pre-entrenamiento de pesos, con imágenes de tamaño 224×224 . Para la optimización deberán utilizar Adam, con los parámetros por defecto. Los modelos son entrenados por aproximadamente 20 épocas. Se recomienda un tamaño de batch de 64.

2.1. Gráficos y Métricas

Para la parte de análisis de resultados, deberán calcular mediciones y generar gráficos que muestren el comportamiento de sus modelos. Así, se pide incluir lo siguiente:

- Gráfico **epoch vs. loss** y **epoch vs. accuracy**, para cada una de las redes descritas.
- Accuracy de cada modelo, total y separado por clases. Un gráfico de barras ayudaría!! Un ejemplo de gráfico esperado se muestra en la Figura 1
- Caracterización de casos difíciles y fáciles.

2.2. Datasets

Las redes descritas anteriormente deberán ser entrenadas en el contexto de clasificación de prendas de vestir. Para este fin, se dispone del dataset *Impresee-Clothing* que puede ser descargado de <https://www.dropbox.com/s/2jt9086zlm9fj5u/impresee-clothing.tar>. Este dataset contiene 121.930 imágenes distribuido en 19 clases. El dataset está dividido en 100.000 imágenes de entrenamiento y 21.930 imágenes de test. Un ejemplo de las imágenes contenidas en este dataset se muestran en Figura 2.

¹ResNet: <https://arxiv.org/pdf/1512.03385.pdf>

²ResNext: <https://arxiv.org/pdf/1611.05431.pdf>

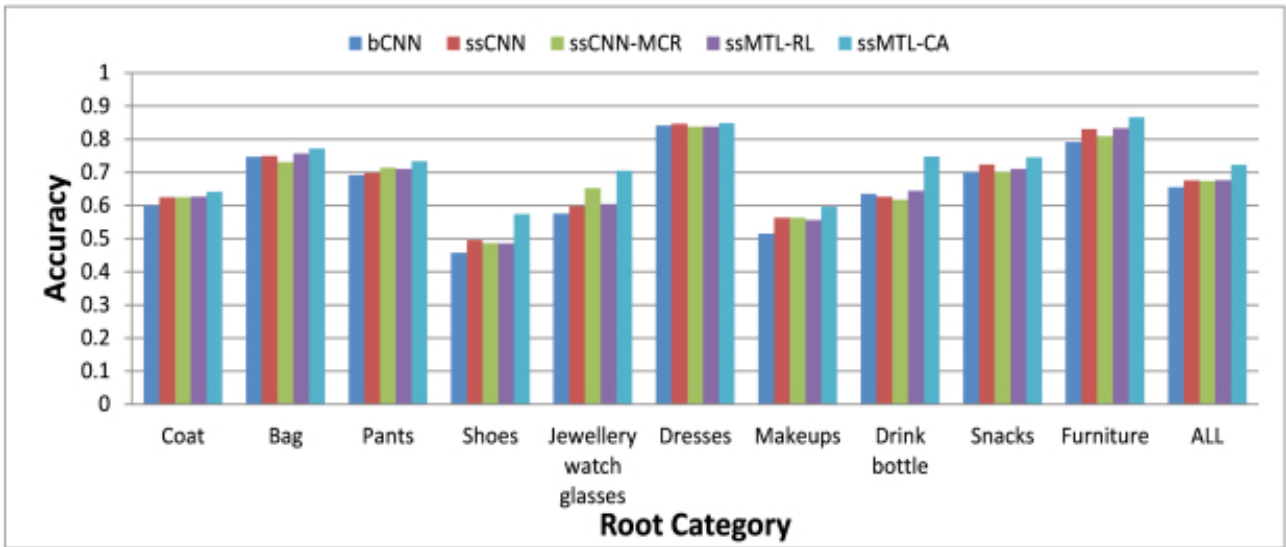


Figura 1: Ejemplos de accuracy por diferentes clases. Los métodos aplicados se diferencian por color. Ojo, el gráfico mostrado corresponde a otro problema, pero sirve para ilustrar lo que se espera en este trabajo.

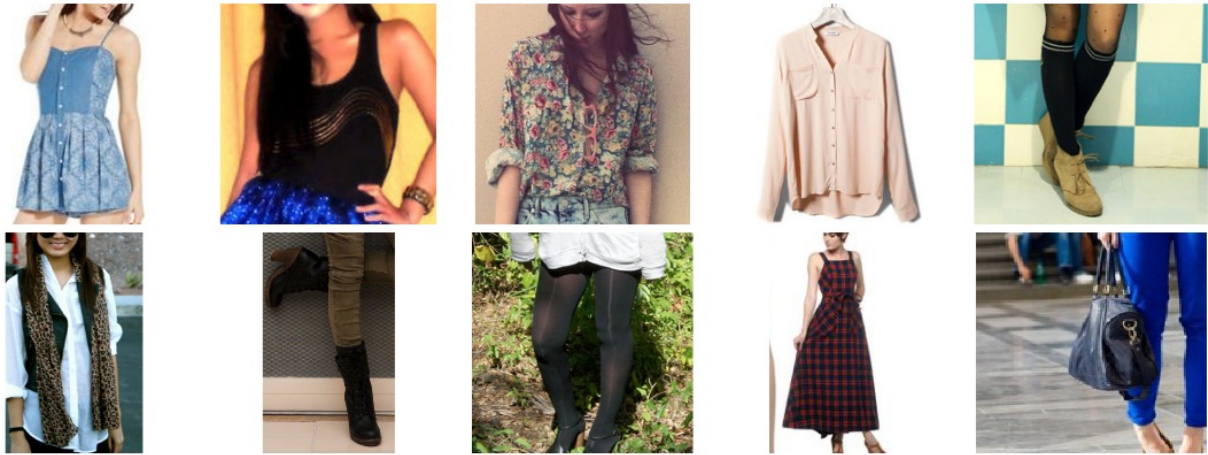


Figura 2: Ejemplos de imágenes del dataset clothing

En caso de presentar problemas de memoria durante el entrenamiento con el dataset completo, se permitirá realizar experimentos con una versión reducida del dataset *Impresee-Clothing*. El dataset reducido contiene 31.977 imágenes para entrenamiento y 1.900 imágenes para test. Este dataset puede ser descargado de <https://www.dropbox.com/s/c3h9e0o6ajydbao/clothing-small.zip>

Es importante que en el informe se indique claramente qué dataset ocuparon para los experimentos.

2.3. Implementación

Para esta tarea se recomienda fuertemente utilizar la implementación de <https://github.com/jmsaavedrar/convnet2>. Aquí encontrarán las arquitecturas AlexNet y ResNet.

2.4. Arquitecturas más pequeñas

En caso de tener problemas de recursos computacionales para entrenar los modelos ResNet-50 y ResNext, se aceptarán experimentos con la versión de ResNet-34 y su correspondiente versión ResNext.

3. Informe

1. **Abstract o Resumen:** es el resumen del trabajo.
2. **Introducción:** Aquí se describe el problema y el contexto. Comente las diferencias entre Alex-Net, ResNet y ResNext. (10 %)
3. **Desarrollo:** Aquí se describe el diseño e implementación de los programas necesarios para realizar sus experimentos. Describa los pasos seguidos para entrenar. Además, aquí es obligatorio que describa detalladamente cómo se implementó ResNext. (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? (40 %).
5. **Conclusiones** (10 %)

4. Entrega

La entrega se realiza por u-cursos hasta el domingo 25 de abril, 2020, 23:50 hrs. Se debe incluir:

1. Código fuente (en Python)
2. Informe