

EL7007-1 - Introducción al Procesamiento Digital de Imágenes

TAREA 3

Profesor de Cátedra: Claudio Pérez F.
Profesor Auxiliar: Jorge Zambrano I.
Semestre: Otoño 2022

EXTRACCIÓN DE CARACTERÍSTICAS Y CLASIFICACIÓN DE IMÁGENES

1. Objetivo:

El objetivo de esta tarea es explorar distintos métodos de extracción de características, desde los algoritmos tradicionales hasta los modelos convolucionales para la clasificación de imágenes. Para ello se trabajará con algunas imágenes de la base de datos **MaskedFace-Net** [1][2][3], que contiene imágenes de rostros con mascarilla bien puestas y mal puestas. Este conjunto de imágenes está basado en la base de datos **Flirck-Faces-HQ** [4], creada originalmente como puntos de referencia para redes generativas adversarias GAN. Los algoritmos de extracción de características que se emplearán son LBP, HOG y dos redes convolucionales a su elección (e.g., ResNet, DensNet).

2. Descripción:

La extracción de características es un paso fundamental dentro de los algoritmos de visión computacional. La idea es recuperar las características más relevantes de una imagen para heredarlos a una próxima etapa, como por ejemplo la clasificación.

Los extractores de características han ido evolucionando con el paso del tiempo, sin embargo, muchos siguen siendo utilizados por su rapidez, simplicidad y efectividad. Con la llegada del “Deep Learning”, se ha presentado una nueva forma de extraer características, aprovechando el poder computacional de la actualidad y la robustez de las bases de datos de entrenamiento.

Esta tarea tiene como finalidad implementar un clasificador de imágenes de rostros en tres categorías: Mascarilla bien colocada, Mascarilla mal colocada y Rostros sin mascarilla. Para esto, se extraerán los vectores de características (embeddings) de las imágenes con extractores como LBP, HOG y la capa final de dos modelos convolucionales. Posteriormente se deberá entrenar un clasificador. Este podría ser un Support Vector Machine (SVM), un Random Forest (RF) o un Multi Layer Perceptron (MLP). El clasificador también queda a su elección.

Recomendamos realizar esta tarea en Python con librerías como OpenCV, skimage, sklearn, PIL, Tensorflow, Keras, Pytorch entre otras.

3. Implementación:

3.1. Bases de Datos:

Se deberá utilizar una subdivisión de la base de datos anteriormente mencionada. En el enlace de Google Drive encontrará las imágenes separadas en tres carpetas. En la Figura 1 se pueden observar ejemplos de cada clase. Utilice un 70% de la base para *entrenamiento* y un 30% para *prueba*.

https://drive.google.com/drive/folders/1w_Q8gYPooSoLNDDhNdRk_y-BvbOjSA9L?usp=sharing



(a) Mascarilla bien colocada



(b) Mascarilla mal colocada



(c) Sin mascarilla

Figura 1: Ejemplos de las clases del dataset entregado. Las imágenes (a) y (b) han sido tomadas de [3]. La imagen (c) ha sido tomada de [4].

3.2. Extractores de características:

Utilice funciones de librerías especializadas para la extracción de características por LBP y HOG. Para el caso de la extracción por CNN, ocupe la última capa de dos modelos de familias diferentes. (Por ejemplo, un modelo de la familia de DenseNet y otro de ResNet).

3.3. Clasificador:

Entrene un clasificador con los embeddings obtenidos de conjunto de *entrenamiento* y pruébelos con los embeddings del conjunto de *prueba*. Detalle en el informe las características de su clasificador. Muestre en el reporte el **Accuracy** obtenido para los 4 casos, junto con su respectiva **Matriz de Confusión** (recuerde especificar bien las etiquetas en la matriz de confusión).

4. Entregables:

Entregue un reporte individual del trabajo junto con su código. Puede presentarse en un .pdf por separado o un archivo ipynb. En el informe también se deben detallar todos los parámetros empleados, ya sean en el método de extracción de características o del clasificador. Muestre ejemplos de clasificaciones correctas e incorrectas, mostrando la etiqueta original y la predicción.

Además, utilice 3 fotografías suyas (con la mascarilla bien y mal colocada, y sin mascarilla) y haga la predicción con el mejor modelo obtenido. Analice esos resultados.

5. Referencias:

- [1] A. Cabani, K. Hammoudi, H. Benhabiles, and M. Melkemi, “MaskedFace-Net – A dataset of correctly/incorrectly masked face images in the context of COVID-19,” *Smart Heal.*, vol. 19, p. 100144, Mar. 2021, doi: 10.1016/J.SMHL.2020.100144.
- [2] H. B. M. M. Karim Hammoudi, Adnane Cabani, “Validating the correct wearing of protection mask by taking a selfie: Design of a mobile application “checkyourmask” to limit the spread of covid-19,” *Computer Modeling in Engineering & Sciences*, vol. 124, no. 3, pp. 1049–1059, 2020.
- [3] cabani/MaskedFace-Net: MaskedFace-Net es un conjunto de datos de rostros humanos con una máscara usada correcta e incorrectamente basada en el conjunto de datos Flickr-Faces-HQ (FFHQ). (github.com).
- [4] NVlabs/ffhq-dataset: Flickr-Faces-HQ Dataset (FFHQ) (github.com).

La fecha de entrega de la tarea será el día 26/05/2022 a las 18:00 hrs por medio de U-cursos.