

Abstract

Se construyo dos redes neuronales convolucionales binarias, una para caras de personas y otra para sombreros, con el fin de localizar caras y sombreros de forma individual en una imagen en particular. Se proceso los resultados de las dos redes neuronales y con cada respuesta se construyo un mapa lógico en el cual se puede localizar la zona donde esta presente el objeto de interés, caras o sombreros. Posteriormente se procesan los dos mapas y del procesamiento se puede etiquetar la imagen indicando si las caras presentes en la imagen están usando o no sombreros.

1 Introducción

Este documento se organizo de la siguiente forma: En la sección 2 se analizan trabajos previos de reconocimiento del contenido de una imagen. En la sección 3 se describen los materiales y metodologías usados para abordar el proyecto. En la sección 4 se profundiza en la metodología usada y en la sección 5 se presentan los resultados obtenidos de esta metodología. En la sección 6 se presentan las conclusiones de los resultados obtenidos.

2 Trabajo Previo

Google Cloud Vision API

Image Recognition Tensorflow

Integrating Humans and Computers for Image and Video Understanding

How Will Google "Read" & Rank Your Images in the Near Future

3 Materiales y Métodos

3.1 Materiales

Las imágenes de las caras y otras imágenes diferentes a caras fueron obtenidas de "Labeled Faces in the Wild Home y Caltech 101". Las imágenes de los sombreros y otras imágenes diferentes a sombreros fueron obtenidas

de "ImageNet" Se uso la biblioteca TensorFlow para el lenguaje de programación Python 3.5

3.2 Métodos

Se uso una metodología basada en redes neuronales convolucionales tomando como base la implementación de Cifar-10, el cual se encarga de clasificar y etiquetar imágenes RGB de 32x32 de 10 clases diferentes. La arquitectura de Cifar-10 consiste en capas de convolución y operaciones no lineales cuyo resultado es procesado por una red neuronal, el procedimiento se describe en:

4 Experimentos

5 Resultados

6 Conclusiones