

Proyecto de Machine Learning

Detección de Dígitos en las Manos

Carlos Cortés

June 7, 2018

1 Objetivo

Clasificar imágenes de manos haciendo las señas de los dígitos del lenguaje de señas y reconocer el dígito que están haciendo.

2 Dataset

La base de datos de imágenes fue tomada del github de la referencia.

Consta de 2062 imágenes etiquetadas del 0-9 [1].

Cada imagen es de 100x100 y en formato Jpg (cada entrada es una triada con los valores RGB) por lo que fue necesario convertirlas a escala de grises.



3 Método

Lo primordial es recordar que las imágenes son matrices con cada entrada siendo el valor del pixel.

Por lo que se puede desdoblar la matriz en un gran vector para aplicarle algoritmos de machine learning con cada imagen como vector instancia.

Se probó con dos métodos:

K-NN: Es muy sencillo de realizar y da buenos resultados.

Regresión Logística: Supuestamente da mejores resultados.

4 Implementación

La Implementación de los algoritmos fue realizada en C++ para obtener mayor velocidad en la clasificación y utilizando bibliotecas con funciones de aprendizaje automático.

4.1 Dependencias

- *Mlpack*.
Biblioteca principal para métodos de Machine Learning.
- *Armadillo*.
Biblioteca para matrices que Mlpack utiliza.
- *OpenCV*.
Biblioteca que utilizo para manipular imágenes y convertirlas en vectores.

4.2 Breve Descripción del Programa

El programa tiene 4 funciones principales:

- **Actualizar la Base de datos.** Lee todas las imágenes de la carpeta *Dataset*, las convierte en una matriz y escribe un archivo *.csv* con los datos y otro con las clases.
- **Modo Interactivo.** Inicia un *prompt* en el que puedes indicarle la ruta de la imagen a clasificar.
- **Test.** Parte el Dataset en conjunto de entrenamiento y de prueba y luego estima el poder predictivo.
- **Cámara.** Toma como input a clasificar la imagen de la cámara.

5 Mejor Puntaje

El mejor puntaje obtenido por el clasificador

Partiendo el Dataset en 1856 de entrenamiento y 206 de prueba.

Con $k = 10$ vecinos más cercanos.

5.1 K-NN

154 casos correctos.

Porcentaje = 74.75%.

Entre (68.85% , 80.65%) al 95% de confianza.

5.2 Regresión Logística

25 casos correctos.

Porcentaje = 12.13%.

Entre (7.69% , 16.57%) al 95% de confianza.

References

- [1] ardamavi. Sign-Language-Digits-Dataset. 2018. Recuperado de <https://github.com/ardamavi/Sign-Language-Digits-Dataset>