

Ingeniería del Conocimiento: Utilizando SIFT, k-means, BOW y SVM para reconocimiento de lugares

Andrés Gómez

I. MÉTODO

Para la solución del problema se partió de un conjunto de imágenes de cada una de las 10 zonas, reunidas por varias personas para llegar a un total de 1430 imágenes.

A cada una de estas se le sacaron los descriptores SIFT, luego se creó con estos descriptores un modelo utilizando k-means de 350 clusters, que se determinó entre un conjunto de clusters de 200 a 500 y dio la mayor presición en el validation set.

Para entrenar el clasificador k-means se tomó solo una muestra de 100000 vectores SIFT aleatorios del training set para ahorrar tiempo de procesamiento.

Luego se formó un histograma utilizando la técnica BOW (Bag of words) contando la frecuencia de cluster identificados en cada una de las imágenes, y estos histogramas se utilizaron para entrenar un modelo SVM (Support Vector Machines). Para el SVM se utilizó el parámetro $C = 10000$ para mejorar la presición en el training set.

Para el entrenamiento se utilizó "mean normalization" para obtener una media de 0 y una desviación estándar de 1. Se tomó la media y la desviación estándar de los histogramas del training set y luego se aplican a los histogramas de nuevas imágenes que se alimentan al clasificador.

El data set fue dividido en training, val y test con porcentajes de 60%, 20% y 20% respectivamente. Las precisiones después de entrenar el clasificador fueron 100% en el training set, 95% en el validation set y 91% en el test set.

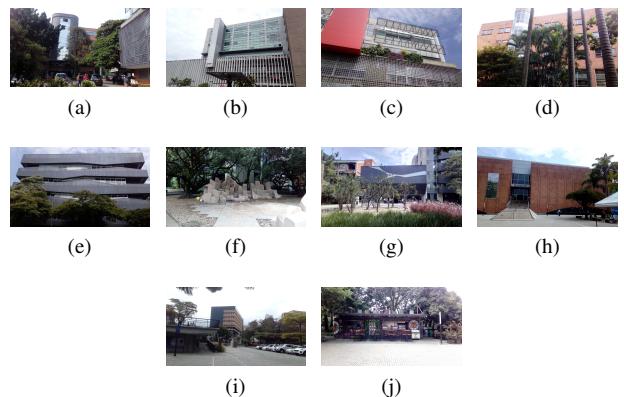


Fig. 2: Muestras de cada uno de los lugares.

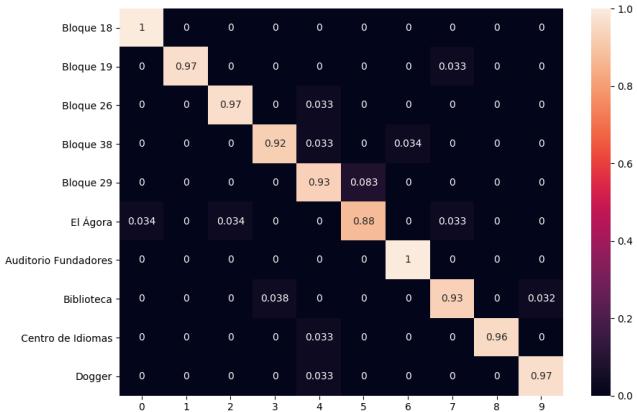


Fig. 1: Matriz de confusión.