Tamaño del conjunto de entrenamiento

Se llevaron a cabo una serie de pruebas con el objetivo de estudiar el número de días necesarios para lograr un buen rendimiento de los clasificadores. Se armaron conjuntos de entrenamiento variando el número de instancias de entrenamiento desde 5 hasta 2400. Para cada uno de estos conjuntos de entrenamiento se entrenó un clasificador y el mismo fue evaluado.

Las **figuras A y B** representan el valor obtenido de correlación y MAE de un clasificador no lineal (SVM) entrenado con diferente número de instancias. Para todas las playas analizadas se observa que el incremento en el tamaño del conjunto de entrenamiento afecta directamente el rendimiento del clasificador, logrando mejoras en términos de correlación y error absoluto medio promedio. Al mismo tiempo se observa que esta mejora tiene un límite, específico de cada playa, a partir del cual un incremento en el tamaño del conjunto de entrenamiento no produce mejorías en el rendimiento del clasificador.

Dentro del rango de las 100 primeras instancias se puede apreciar un crecimiento en el rendimiento de manera acelerada, pero con pequeñas oscilaciones que luego van atenuándose para estabilizarse cerca del valor limite de rendimiento.

El número de instancias necesarias para llegar al límite del rendimiento es particular de cada playa. En la figuras **C**  **y D** se puede observar con más detalle este hecho. Las playas de Sunset Makaha y Makapuu logran estabilizar su rendimiento a partir de las 200 instancias. Mientras que las playas de DH y Ala Moana, logran estabilizar su rendimiento a partir de las 600 instancias.

Del análisis de las pruebas realizadas en esta sección podemos decir que para las playas de Sunset, Makaha y Makapuu, un conjunto de entrenamiento compuesto de 6 a 12 meses de observaciones visuales, nos daría un rendimiento cercano al óptimo. Mientras que para las playas de Diamond Head y Ala Moana es necesario de un conjunto de observaciones entre 12 y 24 meses para obtener un rendimiento cercano al óptimo.

Figura A. Correlación de un clasificador no lineal para diferentes tamaños de conjuntos de entrenamiento en las diferentes playas analizadas.

Figura B. MAE de un clasificador no lineal para diferentes tamaños de conjuntos de entrenamiento en las diferentes playas analizadas

Figura C. Correlación de un clasificador no lineal para diferentes tamaños de conjuntos de entrenamiento en las diferentes playas analizadas.

Figura D. MAE de un clasificador no lineal para diferentes tamaños de conjuntos de entrenamiento en las diferentes playas analizadas