

**Reconocimiento de Patrones, Máster en
Visión Artificial
Universidad Rey Juan Carlos
Curso 2016-2017
Práctica 2: Reducción de dimensionalidad**

1. Normas

La practica es individual y deberá programarse **obligatoriamente en python**.

La memoria de la práctica junto con los **scripts python** se entregarán en el Moodle de la asignatura **La fecha tope de entrega de la práctica es el día del examen de la asignatura**.

En la portada de la memoria deben aparecer el nombre y DNI del autor. No existe restricción alguna en cuanto al formato y extensión de la memoria, aunque se recuerda que la calidad de la memoria es un aspecto de importancia similar a los programas.

2. Planteamiento

Esta práctica tiene como objetivos:

1. Introducir al alumno en el problema de la clasificación de señales de tráfico en imágenes.
2. Aprender a utilizar la técnica de reducción de dimensionalidad supervisada: el Análisis Discriminante Lineal (LDA).
3. Introducir al alumno en el problema de detección de objetos.

2.1. Señales de tráfico

Las señales de tráfico se han diseñado para que sean fácilmente distinguibles del entorno en cualquier condición de luz, color de fondo y climatología.



Figura 1: Ejemplos de señales de tráfico normalizadas y recortadas

Por tanto tienen unas formas geométricas regulares, colores muy vivos y son reflectantes (ver Fig. 1).

Observando la Fig. 1 podemos hacernos una idea sobre el tipo de información que nos permitiría distinguir un tipo de otro e incluso localizarlas en una imagen de la carretera o la calle. En particular, las que hemos seleccionado, se distinguen por la forma geométrica (triángulo, octógono o círculo), por las zonas en las que se divide (borde oscuro y fondo claro o fondo oscuro) o por el color (rojo, blanco, negro y/o azul oscuro). La forma geométrica como característica nos podría llevar a usar la extracción de bordes o los histogramas de orientación de bordes. El color nos podría llevar a utilizar la densidad de probabilidad en un espacio de color como característica discriminante.

En esta práctica utilizaremos, en la parte obligatoria la imagen de la señal en color, con sus tres canales RGB, como característica discriminante (algo relacionado, a la vez, con la forma y las zonas con colores diferentes en que se divide la señal).

2.2. Clasificación basada en color (RGB)

En esta sección construiremos un clasificador de señales de tráfico en 4 clases: Stop, peligro, prohibición y fondo (o no-señal de tráfico). Las imágenes de entrada serán las imágenes de diferentes señales de tráfico que habrá que leer de disco, reducir a un tamaño pequeño, por ejemplo 25×25 píxeles, y vectorizar la imagen con sus tres canales RGB (p.ej. con `numpy.ravel`). Dada la alta dimensionalidad del problema, en el ejemplo de tamaño $3 \times 625 = 1875$ dimensiones, habrá que utilizar una técnica de reducción de dimensionalidad de las estudiadas en clase.

Una posibilidad es emplear Análisis Discriminante Lineal (LDA) para obtener una matriz de proyección, $A_{3 \times 1875}$, que nos lleve de 1875 dimensiones a 3 (dado que tenemos cuatro clases). El resultado esperado de la proyección sobre el espacio tridimensional con $A_{3 \times 1875}$ de las imágenes de diferentes tipos

de señales de tráfico y el fondo. Es este nuevo espacio, y dado que utilizamos LDA, las imágenes de los diferentes tipo de señales deberían quedar separadas de las del fondo y entre sí (si se cumplieren las hipótesis del LDA).

Una vez en un espacio de características transformado, con una dimensión reducida y con buena separabilidad entre las clases, habrá que elegir alguno de los clasificadores estudiados en clase (Euclídeo, bayesiano de mínimo error paramétrico o no paramétrico, K vecinos más próximos, etc.). Para la elección del mejor clasificador se deberá utilizar *validación cruzada* debido a que tenemos un número reducido de ejemplos de entrenamiento. Como vimos en clase el LDA multiclase tiene problemas si la matriz de dispersión intraclass (\mathbf{S}_w) está mal condicionada (no es de rango completo) o está numéricamente cercana a serlo. En clase se vieron algunas soluciones a ese problema (utilización de \mathbf{S}_t y/o PCA+LDA, por ejemplo).

3. Desarrollo

En esta práctica se pide:

3.1. Parte obligatoria (hasta 8 puntos)

1. **Opcional:** Utilizar el script que se proporciona para detección de regiones *MSE*R para obtener ejemplos negativos de imágenes sin señales de tráfico (al menos de las que nos interesan en la práctica). Esto puede ser importante porque vamos a usar el mismo detector de regiones para obtener regiones de alto contraste con el fondo (MSERs), dado que las señales de tráfico lo son.
2. Usar el Análisis Discriminante Lineal (LDA) que proporciona *Scikit-Learn* (`sklearn.discriminant_analysis.LinearDiscriminantAnalysis`) calculando la matriz de proyección LDA, $\mathbf{A}_{3 \times 1875}$. Para el entrenamiento se utilizarán las imágenes en *ImgAlumnos*. Las imágenes se consideraran como un vector columna (de 1875 dimensiones) para el LDA.
3. Probar diferentes clasificadores con las 4 clases en el espacio de características de 3 dimensiones encontrado mediante LDA en el paso anterior. Los clasificadores se evaluarán mediante validación cruzada con 10 grupos (10 *fold*). Elegir el clasificador con mejor resultado (pensar también en la velocidad de ejecución) para construir el detector. **Las**



Figura 2: Ejemplos detección de señales de tráfico utilizando LDA

medidas de rendimiento que se consideren oportunas (matrices de confusión, error, medida F1, etc) deben de aparecer en la memoria.

4. Si la matriz S_w en el análisis discriminante está cercana a no tener inversa (ver `np.linalg.matrix_rank` en numpy) entonces el LDA nos dará una mala solución (las clases estarán muy cercanas en el espacio de dimensión reducida mediante LDA). Se pide también explorar la solución aportada en clase basada en PCA+LDA. Es importante destacar que el conjunto de autovectores de PCA con la que nos quedaremos deberíamos elegirlo mediante alguna de las técnicas de selección de características vista en clase (usando validación cruzada con 10 grupos y teniendo como objetivo la mínima tasa de falsos positivos).
5. Una vez se tiene el mejor clasificador elegido y entrenado, se pide usar el script proporcionado para la detección de objetos mediante *MSER*, y pasarlo a las imágenes de test proporcionadas. Es importante darse cuenta de que la señal de tráfico a veces tiene dos regiones MSER y a veces solo una y puede ser importante modificar ese cuadrado obtenido antes de pasárselo al clasificador. Los rectángulos que son MSER se pasarán a nuestro clasificador para determinar si es o no una señal de tráfico y se pintará sobre la imagen el rectángulo de las que nuestro clasificador de respuesta positiva. La evaluación de la detección de señales de tráfico en este caso será cualitativa (ver Fig.2). **Las imágenes de test con las detecciones se entregarán junto con la memoria.**

3.2. Parte opcional sobre gradientes (hasta 2 puntos adicionales)

En esta parte se pide utilizar características HoG (ver *skimage.feature.hog*) para clasificar. En concreto:

1. Construir un clasificador basado en HoGs con 4 clases en el espacio de características de 3 dimensiones encontrado mediante LDA en el paso anterior. El clasificador se evaluará mediante validación cruzada con 10 grupos (10 *fold*). **Las medidas de rendimiento que se consideren oportunas (matrices de confusión, error, medida F1, etc) deben de aparecer en la memoria.**
2. Una vez se tiene el mejor clasificador elegido y entrenado, se pide usar el script proporcionado para la detección de objetos con MSER, y pasarlo a las imágenes de test proporcionadas. La evaluación de la detección de señales de tráfico en este caso será cualitativa. **Las imágenes de test con las detecciones con estos nuevos descriptores se entregarán junto con la memoria.**

3.3. Parte opcional sobre gradientes (hasta 2 puntos adicionales)

En esta parte se pide utilizar características HoGs (ver *skimage.feature.hog*) con un clasificador SVM lineal.

1. Construir un clasificador SVM lineal por cada señal de tráfico (stop/fondo, peligro/fondo, prohibición/fondo) en el espacio de características HoG. Los clasificadores se evaluarán mediante validación cruzada con 10 grupos (10 *fold*). **Las medidas de rendimiento que se consideren oportunas (matrices de confusión, error, medida F1, etc) deben de aparecer en la memoria.**
2. Una vez se tiene el mejor clasificador elegido y entrenado, se pide usar el script proporcionado para la detección de objetos con MSER, y pasarlo a las imágenes de test proporcionadas. La evaluación de la detección de señales de tráfico en este caso será cualitativa. **Las imágenes de test con las detecciones con estos nuevos descriptores se entregarán junto con la memoria.**

4. Preguntas a responder

Estas preguntas son algunas de las que sería interesante responder. No quiere decir que haya que responder a todas. Pero cuanto mayor sea el número de preguntas respondidas podrá ser mejor la puntuación.

- ¿Cuál es el mejor clasificador para cada tipo de descriptor?
- ¿Cuál es el mejor descriptor de todos para la clasificación de señales de tráfico?
- ¿Qué pasa si hago una combinación de descriptores (RGB + HoGs)? ¿Mejora la clasificación?
- ¿Qué pasa si en la validación cruzada utilizo como objetivo el maximizar la medida F1 (que combina precisión y recall)?

5. Valoración de la práctica

Esta práctica puede puntuarse sobre 12 puntos y hará media con la anterior

La puntuación obtenida vendrá dada por los siguientes criterios:

- Justificación de la elección de la técnica de clasificación.
- Funcionamiento de la clasificación de señales de tráfico.
- Evaluación de la clasificación (matrices de confusión, tasas de acierto y varianza, etc).
- Limpieza del código.
- Calidad de la memoria y de las explicaciones sobre las decisiones tomadas.
- Extensión de las pruebas realizadas (diferentes pruebas realizadas, utilización de imágenes no provistas junto con el enunciado, etc.).

6. Datos proporcionados

Se proporcionan las imágenes de entrenamiento y clasificación en el moodle de la asignatura. Las imágenes vienen de dos retos de clasificación y detección de señales de tráfico (“German Traffic Sign Recognition and Detection Benchmarks”).