



**Extracción de Características en Imágenes 2024-2025**

MASTER CIENCIA DE DATOS

UNIVERSIDAD DE GRANADA

# Práctica 1

MIGUEL GARCÍA LÓPEZ

## Índice

1. Introducción	2
2. Dataset	2
3. Clasificación con HOG	3
3.1. HOG . . . . .	3
3.2. Búsqueda de hiperparámetros . . . . .	4
4. Bibliografía	5

## Índice de figuras

1. Grafo de decisión de tareas. . . . .	2
2. Flujo de cómputo del descriptor HOG. . . . .	3

## Índice de cuadros

1. Parámetros del modelo SVM. . . . .	4
2. Parámetros del descriptor HOG. . . . .	5

## 1. Introducción

En la presente práctica de la asignatura de **Extracción de Características en Imágenes**, se llevarán a cabo una serie de tareas definidas por un grafo de decisión (fig 1). Dado este grafo es necesario seguir el camino hasta al final y allí donde haya una bifurcación, escoger entre una tarea básica o una tarea bonificadora. Las tareas bonificadoras son iguales que las básicas, pero con un toque de dificultad y desarrollo por parte del alumno. De forma resumida, las tareas a realizar son las siguientes:

- **Búsqueda de un conjunto de datos:** El estudiante puede usar el *dataset* **MNIST** por defecto, pero en este caso se ha optado por la tarea complementaria de escoger uno.
- **Clasificación con HOG:** Se entrenará un modelo **SVM** usando el descriptor **HOG** y se realizará un análisis de los resultados del mismo. Además se aplicarán técnicas como validación cruzada y selección de hiperparámetros.
- **Clasificación con LBP:** Se realizará lo mismo que con **HOG** descrito en el apartado anterior. Además, como parte de la tarea complementaria bonificada, se usará una implementación propia de **LBP** para extraer las características del *dataset*.

De las 24000 imágenes se han escogido de forma aleatoria, y teniendo en cuenta equilibrio entre clases, 8000 imágenes para la clasificación con **SVM+LBP** y **SVM+HOG**.

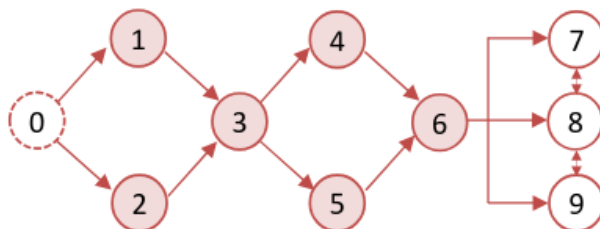


Figura 1: Grafo de decisión de tareas.

## 2. Dataset

Para el *dataset* se ha escogido el conjunto de “Gatos vs Perros” de *Kaggle*. Este se puede encontrar en el siguiente enlace: <https://www.kaggle.com/competitions/dogs-vs-cats>.

El conjunto contiene cerca de 24000 imágenes, la mitad de perros y la mitad de gatos. Este conjunto se compone de imágenes de multitud de resoluciones, por lo que se ha procedido a realizar varias pruebas en el código y se ha llegado a la conclusión de que  $28 \times 28$  píxeles es un tamaño con el que poder trabajar por los siguiente motivos:

- La extracción de descriptores en imágenes de alta resolución lleva a altos tiempos de cómputo.
- La implementación de **LBP** es rápida, pero no tanto como las implementaciones de otros descriptores como **HOG** en *OpenCV*.
- Se han realizado varias pruebas y con tamaños de resolución mucho mayores no se consiguen unos resultados mucho mejores (hasta donde se ha podido comprobar).

Además de lo descrito, se han transformado las imágenes a escala de grises para trabajar con un solo canal.

### 3. Clasificación con HOG

#### 3.1. HOG

El descriptor **HOG** es una técnica ampliamente utilizada en la visión por computador, cuyo objetivo es capturar la estructura local de las imágenes basándose en los gradientes de intensidad. Divide la imagen (fig 2) en celdas pequeñas y calcula un histograma de orientaciones de gradiente dentro de cada celda. Para mejorar la robustez frente a cambios de iluminación, se normalizan los histogramas en bloques de celdas adyacentes.

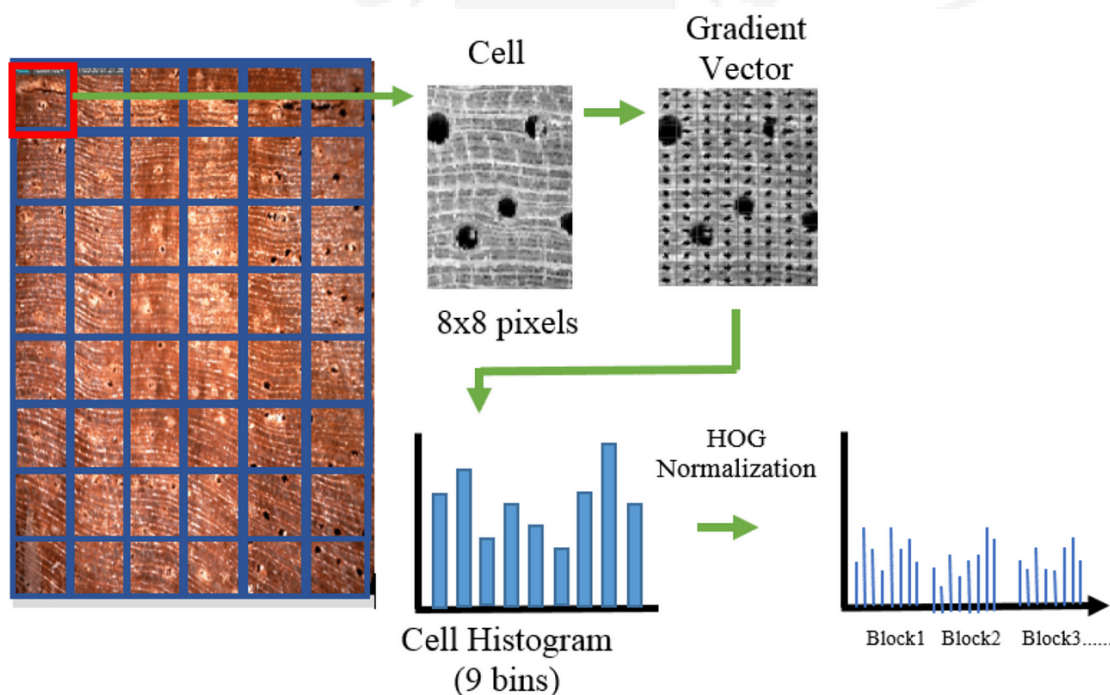


Figura 2: Flujo de cómputo del descriptor HOG.

### 3.2. Búsqueda de hiperparámetros

Se ha implementado una búsqueda de hiperparámetros de forma que es posible buscar solo hiperparámetros del algoritmos **SVM** o, si se selecciona, búsqueda para **SVM** y **HOG**. Hay que tener en cuenta que la búsqueda de hiperparámetros es un proceso muy costoso computacionalmente, por lo que esta búsqueda en lo relativo a **HOG** se ha realizado con 200 imágenes usando el algoritmo **RandomSearch** de *Scikit-Learn*. Las principales ventajas de este algoritmo son:

- Un espacio de búsqueda más amplio al realizar combinaciones aleatorias.
- Más eficiente en espacios de dimensionalidad alta.

Los parámetros optimizados no han sido todos. Inicialmente se realizó un estudio de los parámetros tanto experimental como teórico. Dado ese primer paso se decidió que **winSize** sería del tamaño de la imagen, **blockSize** la mitad y **blockStride** y **cellSize** un cuarto. Se realizan búsquedas sobre los siguientes parámetros:

- **nbins**: Número de histogramas por celda. Valores: {6, 9, 12}.
- **winSigma**: Sigma de la ventana de suavizado gaussiano. Valores: {0,5, 1,0, 2,0, 5,0}.
- **L2HysThreshold**: Umbral para la normalización L2. Valores: {0,1, 0,2, 0,3, 0,4}.
- **signedGradients**: Indicador de gradientes firmados (booleano).
- **gammaCorrection**: Aplicación de corrección gamma (booleano).

Parámetro	Valor
svm_kernel	rbf
svm_gamma	1
svm_C	1

Cuadro 1: Parámetros del modelo SVM.

Los valores escogidos finalmente para los parámetros tanto del descriptor como para el **SVM** se encuentran descritos en las tablas 2,1.

Parámetro	Valor
winSize	(28, 28)
blockSize	(14, 14)
blockStride	(7, 7)
cellSize	(7, 7)
nbins	12
winSigma	5
L2HysThreshold	0.3
signedGradients	True
gammaCorrection	0

Cuadro 2: Parámetros del descriptor HOG.

## 4. Bibliografía