

# DOG & CAT CLASIFICATOR



**2022-2023**

**Odei Hijarrubia Bezares  
Endika Aguirre Eguizabal**

# ÍNDICE:

<b>Introducción y descripción del problema:</b>	<b>2</b>
<b>Extracción de características 1:</b>	<b>4</b>
<b>Extracción de características 2:</b>	<b>5</b>
<b>Clasificación:</b>	<b>6</b>
<b>Resultados:</b>	<b>7</b>
<b>Conclusión:</b>	<b>8</b>
<b>Vídeo:</b>	<b>8</b>

## **Introducción y descripción del problema:**

El objetivo de este trabajo es clasificar imágenes de perros y gatos utilizando técnicas de extracción de características y modelos de clasificación basados en aprendizaje automático.

La fase de extracción de características consiste en transformar las imágenes en un conjunto de características numéricas que sean útiles para la tarea de clasificación. La fase de aprendizaje de los modelos consiste en entrenar un modelo de clasificación utilizando el conjunto de características extraídas y las etiquetas de clase correspondientes para cada imagen.

Compararemos dos técnicas de extracción de características, las cuales explicaremos en detalle más adelante, con un mismo algoritmo de clasificación y se seleccionará el modelo que tenga el mejor rendimiento en la tarea de clasificación de perros y gatos. Las imágenes se tratarán en escala de grises, pero se pueden probar imágenes en color opcionalmente.

## Extracción de características 1:

Para realizar la primera forma de extracción de características, deberemos dividir la imagen en bloques disjuntos de  $N \times N$  píxeles (por ejemplo,  $N = 16$ ). Llamaremos a estos bloques celdas.

Para cada píxel de la celda, compararemos cada uno de sus 8 vecinos, siguiendo la dirección de las agujas del reloj. Si el vecino es mayor que el píxel central, lo sustituiremos por un 1, y si es menor, lo sustituiremos por un 0. Esto generará un número binario que, al transformarlo a decimal, resultará en un número entre 0 y 255.

Modificamos el valor del píxel por el del número decimal obtenido. Una vez realizado esto para cada píxel de la celda, construiremos un histograma de los  $N \times N$  píxeles obtenidos. El valor del histograma será el vector de características asociado a la celda. Si concatenamos todos los histogramas de la imagen siguiendo el orden de las celdas, obtendremos el vector de características de la imagen.

Por ejemplo, si tenemos una imagen de  $128 \times 128$  píxeles y tomamos  $N=16$ , esto generará 64 celdas distintas y el vector de características finales tendrá 16.384 ( $64 \times 256$ ) elementos.

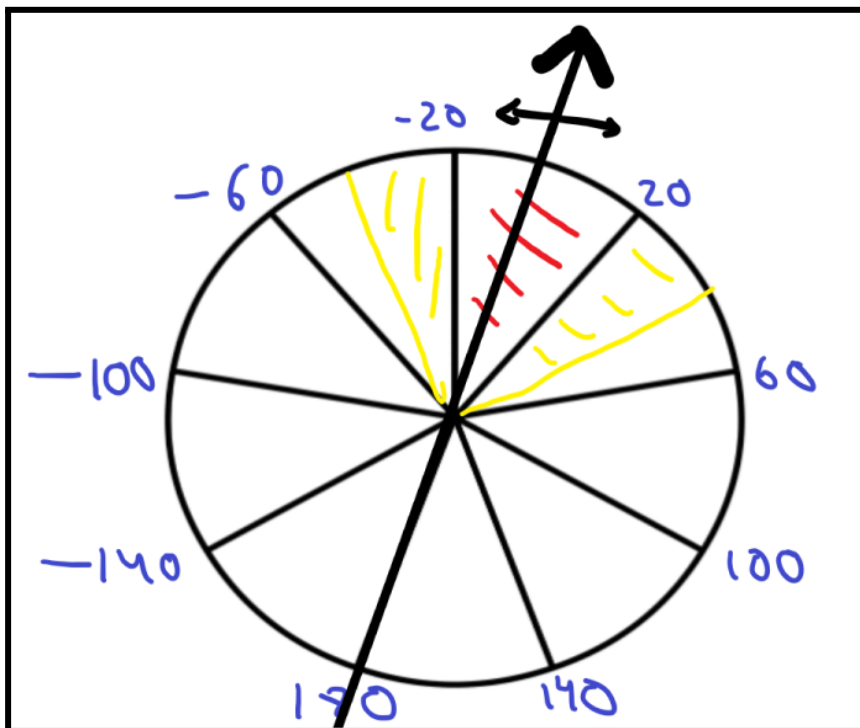
## Extracción de características 2:

Este método de extracción de características involucra tres pasos:

1. Aplicar filtros de convolución para calcular la magnitud y orientación del gradiente.
2. Dividir la imagen en bloques disjuntos y crear histogramas para cada bloque
3. Normalizar los histogramas.

Cada histograma tiene 9 bins o valores y están basados en la orientación del gradiente en lugar de en la intensidad de píxeles. La división es como se muestra en la figura de abajo.

Los histogramas en vez de contar cuántos píxeles tienen la misma orientación, representan la suma de las magnitudes de los píxeles de la misma orientación. Los histogramas se concatenan y se normalizan para obtener el vector de características final de la imagen.



Por ejemplo, si tenemos una imagen de 128x128 píxeles y la dividimos en bloques de 16x16 píxeles, tendremos 64 bloques y el vector de características final tendrá 576 (64x9) elementos.

## Clasificación:

En este trabajo, debemos implementar dos formas de extracción de características de imágenes, y utilizar estas características para entrenar un clasificador de imágenes para distinguir entre perros y gatos.

La primera forma de extracción de características implica dividir la imagen en bloques de  $N \times N$  píxeles, comparar cada píxel con sus vecinos, reemplazar el valor del píxel con un número binario basado en esta comparación y construir un histograma de  $N \times N$  píxeles.

La segunda forma de extracción de características implica aplicar un filtro de derivada parcial a la imagen, calcular la magnitud y orientación del gradiente, dividir la imagen en bloques de  $N \times N$  píxeles, construir un histograma basado en la magnitud del gradiente y normalizar el histograma.

Una vez que hemos extraído las características de las imágenes, con ambos métodos, debemos utilizar un clasificador para entrenar un modelo que pueda predecir si una imagen contiene un perro o un gato. Utilizaremos dos conjuntos de datos, uno con 200 imágenes de entrenamiento y 40 de prueba, y otro con 1000 imágenes de entrenamiento y 200 de prueba.

En este trabajo, hemos utilizado el clasificador SVM (Support Vector Machine) de la librería `scikit-learn` para clasificar imágenes de gatos y perros ya que hemos investigado y hemos visto que este tipo de clasificador da muy buenos resultados en esta clase de problemas de clasificación. Para optimizar los parámetros del SVM hemos utilizado una partición de validación mediante *ShuffleSplit* y *GridSearchCV* para obtener el mejor  $C$  y  $\gamma$  (los parámetros difieren para cada método de extracción de características).

Una vez obtenidos los parámetros, evaluamos la precisión del clasificador SVM y comparamos el rendimiento de ambas.

## Resultados:

Tras haber realizado la optimización de parámetros los resultados obtenidos son:

### Para el primer método:

En este caso hemos puesto valores default y  $\gamma = \text{'auto'}$  debido a que de otra forma obtenemos malos resultados (el optimizador fallaba con este método).

#### Dataset "cat\_dog\_100" :

- Precisión: 95%
- Tiempo de ejecución: 16.30 segundos

#### Dataset "cat\_dog\_500" :

- Precisión: 92,5%
- Tiempo de ejecución: 1 minutos 27.68 segundos

### Para el segundo método:

Parámetros del SVM ->  $C = 5.0$  y  $\gamma = 0.01$

#### Dataset "cat\_dog\_100" :

- Precisión: 92,5%
- Tiempo de ejecución: 16.46 segundos

#### Dataset "cat\_dog\_500" :

- Precisión: 90%
- Tiempo de ejecución: 1 minutos 27.68 segundos

## **Conclusión:**

En este trabajo se han implementado dos formas de extracción de características a partir de imágenes para ser utilizadas en un problema de clasificación supervisada, obteniendo en ambas diferentes pero en general buenos resultados.

Por lo que se puede concluir que el método de extracción de características es un método que funciona muy bien con los clasificadores SVM para este tipo de problema, y también que la selección del método de extracción y de los parámetros utilizados pueden tener un impacto significativo en el rendimiento del clasificador.

## **Video:**

<https://www.youtube.com/watch?v=P1R-ZUXZUCc>