

# Clasificación de imágenes: Adulto, No adulto

Sergio Marrero Marrero

Master Universitario en Sistemas Inteligentes y Aplicaciones Numéricas en la Ingeniería(MUSIANI)

Asignatura: Sistemas de Percepción

Tutor: Modesto Castrillón-Santana

Universidad Las Palmas de Gran Canaria(ULPGC)

1 de junio de 2016

## 1. Introducción

El contexto de este documento es la asignatura Sistemas de Percepción del Máster Oficial de Sistemas Inteligentes y Aplicaciones Numéricas en Ingeniería (SIANI)(año académico 2015/2016). Se ha realizado como tarea final para superar la asignatura un pequeño proyecto de clasificación de un individuo en las clases de *Adulto*, *No adulto*, a partir de una fotografía de su rostro (adulto > 18 años).

## 2. Explicación del conjunto de datos

El conjunto de datos escogido consta de 972 imágenes, de las cuales, 396 imágenes corresponden a adultos (40.74 % del conjunto de datos total), y 576 imágenes a no adultos (59.26 %).

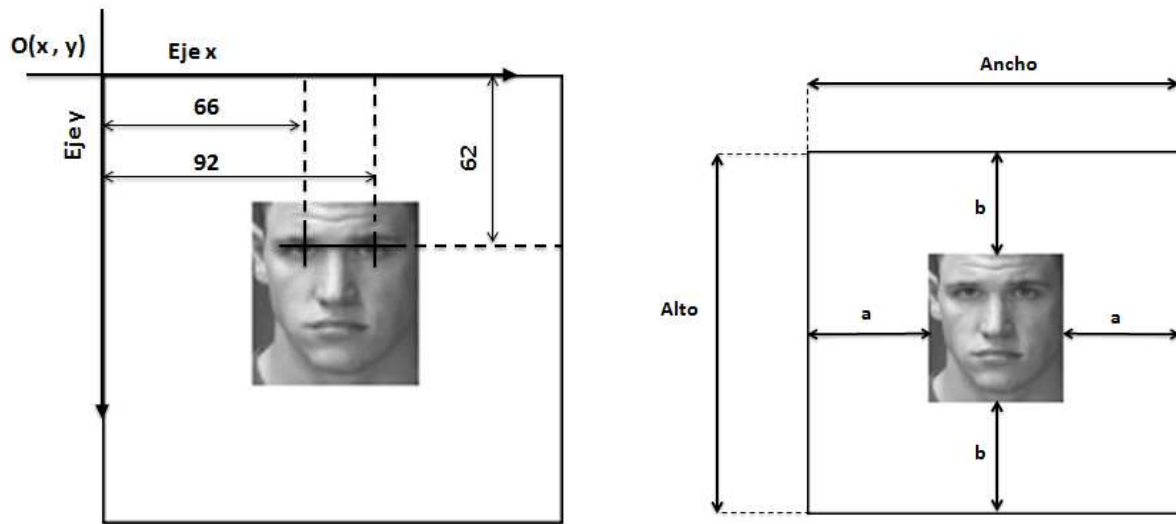
En la figura 1 se aprecian cuatro fotografías correspondientes a este conjunto de imágenes. Las dos de la izquierda pertenecen al subconjunto ‘Adulto’ y las dos de la derecha pertenecen al subconjunto ‘No Adulto’.



Figura 1: Muestras del conjunto de datos. a) y b) Adultos. c) y d): No adultos

Todo el conjunto de imágenes tiene las mismas dimensiones y se han efectuado las debidas rotaciones y traslaciones para que la línea ocular quede paralela al *eje x* y las coordenadas de los ojos queden en las mismas posiciones para cada imagen del conjunto de datos. Precisamente, la coordenada(en pixeles) de los ojos queda situada en las coordenadas  $[(66,62), (66,92)]$ . El eje que referencia dichas coordenadas está situado en la esquina superior izquierda, tal y como señala la figura 2a

El conjunto de imágenes se ha ajustado al rostro, eliminando así el contexto. En la imagen 2b podemos observar el recorte efectuado. El recuadro exterior marca el contorno de la foto original (159 x 155), en el interior de este recuadro se encuentra la porción de la imagen con la que se realizarán los experimentos, los valores a y b representan la cantidad recortada. La dimension del nuevo conjunto de datos es alto=65 *px*, y ancho = 59 *px*. (59 x 65) Los valores señalados en la figura 2b son: Alto=155 *px*, Ancho=159 *px*, a=50 *px*, b=45 *px*.



(a) Posición de los ojos

(b) Acercamiento al rostro

Figura 2: Explicación del conjunto de datos

### 3. Experimentos

#### 3.1. Primera fase

Se usará como descriptores, el espacio LBP(Local Binary Pattern) y el espacio HOG(Histogram of Oriented Gradients).

- LBP: Dentro del espacio LBP se revisarán las subvariantes *LBP uniforme*, *LBP invariante a rotaciones*, y *LBP uniforme e invariante a rotaciones*. Estos tres descriptores serán utilizados para representar las imágenes utilizando cuadrículas de tamaño 1x1 hasta 10x10, aumentando su tamaño de uno en uno.
- HOG: Dentro del espacio HOG se dividirán los sectores en 6 bins, 9 bins y 12 bins. Para cada uno de ellos se utilizarán cuadrículas de tamaño 3x3 hasta 12x12, aumentando su tamaño de uno en uno.

Cada sistema de representación fue entrenado utilizando un sistema de k-folds ( $k=5$ ). Se realizaron 5 tandas para cada experimento, calculándose luego la media de los valores obtenidos. Los clasificadores que se utilizaron para realizar dichos experimentos fueron el SVM (Support Vector Machine) utilizando dos kernel diferentes: Lineal ( $c=100, g=0.07$ ) y Función de Base Radial (Para ambos, los parámetros no especificados se suponen por defecto en la librería *lib-svm-3.21*).

En la tabla 1 se muestran los distintos ‘*Ratios de Acierto*’ obtenidos para las distintas variantes del descriptor LBP. En la primera columna se muestran las distintas cuadrículas que se utilizaron. La segunda, tercera y cuarta columna muestran respectivamente *LBP uniforme*, *LBP invariante a rotaciones* y *LBP uniforme e invariante a rotaciones* usando el clasificador SVM con kernel *Lineal*. Las columnas restantes muestran los mismos resultados pero usando el kernel *Función de Base Radial*. Los resultados de ambos experimentos se muestran en la gráfica a) y la gráfica b) de la figura 3.

En la tabla 2 se muestran los resultados para el descriptor HOC. La estructura de la tabla es idéntica a la tabla 1.

La figura 3 contiene cuatro gráficos. Los gráficos a) y b) se corresponde con la tabla 1 y los gráficos c) y d) se corresponde con la tabla 2.

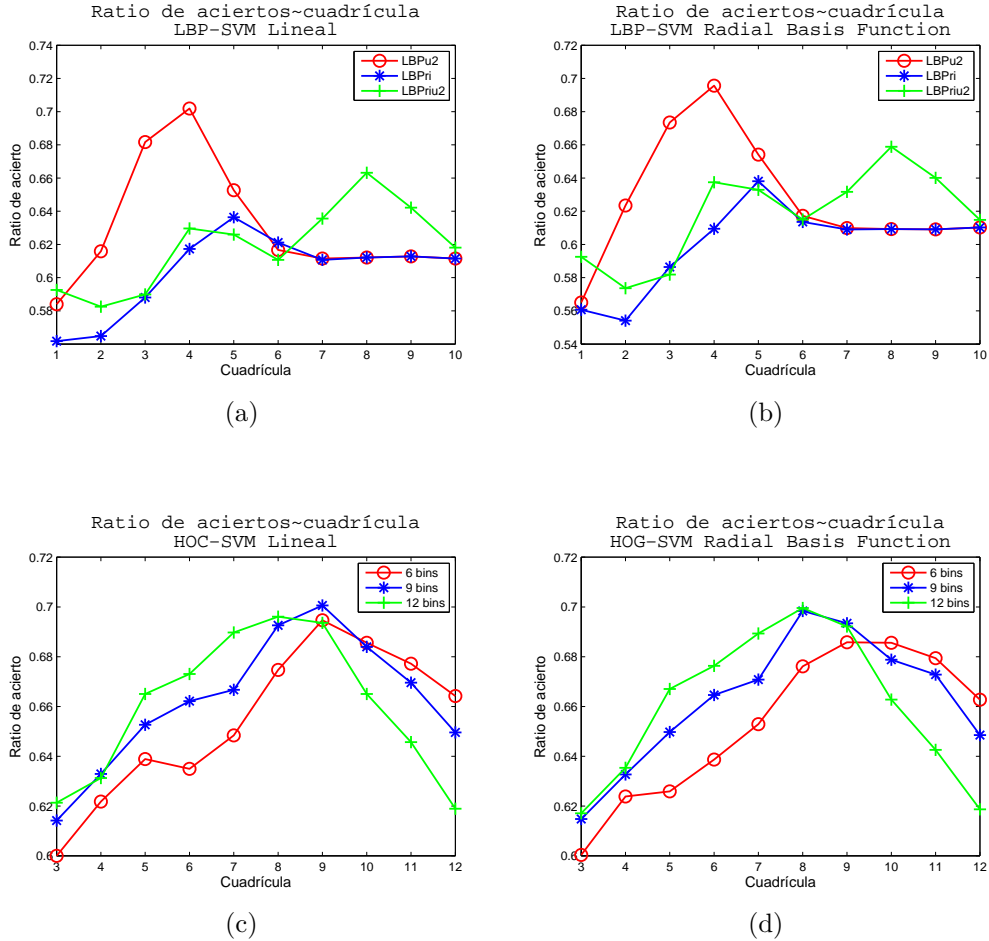


Figura 3: Resultados

	Lineal			Radial		
	U	R	UR	U	R	UR
1x1	58.40	56.17	59.26	56.50	56.07	59.26
2x2	61.59	56.48	58.25	62.34	55.41	57.37
3x3	68.17	58.81	58.99	67.35	58.64	58.19
4x4	70.19	61.73	62.96	69.56	60.94	63.75
5x5	65.27	63.64	62.59	65.41	63.81	63.29
6x6	61.67	62.12	61.07	61.73	61.36	61.50
7x7	61.15	61.07	63.56	60.99	60.90	63.17
8x8	61.21	61.21	66.32	60.93	60.93	65.89
9x9	61.28	61.28	64.22	60.91	60.91	64.01
10x10	61.15	61.15	61.81	61.03	61.03	61.48

Tabla 1: Ratio de aciertos de LBP usando SVM con kernel lineal y radial(expresado en porcentajes)

	Lineal			Radial		
	6	9	12	6	9	12
3x3	60	61.42	62.14	60.04	61.48	61.71
4x4	62.18	63.29	63.13	62.39	63.27	63.54
5x5	63.89	65.27	66.50	62.59	64.98	66.71
6x6	63.5	66.22	67.31	63.87	66.46	67.63
7x7	64.84	66.67	68.98	65.29	67.08	68.93
8x8	67.47	69.26	69.61	67.61	69.84	69.96
9x9	69.46	70.06	69.36	68.58	69.34	69.22
10x10	68.56	68.40	66.50	68.56	67.88	66.28
11x11	67.72	66.96	64.57	67.94	67.28	64.26
12x12	66.42	64.96	61.89	66.27	64.85	61.87

Tabla 2: Ratio de aciertos de HOC usando SVM con kernel lineal y radial(expresado en porcentajes)

Finalmente la tabla 3 muestra los mejores resultados obtenidos para cada sistema de representación y cada kernel.

	Tipo	Lineal	Tipo	Radial
LBP	u-4x4	70.19	u-4x4	69.56
	r-5x5	63.64	r-5x5	63.81
	ur-4x4	62.96	ur-9x9	64.01
HOC	6-9x9	69.46	6-9x9	68.58
	9-9x9	70.06	9-8x8	69.84
	12-8x8	69.62	12-8x8	69.96

Tabla 3: Mejores resultados obtenidos de cada descriptor y cada kernel

### 3.2. Segunda fase

Los dos mejores resultados obtenidos en la fase anterior fueron utilizados para hacer dos experimentos adicionales.

- *Stacking*: Se usará la técnica *stacking* con los dos mejores resultados de los experimentos realizados en el apartado anterior, precisamente con *LBP-uniforme-4x4* y *HOG-9 bins-9x9* en ambos usando como clasificador *SVM-Kernel Lineal*. El *stacking* se realizará con el clasificador *SVM-Lineal*.
- *Random Forest*: Idéntico al item anterior pero usando para cada uno de ellos el clasificador Random Forest.

Los resultados de ambos experimentos se pueden observar en la tabla 4

Clasificador	Ratio de aciertos
Stacking	70.72
RF: LBP-u-4x4	62.75
RF: HOG-9-9x9	66.87

Tabla 4: Ratio de aciertos de *Stacking* y *Random Forest* (expresado en porcentajes)

## 4. Conclusiones

Partiendo del conjunto de datos explicado en la sección 2 de este trabajo, se ha intentado buscar la mejor combinación de descriptores y clasificadores para clasificar el conjunto de datos en *Adulto* y *No adulto*. En una primera fase se han centrado los esfuerzos en observar el ratio de aciertos de distintas tipologías de los descriptores *LBP* y *HOG* frente al clasificador *SVM*, usando *kernel Lineal* y *Función de Base Radial*. En una segunda fase se han usado los dos mejores resultados de estas observaciones para realizar un proceso de *Stacking* por un lado y por otro lado la técnica de *Random Forest*.

Se observa en la tabla 1 y la grafica a) y b) de la figura 3 que el descriptor *LBP-uniforme* es el que mejor resultados obtiene. Los mejores resultados para este se encuentra entorno a cuadrículas de 4x4. Para el descriptor *HOG* (tabla 2 y grafica c) y d) de la figura 3) el óptimo se encuentra utilizando 9 bins y cuadrículas de 9x9. Respecto del clasificador utilizado(*SVM*), los mejores resultados se obtuvieron con *kernel lineal*, sin embargo no se aprecian muchas diferencias entre un kernel y otro.

Finalmente, se han utilizado los mejores resultados para realizar un proceso de *stacking* en el cual se mejoran algo los resultados obtenidos. También se utilizó *Random Forest*, sin embargo en este caso los resultados no mejoran.