

Resultados de los experimentos realizados de reconocimiento de género curso 2012/2013

Y. Álvarez Romero, T. Cristóbal Betancor, G. A. Cedeño Rodríguez,
V. M. Medina García, N. M. Monzón López, H. Pérez de la Nuez,
Á. Ramos de Miguel, R. Torón Domínguez y M. Castrillón Santana
SIANI

Universidad de Las Palmas de Gran Canaria (ULPGC)

mcastrillon@iusiani.ulpgc.es



Figura 1. Imágenes ejemplo de The Images of Groups, tras ser normalizadas y recortadas.

Abstract

Resumen de los resultados obtenidos en el experimento de clasificación de género, por los estudiantes matriculados en la asignatura Sistemas de Percepción del Máster Oficial de Sistemas Inteligentes y Aplicaciones Numéricas en Ingeniería en el curso 2012/2013. Palabras clave: reconocimiento de género, análisis facial.

1.. Introducción

En el contexto de la asignatura Sistemas de Percepción del Máster Oficial de Sistemas Inteligentes y Aplicaciones Numéricas en Ingeniería (SIANI), se ha celebrado entre los estudiantes del curso 2012/2013, una competición de reconocimiento de género.

El conjunto de datos escogido ha sido la división realizada por en el trabajo de Dago Casas [2] haciendo uso de un subconjunto anotado de la base de imágenes *The Images of Groups* [4].

El subconjunto se reduce a aquellas imágenes con al menos 20 píxeles de separación entre los ojos, superando las 14000 imágenes. Para cada muestra se realiza una etapa previa de normalización, que incluye la rotación de la línea ocular, y el escalado a un tamaño 159×155 con el centro de los ojos respectivamente localizados en las posiciones (66, 61) y (94, 61), como se puede ver en las Figuras 1 y 2.

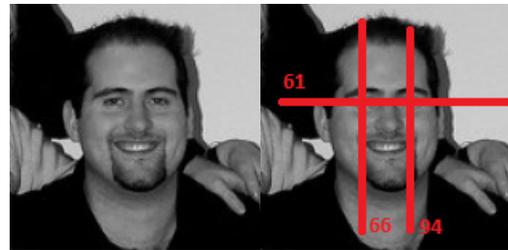


Figura 2. Imagen de ejemplo del conjunto de datos utilizado, a la izquierda, mostrando a la derecha las localizaciones de los ojos.

Los estudiantes parten de código matlab que permite la obtención de características frecuentes en la literatura de análisis facial automático: componentes principales (PCA) [7], histogramas de gradientes orientados (HOG) [3] (implementación del [6]) y patrones binarios locales (LBP) [1]. Además cuentan con el software para minería de datos Weka [5].

2.. Soluciones propuestas

Resumimos a continuación cada una de las soluciones y resultados obtenidos en orden alfabético.

2.1.. Yeray Álvarez Romero

Las imágenes de entrada se recortan de diversa forma para centrarse en la zona facial, o en la ocular conteniendo la zona nasal o sin ella, ver Figura 3.

Como descriptores hace uso de LBP uniforme o HOG, ambos con rejillas 3×3 y 7×7 . Para clasificar emplea Máquinas de vectores soporte (SVM) con diversos kernels y Random Forest (RF). Los resultados resumidos se muestran en el Cuadro 1.



Figura 3. Muestras de las zonas de imagen empleadas por Yeray Álvarez: a) Cara recortada en las dimensiones 90×70 , b) recorte de ojos-nariz en las dimensiones 35×55 y c) recorte de la zona de los ojos en las dimensiones 16×55 .

Cuadro 1. Resultados obtenidos por Yeray Álvarez.

Zona	Clas.	LBP		HOG	
		3x3	7x7	3x3	7x7
Cara	SVM-lin	52.81 %	60.83 %	52.73 %	59.62 %
	SVM-RBF	49.51 %	60.28 %	53.13 %	64.86 %
	SVM-sig	52.97 %	57.31 %	55.24 %	65.49 %
	RF-10	50.41 %	52.04 %	49.66 %	56.32 %
	RF-30	50.31 %	52.82 %	51.40 %	59.44 %
Ojos	SVM-lin	52.81 %	52.55 %	52.73 %	51.28 %
	SVM-RBF	49.51 %	49.51 %	55.13 %	56.13 %
	SVM-sig	52.97 %	53.73 %	55.24 %	56.63 %
	RF-10	50.41 %	51.50 %	49.66 %	51.03 %
	RF-30	50.31 %	50.65 %	51.40 %	52.96 %
Ojos y nariz	SVM-lin	51.91 %	47.04 %	52.87 %	49.68 %
	SVM-RBF	50.42 %	49.51 %	54.65 %	50.87 %
	SVM-sig	50.15 %	49.53 %	54.29 %	50.32 %
	RF-10	50.23 %	50.76 %	53.52 %	51.68 %
	RF-30	50.61 %	49.44 %	54.79 %	51.60 %



Figura 4. Respectivamente: imagen original, reducida al 50 % e imagen parcial, utilizadas en el experimento de Teresa Cristóbal.

2.2.. Teresa Cristóbal Betancor

Como datos de entrada hace uso de la imagen original escalada al 50 % (78×80) y una imagen parcial centrada en el rostro, pero ocupando también el pelo y cuello (155×51), pueden verse en a Figura 4.

Como características hizo uso de LBP uniformes en una rejilla 3×3 , y HOG en una rejilla 7×7 . Para la clasificación, empleó SVM con kernel de base radial (RBF) y Adaboost con un a única división. Su resumen de resultados se presenta en el Cuadro 2.

	LBP		HOG	
	SVM	Adaboost	SVM	Adaboost
Reducida	76.4	75.9	81.7	79.7
Parcial	78.8	76.1	80.9	79.2

Cuadro 2. Cuadro comparativo de las diferentes configuraciones de la imágenes de entrada, en experimento de Teresa Cristóbal.

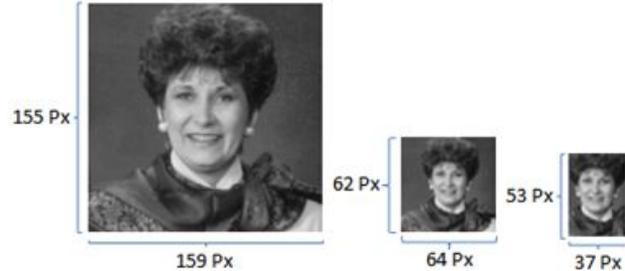


Figura 5. Relación del aspecto final de una imagen ejemplo considerando el redimensionamiento y el recorte para las pruebas, en el experimento de Guillermo Cedeño.

	LBP		HOG	
	SVM	Bagging	SVM	Bagging
Reducida	73.48	72.87	82.11	81.19
Parcial	71.79	71.6	83.49	82.91

Cuadro 3. Cuadro comparativo de las diferentes configuraciones de la imágenes de entrada, en experimento de Guillermo Cedeño.

	LBP		HOG	
	IBK(3)	SVM	IBK(3)	SVM
Reducida	70.83	81.98	76.39	82.82
Ojos	64.0	73.48	71.81	78.34
Cara	78.43	83.76	79.83	83.63

Cuadro 4. Cuadro comparativo de las diferentes configuraciones de la imágenes de entrada, en experimento de Víctor Medina.

2.3.. Guillermo Augusto Cedeño Rodríguez

Redimensiona las imágenes originales al 40 %, trabajando con una imagen 62×64 y una zona recortada centrada en el rostro de 53×37 , ver la Figura 5.

Como características hace uso de LBP uniforme e invariante a rotaciones en una rejilla 5×5 y HOG con una rejilla 7×7 . Clasifica con SVM y Bagging con SVM, siempre con kernel lineal. Los resultados se presentan en el Cuadro 3.

2.4.. Víctor Manuel Medina García

Como datos de entrada hace uso de la imagen escalada al 50 % (78×80), una zona centrada en los ojos (59×16), o en el rostro (61×91).

Las características extraídas de estas imágenes usan LBP uniforme con una rejilla 5×5 y HOG con una rejilla 7×7 y 9 bins. Para clasificación ha probado con IBK ($k = 3$) y SVM con kernel lineal. Los resultados se presentan en el Cuadro 4.

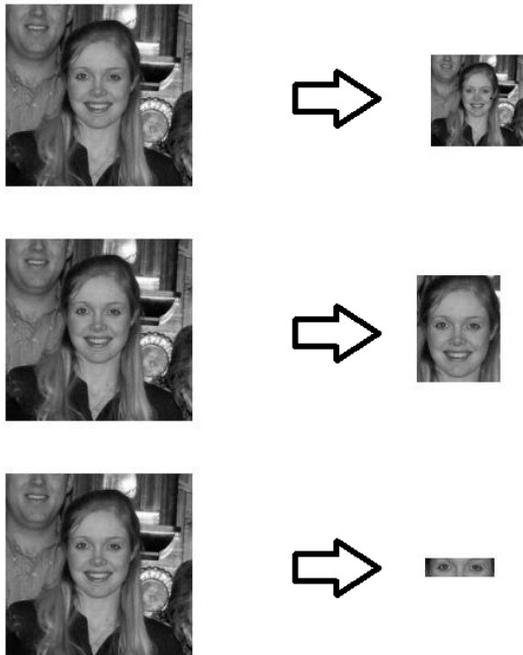


Figura 6. Zonas de la imagen utilizadas en el experimento de Víctor Medina.

2.5.. Nelson Manuel Monzón López

Opta por tomar tres zonas diferentes de la imagen: la zona ocular, facial, y la imagen original escalada un 50 %, ver Figura 7. Como descriptores hace uso de LBP uniforme con una rejilla 3×3 y HOG con rejilla 7×7 . Como clasificadores ha optado por J48, Random Forest (RF) e IBK ($k=3$), incluyendo el resumen de resultados en el Cuadro 5.

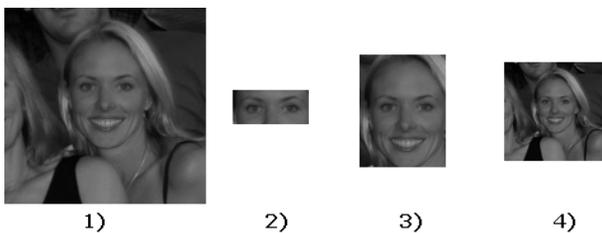


Figura 7. Ejemplos de imágenes del dataset empleado en cada apartado. Se observa la imagen original, la zona de los ojos (25×60), el rostro (80×70) y escalada al 50 %.

2.6.. Héctor Pérez de la Nuez

En su propuesta opta por reducir las imágenes de entrada al 60 %, así como, sin escalar, centrarse en una imagen más centrada en el rostro y en la zona ocular, como se observa en la Figura 8.

Como características hace uso de rejillas HOG 7×7 y

			Total
Reducida	LBP	RF	74.59 %
		J48	64.55 %
		IBK	68.65 %
	HOG	RF	80.72 %
		J48	70.07 %
		IBK	64.55 %
Rostro	LBP	RF	78.21 %
		J48	66.41 %
		IBK	76.45 %
	HOG	RF	81.16 %
		J48	70.26 %
		IBK	79.61 %
Ojos	LBP	RF	73.37 %
		J48	63.17 %
		IBK	70.27 %
	HOG	RF	78.64 %
		J48	67.51 %
		IBK	74.76 %

Cuadro 5. Resumen de resultados obtenidos en la propuesta de Nelson Monzón.



Figura 8. Respectivamente: imagen reducida al 60 % (97×92), imagen de la cara (91×86) y zona ocular (56×26) empleadas.

			Total
Reducida	LBP	Random Forest	68,16 %
		SVM	69,34 %
	HOG	Random Forest	81,91 %
		SVM	82,65 %
Ojos	LBP	Random Forest	71,52 %
		SVM	71,61 %
	HOG	Random Forest	79,24 %
		SVM	80,81 %
Cara	LBP	Random Forest	70,33 %
		SVM	72,18 %
	HOG	Random Forest	83,02 %
		SVM	84,77 %

Cuadro 6. Resumen de resultados obtenidos en la propuesta de Héctor Pérez.

LBP invariante a rotaciones calculado sobre una rejilla 3×3 . En la etapa de clasificación emplea SVM con kernel lineal y Random Forest con 50 árboles. Los resultados se presentan en el Cuadro 6.

2.7.. Ángel Ramos de Miguel

A partir de las imágenes de entrada, realiza un escalado al 20 %, y de esta imagen escalada se selecciona la sección del rostro con el contorno por un lado (45×40), y por otro



Figura 9. Imágenes utilizadas por Angel Ramos.

	PCA	LBP	HOG	PCA+HOG
SVM	72.98	71.04	80.69	81.72
IB1	71.49	64.61	74.91	75.60
IB3	72.33	65.66	76.52	77.17
J48	65.05	61.84	69.93	69.80

Cuadro 7. Resultados empleando la cara completa en el experimento de Ángel Ramos.

	PCA	LBP	HOG	PCA+HOG
SVM	72.16	66.58	78.92	79.58
IB1	70.17	58.59	71.95	72.66
IB3	72.02	59.45	73.80	74.59
J48	67.14	57.24	67.98	68.47

Cuadro 8. Resultados empleando la zona ocular en el experimento de Ángel Ramos.

	LBP 5 × 5	HOG 16 × 16	HOG 32 × 32
SVM lineal	83,12 %	84,51 %	85,88 %
SVM RBF	80,43 %	84,17 %	86,41 %

Cuadro 9. Resumen final comparativo de la propuesta de Rubén Torón, incluyendo las tasas medias según descriptores y clasificadores SVM Lineal y SVM RBF

la zona de los ojos 13×40 , como se aprecia en la Figura 9.

Como clasificadores hace uso de SVM, utilizando un kernel lineal, IB1, IB3 y J48. Los resultados obtenidos se pueden consultar en los Cuadros 7 y 8.

2.8.. Rubén Valentín Torón Domínguez

Esta propuesta hace uso de descriptores calculados sobre las imágenes de entrada sin modificación. Los descriptores utilizados son LBP uniforme para componer un histograma sobre una rejilla 5×5 . También el experimento incluye el uso de características HOG sobre rejillas 32×32 y 16×16 .

Para la clasificación emplea SVM probando con dos tipos de kernels, el lineal y la función de base radial (RBF sus siglas en inglés). Los resultados pueden verse en el Cuadro 9.

3.. Conclusiones

Con la excepción de los resultados presentados por Y. Álvarez, todos los participantes han alcanzado tasas de reconocimiento por encima del 80 %, lo cual, dado el uso de

similares características y resoluciones de imágenes parecen indicar una realización errónea del experimento en ese caso concreto, al lograr escasamente una mejor tasa de reconocimiento que no supera el 65 % en un problema, no olvidemos, biclase.

En el resto de pruebas experimentales, los resultados obtenidos parecen proporcionar una mejor tasa haciendo uso de características basadas en HOG, y preferentemente con clasificadores basados en máquinas de vectores soporte (SVM). En el caso de Nelson Mónico, al no incluir en su comparativa clasificadores basados en SVM, los mejores resultados se obtienen con Random Forest, pero sin lograr las tasas obtenidas por otros participantes haciendo uso de SVM.

En cuanto a la zona de la imagen utilizada, la zona ocular presenta buenas tasas, aún suponiendo una notable reducción del área de interés. Sin embargo no parece haber una conclusión sobre si es mejor hacer uso sólo del rostro o de la imagen contexto. En la mayoría de los experimentos presentados siempre la imagen de contexto incluye una menor resolución del rostro, lo que puede afectar al reconocimiento. Si bien la comparativa no fue realizada por todos, los clasificadores basados en el rostro exclusivamente obtienen mejores resultados para G. Cedeño, V. Medina, N. Monzón y H. Pérez; y a la inversa para T. Cristóbal.

Referencias

- [1] T. Ahonen, A. Hadid, and M. Pietikäinen. Face description with local binary patterns: Application to face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(12), December 2006.
- [2] P. Dago-Casas, D. González-Jiménez, L. Long-Yu, and J. L. Alba-Castro. Single- and cross- database benchmarks for gender classification under unconstrained settings. In *Proc. First IEEE International Workshop on Benchmarking Facial Image Analysis Technologies*, 2011.
- [3] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In C. Schmid, S. Soatto, and C. Tomasi, editors, *International Conference on Computer Vision & Pattern Recognition*, volume 2, pages 886–893, INRIA Rhône-Alpes, ZIRST-655, av. de l'Europe, Montbonnot-38334, June 2005. http://ieeexplore.ieee.org/xpls/abs_all.jsp?isNumber=31472&arNumber=1467360&isnumber=31472&arNumber=1467360.
- [4] A. Gallagher and T. Chen. Understanding images of groups of people. In *Proc. CVPR*, 2009.
- [5] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten. The weka data mining software: An update. *SIGKDD Explorations*, 11(1), 2009.
- [6] O. Ludwig, D. Delgado, V. Goncalves, and U. Nunes. Trainable classifier-fusion schemes: An application to pedestrian detection. In *12th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, 2009.
- [7] M. Turk and A. Pentland. Eigenfaces for recognition. *J. Cognitive Neuroscience*, 3(1):71–86, 1991.