

# ¿Dónde está Wally?

25 de noviembre de 2010

## Autores:

Creo Martínez, Javier	javiercreomartinez@gmail.com
Diego Dios, Pedro José	pedro.diegodios@gmail.com
Pérez López, Carlos	carlospzllz@gmail.com
Puig González, Pablo	pablo.puig@gmail.com

*“Wally es uno entre la multitud, pero esté donde esté tienes que encontrarlo. ¿Por qué? porque está ahí, porque se llama Wally, porque el nombre del juego es “¿dónde está Wally?”. Búscalos en la playa, en la estación, en el camping, en el estadio. Dondequiera que haya una muchedumbre, ahí está Wally, incansable, pasando desapercibido por entre la gente”*

# Contenido

1 Introducción .....	5
2 Descripción del problema .....	6
2.1 Objetivo.....	6
2.2 Resolución .....	6
2.3 Conclusión .....	6
2.4 Validez del Sistema.....	6
4. Estado del arte .....	8
4.1 Estudio del proceso biológico (MIT).....	8
4.2. Trabajos anteriores .....	9
5. Justificación del uso de una RNA.....	10
6. Primera aproximación .....	11
6.1. Descripción .....	11
6.2. Resultados .....	12
6.3. Discusión .....	15
7. Segunda aproximación .....	16
7.1. Descripción .....	16
7.2 Resultados .....	17
7.3 Discusión .....	19
8. Tercera aproximación.....	21
8.1. Descripción .....	21
8.2 Resultados .....	22
8.3 Discusión .....	24
9. Cuarta aproximación .....	25
9.1 Descripción.....	25
9.2 Resultados .....	25
9.3 Discusión .....	27
10. Conclusión .....	29
11. Trabajo futuro .....	30
12. Referencias.....	31



## 1 Introducción

Una red de neuronas artificiales es un sistema de procesado de información perteneciente a la rama de la Inteligencia Artificial orientada a la emulación; así, imita tanto el funcionamiento como la estructura presente en un cerebro animal. Para ello consta de un conjunto de elementos de cómputo interconectados entre sí que reciben información del entorno y producen una salida.

Las ventajas de utilizar redes de neuronas son abundantes, las más interesantes son:

- Aprendizaje: las redes tienen la habilidad de aprender mediante una etapa de aprendizaje. En la que se le da datos de entrada y sus correspondientes salidas esperadas.
- Tolerancia a fallos: las redes pueden seguir respondiendo de manera aceptable aun si se daña parcialmente debido a que almacena información de forma redundante.
- Flexibilidad: pueden manejar cambios no importantes de la información de entrada(ruido ...)
- Tiempo real: pueden ser implementadas en ordenadores o dispositivos electrónicos para obtener respuestas en tiempo real.
- Auto-organización: las redes de neuronas crean su propia representación de la información en su interior así el usuario no tiene que preocuparse de esto.

Su ámbito de aplicación es muy amplio siendo uno de los principales el tratamiento de imágenes; que será, en nuestro caso, el objeto de estudio. Además también se usan mucho en el campo de la clasificación.

En nuestro trabajo trataremos de aplicar esta metodología para resolver el clásico problema de "Buscando a Wally" mediante el análisis de imágenes escaneadas de los libros de la saga.

## **2 Descripción del problema**

### **2.1 Objetivo**

Encontrar a Wally en las imágenes de los conocidos libros infantiles de la saga "¿Dónde está Wally?"(ver fig.1)

### **2.2 Resolución**

Hemos considerado inicialmente entrenar imágenes de Wally de cuerpo completo, en posiciones similares, con la finalidad de que al pasarle al sistema una imagen escaneada del libro, éste nos indique su posición.

Sin embargo, nos hemos encontrado con los siguientes problemas:

- En la mayoría de los paisajes en los que se esconde Wally, éste aparece parcialmente oculto, siendo visible en bastantes casos únicamente su cabeza.
- Los complementos de Wally varían, tapando parte de su vestimenta primaria; así encontramos distintos Wallys equipados con cámara de fotos, mochila y otros objetos que entorpecen el reconocimiento de Wally.
- La similitud con el cuerpo y vestimenta de otros muchos personajes del paisaje resulta en ocasiones abrumadora. En concreto nos encontramos con algún caso extremo en el que todos los personajes de la imagen están vestidos con la misma indumentaria que Wally.

### **2.3 Conclusión**

Basándonos en los problemas encontrados en un estudio preliminar y en soluciones adoptadas por otros trabajos similares, hemos decidido continuar nuestro planteamiento inicial, basándonos en todo el cuerpo del personaje, teniendo en cuenta que en ocasiones es muy complicado de identificar. Emplearemos imágenes no muy complejas dónde Wally aparezca visible de cuerpo entero para facilitar el estudio.

### **2.4 Validez del Sistema**

Valoraremos como resultado satisfactorio todos aquellos casos en los que el sistema encuentre a Wally. Además encontraremos como resultados aceptables la detección de Walda y Odlaw. (ver fig.2) Cualquier otro falso positivo será caso de estudio

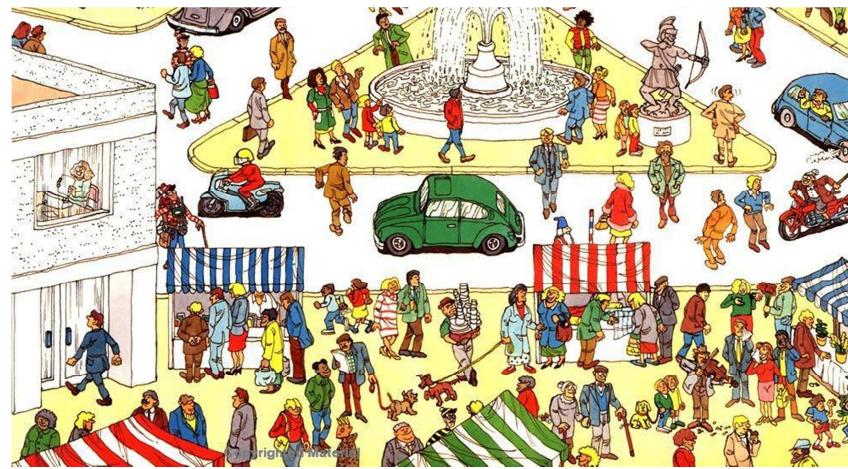


Fig.1: Imagen escaneada del libro que nos sirve como entrada

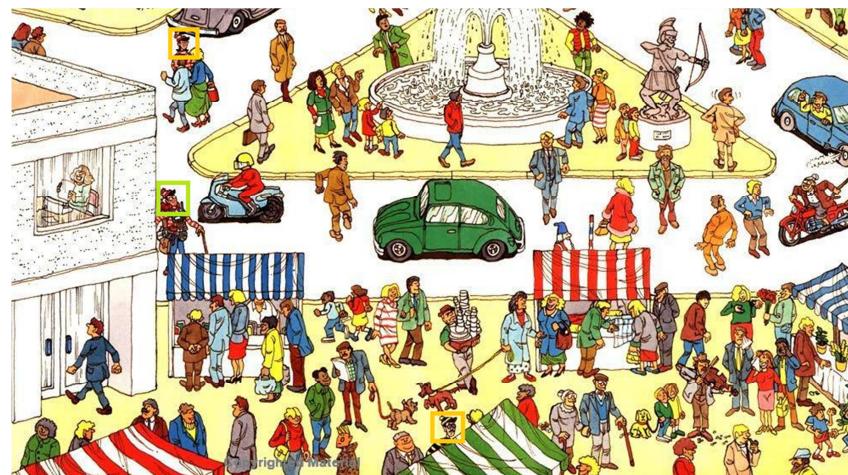


Fig.2: resultado deseado del sistema para la imagen de la figura 1

## 4. Estado del arte

### 4.1 Estudio del proceso biológico (MIT)

Para estudiar la búsqueda de Wally empleando un sistema de neuronas artificiales sería interesante echar primero un vistazo a cómo un sistema de neuronas naturales, un cerebro, realmente lo hace.

El 8 de Junio del 2008 fue publicado un artículo en el Science Daily que explica de manera cercana qué es lo que realmente sucede en nuestras cabezas cuando nos enfrentemos a unas páginas de estos libros abarrotadas de gente en las que una de esas minúsculas personas es Wally.

A cada momento, el mundo bombardea los sentidos con más información de la que el cerebro puede procesar, y durante más de un siglo los científicos y psicólogos han debatido cómo hace el cerebro para filtrar las distracciones y centrar la atención en las cosas que realmente importan.

El profesor Robert Desimone, director del Instituto McGovem dedicado a la investigación cerebral en el MIT, y sus colegas exponen que las neuronas sincronizan sus señales para llamar la atención, al igual que un coro de gente alzando su melodía sobre el ruído de los murmullos en una habitación abarrotada de gente.

Cuando buscamos a Wally entre la gente, ¿qué es lo que realmente estamos haciendo? ¿realiza el cerebro un escaneo de la página espacialmente (procesado secuencial)? ¿O toma el cerebro una imagen total de la página y hace zoom gradualmente cuando encuentra características relevantes como el color o la forma de Wally (procesado paralelo)? En el primer modelo, el punto de luz de la atención se mueve a través de la página, comprobando cada detalle con la imagen mental que tenemos de Wally. Mientras que en el segundo modelo, el color rojo y su forma vienen gradualmente a primer plano y los otros colores y formas pasan a formar parte de un segundo plano.

Durante décadas, estos dos campos han estado claramente diferenciados, pero recientes evidencias hicieron sospechar a algunos científicos que la conducta del cerebro es una combinación de los dos.

La neurona especializada en detectar el rojo emite señales más intensas cuando lo consigue, pero no es sólo esto lo que explica cómo consigue hacerse oír por encima de la cacofonía de las otras neuronas. "No creemos que sea asunto de una sola neurona," postulan. "Es como esa neurona coopera con otras neuronas y hacen oírse sus voces. Cuando las neuronas sincronizan su señal, se incrementa su actividad".

Desimone emplea otra vez la analogía de una habitación llena de gente hablando. Si aleatoriamente personas individuales alzan su voz, la habitación se vuelve

simplemente todavía más ruidosa. En cambio, si un grupo de gente empieza a cantar al unísono, sus voces se alzan por encima del ruído de fondo de la gente.

La sincronización de las señales explica como el cerebro usa el procesamiento paralelo para concentrarse en características relevantes dentro de la compleja escena. Después, el cerebro cambia a procesamiento secuencial o en serie, scrutando cada objeto de manera precisa para hallar el objeto deseado.

Nuestro sistema empleará únicamente procesado secuencial, ya que emular tal y como lo hace el cerebro, además de ser mucho más potente, también sería mucho más complejo.

## 4.2. Trabajos anteriores

El reto de encontrar a Wally, ya se ha intentado resolver anteriormente mediante el uso de RNAs. Es el caso del trabajo "Specific face recognition using a backpropagation neural network". Este trabajo consiste en el reconocimiento de caras en una escena mediante el uso de una red binaria y una red de propagación hacia atrás.

La red neuronal fue construida en MATLAB y fue entrenada con el algoritmo de aprendizaje con propagación hacia atrás hasta un error cuadrático medio menor que  $10^{-5}$ . La arquitectura de cada una de las redes neuronales consistía en 500 entradas, 50 unidades en la capa oculta y una salida. Las 500 entradas venían de dividir la imagen de entrada de 850x1 150 en ventanas de 20x25. Había 50 unidades ocultas porque la red tenía que reconocer las imágenes que son Wally y las que no son Wally. Y una salida porque consiste en decir si Wally está o no en la ventana. Los resultados de este trabajo fueron pobres en cuanto a encontrar a Wally en las imágenes, pero sin embargo daba buen resultado diciendo que imágenes no eran Wally. Una de las razones del problema de encontrar a wally es que éste no está siempre dibujado de la misma forma cambiando mucho de una imagen a otra.

## 5. Justificación del uso de una RNA

Como ya hemos adelantado en la introducción del trabajo, las redes de neuronas poseen una serie de características que la hacen muy interesantes respecto a los métodos tradicionales.

Para nuestro problema en concreto, estas redes poseen unas características que las hacen apropiadas por ejemplo la flexibilidad, Wally aparecerá dibujado de formas bastante diferentes en las imágenes, debido a esto es complicado determinar de manera formal cómo es Wally. Todo el mundo sabe identificar a Wally, pero no sabe cómo definirlo formalmente. Ésta es la principal razón que justifica el uso de una red de neuronas.

Otra de los motivos es que la simplicidad de programarlo, ya que solo nos tenemos que preocupar de escoger el tipo de red, de su entrenamiento y no como distribuir los pesos, que es una tarea más laboriosa. De eso ya se encarga la propia red.

En nuestro caso tenemos relativamente fácil conseguir pares de entradas y salidas para el entrenamiento lo cual es una ayuda para emplear este tipo de redes.

## 6. Primera aproximación

### 6.1. Descripción

Hemos tomado como patrones positivos todos los Wally's presentes en las páginas de tres libros de "Dónde está Wally" escaneados. De todas estas imágenes hemos descartado, en primer lugar, todos aquellos patrones en los que el cuerpo de Wally no era visible, para trabajar en la primera aproximación de resolución del problema.

Como patrones negativos, hemos extraído algunos patrones al azar de las hojas escaneadas así como algún patrón que por sus características podría presentar ciertos problemas, esto es, por similitud de color o forma (en nuestro caso, para la primera aproximación, de color).

Para entrenar la red en esta primera aproximación hemos trabajado con los patrones anteriormente descritos, podemos ver un ejemplo de los patrones en las imágenes inferiores.



Fig.3 patrón positivo



Fig.4 patrón negativo

La red tiene 3 entradas que corresponden al porcentaje de los 3 colores más representativos de wally. Estos 3 colores escogidos han sido: azul de los pantalones, rojo de la camiseta y color carne de la cara y manos de wally. Tiene una sola salida que oscila entre 0 y 1, para que tomemos como positivo la entrada la salida tiene que superar un umbral.

Para hallar el valor de estos colores hemos tomado un valor medio de las los patrones con tonalidades más distantes. Los valores RGB para cada color, y el porcentaje promedio presente en los patrones positivos, han sido los siguientes:

**Rojo: [230 120 120] 8.3%**  
**Azul: [140 175 215] 12.9%**  
**Piel: [255 216 147] 7.6%**

Para obtener los porcentajes de color de los patrones entrenados nos apoyamos en la función “colourRatio” que tiene como parámetros de entrada el patrón en cuestión, el valor del color del que queremos conocer el porcentaje presente y un factor de tolerancia para ese color. En la primera aproximación hemos utilizado un **factor de tolerancia=30** para cada uno de los colores anteriormente cuantificados.

De esta forma para entrenar la red se le pasa como entradas de la red las proporciones de colores de las imágenes.

Los motivos de esta elección es que hemos considerado que los atributos más característicos de Wally son los colores que nunca varían en el personaje (bueno, para ser exactos diremos que varían en tonalidad).

## 6.2. Resultados

Hemos utilizado un total de **24 patrones positivos y 22 patrones negativos**. En estos simplemente se han extraído los colores que comentamos anteriormente sin aplicar ningún filtro y posteriormente normalizados entre 0 y 1. Los porcentajes que hemos utilizado para la entrada, test y validación son 60%,20% y 20% respectivamente.

Para la elección de la arquitectura a utilizar hemos realizado pruebas con diferentes topologías, obteniendo como resultados los visibles en la tabla inferior. Se puede observar que la Arquitectura [8] presenta unos resultados levemente mejores que el resto de arquitecturas. Aunque como podemos observar, la desviación típica es 14,69, con lo que no podemos estar seguros de que sea realmente la mejor arquitectura, para ello habría que realizar un estudio estadístico.

Arquitectura	Entrenamiento		Validación		Test	
	Precisión	$\sigma$	Precisión	$\sigma$	Precisión	$\sigma$
[2]	74.6%	15.02	81.71%	16.3	67.71%	20.57
[3]	77.37%	12.65	83.14%	13.74	74.86%	17.45
[8]	81.13%	13.6	85.14%	15.8	76.86%	14.69
[15]	78.06%	13.13	83.71%	14.14	75.14%	19.75
[25]	79.31%	13.39	80.57%	15.74	72.85%	20.86
[8][5]	82.19%	10.64	85.71%	13.23	73.71%	15.1
[20][5]	80.25%	12.7	83.14%	15.73	73.14%	18.85
[20][12]	81.37%	15.88	83.71%	11.18	72%	20.4

Fig.5: Tabla de topologías

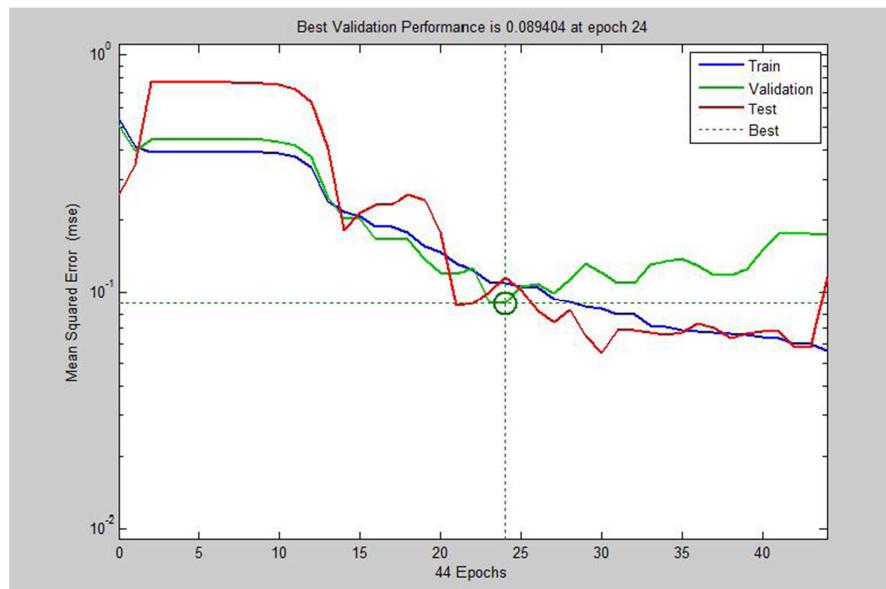


Fig.6: Gráfico de entrenamiento

La figura 6 corresponde con un gráfico de entrenamiento característico de la red.

Y en la figura 7 se muestra la matriz de confusión, en la cual vemos que de los patrones negativos toma 4 como positivos y de los positivos toma 2 como negativos.

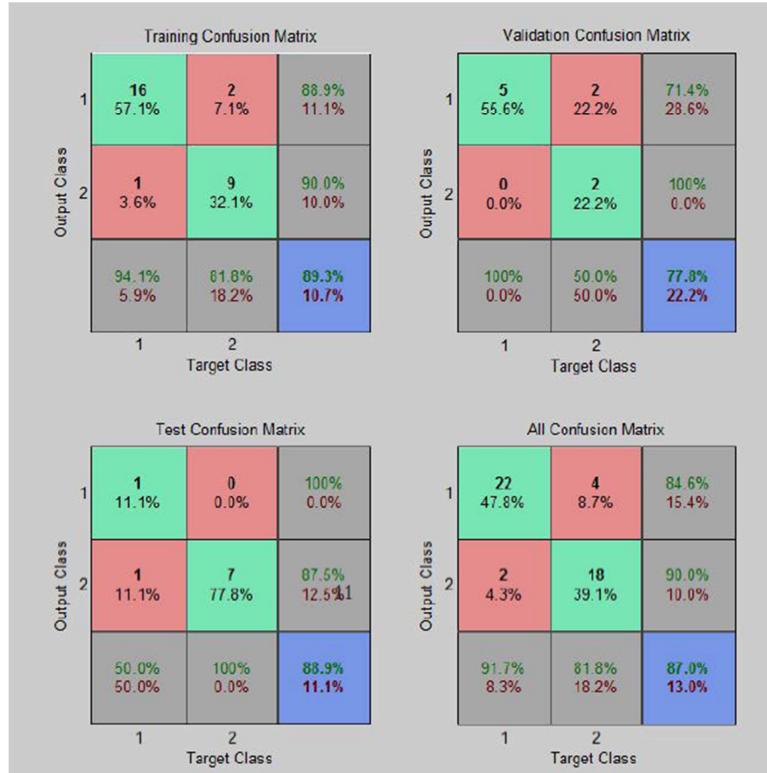


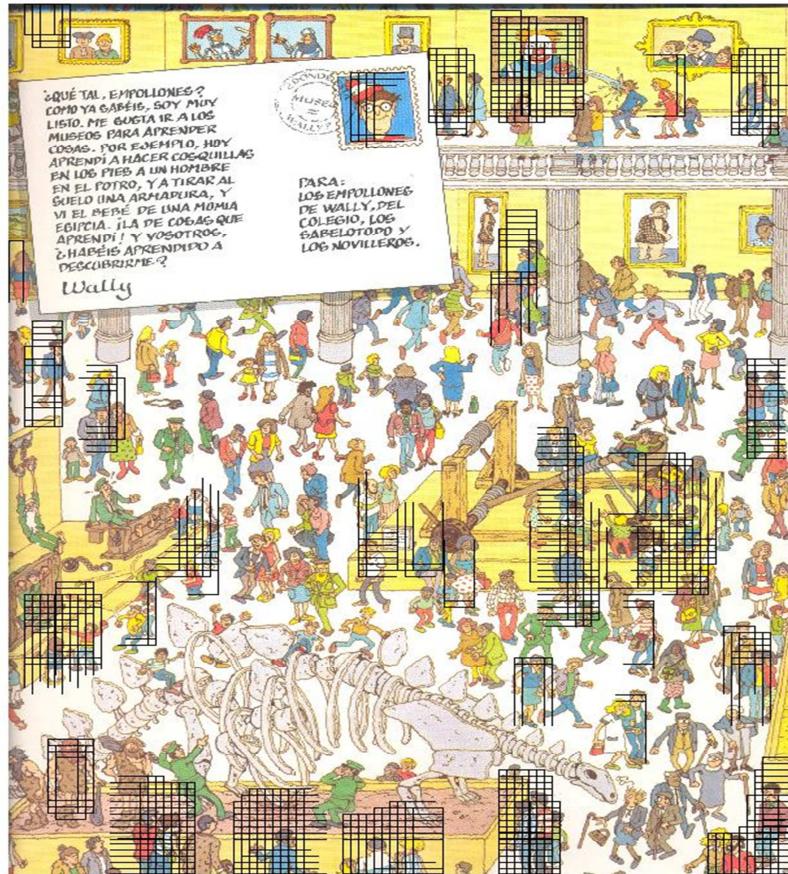
Fig.7: Matriz de Confusión

Una vez entrenada la red con la función `trainingNet` con una topología de 1 capa oculta y 8 neuronas, se procede a aplicar a la red una imagen del libro de Wally con la función de `findWaldo`, la cual debería dar como resultado ideal la posición del personaje buscado, Wally.

En un primer momento la función recorría toda la imagen ventana a ventana tomando como ventana un cuadrado de pequeñas dimensiones. Tras obtener resultados pobres, decidimos cambiar la ventana a un formato más adecuado, similar a los patrones de entrenamiento, es decir, cogimos ventanas rectangulares de proporciones similares. Además para incrementar la velocidad de búsqueda recorremos vectores de ventanas en conjunto devolviendo los resultados también en un vector, en vez de tratar cada ventana por separado.

Cada vez que encontramos un supuesto Wally para la función, mostramos por pantalla las coordenadas del mismo así como el porcentaje de “ser el auténtico Wally” que presenta esa ventana. Al terminar de evaluar la hoja escaneada, obtenemos la misma hoja resaltada con rectángulos negros para cada resultado positivo según el umbral que

hemos adoptado como umbral mínimo (un **umbral=0.5**) a partir del cual consideraremos una ventana como un Wally auténtico (véase figura 8).



*Fig.8: Imagen resultante del sistema para la primera aproximación*

### 6.3. Discusión

Observando los resultados obtenidos, vemos que nuestra red encuentra a Wally en la mayoría de los casos pero también detecta muchos falsos positivos.

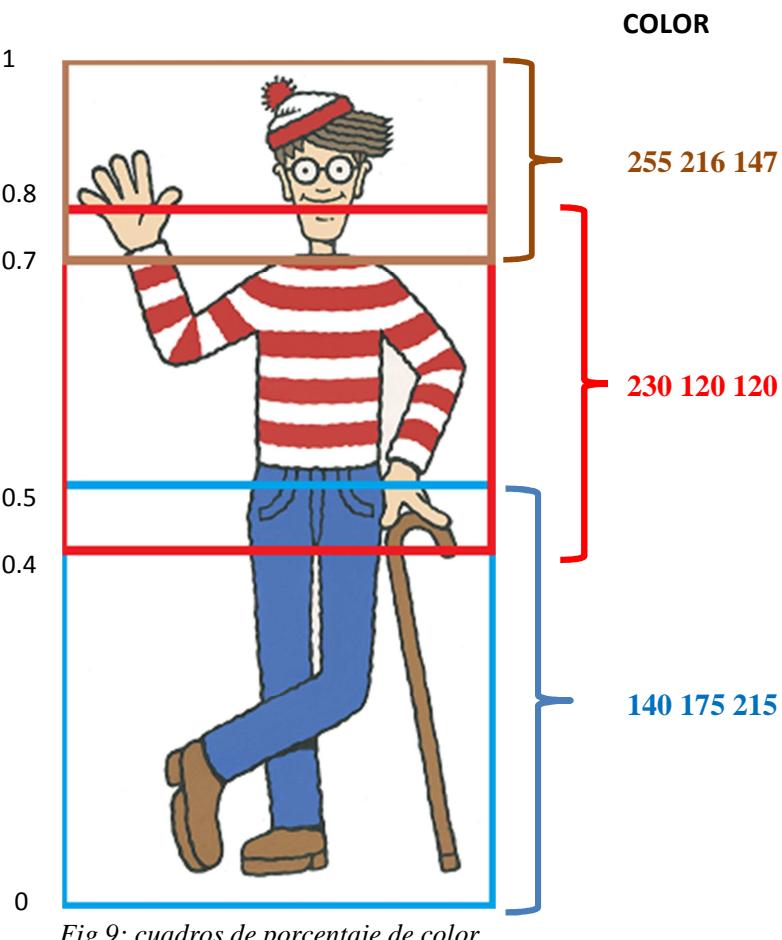
Creemos que esto es debido a que los porcentajes de colores de los falsos Wallys son muy similares a los del verdadero Wally.

Para refinar, en la siguiente aproximación además de mirar el porcentaje de los colores, también vamos a comprobar la situación de estos en la ventana.

## 7. Segunda aproximación

### 7.1. Descripción

Para mejorar la implementación anterior hemos considerado la posición de Wally en las imágenes; esto es, en todas ellas el pantalón de Wally ocupará la parte inferior de la imagen, la camiseta roja y blanca la parte intermedia y la cara la parte superior. Por lo tanto hemos modificado la entrada de la red para indicar qué zonas en los recuadros buscados tendrán uno u otro porcentaje de cada color. Hemos decidido solapar las tres zonas debido a que Wally no siempre está centrado ni siempre presenta el mismo formato en los rectángulos utilizados para su búsqueda, así como en los patrones utilizados en el entrenamiento(ver fig.9).



## 7.2 Resultados

Para esta iteración hemos utilizado los mismos patrones negativos que la 1<sup>a</sup> iteración, y hemos duplicado los patrones positivos utilizando los mismos patrones pero tomando un recuadro más ajustado a su contorno. En estos simplemente se han extraído los colores que comentamos anteriormente sin aplicar ningún filtro y posteriormente normalizados entre 0 y 1. Los porcentajes que hemos utilizado para la entrada, test y validación son 60%, 20% y 20% respectivamente.

De esta forma después del entrenamiento obtenemos la matriz de confusión de la figura 10. El resto de la arquitectura sigue igual que la anterior iteración.

Como se observa, hay 5 falsos positivos por parte de los patrones negativos y 6 falsos negativos por parte de los patrones positivos (ver fig.10).

Tras realizar el estudio de las topologías de red (ver fig.11), decidimos emplear una arquitectura con una capa oculta con 15 neuronas ya que tiene una alta precisión y baja desviación típica. De todas maneras, éstos resultados no son del todo fiables porque la desviación típica todavía es demasiado elevada, y para unos resultados más robustos habría que realizar un estudio estadístico.

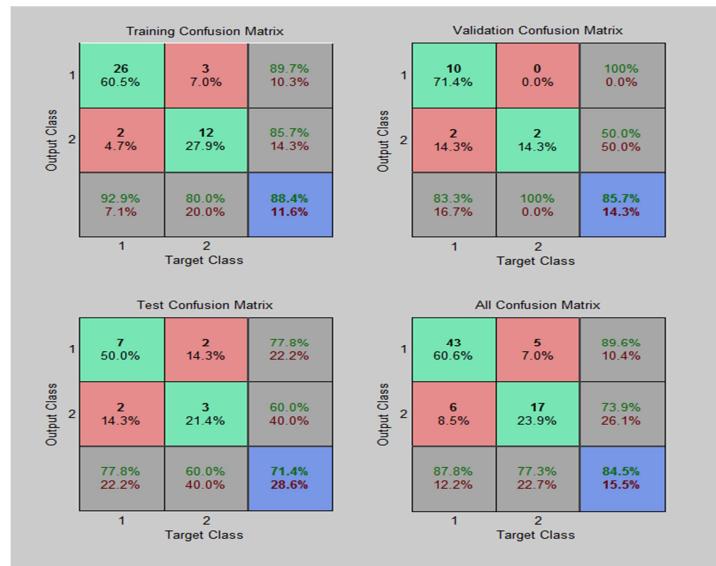


fig.10: Matriz de Confusión

Arquitectura	Entrenamiento		Validación		Test	
	Precisión	$\sigma$	Precisión	$\sigma$	Precisión	$\sigma$
[2]	76.04%	10.82	81.09%	12.02	74.00%	15.15
[3]	79.96%	9.47	86.73%	11.5	75.27%	14.23
[8]	77.63%	12.86	82.36%	14.28	74.91%	14.29
[15]	80.75%	11.28	81.64%	13.31	77.27%	13.02
[25]	78.88%	10.43	83.09%	15.03	74.73%	13.15
[2][2]	74.88%	11.30	82.00%	12.92	74.73%	13.41
[8][5]	77.67%	13.26	82.00%	12.92	73.09%	15.8
[20][5]	83.79%	9.55	82.73%	11.21	74.55%	14.92
[20][12]	82.83%	10.74	82.18%	12.45	73.64%	13.65

fig.11: Tabla de Topologías

El cerebro humano a la hora de realizar la búsqueda de Wally realiza un “escaneado” conjunto de la imagen, para posteriormente centrarse en aquellas zonas que presentan más posibilidades de incluir a Wally, que serán aquellas que tengan colores similares al mismo.

Para adaptar nuestra búsqueda a un proceso más similar al del razonamiento humano, en primera instancia analizamos la imagen focalizando rojos y azules; de las imágenes focalizadas con los rojos por un lado, y con los azules por el otro, nos quedaremos con aquella de ambas que presente menor número de objetos de dicho color. Esta adaptación nos ayuda a realizar una búsqueda mucho más eficiente ya que solo tendremos que buscar secuencialmente dentro de aquellas zonas en las que esté presente alguno de los dos colores, que coinciden por su distribución con nuestro entrañable personaje, Wally.

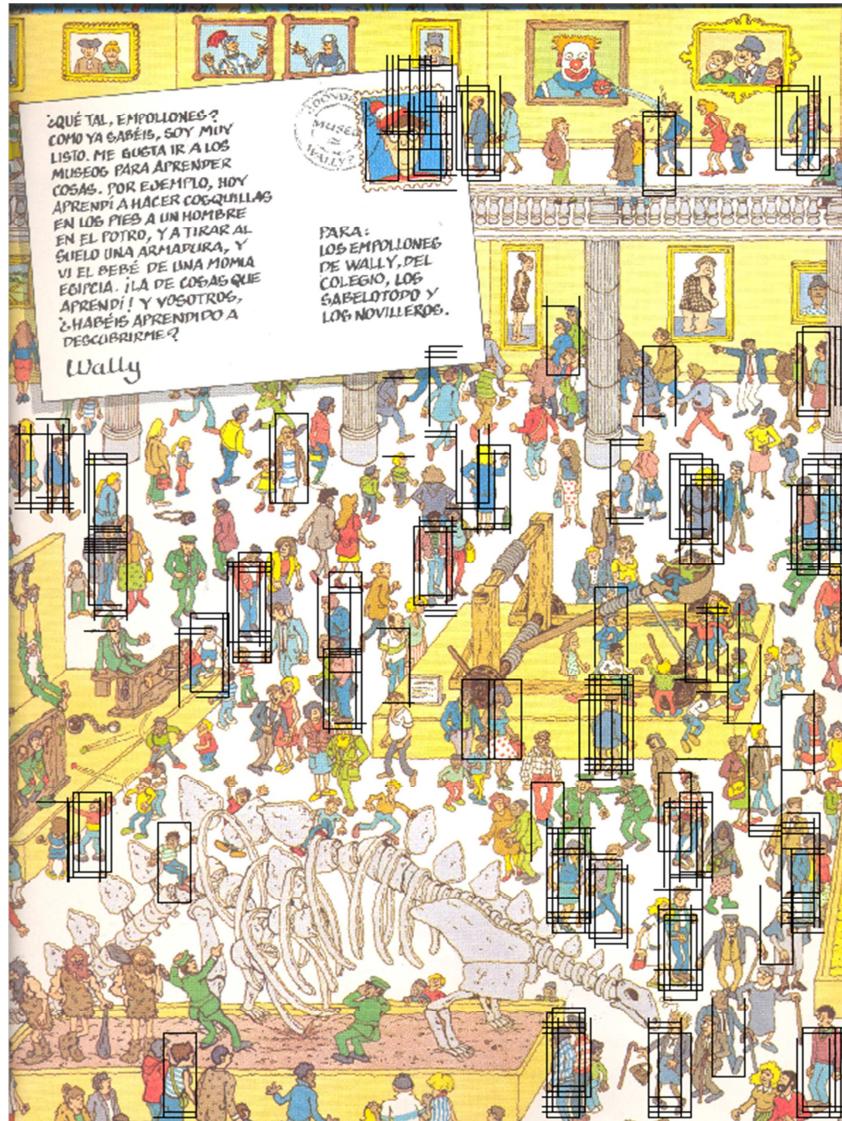


Fig. 12: Salida de la 2<sup>a</sup>aproximación

Se aprecia en la imagen resultante (fig. 12) que hay una mejoría con respecto a la iteración anterior.

### 7.3 Discusión

Las mejorías en esta segunda iteración resultan evidentes. La eficiencia del programa se ha visto enormemente incrementada, encontrando las posibles situaciones de Wally en un tiempo mucho menor que el dedicado en la primera aproximación al problema. Además la discriminación de porcentajes según posición en el recuadro buscado ha reducido y mejorado los recuadros considerados Wally's positivos.

Sin embargo también hemos advertido que la entrada de porcentaje de color que definimos para la piel está creando cierta confusión en la red, por lo tanto para una

tercera aproximación probaremos a eliminar dicha entrada. Además cuando al elegir entre la imagen focalizada en rojo y focalizada en azul, obtenemos como imagen la focalizada en azul, parece que presenta también mayor cantidad de errores; por lo tanto decidimos descartar la focalización en azul y centrarnos únicamente en los rojos de la imagen.

En la siguiente aproximación nuestro objetivo primordial es reducir el número de positivos encontrados tras la ejecución del programa sobre la imagen en la que se encuentra Wally.

## 8. Tercera aproximación

### 8.1. Descripción

En esta aproximación, se ha eliminado el color carne como entrada en el sistema debido a la neutralidad de este color en la mayoría de los escenarios, lo cual produce cierta confusión con otros colores situados en un rango cercano.

Además, también hemos hecho un estudio de los patrones que le damos a la red para entrenar mediante gráficas y gracias a esto conseguimos un refinamiento de dichos patrones mejorándolos y añadiendo nuevos patrones negativos, sobre todo en los que hay un claro color dominante para enseñarle a la red que eso NO es Wally.

El siguiente grafico (fig. 13) corresponde a como dividía antes la red entrenada las entradas..

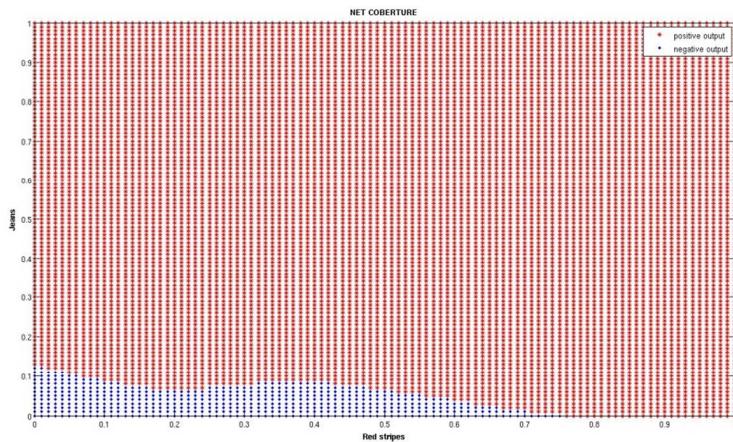


fig. 13

Y esta otra con los patrones corregidos

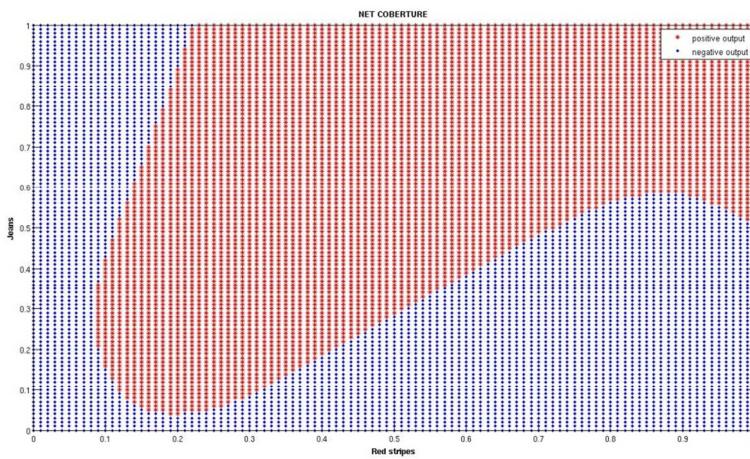


fig. 14

## 8.2 Resultados

Como ya se comentó en esta iteración se ha optimizado los patrones de entrada añadiendo nuevos y quitando otros que empeoraban el resultado. Quedando en 25 positivos y 41 negativos. En estos simplemente se han extraído los colores que comentamos anteriormente sin aplicar ningún filtro y posteriormente normalizados entre 0 y 1. Los porcentajes que hemos utilizado para la entrada, test y validación son 60%,20% y 20% respectivamente.



*fig. 15: Matriz de confusión*

Como se aprecia en la matriz de confusión el entrenamiento de la red presenta una notable mejoría, habiendo solo un patrón negativo que reconoce como positivo.

Arquitectura	Entrenamiento		Validación		Test	
	Precisión	$\sigma$	Precisión	$\sigma$	Precisión	$\sigma$
[2]	79.25%	16.82	85.39%	15.2	75.30%	19.75
[3]	79.76%	21.87	82.76%	19.0	79.57%	19.23
[8]	89.75%	14.66	90.96%	14.46	87.91%	17.79
[15]	88.25%	16.28	88.64%	15.41	85.27%	17.09
[25]	89.88%	14.63	89.69%	12.93	84.33%	15.22
[8][5]	90.57%	15.26	90.30%	13.82	87.09%	18.8
[20][5]	93.85%	13.85	93.53%	9.4	92.75%	13.92
[20][12]	93.53%	14.54	93.53%	13.75	90.31%	15.65

fig. 16: Tabla de Topologías

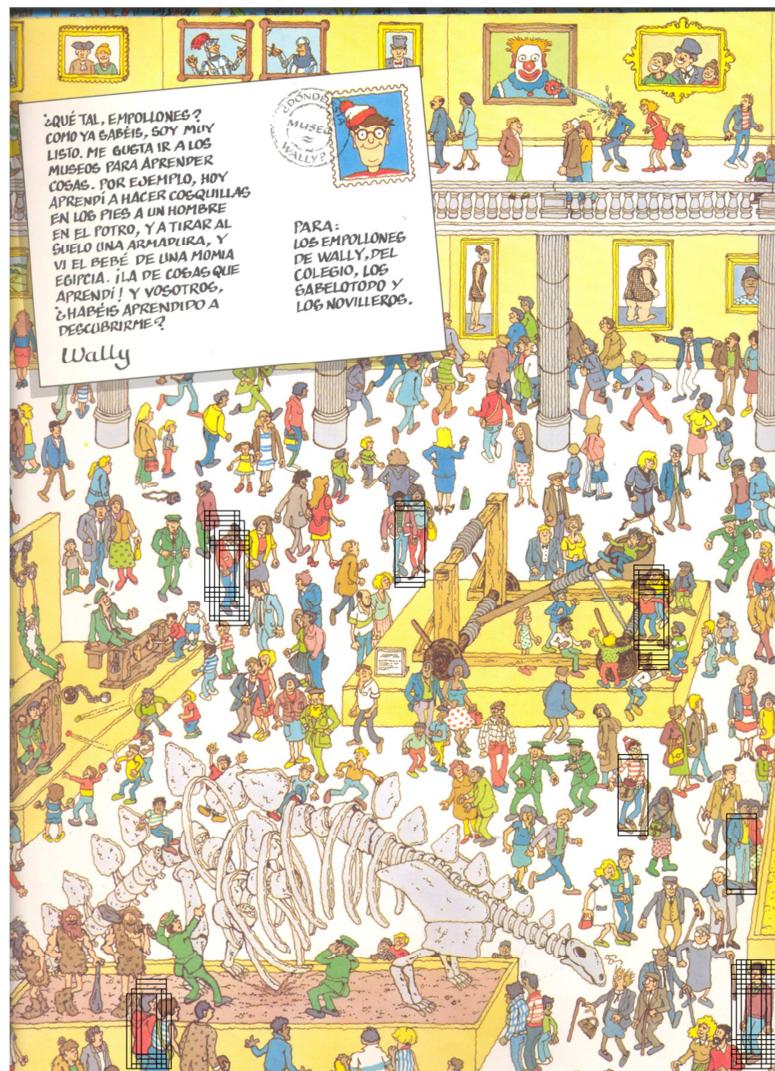


fig.17: Imagen resultante de la aprox. 3

### 8.3 Discusión

A la hora de observar los resultados, nos encontramos con que algunas imágenes que son detectadas como Wally no deberían serlo, ya que aunque presentan en algunos casos gran similitud de color con respecto a Wally, la forma sigue divergiendo del Wally original. Así, cualquier forma que presente una similitud con Wally en la distribución del color será tenida en cuenta como Wally sin importar su forma o características.

Por lo tanto, partiendo de esta idea hemos pensado en trabajar sobre la forma de la imagen, esto es, inclinación y número de objetos en escena.

## 9. Cuarta aproximación

### 9.1 Descripción

La cuarta y última aproximación trabaja sobre la forma de la imagen, observando aspectos característicos de Wally tales como el número de objetos de color rojo que forman su camiseta (cada raya debería ser tomada como un objeto por nuestro sistema). De este modo, una imagen de una persona con pantalones azules y jersey liso de color rojo, en teoría no debería ser tratado como un Wally positivo.

### 9.2 Resultados

En esta iteración utilizamos 18 patrones positivos y 22 negativos. En estos simplemente se han extraído los colores que comentamos anteriormente sin aplicar ningún filtro y posteriormente normalizados entre 0 y 1. Los porcentajes que hemos utilizado para la entrada, test y validación son 60%, 20% y 20% respectivamente

Como se puede observar en la figura 20, hemos conseguido eliminar gran cantidad de falsos positivos, llegando a un resultado de alta eficacia; a decir verdad, a la hora de comenzar a trabajar con los primeros resultados obtenidos al comienzo del trabajo, no pensamos que podríamos conseguir un resultado tan próximo a la identificación única del auténtico Wally.



fig.18: Matriz de confusión

Arquitectura	Entrenamiento		Validación		Test	
	Precisión	$\sigma$	Precisión	$\sigma$	Precisión	$\sigma$
[2]	91.83%	11.43	92.95%	10.59	90.85%	12.31
[3]	97.71%	2.80	97.71%	3.07	96.19%	4.61
[8]	95.90%	8.79	95.71%	8.005	94.76%	9.68
[15]	92.51%	16.06	94.00%	15.80	92.57%	15.75
[25]	92.08%	12.98	93.04%	9.86	90.38%	14.15
[8][5]	94.03%	13.49	96.38%	8.68	92.76%	14.50
[20][5]	97.03%	8.96	96.47%	8.26	95.71%	10.41
[20][12]	97.16%	6.27	96.85%	5.82	95.33%	7.90

fig.19: Tabla de Topologías



*fig.20: Imagen resultado de la aproximación 4*

### 9.3 Discusión

A pesar de la bondad del resultado, como en todo, quedan pequeños flecos por mejorar; el resultado más preocupante en la imagen resultante es el personaje que se encuentra en la parte inferior derecha de la misma, el cual tiene un jersey liso granate en lugar del jersey de listas rojas y blancas que identifica al Wally real. Así mismo, también encontramos otro personaje con sudadera lisa roja en la parte izquierda.

Otro resultado remarcable es el personaje encontrado en la parte inferior de la imagen que presenta un jersey casi idéntico al de Wally y pantalones de un color aproximado al azul; consideramos que este falso positivo como un mal menor difícilmente corregible.

Para rebajar los positivos identificados y considerados inaceptables se ha pensado en descartar en una capa de post-procesado las imágenes que no contengan alguna de las características básicas que identifican a Wally. Consideraremos por ejemplo, que una imagen no puede ser Wally si no presenta cierto porcentaje de color blanco en la porción de imagen tomada como “tronco”.

## 10. Conclusión

En definitiva, hemos conseguido mejores resultados de los esperados. La eficacia de la red de neuronas fue bastante alta y respondió bien ante la mayoría de las imágenes aunque a veces los resultados fuesen algo impredecibles.

Hemos mejorado la eficiencia de nuestro sistema utilizando una capa de pre procesado pero aún se podría mejorar más utilizando otra capa de post procesado, como por ejemplo eliminar de las ventanas identificadas como positivos las que no contengan un porcentaje mínimo de blanco (rayas blancas en la camiseta de Wally).

Aunque inicialmente pensábamos que podríamos afrontar el problema con Wally parcialmente oculto, vemos que realmente supondría una complejidad excesivamente alta.

Para encontrar un resultado óptimo siempre va a ser necesaria una capa de post procesado porque va a haber elementos muy similares a Wally que la red no va a ser capaz de distinguir.

## 11. Trabajo futuro

Como habíamos dicho anteriormente, en futuras aproximaciones sería conveniente añadir una capa de post procesado para eliminar los falsos positivos que la red no fuese capaz de eliminar.

Podríamos mejorar la última aproximación no teniendo en cuenta los objetos pequeños a la hora de contar las rayas del jersey.

Además del color y del número de objetos que hemos usado hasta ahora en nuestro sistema, podríamos fijarnos en que las rayas del jersey de Wally siempre están dispuestas de forma horizontal y añadir una nueva entrada a la red como la inclinación de los objetos rojos en la camisa de Wally.

## **12. Referencias**

[1]Robert Desimone, *Research explains how brain finds Waldo*. *Science Daily*, (8 de Junio del 2005)

[2]Samuel Inverso, *Specific face recognition using a backpropagation neural network*, (21 de Febrero del 2000)