**Examen final del curso**

**PUNTOS TOTALES DE 12**

**1.Pregunta 1**

**Indica cuál es la afirmación INCORRECTA sobre el color de cada píxel en una imagen.**

* El algoritmo *white-patch* persigue eliminar los efectos de cambios de color debidos a la iluminación y a los sensores en una imagen.
* Si se representa el color en sus coordenadas cromáticas estamos eliminando los efectos de cambios de intensidad de la luz que pueden darse dentro de la escena representada en la imagen.
* La distribución espectral de los sensores de la cámara puede tener efecto sobre el color de los píxeles de la imagen.
* ***No depende de la distribución espectral de la luz que llega a los sensores de la cámara.***

**2.Pregunta 2**

**Indica cuál es la afirmación INCORRECTA sobre el color de cada píxel en una imagen.**

* La convolución permite representar cada píxel de la imagen con un vector de caraterísticas locales relacionadas con la forma de un conjunto de filtros.
* Las imágenes RGBD permiten construir descriptores que consideran la información de profundidad de la escena que representa la imagen.
* ***La tripleta RGB de los colores del píxel son un descriptor básico de la imagen que no se ve afectado por cambios de intensidad o color de la luz de la escena.***
* En el método de *Template matching* se utiliza directamente una ventana de la imagen como descriptor del objeto que se quiere detectar.

**3.Pregunta 3**

**Sea***P*(*Y*=1|*X*=*x*;*w*)=(*Logistic*(*wTx*))**y supongamos***Y*∈{0,1}**y***X*∈R*n*+1**. Es decir, nos situamos en el contexto de la clasificación de ventanas basada en un modelo obtenido por regresión logística.**

**Indica la afirmación incorrecta:**

* En el contexto de las redes neuronales, *P*(*Y*=1|*X*=*x*;*w*) se puede tomar como la formalización de una neurona.
* Como descriptor de ventanas (x)(*x*) podemos usar Histogramas de Bloques LBP (HB-LBP), HOG, incluso HOG & HB-LBP simultáneamente.
* La condición *P*(*Y*=1∣*X*=*x*) <*P*(*Y*=0∣*X*=*x*) es equivalente a la condición *wTx*<0.
* ***Como la función de coste J(w)J(w) que optimizamos para obtener el valor de ww es convexa, cualquier conjunto de entrenamiento S={(x′1,y1),...,(x′M,yM)} nos va a servir para obtener un modelo (w) que nos de buenos resultados.***

**4.Pregunta 4**

**Pensando en el descriptor conocido como LBP para clasificar ventanas de una imagen, indica la afirmación incorrecta:**

* ***Los patrones uniformes que hemos visto en el curso, son independientes del tamaño y tipo de la vecindad utilizada en el cálculo del LPB básico.***
* Describir una ventana con un único histograma de códigos LBP es menos preciso que si utilizamos bloques de ventana sobre los que calcular los histogramas (para su posterior concatenación formando el descriptor de toda la ventana).
* Describir una ventana con un histograma de códigos LBP en lugar de con una concatenación (a modo de vector) de los códigos LBP de cada píxel, nos da robustez ante pequeños “movimientos” del contenido de la ventana y, normalmente, resulta en un descriptor de menos dimensiones.
* Los códigos LBP pueden definirse utilizando vecindarios más grandes de 3×3 píxeles.

**5.Pregunta 5**

**Indica la afirmación incorrecta:**

* Para no calcular de forma repetida los histogramas de códigos LBP de los distintos bloques de la imagen, necesitamos seleccionar cuidadosamente los pasos del desplazamiento de la ventana, es decir, Δ*x*&Δ*y*.
* Con el método de la pirámide, si utilizamos más niveles, probablemente el módulo de refinación deberá tratar con más pre-detecciones.
* ***Durante la ejecución del mecanismo de bootstrapping no se utiliza la pirámide con ventana deslizante.***
* Con el método de ventana deslizante, si el desplazamiento de la ventana (Δ*x*&Δ*y*) es más pequeño probablemente el módulo de refinación deberá tratar con más pre-detecciones.

**6.Pregunta 6**

**En el contexto del mecanismo de “pirámide con ventana deslizante” y suponiendo los parámetros siguientes:**

**Tamaño de imagen → 640 ×480 píxeles,  
Tamaño ventana canónica → 64 ×128 píxeles, y  
(Δ*x*,Δ*y*,*s* ) = (8, 8, 1.2)**

**Indica la respuesta correcta:**

* El número de ventanas a clasificar está en el rango de 15.000 a 30.000.
* El número de ventanas a clasificar está en el rango de 30.000 a 45.000.
* El número de ventanas a clasificar no llega a 5.000.
* ***El número de ventanas a clasificar está en el rango de 5.000 a 15.000.***

**7.Pregunta 7**

**Si tenemos una imagen de 64x128 píxeles y utilizamos la siguiente configuración para calcular el descriptor HOG:**

**- Tamaño de celda: 16x16 píxeles  
- Orientación del gradiente con signo  
- Rango de cada intervalo del histograma de orientaciones: 40º  
- Interpolación de orientación y espacial  
- Bloques de 2x2 con un paso de solapamiento igual a 1**

**Con esta configuración, el número final de dimensiones del descriptor será:**

* ***756***
* 288
* 1512
* 576

**7.Pregunta 7**

**Indica cuál de las siguientes afirmaciones es CIERTA respecto al cálculo del descriptor HOG con interpolación de orientación y espacial:**

* En un bloque se acumula la suma normalizada de los histogramas de todas las celdas del bloque.
* La contribución de un píxel al histograma de una celda es inversamente proporcional a la distancia del píxel al borde de la celda.
* Todos los píxeles de una celda tienen una contribución mayor que cero en todos los intervalos del histograma.
* Todos los píxeles de una celda tienen el mismo peso relativo en el cómputo del histograma de la celda.
* ***En el cómputo del histograma de una celda hay que tener en cuenta hay que tener en cuenta todos los píxeles de la celda, y también algunos de los píxeles de las 8 celdas vecinas.***

**8.Pregunta 8**

**Indica cuál de las siguientes afirmaciones es CIERTA respecto a un clasificador SVM:**

* Cualquier cambio en las muestras de entrenamiento provoca también un cambio en los parámetros del hiperplano solución.
* ***Los vectores de soporte son el subconjunto de muestras que determinan el hiperplano solución y quedan determinados por aquellos multiplicadores de Lagrange con valor estrictamente mayor que cero.***
* La condición de margen máximo implica la definición de un hiperplano con la distancia más grande a todas las muestras de entrenamiento.
* La función de maximización del margen depende de los parámetros *w* y *b* del hiperplano de separación.
* La condición de margen máximo implica que todas las muestras positivas deben cumplir la condición wTxi ​ ≥ b, donde *w* y *b* son los parámetros que definen el hiperplano de separación.

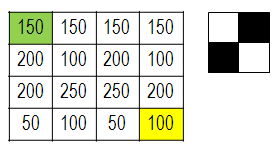
**8.Pregunta 8**

**Indica cuál de las siguientes afirmaciones es CIERTA respecto a un clasificador SVM:**

* Las funciones de kernel nos definen una transformación explícita de las muestras en un nuevo espacio de características donde es posible encontrar una separación lineal de los datos.
* Las variables de holgura relajan la condición de margen máximo, permitiendo que algunas muestras violen la condición del margen y se unan a los vectores de soporte como muestras que se utilizan para determinar los parámetros del hiperplano solución.
* Valores muy altos del factor de regularización C que se introduce con las variables de holgura permiten una mayor tolerancia a errores.
* Las variables de holgura introducen un factor de regularización en la formulación del problema, cuyo valor óptimo se obtiene como resultado del problema de optimización junto con los parámetros w y b del hiperplano.
* ***Utilizando el truco del kernel, la formulación del problema de optimización es la misma, excepto que se sustituye el cálculo del producto escalar entre las muestras por la aplicación de la función de kernel.***

**9.Pregunta 9**

**La tabla siguiente representa una imagen de tamaño 4\times×4. Cada valor de la tabla corresponde al valor de intensidad de cada uno de los píxeles de la imagen. También mostramos un filtro de Haar básico que es el que utilizaremos para los cálculos de esta pregunta.**

****

**Indica cuál de las siguientes afirmaciones es CIERTA:**

* El valor de la imagen integral en el extremo inferior derecho de la imagen (el píxel marcado en amarillo) es igual a 2500.
* El número de características de Haar que se obtienen aplicando el filtro de Haar a todas las escalas posibles (tanto en horizontal como en vertical) sólo en la primera posición de la imagen (el píxel marcado en verde), sin traslación es igual a 9.
* ***El valor del filtro de Haar escalado de forma que ocupe toda la imagen (4×4 píxeles) es igual a 0.***
* El número de características de Haar que se obtienen aplicando el filtro de Haar a la escala básica (2\times×2 píxeles) en todas las posibles posiciones de la imagen es igual a 16.

**9.Pregunta 9**

**Indica cuál de las siguientes afirmaciones sobre las características de Haar y la utilización de la imagen integral es CIERTA.**

* En los filtros extendidos de Haar sólo se incluyen las versiones rotadas de los filtros básicos de Haar.
* ***Utilizando la imagen integral el nº de operaciones necesarias para calcular un filtro de Haar es independiente de la escala y de la posición de la imagen dónde se evalúe.***
* Al utilizar el método de ventana deslizante para detectar objetos con las características de Haar, la ventana de detección se normaliza siempre al tamaño estándar de 24x24 píxeles.
* Utilizando la imagen integral el nº de operaciones necesarias para calcular todos los filtros de Haar es el mismo.
* Al utilizar imágenes integrales para el cálculo del descriptor HOG no se puede aplicar la interpolación de orientación.

**10.Pregunta 10**

**Indica cuál de las siguientes afirmaciones sobre Adaboost y la cascada de clasificadores es FALSA:**

* En cada paso de una cascada de clasificadores, el objetivo es rechazar el máximo número posible de ventanas negativas, aceptando a su vez todas (o la mayoría) de las ventanas positivas.
* Para ajustar el clasificador al objetivo fijado de falsos positivos y negativos para cada nivel de la cascada hay que modificar el umbral de decisión del clasificador “fuerte” introduciendo un margen en la decisión.
* Los clasificadores débiles se aprenden de la misma forma independientemente de si se está usando una cascada de clasificadores o un único clasificador global.
* ***Cada nivel de una cascada de clasificadores se puede entrenar de forma independiente al resto de niveles de la cascada.***
* Si el número de falsos positivos en un nivel de la cascada es demasiado elevado habrá que fijar un valor de margen negativo para reducir el número de muestra que quedan clasificadas como positivas.

**10.Pregunta 10**

**Indica cuál de las siguientes afirmaciones sobre Adaboost es FALSA:**

* Para entrenar un clasificador débil en una iteración del proceso de aprendizaje se analizan los valores de todas las características de Haar para todas las muestras de entrenamiento.
* Un clasificador débil queda determinado por una determinada característica de Haar (tipo de filtro, escala y posición de la imagen), un umbral de decisión y un signo de decisión.
* Los clasificadores débiles con menor error de clasificación tienen mayor peso relativo en el clasificador global final.
* ***En cada iteración, el peso relativo de todas las muestras se modifica por igual en función del error de clasificación del último clasificador débil aprendido.***
* El error de clasificación a minimizar al entrenar un clasificador débil tiene en cuenta la ponderación de las muestras de entrenamiento en esa iteración.

**11.Pregunta 11**

**Indica la afirmación incorrecta.**

* ***En las arquitecturas tipo CNN, las etapas de max-pooling añaden más parámetros a aprender.***
* En el método de generación de candidatos conocido como “selective search”, con una sola ejecución del algoritmo tendremos simultáneamente los candidatos para detectar distintos tipos de objetos y a diferentes escalas.
* Supongamos un par estereoscópico donde una de las cámaras del par se utiliza para detectar objetos en base a su apariencia visual. Entonces podemos usar la información de profundidad del par estereoscópico o bien para reducir el número de ventanas a examinar (candidatos) durante la detección de objetos, o bien para eliminar falsos positivos previamente a la refinación.
* La idea más novedosa del uso de CNNs respecto a otros enfoques como HOG/SVM o Haar/AdaBoost es que los propios descriptores de ventanas también se aprenden.

**12.Pregunta 12**

**Indica la afirmación incorrecta.**

* Para entrenar un DPM, las muestras positivas de entrenamiento (p.e. peatones) solo necesitan venir anotadas con el marco global (*bounding box*), no se necesita la anotación de las partes constituyentes.
* El objetivo de la adaptación de dominio (*domain adaptation*) es reaprovechar trabajo previo, ya sea en forma de datos anotados o de modelos aprendidos.
* ***En el método de Random Forest de Expertos Locales visto en el curso, los expertos locales siguen la misma filosofía que las “reglas simples” de AdaBoost (p.e., umbral en el resultado de un filtro de Haar). Es decir, los expertos locales y esas “reglas simples” pretenden ser clasificadores débiles.***
* El aprendizaje activo (*active learning*) se puede utilizar para obtener muestras difíciles del dominio de destino (*target*), con el soporte de un clasificador desarrollado en el dominio fuente (*source*).