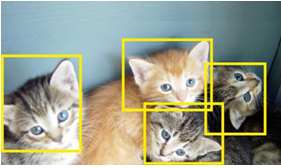
**Cuestionario 6**

**PUNTOS TOTALES DE 10**

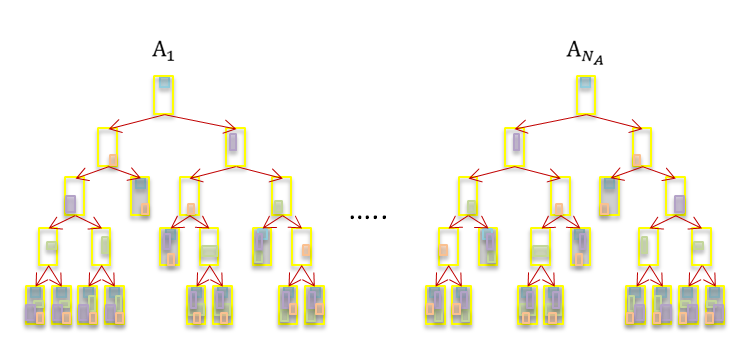
**1.Pregunta 1**

****

**Según las detecciones que se ven en la imagen, indica la afirmación incorrecta**

* Es más probable que se haya utilizado 1 DPM que 1 modelo de partes no deformables (posición fija).
* Es más probable que se haya utilizado 1 DPM que 1 modelo holístico basado en HOG/SVM
* Es más probable que se haya utilizado 1 DPM que 1 modelo holístico basado en Haar/AdaBoost.
* ***La cabeza es un objeto rígido, no tiene sentido usar un DPM.***

**2.Pregunta 2**

****

**Sea el modelo de la figura basado en un *Random Forest* de Expertos Locales, indica la afirmación incorrecta.**

* Cada región seleccionada tiene un clasificador asociado. Puede haber solapamiento entre regiones.
* Dado uno de estos árboles, las regiones concretas (sub-ventanas) que han intervenido en el camino desde la raíz a una hoja es aleatorio.
* El clasificador asociado a cada región no sigue la misma filosofía de los clasificadores débiles usados en AdaBoost, es decir, pretenden ser lo más discriminativos posible.
* ***El número de árboles y su profundidad se aprende (sin necesidad de validación cruzada o hold-out).***

**3.Pregunta 3**

**En el contexto de la adaptación de dominio (*domain adaptation*), indica la afirmación incorrecta.**

* Al cambiar de sensor (cámara) podría ser necesario realizar adaptación de dominio.
* El objetivo es reaprovechar trabajo previo, ya sea en forma de datos anotados o de modelos aprendidos.
* ***En realidad “adaptación de dominio” solo es otro nombre para la utilización conjunta del bootstrapping y el aprendizaje activo.***
* 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*).

**4.Pregunta 4**

**En el contexto de las redes neuronales artificiales, indica la afirmación incorrecta.**

* ***La topología de una red (número de capas y cómo están conectadas) se aprende (sin necesidad de validación cruzada o hold-out).***
* La función de entrada de una neurona consiste en el producto escalar del vector de entradas y el vector de pesos de las entradas.
* El objetivo es aprender representaciones completas, es decir, los descriptores de los datos y los clasificadores basados en ellos.
* La función logística es una posible función de activación de la neurona.

**5.Pregunta 5**

**En el contexto de las CNNs, indica la afirmación correcta.**

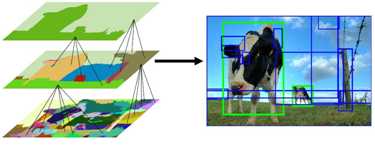
* ***Teóricamente, la primera capa convolucional de una CNN podría aprender un banco de filtros de Haar.***
* Teóricamente, la primera capa convolucional de una CNN podría aprender el descriptor HOG.
* Teóricamente, la primera capa convolucional de una CNN podría aprender el descriptor basado en histogramas de bloques LBP.
* Si en la primera capa de una CNN usamos un banco de K filtros, filtros de M píxeles, y hemos de aprender a partir de ventanas de N píxeles, el número total de parámetros a aprender es N\*K\*M.

**6.Pregunta 6**

**En el contexto del uso de datos multimodales, indica la afirmación incorrecta.**

* ***Si ya tenemos un LIDAR de cuatro planos, una cámara de infrarrojo lejano no aportará información adicional.***
* A partir de una cámara color podremos tener en cuenta RGB y flujo óptico.
* A partir de un par estereoscópico color podremos tener en cuenta RGB, flujo óptico y profundidad.
* A partir de un par estereoscópico color y una cámara de infrarrojo lejano, podremos tener en cuenta RGB, flujo óptico, profundidad y temperatura relativa.

**7.Pregunta 7**

****

**En el contexto de la generación de candidatos y dada la imagen de arriba, indica la afirmación incorrecta.**

* Corresponde a la estrategia de generación de candidatos conocida como “*selective search*”.
* La generación de candidatos en sí misma también requiere aprendizaje.
* Idealmente, con una sola ejecución del algoritmo tendremos simultáneamente los candidatos para detectar distintos tipos de objetos, y a diferentes escalas.
* ***Para este tipo de ventanas candidatas será más conveniente el uso de modelos holísticos que de DPMs.***

**8.Pregunta 8**

**Supongamos que trabajamos con un par estereoscópico, indica la afirmación correcta.**

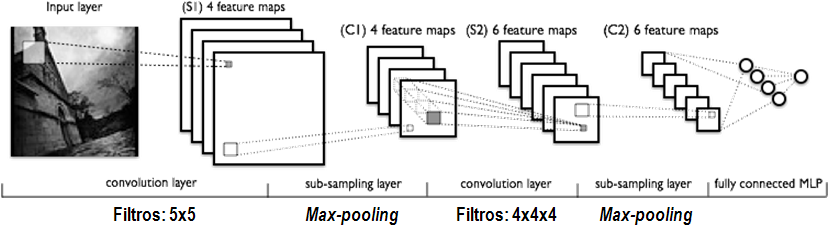
* Asumiendo que usamos el mecanismo de ventana deslizante con pirámide y HOG/SVM para procesar las imágenes provenientes de una de las cámaras, utilizar la información de profundidad para rechazar detecciones ya no tiene sentido.
* Asumiendo que usamos el mecanismo de ventana deslizante y Haar/AdaBoost (escalando los filtros) para procesar las imágenes provenientes de una de las cámaras, utilizar la información de profundidad para rechazar detecciones ya no tiene sentido.
* Asumiendo que usamos el DPM para procesar las imágenes provenientes de una de las cámaras, utilizar la información de profundidad para rechazar detecciones ya no tiene sentido.
* ***Todas las anteriores son incorrectas.***

**9.Pregunta 9**

**Aunque no hemos hablado de tiempos de cálculo explícitamente, según lo que hemos visto de los métodos de generación de candidatos, indica la afirmación correcta.**

* El tiempo de generación de ventanas (candidatos) del método de “*selective search*” es independiente del contenido de la imagen que se procesa (guía tu intuición por la figura de la pregunta anterior).
* En detección de peatones desde un coche, el tiempo de generación de ventanas (candidatos) del método basado en un par estéreo (horizonte + *occupancy grid*) es independiente del contenido de la imagen que se procesa.
* ***El tiempo de generación de ventanas (candidatos) del método ventana deslizante con pirámide es independiente del contenido de la imagen que se procesa.***
* Todas las afirmaciones anteriores son falsas.

**10.Pregunta 10**

****

**Sea la CNN de la figura, con los hiperparámetros indicados (bancos de filtros, tamaños de filtro y *max-pooling*). Supongamos que nos fijamos solo en la parte de los bloques convolunión+sub-sampling, no en la parte “fully connected MLP”. Indica la afirmación correcta respecto al número de parámetros a aprender.**

* Son 89
* ***Son 484***
* No se puede saber sin conocer el tamaño de las celdas sobre las que se aplica *max-pooling*.
* No se puede saber sin conocer el número de píxeles de la imagen de entrada.