



## Control de lectura: Detección de Localización de

1 mensaje

Formularios de Google <forms-receipts-noreply@google.com>  
Para: ysantosa@unsa.edu.pe

vie., 16 de julio de 2021 a la hora 8:29 a. m.

### Google Forms

Gracias por rellenar [Control de lectura: Detección de Localización de](#)

Esto es lo que se recibió.

## Control de lectura: Detección de Localización de

Se ha registrado tu correo ([ysantosa@unsa.edu.pe](mailto:ysantosa@unsa.edu.pe)) al enviar este formulario.

Sobre la RCNN, marque la o las respuestas falsas.



El segundo módulo utiliza 5 capas convolucionales y 2 full conectados para extraer un vector característica de tamaño 4096 único entre todas las regiones propuestas.



En la etapa de test, selecciona 2000 regiones propuestas



Las regiones propuestas son deformadas antes de ingresar a la full conectada de la CNN para que las imágenes tengan el mismo tamaño



si tiene una intersección sobre unión (IoU) solapada con una región seleccionada de mayor puntuación mayor que un umbral aprendido

En la fast RCNN es verdad



La última capa de convolución de la CNN es reemplazada por un max-pooling similar a la SPP-Net



La última capa max pooling de la CNN es reemplazada por un IoU pooling layer

☒ La última capa de max pooling de la CNN es reemplazada por un ROI pooling layer

☐ La última capa de convolución de la CNN es reemplazada por ROI pooling layer

Se tiene una imagen de  $n \times n$ . Se toman  $15 \times 15$  celdas en la imagen. Por otro lado, el número de anchor por celda es de 9 y se tienen en total 20 clases. Luego, podemos afirmar que el vector "y", que representa el ground truth deberá tener una dimensión igual a ?

☐ 7650

☒ 14625

☐ 30375

☐ 50625

Sobre YoloV2 cual de las siguientes afirmaciones es verdadera

☐ Es pre-entrenado en imágenes de alta resolución

☐ Define mejores anchor boxes mediante k-means

☐ Utiliza una técnica de entrenamiento por múltiples escalas

☒ Todas son correctas

¿Cuál o cuáles de las siguientes afirmaciones es una característica de los two-state detector?

☒ Generan ROIs en la primera etapa y, en una segunda etapa, envían esos ROIs al clarificador y al Bounding Box Regressor

☒ Tiene, en general, un mayor accuracy que los enfoques de tipo one-state

☐ La generación de los ROIs se hace con selective search

☐ Fue el primer enfoque para la detección y localización de objetos basados en Deep Learning

De las siguientes afirmaciones, no es verdad.

☐ Los métodos clásicos de regiones propuestas como EdgeBoxes o Selective Search consumen

mucho tiempo,

☐ La faster-RCNN Introduce el concepto de ancor boxes

☐ La faster-RCNN tiene dos módulos. Un Fast-RCNN y un RPN. El RPN propone regiones y le dice a la Fast-RCNN donde prestar atención en la imagen para localizar un objeto, colocar el bounding box y clasificar

☒ Si se sabe que se tiene 20 clases y 9 anchors, entonces la Reg Layer, de la RPN en la Fast-RCNN, es un vector de dimensión 720

☐ Un mapa de características se ha dividido en celdas de  $W \times H$  y se tienen hasta 4 anchors, entonces el número total de anchor generados es  $W \times H \times 4$

### Sobre YOLO, marque la respuesta o las respuestas incorrectas

☐ Permite un entrenamiento end-to-end

☒ Cada bounding box consiste en 4 valores, los cuales son  $(x, y)$  que representa el centro del bounding box,  $w$  que es ancho de la caja y  $h$  que es la altura

☒ Cada celda es la encargada de predecir  $P(C_{\{i\}} | Obj)$ , donde  $C_{\{i\}} \in C$  y  $C$  es el conjunto de clases a predecir

☐ Cada celda predice un número constante de Bounding box y un valor que representa el nivel de confianza para ese bounding box

### No es o no son una limitación de YOLO

☐ Tiene un límite en cuanto al número de objetos cercanos entre sí que puede predecir.

☒ Problemas para detectar objetos grandes que aparecen en grupo

☐ Problemas para generalizar objetos con aspectos inusuales

☐ Cada bounding box sólo puede predecir una clase

☐ La función loss trata los errores producidos por los bounding box grandes de la misma manera que a los bounding box pequeños