

1. What do you think it occurred during this model development (training & evaluation)?
2. How would you fix this behavior? Please provide at least 2 options explaining their pros and drawbacks
3. Extra: Do you know any tracking algos (Deep learning based) that could be used here?

## Respuestas Q1

1. No se han contemplado casos en que el color de la pelota coincida con el color de la cabeza de alguna persona en el campo, ya sea jugador o arbitro. En los datasets de training y evaluation no había imágenes similares para entrenar la red. Solo se ha tenido en cuenta la forma y el color, no se ha valorado la información de los píxeles vecinos.
2. Opción 1: Contemplar los píxeles vecinos para saber si el objeto es una pelota o un jugador y obtener información de los mismos para poder diferenciar la pelota de la cabeza.

**PRO:** Tienes información extra de la imagen para diferenciar entre jugador y pelota. No solo tienes en cuenta el color, sino otros parámetros.

**Drawbacks:** Baja resolución en la imagen que provoca que los bordes de la pelota y de la cabeza del arbitro no sean claros y la red pueda confundirse. Mucha información a tener en cuenta, tiempo de ejecución mayor.

Opción 2: Entrenar la red segmentando, en diferentes partes, toda la imagen para detectar la pelota pudiendo delimitar los bordes del objeto y diferenciarlo del resto de jugadores del campo, teniendo en cuenta que tienes información de toda la imagen (píxeles de alrededor).

**PRO:** Reduces tiempo de ejecución ya que reduces el tamaño de imagen para la segmentación, puedes encontrar la pelota antes y diferencias la pelota teniendo en cuenta información de la forma, color y el resto de píxeles de la imagen.

**Drawbacks:** Debes realizar varias segmentaciones. Dependiendo de la resolución de la imagen y del tamaño de los objetos puede dar lugar a errores.

Opción 3: Dibujar patrones en la pelota visibles para la IA para que sea capaz de diferenciarla del resto de objetos y personas que se encuentren en la imagen.

**PRO:** entrenas la red para que sea capaz de detectar ese patrón.

**Drawbacks:** Mucho tiempo de dedicación para reentrenar red y detectar dicho patrón, crear dataset de entrenamiento y validación...

3. **Poinrend** – utilizado para segmentación semántica y por instancias  
Modelo de DL que realiza la segmentación de imágenes utilizando representaciones puntuales de los objetos que aparecen en la imagen.

Consta de 3 módulos:

- Una estrategia de selección de puntos elige un pequeño número de puntos de valor real sobre los que realizar las predicciones, evitando un cálculo excesivo para todos los píxeles de la cuadrícula de salida de alta resolución.
- Para cada punto seleccionado, se extrae una representación de características a nivel de punto. Las características de un punto de valor real se calculan mediante la interpolación bilineal de  $f$ , donde  $f$  es un mapa de características con 3 canales (CxHxW), utilizando los 4 vecinos más cercanos

del punto que se encuentran en la cuadrícula regular de  $f$ . Como resultado, es capaz de utilizar la información de los subpíxeles codificada en la dimensión del canal de  $f$  para predecir una segmentación que tiene mayor resolución que  $f$ .

- Una cabeza de punto: una pequeña red neuronal entrenada para predecir una etiqueta a partir de esta representación de características a nivel de punto, de forma independiente para cada punto.