

Introducción

El campo de la visión por computador está evolucionando muy rápidamente en los últimos años, permitiendo que los desarrolladores la utilicen para tareas más sofisticadas. Uno de los campos dentro de la visión por computador es el *human pose estimation*, que nos permite localizar y predecir donde se sitúan los puntos de unión del cuerpo de un individuo mediante imágenes. Esto permite aplicar esta tecnología en ámbitos como la animación, el gaming, la salud o el deporte.

Las tareas que realiza una librería de pose estimation se dividen en dos, identificar las personas en la imagen y por cada detección identificar los keypoints correspondientes. Según en el orden que realicen estas tareas podemos diferenciar las librerías que tienen un enfoque top-down o bottom-up. Top-down detecta en primer lugar las personas en la imagen y a continuación predice donde se encuentran los puntos de unión del cuerpo, un ejemplo de librerías que utilizan este enfoque son AlphaPose y Mask RCNN. En cambio, bottom-up localiza en primer lugar estos puntos de unión del cuerpo y a continuación los agrupa creando así todos los puntos de unión del individuo, algunas de las librerías que utilizan este enfoque son OpenPose y DeepCut.

Como podemos ver existen varios enfoques y esto es debido a que en los entornos reales se encuentran con múltiples problemas que necesitan solucionar dependiendo la aplicación que vayan a realizar. Alguno de los problemas más comunes son múltiples personas en una misma imagen, partes del cuerpo de la persona oculta o simplemente mejorar el rendimiento para poder utilizar estas librerías en tiempo real.

En este trabajo nos centraremos en analizar varias librerías de pose estimation, para encontrar la solución que mejor se adapta al ámbito del deporte, con el objetivo final de aplicarlo en un entorno real y crear un dataset público.

Objetivos

El deporte es uno de los ámbitos en los que hoy en día ya se están utilizando técnicas de pose estimation, algunas de sus aplicaciones actuales son obtener estadísticas en tiempo real de los jugadores en un partido de fútbol, ayudar en el entrenamiento de los deportistas o recrear jugadas en tres dimensiones. Debido a las múltiples aplicaciones y los diferentes contextos en los que se realiza el deporte, nuestro principal objetivo será encontrar las librerías de pose estimation que mejor se adaptan al ámbito deportivo.

Con el fin de comprobar que las librerías que mejor se adaptan según sus características al deporte, se realizará una práctica en un entorno real con el objetivo de crear una base de datos y detectar anomalías en la actividad realizada.

Metodología

Planning

Estado del arte

2014 - DeepPose: Human Pose Estimation via Deep Neural Networks [1] (No necesita de modelos gráficos para la predicción del posicionamiento de los puntos de unión)

2015 - Efficient Object Localization Using Convolutional Networks [2]

2016 - Convolutional Pose Machines [3] (Introducción de las CPMs convolutional pose machines en el pose estimation)

2016 - DeepCut: Joint Subset Partition and Labeling for Multi-Person Pose Estimation [4] (Primeras detecciones multi persona)

2018 - Simple Baselines for Human Pose Estimation and Tracking [5]

2018 - DensePose: Dense Human Pose Estimation In The Wild [6] (Mapear todo el humano para obtener de imagen la superficie 3d del cuerpo de la persona)

2019 - OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields [7] (Primeras aplicaciones en tiempo real)

2019 - Human Pose Estimation for Real-World Crowded Scenarios [8] (Solventar el problema de la detección y pose estimation en imágenes con gran densidad de población)

Fuentes de información

[1] - <https://arxiv.org/abs/1312.4659>

[2] - <https://arxiv.org/abs/1411.4280>

[3] - <https://arxiv.org/abs/1602.00134>

[4] - <https://arxiv.org/abs/1511.06645>

[5] - <https://arxiv.org/abs/1804.06208>

[6] - <https://arxiv.org/abs/1802.00434>

[7] - <https://arxiv.org/abs/1812.08008>

[8] - <https://arxiv.org/abs/1907.06922>