

Estimación de Poses en jugadores de Volley

Aplicaciones en Visión Computacional



Problemática





Problemática

Monitorear las posiciones de jugadores de volleyball de forma automática.

Garantizando que la técnica sea correcta en la recepción se contribuye a un mejor rendimiento personal y del equipo y permite evitar lesiones innecesarias por fallos técnicos





Dataset

138 videos correspondientes al movimiento de recepción de jugadores de volleyball.

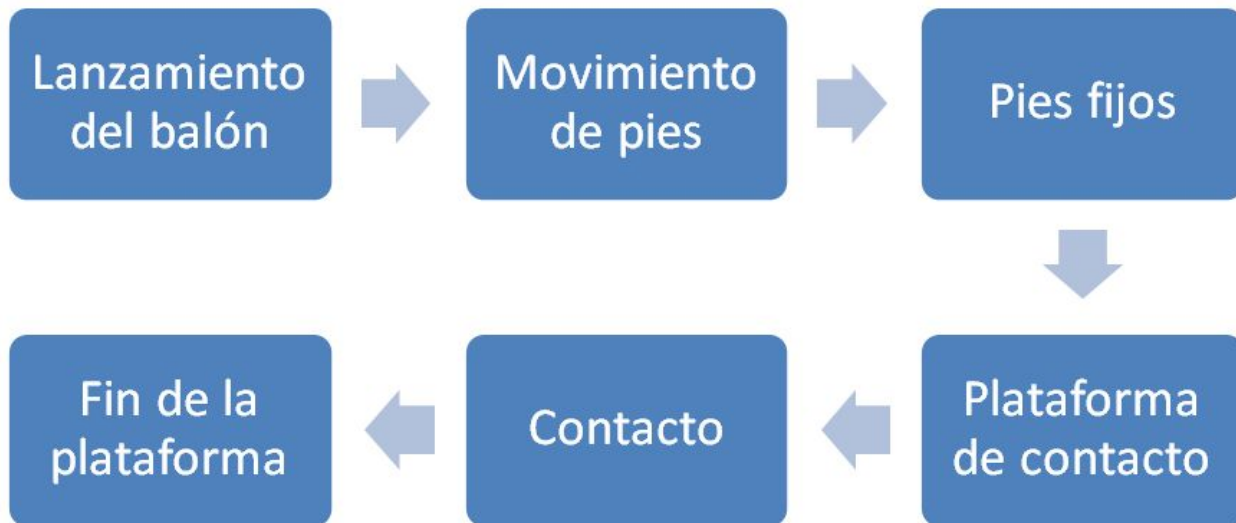
Grabados a 60 fps.

De duración entre 1 a 5 segundos.





Fases de movimientos



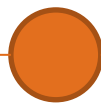


Este proyecto ahonda en el tema de estimación de pose para obtener información sobre la configuración del cuerpo humano, ya sea a través de una representación 2D o 3D de articulaciones o una representación 3D del cuerpo humano en base a una imagen o video.

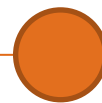
Con esta información, se revisó el estado del arte de esta área de investigación y se realizó la aplicación de un grupo de técnicas recientes que retornan resultados destacables. Por último, estos valores obtenidos a través de la aplicación de Deep Learning fueron interpretados para generar una estimación de cómo varían la posición de las articulaciones a través de la totalidad de los videos.



Revisión de Literatura



Pruebas con detección 2D

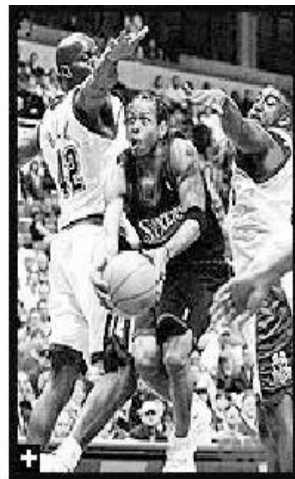


Pruebas con detección 3D

Antecedentes



**3D pose
estimation using
orthographic
camera
assumption
(1985,2000)**



(a)



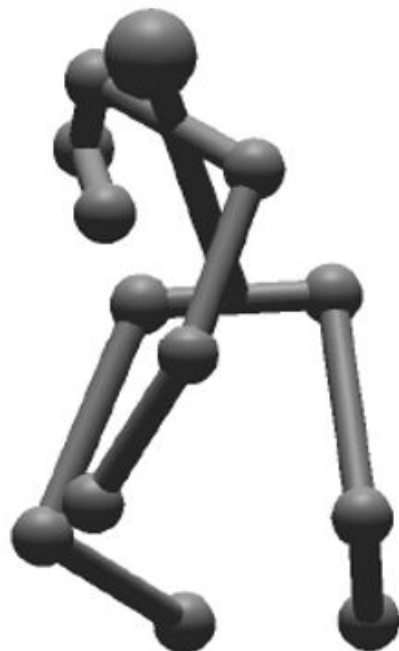
(b)



(c)

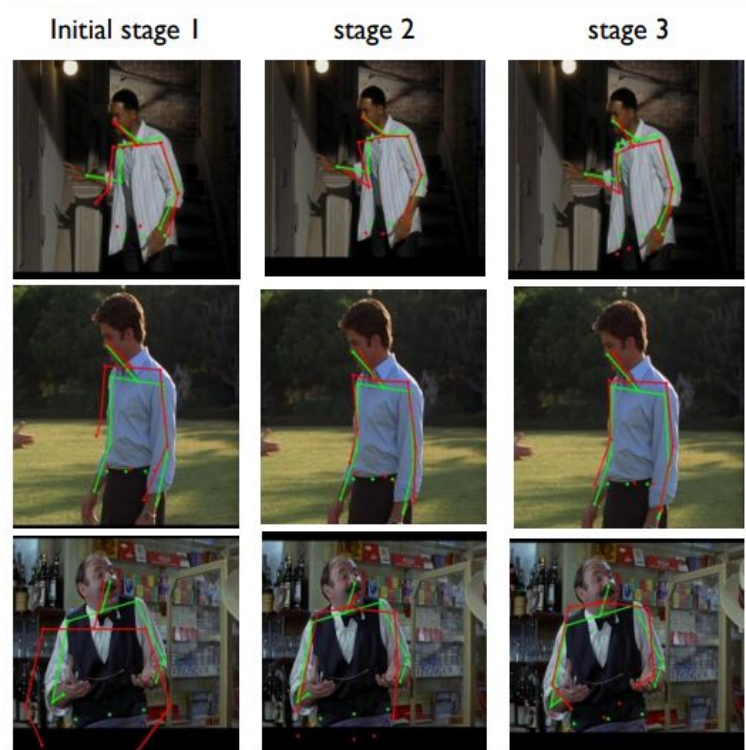


(d)



C. J. Taylor, "Reconstruction of Articulated Objects from Point Correspondences in a Single Uncalibrated Image", Computer Vision and Image Understanding, Vol: 80, No: 10, Pgs: 349-363, October 2000

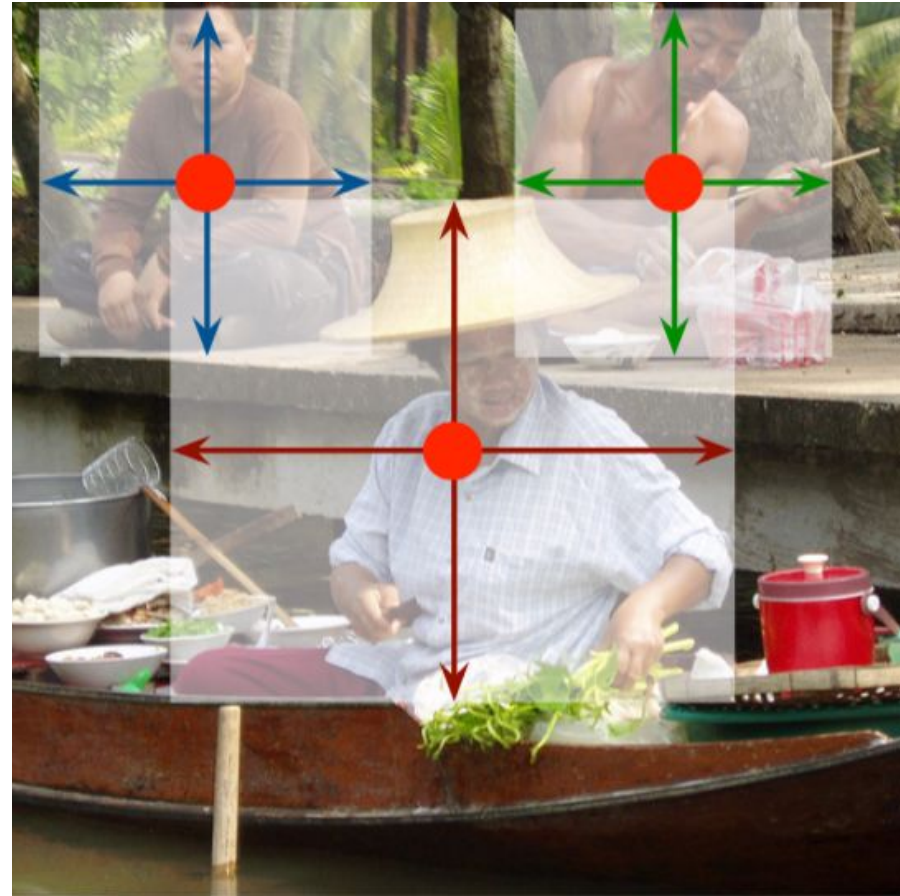
DeepPose: Human Pose Estimation via Deep Neural Networks (CVPR'14)



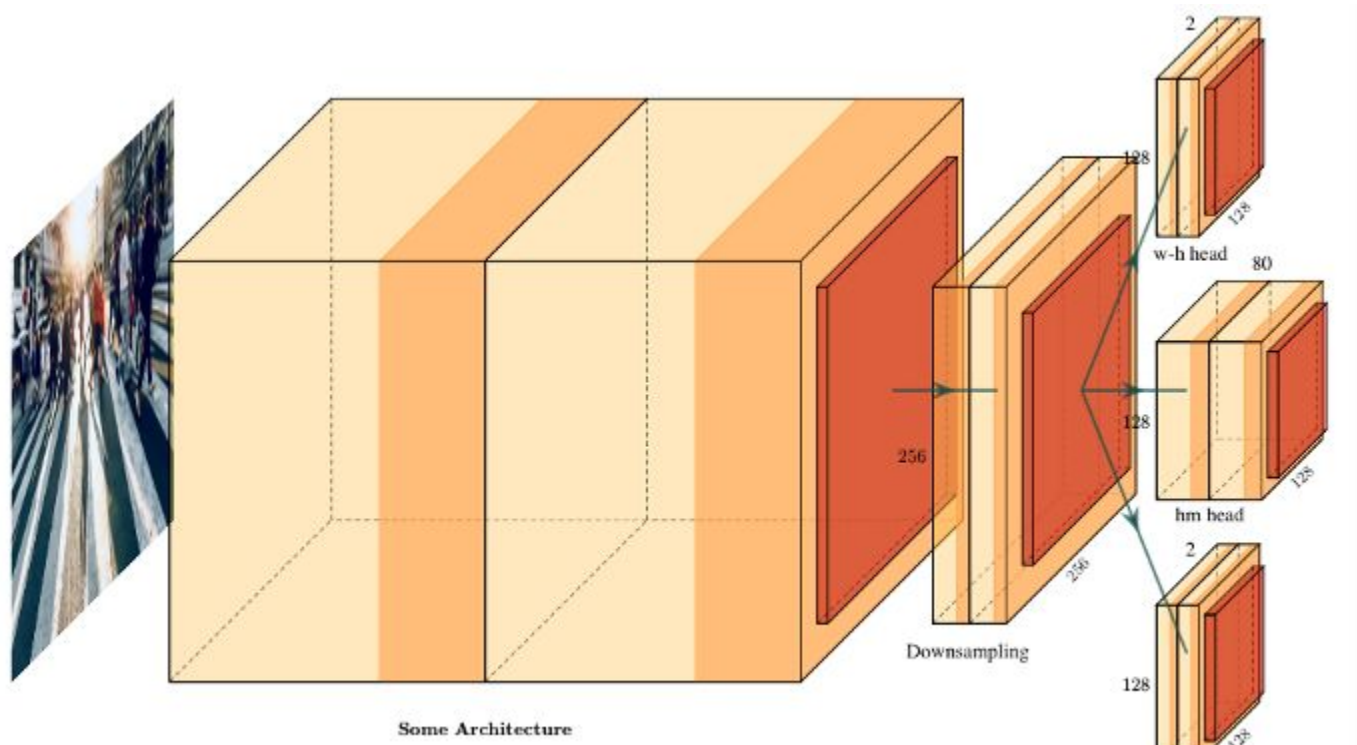
Primera investigación especializada en aplicar Deep Learning para la estimación de poses en humanos.



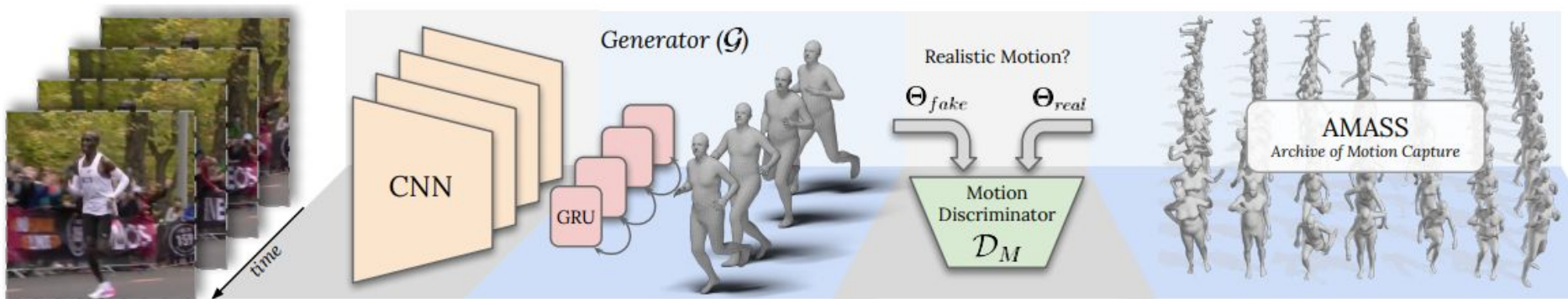
Objects as Points(CenterNet, 2019)



Uso de técnica de detección de keypoints para detectar el bounding box del center point y obtener otras propiedades como el tamaño del bounding box y estimación de poses.



Realiza la estimación de un modelo corporal por cada frame de un video usando una red de generación temporal entrenada con un discriminador de movimiento con acceso a un gran conjunto de movimientos del cuerpo.





Action Recognition



HCI, Game and Animation

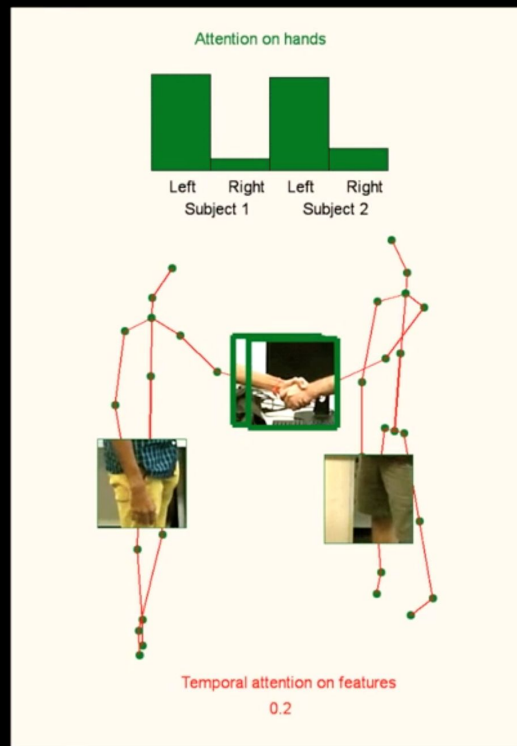


Clothing Parsing
[Yamaguchi et al. CVPR'14]

RGB raw sequence



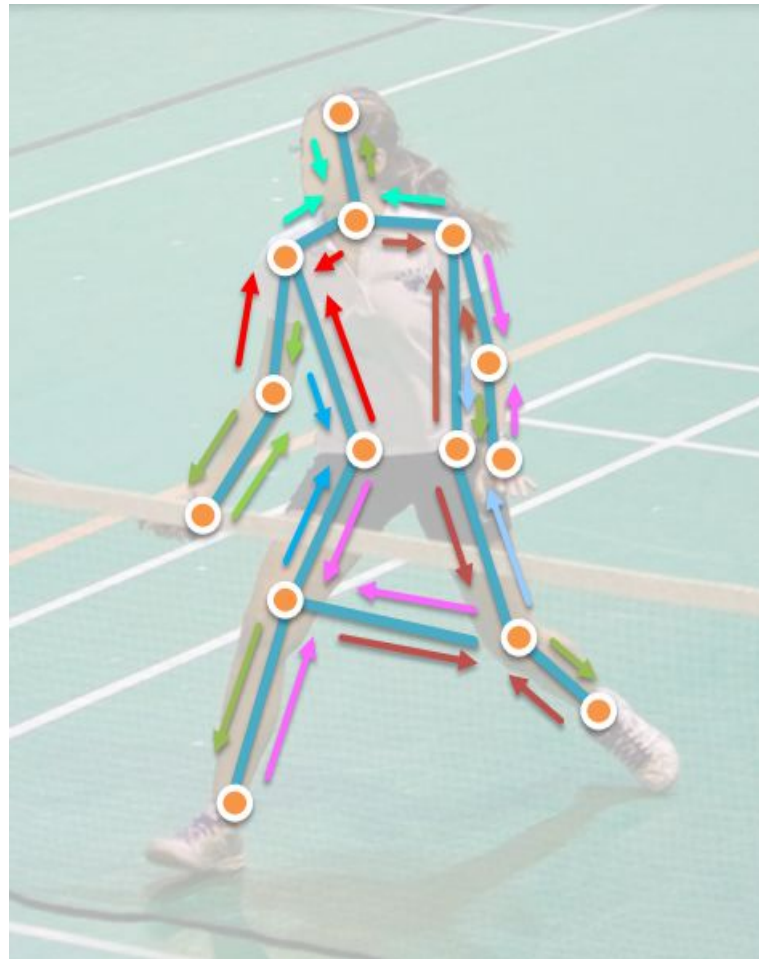
Visualization of the attention process
of our model



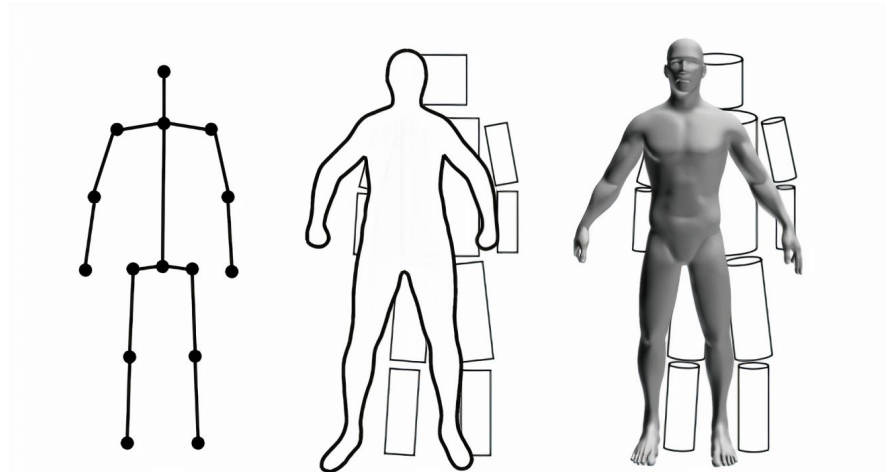
Estimación de pose

Tecnología de Computer Vision que detecta y analiza la postura humana. Su principal componente es el modelado del cuerpo humano.

También es definida como la búsqueda de una pose específica en un conjunto de poses articuladas.



Tipo de modelos de estimación de pose



skeleton-based
model

contour-based
model

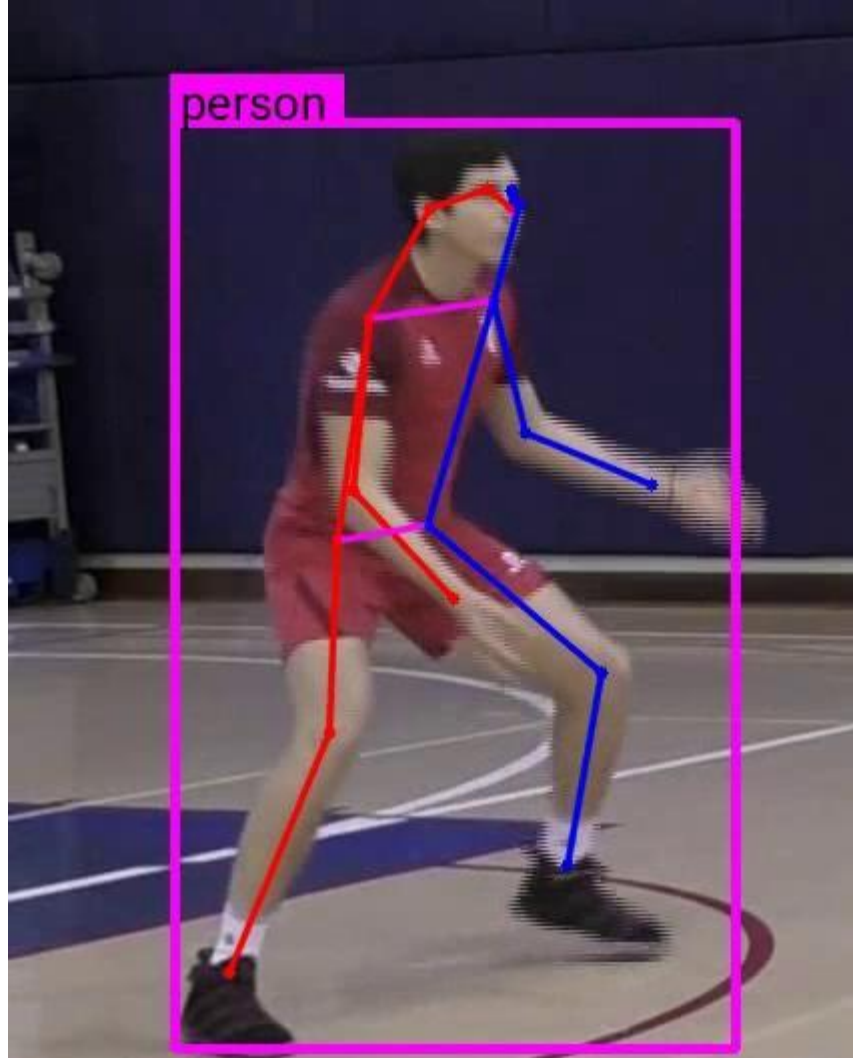
volume-based
model

Centernet





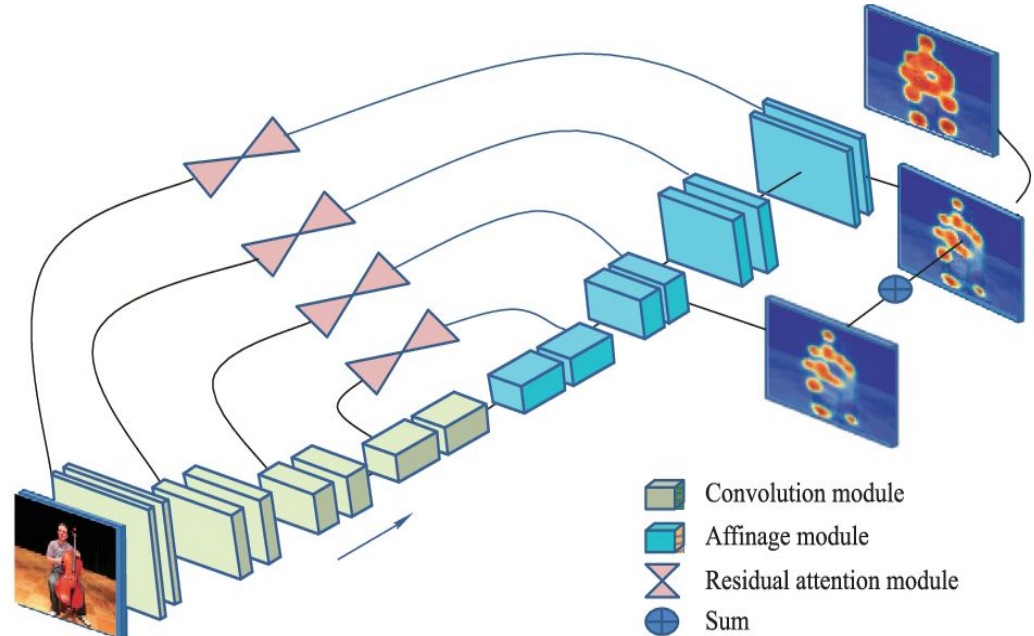
Centernet



Implementación



Keras



hole framework of our method (In comparison with convolutional stacked hourglass network for human pose e

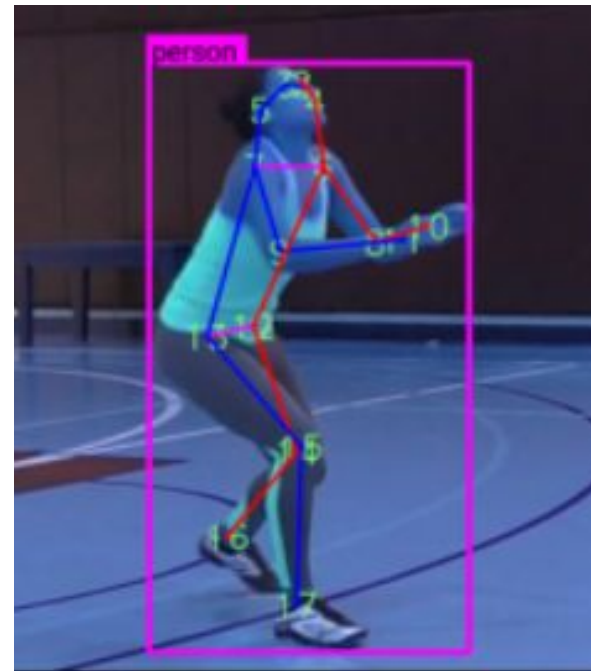
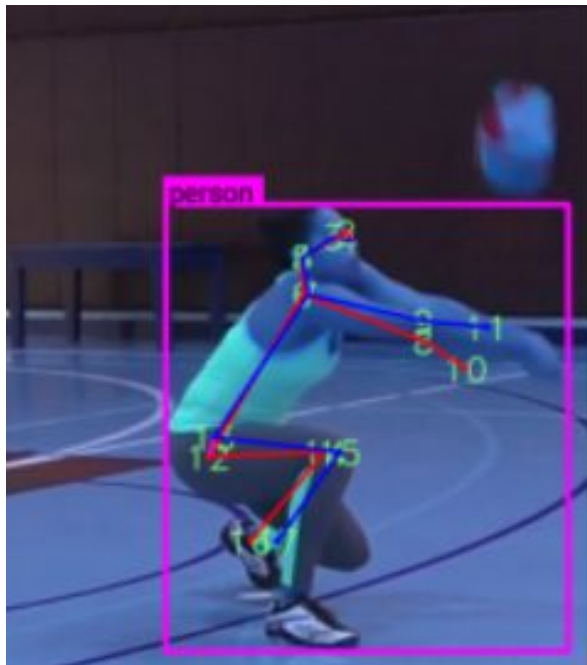


Resultados





Resultados





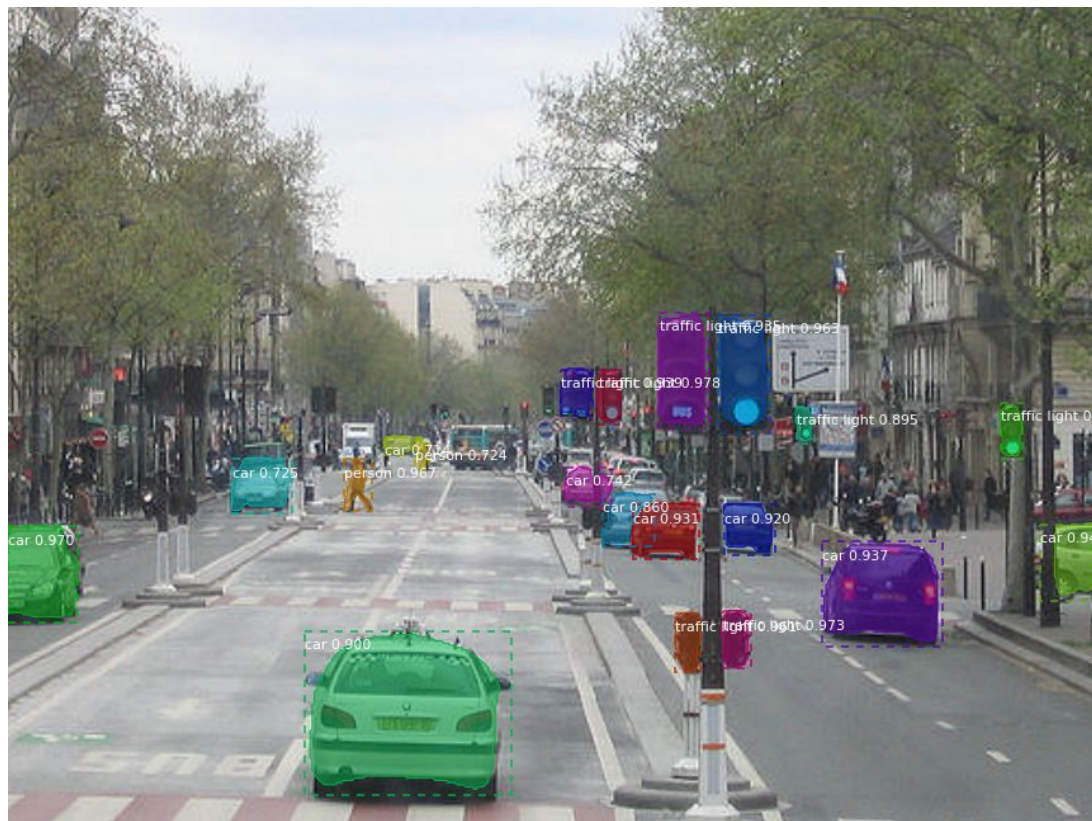
Estimación de ángulos

$$\cos(\theta) = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\| \|\mathbf{B}\|} = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}}$$



Tipo	Lado	Ángulo
Pierna	Derecha	130.22
Pierna	Izquierda	121.52
Brazo	Izquierda	171.5
Brazo	Derecha	156.49
Cintura	Derecha	107.78
Cintura	Izquierda	89.26

Mask RCNN (2017)



“Mask R-CNN for object detection and instance segmentation on Keras and TensorFlow” [\[Link\]](#)

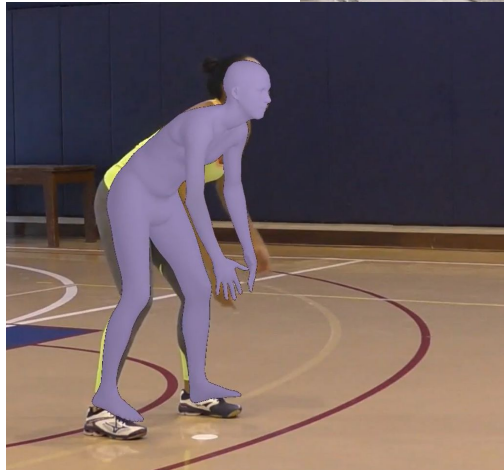


Mask-RCNN

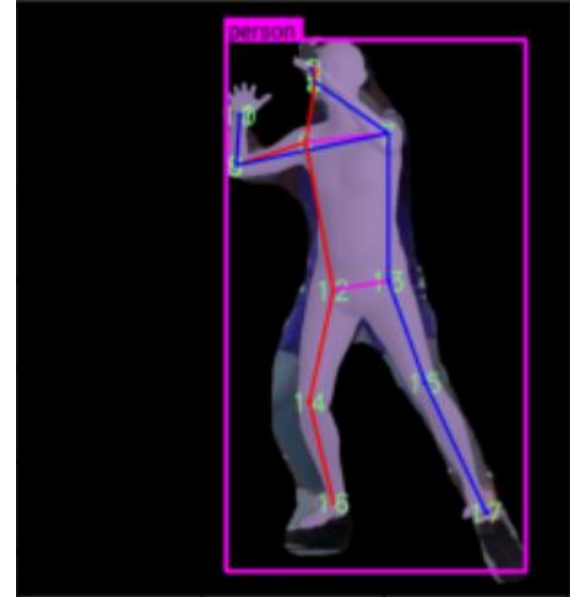
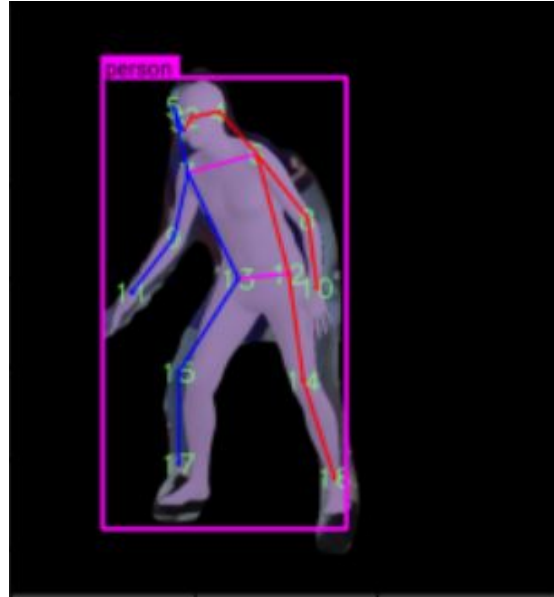
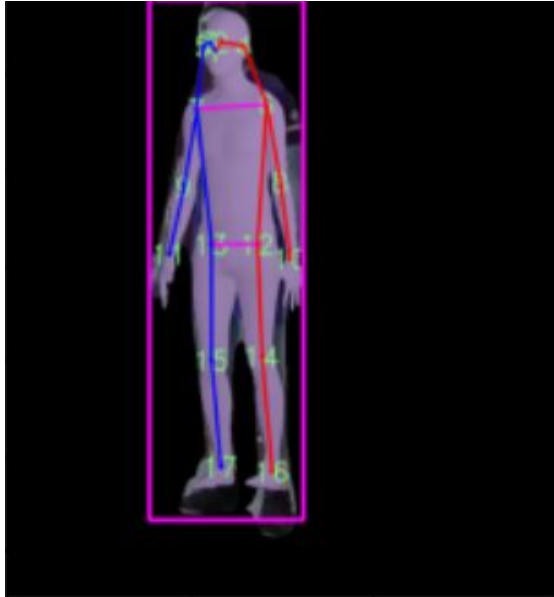
Shiny



VIBE: Video Inference for Human Body Pose and Shape Estimation (2020)



Resultados con Implementación de Centernet



VideoPose3D

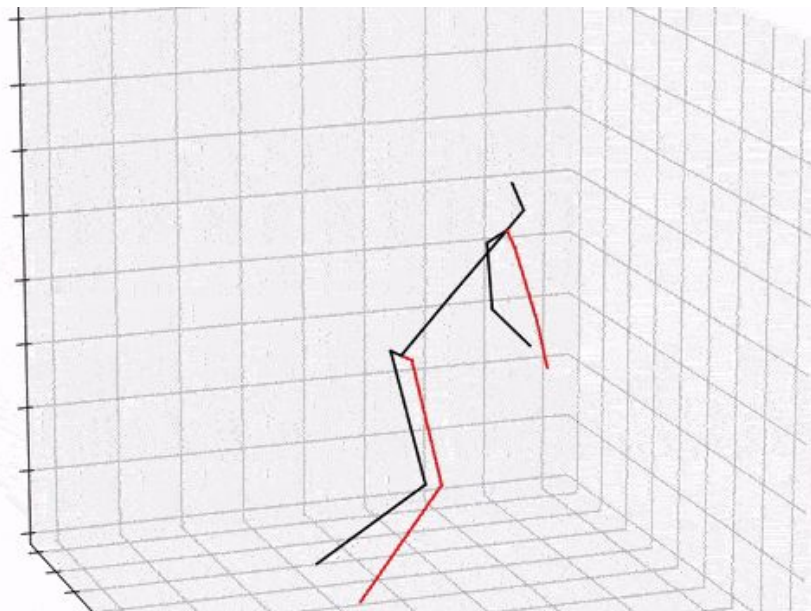




VideoPose3D - Facebook Research

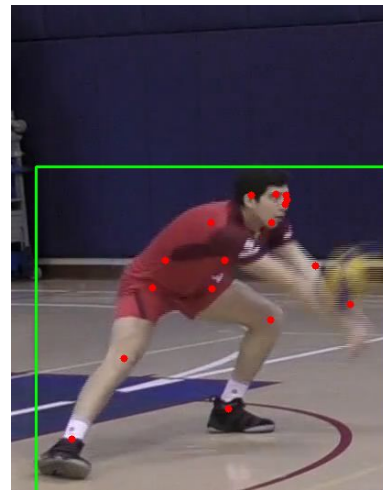
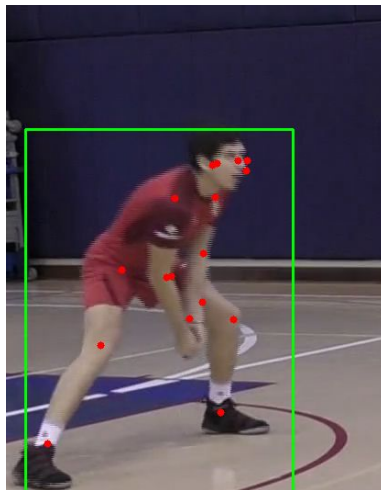
<https://github.com/facebookresearch/VideoPose3D>

Input



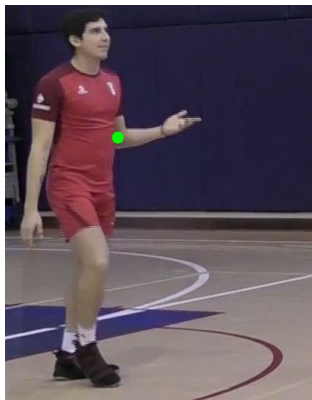
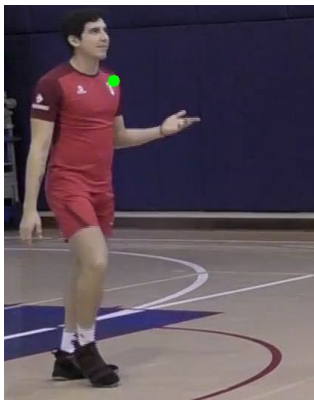
VideoPose3D - Facebook Research

Detección de bounding boxes y keypoints en 2D mediante **Detectron**





VideoPose3D - Facebook Research



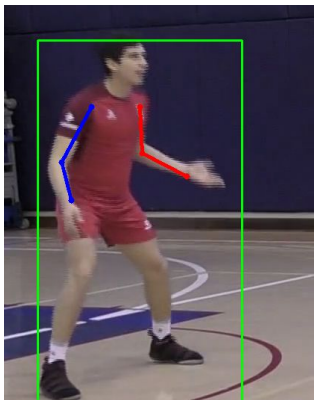
Detectron2

Luego de la extracción de 17 keypoints por frame se asocian los resultados obtenidos a las articulaciones correspondientes.

Por ejemplo, se identifican los keypoints del brazo izquierdo.

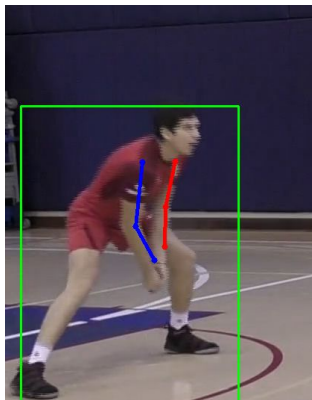


VideoPose3D - Facebook Research



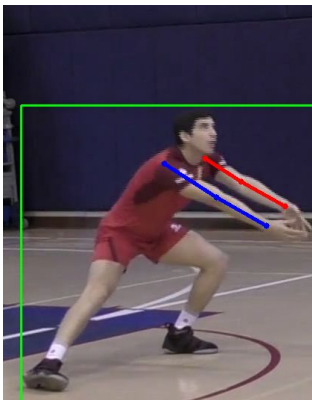
119.06°

136.92°



167.93°

145.02°



176.52°

177.38°

Detectron2

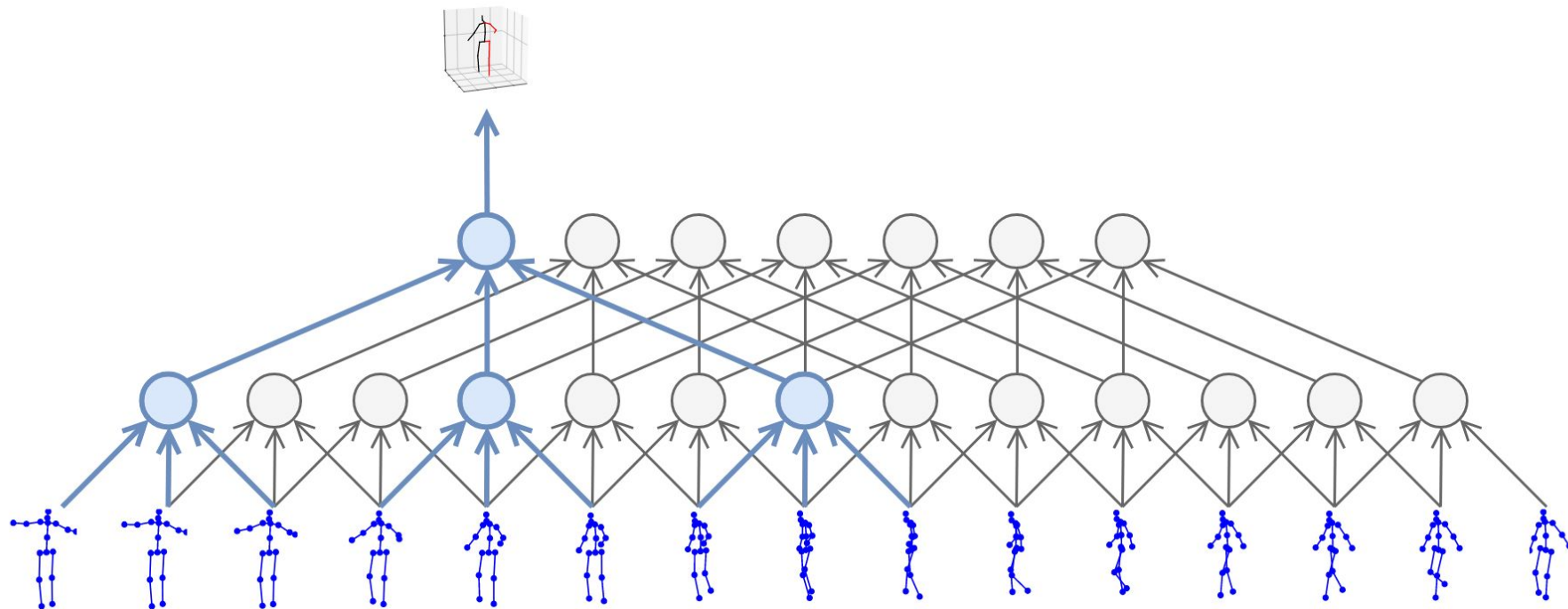
Se puede realizar el seguimiento continuo a la posición de los brazos utilizando los keypoints

El cálculo del **ángulo** entre brazo y antebrazo nos permite evaluar la calidad de recepción



VideoPose3D - Facebook Research

A partir de los keypoints en 2D se estiman los keypoints en 3D usando un modelo temporal





Repositorios

Repositorio principal

<https://github.com/jchaconm/PUCP-VC-Grupo3-Volleyball>

Repositorio experimentos con VideoPose3D

<https://github.com/manuel-montoya-gamio/VideoPose3D>