

Comptador automàtic de repeticions d'exercicis per Visió per Computador



Joan Lara Formoso
2025/2026

Treball Final de Grau en Enginyeria Informàtica
Menció en Tecnologies de la Informació

Introducció i motivació



Amb l'arribada de l'**Intel·ligència artificial**, la tecnologia es troba avançant a gran velocitat a gairebé tots els camps de la nostra vida diària, incloent els esports.

Un dels camps on lentament s'està estandarditzant un ús professional de la tecnologia és el **fitness**, que engloba exercicis aeròbics, l'aixecament de peses, CrossFit i molts altres tipus de competicions similars.

Degut als alts costos d'equips d'anàlisi professionals, en molts d'arbitratges, el comptatge de repeticions es realitza encara amb mesures subjectives i no regulades. En aquest projecte es crea un **comptador de repeticions estandarditzat per a exercicis de fitness amb tecnologia actual de Visió per Computador**, amb un ús senzill per a l'usuari amb capacitat de generalització per a càmeres no professionals.



Resultats

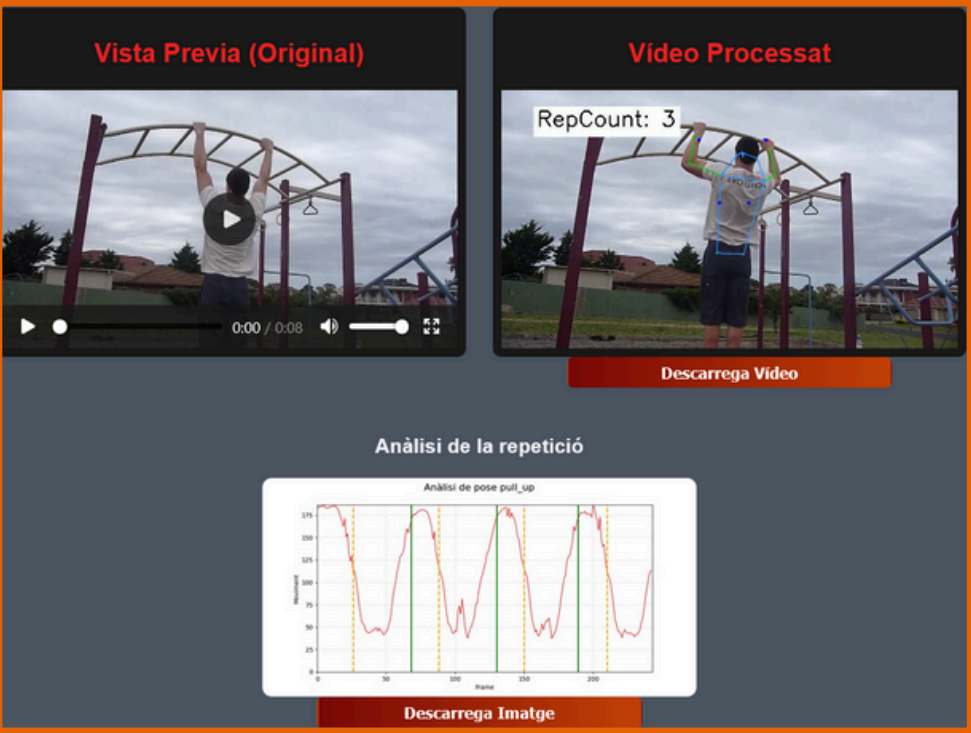


Predicció de tipus d'exercici: Gràcies a l'ús de la xarxa Sapiens, la qual és gran i precisa, per a l'estimació de pose a l'entrenament del LSTM de predicció d'exercici, s'obté un 85% de correctesa a la predicció de l'exercici corresponent, tenint en compte diferents angles de gravació.

Velocitat de processament: S'obté una velocitat mitjana de 32.7 frames per segon (FPS) i 19.5 FPS tenint en compte la generació del vídeo de sortida.

Filtratge d'errors: El comptador de repeticions té en compte els possibles errors de la estimació de pose i l'homografia, i filtra outliers i errors per afegir robustesa.

Ús d'interfície en línia: La eina web ofereix un ús senzill sense necessitat de configuració, pensada per a tot tipus d'usuari sense requerir un format de gravació específic ni instal·lació de cap programa.



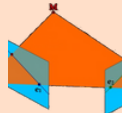
Metodologia



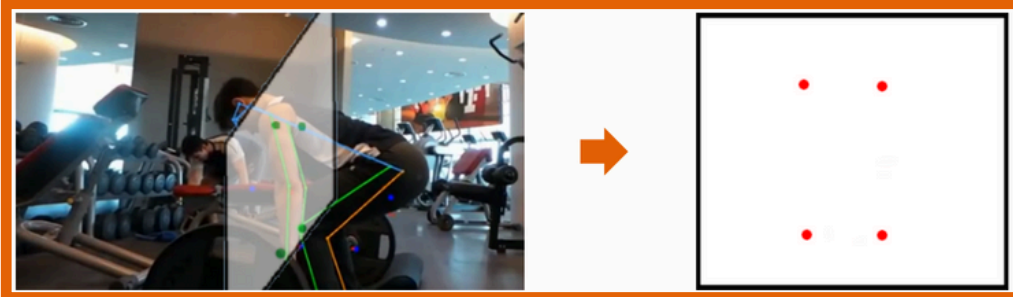
Estimació de pose: A cada frame del vídeo una xarxa neuronal YOLOv11 obté la posició de les articulacions de la persona al vídeo (keypoints).



Predicció d'exercici: Una xarxa neuronal recurrent (RNN) s'encarrega de classificar les seqüències de vídeo segons la classe d'exercici. Aquesta classifica la estimació de pose resultant i no el vídeo directament, per a màxima velocitat!



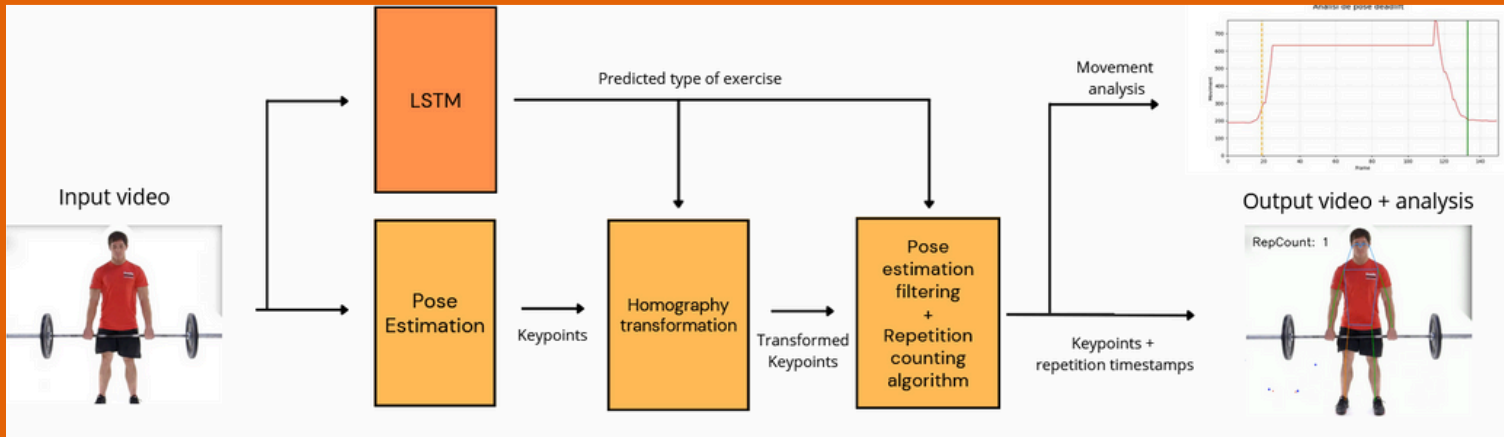
Transformació per homografia: La estimació de pose obtinguda es transforma per a estandarditzar el pla de moviment de l'exercici de l'usuari independentment de l'angle de càmera.



Comptador de repeticions: Amb les articulacions transformades a un pla, s'analiza el moviment vertical entre les articulacions a l'exercici i es determina quan comencen i acaben les repeticions.



Integració amb API web: La eina es pot utilitzar en línia, on l'usuari pot pujar un vídeo per obtenir el seu anàlisi de repeticions i visualitzar el seu historial d'ús.



Conclusions



La característica més destacada de la eina és la **robustesa** cap a diferents angles de càmera.



El mètode emprat per calcular els instants de les repeticions només té en compte el moviment vertical que es realitza al dur a terme un exercici, però un anàlisi més complex podria donar **més aplicacions, com l'anàlisi de postura**.



Aquest mètode pot ser aplicat a l'anàlisi de múltiples exercicis **sense necessitat d'una càmera professional i des d'una sola perspectiva**.

