

Informe inicial 1

Web de comptador automàtic de repeticions d'exercicis amb Visió per Computador

Joan Lara Formoso

3 d'octubre de 2025

1 INTRODUCCIÓ - CONTEXT DEL TREBALL

El camp de la visió per computador, juntament amb el de la Intel·ligència Artificial, ofereix noves i millors eines a un ritme que s'accelera de forma gairebé exponencial. Un dels camps on es troba un gran potencial d'aplicació és en els esports professionals, degut a que molts d'aquests encara depenen d'un biaix humà en la seva avaluació i mesura. És aquí on la detecció automàtica pot cobrir aquest espai, amb models d'anàlisi de pose humana cada vegada més ràpids i precisos. (Fig. 1)

En base a aquesta problemàtica, en aquest treball es buscarà desenvolupar una eina automàtica que durà a terme dues tasques: determinar automàticament l'exercici realitzat i trobar els moments exactes on es comença i on s'acaba una repetició de l'exercici. Es focalitzarà en la precisió dels càlculs i en l'ús senzill per a l'usuari d'aquesta eina. Per això, es desenvoluparà una aplicació web des d'on es podrà fer servir.



Fig. 1: Exemple de Human Pose Estimation amb Sapiens

2 OBJECTIU

En aquesta secció es mostren els objectius als quals s'ha d'arribar per a la finalització del projecte. S'indica, a nivell tècnic, que es requereix per a cada tasca, i a les seccions posteriors s'explicarà la planificació i metodologia que es seguirà per dur-les a terme.

- E-mail de contacte: larafirmosojan@gmail.com
- Menció: Tecnologies de la Informació
- Treball tutoritzat per: Coen Antens (Centre de Visió per Computador)
- Curs 2025/26

- Revisar l'estat de l'art i les tècniques ja utilitzades per resoldre la problemàtica. Recopilació d'informació, mètodes i selecció de models i tècniques a utilitzar.
- Recopilació i formatatge conjunt de dades (dataset). Recopilació d'imatges i vídeos, anotacions i adaptació per a entrenar i inferir amb els diferents models.
- Anotació d'imatges per al model de segmentació (entrenament del model YOLO11seg [8]).
- Extracció de resultats Pose Estimation (ús del model de Meta: Sapiens [7]).
- Entrenament i extracció de resultats de la RNN (Recurrent Neural Network) per a la predicción del tipus d'exercici.
- Combinació del model de segmentació i de pose per construir el comptador de repeticions.
- Creació i configuració d'API web utilitzant frameworks existents (Vue, React, Angular, etc.) amb mesures bàsiques de seguretat, com l'autenticació d'usuaris.
- Unificació de les diferents parts del projecte i integració amb l'API web.

3 METODOLOGIA

Aquest projecte seguirà la metodologia SCRUM. Aquesta és una estratègia de planificació per a projectes utilitzada extensament al camp de l'enginyeria informàtica. Aquesta metodologia divideix els objectius del projecte en "sprints". Per a cada sprint, que comença i acaba en dates preestablertes, s'ha de complir amb unes tasques determinades. Al final d'aquests es fa una "sprint review", és a dir, una avaluació de les tasques acabades així com una reavaluació de com procedir. Addicionalment, a aquest projecte es realitzarà una reunió setmanal per fer un millor seguiment i poder fer canvis menors de direcció abans de les revisions de sprint.

Per seleccionar el programa amb el qual es farà aquest seguiment, s'ha provat d'utilitzar diverses opcions: Trello, Monday i Asana. Després de realitzar la planificació amb les tres, s'ha escollit Asana com a eina. Aquesta ofereix més opcions i detalls a afegir a les tasques, com la prioritat o la

data d'inici, a més de tenir una versió gratuïta del programa que s'adqua a les necessitats d'aquest projecte. Monday no ofereix una versió gratuïta, i Trello no permet els camps a les tasques que sí permet Asana.

4 PLANIFICACIÓ

Les dates dels sprints review coincideixen amb les dels informes de seguiment. Així doncs, l'informe de seguiment inclourà un seguiment de les tasques desenvolupades en el corresponent sprint, així com la direcció a seguir per al següent sprint. La data final de cada sprint és marcada per l'entrega de l'informe de seguiment corresponent; però, per a una millor gestió del temps dins de cada sprint, les principals tasques tenen dates d'entrega que poden ser més flexibles segons el transcurs del projecte.

- 15/9 - 3/10 Kick-off del projecte.
- 4/10 - 12/11 Sprint 1
- 13/11 - 10/12 Sprint 2
- 11/12 - 12/11 Sprint 3

4.1 Cronograma

La següent figura (Fig. 2) inclou la planificació del projecte a la plataforma Asana, on s'indiquen les diferents dates d'entrega.

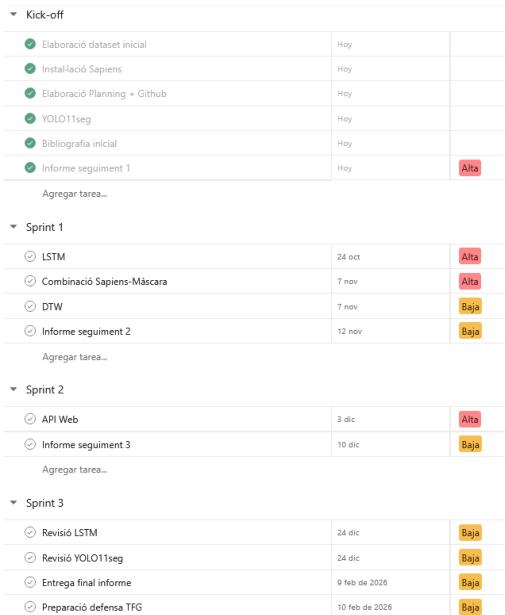


Fig. 2: Planificació Projecte a Asana

A continuació es mostra el diagrama de Gantt conseqüent de l'anterior planificació (Fig. 3). Aquest ha sigut realitzat a la plataforma ProjectLibre. Aquest mostra les dependències que es troben entre les diferents tasques dins de cada sprint.



Fig. 3: Diagrama de Gantt TFG - ProjectLibre

REFERÈNCIES

- [1] Kaggle, Workout/Exercises Video Dataset. [Online]. Available: <https://www.kaggle.com/datasets/hasyimabdillah/workoutfitness-video?resource=download>
- [2] M. Zaher, A. S. Ghoneim, L. Abdelhamid, and A. Atia, “Unlocking the potential of RNN and CNN models for accurate rehabilitation exercise classification on multi-datasets,” *Multimedia Tools and Applications*, vol. 84, pp. 1261–1301, 2025. doi: <https://doi.org/10.1007/s11042-024-19092-0>
- [3] T. Rangari, S. Kumar, P. P. Roy, D. P. Dogra, and B. G. Kim, “Video based exercise recognition and correct pose detection,” *Multimedia Tools and Applications*, vol. 81, pp. 30267–30282, 2022. doi: <https://doi.org/10.1007/s11042-022-12299-z>
- [4] M. Slupczynski, A. Nekhviadovich, N. Duong-Trung, and S. Decker, “Analyzing Exercise Repetitions: YOLOv8-Enhanced Dynamic Time Warping Approach on InfiniteRep Dataset,” in International Workshop on Sensor-Based Activity Recognition and Artificial Intelligence, *Springer Nature Switzerland*, pp. 94–110. Sept. 2024. doi: <https://doi.org/10.1007/978-3-031-80856-2>
- [5] A. Patil, D. Rao, K. Utturwar, T. Shelke, and E. Sarda, “Body posture detection and motion tracking using AI for medical exercises and recommendation system,” in ITM Web of Conferences, vol. 44, p. 03043, 2022. doi: <https://doi.org/10.1051/itmconf/20224403043>
- [6] Q. Yu, H. Wang, F. Laamarti, and A. El Saddik, “Deep learning-enabled multitask system for exercise recognition and counting,” *Multimodal Technologies and Interaction*, vol. 5, no. 9, p. 55, 2021. doi: <https://doi.org/10.3390/mti5090055>
- [7] Meta, Sapiens Pose Estimation. [Online]. Available: <https://github.com/facebookresearch/sapiens>
- [8] Ultralytics, YOLO11 Models. [Online]. Available: <https://docs.ultralytics.com/es/models/yolo11/>