
This is the **published version** of the bachelor thesis:

Castillejo Roig, Lara. *Sistema Automàtic de Detecció d'Errors en Gimnàstica Artística mitjançant Deep Learning*. Treball de Final de Grau (Universitat Autònoma de Barcelona), 2026 (Enginyeria Informàtica)

This version is available at <https://ddd.uab.cat/record/334469>

under the terms of the  license.

Sistema Automàtic de Detecció d'Errors en Gimnàstica Artística mitjançant Deep Learning

Lara Castillejo Roig

19 de juny de 2026

Resum– La gimnàstica artística exigeix una precisió biomecànica estricta, avaluada actualment mitjançant el criteri visual humà segons el Codi de Puntuació de la Federació Internacional de Gimnàstica (FIG) o amb infraestructures d'alt cost. Aquest projecte proposa el desenvolupament d'un sistema automàtic i *low-cost* de detecció d'errors tècnics basat en Visió per Computador i *Deep Learning*. Utilitzant models d'estimació de postura (*Pose Estimation*) sobre vídeos monoculars estàndard, el sistema extreu l'esquelet de l'atleta, en calcula els angles de les articulacions i hi aplica les regles de la FIG per estimar les penalitzacions d'execució (*E-Score*). L'objectiu és democratitzar l'anàlisi esportiva, oferint una eina accessible que generi *feedback* visual objectiu per assistir en els entrenaments i competicions nacionals.

Paraules clau– Visió per Computador, Deep Learning, Estimació de Posi, Gimnàstica Artística, Biomecànica, Anàlisi Esportiu

Abstract– Gymnastics demands strict biomechanical precision, currently evaluated through human visual judgment according to the International Gymnastics Federation (FIG) Code of Points or via high-cost infrastructures. This project proposes the development of an automatic, low-cost technical error detection system based on Computer Vision and *Deep Learning*. Using pretrained pose estimation models on standard monocular videos, the system extracts the athlete's skeleton, calculates joint geometry, and applies FIG rules to estimate execution penalties (*E-Score*). The goal is to democratize sport analysis, offering an accessible tool that generates objective visual feedback to assist in training and national competitions.

Keywords– Computer Vision, Deep Learning, Pose Estimation, Gymnastics, Biomechanics, Sports Analytics

1 INTRODUCCIÓ

ELS Jocs Olímpics de París 2024 van evidenciar una problemàtica històrica en la gimnàstica artística: la immensa dificultat d'arbitrar rutines d'alta velocitat de forma exclusivament manual. La polèmica mediàtica sorgida a la final de terra femenina, centrada en la validació del grau de rotació d'un salt acrobàtic per a la Nota de

Dificultat (*D-Score*)[1], va posar de manifest la pressió i la càrrega cognitiva a la qual estan sotmesos els jutges. Si bé en la dificultat existeix una via de reclamació oficial, els jutges depenen exclusivament de la seva agudesesa visual per detectar petits errors biomecànics en fraccions de segon. Aquest factor humà introduït una inevitable subjectivitat en l'avaluació de la Nota d'Execució (*E-Score*) [2]. Aquesta nota avalua exclusivament la perfecció tècnica de l'exercici: parteix d'una base de 10.0 punts sobre la qual els jutges apliquen deduccions per cada error de postura, flexió o aterratge. Es tracta d'una puntuació crítica que, a diferència de la dificultat, és estrictament inalterable segons el reglament.

Tot i que la intel·ligència artificial ja s'utilitza a competicions Olímpiques mitjançant el sistema multimilionari d'escaneig 3D *Judging Support System*, desenvolupat per

• E-mail de contacte: 1641027@uab.cat
• Menció: Computació
• Treball tutoritzat per: Coen Antens (Ciències de la Computació i Intel·ligència Artificial)
• Curs 2025/26

Fujitsu [3], el seu elevat cost el fa inaccessible fora de l'èlit mundial. Això deixa milers de clubs de tecnificació i atletes amateurs sense eines d'anàlisi biomecànica objectives.

2 OBJECTIUS

L'objectiu principal d'aquest projecte és democratitzar l'accés a l'anàlisi esportiva d'elit mitjançant el desenvolupament d'un sistema d'assistència a l'arbitratge basat en Visió per Computador de baix cost a partir de vídeos monoculars.

Per tal d'assolir aquesta fita, el projecte s'estructura en tres objectius específics fonamentals:

1. **Democratització Tecnològica:** Validar un *pipeline* d'extracció de característiques (*Human Pose Estimation*) capaç de suportar oclusions i inversions complexes utilitzant únicament vídeo estàndard, evitant la dependència d'infraestructures de sensors 3D d'alt cost per a l'esport base.
2. **Objectivació Matemàtica de l'E-Score:** Estandarditzar el judici d'execució traduint el Codi de Puntuació de la FIG [8] en restriccions geomètriques deterministes, recolzades per una classificació algorísmica temporal de les acrobàcies mitjançant xarxes neuronals recurrents.
3. **Interacció Human-in-the-Loop:** Construir una aplicació d'escriptori col·laborativa que actuï com a assistent de precisió. La interfície ha de projectar el *feedback* biomecànic mitjançant un codi de colors intuïtiu, reduint la càrrega cognitiva però mantenint el control absolut de l'avaluació en el jutge humà.

3 ESTAT DE L'ART

La *Human Pose Estimation* (HPE, o Estimació de Pose Humana) és un camp de la Visió per Computador que té com a objectiu localitzar les articulacions clau, anomenades *keypoints*, del cos humà en imatges o vídeos. La resolució d'aquest problema ha evolucionat molt les darreres dècades gràcies als avenços en el reconeixement de patrons i el *Deep Learning*.

3.1 Història de la HPE

Històricament, la comprensió de la figura humana va començar com la detecció d'objectes tradicional (*Object Detection*), on els sistemes es limitaven a emmarcar l'individu dins d'una caps delimitadora (*bounding box*). Com a resposta a la manca d'informació biomecànica d'aquest enfocament, la recerca es va centrar en els mètodes basats en parts (*Part-based models*) [9], que identificaven extremitats de manera independent per després intentar ajuntar-les fins a evolucionar cap a les topologies de punts clau actuals (vegeu la Figura 1).

L'autèntica revolució ha arribat amb les xarxes neuronals profundes, que permeten localitzar directament les coordenades espacials (x, y) de cada articulació [10], un salt tecnològic impulsat inicialment per arquitectures pioneres com *DeepPose* [11].

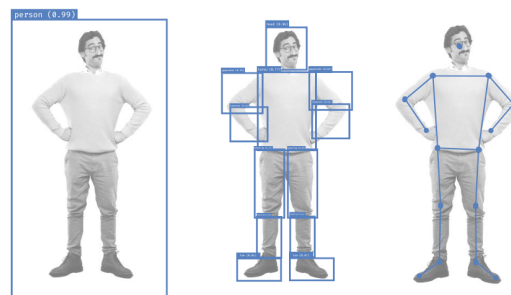


Fig. 1: Evolució cronològica i metodològica dels enfocaments de compressió de la figura humana en Visió per Computador: des de la delimitació per regions (*bounding box*) fins a l'extracció anatòmica de *keypoints*.

L'arquitectura exacta d'aquesta sortida depèn de la topologia anatòmica utilitzada: d'una banda, el format *COCO* [12] defineix un esquelet bàsic de 17 *keypoints*, mentre que el format *BlazePose* l'expandeix fins als 33 *keypoints*, capturant amb detall les mans i els peus (vegeu la Figura 2).

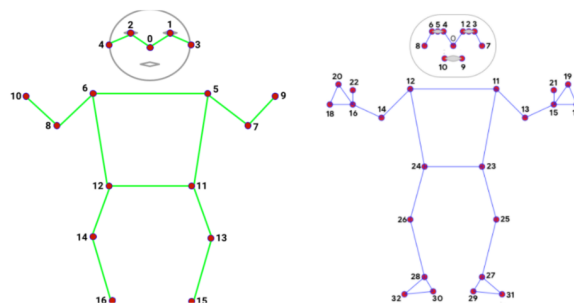


Fig. 2: Comparativa de les topologies anatòmiques estandarditzades en *Human Pose Estimation* (HPE). A l'esquerra, l'esquelet *COCO* (17 punts); a la dreta, l'esquelet *BlazePose* (33 punts).

Els models actuals d'extracció es divideixen en tres grans famílies metodològiques:

- **Models Top-Down basats en CNN:** Detecten primer la persona mitjançant una *bounding box* i posteriorment retallen la imatge per estimar-ne els *keypoints*. El màxim exponent actual d'aquesta eficiència és *YOLO* (en concret, *YOLO26x-pose*) [4], que conviu amb alternatives clàssiques com *HRNet* [13] o *AlphaPose* [14].
- **Models basats en Vision Transformers (ViT):** Arquitectures recents com *Sapiens-2B* [5] de Meta o models previs com *ViTPose* [15] processen la imatge de manera global mitjançant mecanismes d'atenció, eliminant la necessitat de retallades prèvies i entenent millor les postures complexes.
- **Solucions basades en Grafs i Seguiment Temporal:** *Frameworks* com *MediaPipe BlazePose* [7] incorporen filtres temporals que prediuen la posició basant-se en la inèrcia del *frame* anterior per garantir la coherència del moviment.

3.2 Mètriques d'Avaluació: L'indicador OKS

Per mesurar la precisió d'aquests models no es pot utilitzar una distància euclidiana simple, ja que cada articulació té una rellevància i una dificultat de detecció diferent (per exemple, és més fàcil detectar una espatlla que un turmell). Per això, l'estàndard actual és l'*Object Keypoint Similarity (OKS)* [16], que es defineix com:

$$OKS = \frac{\sum_i \exp(-d_i^2 / 2s^2 k_i^2) \delta(v_i > 0)}{\sum_i \delta(v_i > 0)} \quad (1)$$

On d_i és la distància euclidiana entre el punt predit i el real, s és el factor de normalització basat en l'àrea que ocupa la persona en la imatge, i k_i és una constant específica per a cada tipus d'articulació. Com s'il·lustra a la Figura 3, l'OKS defineix un radi d'acceptació al voltant de cada *keypoint* real, on la penalització per desviació es pondera segons la naturalesa de l'articulació.

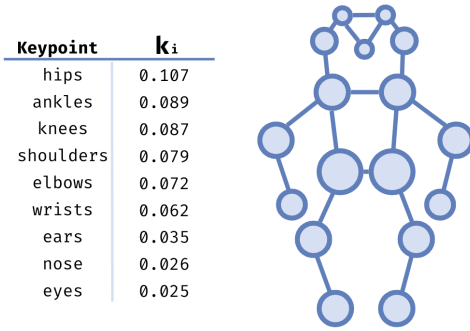


Fig. 3: Representació visual del càlcul de l'OKS. Es defineixen els diferents radis de tolerància segons l'articulació (k_i) escalats per la mida de la persona (s).

Aquesta mètrica permet calcular el *mean Average Precision (mAP)* i es relaciona amb indicadors de distància clau com el *PCK* [17], que són els valors que es comparen en les auditories de models d'aquest projecte.

3.3 Fonaments del Suavitzat Temporal: Filtre Savitzky-Golay

En l'extracció de poses *frame a frame*, el soroll de la predicció genera vibracions artificials (*jitter*). Per solucionar-ho, s'utilitza el filtre de suavitzat temporal de *Savitzky-Golay* [18].

L'algorisme realitza un ajust local d'un polinomi de grau k sobre una petita finestra de filtratge mòbil de mida senar $n = 2m + 1$, on m és el nombre de punts a banda i banda del valor central. Aquest polinomi $p(i)$ representa l'estimació de la coordenada suavitzada en l'instant relatiu i , imitant la trajectòria parabòlica i contínua natural del cos a l'aire, i es defineix com:

$$p(i) = \sum_{j=0}^k c_j i^j \quad (2)$$

L'objectiu del filtre és determinar precisament aquests coeficients c_j per tal de minimitzar l'error quadràtic mitjà

(E) respecte a les coordenades originals x_i dins de la finestra temporal:

$$E = \sum_{i=-m}^m w_i (p(i) - x_i)^2 \quad (3)$$

On w_i representa el pes de cada punt. Aquesta minimització optimitza la variància $\sigma_{\Delta e}^2$ a través dels coeficients, aconseguint estabilitzar la trajectòria de les articulacions sense eliminar la informació clau ni els moviments reals d'alta velocitat de la gimnasta.

3.4 Xarxes Recurrents i Arquitectura LSTM

Per classificar seqüències d'acrobàcies en vídeos, les xarxes neuronals estàndard no serveixen perquè no poden tractar dades seqüencials. Per això s'utilitzen les Xarxes Neuronals Recurrents (*RNN*), que mantenen un estat ocult que actua com a memòria històrica dels *frames* anteriors.

Dins d'aquestes, les xarxes *Long Short-Term Memory (LSTM)* [19] són les més eficaçes perquè evitin que la informació es perdi en seqüències llargues. Com es mostra a la Figura 4, un bloc LSTM controla el flux d'informació mitjançant tres portes:

- **Forget Gate (F_t):** Decideix quina informació antiga ja no és important i s'ha d'esborrar.
- **Input Gate (I_t):** Tria quins nous detalls de la postura actual s'han de guardar a la memòria.
- **Output Gate (O_t):** Filtra la informació acumulada per decidir quina és la predicció de l'acrobàcia en aquell instant.

Aquest mecanisme de portes permet que la xarxa aprengui patrons seqüencials complexos, ignorant el soroll i centrant-se en l'evolució de l'esquelet durant el salt.

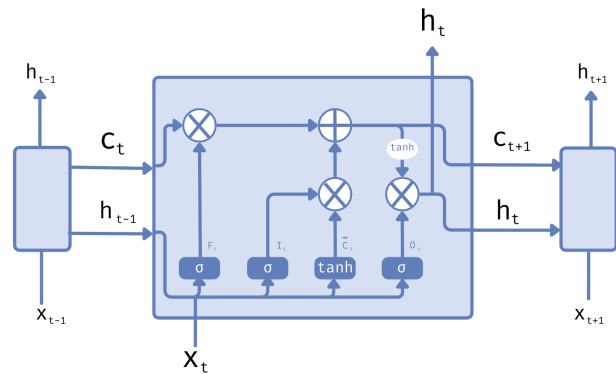


Fig. 4: Arquitectura d'un bloc recurrent LSTM, on es detalla el flux temporal controlat per les *Forget Gate* (F_t), *Input Gate* (I_t), *Output Gate* (O_t) i la cel·la de memòria candidat ($\tilde{C}_t$).

4 METODOLOGIA

El desenvolupament d'aquest projecte segueix una metodologia que combina la gestió de processos àgils amb una selecció tecnològica orientada a la Visió per Computador.

4.1 Metodologia de Desenvolupament

La gestió del projecte segueix una metodologia àgil basada en el sistema *Kanban*, un enfocament iteratiu que aporta la flexibilitat necessària per adaptar-se als reptes tècnics i facilitar el testeig continu del codi. Com a eina central s'utilitza la plataforma *Notion*, on s'estructura el flux de treball en estats de seguiment ("Per fer", "En curs", "Finalitzat").

L'execució es divideix en fases seqüencials, aplicant la regla estricta de complir totes les tasques d'una iteració abans d'avançar a la següent. Així mateix, aquesta plataforma centralitza la gestió documental i el seguiment temporal mitjançant el diagrama de *Gantt*, garantint un control dels terminis acadèmics.

4.2 Metodologia Tècnica

L'entorn tecnològic es fonamenta en *Python* sota una arquitectura modular que separa la visió, la geometria i la interfície. El nucli empra *PyTorch* per a la inferència del model *Sapiens-2B* i l'entrenament de la xarxa *LSTM*, juntament amb *OpenCV* com a motor de visió per al tractament d'imatge i la renderització vectorial de l'esquelet.

Per al processament numèric i l'estabilització del senyal mitjançant el filtre, s'ha utilitzat *SciPy* per al filtre de *Savitzky-Golay*, recolzant-se en *NumPy* i *Pandas*, i *Scikit-Learn* per a poder realitzar la imputació de dades per *KNN*. Finalment, l'eina gràfica interactiva per al jutge es desenvolupa amb *PyQt6*.

Amb l'objectiu de garantir la transparència i la reproducibilitat del sistema, el codi font complet s'ha centralitzat en un repositori públic de GitHub: <https://github.com/loranyeta/gymnastics-error-detection>.

5 PLANIFICACIÓ

El desenvolupament del projecte s'ha estructurat en cinc fases seqüencials, complint els terminis establerts (vegeu el diagrama de *Gantt* a l'Apèndix B):

- **Fase 1: Recerca i Viabilitat.** Definició de l'abast i estudi de l'estat de l'art en models d'estimació de pose (*HPE*).
- **Fase 2: Extracció i Processament.** Implementació de *Sapiens-2B* i filtratge per a l'estabilitat espacial de l'esquelet.
- **Fase 3: Classificació i Lògica.** Entrenament de la xarxa *LSTM* per detectar acrobàcies i programació del motor de deduccions (*FIG*).
- **Fase 4: Interfície i Integració.** Desenvolupament de l'aplicació (*PyQt6*) i avaluació de rutines contínues (*Sliding Window*).
- **Fase 5: Tancament.** Redacció d'aquesta memòria acadèmica i preparació de la defensa.

6 DESENVOLUPAMENT

Aquesta secció detalla l'execució tècnica del projecte, estructurada en el *pipeline* complet de processament, des de l'estandardització de les dades fins al disseny definitiu de la interfície d'usuari.

6.1 Arquitectura de Programari i Estructura de Mòduls

Per garantir la mantenibilitat, l'escalabilitat i una correcta separació de conceptes (*separation of concerns*), el projecte s'ha dissenyat seguint una arquitectura modular. Com es detalla a l'Apèndix A, aquesta organització separa completament les tasques d'inferència de xarxes neuronals profundes, el processament geomètric de senyal i la representació gràfica de la interfície d'usuari. Aquesta decisió de disseny permet que el sistema operi mitjançant un flux de dades estandarditzat, basat principalment en l'intercanvi d'arxius estructurats. D'aquesta manera, el nucli computacional (*backend*), encarregat de la gestió del model de visió i els càlculs trigonomètrics, es manté aïllat de la lògica de presentació (*frontend*). Aquest aïllament resulta fonamental de cara a futures ampliacions, ja que possibilita la substitució del model d'extracció de característiques per versions més modernes sense necessitat de reescriure el motor d'avaluació de l'*E-Score* ni de modificar l'experiència d'usuari. A més a més, facilita el testeig del codi en permetre auditar els errors de visió artificial independentment de l'eina gràfica final.

6.2 Estandardització de l'Esquelet: Migració de Keypoints

Després de seleccionar *Sapiens-2B* com a motor d'extracció (Secció 7.1), va sorgir un repte de compatibilitat estructural. El motor d'avaluació geomètrica de l'*E-Score* s'havia dissenyat per a la topologia de 33 *keypoints* de *MediaPipe*, mentre que *Sapiens* opera amb els formats *COCO*.

Això presentava un doble inconvenient: el format *COCO* estàndard (17 punts) no detecta els peus (indispensables per avaluar obertures i aterratges), mentre que la variant *COCO WholeBody* (133 punts) sobrecarrega el temps d'entrenament de la xarxa recurrent i el còmput global. En canvi, l'estructura de *MediaPipe* (33 punts) ofereix l'equilibri perfecte entre pes i detall anatòmic.

Per resoldre-ho, s'ha implementat un algorisme de mapeig (vegeu la Figura 2 a l'Estat de l'Art). Aquesta funció llegeix el JSON de *Sapiens* i reescriu les coordenades directament sobre la topologia de *MediaPipe*. Aquesta estandardització ha permès aprofitar l'extrema precisió de *Sapiens-2B* sense necessitat de reescriure la lògica geomètrica original, garantint un esquelet lleuger i biomecànicament complet.

6.3 Pipeline d'Extracció i Suavitzat Temporal

S'ha dissenyat un mòdul de processament seqüencial en dues fases per corregir les discontinuïtats i reduir el soroll de l'estimació espacial.

Primer, s'aborda el *frame dropping* (valors *NaN* causats per oclusions severes) mitjançant un *KNNImputer* ($k = 2$). Aquest algorisme reconstrueix els *frames* perduts recolzant-se en la coherència de les dades temporals adjacents, garantint un senyal ininterromput.

A continuació, sobre aquest flux de dades continu, s'aplica el filtre de *Savitzky-Golay*. Inicialment, el seu ús va generar dubtes perquè, en girs d'alta freqüència, l'ajust del po-

linomi provocava un sobre-suavitat (*oversmoothing*) que deformava l'intercanvi ràpid de cames (vegeu la Figura 5).

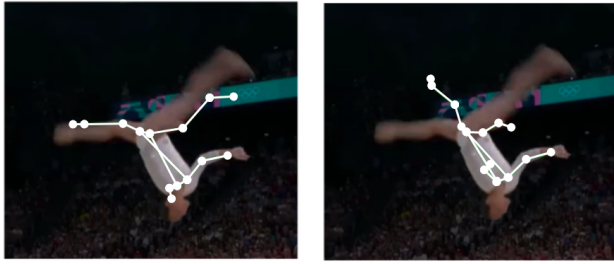


Fig. 5: Exemple d'efecte de sobre-suavitat (*oversmoothing*) en un *frame* concret d'un video de prova en un gir de perfil (dreta) respecte a l'esquelet original (esquerra).

No obstant això, aquesta estructura dual va resultar indispensable. La imputació prèvia del *KNNImputer* assegura que el filtre operi sobre un senyal regular, optimitzant l'eliminació del *jitter* (vibracions del senyal). En acrobàcies de trajectòria fluida, aquest *pipeline* assoleix un alineament biomecànic excel·lent i estable (vegeu la Figura 6).

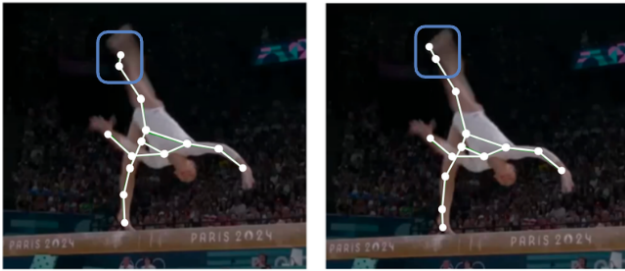


Fig. 6: Comportament òptim del filtre en un salt parabòlic, captat en un *frame* concret d'un dels videos de prova. L'esquelet corregit (dreta) presenta una continuïtat i estabilitat superior al resultat en cru del model (esquerra).

6.4 Creació del Dataset i Classificació

Per construir el *dataset*, s'han seleccionat vídeos del canal *USA Gymnastics* [21] seguint la taxonomia de *FineGym* [22]. L'estudi se centra exclusivament en la barra d'equilibris (*balance beam*), ja que la seva perspectiva de càmera lateral optimitza l'estimació 2D i concentra les principals tipologies de salts. D'aquesta manera, s'han aïllat 72 seqüències aèries classificades en quatre categories (vegeu la Figura 7).



Fig. 7: Representació de les quatre categories de salts analitzades. D'esquerra a dreta: Tuck (salt agrupat); Pike (salt carpat); Split (salt amb obertura lateral); Straddle (salt amb obertura frontal).

Per millorar la generalització del model, s'ha aplicat *Data Augmentation* per simetria (*mirroring*), invertint l'eix X

de les coordenades directament sobre el JSON per ensenyar al model la simetria de l'exercici. Això ha permès duplicar les dades fins a assolir un *dataset* de 144 mostres equilibrades.

Finalment, atès que la fase aèria d'un salt dura aproximadament un segon, les seqüències s'han estandarditzat a blocs fixos de 30 *frames*. Aquesta uniformitat tensorial evita el *padding* excessiu i accelera l'entrenament per lots (*batches*) de la xarxa *LSTM* en *PyTorch*. El model processa les coordenades i transmet el darrer estat ocult a una capa *Softmax* per predir l'acrobàcia.

6.5 Motor de Lògica d'Avaluació i Pelvis Virtual

Un cop la xarxa *LSTM* ha classificat l'acrobàcia, el mòdul *AcrobaticEvaluator* tradueix el Codi de Puntuació de la *FIG* [8] en regles geomètriques computables.

Per garantir una avaluació robusta davant la distorsió de la perspectiva monocular, s'ha dissenyat la Pelvis Virtual (vegeu la Figura 8). Calcular l'angle de cada cama per separat generaria errors segons l'orientació de la gimnasta. Aquesta solució unifica la mesura creant un punt central imaginari (P_i), calculat com el punt mig geomètric entre els malucs esquerre ($P_L = (x_L, y_L)$) i dret ($P_R = (x_R, y_R)$):

$$P_i = \left(\frac{x_L + x_R}{2}, \frac{y_L + y_R}{2} \right) \quad (4)$$

Aquest P_i estableix un origen de coordenades invariant a la rotació de l'atleta o a l'angle de gravació.

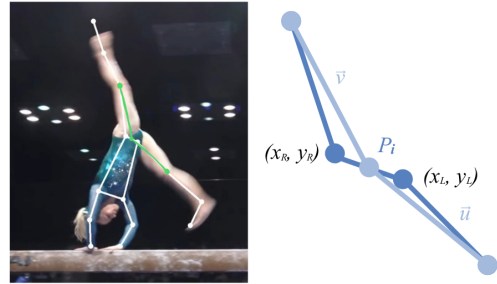


Fig. 8: Validació geomètrica de l'angle d'obertura en un salt *Split*. Es visualitzen els vectors $\vec{u}$ i $\vec{v}$ definits des del punt central de la Pelvis Virtual (P_i) pintats sobre la pose detectada.

Per calcular l'obertura real, el sistema construeix dos vectors ($\vec{u}$ i $\vec{v}$) que connecten aquest punt virtual amb els genolls (K_L i K_R):

$$\vec{u} = K_L - P_i, \quad \vec{v} = K_R - P_i \quad (5)$$

L'angle θ s'obté mitjançant el producte escalar d'ambdós vectors:

$$\theta = \arccos \left(\frac{\vec{u} \cdot \vec{v}}{|\vec{u}| |\vec{v}|} \right) \cdot \frac{180}{\pi} \quad (6)$$

Per adaptar la perfecció teòrica a la variabilitat biomecànica real, s'han definit els llindars de penalització específics de la Taula 1. De forma transversal, s'hi suma l'avaluació de l'extensió dels turmells, amb un límit fix de 130° (exclusivament per aplicar deduccions *Minor* de -0.1).

TAULA 1: LLINDARS ANGULARS I DEDUCCIONS APLICADES (*E-Score*)

Acrobàcia	Criteri	Minor (-0.1)	Medium (-0.3)	Severe (-0.5)
<i>Tuck</i>	<i>Hip</i>	< 65°	< 90°	< 100°
	<i>Knee</i>			
<i>Pike</i>	<i>Hip</i>	< 65°	< 90°	< 100°
	<i>Knee</i>	> 170°	> 160°	> 145°
<i>Split</i>	<i>Opening</i>	> 170°	> 160°	> 135°
	<i>Knee</i>	> 170°	> 160°	> 145°
<i>Straddle</i>	<i>Opening</i>	> 170°	> 160°	> 135°
	<i>Knee</i>	> 170°	> 160°	> 145°

El sistema calcula aquestes desviacions al *peak frame*, les resta de la nota base de 10.0 punts i les visualitza a la interfície amb un codi de colors intuïtiu (verd, taronja i vermell), facilitant-ne la validació al jutge humà.

6.6 Processament Temporal i Mapeig Espacial

Per tal d'adaptar les dades cinemàtiques a les restriccions d'entrada de la xarxa neuronal i permetre una avaluació visual intuïtiva, s'han implementat tres mecanismes fonamentals de processament seqüencial i espacial.

6.6.1 Estandardització Temporal d'Acrobàcies Aïllades

Tot i que la xarxa *LSTM* admet longituds variables, per optimitzar l'entrenament per lots (*batches*) i evitar l'ús ineficient de *padding*, cal garantir una dimensió matricial uniforme. Atès que la fase aèria d'una acrobàcia dura aproximadament un segon, les seqüències s'interpolen linealment fins a assolir una longitud temporal fixa de $N = 30$ frames. Aquest procés redimensiona matemàticament la seqüència afavorint l'estabilitat computacional del model, tot mantenint completament intacta la coherència cinemàtica de les posicions i acceleracions de l'esquelet.

6.6.2 Avaluació Contínua mitjançant Finestra Lliscant

Per avaluar rutines completes en lloc d'exercicis aïllats, s'ha implementat un algorisme de finestra lliscant (*sliding window*). En detectar activitat, el sistema fixa el *peak frame* com a centre i rastreja l'alçada del turmell per acotar l'inici i final del salt. Cada bloc dinàmic resultant s'envia a la xarxa *LSTM*, permetent aïllar i classificar totes les acrobàcies consecutives d'un vídeo (vegeu l'Apèndix D). Aquest procediment es detalla a la secció 6.7.

6.6.3 Mapeig Espacial i Visualització

Finalment, per renderitzar les penalitzacions de forma clara a la interfície, s'ha resolt el repte del mapeig espacial de coordenades. Les dades emmagatzemades originalment al JSON són valors normalitzats relatius al centre del tors (en

un rang aproximat de $[-1, 1]$), un format optimitzat per a l'aprenentatge profund, però no apte per a la representació gràfica directa. Per resoldre-ho, s'ha implementat una funció de desnormalització lineal que tradueix aquestes proporcions als píxels reals de la pantalla:

$$X_{pixel} = (x_{norm} \cdot S) + O_x, \quad Y_{pixel} = (y_{norm} \cdot S) + O_y \quad (7)$$

On S és un factor d'escala constant per amplificar la magnitud de l'esquelet i (O_x, O_y) representen l'*offset* necessari per centrar la figura en el pla bidimensional. Aquesta transformació permet reconstruir l'esquelet virtual amb una proporció biomecànica perfecta, facilitant la il·luminació dinàmica de les articulacions infractores (verd, taronja i vermell) segons la gravetat de la deducció calculada pel motor de regles.

6.7 Segmentació Temporal Dinàmica basada en la Cinemàtica dels Turmells

Un dels reptes identificats durant l'avaluació contínua amb l'algorisme de *sliding window* va ser la generació de falsos positius durant els segments de transició, especialment en els aterratges (*landings*). En aquests instants, la compressió del cos en contactar amb la barra provocava que la xarxa *LSTM* classifiqués el moviment com una acrobàcia activa, aplicant penalitzacions severes de forma errònia.

Per resoldre aquesta limitació sense necessitat de reenrenar el model amb noves classes, s'ha dissenyat un algorisme de segmentació cinemàtica basat en el desplaçament dels *keypoints* dels turmells (punts 27 i 28 de la topologia *BlazePose*). Aquest mètode s'inspira en les tècniques d'actualització de velocitat zero (*ZUPT*) utilitzades en la biomecànica clàssica.

El funcionament de la millora s'estructura de la següent manera:

1. **Detecció del Peak Frame:** La xarxa *LSTM* identifica el *frame* de màxima probabilitat d'una acrobàcia (per exemple, el punt més alt d'un salt *Split*).
2. **Càlcul de Desplaçament Espacial:** A partir d'aquest punt central, el sistema calcula la distància euclidiana Δd dels turmells entre frames consecutius.
3. **Propagació Bidireccional (Forward/Backward):** L'algorisme propaga l'etiqueta de l'acrobàcia cap enrere i cap endavant en el temps. Si el desplaçament Δd dels turmells és elevat, significa que la gimnasta es troba en la fase aèria i l'acrobàcia segueix activa. Si la distància cau per sota d'un llindar establert (indicant que els turmells s'han estabilitzat o la seva velocitat és propera a zero), el sistema interpreta que l'atleta ha contactat amb l'aparell.
4. **Delimitació Aïllada:** En l'instant del contacte, es talla la propagació de l'etiqueta de la *sliding window*.

Aquesta millora permet acotar de forma dinàmica la durada exacta de cada acrobàcia segons l'evolució física real de l'atleta, eliminant completament el soroll auditiu en les recepcions i evitant les penalitzacions artificials per impac-

6.8 Desenvolupament de l'Aplicació d'Escriptori i Experiència d'Usuari (UX)

A partir del prototip conceptual plantejat en fases anteriors, s'ha desenvolupat i consolidat una aplicació d'escriptori totalment funcional mitjançant el *framework PyQt6*. Donat que l'objectiu principal del sistema és assistir els jutges en un entorn on la càrrega cognitiva i la pressió temporal són elevades, el disseny de la interfície s'ha guiat estrictament per principis de *User Experience (UX)* i usabilitat eficient.

S'ha descartat la dependència exclusiva del ratolí en favor d'un flux de treball basat en dreceres de teclat (*shortcuts*), permetent a l'usuari mantenir la vista fixada en el monitor i auditar les rutines de forma ininterrompuda. L'aplicació actua com un copilot: la intel·ligència artificial localitza els *peak frames* i proposa les deduccions, però el jutge manté el control absolut mitjançant un sistema d'interacció ràpida.

7 RESULTATS I DISCUSSIÓ

Aquesta secció presenta la validació experimental del sistema, iniciant amb la comparativa de viabilitat dels models d'extracció i seguint amb l'avaluació quantitativa i qualitativa del classificador recurrent i el motor de puntuació biomecànica sobre rutines reals.

7.1 Resultats de l'Estudi de Viabilitat de Models HPE

Per determinar quina de les arquitectures d'estimació de pose era la més adient per a les exigències de la gimnàstica artística, es va dur a terme un estudi preliminar dividit en dues fases: una avaluació en condicions genèriques i una validació en el domini específic del projecte.

En la primera fase, s'ha avaluat el rendiment teòric dels models sobre una mostra d'un *dataset* bàsic exportat de *Roboflow*, amb anotacions sobre subjectes practicant *running*. S'ha seleccionat aquest conjunt per la linealitat de les postures, garantint un entorn neutral per mesurar la precisió base. Els resultats de la Taula 2 reflecteixen el rendiment en aquestes condicions estàndard.

TAULA 2: AVALUACIÓ PRELIMINAR SOBRE EL DATASET DE *running*

Model	mAP@0.5	mAP@0.75	PCK@0.2
YOLO26x-pose	67.7%	46.2%	88.8%
MediaPipe BlazePose	59.5%	39.3%	81.7%
Sapiens-2B	50.6%	34.5%	76.4%

No obstant això, el rendiment en entorns controlats no es tradueix directament en eficàcia dins d'un domini caracteritzat per l'alta velocitat i la complexitat postural. En la segona fase es va realitzar una prova empírica sobre un fragment de rutina real a la barra d'equilibris. Mitjançant CVAT[20], es va generar un *Ground Truth* manual dels 84 *frames* de la seqüència de prova (vegeu la Taula 3).

TAULA 3: AVALUACIÓ SOBRE FRAGMENT ACROBÀTIC EN BARRA D'EQUILIBRIS

Model	mAP@0.5	mAP@0.75	PCK@0.2
YOLO26x-pose	19.2%	9.3%	44.0%
MediaPipe BlazePose	13.4%	6.9%	35.4%
Sapiens-2B	36.6%	19.2%	72.7%

L'anàlisi comparativa justifica l'elecció final de *Sapiens-2B*. Mentre que *YOLO26x-pose* lidera el banc de proves genèric, pateix una caiguda severa al domini gimnàstic (el seu mAP@0.5 baixa fins al 19.2%). Això demostra que les oclusions i inversions penalitzen greument les arquitectures *Top-Down*. En contraposició, l'enfocament de *Vision Transformers* de *Sapiens-2B* demostra una comprensió estructural global superior, retenint un PCK@0.2 del 72.7% que garanteix la continuïtat de l'esquelet.

7.2 Rendiment del Classificador Recurrent (LSTM)

Per avaluar la capacitat de la xarxa *LSTM* per discriminar entre les quatre categories acrobàtiques (*Pike*, *Split*, *Straddle*, *Tuck*), es va separar el *dataset* final (144 mostres) en un 80% d'entrenament i un 20% de test independent. El rendiment del model s'ha mesurat mitjançant les mètriques clàssiques de Visió per Computador: Precisió (*Precision*), Sensibilitat (*Recall*) i la puntuació *F1-Score*.

A la Taula 4 es detalla el comportament del classificador. Els resultats demostren una alta fiabilitat general, especialment en les classes *Tuck* i *Pike*, on la morfologia de l'esquelet és molt diferent de la resta.

TAULA 4: MÈTRQUES DE RENDIMENT DE L'ARQUITECTURA LSTM PER CLASSE

Categoria	Precision	Recall	F1-Score
<i>Pike</i>	100.0%	100.0%	100.0%
<i>Split</i>	75.0%	33.0%	46.0%
<i>Straddle</i>	33.0%	75.0%	46.0%
<i>Tuck</i>	100.0%	100.0%	100.0%
Mitjana Ponderada	83.0%	76.0%	76.0%

Per analitzar on s'equivoca el model, s'ha generat la matriu de confusió (vegeu la Figura 9). Els indicadors numèrics corroboren de manera empírica la hipòtesi plantejada a la metodologia: el principal repte de la xarxa recau en la confusió entre les classes *Split* i *Straddle* (amb un *F1-Score* simètric del 46.0%). Aquest comportament és biomecànicament lògic, ja que ambdues acrobàcies impliquen una obertura de cames propera als 180° i el model basat en vídeo monocular pateix per distingir la direcció de l'obertura a causa de la pèrdua de profunditat a la perspectiva lateral.

Com a millora futura, la incorporació d'una característica addicional d'orientació (*orientation*), que complementi les mètriques actuals de posició, velocitat, acceleració i angles dins del vector d'entrada de la *LSTM*, ajudaria a solucionar aquest problema, aportant al classificador el context espacial necessari per distingir amb precisió el pla del salt.

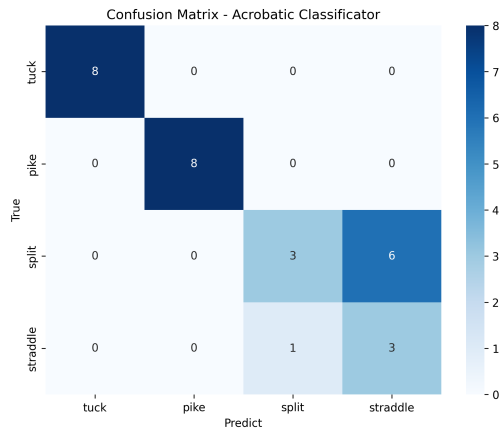


Fig. 9: Matriu de confusió del model *LSTM* sobre el conjunt de test. Es manifesta la lleugera confusió entre les classes d'obertura aèria.

7.3 Interfície d'Assistència a l'Arbitratge i Validació Gràfica

Una vegada validada la classificació i el motor de lògica determinista, s'ha integrat tot el sistema en una aplicació d'escriptori final, dissenyada sota el paradigma *Human-in-the-Loop*.

Per garantir una avaluació fluida i reduir el temps de reacció ergonòmic del jutge, la interfície s'ha dotat d'un sistema de dreceres de teclat (*shortcuts*) orientades a la navegació ràpida entre els punts crítics detectats pel model. Aquestes accions es detallen a la Taula 5.

TAULA 5: DRECERES DE TECLAT PER A LA INTERFÍCIE D'ASSISTÈNCIA A L'ARBITRATGE

Drecera de teclat	Acció assignada
← (Fletxa esq.)	Retrocedir al <i>frame</i> anterior
→ (Fletxa dreta)	Avançar al <i>frame</i> següent
Espai	Reproduir o pausar la seqüència
Z	Saltar a la deducció anterior
X	Saltar a la deducció següent
Ctrl + Z	Desfer l'última acció del jutge
Ctrl + Y	Refer l'última acció desfeta
Ctrl + O	Obrir un nou arxiu d'avaluació
Ctrl + S	Exportar l'informe en format .pdf

L'experimentació visual sobre la interfície amb seqüències reals d'alta velocitat ha permès identificar dues tipologies clares de resultats.

7.3.1 Casos d'Èxit (Funcionament Òptim i Captura del *Peak Frame*)

Les Figures 10 i 11 mostren el comportament ideal de l'aplicació en acrobàcies aïllades. El sistema localitza el *peak frame*, renderitza l'esquelet net i il·lumina en vermell les línies de control que superen els llindars de penalització, facilitant la validació immediata del jutge.

Analitzant les tipologies (Figura 10), el sistema detecta de forma impecable les classes *Pike* i *Tuck*. No obstant

això, la pèrdua de profunditat de la càmera 2D genera confusió en les obertures aèries (Figura 11), arribant a classificar un *Straddle* com a *Split* amb baixa confiança (52%). Malgrat aquest error taxonòmic, la fiabilitat de l'auditoria es manté intacta: ambdues acrobàcies comparteixen la mateixa exigència geomètrica a l'E-Score (obertura de 180°), pel que el càlcul de deduccions no s'altera.

Finalment, per evitar falsos positius en avaluar rutines contínues, es va enriquir el *dataset* de la xarxa *LSTM* afegint-hi seqüències de transició. Això va dotar el classificador del context temporal necessari per discriminar directament els moviments preparatoris d'impuls respecte a l'acrobàcia principal.

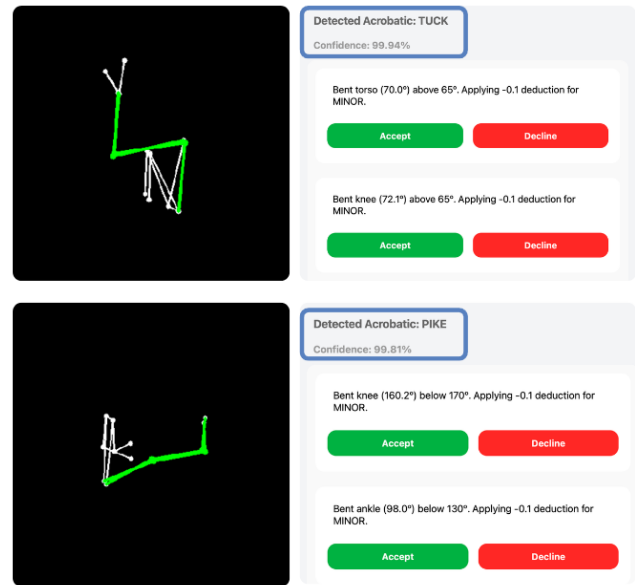


Fig. 10: Visualització a la interfície de salts amb gran diferenciació morfològica. El sistema assoleix un èxit rotund en la classificació i validació geomètrica de les classes *Pike* i *Tuck*.



Fig. 11: Avaluació d'obertures aèries. S'observa l'èxit en la classe *Split* i la confusió en intentar avaluar un *Straddle* amb una baixa confiança (52%).

7.3.2 Limitacions Temporals per Fals Positiu de Transició

Per contra, la Figura 12 mostra un error de detecció causat inicialment per limitacions de segmentació temporal. En aquest escenari, la xarxa *LSTM* genera un fals positiu en classificar un fragment de transició (com l'atterratge de la gimnasta) com si fos una classe acrobàtica activa.

És important matisar que el motor geomètric actua de forma analíticament correcta: calcula els vectors sobre l'esquelet rebut i, en detectar angles completament tancats per l'impacte amb l'aparell, aplica coherentment una deducció de caràcter *Severe* (-0.5). Per tant, la fallada no recau en la lògica matemàtica de les regles, sinó en una classificació errònia d'origen temporal.

Per abordar aquesta problemàtica sense recórrer a la creació d'una nova classe de "Transició", una decisió que hauria requerit la complexitat de reentrenar completament la xarxa *LSTM*, es va treure profit del disseny *Human-in-the-Loop* de l'aplicació. Com que la interfície està concebuda com una ajuda sota el control absolut del jutge, es va implementar l'opció d'invalidar l'avaluació amb un botó de "Fals Positiu", el qual reescriu la classe directament amb "Transició", i descarta totes les possibles penalitzacions.

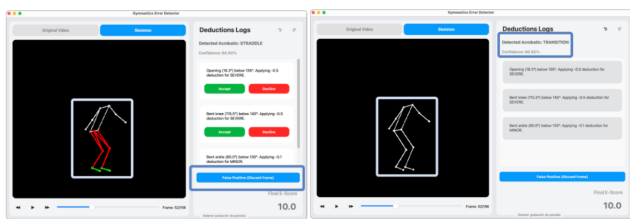


Fig. 12: Visualització a la interfície d'un fals positiu. S'observa l'aplicació d'una penalització severa provocada per una activació errònia durant la fase de recepció.

7.4 Eficiència Computacional

A nivell de rendiment, processar una rutina completa de longitud estàndard (aproximadament 1 minut i 20 segons, equivalent a 2.400 *frames*) amb una GPU *NVIDIA GeForce RTX 2080 Ti* requereix un temps de còmput d'entre 3 i 4 hores. Aquest cost evidencia el compromís inherent entre precisió i velocitat: la complexitat d'utilitzar un model fundacional pesat com *Sapiens-2B* genera un *bottleneck* temporal, però a canvi garanteix la robustesa extrema davant d'oclusions descrita a la Taula 3.

Malgrat aquesta latència, l'execució *offline* compleix amb l'objectiu fonamental del projecte: oferir una solució econòmica i assequible per a l'esport base, evitant infraestructures de maquinari inassolibles com les utilitzades a l'elit esportiva. Com a línies de treball futur per reduir aquest temps de processament, es planteja la migració del sistema cap a GPUs d'última generació (com la sèrie *NVIDIA RTX 50*).

8 CONCLUSIONS

A partir del desenvolupament i les proves de validació realitzades fins a la data, s'ha demostrat la viabilitat tècnica del sistema. Les principals conclusions es resumeixen en els següents punts:

- La integració de *Sapiens-2B* i la xarxa recurrent *LSTM* conforma una arquitectura robusta per a l'anàlisi cinemàtica. Tot i la limitació del model en discriminar algunes obertures aèries (46.0% d'*F1-Score* en *Split* i *Straddle*), aquesta confusió no afecta el funcionament global de l'aplicació. Aquest comportament recau principalment en no haver inclòs l'orientació espacial (*orientation*) de la gimnasta com a paràmetre d'entrada.
- L'èxit operatiu recau en quatre pilars clau: la normalització al tors i el *KNNImputer* per garantir un senyal espacial temporalment robust; la *Pelvis Virtual* per avaluar angles invariants a la perspectiva; i la *Sliding Window* per aïllar dinàmicament els exercicis en rutines contínues.
- Tot i el cost computacional inherent a l'ús de models fundacionals pesats (requerint processament *offline*), el sistema assoleix l'objectiu de democratitzar l'assistència arbitral, oferint una eina viable i de baix cost per a l'esport base davant les inassolibles infraestructures de l'elit.
- S'ha desenvolupat amb èxit una aplicació interactiva sota el paradigma *Human-in-the-Loop*. La intel·ligència artificial actua únicament com a assistent biomecànic per proposar deduccions de l'*E-Score*, delegant sempre la validació i el veredict final al criteri expert del jutge humà.

AGRAÏMENTS

Vull expressar el meu agraïment més sincer al meu tutor, Coen Antens, per creure en aquest projecte des del primer dia i pel seu guiatge acadèmic i tècnic.

També dono les gràcies al CVC i al personal de la Unitat d'Innovació per proporcionar-me durant la meua estada de pràctiques un bon ambient i facilitar-me els recursos computacionals necessaris per dur a terme la inferència del model *Sapiens-2B*.

Finalment, agraeixo a la meua família, parella i amics propers per la seva paciència i el seu suport incondicional durant les in comptables hores de recerca i programació.

ÚS DE LA IA GENERATIVA

En compliment de la normativa d'avaluació del Treball Final de Grau de la UAB, es declara que en l'elaboració d'aquest Informe Final s'han utilitzat eines d'Intel·ligència Artificial generativa (model *Gemini 3.1 Pro*).

D'acord amb les directrius establertes, el seu ús s'ha limitat exclusivament a l'assistència en aspectes formals i tècnics. Concretament, l'eina s'ha emprat per a la revisió ortotipogràfica, la millora de la fluïdesa del text en català, i com a suport tècnic per a la resolució de problemes de formatació i estructuració del codi en llenguatge *LaTeX*.

Es certifica que no s'ha utilitzat cap model generatiu per a la creació dels continguts intel·lectuals sotmesos a avaluació. La definició del context, l'establiment dels objectius, el disseny metodològic, la programació i la planificació tem-

poral del projecte han estat elaborats de manera íntegrament autònoma per l'autora.

REFERÈNCIES

- [1] La Vanguardia, *Una gimnasta rumana recibe la medalla de bronce once días después de la final de los JJOO*, Agost 2024. [En línia]. Disponible a: <https://www.lavanguardia.com/deportes/20240817/9874948/gimnasta-recibe-medalla-bronce-once-dias-final-jjoo.html>
- [2] International Olympic Committee (IOC), *Artistic gymnastics scoring system and judging rules: The D-score and E-score explained*. [En línia]. Disponible a: <https://olympics.com/en/news/artistic-gymnastics-scoring-system-judging-rules-e-d-score>
- [3] Fujitsu, *Judging Support System Co-development with FIG*. [En línia]. Disponible a: <https://www.fujitsu.com/global/themes/data-drive/n/judging-support-system/>
- [4] Ultralytics, *YOLOv8 Pose Task Documentation*. [En línia]. Disponible a: <https://docs.ultralytics.com/tasks/pose/#predict>
- [5] T. Rawal *et al.* (Meta), *Sapiens: Foundation for Human Vision Models*, arXiv preprint arXiv:2408.12569, 2024. Disponible a: <https://arxiv.org/abs/2408.12569>
- [6] Google Research, *MediaPipe: A Framework for Building Perception Pipelines*. Disponible a: <https://arxiv.org/pdf/1906.08172>
- [7] V. Bazarevsky *et al.* (Google), *BlazePose: On-device Real-time Body Pose tracking*, CVPR Workshop on Computer Vision for Augmented and Virtual Reality, 2020. Disponible a: <https://arxiv.org/abs/2006.10204>
- [8] Fédération Internationale de Gymnastique (FIG), *2025-2028 Code of Points (MAG & WAG)*. [En línia]. Disponible a: <https://www.gymnastics.sport/site/rules/>
- [9] X. Li, Y. Liu, N. Dong, S. Qin, and X. Hu, *PartImageNet++ Dataset: Scaling up Part-based Models for Robust Recognition*, en European Conference on Computer Vision (ECCV), 2024. Disponible a: <https://arxiv.org/abs/2407.10918>
- [10] C. Zheng, W. Wu, C. Chen, T. Yang, S. Zhu, J. Shen, N. Kehtarnavaz, and M. Shah, *Deep Learning-Based Human Pose Estimation: A Survey*, ACM Computing Surveys, vol. 56, no. 1, 2023. Disponible a: <https://arxiv.org/pdf/2012.13392>
- [11] A. Toshev and C. Szegedy, *DeepPose: Human pose estimation via deep neural networks*, en Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1653-1660, 2014. Disponible a: <https://arxiv.org/abs/1312.4659>
- [12] T.Y. Lin *et al.*, *Microsoft COCO: Common Objects in Context*, en European Conference on Computer Vision (ECCV), pp. 740-755, 2014. Disponible a: <https://arxiv.org/abs/1405.0312>
- [13] Microsoft Research, *Deep High-Resolution Representation Learning for Human Pose Estimation*. Disponible a: <https://arxiv.org/pdf/1902.09212>
- [14] Hao-Shu Fang, *AlphaPose: Whole-Body Regional Multi-Person Pose Estimation and Tracking in Real-Time*. Disponible a: <https://arxiv.org/pdf/2211.03375>
- [15] ViTAE-Transformer, *ViTPose: Simple Vision Transformer Baselines for Human Pose Estimation*. GitHub. [En línia]. Disponible a: <https://github.com/ViTAE-Transformer/ViTPose>
- [16] D. Maji, S. Nagori, M. Mathew, and D. Poddar, *YOLO-Pose: Enhancing YOLO for Multi Person Pose Estimation Using Object Keypoint Similarity Loss*, en IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2022. Disponible a: <https://arxiv.org/abs/2204.06806>
- [17] Y. Yang and D. Ramanan, *Articulated Human Detection with Flexible Mixtures of Parts*, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 35, no. 12, pp. 2878-2890, 2013. Disponible a: https://www.cs.cmu.edu/~deva/papers/pose_pami.pdf
- [18] Paul W. Oxby, *An Optimal Weighting Function for the Savitzky-Golay Filter*. Disponible a: <https://arxiv.org/pdf/2111.11667>
- [19] C. B. Vennerød, A. Kjærø, and E. S. Bugge, *Long Short-term Memory RNN*, arXiv preprint arXiv:2105.06756, 2021. Disponible a: <https://arxiv.org/pdf/2105.06756>
- [20] CVAT - Open Data Annotation Platform. [En línia]. Disponible a: <https://www.cvat.ai/>
- [21] USA Gymnastics, *USA Gymnastics Official YouTube Channel - Balance Beam Search results*. [En línia]. Disponible a: <https://www.youtube.com/@usagymnastics/search?query=balance%20beam>
- [22] D. Shao *et al.*, *FineGym: A Hierarchical Video Dataset for Fine-grained Action Understanding*. Disponible a: <https://arxiv.org/pdf/2004.06704>

A DIAGRAMA D'ARQUITECTURA MODULAR

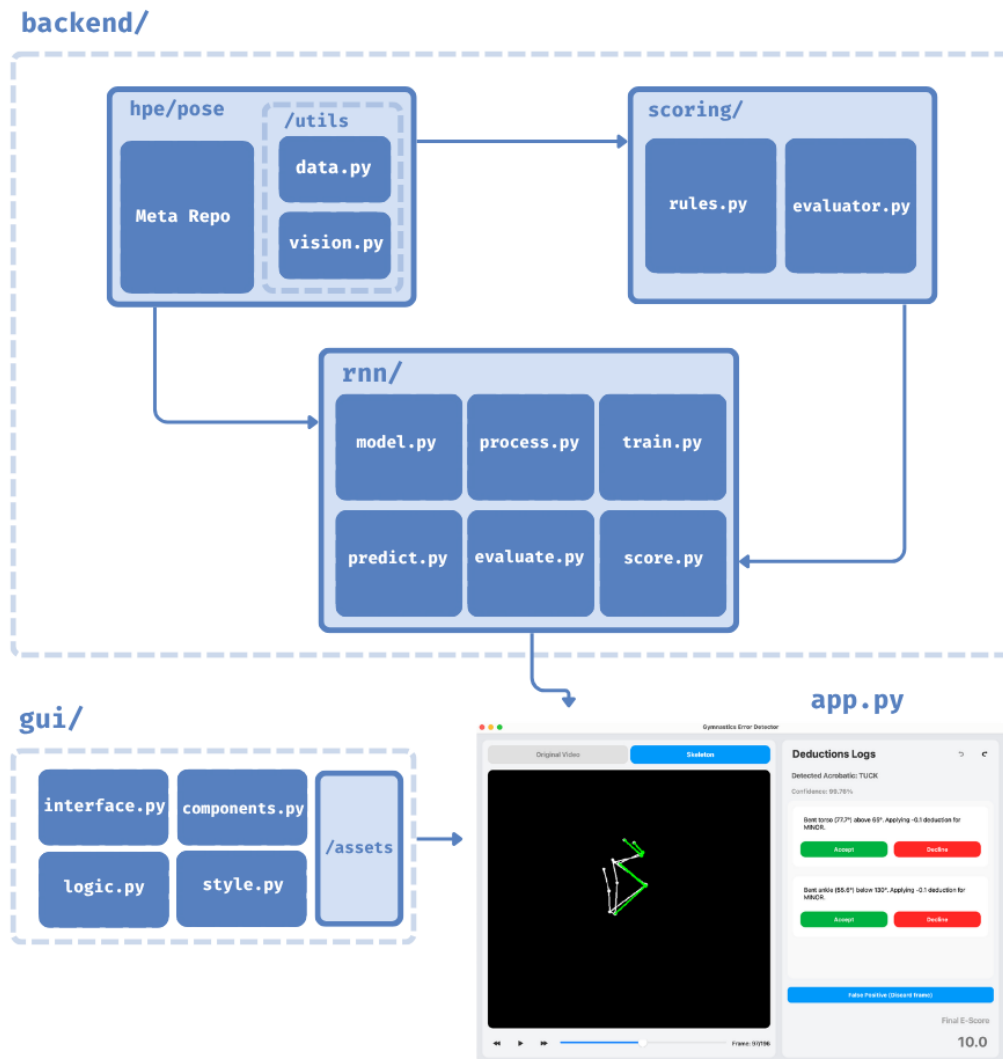


Fig. 13: Diagrama complet d'arquitectura de programari i interacció de mòduls del sistema (HPE, RNN, Scoring i GUI). Elaboració pròpia mitjançant *Canva*.

B CRONOGRAMA DE TASQUES I DIAGRAMA DE GANTT

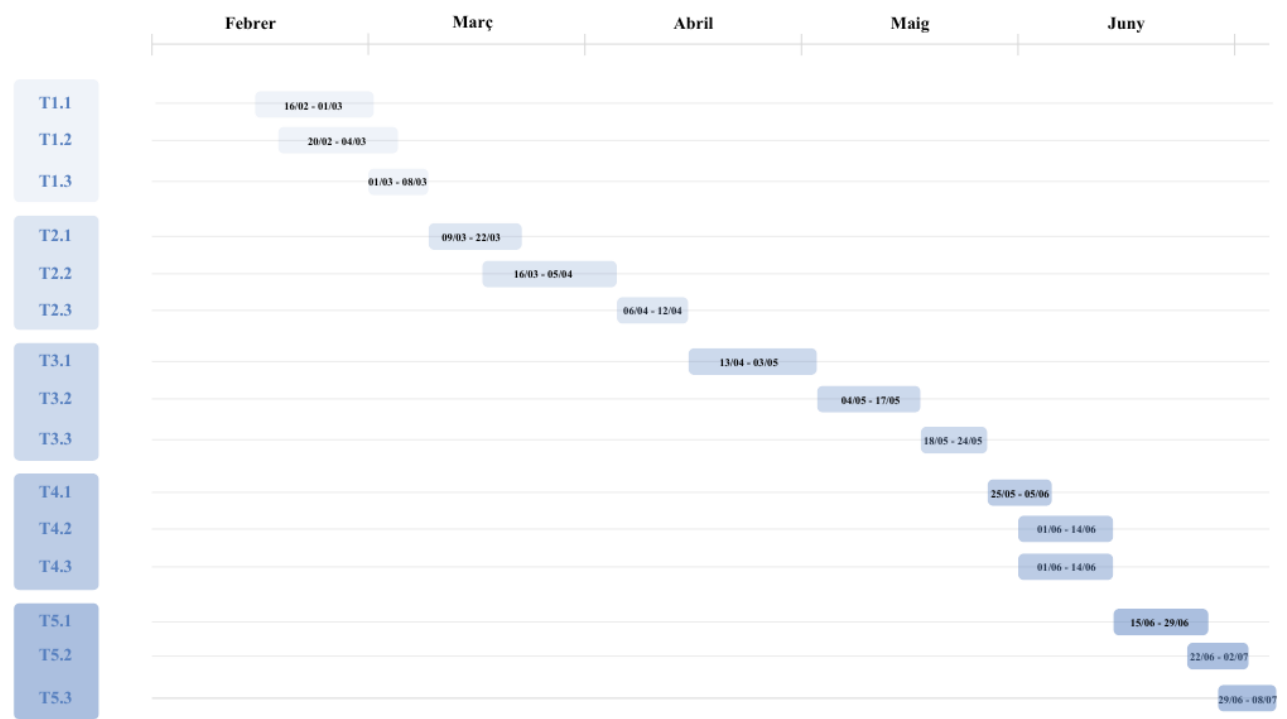


Fig. 14: Cronograma del projecte (Diagrama de Gantt). Elaboració pròpia mitjançant *Canva* a partir de la planificació feta a *Notion*.

TAULA 6: CRONOGRAMA DE TASQUES I FITES D’AVALUACIÓ ACTUALITZAT

Id		Tasca i Fites	Dates
Fase 1: Recerca i Viabilitat			
T1.1		Revisió de l'Estat de l' Art	16/02 - 01/03
T1.2		Estudi de l'abast i domini	20/02 - 04/03
T1.3		Informe Inicial (Fita: 08/03)	01/03 - 08/03
Fase 2: Extracció i Processament			
T2.1		Avaluació YOLO, MediaPipe, Sapiens	09/03 - 22/03
T2.2		Integració Sapiens i Proc. Senyal	16/03 - 05/04
T2.3		Informe Progrés I (Fita: 12/04)	06/04 - 12/04
Fase 3: Classificació i Lògica			
T3.1		Processament de dades i entrenament RNN	13/04 - 03/05
T3.2		Aplicació de la lògica de puntuació (FIG)	04/05 - 17/05
T3.3		Informe Progrés II (Fita: 24/05)	18/05 - 24/05
Fase 4: Interfície i Integració			
T4.1		Generació de l'output visual final (GUI)	25/05 - 05/06
T4.2		Integració de a l'aplicació	01/06 - 14/06
T4.3		Proposta Informe Final (Fita: 14/06)	01/06 - 14/06
Fase 5: Tancament			
T5.1		Dossier TFG (Fita: 29/06)	15/06 - 29/06
T5.2		Pòster (Fita: 02/07)	22/06 - 02/07
T5.3		Defensa (Fita: 02/07 - 08/07)	29/06 - 08/07

C DESKTOP APPLICATION

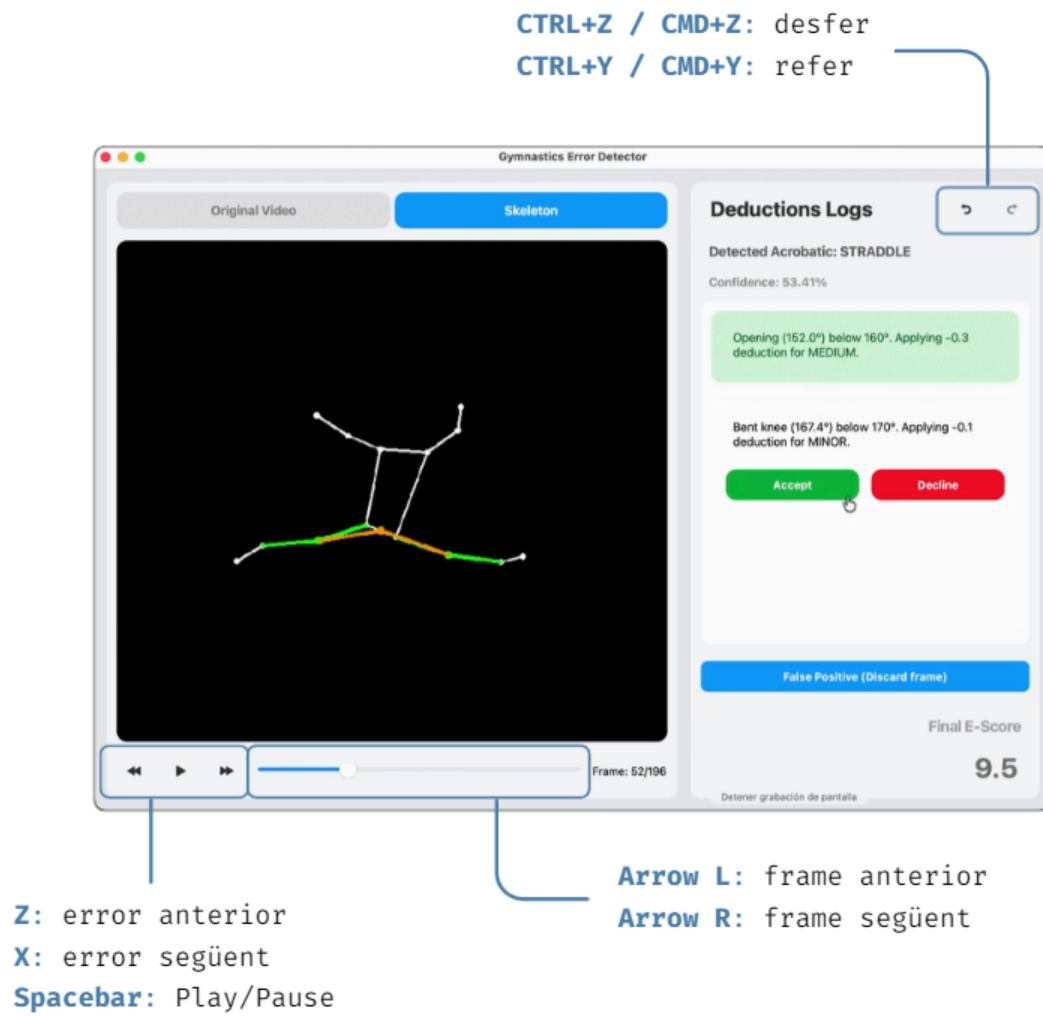


Fig. 15: Interfície gràfica funcional amb el llistat de dreceres de teclat (*shortcuts*).

D SLIDING WINDOW

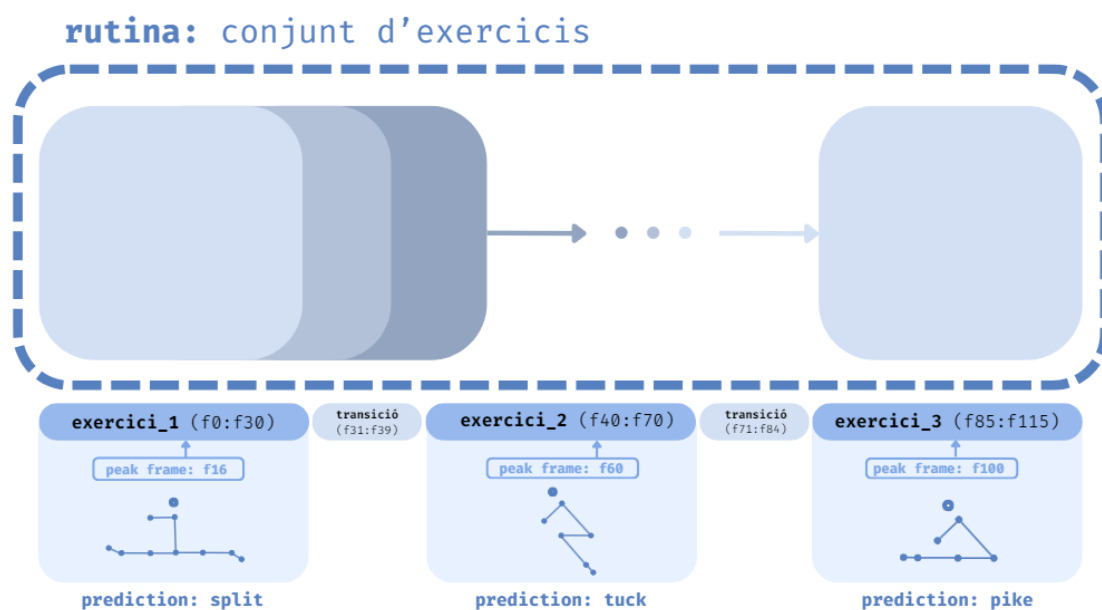


Fig. 16: Exemplificació del funcionament temporal de la *sliding window*, on cada quadrat representa cada *frame* de la seqüència. L'algorisme llisca per la rutina contínua fins a identificar el *peak frame* de cada salt. A partir d'aquest, s'avalua la trajectòria dels turmells per acotar de forma dinàmica l'inici i el final de cada exercici, aïllant-los i descartant les fases de transició.