

UNIVERSITÀ
DI PAVIA

Università degli Studi di Pavia

Facoltà di Ingegneria

Dipartimento di Ingegneria Industriale e dell'Informazione

Corso di Laurea Magistrale in Bioingegneria

Tesi di Laurea Magistrale

Validazione di sistemi markerless di stima della posa per analisi stabilometrica

Candidato/a:
Nicolas Francesco Trombetta
Matricola: 557388

Relatore:
Prof. Nicola Giulietti

Correlatore:
Dott. Davide Todesca

Anno accademico 2025/2026

E quindi uscimmo a riveder le stelle.

Per correr miglior acque alza le vele
omai la navicella del mio ingegno,
che lascia dietro a sé mar sì crudele

Dante Alighieri, *Divina Commedia*
Inferno, XXXIV, 139; *Purgatorio*, I, 1–3

Sommario

La stabilometria valuta il controllo posturale attraverso il centro di pressione (CoP) misurato da una pedana di forza, che però non descrive direttamente il movimento del corpo. I sistemi di motion capture markerless, basati su modelli di stima della posa, permettono di ricostruire il centro di massa (CoM), ma in stazione eretta le oscillazioni del corpo sono dell'ordine dei millimetri, confrontabili con l'errore di tracciamento, e la stima del CoP richiede una doppia derivazione che amplifica il rumore.

Questa tesi ha confrontato con la pedana di forza due sistemi markerless che stimano il CoP a partire dal CoM mediante il modello del pendolo inverso: un sistema multi-camera (YOLO), basato su quattro telecamere sincronizzate, sul modello YOLO26x e su un modello geometrico a solidi del corpo, e un sistema single-camera (SAM), basato su SAM 3D Body. Sedici volontari sani hanno eseguito cinque prove da 30 s (*Occhi Aperti*, *Occhi Chiusi*, test dei limiti di stabilità, *One Foot* e *Tandem*); escluse le sessioni con frame mancanti, ne sono state analizzate 54 su 80. Il CoM è stato filtrato con un passa-basso di Butterworth, con frequenza di taglio scelta confrontando tre criteri (1.0 Hz per YOLO e 1.1 Hz per SAM). L'accordo con la pedana è stato valutato tramite gli errori sugli indici posturografici, le correlazioni, l'RMSE di traiettoria e l'analisi di Bland-Altman.

Il CoP stimato riproduce l'andamento di quello della pedana, con correlazioni per lo più comprese tra 0.7 e 0.9. Come atteso, l'accordo dipende dall'ampiezza dell'oscillazione rispetto al rumore di tracciamento: sull'asse mediolaterale nel test dei limiti di stabilità la correlazione è quasi perfetta (0.999) e gli errori su Range e RMS restano sotto il 13%, mentre in stazione eretta quieta l'errore percentuale può risultare elevato anche quando la differenza assoluta è di circa 1 mm. YOLO ha l'errore percentuale mediano minore in 19 combinazioni indice-esercizio su 30 e l'RMSE di traiettoria minore in 7 su 10, mentre le correlazioni sono equamente divise tra i due sistemi. Gli errori di SAM sono riconducibili alla stima monoculare: un errore di scala lungo l'asse mediolaterale e, lungo l'asse antero-posteriore, che per la telecamera frontale coincide con la profondità, un'oscillazione che la pedana non registra durante l'inclinazione laterale del test dei limiti di stabilità (correlazione mediana 0.16), effetto a cui la triangolazione di YOLO

non è soggetta.

La stima del CoP da video, con quattro telecamere e con una sola, è quindi fattibile e riproduce l'andamento della pedana, soprattutto per oscillazioni ampie, e il sistema multi-camera ne stima meglio l'ampiezza. Allo stato attuale, nella stazione eretta quieta, non consente di sostituire la pedana nella quantificazione precisa degli indici. Sono necessari un campione più ampio, prove più lunghe e, per SAM, vincoli di scala e di profondità.

Indice

Sommario **iii**

1 Introduzione e stato dell'arte **1**
1.1 Il Ruolo della Biomeccanica e della Motion Capture Clinica 1
1.2 Limiti dei Sistemi Optoelettronici 2
1.3 Motion Capture Markerless e Human Pose Estimation 4
1.3.1 Algoritmi di Human Pose Estimation (HPE) 4
1.3.2 Acquisizione Monoculare 9
1.3.3 Sistemi multi-camera e Ricostruzione 3D (5DoF vs 6DoF) 9
1.4 Stima delle Proprietà Inerziali Corporee e Modelli Biomeccanici 11
1.5 Controllo Posturale e Stabilometria 14
1.5.1 Accoppiamento CoM-CoP e Modello a Pendolo Inverso 16
1.6 Obiettivi dello Studio e Struttura della Tesi 17

2 Materiali e metodi **19**
2.1 Pedana di Forza 19
2.1.1 Trasduttore F/T: Axia130-M300 19
2.2 Videocamera IDS U3-3160CP-C-HQ Rev. 2.2 e Ottica 21
2.3 Sistema di Acquisizione 24
2.3.1 Workstation e sincronizzazione 25
2.4 Modelli 26
2.4.1 Razionale Biomeccanico: Il Modello a Pendolo Inverso 26
2.4.2 Modello e Discretizzazione Geometrica per il Sistema multi-camera 27
2.4.3 SAM 3D Body (3DB) per il sistema single-camera 31
2.5 Protocollo 32
2.5.1 Partecipanti 32
2.5.2 Informazioni Antropometriche 33
2.5.3 Protocollo di Acquisizione (Task Sperimentali) 34
2.5.4 Sessioni e Grandezze Confrontate 38
2.6 Elaborazione Dati 40
2.6.1 Calcolo del Centro di Pressione di Riferimento 40
2.6.2 Stima del Centro di Pressione dal Centro di Massa 40
2.6.3 Analisi del Segnale 42
2.6.4 Indici Posturografici 52
2.6.5 Calcolo degli Errori e Aggregazione tra Soggetti 53

2.6.6 Validazione Statistica (Markerless vs Pedana) 54

3 Risultati **57**
3.1 Valori degli indici posturografici 57
3.2 Errori rispetto alla pedana 60
3.3 Correlazioni ed errore di forma della traiettoria 64
3.4 Stima della profondità: SAM sull'asse AP nel *Limite* 67
3.5 Analisi di accordo (Bland-Altman) 68
3.6 Sintesi del confronto tra i due sistemi 74

4 Conclusioni **77**
4.1 Sintesi del Lavoro 77
4.2 Risultati rispetto agli Obiettivi 77
4.2.1 Protocollo 77
4.2.2 Stima del CoM e del CoP 78
4.2.3 Confronto con la Pedana 78
4.2.4 Accordo tra i Sistemi 78
4.3 Interpretazione dei Risultati 79
4.4 Limiti dello Studio 80
4.5 Sviluppi Futuri 81
4.6 Considerazioni Finali 82

Bibliografia **83**

Ringraziamenti **91**

Elenco delle figure

1.1 Marcatori ottici di un sistema marker-based 3
1.2 Workflow di addestramento di un algoritmo di pose estimation 5
1.3 Paradigmi di Human Pose Estimation 6
1.4 Keypoint rilevati dall'algoritmo OpenPose 7
1.5 Rappresentazione dei sei gradi di libertà (6DoF) 10
1.6 Limite dei 5DoF e soluzione a 6DoF 11
1.7 Modello del corpo a solidi geometrici 13
1.8 Schema concettuale del modello SMPL 14
1.9 Margine di stabilità dinamica: CoM, XCoM e BoS 15

2.1 Vista esplosa della pedana di forza 19
2.2 Rappresentazione 3D del Trasduttore F/T: Axia130-M300 20
2.3 Immagine della videocamera IDS con l'ottica montata 22
2.4 Schema del setup di acquisizione 24
2.5 Vista della pedana di forza all'interno del setup sperimentale 25
2.6 Vista delle 4 videocamere all'interno del setup sperimentale 25
2.7 Rappresentazione del momento raddrizzante nel modello a Pendolo Inverso. Per frenare la caduta in avanti, il CoP deve spingersi fisicamente oltre la proiezione a terra del CoM, generando una forza di richiamo verso l'indietro. 26
2.8 Pipeline del sistema di misura markerless 27
2.9 Riduzione dei keypoints da 17 a 13 29
2.10 Discretizzazione del corpo in 10 segmenti 30
2.11 Esecuzione dei task di stazione eretta quieta sulla pedana di forza: a *Occhi Aperti* (a) e a *Occhi Chiusi* (b). 35
2.12 Esecuzione del *Limite* (esempio di sbilanciamento laterale): il soggetto si inclina mantenendo i piedi fermi e il corpo rigido. 36
2.13 Esecuzione dei task a base di appoggio ridotta: *One Foot* (a) e *Tandem* (b). 38
2.14 Segnale grezzo nel dominio del tempo (sessione S2, *Occhi Aperti*). A sinistra, le componenti $X(t)$ (sopra) e $Y(t)$ (sotto); a destra, la traiettoria (X, Y). Sono riportati il CoP della pedana (nero), il CoM di YOLO (blu) e il CoM di SAM (rosso), centrati sulla propria media. La zona grigia indica il primo secondo, scartato nelle analisi. 43

2.15 Spettro di potenza (periodogramma, a sinistra) e PSD di Welch con segmenti da 5 s (a destra) per l'asse X (sopra) e l'asse Y (sotto), per il CoP della pedana, il CoM di YOLO e il CoM di SAM (sessione S2, *Occhi Aperti*). La potenza è in scala logaritmica. 44
2.16 Potenza cumulata in funzione della frequenza per l'asse X (a sinistra) e per l'asse Y (a destra), per il CoP della pedana, il CoM di YOLO e il CoM di SAM (sessione S2, *Occhi Aperti*). Le linee verticali tratteggiate indicano f_{70} di ciascun segnale; la linea orizzontale punteggiata indica il 70%. 45
2.17 Analisi dei residui di Winter. A sinistra, per la sessione S2 (*Occhi Aperti*), residuo RMS in funzione della frequenza di taglio per YOLO e SAM e per gli assi X e Y, con la retta del rumore (tratteggiata), il taglio di Winter (verde) e il taglio adottato (nero). A destra, distribuzione (*box plot*) del taglio di Winter su tutte le sessioni; le linee orizzontali indicano i tagli adottati. 46
2.18 Coerenza tra il CoM (YOLO e SAM) e il CoP della pedana in funzione della frequenza: mediana e intervallo 25–75% su tutte le sessioni e su entrambi gli assi. La linea punteggiata indica la soglia di significatività al 95%; le linee verticali indicano i tagli adottati. 47
2.19 Errore degli indici rispetto alla pedana in funzione della frequenza di taglio, mediana su tutte le sessioni. A sinistra, valore assoluto medio dell'errore percentuale di Sway Area e Path Length (banda: 25–75%; le stelle indicano i minimi). A destra, errore percentuale con segno di Sway Area (linea continua) e Path Length (linea tratteggiata). Le linee verticali indicano i tagli adottati. 49
2.20 Segnale nel dominio del tempo dopo il filtraggio del CoM (YOLO a 1.0 Hz, SAM a 1.1 Hz), con la stessa configurazione della Figura 2.14. A sinistra $X(t)$ e $Y(t)$, a destra la traiettoria (X, Y); il CoP della pedana (nero) non è filtrato. Sessione S2, *Occhi Aperti*; il primo secondo (zona grigia) è scartato. 50
2.21 YOLO: posizione (sopra), velocità (al centro) e accelerazione (sotto) del CoM lungo X (a sinistra) e Y (a destra), grezze (grigio) e filtrate a 1.0 Hz (blu). Nei riquadri, l'RMS dell'accelerazione prima e dopo il filtraggio. Sessione S2, *Occhi Aperti*. 51
2.22 SAM: posizione (sopra), velocità (al centro) e accelerazione (sotto) del CoM lungo X (a sinistra) e Y (a destra), grezze (grigio) e filtrate a 1.1 Hz (rosso). Nei riquadri, l'RMS dell'accelerazione prima e dopo il filtraggio. Sessione S2, *Occhi Aperti*. 52

3.1 Distribuzione completa dell'errore percentuale di Sway Area e Path Length (stessi dati della Tabella 3.3): un punto per sessione, box con mediana e scarto interquartile, per esercizio e per sistema. I punti oltre i baffi (1.5 × IQR) sono etichettati con l'ID del soggetto. 61
3.2 Distribuzione completa dell'errore percentuale di Range e RMS, ML e AP (stessi dati della Tabella 3.4), stessa configurazione della Figura 3.1. . 63
3.3 Distribuzione completa di Pearson, Spearman (ML e AP) e RMSE di traiettoria (corrispondente alla Tabella 3.5), stessa configurazione della Figura 3.1. La linea tratteggiata indica la soglia di correlazione 0.5. 65
3.4 Bland-Altman per Sway Area: media e differenza (sistema − pedana) tra YOLO e pedana (sinistra) e tra SAM e pedana (destra), colorati per esercizio. La linea continua indica il bias, le linee tratteggiate i limiti di accordo al 95%. 70
3.5 Bland-Altman per Path Length, stessa configurazione della Figura 3.4. . 70
3.6 Bland-Altman per Range ML, stessa configurazione della Figura 3.4. . . . 71
3.7 Bland-Altman per Range AP, stessa configurazione della Figura 3.4. . . . 72
3.8 Bland-Altman per RMS ML, stessa configurazione della Figura 3.4. . . . 72
3.9 Bland-Altman per RMS AP, stessa configurazione della Figura 3.4. 73

Capitolo 1

Introduzione e stato dell'arte

1.1 Il Ruolo della Biomeccanica e della Motion Capture Clinica

L'analisi del movimento, in ambito clinico e riabilitativo, viene principalmente utilizzata per studiare i meccanismi neuromuscolari e biomeccanici attraverso il riconoscimento di schemi motori [1]. Individuare piccole alterazioni cinematiche può contribuire alla prevenzione dei danni causati da disturbi muscoloscheletrici e al monitoraggio di patologie neurologiche, come il morbo di Parkinson [2], o gli esiti di un ictus [3]. Tra gli strumenti clinici più diffusi per valutare la mobilità e l'equilibrio vi sono il *Timed Up and Go* e la *Berg Balance Scale*. Il primo valuta l'equilibrio dinamico, mobilità funzionale e rischio di caduta: il paziente, partendo da seduto, si alza, cammina 3 metri, si gira, torna alla sedia e si risiede, e il punteggio dipende dal tempo impiegato [4, 5]. Il secondo, oltre all'equilibrio dinamico e al rischio di caduta, valuta anche l'equilibrio statico, attraverso 14 compiti funzionali della vita quotidiana (e.g. alzarsi da una sedia, stare in piedi a occhi chiusi, reggersi su una gamba), con un punteggio basato sull'autonomia e sulla sua sicurezza nell'esecuzione [6]. La semplicità di questi strumenti li rende ottimi per una valutazione rapida del soggetto ma insufficienti per uno studio nel dettaglio del movimento e della sua complessità [7]. Inoltre, il punteggio della *Berg Balance Scale* può non rilevare cambiamenti modesti dell'equilibrio e, nei pazienti con deficit lievi, può raggiungere facilmente il valore massimo (effetto soffitto) [8, 9]. Con l'esigenza di affiancare alla valutazione clinica misure quantitative, vengono impiegate diverse tecnologie: sensori inerziali (IMU) per analizzare il movimento di segmenti corporei [10, 11], o solette a pressione [12] e le pedane di forza [13] per misurare le forze di reazione al suolo e ricavare il centro di pressione (CoP). L'analisi delle sue oscillazioni durante il man-

tenimento della postura eretta costituisce la base della valutazione stabilometrica [14, 15].

Anche questi strumenti presentano limiti, che devono essere considerati in funzione della tecnologia utilizzata. Nel caso delle IMU, ad esempio, la calibrazione e la deriva influiscono sulla stima del movimento; per i dispositivi indossabili occorre considerare anche il posizionamento e il metodo con cui essi vengono fissati al corpo [1, 11, 16].

Le pedane di forza, pur essendo il riferimento per la misura del CoP in posturografia, presentano a loro volta dei limiti. Costo, portabilità e disponibilità ne limitano la diffusione in ambito clinico e di ricerca [17, 18], e il loro uso resta vincolato al laboratorio, senza possibilità di monitoraggio nella vita quotidiana [11]. Inoltre misurano soltanto le forze scambiate con il suolo, e quindi non forniscono informazioni sul movimento del corpo: il CoP descrive l'azione di controllo posturale ma non il suo esito, cioè la stabilità del centro di massa (CoM), la cui misura richiede sistemi di tracciamento del movimento [19].

Tra questi, negli ultimi anni si ricorre sempre più spesso a sistemi di Motion Capture (MoCap) [1], che ricostruiscono lo scheletro del soggetto; a partire da questo, software dedicati ne stimano la cinematica e la posizione del baricentro [20].

1.2 Limiti dei Sistemi Optoelettronici

Attualmente, i sistemi optoelettronici basati su marcatori retroriflettenti (*marker-based*) rappresentano uno standard di riferimento per l'analisi in laboratorio e in ambito clinico (e.g. Vicon [21] o Qualisys [22]). Questi sistemi sono generalmente affiancati da pedane di forza per la misura delle forze di reazione al suolo (e.g. AMTI [23] o Kistler [24]) [1, 25, 26, 27]. Il principio di funzionamento dei sistemi optoelettronici è ricostruire le traiettorie tridimensionali dei marcatori posizionati in corrispondenza di punti anatomici di riferimento sfruttando più telecamere sincronizzate e di cui sono noti la posizione e i reciproci orientamenti. Questi sistemi sono caratterizzati da alti livelli di precisione che possono raggiungere errori di posizione inferiori a 2 mm [26].

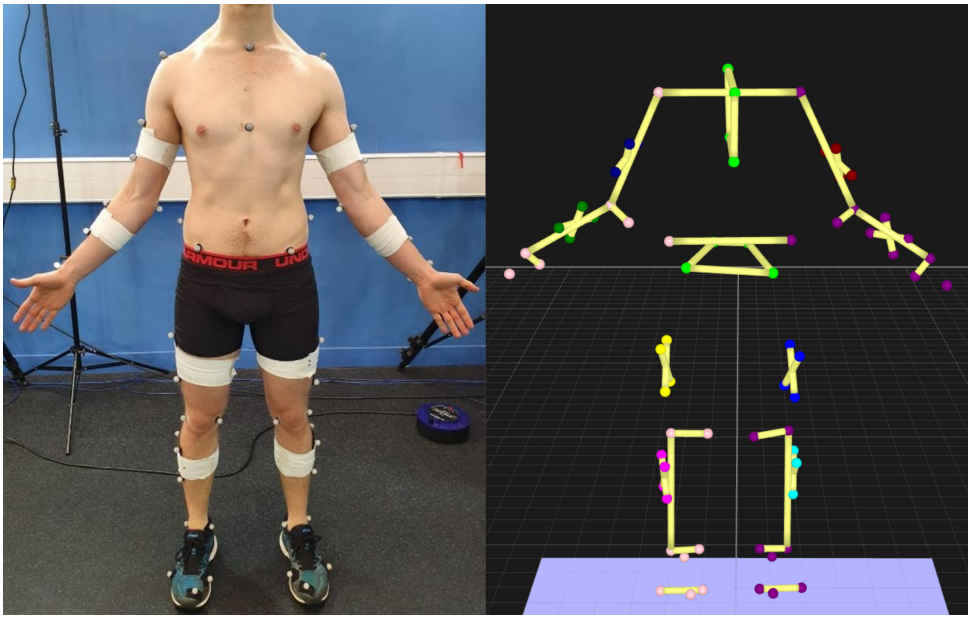


Figura 1.1: Marcatori ottici di un sistema di motion capture marker-based. A sinistra, i marcatori retroriflettenti posizionati sul soggetto; a destra, la loro ricostruzione nello spazio 3D. Figura tratta da Wade et al. [1], distribuita con licenza Creative Commons Attribution 4.0 (CC-BY).

Nonostante la precisione nominale, la letteratura evidenzia diversi problemi metodologici e pratici legati all'approccio marker-based:

1. Artefatti dei Tessuti Molli (Soft Tissue Artifact - STA)

   Pelle, tessuto adiposo e muscoli si muovono in modo indipendente rispetto all'osso sottostante. Questo introduce errori sistematici imprevedibili che non possono essere eliminati in maniera analitica. Studi evidenziano come si possa arrivare a errori di localizzazione del centro articolare fino a 28 mm e a deviazioni angolari sulla rotazione assiale fino a 21.8° [1, 26].

2. Errore umano di palpazione

   La collocazione manuale dei marcatori comporta un errore non trascurabile che dipende dall'operatore e mostra una variabilità inter- e intra-valutatore [28, 29].

3. Oneri logistici

   Il sistema optoelettronico richiede un ambiente dedicato, personale formato e ogni soggetto richiede una preparazione lunga (posizionamento dei marcatori) [1, 20].

Di conseguenza sono valutabili meno soggetti in un dato arco temporale rispetto a un sistema markerless e un uso clinico di routine è meno praticabile.

4. Compromissione della validità ecologica
   Il posizionamento di marcatori, con tute o bande sul corpo, può modificare il gesto, che di conseguenza potrebbe non essere più rappresentativo di quello svolto al di fuori dell'ambiente di misura [20].

1.3 Motion Capture Markerless e Human Pose Estimation

Negli ultimi anni i progressi della visione artificiale e del deep learning, in particolare delle reti neurali convoluzionali, hanno portato allo sviluppo della motion capture senza marker (*markerless*) [1, 25, 26]. Questo approccio evita gli errori legati al posizionamento manuale dei marcatori e al loro scivolamento rispetto all'osso, può ridurre i tempi di preparazione del soggetto [1, 20] e migliorare la validità ecologica della misura, ovvero la sua capacità di rappresentare il movimento così come avviene al di fuori del laboratorio [26, 30]. Resta però la dipendenza dall'aspetto esterno del corpo: i centri articolari sono stimati a partire dall'immagine, quindi pelle e abiti possono influenzare la stima, anche se negli studi che hanno confrontato abiti diversi l'effetto è risultato contenuto [30, 31]. Per questo motivo, nel protocollo sperimentale di questa tesi è stata comunque registrata la vestibilità degli abiti dei partecipanti; la scarsa variabilità del campione (12 partecipanti su 16 con abiti di vestibilità normale) non ha però consentito di valutarne l'effetto (Sezione 2.5.2).

1.3.1 Algoritmi di Human Pose Estimation (HPE)

Gli algoritmi di Human Pose Estimation analizzano i singoli fotogrammi video e individuano configurazioni di pixel (colore, texture, gradienti) che corrispondono ai centri articolari o ad altri punti anatomici (*keypoints*). Le prime applicazioni del deep learning alla stima della posa sono arrivate intorno al 2014. All'inizio non erano pensate per la biomeccanica né per usi clinici: servivano a risolvere problemi più generali di compu-

ter vision, come la videosorveglianza, l'interazione uomo-macchina o le interfacce per i videogiochi [32].

Negli ultimi anni l'interesse verso questi metodi è cresciuto rapidamente, grazie a reti convoluzionali sempre più profonde ed efficienti e alla disponibilità di grandi dataset pubblici annotati manualmente (COCO [33] e MPII [34]) [35, 36]. Queste risorse hanno reso possibile addestrare modelli in grado di generalizzare su soggetti diversi per età, corporatura e condizioni ambientali. L'addestramento avviene in modo supervisionato, secondo il workflow riportato in Figura 1.2.

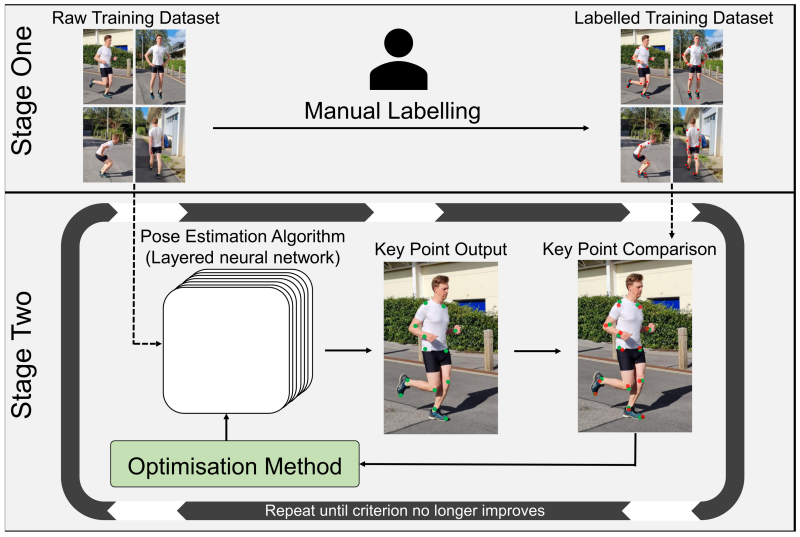


Figura 1.2: Workflow di addestramento di un algoritmo di Human Pose Estimation. Fase 1: creazione di un dataset di addestramento etichettato manualmente. Fase 2: l'algoritmo stima i keypoints sulle immagini non etichettate, li confronta con l'etichettatura manuale e un metodo di ottimizzazione regola i filtri della rete per ridurre lo scarto, ripetendo il processo fino a convergenza. Figura tratta da Wade et al. [1], distribuita con licenza Creative Commons Attribution 4.0 (CC-BY).

Questi strumenti sono poi entrati nella ricerca biomeccanica perché permettevano di superare i limiti logistici dei sistemi a marcatori. Hanno aperto la strada a soluzioni markerless per l'analisi del movimento sia in clinica che nello sport [1, 35].

Ad oggi i vari sistemi si possono dividere in base alla strategia con cui l'algoritmo elabora l'immagine (Figura 1.3) [37]:

- Approccio Top-Down. Lavora in due passaggi. Prima trova le persone nell'immagine e disegna un riquadro (*bounding box*) intorno a ciascuna; poi, dentro ogni

riquadro, stima le articolazioni. È generalmente accurato, ma il tempo di calcolo cresce con il numero di persone presenti nella scena.

- Approccio Bottom-Up. Prima individua tutti i keypoints presenti nell'immagine (e.g. tutti i gomiti e tutti i polsi) e dopo li raggruppa per persona. In scene affollate tende a essere più veloce, perché il lavoro iniziale non dipende dal numero di soggetti.

- Approccio One-Stage. Tutto avviene in un unico passaggio attraverso la rete, senza ritagli o raggruppamenti intermedi. È il più veloce e si presta bene al tempo reale, anche se può risultare meno preciso quando i soggetti sono piccoli o si sovrappongono molto.

Tra i modelli adottati in questa tesi, YOLO26x rientra nell'approccio *one-stage*: rileva le persone e stima i keypoints in un'unica inferenza end-to-end, senza post-elaborazione di tipo NMS (*Non-Maximum Suppression*) [38]. SAM 3D Body segue invece una logica *top-down*, perché elabora il ritaglio di ciascuna persona individuata, ma restituisce una mesh 3D del corpo anziché un insieme di keypoints [39]. Poiché nelle acquisizioni è presente un solo soggetto, il vantaggio dei metodi *bottom-up* e *one-stage* nelle scene affollate non è determinante; la scelta dei modelli è guidata da altri criteri, descritti nel Capitolo 2.

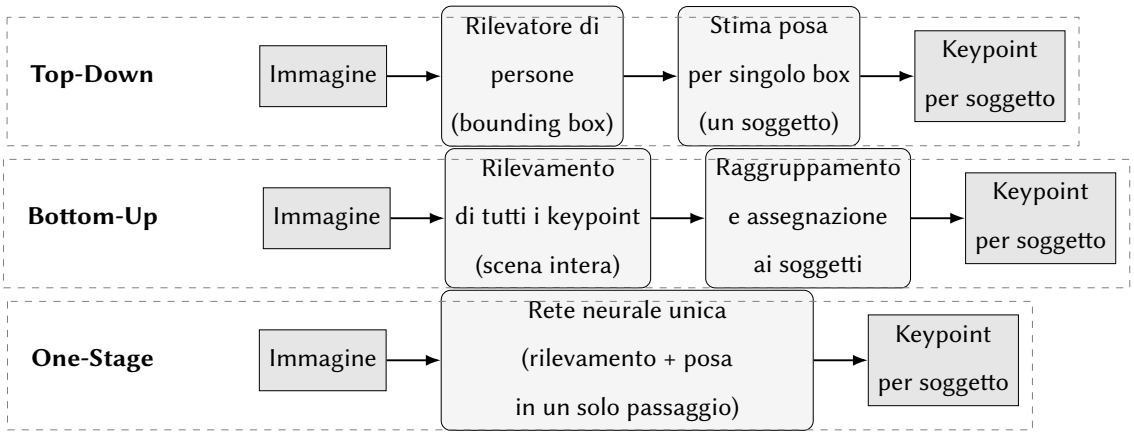


Figura 1.3: Confronto schematico tra i tre paradigmi di Human Pose Estimation: Top-Down (rilevamento della persona seguito dalla stima della posa), Bottom-Up (rilevamento di tutti i keypoints seguito dal raggruppamento per soggetto) e One-Stage (rilevamento e posa in un solo passaggio di rete). Schema originale.

Di seguito sono descritti i modelli, open-source e commerciali, più citati nella letteratura biomeccanica e clinica [1, 26, 35]:

- OpenPose. Modello open-source basato sull'approccio bottom-up. Per capire quali keypoints appartengono alla stessa persona usa i Part Affinity Fields (PAF), campi vettoriali 2D che portano informazione sia sulla posizione sia sull'orientamento degli arti. In questo modo riesce ad associare correttamente le articolazioni (e.g. spalla e gomito) e a gestire anche le scene multi-persona e le occlusioni in modo robusto (Figura 1.4) [1, 40].



Figura 1.4: I 25 keypoints rilevati dall'algoritmo di pose estimation OpenPose [40] applicato a una singola immagine. Figura tratta da Wade et al. [1], distribuita con licenza Creative Commons Attribution 4.0 (CC-BY).

- AlphaPose e HRNet. AlphaPose è un sistema completo di stima della posa e tracking multi-persona basato sull'approccio top-down [41]. HRNet è invece un'architettura di rete che mantiene rappresentazioni ad alta risoluzione lungo tutta l'elaborazione [42], e viene impiegata come backbone all'interno di pipeline top-down. Lavorando su una persona alla volta, già isolata nel suo bounding box, queste soluzioni concentrano le risorse di calcolo sul singolo soggetto e raggiungono un'accuratezza maggiore.

- YOLO-Pose. È un'architettura one-stage costruita sul framework YOLO. Usa l'Object Keypoint Similarity (OKS) come metrica di errore durante l'addestramento: confronta la posizione stimata con quella reale e migliora la precisione senza penalizzare la velocità, consentendo l'elaborazione in tempo reale [43, 37]. A questa famiglia appartiene YOLO26x, il modello adottato in questa tesi per il sistema multi-camera (Sezione 2.4.2).

- DeepLabCut e MediaPipe/BlazePose. Sono reti pensate per usi più specifici. DeepLabCut sfrutta il transfer learning: una rete preaddestrata viene riadattata con un numero ridotto di immagini etichettate dall'utente, così da riconoscere i punti specifici scelti per l'applicazione [44]. BlazePose invece è ottimizzato per funzionare su smartphone con bassa latenza e basso consumo energetico, e per questo trova impiego anche in applicazioni di teleriabilitazione [45, 46].

- Sistemi commerciali allo stato dell'arte (e.g. Theia3D [47]). Sono soluzioni a pagamento che offrono pacchetti completi: stima della posa 2D, calibrazione multi-camera, ricostruzione 3D e spesso anche modellazione muscoloscheletrica. Nel caso di Theia3D, il confronto con un sistema a marker ha mostrato un buon accordo sulle grandezze legate all'equilibrio, come il centro di massa, con bias di pochi millimetri e RMSD di circa 1 cm [20].

- Modelli di ricostruzione della mesh 3D (e.g. SAM 3D Body). Invece di stimare un insieme ridotto di keypoints, ricostruiscono da una singola immagine l'intera

superficie del corpo sotto forma di mesh 3D, appoggiandosi a modelli parametrici della forma umana [39].

1.3.2 Acquisizione Monoculare

Usare una sola telecamera è la soluzione più semplice e meno costosa: si elaborano i fotogrammi con un algoritmo di stima della posa 2D e si ottengono le posizioni dei centri articolari sul piano dell'immagine, espresse in pixel [25]. In ambito biomeccanico, tuttavia, questo approccio presenta diversi limiti.

Il principale è rappresentato dalle auto-occlusioni: se un arto passa davanti a un altro durante il movimento, l'algoritmo può perdere le coordinate del punto oppure scambiare il lato del corpo (il ginocchio destro diventa sinistro e viceversa) [1].

Inoltre, gli angoli articolari calcolabili restano confinati al piano dell'immagine. Per non introdurre errori prospettici bisogna assumere che la telecamera sia sempre perfettamente ortogonale al piano di movimento (sagittale o frontale). Nella pratica clinica e sportiva questa condizione è quasi impossibile da garantire in modo costante [1, 25].

Per superare questi limiti sono stati sviluppati metodi monoculari 3D, che stimano la posa tridimensionale, o l'intera mesh del corpo, a partire da una singola vista, sfruttando le conoscenze sulla forma e sul movimento del corpo umano apprese durante l'addestramento [36, 35]. Resta però un'ambiguità intrinseca sulla profondità e sulla scala, che una sola telecamera non può misurare direttamente [35]. A questa categoria appartiene SAM 3D Body, utilizzato in questa tesi (Sezione 2.4.3) [39].

1.3.3 Sistemi multi-camera e Ricostruzione 3D (5DoF vs 6DoF)

Il sistema multi-camera riduce in modo significativo il problema delle occlusioni [26, 25]. Si usano più telecamere sincronizzate: le predizioni 2D ottenute da punti di vista diversi vengono combinate e le coordinate spaziali si ottengono per triangolazione [1, 48].

Il sistema richiede una calibrazione ottica preliminare. Quella intrinseca corregge le distorsioni delle lenti; quella estrinseca definisce posizione e orientamento di ogni telecamera rispetto a un sistema di coordinate comune nell'ambiente di misura [1].

Il sistema multi-camera adottato in questa tesi segue questo schema, con quattro telecamere sincronizzate e calibrate (Sezione 2.4.2).

Per descrivere in modo completo la cinematica di un segmento corporeo servono sei gradi di libertà (6DoF): le tre traslazioni spaziali e le tre rotazioni attorno agli assi ortogonali, che in ambito clinico corrispondono a flesso-estensione, abduzione-adduzione e rotazione interna-esterna (Figura 1.5) [1].

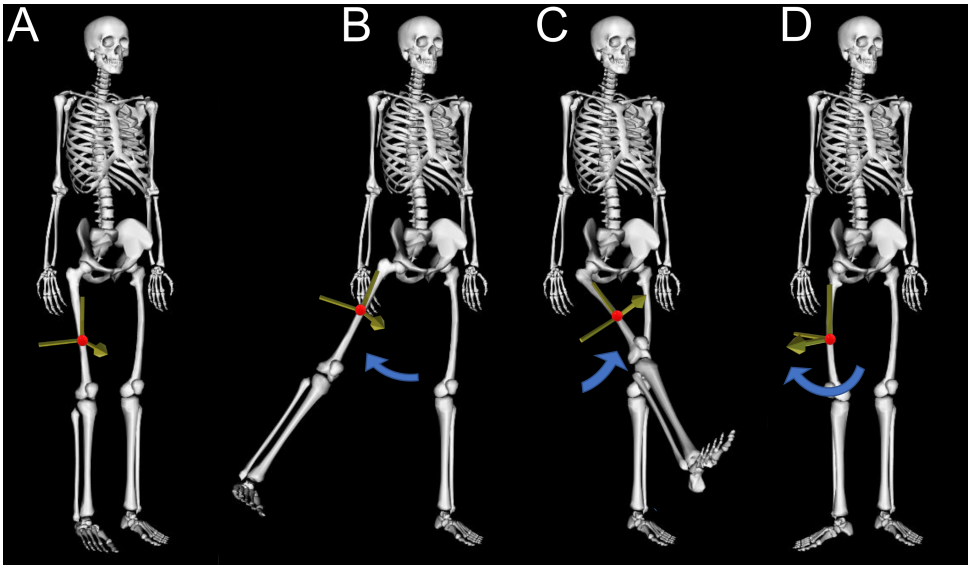


Figura 1.5: Rappresentazione dei sei gradi di libertà (6DoF) necessari per descrivere la posa di un segmento corporeo: (A) postura di riferimento, (B) adduzione/abduzione, (C) flesso-estensione, (D) rotazione attorno all'asse longitudinale. Il punto rosso indica il centro di massa del segmento, le frecce blu la rotazione nei tre piani. Figura tratta da Wade et al. [1], distribuita con licenza Creative Commons Attribution 4.0 (CC-BY).

Nei sistemi markerless la stima di tutti e sei i gradi di libertà resta però problematica, perché gli algoritmi di HPE rilevano in genere solo due keypoints per segmento (e.g. per il femore vengono individuati solamente i keypoints dell'anca e del ginocchio) [1, 26]. Senza un terzo punto non collineare non si riesce a misurare la rotazione intorno all'asse longitudinale del segmento. La descrizione del segmento si ferma quindi a 5 gradi di libertà invece di 6 (Figura 1.6) [1].

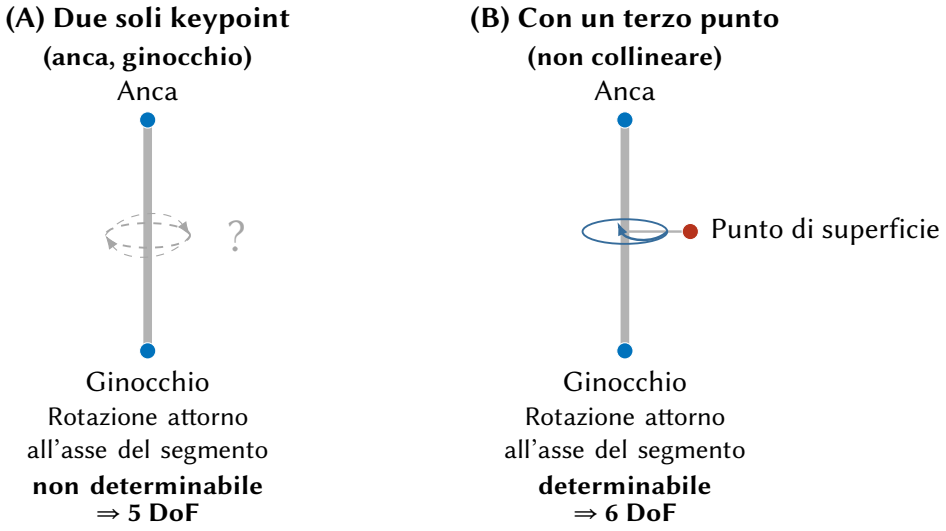


Figura 1.6: Con due soli keypoints per segmento (A) la rotazione attorno all'asse longitudinale del segmento non è determinabile (5 DoF); un terzo punto di superficie non collineare (B), ottenuto ad esempio per *keypoint augmentation*, la rende determinabile (6 DoF). Schema originale.

Per recuperare il sesto grado di libertà si sta lavorando in due direzioni: integrare gli output degli HPE con modelli muscoloscheletrici (e.g. tramite la pipeline Pose2Sim, che scala e guida un modello muscoloscheletrico in OpenSim) oppure usare tecniche di keypoint augmentation. Queste ultime, basate su deep learning, imparano a ricostruire punti di superficie a partire dai centri articolari e li ancorano in modo rigido ai segmenti. In entrambi i casi l'obiettivo è vincolare l'orientamento del segmento attorno al proprio asse, così da poterne stimare la rotazione [1, 26, 48].

Per gli obiettivi di questa tesi il limite dei 5DoF ha però un impatto ridotto: nel modello geometrico adottato (Sezione 2.4.2) ogni segmento è un solido con il centro di massa sul proprio asse longitudinale, per cui la rotazione attorno a tale asse non modifica la posizione del CoM.

1.4 Stima delle Proprietà Inerziali Corporee e Modelli Biomeccanici

Per passare dalla cinematica dei segmenti corporei alla posizione del centro di massa dell'intero corpo e i momenti articolari, è necessario assegnare a ogni segmento una

massa, la posizione del proprio CoM e un tensore di inerzia. Questi valori prendono il nome di Body Segment Inertial Parameters (BSIP) [49]. In particolare, il CoM globale si ottiene come media delle posizioni dei centri di massa dei segmenti, pesata per le rispettive masse:

$$\mathbf{r}_{CoM} = \frac{\sum_i m_i \, \mathbf{r}_i}{\sum_i m_i} \tag{1.1}$$

dove m_i e $\mathbf{r}_i$ sono la massa e la posizione del centro di massa dell'i-esimo segmento.

Tradizionalmente i BSIP sono ricavati da tabelle antropometriche, costruite su misurazioni cadaveriche (e.g. Dempster) o su campioni di soggetti vivi (e.g. le tabelle di McConville e Young, riadattate da Dumas et al.), e successivamente riscalate in base a massa e statura del soggetto [49, 50]. Queste tabelle assumono però distribuzioni di massa regolari e proporzioni fisse, derivate da campioni di numerosità limitata. In presenza di morfologie che si discostano da quelle di riferimento, come nella sarcopenia (perdita di massa muscolare legata all'età) o nell'obesità grave, tessuti molli e tessuto adiposo si distribuiscono in modo diverso: i centri di massa dei segmenti si spostano, i tensori di inerzia cambiano e le stime generiche perdono di affidabilità [50, 51].

Per questo motivo la ricerca si è orientata verso modelli soggetto-specifici non invasivi, che si possono distinguere in due categorie:

- Modelli geometrici (e.g. modello di Yeadon). Approssimano il corpo con una catena di solidi geometrici (ellissoidi, tronchi di cono e altri solidi regolari), senza l'uso di reti neurali (Figura 1.7).

  - *Acquisizione.* Un operatore rileva manualmente misure antropometriche (lunghezze, circonferenze, larghezze, profondità dei vari segmenti) con metri a nastro e calibri.

  - *Pro e contro.* Permettono di ottenere stime soggetto-specifiche senza ricorrere a esami radiologici, con una buona personalizzazione morfologica. La procedura di misura è però lunga, soggetta a errori e richiede personale ad-

destrato, il che la rende poco pratica per valutazioni cliniche rapide o su un numero elevato di soggetti [51, 52].

Il modello geometrico adottato in questa tesi per le analisi effettuate tramite il sistema multi-camera (Sezione 2.4.2) appartiene a questa categoria, ma ricava le dimensioni dei solidi direttamente dai keypoints stimati dal video, superando così la necessità di misure manuali; la massa di ciascun segmento è invece assegnata tramite i coefficienti antropometrici di Dumas et al. [50].

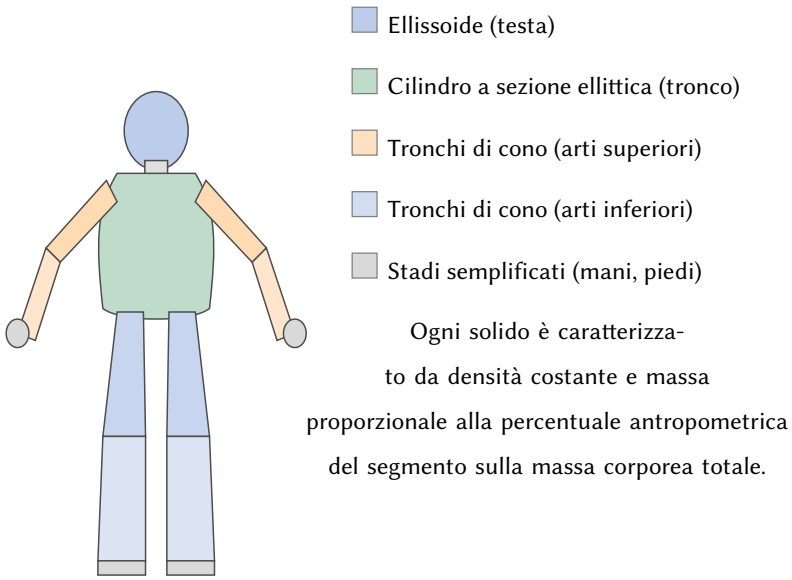


Figura 1.7: Schema illustrativo semplificato di un modello geometrico del corpo umano a solidi regolari (approccio concettualmente analogo a quello di Yeadon [52]): ciascun segmento è approssimato con un solido di densità costante (ellissoide, cilindro, tronco di cono).

- Scansioni 3D e avatar digitali basati su AI (e.g. SMPL). Questi sistemi automatizzano l'estrazione dei dati di superficie, impiegando reti neurali e modelli statistici della forma umana addestrati su migliaia di scansioni 3D ad alta risoluzione.
  - *Acquisizione.* Una rete neurale stima i parametri di forma e di posa da immagini e/o video; a partire da questi, il modello statistico genera una mesh del soggetto ad alta definizione (Figura 1.8).
  - *Pro e contro.* Consentono un'acquisizione senza contatto, rapida e poco onerosa per il soggetto. Restano tuttavia sensibili agli abiti, che alterano la si-

lhouette e quindi la forma della mesh ricostruita, e richiedono una notevole potenza di calcolo [51, 53].

A questa categoria appartiene anche SAM 3D Body, adottato in questa tesi per le analisi effettuate dal sistema single-camera (Sezione 2.4.3), che ricostruisce l'intero volume corporeo del soggetto a partire da una singola telecamera [39].

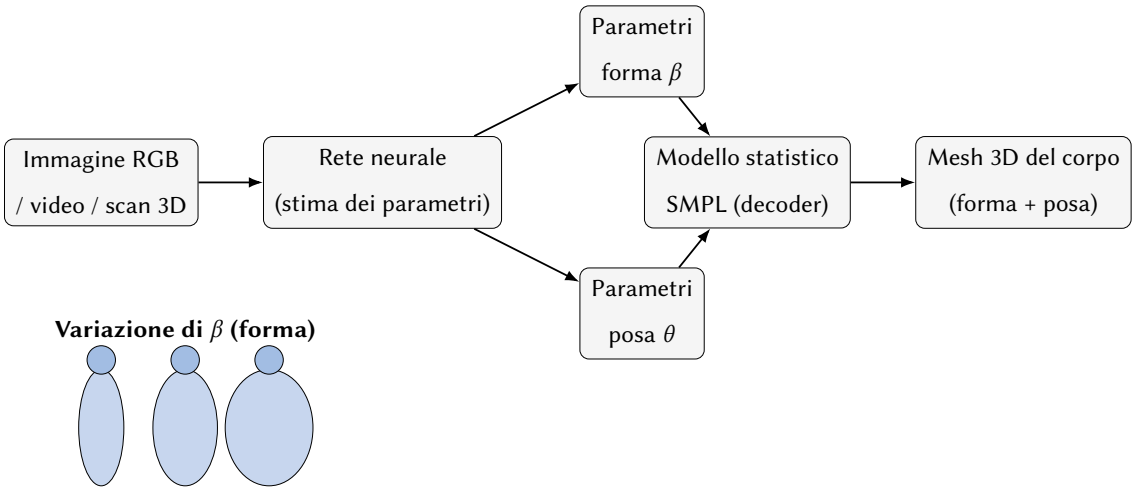


Figura 1.8: Schema concettuale del funzionamento di un modello statistico del corpo come SMPL [53]: una rete neurale stima dai dati di ingresso i parametri di forma (β) e di posa (θ), per generare la mesh 3D del soggetto.

1.5 Controllo Posturale e Stabilometria

La stabilometria valuta la capacità del sistema nervoso centrale di integrare le informazioni visive, vestibolari e propriocettive per mantenere la proiezione del Centro di Massa all'interno della Base di Appoggio (BoS) [13, 54].

Le grandezze di riferimento sono principalmente due:

- Centro di Pressione. È una misura cinetica. Indica il punto in cui agisce il vettore risultante delle forze di reazione al suolo e si ricava dalle forze e dai momenti misurati dalle piattaforme di forza [14, 55].
- Centro di Massa e Extrapolated CoM (XCoM). Queste sono invece misure cinematiche. Il CoM è il punto in cui si può considerare concentrata l'intera massa

corporea; in stabilometria se ne considera la proiezione verticale sul piano di appoggio. Per descrivere la stabilità in condizioni dinamiche, Hof et al. hanno introdotto l'Extrapolated Center of Mass, che aggiunge alla posizione del CoM un termine proporzionale alla sua velocità: il corpo è considerato stabile finché la proiezione dell'XCoM resta all'interno della Base di Appoggio (Figura 1.9) [56]. In questa tesi la grandezza non viene però utilizzata: l'analisi si concentra su CoM e CoP.

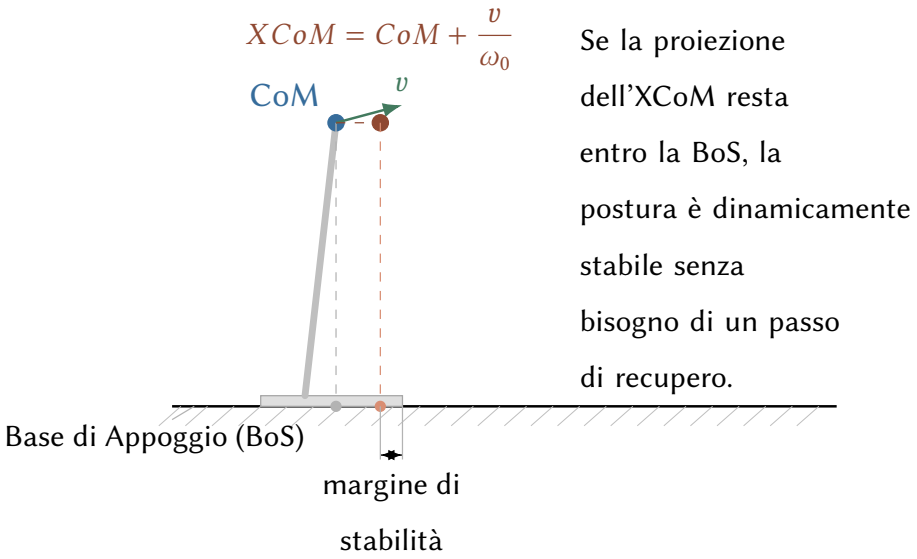


Figura 1.9: Rappresentazione del margine di stabilità dinamica [56]: approssimazione a pendolo inverso, il Centro di Massa che si muove con velocità v; l'Extrapolated Center of Mass ne anticipa la posizione. Finché la proiezione dell'XCoM resta all'interno della Base di Appoggio, la postura è dinamicamente stabile senza necessità di un passo di recupero. Schema originale.

Per la valutazione quantitativa del segnale stabilometrico si utilizzano diverse metriche spaziali e temporali, ognuna delle quali ne descrive un aspetto differente [13, 14, 15]:

- Sway Area (Area di oscillazione). Si calcola generalmente come l'area dell'ellisse di confidenza che contiene il 95% dei punti della traiettoria. Descrive la dispersione dell'oscillazione: un'area maggiore indica una minore capacità del soggetto di mantenere stabile la propria posizione.

- Range Antero-Posteriore (AP). Escursione massima della traiettoria in direzione antero-posteriore. In stazione eretta il controllo in questa direzione è affidato principalmente alla strategia di caviglia [54].
- Range Medio-Laterale (ML). Escursione massima della traiettoria in direzione medio-laterale. In stazione eretta bipodalica il controllo in questa direzione è affidato principalmente ai muscoli abduttori e adduttori dell'anca e dipende dalla larghezza della base di appoggio; valori elevati possono evidenziare asimmetrie tra i due lati [54].
- RMS Antero-Posteriore e Medio-Laterale. Deviazione standard della posizione lungo ciascun asse: misura l'ampiezza tipica dell'oscillazione attorno alla posizione media.
- Path Length (Lunghezza del tracciato). È la distanza totale percorsa dal CoP (o dal CoM) durante tutta la prova. Fornisce una stima dell'attività correttiva svolta dal sistema neuromuscolare per mantenere l'equilibrio.
- Mean Velocity (Velocità media). Si ricava dividendo la lunghezza del tracciato per la durata dell'acquisizione e descrive la rapidità delle correzioni posturali: valori elevati indicano un controllo meno stabile. È tra le metriche più sensibili ai cambiamenti della stabilità posturale, perché distingue meglio di altre le diverse condizioni di prova, di salute e di età [57, 58], e mostra un'affidabilità test-retest in genere buona [15].

1.5.1 Accoppiamento CoM-CoP e Modello a Pendolo Inverso

In stazione eretta bipodalica (*quiet standing*), il corpo può essere visto, in prima approssimazione, come un pendolo inverso rigido che ruota intorno alle caviglie [54, 55]. In questo schema il CoP diventa la variabile di controllo: il sistema neuromuscolare lo modula continuamente per controllare la posizione del CoM.

Sia in direzione antero-posteriore sia in direzione medio-laterale, per arrestare la caduta del CoM il CoP deve essere portato oltre la proiezione a terra del CoM, nella stessa

direzione dello sbilanciamento. Si genera così un momento che inverte l'accelerazione e riporta il corpo verso l'equilibrio (Figura 2.7) [54, 59].

Questa relazione può essere usata anche in senso inverso: noti la traiettoria del CoM e la sua accelerazione, il modello del pendolo inverso consente di stimare il CoP secondo un approccio di dinamica inversa. È il principio su cui si basa la stima del CoP a partire dai sistemi markerless in questa tesi (Sezione 2.6.2).

1.6 Obiettivi dello Studio e Struttura della Tesi

Negli ultimi anni diversi studi hanno mostrato che i sistemi markerless forniscono stime affidabili della cinematica e degli angoli articolari in movimenti ampi, come il cammino o la corsa; la loro applicazione alla stabilometria risulta invece ancora poco affidabile [25, 26, 20].

Il problema di fondo è il rapporto segnale-rumore. Nei gesti dinamici l'errore medio degli algoritmi di Human Pose Estimation, compreso tra circa 1.5 e 3 cm a seconda del modello e della risoluzione, è accettabile; in stazione eretta quieta, invece, le oscillazioni fisiologiche del corpo sono dell'ordine dei millimetri. L'errore di tracciamento rischia quindi di mascherare il segnale reale. La stima del CoP con il modello del pendolo inverso è ancora più delicata, perché richiede l'accelerazione del CoM e quindi una doppia derivazione, che amplifica il rumore [20].

L'obiettivo di questa tesi è validare un sistema markerless per la ricostruzione della traiettoria del CoM e la stima, a partire da essa, del CoP in modo non invasivo, utilizzando come riferimento il CoP misurato da una pedana di forza. Sono stati confrontati due sistemi: un sistema multi-camera, basato su quattro telecamere sincronizzate, sul modello YOLO26x e su un modello geometrico a solidi, e un sistema single-camera basato su SAM 3D Body.

Nel dettaglio, gli obiettivi dello studio presentato sono:

1. definire un protocollo di acquisizione che comprenda sia prove di equilibrio statico (*Quiet Standing* a *Occhi Aperti* e *Occhi Chiusi*), sia prove più impegnative per il

controllo posturale, a base di appoggio ridotta (*Tandem Stance* e *Single-Leg Stance*) o di natura dinamica (*Limits of Stability*);

2. stimare il CoM con entrambi i sistemi e ricavarne il CoP tramite il modello del pendolo inverso, definendo una procedura di filtraggio adeguata alla derivazione del segnale;

3. confrontare il CoP stimato da ciascun sistema con quello misurato dalla pedana di forza, acquisita in modo sincronizzato con le telecamere, attraverso gli indici posturografici descritti in precedenza;

4. quantificare l'accordo tra i sistemi mediante gli errori sugli indici, i coefficienti di correlazione di Pearson e di Spearman, l'RMSE di traiettoria e i grafici di Bland-Altman.

La tesi è organizzata come segue. Il Capitolo 2 descrive la strumentazione utilizzata (pedana di forza, videocamere e computer), i modelli adottati per i due sistemi markerless (YOLO26x per il sistema multi-camera e SAM 3D Body per il sistema single-camera), il protocollo di acquisizione, lo studio del segnale e la sua elaborazione (analisi in frequenza e scelta del filtraggio), la stima del CoP con il pendolo inverso, il calcolo degli indici posturografici e delle misure di accordo. Il Capitolo 3 presenta i risultati dei sistemi confrontati con la pedana: valori assoluti degli indici, errori percentuali, correlazioni, analisi di accordo di Bland-Altman. Il Capitolo 4 conclude la tesi discutendo i risultati, i limiti dello studio e i possibili sviluppi futuri.

Capitolo 2

Materiali e metodi

2.1 Pedana di Forza

La pedana di forza utilizzata è una struttura composta da un insieme di interfacce e da una cella di carico a sei assi, il trasduttore ATI Axia130-M300 [60] (Figura 2.1). La piastra centrale è collegata al trasduttore, misura 50×50 cm ed è circondata da prolungamenti fissi di larghezza 10 cm non accoppiati al sensore.

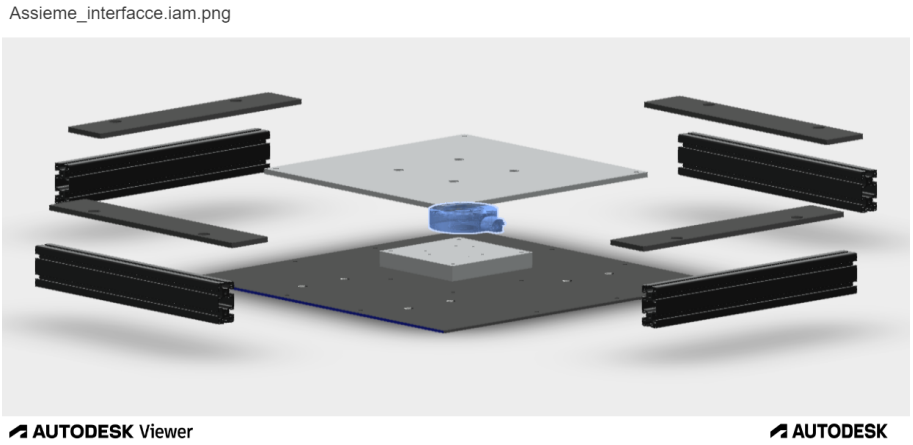


Figura 2.1: Vista esplosa della pedana di forza: interfacce, piastre superiori e inferiori, profilato e supporti laterali. È evidenziato il trasduttore ATI Axia130-M300.

2.1.1 Trasduttore F/T: Axia130-M300

Il sensore è un trasduttore a sei assi ad alte prestazioni (Figura 2.2), impiegato per misurare le tre componenti della forza e i tre momenti rispetto agli assi X, Y e Z; le specifiche tecniche sono riportate nella Tabella 2.1.

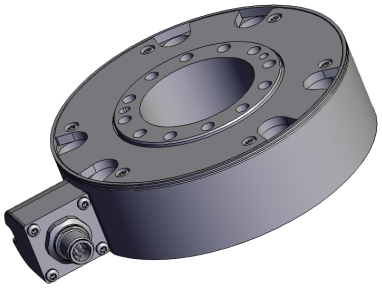


Figura 2.2: Rappresentazione 3D del Trasduttore F/T: Axia130-M300

Sovraccarico monoassiale	
Fxy	±20000 N
Fz	±30000 N
Txy	±1500 Nm
Tz	±1500 Nm
Rigidità (calcolata)	
Forze sull'asse X e sull'asse Y (Kx, Ky)	1,5 x 10^8 N/m
Forza sull'asse Z (Kz)	2,2 x 10^8 N/m
Coppia asse X e asse Y (Ktx, Kty)	2,8 x 10^8 Nm/rad
Coppia sull'asse Z (Ktz)	6,0 x 10^8 Nm/rad
Frequenza di risonanza	
Fx, Fy, Tz	2450 Hz
Fz, Tx, Ty	2900 Hz
Specifiche fisiche	
Peso*	1,88 kg
Diametro*	130 mm
Altezza*	39,2 mm

Tabella 2.1: Specifiche tecniche del trasduttore ATI Axia130-M300 (calibrazione SI-4000-300) [60].

2.2 Videocamera IDS U3-3160CP-C-HQ Rev. 2.2 e Ottica

Il sistema di acquisizione ottica utilizza telecamere industriali con interfaccia USB3, modello IDS U3-3160CP-C-HQ Rev. 2.2 [61] (Tabella 2.2). Il dispositivo è equipaggiato con un sensore CMOS da 2/3 di pollice (modello PYTHON 2000, prodotto da onsemi [62]) basato su architettura *global shutter*, dove l'esposizione simultanea di tutti i pixel del sen-

sore previene le distorsioni spaziali (effetto *rolling shutter*) tipiche della ripresa di corpi in movimento. Il sensore garantisce una risoluzione nativa di 2.30 Megapixel, fornendo matrici d'immagine da 1920×1200 pixel, e supporta una frequenza di acquisizione massima fino a 169 fps, assicurando un'elevata densità di campionamento temporale per i successivi algoritmi di tracciamento.

Ogni videocamera è equipaggiata con un obiettivo industriale Kowa LM6JC [63], con attacco C-mount, lunghezza focale di 6 mm e apertura regolabile da f/1.4 a f/16 (Figura 2.3).



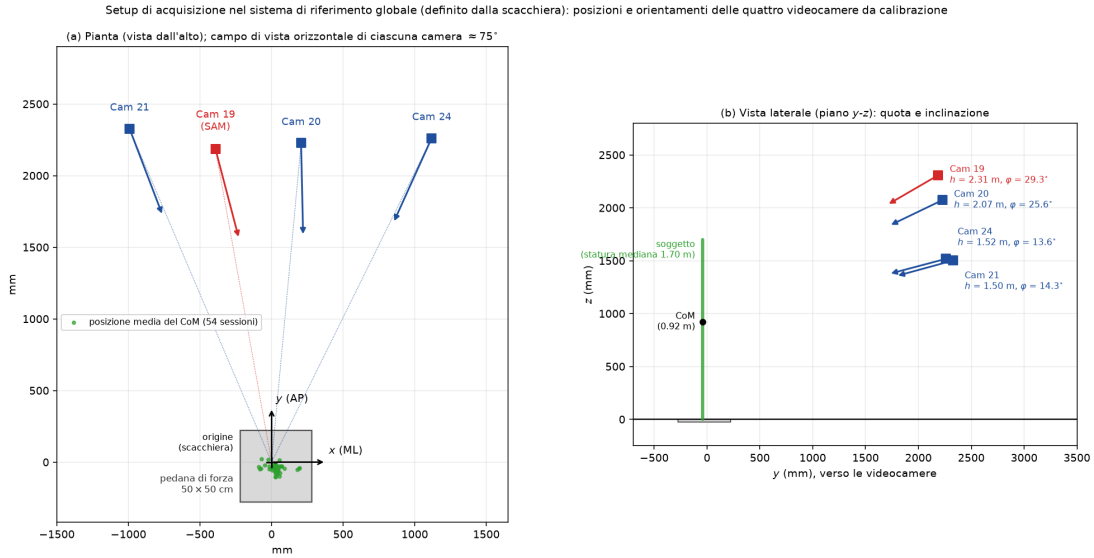
Figura 2.3: Immagine della videocamera IDS con l'ottica montata

Sensor	type CMOS Color
Shutter	Global Shutter
Sensor characteristic	Linear
Pixel Class	2 MP
Resolution	2.30 Mpix
Resolution (h x v)	1920 x 1200 Pixel
Aspect ratio	16:10
ADC	10 bit
Color depth (camera)	12 bit
Optical sensor class	2/3”
Optical Size	9.216 mm x 5.760 mm
Optical sensor diagonal	10.87 mm 2/3"
Pixel size	4.8 µm
CRA (Chief Ray Angle)	4 °
Manufacturer	Onsemi
Sensor Model	NOIP1SE2000A-LTI
Gain (master/RGB)	4x/4x
AOI horizontal	increased frame rate
AOI vertical	increased frame rate
AOI image width / step width	128 / 16
AOI image height / step width	2 / 2
AOI position grid (horizontal/vertical)	16 / 2
Binning horizontal	-
Binning vertical	-
Binning method	-
Binning factor	-
Decimation (subsampling) horizontal	increased frame rate
Decimation (subsampling) vertical	increased frame rate

2.3 Sistema di Acquisizione

Il sistema di acquisizione è stato strutturato con una pedana di forza centrale (Sezione 2.1) e quattro videocamere IDS (Sezione 2.2), poste frontalmente con angolature diverse e altezze leggermente diverse (Figure 2.5 e 2.6).

Le quattro videocamere sono disposte frontalmente al soggetto, tutte dallo stesso lato della pedana, a una distanza orizzontale dall'origine del sistema di riferimento globale compresa tra 2.2 e 2.5 m e a quote diverse (Figura 2.4). Il sistema di riferimento globale ha gli assi paralleli a quelli della pedana: in tutte le sessioni l'asse X corrisponde alla direzione mediolaterale e l'asse Y a quella antero-posteriore, come verificato dall'orientamento della linea delle anche dei soggetti.



Camera	x (mm)	y (mm)	h (mm)	d (mm)	θ (°)	ψ (°)	φ (°)
19 (SAM)	-392	2186	2308	2221	-10.2	+14.7	29.3
20	+206	2227	2073	2236	+5.3	+1.3	25.6
21	-992	2328	1501	2530	-23.1	+20.9	14.3
24	+1118	2261	1518	2522	+26.3	-24.3	13.6

x, y = posizione in pianta; h = altezza; d = distanza orizzontale dall'origine; θ = posizione angolare dall'origine rispetto all'asse y; ψ = orientamento dell'asse ottico rispetto all'asse $-y$ (positivo verso $+x$); φ = inclinazione verso il basso

Figura 2.4: Schema del setup di acquisizione nel sistema di riferimento globale: (a) pianta, (b) vista laterale e tabella dei parametri di ciascuna videocamera ottenuti dalla calibrazione, identici in tutte le sessioni: posizione in pianta (x, y), altezza h, distanza orizzontale d dall'origine, posizione angolare θ rispetto all'asse y, orientamento ψ dell'asse ottico rispetto all'asse $-y$ (positivo verso $+x$) e inclinazione φ verso il basso. La posizione della pedana e quella del soggetto sono stimate dai dati.



Figura 2.5: Vista della pedana di forza all'interno del setup sperimentale



Figura 2.6: Vista delle 4 videocamere all'interno del setup sperimentale

2.3.1 Workstation e sincronizzazione

Le acquisizioni e le elaborazioni dei modelli sono state eseguite con un computer dotato di processore Intel(R) Core(TM) i7-13700 di 13a generazione e scheda grafica NVIDIA GeForce RTX 3060.

Le videocamere sono state sincronizzate via software tramite un segnale digitale periodico a 60 Hz generato da un thread dedicato. Il segnale digitale è stato inviato

parallelamente alle quattro telecamere. L'allineamento tra i rispettivi frame delle telecamere risulta inferiore al millisecondo [64]. Il flusso di dati della pedana, trasmesso a 500 Hz tramite connessione UDP, è stato campionato mediante lo stesso segnale digitale utilizzato per l'acquisizione delle telecamere.

2.4 Modelli

2.4.1 Razionale Biomeccanico: Il Modello a Pendolo Inverso

Per stimare il Centro di Pressione dal Centro di Massa ricostruito dal sistema markerless, e confrontarlo con il CoP misurato dalla pedana, il corpo umano è stato approssimato con un pendolo inverso rigido, incernierato alle caviglie, modello comunemente adottato per descrivere l'equilibrio in stazione eretta [54, 55] (Figura 2.7). Le equazioni utilizzate sono riportate nella Sezione 2.6.2.

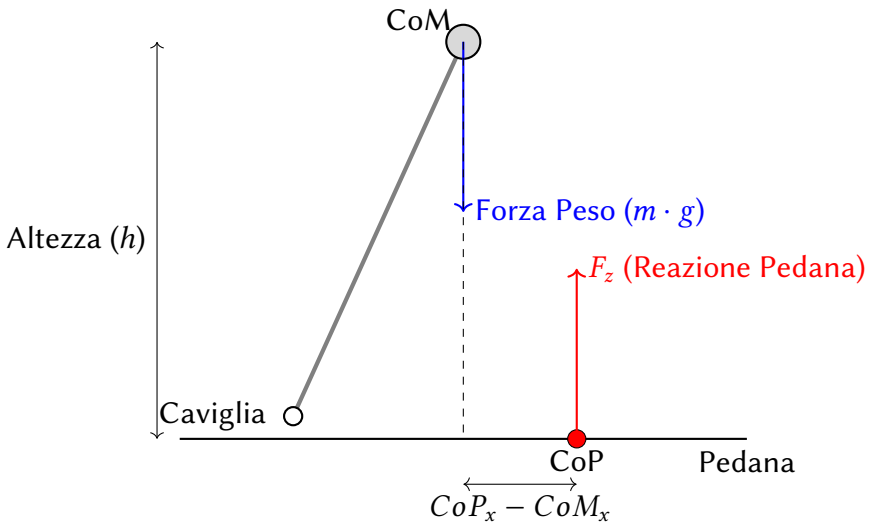


Figura 2.7: Rappresentazione del momento raddrizzante nel modello a Pendolo Inverso. Per frenare la caduta in avanti, il CoP deve spingersi fisicamente oltre la proiezione a terra del CoM, generando una forza di richiamo verso l'indietro.

2.4.2 Modello e Discretizzazione Geometrica per il Sistema multi-camera

La procedura complessiva, dall'acquisizione multi-camera alla stima delle proprietà inerziali del corpo, è schematizzata in Figura 2.8.

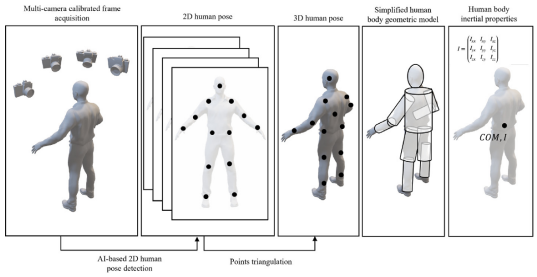


Figura 2.8: Schema della procedura e delle fasi del sistema di misura markerless proposto, dall'acquisizione multi-camera alla stima delle proprietà inerziali del corpo (CoM e tensore d'inerzia), passando per la 2D human pose estimation, la triangolazione 3D e il modello geometrico semplificato del corpo. Adattato da *Giulietti et al.* [64].

YOLO26x

Per l'estrazione della cinematica multi-camera, il presente lavoro impiega YOLO26x, la variante extra-large della recente architettura YOLO26, la più accurata della famiglia a fronte di un costo computazionale maggiore [38], utilizzata con i pesi preaddestrati su COCO e senza riaddestramento. A differenza dei tradizionali rilevatori convoluzionali, YOLO26x adotta un'architettura *dual-head* con inferenza nativa *end-to-end*, eliminando la necessità di post-elaborazione NMS (Non-Maximum Suppression) e rimuovendo il modulo DFL (Distribution Focal Loss), ottenendo una testa di rete più leggera e un intervallo di regressione non vincolato [38].

Nel lavoro in cui sono state confrontate diverse configurazioni di modelli, formati e telecamere su Human3.6M, la più accurata è risultata YOLOv8x-pose con quattro telecamere, con un errore medio per giunto (MPJPE) di 18.2 mm [65]. È stato scelto YOLO26x, variante extra-large della versione più recente della stessa famiglia, che per la stima della posa migliora l'AP di 2.1 punti rispetto a YOLO11x-pose sul set di validazione COCO (71.6 contro 69.5) [38]. Tra le sue novità vi è la *Residual Log-Likelihood Estimation* (RLE), un obiettivo di addestramento che modella l'incertezza di ciascun giunto riducendo il

peso dei keypoints ambigui o occlusi [38]. Si mantiene così la continuità con il sistema presentato da *Giulietti et al.*, in cui si utilizza la stessa famiglia di modelli e la stessa configurazione di camere e keypoints [64].

Il sistema multi-camera adottato da *Giulietti et al.* stima il CoM e il tensore d'inerzia a partire da 13 keypoints 2D triangolati da quattro telecamere sincronizzate, utilizzando il modello YOLOv11x-pose e un modello geometrico semplificato del corpo. Gli autori lo hanno validato tramite uno studio condotto su 15 partecipanti sani in cui si mette a confronto il momento generato dal soggetto attorno all'asse di rotazione di una cella di carico a un grado di libertà, e quello stimato dal modello del sistema. I risultati presentano un errore assoluto medio mediano di 5.88 Nm e un coefficiente di correlazione di Pearson mediano di 0.98. Per queste ragioni, in continuità con quanto già studiato e sviluppato, questa tesi adotta lo stesso sistema di stima, sostituendo il modello AI di stima della posa con YOLO26x.

I keypoints 2D di ciascuna telecamera sono triangolati con una DLT pesata (*weighted direct linear transform*): per ogni keypoint si calcolano più candidati 3D dalle coppie di telecamere, se ne valuta l'errore di riproiezione e si pesa ciascuna triangolazione in base alla coerenza tra le telecamere e con il frame precedente [65, 64]. Nel sistema di *Giulietti et al.* i keypoints con visibilità inferiore a 0.5 sono esclusi dalla triangolazione, che richiede almeno due telecamere; quando ciò non è possibile, la posizione è propagata dal frame precedente [64]. A differenza del sistema di *Giulietti et al.*, in questa tesi non è stato applicato il filtro α-β-γ sui keypoints 2D. Nelle acquisizioni di questa tesi la confidenza dei 13 keypoints utilizzati è sempre superiore a 0.5 (mediana 0.99): il 79% dei keypoints è triangolato con tutte e quattro le telecamere, il 20% con tre, e l'errore di riproiezione medio ha una mediana di 4.3 px. La triangolazione è stata eseguita con un software sviluppato in laboratorio, lo stesso utilizzato per la calibrazione.

Discretizzazione Geometrica del Corpo Umano

Per calcolare la posizione del Centro di Massa, il corpo umano è stato approssimato tramite un modello geometrico a 10 segmenti tridimensionali, le cui dimensioni e posizioni

sono vincolate istante per istante alle coordinate dei 13 *keypoints* estratti dal sistema ottico (Figure 2.9 e 2.10) [64].

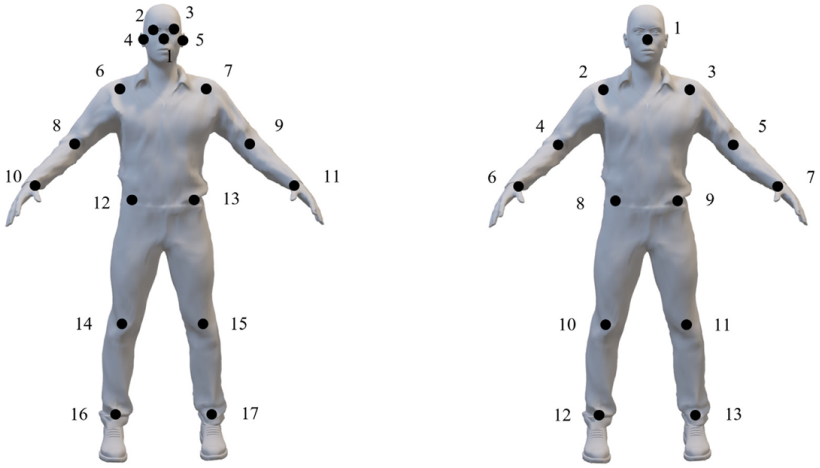


Figura 2.9: Confronto tra i 17 keypoints originali stimati da un modello YOLO-pose e la versione ridotta a 13 keypoints, ottenuta escludendo occhi e orecchie per la difficoltà di rilevamento consistente e la loro limitata rilevanza per le analisi successive. Approccio analogo a quello adottato in questa tesi con YOLO26x. Adattato da *Giulietti et al.* [64].

Geometricamente, la testa e il collo sono fusi in un ellissoide, il tronco è rappresentato come un cilindro a sezione ellittica, mentre gli arti sono rappresentati da tronchi di cono [64]. Mani e piedi, che non hanno keypoints associati, non costituiscono segmenti a sé: le percentuali di massa di Dumas et al. [50] sono assegnate ai soli dieci segmenti e sommano a circa il 100% della massa corporea, per cui la massa di mani e piedi è inglobata negli avambracci e nelle gambe, i cui solidi sono definiti dai keypoints di gomito e polso e di ginocchio e caviglia.

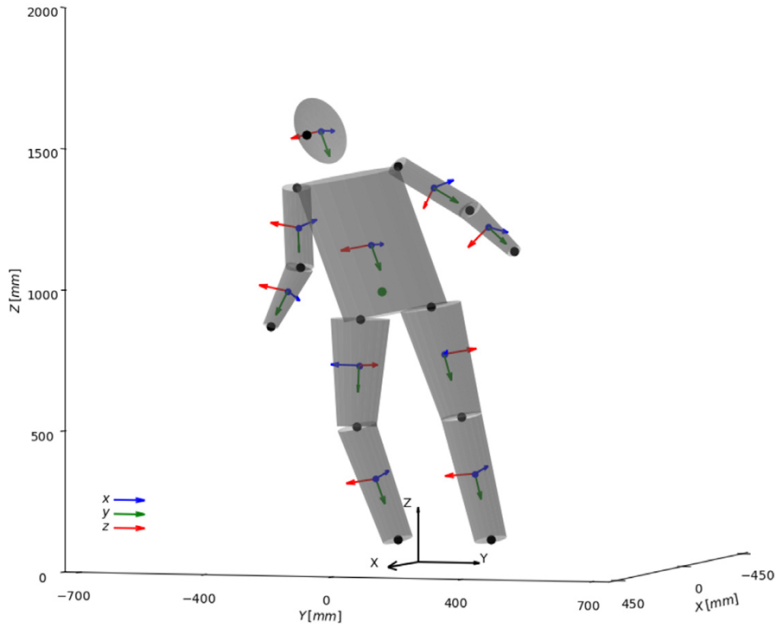


Figura 2.10: Il corpo umano discretizzato in 10 segmenti a partire dalla posizione dei 13 keypoints (in nero). Il CoM di ciascun segmento è rappresentato in blu, con il relativo sistema di riferimento locale; il CoM globale calcolato dall'algoritmo è rappresentato in verde. Adattato da *Giulietti et al.* [64].

La massa di ogni solido viene assegnata come percentuale della massa corporea totale del soggetto, applicando i coefficienti antropometrici differenziati per sesso di Dumas et al. [50]. Questa modellizzazione consente di ricavare analiticamente le coordinate spaziali del CoM globale come media delle posizioni dei CoM dei singoli segmenti, pesata sulle rispettive masse. Indicando con m_i e $\mathbf{r}_i$ la massa e la posizione del centro di massa dell'i-esimo segmento, la posizione del CoM globale è

$$\mathbf{r}_{CoM} = \frac{\sum_i m_i \mathbf{r}_i}{M}, \qquad M = \sum_i m_i. \tag{2.1}$$

Poiché ciascuna massa m_i è una percentuale della massa corporea totale, M si semplifica e la posizione del CoM dipende dalle sole percentuali dei segmenti. La massa corporea totale è stata ricavata dalla componente verticale F_z della forza misurata dalla pedana di forza ($M = F_z/g$ a soggetto fermo).

2.4.3 SAM 3D Body (3DB) per il sistema single-camera

Per il sistema single-camera è stato adottato SAM 3D Body (3DB), un modello che ricostruisce da una singola immagine la mesh 3D dell'intero corpo; il suo vantaggio rispetto ai modelli basati su keypoints risiede proprio nella ricostruzione del volume corporeo [39]. A differenza dei modelli basati su SMPL, 3DB adotta una nuova rappresentazione parametrica della mesh, il *Momentum Human Rig* (MHR), che separa la struttura scheletrica dalla forma della superficie corporea (tessuti molli) [39].

A livello architetturale, SAM 3D Body si distingue per un approccio *promptable* basato su *encoder-decoder*, che permette al sistema di ricevere input ausiliari (come maschere di segmentazione o coordinate 2D) per guidare l'inferenza, garantendo una stima simultanea e coerente della postura del corpo e delle estremità (mani e piedi) all'interno di un unico framework [39].

Nell'uso standard, SAM 3D Body ricava i parametri intrinseci della telecamera da uno stimatore automatico del campo visivo (*Field of View*, FoV) (e.g. MoGe-2) [39]. In questo lavoro, disponendo della calibrazione di laboratorio, i parametri intrinseci sono stati forniti direttamente al modello, senza ricorrere allo stimatore.

Per ogni fotogramma, dopo la correzione della distorsione con la calibrazione di laboratorio, il modello individua il soggetto e ne ricostruisce la mesh in coordinate della telecamera [39]. È proprio la ricostruzione del volume a consentire di ricavare il CoM direttamente dalla mesh, senza un modello geometrico a segmenti né coefficienti antropometrici: il CoM è il baricentro della mesh triangolata, assumendo densità uniforme su tutto il volume. La definizione è quindi diversa da quella del sistema multi-camera e comprende anche mani e piedi con la loro forma reale [39]. Le coordinate del CoM, espresse nel sistema della telecamera, sono poi portate nel sistema di riferimento globale, definito dalla scacchiera posta sulla pedana, mediante i parametri estrinseci della telecamera [64]. Per la stima monoculare la profondità e la scala sono quelle predette dal modello a partire dai parametri intrinseci forniti: non è stato introdotto alcun vincolo aggiuntivo, per esempio sulla statura del soggetto.

Per la pipeline single-camera è stata utilizzata una delle quattro videocamere del

sistema multi-camera (camera 19 in Figura 2.4), posta frontalmente al soggetto a circa 2.2 m di distanza orizzontale dall'origine del sistema di riferimento e a circa 2.3 m di altezza.

2.5 Protocollo

La pedana di forza è stata accesa almeno un'ora prima delle acquisizioni, con lo streaming dei dati attivo, per consentirne la stabilizzazione termica e limitare la deriva del segnale; successivamente, prima di ogni soggetto è stato azzerato l'offset a pedana scarica. Le videocamere sono state pre-calibrate con una scacchiera (*chessboard*) 11×8 con riquadri da 100 mm, utilizzando il software sviluppato in laboratorio, che si basa sull'implementazione OpenCV del metodo di Zhang [66, 64]: dalle immagini della scacchiera si ricavano le matrici dei parametri intrinseci (lunghezza focale, centro ottico e coefficienti di distorsione radiale e tangenziale) ed estrinseci (posizione e orientamento rispetto al sistema globale) di ciascuna videocamera. Con la scacchiera posta sopra la pedana si è inoltre definito il sistema di riferimento globale, con gli assi orientati in modo da farli corrispondere a quelli della pedana [64]. La calibrazione è stata eseguita una sola volta per l'intera campagna di acquisizione: i parametri estrinseci e intrinseci delle videocamere sono identici in tutte le sessioni.

2.5.1 Partecipanti

Hanno partecipato allo studio 16 volontari sani (11 donne e 5 uomini), di età compresa tra 21 e 30 anni. Tutti i partecipanti hanno firmato il consenso informato prima dell'acquisizione (studio approvato dal Comitato Etico dell'Università di Pavia, protocollo n. 230/25, giugno 2025). I volontari sono stati reclutati in base a criteri di inclusione quali una vista normale o corretta e la capacità di stare in piedi senza appoggio. Sono stati esclusi i soggetti con patologie neurologiche, muscolo-scheletriche, cardiovascolari o respiratorie, attuali o pregresse, con interventi ortopedici recenti agli arti inferio-

ri, con sintomi persistenti da traumi pregressi o con storia di commozione cerebrale o cinetosi [64]. Le caratteristiche del campione sono riportate in Tabella 2.3.

Variabile	Valore
Sesso (F/M)	11/5
Età (anni)	24.4 ± 2.2 (21–30)
Massa corporea (kg)	65.1 ± 12.1 (49–86)
Statura (cm)	169.9 ± 10.0 (152–190)
BMI (kg/m^2)	22.4 ± 2.9 (17.4–28.3)
Numero di piede	39.8 ± 3.3 (35–47)
Vestibilità degli abiti (normale/oversize/attillata)	12/2/2

Tabella 2.3: Caratteristiche dei 16 partecipanti. I valori continui sono riportati come media ± deviazione standard (minimo–massimo).

2.5.2 Informazioni Antropometriche

Oltre ai segnali di centro di pressione e di centro di massa, per ogni soggetto sono stati raccolti statura, massa corporea, indice di massa corporea (BMI), numero di piede, sesso, età e vestibilità degli abiti (Tabella 2.3).

Queste variabili sono state registrate perché le misure di oscillazione del CoP non sono indipendenti dalle dimensioni del corpo che le genera: in adulti sani, ad esempio, la lunghezza del percorso e l'area di oscillazione aumentano con la statura, mentre una misura normalizzata rispetto alla base di appoggio non varia, per cui parte delle differenze tra soggetti può essere un artefatto della taglia corporea [67]. L'effetto della massa corporea è invece documentato ma non univoco: nelle donne obese il range di oscillazione cambia in modo diverso nei due piani [68], mentre in un campione di adulti con BMI medio di 27.5 la correlazione tra BMI e oscillazione è debole [69]. Le stesse caratteristiche potrebbero influire anche sull'errore dei sistemi markerless rispetto alla

pedana di forza, perché morfologia e vestibilità degli abiti modificano la silhouette su cui si basano i modelli; gli studi che hanno confrontato abiti diversi riportano tuttavia differenze angolari contenute, circa 2.6° nel cammino e meno di 2.8° nella corsa [30, 31].

Il campione di questa tesi, composto da 16 giovani adulti di cui 12 con abiti di vestibilità normale, ha numerosità e variabilità limitate: per questo l'effetto di queste variabili sull'errore dei due sistemi non è stato analizzato.

2.5.3 Protocollo di Acquisizione (Task Sperimentali)

Ciascun soggetto ha eseguito cinque task motori, della durata di 30 s ciascuno, descritti di seguito.

L'ordine dei task è stato fisso per tutti i soggetti: *Occhi Aperti*, *Occhi Chiusi*, *Limite*, *One Foot* e *Tandem*. Ogni task è stato eseguito una volta, con 30 s di pausa tra un task e l'altro. Per tutti i task a parte *Occhi Chiusi* è stato chiesto al soggetto di fissare un punto posto sulla parete di fronte, a circa 2 m di distanza dalla pedana. Ogni soggetto ha mantenuto indossate le proprie calzature durante le prove. I piedi dei soggetti sono stati posizionati secondo dei riferimenti sulla pedana, con i talloni a circa 20 cm di distanza e le punte dei piedi orientate verso l'esterno di circa 30°.

Task 1 e 2: Quiet Standing (Postura Statica Bipodalica)

Il test è la prova di riferimento per la valutazione posturografica di base [14]. Il soggetto ha mantenuto la postura eretta sulla pedana, con i piedi posizionati in posizioni specifiche e le braccia lungo i fianchi (Figura 2.11).

- Esecuzione: la prova è stata eseguita a *Occhi Aperti* e a *Occhi Chiusi*, su superficie rigida [15].
- Obiettivo: verificare che il sistema markerless sia in grado di rilevare le micro-oscillazioni posturali in condizioni statiche, in cui il segnale utile è dello stesso ordine di grandezza del rumore di misura [55].



(a) *Occhi Aperti* **(b)** *Occhi Chiusi*

Figura 2.11: Esecuzione dei task di stazione eretta quieta sulla pedana di forza: a *Occhi Aperti* (a) e a *Occhi Chiusi* (b).

Task 3: Limits of Stability (Oscillazione Volontaria Massima)

La letteratura evidenzia la necessità di superare la sola analisi statica introducendo perturbazioni dinamiche volontarie [13].

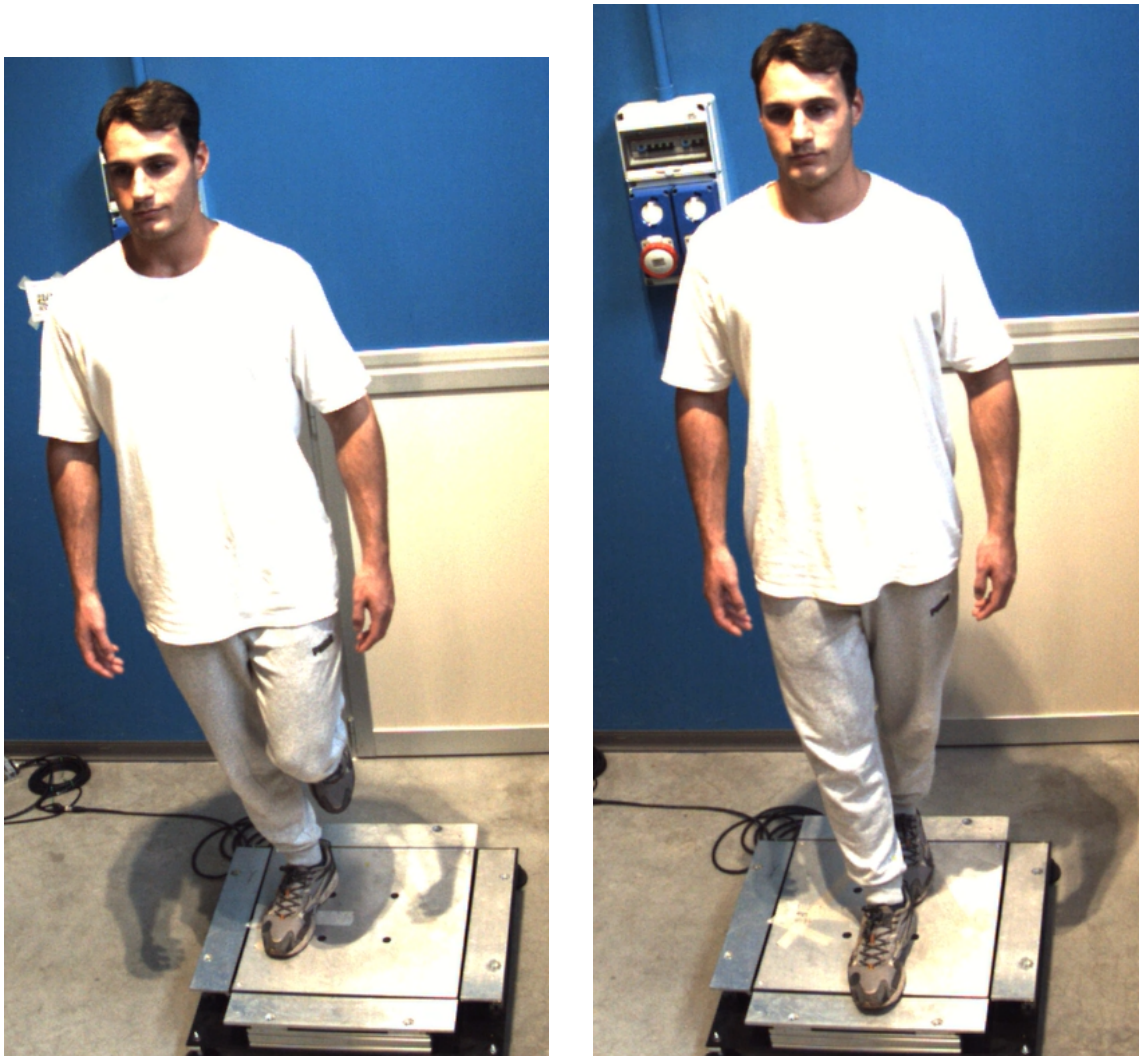
- Esecuzione: il soggetto ha oscillato una volta a sinistra partendo dal centro, rimanendo sbilanciato a sinistra per circa 15 s, per poi spostarsi a destra per il restante tempo del task, mantenendo il corpo rigido il più possibile come un pendolo, senza piegare ginocchia o bacino e mantenendo i piedi fermi (Figura 2.12).
- Obiettivo: ottenere ampie escursioni in direzione medio-laterale, mantenute per circa 15 s su ciascun lato, in cui il segnale utile è molto più grande del rumore di misura, e verificare il comportamento dei due sistemi in un movimento di grande ampiezza, complementare alle prove statiche.



Figura 2.12: Esecuzione del *Limite* (esempio di sbilanciamento laterale): il soggetto si inclina mantenendo i piedi fermi e il corpo rigido.

Task 4 e 5: Single-Leg Stance / Tandem Stance

- Esecuzione: nel *One Foot* il soggetto si è retto sulla sola gamba dominante; nel *Tandem* ha posizionato un piede esattamente davanti all'altro (tacco-punta), col piede destro dietro al sinistro. Ciascuna prova è durata 30 secondi mantenendo, per quanto fosse possibile da parte del soggetto, le braccia lungo i fianchi (Figura 2.13) [14].
- Obiettivo: ridurre la base di appoggio in direzione medio-laterale, generando oscillazioni più ampie e meno regolari, quindi più facilmente distinguibili dal rumore di misura.



(a) *One Foot* **(b)** *Tandem*

Figura 2.13: Esecuzione dei task a base di appoggio ridotta: *One Foot* (a) e *Tandem* (b).

2.5.4 Sessioni e Grandezze Confrontate

Ciascuno dei 16 soggetti ha eseguito i cinque esercizi, per un totale di 80 sessioni acquisite. Sono state escluse 26 sessioni in cui il numero di frame acquisiti differiva di più di 15 da quello atteso (1800, pari a 30 s a 60 fps) a causa di frame mancanti in acquisizione. Le sessioni mantenute perdono al più 15 frame, mentre quelle escluse da 64 a 1190, per cui qualsiasi soglia tra 15 e 63 porta alla stessa selezione. Gli errori sono dovuti a registrazioni terminate in anticipo oppure a frame non acquisiti durante la registrazione. Dal registro dell'acquisizione risulta che 9 sessioni escluse sono registrazioni terminate

in anticipo, 8 sono registrazioni complete con frame non acquisiti e 9 presentano entrambe le situazioni. Rispetto alle sessioni mantenute, quelle escluse hanno un tempo medio di acquisizione per frame più alto (mediana 5.7 contro 1.8 ms) e più frame in coda (8 contro 2); ciò indica un problema tecnico. Le 54 sessioni rimanenti sono state utilizzate in tutte le analisi: *Occhi Aperti* 13, *Occhi Chiusi* 10, *Limite* 12, *One Foot* 11, *Tandem* 8. L'esclusione delle sessioni non è uniforme tra gli esercizi: 8 delle 26 sessioni escluse riguardano il *Tandem*, 6 gli *Occhi Chiusi*, 5 il *One Foot*, 4 il *Limite* e 3 gli *Occhi Aperti*, ma non risulta associata in modo significativo al tipo di esercizio (test chi-quadro, $p = 0.38$): con 16 sessioni per esercizio il test avrebbe individuato un'associazione di questa entità solo in circa un terzo dei casi, per cui un legame con il compito non può essere escluso. I gruppi hanno quindi numerosità diversa (da 8 a 13 sessioni) e sono composti da soggetti in parte diversi.

Per ogni esercizio sono stati ottenuti tre tracciati: il CoP misurato dalla pedana di forza, assunto come riferimento (*gold standard*), e i due CoP stimati dal CoM, uno dalla pipeline multi-camera con YOLO26x e uno dalla pipeline single-camera con SAM 3D Body. Da ciascun tracciato sono stati calcolati gli indici descritti nella Sezione 2.6.4, con cui sono stati confrontati i due sistemi. Nel seguito, per brevità, la pedana di forza è indicata come *pedana* e i due sistemi markerless, multi-camera e single-camera, rispettivamente come *YOLO* e *SAM*. Il confronto con la pedana riguarda il CoP stimato dall'intera procedura (quindi dalla stima del CoM all'applicazione del filtraggio, della doppia derivazione e del modello a pendolo inverso). YOLO e SAM differiscono per modello di stima della posa, per numero di telecamere (quattro contro una) e per la definizione del CoM, dove YOLO lo definisce dal modello geometrico a solidi, mentre SAM come baricentro della mesh. Perciò il confronto riguarda due sistemi completi.

2.6 Elaborazione Dati

2.6.1 Calcolo del Centro di Pressione di Riferimento

Il Centro di Pressione di riferimento è stato calcolato dalle componenti di forza verticale F_z e dai momenti T_x, T_y misurati dalla pedana, secondo le relazioni classiche derivate dall'equilibrio dei momenti alla base d'appoggio:

$$CoP_y = \frac{T_x}{F_z}, \qquad CoP_x = -\frac{T_y}{F_z} \tag{2.2}$$

Il CoP così ottenuto costituisce il riferimento (*ground truth*) rispetto al quale sono stati confrontati i due sistemi di visione.

2.6.2 Stima del Centro di Pressione dal Centro di Massa

Modello del Pendolo Inverso

I sistemi video restituiscono la traiettoria del CoM e non direttamente quella del CoP: per mettere in relazione le due grandezze è stato applicato il modello del pendolo inverso a singolo segmento. Trattando il corpo come un pendolo invertito imperniato alla base d'appoggio, il bilancio dei momenti attorno al CoP fornisce la stima:

$$CoP_x(t) = CoM_x(t) - \frac{\bar{z}}{g}\, a_x(t) \tag{2.3}$$

$$CoP_y(t) = CoM_y(t) - \frac{\bar{z}}{g}\, a_y(t) \tag{2.4}$$

dove $a_x(t)$ e $a_y(t)$ sono le componenti orizzontali dell'accelerazione del CoM, $g = 9.81\ \mathrm{m/s^2}$ l'accelerazione di gravità e $\bar{z}$ la quota del CoM, assunta costante e pari alla media della sua componente verticale lungo la prova: durante le acquisizioni tale componente varia di pochi millimetri (deviazione standard mediana di circa 2 mm sulle 54 sessioni). Il termine correttivo tiene conto del fatto che, quando il corpo accelera orizzontalmente, il punto di applicazione della risultante delle forze a terra si discosta dalla proiezione ver-

ticale del CoM in misura proporzionale all'altezza del CoM stesso e all'accelerazione. L'equazione deriva dal bilancio dei momenti di un pendolo inverso [54, pp. 194–196]. Considerando il moto nel piano verticale che contiene l'asse x, orientato verso destra, misurando le posizioni a partire dal perno e assumendo positive le rotazioni antiorarie, il momento della reazione vincolare R, applicata nel CoP, e quello del peso W, applicato nel CoM, determinano l'accelerazione angolare α del pendolo, di momento d'inerzia I:

$$R\,CoP_x - W\,CoM_x = I\,\alpha. \tag{2.5}$$

Poiché la componente verticale del CoM varia di pochi millimetri, l'accelerazione verticale è trascurabile e $R = W$; inoltre, per piccole oscillazioni, l'accelerazione orizzontale del CoM è $a_x = -d\,\alpha$, con d distanza tra il perno e il CoM. Ne segue

$$CoP_x - CoM_x = -\frac{I}{W d}\,a_x, \tag{2.6}$$

cioè la differenza tra CoP e CoM è proporzionale all'accelerazione del CoM e di segno opposto, come riscontrato da Winter nei dati sperimentali (p. 196). Nella formulazione adottata il corpo è trattato come massa concentrata nel CoM: il momento d'inerzia vale $I = md^2$ e il peso $W = mg$, per cui $I/(Wd) = d/g$. Winter colloca il perno nell'articolazione della caviglia; qui il perno è invece assunto al livello della pedana, dove si trova l'origine del sistema di riferimento, per cui d coincide con la quota del CoM $\bar{z}$ e si ottiene l'Eq. 2.4. La stessa relazione vale lungo l'asse y. Il modello presuppone una strategia di caviglia, valida in postura quieta e per piccole perturbazioni [54, p. 198], in cui il corpo oscilla come un unico segmento rigido attorno alle caviglie. Questa ipotesi è meno adeguata sul piano mediolaterale, dove il controllo è affidato soprattutto ad abduttori e adduttori dell'anca [54, pp. 196–198], e nei task a base d'appoggio ridotta e nel *Limite*, in cui intervengono anche anca e tronco [70]: il corpo non si comporta più come un solo segmento e la stima è meno accurata. Di questo si tiene conto nella discussione dei limiti dello studio.

2.6.3 Analisi del Segnale

Per stimare il CoP dal CoM con il modello del pendolo inverso (Eq. 2.4) serve l'accelerazione del CoM, ottenuta derivando due volte la sua posizione. La posizione ricostruita dal video è però affetta da rumore di misura, e la derivazione numerica amplifica le alte frequenze (proporzionalmente alla frequenza per la velocità e al suo quadrato per l'accelerazione). Il CoM deve quindi essere filtrato prima di essere derivato. La frequenza di taglio del filtro non è stata fissata arbitrariamente, ma scelta a partire dall'analisi del segnale nel dominio del tempo e della frequenza, descritta nei passaggi seguenti.

Le figure di questa sezione si riferiscono a una sessione di esempio (soggetto S2, esercizio *Occhi Aperti*, 30 s acquisiti a 60 Hz), scelta tra quelle che rispettassero i criteri: assenza di frame mancanti o duplicati. Fanno eccezione le Figure 2.17, 2.18 e 2.19, che riassumono tutte le 54 sessioni valide (16 soggetti, 5 esercizi). I segnali sono espressi in millimetri e centrati sulla propria media, perché CoP e CoM sono grandezze diverse riferite a origini diverse; il centraggio non modifica gli indici utilizzati, che sono invarianti rispetto a una traslazione. Il primo secondo di ogni acquisizione è scartato, perché il CoM presenta un transitorio iniziale privo di significato posturale.

Segnale nel dominio del tempo

Il primo passo è l'osservazione del segnale grezzo. La Figura 2.14 riporta a sinistra le componenti $X(t)$ e $Y(t)$, rispettivamente sopra e sotto, e a destra la traiettoria nel piano (X, Y) (*path*), per il CoP della pedana, il CoM di YOLO e il CoM di SAM.

La figura mostra che il segnale della pedana è dominato da componenti a bassa frequenza, mentre ai segnali ricavati dai video è sovrapposto un rumore ad alta frequenza, che rende frastagliate le componenti e le traiettorie.

Segnale nel dominio della frequenza: spettro e PSD

Per separare il segnale utile dal rumore è stato effettuato lo studio nel dominio della frequenza, scomponendo ciascun segnale (centrato sulla media) nelle sue componenti sinusoidali con la trasformata di Fourier. Sono state calcolate due stime della distri-

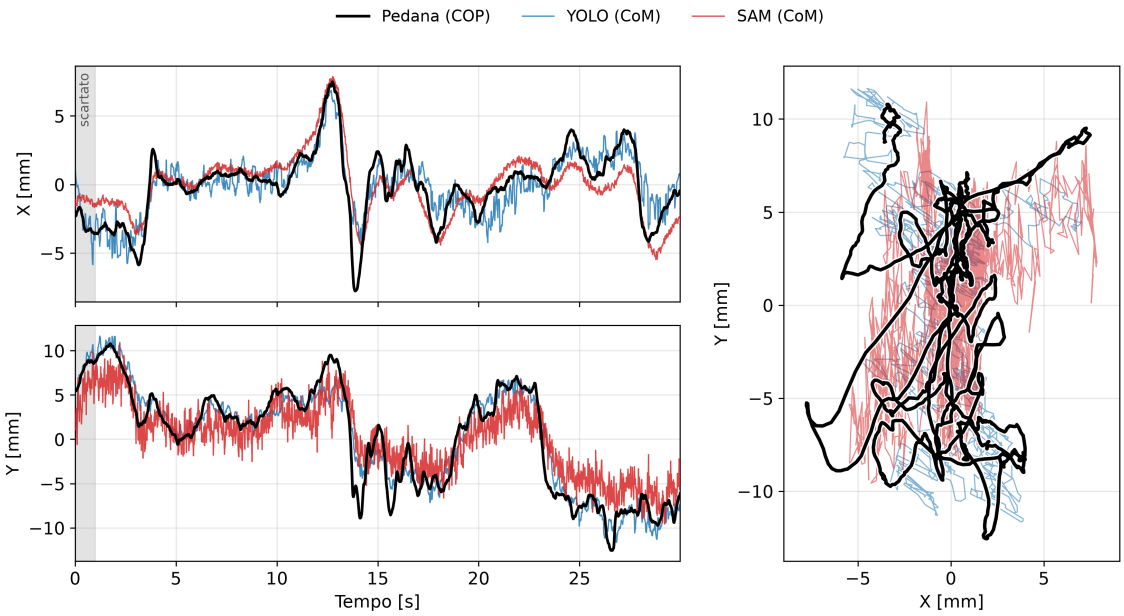


Figura 2.14: Segnale grezzo nel dominio del tempo (sessione S2, *Occhi Aperti*). A sinistra, le componenti $X(t)$ (sopra) e $Y(t)$ (sotto); a destra, la traiettoria (X, Y). Sono riportati il CoP della pedana (nero), il CoM di YOLO (blu) e il CoM di SAM (rosso), centrati sulla propria media. La zona grigia indica il primo secondo, scartato nelle analisi.

buzione della potenza in funzione della frequenza (Figura 2.15). La prima è lo spettro di potenza, calcolato sull'intero segnale con una finestra di Hann: ha la massima risoluzione in frequenza ma una forte variabilità da un punto all'altro. La seconda è la densità spettrale di potenza (PSD) con il metodo di Welch, che divide il segnale in segmenti sovrapposti da 5 s, ne calcola il periodogramma e ne fa la media: il grafico mostra una curva meno rumorosa, a scapito della risoluzione (0.2 Hz), e permette di leggere l'andamento medio della potenza.

Nella figura la potenza si concentra alle basse frequenze e decresce al crescere della frequenza. Ad alta frequenza i segnali ricavati dai sistemi di visione si appiattiscono su un livello di rumore (*noise floor*), mentre la pedana continua a decadere. Il punto in cui la PSD si appiattisce indica dove il rumore prevale sul segnale utile.

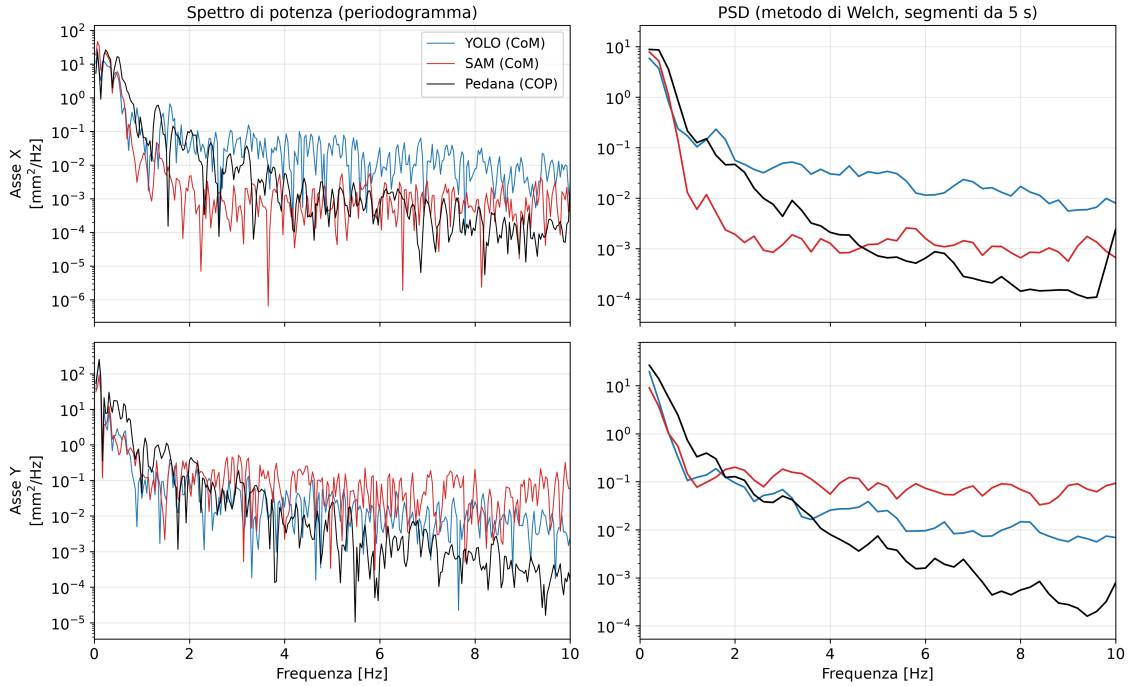


Figura 2.15: Spettro di potenza (periodogramma, a sinistra) e PSD di Welch con segmenti da 5 s (a destra) per l'asse X (sopra) e l'asse Y (sotto), per il CoP della pedana, il CoM di YOLO e il CoM di SAM (sessione S2, *Occhi Aperti*). La potenza è in scala logaritmica.

Potenza cumulata e frequenza f_{70}

Per quantificare la banda occupata dal segnale si calcola la potenza cumulata, ossia la frazione della potenza totale contenuta al di sotto di ciascuna frequenza:

$$C(f) = \frac{\sum_{f_k \le f} P(f_k)}{\sum_k P(f_k)}, \tag{2.7}$$

dove $P(f_k)$ è la PSD. Si definisce f_p la frequenza tale che $C(f_p) = p$. Si è usato $p = 0.7$: f_{70} è la frequenza sotto cui cade il 70% della potenza e fornisce una misura sintetica della banda occupata dal segnale. Si è scelto il 70% invece dell'80% perché f_{80} è instabile in presenza di rumore bianco. In questo calcolo la PSD è stimata con segmenti da 20 s, per avere una risoluzione in frequenza di 0.05 Hz. La Figura 2.16 riporta le curve $C(f)$ e i valori di f_{70}. Nella sessione di esempio f_{70} è compresa tra 0.15 e 0.60 Hz per tutti i segnali: la potenza dei movimenti posturali è concentrata a frequenze molto basse, tutte ben inferiori a 1 Hz, e questo dà l'ordine di grandezza della banda del segnale utile, che

fa da riferimento per le frequenze di taglio valutate nel seguito. La scelta del 70% si vede bene sull'asse Y di SAM, dove la curva supera il 70% poco sopra 0.6 Hz ma raggiunge l'80% solo intorno a 3 Hz, per il rumore presente alle frequenze più alte: con $p = 0.8$ la frequenza ottenuta sarebbe stata determinata dal rumore. f_{70} non fissa quindi il taglio, che è scelto con le tre analisi descritte di seguito.

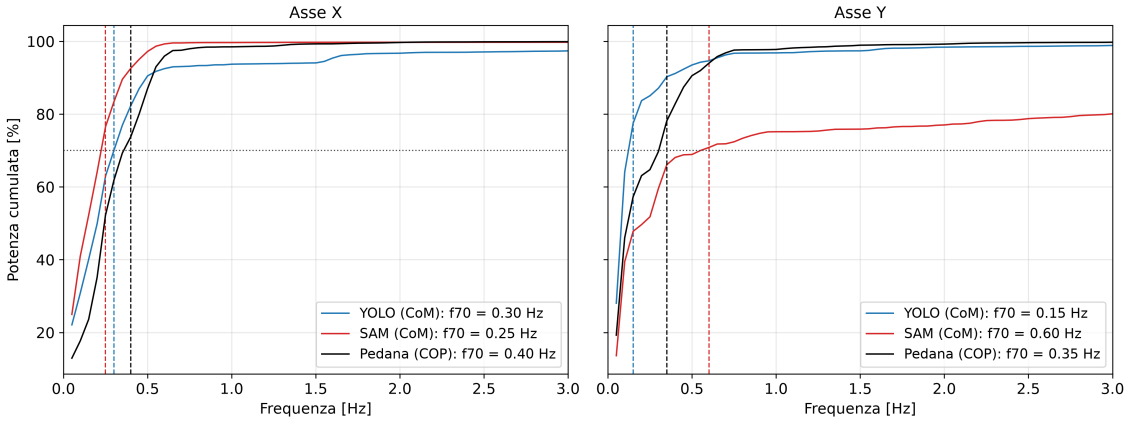


Figura 2.16: Potenza cumulata in funzione della frequenza per l'asse X (a sinistra) e per l'asse Y (a destra), per il CoP della pedana, il CoM di YOLO e il CoM di SAM (sessione S2, *Occhi Aperti*). Le linee verticali tratteggiate indicano f_{70} di ciascun segnale; la linea orizzontale punteggiata indica il 70%.

Scelta della frequenza di taglio

La frequenza di taglio è stata scelta confrontando tre analisi complementari, riportate nelle tre figure che seguono.

Analisi dei residui di Winter. Il metodo di Winter [49] filtra il segnale con frequenze di taglio f_c crescenti e calcola, per ciascuna, il residuo, ossia lo scarto quadratico medio tra il segnale originale x_i e quello filtrato $\hat{x}_i(f_c)$:

$$R(f_c) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \hat{x}_i(f_c))^2}. \tag{2.8}$$

Il residuo diminuisce al crescere di f_c; alle alte frequenze dipende solo dal rumore e diventa approssimativamente lineare. Estrapolando questo tratto lineare fino a $f_c = 0$ si ottiene un livello di riferimento del rumore, e la frequenza di taglio di Winter è quella

in cui il residuo raggiunge tale livello. La Figura 2.17 mostra l'analisi per la sessione di esempio (a sinistra) e la distribuzione dei tagli di Winter su tutte le sessioni (a destra).

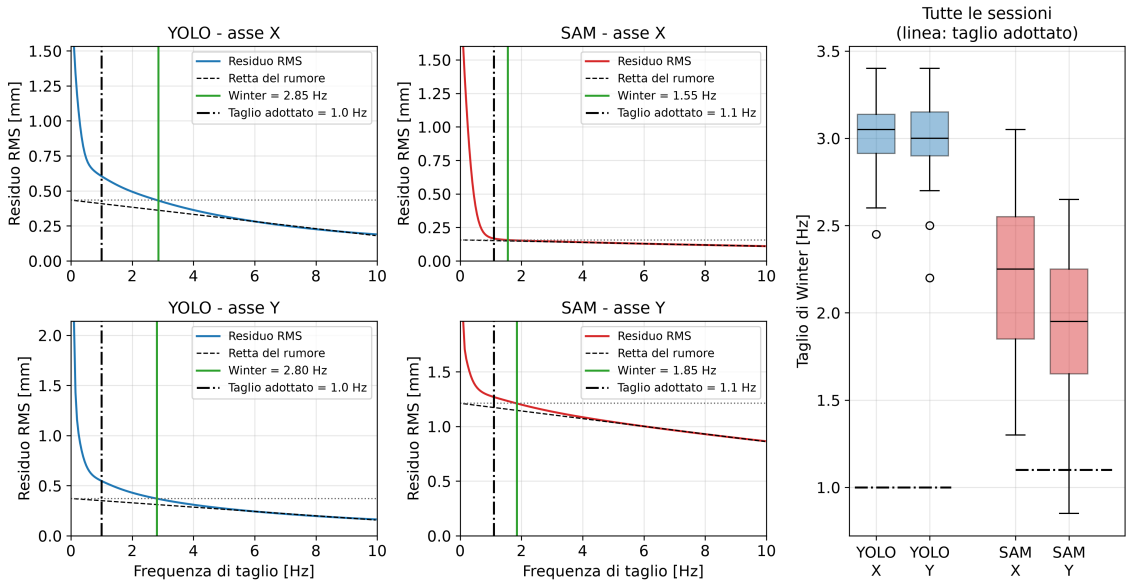


Figura 2.17: Analisi dei residui di Winter. A sinistra, per la sessione S2 (*Occhi Aperti*), residuo RMS in funzione della frequenza di taglio per YOLO e SAM e per gli assi X e Y, con la retta del rumore (tratteggiata), il taglio di Winter (verde) e il taglio adottato (nero). A destra, distribuzione (*box plot*) del taglio di Winter su tutte le sessioni; le linee orizzontali indicano i tagli adottati.

Il taglio di Winter indica fino a quale frequenza la *posizione* del CoM contiene segnale. Il modello del pendolo inverso richiede però l'*accelerazione*, che amplifica il rumore come il quadrato della frequenza: il taglio di Winter va quindi considerato un limite superiore per il taglio da utilizzare.

Coerenza tra CoM e CoP. Il secondo criterio verifica fino a quale frequenza il CoM ricavato dal video contiene informazione sul CoP misurato dalla pedana; i due segnali si riferiscono infatti allo stesso soggetto e agli stessi istanti. A questo scopo si è calcolata la coerenza quadratica tra i due segnali:

$$\gamma_{xy}^2(f) = \frac{|P_{xy}(f)|^2}{P_{xx}(f)\,P_{yy}(f)}, \tag{2.9}$$

dove P_{xy} è la densità spettrale incrociata tra CoM e CoP e P_{xx}, P_{yy} sono le rispettive PSD. Per ogni frequenza la coerenza indica quanto le oscillazioni del CoM a quella frequenza

siano una copia, in fase, di quelle del CoP: vale 1 quando i due segnali sono linearmente legati e 0 quando non c'è relazione. Il calcolo è stato eseguito con segmenti da 5 s su tutte le sessioni e su entrambi gli assi.

Poiché la stima è ottenuta su un numero finito di segmenti, anche due segnali indipendenti presentano una coerenza casuale maggiore di zero. Si è quindi utilizzata una soglia di significatività al 95%, pari a 0.27 con i circa 11 segmenti mediati: è il valore che segnali non correlati superano solo nel 5% dei casi, e al di sotto del quale la coerenza non è distinguibile da zero. La Figura 2.18 riporta la mediana della coerenza e l'intervallo tra il 25° e il 75° percentile, insieme alla soglia.

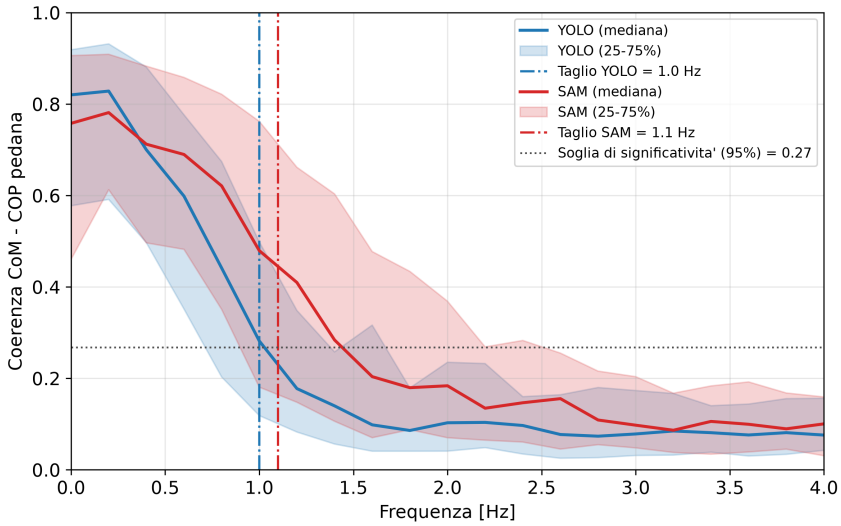


Figura 2.18: Coerenza tra il CoM (YOLO e SAM) e il CoP della pedana in funzione della frequenza: mediana e intervallo 25–75% su tutte le sessioni e su entrambi gli assi. La linea punteggiata indica la soglia di significatività al 95%; le linee verticali indicano i tagli adottati.

La frequenza a cui la coerenza scende sotto la soglia indica il limite oltre il quale il CoM non contiene più informazione sul CoP: un taglio superiore a tale frequenza lascerebbe passare solo rumore, che la derivazione seconda amplifica. Tale frequenza costituisce quindi un limite superiore per il taglio, e non ne fornisce il valore esatto. La riduzione della coerenza non dipende infatti solo dal rumore del video, ma anche dal fatto che il CoM, per la sua natura fisica, non segue le oscillazioni più rapide del CoP.

Errore degli indici in funzione del taglio. Il terzo criterio non analizza i segnali, ma il risultato finale della catena di elaborazione: l'accordo degli indici posturografici del CoP stimato con quelli della pedana. Per ogni sessione e per ogni sistema, per ciascuna frequenza di taglio da 0.4 a 2.6 Hz (passo 0.1 Hz) si è ripetuta la procedura seguente:

1. filtraggio del CoM con quel taglio;
2. derivazione due volte del segnale filtrato e applicazione del modello del pendolo inverso (Eq. 2.4), per ottenere il CoP stimato;
3. calcolo di Sway Area e Path Length del CoP stimato e del CoP della pedana;
4. calcolo dell'errore percentuale di ciascun indice, $100\,(I_{\text{stimato}} - I_{\text{pedana}})/I_{\text{pedana}}$.

Ogni taglio è stato riassunto con un unico punteggio: la media dei valori assoluti dei due errori percentuali, di cui si è presa la mediana sulle sessioni. Più il punteggio è basso, più il CoP stimato riproduce gli indici della pedana. I due indici sono stati scelti perché sensibili a bande di frequenza diverse: la Sway Area dipende soprattutto dall'ampiezza delle oscillazioni, cioè dalle basse frequenze, mentre il Path Length dipende molto dalle alte frequenze, perché il rumore allunga il percorso. Se il taglio è troppo basso il segnale risulta eccessivamente filtrato e gli indici sono sottostimati; se è troppo alto passa rumore che l'accelerazione amplifica con il quadrato della frequenza e gli indici sono sovrastimati.

La Figura 2.19 riporta a sinistra il punteggio in funzione del taglio, con l'intervallo tra il 25° e il 75° percentile e il minimo evidenziato, e a destra gli stessi errori percentuali con il segno, che permettono di vedere se l'indice è sovrastimato o sottostimato. Si è scelto questo punteggio, e non l'errore quadratico medio punto per punto tra le traiettorie, perché quest'ultimo tende a premiare tagli molto bassi: un segnale eccessivamente filtrato coincide più facilmente con un altro segnale filtrato, ma a costo di perdere l'oscillazione reale.

Taglio adottato. Dal confronto dei tre criteri si è adottata una frequenza di taglio per sistema, uguale per tutti gli esercizi e per tutti i soggetti: 1.0 Hz per YOLO e 1.1 Hz per

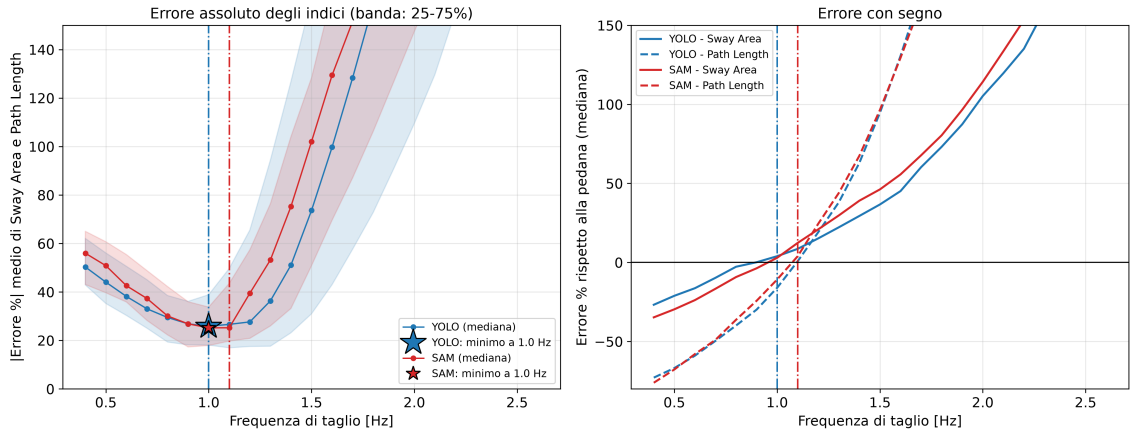


Figura 2.19: Errore degli indici rispetto alla pedana in funzione della frequenza di taglio, mediana su tutte le sessioni. A sinistra, valore assoluto medio dell'errore percentuale di Sway Area e Path Length (banda: 25–75%; le stelle indicano i minimi). A destra, errore percentuale con segno di Sway Area (linea continua) e Path Length (linea tratteggiata). Le linee verticali indicano i tagli adottati.

SAM. In entrambi i casi il valore è compatibile con il minimo del punteggio di Figura 2.19, è coerente con la frequenza a cui la coerenza tra CoM e CoP si esaurisce (Figura 2.18) ed è inferiore ai tagli di Winter (Figura 2.17), come richiesto dall'uso dell'accelerazione. Il taglio è stato scelto sugli stessi dati su cui si valuta poi l'accordo nel Capitolo 3, il che tende a favorire il risultato. Per verificare di quanto, la scelta è stata ripetuta 16 volte, escludendo ogni volta un soggetto (validazione incrociata *leave-one-subject-out*): il taglio è stato scelto sugli altri 15 soggetti, usando il punteggio di Figura 2.19, e poi applicato alle sessioni del soggetto escluso. I tagli ottenuti sono compresi tra 0.9 e 1.2 Hz per YOLO (1.0 Hz in 7 prove su 16) e tra 1.0 e 1.1 Hz per SAM (in tutte le 16 prove), cioè intorno al valore adottato. Sulle sessioni dei soggetti esclusi l'errore percentuale mediano su Sway Area e Path Length (media dei due valori assoluti) è 31.7% per YOLO, contro 25.8% con il taglio scelto su tutti i soggetti, e 25.8% contro 25.2% per SAM. La scelta fatta sugli stessi dati favorisce dunque il risultato di circa 6 punti percentuali per YOLO e di meno di 1 per SAM. Con la media al posto della mediana lo scarto sarebbe maggiore (circa 14 punti per YOLO e 3 per SAM), per il peso di pochi soggetti con errori elevati. La verifica non sostituisce una validazione su dati indipendenti, che richiederebbe un secondo gruppo di soggetti non usato per la scelta.

Filtraggio del segnale

Il filtro utilizzato è un Butterworth passa-basso del secondo ordine, applicato in entrambe le direzioni temporali per ottenere un filtraggio a fase zero (*filtfilt*). La risposta in ampiezza del filtro di Butterworth è $|H(f)| = 1/\sqrt{1 + (f/f_c)^{2n}}$, con $n = 2$ ordine del filtro; la doppia applicazione annulla lo sfasamento, evitando di traslare il segnale nel tempo, ed eleva al quadrato il modulo della risposta (attenuazione di -6 dB, invece di -3 dB, alla frequenza di taglio nominale). Il filtro è applicato alle componenti X e Y del CoM, dopo aver scartato il primo secondo, con frequenza di taglio di 1.0 Hz per YOLO e di 1.1 Hz per SAM. Il CoP della pedana non è filtrato.

La Figura 2.20 riporta il segnale nel dominio del tempo dopo il filtraggio, con la stessa configurazione della Figura 2.14.

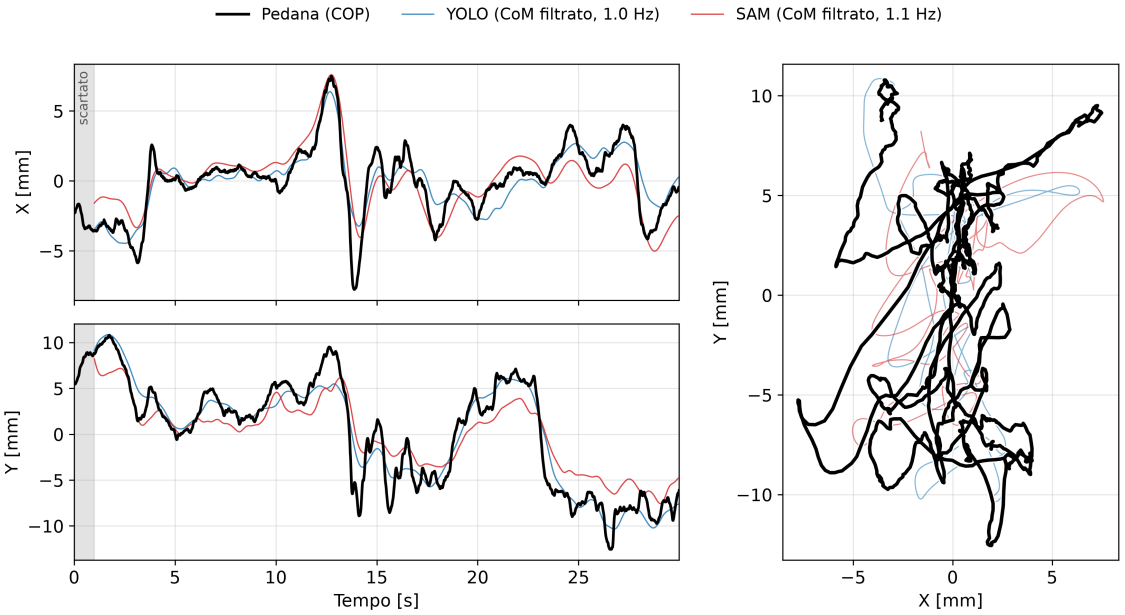


Figura 2.20: Segnale nel dominio del tempo dopo il filtraggio del CoM (YOLO a 1.0 Hz, SAM a 1.1 Hz), con la stessa configurazione della Figura 2.14. A sinistra $X(t)$ e $Y(t)$, a destra la traiettoria (X, Y); il CoP della pedana (nero) non è filtrato. Sessione S2, *Occhi Aperti*; il primo secondo (zona grigia) è scartato.

Le Figure 2.21 e 2.22 mostrano, per YOLO e SAM, il confronto tra il CoM grezzo e il CoM filtrato in posizione, velocità e accelerazione, dove velocità e accelerazione sono ottenute derivando numericamente (differenze finite) rispettivamente una e due volte

la posizione. In ciascun grafico dell'accelerazione è riportato l'RMS dell'accelerazione prima e dopo il filtraggio.

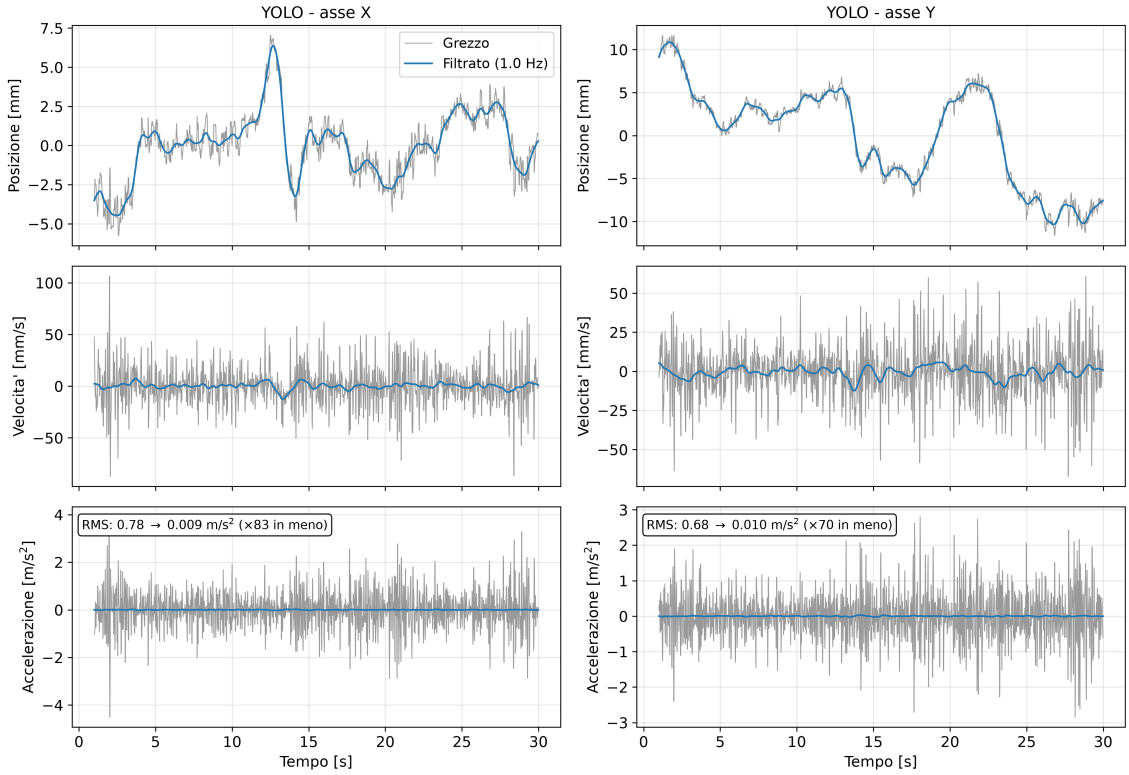


Figura 2.21: YOLO: posizione (sopra), velocità (al centro) e accelerazione (sotto) del CoM lungo X (a sinistra) e Y (a destra), grezze (grigio) e filtrate a 1.0 Hz (blu). Nei riquadri, l'RMS dell'accelerazione prima e dopo il filtraggio. Sessione S2, *Occhi Aperti*.

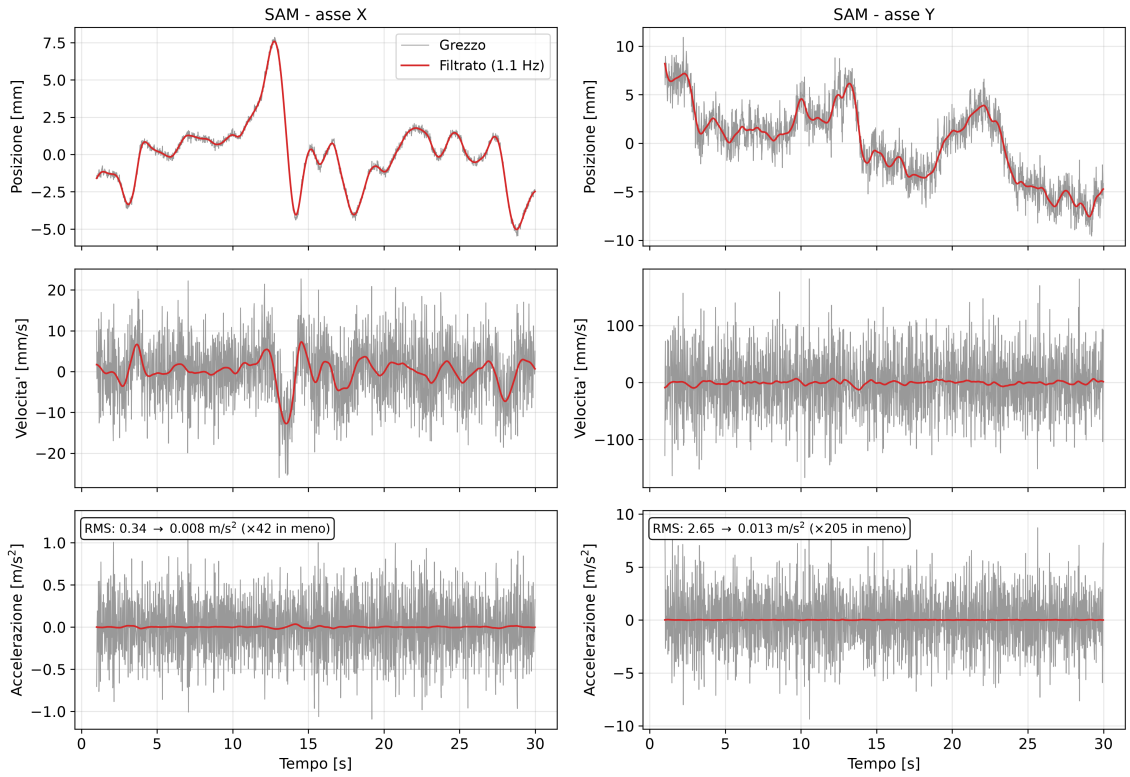


Figura 2.22: SAM: posizione (sopra), velocità (al centro) e accelerazione (sotto) del CoM lungo X (a sinistra) e Y (a destra), grezze (grigio) e filtrate a 1.1 Hz (rosso). Nei riquadri, l'RMS dell'accelerazione prima e dopo il filtraggio. Sessione S2, *Occhi Aperti*.

Il segnale così ottenuto (posizione filtrata e relative derivate) è quello utilizzato come ingresso del modello del pendolo inverso.

2.6.4 Indici Posturografici

Per ciascun tracciato (CoP misurato e CoP stimato da ciascuno dei due sistemi video) sono stati calcolati i seguenti indici standard di analisi posturografica:

- Sway Area: area dell'ellisse di confidenza al 95% dei punti (x, y), ottenuta dagli autovalori λ_1, λ_2 della matrice di covarianza campionaria e dal quantile $\chi^2_{0.95,2}$ della distribuzione chi-quadro a 2 gradi di libertà:

$$A = \pi \, \chi^2_{0.95,2} \sqrt{\lambda_1 \lambda_2} \tag{2.10}$$

- Path Length: lunghezza totale del percorso, calcolata come somma delle distanze euclidee tra campioni consecutivi:

$$L = \sum_{i=1}^{n-1} \sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2} \tag{2.11}$$

- Mean Velocity: rapporto tra la Path Length e la durata dell'analisi (circa 29 s, dopo lo scarto del primo secondo); poiché la durata è la stessa per tutti i sistemi, il suo errore percentuale coincide con quello della Path Length;
- Range_{ML}, Range_{AP}: escursione totale (massimo meno minimo) lungo ciascun asse;
- RMS_{ML}, RMS_{AP}: deviazione standard campionaria della posizione lungo ciascun asse, quale misura dell'ampiezza tipica dell'oscillazione attorno al proprio valore medio.

2.6.5 Calcolo degli Errori e Aggregazione tra Soggetti

Per ciascun indice (Sway Area, Path Length, Mean Velocity, Range e RMS) è stato calcolato l'errore assoluto, come differenza tra il valore ottenuto da ciascun sistema video e il valore di riferimento misurato dalla pedana, e il corrispondente errore percentuale, normalizzato rispetto al valore di riferimento. La normalizzazione percentuale è necessaria poiché la grandezza assoluta degli indici varia molto tra le diverse tipologie di esercizio, fino a uno o due ordini di grandezza per Sway Area e per Range e RMS mediolaterali (il *Limite* produce escursioni fisiologicamente più ampie di una prova di stazione eretta quieta), rendendo il solo errore assoluto non direttamente confrontabile tra compiti motori diversi.

Va precisato che Sway Area, Path Length, Mean Velocity, Range e RMS sono per costruzione invarianti rispetto a una traslazione rigida del tracciato: un'eventuale differenza nell'origine spaziale tra il sistema di riferimento della pedana e quello di ciascuna pipeline video non influisce quindi né sul valore di questi indici né, di conseguenza, sugli errori calcolati a partire da essi.

I risultati per singola sessione sono stati aggregati per tipologia di esercizio riportando, su tutti i soggetti che hanno svolto quel compito motorio, la mediana e lo scarto interquartile (IQR) di ciascun indice e del relativo errore, assoluto e percentuale. Si sono preferite queste statistiche a media e deviazione standard perché, con gruppi di 8–13 sessioni, le distribuzioni degli errori sono asimmetriche e contengono valori anomali, per cui la media risulta poco rappresentativa, come mostrano i box plot del Capitolo 3 (Figure 3.1, 3.2 e 3.3).

2.6.6 Validazione Statistica (Markerless vs Pedana)

L'accordo tra ciascun sistema markerless e la pedana è stato valutato mediante:

- Coefficiente di Correlazione (Pearson e Spearman): calcolato tra le serie temporali del CoP stimato e del CoP misurato, separatamente lungo le direzioni ML e AP, per valutare la concordanza dell'andamento dei due segnali. Nella presentazione dei risultati si riporta, per ciascun esercizio, anche il numero di sessioni con correlazione inferiore a 0.5.
  - Pearson: misura la forza e la direzione della relazione lineare tra due variabili quantitative. Assume valori tra -1 e 1: valori prossimi a 1 indicano una relazione lineare diretta, valori prossimi a -1 una relazione lineare inversa. Valori prossimi a 0 indicano l'assenza di relazione lineare.
  - Spearman: è calcolato sui ranghi dei dati anziché sui valori; non richiede quindi né la normalità dei dati né una relazione lineare, ma valuta se la relazione tra le due variabili è monotona, cioè se all'aumentare dell'una l'altra tende ad aumentare (o a diminuire) in modo costante. È inoltre più robusto agli outlier.

  La correlazione misura la somiglianza dell'andamento, non l'accordo in ampiezza: una correlazione alta può accompagnarsi a errori importanti nell'ampiezza. Per questo l'analisi è affiancata dagli errori sugli indici e dai grafici di Bland-Altman.

- RMSE di traiettoria (ML e AP): radice dell'errore quadratico medio istante per istante tra il CoP stimato e quello misurato, calcolata lungo ciascun asse dopo aver centrato ciascun segnale sulla propria media. La centratura serve per annullare eventuali effetti causati da una traslazione costante sulla misura. Quantifica l'errore medio sulla traiettoria stimata [20, 71].

- Grafici di Bland-Altman: per l'individuazione di eventuali bias sistematici introdotti dall'algoritmo di stima [20]. Per ciascun indice, ogni sessione è rappresentata da un punto: in ascissa la media tra il valore stimato dal sistema markerless e quello misurato dalla pedana, in ordinata la loro differenza (sistema markerless meno pedana). I grafici sono costruiti separatamente per YOLO26x e per SAM 3D Body.

Capitolo 3

Risultati

I risultati sono calcolati sulle 54 sessioni valide del dataset (16 soggetti, 5 esercizi: *Occhi Aperti* $n = 13$, *Occhi Chiusi* $n = 10$, *Limite* $n = 12$, *One Foot* $n = 11$, *Tandem* $n = 8$), con il filtraggio e il taglio descritti nel Capitolo 2 (Butterworth ordine 2, 1.0 Hz per YOLO e 1.1 Hz per SAM). Nel seguito, YOLO indica la pipeline multi-camera con YOLO26x e SAM la pipeline single-camera con SAM 3D Body. Viene quindi valutato il CoP stimato dall'intera procedura descritta nel Capitolo 2 con due sistemi completi, non direttamente il CoM. Come presentato nella Sezione 2.6.5, i valori nelle tabelle sono riportati come mediana [scarto interquartile] perché le distribuzioni degli errori sono asimmetriche.

3.1 Valori degli indici posturografici

La Tabella 3.1 e la Tabella 3.2 confrontano, per ciascun esercizio, i valori assoluti degli indici ottenuti dalla pedana, da YOLO e da SAM. Le tabelle permettono di verificare che i valori ottenuti dai due sistemi video siano compatibili con quelli della pedana, prima di analizzare gli errori percentuali.

Esercizio	Metodo	n	Sway Area [mm^2]	Path Length [mm]	Mean Velocity [mm/s]
Occhi Aperti	Pedana	13	169 [191]	276 [83]	9.5 [2.8]
	YOLO	13	199 [214]	345 [75]	11.9 [2.6]
	SAM	13	222 [252]	425 [275]	14.7 [9.5]
Occhi Chiusi	Pedana	10	73 [147]	315 [120]	10.9 [4.1]
	YOLO	10	106 [111]	401 [77]	13.8 [2.7]
	SAM	10	88 [109]	393 [69]	13.6 [2.4]
Limite	Pedana	12	15966 [5289]	1166 [525]	40.2 [18.1]
	YOLO	12	18118 [7234]	919 [283]	31.7 [9.7]
	SAM	12	28864 [16679]	1230 [588]	42.4 [20.2]
One Foot	Pedana	11	839 [431]	1329 [382]	45.8 [13.2]
	YOLO	11	1122 [554]	923 [549]	31.9 [18.9]
	SAM	11	889 [450]	913 [369]	31.5 [12.7]
Tandem	Pedana	8	770 [386]	1081 [605]	37.3 [20.9]
	YOLO	8	668 [740]	678 [119]	23.4 [4.1]
	SAM	8	909 [280]	886 [265]	30.6 [9.1]

Tabella 3.1: Valori mediani [scarto interquartile] di Sway Area, Path Length e Mean Velocity, per esercizio e per metodo (Pedana, YOLO, SAM).

Esercizio	Metodo	n	Range$_{ML}$ [mm]	Range$_{AP}$ [mm]	RMS$_{ML}$ [mm]	RMS$_{AP}$ [mm]
Occhi Aperti	Pedana	13	8.6 [9.1]	22.2 [11.2]	1.9 [2.5]	4.9 [4.4]
	YOLO	13	12.1 [7.0]	25.6 [21.0]	2.5 [2.0]	5.2 [4.3]
	SAM	13	12.5 [12.2]	21.6 [15.2]	3.0 [2.8]	4.4 [4.3]
Occhi Chiusi	Pedana	10	6.3 [5.1]	22.5 [10.2]	1.1 [1.1]	4.2 [1.3]
	YOLO	10	8.9 [6.6]	22.2 [6.6]	1.5 [1.0]	4.3 [1.3]
	SAM	10	7.6 [7.1]	19.9 [1.2]	1.8 [1.8]	3.8 [0.6]
Limite	Pedana	12	315.3 [39.9]	49.1 [20.2]	129.8 [26.1]	8.3 [3.8]
	YOLO	12	311.0 [43.9]	51.7 [22.9]	132.4 [23.7]	11.1 [4.4]
	SAM	12	344.3 [51.0]	97.5 [26.6]	147.9 [20.1]	26.0 [6.0]
One Foot	Pedana	11	32.7 [7.3]	48.4 [14.4]	5.7 [1.5]	8.7 [1.8]
	YOLO	11	26.1 [8.1]	56.3 [23.0]	5.6 [1.6]	11.4 [4.7]
	SAM	11	30.4 [8.5]	48.3 [26.0]	6.1 [1.0]	9.1 [2.2]
Tandem	Pedana	8	37.4 [4.5]	32.3 [12.6]	6.2 [1.6]	5.0 [2.4]
	YOLO	8	30.0 [7.3]	27.2 [32.0]	5.4 [2.0]	5.2 [6.0]
	SAM	8	36.4 [5.6]	40.0 [20.0]	6.3 [1.2]	6.9 [3.9]

Tabella 3.2: Valori mediani [scarto interquartile] di Range e RMS lungo gli assi mediolaterale e antero-posteriore, per esercizio e per metodo (Pedana, YOLO, SAM).

I valori dei due sistemi video sono dello stesso ordine di grandezza di quelli della pedana in tutti gli esercizi. Per tutti e tre i metodi gli indici seguono l'andamento atteso: sono minimi nelle prove a occhi aperti e a occhi chiusi e aumentano nelle prove più impegnative. Per la pedana, ad esempio, la Sway Area passa da 73–169 mm^2 nelle prove statiche a circa 800 mm^2 in *Tandem* e *One Foot* e a circa 16000 mm^2 in *Limite*.

Poiché dopo le esclusioni i gruppi sono composti da soggetti in parte diversi, i valori di esercizi simili non sono direttamente confrontabili tra loro: ad esempio, la Sway Area

della pedana risulta minore in *Occhi Chiusi* che in *Occhi Aperti*.

3.2 Errori rispetto alla pedana

Per ciascun indice si riporta l'errore percentuale rispetto alla pedana (Sezione 2.6.5), positivo quando il sistema sovrastima, calcolato per ogni sessione e sintetizzato con la mediana e lo scarto interquartile sulle sessioni valide di ciascun esercizio (Tabella 3.3 e Tabella 3.4). Per questo motivo gli errori percentuali non coincidono con le differenze tra le mediane delle Tabelle 3.1 e 3.2: ad esempio, per YOLO a occhi aperti le mediane di Sway Area differiscono del 18%, mentre l'errore percentuale mediano è +1.5%. L'indice Mean Velocity non è riportato separatamente: essendo Path Length diviso per la durata dell'acquisizione, che è la stessa per pedana e sistema video nella stessa sessione, il suo errore percentuale coincide con quello di Path Length.

La mediana indica l'errore sistematico di ciascun sistema, mentre lo scarto interquartile descrive quanto l'errore varia da una sessione all'altra. Come atteso, questa variabilità dipende dall'ampiezza del movimento. Nelle prove a occhi aperti e a occhi chiusi, in cui le oscillazioni sono di pochi millimetri e confrontabili con il rumore di tracciamento, lo scarto interquartile è ampio anche quando l'errore mediano è prossimo a zero (per YOLO a occhi aperti, Sway Area +1.5% [54.9]). Nel test dei limiti, in cui lo spostamento mediolaterale supera i 300 mm, è invece contenuto (per YOLO, Range_{ML} −0.6% [5.0]). Il risultato è coerente con il problema del rapporto segnale-rumore discusso nei capitoli precedenti.

Esercizio	Sistema	n	Sway Area (%)	Path Length (%)
Occhi Aperti	YOLO	13	1.5 [54.9]	21.3 [41.9]
	SAM	13	9.6 [28.8]	56.4 [38.3]
Occhi Chiusi	YOLO	10	27.6 [70.3]	14.9 [54.9]
	SAM	10	7.4 [26.1]	23.3 [32.3]
Limite	YOLO	12	9.0 [10.5]	-30.7 [15.6]
	SAM	12	61.9 [57.6]	-0.8 [13.6]
One Foot	YOLO	11	0.2 [48.8]	-35.0 [32.6]
	SAM	11	-3.7 [24.6]	-36.0 [27.6]
Tandem	YOLO	8	-16.4 [25.7]	-42.7 [27.3]
	SAM	8	13.0 [44.6]	-24.6 [21.7]

Tabella 3.3: Errore percentuale mediano [scarto interquartile] di Sway Area e Path Length rispetto alla pedana, per esercizio e per sistema.

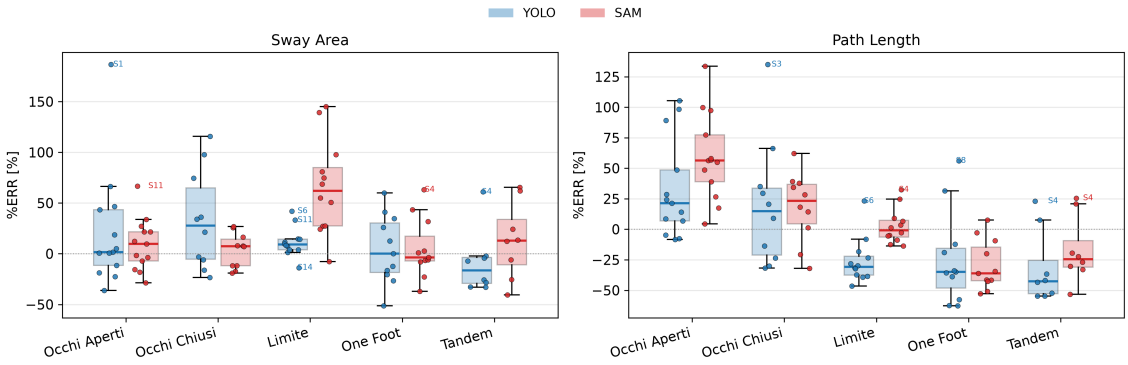


Figura 3.1: Distribuzione completa dell'errore percentuale di Sway Area e Path Length (stessi dati della Tabella 3.3): un punto per sessione, box con mediana e scarto interquartile, per esercizio e per sistema. I punti oltre i baffi (1.5 × IQR) sono etichettati con l'ID del soggetto.

Il segno dell'errore dipende dall'esercizio più che dal sistema. La Path Length è sovrastimata da entrambi i sistemi nelle prove a occhi aperti e a occhi chiusi ed è sot-

tostimata in *One Foot* e *Tandem*; in *Limite* è sottostimata da YOLO (−30.7%), mentre per SAM l'errore è trascurabile (−0.8%). L'andamento è atteso, perché la Path Length è l'indice più sensibile al contenuto in frequenza del segnale. Nelle prove statiche l'oscillazione reale è piccola e il rumore residuo, amplificato dalla doppia derivazione, allunga il percorso stimato. Nelle prove più impegnative il CoP reale contiene correzioni rapide che il filtraggio del CoM a circa 1 Hz attenua. Lo stesso andamento si osserva nell'errore con segno in funzione della frequenza di taglio (Figura 2.19).

Esercizio	Sistema	n	Range$_{ML}$ (%)	Range$_{AP}$ (%)	RMS$_{ML}$ (%)	RMS$_{AP}$ (%)
Occhi Aperti	YOLO	13	-4.0 [11.4]	3.9 [37.9]	3.5 [24.8]	-0.8 [18.8]
	SAM	13	29.9 [18.6]	1.6 [17.4]	41.7 [45.9]	-8.1 [17.0]
Occhi Chiusi	YOLO	10	12.0 [52.9]	2.2 [21.3]	12.4 [55.3]	2.6 [9.1]
	SAM	10	28.8 [13.4]	-9.0 [35.0]	51.0 [43.1]	-13.0 [24.4]
Limite	YOLO	12	-0.6 [5.0]	10.8 [32.5]	1.8 [5.8]	19.8 [30.8]
	SAM	12	12.3 [5.7]	90.2 [65.9]	11.6 [6.5]	223.9 [175.0]
One Foot	YOLO	11	-19.5 [18.6]	23.1 [44.2]	-8.5 [13.3]	16.1 [46.6]
	SAM	11	-4.3 [29.2]	4.2 [36.2]	-0.2 [22.7]	2.3 [11.2]
Tandem	YOLO	8	-20.3 [9.7]	-10.2 [54.2]	-11.4 [7.3]	4.2 [12.1]
	SAM	8	-7.6 [7.9]	16.2 [62.8]	1.4 [6.7]	15.1 [30.9]

Tabella 3.4: Errore percentuale mediano [scarto interquartile] di Range e RMS (ML e AP) rispetto alla pedana, per esercizio e per sistema.

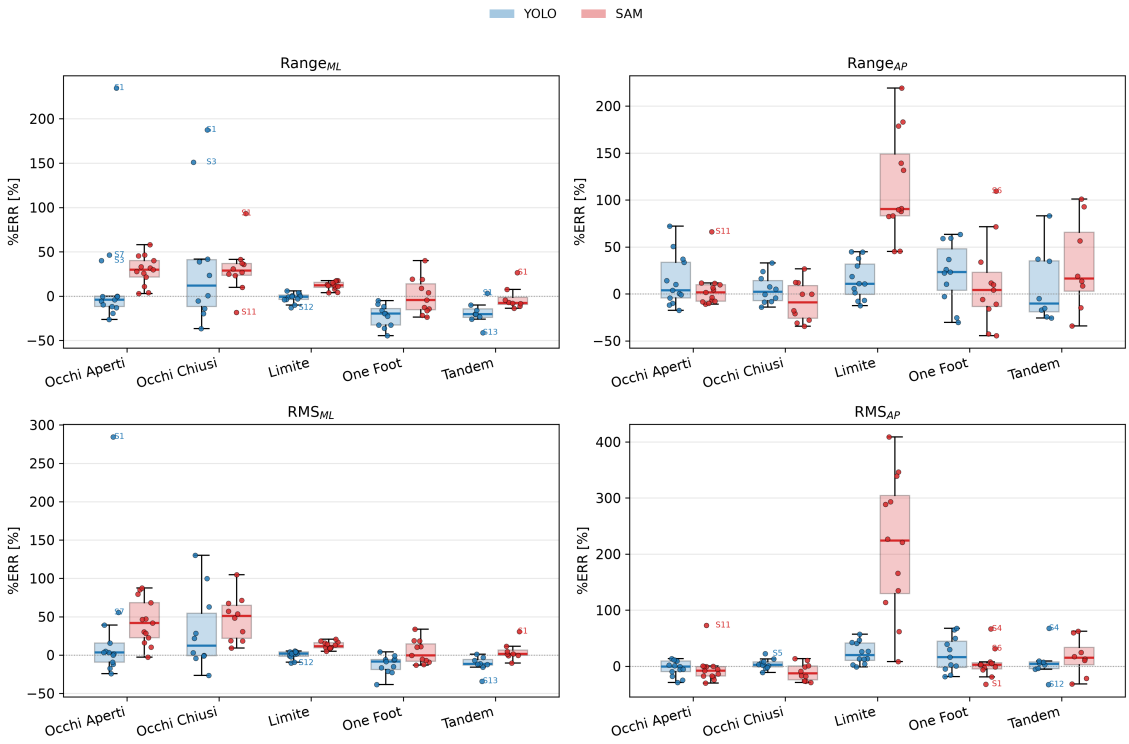


Figura 3.2: Distribuzione completa dell'errore percentuale di Range e RMS, ML e AP (stessi dati della Tabella 3.4), stessa configurazione della Figura 3.1.

Su Range e RMS, che descrivono l'ampiezza dell'oscillazione, YOLO ha un errore mediano entro circa ±23% in tutti gli esercizi e su entrambi gli assi. Il risultato è atteso, perché questi indici dipendono soprattutto dalle basse frequenze, cioè dalla parte del segnale che il sistema video ricostruisce meglio. SAM sovrastima invece l'ampiezza mediolaterale nelle prove statiche (RMS$_{ML}$ +41.7% a occhi aperti e +51.0% a occhi chiusi; Range$_{ML}$ +29.9% e +28.8%). In termini assoluti la differenza è però di circa 1 mm (RMS$_{ML}$ di 3.0 contro 1.9 mm e di 1.8 contro 1.1 mm, Tabella 3.2): l'errore percentuale è elevato perché l'oscillazione mediolaterale in stazione eretta quieta è la più piccola misurata, e su un segnale di pochi millimetri il rumore di una singola telecamera pesa di più. Infine, YOLO sottostima il Range$_{ML}$ nelle prove a base ridotta (−19.5% in *One Foot* e −20.3% in *Tandem*), per lo stesso effetto del filtraggio osservato sulla Path Length: le correzioni rapide tipiche di queste prove vengono attenuate e l'escursione massima risulta ridotta.

Un'eccezione netta è SAM sull'asse AP nel *Limite* (Tabella 3.4, riga *Limite*/SAM): RMS$_{AP}$ ha un errore mediano di +224% [scarto interquartile 175%] e Range$_{AP}$ di +90%

[66%], molto più alti di qualunque altra cella della tabella. Questo caso è approfondito nella Sezione 3.4.

3.3 Correlazioni ed errore di forma della traiettoria

La Tabella 3.5 riporta, per esercizio e sistema, i coefficienti di correlazione di Pearson e di Spearman tra la traiettoria stimata e quella della pedana lungo gli assi ML e AP, insieme all'RMSE istante per istante (in mm) tra le due traiettorie, dopo aver centrato entrambe sulla propria media per non essere influenzati da un'eventuale differenza di origine tra i due sistemi di riferimento. La correlazione misura quanto l'andamento nel tempo della traiettoria stimata dai sistemi assomiglia a quello della pedana. È una misura indipendente dall'ampiezza, quindi da sola non basta a stabilire l'accordo con la pedana, che va valutato assieme agli errori sugli indici (Tabelle 3.3 e 3.4) e ai limiti di accordo di Bland-Altman (Sezione 3.5).

Esercizio	Sistema	n	Pearson$_{ML}$	Pearson$_{AP}$	Spearman$_{ML}$	Spearman$_{AP}$	RMSE$_{ML}$ [mm]	RMSE$_{AP}$ [mm]
Occhi Aperti	YOLO	13	0.88 [0.22] (2/13)	0.81 [0.21]	0.86 [0.23] (2/13)	0.79 [0.24] (1/13)	1.2 [0.2]	3.2 [1.1]
	SAM	13	0.84 [0.09]	0.82 [0.13]	0.82 [0.10]	0.83 [0.14]	1.4 [1.3]	3.3 [1.7]
Occhi Chiusi	YOLO	10	0.80 [0.35] (2/10)	0.87 [0.20]	0.76 [0.34] (3/10)	0.88 [0.21]	1.2 [0.4]	2.4 [1.4]
	SAM	10	0.78 [0.09]	0.83 [0.14]	0.78 [0.15]	0.78 [0.15]	1.0 [0.4]	2.6 [1.2]
Limite	YOLO	12	1.00 [0.00]	0.79 [0.27] (1/12)	0.95 [0.03]	0.79 [0.28] (1/12)	8.3 [4.1]	7.1 [3.1]
	SAM	12	1.00 [0.00]	0.16 [0.90] (9/12)	0.96 [0.01]	0.18 [1.04] (8/12)	18.7 [7.7]	28.1 [8.0]
One Foot	YOLO	11	0.79 [0.14]	0.76 [0.20] (2/11)	0.77 [0.12] (1/11)	0.73 [0.20] (3/11)	4.2 [1.3]	8.4 [4.3]
	SAM	11	0.84 [0.06]	0.83 [0.15]	0.81 [0.07]	0.82 [0.15]	3.4 [1.2]	5.2 [3.2]
Tandem	YOLO	8	0.89 [0.05]	0.55 [0.17] (3/8)	0.88 [0.05]	0.52 [0.20] (3/8)	2.9 [1.3]	5.3 [1.6]
	SAM	8	0.93 [0.09]	0.31 [0.45] (5/8)	0.91 [0.09]	0.33 [0.42] (5/8)	2.9 [1.2]	6.3 [1.3]

Tabella 3.5: Pearson e Spearman (ML e AP) ed RMSE di traiettoria (mm, ML e AP), mediana [scarto interquartile], per esercizio e per sistema. Tra parentesi, dove presente, il numero di sessioni con correlazione inferiore a 0.5 sul totale delle sessioni di quell'esercizio.

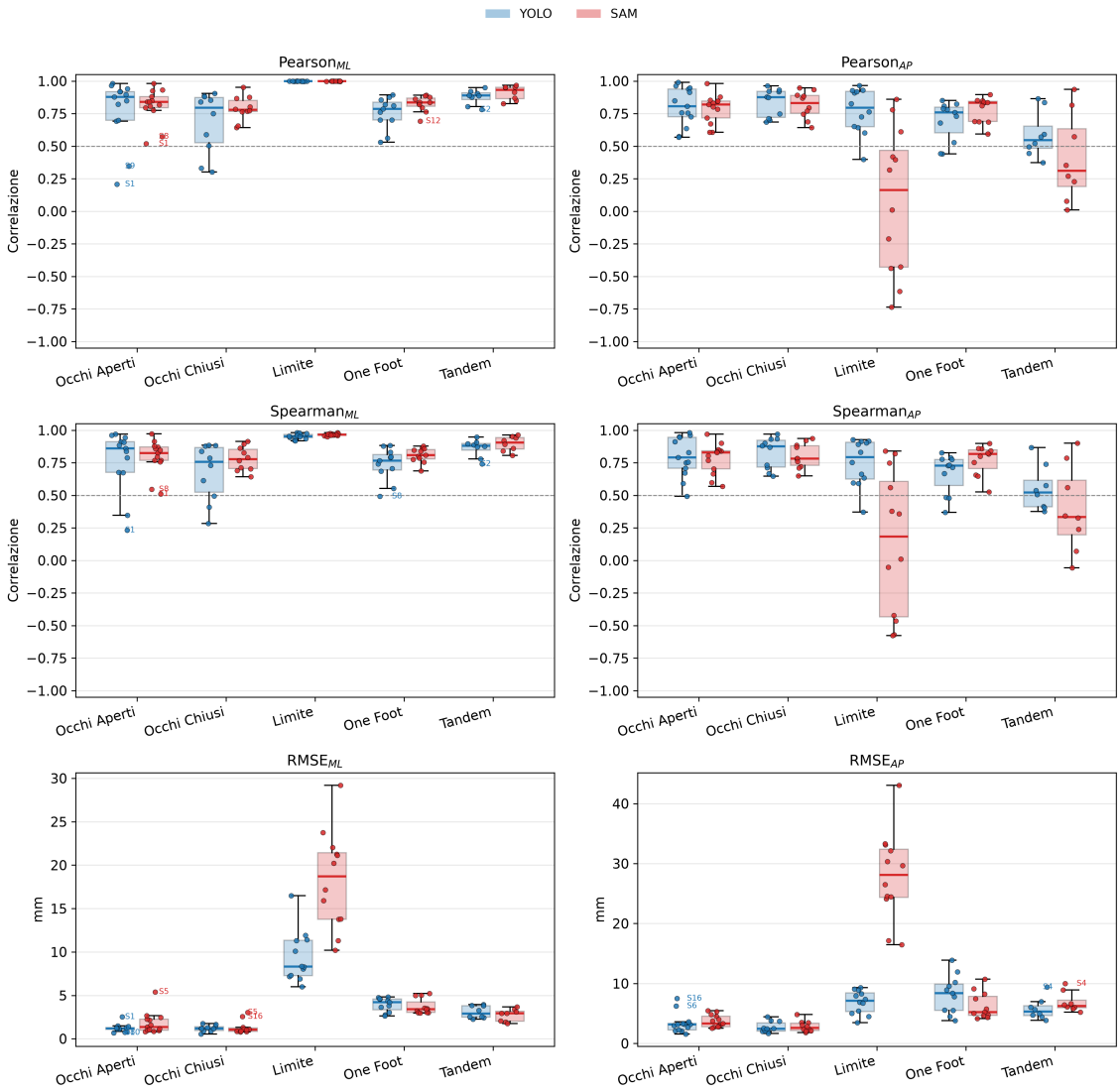


Figura 3.3: Distribuzione completa di Pearson, Spearman (ML e AP) e RMSE di traiettoria (corrispondente alla Tabella 3.5), stessa configurazione della Figura 3.1. La linea tratteggiata indica la soglia di correlazione 0.5.

Le correlazioni sono per lo più comprese tra 0.7 e 0.9, valori che indicano una correlazione alta con la pedana [72]. Come anticipato, una correlazione alta non implica però un accordo in ampiezza: ad esempio, SAM a occhi aperti e a occhi chiusi ha correlazioni ML mediane di 0.84 e 0.78, ma sovrastima l'RMS$_{ML}$ del 42% e del 51% (Tabella 3.4), una differenza che in termini assoluti è di circa 1 mm, come discusso nella sezione precedente. Fa eccezione l'asse ML nel *Limite*, dove la correlazione di Pearson è quasi perfetta per entrambi i sistemi (mediana 0.999, scarto interquartile 0.001): in questo esercizio

il soggetto compie un'unica ampia oscillazione volontaria lungo ML, per cui il segnale utile è molto più grande di qualunque rumore di tracciamento residuo. In questo caso alla correlazione elevata si accompagnano anche errori contenuti su Range_{ML} e RMS_{ML} (entro il 13%, Tabella 3.4), per cui l'accordo è buono anche in ampiezza. Anche la correlazione di Spearman ML resta alta nello stesso esercizio, ma un poco più bassa e variabile (mediana 0.95 per YOLO e 0.96 per SAM, scarto interquartile fino a 0.03), una differenza compatibile con il fatto che il segnale del *Limite* è formato da due plateau (soggetto fermo agli estremi), nei quali i valori differiscono di pochi millimetri e il loro ordine è dominato dal rumore: ciò abbassa la correlazione sui ranghi, mentre quella di Pearson è dominata dalla grande escursione tra i due plateau. L'eccezione opposta è di nuovo l'asse AP di SAM nel *Limite*, discussa di seguito. Anche nel *Tandem* l'asse AP ha correlazioni più basse per entrambi i sistemi (Pearson mediano 0.55 per YOLO e 0.31 per SAM). Il risultato è atteso: con i piedi in tandem il controllo antero-posteriore avviene soprattutto trasferendo il carico tra il piede anteriore e quello posteriore [54], per cui il CoP si sposta lungo AP senza un corrispondente movimento del CoM, che il modello del pendolo inverso a singolo segmento non può riprodurre (Sezione 2.6.2).

Va precisato che mediana e scarto interquartile descrivono solo la fascia centrale delle sessioni: alcune singole sessioni possono avere una correlazione bassa senza che questo si veda dal solo intervallo interquartile. Per questo, in Tabella 3.5, le celle di Pearson e Spearman in cui almeno una sessione è sotto 0.5 riportano tra parentesi quante sessioni, su quante totali; la Figura 3.3 mostra inoltre la distribuzione completa, sessione per sessione. Nella maggior parte dei casi si tratta di 1–3 sessioni isolate. Fanno eccezione l'asse AP nel *Tandem*, sotto soglia in 3 sessioni su 8 per YOLO e in 5 su 8 per SAM per il motivo discusso sopra, e l'asse AP di SAM nel *Limite* (9 sessioni su 12), approfondito nella Sezione 3.4.

Le Figure 3.1, 3.2 e 3.3 mostrano inoltre che gran parte delle sessioni etichettate come outlier non sono distribuite a caso, ma si concentrano su pochi soggetti che ricompaiono su più indici. Il caso più evidente è il soggetto S1: nelle prove statiche (*Occhi Aperti*, *Occhi Chiusi*) compare tra i valori più estremi per YOLO in Sway Area (+187% e +116%), Range_{ML} (+235% e +188%), RMS_{ML} (+285% e +130%) e nella correlazione ML (Pearson

0.208 e 0.332). La causa emerge dal contenuto in frequenza del segnale: in queste due sessioni, la frazione di potenza del CoM di YOLO sotto 1 Hz (media dei due assi) è 0.74 e 0.64, contro una mediana di 0.92 [0.17] su 54 sessioni; il segnale video di quel soggetto contiene quindi più energia sopra 1 Hz del tipico, compatibile con un rumore di tracciamento maggiore che il taglio fisso a 1.0 Hz attenua solo in parte. Lo stesso può essere riscontrato in S3 a *Occhi Chiusi*, che ha il valore più basso di tutte le sessioni (0.51) ed errori anomali su Range_{ML} e Path Length. In questi casi, gli errori più elevati derivano quindi da singole acquisizioni con tracciamento più rumoroso, e non da un limite sistematico del metodo.

3.4 Stima della profondità: SAM sull'asse AP nel *Limite*

Nel *Limite*, tre indicatori mostrano che la stima di SAM lungo l'asse AP si discosta dai valori della pedana più di qualunque altra condizione esaminata:

- l'errore percentuale di RMS_{AP} e Range_{AP} è il più alto tra tutti gli esercizi (+224% e +90%, Tabella 3.4);
- la correlazione di Pearson sull'asse AP ha una mediana di 0.16, con 9 sessioni su 12 sotto 0.5 (Tabella 3.5);
- l'RMSE di traiettoria sull'asse AP è il più alto registrato, 28.1 mm [8.0] (Tabella 3.5), a fronte di un range AP della pedana di 49 mm (Tabella 3.2).

Il risultato è atteso e dipende dalla geometria dell'acquisizione. Nel *Limite* il movimento volontario avviene quasi interamente lungo l'asse ML (Range ML della pedana 315 mm contro 49 mm di Range AP, Tabella 3.2), per cui lungo l'asse AP il segnale posturale reale è piccolo. La telecamera usata da SAM è inoltre posta frontalmente al soggetto (Sezione 2.4.3), quindi l'asse AP coincide con la direzione della profondità, cioè con la grandezza che una stima monoculare non misura direttamente ma stima tramite la forma del corpo [35]. Nelle prove statiche il soggetto è quasi fermo e l'errore sulla profondità risulta pressoché costante, per cui non altera Range e RMS. Durante

l'ampia inclinazione laterale del *Limite* la postura cambia sensibilmente e l'errore sulla profondità varia nel tempo, aggiungendo lungo AP un'oscillazione che la pedana non registra. YOLO, che utilizza la triangolazione multi-camera per ricostruire la posizione, non presenta questo effetto: pur avendo anch'esso una correlazione AP più bassa nel *Limite* (mediana 0.79) che lungo ML (mediana 0.999), mantiene un errore di RMS_{AP} del +19.8% e una correlazione di Pearson AP minima di 0.40. Il caso conferma quindi il vantaggio atteso del sistema multi-camera rispetto alla direzione della profondità della camera usata per la stima monoculare.

3.5 Analisi di accordo (Bland-Altman)

Per ciascun indice, le Figure 3.4–3.9 riportano, sessione per sessione, la media e la differenza tra sistema e pedana su tutti gli esercizi insieme, distinti per colore e simbolo. In ogni grafico la linea orizzontale continua indica il *bias*, cioè la media delle differenze tra i due metodi, la linea punteggiata il riferimento allo zero e le due linee tratteggiate i *limiti di accordo* al 95% (bias ± 1.96 volte la deviazione standard delle differenze). I limiti di accordo non vanno confusi con un intervallo di confidenza sulla media: non descrivono quanto è precisa la stima del bias, ma l'intervallo entro cui ci si aspetta che cada circa il 95% delle differenze *individuali* tra i due metodi.

La Tabella 3.6 riassume, per ciascun indice, il bias e i limiti di accordo, calcolati sull'insieme di tutti gli esercizi. Questi valori rispondono a una domanda diversa da quella delle Tabelle 3.3 e 3.4, che riportano l'errore percentuale separatamente per ciascun esercizio: indicano di quanto, in unità assolute, la stima di una singola sessione può discostarsi dalla pedana, e se il sistema tende a sovrastimare o a sottostimare.

Indice	Sistema	Bias	LoA	Unità
Sway Area	YOLO	408	[-2524, 3340]	mm^2
	SAM	2510	[-9338, 14359]	mm^2
Path Length	YOLO	-194	[-925, 537]	mm
	SAM	-62	[-679, 554]	mm
Mean Velocity	YOLO	-6.7	[-31.9, 18.5]	mm/s
	SAM	-2.1	[-23.4, 19.1]	mm/s
$Range_{ML}$	YOLO	-3.5	[-21.0, 14.0]	mm
	SAM	9.1	[-24.7, 43.0]	mm
$Range_{AP}$	YOLO	5.1	[-15.7, 25.8]	mm
	SAM	13.1	[-37.9, 64.2]	mm
RMS_{ML}	YOLO	-0.1	[-6.1, 5.9]	mm
	SAM	4.1	[-10.1, 18.4]	mm
RMS_{AP}	YOLO	1.0	[-3.2, 5.2]	mm
	SAM	3.9	[-13.2, 21.0]	mm

Tabella 3.6: Bias e limiti di accordo al 95% (bias ± 1.96 SD) tra ciascun sistema e la pedana, per indice, calcolati su tutte le sessioni e tutti gli esercizi insieme.

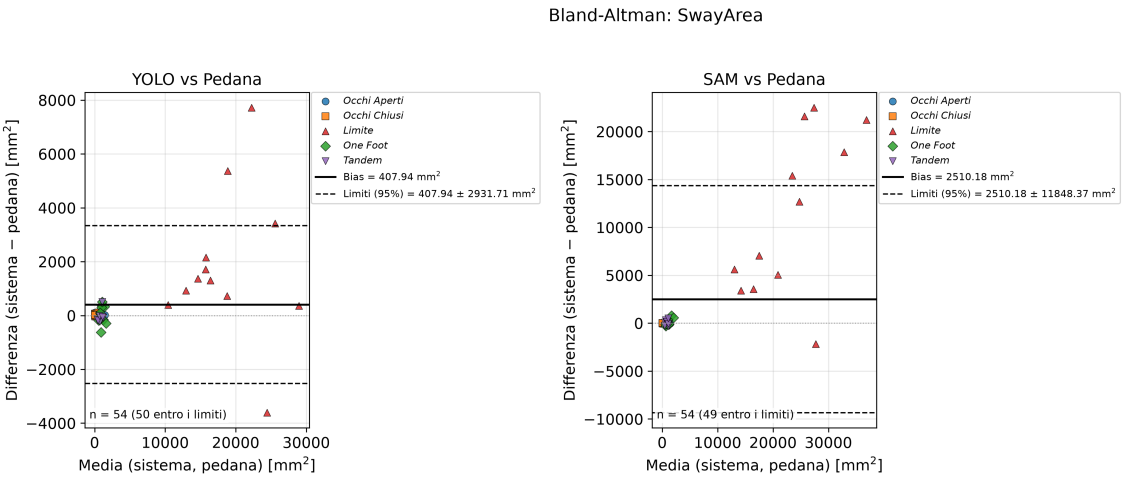


Figura 3.4: Bland-Altman per Sway Area: media e differenza (sistema − pedana) tra YOLO e pedana (sinistra) e tra SAM e pedana (destra), colorati per esercizio. La linea continua indica il bias, le linee tratteggiate i limiti di accordo al 95%.

Per Sway Area il bias è di +408 mm^2 per YOLO e di +2510 mm^2 per SAM. In entrambi i pannelli i punti delle prove statiche, di *One Foot* e di *Tandem* restano concentrati vicino allo zero, mentre quelli del *Limite* si allontanano al crescere della media. Il bias più alto di SAM deriva quasi interamente da questo esercizio, in cui la Sway Area è sovrastimata in 11 sessioni su 12 (errore percentuale mediano +61.9%, Tabella 3.3): poiché l'area dell'ellisse dipende dall'ampiezza dell'oscillazione lungo entrambi gli assi, vi si riflette l'eccesso di ampiezza lungo AP descritto nella Sezione 3.4.

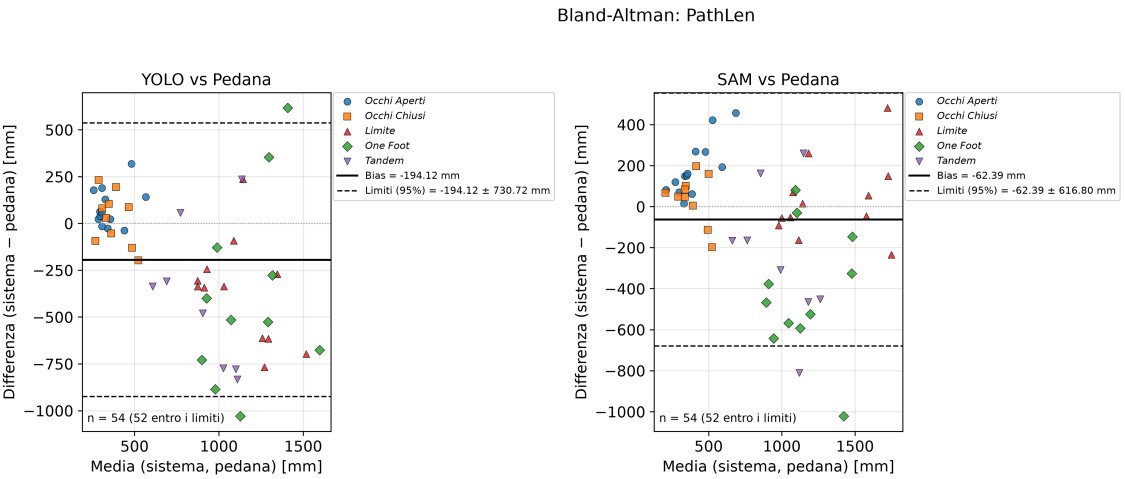


Figura 3.5: Bland-Altman per Path Length, stessa configurazione della Figura 3.4.

Per Path Length il bias è negativo per entrambi i sistemi (−194 mm per YOLO e −62 mm per SAM). I punti seguono il cambio di segno già visto nella Tabella 3.3: per YOLO stanno sopra lo zero in 16 delle 23 sessioni a occhi aperti e a occhi chiusi e sotto lo zero in 26 delle 31 sessioni di *Limite*, *One Foot* e *Tandem*, e le mediane di SAM hanno lo stesso segno di quelle di YOLO in tutti gli esercizi. Il bias complessivo è negativo perché nelle prove più impegnative la Path Length è circa quattro volte maggiore che nelle prove statiche (Tabella 3.1), per cui in millimetri la sottostima prevale sulla sovrastima. La differenza tra i due bias deriva soprattutto dal *Limite*, in cui la Path Length è sottostimata da YOLO (−30.7%) mentre per SAM l'errore è trascurabile (−0.8%). Mean Velocity non è rappresentata in figura: essendo Path Length diviso per la durata dell'acquisizione, ha lo stesso andamento in un'altra scala (bias e limiti di accordo sono nella Tabella 3.6).

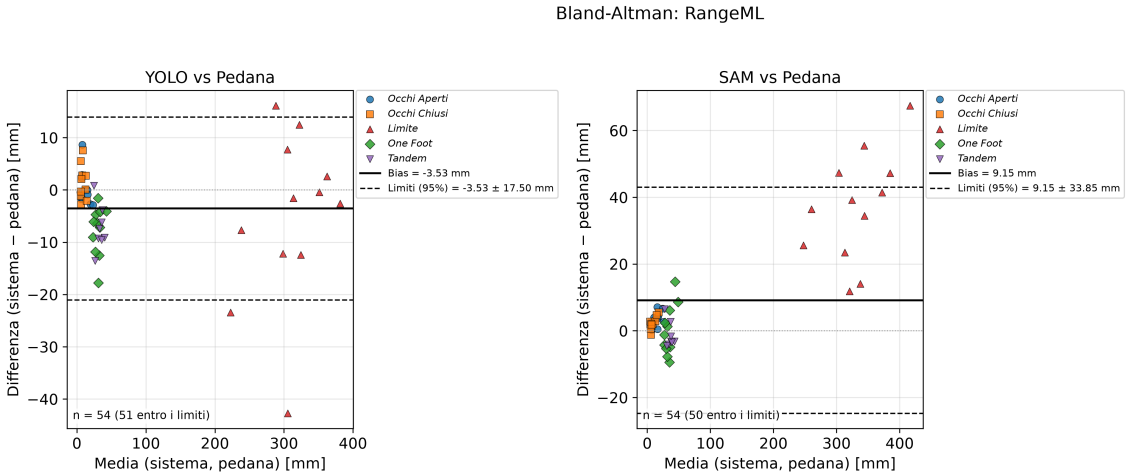


Figura 3.6: Bland-Altman per Range ML, stessa configurazione della Figura 3.4.

Per Range ML il bias è di −3.5 mm per YOLO e di +9.1 mm per SAM. Il bias di SAM deriva dal *Limite*, in cui il Range ML è sovrastimato di una quota quasi costante tra le sessioni (+12.3% [5.7], Tabella 3.4): anche il punto più lontano dallo zero, con una differenza di quasi +70 mm su una media di circa 420 mm, corrisponde a una sovrastima di circa il 18%. Un errore proporzionale così regolare è compatibile con l'ambiguità di scala della stima monoculare [35]: SAM ricava la distanza del soggetto dalla telecamera senza alcun vincolo sulla statura (Sezione 2.4.3), e un errore su questa distanza si traduce

in un errore proporzionale sugli spostamenti nel piano dell'immagine, che un vincolo sulla statura del soggetto potrebbe correggere.

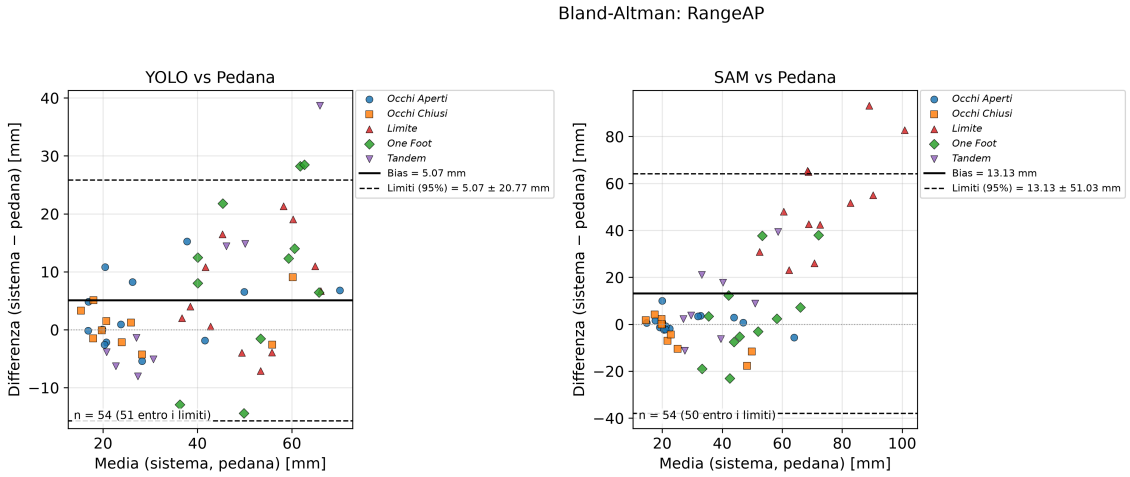


Figura 3.7: Bland-Altman per Range AP, stessa configurazione della Figura 3.4.

Per Range AP il bias è positivo per entrambi i sistemi (+5.1 mm per YOLO e +13.1 mm per SAM). Il segno è atteso: il Range dipende dai due valori estremi del tracciato, e il rumore di tracciamento residuo tende ad ampliarli. Per YOLO l'effetto è di pochi millimetri; per SAM i punti più lontani dallo zero sono quelli del *Limite*, con differenze fino a oltre +90 mm, dovute all'oscillazione aggiunta lungo la direzione della profondità (Sezione 3.4).

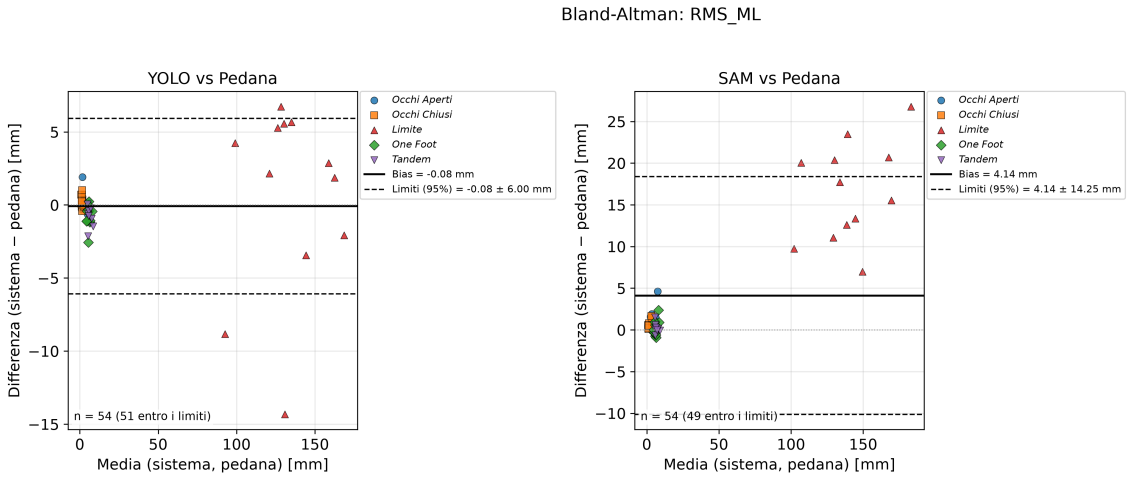


Figura 3.8: Bland-Altman per RMS ML, stessa configurazione della Figura 3.4.

Per RMS ML il bias è praticamente nullo per YOLO (−0.1 mm) e di +4.1 mm per SAM. Nel *Limite* le differenze di YOLO hanno segno variabile (da −14.3 a +6.7 mm), mentre quelle di SAM sono tutte positive (da +7.0 a quasi +27 mm), con un errore percentuale mediano di +11.6% [6.5] (Tabella 3.4): praticamente lo stesso valore del Range ML, come atteso per l'errore di scala descritto sopra.

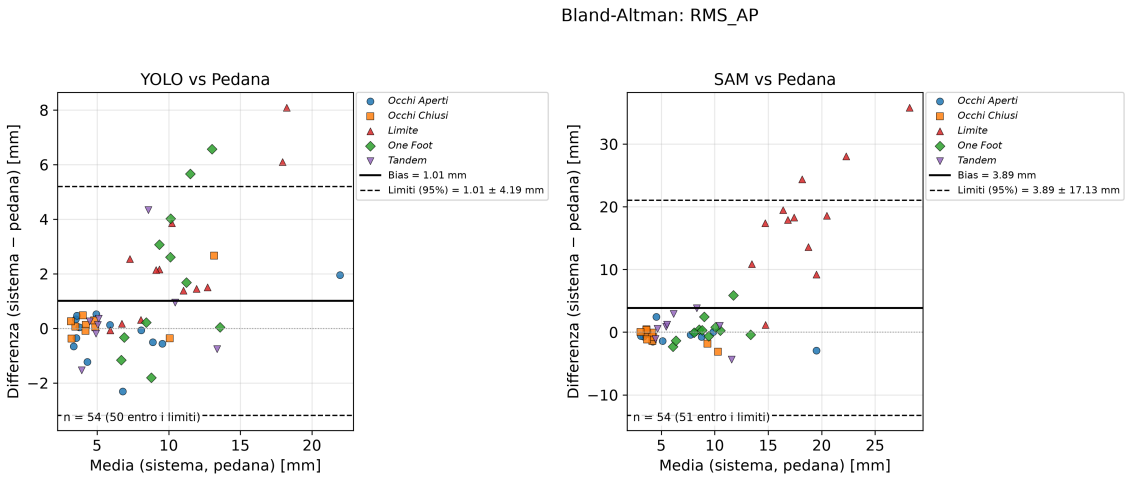


Figura 3.9: Bland-Altman per RMS AP, stessa configurazione della Figura 3.4.

RMS AP è l'indice in cui la differenza tra i due sistemi è più evidente. YOLO resta vicino allo zero anche nel *Limite* (bias +2.5 mm in questo esercizio e +1.0 mm su tutte le sessioni, differenza massima +8.1 mm), mentre per SAM lo scarto cresce con la media, in modo compatibile con l'effetto della profondità (Sezione 3.4). Tre sessioni su 54 superano il limite di accordo superiore (21.0 mm), con differenze da +24 a +36 mm: sono tutte del *Limite*, esercizio in cui tutte le 12 differenze di SAM sono positive, per cui il superamento non è casuale ma riflette l'errore sistematico descritto sopra.

Per Sway Area, Range_{ML} e RMS_{ML} la dispersione delle differenze cresce con la media, più per SAM che per YOLO: i punti del *Limite* si allontanano dalla linea del bias molto più di quelli delle altre prove, che restano concentrati vicino allo zero. L'andamento è atteso, perché nel *Limite* questi indici hanno valori assoluti di uno o due ordini di grandezza superiori a quelli delle prove di stazione eretta quieta, mentre gli errori percentuali variano molto meno (Tabelle 3.3 e 3.4): un errore relativo simile produce quindi uno scarto assoluto molto maggiore. Per la stessa ragione i limiti di accordo della Tabel-

la 3.6, calcolati su tutti gli esercizi insieme, sono determinati soprattutto dalle sessioni del *Limite* e sono molto più ampi della differenza attesa in una prova di stazione eretta quieta. Per Path Length la dispersione è invece maggiore in *One Foot* e *Tandem*, dove la sottostima dovuta al filtraggio è più marcata (−35.0% e −42.7% per YOLO, −36.0% e −24.6% per SAM, Tabella 3.3).

3.6 Sintesi del confronto tra i due sistemi

La Tabella 3.7 riassume, in forma di conteggio, il confronto tra YOLO e SAM sulle mediane per esercizio delle Tabelle 3.3, 3.4 e 3.5: ogni combinazione è una coppia esercizio-indice, distinta per asse dove previsto. YOLO ha l'errore percentuale mediano minore in valore assoluto in 19 combinazioni su 30 e l'RMSE di traiettoria minore in 7 su 10, mentre le correlazioni mediane sono divise in maniera simmetrica tra i due sistemi in analisi. Il confronto è descittivo, senza test di significatività, su 8–13 sessioni per esercizio.

Il risultato è coerente con quanto osservato nelle sezioni precedenti. Errori percentuali e RMSE dipendono dall'ampiezza del tracciato e, per SAM, risentono sia dell'errore di scala lungo ML sia dell'oscillazione aggiunta lungo la profondità, effetti a cui la triangolazione di YOLO non è soggetta. La correlazione, invece, non risente di un errore di scala, e infatti i due sistemi si dividono equamente le combinazioni: YOLO prevale soprattutto sull'asse AP nel *Limite* e nel *Tandem*, cioè lungo la direzione che per la telecamera di SAM coincide con la profondità. Il vantaggio di YOLO riguarda quindi soprattutto la stima dell'ampiezza dell'oscillazione, mentre l'andamento temporale del CoP è ricostruito in modo analogo da entrambi i sistemi.

Criterio	Combinazioni	YOLO	SAM
Errore percentuale mediano minore in valore assoluto	30	19	11
Errore percentuale mediano entro il 30%	30	27	23
Correlazione mediana più alta (Pearson e Spearman)	20	10	10
RMSE di traiettoria mediano minore	10	7	3

Tabella 3.7: Numero di combinazioni in cui ciascun sistema risulta migliore secondo il criterio indicato, sulle mediane per esercizio. Per correlazione e RMSE il confronto usa le mediane non arrotondate a due decimali.

Capitolo 4

Conclusioni

4.1 Sintesi del Lavoro

Questa tesi ha confrontato due sistemi markerless per la stima del centro di pressione con il CoP misurato da una pedana di forza in stabilometria. Sono stati confrontati un sistema multi-camera basato su quattro telecamere sincronizzate, sul modello YOLO26x e sul modello geometrico a solidi, e un sistema single-camera basato su SAM 3D Body. Per entrambi i sistemi il CoP è stato stimato partendo dal centro di massa e applicando il modello del pendolo inverso, dopo un filtraggio passa-basso scelto a 1.0 Hz per YOLO e 1.1 Hz per SAM. L'analisi è stata condotta su 54 sessioni valide di 16 volontari, che hanno eseguito cinque prove da 30 s l'una (*Occhi Aperti* e *Occhi Chiusi* in *quiet standing*, *Tandem*, *One Foot* e *Limite*). I confronti sono stati valutati attraverso gli errori sugli indici posturografici, le correlazioni, l'RMSE di traiettoria e l'analisi di Bland-Altman.

4.2 Risultati rispetto agli Obiettivi

4.2.1 Protocollo

Il protocollo comprendeva esercizi statici (*Occhi Aperti*, *Occhi Chiusi*), a base ridotta (*Tandem*, *One Foot*) e dinamici (*Limite*, il test dei limiti di stabilità), così da valutare i sistemi su oscillazioni di ampiezza molto diversa, da pochi millimetri a oltre 300 mm. Delle 80 sessioni acquisite ne sono state utilizzate 54: le escluse hanno registrazioni incomplete per cause tecniche e non risultano associate in modo significativo al compito, ma i gruppi finali (da 8 a 13 sessioni per esercizio) sono sbilanciati e composti da soggetti in parte diversi.

4.2.2 Stima del CoM e del CoP

Il CoM è stato stimato con entrambi i sistemi e il CoP ricavato con il modello del pendolo inverso. Poiché la doppia derivazione amplifica il rumore, il CoM è stato filtrato con un Butterworth del secondo ordine a fase zero, con un taglio per sistema scelto confrontando tre criteri (1.0 Hz per YOLO e 1.1 Hz per SAM). Ripetendo la scelta escludendo un soggetto alla volta, il taglio resta vicino al valore adottato per entrambi i sistemi: tra 1.0 e 1.1 Hz per SAM e tra 0.9 e 1.2 Hz per YOLO.

4.2.3 Confronto con la Pedana

Il CoP stimato riproduce l'andamento della pedana, con correlazioni per lo più comprese tra 0.7 e 0.9, cioè alte. Gli indici seguono l'andamento atteso tra gli esercizi e hanno lo stesso ordine di grandezza di quelli della pedana. Il segno dell'errore dipende più dall'esercizio che dal sistema: la Path Length, ad esempio, è sovrastimata da entrambi i sistemi nelle prove statiche e sottostimata in *One Foot* e *Tandem*.

4.2.4 Accordo tra i Sistemi

Per quantificare l'accordo tra ciascun sistema e la pedana sono stati usati tre punti di vista complementari: l'errore sugli indici, la somiglianza delle traiettorie nel tempo (correlazioni e RMSE di traiettoria) e l'analisi di Bland-Altman, che riporta le differenze in unità assolute. Per Sway Area, Range e RMS mediolaterali le differenze in unità assolute crescono con l'ampiezza dell'oscillazione, per cui i limiti di accordo, calcolati su tutti gli esercizi insieme, sono determinati soprattutto dal *Limite*. YOLO ha l'errore percentuale mediano minore in 19 combinazioni indice-esercizio su 30 e l'RMSE di traiettoria minore in 7 su 10, mentre le correlazioni mediane sono più alte per ciascun sistema in 10 combinazioni su 20. Il confronto è descrittivo, senza test di significatività, e riguarda due sistemi completi.

4.3 Interpretazione dei Risultati

Dai risultati emerge che l'accordo con la pedana dipende soprattutto dall'ampiezza del segnale rispetto al rumore di tracciamento, il problema di fondo richiamato nel Capitolo 1. Dove l'oscillazione è ampia l'accordo è molto buono: sull'asse mediolaterale nel *Limite* la correlazione è quasi perfetta (0.999) e gli errori su Range e RMS mediolaterali restano sotto il 13%. Nella stazione eretta quieta, invece, l'RMSE di traiettoria (1–3 mm) è dello stesso ordine di grandezza dell'oscillazione misurata dalla pedana (RMS di 1–5 mm), per cui l'errore percentuale può raggiungere decine di punti anche quando la differenza assoluta è minima: la sovrastima di SAM sull'RMS mediolaterale a occhi aperti e a occhi chiusi (+42% e +51%) corrisponde a circa 1 mm.

Le differenze tra i due sistemi hanno cause riconoscibili e in buona parte attese. Per YOLO l'effetto principale è quello del filtraggio: nelle prove più impegnative il taglio a circa 1 Hz attenua le correzioni rapide del CoP, per cui la Path Length risulta sottostimata (−31% nel *Limite*), mentre nelle prove statiche il rumore residuo la allunga. Per SAM gli errori principali derivano dalla stima monoculare. Lungo l'asse mediolaterale, l'ambiguità di scala produce nel *Limite* una sovrastima proporzionale e regolare tra le sessioni (circa +12% su Range e RMS). Lungo l'asse antero-posteriore, che per la telecamera frontale coincide con la direzione della profondità, l'ampia inclinazione laterale del *Limite* introduce un'oscillazione che la pedana non registra (RMS_{AP} +224%, correlazione di Pearson mediana 0.16), mentre la triangolazione di YOLO non ne risente. Nel *Tandem*, infine, l'asse AP ha correlazioni più basse per entrambi i sistemi (0.55 per YOLO e 0.31 per SAM), come atteso per il controllo antero-posteriore basato sul trasferimento di carico tra i due piedi, che il pendolo inverso a singolo segmento non descrive.

Nel complesso, YOLO stima meglio l'ampiezza dell'oscillazione, mentre l'andamento temporale del CoP è ricostruito in modo analogo dai due sistemi: la correlazione non risente dell'errore di scala, e YOLO prevale soprattutto lungo la direzione della profondità di SAM. La stima del CoP dal CoM con sistemi markerless risulta quindi fattibile e riproduce l'andamento della pedana. Poiché la correlazione misura la somiglianza dell'andamento e non l'ampiezza, questo giudizio va letto insieme agli errori sugli indici e

ai limiti di accordo.

4.4 Limiti dello Studio

- Campione Limitato e Sbilanciato: delle 80 sessioni acquisite ne sono state utilizzate 54, con gruppi da 8 a 13 sessioni per esercizio e soggetti in parte diversi da un esercizio all'altro, per cui i confronti tra esercizi non sono appaiati e solo due soggetti hanno tutte e cinque le prove valide. Le esclusioni non risultano associate in modo significativo al compito, ma con 16 sessioni per esercizio un legame non si può escludere.
- Durata e Struttura delle Prove: le prove durano 30 s, mentre per le prove statiche in appoggio bipodalico la letteratura indica un minimo di 90 s per un'affidabilità accettabile [15], ed è quindi possibile che gli indici siano meno ripetibili di quanto sarebbero con acquisizioni più lunghe. Nel *Limite* il movimento è solo mediolaterale, per cui non è stato possibile valutare l'accordo anche per un'ampia escursione antero-posteriore.
- Modello del Pendolo Inverso: il modello presuppone una strategia di caviglia, in cui il corpo oscilla come un unico segmento rigido. Nelle prove a base d'appoggio ridotta intervengono anche anca e tronco e, nel *Tandem*, il trasferimento di carico tra i due piedi, che il modello non può descrivere: è la causa più probabile delle correlazioni più basse sull'asse AP in questo esercizio.
- Scelta del Taglio e Confronto tra i Sistemi: il taglio del filtro è stato scelto sugli stessi dati su cui si valuta l'accordo. La verifica *leave-one-subject-out* mostra che questa scelta favorisce il risultato di circa 6 punti percentuali per YOLO e di meno di 1 per SAM, ma non sostituisce una validazione su dati indipendenti. Il confronto tra YOLO e SAM è descrittivo, senza test di significatività. Le differenze osservate non possono essere attribuite a un unico fattore, dato che riguardano due sistemi completi che differiscono per numero di telecamere, modello corporeo e modello di posa.

- Ipotesi da Verificare: le cause proposte per gli errori di SAM, cioè l'ambiguità di scala lungo l'asse mediolaterale e l'errore sulla profondità lungo l'asse antero-posteriore, sono coerenti con i dati ma non sono state verificate con prove dedicate. Per alcune sessioni con errori elevati, inoltre, non è stata identificata una causa con i dati disponibili.

4.5 Sviluppi Futuri

Prima di proporre un impiego pratico andrebbero affrontati i punti seguenti.

- Campione e Protocollo: uno studio futuro dovrebbe reclutare un campione più ampio e bilanciato tra gli esercizi, e adottare prove più lunghe (almeno 90 s per le prove statiche), con ripetizioni per valutare anche la ripetibilità degli indici. Nel *Limite* sarebbe utile includere oscillazioni antero-posteriori, per verificare l'accordo anche sull'asse AP quando il movimento è ampio.

- Modelli e Analisi: per SAM andrebbero introdotti vincoli di scala e di profondità, ad esempio scalando il modello con la statura e le misure antropometriche del soggetto, e verificata l'ipotesi sulla profondità, ad esempio ripetendo le prove con la telecamera in posizione laterale. Per le prove a base ridotta, un modello a due segmenti (caviglia e anca) potrebbe descrivere meglio il movimento rispetto al pendolo inverso a singolo segmento. La scelta del taglio andrebbe validata su dati indipendenti, e il confronto tra i due sistemi completato con test statistici appaiati su un campione più ampio.

- Il CoM come Indice Stabilometrico: infine, il CoM stimato dal video potrebbe essere usato direttamente come indice di stabilità, senza passare dal CoP e quindi dalla doppia derivazione che amplifica il rumore, da valutare con gli stessi indici e con lo stesso confronto con la pedana.

- Verifica su Soggetti con Controllo Posturale Alterato: il miglior accordo osservato per le oscillazioni ampie suggerisce un possibile utilizzo con soggetti che in genere

presentano oscillazioni più ampie, ma il vantaggio va verificato su popolazioni cliniche.

4.6 Considerazioni Finali

La stima del CoP dal CoM ricavato da video, con quattro telecamere e con una sola, è fattibile e riproduce l'andamento della pedana, soprattutto per oscillazioni ampie e sull'asse mediolaterale. Il sistema multi-camera stima con maggiore accuratezza l'ampiezza dell'oscillazione, mentre il sistema single-camera, più semplice ed economico, ricostruisce l'andamento in modo analogo, ma richiede vincoli di scala e di profondità per stimarne correttamente l'ampiezza. Nella stazione eretta quieta, dove il segnale è di pochi millimetri, gli errori percentuali restano tali da non consentire, allo stato attuale, di sostituire la pedana nella quantificazione precisa degli indici. Il lavoro è quindi una dimostrazione di fattibilità, e le verifiche indicate negli sviluppi futuri definiscono i passi necessari prima di un impiego pratico.

Bibliografia

[1] L Wade et al. «Applications and limitations of current markerless motion capture methods for clinical gait biomechanics». In: *PeerJ* 10 (2022), e12995.

[2] A. Kamieniarz et al. «A posturographic procedure assessing balance disorders in Parkinson's disease: a systematic review». In: *Clinical Interventions in Aging* 13 (2018), pp. 2301–2316. DOI: 10.2147/CIA.S180894.

[3] E. Yu et al. «Evaluation of postural control in quiet standing using center of mass acceleration: comparison among the young, the elderly, and people with stroke». In: *Archives of Physical Medicine and Rehabilitation* 89.6 (2008), pp. 1133–1139. DOI: 10.1016/j.apmr.2007.10.047.

[4] Diane Podsiadlo e Sandra Richardson. «The timed "Up & Go": a test of basic functional mobility for frail elderly persons». In: *Journal of the American Geriatrics Society* 39.2 (1991), pp. 142–148. DOI: 10.1111/j.1532-5415.1991.tb01616.x.

[5] Anne Shumway-Cook, Sandy Brauer e Marjorie Woollacott. «Predicting the probability for falls in community-dwelling older adults using the Timed Up & Go Test». In: *Physical Therapy* 80.9 (2000), pp. 896–903. DOI: 10.1093/ptj/80.9.896.

[6] K. O. Berg et al. «Measuring balance in the elderly: validation of an instrument». In: *Canadian Journal of Public Health* 83.Suppl 2 (1992). PMID: 1468055, S7–S11.

[7] P. Ortega-Bastidas et al. «Instrumented Timed Up and Go Test (iTUG)—More Than Assessing Time to Predict Falls: A Systematic Review». In: *Sensors* 23.7 (2023), p. 3426. DOI: 10.3390/s23073426.

[8] S. Downs, J. Marquez e P. Chiarelli. «The Berg Balance Scale has high intra- and inter-rater reliability but absolute reliability varies across the scale: a systematic review». In: *Journal of Physiotherapy* 59.2 (2013), pp. 93–99. DOI: 10.1016/S1836-9553(13)70161-9.

[9] L. Blum e N. Korner-Bitensky. «Usefulness of the Berg Balance Scale in stroke rehabilitation: a systematic review». In: *Physical Therapy* 88.5 (2008), pp. 559–566. DOI: 10.2522/ptj.20070205.

[10] Isabelle Poitras et al. «Validity and Reliability of Wearable Sensors for Joint Angle Estimation: A Systematic Review». In: *Sensors* 19.7 (2019), p. 1555. DOI: 10.3390/s19071555.

[11] A Karatsidis et al. «Estimation of ground reaction forces and moments during gait using only inertial motion capture». In: *Sensors* 17.1 (2016), p. 75.

[12] A. Forner Cordero, H. J. Koopman e F. C. van der Helm. «Use of pressure insoles to calculate the complete ground reaction forces». In: *Journal of Biomechanics* 37.9 (2004), pp. 1427–1432. DOI: 10.1016/j.jbiomech.2003.12.016.

[13] Angela Julienne, Evi Verbecque e Stéphane Besnard. «Normative Data for Instrumented Posturography: A Systematic Review and Meta-Analysis». In: *Frontiers in Human Neuroscience* 18 (2024), p. 1498107. ISSN: 1662-5161. DOI: 10.3389/fnhum.2024.1498107.

[14] Flavien Quijoux et al. «A Review of Center of Pressure (COP) Variables to Quantify Standing Balance in Elderly People: Algorithms and Open-Access Code». In: *Physiological Reports* 9.22 (2021), e15067.

[15] Alexander Ruhe, René Fejer e Bruce Walker. «The Test-Retest Reliability of Centre of Pressure Measures in Bipedal Static Task Conditions–a Systematic Review of the Literature». In: *Gait & Posture* 32.4 (2010), pp. 436–445.

[16] B. R. Hindle, J. W. L. Keogh e A. V. Lorimer. «Inertial-Based Human Motion Capture: A Technical Summary of Current Processing Methodologies for Spatiotemporal and Kinematic Measures». In: *Applied Bionics and Biomechanics* 2021 (2021), p. 6628320. DOI: 10.1155/2021/6628320.

[17] Sutton B. Richmond et al. «Leveling the playing field: Evaluation of a portable instrument for quantifying balance performance». In: *Journal of Biomechanics* 75 (2018), pp. 102–107. DOI: 10.1016/j.jbiomech.2018.05.008.

[18] Ross A. Clark et al. «Validity and reliability of the Nintendo Wii Balance Board for assessment of standing balance». In: *Gait & Posture* 31.3 (2010), pp. 307–310. DOI: 10.1016/j.gaitpost.2009.11.012.

[19] William V. C. Figtree e Americo A. Migliaccio. «An Inexpensive 6D Motion Tracking System for Posturography». In: *Frontiers in Neurology* 9 (2018), p. 507. DOI: 10.3389/fneur.2018.00507.

[20] Anaïs Chaumeil et al. «Agreement between a Markerless and a Marker-Based Motion Capture Systems for Balance Related Quantities». In: *Journal of Biomechanics* 165 (2024), p. 112018.

[21] Vicon Motion Systems Ltd. *Vicon: sistemi di motion capture*. https://www.vicon.com. Ultimo accesso: 5 ottobre 2026.

[22] Qualisys AB. *Qualisys: sistemi di motion capture e tracking 3D di precisione*. https://www.qualisys.com. Ultimo accesso: 5 ottobre 2026.

[23] Advanced Mechanical Technology, Inc. (AMTI). *AMTI: piattaforme di forza e sistemi di misura della forza multiassiale*. https://www.amti.biz. Ultimo accesso: 5 ottobre 2026.

[24] Kistler Group. *Kistler: biomeccanica e scienze della vita (piattaforme di forza)*. https://www.kistler.com/INT/en/biomechanics-and-life-sciences/C00000187. Ultimo accesso: 5 ottobre 2026.

[25] Y. S. Park et al. «Verification of Markerless Gait Analysis: Multi-Camera and Single-Camera Approaches in Comparison to Marker-Based Gait Analysis». In: *Medicina (Kaunas)* 62 (2026), p. 418. DOI: 10.3390/medicina62020418.

[26] Desmond Boey et al. «Unlocking the potential of video-based markerless motion analysis to study world-class sporting performance». In: *Journal of Sports Sciences* 44.10 (2026), pp. 1261–1274. DOI: 10.1080/02640414.2025.2576412.

[27] R Feng et al. «Estimation of Three-Dimensional Ground Reaction Force and Center of Pressure During Walking Using a Machine-Learning-Based Markerless Motion Capture System». In: *IEEE Transactions on Biomedical Engineering* (2025).

[28] M. H. Schwartz, J. P. Trost e R. A. Wervey. «Measurement and management of errors in quantitative gait data». In: *Gait & Posture* 20.2 (2004), pp. 196–203. DOI: 10.1016/j.gaitpost.2003.09.011.

[29] G. E. Gorton, D. A. Hebert e M. E. Gannotti. «Assessment of the kinematic variability among 12 motion analysis laboratories». In: *Gait & Posture* 29.3 (2009), pp. 398–402. DOI: 10.1016/j.gaitpost.2008.10.060.

[30] V. T. Keller et al. «Clothing condition does not affect meaningful clinical interpretation in markerless motion capture». In: *Journal of Biomechanics* 141 (2022), p. 111182. DOI: 10.1016/j.jbiomech.2022.111182.

[31] N. Lambricht et al. «Impact of Running Clothes on Accuracy of Smartphone-Based 2D Joint Kinematic Assessment During Treadmill Running Using OpenPifPaf». In: *Sensors* 25.3 (2025), p. 934. DOI: 10.3390/s25030934.

[32] Alexander Toshev e Christian Szegedy. «Deeppose: Human pose estimation via deep neural networks». In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2014, pp. 1653–1660.

[33] Tsung-Yi Lin et al. «Microsoft COCO: Common Objects in Context». In: *Computer Vision – ECCV 2014*. Lecture Notes in Computer Science. Springer International Publishing, 2014, pp. 740–755. DOI: 10.1007/978-3-319-10602-1_48.

[34] Mykhaylo Andriluka et al. «2D Human Pose Estimation: New Benchmark and State of the Art Analysis». In: *2014 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2014, pp. 3686–3693. DOI: 10.1109/CVPR.2014.471.

[35] Ce Zheng et al. «Deep Learning-Based Human Pose Estimation: A Survey». In: *ACM Computing Surveys* 56.1 (2023), pp. 1–37. DOI: 10.1145/3603618.

[36] Yucheng Chen, Yingli Tian e Mingyi He. «Monocular human pose estimation: A survey of deep learning-based methods». In: *Computer Vision and Image Understanding* 192 (2020), p. 102897. DOI: 10.1016/j.cviu.2019.102897.

[37] Debapriya Maji et al. «YOLO-Pose: Enhancing YOLO for Multi Person Pose Estimation Using Object Keypoint Similarity Loss». In: *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. IEEE, 2022, pp. 2636–2645. DOI: 10.1109/CVPRW56347.2022.00297.

[38] Glenn Jocher et al. «Ultralytics YOLO26: Unified Real-Time End-to-End Vision Models». In: *arXiv* (2026). DOI: 10.48550/arXiv.2606.03748. arXiv: 2606.03748 [cs.CV].

[39] Xitong Yang et al. «SAM 3D Body: Robust Full-Body Human Mesh Recovery». In: *arXiv preprint arXiv:2602.15989* (2026).

[40] Z Cao et al. «OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields». In: *IEEE TPAMI* 43.1 (2018), pp. 172–186.

[41] HS Fang et al. «AlphaPose: Whole-Body Regional Multi-Person Pose Estimation and Tracking in Real-Time». In: *IEEE TPAMI* 45.6 (2022), pp. 7157–7173.

[42] Ke Sun et al. «Deep High-Resolution Representation Learning for Human Pose Estimation». In: *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2019, pp. 5686–5696. DOI: 10.1109/CVPR.2019.00584.

[43] J Redmon et al. «You only look once: Unified, real-time object detection». In: *CVPR* (2016), pp. 779–788.

[44] A Mathis et al. «DeepLabCut: markerless pose estimation of user-defined body parts with deep learning». In: *Nature Neuroscience* 21 (2018), pp. 1281–1289.

[45] V Bazarevsky et al. «BlazePose: On-device Real-time Body Pose tracking». In: *CVPR Workshop* (2020).

[46] Carolina Clemente et al. «Feasibility of 3D Body Tracking from Monocular 2D Video Feeds in Musculoskeletal Telerehabilitation». In: *Sensors* 24.1 (2024), p. 206. DOI: 10.3390/s24010206.

[47] Theia Markerless. *Theia3D: motion capture markerless*. https://www.theiamarkerless.com. Ultimo accesso: 5 ottobre 2026.

[48] D Pagnon et al. «Pose2Sim: An End-to-End Workflow for 3D Markerless Sports Kinematics». In: *Sensors* 22.7 (2022), p. 2712.

[49] David A Winter. *Biomechanics and motor control of human movement*. John Wiley & Sons, 2009.

[50] R Dumas et al. «Adjustments to McConville's and Young's body segment inertial parameters». In: *Journal of Biomechanics* 40.3 (2007), pp. 543–553.

[51] S Winter et al. «A Dual X-Ray Absorptiometry (DXA) Validated Geometric Model for the Calculation of Body Segment Inertial Parameters of Young Females». In: *Journal of Applied Biomechanics* 37.4 (2018), pp. 321–328.

[52] MR Yeadon. «The simulation of aerial movement–II. A mathematical model of the human body». In: *Journal of Biomechanics* 23.1 (1990), pp. 67–74.

[53] M Loper et al. «SMPL: A skinned multi-person linear model». In: *ACM TOG* 34.6 (2015), pp. 1–16.

[54] DA Winter. «Human Balance and Posture Control during Standing and Walking». In: *Gait & Posture* 3.4 (dic. 1995), pp. 193–214. ISSN: 0966-6362. DOI: 10.1016/0966-6362(96)82849-9. (Visitato il giorno 29/06/2026).

[55] Diego Ramon Gonçalves Gonzalez, Luis Aureliano Imbiriba e Frederico Caetano Jandre. «Comparison of Body Sway Measured by a Markerless Low-Cost Motion Sensor and by a Force Plate». In: *Research on Biomedical Engineering* 37.3 (set. 2021), pp. 507–517. ISSN: 2446-4740. DOI: 10.1007/s42600-021-00161-4. (Visitato il giorno 29/06/2026).

[56] A. L. Hof, M. G. J. Gazendam e W. E. Sinke. «The condition for dynamic stability». In: *Journal of Biomechanics* 38.1 (2005), pp. 1–8. DOI: 10.1016/j.jbiomech.2004.03.025.

[57] T. E. Prieto et al. «Measures of postural steadiness: differences between healthy young and elderly adults». In: *IEEE Transactions on Biomedical Engineering* 43.9 (1996), pp. 956–966. DOI: 10.1109/10.532130.

[58] J. A. Raymakers, M. M. Samson e H. J. J. Verhaar. «The assessment of body sway and the choice of the stability parameter(s)». In: *Gait & Posture* 21.1 (2005), pp. 48–58. DOI: 10.1016/j.gaitpost.2003.11.006.

[59] H Corriveau et al. «Intrasession Reliability of the "Center of Pressure Minus Center of Mass" Variable of Postural Control in the Healthy Elderly». In: *Archives of Physical Medicine and Rehabilitation* 81.1 (2000), pp. 45–48.

[60] ATI Industrial Automation. *F/T Sensor: Axia130-M300*. https://www.ati-ia.com/products/ft/ft_models.aspx?id=Axia130-M300. Ultimo accesso: 7 ottobre 2026.

[61] IDS Imaging Development Systems GmbH. *U3-3160CP Rev. 2.2*. https://en.ids-imaging.com/store/u3-3160cp-rev-2-2.html. Ultimo accesso: 7 ottobre 2026.

[62] onsemi. *PYTHON2000: CMOS Image Sensor, 2.3 MP, Global Shutter*. https://www.onsemi.com/products/sensors/image-sensors/python2000. Ultimo accesso: 7 ottobre 2026.

[63] Kowa Lenses. *LM6JC: 2/3" 6mm MP C-Mount Lens*. https://www.kowa-lenses.com/LM6JC-2-3-6mm-MP-C-Mount-Lens/10221. Ultimo accesso: 11 ottobre 2026.

[64] Nicola Giulietti et al. «A Vision-Based marker-less measurement system for Real-Time Estimation of Human Body Inertial properties». In: *Measurement* 260 (2026), p. 119843. DOI: 10.1016/j.measurement.2025.119843.

[65] Nicola Giulietti et al. «A Real-Time Human Pose Measurement System for Human-In-The-Loop Dynamic Simulators». In: *IEEE Access* 13 (2025), pp. 24954–24969. DOI: 10.1109/ACCESS.2025.3538332.

[66] Z. Zhang. «A flexible new technique for camera calibration». In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22.11 (2000), pp. 1330–1334. DOI: 10.1109/34.888718.

[67] Kevin D. Dames e Sutton B. Richmond. «Tall Tales of Balance: The Influence of Height on Postural Control Measures». In: *Motor Control* 29.2 (apr. 2025), pp. 173–182. ISSN: 1087-1640. DOI: 10.1123/mc.2024-0132.

[68] Joanna Magdalena Cieślińska-Świder e Janusz Wiesław Błaszczyk. «Posturographic Characteristics of the Standing Posture and the Effects of the Treatment of Obesity on Obese Young Women». In: *PloS One* 14.9 (2019), e0220962. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0220962.

[69] Taylor M. Kruse, Andrew L. Shim e Ian J. Bonder. «Does Body Mass Index Affect Postural Control in Adults?» In: *International Journal of Exercise Science* 18.7 (2025), pp. 394–403. ISSN: 1939-795X. DOI: 10.70252/ABWK5726.

[70] Fay B. Horak e Lewis M. Nashner. «Central programming of postural movements: adaptation to altered support-surface configurations». In: *Journal of Neurophysiology* 55.6 (1986), pp. 1369–1381. DOI: 10.1152/jn.1986.55.6.1369.

[71] Glen A Lichtwark et al. «Markerless Motion Capture Provides Accurate Predictions of Ground Reaction Forces across a Range of Movement Tasks». In: *Journal of Biomechanics* 166 (2024), p. 112051.

[72] M. M. Mukaka. «A guide to appropriate use of Correlation coefficient in medical research». In: *Malawi Medical Journal* 24.3 (2012), pp. 69–71.

Ringraziamenti

Desidero ringraziare il Prof. Nicola Giulietti per avere un incredibile energia nello spiegare e per la sua disponibiltà e competenza durante questo lavoro. E ci tengo a esprimere la mia gratitudine a Davide Todesca che mi ha sopportato e supportato anche quando ho fatto errori o non capivo qualcosa e per il suo incredibile impegno nel correggere questo lavoro che scritto da me risulta impegnativo da correggere.

E un ringraziamento particolare va al Prof. Paolo Magni, che con parole sincere e incoraggianti guida i ragazzi in questo percorso difficile. Uno dei professori che ci tiene di più al percorso degli studenti, e spero che mantenga sempre questo impegno nei loro confronti.

Questi cinque anni sono stati pieni di difficoltà non solo nello studio, ma anche nella vita privata. Fin da subito i dubbi mi hanno assalito, i fallimenti sono stati la prima realtà che ho vissuto in questo percorso. si ripetevano le frasi ripetute da chi conosceva il Nicolas di 5 anni fa: "non sei mai andato bene a scuola", "sei un somaro", "ma sei sicuro faccia per te?", ma anche "non dai mai retta, devi imparare ad ascoltare!", e meno male che non do retta, e in questi anni, ho imparato a non dare retta agli s*******, sennò avrei finito con ascoltare anche quelle voci, e non rendermi conto di quanto fantastico io sia e di quante persone fantastiche sia circondato e abbia avuto la fortuna di conoscere in questi anni.

La MAMMA, che è andata avanti inesorabilmente mentre il mondo ci cadeva addosso, che nonostante le comprensibili preoccupazioni al mio voler rimettermi a studiare, mi ha detto: "fallo!" che dopo un anno a non passare esami era sempre lì a incoraggiarmi mentre altri avrebbero già fatto ritirare il figlio. nonostante l'età, il metro e novanta di altezza e gli 86 kili, spero di poter sempre tornare da te ed essere il tuo bambino.

La ZIA LUCIA, sei stata come un terzo genitore, soprattutto nel momento del bisogno, ti ho preso spesso come esempio nell'affrontare le avversità, cercando di imitare la tua energia, grazie di essermi stata vicina e avermi ascoltato e aiutato ogni volta che ne avevo bisogno.

Il PAPÀ, in questi anni, ogni giorno, sono andato avanti con una tormentone in testa

che mi ripetevi sempre: "prima il dovere e poi il piacere", ma a vedere te, questo dovere non finiva mai. Mi hai insegnato che nella vita bisogna sempre mettersi in gioco, non bisogna mai darsi per vinti e grazie a te ho imparato a riconoscere gli s.

La ELISABETTA, la persona in grado di farmi sbottare di più in assoluto, litigi, botte, deturpamenti estetici, vandalismo, bullismo, e questo solo nei primi 15 anni di convivenza, pensate dopo. Andavamo così d'accordo che per smettere di litigare uno dei due è andato in un altro continente. Rimani la persona più importante della mia vita, nonostante i litigi e tutte le volte che ci siamo insultati anche la mamma.

Siamo una famiglia poco convenzionale, disastrati ma questo traguardo l'ho raggiunto con voi, e non vorrei altre persone con il mio stesso sangue.

SIMONE, una persona speciale, che ha aiutato un randagio come me, senza conoscermi minimamente, ed è finito, purtroppo per lui, ad accompagnarmi in questo percorso e a condividere le mie pene. Sei una delle persone migliori che conosco e ti ringrazio per essermi amico, spero di non perderti mai.

PIA, l'amica della quale non sapevo di aver bisogno, la tua gentilezza, bontà e intelligenza mi hanno aiutato in questi anni. Ad ogni problema, c'era una tisana o un caffè pronti ad aspettarmi e tu con loro pronta ad ascoltarmi. Ti voglio un mondo di bene e spero che duri negli anni.

ALESSANDRO, chi l'avrebbe mai detto che sarei diventato così amico di un bergamasco... grazie dell'aiuto e dell'amicizia che mi hai dato in questi anni.

ELENA, ti ringrazio per avermi fatto rubare i tuoi pan di stelle in questi anni, e per avermi fatto sentire sempre molto intelligente, e grazie delle tue premure.

SANTINO e MARIA, ormai andate in coppia, so che Maria non c'è quando santino mi chiede di giocare a WarThunder. Vi ringrazio per essere state persone così presenti e per essere sempre stati un componente di energia nella mia quotidianità.

Il mio compagno di merende, MARCO, c'è chi ha detto che insieme noi due non c'entriamo nulla, e penso che sia anche per questo che siamo un'accoppiata vincente.

THOMMY, ti ringrazio per essere stato una persona così fantastica, e sono incredibilmente felice che questi anni lontani non ci abbiano fatto perdere, ci tengo a ringraziare

anche la Silvia e Mauro che nonostante avessero già troppi fogli problematici tipo te, hanno sempre avuto una posto per me

I ragazzi del GREEN, vi ringrazio per le serata a giochi da tavolo e a sparare m come se non ci fosse un domani.

Le GIRLZ e i miei BOYZ/QUELLI DELLA SCAB..., i miei compagni di avventura, grazie per il supporto, per la compagnia, le uscite e tutto il divertimento che abbiamo avuto assieme, spero di non perdervi mai.

I ragazzi da BRESCIA, Simo, Dani e Michi, che nonostante la lontananza non avete mai smesso di essermi amici e avete continuato ad accogliermi anche se per mesi non ci vedevamo, vi ringrazio per la vostra amicizia e il vostro sostegno.

Le RAGAZZE della bioingegneria, AnnaMaria, Elisa, Gaia, Marika e Pier, grazie per la pizza, la compagnia e la gentilezza che mi avete dimostrato.

I CANI, che forse alcuni, non avevano ben capito che l'obiettivo della pallanuoto è tirare la palla in porta, perché continuavano a mirare alla faccia. Senza di voi questi anni sarebbero stati più vuoti e vi ringrazio per le amicizie che abbiamo coltivato e tutto il tempo passato assieme [sigla dei pirati dei caraibi].

E a tutti voi che siete qua, in questo giorno che conclude un pezzo importante della mia vita vi ringrazio.