Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2026.0429000

Classifying Spatial Navigation Performance in Healthy Adults Using Multimodal Behavioral Analysis of the Walking Corsi Test

MOBINA ZIBANDEHPOOR[1], MEHDI DELROBAEI[1], (Senior Member, IEEE)
[1]Department of Mechatronics, Faculty of Electrical Engineering, K. N. Toosi University of Technology, Tehran, Iran
Corresponding author: Mehdi Delrobaei (delrobaei@kntu.ac.ir).

ABSTRACT Cognitive navigation and spatial memory are essential for interacting with the environment, and their accurate assessment is increasingly important in cognitive research. This study presents a novel, non-invasive framework for evaluating spatial navigation abilities in healthy adults using the Walking Corsi Test (WalCT). Thirty-five participants performed the task while their movements were simultaneously recorded using camera-based pose estimation and IMU-based motion capture systems. A rich set of biomechanical features was extracted from these multimodal data, and participants also completed validated self-report questionnaires assessing spatial abilities. To classify individual navigation performance, two feature selection strategies were implemented: mutual information filtering followed by recursive feature elimination with cross-validation (MI-RFECV), and RFECV alone. The framework achieved high prediction accuracies, with navigation & orientation and spatial anxiety both reaching 85.7%, sense of direction 78.6%, and distance estimation 67.9%. SHAP-based feature importance was applied to interpret the contribution of features within each subscale-specific model. These findings demonstrate the feasibility of combining sensor-based behavioral analysis and machine learning to evaluate spatial navigation abilities in healthy adults and represent steps toward broader research.

INDEX TERMS Spatial navigation, Spatial cognition, Sensor-based assessment, Machine learning, Biomechatronic systems.

I. INTRODUCTION

THE ability to remain cognitively aware of one's environment and navigate effectively is fundamental to human functioning. These capacities support behaviors from basic orientation to complex route planning. At its core is spatial navigation, a higher-order cognitive process that integrates perceptual and decision-making mechanisms to interpret spatial relationships, form mental maps, and adapt to one's surroundings. As a key aspect of daily life, spatial navigation supports autonomy and independence, and its decline can affect quality of life and safety [1], [2], [3].

This capacity relies on cognitive maps, internal models of the environment enabling orientation and movement, encoded by neural systems such as the hippocampus and entorhinal cortex via place and grid cells [4], [5]. Navigation draws on both allocentric and egocentric frames of reference [6], supported by spatial memory from sensory input to large-scale structures [7]. Landmark anchoring enhances map precision, especially in familiar settings [4]. While adaptable, the system may degrade in unfamiliar or ambiguous environments due to unreliable cues or misinterpretation [7], [8].

Despite advances in evaluating spatial navigation, current assessment methods remain limited in validity and applicability. Laboratory-based tasks often lack ecological relevance, failing to capture the complexity of real-world navigation behaviors [9], [10]. Many tests suffer from poor standardization and inconsistent construct validity, complicating cross-study comparisons [11]. Additionally, traditional tools often overlook key spatial skills or confound cognitive performance with motor demands, particularly in clinical populations [12]. Subjective measures such as questionnaires introduce bias, while apparatus-specific effects in experimental models further constrain generalizability [13], [14]. These limitations underscore the need for objective and ecologically valid approaches to assess navigation ability more accurately [11].

Modern technologies offer scalable and ecologically valid alternatives to traditional spatial navigation assessments. Virtual Reality (VR) enables immersive yet controlled simula-

tions of complex real-world environments, such as virtual hospitals and gamified tasks like the "Virtual City Paradigm", thereby increasing ecological validity while maintaining experimental precision [15], [16]. In parallel, wearable sensing technologies provide objective, multimodal data that capture different dimensions of navigational behavior. GPS systems support real-world tracking of outdoor mobility patterns [17], while eye-tracking and electrooculography (EOG) reveal attentional and visual search strategies [18], [19]. Inertial Measurement Units (IMUs) quantify movement and body orientation [20], and EEG recordings provide insights into the neural dynamics underlying spatial processing [15].

The true advantage of these technologies might be found not in using them separately, but in combining them through multimodal fusion [21]. This fusion can simultaneously integrate complementary information across data, features, representations, or decisions, including behavioral, attentional, neural, and contextual data. In particular, combining vision-based and inertial measurements links externally observed movement with body-centered kinematics [22]. For spatial navigation classification, this approach could integrate locomotor trajectories and head orientation with environmental-context measures, thereby providing a more comprehensive representation of navigation performance.

Advancements in sensor-based data and artificial intelligence have shifted the focus from simply assessing navigation outcomes to uncovering the cognitive strategies behind them. Researchers now examine not just whether individuals succeed, but how and why they navigate. Sophisticated analysis methods are essential for interpreting this complex data. Geometric metrics have shown sensitivity to Alzheimer's-related decline [10]. Machine learning and deep learning are increasingly used to model navigation-related cognitive functions, analyze eye movements, and mimic neural activity [23], [24]. These tools help classify cognitive states such as uncertainty [15] and spatial familiarity [20], [13], with Multilayer Perceptrons (MLPs) and Convolutional Neural Networks (CNNs) providing robust, unbiased pipelines [25], [14]. Deep reinforcement learning (DRL) is also being explored to model spatial learning and memory, linking AI with neuroscience [26]. Although effective for specific prediction tasks, current models do not fully capture the dynamic integration of sensory cues, memory, and attention, nor the flexible switching between egocentric and allocentric strategies that characterizes human navigation.

While recent advances in spatial navigation research have introduced sophisticated tools, many existing assessments remain screen-based and overlook the embodied nature of real-world navigation. In this study, we present a novel, non-invasive framework that applies machine learning and feature-level multimodal fusion to kinematic data collected during the Walking Corsi Test (WalCT). Thirty-five participants completed the Mini-Mental State Examination (MMSE), the Santa Barbara Sense of Direction (SBSOD), and the Wayfinding Questionnaire. Their head and body movements were recorded synchronously using IMU-based motion capture and video-based pose estimation, respectively. From these modalities, we independently extracted trajectory and head-motion features and integrated them into a unified feature space before feature selection and classification. Multiple ensemble-based classifiers were subsequently trained and evaluated for each questionnaire subscale, and the best-performing model feature selection configuration was retained. The resulting framework classified individual differences in navigation performance against standardized self-report measures.

II. LITERATURE REVIEW

A variety of self-report instruments have been developed to assess different facets of spatial navigation. The SBSOD Questionnaire, a 15-item instrument, is widely used for evaluating general spatial abilities in healthy individuals via its "Sense of Direction" subscale [27]. The Wayfinding Questionnaire expands on this with 22 items divided into navigation and orientation, distance estimation, and spatial anxiety, making it particularly suitable for clinical populations [28]. Both utilize a 7-point Likert scale. Other tools focus on navigation strategies: the Questionnaire on Spatial Representation (QSR) employs 11 to 16 items to identify survey, route, or landmark-based approaches [29], while the Navigational Strategy Questionnaire (NSQ) uses 44 items to classify styles as egocentric, procedural, or allocentric [30]. Despite their utility, these self-report measures rely on introspective accuracy. Their validity is often reduced in aging or clinical populations by cognitive biases, limited self-awareness, or memory deficits.

In contrast, VR technologies provide immersive, controllable environments for assessing and training navigation. The Spatial Navigation Paradigm (SNaP) framework enables the implementation of diverse paradigms with high experimental control for both novice and expert users. Similarly, the Virtual Navigation Toolbox (VNT) provides a modular, Unity-based platform for constructing custom navigation experiments, such as homing or landmark-sensitivity triangle completion tasks [31]. Innovations like AudioMaze have further expanded this field by introducing auditory-only cues, enabling the collection of EEG and motion capture data during virtual maze exploration [32]. A shared limitation across these tools remains the lack of standardization and the significant technical expertise required, which can hinder reproducibility and accessibility across research settings.

Physical tasks offer a direct means of assessing navigation through real-world movement. The Triangle Completion Task (TCT) involves blindfolded participants traversing two sides of a triangle and attempting to return to the start, providing measures of spatial estimation and body-based cues [33]. Egocentric tasks test the ability to update spatial information from a specific viewpoint within a 3 m $\times$ 2.4 m area, while allocentric tasks require recalling locations based on external landmarks to target environment-based processing [6]. The "Magic Carpet" (MC) extends this by requiring participants to replicate illuminated tile sequences

through whole-body movement, directly linking spatial memory to motor execution [34]. While these tasks provide insight into observable behaviors, they do not permit precise tracking of internal cognitive and motor processes without integrated sensors, restricting the ability to capture subtle, moment-to-moment changes in spatial orientation, attention, and decision-making.

Researchers are increasingly employing sensor-based methodologies to capture navigation across various contexts. One approach involves navigating a controlled, street-like environment where participants search for a hidden goal. This paradigm includes a learning phase and a probe trial, where wall rotation generates a conflict between geometric and landmark cues. The setup integrates a ten-camera optoelectronic motion capture system to record body kinematics and a head-mounted eye-tracker to capture gaze, enabling detailed trajectory reconstruction and visual attention analysis [18]. Conversely, passive video-watching paradigms like the Leiden Navigation Test (LNT) focus on visual-cognitive dimensions. By monitoring ocular activity via EOG headbands [19], researchers can analyze events such as saccades, fixations, and blinks to infer navigation-relevant cognitive functions [35].

Recent developments in pose estimation have introduced powerful tools for objective movement analysis. By extracting skeletal keypoints from video, these methods allow for precise quantification of posture, orientation, and movement trajectories. Such capabilities have been used to characterize imitation impairments and reduced interpersonal synchrony in autism spectrum disorder [36], identify hyperactivity-related motor patterns in ADHD [37], and detect affective indicators like gaze aversion and reduced head mobility in mood disorders [38]. These applications highlight the value of pose estimation for evaluating subtle motor signatures during naturalistic, goal-directed tasks.

Despite these advancements, several methodological limitations remain. Most studies focus on gross body movements, spatial outcomes, or eye-tracking, often overlooking subtle cues like head movement during embodied spatial tasks. Moreover, while pose estimation algorithms have advanced, their application in navigation assessments remains limited. Specifically, in assessments like the WalCT, which provides an ecologically valid measure of spatial working memory, analyses have been restricted to individual trials. This scope overlooks rich, continuous kinematic data that could capture complex cognitive and spatial encoding processes, leaving the embodied and temporal nuances of navigation behavior underexplored.

III. MATERIALS AND METHODS

This section provides an overview of the study's methodology, including participant recruitment and experimental setup, data processing, and the feature extraction and selection procedures for motion capture and video data (Fig. 1).

A. PARTICIPANTS AND PROCEDURE

Thirty-five healthy university students (19 female, 16 male; mean age = 23.4 years, SD = 2.2) from K. N. Toosi University of Technology voluntarily participated in this study. Written informed consent was obtained from all participants, including explicit confirmation that no identifying personal information or images would be published. The study was conducted in accordance with the Declaration of Helsinki, and ethical approval was granted by the Research Ethics Committee (REC) of the Faculty of Electrical Engineering, K. N. Toosi University of Technology (Reference No. 39328-B-D/03-06).

Each participant was individually assessed in a controlled environment to minimize external distractions. To assess spatial abilities, two validated self-reported instruments were administered: the 22-item Wayfinding Questionnaire and the 15-item sense of direction questionnaire. Persian versions of both tools were developed through a rigorous translation and cultural adaptation process [39]. Participants were assigned to high- and low-performance classes for each questionnaire subscale using 50% of the normalized subscale score as the threshold. Normalized scores greater than or equal to 50% were classified as high performance, whereas scores below 50% were classified as low performance. These class assignments were used as the ground-truth labels for model training. Cognitive status was evaluated using the standardized Persian version of the MMSE [40], and all participants scored within the normal range.

The experimental task was the WalCT. Participants observed an illuminated sequence of squares, memorized the order, and reproduced the path by walking to each square, briefly pausing on each, and then returning to the starting point. To minimize data noise, participants were instructed to remain silent and avoid extraneous movements throughout the task [41]. To control for learning effects, each complexity level (ranging from 2 to 9 squares) was administered twice using distinct configurations. One participant was excluded due to incomplete data, resulting in a final sample of 34 participants and 136 valid trials (four trials per participant across the first four sequence levels). Each trial was treated as a distinct sequence-level observation, and the data partitioning described was therefore performed at the sequence level.

B. EXPERIMENTAL SETUP

This study was conducted in a controlled laboratory setting designed to evaluate participants' spatial memory and navigation performance using the WalCT [42]. The recording setup, hardware, software, and embedded procedural steps are described in detail.

1) Test environment and configuration

The setup consisted of nine 30 cm $\times$ 30 cm squares arranged on the floor. Each square featured a central cool-white LED as a visual cue, outlined by thin green plastic strips to enhance visibility. The LED sequences were controlled via an Arduino

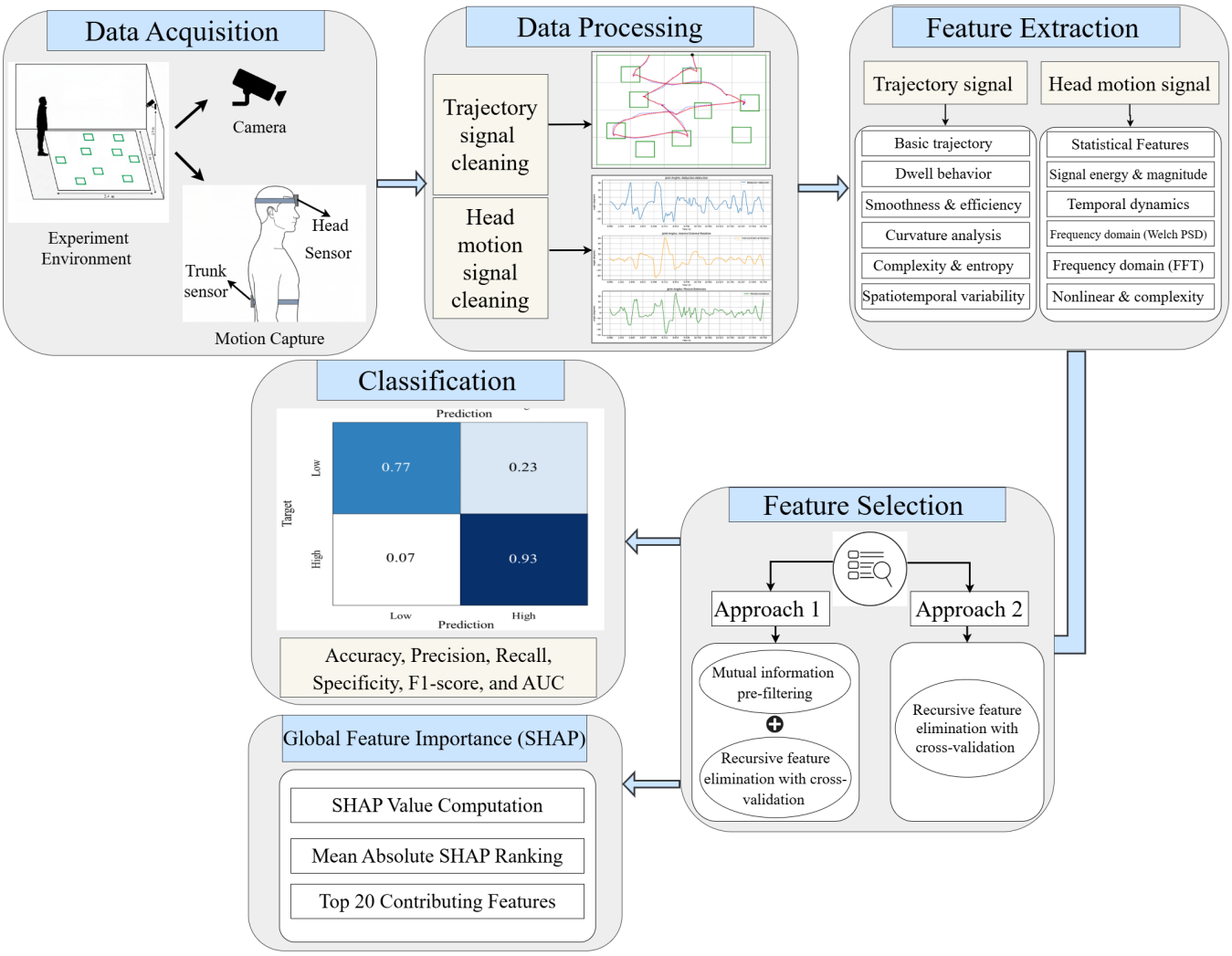


FIGURE 1. **Architecture of the proposed system.**

Uno and custom software, allowing for programmable, real-time stimulus presentation.

2) Data acquisition hardware

Two systems were employed to capture motion data:

- **Video data:** A Microsoft LifeCam Studio webcam recorded body trajectories at 30 FPS. Mounted at a height of 2.5 m on the wall opposite the test area, the camera provided an unobstructed, full-body view of the test floor, minimizing potential occlusion.
- **Motion capture data:** Head and trunk motion were recorded using the BSN-Lab motion capture system (Fig. 2) at a 250 Hz sampling rate (4 ms interval). One sensor was attached to the participant's forehead to capture head displacement and orientation, while a second sensor on the trunk served as a stable reference.

Temporal synchronization between modalities was achieved by interfacing both systems with a single laptop, ensuring all data were recorded against a common system clock. During post-processing, the video frame indices were mapped to absolute timestamps, allowing for direct alignment with the motion capture data stream.

C. DATA PROCESSING

Raw data from the video and motion capture systems were processed through distinct pipelines to prepare them for analysis. Video data underwent a computer vision workflow to correct geometric distortions and isolate spatial trajectories, while motion capture signals were cleaned and corrected to isolate relevant head-movement intervals.

1) Video data

Preprocessing involved several steps to reconstruct movement trajectories within the test area. First, a reference frame was selected to define floor boundaries. To enhance edge visibility for the detection algorithm, frames were converted to grayscale and subjected to histogram equalization. A structure-tensor-based corner detection algorithm localized

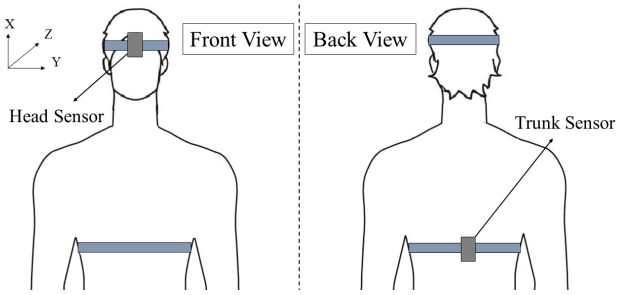


FIGURE 2. **Schematic of the front and back views showing the placement of the motion capture sensors to record head angular movement during the tasks.**

the square boundaries with sub-pixel accuracy, enabling the computation of a perspective transformation matrix to map image coordinates to a real-world reference space [43], [44].

Full-body keypoints were extracted using the OpenPose BODY-25 model [45]. Temporal consistency was maintained via a person-selection algorithm employing position-based matching and confidence scoring. As shown in Fig. 3(a), the model detects and overlays keypoints; the frame-wise average of the left and right toe coordinates was used to represent the participant's trajectory (Fig. 3(b)).

Trajectory validity was ensured through multi-stage filtering. Points outside the designated starting line or the test area (verified via a point-in-polygon algorithm) were excluded. Segments containing low-confidence keypoints were replaced using Hermite interpolation to preserve movement dynamics [46]. To stabilize boundary conditions against tracking artifacts or rotation-induced errors, artificial padding was applied to the beginning and end of each sequence.

The phase concluded with trajectory smoothing using a Kalman filter implemented on a four-dimensional motion model (position x, y and velocity v_x, v_y) [47]. Adaptive noise parameters balanced smoothing against trajectory fidelity to ensure robustness against jitter [48].

2) Motion capture data

Raw signals from the head and trunk sensors were stored as time-stamped files containing angular kinematics across three planes: flexion–extension, abduction–adduction, and internal–external rotation (Fig. 4). Signals were first cleaned of baseline offsets. To isolate head motion relative to the trunk, trunk angles were subtracted from head angles after alignment to a standard anatomical reference frame. Artifacts and signal jumps from transient sensor misalignment were identified using an adaptive threshold (± 3 standard deviations from the moving average) and corrected through cumulative offset adjustment [49].

Next, an adaptive filtering approach removed high-frequency noise [50]. Power spectral density (PSD) was computed using the Welch method [51], which identified primary noise concentration in the 6–12 Hz range. A logarithmic thresholding method defined the cutoff as the lowest frequency where the PSD dropped by more than 3 dB; otherwise, a default 8 Hz cutoff was used [52]. A sixth-order zero-phase Butterworth filter was then applied, utilizing signal reflection at the edges to minimize boundary artifacts [53].

Finally, to address slow-varying trends and drifts, polynomial models (orders 0–10) were fitted to each signal. The optimal model was selected using the Bayesian Information Criterion (BIC) [54], [55] to balance goodness-of-fit with model complexity. The resulting polynomial trend was subtracted, yielding a detrended signal suitable for feature extraction.

D. FEATURE EXTRACTION

Following signal preprocessing, a structured feature extraction process was implemented; the specific features derived from video-based trajectory data are enumerated in Table 1, while those extracted from head motion data are provided in Table 2. This step transformed the preprocessed trajectory and angular displacement signals into quantitative descriptors capturing underlying motor control patterns and movement strategies during spatial navigation tasks.

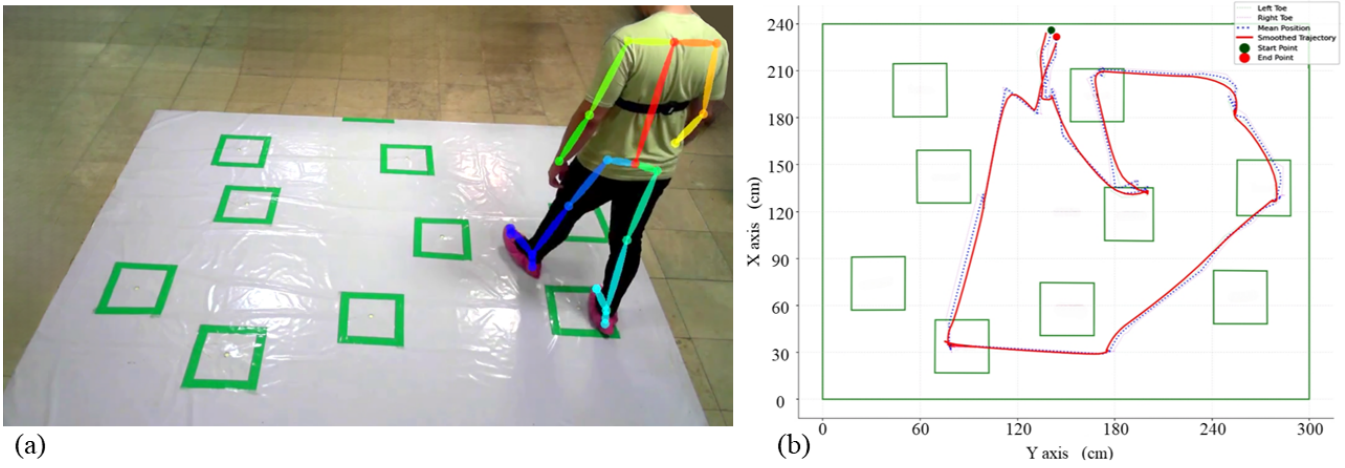


FIGURE 3. **Video data preprocessing pipeline for extracting movement trajectories. (a) Sample video frame demonstrating keypoint detection using OpenPose. (b) Reconstructed trajectory illustrating the final output of the video processing pipeline.**

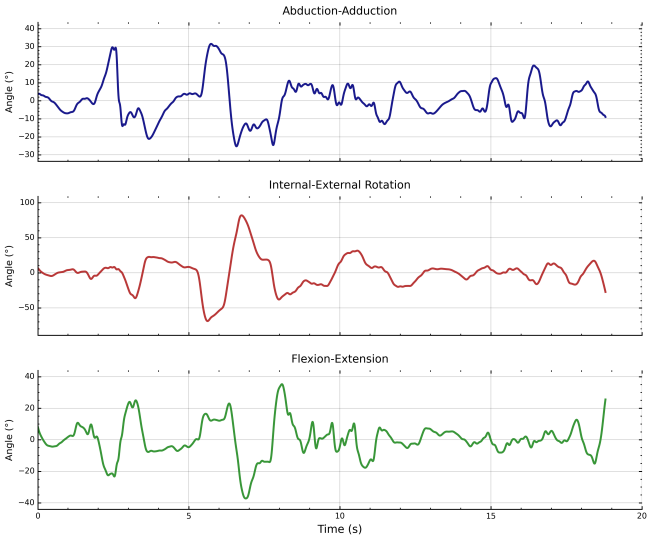


FIGURE 4. **Sample output signals representing participant head angles in the three principal axes: flexion-extension, internal-external rotation, and abduction-adduction.**

1) Pose-based feature extraction

Pose-based features were extracted from participants' trajectories during the WalCT using a multi-dimensional analysis framework. The smoothed trajectory data were processed through six analytical modules, each computing specific aspects of movement behavior. The feature sets and categories are summarized in Table 1.

- **Basic trajectory** [56] metrics quantify fundamental aspects of locomotor behavior. These include classical kinematic descriptors such as total distance traveled, movement duration, and velocity/acceleration profiles. Additional features such as time to peak velocity, deceleration duration, and movement intermittency, derived via prominence-based peak detection, offer insights into rhythm, motor control stability, and initiation dynamics.
- **Dwell behavior** [56] analysis characterizes stopping patterns through a multi-criteria detection algorithm that incorporates spatial constraints (e.g., within target squares), velocity thresholds, and temporal persistence of low-speed frames. These features serve as proxies for cognitive demands, spatial decision-making uncertainty, and task-oriented planning strategies.
- **Smoothness and efficiency** [57], [58] metrics assess the quality and fluency of movement execution. Normalized jerk and RMS jerk capture kinematic smoothness, while path efficiency (ratio of actual to optimal trajectory length) and deviation-based metrics evaluate navigational precision and route adherence.
- **Curvature analysis** [59] applies principles from differential geometry to analyze trajectory shape and turning dynamics. Curvature is computed from the cross-product of velocity and acceleration vectors, normalized by speed cubed, allowing quantification of turning sharpness, corrective maneuvers, and navigational path planning.
- **Complexity and entropy** [59] measures provide indices of movement unpredictability and structural regularity. Approximate entropy (ApEn) reflects the temporal consistency of velocity fluctuations, while the Lyapunov exponent estimates sensitivity to initial conditions and system stability. Path tortuosity and directional change frequency serve as indicators of spatial disorientation or inefficient decision-making.
- **Spatiotemporal variability** [60] features capture intra-individual fluctuations in movement timing and spatial positioning. Metrics such as spatial and temporal variance, time–distance correlation coefficients, and speed consistency indices provide a window into the temporal organization and motor reliability of the trajectory.

This feature extraction framework enables a multi-dimensional characterization of human movement behavior, supporting assessments and fundamental research into relationships between motor execution, spatial cognition, and navigation performance.

For the motion capture data, we derived 57 features from categories designed to capture the temporal, frequency, and

TABLE 1. **Comprehensive feature categories and metrics extracted from video-based walking trajectories.**

Category	Representative Features
Basic trajectory [56]	Total distance, execution time, velocity statistics (avg/max/min/range), acceleration metrics, time to peak velocity, movement intermittency
Dwell behavior [56]	(Total/average) dwell duration, dwell time variability, number of dwelling episodes
Smoothness & efficiency [57], [58]	Normalized jerk, Root Mean Square (RMS) jerk, jerk variability, path efficiency ratio, redundancy index, trajectory deviations (max/mean/total), Root mean square error (RMSE)
Curvature analysis [59]	(Mean/ max/ std/ total) curvature, curvature per meter
Complexity & entropy [59]	Movement entropy, path variance, direction changes, path tortuosity, approximate entropy (ApEn), Lyapunov exponent
Spatiotemporal variability [60]	(Spatial/ temporal) variance, time-distance correlation, speed consistency index

nonlinear characteristics of head movement relative to trunk orientation across each of the three anatomical planes (flexion–extension, abduction–adduction, and internal–external rotation). These features were selected based on established principles in motor behavior analysis, biomedical signal processing, and movement assessment to ensure comprehensive characterization of head movement dynamics. The process yielded a total of 171 features per participant (57 features $\times$ 3 movement planes), providing a rich multidimensional representation of head movement patterns during the tasks.

From the video data, 41 features were extracted from the reconstructed body trajectory. The framework was designed to capture multiple dimensions of human movement behavior during the tasks. These features were organized into six categories, each targeting specific aspects of movement behavior indices.

2) Kinematic feature extraction from motion capture data

To characterize head dynamics during navigation, we extracted feature sets from the preprocessed motion capture signals, which captured angular displacement across three primary planes. As shown in Table 2, the selected features span temporal, frequency, and nonlinear domains to capture a wide range of motion behaviors and motor control characteristics.

- **Basic statistical** [61] descriptors summarize the distributional properties of head orientation signals over time. Metrics such as mean, median, variance, skewness, and kurtosis characterize central tendency and shape, while interquartile range (IQR) and median absolute deviation provide robust dispersion estimates that are less sensitive to outliers than conventional standard deviation.
- **Signal energy and magnitude** [61] metrics quantify the amplitude and energetic intensity of head movements. RMS, SMA, and area under the curve capture the overall magnitude of movement, while signal energy and total power characterize the strength of motor output over time, reflecting the effort associated with head motion.
- **Temporal dynamics** [62] describe the kinematic evolution of movement using first-, second-, and third-order angular derivatives. Angular velocity, acceleration, and jerk offer insights into motion speed, smoothness, and neuromotor control precision. Metrics such as zero and $\pm 15^{\circ}$ threshold crossings assess directional changes, while stability variance reflects consistency in head movement.
- **Frequency domain features (Welch PSD)** [63] employ Welch's method for power spectral density estimation to characterize oscillatory patterns. Dominant and mean frequencies capture primary periodic components, while spectral entropy, bandwidth, and high-frequency content ratio reflect the complexity, variability, and potential noisiness of the movement signals.
- **Frequency domain features (FFT)** [64] utilize Fast Fourier Transform to provide detailed spectral analysis, including spectral moments (skewness, kurtosis), harmonic structure, and DC component analysis. These metrics capture asymmetries, peak sharpness, and the regularity of frequency distributions, offering insight into the rhythmic structure and organization of head motion.
- **Nonlinear and complexity** [65] measures assess the irregularity and information content of movement signals. Shannon entropy quantifies signal unpredictability and randomness, while the Teager–Kaiser Energy Operator (TKEO) detects transient, high-frequency energy bursts indicative of abrupt neuromotor adjustments or cognitively mediated motor responses.

These feature sets enable detailed characterization of motor strategies, movement efficiency, and underlying cognitive and neuromotor control mechanisms. The combination of time-domain, frequency-domain, and nonlinear features provides a multifaceted view of head movement dynamics that can

TABLE 2. Comprehensive feature categories and metrics extracted from head motion signals.

Category	Features
Statistical [61]	Mean, median, minimum, maximum, range, mode, variance, standard deviation, coefficient of variation, RMS, skewness, kurtosis, IQR, median absolute deviation (MAD)
Signal energy & magnitude [61]	Signal energy, Signal Magnitude Area (SMA), area under curve, root mean square deviation (RMSD), total power, peak amplitude measures
Temporal dynamics [62]	Angular velocity statistics (mean/variance/peak), angular acceleration metrics (mean/variance/peak), angular jerk (mean), total angular displacement, stability variance, (zero/threshold ($\pm 15^{\circ}$)) crossings
Frequency domain (Welch PSD) [63]	Total power, dominant frequency, mean frequency, frequency variance, bandwidth (90% energy), spectral flatness, spectral centroid, high-frequency content ratio, spectral entropy
Frequency domain (FFT) [64]	Total power, peak frequency, mean frequency, energy, spectral centroid, spectral variance, spectral flatness, spectral skewness, spectral kurtosis, dominant frequency amplitude, harmonic energy, spectral entropy, high-frequency energy ratio, bandwidth, DC component
Nonlinear & complexity [65]	Shannon entropy, TKEO

reveal subtle differences in navigation-related motor behavior patterns.

E. FEATURE SELECTION

To prepare the dataset for analysis, missing values were imputed according to the distribution of each feature, as assessed using the Kolmogorov–Smirnov test. Normally distributed features were imputed with the mean, whereas non-normally distributed features were imputed with the median to improve robustness. This procedure ensured a complete dataset for subsequent feature selection and classification [66]. The trajectory and head-motion features were then combined into a unified feature set, enabling feature-level fusion of video-based trajectory data and IMU-based head kinematics.

To ensure generalizability and mitigate overfitting in our high-dimensional dataset, we employed two complementary feature selection strategies. These approaches were designed to retain the most discriminative variables while preserving potential inter-feature interactions essential for accurately classifying spatial navigation ability. By applying these strategies across all subscales, we identified the optimal combination of model and feature subset, aiming to maximize classification accuracy and predictive reliability.

- **Mutual information pre-filtering with recursive feature elimination cross-validation (MI-RFECV)**
  This two-stage pipeline integrated a filter-based method with a wrapper-based approach to optimize feature subset selection. First, mutual information (MI) scores were calculated between each feature and the target class labels to quantify statistical dependencies [67]. The top 50 features (approximately 25% of the full feature space) with the highest MI scores were retained [68]. A fixed random state was used to ensure reproducibility. This pre-selection step improved computational efficiency while preserving the most informative features for subsequent analysis.
  Next, recursive feature elimination with cross-validation (RFECV) was independently applied using three ensemble models—Random Forest, XGBoost, and Gradient Boosting—as base estimators. Features were eliminated iteratively with a step size of one, and classification accuracy served as the evaluation metric. To ensure robust validation and mitigate overfitting, a 10-fold K-Fold cross-validation strategy was employed. This systematic procedure yielded compact yet highly discriminative feature subsets tailored to each classifier.
- **RFECV without pre-filtering**
  To preserve potential synergistic interactions among features that might be discarded during initial filtering, the second approach applied RFECV directly to the full feature set. This method also used 10-fold K-Fold cross-validation to ensure reliable performance estimation, with classification accuracy as the primary evaluation metric, in line with best practices in studies prioritizing interpretability, reproducibility, and generalizability [69].

Each feature selection strategy was paired with the three classifiers—Random Forest, XGBoost, and Gradient Boosting—yielding nine model–feature selection pipelines per subscale. To enable fair comparison across configurations, a composite performance score was computed by averaging the evaluation metrics. The optimal feature selection strategy and classifier combination for each subscale was determined as the configuration with the highest average composite score, ensuring robust predictive performance in the binary classification tasks across all dimensions of navigation assessment.

IV. RESULTS

To develop robust classifiers capable of distinguishing between high and low navigational performance and spatial anxiety, we implemented two distinct feature selection strategies prior to model training to address the high dimensionality of the input dataset. Each strategy was combined with three classification algorithms—Random Forest, XGBoost, and Gradient Boosting—paired with different RFECV estimators, resulting in a set of model-feature selection pipelines. Furthermore, hyperparameter tuning was performed for each pipeline to optimize predictive performance. The detailed configuration and evaluation metrics for all models implemented for each subscale are presented in Table 3.

The three classifiers were selected for their complementary tree-based ensemble strategies, which model nonlinear relationships and complex feature interactions in high-dimensional behavioral data. Their ability to handle potentially noisy features and to provide feature-importance estimates also supported consistent feature selection, model comparison, and interpretation across the four subscales.

To address class imbalance within the dataset, we first analyzed the class distribution across the four subscales. We identified a notable class imbalance in the navigation & Orientation, as well as the Sense of direction. For these subscales, an enhanced class weighting strategy was implemented to mitigate bias toward the majority class [70]. Specifically, a boost factor of 2.5 was applied to increase the penalty for misclassification of minority class samples; this value was empirically selected after evaluating multiple boost factors ranging from 2.0 to 5.0 in increments of 0.5. The resulting class weights were integrated into the model training process to support balanced learning.

Given the low sample-to-feature ratio, Gaussian noise augmentation with a factor of 0.05 was applied exclusively to the training data as a regularization strategy. The perturbed samples were intended to improve robustness to small variations in the feature space, while dimensionality was controlled through feature selection [71].

To separate model development from final evaluation, a two-stage sequence-level evaluation protocol was adopted. 20% of the observations were withheld from model development and used only for final performance evaluation. The remaining 80% were used for training and internal validation, within which stratified five-fold cross-validation combined with grid search was applied for model selection and hy-

TABLE 3. Comparative performance analysis of various feature selection and classification model combinations across different spatial navigation subscales. Abbreviations: GB = gradient boosting; XG = XGBoost; RF = random forest; RFECV est = estimator used in recursive feature elimination with cross-validation; Feat = number of selected features; Acc = accuracy; Prec = precision; Rec = recall; Spec = specificity; F1 = F1 score; est/tree = number of estimators or trees; m_d = maximum tree depth. '–' indicates that the parameter is not applicable for that model. Bolded rows indicate the top-performing model within each subscale.

Subscale	Feat method	Classifier	RFECV est	Feat	Acc	Prec	Rec	Spec	F1	est/tree	m_d
Navigation & orientation	**RFECV**	**GB**	**XG**	**4**	**85.7**	**100**	**80.0**	**100**	**88.9**	**100**	**3**
	RFECV	RF	XG	4	82.1	94.1	80.0	87.5	86.5	200	5
	MI-RFECV	GB	XG	38	78.6	93.7	75.0	87.5	83.3	100	3
	MI-RFECV	GB	RF	25	75.0	93.3	70.0	87.5	80.0	50	5
Distance estimation	**MI-RFECV**	**GB**	**XG**	**31**	**67.9**	**66.7**	**71.4**	**64.3**	**69.0**	**100**	**3**
	RFECV	RF	XG	30	64.3	62.5	71.4	57.1	66.7	100	–
	RFECV	XG	XG	30	60.7	61.5	57.1	64.3	59.3	100	6
	RFECV	GB	XG	30	60.7	60.0	64.3	57.1	62.1	200	5
Spatial anxiety	**RFECV**	**RF**	**XG**	**51**	**85.7**	**82.3**	**93.3**	**76.9**	**87.5**	**100**	**--**
	RFECV	RF	GB	80	78.6	76.5	86.7	69.2	81.2	100	5
	RFECV	XG	RF	10	78.6	71.4	100	53.8	83.3	100	6
	MI-RFECV	GB	XG	30	78.3	73.7	93.3	61.5	82.3	50	3
Sense of direction	**MI-RFECV**	**XG**	**RF**	**17**	**78.6**	**89.5**	**80.9**	**71.4**	**85.0**	**100**	**6**
	MI-RFECV	GB	RF	17	75.0	88.9	76.1	71.4	82.0	50	3
	RFECV	XG	XG	32	75.0	79.1	90.4	28.6	84.4	100	6
	MI-RFECV	XG	XG	41	71.4	80.9	80.9	42.9	80.9	200	6

perparameter optimization. Given the limited dataset, cross-validation allowed more efficient use of the available development data. For all classifier models, hyperparameter tuning focused on two critical parameters: the number of estimators (i.e., the number of decision trees in the ensemble) and the maximum depth of each tree. Prior to training, all features were standardized using z-score normalization to ensure consistent scaling across variables.

Model performance was assessed using six metrics: accuracy, precision, recall, specificity, F1-score, and the AUC. To facilitate a holistic comparison across models and feature selection strategies, a composite performance score was computed. Final model rankings were determined based on the scores obtained across the cross-validation folds. The receiver operating characteristic (ROC) curves and corresponding AUC values for the best-performing models across the four subscales are illustrated in Fig. 5, further supporting the consistency and generalizability of the classifiers across multiple validation folds.

The proposed framework demonstrated strong predictive performance across the navigation subscales. Table 3 presents the top four classifier–feature selection combinations, ranked by accuracy, for each subscale. For Navigation & orientation, the combination of RFECV-based feature selection with a Gradient Boosting classifier using the XGBoost RFECV estimator achieved 85.7% accuracy and an F1-score of 88.9%. Similarly, for spatial anxiety, the RFECV–Random Forest pipeline, combined with the XGBoost RFECV estimator, yielded 85.7% accuracy and an F1-score of 87.5%. In the case of Sense of Direction, the MI-RFECV feature selection method paired with an XGBoost classifier and Random Forest RFECV estimator performed best, achieving 78.6% accuracy and an F1-score of 85%. Although distance estimation exhibited comparatively lower performance, the MI-RFECV strategy combined with a Gradient Boosting classifier and XGBoost RFECV estimator still attained 67.9% accuracy and an F1-score of 69%, indicating that the model had greater difficulty capturing the features linked to distance estimation compared to other subscales. These results suggest that the proposed approach is highly effective in predicting navigation & orientation and spatial anxiety, moderately effective for sense of direction, but less suitable for reliably predicting distance estimation, indicating the need for further investigation or alternative modeling strategies for this subscale.

Fig. 6 presents the normalized confusion matrices for the best-performing classification models corresponding to each spatial navigation subscale: (a) navigation & orientation, (b) distance estimation, (c) spatial anxiety, and (d) sense of direction. These matrices demonstrate the classification performance in terms of sensitivity and specificity across both "Low" and "High" target categories. Notably, the models for navigation & orientation and spatial anxiety exhibit high predictive accuracy with minimal false positives and false negatives, whereas the distance estimation subscale shows relatively lower classification consistency.

To further interpret model decisions and understand the

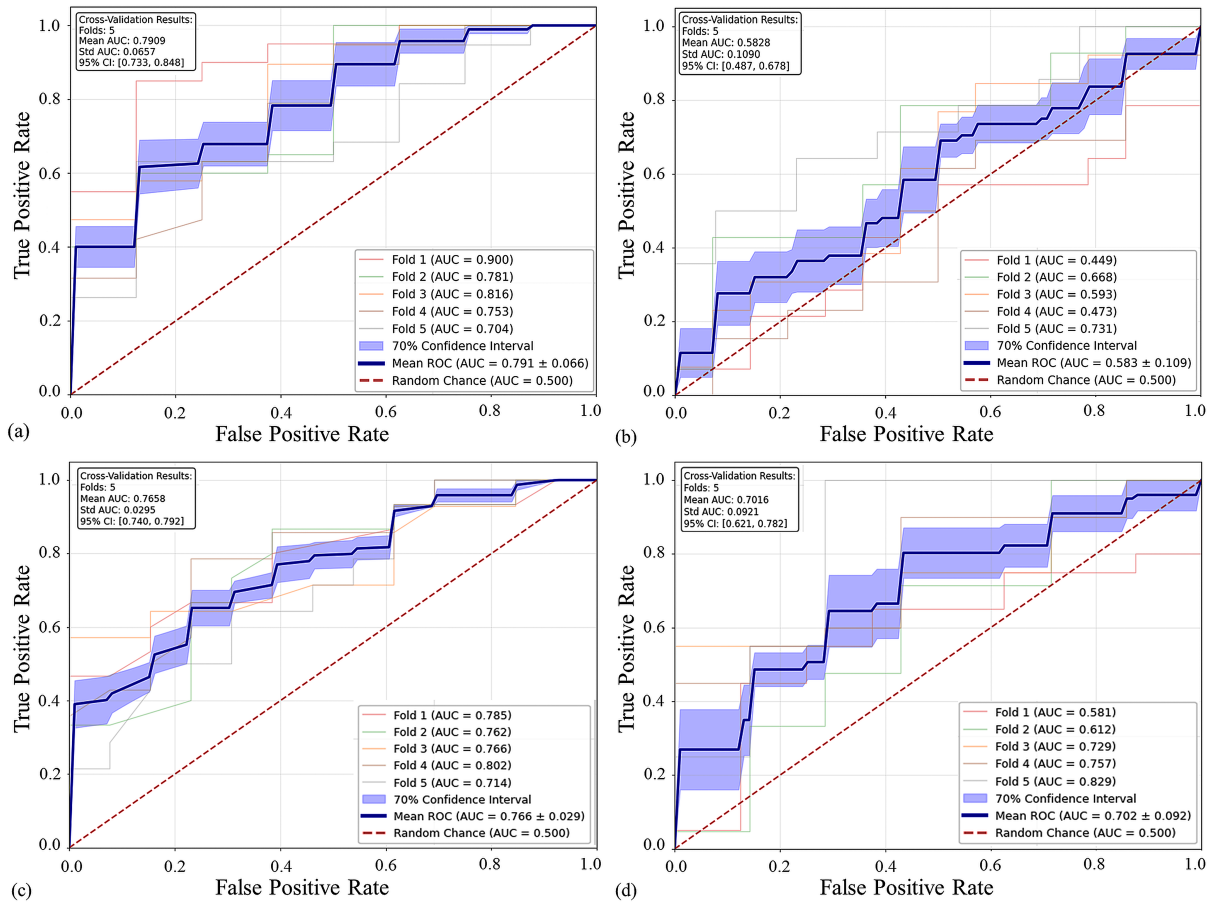


FIGURE 5. ROC curves illustrating the predictive performance of the optimized machine learning models for classifying human spatial navigation abilities across the four subscales: (a) navigation & orientation, (b) distance estimation, (c) spatial anxiety, and (d) sense of direction. The shaded blue regions represent the standard deviation of the ROC curve across folds, indicating model robustness. The dashed diagonal line (y=x) represents the performance of a random classifier. AUC values for each model, reflecting overall discriminative power, are provided in the legend.

contribution of individual features, SHAP (SHapley Additive exPlanations) [72] values were computed. The top 20 features with the highest mean absolute SHAP values, reflecting their overall impact on model outputs, are ranked and visualized in Fig. 7. This analysis provides valuable insight into the most influential variables across different models and validation schemes, offering interpretability to the predictive framework.

V. DISCUSSION

Despite the clinical significance of spatial navigation ability as an early marker of cognitive decline, existing assessment tools are constrained by artificial task environments, subjective self-reports, or an overreliance on outcome metrics that fail to capture the dynamic, movement-based nature of real-world wayfinding. We introduced a movement-based cognitive assessment that captures embodied signatures of spatial navigation through real-time analysis of gait and head dynamics during the WalCT. By integrating video pose estimation with head motion sensors, our approach captures gait, turning, and head stabilization, reflecting embodied cognitive processes of real-life spatial behavior.

Our multimodal framework successfully classified individual differences in spatial navigation performance using behavioral kinematics, achieving high accuracy for navigation & orientation and spatial anxiety (85.7% each), and moderate accuracy for sense of direction (78.6%). These predictions were based not only on task results but also on subtle motor signatures such as gait, head stability, and pausing during natural walking. These findings suggest that affective and cognitive navigation traits are embodied, reflected in movement patterns. Although distance estimation was less accurate, possibly due to reliance on internal representations, the overall performance surpasses that of typical sensor-based cognitive assessments. Brief walking tasks with high-resolution motion data effectively gauge spatial cognition, offering a reliable research baseline [20].

A key consideration in designing this framework was the simultaneous capture of head-motion kinematics and movement trajectories. Evidence shows that visual search and attentional allocation during navigation are not only expressed through eye movements but are also embodied in head movements [73]. Navigational behavior is likewise conveyed through the paths individuals choose, as route selection en-

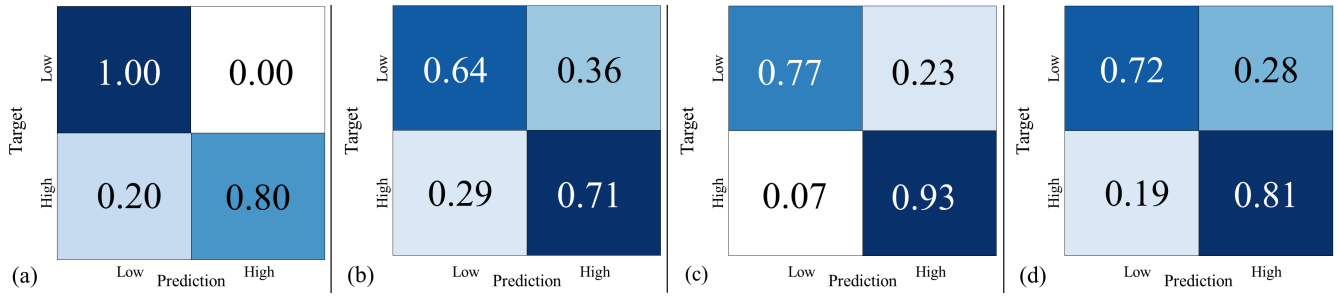


FIGURE 6. **Normalized confusion matrices for the best-performing classification models across four spatial navigation subscales. Each matrix illustrates the model's performance in distinguishing between "Low" and "High" navigation performance categories, displaying the proportions of true positives, true negatives, false positives, and false negatives. Subfigures correspond to: (a) navigation & orientation, (b) distance estimation, (c) spatial anxiety, and (d) sense of direction. The matrices highlight the models' ability to classify individuals in both performance categories correctly.**

codes strategic decisions, environmental understanding, and efficiency trade-offs [74]. Hence, participants' movement trajectories were recorded to extract navigation strategies. In tasks such as the WalCT, individuals reorient their bodies or heads to scan, locate targets, decide, and confirm spatial relationships, reflecting cognitive processes and aiding classification.

In the SHAP-based analysis, the top 20 features ranked for each subscale were examined to assess the contribution of specific feature categories (Fig. 7). For the navigation & orientation subscale, three key features were from temporal characteristics of head kinematics, and one was from the statistical category. The most influential feature was angular acceleration in the flexion-extension, highlighting the role of dynamic head motion in spatial orientation. This aligns with prior research showing head kinematics, especially rotational velocity and acceleration, as proxies for embodied visual attention and spatial reorientation [75]. Notably, the best classification performance was achieved using only four features, all related to head movement, with no trajectory-based features. This suggests that for spatial orientation tasks, sensorimotor cues from head dynamics may be more informative than path-based metrics.

For the spatial anxiety subscale, SHAP analysis indicated that head movement dynamics were central to model predictions: 10 of the top 20 features were statistical descriptors of head kinematics, and 6 were temporal metrics, underscoring the sensitivity of anxiety-related behavior to the magnitude, variability, and timing of head motion. Two trajectory features, specifically average speed and speed variability, were also among the most important. The top-ranked feature was the peak angular displacement along the flexion-extension axis, followed by angular jerk along the abduction-adduction axis. The findings suggest that individuals with high spatial anxiety may show distinct patterns in head movement magnitude and smoothness, possibly reflecting compensatory scanning or increased vigilance. The prominence of head-based features supports the idea that internal motion dynamics, especially those tied to sensorimotor control and bodily awareness, are key markers of anxiety-related spatial processing [76].

For the sense of direction subscale, frequency-domain features were the most prominent among the top 20 SHAP-ranked features. The most influential was spectral entropy in the flexion-extension plane, highlighting the role of signal complexity in this plane. PSD-derived features accounted for eight of the top features, followed by three temporal and two statistical features. Additionally, two trajectory-based features, Lyapunov Exponent and Time to Peak Velocity, were also represented. These results indicate that both the spectral characteristics of head motion and aspects of trajectory variability contribute to individual differences in sense of direction. The top-ranked features reflect the unpredictability and richness of head movement signals, which may correspond to greater environmental scanning or attentional engagement [77].

For the distance estimation subscale, SHAP analysis revealed the most heterogeneous feature distribution, with the top 20 features spanning seven distinct categories, indicating a more diffuse pattern of predictive cues. Statistical and temporal features of head movement were most prominent, each contributing six features, while the median angular displacement in the abduction–adduction plane was identified as the most influential. Despite this variety, the model's performance was lower (67.9% accuracy) compared to other subscales. Given prior evidence of strong to moderate correlations between distance estimation and both navigation & orientation (r = 0.648) and spatial anxiety (r = 0.382) in healthy adults [28], higher accuracy was anticipated, similar to other subscales. The current feature sets may therefore lack sensitivity to the visuospatial transformations and environmental scaling cues essential for accurate distance judgments, highlighting a need for richer sensory or gaze-based features in future models.

Among the 20 highest-ranked features, those capturing motion in the abduction–adduction and then flexion-extension plane appeared with the greatest commonness. This finding is both interpretable and aligned with the naturalistic demands of the WalCT, where spatial cues and task stimuli are positioned on the ground, prompting frequent lateral head tilts and body adjustments [78]. Particularly, abduction–adduction angular jerk emerged as one of the top five features across

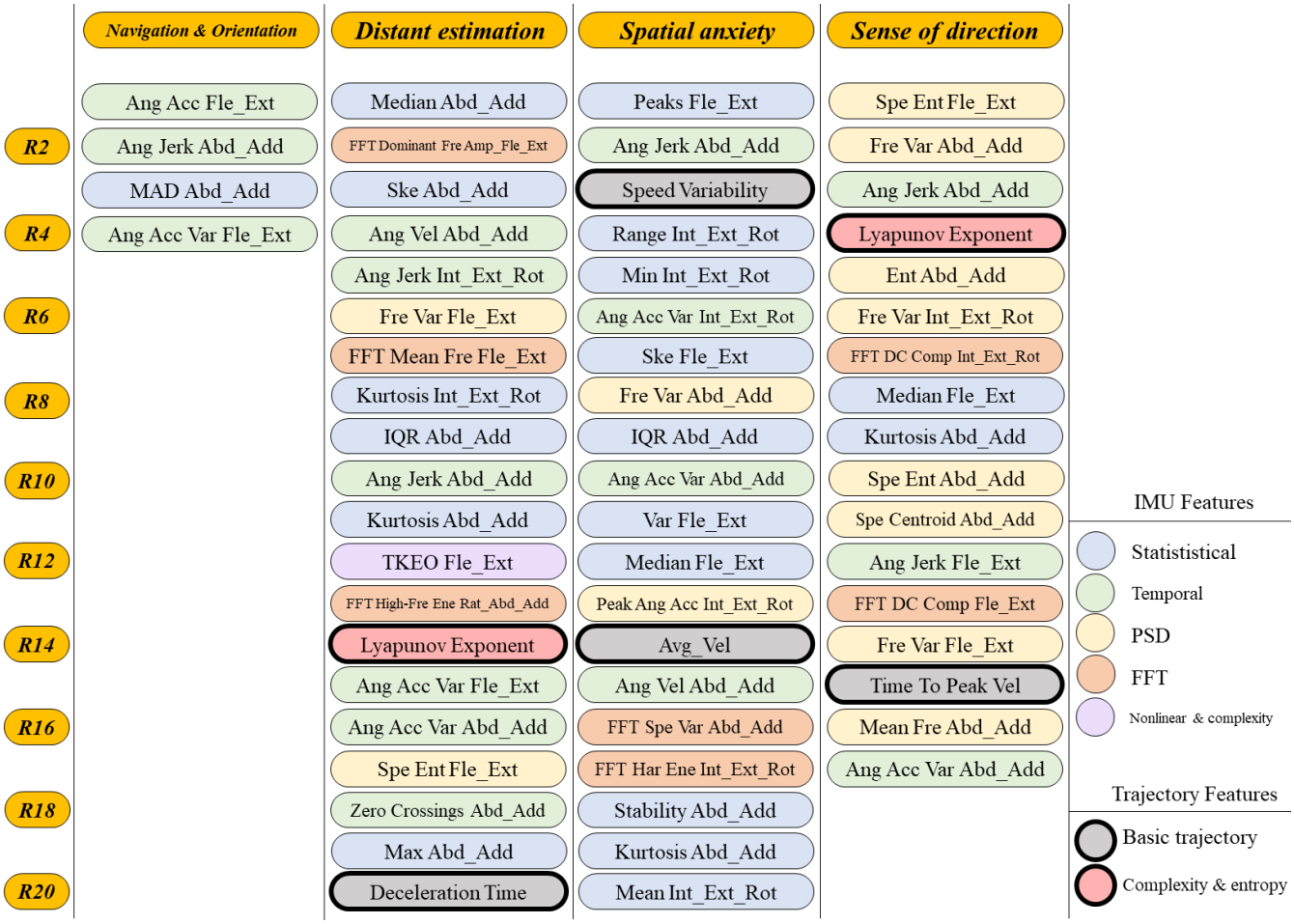


FIGURE 7. **Top 20 contributing features, as determined by SHAP values and employed through feature selection algorithms, for the high-performance models used to classify spatial navigation abilities and spatial anxiety. Trajectory features are indicated by yellow highlighting, differentiating them from head movement features. Full abbreviations: Ang (Angular), Acc (Acceleration), Fre (Frequency), Amp (Amplitude), Max (Maximum), Min (Minimum), Vel (Velocity), Var (Variance), Ent (Entropy), Spe (spectral), Ene (Energy), Ske (Skewness), Comp (Component), Rat (Ratio), Har (Harmonic), Fle_Ext (Flexion-Extension), Abd_Add (Abduction-Adduction), Int_Ext_Rot (Internal-External-Rotation).**

all four subscales, underscoring its sensitivity to subtle navigational behaviors and its potential as a robust marker of spatial navigation skills. Furthermore, abduction–adduction kurtosis and abduction–adduction angular acceleration variation were among the most frequent features, appearing in three subscales, reinforcing the idea that dynamic and higher-order characteristics of lateral head motion are particularly sensitive to variations in spatial navigation ability.

Notably, trajectory-derived features appeared far less frequently than head-movement–based features. When present, they were confined to basic trajectory descriptors and measures from the complexity/entropy category (e.g., Lyapunov Exponent), suggesting that head kinematics carried the dominant predictive signal while trajectory metrics provided only complementary cues. These findings position abduction–adduction kinematics as a core biomechanical indicator for future multimodal navigation assessments and open avenues for refining models by targeting this motion plane more deliberately.

This study has some limitations that inform the interpretation of its findings and point to future research directions. First, the narrow age range limits generalizability because older adults may differ in navigation strategies, sensory integration, gait, and head stabilization. Although the sample was nearly balanced by gender, the sample size was insufficient for reliable gender-specific analysis. As gender-related differences may influence navigation strategy, spatial anxiety, and visuospatial working memory, future studies should examine age and gender effects in larger, more diverse cohorts. Moreover, because data partitioning was performed at the sequence level, the present findings characterize generalization to WalCT sequence observations rather than to participants excluded entirely from model development.

Second, while validated self-report questionnaires were used to define performance groups, they serve as indirect proxies for cognitive traits; future work should validate the framework against direct behavioral or neurophysiological markers of spatial function. Third, the binary classification approach—dichotomizing participants into “high” and “low” performers—simplifies a continuous, multidimensional cog-

nitive ability. Although this binary design simplifies interpretation, it may obscure subtle performance gradients; models using regression or multi-class outputs could provide finer-grained representations of navigation abilities. Additionally, the absence of eye-tracking data limits our ability to distinguish overt visual attention from postural adjustments during navigation, particularly relevant given the lower accuracy for the distance estimation subscale, which may reflect insufficient capture of perceptual-metric components. Nonetheless, these constraints outline clear avenues for refinement, reinforcing the potential of this framework to evolve into a more comprehensive, multimodal tool for real-world cognitive assessment.

Together, the findings of this study establish the WalCT, augmented with multimodal motion capture and SHAP-interpretable machine learning, as a robust and ecologically grounded framework for objectively assessing spatial navigation. These results demonstrate the promise of multimodal movement analysis for advancing research on spatial cognition and lay the groundwork for the future development of digital biomarkers.

VI. CONCLUSION

This study presents a multimodal framework for the classification of spatial navigation abilities in healthy adults by integrating video-based pose estimation with head motion data collected during the WalCT. Notably, head kinematics, with an emphasis on abduction–adduction dynamics, emerged as particularly salient features, capturing nuanced biomechanical markers of spatial cognition that traditional assessments may overlook. Future research may address the limitations of this study by incorporating eye-tracking modalities to enrich understanding of visual search strategies during navigation, adopting multi-class or regression-based modeling approaches to better represent the spectrum of spatial abilities, evaluating alternative machine learning approaches, such as support vector machines (SVM), using the same validation framework to establish a benchmark against the current models, and expanding participant cohorts to include diverse and clinical populations to enhance external validity. Moreover, this study focused exclusively on the first four correctly completed WalCT trials; incorporating analyses of failed or incomplete trials could elucidate error-related behaviors, cognitive strategy failures, and task demands more comprehensively. Collectively, these methodological refinements are anticipated to enhance the interpretive depth, ecological validity, and overall applicability of sensor-driven frameworks for spatial navigation assessment.

ACKNOWLEDGEMENTS

The authors acknowledge the support and encouragement of the Iran National Science Foundation (INSF). The authors also acknowledge the use of AI-assisted tools for language editing and improving the clarity and readability of the manuscript.

DATA AVAILABILITY

The data supporting the findings of this study are publicly available on Figshare [79].

REFERENCES

[1] O. Baumann and J. B. Mattingley, "Extrahippocampal contributions to spatial navigation in humans: A review of the neuroimaging evidence," *Hippocampus*, vol. 31, no. 7, pp. 640–657, 2021.
[2] P. Tragantzopoulou and V. Giannouli, "Spatial orientation assessment in the elderly: a comprehensive review of current tests," *Brain Sciences*, vol. 14, no. 9, p. 898, 2024.
[3] L. Piccardi, R. Nori, J. M. Cimadevilla, and M. Kozhevnikov, "The contribution of internal and external factors to human spatial navigation," *Brain Sciences*, vol. 14, no. 6, p. 585, 2024.
[4] R. A. Epstein, E. Z. Patai, J. B. Julian, and H. J. Spiers, "The cognitive map in humans: spatial navigation and beyond," *Nature Neuroscience*, vol. 20, no. 11, pp. 1504–1513, 2017.
[5] R. S. Koolschijn, P. Parthasarathy, M. Browning, X. Przygodda, L. P. Capitão, W. T. Clarke, T. P. Vogels, J. X. O'Reilly, and H. C. Barron, "Noradrenaline causes a spread of association in the hippocampal cognitive map," *Nature Communications*, vol. 17, no. 1, p. 3961, 2026.
[6] C. Fernandez-Baizan, P. Nuñez, J. L. Arias, and M. Mendez, "Egocentric and allocentric spatial memory in typically developed children: Is spatial memory associated with visuospatial skills, behavior, and cortisol?" *Brain and Behavior*, vol. 10, no. 5, p. e01532, 2020.
[7] S. Becker, "17 memory for space," *The Oxford Handbook of Human Memory, Two Volume Pack: Foundations and Applications*, p. 467, 2024.
[8] W. WU and L. WANG, "The cognitive map and its intrinsic mechanisms," *Advances in Psychological Science*, vol. 31, no. 10, p. 1856, 2023.
[9] S. Penaud, D. Yeh, A. Gaston-Bellegarde, and P. Piolino, "The role of bodily self-consciousness in episodic memory of naturalistic events: an immersive virtual reality study," *Scientific Reports*, vol. 13, no. 1, p. 17013, 2023.
[10] U. Lim, R. L. Cervantes, G. Coughlan, R. Lambiotte, H. J. Spiers, M. Hornberger, and H. A. Harrington, "Geometry of navigation identifies genetic-risk and clinical alzheimer's disease," *medRxiv*, pp. 2023–10, 2023.
[11] D. H. Uttal, K. McKee, N. Simms, M. Hegarty, and N. S. Newcombe, "How can we best assess spatial skills? practical and conceptual challenges," *Journal of Intelligence*, vol. 12, no. 1, p. 8, 2024.
[12] G. Colombo, K. Minta, T. Thrash, J. Grübel, J. Wiener, M. Avraamides, W. R. Taylor, C. Hölscher, and V. R. Schinazi, "Beyond traditional assessments of cognitive status: Exploring the potential of spatial navigation tasks," *Behavior research methods*, vol. 58, no. 5, p. 126, 2026.
[13] H. Liao, W. Zhao, C. Zhang, W. Dong, and H. Huang, "Detecting individuals' spatial familiarity with urban environments using eye movement data," *Computers, Environment and Urban systems*, vol. 93, p. 101758, 2022.
[14] T. Illouz, L. A. B. Ascher, R. Madar, and E. Okun, "Unbiased analysis of spatial learning strategies in a modified barnes maze using convolutional neural networks," *Scientific Reports*, vol. 14, no. 1, p. 15944, 2024.
[15] B. Zhu, J. G. Cruz-Garza, Q. Yang, M. Shoaran, and S. Kalantari, "Identifying uncertainty states during wayfinding in indoor environments: An eeg classification study," *Advanced Engineering Informatics*, vol. 54, p. 101718, 2022.
[16] B. Del Lucchese, V. Belmonti, P. Brovedani, M. C. Caponi, A. Castilla, G. Masi, A. Tacchi, M. Zaoui, G. Cioni, and A. Berthoz, "The virtual city paradigm for testing visuo-spatial memory, executive functions and cognitive strategies in children with adhd: a feasibility study," *Frontiers in Psychiatry*, vol. 12, p. 708434, 2021.
[17] A. Ghosh, V. Puthusseryppady, D. Chan, C. Mascolo, and M. Hornberger, "Machine learning detects altered spatial navigation features in outdoor behaviour of alzheimer's disease patients," *Scientific Reports*, vol. 12, no. 1, p. 3160, 2022.
[18] M. Bécu, D. Sheynikhovich, G. Tatur, C. P. Agathos, L. L. Bologna, J.-A. Sahel, and A. Arleo, "Age-related preference for geometric spatial cues during real-world navigation," *Nature Human Behaviour*, vol. 4, no. 1, pp. 88–99, 2020.
[19] M. Zibandehpoor, F. Alizadehziri, A. A. Larki, S. Teymouri, and M. Delrobaei, "Electrooculography dataset for objective spatial navigation assessment in healthy participants," *Scientific Data*, vol. 12, no. 1, p. 553, 2025.
[20] M. Kattenbeck, I. Giannopoulos, N. Alinaghi, A. Golab, and D. R. Montello, "Predicting spatial familiarity by exploiting head and eye movements during pedestrian navigation in the real world," *Scientific Reports*, vol. 15, no. 1, p. 7970, 2025.

[21] T. Baltrusaitis, C. Ahuja, and L.-P. Morency, "Multimodal machine learning: A survey and taxonomy," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, no. 2, pp. 423–443, 2019.
[22] S. Majumder and N. Kehtarnavaz, "Vision and inertial sensing fusion for human action recognition: A review," *IEEE Sensors Journal*, vol. 21, no. 3, pp. 2454–2467, 2020.
[23] Y. Zhong, Z. Wang, X. Zhao, T. Xu, T. Zhou, and H. Hu, "Electroencephalogram-based multimodal attention level classification using deep learning techniques," *Frontiers in Human Neuroscience*, vol. 20, p. 1791677, 2026.
[24] S. Teymouri, F. Alizadehziri, M. Zibandehpoor, and M. Delrobaei, "Algorithmic derivation of human spatial navigation indices from eye movement data," *arXiv preprint arXiv:2501.10696*, 2025.
[25] I. Annaki, M. Rahmoune, M. Bourhaleb, M. Zaoui, A. Castilla, A. Berthoz, and B. Cohen, "Evaluating the efficiency of multilayer perceptron neural network architecture in classifying cognitive impairments related to human bipedal spatial navigation," in *International Conference on Digital Technologies and Applications*. Springer, 2023, pp. 54–61.
[26] E. Bermudez-Contreras, "Deep reinforcement learning to study spatial navigation, learning and memory in artificial and biological agents," *Biological Cybernetics*, vol. 115, no. 2, pp. 131–134, 2021.
[27] M. Hegarty, A. E. Richardson, D. R. Montello, K. Lovelace, and I. Subbiah, "Development of a self-report measure of environmental spatial ability," *Intelligence*, vol. 30, no. 5, pp. 425–447, 2002.
[28] M. H. Claessen, J. M. Visser-Meily, N. K. de Rooij, A. Postma, and I. J. van der Ham, "The wayfinding questionnaire as a self-report screening instrument for navigation-related complaints after stroke: internal validity in healthy respondents and chronic mild stroke patients," *Archives of Clinical Neuropsychology*, vol. 31, no. 8, pp. 839–854, 2016.
[29] F. Pazzaglia and R. De Beni, "Strategies of processing spatial information in survey and landmark-centred individuals," *European Journal of Cognitive Psychology*, vol. 13, no. 4, pp. 493–508, 2001.
[30] C. A. Lawton and J. Kallai, "Gender differences in wayfinding strategies and anxiety about wayfinding: A cross-cultural comparison," *Sex Roles*, vol. 47, pp. 389–401, 2002.
[31] M. M. Müller, J. Scherer, P. Unterbrink, O. J. Bertrand, M. Egelhaaf, and N. Boeddeker, "The virtual navigation toolbox: Providing tools for virtual navigation experiments," *PLoS ONE*, vol. 18, no. 11, p. e0293536, 2023.
[32] M. Miyakoshi, L. Gehrke, K. Gramann, S. Makeig, and J. Iversen, "The audiomaze: An eeg and motion capture study of human spatial navigation in sparse augmented reality," *European Journal of Neuroscience*, vol. 54, no. 12, pp. 8283–8307, 2021.
[33] E. X. Wei, E. R. Anson, S. M. Resnick, and Y. Agrawal, "Psychometric tests and spatial navigation: Data from the baltimore longitudinal study of aging," *Frontiers in Neurology*, vol. 11, p. 484, 2020.
[34] V. Belmonti, G. Cioni, and A. Berthoz, "Switching from reaching to navigation: differential cognitive strategies for spatial memory in children and adults," *Developmental Science*, vol. 18, no. 4, pp. 569–586, 2015.
[35] F. Mirzapour, R. Mirzapour, M. Zibandehpoor, and M. Delrobaei, "Real-time biometric analysis of eye movement patterns in human spatial navigation," in *2025 29th International Computer Conference, Computer Society of Iran (CSICC)*. IEEE, 2025, pp. 01–06.
[36] Z. Yang, Y. Zhang, J. Ning, X. Wang, and Z. Wu, "Early diagnosis of autism: A review of video-based motion analysis and deep learning techniques," *IEEE Access*, vol. 13, pp. 2903–2928, 2025.
[37] Y. Li, R. Nair, and S. M. Naqvi, "Adhd detection based on human action recognition," *Neuroscience Applied*, vol. 3, p. 104093, 2024.
[38] D. Liu, B. Liu, T. Lin, G. Liu, G. Yang, D. Qi, Y. Qiu, Y. Lu, Q. Yuan, S. C. Shuai *et al.*, "Measuring depression severity based on facial expression and body movement using deep convolutional neural network," *Frontiers in Psychiatry*, vol. 13, p. 1017064, 2022.
[39] M. Zibandehpoor and M. Delrobaei, "Persian version of wayfinding questionnaire," *arXiv preprint arXiv:2412.02143*, 2024.
[40] T. N. Tombaugh and N. J. McIntyre, "The mini-mental state examination: a comprehensive review," *Journal of the American Geriatrics Society*, vol. 40, no. 9, pp. 922–935, 1992.
[41] A. Castilla, A. Berthoz, D. Urukalo, M. Zaoui, A. Perrochon, and T. Kronovsek, "Age and sex impact on visuospatial working memory (vswm), mental rotation, and cognitive strategies during navigation," *Neuroscience Research*, vol. 183, pp. 84–96, 2022.
[42] M. Zibandehpoor and M. Delrobaei, "Multimodal classification of sense-of-direction in healthy individuals using trajectory and head motion during a walking-based spatial task," in *2026 12th International Conference on Web Research (ICWR)*. IEEE, 2026, pp. 64–69.
[43] S. Krig, "Ground truth data, content, metrics, and analysis," in *Computer Vision Metrics: Textbook Edition*. Springer, 2016, pp. 247–271.
[44] M. Zibandehpoor and M. Delrobaei, "Objective assessment of navigation trajectory in walking corsi test using pose estimation algorithms," in *2024 12th RSI International Conference on Robotics and Mechatronics (ICRoM)*. IEEE, 2024, pp. 307–312.
[45] Z. Cao, G. Hidalgo, T. Simon, S.-E. Wei, and Y. Sheikh, "Openpose: Realtime multi-person 2d pose estimation using part affinity fields," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 1, pp. 172–186, 2019.
[46] B. Zaman, D. Marijan, and T. Kholodna, "Interpolation-based inference of vessel trajectory waypoints from sparse ais data in maritime," *Journal of Marine Science and Engineering*, vol. 11, no. 3, p. 615, 2023.
[47] Y. Liu, Z. Zhou, and Y. Wang, "A tracking method of multi-workers onsite with kalman filter and openpose," in *ICCREM 2021*, 2021, pp. 272–280.
[48] J. Li, Q. Miao, Z. Zou, H. Gao, L. Zhang, Z. Li, and N. Wang, "A review of computer vision-based monitoring approaches for construction workers' work-related behaviors," *IEEE Access*, vol. 12, pp. 7134–7155, 2024.
[49] M. K. Islam, A. Rastegarnia, and S. Sanei, "Signal artifacts and techniques for artifacts and noise removal," in *Signal Processing Techniques for Computational Health Informatics*. Springer, 2020, pp. 23–79.
[50] C. E. Cruttenden, J. M. Taylor, M. Ahmadi, Y. Zhang, X.-H. Zhu, W. Chen, and R. Rajamani, "Reference-free adaptive filtering of extracellular neural signals recording in ultra-high field magnetic resonance imaging scanners: Removal of periodic interferences," *Biomedical Signal Processing and Control*, vol. 71, p. 102758, 2022.
[51] P. Welch, "The use of fast fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms," *IEEE Transactions on audio and electroacoustics*, vol. 15, no. 2, pp. 70–73, 1967.
[52] E. M. Taranta, S. L. Koh, B. M. Williamson, K. P. Pfeil, C. R. Pittman, and J. J. LaViola, "Pitch pipe: An automatic low-pass filter calibration technique for pointing tasks," *Proceedings of Graphics Interface*, 2019.
[53] K. Rácz and R. M. Kiss, "Marker displacement data filtering in gait analysis: A technical note," *Biomedical Signal Processing and Control*, vol. 70, p. 102974, 2021.
[54] G. Schwarz, "Estimating the dimension of a model," *The Annals of Statistics*, pp. 461–464, 1978.
[55] A. K. Nandi, "Theoretical and experimental results on time-series representations of polynomials," *Frontiers in Signal Processing*, vol. 5, p. 1582043, 2025.
[56] H. Masum, S. Chattopadhyay, R. Ray, and S. Bhaumik, "Measurement of walking speed from gait data using kurtosis and skewness based approximate and detailed coefficients," *IET Science, Measurement & Technology*, vol. 12, no. 4, pp. 521–527, 2018.
[57] P. Gulde and J. Hermsdörfer, "Smoothness metrics in complex movement tasks," *Frontiers in Neurology*, vol. 9, p. 615, 2018.
[58] B. Rohrer, S. Fasoli, H. I. Krebs, R. Hughes, B. Volpe, W. R. Frontera, J. Stein, and N. Hogan, "Movement smoothness changes during stroke recovery," *Journal of neuroscience*, vol. 22, no. 18, pp. 8297–8304, 2002.
[59] R. Oves Garcia, E. F. Morales, and L. E. Sucar, "Second-order motion descriptors for efficient action recognition," *Pattern Analysis and Applications*, vol. 24, no. 2, pp. 473–482, 2021.
[60] L. Alessandretti, P. Sapiezynski, S. Lehmann, and A. Baronchelli, "Multi-scale spatio-temporal analysis of human mobility," *PloS ONE*, vol. 12, no. 2, p. e0171686, 2017.
[61] J. Margarito, R. Helaoui, A. M. Bianchi, F. Sartor, and A. G. Bonomi, "User-independent recognition of sports activities from a single wrist-worn accelerometer: A template-matching-based approach," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 4, pp. 788–796, 2015.
[62] M. Janidarmian, A. Roshan Fekr, K. Radecka, and Z. Zilic, "A comprehensive analysis on wearable acceleration sensors in human activity recognition," *Sensors*, vol. 17, no. 3, p. 529, 2017.
[63] M. Gießler, B. Waltersberger, T. Götz, and R. Rockenfeller, "A multi-method framework for establishing an angular acceleration reference in sensor calibration and uncertainty quantification," *Communications Engineering*, vol. 4, no. 1, p. 65, 2025.
[64] B. Bruno, F. Mastrogiovanni, A. Sgorbissa, T. Vernazza, and R. Zaccaria, "Analysis of human behavior recognition algorithms based on acceleration data," in *2013 IEEE International Conference on Robotics and Automation*. IEEE, 2013, pp. 1602–1607.
[65] P. Maragos, J. F. Kaiser, and T. F. Quatieri, "Energy separation in signal modulations with application to speech analysis," *IEEE Transactions on Signal Processing*, vol. 41, no. 10, pp. 3024–3051, 2002.

[66] Z. Zhang, "Missing data imputation: focusing on single imputation," *Annals of Translational Medicine*, vol. 4, no. 1, p. 9, 2016.
[67] J. Brownlee, "Information gain and mutual information for machine learning," 2019.
[68] M. A. Salam, S. Basu Roy, and G. Das, "Efficient approximate top-k mutual information based feature selection," *Journal of Intelligent Information Systems*, vol. 61, no. 1, pp. 191–223, 2023.
[69] M. Awad and S. Fraihat, "Recursive feature elimination with cross-validation with decision tree: Feature selection method for machine learning-based intrusion detection systems," *Journal of Sensor and Actuator Networks*, vol. 12, no. 5, p. 67, 2023.
[70] M. Pirizadeh, H. Farahani, and S. R. Kheradpisheh, "Imbalance factor: a simple new scale for measuring inter-class imbalance extent in classification problems," *Knowledge and Information Systems*, vol. 65, no. 10, pp. 4157–4183, 2023.
[71] J. Brownlee, "Train neural networks with noise to reduce overfitting," *Machine Learning Mastery*, vol. 11, 2019.
[72] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," *Advances in neural information processing systems*, vol. 30, 2017.
[73] N. Stein, T. Watson, M. Lappe, M. Westendorf, and S. Durant, "Eye and head movements in visual search in the extended field of view," *Scientific Reports*, vol. 14, no. 1, p. 8907, 2024.
[74] T. T. Brunyé, Z. A. Collier, J. Cantelon, A. Holmes, M. D. Wood, I. Linkov, and H. A. Taylor, "Strategies for selecting routes through real-world environments: Relative topography, initial route straightness, and cardinal direction," *PLoS ONE*, vol. 10, no. 5, p. e0124404, 2015.
[75] K. Pickard, M. J. Davidson, S. Kim, and D. Alais, "Incongruent active head rotations increase visual motion detection thresholds," *Neuroscience of Consciousness*, vol. 2024, no. 1, p. niae019, 2024.
[76] I. Bouziane, M. Das, K. J. Friston, C. Caballero-Gaudes, and D. Ray, "Enhanced top-down sensorimotor processing in somatic anxiety," *Translational Psychiatry*, vol. 12, no. 1, p. 295, 2022.
[77] L. Mikula, S. Mejía-Romero, R. Chaumillon, A. Patoine, E. Lugo, D. Bernardin, and J. Faubert, "Eye-head coordination and dynamic visual scanning as indicators of visuo-cognitive demands in driving simulator," *PLoS ONE*, vol. 15, no. 12, p. e0240201, 2020.
[78] L. Piccardi, G. Iaria, M. Ricci, F. Bianchini, L. Zompanti, and C. Guariglia, "Walking in the corsi test: which type of memory do you need?" *Neuroscience letters*, vol. 432, no. 2, pp. 127–131, 2008.
[79] M. Zibandehpoor and M. Delrobaei, "MoWalCT Dataset: Head-Trunk Motion and Walking Trajectories During the Walking Corsi Task," 8 2026,. [Online]. Available: https://figshare.com/articles/dataset/MoWalCT_Dataset_Head-Trunk_Motion_and_Walking_Trajectories_During_the_Walking_Corsi_Task/33152837



MOBINA ZIBANDEHPOOR received the B.Sc. degree in Electrical Engineering and the M.Sc. degree in Mechatronics Engineering from K. N. Toosi University of Technology, Tehran, Iran, in 2021 and 2025, respectively. She is currently a member of the Mechatronics Laboratory at the Faculty of Electrical Engineering at K. N. Toosi University of Technology. Her research focuses on integrating pose estimation, eye-tracking, and behavioral data to study cognitive processes. Her interests lie in applying advanced technologies to human-centered systems, aiming to bridge engineering with cognitive and affective science.



MEHDI DELROBAEI (Senior Member, IEEE) received his Ph.D. degree in electrical engineering from Western University, London, ON, Canada, in 2010. From 2011 to 2016, he completed his postdoctoral training at Western University and the Lawson Health Research Institute, London, ON. He is currently an Associate Professor of mechatronics and the Director of the Mechatronics Laboratory at the Faculty of Electrical Engineering at K. N. Toosi University of Technology. His research interests include biomechatronic systems, remote patient monitoring, cognitive assessment, and wearable assistive technologies.