

Tactile-Based Human Intent Recognition for Robot Assistive Navigation

Shaoting Peng^{1*}, Dakarai Crowder², Wenzhen Yuan², Katherine Driggs-Campbell¹

Abstract—Robot assistive navigation (RAN) is critical for enhancing the mobility and independence of the growing population of mobility-impaired individuals. However, existing systems often rely on interfaces that fail to replicate the intuitive and efficient physical communication observed between a person and a human caregiver, limiting their effectiveness. In this paper, we introduce Tac-Nav, a RAN system that leverages a cylindrical tactile skin mounted on a Stretch 3 mobile manipulator to provide a more natural and efficient interface for human navigational intent recognition. To robustly classify the tactile data, we developed the Cylindrical Kernel Support Vector Machine (CK-SVM), an algorithm that explicitly models the sensor’s cylindrical geometry and is consequently robust to the natural rotational shifts present in a user’s grasp. Comprehensive experiments were conducted to demonstrate the effectiveness of our classification algorithm and the overall system. Results show that CK-SVM achieved superior classification accuracy on both simulated (97.1%) and real-world (90.8%) datasets compared to four baseline models. Furthermore, a pilot study confirmed that users more preferred the Tac-Nav tactile interface over conventional joystick and voice-based controls. Code and video are available at: <https://sites.google.com/view/tac-nav/home>.

I. INTRODUCTION

Robot assistive navigation (RAN), in which a robot provides physical support while individuals move from one location to another, plays a crucial role for people with mobility impairments [1]. In the U.S., 12.2% of adults live with a mobility disability, and the aging population suggests an increasing need for navigation assistance [2], [3]. To address mobility challenges and increase the independence of these populations, researchers have been focusing on developing assistive robotic systems that can help navigation in both indoor and outdoor environments, such as augmented canes [4], [5], wheelchairs [1], [6] and walkers [7], [8]. Compared to a cane or wheelchair, walker devices provide adequate physical support while maintaining a degree of user independence, and are also helpful for building endurance and regaining muscle strength, especially during recovery from surgery, injury, or illness [9], [10]. In this work, we explore interaction mechanisms for a robotic walker, which often takes the form of a mobile robot with a handle [11], for assisted navigation.

In the development of effective RAN, a key challenge is human intent recognition, which is fundamental to ensuring the safety and fluidity of the human-robot interaction. The ability to accurately and efficiently detect or input a user’s movement intent is highly reliant on the sensing and user



Fig. 1. **Intuition of our work:** Mobility-impaired people convey their moving intents by applying specific grasping force patterns on the arm of caregiver. We adapt this intuitive and natural way of interaction to the robot assistive navigation by putting a tactile skin on the arm of a Stretch robot.

interface. Existing interfaces range from simple button arrays and joysticks [12] to more complex biosignal sensors like electromyography (EMG) or electroencephalography (EEG) [6]. However, each approach presents trade-offs. A button array, while intuitive, may divert the user’s attention from their surroundings, whereas an EEG sensor, while avoiding physical distraction, can be prone to signal noise and require extensive calibration. Our goal in this work is to mitigate the intent recognition gap by designing an algorithm and system that is natural, accurate, and robust, fostering a human-centric interaction where the user feels they are collaborating with a partner rather than merely operating a machine.

We are inspired both by cane and walker design as well as how caregivers assist individuals with mobility challenges. We observe that information is conveyed through physical contact, in that users often apply specific grasping force patterns to a caregiver’s arm to convey their intentions: a slight twist to indicate a turn or a pull to signal a stop (Fig. 1 left). To replicate this natural and intuitive dynamic, we propose a novel interface that uses a tactile sensor patch that functions like robotic skin, enabling the robot to “feel” the user’s nuanced touch gestures. This sensor is further integrated onto a cylindrical handle mounted on a Stretch 3 mobile manipulator (Fig. 1 right). While a natural design for grasping the robot, determining human intent for this form factor proves challenging for traditional algorithms. To effectively classify the tactile data for human mobility intent recognition, we developed a novel *Cylindrical Kernel Support Vector Machine (CK-SVM)*. This algorithm is designed to leverage the sensor’s cylindrical geometry by enhancing the standard Radial Basis Function Support Vector

* Corresponding author: peng33@illinois.edu

¹Electrical & Computer Engineering, University of Illinois Urbana-Champaign, IL, 61801, USA

²Siebel School of Computing and Data Science, University of Illinois Urbana-Champaign, IL, 61801, USA

Machine (RBF-SVM) to be robust against slight rotational shifts in the user’s grasp, thereby increasing classification accuracy. Finally, we integrated the sensor and CK-SVM into a complete RAN system named *Tac-Nav*. We performed a comprehensive evaluation, benchmarking the CK-SVM against standard classifiers such as RBF-SVM and Convolutional Neural Network (CNN) using both simulated and real-world tactile data. Furthermore, we conducted a pilot study with 5 subjects to compare the performance and user experience of *Tac-Nav* against conventional joystick-based and voice-controlled RAN systems. The results demonstrate the superior classification accuracy of CK-SVM and highlight the benefits of *Tac-Nav* in providing a more effective and user-preferred navigation experience.

We present the following contributions:

- A novel Cylindrical Kernel SVM (CK-SVM) designed for our tactile sensor, which improves its robustness to rotational shifts in the user’s grasp.
- The *Tac-Nav* system, which integrates our sensor and algorithms on a mobile manipulator for intuitive, tactile-based assistive navigation.
- A comprehensive evaluation of CK-SVM against other classifiers using real and simulated data, and a pilot study comparing *Tac-Nav* to two other RAN systems.

II. RELATED WORK

A. Robot Assistive Navigation

Robot assistive navigation (RAN) systems can be broadly categorized based on their target users: those designed for individuals with visual impairments and those for individuals with mobility impairments. For the visually impaired, the primary goal of a RAN system is to facilitate wayfinding to a target destination while ensuring obstacle avoidance [13]. For example, Balatti et al. employed a Franka manipulator on a mobile base with an impedance controller to achieve adaptive pulling [14]. The DRAGON system mounted a handle on a TurtleBot and implemented a dialogue system to enable navigation through natural language commands [11], [15]. Other research has focused on augmenting the traditional white cane with advanced sensing and feedback capabilities [16], [17], [18]. For individuals with mobility impairments, RAN systems are primarily aimed at providing physical support and reducing the effort required to perform tasks such as maintaining balance [19]. The “NavChair” was one of the first shared-control wheelchair systems [20], and its development was followed by numerous other smart wheelchair projects [21], [22]. Other prominent forms of RAN include robotic walkers [23] and exoskeletons [24]. Research in RAN often explores various human input modes (e.g., joystick [25], brain-machine interface (BMI) [26]), feedback modalities (e.g., audio [11], haptics [27]), and sensor integrations (e.g., camera, LiDAR, IMU [28]). Notably, tactile sensing¹ has been underexplored. While prior work

¹It is worthwhile distinguishing between *tactile input / tactile sensor*, through which the user conveys intent to the system via physical contact, and *vibrotactile output / tactile actuator*, by which the system delivers feedback or alerts to the user via vibration.

such as [8] has used simple pressure sensors for singular tasks like fall detection, they lack the data richness required for complex command recognition. Similarly, force-torque sensors [29], [30] often require rigid, embodiment-specific mechanical integration that limits their versatility across different robots. In contrast, our soft tactile skin is universal and can be applied to diverse surfaces such as robot arms, custom handles, or even humanoid frames without extensive hardware modification, leaving its full potential for assistive navigation yet to be explored.

B. Assistive Human Intent Recognition

Accurate and efficient recognition of human intent is crucial for any assistive robot. Various algorithms have been proposed for different sensor types, yet limitations persist. The DRAGON system, for instance, maps voice commands to landmarks using a fine-tuned CLIP model, but its speech-to-text performance can degrade in noisy environments [11]. Jain et al. proposed a recursive Bayesian filtering approach using non-verbal observations, but such probabilistic modeling can impose restrictive assumptions on features and environments [31]. Shen et al. placed three resistive force sensors on a lower extremity assistive device and used Principal Component Analysis (PCA) followed by an SVM to classify motion intent, though this required specialized and expensive hardware [32]. Ranganeni et al. used a one-dimensional torque sensor to detect a user’s turning intent, but the sensor could not provide richer contextual information [29]. Alternatively, EEG sensors can provide rich brain signals. For example, Choi et al. used sample covariance matrices from EEG signals to classify grasping intent, and Perrin et al. used EEG to confirm or reject proposed actions [6], [33]. However, EEG signals are notoriously noisy, particularly when the user is in motion. Other work has explored deep learning models with multi-modal sensor inputs to learn underlying human intent [34], [35]. While these methods can achieve superior performance, their lack of interpretability can be a significant concern in safety-critical assistive applications. In this work, we achieve robust human intent recognition using a classical SVM enhanced with a novel, interpretable kernel, applied to high-dimensional tactile data.

C. Tactile Gesture Classification

To enable robots to understand human intent via touch, researchers typically employ gesture-classification methods, most often using a pressure-sensing grid as the sensing apparatus. Traditional approaches have frequently relied on classical machine learning models like boosting, random forests, and SVMs [36], [37]. For instance, the CoST social gesture dataset was classified using SVMs and Bayesian classifiers on extracted features such as mean pressure and gesture duration [38]. Similarly, other work has used random forests to classify different types of infra hug gestures [39]. While these methods are still in use, there is a growing trend toward using deep learning to classify touch gestures [40], [41], [42]. Some of these advanced methods enhance recognition by

extracting spatiotemporal features from wavelet coefficients [43], and another novel approach used microphone arrays as a form of robotic skin, feeding spectrograms into a convolutional neural network [44]. Although numerous methods have been developed, they are often highly specialized for the specific gestures and sensor modalities for which they were designed.

III. METHOD AND SYSTEM

This section details our methodology, beginning with CK-SVM, our main algorithmic contribution. We then describe the implementation of the complete Tac-Nav system, covering the sensor design and hardware configuration in Section III-B.1. This is followed by a description of the software pipeline in Section III-B.2, which includes data collection, model training, and real-time control logic.

A. Cylindrical Kernel SVM

1) *Preliminary – Support vector machine with radial basis function kernel (RBF-SVM)*: The Support Vector Machine (SVM) is a powerful supervised learning algorithm, and its effectiveness can be extended to non-linearly separable datasets through the use of the kernel trick [45]. A popular and effective choice for this is the Radial Basis Function (RBF) kernel, which implicitly maps input features into a high-dimensional space where a linear separation becomes feasible. This kernel function measures the similarity between two feature points x_1 and x_2 using a Gaussian function:

$$K(x_1, x_2) = \exp(-\gamma \cdot \|x_1 - x_2\|^2) \quad (1)$$

In this equation, $\|x_1 - x_2\|^2$ represents the squared Euclidean distance between the two feature vectors. The hyperparameter γ acts as the length-scale, controlling the “spread” or influence of the kernel.

With the kernel defined, the SVM decision function for classifying a new sample feature x is formulated as:

$$y = \text{sign} \left(\sum_{i=1}^N (\alpha_i \cdot y_i \cdot K(x, x_i)) + b \right) \quad (2)$$

where y is the predicted label, N is the number of support vectors, α_i are the learned Lagrange multipliers, y_i are the corresponding class labels of the support vectors, and b is the bias term. This function computes a weighted sum of the kernel-based similarities between the input sample and the support vectors to make a final classification. In the context of this work, it maps complex tactile features to their corresponding labels: “turn left”, “turn right”, “speed up”, “stop”, and “neutral”.

2) *Cylindrical Kernel SVM (CK-SVM)*: Different human intents correspond to distinct grasping position and local patterns across the tactile handle. However, because we do not restrict human holding gesture and there is grasping position inconsistency across users and trials, these patterns are subject to rotational shifts along the cylindrical perimeter of the handle, even when the intended action is the same. If a standard SVM classifier with an RBF kernel is used on user data from the cylindrical form factor of our tactile sensor

(as shown in Fig. 2 and explained in Section III-B.1), intent recognition is not reliable (accuracy of approximately 0.37 for five intent classes). An analysis of this result revealed the following key observation:

The standard kernel fails because its “flattening” operation destroys the tactile sensor’s cylindrical topology, which causes it to misinterpret the slight rotational shifts in a user’s grasp as large, unrelated data changes.

To be interpreted correctly, this tactile feature should be viewed on its native 3D cylindrical manifold. A standard RBF kernel, however, requires the feature to be “flattened” into a one-dimensional vector. This operation destroys the essential 3D topology, causing a minor rotational shift, particularly the one that “wraps around” from the last row of the sensor to the first, to make two similar grasp patterns appear drastically different in their vector representations. Although the change is insignificant on the sensor’s 3D surface, the large Euclidean distance in the flattened space leads the RBF kernel to misinterpret the patterns as highly dissimilar, resulting in poor classification performance.

To address the misinterpretation of rotational shifts, we introduce the **Cylindrical Kernel**, a modification of the standard RBF kernel designed to respect the sensor’s cylindrical topology. Our kernel retains the Gaussian form but replaces the Euclidean distance with a custom **Cylindrical Distance** metric, $d_C(x_1, x_2)$:

$$K(x_1, x_2) = \exp(-\gamma \cdot d_C(x_1, x_2)) \quad (3)$$

The Cylindrical Distance finds the minimum cost to align two tactile feature maps, x_1 and x_2 , by rotationally shifting one relative to the other:

$$d_C(x_1, x_2) = \min_{s \in \{0, \dots, k-1\}} (\|x_1 - \tau_s(x_2)\|_F^2 + \Lambda(s)) \quad (4)$$

This metric consists of two key components:

- 1) $\|x_1 - \tau_s(x_2)\|_F^2$: The squared Frobenius distance between the first feature map and the second map after it has been rotationally shifted by s rows.
- 2) $\Lambda(s)$: A penalty function that discourages large shifts.

The shift operation and penalty function are defined as:

$$[\tau_s(x)]_{i,j} \equiv x'_{i,j} = x_{((i-s) \bmod k),j} \quad (5)$$

$$\Lambda(s) = \exp(\min(s, k-s)/\delta) - 1 \quad (6)$$

Equation 5 details the circular shift for each element at row i and column j of the feature map, where k is the total number of rows on the sensor. Equation 6 defines the exponential penalty for a given shift s , where the $\min(s, k-s)$ term measures the shortest rotational distance around the cylinder. This penalty ensures that two patterns are considered different even if they match perfectly after a large rotational shift, preventing misclassifications between similar grasp patterns that are far apart on the handle. δ is a tunable hyperparameter.

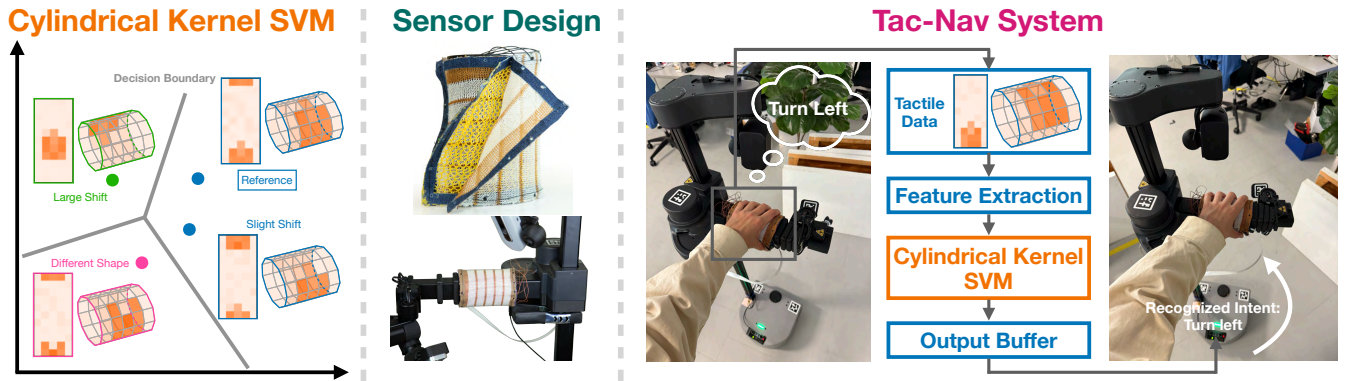


Fig. 2. **Left: CK-SVM illustration.** The CK-SVM classifies tactile patterns based on their similarities. Given a **reference** tactile data, **patterns with slight shifts** are classified to be the same class, while either **patterns with large shifts** or **data with different patterns** are classified to be different classes. Draw in 2D for visualization simplicity. **Middle: Sensor Design.** The three-layer knitted tactile sensor and its mounting on the arm of the Stretch 3 robot. **Right: Tac-Nav system overview.** For a given intent such as “turn left”, the system reads tactile data, extracts features that preserve the cylindrical topology, classifies the features using the CK-SVM, and sends the final recognized intent to the robot after filtering by an output buffer.

By explicitly capturing the cylindrical geometry of the tactile sensor, our proposed CK-SVM is theoretically poised to increase classification accuracy. We validate this hypothesis through a comprehensive evaluation of CK-SVM on both simulated and real-world data, presented in Section IV.

B. Tac-Nav System Implementation

1) *Hardware Configuration:* To obtain user tactile data, a machine knitted resistive-based sensor was created. Using machine knitting allows for quick and cheap sensor fabrication allowing for a faster sensor iteration cycle. Since people touch fabric every day, creating a sensor using a similar material makes the sensor more comfortable and familiar.

Our sensor adopts the three-layer architecture proposed by Si et al. [46], comprising two conductive layers patterned with orthogonal stripes and a central non-conductive mesh. The resulting sensing array spans a 11×5 grid, offers a spatial resolution of approximately 15.5 mm^2 , and operates at a sampling rate of about 50 Hz. The sensor has a height of 15 cm and a diameter of 7.4 cm. The inner mesh was created with A and E nylon yarn and the insulating stripes between each conductive strip is made of rayon yarn. The conductive stripes are composed of stainless steel infused yarn which had a resistance of around $2.5 \text{ M}\Omega$. A NUCLEO-F030R8 board was used.

Then, this tactile sensor array is mounted on a 3D-printed cylindrical holder. This assembly is affixed to the horizontal arm of a Hello Robot Stretch 3, as shown in Figure 2. Mounting the sensor on the adjustable arm, rather than the main vertical body, offers the ergonomic benefit of allowing users to easily set the handle to their preferred height. To ensure a clean and unobtrusive setup, the control board is housed underneath the arm, with wiring routed along the arm to the main body.

2) *System Pipeline:* Our system is designed to recognize a set of five human intents: {“turn left”, “turn right”, “speed up”, “stop”, “neutral”}. The software pipeline, from data collection to real-time control, is as follows:

a) *Data Collection and Training:* To minimize user burden, we collected a concise dataset for each participant,

consisting of 40 one-second samples for each of the five intents. The users are encouraged to use their natural holding pose for each intent instead of a prefixed grasping pose to account for distinct user preferences. By fixing the sampling frequency to be 45 Hz, the preliminary dataset for each subject has a shape of $(200, 45, 11, 5)$. From each one-second sample, we extract features including the mean, max, standard deviation, and spatial gradient of the tactile pressures, which collapses the time-series data into four 2D feature maps that preserve the sensor’s cylindrical topology, resulting in a final dataset for each subject of shape $(200, 4, 11, 5)$. Then we shuffle the dataset and split 80% to be the training set, and the rest for test set. Our CK-SVM model is then trained on the training set, with a total training time of less than four seconds per subject.

b) *Real-Time Classification:* During operation, the system uses a sliding window approach for real-time intent classification. A one-second window, consistent with the training data, slides in 100-millisecond increments. To enhance system robustness and prevent reactions to occasional misclassifications, we implement an *output buffer* that stores the latest seven predictions. A new command is sent to the robot only when all seven intent classification results in the buffer are identical; otherwise, a “neutral” command is issued, which either stops turning or maintains the current moving velocity.

c) *Robot Control:* We use ROS2 to communicate the final classified intent to the Stretch 3 robot. For robot motion, the turning angular speed is set to 0.15 radians per second, and each “speed up” command increases the robot’s forward velocity by 0.01 meter per second, up to a maximum of 0.15 meter per second. The “stop” command brings the robot to an immediate halt.

IV. INTENT CLASSIFICATION EVALUATION

To evaluate the performance of our proposed CK-SVM, we compare it against the following four baseline classifiers:

1) **Standard SVM with an RBF kernel.** The RBF-SVM introduced in Section III-A.1 is chosen to provide a di-

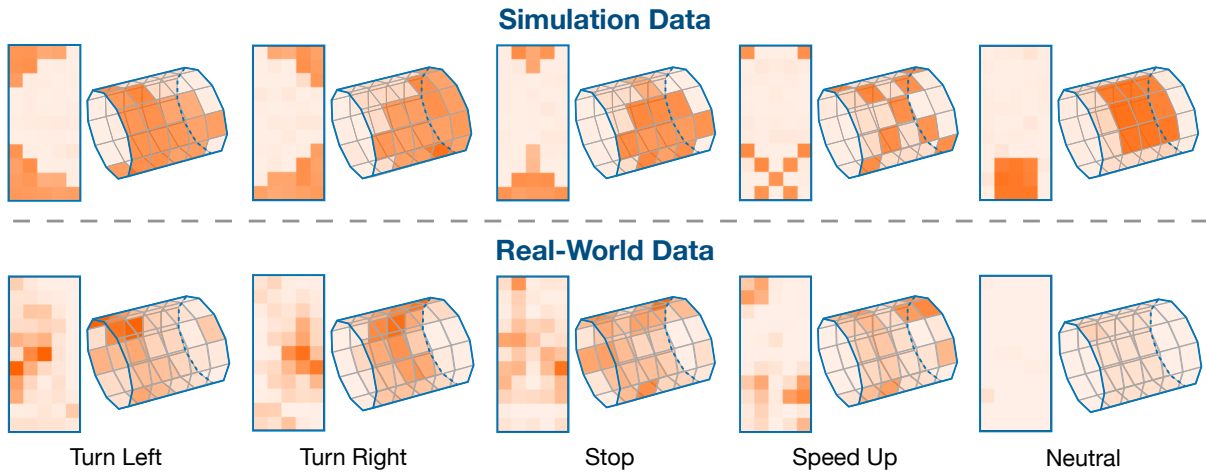


Fig. 3. **Tactile data visualization for all intents.** Top: designed simulated base patterns. Bottom: mean patterns across trials from one user.

TABLE I

Classification accuracies (mean \pm std) of four baseline methods and CK-SVM on simulated and real-world data.

	RBF-SVM	MLP	MDCM	CNN	CK-SVM
Simulated Dataset	0.7708 \pm 0.0579	0.7375 \pm 0.0818	0.7417 \pm 0.0847	0.8125 \pm 0.0209	0.9708 \pm 0.0188
Real-World Dataset	0.3708 \pm 0.0219	0.3807 \pm 0.0113	0.5683 \pm 0.1250	0.7599 \pm 0.0782	0.9081 \pm 0.0198

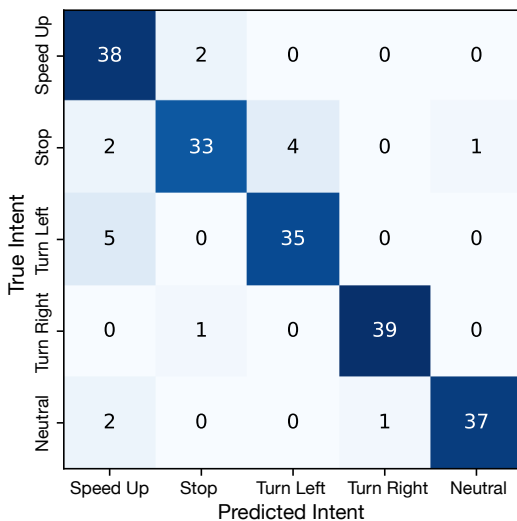


Fig. 4. Aggregated confusion matrix for CK-SVM on real-world dataset, summing the individual ones from all subjects.

rect comparison with our Cylindrical Kernel, isolating the benefit of our geometry-aware design.

- 2) **Multilayer Perceptron (MLP).** The MLP is selected as a powerful, general-purpose non-linear classifier.
- 3) **Minimum Distance to Covariance Mean (MDCM).** MDCM is a common method for biosignal classification, making it a good candidate because tactile sensor data can be viewed as a multi-channel time-series signal. Specifically, for each class k , a sample covariance matrix centroid M^k is calculated from the training data. For a new data sample X , the predicted class \hat{k} is given by the nearest centroid: $\hat{k} = \arg \min_k d(X, M^k)$, where the distance metric d is the geodesic distance between two sample covariance matrices on their Riemannian manifold. Please refer to [47] for more details.

- 4) **Convolutional Neural Network (CNN).** The CNN is chosen as a strong baseline due to its inherent ability to learn local spatial patterns directly from the 2D tactile data. However, it lacks interpretability and typically requires a larger dataset to train effectively and avoid overfitting, which may increase human burden.

We evaluate the overall classification accuracies of all methods on both simulated dataset and real-world dataset. Based on the later dataset we also provide the confusion matrix to gain deeper insights into our models.

A. Simulated Data Results

To validate our method in a controlled setting, we first conducted a proof-of-concept experiment on simulated data. We generated a synthetic dataset by creating five distinct local patterns, each corresponding to one of the five intents, shown as the top row of Fig. 3. The final dataset was produced by applying random rotational shifts and adding Gaussian noise. The performance of the baseline methods and our CK-SVM on this dataset is presented in the first row of Table I. While all classifiers achieved respectable accuracies above 0.73, our CK-SVM demonstrated superior performance, attaining a classification accuracy of 0.9708 with the lowest standard deviation of 0.0188. This result highlights its robustness to the specific geometric challenges it was designed to overcome.

B. Real-World Data Results

We next evaluated the classifiers on real-world data collected from five human subjects, following the protocol described in Section III-B.2.a. To provide an intuitive understanding of this data, the bottom row in Fig. 3 visualizes the mean tactile activation maps for each intent, averaged across all trials from one user. The visualizations reveal intuitive

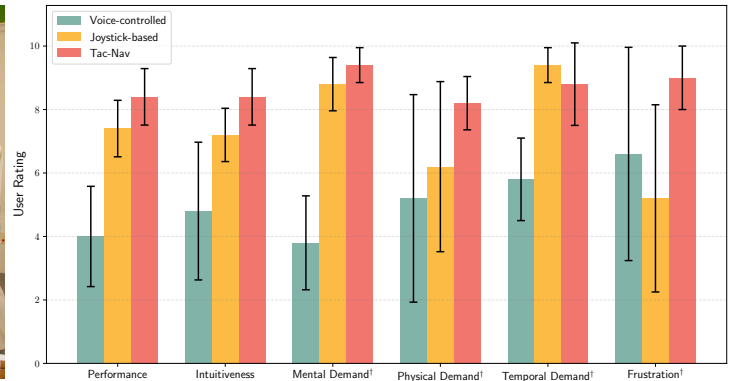
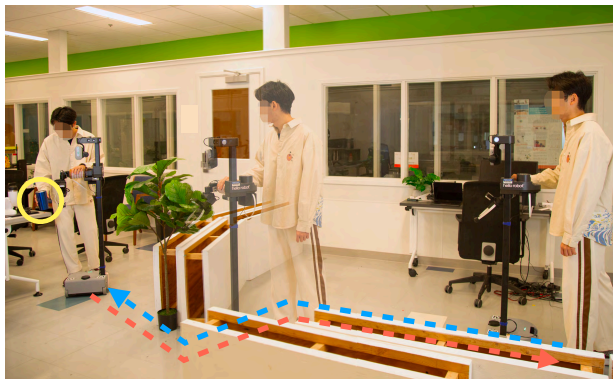


Fig. 5. **Left: Experiment setup.** The user starts from right, follows the blue arrow to pick up the water bottle on the left, and returns by following the red arrow. **Right: Pilot study results.** To unify the visual representation of our metrics, “mental demand”, “physical demand”, “temporal demand”, and “frustration” were inverted, where a higher score indicates a more positive outcome (e.g., easy to use).

patterns; for instance, the left side of the sensor shows higher activation for the “turn left” intent, while the right side is more active for “turn right”. However, as expected, these real-world patterns are substantially more complex and less distinct than those in the simulated dataset. The quantitative results, summarized in Table I, show a general performance decrease compared to the simulated data due to this increased complexity. The RBF-SVM, MLP, and MDCM models struggled, achieving accuracies below 0.6. While the CNN performed better with an accuracy over 0.75, its large standard deviation indicates inconsistent performance across subjects. In contrast, our CK-SVM achieved the highest classification accuracy of over 0.9 with a low standard deviation of 0.02, demonstrating both the superior performance of our cylindrical kernel and its robustness across different users. To gain some deeper insights of the CK-SVM’s performance, we summed over the classification results and the corresponding true labels for all users, and provided the aggregated confusion matrix as shown in Fig. 4. This matrix reveals that the “speed up” intent has a noticeably low precision of approximately 0.81. We attribute this to the physical nature of the interaction, as the user’s palm remains in constant contact with the sensor; this means that even during a turning or “stop” action, some forward pressure is often present, leading to a higher rate of false positives for the “speed up” class.

V. REAL-WORLD NAVIGATION EXPERIMENTS

Beyond offline classification accuracy, we evaluated the complete Tac-Nav system by conducting a pilot study with five subjects. This approach allowed us to gain a deeper understanding of the system’s real-world performance and usability by collecting a range of user experience metrics.

1) *Experimental Setup:* The experiment for each user was conducted in two phases, illustrated as the left image of Fig. 5. Before the experiments, participants were asked to imagine they had limited mobility, so they would rely on the Stretch 3 robot as the form of a cane or a walker, similar to an older adult requiring such walking balance assistance. After the initial data collection and model training period as introduced in Section III-B.2.a, users are required to

complete a navigation task mimicking a common in-home scenario: navigating to a location to retrieve a cup of water and returning. In our setup, the users start from the right side of the image, follow the right-to-left blue arrow to reach the table on the left, get the blue water bottle circled in yellow, and follow the left-to-right red arrow to get back to the starting place. The path was designed to include a narrow passage that necessitated both left, right turns and stop. For the return journey, users would have to hold the water bottle while walking and maintain balance.

2) *Comparisons & Evaluation Metrics:* To evaluate the performance of our Tac-Nav system, we designed two alternative control systems for comparison: one based on natural language voice commands and another using a joystick.

For the voice-controlled system, we use python *SpeechRecognition* as voice-to-text model. A “turn left” or “turn right” instruction causes the robot to stop and turn 15 degrees in the specified direction. The “speed up” command increases the robot’s velocity by 0.07 meter per second, with a maximum speed of 0.14 meter per second. The “stop” command brings it to a complete halt. For the joystick-based system, the controller was mounted on the robot’s shoulder for easy access with the user’s left hand, allowing their right hand to hold the robot’s arm for balance. The control logic was standard: pushing forward commanded forward motion at a velocity of 0.1 meter per second, while left and right inputs produced an angular velocity of around 30 degree per second. Pushing backward was disabled, and releasing the joystick stopped the robot.

User experience with all three systems was evaluated using a questionnaire based on the NASA Task Load Index (TLX) [48], which assesses “mental demand”: How mentally demanding the task was, “physical demand”: how physically demanding the task was, “temporal demand”: how hurried or rushed the pace of the task was, “performance”: how successful the user was in accomplishing what she was asked to do, and “frustration”: how insecure, discouraged, irritated, stressed, and annoyed the user was. We augmented this with an “intuitiveness” metric to measure how natural and intuitive the user perceived each interface for controlling the robot. After using all three systems, participants provided 10-point

scale ratings for each metric, followed by a short interview to gather qualitative feedback.

3) *Pilot Study Results*: The pilot study results are presented in Fig. 5. The voice-controlled system, joystick-based system, and our Tac-Nav system are colored in green, yellow, and red, respectively. The error bars represent the standard variance across all participants. The ‘†’ sign after “mental demand”, “physical demand”, “frustration”, and “temporal demand” indicates the scores are inverted for a unified metric visual representation, where a higher score indicates a higher user preference (i.e., low demand or frustration).

The results show a strong user preference for our Tac-Nav system. It outperformed both the joystick-based and voice-controlled systems in five out of the six metrics, exhibiting the highest average scores and the lowest variance across subjects, which demonstrates its superior overall performance and consistent user experience. Notably, for “physical demand” and “frustration”, Tac-Nav scored more than two points better than the alternatives. Also, the large standard variances of joystick-based system reflect a split in user experience: while two participants who successfully complete the task found the task to be manageable, three users commented that using the joystick while holding the cup of water was physically demanding and frustrating, especially when water was occasionally spilled. In contrast, our Tac-Nav system provided a more natural interaction, enabling users to simultaneously control the robot’s moving direction and maintain balance with one hand while carrying the water bottle with the other. The only metric where the joystick performed slightly better was “temporal demand” (by 0.5 points). This is attributable to the joystick’s direct real-time control, whereas the output buffer in Tac-Nav introduces a minimum delay of 0.2 seconds, with occasional classification noise at the start of a movement increasing the stabilization time to nearly one second. However, all five subjects stated this delay was acceptable, especially given the significant improvement in reliability that the buffer provided. Regarding the voice-based controller, four out of five users found the speech recognition to be imperfect and noted that issuing frequent verbal commands, particularly for large turns, could be tedious.

VI. CONCLUSION AND DISCUSSION

A. Conclusion

Inspired by the way caregivers interpret physical grasping cues from those they assist, this work presents Tac-Nav, a robot assistive navigation system that uses a tactile skin interface for intuitive human intent recognition. To address the challenge of natural variations in a user’s grasp, we developed the Cylindrical Kernel Support Vector Machine (CK-SVM), an algorithm specifically designed to be robust to rotational shifts by explicitly modeling the sensor’s cylindrical topology. Integrated into a real-time system with a filtering output buffer, the CK-SVM enables Tac-Nav to provide fluid and robust mobility support based on the user’s touch. Intent classification evaluation results demonstrate the superiority of our approach, with the CK-SVM achieving

highest classification accuracy than baseline models on both simulated and real-world data. Furthermore, a real-world navigational pilot study confirmed the high usability and user preference for Tac-Nav over conventional voice-controlled and joystick-based systems, validating the effectiveness of our tactile-based system.

B. Limitations and Future Work

While this work demonstrates the benefits of our tactile-based robot assistive navigation system, we acknowledge several limitations that inform our future work. First, our current CK-SVM is specifically designed for cylinder-shaped sensors, which we will explore generalizing our kernel design to other common handle geometries. Besides, our findings on user preference are preliminary based on a small pilot study. In the future we will expand this into a comprehensive clinical trial with mobility-impaired individuals to validate the system’s long-term usability. Finally, the rich data from the tactile skin is currently only used for mobility intent recognition of five classes; we intend to leverage its full potential to enable new functionalities, such as inferring the user’s physical state (e.g., fatigue or confidence) from their grip forces and patterns, or enhancing safety by detecting unstable grips indicative of a potential fall.

ACKNOWLEDGMENT

This work was supported by the National Science Foundation under Grant CCF 2236484. We would like to thank Haonan Chen and Kaiwen Hong for their valuable feedback during proofreading.

REFERENCES

- [1] D. Bell, J. Borenstein, S. Levine, Y. Koren, and J. Jaros, “An assistive navigation system for wheelchairs based upon mobile robot obstacle avoidance,” in *Proceedings of the 1994 IEEE International Conference on Robotics and Automation*, 1994, pp. 2018–2022 vol.3.
- [2] Centers for Disease Control and Prevention, “Disability impacts all of us,” <https://www.cdc.gov/disability-and-health/media/pdfs/disability-impacts-all-of-us-infographic.pdf>, July 2024.
- [3] V. A. Freedman, J. C. Cornman, and J. D. Kasper, “National health and aging trends study: Trends dashboards,” <https://www.nhats.org/researcher/trends-dashboards>, 2021.
- [4] R. Pyun, Y. Kim, P. Wespe, R. Gassert, and S. Schneller, “Advanced augmented white cane with obstacle height and distance feedback,” in *2013 IEEE 13th International Conference on Rehabilitation Robotics (ICORR)*, 2013, pp. 1–6.
- [5] P. Slade, A. Tambe, and M. J. Kochenderfer, “Multimodal sensing and intuitive steering assistance improve navigation and mobility for people with impaired vision,” *Science Robotics*, vol. 6, no. 59, p. eabg6594, 2021.
- [6] X. Perrin, R. Chavarriaga, F. Colas, R. Siegwart, and J. del R. Millán, “Brain-coupled interaction for semi-autonomous navigation of an assistive robot,” *Robotics and Autonomous Systems*, vol. 58, no. 12, pp. 1246–1255, 2010, intelligent Robotics and Neuroscience.
- [7] A. Morris, R. Donamukkala, A. Kapuria, A. Steinfeld, J. Matthews, J. Dunbar-Jacob, and S. Thrun, “A robotic walker that provides guidance,” in *2003 IEEE International Conference on Robotics and Automation (Cat. No.03CH37422)*, vol. 1, 2003, pp. 25–30 vol.1.
- [8] X. Zhao, Z. Zhu, M. Liu, C. Zhao, Y. Zhao, J. Pan, Z. Wang, and C. Wu, “A smart robotic walker with intelligent close-proximity interaction capabilities for elderly mobility safety,” *Frontiers in Neuroinformatics*, vol. Volume 14 - 2020, 2020.
- [9] M. Sehgal, S. Tsin, T. Kroll, T. Piro, and D. O. Lawson, “Mobility assistive device use in older adults,” *American Family Physician*, vol. 103, no. 12, pp. 737–744, jun 2021.

- [10] L. Resnik, S. Allen, L. Klinger, E. Necessary, and A. Sayers, "Perspectives on use of mobility aids in a diverse population of seniors: implications for intervention," *Disability and Health Journal*, vol. 2, no. 2, pp. 77–85, apr 2009.
- [11] S. Liu, A. Hasan, K. Hong, R. Wang, P. Chang, Z. Mizrachi, J. Lin, D. L. McPherson, W. A. Rogers, and K. Driggs-Campbell, "Dragon: A dialogue-based robot for assistive navigation with visual language grounding," *IEEE Robotics and Automation Letters*, vol. 9, no. 4, pp. 3712–3719, 2024.
- [12] S. Levine, D. Bell, L. Jaros, R. Simpson, Y. Koren, and J. Borenstein, "The navchair assistive wheelchair navigation system," *IEEE Transactions on Rehabilitation Engineering*, vol. 7, no. 4, pp. 443–451, 1999.
- [13] J. Pohovey, M. Lusardi, A. Hasan, S. Liu, A. Schreiber, S. A. Olatunji, W. A. Rogers, and K. Driggs-Campbell, "Beyond canes and guide dogs: A review of 40 years of robotics for wayfinding, navigating, and orienting assistance for people with visual impairments," August 2025.
- [14] P. Balatti, I. Ozdamar, D. Sirintuna, L. Fortini, M. Leonori, J. M. Gandarias, and A. Ajoudani, "Robot-assisted navigation for visually impaired through adaptive impedance and path planning," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*, 2024, pp. 2310–2316.
- [15] S. Liu, A. Hasan, K. Hong, C.-K. Yao, J. Lin, W. Liang, M. A. Bayles, W. A. Rogers, and K. Driggs-Campbell, "Designing a wayfinding robot for people with visual impairments," 2023.
- [16] G. A. Kebede and Y. K. Shiferaw, "Assistive smart cane technology for visually impaired peoples: A review," in *Artificial Intelligence and Digitalization for Sustainable Development*, B. H. Woldegiorgis, K. Mequanint, M. A. Bitew, T. B. Beza, and A. M. Yibre, Eds. Cham: Springer Nature Switzerland, 2023, pp. 196–208.
- [17] C. Mai, D. Xie, L. Zeng, Z. Li, Z. Li, Z. Qiao, Y. Qu, G. Liu, and L. Li, "Laser sensing and vision sensing smart blind cane: A review," *Sensors*, vol. 23, no. 2, 2023.
- [18] C.-E. Panazan and E.-H. Dulf, "Intelligent cane for assisting the visually impaired," *Technologies*, vol. 12, no. 6, 2024.
- [19] A. Nanavati, V. Ranganeni, and M. Cakmak, "Physically assistive robots: A systematic review of mobile and manipulator robots that physically assist people with disabilities," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 7, no. Volume 7, 2024, pp. 123–147, 2024.
- [20] S. Levine, D. Bell, L. Jaros, R. Simpson, Y. Koren, and J. Borenstein, "The navchair assistive wheelchair navigation system," *IEEE Transactions on Rehabilitation Engineering*, vol. 7, no. 4, pp. 443–451, 1999.
- [21] R. C. Simpson, "Smart wheelchairs: A literature review," *Journal of Rehabilitation Research and Development*, vol. 42, no. 4, pp. 423–436, jul 2005.
- [22] J. Leaman and H. M. La, "A comprehensive review of smart wheelchairs: Past, present, and future," *IEEE Transactions on Human-Machine Systems*, vol. 47, no. 4, pp. 486–499, 2017.
- [23] S. Page, L. Saint-Bauzel, P. Rumeau, and V. Pasqui, "Smart walkers: an application-oriented review," *Robotica*, vol. 35, no. 6, p. 1243–1262, 2017.
- [24] C. J. Siviý, M. A. G. G. Versteeg, A. D. D'Arpino, and C. J. Walsh, "Opportunities and challenges in the development of exoskeletons for locomotor assistance," *Nature Biomedical Engineering*, vol. 7, no. 4, pp. 456–472, 2023.
- [25] H. Kivrak, F. Cakmak, H. Kose, and S. Yavuz, "Social navigation framework for assistive robots in human inhabited unknown environments," *Engineering Science and Technology, an International Journal*, vol. 24, no. 2, pp. 284–298, 2021.
- [26] S. A. S. Bellary, D. Grabowsky, and J. M. Conrad, "Indoor navigation for assistive robots using eeg signals as feedback," in *2020 Southeast-Con*, 2020, pp. 1–8.
- [27] M. F. Jiménez, R. C. Mello, T. Bastos, and A. Frizera, "Assistive locomotion device with haptic feedback for guiding visually impaired people," *Medical Engineering & Physics*, vol. 80, pp. 18–25, 2020.
- [28] C.-L. Lu, Z.-Y. Liu, J.-T. Huang, C.-I. Huang, B.-H. Wang, Y. Chen, N.-H. Wu, H.-C. Wang, L. Giarré, and P.-Y. Kuo, "Assistive navigation using deep reinforcement learning guiding robot with UWB/Voice beacons and semantic feedbacks for blind and visually impaired people," *Frontiers in robotics and AI*, vol. 8, p. 654132, 2021.
- [29] V. Ranganeni, M. Sinclair, E. Ofek, A. Miller, J. Campbell, A. Kolobov, and E. Cutrell, "Exploring levels of control for a navigation assistant for blind travelers," in *Proceedings of the 2023 ACM/IEEE International Conference on Human-Robot Interaction*, ser. HRI '23. New York, NY, USA: Association for Computing Machinery, 2023, p. 4–12.
- [30] G. Suter, D. C. Guastella, F. Cancelliere, and G. Muscato, "Adaptive navigation of a smart walker with shared control," *Actuators*, vol. 14, no. 5, 2025.
- [31] S. Jain and B. Argall, "Probabilistic human intent recognition for shared autonomy in assistive robotics," *J. Hum.-Robot Interact.*, vol. 9, no. 1, Dec. 2019.
- [32] B. Shen, J. Li, F. Bai, and C.-M. Chew, "Motion intent recognition for control of a lower extremity assistive device (lead)," in *2013 IEEE International Conference on Mechatronics and Automation*, 2013, pp. 926–931.
- [33] H. J. Choi, S. Das, S. Peng, R. Bajcsy, and N. Figueroa, "On the feasibility of eeg-based motor intention detection for real-time robot assistive control," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*, 2024, pp. 5592–5599.
- [34] H. Sun, X. Gu, Y. Zhang, F. Sun, S. Zhang, D. Wang, and H. Yu, "An enhanced reset deep learning method for multimodal signal-based locomotion intention recognition," *Biomedical Signal Processing and Control*, vol. 101, p. 107254, 2025.
- [35] S. Liu, P. Chang, Z. Huang, N. Chakraborty, K. Hong, W. Liang, D. L. McPherson, J. Geng, and K. Driggs-Campbell, "Intention aware robot crowd navigation with attention-based interaction graph," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*, 2023, pp. 12 015–12 021.
- [36] Y. F. A. Gaus, T. Olugbade, A. Jan, R. Qin, J. Liu, F. Zhang, H. Meng, and N. Bianchi-Berthouze, "Social touch gesture recognition using random forest and boosting on distinct feature sets," in *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction*. New York, NY, USA: ACM, Nov. 2015.
- [37] A. Flagg and K. MacLean, "Affective touch gesture recognition for a furry zoomorphic machine," in *Proceedings of the 7th International Conference on Tangible, Embedded and Embodied Interaction*. New York, NY, USA: ACM, Feb. 2013.
- [38] M. M. Jung, R. Poppe, M. Poel, and D. K. J. Heylen, "Touching the void – introducing CoST," in *Proceedings of the 16th International Conference on Multimodal Interaction*. New York, NY, USA: ACM, Nov. 2014.
- [39] A. E. Block, H. Seifi, O. Hilliges, R. Gassert, and K. J. Kuchenbecker, "In the arms of a robot: Designing autonomous hugging robots with intra-hug gestures," *ACM Trans. Hum. Robot Interact.*, vol. 12, no. 2, pp. 1–49, Jun. 2023.
- [40] D. Jung, C. Gu, J. Park, and J. Cheong, "Touch gesture recognition-based physical human–robot interaction for collaborative tasks," *IEEE Transactions on Cognitive and Developmental Systems*, vol. 17, no. 2, pp. 421–435, 2025.
- [41] S. Albawi, O. Bayat, S. Al-Azawi, and O. N. Ucan, "Social touch gesture recognition using convolutional neural network," *Comput. Intell. Neurosci.*, vol. 2018, p. 6973103, Oct. 2018.
- [42] D. Crowder, K. Vandyck, X. Sun, J. McCann, and W. Yuan, "Social gesture recognition in spHRI: Leveraging fabric-based tactile sensing on humanoid robots," 2025.
- [43] Y.-K. Li, Q.-H. Meng, and H.-W. Zhang, "Touch gesture recognition using spatiotemporal fusion features," *IEEE Sensors Journal*, vol. 22, no. 1, pp. 428–437, 2022.
- [44] M. J. Yang, K. Park, W. D. Kim, and J. Kim, "Robotic skin mimicking human skin layer and pacinian corpuscle for social interaction," *IEEE ASME Trans. Mechatron.*, vol. 29, no. 4, pp. 2709–2719, Aug. 2024.
- [45] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, p. 273–297, Sep. 1995.
- [46] Z. Si, T. C. Yu, K. Morozov, J. McCann, and W. Yuan, "RobotSweater: Scalable, generalizable, and customizable Machine-Knitted tactile skins for robots," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, May 2023.
- [47] A. Barachant, S. Bonnet, M. Congedo, and C. Jutten, "Multiclass Brain-Computer Interface Classification by Riemannian Geometry," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 4, pp. 920–928, Mar. 2012.
- [48] S. G. Hart and L. E. Staveland, "Development of nasa-tlx (task load index): Results of empirical and theoretical research," in *Human Mental Workload*, ser. Advances in Psychology, P. A. Hancock and N. Meshkati, Eds. North-Holland, 1988, vol. 52, pp. 139–183.