Don't Park There! Learning Socially-Appropriate Robot Parking Spots in the Home

De'Aira Bryant
Amazon Lab126 Consumer Robotics
Sunnyvale, CA, USA
bryadeai@amazon.com

Apaar Sadhwani
Amazon Lab126 Consumer Robotics
Sunnyvale, CA, USA
apaar@amazon.com

Hanxiao Fu
Amazon Lab126 Consumer Robotics
Sunnyvale, CA, USA
hanxiaof@lab126.com

William D. Smart*
Oregon State University
Corvallis, OR, USA
bill.smart@oregonstate.edu

Dylan F. Glas
Amazon Lab126 Consumer Robotics
Sunnyvale, CA, USA
dgglas@amazon.com

Abstract

As autonomous social robots become more prevalent in home environments, they must decide where to position themselves within many different types of rooms or spaces, balancing accessibility with staying out of the way. This paper presents a machine learning approach to modeling user preferences for robot parking spots in the home using standard 2D occupancy maps. Our method learns spatial patterns from the information available in the occupancy maps and user-annotated floorplans without requiring specialized inputs. We evaluate the approach using floorplan data from 84 users who provided parking spot preferences after living with and evaluating a social robot in their homes for at least two weeks. Our method significantly outperforms a state-of-the-art baseline focused exclusively on avoiding walking paths. We demonstrate how the approach extends to additional map features and share insights about the types of preference patterns learned by the model. This contribution provides a framework that can incorporate new environmental inputs as robot perception capabilities evolve.

CCS Concepts

• **Computer systems organization** → **Robotic autonomy**; • **Computing methodologies** → **Robotic planning**; • **Human-centered computing** → **Empirical studies in HCI**.

Keywords

human-robot interaction, social robot navigation, spatial machine learning, home robots, robot parking, parking spot selection

ACM Reference Format:
De'Aira Bryant, Apaar Sadhwani, Hanxiao Fu, William D. Smart, and Dylan F. Glas. 2026. Don't Park There! Learning Socially-Appropriate Robot Parking Spots in the Home. In *Proceedings of the 21st ACM/IEEE International Conference on Human-Robot Interaction (HRI '26), March 16–19, 2026, Edinburgh, Scotland Uk.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3757279.3785648

*William D. Smart is also affiliated with Amazon Lab126, Sunnyvale, CA, USA

Figure 1: Left: The Astro home robot, Right: user-annotated parking preference map (yellow = good, red = bad).

1 Introduction

At the California Computer History Museum's "The First 2000 Years of Computing" exhibit, in the section on robotics and artificial intelligence, sits a 1984 home robot kit, Hero Jr., designed to "roam hallways, play games, act as an alarm clock, and even seek to remain near human companions" [25, 35]. This early vision recognized that social robots need the flexibility to position themselves wherever interaction might naturally occur. While the technology needed to deliver these user experiences more seamlessly wasn't yet ready [50], it is notable that most modern robots more closely follow the vision presented in Isaac Asimov's The Robots of Dawn (1983), in which robots waited inside designated wall niches, akin to the fixed charging docks of today's robots [2]. While a dock provides a utilitarian means to keep robots out of the way when not in use, we argue that social robots living with users in the home will require a dynamic approach to parking spot selection [30], in order to provide opportunities for spontaneous social interaction in a variety of rooms and spaces. This evolution in robot behavior presents a fundamental challenge: **selecting socially appropriate parking spots that balance accessibility and unobtrusiveness.**

Poor parking spot selection can create immediate safety risks like trip hazards or blocked pathways, while also affecting long-term robot acceptance and adoption. A robot that frequently obstructs movement, blocks access to furniture or appliances, or positions itself where it is difficult to find when needed may quickly become more of a nuisance than a help. These challenges are particularly acute in home environments where the functional and social use of space may be influenced by many factors (*e.g.*, the size of the home, its layout, pedestrian traffic patterns, etc.) and vary significantly between households [4, 9, 18].

HRI '26, Edinburgh, Scotland Uk

ACM ISBN 979-8-4007-2128-1/2026/03
https://doi.org/10.1145/3757279.3785648

Recent advances in sensing, perception, and computational modeling have made the vision of socially intelligent home robots feel closer to reality than ever before with several robots emerging in both academic research [12, 14, 27, 41] and commercial products [30] (Figure 1, left). However, while decades of research have advanced robots' ability to navigate through dynamic environments, this work has primarily focused on reaching known destinations. A related challenge that has been less explored is determining how to select the parking spot in the first place, which often requires an understanding of the social use of the space.

Glas and Smart argued that "the most fundamental factor in determining acceptability of a parking spot is whether it blocks any of the main walking paths in a space" and presented a method using geometry to infer primary pedestrian walking paths for identifying out-of-the-way parking spots [17]. While this heuristic approach provides a valuable foundation, the complexity of home environments suggests that additional spatial patterns and relationships may also influence social acceptability of parking spots [17].

To address these complexities, this paper presents a machine learning approach to model user preferences for robot parking spots using standard occupancy maps. Our key contributions include:

(1) **A flexible ML framework** for learning social acceptability of parking spots from standard robot occupancy maps;
(2) **A specialized loss function and sampling strategy** to balance overall parking spot quality estimation with region-level spot selection; and
(3) **Experimental validation of our approach** using data from 84 home robot users, revealing key patterns about human expectations for robot parking behavior.

2 Related Work

Understanding where a robot should park in a home environment intersects multiple areas of research including human-robot interaction [30], socially aware robot navigation [43], and machine learning from spatial data [37].

2.1 Social Robots in Home Environments

The potential for social robots to operate and interact with people in home environments has been extensively studied in HRI research, with applications ranging from household tasks [16], and educational assistance [6], to pet care [51], and companionship [33]. Recent studies have found that home robots engaging in proactive behaviors and socially-aware repositioning during idle time lead to higher levels of user engagement, acceptance, and prolonged use [11, 30, 32, 42]. This suggests that effective parking strategies are essential for maintaining long-term robot adoption. To achieve this, robots will need to develop both spatial and social awareness of their surroundings. While significant progress has been made in robot navigation and mapping, understanding the social use of space remains a key challenge for autonomous home robots [8, 50].

2.2 Social Navigation & Positioning

Prior work in social navigation has explored various aspects of robot movement and positioning in human environments. Researchers have investigated personal space preservation [21], human trajectory prediction [46], socially-aware approach paths [3], conversational positioning [5, 24, 38], and approach techniques for moving users [40]. Particularly relevant to our work are approaches that model patterns of human space utilization. Kanda et al. [26] developed spatiotemporal maps and clustering models of dynamic motion to understand aggregate patterns of human behavior, which were used to customize robot patrolling and approaching behavior. While these works provide valuable insights into human-relative positioning, they primarily focus on active interaction scenarios rather than proactive positioning without explicit instruction.

The challenge of selecting parking locations has received less attention, as many current home robots simply return to a charging dock after task completion [17, 47]. While prior work has addressed dock placement from an energy efficiency perspective [29], our work focuses on social factors rather than energy constraints.

Kitade et al. explored robot waiting positions in shopping malls using computational features like distance to walls and shops [28]. Glas and Smart developed a method for selecting parking spots based on a model of mutual passing along likely walking paths estimated from occupancy maps [17]. However, modern sensing capabilities can recognize more features in the environment, through segmentation of walls and clutter [34], detection of architectural features like doors and stairs [45], and semantic scene understanding capturing furniture, appliances, and other objects [44, 49]. We aim to use this rich environmental context to support machine learning for socially-aware parking.

2.3 Machine Learning with Spatial Data

Applying machine learning to spatial data presents unique challenges related to spatial dependence, scale, and feature engineering [37]. Tobler's First Law establishes that spatial correlation typically decreases with distance [36], suggesting that parking spot quality is most influenced by immediate surroundings rather than distant features. For example, whether a spot is positioned between two pieces of furniture likely has significant impact on its appropriateness for parking, while the presence of a bed in a distant room has little relevance. This principle has motivated patch-based sampling to emerge as a common practice for capturing these local spatial relationships, enabling models to focus on the most relevant contextual features across various spatial analysis domains [13, 23].

The effectiveness of patch-based methods depends heavily on appropriate patch size selection [37]. While satellite imagery applications often require large patches to capture meaningful patterns [23], medical imaging may use small patches for cell-level analysis [22]. In the context of robot parking, too large a patch size can introduce noise by capturing unrelated areas in the home. For example, a patch surrounding a parking spot in the living room might unnecessarily include a nearby bedroom that has no effect on the spot's appropriateness for parking. Thus, appropriate patch size is one question we investigate in this work.

Another key consideration in our application is **scale preservation**. Unlike many image recognition tasks where scale invariance is desirable, robot parking requires preserving precise spatial scale where each cell in a patch corresponds to specific real-world dimensions. To process this type of spatial data, convolutional neural

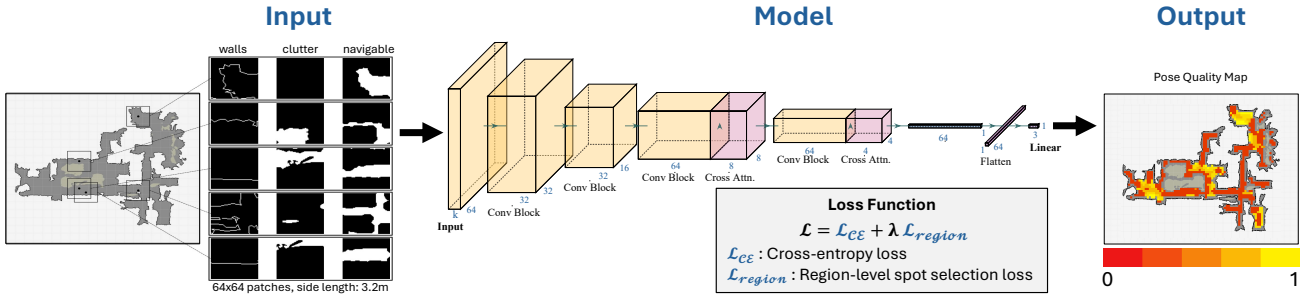

Figure 2: Learning robot parking preferences through feature patches from occupancy maps (left), a hybrid CNN–attention model with specialized loss functions (center), and resulting parking spot quality map (right).

networks (CNNs) are well-suited for learning hierarchical local features through convolutional filters, and—with fixed-scale patch sampling and limited pooling—can maintain the critical relationship between physical dimensions and spatial patterns [37]. At the same time, patch-level decisions can depend on contextual cues that extend beyond local receptive fields (*e.g.*, doorways or circulation paths spanning adjacent rooms). To capture such longer-range dependencies without sacrificing scale fidelity, we employ a hybrid CNN–attention approach in which a convolutional backbone provides local spatial encodings and cross-attention integrates broader contextual information across the patch. Building on these principles, we present an approach that applies patch-based sampling with scale-preserving CNNs augmented by attention to learn patterns of socially acceptable robot parking locations from annotated home floorplans.

3 Learning Robot Parking Preferences

We frame the task of modeling socially appropriate robot parking spots as generating a pose quality map of the environment where higher values represent more suitable parking spots. This representation allows a robot to efficiently query the estimated social appropriateness of any potential parking pose during operation after the map has been generated. Following this formulation, we first identify a discrete set of allowed poses P over a given floorplan f. Next, our approach generates a pose quality map $\hat{U}$ over this set. A mobile robot can then estimate the social appropriateness of any potential pose $p = (p_x, p_y)$ by querying this map as $\hat{U}(p)$. Note that we use "pose" and "parking spot" interchangeably throughout.

Most homes are naturally organized into distinct rooms or regions based on functional, social, and cultural patterns of space utilization [7]. While these divisions aren't always strict (*e.g.*, open floor plans, partially divided rooms, etc.), they provide a natural framework for evaluating poses within meaningful spaces. In this work, we assume these regions have been defined in the representation of the floorplan with each cell being assigned to a named region. A key challenge specific to learning parking preferences is that a robot must not only estimate the quality of individual poses, but must also reliably identify a high quality pose within each region. We define **region-level spot selection** as the task of identifying the most appropriate pose within a given region r from the set of available poses in that region.

One typical way to measure accuracy of the pose quality map would be to simply evaluate pairwise error between predicted scores and ground truth annotations across the entire floorplan. This *spot-level* metric would provide a broad evaluation of model fit, but it would not necessarily provide an accurate measure of the quality of the top poses selected by the model (*i.e.*, the poses that would be observable to users), as these are selected relative to other poses in the region. This calls for using a *region-level* metric that measures, say, the ground truth quality of the chosen pose within the region. This misalignment between spot-level and region-level metrics can open a significant gap—for instance, a pose quality map might achieve low spot-level error yet still incur high region-level error by failing to identify the best poses within each region. While training the model to optimize the region-level metric most closely aligns with model usage, it lacks the fine-grained supervision of the spot-level metric which utilizes our detailed pixel-level ground truth annotations. Therefore, our overall objective is to optimize two complementary goals: (1) accurately estimating pose quality across the entire floorplan (spot-level), and (2) correctly identifying the most appropriate poses within each region (region-level).

To achieve our dual spot-level and region-level objectives, we developed a specialized approach that incorporates both considerations during training and evaluation. In the following sections, we describe our flexible ML framework (**C1**) that learns user parking preferences from floorplan features.

3.1 Input Features and User Annotations

Our approach builds upon standard occupancy maps that most mobile robots use for navigation, which represent basic information about free and occupied space in the environment. From these maps, we can infer fundamental spatial features such as walls, clutter, and navigable space. As we investigate later, other semantic features, such as doors, cabinets, or appliance locations, can be incorporated if they are available via the robot's perception systems. We represent this environmental information as a set of 2D feature layers that maintain the spatial relationships and physical scale of the environment. These layers are represented as a discrete grid where each cell represents a fixed physical dimension (*e.g.*, 5cm × 5cm), ensuring that spatial patterns learned by the model correspond to real-world measurements.

Let $X = \{X^{(1)}, X^{(2)}, ..., X^{(k)}\}$ represent a set of k concatenated 2D feature channels extracted from the floorplan f. Here each

channel has the same spatial extent as the floorplan and captures salient information about the environment (*e.g.*, walls, clutter, navigable space). The feature set X is flexible and can incorporate any spatially-encoded information available to the robot.

For each floorplan f, our ground truth training data consists of human-annotated robot parking preferences U^* over the entire floorplan. Specifically, these annotations are not limited to the best spots within each region, which allows us to gain a fine-grained understanding of user preferences. For any pose $p \in P$, the user preference $U^*(p)$ is categorized as 0 (bad), $\frac{1}{2}$ (neutral), or 1 (good). Section 4.1 provides further details on this data collection.

3.2 Method Overview

We train a hybrid CNN–attention model M to predict the quality of a given pose p by consuming a square $s \times s$ patch X_p centered on that pose:

$$\hat{U}(p) = M(X_p) \tag{1}$$

Note that X_p is extracted from X and contains k feature channels, as described in Section 3.1. The model M is trained to optimize a weighted combination of the spot-level and region-level metrics. Such a model would typically be trained over several steps, using Gradient Descent on a batch of randomly sampled poses at each iteration. This strategy suffices to optimize the spot-level metric. Optimizing the region-level metric, however, requires simultaneous access to multiple within-region poses at each step, so the model can learn to identify the best amongst them. This necessitates a 2-step sampling strategy to construct each batch: first sample the regions, and then sample multiple poses for each region. The loss functions corresponding to the spot-level and the region-level metrics are averaged over the spots and the regions within each batch, respectively, before taking their weighted combination to compute the batch loss at training. Finally, once the model is trained, a single pose (or an arbitrary batch of poses) may be used at inference. We next describe each step of our methodology in detail.

3.3 Patch Sampling Strategy

Pose Generation. Given a floorplan f with cell resolution m meters and a navigability matrix N where $N(x, y) \in \{0, 1\}$, we first create the set of all parking spots P by generating poses every m meters along both axes:

$$P = \{p = (p_x, p_y) | N(p_x, p_y) = 1, p \in f\} \tag{2}$$

Additionally, for each region r on the floorplan, we construct the set of allowed poses within the region $P_r = r \cap P$.

Patch Extraction. For each pose $p \in P$, we extract a square patch of size $s \times s$ cells (representing $s \cdot m \times s \cdot m$ meters) centered at p from each feature channel in X. This patch represents a window view of the environmental features near the pose of interest, limiting the model's evaluation to only the local spatial context to determine the pose's appropriateness for parking. The patch size selection trades off between capturing meaningful local context and minimizing the influence of distant, irrelevant features. Using larger patch sizes also presents risks of overfitting and increased model complexity, especially if training data is limited. The multi-channel patch tensor for each pose is $X_p \in \mathbb{R}^{s \times s \times k}$, where k is the number of feature channels in X. Figure 2 illustrates this patch extraction process, showing sample patches extracted from different locations on a floorplan across three feature channels.

Patch Sampling. At training, the model M is trained in several steps, employing a batch of B sampled patches at each iteration. Our sampling strategy is guided by the choice of our objective function, which factors in both spot and region level considerations. The region-level metric incentivizes the model to correctly identify the best patch amongst those sampled from the same region. This requires sampling multiple patches within each region at training. We employ a 2-step strategy to operationalize this: first sample B_r regions over the entire train set, and next sample B_p poses uniformly from each sampled region r. Here, the regions are sampled in proportion to their size $|P_r|$, effectively ensuring a uniform distribution over the poses in the train set. Note that the overall batch size is $B = B_r \cdot B_p$, where $B_p > 1$ ensures we consider multiple within-region poses for computing the region-level metric.

3.4 Model Architecture

We frame pose quality estimation as a three-class supervised classification problem, with classes bad, neutral, and good, representing socially inappropriate, acceptable, and desirable poses, respectively. We choose classification over regression since parking preferences are often expressed as categorical judgments rather than continuous values. Furthermore, we opted not to use ordinal classification methods (such as CORAL [10]), as standard categorical classification better captures bimodal user preferences, *e.g.*, when a patch is considered desirable by some users but inappropriate by others.

We implement this classification framework using a hybrid CNN–attention model designed to process multi-channel patches of the environment, described in Algorithm 1. The network is designed to scale with patch size to ensure appropriate receptive fields for different spatial scales. As patch size increases, a deeper network is needed to allow each output neuron to be influenced by a larger portion of the input space, enabling the model to capture spatial relationships across the entire patch. We adopt a standard CNN backbone augmented with cross-attention, as this combination effectively captures local spatial features within regions while also modeling longer-range contextual dependencies across regions.

The model outputs class confidences $\mathbf{c}(p) = [c_1(p), c_2(p), c_3(p)]$ through a softmax layer, representing probabilities of bad, neutral, and good parking quality respectively. These are converted to a final quality score $\hat{U}(p) \in [0, 1]$ using weighted averaging:

$$\hat{U}(p) = 0 \cdot c_1(p) + \frac{1}{2} \cdot c_2(p) + 1 \cdot c_3(p) \tag{3}$$

3.5 Training Objective

In line with recent works [15, 48], we seek to align the training objective of our model with the decision-making it drives. We combine a per-patch classification loss with a differentiable surrogate loss for region-level spot selection. For each patch p, the model outputs class probabilities $\mathbf{c}(p)$ and a scalar prediction $\hat{U}(p) \in [0, 1]$ via Eq. (3). Ground-truth supervision is provided by $U^*(p)$, introduced in Section 3.1.

Spot-level objective. Let $y(p) \in \{1, 2, 3\}$ be the class index obtained by mapping the ground truth annotation $U^*(p) \in \{0, \frac{1}{2}, 1\}$

Algorithm 1 CNN Architecture for Patch Processing

Require: Input patch $X_p \in \mathbb{R}^{s \times s \times (k)}$
Ensure: Pose quality score $\hat{U}(p)$
$h \leftarrow X_p$ {Initialize hidden state}
$n_{blocks} \leftarrow \lfloor \log_2(s) \rfloor - 2$ {*e.g.*, 4 blocks for s=64, 3 for s=32}
for $i \leftarrow 1$ to n_{blocks} **do**
  $h \leftarrow \text{Conv2D}(h, \text{filters} = 32 \cdot 2^{\lfloor i/2 \rfloor})$
  $h \leftarrow \text{BatchNorm}(h)$
  $h \leftarrow \text{ReLU}(h)$
  if $i \geq n_{blocks} - 1$ **then**
    $h \leftarrow h + \text{CrossAttention}(h)$
  end if
end for
$h \leftarrow \text{GlobalAvgPool}(h)$
$\mathbf{c} \leftarrow \text{Softmax}(\text{Dense}(h))$ {Class probabilities}
$\hat{U}(p) \leftarrow 0 \cdot c_1 + \frac{1}{2} \cdot c_2 + 1 \cdot c_3$ {Weighted quality score}

return $\hat{U}(p)$

to its categorical label. We train the classifier with cross-entropy, using the class $y(p)$ as ground truth. The loss averages over all B patches in the batch:

$$\mathcal{L}_{\text{CE}} = -\frac{1}{B} \sum_{p} \log c_{y(p)}(p). \tag{4}$$

This cross-entropy loss ($\mathcal{L}_{\text{CE}}$) allows the model to harness the fine-grained user preferences captured by our annotated data.

Region-level objective. The model aids in region-level spot selection by selecting the pose with the highest predicted score. We create a differentiable surrogate for this argmax operation by employing a soft relaxation. Specifically, we use the standard softmax distribution over the predicted scores $\hat{U}(p)$ for all the sampled poses within one region. This can be seen as the top-1 special case of NeuralSort [19] and SoftStart [39].

Let S_r denote the set of sampled poses for region r in the current batch. Within each region r, we form a temperature-controlled soft selection over its poses using the model's scalar predictions $\hat{U}(p)$:

$$\alpha_{p|r}(\tau) = \frac{\exp\big(\hat{U}(p)/\tau\big)}{\sum_{q \in S_r} \exp\big(\hat{U}(q)/\tau\big)}, \qquad p \in S_r, \tag{5}$$

with temperature $\tau > 0$ (higher τ yields smoother weights; lower τ approaches a hard argmax). The soft choice is then evaluated against the region's ground truth. The region-level loss averages the squared gap over the B_r regions:

$$\mathcal{L}_{\text{region}} = \frac{1}{B_r} \sum_{r} \left(\max_{p \in S_r} U^*(p) - \sum_{p \in S_r} \alpha_{p|r}(\tau) \cdot U^*(p) \right)^2. \tag{6}$$

Overall objective. The final loss is a weighted combination of the two terms:

$$\mathcal{L} = \mathcal{L}_{\text{CE}} + \lambda \cdot \mathcal{L}_{\text{region}}, \tag{7}$$

where λ balances spot-wise accuracy against reliable region-level spot selection.

3.6 Quality Map Construction

Given a floorplan f, we generate a quality map by evaluating each navigable pose $p \in P$:

$$\hat{U} = \{\hat{U}(p) \mid p \in P\} \tag{8}$$

This map can be regenerated whenever the floorplan is updated to reflect changes in the environment.

4 Evaluation

To evaluate our proposed method for robot parking spot prediction, we conducted a user data collection to obtain annotated floorplan data from 84 participants with direct robot experience in their homes (**C3**), which was then used to train and evaluate the model.

We investigated four research questions related to the implementation and effectiveness of the proposed model. As a baseline for comparison, we used the approach from Glas and Smart [17] that estimates parking spot quality based on avoidance of estimated walking paths, which we will designate "WP".

(1) **Patch Size:** How much nearby spatial context (patch size) is most effective to learn human parking spot preferences?
(2) **Improvement Over Baseline:** How effectively does our ML model capture user preferences for robot parking spots? How does performance compare to the state-of-the-art baseline?
(3) **Learned Spatial Patterns:** What high-level spatial patterns about good parking locations has the model learned?
(4) **Approach Extensibility:** How does incorporating additional semantic features affect model performance and our understanding of parking spot selection?

4.1 Data Collection

Our study uses floorplan data and robot parking preference annotations from 84 employees of a technical company who had experience with the Astro robot [1] in their homes for at least two weeks. 16.7% of participants were unpaid volunteers and 83.3% were paid participants recruited through the company's beta testing program. This data collection was conducted in accordance with our company's legal, ethical, and privacy policies, with informed consent obtained from all participants. From each home, we collected the robot's occupancy map and used it to generate a set of potential parking poses. These poses were sampled uniformly across all navigable space (areas accessible to the 0.5m-wide Astro robot) at $m = 0.25$m intervals, as shown in Figure 3.

As part of the standard robot setup process, participants divided their homes into regions corresponding to rooms or other semantically meaningful spaces. To collect robot parking preferences, we presented each participant with a 2D visualization of their home floorplan and asked them to identify good and bad areas for robot parking, with instructions to mark at least one good parking area in each room where possible. The 2D representation captured (x, y) positions only. Orientation can be determined post-hoc based on geometry or user locations [30].

We processed the annotated floorplans to align them with the corresponding occupancy maps, assigning scores of 1 to areas marked as good for parking, 0 to areas marked as bad, and 0.5 to unmarked areas. The resulting dataset consists of 84 de-identified floorplan maps containing 47,618 unique poses grouped into 728 regions. The

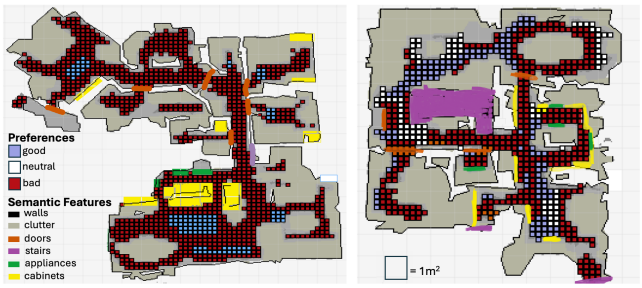

Figure 3: Example floorplans with user-annotated parking preferences and semantic features on maps. Each colored square represents a viable pose in navigable space.

distribution of scores (mean = 0.40, SD = 0.14) reflects the inherent scarcity of good parking spots in home environments, where many areas are unsuitable for robot parking due to spatial and social constraints (see Table 1).

4.2 Experimental Design

To evaluate RQ1 on patch size, we generated 7 variants of the feature dataset X using patch sizes in $\{2, 4, 8, 16, 32, 64, 128\}$ (*i.e.*, ranging from 0.1m to 6.4m width on a side, centered on a potential parking spot). A model was designed using the architecture described in Section 3.5 for each dataset variant with its respective patch size. The temperature parameter τ was annealed from 1.0 to 0.1 over the first 20 epochs of training and the region-level loss weight λ was set to 1.0.

For RQ2, we compared the best performing model to two baselines, a state-of-the-art walking path avoidance model (WP) [17], and random spot selection (RS).

For RQ3, we analyzed the spatial patterns learned by the model through clustering analysis of poses where it significantly outperformed the baseline.

Finally, we evaluated the model's extensibility (RQ4). Our base model used three basic environmental features available from standard robot occupancy maps: walls, clutter, and navigable space, where "clutter" represents any obstacle lower than ceiling height (*e.g.*, furniture, counters). We then augmented this base feature set with additional semantic information: doors, stairs, appliances, and cabinets. We evaluated the model's performance with these richer environmental representations and identified which semantic features contributed most through an ablation study.

Table 1: Summary statistics of the parking preference dataset (mean ± standard deviation where applicable).

Metric	Total	Per Floorplan	Per Region
Poses	47,618	566.9 ± 302.0	65.4 ± 63.0
Score Dist.			
Bad (0)	42.1%	40.8 ± 19.8%	49.9 ± 37.8%
Neutral (0.5)	35.0%	34.0 ± 24.6%	32.9 ± 34.6%
Good (1)	22.9%	25.2 ± 18.2%	17.2 ± 24.9%
Mean score	0.40 ± 0.39	0.42 ± 0.14	0.34 ± 0.27

4.3 Evaluation Metrics

Following [17], we evaluate our approach using three complementary metrics that capture different aspects of model performance. Each metric is computed using the user preference map U^* and our computed quality map $\hat{U}$ containing the predicted scores from model M. All scores are scaled to $[0, 1]$, with higher values indicating better/more preferred positions. Poses are organized into user-defined regions R, where each pose p belongs to one region $r \in R$. The metrics are:

- **Estimation Accuracy (EA):** Measures overall fit between predicted and user preferences across all poses:

$$EA = 1 - \frac{1}{n}\sum_{i=1}^{n} |U^*(i) - \hat{U}(i)| \tag{9}$$

  EA ranges from 0 to 1, with 1 indicating perfect alignment between model predictions and user preferences. This metric provides a global view of model performance across all poses.
- **Spot Quality (SQ):** Evaluates the quality of the parking poses with the highest predicted score in each region:

$$SQ = \frac{1}{|R|}\sum_{r\in R} U^*(\arg\max_{p\in P_r} \hat{U}(p)) \tag{10}$$

  where P_r is the set of poses in region r. SQ ranges from 0 to 1, but notably, the maximum achievable score may be less than 1 in regions where the best available pose(s) were marked as neutral (0.5) rather than good (1).
- **Region Agreement (RA):** Agreement between model and user preferences on whether regions should contain parking spots:

$$PA = \frac{\text{number of regions where model and user agree}}{|R|} \tag{11}$$

  where agreement occurs when both the model's maximum score in the region ($\max_{p\in P_r} \hat{U}(p)$) and the user's maximum score in the region ($\max_{p\in P_r} U^*(p)$) are either both ≥ 0.5 or both < 0.5.

These metrics collectively provide a comprehensive evaluation of the model's performance. While Estimation Accuracy provides a overall view of model performance across the entire floorplan, it doesn't fully capture the nuances of parking spot selection. In practice, a robot may only need to identify one to a few good parking spots in each region, so the accuracy of the poses with the highest predicted scores are important. To address these limitations, we use Spot Quality to evaluate the model's ability to identify the best parking spot in each region, and Region Agreement to assess how well the model identifies parkable and non-parkable regions across the home. Together, these metrics provide a more comprehensive evaluation of the model's performance in real-world scenarios.

5 Results

5.1 Effect of Patch Size on Model Performance

As shown on the left of Figure 4, model performance improved with patch size, achieving best overall results at 64×64 pixels (3.2m side length), beyond which performance degraded.

Patches below 16 pixels (0.8m side length at 5cm resolution) appeared to lack sufficient context to learn meaningful patterns. Performance improved with increasing patch size up to 3.2m as more relevant features (e.g., from small furniture and doorways up to walking paths and room shapes) became visible, then declined at

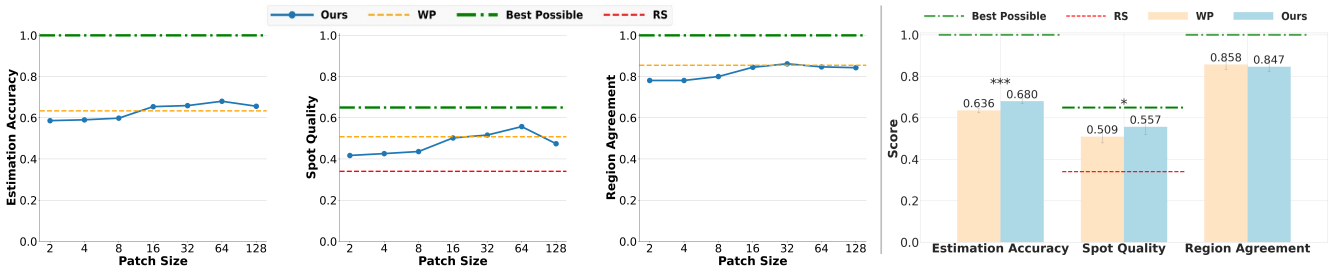

Figure 4: Left: Performance across different patch sizes for our three evaluation metrics, with walking path baseline (WP, orange), random selection baseline (RS, red), and Ours (blue). Right: Performance comparison between 64 patch size model and baselines. Green line indicates maximum achievable scores for each metric (1.0 for EA and RA; 0.65 average for SQ, following [17]). Error bars show 95% bootstrapped confidence intervals. Statistical significance: *** p < 0.001, * p < 0.05.

128 pixels (6.4m), likely due to noise from irrelevant distant areas and increased model complexity. Based on this observation, we recommend tuning patch size to capture approximately 3.2m of real-world context for home environments.

5.2 Comparison with Baselines

Next, we evaluate how well the best trained model captures user preferences in comparison to two baselines, the walking path (WP) approach [17], and random spot selection (RS). The WP baseline provides predictions across the full floorplan, enabling comparison on all metrics, while RS is only meaningful for spot quality as it simply selects poses at random within each region.

As shown in Figure 4 (right), Wilcoxon signed-rank tests reveal that the ML model significantly outperforms the WP baselines in estimation accuracy ($W = 724.5$, $p < 0.001$, Cohen's $|d| = 0.553$, large effect) and spot quality ($W = 1199.0$, $p < 0.05$, Cohen's $|d| = 0.222$, small effect). No significant difference was observed for region agreement ($W = 317.5$, $p = 0.608$). For spot quality, both ML and WP approaches significantly outperformed random spot selection (RS) with small effects (ML vs RS: Cohen's $|d| = 1.08$; WP vs RS: Cohen's $|d| = 0.99$, both $p < 0.001$), demonstrating that both methods were successful in identifying appropriate parking poses.

5.3 Analysis of Learned Spatial Patterns

The WP approach primarily identifies narrow corridors and bottlenecks, so our aim in RQ3 was to understand what spatial patterns beyond this our model had learned. We manually identified map areas where the ML model's predictions outperformed the WP baseline in capturing user preferences. From these examples, we broadly identified three categories where our model displayed noticeable improvement across multiple floorplans, illustrated in Figure 5:

(1) **Hallways**: Even wide hallways make poor parking locations. Although the WP algorithm generally identified hallways as bad spots to park, variations in hallway width (real or perceived) sometimes led to a mix of negative and neutral ratings, whereas our algorithm often cleanly marked entire hallways as bad, likely due to visually recognizing the feature as a hallway.
(2) **Transitional Spaces**: Sometimes major rooms or functional spaces were connected by an opening or short passage which is wide enough that WP judges it as good for parking. However, we observed several cases where users marked these connecting areas as bad, which our model successfully captured.
(3) **Dead Ends**: Enclosed spaces such as nooks, tight corners, bathrooms, and utility areas were sometimes marked as bad by users. We observed cases where these dead ends were considered wide enough by the WP approach but correctly evaluated as bad by our model, possibly based on the room shapes and sizes.

In each of these scenarios, while the WP baseline measured adequate physical space for a human and robot to pass each other, users rated these locations poorly, ostensibly for social or functional reasons. Our model appears to have captured enough spatial context to better predict user preferences in these cases.

5.4 Incorporating Additional Semantic Features

One advantage of our approach is that it is straightforward to extend the model with new types of feature layers. To evaluate this extensibility (RQ4), we augmented the feature set with additional semantic information for 46 of the 84 total user map annotations, with examples shown in Figure 3. We hypothesized that features like doors, stairs, appliances, and cabinets would be useful in identifying bad spots for parking, so we asked users to manually label these features on their floorplan maps during the annotation process, providing additional feature layers for model training.

After retraining the model with these additional features, we again observed significant improvement over the WP baseline in estimation accuracy ($W = 708.0$, $p < 0.001$, Cohen's $|d| = 0.553$, medium effect) and spot quality ($W = 942.5$, $p < 0.01$, Cohen's $|d| = 0.337$, small effect). No significant difference was observed for region agreement ($W = 615.0$, $p = 0.063$). Compared to the RS baseline, we observed an even stronger effect (Cohen's $|d| = 1.32$, $p < 0.0001$) than the model with only the basic features ($|d| = 1.08$).

When we compared the model with the added features to model with the original feature set, we found no significant differences for any of the three metrics, though we observed a slight increase in spot quality (from 0.565 ± 0.241 to 0.583 ± 0.255, Wilcoxon signed-rank test: $W = 1604.000$, $p = 0.790$). Since only 54% of our dataset was augmented with the additional features, we were not expecting large improvements in overall performance.

We conducted a feature importance analysis using an ablation study, measuring the observed change in overall estimation accuracy (Figure 6). Among the semantic features, doors emerged as

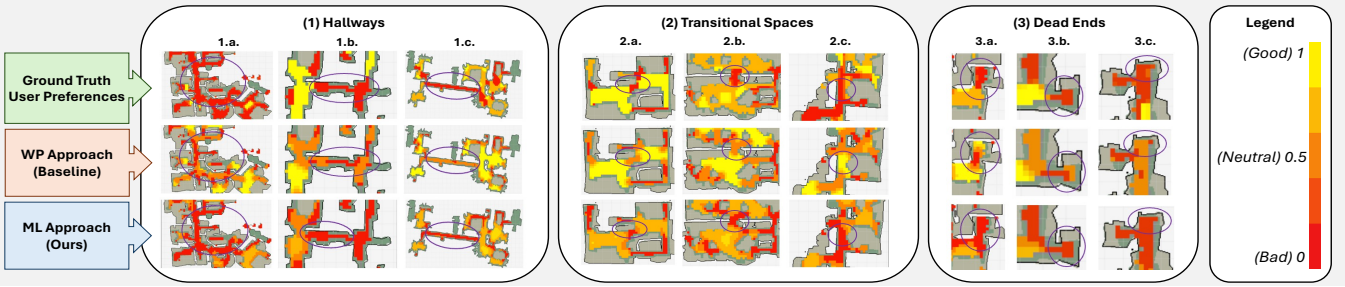

Figure 5: Selected scenarios where ML outperformed WP (three examples per pattern). Rows: user preferences (top), baseline WP pose quality maps (middle), ML pose quality maps (bottom). The insights of interest are circled.

the most impactful. Doors were also the most frequent feature in our dataset (averaging 4.85 per map, std: 3.25), while stairs were the least frequent (averaging 0.89 per map, std: 1.20). This suggests that door recognition capabilities could be particularly valuable for autonomous robot parking systems, aligning with prior work emphasizing the importance of keeping doorways clear [8].

6 Discussion and Limitations

Real-time data. The method developed here is limited to static inputs. By contrast, the majority of research into socially-aware navigation focuses on the locations and behavior of humans [43]. A natural next step for this work would be to incorporate real-time data about human locations and behaviors to dynamically choose parking spots in a socially-aware way.

Robot platform. As with many initial studies in robot behavior learning, our evaluation focused on a single robot platform to establish foundational principles that could later extend to other form factors. We expect that actual parking spot preferences may differ for robots of different sizes or shapes, although the proposed technique should still be an effective way to model them.

User diversity. To ensure realism of the collected preference data, we needed to structure our experimental validation around map annotations from users with direct robot experience, although this limited the potential diversity of our dataset. It would be interesting to collect more data across diverse users and environments (varying home layouts, sizes and household compositions), although doing this in practice presents significant logistical challenges [31].

Individual preferences. In our data collection we observed differences in some individual preferences, visible in the two examples shown in Figure 3. Selected post-hoc followup revealed some users strongly preferred wall-adjacent parking to minimize obstruction (right), *"Always [park] near walls to be out of the way, yet accessible. Avoid being a trip hazard"*, while others valued visibility and interaction potential (left), *"[Parking in open space], that is my preference. I prefer Astro is visible when people are around"*, and were less sensitive to the robot being in the way, *"I don't mind him [parking] in the middle of the room if I have enough space to walk"*. While these responses support individual differences, our floorplan data alone cannot disentangle whether variance stems from personal preferences or environmental characteristics not captured in the maps.

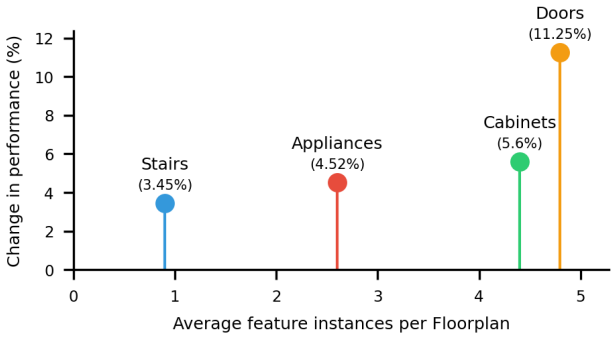

Figure 6: Impact of semantic features on model performance vs. frequency per floorplan, measured through ablation.

Future work could build on these findings to support adaptation to individual or household-specific parking behavior [20].

7 Conclusion

The vision of social robots operating autonomously in the home environment has persisted since early prototypes like the Hero Jr. robot kit in 1984, which highlighted the importance of robots being accessible and ready for use without explicit instruction [25].

In this work, we presented a flexible machine learning approach to modeling the social appropriateness of robot parking in the home. We designed a learning framework and novel loss function that balance accurate pose quality estimation across the floorplan with effective region-level spot selection. Our approach was validated using a dataset of robot parking preferences collected from 84 users with at least 2 weeks of experience using the Astro robot in their homes. Our results demonstrated that the ML approach significantly outperformed a state-of-the-art baseline [17] in estimation accuracy and spot quality. We presented examples of features learned by the system, and we demonstrated its extensibility by incorporating additional semantic features, leading to incremental improvements in model performance. As robot sensing capabilities continue to evolve, our approach presents an adaptable and extensible framework that can support space selection for a variety of intelligent social robot behaviors, not limited to parking. Four decades after Hero Jr., this work represents another step toward realizing the vision of robots that can seamlessly coexist in our home environments.

References

[1] 2021. Amazon Astro, household robot for home monitoring. https://www.amazon.com/Introducing-Amazon-Astro/dp/B078NSDFSB

[2] Isaac Asimov. 1983. *The Robots of Dawn.* Doubleday, Garden City, N.Y.

[3] Eleanor Avrunin and Reid Simmons. 2014. Socially-appropriate approach paths using human data. In *The 23rd IEEE International Symposium on Robot and Human Interactive Communication*. IEEE, 1037–1042. doi:10.1109/ROMAN.2014.6926389

[4] Lynne Baillie and David Benyon. 2008. Place and technology in the home. *Computer Supported Cooperative Work (CSCW)* 17, 2 (2008), 227–256. doi:10.1007/s10606-007-9063-2

[5] Hrishav Bakul Barua, Theint Haythi Mg, Pradip Pramanick, and Chayan Sarkar. 2024. Enabling Social Robots to Perceive and Join Socially Interacting Groups using F-formation: A Comprehensive Overview. *ACM Transactions on Human-Robot Interaction* 13, 4 (2024), 1–48. doi:10.1145/3682072

[6] Tony Belpaeme, James Kennedy, Aditi Ramachandran, Brian Scassellati, and Fumihide Tanaka. 2018. Social robots for education: A review. *Science robotics* 3, 21 (2018), eaat5954. doi:10.1126/scirobotics.aat5954

[7] Richard Bormann, Florian Jordan, Wenzhe Li, Joshua Hampp, and Martin Hägele. 2016. Room segmentation: Survey, implementation, and analysis. In *2016 IEEE international conference on robotics and automation (ICRA)*. IEEE, 1019–1026. doi:10.1109/ICRA.2016.7487234

[8] De'Aira Bryant, Tiago Etiene, Ayanna Howard, William D Smart, and Dylan F Glas. 2023. Teaching a robot where to park: A scalable crowdsourcing approach. In *2023 32nd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 226–233. doi:10.1109/RO-MAN57019.2023.10309612

[9] Bengisu Cagiltay, Hui-Ru Ho, Joseph E Michaelis, and Bilge Mutlu. 2020. Investigating family perceptions and design preferences for an in-home robot. In *IDC '20: Proceedings of the Interaction Design and Children Conference*. 229–242. doi:10.1145/3392063.3394411

[10] Wenzhi Cao, Vahid Mirjalili, and Sebastian Raschka. 2020. Rank consistent ordinal regression for neural networks with application to age estimation. *Pattern Recognition Letters* 140 (2020), 325–331. doi:10.1016/j.patrec.2020.11.008

[11] Maartje de Graaf. 2015. *Living with robots: Investigating the user acceptance of social robots in domestic environments.* Ph. D. Dissertation. University of Twente. doi:10.3990/1.9789036538794

[12] Maartje M. A. de Graaf, Somaya Ben Allouch, and Jan AGM Van Dijk. 2016. Long-term evaluation of a social robot in real homes. *Interaction studies* 17, 3 (2016), 462–491. doi:10.1075/is.17.3.08deg

[13] Kevin T Decker and Brett J Borghetti. 2025. A Survey of Sampling Methods for Hyperspectral Remote Sensing: Addressing Bias Induced by Random Sampling. *Remote Sensing* 17, 8 (2025), 1373. doi:10.3390/rs17081373

[14] Andrea Eirale, Mauro Martini, Luigi Tagliavini, Dario Gandini, Marcello Chiaberge, and Giuseppe Quaglia. 2022. Marvin: An innovative omni-directional robotic assistant for domestic environments. *Sensors* 22, 14 (2022), 5261. doi:10.3390/s22145261

[15] Adam N Elmachtoub and Paul Grigas. 2022. Smart "predict, then optimize". *Management Science* 68, 1 (2022), 9–26. doi:10.1287/mnsc.2020.3922

[16] Paolo Fiorini and Erwin Prassler. 2000. Cleaning and household robots: A technology survey. *Autonomous robots* 9, 3 (2000), 227–235. doi:10.1023/A:1008954632763

[17] Dylan F Glas and William D Smart. 2024. Where can I park my robot? Modeling out-of-the-way parking spots in the home using room geometry. In *2024 33rd IEEE International Conference on Robot and Human Interactive Communication (ROMAN)*. IEEE, 2079–2086. doi:10.1109/RO-MAN60168.2024.10731293

[18] Lindsay T Graham, Samuel D Gosling, and Christopher K Travis. 2015. The psychology of home environments: A call for research on residential space. *Perspectives on Psychological Science* 10, 3 (2015), 346–356. doi:10.1177/1745691615576761

[19] Aditya Grover, Eric Wang, Aaron Zweig, and Stefano Ermon. 2019. Stochastic optimization of sorting networks via continuous relaxations. *arXiv preprint arXiv:1903.08850* (2019).

[20] Mehdi Hellou, Norina Gasteiger, Jong Yoon Lim, Minsu Jang, and Ho Seok Ahn. 2021. Personalization and localization in human-robot interaction: A review of technical methods. *Robotics* 10, 4 (2021), 120. doi:10.3390/robotics10040120

[21] Yutaka Hiroi and Akinori Ito. 2019. A pedestrian avoidance method considering personal space for a guide robot. *Robotics* 8, 4 (2019), 97. doi:10.3390/robotics8040097

[22] Le Hou, Dimitris Samaras, Tahsin M Kurc, Yi Gao, James E Davis, and Joel H Saltz. 2016. Patch-based convolutional neural network for whole slide tissue image classification. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2424–2433. doi:10.1109/CVPR.2016.266

[23] Jingwen Hu, Gui-Song Xia, Fan Hu, and Liangpei Zhang. 2015. A comparative study of sampling analysis in the scene classification of optical high-spatial resolution remote sensing imagery. *Remote Sensing* 7, 11 (2015), 14988–15013. doi:10.3390/rs71114988

[24] Takamasa Iio, Yuichiro Yoshikawa, and Hiroshi Ishiguro. 2024. Multi-robot cooperative behavior for reducing unnaturalness of starting a conversation. *Advanced Robotics* 38, 7 (2024), 465–481. doi:10.1080/01691864.2023.2290151

[25] Ron Johnson. 1983. Here comes the HERO: Designed as a training device, the HERO I, the world's first mass-produced personal robot, can walk, talk, and guard a room. *IEEE Potentials* 2, Fall (1983), 38–40. doi:10.1109/MP.1983.6499641

[26] Takayuki Kanda, Dylan F Glas, Masahiro Shiomi, and Norihiro Hagita. 2009. Abstracting people's trajectories for social robots to proactively approach customers. *IEEE Transactions on Robotics* 25, 6 (2009), 1382–1396. doi:10.1109/TRO.2009.2032969

[27] Cory D Kidd and Cynthia Breazeal. 2008. Robots at home: Understanding long-term human-robot interaction. In *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 3230–3235. doi:10.1109/IROS.2008.4651113

[28] Takuya Kitade, Satoru Satake, Takayuki Kanda, and Michita Imai. 2013. Understanding suitable locations for waiting. In *2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 57–64. doi:10.1109/HRI.2013.6483502

[29] Tanmoy Kundu and Indranil Saha. 2023. Approximation algorithms for charging station placement for mobile robots. In *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 4770–4776. doi:10.1109/IROS55552.2023.10341361

[30] Jin Joo Lee, Amin Atrash, Dylan F Glas, and Hanxiao Fu. 2023. Developing autonomous behaviors for a consumer robot to be near people in the home. In *2023 32nd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 197–204. doi:10.1109/RO-MAN57019.2023.10309443

[31] Iolanda Leite, Carlos Martinho, and Ana Paiva. 2013. Social robots for long-term interaction: a survey. *International Journal of Social Robotics* 5, 2 (2013), 291–308. doi:10.1007/s12369-013-0178-y

[32] Phoebe Liu, Dylan F Glas, Takayuki Kanda, and Hiroshi Ishiguro. 2018. Learning proactive behavior for interactive social robots. *Autonomous Robots* 42, 5 (2018), 1067–1085. doi:10.1007/s10514-017-9671-8

[33] Tamara Lorenz, Astrid Weiss, and Sandra Hirche. 2016. Synchrony and reciprocity: Key mechanisms for social companion robots in therapy and care. *International Journal of Social Robotics* 8, 1 (2016), 125–143. doi:10.1007/s12369-015-0325-8

[34] Matteo Luperto, Tomasz Piotr Kucner, Andrea Tassi, Martin Magnusson, and Francesco Amigoni. 2022. Robust Structure Identification and Room Segmentation of Cluttered Indoor Environments From Occupancy Grid Maps. *IEEE Robotics and Automation Letters* 7, 3 (2022), 7974–7981. doi:10.1109/LRA.2022.3186495

[35] Allison C Marsh. 2013. Revolution: the first 2,000 years of computing: the computer history museum, mountain view, California. *Technology and Culture* 54, 3 (2013), 640–649. doi:10.1353/tech.2013.0094

[36] Harvey J Miller. 2004. Tobler's first law and spatial analysis. *Annals of the association of American geographers* 94, 2 (2004), 284–289. doi:10.1111/j.1467-8306.2004.09402005.x

[37] Behnam Nikparvar and Jean-Claude Thill. 2021. Machine learning of spatial data. *ISPRS International Journal of Geo-Information* 10, 9 (2021), 600. doi:10.3390/ijgi10090600

[38] Sai Krishna Pathi, Annica Kristofferson, Andrey Kiselev, and Amy Loutfi. 2019. Estimating optimal placement for a robot in social group interaction. In *2019 28th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 1–8. doi:10.1109/RO-MAN46459.2019.8956318

[39] Sebastian Prillo and Julian Eisenschlos. 2020. Softsort: A continuous relaxation for the argsort operator. In *International Conference on Machine Learning*. PMLR, 7793–7802.

[40] Satoru Satake, Takayuki Kanda, Dylan F Glas, Michita Imai, Hiroshi Ishiguro, and Norihiro Hagita. 2012. A robot that approaches pedestrians. *IEEE Transactions on Robotics* 29, 2 (2012), 508–524. doi:10.1109/TRO.2012.2226387

[41] Nur Muhammad Mahi Shafiullah, Anant Rai, Haritheja Etukuru, Yiqian Liu, Ishan Misra, Soumith Chintala, and Lerrel Pinto. 2023. On bringing robots home. *arXiv preprint arXiv:2311.16098* (2023).

[42] Yoshiaki Shiokawa, Winnie Chen, Aditya Shekhar Nittala, Jason Alexander, and Adwait Sharma. 2025. Beyond Vacuuming: How Can We Exploit Domestic Robots' Idle Time?. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. 1–17. doi:10.1145/3706598.3714266

[43] Phani Teja Singamaneni, Pilar Bachiller-Burgos, Luis J Manso, Anaís Garrell, Alberto Sanfeliu, Anne Spalanzani, and Rachid Alami. 2024. A survey on socially aware robot navigation: Taxonomy and future challenges. *The International Journal of Robotics Research* 43, 10 (2024), 1533–1572. doi:10.1177/02783649241230562

[44] Jared Strader, Nathan Hughes, William Chen, Alberto Speranzon, and Luca Carlone. 2024. Indoor and Outdoor 3D Scene Graph Generation Via Language-Enabled Spatial Ontologies. *IEEE Robotics and Automation Letters* 9, 6 (2024), 4886–4893. doi:10.1109/LRA.2024.3384084

[45] José Armando Sánchez-Rojas, José Aníbal Arias-Aguilar, Hiroshi Takemura, and Alberto Elías Petrilli-Barceló. 2021. Staircase Detection, Characterization and Approach Pipeline for Search and Rescue Robots. *Applied Sciences* 11, 22 (2021). doi:10.3390/app112210736

[46] Tommaso Van Der Meer, Andrea Garulli, Antonio Giannitrapani, and Renato Quartullo. 2025. A Comparative Study of Human Motion Models in Reinforcement Learning Algorithms for Social Robot Navigation. *ACM Transactions on Human-Robot Interaction* (2025). doi:10.1145/3746463

[47] Florian Vaussard, Julia Fink, Valérie Bauwens, Philippe Rétornaz, David Hamel, Pierre Dillenbourg, and Francesco Mondada. 2014. Lessons learned from robotic vacuum cleaners entering the home ecosystem. *Robotics and Autonomous Systems* 62, 3 (2014), 376–391. doi:10.1016/j.robot.2013.09.014

[48] Bryan Wilder, Bistra Dilkina, and Milind Tambe. 2019. Melding the data-decisions pipeline: Decision-focused learning for combinatorial optimization. In *Proceedings of the AAAI conference on artificial intelligence*, Vol. 33. 1658–1665. doi:10.1609/aaai.v33i01.33011658

[49] Joey Wilson, Marcelino Almeida, Min Sun, Sachit Mahajan, Maani Ghaffari, Parker Ewen, Omid Ghasemalizadeh, Cheng-Hao Kuo, and Arnie Sen. 2025. Modeling Uncertainty in 3D Gaussian Splatting through Continuous Semantic Splatting. In *2025 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 3284–3290. doi:10.1109/ICRA55743.2025.11128341

[50] Junpei Zhong, Chaofan Ling, Angelo Cangelosi, Ahmad Lotfi, and Xiaofeng Liu. 2021. On the gap between domestic robotic applications and computational intelligence. *Electronics* 10, 7 (2021), 793. doi:10.3390/electronics10070793

[51] Yuxuan Zhou, Jiarui Ding, Chaoyi Li, Ning Xie, and Xinran Wang. 2024. Analysis of the Pet Companion Robot Market. *Advances in Economics, Management and Political Sciences* 106 (2024), 231–235. doi:10.54254/2754-1169/106/20241679

Received 2025-09-30; accepted 2025-12-01