This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/MCE.2021.3101060, IEEE Consumer Electronics Magazine

KF-Loc: A Kalman Filter and Machine Learning Integrated Localization System Using Consumer-Grade Millimeter-wave Hardware

Abhishek Vashist
Rochester Institute of Technology

Maojia Patrick Li
Rochester Institute of Technology

Amlan Ganguly
Rochester Institute of Technology

Sai Manoj P D
George Mason University

Clark Hochgraf
Rochester Institute of Technology

Raymond Ptucha
Rochester Institute of Technology

Andres Kwasinski
Rochester Institute of Technology

Michael E. Kuhl
Rochester Institute of Technology

With the ever-increasing demands of e-commerce, the need for smarter warehousing is increasing exponentially. Such warehouses requires industry automation beyond Industry 4.0. In this work, we use consumer-grade millimeter-wave (mmWave) equipment to enable fast, and low-cost implementation of our localization system. However, the consumer-grade mmWave routers suffer from coarse-grained channel state information due to cost-effective antenna array design limiting the accuracy of localization systems. To address these challenges, we present a Machine Learning (ML) and Kalman Filter (KF) integrated localization system (KF-Loc). The ML model learns the complex wireless features for predicting the static position of the robot. When in dynamic motion, the static ML estimates suffer from position mispredictions, resulting in loss of accuracy. To overcome the loss in accuracy, we design and integrate a KF that learns the dynamics of the robot motion to provide highly accurate tracking. Our system achieves centimeter-level accuracy for the two aisles with RMSE of 0.35m and 0.37m, respectively. Further, compared with ML only localization systems, we achieve a significant reduction in RMSE by 28.5% and 54.3% within the two aisles.

I. Introduction

In the recent decade, advances in industrial automation have gained a lot of attention from academia and industry. Leading to the evolution of the fourth industrial revolution, known as Industry 4.0 [1]–[3]. The advances are made in the computation and the communication aspects of industrial automation targeting robotics and material handling. Realizing Industry 4.0 requires the design of autonomous robots working in coordination to perform various tasks. One of the preliminary requirements for the robot is to localize itself within the environment with very high accuracy for coordination and carrying out tasks. Localization requires information from various kinds of sensors. Sensors such as LiDAR, vision and wireless are most commonly used for localization [4]–[7]. Authors in [8] detailed discussion regarding advanced driver assistance systems utilizing wireless, radar, and LiDAR-based hardware and software components for self-driving vehicles. They present the use of different consumer sensors for tasks like localization, mapping, and navigation. In [9] authors use deep learning for designing smarter and intelligent consumer devices and software services.

In recent years, with the maturity of consumer-grade sensor technology, localization algorithms are heavily studied. However, reliable localization is still challenging for accurate and reliable autonomous design. The challenges arise due to the unpredictable environmental conditions for both indoor and outdoor applications. Traditional approaches to wireless localization use triangulation and trilateration to estimate the position of the client. Using such an approach for localization requires accurate knowledge of the wireless channel

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/MCE.2021.3101060, IEEE Consumer Electronics Magazine

TABLE I
PERFORMANCE COMPARISON WITH DIFFERENT WIRELESS BASED LOCALIZATION TECHNIQUES

Sensor	Performance	Advantages	Disadvantages
GPS	Around 10m	Low cost sensor off-the shelf localization	NLoS in indoor environment Significant loss in accuracy for indoor application
Camera [4]	Mean error of 0.75m	Low cost sensor High accuracy can be achieved	Susceptible to environment changes Require more complex localization system design
LiDAR [5]	0.07m to 0.03m RMSE	High Accuracy Obstacle avoidance ability	Computationally very expensive High cost sensor
UWB [6]	15cm RMSE	Low cost sensor Low power consumption	High dependence on sensor placement Need custom hardware design

model is required. Millimeter-wave channel in an indoor environment suffers from multi-path propagation [10]. To model the wireless channel, detailed knowledge of the electromagnetic characteristics from all scatters is required. Such modeling of indoor wave propagation at 60 GHz is non-trivial and involves custom hardware design. For this, the mmWave equipment used in our approach are the consumer-grade routers capable of communicating at 60 GHz frequency. Such consumer hardware is designed to communicate and provides coarse-grained channel state information due to cost-effective antenna array design. We propose a machine learning and Kalman filter-based indoor localization system to alleviate the challenges associated with channel modeling and irregular signal strength. The system learns the mapping between the complex wireless radio features and the distances to provide dynamic position tracking.

Many data augmentation techniques are heavily employed for machine learning applications like image classification, object detection, and object recognition, where the images in the training datasets are rotated, and cropped. The augmentation is performed for two primary reasons: first, to increase the size of the training dataset without physically collecting more data, which saves the data collection time, and second, to increase the robustness of the ML model for inference on unseen test data and has shown drastic and significant improvement in the learning capabilities of neural networks. Most of the data augmentation techniques are for image/vision-based datasets [11]. In our design, the features are the wireless SNR information and not image pixels and cannot be used on wireless features. To improve ML-based localization models, we propose a new data augmentation technique for wireless features. To perform data augmentation, we synthetically generate another set of radio map, assuming a large-scale shadow fading in the 60GHz propagation model [10].

Table I summarizes the localization performance using different sensor modalities, namely, Global Positioning System (GPS), LiDAR, and vision. Simultaneous Localization and Mapping (SLAM) approaches use information from sensors, and depending on the environment, high accuracy can be achieved. The LiDAR-based SLAM approach, as well as vision-based systems, are computationally costly. The data generated is either a point cloud of all the distances or video streams from hi-fidelity cameras respectively [7]. To alleviate these concerns, high-performance computational devices such as Graphics Processing Units (GPUs) are needed on the robot to support the localization. While custom mmWave wireless equipment could be used as wireless beacons, designing such hardware at mmWave frequencies is significantly more challenging and is beyond the scope of this work. Further, to enable a low barrier to entry of automation for small-scale and local warehouses without compromising the system's accuracy and performance, we have introduced the use off-the shelf mmWave routers. Our approach is designed and evaluated for a working warehouse. The methodology presented can be extended to any indoor

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/MCE.2021.3101060, IEEE Consumer Electronics Magazine

environment for agent localization utilizing wireless features. The exact results may vary, but the methodology can be adopted across different indoor environments. Further, we have discussed the network training details and AP arrangements process in Section III and IV. The evaluation details can be used to reproduce or re-design the proposed system for different indoor environments.

The contributions of this work are outlined as follows:

- We present design of an indoor warehouse localization system using consumer-grade off-the-shelf 60 GHz wireless routers. For this, we use Signal-to-Noise-Ratio (SNR) as a feature from consumer-grade wireless Access Points (APs) in ML based localization algorithm.
- We introduce a method to deal with the missing feature information from the APs during the training and inference of the ML models, as consumer-grade APs can lose connectivity intermittently, which can cause severe performance degradation. This approach to imputation is made by modeling the SNR characteristics using the collected dataset and further augmenting it with synthetic data.
- ML models are trained using static data and cannot learn the robot's dynamics during movement. Due to the robot's motion, the ML output is susceptible to run-time mispredictions. To overcome these challenges, KF is designed and integrated with the ML model. KF learns the motion behavior of the robot and combines it with the ML position output to provide highly accurate and smooth run-time tracking of the robot. With KF integration overhead, our system still achieves real-time performance at inference.

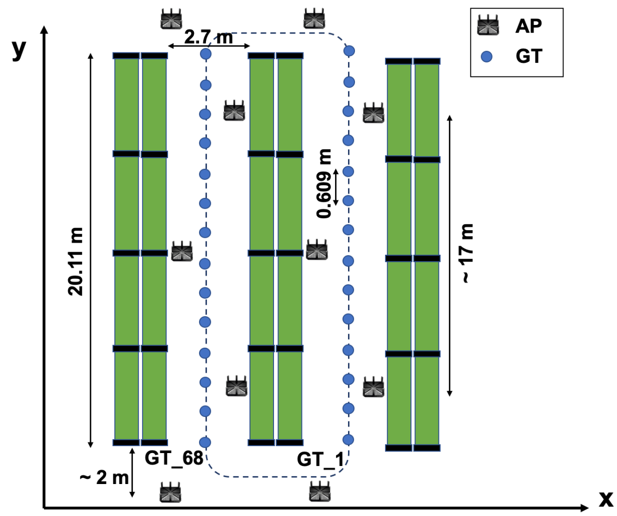

Fig. 1. Warehouse layout with two aisles for our system evaluation

II. Related Work

Indoor wireless localization at mmWave frequencies has been a very active research area in literature and industries recently. One of the major contributing factors is the transition of various industrial sectors towards automation, which is regarded as the next generation of the industrial revolution, Industry 4.0. This section discusses the recent work done in indoor wireless localization using mmWave wireless technology using machine learning and filtering approaches.For localization in indoor environments such as warehouses, wireless sensor-based approaches are vastly investigated [12]–[15] and have been the preferred positioning approach for the indoor environment due to the low cost, easy deployment, and power efficiency. Wireless localization techniques can be broadly classified into two technologies. The first technique uses the channel propagation model to estimate the distance to the Client using the signal strength information and then using the known location positions of the APs and the distances, trilateration is used to predict the location of the Client. The second technique estimates the position of the Client by matching the known signal strength from the APs. This is done by collecting the signal strength information at many different locations within the environment and then using it as a database for matching. This technique is known as fingerprinting.

In [16], the authors present a moving average based k-Nearest neighbor approach for the fingerprint-based localization. The accuracy achieved by their system is low. Machine learning and deep learning-based techniques have been investigated for fingerprint-based localization in recent years to provide high-accuracy localization prediction. Authors in [17], [18] designed support vector regression for localization using the received signal strength indicator (RSSI) fingerprinting as input features. In [19] authors proposed to use AP

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/MCE.2021.3101060, IEEE Consumer Electronics Magazine

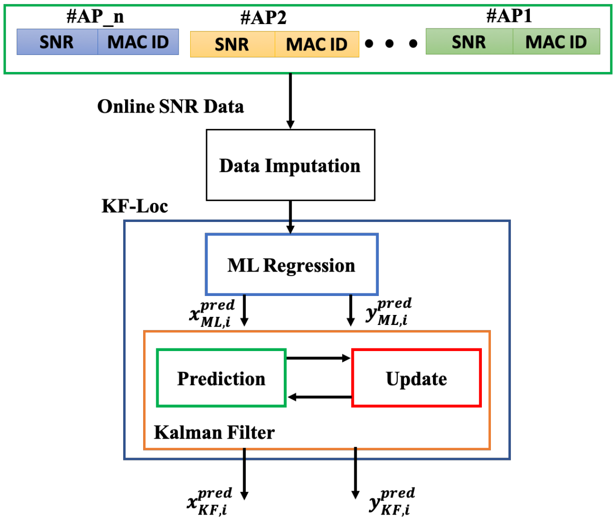

Fig. 2. Proposed kalman filter integrated machine learning architecture

#MAC_1			...	#MAC_n			X	Y
S_1	...	S_36	...	S_1	...	S_36		
0	...	15	...	15	...	10	X1	Y1
0	...	0	...	0	...	0	X2	Y2
...	...	...	...	...	...	...	...	...
12	...	17	...	20	...	19	Xm	Ym

#MAC_1			...	#MAC_n			X	Y
S_1	...	S_36	...	S_1	...	S_36		
$Mean^1_1$	...	15	...	15	...	10	X1	Y1
$Mean^1_2$	...	$Mean^{36}_2$	...	20	...	24	X2	Y2
...	...	...	...	...	...	...	...	...
12	...	17	...	$Mean^1_m$	...	$Mean^{36}_m$	Xm	Ym

Fig. 3. Output data from Imputation Unit

selection and data filtering approach to localize in an indoor environment. They show the improvement due to EM-based filter and feature matching using the Bayesian network. Authors in [20] have presented a deep reinforcement learning framework for indoor localization using Bluetooth devices. The idea is to mitigate the data collection step for training the ML model and use reinforcement learning algorithms to learn the location predictions. Their approach shows low accuracy in testing by achieving a Root Mean Square Error (RMSE) of 12.2m. In [21] authors have used the same mmWave routers in an indoor office environment and have shown an accuracy of 98.8%. Still, only 7 locations are used for the training and testing in their work, which is very coarse-grained localization performance. Further in [21] the task of localization is static as the robot is not in dynamic motion compared to our approach.

III. Localization Architecture

The machine learning predictions are static and can only provide output independent of time. In our system, the client is a moving robot inside a warehouse, illustrated in Fig. 1. Miss-predictions by ML model can negatively impact system autonomy. We enhance the localization capability by augmenting the ML location prediction with a Kalman filter to capture the time-dependent dynamics. KF uses inputs from various sensors, including the robot's state as provided by the odometry from the robotic platform. In our implementation of KF, we use the output from the ML prediction model instead of giving raw sensor output. We provide an integration where the KF utilizes the static location estimates from the trained ML model and infers the robot's time-dependent state by estimating velocity components in the 2D cartesian coordinate system. The system architecture is shown in Fig. 2. This final time-dependent output is used for the run-time tracking of the robot. Integrating KF significantly reduces the ML misprediction penalty that can result in wrong location estimates. Such miss-predictions in ML are caused due to obstructions and fluctuations of wireless signals at the client. KF-Loc provides highly accurate and robust tracking of an agent in a complex warehouse environment by utilizing a consumer-grade mmWave sensor.

A. Data Imputation and Augmentation

Wireless features are Signal-to-Noise Ratio (SNR) signals from APs as captured by the Client. The Client, for communication purposes, only selects the AP with the highest signal strength. We enable the Client to capture all available SNR signals by writing in-house firmware modification scripts. To build the dataset, we collect SNR features within the two aisles of the warehouse. The ML model uses the dataset to train network parameters during training stage. The wireless features are susceptible to obstacles and suffer from multipath reflections which can cause loss of connectivity with the receiver, which adds to the feature unreliability. Machine learning models will fail to generalize if the input features show such behavior. Further, there can be situations when the SNR signals from the sectors of APs are missing. The missing sector information and APs (features) are inconsistent and can cause the ML model to not generalize with high robustness for unseen test data, leading to severe performance degradation. To overcome the before-mentioned negative impact on performance, we introduce mean imputation and synthetic data generation technique as a pre-processing step before training the ML model.

During mean imputation, missing wireless features in the collected dataset are substituted with the respective mean value at the given location. Fig. 3 illustrates the imputation process where a zero value indicates the missed feature during data collection routine. The imputation step substitutes the missing features by their mean values. Imputation makes the dataset more consistent for the ML models, as ML models require input dimensions to be consistent throughout training and inference.

When considering the mean received SNR, we consider the log-normal channel model. The propagation model indicates that the log received signal power will decrease linearly with log distance and, superimposed, will present a random variation due to the large-scale shadow fading effect. This random variation can be modeled as a Gaussian random variable (when SNR or power are measured in the log scale of decibels). With fixed background noise power, the SNR measured in decibels can therefore be modeled as a linear decrease due to log-distance plus a random variation that follows a zero-mean Gaussian distribution. Hence, to create augmented dataset, we model the mean of SNR value from an AP j for sector i, $\mu^j_{i,k}$, and standard deviation $\sigma^j_{i,k}$ from the collected dataset at location k. This captures both the channel shadow fading as well as noise power. In this way, without collecting another set of training data we artificially create a new synthetic dataset and reduce the data collection time which can be fairly significant for large warehouses.

In this manner, we augment the on-site data with synthetic data and this enables us to improve the learning capability of the ML models and the robustness of the trained models when performing the test-time inference. In the combined dataset, we additionally introduce random drop-off of wireless SNR features by randomly selecting APs and the number of features to drop off. This is done to introduce more variability in the training dataset to capture the random signal fluctuations and scenarios where an AP can lose connection with the Client due to obstruction in the environment.

$$\hat{x_k} = Fx_{k-1} + Bu_{k-1} \tag{1}$$
$$\hat{P}_k = FP_{k-1}F^T + Q \tag{2}$$
$$y = z_k - H\hat{x_k} \tag{3}$$
$$S = H\hat{P}_kH^T + R \tag{4}$$
$$K = \hat{P}_kH^TS^{-1} \tag{5}$$
$$x_k = \hat{x_k} + Ky \tag{6}$$
$$P_k = (1 - KH)\hat{P}_k \tag{7}$$

Our training dataset is collected across different days and working hours. The rational behind collecting data across multiple days and multiple scans is to capture the random variations that occurs in the wireless signals across time and space. This routine makes the training data more feature rich and that significantly helps the ML models to generalize with high accuracy to unseen data during the test time. To test the trained models, another set of data is collected at same locations and during the testing, the ML models are fed only this test data and the model's output is the predicted location of the Client.

B. Machine Learning Prediction Model

We train a regression-based ML localization model. The regression approach uses the outputs as the 2D coordinates of the robot within the warehouse. Here, the outputs of the neural network are continuously valued with two output heads. Training ML regression model is done by considering both x and y positions as the GT in the training dataset. The dataset consists of X_{fet} number of input features, where X_{fet} consists of SNR from all the sectors. The GT for each training sample is the known position location at which data collection

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/MCE.2021.3101060, IEEE Consumer Electronics Magazine

Fig. 4. Warehouse aisle for experimental testbed

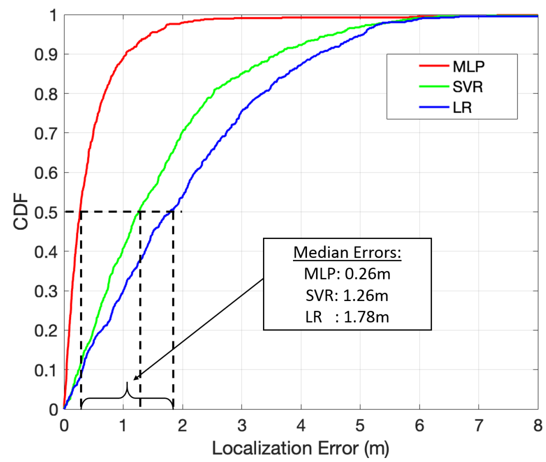

Fig. 5. Localization performance for different ML models

is performed. The input dimension of our dataset is $N \times X_{fet}$, where N is the number of training data points. The dimension of GT in the dataset are $N \times 2$, as for each training sample, we have a position in two-dimension space on the warehouse floor.

C. Kalman Filter Design

Kalman filter estimates the state of a robot based on the prediction-update cycle. KF is a linear recursive estimator that minimizes the mean square error of the estimated parameters. In Kf-Loc, input to the KF is noisy position prediction from the ML model, i.e., (x,y) coordinates of the robot. The KF filters the noisy coordinates to generate more accurate 2D tracking. At the start, different parameters of KF are initialized. These parameters include the 60 GHz sensor's noise covariance matrix (R), measurement matrix (H), robot's state vector (x), state transition matrix (F), and covariance matrix (P). We assume the process covariance (Q) to be zero as we perform a constant velocity-based linear tracking of the robot.

In the first step, the KF predicts the robot's state and error covariance for the next step; this is represented using (1), (2). B and u represent the control input matrix and control vector, respectively. Next, the KF performs the update process where based on the received measurements, predicted states and covariance estimates are updated. The update process is shown mathematically by (6), (7). During the prediction stage, the uncertainty of the robot's position increases as the agent gains no information. While during the update step, the agent gains information through the sensors' measurements and becomes more confident regarding its state.

KF considers the position probability of the robot to be Gaussian probability density function (PDF) which can be characterized using the mean and the standard deviation. The process starts by initializing the belief of the robot at the start time. For this, we take in our initial prediction output from the ML model. Next, using a motion model for the system, which is a constant velocity model in our design, we estimate the robot's location in the next time step. During the update-cycle of KF, after the time step has elapsed, we obtain the measurement reading from the ML model and update our previously predicted belief of the robot's state. It is to note that the robot's position, i.e., the state and the measurements, is modeled as Gaussian PDF. The tuning of R and Q parameters is done based on the sensor.

IV. Experimental Evaluation

In this section, we provide a detailed evaluation of the proposed KF localization framework. For the experimental testbed, we select a working warehouse [22] as shown in Fig. 4. Within the two aisles of the warehouse, we have mounted ten 60 GHz routers on the ceiling configured as APs. The placement of the routers is done in a zig-zag manner. On the robot, we place a 60 GHz router configured as Client. For APs and Client, we have used TP-Link AD7200 consumer grade routers. For the data collection, we collect the SNR data from all the

TABLE II
ML REGRESSION MODEL PERFORMANCE WITH DATA IMPUTATION

ML Model	Configuration	Augmentation	
		With	Without
LR	Linear	(0.51m, 2.6m)	(0.58m, 2.84m)
SVR	Polynomial Kernel	(0.71m, 2.17m)	(0.48m, 2.33m)
MLP	200, 200, 200	(0.15m, 0.91m)	(0.27m, 1.54m)

available APs. We collect the training data within the two aisles by moving the robot 0.609m (2-feet) in the y-dimension of the aisle and keeping the x-dimension fixed. For both the aisles, we collect two different training datasets. For test-time evaluation, another separate held-out dataset consisting of 674 test samples is collected. We assume the robot moves in one dimension, which is the y-dimension for localization and tracking. This is realistic in many practical warehouses for autonomous agents, where only one agent is allowed to move in a single-aisle at any given time [3].

A. Machine Learning Model Analysis

The optimized ML localization model is a multi-layer neural network with three hidden layers consisting of 200 neurons. The ML prediction is a regression model, where the output is the 2D prediction of the robot's position inside the warehouse. For model selection, we train and evaluate different regression models of varying complexity. The three models that we experiment with are Linear Regression (LR), Multi-Layer Perceptron (MLP), and Support Vector Regression (SVR). These ML models are considered as they represent varying complexity in terms of computation and learning capability. We train and optimize each regression model individually on the training dataset and use the held-out test dataset to evaluate the performance.

TABLE III
PERFORMANCE COMPARISON WITH DIFFERENT ML MODELS

ML regression	Configuration	RMSE-X	RMSE-Y
LR	Linear Model	0.51m	2.60m
SVR	Polynomial kernel	0.71m	2.17m
MLP	**200, 200, 200**	**0.15m**	**0.91m**

Table II presents the performance comparison between regression models trained with and without data augmentation. MLP achieves significant performance improvement, with RMSE of 0.19m and 0.92m in x-and y location prediction, respectively. This shows 29.3% and 40.25% improvement when compared with training without data augmentation. For the other two regression models, LR and SVR, performance improvement of 8.4% and 6.8% is achieved in y-dimension, respectively. This improvement in performance is seen due to robust training of the ML models with the synthetic data augmentation and mean imputation. Overall, the data augmentation approach that is introduced in our localization system benefits ML models. In the subsequent sections, we have used data augmentation during the training of ML models.

Table III illustrates the performance of the three ML regression models. It is seen that for LR model has poor performance with Root Mean Square Error (RMSE) in y-dimension of 2.6m. The results show that using a simple and less complex learning model, the complex behavior of the 60 mmWave wireless features with distances cannot be learned at high accuracy and requires more complex learning models. SVR and MLP, both complex models, achieves better performance compared to LR. MLP with three hidden layers and 200 neurons each achieves significantly higher accuracy with RMSE of 0.15m and 0.91m in the x and y position, respectively. Fig. 5 shows the localization error performance in terms of the Cumulative Distribution Function (CDF) plot, where three different regression models are compared. It is seen from Fig. 5 that MLP also achieves the lowest median error of 0.26m when compared to the SVR and LR regression models. Based on these performance metrics, we have selected the three-layer MLP as our optimized ML localization model.

The robot only moves in y-dimension within the aisle, but our ML model predicts both the x- and y-position of the robot. This is required as the x-position of the robot will be different for different aisles, and to differentiate the location of the robot within the multiple aisles, 2D-position output is

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/MCE.2021.3101060, IEEE Consumer Electronics Magazine

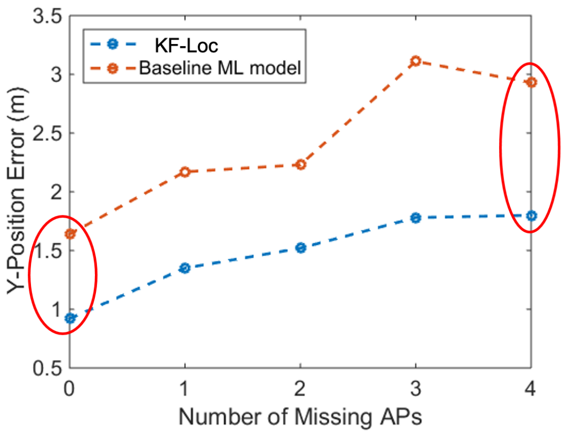

Fig. 6. Robustness performance of Multi-Loc system with baseline ML model for regression

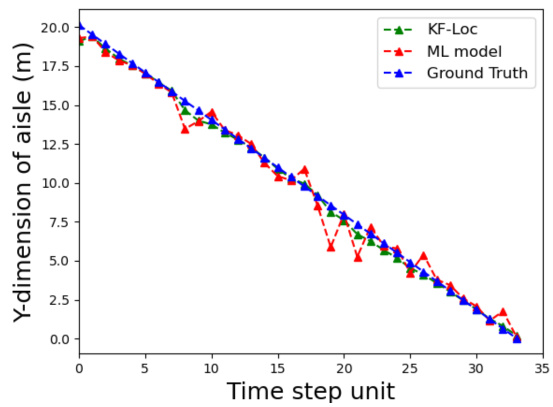

Fig. 8. Comparison of y-position between KF-Loc and ML model in aisle-2

required. Next, we will analyze how KF integration overcomes the fluctuations in predictions and increases localization accuracy.

B. Robustness with Data Augmentation

This section compares the performance of the KF-Loc system with ML localization model trained without data augmentation and KF integration. The ML model is the same MLP that we use in the KF-Loc. To evaluate the robustness performance, we randomly switch off data corresponding to APs, introducing missing or corrupt wireless signals in the test dataset, capturing the scenario where APs can lose connection with Client. Fig 6 illustrates the performance where the x-axis represents the number of missing APs, and the y-axis shows the corresponding RMSE in y-dimension.

We see the KF-Loc system maintains low positioning error compared to the ML-based system with RMSE of 1.98m and 2.93m, respectively. The significant decrease in performance for ML-based systems is due to the low robustness of the trained model. In comparison, KF-Loc maintains high tolerance to signal fluctuations due to the augmentation-based training described in Section III-A. Data augmentation makes the localization system highly robust against random fluctuations and loss of connectivity with the wireless signals. Such scenarios are common in real-world deployment using ML based localization systems, and the ability to generalize with high robustness is a critical requirement. Such robustness performance analysis is missing in related state-of-the-art localization models presented in Table IV.

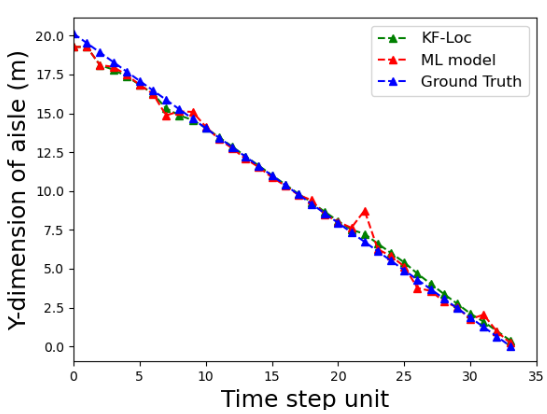

Fig. 7. Comparison of y-position between KF-Loc and ML model in aisle-1

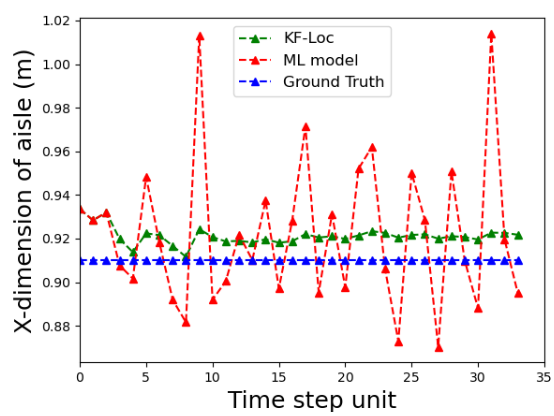

Fig. 9. Comparison of x-position between KF-Loc and ML model in aisle-1

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/MCE.2021.3101060, IEEE Consumer Electronics Magazine

TABLE IV
PERFORMANCE COMPARISON WITH DIFFERENT WIRELESS BASED LOCALIZATION TECHNIQUES

Work	Wireless	Frequency	Environment	Methodology	Performance
Bahl [23]	RF-based	2.4 GHz	Indoor	KNN	2m-3m
Laoudias [24]	WiFi	2.4 GHz	Indoor	ANN	Mean error of 3.4m
Yang [25]	WiFi	2.4 GHz	Indoor	WiFi Fingerprinting	Mean error of 5.88m
Kanhere [26]	mmWave	28 GHz	Indoor	Fusion of AoA and received power	Mean error of 1.86m
Kanhere [26]	mmWave	28 GHz	Outdoor	Fusion of AoA and received power	Mean error of 34m
Bielsa [27]	mmWave	60 GHz	Indoor	Particle filter	Median error of 1.1m to 1.4m
Wei [28]	mmWave	60 GHz	Outdoor	DoA based WKNN fingerprint	Mean error 1.32m
Walaa [29]	mmWave	28 GHz	Outdoor	Different ML models	MAE of 3m-33m
Li [30]	WiFi	5 GHz	Indoor	Particle Swarm Optimization	Median error 1.5m
Vashist [31]	mmWave	60 GHz	Indoor	MLP fingerprint	RMSE 0.84m
Our work	**mmWave**	**60 GHz**	**Indoor**	**KF and ML integrated (KF-Loc)**	**RMSE of 0.35m and 0.37m**

C. Kalman Filter Integrated Tracking

For KF system design, we define the robot state as a three-dimension vector defining the position in 2D Cartesian coordinate frame and its velocity in the y-dimension. The covariance matrix, P, is a 3x3 matrix initialized with very high uncertainty for velocity. The measurement matrix, H, is a 2x3 matrix, initialized with 1 for the position, and the measurement covariance, R, is a 2x2 matrix initialized with values 0.01. The ML prediction output is integrated with the KF to provide the 2D tracking. The optimized model designed in section IV-A is used as the ML model, which is MLP consisting of three hidden layers with 200 neurons each.

We evaluate the tracking of the robot in two different aisles of the warehouse. Figs. 7 and 8 show the performance comparison between the KF-Loc, in green, and ML model, in red, tracking in y-position for aisle-1 and aisle-2 respectively. The GT position is shown in blue for both figures. The figure shows fluctuations in the position estimates for the ML model with agent's motion. These fluctuations are due to mispredictions by the ML model. Signal fluctuations due to agent motion cause mispredictions.

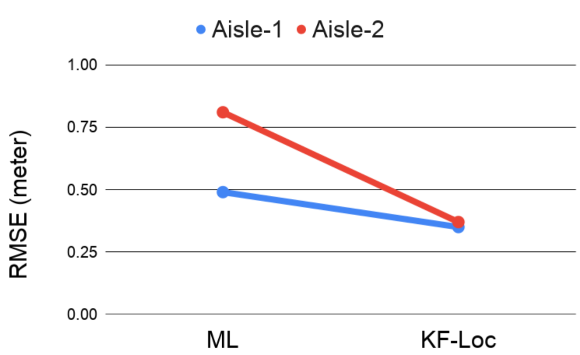

Fig. 10. Error comparison between KF-Loc and ML model in both the aisles

Similarly, Fig. 9 shows the position estimates of the robot along the x-position within the aisle. For location estimation, the KF-Loc provides a smoother localization and tracking performance. This is achieved as the KF within the KF-Loc system can filter and smooth out the raw position prediction from the ML model and capture the robot's dynamic motion by estimating its velocity. This is where the KF provides powerful integration with the ML model by estimating the robot's dynamic state variable, in our case, the velocity, without explicitly being provided with velocity output from the robot.

Performance comparison between the KF-Loc and ML based localization model is made by measuring the Root Mean Square Error (RMSE) of the robot in both the aisles individually. Figure 10 plots the robot's RMSE in y-position for both aisle-1 and aisle-2. For aisle-1, performance improvement with 28.5% reduction in RMSE error is achieved, and for aisle-2, we see more significant performance improvement with 54.3% reduction in RMSE error. Both the aisles KF-Loc achieve centimeter-level localization accuracy with RMSE of 0.35m and 0.37m, respectively. KF-Loc reduces the negative

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/MCE.2021.3101060, IEEE Consumer Electronics Magazine

impact of ML mispredictions by learning the motion dynamics of the robot.

D. Comparison and Overhead Analysis

In this section, we compare the performance of KF-Loc with different wireless-based localization systems. Table IV compares KF-Loc with different state-of-the-art approaches using millimeter wave and sub 5GHz features. From Table IV we see that mmWave approaches show better performance compared to sub-GHz wireless systems. This is due to the shorter wavelength of the mmWave results in a high-resolution radio map. ML models use this feature-rich radio map and result in more accurate modeling. Authors in [23]–[25] uses lower 2.4GHz frequency and provides estimation using K-nearest neighbor [23], artificial neural network (ANN) [24], and fingerprinting matching [25] based approaches. The results presented showed sub-meter level accuracy. As these approaches use relatively simpler models and sub-GHz wireless features, such systems' performance is lower than the KF-Loc localization approach.

Work presented in [26]–[28], [31] utilizes millimeter-wave frequencies varying from 28GHz to 60GHz. They achieve meter-level accuracy in localization. In [26] Angle of Arrival (AoA) approach is used for position estimation of the client. Such approaches to localization require precise knowledge of AP placement. Such precise placement is hard to achieve, and the arrangement directly impacts performance. ML approaches do not require knowledge of AP location and are more robust to change in the environment. Authors in [28] utilizes fingerprint-based approach similar to ours. In [28] the estimation is performed using weighted K-nearest neighbor; KNNs are computationally inexpensive but suffer from lower performance and are highly susceptible to noise and signal variations.

Authors in [29] use 28GHz wireless with fingerprinting approach and evaluate 13 different ML models. They achieve meter-level accuracy in an outdoor simulation environment. Further, the results presented are simulation-based. Li [30] uses 5GHz WiFi-based location estimation using Particle Swarm Optimization (PSO) approach. All mentioned approaches do not show and discuss the impact of missing features or signal loss on the localization system. Such scenarios are common, as ML-based methods are data-dependent, and such signal fluctuations or AP failure can cause the system to fail. In our approach, with the introduction of data imputation and synthetic data for training, we show the high robustness of the KF-Loc system under these scenarios.

With the integration of KF, additional timing overhead is introduced. In our design, the KF consumes a total of 0.64ms for each input processing. When compared to the run-time performance, the system does not contribute to significant overhead. Further, the ML model is an MLP consisting of two intermediate fully connected layers that do not have additional computational complexity compared to deeper CNN architectures. Additionally, the agent in use comprises a computing system with GPU. The ML model utilizes the GPU parallelism to provide a processing rate of 4ms. The total processing delay of the KF-Loc is 4.64ms, which is significantly lower and can easily provide the required real-time performance. Localization and ML systems designed for autonomy should have latency in the range of a few hundred milliseconds as surveyed by Google, and Mercedes Benz for self-driving systems [32]. In [32] the evaluated reaction time is estimated around 0.83s and Kf-Loc latency is only 4.64ms, which is significantly lower than desired latency.

V. Conclusion

We present the design and implementation of a robust localization system, *KF-Loc* for indoor warehouses using 60 GHz routers. We introduce the use of the consumer-grade 60 GHz wireless routers for providing high accuracy localization performance using off-the-shelf mmWave routers. In our system, complex mmWave features are learned by a regression-based ML model, providing static position predictions of the robot. Kalman filter is designed to improve the ML prediction error during the robot motion. KF improves the motion tracking of the robot by removing the fluctuations due to mispredictions in ML output. ML models are susceptible to fluctuations in input features causing severe performance degradation. To address robustness, we present a data imputation and augmentation for wireless features. Our results show for the worst-case AP dropout, KF-Loc achieves 1.4X less degradation in localization performance. To test the practicality of our system, we deploy and test our system within two aisles in a functional warehouse. KF-Loc achieves centimeter-level accuracy in our test setup of two aisles with RMSE of 0.35m and 0.37m, respectively. Further, compared with the static ML localization system, our proposed system shows significant performance improvement by achieving 28.5% and 54.3% improvement in RMSE error for the two test aisles.

REFERENCES

[1] D. Gorecky, M. Schmitt, M. Loskyll, and D. Zühlke, "Human-machine-interaction in the industry 4.0 era," in *IEEE Int. Conf. on Industrial Informatics (INDIN)*, 2014.

[2] A. Diez-Olivan et al., "Data fusion and machine learning for industrial prognosis: Trends and perspectives towards industry 4.0," *Information Fusion*, vol. 50, pp. 92–111, 2019.

[3] T. Le-Anh and R. de Koster, "A review of design and control of automated guided vehicle systems," ERIM at Erasmus University Rotterdam, Tech. Rep. ERS;2004-030-LIS, 2004.

[4] C. Li, B. Dai, and T. Wu, "Vision-based precision vehicle localization in urban environments," in *2013 Chinese Automation Congress*, Nov 2013, pp. 599–604.

[5] A. Y. Hata and D. F. Wolf, "Feature detection for vehicle localization in urban environments using a multilayer LIDAR," *IEEE Trans. on Intelligent Transportation Systems*, vol. 17, no. 2, pp. 420–429, Feb 2016.

[6] S. Krishnan, P. Sharma, Z. Guoping, and O. H. Woon, "A uwb based localization system for indoor robot navigation," in *2007 IEEE International Conference on Ultra-Wideband*, 2007, pp. 77–82.

[7] S. J. Velat, J. Lee, N. Johnson, and C. D. Crane, "Vision based vehicle localization for autonomous navigation," in *2007 International Symposium on Computational Intelligence in Robotics and Automation*, 2007, pp. 528–533.

[8] V. K. Kukkala, J. Tunnell, S. Pasricha, and T. Bradley, "Advanced driver-assistance systems: A path toward autonomous vehicles," *IEEE Consumer Electronics Magazine*, vol. 7, no. 5, pp. 18–25, 2018.

[9] J. Lemley, S. Bazrafkan, and P. Corcoran, "Deep learning for consumer devices and services: Pushing the limits for machine learning, artificial intelligence, and computer vision." *IEEE Consumer Electronics Magazine*, vol. 6, no. 2, pp. 48–56, 2017.

[10] T. S. Rappaport, Y. Xing, G. R. MacCartney, A. F. Molisch, E. Mellios, and J. Zhang, "Overview of millimeter wave communications for fifth-generation (5g) wireless networks—with a focus on propagation models," *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 12, pp. 6213–6230, Dec 2017.

[11] L. Perez and J. Wang, "The effectiveness of data augmentation in image classification using deep learning," *CoRR*, vol. abs/1712.04621, 2017. [Online]. Available: http://arxiv.org/abs/1712.04621

[12] Z. Farid, R. Nordin, and M. Ismail, "Recent advances in wireless indoor localization techniques and system," *Journal Comp. Netw. and Communic.*, vol. 2013, pp. 185 138:1–185 138:12, 2013.

[13] X. Guo, L. Li, N. Ansari, and B. Liao, "Accurate wifi localization by fusing a group of fingerprints via a global fusion profile," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 8, pp. 7314–7325, Aug 2018.

[14] D. Liang, Z. Zhang, and M. Peng, "Access point reselection and adaptive cluster splitting-based indoor localization in wireless local area networks," *IEEE Internet of Things Journal*, vol. 2, no. 6, pp. 573–585, Dec 2015.

[15] L. Chen, K. Yang, and X. Wang, "Robust cooperative wifi fingerprint-based indoor localization," *IEEE Internet of Things Journal*, vol. 3, no. 6, pp. 1406–1417, Dec 2016.

[16] M.-S. Choi and B. Jang, "An accurate fingerprinting based indoor positioning algorithm," 2017.

[17] M. Brunato and R. Battiti, "Statistical learning theory for location fingerprinting in wireless lans," *Computer Networks*, vol. 47, no. 6, pp. 825 – 845, 2005.

[18] Z. Wu, C. Li, J. K. Ng, and K. R. p. h. Leung, "Location estimation via support vector regression," *IEEE Transactions on Mobile Computing*, vol. 6, no. 3, pp. 311–321, 2007.

[19] F. Zhao, T. Huang, and D. Wang, "A probabilistic approach for wifi fingerprint localization in severely dynamic indoor environments," *IEEE Access*, vol. 7, pp. 116 348–116 357, 2019.

[20] Y. Li, X. Hu, Y. Zhuang, Z. Gao, P. Zhang, and N. El-Sheimy, "Deep reinforcement learning (drl): Another perspective for unsupervised wireless localization," *IEEE Internet of Things Journal*, pp. 1–1, 2019.

[21] T. Koike-Akino, P. Wang, M. Pajovic, H. Sun, and P. V. Orlik, "Fingerprinting-based indoor localization with commercial mmwave wifi: A deep learning approach," *IEEE Access*, Apr. 2020. [Online]. Available: https://www.merl.com/publications/TR2020-054

[22] "Simcona electronics corporation," https://www.simcona.com/, accessed: 2010-09-13.

[23] P. Bahl and V. N. Padmanabhan, "RADAR: an in-building rf-based user location and tracking system," in *IEEE INFOCOM*, 2000.

[24] C. Laoudias, P. Kemppi, and C. G. Panayiotou, "Localization using radial basis function networks and signal strength fingerprints in wlan," in *GLOBECOM 2009*, Nov 2009, pp. 1–6.

[25] Z. Yang, C. Wu, and Y. Liu, "Locating in fingerprint space: Wireless indoor localization with little human intervention," in *International Conference on Mobile Computing and Networking*, 2012.

[26] O. Kanhere and T. S. Rappaport, "Position locationing for millimeter wave systems," *CoRR*, 2018.

[27] G. B. et al., "Indoor localization using commercial off-the-shelf 60 ghz access points," in *IEEE INFOCOM 2018 - IEEE Conference on Computer Communications*, April 2018, pp. 2384–2392.

[28] Z. Wei, Y. Zhao, X. Liu, and Z. Feng, "Doa-lf: A location fingerprint positioning algorithm with millimeter-wave," *IEEE Access*, 2017.

[29] W. Y. Al-Rashdan and A. Tahat, "A comparative performance evaluation of machine learning algorithms for

fingerprinting based localization in dm-mimo wireless systems relying on big data techniques," *IEEE Access*, vol. 8, pp. 109 522–109 534, 2020.

[30] Z. Li, Z. Tian, Z. Wang, and Z. Zhang, "Multipath-assisted indoor localization using a single receiver," *IEEE Sensors Journal*, vol. 21, no. 1, pp. 692–705, 2021.

[31] A. Vashist, D. R. Bhanushali, R. Relyea, C. Hochgraf, A. Ganguly, P. D. Sai Manoj, R. Ptucha, A. Kwasinski, and M. E. Kuhl, "Indoor wireless localization using consumer-grade 60 ghz equipment with machine learning for intelligent material handling," in *2020 IEEE International Conference on Consumer Electronics (ICCE)*, 2020, pp. 1–6.

[32] V. V. Dixit, S. Chand, and D. J. Nair, "Autonomous vehicles: Disengagements, accidents and reaction times," *PLOS ONE*, vol. 11, no. 12, pp. 1–14, 12 2016.

Abhishek Vashist is currently pursuing his Ph.D. degree in the Department of Computer Engineering at Rochester Institute of Technology, Rochester, NY, USA. His research interest includes millimeter wave communication, 60 GHz, autonomous vehicles, machine learning, and localization systems. His work has also won the best paper award at ICCE 2020. Contact him at av8911@rit.edu.

Maojia Patrick Li is currently pursuing his Ph.D. degree at Rochester Institute of Technology, Rochester, NY, USA. His research interest includes discrete-event simulations, automated material handling systems, and deep learning. Contact him at mxl8487@rit.edu.

Amlan Ganguly is currently an Associate Professor in the Department of Computer Engineering at Rochester Institute of Technology. He received his PhD from Washington State University, USA in 2005. His research interests are scalable interconnection architectures and novel datacenter networks. Contact him at axgeec@rit.edu.

Sai Manoj P D is an assistant professor at George Mason University. He received his Ph.D. in Electrical and Electronics Engineering from Nanyang Technological University, Singapore, 2015. His research interests include adversarial machine learning, self-aware SoC design. His works won the best paper award in ICDM and ICCE. Contact him at spudukot@gmu.edu.

Clark Hochgraf is an Associate Professor in the Department of Electrical, Computer, and Telecommunications Engineering Technology at Rochester Institute of Technology. He earned his Ph.D. in Electrical Engineering from the University of Wisconsin-Madison in 1997. His areas of research interest include intelligent vehicles, power conversion, and education of engineers. Contact him at cghiee@rit.edu.

Raymond Ptucha is an Assistant Professor in Computer Engineering and the Director of the Machine Intelligence Laboratory at the Rochester Institute of Technology. His research includes machine learning, computer vision, and robotics, with a specialization in deep learning. Contact him at rwpeec@rit.edu

Andres Kwasinski received his Ph.D. degree in Electrical and Computer Engineering from the University of Maryland at College Park, 2004. He is currently a Professor at Rochester Institute of Technology. Contact him at axkeec@rit.edu.

Michael E. Kuhl is a Professor in the Department of Industrial and Systems Engineering at Rochester Institute of Technology. He earned his Ph.D. in Industrial Engineering from North Carolina State University in 1997. His areas of research interest including simulation modeling and analysis. Contact him at mekeie@rit.edu.