


RESEARCH

Open Access

Using vision-based object detection for link quality prediction in 5.6-GHz channel



Riichi Kudo , Kahoko Takahashi, Takeru Inoue and Kohei Mizuno

* Correspondence: riichi.kudou.cw@hco.ntt.co.jp
NTT Network Innovation
Laboratories, Hikarinooka,
Yokosuka-Shi, Kanagawa 239-0847,
Japan

Abstract

Various smart connected devices are emerging like automated driving cars, autonomous robots, and remote-controlled construction vehicles. These devices have vision systems to conduct their operations without collision. Machine vision technology is becoming more accessible to perceive self-position and/or the surrounding environment thanks to the great advances in deep learning technologies. The accurate perception information of these smart connected devices makes it possible to predict wireless link quality (LQ). This paper proposes an LQ prediction scheme that applies machine learning to HD camera output to forecast the influence of surrounding mobile objects on LQ. The proposed scheme utilizes object detection based on deep learning and learns the relationship between the detected object position information and the LQ. Outdoor experiments show that LQ prediction proposal can well predict the throughput for around 1 s into the future in a 5.6-GHz wireless LAN channel.

Keywords: Link quality prediction, Machine learning, Object detection, Wireless LAN

1 Introduction

Connected devices are improving our lives, and their roles are expanding with the rapid enhancement of wireless communication systems [1]. Some connected devices being developed operate autonomously such as self-driving cars, patrol robots, transportation vehicles, and remote-controlled construction machines. These autonomous operation devices need to perceive the surrounding environment by using their vision systems to conduct their missions without trouble such as colliding with things. Advances in deep learning technologies enable the connected devices to perceive the environment information by using their cameras and sensors [2, 3]. The accurate environment information thus created for is also useful for enhancing wireless communications.

The enhancement of wireless communications is also being accelerated by demands of greater wireless link capacities, latencies, and reliabilities [4]. The frequency of the wireless communications is expanded to super high frequency (SHF) band (3~30 GHz) and extremely high frequency (EHF) (30 GHz~3 THz) to obtain larger capacity [5–7]. As the radio frequency increases, the influence of the surrounding environment



© The Author(s). 2020 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

strengthens [8]. To counter the wireless link quality (LQ) variations caused by the surrounding environment changes, LQ prediction plays an important role.

LQ prediction methods have been proposed to guarantee the quality of service (QoS) in wireless communication systems. Machine learning-based LQ prediction has been widely studied as surveyed in [9]. The network performance prediction and related deep learning technologies for solving mobile networking problems are summarized in [10], and the case studies and intelligent decision-making use of machine learning are described in [11]. The prediction of channel state information (CSI) was proposed in [10–14]. CSI is predicted by the position of the terminal, temperature, humidity, and weather in [12]. The prediction scheme in [13] enables CSI estimation with minimal pilot overhead. The CSI yielded by massive numbers of antennas was used to predict the channel statistical characteristics in millimeter wave (mmWave) environments in [14]. Xu et al. [15] showed that the future network performance could be predicted by using the appropriate metrics in major cellular networks. Wei et al. [16] focused on the transportation mode of the holder of the smart phone and predicted throughput by using the moving history. The LQ related to autonomous connected devices has been studied for connected vehicles [17] and unmanned aerial vehicles (UAVs) [18]. Yang et al. [17] proposed resource management to realize ultra-reliable and low-latency wireless communication for connected vehicles, and Almeida et al. [18] proposed a quality of service estimator using UAV base station positions and user traffic demand. Since these works do not consider the influence of the surrounding objects, the LQ change caused by large mobile objects such as trucks cannot be predicted.

To consider surrounding mobile objects, LQ prediction based on cameras and sensors was studied in [8, 19–21]. RadMAC in [8] used radar to detect human obstacles and switched the beam pattern of 60-GHz channel in a 4.65 m × 5.95 m room, and Wang et al. [19] proposed the mmWave beam prediction by using receiver location and surrounding vehicles. Papers [20, 21] proposed machine learning-based LQ prediction using depth cameras in a 60-GHz channel, and the LQ degradation created by human body blocking was predicted for a transmitter and receiver pair separated by 4 m. Although the impact of mobile objects has been studied in UHF channels, the mobile objects impact LQ even in microwave channels such as SHF. Since SHF has wider service areas than UHF band, the vision system must recognize mobile objects in wider regions.

Object detection technologies have been advanced significantly thanks to deep neural networks [22, 23]. The precision accuracy and inference speed of object detection have been improving every year, and many significant reports are now available. The typical object detection output is the object bounding-box with classification result. The object bounding-box indicates the position and size of a rectangular area delineating an object in the image. Classifications are derived from annotated images, and some detection algorithms also output confidence scores.

Autonomous connected devices have driven advances in object detection performance. Therefore, we present an LQ prediction scheme that uses the vision-based advanced object detection methods. The LQ prediction proposal uses object bounding-boxes and their classification provided by advanced object detection algorithms. Experiments using HD cameras and wireless LAN systems are conducted in an outdoor environment, and LQ prediction performance is evaluated by using the relationship between

the object bounding-box information and the throughput of SHF channels (5.6 GHz). The LQ prediction proposal consists of two-step machine learning: the first step is object detection using deep learning and the second step uses random forest regression to predict the future LQ.

The key contribution of this paper is to introduce the two-step LQ prediction that can take advantage of subsequent advances in object detection algorithms and yield explainable AI results by using the bounding-box information. It is confirmed that vision-based LQ prediction can accurately predict the throughput 1 s into the future in low-SHF channels (5.6 GHz) by using HD camera images captured in an outdoor environment in which various moving objects were present; the distance between the transmitter and receiver is 42 m.

The rest of this paper is organized as follows. Section 2 describes the target region of the vision-based LQ prediction and the system model. Section 3 details the LQ prediction process. Section 4 shows the outdoor experiment setup. Section 5 details the performance evaluation results and the discussions, and Section 6 concludes this paper.

2 System model

To satisfy the various requirements posed by wireless access services, LQ variation should be predicted and countered as needed. LQ prediction is one foundation technology of advanced wireless access management. The base station (BS) communicates with the terminal-connected devices by wireless access, such as IEEE 802.11, long-term evolution (LTE), and 5G. The LQ of the wireless access is determined by various factors, which are categorized into two parts: communication network and radio-wave propagation condition. The communication network factors include the transmission power, transmission beamforming, traffic, interference, modulation scheme, and error correction code. Their relationship with LQ has been well studied, and LQ can be improved by advanced signal processing. The other aspect is the radio-wave propagation condition between the transmitter and receiver. We separate the radio-wave propagation condition into three categories: surrounding mobile objects, static environment, and connected device status. The factors of static environment and connected device status such as position have been studied by papers [24, 25]. The remaining factor, the influence of the surrounding mobile object, was the last piece to achieve advanced wireless access management based on the accurate LQ prediction.

The target radio frequency is important in considering the LQ prediction. Oguma et al. [20] proposed LQ prediction for the millimeter wave communications by using depth cameras. However, LQ prediction for the wireless systems in SHF is needed because the major wireless systems are being operated in SHF. In SHF wireless systems, the service area is likely to be greater than that of the millimeter wave wireless systems and the influence of the mobile object varies widely depending on the size, movement, and type of the mobile objects. The object type denotes the category of the object such as car, truck, and person. The object detection algorithms that use the vision information obtained by the cameras and sensors are promising to provide accurate information of the surrounding mobile objects. The state-of-art object detection achieves accurate and real-time operations. In this paper, we used the leading object detection algorithms, M2Det [26] and YOLO v3 [27], and these algorithms can process each frame in less than 100 ms.

Since the movement of a physical object does not change in periods of the order of milliseconds, the vision-based object detection is expected to be used to predict LQ with lead time of around 1 s. Such a lead time allows negative changes in LQ to be countered effectively. The connected device switches to an expensive but more robust wireless link only when the predicted LQ degradation is excessive. Since Lauridsen et al. [28] showed that when transferring from the idle state to connected state in LTE networks the round trip time of ping packets can be several hundreds of milliseconds, long-term LQ prediction is attractive. In the other approach, the data-rate of the video streaming service is decreased to avoid fatal errors such as monitoring video freeze. Furthermore, this might, in combination with position information, yield enhanced movement control of the autonomous robot to optimize LQ.

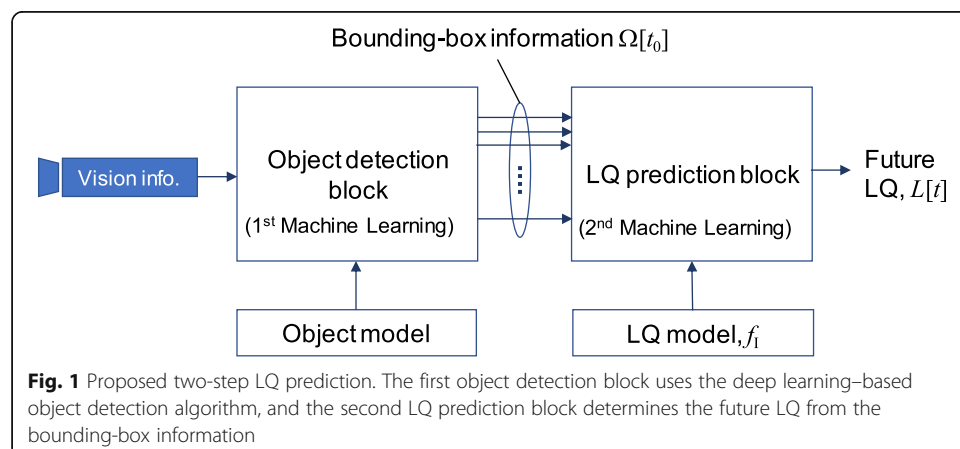
In this paper, we consider that the wireless environments around connected devices are recognized using the images obtained by cameras and sensors, and the LQ of wireless communication can be estimated using the recognized environmental information. Thus, we define our problem as accurately estimating the LQ about 1 s into the future by using past images taken by cameras. To simplify the scenario, we use a fixed transmitter and a fixed receiver with a dedicated wireless channel between them. In this system model, the wireless channel is disturbed by mobile objects and the objects are found in the camera images. The target LQ at timing t are taken to be $L[t]$. The camera image images, $\Omega[t_0]$, are obtained from the images at the current timing t_0 and past timings. The relationship between target LQ and the image features is denoted by introducing function f_1 as follows,

$$L[t] = f_1(\Omega[t_0]) \quad (1)$$

From the viewpoint of machine learning, our problem is to construct function f_1 , given training dataset $(L[t], \Omega[t_0])$.

3 LQ prediction method

Figure 1 shows the structure of the proposed LQ prediction. The proposal uses two-step machine learning: the first step realizes the object detection, and the second step predicts the LQ using the bounding-box information, which consists of the object category, position, and size in the processed image. The object detection block provides the bounding-box information of the surrounding objects by using vision information



from cameras and pre-trained object models. The second machine learning block predicts the future LQ, $L[t]$, using the LQ model and the bounding-box information. The LQ model and the bounding-box information are the function f_l and camera image features $\Omega[t_0]$ in Eq. 1. The object model for the first block and the LQ model for the second block were pre-trained by using the MS COCO dataset [29] and the measured dataset of $(L[t], \Omega[t_0])$, respectively. In this paper, random forest regression is used to implement the second machine learning.

The benefits of the proposal are that it allows us to take advantage of subsequent advances in object detection algorithms and that we can understand what condition alters the LQ in the environment. The detection precision of object prediction using camera images is improving continually [23], and novel object detection schemes are expected to emerge in the future. Since the object bounding-box is used to measure object detection performance [23], it is expected that the future object detection algorithms will also provide the object bounding-box information and this approach allows us to benefit from future enhancements in object detection. LQ prediction must have explainable features to encourage the development of the technologies. By using the object bounding-box information, what condition impacts the LQ can be evaluated. Furthermore, we can separately consider the performance of the object detection block and the LQ prediction block.

3.1 LQ definition

This paper takes normalized throughput as LQ to focus on the LQ variation caused by the surrounding mobile objects. The downlink throughput was measured using User Datagram Protocol (UDP) traffic with full buffer condition. The throughput at timing t_i , $R[t_i]$, was obtained every ΔT , ($t_i - t_{i-1} = \Delta T$) as the bit rate from t_{i-1} to t_i ,

$$R[t_i] = \frac{1}{\Delta T} \sum_{t_i - \Delta T < t \leq t_i} B[t], \quad (2)$$

where $B[t]$ is the bit amount of the UDP packets successfully received at timing t . To focus on dynamic throughput changes, the normalized throughput, $\tilde{R}[t_i]$, is defined as

$$\tilde{R}[t_i] = \frac{R[t_i]}{\text{Median}(R[t])_{t_i - T_A \leq t < t_i}}, \quad (3)$$

where Median() denotes the function that calculates the median value and T_A is the averaging time window (set to be greater than ΔT). Since the measured throughput contains extremely low values but with at very low probability, we adopt the median value instead of averaging to alleviate the influence of outliers. $\tilde{R}[t_i]$ ($i > 0$) is used as the LQ, $L[t]$, in Eq. 1. To consider the LQ prediction performance corresponding to the difference between the target LQ timing and the current timing, the lead time, T_F , is defined as $T_F = t_i - t_0$. In this paper, the time interval for the normalized throughput, ΔT , and the averaging time window, T_A , are set to 0.5 [s] and 30.0 [s], respectively. Thus, the target normalized throughputs with the lead time, T_F , of 1.0 and 1.5 [s] are given by $\tilde{R}[t_2]$ and $\tilde{R}[t_3]$, respectively.

3.2 Object detection

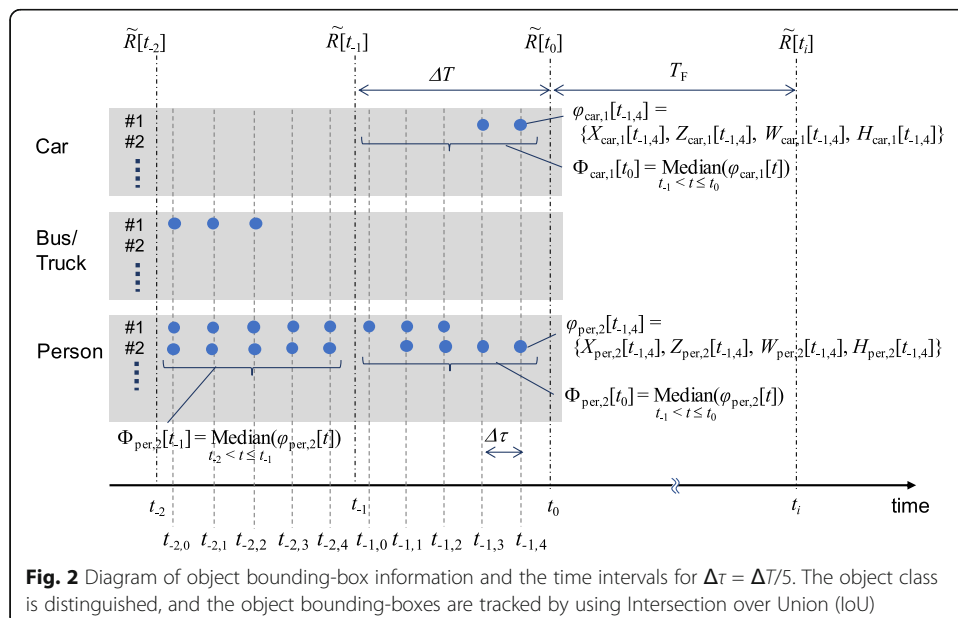
The first machine learning block outputs detected object categories and object bounding-box information. Figure 2 shows the generation of object bounding-box information in the time domain. The object bounding-box acquisition timing is assumed to be asynchronous to LQ acquisition timing. Thus, term $t_{n,m}$ is defined as the m -th object bounding-box acquisition timing in the time window from t_n to t_{n+1} . The time interval of object bounding-box acquisition, $\Delta\tau = t_{n,m} - t_{n,m-1}$, is set to 0.1 [s].

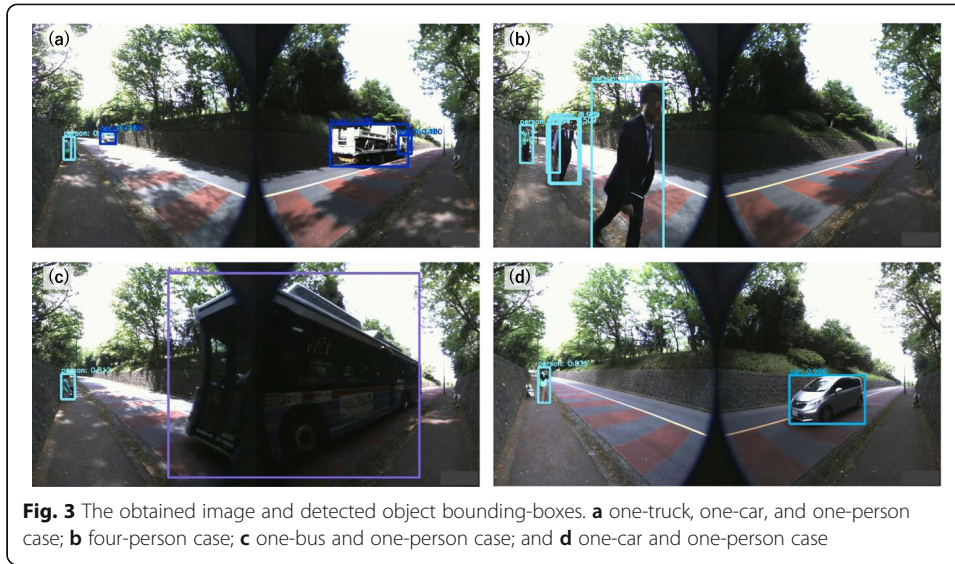
The object bounding-box information consists of object category, detection reliability, position, and size. Figure 3 shows an example of the object bounding-box data as gathered in the outdoor experiment. We can see that mobile objects are detected by rectangular bounding-boxes. The positions and sizes were obtained as

$$\phi_{\chi,j}[t_{n,m}] = \{X_{\chi,j}[t_{n,m}], Z_{\chi,j}[t_{n,m}], W_{\chi,j}[t_{n,m}], H_{\chi,j}[t_{n,m}]\}, \tag{4}$$

where the term χ denotes the object class defined in MS COCO dataset, the term j is the object serial number belonging to the same object class, and $\{X_{\chi,j}[t_{n,m}], Z_{\chi,j}[t_{n,m}], W_{\chi,j}[t_{n,m}], H_{\chi,j}[t_{n,m}]\}$ are, respectively, x -axis position, z -axis position, width, and height, of the j -th object of the object class, χ , at the timing $t_{n,m}$. In this paper, object class χ consists of “car,” “bus (truck),” and “person.” Since the observed number of “truck” was only two in the experiments, “truck” was merged into “bus” class. The object class, “all,” which contains “car,” “bus,” “truck,” and “person” object classes, is defined to evaluate the effectiveness of object detection. The performance of the LQ prediction using the bounding-box information of the object class “all” corresponds to using bounding-box information without the object categories.

Since the object detection block provides several object bounding-boxes for the same object, the overlapped objects are deleted by using the Intersection over Union (IoU), which is given by





$$IoU = \frac{A_{intersection}}{A_{union}}, \tag{5}$$

where $A_{intersection}$ and A_{union} are the overlapping area and total area of the bounding-boxes, respectively. To track the same object over consecutive timing intervals, IoU for the past 2 frames were calculated, and objects whose IoU is greater than 0.6 are recognized as the same object. The more recent object recognized as the same object is assigned the same serial number as the earlier object. Since all combinations of the bounding-boxes are checked by Eq. 5, it is not possible for the same object to belong to several categories.

Since the object bounding-box information includes a reliability score that ranges from 0 to 1.0, we chose the objects whose reliable scores are greater than threshold S_{recog} . Increasing threshold S_{recog} reduces the number of detected objects. Although higher thresholds prevent misrecognition, the detect detection can be delayed. The numbers of the object class χ was defined as N_{χ} , which depends on the object detection algorithm and threshold S_{recog} .

3.3 Object bounding-box information as input features

Since the throughputs are obtained every ΔT , the object bounding-box information is translated to use as input features for the LQ prediction block. Since the time interval of the object bounding-box acquisition, $\Delta \tau$, is shorter than that of the LQ acquisition, the bounding-box features, $\Phi_{\chi,j}[t_n]$, are calculated as median values of the positions and sizes of the object bounding-boxes for LQ acquisition timing between $t_n - 1$ and t_n .

$$\begin{aligned} \Phi_{\chi,j}[t_n] &= \{ \tilde{X}_{\chi,j}[t_n], \tilde{Z}_{\chi,j}[t_n], \tilde{W}_{\chi,j}[t_n], \tilde{H}_{\chi,j}[t_n], \} \\ &= \{ \underset{t_n-1 < t \leq t_n}{Median}(X_{\chi,j}[t]), \underset{t_n-1 < t \leq t_n}{Median}(Z_{\chi,j}[t]), \underset{t_n-1 < t \leq t_n}{Median}(W_{\chi,j}[t]), \underset{t_n-1 < t \leq t_n}{Median}(H_{\chi,j}[t]) \} \end{aligned} \tag{6}$$

Furthermore, the delta-value bounding-box features, $\Delta\Phi_{\chi,j}$ and $\Delta'\Phi_{\chi,j}$, are also calculated to consider the velocity of the mobile objects. $\Delta\Phi_{\chi,j}$ and $\Delta'\Phi_{\chi,j}$ are given by

$$\begin{aligned}\Delta\Phi_{X,j}[t_n] &= \{\Delta X_{X,j}[t_n], \Delta Z_{X,j}[t_n], \Delta W_{X,j}[t_n], \Delta H_{X,j}[t_n]\} \\ &= \left\{ \frac{X_{X,j}[t_n - 1, \beta] - X_{X,j}[t_n - 1, \alpha]}{t_n - 1, \beta - t_n - 1, \alpha}, \frac{Z_{X,j}[t_n - 1, \beta] - Z_{X,j}[t_n - 1, \alpha]}{t_n - 1, \beta - t_n - 1, \alpha}, \frac{W_{X,j}[t_n - 1, \beta] - W_{X,j}[t_n - 1, \alpha]}{t_n - 1, \beta - t_n - 1, \alpha}, \frac{H_{X,j}[t_n - 1, \beta] - H_{X,j}[t_n - 1, \alpha]}{t_n - 1, \beta - t_n - 1, \alpha} \right\},\end{aligned}\quad (7)$$

$$\begin{aligned}\Delta'\Phi_{X,j}[t_n] &= \{\Delta'X_{X,j}[t_n], \Delta'Z_{X,j}[t_n], \Delta'W_{X,j}[t_n], \Delta'H_{X,j}[t_n]\} \\ &= \left\{ \frac{\tilde{X}_{X,j}[t_n] - \tilde{X}_{X,j}[t_n - 1]}{\Delta T}, \frac{\tilde{Z}_{X,j}[t_n] - \tilde{Z}_{X,j}[t_n - 1]}{\Delta T}, \frac{\tilde{W}_{X,j}[t_n] - \tilde{W}_{X,j}[t_n - 1]}{\Delta T}, \frac{\tilde{H}_{X,j}[t_n] - \tilde{H}_{X,j}[t_n - 1]}{\Delta T} \right\},\end{aligned}\quad (8)$$

where $\Delta\Phi_{X,j}[t_n]$ is the delta value of the object bounding-box information between t_{n-1} and t_n , the timings $t_{n-1, \alpha}$ and $t_{n-1, \beta}$ are the first and last acquisition timing of the bounding-box information in the time region $t_{n-1} < t \leq t_n$, and $\Delta'\Phi_{X,j}[t_n]$ is the delta values of $(\Phi_{X,j}[t_n] - \Phi_{X,j}[t_{n-1}])$. $\Delta\Phi_{X,j}$ can be obtained when there are at least two pieces of object bounding-box information between t_{-1} and t_n , and $\Delta'\Phi_{X,j}$ requires the object bounding-box information in the previous time slot t_{n-1} . In Fig. 2, $\Delta\Phi_{\text{per},2}[t_0]$ and $\Delta'\Phi_{\text{per},2}[t_0]$ are calculated as $(\Phi_{\text{per},2}[t_{-1,4}] - \Phi_{\text{per},2}[t_{-1,1}])/(3\Delta\tau)$ and $(\Phi_{\text{per},2}[t_0] - \Phi_{\text{per},2}[t_{-1,1}])/\Delta T$, respectively. The terms α and β for $\Delta\Phi_{\text{per},2}[t_0]$ are 1 and 4, respectively, since the object bounding-box information of the second person object was not observed at $t_{-1,0}$. $\Delta'\Phi_{X,j}$ corresponds to a longer time average than $\Delta\Phi_{X,j}$.

3.4 LQ prediction block

The second machine learning block predicts the future LQ by using the object bounding-box information. In this paper, the random forest regression with 500 decision trees is used to evaluate the LQ prediction performance. The input features for the LQ prediction block, $\Omega[t_0]$, were chosen from the bounding-box features $\Phi_{X,j}[t_n]$, $\Delta\Phi_{X,j}[t_n]$, and $\Delta'\Phi_{X,j}[t_n]$, where $n \leq 0$. The LQ model f_I in Eq. 1 is pre-trained by using the dataset of $(\tilde{R}[t_i], \Omega[t_0])$, and the future normalized throughput is given by

$$\hat{R}[t_i] = f_I(\Omega[t_0]).\quad (9)$$

In random forest regression, the output of function f_I is obtained as an average of outputs of 500 decision trees. The prediction error, $E[t_i]$, is defined as

$$E[t_i] = |\tilde{R}[t_i] - \hat{R}[t_i]|,\quad (10)$$

where $|A|$ denotes the absolute value of A .

For performance comparison, the prediction performance of the LQ prediction using past LQ features, $\Theta[t_0]$, which are chosen from the past normalized throughput information $\tilde{R}[t_n]$ ($n \leq 0$), is also evaluated. The relationship between the target normalized throughput and the past normalized throughputs is pre-trained as function f_L by using dataset $(\tilde{R}[t_i], \Theta[t_0])$, while the relationship between the target normalized throughput and both bounding-box information and past normalized throughputs is also pre-trained by using datasets $(\tilde{R}[t_i], \{\Omega[t_0], \Theta[t_0]\})$ as function $f_{I\&L}$. The prediction errors are evaluated by the predicted normalized throughputs $f_L(\Theta[t_0])$ and $f_{I\&L}(\Omega[t_0], \Theta[t_0])$.

4 Experiment and dataset

The experiments evaluated the LQ prediction performance in an actual outdoor environment. The major parameters are shown in Table 1. The throughputs in 5.660-GHz

channels were measured in IEEE 802.11 ac [30]. No interference signals were observed in this environment. The bandwidth was set to 20 MHz. The normalized throughput, $\tilde{R}[t_n]$, was measured every 0.5 ms ($\Delta T = 0.5$ s), and time interval for image acquisition, $\Delta\tau$, was set to 0.1 s, which corresponds to 10 frames per second (FPS). Figure 4 shows a photo of the connected device. The environment surrounding the connected device was captured by 2 HD cameras with fisheye lens, providing a 360° view. The cameras and laptop PC with LQ measurement function were set at 1.2 m and 0.4 m height, respectively. A map of the experiment environment is shown in Fig. 5. A road and sidewalk lay between the connected device and the base station, separated by 42 m, and vehicles and pedestrians passed through the area.

Figure 6 shows examples of the images captured and the coordinate system. The object was detected by the object detection block, and the object bounding-box information was used in the LQ prediction block. The ranges of the x -axis and z -axis were set to from -1 to 1 and from 0 to 1 , respectively. To evaluate LQ prediction performance in the event of surrounding mobile objects, we defined an object transit event. In this paper, the output of the object detection block was used as the dataset to evaluate the LQ prediction performance in the LQ prediction block. The object transit event is the timing at which some object was detected in the transit window with the x -axis boundaries of -0.15 to 0.35 , see Fig. 6. The dataset for the LQ prediction block was generated at the transit event timing, which is the period from 5 s before the transit event to 5 s after the transit event for all objects: “car,” “bus (truck),” and “person.” The vehicle event and person event correspond to the transit event timing of vehicle-related objects (“car” and “bus (truck)”), and people (“person”), respectively. By using the dataset, we generated the LQ model and evaluated the LQ prediction performance.

4.1 Dataset and LQ prediction performance evaluation

The dataset for the LQ prediction block totaled 3490 s of data, containing transit events of 288 cars, 20 buses/trucks, and 36 persons. The vehicle-event data totaled 3061.5 s, while the person-event data held 976 s. The dataset includes 547.5-s data corresponding to both vehicle event and person event. LQ prediction performance was evaluated by the metric of k -cross validation. The dataset was divided into 10 parts, and 9 of tenths were used for training to generate the LQ model. LQ prediction was conducted using the remaining one-tenth dataset.

Table 1 Outdoor experiment parameters

Wireless systems	IEEE 802.11 ac
Radio frequency	5.660 MHz
Frequency bandwidth	20 MHz
Link quality indicator	Normalized throughput (every 0.5 s)
Antenna height of connected device	0.4 m
Distance from the base station	42 m
Camera	High-definition camera (1920 × 1080 pixel)
Frames per second	10
Lens	Focal distance: 1 mm (fisheye)
Camera height	1.2 m

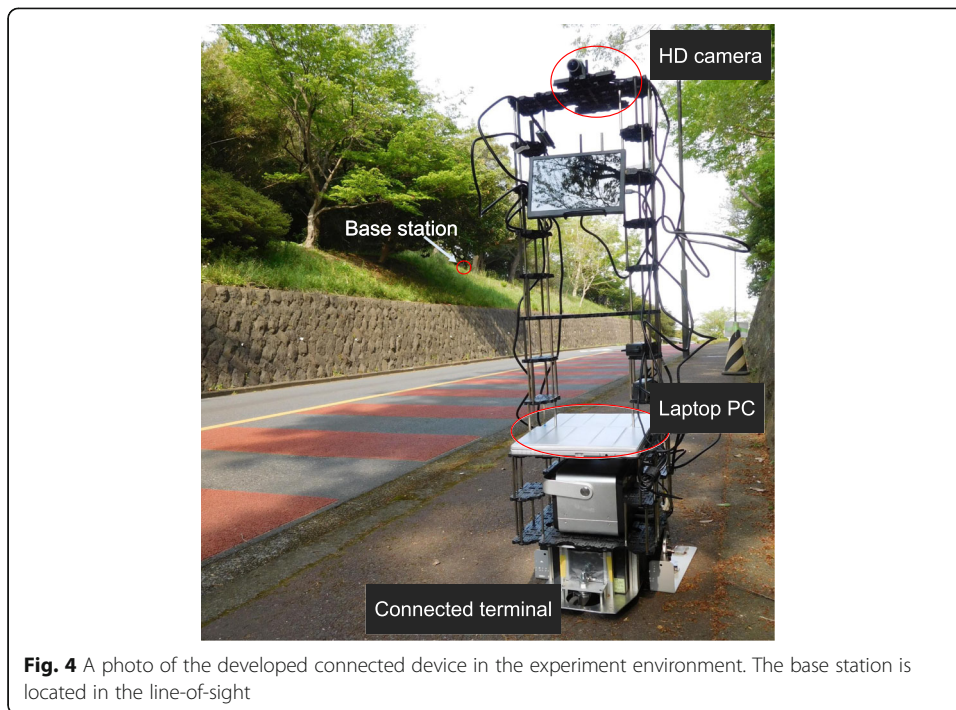


Fig. 4 A photo of the developed connected device in the experiment environment. The base station is located in the line-of-sight

4.2 Measured normalized throughput

The cumulative distribution functions of the normalized throughput, $\bar{R}[t_i]$, of all-transit-event timing, vehicle-event timing, and person-event timing are shown in Fig. 7. We can see that the distribution is far from Gaussian; this is considered to be due to the moving objects. The probabilities of the normalized throughputs falling below 0.8 were 0.113, 0.105, and 0.280 for

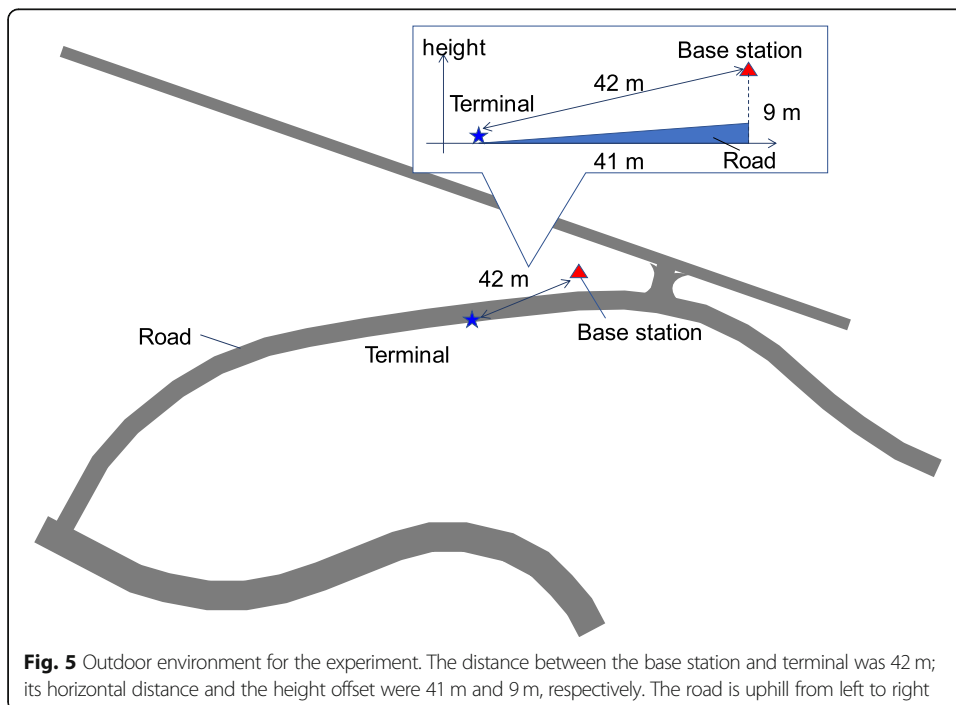


Fig. 5 Outdoor environment for the experiment. The distance between the base station and terminal was 42 m; its horizontal distance and the height offset were 41 m and 9 m, respectively. The road is uphill from left to right

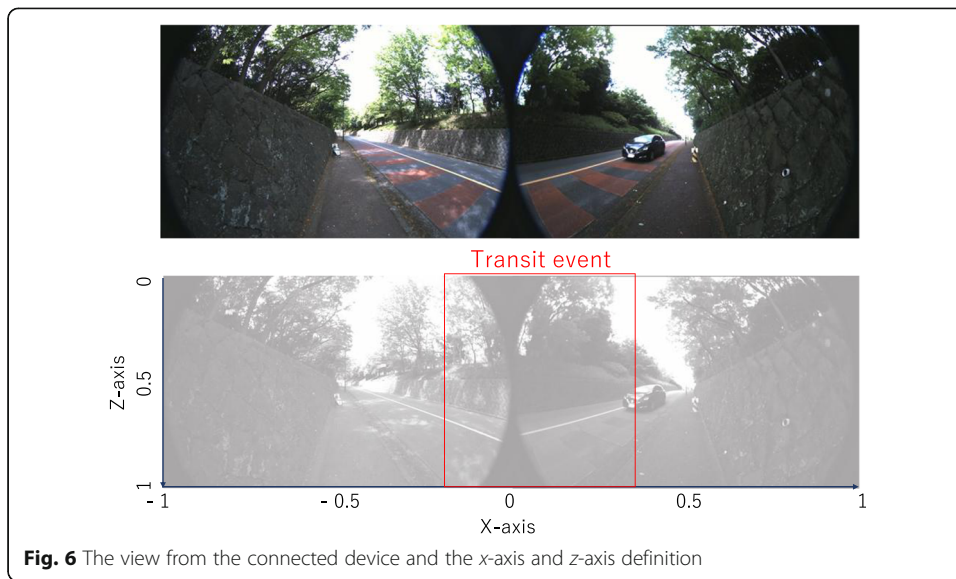


Fig. 6 The view from the connected device and the x-axis and z-axis definition

all event, vehicle event, and person event, respectively. Since the time when the moving persons affect the normalized throughput is longer than that caused by the vehicles because of their low moving speed, they have high probability of low normalized throughput.

4.3 Object detection and object bounding-box

M2Det in 2019 [22] and YOLO v3 in 2018 [23], which are used as the object detection block in Fig. 1, are state-of-the-art detectors based on deep neural networks. The image processing speed and average precisions of M2Det are stated to be 30 ms and 37.6 in [22], while those of YOLO v3 are 51 ms and 33.0 in [23]. The average precision of the object detection denotes the detection performance, and M2Det has better performance than YOLO v3. Both object detectors output object bounding-boxes, their categories,

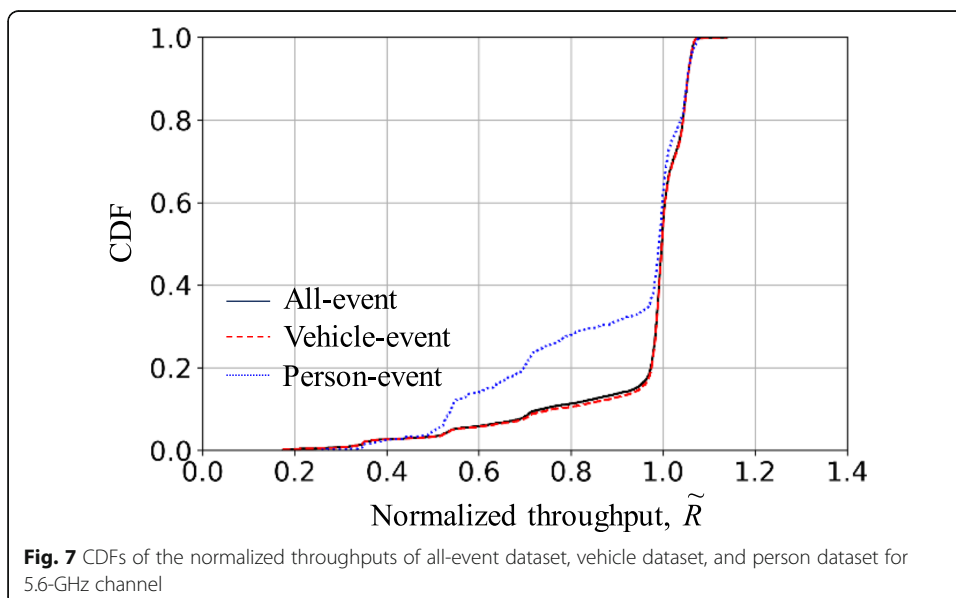


Fig. 7 CDFs of the normalized throughputs of all-event dataset, vehicle dataset, and person dataset for 5.6-GHz channel

and reliability scores. We used the bounding-box information whose reliability score is greater than the threshold value, S_{recog} . Thus, the number of detected objects depends on S_{recog} . Table 2 shows the maximum number of objects detected by M2Det [26] and YOLO v3 [27] for several threshold values, S_{recog} . We checked the maximum number of objects by watching the video, and the number of the maximum numbers of “car,” “bus (truck),” and “person” were 2, 1, and 4, respectively. Therefore, the large number of the detected objects means the object detection block generated unnecessary bounding-boxes. We can see the object number increases as the threshold was set to be low. We can see that the error of the detected number becomes 1 or 0 when the threshold S_{recog} is 0.5 and 0.8, respectively, for M2Det and YOLO v3.

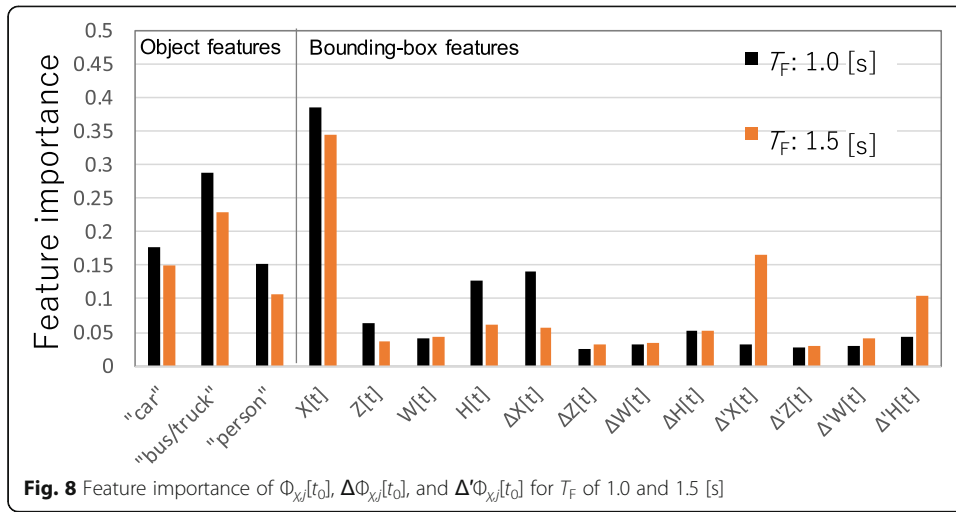
4.4 Feature importance evaluation

Since the LQ prediction block uses the object bounding-box information to predict the LQ, the relationship between the elements of $\Phi_{x,j}[t_0]$, $\Delta\Phi_{x,j}[t_0]$, and $\Delta'\Phi_{x,j}[t_0]$ and LQ was studied using the feature importance of random forest regression [31]. The random forest regression used 500 decision trees, and the normalized throughput with the lead time T_F of 1.0 [s] and 1.5 [s], $\tilde{R}[t_2]$ and $\tilde{R}[t_3]$, were used as target parameters. $\Phi_{x,j}[t_0]$, $\Delta\Phi_{x,j}[t_0]$, and $\Delta'\Phi_{x,j}[t_0]$ were generated by M2Det with S_{recog} of 0.5. The number of input features from $\Phi_{x,j}[t_0]$, $\Delta\Phi_{x,j}[t_0]$, and $\Delta'\Phi_{x,j}[t_0]$ was 132, where N_{car} , N_{bus} , and N_{person} were 4, 2, and 5, respectively. Figure 8 shows a summation of the feature importance corresponding to the object categories, and the bounding-box component elements, in the dataset. The object “bus (truck)” has the highest importance among the objects. The largest metallic structure has the biggest impact on the normalized throughput. Among the bounding-box information, the x -axis position has the largest weight since all objects came from left to right or from right to left. The feature importance of $\Delta\Phi_{x,j}[t_0]$ and $\Delta'\Phi_{x,j}[t_0]$ depends on T_F . When T_F is 1.0 [s], $\Delta X_{x,j}$ and $\Delta H_{x,j}$ in $\Delta\Phi_{x,j}[t_0]$ were the largest among $\Delta\Phi_{x,j}[t_0]$ and $\Delta'\Phi_{x,j}[t_0]$. However, $\Delta'X_{x,j}$ and $\Delta'H_{x,j}$ of $\Delta'\Phi_{x,j}[t_0]$ became the largest for $T_F = 1.5$ [s]. This means that the instant information is needed to accurately predict the near future condition, and the accuracy of the delta is more important for predictions with greater lead times.

Figure 9 plots the feature importance for the target normalized throughputs with T_F of 1.0 [s] when the vehicle event and person event were picked up from the dataset. The feature importance of boundary-box features in the vehicle event was similar to that of the all-event dataset, and the x -axis information is the most important. In the person event, height information $H_{x,j}$ is more important than the x -axis position, $X_{x,j}$. This is because the width information of people is unstable compared to that of the vehicle. People with their arms spread wide are detected as large width structures, and the x -axis position can be biased.

Table 2 Maximum number of recognized objects (“car,” “bus (truck),” and “person”)

Threshold, S_{object}	M2Det	YOLO v3	Manually counted
0.1	19 cars, 4 buses, 19 persons	8 cars, 4 buses, 14 persons	2 cars, 1 bus, 4 persons
0.3	3 cars, 2 buses, 5 persons	3 cars, 2 buses, 8 persons	
0.5	2 cars, 2 buses, 4 persons	2 cars, 2 buses, 4 persons	
0.8	2 cars, 2 buses, 4 persons	2 cars, 1 bus, 4 persons	



4.5 Input feature set for LQ prediction block

The input feature sets for LQ prediction block were generated by

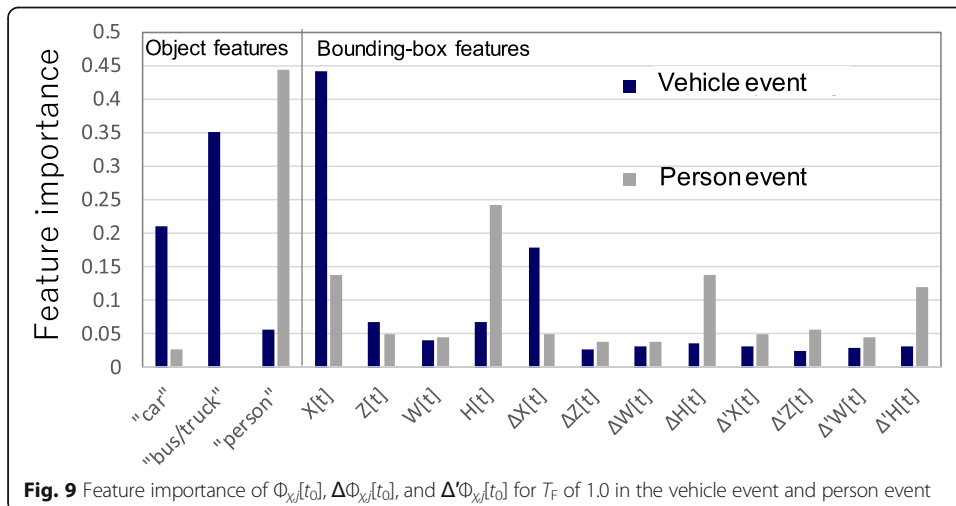
$$\Omega_{BB}[t_0] = \{ \Phi_{\chi,j}[t_0] \} \text{ where } \chi \in \{ \text{"car"}, \text{"bus (truck)"}, \text{"person"} \} \text{ and } 1 \leq j \leq N_{\chi}, \quad (9)$$

$$\Omega_{BBA}[t_0] = \{ \Phi_{\chi,j}[t_0] \} \text{ where } \chi \in \{ \text{"all"} \} \text{ and } 1 \leq j \leq N_{\text{all}}, \quad (10)$$

$$\Omega_{BV}[t_0] = \{ \Phi_{\chi,j}[t_0], \Delta X_{\chi,j}[t_0], \Delta H_{\chi,j}[t_0], \Delta'X_{\chi,j}[t_0], \Delta'H_{\chi,j}[t_0] \} \quad (11)$$

where $\chi \in \{ \text{"car"}, \text{"bus (truck)"}, \text{"person"} \} \text{ and } 1 \leq j \leq N_{\chi}$.

$\Omega_{BB}[t_0]$ is the basic boundary-box information set of the objects, "car," "bus (truck)," and "person." The feature number is given by $4 \times (N_{\text{car}} + N_{\text{bus}} + N_{\text{person}})$. $\Omega_{BBA}[t_0]$ is the boundary-box information with single object class "all." The feature number is $4 \times N_{\text{all}}$. $\Omega_{BV}[t_0]$ is the advanced boundary-box information that contains the delta values, $\Delta X_{\chi,j}$, $\Delta H_{\chi,j}$, $\Delta'X_{\chi,j}$, and $\Delta'H_{\chi,j}$ where $\chi \in \{ \text{"car"}, \text{"bus (truck)"}, \text{and "person"} \}$. $\Delta Z_{\chi,j}$, $\Delta W_{\chi,j}$, $\Delta'Z_{\chi,j}$, and $\Delta'W_{\chi,j}$ were not used because of their low feature importance.



As conventional LQ prediction approach, past LQ information use [15] was also evaluated. The input feature set of current and past LQ information is given by

$$\Theta_{\text{TH}}[t_0] = \{\tilde{R}[t_0], \tilde{R}[t_{-1}], \tilde{R}[t_{-2}], \tilde{R}[t_{-3}], \tilde{R}[t_{-4}], \tilde{R}[t_{-5}]\}, \quad (12)$$

where $\tilde{R}[t_{-5}]$ corresponds to 2.5 s past normalized throughput. The effectiveness of the past feature use is discussed in Section 5.3. Furthermore, the input feature set for the combination of the object bounding-box and past LQ was also generated as

$$\Omega_{\text{BVT}}[t_0] = \{\Omega_{\text{BV}}[t_0], \Theta_{\text{TH}}[t_0]\} \quad (13)$$

4.6 Calculation complexity in LQ prediction block

The computation time of LQ prediction block was evaluated by using the all-event dataset of 3490 s. The training dataset ($\tilde{R}[t_i], \Omega_{\text{BB}}[t_0]$), which consists of 3141 s (nine tenths of all the data), is used to construct function f_1 , and the predicted throughput data $\hat{R}[t_i]$ is calculated by using the remaining 349-s data (one-tenth of all the data). It takes 0.51 s to provide the normalized throughput $\tilde{R}[t_i]$ for a 349-s data by using the LQ model. One-second bounding-box information can be processed in 1.4 ms by random forest regression. This shows that the dominant computation load of the two-step LQ prediction proposal is the object detection block. Regarding the training computation load for LQ model, it takes 218 s to generate function f_1 by using the training dataset.

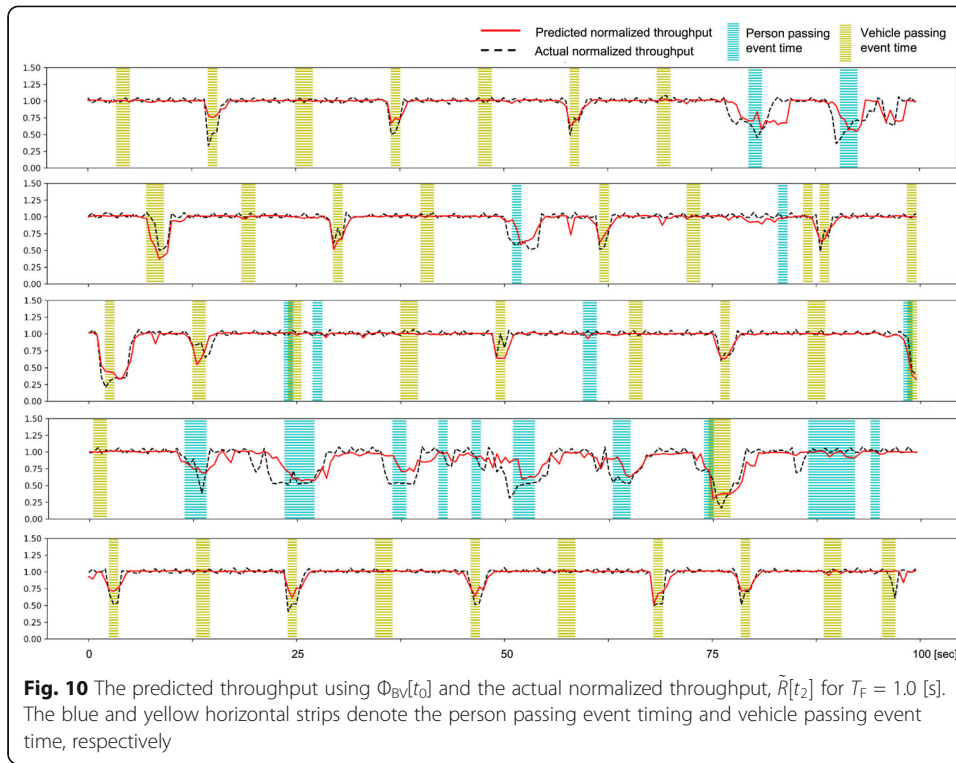
5 Results and discussion

5.1 LQ prediction in time domain

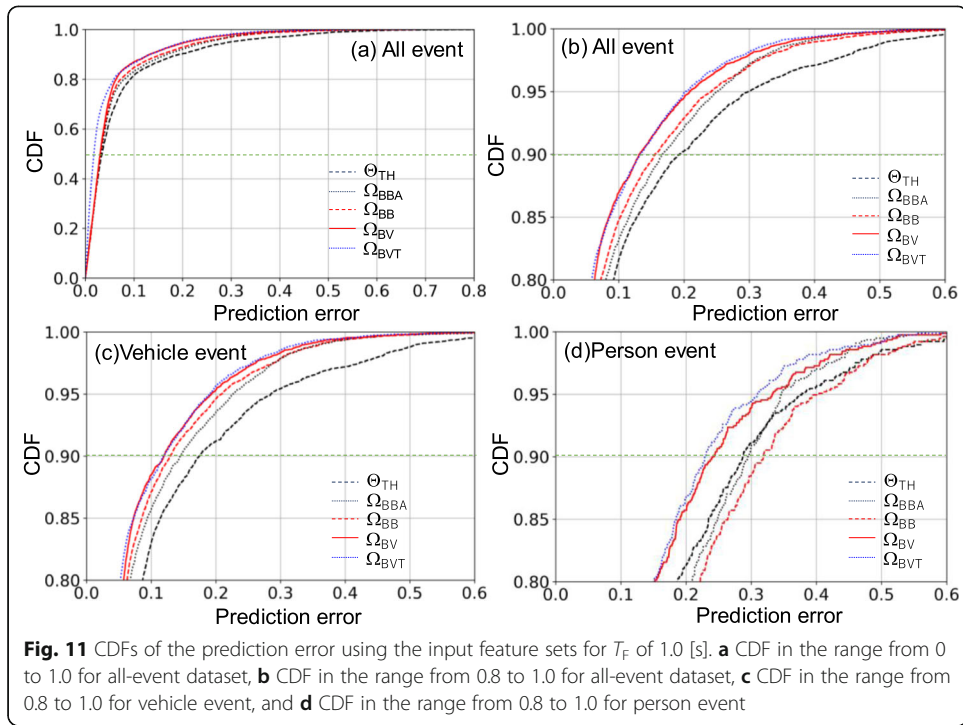
The normalized throughput was predicted by using the input feature set, $\Omega_{\text{BV}}[t_0]$, provided by M2Det with S_{recog} of 0.5 and T_{F} of 1.0 [s]. Figure 10 shows the measured and predicted throughput in the 5.6-GHz channel for the 500-s dataset, and the red solid line and black dashed line correspond to the predicted throughput and actual throughput, respectively. If the prediction is perfect, the lines overlap. The timing of the objects being present in the transit window shown in Fig. 6 is indicated as horizontal stripes. The yellow and blue stripes indicate the transit events of vehicles and persons; they were detected by M2Det with $S_{\text{recog}} = 0.3$. The threshold setting and dependencies on the object detection algorithms are discussed in Section 5.5. Figure 10 shows that the throughput degradation of the 5.6-GHz channel was predicted by using the 1 s past boundary-box information. In particular, the vehicle-related throughput degradation was more clearly predicted than people movement. This is because vehicle movement was stable over time while the people walking around the terminal changed movement speed and body position more freely.

5.2 CDF of the prediction error

Figure 11 shows the CDFs of the prediction errors for the 5.6-GHz normalized throughput with $T_{\text{F}} = 1.0$ [s] using the input feature set, Ω_{BB} , Ω_{BBA} , Ω_{BV} , Ω_{BVT} , and Θ_{TH} . The object bounding-boxes were detected by M2Det with S_{recog} of 0.3. The numbers of input features of Ω_{BB} , Ω_{BBA} , Ω_{BV} , Ω_{BVT} , and Θ_{TH} were 52, 32, 104, 110, and 6, respectively. Figure 11 a and b show the CDFs of the prediction errors for the



probability range from 0 to 1.0 and 0.8 to 1.0, respectively. The distribution from 0 to 80% in CDF mainly corresponds to the timing when the moving object does not affect the LQ while that from 80 to 100% denotes that the moving object does impact the LQ. The horizontal distribution at 99.99% value of the CDF denotes the LQ change which cannot be predicted by using the proposed LQ prediction. Since Fig. 7 shows that about 20% of the normalized throughputs of the all-event dataset are degraded by the mobile object, the highest 20% of the prediction error (80 to 100%) is considered to correspond to the LQ degradation caused by the mobile object. Thus, we focus on the 90% value of the CDF of the prediction error as the middle value between 80 and 100%. Figure 11 a shows that the prediction performance using Ω_{BVT} has the best performance and the median value of Ω_{BVT} is 49.3% less than that of Θ_{TH} . Figure 11 b shows 90% values of Ω_{BVT} and Ω_{BV} were 31.1% and 31.9% less than those of Θ_{TH} . The prediction performance using Ω_{BV} was slightly better than that using Ω_{BVT} at 90% outage. This indicated that past LQ information did not contribute to the prediction performance for the 5.6-GHz channel with $T_F = 1.0$ [s] when Ω_{BV} was available. The 90% outage values of Ω_{BB} and Ω_{BBA} were 19.5% and 13.3% less than those of Θ_{TH} , respectively. Thus, the object classification improves 6.2% at 90% value (Ω_{BB} over Ω_{BBA}), and the delta value information improves 12.4% (Ω_{BV} over Ω_{BB}). Figure 11 c and d show the CDFs of the prediction errors corresponding to vehicle event and person event, respectively, for the range from 0.8 to 1.0. In the vehicle event (c), the 90% outage values for Ω_{BVT} and Ω_{BV} were 31.6% and 31.4%, respectively, less than those for Θ_{TH} . In the person event (d), the 90% outage values for Ω_{BVT} and Ω_{BV} were 16.2% and 20.3%, respectively, less than those for Θ_{TH} . This suggests that LQ prediction is more effective for vehicle transit events than for person transit events. In the person event,



the accuracy of the LQ prediction using Ω_{BB} and Ω_{BBA} was less than that of Θ_{TH} -based LQ prediction, and the prediction error using Ω_{BB} was greater than that using Ω_{BBA} . This is caused by the shortage of training data of person transit event. Considering all objects as a single category in Ω_{BBA} yielded more efficient training of the LQ prediction model in the second machine learning block for our datasets.

5.3 Past information use

LQ prediction performances with past information was evaluated for the lead time T_F of 1.0 [s] by using M2Det with S_{recog} of 0.3. In this evaluation, the input feature set from $t_0 - T_{\text{past}}$ to t_0 was used. Thus, the lines of Ω_{BV} and Θ_{TH} denote LQ prediction performance with $\{\Omega_{BV}[t_0], \dots, \Omega_{BV}[t_0 - T_{\text{past}}]\}$ and $\{\tilde{R}[t_0], \dots, \tilde{R}[t_0 - T_{\text{past}}]\}$. The numbers of input features of Ω_{BV} and Θ_{TH} were 104, 208, ..., 624, and 1, 2, ..., 6, for T_{past} of 0, 0.5, ..., 2.5 [s], respectively. Figure 12 shows that LQ prediction using past LQ information improved the prediction performance while the past bounding-box information yielded no improvement. This shows that the latest object bounding-box information is the most important features in predicting LQ.

5.4 Lead time dependency

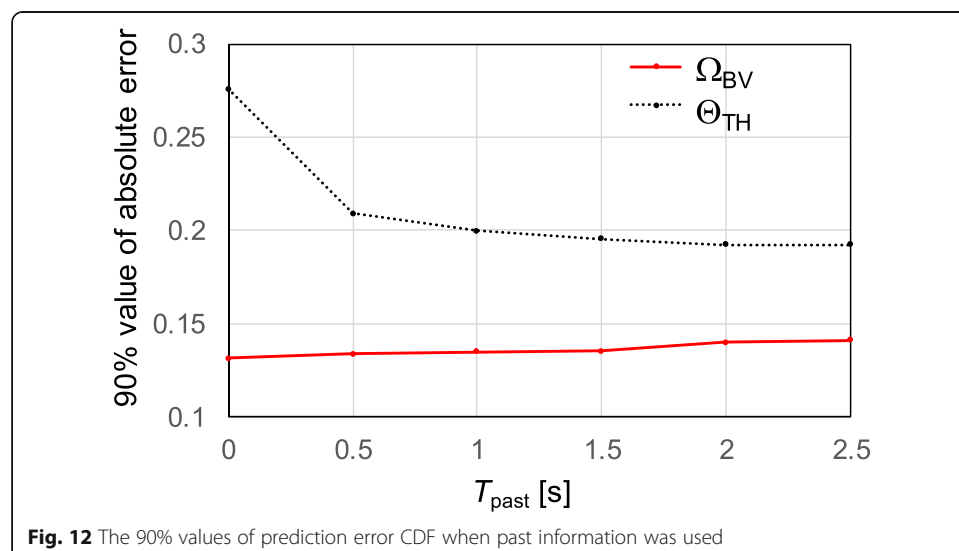
The prediction performance against lead time T_F was evaluated by using M2Det with S_{recog} of 0.3. Figure 13 a, b, and c plot the 90% values of LQ prediction with Θ_{TH} , Ω_{BV} , and Ω_{BVT} versus T_F for all-transit events, vehicle events, and person events, respectively. Although the prediction error with Ω_{BV} increases as T_F becomes large, the rate of degradation in LQ prediction performance with Ω_{BV} and Ω_{BVT} was gentle and the

prediction error for T_F of 2.0 [s] was much better than that with Θ_{TH} for T_F of 1.0 [s]. We can see that old LQ information was less effective for lead times greater than 1 s since there was no advantage to using Ω_{BVT} rather than Ω_{BV} .

5.5 Impact of detection threshold and algorithm

The proposed LQ prediction adopts the two-step machine learning, and each machine learning block must prepare its own model. Since the pre-trained model has much higher training cost than the prediction phase, the dependencies on the object detection model and detection threshold setting must be evaluated. If the relationship between the bounding-box information and the LQ depends on the detection algorithm and threshold setting, the second LQ prediction block must prepare all combinations to cover all possible object detection algorithms and threshold settings. The resulting computation load for training will be significant. On the other hand, if the dependency on object detection algorithms and threshold setting is not critical, the prediction model of the second machine learning block can be common and it is also expected that the second machine learning block can be developed independently by using the bounding-box information. Furthermore, it is also important to confirm that better object detection algorithms will enhance the LQ prediction performance. Therefore, the impact of the object detection algorithm used and threshold S_{recog} was evaluated to confirm that the proposal can take advantage of advances in object detection algorithms.

The bounding-box information was generated by using M2Det and YOLO v3 with S_{recog} values of 0.1, 0.3, 0.5, and 0.8, and 8 all-event datasets were generated for the input feature set Ω_{BV} . Eight all-event datasets were divided into 10 sub-datasets to conduct the k-cross validation and duplicated for 10 test sub-datasets and 10 training sub-datasets. Then, 9 of the training sub-datasets were used as the training data to make the LQ prediction model, and the normalized throughput was determined for the remaining one of the test sub-datasets, which corresponds to different timing.



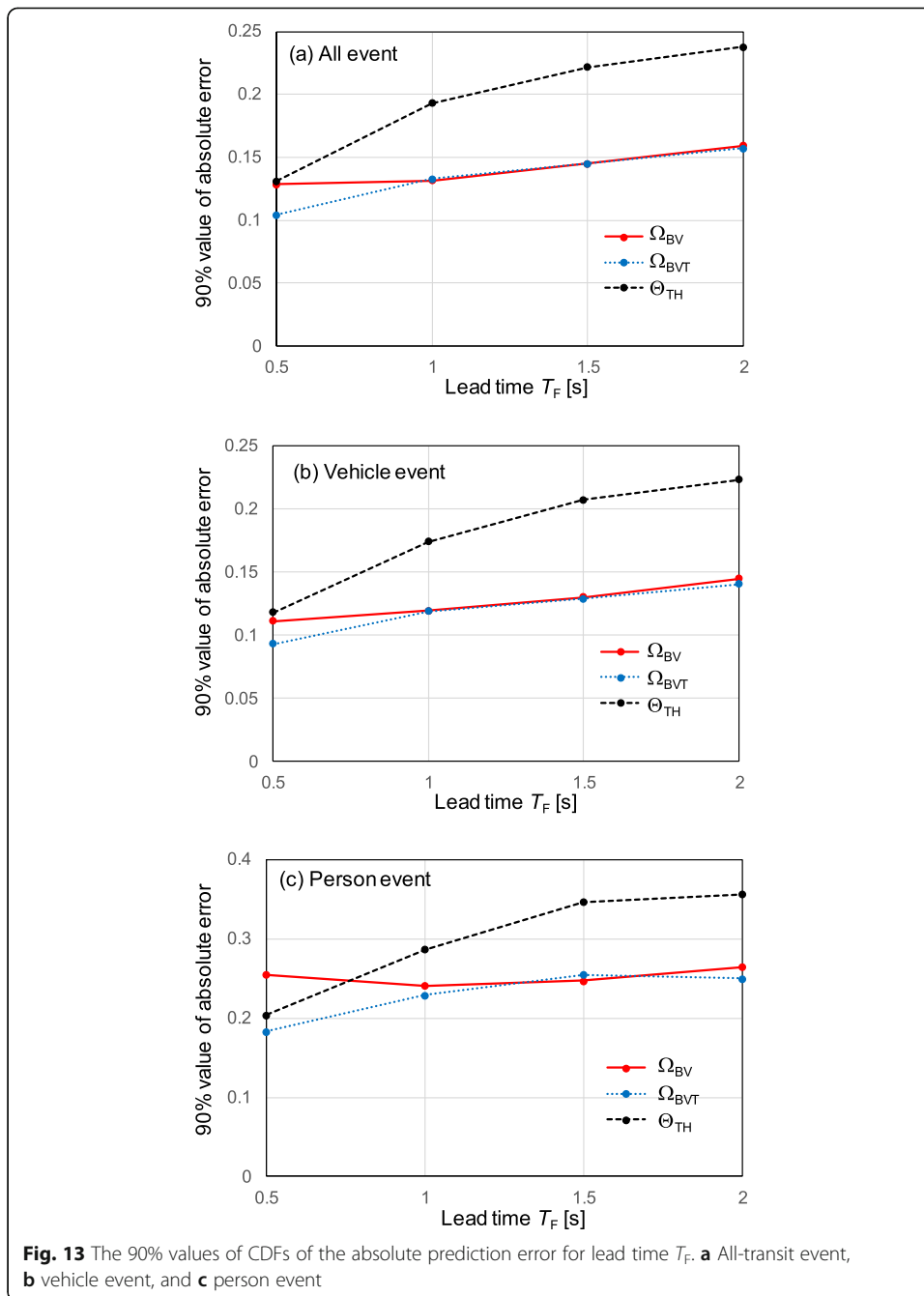


Table 3 shows the 90% outage value of the prediction error corresponding to all the combination of the object detection algorithm pairs. Since 90% absolute error of LQ prediction with old LQ information was 0.183, some of the combinations were worse than that with Θ_{TH} . We can see that M2Det with S_{recog} of 0.8 for training and M2Det with S_{recog} with 0.1 for test provided bad prediction performances, and the same algorithm combination with S_{recog} of 0.3 and 0.5 provided the best performances. Although almost all combinations outperformed LQ prediction with the old LQ information, using the same algorithm for training and test yielded better performance than using the different algorithms for test and training. Since M2Det has better detection

Table 3 The 90% value of absolute error with Ω_{BY} using different object detection algorithm and threshold

		YOLO v3				M2Det				Training Average
		0.1	0.3	0.5	0.8	0.1	0.3	0.5	0.8	
YOLO v3	0.1	0.152	0.137	0.137	0.147	0.193	0.139	0.142	0.148	0.149
	0.3	0.162	0.148	0.139	0.148	0.197	0.148	0.147	0.148	0.155
	0.5	0.173	0.140	0.136	0.141	0.203	0.142	0.141	0.144	0.153
	0.8	0.182	0.168	0.153	0.143	0.203	0.165	0.159	0.148	0.165
M2Det	0.1	0.161	0.147	0.148	0.158	0.156	0.136	0.136	0.145	0.148
	0.3	0.179	0.161	0.150	0.154	0.186	0.131	0.136	0.149	0.156
	0.5	0.187	0.169	0.160	0.154	0.194	0.142	0.138	0.142	0.161
	0.8	0.217	0.194	0.181	0.163	0.221	0.162	0.157	0.143	0.180
Test Average		0.177	0.158	0.150	0.151	0.194	0.146	0.145	0.146	

accuracy than YOLO v3, the best 90% error of M2Det (S_{recog} of 0.3) is better than that of YOLO v3 (S_{recog} of 0.5). This shows that the LQ prediction scheme can take advantage of advances in object detection algorithms if the LQ prediction model is updated by using the advanced object detection algorithms.

The averaged values of each column and row are also listed in Table 3. The averaged prediction error for training sub-datasets decreases as S_{recog} is set to a lower value. On the other hand, the detection accuracy becomes more important for the prediction phase. The averaged prediction error for the test sub-dataset shows that S_{recog} of 0.5 yielded the lowest error for both M2Det and YOLO v3. When training the LQ prediction model, many mobile objects should be used even if the misrecognition number increases.

6 Conclusion

This paper presented a wireless link quality prediction scheme that uses the two-step machine learning; the first machine learning block realizes object detection while the second block predicts the future LQ using bounding-box information. Although the structure is simple, the proposed LQ prediction can well predict throughputs with lead times of more than 1 s. Proof of concept experiments were conducted in 5.6-GHz WLAN channels, and the relationship between the type of passing object and its impact of measured throughput was shown. Performance evaluation in the 5.6-GHz channel clarified the dependency on the future time, the input feature sets, and the advantages compared to LQ prediction based on the past throughput information. By using the object bounding-box information, the 90% values of the absolute prediction error in the proposed LQ prediction were 31.1% less than those of the LQ prediction using past LQ information. By using the LQ prediction, the connected device side can recognize the surrounding environment precisely. So far, wireless management techniques have been developed on the network side as in LTE and 5G because the network side can monitor the data traffic and obtain various types of information. The network side has much more abundant information than the terminal side, while the connected device wins in terms of freshness of obtained information. The vision of the smart connected devices is expected to be one of keys for raising the next-generation wireless systems to a whole new level of service.

Abbreviations

5G: 5th generation; BS: Base station; CSI: Channel state information; EHF: Extremely high frequency; HD: High-definition; IoU: Intersection over Union; LAN: Local area network; LIDAR: Light detection and ranging; LQ: Link quality; LTE: Long-term evolution; mmWave: Millimeter wave; QoS: Quality of service; SHF: Super high frequency; UAV: Unmanned aerial vehicle; WLAN: Wireless local area network

Acknowledgements

Not applicable.

Authors' contributions

RK designed and performed the experiments. RK and KT analyzed the experiment results. All the authors participated in writing the article and revising the manuscript. The authors read and approved the final manuscript.

Availability of data and materials

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Received: 29 December 2019 Accepted: 2 October 2020

Published online: 19 October 2020

References

1. N. Al-Falahy, O.Y. Alani, Demonstration of 5G connected cars. *IT Professional*. **19**(1), 12–20 (Feb. 2017)
2. T.J. Lee, C.H. Kim, D.I.D. Cho, A monocular vision sensor-based efficient SLAM method for indoor service robots. *IEEE Trans. on industrial electronics* **66**(1), 318–328 (Jan. 2019)
3. G. Peng, W. Zheng, Z. Lu, J. Liao, L. Hu, G. Zhang, and D. He, "An improved AMCL algorithm based on laser scanning match in a complex and unstructured environment," *Hindawi, Complexity*, vol. 2018, article ID 2327637, Dec. 2018.
4. NGMN Alliance, "Perspectives on Vertical Industries and Implications for 5G," *white paper*, Sept. 2016.
5. A. Osseiran, Jose F. Monserrat, and Patrick Marsch, eds. *5G mobile and wireless communications technology*. Cambridge University Press, 2016.
6. B. Bellalta, IEEE 802.11ax: high-efficiency WLANs. *IEEE Wireless Communications* **23**(1), 38–46 (Feb. 2016)
7. Y. Ghasempour, G.R.C.M. da Silva, C. Cordeiro, E.W. Knightly, IEEE 802.11ay: next-generation 60GHz communication for 100 Gb/s Wi-Fi. *IEEE Communication Magazine* **55**(12), 186–192 (Oct. 2017)
8. L. Simić, J. Arnold, M. Petrova, and P. Mähänen. "RadMAC: radar-enabled link obstruction avoidance for agile mm-wave beamsteering." In *Proceedings of the 3rd Workshop on Hot Topics in Wireless*, pp. 61–65. ACM, 2016.
9. G. Cerar, H. Yetgin, M. Mohorčič, and C. Fortuna, "Machine Learning for Link Quality Estimation: A Survey," *arXiv Prepr.* 1812.08856, Dec. 2018.
10. C. Zhang, P. Patras, H. Haddadi, Deep learning in mobile and wireless networking: survey. *IEEE Communications surveys & tutorials* **21**(3), 2224–2287 (Mar. 2019)
11. H. Yang, X. Xie, M. Kadoch, "Machine learning techniques and a case study for intelligent wireless networks," *IEEE Network* (Jan. 2020). <https://doi.org/10.1109/MNET.001.190035>
12. C. Luo, J. Ji, Q. Wang, X. Chen, P. Li, Channel state information prediction for 5G wireless communications: a deep learning approach. *IEEE Transactions on Network Science and Engineering* **7**(1), 227–236 (Jun. 2018)
13. R. Liao, H. Wen, J. Wu, H. Song, F. Pan, and L. Dong, "The Rayleigh fading channel prediction via deep learning," *Hindawi, wireless communications and mobile computing*, vol. 2018, article ID 6497340, Jul. 2018.
14. L. Bai, C. Wang, J. Huang, Q. Xu, Y. Yang, G. Goussetis, J. Sun, and W. Zhang, "Predicting wireless mmWave massive MIMO channel characteristics using machine learning algorithms," *Hindawi, wireless communications and mobile computing*, vol. 2018, article ID 9783863, Aug. 2018.
15. Q. Xu, S. Mehrotra, Z.M. Mao, and J. Li, "PROTEUS: network performance forecast for real-time, interactive mobile applications," in *Proceeding of the 11th annual international conference on Mobile systems, applications, and services*, ACM, Jun. 2013.
16. B. Wei, K. Kanai, W. Kawakami, and J. Katto, "HOAH: a hybrid TCP throughput prediction with autoregressive model and hidden Markov model for mobile network," *IEICE transaction Communication*. vol. E101-B, no. 7. Jul. 2018.
17. H. Yang, X. Xie, M. Kadoch, Intelligent resource management based on reinforcement learning for ultra-reliable and low-latency loV communication networks. *IEEE Transactions on Vehicular Technology* **68**(5), 4157–4169 (May 2019)
18. E. N. Almeida, K. Fernandes, F. Andrade, P. Silva, R. Campos, and M. Ricardo, "A machine learning based quality of service estimator for aerial wireless networks," in *Proceeding of 2019 International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, Oct. 2019.
19. Y. Wang, M. Narasimha, R.W. Heath Jr., in *Proceeding of IEEE 19th international workshop on signal processing advances in wireless communications (SPAWC)*. MmWave beam prediction with situational awareness: a machine learning approach (Jun. 2018)
20. Y. Oguma, R. Arai, T. Nishio, K. Yamamoto, M. Morikura, in *Proceeding of the IEEE Global Commun. Conf. (GLOBECOM)*. Proactive base station selection based on human blockage prediction using RGB-D cameras for mmWave communications (Dec. 2015)
21. T. Nishio, H. Okamoto, K. Nakashima, Y. Koda, K. Yamamoto, M. Morikura, Y. Asai, R. Miyatake, Proactive received power prediction using machine learning and depth images for mmWave networks. *IEEE Journal on Selected Areas in Commun.* **37**(11), 2413–2427 (Nov. 2019)
22. L. Liu, W. Ouyang, X. Wang, P. Fieguth, J. Chen, X. Liu, M. Pietikainen, Deep learning for generic object detection: a survey. *International Journal of Computer Vision* **128**, 261–318 (Oct. 2019)

23. L. Jiao, F. Zhang, F. Liu, S. Yang, L. Li, Z. Feng, R. Qu, A survey of deep learning-based object detection. *IEEE Access* **7**, 128837–128868 (Sep. 2019)
24. C.J. Lowrance, and A.P. Lauf, "An active and incremental learning framework for the online prediction of link quality in robot networks," *Engineering Applications of Artificial Intelligence*, No. 77 pp. 197–211, Jan. 2019. .
25. K. Katsaros, M. Dianati, R. Tafazolli, R. Kernchen, in *Proceeding of IEEE Vehicular Networking Conference (VNC)*. CLWPR – a novel cross-layer optimized position based routing protocol for VANETs (Nov. 2011)
26. Q. Zhao, T. Sheng, Y. Wang, Z. Tang, Y. Chen, L. Cai, H. Ling, M2det: A single-shot object detector based on multi-level feature pyramid network. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 9259–9266 (Jul. 2019)
27. J. Redmon and A. Farhadi, "Yolov3: An Incremental Improvement," CoRR, vol. abs/1804.02767, 2018.
28. M. Lauridsen, L.C. Gimenez, I. Rodriguez, T.B. Sorensen, P. Mogensen, From LTE to 5G for connected mobility. *IEEE Communications Magazine* **55**(3), 156–162 (March 2017)
29. T. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Doll'ar, and L. Zitnick, "Microsoft COCO: common objects in context," *Computer Vision – ECCV 2014. ECCV 2014. Lecture Notes in Computer Science*, vol 8693, pp. 740–755, 2014.
30. WG802.11-Wireless LAN Working Group, IEEE P802.11 ac Draft 5.0: Enhancements for very high throughput for operation in bands below 6 GHz. IEEE Ongoing Project, IEEE Standards, Piscataway, NJ, USA, 2013.
31. <https://scikit-learn.org/stable/>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ [springeropen.com](https://www.springeropen.com)
