VLC localization: deep learning models by Kalman filter algorithm combined with RSS

Wessam M. Salama1 · Moustafa H. Aly2 · Eman S. Amer3

Received: 30 April 2022 / Accepted: 7 July 2022 / Published online: 3 August 2022
© The Author(s) 2022

Abstract
In this paper, a new framework is presented for indoor visible light communication (VLC) system, based on Yolo v3, EfficientNetB3, and DenseNet121 deep learning (DL) models, as well as an optimization strategy. The proposed framework consists of two steps: data collecting and DL model training. To start, data is acquired using MATLAB and Kalman Filtering (KF) with averaging approaches. Second, the received signal strength (RSS) is employed as the DL models input, with the Cartesian coordinates as the DL models output. The averaging RSS approach combined with KF algorithm are used in the suggested framework. This work introduces the impacts of Non-Line-of-Sight (NLoS) for initial reflection and Line-of-Sight (LoS) based on the three mentioned DL models. Furthermore, we used Bayesian optimization and automatic hyper-parameter (HP) optimization to increase system efficiency and to reduce positioning error in DL models. The obtained results show that the models outperform existing the HP-RSS-KF-LoS-DL models in terms of localization error when compared to traditional RSS signal-based localization techniques. Many performance indicators are considered to evaluate the proposed framework resiliency, including accuracy (ACC), area under the curve (AUC), sensitivity (Se), and precision (Pr), as well as F1-score, root mean square error (RMSE), training, and testing time. The DL models are generated and trained using Python software on a Kaggle Notebook GPU cloud (2 CPU cores and 13 GB RAM). The achieved results are: 99.99% ACC, 99.98% AUC, 98.88% Se, 98.98% Pr, 99.97% F1-score, 0.112 cm RMSE, and 0.29 s testing time. The proposed system could be easily deployed for autonomous applications, based on the analysis of the experimental data. Several applications can be used depending on enhancing the localization of VLC system in military systems, underwater systems, and indoor systems like hospitals, hotels, libraries and malls.

Keywords Visible light communication (VLC) · Kalman filter (KF) · Deep learning (DL) · Received signal strength (RSS)
1 Introduction

VLC represents a revolution in the communication systems (Naveed et al. 2015). Indeed, it provides high data transmission rates and bandwidth along with illumination in the indoor environment. In addition, compared with other techniques including radio frequency (RF) systems, VLC has a high potential in positioning. Specifically, RF techniques have lack positioning accuracy and coverage and high interference problems (Matheus et al. 2019 and Mousa et al. 2018).

The use of received signal strength (RSS) signals for indoor localization is a promising technology. Localization has recently become increasingly granular in the commercial and scientific spheres. It encompasses a wide range of monitoring, surveillance, and tracking applications (Akter et al. 2018). In general, there are two types of localization techniques: range-based and range-free (Singh et al. 2015). The range-based schemes, on average, achieve more accuracy than range-free methods. Range-based localization is used in a variety of technologies, including time of arrival (TOA), angle of arrival (AOA), and RSS approaches. Both TOA and AOA approaches deliver great accuracy at a significant cost and complexity. The RSS approach provides moderate accuracy at a cheap cost (Koyuncu et al. 2010). The average RSS-based positioning strategy reduces the obstacles in indoor localization as compared to traditional localization approaches that employ raw RSS signals.

Extended Kalman filter (KF) in the positioning system has been demonstrated in (Vatansever et al. 2017; Zhitian et al. 2017; Eroglu et al. 2019; Shawky et al. 2020). The KF can be used in VLC localization for increasing the accuracy of the system. Convolutional neural network (CNNs) are used to model problems involving spatial image inputs and provide precise image classification results.

The DL models have recently been successfully used in a wide range of data-intensive applications, including robotics, tracking, navigation, object recognition, medical diagnosis, image processing (Hossain et al. 2019). The difficulty of training the DL models and the vast amount of training data are the main challenges of DL models -based localization systems (Yasir et al. 2015).

The localization method introduced in (Hoang et al. 2019) is based on the advantages of DL models. These technologies increase the performance of localization and produced significant localization results for real-time implementation.

A hyper-parameter (HP) is a parameter in machine learning that must be fixed before the training process can begin. As a result, unlike the value of parameters (e.g., weights) that may be taught during the training process, HPs (e.g., learning rate, batch size, and number of hidden nodes) cannot be learned during the learning process. HPs can affect the quality of the model produced by the training process as well as the algorithm’s time and memory requirements (Yang et al. 2020). As a result, HP must be fine-tuned to provide the best results for a given situation.

The impact of training techniques in an artificial neural network (ANN) equalizer in VLC systems employing a nature light source was investigated by (Chaleshtori et al. 2020). In Musumti et al. (2018), authors looked on the design and implementation of machine learning-based demodulation algorithms in VLC systems physical layer. Irshad et al. created a decision tree approach for indoor localization in VLC networks and compared it to various machine learning classifiers (Irshad et al. 2021). Alonso-González et al. introduced an indoor fingerprinting positioning estimate method depending on DL models to predict the device position in a 3D environment (González et al. 1998).
This paper aims to improve the performance of localization for VLC indoor systems by DL models that are used as prediction techniques to estimate a two-dimensional positioning system. The proposed system uses averaging RSS to estimate the Cartesian coordinates \((x, z)\). The received signal power is used as the DL models input. In the averaging method, the position of the receiver is estimated utilizing the RSS method for several times for number of samples and then, proposed technique take averaging for over samples. In a previous work, E. Shawky et al. worked on improving the visible light communication localization system using Kalman filtering with averaging, without deep learning. In our work, we use the main DL models to enhance the performance of localization in indoor systems, leading to more accurate and low-cost indoor localization technology. The KF is used to predict the power of receiver for certain number of samples, using the RSS method in averaging the received power. The accuracy of the positioning systems could be increased by using KF. KF algorithm is applied by adjusting the values of KF parameters to the user to include the information signal in positioning technique. The proposed methods are analyzed in a mathematical form, considering both NLoS for first-reflection and LoS propagation. Moreover, an HP approach based on Bayesian optimization is applied to improve our frame work performance.

The accuracy is the main factor to evaluate the performance of the proposed techniques, in addition to AUC, Se, Pr and F1-score, RMSE, training and testing time. Our proposed system is cheap and is featured with high performance and low computational complexity that achieves the hardware feasibility of the system.

The main contributions of the paper are summarized as follows:

We propose the design and analysis of an HP approach based on Bayesian optimization for indoor localization using DL models.

Two techniques are used for localizing the real track of receiver: RSS averaging technique and KF with average RSS technique in both LoS and NLoS links.

When compared to standard localization techniques, the proposed HP-RSS-KF-LoS- DL models methodology achieves a greater localization accuracy and reduced error.

This paper is structured as follows. The methodology of our framework is described in Sect. 2. The obtained results are displayed and discussed in Sec. 3 to evaluate the performance and robustness of the system. Finally, Sect. 4 is devoted to the main conclusions.

## 2 Methodology

### 2.1 Indoor model

In a typical room, we introduce the optical indoor VLC for both NLoS and LoS propagation. We consider 4 LEDs transmitters at ceiling, located at \(T_{x,i} = (x_i, y_i, z_i)\), \(i \in \{1, 2, 3, 4\}\), and one receiver as photodetector (PD), at \(R_x = (x_0, y_0, z_0)\).

#### 2.1.1 LoS link

\(H^i_{\text{LoS}}\) is the optical gain link of LoS for \(i\) LED to the PD and can be expressed as (Ghassemlooy et al. 2013)

\[
H^i_{\text{LoS}} = \frac{m + 1}{2\pi d_i^2} \cos^n(\psi_i) A_R \cos(\varphi_i) T_s(\varphi_i) g(\varphi_i)
\]  

(1)
where \( m \) represents the Lambertian order, \( \psi_i \) is the incidence angle, \( \varphi_i \) is the irradiance angle, \( d_i \) is the distance between the receiver and transmitter \( i \), \( T_s(\cdot) \) and \( g(\cdot) \) are, respectively, the gains of the optical filter and concentrator at the receiver (assumed as unity gain), and \( A_R \) is the effective area of the PD.

### 2.1.2 NLoS link

For the gain of the first reflection for NLoS, and the reflected point at \( p = (x, y, z) \), the gain for transmitter \( i \) can be obtained as (Huang et al. 2017)

\[
H_{iNLOS}^p = \frac{m + 1}{2\pi D_{ip,1}^2 D_{p,2}^2} \cos^m(\varphi_i) \cos^2(\varphi_i) \cos(\varphi_p) T_s(\varphi_p) g(\varphi_p) A_R
\]

where the \( H_{iNLOS} \) represents the gain of NLoS given by summing the reflectors for all the four walls of the room (Shchekotov 2014 and Welch et al. 2006), \( D_{ip,1} \) represents the farness between reflected point \( p \) and the transmitter, \( D_{p,2} \) represents the distance between transmitter and the reflected point \( p \), and both of \( \beta_p \) and \( \alpha_p \) are the irradiance and incidence angles at the reflection point on the wall, respectively. The receiver \( R_{ip} \), \( \varphi_{ip} \), and \( \theta_p \) are the NLoS irradiance and incidence angles related to \( p \), respectively. \( \rho \) is the reflectivity of the wall and \( dAp \) expresses the reflected area for \( p \) on the wall. Figure 1 shows the LoS/ NLoS channel model for indoor VLC system.

### 2.2 Localization method utilizing averaging RSS technique

The traditional trilateration localization method is applied to obtain the receiver location, by using the RSS technique from 3 LEDs transmitters (Teruyama et al. 2013). The approach is averaging the predicted receiver location over a specific value of the estimations to decrease the error of the localization. Figure 2 shows the block diagram that demonstrates this proposal.

Using Eq. (1) and the RSS technique, the received LoS power from transmitter \( i \in \{1, 2, 3, 4\} \) can be written as

![Fig. 1 LoS/NLoS configurations in indoor VLC system](image-url)
where $\text{PT}_i$ represents the power of transmitted $i^{th}$ LED.

Here, we assumed $\varphi_i = \psi_i$, which is calculated in (Huang et al. 2017) as

$$\cos(\varphi_i) = \frac{V}{d_i}$$

where $V$ is the height between transmitter and receiver and is assumed constant. The distance between transmitter $i$ and receiver can be obtained as (Shawky et al. 2020)

$$d_i = \sqrt{(m + 1)V^{m+1}A_RPT_i}{2\pi PR_i}$$

The total power, $PR_i$, collected at the receiver is obtained with considering the effect of NLoS path by modifying Eq. (3) to be:

$$PR_i = \left(H_{\text{LoS}} + H_{\text{NLoS}}^{i}\right)PT_i$$

### 2.3 Localization using KF in conjunction with averaging

The KF is used to enhance the prediction of the receiver localization. First step, KF estimates number N of samples for the predicted received powers. Secondly, the averaging of the estimated power is computed. Utilizing the average of power estimation, the estimated location of the receiver is calculated using the RSS method. The block diagram of KF with averaging technique is shown in Fig. 3.

The flowchart, in Fig. 4 illustrates the stages of utilizing KF with average (AVG) technique. The KF algorithm recursively estimates the state of the variables in this system in two phases; prediction and measurement (Chen et al. 2021 and Bo Liu 2021).
2.4 Kalman algorithm

The channel is modeled to be as an auto-regressive (AR) process in the state space model. The AR models and processes operate with the premise past values taking the effect of current values. The scheme depends on the idea of using KF to enhance the accuracy of the estimation. In the KF, the state vector is denoted as $x$. This vector state measures the power received and number of samples that are utilizing in the process, depending on the estimation at $k - 1$, and the state $x_{k-1/k-1}$. The following $k$ of the dynamics system, $x_{k/k-1}$, is calculated in predict and measurement stages as follows.

First: Predict step:

$$ x_{k/k-1} = F_k x_{k-1/k-1} + v_k $$  \hspace{1cm} (7)

where $F_k$ is the transition state matrix and $v_k$ is the process white noise. The corresponding state for the matrix covariance is given by (E. Shawky et al. 2020)

$$ P_{k/k-1} = F_k P_{k-1/k-1} F_k^T + Q_k $$  \hspace{1cm} (8)

where $Q_k$ represents the process noise of the covariance.

Measuring step:

The updated variable state, $x_{k/k}$, and updated covariance state $P_{k/k-1}$ can, respectively, be represented by

$$ x_{k/k} = x_{k/k-1} + K_k (z_k - H_k x_{k/k-1}) $$  \hspace{1cm} (9)

$$ P_{k/k} = (I - K_k H_k) P_{k/k-1} $$  \hspace{1cm} (10)

where $K_k$ represents the Kalman gain, and $H_k$ denotes the observation model given by:

$$ K_k = P_{k/k-1} H_k^T S_k^{-1} $$  \hspace{1cm} (11)

Here, $z_k$ is the measurement vector given by

$$ z_k = x_k^T + w_k $$  \hspace{1cm} (12)

where $w_k$ is the measurement noise.
Fig. 4 Flowchart for combining the KF algorithm and AVG method

Also, $S_k$ represents the innovation matrix, which is correlated with the covariance of the state variables to measurement vector as:

$$ S_k = \left( H_k P_{k-1} H_k^T \right) + R_k $$  

where $R_k$ is the covariance of the observation noise.
2.5 Proposed DL models based indoor positioning system

The DL models incorporates the benefits of optimization approach to enhance the system performance (Chen et al. 2021). We build a hybrid network, HP-RSS-KF-LoS-DLM, which employs an optimized DL models. We provide our technique for getting the optimal HP-RSS-KF-LoS-DLM configurations for target localization in this section. As previously mentioned, the first stage is data set gathering utilizing MATLAB. The second stage involves the optimization approach based on Bayesian optimization and DL models with the predicted data set using Python software. The proposed localization method identifies the user position using different strategies: average RSS based on DL models, average RSS with KF based on DL models, and the Bayesian approach hybrid with the DL models for optimization process to enhance our frame work performance. The proposed HP-RSS-KF-LOS-DL models based localization technique is illustrated in Fig. 5. The suggested system starts with gathering the RSS data; a normalization technique is used to RSS data to center it to a mean value, $\mu$, to enhance and minimize redundancy. Finally, three different DL models, Yolo v3 (Adarsh et al. 2020), EfficientNetB3 (Ganesh et al. 2022) and DenseNet121 (Nandhini et al. 2022), are used to train the data.

2.6 DL models with hyper-parameter optimization

As previously stated, the choice of HPs affects the performance of a model, and determining the ideal value for each HP is not easy. As a result, we apply Bayesian optimization to adjust the suitable HP for the used DL models to check if it brings any benefit. Both the Adam optimizer (Jais et al. 2019) and the Stochastic Gradient Descent (SGD) optimizer (Ratre 2020) are subjected to HP tuning. The best combination of Adam optimizer with yolo v3 is a learning rate value of 0.001954, beta 1 value of 0.854 which gives a loss metric 2.34, while the best combination of SGD optimizer is a learning rate value of 0.01821, and a tuned momentum value of 0.962 which gives a loss metric of 2.01 as shown in Table 1. All values represented in Table 1 are obtained based on the authors trials for the different algorithms to get the optimum performance. Moreover, a batch size of 64/32/16 and number of epoch 100, 150, 200, are used for Yolo v3, EfficientNetB3 and DenseNet121, respectively.

![Diagram](image.png)

Fig. 5 Proposed DL models based indoor positioning system
| Setup               | Default learning rate | tuned learning rate | default beta 1 | tuned beta 1 | default momentum | tuned momentum | loss metric with default | loss metric with tuned parameters |
|---------------------|-----------------------|--------------------|----------------|--------------|-----------------|---------------|--------------------------|----------------------------------|
| Yolo v3 + Adam      | 0.001                 | 0.001954           | 0.9            | 0.854        | NA              | NA            | 2.32                     | 2.34                              |
| Yolo v3 + SGD       | 0.01                  | 0.01821            | NA             | 0.9          | 0.962           | 0.94          | 2.12                     | 2.01                              |
| EfficientNetB3 + Adam | 0.001               | 0.001362           | 0.9            | 0.821        | NA              | NA            | 2.41                     | 2.22                              |
| EfficientNetB3 + SGD | 0.01                 | 0.01524            | NA             | 0.9          | 0.874           | 0.867         | 2.21                     | 2.14                              |
| DenseNet121 + Adam  | 0.001                 | 0.001698           | 0.9            | 0.832        | NA              | NA            | 2.45                     | 2.31                              |
| DenseNet121 + SGD   | 0.01                  | 0.01247            | NA             | 0.9          | 0.865           | 0.861         | 2.31                     | 2.11                              |
It is observed that the loss metric changes from 2.22 to 2.34 for the Adam optimizer and from 2.01 to 2.14 for SGD optimizer. This indicates that the SDG optimizer is better than Adam optimizer in these datasets.

2.7 Localization process

In Fig. 5, the localization method first employs the training RSS to train the DL models. Following model weight initialization, the system employs testing RSS for localization. The DL model gathers the RSS information from the spatial domain and then start the testing phase. The DL models predict user locations by using information in the temporal domain. The DL models output is the user $x$ and $y$ position values.

3 Results and discussion

3.1 Evaluation metrics

In order to achieve the superb robustness of proposed technique, various DL models are utilized. Here, we evaluate the performance of indoor localization for several DL models based on different strategies.

The metric evaluation depends mainly on calculating four parameters: the number of true positives (TP), true negatives (TN), false negatives (FN), and false positives (FP). The classification performance is identified in terms of ACC, $Se$ or recall, $Pr$, F1-score, $AUC$, RMSE and computational time. The ACC is used to evaluate the rate of correct classification, $Pr$ is the positive predictive value that matches the original value, and $Se$ is the true positive values. The F1-score is the harmonic mean of $Pr$ and $Se$. It represents a more generalized form for balancing both $Pr$. The $AUC$ measures the entire two-dimensional area underneath the entire ROC curve. The RMSE is an error metric that obtains a cumulative estimate of error. It is evaluated as the square root of the arithmetic mean of squares of error in our dataset. It provides an aggregate measure of performance across all possible classification thresholds. All these metrics are defined as follows (Muschelli 2020)

$$ACC = \frac{TP + TN}{TP + FN + TN + FN}$$  \hspace{1cm} (14)

$$Pr = \frac{TP}{TP + FP}$$ \hspace{1cm} (15)

$$Se = \frac{TP}{TP + FN}$$ \hspace{1cm} (16)

$$F1 = \frac{2(Pr \times Se)}{Pr + Se}$$ \hspace{1cm} (17)

$$RMSE = \sqrt{\frac{1}{k} \sum_{j=1}^{k} \left( \hat{x}_j - x_j \right)^2 + \left( \hat{y}_j - y_j \right)^2}$$ \hspace{1cm} (18)

where $(\hat{x}_j, \hat{y}_j)$ and $(x_j, y_j)$ refer to $j^{th}$ estimated and true locations, respectively, and $k$ is the number of dataset points.

Springer
Now, simulation results for the proposed algorithm are presented and compared with that of the traditional systems. The main parameters used in the simulations for the VLC link are listed in Table 2.

We start our simulation with HP tweaking to assess the performance and correctness of the proposed DL models. In HP tuning, we train the model with various HP settings.

| Table 2 | Simulation parameters (Shawky et al. 2020) |
|---------|-------------------------------------------|
| Parameter | Value | |
| Dimensions of the room | $5 \times 5 \times 3$ m$^3$ | |
| Number of transmitters | 4 | |
| Total transmitted power per Tx | 30 W | |
| LEDs' locations | $(1.25,1.25,3),(1.25,3.75,3),(3.75,1.25,3),(3.75,3.75,3)$ m | |
| FoV of PD | 70o | |
| Signal to noise ratio (SNR) | 20 dB | |
| PD active area | $1$ cm$^2$ | |
| Wall reflectivity $\rho$ | 0.8 | |
| Number of samples | 50 (0.1 ms between each sample) | |
| Range of receiver in room | 1–3.5 m over both x–y axes | |

| Table 3 | Time for all models to be trained and tested |
|---------|-------------------------------------------|
| Model | Training time (s) | Testing time (s) |
| Average RSS technique | | |
| Yolo V3 | 150 | 0.37 |
| EfficientNetB3 | 120 | 0.20 |
| DenseNet121 | 200 | | |
| Average RSS technique + KF | | |
| Yolo V3 | 170 | 0.32 |
| EfficientNetB3 | 200 | 0.27 |
| DenseNet121 | 220 | | |
| Average Technique + Bayesian optimization | | |
| Yolo V3 | 185 | 0.36 |
| EfficientNetB3 | 165 | 0.29 |
| DenseNet121 | 225 | | |
| Average Technique + KF + Bayesian optimization | | |
| Yolo V3 | 140 | 0.29 |
| EfficientNetB3 | 115 | 0.23 |
| DenseNet121 | 195 | 0.34 |
to find the optimal values that offer the best model performance. The HP values utilized in the proposed model are previously summarized in Table 1.

Table 3 illustrates that the proposed HP-RSS-KF-LoS- DL models model training and testing duration is less than the other models. However, when examining the suggested DL models localization capability, these computational durations are reasonable for indoor localization (Chatterjee et al. 2019).

Table 4 displays different strategies, average RSS technique, average RSS technique with KF, average RSS technique based on DL models and average technique with KF based on HP-RSS-KF-LoS- DL models with LoS and NLoS.

It is observed from the experimental results that the HS-RSS-KF-LoS- DL models, YoloV3, achieves the best performance with 99.99% accuracy, 99.98% AUC, 98.88% sensitivity, 98.98% precision, 99.97% F1-score and 0.112 RMSE. We would like to notify that the obtained RMSE is related to Yolo V3 model that is concluded to have the superior probabilities of performance.

Our proposed frame work is compared with others in the literature as introduced in Table 5. The results reveal that our proposed framework achieves superior performance in ACC, Pr, AUC, Se, F1-score and RMSE. The comparison depicts an accuracy enhancement of 1.29% to 4.04% and an RMSE enhancement of 3.89% to 21.59%. The other evaluation indicators are also better in our work, with less percentage ratios. We note also that, sensitivity is not found in literature. All of this gives a superiority of our work.

4 Conclusion

In this paper, we introduced multi-techniques to enhance the localization using RSS average technique and KF with average RSS technique with both effects of LoS and NLoS links. The output of these techniques \((x, y)\) of the estimated track of the receiver was the input of DL models -based localization system for indoor VLC system. The viability of employing the average RSS hybrid with KF for indoor localization is tested experimentally. It is observed that the HP optimization plays an important role in improving the performance of our proposed framework. The suggested HP-RSS-KF-LoS-DLM-based localization system achieves a reasonable localization accuracy for indoor localization, according to the findings of our trials.

Compared with previously published work, our proposed frame work is found to have better performance. It achieves accuracy of 99.99% accuracy, 99.98% AUC, 98.88% sensitivity, 98.98% precision, 99.97% F1-score and 0.112 RMSE and 0.29 s for testing time.

Accordingly, our proposed system is featured with high accuracy, low complexity and small error distance at very small training time. This makes it appropriate to be included in mobile devices. Therefore, the proposed system can be scaled and applied with any VLC environment just by estimating their RSS values. Moreover, identical idea can be utilized with building 3D localization system.
Table 4 Performance of different strategies based on DL models

Average RSS technique + Kalman filter + light of Sight + DL models + Bayesian Optimization (HP)
(RSS + KF + LoS + DLM + HP)

| Model                | Accuracy % | AUC % | Sensitivity % | Precision % | F1-score % | RMSE (cm) |
|----------------------|------------|-------|---------------|-------------|------------|-----------|
| Yolo V3              | 99.99      | 99.98 | 98.88         | 98.98       | 99.97      | 0.112     |
| EfficientNetB3       | 97.99      | 97.99 | 97.98         | 97.88       | 97.99      | 0.214     |
| DenseNet121          | 96.49      | 96.59 | 96.78         | 96.77       | 96.87      | 0.241     |

RSS technique

| Model                | Accuracy % | AUC % | Sensitivity % | Precision % | F1-score % | RMSE (cm) |
|----------------------|------------|-------|---------------|-------------|------------|-----------|
| Yolo V3              | 85.74      | 85.89 | 85.87         | 85.77       | 85.88      | 0.234     |
| EfficientNetB3       | 84.65      | 84.45 | 84.52         | 84.41       | 84.21      | 0.254     |
| DenseNet121          | 82.35      | 82.85 | 82.45         | 82.58       | 82.64      | 0.261     |

RSS technique + DL models

| Model                | Accuracy % | AUC % | Sensitivity % | Precision % | F1-score % | RMSE (cm) |
|----------------------|------------|-------|---------------|-------------|------------|-----------|
| Yolo V3              | 87.87      | 87.74 | 87.65         | 87.85       | 87.96      | 0.245     |
| EfficientNetB3       | 85.75      | 85.78 | 85.79         | 85.88       | 85.76      | 0.267     |
| DenseNet121          | 83.99      | 83.88 | 83.87         | 83.89       | 83.77      | 0.274     |

HP-RSS technique + DL models

| Model                | Accuracy % | AUC % | Sensitivity % | Precision % | F1-score % | RMSE (cm) |
|----------------------|------------|-------|---------------|-------------|------------|-----------|
| Yolo V3              | 89.98      | 89.84 | 89.98         | 89.78       | 89.77      | 0.267     |
| EfficientNetB3       | 86.85      | 86.99 | 86.85         | 86.88       | 86.96      | 0.321     |
| DenseNet121          | 84.78      | 84.88 | 84.87         | 84.89       | 84.87      | 0.365     |

Average RSS technique

| Model                | Accuracy % | AUC % | Sensitivity % | Precision % | F1-score % | RMSE (cm) |
|----------------------|------------|-------|---------------|-------------|------------|-----------|
| Yolo V3              | 90.21      | 90.41 | 90.54         | 90.45       | 90.47      | 0.321     |
| EfficientNetB3       | 88.65      | 88.54 | 88.78         | 88.98       | 88.77      | 0.357     |
| DenseNet121          | 86.65      | 86.87 | 86.85         | 86.78       | 86.96      | 0.378     |

Average RSS technique + DL models

| Model                | Accuracy % | AUC % | Sensitivity % | Precision % | F1-score % | RMSE (cm) |
|----------------------|------------|-------|---------------|-------------|------------|-----------|
| Yolo V3              | 94.84      | 94.98 | 94.87         | 94.88       | 94.89      | 0.241     |
| EfficientNetB3       | 92.68      | 92.77 | 92.78         | 92.79       | 92.87      | 0.312     |
| DenseNet121          | 90.65      | 90.45 | 90.78         | 90.63       | 90.78      | 0.351     |

Average RSS technique + DL models + HP

| Model                | Accuracy % | AUC % | Sensitivity % | Precision % | F1-score % | RMSE (cm) |
|----------------------|------------|-------|---------------|-------------|------------|-----------|
| Yolo V3              | 95.87      | 95.84 | 95.73         | 95.77       | 95.78      | 0.201     |
| EfficientNetB3       | 93.89      | 93.88 | 93.99         | 93.99       | 93.87      | 0.325     |
| DenseNet121          | 91.86      | 91.89 | 91.88         | 91.89       | 91.78      | 0.345     |

Kalman filter technique

| Model                | Accuracy % | AUC % | Sensitivity % | Precision % | F1-score % | RMSE (cm) |
|----------------------|------------|-------|---------------|-------------|------------|-----------|
| Yolo V3              | 88.54      | 88.43 | 88.52         | 88.42       | 88.62      | 0.263     |
| EfficientNetB3       | 84.41      | 84.88 | 84.79         | 84.85       | 84.88      | 0.301     |
| DenseNet121          | 82.87      | 82.89 | 82.88         | 82.87       | 82.89      | 0.341     |

Kalman filter technique + DL models

| Model                | Accuracy % | AUC % | Sensitivity % | Precision % | F1-score % | RMSE (cm) |
|----------------------|------------|-------|---------------|-------------|------------|-----------|
| Yolo V3              | 89.99      | 89.98 | 89.88         | 89.87       | 89.87      | 0.278     |
| EfficientNetB3       | 85.98      | 85.99 | 85.89         | 85.98       | 85.97      | 0.299     |
| DenseNet121          | 83.99      | 83.88 | 83.98         | 83.88       | 83.88      | 0.321     |

Kalman filter technique + DL models + HP

| Model                | Accuracy % | AUC % | Sensitivity % | Precision % | F1-score % | RMSE (cm) |
|----------------------|------------|-------|---------------|-------------|------------|-----------|
| Yolo V3              | 92.87      | 92.87 | 92.86         | 92.82       | 92.88      | 0.199     |
| EfficientNetB3       | 90.78      | 90.84 | 90.89         | 90.88       | 90.95      | 0.214     |
| DenseNet121          | 89.89      | 89.78 | 89.88         | 89.96       | 89.87      | 0.265     |

Average RSS technique + KF + LoS

| Model                | Accuracy % | AUC % | Sensitivity % | Precision % | F1-score % | RMSE (cm) |
|----------------------|------------|-------|---------------|-------------|------------|-----------|
| Yolo V3              | 94.88      | 94.99 | 94.87         | 94.78       | 94.96      | 0.287     |
| EfficientNetB3       | 92.65      | 92.96 | 92.69         | 92.78       | 92.87      | 0.354     |
Table 4 (continued)

Average RSS technique + Kalman filter + light of Sight + DL models + Bayesian Optimization (HP)  
(RSS + KF + LoS + DLM + HP)

| Model                              | Accuracy % | AUC % | Sensitivity % | Precision % | F1-score % | RMSE (cm) |
|------------------------------------|------------|-------|---------------|-------------|------------|-----------|
| DenseNet121                         | 90.65      | 90.68 | 90.78         | 90.88       | 90.88      | 0.324     |
| Average RSS technique + KF + LoS + NLoS |            |       |               |             |            |           |
| Yolo V3                            | 92.58      | 92.54 | 92.57         | 92.65       | 92.86      | 0.302     |
| EfficientNetB3                     | 90.82      | 90.76 | 90.89         | 90.99       | 90.87      | 0.387     |
| DenseNet121                         | 89.88      | 89.78 | 89.87         | 89.88       | 89.87      | 0.398     |
| Average RSS technique + KF + LoS + DL models |            |       |               |             |            |           |
| Yolo V3                            | 96.99      | 96.88 | 96.87         | 96.77       | 96.88      | 0.341     |
| EfficientNetB3                     | 93.85      | 93.98 | 93.87         | 93.98       | 93.89      | 0.412     |
| DenseNet121                         | 92.99      | 92.88 | 92.89         | 92.89       | 92.98      | 0.458     |
| Average RSS technique + KF + LoS + NLoS + DL models |            |       |               |             |            |           |
| Yolo V3                            | 94.87      | 94.78 | 94.69         | 94.77       | 94.65      | 0.298     |
| EfficientNetB3                     | 93.96      | 93.88 | 93.87         | 93.89       | 93.78      | 0.354     |
| DenseNet121                         | 92.96      | 92.98 | 92.88         | 92.87       | 92.88      | 0.374     |
| Average RSS technique + KF + LoS + NLoS + DL models + HP |            |       |               |             |            |           |
| Yolo V3                            | 97.88      | 97.98 | 97.89         | 97.87       | 97.89      | 0.265     |
| EfficientNetB3                     | 94.54      | 94.86 | 94.69         | 94.88       | 94.87      | 0.321     |
| DenseNet121                         | 92.85      | 92.86 | 92.87         | 92.87       | 92.89      | 0.385     |
| Average RSS technique + KF + DL models |            |       |               |             |            |           |
| Yolo V3                            | 95.45      | 95.59 | 95.54         | 95.56       | 95.65      | 0.254     |
| EfficientNetB3                     | 93.54      | 93.68 | 93.48         | 93.44       | 93.45      | 0.321     |
| DenseNet121                         | 91.89      | 91.88 | 91.98         | 91.87       | 91.88      | 0.369     |
| Average RSS Technique + KF + DL models + HP |            |       |               |             |            |           |
| Yolo V3                            | 97.12      | 97.99 | 97.54         | 97.66       | 97.87      | 0.335     |
| EfficientNetB3                     | 94.99      | 94.68 | 94.87         | 94.99       | 94.96      | 0.421     |
| DenseNet121                         | 93.66      | 93.65 | 93.87         | 93.78       | 93.88      | 0.487     |
### Table 5  Comparison between our framework and others in the literature

| Model                                                                 | Accuracy % | AUC %  | Sensitivity % | Precision % | F1- score % | RMSE (cm) |
|-----------------------------------------------------------------------|------------|--------|---------------|-------------|-------------|-----------|
| Average RSS technique + KF + light of sight + DL models + LSTM (RSS-KF-LoS-DLM-LSTM) based on ResNet50 (Present Work) | 99.99      | 99.98  | 98.88         | 98.98       | 99.97       | 0.112     |
| Neural Networks (Xue et al. 2018)                                     | 98.7       | 99.1   | NA            | 98.7        | 98.7        | NA        |
| Enhanced J48 tree (Chen et al. 1998)                                  | 98.15      | NA     | NA            | NA          | NA          | NA        |
| ICT-Net (Ghonim et al. 2021)                                          | 97.14      | NA     | NA            | NA          | NA          | NA        |
| Convolutional Neural Networks (Irshad et al. 2021)                     | 97.81%     | NA     | NA            | NA          | NA          | NA        |
| Deep Learning (Chatterjee et al. 2019)                                | 95.95%     | NA     | NA            | NA          | NA          | NA        |
| k-NN (Z. Chen et al. 2020)                                            | 99.33%     | NA     | NA            | NA          | NA          | NA        |
| RSS + TDOA (Turgut al. 2019)                                          | NA         | NA     | NA            | NA          | NA          | 5.81      |
| Fingerprint (Tran et al. 2019)                                        | NA         | NA     | NA            | NA          | NA          | 21.7      |
| RSS (Karmy et al. 2020)                                               | NA         | NA     | NA            | NA          | NA          | 4         |
Funding Open access funding provided by The Science, Technology & Innovation Funding Authority (STDF) in cooperation with The Egyptian Knowledge Bank (EKB).

Declarations

Conflict of interest Authors also declare no conflict of interest.

Ethical approval All authors of this paper have read and approved the final version submitted.

Human or animal rights All authors of this research paper have directly participated in the planning, execution, or analysis of this study.

Consent for publication The contents of this manuscript have not been copyrighted or published previously: The contents of this manuscript are not now under consideration for publication elsewhere; The contents of this manuscript will not be copyrighted, submitted, or published elsewhere, while acceptance by the Journal is under consideration; There are no directly related manuscripts or abstracts, published or unpublished, by any authors of this paper.

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/.

References

Adarsh, P., Rathi, P., Kumar, M.: “YOLO v3-Tiny: Object detection and recognition using one stage improved model,” 2020 IEEE 6th International Conference on Advanced Computing and Communication Systems (ICACCS), pp. 687–694, (2020)

Akter, M., Rahman, M.O., Islam, M.N., Hassan, M.M., Alsanad, A., Sangaihah, A.K.: Energy-efficient tracking and localization of objects in wireless sensor networks. IEEE Access 6, 17165–17177 (2018)

Chaleshtori, Z.N., Burton, A., Zvanovec, S., Ghassemlooy, Z.F., Chvojka, P.: Comprehensive optical and electrical characterization and evaluation of organic light-emitting diodes for visible light communication. Optic. Eng. (2020). https://doi.org/10.1117/1.OE.59.4.046106

Chatterjee, B., Poullis, C., “On building classification from remote sensor imagery using deep neural networks and the relation between classification and reconstruction accuracy using border localization as proxy.” In 16th Conference on Computer and Robot Vision (CRV) (IEEE, 2019), Kingston, QC, Canada, pp. 41–48, (2019)

Chen, Y.C., Li, D.C.: Selection of key features for PM2. 5 prediction using a wavelet model and RBF-LSTM. Appl. Intell. 51(4), 2534–2555 (2021)

Chen, S.H., Hwang, S.H., Wang, Y.R.: An RNN-based prosodic information synthesizer for Mandarin text-to-speech. IEEE Trans. Speech Audio Process. 6(3), 226–239 (1998)

Chen, Z., AlHajri, M.I., Wu, M., Ali, N.T., Shubair, R.M.: A novel real-time deep learning approach for indoor localization based on RF environment identification. IEEE Sensors Lett. 4(6), 1–4 (2020)

Eroglu, Y. S., Erden, F., Guvenc, I.: "Adaptive Kalman tracking for indoor visible light positioning." MILCOM 2019-2019 IEEE Military Communications Conference (MILCOM), pp. 331–336, (2019)

Ganesh, M., Dulam, S., Venkatasubbu, P.: Diabetic retinopathy diagnosis with inception ResNetV2, Xception, and efficient NetB3. In: Raje, R.R., Hussain, F., Kannan, R.J. (eds.) Artificial Intelligence and Technologies, pp. 405–413. Springer, Singapore (2022)

Ghassemlooy, Z., Popoola, W., Rajbhandari, S.: Optical Wireless Communications: System and Channel Modelling with MATLAB. CRC Press, New York, USA (2013)
Ghonim, A.M., Salama, W.M., El-Fikky, A.E.R.A., Khalaf, A.A., Shalaby, H.M.: Underwater localization system based on visible-light communications using neural networks. Appl. Opt. 60(13), 3977–3988 (2021)

González, A., Steffen, K.L., Lynch, J.P.: Light and excess manganese: implications for oxidative stress in common bean. Plant Physiol. 118(2), 493–504 (1998)

Hoang, M.T., Yuen, B., Dong, X., Lu, T., Westendorp, R., Reddy, K.: Recurrent neural networks for accurate RSSI indoor localization. IEEE Internet Things J. 6(6), 10639–10651 (2019)

Hossain, S., Lee, D.J.: Deep learning-based real-time multiple-object detection and tracking from aerial imagery via a flying robot with GPU-based embedded devices. Sensors (2019). https://doi.org/10.3390/s19153371

Huang, C., Zhang, X.: “LoS-NLoS identification algorithm for indoor visible light positioning system,” In 2017 20th International Symposium on Wireless Personal Multimedia Communications (WPNC), pp. 575–578 (2017)

Irshad, M., Liu, W., Wang, L., Khalil, M.U.R.: Cogent machine learning algorithm for indoor and underwater localization using visible light spectrum”. Wirel. Pers. Commun. 116(2), 993–1008 (2021)

Jais, I.K.M., Ismail, A.R., Nisa, S.Q.: Adam optimization algorithm for wide and deep neural network. Knowl. Eng. Data Sci. 2(1), 41–46 (2019)

Karmy, M., El-Sayed, S., Zekey, A.: Performance enhancement of an indoor localization system based on visible light communication using RSSI/TDOA hybrid technique. J. Commun. 15(5), 379–389 (2020)

Koyuncu, H., Yang, S.H.: “A survey of indoor positioning and object locating systems. IJCSNS Int. J. Comput. Sci. Netw. Secur. 10(5), 121–128 (2010)

Li, Z., Feng, L., Yang, A.: Fusion based on visible light positioning and inertial navigation using extended kalman filters. IEEE Sensors 17, 10–93 (2017)

Liu, B.: Understanding the loss landscape of one-hidden-layer ReLU networks. Knowl. Based Syst. (2021). https://doi.org/10.1016/j.knosys.2021.106923

Matheus, L.E.M., Vieira, A.B., Vieira, L.F.M., Vieira, M.A.M., Gnawali, O.: Visible light communication: concepts, applications and challenges. IEEE Commun. Surv. Tutorials 21, 3204–3237 (2019)

Mousa, F., Almaadeed, N., Busawon, K., Bouridane, A., Binns, R., Elliot, I.: Indoor visible light communication localization system utilizing received signal strength indication technique and trilateration method. Optic. Eng. (2018). https://doi.org/10.1117/1.OE.57.1.016107

Muschelli, J.: ROC and AUC with a binary predictor, a potentially misleading metric. J. Classif. 37(3), 696–708 (2020)

Musumeci, F., Rottondi, C., Nag, A., Macaluso, I., Zibar, D., Ruffini, M., Tornatore, M.: An overview on application of machine learning techniques in optical networks. IEEE Commun. Surv. Tutorials 21(2), 1383–1408 (2018)

Nandhini, S., Ashokkumar, K.: An automatic plant leaf disease identification using DenseNet-121 architecture with a mutation-based henry gas solubility optimization algorithm. Neural Comput. Appl. (2022). https://doi.org/10.1007/s00521-021-06714-z

Naveed, H., Aqsa, N., Adeel, P.M., Tariq, I., Chau, Y.: Indoor po-sitioning using visible LED lights: a survey. ACM Comput. Surv. 48, 1–20 (2015)

Ratre, A.: Stochastic gradient descent–whale optimization algorithm-based deep convolutional neural network to crowd emotion understanding. Comput. J. 63(2), 267–282 (2020)

Shawky, E., El-Shimy, M., Mokhtar, A., El-Badawy, E.A., Shalaby, H.M.H.: Improving the visible light communication localization system using Kalman filtering with averaging. J. Opt. Soc. Am. B 37, A130–A138 (2020)

Shchekotov, M.: “Indoor localization method based on Wi-Fi trilateration technique,” Proc. 16th Conf. Fruct. Assoc. (ACP 2018) pp. 177–179 (2014)

Singh, S.P., Sharma, S.C.: Range free localization techniques in wireless sensor networks: a review. Procedia Comput. Sci. 57, 7–16 (2015)

Teruyama, Y., Watanabe, T.: Effectiveness of variable-gain Kalman Filter based on angle error calculated from acceleration signals in lower limb angle measurement with inertial sensors. Comput. Math. Methods Med. 10, 1–12 (2013)

Tran, H.Q., Ha, C.: Fingerprint-based indoor positioning system using visible light communication—a novel method for multipath reflections. Electronics 8(1), 63–66 (2019)

Turgut, Z., Üstebay, S., Zeynel Gürkaş Aydin, G., Serbış, A.: Deep learning in indoor localization using WiFi. In International Telecommunications Conference, Springer, Singapore, pp. 101–110 (2019)

Vatansever, Z., Brandt-Pearce, M.: “Visible light positioning with diffusing lamps using an extended kalman filter,” in 2017 IEEE Wireless Communications and Networking Conference (WCNC), San Francisco, CA, USA, pp. 1–6, (2017)
Welch, G., Bishop, G.: “An introduction to the Kalman filter,” Tech. Rep. 95–041, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA, (2006)
Xue, H., Huynh, D.Q., Reynolds, M.: “SS-LSTM: A hierarchical LSTM model for pedestrian trajectory prediction,” In 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), Lake Tahoe, NV, USA, pp. 1186–1194, (2018)
Yang, L., Shami, A.: On hyper-parameter optimization of machine learning algorithms: theory and practice. Neurocomputing 415, 295–316 (2020)
Yasir, M., Ho, S.W., Vellambi, B.N.: Indoor position tracking using multiple optical receivers. J. Lightwave Technol. 34(4), 1166–1176 (2015)

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

Authors and Affiliations

Wessam M. Salama1 · Moustafa H. Aly2 · Eman S. Amer3

Wessam M. Salama
wessam.salama@pua.edu.eg

Eman S. Amer
eman.shawky@alexu.edu.eg

1 Faculty of Engineering, Pharos University, Alexandria, Egypt
2 Arab Academy for Science, Technology and Marine Transport, Alexandria 1029, Egypt
3 Higher Institute for Engineering and Technology, King Mariout, Alexandria, Egypt