Indoor Space Classification Using Cascaded LSTM

Rohan Kumar Yadav, Bimal Bhattarai, Lei Jiao, Morten Goodwin, and Ole-Christoffer Granmo
Department of Information and Communication Technology,
University of Agder, 4879, Grimstad, Norway
{rohan.k.yadav, bimal.bhattarai, lei.jiao, morten.goodwin, ole.granmo}@uia.no

***Abstract*—Indoor space classification is an important part of localization that helps in precise location extraction, which has been extensively utilized in industrial and domestic domain. There are various approaches that employ Bluetooth Low Energy (BLE), Wi-Fi, magnetic field, object detection, and Ultra Wide Band (UWB) for indoor space classification purposes. Many of the existing approaches need extensive pre-installed infrastructure, making the cost higher to obtain reasonable accuracy. Therefore, improvements are still required to increase the accuracy with minimum requirements of infrastructure. In this paper, we propose an approach to classify the indoor space using geomagnetic field (GMF) and radio signal strength (RSS) as the identity. The indoor space is an open big test bed divided into different indiscernible subspace. We collect GMF and RSS at each subspace and classify it using cascaded Long Short Term Memory (LSTM). The experimental results show that the accuracy is significantly improved when GMF and RSS are combined to make distinct features. In addition, we compare the performance of the proposed model with the state-of-the-art machine learning methods.**

***Index Terms*—geomagnetic field, radio signal strength, LSTM, deep recurrent neural network**

I. Introduction

Location identification is more and more important for informational services, real time tracking, address extraction and other entertainment purposes. Global Positioning System (GPS) is limited to outdoor area, and therefore for indoor positioning, other techniques must be employed, such as BLE, GMF, UWB, and Wi-Fi [1]. Those techniques can provide, in the indoor environment, necessary information to be utilized for positioning purpose. Once the information are gathered, the position algorithms estimate the position of an object. In addition to traditional positioning algorithms, such as triangular positioning, AI based solutions become more and more popular, among which, the fingerprinting method is one of the most effective approach because of its low complexity in smartphone based localization [2]. Fingerprinting method has two phases: (1) offline phase, which is the data collection stage and (2) online phase, where the live data is compared with stored data to give final localization. Despite of having various methodologies to realize positioning, there are still room to improve the accuracy with reduced cost. In this article, we are aiming to increase positioning accuracy and reduce cost in a better manner, using the latest AI-based techniques.

RSS is the most used metrics for positioning because of its easy availability from BLE and Wi-Fi. However, it has certain drawbacks like instability due to multi-path effect and device heterogeneity. Hence, the channel state information (CSI) comes to the picture that provides much more information from multiple sub-carriers and antennas [3] [4]. Despite of its better performance, it is a device specific metrics and thus it is less practical choice for positioning as compared to RSS. There are two major types of positioning approaches: deterministic and probabilistic. The most common approach for deterministic methods are K-nearest neighbor (KNN) and its variants [5] [6]. For probabilistic method, it adopts the statistical evaluation between trained and measured RSS using Bayes rule [7]. In addition, the fusion of deterministic approach and probabilistic approach are usually employed for fine grained location extraction [8]. Although there are various measures to deal with the ambiguity nature of RSS, there still lacks the proper technique to completely discard the RSS fluctuation. Hence, deep learning has been utilized for learning the pattern in RSS to classify various location [9]. More specifically, the location classification can be enhanced using Recurrent Neural Network (RNN) by making use of sequential relation of RSS measurements [10].

Different from RSS, which is generated by pre-installed devices, GMF is a magnetic force that surrounds the earth. It is generated by earth's rotation and the movement of molten iron in the earth's core, which is omnipresent. The GMF generated are not uniform inside the building areas due to materials like iron, steel, concrete, and other equipment [11]. Due to these interference in the GMF inside buildings, it can be used as one of the technologies of location based service (LBS). Basically, GMF based LBS heavily relies on deep learning due to randomness in data unlike RSS. Many existing works that is based on magnetic field are related to robot navigation that uses Simultaneous Localization and Mapping (SLAM) [12]. The ubiquitous nature of GMF has shown good performance in the indoor landmark classification using RNN [13].

To improve the classification accuracy, there are several existing studies that adopt an integration of various sensors, such as GMF, light sensors, Inertial Measurement Unit (IMU), and visual sensors [14] [15]. Inspired by those studies, in this paper, we propose a hybrid model based on GMF data and RSS data as references, which most of the modern smart phones have, for indoor space classification using cascaded Long Short-Term Memory (LSTM). We have witnessed that LSTM is suitable for long sequence data that captures the sequential pattern from the given features [16]. Similarly, such property of LSTM is also applicable in GMF for landmark classification [17]. Therefore, in this article, we are motivated to employ those two types of information and eventually show that using RSS along with GMF can give a better distinctive feature for indoor space classification and can increase the accuracy with significant margin. In more details, we make use of cascaded LSTM that combines both unidirectional and bidirectional model of LSTM to classify the space based on variable length input sequence of GMF and RSS, which is to be detailed below.

The remaining sections of the paper are organized as follows. In Section II, the proposed models are presented. In Section III, the proposed LSMT is examined extensively. The paper concludes in Section IV.

II. Proposed Model

In this section, we will present in detail the network structure, the data collection procedure, and the training and validation procedure for the RNN based solution.

A. Neural Network Structure

Recurrent neural network (RNN) is a type of deep learning technique that is not only dependent on current input but also on the previous input. Basically, it is applicable to the scenario where the data have a sequential correlation. However, when dealing with a long sequence of data, it has a problem of vanishing and exploiting gradient. To overcome this effect, an LSTM is used which has an internal memory states that adds forget gate. This gate controls the time dependence and the effect of previous input. There are other variations, like BiRNN and BiLSTM, which not only reflect previous inputs but also consider the future inputs of a particular time frame. In this study, inspired by [17], we propose the cascaded unidirectional LSTM and bidirectional LSTM (BiLSTM) RNN model as shown in Fig. 1. The model consists the first layer of bidirectional RNN combined with a unidirectional RNN layer. The bidirectional LSTM consists of forward and backward track for learning patterns in both directions.

The Eq. (1) and Eq. (2) show the operations of forward and backward track.

$$O_n^{f1}, h_n^{f1}, i_n^{f1} = L^{f1}(i_{n-1}^{f1}, h_{n-1}^{f1}, x_n : P^{f1}), \tag{1}$$

$$O_n^{b1}, h_n^{b1}, i_n^{b1} = L^{b1}(i_{n-1}^{b1}, h_{n-1}^{b1}, x_n : P^{b1}), \tag{2}$$

where $O_n^{f1}, h_n^{f1}, i_n^{f1}$ and $O_n^{b1}, h_n^{b1}, i_n^{b1}$ are the output, the hidden state, the internal state of the current state for forward and backward LSTM track respectively. x_n is the input sequence, P is the LSTM cell parameter. The output from both tracks are combined as in Eq. (3) and forwarded into the second layer.

$$O_n^{1} = O_n^{f1} + O_{N-n+1}^{b1}. \tag{3}$$

Bidirectional RNN is followed by unidirectional RNN, which transforms data into a more abstract form and aids in learning spatial dependencies [18]. The output from the unidirectional layer is obtained using Eq. (4).

$$O_n^{l}, h_n^{l}, i_n^{l} = LSTM^{l}(i_{n-1}^{l}, h_{n-1}^{l}, O_n^{l-1}; P^{l}), \tag{4}$$

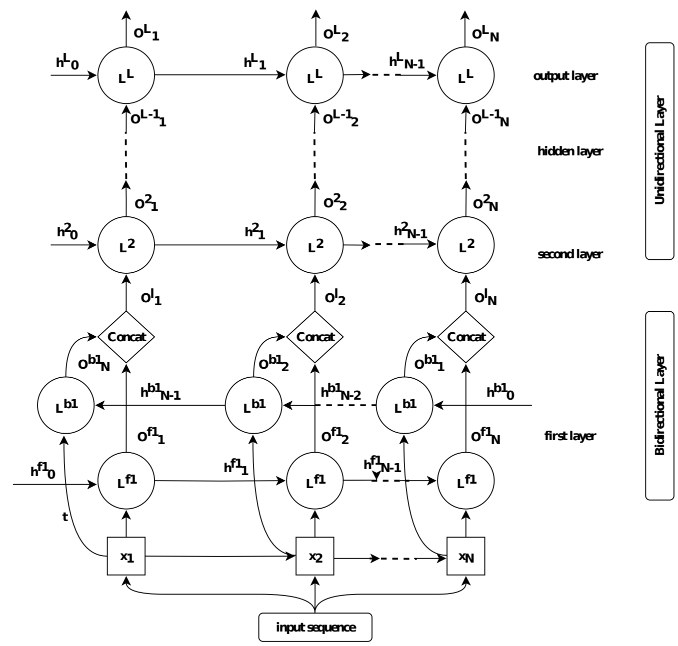

Fig. 1. Cascaded unidirectional and bidirectional LSTM-based DRNN model.

where the output from the lower layer O_n^{l-1} is combined with previous internal state i_{n-1}^{l} and hidden state h_{n-1}^{l} to obtain output O_n^{l} of layer l, and P^{l} represents a parameter of the LSTM cells. The input data contains a sequence of GMF and RSS samples $(x_1, x_2, ..., x_N)$, where each feature x_n is observed at time n $(n = 1, 2, ..., N)$. The data is primarily divided into windows of time segment N and fed into the cascaded LSTM. At the output, we obtain a prediction scores vector for each time step $(O_1^{L}, O_2^{L}, ..., O_N^{L})$. The overall prediction score is obtained by combining the prediction scores vector for the entire window N. The fusion of the scores is performed by applying the sum rule as shown in Eq. (5), which performs better than other methods used in [19]. The prediction scores are finally converted into probabilities by a softmax layer over Y.

$$Y = \frac{1}{N} \sum_{(n=1)}^{N} O_n^{L}. \tag{5}$$

B. Data Collection

The data collections are carried out in a rectangular indoor space using smartphone sensors. The test space contains elements that distort magnetic field like iron doors, metals, steel tables and chairs. It is also equipped with 6 BLE devices as shown in the Fig 2. The GMF and RSS data are collected with android application developed using a smartphone. These features designed for GMF are precisely selected to create a distinctive pattern in each subspace. The structure for GMF data collections are shown in Table I. RSS are collected from 6 BLE devices that are attached to the wall, namely BLE1 to BLE6. These data are combined together to obtain final data set that has 13 features for each subspace. The data set

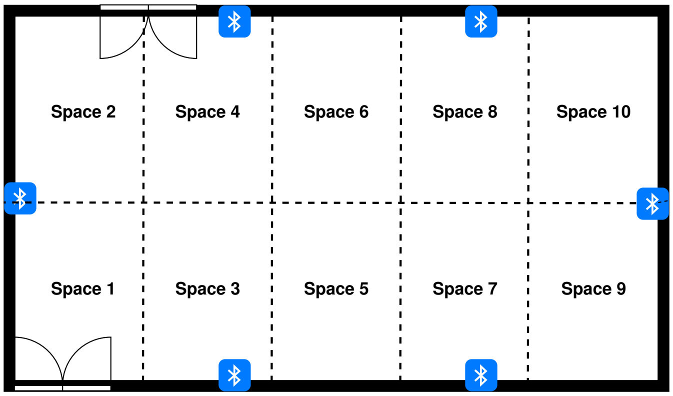

Fig. 2. Floor model where GMF and RSS data are collected.

has around 3000 number of samples for the combined GMF and RSS in each subspace, altogether making it around 30,000 samples.

TABLE I
STRUCTURE OF INPUT FOR GMF

GMF Data						
X-AXIS	Y-AXIS	Z-AXIS	AVG	X-ROT	Y-ROT	Z-ROT

C. Training and Validation Methodology

We divided our data set in the ratio of $8:2$, where 80 % is used for the training and the remaining 20 % is used for the testing. The mean cross-entropy between actual and predicted labels were calculated using cost function $L(O, O^{pred})$ [20], which can be calculated as:

$$L(O, O^{pred}) = -\sum_{(c=1)}^{k} o_c \log_2(p_c), \tag{6}$$

where O and O^{pred} are the actual and predicted class respectively. o_c and p_c are the prediction label and probability of each class respectively.

The cost function, $L(O, O^{pred})$, is minimized using Adam optimization, which uses back-propagation of gradient to update the model parameters. To avoid over-fitting, we have employed dropout as a regularization technique. The output of the final LSTM layer is passed to the soft-max classifier, which converts the output predictions into respective probabilities. Since the input data is segmented with constant N, the model is able to generalize and learn the patterns in the data quickly. Also, the use of mini-batch processing deals with efficient memory utilization and gradient explosion problem. Nevertheless, the training time is slightly increased due to the large number of batches utilized in this work.

III. RESULT AND DISCUSSION

In this section, we detail the performance of proposed model of cascaded LSTM. There are 10 indoor subspace to be classified and the features used are combination of GMF and RSS as mentioned earlier. Fig. 3 shows the training and testing accuracy for our proposed model. It can be seen that the accuracy increases as the model generalizes the data. Similarly, the cost function decreases as the model reaches optimal learning checkpoint, which also means the model deals effectively with data over-fitting. Fig. 4 shows the training and testing cost of out model. If we use the GMF only as the input features for classifying those 10 indoor subspace, we get relatively less accuracy than the proposed method. The accuracy and cost of the GMF based classification using cascaded LSTM are shown in Fig. 5 and 6. However, we can see that the performance of RSS based classification performs significantly lower than GMF data using cascaded LSTM. The RSS data collected at each subspace heavily affected by multi-path and fading effect. Hence, the accuracy and cost looks slightly lower on this case as shown in Fig. 7 and 8. The training accuracy for RSS only and GMF only data are 96.10% and 99.2% respectively. Here, it is clear that the cascaded

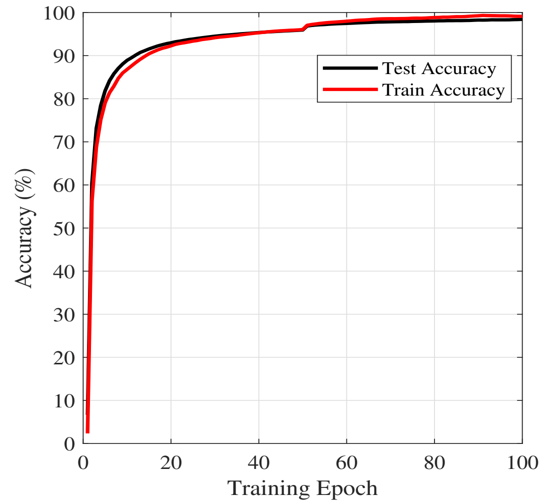

Fig. 3. The accuracy of our proposed data (GMF and RSS) for cascaded LSTM over mini-batch training iterations.

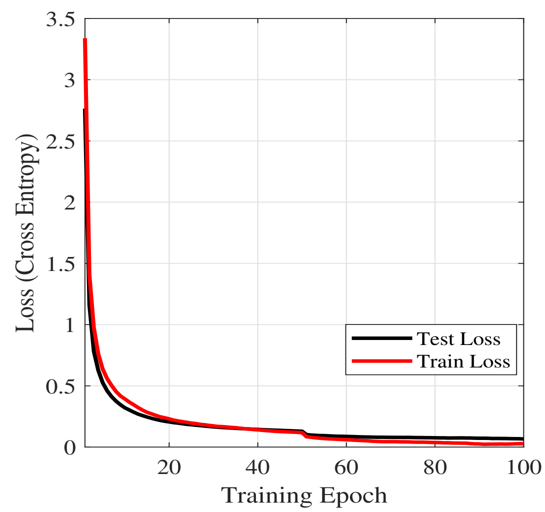

Fig. 4. The cost of our proposed model for the designed data set over mini-batch training iterations.

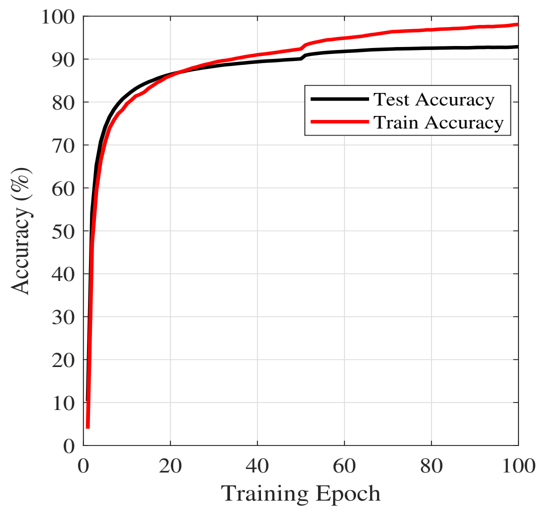

Fig. 5. The accuracy of GMF data for cascaded LSTM over mini-batch training iterations.

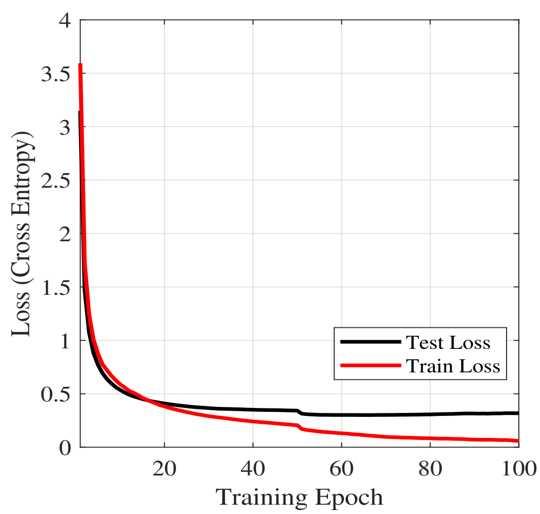

Fig. 6. The cost of GMF data for cascaded LSTM over mini-batch training iterations.

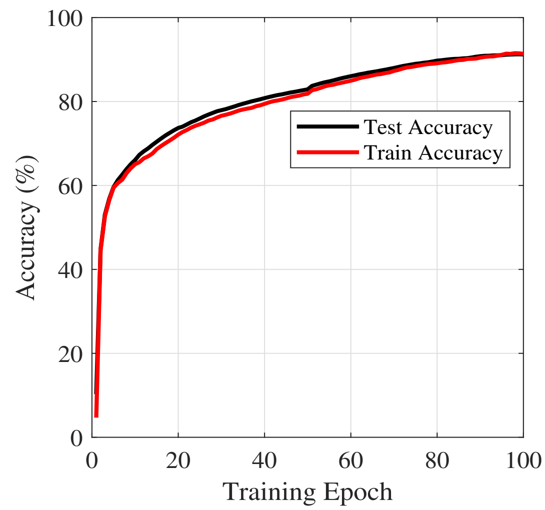

Fig. 7. The accuracy of RSS data for cascaded LSTM over mini-batch training iterations.

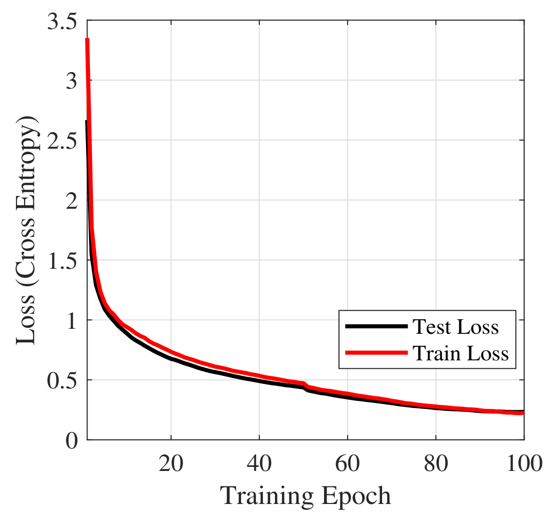

Fig. 8. The cost of RSS data for cascaded LSTM over mini-batch training iterations.

model is able to learn the GMF quite efficiently compared to RSS only data within 100 epochs. Also the losses for GMF data are less compared to RSS data. However, in the case of GMF and RSS data together, the training and testing accuracy reaches up to 100% and 98.95% respectively. Altogether, we can see that the proposed combined data outperforms both individually used data. The accuracy and its loss for training and testing phases are shown in table II and III respectively. In addition, the confusion matrix in Fig. 9 gives an overview of the classification result for the proposed method in test set. It shows the per class precision and recall results as well.

We compared our model with other machine learning methods such as logistic regression, support vector machine (SVM), decision tree (DT), and Gaussian Naïve Bayes (GNB). Our model surpasses other machine learning methods in terms of accuracy on proposed combined GMF and RSS data, which can be shown in Fig. 10. Although some machine learning methods perform close to our model, the data modelling capacity of shallow structure methods fails to capture reliable features as size of the data set increases. Therefore, we believe that if a large number of samples are present, our model can still perform better than other methods.

IV. CONCLUSION

In this paper, we propose cascaded unidirectional and bidirectional model for indoor space classification using GMF and RSS data. We experimentally collect data at various indoor subspaces and then propose an LSTM based structure for learning. We evaluate the performance of our model and compare it with other approaches. The result shows that our model, when both GMF data and RSS data are employed, outperforms the other evaluated methods. The improved performance is

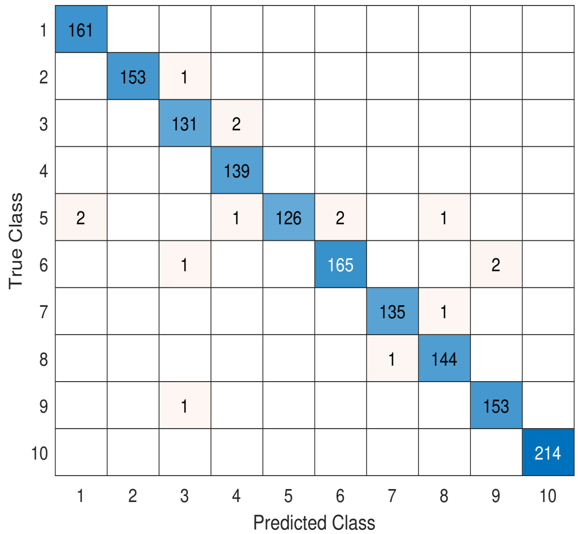

Fig. 9. Confusion matrix of the test data for proposed method.

TABLE II
TRAINING AND TESTING ACCURACY

Methods	Train Accuracy	Test Accuracy
RSS	96.10%	92.01%
GMF	99.2%	93.49%
GMF + RSS	100%	98.95%

mainly due to the capability of our model for extracting distinctive features.

In the future, the study can be carried out when magnetic sensor is better calibrated for building a more accurate data set. Besides, the number of subspaces can be increased in bigger indoor place for commercial and industrial application.

TABLE III
TRAINING AND TESTING LOSS

Methods	Train Loss	Test Loss
RSS	0.22	0.21
GMF	0.02	0.34
GMF + RSS	0.012	0.45

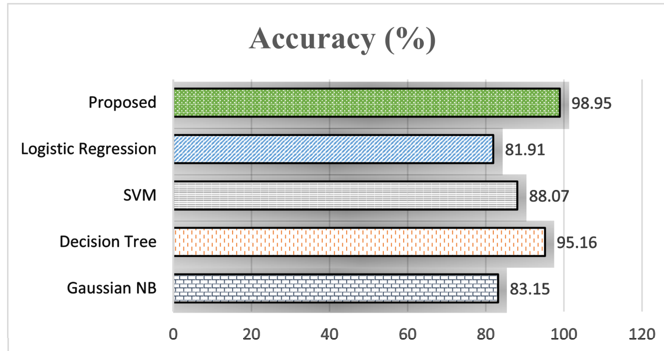

Fig. 10. Comparison with several Machine learning methods.

REFERENCES

[1] Gabriel Deak, Kevin Curran, and Joan Condell. A survey of active and passive indoor localisation systems. *Computer Communications*, 35:1939–1954, 2012.

[2] Santosh Subedi and Jae-Young Pyun. Practical fingerprinting localization for indoor positioning system by using beacons. *J. Sensors*, 2017:9742170:1–9742170:16, 2017.

[3] Xuyu Wang, Lingjun Gao, and Shiwen Mao. Biloc: Bi-modal deep learning for indoor localization with commodity 5ghz wifi. *IEEE Access*, 5:4209–4220, 2017.

[4] Xuyu Wang, Lingjun Gao, Shiwen Mao, and Santosh Pandey. Csi-based fingerprinting for indoor localization: A deep learning approach. *IEEE Transactions on Vehicular Technology*, 66:763–776, 2016.

[5] Yadav Rohan Kumar and Jae-Young Pyun. Ble beacon based indoor positioning using variance-weighted knn. In *Proceedings of Symposium of the Korean Institute of communications and Information Sciences: 2018 Korea Institute Of Communication Sciences*, pages 128–129, Seoul, South Korea, 2018.

[6] Xingbin Ge and Zhiyi Qu. Optimization wifi indoor positioning knn algorithm location-based fingerprint. *2016 7th IEEE International Conference on Software Engineering and Service Science (ICSESS)*, pages 135–137, 2016.

[7] Moustafa Youssef and Ashok K. Agrawala. The horus wlan location determination system. In *MobiSys '05*, 2005.

[8] Rohan Kumar Yadav, Bimal Bhattarai, Hui-Seon Gang, and Jae-Young Pyun. Trusted k nearest bayesian estimation for indoor positioning system. *IEEE Access*, 7:51484–51498, 2019.

[9] Xingli Gan, Baoguo Yu, Lu Huang, and Yaning Li. Deep learning for weights training and indoor positioning using multi-sensor fingerprint. *2017 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, pages 1–7, 2017.

[10] Minh Tu Hoang, Brosnan Yuen, Xiaodai Dong, Tao Lu, Robert Westendorp, and Kishore Reddy. Recurrent neural networks for accurate rssi indoor localization. *IEEE Internet of Things Journal*, 6:10639–10651, 2019.

[11] K. Yamazaki, Koichi Kato, Kazuhiro Ono, Haruhisa Saegusa, Yasuhiro Iida, S. Yamamoto, Kazutoshi Tokunaga, K. Ashiho, Kent Fujiwara, and Norio Takahashi. The analysis of magnetic disturbance due to buildings. *Digest of INTERMAG 2003. International Magnetics Conference (Cat. No.03CH37401)*, pages CR–06, 2003.

[12] Ilari Vallivaara, Janne Haverinen, Anssi Kemppainen, and Juha Röning. Magnetic field-based slam method for solving the localization problem in mobile robot floor-cleaning task. *2011 15th International Conference on Advanced Robotics (ICAR)*, pages 198–203, 2011.

[13] Ho Jun Jang, Jae Min Shin, and Lynn Choi. Geomagnetic field based indoor localization using recurrent neural networks. *GLOBECOM 2017 - 2017 IEEE Global Communications Conference*, pages 1–6, 2017.

[14] Xuyu Wang, Zhitao Yu, and Shiwen Mao. Deepml: Deep lstm for indoor localization with smartphone magnetic and light sensors. *2018 IEEE International Conference on Communications (ICC)*, pages 1–6, 2018.

[15] Zhenguang Liu, Luming Zhang, Qi Liu, Yifang Yin, Li Cheng, and Roger Zimmermann. Fusion of magnetic and visual sensors for indoor localization: Infrastructure-free and more effective. *IEEE Transactions on Multimedia*, 19:874–888, 2017.

[16] Pengfei Liu, Xipeng Qiu, Xinchi Chen, Shiyu Wu, and Xuanjing Huang. Multi-timescale long short-term memory neural network for modelling sentences and documents. In *EMNLP*, 2015.

[17] Bimal Bhattarai, Rohan Kumar Yadav, Hui-Seon Gang, and Jae-Young Pyun. Geomagnetic field based indoor landmark classification using deep learning. *IEEE Access*, 7:33943–33956, 2019.

[18] Alex Graves, Abdel-rahman Mohamed, and Geoffrey E. Hinton. Speech recognition with deep recurrent neural networks. *CoRR*, abs/1303.5778, 2013.

[19] Yonghui Wu et al. Google's neural machine translation system: Bridging the gap between human and machine translation, 2016.

[20] Zengqiang Chen and Cheng long Wang. Modeling rfid signal distribution based on neural network combined with continuous ant colony optimization. *Neurocomputing*, 123:354–361, 2014.