NeuralIO: Indoor Outdoor Detection via Multimodal Sensor Data Fusion on Smartphones

Long Wang(✉), Lennard Sommer, Till Riedel, Michael Beigl, Yexu Zhou, and Yiran Huang

Karlsruhe Institute of Technology, TECO, Karlsruhe, Germany
{wanglong,sommer,riedel,beigl,zhou,yhuang}@teco.edu

Abstract. The Indoor Outdoor (IO) status of mobile devices is fundamental information for various smart city applications. In this paper we present NeuralIO, a neural network based method to deal with the Indoor Outdoor (IO) detection problem for smartphones. Multimodal data from various sensors on a smartphone are fused through neural network models to determine the IO status. A data set consisting of more than 1 million samples is constructed. We test the performance of an early fusion scheme in various settings. NeuralIO achieves above 98% accuracy in 10-fold cross-validation and above 90% accuracy in a real-world test.

Keywords: Indoor outdoor detection · Multimodal data fusion · Neural network model

1 Introduction

The past decade has witnessed the flourishing of the Internet of Things (IoT) and its applications in urban spaces. The widespread deployment of IoT devices and the rise of the smart cities are giving birth to an increasing number of smart applications [3,6,9,12]. Context status is critical and fundamental information for ubiquitous computing systems and context-aware IoT applications [10,25]. "Context" consists of a wide range of aspects such as location, time, surrounding environment and so on. The rapid growth of smartphones is driving the increasing interest in context-aware applications [15,16,19].

One of the most fundamental contextual information is whether the device is in an indoor or outdoor environment. It makes a significant difference if a user is standing in front of a shopping mall or in a shopping mall. Further, the availability and capabilities of different technologies vary considerably between these two environments. The knowledge about the Indoor Outdoor (IO) status enables the choice of appropriate technologies, which leads to a better user experience. For instance, the device can trigger a reminder, change the working mode, and

H. Santos et al. (Eds.): SmartCity 360 2019, LNICST 323, pp. 127–138, 2020.
https://doi.org/10.1007/978-3-030-51005-3_13

switch between GPS based navigation and indoor navigation schemes when the user enters or leaves an indoor environment. Further, the device can save energy by turning off the GPS module in indoor environments such as a metro station. Existing IO detection approaches commonly use GPS signal [7,8,18,26], wireless signal [5,22,27,29] and other sensor data [2,11,17,20,28,28] to determine IO status.

Due to the rich characteristics of natural phenomena, it is rare that a single modality provides comprehensive knowledge of the phenomenon of interest [13]. The increasing availability of multiple sensing modalities on smartphones offers us more freedom to recognize the context. The capability of neural network models bas been proven superior in solving increasingly complex machine learning problems, which often involve multiple data modalities [21].

We propose NeuralIO to detect the Indoor Outdoor status of smartphones through multimodal sensor data fusion using neural network models. We create a data set containing more than 1 million labelled samples by 9 users. 9 different sensing modalities are covered in the data set, which are accelerometer, GPS, light, magnetic, proximity, cellular signal strength, sound level, temperature and WiFi. We test the performance of an early fusion scheme in various settings.

To summarize, the contributions in this paper are as follows:

1. We apply neural network models to the IO detection problem and provide a comprehensive analysis.
2. We implement an Android app for data collection and conduct experiments to collect data samples in various real daily scenarios. A data set consisting of more than 1 million labeled data samples is constructed.
3. We evaluate the performance of an early fusion scheme on the data set through cross-validation and a real-world test. Above 98% accuracy is achieved in the cross-validation and above 90% accuracy is achieved in the real-world test.

The rest of the paper is organized as follows: Sect. 2 presents related work. Different fusion schemes are introduced in Sect. 3. The experiment and data collection is described in Sect. 4 and evaluation results are presented in Sect. 5. We conclude our work in Sect. 6.

2 Related Work

2.1 GPS Based Methods

GPS signal is highly dependent on the line-of-sight (LOS) paths between the device and GPS satellites. It is well known that GPS signals are poor in indoor environments as the LOS paths of GPS signals are blocked. In contrast, the LOS paths are not blocked in most outdoor scenarios. On the basis of these facts, the localization accuracy of GPS or the availability of GPS signal has been exploited to determine whether a device was in an indoor or outdoor environment [7,8,18,26].

Despite the intuitive nature and easy implementation of GPS based methods, they suffer from several disadvantages. Radu et.al. identified the GPS chipset as the sensor with the highest power consumption among the evaluated sensors [20]. The battery capacity is still limited in state-of-the-art mobile phones and most users dislike applications which drain the battery. Secondly, the intuition behind these methods is not always reliable. For instance, the GPS signal is reasonably strong if a device is in an indoor environment with large windows. In contrast, the GPS signal can be blocked by surrounding mountains if the device is in a valley. Under these circumstances, GPS based methods may give misleading results. A third disadvantage is that it normally takes around one minute to launch a GPS module, making GPS-based methods unsuitable for real-time applications.

2.2 Wireless Signals

Shtar et al. [22] presented a method for continuous indoor outdoor environment detection on mobile devices based solely on WiFi fingerprints and assumed no prior knowledge of the environment. The model trained with the data collected for just a few hours on a single device was applicable for unknown locations and new devices. WifiBoost [5] made use of a machine learning meta-algorithm that combined a sufficiently large ensemble of simple classifiers (so-called weak learners) to improve the overall performance. An average error rate of around 2.5% was achieved in the evaluation. However, a classifier needed to be created for each building and the surrounding area through measurements and labeling of each measurement point, especially in cases where there was no previous fingerprinting database. Building such a database is not a trivial task.

Wang et al. [27] applied a machine learning algorithm to classify the signal strength of neighboring cellular base stations in different environments and identified the current context by signal pattern recognition. Accuracy of 100% was reported for the identification of open outdoors, semi-outdoors, light indoors, and deep indoors.

In [29], low-power iBeacon technology was leveraged to develop an accurate, fast response and energy-efficient scheme for indoor outdoor detection. The transitions between outdoors and indoors were detected by comparing the Received Signal Strength of two pre-deployed Bluetooth beacons at two sides of each entrance.

2.3 Multiple Sensors

Since a single sensor might not be able to tackle with all application scenarios, data from multiple sensors such as accelerometer, proximity and light sensor, wireless receiver and magnetometer were exploited for IO detection [2,11,17,20,28]. IODetector [28] combined data from three lightweight sensors (light sensor, cell tower signal strength and magnetic sensor) to develop an extensible IO detection framework which did not require a training phase. Although acceptable error rates were achieved, Radu et al. [20] criticized IODetector for its hard-coded thresholds which might not work with new devices and new environments. As an alternative, Radu et al. proposed a semi-supervised training method to improve IO detection accuracy across different devices and environments.

2.4 Other Methods

In [14], the embedded digital camera on a mobile phone was utilized for IO detection. The developed gentle boosting classifier achieved an error rates of 1.7% for indoor and 10.8% for outdoor scenes. Beside, a feed forward neural network was trained with GIST feature of images to address the IO detection problem [24]. These methods help in generating semantic IO labels for images, but do not work for tracking and other real time application cases.

Sung et al. [23] developed a sound based IO detection method using chirp signal. A simple classifier was developed with a static threshold. However, this work was rather simple and straightforward, and no comprehensive analysis was provided. Wang et al. conducted a comprehensive study on an audio based IO detection method. The method was evaluated in various scenarios with different probing signals (MLS and chirp), noise levels and device types.

3 Fusion Scheme

Neural networks offer the flexibility of implementing multimodal sensor fusion as either early, late or intermediate fusion [21].

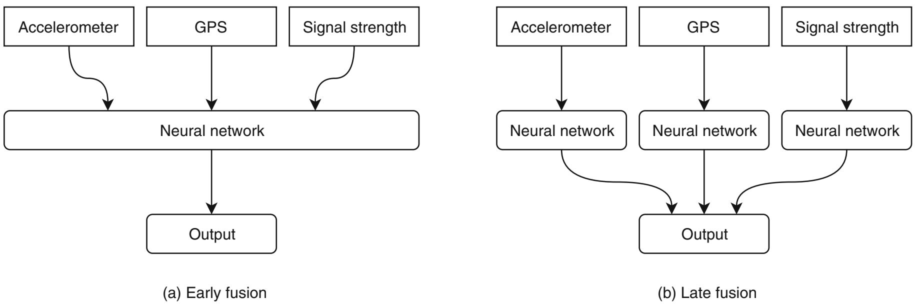

Fig. 1. Schema of early fusion and late fusion schemes based on neural networks

As shown in Fig. 1a), in early fusion scheme data from multiple sources are intergrated into a single feature vector to serve as the input of one machine learning model. In contrast, late fusion scheme aggregates decisions from multiple models which are trained separately on their own modality as shown in Fig. 1b). This fusion scheme is often favored because errors from multiple classifiers tend to be uncorrelated and the method is feature independent [21].

For traditional machine learning methods, it is typically necessary to manually extract features from each modality which is not only time-consuming but also challenging. Neural networks are known for being able to learn features automatically. In this paper we choose to use an Feedforward Neural Network (FNN) model to conduct early fusion for indoor outdoor detection problem.

4 Experiment and Data Collection

4.1 App Design and Implementation

We have developed an Android app for data collection. The app needs to access multiple sensors on the smartphone and save the sensor readings to a database. The collected data contains: battery temperature, luminance, magnetic flux density, proximity, cellular signal strength, cellular network bit error rate, an abstract level for the overall signal strength ranging from one to four number of WiFi networks around the user, the highest signal strength of the WiFi networks around the user, number of GPS satellites, GPS accuracy in meters, GPS signal-to-noise ratio, ambient noise level. Additionally, some anonymous information about the device is also recorded to distinguish different data traces.

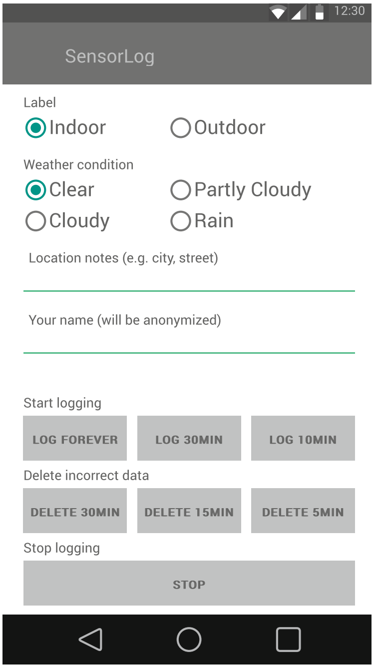

Fig. 2. Screenshot of the developed Android app

Figure 2 shows a screenshot of the developed app. The users specifies whether they are indoor or outdoor and inputs the current weather condition. Then, they have the option to provide notes on the location and their name. The users start the logging period for either 10 min, 30 min or an unlimited amount of time. If, for example, the users walk indoors while logging data labeled outdoor they have the option to invalidate the last 5, 15 or 30 min of the collected data. The users can stop the logging process at any time.

The application collects the specified information every 200 ms as one json object. The data is then sent to an instance of the Firebase Realtime Database (DB) [1]. This ensures that every user directly writes to the same database and no data is saved locally on the user's device. From there, the data can be downloaded for further processing.

4.2 Data Collection

The smartphone application is handed out to multiple participants for data collection. The users are instructed about the application and how to use it. The data collection runs for four weeks, users are free to choose the time and environment for data logging. Figure 3 shows the typical data logging scenario.

Fig. 3. Data logging process. The picture on the left shows how a user configuring the data logging session and the picture on the right depicts data logging inside a pocket.

The resulting dataset consists of 1,038,678 samples which is around 58 h of data. 99.49% of the data is collected by four users. The remaining 0.51% of the data was collected by 5 other users. Overall, the distribution of indoor to outdoor samples is 57.61% to 42.39%. The distribution before cleaning for different smartphones is illustrated in Fig. 4. Different smartphones also represent different users.

4.3 Preprocessing

By removing the samples that were invalidated by the users themselves, 1,019,091 samples are left which is equivalent to about 56.5 h of data. However, not every collected sample is completed due to various reasons. After removing the incomplete samples, the resulting dataset includes 623,320 samples, which is equivalent to around 34.5 h of data. The balance between indoor and outdoor samples is now 43.98% to 56.02%.

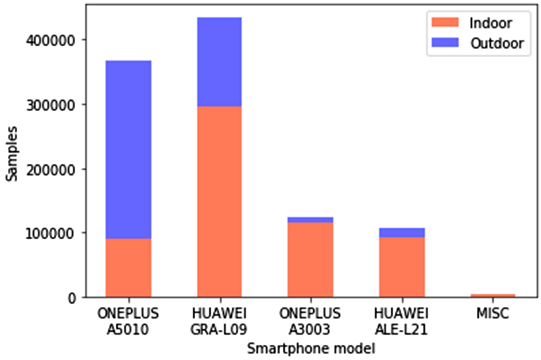

Fig. 4. Distribution before cleaning

5 Evaluation

5.1 Cross Validation

We use 10-fold cross-validation to evaluate the performance of the constructed model. We have tried various numbers of hidden units and hidden layers. Finally we got a good balance between performance and model complexity by using the architecture in Fig. 5. The input layer with 24 input nodes is omitted due to limited space. There are four hidden layers with 10, 5, 4, 3 hidden units with Relu as activation function. The output unit uses the sigmoid function as the activation function. As shown in Table 1, the results from 10-fold cross-validation demonstrate that the model performs very well in 9 out of 10 folds, in the 5th fold the model only achieves an accuracy of 0.73. This is probably due to the loss function becoming trapped at a local minimum.

Table 1. Results of 10-fold cross-validation. Precision and recall are for the outdoor label.

	1	2	3	4	5	6	7	8	9	10
Accuracy	0.98	0.99	0.99	0.99	0.73	0.98	0.99	0.99	0.99	0.98
Precision	0.98	0.99	0.99	0.99	0.73	0.99	0.99	0.99	0.99	0.99
Recall	1	0.99	0.99	0.99	1	0.99	0.99	0.99	0.99	0.98

5.2 Real-World Test

To verify the performance of the model in the real world, we tested the trained FNN model on a real-world dataset. The real-world dataset was recorded around two months later than the training dataset. During the collection of the data

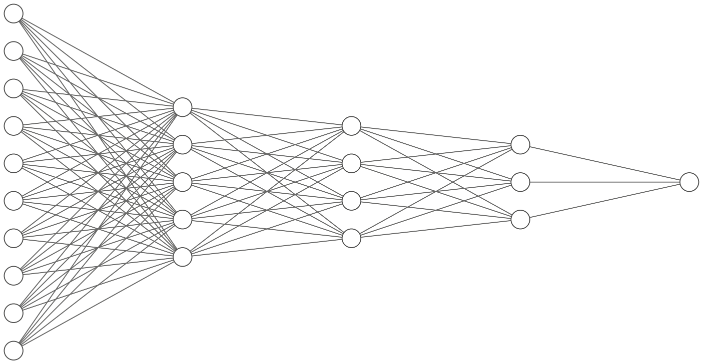

Fig. 5. FNN model architecture with four hidden layers and one output layer. Note that the input layer (with 24 input nodes) is not displayed due to limited space. The activation function of all hidden units is the Relu function. The activation function for the output node is the sigmoid function.

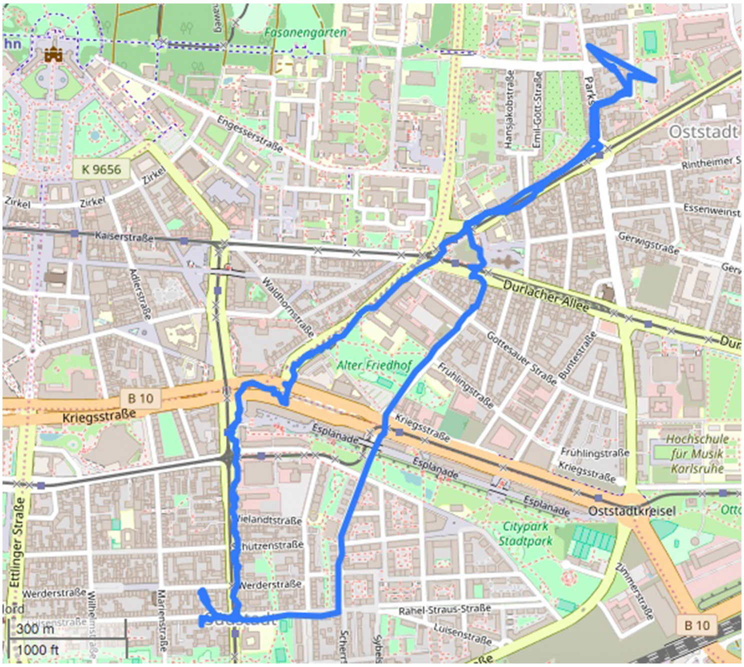

Fig. 6. Real-world test path

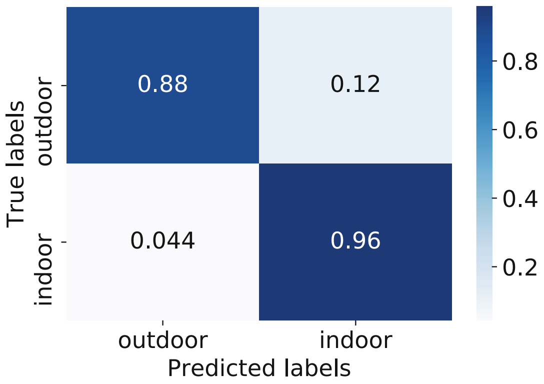

Fig. 7. Confusion matrix

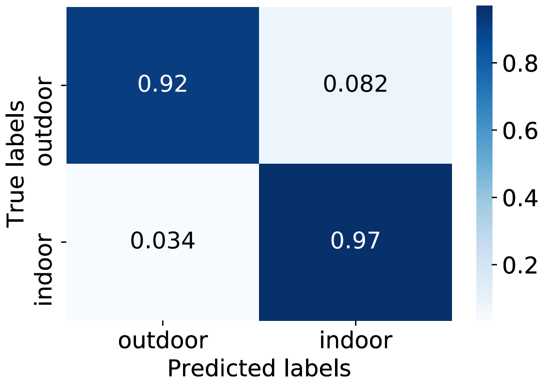

Fig. 8. Confusion matrix for model with majority voting

set, the user walked through the city as depicted in Fig. 6. The trace covers indoor environments such as campus buildings and shopping malls, and outdoor environments such as streets.

The confusion matrix is shown in Fig. 7. Generally the model perform quite well in real-world test with an overall accuracy of 91%. Specifically, the model can recognize indoor cases with a precision of 96% with only 4% falsely classified as outdoor. The model achieves a precision of 88% for outdoor cases with 12% of all outdoor cases falsely classified as indoors. The model shows good generalizability on new data set.

To investigate the cause of misclassification problem of the model, we plot the labels of all data entries against the index in Fig. 9. As shown in Fig. 9, there are some isolated misclassifications for both indoor and outdoor cases. Considering the common sense that it is very rare for people to switch between indoor and outdoor states in a short time period (for instance 2 s), we can use a majority voting strategy with a sliding window to filter out the isolated misclassification

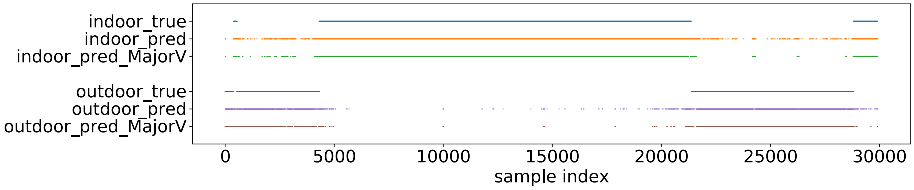

Fig. 9. Ground truth, prediction results without/with majority voting strategy.

cases. The basic idea is that the indoor/outdoor state is not only determined by the input data, but also depends on the previous predicted labels in the sliding window. As shown in Fig. 9, there are fewer isolated misclassification cases after applying the majority voting strategy with a sliding window of 10. The confusion matrix in Fig. 8 also shows an increase in the precision for both indoor and outdoor cases.

6 Conclusion and Discussion

We developed NeuralIO, a neural network based multimodal fusion method for indoor outdoor detection problem on smartphones. A data set consisting of more than 1 million data samples was constructed. 9 different sensing modalities were covered in the data set. We built a feed forward neural network model for early fusion of all available raw data. Cross-validation and real-world test have shown its feasibility for indoor outdoor detection and generalizability on a new data set.

Due the length limit, we did not investigate the late, intermediate fusion scheme and other neural network models. We reserve them as future work.

Acknowledgements. This work was partially funded as part of BMVI project "SmartAQnet" [4], grant no. 19F2003B.

References

1. Firebase realtime database documentation. https://firebase.google.com/docs/database/. Accessed 07 June 2019
2. Ali, M., ElBatt, T., Youssef, M.: Senseio: realistic ubiquitous indoor outdoor detection system using smartphones. IEEE Sens. J. **18**(9), 3684–3693 (2018)
3. Amft, O., Van Laerhoven, K.: What will we wear after smartphones? IEEE Pervasive Comput. **4**, 80–85 (2017)
4. Budde, M., et al.: SmartAQnet: remote and in-situ sensing of urban air quality. In: Proceedings of the SPIE 10424, Remote Sensing of Clouds and the Atmosphere XXII, 104240C (2017)
5. Canovas, O., Lopez-de Teruel, P.E., Ruiz, A.: Detecting indoor/outdoor places using wifi signals and adaboost. IEEE Sens. J. **17**(5), 1443–1453 (2017)

6. Chen, Y., Yonezawa, T., Nakazawa, J., Tokuda, H.: Evaluating the spatio-temporal coverage of automotive sensing for smart cities. In: 2017 Tenth International Conference on Mobile Computing and Ubiquitous Network (ICMU), pp. 1–5. IEEE (2017)
7. Chintalapudi, K., Padmanabha Iyer, A., Padmanabhan, V.N.: Indoor localization without the pain. In: Proceedings of the Sixteenth Annual International Conference on Mobile Computing and Networking, pp. 173–184. ACM (2010)
8. Cho, H., Song, J., Park, H., Hwang, C.: Deterministic indoor detection from dispersions of GPS satellites on the celestial sphere. In: The 11th International Symposium on Location Based Services (2014)
9. Franke, T., Lukowicz, P., Blanke, U.: Smart crowds in smart cities: real life, city scale deployments of a smartphone based participatory crowd management platform. J. Internet Serv. Appl. **6**(1), 1–19 (2015). https://doi.org/10.1186/s13174-015-0040-6
10. Ishida, Y., Thepvilojanapong, N., Tobe, Y.: Winfo+: identification of environment condition using walking signals. In: Tenth International Conference on Mobile Data Management: Systems, Services and Middleware, MDM 2009, pp. 508–512. IEEE (2009)
11. Jia, M., Yang, Y., Kuang, L., Xu, W., Chu, T., Song, H.: An indoor and outdoor seamless positioning system based on android platform. In: Trustcom/BigDataSE/I SPA, 2016 IEEE, pp. 1114–1120. IEEE (2016)
12. Khaled, A.E., Helal, A., Lindquist, W., Lee, C.: IoT-DDL-device description language for the "T" in IoT. IEEE Access **6**, 24048–24063 (2018)
13. Lahat, D., Adali, T., Jutten, C.: Multimodal data fusion: an overview of methods, challenges, and prospects. Proc. IEEE **103**(9), 1449–1477 (2015)
14. Lipowezky, U., Vol, I.: Indoor-outdoor detector for mobile phone cameras using gentle boosting. In: 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 31–38. IEEE (2010)
15. Maeda, H., Sekimoto, Y., Seto, T.: An easy infrastructure management method using on-board smartphone images and citizen reports by deep neural network. In: Proceedings of the Second International Conference on IoT in Urban Space, pp. 111–113. ACM (2016)
16. Maeda, H., Sekimoto, Y., Seto, T.: Lightweight road manager: smartphone-based automatic determination of road damage status by deep neural network. In: Proceedings of the 5th ACM SIGSPATIAL International Workshop on Mobile Geographic Information Systems, pp. 37–45. ACM (2016)
17. Nakamura, Y., Ono, M., Sekiya, M., Honda, K., Takahashi, O.: Indoor/outdoor determination method using various sensors for the power saving of terminals in geo-fencing. In: Proceedings of the 2015 International Workshop on Informatics (2015)
18. Okamoto, M., Chen, C.: Improving GPS-based indoor-outdoor detection with moving direction information from smartphone. In: Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers, pp. 257–260. ACM (2015)
19. Perttunen, M., et al.: Distributed road surface condition monitoring using mobile phones. In: Hsu, C.-H., Yang, L.T., Ma, J., Zhu, C. (eds.) UIC 2011. LNCS, vol. 6905, pp. 64–78. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-23641-9_8

20. Radu, V., Katsikouli, P., Sarkar, R., Marina, M.K.: A semi-supervised learning approach for robust indoor-outdoor detection with smartphones. In: Proceedings of the 12th ACM Conference on Embedded Network Sensor Systems, pp. 280–294. ACM (2014)
21. Ramachandram, D., Taylor, G.W.: Deep multimodal learning: a survey on recent advances and trends. IEEE Signal Process. Mag. **34**(6), 96–108 (2017)
22. Shtar, G., Shapira, B., Rokach, L.: Clustering Wi-Fi fingerprints for indoor-outdoor detection. Wirel. Netw. **25**, 1–19 (2018)
23. Sung, R., Jung, S., Han, D.: Sound based indoor and outdoor environment detection for seamless positioning handover. ICT Express **1**(3), 106–109 (2015)
24. Tahir, W., Majeed, A., Rehman, T.: Indoor/outdoor image classification using gist image features and neural network classifiers. In: 2015 12th International Conference on High-Capacity Optical Networks and Enabling/Emerging Technologies (HONET), pp. 1–5. IEEE (2015)
25. Uehara, Y., Mori, M., Ishii, N., Tobe, Y., Shiraishi, Y.: Step-wise context extraction in AoK mule system. In: Proceedings of the 4th International Conference on Embedded Networked Sensor Systems, pp. 379–380. ACM (2006)
26. Wang, H., Sen, S., Elgohary, A., Farid, M., Youssef, M., Choudhury, R.R.: No need to war-drive: Unsupervised indoor localization. In: Proceedings of the 10th International Conference on Mobile Systems, Applications, and Services, pp. 197–210. ACM (2012)
27. Wang, W., Chang, Q., Li, Q., Shi, Z., Chen, W.: Indoor-outdoor detection using a smart phone sensor. Sensors **16**(10), 1563 (2016)
28. Zhou, P., Zheng, Y., Li, Z., Li, M., Shen, G.: IODetector: a generic service for indoor outdoor detection. In: Proceedings of the 10th ACM Conference on Embedded Network Sensor Systems, pp. 113–126. ACM (2012)
29. Zou, H., Jiang, H., Luo, Y., Zhu, J., Lu, X., Xie, L.: Bluedetect: an ibeacon-enabled scheme for accurate and energy-efficient indoor-outdoor detection and seamless location-based service. Sensors **16**(2), 268 (2016)