2019 IEEE Intelligent Transportation Systems Conference (ITSC)
Auckland, NZ, October 27-30, 2019

WiFi-based urban localisation using CNNs

N. Hernández, H. Corrales, I. Parra, M. Rentero, D.F. Llorca, M.A. Sotelo, Fellow, IEEE[1]

Abstract— **The continuous expanding scale of WiFi deployments in metropolitan areas has made possible to find WiFi access points at almost any place in our cities. Although WiFi has been mainly used for indoor localisation, there is a growing number of research in outdoor WiFi-based localisation. This paper presents a WiFi-based localisation system that takes advantage of the huge deployment of WiFi networks in urban areas. The idea is to complement localisation in zones where the GPS coverage is low, such as urban canyons. The proposed method explores the CNNs ability to handle large amounts of data and their high accuracy with reasonable computational costs. The final objective is to develop a system able to handle the large number of access points present in urban areas while preserving high accuracy and real time requirements. The system was tested in a urban environment, improving the accuracy with respect to the state-of-the-art and being able to work in real time.**

I. Introduction

Self driving cars research has ushered in a great development in a number of areas, being localisation one of the most outstanding ones. Major car makers are investing significant efforts on accurate maps building for automated driving purposes. Those maps usually contain enriched information regarding the geometrical configuration of the environment and any other salient feature that can help on the localisation process using vision [1] or LiDAR [2] [3].

Regarding automated driving, there is a debate in the scientific community about the trade-off between local perception capability and map dependence. However, map-based localisation is still a crucial and necessary element in today's automated driving systems.

Although WiFi has been mainly used for indoor localisation, in last few years, the number and quality of the research in outdoor WiFi-based localisation has been continuously growing. The continuous expanding scale of WiFi deployments in metropolitan areas has made possible to perform accurate GPS-free localisation solely based on the existing WiFi infrastructure [4]. In addition, some new crowdsensing projects allow for a collaborative mapping of metropolitan areas reducing the traditionally big survey effort [5].

In this paper, we present a preliminary work for WiFi-based urban localisation using CNNs (Convolutional Neural Networks) (Fig. 1). This paper is a natural evolution of our previous work [6] in which we developed a continuous space estimator (CSE) with mean localisation errors under 4 m using the Received Signal Strength (RSS) from WiFi APs (Access Points). This continuous space estimator used an SVR (Support Vector Regression) [7] algorithm to estimate virtual fingerprints not available in the fingerprint database, thus reducing the site survey effort. However, the computational cost of searching for the best position along the estimated surfaces was very high ($\sim 200s$ per sample) making it impossible for a real time application. In this paper, we want to explore the CNNs ability to deal with huge amounts of data and extend our previous system to WiFi-based urban localisation able to complement autonomous driving map-based localisation.

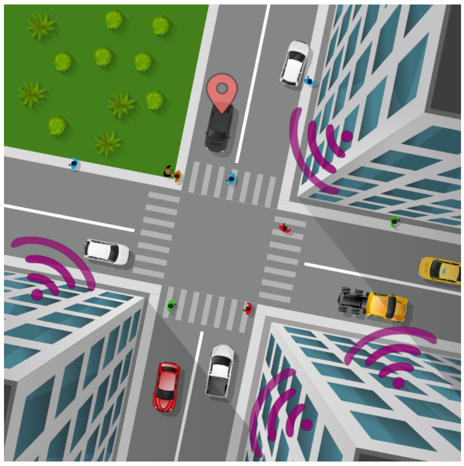


Fig. 1. WiFi-based urban localisation.

In the future, we plan to include this WiFi information in our enriched 3D map [8] to help with the localisation in urban canyons where the GPS coverage is low (Fig. 2).

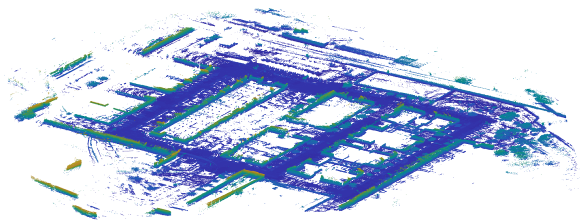


Fig. 2. 3D enriched map [8].

The remainder of the paper is organised as follows: Section II introduces the state-of-the-art, in Section III the proposed localisation algorithm is described. Section IV analyses the experimental results. Finally, Section V presents the conclusions and future work.

[1] All authors are with the Computer Engineering Department, Universidad de Alcalá, Spain noelia.hernandez, hector.corrales, ignacio.parra{@uah.es}

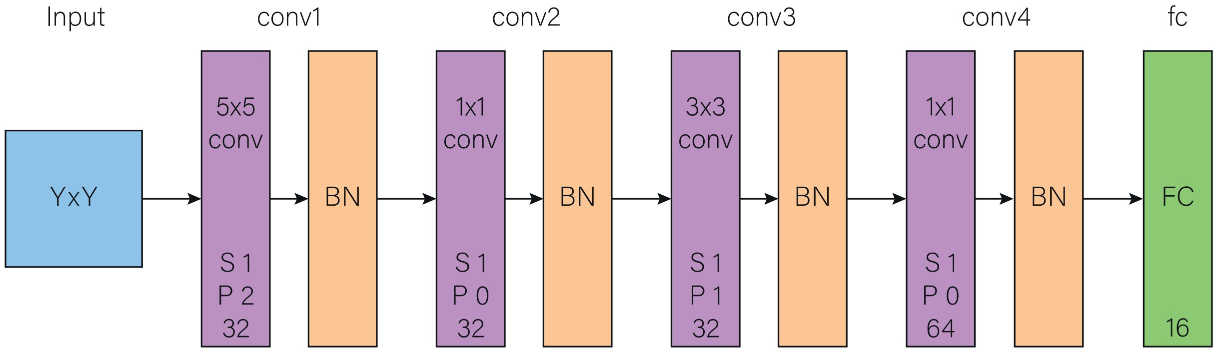


Fig. 3. Proposed CNN architecture based on ResNet bottleneck units.

II. Related Work

WiFi is one of the most widely used technologies for indoor localisation where it has been proved accurate and reliable. However, some recent work [4], [9] propose the use of WiFi for outdoor localisation. This new line of research is motivated by the low performance of the GPS in urban canyons and the high cost of differential GPSs.

Both indoor and outdoor WiFi localisation is usually approached in two different ways: using range-based (propagation model-based) methods or using fingerprinting-based methods.

On the one hand, range-based methods [10], [11] take advantage of the way the signal is propagated over the environment, being able to transform the RSS from an AP into a distance to that AP. These methods have the advantage of being quick to implement, requiring low maintenance and being valid under any circumstances. Their main disadvantage is the need of knowing the exact location of all the APs. This location is used to apply lateration algorithms to estimate the real position of the device. Another disadvantage is the need of Line-Of-Sight (LOS) conditions to obtain high localisation accuracy.

On the other hand, fingerprint-based localisation methods [12], [13], [14], [15] follow a two-stages approach: a training stage and a localisation stage. During the training stage, the environment is divided in cells and the RSS from all the visible APs is collected from each one of them. Then, during the localisation stage, new measurements are compared with the stored ones (usually through classification algorithms) to estimate the position of the device. The main advantage of these methods is that once the system is trained, they faithfully represent the RSS on the different cells, obtaining high localisation accuracy. Their main disadvantage is the need of site-surveying the environment and the high maintenance costs.

When analysing the main characteristics of these methods, it can be noted that obtaining the exact location of the APs on an urban environment is not feasible: APs are mainly located inside private residential buildings. For the same reason, LOS conditions can not be achieved. As a consequence, it is very difficult to obtain high outdoor localisation accuracy using range-based methods. That lead us to select fingerprint-based methods, even though they require a high site-survey effort.

Recent projects, such as Radiocells.org [5], have proposed to collect crowd-sourced RSS data making it publicly available. Even though Spain region is not currently exhaustively covered, we believe that in the near future this kind of public databases will grow, making it easier to implement fingerprint-based localisation methods in wider urban areas. In addition, some methods as the author's previous work [6] seek to reduce the site-survey effort by using regression algorithms to generate "virtual" fingerprints. This method has been tested in a medium-size environment obtaining low localisation errors at the cost of a high computational time (localisation in an environment around 2500 m^2 with 100 APs takes 42.4 seconds per request).

When looking for a new localisation system for autonomous vehicles, two of the most important characteristics, besides obtaining a low localisation error, is that it must be available for large areas and work in real time, so the previously proposed system can not be used to that end. That is the reason why in this paper we want to explore a new localisation method based on CNNs known by their high performance, ability to handle large amounts of data and reasonable computational costs.

III. Urban WiFi Localisation

In this paper, we propose a CNN-based localisation system using RSS fingerprints collected from the available WiFi APs. Given that CNNs are designed to work with images, in the training stage the collected images must be transformed as will be exposed in Section III-C. Then, during the localisation stage the trained CNN will be used to estimate the vehicle position using WiFi RSS measurements collected online.

A. CNN description

There are a huge amount of CNN models for the task of classification/prediction. Some of the most significant ones are AlexNet [16], VGG [17], GoogLeNet [18] or ResNet [19]. All this models have in common that they were designed with Imagenet [20] in mind, a large scale dataset with 1000 different classes. Using one of these deep models could lead us to overfitting, not to mention the enormous computational cost and time required.

In this work, we propose a custom CNN based on ResNet bottleneck units, to take advantage of its efficiency and accuracy. Our CNN is formed by 4 convolutional layers, each one followed by a batch normalisation layer to end with a fully connected layer. The full architecture is shown in Fig. 3. Net weights are initialised with a normal distribution $\mathcal{N}(0, 0.01)$. Finally, the system is trained for 20 epochs.

B. Data Collection

The experiments took place at Alcalá de Henares (Madrid, Spain), covering 27000 m^2 on a urban area of the city centre. This is an urban canyon characterised by a low coverage of GPS signal and a high number of WiFi APs, being all of them deployed by the residents. We have no information about their exact locations or configurations.

The WiFi measurements were collected at 16 evenly distributed locations, approximately 50 m apart (Fig. 4). The measurements were collected on three different runs (train, validation and test), at least three days apart from each other using our automated vehicle (DRIVERTIVE) [21] (Fig. 5).

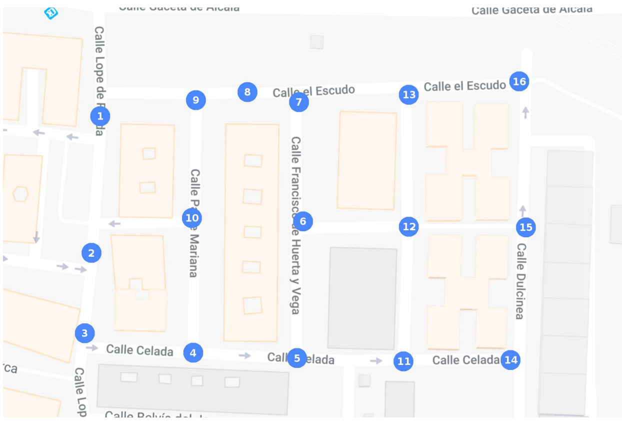


Fig. 4. Experimental environment located in a residential area of Alcalá de Henares (Madrid): Measurement locations are represented with blue circles numbered from 1 to 16.

Fig. 5. DRIVERTIVE vehicle.

Two WiFi interfaces were used to collect the data: an external WiFi antenna mounted on the roof of the car and an embedded WiFi interface of a Toshiba laptop located inside the car. The reason for this was to compare performance when the antenna was located inside or outside the car.

Both antennas are configured to collect data on the 2.4 GHz frequency band, ranging from -99 dBm with a resolution of 1 dBm. The acquisition frequency was 1 Hz for the Toshiba interface and 0.5 Hz for the external device.

At each position, 60 and 30 samples were collected for the Toshiba and the external interface respectively. This way, the time required to collect the data from both interfaces was the same. Each one of the datasets (train, validation and test) are thus composed of 960 samples for the Toshiba interface and 480 for the external device.

C. Data pre-processing

Before we can train our CNN we have to adapt the measured RSS to a format that fit its architecture. Each RSS from an AP has to be pre-processed as shown in Fig. 6.

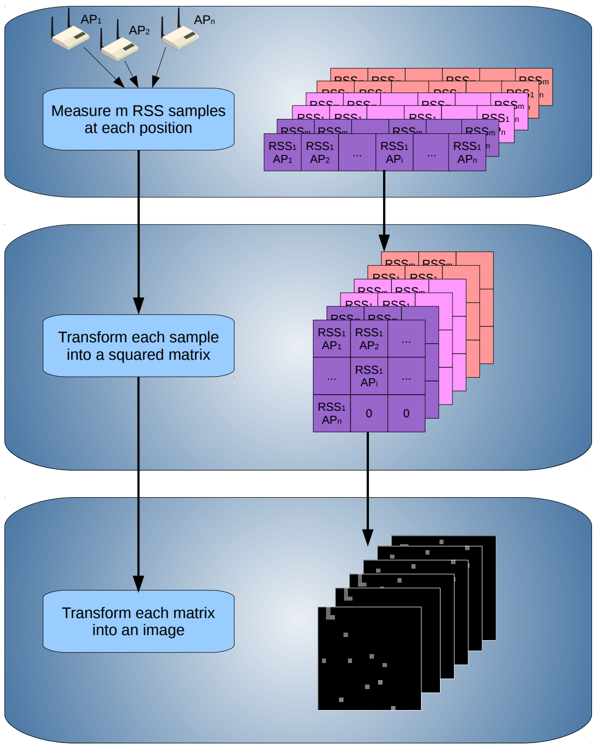


Fig. 6. Collecting and pre-processing data to convert RSS samples into images.

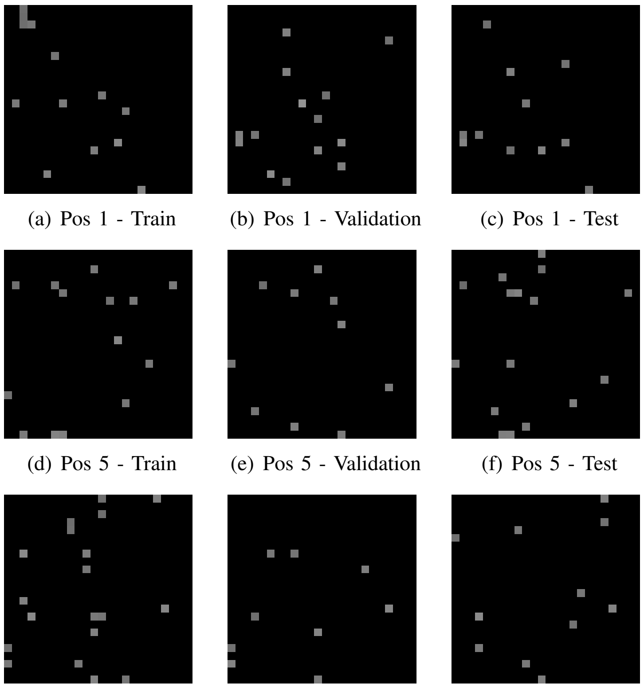
First, each RSS sample is shaped into a square image. To do so, APs order will be arbitrarily selected. Then, the data for each sample will be re-arranged into an Y by Y square image following the selected order, with

$$Y = \left\lceil \sqrt{N_{AP}} \right\rceil \tag{1}$$

being N_{AP} the number of APs seen during the data collection.

In addition, measured RSS must be transformed into pixel intensity values (i.e. from 0 to 255). This transformation is performed by adding an offset to the RSS. Given that the collected RSS ranges from -99 dBm to -50 dBm, we have selected the offset to be +200, resulting in medium-ranged image intensity values (101 to 150). Also, when there is no available RSS from an AP, the corresponding pixel intensity will be set to 0, allowing for high differentiation between non-seen APs and low RSS APs. When the total number of APs is not a perfect square, some empty pixels have to be added in order to create a square image. These pixels are also set to 0.

In the experiments, 536 APs were detected using the Toshiba interface and 596 using the external interface. This way, the images for the Toshiba dataset will be square images of 24×24 pixels and for the external interface of 25×25 pixels. Fig. 7 shows some examples of the images used to train, validate and test the CNN.



(a) Pos 1 - Train (b) Pos 1 - Validation (c) Pos 1 - Test

(d) Pos 5 - Train (e) Pos 5 - Validation (f) Pos 5 - Test

(g) Pos 13 - Train (h) Pos 13 - Validation (i) Pos 13 - Test

Fig. 7. Example of training, validation and test Toshiba interface images.

These images can be interpreted as heat map images, in which each pixel brightness represents the RSS from a different AP collected at a certain position. The lighter the pixel, the higher the RSS from the AP.

IV. System Evaluation

This section describes the experimental results obtained using the CNN as described in the previous section. After the off-line training and validation of the CNN, the system is tested using the third dataset which contains unseen data.

Tables I and II show the accuracy and mean error of the proposed system compared with two state-of-the-art methods.

Firstly, RADAR [12] which uses K-Nearest Neighbours (KNN) [22] and is a commonly used baseline to compare WiFi localisation methods [13], [23].

Secondly, Jedari et. al [24] method, which uses Random Forest (RF), a classification algorithm recognised by its high accuracy when used for localisation systems. Both methods were implemented using the KNN and RF algorithm versions provided by the data mining tool Weka [25], [26].

As can be seen in the results, the accuracy is higher than 93% independently of the method. This is an expected behaviour, as the positions are not very close to each other (minimum of 35 m). Even in this case, it can be seen that the proposed CNN improves the localisation accuracy, reaching a 100% using the Toshiba laptop interface and 98.69% using the external interface which lead us to think that the system has room for improvement in more challenging environments. The mean localisation error is also reduced for both interfaces, specially when compared with the RADAR system.

It is worth noticing that, despite the 100% of accuracy, train and test datasets were collected one week apart in time, thus ruling out the possibility of overlearning.

TABLE I

Summary of the results - Accuracy

Method	Accuracy (%)	
	Toshiba interface	External interface
CNN	**100%**	**98.69%**
RADAR [12]	95.34%	93.23%
RF [24]	96.51%	96.28%

TABLE II

Summary of the results - Mean error

Method	Mean error (m)	
	Toshiba interface	External interface
CNN	**0 m**	**1.02 m**
RADAR [12]	4.29 m	2.92 m
RF [24]	1.75 m	1.74 m

Table III shows the training and test times. Note that the times are different depending on the interface used because both, the number of detected APs and the number of samples is different for each one of them (536 APs and 960 samples using the Toshiba interface and 596 APs and 480 samples using the external interface). This results in a different number of input features for the classifiers (or different size of images for the CNN) and a different number of input samples to train the system depending of the number of collected samples. The training times include training and validation over the complete datasets (960 or 480 samples) while test times are per sample (one location estimation).

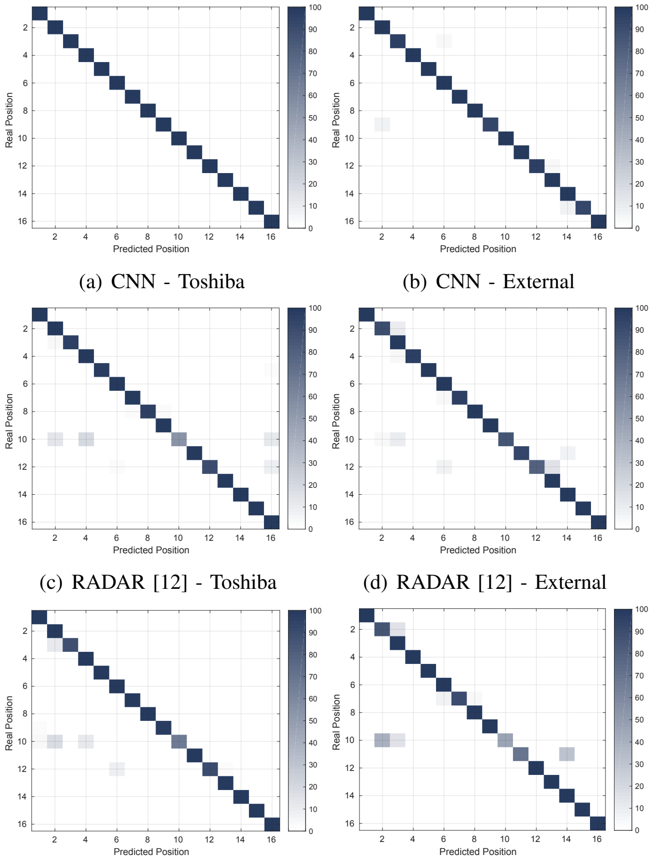
CNN and RF training times do not significantly vary with the interfaces, but testing times doubles for the external interface because the number of APs seen is higher (which means bigger feature vectors for the RF and bigger images for the CNN). RADAR system is highly dependent on the number of samples because it compares the new samples with all the stored ones. So, an increase on the number of inputs or the number of samples increases the computational time. However, for this environment, the real time requirements are met using any of the tested systems since each location estimation time is lower than the measuring frequency (1 second or 2 seconds for 1 Hz or 0.5 Hz interfaces respectively). If we extend these systems to wider environments, and assuming equal APs density, RADAR and RF will be highly penalised in their testing times while the CNN solution will remain unaffected.

The proposed system obtains high accuracy being able to run on real time for a big environment (27000 m^2 with more than 500 APs), which is an improvement when compared with the previously proposed method which needs 220 seconds per location request in this scenario.

TABLE III

SUMMARY OF THE RESULTS - TRAINING AND TEST TIMES

Method	Time (ms)			
	Toshiba interface		External interface	
	Train	Test	Train	Test
CNN	**1293 ms**	7.184 ms	**1260 ms**	14.541 ms
RADAR [12]	21596 ms	13.284 ms	6932 ms	9.802 ms
RF [24]	3562 ms	**1.210 ms**	2532 ms	**2.158 ms**



(a) CNN - Toshiba (b) CNN - External

(c) RADAR [12] - Toshiba (d) RADAR [12] - External

(e) RF [24] - Toshiba (f) RF [24] - External

Fig. 8. Confusion matrix.

Finally, Fig. 8 shows the confusion matrices for both interfaces. It details the predicted positions by the system related to the groundtruth (the positions where the vehicle really was). As can be seen, most of the localisation errors occur within neighbour locations, especially using the CNN.

V. Conclusions and future work

This work has presented a preliminary proposal to estimate the location of an autonomous vehicle using the measurements collected using only the RSS from a WiFi interface by means of a CNN. To do so, the proposed system transforms the RSS samples into square images to take advantage of the good performance of CNNs. The final goal is to apply this system to a continuous space estimator (CSE) and reduce its computational cost.

The proposal was tested in an urban environment under real traffic conditions, improving the accuracy with respect to state-of-the-art methods and being able to work in real time. During the experimentation we have found that urban areas have enough WiFi APs to provide with accurate and reliable outdoor localisation.

In the future, we plan to include WiFi localisation information into enriched 3D maps for automated driving to support the localisation process. In addition, the previously proposed WiFi localisation system [6] (CSE) will be used to generate "virtual" positions that will be used to train the CNN obtaining more dense environments. This way, we expect to increase the accuracy while maintaining the real time requirements.

Finally, we plan to test the system on wider environments using crowdsensing available databases such as Radiocells.org [5].

Acknowledgement

This work was funded by PCIN-2017-086 (Acción Complementaria Internacional, MINECO), from the Electronic Component Systems for European Leadership Joint Undertaking under grant agreement No 737469 (AutoDrive Project). This Joint Undertaking receives support from the European Union Horizon 2020 research and innovation programme and Germany, Austria, Spain, Italy, Latvia, Belgium, Netherlands, Sweden, Finland, Lithuania, Czech Republic, Romania, Norway. This project has also received funding by DPI2017-90035-R (Spanish Ministry of Economy).

References

[1] J. Ziegler, P. Bender, M. Schreiber, H. Lategahn, T. Strauss, C. Stiller, T. Dang, U. Franke, N. Appenrodt, C. G. Keller, E. Kaus, R. G. Herrtwich, C. Rabe, D. Pfeiffer, F. Lindner, F. Stein, F. Erbs, M. Enzweiler, C. Knöppel, J. Hipp, M. Haueis, M. Trepte, C. Brenk, A. Tamke, M. Ghanaat, M. Braun, A. Joos, H. Fritz, H. Mock, M. Hein, and E. Zeeb, "Making Bertha drive - An autonomous journey on a historic

route," *IEEE Intelligent Transportation Systems Magazine*, vol. 6, no. 2, pp. 8–20, Summer 2014.

[2] "Google self-driving car project," accessed on April 2019. [Online]. Available: https://www.google.com/selfdrivingcar/

[3] S. Kim, G. Gwon, W. Hur, D. Hyeon, D. Kim, S. Kim, D. Kye, S. Lee, S. Lee, M. Shin, and S. Seo, "Autonomous campus mobility services using driverless taxi," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 12, pp. 3513–3526, Dec 2017.

[4] J. Wang, N. Tan, J. Luo, and S. J. Pan, "WOLoc: WiFi-only outdoor localization using crowdsensed hotspot labels," in *IEEE INFOCOM 2017 - IEEE Conference on Computer Communications*, 2017, pp. 1–9.

[5] Radiocells.org, "Community project to collect information on cell tower and WiFi base stations," 2009, accessed on April 2019. [Online]. Available: https://radiocells.org/

[6] N. Hernández, M. Ocaña, J. M. Alonso, and E. Kim, "Continuous space estimation: Increasing WiFi-based indoor localization resolution without increasing the site-survey effort," *Sensors*, vol. 17 (1), no. 147, pp. 1–23, 2017.

[7] H. Drucker, C. J. C. Burges, L. Kaufman, A. Smola, and V. Vapnik, "Support vector regression machines," *Advances in Neural Information Processing Systems*, vol. 9, no. 9, pp. 155–161, 1997.

[8] N. Hernández, I. G. Daza, C. Salinas, I. Parra, J. Alonso, D. F. Llorca, and M. A. Sotelo, "Intelligent feature selection method for accurate laser-based mapping and localisation in self-driving cars," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2018, pp. 1–6.

[9] N. Hernández, A. Hussein, D. Cruzado, I. Parra, and J. M. Armingol, "Applying low cost WiFi-based localization to in-campus autonomous vehicles," in *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, 2017, pp. 1–6.

[10] K. Chintalapudi, A. Padmanabha Iyer, and V. N. Padmanabhan, "Indoor localization without the pain," in *Proceedings of the Annual International Conference on Mobile Computing and Networking*, 2010, pp. 173–184.

[11] J. Yang, H. Lee, and K. Moessner, "Multilateration localization based on singular value decomposition for 3D indoor positioning," in *Proceedings of the 2016 International Conference on Indoor Positioning and Indoor Navigation*, 2016, pp. 1–8.

[12] P. Bahl and V. N. Padmanabhan, "RADAR: An in-building RF-based user location and tracking system," in *Proceedings of the Annual Joint Conference of the IEEE Computer and Communications Societies*, 2000, pp. 775–784.

[13] M. Youssef and A. Agrawala, "The Horus location determination system," *Wireless Networks*, vol. 14, no. 3, pp. 357–374, 2008.

[14] W. Li, D. Wei, H. Yuan, and G. Ouyang, "A novel method of WiFi fingerprint positioning using spatial multi-points matching," in *Proceedings of the 2016 International Conference on Indoor Positioning and Indoor Navigation*, 2016, pp. 1–8.

[15] J. Kim and D. Han, "Passive WiFi fingerprinting method," in *2018 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, 2018, pp. 1–8.

[16] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems 25*, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, Eds., 2012, pp. 1097–1105.

[17] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *CoRR*, vol. abs/1409.1556, 2014.

[18] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 1–9.

[19] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778.

[20] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei, "Imagenet large scale visual recognition challenge," *International Journal of Computer Vision*, vol. 115, no. 3, pp. 211–252, Dec 2015.

[21] I. Parra, R. Izquierdo, J. Alonso, A. García-Morcillo, D. Fernández-Llorca, and M. A. Sotelo, "The experience of DRIVERTIVE - DRIVERless cooperaTIve VEhicle-Team in the 2016 GCDC," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 4, pp. 1322–1334, 2018.

[22] D. Kibler and D. Aha, "Learning representative exemplars of concepts: An initial case study," in *Proceedings of the International Workshop on Machine Learning*, 1987, pp. 24–30.

[23] B.-F. Wu, C.-L. Jen, and K.-C. Chang, "Neural fuzzy based indoor localization by Kalman filtering with propagation channel modeling," in *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, 2007, pp. 812–817.

[24] E. Jedari, Z. Wu, R. Rashidzadeh, and M. Saif, "Wi-Fi based indoor location positioning employing random forest classifier," in *Proceedings of the 2015 International Conference on Indoor Positioning and Indoor Navigation*, 2015, pp. 1–5.

[25] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The weka data mining software: An update," *SIGKDD Explorations*, vol. 11, no. 1, pp. 10–18, 2009.

[26] I. H. Witten, E. Frank, and M. A. Hall, *Data Mining: Practical machine learning tools and techniques*, 3rd ed., ser. Data Management Systems Series. Morgan Kaufmann, 2011.

Figure 1 is based on a design by macrovector / Freepik.