Radio map reconstruction with deep neural networks in a weakly labeled learning context with use of heterogeneous side information

Aleksandra Malkova, Massih-Reza Amini, Benoît Denis, Christophe Villien

Radio Map Reconstruction with Deep Neural Networks in a Weakly Labeled Learning Context with use of Heterogeneous Side Information

Aleksandra Malkova[†], Massih-Reza Amini[†],
Benoit Denis[‡], Christophe Villien[‡]

[†] Université Grenoble Alpes, LIG-APTIKAL,
700 avenue centrale, 38400 Saint-Martin d'Hères
[‡] CEA-Leti, 17 Av. des Martyrs, 38054 Grenoble,

E-mail: Aleksandra.Malkova@univ-grenoble-alpes.fr

Abstract

Wireless localization is an essential component in Internet of Things applications. In this paper, we address the generalization problem of RSS map reconstruction for each city (usually, over gateways), relying on location-dependent radio measurements and taking advantage of side knowledge about the local region; like, city plan, terrain height, gateway position. We employ Neural Architecture Search to find an optimized Neural Network model with the best architecture for each of the supposed settings, depending on the amount of such prior side information. We show that using additional side information improves the final accuracy of the Received Signal Strength map reconstruction, especially in sub-areas close to the gateways where larger variations of the average received signal power are usually observed (ultimately, with a prominent beneficial impact onto positioning performance accordingly).

1 Introduction

Retrieving the exact position of the connected objects has become an important feature of the Internet of Things (IoT). Such connected objects have indeed been widespread over the last few years thanks to the low cost of the radio integrated chips and sensors and their possibility of being embedded in plurality of the devices. For the location-dependent application and services these abilities to associate accurate location with physical data gives huge opportunities. One of the localization techniques is Global Positioning System (GPS) which has been widely used over the past decades, but it suffers from high energy consumption which is not suitable for IoT applications. As an alternative, one can opportunistically measure location-dependent radio metrics, like Received Signal Strength (Indicator) (RSSI), Time (Difference) of Arrival (TDoA), Angle of Arrival (AoA), etc. Based on these metrics, there exist several methods to determine the node position: trilateration, triangulation, proximity detection, or fingerprinting. We will focus on fingerprinting [1] which requires (ideally) full map of mentioned above radio metrics, covering the zone of interest. However, collecting metrics in each point of the zone of interest is

impractical and time costly in real-world scenarios, therefore most approaches rely on sparse and non-uniformly distributed measurements. In this sense, classical map interpolation techniques such as Radial Basis Function (RBF) [2] or kriging [3] are used. Although these methods are relatively fast, they are quite weak in retrieving and predicting the complex and heterogeneous spatial patterns that are usually observed in real life signals. Another approach consists on deterministic simulation such as Ray-Tracing (RT) tools [4]. Given some real field measurements and then calibrated over them, these models predict the radio propagation while simulating electromagnetic interactions with the environment. These technologies, however, need a complete description of the environment (properties of the materials of the obstacles, shape, etc.); moreover, they are complex computationally. Recently, studies have employed machine learning for this task by considering radio maps as images and adapting neural network (NN) models that have been proposed for image completion. These models are based on the fully generated dataset by RT tools for predicting the signal propagation given the buildings mask and position of the transmitter [5]; or predicting the received power value for the LTE (Long Term Evolution) signal with use of additional information and NNs [6] with handcrafted structures. In this work we will focus on the RSS map reconstruction, where only small amount of ground truth GPS-tagged measurements are available preventing to use existing NN models with handcrafted architectures and for which RT models could not be applied (due to the lack of information or high computational complexity). Our approach is based on Neural Architecture Search (NAS) [7] which aims to find an optimized NN model for this task – it is feasible to learn model parameters while exploring the architecture. In this paper, we provide:

1. A unified framework with the use of side information, for which we study the generalization ability of a neural network model which architecture is optimized over labeled and unlabeled data. This is an extension of the work of [8], where a NN was developed to exclusively find the RSSI measurements for a given unlabeled set without the use of side-information.

2. Empirical evaluation over two large-scale RSSI collections showing that the proposed approach is highly competitive compared to the SotA models in terms of quality metrics.

2 NAS with Genetic Algorithm for RSSI map reconstruction using side information

Additional information could be represented in different manners, and they could be included into the algorithm in a variety of ways, such as independent channels, parallel channels inputs, directly in the learning goal, or in the ranking metric during model selection. We adapted the proposed algorithm presented in [8] for multi-channel input by combining additional context information with the data in the model's input; and we assessed the model's performance on unseen base stations that were not utilized in the learning process.

Following [9, 10], we suppose to have a small set of n available base stations $(X^j)_{1 \leqslant j \leqslant n}$. For each given matrix of base station $X^j; j \in \{1, \ldots, n\}$, let $Y^j \in \mathbb{R}^{H \times W}$ be its corresponding 2D matrix of signal strength values measurements, where $H \times W$ is the size (in number of elements in a grid) of the zone of interest. In practice, we have access only to some ground truth measurements Y^j_m, meaning that $Y^j_m = Y^j \odot M^j$, with $M^j \in \{0,1\}^{H \times W}$ a binary mask of available measurements, and $\odot$ is the Hadamard's product. Here we suppose sparsity meaning that the number of non-null elements in Y^j_m is much lower than the overall size $H \times W$. For each base station X^j we estimate unknown measurements $\tilde{Y}^j_u$ in Y^j with a RBF interpolation given (X^j_m, Y^j_m), so that we have a new subset $(X^j_u, \tilde{Y}^j_u)$, where $X^j_m = X^j \odot M^j$ is the associated 2D node locations of Y^j_m in X^j, and the values in $\tilde{Y}^j_u$ are initially given by RBF predictions on X^j_u

corresponding to the associated 2D node locations (or equivalently, the cell/pixel coordinates) with respect to the base station X^j which do not have measurements. In our semi-supervised setting, the values for unknown measurements in $\tilde{Y}_u^j$ will evolve by using the predictions of the current NN model during the learning process.

We further decompose the measurements set Y_m^j into two parts: Y_ℓ^j (for *training*), Y_v^j (for *validation*), such that $Y_\ell^j \oplus Y_v^j = Y_m^j$, where $\oplus$ is the matrix addition operation. Let X_ℓ^j, X_v^j be the associated 2D node locations pf Y_ℓ^j and Y_v^j in X^j.

In our experiments the number of base stations n is small, so in order to increase the size of labeled and pseudo-labeled training samples, we cut the initial measurements maps $(Y_m^j \oplus \tilde{Y}_u^j)_{1 \leqslant j \leqslant n}$ into smaller matrices which resulted into the sets $(S^{j,i})_{\substack{1 \leqslant j \leqslant n \\ 1 \leqslant i \leqslant m_j}}$ where the sets $S^{j,i} \subseteq Y_m^j \oplus \tilde{Y}_u^j; \forall i \in \{1, \ldots, m_j\}$ are shifted with overlapping of the points. Each submatrix $S^{j,i}$ is hence divided into labeled, $S_\ell^{j,i} \cup S_v^{j,i}$, and pseudo-labeled (first interpolated points using RBF and then using the predictions of the current NN model) $S_{\tilde{Y}_u}^{j,i}$. To each submatrix $S^{j,i}$ corresponds a 2D location $X^{j,i} \subset X$. Figure 1 gives a pictorial representation of the notations.

Our objective is to find the optimal architecture for encoder-decoder architecture with skip connections using these side-information and study the generalization ability of obtained models for RSSI map reconstruction.

Side information. As additional context (or side) information, we have considered: (i) information about buildings presence, which was taken from the open-source OpenStreetMap

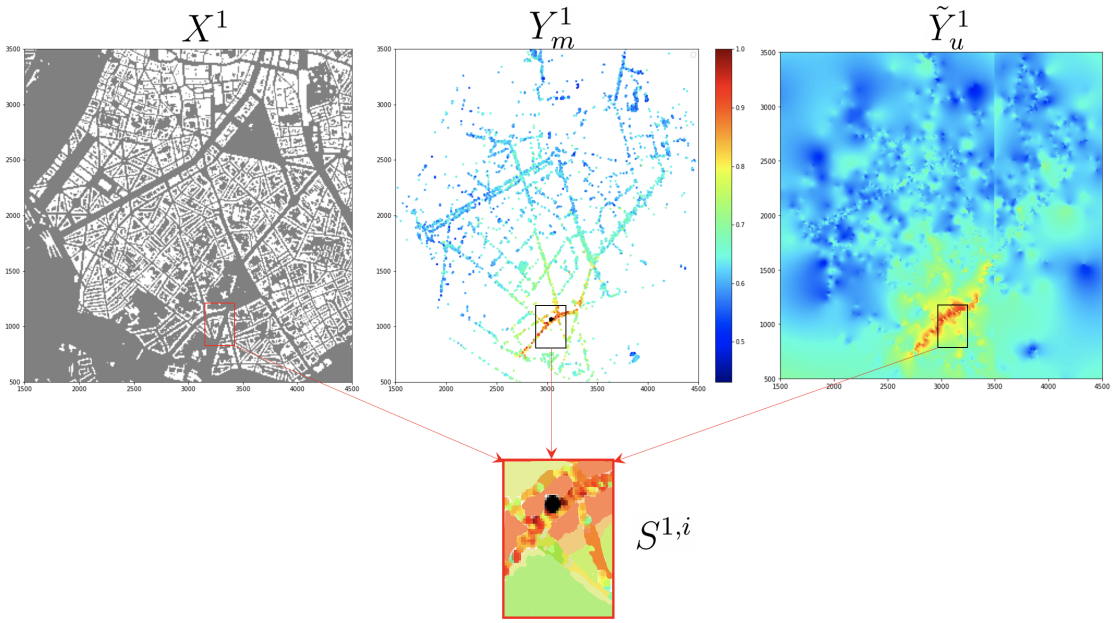

Figure 1: An example of constituting the training sets for one base station. X^1 corresponds to 2D node locations, buildings are shown in white. Y_m^1 is RSSI map (true measurements); the base station is shown by a black circle,and $\tilde{Y}^1$ corresponds to interpolated points found by RBF. Colors depict the strength of the signal from dark red (highest) to deep blue (lowest) RSSI values. $S^{1,i} = S_\ell^{j,i} \cup S_v^{j,i} \cup S_{\tilde{Y}_u}^{j,i}$ is one sub-matrix of partially labeled training data found from $Y_m^1 \cup \tilde{Y}^1$.

Algorithm 1: $\texttt{SL}_{\texttt{NAS}}{}^{ind}$

Input: A labeled training set with given measurements: $(X_m^j, Y_m^j)_{1\leqslant j\leqslant n}$ and an unlabeled set $(X_u^j)_{1\leqslant j\leqslant n}$

Init: Using $(X_m^j, Y_m^j)_{1\leqslant j\leqslant n}$, find interpolated measurements $(\tilde{Y}_u^j)_{1\leqslant j\leqslant n}$ over $(X_u^j)_{1\leqslant j\leqslant n}$ using the RBF interpolation method;

Step 1: Cut the initial measurements maps $(Y_m^j \oplus \tilde{Y}_u^j)_{1\leqslant j\leqslant n}$ into smaller matrices: $(S^{j,i})_{\substack{1\leqslant j\leqslant n\\ 1\leqslant i\leqslant m_j}}$.

Step 2: Search the optimal NN architecture using $(S^{j,i})_{\substack{1\leqslant j\leqslant n\\ 1\leqslant i\leqslant m_j}}$;

Scenario 1: Find the parameters $\theta_1^\star$ of the NN model f_θ :

$$\theta_1^\star = \text{argmin}_\theta \, \mathcal{L}(f_\theta, S_\ell \cup S_{\not\ell}^{j,i}) \quad \# \text{ (Eq. 1)};$$

Scenario 2: Apply $f_{\theta_1^\star}$ on unlabeled data and obtain new pseudo-labeled measurements $S_{\not\ell}^{j,i}$ and find the new parameters $\theta_2^\star$ of the NN model f_θ :

$$\theta_2^\star = \underset{\theta}{\text{argmin}}\, \mathcal{L}(f_\theta, S_\ell \cup S_{\not\ell}^{j,i})$$

Output: $f_{\theta_1^\star}$ for scenario 1 or $f_{\theta_2^\star}$ for scenario 2.

dataset; (ii) amount of crossed buildings by signal from base station (BS) to each point of the map; (iii) information about distance from the BS; (iv) information about the relief represented by DSM (digital surface model). This information was taken from the open-source dataset provided by Japan Aerospace Exploration Agency with 30m accuracy.

Neural architecture search procedure. NAS is performed with a Genetic algorithm is similar to one presented in [8] as it gave better performance result in terms of obtained accuracy.

From the sets $(S^{j,i})_{\substack{1\leqslant j\leqslant n\\ 1\leqslant i\leqslant m_j}}$ we use an evolutionary algorithm similar to [11] for searching the most efficient architecture represented as a Direct Acyclic Graph (DAG). Here, the validation sets $(S_v^{j,i})_{\substack{1\leqslant j\leqslant n\\ 1\leqslant i\leqslant m_j}}$ are put aside for hyperparameter tuning. The edges of this DAG represent data flow with only one input for each node, which is a single operation chosen among a set of candidate operations. We consider usual operations in the image processing field, that are a mixture of convolutional and pooling layers. We also consider three variants of 2D convolutional layers with kernels of size 3, 5 and 7, and two types of pooling layers that compute either the average or the maximum on the filter of size 4.

Candidate architectures are built from randomly selected operations and the corresponding NN models are trained over the set $(S_\ell^{j,i})_{\substack{1\leqslant j\leqslant n\\ 1\leqslant i\leqslant m_j}}$ and its (possible) combinations with side information. The resulted architectures are then ranked according to pixel-wise error between the interpolated result of the outputs over $(S_v^{j,i})_{\substack{1\leqslant j\leqslant n\\ 1\leqslant i\leqslant m_j}}$ and interpolated measurements given by RBF interpolation by filtering out the buildings. As error functions, we have considered the Mean Absolute Error (MAE) or its Normalized version (NMAE) where we additionally weight the pixel error according to the distance matrix value. Best ranked model is then selected for mutation and placed in the trained population. The oldest and worst in the rank are then removed to keep the population size equal to 20 models.

Once the NN model with the optimized parameters are found by NAS, f_θ, we consider the

following two scenarios for learning its corresponding parameters θ by minimizing

$$\mathcal{L}(f_\theta, S_\ell \cup S_{\mathcal{U}}^{j,i}) = \frac{1}{n}\sum_{j=1}^{n}\frac{1}{m_j}\sum_{i=1}^{m_j}\left[\frac{1}{|S_\ell^{i,j}|}\sum_{(x,y)\in S_\ell^{i,j}}\ell(y, f_\theta(x)) + \frac{1}{|S_{\mathcal{U}}^{j,i}|}\sum_{(x,\tilde{y})\in S_{\mathcal{U}}^{j,i}}\ell(\tilde{y}, f_\theta(x))\right] \quad (1)$$

These two scenarios relate to obtaining model parameters on *labeled and pseudo-labeled* measurements using just RBF interpolated data (scenario 1) or predictions from a first model learnt on these data (scenario 2). The overall learning process is depicted in Algorithm 1.

3 Experimental Results

Datasets. In the area of data-based research and in the field of machine learning singularly, it is usually hard to find large open-source datasets made of real data. In some works however, alternatively (or as a complement) to using real data, synthetic data can be generated, for instance through deterministic simulations (as already mentioned in Introduction). As the two possible testing environments we considered two distinct datasets of LoRa signals in urban environments: Grenoble (which was collected during this work) with 4 BSs and Antwerp [12] with 9 BSs.

The size of maps are 700×700 pixels for the dataset collected in Antwerp and 368×368 for Grenoble. For that, we aggregated and averaged the power of collected measurements in cells/pixels of size 10 meters ×10 meters by the measured distance from base station location based on local ENU coordinates for Grenoble and Antwerp. After the aggregation, for Grenoble maps the amount of available measurements in the zone of interest varies from 2728 to 7266,

city	name	all points	train points	validation points	status
Grenoble	BS_1	6264	5591	673	train
	BS_2	2728	2448	280	train
	BS_3	7266	6516	750	train
	BS_4	6836	6096	740	test
Antwerp	BS'_1	6060	5440	620	train
	BS'_2	5606	5034	572	train
	BS'_3	7548	6785	763	train
	BS'_4	2539	2276	263	train
	BS'_5	2957	2667	290	train
	BS'_6	4940	4453	487	train
	BS'_7	3154	2829	325	train
	BS'_8	8277	7455	822	train
	BS'_9	4335	3888	447	test

Table 1: Summary over the settings for different cities: total amount of available measurements, points used as an input to the models, validation (test) points that were used also in the computation of the loss (during the evaluation)

and for the Antwerp base stations the amount is from 2539 to 8277 (the details are shown in Table 1). As we also consider the generalization task, the algorithm should learn from all the available base stations data simultaneously, but to test its accuracy, we left from each city one base station that was not used in the training process.

In our settings, we only have access to several base stations lacking several orders of magnitude in size compared to aforementioned datasets. To artificially overcome this drawback, we created submatrices of the original images by cutting them into smaller ones (we tested over 96 by 96 pixels size because of memory issues during learning of the neural network for the storing of the model weights). We also added the flipped and mirrored images and we also did a shift in 20 pixels meaning that in our dataset there were overlapping between the images. Moreover, if the amount of pixels with measurements in the initial cutted image was high enough (more than 3% of the presented pixels) then we masked out the randomly sampled rectangle of presented measurements similar to the cutout regularization ([13]). By doing this we force the algorithm to do the reconstructions in the zones without measurements (not only locally) and be more robust to the amount of input data. Matrices of the side information were used in the models as additional channels concatenated with measurements map. Before feeding the data into the algorithm, all the values have been normalized between 0 and 1 in each channel separately before cutting them into smaller sizes to feed into the models.

Evaluation of the results over held out base stations. To evaluate the result we left one base station out of the initial set of each city to compare further the models performances with baselines, namely test Antwerp and test Grenoble. To do this, all the points were divided into two parts, namely train and test points for 90% and 10% respectively. Moreover, to highlight the importance of the zones close to base station (as it was mentioned in the Introduction) we compare the performance of the algorithms over different considered circles around the base station location, namely 200 meters and 400 meters radiuses. We considered state-of-the-art interpolation approaches which are: Total Variation (TV) in-painting by solving the optimization problem, Radial basis functions (RBF) with linear kernel that was found the most efficient, and the k Nearest Neighbours (kNN) regression algorithm. The evolutionary algorithm in the model search phase was implemented using the NAS-DIP package [14]. All experiments were run on NVIDIA GTX 1080 Ti 11GB GPU.

In our experiments, we are primarily interested in addressing the following two questions: (a) does the use of side contextual information aid in the more accurate reconstruction of RSSI maps?; and; (b) to what extent is the search for an optimum NN design effective in the two scenarios considered (Algorithm 1)?

We begin by envisaging scenario 1 of Algorithm 1 and investigating the impact of side information on the performance of the optimized NN model discovered by NAS.

Figure 2 shows the evolution of MAE inside a circular zone of varied radius for $f_{\theta_1^\star}$ trained just on measurements and measurements with distance maps (left) and $f_{\theta_1^\star}$ trained on measurements, distance maps, and building counts or elevation (right) on the test base stations of Antwerp BS_9'. The use of distance maps to supplement measurements improves predictions, which is consistent with our earlier findings. When the third side-information is included, such as height or building counts, we find that the elevation yields better signal estimations than the latter. This is understandable because signal transmission can be severely slowed by building heights.

As a best model obtained by Algorithm 1, scenario 1 we consider the case with three input channels: measurements, distances and elevations and present comparative results with other baselines in Table 2. The lowest errors are shown in boldface. The symbol $^\downarrow$ denotes that the error is significantly greater than the best result using the Wilcoxon rank sum test with a p-value threshold of 0.01. According to these findings, $f_{\theta_1^\star}$ outperforms other state-of-the-art models as well as the UNet model with a handcrafted architecture. These results suggest that the search

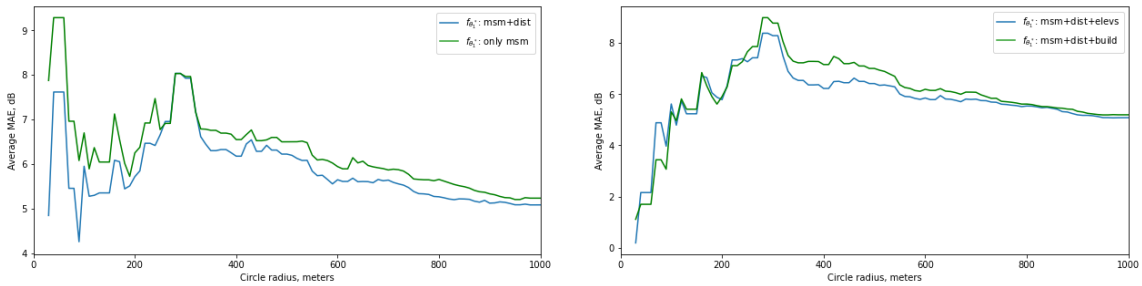

Figure 2: Cumulative MAE distribution of $f_{\theta_1^\star}$ (scenario 1 of Algorithm 1) according to the distance to the test base station for the city of Antwerp BS'_9; with measurements and measurements with distance maps (left) and measurements, distance maps, and building counts or elevation (right).

of an optimal NN model with side-information has strong generalization ability for RSSI map reconstruction.

Model	MAE, dB, 200m	MAE, dB, 400m
RBF	$8.34^{\downarrow}$	$7.04^{\downarrow}$
kNN	$7.98^{\downarrow}$	$7.08^{\downarrow}$
TV	$7.50^{\downarrow}$	$6.97^{\downarrow}$
UNet	$6.37^{\downarrow}$	$6.81^{\downarrow}$
$f_{\theta_1^\star}$	**5.88**	**6.37**

Table 2: Comparisons between baselines in terms of MAE with respect to the three distances to Antwerp' test base station BS'_9. Best results are shown in bold.

Figure 3 (left) depicts the average MAE in dB of all models as well as the NN model $f_{\theta_2^\star}$ corresponding to scenario 2 of Algorithm 1, with respect to the distance to the test base Station BS'_9 for the city of Antwerp. For distances between 200 and 400 meters, $f_{\theta_2^\star}$ consistently outperforms in terms of MAE. Figure 3 (right) presents the average MAE in dB of all models with respect to the distance to the test base station BS4 for the city of Grenoble. These results are consistent with those obtained over the city of Antwerp.

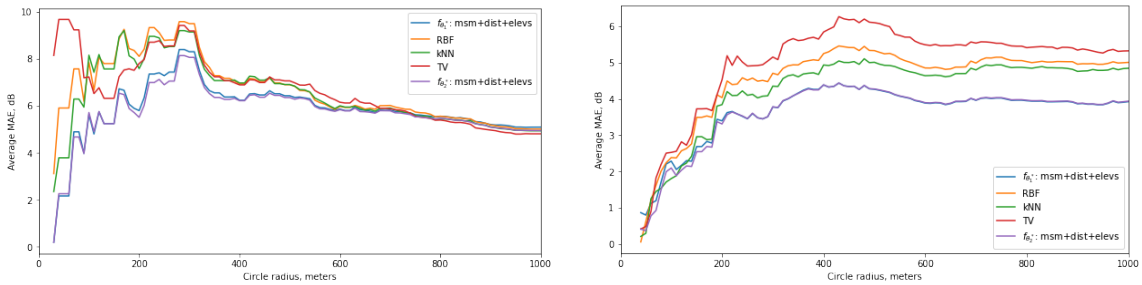

Figure 3: Average MAE in dB of all models with respect to the distance to the test base Station BS'_9 for the city of Antwerp (left) and Average MAE in dB of all models with respect to the distance to the test base Station BS_4 for the city of Grenoble (right).

4 Conclusions

In this paper, we proposed to find an optimal Neural Network model for RSSI map reconstruction using NAS, showing its generalization ability on two distinct datasets. As a future work, we propose to model the pseudo-labeling errors in the predictions [15] for NAS in RSSI.

References

[1] Wongeun Choi, Yoon-Seop Chang, Yeonuk Jung, and Junkeun Song. Low-power lora signal-based outdoor positioning using fingerprint algorithm. *ISPRS International Journal of Geo-Information*, 7(11), 2018.

[2] A. Enrico and C. Redondi. Radio Map Interpolation Using Graph Signal Processing. *IEEE Communications Letters*, 22(1):153–156, 2018.

[3] M.A. Oliver and R. Webster. Kriging: a method of interpolation for geographical information systems. *International Journal of Geographical Information System*, 4(3):313–332, 1990.

[4] S. Sorour. and et al. RSS Based Indoor Localization with Limited Deployment Load. In *Proc. IEEE Global Communications Conference (GLOBECOM'12)*, pages 303–308, 2012.

[5] R. Levie, Ç. Yapar, G. Kutyniok, and G. Caire. Pathloss prediction using deep learning with applications to cellular optimization and efficient d2d link scheduling. In *ICASSP*, pages 8678–8682, 2020.

[6] Tatsuya Nagao and Takahiro Hayashi. Study on radio propagation prediction by machine learning using urban structure maps. In *2020 14th European Conference on Antennas and Propagation (EuCAP)*, pages 1–5, 2020.

[7] Thomas Elsken, Jan Hendrik Metzen, and Frank Hutter. Neural architecture search: A survey. *Journal of Machine Learning Research*, 20(55):1–21, 2019.

[8] Aleksandra Malkova, Loïc Pauletto, Christophe Villien, Benoît Denis, and Massih-Reza Amini. Self-learning for received signal strength map reconstruction with neural architecture search. Lecture Notes in Computer Science, pages 515–526. Springer International Publishing, 2021.

[9] Massih R. Amini, Nicolas Usunier, and François Laviolette. A transductive bound for the voted classifier with an application to semi-supervised learning. In *Advances in Neural Information Processing Systems*, pages 65–72, 2009.

[10] Yury Maximov, Massih-Reza Amini, and Zaïd Harchaoui. Rademacher complexity bounds for a penalized multi-class semi-supervised algorithm. *J. Artif. Intell. Res.*, 61:761–786, 2018.

[11] E. Real, A. Aggarwal, Y. Huang, and Q. V Le. Regularized evolution for image classifier architecture search. In *AAAI*, volume 33, pages 4780–4789, 2019.

[12] M. Aernouts, R. Berkvens, K. Van Vlaenderen, and M. Weyn. Sigfox and lorawan datasets for fingerprint localization in large urban and rural areas. *Data*, 3(2), 2018.

[13] Terrance Devries and Graham W. Taylor. Improved regularization of convolutional neural networks with cutout. *CoRR*, abs/1708.04552, 2017.

[14] Kary Ho, Andrew Gilbert, Hailin Jin, and John Collomosse. Neural architecture search for deep image prior, 2020.

[15] Anastasia Krithara, Massih-Reza Amini, Jean-Michel Renders, and Cyril Goutte. Semi-supervised document classification with a mislabeling error model. In *30th European Conference on Information Retrieval*, pages 370–381, Glasgow, 2008.