IRGAN : cGAN-based Indoor Radio Map Prediction

Cheick Tidiani CISSE
UTBM, CNRS, Institut FEMTO-ST
Orange Lab, Orange
F-90000 Belfort, France
cheick.cisse@utbm.fr

Oumaya BAALA
UTBM, CNRS, Institut FEMTO-ST
F-90000 Belfort, France
oumaya.baala@utbm.fr

Valéry GUILLET
Orange Lab, Orange
F-90000 Belfort, France
valery.guillet@orange.com

François SPIES
Université de Franche-Comté, CNRS, institut FEMTO-ST
F-25200 Montbéliard, France
francois.spies@univ-fcomte.fr

Alexandre CAMINADA
I3S, COATI
Univ. Côte d'Azur, CNRS, Inria
F-06000 Nice, France
alexandre.caminada@unice.fr

***Abstract*—Radio map or radio coverage prediction in indoor and outdoor remains a challenge of great interest due to the large number of applications it allows. Many techniques such as data-driven interpolation methods, model-based data fitting methods, model-based prediction methods exist. However, each of them has their limitations (computation time, accuracy, level of input information required, generalization to different environments). Also indoors seems less studied than outdoors.**

In this paper, a multi-material model and its enhanced version boosted by attention mechanism for indoor radio map prediction are introduced. The proposed models are based on conditional Generative Adversarial Networks (cGANs) and take floor plan images as input. We also propose a new method for floor plan images preprocessing by segmentation to characterize the materials in the environment of interest, which alleviates user's effort. The validity and efficiency of our method in generating high quality radio maps in different environments and its ability to consider the electromagnetic properties of materials are verified on two simulated datasets. Numerical results show that our approach outperforms state of the art methods.

***Index Terms*—Radio map prediction, path loss, indoor planning, propagation, conditional generative adversarial networks, deep learning, floor plan segmentation**

I. Introduction

The development of the 5G and Internet of Thing (IoT) sectors have led to an increase in the number of telecommunications services and networks users. Network planning, optimization is therefore becoming a very important issue for a good Quality of Service (QoS). For example, one can highlight that the need for good QoS in indoor and small offices is increasingly important since the widespread use of fiber access. In telecommunications and wireless networks, QoS is paramount for user experience. Many criteria are known to have an impact on QoS such as the choice of technology (e.g. Wi-Fi, 5G), the environment, the coverage which depends on the placement of the access points (APs) and many others. Since coverage is a key factor, we need to know how to characterize and improve it. The path loss function related to the coverage provides the radio map representing the distribution of a signal through an environment. This map depends on several factors like walls, doors, building materials and many other obstacles that can block the waves. Thus, a detailed description of the environment leads to better quality radio maps. Predicting a reliable radio map in a few milliseconds is possible with deep learning methods and thus, allowing many applications such as network optimization, network planning, interference management, power control, real-time radio map visualization, fingerprint based localization and many others [1][2][3][4]. With conventional methods like ray tracing (RT) widely used, expert assistance is needed especially for the characterization of the materials, simulation parameters selection, optimal locations for AP placement, etc. All these tasks are time consuming and very expensive for telecommunications operators.

Several works have been proposed in the context of radio map prediction (RMP) to alleviate the user effort in particular in term of simplicity of use. There are five main categories of path loss models for RMP:

Empirical path loss models, also called statistical models, combine measurements with mathematical equations to predict path loss but do not consider detailed floor plans or building representation. These models have two advantages: speed of execution and limited dependence on detailed knowledge of the terrain. However, they are not very accurate. The log-distance model [5] and COST-231 [6] are such examples.

Semi-deterministic path loss models consider measurements at given positions to calibrate the model parameters and a few information on the environment in order to predict the path loss. Motley Keenan's [7] and Dominant Path Model (DPM) [8] models are examples of this category. These models represent a good balance between simplicity and precision.

Deterministic path loss models turn out to be very accurate, but require high computing resources and a detailed description of the environment. We can quote for example the ray model [9] which requires a 3D description of the propagation environment.

ISBN 978-3-903176-57-7

Non-generative Deep learning-based models, among the approaches in this category, some are based on Multilayer perceptron (MLP) and others on Convolutional Neural Networks (CNNs). The former are limited and not generalizable to new environments and seem to be unaware of the physical structure of the environment. This is a simple regression problem [10]. The latter stand out by their ability to generalize to other environments and to consider the type of materials and the physical dimensions of the environment.

At a first level of complexity in deep learning based propagation modeling, Teganya *et al.* [11] proposed an Auto-Encoder for radio map completion. [11] evaluates the impact of the number of input measurement points in completion task. In the same vein, Hashimoto *et al.* [12] proposed a CNN-based spatial interpolation method for estimating the radio map that is environmentally independent.

At a second level of complexity, Levie *et al.* [13] proposed two UNet based model, the first one for solving radio map prediction given environment geometry and AP location and the second for radio map completion task given some sparse measurements as additional input. The authors also proposed a new public dataset for outdoor called RadioMapSeer[1] and demonstrated the superiority of their methods to prior state-of-art. The work in [14] propose a more performing UNet based model by replacing the original UNet pooling layers by strided convolutions and combine them with inception block. The work in [15] is, to the best of our knowledge, the first to use visual transformer (ViT) [16] for radio map prediction in urban scenarios. The authors replaced transformer embedding by a new consistent positional embedding for radio map prediction. Concretely, they rely on the capacity of transformers to capture long-range dependencies but still uses UNet and CNNs as baseline. All these methods were developed for outdoor applications. Note that CNNs based methods are generally faster than transformers based methods in term of inference time.

More recently, Bakirtzis *et al.* [10] proposed an efficient way to predict radio map based on UNet architecture with atrous convolutions and trained on ray tracing dataset. The proposed method is designed for indoor and can deal with different floor plans but still requires a user with radio engineering and radio design expertise.

Generative Deep learning based models, the development of generative models, more precisely cGANs, allowed many advances in image-to-image translation problems. The authors in [17] are among the first to use cGANs for radio map prediction simultaneously with the automatic placement of APs from floor plan images as an image-to-image translation task. To achieve this goal, the authors rely on the cGANs model (Pix2Pix) proposed in [18]. In their work, the authors indicate that training the model under the cGANs framework (with a discriminator) is better than training the model without discriminator for RMP task. They also raise the difficulty of their model to handle complex floor plans containing various non-uniform or unseen elements such as texts and annotations. The authors in [19] propose an access-point-centered window radio map generation network to shorten the training time in which the floor plan materials are one-hot encoded. However, the study uses only one environment for learning and testing and does not give any prof of generalizability to other floor plans. The authors in [20] and [21] also use a cGANs for respectively radio map prediction and completion but for outdoor.

Through the proposed approaches, some suffer from generalization, others do not really consider the geometry of the map and are unable to handle complex floor plans with texts and annotations. Moreover, most of them are designed for outdoor and require good knowledge in deep learning and radio propagation. Also the analysis of the proposed methods is generally limited to ray tracing dataset and most of them are UNet-based models.

Tab. I summarises different models presented above. The models are classified according to four aspects: environment (Indoor or Outdoor), cGANs based (yes or no), generalizability (low, medium, high) of the model to other environments and APs locations, and finally the use case considered: radio map prediction (RMP), radio map completion (RMC) or automatic cells placement (ACP).

TABLE I: Models comparison

Model	Env	cGANs	Generalization	Purpose
[17]	Indoor	yes	medium	ACP
[19]	Indoor	yes	low	RMP
[10]	Indoor	no	high	RMP
[22], [13], [14]	Outdoor	no	medium	RMP
[20]	Outdoor	yes	medium	RMP
[11], [12]	Outdoor	no	low	RMC
[21]	Outdoor	yes	high	RMC
[23]	Outdoor	yes	medium	RMC
[13]	Outdoor	no	medium	RMC

In this paper, we propose two models for radio maps prediction named IRGAN (Indoor Radio cGAN) and the enhanced version of IRGAN named E-IRGAN, which are based on cGANs framework. Our models are designed to deal with different indoor environments and AP locations. We introduce the use of attention mechanism [24] coupled to dilated convolution also called atrous convolution layers to boost our baseline method performance. To make our models helpful for non-experts, we propose a floor plan pre-processing method based on segmentation. The proposed framework shows the highest accuracy among state-of-the-art (SOTA) models.

The main contributions of our work are as follows.

- We introduce a new way to process floor plan images through segmentation and a technique for a consistent encoding of building materials that allows a better description of environment. The segmentation model helps cleaning noise and extracting useful part (walls, doors) of floor plan images.
- We design new cGANs architectures for accurate indoor radio map prediction capable of operating on new en-

[1] $https://RadioMapSeer.github.io$

vironments with different AP locations based solely on floor plan image, building physical dimensions and AP location as inputs.
- We propose a dataset for floor plan images segmentation. We show the ability of our models to deal with different types of building materials and present a deep analysis of their behavior according to materials.
- Unlike prior studies [17] [15] [22], we consider two simulated datasets developed specifically for home and small offices environments in order to show the effectiveness of our method.

The rest of the paper is organized as follows. The proposed neural networks architectures are described in Section II. The loss functions are discussed in Section III. The experiments are detailed in Section IV where the datasets and pre-processing methods are introduced. The results are presented and analyzed in Section V. Finally, Section VI highlights our main contributions and provides guidelines for future work.

II. Proposed Networks Architecture

Generative Adversarial Networks (GAN) are powerful generative models in image-to-image translation tasks such as segmentation [25], colorization [18], super-resolution [26], filling missing image data and many others. A conditional Generative Adversarial Networks (cGAN) is composed of 2 models as illustrated in Fig. 1: a generator that learns to generate data and a discriminator that gives feedback to the generator about the quality of the generated data. Thus, the quality of the generated data depends not only on the capacity of the generator but also on that of the discriminator. In our work we rely on the well-known PatchGAN [18] as discriminator and test several generators.

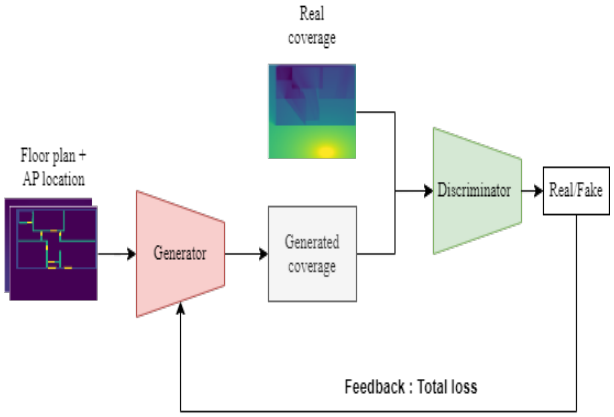

Fig. 1: cGANs Framework

In the search for a new high-performance generator, we tested the UNet-based cGANs as proposed in [18]. First, we proposed a ResNet-like architecture as baseline named IRGAN which is inspired by the work in [27] with few modifications. Then, we proposed a new breakthrough architecture named Enhanced IRGAN (E-IRGAN) combining the advantages of dilated convolution layers, skips connection and attention mechanism. Dilated Convolutions are a type of convolution that expands the kernel size by inserting holes between the kernel elements. This way, we increase the receptive field without increasing the number of parameters compared to standard convolutions. The parameter named dilation rate indicates how much the kernel is widened (dilated). The attention mechanism helps to decode and focus on the most relevant part of a feature map.

To compare our models to state of the art, we have included in our study two other recent proposed architectures namely SegNet [10] and DeepRay [22] for radio map prediction.

IRGAN and E-IRGAN have several common elements. Referring to Fig. 2 and 3, the decoder/encoder, and the initial blocks are very similar: While IRGAN uses InstanceNormalization (IN), E-IRGAN uses BatchNormalization (BN) after each convolution layer. In addition, we apply L1 kernel regularization on the convolution layers in E-IRGAN which results in a clear improvement of the performance.

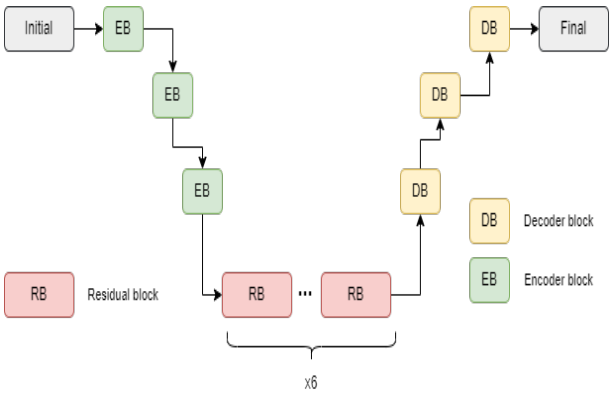

Fig. 2: IRGAN generator architecture.

The initial block is composed of reflection padding + Conv2D + IN/BN + Leaky ReLU. The down sampling blocks (encoder block) are made up by Conv2D + IN/BN + Leaky ReLU. The up sampling blocks (decoder block) are formed by two Conv2D + IN/BN + Leaky ReLU. The final block is composed by one Conv2D followed by Tanh activation function. Gate blocks are composed of Conv2D + BN + Leaky ReLU. Residual blocks are a sequence of two times Conv2D + IN + Leaky ReLU. Fig. 4 and 5 shows respectively the architecture of attention blocks and dilated convolution blocks.

The proposed model architecture is the result of an experimental test process. We started from a baseline model that was experimentally improved by applying several operations of adding, modifying, and deleting layers.

For the segmentation task, we chose the $DeepLabV3+$ model which performs well on semantic segmentation benchmarks. The implementation can be found on keras website[2]. We do not give here more details on the segmentation model but it should be noted that automatic floor plan processing is not a straightforward task due to the amount of floor plan

[2] $https://keras.io/examples/vision/deeplabv3\_plus/$

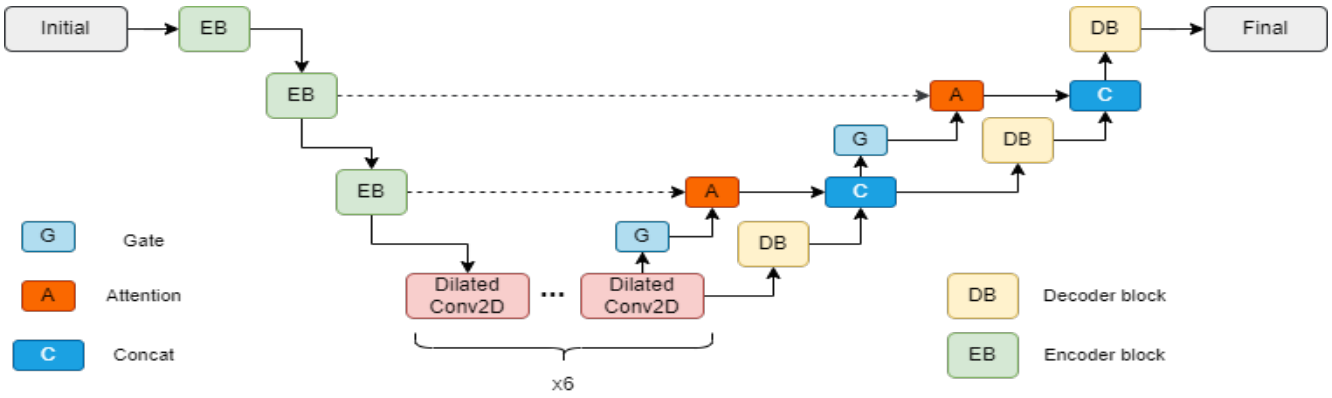

Fig. 3: E-IRGAN generator architecture.

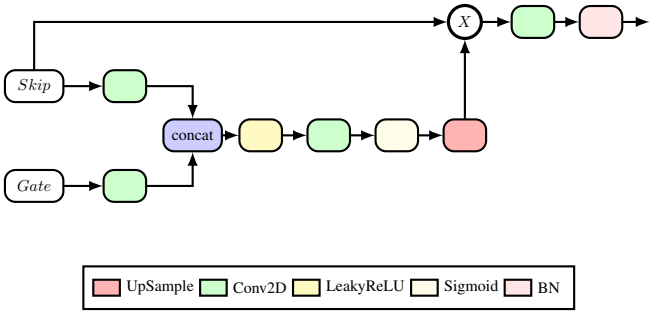

Fig. 4: Attention block.

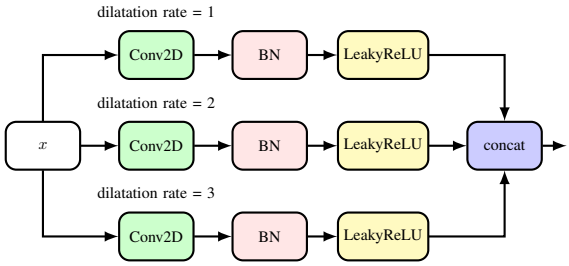

Fig. 5: Dilated conv2D block.

styles, resolutions etc [28]. Once implemented $DeepLabV3+$ model showed good performance on our dataset.
It is worthy to note that our radio map prediction models can be used without our segmentation model as long as the input floor plan image has been pre-processed as detailed in section IV-B.

III. Loss functions

The objective of a cGAN can be expressed by (1). Previous approaches highlighted the importance of mixing the GANs objective with more objective functions such as L2 [29], and L1 [18]. The learning of the discriminator (D) remains unchanged. The generator (G) is given the task of not only fooling the discriminator but also predicting near ground truth images by conforming to the additional losses. In the same way, the objective of the proposed cGAN models is composed of 3 loss functions presented below.

A. Adversarial loss

With x as the input, y the target output and $\mathbb{E}$ the expectation, the adversarial loss of a cGANs can be expressed as:

$$L_{cGAN}(G,D) = \mathbb{E}_{x,y} log(D(x,y)) + \mathbb{E}_{x} log(1 - D(x, G(x))) \quad (1)$$

We are in a minimax game where the generator tries to confuse the discriminator by predicting data close to the real one, and the discriminator tries to discriminate between generated (fake) and real data. The equilibrium is reached when the generator succeeds in generating data in such a way that the discriminator can no longer distinguish between real data and generated data.

B. L1 loss

We use L1 as in [18] to encourage the generator to produce images that are close to the target in addition to its task of fooling the discriminator. The L1 loss can be expressed as:

$$L_1 = \mathbb{E}_{x,y\ p_{data}(x,y)}\{||y - G(x)||_1\} \quad (2)$$

C. Structural similarity (SSIM)

SSIM [30] calculates the similarity between two images. We use a loss function based on SSIM to reinforce the similarity of structure between the predicted radio map and the ground truth radio map which is given by

$$L_{SSIM} = 1 - SSIM(y, G(x)) \quad (3)$$

D. Total Loss

Finally, the total loss of our cGAN models is a weighted sum of the above loss functions and is given by the following equation:

$$L = L_{cGAN}(G,D) + \lambda_1 L_1 + \lambda_2 L_{SSIM} \quad (4)$$

Experimentally the values $\lambda_1 = 10$ and $\lambda_2 = 10$ turn out to be good.

IV. Experiments

At this stage of work, it is now necessary to introduce the datasets description, the input data processing and the implementation details.

A. Datasets description

Making radio measurements in various environments is time-consuming and very expensive. For the study purposes, due to the unavailability of public datasets for indoor unlike outdoor, we opted for creating simulated datasets based on very accurate path loss models while waiting for our own real-world measurements. It is worth noting that fingerprinting datasets [31] for localization are not suitable for our study because they contain neither the floor plan of the building nor the position of the APs and their characteristics. Also the great dispersion of Received Signal Strength (RSS) according to the devices used for measurement, as studied in [32], makes useless any dataset without accurate information on the measurements process and the receiving device. As part of the training testing and analysis of our deep learning models, we set up two simulated datasets with various building materials typical for home and small office environments. The datasets are based on Ray Tracing (RT) and Dominant Path Model (DPM), two propagation models provided by the simulation software Winprop [33], [34]. Note that the RT is a deterministic propagation model and the DPM is a semi-deterministic propagation model. Fig. 6 shows RT and DPM simulation results on the same environment with the same AP locations. Depending on the surface of the floor plans, the average prediction time of the radio map with the RT model is about 300 times greater than the DPM model. For the training of the segmentation model, we set up an annotated dataset whose floor plans come mostly from the CubiCasa5K[3] public dataset [35]. Our segmentation dataset generated contains around 200 annotated floor plans.

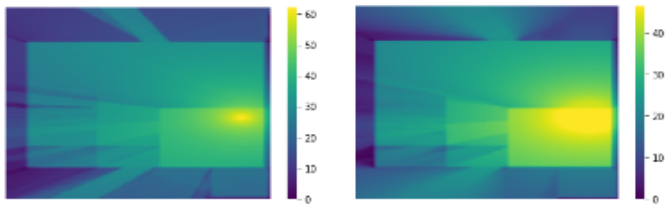

Fig. 6: Illustration of the level of details between RT and DPM models. From left to right : RT truth path loss, DPM truth path loss in dB.

To generate the simulated datasets, we defined some settings summarized in Tab. II. The path loss is calculated at the height of 1.2 m as to mimic a phone in hand. The frequency used is 2.4 GHz. In terms of type of materials for the walls and doors description, we use two types of materials for internal and external walls. We set the material type of doors to wood. The resolution of the generated radio coverage maps is 0.05 m.

[3] $https://github.com/CubiCasa/CubiCasa5k$

In other words, we generated simulated measurements every 5 cm from the AP position, which is a very fine resolution. Our datasets include 24 different floor plans. Each floor plan includes 4x26 radio maps with different APs locations. We consider 4 different scenarios in terms of floor plan materials composition: concrete-brick, concrete-plaster, brick-brick, brick-plaster. For each scenario we generate 26 radio maps among which 16 AP locations are the same in the remaining scenarios of the concerned floor plan. Among the 16 locations, 10 are inside the building and 6 outside, the idea behind is to characterize the impact of all building material types. The other 10 locations are random. We selected 21 floor plans for training, and the 3 remaining for testing. So, we obtain 312 (26x4x3) samples for test purpose. It is noteworthy that the floor plans used for the test stage are different from those used in the learning stage with different AP locations. To increase the number of samples, we apply random data augmentation during the training by flipping and rotating the images.

TABLE II: Dataset settings

General settings			
Tx height	1 m	**Rx height**	1.2 m
Frequency	2.4 GHz	**Resolution**	0.5 cm
Inner walls	Brick - 10 cm Plaster - 2.6 cm	**Outer walls**	Concrete - 20 cm Brick - 20 cm
Propagation Model	DPM, RT	**Signal**	Path Loss
Min surface	103 m^2	**Max surface**	500 m^2

B. Input feature engineering

Knowing that the performance of deep learning models depends a lot on the quality of the data but also on the feature representation, we decide to perform a few pre-processing steps illustrated in Fig. 7. First, we use our segmentation model to extract doors, internal and external walls from raw floor plan images. The segmentation model enables to do away with plan vectorization and use only simple floor plan scans or photos. Then, to integrate the location of the AP and the physical dimensions of the building, we set up one more channel called free-space channel. We set to 1 the pixel corresponding to the AP location, and the other pixels values are calculated from the free-space path loss equation: $20 * log10(d) + 20 * log10(f) + 32.44$ where d is the distance in meters between the transmitter and the pixel concerned, and f the frequency. Regarding the plan encoding, the pixels corresponding to each materials are encoded by the value of their permittivity. As an example, pixels belonging to a concrete wall takes the permittivity of concrete as value. The permittivity for concrete, brick, plaster, wood is respectively 5.24, 3.91, 2.73, 1.99. Note that the thickness of the same material is different when inside or outside. To take walls thickness into account, the value of the permittivity is multiplied by 2 when it is an external wall. Pixels identifying background are set to 0. Added to the free-space channel, the radio map generator in Fig. 7 has two channels as an input and one for the output. As deep learning models have fixed input size, we resize all images to 256x256 and normalize the

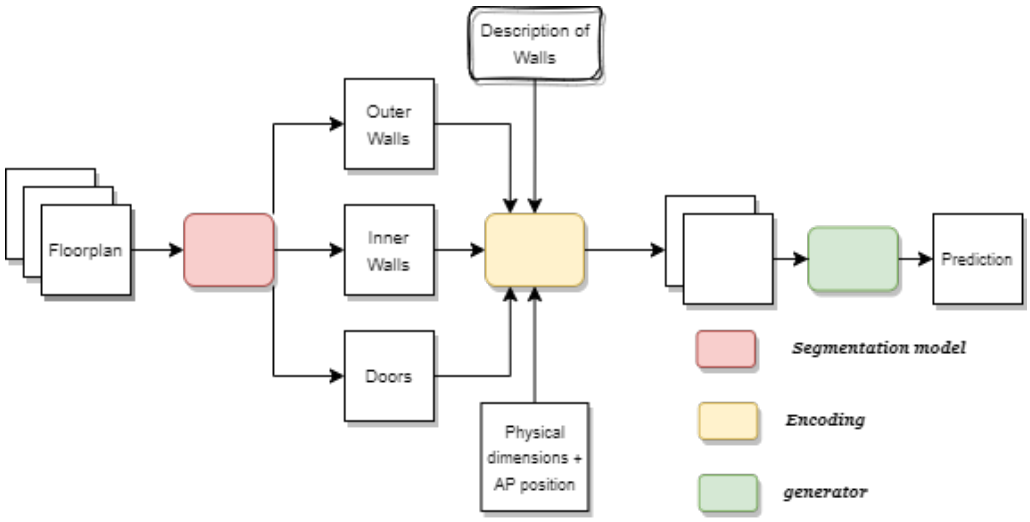

Fig. 7: Illustration of input processing and how it is fed to the generator.

range of pixels value to [-1,1]. For visualization purposes, we inverted the value scale of the radio map in the figures as depicted in Fig. 8.

C. Implementation Details

In our attempt to be fair, we followed the implementation details of the authors for SegNet [22], and DeepRay [10] models. However the batch size is set to 5 for all the models.

To train the cGANs models we use Adam Optimizer [36] with $\beta_1 = 0.5$ and $\beta_2 = 0.99$. The learning rate is set to $1e-4$ and decrease progressively to $1e-8$. The learning rate decreases by factor 0.5, when the error on test set does not decrease after two epochs. We fix the number of epochs to 80. Models implementation are based on TensorFlow [37]. The training of the models is realized on HP Z8 workstation with 2 NVIDIA Quadro P5000 GPUs.

V. Analysis and Results

This section is dedicated to the evaluation of the performance of the tested models.

In order to assess the models, we use some conventional performance metrics for image quality assessment: Peak Signal-to-Noise Ratio (PSNR) which is the ratio of the highest possible signal power to the noise power, SSIM [30] which measures the similarity between two given images. SSIM is bounded between 0 and 1 and SSIM values close to 1 mean high similarity, and Root Mean Square Error (RMSE) in dB which is a standard metric used in other work [10][14][22] to evaluate models performance.

Tab. III summarizes the performance of the tested models on DPM and RT dataset. Train and test set contain respectively 2184 and 312 samples. Models with $*$ have been trained under cGANs framework. As shown in Tab. III, the E-IRGAN* model achieves the best performance in both datasets followed by IRGAN*. We can note with the UNet and UNet* that training a model under cGANs framework yields to better performance in RMP task.

Fig. 13 shows E-IRGAN prediction results on different floor plans. One can observe the consistency and the capacity of our model to deal with unseen floor plans. We can also observe that the model behaves according to the type of building materials, which can be seen in the surroundings of the doors in Fig. 13.

Tab. IV and V show respectively the sensitivity of IRGAN and E-IRGAN models according to the type of floor plan materials. Numerical results show that the more the materials used are attenuating the more the model makes errors. Indeed with strong attenuating materials, the local perturbations around the walls become stronger resulting in more errors. To bring more elements of analysis in Fig. 9 and 10, we calculate the accuracy of the models on 6 thresholds of tolerance from 1 to 6 dB. This accuracy is determined by calculating the absolute errors map between the reference image and the prediction, and then dividing the number of pixels under the threshold by the total number of pixels of the radio map. It appears E-IRGAN performs the best whatever is the tolerance threshold, while IRGAN overcomes the performance of the other models only for a tolerance threshold value greater than 3.

As illustrated in Fig. 3, the number of parameters, the complexity and the performance of E-IRGAN model rely on two elements: the number of encoding/decoding blocks and the number of dilated blocks. Fig. 11 provides the RMSE function of the number of encoder blocks and the number of dilated blocks. We can observe performance improvement with the increase of these 2 parameters. The cost of this performance improvement is an increase in the number of parameters and the complexity of the model. Note that for this part, the model is trained only for 30 epochs.

Fig. 12 shows the path loss predictions of tested models on DPM dataset. We can evaluate that in the context of cGANs, all the models perform better and produce smoother radio maps. Especially, IRGAN and E-IRGAN outperform the other models.

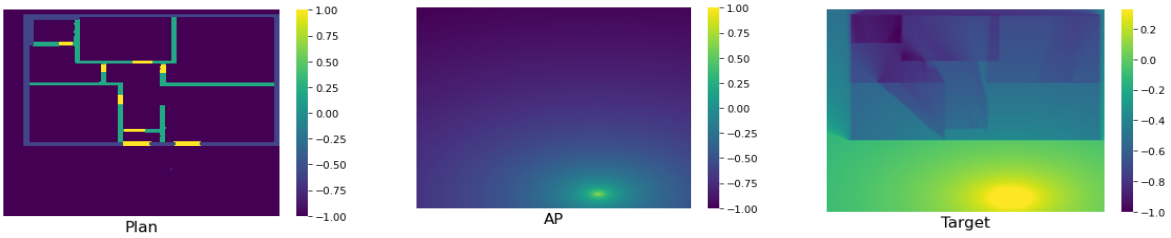

Fig. 8: Example of pairs from DPM dataset. From left to right: Encoded floor plan, Access point and physical dimension, target.

TABLE III: Models performance on RT and DPM datasets.

	RT						DPM						#parameters	Inference
Models	Test set			Train set			Test set			Train set				time
	PSNR	SSIM	RMSE	PSNR	SSIM	RMSE	PSNR	SSIM	RMSE	PSNR	SSIM	RMSE		seconds
SegNet	24.15	0.922	6.23	26.10	0.936	5.83	25.96	0.944	4.72	26.88	0.951	4.81	**15 M**	0.17
DeepRay	25.83	0.923	4.95	26.52	0.928	5.47	25.24	0.936	5.18	25.98	0.946	5.34	21 M	0.19
UNet	25.11	0.927	5.84	25.93	0.933	5.93	26.29	0.950	4.57	28.88	0.960	3.79	54 M	0.19
UNet*	28.54	0.956	3.82	35.86	0.979	**1.86**	28.24	0.969	3.72	27.9	0.968	4.36	54 M	0.19
IRGAN*	28.88	0.963	3.52	32.38	0.978	2.78	28.32	0.974	3.64	34.55	0.989	2.12	32 M	**0.13**
E-IRGAN*	**30.77**	**0.963**	**2.94**	**36.11**	**0.985**	1.95	**30.47**	**0.975**	**3.02**	**37.55**	**0.992**	**1.50**	39 M	0.26

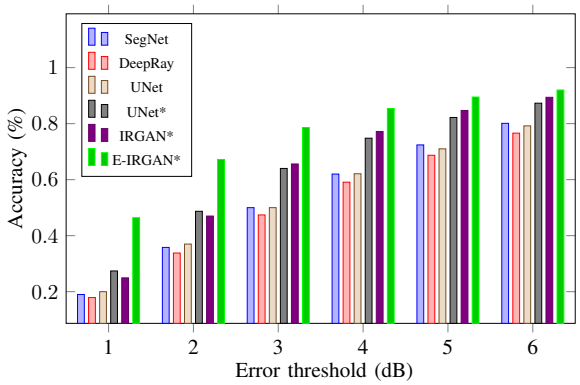

Fig. 9: Models accuracy on DPM test set.

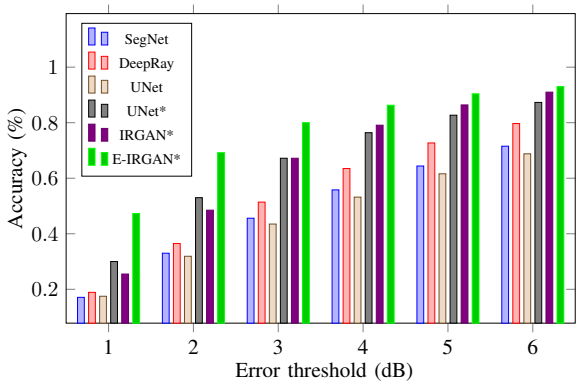

Fig. 10: Models accuracy on RT test set.

TABLE IV: IRGAN performance according to materials type on DPM test set

Materials (outer - inner)	PSNR	SSIM	RMSE
brick - brick	28.62	0.975	3.33
brick - plaster	29.62	0.979	2.78
concrete - brick	27.29	0.969	4.38
concrete - plaster	27.62	0.952	4.08

TABLE V: E-IRGAN performance according to materials type on DPM test set

Materials (outer - inner)	PSNR	SSIM	RMSE
brick - brick	31.56	0.964	2.49
brick - plaster	32.43	0.972	2.08
concrete - brick	29.34	0.955	3.79
concrete - plaster	29.76	0.960	3.41

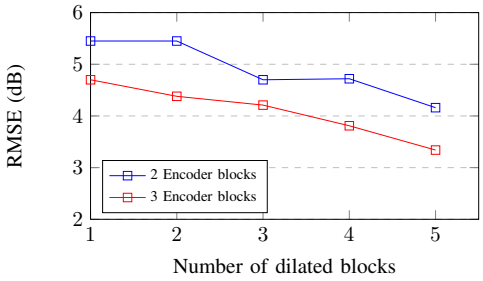

Fig. 11: Complexity analysis.

VI. Conclusion

In this work, two new radio map prediction models named IRGAN and E-IRGAN are proposed with the aim to ease the radio map generation for the non-expert. IRGAN and E-IRGAN are powered by a segmentation model, and have several advantages: ease of use (no need to enter and draw floor plans), speed compared to conventional methods (less than 1 second compared to several minutes for RT and DPM simulated radio maps), accuracy, multi-material capability, and high generalization capability. The introduced segmentation method is designed to reduce input noise and can be seen as a dimension reduction given the diversity of floor plans images we have as input (scan, drawing, photos etc). Two additional loss functions are used to efficiently guide the learning process.

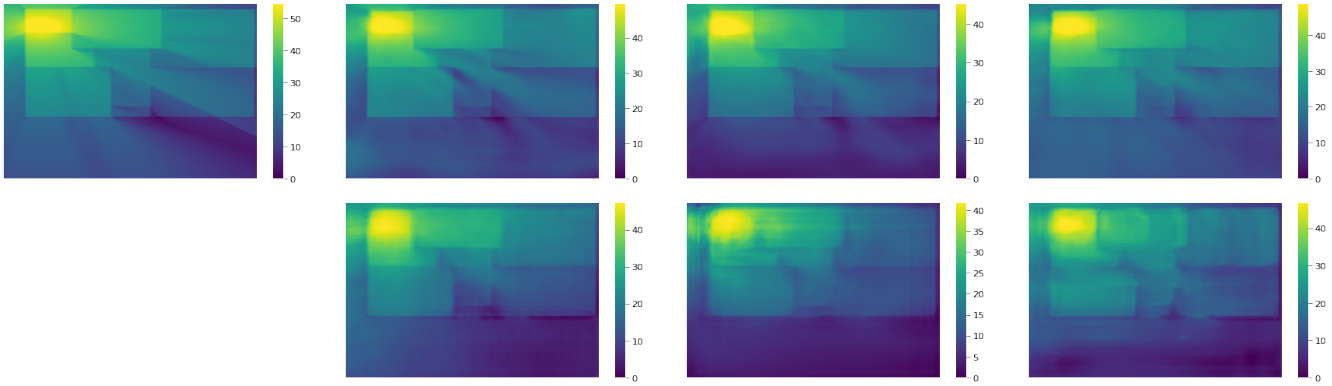

Fig. 12: Models predictions on DPM test set. First row, from left to right : target, E-IRGAN*, IRGAN*, UNet*. Second row, from left to right : UNet, DeepRay, SegNet

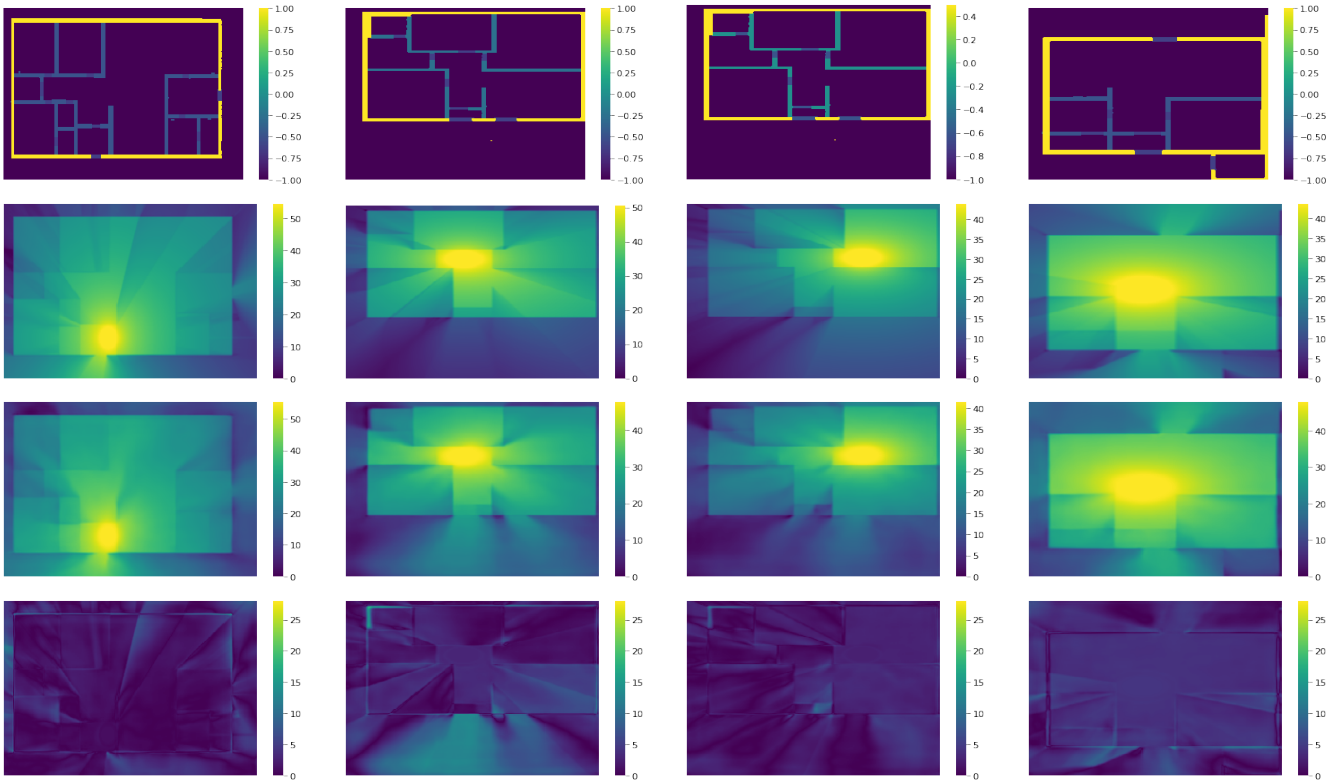

Fig. 13: E-IRGAN predictions on different environments. First row: floor plans, second row: targets, third row: predictions, last row: absolute error maps

The proposed models generate high quality radio maps from only floor plan images, transmitter location and descriptive information such as physical dimensions of the building, and types of materials of the floor plan. Several experiments are conducted based on a variety of environments to show the efficiency of the proposed models to consider different types of building materials. We obtain a RMSE of about ~ 3 dB, which is an improvement of approximately 1.5 dB compared to state of art. In our future works, we plan to extend and train our model on more frequencies and fine tune it to radio measurements.

REFERENCES

[1] B. Berruet, O. Baala, A. Caminada, and V. Guillet, "An evaluation method of channel state information fingerprinting for single gateway indoor localization," *Journal of Network and Computer Applications*, vol. 159, p. 102591, 2020.

[2] T. Van Chien, T. N. Canh, E. Björnson, and E. G. Larsson, "Power control in cellular massive mimo with varying user activity: A deep learning solution," *IEEE Transactions on Wireless Communications*, vol. 19, no. 9, pp. 5732–5748, 2020.

[3] S. Bi, J. Lyu, Z. Ding, and R. Zhang, "Engineering radio maps for wireless resource management," *IEEE Wireless Communications*, vol. 26, no. 2, pp. 133–141, 2019.

[4] I. Kakalou, K. Psannis, S. K. Goudos, T. V. Yioultsis, N. V. Kantartzis, and Y. Ishibashi, "Radio environment maps for 5g cognitive radio

network," in *2019 8th International Conference on Modern Circuits and Systems Technologies (MOCAST)*. IEEE, 2019, pp. 1–4.

[5] P. Series, "Propagation data and prediction methods for the planning of indoor radiocommunication systems and radio local area networks in the frequency range 300 mhz to 450 ghz," *Electronic Publication: Geneva, Switzerland*, 2019.

[6] P. E. Mogensen and J. Wigard, "Cost action 231: Digital mobile radio towards future generation system, final report." in *Section 5.2: On antenna and frequency diversity in GSM. Section 5.3: Capacity study of frequency hopping GSM network*, 1999.

[7] J. M. Keenan and A. J. Motley, "Radio coverage in buildings," *British telecom technology Journal*, vol. 8, no. 1, pp. 19–24, 1990.

[8] R. Wahl, G. Wölfle, P. Wertz, P. Wildbolz, and F. Landstorfer, "Dominant path prediction model for urban scenarios," in *14th IST Mobile and Wireless Communications Summit*, 2005.

[9] K. Rizk, J.-F. Wagen, and F. Gardiol, "Two-dimensional ray-tracing modeling for propagation prediction in microcellular environments," *IEEE Transactions on Vehicular Technology*, vol. 46, no. 2, pp. 508–518, 1997.

[10] S. Bakirtzis, J. Chen, K. Qiu, J. Zhang, and I. Wassell, "Em deepray: An expedient, generalizable and realistic data-driven indoor propagation model," *IEEE Transactions on Antennas and Propagation*, 2022.

[11] Y. Teganya and D. Romero, "Deep completion autoencoders for radio map estimation," *IEEE Transactions on Wireless Communications*, 2021.

[12] R. Hashimoto and K. Suto, "Sicnn: Spatial interpolation with convolutional neural networks for radio environment mapping," in *2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIC)*. IEEE, 2020, pp. 167–170.

[13] R. Levie, Ç. Yapar, G. Kutyniok, and G. Caire, "Radiounet: Fast radio map estimation with convolutional neural networks," *IEEE Transactions on Wireless Communications*, vol. 20, no. 6, pp. 4001–4015, 2021.

[14] O. Ozyegen, S. Mohammadjafari, M. Cevik, J. Ethier, A. Basar *et al.*, "An empirical study on using cnns for fast radio signal prediction," *SN Computer Science*, vol. 3, no. 2, pp. 1–17, 2022.

[15] Y. Tian, S. Yuan, W. Chen, and N. Liu, "Radionet: Transformer based radio map prediction model for dense urban environments," *arXiv preprint arXiv:2105.07158*, 2021.

[16] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly *et al.*, "An image is worth 16x16 words: Transformers for image recognition at scale," *arXiv preprint arXiv:2010.11929*, 2020.

[17] C.-H. Liu, H. Chang, and T. Park, "Da-cgan: A framework for indoor radio design using a dimension-aware conditional generative adversarial network," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 2020, pp. 498–499.

[18] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1125–1134.

[19] W.-Y. Kim, S.-H. Tae, and D.-H. Seo, "Access-point centered window-based radio-map generation network," *Sensors*, vol. 21, no. 18, p. 6107, 2021.

[20] S. K. Vankayala, S. Kumar, I. Roy, D. Thirumulanathan, S. Yoon, and I. S. Kanakaraj, "Radio map estimation using a generative adversarial network and related business aspects," in *2021 24th International Symposium on Wireless Personal Multimedia Communications (WPMC)*. IEEE, 2021, pp. 1–6.

[21] S. Zhang, A. Wijesinghe, and Z. Ding, "Rme-gan: A learning framework for radio map estimation based on conditional generative adversarial network," *arXiv preprint arXiv:2212.12817*, 2022.

[22] K. Qiu, S. Bakirtzis, H. Song, J. Zhang, and I. Wassell, "Pseudo ray-tracing: Deep leaning assisted outdoor mm-wave path loss prediction," *IEEE Wireless Communications Letters*, vol. 11, no. 8, pp. 1699–1702, 2022.

[23] Z. Li, J. Cao, H. Wang, and M. Zhao, "Sparsely self-supervised generative adversarial nets for radio frequency estimation," *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 11, pp. 2428–2442, 2019.

[24] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," *Advances in neural information processing systems*, vol. 30, 2017.

[25] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.

[26] W. Yang, X. Zhang, Y. Tian, W. Wang, J.-H. Xue, and Q. Liao, "Deep learning for single image super-resolution: A brief review," *IEEE Transactions on Multimedia*, vol. 21, no. 12, pp. 3106–3121, 2019.

[27] T.-C. Wang, M.-Y. Liu, J.-Y. Zhu, A. Tao, J. Kautz, and B. Catanzaro, "High-resolution image synthesis and semantic manipulation with conditional gans," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 8798–8807.

[28] Z. Xu, C. Yang, S. Alheejawi, N. Jha, S. Mehadi, and M. Mandal, "Floor plan semantic segmentation using deep learning with boundary attention aggregated mechanism," in *2021 4th International Conference on Artificial Intelligence and Pattern Recognition*, 2021, pp. 346–353.

[29] D. Pathak, P. Krahenbuhl, J. Donahue, T. Darrell, and A. A. Efros, "Context encoders: Feature learning by inpainting," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2536–2544.

[30] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in *The Thrity-Seventh Asilomar Conference on Signals, Systems & Computers, 2003*, vol. 2. Ieee, 2003, pp. 1398–1402.

[31] A. N. Nor Hisham, Y. H. Ng, C. K. Tan, and D. Chieng, "Hybrid wi-fi and ble fingerprinting dataset for multi-floor indoor environments with different layouts," *Data*, vol. 7, no. 11, p. 156, 2022.

[32] G. Lui, T. Gallagher, B. Li, A. G. Dempster, and C. Rizos, "Differences in rssi readings made by different wi-fi chipsets: A limitation of wlan localization," in *2011 International conference on localization and GNSS (ICL-GNSS)*. IEEE, 2011, pp. 53–57.

[33] R. Hoppe, G. Wölfle, and U. Jakobus, "Wave propagation and radio network planning software winprop added to the electromagnetic solver package feko," in *2017 International Applied Computational Electromagnetics Society Symposium-Italy (ACES)*. IEEE, 2017, pp. 1–2.

[34] V. Jodalen, "An evaluation of the radio propagation models available in winprop from awe communicatons," 2015.

[35] A. Kalervo, J. Ylioinas, M. Häikiö, A. Karhu, and J. Kannala, "Cubicasa5k: A dataset and an improved multi-task model for floorplan image analysis," in *Scandinavian Conference on Image Analysis*. Springer, 2019, pp. 28–40.

[36] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[37] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin *et al.*, "Tensorflow: Large-scale machine learning on heterogeneous systems," 2015.