Prediction of Optimal Locations for 5G Base Stations in Urban Environments using Neural Networks and Satellite Image Analysis

Iván García-Aguilar[1,2], Jesús Galeano-Brajones[3], Francisco Luna-Valero[1], Javier Carmona-Murillo[3], and Jose David Fernández-Rodríguez[1,2] Rafael M. Luque-Baena[1,2]

[1] ITIS Software. University of Málaga. C/ Arquitecto Francisco Peñalosa, 18. 29010 Málaga, Spain.

[2] Biomedical Research Institute of Málaga (IBIMA), C/ Doctor Miguel Díaz Recio, 28, Málaga, 29010, Spain.

[3] University of Extremadura, Department of Computing and Telematics Engineering, C/ Sta. Teresa Jornet, 38, Mérida, 06800, Spain

Abstract. Deploying 5G networks in urban areas is crucial for meeting the increasing demand for high-speed, low-latency wireless communications. However, the complex topography and diverse building structures in urban environments have challenges in identifying suitable locations for base stations. This research explores leveraging deep learning neural networks to analyze satellite imagery, creating a predictive tool for identifying potential rooftop locations. Integrating these predictions into a user-friendly desktop application simplifies the site selection process, reducing the need for costly and labor-intensive site visits in 5G network deployment. This approach democratizes the deployment process, making it accessible to a broader audience. The combination of advanced technology and satellite imagery offers a promising solution to efficiently deploy 5G base stations in urban landscapes, contributing to the widespread adoption of this technology in densely populated areas and advancing 5G connectivity globally.

Keywords: Convolutional Neural Network (CNN) · Deep Learning (DL) · 5G Deployment.

1 Introduction

The demand for high-speed communications has led to the implementation of 5G networks, offering fast data speeds, low latency, and reliable communication. Despite its potential, the success of the 5G revolution relies heavily on the effective deployment of 5G base stations in densely populated urban environments. Growing urban areas and the increasing demand for wireless connectivity necessitate addressing challenges in identifying suitable locations, particularly on building rooftops. Urban environments, with their intricate topography and diverse building structures, present significant obstacles. Identifying suitable locations for 5G base stations involves geospatial assessments and meticulous planning, which consume considerable resources and time. Inefficiencies in this process can impede 5G expansion in urban areas and substantially raise deployment costs.

Optimizing base station deployment is crucial for efficiency and cost reduction. Articles like [1] focus on the transition to 5G wireless networks, while others such as [2,3] emphasize energy consumption and deployment optimization. In [4], energy efficiency in 5G networks, especially in rural areas, is highlighted. Qi Wang et al. [5] address challenges in urban 5G network deployment, emphasizing issues with millimeter wave signals. The main challenge is deploying an ultra-high density of base stations (BSs) for satisfactory communication coverage. [6] focuses on implementing 5G base stations for optimal signal coverage and cost management. Deep learning, particularly convolutional neural networks, has revolutionized image analysis and object segmentation approaches [7,8,9]. Notably, models like U-Net [10], FPN [11], PSPNet [12], and Mask R-CNN [13] focus on segmentation. These models, exemplified in [14], directly identify specific elements in satellite images, representing a significant advancement in computer vision. They are essential for simplifying and optimizing the implementation of 5G networks in urban environments.

In this work, a solution is presented that leverages the capabilities of Convolutional Neural Networks (CNN) in deep learning to address the challenge of identifying suitable locations for 5G base stations in urban environments. This approach aims to simplify selecting installation locations, thereby reducing the dependence on costs related to site visits or time-consuming tasks, ultimately democratizing the implementation process. This not only accelerates the deployment of 5G base stations but also makes it accessible to a broader audience, paving the way for widespread 5G connectivity in densely populated urban areas.

The remainder of the article is organized as follows. In Section 2, the proposed methodology is presented. Subsequently, in Section 3, the conducted experiments are detailed, including the selected models, the applied dataset, and the evaluated metrics. Finally, in Section 4, the conclusions are elaborated.

2 Methodology

The suggested method for predicting optimal 5B base station locations in urban settings comprises two subsystems. Initially, a satellite image is segmented with a preprocessing phase to align it with the chosen model (see Subsection 2.1). The resulting segmented image is then input into the base station deployment simulator, guiding the deployment strategy (outlined in Subsection 2.2).

2.1 Segmentation of satellite images

This subsection describes the procedure to obtain the segmentation of a satellite image using convolutional neural networks. Our proposal begins with a dataset composed of several annotated images, which were processed individually to train the model in the next step.

$$\mathbf{D} = \{(\mathbf{F}_l) \mid l \in \{1, ..., N\}\} \quad (1)$$

D represents the labeled dataset, while $\mathbf{F}_l$ denotes each individual frame. N corresponds to the total number of images that compose the entire dataset. The starting point is to train a deep convolutional neural network, designated as $\mathcal{G}$ for object segmentation. A strategy based on K-fold is considered, where we have K subsets of data divided into training, validation, and testing.

In the second step, image transformations are employed to enhance training through the utilization of the Data Augmentation technique. To address this challenge, the Albumentations library[4] has been employed, allowing a series of transformations to be applied to the original images. These transformations have been applied to the dataset to generate an expanded set of transformed data without incurring model overfitting. Formally, the process can be expressed as follows:

$$D' = \{X'_1, X'_2, \dots, X'_m\}, where X'_j = T(X_i) \tag{2}$$

D' represents the extended dataset of transformed images, and X'_j are the transformed images obtained by applying the transformation T to the original image X_i. This results in a larger dataset without increasing the labeling workload. Additionally, normalization of training images is conducted to enhance model training. The objective of this normalization is to eliminate differences in magnitude among the various image features. This is achieved through the standardization technique, expressed as:

$$X_{normalized} = \frac{X - \mu}{\sigma} \tag{3}$$

where X represents an image in the training set, μ is the mean of the image features and σ is the standard deviation of the image features. Z-score normalization is applied to each pixel in the image, making the mean zero and the standard deviation one. This ensures that all image features have a consistent scale, reducing noise during model training and improving result accuracy.

At this point, the third step is based on the training of the G model, with the objective of performing the segmentation of satellite images for the identification of buildings. For this purpose, a series of parameters related to this task were determined previously. The loss function is defined as follows:

$$L(\theta_M) = \frac{1}{N_{training}} \sum_{i=1}^{N_{training}} IoU(\mathbf{Y}_i, f_M(\mathbf{X}_i; \theta_G)) \tag{4}$$

where $N_{training}$ is the number of examples that compose the training subset. Y_i are the Ground truth annotations for the example i, $f_M(\mathbf{X}_i; \theta_G)$ the prediction determined by the model G for the example i. θ_G represents the parameters of the model G. After defining the loss function, the model is optimized. In this case, the ADAM optimizer is selected. The formula for the optimization of the model parameters is as follows:

$$\theta^*_M = \arg\min_{\theta_M} L(\theta_M) \qquad G' = Fine\_Tuned\{G\} \tag{5}$$

[4] `https://albumentations.ai`

The set of parameters is searched for θ_M which minimizes the loss function L. Once this is defined, the training of the segmentation model G is performed. To enhance the result of processing a non-square image that has been resized to 288x288 pixels, a sliding window technique is employed to divide the image into small regions and make predictions in each of them. The requirement to transform the input image was due to the restrictions imposed by the configuration and the data set used during the model training. In this context, the model has been designed to process images with only specific dimensions, 288x288 pixels. This limitation imposed by the model architecture implies that any input image not complying with this dimension must transform to fit the model requirements. Subsequently, the information from all predictions is combined to obtain the final segmentation results.

Firstly, the image is divided into small windows of size 288x288 pixels with a displacement S of 20 pixels. This can be mathematically represented as:

$$W_{i,j} = F(I_{val}, i, j) \qquad S_{i,j} = G'(W_{i,j}) \tag{6}$$

where $W_{i,j}$ is a sub-image extracted from the validation image I_{val} at position (i, j), and F is the function that extracts a region from the image. Next, for each sub-image $W_{i,j}$, a prediction is made using the trained model G', resulting in a segmentation for that region, $S_{i,j}$. Subsequently, the information from all segmentations $S_{i,j}$ are combined to obtain the final segmentation S_{final}. A unified mask is generated by summing the pixel values of masks associated with detections. Pixels in the consolidated mask are marked as true when the corresponding sums exceed zero.

2.2 Base station deployment

In this phase, an Ultra-Dense Networks (UDNs) simulator based on the model is employed [15]. The objective is to simulate the deployment of base stations in an ultra-dense network. The deployment process is carried out randomly using Poisson Point Processes, and it takes into account social attractors in the network [16]. Poisson Point Processes are a mathematical tool for modeling the distribution of random events in space. In the context of base station deployment, they are used to model the random locations of base stations within a geographic region. This can be expressed through the intensity of the Poisson point process formula:

$$\lambda(x, y) = \frac{Expected\ number\ of\ base\ stations}{dA} \tag{7}$$

where $\lambda(x, y)$ is the intensity of the point process at coordinates (x, y) and dA represents an elemental area in the geographic plane.

In the base station deployment simulation, social attractors are also taken into account. These social attractors may represent areas with a higher density of users or activities, influencing the placement of base stations. The influence of social attractors on deployment intensity can be expressed as:

$$\lambda_s(x, y) = \lambda(x, y) \times \alpha_s \tag{8}$$

where $\lambda_s(x, y)$ is the intensity of the point process adjusted for social attractors at coordinates (x, y) and α_s is a factor representing the influence of social attractors on base station density. This approach allows for modeling different severity levels in the heterogeneity of the traffic demands within the network.

3 Experiments

The objective of this study is to evaluate the effectiveness of the proposed methodology based on the segmentation done by the convolutional neural networks, and the solution obtained. The study involves a comparison of several models applying one specific dataset. Below is a detailed description of each of the architectures used to evaluate their performance, as well as information on the dataset, the metrics used and the results obtained.

3.1 Convolutional Neural Networks:

The presented approach is adaptable to any neural DCNN-based object segmentation model. For conducting the experiments, three architectures were selected:

- U-Net [10]: Widely used CNN for image segmentation, known for its U-shaped structure that captures intricate details and high-level patterns.
- FPN (Feature Pyramid Network) [11]: Applied in object detection and image segmentation, FPN enhances detection across scales by combining features from different pyramid levels, facilitating precise segmentation of objects of varying sizes.
- PSPNet (Pyramid Scene Parsing Network) [12]: An image scene segmentation model utilizing a feature pyramid to analyze images at various scales, contributing to comprehensive scene understanding and context extraction.

The pre-trained models were selected from the *SMP* (Segmentation Models Pytorch) library , and were previously trained on the COCO dataset (Common Objects in Context) [17]. The COCO dataset offers diverse, challenging images with various object scales and compositions, making it valuable for training real-world object detection models widely used in deep learning applications.

3.2 Dataset

The dataset used consists of images which are labelled for the segmentation task. The choice of this dataset is based on its size, which is sufficiently large to achieve optimal results in the project. This dataset was obtained from the crowdAI platform. The dataset is linked to the following challenge [5]. In this challenge, a dataset of RGB satellite images, along with annotations indicating the location of buildings in each image, is provided. The goal of this challenge is to train a model capable of detecting all buildings in a new image. To work with this dataset, this is divided into three essential parts: training, testing, and validation. Since this project employs the K-Fold technique, the data is distributed as follows in each experiment: 75% for training and 25% for testing. Additionally, 20% of the training dataset is set aside for validation purposes.

Fig. 1. Image extracted from the dataset along with its segmentation.

In total, the training dataset for the neural network consists of 280,741 satellite images, each with dimensions of 300x300 pixels in RGB format. These training images are provided with annotations in MS-COCO format, presented in a JSON file. Figure 1 shows the annotations based on the segmentation provided by each of the images.

3.3 Evaluation

The principal metric for evaluation is Intersection over Union (IoU), a widely accepted measure for assessing segmentation model accuracy. IoU calculation involves matching model detections with Ground Truth (GT) annotations using established criteria, comparing overlap between detected and actual building areas. This process identifies true positives, false positives, and false negatives. An accuracy-recall curve is generated based on these values, illustrating how accuracy varies with adjustments to confidence thresholds in model detections. Lastly, the average IoU for the building class is calculated, representing an overall measure of the model's accuracy in segmenting buildings across various confidence thresholds.

3.4 Results

The results section is split into two main parts. The first part covers the quantitative and qualitative outcomes of training the segmentation model, including objective metrics and visual examples demonstrating its effectiveness. The second part examines how precise building segmentation has impacted the planning and success of antenna deployment in urban environments, providing a qualitative analysis to assess its real-world applicability.

Training Segmentation Models:

The U-Net model outperforms the FPN and PSPNet models across all evaluation stages, as shown in Table 1. During training, U-Net achieved a low loss

[4] `https://github.com/qubvel/segmentation_models.pytorch`
[5] `https://www.crowdai.org/challenges/mapping-challenge`

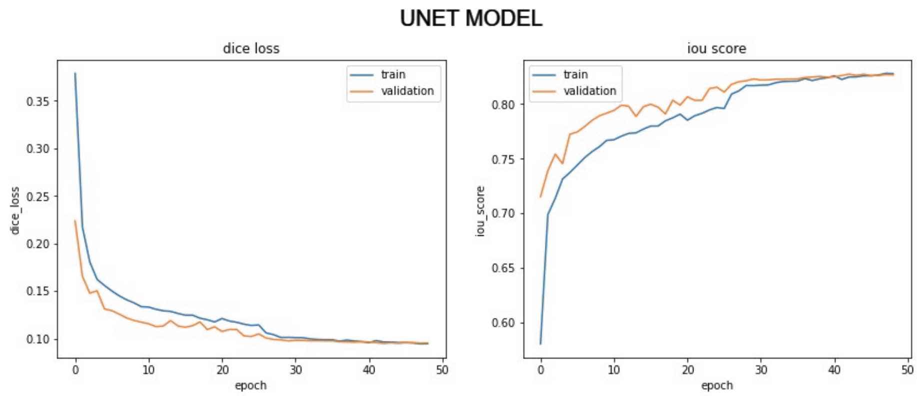

Fig. 2. Evolution of loss function and IoU values for each of the evaluated stages and models.

function of 0.0935, indicating effective adjustment to the data, and an IoU value of 0.8307, indicating strong segmentation capability. In validation, U-Net maintained high performance with a loss function of 0.0983 and an IoU value of 0.8228. In the test stage, U-Net displayed a loss function of 0.1214 and an IoU value of 0.8002, confirming its accuracy and consistency. Conversely, FPN and PSPNet models performed similarly, with comparable loss function and IoU values. However, in the test stage, both experienced a decline in performance. FPN recorded a loss function of 0.1214 and an IoU value of 0.8002, while PSPNet had a slightly higher loss function of 0.1383 and an IoU value of 0.7758, indicating slightly inferior performance compared to U-Net. Figure 2 illustrates the evolution of loss function and IoU values for each stage, highlighting U-Net's suitability for image segmentation in this context.

Model	Stage	Loss Function	IoU
U-Net	Training	0.0935	0.8307
	Validation	0.0983	0.8228
	Test	0.1214	0.8002
FPN	Training	0.09351	0.8306
	Validation	0.0983	0.8228
	Test	0.1214	0.8002
PSP	Training	0.1205	0.7880
	Validation	0.1156	0.7950
	Test	0.1383	0.7758

Table 1. Results obtained based on the training of the selected models in the different stages.

Since the U-Net model demonstrated superior performance in the various stages of training, validation, and testing compared to the other models, cross-validation training was conducted to assess its robustness and generalization further. These results are presented in Table 2. In the case of cross-validation

Stage	Loss Function	IoU
Training	0.0940	0.8287
Validation	0.0818	0.8490
Test	0.1103	0.8171

Table 2. Results of U-Net model training and evaluation with cross validation.

training, a slightly higher loss function is obtained, in this case being 0.0940 compared to the non-cross-validated result of 0.0935. The IoU value decreased slightly to 0.8287 from 0.8307, suggesting that the cross-validated model may incur a minor penalty in terms of accuracy. In the validation stage, the cross-validated model demonstrated improved performance compared to the non-cross-validated model. The loss function reduced from 0.0983 to 0.0818, and the IoU value increased from 0.8228 to 0.8490. This indicates that the model can generalize and achieve a more robust accuracy. In the test stage, the U-Net model with cross-validation maintains a strong performance with a loss function of 0.1103 and an IoU value of 0.8171. While a slight penalty is observed in the training stage, this is more than compensated for in terms of accuracy on unseen data, thus determining that the U-Net model with cross-validation is the preferred choice for the image segmentation task in this context.

Stations Deployment:

The primary motivation lies in the need to develop a specialized neural network for detecting terraces and rooftops to simulate the deployment of 5G base stations accurately. The UDN simulator used is based on an ultra-dense network model, where the deployment of base stations is randomly carried out through Poisson Point Processes (PPP). This model considers social attractors in the network, such as offices and shopping centers, to model a realistic traffic demand distribution. The rooftop and terrace detection neural network plays a crucial role by providing detailed information about urban topology and the presence of relevant structures, such as buildings and rooftops. This information is integrated with the UDN simulator, allowing for more realistic planning of 5G base station deployment based on the location of these structures. Precisely detecting rooftops and terraces with the neural network ensures that 5G base stations are placed in viable locations so as to simulate real-world deployments, that is, avoding placing a base station in the middle of a street. The collaboration between the detection technology and the UDN simulator enables the evaluation and optimization of next-generation network deployments in urban environments, considering social attractor factors and specific geographic features. Figure 3, presented below, visually exemplifies the autonomous deployment conducted as a result of this work. A red mask identifies the segmentation of the buildings. Each of the dots in the image indicates the position where the base stations will be installed for deployment.

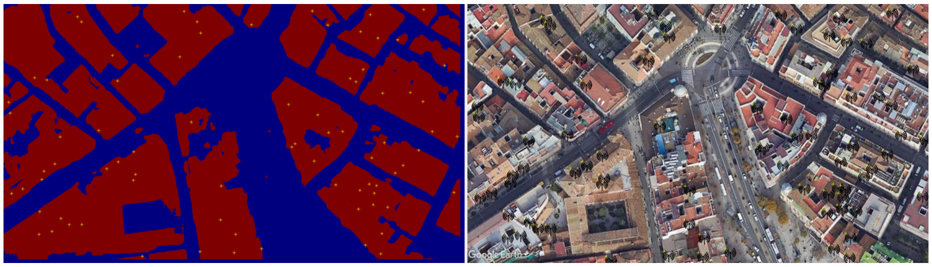

Fig. 3. Base Stations Previsualization.

4 Conclusions

According to the work presented, a significant improvement has been achieved in the planning and deployment of 5G base stations in urban and suburban environments through the implementation of a specialized neural network for the detection of rooftops and terraces. Based on the various segmentation models evaluated, U-Net has demonstrated its ability to achieve precise segmentation of structures, with an Intersection over Union (IoU) value consistently high during the training stages (average of 0.8268), validation (average of 0.8443), and testing (average of 0.8058). Furthermore, the application of cross-validation has led to an improvement in performance during the validation stage, with an average loss function of 0.0818 and an average IoU value of 0.8490, confirming the model's generalization capability. Integrating the neural network with the Unified Data Networks (UDN) simulator has resulted in an autonomous deployment solution, providing an efficient and precise tool for planning and deploying base stations in urban areas. As for future lines of research, expansion of the neural network for the detection of other urban elements or research into the optimization of real-time data-based 5G base station deployment algorithms is identified as areas for continued improvement and optimization of the proposed solution.

Acknowledgements

This work is partially supported by the Ministry of Science and Innovation of Spain under grants PID2022-136764OA-I00, TED2021-131699B-I00 (MCIN / AEI / 10.13039 / 501100011033, FEDER) and PID2020-112545RB-C54, by the University of Málaga (Spain) under grants B1-2021_20, B4-2023_13, B1-2022_14 and by the Fundación Unicaja under project PUNI-003_2023.

References

1. Xiaoyuan Cheng, Yukun Hu, and Liz Varga. 5g network deployment and the associated energy consumption in the uk: A complex systems' exploration. *Technological Forecasting and Social Change*, 180:121672, 2022.

2. Qingwen Zeng. Optimization of millimeter-wave base station deployment in 5g networks. In *2022 Thirteenth International Conference on Ubiquitous and Future Networks (ICUFN)*, pages 117–121, 2022.
3. Zahra Mohammadnejad, Hamza Mohammed Ridha Al-Khafaji, Amin Salih Mohammed, and Sura Rahim Alatba. Energy optimization for optimal location in 5g networks using improved barnacles mating optimizer. *Physical Communication*, 59:102068, 2023.
4. Alexander Karlsson, Osama Al-Saadeh, Anton Gusarov, Renuka Venkata Ramani Challa, Sibel Tombaz, and Ki Won Sung. Energy-efficient 5g deployment in rural areas. In *2016 IEEE 12th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, pages 1–7, 2016.
5. Qi Wang, Xiang Zhao, Zhantian Lv, Xiaoya Ma, Ruitian Zhang, and Yifan Lin. Optimizing the ultra-dense 5g base stations in urban outdoor areas: Coupling gis and heuristic optimization. *Sustainable Cities and Society*, 63:102445, 2020.
6. Chia-Hung Wang, Chia-Jung Lee, and Xiaojing Wu. A coverage-based location approach and performance evaluation for the deployment of 5g base stations. *IEEE Access*, 8:123320–123333, 2020.
7. Iván García-Aguilar, Jorge García-González, Rafael Marcos Luque-Baena, and Ezequiel López-Rubio. Automated labeling of training data for improved object detection in traffic videos by fine-tuned deep convolutional neural networks. *Pattern Recognition Letters*, 167:45–52, 2023.
8. Ravpreet Kaur and Sarbjeet Singh. A comprehensive review of object detection with deep learning. *Digital Signal Processing*, 132:103812, 2023.
9. Iván García-Aguilar, Jorge García-González, Rafael M. Luque-Baena, Ezequiel López-Rubio, and Enrique Domínguez. Optimized instance segmentation by super-resolution and maximal clique generation. *Integrated Computer-Aided Engineering*, 30(3):243–256, May 2023.
10. Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation, 2015.
11. Tsung-Yi Lin, Piotr Dollár, Ross Girshick, Kaiming He, Bharath Hariharan, and Serge Belongie. Feature pyramid networks for object detection, 2017.
12. Hengshuang Zhao, Jianping Shi, Xiaojuan Qi, Xiaogang Wang, and Jiaya Jia. Pyramid scene parsing network, 2017.
13. Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross Girshick. Mask r-cnn. In *2017 IEEE International Conference on Computer Vision (ICCV)*, pages 2980–2988, 2017.
14. Loretta Ichim and Dan Popescu. Road detection and segmentation from aerial images using a cnn based system. In *2018 41st International Conference on Telecommunications and Signal Processing (TSP)*, pages 1–5, 2018.
15. Francisco Luna, Pablo H. Zapata-Cano, Juan C. González-Macías, and Juan F. Valenzuela-Valdés. Approaching the cell switch-off problem in 5g ultra-dense networks with dynamic multi-objective optimization. *Future Generation Computer Systems*, 110:876–891, 2020.
16. Meisam Mirahsan, Rainer Schoenen, and Halim Yanikomeroglu. Hethetnets: Heterogeneous traffic distribution in heterogeneous wireless cellular networks. *IEEE Journal on Selected Areas in Communications*, 33(10):2252–2265, 2015.
17. Tsung-Yi Lin, Michael Maire, Serge Belongie, Lubomir Bourdev, Ross Girshick, James Hays, Pietro Perona, Deva Ramanan, C. Lawrence Zitnick, and Piotr Dollár. Microsoft coco: Common objects in context, 2015.