

# Deep Learning for On-Street Parking Violation Prediction

Nikolaos Karantaglis, Nikolaos Passalis, and Anastasios Tefas

Dept. of Informatics

Aristotle University of Thessaloniki

Thessaloniki, Greece

Email: {nikgeokar, passalis, tefas}@csd.auth.gr

**Abstract**—Illegal parking along with the lack of available parking spaces are among the biggest issues faced in many large cities. These issues can have a significant impact on the quality of life of citizens. On-street parking systems have been designed to this end aiming at ensuring that parking spaces will be available for the local population, while also providing easy access to parking for people visiting the city center. However, these systems are often affected by illegal parking, providing incorrect information regarding the availability of parking spaces. Even though this can be mitigated using sensors for detecting the presence of cars in various parking sectors, the cost of these implementations is usually prohibiting large. In this paper, we investigate an indirect way of predicting parking violations at a fine-grained level, equipping such parking systems with a valuable tool for providing more accurate information to citizens. To this end, we employed a Deep Learning (DL)-based model to predict fine-grained parking violation rates for on-street parking systems. Moreover, we developed a data augmentation and smoothing technique for further improving the accuracy of DL models under the presence of missing and noisy data. We demonstrate, using experiments on real data collected in Thessaloniki, Greece, that the developed system can indeed provide accurate parking violation predictions.

## I. INTRODUCTION

Searching for a parking space in a large city is often very challenging and frustrating for drivers, especially in crowded areas like the city’s centre or around of points of interest. To address this issue, many cities have implemented intelligent on-street parking systems [1]. In such systems, available parking spaces are often divided into sectors, with each sector being composed of several parking slots. Drivers usually manually update the system by stating the slot/sector where they have parked, when paying the parking fee. In this way, the parking system is updated and can provide useful information to drivers in order to guide them to sectors that have free slots available. However, when drivers park without providing the necessary information to the system, i.e., usually to avoid paying the parking fee, the system no longer has up-to-date information regarding the availability of parking slots and cannot provide useful information to other drivers. During peak hours, such phenomena can often lead to a significant degradation in the system’s performance, which lowers the trust of drivers to such systems and discourages their use.

The aforementioned issue can be addressed by installing street occupation sensors that can detect the presence of cars in various parking sectors, e.g., either street sensors buried under the road or vision-based systems [2]. However, the

cost of installing and maintaining such complicated systems in large cities can outweigh the benefits of their installation. Furthermore, such systems tend to work better in more constrained settings, e.g., in closed spaces, instead of public parking systems that can span an entire city. In this work, we investigate whether it is possible to acquire such information using *indirect* measurements, such as weather information, time and date, as well as historical data. More specifically, we aim to predict illegal parking rates in different parking sectors around a city. This information, along with the existing information held by such systems (i.e., cars that are legally parked) would allow for upgrading their operation with minimal cost. Indeed, such an indirect prediction system would allow for providing corrections to the number of available parking slots, taking into account slots that might have been taken up by illegal parking, redirecting users towards sectors that are more probable to have free parking slots.

To this end, in this work we aim on developing a machine learning-based approach for predicting illegal parking in various sectors of an on-street parking system. However, developing such a system poses significant challenges regarding several aspects of its design, ranging from how the input data are gathered and encoded to how the ground truth parking violation data are collected. Despite the importance of this problem, there is a relatively small number of works attempting to address this issue. For example, in [3] parking violations are predicted using random forests and existing open data, while [4] tackles the problem of predicting double parking events. However, the latter only focuses on double parking events, while the former, despite solving a similar problem, provides coarse information. In contrast, this paper focuses on solving a fine-grained variant of the on-street parking violation prediction providing information for hundreds of sectors. Other approaches, such as [5], [6], attempt to directly predict the number of available parking spaces instead of focusing only on the illegal parking rate, potentially solving a significantly harder problem. It is worth noting that most of these approaches are often tailored to the specifics of each parking system, available annotation type, and city structure, so it is not always possible to directly apply them in other scenarios.

The main aim of this work is to develop and evaluate a Deep Learning (DL)-based on-street parking violation prediction system that works integrated with on-street parking systems. DL has been shown to outperform other approaches for a wide variety of tasks [7], yet its application to novel

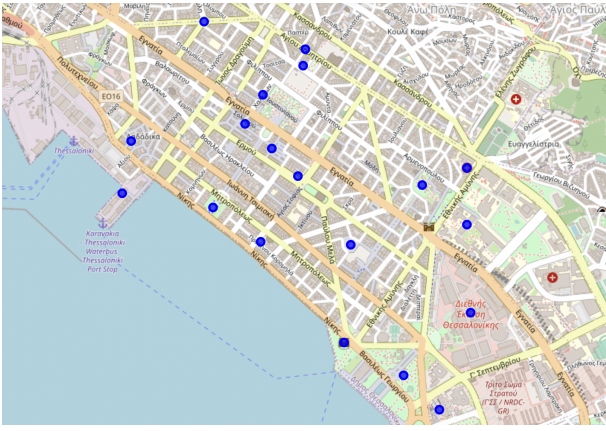


Fig. 1. A total of 19 PoIs were defined and used for representing various sectors of the used on-street parking system.

domains remains challenging. Among the main challenges that we face is the existence of sparse annotations, i.e., information regarding illegal parking is only available when the police scan for such cases. This poses significant challenges since annotations can be especially noisy, i.e., having information for a sector but not for neighboring ones. Directly using such information in DL-based systems is challenging, since it is known that DL models can be affected by such types of annotations, which often leads to applying label smoothing and distillation methods to mitigate these effects [8], [9], [10]. To address these issues, we developed a novel data smoothing and augmentation approach that assigns existing ground truth annotation to spatial sectors and time slots, leading to significant performance improvements. To the best of our knowledge, this is the first time that such an approach is applied in the context of an on-street parking violation prediction system. Other challenges include imprinting input data, such as the geographical location of each sector, time and other information, such as pandemic measures (i.e., due to COVID-19) in an efficient manner. Finally, we extensively evaluate the proposed method using data from the THESi system implemented in Thessaloniki, Greece.

The rest of the paper is structured as follows. Section II introduces the proposed method, detailing all important aspects of data pre-processing and model construction. Then, the experimental evaluation is provided in Section III. Finally, conclusions are drawn in Section IV.

## II. PROPOSED METHOD

As described before, parking slots in public on-street parking systems are divided into sectors. These sectors consist the fundamental block for which we predict the rate of parking violations. Given that the location and capacity of each sector are known, we can directly use the parking violation predictions for each sector in order to guide drivers towards sectors where parking slots will be most probably available. The first challenge we faced was to represent the geographical place of a sector in a way that would allow DL models to generalize. To this end, we avoided directly using a global coordinate system and we opted for indirectly encoding sector information using distances from several Points of Interest (PoI). A PoI is often visited by citizens on a large scale and, as a result, it is crowded

most hours of the day. A PoI could be a sight, a central park, a museum, the city hall, and others. Given that traffic and demand for parking slots are higher across different PoIs, this way of encoding the information can allow a DL model to associate PoIs with expected parking violation rates, as well as encode sectors using distances from these points. To this end, we measured the distance between the center of each sector and the selected PoIs and used this vector to represent each sector. The way this information was encoded is shown in Fig. 1, where blue circles represent the PoIs and the red circle shows an example of a sector. A total of 19 points of interest were employed, leading to a 19-dimensional representation for each sector. The capacity of each sector was also included as a feature.

Furthermore, we also included information regarding the current weekday (Monday to Sunday), day (1 to 31), and month (January to December) in order to allow the DL model to encode historical information regarding the expected parking violations. To capture the periodic nature of these features and prevent discontinuities in the features we opted for using sine-based encoding for all of these features. For example, the weekday, denoted by  $x_w$  was encoded as:

$$x_w = \sin\left(2\pi \frac{w}{7}\right), \quad (1)$$

where  $w$  denotes the current day using sequential integer encoding, i.e., 0 for Monday, 1 for Tuesday, etc. Similarly, we extracted a feature for the current day  $x_d$  as:

$$x_d = \sin\left(2\pi \frac{d-1}{N_d}\right), \quad (2)$$

where  $d$  is the current day (from 1 to 31) and  $N_d$  is the number of days for the current month. Moreover, the month feature  $x_m$  was calculated as:

$$x_m = \sin\left(2\pi \frac{m-1}{12}\right), \quad (3)$$

where  $m$  is the current month (ranging from 1 to 12). As we explain later in this section, the prediction horizon is one-hour. Therefore, each day was also divided into time slots (each with one-hour duration) and this information was also included the extracted features. Note that time slots do not span the whole day, since parking control is enforced from 7:00 until 19:00. Therefore, we did not use sine encoding for these features and we directly imprinted the temporal location of each time slot as its distance from the time when parking control is enforced.

Apart from the location of each sector and time features, weather conditions also often affect traffic and demand for parking spaces. Therefore, we used the hourly values of temperature and humidity. Temperature is represented in Celsius degrees, while humidity is provided as a percentage of the maximum possible humidity. Given that weather features usually do not immediately affect the behavior of drivers, i.e., if the temperature rises, traffic is not immediately affected, we opted for encoding weather features as the average of windows. To this end, the current temperature feature  $x_T$  for a time step  $t$  was calculated as:

$$x_T = \frac{1}{W} \sum_{i=1}^W T_{t-i}, \quad (4)$$

where  $T_t$  represents the temperature at the  $t$ -th time step and  $W$  denotes the window size used for the averaging. Note that we omitted the time step index  $t$  from the feature notation to

reduce clutter. In this work, we used the current and 5 previous temperature measurements, i.e.,  $W = 6$ . Humidity features  $x_h$  were similarly calculated.

Finally, two additional features we employed to enrich the information that is available to the DL model. The first one is a binary variable that is used to indicate whether the current day is a public holiday. The other feature was used to indicate whether the time slot belongs to a period of a pandemic or not. We included this feature since we observed that during the recent COVID-19 pandemic it was observed that traffic was affected. Even though we used this feature to only mark time slots that belong to COVID-19 pandemic, we suspect that this feature might be also useful during subsequent flu seasons due to increased public awareness.

All features were normalized using standarization, i.e., subtracting their mean and dividing by the standard deviation. The training set statistics were used for the normalization process. Also, the proposed prediction system operates on one-hour time slots aiming to predict the parking violation rate for each time slot and sector. The target parking violation rate is calculated as:

$$p = \frac{N_s}{N_c}, \quad (5)$$

where  $p$  denotes the target violation rate to be predicted,  $N_s$  is the number of illegally parked cars and  $N_c$  is the capacity of the specific sector at a specific time slot (again we omit the sector and time slot indices to avoid cluttering the used notation). The number of illegally parked cars is calculated from police scan data. During such scans, a police officer scans cars' license plates that are parked in a sector. The scanner then checks if the parking fee has been paid (or if the car can legally park there without paying a fee) for each parked car. Then, this information is used to calculate the number of illegally parked cars within each sector and time slot. Then, each scan can be assigned to the time slot that it is nearest to the center of each time slot.

Even though a large number of scans exist, the ground truth data remain sparse. For example, scans may concern only a specific sector and no information might be available for neighboring sectors at the same time. Furthermore, scans might span more than one time slots, often leaving the first and last time slots with very few scans, providing false information regarding the ground truth parking violation rate. To overcome this problem, we propose to distribute each scan to time slots assuming that the violation rate follows a smooth curve. This can mitigate annotation sparsity and discontinuity issues, improving the performance of DL models, as also shown in Section III.

To this end, we applied a Gaussian-based smoothing scheme. After calculating the violation rate for each scanning session, we distributed the observed violation rate to nearby time slots using the following equation:

$$y = \frac{1}{|\mathcal{S}|} \sum_{s \in \mathcal{S}} \exp\left(\frac{-d_s}{\sigma}\right) p_s, \quad (6)$$

where  $s$  denotes a scanning session,  $\mathcal{S}$  consists of scanning sessions overlapping or being nearby (by a difference of one time slot) to the current time slot,  $d_s$  is the distance in minutes of scanning session  $s$  from the current time slot's center,  $p_s$  is the violation rate calculated for scanning session  $s$  and  $|\mathcal{S}|$

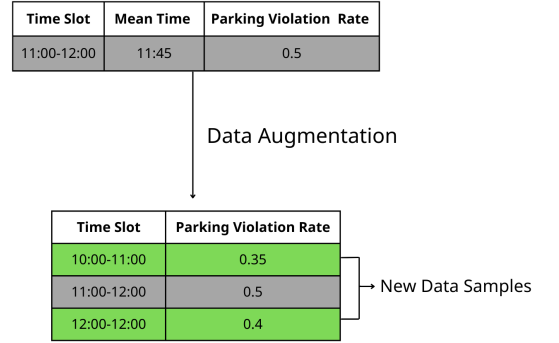


Fig. 2. The proposed data augmentation and smoothing method allows for assigning observed violation rates into nearby time slots, reducing the impact of sparse annotations and missing data scans.

denotes the cardinality of set  $\mathcal{S}$ , i.e., the number of scanning sessions that can be assigned to the current time slot. The parameter  $\sigma$  controls the distribution process, i.e., a larger value distributes the violations to the nearby time slots, while smaller values allow for more steep changes in the parking violation rate between slots. A value of  $d = 210$  minutes was used for the models developed in this paper. For example, if we observe a rate of thirty five percentage (35%) of parking violations in a sector at twelve and fifty (12:50), i.e., in “12:00-13:00” time slot, then we can also assume that this rate should be also partially assigned to the next time slot of “13:00-14:00”, as shown in Fig. 2. As described before, this allows for filling potential gaps in the data, reducing the impact of sparse annotations and missing scans.

#### A. Model architecture

In this we work, we employed a residual DL architecture for predicting the parking violation rate in different sectors. The employed architecture is composed of six hidden layers (each one with 512, 256, 128, 64, 128 and 32 neurons respectively) and one output layer. The ReLU activation function was used for the hidden layer [11], while a residual link was used between the third and fifth hidden layers [12]. Using this link allows for improving the prediction accuracy over architectures that do not use residual connections. Also, further increasing the size of the employed architecture is not expected to lead to further accuracy improvements. The last layer employs a sigmoid activation function, since it is used to predict the parking violation rate, which is always bounded between 0 and 1. To avoid pushing the sigmoid function to its extremes, we re-normalized targets between 0.1 and 0.9. The Adamax optimization algorithm was used for training the model [13]. The mean squared error was used as the loss function. Finally, the learning rate was initialized at 0.001 and an exponential learning rate decay strategy with a decay rate equal to 0.25 was used for performing learning rate scheduling [14].

### III. EXPERIMENTS

For all the conducted experiments, we used data collected from Thessaloniki's on-street public parking system called THESi<sup>1</sup>. THESi consists of about 4700 parking slots in the

<sup>1</sup><https://thesi.gr>

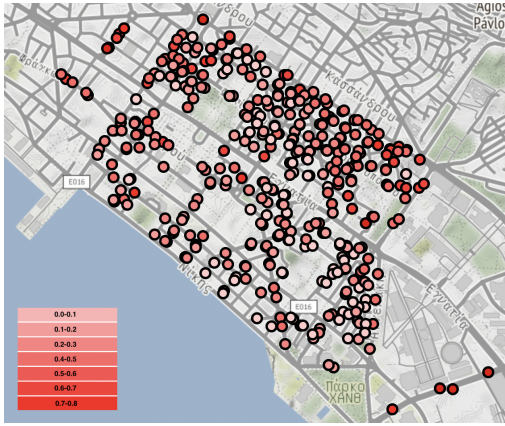


Fig. 3. Spatial distribution of parking violations. Darker values indicate higher average parking violation rate per sector

city’s centre, which are distributed into 396 separate sectors. The aim of the proposed method is to predict the parking violation rate for each of these sectors. We used historical data that consist of scans conducted by the city’s traffic police, as well as weather data from OpenWeather<sup>2</sup>. A total of 3.8 million scans that were captured through 300,000 checks were used in the conducted evaluation. Furthermore, we manually defined 19 PoIs used for encoding the distances for each sector, as described in Section II. Also, we used 80% of the data to train the model, while the remaining 20% was used to test the model. The mean absolute error (MAE) was used as the evaluation metric.

We provide a brief data analysis before using the collected data for predicting parking violations. Parking violation rate ranges between zero (0) and one (1), while the average parking violation rate is 0.41. Given that an average sector has eleven parking slots, this parking violation rates means that three to four (3-4) parking slots are expected to be illegally occupied. We also examined the spatial distribution of parking violations in Fig. 3, where darker colors indicate a higher parking violation rate. The higher mean parking violation of a sector is 0.78, while the lower is 0.1. We observe a higher parking violation rate in the upper region of the city. Another observation is that parking violations are higher in remote sectors. This can be probably explained since drivers might believe that these sectors are not as frequently scanned for illegal parking.

The evaluation results are reported in Table I. The baseline model without any smoothing achieves a MAE of 0.175, which demonstrates that we can accurately predict the parking violation rate using the proposed method, since the MAE of a baseline average value predictor (using the average violation rate as the prediction) is 0.251. Furthermore, using the proposed data augmentation and smoothing method for the train set allowed for further reducing the error to 0.169 on the original test set, while the error dropped to 0.146 when the data smoothing methodology was applied to the whole dataset. These results demonstrated the importance of employing the proposed data augmentation and smoothing approach to fill missing data and reduce data discontinuities.

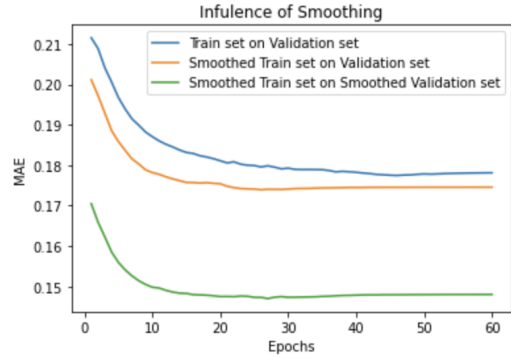


Fig. 4. Validation error during training when using two different setups and evaluation sets.

TABLE I. PARKING VIOLATION RATE PREDICTION EVALUATION

Method	Raw Test Set (MAE)	Smoothed Test Set (MAE)
Without Smoothing	0.175	0.173
With Smoothing	<b>0.169</b>	0.146

We also conducted a 4-fold cross validation experiments, where we monitored the validation MAE during the training epochs, to select the most appropriate number of training epochs for each model. The results are provided in Fig. 4. Note that using the smoothed training set leads to significant improvements over directly using the raw training set. We also plotted the MAE for the smoothed validation set, where again we observe that the training process converges smoothly, while achieving significantly better MAE, confirming the results reported in Table I.

#### IV. CONCLUSIONS

In this work, we presented and evaluated a DL-based model for on-street parking violation prediction. Several challenges were faced during the development of such a system, ranging from data sparsity challenges to encoding data in the most appropriate format for the employed DL model. To address these issues, we developed a novel approach that assigns the existing ground truth annotation to spatial sectors and time slots, leading to significant performance improvements. Other challenges include imprinting input data, such as the geographical location of each sector, time, and other information, such as pandemic measures (i.e., due to COVID-19) in an efficient manner. We evaluated the proposed method using novel data from the THESi system implemented in Thessaloniki, Greece, while we also presented various aspects of the employed data to provide more insight into this especially challenging problem.

There are several interesting future research directions. First, graph neural networks can be used to better model spatial relations between sectors [15]. Furthermore, traffic on city streets can affect the rate of parking violations, hinting that this could be an additional feature that can be included to further improve the performance of the presented model. Furthermore, even though city traffic information can be a valuable predictor, this information typically is not available for all sectors/roads, requiring the use of methodologies that can work even under the presence of partial information and/or employ methods for adaptively normalizing the input data and adjusting to the input distribution [16].

<sup>2</sup><https://openweathermap.org>



## REFERENCES

- [1] F. Vital, P. Ioannou, and A. Gupta, "Survey on intelligent truck parking: Issues and approaches," *IEEE Intelligent Transportation Systems Magazine*, vol. 13, no. 4, pp. 31–44, 2020.
- [2] C.-F. Yang, Y.-H. Ju, C.-Y. Hsieh, C.-Y. Lin, M.-H. Tsai, and H.-L. Chang, "iparking—a real-time parking space monitoring and guiding system," *Vehicular Communications*, vol. 9, pp. 301–305, 2017.
- [3] S. Gao, M. Li, Y. Liang, J. Marks, Y. Kang, and M. Li, "Predicting the spatiotemporal legality of on-street parking using open data and machine learning," *Annals of GIS*, vol. 25, no. 4, pp. 299–312, 2019.
- [4] J. Gao and K. Ozbay, "A data-driven approach to predict double parking events using machine learning techniques," in *Proceedings of the Transportation Research Board's Annual Meeting*, 2017, pp. 8–12.
- [5] T. Tiedemann, T. Vögele, M. M. Krell, J. H. Metzen, and F. Kirchner, "Concept of a data thread based parking space occupancy prediction in a berlin pilot region," in *Workshops at the Twenty-Ninth AAAI Conference on Artificial Intelligence*, 2015.
- [6] X. Ye, J. Wang, T. Wang, X. Yan, Q. Ye, and J. Chen, "Short-term prediction of available parking space based on machine learning approaches," *IEEE Access*, vol. 8, pp. 174 530–174 541, 2020.
- [7] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
- [8] M. Lukasik, S. Bhojanapalli, A. Menon, and S. Kumar, "Does label smoothing mitigate label noise?" in *Proceedings of the International Conference on Machine Learning*, 2020, pp. 6448–6458.
- [9] A. Tsantekidis, N. Passalis, and A. Tefas, "Improving deep reinforcement learning for financial trading using neural network distillation," in *Proceedings of the IEEE International Workshop on Machine Learning for Signal Processing*, 2020, pp. 1–6.
- [10] N. Passalis, M. Tzelepi, and A. Tefas, "Chapter 8 - knowledge distillation," in *Deep Learning for Robot Perception and Cognition*, A. Iosifidis and A. Tefas, Eds. Academic Press, 2022, pp. 165–186.
- [11] X. Glorot, A. Bordes, and Y. Bengio, "Deep sparse rectifier neural networks," in *Proceedings of the International Conference on Artificial Intelligence and Statistics*. JMLR Workshop and Conference Proceedings, 2011, pp. 315–323.
- [12] K. He, X. Zhang, S. Ren, and J. Sun, "Identity mappings in deep residual networks," in *Proceedings of the European Conference on Computer Vision*, 2016, pp. 630–645.
- [13] R. Llugsi, S. E. Yacoubi, A. Fontaine, and P. Lupera, "Comparison between adam, adamax and adam w optimizers to implement a weather forecast based on neural networks for the andean city of quito," in *2021 IEEE Fifth Ecuador Technical Chapters Meeting (ETCM)*, 2021, pp. 1–6.
- [14] F. Chollet *et al.*, "Keras," <https://keras.io>, 2015.
- [15] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and S. Y. Philip, "A comprehensive survey on graph neural networks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 1, pp. 4–24, 2020.
- [16] N. Passalis, A. Tefas, J. Kannianen, M. Gabbouj, and A. Iosifidis, "Deep adaptive input normalization for time series forecasting," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 31, no. 9, pp. 3760–3765, 2019.