IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 2 August 2023, accepted 7 September 2023, date of publication 12 September 2023, date of current version 21 September 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3314660

RESEARCH ARTICLE

Predicting Free Parking Slots via Deep Learning in Short-Mid Terms Explaining Temporal Impact of Features

STEFANO BILOTTA, LUCIANO ALESSANDRO IPSARO PALESI, AND PAOLO NESI, (Member, IEEE)
Distributed Systems and Internet Technologies Laboratory, Department of Information Engineering, University of Florence, 50121 Florence, Italy
Corresponding author: Paolo Nesi (paolo.nesi@unifi.it)

This work was supported in part by Ministero Industria Università e Ricerca (MIUR), University of Florence; and in part by the National Center on Sustainable Mobility, Centro Nazionale MObilità SosTenibile (MOST).

ABSTRACT Looking for available parking slots has become a serious issue in urban mobility, since it influences traffic and emissions. This paper presents a set of metrics and techniques to predict the number of available parking slots in off-street parking facilities. This study deals with deep learning model solutions according with a mid-term prediction of 24 hours, every 15 minutes. Such a mid-term prediction can be useful for citizens who need to plan a car transfer well in advance and to reduce as much as possible any computational effort. Since most solutions in literature are focused on 1-hour ahead prediction, the proposed solution has been also tested in these conditions. The proposed solution is based on Convolutional Bidirectional LSTM models. Results have been compared in terms of precision metrics based both on occupancy and free slots. The paper also provides a framework to pass from an assessment model based on occupancy to models based on free slots and vice-versa. The obtained results have improved those already available in literature. A formal study has been conducted to perform feature relevance analysis by using explainable AI technique based on gradient and integrated gradient and proposing new heatmaps which highlighted the difference from LSTM and Bidirectional LSTM, feature relevance (base line, weather, traffic, etc.) and the impact of seasonality on predictions, namely the temporal relevance of features. The comparison has been performed on the basis of data collected in garages in the area of Florence, Tuscany, Italy by using Snap4city platform and infrastructure.

INDEX TERMS Smart city, available parking lots, prediction model, machine learning, deep learning, explainable AI.

I. INTRODUCTION

Traffic management and sustainable mobility are central topics for intelligent transportation systems (ITS) so as to monitor and reduce vehicular traffic congestion [1], [2] and emissions [3], [4], [5]. Services providing available parking slots (in real time or as predictions) are becoming relevant for urban mobility management due to the increment of vehicles which need to park in cities. Drivers do waste a considerable amount of time while trying to find a vacant parking lot, especially during peak hours and in specific urban areas (e.g., hospitals, stations, parks, sport stadium). Searching for available parking spots can be a time-consuming task that simultaneously increases traffic congestion, thus leading to a peak of 25-40% of the traffic flow [6], [7] and greenhouse gas pollution.

Parking slots can be located on the street (they are called *on-street parking*) or in parking garages with gates (named as *off-street parking*). Searching for an available parking space has a harmful impact on both transportation system efficiency within the urban tissue and sustainability. Actually, any car parking searching activity generates unnecessary traffic workload and may affect the environment negatively due to increased vehicle emissions. These issues are surely valid

The associate editor coordinating the review of this manuscript and approving it for publication was Bo Pu.

when it comes to searching *on-street* parking, and they can be also considered in the context of off-street parking. More precisely, *off-street* parking with gates can be full in certain area and time slots; while in other areas, they may become full unexpectedly or due to conditions unknown to drivers. Recently, it is possible to collect real-time parking information - i.e., capacity, garage prices, number of empty parking slots in the silos or in the area - in order to realize predictive models. In terms of prediction models, there is a substantial difference between parking garages (i.e., off-street) and on-street parking. In fact, in parking garages, the total number of available slots can be estimated by considering the total tickets produced at the entrance gate, and the number of outputs from exits. On the other hand, as to on-street parking, it could be necessary to detect the occupancy by means of some distributed sensor systems. Thus, two distinct research lines can be found in the literature [8], [9], focusing on either on-street parking prediction or predicting free/available parking slots inside garages (off-street).

The car park selection performed by a driver is influenced by multiple factors, for instance: walking distance to destination, driving and waiting time, parking fees, service level, parking size, safety [10], [11], availability and accessibility [12]. In particular, two important attributes in a parking decision-making process of any driver are: the number of available parking spaces (if known), and past experience in finding available slots.

In this paper, a solution to predict the number of available/free parking slots (not occupied) in the parking garages with gates (e.g., silos, or on flat, or under railway station) is proposed. Predictions can be usually short-term (from 15 to 60 minutes in advance), and mid-term as 24 hours (1 day in advance). Most of the studies in literature have considered 1-hour time horizon, as short-term predictions to reduce local traffic and redirect drivers towards available services at short distance. Nevertheless, literature is quite poor with respect to mid-term predictions which can be more useful for citizens to program a dedicated car trip. For example, to know the number of available hospital parking slots 1-day in advance could be of significant value for patients who have scheduled medical examinations (or need to visit an admitted family member) at the hospital, so as to program within a day in advance any related transfer by car. Moreover, the predictive capability may be influenced by different characteristics of parking place locations. More precisely, different behaviors are registered in terms of free spaces for different kinds of parking places serving different areas: suburb hospital, shopping places area as well as down-town entertainment areas within historical and pedestrian city area; plus any possible combination of those aspects. As a result, it is expected that the produced predictions should be more accurate for cases presenting regular free slots trends and seasonality (daily, weekly, monthly, etc.), thus resulting in more predictable services, with respect to cases having strongly randomized behavior and trends. A partial solution to improve precision could be to consider additional information and real time data variables such as: the description of the parking area neighborhoods, the real time or predictions about traffic flow, and the information related to weather conditions and/or forecasts [13], [14], [15].

This paper is focused on presenting an approach for short- and mid-term prediction of the number of free slots on off-street parking area which overcome the solutions from literature in terms of precision in critical conditions. The solution proposed in this paper is based on convolutional bidirectional deep learning. In addition, the paper is presenting a new approach for reasoning on feature relevance of deep learning which addressed the aspects of feature and prediction seasonality, thus, contributing to improve techniques for explainable artificial intelligence, XAI, with feature relevance magnitude and time.

The next subsection is devoted to recall the related work to allow readers to better contextualize this research topic. After that, a more detailed description of the research with its aim and goals is reported. Finally, the paper structure is presented and commented.

A. RELATED WORKS

In literature, the problem of parking predictions has been addressed through different approaches, most of the recent proposals are using deep learning techniques [13], [16], [17], [18]. Most of them are grounded on time-series predictions, and in particular with deep recurrent neural networks, because of their capability of exploiting previous data observations. Data are typically collected from parking slots in constrained areas (*off-street* parking), for example, parking garages/facilities with gates where the number of offered slots is typically high, and the whole status is clearly reported in real time at the garage entrance gate. Therefore, they offer a strong appeal to drivers that may arrive all together, especially when parking facilities are located closer to attraction centers such as commercial centers, hospitals, railway stations, theatres, and multiservice areas. Differently, the turnover is faster for *on-street* parking, which may be reserved to specific categories (e.g., resident, shops, wheelchairs). The availability of the remaining free slots *on-street* is more unpredictable than the availability on *off-street* parking areas no matter if public or private. The parking *occupancy* is defined as the fraction of slots that are occupied by vehicles as part of the total number of potentially available parking slots for the parking area/facility. A number of representative papers dealing with various aspects of *on-street* and *off-street* parking are reviewed and discussed in details in [9].

In **Table 1**, a comparative summary of the state-of-the-art solutions is reported for slots predictions in off-street parking facilities, since this is the focus of the paper. Such a comparison highlights the predictive goals, the adopted features, the used techniques, and the obtained results in terms of metrics: *RMSE* (Root Mean Square Error), *MAPE* (Mean Absolute Percentage Error), *MSE* (Mean Squared Error) and *MAE* (Mean Absolute Error). The definition of some of these metrics is reported later in this paper, others are discussed here as follows.

Please note that, in some cases reported in the literature, assessment metrics have been provided in terms of occupancy (defined as *(parking capacity - number of free slots) / parking capacity ∗ 100)*, in others in terms of free number of slots. The two mentioned assessment approaches are not equivalent. For example, in terms of *MAE*:

$$MAE = \frac{\sum_{i=1}^{n} |obs_i - pred_i|}{n} \tag{1}$$

According to the definition of *occupancy*, the estimation of *MAE* occupancy, *MAEo*, depends on the parking size, that is, the *Capacity*:

$$MAEo = \frac{\sum_{i=1}^{n} |obs_i - pred_i|}{n\ Capacity} 100 \tag{2}$$

Thus, the relationship between *MAE* assessed on free slots, *MAEf*, and *MAEo* results to be:

$$MAEf = \frac{MAEo\ Capacity}{100}. \tag{3}$$

For *MSE*:

$$MSE = \frac{\sum_{i=1}^{n} (obs_i - pred_i)^2}{n} \tag{4}$$

We have the following relationship between *MSE* assessed on free slots, *MSEf*, and that based on occupancy, *MSEo*:

$$MSEf = \frac{MSEo\ Capacity^2}{100^2} \tag{5}$$

Similarly, for *RMSE* which is calculated as: $RMSE = \sqrt{MSE}$, thus:

$$RMSEf = \frac{RMSEo\ Capacity}{100} \tag{6}$$

Unfortunately, similar relationships cannot be derived for other metrics (*MAPEo*, *MAPEf*), since they are un-linearly dependent on the parking *Capacity*. Therefore, the comparison of results by using a set of standard metrics should be carefully performed. Reasons are: (i) the usage of non-comparable metrics depending on *capacity* and lack of details, (ii) the usage of different parking data sets, (iii) the adoption of additional variables such as weather, traffic, etc., (iv) the computation of assessment metrics as average in the day or week period, instead of providing precision metrics in critical conditions, which typically occur when the parking facility risks to become empty (the observation becomes close to zero). Thus, when performing the analysis of the state of the art, the identification of the best model has to take into account the results obtained in the same paper, using the same metrics and data. Moreover, an additional analysis should be derived from the explainability of machine learning and deep learning techniques, in relation to additional variables and seasonality, as discussed hereafter.

In [14], classic machine learning and predictive models have been used for off-street parking, such as Bayesian Regularized Artificial Neural Network (BRANN) [19], Support Vector Regression (SVR) [20], Recurrent Neural Network (RNN) [21], and Autoregressive Integrated Moving Average (ARIMA) [22]. The adopted feature space has included historical data, seasonal information (day, day of the week, etc.), weather aspects, and traffic flow data. Both model and comparison assessment has been performed to provide short- and mid-term predictions, every 15 minutes for the next 24 hours, in different parking context (e.g., parking lots serving markets, hospital, railway stations). The metric to assess such a performance was mainly *MASEf* (Mean Average Scaled Error) on free slots which can produce reliable assessment without falling in singularity when the number of free slots is close to zero. Experiments have demonstrated a better performance for BRANN with respect to other models. More recently, in [23], models based on Neural Network (NN) [25], Convolutional NN (CNN) [24], and Random Forest (RF) [26], have been compared with the aim of predicting off-street occupancy in the range of 15'-60' minutes in advance. The exploited features are similar to the ones presented in [14].

The results in [23] could demonstrate the validity of NN for any occupancy prediction (from 15 to 60 minutes in advance) in terms of *MSEo*, *MAEo* and *RMSEo* (which are dependent on the size of the parking areas), with respect to CNN, RF. In [23], the model performance has been evaluated in the parking areas of Arnhem (NL) having capacity of 1050 spaces, not reported in the paper.

In [27], a comparison of RF and CatBoost [65] has been proposed as to parking predictions. CatBoost is based on the ensemble learning method called gradient boosting and it combines different learners to get a stronger learner. The assessment has been conducted in the city of Split where parking areas are equipped with sensors and the prediction is related to the concerned area. The city of Split has 50 parking areas equipped with ground parking sensors, or a total 1516 parking areas. The study has been conducted on 44 areas and the CatBoost resulted to be the best model for predicting park utilization/occupancy. Moreover, a feature relevance analysis has identified the most important features: the historical data of parking capacity, weather conditions. In [27], the model performance has been evaluated in parking areas having an average capacity of 25 slots.

In [28], a set of statistical models, machine learning and deep learning approaches have been tested: vector regressive (VAR) [29], Gated Recurrent Units (GRU) [30] which is a class of RNN, Graph Convolutional Neural Network (GCNN) [31]. The assessment performed on a number of data sets has identified the GCNN as the best results in terms of *RMSEo* for predicting short-term occupancy of parking areas. Models have been tested on three different datasets: 76 on-street parking areas in Italy with an average capacity of 32.69 lots; 420 on-street parking area in San Francisco with an average capacity of 9.29 stalls; 17 off-street parking area in Birmingham with an average size/capacity of 676.88 stalls. In this case, the model has considered temporal features and a distance among stalls for on-street parking and an interesting

TABLE 1. Related work solutions for off-street parking predictions.

<table>
<tr><th>Authors</th><th>Target</th><th>Features</th><th>Models</th><th>Best Model / Results / performance</th></tr>
<tr><td>J. C. Provoost et al. [23] (2020)</td><td>Off-street occupancy prediction: 15'-60' in advance (capacity 1050)</td><td>Parking data, traffic flow data, weather data</td><td>NN, CNN, RF (predictions at 60')</td><td><table><tr><th></th><th>NN</th><th>CNN</th><th>RF</th></tr><tr><td>MSEo</td><td>7.18</td><td>8.27</td><td>7.98</td></tr><tr><td>MAEo</td><td>1.91</td><td>2.20</td><td>1.92</td></tr><tr><td>RMSEo</td><td>2.68</td><td>2.88</td><td>2.82</td></tr></table></td></tr>
<tr><td>G. Jelen et al. [27] (2021)</td><td>occupancy of parking areas prediction: 60' in advance (average capacity: 25)</td><td>Ground parking sensors data, weather data</td><td>catBoost, RF (predictions at 60')</td><td><table><tr><th></th><th>CatBoost</th><th>RF</th></tr><tr><td>MAEo</td><td>6.71</td><td>6.85</td></tr><tr><td>MSEo</td><td>78.62</td><td>90.57</td></tr></table></td></tr>
<tr><td>C. Lucchese et al. [28] (2022)</td><td>Occupancy on/off-street parking prediction: 15', 30', 45' and 60' (average capacity: 678)</td><td>Spatio-temporal parking data</td><td>VAR, GRU, GCNN (predictions at 60')</td><td><table><tr><th>GCNN</th><th>RMSEo 60'</th></tr><tr><td>Parking in Italy</td><td>0.073</td></tr><tr><td>Parking San Francisco</td><td>0.131</td></tr><tr><td>Parking Birmingham</td><td>0.042</td></tr></table></td></tr>
<tr><td>M. K. Mufida et al. [33] (2021)</td><td>Occupancy off-street parking prediction: 10' in advance (up to 120')</td><td>Parking data (Euralille capacity: 2600, Gare Lille: 520)</td><td>SVR, MLR, SARIMAX, RNN-LSTM</td><td><table><tr><th>RNN-LSTM</th><th>Euralille</th><th>Gare Lille</th></tr><tr><td>RMSEo</td><td>0.126</td><td>0.556</td></tr><tr><td>MAPEo</td><td>0.66</td><td>0.03</td></tr></table>MAPEo > 6% at 60'</td></tr>
<tr><td>C. Zeng et al. [16] (2022)</td><td>Availability of free slots off-street parking prediction (underground car park): 10' in advance</td><td>Parking data (capacity: 250)</td><td>LSTM, CONV-LSTM, BI-LSTM, CONV-BI-LSTM, DWT-Bi-LSTM</td><td><table><tr><th></th><th>CONV-BI-LSTM, 10'</th><th>DWT-Bi-LSTM 10'</th></tr><tr><td>RMSEf</td><td>10.27</td><td>7.42</td></tr><tr><td>MSEf</td><td>105.65</td><td>55.16</td></tr><tr><td>MAEf</td><td>7.53</td><td>4.64</td></tr></table></td></tr>
<tr><td>Y. Feng et al. [17] (2022)</td><td>Availability of free slots off-street parking prediction: 5',15', 30', 45', 60' in advance</td><td>Spatio-temporal parking data (St7 capacity: 800)</td><td>LSTM, GRU, SVR, KNN, etc., and dConvLSTM-DCN (predictions at 60')</td><td><table><tr><th>(St7)</th><th>dConvLSTM-DCN, 60'</th></tr><tr><td>RMSEf</td><td>24.60</td></tr><tr><td>MAEf</td><td>17.69</td></tr><tr><td>MAPEf</td><td>7.28</td></tr></table></td></tr>
<tr><td>R. K. Kasera, T. Acharjee [18] (2022)</td><td>Off-street occupancy rate prediction: time window not formally defined</td><td>Parking data (capacity from 300 to 480)</td><td>ARIMA, MLP, CNN, LSTM, GRU, CNN-LSTM</td><td><table><tr><th></th><th>CNN-LSTM k-Avg. Percent</th></tr><tr><td>RMSEo</td><td>28.65</td></tr><tr><td>MAEo</td><td>8.54</td></tr></table></td></tr>
<tr><td>E. S. Fokker et al. [13] (2022)</td><td>Occupancy off-street parking prediction: 60' and 6 months ahead (capacity from 260 to 1000)</td><td>Parking data, weather data, events, parking fee, public transport lines data</td><td>SARIMAX, ETS (6 months), LSTM (predictions at 60') (5) is a parking which may go to 0 free slots in daily hours</td><td><table><tr><th>Diff. Parking</th><th>LSTM</th><th>SARIMAX</th></tr><tr><td>RMSEo (1)</td><td>0.202</td><td>28.03</td></tr><tr><td>RMSEo (2)</td><td>1.104</td><td>43.73</td></tr><tr><td>RMSEo (3)</td><td>1.061</td><td>51.89</td></tr><tr><td>RMSEo (4)</td><td>0.656</td><td>51.20</td></tr><tr><td>RMSEo (5)</td><td>2.179</td><td>60.25</td></tr><tr><td>RMSEo (6)</td><td>1.803</td><td>70.76</td></tr></table></td></tr>
</table>

feature relevance analysis on distance for on-street parking has been provided.

Most recent studies are focused on the application of hybrid deep learning techniques, in particular Long Short-Term Memory (LSTM) models [32], [34], by using the (baselines) features related to parking data time-series only. In [33], SVR, Multiple Linear Regression (MLR) [35], SARIMAX (which is seasonal ARIMA with exogenous variables) [36], and RNN-LSTM [37] have been compared. The experiments performed on several off-parking structure have demonstrated that RNN-LSTM outperformed the other models. The used dataset contained data regarding occupancy of 27 parking lots, representing a total of 18180 parking areas. The performance assessment has been conducted in Parking Euralille (center commercial) and Parking Gare Lille Europe (railway station) in the North of France, with a capacity close to 2500 and 500 stalls, respectively.

In [16], several models have been tested: LSTM, CONV-LSTM (convolutional LSTM), BI-LSTM (Bidirectional LSTM), CONV-BI-LSTM, and DWT-BI-LSTM which integrated a wavelet transform with BI-LSTM, as proposed in [38]. Such a transform divides the entire time domain into equally spaced local regions, each of which is approximated as smooth, and then the Fourier series is computed. It decomposes the signal into multiple wavelet functions according to two parameters, displacement and scale factor, in order to carry out a better extraction of time-frequency components and reduce any noise effect in the data for short-terms prediction. The described techniques are used to predict free slots of off-street parking in the next 10 minutes using the parking historical data. The used dataset includes data about the occupancy of 2 off-street parking in Chongqing. In [16], the model performance is evaluated in one of them, having its capacity close to 250 spaces.

In [17], few models have been compared: LSTM, GRU, SVR, KNN, etc., and the proposed dConvLSTM-DCN (dual Convolutional Long Short-Term Memory with Dense Convolutional Network, DCN). The latter turned out to be better to predict time slots from 5' to 60' minutes. It consists of two parallel ConvLSTM components and a DCN [39] to fully exploit the spatial-temporal correlations in the historical data related to free parking spaces considering 9 parking lots. In such a study the target is the availability of prediction related to free off-street parking spaces, up to one hour in advance. In [17], the used dataset contains the free parking spaces of 9 off-street parking lots in California. The model performance comparison is evaluated on parking (St7) with capacity close to 800 slots.

In [18] a hybrid CNN-LSTM model is proposed with multilevel parking occupancy percentage prediction via different time steps (up to 100) where no additional information in terms of external data is considered. However, the adopted model has a best fit with respect to a set of compared models: ARIMA, MLP, CNN, LSTM and GRU for off-street occupancy rate prediction in terms of *RMSEo* and *MAEo*. The used dataset contains data about the occupancy rate of an off-street multilevel parking which is formed by 3 floors having the following capacity: 300, 317 and 480 slots, and it is not clear which parking area results are referring to. The prediction is conducted via time steps (from 1 to 100 time-steps) and a time window is not formally defined. As results, it is reported the average percentage of parking occupancy rate accuracy at k time steps (k-Average Percent).

In [13], the used dataset is enriched with weather data, events, parking fees and public transport lines data leading to both short-term prediction (60 minutes in advance) and long-term prediction (6 months ahead) for 57 off-street parking occupancy. In such a work, the comparison of models is among: SARIMAX (seasonal ARIMA with exogenous variable) [40], ETS (exponential smoothing) [41], LSTM. As described in [13], external variables, which affected parking occupancy, are related to weather conditions and events.

In [13], the used dataset contains data about the occupancy of 57 off-street parking lots in Amsterdam. The model performances have been evaluated for 6 of them having the capacity respectively close to: (1) 260, (2) 360, (3) 450, (4) 350, (5) 400, (6) 1000 (numbers corresponding to the assessment in **Table 1**).

B. PAPER'S AIM AND ORGANIZATION

This work is focused on presenting research results related to a solution for predicting the number of available parking slots (free slot of off-street parking facilities) for garages in the city of Florence. Prediction of available parking spaces is a complex non-linear process whose dynamic changes involve multiple factors. Parking facilities provide several different working conditions. Some of them are dedicated to a specific facility (football stadium, hospital), others are meant for multipurpose (station, expo, downtown, etc.), and others are located on city outskirts. Variability and performance are one of the main problems to be addressed, together with the precision in critical time slots, which is when the parking is getting full, running out of available slots.

The major focus of this paper is on:

- Identification of the best prediction models among a number of machine and deep learning techniques covering what has been presented in the literature, for example: BRANN, RNN, CNN-GRU, CNN-LSTM, and CNN-BI-LSTM.
- The production of predictions, not only few minutes in advance, but for the **next 24 hours, with 15-minute sampling** with satisfactory precision. This would drastically reduce the computing prediction costs, since prediction is performed only once per day and not every few minutes, as typically proposed in literature. In addition, it may be of help for drivers planning their travel the day before. Most of state-of-the-art solutions produce predictions 1-hour ahead.
- Clarify how the different solutions can be compared one another and compare the results on the basis of assessment metrics. Compare the results in terms of Occupancy rate and Free slots according to their corresponding metrics, in the 1-hour slot prediction.
- Provide prediction results with respect to critical conditions for parking occurring when the parking facility is almost empty or full. This factor has been neglected in most solutions of related literature.
- Propose explainability techniques (XAI, explainable artificial intelligence) for assessing feature relevance (among the several ones identified in the literature) in terms of their magnitude and temporal impact, or seasonality. This approach has been adopted for the proposed deep learning solution and can be used for almost all ML and AI models in literature.

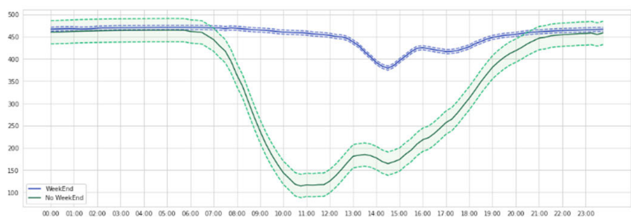
The proposed prediction model has been created in the context of a national center on sustainable mobility (MOST, in Italy) within the spoke on urban mobility and funded by the Ministry of Research, and by exploiting data and facilities of Snap4City, https://www.snapcity.org, infrastructure in the Florence / Tuscany area, Italy for Smart City [42], [43].

The paper is organized as follows. Section II considers the data description and the features definition applied to our field of research. In Section III, the deep learning model definitions are presented with their hyper-parameters tuning. Section IV presents the obtained results with short and mid-terms prediction of 1-hour and 24-hours in advance, respectively. The feature relevance analysis for the best model is described in Section V by applying gradient-based techniques. Finally, conclusions are drawn in Section VI.

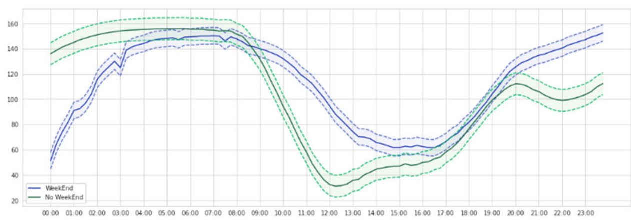
II. DATA DESCRIPTION AND FEATURE DEFINITION

The main goal of current work is to find a solution to predict the number of available parking slots (free slots) in parking garages/facilities, for example controlled by a gate (off-street parking facilities). This study concerns off-street parking located in the municipality of Florence so as to identify a common predictive and flexible model for different parking areas. Such areas are of different capability and provide different behaviors in different days of the week, as well as moments of the day. Some of them may experience critical conditions when the available parking slots are close to zero. **Figure 1** reports the typical daily trends of available slots for the considered parking garages, where workdays and weekends are examined. Trends are significantly different, and their related behaviors depend on the city areas and services. Therefore, there is clearly a daily and weekly seasonality. More precisely, we have considered three different representative parking areas: a suburb hospital parking location (namely *Careggi car park*) the behavior of which is reported in **Figure 1(a)**, with its corresponding confidential trend considering the related standard deviation, and a capacity of 514 slots; a central parking location (namely *Beccaria car park*) which is very close to the historical and pedestrian city area, having its related behavior reported in **Figure 1(b)** and a capacity of 203 slots; a parking garage located in the historical center of Florence (namely *S. Lorenzo car park*) having its related behavior reported in **Figure 1(c)** and a capacity of 179 slots. Please note that trends (working days and weekends) have been reported with their standard deviations. This permits to stress their corresponding noise level. Among them the noisiest one is *S. Lorenzo car park* due to the daily hours of market activities and the evening hours of *movida* and restaurants.

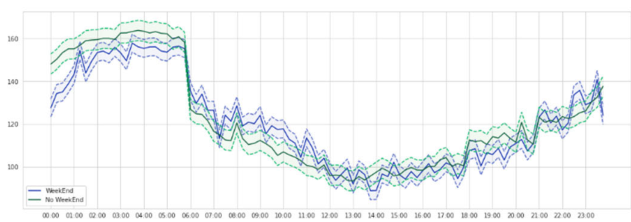
The used data refer to the period from March 1, 2022, to June 5, 2022. For each parking facility, the number of available slots has been checked and registered every 15 minutes.



(a) Careggi car park (hospital), capacity: 514



(b). Beccaria car park (downtown access), capacity: 203



(c) S. Lorenzo car park (old town access), capacity: 179

FIGURE 1. Typical daily trends (and std. dev) of available lots for different parking garages in Florence, where workdays and weekends are examined.

According to the state-of-the-art analysis, three groups of features have been identified as possible predictive metrics and they are briefly discussed in the following (see **Table 2**).

The features related to the *Baseline* category consider measures related to direct observation of parking data and derived information over time. This category is ordered on the basis of date and time when measures are taken. The latter include number of available slots, working day or not, etc. Values are recorded every 15 minutes. These variables are used to consider the data seasonality, which may have different trends – i.e., the workdays with respect to the weekends, etc. Two other features have been included in the model:

- POD: difference between the actual and previous number of available spaces at the same time, recorded one week before;
- SOD: difference between the actual number of parking spaces and the next one at the same time, recorded one week before.

TABLE 2. Overview of the exploited features that can be of help in describing the context of parking usage with their: category, features and description.

Category	Features	Description of features variable
Baseline features of free slot data	Free parking slots	Real number of available slots recorded every 15 minutes
	dateTime	Date and time (day, hour and minutes)
	dayWeek	0 for working days, 1 else
	Previous observation's difference (POD)	Difference between the number of free spaces at time i and number of free spaces at time ($i-15$ minutes) recorded in the previous week
	Subsequent observation's difference (SOD)	Difference between the number of free spaces at time i, and the number of free spaces at time ($i+$ 15 minutes) recorded in the previous week
Weather features	Temperature	City temperature measured one hour earlier than Time (°C)
	Humidity	City humidity measured one hour earlier than Time (%)
Traffic Sensors features	Average Vehicle Speed	Average speed of vehicles on the road nearest to the parking, over one-hour period (km/h)
	Vehicle Flow	Number of vehicles passing nearest to the parking, over one-hour period
	Average Vehicle Time	Average of distance between vehicles, over one-hour period
	Vehicle Concentration	Number of vehicles per kilometer, over one-hour period

Features belonging to *Weather* are also collected every 15 minutes, that is, *temperature* and *humidity*. According to our analysis, the most significant values are those related to the hour before the parking time in the context of *1-hour* prevision time horizon. Therefore, in order to predict the number of free slots in a garage at 3 pm, weather features at 2 pm are relevant. In fact, weather conditions typically affect decisions, when it comes to taking the car or the public transportation. For example, the expected behavior held by citizens when it rains (according to an appropriated value of humidity and temperature), is to drive a car, instead of the motorcycle. By doing so, more parking lots will be occupied. On this line, one would suppose exploit long term weather forecasts (6 hours or days in advance), since they could also impact on decisions (they are accessible on the Snap4City smart city platform).

Features related to *Traffic Sensors* refer to values of traffic recorded by sensors which are in the neighborhood of the parking area, and mainly on the streets that can be used to reach the area [44]. Not unlike from weather data, traffic sensors values seem to be relevant if available for the previous hour with respect to the time of prediction. Typical values are the ones related to vehicle flow, concentration, average time and average speed [45]. They are estimated every 15 minutes. These metrics adopted for traffic flow estimation are those that typically are accessible from city traffic flow sensors. In this context, the value of traffic flow is used for assessing traffic conditions, and thus average values are satisfactory. Traffic sensors which are relevant for each garage may be one or more and they should be chosen considering the direction of travel and the most likely route to reach the garage. On the other hand, for other applications, such as routing path finding or what-if analysis in traffic reconstruction, more precise data and predictions should be used [46], [47], [48], [49].

III. DEFINITION OF THE PREDITING MODELS

This study can be divided into different steps: data collection, pre-processing, model training, evaluation, and deployment. In this process a large amount of data about parking occupancy must be collected, cleaned, normalized and imputed. Data missing is an unavoidable problem when dealing with real-world sensors. Parking data sensors may suffer of problems such as detector malfunction and communication failure, while there could be also some problems during every data acquisition process. All these problems can affect the monitoring of the parking state and may limit the predictive capability of the predictive models at runtime. In general, approaches of data imputation for producing surrogate data may help in creating dense data in training and execution [50].

The settings used the temporal data explicitly as the input for models, therefore it was necessary to have observations consistent and complete. Temporal data has been so rearranged to have 96 timestamps per day (24 hours and samples every 15 minutes). Missing observations of parking data have been imputed as follows: at a timestamp t of the day d a missing observation has been imputed using the average of the timestamp t of the most similar 3 days to the same day d. The model used for the imputation is K-NearestNeighbours [51] with k equal to 3, where the similarity is computed as the Euclidean distance with the nearest neighbors. More precisely, in training, we consider the imputation of missing free car parking data which corresponds to 1.1% of the measured (original) dataset.

In our general framework, different approaches are tested: BRANN, RNN, CNN-LSTM, CNN-BI-LSTM and CNN-GRU models applied well to the features presented above. More precisely, BRANN is the model producing the most reliable results according to the previous study carried out in [14] by considering data in the same car parks in Florence as it occurs in this study. Then, we have considered the most recent deep learning techniques by applying CNN-LSTM, CNN-BI-LSTM and CNN-GRU models to the datasets to improve the prevision performance.

A. ARTIFICIAL NEURAL NETWORKS WITH BAYESIAN REGULARIZATION

The Artificial Neural Network (ANN) is a supervised learning technique and it inspired by theories about how the human brain works [52], [53], [54]. Usually, ANNs tends to overfit, which in substance means to have trained the NN (Neural Network) to fit the noise trend without producing a good generalization, as expected by the ANN. However, Bayesian Regularized ANNs (BRANNs) attempts to overcome the overfitting problem by incorporating Bayes' modeling into the regularization scheme [19]. In general, the risk of overfitting increases when a neural network grows through additional hidden layer neurons. BRANN approach avoids the overfitting because regularization pushes unnecessary weights towards zero. On such grounds BRANN method is more robust and efficient than classical ANNs and network weights are typically more significant in modeling the phenomena [19]. BRANN model fits a three-layer neural network as described in [55] and [66]. The layer weights the network, which is initialized by the Nguyen-Widrow initialization method [56], and thus, the model is given by:

$$
\begin{aligned}
&y_i = g\left(x_i\right) + e_i \\
&y_i = \sum\nolimits_{k=1}^{s} w_k g_k \left(b_k + \sum\nolimits_{j=1}^{p} x_{ij}\beta_j^{[k]}\right) + e_i, \quad i = 1, \ldots, n
\end{aligned} \tag{7}
$$

where:

- $e_i \sim N\left(0,\sigma_e^2\right)$;
- s is the number of neurons;
- w_k is the weight of the k-th neuron, $k = 1, \ldots, s$;
- b_k is a bias for the k-th neuron, $k = 1, \ldots, s$;
- $\beta_j^{[k]}$ is the weight of the j-th input to the net, $j = 1, \ldots, p$;
- $g_k(\cdot)$ is the activation function: in this case:

$$
g_k(x) = \frac{e^{2x} - 1}{e^{2x} + 1} \tag{8}
$$

The objective function consists in minimizing $F = \alpha E_W + \beta E_D$, where E_W is the sum of squares of network parameters (weight and bias), and E_D is the error (sum of squares), α and β are the objective function parameters.

B. RECURRENT NEURAL NETWORK

Neural Networks have been the focus of great interest for many decades, due to the desire to understand the human brain and to build learning machines. Recurrent Neural Networks (RNNs) are basically a Feedforward Neural Network with a recurrent loop [21]. RNNs are considered a powerful model for sequential data and they are applied to a wide variety of problems involving time sequences of events and ordered data. RNN are neural networks that consists in a hidden state h and an output y that operates on a sequence of variables $\boldsymbol{x} = (x_1, \ldots, x_T)$. At each time step t, the hidden state of the RNN is updated by $h_{(t)} = f\left(h_{(t-1)}, x_t\right)$, where f is a non-linear activation function. Note that, while in principle the recurrent network is a simple and powerful model, in practice, it is hard to train properly [57].

C. CNN-LSTM

Long short-term memory (LSTM) network is an improved recurrent neural network (RNN). Many studies have shown that CNN-LSTM, which is based on a combination of a CNN layer followed by an LSTM layer, shows better prediction performance than single CNN or LSTM models [58], [59], [60]. A CNN layer normally computes input data via a convolution layer and a pooling layer and then computes the characteristic descent. The result value is obtained based on the characteristic descent. As soon as the CNN layer computation is completed, the outcome is calculated via the LSTM layer, and result values are determined. The addition of a convolutive layer is grounded on the advantage of combining powerful feature extraction of CNN with LSTM capability in capturing temporal dependencies. Two main customization techniques can be considered to optimize the model performance. They concern both architecture of the CNN-LSTM network and tuning of the hyper-parameters used during the training phase. Generally, to determine the LSTM network structure, the following variables are considered: number of layers, number of neurons and dropout function.

Moreover, as to hyper-parameters tuning of the mentioned network, several items also are considered, namely: type of optimizer, learning rate, momentum, etc. Regarding CNN network, hyper-parameters tuning is primarily related to the following items: number of filters and kernel size. In the case of study, CNN-LSTM has a 6-layers architecture structured such as:

- the first two layers are made of CNN unit layer and average pooling layer;
- the other four layers are made of LSTM network having 2 hidden layers, a dropout layer and a last layer which is a Dense layer with 96 units.

The implementation of recurrent neural networks is stateful with a number of timesteps considered equal to 96, which corresponds to the data of one day prior (24 hours) to the observation/prediction time in the case of mid-term prevision of 24 hours in advance.

The training process has been made with early stopping with patience set to 60 and weights restored to the best model. Hyperparameters, that have been optimized through a Random Search, are reported in **Table 3**.

The structure of the defined CNN-LSTM network is made up of the following 2 components:

- The first component is made up of a Convolutional 1dimensional layer with 16 filters and a kernel size of 3, and a AveragePooling 1dimensional layer.
- The second component consists in LSTM layers, in particular: 2 hidden layers with 1024 and 32 units, respectively, a dropout layer of 0.3 and a dense layer with 96 units.

The considered activation function is ReLu. The used optimizer is Adam Optimizer with learning rate equals to 0.001 and momentum equals to 0.9. *MAE* was selected as the loss function to be monitored during optimization. The batch size has been set to 16 and the number of epochs was set to a maximum value of 600, because the training strategy used the Early Stopping method with patience parameter set to 60 to determine the optimum epoch number, restoring the weights of the best model at the end of the learning process.

TABLE 3. Hyperparameter optimized.

hyperparameters	value
LSTM units number	256, 512, 1024, 2048
LSTM-1 units number	16, 32, 64, 128
Filters number	4, 8, 16, 32, 64
Kernel size	2, 3, 4, 5
Dropout rate	0.0, 0.1, 0.2, 0.3, 0.4, 0.5 0.6
Learning rate	0.1, 0.01, 0.001, 0.0001
Momentum	0.0, 0.2, 0.4, 0.6, 0.8, 0.9, 0.99
Optimizer	SGD, RMSprop, Adam

D. CNN-BI-LSTM

The LSTM layers in the above model are improved by means of a *Bidirectional* approach. The idea of Bidirectional LSTMs (BI-LSTM) is to aggregate input information in the past and future of a specific time step in LSTM models. In BI-LSTM, at any point in time, it is possible to preserve information from both past and future. In particular, the structure of the defined CNN-BI-LSTM network is made up of the following 2 components:

- The first component is made up of a Convolutional 1dimensional layer with 16 filters and a kernel size of 3, and a AveragePooling 1dimensional layer.
- The second component consists in BI-LSTM layers, in particular: 2 hidden layers with 2048 (1024 +1024) and 64 (32 + 32) units, respectively, a dropout layer of 0.3 and a dense layer of 96 units.

All model settings in CNN-LSTM are preserved in the CNN-BI-LSTM model.

E. CNN-GRU

For the sake of completeness, we have also considered the described model for Gated Recurrent Units (GRUs) application. GRUs are a gating mechanism in recurrent neural networks introduced in [61] and they are considered as a LSTM variation, because both are designed similarly. In particular, GRU can be seen as a long short-term memory (LSTM) with a forget gate, then GRU has fewer parameters than LSTM, as it lacks an output gate.

F. CONSIDERATIONS

Actually, LSTM, BI-LSTM and GRU are improved recurrent neural network (RNN) and we are going to compare them in our case of study preserving the same model setting (as described above). In short, the structure of the advanced RNN network has a CNN component, 2 (LSTM/BI-LSTM/GRU) layers, a dropout layer of 0.3 and a dense layer of 96 units, respectively. The architecture of the adopted CNN-BI-LSTM model and the related graphical representation is depicted in **Figure 2**.

Layer (type)	Output Shape	Param #
conv1d (Conv1D)	(None, 166, 16)	496
average_pooling1d (AveragePooling1D)	(None, 83, 16)	0
bidirectional (Bidirectional)	(None, 83, 2048)	8527872
dropout (Dropout)	(None, 83, 2048)	0
bidirectional (Bidirectional)	(None, 64)	532736
dense (Dense)	(None, 96)	6240

Total params: 9,067,344
Trainable params: 9,067,344
Non-trainable params: 0

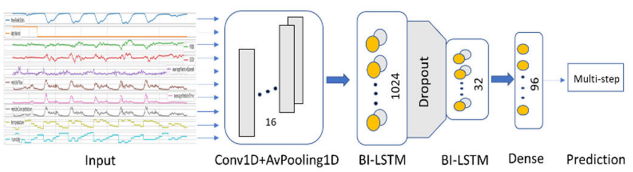


FIGURE 2. The adopted CNN-BI-LSTM model architecture.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

According to the data and considerations reported above, the identified challenge was to create a model and tools able to predict the number of free slots in off-street parking facilities with a resolution of 15 minutes for the next 24 hours, primarily. Each predictive model produces 96 predictions for the next day and it can be executed for updating them every 24 hours. Therefore, updating every 15', not even every hour, would not be strictly needed.

As usual, the dataset is split in several parts used for model training in the learning phase and testing it, that is comparing the forecasts obtained with the trained model with the actual data contained in the testing set. As a training dataset we have elected a sample of three months, from March 1st, 2022, to June 1st, 2022. The test set has been composed by 96 daily observations (every 15 minutes) recorded during the weeks from May 23th (Monday) to June 5th (Sunday).

A. ERROR MEASUREMENT DEFINITION

As discussed in the Introduction, to calculate the prediction error, in literature most researchers have adopted the occupancy approach using *MAEo*, *MSEo* and *RMSEo*. In this paper, we prefer the use of metrics based on free slots to produce measures which are independent of the parking area capacity.

The identification of the model for measuring the error is very relevant, since it has to work well close to zero. This is due to the fact that on the particular issue of street-parking predictions, critical cases occur when the available parking slots are close to zero. For this reason, we have chosen the Mean Absolute Scaled Error (*MASE*) by Hyndman and Koehler, 2006 in [62]. The Mean Absolute Scaled Error is calculated as follows:

$$MASE = mean\left(|q_t|\right), \quad t = 1, \ldots, n \tag{9}$$

and

$$q_t = \frac{obs_t - pred_t}{\frac{1}{n-1}\sum_{i=2}^{n}|obs_i - obs_{i-1}|} \tag{10}$$

where:

- obs_t = observation at time t
- $pred_t$ = prediction at time t
- n is the number of the values predicted over all test sets (96 daily observations per 7 days).

Note that, as it can be easily verified, *MASEf* is identical to *MASEo* and is clearly independent of both scale of the data and capacity. When *MASE* is used for comparing predictive models, the best model is the one presenting the smaller *MASE*. Therefore, the *MASE* should be the best solution to compare solutions assessed on the basis of occupancy or free slots.

Additional metric, with respect to those presented above and, in the introduction, can include Mean Absolute Percentage Error (*MAPE*), which is calculated as follows:

$$MAPE = \frac{\sum_{i=1}^{n}\left|\frac{obs_i - pred_i}{obs_i}\right|}{n} * 100. \tag{11}$$

Relationship among metrics *MAPEo*, *MAPEf* cannot be easily computed a posteriori, since they are un-linearly dependent on the parking Capacity. Also, R-squared (*R2*) is not linear and for this reason has not been taken into account in the comparison.

B. PREDICTION MODEL RESULTS

The comparison has been carried out by considering BRANN, RNN, CNN-GRU, CNN-LSTM and CNN-BI-LSTM on the set of car parks in Florence, a set composed by (a) *Careggi* car park (Hospital) having its capacity closed to 514 spaces; (b) *Beccaria* car park (market and downtown access) having its capacity closed to 203 spaces; (c) *S. Lorenzo* car park (historical center) having its capacity closed to 179 spaces. As a result, **Table 4** reports the comparison in terms of *MASE*, *MAE* and *RMSE* over the predicted week, considering a target of 24 hours in advance, for the number of free parking spaces applied to different locations. Moreover, a specific estimation for daily periods: morning (from 06:00:00 to 11:59:59), afternoon (from 12:00:00 to 17:59:59), evening (from 18:00:00 to 23:59:59) and night (from 00:00:00 to 05:59:59) is also considered in order to evaluate the daily moment having greater accuracy. The most critical day window is the **afternoon** to get access to restaurant areas and to pay a visit to the hospital.

The comparison of predictive models has been estimated on a training period of 3 months and the dataset is scaled with respect to the mean value. Error metrics have been estimated on a testing period of 1 week after the 23rd of May 2022 by considering the next 24 hours. More precisely, each input (test) set gives a prevision according to the next 96 timestamps (having 15 minutes of resolution) starting from the next hour related to the last time observation in the considered input set. Each input test set admits 1 week data observation sampled each hour (168 samples or timestamps) and two consecutive input tests are progressively defined. Given two consecutive input sets A, B then A is defined in the time interval $[t_0, \ldots, t_{167}]$ and B is defined in the time interval $[t_0 + 15', \ldots, t_{167} + 15']$. Then, each input test admits a prevision in different moment of the day. In order to evaluate the error according to a specific day moment, we are going to collect together previsions belonging to a specific time window according with morning, afternoon, evening and night.

The comparison has highlighted that CNN-BI-LSTM approach produced the most reliable results. Please note that metrics are reported according to the daily time slots, where the most critical one is the afternoon since in that slot some of the parking facilities may reach zero free slots.

The above-presented results are related to mid-term forecasts for the next 24 hours in terms of *MASE*, *MAEf*, *MSEf* and *RMSEf*. The computing of predictions for each hour for the next 24 hours reduces computational costs and energy. When data are very noisy, the results obtained with CNN-BI-LSTM are not the best ones, but still comparable with best results. In the critical daily moment, which is the afternoon, best results are obtained in most cases in terms of *MASE* and *MAEf* by CNN-BI-LSTM.

In **Table 5**, according to literature, we have assessed the results with respect to 1-hour horizon consisting in the first timestamp prevision of our output model, at 24 hours. Also in this case, CNN-BI-LSTM turned out to be the best model for almost all parking cases. Good results have been also obtained by CNN-GRU.

Furthermore, the studied models have been also trained for producing *1-hour* prevision, by modifying the number of units of the dense (last) layer, appropriately. So that, we specifically trained the models to produce *1-hour* prediction only. Results are reported in **Table 6**. In this case, the CNN-BI-LSTM turned out to be unequivocally the best model. Moreover, as expected, the estimation of a specifically trained model has produced better results than the model trained for 24hours predictions. On the other hand, differences are limited. CNN-BI-LSTM model obtained a *MAEf* equals to 14.8, 10.3 and 9.6 and *RMSEf* equals to 17.7, 12.3 and 11.7 for *Careggi*, *Beccaria* and *S. Lorenzo* car parks respectively (in the context of 1-hour forecast target).

According to the specific error estimation for daily periods (morning, afternoon, evening and night) we have obtained

TABLE 4. Comparison among predictive models for 24 hours in advance, every 15 minutes. Darker cells are the ones presenting better values. The metrics assessment has been performed over 1 week, with 15 minutes step.

Error Comparison	Forecasting Techniques				
	BRANN	RNN	CNN-GRU	CNN-LSTM	CNN-BI-LSTM
	Careggi car park (hospital)				
MASE Night	34.2	27.0	24.0	20.4	20.7
MASE Morning	8.6	7.2	5.3	6.0	4.7
MASE Afterno.	10.9	7.6	5.6	6.8	5.2
MASE Evening	10.1	6.4	7.4	6.2	6.0
MASE (total)	10.6	8.7	7.1	7.5	6.2
	Beccaria car park (downtown)				
MASE Night	8.6	12.1	14.0	8.5	11.9
MASE Morning	7.0	5.7	4.2	3.8	3.2
MASE Afterno.	10.2	8.2	7.7	7.8	6.8
MASE Evening	6.2	6.2	6.1	6.2	5.9
MASE (total)	8.2	7.6	6.9	6.3	6.0
	S. Lorenzo car park				
MASE Night	13.7	23.8	21.1	23.4	16.0
MASE Morning	9.2	6.1	5.7	5.8	6.0
MASE Afterno.	5.0	4.4	5.6	4.6	5.3
MASE Evening	3.8	4.0	3.1	3.7	2.9
MASE (total)	6.9	6.4	6.2	6.3	5.8

Error Comparison	Forecasting Techniques				
	BRANN	RNN	CNN-GRU	CNN-LSTM	CNN-BI-LSTM
	Careggi car park (hospital)				
MAEf Night	27.6	28.7	25.6	21.7	22.0
MAEf Morning	116.4	97.1	71.5	80.5	63.6
MAEf Afterno.	111.8	83.4	61.9	74.4	57.5
MAEf Evening	58.7	31.0	35.8	30.2	29.3
MAEf (total)	78.6	60.0	48.7	51.7	43.1
	Beccaria car park (downtown)				
MAEf Night	12.3	12.5	14.4	8.7	12.2
MAEf Morning	38.8	30.6	22.3	20.4	17.3
MAEf Afterno.	38.7	33.5	31.4	31.8	27.7
MAEf Evening	30.8	20.2	19.9	20.1	19.2
MAEf (total)	30.2	24.2	22.0	20.2	19.1
	S. Lorenzo car park				
MAEf Night	14.3	22.1	19.6	21.7	14.9
MAEf Morning	39.1	27.6	25.9	26.5	27.2
MAEf Afterno.	18.4	16.9	21.4	17.7	20.3
MAEf Evening	17.6	15.1	11.8	14.1	11.2
MAEf (total)	22.3	20.4	19.6	20.0	18.4

Error Comparison	Forecasting Techniques				
	BRANN	RNN	CNN-GRU	CNN-LSTM	CNN-BI-LSTM
	Careggi car park (hospital)				
RMSEf Night	38.2	36.1	32.5	28.3	27.4
RMSEf Morning	152.6	123.5	102.7	100.4	93.8
RMSEf Afterno.	134.1	103.0	78.9	94.6	86.0
RMSEf Evening	72.4	40.0	46.7	39.2	39.1
RMSEf (total)	109.5	84.8	70.7	73.1	68.0
	Beccaria car park (downtown)				
RMSEf Night	15.6	15.7	17.2	11.7	16.7
RMSEf Morning	45.7	39.6	29.0	28.5	23.7
RMSEf Afterno.	45.1	42.6	39.5	37.8	36.1
RMSEf Evening	38.6	25.8	25.0	24.9	24.1
RMSEf (total)	38.3	32.8	28.8	27.4	26.1
	S. Lorenzo car park				
RMSEf Night	17.2	24.7	11.5	23.6	17.8
RMSEf Morning	44.8	31.0	29.1	29.6	30.8
RMSEf Afterno.	23.1	21.0	26.6	23.2	25.8
RMSEf Evening	23.4	18.6	14.9	16.9	14.4
RMSEf (total)	29.1	24.3	23.6	23.7	23.1

TABLE 5. Comparison among predictive models for 1 hour in advance, assessing precision of the first hour of the 24 hours prevision target model.

Error Comparison	Forecasting Techniques				
	BRANN	RNN	CNN-GRU	CNN-LSTM	CNN-BI-LSTM
	Careggi car park				
MASE (1H/24)	10.6	6.6	7.4	5.9	5.5
MAEf (1H/24)	78.6	45.8	51.2	40.6	38.0
RMSEf (1H/24)	109.5	71.7	75.1	56.0	44.4
	Beccaria car park				
MASE (1H/24)	8.2	5.7	5.9	5.2	4.9
MAEf (1H/24)	30.2	18.5	19.3	17.1	16.0
RMSEf (1H/24)	38.3	25.2	24.3	23.0	20.6
	S. Lorenzo car park				
MASE (1H/24)	6.9	6.9	5.8	6.5	5.7
MAEf (1H/24)	22.3	22.3	18.7	21.1	18.3
RMSEf (1H/24)	29.1	26.0	22.6	24.9	23.0

TABLE 6. Comparison among predictive models for 1 hour in advance, adopting a specific trained model.

Error Comparison (1H)	Forecasting Techniques				
	BRANN	RNN	CNN-GRU	CNN-LSTM	CNN-BI-LSTM
	Careggi car park				
MASEf (1H)	10.6	2.4	1.8	2.2	1.1
MAEf (1H)	78.6	32.5	24.8	29.5	14.8
RMSEf (1H)	109.5	43.7	31.5	36.5	17.7
	Beccaria car park				
MASEf (1H)	8.2	3.3	2.7	3.3	2.0
MAEf (1H)	30.2	16.4	13.6	16.7	10.3
RMSEf (1H)	38.3	21.0	17.1	25.0	12.3
	S. Lorenzo car park				
MASEf (1H)	6.9	3.4	2.8	2.2	2.0
MAEf (1H)	22.3	16.6	13.7	10.8	9.6
RMSEf (1H)	29.1	19.3	16.5	15.1	11.7

Table 7 for *MAEf* estimations and **Table 8** for *MAEo* estimations, according to the relation between *MAEo* and *MAEf* reported in Section I-A.

TABLE 7. Daily ***MAEf*** for CNN-BI-LSTM models for 1 hour in advance, adopting a specific trained model.

Error Comparison	CNN-BI-LSTM (1 hour direct)		
	Careggi car park	Beccaria car park	San Lorenzo car park
MAEf Night	15.8	10.6	9.0
MAEf Morning	14.2	8.4	9.9
MAEf Afternoon	9.4	8.9	10.2
MAEf Evening	18.7	14.1	9.7
MAEf (daily)	14.8	10.3	9.6

As it can be noticed, best results have been registered for *Careggi* park during morning and afternoon daily periods, which are the most critical ones. *S. Lorenzo* provided very noisy trends, thus resulting in less predictable values.

C. LITERATURE COMPARISON DISCUSSION

As mentioned in the Introduction, the comparison of results with respect to literature has to be carefully performed. Since

TABLE 8. Daily ***MAEo*** for CNN-BI-LSTM models for 1 hour in advance, adopting a specific trained model.

Error Comparison	CNN-BI-LSTM (1 hour direct)		
	Careggi car park	Beccaria car park	San Lorenzo car park
MAEo Night	3.07	5.22	5.03
MAEo Morning	2.76	4.14	5.53
MAEo Afternoon	1.83	4.38	5.70
MAEo Evening	3.64	6.95	5.42
MAEo (daily)	2.88	5.07	5.36

in most cases we registered: (i) the usage of un-linear metrics or the usage of metrics depending on capacity without providing details, (ii) the usage of different parking data sets (different noise level and seasonality, see explainability hereafter), which in some cases also means to adopt additional variables such as weather, traffic, etc., (iii) the computation of assessment metrics as average in the day or week period, instead of providing them in the daily hours and thus in critical conditions.

Our target is related to the availability of free parking slots prediction in off-street parking which is formally different from occupancy prediction of free parking spaces in off-street parking. The most recent deep learning models have been applied to *1-hour* short-term parking prediction or shorter, which is very computationally expensive.

For these reasons, the performance comparison in terms of precision can be carried out only with respect to a limited number of state-of-the-art results. With the aim of comparing results, it should be noticed that [32], [33] and [Jelen] provided relevant values for error percentage at 60' with respect those presented in **Table 8** with our CNN-BI-LSTM, and in particular *MAEo* of 6.71 for [32], [33], and a *MAPEo* $>$ 6% for [27]. Moreover, according to **Table 8**, the proposed CNN-BI-LSTM model provides a *MAEo* of 1.83 for critical hours, thus overcoming the NN solution of [23] providing *MAEo* of 1.91, and those of [17]. In [17], as reported in **Table 1**, ConvLSTM-DCN model for the availability of free parking spaces prediction at *1-hour* in off-street context has been proposed. Thus, applied on not very noisy car parks, it could obtain a larger absolute error as *MAEf*, and comparable results normalizing its results computing *MAEo* with the information provided.

On the other hand, for mid-term predictions of 24-hours, the literature does not provide specific results, except for the analysis of [32], [33] which detected a strong increment of error in increasing the prediction time. As it can be easily obtained by Table 4 and the definition of *MAEo*, the results for H24 prediction are: a *MAEo* of 8.39 for *Careggi* park for the whole day, which is good compromise to avoid computing 24 or 96 estimations, over 24 hours.

D. COMPUTATIONAL PERFORMANCE ANALYSIS

In **Table 9**, the comparison of execution time performance of the used techniques in the case of 24 hours predictions models of **Table 4** is reported. The case of study in Beccaria car park is considered in the mentioned Table. Deep learning solutions have been executed on GPU as NVIDIA Quadro GV100 with 32GByte Ram, which has 5120 CUDA Cores, FP64 perf as 7.4 TFLOPS.

Note that, CNN-BI-LSTM model provides a training execution time shorter than the CNN-LSTM model due to the used training strategy involving the Early Stopping method with patience parameter set to 60 to determine the optimum epoch number. In that case, CNN-BI-LSTM model admits the optimum epoch number equal to 188 with a training execution time of 1162.32s, while CNN-LSTM model admits the optimum epoch number equal to 378 with a training execution time of 989.95s. Then, for a single epoch CNN-BI-LSTM model is more time consuming than CNN-LSTM, as already seen in the results related to the prediction execution.

TABLE 9. Comparison of performance of the deep learning techniques in the context of parking prediction.

Model	Training execution (s)	Prediction execution (s), 24 hours, 96 predictions
BRANN	137.66	0.032
RNN	327.41	0.836
CNN-GRU	523.56 (130 epochs)	5.191
CNN-LSTM	1162.32 (378 epochs)	3.926
CNN-BI-LSTM	989.95 (188 epochs)	7.386

V. PREDICTION MODEL EXPLAINABILITY

The level of relevance of features and time lags for the multi-step prediction can be analyzed by using the gradient and integrated gradient technique [63]. Gradient, denoted by G, is a fundamental concept in machine learning to identify both direction and magnitude of the maximum growth of a function at a specific point. To calculate the Gradients and the Integrated Gradients, we created x_{test} dataset consisting of 672 arrays of timeseries (from data of Table 2) starting from Sunday to Saturday. Each array includes a timeseries of 1-week data observations sampled each hour (168 samples / timestamps).

Two consecutive arrays A, B in the x_{test} set of arrays, are translated of 15 min. Then, A is defined in the time interval $[t_0, \ldots, t_{167}]$ and B is defined in the time interval $[t_0 + 15', \ldots, t_{167}+15']$. Let Γ be the set of features such that $\Gamma = \{POD, freeParkSlots, SOD, averageVehicleSpeed, vehicleFlow, averageVehicleTime, vehicleConcentration, temperature, humidity, dayWeek\}$ from **Table 2**, then we are going to consider the $j-$th feature X^j in Γ and $X^j_{t,i}$ is the related $i-$th observation in the $t-$th week of the $j-$th feature, with $0 \leq t \leq 671$, $0 \leq i \leq 167$ and $0 \leq j \leq 9$. Thus, x_{test} can be represented as follows:

$$x_{test} = \bigcup\nolimits_{j=0,\ldots,9} \begin{bmatrix} X^j_{0,0} & \cdots & X^j_{0,167} \\ \vdots & \ddots & \vdots \\ X^j_{671,0} & \cdots & X^j_{671,167} \end{bmatrix}. \quad (12)$$

The visualization of the gradient helps us to understand how the steps of the input features $X^j_{t,i}$ affect the output of the model at each timestep i. In a nutshell, the gradient provides a representation of the input areas that are important for the model in predicting.

Formally, for each $X^j \in \Gamma$, we have:

$$G^j{}_{t,i} = \frac{\partial F}{\partial X^j}\left(X^0_{t,i}, \ldots, X^9_{t,i}\right), \quad (13)$$

where F is the output function and $G^j{}_{t,i}$ is the related $i-$th gradient in the $t-$th week of the $j-$th feature, with $0 \leq t \leq 671$, $0 \leq i \leq 167$ and $0 \leq j \leq 9$.

Figure 3 shows the gradient map for features, while considering CNN-BI-LSTM and CNN-LSTM models, respectively. The point having coordinates (t, i) in the $j-$th feature represents the $i-$th Gradient value according to $t-$th week. The steps of data input which influence more positively the prediction are reported in green, in red are those that negatively influence the prediction and in white are the steps of input having a low influence on the prediction output.

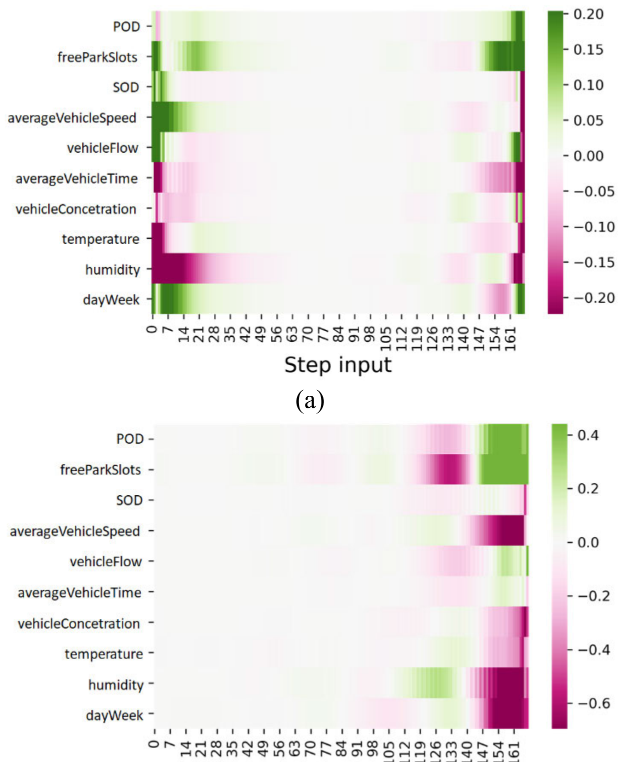


FIGURE 3. Gradient for features for the (a) CNN-BI-LSTM and (b) CNN-LSTM models. In green, red and white are reported the steps influencing positively, negatively and marginally the predictions, respectively. (Careggi Car Park).

According to **Figure 3(a)** with bidirectional layer, the features having the greatest influence at a higher value of the prediction are *freeParkSlots*, *averageVehicleSpeed, vehicleflow* and *dayWeek*. While features affecting lower values of prediction are: *humidity*, *temperature* and *averageVehicleTime*. *Humidity* and *freeParkSlots* are the features for which a larger number of steps affected the prediction of the x_{test} considered. The analysis of **Figure 3(b)** without the bidirectional layer, highlighted that predictions are obtained in a quite different manner, if considering positively *POD* and f*reeParkSlots*, and negatively Humidity and *DayWeek*. Please note that also scales are different in the two graphs of **Figure 3**. In **Figure 3(a)**, the first 30 steps and the last 45 steps of the input have greater importance in the forecast than the middle steps. This is related to the bidirectional approach of taking data. In fact, **Figure 3(b)**, presents relevant features only in the last hours. Thus, the white central area the **Figure 3(a)** graph could be interpreted as follows: data older than 45 hours could be less relevant for the models.

Figure 4 shows the normalized cumulated gradients for all features, as a function of hours slots for the CNN-BI-LSTM and CNN-LSTM models, respectively. On the x-axis, the $i-$th step of the input week hours is shown and on the y-axis the value of the normalized cumulated gradient.

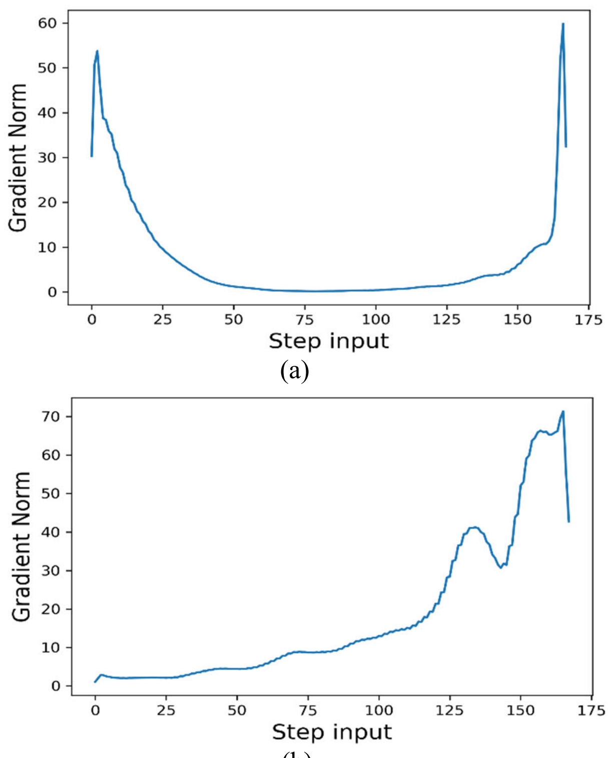


FIGURE 4. Normalized cumulated gradient plot for the CNN-BI-LSTM and CNN-LSTM models, from 1 to 168 samples, Careggi car park.

Again, a different behavior is shown for bidirectional and non-bidirectional models. In fact, as to **Figure 4(b)**, the input steps most influential to the prediction are the ones closest to the prediction, thus the last 75 steps of the input week. These steps are much more influential than in the bidirectional model, because they generate larger gradients. Thus, the structure of CNN-BI-LSTM model allows us to capture information on the weekly seasonality of the amount of parking predicted and may need a shorter time window data to provide predictions.

A. ANALYSIS VIA INTEGRATED GRADIENTS

By estimating the Integrated Gradients, *IG* [64], is possible to determine the importance of individual input features for each time step of the output/prediction. The IG estimates the gradient of the output/prediction with respect to the interpolated input/features. Thus, features playing a relevant role in certain prediction sequences are identified. The idea behind *IG* is to calculate the weighted average of the gradients of the output/prediction function with respect to the input/features using a reference baseline.

$$IG^{j}{}_{t,i} = \left(\left(X^0_{t,i}, \ldots, X^9_{t,i}\right) - \left(b^0_{t,i}, \ldots, b^9_{t,i}\right)\right) \times \int \frac{\partial F}{\partial X^j}\left(\left(b^0_{t,i}, \ldots, b^9_{t,i}\right) + \alpha\left(\left(X^0_{t,i}, \ldots, X^9_{t,i}\right) - \left(b^0_{t,i}, \ldots, b^9_{t,i}\right)\right)\right) d\alpha, \tag{14}$$

where: b is the baseline, F is the output function and α is a scale value between 0 and 1 that allows interpolation between the baseline and the input. **Figure 5** shows the *IG* maps for the features and the corresponding time trend, in a week from Sunday to Saturday (right). The prediction is for the next Sunday.

Figure 5 reports the typical *IG* heatmaps computed for the CNN-BI-LSMT model on *Careggi* car park. On the other hand, it can be applied as a generic XAI approach for almost all ML and AI models in the literature.

For each $j-$th map, on the x-axis there are the observations considered in our test dataset, and on the y-axis there are the predictions considering 1-day observation with time lags of 15 minutes (96 timestamps, 24×4 time slots of 15'). In green are the steps of the output model that are most positively influenced in the prediction by means of the $j-$th feature, in red are the steps that are negatively influenced and in white the steps that are slightly influenced.

From the *IG* heatmaps, we can make several considerations. Please note that, the heatmaps of the features do not have all the same scale; *Humidity* and *freeParkSlots* are those with large magnitude. *Humidity* influences the prediction steps from step 25 up to step 50.

From almost all heatmaps, what is self-evident is the presence of a week seasonality and a difference in relevance with a daily seasonality (along y). A different week seasonality has been registered for *freeParkSlots* with respect to *vehicleConcentration* and *vehicleFlow*, which are data coming from traffic flow sensors. Moreover, it is also confirmed that the predictions on Sunday are based on the values of previous Saturday, Friday, and Sunday (see *averageVehicleTime*). The *averageVehicleTime* feature has a greater influence on the first part of the forecast series. *Temperature* always affects negatively the predictions, with more relevance in the central part of the day. Features such as *vehicleFlow*, *average VehicleSpeed*, *averageVehicleTime*, *vehicleConcentration* and *temperature* have a greater influence in the prediction from step 40 to step 65.

This approach has been adopted for the proposed deep learning solution CNN-BI-LSMT and could be applied for

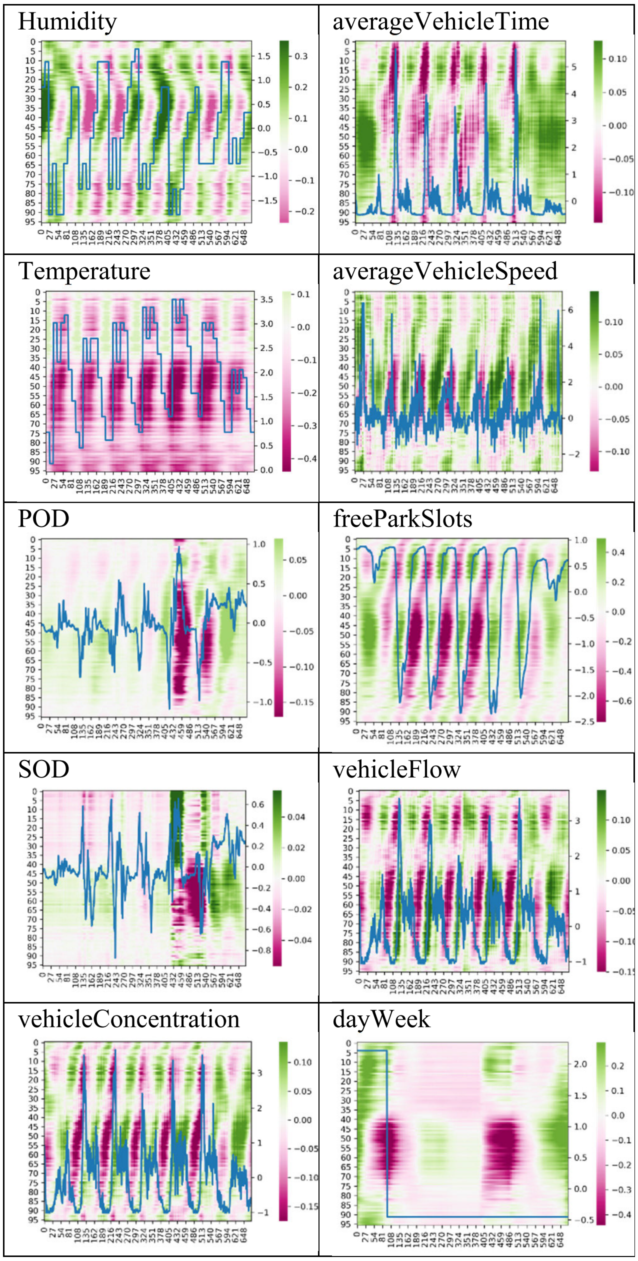


FIGURE 5. Integrated Gradient for predictions with respect to the features, for *Careggi* Car Park. In green, red and white are reported the steps influencing positively, negatively and marginally the predictions, respectively. In blue is the time trend of the feature.

almost all ML and AI models in literature, with some differences in the interpretations as occurred in the gradient.

VI. CONCLUSION

The Availability off-street parking spaces prediction is a complex non-linear process involving multiple kinds of factors, as the variety of parking areas (downtown, on hospital and others on the periphery, close to theaters, airports, etc.).

The present work has been focused to find a solution for predicting the number of available parking slots in off-street case in the city of Florence for the next 24 hours in advance, every 15 minutes. Such a mid-term prevision could be very useful for drivers to plan their travel and parking the day before. Moreover, the 24H predictions allow to drastically reduce the computational costs. To this aim, we have considered and compared a number of different techniques and metrics to assessment them. We discovered in the literature two main approaches for parching prediction assessment. One based on metrics for assessing *occupancy rate* as percentage of fullness, and the other of absolute measure of the error in *free slot* estimation. The assessment approaches have been analyzed and compared to produce a framework which allows to compare the results obtained and measured according to these different methods.

According to the analysis a number of techniques have been compared: BRANN, RNN, CNN-GRU, CNN-LSTM, and CNN-BI-LSTM to identify the best solution in terms of precision, especially for the estimation of free slot in critical conditions (when the free slot risk to become zero). The solution identified results the better ranked in these conditions 1-hour in advance, and in producing prediction 24hours in advance (mid-terms). The comparison has been performed considering several metrics according to the two different approaches: occupancy rate and free slots.

In addition, the paper has performed a feature relevance analysis to identify the most relevant features and their impact over time since some of them provide different seasonality, and also the predictions are affected by seasonality over the week and the 24 hours. To this end, features such as the historical data, the weather conditions and the traffic flow data have been exploited and analyzed. In almost all predictive models, the historical data, traffic flow sensors and weather have demonstrated high predictive capabilities in explaining the number of free parking slots. The research documented in this paper demonstrated by using the gradient the differences from CNN-LSTM and CNN-BI-LSTM. And by using the integrated gradient and a new heatmap representations impact of seasonality in the parking predictions. This approach can be used for almost all ML and AI models in the literature.

The prediction model proposed has been created by exploiting data in the Snap4city platform and infrastructure in Florence and Tuscany area, Italy.

ACKNOWLEDGMENT

The authors would like to thank the many developers of snap4city platform. Snap4City (https://www.snap4city.org) is an infrastructure and an open platform and technology of the DISIT Laboratory, UNIFI.

REFERENCES

[1] T. Afrin and N. Yodo, "A survey of road traffic congestion measures towards a sustainable and resilient transportation system," *Sustainability*, vol. 12, no. 11, p. 4660, Jun. 2020.

[2] S. Bilotta and P. Nesi, "Traffic flow reconstruction by solving indeterminacy on traffic distribution at junctions," *Future Gener. Comput. Syst.*, vol. 114, pp. 649–660, Jan. 2021.

[3] R. Smit, L. Ntziachristos, and P. Boulter, "Validation of road vehicle and traffic emission models—A review and meta-analysis," *Atmos. Environ.*, vol. 44, no. 25, pp. 2943–2953, Aug. 2010.

[4] S. Bilotta and P. Nesi, ‘‘Estimating CO_2 emissions from IoT traffic flow sensors and reconstruction,’’ *Sensors*, vol. 22, no. 9, p. 3382, Apr. 2022.

[5] C. Badii, S. Bilotta, D. Cenni, A. Difino, P. Nesi, I. Paoli, and M. Paolucci, ‘‘High density real-time air quality derived services from IoT networks,’’ *Sensors*, vol. 20, no. 18, p. 5435, Sep. 2020.

[6] D. C. Shoup, ‘‘Cruising for parking,’’ *Transp. Policy*, vol. 13, no. 6, pp. 479–486, Nov. 2006.

[7] M. C. Bell and F. Galatioto, ‘‘Simulation of illegal double parking: Quantifying the traffic and pollutant impacts,’’ in *Proc. 4th Int. SIIV Conf.*, 2007, pp. 12–14.

[8] C. Tiexin, T. Miaomiao, and M. Ze, ‘‘The model of parking demand forecast for the urban CCD,’’ *Energy Proc.*, vol. 16, pp. 1393–1400, Jan. 2012.

[9] M. Aljohani, S. Olariu, A. Alali, and S. Jain, ‘‘A survey of parking solutions for smart cities,’’ *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 8, pp. 10012–10029, Aug. 2022.

[10] S. An, B. Han, and J. Wang, ‘‘Study of the mode of real-time and dynamic parking guidance and information systems based on fuzzy clustering analysis,’’ in *Proc. 3rd Conf. Mach. Learn. Cybern.*, Shanghai, China, 2004, pp. 2790–2794.

[11] W. H. K. Lam, Z.-C. Li, H.-J. Huang, and S. C. Wong, ‘‘Modeling time-dependent travel choice problems in road networks with multiple user classes and multiple parking facilities,’’ *Transp. Res. B, Methodol.*, vol. 40, no. 5, pp. 368–395, Jun. 2006.

[12] Z. S. Qian and R. Rajagopal, ‘‘Optimal parking pricing in general networks with provision of occupancy information,’’ *Proc., Social Behav. Sci.*, vol. 80, pp. 779–805, Jun. 2013.

[13] E. S. Fokker, T. Koch, M. van Leeuwen, and E. R. Dugundji, ‘‘Short-term forecasting of off-street parking occupancy,’’ *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2676, no. 1, pp. 637–654, Jan. 2022.

[14] C. Badii, P. Nesi, and I. Paoli, ‘‘Predicting available parking slots on critical and regular services by exploiting a range of open data,’’ *IEEE Access*, vol. 6, pp. 44059–44071, 2018.

[15] P. M. Calvo, B. Schotten, and E. R. Dugundji, ‘‘Assessing the predictive value of traffic count data in the imputation of on-street parking occupancy in Amsterdam,’’ *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2675, no. 12, pp. 330–341, Dec. 2021.

[16] C. Zeng, C. Ma, K. Wang, and Z. Cui, ‘‘Predicting vacant parking space availability: A DWT-Bi-LSTM model,’’ *Phys. A, Stat. Mech. Appl.*, vol. 599, Aug. 2022, Art. no. 127498.

[17] Y. Feng, Y. Xu, Q. Hu, S. Krishnamoorthy, and Z. Tang, ‘‘Predicting vacant parking space availability zone-wisely: A hybrid deep learning approach,’’ *Complex Intell. Syst.*, vol. 8, no. 5, pp. 4145–4161, Oct. 2022.

[18] R. K. Kasera and T. Acharjee, ‘‘Parking slot occupancy prediction using LSTM,’’ *Innov. Syst. Softw. Eng.*, pp. 1–13, Sep. 2022, doi: 10.1007/s11334-022-00481-3.

[19] F. R. Burden and D. A. Winkler, ‘‘Robust QSAR models using Bayesian regularized neural networks,’’ *J. Med. Chem.*, vol. 42, no. 16, pp. 3183–3187, Aug. 1999.

[20] H. Drucker, C. J. C. Burges, L. Kaufman, A. Smola, and V. Vapnik, ‘‘Support vector regression machines,’’ in *Advances in Neural Information Processing Systems*, vol. 9, M. Mozer, M. Jordan, and T. Petsche, Eds. Cambridge, MA, USA: MIT Press, 1997, pp. 155–161.

[21] J. T. Connor, R. D. Martin, and L. E. Atlas, ‘‘Recurrent neural networks and robust time series prediction,’’ *IEEE Trans. Neural Netw.*, vol. 5, no. 2, pp. 240–254, Mar. 1994.

[22] G. E. P. Box, G. M. Jenkins, and G. C. Reinsel, *Time Series Analysis: Forecasting and Control*, 3rd ed. Englewood Cliffs, NJ, USA: Prentice-Hall, 1994.

[23] J. C. Provoost, A. Kamilaris, L. J. J. Wismans, S. J. van der Drift, and M. van Keulen, ‘‘Predicting parking occupancy via machine learning in the web of things,’’ *Internet Things*, vol. 12, Dec. 2020, Art. no. 100301.

[24] Y. LeCun and Y. Bengio, ‘‘Convolutional networks for images, speech, and time series,’’ in *The Handbook of Brain Theory and Neural Networks*, vol. 3361, no. 10. Cambridge, MA, USA: MIT Press, 1995, p. 1995.

[25] J. Schmidhuber, ‘‘Deep learning in neural networks: An overview,’’ *Neural Netw.*, vol. 61, pp. 85–117, Jan. 2015.

[26] L. Breiman, ‘‘Random forests,’’ *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.

[27] G. Jelen, V. Podobnik, and J. Babic, ‘‘Contextual prediction of parking spot availability: A step towards sustainable parking,’’ *J. Cleaner Prod.*, vol. 312, Aug. 2021, Art. no. 127684.

[28] C. Lucchese, G. Callegher, M. Modenese, and S. Dassiè, ‘‘A comparison of spatio-temporal prediction methods: A parking availability case study,’’ in *Proc. 37th ACM/SIGAPP Symp. Appl. Comput.*, Apr. 2022, pp. 1013–1020.

[29] T. Rajabioun and P. A. Ioannou, ‘‘On-street and off-street parking availability prediction using multivariate spatiotemporal models,’’ *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 5, pp. 2913–2924, Oct. 2015.

[30] K. Cho, B. van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, ‘‘Learning phrase representations using RNN encoder–decoder for statistical machine translation,’’ 2014, *arXiv:1406.1078*.

[31] Y. Seo, M. Defferrard, P. Vandergheynst, and X. Bresson, ‘‘Structured sequence modeling with graph convolutional recurrent networks,’’ in *Neural Information Processing: 25th International Conference, ICONIP 2018, Siem Reap, Cambodia, December 13–16, 2018, Proceedings, Part I 25*. Springer, 2018, pp. 362–373.

[32] F. A. Gers, D. Eck, and J. Schmidhuber, ‘‘Applying LSTM to time series predictable through time-window approaches,’’ in *Proc. Int. Conf. Artif. Neural Netw.* Berlin, Germany: Springer, Aug. 2001, pp. 669–676.

[33] M. K. Mufida, A. A. E. Cadi, T. Delot, and M. Trépanier, ‘‘Towards a continuous forecasting mechanism of parking occupancy in urban environments,’’ in *Proc. 25th Int. Database Eng. Appl. Symp.*, Jul. 2021, pp. 263–272.

[34] S. Hochreiter and J. Schmidhuber, ‘‘Long short-term memory,’’ *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997.

[35] Z. Zhao, Y. Zhang, and Y. Zhang, ‘‘A comparative study of parking occupancy prediction methods considering parking type and parking scale,’’ *J. Adv. Transp.*, vol. 2020, Feb. 2020, Art. no. 5624586.

[36] F. R. Alharbi and D. Csala, ‘‘A seasonal autoregressive integrated moving average with exogenous factors (SARIMAX) forecasting model-based time series approach,’’ *Inventions*, vol. 7, no. 4, p. 94, Oct. 2022.

[37] S. Yang, W. Ma, X. Pi, and S. Qian, ‘‘A deep learning approach to real-time parking occupancy prediction in transportation networks incorporating multiple spatio-temporal data sources,’’ *Transp. Res. C, Emerg. Technol.*, vol. 107, pp. 248–265, Oct. 2019.

[38] L. F. He, X. B. Wang, Z. X. He, and C. F. Li, ‘‘Wavelet-based denoising of MT time series,’’ *Seismol. Geol.*, vol. 23, no. 2, pp. 222–226, 2001.

[39] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, ‘‘Densely connected convolutional networks,’’ in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 2261–2269.

[40] M. Cools, E. Moons, and G. Wets, ‘‘Investigating the variability in daily traffic counts through use of ARIMAX and SARIMAX models: Assessing the effect of holidays on two site locations,’’ *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2136, no. 1, pp. 57–66, Jan. 2009.

[41] G. Jain and B. Mallick, ‘‘A study of time series models ARIMA and ETS,’’ Available SSRN 2898968, 2017.

[42] C. Badii, E. G. Belay, P. Bellini, D. Cenni, M. Marazzini, M. Mesiti, P. Nesi, G. Pantaleo, M. Paolucci, S. Valtolina, M. Soderi, and I. Zaza, ‘‘Snap4City: A scalable IoT/IOE platform for developing smart city applications,’’ in *Proc. IEEE SmartWorld, Ubiquitous Intell. Comput., Adv. Trusted Comput., Scalable Comput. Commun., Cloud Big Data Comput., Internet People Smart City Innov. (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)*, Oct. 2018, pp. 2109–2116.

[43] P. Bellini, S. Bilotta, A. L. I. Palesi, P. Nesi, and G. Pantaleo, ‘‘Vehicular traffic flow reconstruction analysis to mitigate scenarios with large city changes,’’ *IEEE Access*, vol. 10, pp. 131061–131075, 2022.

[44] P. Bellini, S. Bilotta, P. Nesi, M. Paolucci, and M. Soderi, ‘‘Real-time traffic estimation of unmonitored roads,’’ in *Proc. IEEE 16th Int. Conf. Dependable, Autonomic Sec. Comput., 16th Int. Conf. Pervasive Intell. Comput., 4th Int. Conf Big Data Intell. Comput. Cyber Sci. Technol. Congr. (DASC/PiCom/DataCom/CyberSciTech)*, Athens, Greece, Aug. 2018, pp. 935–942.

[45] S. Bilotta, P. Nesi, and I. Paoli, ‘‘Real-time system for short- and long-term prediction of vehicle flow,’’ in *Proc. IEEE Int. Conf. Smart Data Services (SMDS)*, Oct. 2020, pp. 97–104.

[46] B. Yang, J. Dai, C. Guo, C. S. Jensen, and J. Hu, ‘‘PACE: A path-centric paradigm for stochastic path finding,’’ *VLDB J.*, vol. 27, no. 2, pp. 153–178, Apr. 2018, doi: 10.1007/s00778-017-0491-4.

[47] J. Hu, B. Yang, C. Guo, and C. S. Jensen, ‘‘Risk-aware path selection with time-varying, uncertain travel costs: A time series approach,’’ *VLDB J.*, vol. 27, no. 2, pp. 179–200, Apr. 2018, doi: 10.1007/s00778-018-0494-9.

[48] J. Dai, B. Yang, C. Guo, C. S. Jensen, and J. Hu, ‘‘Path cost distribution estimation using trajectory data,’’ *Proc. VLDB Endowment*, vol. 10, no. 3, pp. 85–96, Nov. 2016, doi: 10.14778/3021924.3021926.

[49] S. Bilotta, E. Collini, P. Nesi, and G. Pantaleo, "Short-term prediction of city traffic flow via convolutional deep learning," *IEEE Access*, vol. 10, pp. 113086–113099, 2022.
[50] Z. Cui, R. Ke, Z. Pu, and Y. Wang, "Stacked bidirectional and unidirectional LSTM recurrent neural network for forecasting network-wide traffic state with missing values," *Transp. Res. C, Emerg. Technol.*, vol. 118, Sep. 2020, Art. no. 102674.
[51] O. Troyanskaya, M. Cantor, G. Sherlock, P. Brown, T. Hastie, R. Tibshirani, D. Botstein, and R. B. Altman, "Missing value estimation methods for DNA microarrays," *Bioinformatics*, vol. 17, no. 6, pp. 520–525, Jun. 2001.
[52] C. Bishop, *Neural Networks for Pattern Recognition*. Oxford, U.K.: Oxford Univ. Press, 1995.
[53] B. D. Ripley, "Statistical ideas for selecting network architectures," in *Neural Networks: Artificial Intelligence and Industrial Applications: Proceedings of the Third Annual SNN Symposium on Neural Networks, Nijmegen, The Netherlands, 14–15 September 1995*. London, U.K.: Springer, 1995, pp. 183–190.
[54] M. Titterington, "Neural networks," *Wiley Interdiscipl. Rev., Comput. Statist.*, vol. 2, no. 1, pp. 1–8, 2010.
[55] D. J. C. MacKay, "A practical Bayesian framework for backpropagation networks," *Neural Comput.*, vol. 4, no. 3, pp. 448–472, May 1992.
[56] D. Nguyen and B. Widrow, "Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights," in *Proc. Int. Joint Conf. Neural Netw.*, Jun. 1990, pp. 21–26.
[57] R. Pascanu, T. Mikolov, and Y. Bengio, "On the difficulty of training recurrent neural networks," in *Proc. Int. Conf. Mach. Learn.*, 2013, pp. 1310–1318.
[58] T. Li, M. Hua, and X. Wu, "A hybrid CNN-LSTM model for forecasting particulate matter ($PM_{2.5}$)," *IEEE Access*, vol. 8, pp. 26933–26940, 2020.
[59] I. E. Livieris, E. Pintelas, and P. Pintelas, "A CNN–LSTM model for gold price time-series forecasting," *Neural Comput. Appl.*, vol. 32, no. 23, pp. 17351–17360, Dec. 2020.
[60] T.-Y. Kim and S.-B. Cho, "Predicting residential energy consumption using CNN-LSTM neural networks," *Energy*, vol. 182, pp. 72–81, Sep. 2019.
[61] K. Cho, B. van Merriënboer, D. Bahdanau, and Y. Bengio, "On the properties of neural machine translation: Encoder–decoder approaches," in *Proc. 8th Workshop Syntax, Semantics Struct. Stat. Transl. (SSST)*, 2014, pp. 103–111.
[62] R. J. Hyndman and A. B. Koehler, "Another look at measures of forecast accuracy," *Int. J. Forecasting*, vol. 22, no. 4, pp. 679–688, Oct. 2006.
[63] H. Choi, C. Jung, T. Kang, H. J. Kim, and I.-Y. Kwak, "Explainable time-series prediction using a residual network and gradient-based methods," *IEEE Access*, vol. 10, pp. 108469–108482, 2022.
[64] M. Sundararajan, A. Taly, and Q. Yan, "Axiomatic attribution for deep networks," in *Proc. 34th Int. Conf. Mach. Learn. (ICML)*, vol. 70, Sydney, NSW, Australia, 2017, pp. 3319–3328.
[65] J. T. Hancock and T. M. Khoshgoftaar, "CatBoost for big data: An interdisciplinary review," *J. Big Data*, vol. 7, no. 1, pp. 1–45, Dec. 2020.
[66] F. D. Foresee and M. T. Hagan, "Gauss–Newton approximation to Bayesian learning," in *Proc. Int. Conf. Neural Netw.*, vol. 3, 1997, pp. 1930–1935.

STEFANO BILOTTA is currently a Researcher with the DIMAI Department, University of Florence, and the DISIT Laboratory collaborator with the DINFO Department. He has been involved in many projects, such as Sii-Mobility, Trafair, and Snap4city. His current research interests include traffic flow reconstruction algorithms, mobility, parallel solutions, dynamic systems, machine learning, languages, and coding theory.

LUCIANO ALESSANDRO IPSARO PALESI is currently pursuing the Ph.D.Eng. degree with the DISIT Laboratory, DINFO Department, University of Florence. He has been involved in many projects, such as Sii-Mobility, Feedback, and Snap4City. His current research interests include deep learning, mobility, and data models.

PAOLO NESI (Member, IEEE) is currently a Full Professor with the DINFO Department, University of Florence, and the Chief of the DISIT Laboratory. He has been the coordinator of several research and development multi-partner international research and development projects. He has published more than 400 papers in international journals and conferences. His current research interests include machine learning, massive parallel and distributed systems, physical models, the IoT, mobility, big data analytics, AI/XAI, semantic computing, formal models, machine learning, and data privacy. He has been the chair of several international conferences.

Open Access funding provided by 'Università degli Studi di Firenze' within the CRUI CARE Agreement