Proceedings of the Sixth International Conference on
Railway Technology: Research, Development and Maintenance
Edited by: J. Pombo
Civil-Comp Conferences, Volume 7, Paper 10.9
Civil-Comp Press, Edinburgh, United Kingdom, 2024
ISSN: 2753-3239, doi: 10.4203/ccc.7.10.9

An Examination of Machine Learning Methods for Predicting Station Departure Delay Time Based on Historical Train Traffic Records

T. Fukuda, S. Takahashi and H. Nakamura

College of Science and Technology, Nihon University, Japan

Abstract

In the urban railways of the metropolitan area, train delays during morning rush hours have become a significant issue. This problem isn't merely caused by large, sudden delays, but also by an accumulation of minor delays, which can escalate into substantial setbacks. This presents a disadvantage not only for the operators managing the train services but also for the users. Recent years have seen active reporting of studies predicting delay times. However, these studies do not take into account the accumulative nature of delays because the data input to the learning machine is only for stations behind the target station. Therefore, this paper aims to predict the delay time that occurs when each train departs from a station. To do so, we constructed a network that takes into account the propagation of delays by using data from stations in front of the target station as input. We will provide an overview of the network we constructed and report on its predictive accuracy.

Keywords: train, machine learning, simulator, delay countermeasures, train scheduling, robustness.

1 Introduction

One of the issues facing urban rail transit in the metropolitan area is the delays during the morning commuting hours. The number of passengers commuting or going to school is the highest of the day from around 6 to 8 in the morning, during which various measures are taken by operators to increase transport capacity. Among these, reducing the train headway is a crucial measure that directly contributes to enhancing transport capacity and significantly affects passenger convenience. However, while decreasing train headway improves transport capacity, it also exacerbates the impact of delays.

Many conventional rail lines operate under a system known as fixed block signaling, where only one train is allowed to run within a designated section of the track to ensure safety. When the trailing train gets too close to the leading train, reducing the train interval, a mechanism is employed to limit the speed of the trailing train to maintain safety.

When attempting to enhance transport capacity under fixed block signaling by reducing the intervals between train operations, the impact on the following train is significantly larger if, for example, the leading train is delayed when leaving a station, compared to when the intervals are wider. As a result, delays occurring during the morning commuting hours can take a long time to resolve. To alleviate these increased delays, measures such as reducing train services or adjusting train intervals are implemented, leading to a decrease in transport capacity and customer satisfaction. Therefore, achieving operations with fewer delays is strongly desired to enable high-capacity transport.

On the other hand, when implementing measures against delays, a method that uses simulators to check the impact in advance is employed [1–3]. When simulating train operations with existing train operation simulators, basic data such as route data and vehicle data, along with the stop times at stations, are inputted for the simulation. Generally, the stop times at stations that are inputted are the actual data from when trains were operated (hereafter referred to as actual stop times). Therefore, as long as the actual stop times, which do not incorporate changes due to delay measures, are used, there were limitations in discussing the effectiveness of the delay measures even if changes occur in train operations due to those measures.

Therefore, the authors believed that if it becomes possible to predict the changes in train operation times, such as those due to the addition of platforms or changes in congestion rates, using machine learning [4–6], the effects of delay countermeasures could be verified. In this paper, with the future application to verify the effectiveness of delay countermeasures in mind, a predictive method using Convolutional Neural Networks (CNN) was proposed, aiming initially to improve the accuracy of delay estimation in the current timetable.

2 Analysis of the historical train traffic records

In this chapter, we describe the historical train traffic records and discuss the characteristics of delays that can be identified through the analysis of these records.

2.1 Historical train traffic records

Historical train traffic records are documents that record all train movements on a given day of operation, including the ”train name, train type, arrival times at each station, departure times from each station, and arrival track number.”

The historical train traffic records used in this paper are estimated using information from the activation circuits closest to the station. In railways, the arrival and departure times at stations are defined as the times when a train stops and starts moving at the station, respectively. However, because the train's speed is not captured in real-time, these times are calculated using activation circuits. An activation circuit is a type of sensor system installed on railway tracks to detect the presence and location of a train. By dividing the tracks into sections and flowing electric current through each, they function as electrical circuits. When a train enters a section, it short-circuits the circuit between the rails with its axles, allowing the system to detect the train's presence. This system enables the detection of the times a train enters or exits a section, allowing for the calculation of the train's arrival and departure times using the following equation:

$$\text{Train's arrival and departure time} = \text{Activation circuit's drop time} + t \tag{1}$$

Because the data is calculated based on information from the activation circuits before and after the station, the arrival and departure times are recorded in seconds, resulting in data with high temporal resolution. In this paper, these data are compared with the train operation timetable to calculate the arrival delay time, departure delay time, and travel time between stations at each station, which are then used as training data. The calculated data are shown below, all in seconds.

1. Arrival delay = Actual arrival time - Scheduled arrival time
2. Departure delay = Actual departure time - Scheduled departure time
3. Station occurrence delay = Departure delay time - Arrival delay time
4. Station Stopping Time = Departure time - Arrival time
5. Station travel time = Station arrival time - Previous station departure time
6. Station travel delay = Arrival delay time - Departure delay time at the previous station
7. Arrival interval = Actual arrival time - Departure time of the preceding train

2.2 The characteristics of train delays

Delays can be divided into three states based on the time of day. The graph shown in Figure 1 is used to check the increasing trend of delays at a particular station over a month, with the vertical axis representing departure delay time and the horizontal axis representing time. The data's variability is represented with a box-and-whisker plot, and the data points indicated with circles are considered noise, such as sudden delays or adjustments in train intervals. From Figure 1, it can be seen that the transition of delays includes:

1. A time period where initial delays occur and delay increases
2. A time period where the delay remains constant after increasing
3. A time period where the increased delay starts to decrease

These three time periods help in understanding the behavior of delays.

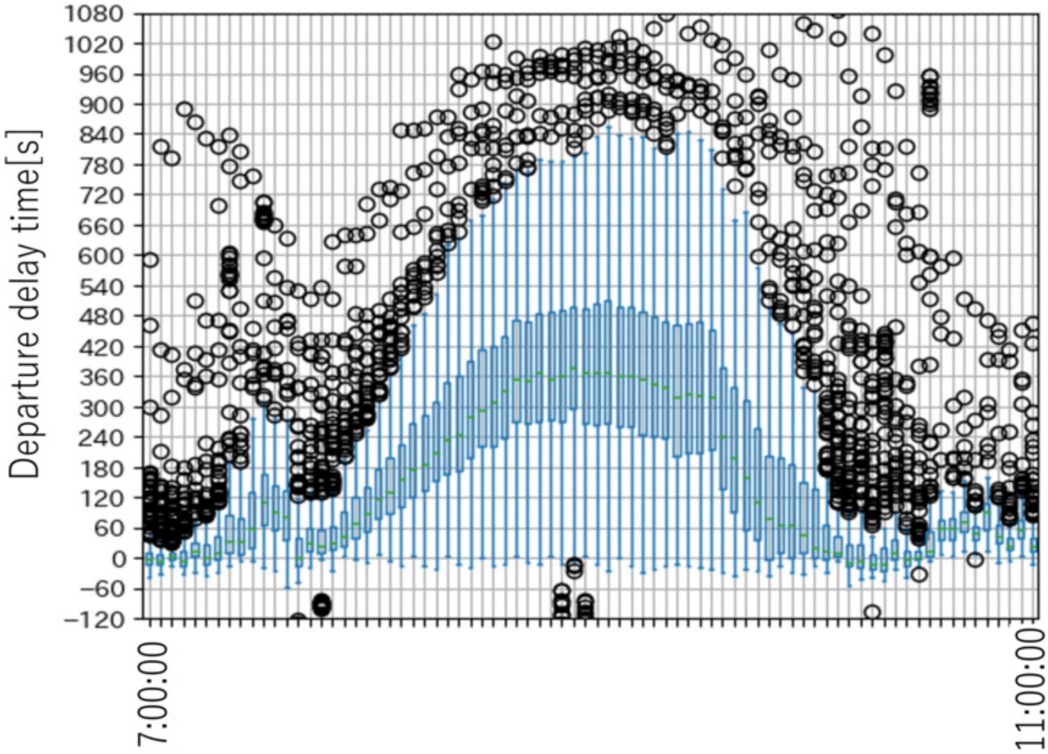

Figure 1: Delay trends.

Before the time period (1), various delays occur, but there is no trend of increasing delays. However, upon entering time period (1), a trend of increasing delays can be observed. This period corresponds to when the number of passengers increases, leading to shorter intervals between trains. Consequently, boarding and alighting times at stations tend to increase, causing dwell times to exceed the standard and making delays more likely to increase, as observed in trend (1). During time period (2), the increased delays remain constant without further increase. This period involves operational adjustments, such as modifying train intervals, to prevent further delay increments. In time period (3), a trend of decreasing delay times, which had been constant, can be observed. This is when the number of passengers decreases, leading to longer intervals between train operations. During this period, operational strategies to control delays and the wider train intervals contribute to the rapid convergence of increased delays.

An example of a timetable that shows the delays occurring in time period (1) is presented in Figure 2. Figure 2 illustrates the propagation of delays using a chromatic diagram, where bluer colors indicate lesser delays, and the colors shift towards red as the delays increase. From Figure 2, it can be seen that delays start at a train at a station in the lower-left part of the diagram, and the delay propagates as the next train's travel speed decreases.

The delays addressed in this paper are those propagating delays occurring in time period (1) as shown in Figure 2.

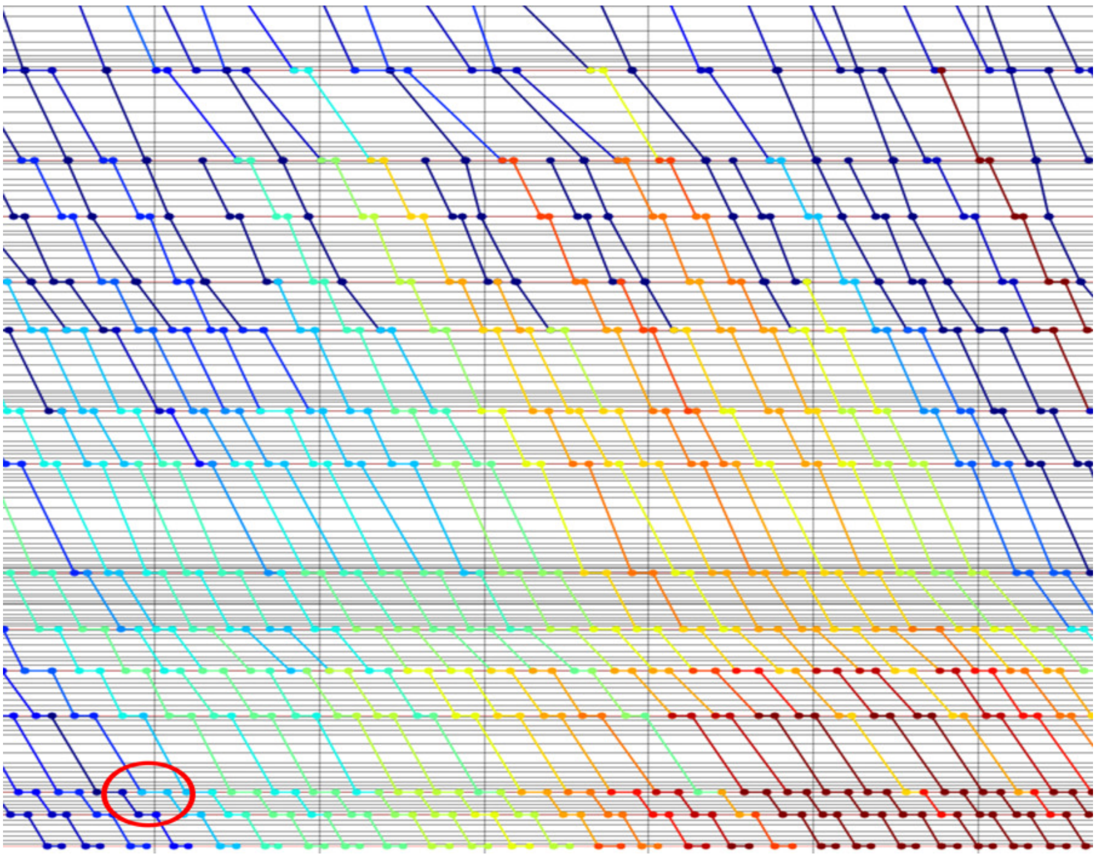

Figure 2: Chromatic Diagram.

3 Machine Learning-based Delay Prediction

In this chapter, we discuss the estimation of delay times using machine learning.

3.1 Configuration of the Machine Learning Model and Input Data

The network used in this paper is a CNN. While many studies on train delay prediction use LSTM, the propagation of delays is not occurring at individual train units but is influenced by the preceding trains. Therefore, we adopted CNN, which can learn the relationships in individual data spaces, assuming the data we handle are discrete. To learn the features of delay propagation, the data from the sections that the preceding trains have run, rather than the sections the train has just passed, are important. In reality, there are sections where the preceding train has not yet arrived at the station ahead when the target train arrives at a station. However, this study aims to evaluate delay countermeasures on a simulator, so the input data were created assuming that all data on the preceding trains are available.

The network used in this paper is shown in Figure 3. The standard network (Model 1) consists of three 3D convolutional layers and four fully connected layers, using only data from the stations behind the target station as input. We constructed three networks that take data from both the stations behind and ahead of the target station for Model 1 and compared their accuracy. The differences in the models and input data are shown in Table 1. The trains used as input include the five trains preceding the target train and the target train itself, totaling six trains. For Models 2 to 4, data from stations ahead of the target station are used as explanatory variables. Model 2 was constructed using LSTM, which is commonly used in existing research. Note that the input to LSTM is variable-length, so data not known at the prediction point are not input. Models 3 and 4 are both constructed using CNN, but Model 3 fills in the unknown parts of the data for the explanatory variables at the forward stations with zeros, assuming the delay is zero. On the other hand, Model 4 assumes that all information about the preceding trains is known, based on the assumption that it will be used in simulations. For all models, the mean absolute error is used as the loss function, and adam is used as the optimization algorithm.

The delay targeted in this study is the one that gradually increases. Therefore, sudden large delays are not the focus as they can act as learning noise. Only the data within the range shown in Table 2, which excludes such large delays, was used.

The data used consists of all 1,600 records from trains running between 7 am and 9 am on weekdays in the fiscal year 2019, after excluding the noise data. Of these, 80% were used as training data and 20% as test data to compare prediction accuracy. The test data was not selected consecutively but was chosen randomly.

Name	Model 1	Model 2	Model 3	Model 4
Structure	CNN	LSTM	CNN	CNN
Number of trains	6			
Stations including	Rear: 3 Forward: 0	Rear: 3 Forward: 3	Rear: 3 Forward: 3 (with 0 padding)	Rear: 3 Forward: 3 (Non 0 padding)
Dataset	Arrival Delay Departure delay Station occurrence delay Station Stopping Time Station travel time Station travel delay Arrival interval			

Table 1: Model variations.

Feature	Max[s]	Min[s]
Arrival delay	400	-100
Departure delay	400	0
Station occurrence delay	100	-100
Station Stopping Time	200	0
Station travel time	250	0
Station travel delay	100	-100
Arrival interval	500	0

Table 2: Data range.

3.2 Prediction accuracy of each model

The accuracy of each model is shown in Table 3. We prepared seven stations (A to G) along a metropolitan area line and compared the predicted accuracy for each station. At each validation station on the target line, there are no overtakings or meetups, and all trains stop at every station. The target time frame is the morning rush hour, during which trains run at the shortest intervals of the day. To evaluate prediction accuracy, we compared the accuracy within an absolute difference of 5 seconds between actual and predicted values.

From Table 3, comparing the reference Model 1 with Model 2, which uses LSTM, it was found that Model 1, which employs CNN, has a higher accuracy by more than 10 points at five out of seven stations. At the remaining two stations, the accuracy is approximately 7 points better. The lower accuracy of Model 2 compared to other models suggests that delay propagation is influenced by preceding and following trains, not occurring independently within each train unit.

Station	Model 1[%]	Model 2[%]	Model 3[%]	Model 4[%]
A	58.3	36.4	57.6	57.9
B	43.9	36.8	47.5	47.3
C	43.9	30.3	48.5	47.8
D	43.9	36.5	50.2	51.1
E	58.5	36.7	54.3	57.1
F	69.5	39.1	65.2	67.8
G	62.4	52.5	60.6	65.0

Table 3: Prediction accuracy of the models.

Since CNN was effective in learning the relationships between train delays, we examined the differences in the input explanatory variables. Comparing Models 1 and 3, an improvement in accuracy was observed at three out of seven stations, with a maximum improvement of about 6 points. However, accuracy worsened at other stations, especially beyond station E, where some stations showed a 4-point decrease. The difference between stations B, C, D, where accuracy improved, and stations E, F, G, where it deteriorated, might be due to the presence or absence of connections with other lines. Since the model does not include congestion rate as an explanatory variable, it does not account for passenger fluctuations due to connection timings. Thus, stations where congestion is a more likely cause of delays than preceding trains may experience noise when inputting information from forward stations. Model 3 assumes zero for unknown information of preceding trains at the station's arrival time, which could have introduced strong noise effects. Comparing Model 3 with Model 4, which assumes all preceding train information is known, only two stations showed worsened prediction accuracy, while the rest were equivalent or improved. Notably, the improved stations were E, F, and G, suggesting that unknown information from forward stations at the arrival time was crucial.

From the above, it is evident that CNN can capture the characteristics of propagating delays and has demonstrated the potential to be effective in predicting train delays. Furthermore, it has been clarified that data from stations ahead of the target station for the train in question are effective explanatory variables for delay prediction at many stations.

3.3 Prediction accuracy for data with different timetables

This study aims to pre-verify the effectiveness of delay countermeasures. Therefore, it is necessary to predict station stopping times after delay countermeasures, such as changes in operation time due to additional platforms or changes in block ratios. To evaluate the performance of this network, we used Models 1 and 4, which were trained with data from previous chapters, to determine the prediction accuracy for data with different timetables. Table 4 shows the results.

Station	Model 1[%]		Model 4[%]	
	Learned	Unknown	Learned	Unknown
a	58.3	54.0	57.9	56.6
b	43.9	48.4	47.3	50.5
c	43.9	46.5	47.8	45.6
d	43.9	54.8	51.1	53.5
e	58.5	53.8	57.1	54.7
f	69.5	64.4	67.8	62.4
g	62.4	60.7	65.0	57.5

Table 4: Prediction accuracy of Models 1 and 4 for data with different timetables.

From Table 4, it can be seen that each model is able to maintain prediction accuracy for an unknown timetable. One possible reason for maintaining accuracy is that station-specific data was not used as explanatory variables. Including fixed times such as travel time between stations or the scheduled stopping time at stations, which are determined by the timetable, would not allow for adjustments when the timetable changes. However, by not including those data and using values calculated based on the timetable, it was possible to learn solely the influence from other trains. On the other hand, there are stations where prediction accuracy has decreased. The explanatory variables used in this paper do not include data obtained from sources other than historical train traffic records. Therefore, by adding direct elements related to passenger flow such as congestion rates or weather conditions, there may be potential to further improve accuracy.

4 Concluding remarks

In this paper, with the goal of developing a train operation simulator for quantitatively evaluating delay countermeasures before implementation, we proposed a machine learning model and input data structure to predict train delays occurring at stations, and compared the accuracy of delay predictions for small, cumulative delays occurring during the morning rush hour.

As a result, CNN demonstrated its ability to capture the characteristics of delay propagation, with an average prediction accuracy of about 50% when the absolute difference between actual and predicted values was allowed to be 5 seconds, indicating it can better capture delay propagation features than LSTM. Furthermore, including stations ahead of the target station as explanatory variables improved prediction accuracy by up to 10 points, confirming the effectiveness of the proposed method based on the characteristics of delay propagation.

Additionally, the model was able to predict with accuracy without significant degradation for unknown timetables. Therefore, incorporating this learning machine into the simulator could enable more detailed verification of the effectiveness of delay countermeasures. However, the presence of stations where prediction accuracy did not improve using the proposed method suggests that there may be station-specific factors affecting delays.

In the future, we will evaluate the effectiveness of delay countermeasures by integrating this machine learning method into simulations.

References

[1] T. Kunimatsu, C. Hirai and N. Tomii: “Train Timetable Evaluation from the Viewpoints of Passengers by Microsimulation of Train Operation and Passenger Flow”, *IEEJ Trans. IA*, vol. 130, No. 4. p. 459-467 (2010) (in Japanese)

[2] A. Yamamura and N. Tomii: “Train Traffic Simulation based on Analysis of Historical Train Traffic Records for Dense Railway Networks”, *IEEJ Trans. IA*, vol. 139, No.2, pp.206-214 (2019)

[3] P. Zhou, L.Chen, X.Dai, B.Li and T.Chai: “Intelligent prediction of train delay changes and propagation using RVFLNs with improved transfer learning and ensemble learning”, IEEE Transactions on Intelligent Transportation Systems, vol.22, no.12, pp.7432–7444 (2021)

[4] D. Tatsui, K. Nakabasami, T. Kunimatsu, T. Sakaguchi and S. Tanaka: “Prediction Method of Train Delay Using Deep Learning Technique”, RTRI Report, vol.35, No.3, p.23-28 (2021) (in Japanese)

[5] A. Graves, S. Fernández and J. Schmidhuber: “Bidirectional LSTM Networks for Improved Phoneme Classification and Recognition”, ICANN 2005, pp.799–804 (2005)

[6] Y. LeCun, L. Bottou, Y. Bengio and P. Haffner: “Gradient based learning applied to document recognition”, Proceedings of the IEEE, vol.86, no.11, pp.2278–2324 (1998)

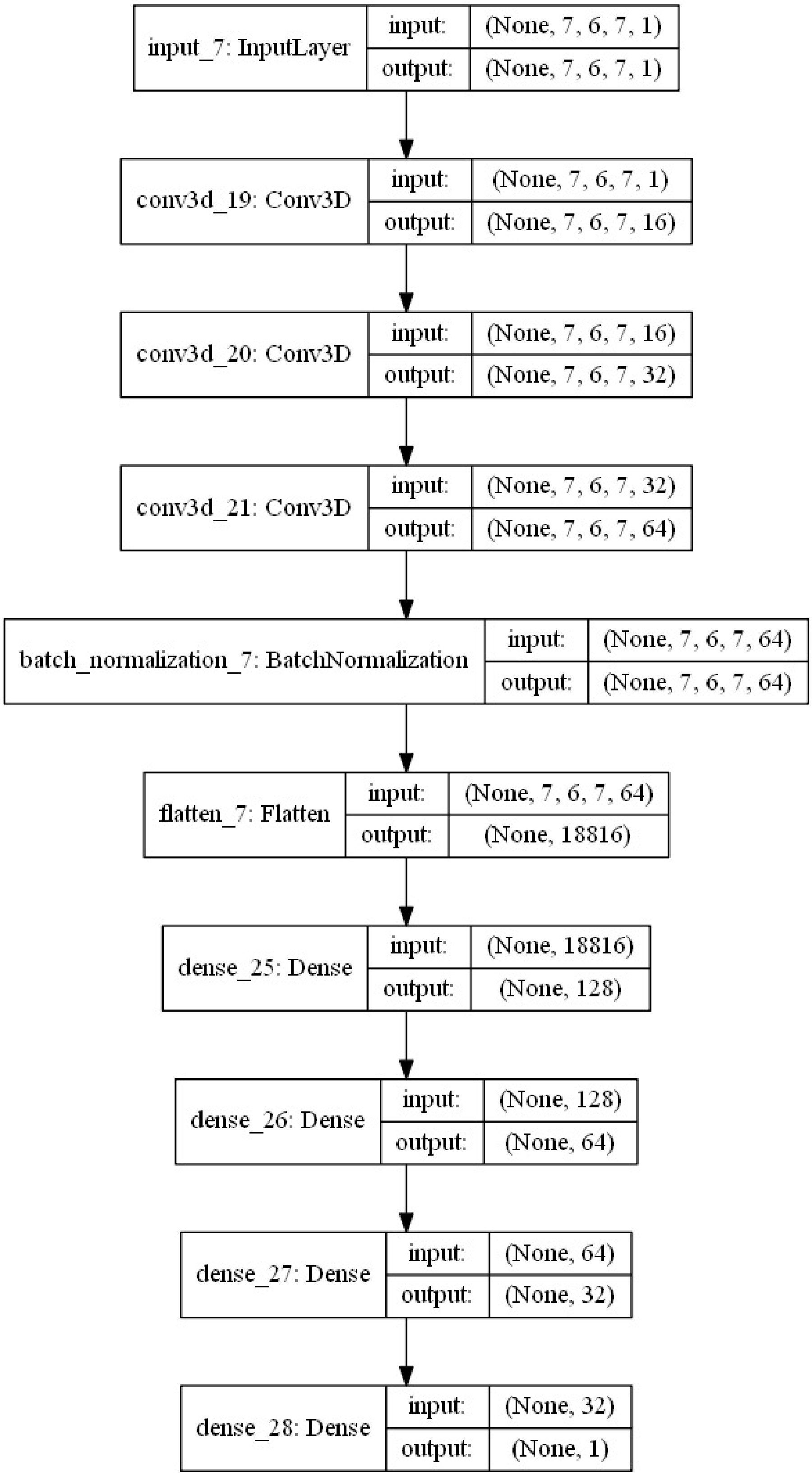

Figure 3: CNN model.