Hybrid Deformable Convolutional with Recurrent Neural Network for Optimal Traffic Congestion Prediction: A Fuzzy Logic Congestion Index System

Sara Berrouk, Abdelaziz El Fazziki, Mohammed Sadgal
Computer Science Engineering Laboratory
Faculty of Sciences, Cadi Ayyad University, Marrakech, Morocco

Abstract—In the field of Intelligent Transportation Systems (ITs), traffic congestion is considered as an important problem. Traffic blockage usually affects the quality of time, travel time, economy of the country, and transportability of people. The information of traffic congestion is collected and analyzed in ITs, and the methods to prevent the traffic congestion are predicted. However, the tackling of huge data is still challenging. The rapid increase in vehicle usage and road construction has resulted in traffic congestion. Various studies are undergone in ITs to recognize the traffic management system by adopting few resources. Real time-based traffic services are implemented to prevent the traffic congestion in existing areas. These services provide high expense accuracy. This paper plans to develop a new technique to predict the traffic congestion using improved deep learning approaches. At first, the benchmark dataset is gathered and the pre-processing of data is performed with removing the bad data, organizing the raw data, and filling the null values. The optimized weighted features are selected from the pre-processed data by adopting a new meta-heuristic Hybrid Jaya Harris Hawk Optimization (HJHHO) algorithm. The prediction of congestion parameters such as speed reduction rate, very low speed rate, and volume to capacity ratio of vehicles are performed by the proposed Improved Deformable Convolutional Recurrent Network (IDCRN) prediction model. These predicted measures are subjected to fuzzy interference system for congestion index computation. From the experimental analysis, it has proved that the proposed method has reduced the error rate while comparing with other deep learning and machine learning approaches.

Keywords—*Optimal traffic congestion prediction; deformable convolutional network; recurrent neural network; Hybrid Jaya Harris Hawk Optimization Algorithm; congestion index computation; fuzzy interference system*

I. INTRODUCTION

Population growth and the rapid urbanization in economy have increased the traffic clogging drastically in most developing cities and large areas in all over the world. It also increases both the road rage statistics and commute time. Consequently, the cause of accidents is increasing due to high traffic congestion [1]. Nowadays, the study for traffic management is highly significant among the researchers. The traffic clogging can be prevented by developing the infrastructure of transportation, which is expensive or else by organizing possible traffic schemes, like analyzing the blocking pattern or prediction of traffic congestion in short

term that can be more effective for road networks in a short time [2].

The timely congestion prediction models focus on the traffic factors such as speed, volume, and traffic stream on group of roads, one way road and especially small rural roads due to the unavailability of data. The limitation on predicting the road networks causes trouble for both the traffic agencies and commuters [3]. The researchers use the data, which is gathered from sensors like CCTV cameras, road sensors etc. These road sensors are located everywhere in road and even in vehicle networks to operate in all routes. This type of data is inconvenient because the process of operating, installation and maintenance is difficult, and it is expensive as it requires permission to access the data for third parties [4]. The real time traffic information like the average speed and blocking level on the section of roadway are provided in the web services such as Seoul Transportation Operation, Google Traffic, Information Service (TOPIS), Baidu Map and Bing Map [5]. Therefore, these web services are freely accessible and offer traffic congestion-related information to the majority cities all over the world, but only few studies are available based on this web services. The data requires multiple inputs, which are in time series, the processing of numerous input traffic images is complex due to high cost [6].

In past years, the data-driven methods were followed by the researchers to develop the mathematical and statistical models to detect the relation of time series in traffic data, also known as parametric method [7]. Mainly, the existing works were related to the presumption of stationary and linearity to capture the prediction trends such as smoothing model, error component method, and historical average method [8]. The seasonal pattern and long-term trends are decomposed to detect the pattern using Autoregressive Integrated Moving Average (ARIMA) model. However, it is unable to focus on the mean value of time series and incapable to predict the intense. The researchers began to concentrate on machine learning parametric models due to the limitations in parametric model [9]. The nonparametric model depends on the training data to establish parameter numbers and structural model[10]. The machine learning approach called K-Nearest Neighbor (KNN) predicts the traffic stream by searching the closer neighbor matching to the present data from the traditional database. The Support Vector Machine (SVM) approach minimizes the structural risk and has advantages in high dimensional data and minute samples [11]. The Bayesian Network (BN) approach

detects the issues in partial data, which is related to the message transferring mechanism [12]. Similarly, the deep learning approaches are also applied in these fields like time series production, recognition and translation, etc. because of the ability in data mining [13]. However, deep learning approach is difficult to use and preprocessing the output and input series in high dimensional data is very slow. The other deep learning-based approach is Convolution Neural Networks (CNN) that is used to extract the significant data from the raw data [14] [15][16]. The hybrid network attains promising results than simple networks due to its capability in mining the attributes from the input traffic data [17]. Therefore, a hybrid networks method shows better results in predicting the traffic congestion. The traffic congestion prediction with hybrid network contains very few works due to the shortage of high-quality traffic congestion data.

The major contribution of the suggested traffic congestion prediction is given here:

- To design a new model to predict the traffic congestion by employing the traffic dataset with new hybrid algorithm for feature selection and hybrid architecture for prediction with congestion index computation by fuzzy interference system.

- To select the weighted features with HJHHO algorithm by optimizing the weight and selecting the features from the pre-processed data through minimizing the variance of features.

- To evaluate a proposed IDCRN prediction with Deformable Convolution Network (DCN) and Recurrent Neural Network (RNN) architecture by optimizing hidden neuron count of DCN, learning rate of DCN, epoch count of DCN and further, hidden neuron counts of RNN using HJHHO algorithm. The objective function of IDCRN is to minimize the MAE and RMSE.

- To compute congestion index computation by considering the inputs as the speed reduction rate, very low speed rate and volume to capacity rate by the well-performing fuzzy interference system.

- To determine the performance of the suggested model with other existing approaches by evaluating different measures to examine the outcomes of the proposed model.

The other sections of this paper are illustrated here. Section II represents the related works. Section III denotes the architectural view of optimal traffic congestion prediction with new hybrid heuristic algorithm with fuzzy logic system. Section IV denotes the enhanced traffic congestion prediction with new optimized weighted feature selection. Section V depicts the optimal traffic congestion prediction by hybrid deep learning with traffic congestion index computation by fuzzy model. Section VI denotes the result and discussion. Section VII indicates the conclusion.

II. LITERATURE SURVEY

A. Related Work

In 2018, Tseng et al. [18] have applied Apache storm in real time traffic congestion prediction scheme by analyzing various data such as rainfall volume, road density and traffic incidents. The new SVM-based Real-Time Highway Traffic Congestion Prediction (SRHTCP) method has gathered the traffic incident data reported in roadways from the Taiwan police broadcasting service and the weather reports were collected from Taiwan Central Weather Bureau. Here, the fuzzy set theory was used to estimate the traffic congestion in real time with regarding rainfall volume, road density, and road traffic incidents. The road speed for next timeline was predicted with SRHTCP method by analyzing the weather data and traffic flow data. The suggested SRHTCP scheme has achieved better accuracy prediction than other prediction models based on weighted exponential moving average approach.

In 2020, Ranjan et al. [19] have proposed three techniques to predict the traffic congestion like, an effective and low-cost data attainment scheme by captivating a snap of traffic clogging from the online open source web service TOPIS and a hybrid deep learning approach to predict the traffic congestion level by extracting the temporal and spatial information. The relationship of temporal and spatial for predicting the traffic congestion level was analyzed efficiently and effectively by the proposed model. This suggested model effectively has outperformed other Deep Neural Network (DNN) models.

In 2020, Shin et al. [20] have proposed a deep learning approach to predict the traffic congestion and in addition to correct the missing spatial and temporal values. The proposed model was pre-processed with outlier removal method using the Median Absolute Deviation (MAD) of traffic data. Therefore, the spatial and temporal values were corrected using spatial and temporal trends and pattern data. The proposed prediction model was combined with LSTM for learning the data in time series. The suggested model was compared over other existing model and has achieved better prediction results.

In 2019, Zhao et al. [21] have proposed an optimized GRU architecture to predict the speed of trucks on urban roads under non-periodic congested situations. The driving data of struck in Beijing Road was gathered. To get rid of the unwanted data, the pre-processing and screening of data were approached, and then, sequence of traffic speed was extracted. The learning rate was not adjusted by the Suitable Development Goals (SGD) weight optimization algorithm. The weight was optimized by Adadelta, Rmsprop and Adam in this proposed GRU model. The accuracy of the suggested model was verified based on four scenarios, such as accident, workday, rainy and weekend.

In 2020, Zaki et al. [22] have suggested a new Hidden Markov Model to determine the traffic stages in two dimensional (2D) spaces during the high peak hours. The proposed model has captured variance in traffic pattern data using the contrast and mean speed. The proposed model has enhanced the prediction error than other neuro fuzzy and HMM approaches. In 2019, Wena et al. [23] have suggested a Hybrid Temporal Association Rules Mining mode to predict the traffic clogging. The traffic states were predicted using DBSCAN algorithm, which has suitable rules in analyzing the traffic congestion in road ways. The temporal associated rules in traffic states were extracted using genetic algorithm based temporal association rules mining algorithm. The classification process was used to predict the traffic congestion level. The prediction and the stimulation tests were studied in different sizes of road ways. The stimulation results have determined that the suggested model has predicted the traffic congestion with high precision.

In 2016, Li et al. [24] have proposed an adaptive real time prediction model. This scheme has comprised a traffic pattern recognition algorithm related to an adaptive threshold calibration method, the adaptive K-means clustering and a 2D speed prediction method. The patterns from traffic data were obtained by the adaptive K-means clustering. The adaptive threshold calibration method was applied to recognize the traffic congestion and prediction. The obtained result has shown that the adaptive K-means clustering has recognized the traffic pattern better than Gaussian Mixture method. The proposed model has performed well in real time application of traffic congestion prediction and gained better accuracy.

In 2019, Song et al. [25] have applied the k-means clustering algorithm to classify the spatio-temporel distribution of congested roadways. Then, the spatiotemporal features were mined to extract the potential factors using geo-detector. The congested patterns were selected from six inter- regional and intra-regional roads on weekdays. The public properties like tourist spots, hospitals and green spots were often congested in off peak and peak hours. The result has suggested that the roads built-in high-density areas could reduce the repeated trips in center of the city. The land utilizing plan has involved with a detailed design of the environment to enhance various travel approaches in order to reduce the traffic congestion and increase the effectiveness of traffic. More advanced techniques were applied in land use plan and traffic congestion based on multi source real time data.

B. Problem Statement

Multi source data are collected and evaluated to predict and prevent the traffic congestion and traffic stream in transportation system. The training of all this data is still challenging. Various studies were undergone to handle this type of huge road data, which prevented the real time traffic congestion by using different techniques. Table I shows the features and challenges of traditional congestion prediction on traffic flow methods. SVM [18] returns better prediction accuracy and the car speed of the proceeding time period is predicted by the SRHTCP model. But, used open datasets are not verified using t-test technique. CNN and LSTM [19] train the image data using a large resolution on a smaller resource and better performance is achieved with respect to the computing time. Still, it does not enhance the performance of the model by including external factors such as weather information for every road. LSTM [20] solves the long term dependency problem and time-series features associated with the traffic data are learnt. Yet, it does not enhance the accuracy in the urban areas and low-speed regions. HMM [21] enhances the recovery vehicle count in a specific stretch of road at particular times and also supports the enhancement strategies and the traffic management on the longer term like ramp metering. But, the technique of contrast is not verified on extra datasets. GRU [22] offers an efficient information service for truck drivers and simultaneously meets the efficiency and accuracy of matching. Still, the applicability is not considered for the real traffic systems. GA [23] minimizes the correlation in the environments and also predicts the traffic congestion with a low error. Yet, it does not automatically choose the best parameter values. Adaptive K-means cluster [24] is insensitive for the actual congestion probability and also enhances the prediction performance in the case of real time. But it does not consider the external influencing factors like mega-event and weather for enhancing the traffic mode recognition. Mining technique [25] improves the accessibility for distinct travel modes and also identifies the potential urban form factors and the traffic congestion hotspots. Still, it does not consider the useful knowledge for the improved decision making using novel approaches. Thus, it is necessary to introduce novel deep learning methods for predicting the congestion in the traffic flow management system in order to reduce the error and enhance the accuracy of the overall system.

TABLE I.   FEATURES AND CHALLENGES OF TRAFFIC CONGESTION PREDICTION MODELS

Author [citation]	Methodology	Features	Challenges
Tseng et al. [18]	SVM	The car speed of the proceeding time period is predicted by the SRHTCP model. It returns better prediction accuracy.	The used open datasets are not verified using t-test technique.
Ranjan et al. [19]	CNN and LSTM	Better performance is achieved with respect to the computing time. The image data is trained using a large resolution on a smaller resource.	The performance of the model is not enhanced by including external factors such as weather information for every road.
Shin et al. [20]	LSTM	The time-series features associated with the traffic data are learnt. The long term dependency problem is solved.	The accuracy is not enhanced in the urban areas and low-speed regions.
Zhao et al. [21]	HMM	It supports the enhancement strategies and the traffic management on the longer term like ramp metering. The recovery vehicle count is enhanced in a specific stretch of road at particular times.	The technique of contrast is not verified on extra datasets.
Zaki et al. [22]	GRU	It simultaneously meets the efficiency and accuracy of matching. An efficient information service is offered for truck drivers.	The applicability is not considered for the real traffic systems.
Wena et al. [23]	GA	The traffic congestion is predicted with a low error. The correlation in the environments is minimized.	The best parameter values are not automatically chosen.
Li et al. [24]	Adaptive K-means cluster	The prediction performance in the case of real time is enhanced. It is insensitive for the actual congestion probability.	The external influencing factors like mega-event and weather are not considered for enhancing the traffic mode recognition.
Song et al. [25]	Mining technique	It identifies the potential urban form factors and the traffic congestion hotspots. The accessibility is improved for distinct travel modes.	The useful knowledge for the improved decision making is not considered using novel approaches.

III. ARCHITECTURAL VIEW OF OPTIMAL TRAFFIC CONGESTION PREDICTION WITH NEW HYBRID HEURISTIC ALGORITHM WITH FUZZY LOGIC SYSTEM

A. Proposed Model and Description

Traffic congestion is one of the most important issues in all over the world. The current infrastructure is unable to cope with new traffic applications. The small spaces and other construction activities influence the traffic congestion. Due to the cause of traffic blockage, the fuel costs and the travel time of employers and distributing workers is affected. Traffic congestion is defined as the transportation vehicles surpass the capacity of roadway in peak time. The congestion indicators are mostly used to evaluate the traffic congestions in the urban road routes. Millions of peoples are affected by traffic congestion. This also causes noise and air pollution in whole surroundings. The impact of traffic blocking can be associated to fuel price raise, environment related matters, and transits cost. Various studies and researchers were undergone to overcome the traffic congestion. The timely prediction of traffic congestion in real time can prevent the unnecessary blockage. Deep learning and machine learning approaches were implemented to predict the traffic congestion. Machine learning-based model is most popular than other non-parametric models. It analyses the traffic patterns with low restrictions and gives better prediction results. Deep learning approaches are discussed to predict the real time traffic congestion. The traffic data are huge data, which are difficult to train. In this case, various techniques were executed to train the huge data volume and to enhance the prediction accuracy. The architectural diagram for the proposed traffic congestion prediction is given in Fig. 1.

The proposed traffic congestion prediction model covers five main phases (a) data collection (b) Pre-processing (c) Feature selection (d) Prediction and (e) Congestion index computation. The benchmark dataset is gathered from Radar traffic counts, which is publicly available. These datasets are pre-processed by three methods such as removal of bad data, organizing the raw data and filling the null values. In this pre-processing phase, the raw data are cleaned by eliminating the unwanted data and filling the missing data. The pre-processed data is inputted to the optimized weighted feature selection phase, where the weighted feature selection is enhanced by hybrid HJHHO algorithm by optimizing the weight and features. The selected features are given to the classification phase. The features are predicted with DCN and RNN architecture by optimizing the hyperparameters like hidden neuron count, learning rate and epoch count using HJHHO algorithm. As the prediction uses hybridization of two deep leading models with architecture optimization, the proposed model is termed as IDCRN. The main purpose of the proposed prediction is to minimize the error rate and maximize the prediction performance. The predicted parameters such as speed reduction rate, low speed rate, and volume to capacity rate are attained from IDCRN. Finally, the predicted parameters are inputted to fuzzy interference system for congestion index computation such as high, low, moderate and very high.
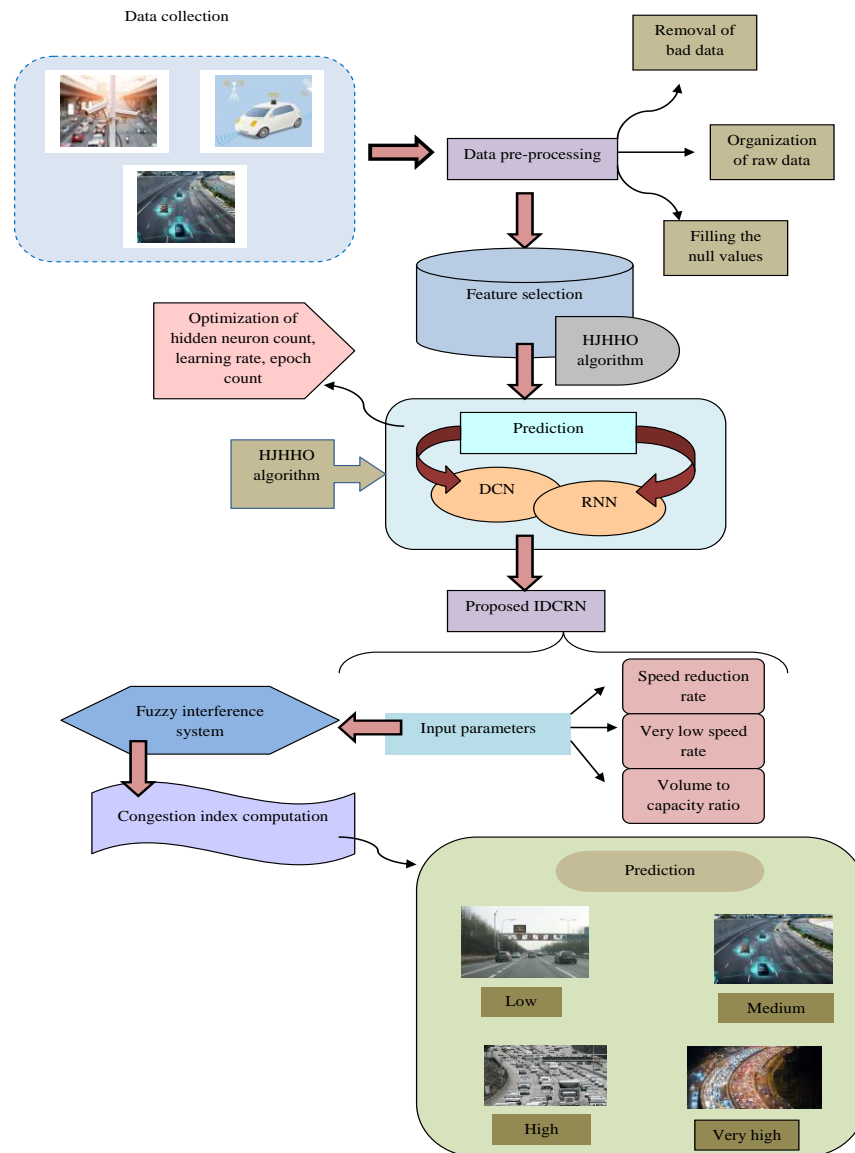
Fig. 1.   Architectural Illustration of the Proposed Traffic Congestion Prediction Method with Congestion Index Computation.

B. Dataset Description

The benchmark dataset Radar Traffic counts are collected from "https://data.austintexas.gov/Transportation-and-Mobility/Radar-Traffic-Counts/i626-g7ub"- Access date: 24-12-2021. The traffic speed and count are gathered from various Wavetronix radar sensors, which is taken from the city of Austin. The dataset contains the hourly transportation of vehicles with 6.83 million of rows and 17 columns. The 70% of data is used for training and 30% of data is used for testing.

The collected input data is determined as T_f^{ip}, where $f = 1, 2, ..., F$, here the term f is indicated as the total number of gathered dataset.

C. Data Pre-processing

The process of converting the collected raw data into suitable format is known as data pre-processing. The gathered raw data are pre-processed to remove the unwanted data and to eliminate the noises before utilization. The pre-processing of raw data is essential for accurate precision analysis. In this proposed work, the pre-processing phase is performed with removal of bad data, organizing the raw data, and filling the null values.

Removing bad data: The input data T_f^{ip} is given to bad data removal process. Data cleaning or removal of bad data is a process to remove the unwanted data in the gathered dataset. The conversion of data from a structure or format to another is known as data transformation. It is essential to remove the bad data before utilization. In this process, the corrupted data, duplicate data, missing data are removed or fixed from the dataset. The presence of duplicate or unwanted data results in producing error in the prediction performance. The removal of bad data is denoted as T_f^{bad}.

Organizing the raw data: The data gathered are mostly unorganized or non-systematic, which are known as raw data. The deconstruction analyze method is used to manipulate or organize the data. The obtained raw data are in the form of recorded values and the systematic process of organizing them is referred as raw data organization. The organized data is denoted as T_f^{org} .

Filling the null values: The traffic data are usually affected in two different categories. First, the data are missed in certain time periods and locations. The entire data is necessary for the prediction and modelling of transportations. Second is the loss of statistical information. It causes violation of missing traffic patterns. The null values are filled with appropriate methods. When the missing values are to predict, the 0 or NA is used instead of the missing values. Filling the null values generates robust data models. Finally, the pre-processed data is denoted as T_f^{pre} and it is applied to feature selection phase.

IV. ENHANCED TRAFFIC CONGESTION PREDICTION WITH NEW OPTIMIZED WEIGHTED FEATURE SELECTION

A. Optimal Weighted Feature Selection

The pre-processed data T_f^{pre} are given to HJHHO algorithm for the feature selection. From the evaluated mining features, the optimized weighted features are obtained by multiplying with the optimized weight with the extracted features Fe_d^{sel} as given in Eq. (1).

$$Fe_{d*}^{sel*} = OP_d \times Fe_d^{sel}$$

(1)

In the above equation, the optimized weight is indicated as $OP_d, d = 1, 2, ..., ND$, where ND indicates total number of features being considered and the term Fe_{d*}^{sel*} is denoted as weighted optimal features. The weight is optimized with HJHHO algorithm in the range of $[0,1]$. The maximum iteration is fixed as 10. The main objective function of the weighted feature selection is to optimize the weight by minimizing the variance of the selected features. The number of optimal features selected from the extracted features is counted as 10.

$$objfn1 = \underset{\{OP_d, Fe_d^{sel}\}}{\arg\max} \left(\frac{1}{\text{var}} \right)$$

(2)

In Eq. (2), the term var refers to the variance, $objfn1$ is the objective function, the variance is defined as "a statistical measurement of the spread between numbers in a data set. More specifically, variance measures how far each number in the set is from the mean and thus from every other number in the set" as given in Eq. (3).

$$Var = \sigma^2 = \frac{\sum_{d=1}^{r}(y_d - \bar{y})^2}{r-1}$$

(3)

Here, the term y_d denotes the dth data feature, $\bar{y}$ represents the data features and r denotes the total number of data features. Consider, the optimal weighted features as $Fea_{d*}^{sel*} = Fe_1^{sel}, Fe_2^{sel}, Fe_3^{sel}, ... Fe_{ND^*}^{sel}$, where the total number of optimized weighted features are represented as ND^*.

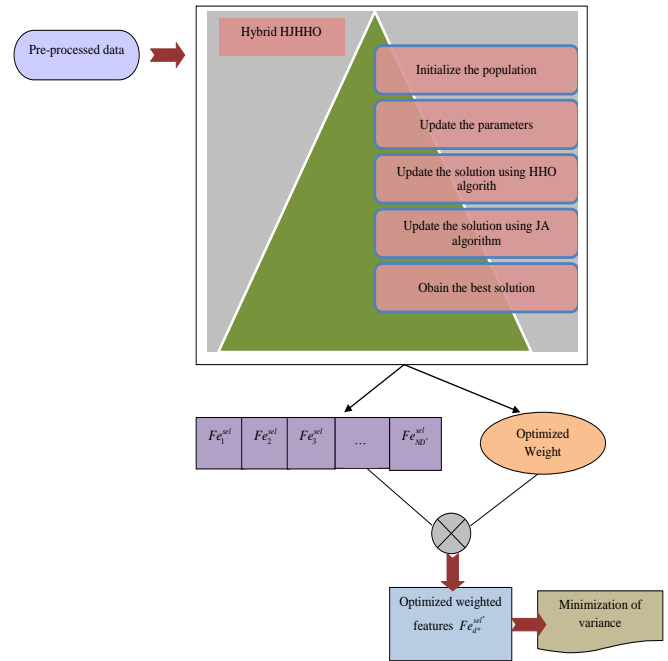The representation for optimal weighted feature selection is given in Fig. 2.

Fig. 2.   Representation of Optimal Weighted Feature Selection.

B. Proposed HJHHO

The proposed HJHHO algorithm is used for selecting the optimized weighted feature selection from data by optimizing the selecting the features and weight and improving the hybrid based IDCRN prediction performance by optimizing the hidden neuron count of DCN, learning rate of DCN, epoch count of DCN and further, hidden neuron counts of RNN.

HHO [26] algorithm is flexible and used in various optimization problems to attain optimal solutions. It has high efficiency and predicted the failure probability using key factors. The exploration phase is maximized using compound agents. However, the HHO algorithm suffers from challenges like population diversity problem and local optima in high dimensional issues. To over these challenges in HHO, JA algorithm is used. JA [27] algorithm is easy and simple to implement and it contains few parameters in a single phase. It is applied to resolve various problems in optimization and finds optimal solutions within less computational time. The proposed hybrid HJHHO algorithm enhances the problems in optimization and more efficient than other optimization algorithms. If escaping energy $|B| \geq 1$, the position is updated using HHO through exploration phase, otherwise the position is updated using JA algorithm.

HHO algorithm is inspired by exploring and attacking behavior of Harris hawks. The HHO algorithm is gradient free population-based optimization technique, which can be proposed to all kinds of optimization problems. The exploration phase of HHO is utilized in HJHHO. Harris hawks can spot and track their prey with their powerful eyes. The prey is not visible often. The hawk waits in the desert spot for several hours to observe, monitor, and to detect the prey. The candidate solutions are denoted as Harris hawks and the best candidate solution is referred as the intentional prey. The hawk waits in a location for hours to detect the prey in terms of two strategies. The term r is denoted as the strategy of each perching, hawks perch depend on the position of their family members and the rabbit. If the condition $r < 0.5$, they perch on random trees is given in Eq. (4).

$$C(n+1) = \begin{cases} C_{rand}(n) - m_1 \left| C_{rand}(n) - 2m_2 C(n) \right| & r \geq 0.5 \\ \left(C_{rabbit}(n) - C_k(n) \right) - \\ m_3 \left(FS + m_4 (WS - FS) \right) & r < 0.5 \end{cases} \tag{4}$$

In above equation, the term $C(n+1)$ is denoted as the position of hawks in next iteration n, $C_{rabbit}(n)$ denotes the position of rabbit, the term $C(n)$ is the current position of hawks, the random numbers are represented as m_1, m_2, m_3, m_4, and r in $(0,1)$, the random numbers are upgraded in all iterations. The terms WS and FS are the lower and upper bounds of the variables, $C_{rand}(n)$ denotes the selected hawks from the present population, and C_k represents the average position of hawks from the present population.

At first, the solutions are generated depend up on its random location of hawks. Secondly, the location difference of the best candidate and the average position of the other hawks add a component depend on the ranges of variable. The scaling component m_3 increases the nature of rules once when m_4 get closer to the value 1 and similar distribution takes place. Based on this rule, a randomly scaled movement is added to FS. Then, a randomly scaled coefficient is considered to the component to offer more diversity trends and to explore various region of the attribute space. A simple rule is utilized to imitate the behavior of hawks. The average position of the hawks is represented in Eq. (5).

$$C_k(n) = \frac{1}{A} \sum_{0=1}^{A} C_o(n) \tag{5}$$

Here, the term $C_k(n)$ denotes the position of every hawk in iteration n and A refers to the total number of hawks. The average position is attained in various ways, but the simplest rule is utilized here. The HHO algorithm transfers the exploration phase to exploitation phase based on their escaping behavior. The escaping energy of the prey is denoted in Eq. (6).

$$B = 2B_h \left(1 - \frac{n}{N} \right) \tag{6}$$

Here, the term B denotes the escaping energy of the prey; the maximum iteration is represented as N and B_h is the initial stage of energy. The value B_h differs in interval (-1, 1) at all iteration. If the B_h values decrease from 0 to -1, that means the rabbit is failing, and when the B_h value maximizes from 0 to 1, that means the rabbit is strengthening. When the escaping energy of the rabbit is $|B| \geq 1$, the hawks search the rabbit in various locations, hence the exploration phase takes place and when $|B| < 1$ JA algorithm is used in exploitation phase.

JA algorithm [27] is very simple and used to resolve all optimization problems. The main purpose of JA algorithm is when a best solution is attained once for a certain problem, the results are obtained by 0eliminating the worst solutions immediately. JA algorithm obtains the optimal solution by ignoring the worst solution. It resolves the constrained and unconstrained optimization problems. Let $e(z)$ is denoted as the minimized major function. There are h number of intended variables $(w = 1, 2, ..., h)$ in all iteration n, o represents the candidate solution. The best candidate attains the best value of $e(z)$ $\left(e(z)_{best} \right)$ in all candidate solution and further, the worst solution attains the worst value of $e(z)$ $\left(e(z)_{worst} \right)$ in whole candidate solution. If $Z_{k,d,v}$ is the given value of kth candidate in iteration n, the modified value is denoted in Eq. (7).

$$Z'_{k,d,v} = Z_{k,d,v} + m1_{k,v} \left(Z_{k,best,v} - \left| Z_{k,d,v} \right| \right)$$
$$- m2_{k,v} \left(Z_{k,worst,v} - \left| Z_{k,d,v} \right| \right) \tag{7}$$

In above equation, the term $Z_{k,best,v}$ represents best candidate of variable k and the term $Z_{k,worst,v}$ is denotes the worst candidate of variable k. The term $Z'_{k,d,v}$ is the upgraded value of the terms $m1_{k,v}$, $Z_{k,d,v}$, and $Z_{k,d,v}$ are the two number of kth candidate, in iteration at range [0,1]. The term $m1_{k,v} \left(Z_{k,best,v} - \left| Z_{k,d,v} \right| \right)$ indicates the inclination of the solutions to neglect the worst solution and the term $m2_{k,v} \left(Z_{k,worst,v} - \left| Z_{k,d,v} \right| \right)$ denotes the tendency to avoid the worst solution. At the end of iteration, the attained functional values are given to the next iteration. The pseudo code for designed HJHHO algorithm is given in algorithm 1.

Algorithm 1: Designed HJHHO

---

initialize of population
calculate the fitness of hawks
    update the position of the hawks using Eq. (4)
    update the position of rabbit using Eq. (4)

    Determine the parameters B escaping energy

If $|B| \geq 1$
    Update the solution using HHO
        update the position using exploration phase using Eq. (5)
else
    Update the solution using JA
    update the position using Eq. (7)
end for

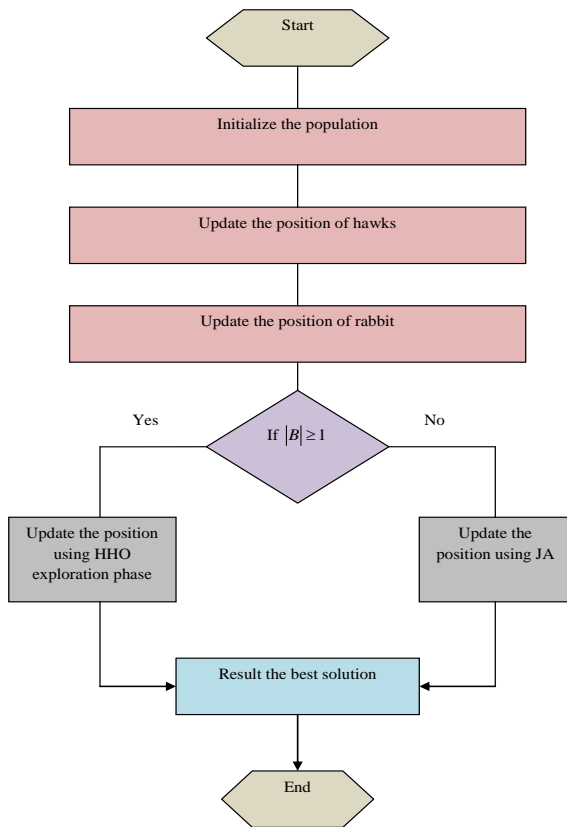The flowchart of designed HJHHO algorithm is given in Fig. 3.

Fig. 3. The Flowchart for Designed HJHHO Algorithm.

V. OPTIMAL TRAFFIC CONGESTION PREDICTION BY HYBRID DEEP LEARNING WITH TRAFFIC CONGESTION INDEX COMPUTATION BY FUZZY MODEL

A. Proposed IDCRN

The proposed IDCRN is a hybridization of improved DCN and RNN, which is used to predict the traffic congestion parameters in the proposed mode. The main function of IDRCN is to predict the data by optimizing hidden neuron count of DCN, learning rate of DCN, epoch count of DCN and further, hidden neuron counts of RNN by HJHHO algorithm. The proposed IDCRN minimizes the MAE and RMSE measures, thus reducing the error rate.

DCN [28] architecture is used to extract the deep features and predict the data. The offset vector for each sampling is introduced with DCN. However, generalization ability of DCN can be reduced in some extent. In addition, the RNN are used in time series prediction as it retains information about previous input and it is capable to remember all information throughout time. However, the computation of RNN is slow and it is difficult to train. To overcome the drawbacks in CNN and RNN, the hybrid IDCRN model is introduced. The parameters related to the traffic congestion index are predicted using IDCRN.

DCN are used to extract the input feature maps, where the field of offset is calculated with convolution networks. The offset obtained from the additional convolution network is inputted to the original convolution network. The present location of the random sampling is realized by the kernels and the location is not inadequate for the standard grid. Each offset location is learned by the network rather than the convolutional kernels. The end-to-end spatial transformations are effectively and easily realized by DCN. To analyze the 2D convolutions, the regular grid N is used to model the input feature map y and the addition of sample values is weighed by a. The dilation and size of receptive fields are defined with the grid N. The convolution kernel 3×3 with dilation 1 and $N = \{(-1,-1),(-1,0),...,(0,1),(1,1)\}$ is taken. In each location r_o, the output feature map x is given as represented in Eq. (8).

$$x(r_o) = \sum_{r_s} ar_s \cdot y(r_o + r_s) \tag{8}$$

Here, the term r_s computes the location in N. In DCN, the regular grid N is enhanced with the offsets $\{\Delta r_o | n1,..,N\}$, where $S = |N|$. The modified equation is given in Eq. (9).

$$x(r_o) = \sum_{r_s} ar_s \cdot y(r_o + r_s + \Delta rs) \tag{9}$$

Here, the term $r_s + \Delta rs$ represents the irregular location of offset, in which, Δrs is naturally fractional. The bilinear interpolation is given in Eq. (10).

$$y(r) = \sum_t K(t,r) \cdot y(t) \tag{10}$$

Here, the term $r = r_o + r_s + \Delta rs$ indicates the fractional location, t denotes all spatial location present in feature map y, the term $K(\cdot,\cdot)$ denotes the bilinear interpolation of kernel. The bilinear interpolation defines the linear interpolation in Y direction and X direction. The term K is divided into 2D dimensional kernel is referred in Eq. (11).

$$K(t,r) = k(t_y, r_y).k(t_x, r_x) \tag{11}$$

Hence, the weighted feature data are predicted using DCN architecture.

RNN [is another type of Artificial Neural Networks (ANN), which is applied in prediction of sequential traffic data by using previous traffic data. RNN generates the sequential information and each sequence of the element executes the similar task. The information is recorded in the memory of RNN. The input sequence of RNN is represented as c, which analyses the state t_{xx} at time xx. The different function v is applied with the previous state t_{xx-1}. The term t_{xx-1} does not contain the previous information. The RNN weight parameters are denoted as B, D and Y respectively. The RNN formula is given in Eq. (12) and Eq. (13).

$$t_{xx-1} = t_{xx-1} * Y + c_{xx} * B \tag{12}$$

$$o_{xx} = t_{xx} * D \tag{13}$$

The RNN model reduces the loss function and the errors in the predictions are compared with the actual values. The loss function is denoted as w, with regards the output state at given time xx. The correlation of gradient parameters is given here.

$$\frac{\partial w}{\partial_{xx-1}} = \frac{\partial w}{\partial_{xx}} Y \tag{14}$$

$$\frac{\partial w}{\partial B} = \sum_{xx=0}^{E} \frac{\partial w}{\partial_{xx}} c_{xx} \tag{15}$$

$$\frac{\partial w}{\partial Y} = \sum_{xx=0}^{e} \frac{\partial w}{\partial_{xx}} t_{xx-1} \tag{16}$$

Finally, the prediction data outcomes are attained from RNN. The optimized weighted features are inputted in DCN and RNN and the parameters such as speed reduction rate, volume to capacity rate and very low speed rate are predicted. The average from both the prediction models is utilized to perform the traffic congestion prediction.

B. Objective Function for IDCRN-based Parameter Prediction

The proposed IDCRN-based parameter prediction approach is used to predict the traffic congestion. The data is predicted using DCN and RNN. The parameters such as volume, time and speed of vehicles are inputted to IDCRN. The obtained output parameters such as volume to capacity rate, speed reduction rate and very low speed rate from IDCRN are subjected to fuzzy interference system to predict the congestion index computation in traffic flow. The main aim of the proposed prediction model is to optimize the hidden neuron count of DCN, learning rate of DCN, epoch count of DCN and further, hidden neuron count of RNN by minimizing the MAE and RMSE is given in Eq. (17).

$$objfn2 = \underset{\{HN,LR,EN,HRN\}}{\arg\max} \left(MAE + RMSE \right) \tag{17}$$

In above equation, the term HN denote the hidden neuron count of DCN that is ranging from $[5,255]$, LR denotes the learning rate of DCN lies in the range of $[0.01,0.99]$, EN represents the epoch count of DCN with the bounding limit of $[2,20]$ and further, the hidden neuron count of RNN HRN ranging from $[5,255]$. The term MAE indicates Mean Absolute Error and $RMSE$ denotes Root Mean Square Error. MAE is known as the "measure of errors between paired observations expressing the same phenomenon" as shown in Eq. (18).

$$MAE = \frac{\sum_{o=1}^{r} |xx_o - yy_o|}{r} \tag{18}$$

RMSE "is one of the most commonly used measures for evaluating the quality of predictions. It shows how far predictions fall from measured true values using Euclidean distance" as shown in Eq. (19).

$$RMSE = \frac{\sqrt{\sum_{o=1}^{r} \|yy_o - xx_o\|^2}}{r} \tag{19}$$

In above equations, term xx is represented as actual value and yy is denoted as forecasted value. The congestion measures like speed reduction rate, volume to capacity rate and very low speed rate are inputted to fuzzy interference for congestion index computation, each of these parameters indicates the traffic flow rates accurately. The fuzzy interference gives less error rate. The architectural diagram for IDCRN-based prediction parameters is given in Fig. 4.

C. Congestion Index Computation by Fuzzy Interference System

The training of different congestion measures has individual advantages and drawbacks. The congestion is an incident which is caused by various factors and the efforts are incorporated in different measures. The traffic congestion measurement is significant to detect the passenger perception. The boundary between each passenger differs due to the travel situations. These limitations are considered and fuzzy interference system is incorporated to detect the indistinct boundary in a set, and to find the solution uncertainty problems. The values of input parameters are calculated, categorized into various groups, determined different stages of traffic congestion and then finally established the congestion index computation.
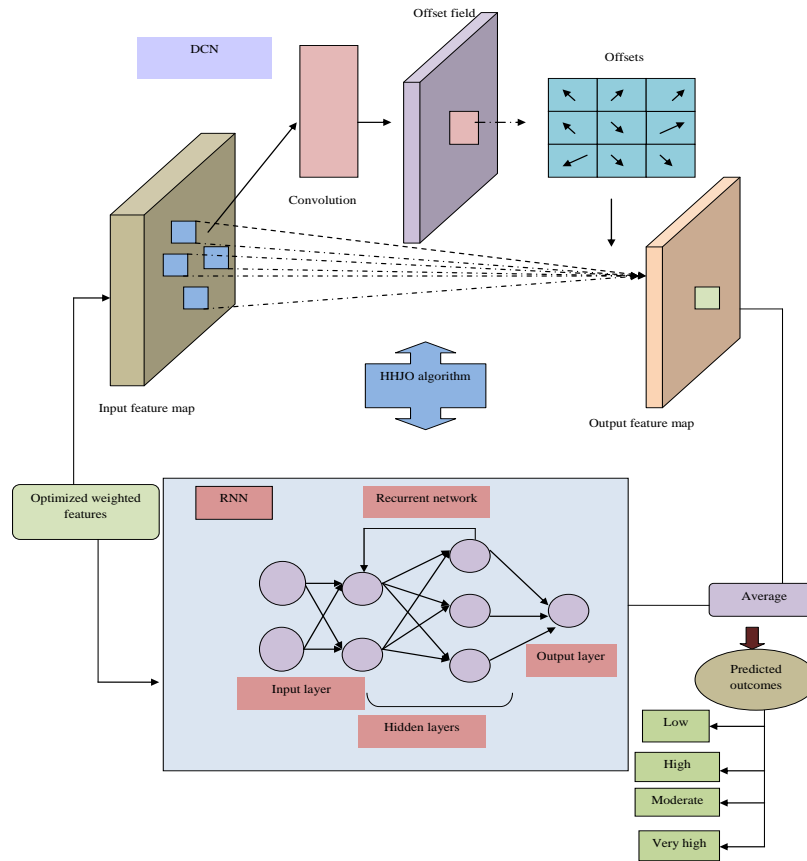
Fig. 4.   Architectural diagram for IDCRN-based Parameter Prediction

Input parameters: The input parameters are observed from proposed IDCRN and then combined to form a single fuzzy measure. The three parameters such as speed reduction rate, volume to capacity ratio and very low speed rate are compared with traffic volume and traffic travel time to roadway capacity rate. These three parameters are calculated separately based on the gathered data and combined according to fuzzy interference system rule. The traffic condition is represented with these three input parameters, by varying the volume to capacity, average speed and the variation in speed.

Speed reduction rate: Speed reduction rate is used to minimize the speed of vehicles that causes traffic congestion. The congestion in various routes during peak and non-peak situation are compared with speed reduction rate as given in Eq. (20).

$$Speed \ reduction \ rate = \frac{\left(NPS - PS\right)}{NPS} \tag{20}$$

In above Eq. (20), the term NPS denotes the non-peak flow speed and PS denotes the peak flow speed. The speed obtained value ranges from [0, 1], where 1 is considered as the worst state when the peak flow speed is nearer to 0. The 0 is denoted as best state when the peak range is equal to or larger than non-peak flow speed.

Very low speed rate: It is defined as the amount of travelling time at very low speed when compared over the total travel period as given in Eq. (21).

$$Very \ low \ speed \ rate = \frac{TSD}{TT} \tag{21}$$

In Eq. (21), the term TSD is represented as times spend in delay and the term TT denotes total travel time. The obtained value ranges from [0, 1], where 0 denotes the best state with no delay and 1 denotes the worst state with delayed travel time. The delay is referred as the time travel speed, which must be lesser than 5km/hr.

Volume to capacity ratio: It calculates all traffic block during peak hour situations. The volume is computed by taking the vehicles unit per hour, when the capacity of the roadway is increased as given in Eq. (22).

$$Volume \ to \ capacity \ ratio = \frac{VP}{CR} \tag{22}$$

In Eq. (22), the term VP denotes the volume of vehicle in peak hour and CR represents the capacity of roadway. The obtained value ranges from [0, 3], where the value closer to 0 is the best condition, when the capacity to ratio is minimized and value 3 is the worst condition when the huge volume of

vehicles moving towards the roadway by comparing to its capacity.

Output parameters: The output process is termed as congestion index, which is grouped into four phases low, high, very high and moderate. The function condition is given in a scale range from 0 to 1, where 0 is denoted as good and 1 as bad. The congestion condition of the four phases is determined in terms of this scale value.

Rules: The task of compiling three inputs and evaluating the traffic congestion measures are executed with a fuzzy interference system. When the speed reduction rate is considered as S, the very low speed rate is denoted as R, the volume to capacity ratio is indicated as V, and the congestion is referred as the term C. Here, S, R, V and C denote the degree of congestion. In the following rule, IF part is known as antecedent and THEN part is known as consequent. The exemplary representation of the generated fuzzy rules is given below.

- IF the speed reduction rate is low, AND the volume to capacity ratio is low, AND the very low speed rate is low, THEN, the congestion is low.

- IF the speed reduction rate is high, the volume to capacity ratio is high, AND the very low speed rate is high, THEN, the congestion is very high.

In the same way, totally, 54 numbers of rules are considered, among them three intense conditions are appropriate for using the combination of three categories such as volume to capacity ratio, speed reduction rate and very low speed. Among these rules, 38 numbers of suitable rules are

evaluated in this process to get required output. In this way, the traffic congestion index is computed. The architectural illustration for congestion index computation with fuzzy interference system is given in Fig. 5.

D. Output of Congestion Index by Fuzzy Interference System

The output of congestion index by fuzzy interference system in terms of membership function is given in Fig. 6.
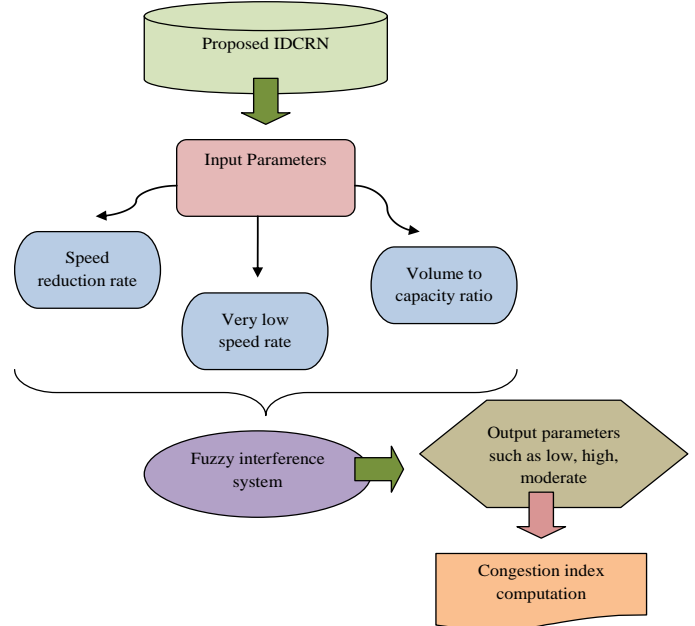

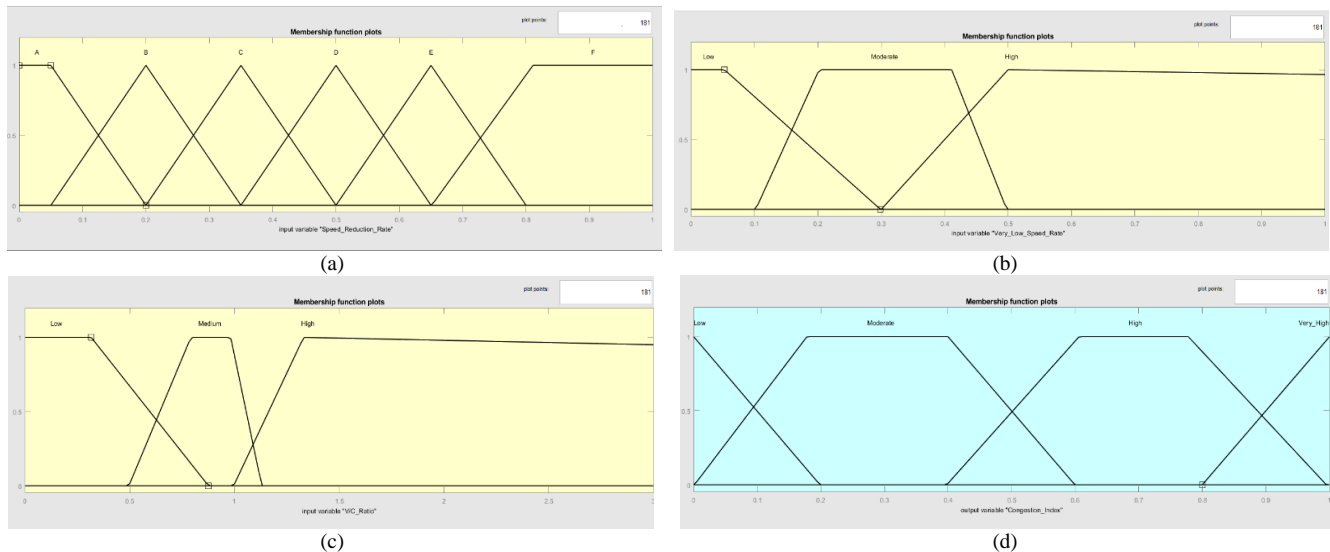Fig. 5.   Architectural Illustration of Congestion Index Computation with Fuzzy Interference System.

(a)

(b)

(c)

(d)

Fig. 6.   Input and Output of Fuzzy Interference System for "(a) Speed Reduction Rate, (b) Very Low Speed Rate, (c) Volume to Capacity Ratio and (d) Final Congestion Index.

VI. RESULT AND DISCUSSION

A. Experimental Setup

The proposed model traffic congestion prediction was implemented in python and the experimental results were evaluated. The performance of the suggested model was compared over various existing models in terms of different measures. The suggested model was compared with different heuristic algorithms and prediction approaches. The maximum iteration was 10 and the number of populations was 10. The proposed is compared over various heuristic algorithms like GWO-IDCRN [29], WOA-IDCRN [30], HHO-IDCRN [26], JA-IDCRN [27] and prediction models like LSTM [31], CNN [32], RNN [33], and CNN-RNN [34].

B. Performance Metrics

The performance measures used for traffic congestion prediction are given here.

1) L1-Norm "is the sum of the magnitudes of the vectors in a space. It is the most natural way of measure distance between vectors that is the sum of absolute difference of the components of the vectors" as shown in Eq. (23).

$$L1-Norm = \|yy\|_1 = \sum_{o=1}^{rs} |yy_o| \tag{23}$$

2) L2-Norm "is also known as the Euclidean norm. It is the shortest distance to go from one point to another" as shown in Eq. (24).

$$L2-Norm = \|yy\|_2 = \sqrt{\sum_{o=1}^{rs} yy_o^2} \tag{24}$$

3) L-infinity Norm "is the vector space of essentially bounded measurable functions with the essential supreme norm and only the largest element has any effect" as shown in Eq. (25).

$$L-infinity\ norm = \|yy\|_\infty = \max_{1 \leq o \leq rs} |yy_o| \tag{25}$$

4) MAE: it is given in Eq. (18).

5) Mean Absolute Scaled Error (MASE) "is a measure of the accuracy of forecasts. It is the mean absolute error of the forecast values, divided by the mean absolute error of the in-sample one-step naive forecast" as shown in Eq. (26).

$$MASE = \frac{1}{RS} \sum_{o=1}^{rs} u_o \tag{26}$$

6) Mean Effective Pressure (MEP) "is a theoretical parameter used to measure the performance of an internal combustion engine" as shown in Eq. (27).

$$MEP = \frac{100\%}{rs} \sum_{o=1}^{rs} \frac{xx_o - yy_o}{xx_o} \tag{27}$$

7) RMSE: It is given in Eq. (19).

8) Symmetric Mean Absolute Percentage Error (SMAPE) "is used to measure the predictive accuracy of models" as shown in Eq. (28).

$$SMAPE = \frac{1}{rs} \sum_{o=1}^{rs} \frac{|xx_0 - yy_o|}{(|xx_0| + |yy_o|)/2} \tag{28}$$

C. Performance Analysis on Heuristic Algorithms

The performance analysis of the suggested model HJHHO-IDCRN algorithm is compared with other meta-heuristic algorithms by varying the learning percentage is given in Fig. 7. The MAE of proposed HJHHO-IDCRN algorithm is 1% higher than GWO-IDCRN, 3% higher than WOA, 1% higher than HHO-IDCRN and 4% higher than JA-IDCRN at learning percentage 60. The MEP measure had attained higher results over other heuristic algorithms, thus the proposed HJHHO-IDCRN algorithm is 3% superior to GWO-IDCRN, 2% superior to WOA-IDCRN, 4% superior to HHO-IDCRN and 2% superior to JA-IDCRN at learning percentage 40. The performance of the suggested algorithm has reduced the error rate and obtained high prediction accuracy than other existing algorithms.

D. Performance Measures on Prediction Models

The performance analysis of the proposed HJHHO-IDCRN algorithm is compared with other prediction models by varying learning percentage is given in Fig. 8. At learning percentage 55, the L-infinity form measures of proposed HJHHO-IDCRN algorithm is 12% higher than LSTM, 16% higher than CNN, 11% higher than RNN and 16% higher than CNN-RNN. The MAE measures of HJHHO-IDCRN algorithm are 2% is improved than LSTM, 1% improved than CNN, 3% improved than RNN and 5% improved than CNN-RNN at learning percentage 75. The MEP measures had obtained higher results over other heuristic algorithms, thus the proposed HJHHO-IDCRN algorithm is 8% superior to LSTM, 4% superior to CNN, 5% superior to RNN and 6% superior to CNN-RNN at learning percentage 85. The proposed algorithm has gained high prediction accuracy than other prediction models and reduced the error rate.
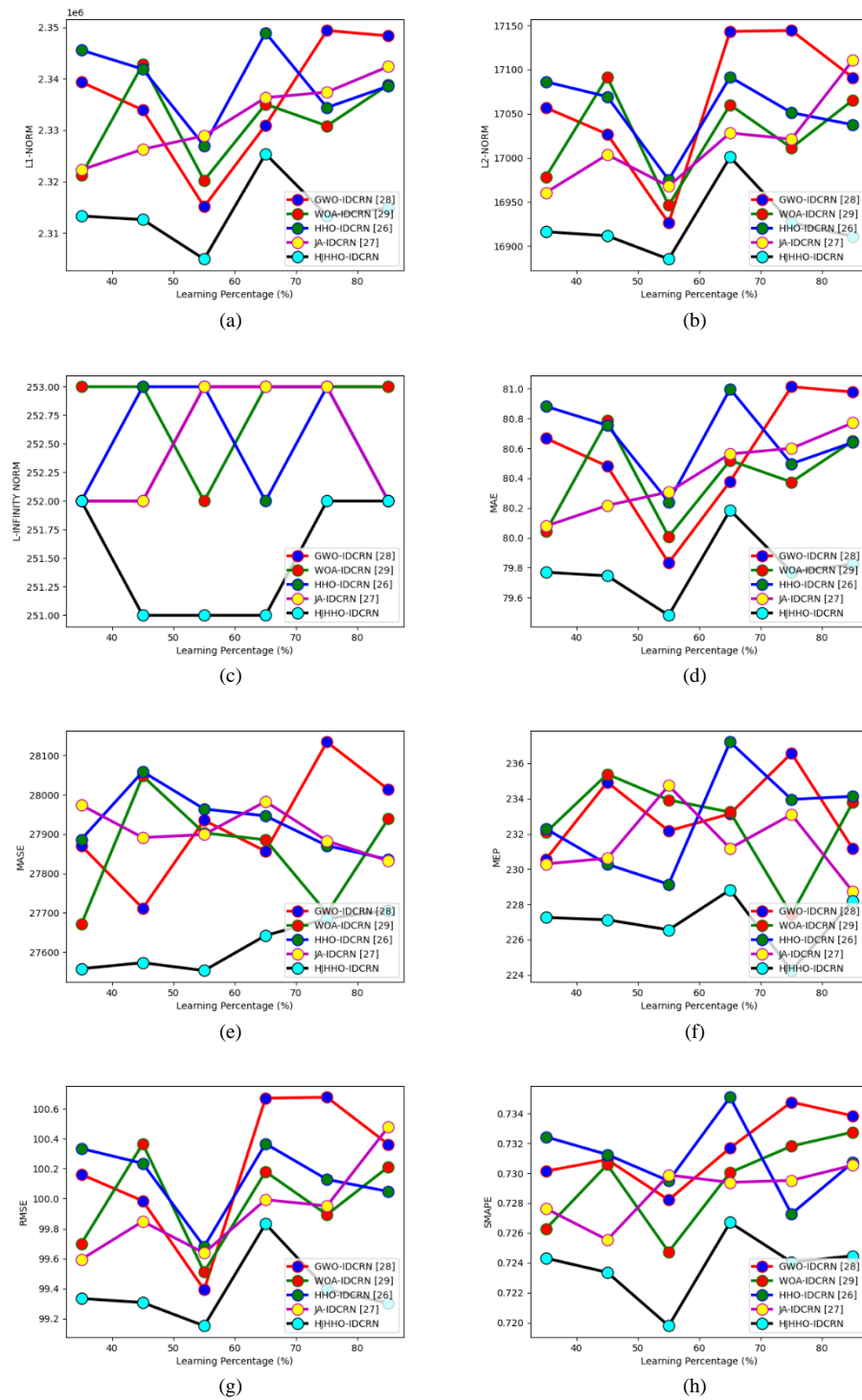
Fig. 7. Analysis of Traffic Congestion Prediction over other Heuristic Algorithms by Varying Learning Percentage in Terms of (a) L1Norm (b) L2-Norm (c) L-Infinity Norm (d) MAE (e) MASE (f) MEP (g) RMSE (h) SMAPE.
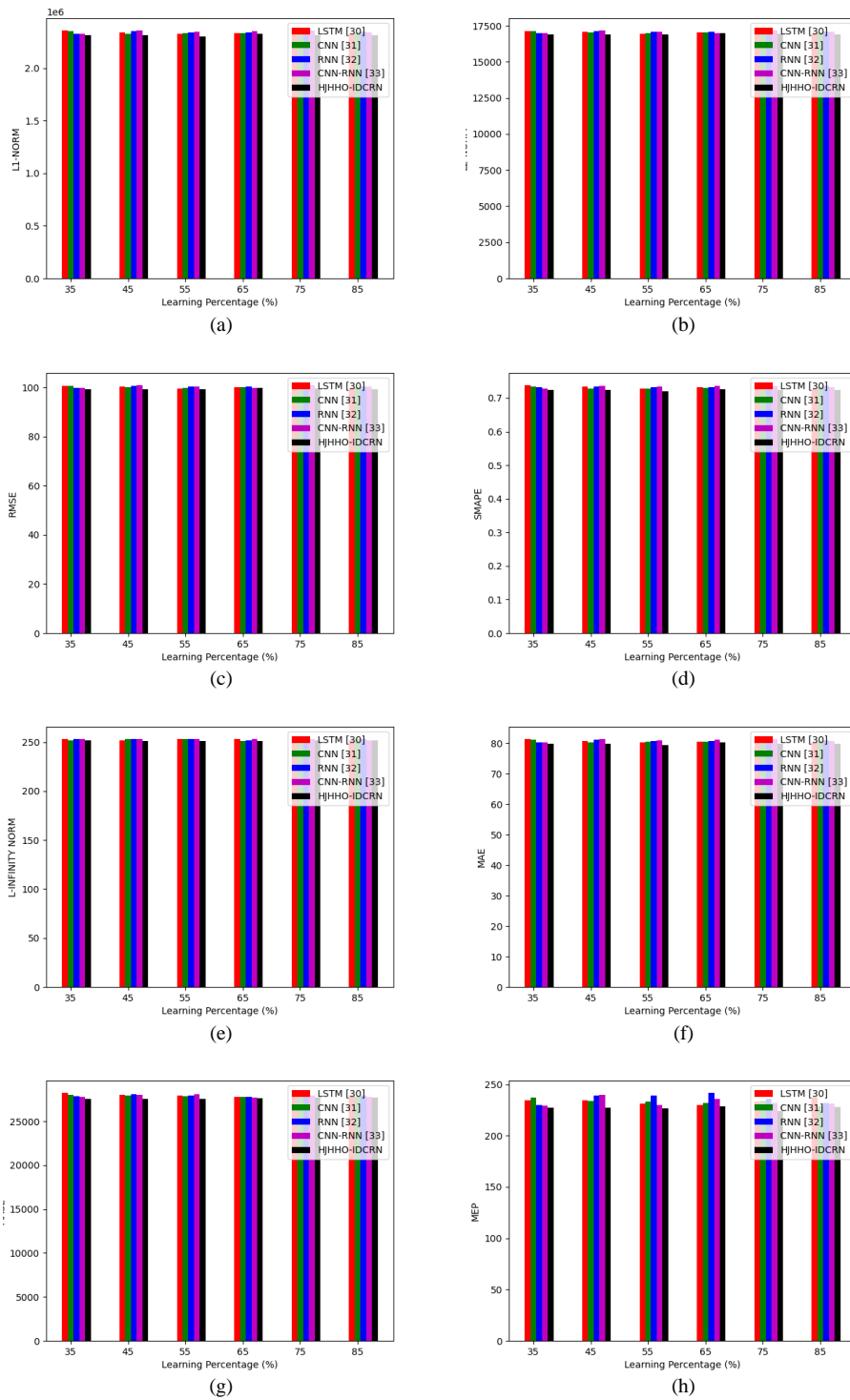
Fig. 8. Analysis of Traffic Congestion Prediction over other Prediction Models by Varying Learning Percentage in Terms of (a) L1Norm (b) L2-Norm (c) L-Infinity Norm (d) MAE (e) MASE (f) MEP (g) RMSE (h) SMAPE.

E. Overall Performance Analysis on Algorithms

The overall performance analysis of traffic congestion is evaluated with HJHHO-IDCRN is illustrated in Table II. The SMAPE measures of the proposed HJHHO-IDCRN algorithm is 5% higher than GWO-IDCRN, 3% higher than WOA-IDCRN, 1% higher than HHO-IDCRN and 6% higher than JA-IDCRN. The RMSE measures of proposed HJHHO-IDCRN algorithm is 15% superior to GWO-IDCRN, 11% superior to WOA-IDCRN, 14% superior to HHO-IDCRN and 9% superior to JA-IDCRN. The MAE measures of proposed HJHHO-IDCRN algorithm is 3% higher than GWO-IDCRN, 1% higher than WOA-IDCRN, 5% higher than HHO-IDCRN and 4% higher than JA-IDCRN. Therefore, the outcomes of the proposed HJHHO-IDCRN algorithm obtain better results than other heuristic algorithms while evaluating with all the measures.

F. Overall Performance Analysis on Prediction Models

The overall performance measures of the proposed HJHHO-IDCRN algorithm are compared with various prediction models and the prediction measures are given in Table III. The MASE measures of proposed HJHHO-IDCRN

algorithm is 2% better then LSTM, 4% better than CNN, 3% better than RNN and 6% better than CNN-RNN. The MEP measures of proposed HJHHO-IDCRN algorithm is 11% better then LSTM, 22% better than CNN, 16% better than RNN and 15% better than CNN-RNN. The MEP measures had gained higher results over other prediction models, thus the proposed HJHHO-IDCRN algorithm is 15% superior to LSTM, 22% superior to CNN, 15% superior to RNN and 16% superior to CNN-RNN The suggested HJHHO-IDCRN algorithms give better prediction performance and reduce the error rate.

G. Results on Traffic Congestion Index

The obtained congestion parameter results are illustrated in Table IV. The three parameters capture the real condition of traffic congestion. The results have shown when a particular input parameter is high, then the other two parameters are on lower sate. This state defines the moderate condition. When all the three parameters are on higher state, it results in high congestion result. In other hand, when all three parameters are in lower state, the congestion index results in very low values. The proposed model gives accurate results in traffic congestion prediction while comparing to other existing methods.

TABLE II. OVERALL PERFORMANCE OF TRAFFIC CONGESTION PREDICTION OVER OTHER META-HEURISTIC ALGORITHMS

Algorithms	GWO-IDCRN [29]	WOA-IDCRN[30]	HHO-IDCRN[26]	JA-IDCRN[27]	HJHHO-IDCRN
MEP	236.5818	227.4271	233.9635	233.1066	224.2438
SMAPE	0.734767	0.731819	0.727276	0.729521	0.724051
MASE	28134.79	27699.09	27870.95	27882.89	27685.07
MAE	81.01386	80.37438	80.49721	80.60052	79.77148
RMSE	100.6757	99.89368	100.1283	99.95179	99.39565
L1-NORM	2349402	2330857	2334419	2337415	2313373
L2-NORM	17144.45	17011.28	17051.24	17021.18	16926.47
L-INFINITY NORM	253	253	253	253	252

TABLE III. OVERALL PERFORMANCE OF TRAFFIC CONGESTION PREDICTION OVER OTHER PREDICTION MODELS

Algorithms	LSTM[31]	CNN[32]	RNN[33]	CNN-RNN[34]	HJHHO-IDCRN
MEP	233.3224	233.7626	235.8281	232.0925	224.2438
SMAPE	0.728538	0.732675	0.732912	0.735904	0.724051
MASE	27856.46	28007	28206.31	27931.84	27685.07
MAE	80.18579	80.69983	81.0989	81.31279	79.77148
RMSE	99.61117	100.3257	100.5825	100.767	99.39565
L1-NORM	2325388	2340295	2351868	2358071	2313373
L2-NORM	16963.17	17084.86	17128.58	17159.99	16926.47
L-INFINITY NORM	252	253	252	253	252

TABLE IV.   RESULTS ON TRAFFIC CONGESTION INDEX

Speed reduction rate	Volume to capacity ratio	Very low speed rate	Congestion index
0.669068	0.302595	0.282547	0.467693
0.669252	0.304988	0.294872	0.412533
0.68714	0.448816	0.294838	0.531245
0.671981	0.319807	0.239569	0.765443
0.662298	0.740872	0.361626	0.678416
0.66062	0.416873	0.537614	0.504879
0.658038	0.174094	0.277427	0.43729
0.691069	0.394865	0.398412	0.268478
0.667211	0.122501	0.398846	0.309538
0.671858	0.359507	0.247754	0.60257
0.666126	0.246816	0.274632	0.610337
0.683519	0.244399	0.281251	0.586398
1	0.546707	1	0.380909
0.654448	0.38242	0.952603	0.450354
0.704439	0.276988	0.366578	0.343328
0.668823	0.300416	0.247549	0.348988
0.680675	0.393189	0.398136	0.440765
0.657003	0.98637	0.511805	0.254022

VII. CONCLUSION

In this suggested work, a hybrid-based deep learning approach with congestion index computation by fuzzy model was implemented to predict the traffic congestion. The traffic data was gathered from a publicly available database. The dataset was pre-processed with three phases to remove the unwanted data and to fill the missing data by removing bad data, organizing the raw data and filling the null values techniques. The weighted feature selection of data was performed with proposed HJHHO algorithm by optimizing the weight and minimizing the variance. The prediction phase was done with proposed IDCRN. The main purpose of prediction phase was to minimize the MAE and EMSE measures by optimizing the hidden neuron count of DCN, learning rate of DCN, epoch count of DCN and further, hidden neuron counts of RNN. The congestion parameters such as speed reduction, volume to capacity ratio and very low speed rate obtained from IDCRN were subjected to fuzzy interference system for congestion index computation. Finally, the proposed HJHHO-IDRCN algorithm has improved the performance of prediction. The performance analysis has shown the MAE measures of HJHHO-IDRCN algorithm was 5% higher than GWO-IDRCN, 3% higher than WOA-IDRCN, 1% higher than HHO-IDRCN and 6% higher than JA-IDRCN. Thus, the suggested HJHHO-IDRCN algorithm has attained better prediction in traffic congestion. In future works, the proposed model will be run over different datasets of different road networks and furthermore it will be deployed in route planning solutions.

REFERENCES

[1] C. Onyeneke, "Modeling the Effects of Traffic Congestion on Economic Activities - Accidents, Fatalities and Casualties," Biomed. Stat. Informatics, vol. 3, no. 2, p. 7, 2018, doi: 10.11648/j.bsi.20180302.11.

[2] Z. Huang, J. Xia, F. Li, Z. Li, and Q. Li, "A Peak Tra ffi c Congestion Prediction Method Based on Bus Driving Time," entropy, vol. 21, no. 709, pp. 1–18, 2019.

[3] M. Chen, X. Yu, and Y. Liu, "PCNN: Deep Convolutional Networks for Short-Term Traffic Congestion Prediction," IEEE Trans. Intell. Transp. Syst., vol. 19, no. 11, pp. 3550–3559, 2018, doi: 10.1109/TITS.2018.2835523.

[4] A. Elfar, A. Talebpour, and H. S. Mahmassani, "Machine Learning Approach to Short-Term Traffic Congestion Prediction in a Connected Environment," Transp. Res. Rec., vol. 2672, no. 45, pp. 185–195, 2018, doi: 10.1177/0361198118795010.

[5] X. Ma, H. Yu, Y. Wang, and Y. Wang, "Large-scale transportation network congestion evolution prediction using deep learning theory," PLoS One, vol. 10, no. 3, pp. 1–17, 2015, doi: 10.1371/journal.pone.0119044.

[6] S. Zhang, Y. Yao, J. Hu, Y. Zhao, S. Li, and J. Hu, "Deep autoencoder neural networks for short-term traffic congestion prediction of transportation networks," Sensors (Switzerland), vol. 19, no. 10, pp. 1–19, 2019, doi: 10.3390/s19102229.

[7] S. Ye, "Research on Urban Road Traffic Congestion Charging Based on Sustainable Development," Phys. Procedia, vol. 24, pp. 1567–1572, 2012, doi: 10.1016/j.phpro.2012.02.231.

[8] C. Wang, M. A. Quddus, and S. G. Ison, "Impact of traffic congestion on road accidents: A spatial analysis of the M25 motorway in England," Accid. Anal. Prev., vol. 41, no. 4, pp. 798–808, 2009, doi: 10.1016/j.aap.2009.04.002.

[9] J. Zhao, Y. Gao, Z. Bai, H. Wang, and S. Lu, "Traffic speed prediction under non-recurrent congestion: based on lstm method and beidou navigation satellite system data," IEEE Intell. Transp. Syst. Mag., vol. 11, no. 2, pp. 70–81, 2019, doi: 10.1109/MITS.2019.2903431.

[10] S. Sun, J. Chen, and J. Sun, "Traffic congestion prediction based on GPS trajectory data," Int. J. Distrib. Sens. Networks, vol. 15, no. 5, 2019, doi: 10.1177/1550147719847440.

[11] M. Aftabuzzaman, G. Currie, and M. Sarvi, "Evaluating the Congestion Relief Impacts of Public Transport in Monetary Terms," J. Public Transp., vol. 13, no. 1, pp. 1–24, 2010, doi: 10.5038/2375-0901.13.1.1.

[12] M. Bani Younes and A. Boukerche, "A performance evaluation of an efficient traffic congestion detection protocol (ECODE) for intelligent

transportation systems," Ad Hoc Networks, vol. 24, no. PA, pp. 317–336, 2015, doi: 10.1016/j.adhoc.2014.09.005.

[13] X. Yang, Yuanfeng and Cui, Z. and Wu, J. and Zhang, G. and Xian, "Fuzzy c-means clustering and opposition-based reinforcement learning for traffic congestion identification," J. Inf. Comput. Sci., vol. 9, pp. 2441–2450, 2012.

[14] Y. Gu and L. Xie, "Fast sensitivity analysis approach to assessing congestion induced wind curtailment," IEEE Trans. Power Syst., vol. 29, no. 1, pp. 101–110, 2014, doi: 10.1109/TPWRS.2013.2282286.

[15] H. Shankar, P. L. N. Raju, and K. R. M. Rao, "Multi Model Criteria for the Estimation of Road Traffic Congestion from Traffic Flow Information Based on Fuzzy Logic," J. Transp. Technol., vol. 02, no. 01, pp. 50–62, 2012, doi: 10.4236/jtts.2012.21006.

[16] Y. Zhang, N. Ye, R. Wang, and R. Malekian, "A method for traffic congestion clustering judgment based on grey relational analysis," ISPRS Int. J. Geo-Information, vol. 5, no. 5, pp. 1–15, 2016, doi: 10.3390/ijgi5050071.

[17] J. F. W. Zaki, A. M. T. Ali-Eldin, S. E. Hussein, S. F. Saraya, and F. F. Areed, "Time aware hybrid hidden markov models for traffic congestion prediction," Int. J. Electr. Eng. Informatics, vol. 11, no. 1, pp. 1–17, 2019, doi: 10.15676/ijeei.2019.11.1.1.

[18] F. H. Tseng, J. H. Hsueh, C. W. Tseng, Y. T. Yang, H. C. Chao, and L. Der Chou, "Congestion prediction with big data for real-time highway traffic," IEEE Access, vol. 6, pp. 57311–57323, 2018, doi: 10.1109/ACCESS.2018.2873569.

[19] N. Ranjan, S. Bhandari, H. P. Zhao, H. Kim, and P. Khan, "City-wide traffic congestion prediction based on CNN, LSTM and transpose CNN," IEEE Access, vol. 8, pp. 81606–81620, 2020, doi: 10.1109/ACCESS.2020.2991462.

[20] D. H. Shin, K. Chung, and R. C. Park, "Prediction of Traffic Congestion Based on LSTM through Correction of Missing Temporal and Spatial Data," IEEE Access, vol. 8, pp. 150784–150796, 2020, doi: 10.1109/ACCESS.2020.3016469.

[21] J. Zhao et al., "Truck traffic speed prediction under non-recurrent congestion: Based on optimized deep learning algorithms and GPS Data," IEEE Access, vol. 7, pp. 9116–9127, 2019, doi: 10.1109/ACCESS.2018.2890414.

[22] J. F. Zaki, A. Ali-Eldin, S. E. Hussein, S. F. Saraya, and F. F. Areed, "Traffic congestion prediction based on Hidden Markov Models and contrast measure," Ain Shams Eng. J., vol. 11, no. 3, pp. 535–551, 2019, doi: 10.1016/j.asej.2019.10.006.

[23] F. Wen, G. Zhang, L. Sun, X. Wang, and X. Xu, "A hybrid temporal association rules mining method for traffic congestion prediction," Comput. Ind. Eng., vol. 130, no. 6, pp. 779–787, 2019, doi: 10.1016/j.cie.2019.03.020.

[24] F. Li, J. Gong, Y. Liang, and J. Zhou, "Real-time congestion prediction for urban arterials using adaptive data-driven methods," Multimed. Tools Appl., vol. 75, no. 24, pp. 17573–17592, 2016, doi: 10.1007/s11042-016-3474-3.

[25] J. Song, C. Zhao, S. Zhong, T. A. S. Nielsen, and A. V. Prishchepov, "Mapping spatio-temporal patterns and detecting the factors of traffic congestion with multi-source data fusion and mining techniques," Comput. Environ. Urban Syst., vol. 77, no. July, 2019, doi: 10.1016/j.compenvurbsys.2019.101364.

[26] A. A. Heidari, S. Mirjalili, H. Faris, I. Aljarah, M. Mafarja, and H. Chen, "Harris hawks optimization: Algorithm and applications," Futur. Gener. Comput. Syst., vol. 97, no. March, pp. 849–872, 2019, doi: 10.1016/j.future.2019.02.028.

[27] R. Venkata Rao, "Jaya: A simple and new optimization algorithm for solving constrained and unconstrained optimization problems," Int. J. Ind. Eng. Comput., vol. 7, no. 1, pp. 19–34, 2016, doi: 10.5267/j.ijiec.2015.8.004.

[28] N. A. M. Yusof et al., "Deep convolution neural network for crack detection on asphalt pavement," J. Phys. Conf. Ser., vol. 1349, no. 1, 2019, doi: 10.1088/1742-6596/1349/1/012020.

[29] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey Wolf Optimizer," Adv. Eng. Softw., vol. 69, pp. 46–61, 2014, doi: 10.1016/j.advengsoft.2013.12.007.

[30] B. M. Nguyen, T. Tran, T. Nguyen, and G. Nguyen, "Hybridization of Galactic Swarm and Evolution Whale Optimization for Global Search Problem," IEEE Access, vol. 8, pp. 74991–75010, 2020, doi: 10.1109/ACCESS.2020.2988717.

[31] X. Tang, "Large-scale computing systems workload prediction using parallel improved LSTM neural network," IEEE Access, vol. 7, pp. 40525–40533, 2019, doi: 10.1109/ACCESS.2019.2905634.

[32] R. Toncharoen and M. Piantanakulchai, "Traffic State Prediction Using Convolutional Neural Network," Proceeding 2018 15th Int. Jt. Conf. Comput. Sci. Softw. Eng. JCSSE 2018, pp. 1–6, 2018, doi: 10.1109/JCSSE.2018.8457359.

[33] M. Jaihuni et al., "A novel recurrent neural network approach in forecasting short term solar irradiance," ISA Trans., vol. 121, no. xxxx, pp. 63–74, 2022, doi: 10.1016/j.isatra.2021.03.043.

[34] B. Karthika, N. Umamaheswari, and R. Venkatesh, "A research of traffic prediction using deep learning techniques," Int. J. Innov. Technol. Explor. Eng., vol. 8, no. 9 Special Issue 2, pp. 725–728, 2019, doi: 10.35940/ijitee.I1151.0789S219.