A New Ensemble Reinforcement Learning Strategy for Solar Irradiance Forecasting using Deep Optimized Convolutional Neural Network Models

Seyed Mohammad J. Jalali
IISRI
Deakin University
Victoria, Australia
sjalali@deakin.edu.au

Mahdi Khodayar
Department of Comp. Sci.
University of Tulsa
Tulsa, OK, USA
mahdi-khodayar@utulsa.edu

Sajad Ahmadian
Faculty of Inf. Tech.
Kermanshah University of Technology
Kermanshah, Iran
s.ahmadian@kut.ac.ir

Miadreza Shafie-khah
School of Tech. and Innov.
University of Vaasa
Vaasa, Finland
mshafiek@univaasa.fi

Abbas Khosravi
IISRI
Deakin University
Victoria, Australia
abbas.khosravi@deakin.edu.au

Syed Mohammed S. Islam
School of Sci.
Edith Cowan University
Joondalup, WA, Australia
syed.islam@ecu.edu.au

Saeid Nahavandi
IISRI
Deakin University
Victoria, Australia
saeid.nahavandi@deakin.edu.au

João P. S. Catalão
INESCTEC
FEUP
Porto, Portugal
catalao@fe.up.pt

2021 International Conference on Smart Energy Systems and Technologies (SEST) | 978-1-7281-7660-4/21/$31.00 ©2021 IEEE | DOI: 10.1109/SEST50973.2021.9543462

***Abstract*—Solar irradiance forecasting is a major priority for the power transmission systems in order to generate and incorporate the performance of massive photovoltaic plants efficiently. As such, prior forecasting techniques that use classical modelling and single deep learning models that undertake feature extraction procedures manually were unable to meet the output demands in specific situations with dynamic variability. Therefore, in this study, we propose an efficient novel hybrid solar irradiance forecasting based on three steps. In step I, we employ a powerful variable input selection strategy named as partial mutual information (PMI) to calculate the linear and non-linear correlations of the original solar irradiance data. In step II, unlike the traditional deep learning models designing their architectures manually, we utilize several deep convolutional neural network (CNN) models optimized by a novel modified whale optimization algorithm in order to compute the forecasting results of the solar irradiance datasets. Finally in step III, we deploy a deep Q-learning reinforcement learning strategy for selecting the best subsets of the combined deep optimized CNN models. Through analysing the forecasting results over two USA solar irradiance stations, it can be inferred that the proposed optimized deep RL-ensemble framework (ODERLEN) outperforms other powerful benchmarked algorithms in different time-step horizons.**

***Index Terms*—Solar irradiance forecasting, Deep neural networks, Evolutionary computation, Ensemble strategy, Deep reinforcement learning.**

Nomenclature

$IWOA$ Improved whale optimization algorithm
MSE mean square error
$ODERLEN$ optimized deep RL-ensemble
PMI Partial mutual information

I. Introduction

Due to the loss of all types of renewable energies coming from solar energies, global climate change is growing stronger and the application of renewable energy has gained greater public attention. Generally, all types of renewable energies are various forms of solar energy, excluding geothermal and tidal energy. Photovoltaic (PV) generation, including network-connected PV systems and isolated PV panels, have developed rapidly in recent years all throughout the planet, and this growth will continue to expand in the future as one of the most common types of solar energy applications [1]. In the past decade, annual growth rates have been over 40% on average which makes PVs as one of the most key emerging renewable energy markets. Several market leaders predicted that revenues in the photovoltaics, innovations, and industries will be doubled, from 35-40 billion euros in 2010 to 70 billion euros in 2015 [2]. Rapid development of solar photovoltaic energy, more accurate and reliable modeling as well as predictions of solar irradiance are expected increasingly. On the other hand, the main attributes of the solar resources that need to have well characterized in order to operate effectively in photovoltaic and concentrated solar energy plants are, in particular, global horizontal irradiance (GHI). The short-term variability of the solar irradiance in changing weather is causing considerable difficulty in the balance of the power and changes of the regional power systems due to the power supply from the grid-connected photovoltaic plant following the fluctuation of solar irradiation [1]. The grid-connected power prediction of PV plants is an essential step towards solving this issue. This can offer beneficial inputs and raw data for different regional power system operation activities such as optimum energy flow, economical transportation management of grid connection and safety evaluation. Thus, accurate solar irradiance forecasting will increase the degree of transparency, safety and contribute to more economical operational electricity grid preferences. In addition, it is necessary for the established generators to schedule various energy plants for maintaining resources, and provide more details upon on the solar energy trades. Deep neural network (DNN) technologies are one of the sub fields of artificial neural networks (ANNs) that have gained huge attention from the power systems domain in the recent years [3], [4]. Among different DNN models, convolutional neural network (CNN) CNN is a promising technique for time series-based solar irradiance forecasting [5]. In [6], the authors proposed an efficient deep learning algorithm based on the CNN and long short-term memory (LSTM) for solar irradiance forecasting with a CEEMDAN model in order to decompose the solar dataset in different climate regions. This algorithm outperforms other state-of-the-art algorithms. Moreover, [7] a CNN algorithm is developed for forecasting 5 to 20 minutes ahead of solar irradiance by a dataset of collected sky images. By assessing the efficiency of utilized CNN with other compared forecasting models, the experimetnal results show that the proposed CNN algorithm is more effective for short-term solar irradiance forecasting. However, obtaining a fine-tuned set of hyperparameters for designing the CNN architectures is a challenging task which is time-consuming and generally conducts through the trial and error procedures [8]. To overcome this issue, evolutionary algorithms can be employed to find the optimal values of these hyperparameters [9]. Therefore, in this paper, we optimize the hyperparameters and architecture of CNNs using a novel optimization algorithm that we call the improved whale optimization algorithm (IWOA). After executing the IWOA in order to obtain the optimal CNN architectures, the forecasting results of all these optimized models are integrated into an ensemble strategy based on deep Q-learning reinforcement learning algorithm to output the best forecasting GHI performance. We name this optimized deep RL ensemble model as ODERLEN. The experimental results confirm that ODERLEN exhibits the best performance among other strong benchmarked algorithms for solar GHI prediction.

The rest of the paper is organized as follows: in Section II, the proposed deep hybrid ensemble model (ODERLEN) is represented in details. The description of solar irradiance datasets and the initialized setting for the proposed ODERLEN model is presented in Section III. Section IV denotes to the discussion of the obtained results from the proposed model and finally, the paper is concluded in Section V.

II. Methodology

In this section, we introduce our novel three-stage hybrid framework comprehensively. - *First step:* We first utilize the partial mutual information (PMI) strategy as a powerful input variable selection for the deep CNN models. PMI approach is fundamentally similar to the partial correlation-based model however it incorporates mutual information instead of sequential relation to select the data inputs [10]. The value of PMI is a closely related entropy between the output Y and the candidate C_j, which is, therefore, not already in S, represented by $MI\left(C_j : Y \mid S\right)$. Conditional expectation of x given S, is provided by the following formula as the non-parametric regression kernel approximation:

$$E[x \mid S=s]=\frac{1}{n} \frac{\sum_{i=1}^n x_i K_h\left(s-s_i\right)}{\sum_{i=1}^n K_h\left(s-s_i\right)} \tag{1}$$

where the selected input set is denoted by S, n represents the total number of samples, x represents the c_y or y, and the Gaussian kernel function (K_h) is expressed by:

$$K_h\left(x-x_1\right)=\frac{1}{(\sqrt{2 \pi h})^d \sqrt{|\sigma|}} \exp \left(-\frac{\left(x-x_i\right)^T \sigma^{-1}\left(x-x_i\right)}{2 h^2}\right) \tag{2}$$

where σ represents a matrix of covariance sample, and the dimensionality of x is denoted by d. The kernel bandwith which is represented by h is described by the following formula:

$$h=\left(\frac{4}{d+2}\right)^{\frac{1}{d+4}} n^{\frac{-1}{d+4}} \tag{3}$$

Therefore, the PMI value is computed by the following formula:

$$PMI\left(C_j ; Y \mid S\right)=MI(u ; v) \approx \frac{1}{n} \sum_{i=1}^n \log_e\left[\frac{f(u, v)}{f(u) f(v)}\right] \tag{4}$$

where $u=Y-Y(S)$ and $v=C_j-\widehat{C}_j(S)$ in which S shows a set of selected inputs and j represents a set of candidate inputs by using the estimators of the non-parametric kernel indicated by $C_1(S)=E\left[c_j \mid S=s\right]$ and $\varphi(S)=E[y \mid S=s]$. It should be noted that variable φ denotes to the sample covariance matrix, f(u, v), f(u) and f(v) represent probability density functions of u, v, and joint u and v, respectively, and the variables of $\widehat{C}_j$ and $\hat{Y}$ represent the linear least squares regression estimates for the C_j and Y. - *Second step:* The second step optimizes the hyperparameters of the base CNN models with a novel optimization model based on the theorem of evolutionary computation. Since whale optimization algorithm (WOA) has shown excellent performance in numerous engineering applications [11], we further improve its capabilities using the chaotic map (CM) [12] and evolutionary boundary constraint handling (EBCH) [13] strategies. The following formulas are presented throughout the optimisation to mathematically model the surrounding phenomenon.

$$\overrightarrow{D}=\left|\overrightarrow{C} \cdot \overrightarrow{X^*}(t)-\overrightarrow{X}(t)\right| \tag{5}$$

$$\overrightarrow{X}(t+1)=\overrightarrow{X^*}(t)-\overrightarrow{A} . \overrightarrow{D} \tag{6}$$

where the current iteration is represented by t, $\overrightarrow{A}$ and $\overrightarrow{C}$ denote to coefficient vectors, $X(t)$ denotes to the position vector (a random whale) and X^* represents the optimal solution position vector that has been so far achieved. The $\overrightarrow{A}$ and $\overrightarrow{C}$ coefficient vectors are determined according to:

$$\overrightarrow{A} = 2\overrightarrow{a} \cdot \overrightarrow{r}-\overrightarrow{a} \tag{7}$$

$$\overrightarrow{C} = 2 \cdot \overrightarrow{r} \tag{8}$$

where $\overrightarrow{a}$ is declined linearly from 2 to 0 during iterations, and r represent a random vector in the interval of [0, 1]. The procedure of updating position of each search agents based on the spiral (to simulate bubble-net attacking mechanism of humpback whales) is given numerically as follows:

$$\overrightarrow{D'}=\left|\overrightarrow{X^*}(t)-\overrightarrow{X}(t)\right| \tag{9}$$

$$\overrightarrow{X'}(t+1) = \overrightarrow{D'} \cdot e^{bl} \cdot \cos(2 \pi l)+\overrightarrow{X^*}(t) \tag{10}$$

where $\overrightarrow{D'}$ determines the difference to the best solution from the ith search agent, b denotes to a parameter used to describe logarithmic spiral modes and l represents a randomized value of the [-1,1] range. for further simplification, we generally suppose that Eq. (2) or Eq. (6) would update the position of the search agents, each having a 50% probability, that can be given by the following mathematical formulas:

$$\overrightarrow{X}(t+1)=\begin{cases}\overrightarrow{X^*}(t)-\overrightarrow{A} \cdot \overrightarrow{D} & if\ p<0.5 \\ \overrightarrow{D'} \cdot e^{bl} \cdot \cos(2 \pi l)+\overrightarrow{X^*}(t) & if\ p>0.5\end{cases} \tag{11}$$

where the variable p represents a random value within the interval of [0, 1]. Using the chaotic circle map, we enforce the improvement to the traditional WOA. Compared to WOA, the proposed chaotic operator can prevent local Optima from being stuck and also increase its convergence speed. The proposed chaotic WOA is developed using the chaotic circle operator to model the value of p parameter as follows:

$$p_{i+1}=\bmod \left(p_i+b-(a / 2 \pi) \sin(2 \pi p_i), 1\right); \quad a=0.5, \quad b=0.2 \tag{12}$$

where the initialized values for the parameters a and b are chosen based on the reference WOA article [11]. The position of search agents can be constantly updated for every iteration in the WOA algorithm. After the updating procedure is completed, search agents that go outside the search area are considered useless for the obtained best solutions. The search agents therefore, be redirected to the search area. The technique can overcome this issue is named as evolutionary boundary constraint handling (EBCH). Once this approach has been integrated into WOA, the compatibility to exploration and exploitation is effectively improved. The mechanism of EBCH updating is as formulated as follows:

$$f(ob_i \to x_i) = \begin{cases} \xi \cdot \mathrm{lb}_i + (1-\xi)x_i^b & \text{if } ob_i < \mathrm{lb}_i \\ \varrho \cdot \mathrm{ub}_i + (1-\varrho)x_i^b & \text{if } ob_i > \mathrm{ub}_i \end{cases} \tag{13}$$

where ξ and ϱ represent two random variables in the range of [0,1], ob_i denotes to the search agent that is from the out-of-bound, x_i^b represents the ith best search agent and the upper bound and lower bound of the search space is presented by ub_i and lb_i denote, respectively. The flowchart of the proposed chaotic-EBCH WOA called as improved WOA (IWOA) is illustrated in Fig. **??**. Now, it is the time to optimize the hyperparameters of CNN models using the IWOA strategy. Eight important CNN hyperparameters as tabulated in Table I which have the critical role in the architectural design of CNNs are optimized by the IWOA algorithm. Thus, in the solution space of the IWOA, each one of the eight hyperparameters of a CNN model is interpreted as a eight dimensional vector. The continuous variables in the CNN hyperparameters are transferred as $D = [Hyp_1, Hyp_2, ..., Hyp_n]$ into a discreet search space. The following formulas are taken into account for formulating the discretization model:

$$\Lambda = 1 + n \times K \tag{14}$$

$$\omega = min(\lfloor \Lambda \rfloor, n) \tag{15}$$

where K is a continuous variable in the $[0, 1]$ exploration range for the search space, Λ is a mapping operator of K to $[1, n+1]$ and ω represents another Λ mapping operator to $[1, 2, 3, ..., n]$ interval. Any integer value that belongs to the continuous dimension of the solution can thus be determined using the following equation:

$$X_{ij} = H_\omega \tag{16}$$

where X_{ij}, where $i = 1..n$ and $j = 1..8$ represents a 8-dimensional vector standing for the i^{th} solution encoding the eight CNN hyperparameters and H_ω denotes to these hyperparameters mapped from discretization manner. Based on the obtained values of hyperparameters, we consider a fitness function to test the efficiency of the configured CNN architectures for GHI forecasting. In this regard, we take into consideration the mean square error (MSE) to calculate the fitness value of each IWOA solutions given by the following equation:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - y_i')^2 \tag{17}$$

where y_i and y_i' are the actual and forecasted GHI values by the CNN model. The aim of the proposed method is therefore to elicit the lowest MSE value solution containing the CNN hyperparameter optimal values. This means that a CNN model has been achieved with the best performance in the GHI test set. - *Third step:* By obtaining the best CNN models based on the optimal hyperparameters, the reinforcement learning strategy is adopted in the third stage to achieve an optimum sub-set of regression models which can be deployed in ensemble strategy as the final selected base regression model. Reinforcement learning is recognized as a dynamic technique of learning that focuses on the mechanism of environment interactions. The key concept behind this learning model is to determine the best solution using the trials and errors. The Q-learning method is among the most common reinforcement learning algorithms based on Q values being updated in the environment. We utilize this approach to identify ensemble with the proposed method because of flexibility and effective convergence of the Q-learning mechanism. In other words, we employ Q-learning to achieve an optimal subset of optimized CNN models in order to enhance the accuracy of the proposed ensemble model.

We apply the bagging mechanism to create a collection of optimized CNN regression models ($Bagg_M = \{b_1, b_2, \ldots, b_M\}$) where M denotes to number of optimized CNN models and the principle is to use reinforcement training strategy to choose an optimal sub-group of the optimized CNN models called $Bagg'_{M'}$. We require to specify the states set for S and the actions set for A to undertake reinforcement learning strategy. Each state in the proposed model is represented by a $s_t = [L_t, MSE_t]$ tuple in which L_t refers to the number of selected CNN models optimized by the evolutionary algorithm in the s_t state, and MSE_t belongs to the MSE as the error metric of the ensemble regression model built by the selected optimized CNN regression models. The outputs of the optimized CNN models are combined based on a simple voting approach to measure the output performance of the ensemble model. The proposed method is designed to ensure that an ensemble regression model with minimum error is obtained, using the Q-Learning algorithm by choosing an efficient subgroup of the optimized CNN models. The reinforcement Learning algorithm maps the state to an action described as the $\Pi : S \to A$ policy [14], [15]. The Q-Learning approach is founded on a function-value named as $Q^\Pi(s_t, a_t)$, where the states and actions are respectively represented by s_t and a_t. The overall estimated discount reward value depending on the optimal Π strategy is defined by the following equation:

$$Q^\pi(\mathbf{s}, a) = E\left(\sum_{k=0}^{\infty} \gamma^i r_t \mid \mathbf{s}_0 = \mathbf{s}, a_0 = a, \pi\right) \tag{18}$$

where γ corresponds to the discount coefficient of $0 \leq \gamma \leq 1$. Using the Q-learning algorithm, the optimized action-value function $Q^*(s, a)$ can be quantified by:

$$Q^*(\mathrm{s}, a) = E\left(r_{t+1} + \gamma \max_{a'} Q^*\left(\mathrm{s}_{t+1}, a'\right) \mid \mathrm{s}_t = \mathrm{s}, a_t = a\right) \tag{19}$$

The technique proceeds with a series of episodes in which the action value function Q is updated based on the following formula to determine an optimal search strategy:

$$\begin{aligned} Q\left(\mathbf{s}_t, a_t\right) \leftarrow Q\left(\mathbf{s}_t, a_t\right) + \alpha\left[r_{t+1} + \lambda \max_a Q\left(\mathbf{s}_{t+1}, a\right)\right. \\ \left. - Q\left(\mathbf{s}_t, a_t\right)\right] \end{aligned} \tag{20}$$

The desirable subgroup of the optimized CNN regression models $Bagg'_{L'}$ can be defined based on the best result achieved, after the Q-learning technique has been executed. This optimum subset is subsequently applied to yield the proposed ensemble GHI time series forecasting strategy as a final frame for the optimized CNN models. We show the entire framework of our optimized deep RL ensemble models briefly named as ODERLEN in Fig. 1 and from technical point of view in Algorithm 1.

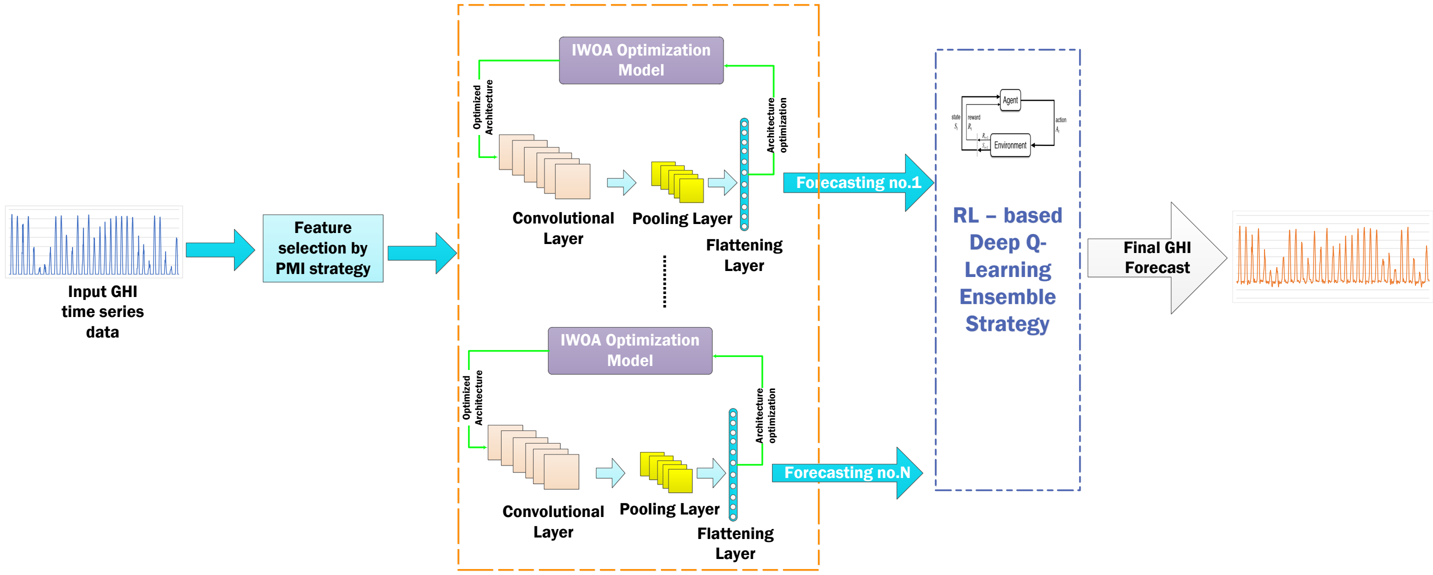

Fig. 1: An overall schema of the novel ODERLEN framework.

Algorithm 1 The pseudo-code of the proposed deep ODERLEN model for solar GHI forecasting.

```
Input: N (Population size), GEN (Maximum number of generations), and L (Number of base regression models).
Output: Predicted GHI values.
Begin algorithm:
Split GHI dataset into training Tr and testing Te sets;
Generate a set bag of base regression models Bagg_M = {b_1, b_2, ..., b_M};
Set m = 1;
while (m < L) do
    Generate a random initial population X_i (i=1,2,..., N);
    Set g=1;
    while (g < GEN) do
        Set a CNN model for each solution based on their hyperparameter values;
        Calculate the population fitness using Eq.(17) as the MSE of CNN algorithm obtained by Tr set;
        for each search agent do
            Update a, A, C and p using Eq.(12) ;
            if (|A|<1) then
                Update the position of search agent using Eq.(5);
            else if (|A|≥1) then
                Update the position of the search agent by the Eq.(11);
            end if
        end for
        Perform the EBCH operator using Eq.(13);
        Calculate the fitness of all search agents and update X* if a better solution is found.
        Set g=g+1;
    end while
    Consider the b_m CNN regression model with hyperparameters obtained by the best search agent;
    Set m=m+1;
end while
Perform deep Q-learning model over the CNN regression model Bagg_M in order to chose an optimal subset of the regression models indicated by Bagg'_M';
Apply the ensemble strategy to forecast the GHI points in the test set Te using the selected regression models Bagg'_M';
Return the predicted GHI as the output;
End algorithm
```

III. Experimental setups and solar irradiance datasets

In order to execute ODERLEN framework, we program it with python using the powerful deep learning libraries including Keras and TensorFlow using a machine with one 16 GB RAM, one GPU of GeForce GTX 1080 Ti and the Ubuntu operating system. For configuring the evolutionary IWOA, we chose the number of population to 20 and maximum iteration number to 30. It should be mentioned that the proposed ODERLEN model and the other competitor models are executed 10 times. Furthermore, the CNN hyperparameters that are evolved with IWOA are described in Table I. We optimize eight hyperparameters during the constructing of the optimal CNN architectures including number of filters (N_f), kernel size (K_s), maxpooling size (MP_s), batch size (B_s), number of convolutional layers (N_c), number of epochs (N_e), dropout rate (D_r), and learning rate (L_r). Moreover, we chose powerful Adam optimizer and Relu as the activation function during the training process. For RL ensemble configurations, we set the number of episodes in the Q-learning algorithm to 200 and consider 15 base optimized CNN regression models for selecting an optimal subset of regression models with deep RL ensemble strategy. In order to show the competitiveness of our proposed ODERLEN hybrid framework, we compare it with five hybrid powerful state of the art deep learning models including adaptive hybrid model (AHM) [16], hybrid feature selection method (HFS) [17], Outlier-robust hybrid model (ORHM) [18], novel hybrid deep neural network model (NHDNNM) [19], OHS-LSTM [20] that have shown their strength in time-series forecasting problems. Also, in order to demonstrate the search capabilities of our proposed ODERLEN algorithm, we use powerful evolutionary algorithms such as genetic algorithm (GA), particle swarm optimization (PSO), ant colony optimization algorithm (ACO), biogeography-based optimization (BBO) and whale optimization algorithm (WOA) on the framework proposed in this work. These refer to cases in which the WOA algorithm is replaced by another optimisation method (GA, PSO, etc.) in the ODERLEN framework.

TABLE I: The symbols and corresponding values of the optimized CNN hyperparameters with the evolutionary IWOA model.

Symbol	Value
N_f	[1, 600]
K_s	[1, 30]
N_e	[1, 500]
N_c	[1, 2,..., 20]
MP_s	[1, 30]
B_s	[10, 20,..., 400]
D_r	[0.2, 0.25,..., 0.65]
L_r	[0.001, 0.006,.., 0.1]

We use National Renewable Energy Laboratory [21] solar irradiance datasets collected from two solar stations located in Los Angeles and Phoenix in the western side of United states. Each of these two solar stations contain 8760 GHI time series data points in one-hour intervals for the year 2018 in which the GHI values are normalised. We consider 75% of each dataset as the training set and the remaining 25% is allocated to test set. We should note that the 25/75% split of the training and test sets is not random, thus that both sets are made of distinct days. Based on these two sets of solar irradiance data, GHI has an increase from 8:00 to 13:00, and then, has a decrease until it meets zero from about 18:00 to 20:00. For selecting the input features of Deep optimized CNNs, we utilize the partial mutual information strategy (PMI). The PMI values considered above of the threshold equal to $\Xi = 0.4$ is chosen which results in 57 input GHI features for training of the deep learning models. The root mean square error (RMSE) and mean absolute error (MAE) which are widely used evaluation metrics in the literature are considered to calculate the accuracy of the forecasting models.

IV. Experimental Results

In this section, we test the forecasting performance of our proposed algorithm compared to other powerful benchmark algorithms. Tables II and III show that in all testing GHI data sets, the RMSE and MAE of our proposed model are smaller than all other ten benchmark models.

For instance, we notice that the RMSE OF 1-step ahead prediction in Phoenix station for the proposed ODERLEN equal to 0.034237 significantly outperforms in comparison with HFS as the closest competitor algorithm equal to 0.034436 , while the MAPE of the proposed model equal to 0.015165 significantly outperform the GA-RL-Ens model equal to 0.015421 as the most compatibale model to ORDELEN. Furthermore, for the LA test dataset for the proposed method with a RMSE value of 0.033629 and also for the metric error MAE with a value of 0.016869, the ODERLEN model has the lowest value among all the ten powerful competitive methods compared. It can be noted that in comparison with AHM, HFS, ORHM, NHDNNM and HOS-LSTM as the state of the art models for time series energy forecasting problems, we find that the accuracy of our proposed ODERLEN model is much more reliable than these recent powerful models. Besides, the predicted performance of the proposed model is still good compared to other hybrid evolutionary-RL ensemble models. These results indicate that the ODERLEN algorithm can capture complex solar GHI characteristics compared to other hybrid deep optimized CNNs by RL ensemble models.

If a forecasting algorithm can match the actual and predicted points optimally, it indicates the obvious strength of that method. As can be seen from Figs. 2 and 3, our proposed ODERLEN algorithm matches well the blue dots that represent the actual data points and the red dots represent the predicted data points for each of the next four different time-steps. The important point should be noticed from Tables II and III, as well as Figs. 2 and 3, show is that as the time-steps of the next step increases, the prediction made by the our proposed ODERLEN model becomes more and more difficult, and in this case, our algorithm has the least amount of error metric values in comparison to other ten benchmark models. In Figs. 5 and 4, the diagrams generated by the proposed ODERLEN algorithm using the next four different forecasting time-steps are demonstrated. The obvious highlight from these diagrams refers to the high convergence speed of the ODERLEN model in order to find the optimal solution, which is due to the use of two efficient operators including chaotic circle map and EBCH on the WOA algorithm.

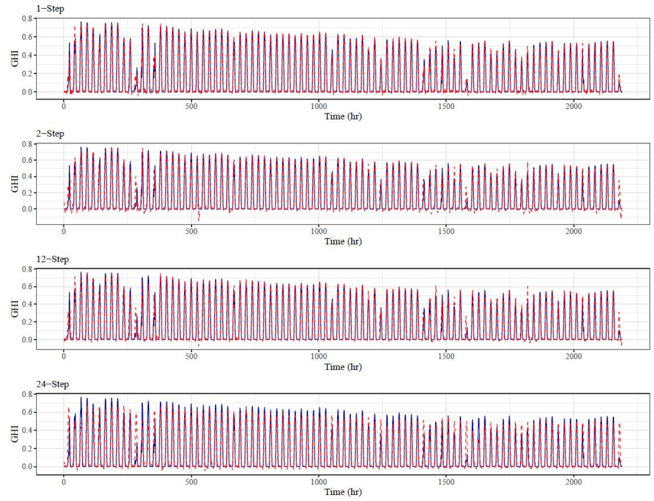

Fig. 2: Actual vs predicted points for Phoenix GHI test dataset.

In summary, this work demonstrates that by applying two efficient chaotic circle map and EBCH operators on the WOA algorithm, the hyperparameters of the deep CNN models are perfectly optimized, and also, the deep RL Q-learning algorithm selects the best subset of optimal solutions for solar irradiance GHI forecasting, which indicates the strength performance of our proposed ODERLEN hybrid model.

TABLE II: The performance results of RMSE and MAE metrics for Phoenix dataset.

Model	1-step		2-step		12-step		24-step	
	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
AHM	0.037729	0.020656	0.049905	0.024177	0.074583	0.034471	0.093377	0.045822
HFS	0.034436	0.016075	0.051074	0.024943	0.078714	0.034137	0.095067	0.044216
ORHM	0.035758	0.018975	0.053833	0.026384	0.073988	0.032091	0.094503	0.043452
NHDNNM	0.035562	0.015896	0.050868	0.023332	0.077522	0.032355	0.095013	0.042383
OHS-LSTM	0.036146	0.019742	0.051916	0.024527	0.072677	0.033397	0.091561	0.043514
GA-RL-Ens	0.034745	0.015421	0.051341	0.024664	0.074939	0.033543	0.094074	0.044231
PSO-RL-Ens	0.035228	0.016573	0.051968	0.023416	0.079238	0.035511	0.092422	0.045645
ACO-RL-Ens	0.039853	0.019392	0.060154	0.034946	0.080012	0.034056	0.095969	0.043643
BBO-RL-Ens	0.038241	0.020397	0.049899	0.023672	0.072099	0.032611	0.094288	0.042522
WOA-RL-Ens	0.035012	0.016592	0.053267	0.025107	0.083558	0.039762	0.096543	0.048172
ODERLEN	**0.034237**	**0.015165**	**0.049541**	**0.023021**	**0.070361**	**0.031977**	**0.091424**	**0.041545**

TABLE III: The performance results of RMSE and MAE metrics for Los Angeles dataset.

Model	1-step		2-step		12-step		24-step	
	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
AHM	0.034389	0.018903	0.053352	0.033071	0.067799	0.036235	0.082938	0.040784
HFS	0.034481	0.017176	0.047231	0.029983	0.067668	0.041654	0.083894	0.046652
ORHM	0.034641	0.017746	0.045537	0.024858	0.069499	0.034732	0.085318	0.038183
NHDNNM	0.039785	0.027732	0.046621	0.026433	0.076965	0.042568	0.092854	0.047328
OHS-LSTM	0.034156	0.017949	0.048055	0.028697	0.067045	0.034067	0.082507	0.040785
GA-RL-Ens	0.033988	0.017232	0.053572	0.034503	0.066893	0.031952	0.081635	0.036333
PSO-RL-Ens	0.036811	0.020526	0.046609	0.024815	0.075352	0.041173	0.089168	0.046892
ACO-RL-Ens	0.036782	0.022753	0.060726	0.041971	0.067655	0.033035	0.082322	0.039572
BBO-RL-Ens	0.034034	0.017022	0.046788	0.023352	0.066531	0.035568	0.085281	0.039342
WOA-RL-Ens	0.034246	0.018471	0.062522	0.038106	0.078492	0.042145	0.092705	0.048375
ODERLEN	**0.033629**	**0.016869**	**0.044447**	**0.023143**	**0.066474**	**0.031438**	**0.081286**	**0.035824**

V. Conclusion

In this research work, a hybrid model adaptive to three stages including PMI input selection strategy, the optimized deep CNNs and deep Q-learning RL algorithm called ODERLEN for predicting solar irradiance is introduced. To validate the efficacy of the hybrid ODERLEN model, the GHI data points obtained from two

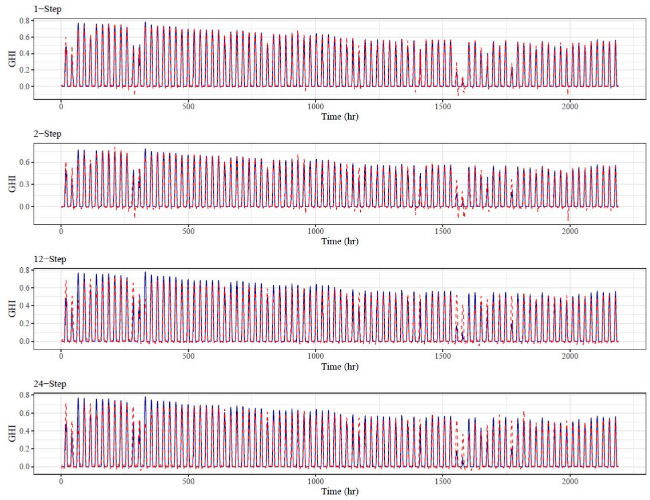

Fig. 3: Actual vs predicted points for LA GHI test dataset.

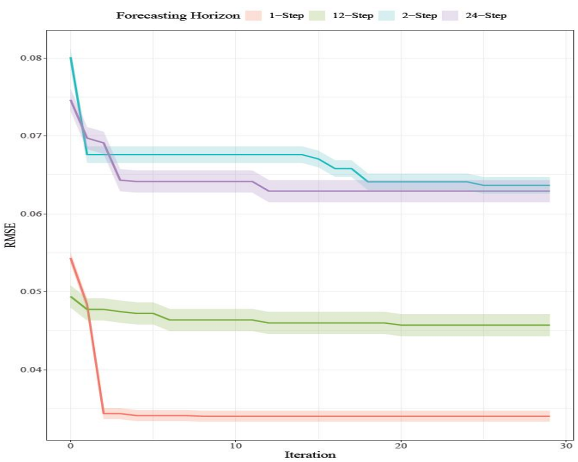

Fig. 4: The convergence curves of the ODERLEN model for legend horizons of LA dataset.

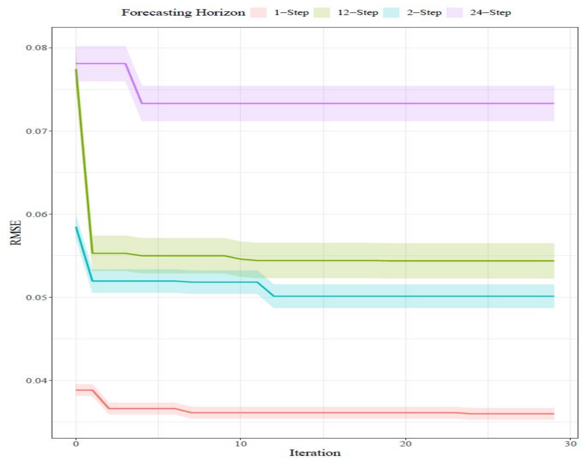

Fig. 5: The convergence curves of the ODERLEN model for legend horizons of Phoenix dataset.

separate solar stations which are near to the cities of Phoenix and Los Angeles (LA) are utilized. Meanwhile, in different time-span scenarios, ten well-known and recently published algorithms are fairly compared with our proposed framework. The experimental findings show that our proposed ODERLEN model can increase the forecasting accuracy noticeably and indicate that the model proposed is also more robust than other compared benchmark models.

REFERENCES

[1] M. Khodayar, M. E. Khodayar, and S. M. J. Jalali, "Deep learning for pattern recognition of photovoltaic energy generation," *The Electricity Journal*, vol. 34, no. 1, p. 106882, 2021.

[2] A. Jäger-Waldau, "Pv status report 2011. research, solar cell production and market implementation of photovoltaics. jrc scientific and technical reports, eur 24807 en," *Institute for Energy/Joint Research Centre/European Commission. Office for Official Publications of the European Communities, Luxembourg. Available at http://ie. jrc. ec. europa. eu. Accessed Mar*, 2012.

[3] S. M. J. Jalali, S. Ahmadian, M. Khodayar, A. Khosravi, V. Ghasemi, M. Shafie-khah, S. Nahavandi, and J. P. Catalão, "Towards novel deep neuroevolution models: chaotic levy grasshopper optimization for short-term wind speed forecasting," *Engineering with Computers*, pp. 1–25, 2021.

[4] S. M. J. Jalali, P. M. Kebria, A. Khosravi, K. Saleh, D. Nahavandi, and S. Nahavandi, "Optimal autonomous driving through deep imitation learning and neuroevolution," in *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*. IEEE, 2019, pp. 1215–1220.

[5] S. M. J. Jalali, S. Ahmadian, A. Kavousi Fard, A. Khosravi, and S. Nahavandi, "Automated deep cnn-lstm architecture design for solar irradiance forecasting," *IEEE Transactions on Systems, Man, and Cybernetics - Systems*, vol. 1, pp. 1–12, 2021.

[6] P. Kumari and D. Toshniwal, "Long short term memory–convolutional neural network based deep hybrid approach for solar irradiance forecasting," *Applied Energy*, vol. 295, p. 117061, 2021.

[7] A. Ryu, M. Ito, H. Ishii, and Y. Hayashi, "Preliminary analysis of short-term solar irradiance forecasting by using total-sky imager and convolutional neural network," in *2019 IEEE PES GTD Grand International Conference and Exposition Asia (GTD Asia)*. IEEE, 2019, pp. 627–631.

[8] S. M. J. Jalali, S. Ahmadian, A. Khosravi, M. Shafie-khah, S. Nahavandi, and J. P. Catalao, "A novel evolutionary-based deep convolutional neural network model for intelligent load forecasting," *IEEE Transactions on Industrial Informatics*, 2021.

[9] S. J. Mousavirad, S. M. J. Jalali, S. Ahmadian, A. Khosravi, G. Schaefer, and S. Nahavandi, "Neural network training using a biogeography-based learning strategy," in *International Conference on Neural Information Processing (ICONIP)*. Springer, 2020, pp. 147–155.

[10] E. Snieder, R. Shakir, and U. Khan, "A comprehensive comparison of four input variable selection methods for artificial neural network flow forecasting models," *Journal of Hydrology*, vol. 583, p. 124299, 2020.

[11] S. Mirjalili and A. Lewis, "The whale optimization algorithm," *Advances in engineering software*, vol. 95, pp. 51–67, 2016.

[12] M. Kohli and S. Arora, "Chaotic grey wolf optimization algorithm for constrained optimization problems," *Journal of computational design and engineering*, vol. 5, no. 4, pp. 458–472, 2018.

[13] A. H. Gandomi and A. R. Kashani, "Evolutionary bound constraint handling for particle swarm optimization," in *2016 4th International Symposium on Computational and Business Intelligence (ISCBI)*. IEEE, 2016, pp. 148–152.

[14] M. Sun, C. Feng, and J. Zhang, "Multi-distribution ensemble probabilistic wind power forecasting," *Renewable Energy*, vol. 148, pp. 135–149, 2020.

[15] C. Feng, M. Sun, and J. Zhang, "Reinforced deterministic and probabilistic load forecasting via *q*-learning dynamic model selection," *IEEE Transactions on Smart Grid*, vol. 11, no. 2, pp. 1377–1386, 2019.

[16] J. Zhang, Z. Tan, and Y. Wei, "An adaptive hybrid model for short term electricity price forecasting," *Applied Energy*, vol. 258, p. 114087, 2020.

[17] X. Zhang, J. Wang, and Y. Gao, "A hybrid short-term electricity price forecasting framework: Cuckoo search-based feature selection with singular spectrum analysis and svm," *Energy Economics*, vol. 81, pp. 899–913, 2019.

[18] J. Wang, W. Yang, P. Du, and T. Niu, "Outlier-robust hybrid electricity price forecasting model for electricity market management," *Journal of Cleaner Production*, vol. 249, p. 119318, 2020.

[19] C.-J. Huang, Y. Shen, Y.-H. Chen, and H.-C. Chen, "A novel hybrid deep neural network model for short-term electricity price forecasting," *International Journal of Energy Research*, 2020.

[20] S. Zhou, L. Zhou, M. Mao, H.-M. Tai, and Y. Wan, "An optimized heterogeneous structure lstm network for electricity price forecasting," *IEEE Access*, vol. 7, pp. 108 161–108 173, 2019.

[21] M. Sengupta, Y. Xie, A. Lopez, A. Habte, G. Maclaurin, and J. Shelby, "The national solar radiation data base (nsrdb)," *Renewable and Sustainable Energy Reviews*, vol. 89, pp. 51–60, 2018.