Timeseries based deep hybrid transfer learning frameworks: a case study of electric vehicle energy prediction

Paul Banda, Muhammed A. Bhuiyan, Kazi N. Hasan, Kevin Zhang, and Andy Song

RMIT University, School of Engineering, Melbourne 3000, Australia

Abstract. The problem of limited labelled data availability causes underfitting, which negatively affects the development of accurate time series based prediction models. Two-hybrid deep neural network architectures, namely the CNN-BiLSTM and the Conv-BiLSTM, are proposed for time series based transductive transfer learning and compared to the baseline CNN model. The automatic feature extraction abilities of the encoder CNN module combined with the superior recall of both short and long term sequences by the decoder LSTM module have shown to be advantageous in transfer learning tasks. The extra ability to process in both forward and backward directions by the proposed models shows promising results to aiding transfer learning. The most consistent transfer learning strategy involved freezing both the CNN and BiLSTM modules while retraining only the fully connected layers. These proposed hybrid transfer learning models were compared to the baseline CNN transfer learning model and newly created hybrid models using the R^2, MAE and RMSE metrics. Three electrical vehicle data-sets were used to test the proposed transfer frameworks. The results favour the hybrid architectures for better transfer learning abilities relative to utilising the baseline CNN transfer learning model. This study offers guidance to enhance time series-based transfer learning by using available data sources.

Keywords: Hybrid Deep Learning · Electric Vehicle Load Prediction · Transfer Learning

1 Introduction

Due to the environment-friendly policies and emission reduction schemes, the transportation sector is expected to go through a significant transformation with the adoption of many electric vehicles (EVs) into the existing vehicle fleet [8]. The usage of EVs is still in the early stages; thus, the EV charging demand data is scarce. Hence, it is challenging to perform the EV impact studies with the EV data that accurately represents the EV charging profiles [9]. Moreover, the deployment of more public EV charging stations with a large capacity charging requirement would pose a new challenge to the secure operation of the power grid [21]. In this perspective, adequately capturing the EV charging demand and

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77977-1_20

accurately predict the near future charging demand is critical for the electrical power grid's secure operation.

The implementation of forecasting techniques for EV charging demand prediction has been conducted by using fuzzy clustering and back-propagation neural network models in [22], by employing evolutionary optimisation techniques in [14] and by examining various data mining techniques [20]. An unsupervised algorithm was proposed in [13] for non-intrusive EV load extraction. A comparison of the traditional time series methods and machine learning methods for EV load prediction was presented in [4]. Machine learning methods were concatenated to give an ensemble method that performed better than individual machine learning methods in [1] for residential EV load prediction. A comparative assessment of the deep learning methods for EV charging demand prediction has been performed in [23]. The sub-hourly and hourly EV charging load prediction models have been developed using a hybrid lion algorithm consist of convolutional neural network (CNN) and long-short-term memory (LSTM) inspired models in [12]. The above studies have demonstrated effective EV prediction methods; however, the only drawback is that they require independent and identical distribution to exist in a data-set and that they must be enough training data to learn a good model. To counter these limitations, we introduce transfer learning for EV load prediction, which is not bound by the mentioned limitations.

1.1 Transfer Learning Models Review

Transfer learning using deep learning models has received great attention in various application domains such as image processing [18], time series classification, [7], natural language processing tasks [2], [17] and building energy forecasting [5]. The CNN model has been the dominant model facilitating transfer learning in most studies.

In non-transfer learning studies [10],[19],[15], hybrid deep learning models have been shown to outperform the conventional CNN model as they leverage on the advantages of the encoder and decoder modules during operations. In the same line of thought, the authors propose using hybrid deep learning model architectures, namely, the CNN-BiLSTM and Conv-BiLSTM models for transfer learning and compare their performance to the commonly used CNN transfer learning model. Furthermore, most transfer learning studies seem to report transfer methods is isolation, without showing a comparative assessment to determine which transfer learning method shows superior results, an observation which is addressed in this study. The Conv-BiLSTM model can capture salient spatial features using a convolution operator within the LSTM cell on multiple-dimensional data and has the extra ability to process in both forward and backward directions (bidirectional) is helpful for transfer learning. The CNN-BiLSTM model can leverage the automatic feature extraction advantages of the encoder CNN module and the superior recall of both short and long term sequences by the decoder LSTM module, which processes both forward and backward directions.

The study proposes the implementation of CNN-BiLSTM and Conv-BiLSTM hybrid deep learning architectures for time series transfer learning. Also, the

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77977-1_20

study introduces transductive transfer learning for electrical vehicle load prediction to enhance prediction efforts in limited labelled EV load data situations.

The rest of the paper is organised as follows: Section 2 briefly introduces transfer learning and the proposed hybrid transfer learning models, Section 3 describes the three case study data-sets, Section 4 discusses the transfer learning results and discussion, and Section 5 highlights the conclusions and future works.

2 Methodology

This section provides a brief description of the application of transfer learning for EV load prediction. It also discusses the proposed hybrid architectures and how they are modified for implementation in transfer learning.

2.1 Transfer Learning

Given a source domain EV data-set (data-rich), a target domain (limited labelled data) EV data-set and a learning task, transfer learning seeks to improve the learning of the target predictive function in the target domain EV data-sets (slow and fast commercial EV charging stations (CEVCS)) using the knowledge learnt from the source domain (residential EVCS) data-set. The formal expression to define transfer learning is given as; A domain comprises feature space X and label space Y, thus given a source domain

$$D_s = \{x_S^i, {y^i}_S\}_{i=1}^{N_S} \tag{1}$$

and a target domain

$$D_T = \{x_T^i, {y^i}_T\}_{i=1}^{N_T} \tag{2}$$

where $N_S > N_T$ and N is the labelled data size. It is challenging for a model to learn well using little data in the target domain. Since the source and target domains have different data distributions, it is unlikely for a model trained on the source domain to predict the target domain's test data-set accurately. Instead of creating two separate models for the source and target domain, as usually done in traditional machine learning, transfer learning seeks to utilise the knowledge learnt on the source domain to help predict the data domain.

The proposed hybrid transfer learning workflow illustrated in Fig. 1 applies to both the CNN-BiLSTM and Conv-BiLSTM hybrid models.

Following the above formal transfer learning definition, the transfer learning procedure implementation in this study is thus summarised as below;

(i) Pre-process data and develop the hybrid deep neural networks for source domain (residential EVCS) prediction with full data complement, (ii) Train the pre-trained models (fine-tuning) with limited data from the target domain (slow and fast CEVCS) (iii) Develop new neural networks models from scratch for the target stations with limited data (same data size as used in step II), (iv) Compare the pre-trained model's (ii) results to the new target models developed from scratch for both slow and fast CEVCS.

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77977-1_20

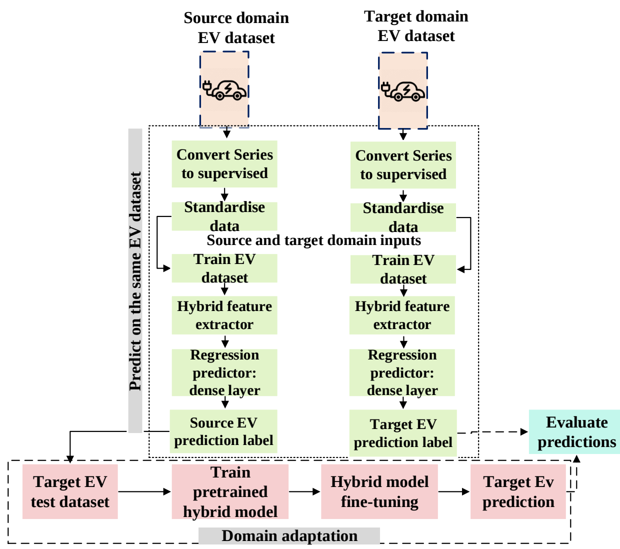


Fig. 1: An illustration of the implemented hybrid deep transfer learning workflow.

2.2 Hybrid Deep Learning Models for Transfer Learning

This section describes the proposed models for transfer learning using the available three EV data-sets. Hybrid deep learning models are designed by the fusion of conventional deep learning models that combine multiple models by diversifying the input features and varying the initialisation of the neural network's weights.

2.3 Baseline Convolutional Neural Networks (CNN)

Originally designed to handle image data, the CNN model has achieved a state of art results in image classification and object recognition, among other tasks. The CNN model can extract useful features automatically from time series data by treating a sequence of observations as a one-dimensional image. The CNN model is relatively easy and faster to train because the weights are less than those of a fully connected artificial neural network architecture. A predetermined number of historical EV energy consumption observations are fed into the convolutional layers, which perform a one-dimensional convolution operation on this data, whose output is then passed to the fully connected layers for final processing. A CNN operation is described in detail in [11];

2.4 Hybrid Convolutional Bidirectional Long Short-Term Memory (Conv-BiLSTM)

The hybrid Conv-BiLSTM does a convolution operation within the LSTM cells. The Conv-BiLSTM layer is a recurrent layer (same as LSTM) that replaces the usual matrix multiplication operation with a convolutional process. The convolution operator is applied directly to read input into the BiLSTM cells, that is, during the input-to-state transitions and during the state-to-state transitions [16]. The Conv-BiLSTM compresses the EV sequence into a hidden state tensor decodable by an LSTM layer that processes this input in both forward and backward directions (bidirectional), forwarding its output to the fully connected layer for final prediction. The critical equations of the Conv-BiLSTM cell gates are given in equations (3-7) below;

$$I_t = \sigma(W_{XI} * X_t + W_{HI} * H_{t-1} + W_{CI} \circ C_{T-1} + b_I) \tag{3}$$

$$O_t = \sigma(W_{XO} * X_t + W_{HO} * H_{T-1} + W_{CO} \circ C_t + b_0) \tag{4}$$

$$F_t = \sigma(W_{XF} * X_t + W_{HF} * H_{t-1} + W_{CF} \circ C_{t-1} + b_F) \tag{5}$$

$$C_t = F \circ C + i_t \circ (W_{XC} * x_t + W_{HC} * h_{t-1} + b_C) \tag{6}$$

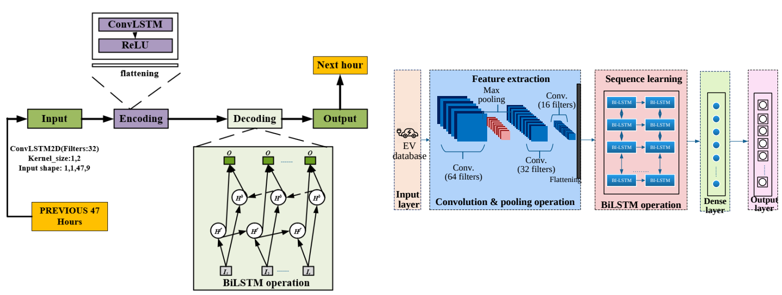
$$H_t = O \circ tanh(C_t) \tag{7}$$

The convolutional product and element-wise multiplication operations are denoted by "$*$" and "$\circ$" respectively. In equations (3-5), I_t, F_t, and O_t are the input, forget, and output gate, respectively. W represents the weight matrix, x_t is the current input data, h_{t-1} is the previously hidden output, and C_t denotes the cell state at timestep t. The traditional LSTM equations use convolution operation (*), in comparison, the Conv-BiLSTM uses matrix multiplication between W and X_t ,h_{t-1} for every gate. This matrix multiplication replaces the fully connected layer with a convolutional layer, leading to a reduced number of weight parameters in the model. The Hybrid Conv-BiLSTM model is implemented to capture the advantages of the CNN and BiLSTM techniques to improve the overall prediction accuracy. The expected input shape into the Conv-BiLSTM model must be of the form [samples, timesteps, rows, columns, channels]. When using the Conv-BiLSTM, the previous EV timeseries of 47 (lags) is split such that it has one row of 47 timesteps, and the Conv-BiLSTM performs the convolutional operation on this particular row. This sequence design operation results in a 5D input tensor with shape [s, 1, 1, 47, 9] denoting sample, timestep, rows, columns and channels, respectively, as shown in Fig. 4(a). Finally, the hidden layer of the Conv-BiLSTM encoder is defined and flattened in readiness for decoding using the BiLSTM operation. The last layer is made up of the fully connected (dense layer) with 200 neurons for processing the output from the BiLSTM operation.

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77977-1_20

2.5 Hybrid Bidirectional Deep Convolutional Neural Network Long Short-Term Memory (CNN-BiLSTM)

The CNN-BiLSTM is a hybrid model that combines a 1D-CNN with a BiLSTM model to solve the vanishing gradient problem. This hybrid system consequently captures the advantages of both models. In this hybrid system, the 1D-CNN acts as an encoder responsible for interpreting the input sequence from the EV time series. The CNN encoder model then outputs a sequence that is passed on to the bidirectional BiLSTM model (decoder) for interpretation. The encoder CNN model does the convolutional operation and outputs a sequence. Naturally, the LSTM has an inherent strong ability to remember the structure of short and long term sequences; thus, by combining the BiLSTM with the CNN, which automatically learns features from sequence data, the hybrid CNN-BiLSTM model offers improved accuracy. The CNN and BiLSTM model hybrid structure expect the input EV demand data to have a 3-dimensional form (sample, input and timestep). Fig. 4(b) presents an illustration of hybrid CNN-BiLSTM workflow. In this implementation, the historical EV time series of 47 hours is input into the CNN encoder architecture for reading. The first convolutional layer reads across this input sequence using a filter of size three timesteps (1x3) and then projects its output onto 32 feature maps. The second convolutional layer reads the previous layer's output and uses the max-pooling layer to simplify the feature maps sizes by preserving the maximum possible amount of information (signal). The final extracted features map from the max-pooling layer is then flattened for use with the BiLSTM decoding module, which processes the cell state in both forward and backward directions before passing the output to the fully connected layer prediction.



(a) ConvBiLSTM (b) CNN-BiLSTM

Fig. 2: Proposed Hybrid models for EV charging load transfer learning implementation

3 Data-sets Description

The EV datasets(https://data.dundeecity.gov.uk/dataset/ev-charging-data), collected over an entire year from 01/09/2017 to 31/08/2018 equates to 8760 data

DOI: 10.1007/978-3-030-77977-1_20

points, representing hourly EV charging power (in kilo Watt, kW) of a year. As single residential charging stations have small data sizes with highly random data, collective residential EV charging profiles were created by grouping data from 100 single residential profiles located in relative proximity to each other. This was done to ascertain the EV charging behavior at distribution substation level. Additionally, slow commercial EV charging station profile represents data from a Health Care Centre, and fast commercial EV charging station profile represents data from a multi-story car park, which are referred throughout this report as slow commercial EV charging station (slow CEVCS) and fast commercial EV charging station (fast CEVCS), respectively. These commercial profiles differ in charging capacities; slow charging (6-10kW) and fast charging (22-70kW). The EV energy consumption profiles for the three types of charging stations, namely residential EVCS, slow CEVCS and fast CEVCS, are illustrated in Fig. 3.

The residential EVCS consumption patterns are similar throughout the seven days of the week, steadily rising from 6 am up until 7 pm, remaining low in the night when users are sleeping. The slow commercial CEVCS demand has two distinct consumption patterns within the week: weekdays and weekends. As seen in Fig. 3, the EV charging demand is higher during the weekdays and lower during the weekends. During the weekends it is low throughout the whole day because of less demand. The fast CEVCS offer quick charging (about 30mins) for users, resulting in the irregular demand with no distinct pattern. Predicting the EV demand for the fast charging station is expected to be challenging, given the irregular consumption patterns. Consequently, transfer learning can be expected to be problematic for fast CEVCS relative to the slow CEVCS.

3.1 Input Processing

The EV time series data must be transformed into supervised learning to allow reading by the model. The EV sequence is reframed into pairs of input and output variables using the sliding window method. The inputs comprise values of the current EV observation t and previous EV observations at times $t-1$,$t-2$....$t-n$ which predict the output EV observation at time $t+1$.The length of the sliding window is defined based on [6],[3].

3.2 Data Standardisation

The differences in numerical ranges between the input and output values require that the data-set be standardised. By standardisation, each feature is scaled to have a distribution that is centred around 0, having a standard deviation of 1. Standardisation allows for comparability among input data-sets, and it also enhances the training efficiency of models since the numerical condition of optimisation is improved. The mean and standard deviation for each feature is calculated, then the feature is scaled using equation 8.

$$z = (x_i - \mu)/\sigma \tag{8}$$

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77977-1_20

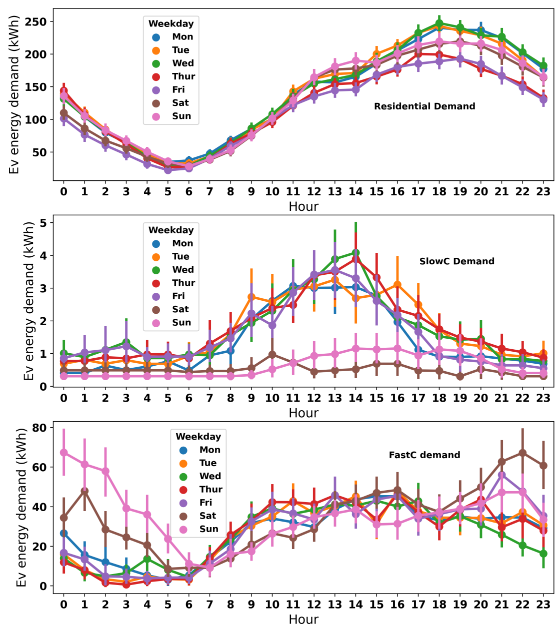


Fig. 3: EV energy consumption weekly profile by the hour.

where, z represents the standardized value, x_i is the observed energy consumption value, μ is the mean, and σ is the standard deviation of the EV energy data-sets.

3.3 Evaluation Metrics

An evaluation of the models' skill is done using the mean absolute error MAE, root-mean-square error RMSE, and R-squared R^2 metrics. The R^2 metric describes the proportion of variance of the input variable that the regression model explains. The MAE metric calculates the positive mean error value for the test data. The root-mean-square error (RMSE), which calculates the square root of the means square error, is a strict metric that penalises significant errors; thus, it is a reliable metric. One must look at these metrics concurrently when selecting the superior model.These performance evaluation metrics are calculated using equations (9-11) below

$$R^2 = 1 - \frac{\sum(y - y')^2}{\sum(y - \bar{y}')^2}, \tag{9}$$

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77977-1_20

$$RMSE = \sqrt{\Sigma \frac{(y' - y)^2}{N}}, \tag{10}$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} |y' - y| \tag{11}$$

where y is the measured EV energy consumption value, y' is the predicted EV energy consumption value, and N represents the number of data pairs considered.

4 Results and Discussion

This study investigates the possibility of knowledge transfer from; (1) residential EVCS to slow CEVCS and (2) residential EVCS to fast CEVCS to enhance prediction in the target datasets with limited labelled data. The baseline transfer learning-based model (CNN-T) is compared against the proposed hybrid transfer learning-based models (CNN-BiLSTM-T and Conv-BiLSTM-T) to determine which model is more superior. The newly created (CNN-N), CNN-BiLSTM-N and Conv-BiLSTM-N non-transfer learning models are also compared against their transfer learning-based counterpart models. If a new model without transfer learning can outperform the transfer learning-based models, we can conclude that transfer learning is not practical. Thus, we may need to fine-tune the transfer learning model or abandon it altogether. However, if the transfer learning-based models outperform the new model (which is the expected outcome), we can conclude that transfer learning is beneficial.

Table 1: Comparing tested models for residential commercial charging load prediction

MODEL	**RMSE (kWh)**	**MAE (kWh)**	R^2
CNN	27.87	22.13	0.86
CNN-BiLSTM	27.17	21.25	0.87
Conv-BiLSTM	26.96	21.12	0.87

4.1 Training with Full Data Complement

The results of the initial training of the three models on residential EVCS before transfer learning are shown in table 1. Lag 47 provided the best EV prediction results compared to other tested lags. The results in table 1 are in reference to lag 47; that is, the previous 47 hours of EV energy consumption are used as input for predicting the next hour of EV energy consumption. The tested models demonstrated almost similar performance in predicting the next hour charging

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77977-1_20

load for the residential EVCS (source domain), with the CNN, CNN-BiLSTM and Conv-BiLSTM models having comparable RMSE scores of 27.87 kWh, 27.17 kWh and 26.96 kWh, respectively.

4.2 Comparing Transfer Learning and Non-transfer Learning-based Methods

This section compares the three transfer learning-based models' performance against their newly created counterparts for residential EVCS to slow CEVCS and residential EVCS to fast CEVCS transfer learning. The results from this part of the investigation are shown in Tables 2 and 3 and illustrated in Fig 4 (a) and (b).

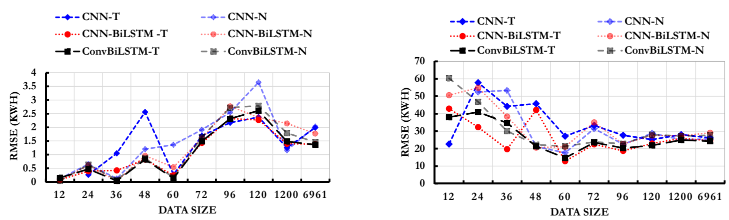
Table 2: Comparing RMSE values for transfer and non-transfer learning-based model for residential EVCS to slow CEVCS

Data size	**CNN-T**		**CNN-N**		**CNN-BiLSTM-T**		**CNN-BiLSTM-N**		**Conv-BiLSTM-T**		**Conv-BiLSTM-N**	
	rmse	mae	rmse	mae	rmse	mae	rmse	mae	rmse	mae	rmse	mae
12	44.87	39.23	0.12	0.10	0.08	0.07	0.06	0.05	0.15	0.14	0.04	0.04
24	0.26	0.25	0.65	0.65	0.39	0.39	0.61	0.61	0.47	0.47	0.63	0.63
36	1.05	1.04	0.12	0.11	0.42	0.41	0.03	0.03	0.04	0.03	0.08	0.07
48	2.56	2.00	1.20	1.08	0.82	0.04	0.96	0.84	0.82	0.48	0.86	0.66
60	0.27	0.26	1.36	1.31	0.24	0.19	0.54	0.53	0.13	0.11	0.16	0.15
72	1.69	1.50	1.91	1.58	1.43	1.11	1.60	1.29	1.49	1.24	1.65	1.33
96	2.17	1.21	2.55	1.67	2.32	1.32	2.77	1.70	2.32	1.27	2.72	1.78
120	2.37	1.58	3.65	3.43	2.27	1.62	2.36	2.09	2.61	1.64	2.79	2.38
1200	1.25	0.65	1.16	0.69	1.40	0.62	2.14	1.14	1.48	0.81	1.79	0.9
6961	1.99	1.17	2.03	1.13	1.38	0.98	1.77	1.14	1.36	0.98	1.46	0.88

The results from this part of the investigation indicate the potential of transferring learning in aiding model prediction in cases of limited labelled data. In the residential EVCS to slow CEVCS, the table and charts show competitive performance by all tested models. Only at data instance 36 do the newly-created CNN-N, and CNN-BiLSTM-N models seem to present superior results than the transfer based hybrid models, with the latter presenting equal MAE scores with the Conv-BiLSTM-T model. In the rest of the tested data instances, the proposed transfer learning-based hybrid models, particularly the Conv-BiLSTM-T model presenting superior results.

In the residential EVCS to fast CEVCS transfer scheme, hybrid transfer models continue to show dominance over newly created counterpart models. The CNN-N and the CNN-BiLSTM-N supersede counterpart models at data instance 48. Besides this instance, transfer learning models continue to dominate in both the limited data instances and beyond.

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77977-1_20



(a) Residential to slow CEVCS (b) Residential to fast CEVCS

Fig. 4: An RMSE comparison of the implemented non transfer and transfer learning-based models

Table 3: A comparison of the transfer learning and non-transfer learning-based model for residential EVCS to fast CEVCS

Data size	**CNN-T**		**CNN-N**		**CNN-BiLSTM-T**		**CNN-BiLSTM-N**		**Conv-BiLSTM-T**		**Conv-BiLSTM-N**	
	rmse	mae	rmse	mae	rmse	mae	rmse	mae	rmse	mae	rmse	mae
12	22.64	18.54	143.64	136.50	42.90	36.33	50.63	46.08	37.97	33.40	60.27	56.44
24	57.77	51.74	52.48	47.50	32.29	29.58	54.90	50.27	40.98	38.10	46.82	41.29
36	44.24	36.35	53.35	50.54	19.70	18.70	38.36	34.05	34.73	30.32	30.04	34.63
48	45.77	39.07	21.23	16.31	42.14	36.02	20.64	16.30	21.34	17.88	22.32	17.00
60	27.13	19.05	17.63	16.71	12.82	10.21	21.08	19.89	14.70	12.43	21.35	20.10
72	33.30	26.81	31.46	25.67	22.51	17.92	35.03	29.23	23.52	18.39	23.84	19.06
96	27.67	20.95	22.46	18.75	18.73	15.24	22.77	18.02	20.33	17.27	22.80	18.75
120	25.18	20.06	28.82	24.68	22.95	18.01	27.56	23.95	21.85	18.11	28.02	24.55
1200	28.03	21.38	25.13	19.60	25.90	20.28	27.42	20.69	24.92	18.41	27.20	20.68
6961	26.20	20.13	25.50	19.34	24.58	19.61	29.08	21.79	24.30	19.27	27.68	20.78

4.3 Comparing Transfer Learning-based Models

In another part of the investigation, transfer learning-based models were compared to determine which methods are most effective during transfer learning. It is observed that hybrid-based transfer learning models outperform the baseline CNN-T model at most tested data sizes, which point to the hybrid architectures' effectiveness during transfer learning.

As seen in Table 4, there is dominance by the proposed hybrid transfer learning-based models over the baseline CNN-T model. As clearly seen in the illustration given in Fig. 5, the CNN-T model at all instances recorded higher RMSE values in both tested transfer learning schemes, that is, from residential EVCS to slow CEVCS and from residential EVCS to fast CEVCS. When considering the hybrid transfer learning-based models, it is observed that in small data settings, the Conv-BiLSTM-T model tend to show superior performance over its counterpart CNN-BiLSTM-T hybrid model.

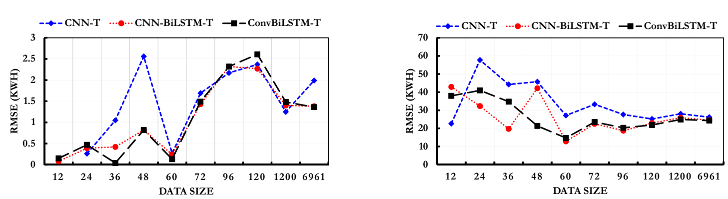
ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77977-1_20

Table 4: Comparing the implemented transfer learning-based models from residential EVCS to slow CEVCS and to fast CEVCS

Data size	**CNN-T**		**CNN-BiLSTM-T**		**Conv-BiLSTM-T**		**CNN-T**		**CNN-BiLSTM-T**		**Conv-BiLSTM-T**	
	rmse	mae	rmse	mae	rmse	mae	rmse	mae	rmse	mae	rmse	mae
12	44.87	39.23	0.08	0.07	0.15	0.14	22.64	18.54	42.90	36.33	37.97	33.40
24	0.26	0.25	0.39	0.39	0.47	0.47	57.77	51.74	32.29	29.58	40.98	38.10
36	1.05	1.04	0.42	0.41	0.04	0.03	44.24	36.35	19.70	18.70	34.73	30.32
48	2.56	2.00	0.82	0.04	0.82	0.48	45.77	39.07	42.14	36.02	21.34	17.88
60	0.27	0.26	0.24	0.19	0.13	0.11	27.13	19.05	12.82	10.21	14.70	12.43
72	1.69	1.50	1.43	1.11	1.49	1.24	33.30	26.81	22.51	17.92	23.52	18.39
96	2.17	1.21	2.32	1.32	2.32	1.27	27.67	20.95	18.73	15.24	20.33	17.27
120	2.37	1.58	2.27	1.62	2.61	1.64	25.18	20.06	22.95	18.01	21.85	18.11
1200	1.25	0.65	1.40	0.62	1.48	0.81	28.03	21.38	25.90	20.28	24.92	18.41
6961	1.99	1.17	1.38	0.98	1.36	0.98	26.20	20.13	24.58	19.61	24.30	19.27

4.4 Summary of Findings

Transfer learning models dominate the newly created (CNN-N) between data sizes of 24 hrs and 60 hrs and beyond, except at isolated instances (12 hrs, 24 hrs and 48 hrs). A 12-hour sized sample did not provide enough information to enhance transfer learning in most tested cases, explaining the relatively high error values recorded at this testing point by transfer learning models. Both hybrid deep transfer models prove robust models, with consistent superior performance over the CNN-N and CNN-T models for both residential EVCS to slow CEVCS transfer and residential EVCS to fast CEVCS transfer learning tasks. The hybrid CNN-BiLSTM-T and the Conv-BiLSTM-T show interchangeable performance between themselves in tested transfer learning tasks. That is, no outright dominance by either model in the critical transfer window and beyond.



(a) Residential to slow CEVCS (b) Residential to fast CEVCS

Fig. 5: An RMSE comparison of the implemented transfer learning-based models

The above observations confirm the benefits of transfer learning for electrical vehicle knowledge transfer. In cases of little available labelled data at either

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77977-1_20

the slow or fast CEVCS, a model trained at the residential EVCS can improve prediction efforts at these target data-sets.

5 Conclusions and Future Works

The CNN model is a standard model for transfer learning due to its enabling properties. Most research efforts study the effective strategies of improving transfer learning using the CNN model. However, little is known about the performance of hybrid deep learning models as mediums for transfer learning. Experiments were set up using electric vehicle data-sets to determine the performance of hybrid CNN-BiLSTM and Conv-BiLSTM models against the commonly used CNN-based model in the transfer learning tasks. The experimental results confirmed the superiority of hybrid deep learning transfer learning-based models over the conventional CNN transfer learning model. These results show that the hybrid structure of the implemented models is beneficial for the transfer learning tasks. As such, it can be concluded that the use of the hybrid CNN-BiLSTM and Conv-BiLSTM models for time series data-sets can improve the performance of the models in transfer learning settings. This study is also a pioneer transfer learning study to electric vehicles prediction literature. Future works would involve pre-processing input data to improve the transfer learning performance, such as the weighting of the samples selected for transfer learning; that way, only important data pairs are chosen, negating instances of negative transfer learning. Source domain models can be improved by considering more input data to enhance their generalisation capacity.

References

1. Ai, S., Chakravorty, A., Rong, C.: Household ev charging demand prediction using machine and ensemble learning. In: 2018 IEEE International Conference on Energy Internet (ICEI). pp. 163–168. IEEE (2018)
2. Bahdanau, D., Cho, K., Bengio, Y.: Neural machine translation by jointly learning to align and translate. arXiv preprint arXiv:1409.0473 (2014)
3. Brownlee, J.: Deep learning for time series forecasting: predict the future with MLPs, CNNs and LSTMs in Python. Machine Learning Mastery (2018)
4. Buzna, L., De Falco, P., Khormali, S., Proto, D., Straka, M.: Electric vehicle load forecasting: A comparison between time series and machine learning approaches. In: 2019 1st International Conference on Energy Transition in the Mediterranean Area (SyNERGY MED). pp. 1–5. IEEE (2019)
5. Fan, C., Sun, Y., Xiao, F., Ma, J., Lee, D., Wang, J., Tseng, Y.C.: Statistical investigations of transfer learning-based methodology for short-term building energy predictions. Applied Energy **262**, 114499 (2020)
6. Fan, C., Wang, J., Gang, W., Li, S.: Assessment of deep recurrent neural network-based strategies for short-term building energy predictions. Applied energy **236**, 700–710 (2019)
7. Fawaz, H.I., Forestier, G., Weber, J., Idoumghar, L., Muller, P.A.: Transfer learning for time series classification. In: 2018 IEEE international conference on big data (Big Data). pp. 1367–1376. IEEE (2018)

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77977-1_20

8. Grubb, M., Vrolijk, C., Brack, D.: Routledge Revivals: Kyoto Protocol (1999): A Guide and Assessment. Routledge (2018)
9. Hilson, D.: Managing the impacts of renewably powered electric vehicles on distribution networks. Tech. rep., Technical Report
10. Kim, T.Y., Cho, S.B.: Predicting residential energy consumption using cnn-lstm neural networks. Energy **182**, 72–81 (2019)
11. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. Communications of the ACM **60**(6), 84–90 (2017)
12. Li, Y., Huang, Y., Zhang, M.: Short-term load forecasting for electric vehicle charging station based on niche immunity lion algorithm and convolutional neural network. Energies **11**(5), 1253 (2018)
13. Munshi, A.A., Mohamed, Y.A.R.I.: Unsupervised nonintrusive extraction of electrical vehicle charging load patterns. IEEE Transactions on Industrial Informatics **15**(1), 266–279 (2018)
14. NIU, D., MA, T., WANG, H., LIU, H., HUANG, Y.: Short-term load forecasting of electric vehicle charging station based on kpca and cnn parameters optimized by nsga. Electric Power Construction p. 03 (2017)
15. Sajjad, M., Khan, Z.A., Ullah, A., Hussain, T., Ullah, W., Lee, M.Y., Baik, S.W.: A novel cnn-gru-based hybrid approach for short-term residential load forecasting. IEEE Access **8**, 143759–143768 (2020)
16. Schuster, M., Paliwal, K.K.: Bidirectional recurrent neural networks. IEEE transactions on Signal Processing **45**(11), 2673–2681 (1997)
17. Sutskever, I., Vinyals, O., Le, Q.V.: Sequence to sequence learning with neural networks. arXiv preprint arXiv:1409.3215 (2014)
18. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A.: Going deeper with convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 1–9 (2015)
19. Ullah, F.U.M., Ullah, A., Haq, I.U., Rho, S., Baik, S.W.: Short-term prediction of residential power energy consumption via cnn and multi-layer bi-directional lstm networks. IEEE Access **8**, 123369–123380 (2019)
20. Xydas, S., Marmaras, C., Cipcigan, L.M., Hassan, A., Jenkins, N.: Electric vehicle load forecasting using data mining methods (2013)
21. Zhang, C., Yang, Z., Li, K.: Modeling of electric vehicle batteries using rbf neural networks. In: 2014 International Conference on Computing, Management and Telecommunications (ComManTel). pp. 116–121. IEEE (2014)
22. Zhang, W.G., Xie, F.X., Huang, M., Li, J., Li, Y.F.: Research on short-term load forecasting methods of electric buses charging station. Power System Protection and Control **41**(4), 61–66 (2013)
23. Zhu, J., Yang, Z., Mourshed, M., Guo, Y., Zhou, Y., Chang, Y., Wei, Y., Feng, S.: Electric vehicle charging load forecasting: A comparative study of deep learning approaches. Energies **12**(14), 2692 (2019)

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77977-1_20