Battery Health Prediction using Deep Hybrid Learning

Rahul Anil Nair[1]
Department of Data Science and Business Systems,
School of Computing,
College of Engineering and Technology, SRM Institute of Science and Technology
Kattankulathur, Tamil Nadu, India
rahulanilnair1220@gmail.com

Dev Karan Suresh[1]
Department of Data Science and Business Systems
School of Computing
College of Engineering and Technology, SRM Institute of Science and Technology
Kattankulathur, Tamil Nadu, India
devkaran7501@gmail.com

Rajeswari D[2*]
Department of Data Science and Business Systems
School of Computing
College of Engineering and Technology, SRM Institute of Science and Technology
Kattankulathur, Tamil Nadu, India
drajiit@gmail.com
*Corresponding author

***Abstract*— The type and condition of the road the vehicle travels; how long the battery is used for are just a few of the variables that affect the battery life of electric cars (EVs). These elements might affect the battery's lifespan by causing it to drain more quickly than usual. By understanding the various factors that impact battery life, work can be done to mitigate these issues and extend the useful life of the battery. To find the best circumstances for increasing the battery's longevity, this research paper is concentrated on examining the impacts of road conditions on battery life. The previous article classified routes with an accuracy of 83.56% based on battery health while arriving at the destination [1]. This article examines EV routed and categorises them on the basis of the number of brakes applied upon arrival at the destination. By doing so, it aims to develop recommendations for drivers on how to optimize their driving habits to extend the battery's lifespan. Observations indicate that bad routes generally involve a lot more brakes as compared to good routes. Vehicle speed has a negative correlation with heater power in a bad route and a positive correlation in a good route. Battery voltage and Battery state-of-charge have opposing correlations in both categories. In this research, a prediction of the route category was attempted by using Deep Hybrid Learning to integrate three pre-trained models., VGG16, Resnet50 and VGG19 with two boosting algorithms:- XGBoost and AdaBoost. The accuracy of the six deep hybrid learning models in predicting the category of these routes are as follows:-0.87625, 0.84875, 0.80125, 0.7925, 0.8425 and 0.855.**

Keywords*— *Battery Health, Route System, Image categorization, Electric Vehicle, Deep Hybrid Learning, Visual Geometry Group Neural Network, Residual Neural Network, Extreme Gradient Boost, Adaptive Boost

I. Introduction

Electric Vehicles (EV) have witnessed a surge in popularity in the last decade. This is a result of a combination of climate consciousness of the public and government support. The annual sale of EVs increased from just over two thousand to over 753 thousand worldwide. It is estimated that by 2040, EVs will account for 12-28% of the global fleet [2]. Because EVs are sustainable, that is, they are powered by renewable energy sources such as wind or solar, the cost of electricity is cheaper than the cost of fossil fuels [3]. While EV market penetration is increasing, substantial difficulties remain for these cars to attain their full potential. These barriers include battery capacity, cost, charging time, and charging station accessibility. When completely charged, EV batteries now have a limited range, and as the range improves, so does the cost of the battery. [4].

Lithium-ion batteries (LIBs) remains a promising energy storage solution, thanks to its qualities such as high energy density, low self-discharge, minimal memory effect, and extended lifespan. Notably, their high-energy density makes them the preferred power source for EVs [5]. A review of the degradation of LIBs was first published in 2005 [6], but as the applications of LIBs have increased, more research is necessary with a focus on a certain application. Because EV batteries are costly, EV owners may be concerned about battery degradation. When it comes to electric vehicles, battery deterioration is route dependent [7]. The battery can deteriorate in a variety of ways depending on the traveling, preservation, and charging conditions. In the literature, many simulated driving profiles have been proposed to analyse the range and durability of EV batteries. However, there has been no report comparing these profiles to actual degradation during driving. Given the importance of understanding deterioration route dependence, it is critical to assess the correctness of these profiles in modelling real-world driving behaviour [8].

This research intends to study real driving profiles and study the effect of path and road conditions on the degradation of the LIB with the help of deep hybrid learning models. This may help with prolonging battery life and reducing the anxiety surrounding the maintenance cost of EVs.

II. LITERATURE SURVEY

A. Hoke, A. Brissette et al. [9] developed a LIB lifetime model which considered numerous elements like state of charge profile, daily depth of discharge and temperature to estimate fade in both energy capacity and power. This model was validated by comparing it to a comprehensive technique evolved at the National Renewable Energy Laboratory.

979-8-3503-2284-2/23/$31.00 ©2023 IEEE
DOI 10.1109/ICPCSN58827.2023.00072

According to this study, power fade is not the decisive factor for battery life, rather capacity loss is. This means that the battery provides the same amount of energy but lasts less time. The patterns are applicable to both plug-in hybrid electric vehicles and electric vehicles because they are displayed across a wide range of battery sizes.

M. Xu, T. Wu et al. [10] considered fleet size as one of the factors affecting EV battery degradation. They proposed that the fleet size be determined by the on-demand charging technique to slow down battery degradation and save expenses. The authors created a mixed-integer nonlinear programming model and validated it using EVCARD, a Chinese electric car sharing operator. Furthermore, an examination of how the performance of one-way electric CSSs is influenced by many crucial elements such as the daily fixed cost of EV and pricing of the battery, efficiency of battery cycle, service charge, and relocation cost was carried out.

M. Jafari, L. E. Brown et al. [11] created a hierarchical Bayesian model for the estimation of fade in capacity in EVs. The authors noted that deterministic models would not be reliable as battery degradation depends on external factors such as the driver's usage pattern. As a result, they concluded that a probabilistic model would be more suited since it can account for uncertainties and produce more accurate estimations. On an 85-15 training to testing data split, they employed the Metropolis-Hastings Markov chain Monte Carlo sampling approach. They further validate their model by showing three case studies of different drivers where their model gives consistent results regardless of varying inputs.

G. Saldaña, J. I. S. Martín et al. [12] developed a model which can determine the LIBs State of Health (SOH) at every moment to develop operating strategies for lifetime maximization. The authors noted that the temperate and C-rate were the most influential factors for the degradation of the battery. They also developed a model which took the environment into consideration. They split the environment into three categories: high demanding, low demanding, and average demanding environments based on the type of road and traffic on that road. They found that the batteries degraded faster in high demanding areas such as the highway, while in low demanding areas the degradation was 16% slower.

III. PROPOSED SYSTEM

The original dataset includes an EV's latitude, longitude, Battery SOC, Air Conditioning Power, Heater Power, Battery Voltage, Battery Current, and Vehicle Speed [km/hr] from start to finish. Using the latitudes and longitudes provided in the information, satellite photos of the vehicle paths were created using Google Earth. Based on the number of brakes, these photos are separated into 'best' and 'worst' routes, which are then used to train and test six deep hybrid learning models, each of which reports varying accuracies when analyzing the test photographs.

A. DATASET DESCRIPTION

"A Large-Scale Dataset for Vehicle Energy Consumption Research – Vehicle Energy Dataset (VED)" [13], an IEEE transaction article, was utilized to produce the original dataset employed for training and testing the models. This dataset contains various characteristics, which are shown in the table below. From November 2017 to November 2018, it comprises data for different PHEV/EVs, HEVs, and petrol automobiles. It is shown in Fig 1 below.

#	Column	Dtype
1	DayNum	float64
2	VehId	int64
3	Trip	int64
4	Timestamp(ms)	int64
5	Latitude[deg]	float64
6	Longitude[deg]	float64
7	Vehicle Speed[km/h]	float64
8	MAF[g/sec]	float64
9	Engine RPM[RPM]	float64
10	Absolute Load[%]	float64
11	OAT[DegC]	float64
12	Fuel Rate[L/hr]	float64
13	Air Conditioning Power[kW]	float64
14	Air Conditioning Power [Watts]	float64
15	Heater Power [Watts]	float64
16	HV Battery Current[A]	float64
17	HV Battery SOC[%]	float64
18	HV Battery Voltage[V]	float64
19	Short Term Fuel Trim Bank 1[%]	float64
20	Short Term Fuel Trim Bank 2[%]	float64
21	Long Term Fuel Trim Bank 1[%]	float64
22	Long Term Fuel Trim Bank 2[%]	float64
23	Category	object

Fig 1. Description of data

The coordinates of the electric vehicles were plotted and saved as photos employing the dataset for the whole year of electric vehicles with the ids of '10,' '455,' and '541' and Google Earth. The dataset yielded 461 route photos, 137 of which were "best" images and 324 of which were "worst." These photos were then supplemented to guarantee that the training and test datasets had an identical amount of samples. These final photos served as the pre-trained models' dataset and were used to train and test the models. The proposed dataset features are given below: -

Best images: - The 'best' photos represent routes where number of brakes used was fewer than 100 during the entire journey. Fig 2 depicts an example of a best path picture.

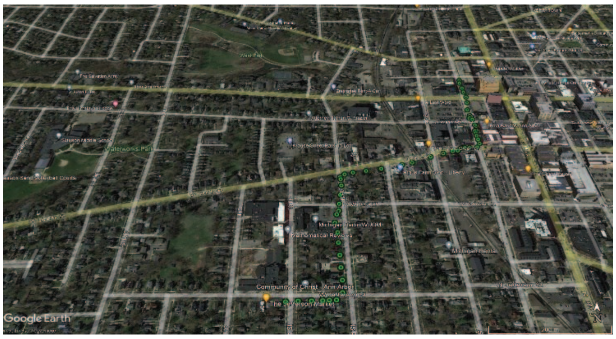

Fig 2. Best Path

Worst images: - The 'worst' photos represent the routes where the number of brakes used surpassed 100 during the voyage. Fig 3 shows an example of a worst path picture.

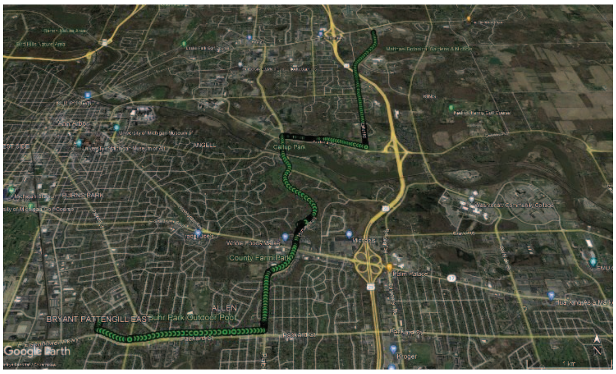

Fig 3. Worst Path

Latitude: Provides distance north or south of the equator of a vehicle at a timestamp given in milliseconds. The location illustrated in these route charts is Ann Arbor, United States of America.
Longitude: - Returns the distance east or west of a vehicle in relation to the prime meridian at the time stamp.
Battery SOC: -defines the whole battery capacity which is now accessible in proportion to the rated capacity.
Vehicle Speed [km/hr]: Returns the speed of the vehicle at various timestamps.
Heater Power [Watts]: - Provides the heat generated by the vehicle's engine at various timestamps throughout the route.
Air Conditioning Power [Watts]: Specifies the amount of air conditioning power generated by the vehicle at various points throughout the journey.
HV Battery Current:- Gives the current generated by the vehicle at different timestamps throughout the journey.
HV Battery Voltage:- Gives the voltage of the battery of the EV at different timestamps throughout the journey.

B. PREPROCESSING

To generate the final dataset, the original dataset was pre-processed in phases. First, the electric cars' vehicle ids were obtained, and the routes of these vehicles were split and stored in new data frames that were further segregated based on the route id.
These data frames were inserted into Google Earth Pro where satellite images of the routes were obtained. These were then. segregated into best and worst images based on the number of brakes applied throughout the journey. This was calculated by counting the instances where the current value of the Vehicle speed was greater than the succeeding value. The function developed for this purpose is depicted in Fig 4.

```
def countbrake(df):
    count=0
    lis=df['Vehicle Speed[km/h]'].tolist()
    for i in range(0,len(lis)-1):
        if(lis[i+1]<lis[i]):
            count=count+1
    return count
```

Fig 4. Function to count number of brakes in the route.

Both the best and worst images were then augmented by employing various image augmentation functions, demonstrated in Fig 5, and stored in separate folders for further use. This included rotation, shearing, zooming, horizontal flip and adjusting the range of brightness. Out of the augmented and original images, 4000 images were selected, out of which 2000 were best images and 2000 were worst. The image augmentation function used is depicted in Fig 5.

```
Image Augmentation Function

from keras. preprocessing. image import ImageDataGenerator

from skimage import io

datagen = ImageDataGenerator (

    rotation_range = 180,

    shear_range = 0.2,

    zoom_range = 0.2,

    horizontal_flip = True,

    brightness_range = (0.5, 1.5))
```

Fig 5. Image Augmentation Function with parameters

3200 images of both categories were used to train the models while the remaining 800 were used as test data. The entire process is depicted in the form of an architectural diagram in Fig 5.

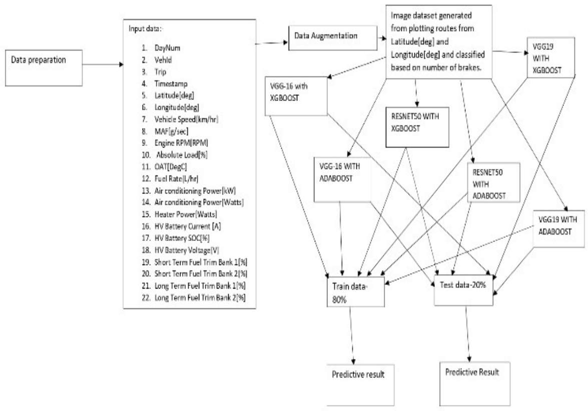

Fig 6. Architectural diagram

B. MODEL DESCRIPTION

Six deep hybrid learning models were made to analyze how accurately the category of the given routes can be predicted. Deep hybrid learning models involve combining a pre-trained model with a boosting algorithm like XGBoost or AdaBoost to enhance the accuracy of the predictions and reduce overfitting or underfitting.
The pre-trained models used for this purpose are as follows: - VGG16, Resnet50 and VGG19. Each of these pre-trained models are combined with XGBoost and AdaBoost and the results are represented by a confusion matrix.
These are the models used for implementation in our paper.
These models were used to further enhance the accuracy of our predictions and avoid overfitting and underfitting. They are given as follows: -

VGG16

A pre-trained model that was implemented in this research. It is a 16-layer convolutional neural network containing convolutional and pooling layers [14]. It is pre-trained since it was trained using the ImageNet database. The VGG16 weights can be either 'imagenet', or None. In this paper, VGG-16 is paired with both XGBoost and AdaBoost to calculate the accuracy of predictions in a deep hybrid learning model.

Resnet50

Another pre-trained model that was picked to be implemented in this paper is Resnet50 [15]. It, too, is a convolutional neural network with 50 layers, like VGG-16. It has 50 layers, of which 48 are convolutional layers, one is a MaxPool layer, and one is an average pool layer. The Resnet50 weights can be 'imagenet,', or None. It is commonly used in applications related to computer vision. In this paper, Resnet50 is paired with both XGBoost and AdaBoost to calculate the accuracy of predictions in a deep hybrid learning model.

VGG-19

It is another pre-trained model that was implemented by pairing it with XGBoost and AdaBoost. The only difference between VGG-19 and VGG-16 is the additional three fully connected convolutional layers. Compared with the previous methods, it has faster training time and fewer training samples per time. Just like VGG-16 and Resnet-50, it is commonly used in applications related to image classification.

XGBoost

It is an efficient, open-source gradient boosting tree algorithm. It is a supervised learning algorithm which aims to predict the target by combining the estimated of a set of simpler, weaker models [16]. It pushes the limitations of computational power for boosted tree algorithms. It allows for fine-tuning of various parameters to optimize performance of the model. The algorithm's method is depicted in Fig 7[17].

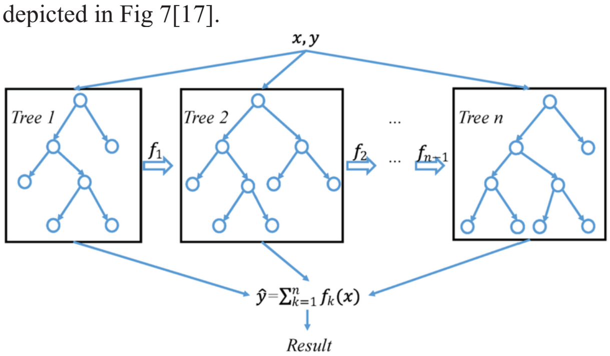

Fig 7. Structure of XGBoost

AdaBoost

An ensemble approach to machine learning. The weights of the model are redistributed with each occurrence, with greater weights being applied to erroneously identified instances [18]. It is mostly used to improve the performance of machine learning algorithms, particularly for poor learners. The algorithm's method is depicted in Fig 8[19].

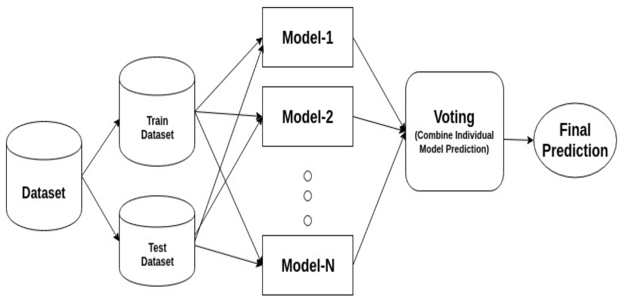

Fig 8. Structure of AdaBoost

IV. RESULTS AND DISCUSSION

The influence of road conditions on the vehicle's battery health can be seen in the characteristic features denoted by the best and worst routes as follows: -

1. Most of the best routes have a positive correlation between Vehicle Speed[km/hr] and Heater Power [Watts]. This may be because less brakes are involved throughout the journey which may lead to the heater power being consistent while vehicle speed increases gradually.
   The opposite result can be viewed in the case of the worst routes, with more negative correlations between Vehicle Speed [km/hr] and Heater Power [Watts]. This may be because more brakes are involved throughout the journey, resulting in more heat being produced by the vehicle with vehicle speed decreasing along with it. The above cases can be seen in Fig 9 and 10 for best and worst routes respectively. The figures were obtained by plotting the list of the correlations, stored separately for best and worst routes.

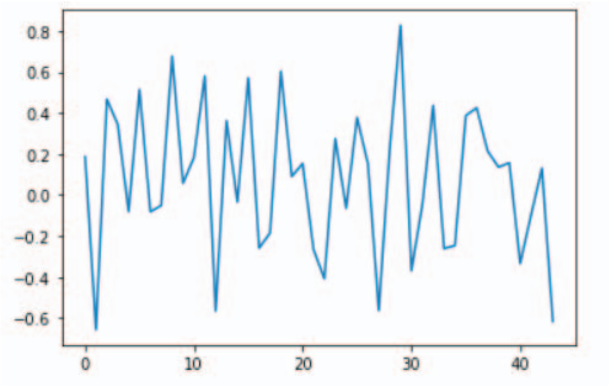

Fig 9. Correlations b/w Vehicle Speed and Heater Power-Best Route

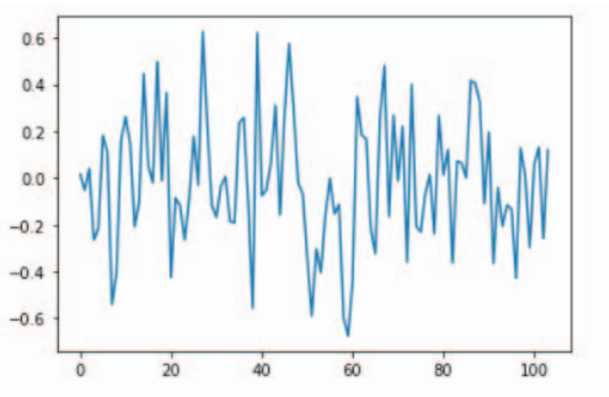

Fig 10. Correlations b/w Vehicle Speed and Heater Power-Worst Route

2. Both the best routes and worst routes have more positive correlations between Vehicle Speed[km/hr] and HV Battery SOC than negative, but the maximum positive correlation between Vehicle Speed[km/hr] and HV Battery is 0.9239962696790317 for best routes while it is 0.7434561414630134 for worst routes.
   The higher positive correlation in best routes may be because while vehicle speed increases, there is not much of a dip in Battery SOC as compared to worst routes due to there being less occurrences of braking and shorter distances. The above cases have been depicted in Fig 11 and 12 for best and worst routes respectively. The figures were obtained by plotting the list of the correlations, stored separately for best and worst routes.

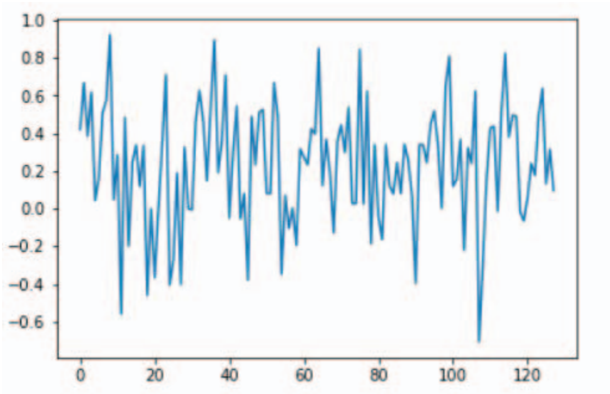

Fig 11. Correlation b/w Vehicle Speed and HV Battery SOC-Best Route

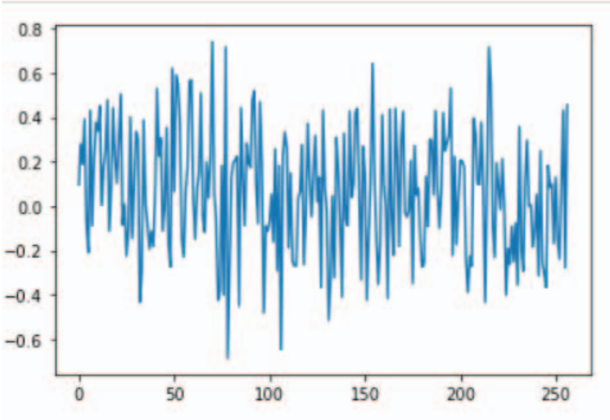

Fig 12. Correlation b/w Vehicle Speed[km/hr] and HV Battery state of charge-Worst Route

3. Most of the best routes have a positive correlation between Battery Voltage and Battery SOC. The opposite result can be viewed in the case of the worst routes, with more negative correlations between Battery Voltage and Battery SOC. Even though Vehicle Speed[km/hr] has a negative correlation with Battery Voltage in both cases, there is not much of a dip in Battery SOC in best routes, as compared to worst routes. Since Battery Voltage is decreasing throughout the journey, the dip in Battery SOC could be considered as the cause for the opposing results in both categories. The above cases have been demonstrated in Fig 13 and 14 for best and worst routes respectively. The figures were obtained by plotting the list of the correlations, stored separately for best and worst routes.

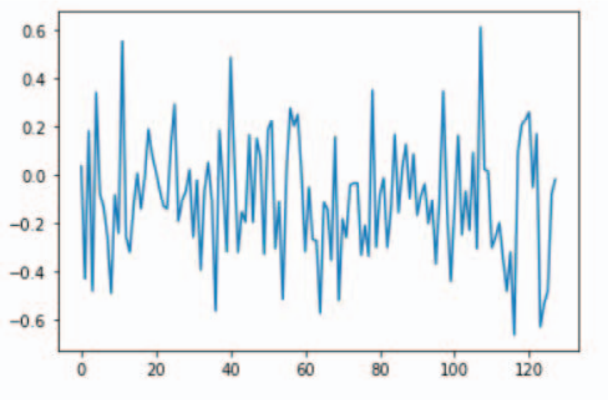

Fig 13. Correlation b/w Voltage and Battery State of Charge-Best Route

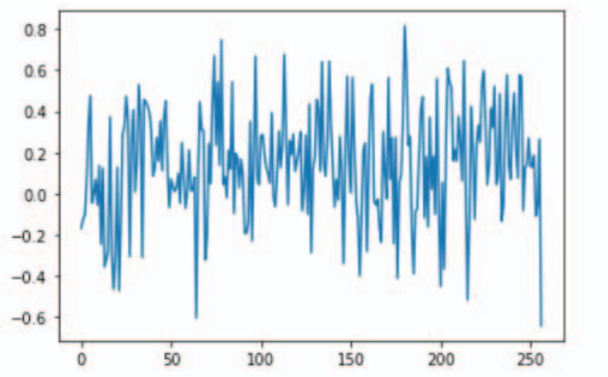

Fig 14. Correlation b/w Voltage and Battery State of Charge-Worst Route

The accuracy of the predictions made by extracting features using three pre-trained models and putting them through two boosting algorithms, resulting in six deep hybrid learning models can be given as follows: -

Pre-trained models/Boosting algorithms	XGBoost (Extreme Gradient Boosting)	AdaBoost (Adaptive Boosting)
VGG-16	0.87625	0.84875
Resnet50	0.80125	0.7925
VGG-19	0.8425	0.855

Table 1. Results of accuracy of predictions by models

The performance of the six deep hybrid learning models can be further demonstrated by the confusion matrices of these models with the pre-trained models listed vertically and the boosting algorithms listed horizontally, depicted in Fig 15.

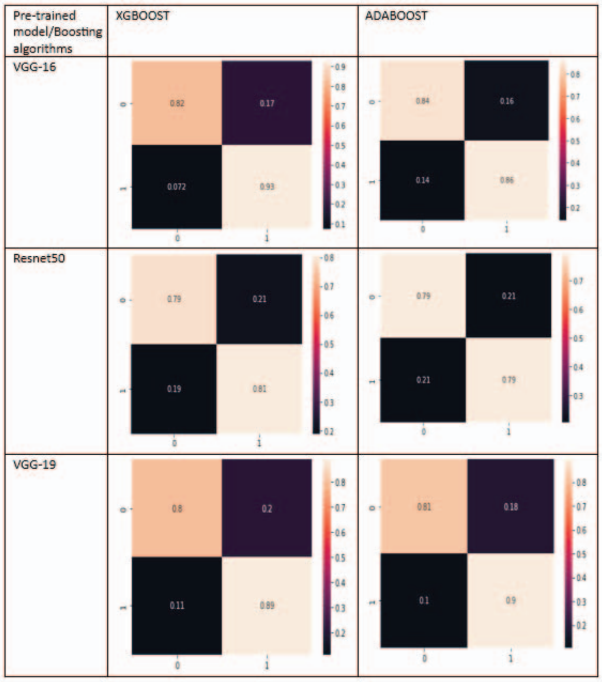

Fig 15. Confusion Matrices of the deep hybrid learning models

The results show that all six models have been well trained. They can be used to correctly predict whether the route is the best or worst for the battery health of an electric vehicle. Different parameters were run before arriving at the current results, which are the best-case scenario for each model. These models can be used for testing routes in areas different from those mentioned in the database used in this paper. It would be possible to further increase the performance of the model by subjecting it to situations that may not have been present in the existing database. The data pre-processing and model creation were done in Python on a Jupyter notebook with an i7 CPU, 16 GB RAM, and Windows 11.

V. CONCLUSION

This study used Google Earth to extract satellite photos of the routes from the VED dataset. The images were classified as best and worst routes according to the number of brakes applied throughout the journey. Correlations between the different features were taken to get a better idea about the characteristics of these routes. Six deep hybrid learning models were trained to predict the category of the routes and were tested successfully.

In the future, it will be possible to employ this research to categorize routes in different terrains, to be able to further analyze the influence of the route conditions on an electric vehicle and improve the performance of the model to be tested.

REFERENCES

1. R. A. Nair, D. K. Suresh, and R. D, "Predicting the Impact of Road Conditions on Battery Health Via Transfer Learning," 2023 Third International Conference on Artificial Intelligence and Smart Energy (ICAIS), Coimbatore, India, 2023, pp. 1194-1200, doi:10.1109/ICAIS56108.2023.10073732.
2. Nikita O. Kapustin, Dmitry A. Grushevenko, Long-term electric vehicles outlook and their potential impact on electric grid, Energy Policy, Volume 137, 2020
3. Granovskii, M.; Dincer, I.; Rosen, M.A. Economic and environmental comparison of conventional, hybrid, electric and hydrogen fuel cell vehicles. J. Power Sources 2006, 159, 1186–1193
4. D. Pevec, J. Babic, and V. Podobnik, “Electric Vehicles: A Data Science Perspective Review,” Electronics, vol. 8, no.10,p.1190,Oct.2019,doi: 10.3390/electronics8101190.
5. T. Kim, W. Song, D.-Y. Son, L. K. Ono, and Y. Qi, “Lithium-ion batteries: outlook on present, future, and hybridized technologies,” Lithium-ion batteries: outlook on present, future, and hybridized technologies - Journal of Materials Chemistry A (RSC Publishing), Jan. 17, 2019
6. J. Vetter, P. Novák, M. R. Wagner, C. Veit, K.-C. Möller, J. O. Besenhard, M. Winter, M. Wohlfahrt-Mehrens, C. Vogler and A. Hammouche, Ageing mechanisms in lithium-ion batteries, J. Power Sources, 2005, 147(1–2), 269–281
7. Dubarry, M.; Baure, G.; Devie, A. Durability and Reliability of EV Batteries under Electric Utility Grid Operations: Path Dependence of Battery Degradation. J. Electrochem. Soc. 2018, 165, A773–A783
8. G. Baure and M. Dubarry, “Synthetic vs. Real Driving Cycles: A Comparison of Electric Vehicle Battery Degradation,” Batteries, vol. 5, no. 2, p. 42, May 2019, doi: 10.3390/batteries5020042.
9. A. Hoke, A. Brissette, K. Smith, A. Pratt and D. Maksimovic, "Accounting for Lithium-Ion Battery Degradation in Electric Vehicle Charging Optimization," in IEEE Journal of Emerging and Selected Topics in Power Electronics, vol. 2, no. 3, pp. 691-700, Sept. 2014, doi: 10.1109/JESTPE.2014.2315961.
10. M. Xu, T. Wu, and Z. Tan, “Electric vehicle fleet size for carsharing services considering on-demand charging strategy and battery degradation,” Transportation Research Part C: Emerging Technologies, vol. 127, p. 103146, Jun. 2021, doi: 10.1016/j.trc.2021.103146.
11. M. Jafari, L. E. Brown and L. Gauchia, "Hierarchical Bayesian Model for Probabilistic Analysis of Electric Vehicle Battery Degradation," in IEEE Transactions on Transportation Electrification, vol. 5, no. 4, pp. 1254-1267, Dec. 2019, doi: 10.1109/TTE.2019.2956350.
12. G. Saldaña, J. I. S. Martín, I. Zamora, F. J. Asensio, O. Oñederra and M. González, "Empirical Electrical and Degradation Model for Electric Vehicle Batteries," in IEEE Access, vol. 8, pp. 155576-155589, 2020, doi: 10.1109/ACCESS.2020.3019477.
13. G. Oh, D. J. Leblanc and H. Peng, "Vehicle Energy Dataset (VED), A Large-Scale Dataset for Vehicle Energy Consumption Research," in IEEE Transactions on Intelligent Transportation Systems, vol. 23, no. 4, pp. 3302-3312, April 2022, doi: 10.1109/TITS.2020.3035596.
14. A. A. Rangari, S. Das and R. D, "Cursive Handwriting Recognition Using CNN with VGG-16," 2023 International Conference on Artificial Intelligence and Knowledge Discovery in Concurrent Engineering (ICECONF), Chennai, India, 2023, pp. 1-6, doi: 10.1109/ICECONF57129.2023.10083561.
15. R. Mathur, T. Chintala and D. Rajeswari, "Identification of Illicit Activities & Scream Detection using Computer Vision & Deep Learning," 2022 6th International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 2022, pp. 1243-1250, doi: 10.1109/ICICCS53718.2022.9787991.
16. D. Sikka and R. D, "Basketball Win Percentage Prediction using Ensemble-based Machine Learning," 2022 6th International Conference on Electronics, Communication and Aerospace Technology, Coimbatore, India, 2022, pp. 885-890, doi: 10.1109/ICECA55336.2022.10009313.
17. Wang, Y., Pan, Z., Zheng, J. *et al.* A hybrid ensemble method for pulsar candidate classification. *Astrophys Space Sci* 364, 139 (2019). https://doi.org/10.1007/s10509-019-3602-4
18. D. Sikka, Shivansh, R. D and P. M, "Prediction of Delamination Size in Composite Material Using Machine Learning," 2022 International Conference on Electronics and Renewable Systems (ICEARS), Tuticorin, India, 2022, pp. 1228-1232, doi: 10.1109/ICEARS53579.2022.9752123.
19. .