This is the Pre-Published Version.

The following publication Y. Wang and I. W. -H. Ho, "Joint Deep Neural Network Modelling and Statistical Analysis on Characterizing Driving Behaviors," 2018 IEEE Intelligent Vehicles Symposium (IV), Changshu, China, 2018, pp. 2060-2065 is available at https://doi.org/10.1109/IVS.2018.8500376.

Joint Deep Neural Network Modelling and Statistical Analysis on Characterizing Driving Behaviors

Yuhao Wang[1] and Ivan Wang-Hei Ho[1,2]
[1]Department of Electronic and Information Engineering, The Hong Kong Polytechnic University, Hong Kong
[2]The Hong Kong Polytechnic University Shenzhen Research Institute, Shenzhen, China
yohanna.wang@connect.polyu.hk, ivanwh.ho@polyu.edu.hk

Abstract— **Google defines the concept of autonomous driving as one of the applications of big data. Specifically, with the input sensor data, the autonomous vehicles can be provided with the semantic-level driving characteristics for an accurate and safe driving control. However, both the enumeration of handcrafted driving features with expert knowledge and the feature classification with machine learning for characterizing driving behaviors is lack of practicability under a complex scale. Therefore, this study focuses on detecting the sematic-level driving behaviors from large-scale GPS sensor data. Specifically, we classified different driving maneuvers from a huge amount of dataset through a layer-by-layer statistical analysis method. The identified maneuver information with the corresponding driver ID is useful for the supervised learning of high-level feature abstraction with neural network. With the aim of analyzing the sensory data with deep learning in a consumable form, we propose a joint histogram feature map to regularize the shallow features in this paper. Besides, extensive simulation is conducted to evaluate different machine learning and deep learning methodologies for optimal driving behavior characterization. Overall, our results indicate that Deep Neural Network (DNN) is suitable for the driving maneuver classification task with more than 94% accuracy, while Long Short-term Memory (LSTM) neural network performs well with a 92% accuracy in identifying a specific driver. However, LSTM shows degraded accuracy when the scale of the identification task becomes larger. In this case, a hierarchical deep learning model is proposed, and simulation results show that the combination of DNN and LSTM in this hierarchical model can well maintain the prediction accuracy even when the scale of the recognition task is four times larger.**

Keywords- Deep Learning; Machine Learning; Statistical Analysis; Driving Behaviour Classification; Autonomous Driving

I. INTRODUCTION

With the analysis and abstraction of driving knowledge, cars can not only identify the numerical mobility models of nearby vehicles, but also analyze their own driving behaviors, and take into account this information as the feedback control. When the data are continuously being acquired, cars can also rethink of their previous-made decisions for a better control. However, before the realization of fully autonomous driving without human control, the 'autonomous mode disengagement' problem [1] still widely occurs when failure operations are detected, or the human driver takes control. Therefore, to train an autonomous driving system, on one hand, a better understanding of the sematic level driving features from raw sensor data is necessary. On the other hand, a seamless driving feature identification model with artificial intelligence for the real-time driving behavior analysis is of paramount importance.

To characterize the uniqueness of driver behaviors from vehicle sensor data, the classification of driving maneuvers and the identification of the driving characteristics is warranted. Previous works mainly focused on the enumeration of handcrafted features as many as possible [2-4], and feature selection follows by for an accurate understanding of driver behavior. Meanwhile, [5] considered both geometric and structural complexity as the classification features to detect the unknown movement types with machine learning techniques. Eren at al. [6] used supervised learning with both moving average and empirical threshold to distinguish between risk and safe drivers. When it comes to the high-level driver identification task, Abdul at al. in [7] exploited the Cerebellum Model Articulation Controller (CMAC) to identify the driver behavior profile, while [8] proposed a scoring function to represent user's driving behaviors. However, the feature enumeration-based method is lack of efficiency and limited by domain-specific knowledge [4]. On the other hand, with the non-parametric control of deep learning [9], e.g., with the use of hierarchical architecture to learn the object at different levels of expression, abstracted features can be extracted out without complex manually-edited feature extractor. Previous works on deep learning assisted autonomous driving control are also numerous. For example, NVIDIA created PilotNet [10], which is a neural network based system and can output steering angles for the driving control. Also, the RNN based vehicle speed estimator for both the system identification and control dynamics is proposed in [11]. Meanwhile, Long Short-term Memory (LSTM) based RNN is utilized in trajectory classification tasks [12] with human annotation labels for supervised learning. However, the trajectory classification from raw GPS sensor data is not sufficient to deal with specific driver identification, and the manually annotated work is also costly and impractical in many data-driven tasks.

In this paper, we propose a joint neural network modeling and statistical analysis framework for characterizing different driving behaviors. In the first stage, the proposed system can classify different driving maneuvers using statistical analysis method from a massive amount of driving trajectories. In the second stage, the analyzed driving maneuvers and the driver identity will be served as both the general and specific high-level feature labels for the supervised learning of characterizing driving behaviors. By means of analyzing and comparing various machine learning and deep learning models, we conclude that the general high-level features are suitable to be classified by the fully connected deep neural network, while the specific high-level driver identification task can be handled well by the Long Short-term Memory (LSTM) neural network. Last but not least, with the analyzed learning models above, we further proposed a hierarchical deep learning model, which aims to conquer the degraded learning ability for the specific high-level feature abstraction on a larger scale. In short, the major contributions in this work are three folds:

1. We propose a statistical analysis method to characterize different types of driving maneuvers from a huge amount of GPS trajectories;
2. We propose a joint-histogram based feature map construction method. And we compared various machine learning and deep learning methods for identifying both the general and specific driving behaviours. Specifically, we find that fully connected DNN can get a classification accuracy of more than 94% in the driving maneuver classification task, while LSTM shows a 92% accuracy to identify a specific driver based on their driving trajectories;
3. The hierarchical deep learning model in the final part of this work can enlarge the driver identification scale and conquer the problem of the degraded classification accuracy in LSTM.

II. Proposed Methodologies

A. System Architecture

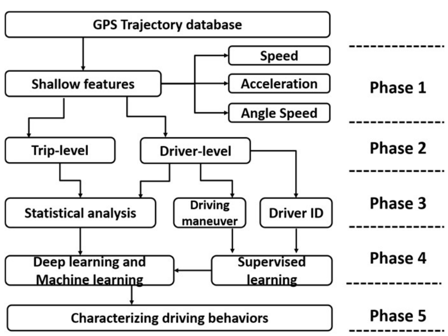

Figure 1. System architecture

The general architecture of the proposed system is shown above in fig. 1 and can be divided from top to down into five phases. Specifically, the shallow driving features, including speed, acceleration and angle speed, can be calculated from a large GPS trajectory database. After that, both the statistical analysis of driving maneuvers and the driver identity will be treated as the sematic-level driving features. To successfully identify a specific driver and recognize their driving maneuvers, we will introduce a joint histogram feature map as the input to characterize the driving features with the assistance of the artificial neural network. Moreover, based on the prediction accuracy of various neural networks, we introduced a hierarchical deep learning model for the large-scale driver identification task.

B. Driving maneuver analysis based on GPS data

To abstract vehicle motions into high-level driving maneuvers, local driving characteristics (such as the rapid steering, quickly stepping, high speed, steady deceleration) can be summarized as the variation of the performance state of multiple superficial driving features, while the global driving behaviors are the comprehensive consideration of the local characteristics. Therefore, we classified the general high-level feature of driving maneuvers into four types, namely, expert, cautious, normal, and reckless. From an omni-bearing analysis, an expert driver can well maintain a steady-state driving behavior within a long duration, while a cautious driver is slightly inferior when compared to the expert driver, but not very far. On the contrary, a reckless driver has more frequent aggressive maneuvers while a normal driver behaves more stable. Therefore, to adequately estimate the drivers' driving behaviors from the GPS trajectory dataset, we consider ten trip-level features (including both the x and y directional velocities (accelerations) and the overall velocity (acceleration), jerk, angular speed, absolute angular speed, and angle) to determine the effect of their statistical characteristics on the driving maneuvers.

As show in fig. 2, the input GPS data flow contains more than 2700 drivers, each with 200 trips and the GPS trajectories are recorded in every second. From each trip, the ten-dimensional shallow driving features mentioned above are calculated. With 200 trips for each driver, the overall tensor size of the processed dataset can be expressed as $\mathrm{x} \in \mathbb{R}^{features \times duration \times trips \times drivers}$. However, the *duration* feature is different from trip to trip. To weaken the trip-duration differentiation problem and quantify the feature stability for the analysis, further process is conducted within each feature row of the trip duration in term of the standard deviation with (1), and x represents any of the ten shallow features. Besides, to summarize the overall driving performance, we further reduced the *trip* dimension with the coefficient of variation in (2).

$$SD(x) = \sqrt{\frac{\sum_{i=1}^{duration}(x_i - \overline{x_i})}{duration-}} \quad (1)$$

$$CV(x) = (SD(SD(x))/\overline{SD(x)}) \times 100\% \quad (2)$$

Specifically, a CV value close to zero represents stable performance of a specific shallow feature, while a large value characterizes reckless driving performance. The tensors now have a dimension of $CV(x) \in \mathbb{R}^{features \times drivers}$. To further abstract the driving maneuvers, we choose the median value d_i in each row of the feature matrix $CV(x)$ as the decision boundaries in (3). Therefore, the processed elements in $CV(x)$ can be either 0 or 1, and thus the binary decision feature map is established.

$$CV(x) = \begin{cases} 0, & x < d_i \\ 1, & x \geq d_i \end{cases} \quad (3)$$

Meanwhile, the ten columns in this feature map is formed in order with the feature vectors below in (5–7), and we classified these ten features into three major categories in (4) to determine the driving maneuvers, namely, speed, acceleration, and angle.

$$\mathrm{maneuvers} = [\mathrm{speed}, \mathrm{acceleration}, \mathrm{angle}] \quad (4)$$

$$\mathrm{speed} = [\mathrm{vel}_x, \mathrm{vel}_y, \mathrm{velocity}] \quad (5)$$

$$\mathrm{acceleration} = [\mathrm{accel_x}, \mathrm{accel}_y, \mathrm{accel}, \mathrm{jerk}] \quad (6)$$

$$\mathrm{angle} = [\mathrm{angSpeed}, \mathrm{absAngSpeed}, \mathrm{angles}] \quad (7)$$

The three elements in (4) can be either 0 or 1 and the driving maneuver can be determined by mapping (4) into the look-up-table in Table 1. Each of the elements in (4) is

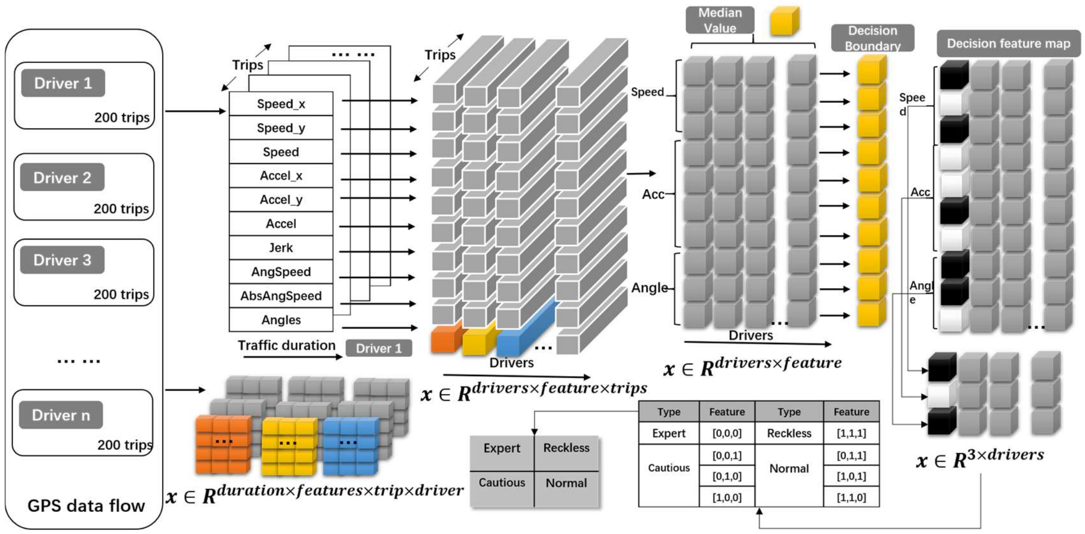

Figure 2. Driving maneuver analysis

determined by (5–7), e.g., for the *speed* feature in (5), the corresponding element value in (4) (0 or 1) is determined from the feature space of $speed \in \mathbb{R}^3$ and have 2^3 sub-features in total. The 2^3 sub state forms the (3×8) zero-one matrix in (8). In detail, to determine whether the *speed* feature is 0 or 1 in (9), we match (5) with the zero-one matrix in (8). Meanwhile, to strengthen the robustness of the detection results, we allow at most one unstable element (i.e., one 1-value among the three elements in (5)) to be classified as stable speed in (4) (i.e., speed=0 in (9)). Therefore, as shown in (9), if the vector $[\mathrm{vel}_x, \mathrm{vel}_y, \mathrm{velocity}]$ contains all-zero or only one 1-value, then the *speed* element in (4) equals 0, which means the overall performance is stable. On the contrary, the *speed* equals 1 and represents the unstable driving maneuvers.

$$\mathrm{substate}_{n\times m} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 \end{bmatrix} \tag{8}$$

$$\begin{cases} speed = 0,\ if\ [\mathrm{vel}_x, \mathrm{vel}_y, \mathrm{velocity}]^T = \mathrm{substate}_{n\times m}[:, i](i \in [1,4]) \\ speed = 1,\ if\ [\mathrm{vel}_x, \mathrm{vel}_y, \mathrm{velocity}]^T = \mathrm{substate}_{n\times m}[:, i](i \in [5,8]) \end{cases} \tag{9}$$

Similarly, the *angle* feature in (7) can also be processed by the same token. For the $acceleration \in \mathbb{R}^4$ in (6), it has 2^4 substates. The corresponding element in (4) can be determined by (10) and (11) as follows. And we allow at most two unstable elements (except the *jerk* feature) to be classified as stable state for the acceleration element in (4)

$$\mathrm{AccSubstate}_{\mathrm{n\times m}} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 1 & 1 \end{bmatrix} \tag{10}$$

$$\begin{cases} Acc = 0,\ if\ [\mathrm{accel_x}, \mathrm{accel}_y, \mathrm{accel}, \mathrm{jerk}]^T = Acc\mathrm{substate}_{\mathrm{n}\times m}[:, i](i \in [1,8]) \\ Acc = 1,\ if\ [\mathrm{accel_x}, \mathrm{accel}_y, \mathrm{accel}, \mathrm{jerk}]^T = Acc\mathrm{substate}_{\mathrm{n}\times m}[:, i](i \in [9,16]) \end{cases} \tag{11}$$

After that, we can determine a (1×3) decision vector in (4) based on the statistical analysis. Finally, we map this maneuver vector into the self-defined driving maneuver look-up-table in Table 1, so that the feature characterization process from the GPS trajectory to the high-level driving maneuvers can be achieved.

In short, this statistical analysis system can process the string of driving events to derive the high-level driving maneuvers. The automation of this task carries numerical benefits, such as the evaluation of the drivers' driving ability, safety rating, and the feedback control in autonomous driving. However, a representative understanding and summarization of the inner properties with statistical analysis needs to be developed from the processing of the massive datasets. Specifically, in this work, the classification results are analyzed from 547,200 trips. Nevertheless, for prior works on driving maneuver identification from real-time sensor data, the statistical analysis from both a huge amount of dataset and numerous drivers is not available. Therefore, with the hierarchical feature abstraction process in deep learning, we can utilize the statistical analyzed driving maneuvers as labels, so that the trip-level dataset can have stronger classification ability with the aid of supervised learning. Meanwhile, the specific high-level feature of the driver-id (a trip from a corresponding driver) can also be served as the label in the driver-identification task. All in all, as we will prove in the remainder of this paper, these two high-level features can be abstracted from GPS trajectories with the artificial neural network to benefit the overall driving behavior characterization process.

TABLE I. DRIVING MANEUVER LOOK-UP-TABLE

Type	Feature	Type	Feature
Expert	[0,0,0]	Reckless	[1,1,1]
Cautious	[0,0,1]	Normal	[0,1,1]
	[0,1,0]		[1,0,1]
	[1,0,0]		[1,1,0]

C. Feature map construction

Note that this work is not only about statistical driving feature classification, most importantly, we aim to propose a working methodology to train autonomous vehicles so that they can have stronger ability to recognize the statistical-analyzed driving maneuvers as well as to identify a specific driver under different driving conditions with a robust performance. However, directly feeding the raw GPS sensor data into the neural network for a workable driver identification or maneuver classification model is not practical. This is because GPS data only records the driving coordinates and time stamps, while there are more abstracted features need to be discovered from high-dimensional data [9]. In other words, the transformation from raw GPS sensor data $\mathcal{X}$ in (12) to numerical shallow features, which can possibly measure each individual driver, is necessary.

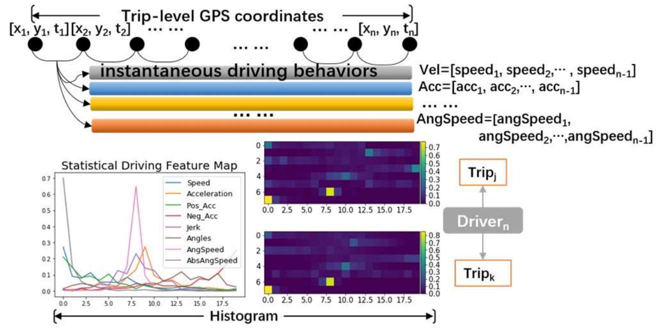

Figure 3. Joint-histogram feature map construction

As shown in fig. 3, after the transformation of the instantaneous driving behaviors from every adjacent GPS coordinates, we obtain the matrix $\mathcal{R}$ in (12), while every column in $\mathcal{R}$ corresponds a set of parameters of a specific driving feature to form the high-dimensional dataset. The subscript n represents the trip duration, and a trip can be formed with a varied length of n. To construct the feature dataset and separate it into both training (85%) and testing (15%) afforded by the use of the artificial neural network, we need to regularize the matrix $\mathcal{R}$ from different trips to a consumable form in the so called 'feature map'.

$$\mathcal{X} = \begin{bmatrix} x_1 & y_1 & t_1 \\ x_2 & y_2 & t_2 \\ x_3 & y_3 & t_3 \\ \cdots & \cdots & \cdots \\ x_n & y_n & t_n \end{bmatrix} \rightarrow \mathcal{R} = \begin{bmatrix} speed_1 & acc_1 & pos_{acc_1} & neg_{acc_1} & jerk_1 & \cdots & angSpped_1 \\ speed_2 & acc_2 & pos_{acc_2} & neg_{acc_2} & jerk_2 & \cdots & angSpeed_2 \\ speed_3 & acc_3 & pos_{acc_3} & neg_{acc_3} & jerk_3 & \cdots & angSpeed_3 \\ \cdots & \cdots & \cdots & \cdots & \cdots & & \cdots \\ speed_{n-1} & acc_{n-1} & pos_{acc_{n-1}} & neg_{acc_{n-1}} & jerk_{n-1} & \cdots & angSpeed_{n-1} \end{bmatrix} \quad (12)$$

For this reason, to identify a specific driver and understand their driving behaviors in an efficient and effective way, we proposed a powerful joint-histogram feature map and considered the trip-level driving behaviors. In detail, to recognize a specific driver from numerical driving trips, the characteristics of their driving feature dynamic-state (such as whether they often drive with high-speed, go through an intersection with a sharp turn and high acceleration or braking) is an important clue. Based on these, our joint histogram feature map is constructed with (13).

$$\text{newFeature}_\text{i} = \text{histogram}(normalization(feature_i), 20) \quad (13)$$

Specifically, the eight shallow features (*speed, acceleration, positive acceleration, negative acceleration, jerk, angle, angle speed and absolute angle speed*) within each trip are firstly being processed with the minimum-maximum normalization within a range of [0, 1]. Then the histogram of each of these features is generated based on twenty bins (such number of bins is identified through our experimental study as more bins cannot increase the classification accuracy while less number of bins is not representative enough). Besides, since the histogram results represent the cumulative frequency of a local trip, which is highly related to the trip duration. In order to relieve the effect of the varied trip duration, we further processed the count value of the histogram results to rate representation as $newFeature_j / \sum_{j=1}^{20} newFeature_j$. In fact, this joint histogram feature map groups the driving feature dynamics in microscopic perspective. Therefore, the driver's driving preference can be well represented, which greatly facilitate the characterization of driving behaviors.

D. High-level feature abstraction with deep learning

For a better control and decision making in autonomous driving, discovering useful intricate structure from high-dimensional big data is necessary. In this part, we intend to understand the semantic-level driving features from raw sensor data for autonomous driving control. In detail, with the analyzed driving maneuver labels proposed in Section B, the driver-id, and the constructed joint histogram feature map in Section C, we analyze different deep learning models from different aspects.

1) Fully connected Deep Neural Network

Deep Neural Network (DNN) makes use of a hierarchical architecture to learn the object at different levels of expression. In detail, by studying the non-linear network structure, the approximation of a complex function can be realized, and the original data can be expressed in a new feature space. For this reason, the working way of the DNN makes the driving characteristic classification problem easier to be implemented. For example, the fatigue driver detection with neural networks was implemented in [13]. A pattern recognition approach with the multilayer perception artificial neural network in [14] also showed an extraordinary performance to classify different skills of drivers. In our work, a five-layer fully connected DNN is designed with three hidden layers and the uniform kernel initializer. The activation function in each of the hidden layer is the recitified linear units (ReLUs) and the total number of the training epochs is 1000. Besides, the batch size in our work is 32 and the RMSProp optimizer with a learning rate of 10^{-3} is considered. Furthermore, the dense layer contains 50 neurons for the driver identification task and 4 neurons for driving maneuver classification.

2) Recurrent Neural Network

In traditional neural network models, such as the DNN we implemented above, the neurons within each layer are not connected. This made the DNN incapable of handling many issues, especially the prediction and feature classification on the sequence-dominated datasets. For this reason, the Recurrent Neural Network (RNN) was designed with the hidden units as the 'state vector' and responsible for maintaining the historical information [9]. In detail, RNN can memorize the information ahead and apply it to the calculation of the current state, so that nodes in each hidden layer are connected. In the application of autonomous driving, both the steering dynamic prediction [15] and the vehicle speed estimation [11] are very effective with RNN. Therefore, in this paper, we considered a two-hidden layer RNN on characterizing driving performance. Specifically, the parameters are tuned with cross validation, and we set the learning rate to 10^{-6}. By initializing the weights appropriately, we followed the idea of Le et al. [16] and considered the ReLU in these hidden layers, while the softmax is appended in the last layer.

3) Long Short-term Memory Network

However. it is difficult to learn and store long sequence information with RNN, and training an RNN can be problematic due to the shrink or grow of the backpropagation gradients in each step [9]. To solve this problem, as a subclass of RNN, the Long Short-term Memory (LSTM) network is proposed to replace the normal RNNs' hidden units with 'memory cells'. With a more sophisticated structure, LSTM can selectively pass information with its purpose-built memory cells to make it suitable for exploiting long-range contents. Therefore, LSTM is ubiquitous in the sequence -oriented applications, such as the trajectory classification task in [12] to identify two types of underwater vehicle with supervised learning. In our experiment of the LSTM neural network, similar to our RNN structure, we also have a two-hidden layer network, but with a 0.3% dropout included to avoid overfitting. The output is a dense layer with the softmax activation, and the learning rate in the LSTM is 10^{-6}.

III. SIMULATION RESULTS AND DISCUSSION

For the experimental studies, the dataset we adopt in this work is from the Kaggle competition in 2015 of the driver telematic analysis [17]. This dataset contains over 50,000 driver trips from more than 2700 drivers and 200 trips under each driver. Besides, a small and random number of false trips (trips that do not belong to a particular driver) exist in every driver's profile. We consider these trips as the noise. In our work, we randomly select 50 drivers from this dataset for the sematic-level driver identification as well as driving maneuver classification. Specifically, machine learning methods including the Principal Component Analysis (PCA) [18], t-Distributed Stochastic Neighbor Embedding (t-SNE) [19], Support Vector Machine (SVM) [20] and Random Forest [21], as well as deep learning methods including DNN, RNN and LSTM, are all implemented for comparison. The experimental results in Table 2 shows the trip-level accuracy (a trip-level feature map being successfully recognized), top-5 trip accuracy (the accuracy for the top-5 predictions) and the driver-level accuracy (the weighted vote of the majority prediction results from all the trip-level accuracy for a specific driver).

TABLE II. EXPERIMENT RESULTS

Feature map types	Methods	Driver identification accuracy			Driving maneuver accuracy
		Trip	Top-5	Driver	
Joint-histogram feature map	PCA&SVM	3.13%	---	2%	20%
	T-SNE&SVM	6.13%	---	10%	44.24%
	Random Forest	17.73%	---	18%	39.8%
	DNN	12.35%	35.10%	14%	94.66%
	RNN	28.62%	58.98%	66%	68.01%
	LSTM	36.54%	67.91%	92%	72.38%

On one hand, when considering all the machine learning methods, not surprisingly, the PCA with SVM behaves worst in both the driver identification and driving maneuver classification tasks. This is because after the reduction of dimension, PCA is still able to maintain the intrinsic information and measures their importance in the projection direction. However, especially for the four-type driving maneuver classification task (where random guess probability is about 25%) with 8,500 training feature maps, such projection step can cause the data points mix together, and therefore the classification accuracy of SVM (20% according to Table 2) is even worse than random guess. Hence, we considered the t-SNE for multi-dimensional reduction as well as enlarging the inter-cluster distance for a better classification. The SVM method, which is efficient in the high-dimensional space classification process, follows. However, for the 50-driver identification task, it can only get a trip-level prediction accuracy of more than 4% better than random guess. The driving maneuver classification accuracy with t-SNE and SVM outperforms the PCA based method by almost 25%, but still far from satisfaction. Besides, the random forest method behaves slightly worse in the driving maneuver identification task, but it performs better in the driver identification problem than the other two machine learning methods above. The reason is that the voted combination of different decision trees inside the random forest increases the identification probability of characteristics from different drivers.

On the other hand, deep learning methods show a better overall performance when compared to the shallow machine learning methods. Specifically, with the driving maneuver identification task, we found that a five-layer DNN can get the best classification accuracy of more than 94%, while the RNN takes longer training time and only gets an accuracy of around 68%. With a more complex network structure, the LSTM behaves better than RNN, but still 22% lower than the DNN classification accuracy. This is because the DNN analyzes the driving behaviors through a layer-by-layer information compression. It cares less about the detailed sequential driving feature information, but talented at mapping the driving maneuvers into a more separable feature space, which can greatly benefit the general high-level feature recognition task. However, this also leads to the weak classification accuracy of DNN in the specific driver identification problem. As we can see in Table 2 that it can only achieve 14% driver-level identification accuracy, which is even worse than the random forest method. On the contrary, the RNN can get a much better identification accuracy of 66%, while the enlarged sequence memorization ability in LSTM boosts the driver identification accuracy to the highest 92%. This is because to identify a specific driver from the joint-histogram feature maps, the drivers' signatures, such as the hard braking from high-speed driving of a reckless driver and stable acceleration dynamics with moderate or high-speed driving in most of the driving trajectories from an expert driver, all behave as the sequential correlation inside each of the feature maps, while both RNN and LSTM are designed in a sequenced-learning manner. All in all, as shown in Table 2, the deep LSTM neural network is suitable for the high-level driver-identification task, and DNN is more talented in mapping the general driving maneuver features into a more separable space.

IV. HIERARCHICAL DEEP LEARNING MODEL

In addition, it is necessary to further stretch our driver-identification task into a larger scale. However, as the most effective model in Table 2, LSTM shows a degraded performance in Table 3 when we enlarge the pool of drivers to 200. To deal with this problem, inspired by the working of deep neural network with its level by level study ability from lower features for a complicated problem, we propose a hierarchical deep learning model to enlarge the model's generalization ability. Specifically, we separate the driver identification task into two parts: the general high-level feature of the driving maneuver classification from shallow feature maps, and the specific high-level driver identification in each category of the classified driving maneuvers.

Specifically, with the same DNN and LSTM network structure in Section II above, we firstly classify the source dataset into four groups with respect to the four driving maneuvers with the DNN. Later on, for different driver types, we then train four LSTM classification models (same as the network structure in Section 2, but trained for the driver identification of each driving maneuvers) to adapt to the four maneuver groups' characteristics respectively. In short, with the experiment results in Table 3, compared to the LSTM alone, the hierarchical deep learning model behaves better in the 50-driver identification task, while more than two times better in the top-5 identification accuracy for the 200-driver case.

TABLE III. A COMPARISON OF THE DRIVER IDENTIFICATION ACCURACY

Accuracy	LSTM		Hierarchical deep learning	
	Trip-level accuracy	Top-5 accuracy	Trip-level accuracy	Top-5 accuracy
50 drivers	36.54%	67.91%	42.08%	78.55%
200 drivers	17.04%	36.64%	41.52%	73.38%

V. CONCLUSION AND FUTURE WORK

In this paper, we have classified four types of driving maneuvers from a huge amount of GPS using the statistical analysis methods. To make the driving behavior characterization task suitable for real-time applications of autonomous driving, and get benefits from the statistical analysis of sensor data, we set both the driving maneuver and driver-id as the supervised training labels for the sematic-level driving feature characterization with the artificial neural networks. Besides, we have proposed a joint-histogram feature map to regularize the shallow features from raw GPS sensor data to a consumable form, so that these feature maps can be served as the input for the deep-learning assisted driving feature identification tasks. Furthermore, we have compared the effectiveness of various deep learning and machine learning models, and found that the LSTM neural network has the best performance at driver identification with a 92% driver-level prediction accuracy, while a five-layer DNN shows an extraordinary performance with more than 94% accuracy in the general driving maneuver classification task. However, the LSTM model also shows a degraded feature recognition accuracy with the enlarged data size. To conquer this issue, we have considered the general high-level driving maneuvers as the intermediate feature between the shallow feature map and the specific driver identification, so that a hierarchical deep learning model is proposed with the DNN for maneuver group classification and LSTMs for grouped driver identification. This hierarchical model can expand the recognition scope. Besides, since the dataset used in this paper is lack of the real-world road topology as well as the traffic condition, which limits the characterization of driving features from many aspects. For future work, we intend to stretch this work with a great amount of realistic dataset collected from sensors on on-road vehicles, so that the driving features can be characterized under a more sophisticated condition.

ACKNOWLEDGMENT

The work of I. W.-H. Ho is supported in part by the Early Career Scheme (Project No. 25200714) established under the University Grant Committee (UGC) of the Hong Kong Special Administrative Region (HKSAR), China. The work of Y. Wang is supported in part by the National Natural Science Foundation of China (Project No. 61401384); and by The Hong Kong Polytechnic University (Projects G-YBK6, 4-BCCH, G-YBXJ).

REFERENCES

[1] V. V. Dixit, S. Chand, and D. J. Nair, "Autonomous vehicles: disengagements, accidents and reaction times," *PLoS one*, vol. 11, no. 12, p. e0168054, 2016.

[2] C. G. Q. M, J. O. López, and A. C. C. Pinilla, "Driver behavior classification model based on an intelligent driving diagnosis system," *in 2012 15th International IEEE Conference on Intelligent Transportation Systems*, pp. 894-899, 2012.

[3] J. Lu, D. Filev, K. Prakah-Asante, F. Tseng, and I. V. Kolmanovsky, "From vehicle stability control to intelligent personal minder: Real-time vehicle handling limit warning and driver style characterization," *in 2009 IEEE Workshop on Computational Intelligence in Vehicles and Vehicular Systems*, pp. 43-50, 2009.

[4] N. Lin, C. Zong, M. Tomizuka, P. Song, Z. Zhang, and G. Li, "An overview on study of identification of driver behavior characteristics for automotive control," *Mathematical Problems in Engineering*, vol. 2014, 2014.

[5] X. Li, "Using complexity measures of movement for automatically detecting movement types of unknown GPS trajectories," *American Journal of Geographic Information System*, vol. 3, no. 2, pp. 63-74, 2014.

[6] H. Eren, S. Makinist, E. Akin, and A. Yilmaz, "Estimating driving behavior by a smartphone," *in 2012 IEEE Intelligent Vehicles Symposium*, pp. 234-239, 2012.

[7] W. Abdul, W. Toh Guang, and K. Norhaslinda, "Understanding driver behavior using multi-dimensional CMAC," *in 2007 6th International Conference on Information, Communications & Signal Processing*, pp. 1-5, 2007.

[8] G. Castignani, T. Derrmann, R. Frank, and T. Engel, "Smartphone-Based Adaptive Driving Maneuver Detection: A Large-Scale Evaluation Study," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 9, pp. 2330-2339, 2017.

[9] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436-444, 2015.

[10] M. Bojarski et al., "Explaining How a Deep Neural Network Trained with End-to-End Learning Steers a Car," *arXiv preprint* arXiv:1704.07911, 2017.

[11] A. Haddoun, M. E. H. Benbouzid, D. Diallo, R. Abdessemed, J. Ghouili, and K. Srairi, "Modeling, analysis, and neural network control of an EV electrical differential," *IEEE Transactions on industrial electronics*, vol. 55, no. 6, pp. 2286-2294, 2008.

[12] M. D. L. Alvarez, H. Hastie, and D. Lane, "Navigation-Based learning for survey trajectory classification in autonomous underwater vehicles," *in 2017 IEEE 27th International Workshop on Machine Learning for Signal Processing (MLSP)*, pp. 1-6, 2017.

[13] M. M. Bundele and R. Banerjee, "Detection of fatigue of vehicular driver using skin conductance and oximetry pulse: a neural network approach," *in Proceedings of the 11th International Conference on Information Integration and web-based applications & services*, pp. 739-744: ACM, 2009.

[14] Y. Zhang, W. C. Lin, and Y. K. S. Chin, "A Pattern-Recognition Approach for Driving Skill Characterization," *IEEE Transactions on Intelligent Transportation Systems*, vol. 11, no. 4, pp. 905-916, 2010.

[15] G. Garimella, J. Funke, C. Wang, and M. Kobilarov, "Neural network modeling for steering control of an autonomous vehicle," *in 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2017, pp. 2609-2615.

[16] Q. V. Le, N. Jaitly, and G. E. Hinton, "A simple way to initialize recurrent networks of rectified linear units," *arXiv preprint* arXiv:1504.00941, 2015.

[17] (2015). Kaggle: Driver Telematics Analysis. Available: https://www.kaggle.com/c/axa-driver-telematics-analysis

[18] S. Wold, K. Esbensen, and P. Geladi, "Principal component analysis," *Chemometrics and intelligent laboratory systems*, vol. 2, no. 1-3, pp. 37-52, 1987.

[19] L. v. d. Maaten and G. Hinton, "Visualizing data using t-SNE," *Journal of Machine Learning Research*, vol. 9, no. Nov, pp. 2579-2605, 2008.

[20] C. Cortes and V. Vapnik, "Support vector machine," *Machine learning*, vol. 20, no. 3, pp. 273-297, 1995.

[21] M. R. Segal, "Machine learning benchmarks and random forest regression," *Center for Bioinformatics & Molecular Biostatistics*, 2004.