



HAL
open science

Prediction of human driving behavior using deep learning: a recurrent learning structure

Hugo Pousseur, Alessandro Corrêa Victorino

► **To cite this version:**

Hugo Pousseur, Alessandro Corrêa Victorino. Prediction of human driving behavior using deep learning: a recurrent learning structure. 25th IEEE International Conference on Intelligent Transportation Systems (ITSC 2022), Oct 2022, Macau, China. pp.647-653, 10.1109/ITSC55140.2022.9922144 . hal-04041186

HAL Id: hal-04041186

<https://hal.science/hal-04041186>

Submitted on 22 Mar 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Prediction of human driving behavior based on a recurrent LSTM model using different types of data

Hugo POUSSEUR¹ and Alessandro CORREA VICTORINO²

Abstract—Predicting the intentions of the human and the machine on a near future is required to the human-machine shared control of automated intelligent vehicles. The autonomous system is able to inform about its future intentions, however it is not possible for the human to provide this information, it is necessary then to predict it. This paper proposes a deep learning methodology to predict human navigation intentions in a time horizon of a few seconds, using a recurrent neural network (RNN) architecture based on the Long Short-Term Memory (LSTM) architecture. Taking as input various preprocessed and non-preprocessed data, generated by embedded sensors and the intrinsic data of the vehicle, the proposed model predicts the future linear and angular velocities of the vehicle. The model was trained and tested on a dataset created from real data from our cars equipped with sensors (LiDAR, camera), in different scenarios and road types. Furthermore, a data sensitive study is presented evaluating the effects of missing data in the learning process.

I. INTRODUCTION

In 2021, the number of deaths due to road accidents was 1.3 million¹ over the world. Of all road accidents, 90% are due to human error [6]. The development of autonomous cars is intended to increase safety on board vehicles and consequently reduce the number of deaths while improving driving comfort [11].

The complexity of the road traffic and the environment in which an autonomous car must evolve, creates an obstacle to the realization of the autonomous car in all situations without human supervision [23].

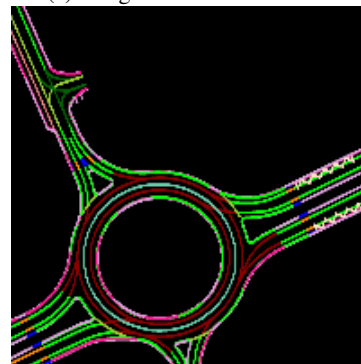
However, it is possible to increase driver safety by assisting the driver. The Advanced Driver-Assistance Systems (ADAS) [18],[10] cars can assist the driver by adjusting the vehicle speed to avoid collisions or to warn the driver of a risk (presence of a vehicle in a blind spot or a lane departure). We can extend this assistance of the autonomous system to the cooperation with the driver. In this case, the autonomous system should not only supervise the human in some tasks but should be able to drive alongside him, thus realizing a shared navigation, a concept already known in the fields of robotics, that is the shared control [1]. In this situation, according to the initial definition of shared control [21], the human and the autonomous system work on the same task at the same time. The result of the control is a weighting between the controls of each of these entities, this weighting is based on an arbitration principle defining a form of negotiation between human and the autonomous system.

Heuristics and Diagnosis of Complex Systems), CS60319 - 60203 Compiègne Cedex, FRANCE, ¹ hugo.pousseur@hds.utc.fr, ² alessandro.victorino@hds.utc.fr

¹<https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries>



(a) Image from camera front.



(b) Map image generated.

Fig. 1: (b) Map generation from the (a) situation.

Thus, each of the entities must define a control input (linear and angular velocity is considered here), then a weighting will be carried out according to criteria such as safety and comfort.

However, this arbitration cannot depend solely on the command at the present time because:

- It is poor in information, not providing real information about the intention of each of the entities in the near future; it is therefore difficult to judge the quality of the order according to a criterion;
- It is outdated at the time of its processing by the fusion system;

We need to know for each entity its intentions on a short time horizon (a few seconds). For the autonomous system we can ask its intentions, it is however more difficult to ask to the human its intentions. But we can predict them from the previous information (previous vehicle dynamic states and sensors data). The objective of this paper is to implement a method for predicting human behavior on the ego-vehicle by exploiting different types of data with a deep learning model based on an LSTM architecture. These data can be raw data or not, the idea is to build a single model able to process these different data. Moreover, this model is exploited to illustrate the importance of data according to the situation

encountered by the car.

The article is structured as follows, the section II introduces related works, the section III explains the methodology of our work defining the data and the model used, the section IV shows the validation process and results then the section V discusses these results. Then, in this section, by revisiting previous works in a survey like description, we formally define driver intentions, the goal of the principal human prediction models.

II. RELATED WORKS

The prediction of the human driver is a task that seems feasible, because driving is constrained by the geometry of the road, the rules of the road but also by the intentions of other drivers, and the driver's controls can also indicate his intentions, for example, if a driver activates his right turn signal, it is very likely that he will make a right turn in a short period of time.

A. Driver Intentions Definitions

Driving intentions can be translated by a succession of maneuvers, trajectories or more generally by states of the vehicle. The article [17] proposes a generic formulation about driving intentions for the vehicle.

$$S = \{s_t, s_{t+1}, \dots, s_{t+m}\} \quad (1)$$

Where s_t is the state (velocity or position) at time step t .

The literature focuses on two types of predictions the behavior of the cars around the driven car [17], and directly the driven car [25].

B. Driver Prediction Target

Driving prediction studies focus on predicting the behavior of cars around the ego-vehicle [24] [19]. The driving of a vehicle depends on the infrastructure of the road, the rules of the road but also the behavior of other vehicles, for this reason the prediction of behaviors is important and allows taking into account the intentions of the vehicles around in the command generated by the autonomous system. In this case, the prediction is based on extrinsic vehicle data.

Other studies focus their research on the prediction of the ego-vehicle [25] [15]. In this case, the prediction uses the extrinsic data of the car but also the intrinsic data (e.g. used controls, dynamic state of the vehicle). The prediction of human driving can be used for ecological reasons, in fact to predict the behavior of the car that is driven in order to anticipate the future speeds of it and thus adopt the best strategy on the change of speed by saving fuel consumption. [15]. Moreover, prediction can also be used, in our case study, to make shared control [12]. With this prediction, the system performing the command fusion between the autonomous system and the driver can interpret the driver's choices and act accordingly.

C. Models And Data Used In Human Behavior Prediction

1) Models:

a) *Parametric and Non-parametric*: The human driver predictions models are divided into two categories parametric models and non-parametric models. The parametric model regroups physic models[14]. These models are based on theoretical concepts. These models have the particularity of being fast in execution, but are nevertheless too rigid models that are not able to adapt to a set of diverse and complex situations.

The non-parametric models regroups models based on data. Due to the complexity of the problem and the very rapid development of deep learning in recent years, research on the prediction of human driving has been strongly oriented towards the use of neural networks [14]. Thus, the model, unlike parametric-model based solutions, is able to make a complex representation of the data and finding relationships, it is however important to note that in this case the model may be biased by its learning [7].

Because of its temporality, the prediction of driver intentions could be modeled by a recurrent neural network [2]. Thus, the model does not only depend on current data but also on previous data and is able to establish relationships between these data. Avoiding the problems associated with vanishing/exploding gradient learning [8], recurrent networks are replaced by LSTM cells [9].

b) *Classification and Regression*: The prediction of human driving can be expressed in several forms, the next manoeuvre that the driver will perform ([27],[20]) or the state of the vehicle in the near future ([26],[4]). Depending on the nature of the prediction, the model must be adapted, the prediction of the maneuvers is usually based on classification, in this case the model is able to classify among a fixed number of maneuvers that the driver will perform. In the case of vehicle state prediction, the model estimates a sequence of states (position or speed); thus, the model must performs a regression and gets as close as possible to the numerical values (defining the vehicle state).

Some models [13] combine the two types of prediction by giving for each manoeuvre a sequence of possible states of the car.

2) Data:

a) *Data exploitable for the prediction*: The prediction of driving behavior depends on several factors:

- state of the environment (e.g. obstacles, road topology, driving of other cars);
- state of the car (e.g. speed, current acceleration);
- state of the controls (e.g. turn signal, steering wheel angle);
- driver status and profile (e.g. driver fatigue);

b) *Raw and semantic data*: Information can more or less reflect semantic information. A semantic data allows to provide a richer and more interpretable information (more meaningful on what it describes) as opposed a raw data is less interpretable without an extraction process. Raw data provide complete information but requires more processing load from the model, which usually has an impact on the execution time.

TABLE I: Sensors information.

sensor	frequency	additional information
LiDAR	10 Hz	40 layers 130 000 points per acquisition
camera	10 Hz	1280*720 resolution
GPS	50 Hz	cm precision
BUS-CAN reader	100 Hz	

D. Contributions

Most publications use high-level information in their model by adding a preprocessing phase to extract richer information. For example, the paper [25] does not use radar data directly, but extracts the position and speed of the vehicle in front. In the proposed approach, a network is created with multi-inputs models that can combine raw data with higher level data and to directly train the whole model without under-training. The paper [5] does not exploit directly raw data but a grid map representation. In addition, this architecture is also used to analyze the dependency of the model on different types of data depending on the situation.

III. METHODOLOGY

A. Our Problem Definition

The human driver prediction is exploited in order to predict the futur vehicle states or the futur manoeuvres of the vehicle. In our situation, the model has to predict the futur velocities (linear and angular) of the vehicle on a short time horizon because we want to do shared driving on local navigation, especially on lane tracking, in this context it is more important to know the future velocities.

Our problem can be defined as:

$$\begin{aligned}
 H_{\Theta}(X_t, X_{t-1}, X_{t-2}, \dots, X_{t-n}) &= (y_{t+1}, y_{t+2}, \dots, y_{t+m}) \\
 &= \begin{bmatrix} (v_{t+1}, w_{t+1}) \\ (v_{t+2}, w_{t+2}) \\ \dots \\ (v_{t+m}, w_{t+m}) \end{bmatrix}^T
 \end{aligned} \tag{2}$$

Where H_{Θ} represents the RNN system with Θ parameters, v_{t+1} linear velocity, w_{t+1} angular velocity at the time $t + 1$ and $(X_t, X_{t-1}, X_{t-2}, \dots, X_{t-n})$ represents the sequential input used in the learning phase of the model. The horizon time of the prediction is 5 seconds. This value was chosen based on previous studies [17] showing that the prediction error was significant beyond 5 seconds, and in the context of local driving behavior, a prediction of 5 seconds seems to be sufficient.

B. Data

The data used to train the model was acquired by a vehicle equipped with sensors. We were able to create a dataset of part of the city Compiègne in France.

1) *Vehicle Acquisition:* We have a vehicle with onboard sensors to collect data, the following Table I, gives information about these sensors.

2) *Roadmap And Data Acquisition:* A road map was established upstream to ensure the diversity of the dataset. The dataset was built with the following characteristics:

- Lanes: 1 and 2;
- Curved and straight curve;
- Speed limit: 30, 50, 70 and 90 km/h;
- With/without roundabouts;
- With/without intersections;

We save the following data:

- Environment:
 - LiDAR [points cloud];
 - Front camera [image];
- Vehicle dynamic state:
 - Velocity linear [m/s];
 - Velocity angular [rad/s];
 - Acceleration linear [$m.s^{-2}$];
 - Acceleration angular [$rad.s^{-2}$];
 - GPS positions [latitude and longitude];
- Vehicle controls state:
 - Steering wheel angle [rad];
 - Flashlight states [state value];

The recorded data are saved at 10 Hz. The training dataset is composed of about 32000 acquisitions regrouping what represents 3200 seconds of acquisitions.

3) *Data Processing:* All data are normalized to avoid scale influence and to improve the convergence of the model. Each of these data is normalized by the min-max rescaling normalization method. In addition to normalization, some of these data are pre-processed before they are submitted to the model.

a) *Rolling Data:* The acquisition of temporal sensor data, such as speed, acceleration, and steering wheel angle, adds noise to data. In order to prevent sensitivity to acquisition noise, we applied a centered moving average [3].

b) *Map Generation:* The position of the vehicle is accurate to the centimeter, we can use the information from a detailed map (such as lane positions, markers positions). By combining this information, we can generate an image with the geometry of the road around the vehicle. The image is built by projecting this information on a map, for this we created a virtual camera placed above the car, and we project the map information in the image frame. Representing the edges of the road, the lanes and the information on the ground. The figure 1 shows an example of generated map.

c) *LiDAR Data Reduction:* Per acquisition, the LiDAR gives us about 130000 points, it is not possible to exploit these points cloud directly into the learning system. The LiDAR points cloud data are projected into a 2-dimensional space by filtering out some points considered too high or too low and points behind the vehicle are also filtered. Then we average these points to reduce the final number of points.

C. The RNN LSTM Learning Model Structure

The proposed prediction model takes into account different inputs, different modalities of data, as shown in previous section III-B. The global architecture is schematized in

figure 2, showing the input and output models, are listed as follows:

1) *Input Models*: We set up a model for each type of input that will pre-process this data, by developing a Tensorflow-based framework. For each prediction, we submit 50 previous data per input model (= 5s of data). Depending on the nature of the input data, we can take two different approaches:

a) *Raw Data*: The raw data cannot be injected directly into the final model because it represents too much information of little significance, which would risk drowning the other more significant data. We decided to process on the data upstream, in order to compress the information as much as possible. This compression is based on an encoding model, like those found in auto-encoder models[16], the final model is given the encoded data of the encoding part. Among the list of data that we exploit (Section III-B.2), we have to encode data from the map image and from the LiDAR. Data from the map image are encoded by a VGG16 model [22] (convolution2D/pool2D layers then fully connected to dense layers) and the LiDAR is encoded by a similar model but in 1 dimension. Note that for the map, we use a single data and not a time series, as we consider that only the last data is needed for the model to predict the future behavior of the car. Therefore, we repeat the encoded vector so that the output can be adapted to other models.

b) *High Data*: The data does not need to be transformed and can be exploited directly. In the case a model upstream is not required, but it is used, in order to augment the input vector to balance it with the other outputs of the input models. We do not use the GPS data directly as model input, but it was used to build the road map.

2) *Output Model*: All the outputs of the input models are concatenated into a single time vector tensor. This vector is then submitted to a recurrent model composed of 2 LSTM layers (100 units) fully connected to 3 dense layers (1000, 200 and 100 units respectively). The output is reshaped into a vector of 50 velocities couples (v_t, w_t) , the delta time between 2 couples is 100 ms, then the total time of the prediction is 5 seconds.

IV. METHODOLOGY VALIDATION

A. Dataset Test

The test dataset was constructed to test several situations to best assess the fit of the proposed model to the situation. The table II shows information about these tests.

B. Metrics

We used different metrics (errors and accuracies) to evaluate the quality of our model. We computed the different metrics on the first N predictions, this way we can evaluate the quality of the prediction on several horizon times (1,2,3 and 5 seconds).

- Mean Squared Error (limit to N first time predictions):

$$MSE_N = \frac{1}{N} \sum_{i=1}^N (Y_i - \hat{Y}_i) \quad (3)$$

- Root Mean Absolute Error on index j (v index or w index):

$$RMSE_N^j = \sqrt{\frac{1}{N} \sum_{i=1}^N |Y_{(i,j)} - \hat{Y}_{(i,j)}|} \quad (4)$$

- R^2 on index j (v index or w index):

$$(R^2)_N^j = 1 - \frac{\sum_{i=1}^N (Y_{(i,j)} - \hat{Y}_{(i,j)})}{\sum_{i=1}^N (Y_{(i,j)} - \bar{Y}_{(i,j)})} \quad (5)$$

- distance (limit to N predictions):

$$d = \|\text{pos}_N(Y_N) - \text{pos}_N(\hat{Y}_N)\|_2 \quad (6)$$

Where pos_i is the relative position of the vehicle after applied i first velocities. The position is estimated from velocities (in considering the d_t interval time between each velocity couple).

Let k the size of the batch tested and metric_N the metric tested depending on N , then the final metric is the average on this batch:

$$\text{metric}_{N,\text{batch}} = \frac{1}{k} \sum_{i=1}^k \text{metric}_N(H_{\Theta}(\text{batch}[i])) \quad (7)$$

C. Results

1) *Model Validation*: The model was tested on the dataset test, the Figure 3 shows an example of the model prediction realized on the test "roundabout".

The following Tables III and IV shows metrics (defined in section IV-B) results per test perform by the model. We averaged each metric over several N first elements showing the quality of our model over different horizon times, and thus we can see the evolution of the prediction.

2) *Sensitivity Of The Model To Data*: In addition to evaluating the quality of our model, we can examine the sensitivity of our model to the data. To do this, we have tests in which we hide the information of some data by placing them in a neutral state. The idea is to interpret in which situation the model exploits these data. The following table V shows the neutral data corresponding for each data tested. The tables VI and VII shows the result metrics (for $N = 50$) for each test defined in section IV-A.

V. DISCUSSION

A. Test Results Analyze

The results of the tests allow us to conclude that the model is able to predict on a very close time horizon the human driver intentions based on data intrinsic of the vehicle (sensors and vehicle information). The average error on the linear velocity is less than 3 km/h in all the different tests and for a horizon time of 5 seconds. In a general way (excluding roundabout test) we can make a remark that this prediction degrades according to the duration of the horizon, more the duration of the horizon is important more the quality of the prediction will be impacted. Indeed, the angular velocity prediction error is 2x larger at 2 seconds than at 1 second, while the linear velocity prediction error has only increased

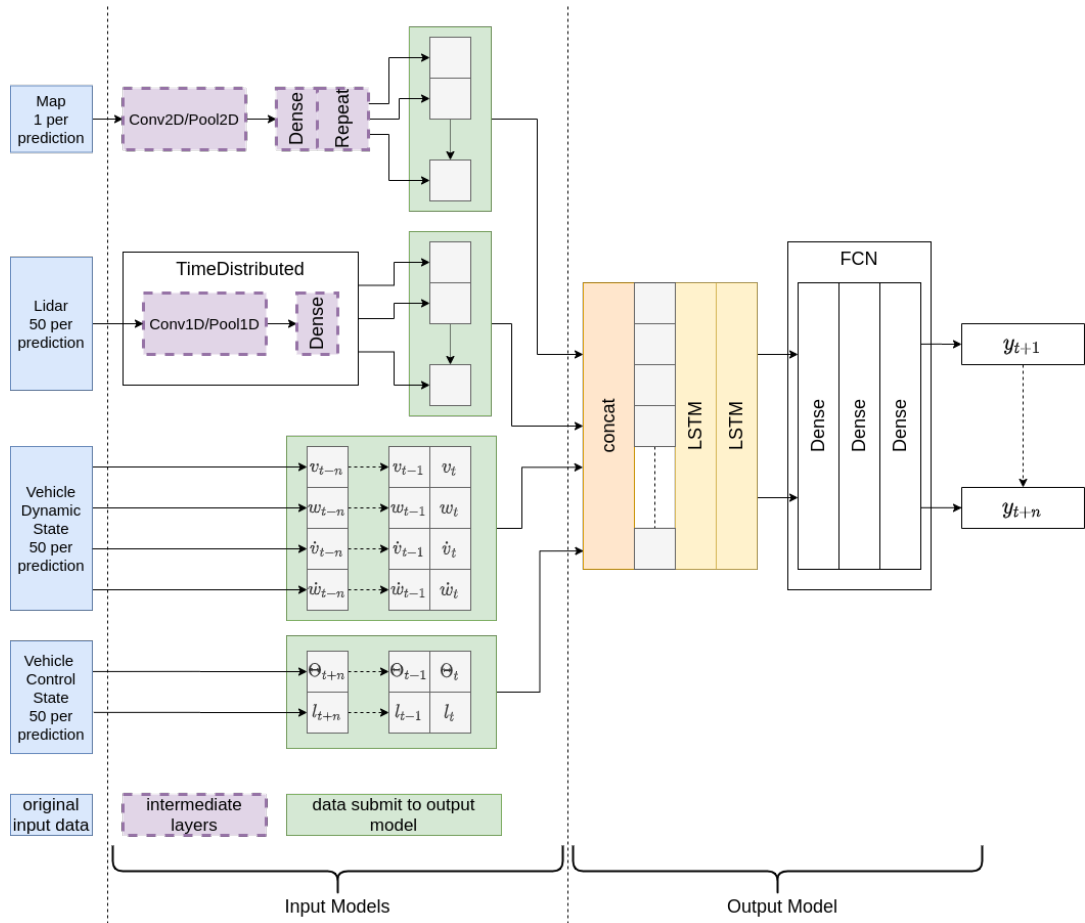
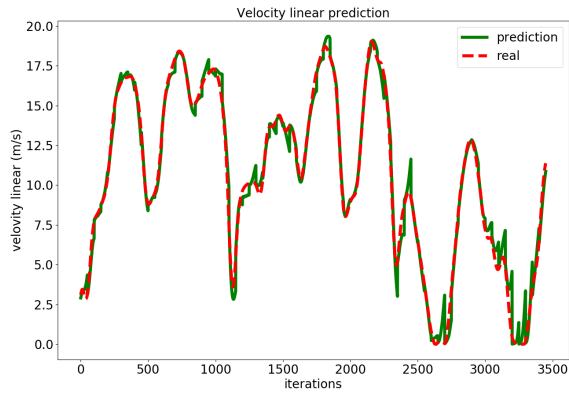


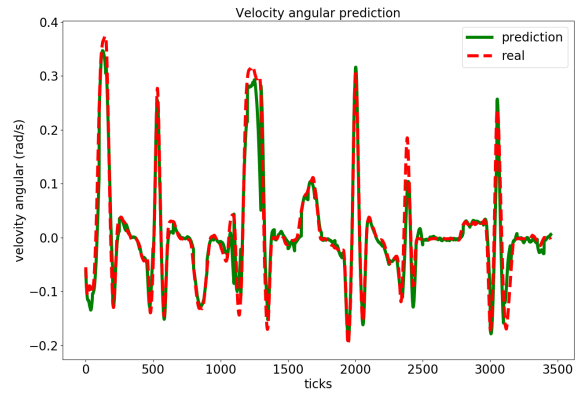
Fig. 2: Global model architecture.

TABLE II: Tests information.

test name	n roundabouts	n intersections	speed limit	distance	time record	n lanes
roundabout	6	1	70 km/h	4 km	378 s	2
city	4	7	50 km/h	4 km	519 s	1
speed (1 lane)	2	0	70 km/h	2 km	116 s	1
speed (2 lanes)	0	0	90 km/h	2 km	84 s	2



(a) Velocity linear prediction.



(b) Velocity angular prediction.

Fig. 3: Predictions realized by our model on the test "roundabout".

TABLE III: Average result of error metrics (of N first elements predicted) per test, the percentage represents the difference with the error for $N = 10$.

metric	N first elements	test name							
		roundabout		city		speed (1 lane)		speed (2 lanes)	
MSE	10 (=1s)	0,000145	0%	0,000226	0%	0,000376	0%	5,40E-05	0%
	20 (=2s)	0,00024	+66%	0,000514	+127%	0,000742	+97%	5,90E-05	+9%
	30 (=3s)	0,000367	+153%	0,001001	+343%	0,000898	+139%	7,10E-05	+31%
	50 (=5s)	0,000907	+526%	0,002734	+1110%	0,001284	+241%	0,000138	+156%
RMSE v (km/h)	10 (=1s)	0,579851	0%	0,881858	0%	0,945485	0%	0,937452	0%
	20 (=2s)	0,694269	+20%	1,055049	+20%	1,169119	+24%	1,15546	+23%
	30 (=3s)	0,907557	+57%	1,403347	+59%	1,580911	+67%	1,552372	+66%
	50 (=5s)	1,680842	+190%	2,69627	+206%	2,972805	+214%	2,893665	+209%
RSME w (rad/s)	10 (=1s)	0,013256	0%	0,012517	0%	0,014425	0%	0,014175	0%
	20 (=2s)	0,015753	+19%	0,017539	+40%	0,021773	+51%	0,021419	+51%
	30 (=3s)	0,018597	+40%	0,022506	+80%	0,028685	+99%	0,027998	+98%
	50 (=5s)	0,029727	+124%	0,03423	+173%	0,04324	+200%	0,041808	+195%
distance (m)	10 (=1s)	0,154637	0%	0,208928	0%	0,246685	0%	0,187872	0%
	20 (=2s)	0,404201	+161%	0,538614	+158%	0,694432	+182%	0,408118	+117%
	30 (=3s)	0,812952	+426%	1,200391	+475%	1,400716	+468%	0,638761	+240%
	50 (=5s)	2,813853	+1720%	3,949019	+1790%	3,672162	+1389%	1,431308	+662%

TABLE IV: Average result of accuracy metrics (of N first elements predicted) per test, the percentage represents the difference with the error for $N = 10$.

metric	N first elements	test name							
		roundabout		city		speed (1 lane)		speed (2 lanes)	
R2 v	10 (=1s)	0,998685	0%	0,995846	0%	0,996815	0%	0,975354	0%
	20 (=2s)	0,998151	-0%	0,993452	-0%	0,995105	-0%	0,972998	-0%
	30 (=3s)	0,996847	-0%	0,98746	-1%	0,992759	-0%	0,967615	-1%
	50 (=5s)	0,987465	-1%	0,954953	-4%	0,984349	-1%	0,940597	-4%
R2 w	10 (=1s)	0,986572	0%	0,969484	0%	0,950148	0%	0,367955	0%
	20 (=2s)	0,974036	-1%	0,921467	-5%	0,875317	-8%	0,398099	+8%
	30 (=3s)	0,959301	-3%	0,844294	-13%	0,805404	-15%	0,366389	-0%
	50 (=5s)	0,921706	-7%	0,593164	-39%	0,723784	-24%	0,061333	-83%

TABLE V: Neutral data corresponding.

data	neutral data
map image	zeros vector (= black image, map missing)
scan	ones vector (= no obstacle)
control state	zeros vector (= flashlight never on and steering wheel angle fixed to middle position)

by 20/25%. The error on the angular velocity prediction increases faster than the error on the linear velocity.

The model is quite capable of making predictions in situations with roundabouts, the "roundabout" test shows that the model makes good predictions (linear and angular) without degrading over time.

The quality of the angular prediction on the fast lane (2 lanes) is very poor, this can be explained by a variation of the angular velocity which is very low and that our model is not able to have this accuracy.

The R2 scores confirm the analysis and show us that the quality of the angular velocity prediction deteriorates rapidly compared to the linear velocity prediction. We notice that the R2 for angular velocity in the city test decreases rapidly with time, this reflects that the quality of prediction deteriorates rapidly with time.

B. Sensitive Data Analyze

The second study allows us to interpret the sensitivity of the model to the type of data according to the tests. From results (tables VI and VII), we can notice that the

model needs to know the map data to improve the angular velocity prediction, especially in urban areas. This means that the model has inferred the road geometry in the angular velocity prediction. The control state data of car flashlight and steering wheel angle are important for linear and angular velocity prediction. Its absence induces a prediction error on the linear velocity of at least 2 times larger than the reference prediction. In contrast, we find that the absence of LiDAR data does not create any perturbation on the model prediction.

VI. CONCLUSION

In order to predict human behavior, we built a recurrent model (LSTM) able to predict the intentions of human, described by the futur velocities (linear and angular) of the vehicle. The architecture in multi-inputs model allowing the inclusion of different types of data.

The results show us that the model has successfully interpreted these data and is able to predict human behavior. In addition, the model allowed us to interpret how the model interprets the data and to understand in what situation the model uses these data, and understand the data important to make this prediction.

In the future work, we want to build a similar model capable of predicting less predictable and more dangerous behaviors, we plan to include simulated data to add dangerous driving data.

TABLE VI: Average RMSE error results with neutral data injection, the percentage represents the difference with the reference data.

	rounabout		city		speed (1 lane)		speed (2 lanes)	
	RMSE v	RMSE w	RMSE v	RMSE w	RMSE v	RMSE w	RMSE v	RMSE w
reference	1.680842	0.029727	2.69627	0.03423	2.972805	0.04324	2.893665	0.041808
neutral map	62%	102%	23%	108%	14%	76%	14%	76%
neutral scan	26%	4%	20%	4%	14%	1%	14%	1%
neutral state	237%	62%	132%	46%	109%	31%	112%	31%

TABLE VII: Average R2 score results with neutral data injection, the percentage represents the difference with the reference data.

	rounabout		city		speed (1 lane)		speed (2 lanes)	
	R2 v	R2 w	R2 v	R2 w	R2 v	R2 w	R2 v	R2 w
reference	0,987465	0,921706	0,954953	0,593164	0,984349	0,723784	0,940597	0,061333
neutral map	-1%	-33%	-1%	-127%	0%	-20%	-8%	-4738%
neutral scan	-1%	-1%	-1%	0%	0%	1%	-2%	-545%
neutral state	-8%	-12%	-12%	-40%	-7%	-13%	-68%	-3226%

REFERENCES

- [1] David A Abbink, Senior Member, Tom Carlson, Mark Mulder, Joost C F De, Farzad Aminravan, Tricia L Gibo, and Erwin R Boer. A Topology of Shared Control Systems – Finding Common Ground in Diversity. *IEEE Transactions on Human-Machine Systems*, 48(5):509–525, 2018.
- [2] Florent Alché and Arnaud de La Fortelle. An lstm network for highway trajectory prediction. In *2017 IEEE 20th international conference on intelligent transportation systems (ITSC)*, pages 353–359. IEEE, 2017.
- [3] Florent Alche and Arnaud De La Fortelle. An LSTM network for highway trajectory prediction. *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, 2018-March:353–359, 2018.
- [4] Shengzhe Dai, Li Li, and Zhiheng Li. Modeling vehicle interactions via modified lstm models for trajectory prediction. *IEEE Access*, 7:38287–38296, 2019.
- [5] Nachiket Deo and Mohan M Trivedi. Convolutional social pooling for vehicle trajectory prediction. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 1468–1476, 2018.
- [6] Thomas A Dingus, Feng Guo, Suzie Lee, Jonathan F Antin, Miguel Perez, Mindy Buchanan-King, and Jonathan Hankey. Driver crash risk factors and prevalence evaluation using naturalistic driving data. *Proceedings of the National Academy of Sciences*, 113(10):2636–2641, 2016.
- [7] Sorin Grigorescu, Bogdan Trasnea, Tiberiu Cocias, and Gigel Macesanu. A survey of deep learning techniques for autonomous driving. *Journal of Field Robotics*, 37(3):362–386, 2020.
- [8] Roger Grosse. Lecture 15: Exploding and vanishing gradients. *University of Toronto Computer Science*, 2017.
- [9] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9:1735–80, 12 1997.
- [10] Rasheed Hussain, JooYoung Lee, and Sherali Zeadally. Autonomous cars: Social and economic implications. *IT Professional*, 20(6):70–77, 2018.
- [11] Kichun Jo, Junsoo Kim, Dongchul Kim, Chulhoon Jang, and Myoung-ho Sunwoo. Development of autonomous car—part i: Distributed system architecture and development process. *IEEE Transactions on Industrial Electronics*, 61(12):7131–7140, 2014.
- [12] Shriram C Jugade, Alessandro C Victorino, Véronique B Cherfaoui, and Stratis Kanarachos. Sensor based prediction of human driving decisions using feed forward neural networks for intelligent vehicles. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, pages 691–696. IEEE, 2018.
- [13] Namhoon Lee, Wongun Choi, Paul Vernaza, Christopher B Choy, Philip HS Torr, and Manmohan Chandraker. Desire: Distant future prediction in dynamic scenes with interacting agents. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 336–345, 2017.
- [14] Stéphanie Lefèvre, Dizan Vasquez, and Christian Laugier. A survey on motion prediction and risk assessment for intelligent vehicles. *ROBOMECH journal*, 1(1):1–14, 2014.
- [15] Jing Lian, Shuang Liu, Linhui Li, Xuanzuo Liu, Yafu Zhou, Fan Yang, and Lushan Yuan. A mixed logical dynamical-model predictive control (mld-mpc) energy management control strategy for plug-in hybrid electric vehicles (phevs). *Energies*, 10:74, 01 2017.
- [16] Qinxue Meng, Daniel Catchpoole, David Skillicom, and Paul J. Kennedy. Relational autoencoder for feature extraction. In *2017 International Joint Conference on Neural Networks (IJCNN)*, pages 364–371, 2017.
- [17] Sajjad Mozaffari, Omar Y. Al-Jarrah, Mehrdad Dianati, Paul Jennings, and Alexandros Mouzakitis. Deep Learning-Based Vehicle Behavior Prediction for Autonomous Driving Applications: A Review. *IEEE Transactions on Intelligent Transportation Systems*, pages 1–15, 2020.
- [18] Ryosuke Okuda, Yuki Kajiwara, and Kazuaki Terashima. A survey of technical trend of adas and autonomous driving. In *Technical Papers of 2014 International Symposium on VLSI Design, Automation and Test*, pages 1–4. IEEE, 2014.
- [19] Seong Hyeon Park, ByeongDo Kim, Chang Mook Kang, Chung Choo Chung, and Jun Won Choi. Sequence-to-sequence prediction of vehicle trajectory via lstm encoder-decoder architecture. In *2018 IEEE Intelligent Vehicles Symposium (IV)*, pages 1672–1678. IEEE, 2018.
- [20] Derek J Phillips, Tim A Wheeler, and Mykel J Kochenderfer. Generalizable intention prediction of human drivers at intersections. In *2017 IEEE intelligent vehicles symposium (IV)*, pages 1665–1670. IEEE, 2017.
- [21] Thomas B Sheridan and William L Verplank. Human and computer control of undersea teleoperators. Technical report, Massachusetts Inst of Tech Cambridge Man-Machine Systems Lab, 1978.
- [22] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.
- [23] Jianqiang Wang, Heye Huang, Keqiang Li, and Jun Li. Towards the unified principles for level 5 autonomous vehicles. *Engineering*, 7(9):1313–1325, 2021.
- [24] Long Xin, Pin Wang, Ching-Yao Chan, Jianyu Chen, Shengbo Eben Li, and Bo Cheng. Intention-aware long horizon trajectory prediction of surrounding vehicles using dual lstm networks. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, pages 1441–1446. IEEE, 2018.
- [25] Kyuhwan Yeon. EGO-VEHICLE SPEED PREDICTION USING A LONG SHORT-TERM MEMORY BASED RECURRENT NEURAL NETWORK. *International Journal of . . .*, 13(2):293–300, 2019.
- [26] Tianyang Zhao, Yifei Xu, Mathew Monfort, Wongun Choi, Chris Baker, Yibiao Zhao, Yizhou Wang, and Ying Nian Wu. Multi-agent tensor fusion for contextual trajectory prediction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 12126–12134, 2019.
- [27] Alex Zyner, Stewart Worrall, James Ward, and Eduardo Nebot. Long short term memory for driver intent prediction. In *2017 IEEE Intelligent Vehicles Symposium (IV)*, pages 1484–1489. IEEE, 2017.