Empowering Advanced Driver-Assistance Systems from Topological Data Analysis

Tarek Frahi [1],*, Francisco Chinesta [1], Antonio Falcó [2] , Alberto Badias [3], Elias Cueto [3] , Hyung Yun Choi [4], Manyong Han [5] and Jean-Louis Duval [6]

[1] PIMM Lab, Arts et Metiers Institute of Technology, 151 boulevard de l'Hopital, 75013 Paris, France; Francisco.Chinesta@ensam.eu

[2] Departamento de Matemáticas, Física y Ciencias Tecnológicas, Universidad Cardenal Herrera-CEU, CEU Universities, San Bartolome 55, 46115 Alfara del Patriarca, Valencia, Spain; afalco@uch.ceu.es

[3] I3A, Aragon Institute of Engineering Research, Universidad de Zaragoza, 50018 Zaragoza, Aragon, Spain; abadias@unizar.es (A.B.); ecueto@unizar.es (E.C.)

[4] Department of Mechanical and System Design Engineering, Hongik University, 94 Wausanro, Mapogu, Seoul 04066, Korea; hychoi@hongik.ac.kr

[5] Digital Human Lab, Hongik University, 94 Wausanro, Mapogu, Seoul 04066, Korea; myhan0521@gmail.com

[6] ESI Group, 3bis rue Saarinen, CEDEX 1, 94528 Rungis, France; Jean-Louis.Duval@esi-group.com

* Correspondence: tarek.frahi@ensam.eu

Abstract: We are interested in evaluating the state of drivers to determine whether they are attentive to the road or not by using motion sensor data collected from car driving experiments. That is, our goal is to design a predictive model that can estimate the state of drivers given the data collected from motion sensors. For that purpose, we leverage recent developments in topological data analysis (TDA) to analyze and transform the data coming from sensor time series and build a machine learning model based on the topological features extracted with the TDA. We provide some experiments showing that our model proves to be accurate in the identification of the state of the user, predicting whether they are relaxed or tense.

1. Introduction

While there have recently been considerable advances in self-driving car technology, driving still relies mainly on human factors. Even in self-driving mode, human drivers must often make decision in a fraction of a second to avoid accidents. Therefore, it is still of utmost importance to develop systems capable of discerning if the human driver is attentive or not to the road conditions. In general, the so-called advanced driver assistance systems (ADAS) [1,2] are systems that are able to improve the driver's performance, among which, adaptive speed limiters, pedestrian detectors [3], and cruise controllers are some of the most popular systems. Fatigue alerting systems are among the most useful among ADAS systems, and the aim of this work is to contribute to the development of such a system based on a systematic analysis of drivers in actual driving conditions.

The estimation of the driver's condition (degree of attention to the road, fatigue, etc.) is a very important factor to ensure safety in driving [4,5]. A recent review on the topic can be found in [6]. The goal of this work is to extract behavior patterns from car user data to be able to accurately estimate their state. We used data obtained by the laboratory of prof. Hyung Yun Choi at Hongik University in Seoul. His experiment involved the application of mechanical stimulation to people seated in an automobile.

Our main goal is to extract patterns of behavior from experimental data so as to allow us to learn the most relevant factors affecting driver's attention to the situation of the road.

In the present work, we combine some tools from Morse theory [7] and topological data analysis (TDA) with all of the associated concepts and methods (e.g., Betti numbers,

homology persistence, barcodes, persistence images, etc.) [8], most of them introduced and employed later in order to analyze and classify the experimental data. This allows us to introduce concepts as barcodes, that is, persistent and life-time diagrams in a similar way to how they are used in persistent homology. Our main goal is to predict car user behavior following a supervised approach [9]. Instead of considering an original sensor signal as the quantity of interest, we focus on its topological features. In this sense, the framework proposed in this paper allows us to unveil the true dimensionality of data or, in other words, the actual number of factors affecting driver's performance. Thus, we model a sensor signal as a dynamical system, and, therefore, our approach seems to be better at describing its properties, or rather its variations, such as extrema, patterns, and self-similarity, than other approaches. We note that our approach is, in some senses, similar to that followed by Milnor and Thurston [10] in the study of the combinatorial properties of dynamical systems by combining tools from automata theory.

The structure of the paper is as follows: In Section 2, we describe the material and methods employed in this work. Particular attention is paid to the process of data acquisition and the description of time series and data curation. In Section 3, we present the main results of this work, and we discuss the main consequences in Section 4. As a complement, in Appendix A, we thoroughly illustrate the process of computing persistence images for the data of interest.

2. Material and Methods

In this section, we describe the collection and preprocessing of the experimental data. In Section 2.1, we describe the data acquisition, and in Section 2.2, we provide a description of the time series. Section 2.3 is devoted to data preprocessing. The mathematical tools used to describe the times series at a topological level are explained in Section 2.4. Finally, the image classification methodology is given in Section 2.5.

2.1. Data Acquisition

Our proposed predictor directly uses the data collected from the experiments. The data acquisition process involves measuring the response of human behavior when an excitation is applied to the seat. Figure 1 shows the location of the sensors in the experiments.
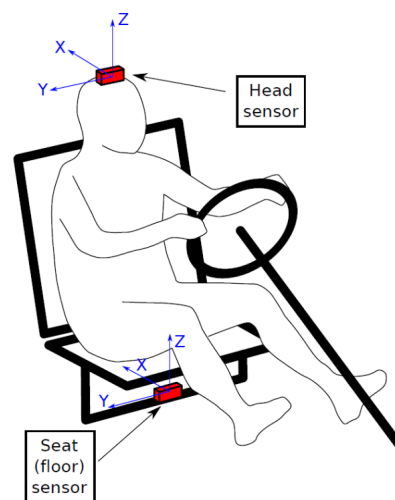
Figure 1. Scheme of the data acquisition process showing the location of the sensors.

The excitation signal is an angular acceleration imposed on the seat of the user. This acceleration is an oscillating chirp function with a frequency range of 1 to 7.5 Hz on the X axis in rotation. The linear acceleration $\mathbf{a} = (a_x, a_y, a_z)$ and angular velocity $\boldsymbol{\omega} = (\omega_x, \omega_y, \omega_z)$ were measured in both the head and the seat by two IMU (Shimmer inertia measurement

unit (IMU) sensors) at 256 Hz. By observing the floor excitation signals, we noted that the excitation is purely rotational around the X-axis—see Figure 2.
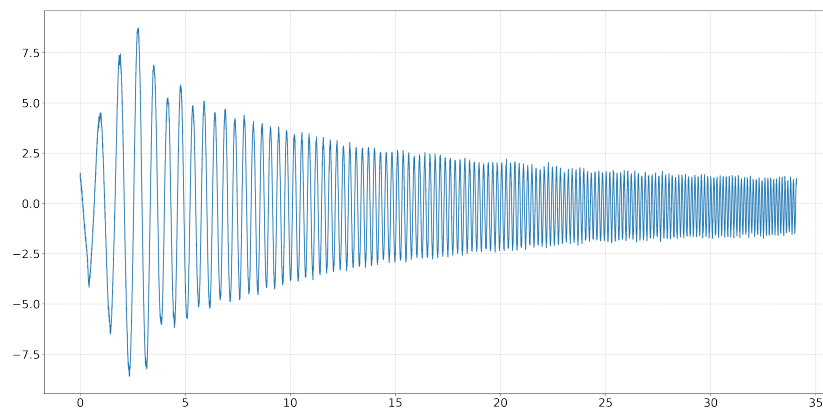

Figure 2. Floor excitation: X-axis angular velocity time series.

Several experiences were conducted by nine people by taking into account a set of six fixed states: driver, passenger, tense person, relaxed person, rigid seat, and SAV (sport activity vehicle seat). In particular, for each individual, eight experiments for the six available states were performed:

$$
\left\{
\begin{array}{cl}
\textbf{Class} & \textbf{Label} \\
1 & \text{SAVRelaxedPassager} \\
2 & \text{SAVTensePassager} \\
3 & \text{SAVRelaxedDriver} \\
4 & \text{SAVTenseDriver} \\
5 & \text{RigidRelaxedPassager} \\
6 & \text{RigidTensePassager} \\
7 & \text{RigidRelaxedDriver} \\
8 & \text{RigidTenseDriver}
\end{array}
\right.
$$

As a consequence, we worked with a sample of 72 experiences, each of them encoded in a time series (as we explain later). Our goal is to classify the behavior of a generic driver, assigning one of the two states (tense or relaxed) by using the sensor data.

2.2. Time Series Description

The data acquired from sensors (see Figures 3 and 4) were stored into six-dimensional time series, for both linear acceleration and angular velocity of the head movement. The sampling frequency of the data was 256 Hz, and the duration of the experiment was 34 s; hence, the resulting data dimensionality is $256 \times 34 = 8704$. For each times series, where $1 \leq t \leq 8704$, we constructed three new times series called the sliding window, embedding a length of 5800. The first one is given by the times values from $t = 1$ to $t = 5800$, the second is given by the times values from $t = 1450$ to $t = 7250$, and, to conclude, the third time window is defined as from $t = 2904$ to $t = 8704$. Each element in the sample ($1 \leq i \leq 72$) was encoded by means of three six-dimensional time series representing each of the three sliding windows that we represent in matrix form as follows:

$$
TS_{3(i-1)+1} = \begin{bmatrix}
a_x^\ell(1) & a_x^\ell(2) & \cdots & a_x^\ell(5800) \\
a_y^\ell(1) & a_y^\ell(2) & \cdots & a_y^\ell(5800) \\
a_z^\ell(1) & a_z^\ell(2) & \cdots & a_z^\ell(5800) \\
\omega_x^\ell(1) & \omega_x^\ell(2) & \cdots & \omega_x^\ell(5800) \\
\omega_y^\ell(1) & \omega_y^\ell(2) & \cdots & \omega_y^\ell(5800) \\
\omega_z^\ell(1) & \omega_z^\ell(2) & \cdots & \omega_z^\ell(5800)
\end{bmatrix}, \quad
TS_{3(i-1)+2} = \begin{bmatrix}
a_x^\ell(1450) & a_x^\ell(1451) & \cdots & a_x^\ell(7251) \\
a_y^\ell(1450) & a_y^\ell(1451) & \cdots & a_y^\ell(7251) \\
a_z^\ell(1450) & a_z^\ell(1451) & \cdots & a_z^\ell(7251) \\
\omega_x^\ell(1450) & \omega_x^\ell(1451) & \cdots & \omega_x^\ell(7251) \\
\omega_y^\ell(1450) & \omega_y^\ell(1451) & \cdots & \omega_y^\ell(7251) \\
\omega_z^\ell(1450) & \omega_z^\ell(1451) & \cdots & \omega_z^\ell(7251)
\end{bmatrix}
$$

and

$$TS_{3i} = \begin{bmatrix} a_x^\ell(2903) & a_x^\ell(2905) & \cdots & a_x^\ell(8704) \\ a_y^\ell(2903) & a_y^\ell(2905) & \cdots & a_y^\ell(8704) \\ a_z^\ell(2903) & a_z^\ell(2905) & \cdots & a_z^\ell(8704) \\ \omega_x^\ell(2903) & \omega_x^\ell(2905) & \cdots & \omega_x^\ell(8704) \\ \omega_y^\ell(2903) & \omega_y^\ell(2905) & \cdots & \omega_y^\ell(8704) \\ \omega_z^\ell(2903) & \omega_z^\ell(2905) & \cdots & \omega_z^\ell(8704) \end{bmatrix}.$$

Here, the matrices have a size of 6×5800 and $1 \leq i \leq 72$. This allows us to represent the information by using a third-order tensor, namely, $\mathcal{Z} \in \mathbb{R}^{216 \times 6 \times 5800}$ defined by

$$\mathcal{Z}_{i,j,k} := (TS_i)_{j,k}$$

for $1 \leq i \leq 216$, $1 \leq j \leq 6$ and $1 \leq k \leq 5800$. We can identify $Z_i = TS_i$ for $1 \leq i \leq 216$.

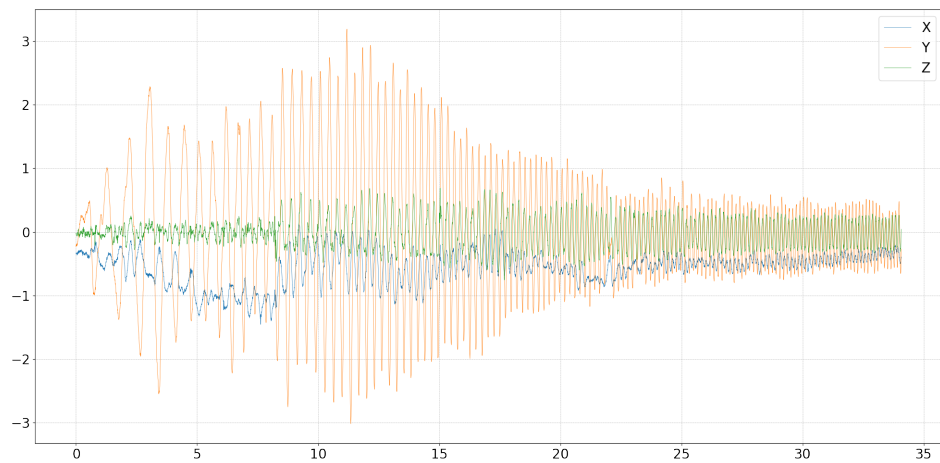**Figure 3.** Sensor data: linear acceleration time series.
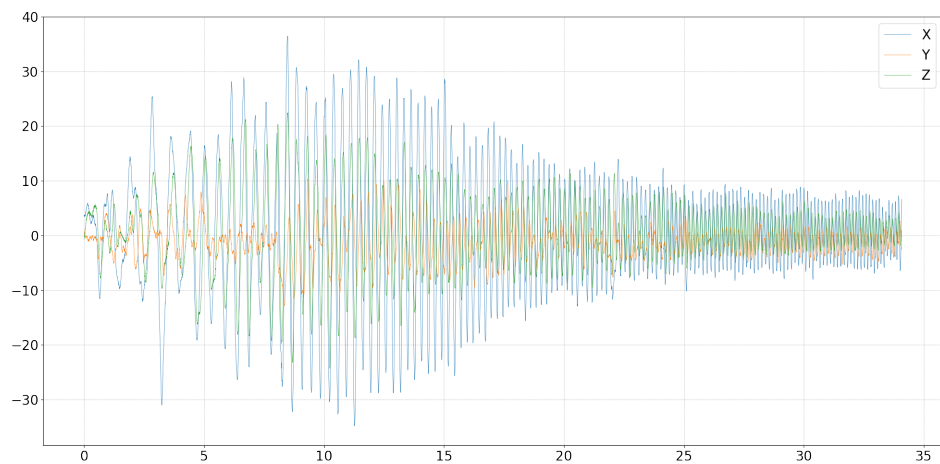

Figure 4. Sensor data: angular velocity time series.

2.3. Data Preprocessing

In order to obtain a single series for each observation, we concatenated all of the 6 time series (linear accelerations and angular velocities) for each observation horizontally and then created a data frame by stacking the 216 in sample observations.

The concatenation operation on the multidimensional time series collapsed the last two dimensions into one dimensional arrays with a length of $5800 \times 6 = 34{,}800$. The result is the two-dimensional table of concatenated time series

$$D = \begin{bmatrix} vec(\mathcal{Z}_{1,:,:}) \\ \dots \\ vec(\mathcal{Z}_{216,:,:}) \end{bmatrix} \in \mathbb{R}^{216 \times 34800}.$$

We chose not to filter the signals because the topological sub-level set method should filter the high-frequency features naturally. We also chose to keep working on acceleration signals in order to avoid signal deviations after two integrations in time so as to obtain positions, the sensors not always keeping a zero mean height. Thus, the approach is completely (topologically) data-based.

The six time series $\mathcal{Z}_i$ of each observation were collapsed into a single concatenated time series with a size of 34,800—see Figure 5. The concatenated time series for the 216 observations were then stacked to create the dataset D with a size of $216 \times 34{,}800$. We also used binary labels in the chained time series $\mathcal{Z}_i$ on the two target classes that we were interested in. In particular, we wrote $\mathcal{Z}_i^{(\alpha)}$ where α is "0" for a relaxed driver and "1" for a tense one.
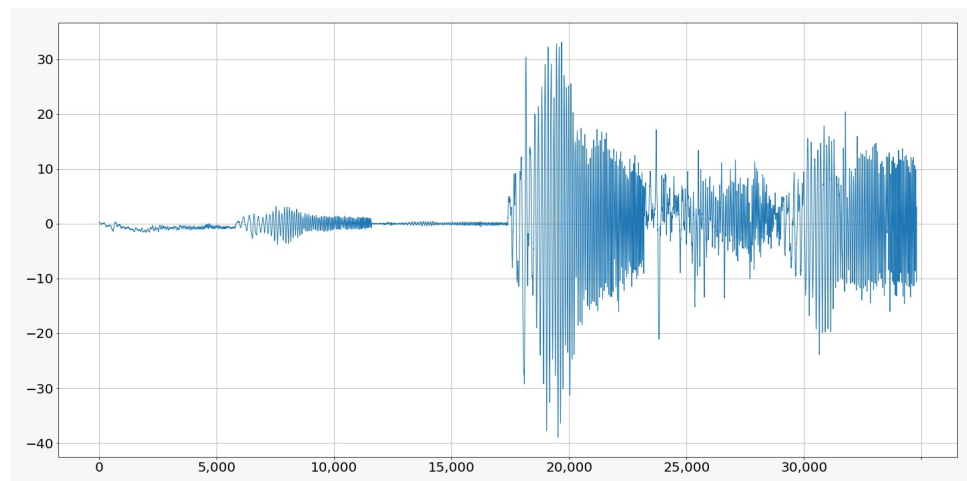

Figure 5. Tensor reduction of a sensor time series.

2.4. Extracting Topological Features from a Time Series

The idea to extract the topological information regarding the times series is to consider each sample observation as a piecewise linear continuous map from a closed interval to the real line. Therefore, we used a construction closely related to the Reeb graph [11] used in Morse theory to describe the times series at the topological level.

To this end, we consider the time series x_t for $0 \le t \le N - 1$ ($N \ge 3$) given by a vector

$$\mathbf{X} = (x_0, x_1, \dots, x_{N-1}) \in \mathbb{R}^N.$$

we can view $\mathbf{X}$ as a function also denoted by $\mathbf{X} : \{0, 1, \dots, N-1\} \longrightarrow \mathbb{R}$ defined by $\mathbf{X}(i) = x_i$ for $0 \le i \le N - 1$. Here, to study the topological features of $\mathbf{X}$ we use the sub-level set of a piecewise-linear function $f_{\mathbf{X}} : \mathbb{R} \longrightarrow \mathbb{R}$ associated with $\mathbf{X}$ satisfying that $f_{\mathbf{X}}(i) = \mathbf{X}(i) = x_i$ for $0 \le i \le N - 1$.

To construct this function, we consider the basis functions $\{\varphi_0, \dots, \varphi_{N-1}\}$ of continuous functions $\varphi_i : \mathbb{R} \longrightarrow \mathbb{R}$ defined by

$$\varphi_i(s) := \begin{cases} s - i + 1 & \text{if} \quad i - 1 \le s \le i \\ i + 1 - s & \text{if} \quad i \le s \le i + 1 \\ 0 & \text{if} \quad s \notin]i - 1, i + 1[\end{cases}$$

where $i = 1, \ldots, N - 2$ and

$$\varphi_0(s) := \begin{cases} 1 - s & \text{if} \quad 0 \leq s \leq 1 \\ 0 & \text{if} \quad s \in [0, 1] \end{cases}$$

$$\varphi_{N-1}(s) := \begin{cases} s - N + 2 & \text{if} \quad N - 2 \leq s \leq N - 1 \\ 0 & \text{if} \quad s \notin [N - 2, N - 1[\end{cases}$$

This allows us to construct a piecewise continuous map $f_{\mathbf{X}} : \mathbb{R} \longrightarrow \mathbb{R}$ by

$$f_{\mathbf{X}}(s) = \sum_{j=0}^{N-1} x_j \varphi_j(s),$$

and also to endow $\mathbb{R}^N$ with a norm given by

$$\|\mathbf{X}\| := \|f_{\mathbf{X}}\|_{L^2(\mathbb{R})} = \left(\int_{-\infty}^{\infty} |f_{\mathbf{X}}(s)|^2 ds \right)^{1/2}.$$

In particular, we prove the following result, which helps us to identify the time series given by the vector $\mathbf{X}$ in $\mathbb{R}^N$ with the function $f_{\mathbf{X}}$ in $L^2(\mathbb{R})$.

Proposition 1. *The linear map* $\Phi : (\mathbb{R}^N, \| \cdot \|) \longrightarrow (L^2(\mathbb{R}), \| \cdot \|_{L^2(\mathbb{R})})$ *given by* $\Phi(\mathbf{X}) = f_{\mathbf{X}}$ *is an injective isometry between Hilbert spaces. Furthermore,* $\Phi(\mathbb{R}^N)$ *is a closed subspace in* $L^2(\mathbb{R}^N)$.

Proof. The map is clearly isometric and injective because $\{\varphi_0, \ldots, \varphi_{N-1}\}$ is a set of linear independent functions in $L^2(\mathbb{R})$. $\square$

Here, we describe the maps $f_{\mathbf{X}} \in \Phi(\mathbb{R}^N)$ at the combinatorial level using the connected components (intervals) associated with its λ sub-level sets

$$\mathcal{LS}_\lambda(f_{\mathbf{X}}) := \{x \in [0, N - 1] : f_{\mathbf{X}}(x) \leq \lambda\}$$

for $\lambda \in \mathbb{R}$. For this purpose, we introduce the following distinguished objects related to the $\text{supp}(f_{\mathbf{X}}) = [0, N - 1] \subset \mathbb{R}$ of $f_{\mathbf{X}}$:

- The nodes or vertices denoted by

$$\mathcal{V} := \{[0], [1], \ldots, [N - 1]\}$$

that represent the components of the vector $\mathbf{X}$,;

- The faces denoted by

$$\mathcal{F} := \{[0, 1][1, 2], \ldots, [N - 2, N - 1]\}$$

that represent the intervals used to construct the connected components of the sub-level sets of the map $f_{\mathbf{X}}$. Recall that we consider

$$[i, i + 1] := \{z \in \mathbb{R} : z = \mu \, x_{i+1} + (1 - \mu)x_i, \, 0 \leq \mu \leq 1\} \subset \mathbb{R}.$$

Let

$$\lambda_{\max} = \max_{s \in [0, N-1]} f_{\mathbf{X}}(s) = \max_{0 \leq i \leq N-1} \mathbf{X}(i),$$

and

$$\lambda_{\min} = \min_{s \in [0, N-1]} f_{\mathbf{X}}(s) = \min_{0 \leq i \leq N-1} \mathbf{X}(i).$$

For each $\lambda_{\min} \leq \lambda \leq \lambda_{\max}$, we introduce the following symbolic λ sub-level set for the map $f_{\mathbf{X}}$:

$$LS_\lambda(f_{\mathbf{X}}) := \{\sigma \in \mathcal{F} : f(\sigma) \leq \lambda\}$$

For $\lambda_{\min} \leq \lambda \leq \lambda' \leq \lambda_{\max}$, it holds

$$LS_\lambda(f_{\mathbf{X}}) \subset LS_{\lambda'}(f_{\mathbf{X}}).$$

Our next goal was to quantify the evolution of the above symbolic λ sub-level with. To this end, we introduce the notion of feature associated with the λ sub-level set $LS_\lambda(f_{\mathbf{X}})$.

We define the set of features for functions in $\Phi(\mathbb{R}^N)$ as

$$\mathfrak{F}(\Phi(\mathbb{R}^N)) := \{[i,j] \subset \mathbb{R} : 0 \leq i < j \leq N-1\}.$$

We note that $LS_\lambda(f_{\mathbf{X}}) \subset \mathcal{F} \subset \mathfrak{F}(\Phi(\mathbb{R}^N))$. Then next definition introduces the notion of features for a symbolic λ sub-level set as the interval of $\mathfrak{F}(\Phi(\mathbb{R}^N))$ constructed by a maximal union of faces of $LS_\lambda(f_{\mathbf{X}})$.

Definition 1. *We suggest that $\mathbb{I} \in \mathfrak{F}(\Phi(\mathbb{R}^N))$ is a feature for the symbolic λ sub-level set $LS_\lambda(f_{\mathbf{X}})$ if there exists $\mathbb{I}_1, \ldots, \mathbb{I}_k \in LS_\lambda(f_{\mathbf{X}})$ such that $\mathbb{I} = \bigcup_{j=1}^k \mathbb{I}_k$ and for every $\mathbb{J} \in LS_\lambda(f_{\mathbf{X}})$ such that $\mathbb{J} \neq \mathbb{I}_i$ for $1 \leq i \leq k$ it holds that $\mathbb{I} \cap \mathbb{J} = \emptyset$. We denote by $\mathfrak{F}(LS_\lambda(f_{\mathbf{X}}))$ the set of features for the λ sub-level set $LS_\lambda(f_{\mathbf{X}})$.*

A feature for a λ sub-level set $LS_\lambda(f_{\mathbf{X}})$ is the maximal interval of $\mathfrak{F}(\Phi(\mathbb{R}^N))$ that we can construct by unions of intervals in $LS_\lambda(f_{\mathbf{X}})$. To illustrate this definition, we give the following example:

Example 1. *Let us consider the time series*

$$\mathbf{X} = (11, 14, 9, 7, 9, 7, 8, 10, 9).$$

This allows us to construct the map $f_{\mathbf{X}}$ as shown in Figure 6. Then, $\lambda_{min} = 7$ and $\lambda_{\max} = 14$, and we have the following symbolic λ sub-level sets.

$$\begin{aligned}
LS_{\lambda=7}(f_{\mathbf{X}}) &= \emptyset \\
LS_{\lambda=8}(f_{\mathbf{X}}) &= LS_{\lambda=7}(f_{\mathbf{X}}) \cup \{[5,6]\} \\
LS_{\lambda=9}(f_{\mathbf{X}}) &= LS_{\lambda=8}(f_{\mathbf{X}}) \cup \{[2,3],[3,4],[4,5]\} \\
LS_{\lambda=10}(f_{\mathbf{X}}) &= LS_{\lambda=9}(f_{\mathbf{X}}) \cup \{[6,7],[7,8]\} \\
LS_{\lambda=11}(f_{\mathbf{X}}) &= LS_{\lambda=10}(f_{\mathbf{X}}) \\
LS_{\lambda=12}(f_{\mathbf{X}}) &= LS_{\lambda=11}(f_{\mathbf{X}}) \\
LS_{\lambda=13}(f_{\mathbf{X}}) &= LS_{\lambda=11}(f_{\mathbf{X}}) \\
LS_{\lambda=14}(f_{\mathbf{X}}) &= LS_{\lambda=11}(f_{\mathbf{X}}) \cup \{[0,1]\}.
\end{aligned}$$

This allows us to compute the available features for each λ-value:

	$\lambda = 7$	$\lambda = 8$	$\lambda = 9$	$\lambda = 10$	$\lambda = 11$	$\lambda = 12$	$\lambda = 13$	$\lambda = 14$
$\mathfrak{F}(LS_\lambda(f_{\mathbf{X}}))$	$\emptyset$	$[5,6]$	$[2,6]$	$[2,8]$	$[2,8]$	$[2,8]$	$[2,8]$	$[0,8]$

Let $\mathfrak{F}(f_{\mathbf{X}})$ be the whole set of features for $f_{\mathbf{X}}$, that is,

$$\mathfrak{F}(f_{\mathbf{X}}) = \{\mathbb{I} : \mathbb{I} \in \mathfrak{F}(LS_\lambda(f_{\mathbf{X}})) \text{ for some } \lambda_{\min} \leq \lambda \leq \lambda_{\max}\}.$$
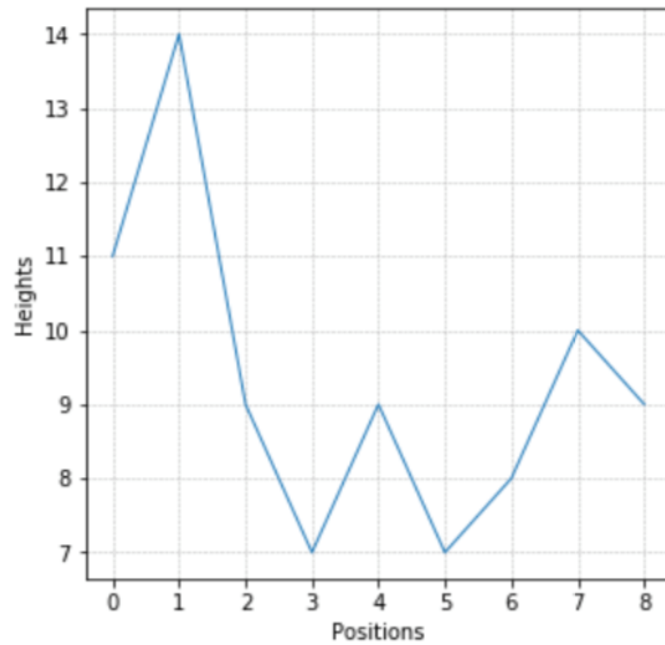
Figure 6. The map $f_{\mathbf{X}}$ for $\mathbf{X} = (11, 14, 9, 7, 9, 7, 8, 10, 9)$.

Example 2. *From Example 1, we obtain*

$$\mathfrak{F}(f_{\mathbf{X}}) = \{[5,6], [2,6], [2,8], [0,8]\}.$$

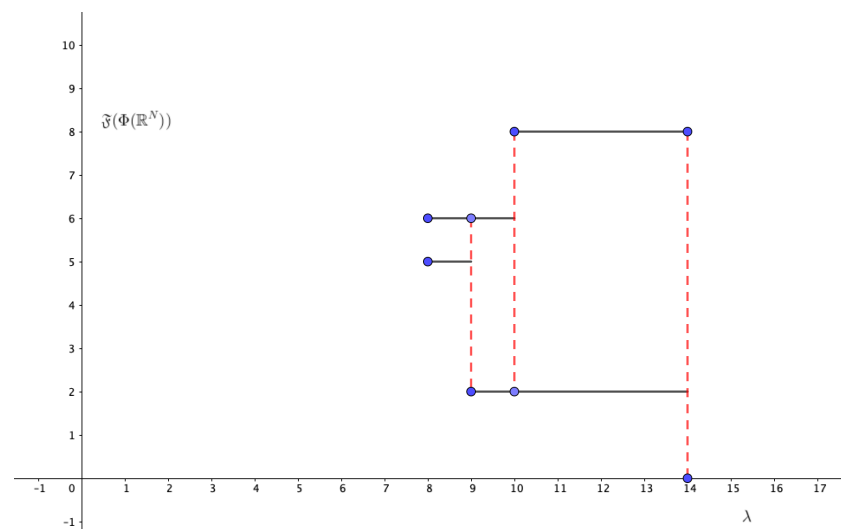*We can represent the map $\lambda \mapsto LS_\lambda(f_{\mathbf{X}})$ from $[\lambda_{\min}, \lambda_{\max}]$ to $\mathfrak{F}(f_{\mathbf{X}})$ as shown in Figure 7.*

Figure 7. The map $\lambda \mapsto LS_\lambda(f_{\mathbf{X}})$ for $\mathbf{X} = (11, 14, 9, 7, 9, 7, 8, 10, 9)$.

Let $\mathbb{I} \in \mathfrak{F}(f_{\mathbf{X}})$; in order to quantify the persistence of this particular feature for the map $f_{\mathbf{X}}$, we use the map $\lambda \mapsto LS_\lambda(f_{\mathbf{X}})$ from $[\lambda_{\min}, \lambda_{\max}]$ to $\mathfrak{F}(f_{\mathbf{X}})$. To this end, we introduce the following definition: the birth point of the feature $\mathbb{I}$ is defined by

$$a(\mathbb{I}) = \inf\{\lambda : \mathbb{I} \in \mathfrak{F}(LS_\lambda(f_{\mathbf{X}}))\}$$

and the corresponding death point by

$$b(\mathbb{I}) = \sup\{\lambda : \mathbb{I} \in \mathfrak{F}(LS_\lambda(f_{\mathbf{X}}))\}.$$

In particular, we note that $a([0, N - 1]) = \lambda_{\max}$ (see Figure 7). Since $a(\mathbb{I}) \leq b(\mathbb{I}) < \infty$ holds for all $\mathbb{I} \in \mathfrak{F}(f_{\mathbf{X}}), \mathbb{I} \neq [0, N - 1]$, we call the finite interval $[a(\mathbb{I}), b(\mathbb{I})]$ *the barcode of the feature* $\mathbb{I} \in \mathfrak{F}(f_{\mathbf{X}}) \setminus \{[0, N - 1]\}$.

Example 3. *From Example 1 we consider the features* $[5, 6] \in LS_{\lambda=8}(f_{\mathbf{X}})$, $[2, 6] \in LS_{\lambda=9}(f_{\mathbf{X}})$, *and* $[2, 8] \in LS_{\lambda=10}(f_{\mathbf{X}})$. *Then, the feature* $[5, 6]$ *has its birth point at* $a([5, 6]) = 8$ *and its death point at* $b([5, 6]) = 9$; *the feature* $[2, 6]$ *has its birth point at* $a([2, 6]) = 9$ *and its death point at* $b([2, 6]) = 10$. *Finally, the feature* $[2, 8]$ *has its birth point at* $a([2, 8]) = 10$ *and its death point at* $b([2, 8]) = 14$. *As a consequence, the set*

$$\mathcal{B}(f_{\mathbf{X}}) := \{([5, 6]; 8, 9), ([2, 6]; 9, 10), ([2, 8]; 10, 14)\}$$

of features and its corresponding barcodes contain the relevant information of the shape of $f_{\mathbf{X}}$ *(see Figure 7).*

Thus, we define the set of barcodes for $f_{\mathbf{X}}$ by

$$\mathcal{B}(f_{\mathbf{X}}) = \{(\mathbb{I}; a(\mathbb{I}), b(\mathbb{I})) : \mathbb{I} \in \mathfrak{F}(f_{\mathbf{X}}) \setminus \{[0, N - 1]\}\}$$

and its persistence diagram as

$$\mathcal{PD}(f_{\mathbf{X}}) := \left\{ (a(\mathbb{I}), b(\mathbb{I})) \in \mathbb{R}^2 : \mathbb{I} \in \mathfrak{F}(f_{\mathbf{X}}) \setminus \{[0, N - 1]\} \right\}$$

(see Figure 8). An equivalent representation of the persistence diagram is the life-time diagram for $f_{\mathbf{X}}$, which is constructed by means of a bijective transformation $T(a, b) = (a, b - a)$, acting over $\mathcal{PD}(f_{\mathbf{X}})$, that is,

$$\mathcal{LT}(f_{\mathbf{X}}) := \left\{ (a(\mathbb{I}), b(\mathbb{I}) - a(\mathbb{I})) \in \mathbb{R}^2 : \mathbb{I} \in \mathfrak{F}(f_{\mathbf{X}})) \setminus \{[0, N - 1]\} \right\};$$
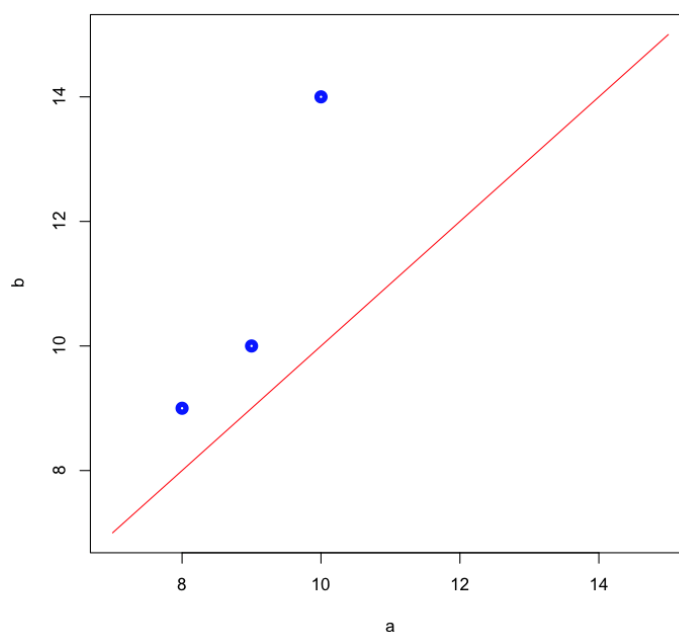
see Figure 9.

Figure 8. Persistence diagram for the map $f_{\mathbf{X}}$ when $\mathbf{X} = (11, 14, 9, 7, 9, 7, 8, 10, 9)$.
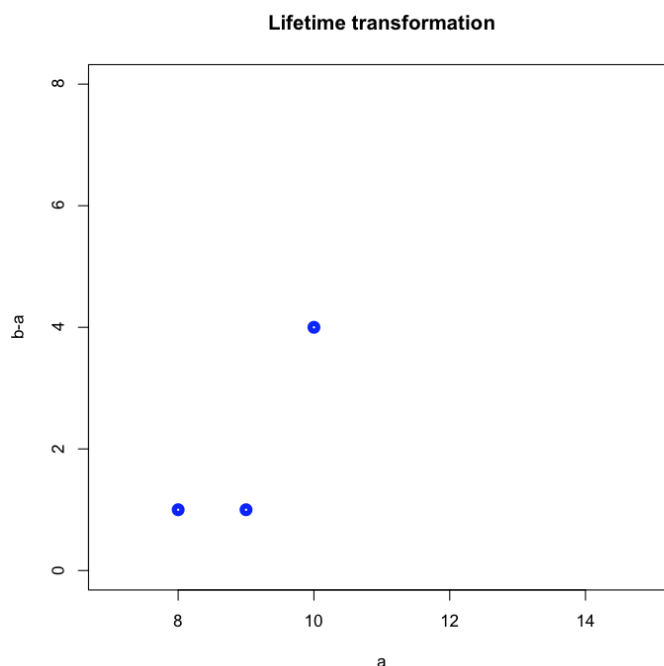
Figure 9. Life-time diagram for the map $f_\mathbf{X}$ when $\mathbf{X} = (11, 14, 9, 7, 9, 7, 8, 10, 9)$.

In order to determine the grade of similarity between two barcodes from two different time series, we need to set a similarity metric. To this end, we construct the persistent image for $f_\mathbf{X}$ as follows: we observe that $\mathcal{LT}(f_\mathbf{X})$ is a finite set of points, namely,

$$\mathcal{LT}(f_\mathbf{X}) = \{(a_1, b_1 - a_1), \ldots, (a_k, b_k - a_k)\}$$

for some natural numbers $k \geq 1$ and such that $b_1 - a_1 \leq b_2 - a_2 \ldots \leq b_k - a_k$. Then, we consider a non-negative weighting function $w : \mathcal{LT}(f_\mathbf{X}) \longrightarrow [0, 1]$ given by

$$w(a_i, b_i - a_i) = \frac{b_i - a_i}{b_k - a_k} \text{ for } 1 \leq i \leq k.$$

Finally, we fix M, a natural number, and take a bivariate normal distribution $g_u(x, y)$ centered at each point $u \in \mathcal{LT}(f_\mathbf{X})$ and with its variance $\sigma \, id = \frac{1}{M} \max_{1 \leq i \leq k}(b_i - a_i) \, id$, where id is the 2×2 identity matrix. A persistence kernel is then defined by means of a function $\rho_\mathbf{X} : \mathbb{R}^2 \to \mathbb{R}$, where

$$\rho_\mathbf{X}(x, y) = \sum_{u \in \mathcal{LT}(f_\mathbf{X})} w(u) g_u(x, y). \tag{1}$$

We associate with each $\mathbf{X} \in \mathbb{R}$ a matrix in $\mathbb{R}^{M \times M}$ as follows: let $\varepsilon > 0$ be a non-negative real number that is sufficiently small, and then consider a square region $\Omega_{\mathbf{X}, \varepsilon} = [\alpha, \beta] \times [\alpha^*, \beta^*] \subset \mathbb{R}^2$, covering the support of $\rho_\mathbf{X}(x, y)$ (up to a certain precision), such that

$$\iint_{\Omega_{\mathbf{X}, \varepsilon}} \rho_\mathbf{X}(x, y) \, dx \, dy \geq 1 - \varepsilon$$

holds. Next, we consider two equispaced partitions of the intervals

$$\alpha = p_0 \leq p_1 \ldots \leq p_M = \beta \text{ and } \alpha^* = p_0^* \leq p_1^* \ldots \leq p_M^* = \beta^*.$$

Now, we put

$$\Omega_{\mathbf{X},\varepsilon} = \bigcup_{i=0}^{M-1} \bigcup_{j=0}^{M-1} [p_i, p_{i+1}] \times [p_j^*, p_{j+1}^*] = \bigcup_{i=0}^{M-1} \bigcup_{j=0}^{M-1} P_{i,j}$$

The persistence image of $\mathbf{X}$ associated with the partition $\mathcal{P} = \{P_{i,j}\}$ is then described by the matrix given by the following equation:

$$PI(\mathbf{X}, M, \mathcal{P}, \varepsilon) = \left(\iint_{P_{i,j}} \rho_{\mathbf{X}}(x,y)dxdy \right)_{i=0,j=0}^{i=M-1,j=M-1} \in \mathbb{R}^{M \times M}. \tag{2}$$

2.5. Classification

Image classification is a procedure that is used to automatically categorize images into classes by assigning to each image a label representative of its class. A supervised classification algorithm requires a training sample for each class, that is, a collection of data points whose class of interest is known. Labels are assigned to each class of interest. The classification problem applied to a new observation is thus based on how close a new point is to each training sample. The Euclidean distance is the most common distance metric used in low-dimensional datasets. The training samples are representative of the known classes of interest to the analyst. In order to classify the persistence diagrams, we can use any state-of-the-art technique. In our case, we considered the random forest classification.

Recall that we conducted 9 different experiments, with 24 samples associated with each one of them corresponding to 3 samples for each of the different experimental conditions: relaxed rigid driver, relaxed rigid passenger, relaxed SAV driver, relaxed SAV passenger, tense rigid driver, tense rigid passenger, tense SAV driver, and tense SAV passenger. Their respective labels are $\{0, 0, 0, 0, 1, 1, 1, 1\}$. Therefore, we designed the following training validation process: The model is trained over 144 samples and evaluated over the remaining unseen 72 experiments (two-to-one training-to-testing ratio). The split between training and sampling is achieved using random shuffling and stratification to ensure balance between the classes. In order to improve the evaluation of the model generalizability, we also performed a cross-validation procedure following a *leave-one-out* strategy, consisting of iteratively training over the full dataset except one sample that was left out and used to test and score the model. We used the *accuracy* metric to evaluate the classification model. We can represent the performance of the model using the so-called confusion matrix: a 2D entries table where elements account for the number of samples in each category, with the first axis representing the true labels and the second axis the predicted labels. We also computed the different classification metrics to obtain a more detailed reporting of the model performances.

3. Results

The trained random forest classifier model for the persistence images has a notably high accuracy score on the training dataset (144) for both approaches and high accuracy for the testing dataset (72 samples). This suggests strong differentiation of the images with the respect to their generating signals, see Figure 10. The scores on the training and testing are 93 and 83%, respectively. The leave-one-out cross-validation achieved a score of 81%, indicating a good variance–bias trade-off and good generalization potential of the model.
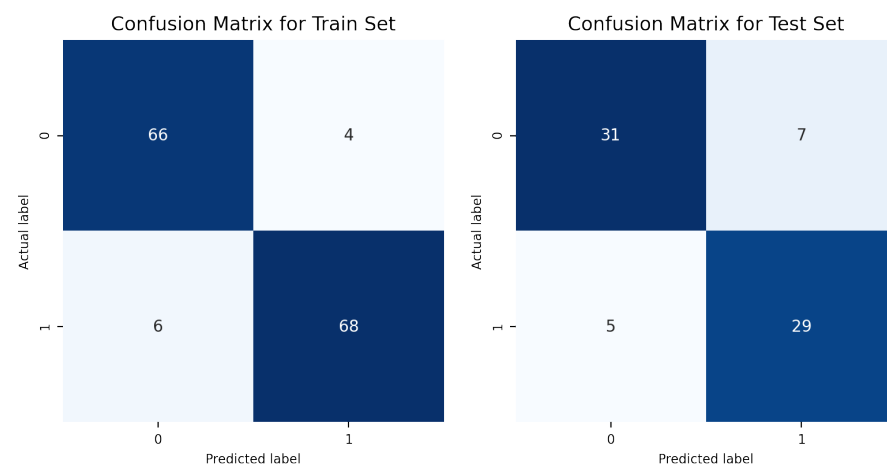
Figure 10. Model performance for prediciting the attention state.

4. Discussion

The combination of Morse theory and topological data analysis allows us to extract information from real data without the need for smoothness or regularity assumption on the time series. In our case, input data for each experiment were reduced from six-sensor time series of measurements to one single image containing the persistent pattern for attention to the road. Using the obtained persistence images as the new inputs, supervised learning proved to successfully predict the attention state of the driver or passenger.

The procedure used and described in this paper does not involve any additional pre-processing of the sensor data; is robust to noise and degraded signals; and supports large quantities of data, which makes it efficient and scalable.

It is important to highlight the fact that while the proposed methodology based on the TDA (successfully applied in large datasets [9]) seems general and powerful and it was able to extract the main data features, the validity of the driver behaviors observed in the analyzed dataset should be carefully checked due to the overly reduced dataset employed (limited to nine individuals) that does not allow for the full validation of prediction robustness.

Institutional Review Board Statement: Ethical review and approval were waived for this study due to the research presents no risk of harm to subjects and only collects non-personalized anonymized data.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Data are available under request.

Appendix A

We can illustrate the process of computing the persistence diagrams, the lifetime diagrams, and the persistence images for the driver time series for each experimental setup:

1. Relaxed driver with SAV seat;
2. Relaxed driver with rigid seat;
3. Relaxed passenger with SAV seat;
4. Relaxed passenger with rigid seat;
5. Tense driver with SAV seat;
6. Tense driver with rigid seat;
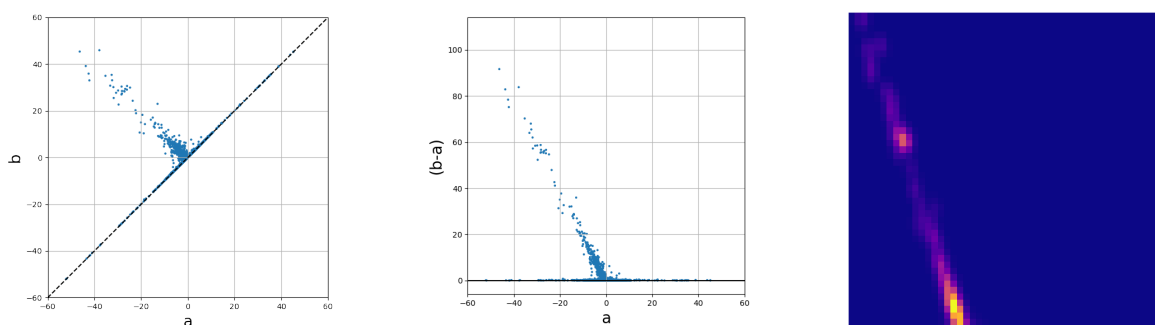7. Tense passenger with SAV seat;
8. Tense passenger with rigid seat.

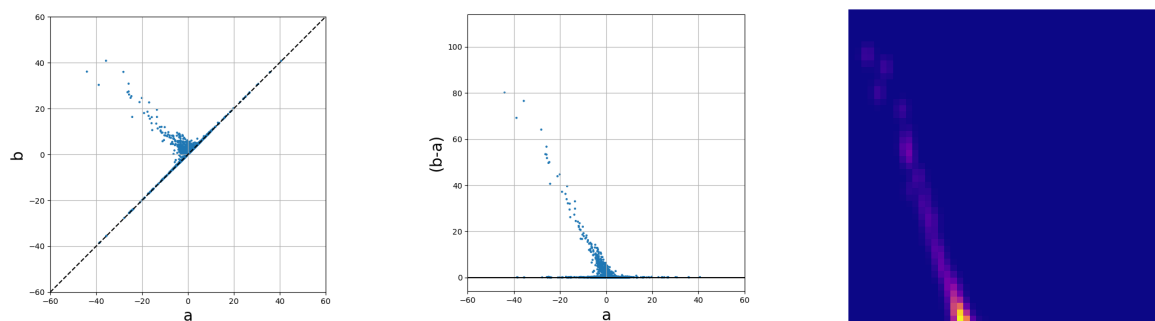**Figure A1.** Relaxed driver with SAV seat.
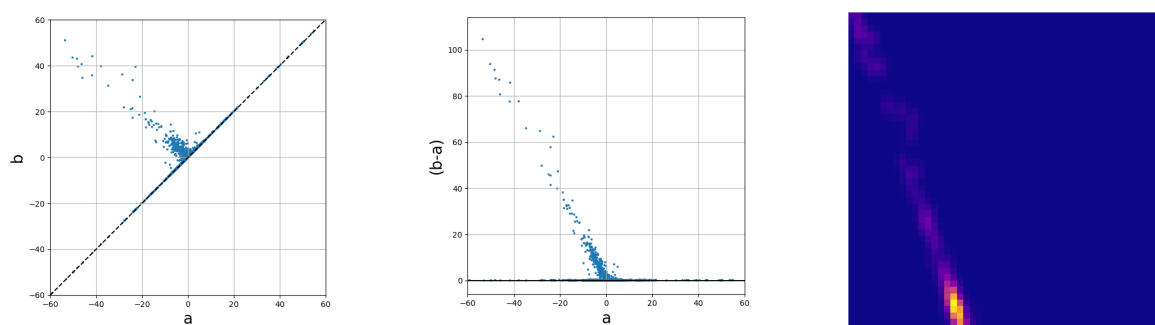
Figure A2. Relaxed driver with rigid seat.

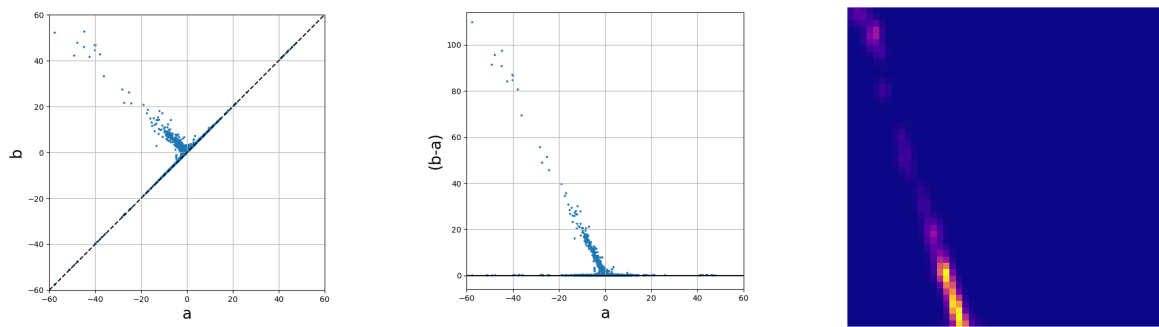**Figure A3.** Relaxed passenger with SAV seat.

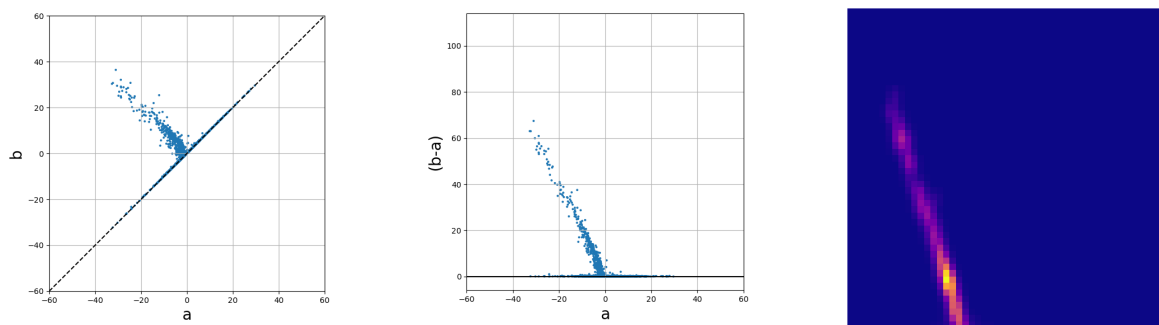**Figure A4.** Relaxed passenger with rigid seat.
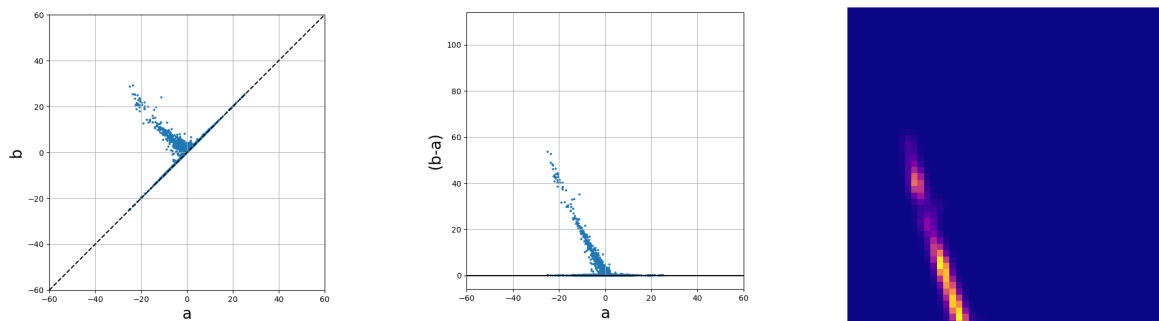

Figure A5. Tense driver with SAV seat.
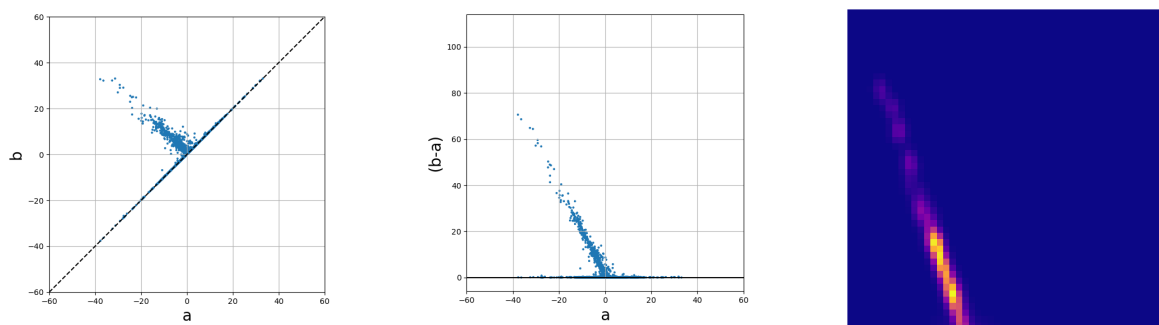

Figure A6. Tense driver with rigid seat.

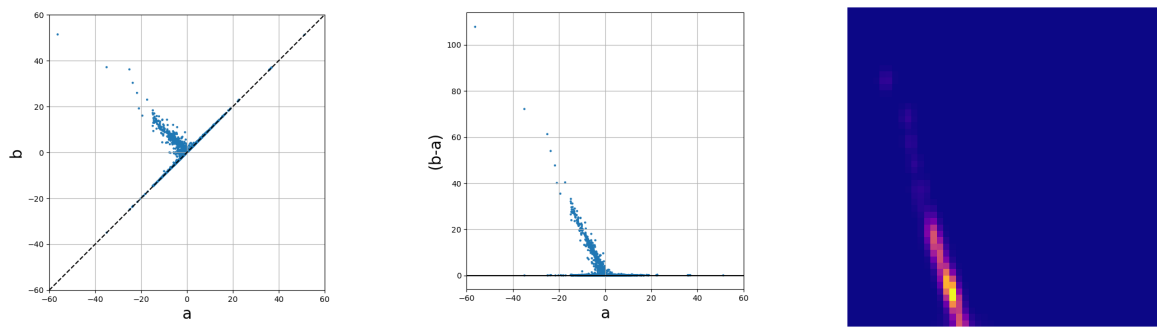**Figure A7.** Tense passenger with SAV seat.

Figure A8. Tense passenger with rigid seat.

Appendix B

To better evaluate a classification model, we are interested in quantities that express how often a sample is correctly or wrongly labelled into a particular class over all the samples and all the classes:

- A *True positive* (TP): the correct prediction of a sample into a class;
- A *True negative* (TN): the correct prediction of a sample out of a class;
- A *False positive* (FP): the incorrect prediction of a sample into a class;
- A *False negative* (FN): the incorrect prediction of a sample out of class.

Therefore, we can examine in more detail the classification model performance using the following metrics:

- The *precision* P is the number of correct positive results divided by the number of all positive results.

$$P = \frac{TP}{TP + FP} \tag{A1}$$

- The *recall* R is the number of correct positive results divided by the number of all relevant samples.

$$R = \frac{TP}{TP + FN} \tag{A2}$$

- The F-1 score is the harmonic mean of precision and recall.

$$F1 = 2 \times \frac{P \times R}{P + R} \tag{A3}$$

- The *accuracy* A is the number of correct predictions over the number of all samples.

$$A = \frac{TP + TN}{TP + TN + FP + FN} \tag{A4}$$

We can summarize the presented metrics for our model in the following two reports:

	precision	recall	f1-score	support
0	0.94	0.92	0.93	72
1	0.92	0.94	0.93	72
accuracy			0.93	144
macro avg	0.93	0.93	0.93	144
weighted avg	0.93	0.93	0.93	144

	precision	recall	f1-score	support
0	0.82	0.86	0.84	36
1	0.85	0.81	0.83	36
accuracy			0.83	72
macro avg	0.83	0.83	0.83	72
weighted avg	0.83	0.83	0.83	72

(**a**) Training set. (**b**) Testing set.

Figure A9. Classification report.

References

1. Paul, A.; Chauhan, R.; Srivastava, R.; Baruah, M. *Advanced Driver Assistance Systems (No. 2016-28-0223)*; SAE Technical Paper; SAE International: Warrendale, PA, USA, 2016.
2. Shaout, A.; Colella, D.; Awad, S.S. Advanced Driver Assistance Systems-Past, Present and Future. In Proceedings of the 2011 Seventh International Computer Engineering Conference (ICENCO'2011), Cairo, Egypt, 27–28 December 2011; pp. 72–82.
3. Geronimo, D.; Lopez, A.M.; Sappa, A.D.; Graf, T. Survey of pedestrian detection for advanced driver assistance systems. *IEEE Trans. Pattern Anal. Mach. Intell.* **2009**, *32*, 1239–1258. [CrossRef]
4. Lindgren, A.; Chen, F. State of the art analysis: An overview of advanced driver assistance systems (adas) and possible human factors issues. *Hum. Factors Econ. Asp. Saf.* **2006**, *38*, 50.
5. Izquierdo-Reyes, J.; Ramirez-Mendoza, R.A.; Bustamante-Bello, M.R. A study of the effects of advanced driver assistance systems alerts on driver performance. *Int. J. Interact. Des. Manuf.* **201**8, *12*, 263–272. [CrossRef]
6. Sikander, G.; Anwar, S. Driver fatigue detection systems: A review. *IEEE Trans. Intell. Transp. Syst.* **2018**, *20*, 2339–2352. [CrossRef]
7. Milnor, J. *Morse Theory*; Princeton University Press: Princeton, NJ, USA, 1963.
8. Epstein, C.; Carlsson, G.; Edelsbrunner, H. Topological data analysis. *Inverse Probl.* **2011**, *27*, 120201.
9. Frahi, T.; Argerich, C.; Youn, M.; Chinesta, F.; Falco, A. Tape surfaces characterization with persistence images. *AIMS Mater. Sci.* **2020**, *7*, 364–380. [CrossRef]
10. Milnor, J.W.; Thurston, W. On iterated maps of the interval, Dynamical systems (College Park, MD, 1986–87). In *Lecture Notes in Mathematics*; Springer: Berlin/Heidelberg, Germany, 1988; pp. 465–563.
11. Reeb, G. Sur les points singuliers d'une forme de Pfaff complètement intégrable ou d'une fonction numérique. *C. R. Acad. Sci. Paris* **1946**, *222*, 847–849.