Segment-based CO$_2$ Emission Evaluations from Passenger Cars based on Deep Learning Techniques

Naghmeh Niroomand$^1$, Christian Bach$^1$, and Miriam Elser$^1$

$^1$Automotive Powertrain Technologies Laboratory, Swiss Federal Laboratories for Materials Science and Technology, Dübendorf, CH-8600, Switzerland

Corresponding author: Naghmeh Niroomand (email: naghmeh.niroomand@empa.ch).

ABSTRACT The overall level of emissions from the Swiss passenger cars is strongly dependent on the fleet composition. Despite technology improvements, the Swiss passenger cars fleet remains emissions intensive. To analyze the root of this problem and evaluate potential solutions, this paper applies deep learning techniques to evaluate the inter-class (namely micro, small, middle, upper middle, large and luxury class) and intra-class (namely sport utility vehicle and non-sport utility vehicle) differences in carbon dioxide (CO$_2$) emissions. This paper takes full use of novel semi-supervised fuzzy C-means (SSFCM), random forest and AdaBoost models as well as model fusion to successfully classify passenger vehicles and enable segment-based CO$_2$ emission evaluations.

INDEX TERMS CO$_2$ emissions, Feature learning, Semi-supervised deep learning, Vehicle classification. 

I. INTRODUCTION

More than 5 years after the adoption of the Paris agreement, which aims to limit global warming to below 2 °C (preferably 1.5 °C), global greenhouse gas emissions continue growing steadily [1]. According to the 2016 EU Reference Scenario, without an ambitious commitment towards decarbonization, transport related carbon dioxide (CO$_2$) emissions are expected to decrease only by 8% between 2010 and 2050 and will reach their largest share by the end of the projection period (2050) [2-3]. Underlying this limited decrease are a significant increase in the number of passenger cars, the slow market penetration of electric cars and a limited shift towards alternative fuels.

Switzerland is responsible for less than 0.2% of global man-made fossil CO$_2$ emissions [4]. However, the transport sector represents the largest consumer of fossil fuels in Switzerland and caused about 32% of Switzerland’s CO$_2$ emissions in 2019 (i.e., around 15 million tonnes CO$_2$ eq., excluding international aviation and shipping). Road transport was responsible for 98% of these emissions, with only small contributions from national shipping (0.8%), aviation (0.8%) and rail transport (0.2%). Among the different forms of road transport, passenger cars accounted for almost two thirds of the total emissions (73%), followed by freight transport (21%), buses (3%) and motorcycles (2%) [5]. Therefore, in order to meet the CO$_2$ reduction targets of Switzerland [6], fossil energy carriers in the mobility sector have to be substituted by ones based on renewable energies and the overall energy consumption has to be reduced. Although substituting fossil energy by renewable ones is essential to meet the CO$_2$ reduction targets, decisions about investments and new policies are not moving fast enough to decarbonize the economy in compliance with the Paris agreement.

On the other hand, during the last decades there have been large technical and dimensional changes in new passenger vehicles, mostly related to technology improvements and intra-class variations. Particularly relevant are the changes in the dimensions of the vehicle segments (i.e., increased size of the vehicles in most segments), within single vehicle segments (i.e., increased share of SUVs), and other design parameters like increased efficiency of the engine and engine displacement down-sizing. Understanding the impact of these changes on the fuel consumption and CO$_2$ emissions is crucial to develop successful strategies to decarbonize the road transport.

Since the division of vehicles into segments by experts is not standardized and therefore not always uniform, and some vehicle models have recently positioned themselves as "crossovers" between established vehicle categories [7-8], it has become increasingly difficult and inaccurate to segment
the vehicle population using conventional classification methods. Using mathematical approaches, vehicles can be uniformly divided into segments based on similarity features. The development of a mathematical approach to accurately segment passenger vehicles is essential for determining the real \( \text{CO}_2 \) emissions from road traffic in the future. While road traffic has so far had its own energy system, which was comparatively easy to assess in terms of \( \text{CO}_2 \) emissions, increasing electrification of road traffic will difficult the distinction of energy consumption from road traffic and other stationary energy uses. Moreover, the estimated overall impact of the introduction of the world harmonized light-duty vehicles test procedure (WLTP) on average \( \text{CO}_2 \) emissions is in the order of 15-25%, which would lead to on average 18-30 g/km higher \( \text{CO}_2 \) emissions for the new passenger cars [9-12] (Fig. 1). Moreover, due to the limited informative value of the \( \text{CO}_2 \) type approval values on the real \( \text{CO}_2 \) emissions, the wide margin of uncertainty regarding vehicle classification and the type approval extension based on the new definitions in the test protocol [11], this segmentation is an important step on the way to a new \( \text{CO}_2 \) assessment of road traffic.

In this study, by segmenting the passenger vehicles based on technical and dimensional characteristics, we aim to better understand the impact of inter-class (between classes of a multi-class) and intra class (within each class) variations to the passenger vehicle fleet \( \text{CO}_2 \) footprint [14]. In our approach, several semi-supervised clustering algorithms are compared and used to predict labels from unsupervised clustering algorithms based on a feature learning technique, which is a highly useful method for representation learning with high-dimensional datasets containing high-level uncertainties [15-24]. This paper is an extension of a previous work originally focused on developing a machine learning based methodology for the mathematical inter-class and intra-class segmentation of passenger vehicles. Here we improved the classification performance of this method by adding emission and technical features as an input. Based on this novel approach, we can then predict accurate segment-based \( \text{CO}_2 \) emissions, which allows for detailed analyses of the main factors influencing the average fleet \( \text{CO}_2 \) emissions. Our results show that the proposed method is a viable and effective to categorize vehicles based on their technical, emission and dimensional features.

Section II briefly introduces the Swiss transportation system. Section III presents the related research. Section IV describes the methods. Section V provides concise details on the used datasets, the algorithms, the performed experiments and the discussion of the results and last, section VI provides the majors findings of our work and recommendations for further research.

II. SWISS TRANSPORT SECTOR AND \( \text{CO}_2 \) EMISSIONS

In terms of mobility, Switzerland can be divided in three main regions, namely urban, suburban, and rural areas. There are major differences in the sustainability challenges posed within these regions due to the urbanization. Fig. 2 illustrates that the growth of the number of cars has placed additional pressures on traffic congestion and parking spaces, particularly in higher density areas. This creates opportunities for offering alternatives to cover the existing transportation needs, including public transit network and shared mobility services. In contrast, in rural areas, which represent about one third of the total Swiss population, due to the lack of attractive and feasible transport alternatives, private automobile remains the most common form of transportation. In addition, despite the high rate of the population accepting public transport modes in Switzerland (59%), two thirds of the total passenger kilometers are still completed by car [25].

Over 6.5 million motor vehicles were registered in 2018 in Switzerland, more than 4.6 million of which were passenger cars. Around one third of all passenger cars were more than ten years old and over 1.6 million cars were completely outdated. Only 0.4% of all passenger cars had purely electric propulsion systems. Among the new registrations, petrol was the most common fuel (68%), while almost 30% of the sold vehicles had a diesel engine [28]. The mean type approval \( \text{CO}_2 \) emissions of newly registered cars continuously dropped from around 190 g CO\(_2\)/km in 2003 to around 134 g CO\(_2\)/km in 2016. After this steady decline, the mean \( \text{CO}_2 \) emissions of the new registrations rose again to 137.8 g CO\(_2\)/km in 2018. As a result, the specified target value of 130 g CO\(_2\)/km that came into force in 2012 was not entirely met. Fig. 3 shows that new cars in the Southeast of the country are generally less fuel efficient and produce more \( \text{CO}_2 \) emissions compared to the new cars in the Northwest.

It is forecasted that the share of electric vehicles in the Swiss passenger car fleet will increase from 1.5% in 2021 to 38-74% in 2050, depending on the considered scenario [29]. In addition to the considerable savings in terms of fossil fuel consumption, the increasing share of electric vehicles will drastically reduce the \( \text{CO}_2 \) emissions compared to vehicles powered by fossil fuels [30].

III. RELATED WORK

Over the last decades, as a result of new European \( \text{CO}_2 \) regulations, car manufacturers have used a whole range of technical and dimensional solutions to meet specific annual \( \text{CO}_2 \) emission targets. The review of related literature shows that the large changes in the passenger car models over time poses an additional challenge to the accurate vehicle classification [31-34].

In spite of the partial achievement of the targets based on the type approval \( \text{CO}_2 \) emissions (laboratory tests), real-world \( \text{CO}_2 \) emissions have decreased by only about 10% [35]. Subsequently, the gap between the calculated and real-world \( \text{CO}_2 \) emissions has widened from 9% to 42%, resulting in 31 g CO\(_2\)/km of fake emissions savings [36-37]. This gap varies considerably across countries due to the significant variation within vehicle classes. However, the lack of a standard classification method, hinders the comparison of these results between different countries [38-39].
In order to close the gap between the CO₂ emissions results estimated by two major techniques (top-down approaches focusing on fuel market interactions and bottom-up approaches focusing on technological details), researchers have developed multiple simulation programs, such as greenhouse gas emission models and vehicle energy calculation tools, for the compilation of emission inventories [40-45]. From this point of view, the simulation is useful to compensate the limitations of the laboratory test methods. For example, Seo et al. [46] developed a vehicle type classification simulation method using a bottom-up approach to calculate national CO₂ emissions. This study concluded that CO₂ emissions of medium and heavy-duty vehicles (MHDV) represented 25.5% of the total on-road emissions, although only 4.2% of all vehicles were MHDV. Jimenez et al. [47] reviewed the influence of vehicle classification, vehicle characteristics, vehicle brand and registration year on the real-world CO₂ emissions. They employed a database of 650 passenger cars. This study explained the impact of these factors on the gap between real-world and type-approval emission values. Ntziachristos et al. [48] reported that the deviations in fuel consumption are directly reflected in CO₂ emissions. This study computed the observed 11% gap for in-use petrol and 16% for in-use diesel with the type-approval procedure by controlling engine capacity, vehicle mass and power. They used a database of 924 passenger cars from Europe. The results indicated that the large vehicle class has the highest deviation in test score.

All these studies show that simulation techniques are capable to overcome some of the limitations faced with fuel-based approach in terms of estimating the CO₂ emissions of each vehicle class. However, the simulation techniques cannot consider intra-class variations in CO₂ emissions, they are difficult to use when conducting a detailed analysis and they require expert knowledge.

Lately, feature learning techniques have shown an outstanding performance for addressing uncertainty problems for clustering and classification [19, 49-55]. The classification performance highly depends on a quality of features generated from the data as input to the classifier process. However, only a limited number of studies have combined feature learning techniques to improve the classification performance on a high dimensional dataset and predict the vehicle CO₂ emissions. For more details about the feature learning techniques, refer to the article by He et al. [56], in which the authors implemented feature learning classification to analyze vehicular emissions.

In particular, they applied decision tree, random forest, AdaBoost, and XgBoost models based on the fuel type and registration date. This study achieved a prediction accuracy of 70% by artificially controlling the registration date for different users. Saleh et al. [57] used deep learning with a support vector machine (SVM) model to predict CO₂ emissions by monitoring energy consumption. The low value of Root Mean Square Error of the model indicates the high accuracy of the prediction. Ghahramani et al. [58] proposed an unsupervised learning approach to estimate CO₂ emissions from road transport with a focus on taxi trips. This study identified the most polluting trips and the vehicles associated with these trips in order to replace them with alternative alternatives powertrains, such as electric vehicles.

The classification method proposed in this paper is a new semi-supervised clustering scheme (SSFCM) that incorporates semi-supervised information in fuzzy C-means (FCM) algorithm to considerably improve its effectiveness [59-63]. In this field, Jiang et al. [64] combined several feature extraction methods with a support vector machine classifier to group the vehicles in six categories "large bus", "passenger car", "motorcycle", "minibus", "truck" and "van". This study achieved a classification accuracy of 97.4%. Balid et al. [65] implemented deep learning-based classification using the vehicle length as a key feature. Their method classifies vehicles into passenger vehicles, single unit trucks, combination trucks, and multi-trailer trucks and achieved a classification accuracy of 97%. Maungmai and Nuthong [66] used a convolutional neural network method to classify the vehicles type as "small", "medium", "large", and "unknown", and vehicle color as "black", "blue", "white", "green", "yellow", "red", and "unknown". The results comparison shows that, using decision trees, random forest, and densely deep neural network classifier, the classification accuracy of vehicle type and vehicle color increased by 1.8% and 0.8%, respectively. Dong et al. [67] proposed a vehicle type classification method using a semi-supervised convolutional neural network using high-resolution vehicle frontal view images. The algorithm achieved 88.1% accuracy.

### IV. MATERIALS AND METHODS

#### A. SEMI-SUPERVISED CLUSTERING

Semi-supervised clustering aims to boost the accuracy of the defined clusters by identifying better clusters than the ones obtained from the unsupervised learning algorithm [19, 68-75]. Typically, semi-supervised clustering methods result in a worse representation of the results in the original feature space. To make the semi-supervised clustering more efficient, it is reasonable to combine semi-supervised clustering with deep feature learning [76-79]. The framework of the proposed clustering approach is depicted in Fig. 4.

Unlike the most widely used approaches in semi-supervised clustering based on the feature extraction technique, we consider three types of information (diffusion labels, extracted core data, and extracted feature vectors) in order to improve the classification accuracy and mitigate the class imbalance and multi-class overlapping problems. This framework includes three main layers. First, labeled data is divided into train set and test set in order to build a classifier and evaluate its output, respectively. Then, recordings from the train set along with the unlabeled data are used as input to the feature learning process. The output of the feature
learning step are the cluster centroids that are used to project data from train and test sets into a new learnt space and extract feature vectors in the feature extraction step. In the classification step, AdaBoost [80], Random Forest [81], and SSFCM models are built on the vectors of the train set and then used to predict the labels for the feature vectors of the test set. Finally, the performance parameters of the three single models are compared to the model fusion ones to evaluate their performances in terms of data classification and prediction.

B. SEMI-SUPERVISED FUZZY C-MEAN CLUSTERING

Fuzzy C-means (FCM), as an overlapping clustering algorithm, is one of the most popular fuzzy clustering methods [82]. FCM is a soft clustering algorithm, meaning that each data point has a probability of belonging to each cluster with partial membership values ranged from 0 to 1. However, due to the non-convexity of its objective function, it may fall into a local optimal solution during optimization. To address this issue, we propose a semi supervised fuzzy C-means clustering (SSFCM) that incorporates deep feature learning in FCM to further improve its effectiveness and eliminate redundant information [83-85]. This method aims to minimize the objective function (J) as follows:

\[
\text{Min } J(X; U, V) = \sum_{k=1}^{N} \sum_{i=1}^{C} u_{ki} m \| X_k - v_i \|_A^2 \quad (1 \leq m < \infty) \tag{1}
\]

s. t. \( \sum_{i=1}^{C} u_{ki} = 1 \) \( (0 \leq u_{ki} \leq 1) \) \( \tag{2} \)

\[
u_i = \frac{\sum_{k=1}^{N} u_{ki} x_k}{\sum_{k=1}^{N} u_{ki}} \tag{3}
\]

\[
u_{ki} = \frac{1}{\sum_{j=1}^{C} (p_{kj}A)^{(m-1)/m}} \tag{4}
\]

\[
D_k^2 = \| X_k - v_i \|_A^2 = (X_k - v_i)^T A (X_k - v_i) \tag{5}
\]

Where \( N \) is number of data elements, \( C \) is the number of clusters; \( X_i \) represents the data k of \( X = \{X_1, X_2, X_3, ..., X_N\} \) in the \( p \)th cluster; \( u_{ki} \) represents the weighted squared errors function known as membership function; \( m \) is a weighting exponent that determines the degree of fuzziness and that was set to 2 in order to ensure high membership values for each data point to its closest cluster; \( A \) is a positive and symmetric (n \times n) weight matrix; \( U \) is the fuzzy partition matrix of the dataset \( X \) into c cluster; \( v_i \) is vectors of center in \( i \)th cluster; \( K \) denotes the features, and \( \| X_k - v_i \|_A^2 \) denotes to the Euclidean distance function and it is computed in the \( A \) norm between \( j \)th data and \( i \)th cluster center.

The SSFCM methodology is composed of the following four steps. First, with algorithm 1 we find the FCM memberships and centroids:

**Algorithm 1 Fuzzy C-means Membership and Centroid**

**Input:** N data elements \( X = \{X_1, X_2, ..., X_N\} \), weight matrix \( A \), number of clusters \( C \), degree of fuzziness \( (m=2) \), max iteration number \( (T) \), error threshold \( (\varepsilon) \)

**Output:** \( u_{ki}, v_i \)

1. Set \( t = 0 \)
2. Initialize \( v_i \)
3. Update \( t = t + 1 \)
4. Compute \( u_{ki} \)
5. Compute \( v_i \)
6. If \( t > T \) or \( \| u_k - u_{k-1} \| < \varepsilon \) then stop; otherwise
7. Repeat from step 3.

Next, algorithm 2 is used to calculate deep FCM memberships and centroids:

**Algorithm 2 The Training Strategies for Deep Fuzzy C-means**

**Input:** N data elements \( X = \{X_1, X_2, ..., X_N\} \), number of clusters \( C \), clusters feature \( K \), labeled dataset \( L \), unlabeled dataset \( UN \), membership degree \( U \), max iteration number \( (T) \), error threshold \( (\varepsilon) \)

**Output:** \( u_{kiL}, u_{kiUNL}, v_k^L, v_k^UNL \)

1. Initialize \( v_k^L \) (random for labeled data)
2. Update \( t = t + 1 \)
3. Compute \( u_{kiL}, u_{kiUNL} \)
4. Compute \( v_k^{L+1}, v_k^{UNL+1} \)
5. If a stopping criterion, \( t > T \) or \( \| J_t - J_{t-1} \| < \varepsilon \), is fulfilled for all labeled and unlabeled objective functions separately then stop; otherwise
6. Repeat from step 3.

Then, using algorithm 3 we select the features \( s \subseteq K \) by using the random oversampling (ROS) technique. The purpose of the ROS approach is to maintain a balance between the feature subsets of labeled classes and unlabeled data elements [86-87].

**Algorithm 3 Feature Extraction of Deep Fuzzy C-means**
To assess the performance of the different algorithms, we compute the confusion matrix and use it to determine the

C. STATE-OF-THE-ART METHODS

Two ensemble learning methods, Random Forest and AdaBoost, are used to enhance the accuracy and performance of the classification [89-90]. The Random Forest model is a parallel learning process that uses a bagging technique for the data training [91]. This data sampling technique aims to reduce the variance and bias in the model by generating multi-sets (multiple decision trees) for training from the original data. In the parallel process none of these decision trees is dependent on other trees.

Algorithm 6 AdaBoost Classifier

Input: Data X whose number of elements N, training set (S), decision tree in forest (B), subsample size (µ), max iteration number (T)
Output: Set Q = ∅
1. for t ∈ {1, . . . , T} do
2. Initialize data weights {Dn} to 1/N
3. find best weak classifier ym(x) by minimizing weighted error function Jn:
   \[ J_m = \sum_{n=1}^{N} D_n^{(m)} 1[y_m(x_n) ≠ t_n] \]
4. Compute \(err_m = \sum_{n=1}^{N} D_n^{(m)} 1[y_m(x_n) ≠ t_n]/\sum_{n=1}^{N} D_n^{(m)} \)
5. assign weight \( \alpha_m = \log (\frac{1-err_m}{err_m}) \) to classifier \( y_m(x) \)
6. update the data weights: \( D_n^{(m+1)} = D_n^{(m)} \exp[\alpha_m 1[y_m(x_n) ≠ t_n]] \)
7. Normalize \( D_n^{(m+1)} \) to be proper distribution
8. Make prediction using the final model: \( Y_m(x) = sign (\sum_{m=1}^{M} \alpha_m y_m(x)) \)

D. Performance Measure

To assess the performance of the different algorithms, we compute the confusion matrix and use it to determine the
precision ($P_i$), recall ($R_i$), F-Measure and adjusted rand index (ARI) as given in the following:

\[
P_i = \frac{TP_i}{TP_i + FP_i} \quad (1 \leq i \leq c)
\]

\[
R_i = \frac{TP_i}{TP_i + FN_i} \quad (1 \leq i \leq c)
\]

\[
F - \text{Measure} = \frac{2P_iR_i}{P_i + R_i}
\]

\[
RI = \frac{TP+TN}{TP+TN+FP+FN} \quad (0 \leq RI \leq 1)
\]

\[
ARI = \frac{RI-E[RI]}{max(RI)-E[RI]} \quad (-1 \leq ARI \leq 1)
\]

Here, $TP_i$ (True Positives) is the proportion of data points classified correctly to each class $i$; $FN_i$ (False Negatives) is the proportion of data points that are not classified to class $i$ but actually belong to class $i$; $TN_i$ (True Negatives) is the proportion of data points that are correctly not assigned to class $i$; $FP_i$ (False Positives) is the proportion of the data points that are incorrectly assigned to class $i$.

E. Model Fusion

The Model fusion method is a deep learning process, by which different classification predictive modeling algorithms associated with individual weights are trained and combined in order to enhance the final estimation. This method turns out to be a stronger meta-classifier as it combines different classification models using a majority voting classifier estimator, partially overcoming the weaknesses of single classifiers and achieving higher classification accuracy. The commonly used voting classifiers include the hard voting classifier and soft voting classifier. The hard voting classifier takes the majority vote applying equal weights to each classifier (mode of all the predicted labels is taken) while the soft voting classifier takes the majority vote based on applying different weights to each classifier (probability of all the predicted labels is taken) [56, 93]. The voting classifier predictions can be defined as:

\[
H_{vote}(x) = \max\{\sum_{j \leq T} \text{lab}(x, j, 1), \sum_{j \leq 2} \text{lab}(x, j, 2), ..., \sum_{j \leq K} \text{lab}(x, j, c)\} \quad (1 \leq j \leq T) \quad (1 \leq c \leq K)
\]

\[
S_{vote}(x) = \max\left\{\frac{\sum_{j \leq 1} p(x, j, 1)}{n_T}, \frac{\sum_{j \leq 2} p(x, j, 2)}{n_T}, ..., \frac{\sum_{j \leq c} p(x, j, c)}{n_T}\right\} \quad (1 \leq j \leq T)
\]

where $H_{vote}(x)$ denotes the vote result of hard voting, lab $(x, j, c)$ is an indicator function that shows if $x$ belongs to label $c$ calculated by $j$th classifier, $S_{vote}(x)$ is the vote result of hard voting, $p(x, j, c)$ is the probability for classifier of exceeding some threshold values, $n_T$ refers to the total number of classifiers and $k$ is the number of labels.

V. EXPERIMENTS

A. DATA PREPARATION

The core dataset of this work is the Swiss Motor Vehicle Information System (MOFIS) [94], which contains over 6.5 million passenger vehicles along with their type approval numbers, geometric and weight properties, ownership details, technical information and date of registration. In addition, we also use vehicle technical specifications and vehicle expert segmentation data from the Technical Type Approval Information from the Federal Roads Office (ASTRA) [28] and a Vehicles Expert Partner [95], respectively.

The data-mining framework consists of three main components: filtering of raw data, extraction of the vehicle sample by registration year, and identification of suitable clustering attributes. In a first step we filter the dataset by removing the vehicles that do not meet the definitions of typical passenger cars, such as small pickup trucks, standard pickup trucks, vans, special purpose vehicles (SPVs), sports cars and multi-purpose vehicles (MPVs) [18]. By considering the goal of this paper, the dataset is then separated into two parts, a training part and a testing part. The training dataset contains 308,824 new registered passenger cars in 2018 along with 30 features including emissions (carbon dioxide (CO$_2$), carbon monoxide (CO), nitrogen oxides (NOx), particulate matter (PM2.5), etc.), weight properties, dimensional features (length, height, width, axle, etc.) and vehicle technical specifications (power, engine capacity, drive, torque, etc.) for each car. It is important to note that we used two different values of CO$_2$ emissions, namely the average type approval values provided in the ASTRA database (measured CO$_2$) and the vehicle specific type approval values that also consider the vehicle weight and the gearbox and are reported in the MOFIS database (calculated CO$_2$).

B. EXPERIMENTAL SETUP AND RESULTS

In the first step of the learning process the training dataset is considered to contain two types of patterns: unlabeled and labeled data. The labeled dataset results from applying the unsupervised FCM clustering algorithm to the total 366 unique new registered passenger cars (based on make, model and manufacturer code) based on the dimensional features: micro class containing 18 samples, small class containing 50 samples, middle class containing 110 samples, upper middle class containing 84 samples, large class and luxury class containing 104 samples. The average accuracy rate and adjusted rand index of the FCM clustering algorithm in comparison to the Swiss expert classification was approximately 79% and 75%, respectively [18]. Due to some limitations of the unsupervised FCM clustering algorithm, only the labeled data with true labels (in vehicle class, measured CO$_2$ and calculated CO$_2$) with a membership degree greater than 0.95 were used as the core dataset to extract the accurate classification of misclassified samples and provide the base for the later step of training. Following
this, we selected 10% of the data from each class as training labeled samples.

The preliminary statistical analysis on the correlation between the emissions, vehicle segments, sub-segments and preselected influencing factors demonstrated a high correlation between the features. In the feature learning process, the unlabeled data and the previous labeled data along with the labels of the core dataset are used as input. Each group of labeled and unlabeled data has a set of features in common. In order to eliminate multicollinearity, principal component analysis (PCA) was performed on the data. Prior to model development, new features are extracted to reduce the number of features. In the feature extraction step, the cluster centroids are defined using algorithm 2 and each patch is transformed to a feature vector.

In the feature selection step (algorithm 3), the resampling (ROS) technique is used in order to increase the number of extracted features from minority groups until it equals the number of features in the majority group. Then, algorithm 4 (based on the Euclidean distance) is used to select the best features and remove redundancy from the feature vector. After we initialized all parts, pseudo labels of labeled data are assigned to the unlabeled data in the training data. Following, this unlabeled data with pseudo labels is used to pre-train the SSFCM, random forest (algorithm 5) and AdaBoost (algorithm 6) classification algorithms by extracting discriminative features. Finally, model fusion is applied using only the labeled data with true labels.

The experimental results show that the single clustering models using SSFCM, random forest and AdaBoost algorithms and the fusion model all enhance the classification accuracy in comparison to the traditional FCM algorithm (overall accuracy of 79%). Among them, the soft voting fusion model and the SSFCM provide the most accurate results, 94.2 and 95.4% respectively. The F-measure value (F1), which represents the model performance, is 91.6% for the SSFCM clustering algorithm and 91.5% for the fusion model with soft voting (Table 1).

From the results of the model fusion, we extract the final features reported in Fig. 5 which we use to re-run the single algorithms and select the final classification model. The underlying assumption of feature extraction is that it leads to improved classification results in comparison to the initial classifier’s predictions with the original features. To verify that this assumption holds for our task, we use the prediction accuracy and other verification measures to check the classification performance of traditional FCM with the original features and the SSFCM, random forest and AdaBoost algorithms with feature extraction. It can be seen in Table 2 that, compared to the FCM-based classifier, the use of feature extraction techniques increases the classification performance. Among all tested approaches, SSFCM provides the best results in terms of both prediction accuracy (95.2%) and verification measures (90.4%) and is therefore selected as our final classification model.

The experimental results demonstrate that the SSFCM algorithm can extract richer information from the vehicle dataset and obtain more discriminative recognition rates than other classifiers do. Therefore, the proposed approach can not only effectively address the problem of multi-class imbalanced data but also improve the prediction performance.

C. DISCUSSIONS

Using the SSFCM model, we estimate the average CO₂ emissions of all new passenger vehicles registered in Switzerland in 2018 to be 138.9 g CO₂/km, which only deviates by 1.1% from the official estimate of the Swiss Federal Office of Energy (SFOE) of 137.8 g CO₂/km [96]. Moreover, for all 26 Swiss Cantons, we find that the correlation between our estimates and those from the SFOE are very high (R² > 0.95). Thus, although slightly different approaches were used to estimate the CO₂ emissions in both cases, the results are highly correlated.

The overall level of emissions from the Swiss passenger cars is strongly affected by the fleet composition, which is shifting in time between classes (from the upper-middle class to the large and luxury classes) and within each class (from non-SUV to SUV).

Fig. 6 shows the distribution of the CO₂ emissions among the different vehicle classes. Both, the median CO₂ emissions and the spread around the median, show a clear upwards trend with the size of the vehicle class.

Fig. 7 shows the average CO₂ emissions (in g CO₂/km) calculated for each vehicle segment resulting from our inter- and intra-class classification. For each vehicle class we report the results based on the interquartile range distributions of the engine power. Overall, we see a significant variation of the CO₂ emissions between vehicle classes, sub-classes and power ranges. Comparing the different vehicle segments, we see that the CO₂ emissions increase with the vehicle size. Moreover, SUV vehicles tend to have significantly higher emissions than non-SUVs. In terms of engine power, it can be seen that within each class, an increase in the engine power generally leads to significantly higher CO₂ emissions.

Fig. 8 shows the spatial distribution of the share of SUVs within the different vehicle segments. It can be seen that the share of SUVs is the lowest for the micro and small classes (between 0 and 20%), followed by the middle class (between 20 and 35%) and the large and luxury class (between 20 and 45%), and is the highest for the upper-middle class (between 30 and 50%). Moreover, we observe that in general the share of SUVs within each class is higher in the southern and center cantons.

Fig. 9 reports the spatial distribution of the average CO₂ emissions for each vehicle class and sub-class. As previously observed, the results demonstrate a high variability in the CO₂ emissions between the different inter and intra-classes. Among the different segments, the CO₂ emissions increase with the size of the vehicle, from around 100 g CO₂/km for
the micro class to around 200 g CO₂/km for the large & luxury class. Moreover, SUV vehicles exhibit generally higher CO₂ emissions than non-SUVs. This difference is extremely remarkable in the case of micro class vehicles, although, as shown in Fig. 7, the share of SUVs is very small in this case. However, this indicates that shifting from a middle class non-SUV vehicle to a micro class SUV could lead to an increase in the CO₂ emissions. In general, the spatial distribution of the CO₂ emissions for each vehicle class and sub-class is quite homogeneous and we do not see any significant trends between different regions of the country.

VI. CONCLUSION

To this paper develops a novel approach to mathematically segment new registered passenger cars and assess the segment-based spatial distributions of the CO₂ emissions. A variety of semi-supervised clustering algorithms are adopted to classify a dataset of new registered passenger cars based on multiple technical, dimensional and emission features. Among all tested classifiers, the SSFCM technique was the most accurate, providing a classification accuracy of about 90.4% for verification measures.

The proposed approach enables accurate automated vehicle classification of large databases, which in turn facilitates the analysis of fleet changes. Another important advantage of the clustering based mathematical segmentation is that it removes the subjectivity factors affecting expert-based segmentations, reducing classification errors and making databases from across the world comparable. Finally, the automatized clustering approach also reduces classification costs and training time.

Despite technology improvements, the Swiss passenger car fleet remains emission intensive. Our results indicate large variabilities in the average CO₂ emissions of different vehicle classes. While a shift of the fleet towards smaller vehicles is likely to diminish CO₂ emissions, the emissions intensity could be more effectively reduced by shifting the vehicles proportion within each class (e.g., switching from SUV to non-SUV or to lower power vehicles in the same vehicle class). Therefore, the combination of the inter-class and intra-class classification provides crucial insights for developing fleet transformation strategies to decarbonize the passenger vehicle fleet. A further area of potentially fruitful research would be to use CO₂ estimates from real world measurements instead of type approval values for a more precise evaluation of the fleet CO₂ emissions.

ACKNOWLEDGMENT

The authors thank the Federal Roads Office (FEDRO) for providing the Swiss Vehicle Information System (MOFIS) data and the technical dataset and the Vehicle Expert Partner (Au-to-i-dat) for providing the expert segmentation data. In addition, the authors would like to express their gratitude to Philippe Zimmermann from the Automotive Powertrain Technologies Laboratory (Empa) for his support in CO₂ emissions estimation.

REFERENCES

[1] The Paris Agreement | UNFCCC. https://unfccc.int/process-and-meetings/the-paris-agreement/the-paris-agreement. Accessed: Sep. 2021.
[2] EU Reference Scenario 2016, Energy, Transport and GHG Emissions. Trends to 2050, Euro-pan Commission, 2016. https://ec.europa.eu/energy/sites/ener/files/documents/ref2016_report_final-web.pdf
[3] P. Capros, D. Vita, A. Tasios, N. Siskos, P. Kannavou, M. Petropoulos, A. Evangelopoulou, S. Zampara, et al., EU Reference Scenario 2016 - Energy, transport and GHG emissions Trends to 2050. European Commission Directorate - General for Energy, Directorate - General for Climate Action and Directorate - General for Mobility and Transport, 2016, Luxembourg.
[4] International Energy Agency (IEA) 2018, World Energy Outlook 2018, IEA, Paris.
[5] Federal Office for the Environment (FOEN) 2021, https://www.bafu.admin.ch/bafu/en/home/topics/climate/state/data/greenhouse-gas-inventory/transport.html, Accessed: Dec. 2021
[6] Federal Energy Research Commission CORE, “Energy research Masterplan of the Federal Government 2021–2024”, 2020, https://pubdb.bfe.admin.ch/en/publication/download/10329.
[7] N. Niroomand, C. Bach, M. Elser, “Vehicle Dimensions Based Passenger Car Classification Using Fuzzy and Non-Fuzzy Clustering Methods”, Transportation Research Record, 2021. DOI: 10.1177/03611981211010795
[8] N. Niroomand, C. Bach and M. Elser, "Robust Vehicle Classification Based on Deep Features Learning," in IEEE Access, vol. 9, pp. 95675-95685, 2021, doi: 10.1109/ACCESS.2021.3094366.
[9] J. Krause, C. Thielen, M. Tsokolis, Z. Samaras, C. Rota, A. Ward, P. Prenninger, T. Coosemans, N. Neugebauer, W. Verhoeven. “EU road vehicle energy consumption and CO₂ emissions by 2050 – Expert-based scenarios”, Energy Policy, Vol. 138, pp. 111224, 2020.
[10] N. Zacharof, G. Pantosas, B. Ciuffo, S. Tsikamakis, K. Anagnostopoulos, M. Marotta, et al., “Review of in use factors affecting the fuel consumption and CO₂ emissions of passenger cars Euro Commission”, Luxembourg 2015.
[11] D. Tsokolis, S. Tsikamakis, A. Dimaratos, G. Pantosas, P. Pistikopoulos, B. Ciuffo, et al., “Fuel consumption and CO₂ emissions of passenger cars over the new worldwide harmonized test protocol”. Appl Energy, vol. 179, pp. 1152-1165, 2016.
[12] J. Pavlovic, A. Manea, B. Ciuffo, S. Serra, G. Pantosas, K. Anagnostopoulos, et al., “Correction of test cycle tolerances: evaluating the impact on CO₂ results”, Transp Res Procedia, vol. 14, pp. 3099-3108, 2016.
[13] A. Dimaratos, D. Tsokolis, G Fontaras, S Tsikamakis, B. Ciuffo, Z Samaras. “Comparative evaluation of the effect of various technologies on light-duty vehicle CO₂ emissions over NEDC and WLTP”. Transp Res Procedia, vol. 14, pp. 3169-3178, 2016.
[14] EC, A European Strategy for Low-Emission Mobility. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions 2016, COM (2016) 501 final
[15] S.L. Shi, Q. Zhong, J.M. Xu, “Robust Algorithm of Vehicle Classification”, Software Engineering, Artificial Intelligence, Networking, and Parallel/ Distributed Computing, 2007. SNPD
[16] W. Shi, Y. Gong, C. Ding, Z. Ma, X. Tao, and N. Zheng, “Transductive semi-supervised deep learning using min-max features,” in Proc. Eur. Conf. Comput. Vis. (ECCV), vol. 11209. Cham, Switzerland: Springer, pp. 299-315, 2015.
[17] X. Zhu, “Semi-Supervised learning literature survey”, 2008. http://pages.cs.wisc.edu/~jerryzhu/research/ssl/semireview.html.
[18] L. Zhuo, L.Y. Jiang, et al., “Vehicle Classification for Large-Scale Traffic Surveillance Videos Using Convolutional Neural Networks”. 2016.
[19] G. Forestier and C. Wenmert, “Semi-supervised learning using multiple clusterings with limited labeled data,” Inf. Sci., vols. 361-362, pp. 48-65, Sep. 2016.
[20] Chapelle, O., Schölkopf, B., & Zien, A. “Semi-supervised learning (1st ed.)”, Cambridge: The MIT Press. 2006.
[21] S. Melacci and M. Belkin, “Laplacian support vector machines trained instead of type approval values for a more precise evaluation of the micro class to around 200 g CO₂/km for the large & luxury class. Moreover, SUV vehicles exhibit generally higher CO₂ emissions than non-SUVs. This difference is extremely remarkable in the case of micro class vehicles, although, as shown in Fig. 7, the share of SUVs is very small in this case. However, this indicates that shifting from a middle class non-SUV vehicle to a micro class SUV could lead to an increase in the CO₂ emissions. In general, the spatial distribution of the CO₂ emissions for each vehicle class and sub-class is quite homogeneous and we do not see any significant trends between different regions of the country. the micro class to around 200 g CO₂/km for the large & luxury class. Moreover, SUV vehicles exhibit generally higher CO₂ emissions than non-SUVs. This difference is extremely remarkable in the case of micro class vehicles, although, as shown in Fig. 7, the share of SUVs is very small in this case. However, this indicates that shifting from a middle class non-SUV vehicle to a micro class SUV could lead to an increase in the CO₂ emissions. In general, the spatial distribution of the CO₂ emissions for each vehicle class and sub-class is quite homogeneous and we do not see any significant trends between different regions of the country.
D.P. Van Vuuren, M. Hoogwijk, T. Barker, K. Riahi, S. Boeters, J. S.D. Tuladhar, M., Yuan, P. Bernstein, W.D. Montgomery, A. Smith, down versus bottom-up projections of China's regional energy use and CO₂ emissions test. Environ. Res., vol. 176, pp: 108572, 2019.

L. Ntziaichristos, G. Mellios, D. Tsokolis, M. Keller, S. Hausberger, N.E. Ligterink, P. Dila. In-use vs. type-approval fuel consumption of current passenger cars in Europe Energy Policy, 2014, 67, pp: 403-411.

J. Li, J. Mariani, T. Kanade, “Discriminative semi-supervised feature selection via manifold regularization”, IEEE Trans. Neural Netw. vol. 21, no.7, pp: 1033–1047, 2010.

Z. Xu, I. King, M.R. Lyu, R. Jin, “Discriminative semi-supervised feature selection based on manifold regularization”, IEEE Trans. Neural Netw. vol. 21, no.7, pp: 1033–1047, 2010.

F.D. Torre, T. Kanade, “Discriminative clustering for Imbalanced Multi-Class Classification,” in IEEE Access, vol. 7, pp. 28100-28112, 2019, doi: 10.1109/ACCESS.2019.2901860.

Q. Li, F. Qiao, L. Yu, “A Machine Learning Approach for Light-Duty Vehicle Idling Emission Estimation Based on Real Driving and Environmental Information”. Journal of Environmental Pollution and Climate Change, vol. 1, 2017.

Z. He, G. Ye, H. Jiang, Y. Fu, “Vehicle Emission Detection in Data-Driven Methods”, Mathematical Problems in Engineering, vol. 2020, Article ID 4875310, 13 pages, 2020. https://doi.org/10.1155/2020/4875310.

M. Ghahramani and F. Pilla, “Analysis of Carbon Dioxide Emissions from Road Transport Using Taxi Trips,” in IEEE Access, vol. 9, pp: 98573-98580, 2021, doi: 10.1109/ACCESS.2021.3096279.

Padmapriya G. and Duraiswamy, K. “Association of deep learning from Road Transport Using Taxi Trips,” in IEEE Access, vol. 9, pp: 98573-98580, 2021, doi: 10.1109/ACCESS.2021.3096279.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/
[66] W. Maungmai, C. Nuthong, “Vehicle Classification with Deep Learning” In IEEE 4th International Conference on Computer and Communication Systems, ICCCSS 2019, Singapore, February vol. 23-25, pp 294-298, 2019.

[67] Z. Dong, Y. Wu, M. Pei et al., “Vehicle type classification using a semi supervised convolutional neural network.” IEEE Trans. Intelligent Transportation Systems, vol. 16, no. 4, pp: 2247-2256, 2015.

[68] M. Bilenko, S. Basu, R.J. Mooney, “Integrating constraints and metric learning in semi-supervised clustering”, in: Proceedings of the International Conference on Machine Learning, pp. 81–88, 2004.

[69] Y. Ren, X. Hu, K. Shi, G. Yu, D. Yao, Z. Xu, “Semi-supervised den peak clustering with pairwise constraints”, in: Proceedings of the Fifteenth Pacific Rim International Conference on Artificial Intelligence, pp. 837– 850, 2018.

[70] N. Grira, M. Crucianu, N. Boujemaa, “Active semi-supervised fuzzy clustering”, Pattern Recognition, vol. 41, no.5, 1834-1844, 2008.

[71] N. Grira, M. Crucianu, N. Boujemaa, “Unsupervised and semi-supervised clustering: A brief survey”, In A Review of Machine Learning Techniques for Processing Multimedia Content, Report of the MUSCLE European Network of Excellence, 2004.

[72] W. Qiu, “Based on Similarity Metric Learning for Semi-Supervised Clustering”, Sensors and Transducers Journal, vol. 177, no.8, pp. 238-245, 2014.

[73] Y. Qin, S. Ding, L. Wang, Y. Wang, “Research Progress on Semi-Supervised Clustering”, Cognitive Computation, vol. 11, pp: 599-612, 2019.

[74] Y. Re, X. Hu, K. Shi, G. Yu, D. Yao, Z. Xu, “Semi-supervised denpeak clustering with pairwise constraints”, in: Proceedings of the Fifteenth Pacific Rim International Conference on Artificial Intelligence, pp. 837–850, 2019.

[75] J. Forestier, C. Wenmert, “Semi-supervised learning using multiple clusterings with limited labeled data”, Information Sciences, vol. 361–362, pp: 48-65, 2016.

[76] A. Arshad, S. Riaz, L. Jiao, and A. Murthy, “Semi-supervised deep fuzzy c-mean clustering for software fault prediction,” IEEE Access, vol. 6, pp. 25675-25685, 2018.

[77] A. Arshad, S. Riaz, L. Jiao, and A. Murthy, “The empirical study of semi-supervised deep fuzzy c-mean clustering for software fault prediction,” IEEE Access, vol. 6, pp. 47047-47061, 2018.

[78] W. Shi, Y. Gong, C. Ding, ZMX. Tao, N. Zheng, “Transductive semi-supervised deep learning using min-max features” Proceedings of the European Conference on Computer Vision (ECCV), 299-315, 2018.

[79] G. Chen, “Deep transductive semi-supervised maximum margin clustering Compute”. Res. Repos, pp 1-14. arXiv: 1501.06237.

[80] S. Wang and X. Yao, “Multiclass imbalance problems: Analysis and potential solutions,” IEEE Trans. Syst., Man, Cybern., B (Cybern.), vol. 42, no. 4, pp. 1119-1130, Aug. 2012.

[81] Verikas, A. Gelzinis, A. Bacauskiene, M. “Mining data with random forests: A survey and results of new tests”. Pattern Recognit., vol. 44, pp: 330–349, 2011.

[82] J.C. Dunn “A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters”. Journal of Cybernetics, vol. 3, no. 3, 32 – 57, 1973.

[83] A. Arshad, S. Riaz and L. Jiao, “Semi-Supervised Deep Fuzzy C-Mean Clustering for Imbalanced Multi-Class Classification,” in IEEE Access, vol. 7, pp. 28100-28112, 2019, doi: 10.1109/ACCESS.2019.2901860.

[84] S. Riaz, A. Arshad and L. Jiao, “A Semi-Supervised CNN With Fuzzy Rough C-Mean for Image Classification”, in IEEE Access, vol. 7, pp. 49641-49652, 2019, doi: 10.1109/ACCESS.2019.2910406.

[85] Y. Ren, X. Hu, K. Shi, G. Yu, D. Yao, Z. Xu, “Semi-supervised deep clustering with pairwise constraints, in: Proceedings of the Fifteenth Pacific Rim International Conference on Artificial Intelligence, pp. 837–850, 2018.

[86] L. Wang, M. Han, X. Li, N. Zhang and H. Cheng, ”Review of Classification Methods on Un-balanced Data Sets,” in IEEE Access, vol. 9, pp. 64606-64628, 2021, doi: 10.1109/ACCESS.2021.3074243.

[87] T. M. Khoshgoftar, C. Seiffert, J. V. Hulse, A. Napolitano, A. Folleco, “Learning with limited minority class data”. Int. Conf. Mach. Learn. Appl., pp. 348-353, 2007.

[88] R. Pu, “Semi-Supervised learning literature survey”, 2008 http://pages.cse.wisc.edu/~jerryzhu/research/sis/semioverview.html

[89] T. Hasanin, T. M. Khoshgoftar, J. Leevy, N. Seliya, “Investigating Random Under-sampling and Feature Selection on Bioinformatics Big Data”. 2019 IEEE Fifth International Conference on Big Data Computing and Service and Applications (BigDataService). pp. 346-356, doi: 10.1109/BigDataService.2019.00063.

[90] B. Handaga, M. Mat Deris, “Similarity Approach on Fuzzy Soft Set Based Numerical Data Classification. In: Zain J.M., Wan Mohd W.M., El Qwasmeh E. (eds) Software Engineering and Computer Systems. ICSECS 2011. Communications in Computer and Information Science, vol 180. Springer, Berlin, Heidelberg. 2011, doi:10.1007/978-3-642-22191-0_50

[91] S. W. Kwok and C. Carter, “Multiple decision trees,” Uncertainty in Artificial Intelligence, vol. 9, pp. 327–335, 1990.

[92] R. Kumar and R. Verma, “Classification algorithms for data mining: a survey,” International Journal of Innovations in Engineering and Technology (IJIET), vol. 1, no. 2, pp. 7–14, 2012.

[93] C. Macdonald and I. Ounis, “Voting for candidates: adapting data fusion techniques for an expert search task,” in Proceedings of the 15th ACM International Conference on Information and Knowledge Management, pp. 387–396, Arlington, VA, USA, November 2006.

[94] This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/
### TABEL 1. Performance of tested models on dataset with labeled rate of 10% from each class

| Method       | Accuracy rate | Precision rate | Recall  | F1    |
|--------------|---------------|----------------|---------|-------|
| SSFCM        | 0.954         | 0.953          | 0.881   | 0.916 |
| AdaBoost     | 0.891         | 0.871          | 0.823   | 0.846 |
| Random Forest| 0.902         | 0.89           | 0.86    | 0.875 |
| Hard Voting  | 0.921         | 0.935          | 0.871   | 0.902 |
| Soft Voting  | 0.942         | 0.956          | 0.878   | 0.915 |

### TABLE 2. Prediction accuracy and verification clustering results

| Method       | Training accuracy | Test accuracy |
|--------------|-------------------|---------------|
| SSFCM        | 0.952             | 0.904         |
| AdaBoost     | 0.781             | 0.715         |
| Random Forest| 0.903             | 0.837         |
FIGURE 1. Type approval CO$_2$ emission measurements are conducted for new vehicles using regulatory cycles. Under the New European Driving Cycle (NEDC), the average level of CO$_2$ emissions from new passenger cars registered may not exceed 95 g CO$_2$/km in 2021. This corresponds to a WLTP (world harmonized light-duty vehicles test procedure) target of 118 g CO$_2$/km. The switch to the WLTP procedure results on average on 20% higher emissions than with the NEDC [13].

FIGURE 2. Distribution of population and key parameters describing road mobility in urban, suburban and rural areas in Switzerland. Source: Statistik der Schweizer Städte [26], Global Traffic Scorecard (Inrix) [27].
FIGURE 3. Average CO₂ emissions (in g CO₂/km) of new passenger car registrations in Switzerland in 2018.

FIGURE 4. The structure of the proposed semi-supervised deep learning approach.
FIGURE 5. Importance of features of the proposed semi-supervised clustering algorithms

FIGURE 6. Distribution of the CO₂ emissions among the different vehicle classes.
FIGURE 7. New registered passenger cars average CO₂ emissions intensity based on the interquartile range power range (Q) by both vehicle inter-class and intra-class classification in 2018.

FIGURE 8. Spatial distribution of the share of SUV vehicles (in %) among different vehicle classes.
FIGURE 9. Spatial distribution of the average CO₂ emissions (in g CO₂/km) among different vehicle classes and sub-classes.