Enhancing Vehicle Identification in Challenging Conditions Through Fine-Grained Classification

Eduardo Santos*,†, Gabriel E. Lima†, Rayson Laroca‡,†, Eduil Nascimento Jr.*, and David Menotti†

*Department of Technological Development and Quality, Paraná Military Police, Curitiba, Brazil

†Department of Informatics, Federal University of Paraná, Curitiba, Brazil

‡Postgraduate Program in Informatics, Pontifical Catholic University of Paraná, Curitiba, Brazil

*{ed.santos,eduiljunior}@pm.pr.gov.br †{gelima,menotti}@inf.ufpr.br ‡rayson@ppgia.pucpr.br

***Abstract*—Automatic License Plate Recognition (ALPR)-based Vehicle Identification Systems are vital for modern traffic management and law enforcement, yet they frequently encounter challenges in real-world scenarios, leading to recognition errors. Our ongoing research investigates a novel method to improve ALPR accuracy through the integration of Fine-grained Vehicle Classification (FGVC), with a particular focus on vehicle make identification. By cross-referencing identified vehicle brands with registered data, we aim to enhance the reliability of ALPR-based vehicle identification systems. Nonetheless, initially, our work-in-progress is concentrated on refining FGVC techniques to facilitate their integration with ALPR. We assess four deep learning models for vehicle make classification and explore methods such as selective prediction and a new class reduction approach. Preliminary results from our experiments on a modified Brazilian vehicle dataset indicate that combining these methods significantly boosts vehicle make identification accuracy. This improved classification approach is anticipated to reduce false positives and increase recognition rates under challenging conditions. Future efforts are going to be directed towards integrating these observed enhancements into ALPR-based vehicle identification systems to further improve their performance in real-world applications.**

I. Introduction

Vehicle identification systems based on ALPR are integral to Intelligent Transportation Systems, supporting a broad spectrum of applications, such as toll collection, vehicle access control, and criminal forensics [1], [2]. Despite their critical role, these systems continue to encounter challenges in real-world conditions, including poor lighting, obscured license plates, and non-standard plate designs [3], [4]. Notably, the Military Police of Paraná in Brazil has reported frequent inaccuracies in their ALPR system during field deployments.

To address these persistent challenges, we propose incorporating vehicle make information as a cross-validation measure for license plate recognition in vehicle identification tasks. This approach can be especially valuable in challenging scenarios where errors occur in license plate recognition, allowing the vehicle make data to validate results and narrow the search space for identifying the correct license plate. Vehicle make identification is particularly effective because, while less complex than vehicle model classification, it remains sufficiently descriptive to help correct misrecognized license plates.

The identification of a vehicle's make is a key area of research within FGVC, a field dedicated to classifying vehicles based on attributes such as type, make, model, and year [5]–[7]. Recent advancements in deep learning have significantly improved FGVC-related tasks. However, integrating these classification systems with ALPR to enhance vehicle recognition accuracy remains underexplored. Challenges in merging these fields stem from the differing scenarios they address, with ALPR systems often being deployed in more challenging environments. Therefore, it is vital to advance FGVC not only to achieve better classification outcomes but also to complement license plate recognition methodologies.

With this in mind, our research focuses on developing and evaluating a vehicle make classification system specifically designed for integration with ALPR. We explore four deep learning architectures and investigate methods to enhance classification accuracy and reliability. This study focuses on two approaches: (i) selective prediction, where the model abstains from making a prediction when confidence is low [8]; and (ii) class reduction, a method that simplifies the number of make categories to boost accuracy without significantly compromising the problem's representativeness.

For our experiments, we expanded the Rodosol-ALPR [9] dataset by incorporating vehicle make information. This dataset was chosen because of its widespread use in ALPR research [3], [10], [11], its realistic representation of intelligent transportation systems scenarios, and its license plate annotations, which support effective FGVC labeling. We first assess the performance of the models in terms of accuracy, precision, and recall. Afterward, we investigate the impact of selective prediction and class reduction techniques using the top-performing model.

The remainder of this work is organized as follows. Section II gives an overview of related works. Section III describes the data preparation process. Section IV details the experiments and results. Section V concludes the work and explores potential directions for future research.

II. Related Work

Vehicle make classification on FGVC has emerged as a prominent area of research in computer vision, with recent advancements significantly enhancing intelligent transportation systems. For instance, the use of an adaptive attention mechanism in Convolutional Neural Networks (CNNs) to focus on critical regions of a vehicle has led to a 94.1% accuracy rate in vehicle make classification on the CompCars dataset [12]. Similarly, a part-guided attention mechanism, which autonomously distinguishes vehicle parts, has achieved an impressive 97.9% accuracy on the same dataset [5].

Another innovative approach was presented by Sochor et al. [13], which employs a multi-task learning framework to simultaneously classify vehicle make, model, and year. Their method reached a stunning 97.6% accuracy in make classification on the Stanford Cars dataset. In a different vein, the hierarchical vision transformer model leverages a hierarchical structure to capture features at various levels of granularity, achieving 96.8% accuracy in make classification on the VehicleID dataset [14].

Contrastive learning has also made strides in the field, as exemplified by the contrastive vehicle classification learning model [15]. By employing hard negative sampling, this model enhances the discriminative power of feature representations, achieving a 95.3% accuracy rate in make classification on the BoxCars116k dataset. Similarly, the multi-scale fusion network model [16] integrates features from different scales to capture both fine details and broader context, attaining a 98.2% accuracy on the CompCars dataset.

Few-shot learning techniques have been explored to enable the classification of new vehicle makes with minimal training data. For example, a prototype-based meta-learning approach achieved a 91.7% accuracy with only five examples per class on the VehicleID dataset [17]. Additionally, Generative Adversarial Networks (GANs) have been employed for data augmentation, with the vehicle GAN method [18] generating synthetic vehicle images that improved make classification accuracy by 3.5% on the Stanford Cars dataset.

These advancements in fine-grained vehicle classification, particularly in make classification, offer significant opportunities to enhance vehicle identification systems by fusing their results with ALPR. However, despite the promising outcomes reported by existing research, many studies are often carried out under controlled conditions that do not fully capture the diverse and challenging conditions encountered in real-world scenarios where ALPR systems are typically deployed.

III. Data Preparation

The RodoSol-ALPR dataset [9] comprises 10,000 images of cars and 10,000 images of motorcycles, all captured by stationary cameras at toll booths along a Brazilian highway (see Fig. 1). To adapt the dataset for vehicle make classification, several preprocessing steps were required. This section outlines the data preparation process applied to the dataset to ensure its suitability for the proposed experimental research.

Inspired by [19], all motorcycle images were excluded due to the limited research on accurate make identification for motorcycles, making it an inconclusive area of study. To maintain the focus on well-established tasks, only car images were retained. The images were then standardized to ensure that they depicted vehicles without background information. Since the dataset did not include bounding box annotations, the YOLOv10 [20] model was employed for vehicle detection.

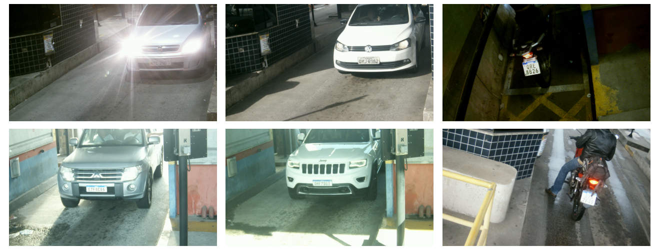


Fig. 1. Sample images from the RodoSol-ALPR dataset [9], showcasing its diversity of vehicle types and varying lighting conditions. For this figure, the images have been slightly resized for better viewing.

Next, the images underwent manual curation to ensure their suitability for the research problem. This process involved grouping multiple images of the same vehicle based on license plate annotations and selecting those that exhibited variations in lighting, angle, or other features [21]. Images with significant occlusions, such as those showing only the vehicle's bumper, were excluded to ensure accurate make classification.

For annotation purposes, license plate data was used to retrieve vehicle information from Brazil's National Traffic Secretariat (SENATRAN) database, enabling the automatic assignment of vehicle make labels. Fewer than 5% of the recovered license plates lacked consistent information, requiring manual annotation. Images with ambiguous make identification were excluded to minimize errors.

The final dataset, referred to as the *VehicleMake* dataset, comprises 9,553 images categorized into 29 vehicle make classes. This includes prominent brands such as Chevrolet, Fiat, Ford, Honda, Toyota, and Volkswagen. The class distribution, shown in Fig. 2, reflects the prevalence of specific makes in Brazilian traffic [22]. Fig. 3 showcases examples of images from the most common classes.

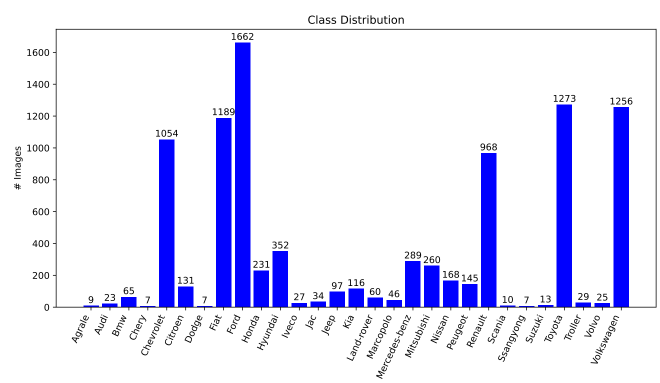


Fig. 2. Distribution of classes in the VehicleMake dataset.

Naturally, the selected images from the RodoSol-ALPR dataset, along with their corresponding vehicle make annotations and our division into training, test and validation subsets, will be made available upon request.

IV. Experiments and Results

This section provides an overview of the experimental research conducted, detailing each experiment and analyzing

Fig. 3. Examples of images from the *VehicleMake* dataset. The corresponding vehicle make annotation is shown above each image. In this figure, the images were resized to a uniform size for better viewing.

its results. Each experiment was repeated five times with different dataset splits, and the results were averaged across these iterations. The baseline metrics assessed include Top-1 accuracy, Top-2 accuracy, precision, recall, and F1-score, all reported using class macro averaging. Further exploration with the best-performing model focused on the correct classification rate and did not include precision, recall, or F1-score metrics.

The dataset images were split into three subsets: 75% for training, 12.5% for validation, and 12.5% for testing. Each subset maintains the original class distribution found in the entire dataset. When the exact class percentages did not allow for precise splitting, any excess images were randomly allocated to either the validation or test subsets.

The remainder of this section is organized as follows. Section IV-A outlines the baseline experiment, where four deep learning models were trained for vehicle make classification. Section IV-B details the class reduction experiment performed with the best-performing baseline model. Section IV-C covers the selective prediction approach applied to this model and evaluates the effectiveness of combining online class reduction with selective prediction.

A. Baseline Experiment

The baseline experiment involved training four deep learning models for vehicle make classification: ResNet-34 [23], MobileNetV3 [24], EfficientNetV2 [25], and ViT b16 [26]. These models were chosen due to their effectiveness in image classification and fine-grained tasks [27], [28], as well as the availability of pre-trained implementations that enhance reproducibility. A transfer learning approach was used, initializing the models with pre-trained weights from ImageNet [29] and modifying the final fully connected layer to match the number of vehicle make classes. Only these final layers were trained.

The Adam optimizer was employed with $\beta_1 = 0.9$, $\beta_2 = 0.999$, a batch size of 128, a weight decay of 10^{-5}, and an initial learning rate of 10^{-4}. A learning rate reduction scheme was implemented with a patience value of 10 epochs and a reduction factor of 0.1. Training was carried out for up to 400 epochs, with early stopping applied if no improvement was observed over 15 epochs. Cross-entropy loss was used as the loss function. Two training protocols were employed: *(p1)* training with data augmentation only, and *(p2)* utilizing data augmentation techniques to increase the representation of minority classes within batches by generating synthetic data, known as oversampling. This approach can improve performance by balancing the dataset distribution [30].

The data augmentation process began by resizing images to 224×224 pixels to align with the model's input size requirements. A variety of affine transformations were then applied, including rotations up to $180°$, scaling between 0.9 and 1.3, and shearing up to $15°$, all with a 50% probability. Additionally, brightness and contrast were randomly adjusted within a range of 0.2, with a 30% probability, and images were blurred using a generalized normal filter with randomly selected parameters, with a 40% probability. There was also a 25% chance that a random 72×72 pixel region of the image would be filled with random noise. Finally, each image was normalized using the mean and standard deviation values from ImageNet [29].

Table I summarizes the performance of each model across the dataset using the training protocols *(p1)* and *(p2)*. Models trained under the second protocol generally achieved higher accuracy compared to those trained under the first, although it comes with a slight trade-off in precision. The ViT b16 model delivered the best performance, reaching a Top-1 accuracy of 65.4% and a Top-2 accuracy of 73.8% under protocol *(p2)*. These results demonstrate strong generalization capabilities, even in the challenging task of FGVC. Consequently, this model was the selected one for further exploration in the following subsections.

TABLE I
GLOBAL METRICS ACHIEVED BY ALL MODELS ON THE VEHICLE MAKE CLASSIFICATION TASK (AVERAGED OVER FIVE RUNS). PROTOCOL *(p2)* INCLUDES OVERSAMPLING OF MINORITY CLASSES, WHILE PROTOCOL *(p1)* DOES NOT.

Protocol	Model	Top-1	Top-2	Precision	Recall	F1
(p1)	ViT b16	**55.3%**	**62.6%**	**63.9%**	**55.3%**	**57.4%**
	ResNet-34	38.7%	47.8%	49.3%	38.7%	41.1%
	EfficientNetV2	39.3%	49.1%	45.8%	39.3%	39.5%
	MobileNetV3	40.9%	50.9%	52.2%	40.9%	43.5%
(p2)	ViT b16	**65.4%**	**73.8%**	**53.0%**	**65.4%**	**56.8%**
	ResNet-34	49.4%	61.8%	33.9%	49.4%	36.9%
	EfficientNetV2	49.4%	60.2%	31.7%	49.4%	33.8%
	MobileNetV3	50.7%	61.8%	37.7%	50.7%	41.2%

B. Class Reduction Experiment

The class reduction approach aims to improve the accuracy of the best-performing model by simplifying the vehicle make classification problem while preserving data representativity essential for an ALPR system. To achieve this, we reduced the number of classes from 29 to 11. The top-10 most prevalent vehicle makes in the Brazilian context [22] were retained as individual classes, while the remaining makes were grouped into a single class labeled "Others." Table II presents the final dataset classes used for the class reduction experiments, along with the total number of images for each class.

TABLE II
CLASS DISTRIBUTION CONSIDERED FOR THE CLASS REDUCTION EXPERIMENTS AND THE RESULTING DISTRIBUTION OBTAINED.

Class	Images
Chevrolet	1,054
Fiat	1,189
Ford	1,662
Honda	231
Hyundai	352
Jeep	97
Nissan	168
Renault	968
Toyota	1,273
Volkswagen	1,256
Others	1,303

The experiment employs two evaluation methodologies:

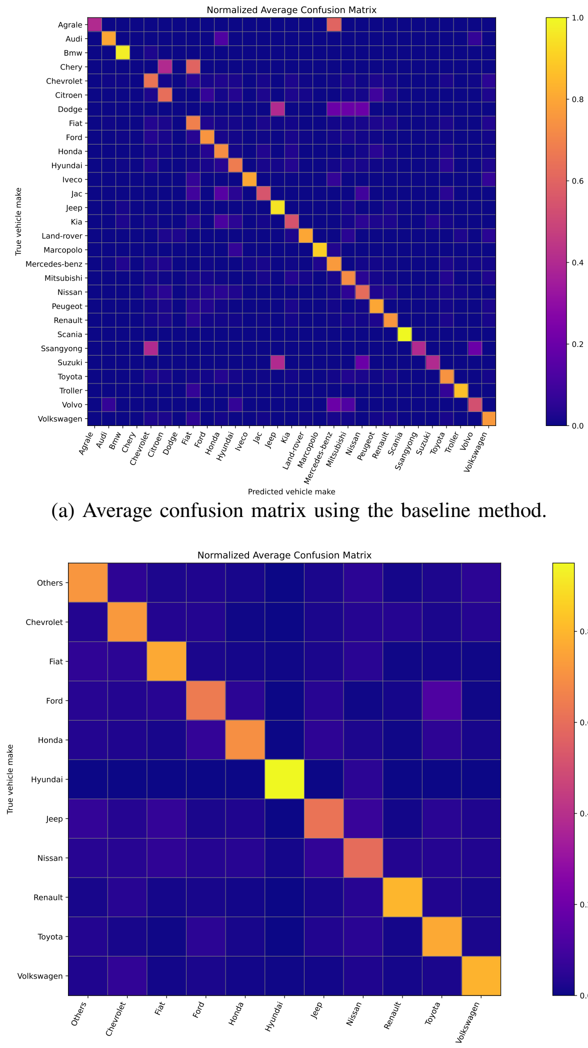
1) **Static Class Reduction:** This method follows the baseline approach but incorporates the new class definitions as specified in Table II. Only the best-performing model, ViT b16 with protocol *p2*, is retrained;
2) **Online Class Reduction:** Rather than retraining the best-performing model, this method adjusts the model's output exclusively during the evaluation phase, based on the new class definitions. For instance, if the ground truth for a sample is "Audi" and the model predicts "BMW," the prediction is deemed correct as both brands fall under the "Others" category.

Table III compares the baseline results with those obtained using the best-performing model under the class reduction approaches. Both static and online methods improved the baseline accuracy by simplifying the vehicle make classification task. Given the similarity of the results from these approaches, we performed a paired t-test at a 5% significance level to determine any statistical differences. The results indicate that, on average, the static method produces more accurate predictions than the online method, confirming a statistically significant difference.

TABLE III
GLOBAL TOP-1 AND TOP-2 ACCURACY ACHIEVED BY VIT B16 FOR VEHICLE MAKE CLASSIFICATION USING STATIC AND ONLINE CLASS REDUCTION APPROACHES (AVERAGED OVER FIVE RUNS).

Method	Top-1	Top-2
Baseline	65.4%	73.8%
Static class reduction	**73.4%**	**85.0%**
Online class reduction	71.1%	81.5%

Fig. 4 shows the average confusion matrices for the baseline experiment (Fig. 4a) and the static class reduction experiment (Fig. 4b), which yielded the best performance as shown in Table III. The matrices reveal that when all original classes are considered, there are more incorrect predictions, particularly for the less-represented classes. In contrast, by reducing the number of classes and consolidating those not among the top 10 most prevalent vehicle makes in Brazil, misclassifications are reduced. This grouping helps minimize errors associated with these less frequent classes.



(a) Average confusion matrix using the baseline method.

(b) Average confusion matrix using the static class reduction method.

Fig. 4. Average confusion matrices for the baseline experiment (a) and the static class reduction experiment (b). The brightness of each pixel directly corresponds to the normalized frequency of samples at that matrix position, with brighter pixels indicating higher frequencies.

C. Selective Prediction Experiment

Selective prediction allows a model to reject predictions based on certain criteria [8]. In this experiment, we implemented a naive softmax-response rejection method [8] during the evaluation phase of the baseline best-performing model. This straightforward approach neither alters the model's architecture nor its training process. It leverages the softmax probability distribution to decide whether to accept or reject predictions based on whether the highest probability surpasses a predefined threshold. We tested nine threshold values ranging from 0.1 to 0.9, in increments of 0.1. To further explore selective prediction, we conducted an additional experiment combining it with the online class reduction method, as both techniques are applied solely during the evaluation phase.

Table IV compares the baseline results with those obtained using selective prediction (both with and without online class reduction) for the ViT b16 model at various confidence thresholds. It also presents the total number of rejected images and the number of originally correct predictions that were incorrectly rejected, expressed as both absolute values and percentages of the rejections from the test subset. The data indicate that applying the selective prediction method enhances accuracy compared to the baseline results. Moreover, combining selective prediction with online class reduction yields even higher accuracy. However, higher thresholds result in increased rejection rates and more correct predictions being incorrectly rejected, thereby limiting further improvements in accuracy.

Finding a balance between accuracy improvements and acceptable rejection rates is crucial. A threshold of 0.4 provides an optimal trade-off for the selective prediction approach, improving accuracy while maintaining a reasonable rejection rate. This method outperforms the baseline results without significantly compromising correct predictions. By combining the selective prediction approach with the online class reduction method, the best performance was achieved at a threshold of 0.5, yielding Top-1 and Top-2 accuracies of 90.2% and 94.6%, respectively. Although this threshold does not offer the highest accuracy, it maintains a rejection rate below 45%.

TABLE IV
GLOBAL TOP-1 AND TOP-2 ACCURACY VALUES ACHIEVED BY VIT B16 ON THE VEHICLE MAKE CLASSIFICATION TASK USING ONLY SOFTMAX RESPONSE REJECTION, AND COMBINING IT WITH ONLINE CLASS REDUCTION APPROACH (AVERAGED OVER FIVE RUNS). REJECTION RATES ARE PRESENTED AS BOTH ABSOLUTE NUMBERS AND PERCENTAGES.

Method	Minimum confidence	Rejected images	Correct predictions incorrectly rejected	Top-1	Top-2
Baseline	–	–	–	65.4%	73.9%
Selective Prediction	0.1	0 / 0.0%	0 / 0.0%	65.4%	73.9%
	0.2	37 / 3.1%	8 / 23.0%	66.7%	75.0%
	0.3	207 / 17.3%	62 / 30.1%	71.9%	79.5%
	0.4	382 / 32.0%	145 / 38.0%	77.0%	82.5%
	0.5	534 / 44.7%	239 / 44.8%	75.4%	79.7%
	0.6	655 / 54.9%	328 / 50.1%	75.8%	78.5%
	0.7	770 / 64.5%	426 / 55.3%	75.8%	77.0%
	0.8	876 / 73.4%	523 / 59.7%	75.3%	75.8%
	0.9	1013 / 84.8%	656 / 64.8%	71.3%	71.5%
Selective Prediction + Online Class Reducing	0.1	0 / 0.0%	0 / 0.0%	71.1%	81.6%
	0.2	37 / 3.1%	10 / 0.9%	72.6%	83.0%
	0.3	207 / 17.3%	71 / 5.9%	78.6%	87.6%
	0.4	382 / 32.0%	159 / 13.4%	84.7%	91.8%
	0.5	534 / 44.7%	255 / 21.3%	90.2%	94.6%
	0.6	655 / 54.9%	345 / 28.9%	93.4%	95.7%
	0.7	770 / 64.5%	445 / 37.3%	95.8%	97.4%
	0.8	876 / 73.4%	543 / 45.5%	94.3%	94.7%
	0.9	1013 / 84.8%	677 / 56.7%	88.5%	88.8%

V. CONCLUSIONS

Our ongoing research seeks to enhance vehicle identification in challenging conditions by integrating Automatic License Plate Recognition (ALPR) systems with Fine-grained Vehicle Classification (FGVC) methods. Preliminary findings indicate that FGVC needs further refinement to be effectively integrated with ALPR systems. This study specifically examines two approaches to enhance vehicle make classification: selective prediction and class reduction.

The results show that both approaches can improve classification accuracy. Initially, it was observed that online class reduction – where the model's output is adjusted to consider a reduced set of vehicle make classes – yields slightly worse results to training a new model with these classes (static class reduction). However, the selective prediction approach needs more careful consideration. It requires a precise balance between rejection rates and accuracy to avoid excessive rejection of correct predictions. In conclusion, combining both methods appears promising for achieving better FGVC results and an improved integration with ALPR systems.

Future research should focus on refining selective prediction techniques to reduce the rejection of correct predictions, potentially through methods like learning with rejection [31] and confidence calibration [32]. Developing an integrated system for ALPR and FGVC system will be critical to assess how enhanced vehicle make classification can improve license plate recognition in challenging real-world scenarios. Additionally, further exploring the impact of environmental conditions and image quality on classification performance will be essential.

ACKNOWLEDGMENTS

This study was financed in part by the *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES)* - Finance Code 001, and in part by the *Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq)* (# 315409/2023-1). We thank the support of NVIDIA Corporation with the donation of the Quadro RTX 8000 GPU used for this research.

REFERENCES

[1] X. Ke, G. Zeng, and W. Guo, "An ultra-fast automatic license plate recognition approach for unconstrained scenarios," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, pp. 5172–5185, 2023.

[2] R. Laroca, L. A. Zanlorensi, V. Estevam, R. Minetto, and D. Menotti, "Leveraging model fusion for improved license plate recognition," in *Iberoamerican Congress on Pattern Recognition*, Nov 2023, pp. 60–75.

[3] V. Nascimento *et al.*, "Super-resolution of license plate images using attention modules and sub-pixel convolution layers," *Computers & Graphics*, vol. 113, pp. 69–76, 2023.

[4] C. He, D. Wang, Z. Cai, J. Zeng, and F. Fu, "A vehicle matching algorithm by maximizing travel time probability based on automatic license plate recognition data," *IEEE Transactions on Intelligent Transportation Systems*, vol. 25, no. 8, pp. 9103–9114, 2024.

[5] H. Wang, J. Peng, Y. Zhao, and X. Fu, "Multi-path deep CNNs for fine-grained car recognition," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 10, p. 10484–10493, 2020.

[6] I. O. Oliveira *et al.*, "Vehicle-Rear: A new dataset to explore feature fusion for vehicle identification using convolutional neural networks," *IEEE Access*, vol. 9, pp. 101 065–101 077, 2021.

[7] S. Wolf, D. Loran, and J. Beyerer, "Knowledge-distillation-based label smoothing for fine-grained open-set vehicle recognition," in *IEEE/CVF Winter Conference on Applications of Computer Vision Workshops (WACVW)*, 2024, pp. 330–340.

[8] Y. Geifman and R. El-Yaniv, "Selective classification for deep neural networks," in *International Conference on Neural Information Processing Systems (NeurIPS)*, 2017, p. 4885–4894.

[9] R. Laroca *et al.*, "On the cross-dataset generalization in license plate recognition," in *International Conference on Computer Vision Theory and Applications (VISAPP)*, Feb 2022, pp. 166–178.

[10] R. Laroca, M. Santos, V. Estevam, E. Luz, and D. Menotti, "A first look at dataset bias in license plate recognition," in *Conference on Graphics, Patterns and Images (SIBGRAPI)*, Oct 2022, pp. 234–239.

[11] Q. Liu, Y. Liu, S.-L. Chen, T.-H. Zhang, F. Chen, and X.-C. Yin, "Improving multi-type license plate recognition via learning globally and contrastively," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–11, 2024, early Access.

[12] X. Ke and Y. Zhang, "Fine-grained vehicle type detection and recognition based on dense attention network," *Neurocomputing*, vol. 399, pp. 247–257, 2020.

[13] J. Sochor *et al.*, "BoxCars: Improving fine-grained recognition of vehicles using 3-D bounding boxes in traffic surveillance," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, p. 97–108, 2018.

[14] D. ul Khairi, F. Ayaz, N. Saeed, K. Ahsan, and S. Z. Ali, "Analysis of deep convolutional neural network models for the fine-grained classification of vehicles," *Future Transportation*, vol. 3, pp. 133–149, 2023.

[15] Y. Wang *et al.*, "Transformer based neural network for fine-grained classification of vehicle color," in *International Conference on Multimedia Information Processing and Retrieval (MIPR)*, 2021, pp. 118–124.

[16] D. Avianto, A. Harjoko, and Afiahayati, "CNN-based classification for highly similar vehicle model using multi-task learning," *Journal of Imaging*, vol. 8, no. 11, p. 293, 2022.

[17] J. He *et al.*, "TransFG: A transformer architecture for fine-grained recognition," *AAAI Conference on Artificial Intelligence*, vol. 36, no. 1, pp. 852–860, 6 2022.

[18] Y. H. Huan Ma, "Logo recognition of vehicles based on deep convolutional generative adversarial networks," *Journal of Measurements in Engineering*, vol. 12, no. 2, pp. 353–365, 2024.

[19] G. E. Lima, R. Laroca, E. Santos, E. Nascimento Jr., and D. Menotti, "Toward enhancing vehicle color recognition in adverse conditions: A dataset and benchmark," in *Conference on Graphics, Patterns and Images (SIBGRAPI)*, Sept 2024, pp. 1–6, in Press.

[20] Ultralytics, "YOLOv10," 2024, accessed: 2024-07-28. [Online]. Available: https://docs.ultralytics.com/pt/models/yolov10/

[21] R. Laroca *et al.*, "Do we train on test data? The impact of near-duplicates on license plate recognition," in *International Joint Conference on Neural Networks (IJCNN)*, June 2023, pp. 1–8.

[22] M. Celestino, https://www.webmotors.com.br/wm1/noticias/10-marcas-que-mais-venderam-carros-na-decada, 2021, accessed: 2024-07-30.

[23] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778.

[24] A. Howard *et al.*, "Searching for MobileNetV3," in *IEEE/CVF International Conference on Computer Vision (ICCV)*, 2019, pp. 1314–1324.

[25] M. Tan and Q. Le, "EfficientNetV2: Smaller models and faster training," in *International Conference on Machine Learning (ICML)*, 2021, pp. 10 096–10 106.

[26] A. Dosovitskiy *et al.*, "An image is worth 16x16 words: Transformers for image recognition at scale," in *International Conference on Learning Representations (ICLR)*, 2021, pp. 1–12.

[27] A. Hassan, M. Ali, N. M. Durrani, and M. A. Tahir, "An empirical analysis of deep learning architectures for vehicle make and model recognition," *IEEE Access*, vol. 9, pp. 91 487–91 499, 2021.

[28] X.-S. Wei, Y.-Z. Song, O. M. Aodha, J. Wu, Y. Peng, J. Tang, J. Yang, and S. Belongie, "Fine-grained image analysis with deep learning: A survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 12, pp. 8927–8948, 2022.

[29] J. Deng, W. Dong, R. Socher, L. J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2009, pp. 248–255.

[30] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: synthetic minority over-sampling technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321–357, 2002.

[31] C. Cortes, G. DeSalvo, and M. Mohri, "Learning with rejection," in *Algorithmic Learning Theory*. Springer International Publishing, 2016, pp. 67–82.

[32] C. Guo, G. Pleiss, Y. Sun, and K. Q. Weinberger, "On calibration of modern neural networks," in *International Conference on Machine Learning (ICML)*, 2017, p. 1321–1330.