engineering proceedings

MDPI

Proceeding Paper

Semantic Classification of Car Styling Using Machine Learning †

Hung-Hsiang Wang and Yen-Ting Lu *

Department of Industrial Design, National Taipei University of Technology, Taipei 10608, Taiwan; wanghh@ntut.edu.tw

* Correspondence: t112588012@ntut.org.tw

† Presented at the 2024 IEEE 7th International Conference on Knowledge Innovation and Invention, Nagoya, Japan, 16–18 August 2024.

Abstract: Product semantics is essential for car styling because it shapes how consumers perceive and interact with cars, influences user experiences, and allows for product differentiation. Although many AI tools are available to assist car designers, research on applying machine learning techniques to evaluate product semantics is rare. Therefore, we developed a classification model that helps designers identify and evaluate the semantics conveyed by car styling using the Waikato Environment for Knowledge Analysis (WEKA), a machine learning tool. We used Python web scraping to collect isometric drawings and introductory articles of 1320 SUV cars of various brands from 2009 to 2024 via websites such as Car Body Design and Car Design News. We also summarized four semantic types of car styling, namely "aggressive", "sporty", "clean", and "off-road", to create the dataset. We used WEKA image classification to randomly select 792 (60%) images from the dataset to train a classification model of car styling semantics. The remaining 528 images (40%) were used for verification. The classification model trained with the Binary Pattern Pyramid Filter and the Random Forest classifier achieved an accuracy of 84.6%. The model was evaluated in terms of whether 10 SUVs created by 10 graduate design students using AI conveyed the anticipated product semantics. Seven of the ten SUVs were correctly classified and the rest were not. All of the participants agreed that the predictions were satisfactory. However, it is necessary to improve the accuracy of each semantic classification, especially the "clean" type. The results of this study demonstrate the capability of machine learning to identify the semantics of car styling effectively, improve the communication and evaluation of product semantics by designers in the design process, and create a car styling with a good appearance that resonates with consumers.

Keywords: product semantics; car styling; industrial design; machine learning; image classification

check for updates

Academic Editors: Teen-Hang Meen, Chun-Yen Chang and Cheng-Fu Yang

Published: 24 February 2025

Citation: Wang, H.-H.; Lu, Y.-T. Semantic Classification of Car Styling Using Machine Learning. *Eng. Proc.* **2025**, *89*, 13. https://doi.org/10.3390/engproc2025089013

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Among consumers' demands for car appearance, practicality, aesthetics, style, and individuality are critical. These demands drive car manufacturers to continuously update their designs to meet the market's requirements for beauty and style [1]. Therefore, designers must grasp the semantic design language that consumers associate with different cars. However, this largely relies on the personal experience of the designers. To date, there is still limited research on how to effectively identify the product semantics of car designs. Hence, we explored related literature utilizing machine learning to analyze product semantics and accurately identify car designs.

In contemporary product design, the development of deep learning technology has empowered designers to predict consumer preferences and identify product typology through

image classification techniques. Researchers have utilized the Waikato Environment for Knowledge Analysis (WEKA) to predict the innovation threshold of product appearance [2]. This method enables designers to innovate product designs based on consumer preferences while maintaining the iconic features of the brand. With this technology, designers enhance the market appeal of products and effectively respond to market demands and consumer expectations.

Reference [3] explored the application potential of decision tree classifiers in WEKA software in architecture, particularly for the classification of cultural heritage images. The study aimed to evaluate the performance of different decision tree algorithms in handling small-scale architectural image sets. The research selected the Fast and Compact Transform-Histogram Filter (FCTH Filter), Edge Histogram Filter, and Discrete Cosine Transform (DCT) as feature extraction processing techniques and applied them to cultural heritage images such as archaeological sites, murals, and monasteries. The Random Forest (RF) algorithm exhibited significant superiority in these tasks. The effectiveness of WEKA in learning and predicting architectural styles was confirmed by demonstrating the wide applicability of this technology, providing a new technical approach for the automatic classification of architectural styles in the future.

Previous studies demonstrate the application achievements of WEKA in design, indicating that WEKA has significant development potential. From intricate product design to grand architectural modeling, WEKA provides solutions for high recognition. This finding reveals the effectiveness of machine learning technology in design innovation and opens up new possibilities for technical application in design.

Based on the previous results, we combine automotive styling and design semantics to develop a new evaluation method and enhance the judgment and cognition of automotive designers. Using this method, they can understand their design styles and provide feedback effectively through interactions with their instructors. We employed deep learning techniques to learn the semantics of automobiles, analyze the car designs of students, and map them to specific stylings. Then, a framework that supports students' creative thinking and provides a powerful teaching aid for instructors was created. Additionally, we explored the feedback from students regarding their experience with the system to assess its overall acceptance and practicality.

We categorized the collected sports utility vehicle (SUV) images into four semantic types: "aggressive", "sporty", "off-road", and "clean". WEKA machine learning was used to create a model that predicts the semantics of car styling. In automotive design, the discussion on adjectives describing styling and their semantic aspects is insufficient. Therefore, by applying machine learning technology, designers can analyze and evaluate the semantic features of car styling to provide precise semantic assessments during the design process.

2. Related Work

2.1. Semantics of Automotive Styling

In the automotive industry, aesthetics influence 60% of consumer purchasing decisions [4], demonstrating that automotive styling significantly affects buying intentions and the automotive industry at large. Therefore, automotive companies need to invest heavily in design to markedly impact market performance [5]. This underscores the critical role of automotive styling in the automotive sector. The significance of automotive semantics within semiotics cannot be overlooked. The positioning of automotive semantics in the semantic space is crucial for consumers to recognize different years and models of vehicles [6]. Even though vehicles' exteriors and technical specifications evolve, the continuity of semantics helps consumers identify and remember specific cars. Manufacturers adapt to

market changes and consumer expectations by introducing new model semantics, changing existing names, and highlighting the key role of semantics in brand management. Overall, the role of automotive semantics in brand recognition, consumer perception, and market strategy is vital for further exploration and research. The importance of automotive styling and the critical impact of automotive semantics on consumers are emphasized. The significant influence of automotive semantics in the automotive industry is revealed, especially in highly technical and innovation-driven car design.

2.2. Semantics in Automotive Design Education

In product design education, particularly in automotive styling design education, applying semantics is important in the automotive design process [7]. Teaching helps students to understand design forms, convey complex shapes, and avoid conceptual compromises using 3D software through the appropriation of complex historical semantic elements. This pedagogical approach enhances students' abilities in terms of their understanding of design aesthetics and increases the number of sketch iterations and the overall aesthetic quality of the products [7]. Reference [8] demonstrated that the principles of product semantics significantly improve students' clarity in expressing design intentions, indicating that product semantics contributes to commercial success as an essential component of the educational process. These studies revealed the profound impact of product semantics in design practice and academic research.

2.3. Image Classification Using Machine Learning

Machine learning techniques are widely used in automotive styling and consumer preference-oriented research to predict aesthetic scores for car designs and automatically generate innovative and attractive designs. By training models, different styles of car designs are analyzed and generated to help designers improve efficiency in screening designs and enhance the innovativeness and marketability of designs [9]. WEKA is used to predict the threshold of product appearance innovation [2]. Designers have used this method to innovate product designs based on consumer preferences while maintaining brand characteristics. This technology increases the marketability of products and helps designers effectively meet market demands and consumer expectations. In classifying automotive and product styling, machine learning effectively recognizes differences between brands, serving as an expert system for brand style classification [10,11]. To improve the accuracy of machine learning, machine learning is used to investigate the car front style [12].

Machine learning learns automotive and product styling using image recognition. However, few studies have explored product semantics. Thus, we explored the accuracy of machine learning in identifying automotive styling semantics to develop a model to enhance communication efficiency and creative processes for creators, students, and related departments.

3. Methods

This research was conducted in two stages. In the first stage, we built and validated a model. First, the required images of automotive designs were collected and preprocessed. Then, these images were split into training and testing sets in a 6:4 ratio. Then, the WEKA image recognition functionality was used to train the model in the training set. Finally, the model generated from the training dataset was validated using the images in the testing dataset to evaluate accuracy. In the second stage, the trained model was applied to analyze and evaluate students' automotive design works. This process provided students with feedback on their design and guided their learning progress and improvement in innovation capabilities. The model was used to explore the practicality and challenges

of machine learning technology in art design evaluation and provide a theoretical and empirical basis for future research and teaching on related technologies.

3.1. Data Collection

To conduct an in-depth image and text analysis, Python's web scraping techniques were used to collect 1320 SUV images from 2009 to 2024 (Figure 1) from two websites related to automotive design [13,14]. Adjectives related to the semantics of these images were determined. The semantics associated with each image were derived from car photos and articles on the websites and were filtered according to the advice of two experts. The finalized keywords included "aggressive", "sporty", "clean", and "off-road".



Figure 1. A total of 1320 SUV car images obtained from Car Body Design and Car Design News (placed the remaining images on the cloud drive).

3.2. Data Preprocessing

We processed the images to reduce the background interference and eliminate the effect of color differences on the analysis results. We utilized Python to preprocess the images. We applied the back removal process to improve the focus clarity and enhance the classification accuracy. The pixels of the images were set between 580 and 1100. Pixel normalization was conducted to ensure that all the images had the same resolution and size, thus enabling the algorithm to process and analyze them accurately. To minimize the interference of color image recognition, the images of cars with white bodies were purposely chosen for this study to minimize the interference of color. The surface property of the white color presents the details of the car body shadow clearly, which is important for highlighting the vehicle's shape and styling features. Finally, the scaling and deformation of the images were performed to increase data diversity, reduce over-simulation, simulate realistic conditions, and improve the robustness of the model. The process ensured the

quality of the data and the accuracy and reliability of the subsequent analysis and the analysis of the visual features of the car.

3.3. Feature Extraction

We used WEKA machine learning software for image feature extraction and classification. We extracted features from the preprocessed images and tested four different image classifiers, namely the Binary Pattern Pyramid Filter (BPPF), Edge Histogram Filter, FCTH Filter, and Gabor Filter, which focus on the texture and feature analysis of images. In the experiment, 1320 images were used, of which 60 and 40% were allocated to the training and test datasets. According to the machine learning results of the WEKA software, BPPF showed superiority in image classification, with a classification accuracy of 85.8%. The high performance of BPPF in processing image texture features provides a reference for future research (Figure 2).

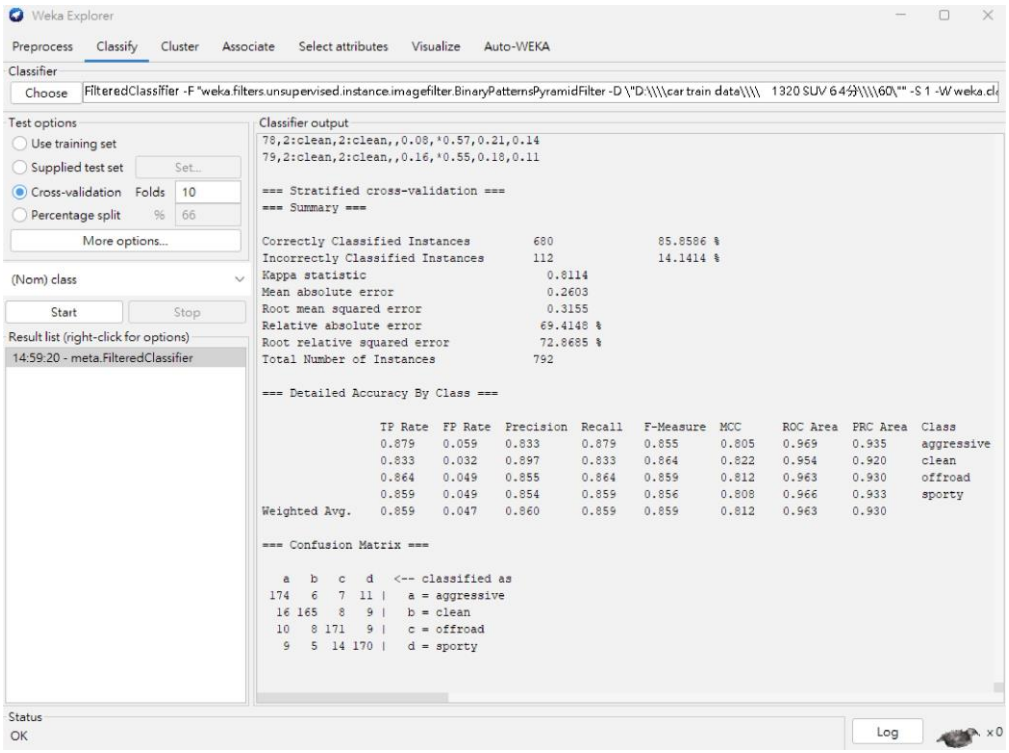


Figure 2. Prediction results of the style semantic model using WEKA.

3.4. Verification and Testing

To validate the generalization ability of the model, 40% of the 1320 images were allocated to the test dataset. We used 528 SUV images. These images and text descriptions were preprocessed for testing. Then, the 528 images were classified into styles using BPPF (Figure 3). A total of 447 images were correctly recognized, with an accuracy rate of 84.6% in the test dataset. The accuracy and effectiveness of the BPPF classifier were validated in recognizing the semantics of visual styles.

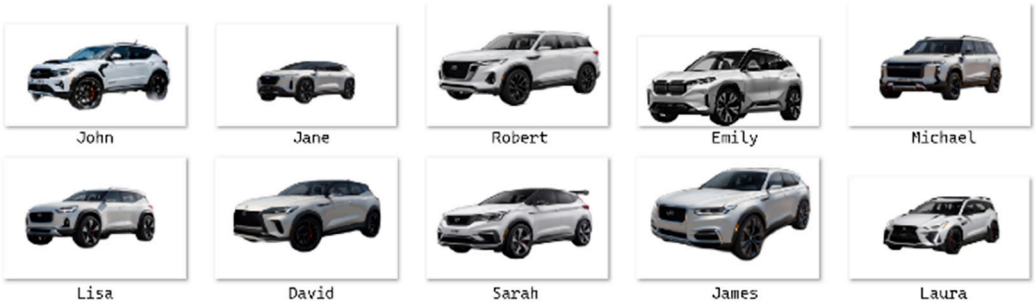


Figure 3. Car designs created by automotive design students.

3.5. Application and Evaluation

We applied the trained model to evaluate the SUV designs of 10 automotive design students using AI (Figure 3). The WEKA-trained automotive semantic model automatically categorized the style semantics of these students' works. Based on the categorization results, the students' designs were evaluated as different automotive semantic styles (Table 1, Figure 4). The results reflected the stylistic diversity of the students' designs and demonstrated the model's ability to identify and evaluate different semantic styles. Feedback on the categorization results was collected and summarized to verify the model's usefulness and applicability.

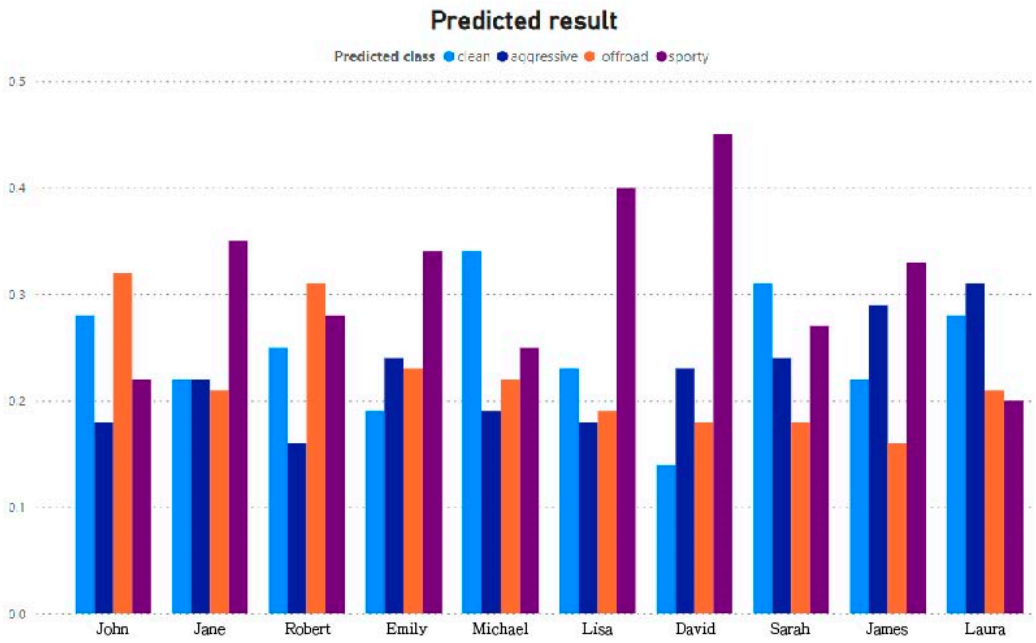


Figure 4. Prediction of semantics of car design works through WEKA and visualization with bar charts.

Table 1. Predicting semantics of car design works using WEKA.

Student	Aggressive	Clean	Off-Road	Sporty
John	0.18	0.28	**0.32**	0.22
Jane	0.22	0.22	0.21	**0.35**
Robert	0.16	0.25	**0.31**	0.28
Emily	0.24	0.19	0.23	**0.34**
Michael	0.19	**0.34**	0.22	0.25
Lisa	0.18	0.23	0.19	**0.40**

Table 1. *Cont.*

Student	Aggressive	Clean	Off-Road	Sporty
David	0.23	0.14	0.18	**0.45**
Sarah	0.24	**0.31**	0.18	0.27
James	0.29	0.22	0.16	**0.33**
Laura	**0.31**	0.28	0.21	0.20

The bolded values indicate the proportion of each car image classified under a particular style. The highest value is shown as Weka's final classification result.

We collected feedback from the students, including comments on the categorization and perceptions of their designs. Semi-structured interviews were conducted with 10 graduate students aged 18−24 years old with a background in automotive design. The following quotations were obtained from their feedback.

John: "*I designed an off-road and sporty SUV, but the model categorized it as off-road and partly sporty and simple. I think the model categorization is quite close to the styling style that I wanted to achieve, only that the simplicity component is higher than what I wanted to achieve*".

Jane: "*I think this car is appropriately categorized as 'sporty', but as far as 'off-road' is concerned, I think only the width of the car matches this category. As for the rating of 0.22 for 'clean', I don't think it's quite accurate because the surface design is not particularly clean. Lastly, I think the rating of "aggressive" is similar to Weka's evaluation, and it is quite appropriate*".

Robert: "*I think it is reasonable for my car to be categorized as 0ff-road by WEKA, including its clean and sporty data, which are very similar to what I wanted to design in the beginning." Student 4: "I think WEKA has a good idea to categorize my car as 0ff-road*".

Emily: "*I think WEKA's judgment of sporty is very much in line with my creative concept, except that I think the semantic meaning of clean, off-road should be a little lower, and should not add up to more than 0.5, and the judgment of aggressive as 0.24 is quite in line with my thoughts*".

Michael: "*The design direction of this car is off-road, but clean is the highest, although he has clean appearance line, but I think off-road will be the highest right, the other sporty and aggressive I am not too satisfied*".

Lisa: "*The sporty is too much higher than I expected, I think it's about 0.25, off-road and clean together is more than 0.65, and aggressive can be reduced to 0.1, which will be better*".

David: "*This car's predicted solution is very sporty and aggressive as I thought it would be, but I think the clean could be a little lower to 0.1 and the off-road to 0.22*".

Sarah: "*I think it's quite inaccurate to classify this car as clean, it's more of an aggressive car, and the sporty value is quite consistent. The sporty value is quite consistent, and off-road can be lowered to 0.3, which would be better*".

James: "*I think it's quite reasonable to classify this car as sporty, I think off-road needs to be higher. aggressive needs to be much lower. Besides, I think it's quite appropriate that clean is not classified*".

Laura: "*I think it is very appropriate for this car to be classified as aggressive by WEKA, but the clean part can be lowered to 0.2, the off-road can be lowered to 0.3, and the sporty part can be raised even more*".

The observations below are based on the comparison of each student's self-assessment of their design with the results of the WEKA model categorization.

3.5.1. Model Categorization

John felt that the model rated the "sportiness" and "simplicity" of his design higher, but he emphasized the off-road nature more. There was a difference between the students' understanding of "simplicity" and the model's. Jane thought that the width of the car was consistent with the "off-road" feature, but was surprised by the low rating of "clean" as the surface of the design was not simple enough. There was an interpretive difference in

the definition of "clean" between the students and the model. Robert was satisfied with the scores of "off-road", "clean", and "sporty" given by the model, which indicated a high degree of agreement between his design intention and the model's understanding.

Emily was satisfied with the model's "sporty" rating and wondered why "clean" and "off-road" were rated lower, suggesting that less emphasis was placed on these elements than on the model's assessment. Michael, Lisa, and David expressed their dissatisfaction with the "clean" rating, which they felt did not match the off-road style of their design, highlighting the need for further clarification on how to balance design elements in teaching. Sarah thought that "clean" was over-rated and expected an "aggressive" style, which indicated that the students had a clearer perception of the aggressive features of the vehicle design. James and Laura were satisfied with or accepted the ratings of "sporty" and "aggressive" but were dissatisfied with the ratings of other elements such as "off-road", showing differences in the expectations of particular design elements.

3.5.2. Strengths and Weaknesses of WEKA Model

1. Participation and cognitive reflection

The students described their views on their car designs in detail, including their understanding of various design elements such as "sporty", "clean", "off-road", and "aggressive". Initiative and self-reflection are important in the design process as they prompt students to think about design intention and realization and make self-assessments and adjustments to the design.

2. Consistency

The WEKA model accurately reflected the students' design intentions, such as in the case of Robert, whose model ratings for "off-road", "clean", and "sporty" were highly consistent with the design objectives. In addition, the high rating of "sporty" for Emily showed the accuracy of the model in evaluating this design element. Laura's rating of "aggressive" showed that the model reasonably assessed the design. The model was validated in understanding and evaluating design elements and its ability to match the students' perceptions, which strengthens the students' trust in the use of the model and demonstrates the practical value of the model in real-world applications.

3. Inconsistency

Discrepancies between students' understandings of particular design elements and the model's ratings can lead to misunderstandings in the design intentions. For example, John, Jane, David, and Sarah were dissatisfied with the "clean" scale, indicating that there was an inconsistency between the students and the model in terms of the meaning of this term. This also suggests that "clean" is the main direction for the subsequent optimization of the training model.

4. Advantages of WEKA model

The WEKA model evaluated design elements to help students understand and master abstract concepts in design theories and apply these theories in practice. The students' feedback showed that the model effectively reflected the students' design intentions by evaluating design elements such as "sporty" and "aggressive" with consistency and accuracy. This enhanced the students' understanding of design and facilitated their self-assessment and improvement of the design effects. In addition, there were discrepancies between the model's scores and the students' self-assessments. Based on these, teachers can guide students to have in-depth discussions on the relationship between design concepts and results to understand design elements and the perception of design. This interactive learning process enhances students' critical thinking skills and strengthens their problem-solving skills.

The advantage of the WEKA model in design education is that it provides a real-time, specific, and accurate design assessment, which significantly impacts students' learning efficiency and design ability.

4. Conclusions

In this study, the WEKA model was established to categorize the design semantics of SUV cars for application in automotive design education. A high accuracy of 84.6% was achieved after analyzing 1320 SUV model images and model training, which demonstrates the ability of machine learning technology to recognize complex design semantics. In addition, the evaluation and application of the technology in students' design work enhanced their in-depth understanding of the design semantics and facilitated their effective communication and feedback on their work. The results emphasize the importance of digital tools in enhancing educational interactivity and student creativity and demonstrate the potential for broader application of technology in professional design education and practice.

A total of 7 out of 10 students recognized that the model helped them to better understand and apply the design concepts, and to evaluate and adjust their designs. However, there were challenges, especially in the ambiguity between students' perceptions and the model's assessment for several design elements such as "clean". The model needs to be optimized and adjusted in these aspects. When the model's assessment was inconsistent with the student's self-assessment, the students were prompted to think deeply, which provided the teacher with opportunities to guide discussion and critical thinking training. The WEKA model has shown pedagogical potential in design education and the potential to develop students' design skills and creativity. Future improvements are needed to improve the assessment accuracy and semantic richness of the model to more accurately distinguish semantic meanings to increase applicability.

Machine learning technology can be used to identify the semantics of automotive design. Through technological innovation and practical application, more innovation and progress can be achieved in design in the future.

Author Contributions: Conceptualization, H.-H.W.; method, H.-H.W.; resources, Y.-T.L.; data curation, Y.-T.L.; validation, H.-H.W.; writing—original draft preparation, Y.-T.L.; writing—review and editing, H.-H.W.; experiments, Y.-T.L.; supervision, H.-H.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by NSTC, Taiwan, R.O.C., through grant 111-2410-H-027-019- MY2.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are unavailable due to privacy reasons.

Acknowledgments: The authors gratefully acknowledge the financial support of NSTC, Taiwan, R.O.C., through grant 111-2410-H-027-019-MY2.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Gartman, D. *Auto-Opium: A Social History of American Automobile Design*; Routledge: London, UK, 2013.
2. Wang, H.-H.; Liu, C.-K.; Yu, S.-B. Predicting Consumer Preferences by the Deformation Threshold of Product Appearance. *Eng. Proc.* **2023**, *55*, 67. [CrossRef]
3. Jankovic, R. Classifying cultural heritage images by using decision tree classifiers in WEKA. In Proceedings of the 1st International Workshop on Visual Pattern Extraction and Recognition for Cultural Heritage Understanding Co-Located with 15th Italian Research Conference on Digital Libraries (IRCDL 2019), Pisa, Italy, 30 January 2019; pp. 119–127.

4. Kreuzbauer, R.; Malter, A.J. Embodied cognition and new product design: Changing product form to influence brand categorization. *J. Prod. Innov. Manag.* **2005**, *22*, 165–176. [CrossRef]
5. Rubera, G. Design innovativeness and product sales' evolution. *Mark. Sci.* **2015**, *34*, 98–115. [CrossRef]
6. Aronoff, M. Automobile semantics. *Linguist. Inq.* **1981**, *12*, 329–347.
7. Berry, J. Overcoming the complexity of teaching form and detail aesthetics in product design education. In Proceedings of the DS 46: Proceedings of E&PDE 2008, the 10th International Conference on Engineering and Product Design Education, Barcelona, Spain, 4–5 September 2008; pp. 156–161.
8. Huang, C. Design for Ease of Use: Product Semantics and Design Education. Doctoral Dissertation, Manchester Metropolitan University, Manchester, UK, 1996.
9. Burnap, A.; Hauser, J.R.; Timoshenko, A. Product aesthetic design: A machine learning augmentation. *Mark. Sci.* **2023**, *42*, 1029–1056. [CrossRef]
10. Wang, H.-H.; Lin, Y.-Y.; Huang, H.-T. Application of Typicality in Predicting Product Appearance. *Eng. Proc.* **2023**, *55*, 66. [CrossRef]
11. Wang, H.H.; Chen, C.P. STYRULE-A Classification System of Car Styling. *J. Des.* **2009**, *9*, 107–121.
12. Ma, L.; Wu, Y.; Li, Q.; Yuan, X. Recognition of car front facing style for machine-learning data annotation: A quantitative approach. *Symmetry* **2022**, *14*, 1181. [CrossRef]
13. Car Body Design. Available online: https://www.carbodydesign.com/ (accessed on 1 May 2024).
14. Car Design News. Available online: https://www.cardesignnews.com/ (accessed on 1 May 2024).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.