Computers in Industry 164 (2025) 104168

Contents lists available at ScienceDirect

Computers in Industry

journal homepage: www.sciencedirect.com/journal/computers-in-industry

Computers as co-creative assistants. A comparative study on the use of text-to-image AI models for computer aided conceptual design

Jorge Alcaide-Marzal *, Jose Antonio Diego-Mas

Human-Centered Technology Research Institute (Human-Tech), Universitat Politècnica de València, Camino de Vera s/n, Valencia 46022, Spain

ARTICLE INFO

Keywords:
Computer Aided Conceptual Design
Artificial intelligence models
Computational creativity
Human-Computer Interaction
Product Design

ABSTRACT

This preliminary research presents a comparative study between Text-to-Image AI models and Shape Grammars, one of the main generative approaches to Computer Aided Conceptual Design. The goal is to determine to which extent AI models can reproduce or complement the performance of grammar algorithms as creative support tools for shape exploration in conceptual product design. Workflows, advantages and limitations are identified through a comprehensive practical comparison example. The results show many similarities regarding generative capabilities and highlight several advantages of Text-to-Image AI models, including an easier way of capturing product grammars and a wider and more immediate range of further applications. In contrast, Shape Grammars approach proved more solid in aspects related to exploration workflows and cognitive stimulation. These results encourage the research on new ways to address the interaction between designers and AI generative models, combining the AI potential with well-established generative strategies.

1. Introduction

Computers have been used as support tools for product design for more than 90 years. However, it is well known that the impact of this use throughout the design process has not been the same in every stage. Early CAD tools were focused on facilitating product representation and analysis, thus being very suitable for embodiment and detail stages, but not for the conceptual one, in which design problems are still ill-defined and the information available is ambiguous (Mothersill and Bove, 2018). The obstacles and requirements to adapt digital tools to early stages of the design process have been thoroughly studied (Abdalla et al., 2021; Bernal et al., 2015; Bonnardel and Zenasni, 2010; Lubart, 2005; van Dijk, 1995). These studies point out the need for knowledge-based computer support systems, focused on cognitive aspects rather than modelling tasks.

In fact, the incorporation of AI-powered systems has been considered since the early days of CAD, but then the available technology could not provide suitable support to theoretical frameworks (Forbus, 1988; Jiaoying et al., 1987; MacCallum, 1990). As technology has evolved, traditional rigid CAD systems have shifted to more flexible tools and computational systems capable of providing support in tasks all along the whole design process (Bernal et al., 2015). A wide range of studies about generative tools for conceptual design have been published (Mountstephens and Teo, 2020), although there exists still a significant gap between academic approaches and industrial applications (Horváth, 2000).

Among them, AI generative models have experienced a highly dynamic development in several fields these recent years (Yüksel et al., 2023). Particularly, Text-to-Image AI models constituted a shocking release along 2022, radically changing graphic design and visual disciplines (Oppenlaender, 2022). Although it is possible to find some examples of their application to product design by freelance pioneer designers and firms, there are not many studies on this issue yet.

This work constitutes a preliminary study to evaluate the possible ways in which product designers and other practitioners may make use of these generative technologies to empower their shape exploring creative processes in conceptual design stages. Text-to Image AI models are a very recent technology, and therefore scarcely studied from an academic point of view and barely implemented in industry. The few (but valuable) studies that exist in this regard have focused mainly on analysing the creative performance of users when using the tool as a source of inspiration as a support for formal exploration.

The present work proposes a different approach and is part of a series of studies in which the incorporation of generative models as an integrated part of existing design methods is analysed. Specifically, in this paper, we explore the workflow and performance of a Text-to-Image AI

* Corresponding author.
E-mail address: jalcaide@dpi.upv.es (J. Alcaide-Marzal).

https://doi.org/10.1016/j.compind.2024.104168
Received 30 April 2024; Received in revised form 15 July 2024; Accepted 30 August 2024
Available online 7 September 2024

model (Stable Diffusion) as compared to that of Shape Grammars in a particular case study. The aim is to identify through a practical application to what extent the AI model can be used to perform similar exploratory design tasks, and how designers could adapt their work to include these generative procedures in their design toolbox. To our knowledge, there are no direct applications of product Shape Grammar strategies applied through generative AI models till now. This approach intends to propose new research lines based on incorporating AI agents to existing and sound methods in design practice.

Using as a comparison an exhaustive previous study based on the analysis of the Buick grammar along more than 60 years, this work initially calibrates the ability of the generative artificial intelligence model to reproduce the grammar formal elements, it is, the geometric shapes conforming the different car parts. The Buick study is then taken as an example and all the creative processes based on Shape Grammars shown in the original paper are replicated using Stable Diffusion. The aim is to determine whether a designer skilled in the use of Shape Grammars could use Stable Diffusion as a tool to support the formal exploration similarly. In other words, whether it is possible to reproduce the exploratory creative processes enabled by Shape Grammars through generative artificial intelligence models.

This work contributes to the field of Computer Aided Conceptual Design by providing a comprehensive comparison of technical workflows between Shape Grammars as a generative paradigm for creativity and Text-to-Image generative models. This kind of comparisons is needed to fully understand the benefits that AI models may contribute to complement, enhance or replace current creative workflows. As aforementioned, while already exist some research papers on the use of these models for concept ideation, their approach is mainly based on using image generation for inspiration. Other different perspectives must be studied to effectively deploy the full potential of this technology as true computer design assistants along the whole design process. In this sense, this work also offers a first test of different uses for tools of a Text-to-Image AI model as resources for creative exploration through an existing procedure (Shape Grammars), thus enabling different exploratory strategies to make the use of AI models more goal focused.

The rest of the paper is organised as follows. Section 2 presents a brief evolution of the generative tools applied in conceptual design, specially highlighting the recent development of generative artificial intelligence models focused on producing images from text inputs. Section 3 describes the method followed to conduct the exploratory study. Section 4 comprehensively describes the application of Text-to-Image techniques to replicate/complement the Shape Grammar exploration process. The most relevant results, the limitations of the study and possible future lines are discussed in Section 5. The use of Stable Diffusion as a tool to explore formal grammar in the case study has proved very satisfying. The possibility of using a model that reproduces with acceptable fidelity characteristics from a wide range of historic eras, the ease of using these characteristics in the study of new product concepts and the ability to combine characteristics efficiently are highlighted. The main weakness detected is the difficulty in ensuring concept generation within the solution space defined by the product's grammar. Finally, Section 6 provides some concluding reflections.

2. Generative tools for Computer Aided Conceptual Design

2.1. The Evolution of CACD Tools

The conceptual stage of the design process is intensively creative (Dorst and Cross, 2001; J. Gero, 1996; Sarkar and Chakrabarti, 2011). Even though the cognitive processes that take part in this stage have been deeply studied, there is still a need for a clear comprehension of the phenomenon in order to provide effective computer tools to designers (Dinar et al., 2015; Jin and Benami, 2010). Traditionally, the complex nature of these processes has imposed a barrier to computer tools supporting design activities (Vuletic et al., 2018).

An evolution of the implementation of computer tools in different stages of the design process can be found in (Chandrasegaran et al., 2013). In the early days of CAD, these tools focused primarily on creating and manipulating information about 2D and 3D shapes. This allowed CAD/CAM/CAE systems to be easily incorporated into the embodiment/detail design stage. However, the rigidness of use and the need for objective and precise information of these early CAD systems posed a considerable obstacle to their use in the conceptual design stage (Company et al., 2009; Mothersill and Bove, 2018).

Conceptual tasks needed, on the contrary, computer tools able to support creative processes (Woodbury, 1990; Tay and Gu, 2002). In fact, 30 years ago, studies on the application of artificial intelligence to the understanding of design processes were already being proposed (J. S. Gero and Maher, 1993). Since these first studies, many computational models representing design activities have been suggested (J. S. Gero, 2000; Goldberg, 1991; Maher and Tang, 2003; Mekern et al., 2019; Pineda, 1993; Tay and Gu, 2002). This view helped to develop early digital tools in which computers were intended to play a more relevant role than just geometric representation support systems, such as Sketchpad (Sutherland, 1964), the Electronic Cocktail Napkin (Gross, 1996), the computer-aided design conversation system proposed by (Lawson and Loke, 1997), or SketchREAD (Alvarado and Davis, 2007). Limited by the available technology, these pioneer proposals were then scarce and not very widespread.

The studies on computer systems suitable to support conceptual design evolved as technological improvements allowed for more flexible digital tools. Progressively, the introduction of more complex systems allowed the incorporation of cognitive design assistants (Huet et al., 2021), the practical implementation of algorithmic approaches (Ekströmer and Wever, 2019) and the management of data driven product design (Briard et al., 2023). The term Generative Design is often associated with this shift of paradigm from computer as a tool to computer as an agent or a design assistant (Tufarelli and Cianfanelli, 2022). An important aspect of generative systems is the ability to produce and evaluate a wide range of design options. Algorithms allow the automatic production of multiple design alternatives based on specific criteria and constraints. These systems can explore a vast design space and provide designers with a multitude of options to consider, allowing for more creative solutions and overcoming some of the creativity blockers that designers suffer, such as fixation. (Crilly, 2015; Jansson and Smith, 1991). The possibility of procedurally generating multiple and varied solutions allows designers to explore alternatives beyond the fixation area, favouring emergence (Alcaide-Marzal et al., 2020; Hyun and Lee, 2018; Karimi et al., 2020; Yüksel et al., 2023).

Thus, during the last two decades, researchers have proposed and analysed the application of many different generative solutions for shape exploration focused on conceptual product design. A thorough review of these generative tools is conducted in (Mountstephens and Teo, 2020). The authors analyse 37 generative product systems based on Shape Grammars, L-systems, Genetic Algorithms, Swarm Intelligence, Parametric CAD and GANs.

Shape Grammars was one of the first generative approaches to product design exploration. Proposed in 1972 by Stiny and Gips (Stiny and Gips, 1972), Shape Grammars are rule-based systems used in computational design that use formal grammar rules to generate or modify geometric shapes. A Shape Grammar system consists of a set of transformation rules, which define how a shape can be transformed or modified, and a generation procedure, which selects and applies these transformations. Starting from an initial shape, transformation rules are applied to generate new shapes. By recursively applying these rules, designers explore the solution space generating concept variations. An example of this process is shown in Fig. 1.

Shape Grammars have been commonly used in computer-aided design (CAD), architecture, and urban planning to create designs and layouts following a set of predefined rules, but also in product design (Agarwal and Cagan, 1996; Barros et al., 2015; Cui and Tang, 2013;

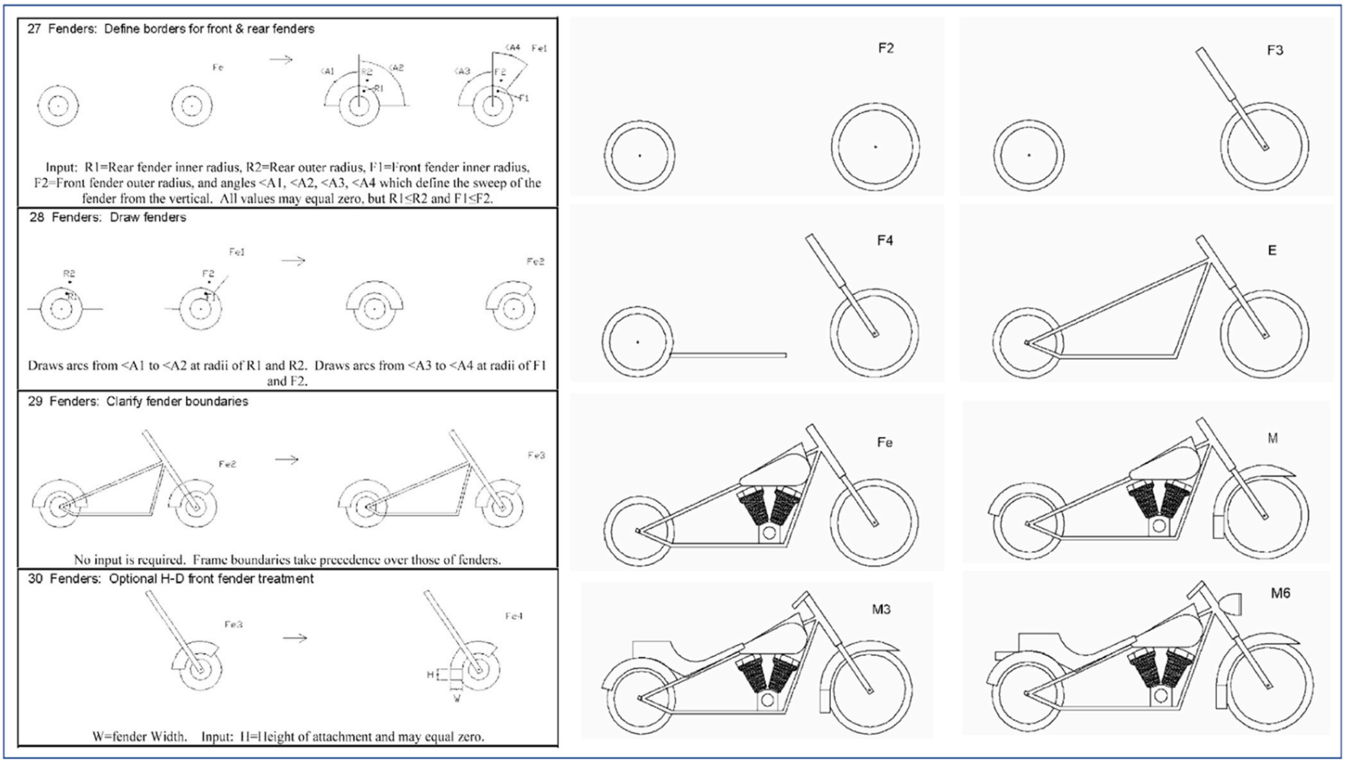

Fig. 1. An example of definition and application of Shape Grammars. The Harley-Davidson grammar (Pugliese and Cagan, 2002). At the left of the picture, some of the defined rules to generate fender shapes. At the right, successive rules are applied to produce a motorbike design within the Harley brand grammar.

Kielarova et al., 2013; Orsborn et al., 2006; Pugliese and Cagan, 2002), being one of their strengths the capability of identifying relevant visual features to convey a brand image or a specific iconic style (McCormack et al., 2004; McKay et al., 2006).

Despite being a powerful framework for conceptual shape exploration, Shape Grammars present several limitations that have prevented a wider development. (Gu and Behbahani, 2021) summarise some of them: Disconnection between the approach of abstract academic research and that of practical professional applications, lack of commercial packages or presence in CAD platforms and low flexibility to adapt to designers' workflows. Nevertheless, Shape Grammars continue being a fruitful field of research and a basic paradigm for other generative variants (Eloy et al., 2018).

Among generative methods, GANs and other related AI algorithms have received increasing attention in recent years, and many studies have been published describing different AI applications to numerous and varied design situations (Chiarello et al., 2021). A thorough compilation of these applications can be found in (Yüksel et al., 2023). The relevance of this field of research has led some authors to use specifically the term "artificial intelligence aided conceptual design" (Isgrò et al., 2022; Xin and Zhao, 2021; Yang et al., 2023).

More recently, AI generative models focused on visual arts such as automatic image or video generation have experimented a breath-taking development (Oppenlaender, 2022). This trend has been further powered by the rising and public release of several Text-to-Image AI models (algorithms able to generate high quality images from text inputs), enabling a radical change and probably introducing a new paradigm which will affect all creative disciplines, including product design. The next section focuses on this question.

2.2. Text-to-Image AI models and current applications in Product Design

The evolution of image generative models has experienced significant advancements over the last decade, especially with the advent of Generative Adversarial Networks (GANs). Previously, image generation was addressed using non-Deep Learning models such as Gaussian Mixture Models (GMMs) and Hidden Markov Models (HMMs) (Permuter et al., 2006). Later, Variational Autoencoders (VAEs) introduced a probabilistic framework combining deep neural networks and variational inference (Elasri et al., 2022).

GANs were proposed by (Goodfellow et al., 2014), and revolutionised image generation algorithms. GANs consist of a generator and a discriminator network, trained adversarially. The generator produces images to fool the discriminator, while the discriminator tries to distinguish real from fake images. GANs achieved remarkable results in generating realistic images and opened the door to numerous applications. such as BigGAN (Brock et al., 2018), StyleGAN (Karras et al., 2018) or GauGAN (Park et al., 2019). (Zhou et al., 2021) provide a comprehensive review of GAN-based Text-to-Image algorithms. GANs are of course not only applied to image generation, and they have been used in different approaches to conceptual product design tasks, as aforementioned (Yüksel et al., 2023).

The introduction of Denoising Diffusion Probabilistic Models (Ho et al., 2020; Sohl-Dickstein et al., 2015), or simply diffusion models, was another crucial step in the development of image generation AI models. They proved more effective than GANs for image synthesis (Dhariwal and Nichol, 2021), and were further improved by the use of latent spaces (Rombach et al., 2021).

Diffusion models became very popular by 2022, when some of the now most used AI models were made public. DALL-E2 (openai.com, n.d.), Imagen (Saharia et al., 2022), Midjourney (midjourney.com, n.d.), and Stable Diffusion (stability.ai, n.d.). Open-source models (Stable Diffusion and Craiyon, a mini version of DALL-E) have fostered the development of third-party integrations, and rapidly many 2D graphic software and 3D systems incorporated AI tools through plugins.

Although it is possible to find several examples of firms, practitioners and freelance designers already testing or using Text-to-Image AI models in their current workflow, the applications of this technology to product design are not very numerous yet in the research literature. An approximation to their use in different creative disciplines is provided by (Ko et al., 2022). A particular review for applications in architecture can be found in (Castro Pena et al., 2021).

One of the pioneer industries adopting AI image models for concept exploration is fashion design. (Zhu et al., 2017) demonstrate a tool which generates a new outfit given an input image of a person and a sentence describing the new appearance, while keeping the initial pose. Deepwear, a system using GANs for clothes design, was proposed by (Kato et al., 2018). Similarly, (Ak et al., 2020) proposed e-AttnGAN. (Jeon et al., 2021) developed an AI-based creativity support system capable of mixing styles and trends called FashionQ. In (H. Liu et al., 2023), a Sketch-to-Image algorithm is proposed and tested on several examples, such as shoes and handbags. A thorough review of AI techniques applied to fashion design can be found in (Mohammadi and Kalhor, 2021).

To a lesser extent, researchers started to use GAN-based image generators for preliminary product exploration. In (Radhakrishnan et al., 2018), a system was proposed to produce aesthetic solutions for car design from sketches provided by designers. The GAN was trained to interpret the strokes of sketches and generate a final image with different colours and perspectives. (Heyrani Nobari et al., 2021) propose CreativeGAN, an adaptation of GANs specifically oriented to product design. A case study based on bicycle design is demonstrated.

Some works using new diffusion models for product design have also been already proposed. In (Tholander and Jonsson, 2023), a combination of GPT-3 and DALL-E is used during a workshop to obtain and test ideas for a design task. DALL-E is used for concept visualisation. (Chiou et al., 2023) used Midjourney as a creativity support tool for design tasks. Three kinds of images (concept, scenario and form) are detected as source of inspiration at different levels.

A very interesting experiment, closely related to industrial design practice, was proposed by (V. Liu et al., 2022). The researchers integrated DALL-E into Fusion 360, thus providing it with an idea assistant. The application was tested among 13 designers, who performed two design tasks using an implementation of text-to-image algorithms in Fusion 360 CAD package. A module connecting prompting to DALL-E with Fusion 360 environment facilitates the use of render images as initial inputs for DALL-E, as well as product image generation. A similar experiment in the field of architecture. but comparing Midjourney, DALL-E and Stable Diffusion, is presented in (Paananen et al., 2023).

Given the enormous potential of Text-to-Image technologies as creativity support systems for product shape exploration, there is still a lack of studies in this field. The immediate utility of a tool capable of generating on demand images is obvious, but just generating images, even if they achieve high quality standards, may not fit the required creative level (Basalla et al., 2022). It is the “beyond average” approach described in (Mothersill and Bove, 2019). We need to investigate different uses others than this immediate one, considering all that we already know about conceptual design and the cognitive processes involved. For instance, in (Padiyath and Magerko, 2021) a tool called desAIner is proposed. It is a creativity support system trained with images of high fashion designer to purposedly generate bad compositions, in order to stimulate creativity through ambiguity. This kind of potential uses of these technologies, how they relate to existing approaches and how they affect creativity performance of designers are to be studied.

3. Material and methods

The goal of this research is to identify, through a comparative analysis, procedures to use Text-to-Image AI models for shape exploration in the conceptual design stage. It aims to determine if the use of one of these models can reproduce or complement the performance of an existing generative approach.

As aforementioned, the Shape Grammar approach is a very powerful instrument to capture the product visual language and to allow the use of computer support of concept design. However, it presents several drawbacks that make its use complex and hard to generalise. We analyse in this study the use of AI models as a complement or supporting tool for some of the tasks performed by Shape Grammar procedures.

The AI model used in this paper is Stable Diffusion. The reasons for choosing this model over all the ones available is twofold: it is open source, and it can be installed locally, which frees the user from depending on remote servers to operate. It also enables other possibilities, such as locally training the model, which is very valuable to make it work in a more specific context, although we have not used this option in the present work.

The behaviour of Stable Diffusion generations depends on the specific model (called checkpoint) chosen to work with. Checkpoints are the files containing trained Stable Diffusion weights. Stable Diffusion has received several official version checkpoints, from V1.x versions to SDXL and SD V3, which are general-purpose models. Based on them, many other unofficial ones have been released, trained with specific datasets to generate a particular style or kind of images. Checkpoints used in this study are Stable Diffusion V1.5, Stable Diffusion V1.5 Inpainting and Deliberate V2.0, a variation of V1.5 presenting a more realistic tuning.

The use of Stable Diffusion has involved the use of the following options:

a) Text-to-image: The basic module where Stable Diffusion is given the prompt or textual description of what is to be represented. Generally, we have used very simple and concise prompts. More elaborated prompts (with terms such as “digital rendering, vehicle illustration, concept art, car design”) were used in subsection 3.2.2. to allow for variability and different styles for solutions. Prompting is affected by several parameters, but typically Classifier Free Guidance (CFG) Scale, which controls to which extent the generated image must conform to the prompt.
b) Image-to-image (Img2Img): this algorithm uses an initial image and produce variations according to the given prompt. We will refer to these images as “starting images”, as they are used as a starting point in the process. It is possible to define which part of the image will be affected by the generation by using a mask tool called Inpaint. We used Inpaint for guiding Stable Diffusion in the task of emulating Shape Grammar iterative process. Once one or some features were achieved, inpainting facilitated altering only specific areas. Img2Img is affected by several parameters, but mainly CFG Scale and Denoising (which controls how much the generated image will respect the original one).
c) ControlNet (Zhang and Agrawala, 2023) uses an image to provide a structure which Stable Diffusion will use when generating an image according to the prompt. We will call these images “conditioning images”. We have used them to force pure front views and to look for specific features. The effect of ControlNet may be modulated by several parameters, such as the Control Weight, which determines how much the generated image will be restricted by the content of the conditioning one.

The case study has been structured according to the original paper and the Shape Grammar workflow. The first step was the identification of the corresponding grammar contained in the Stable Diffusion model. As we intended to use a generic checkpoint with no specific training, we needed to verify that Stable Diffusion could represent adequately all Buick grammar elements. Secondly, we used Stable Diffusion to generate solutions similar to those of the original paper. Finally, we tested its specific capabilities for concept generation and shape exploration.

4. Case study: The Buick Language

A thorough case of shape grammar application is described in (McCormack et al., 2004), focused on Buick front views as a way to convey brand image. The authors examine Buick front view geometries for 13 thematic eras ranging from 1939 to 2002 plus concept designs, and then extract the Buick grammar from them. This comprehensive grammar is validated reproducing existing Buick models, and then used as a creative exploratory tool to generate new designs. The authors point out some further possibilities such as ideating new Buick concepts (cross-over vehicles, for instance) maintaining the brand essence. Their study demonstrates the power of shape grammars to capture the essence of a product visual communication. By extracting and making operable the representative visual characteristics of Buick brand, designers may use them to explore new designs conveying that brand.

In the present case study, we carry out a comparative analysis between this process and the workflow needed to obtain similar results using Stable Diffusion. The objective is to examine how generative text-to-image models can be used in a similar framework to that of other exploratory processes, in this case the use of Shape Grammars for the generation of concepts while maintaining a brand image. After performing this comparison, other possible paths of formal exploration enabled by Stable Diffusion are analysed. An outline of the process is shown in Fig. 2.

4.1. Extraction of the Buick grammar

The authors in (McCormack et al., 2004) conducted a thorough and interesting study of the evolution of Buick designs from the 1930's to 2002 to capture significant visual features related to that brand. Even though the paper does not describe the whole grammar, the authors provide many different rules for relevant features (Fig. 3), as well as some valuable sketches and verbal description. Specifically, they describe in the paper 13 rules for the construction of the grill, 3 rules for the emblem, 3 rules for the middle hood, 9 rules for the center hood, 7 rules for the outer hood, 5 rules for the fender, 5 rules for the hood flow line, one rule for the roof, 3 rules for the headlights and a set of additional rules for the adjustment of shapes, up to a total of 63 rules.

Using this information, we have represented the main features of each era to evaluate the performance of Stable Diffusion reproducing the Buick language (Fig. 4). We have omitted in the study the 13th era, corresponding to "concept" designs, to keep the substantial text of the Stable Diffusion prompts homogeneous and interpretable. The word "concept" is difficult to associate with a particular era and results using it showed varied styles.

The equivalent process to the extraction of a product grammar in the case of a Text-to-Image AI model would be training the model to recognise and reproduce the relevant product visual features. As previously explained, our hypothesis was that Stable Diffusion could perform reasonably well with no additional training, so we considered that the model was potentially able to represent the whole Buick grammar. I.e., the Buick grammar was already "extracted" and contained in the Stable Diffusion model.

4.2. Validation of the Buick grammar

When a product grammar is obtained, the first way to validate it is trying to reproduce existing products by means of that grammar. In (McCormack et al., 2004) this is carried out by generating the front views of the 2002 Buick Regal and the 2002 Buick LeSabre, by using 14 and 13 transformation rules respectively. Fig. 5 represents the process corresponding to the 2002 Buick Regal

The equivalent process using AI models would be trying to make them produce identifiable pictures of existing products. Thus, we tried directly to represent the main 12 eras of Buick language by prompting in Stable Diffusion. The checkpoint used was Deliberate V2. We used very simple prompts describing the product, with the structure: "a front view of a YEAR Buick car" and a CFG of 7. A conditioning image of a modern Buick was used with *depth-midas* algorithm and a Control Weight of 0.6, just to force pure front views.

To assess the consistency of the produced solutions, 50 images of each era were created, and the features of each generation in each era were analysed. Fig. 6 displays a set of 10 generations. Although obtaining a similar set of images is very straightforward by using Stable Diffusion and the same parameter configuration, the authors can provide this information upon request. We have checked if the generated solution displayed the expected feature according to the era in each category, if it displayed a feature of a different era or if it displayed a non-Buick feature. Some features are similar across eras, as described in (McCormack et al., 2004, Figs. 7 and 8), so if in an era, a concept displayed one of these features unexpectedly, it was assigned to the closest previous era. The results are shown in Fig. 7.

To perform this analysis, we just considered the following grammar features described by the authors: fenders, outer and middle hood sections, hood flow line and grill. The study conducted by the authors to

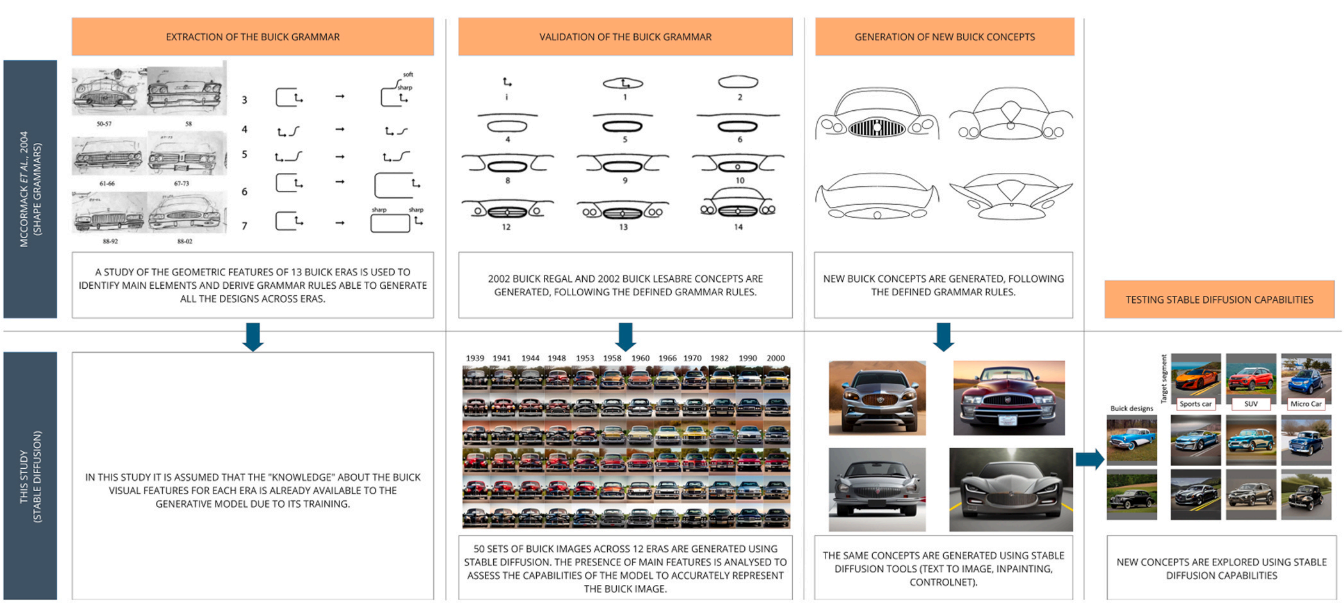

Fig. 2. Graphical outline of the case study. The process carried out by McCormack et al. is described in the upper row, while the equivalent actions using Stable Diffusion are represented in the lower row.

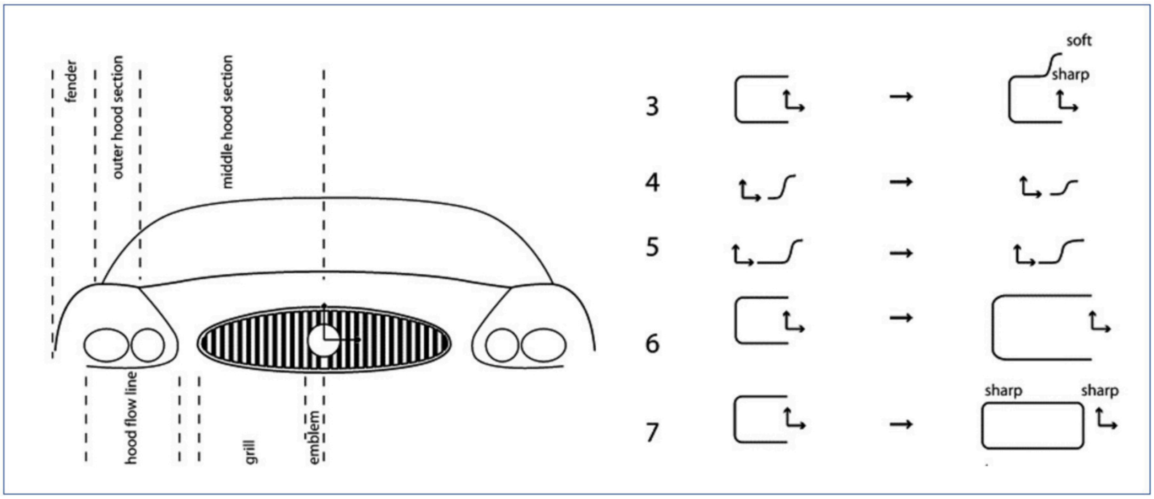

Fig. 3. Buick grammar features and some transformation rules for generating grill variations (McCormack et al., 2004).

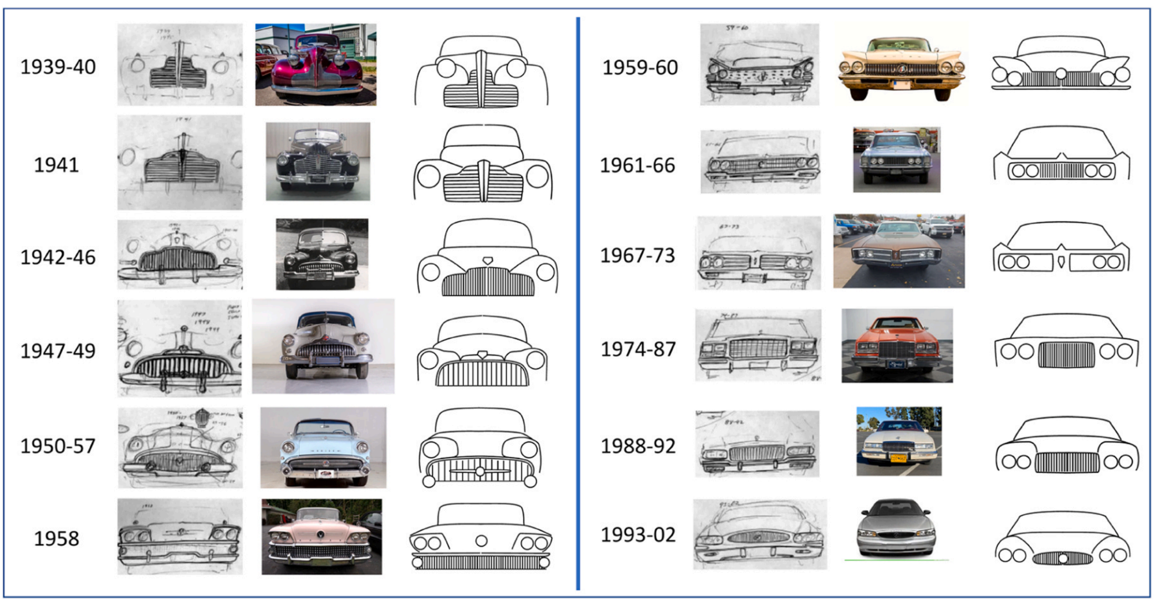

Fig. 4. Representation of Buick grammars from 1939 to 2002. Based in McCormack et al. (2004). As the complete formal grammar is not provided for each era, some of the elements have been derived from different parts of the text and images of correspondent Buick models.

obtain the Buick grammar found roof and headlights less relevant to the Buick brand image, and only some simple rules to build these elements are included. Therefore, they have not been considered in the analysis. We also dispensed with the emblem, as in the original paper it is simplified to a circle that just varies its diameter and vertical position.

Stable Diffusion proved very consistent in many of them, especially 1950–57 and 1974–87 (100 % and 99 % of correct representations). The Earl period (1939–1957) is very accurately represented, except for the grills in 1939–40 era. The shape of those grills is conformed over the sinuous surface of the front design and probably this has made difficult for the AI model to infer the actual frontal silhouette from the training pictures. Many of the other features, particularly the characteristic flow lines of the bumped hood and the round fenders, are represented very consistently. The 1967–2002 period is also very faithfully depicted.

The most difficult period to reproduce by Stable Diffusion has been 1958–66, with grades between 62 % and 76 %. The AI model has not adequately captured the distinctive aesthetic of 1958–1960 Buicks, with the sharp V-shaped sides conformed by the outer hood section, hood flow line and fender. In special, the 1958 unique rectangular grill covering the whole low part of the front has not been displayed in any of the images. In the case of 1961–66 grammars, the outer hood section and fender have been poorly represented, being much more consistent in the following era, which depicts similar features.

Overall feature representation is described in Fig. 8. Dark brown part of the bars corresponds to percentage of expected presence, i.e., features of an era present in pictures for that era. Light brown part corresponds to unexpected presence, features of an era present in pictures for another era. As shown, the overall presence of features is high, being grills and to a lesser extent outer hood sections the features less easily represented.

Evidently, this test gives only a rough description of the “knowledge” of Stable Diffusion about Buick features, and only for the checkpoint used. However, for the purposes of this study, we can conclude that this AI model is able to generate all the elements present in the grammar extracted by the authors in (McCormack et al., 2004).

In this sense, the Text-to-Image approach presents significant benefits when it comes to identifying brand identity traits. While shape grammars use a direct human identification of these features, a generative AI model can be trained to recognise that structure from a set of

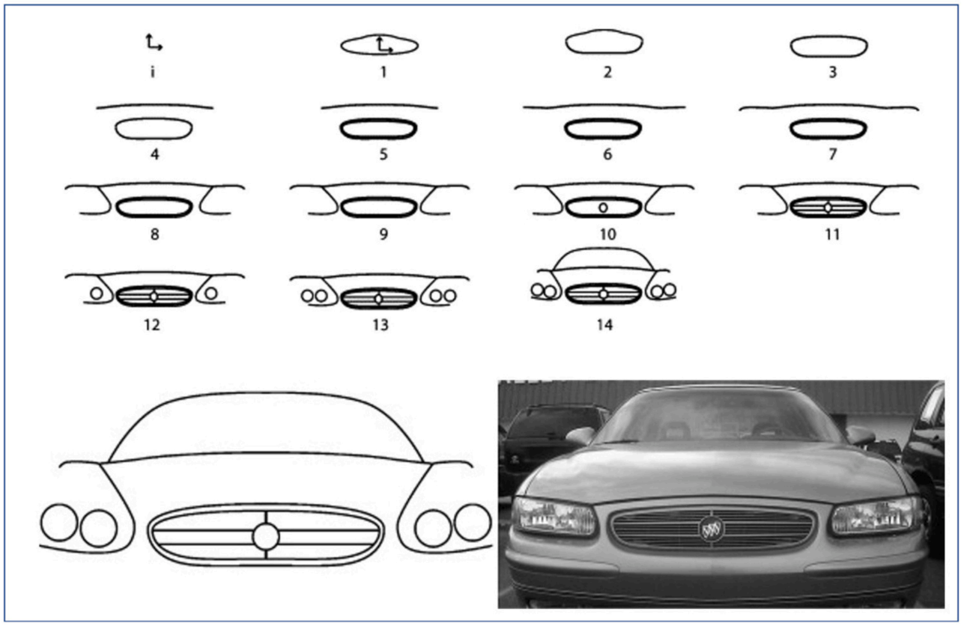

Fig. 5. Process to obtain the 2002 Buick Regal design using 14 transformation rules from the Buick grammar.

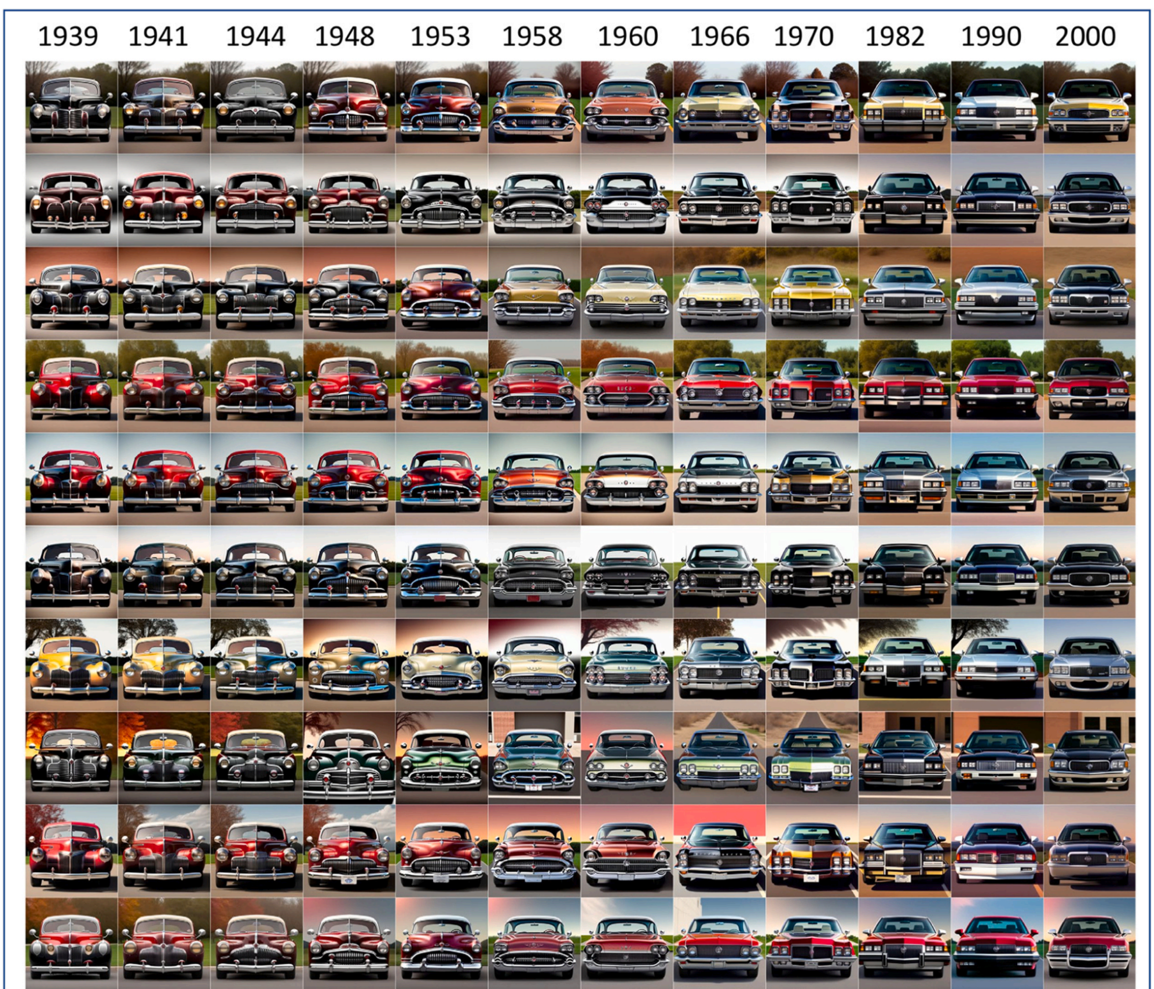

Fig. 6. A sample of 10 generations of Buick pictures over 12 eras produced using Stable Diffusion.

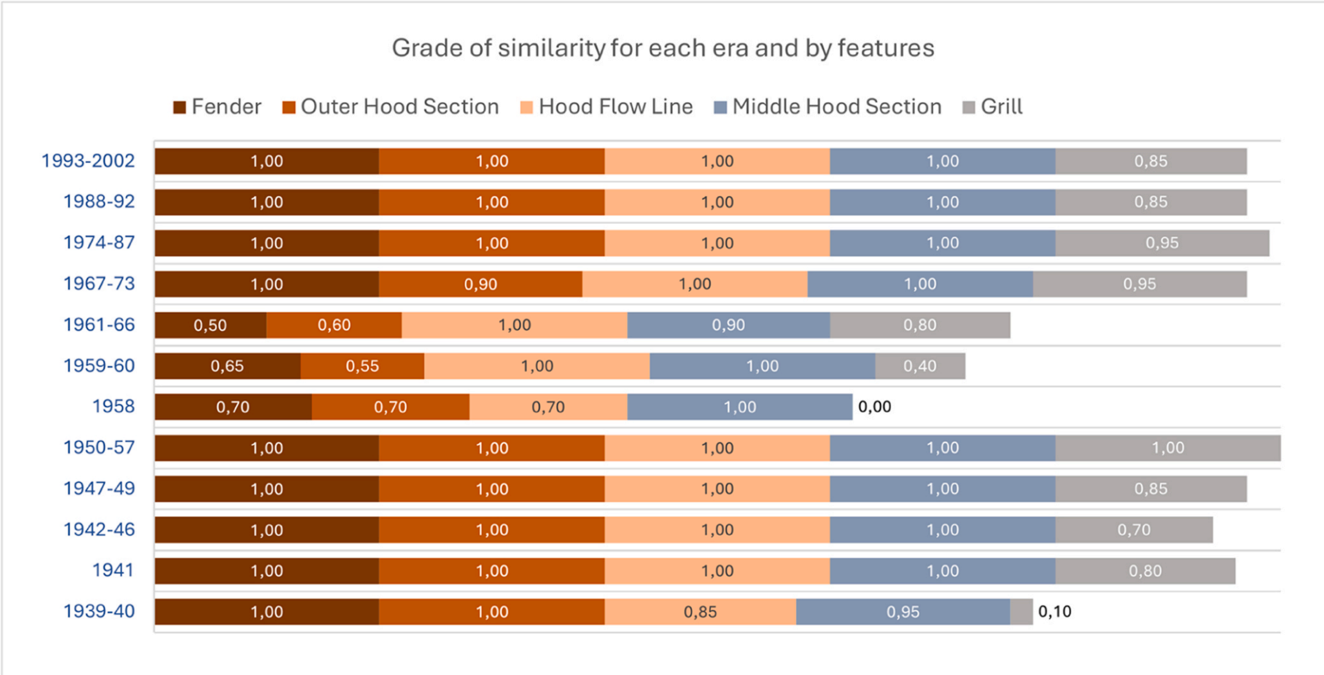

Fig. 7. Grade of similarity of Stable Diffusion grammar representations for each era and feature. Presence of each feature ranges from 0 to 1. The overall fidelity ranges from 0 to 5.

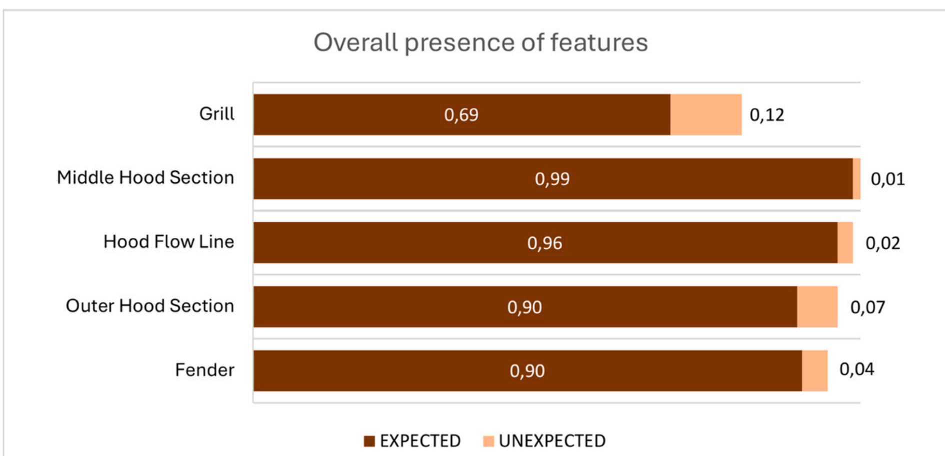

Fig. 8. Overall presence of each grammar element in generated pictures.

pictures. Instead of being processed in terms of grammar rules, this information is applied through generative AI semantic tools.

Even more, it is worth noting that a very powerful difference with Shape Grammars is that a generative AI model can directly utilise the grammar elements of a specific design from one single starting picture representing the desired aesthetic or brand image, using it to produce direct adaptations for the new concept. As an example, in Fig. 9a picture of a 1959 Buick Electra is used as conditioning image, combined with the prompt "front view of a modern SEGMENT car" for sport, SUV and executive segments.

The generation of these modern adapted Buick designs is immediate. This way, it is possible to consistently generate any variation from a starting point of any era. This procedure will be utilised in the next section.

4.3. Generation of new Buick concepts

The application of a shape grammar requires using the inferred grammar rules to produce new concepts, by means of an iterative process. Starting from an initial feature, grammar rules are applied iteratively to build the final object. In each step, choosing between different rules produce different alternatives.

As aforementioned, it is impossible to define such a procedure using Text-to-Image AI models. Alternatively, designers have some tools available in Stable Diffusion to generate new concepts using the information related to Buick brand. In this section we study these tools as compared to the mentioned Shape Grammar process. In (McCormack et al., 2004), after validating the grammar by generating existing Buick concepts, the authors first produce four new concepts demonstrating the capabilities of the approach to instil the brand identity into new original designs (Fig. 10). We will use Stable Diffusion to try to replicate these designs, which here we will call "target grammars", to compare the AI

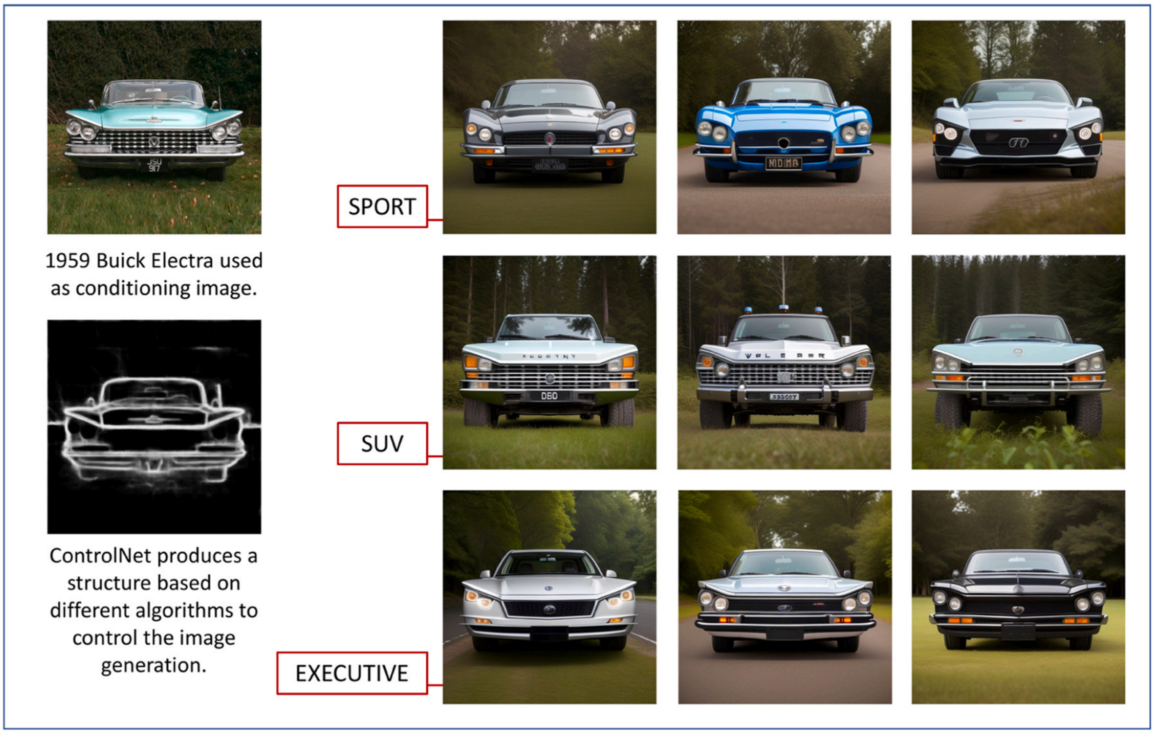

Fig. 9. Using an actual Buick image to instil its grammar into new concepts.

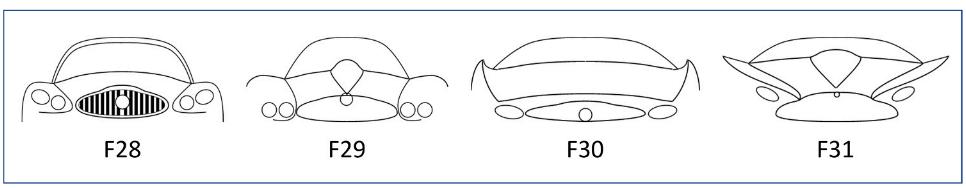

Fig. 10. First new concepts produced using the Buick grammar in (McCormack et al., 2004).

model workflow with the Shape Grammar approach.

We must keep in mind that this process is fictitious, in the sense that we are looking for a predefined target design instead of exploring possibilities using grammar elements. The comparison must be understood in terms of generation capabilities and design freedom. This issue will be considered in the discussion.

The first target grammar is depicted in the corresponding Figure 28 in (McCormack et al., 2004). From now on, we will refer to figures of that paper as F(Nº), to avoid confusion with the figures of the present one. So Figure 28 in (McCormack et al., 2004) will be referred as F28.

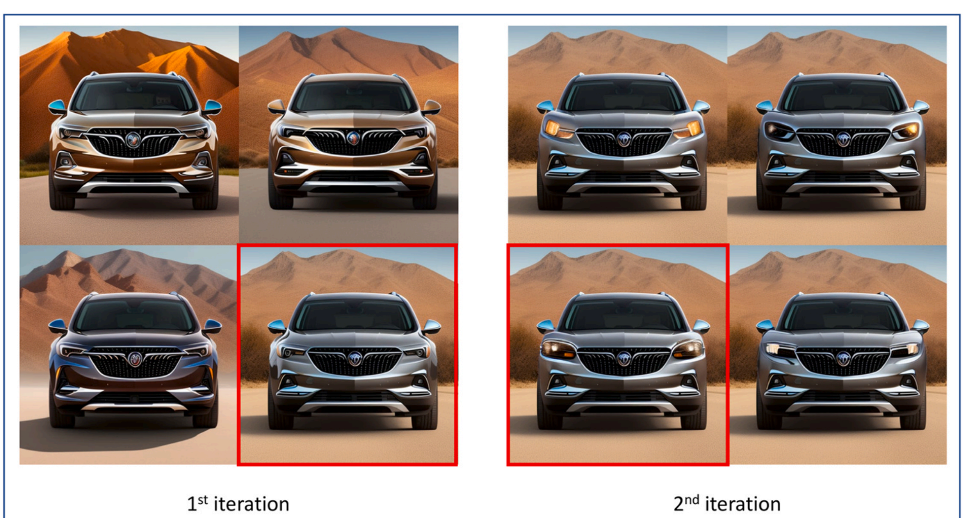

Fig. 11. First steps generating a Buick SUV. 4 concepts were produced in each iteration.

This concept represents a small SUV design derived from the grammar of a 2002 Buick LeSabre. Firstly, we generated a set of 4 concepts using a general prompt: "*a front view of a modern Buick SUV*". The same conditioning image used in the previous section was used here to ensure a similar point of view. Results are shown in the left part of Fig. 11 (first iteration).

All four alternatives made a good starting point as Buick SUVs. The fourth concept, which showed a more elegant grill, was chosen. It is worth noting that the words "modern Buick" have rendered front grills similar to those of the 1993–2002 era, and also close to the one used in F28 grammar.

The selected concept was used for inpainting. Four more concepts were then produced, using the prompt "Buick hood and fenders" and masking the respective areas. The goal was achieving rounded and slightly prominent fenders (Fig. 11, second iteration). The third of those concepts displayed a bit wider fender, due to the headlights design, and was chosen for the next iteration.

This time the inpaint is applied over the hood, trying to get a bulgier middle hood (Fig. 12, third iteration). The results were noticeable only in two of the concepts. The fourth concept was chosen for inpainting the grill. To attain the modern looking of the grill, the prompt "2002 Buick grill, chrome border" was used. The right side of Fig. 13 shows the concepts produced.

Variations in the grill were harder to attain, due mostly to limitations of the AI model to produce slight modifications on particular elements. Therefore, most of the times we did not expect Stable Diffusion to represent the exact geometry depicted in the grammar target, and a reasonable approximation and identification with a Buick grill was considered valid, being aware of the limitations in this sense. The second concept from the fourth iteration is chosen as the final one. The lines of the middle and outer hood section, the fender and the emblem closely resemble those of F28, whereas the grill and partially the hood flow line are less similar.

Interestingly, many different versions of Buick adaptations emerged along the exploration process in Stable Diffusion. That is not the case for the Buick grammar, as shown in F24 and F26 in the original paper. Shape Grammar processes are iterative, so the solution is built by increasing recursive applications of rules from the grill out (Fig. 13). Therefore, only one solution is in progress all the time, although of course many options can be explored by selecting different rules in each step.

The procedure followed using Stable Diffusion could instead be comparable to that of using Shape Grammars to find a new Buick model starting not from scratch, but from the grammar of an existing one. The application of rules over a whole defined Buick design would generate successive different versions of Buick concepts.

Concepts F29, F30 and F31 were obtained by using different conditioning images. F29 features a very prominent outer hood and fender lines that exaggerate the width of the vehicle, like the 1939–1946 designs. F30 uses a variation of the fender from 60's Buicks, while F31 presents a strong modification of hood flow lines that loosely resembles a sport version of 59–60 Buicks. In this case, the hood rule is applied producing an exaggerated sharp effect in the silhouette.

All these concepts depict varied and different silhouettes which are difficult to achieve using just prompting and inpainting. Here we find a current limitation of the AI model. Using Shape Grammars, it is easy to explore even "extreme" solutions within the grammar space, but in the case of the Text-to-Image AI model, some additional information is needed to direct it towards this kind of solutions, as probably they will not be produced spontaneously. Therefore, the designer must define in advance some features, thus partially losing unexpectedness.

There are several ways to address this issue within Stable Diffusion. As described in the previous section, in this case study we used mainly conditioning images through ControlNet, following these steps:

1) Search for a suitable picture of a Buick, closely resembling the main silhouette to achieve or susceptible of producing a close one.
2) Use this picture to condition the generation, by means of ControlNet algorithm.
3) Use a prompt related to Buick or to the required type of car.
4) Refine the results as in the previous examples.

In the case of F29 target grammar we selected as conditioning image a picture of a 1945 Buick Super, suitable for generating the initial silhouette. We used two different prompts: "a front view of a 2000 Buick SUV", and "a front view of a 2000 Buick executive", to explore possible alternatives that could fit the target grammar. Some of the results are shown in Fig. 14.

Many of these concepts already present some features close to that of the target grammar: fenders and outer hoods (produced by the conditioning image), but also the hood flow lines, some similar grills and the placement of headlights. We have selected the C option for executive concepts to show an example of the rest of the process. It consisted of generating a bulgier hood using prompts related to 1949 Buicks, something that proved to work well most of the time, and trying to get a closer grill (Fig. 15). The exact central feature of the hood was not easy to produce, by the same reason that specific grills are not either. But the overall appearance of the concept perfectly represents a modern Buick

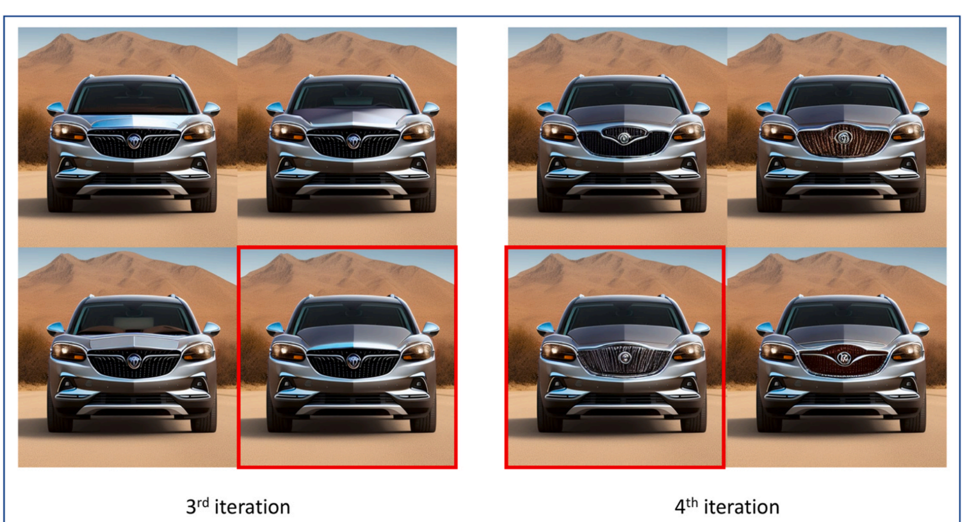

Fig. 12. Producing variations in the hood and grill.

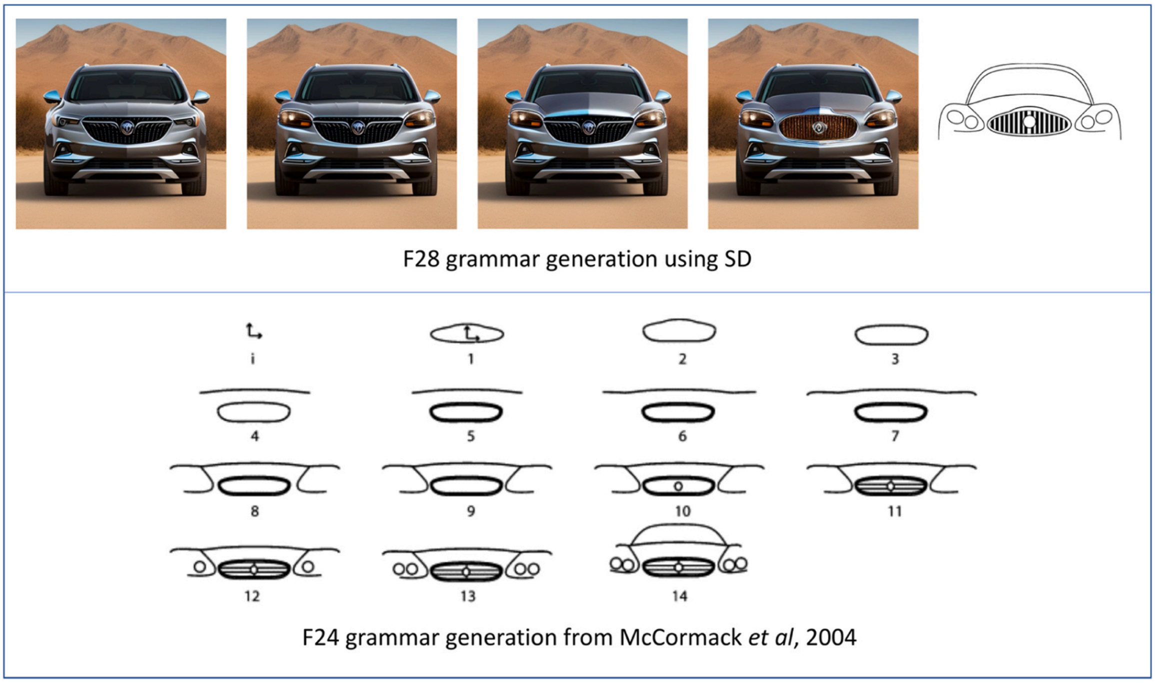

Fig. 13. Sequence for the generation of F28 Buick concept versus generation of F24 (2002 Buick Regal) shown in McCormack et al., (2004). In each step of the F28 generation, a Buick concept is produced, while in F24 the concept is built step by step.

Fig. 14. Several SUV and executive concepts for the F29 target grammar, using a picture of a 1945 Buick Super as conditioning image.

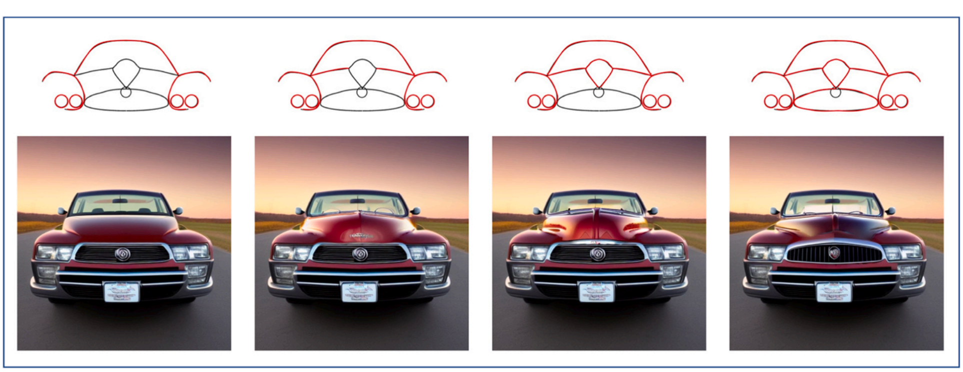

Fig. 15. F29 generation process. Above we include the F29 target grammar, highlighting those elements that are increasingly achieved in each step.

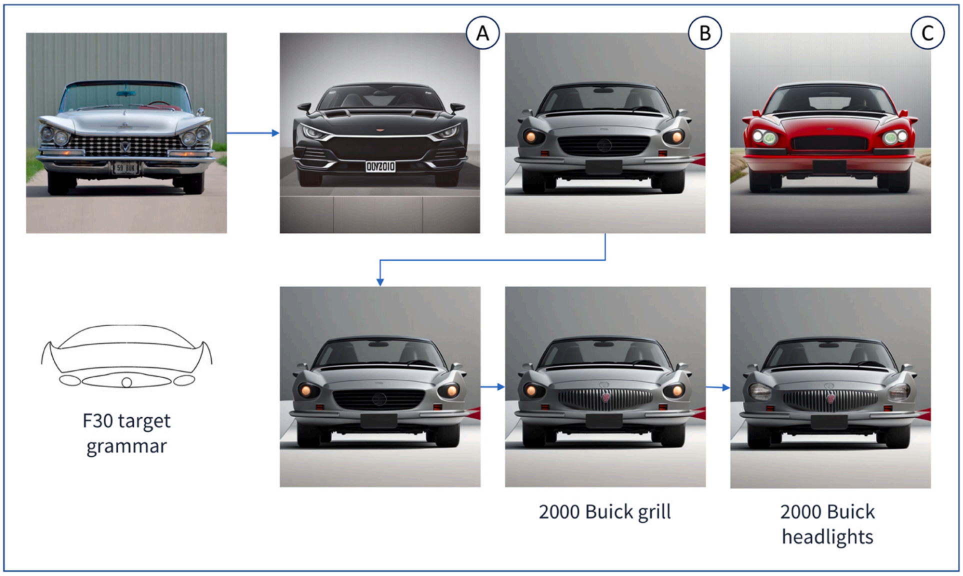

Fig. 16. F30 generation process. Some concepts obtained using "a front view of a modern sport car" prompt and below, adjustments on one of them to achieve the target grammar.

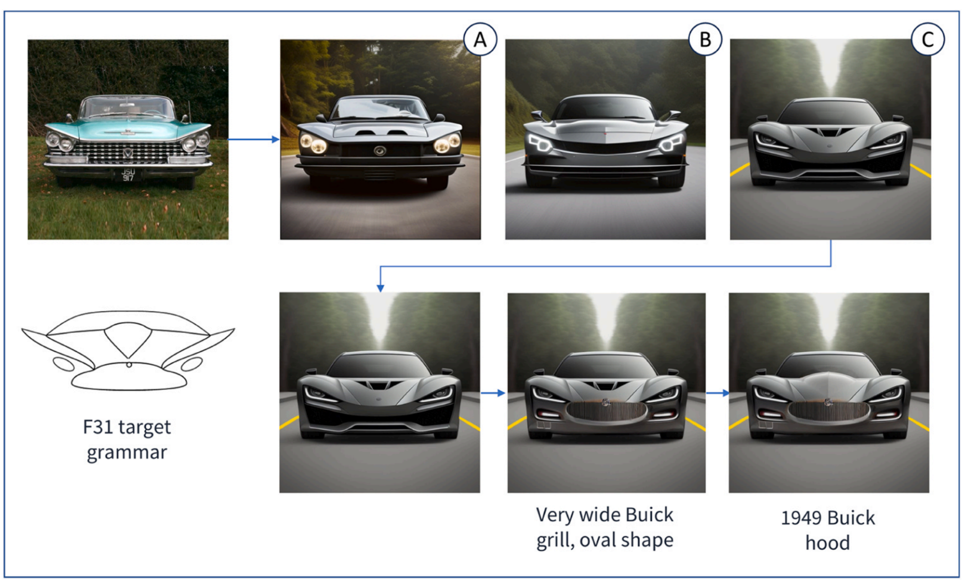

Fig. 17. F31 generation process. As in previous picture, some concepts obtained using "a front view of a modern sport car" prompt and below, adjustments on one of them to achieve the target grammar.

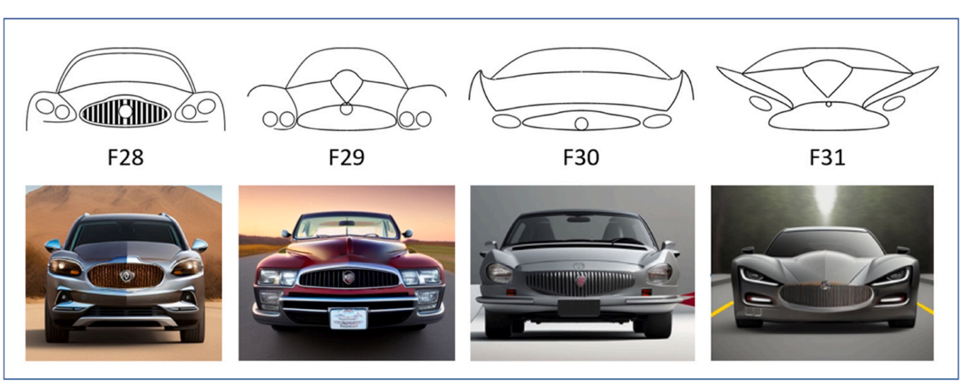

Fig. 18. In the upper row, concepts F28 to F31 generated in McCormack et al. using the Buick grammar. Below, equivalent concepts generated using Stable Diffusion.

design following the proposed lines in F29 grammar.

Finally, Fig. 16 and Fig. 17 shows the corresponding process for F30 and F31 according to this approach. To incorporate the special outer hoods and fenders of these solutions, we used a picture of a 1959 Buick LeSabre for F30, and one of a 1959 Buick Electra for F31. We used these two pictures both as starting picture and as conditioning picture. Then we asked for "a front view of a modern sport car" in the prompt.

Grills are still away from the target grammars, but the overall results are very reasonable. However, the main silhouette has always been predefined by us, to force the desired appearance. This may represent a disadvantage with respect to Shape Grammars, as we will discuss in the next section. On the other side, pictures generated this way always resulted very close to the idea of an updated Buick, just as in the first example we showed in Fig. 9, even though several iterations had to be performed to make all the features of the target grammar emerge or get close to that of the target one. Fig. 18 summarizes the four concepts obtained by Stable Diffusion for concepts F28-F31.

Besides these four concepts, another sample of eight designs is afterwards presented in (McCormack et al., 2004). As in the previous examples, we used different conditioning images to generate these designs The results are shown in Fig. 19.

Again, letting apart the case of the grills, Stable Diffusion produced interesting concepts whose grammars were close enough to the target ones.

4.4. Exploring Stable Diffusion possibilities

In the previous section, Stable Diffusion has been used to reproduce predefined target designs. The goal was comparing both the capabilities and the workflows of the AI model against those of Shape Grammar to get the same results. In this section we describe some other tests without that restriction, to evaluate the performance of the AI model when freely exploring shapes.

4.4.1. Generating full views of the concepts

A considerable advantage of using AI Text2Img versus Shape Grammars is that, once trained, the AI model can capture the whole product visual identity. Understandably, in (McCormack et al., 2004), the analysis is limited to the iconic front view to simplify the grammar extraction and usage. Therefore, designers can only explore front view designs. This is a valid approach and yields fruitful results. However, Stable Diffusion can reproduce not only front views, but any view of Buick concepts. In fact, we have had to force the front view representation, as Stable Diffusion tried by defect to generate perspective views. This way, if a full view of any car is used to guide the Stable Diffusion generation, a more complete depiction of the solution is obtained.

Moreover, as described before in the example of Fig. 6, a designer could directly use pictures of particular Buick designs to generate new updated concepts partially carrying the Buick aesthetic, even without asking for a Buick vehicle (although it will help, particularly to attain Buick grills). This would be an immediate way to use Buick grammar elements in new designs. Two of these examples are depicted in Fig. 20 and Fig. 21, in which three concepts for sport cars, SUVs and executive ones are generated from pictures of actual Buick designs by using prompts related to each segment.

4.4.2. Blending styles from different eras or segments

Mixing Buick grammars from different eras is also straightforward, by using one of them as the starting image and the other one as the conditioning one (Fig. 22). A Buick related prompt (Buick sport car, for instance) helps guiding the process towards the desired aesthetic.

Following this strategy, the designer could play with different Buick designs to produce concepts with mixed grammar elements and then use them as starting or conditioning pictures to obtain new updated concepts as in Fig. 20 and Fig. 21. This combination of features from different source products can be related to morphing techniques (Chen et al., 2003; Hsiao and Liu, 2002), as well as to the product genetics approach (Hsiao et al., 2010), also typically related to Shape Grammars.

In this sense, the same procedure may be applied to mix Buick cars with whatever other kind of vehicle, aesthetics or style. Using different pictures as starting and conditioning images allows for a product genetics way of shape exploration that is very interesting and highly productive. In Fig. 23, several images of Buick designs are combined with pictures of cars from different segments (sports car, SUV, micro car and MPV) to produce updated versions. Pictures of the target vehicles were used as conditioning images to generate Buicks with the silhouette of each segment, and prompts related to the era of each Buick aided to instil relevant features into the final concept. Different effects were attained depending on However, controlling which features were inherited form each parent was not intuitive and overall results were often unpredictable, something that could indeed favour emergence.

Similar (and sometimes faster) results can be achieved simply by using a starting image and a Buick related prompt, the inverse process of the one used in Fig. 20 and Fig. 21. In Fig. 24, several new Buick concepts displaying some cues from past Buick designs have been obtained

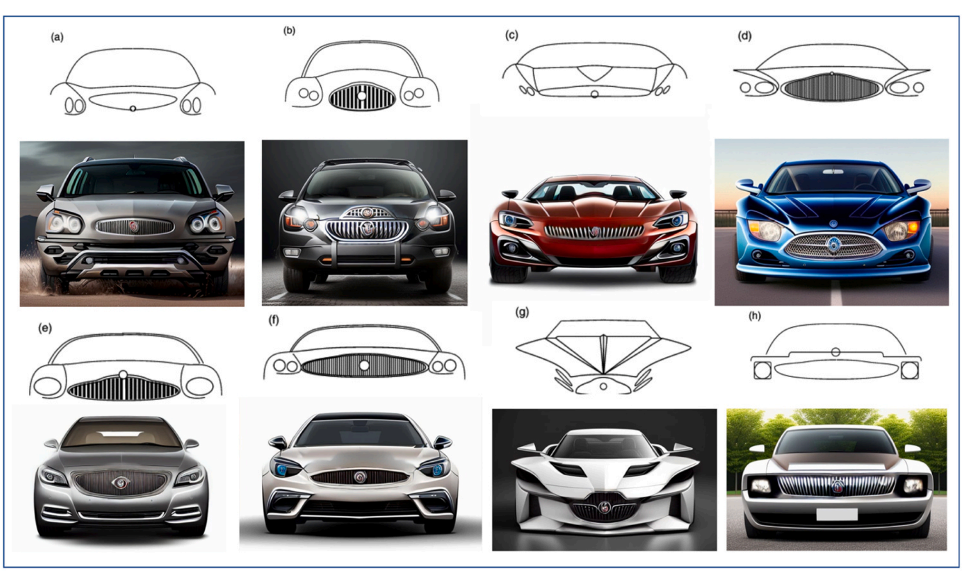

Fig. 19. Buick concepts produced by Stable Diffusion for solutions depicted in F34.

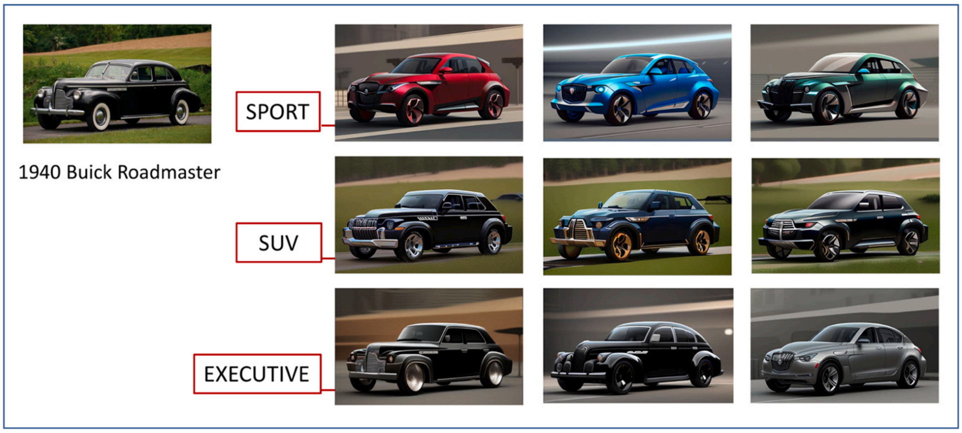

Fig. 20. Updated concepts for sport, SUV and executive Buicks from Buick 1940 Roadmaster grammar.

Fig. 21. Updated versions for sport, SUV and executive Buicks from Buick 1956 Special Coupé grammar.

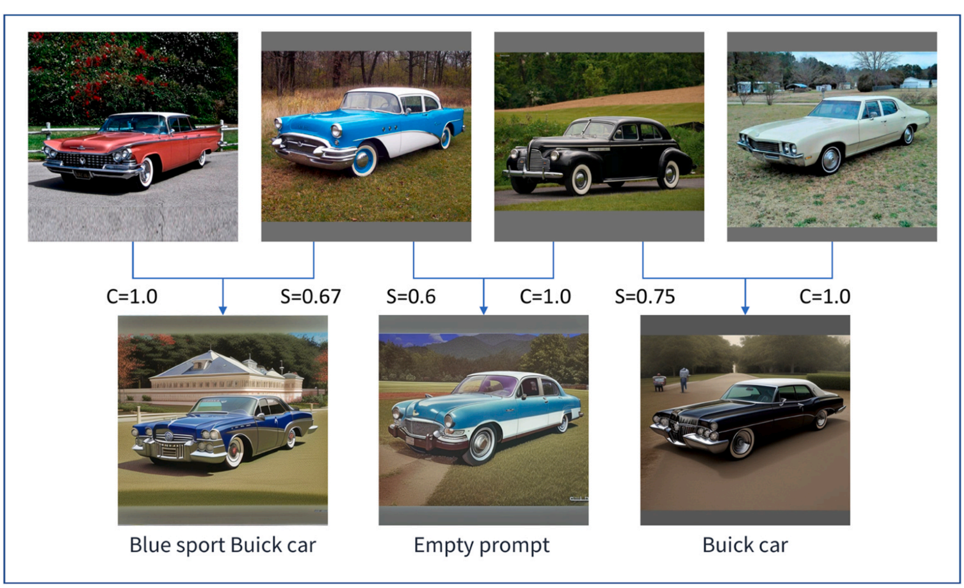

Fig. 22. Some examples of grammar mixing. The upper pairs of Buick images produce the lower ones. Starting (S) and conditioning (C) images are shown along with the Denoising Strength and Control Weight. Prompts used in each case are included below.

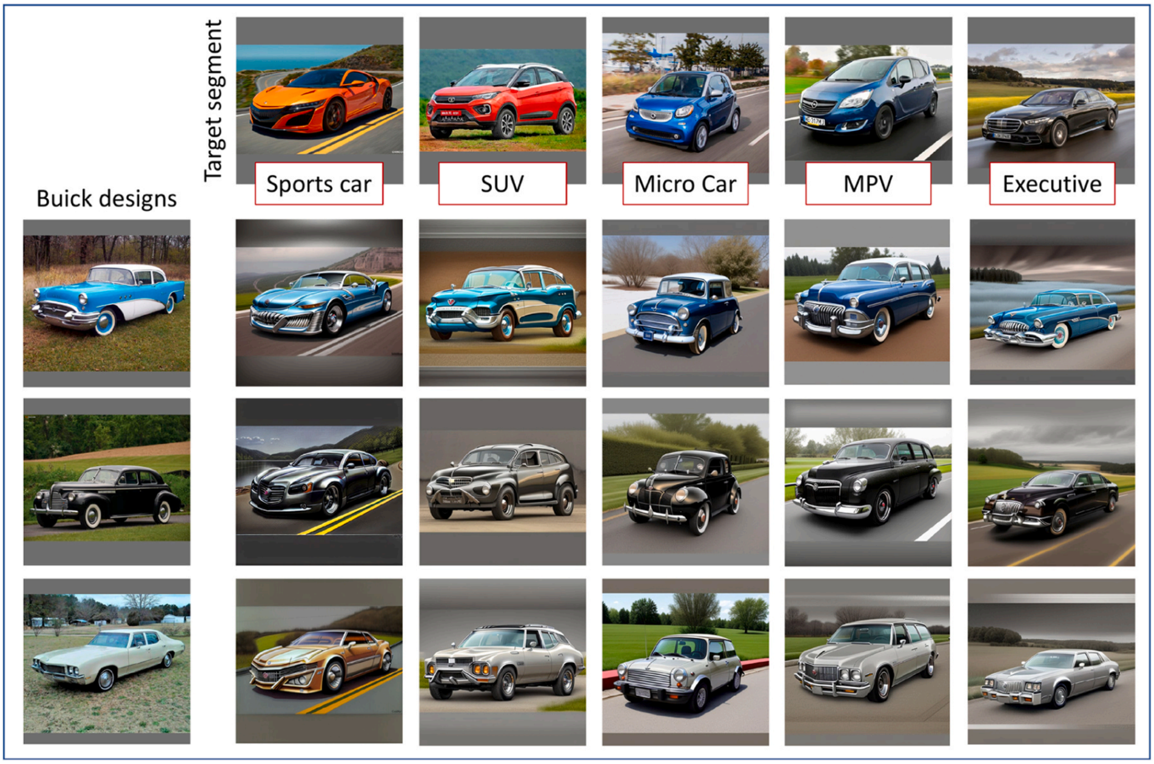

Fig. 23. Combination of Buick car designs with vehicles from different segments.

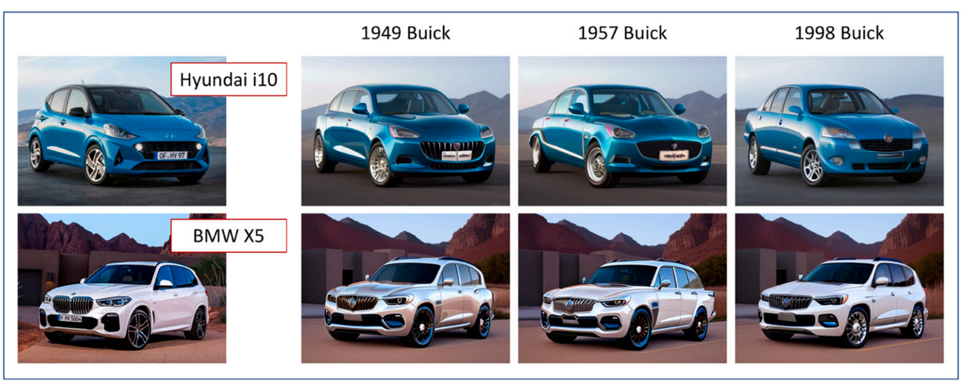

Fig. 24. Buick concepts produced by using Hyundai i10 and BMW X5 as starting image and "YEAR Buick car" as prompt. Concepts for years 1949, 1957 and 1998 were generated.

from pictures of other vehicles. Using a picture of a Hyundai i10 and a BMW X5, a Buick city car and a robust Buick SUV respectively are generated with aesthetics from different eras. Designs corresponding to updated 1949 cars portrait the characteristic bulging hood and prominent fenders, as well as typical curved roofs of those Buicks. The 1957 versions are streamlined, displaying the distinctive outer hood lines, rear fins and a box shaped roof. The city car version even features the classic side line of these Buicks, although going downwards instead of upwards due to the structure of the starting picture. Concepts for 1998 style present the dynamic inclined hood and softer lines of that era, such as the smooth headlights and moderated grills.

4.4.3. *Using affective terms in prompts*

Stable Diffusion allows for many other ways of exploring concepts that could complement the use of Shape Grammars. As a final example, we explored the feasibility of implementing affective terms to be used in the prompt. In (McCormack et al., 2004, p. 12), the authors suggest the possibility of including descriptive terms to specify a desired emotional message: *"Rules can be accompanied by a descriptive term, which would express the emotional impact of applying the rule"*. Although the paper does not provide examples of this application to Buick designs, we conducted several generation tests to check the effect of some affective terms in the design outcomes (Fig. 25 and Fig. 26). Words used for the generations were extracted from a study about affective terms in car design conducted in (Helander et al., 2013).

The analysis of the AI model response to affective terms falls out of the scope of this study and will be a subject for another work. However, it is possible to appreciate some coherence in the resulting images, as the dynamic style for "sporty" and "sexy", the elongated and stable design for "elegant" or the voluminous concepts for "aggressive". Interestingly, the term "rugged" generally produced shabby old vehicles.

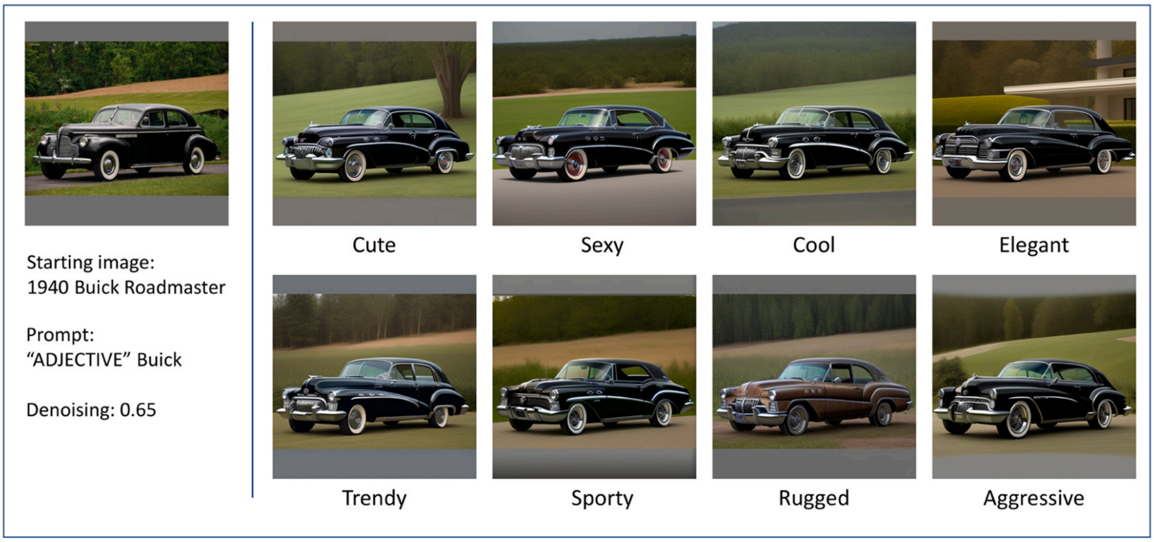

Fig. 25. Buick concepts generated by using simple prompts consisting of affective terms.

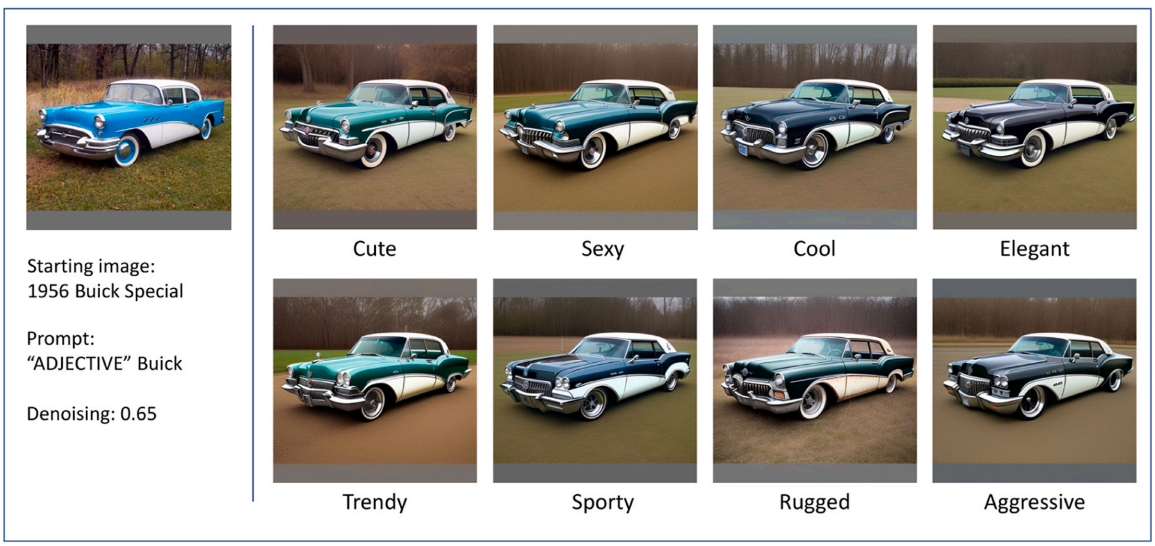

Fig. 26. Buick concepts generated by using simple prompts consisting of affective terms.

5. Discussion

The results of this exploratory study show a high potential of Text-to-image AI models as ready-to-use conceptual support tools offering a wide variety of resources more than the immediate "inspirational image" one. The study offers an example of new ways to address the human-computer co-creative activity in tasks related to conceptual design and shape exploration.

5.1. Exploration capabilities

Stable Diffusion has performed all the test tasks efficiently and with no need for additional training. Both the extraction/assimilation of the Buick grammar and its use in design tasks have been successfully tested. Some limitations have been found when dealing with detail or specific elements, whose variations are not easy to achieve and control, as well as when trying to explore very extreme variants within the Buick brand space.

As compared to the current development state of Stable Diffusion, Shape Grammar provides a more controllable exploratory mechanism if a brand image is to be maintained. The Buick grammar include both feature creation and feature modification rules. Creation actions are easily reproduced by Stable Diffusion. However, Shape Grammar allows for a continuous exploration of the space of solutions, once the constituent elements of a given aesthetic have been obtained. The designer may use the modification rules to generate variations of a particular solution, adjusting the features to obtain for example a stronger look through wider fenders or a more dynamic one lowering or sharpening them. This kind of shape exploration, specially at a particular feature level, is not easy to achieve in the case of Stable Diffusion. On one hand, some features not so prevalent in the Buick grammar contained in Stable Diffusion are not likely to emerge by mere prompting. This was the case of specific outer hoods or some grills in the new concepts. On the other hand, it is not immediate to get variations at feature level with mere descriptions. Generic AI models are trained using very large sets of images of existing objects paired with their descriptions. Presumably, most of these descriptions are not so deeply detailed, and specific training will be necessary to handle this kind of exploration. Currently, generating coherent variations of particular features is possible, but it needs elaboration, and the results not always lay within the grammar space.

Indeed, one of the strengths of the Buick shape grammar is that it plays a control role for design solutions, as the main goal of its use is here to ensure that Buick brand image is present in every generation. This is not the case of Stable Diffusion. It can represent all the relevant features of a Buick vehicle and convincing past and updated designs. However, pushing the exploration further may result in solutions too far away from the Buick brand, because it lacks this mechanism that Shape Grammar provides. An alternative to it could be a purposeful use of conditioning images.

In any case, this may also be an advantage. While Stable Diffusion present these limitations within the Buick grammar, it also permits a wider field of experimentation. As pointed out in Section 1, the generation of images depicting definitive designs may not be the ideal use of this technology. Probably, designers could find another kind of inspiration being exposed to stimuli that needed further work. In fact, the concepts produced by Shape Grammars are incomplete as well, and the designer must finish them by developing a solution based on their schematic lines. Ethnographic or other kind of in-depth studies with practical design cases should be conducted to address these questions.

An interesting research exercise to verify the performance of Stable Diffusion (or any other model) capturing the brand grammar in extreme situations would be contrasting the ability of the generated images to convey the Buick brand by users' evaluation. This external validation is not performed in (McCormack et al., 2004) either, and falls out of the scope of this work, but will be addressed in future studies. That study could be extended to other products and brands.

5.2. Limitations of the study

The previous considerations must be framed in the context of a study purposedly restricted in several ways, which will be discussed in the following paragraphs.

Firstly, just one Text-to-Image AI model (Stable Diffusion) has been used to produce the images. There are many other models and diverse checkpoints which might yield different results. As referred in the first section, there are already some comparative studies between models, and probably several more will be necessary to determine the best way to adapt this approach to product conceptualization.

The study has been conducted with no training of the AI model for specific goals. Although, as an out-of-the-box tool, Stable Diffusion has produced a satisfactory representation of most of the features of different Buick eras, some limitations have been detected. Problems such as generating a variety of grills could probably have been overcome by using a more focused checkpoint trained with Buick images, as performance of a Text-to-Image AI model representing a particular object depends on how it has been trained with respect to that object. A series of studies testing the response of AI models to products other than cars could provide insight on the need for further training in specific cases. Likewise, a more thorough study using specifically trained checkpoints for conceptual design situations (styles, trends, affective terms) could help to further determine the AI models capabilities as design assistants.

On the other hand, the use of Stable Diffusion has been limited to the actions described in Section 2. We have not used any other available workflow, as our purpose was to test the use of the technology alone. However, Stable Diffusion and other AI models can be used in combination with 2D and 3D third party applications, something which hugely multiplies the exploratory power of this technology and provides much more control to the designer. It is possible start drawing a rough sketch or build some preliminary 3D volumes and let Stable Diffusion create images based on them. Many of the grammar features that were hard to achieve during the experimental exercise could had been more precisely controlled if complementary software had been utilised.

Moreover, another restriction was imposed on the use of prompts. Prompts are (by definition) the core interaction with a Text-to-Image AI model, so the resulting images are greatly affected by their composition on the resulting images is essential. However, we considered the analysis of the use of prompts as a different research question, limiting their use to very simple statements.

Finally, the study is based on a comparison with Shape Grammars conducted by our research group as a first approach to the application of AI models in conceptual design. This has provided us with a framework to conduct specific tests to assess the performance of the AI model. However, the exercise conducted here may be replicated for other CACD approaches, thus creating a map of relationships with existing digital tools that could incorporate these AI models in different ways. Multimodal agents are already being tested, and new progress will enable diverse uses by combining existing and emerging procedures.

5.3. Future research

An example of these combinations would be using generative modelling software and Text-to-Image AI models to overcome the problem of the AI model with specific design details, something we are working on. The production of generative controlled silhouettes (or even 3D models) within the desired grammar, combined with the further use of Stable Diffusion to complete them and generate pictures of design variants is a promising approach to the incorporation of AI models to current techniques.

With regards to the nature of the potential human-computer interaction in terms of cognitive processes, an important issue that researchers and designers will have to analyse is how this new approach will affect creative performance. As the introduction of AI models in the conceptual workflow will surely modify these processes, a proper knowledge about how this could occur and how to address the interaction to make it as efficient as possible is needed.

For instance, during the process of extraction and application of shape grammars, designers need to find relevant shapes for the design purpose of the analysed object. This task allows for emergence. In the case of AI models, there is not such room for emergence. Once the model has been trained, shapes are already identified by that model, not for the human designer, and therefore this part of the process is fully automated. So is the generation of new concepts, and because of that, the role of designers in this process needs to adapt to take full advantage of the tool. Although this automation presents many advantages, it is to be analysed how this workflow influences the creative performance of the designer. Intuitively, new ways for emergence may arise, as the images generated by the AI model may facilitate reinterpretation or other cognitive processes. For example, Stable Diffusion has proved very effective exploring combination mechanisms. However, more in-depth studies could help to understand the implications of the use of AI tools and its impact on creative processes, both at an individual level and on the discipline itself. These experiments will require the participation of different stakeholders such as practitioners, scholars, educators and students, in order to gain the knowledge to properly implement this technology in design teaching and practice.

Finally, as generative models become more efficient interpreting and processing training data, an interesting line of research will be to explore their capabilities handling additional requirements other than formal ones. Multimodal models are already able to discuss some functional aspects of a given design just from a representing image. In the case of automotive design, for instance, functional aspects interpretable from images, such as aerodynamic considerations, safety, structural issues or comfort and ergonomics may be incorporated into interactive processes to generate solutions within a wider design scope. Generative models producing 3D objects from text are already available, and although their results are very preliminary, they allow for conceptual tests regarding functional aspects. An integration of all these approaches would enable designers to handle different design requirements with computational generative support.

6. Conclusion

The use of Text-to-Image AI models as tools supporting product shape exploration is very recent. This preliminary study, focused on a very basic use of one of these models, aimed at testing their potential through a comparison with a very solid approach to generative design such as Shape Grammars.

The results suggest that this technology presents an enormous potential in the field of product design. The implementation of AI models for product shape exploration can be very productive, especially for fast idea visualisation. The performance of Stable Diffusion replicating a Shape Grammar practical application has been very notable, and some related creativity procedures, such as combination and product genetics, have been successfully tested.

Several limitations have also been found, mainly related to lack of variety when exploring specific product features or otherwise the need for a substantial human intervention to direct the exploration process, thus partially losing creative support. Compared to Shape Grammars, it has also been difficult to reproduce the definition of a grammar space enclosing the Buick brand.

Even though some of these limitations will presumably diminish as algorithms performance improves, they can also be overcome by combining several generative approaches. Further research in this field may shed light on the most efficient ways to incorporate Text-to-Image and other kind of AI models to the conceptual design workflow.

Funding sources

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CRediT authorship contribution statement

Jose Antonio Diego-Mas: Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Conceptualization. **Jorge Alcaide-Marzal:** Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

Data will be made available on request.

References

Abdalla, S.B., Rashid, M., Ara, D.R., 2021. Plausibility of CAAD in conceptual design: challenges in architectural engineering for early-stage digital design tools. J. Archit. Eng. *27* (2), 04021004 https://doi.org/10.1061/(ASCE)AE.1943-5568.0000457.

Agarwal, M., Cagan, J., 1996. A blend of different tastes: the language of coffeemakers. Environ. Plan. B: Plan. Des. *25* (2), 205–226. https://doi.org/10.1068/b250205.

Ak, K.E., Lim, J.H., Tham, J.Y., Kassim, A.A., 2020. Semantically consistent text to fashion image synthesis with an enhanced attentional generative adversarial network. Pattern Recognit. Lett. *135*, 22–29. https://doi.org/10.1016/J.PATREC.2020.02.030.

Alcaide-Marzal, J., Diego-Mas, J.A., Acosta-Zazueta, G., 2020. A 3D shape generative method for aesthetic product design. Des. Stud. *66* https://doi.org/10.1016/j.destud.2019.11.003.

Alvarado, C., Davis, R., 2007. SketchREAD: A multi-domain sketch recognition engine. ACM SIGGRAPH 2007 Courses 34–es. https://doi.org/10.1145/1281500.1281545.

Barros, M., Duarte, J.P., Chaparro, B.M., 2015. A grammar-based model for the mass customisation of chairs: modelling the optimisation part. Nexus Netw. J. *17* (3), 875–898. https://doi.org/10.1007/s00004-015-0265-5.

Basalla, M., Schneider, J., & Brocke, J. (2022). *Creativity of Deep Learning: Conceptualization and Assessment*. 99–109. https://doi.org/10.5220/0010783500003116.

Bernal, M., Haymaker, J.R., Eastman, C., 2015. On the role of computational support for designers in action. Des. Stud. *41*, 163–182. https://doi.org/10.1016/J.DESTUD.2015.08.001.

Bonnardel, N., Zenasni, F., 2010. The impact of technology on creativity in design: An enhancement? Creat. Innov. Manag. *19* (2), 180–191. https://doi.org/10.1111/J.1467-8691.2010.00560.X.

Briard, T., Jean, C., Aoussat, A., Véron, P., 2023. Challenges for data-driven design in early physical product design: a scientific and industrial perspective. Comput. Ind. *145*, 103814 https://doi.org/10.1016/J.COMPIND.2022.103814.

Brock, A., Donahue, J., Simonyan, K., 2018. Large Scale GAN Train. High. Fidel. Nat. Image Synth. ⟨https://arxiv.org/abs/1809.11096v2⟩.

Castro Pena, M.L., Carballal, A., Rodríguez-Fernández, N., Santos, I., Romero, J., 2021. Artificial intelligence applied to conceptual design. A review of its use in architecture. Autom. Constr. *124*, 103550 https://doi.org/10.1016/J.AUTCON.2021.103550.

Chandrasegaran, S.K., Ramani, K., Sriram, R.D., Horváth, I., Bernard, A., Harik, R.F., Gao, W., 2013. The evolution, challenges, and future of knowledge representation in product design systems. Comput. -Aided Des. *45* (2), 204–228. https://doi.org/10.1016/J.CAD.2012.08.006.

Chen, L.-L., Wang, G.F., Hsiao, K.-A., Liang, J., 2003. Affective product shapes through image morphing. Proc. 2003 Int. Conf. Des. Pleasurable Prod. Interfaces DPPI 03 11. https://doi.org/10.1145/782896.782900.

Chiarello, F., Belingheri, P., Fantoni, G., 2021. Data science for engineering design: State of the art and future directions. Comput. Ind. *129* https://doi.org/10.1016/j.compind.2021.103447.

Chiou, L.-Y., Hung, P.-K., Liang, R.-H., Wang, C.-T., 2023. Designing with AI: An Exploration of Co-Ideation with Image Generators. Proc. 2023 ACM Des. Interact. Syst. Conf. 1941–1954. https://doi.org/10.1145/3563657.3596001.

Company, P., Contero, M., Varley, P., Aleixos, N., Naya, F., 2009. Computer-aided sketching as a tool to promote innovation in the new product development process. Comput. Ind. *60* (8), 592–603. https://doi.org/10.1016/j.compind.2009.05.018.

Crilly, N., 2015. Fixation and creativity in concept development: The attitudes and practices of expert designers. Des. Stud. *38*, 54–91. https://doi.org/10.1016/J.DESTUD.2015.01.002.

Cui, J., Tang, M.X., 2013. Integrating shape grammars into a generative system for Zhuang ethnic embroidery design exploration. CAD Comput. Aided Des. *45* (3), 591–604. ⟨http://www.sciencedirect.com/science/article/pii/S0010448512001704⟩.

Dhariwal, P., Nichol, A., 2021. Diffusion models beat GANs on image synthesis. Adv. Neural Inf. Process. Syst. *11*, 8780–8794. ⟨https://arxiv.org/abs/2105.05233v4⟩.

Dinar, M., Shah, J.J., Cagan, J., Leifer, L., Linsey, J., Smith, S.M., Hernandez, N.V., 2015. Empirical studies of designer thinking: past, present, and future. J. Mech. Des., Trans. ASME *137* (2). https://doi.org/10.1115/1.4029025/376148.

Dorst, K., Cross, N., 2001. Creativity in the design process: co-evolution of problem-solution. Des. Stud. *22* (5), 425–437. https://doi.org/10.1016/S0142-694X(01)00009-6.

Ekströmer, P., Wever, R., 2019. Exploring computer aided design tools for design ideation. Des. J. *22*, 10–12. https://doi.org/10.1080/14606925.2019.1595031.

Elasri, M., Elharrouss, O., Al-Maadeed, S., Tairi, H., 2022. Image generation: a review. Neural Process. Lett. https://doi.org/10.1007/s11063-022-10777-x.

Eloy, S., Pauwels, P., Economou, A., 2018. Advances in implemented shape grammars: solutions and applications. Artif. Intell. Eng. Des., Anal. Manuf. *32* (2), 131–137. https://doi.org/10.1017/S0890060417000634.

Forbus, K.D., 1988. Intelligent computer-aided engineering. AI Mag. *9* (3), 23-23. https://doi.org/10.1609/AIMAG.V9I3.940.

Gero, J., 1996. Creativity, emergence and evolution in design: concepts and framework. Knowl. Based Syst. *9* (7), 435–448. ⟨https://www.sciencedirect.com/science/article/pii/S0950705196010544⟩.

Gero, J.S., 2000. Computational Models of Innovative and Creative Design Processes. Technol. Forecast. Soc. Change *64* (2), 183–196. https://doi.org/10.1016/S0040-1625(99)00105-5.

Gero, J.S., Maher, M.Lou, 1993. Modeling Creativity and Knowledge-Based Creative Design. In *Modeling Creativity and Knowledge-Based Creative Design*. Psychology Press. https://doi.org/10.4324/9780203772881.

Goldberg, D.E., 1991. Genetic Algorithms as a Computational Theory of Conceptual Design. Appl. Artif. Intell. Eng. VI 3–16. https://doi.org/10.1007/978-94-011-3648-8_1.

Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A.C., Bengio, Y., 2014. Generative Adversarial Nets. Adv. Neural Inf. Process. Syst. (*27 (NIPS)* ⟨https://proceedings.neurips.cc/paper/2014/hash/5ca3e9b122f61f8f06494c97b1afccf3-Abstract.html⟩.

Gross, M.D., 1996. The Electronic Cocktail Napkin?a computational environment for working with design diagrams. Des. Stud. *17* (1), 53–69. https://doi.org/10.1016/0142-694X(95)00006-D.

Gu, N., Behbahani, P.A., 2021. Shape Grammars: A Key Generative Design Algorithm. Handb. Math. Arts Sci. 1–21. https://doi.org/10.1007/978-3-319-57072-3_7.

Helander, M.G., Khalid, H.M., Lim, T.Y., Peng, H., Yang, X., 2013. Emotional needs of car buyers and emotional intent of car designers. Theor. Issues Ergon. Sci. *14* (5), 455–474. https://doi.org/10.1080/1463922X.2012.656152.

Heyrani Nobari, A., Rashad, M.F., & Ahmed, F. (2021). *CreativeGAN: Editing Generative Adversarial Networks for Creative Design Synthesis*. https://doi.org/10.1115/DETC2021-68103.

Ho, J., Jain, A., & Abbeel, P. (2020). Denoising Diffusion Probabilistic Models. *ArXiv*.

Horváth, I., 2000. Conceptual design: inside and outside. Proc. 2nd Int. Semin. Workshop Eng. Des. Integr. Prod. 63–72. ⟨https://scholar.google.com/scholar_lookup?title=Conceptual⟩ design%3A inside and outside&publication_year=2000&author=I. Horváth.
Hsiao, S.-W., Chiu, F.-Y., Lu, S.-H., 2010. Product-form design model based on genetic algorithms. Int. J. Ind. Ergon. *40* (3), 237–246. https://doi.org/10.1016/J.ERGON.2010.01.009.
Hsiao, S.-W., Liu, M.C., 2002. A morphing method for shape generation and image prediction in product design. Des. Stud. *23* (6), 533–556. https://doi.org/10.1016/S0142-694X(01)00028-X.
Huet, A., Pinquié, R., Véron, P., Mallet, A., Segonds, F., 2021. CACDA: A knowledge graph for a context-aware cognitive design assistant. Comput. Ind. *125*, 103377 https://doi.org/10.1016/J.COMPIND.2020.103377.
Hyun, K.H., Lee, J.-H., 2018. Balancing homogeneity and heterogeneity in design exploration by synthesizing novel design alternatives based on genetic algorithm and strategic styling decision. Adv. Eng. Inform. *38*, 113–128. ⟨https://www.sciencedirect.com/science/article/pii/S1474034617305657#b0160⟩.
Isgrò, F., Ferraris, S.D., Colombo, S., 2022. AI-Enabled Design Tools: Current Trends and Future Possibilities. [.] Des.: Reinventing Des. Modes 2836–2847. https://doi.org/10.1007/978-981-19-4472-7_183.
Jansson, D.G., Smith, S.M., 1991. Design fixation. Des. Stud. *12* (1), 3–11. https://doi.org/10.1016/0142-694X(91)90003-F.
Jeon, Y., Jin, S., Shih, P.C., Han, K., 2021. Fashionq: an ai-driven creativity support tool for facilitating ideation in fashion design. Conf. Hum. Factors Comput. Syst. - Proc. https://doi.org/10.1145/3411764.3445093.
Jiaoying, S., Feng, L., Ning, Z., 1987. Artificial intelligence in computer aided design. Comput. Ind. *8* (4), 277–282. https://doi.org/10.1016/0166-3615(87)90049-2.
Jin, Y., Benami, O., 2010. Creative patterns and stimulation in conceptual design. AI EDAM *24* (2), 191–209. https://doi.org/10.1017/S0890060410000053.
Karimi, P., Rezwana, J., Siddiqui, S., Lou Maher, M., Dehbozorgi, N., 2020. Creative sketching partner: an analysis of human-AI Co-Creativity. Proc. 25th Int. Conf. Intell. Use Interfaces *10*. https://doi.org/10.1145/3377325.
Karras, T., Laine, S., Aila, T., 2018. A style-based generator architecture for generative adversarial networks. *2019 IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, *2019-June* 4396–4405. https://doi.org/10.1109/CVPR.2019.00453.
Kato, N., Muramatsu, N., Osone, H., Ochiai, Y., Sato, D., 2018. DeepWear: a case study of collaborative design between human and artificial intelligence. TEI 2018 - Proc. 12th Int. Conf. Tangible, Embed., Embodied Interact., 2018-Jan. 529–536. https://doi.org/10.1145/3173225.3173302.
Kielarova, S.W., Pradujphongphet, P., Bohez, I.L.J., 2013. An approach of generative design system: jewelry design application. IEEE Int. Conf. Ind. Eng. Eng. Manag. 1329–1333. https://doi.org/10.1109/IEEM.2013.6962626.
Ko, H.-K., Park, G., Jeon, H., Jo, J., Kim, J., & Seo, J. (2022). *Large-scale Text-to-Image Generation Models for Visual Artists' Creative Works*. 15. https://doi.org/10.1145/3581641.3584078.
Lawson, B., Loke, S.M., 1997. Computers, words and pictures. Des. Stud. *18* (2), 171–183. https://doi.org/10.1016/S0142-694X(97)85459-2.
Liu, H., Xu, Y., Chen, F., 2023. Sketch2Photo: synthesizing photo-realistic images from sketches via global contexts. Eng. Appl. Artif. Intell. *117*, 105608 https://doi.org/10.1016/j.engappai.2022.105608.
Liu, V., Vermeulen, J., Fitzmaurice, G., Matejka, J., 2022. 3DALL-E: Integr. Text. -to-Image AI 3D Des. Work. *18* https://doi.org/10.48550/arxiv.2210.11603.
Lubart, T., 2005. How can computers be partners in the creative process: Classification and commentary on the Special Issue. Int. J. Hum. -Comput. Stud. *63* (4–5), 365–369. https://doi.org/10.1016/j.ijhcs.2005.04.002.
MacCallum, K.J., 1990. Does intelligent CAD exist? Artif. Intell. Eng. *5* (2), 55–64. https://doi.org/10.1016/0954-1810(90)90002-L.
Maher, M.Lou, Tang, H.H., 2003. Co-evolution as a computational and cognitive model of design. Res. Eng. Des. *14* (1), 47–64.
McCormack, J.P., Cagan, J., Vogel, C.M., 2004. Speaking the Buick language: capturing, understanding, and exploring brand identity with shape grammars. Des. Stud. *25* (1), 1–29. https://doi.org/10.1016/S0142-694X(03)00023-1.
McKay, A., Ang, M.C., Chau, H.H., de Pennington, A., 2006. Combining Evolutionary Algorithms And Shape Grammars To Generate Branded Product Design. In *Design Computing and Cognition '06*. Springer, Netherlands, pp. 521–539. https://doi.org/10.1007/978-1-4020-5131-9_27.
Mekern, V., Hommel, B., Sjoerds, Z., 2019. Computational models of creativity: a review of single-process and multi-process recent approaches to demystify creative cognition. In: In *Current Opinion in Behavioral Sciences*, Vol. 27. Elsevier Ltd., pp. 47–54. https://doi.org/10.1016/j.cobeha.2018.09.008
midjourney.com. (n.d.). Midjourney.com. Retrieved May 25, 2023, from ⟨https://www.midjourney.com/⟩.
Mohammadi, S.O., Kalhor, A., 2021. Smart Fash.: A Rev. AI Appl. Fash. Appar. Ind. https://doi.org/10.48550/arxiv.2111.00905.
Mothersill, P., Bove, V.M., 2018. An Ontology of Computational Tools for Design Activities. DRS Bienn. Conf. Ser. *3*, 25–28. https://doi.org/10.21606/DRS.2018.456.
Mothersill, P., Bove, V.M., 2019. Beyond average tools. On the use of 'dumb' computation and purposeful ambiguity to enhance the creative process. Des. J. *22* (sup1), 1147–1161. https://doi.org/10.1080/14606925.2019.1594981.
Mountstephens, J., Teo, J., 2020. Progress and challenges in generative product design: A review of systems. Computers *9* (4), 1–23. https://doi.org/10.3390/COMPUTERS9040080.
openai.com. (n.d.). DALL·E 2. Retrieved May 25, 2023, from ⟨https://openai.com/product/dall-e-2⟩.
Oppenlaender, J., 2022. The Creativity of Text-to-Image Generation. Proc. 25th Int. Acad. Mindtrek Conf. https://doi.org/10.1145/3569219.3569352.
Orsborn, S., Cagan, J., Pawlicki, R., Smith, R.C., 2006. Creating cross-over vehicles: Defining and combining vehicle classes using shape grammars. AI EDAM *20* (03). https://doi.org/10.1017/S0890060406060185.
Paananen, V., Oppenlaender, J., Visuri, A., 2023. Using Text. -to-Image Gener. Archit. Des. Ideation. *ArXiv*, *abs/2304.1*. ⟨https://arxiv.org/pdf/2304.10182.pdf⟩.
Padiyath, A., Magerko, B., 2021. Des ainer: Exploring the use of "bad" generative adversarial networks in the ideation process of fashion design. ACM International Conference Proceeding Series. https://doi.org/10.1145/3450741.3466636.
Park, T., Liu, M.Y., Wang, T.C., Zhu, J.Y., 2019. Semantic Image Synthesis With Spatially-Adaptive Normalization. *2019 IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, *2019-June* 2332–2341. https://doi.org/10.1109/CVPR.2019.00244.
Permuter, H.H., Francos, J.M., Jermyn, I.H., 2006. A study of Gaussian mixture models of color and texture features for image classification and segmentation. Pattern Recognit. https://doi.org/10.1016/J.PATCOG.2005.10.028.
Pineda, L.A., 1993. On computational models of drafting and design. Des. Stud. *14* (2), 124–156. https://doi.org/10.1016/0142-694X(93)80044-D.
Pugliese, M.J., Cagan, J., 2002. Capturing a rebel: modeling the Harley-Davidson brand through a motorcycle shape grammar. Res. Eng. Des. https://doi.org/10.1007/s00163-002-0013-1.
Radhakrishnan, S., Bharadwaj, V., Manjunath, V., Srinath, R., 2018. Creative intelligence – Automating car design studio with generative adversarial networks (GAN). *Lect. Notes Comput. Sci. (Incl. Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinforma.)*, *11015 LNCS* 160–175. https://doi.org/10.1007/978-3-319-99740-7_11.
Rombach, R., Blattmann, A., Lorenz, D., Esser, P., Ommer, B., 2021. High-Resolution Image Synthesis with Latent Diffusion Models. 2022 IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR) 10674–10685. https://doi.org/10.1109/CVPR52688.2022.01042.
Saharia, C., Chan, W., Saxena, S., Li, L., Whang, J., Denton, E.L., Ghasemipour, S.K.S., Ayan, B.K., Mahdavi, S.S., Lopes, R.G., Salimans, T., Ho, J., Fleet, D.J., Norouzi, M., 2022. Photorealistic text-to-image diffusion models with deep language understanding. ArXiv. https://doi.org/10.48550/ARXIV.2205.11487.
Sarkar, P., Chakrabarti, A., 2011. Assessing design creativity. Des. Stud. *32* (4), 348–383. https://doi.org/10.1016/j.destud.2011.01.002.
Sohl-Dickstein, J.N., Weiss, E.A., Maheswaranathan, N., & Ganguli, S. (2015). Deep Unsupervised Learning using Nonequilibrium Thermodynamics. *ArXiv*.
stability.ai (n.d.). Stable Diffusion. Retrieved May 25, 2023, from ⟨https://stability.ai/blog/stable-diffusion-public-release⟩.
Stiny, G., Gips, J., 1972. Shape grammars and the generative specification of painting and sculpture. Inf. Process. 71 Proc. IFIP Congr. *2*, 1460–1465.
Sutherland, I.E., 1964. Sketchpad a man-machine graphical communication system. SIMULATION *2* (5), R-3–R-20. https://doi.org/10.1177/003754976400200514.
Tay, F.E.H., Gu, J., 2002. Product modeling for conceptual design support. Comput. Ind. *48* (2), 143–155. https://doi.org/10.1016/S0166-3615(02)00014-3.
Tholander, J., Jonsson, M., 2023. Design ideation with ai - sketching, thinking and talking with generative machine learning models. Proc. 2023 ACM Des. Interact. Syst. Conf. 1930–1940. https://doi.org/10.1145/3563657.3596014.
Tufarelli, M., Cianfanelli, E., 2022. Generative product design processes: humans and machines towards a symbiotic balance. Proc. Des. Soc. *2*, 1787–1794 https://doi.org/DOI: 10.1017/pds.2022.181.
van Dijk, C.G.C., 1995. New insights in computer-aided conceptual design. Des. Stud. *16* (1), 62–80. https://doi.org/10.1016/0142-694X(95)90647-X.
Vuletic, T., Duffy, A., Hay, L., McTeague, C., Pidgeon, L., Grealy, M., 2018. The challenges in computer supported conceptual engineering design. Comput. Ind. *95*, 22–37. https://doi.org/10.1016/J.COMPIND.2017.11.003.
Woodbury, R., 1990. Searching for designs: paradigm and progress. Build. Environ. https://doi.org/10.1184/R1/6076178.v1.
Xin, H., Zhao, D., 2021. Trend of Artificial Intelligence Aided Industrial Design. Lect. Notes Netw. Syst. *261*, 420–428. https://doi.org/10.1007/978-3-030-79760-7_50/FIGURES/4.
Yang, M., Jiang, P., Zang, T., Liu, Y., 2023. Data-driven Intell. Comput. Des. Prod.: Method, Tech., Appl.
Yüksel, N., Börklü, H.R., Sezer, H.K., Canyurt, O.E., 2023. Review of artificial intelligence applications in engineering design perspective. Eng. Appl. Artif. Intell. *118*, 105697 https://doi.org/10.1016/J.ENGAPPAI.2022.105697.
Zhang, L., Agrawala, M., 2023. Adding Conditional Control to Text-to-Image Diffusion Models. ArXiv. https://doi.org/10.48550/arxiv.2302.05543.
Zhou, R., Jiang, C., Xu, Q., 2021. A survey on generative adversarial network-based text-to-image synthesis. Neurocomputing *451*, 316–336. https://doi.org/10.1016/J.NEUCOM.2021.04.069.
Zhu, S., Fidler, S., Urtasun, R., Lin, D., Loy, C.C., 2017. Be Your Own Prada: Fashion Synthesis with Structural Coherence. 2017 IEEE Int. Conf. Comput. Vis. (ICCV) 1689–1697. ⟨https://arxiv.org/abs/1710.07346⟩.