

A Data-Driven Interactive System for Aerodynamic and User-centred Generative Vehicle Design

Muhammad Usama

Department of Mechanical Engineering
HITEC University
Taxila, Pakistan
musamaa482@gmail.com

Aqib Arif

Department of Mechanical Engineering
HITEC University
Taxila, Pakistan
aqibarifsadozai@gmail.com

Farhan Haris

Department of Mechanical Engineering
HITEC University
Taxila, Pakistan
Fh.haris1234@gmail.com

Shahroz Khan

*Department of Naval Architecture,
Ocean and Marine Engineering*
University of Strathclyde
Glasgow, United Kingdom
shahroz.khan@strath.ac.uk

S. Kamran Afaq

Department of Mechanical Engineering
HITEC University
Taxila, Pakistan
kamran.afaq@hitecuni.edu.pk

Shahrukh Rashid

Department of Mechanical Engineering
University at Buffalo
Buffalo, United States
srashid@buffalo.edu

Abstract—In this work, we propose a data-driven design pipeline for quick design exploration of performance and appearance guided alternatives for vehicle design. At the heart of our system is a machine learning-based generative design method to provide users with a set of diverse optimal design alternatives and an interactive design technique to induce users' preference into the design exploration. The generative design method is structure on two search process, qualitative and quantitative. To avoid the curse of dimensionality, the qualitative search process first builds up a lower-dimensional representation of a given design space, which is then explored using the unsupervised k-means clustering to synthesise a representative set of user-preferred designs. The quantitative search process explores the design space to find an optimal design in terms of performance criterion such as drag coefficient. To reduce the computational complexity, instead of evaluating drag via Computational Fluid Dynamics simulations, a surrogate model is developed to predict the drag coefficients. The designs generated after the generative design step are presented to the user at the interactive step, where potential regions of the design space are identified around the user-selected designs. Afterwards, a new design space is generated by removing the non-preferred regions, which helps to focus the computational efforts on exploring the user preferred regions of the design space for a design tailored to the user's requirements. We demonstrated the performance of the proposed approach on a two-dimensional side silhouette of a sport-utility vehicle.

Index Terms—Generative Design, Interactive Design, Machine Learning, Computational Fluid Dynamics, Drag Coefficient

I. INTRODUCTION

To date, vehicle form design has been an important and critical element of the automotive industry. People's enthusiasm for automotive vehicles, especially cars, has forced manufacturers to come up with the latest designs almost every year. In a world where competition never seems to end, manufacturers put all their efforts to create designs that can catch the attention of maximum customers. Although the newest technology and functional features interest the customers, the most important characteristic is the vehicle's

Aesthetics (i.e., its physical appearance). Therefore, at the design phase, a vehicle shape should not only be optimised according to its performance or functionality but also based on its external appearance [1]. However, design aesthetics and its performance can be conflicting criteria, as a design may have a better appearance but poor performance or vice versa.

Therefore, engineers and designers need to work together from the conceptual phase, where designers enhance the design's appearance and engineers work on meeting the performance criteria. This collaboration between designers and engineers at the conceptional phase usually results in a compromised solution/design that can achieve the maximum satisfaction of its target customers in terms of appearance and performance, thus ensuring a prosperous market place. In order to expedite this process, it essential to have a robust computational design tool to facilitate designers and engineers at the conceptual vehicle design stage to synthesise a diverse set of design ideas, which are driven by designers' aesthetic perception and engineers' performance requirements, for a given exemplar design or sketch. Therefore, in this paper, we aim to take the next step in the computer-aided preliminary vehicle design by introducing a computational design pipeline structured on machine learning-based interactive and generative techniques. On the one hand, generative design provides a promising way to effectively explore design space to create alternative designs based on the specific performance objective defined by a user. On the other hand, interactive design techniques help users to involve their aesthetic judgement during the design formulation.

Recently, interactive design tools [2] gained popularity to involve designers in the computational design process. In interactive design, designers are involved at each step of the design exploration, thus guiding the search process towards the exploration of user-preferred regions of design space. The interaction with the designer facilitates the generation of a

solution that incorporates his/her intuition without explicitly codifying them into the design process. Whereas generative design [3] is an evolutionary goal-oriented design process, which facilitates novice and experienced designers to generate designs, which otherwise may not be possible to be manually created. Instead of a single solution, the generative design makes potentially various solutions satisfying the given design requirements and constraints, thereby facilitating designers with the comfort of selecting a solution that best satisfies their needs. Even for the most experienced designers, their intuition can be limited when manually exploring an unprecedented ample design space.

The remainder of the paper is organised as follows: section II and III give a comprehensive review of the related works and overview of the proposed method, respectively. Section IV elaborates the proposed iterative and generative design techniques along with the formulation of the drag coefficient prediction model. The numerical results of the proposed technique are given in Section V. Finally, concluding remarks and opportunities for future work are presented in Section VI.

II. LITERATURE REVIEW

In this section, we discuss some of the significant previous works in generative and interactive techniques, which are explicitly proposed for computational vehicle design or used vehicle design as a test case for their system's Validation. Here, we also highlight some other recent studies on user-centred vehicle design and manufacturing.

A. Interactive Vehicle Design

Interactive design tools have gained popularity to involve users in the computational design exploration of CAD shapes. Mostly, the interaction with the users is carried out either with interactive interfaces [4]–[7] or with the integration of meta-heuristic optimisation [2], [8]–[11] to semi-automate the exploration process. In the latter approaches, users are involved at each iteration of an optimiser and guide the optimisation process towards exploring user-centred designs. Interactive Genetic Algorithms (IGA) [12], [13] are good examples of the most commonly used interactive meta-heuristic. In IGA, a user guides the optimisation process towards exploring promising regions of the design space by interactively involving each iteration/generation. An initial population of randomly sampled designs is first created, and the user performs interaction either by selecting the desired design [10], [14] or by rating all designs [12]. An iteration is then performed to generate designs similar to the selected or highly-rated design(s) and the creation of a similar design is usually done based on a distance-based metric [10]. This interactive-iterative process continues until the desired design is created/searched.

However, as in IGA, designs generated after an iteration are based on the user's selection in the previous step, therefore, starting the interactive process with the randomly generated designs, which are mostly clustered and can restrain the user from exploring a diverse set of designs [13]. Furthermore, distance-based exploration criterion can force the optimiser

to converge to similar designs at a fast rate, which leaves potential designs unexplored. An IGA-based system was proposed by Cluzel et al. [8] for 2D vehicle design, which creates designs as genotypes using a Fourier decomposition. Cluzel et al. proposed four different operators to converge towards the user-centred designs. As no diversity metric was integrated during the exploration process, therefore, the design flexibility of Cluzel et al.'s system is less compared to the proposed one. The ability of [8] in terms of exploring diverse design is also limited as the interactive process starts with randomly generated designs.

B. Generative Vehicle Design

Limited creativity has forced designers to automate the design process. Generative design is one of the alternatives to traditional design techniques. Until now, many advancements have been made in the field of generative design, and it always seems to be evolving. Here, we mention some of the most relevant works, which proposed generative design systems for parametric design exploration of vehicle silhouettes.

Orsborn et al. [15] proposed a shape grammar based generative system for 2D vehicle design. Compare to parametric design, the shape grammar-based exploration can generate diverse design only if appropriate grammar rules are set, otherwise, infeasible designs are likely to be generated. Moreover, the process of setting the grammar rules is cumbersome, especially for novice users [1]. Hyun and Lee [16] proposed a design exploration technique based on a genetic algorithm to synthesise design alternatives for existing cars shapes based on a similarity metric, which was learned from styling strategy of different car brands. Recently, Dogan et al. [17] proposed a generative design system, which samples design alternative for a given product profile using similarities and primitive shapes. To create diverse designs, [17] utilised Modified Hausdorff distance between the biarc approximations of the alternatives. However, in [17] no performance metric was integrated during the design exploration. A particle tracing-based shape exploration approach was proposed by Gunpinar and Gunpinar [18]. [18] considered designs as particles and algorithms trace the Latin-hypercube designs in the design space. Despite being efficient techniques, [3] and [18] can only create a design on qualitative criteria i.e., designs' form appearance. In another study, Gunpinar et al. [19] proposed a performance-driven generative design system for parametric design exploration of car side silhouettes.

III. METHOD OVERVIEW

A complete overview of the proposed method is given in Fig. 1. The main components of the system are qualitative, quantitative search processes and design space refinement. The qualitative and quantitative search processes are part of the generative design method, whereas design space refinement helps to induce users' design intention into the exploration process at the interactive step.

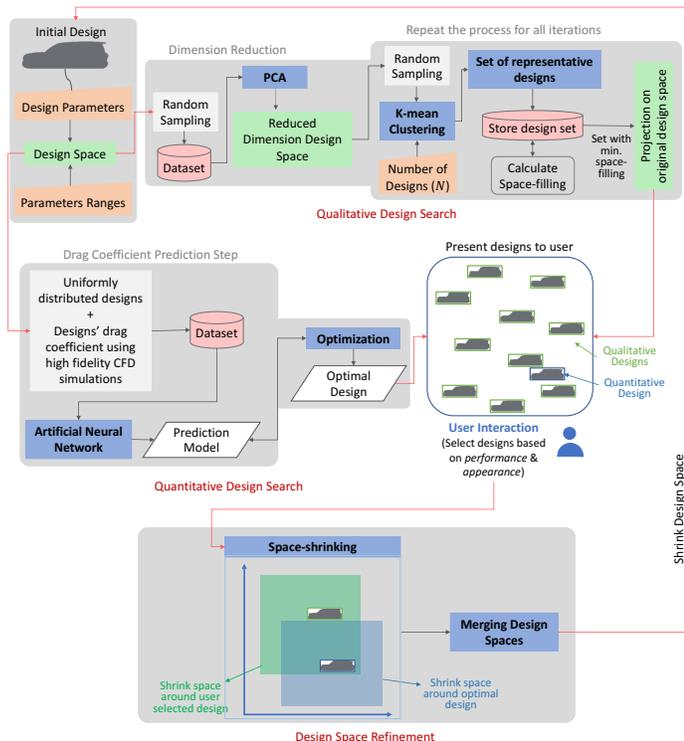


Fig. 1. The overall workflow of the proposed method. The whole process is divided into three parts, quantitative and qualitative search processes, representing generative design method, and design space refinement followed by the user design selection to converge the search process towards the user-centred and performance-oriented design.

A. Qualitative Search

The qualitative search process assists users to effectively explore a set of diverse design possibilities, thereby giving user the ability to start the interactive design process with the most suitable design. The objective of qualitative search process is to find a set of uniformly distributed N designs, where N is a user-defined parameter. This search process is based on Principal Component Analysis (PCA) [20], [21] and k-means clustering [22] approach. First, a dataset is created by randomly sampling a set of designs from a design space, which is formed with design parameters and their bounding limits. This dataset is then used to train PCA, which generates a lower-dimensional representation of the original design space. The reduced dimensional design space is referred to as subspace and the whole search process is carried in this space. To search represented designs, first designs are sampled from this subspace and inputted to a k-means cluster algorithm. This segments the sampled designs into N clusters and the centroid of each cluster is taken as a final set of designs. Fig. 2(a) and (b) shows the clustering of randomly generated designs and the centroid of the clusters, respectively. A space-filling property [23] of the centroid designs is also evaluated to measure their uniform distribution within the design space. The sampling and clustering process is repeated several times and the centroid set giving better space-filling is selected as

the final design. These designs are then shown to the user after projecting onto the original design space. Furthermore, it can be observed that compared to the randomly generated designs, which are shown in Fig. 2(c), designs generated by the proposed approach (see Fig. 2(b)) are evenly distributed covering all the design possibilities a given design space can offer. Therefore, the interactive process started with designs Fig. 2(c) can better scan the design space for potential designs.

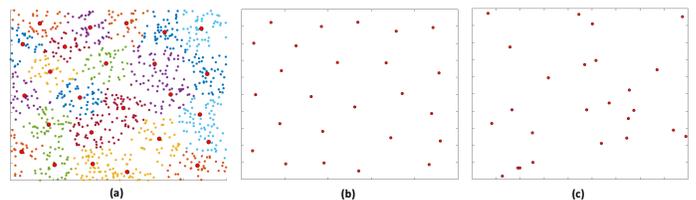


Fig. 2. (a) Clustered designs and (b) centroid of the cluster with the better space-filling property. The points in the plots represent the design in a two-dimensional space. Compare to randomly generated designs (a), the design generated by the proposed qualitative search process (b) are uniformly distributed.

B. Quantitative Search

During the design process, the decision on the selection of design is made on the basis of both design’s appearance and performance. Therefore, we also search the design with the help of an optimiser based on quantitative performance criterion such as drag coefficient (C_d). However, to find a quantitative/optimal design, one with a minimum C_d value, optimiser requires to evaluate C_d of various designs in each iteration. As the calculation of C_d via Computational Fluid Dynamics (CFD) is computationally expensive. Therefore, running these analyses during each interaction can result in higher waiting times, thereby causing user fatigue, which should be avoided as much as possible during the interactive design process [24]. Therefore, to evaluate C_d , a machine learning-based prediction model is developed and integrated with optimisation and also used to evaluate C_d of designs obtained from the qualitative search.

C. Design Space Refinement

In the interactive design process, the designs generated for each interaction should reflect the user’s design selection at the previous interaction, so at the end of the interactive process, the user is able to generate a preferred design. In this work, this is achieved by refining the input design space at each interaction while taking into account the user’s design preference. A Space Shrinking Technique (SST) [13] is utilised, which detects non-potential regions based on the selected designs and then removes these regions to create a new design space. In the proposed approach, the design space is not only shrunk around the user preferred designs but also based on the quantitative design, so that new space contains both potential regions related to user-preference and design’s physical performance. However, original SST works only for a single selected design, therefore, we proposed a space merging

step. At first design space is shrunk around each selected qualitative and quantitative design, and then all the design spaces are merged together to create a single shrunk space.

IV. PROBLEM FORMULATION

An initial parent design is generated for a given exemplar, which can be a 2D profile of an existing vehicle or sketches of a new design. Afterwards, its outer profile is detected and segmented into different features similar to [9], which divides the vehicle's side profile into different characteristic features such as the front bumper, grill, hood line, front windshield, roof, rear windshield, trunk, and back bumper. After the design segmentation, each feature is traced and parametrised with quadratic and cubic Bezier curves and the control points position of these curves, along x-axis and y-axis, are taken as design parameters to build the design space.

Let an initial design m is segmented into v design features and the j^{th} feature is represented with a Bezier curve $P(u)^j = \sum_{i=0}^{n_j} p_{i,j} b_{i,n_j}(u)$ of degree n_j , where $0 \leq u \leq 1$. Here, $p_{i,j}$ and $b_{i,n_j}(u)$, respectively, are its i^{th} control point and corresponding basis function, which is defined as $b_{i,n_j}(u) = C(n_j, i) u^i (1-u)^{n_j-i}$, where $C(n_j, i) = n_j! / (i!(n_j-i)!)$. In other terms, the design, m , is represented with a feature set $\mathcal{F} = \{P(u)^j, j = 1, 2, \dots, v\}$ consisting of v Bezier curves and each curve is either quadratic ($n_j = 2$) or cubic ($n_j = 3$) having three and four control points, respectively. The total number of control points e in \mathcal{F} are $e = \sum_{j=1}^v (n_j + 1)$ and as each control point is represented with two coordinate directions (x- and y-axis), therefore, the total number of design parameters are $p = 2e$. The parametric set for m is represented as $\mathbf{x}_m = \{x_{m,k}, k = 1, 2, \dots, p\} \in \mathcal{X} \subset \mathbb{R}^p$. Here, the design space, \mathcal{X} , is a subset of \mathbb{R}^p and is bounded by the lower \mathbf{x}_m^l and upper \mathbf{x}_m^u bounds of design parameters, such as $\mathcal{X} := \{x_{m,k}^l \leq x_{m,k} \leq x_{m,k}^u, \forall k \in \{1, 2, \dots, p\}\}$.

A. Qualitative Search Process

The objective for the qualitative search process is to generatively explore design space \mathcal{X} to find a qualitative set (\mathcal{N}), consisting of N uniformly distributed vehicle designs (i.e., $\mathcal{N} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\} \in \mathcal{X}$), where N is user-defined parameter and each design in \mathcal{N} represents a specific location in \mathcal{X} . To find \mathcal{N} , a set $\mathcal{P} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{z_p}]^T \in \mathcal{X} \subset \mathbb{R}^p$ consisting of z_p randomly sampled design from \mathcal{X} is generated using (1), whose columns represent p design parameters and rows are z_p sampled designs. In (1), \mathbf{x}_r is the r^{th} random design in \mathcal{P} generated between \mathbf{x}_m^l and \mathbf{x}_m^u of \mathcal{X} and $\text{rand}()$ is a random number between 0 and 1 drawn from a uniform distribution (i.e., $\text{rand}() \sim U(0, 1)$).

$$\mathbf{x}_r = (\mathbf{x}_m^u - \mathbf{x}_m^l) \times \text{rand}() + \mathbf{x}_m^l \quad (1)$$

First, \mathcal{P} is centred after subtracting the mean of the columns $\boldsymbol{\mu} = \{\mu_1, \mu_2, \dots, \mu_p\}$ of \mathcal{P} from each of its row, i.e., $\mathcal{P} = [(\mathbf{x}_1 - \boldsymbol{\mu}), (\mathbf{x}_2 - \boldsymbol{\mu}), \dots, (\mathbf{x}_{z_p} - \boldsymbol{\mu})]^T$. A lower-dimensional representation of \mathcal{X} , named as a subspace of \mathcal{X} , is created by applying PCA on \mathcal{P} . PCA is a data-driven dimension

reduction technique, which learns the lower dimension representation of design space by projecting dataset sampled from that space in a linear subspace, and the basis of this space is defined by the eigenvectors, also refer as principal components ($\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_p$), of the $[p \times p]$ covariance matrix $C = \frac{1}{z_p} \mathcal{P}^T \mathcal{P}$. The eigenvector obtained from this matrix has the property to maximise the geometric variance of designs in \mathcal{P} projected on them and to minimise the squared distance between the original designs and their relative projections. The principal components are defined by the solution of the eigenproblem given in (2).

$$\mathbb{C} \mathbf{w}_i = \lambda_i \mathbf{w}_i, \forall i \in \{1, 2, \dots, p\} \quad (2)$$

The eigenvalues $\Lambda = \lambda_i, i = 1, 2, \dots, p$ with $\lambda_i > \lambda_{i+1}$ represents the variance resolved along the relative eigenvector $\mathbf{W} = \{\mathbf{w}_i, i = 1, 2, \dots, p\}$ having $(\mathbf{w}_i^T \mathbf{w}_i = 1)$. The dimensionality of the subspace is lower than the original design. From this property of \mathbf{W} , a subset \mathbf{Z} consisting of first p' columns of \mathbf{W} are used to compute a reduced dimensional representation of a design $\mathbf{x}_m, \mathbf{d}_m = \mathbf{x}_m \mathbf{Z}$, where the matrix \mathbf{Z} has dimension $\text{dim}(\mathbf{Z}) = p \times p'$. Here, first p' columns of \mathbf{W} capture the α per cent of the geometric variance of the original set. The re-projection of \mathbf{d}_m on the original space is given as $\hat{\mathbf{x}}_m = \mathbf{d}_m \mathbf{Z}^T + \boldsymbol{\mu}$ and the projection $\hat{\mathbf{x}}_m$ represents the minimum squared error approximation of the relative \mathbf{x}_m ($\hat{\mathbf{x}}_m = \mathbf{x}_m \mathbf{Z} \mathbf{Z}^T + \boldsymbol{\mu}$). Similarly, to find the lower dimension representation of \mathcal{X} , its bounding limits \mathbf{x}_m^l and \mathbf{x}_m^u are project onto the subspace as $\mathbf{d}_m^l = \mathbf{x}_m^l \mathbf{Z}$ and $\mathbf{d}_m^u = \mathbf{x}_m^u \mathbf{Z}$. Afterwards, a subspace ($\text{subspace}(\mathcal{X})$) of \mathcal{X} is defined as in (3).

$$\text{subspace}(\mathcal{X}) := \{d_{m,k}^l \leq d_{m,k} \leq d_{m,k}^u, \forall k \in \{1, 2, \dots, p'\}\} \subset \mathbb{R}^{p'} \quad (3)$$

Afterwards, another dataset $\mathcal{D} = [d_1, d_2, \dots, d_{z_d}]^T \in \text{subspace}(\mathcal{X})$ is created by randomly sampling designs from the subspace of \mathcal{X} . Then k-means clustering is applied on \mathcal{D} , which divides it into N number of clusters, $S = \{S_1, S_2, \dots, S_N\}$. The centroid of each cluster is calculated by taking the mean of all designs in each cluster. Consider, an i^{th} cluster $S_i = \{d_{i,j}, j = 1, 2, \dots, \varphi_i\}$ contains φ_i designs. The centroid design, x_{s_i} , of S_i is found using $d_{s_i} = \frac{1}{\varphi_i} \sum_{j=1}^{\varphi_i} d_{i,j}$.

The centroid design for all N clusters are inputted to set $\mathcal{N}' = \{\mathbf{d}_{s_1}, \mathbf{d}_{s_2}, \dots, \mathbf{d}_{s_N}\}$ and projected back to the original design space to generate $\mathcal{N} = \{\mathbf{x}_{s_1}, \mathbf{x}_{s_2}, \dots, \mathbf{x}_{s_N}\}$ as $\mathcal{N} = \mathcal{N}' \mathbf{Z}^T + \boldsymbol{\mu}$. Furthermore, to check the uniform distribution of designs in \mathcal{N} , their space-fillingness ($\mathcal{K}(\mathcal{N})$) is calculated using (4), which follows a physical analogy: *Molecules in space exert repulsive forces on each other that lead to potential energy in a space. These molecules are in equilibrium in case of minimum potential energy* [23]. Therefore, minimising $\mathcal{K}(\mathcal{N})$ crates uniformly distributed designs.

$$\mathcal{K}(\mathcal{N}) = \sum_{p=1}^{N-1} \sum_{q=p+1}^N \frac{1}{\|x_p - x_q\|_2^2} \quad (4)$$

$\|x_p - x_q\|_2$ in (4) is the Euclidean-norm for design p and q over \mathbb{R}^P . This process is repeated for I iterations. At each iteration, z_d random designs are generated, and after implementing the k-means clustering, the centroid designs are obtained to create a new \mathcal{N} and its space-filling value is calculated. Therefore, at the end of interactions, we have I different sets ($\mathcal{N}_1, \mathcal{N}_2, \dots, \mathcal{N}_I$), and a set giving the minimum values of $\mathcal{K}(\mathcal{N})$ is selected as a final optimal one (\mathcal{N}_{op}). Finally, along with a quantitative design (\mathbf{x}_{op}), the designs in \mathcal{N}_{op} are presented to the user at the interactive design phase. From different experiments conducted in the study, it was observed that minimum values of z_p and z_d should be $50Np$ and $2Np$, respectively.

B. Quantitative Search Process

The quantitative search process is carried out to explore the design space for quantitative design (\mathbf{x}_{op}), which is optimal in term of its drag coefficient. The optimisation is performed with the help of a meta-heuristic optimisation technique, called Jaya Algorithm (JA) [25]. JA is selected because of its robustness and tuning parameter-free nature. It should be noted that the physical analyses, such as calculation of C_d for the vehicle, require high-fidelity CFD simulations to be performed. However, as mentioned before, running these computationally expensive analyses make the user wait for a long time before performing the next interaction, which may cause user fatigue or loss of concentration [24], thereby, resulting in directing the exploration process towards the non-preferred regions. Therefore, inspired by [19], we have developed an empirical prediction model using machine learning method, such as an artificial neural network (ANN), to predict the drag coefficient.

To construct this model, a database consisting of 500 uniformly distributed was generated. Afterwards, these designs were imported to a commercial solver, ANSYS®, as a planar object for the calculation of drag coefficient through CFD analyses, where the computational domain was setup as $16H \times 5L$. Here, L and H represents the overall length and height of the design. For boundary conditions the airspeed at the inlet is set at $40m/s$ and $0 Pa$ pressure is taken at outlet. A fine mesh of maximum relevance with proximity and curvature feature is created with inflation layers = 5 and growth Rate = 1.2 and the number of iterations for solution convergence are set to 1000. The final trained model is then integrated within the pipeline to calculate C_d for designs generated during quantitative search process but also for designs obtained at the end of the qualitative search.

C. Design Space Refinement

After performing I iterations the qualitative design set ($\mathcal{N}_{op} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N]$) and a quantitative design, \mathbf{x}_{op} , are obtained. The designs in \mathcal{N} are sorted in ascending order based on their C_d close to the \mathbf{x}_{op} . Afterwards, the sorted set $\mathcal{N}_{op} = [\mathbf{x}_{op}, \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N]$ containing both quantitative and qualitative designs are presented to the user for interaction, where he/she selects a single or multiple designs based on their performance and appearance. The sorted designs in \mathcal{N}_{op}

help the user to judge which design has performance similar to \mathbf{x}_{op} and aids the user to select a trade-off design between performance and appearance. After the user interaction, the design space \mathcal{X} is shrunk towards the user-selected design(s) and \mathbf{x}_{op} , thus generating a new design space for each of the selected design and for \mathbf{x}_{op} . These design spaces are merged together to recreate a new single design space. The design space refinement is carried using a Space Shrinking Technique [13], which creates a new design space using (5). Below, we first summarise the space shrinking technique for a single design then we explain our design space merging strategy.

Let at the T^{th} interaction, a design set, $\mathcal{N}_{op}^T = [\mathbf{x}_{op}, \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N]$, after performing I iterations and from \mathcal{N}_{op}^T , the user selects a t^{th} design \mathbf{x}_t , where T is an integer. After the user selection, a new design space, \mathcal{X}^T , is formed while shrinking the previous design space (\mathcal{X}^{T-1}) based on a selected design $\mathbf{x}_t = (x_{t,1}, x_{t,2}, \dots, x_{t,n})$ and used to generate designs for the next interaction. The shrinking of design space is performed by calculating new lower (\mathbf{x}_m^l) and upper (\mathbf{x}_m^u) bounds using (5) for $\mathcal{X}^T := \{x_{m,k}^l \leq x_{t,k} \leq x_{m,k}^u, \forall k \in \{1, 2, \dots, n\}\}$

$$\begin{cases} x_{m,k}^l = x_{m,k}^u + \left| \frac{\kappa^T}{T} \ln \left(\frac{x_{t,k}}{x_{m,k}^l} \right) \right|, \\ x_{m,k}^u = x_{m,k}^u + \left| \frac{\kappa^T}{T} \ln \left(\frac{x_{t,k}}{x_{m,k}^u} \right) \right|, \\ \forall k \in \{1, 2, \dots, n\} \end{cases} \quad (5)$$

Here, κ^T is the shrink rate initialised by the user in the T^{th} interaction and ranges between $0 < \kappa^T \leq 1$. When κ^T is zero, the space shrinking process terminates. Similarly, at the T^{th} interaction, new design spaces are created for all the selected designs and \mathbf{x}_{op} . Thereafter, these design spaces are merged together to create a new one for the next interaction.

1) *Design Space Merging*: Let us again consider that the user selects two design \mathbf{x}_{t1} and \mathbf{x}_{t2} from \mathcal{N}_{op}^T . Now three different design space, (\mathcal{X}_{t1}^T , \mathcal{X}_{t2}^T and \mathcal{X}_{op}^T) will be created, two around the user-selected designs and one around the qualitative design, \mathbf{x}_{op} , using (5). $[x_{t1,k}^l, x_{t1,k}^u]$ and $[x_{t2,k}^l, x_{t2,k}^u]$ are the lower and upper bounds of \mathcal{X}_{t1}^T , \mathcal{X}_{t2}^T and \mathcal{X}_{op}^T , respectively. Afterwards, three design spaces are merged together to create a single design space \mathcal{X}^T using (6), which takes the lowest value of lower limits, $x_{t1,k}^l$, $x_{t2,k}^l$ and $x_{op,k}^l$, and the highest value of upper limits, $x_{t1,k}^u$, $x_{t2,k}^u$ and $x_{op,k}^u$.

$$\begin{cases} x_{m,k}^l = \min(x_{t1,k}^l, x_{t2,k}^l, x_{op,k}^l), \\ x_{m,k}^u = \max(x_{t1,k}^u, x_{t2,k}^u, x_{op,k}^u), \\ \forall k \in \{1, 2, \dots, n\}. \end{cases} \quad (6)$$

V. RESULTS AND DISCUSSIONS

In this section, we first present the test case and results to validate the quantitative and qualitative search process via different experiments. Then, we provide some of the CFD results, followed by the Validation of our machine learning C_d prediction model and design space refinement technique.

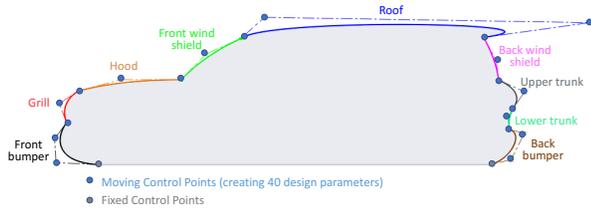


Fig. 3. Test case model of SUV used for the experiments in this work.

A. Example Design Case

A Sport Utility Vehicle (SUV) design, shown in Fig. 3, was selected to test the workability and reliability of the proposed system. The side profile of the SUV was segmented into nine features ($v=9$) and quadratic Bezier curves were used to parametrise front bumper, grill, front and rear windshields and trunk, whereas, roof, front and rear bumper were parametrised with cubic Bezier curves. The total number of design parameters are $p = 40$, which creates a 40-dimensional design space.

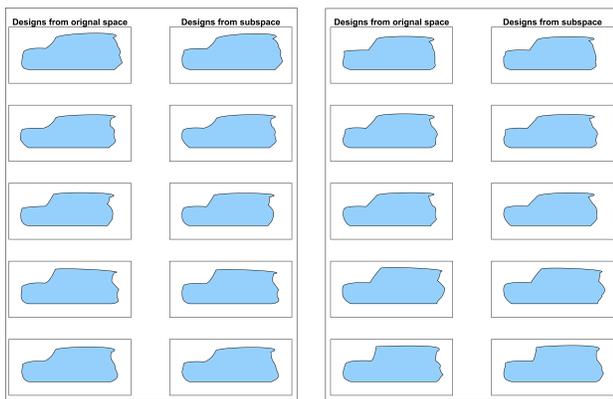


Fig. 4. Designs generated from the original 40-dimensional design space and their reconstruction from 24-dimensional subspace.

B. Results of Qualitative Search

The qualitative search process starts by creating a lower-dimensional representation of a given design space using PCA, which were applied on set \mathcal{P} consisting of $z_p = 5000$ randomly sampled designs and α was set to 90%. First, $p' = 24$ principal components captured the 90.23% of the geometric variance, resulting in the reduction of 16 dimensions. Afterwards, to check the reconstruction accuracy of trained PCA from 24-dimensional subspace to original 40-dimensional design space, we generated 10 designs from the original space and then reconstructed from the subspace, which is shown in Fig. 4. The left columns show the original designs, and the right columns shows their reconstruction from the subspace. It can be observed that these are very similar to each other.

Fig. 5 shows the designs generated from the proposed qualitative search. These designs were searched in a 24-dimensional subspace space after performing 30 iterations ($I = 30$) and were projected back to the original design space

for visualisation. It can be observed that these designs are distinct from each other, which can help users to start the interactive process with a design that meets his/her requirements. To further validate the proposed method, we also compared space-fillingness of the designs in Fig. 5 ($\mathcal{K}(\mathcal{N}) = 167.497$) with randomly generated designs ($\mathcal{K}(\mathcal{N}) = 190.018$). The lower space-filling value of designs generated by the proposed technique indicates that these are uniformly distributed on the design space compared to the randomly generated designs.

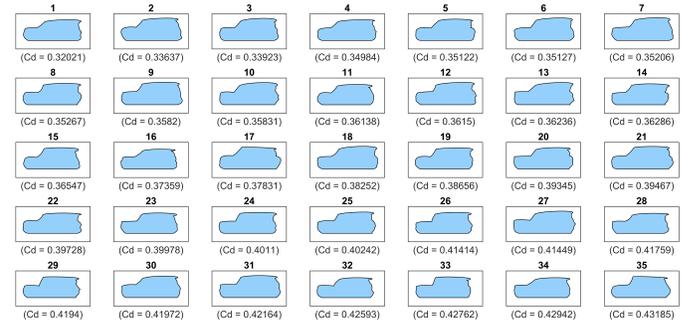


Fig. 5. Design alternatives generated using qualitative search process after performing $I=30$ iterations.

C. Results of Quantitative Search

In this subsection, we first validated the accuracy of the proposed machine learning-based drag coefficient prediction model. Afterwards, we provide the results of optimisation to find a quantitative design (one having a minimum C_d). The training of drag prediction model involves two stages. In the first stage, C_d of the designs in the training dataset was determined using the CFD Analyses and in the second stage training dataset was used to develop a prediction model, and its results were validated to that of obtained using ANSYS®.

After creating a dataset, the feed-forward neural network was used to build the prediction model. Dataset was partitioned into three parts, 70% of it was used as a training set, whereas 15% was used for Validation to measure network generalisation and to halt training when generalisation stops improving. The remaining 15% dataset was used as a test dataset, which has no effect on the training and provides an independent measure of the network during and after the training.

1) *Selection of Neural Network Architecture:* The selection of Neural Network architecture was done based on hit and trial method where the number of hidden layers and number of neurons was varied and the architecture producing the best performance was selected and integrated into our tool. The training was performed with the Levenberg-Marquardt algorithm. Various architectures were tested and their performance was analysed so that the best one among them could be selected for final use. A neural network with optimal performance is the one in which the plots are neither under fitted nor over-fit. Under-fitting occurs when the model has not learned enough from the training data and over-fitting occurs when it captures the noise of the data. The final selected

architecture consisted of three hidden layers containing 18, 16 and 14 neurons and its performance plot are shown in Fig. 6(a) and (b), respectively.

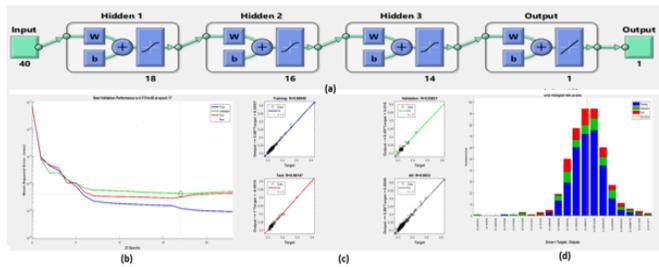


Fig. 6. (a) Selected Architecture (b) Performance Plot, (c) error Histogram and (d) regression Plots.

In Fig. 6(b), the small green circle indicates that the minimum mean squared-error occurred at the 17th iteration and the training did continue for 6 more iterations for validation check in search of lowest error before it was finally stopped at 23rd iteration. The trained model is more accurate if it has a smaller value of mean square error. Fig. 6(c) represents the Error Histogram showing the error range for the number of samples. At last, the plots of linear regression have been shown in Fig. 6(d), where the value of R ranging from 0 to 1 indicates the linear relationship between the outputs and targets with 0 being the worst and 1 being the best fit. The shown plot represents a very good fit as the value of R is 0.983, which is very close to 1.

2) *Validation of the selected model:* To further validate the model, we sampled $N=20$ designs and compared their actual C_d values (obtained from ANSYS®) with ones obtained from the prediction model. Fig. 7(a) and (b) shows their actual and predicted C_d values and absolute percentage error, respectively. From these figures, it can be seen that the highest and lowest Absolute Percent Error (E_{abs}) is 22.84 and 0.85, respectively, and the average E_{abs} is 9.85. This demonstrates that the prediction model can give a good estimation of C_d at the extensively reduced computational cost.

3) *Optimisation Results:* As mentioned previously, the prediction model was integrated with optimisation to search quantitative design (one having least C_d) from \mathcal{X} using the Jaya Algorithm. To start the optimisation, an initial population of 50 designs was created and 300 iterations (i.e., $I = 300$) were performed. Designs having least and maximum C_d during the optimisation process are shown in Fig. 8(a) and (b). It can be observed that the design with a smaller value of C_d has smooth transitions between edges, whereas the design with a large value of C_d has a boxy shape with sharp corners.

D. Results of Interactive Search

To test the performance of the interactive design technique, a design set \mathcal{N}_{op} (shown in Fig. 9) containing a quantitative design and $N = 20$ qualitative designs were created in each interaction after performing 30 iterations. These designs were presented to the user along with their predicted C_d values and

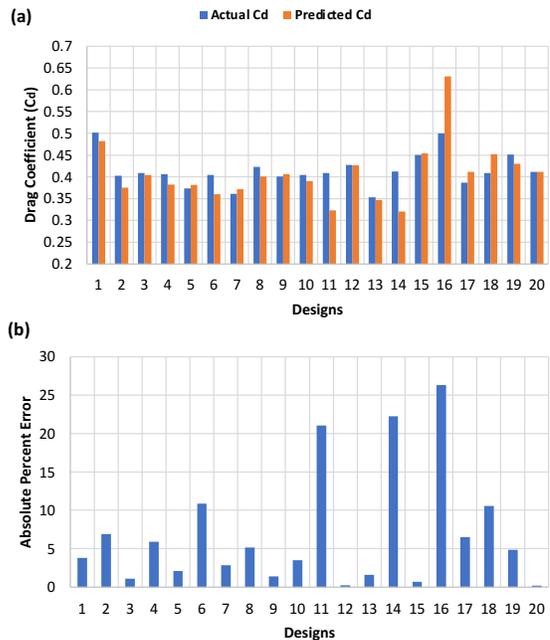


Fig. 7. (a) Plots showing (a) actual and predicted C_d values of designs and (b) their absolute per cent error.

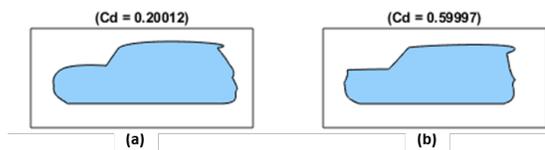


Fig. 8. Designs generated with a minimum (a) and maximum (b) C_d values during optimisation.

interactions were then proceeded with an objective to select a design having a trade-off between performance and appearance. Fig. 9 shows the designs generated in 6 interactions and the first design of each interaction is the quantitative design with least C_d value. For the first three interactions, space shrinking parameter λ was set to 0.5, 0.7, 0.8, respectively, and for the rest of the interactions, it was set to $\lambda = 1.0$. It can be observed that as the interaction process continues, the designs started to converge (i.e., get similar), in terms of appearance and C_d , to the selected designs. It is noteworthy that after certain interactions, the designs generated are a combination of both the selected and quantitative designs. It is because the design space created after each interaction is formed based on both qualitative and quantitative designs.

VI. CONCLUSIONS AND FUTURE WORK

A data-driven interactive-generative design exploration system was proposed to generate optimal and user-centred design alternatives of a given example vehicle silhouette. Features of a given example shape were first parametrised using cubic and quadratic Bezier Curves. For the parametric design, various diverse and optimal set of design alternatives are generated at the generative design step using a qualitative and quantitative

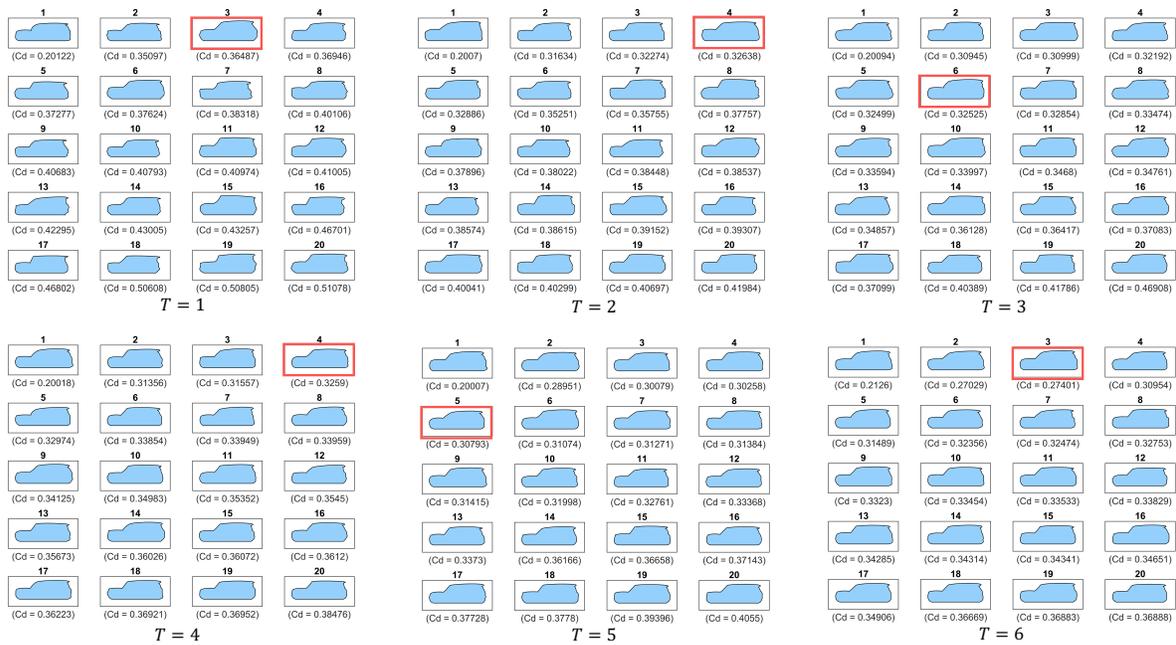


Fig. 9. Designs generated during 6 interactions between a user and the proposed system.

search process. To reduce the computational complexity at the generative step and to avoid user fatigue caused by long waiting time during the interactive process, a prediction model was developed using a feed-forward neural network and connected with qualitative and quantitative search to evaluate drag coefficients of the designs explored during these processes. The designs generated after the qualitative and quantitative search were shown to the user at the interactive step. From these designs, the user selects a single or multiple designs based on its performance and appearance. Afterwards, based on the selected designs space shrinking technique is applied to create a new design space for each of the selected and quantitative design by removing the non-potential regions. These design spaces were then merged together to generate a new space and used to create designs for the next interaction.

In the future, we would like to develop a topology optimisation-based generative model, which will eliminate the need for construction of an initial parametric model using Bezier Curves. The accuracy of the prediction model will be further enhanced by using deep learning and extrapolation models.

REFERENCES

- [1] S. Khan, E. Gunpinar, M. Moriguchi, and H. Suzuki, "Evolving a psycho-physical distance metric for generative design exploration of diverse shapes," *Journal of Mechanical Design*, vol. 141, no. 11, 2019.
- [2] J. Kelly, P. Y. Papalambros, and C. M. Seifert, "Interactive genetic algorithms for use as creativity enhancement tools," in *AAAI spring symposium: creative intelligent systems*, 2008, pp. 34–39.
- [3] S. Khan and M. J. Awan, "A generative design technique for exploring shape variations," *Advanced Engineering Informatics*, vol. 38, pp. 712–724, 2018.
- [4] A. Schulz, J. Xu, B. Zhu, C. Zheng, E. Grinspun, and W. Matusik, "Interactive design space exploration and optimization for cad models," *ACM Transactions on Graphics (TOG)*, vol. 36, no. 4, pp. 1–14, 2017.

- [5] N. Umetani, "Exploring generative 3d shapes using autoencoder networks," in *SIGGRAPH Asia 2017 Technical Briefs*, 2017, pp. 1–4.
- [6] N. Umetani and B. Bickel, "Learning three-dimensional flow for interactive aerodynamic design," *ACM Transactions on Graphics (TOG)*, vol. 37, no. 4, pp. 1–10, 2018.
- [7] A. Badías, S. Curtit, D. González, I. Alfaro, F. Chinesta, and E. Cueto, "An augmented reality platform for interactive aerodynamic design and analysis," *International Journal for Numerical Methods in Engineering*, vol. 120, no. 1, pp. 125–138, 2019.
- [8] F. Cluzel, B. Yannou, and M. Dihlmann, "Using evolutionary design to interactively sketch car silhouettes and stimulate designer's creativity," *Engineering Applications of Artificial Intelligence*, vol. 25, no. 7, pp. 1413–1424, 2012.
- [9] E. Gunpinar, S. E. Ovrur, and S. Gunpinar, "A user-centered side silhouette generation system for sedan cars based on shape templates," *Optimization and Engineering*, vol. 20, no. 3, pp. 683–723, 2019.
- [10] E. Poirson, J.-F. Petiot, L. Boivin, and D. Blumenthal, "Eliciting user perceptions using assessment tests based on an interactive genetic algorithm," *Journal of Mechanical Design*, vol. 135, no. 3, 2013.
- [11] E. Gunpinar and S. Khan, "A multi-criteria based selection method using non-dominated sorting for genetic algorithm based design," *Optimization and Engineering*, vol. 21, no. 4, pp. 1319–1357, 2020.
- [12] A. M. Brintrup, J. Ramsden, H. Takagi, and A. Tiwari, "Ergonomic chair design by fusing qualitative and quantitative criteria using interactive genetic algorithms," *IEEE Transactions on Evolutionary Computation*, vol. 12, no. 3, pp. 343–354, 2008.
- [13] S. Khan, E. Gunpinar, and B. Sener, "Genyacht: An interactive generative design system for computer-aided yacht hull design," *Ocean Engineering*, vol. 191, p. 106462, 2019.
- [14] S. Khan, E. Gunpinar, and M. Moriguchi, "Customer-centered design sampling for cad products using spatial simulated annealing," *Proceedings of CAD*, vol. 17, pp. 100–103, 2017.
- [15] S. D. Orsborn, R. Pawlicki, R. Smith, and J. Cagan, "Creating cross-over vehicles: Defining and combining vehicle classes using shape grammars," *Ai Edam*, vol. 20, no. 3, pp. 217–246, 2006.
- [16] K. H. Hyun and J.-H. Lee, "Balancing homogeneity and heterogeneity in design exploration by synthesizing novel design alternatives based on genetic algorithm and strategic styling decision," *Advanced Engineering Informatics*, vol. 38, pp. 113–128, 2018.
- [17] K. M. Dogan, H. Suzuki, E. Gunpinar, and M.-S. Kim, "A generative sampling system for profile designs with shape constraints and user evaluation," *Computer-Aided Design*, vol. 111, pp. 93–112, 2019.

- [18] E. Gunpinar and S. Gunpinar, "A shape sampling technique via particle tracing for cad models," *Graphical Models*, vol. 96, pp. 11–29, 2018.
- [19] E. Gunpinar, U. C. Coskun, M. Ozsipahi, and S. Gunpinar, "A generative design and drag coefficient prediction system for sedan car side silhouettes based on computational fluid dynamics," *Computer-Aided Design*, vol. 111, pp. 65–79, 2019.
- [20] H. Abdi and L. J. Williams, "Principal component analysis," *Wiley interdisciplinary reviews: computational statistics*, vol. 2, no. 4, pp. 433–459, 2010.
- [21] S. Khan, A. Serani, M. Diez, and P. Kaklis, "Physics-informed feature-to-feature learning for design-space dimensionality reduction in shape optimisation," in *AIAA Scitech 2021 Forum*, 2021, p. 1235.
- [22] J. A. Hartigan and M. A. Wong, "Algorithm as 136: A k-means clustering algorithm," *Journal of the royal statistical society. series c (applied statistics)*, vol. 28, no. 1, pp. 100–108, 1979.
- [23] S. Khan and E. Gunpinar, "Sampling cad models via an extended teaching–learning-based optimization technique," *Computer-Aided Design*, vol. 100, pp. 52–67, 2018.
- [24] S. Wang and H. Takagi, "Improving the performance of predicting users' subjective evaluation characteristics to reduce their fatigue in iec," *Journal of physiological anthropology and applied human science*, vol. 24, no. 1, pp. 81–85, 2005.
- [25] R. V. Rao, *Jaya: an advanced optimization algorithm and its engineering applications*. Springer, 2019.