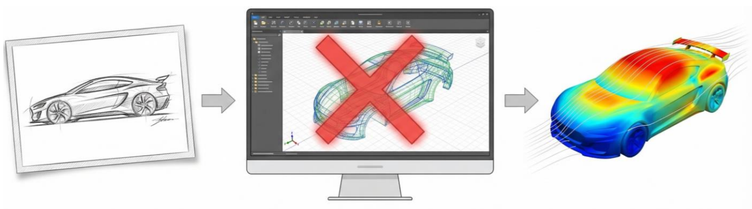
Vehicle Surface Pressure Prediction from 2D Sketches via a Pre-Trained Diffusion Model

Shunya Nakamura, Osamu Ito
Honda Motor Co., Ltd.

1 Introduction

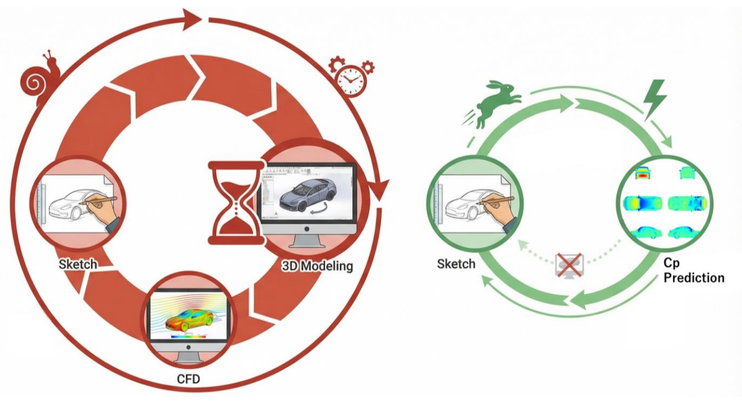
Problem

In automotive design, aesthetics and aerodynamics are both essential.
Aesthetics drive sales, while aerodynamics affect fuel efficiency and stability.
However, current evaluation methods require 3D geometry.
In the early concept stage, designers work with 2D sketches ,but cannot evaluate surface pressure since no 3D models exist yet.



Motivation

f aerodynamics could be assessed at the sketch stage, designers could make informed decisions early, when flexibility is highest.
This reduces costly revisions later.



2 Method

Problem Formulation

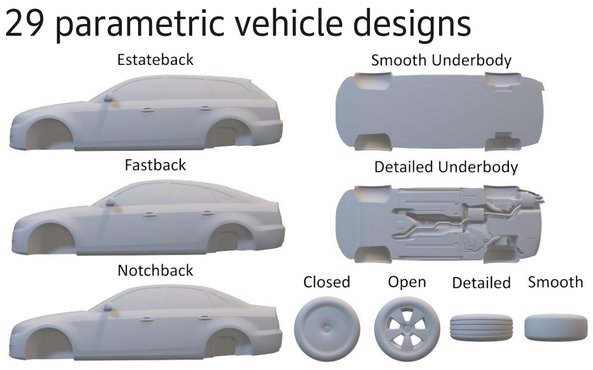
Learn a mapping from 2D sketches to surface pressure (Cp) distributions, both as 1024 × 1024 RGB images in a fixed 6-view grid.

Model

Qwen-Image-Edit-2511 (20B MM-DiT) with dual encoders (VLM + VAE), fine-tuned via LoRA.

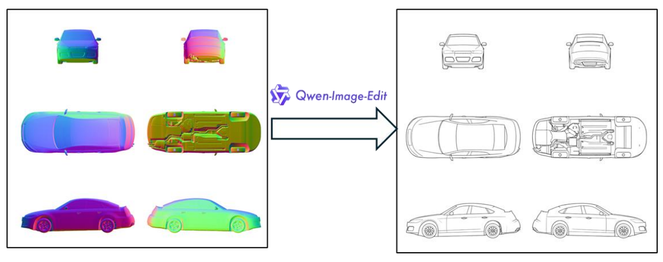
Dataset: DrivAerNet++ [1]

8,129 parametric vehicle designs



Sketch Generation

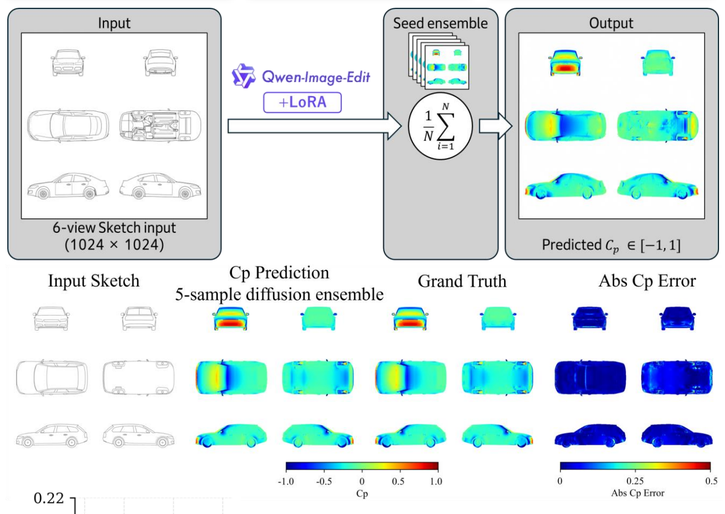
Multi-view sketches are generated from normal maps using Qwen-Image-Edit.

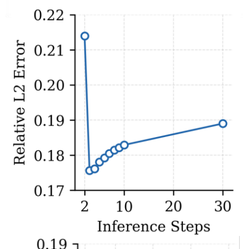


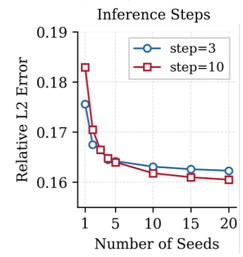
[1] Elrefaie et al., "DrivAerNet++: A Large-Scale Multimodal Car Dataset with CFD Simulations", NeurIPS Datasets & Benchmarks, 2024.

3 Key Results

0.165	0.955	0.0252
Rel L2	R^2	MAE





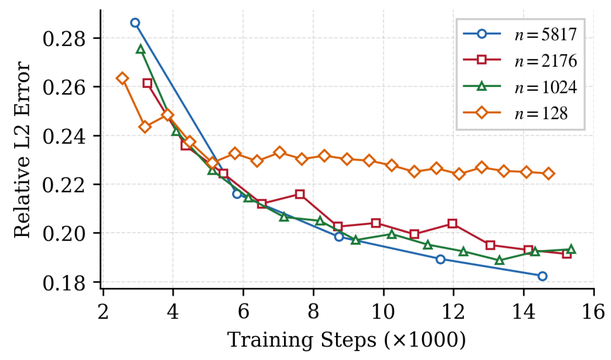


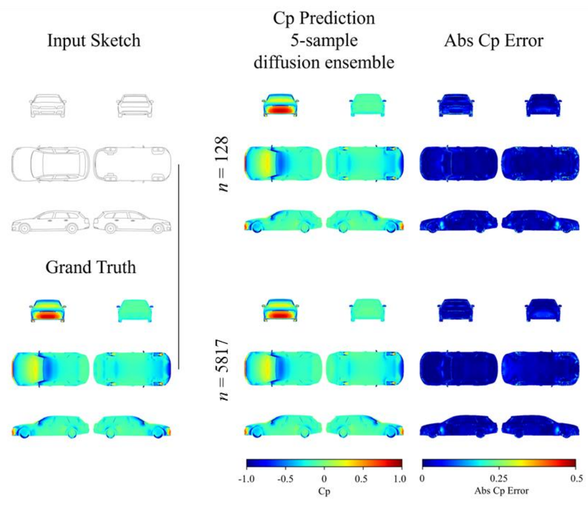
Setting	Value
Training samples	5,817
Hardware	RTX 6000 Ada × 2
Training time	~36h
Inference time	~13s/sample
LoRA rank	32
Diffusion steps	3
CFG scale	4.0
Ensemble size	5

Fewer diffusion steps perform better
→ steps=3 vs. 30 (Rel L2: 0.176 vs. 0.190)
Ensemble improves accuracy
→ 1 sample: 0.176
→ 5 samples: 0.164
→ Diminishing returns beyond 5

4 Data Efficiency

Pre-trained model + LoRA enables efficient learning.
Training data progressively reduced via stratified sampling preserving category balance.



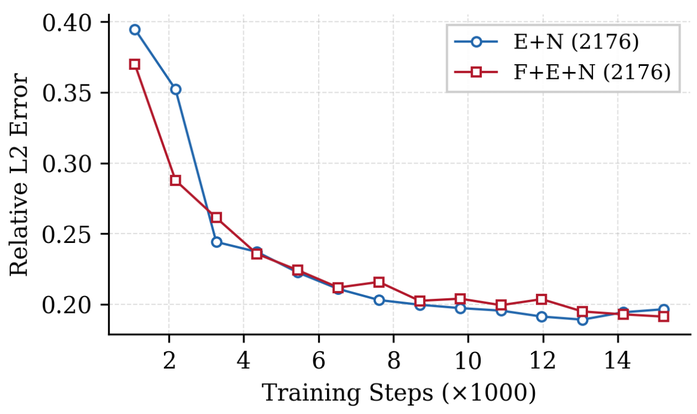


5 Cross-Category Generalization

DrivAerNet++ has 3 base categories:
Fastback (F), Estateback (E), Notchback (N).
We train on E+N only and evaluate on unseen Fastback.

For F_S (Smooth underbody), almost no degradation.
For F_D (Detailed underbody, absent from training), significant degradation.
Cross-category difficulty stems from absent shape features, not category labels.

Training	Test	Rel L2	R^2	MAE
E + N	F_S	0.186	0.943	0.030
E + N	F_D	0.410	0.856	0.066
All	All	0.180	0.943	0.029



6 Uncertainty Quantification

Diffusion Ensemble UQ

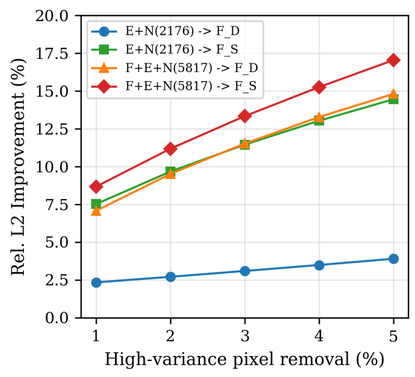
Multiple predictions from different random seeds provide uncertainty estimates via prediction variance.
We analyze correlation between ensemble standard deviation and absolute error across all pixels within vehicle body regions.

Variance Filtering

Excluding high-variance pixels improves accuracy. Removing 5% of high-variance pixels yields 14–17% Rel L2 improvement under in-distribution conditions.

Key Insight

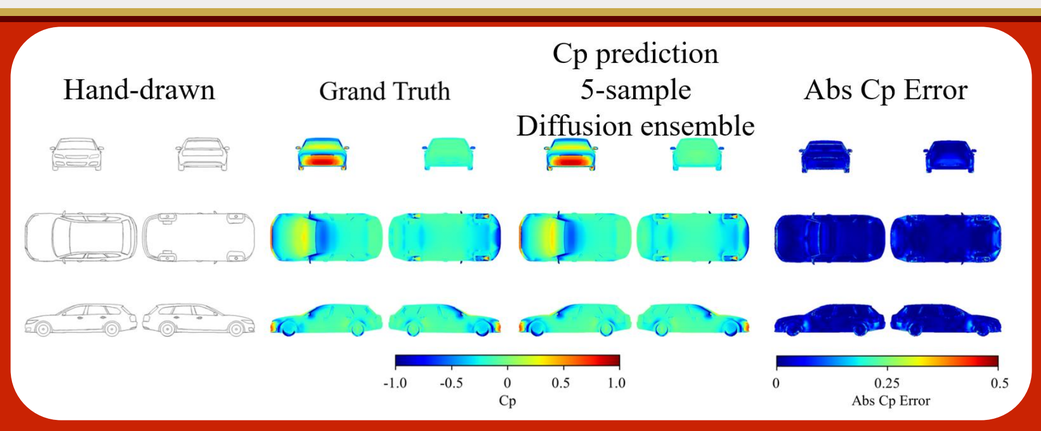
For shapes absent from training data (F_D with E+N training), the model "doesn't know what it doesn't know" — variance does not function as error predictor



7 Conclusion

We proposed an end-to-end framework utilizing Qwen-Image-Edit-2511 for predicting surface pressure coefficient distributions from 2D sketches, demonstrating aerodynamic evaluation without 3D geometry.

- train 5718 samples and 5-sample ensemble→Rel L2 = 0.165, R^2 = 0.955
- train 128 samples (2.2%) → Rel L2 = 0.218
- Hand-drawn sketch → Rel L2 = 0.219



Hand-drawn sketch → Cp prediction result