3D Super-Resolution Model for Vehicle Flow Field Enrichment

Thanh Luan Trinh[1,2,*] Fangge Chen[1,†] Takuya Nanri[1] Kei Akasaka[3]

[1]Mobility & AI Laboratory, Nissan Motor Co., Ltd.
[2]Yokohama National University, Japan
[3]Integrated CAE and PLM Department, Nissan Motor Co., Ltd.

luan.trinh.t@gmail.com {fanggechen, t-nanri, kei-akasaka}@mail.nissan.co.jp

Abstract

In vehicle shape design from an aerodynamic performance perspective, deep learning methods enable us to estimate the flow field in a short period. However, the estimated flow fields are generally coarse and of low resolution. Therefore, a super-resolution model is required to enrich them. In this study, we propose a novel super-resolution model to enrich the flow fields around the vehicle to a higher resolution. To deal with the complex flow fields of vehicles, we apply the residual-in-residual dense block (RRDB) as the basic network-building unit in the generator without batch normalization. We then apply the relativistic discriminator to provide better feedback regarding the lack of high-frequency components. In addition, we propose a distance-weighted loss to obtain better estimation in wake regions and regions near the vehicle surface. Physics-informed loss is used to help the model generate data that satisfies the physical governing equations. We also propose a new training strategy to improve learning effectiveness and avoid instability during training for our enrichment task. Experimental results demonstrate that the proposed method outperforms the previous study in vehicle flow field enrichment tasks by a significant margin.

1. Introduction

Flow fields, including velocity and pressure fields, help design vehicles with a desirable aerodynamic performance that is directly connected to steering stability and energy efficiency. For example, designers frequently use the pressure field to analyze the pressure distribution on a vehicle surface to lower the air resistance of the vehicle. Generally, flow fields are computed using computational fluid dynamics (CFD) simulations based on the Navier-Stokes equations [1] and the lattice Boltzmann methods [2]. CFD simulations developed over the years have produced highly accurate results and contributed to the development of the current automotive industry. However, a CFD simulation requires several days to complete and incurs a huge computational cost. With the increasingly competitive automotive industry, it is crucial to develop new designs within a short period. Therefore, a new approach with rapid computation speed is urgently needed.

Recently, several methods based on convolutional neural networks (CNNs) [3, 4, 5, 6, 7, 8, 9, 10, 11] have been used for flow field estimation. A typical feature of these methods is that they provide relatively accurate results within seconds with pre-trained models, which is a significant benefit over CFD simulation. Most flow field estimation methods use low-resolution data to train estimation models for hardware limitations. As a result, the output is also low resolution, making it difficult for vehicle designers to understand the more detailed aerodynamic performance and possible improvements of vehicles. Therefore, a super-resolution model is required to enrich the low-resolution flow fields around vehicles. Several CNN-based methods [12, 13, 14, 15, 16] have been also proposed for image super-resolution tasks. However, these methods are trained with L1/L2 loss functions so that the averaged loss values are used to update the model. This smooths the regions with high-frequency components which always provide larger loss values. Consequently, these methods have difficulty generating high-frequency components. To address this issue, both Ledig *et al*. [17] and Wang *et al*. [18] improved the high-resolution generation ability of high-frequency components by adding the mechanism of generative adversarial network (GAN [19]) and showed that the generator can receive feedback from the discriminator about lacking the high-frequency components in the generated results.

To deal with the flow field enrichment task, several super-resolution methods with GAN mechanism [20, 21, 22] have been proposed inspired by the aforementioned image super-resolution methods. These methods trained with

*Work done during an internship at Mobility & AI Laboratory, Nissan Motor Co., Ltd.
†Corresponding author.

the additional physics-informed loss functions designed from fluid mechanics such as the continuity equation. The results of these methods showed the potential of generating higher resolution flow fields that also satisfy the physical governing equations. However, the vehicle flow fields we handle have a larger size compared to the flow field data used in these methods. For this reason, we need to expand the super-resolution model to fit our data. Besides, the following three problems, more or less, within these methods also need to be addressed.

1. Using batch normalization [23] layers in the generator tends to create undesirable artifacts and noises on generated results of the test set [16, 18].
2. Using mean square error (MSE) as a loss function tends to cause an over-smoothing problem [8] and results in poor generation ability on areas near the vehicle surface since MSE considers each pixel equally.
3. Suddenly adding the discriminator into the training process after training the generator independently degrades training effectiveness since the learning purpose is also suddenly changed.

In this study, we propose a super-resolution model with GAN mechanism to enrich the flow fields around the vehicle to a higher resolution. We employ TEGAN (Subramaniam *et al*. [20]) as a baseline and aim to improve the generation abilities to address the aforementioned problems. Following the ESRGAN (Wang *et al*. [18]), we apply a residual-in-residual dense block (RRDB) without batch normalization layers as the basic network-building unit to avoid undesirable artifacts and make use of the relativistic discriminator (Alexia *et al*. [24]) to help the model compensate the lacking high-frequency components better. Before outputting the generated high-resolution flow fields, the upsampled results are multiplied by a binary mask to improve training effectiveness because the flow fields values inside the vehicles are always 0. Next, we propose distance-weighted loss instead of using mean absolute error (MAE) or MSE as a part of loss functions to obtain better results in wake regions around the rear of the vehicle and regions near the vehicle surface.

As mentioned above, the abrupt change by suddenly adding the discriminator into the training process may degrade training effectiveness. To address this, we propose a new training strategy for our flow field enrichment task called scheduled-learning. In this strategy, the importance of each loss function changes during training, and the overall learning purpose changes slowly to increase training effectiveness and stability.

Experimental results show that the proposed method outperforms the previous methods such as bicubic interpolation (Khaledyan *et al*. [25]) and other CNN-based methods

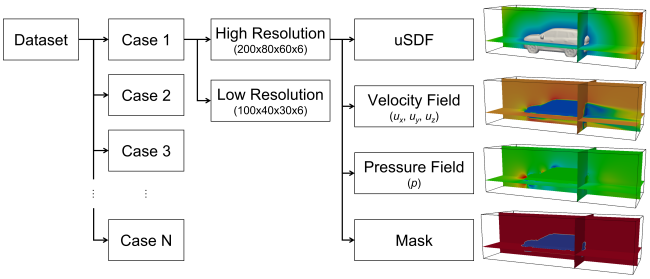

Figure 1. The structure of our dataset.

in vehicle flow field enrichment tasks. The output high-resolution flow fields were surprisingly improved by adding the proposed elements to the previous works, especially in high-frequency components of flow fields. Moreover, the experimental results showed the effectiveness of each proposed element.

To the best of our knowledge, this is the first attempt to enrich the 3D vehicle flow fields with a super-resolution model. It only takes 0.2 seconds to output the higher resolution flow fields from the lower ones which can be obtained from the CNN-based methods. Vehicle designers benefit from the proposed method to discuss the more detailed shape design in a second-unit period rather than wait for CFD simulation which needs at least 1 day in the current process.

2. Dataset

To evaluate the performance of the proposed method, following Chen *et al*. [8], we generated a dataset for our flow field enrichment task, which includes 1,121 cases of vehicles from many different categories such as sedans and SUVs. Figure 1 shows the structure of the dataset. We placed vehicles in a $10m \times 4m \times 3m$ Cartesian space and used the CFD software to estimate the flow fields. In the CFD simulations, the speed of the vehicles was set to 120 km/h. For high-resolution data, we sampled the aforementioned space with a 0.05 m pitch, which resulted in the size of the high-resolution data being $200 \times 80 \times 60$. For low resolution, the sample size was 0.1 m; therefore, the data size was $100 \times 40 \times 30$. Each case of high- or low-resolution data was combined with four components.

- **uSDF**(un-Signed Distance Function): A matrix containing the distance from an arbitrary point in space to the vehicle surface, which is used in the proposed new loss function introduced in Section 3.3.
- **Velocity Fields**: The speed of the air moving around the vehicle is expressed separately under the three components u_x, u_y, and u_z.
- **Pressure Fields**: Air pressure at each point in the space around the vehicle.

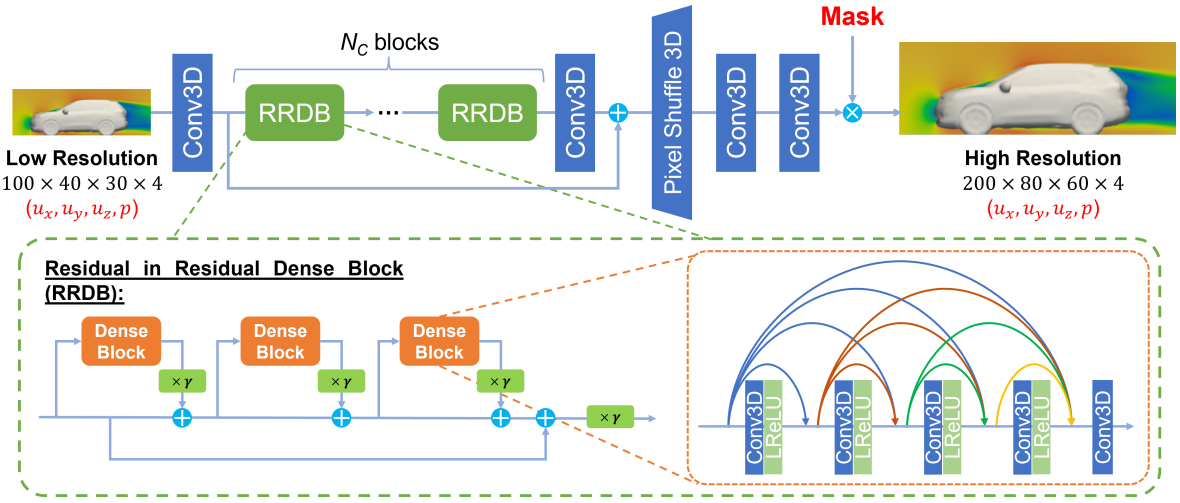

Figure 2. Architecture of the generator, including N_C RRDB blocks. Each RRDB contains N_{DB} dense blocks. γ is the residual scaling parameter. In this study, we set $N_C = 6$, $N_{DB} = 3$ and $\gamma = 0.2$.

- **Mask**: Binary mask for each vehicle, where 1 indicates the region that allows air to pass through and 0 means that the air cannot reach. In this study, the mask plays an important role in reducing the computational cost and improving the model's performance. The details are discussed in Section 3.

3. Proposed Method

Our main aim is to propose a super-resolution model for our vehicle flow field enrichment task. We apply TEGAN as the baseline and improve the generation ability by making use of our proposed method. In this section, we first describe our proposed network architecture. Next, we introduce our proposed loss function for improving the generation accuracy. Finally, we introduce a new training strategy to increase training effectiveness and stability.

3.1. Network Architecture

Figure 2 shows the architecture of the generator of the super-resolution model used in this work. Similar to the previous work TEGAN, we build our model based on SRResNet [17]. To adjust to our 3D flow field enrichment task, 2D convolutional layers are replaced with 3D convolutional (Conv3D) layers for handling the 3D flow field data. However, following Wang *et al*. [18], we also remove batch normalization layers from the both generator and discriminator to avoid undesirable artifacts and improve the generalization ability of the model, which is also able to save the computation cost because handling 3D data need a larger size of memory. We also borrow the idea of residual-in-residual dense block (RRDB) from [18]. The RRDB, which combines multilevel residual networks [26] and dense connections [27], helps the model extract more important features from low-resolution flow fields before upsampling them to high-resolution data with a 3D pixel shuffle layer [28]. The last two Conv3D layers in Figure 2 readjust the features and reduce the checkerboard artifact caused by the pixel shuffle layer. Finally, the generated high-resolution flow fields are multiplied by a binary mask element-wise. Because most parts inside the vehicle do not have air flowing through, adding a binary mask reduces the computational complexity to be equivalent to the size of the vehicle. Adding a binary mask also suppresses gradients in areas where the mask value is 0. Therefore, after a few initial epochs, the model primarily focuses on regions with a mask value of 1; this means that adding a mask indirectly increases the overall accuracy of the model.

3.2. Relativistic Average Discriminator

The content loss, which will be discussed in the next section, helps the model generate super-resolution data that are close to real data (high-resolution ground truth). However, training the model with the content loss only will cause the lack of high-frequency components. Schwarz *et al*. [29] found that discriminator can provide a meaningful training signal for the spectral statistics at high frequencies. Therefore, we train the model with GAN mechanism and use a discriminator to focus on the feedback about the flow field structure and the lack of high-frequency components. We employed the relativistic average discriminator (Alexia *et al*. [24]) for GAN training. In a standard GAN, the discriminator predicts whether an image is real or fake based solely on that image. By contrast, a relativistic average discriminator attempts to predict the probability that the given real

data are more realistic than fake data on average, which increases the ratio of the feedback about the incorrectly generated flow field structure and the lack of high-frequency components in the final feedback comparing with a standard discriminator.

3.3. Loss Function

In a word, there are three types of loss functions used for training the generator of the super-resolution model. First, a content loss ($L_{content}$) helps the generated high-resolution data gradually get close to the ground truth, which is usually MSE or MAE. In this study, we propose a new content loss called distance-weighted loss to improve the generation ability in the wake regions and regions near the vehicle surface. Second, we employ the relativistic average discriminator to provide an adversarial loss ($L_{adversarial}$) for improving generation ability of high-frequency components. In addition, a physical loss ($L_{physical}$) is added to ensure that the generated results obey the physical laws. Details are discussed in each subsection.

3.3.1 Distance-Weighted Loss

MSE and MAE, often used as the content loss for super-resolution models, are calculated as follows in our task:

$$L_{\text{MSE}} = \frac{1}{\text{WHL}} \sum_{x,y,z=0}^{\text{W, H, L}} \left(\mathbf{I}^{\text{HR}}_{x,y,z} - \mathbf{I}^{\text{SR}}_{x,y,z}\right)^2 \tag{1}$$

$$L_{\text{MAE}} = \frac{1}{\text{WHL}} \sum_{x,y,z=0}^{\text{W, H, L}} \left|\mathbf{I}^{\text{HR}}_{x,y,z} - \mathbf{I}^{\text{SR}}_{x,y,z}\right| \tag{2}$$

where W, H, L denote the width, height, and length of the flow field space respectively. As we can see from these two equations, both MSE and MAE consider the weight of each pixel is equivalent; this may be useful for image super-resolution tasks because the importance of each pixel to be equivalent. However, in vehicle flow field enrichment tasks, the flow fields near the vehicle surface and wake regions are more important than those in the other regions. Furthermore, the computational cost can be reduced by ignoring airless areas. To adjust this problem, we use the binary mask and uSDF described above to define a new content loss called distance-weighted loss as follows:

$$L_{content} = \frac{1}{\text{WHL}} \sum_{x,y,z=0}^{\text{W, H, L}} \frac{\alpha_1 \text{mask}_{x,y,z}}{\alpha_2 + \text{uSDF}_{x,y,z}} \left(\mathbf{I}^{\text{HR}}_{x,y,z} - \mathbf{I}^{\text{SR}}_{x,y,z}\right)^2 \tag{3}$$

Equation 3 can be regarded as a weighted squared error, where α_1 and α_2 control the scale of the weight in each pixel. With the appearance of uSDF in the denominator, the closer the region is to the vehicle surface, the more important it will be. By contrast, the influence of the areas far from the vehicle surface is small. However, Chen *et al*. [8] showed that the too small weights will cause the model to lose the generation ability in the corresponding regions. Therefore, in this study, both α_1 and α_2 are set to 5.5 so that the weight of each pixel decreases from 1 to 0.5 when moving from the vehicle surface to the edge. Meanwhile, with the binary mask in the numerator, regions that do not allow air to pass through will be ignored. Using the binary mask, the weights of regions not allowing air to pass through are set to 0. Figure 3 shows an example that visualizes the weights on the central cross-section in y-axis direction.

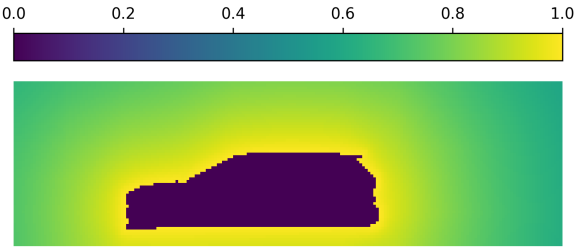

Figure 3. An example of distance-weighted loss.

3.3.2 Physical Loss

Since the flow of vehicles can be regarded as incompressible flow, the divergence of the velocity field should equal to 0, which is also called the continuity equation. In addition, by substituing the continuity equation for the momentum equation derived from Navier-Stokes equation, we obtain the pressure Poisson equation which can be used as the other physical constraint because the right side of the equation is 0. Therefore, similar to Subramaniam *et al*. [20], we use the residual of the continuity equation and the pressure Poisson equation as the physical constraints of the generator outputs. This physical loss is calculated as follows:

$$\begin{aligned} L_{physical} = {} & \lambda_C \underbrace{|\nabla \cdot \mathbf{u}|}_{\text{left side of the continuity equation}} \\ & + (1-\lambda_C) \underbrace{|\nabla^2 p + \rho \nabla \cdot (\mathbf{u} \cdot \nabla \mathbf{u})|}_{\text{left side of the pressure Poisson equation}} \end{aligned} \tag{4}$$

where $\mathbf{u}$ ($[u_x, u_y, u_z]^T$) and p denote the velocity vector and the kinematic pressure stored in each pixel of velocity field and pressure field respectively. ρ is the density of flow which is a constant value. In our case, ρ is set to $1.205\ \text{kg/m}^3$. λ_C is a hyperparameter in range $[0, 1]$ used to control the balance between the continuity equation loss and the pressure Poisson equation loss. Since both of them have the same importance, we set $\lambda_C = 0.5$.

3.4. Scheduled Learning

By combining the aforementioned three loss functions, the total loss for the generator should be described as follows:

$$L_G = w_c L_{content} + w_p L_{physical} + w_a L_{adversarial} \quad (5)$$

where w_c, w_p, w_a are hyperparameters representing the weight of each element. Generally, the training process will be stabilized by choosing suitable and fixed hyperparameters from the beginning. However, optimizing all the elements with fixed weights from the beginning may increase the training time. An illustrative example is to optimize $L_{content}$ and $L_{physical}$ simultaneously. Optimizing $L_{physical}$ from the beginning is unnecessary because optimizing $L_{content}$ independently helps the model generate data similar to high-resolution data; this indirectly helps the model automatically optimize a part of $L_{physical}$ without using $L_{physical}$ as feedback. Optimizing $L_{content}$ independently to a certain level and then slowly adding $L_{physical}$ to continue the training process will make the model train faster and more stable. The same applies to $L_{adversarial}$. In Subramaniam *et al*. [20], the generator was first trained without $L_{adversarial}$. The discriminator was then trained with hundreds of iterations independently. Finally, the generator and discriminator were trained together with the active adversarial loss until the losses were saturated. This learning strategy can improve the learning speed of the model. However, the learning curve of the $L_{physical}$ in this study shows that $L_{physical}$ suddenly spikes after adding $L_{adversarial}$; this happens because of a sudden change in the learning purpose. In some cases, the model may get stuck in the local minima or even does not converge because of this sudden change. In this study, the total loss for the generator is calculated as follows to avoid significant changes in the learning purpose and to stabilize learning as well.

$$\begin{aligned} L_G &= L_{content} + w_1 s_1 L_{physical} + w_2 s_2 L_{adversarial} \\ \text{where } \; w_i &= \frac{e^{\beta_i \times \text{epoch}} - 1}{e^{\beta_i \times \text{max\_epoch}}}, \;\; i \in \{1, 2\} \end{aligned} \quad (6)$$

where s_1, s_2 are hyperparameters in range $[0, 1]$ used to rescale $L_{physical}$ and $L_{adversarial}$, w_1, w_2 are used to control the weight of each loss element. In this work, s_1, s_2 were set to 10^{-2} and 1. β_1, β_2 were set to 10^{-3} and 5×10^{-3}.

Figure 4 shows an example in which $L_{content}$ is the primary loss function of the generator, and the weights of other loss functions slowly increase during training. In the early stages of training, the weight of $L_{physical}$ is minor. Therefore, the focus is on generating data close to high-resolution ground truth. As the training process progresses, the weight of $L_{physical}$ increases, and the focus on whether the generated results are correct according to the physical laws also increases. Finally, when the generator has learned to a certain level, and the importance of $L_{adversarial}$ is large enough, then $L_{adversarial}$ acts as feedback for the generator to focus on supplementing the missing parts, especially high-frequency components, and errors in the structure of the flow fields. This strategy increases training stability because the learning purpose is changed slowly. The computational cost is also reduced because optimizing $L_{content}$ indirectly optimizes the other two loss functions. $L_{physical}$ and $L_{adversarial}$ are gradually added after $L_{content}$ has been optimized to a certain level to help the generator learn the remaining tasks that cannot be completed with $L_{content}$ alone.

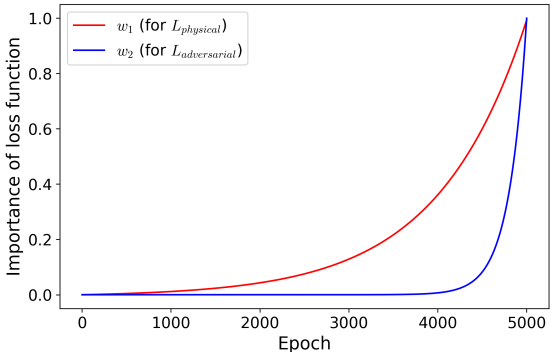

Figure 4. An example of scheduled learning

4. Experiments

We train all the networks on two NVIDIA A100 GPUs using the dataset described in Section 2. We split the dataset into training and test sets which include 1,011 and 110 cases respectively. As introduced in the previous section, rescaling parameters s_1 and s_2 are set to 10^{-2} and 1, scheduled learning parameters β_1 and β_2 are set to 10^{-3} and 5×10^{-3}. We use Adam optimizer (Kingma and Ba [30]) with a learning rate of 10^{-4} to train the model. For comparison purposes and confirming the effectiveness of the proposed elements, we reimplement the TEGAN that fits our flow field enrichment task and its variants by adding scheduled learning, binary mask, and distance-weighted loss sequentially. We train them under the same conditions.

4.1. Super-resolution Results and Comparisons

By training with the training set, it is possible to work on a $2\times$ upsampling flow field enrichment task, which generates the high-resolution ($200 \times 80 \times 60$) 3D flow fields from the low-resolution ones ($100 \times 40 \times 30$). In our case,

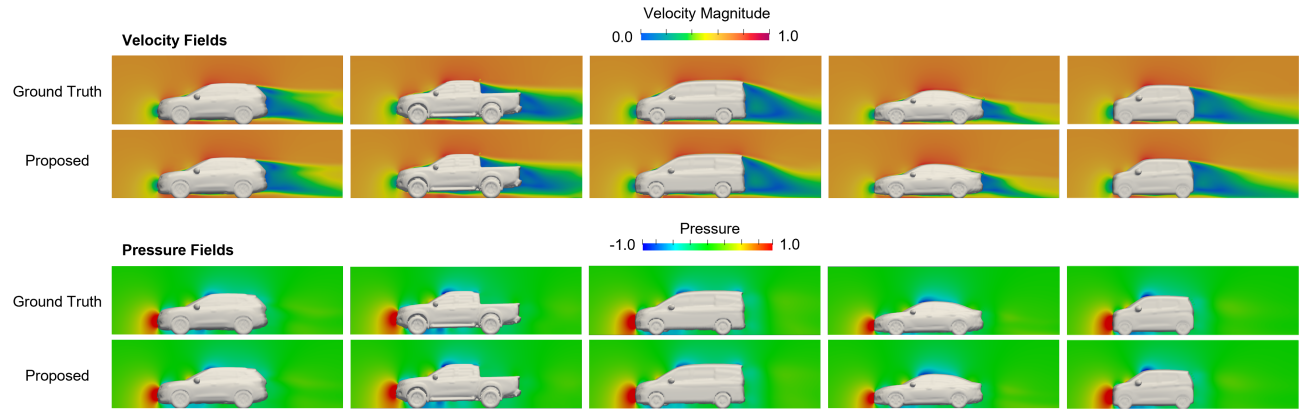

Figure 5. Comparisons of velocity magnitude and pressure between generated high-resolution flow fields and the ground truth in central cross-section. We show several cases of vehicles in the test set.

Masked MAE on the test set								
		2× upsampling					4× upsampling	
		Baseline (TEGAN)	+ SL	+ SL, Mask	+ SL, Mask, DwLoss	**Proposed**	Baseline (TEGAN)	**Proposed**
Velocity Fields (m/s)	u_x	0.501	0.485	0.454	0.399	**0.369**	0.656	**0.361**
	u_y	0.252	0.249	0.241	0.215	**0.203**	0.346	**0.221**
	u_z	0.247	0.246	0.228	0.191	**0.184**	0.319	**0.202**
Pressure Fields (Pa)	p	4.499	4.215	3.241	2.873	**2.671**	5.063	**2.742**

Table 1. Comparisons of Masked MAE on test set when training in different conditions. SL: Scheduled-learning, Mask: Binary mask on output, DwLoss: Distance-weighted loss.

training a super-resolution model takes about 10 days, while generating high-resolution flow fields from low-resolution only takes 0.2 seconds. By combining with the CNN-based flow field estimator like [8], it is possible to obtain higher-resolution flow fields from a vehicle shape in 15 seconds. Vehicle designers can discuss vairous ideas of shape deisgn with the more detailed aerodynamic performance in a short period, which is much faster than the current working process that spends at least one day for one shape revision.

Figure 5[1] shows the central cross-sections in y-axis direction of several vehicles evaluated on the test set, including both velocity magnitude and pressure which visualize the velocity field and pressure field respectively. We find the proposed method can work well on various categories of vehicles and the generated high-resolution flow fields are very close to the ground truth. It is difficult to find the differences by comparing with the ground truth in the figure.

[1]The velocity magnitude and pressure value can only be shown in the scaled unit-less ranges in terms of data provider's policy. However, the differences are allowed to show the actual ranges.

To verify the effectiveness of each element in the proposed method, we reimplement TEGAN as baseline and add those elements sequentially for comparison. All of the models are trained with the same training set and conditions for 2× upsampling flow field enrichment task. In this study, we use the masked mean absolute error (Masked MAE) in Equation (7) as the criterion for a quantitative comparison; this can be considered a variant of L1 loss when the errors have been averaged over the total number of masks, not over the total number of pixels. It eliminates areas with a mask value of 0, where no air flows, and only calculates the error in the airspace; this helped us obtain a more accurate view of the performance of each model.

$$\text{Masked MAE} = \frac{\sum_{x,y,z=0}^{\text{W,H,L}} |\mathbf{I}^{\text{HR}}_{x,y,z} - \mathbf{I}^{\text{SR}}_{x,y,z}|}{\sum_{x,y,z=0}^{\text{W,H,L}} \text{mask}_{x,y,z}} \tag{7}$$

Table 1 shows the results evaluated on the test set. We find that adding the proposed elements to the baseline surprisingly improves the performance. As expected, sched-

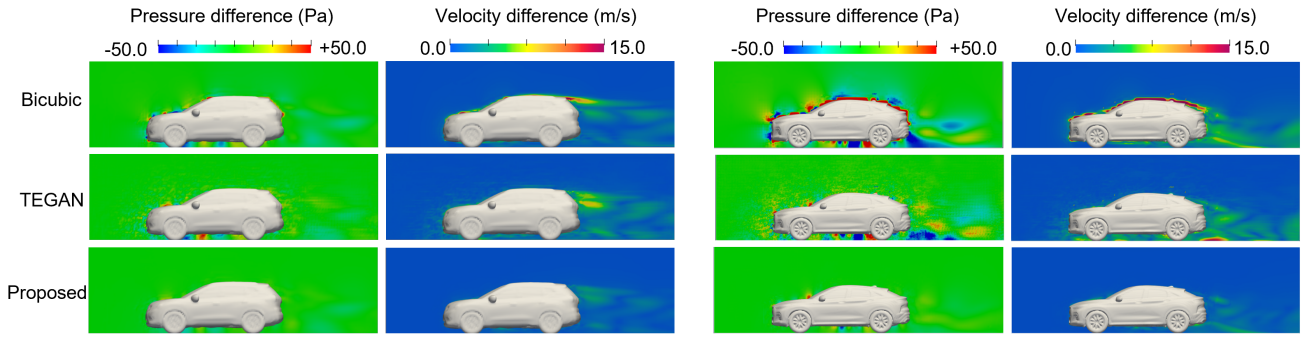

Figure 6. Differences between ground truth and generated high-resolution flow fields by different methods. We provide the comparisons on other vehicle cases in supplementary material Section A.

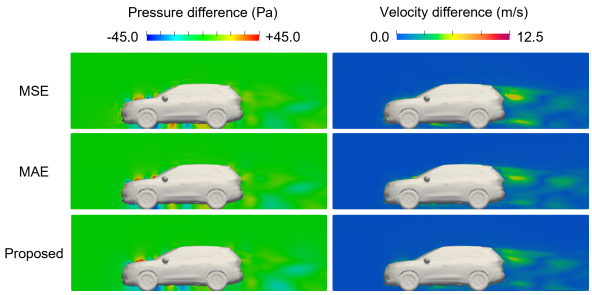

Figure 7. Differences between ground truth and generated results by the model trained with different content loss functions. See additional evaluations in supplementary material Section B.

Masked MAE on the test set				
		Loss functions		
		MSE	MAE	**Proposed**
Velocity Fields (m/s)	u_x	0.392	0.422	**0.369**
	u_y	0.222	0.219	**0.203**
	u_z	0.197	0.199	**0.184**
Pressure Fields (Pa)	p	2.962	2.834	**2.671**

Table 2. Comparisons of Masked MAE on test set when training with different loss functions

uled learning helps the model achieve better results by shifting the importance of $L_{physical}$ and $L_{adversarial}$ to the later part of the training process. Focusing on optimizing all loss functions from the beginning is not necessary because optimizing $L_{content}$ indirectly optimizes other loss functions. Moreover, optimizing all functions from the beginning also makes the model prone to falling into the local minima. Mask and distance-weighted loss also significantly improve accuracy because the mask eliminates air-flowless parts, whereas distance-weighted loss helps the model focus on parts that are difficult to reproduce into high-resolution data, especially in areas close to the surface and inside the vehicle. The proposed method, with a combination of the aforementioned elements and the proposed architecture for the generator and discriminator, outperforms the previous methods with a significant gap in accuracy.

Figure 6 shows the differences of the same central cross-section between ground truth and high-resolution generated flow fields by different methods on an SUV case in the test set. Bicubic interpolation and TEGAN cannot provide reliable results of flow fields close to the vehicle surface such as the pressure fields above the engine hood and the velocity field in the wake region around the rear of vehicles. In contrast, the proposed method is able to generate flow fields the more accurately in those regions which are crucial when assessing the vehicle aerodynamic performance.

To further evaluate the effectiveness of the proposed method, we prepared a small dataset for $4\times$ enrichment task which includes 56 cases of training data. We trained both the proposed method and TEGAN to $4\times$ upsample from the size of $100 \times 40 \times 30$ to $400 \times 160 \times 120$ and evaluated on a small test set which includes 12 cases. Both Figure 6 and Table 1 show similar results as $2\times$ enrichment task. In regards to the results of the proposed method, although there are some errors present in the front windshield and rear of the vehicle, the overall upsampled results using the proposed method are significantly superior compared to the previous methods.

4.2. Ablation Study of Loss Function

We consider that the proposed distance-weighted loss function plays an imperative role in training a more accurate

Masked MAE on the test set			
		Loss functions	
		w/o $L_{physical}$	w/ $L_{physical}$
Velocity Fields (m/s)	u_x	0.399	**0.369**
	u_y	0.217	**0.203**
	u_z	0.201	**0.184**
Pressure Fields (Pa)	p	2.711	**2.671**

Table 3. Comparisons of Masked MAE on test set when training with and without physical loss.

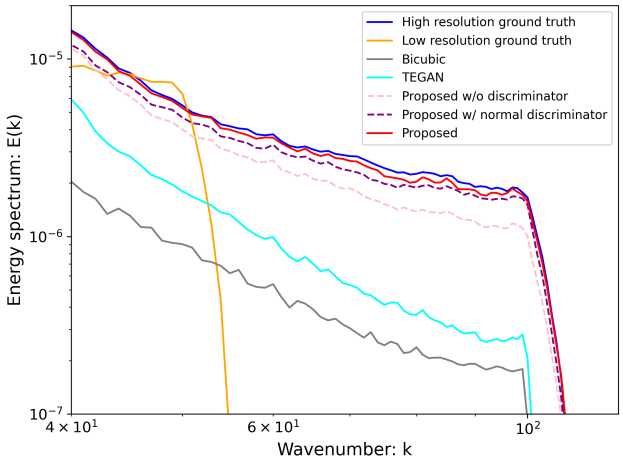

Figure 8. Comparison of the velocity energy spectra in high-frequency range for the different methods. (We provide results on other vehicle cases in supplementary material Section C.)

super-resolution model. To verify its effectiveness, we train the proposed method by replacing the proposed content loss function with MSE and MAE respectively, which are usually used in super-resolution tasks. Figure 7 shows the central cross-sections of the differences in the same SUV case. Compared to MAE and MSE, using distance-weighted loss helps the model generate the more accurate flow fields near the vehicle surface which are considered very important. Table 2 shows the results of the quantitative evaluation on the test set. The distance-weighted loss outperforms MAE and MSE in the results for all components of flow fields. In addition, the results in Table 3 show that the physical loss also contributes to improving performance. The physical loss not only helps produce physically correct results but also increases the accuracy of flow field enrichment on the test set.

4.3. High-frequency Components Analysis

We claimed that the proposed method helps the super-resolution model have a better ability to generate high-frequency components of flow fields. To verify our consideration, we calculate the energy spectra [31] of velocity fields. Figure 8 shows the results of the same SUV case used in Section 4.1 in the high-frequency range. The horizontal axis represents the wavenumber which is proportional to frequency. The vertical axis represents the energy of each wavenumber component. It is easy to find that the high-resolution flow field has the more high-frequency components than the low-resolution one. This partly shows the reason why vehicle designers need the high-resolution flow fields. These high-frequency components show the more details that help analyze the aerodynamic performance more accurately. The energy spectrum of the proposed method is very close to the high-resolution ground truth, which outperforms bicubic interpolation and TEGAN.

As shown in Figure 8, using bicubic interpolation causes large differences appeared in the high-frequency range. TEGAN, with the physical loss and GAN mechanism, improves generation ability of high-frequency components. However, the proposed method, with many improvements in the network architecture, loss functions, and training strategy, achieves a better match to real data than the previous methods in the high-frequency range. We can also confirm the effectiveness of adding GAN mechanism and selecting the relativistic averaged discriminator by comparing with the results of the super-resolution model trained with the normal discriminator [19] and without GAN mechanism separately. The results in Figure 8 clearly show that the GAN mechanism contributes to the generation ability of high-frequency components. Furthermore, the relativistic averaged discriminator performs better than the normal discriminator on high-frequency components generation.

5. Conclusion

In this study, we proposed a super-resolution model for 3D vehicle flow field enrichment task based on 3D CNN and GAN mechanism. We applied RRDB without batch normalization layers as the basic network-building unit of the generator and applied a binary mask multiplying the generated high-resolution flow fields for lower computational cost. The relativistic discriminator was used to provide better feedback regarding the lack of high-frequency components. To obtain better results in wake regions and regions near vehicle surfaces, a distance-weighted loss function was proposed. To imporve the effectiveness of training the model, we proposed a scheduled learning for our task. Experimental results showed that the proposed method outperformed the previous works in the vehicle flow field enrichment task. The proposed method also captured the details in those regions we concern better and had a better generation ability of high-frequency components.

References

[1] Yue-Hong Qian, Dominique d'Humières, and Pierre Lallemand. Lattice bgk models for navier-stokes equation. *EPL (Europhysics Letters)*, 17(6):479, 1992. 1

[2] Guy R McNamara and Gianluigi Zanetti. Use of the boltzmann equation to simulate lattice-gas automata. *Physical review letters*, 61(20):2332, 1988. 1

[3] Xiaoxiao Guo, Wei Li, and Francesco Iorio. Convolutional neural networks for steady flow approximation. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 481–490, 2016. 1

[4] Saakaar Bhatnagar, Yaser Afshar, Shaowu Pan, Karthik Duraisamy, and Shailendra Kaushik. Prediction of aerodynamic flow fields using convolutional neural networks. *Computational Mechanics*, 64(2):525–545, 2019. 1

[5] Nils Thuerey, Konstantin Weißenow, Lukas Prantl, and Xiangyu Hu. Deep learning methods for reynolds-averaged navier–stokes simulations of airfoil flows. *AIAA Journal*, 58 (1):25–36, 2020. 1

[6] Donglin Chen, Xiang Gao, Chuanfu Xu, Shizhao Chen, Jianbin Fang, Zhenghua Wang, and Zheng Wang. Flowgan: A conditional generative adversarial network for flow prediction in various conditions. In *Proceedings of the IEEE 32nd International Conference on Tools with Artificial Intelligence (ICTAI)*, pages 315–322, 2020. 1

[7] Asad Abbas, Ashkan Rafiee, Max Haase, and Andrew Malcolm. Geometric convolutional neural networks–a journey to surrogate modelling of maritime cfd. In *Proceedings of the 9th Conference on Computational Methods in Marine Engineering*, 2021. 1

[8] Fangge Chen and Kei Akasaka. 3d flow field estimation around a vehicle using convolutional neural networks. In *Proceedings of the 32nd British Machine Vision Conference*, 2021. 1, 2, 4, 6

[9] Hao Ma, Yuxuan Zhang, Nils Thuerey, Xiangyu Hu, and Oskar J Haidn. Physics-driven learning of the steady navier-stokes equations using deep convolutional neural networks. *Communications in Computational Physics*, 32(3):715–736, 2022. 1

[10] Sam Jacob Jacob, Markus Mrosek, Carsten Othmer, and Harald Köstler. Deep learning for real-time aerodynamic evaluations of arbitrary vehicle shapes. *SAE International Journal of Passenger Vehicle Systems*, 15(2):77–90, 2022. 1

[11] Yuuichi Asahi, Sora Hatayama, Takashi Shimokawabe, Naoyuki Onodera, Yuta Hasegawa, and Yasuhiro Idomura. Amr-net: Convolutional neural networks for multi-resolution steady flow prediction. In *Proceedings of the IEEE International Conference on Cluster Computing (CLUSTER)*, pages 686–691, 2021. 1

[12] Chao Dong, Chen Change Loy, Kaiming He, and Xiaoou Tang. Image super-resolution using deep convolutional networks. *IEEE transactions on pattern analysis and machine intelligence*, 38(2):295–307, 2015. 1

[13] Chao Dong, Chen Change Loy, and Xiaoou Tang. Accelerating the super-resolution convolutional neural network. In *Proceedings of the 3rd International Conference on Learning Representations (ICLR)*, 2016. 1

[14] Ying Tai, Jian Yang, and Xiaoming Liu. Image super-resolution via deep recursive residual network. *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3147–3155, 2017. 1

[15] Wei-Sheng Lai, Jia-Bin Huang, Narendra Ahuja, and Ming-Hsuan Yang. Deep laplacian pyramid networks for fast and accurate super-resolution. *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 136–144, 2017. 1

[16] Bee Lim, Sanghyun Son, Heewon Kim, Seungjun Nah, and Kyoung Mu Lee. Enhanced deep residual networks for single image super-resolution. *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 624–632, 2017. 1, 2

[17] Christian Ledig, Lucas Theis, Ferenc Huszár, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, et al. Photo-realistic single image super-resolution using a generative adversarial network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4681–4690, 2017. 1, 3

[18] Xintao Wang, Ke Yu, Shixiang Wu, Jinjin Gu, Yihao Liu, Chao Dong, Yu Qiao, and Chen Change Loy. Esrgan: Enhanced super-resolution generative adversarial networks. In *Proceedings of the European Conference on Computer Vision Workshops (ECCVW)*, 2018. 1, 2, 3

[19] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In *Advances in Neural Information Processing Systems*, 2014. 1, 8

[20] Akshay Subramaniam, Man Long Wong, Raunak D. Borker, Sravya Nimmagadda, and Sanjiva K. Lele. Turbulence enrichment with physics-informed generative adversarial network. In *Proceedings of the Third Workshop on Machine Learning and the Physical Sciences (NeurIPS)*, 2020. 1, 2, 4, 5

[21] Mathis Bode, Michael Gauding, Zeyu Lian, Dominik Denker, Marco Davidovic, Konstantin Kleinheinz, Jenia Jitsev, and Heinz Pitsch. Using physics-informed enhanced super-resolution generative adversarial networks for subfilter modeling in turbulent reactive flows. *Proceedings of the Combustion Institute*, 38(2):2617–2625, 2021. 1

[22] Ludovico Nista, Christoph Karl David Schumann, Gandolfo Scialabba, Temistocle Grenga, Antonio Attili, and Heinz Pitsch. The influence of adversarial training on turbulence closure modeling. In *AIAA SCITECH 2022 Forum*, 2022. 1

[23] Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv:1502.03167*, 2015. 2

[24] Alexia Jolicoeur-Martineau. The relativistic discriminator: a key element missing from standard gan. In *Proceedings of the International Conference on Learning Representations*, 2019. 2, 3

[25] Donya Khaledyan, Abdolah Amirany, Kian Jafari, Mohammad Hossein Moaiyeri, Abolfazl Zargari Khuzani, and Najmeh Mashhadi. Low-cost implementation of bilinear and bicubic image interpolation for real-time image super-resolution. In *Proceedings of IEEE Global Humanitarian Technology Conference (GHTC)*, pages 1–5, 2020. 2

[26] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016. 3

[27] Gao Huang, Zhuang Liu, Laurens Van Der Maaten, and Kilian Q Weinberger. Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4700–4708, 2017. 3

[28] Wenzhe Shi, Jose Caballero, Ferenc Huszár, Johannes Totz, Andrew P Aitken, Rob Bishop, Daniel Rueckert, and Zehan Wang. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1874–1883, 2016. 3

[29] Katja Schwarz, Yiyi Liao, and Andreas Geiger. On the frequency bias of generative models. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2021. 3

[30] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In *Proceedings of the 3rd International Conference on Learning Representations (ICLR)*, 2015. 5

[31] Farshad Navah, Marta de la Llave Plata, and Vincent Couaillier. A high-order multiscale approach to turbulence for compact nodal schemes. *Computer Methods in Applied Mechanics and Engineering*, 363:112885, 2020. 8