Article

A Combined Model of Diffusion Model and Enhanced Residual Network for Super-Resolution Reconstruction of Turbulent Flows

Jiaheng Qi [1] and Hongbing Ma [1,2,*]

[1] School of Intelligence Science and Technology, Xinjiang University, Urumqi 830046, China; 107552101682@stu.xju.edu.cn
[2] Department of Electronic Engineering, Tsinghua University, Beijing 100084, China
* Correspondence: hbma@tsinghua.edu.cn

Abstract: In this study, we introduce a novel model, the Combined Model, composed of a conditional denoising diffusion model (SR3) and an enhanced residual network (EResNet), for reconstructing high-resolution turbulent flow fields from low-resolution flow data. The SR3 model is adept at learning the distribution of flow fields. The EResNet architecture incorporates a long skip connection extending from the input directly to the output. This modification ensures the preservation of essential features learned by the SR3, while simultaneously enhancing the accuracy of the flow field. Additionally, we incorporated physical gradient constraints into the loss function of EResNet to ensure that the flow fields reconstructed by the Combined Model are consistent with the direct numerical simulation (DNS) data. Consequently, the high-resolution flow fields reconstructed by the Combined Model exhibit high conformity with the DNS results in terms of flow distribution, details, and accuracy. To validate the effectiveness of the model, experiments were conducted on two-dimensional flow around a square cylinder at a Reynolds number (Re) of 100 and turbulent channel flow at Re = 4000. The results demonstrate that the Combined Model can reconstruct both high-resolution laminar and turbulent flow fields from low-resolution data. Comparisons with a super-resolution convolutional neural network (SRCNN) and an enhanced super-resolution generative adversarial network (ESRGAN) demonstrate that while all three models perform admirably in reconstructing laminar flows, the Combined Model excels in capturing more details in turbulent flows, aligning the statistical outcomes more closely with the DNS results. Furthermore, in terms of L2 norm error, the Combined Model achieves an order of magnitude lower error compared to SRCNN and ESRGAN. Experimentation also revealed that SR3 possesses the capability to learn the distribution of flow fields. This work opens new avenues for high-fidelity flow field reconstruction using deep learning methods.

Keywords: deep learning; super-resolution reconstruction of turbulent flows; diffusion model; residual network

MSC: 68T07

1. Introduction

High-precision flow field data have significant impacts on many areas, such as fluid analysis, weather forecasting, and the optimization design of a flying wing [1]. Reconstructing high-fidelity flow fields represents a significant challenge within the realms of fluid dynamics and computational fluid dynamics (CFDs). It is expensive to obtain high-precision turbulent flow fields using DNS [2] or experimental methods, due to its chaotic behavior with multiple spatiotemporal scales. To rapidly predict high-fidelity flow field data, Shi et al. [3] developed an adjoint method. Compared with the direct iterative reverse automatic differentiation (RAD) method, the adjoint method gives an 84.6% saving in time. Deep learning has a strong nonlinear fitting ability [4] and can mine useful information

from massive, existing flow field data [5]. In this work, we aim to develop a deep learning model capable of reconstructing high-fidelity fluid data from low-fidelity data. Given that low-fidelity data can be generated with fewer computational resources or more readily acquired via experimental methods, employing a deep learning model to reconstruct high-fidelity data from such low-fidelity sources can markedly diminish the computational expenses associated with obtaining high-quality flow fields.

Inspired by various research advances in the super-resolution deep learning of images, such as the convolutional neural network-based (CNN-based) methods [6–10], the generative adversarial network-based (GAN-based) methods [11–13], the transformer-based methods [14,15], and the diffusion-based methods [16–18], several neural network models have been proposed to reconstruct a high-resolution (HR) flow field. We categorize the recent super-resolution (SR) methods for flow fields into three categories:

1. Direct-mapping models: The direct-mapping models are always CNN-based models which are trained to directly minimize the reconstruction loss between the ground truth data and reconstructed flow in the sense of an L_p norm. The purpose of the direct-mapping models is to reconstruct a flow field that is numerically close to the real flow field. Many researchers have reconstructed HR flow by direct-mapping methods. Fukami et al. [19,20] proposed a hybrid downsampled skip-connection multi-scale (DSC/MS) model, which reconstructed HR flow data from grossly under-resolved input data both in space and time. Onishi et al. [21] proposed a CNN-based super-resolution (SR) method to reconstruct HR data from low-resolution (LR) urban meteorological simulation data, and the reconstruction effect is much better than that of traditional interpolation methods. A multiple path super-resolution convolutional neural network (MPSRC) was proposed by Kong et al. [22,23] to fully capture the spatial distribution features of temperature and supersonic flow field. The results demonstrated that the MPSRC can provide a better reconstruction result with a lower mean square error and a higher peak signal-to-noise ratio than CNN. Liu et al. [24] proposed a multiple temporal path CNN to fully capture the temporal information from consecutive fluid fields, which can reconstruct more details and improve the spatial resolution compared to static CNN.

2. Direct-mapping models with judgment: The direct-mapping models with judgment are always GAN-based models which consist of a generator and a discriminator. The generator is trained not only to minimize the reconstruction loss, but also to maximize the loss of discriminator. The purpose of the direct-mapping models with judgment is to generate a flow field that is numerically close to the real flow field and to make the discriminator mistakenly believe that it is the real flow field. Yousif et al. [25,26] proposed a multiscale ESRGAN model which can reconstruct high-fidelity turbulent flow with extremely coarse data and predict the HR turbulent velocity fields of the turbulent channel flow with a different Reynolds number without retraining the network parameters. Xu et al. [27] presented a 3D-SRGAN to learn the topographic information and infer a high-resolution 3D turbulent flame structure with a given LR counterpart. Yu et al. [28] proposed the 3D-ESRGAN with tricubic interpolation-based transfer learning to reconstruct 3D HR turbulent flows with limited training data. Xie et al. [29] presented the first approach to synthesize four-dimensional physics fields. They generated consistent and detailed results by using a novel temporal discriminator based on a conditional GAN that is designed for the inference of three-dimensional volumetric data.

3. Pure diffusion-based models: The pure diffusion-based models are denoising diffusion probabilistic models which include two processes: diffusion process (adding noise into the HR data) and inverse process (denoising from noisy data). Different from the previous two kinds of models, DDPM cannot reconstruct the HR flow field according to the LR flow field at one time but needs to denoise the concatenation of the LR flow field and the noisy flow field T (T is a hyperparameter) times to reconstruct the HR flow field. The goal of this model is to reconstruct an HR flow field that conforms to

the probability distribution of the ground truth data. Shu et al. [30] presented the first diffusion-based model, a physics-informed diffusion model, which could produce accurate reconstruction results for two-dimensional (2D) Kolmogorov flow with a Reynolds number of 1000.

The above deep learning models have achieved encouraging results in flow field SR, but they still have some problems. Direct-mapping models and pure diffusion-based models represent two extremes. The former are merely numerical approximations without capturing the distribution of flow fields. Conversely, the latter focus exclusively on learning the distribution of flow fields, yet their accuracy in reconstruction requires enhancement. Furthermore, adding physical constraints to the pure diffusion-based models is challenging because the diffusion model predicts the added noise rather than directly predicting the flow field. For direct-mapping models with judgment, training GANs is notably difficult, with the potential of converging to local minima. At these local minima, the HR flow fields generated by GANs fail to accurately capture the true statistics of the training data. To solve the aforementioned problems, we introduce an easily trainable Combined Model that integrates SR3 and EResNet for the purpose of the super-resolution reconstruction of flow fields. The SR3 model is utilized to comprehend the distribution of the flow field. The EResNet further enhances precision based on the outcomes of SR3. Consequently, the HR flow fields constructed by the Combined Model not only achieve high precision but also agree the probability distribution of the DNS flow field. Experimental results demonstrate that the Combined Model can successfully reconstruct HR turbulent flow fields from LR flow data, achieving a high level of concordance with the DNS results.

2. Methods

We began with a dataset of input–output flow field pairs, denoted as $\mathcal{D} = \left\{ x^i, y^i \right\}_{i=1}^{N}$, where y^i represents the HR flow field and x^i represents the corresponding LR flow field. The purpose of this work was to reconstruct HR flow y from the corresponding LR data x. These LR data were obtained by downsampling high-resolution data. To achieve this objective, we proposed a method that proceeds in two steps: first, learning the distribution of the flow field; second, enhancing the accuracy of the flow field based on the results in the first step,

$$\breve{y} = F_2(F_1(x, \theta_1), \theta_2) \tag{1}$$

where $\breve{y}$ is the SR reconstructed result. F_1 denotes the model used in the first step to learn the distribution of the flow field. F_2 denotes the model used in the second step to improve the accuracy of the flow field. θ_1 and θ_2 denote the learnable parameters of the F_1 and F_2, respectively.

Suppose a training dataset contains some flow field pairs, e.g., $D_{train} = \left\{ x^i, y^i \right\}_{i=1}^{M}$ ($M < N$), the training process of the model is equivalent to solving an optimization problem,

$$\theta = argmin_{\theta_1, \theta_2} \frac{1}{M} \sum_{i=1}^{M} \mathcal{L}_{\theta_2}(\overline{y} \sim P_{\theta_1}(y^i \mid x^i), y^i) \tag{2}$$

where $\mathcal{L}$, θ_2 denote the loss function and learnable parameters of the F_2; $P_{\theta_1}(y \mid x)$ denotes the probability distributions of y given x learned by F_1. θ_1 and $\overline{y}$ denote the learnable parameters and the reconstructed results of the F_1, respectively. θ denotes all the learnable parameters of the model, where $\theta = \theta_1 + \theta_2$. Once the training process is complete, θ remains fixed during testing.

Therefore, we proposed the Combined Model to achieve the above objectives. The Combined Model consisted of two parts: SR3 and EResNet. The overall framework of the Combined Model is shown in Figure 1. First, the input low-resolution (LR) flow field is processed by SR3 to generate a preliminary flow field consistent with the true flow field distribution, and then the residual network further improves the accuracy of the flow

field to obtain the final flow field result. Next, a detailed description of SR3 and EResNet is provided.
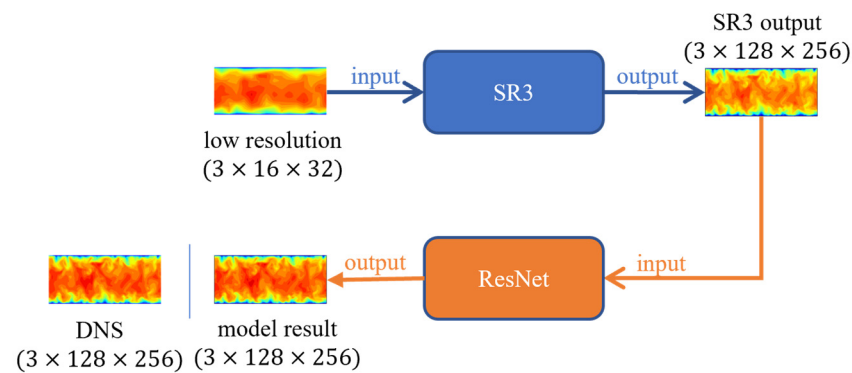

Figure 1. The framework of the Combined Model. For the sake of convenience, we showed only one of the three velocity fields.

2.1. Conditional Denoising Diffusion Model

The SR3 model is a generative model and has some improvements based on the denoising diffusion probabilistic model (DDPM). SR3 exhibits outstanding performance in the super-resolution reconstruction of faces and natural images. This example demonstrates the potential of a DDPM-based model in reconstructing HR flow data from LR flow data. An overview of the primary HR flow data reconstruction framework with SR3 is shown in Figure 2. More details of the SR3 model and how it is used for flow data reconstruction are provided in the following subsections.
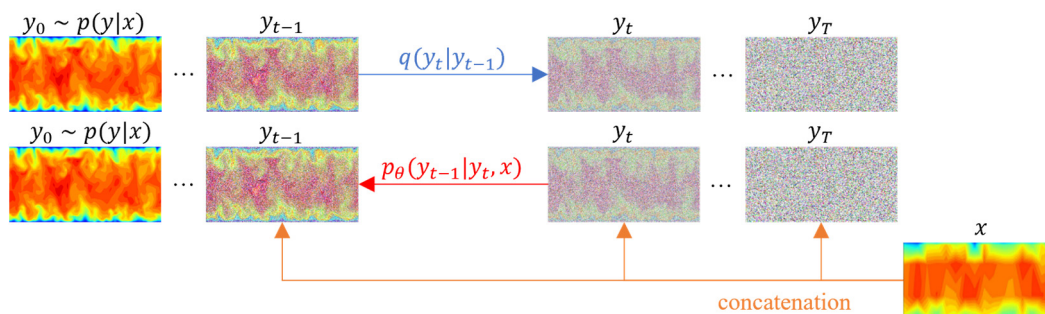

Figure 2. There are two processes involved in the SR3 model: 1. the forward diffusion process q (left to right) gradually adds Gaussian noise to the HR flow; 2. the reverse inference process p (right to left) iteratively denoises the noisy flow, conditioned on an LR flow. For the sake of convenience, we showed only one of the three velocity fields.

2.1.1. Forward Diffusion Process

Following [16,31,32], the definition of the forward Markovian diffusion process q:

$$q(y_{1:T} \mid y_0) := \prod_{t=1}^{T} q(y_t \mid y_{t-1}) \tag{3}$$

$$q(x_t \mid x_{t-1}) = N(x_t, \sqrt{1 - \beta_t} x_{t-1}, \beta_t I) \tag{4}$$

where the scalar parameters $\beta_{1:T}$ are hyperparameters, subject to $0 < \beta_t < 1$, which determine the variance in the noise added at each iteration. The distribution of y_t given y_0 can be characterized by marginalizing out the intermediate steps as:

$$q(y_t \mid y_0) = \mathcal{N}(y_t \mid \sqrt{\gamma_t} y_0, (1 - \gamma_t)\mathbf{I}) \tag{5}$$

where $\gamma_t = \prod\limits_{i=1}^{t} \alpha_i, \alpha_t = 1 - \beta_t$. Furthermore, with some algebraic manipulation and completing the square, one can derive the posterior distribution of y_{t-1} given (y_t, y_0) as:

$$
\begin{aligned}
q(y_{t-1} \mid y_0, y_t) &= \mathcal{N}(y_{t-1} \mid \mu, \sigma^2 \mathbf{I}) \\
\mu &= \frac{\sqrt{\gamma_{t-1}}(1-\alpha_t)}{1-\gamma_t} y_0 + \frac{\sqrt{\alpha_t}(1-\gamma_{t-1})}{1-\gamma_t} y_t \\
\sigma^2 &= \frac{(1-\gamma_{t-1})(1-\alpha_t)}{1-\gamma_t}.
\end{aligned}
\tag{6}
$$

2.1.2. Optimizing the Denoising Model

As suggested by Meng et al. [33], the Markovian property of the backward inference process implies that the process to generate y_0 does not have to start from y_T but can start from y_t given by adding noise into y_0 t times, $t \in \{1, 2, \ldots, T\}$. We optimized a neural denoising model f_{θ_1} that takes as input LR flow x and a noisy flow $\widetilde{y}$,

$$
\widetilde{y} = \sqrt{\gamma} y_0 + \sqrt{1-\gamma} \epsilon, \quad \epsilon \sim \mathcal{N}(\mathbf{0}, \mathbf{I})
\tag{7}
$$

and aims to recover the HR flow y_0. In addition to an LR flow x and a noisy flow $\widetilde{y}$, the denoising model f_{θ_1} takes as input the sufficient statistics for the variance of the noise γ and is trained to predict the noise ϵ added to y_0. The SR3 architecture, which is a U-Net [34] augmented with self-attention, is depicted in Figure 3. The proposed objective function (refer to Appendix A) for training f_{θ_1} is:

$$
\mathbb{E}_{(\mathbf{x},\mathbf{y})} \mathbb{E}_{\epsilon,\gamma} \left\| f_{\theta_1}(x, \sqrt{\gamma} y_0 + \sqrt{1-\gamma} \epsilon, \gamma) - \epsilon \right\|_2^2
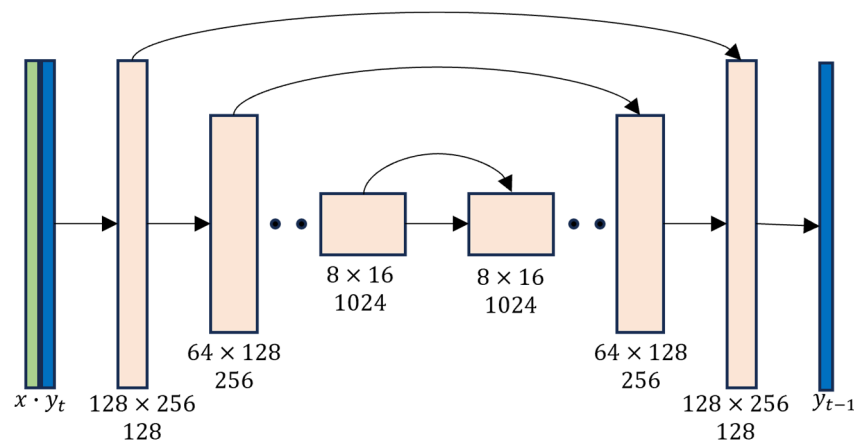\tag{8}
$$

Figure 3. Depiction of U-Net architecture of SR3. The LR input flow x up-samples to the target resolution using bicubic interpolation and concatenates with the noisy high resolution output flow y_t.

During training, to condition the model on the input x, we up-sampled the LR flow to the target resolution using bicubic interpolation and concatenated with $\widetilde{y}$ along the channel dimension. Generally, larger values of T lead to better models, although they also result in a longer backward inference process time. We set T to 200 and adopted a linear noise schedule.

2.1.3. Reverse Inference Process

Inference under our model is defined as a reverse Markovian process, which goes in the reverse direction of the forward diffusion process, starting from the concatenation of LR flow and noisy flow:

$$p_{\theta_1}(y_{0:T} \mid x) = p(y_T)\prod_{t=1}^{T} p_{\theta_1}(y_{t-1} \mid y_t, x) \tag{9}$$

$$p(y_T) = \mathcal{N}(y_T \mid \mathbf{0}, \mathbf{I}) \tag{10}$$

If the noise variances in the forward process steps are set as small as possible, i.e., $\alpha_{1:T} \approx 1$, the optimal reverse process $p_{\theta_1}(y_{t-1} \mid y_t, x)$ will approximate a Gaussian distribution [31].

$$p_{\theta_1}(y_{t-1} \mid y_t, x) = \mathcal{N}(y_{t-1} \mid \mu_{\theta_1}(x, y_t, \gamma_t), \sigma_t^2 \mathbf{I}) \tag{11}$$

Recall that the denoising model f_{θ_1} is trained to estimate ϵ, given any noisy flow $\widetilde{y}$ including y_t. Thus, given y_t, we approximated y_0 by rearranging the terms in (4) as:

$$\hat{y}_0 = \frac{1}{\sqrt{\gamma_t}}\left(y_t - \sqrt{1 - \gamma_t}f_{\theta_1}(x, y_t, \gamma_t)\right) \tag{12}$$

Following [32], we substituted our estimate $\hat{y}_0$ into the posterior distribution of $q(y_{t-1} \mid y_0, y_t)$ in (4) to parameterize the mean of $p_{\theta_1}(y_{t-1} \mid y_t, x)$ as:

$$\mu_{\theta_1}(x, y_t, \gamma_t) = \frac{1}{\sqrt{\alpha_t}}\left(y_t - \frac{1 - \alpha_t}{\sqrt{1 - \gamma_t}}f_{\theta_1}(x, y_t, \gamma_t)\right) \tag{13}$$

Ho et al. [32] achieved the best results by fixing the variance $\sigma_t^2 I$ rather than learning it. We set the variance of the $p_{\theta_1}(y_{t-1} \mid y_t, x)$ to $1 - \alpha_t$. Following this parameterization, each iteration of iterative refinement under our model takes the form:

$$y_{t-1} \leftarrow \frac{1}{\sqrt{\alpha_t}}\left(y_t - \frac{1 - \alpha_t}{\sqrt{1 - \gamma_t}}f_{\theta_1}(x, y_t, \gamma_t)\right) + \sqrt{1 - \alpha_t}\epsilon_t \tag{14}$$

We can estimate the HR flow y_0 by iterating.

2.1.4. SR3 Learns the Distribution of Flow Field

Currently, most reconstruction models are direct-mapping models that learn the mapping between LR flow and HR flow by minimizing the reconstruction loss. Unlike the direct-mapping models, the SR3 model could learn the overall distribution of the flow field rather than a simple numerical approximation. Figure 4 demonstrates that during the early stages of training, the SR3 model can learn the general outline and details of the flow field. As the number of training epochs increases, the flow field generated by the SR3 model becomes more similar to the true flow field. In contrast to other models, a diffusion-based model is not trained to minimize the reconstruction loss directly using an L_p norm. The model is trained to minimize the KL-divergence between the forward and backward diffusion processes. This may be one reason why the SR3 model can learn the flow field distribution so effectively.
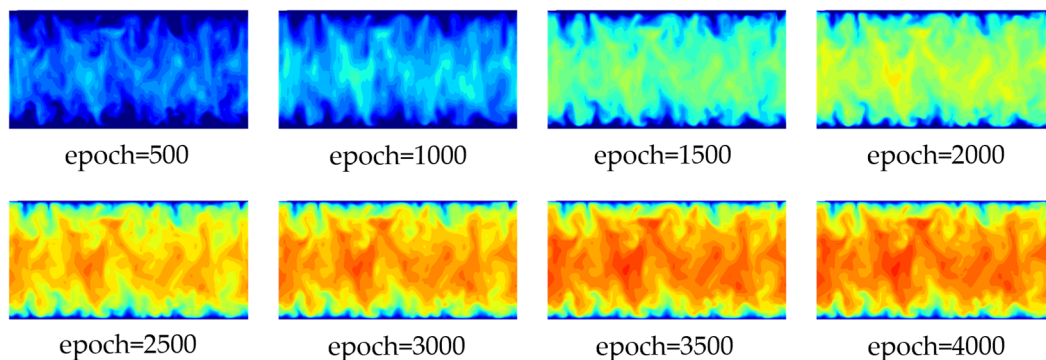

epoch=500   epoch=1000   epoch=1500   epoch=2000

epoch=2500   epoch=3000   epoch=3500   epoch=4000

Figure 4. Reconstructed flow field of the model after different epochs of training.

2.1.5. Reasons to Improve SR3 Flow Field Precision

Figure 4 demonstrates no significant improvement between the 3500th and 4000th training epoch, indicating that the model has converged. However, there is potential for additional refinement in the flow field accuracy. The reasons for this problem are as follows:

1. The input noisy flow y_t includes the information of y_0 during the training of the SR3 model. When using the pretrained SR3 model to generate the HR flow, the input data is pure Gaussian noise and does not contain any information of y_0.
2. In the reverse inference process, we used the mean of $q(y_{t-1} \mid y_0, y_t)$ to replace the mean of the $p_{\theta_1}(y_{t-1} \mid y_t, x)$.

Therefore, there is a deviation between the HR flow field $\bar{y}_0$ generated by the SR3 model and the real HR flow field y_0.

2.2. EResNet

2.2.1. EResNet Architecture

He et al. [35] proposed a residual network to address the degradation of the deep network. The network is composed of residual blocks, as illustrated in Figure 5. A residual block can be represented as:

$$x_{l+1} = x_l + \mathcal{F}(x_l, W_l) \tag{15}$$

where x_l, x_{l+1}, W_l denote the input, output, and weights of the residual block, respectively. $\mathcal{F}(x_l, W_l)$ denotes the residual between the input and the label.
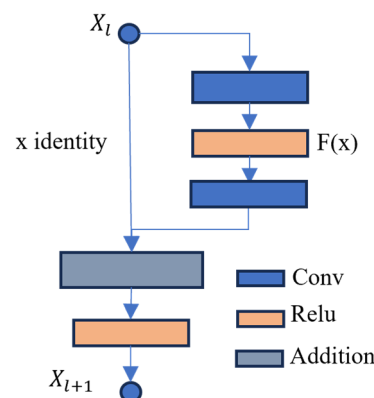

Figure 5. Residual learning: a building block.

Figure 6 illustrates how EResNet maps information directly and constantly to the next layer of the model through shortcut connections. The shortcut connections in this network allow information to bypass the nonlinear layers, helping to solve the problem of vanishing gradients. The loss function of EResNet is denoted as:

$$\mathcal{L} = MSE(y - y^*) + \lambda L_{gradient} \tag{16}$$

where $L_{gradient}$ is a physical constraint to ensure that the gradient of the flow field generated by the Combined Model is consistent with the real flow field, $\lambda = 0.001$. y and y^* denote the ground truth data and the reconstructed flow by EResNet, respectively.
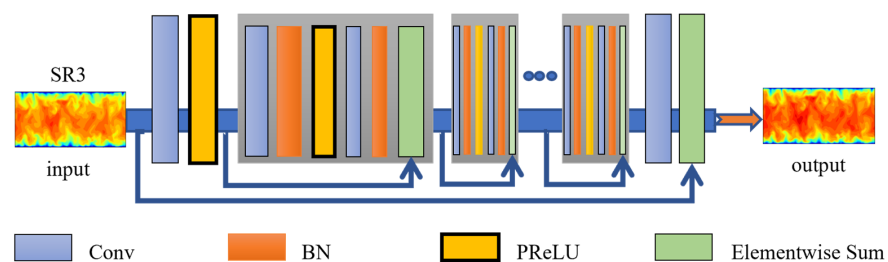

Figure 6. The architecture of EResNet.

2.2.2. EResNet Closes the Deviation between the SR3 Result and Ground Truth Data

In a general EResNet, there are only skip connections between residual blocks to prevent gradient vanishing. In EResNet, we added a long skip connection from the input to the output. As shown in Figure 5, to ensure that Combined Model maintains the flow field distribution obtained by SR3, the flow field reconstructed by SR3 can be directly integrated into the output of EResNet via skip connections.

$$y^* = \overline{y}_0 + \mathcal{F}(\overline{y}_0, \theta_2) \tag{17}$$

where $\mathcal{F}(\overline{y}_0, \theta_2)$ denotes the residual between the $\overline{y}_0$ and the y_0. Therefore, EResNet can learn the deviation δ between $\overline{y}_0$ and y_0. As demonstrated in Figure 7a, the flow field reconstructed by the Combined Model is more accurate than the flow field reconstructed by SR3, indicating an improvement in accuracy with the use of the Combined Model. Figure 7b illustrates a discrepancy between the probability density function (PDF) of the flow field reconstructed by SR3 and the DNS results. In contrast, the PDF of the flow field reconstructed by the Combined Model aligns with the DNS outcomes, as depicted in Figure 7c.
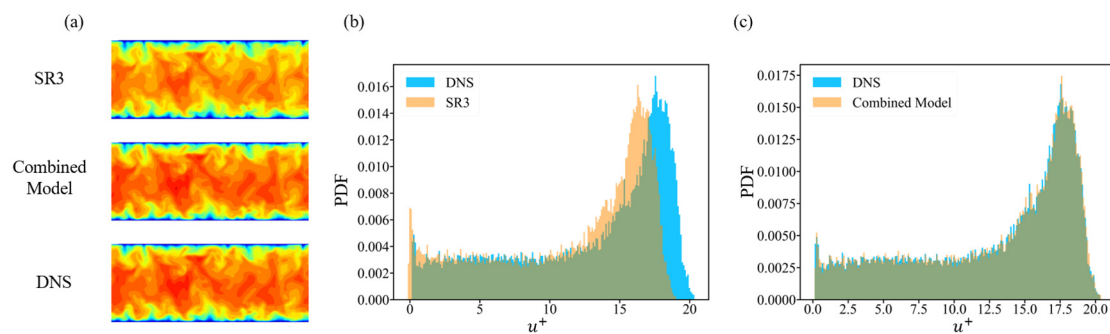

Figure 7. Comparison between SR3, Combined Model, and DNS. (**a**) Instantaneous velocity flow fields generated by SR3, Combined Model, and DNS results. (**b**) The corresponding PDF of SR3 and DNS results. (**c**) The corresponding PDF of Combined Model and DNS results. The green parts represent the overlap between the PDF of the reconstructed flow field and the DNS results.

3. Results

In this study, the effectiveness of the Combined Model for the super-resolution reconstruction of the turbulent flow field was evaluated by super-resolution reconstruction of the 2D laminar flow around a square cylinder and the turbulent channel flow in visualization and statistics.

3.1. Dataset

The 2D laminar flow around a square cylinder with Re = 100 and the turbulent channel flow with Re = 4000 from Yousif and Yu [25] are considered in this study for the training and testing of our model. We obtained 1000 pairs of low- and high-resolution flow field data each for the 2D laminar flow around a square cylinder and the turbulent channel flow (2D slices), respectively. The HR flow field size are both (128 × 256). Following [25], we select points at coarseness level (16 × 32) to obtain the LR flow field. The data are divided into training, validation, and testing datasets in the ratio of 8:1:1, respectively.

3.2. The 2D Laminar Flow around a Square Cylinder

In this section, we explore the capability of the Combined Model in reconstructing HR flow fields of 2D laminar flow around a square cylinder from coarse data. It is important to note that all the results are derived from test data not included in the training set. The reconstructed instantaneous velocity fields (U and V) and pressure fields generated by various models, alongside the DNS results, are presented in Figure 8. Here, the veloc-

ity components are normalized by U_∞, and the dimensionless pressure is expressed as $C_P = (P - P_\infty)/0.5\rho U_\infty$, where P_∞ represents the freestream pressure and ρ denotes the density. It is observed that the instantaneous flow fields of 2D laminar flow around a square cylinder, as reconstructed by SRCNN, ESRGAN, and the Combined Model, correspond closely with the DNS results.
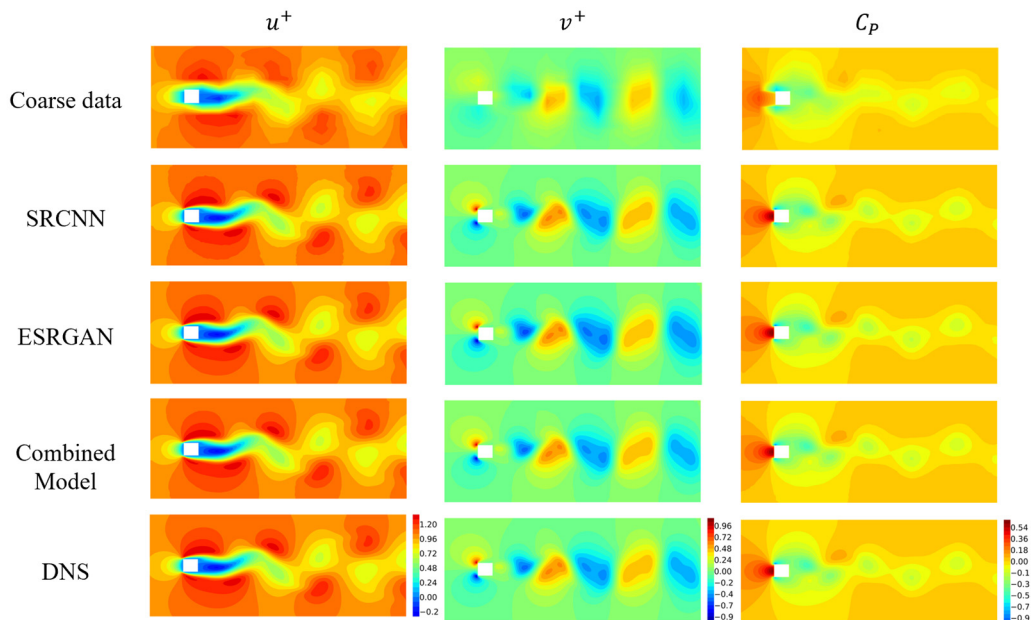

Figure 8. Contours of the two-dimensional flow field around a square cylinder reconstructed by different models, as well as the LR flow field and corresponding DNS results.

To further validate that the 2D flow field around a square cylinder reconstructed by the Combined Model statistically conforms to the DNS results, we compared the PDFs of the flow fields generated by different models with the results of DNS. Figure 9 shows the PDFs of the reconstructed velocity components and pressure generated by SRCNN, ESRGAN, and the Combined Model. The reconstruction results are in outstanding conformity with the DNS-derived results, demonstrating the capability of these three models to accurately reconstruct HR laminar flow around a square cylinder.

To investigate the flow characteristics, the profiles of the mean streamwise velocity and mean pressure are derived from 1000 reconstructed HR flow. As illustrated in Figure 10a,b, the mean streamwise velocity and pressure profiles from all three models demonstrate commendable agreement with the results obtained via DNS.

The L_2 norm is employed to quantify the pointwise error between the reconstructed flow and DNS results at each grid point, where n, m, y_j, $\breve{y}_j$ represent the total number of samples, the total number of grid points per sample, the DNS flow, and the reconstructed results, respectively:

$$D_{L_2}(\breve{y}, y) = \frac{1}{n}\sum_{i=1}^{n}\sqrt{\frac{1}{m}\sum_{j=1}^{m}(\breve{y}_j - y_j)^2},$$ (18)

The pointwise error for the case of flow around a square cylinder reconstructed by different models is shown in the Table 1. In terms of L_2 loss, SRCNN, ESRGAN, and Combined Model have similar performance, which indicates the strength of data-driven learning-based methods. Our method is slightly superior to the other two by a margin.
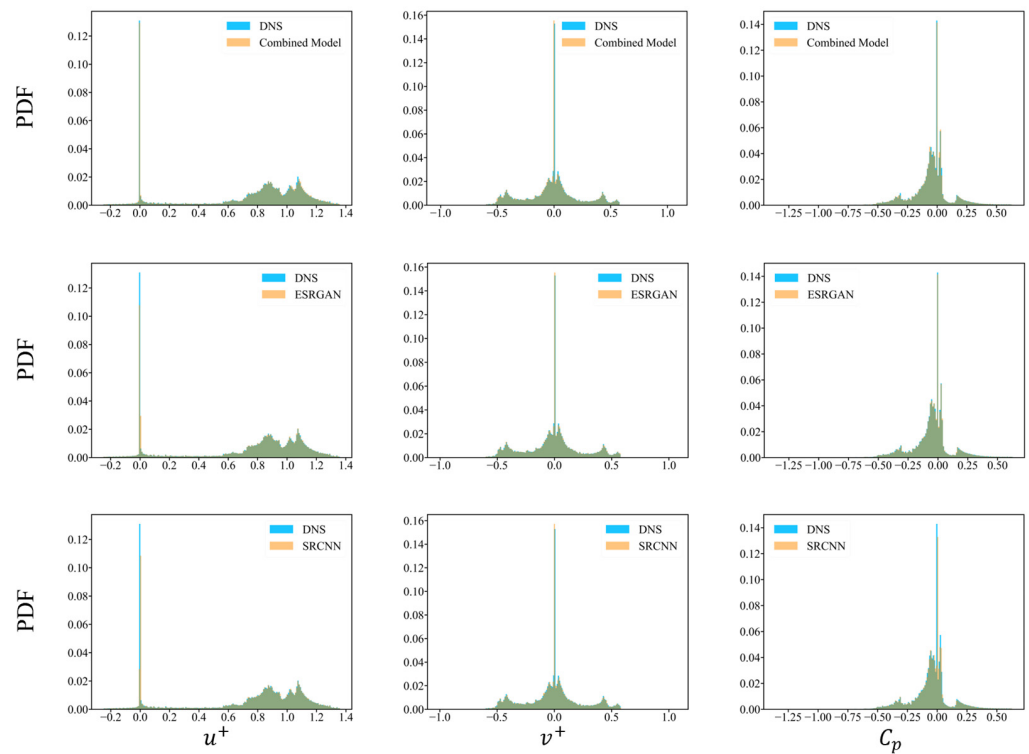
Figure 9. Probability density functions of the reconstructed velocity components and pressure generated by different reconstruction methods and DNS result for 2D flow around a square cylinder.
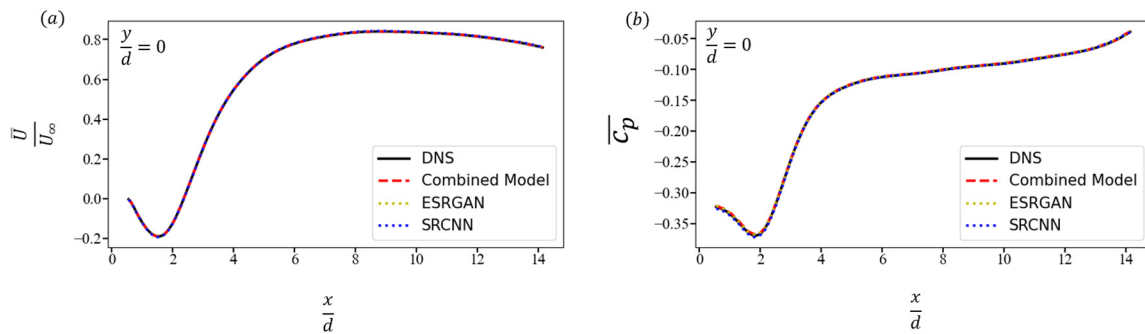

Figure 10. The mean streamwise velocity (**a**) and pressure (**b**) profiles generated by different reconstruction methods and DNS result for the case of 2D laminar flow around a square cylinder.

Table 1. L_2 norms of the reconstructed flow fields from Combined Model, ESRGAN, and SRCNN in the case of flow around a square cylinder.

Flow	Combined Model	ESRGAN	SRCNN
U^+	0.00141	0.00175	0.00453
V^+	0.00160	0.00179	0.00324
C_p	0.00177	0.00189	0.00245

3.3. Turbulent Channel Flow

In this section, the capability of the Combined Model to reconstruct high-resolution turbulent flow fields is validated using a plane perpendicular to the streamwise direction in the turbulent channel flow scenario, specifically the $(y - z)$ plane. The reconstructed instantaneous velocity fields (U^+, V^+, W^+) reconstructed by SRCNN, ESRGAN, and the Combined Model are shown in Figure 11. The results show that compared to SRCNN and

ESRGAN, the turbulent channel flow field reconstructed by our model is more consistent with the DNS results. Furthermore, the flow fields in Figure 11 also clearly show that the Combined Model supplements more details and reconstructs finer structures than the SRCNN and ESRGAN.
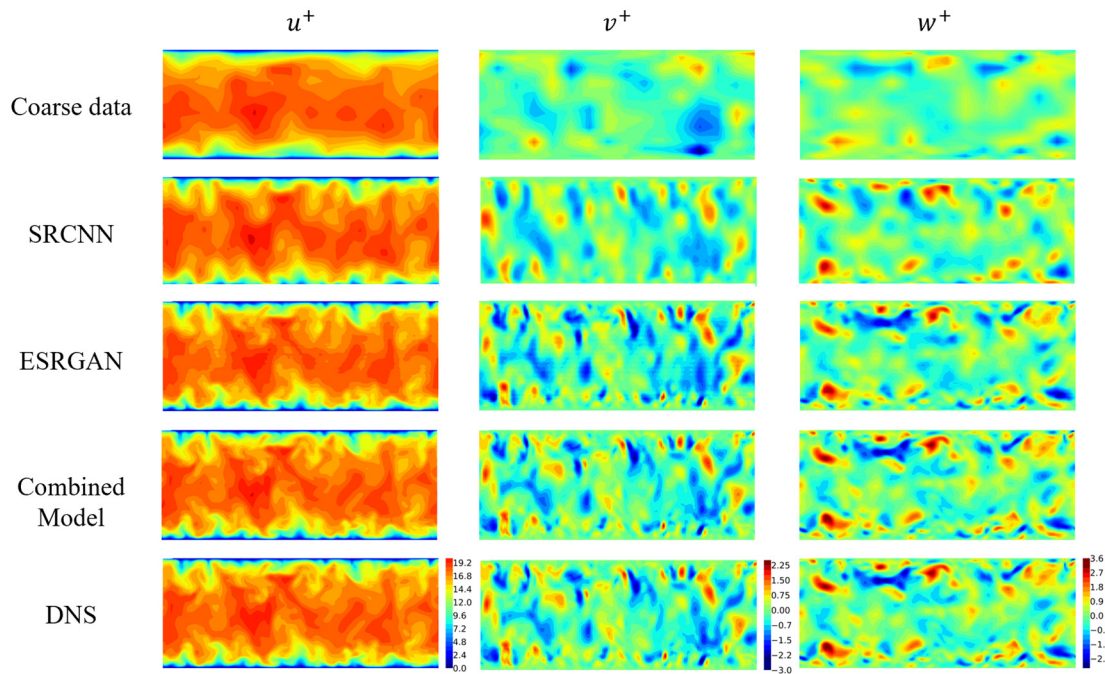

Figure 11. Reconstructed instantaneous velocity fields for the case of turbulent channel flow.

To further verify that the results of the reconstructed turbulent channel flow velocity field of our proposed model are in good agreement with DNS and superior to SRCNN and ESRGAN, we calculated the PDFs for different velocity components of turbulent channel flow, and the results are shown in Figure 12. The results show that the PDF of the velocity component of the Combined Model reconstruction is in good agreement with the flow direction, wall direction, and radial velocity results obtained by DNS. However, there is a bias between the PDF and DNS results of the ESRGAN reconstruction velocity component, especially the wall and span velocities. The deviation between the PDF and DNS results of the SRCNN reconstruction speed component is more obvious, and only the reconstruction effect of flow velocity is close to the DNS results. By comparing the velocity flow field PDF reconstructed by the three models, it is found that the turbulent channel velocity field reconstructed by the Combined Model is closest to the DNS results, indicating that the ability of the Combined Model to reconstruct the turbulent flow field is better than that of ESRGA and SRCNN.

Figure 13 shows the comparison of the statistics of the flow velocity field of the high-precision turbulent channel generated by different models and the statistics of the DNS results. It can be found that the root-mean-square (RMS) profiles of the velocity fields in all three directions of the turbulent channel flow generated by our model are in good agreement with the DNS results, and the performance is better than that of ESRGAN and SRCNN. However, the RMS profiles of the streamwise and wall-normal velocity components (u_{rms}^+, v_{rms}^+) of the ESRGAN results are in good agreement with the DNS results, as shown in Figure 13a,b, and the RMS profile of the spanwise velocity component (w_{rms}^+) is quite different from the DNS results, as shown in Figure 13c. The RMS profiles of the velocity fields in the three directions generated by SRCNN are quite different from the DNS results.
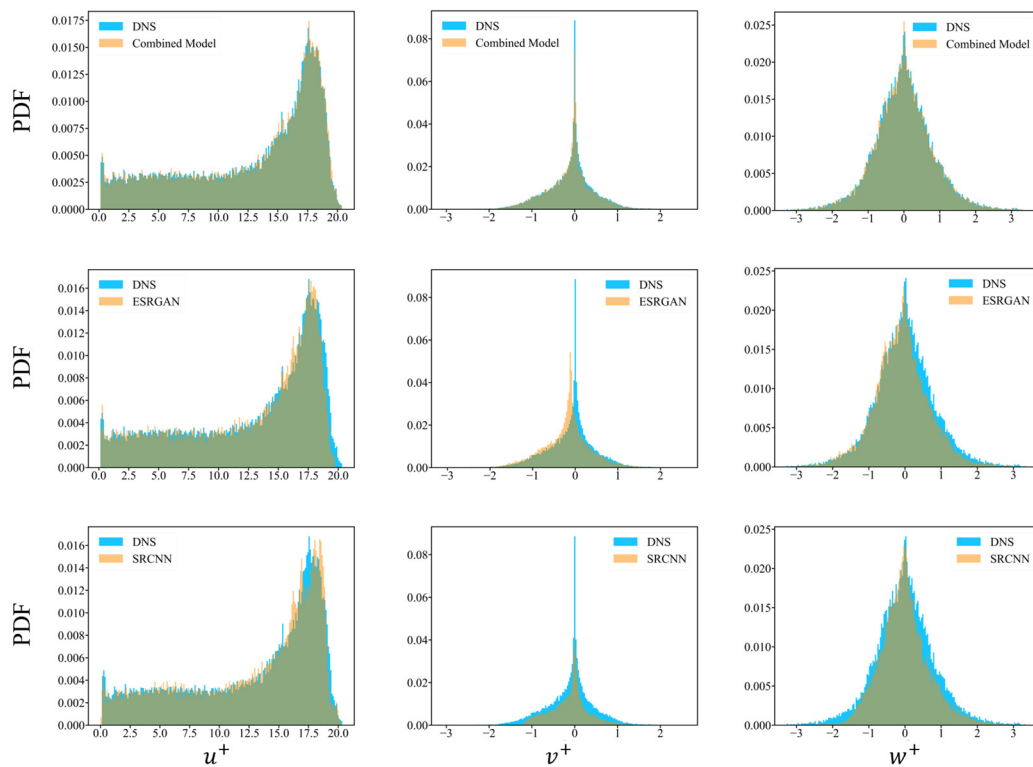
Figure 12. Probability density functions of the reconstructed velocity components and pressure generated by different reconstruction methods and DNS result for the case of turbulent channel flow.
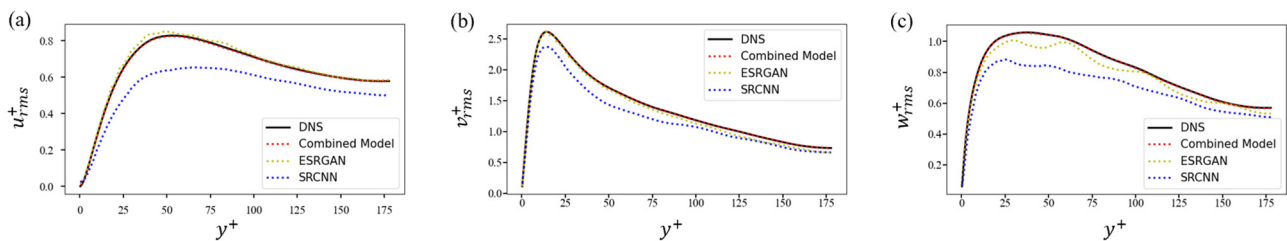

Figure 13. Statistics of different reconstruction methods and reference result for the case of turbulent channel flow: (**a**) RMS profile of the streamwise velocity; (**b**) RMS profile of the wall-normal velocity; (**c**) RMS profile of the spanwise velocity.

The L2 error norms for the models are listed in Table 2. It can be seen that the error norms of the Combined Model are much smaller than SRCNN and ESRGAN. The L2 error norm of the Combined Model in the streamwise velocity is lower by about 77% and 88% than SRCNN and ESRGAN. As for the wall-normal velocity, the Combined Model is lower by about 75.4% and 86.7% than SRCNN and ESRGAN. As for the spanwise velocity, the Combined Model is lower by about 73.3% and 86.6% than SRCNN and ESRGAN. Therefore, the MTPC is the most accurate model.

Table 2. L_2 norms of the reconstructed flow fields from Combined Model, ESRGAN, and SRCNN in the case of turbulent channel flow.

Flow	Combined Model	ESRGAN	SRCNN
U^+	0.00585	0.02540	0.04839
V^+	0.00300	0.01219	0.02241
W^+	0.00373	0.01394	0.02781

4. Conclusions

In this study, we propose a Combined Model of a diffusion model (SR3) and an enhanced residual network (EResNet) for reconstructing high-resolution (HR) instantaneous turbulent flow field from low-resolution (LR) data. We divided an HR flow field reconstruction task into two steps by combining SR3 and EResNet: SR3 reconstructs an initial flow field consistent with the real flow distribution, while EResNet improves the accuracy of flow fields while preserving its distribution. To maximize the preservation of the distribution of flow fields in the second step, we propose an EResNet with long skip connections, directly linking the flow field generated by SR3 to the output layer of EResNet. Additionally, we have incorporated physical constraints into the loss function of EResNet to generate more realistic flow fields. The capabilities of the model to reconstruct laminar flows around a two-dimensional square cylinder at a Reynolds number (Re) of 100 and wall turbulence in channel flows at Re = 4000 were evaluated using data from direct numerical simulations (DNSs). Compared with other deep learning-based reconstruction methods, our model has a slight advantage in reconstructing the flow around a square cylinder, but in the more complex turbulent channel flow reconstruction, our model can generate an HR flow field with higher accuracy, more detail, and more consistent statistical results with DNS results. The experiment results show that as for L2 error norms, our model results are lower by more than 70% and 80% than the enhanced super-resolution generative adversarial network (ESRGAN) and super-resolution convolutional neural network (SRCNN).

Due to the diffusion model requiring T iterations (in this study, T = 200) to generate high-resolution flow fields, the reconstruction time for our model is longer compared to SRCNN-based and SRGAN-based models. With identical model parameters, the time to reconstruct high-precision flow fields is approximately T times that of the other models. To overcome this limitation, we can consider applying acceleration strategies for sampling process of diffusion model to our model in the future.

Author Contributions: Conceptualization, J.Q. and H.M.; methodology, J.Q.; software, J.Q.; validation, J.Q. and H.M.; formal analysis, J.Q. and H.M.; investigation, J.Q.; data curation, J.Q.; writing—original draft preparation, J.Q.; writing—review and editing, H.M.; visualization, J.Q.; supervision, H.M. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: Data are available upon request.

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

Nomenclature

y	high resolution flow
x	low resolution flow
y_t	noisy flow field
y^i	the i-th high-resolution flow field sample
x^i	the i-th low-resolution flow field sample
F	functional mapping
T	the number of iterations in the reverse process
Greek letters	
θ	model parameters
λ	the coefficient of the gradient loss function
Abbreviations	
EResNet	enhanced residual network
CFD	computational fluid dynamics
DNS	direct numerical simulation
CNN	convolutional neural network
GAN	generative adversarial network
HR	high-resolution
LR	low-resolution

SR	super-resolution
PDF	probability density function
RMS	root-mean-square

Appendix A

Given the known training sample data, the parameters of the model can be estimated using maximum likelihood estimation,

$$\mathcal{L} = \mathbb{E}_q[-\log p_\theta(y_0 \mid x)] \tag{A1}$$

For neural network models, the general optimization approach is to find the minimum of the network model through the loss function, and maximizing expectation is not very effective. Therefore, by changing the approach to seek the maximum likelihood estimation of the model, it is equivalent to minimizing the variational upper bound of the negative log-likelihood L_{vlb},

$$\mathbb{E}[-\log p_\theta(y_0 \mid x)] \leq \mathbb{E}_q[-\log \frac{p_\theta(y_{0:T} \mid x)}{q(y_{1:T} \mid y_0)}] =: L_{vlb} \tag{A2}$$

Rewrite the L_{vlb} according to the KL divergence,

$$L_{vlb} = L_0 + L_1 + \ldots + L_{T-1} + L_T \tag{A3}$$

where

$$L_0 = -\log p_\theta(y_0 \mid y_1, x) \tag{A4}$$

$$L_{t-1} = D_{KL}(q(y_{t-1} \mid y_t, y_0) \| p_\theta(y_{t-1} \mid y_t, x)) \tag{A5}$$

$$L_T = D_{KL}(q(y_T \mid y_0) \| p(y_T)) \tag{A6}$$

In L_{vlb}, each term (except L_0) compares two Gaussian distributions, and hence they can be calculated in closed form. L_T is a constant and can be ignored during training since q has no learnable parameters, and y_T is the Gaussian noise. Ho et al. in 2020 modeled L_0 using a separate discrete decoder, which is derived from $\mathcal{N}(y_0; \theta_\theta(y_1, 1), \mathbf{\Sigma}_\theta(y_1, 1))$

Recall Equation (11), the variance $\sigma_t^2 \mathbf{I}$ of $p_{\theta_1}(y_{t-1} \mid y_t, x)$ has been replaced by $1 - \alpha_t$. The KL divergence L_{t-1} can be converted into

$$L_{t-1} \propto \|\widetilde{\mu}_t(y_t, y_0) - \mu_\theta(y_t, x, t)\|^2 \tag{A7}$$

where $\widetilde{\mu}_t(y_t, y_0)$ is the mean of $q(y_{t-1} \mid y_t, y_0)$. $\mu_\theta(y_t, x, t)$ is the training objective of the model.

In practical tests, researchers found that training the model to predict the noise component at any given time step *t* yields better results. Thus, we obtained an objective function for predicting noise:

$$L(\theta) := \mathbb{E}_{t,x,y_0,\epsilon}[\|\epsilon - \epsilon_\theta(\sqrt{\gamma}y_0 + \sqrt{1-\gamma}\epsilon, x, t)\|^2] \tag{A8}$$

References

1. Shi, Y.; Lan, Q.; Lan, X.; Wu, J.; Yang, T.; Wang, B. Robust optimization design of a flying wing using adjoint and uncertainty-based aerodynamic optimization approach. *Struct. Multidiscip. Optim.* **2023**, *66*, 110. [CrossRef]
2. Moin, P.; Mahesh, K. Direct numerical simulation: A tool in turbulence research. *Annu. Rev. Fluid Mech.* **1998**, *30*, 539–578. [CrossRef]
3. Shi, Y.; Song, C.; Chen, Y.; Rao, H.; Yang, T. Complex standard eigenvalue problem derivative computation for laminar–turbulent transition prediction. *AIAA J.* **2023**, *61*, 3404–3418. [CrossRef]
4. LeCun, Y.; Bengio, Y.; Hinton, G.J.N. Deep learning. *Nature* **2015**, *521*, 436–444. [CrossRef] [PubMed]
5. Kutz, J.N. Deep learning in fluid dynamics. *J. Fluid Mech.* **2017**, *814*, 1–4. [CrossRef]

6.   Dong, C.; Loy, C.C.; He, K.; Tang, X. Learning a deep convolutional network for image super-resolution. In *Computer Vision–ECCV 2014: Proceedings of the 13th European Conference, Zurich, Switzerland, 6–12 September 2014*; Proceedings, Part IV 13; Springer: Berlin/Heidelberg, Germany, 2014; pp. 184–199.

7.   Dong, C.; Loy, C.C.; Tang, X. Accelerating the super-resolution convolutional neural network. In *Computer Vision–ECCV 2016: Proceedings of the 14th European Conference, Amsterdam, The Netherlands, 11–14 October 2016*; Proceedings, Part II 14; Springer: Berlin/Heidelberg, Germany, 2016; pp. 391–407.

8.   Kim, J.; Lee, J.K.; Lee, K.M. Accurate image super-resolution using very deep convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 1646–1654.

9.   Zhang, Y.; Li, K.; Li, K.; Wang, L.; Zhong, B.; Fu, Y. Image super-resolution using very deep residual channel attention networks. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 286–301.

10.   Sha, F.; Zandavi, S.M.; Chung, Y.Y. Fast deep parallel residual network for accurate super resolution image processing. *Expert Syst. Appl.* **2019**, *128*, 157–168. [CrossRef]

11.   Ledig, C.; Theis, L.; Huszár, F.; Caballero, J.; Cunningham, A.; Acosta, A.; Aitken, A.; Tejani, A.; Totz, J.; Wang, Z. Photo-realistic single image super-resolution using a generative adversarial network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 4681–4690.

12.   Wang, X.; Yu, K.; Wu, S.; Gu, J.; Liu, Y.; Dong, C.; Qiao, Y.; Change Loy, C. Esrgan: Enhanced super-resolution generative adversarial networks. In Proceedings of the European Conference on Computer Vision (ECCV) Workshops, Munich, Germany, 8–14 September 2018.

13.   Vu, T.; Luu, T.M.; Yoo, C.D. Perception-enhanced image super-resolution via relativistic generative adversarial networks. In Proceedings of the European Conference on Computer Vision (ECCV) Workshops, Munich, Germany, 8–14 September 2018.

14.   Yang, F.; Yang, H.; Fu, J.; Lu, H.; Guo, B. Learning texture transformer network for image super-resolution. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 5791–5800.

15.   Lu, Z.; Li, J.; Liu, H.; Huang, C.; Zhang, L.; Zeng, T. Transformer for single image super-resolution. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 457–466.

16.   Saharia, C.; Ho, J.; Chan, W.; Salimans, T.; Fleet, D.J.; Norouzi, M. Image super-resolution via iterative refinement. *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**, *45*, 4713–4726. [CrossRef]

17.   Ho, J.; Saharia, C.; Chan, W.; Fleet, D.J.; Norouzi, M.; Salimans, T. Cascaded diffusion models for high fidelity image generation. *J. Mach. Learn. Res.* **2022**, *23*, 2249–2281.

18.   Li, H.; Yang, Y.; Chang, M.; Chen, S.; Feng, H.; Xu, Z.; Li, Q.; Chen, Y. Srdiff: Single image super-resolution with diffusion probabilistic models. *Neurocomputing* **2022**, *479*, 47–59. [CrossRef]

19.   Fukami, K.; Fukagata, K.; Taira, K. Super-resolution reconstruction of turbulent flows with machine learning. *J. Fluid Mech.* **2019**, *870*, 106–120. [CrossRef]

20.   Fukami, K.; Fukagata, K.; Taira, K. Machine-learning-based spatio-temporal super resolution reconstruction of turbulent flows. *J. Fluid Mech.* **2021**, *909*, A9. [CrossRef]

21.   Onishi, R.; Sugiyama, D.; Matsuda, K. Super-resolution simulation for real-time prediction of urban micrometeorology. *SOLA* **2019**, *15*, 178–182. [CrossRef]

22.   Kong, C.; Chang, J.-T.; Li, Y.-F.; Chen, R.-Y. Deep learning methods for super-resolution reconstruction of temperature fields in a supersonic combustor. *AIP Adv.* **2020**, *10*, 115021. [CrossRef]

23.   Kong, C.; Chang, J.; Wang, Z.; Li, Y.; Bao, W. Data-driven super-resolution reconstruction of supersonic flow field by convolutional neural networks. *AIP Adv.* **2021**, *11*, 065321. [CrossRef]

24.   Liu, B.; Tang, J.; Huang, H.; Lu, X.-Y. Deep learning methods for super-resolution reconstruction of turbulent flows. *Phys. Fluids* **2020**, *32*, 025105. [CrossRef]

25.   Yousif, M.Z.; Yu, L.; Lim, H.-C. High-fidelity reconstruction of turbulent flow from spatially limited data using enhanced super-resolution generative adversarial network. *Phys. Fluids* **2021**, *33*, 125119. [CrossRef]

26.   Yousif, M.Z.; Yu, L.; Lim, H.-C. Super-resolution reconstruction of turbulent flow fields at various Reynolds numbers based on generative adversarial networks. *Phys. Fluids* **2022**, *34*, 015130. [CrossRef]

27.   Xu, W.; Luo, W.; Wang, Y.; You, Y. Data-driven three-dimensional super-resolution imaging of a turbulent jet flame using a generative adversarial network. *Appl. Opt.* **2020**, *59*, 5729–5736. [CrossRef]

28.   Yu, L.; Yousif, M.Z.; Zhang, M.; Hoyas, S.; Vinuesa, R.; Lim, H.-C. Three-dimensional ESRGAN for super-resolution reconstruction of turbulent flows with tricubic interpolation-based transfer learning. *Phys. Fluids* **2022**, *34*, 125126. [CrossRef]

29.   Xie, Y.; Franz, E.; Chu, M.; Thuerey, N. tempoGAN: A temporally coherent, volumetric GAN for super-resolution fluid flow. *ACM Trans. Graph. (TOG)* **2018**, *37*, 1–15. [CrossRef]

30.   Shu, D.; Li, Z.; Farimani, A.B. A physics-informed diffusion model for high-fidelity flow field reconstruction. *J. Comput. Phys.* **2023**, *478*, 111972. [CrossRef]

31.   Sohl-Dickstein, J.; Weiss, E.; Maheswaranathan, N.; Ganguli, S. Deep unsupervised learning using nonequilibrium thermodynamics. In Proceedings of the International Conference on Machine Learning, Lille, France, 6–11 July 2015; pp. 2256–2265.

32.   Ho, J.; Jain, A.; Abbeel, P. Denoising diffusion probabilistic models. *Adv. Neural Inf. Process. Syst.* **2020**, *33*, 6840–6851.

33.   Meng, C.; Song, Y.; Song, J.; Wu, J.; Zhu, J.-Y.; Ermon, S. Sdedit: Image synthesis and editing with stochastic differential equations. *arXiv* **2021**, arXiv:2108.01073.

34. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: Proceedings of the 18th International Conference, Munich, Germany, 5–9 October 2015*; Proceedings, Part III 18; Springer: Berlin/Heidelberg, Germany, 2015; pp. 234–241.

35. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.