Increasing Image Resolution for Fluid Flows with Physics-Aware Deep Learning

Junyi Guo[1], Adrien Sade[2], Yinxiao Zhang[1], Pu Ren[2], Michael W Mahoney[2,3], and N. Benjamin Erichson[2]

[1]Department of Mechanical Engineering, UC Berkeley
[2]Department of Industrial Engineering and Operations Research, UC Berkeley
[3]Lawrence Berkeley National Laboratory
[4]ICSI and Department of Statistics, UC Berkeley

Contents

Executive Summary Single-Image Super-Resolution is the action to increase the resolution of a low-quality image to a high-resolution image, given only this low-quality image. This is a hard problem as it requires to produce new data and might have an infinite number of solutions. Over the years, scientists have been studying this question and developed various models. Using

Deep Learning, they built models based on convolutional neural network, recurrent neural network, generative network or more recently transformers. In this paper, we build on these models to design a neural network specifically for increasing the resolution of low-quality images of fluid flows. If done correctly it can have a wide variety of applications including the improvement of weather forecasting or quality of fluid simulations for entertainment. To achieve the best results, we use the physics equations governing the fluids' behaviour and embody them inside our model. In short, we design a physics-aware neural network for fluid flows' snpashots super-resolution.

1 Introduction

High-resolution data is increasingly demanded across a variety of industries, including the entertainment industry, medical imaging and weather forecasting. Given the large number of applications: simulating blood flow in arteries, predicting ocean current behavior, or enhancing visual effects in movies and video games, the fluid flows data specifically has always been a major interesting field. Unfortunately, the enormous processing cost associated with modelling the Navier-Stokes equations, which govern fluid flows, makes it difficult to obtain high-resolution photographs of fluid flows only by solving them numerically.

A way to produce high-resolution data, is to increase the resolution of already existing low-resolution data. Taking a low-resolution image as the input to produce a high-resolution image as the output, is what we call Single Image Super Resolution [1] [17] [18]. Such a problem is ill-posed, because given only the low-resolution image, an infinite number of high-resolution images could fit, and that makes it difficult to solve.

In this research project, we suggest a physics-aware deep learning model, Phy-Swin-IR for increasing fluid flow snapshot resolution. Our goal is to create a neural network model that can process low-resolution snapshots of fluid flows as input and produce high-resolution pictures that accurately depict the flow's underlying physics.

Figure 1: The model takes a low-resolution fluid snapshot in input and outputs the corresponding high-resolution snapshot

To achieve our goal, we have thoroughly studied existing deep learning-based approaches for super-resolution, including architectures based on convolutional neural networks (CNNs) [3] [14] [4] [8] [15], generative adversarial networks (GANs) [5] [10], recurrent neural networks (RNNs) [6] [11], and transformers [12]. We have re-implemented and trained these models on our fluid flow data to compare their performance with our proposed model. Our model is designed to not only generate high-resolution images but also to capture the underlying physics of the fluid flow, which is essential for accurate simulations and predictions.

In this report, we outline our suggested approach for super-resolving fluid flows using deep learning. We first propose an overview of the important milestones in the previous work done around single image super resolution [3] [10] [12]. We then present the model designed to increase the resolution of fluid flow snapshots in great details. We also explain the different experiments carried out to ensure and measure the efficiency of our model. And finally, we discuss the limitations of our work and its possible extensions in the last part.

Our research has a wide range of possible applications, including those in the entertainment, medical imaging, and forecasting the weather. To enhance the precision of weather forecasts, our model can be used to create high-resolution visualizations of air flows and ocean currents. Our model can be utilized in medical imaging to simulate blood flow in arteries and increase the

precision of medical diagnostics. Furthermore, the data storage industry as a whole may undergo a transformation thanks to our deep learning-based method. We can decrease storage requirements without sacrificing quality by lowering the quality of the data and using the model to boost resolution, which will result in considerable cost savings and improved productivity. This last example is of course only hypothetic, but as the first step of a barely studied field, we like to think that this is a possible application which could be a great revolution for the digital world.

We would like to emphasize that our project is a research project and does not aim at creating a final product in the time-span allocated. However, we believe that our work can have a significant impact on various fields and can pave the way for future research in this area. As far as we know, our research is one of the first attempts to increase the resolution of fluid flows using deep learning, and we are excited to explore the potential applications of our approach.

Overall, this research project presents a promising approach to increasing the resolution of fluid flow snapshots using deep learning. We believe that our work can contribute to the advancement of this exciting area of research and can have a significant impact on various fields.

2 Overview of Super-Resolution methods

2.1 Traditional method: Bicubic Interpolation

Bicubic interpolation [7] was the first method introduced to increase the resolution of images. It was used for the first time in the super resolution context in 1981. It was laying the foundations of the super resolution field. This model is only using mathematical functions to approximate the value of the missing pieces in the high-resolution picture - namely - it interpolated the missing values with the original neighbors. Given the simplicity of such a model, it yielded relatively poor results in terms of metrics as well as for the human perception, creating exclusively blurry images.

Up until 2013, models have tried to optimize and complexify this method [13], improving slightly the results. But in the end, the general idea remained the same and people were using the bicubic interpolation as their base to build upon.

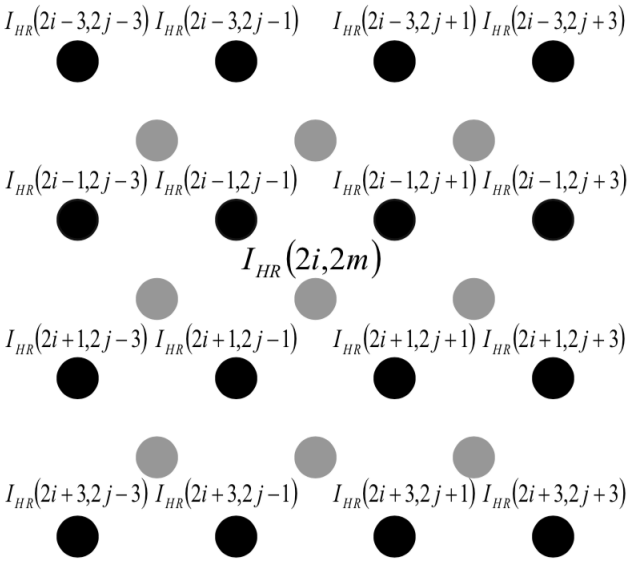

Figure 2: Advanced interpolation method [13] - Grey point in the center $I_{\mathrm{HR}}(2i, 2m)$ results from averaging its neighbors with different weights for greys and blacks.

2.2 Deep Leanring method: SRCNN

The first model of super resolution using deep learning was published at 2014. SRCNN [3] (for Super-Resolution CNN) is a very simple model based on the use of convolutional neural networks (CNNs). CNNs were brand new in the world of deep learning and, when we know their efficiency working with image-data, it was only natural to try them out in the context of super resolution.

The model is built in three parts. The first one is designed to extract the local information of the input image. The second part is dedicated to creating a non-linear mapping to the high-resolution space dimension. Finally, the last part reconstructs the target image. The model is only using two CNNs, in other words, very few parameters.

From this paper on, all relevant super resolution models we came across used deep learning and CNNs in their architecture. Some built directly on top of the idea of SRCNN: ESPCN [14] is one of them, it is basically following the same architecture as SRCNN but chooses to upsample the image at the end of the process with a deconvolution layer (also known as pixel-shuffle layer) [19]. This

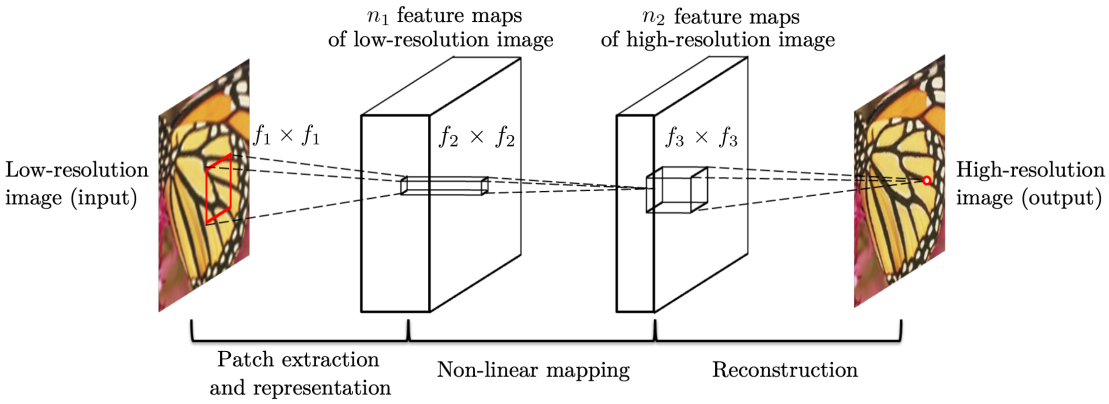

Figure 3: SRCNN Model Architecture [3] - Based on a sequential 3-layer CNN

inverted process allows for a greatly reduced computational time, such that most later papers will use this trick [4].

Other papers tried new layouts and features, still building on the work of SRCNN, but with a new level of complexity. A good example of those is the VDSR model [8] which is based on the VGG architecture [15] while using the SRCNN idea. It's a deep network that tries to learn the residual of the output image and the interpolated input (with a skip connection), instead of studying the direct mapping.

2.3 Complex deep learning methods: SRResNet & SRGAN

One of the complex improvement that has been made is the use of residual skip connections in the architecture of the model. The SRResNet model [10], is using a large number of residual (convolutional) blocks which are fed with the direct and skip connections. There are multiple skip connections throughout the model to help it keep track of the global information while processing the local information. The model is widely complexifies with the use of those blocks and connections, in addition to layers of batch normalization and deconvolution. Overall, when it came out the model was producing state-of-the-art results for image super resolution.

The SRGAN model [10] is essentially the same model as SRResNet but based on a GAN process. A Generative Adversarial Network (GAN) [5] is built with two components. The first component, the generator, is designed to create new images out of noise (that is to say, out of nothing). It aims at creating the most realistic images possible. The second component is dedicated to checking how likely is an image to be real. Given an image input, it will yield a probability of the input being fake. To make a GAN efficient, we train the generator and the discriminator alternatively in competition with each other: the generator trains to create more and more realistic image with the objective to pass the discriminator's test (it is trying to design images so realistic that the discriminator predicts it to be real), conversely the discriminator continuously improve to always predict the images from the generator to be fake even when they become highly realistic. Doing so, we can reach very high levels of performance. The SRGAN model is based on this GAN architecture and on the SRResNet's, for single-image super-resolution and yields way better results than the original SRResNet.

The GAN model is really advanced and provides impressive results when correctly trained. However, it is also extremely hard to train. Some researchers have been studying methods to improve the stability of GANs' training [16], but in the end it makes it difficult to work with.

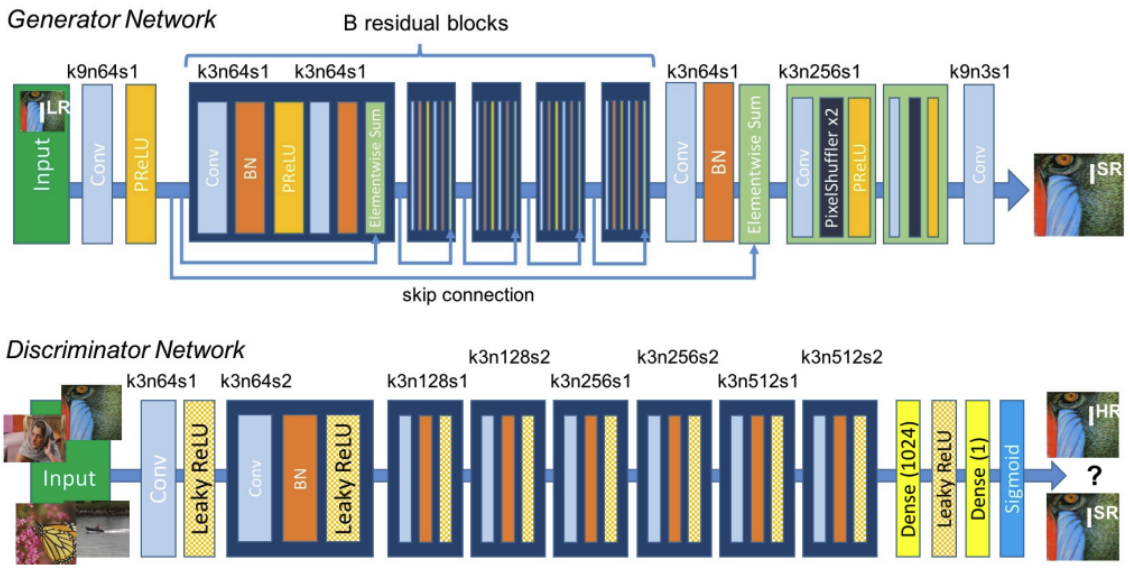

Figure 4: SRGAN Model Architecture [10] - Top: Generator Architecture; Bottom: Discriminator Architecture

2.4 State-of-the-art method: SwinIR

SwinIR [12], published in 2021, is the current state-of-the-art neural network model for single image super resolution. This model introduces a new architecture based on transformers, while using the previous work on super resolution (CNNs for feature extraction, deconvolution layer for upsampling at the end of the model, skip connections to keep track of the global information). Given the great results, and easy to use architecture of the SwinIR model, we decided to build our own super resolution model for fluid flows snapshots upon this one. The architecture will be depicted more precisely in the next section.

3 Method

3.1 SwinIR model

At the core of the SwinIR model lies the shifted window-based self-attention mechanism. This mechanism is designed to reduce the computational complexity of traditional self-attention mechanisms by dividing the input into non-overlapping windows and applying self-attention within each window. The windows are then shifted by half of their size in both height and width dimensions, and self-attention is applied again. This process is repeated, ensuring that all pixels are attended to by the surrounding pixels. The resulting model is efficient, flexible, and scalable, making it suitable for handling a wide range of image restoration tasks.

In the SwinIR model, the input image is first divided into non-overlapping patches, which are then linearly embedded into a sequence of patch tokens. These patch tokens are then processed through a series of Swin Transformer layers. Each layer comprises several stages, including a patch merging stage, a shifted window self-attention stage, and a patch splitting stage. The patch merging stage reduces the spatial resolution of the input while increasing the channel dimension, whereas the patch splitting stage reverses this process, restoring the original spatial resolution.

During the shifted window self-attention stage, the model learns to capture both local and long-range dependencies in the input image. This is accomplished through the use of multi-head self-attention mechanisms, which allow the model to attend to different spatial locations and capture various levels of semantic information. Additionally, the model employs position-sensitive layer normalization, which takes into account the spatial information of the input image, ensuring that the model remains sensitive to the spatial structure of the image throughout the restoration process.

The SwinIR model also features skip connections between the input and output of each Swin Transformer layer. These skip connections facilitate the flow of information across layers, enabling the model to learn more efficiently and maintain better gradient flow during the optimization process. Furthermore, the model employs residual connections within each layer, allowing it to learn incremental improvements over the initial input representation.

3.2 The physics informed loss function

In our practice, we discovered most deep learning super-resolution models, including the state-of-the-art SwinIR model, are fail to adhere to the fundamental principles of fluid dynamics, leading to physically inconsistent results. To address this issue, we propose incorporating the governing equations of fluid dynamics as physics-based constraints within the loss function (Equation B.1) of our deep learning model. By doing so, we aim to obtain predictions that are not only accurate but also conform to the underlying physical laws. In order to achieve this, we employ gradient-free convolutional filters, which enable the calculation of spatial derivatives within the Navier-Stokes equations. The weights of the convolutional kernels are determined by finite difference scheme. The details of mathematics derivation can be found in Appendix A.

3.3 Evaluation Metrics

To evaluate the performance of our models comprehensively, we utilize three widely-used metrics in addition to the data loss and physics loss (defined in Appendix B): the peak signal-to-noise ratio (PSNR), structural similarity index measure (SSIM), and relative Frobenius error (RFNE).

PSNR measures the fidelity of the restored image compared to the original, defined as the ratio between the maximum possible power of a signal and the power of the corrupting noise, expressed in decibels (dB). Higher PSNR values signify better image quality and more accurate restoration.

SSIM, a more sophisticated evaluation metric, accounts for structural information, luminance, and contrast between the original and restored images. Designed to better correspond with the human visual system's perception of image quality, SSIM considers spatial dependencies and structural relationships between image regions. SSIM values range from -1 to 1, where a value of 1 represents a perfect match between the original and restored images. Consequently, higher SSIM values indicate better image quality and more accurate restoration.

RFNE, another evaluation metric, measures the difference between original and restored images in terms of their matrix representations. Defined as the ratio of the Frobenius norm of the difference between the original and restored matrices to the Frobenius norm of the original matrix, RFNE utilizes the Frobenius norm, which is the square root of the sum of the squared matrix elements. Lower RFNE values suggest better image quality and more accurate restoration.

4 Experiments

4.1 Experiment set-up

ur experiment is conducted on numerically simulated Krichnan Turbulence data with Re=16000. We selected two sub-regions with a spatial resolution of (128,128) as our training and evaluation sets, respectively.

We conducted two sets of experiments with different scale factors. In the first experiment, we used a scale factor of x8, which means that the input resolution was (16,16). In the second experiment, we used a scale factor of x16, which means that the input resolution was (8,8).

The purpose of the experiment was to evaluate the performance of our P-SwinIR model in comparison to the traditional SwinIR model on various metrics, including SSIM, PSNR, RFNE, MSE, and Physics Loss. The results of the experiment are presented in Table 1.

4.2 Results

For a scale factor of 8, our P-SwinIR model outperforms the SwinIR model in terms of SSIM, PSNR, RFNE, and MSE, with improvements of 0.003, 1.34, 0.014, and 0.069, respectively. Notably, the P-SwinIR model significantly reduces the Physics Loss to 0.137, which is nearly five times better than the SwinIR model's Physics Loss of 0.613. These results indicate that our P-SwinIR model produces more accurate and visually appealing images while also adhering to the laws of physics.

At a larger scale factor of 16, our P-SwinIR model performs significantly better than the SwinIR model in terms of PSNR, SSIM, RFNE, MSE, and Physics Loss. Specifically, our model exhibits a ten-fold improvement in PSNR compared to the SwinIR model, which is a significant improvement. Additionally, our P-SwinIR model reduces the Physics Loss to 0.269, which is less than half of the SwinIR model's Physics Loss of 0.558. These results demonstrate that our P-SwinIR model produces more accurate and visually appealing images while also ensuring adherence to physics laws.

Overall, the results suggest that our P-SwinIR model is a promising approach for image processing and reconstruction tasks, particularly in challenging scenarios where larger scale factors are required. Moreover, the model's ability to reduce the Physics Loss highlights its utility in applications that demand adherence to physical laws.

Model	Scale Factor	λ	SSIM ↑	PSNR ↑	RFNE ↓	MSE ↓	Physics Loss ↓
P-SwinIR (Ours)	x8	0.01	0.990	33.047	0.126	0.293	0.137
SwinIR	x8	0	0.987	31.707	0.140	0.362	0.613
P-SwinIR (Ours)	x16	0.01	0.931	22.441	0.319	1.879	0.269
SwinIR	x16	0	0.926	21.56	0.336	2.095	0.558

Table 1: Comparison of models on various metrics.

5 Discussion

5.1 Limitations & Future Work

Our current study is limited by the computational resources available, focusing on low spatial-resolution data (128,128) with two sub-regions. As a result, the impact of higher resolution on our model's performance, whether positive or negative, remains undetermined. Future research should investigate this issue using higher-resolution datasets, such as 1024 by 1024, and varying Reynolds numbers. By doing so, we can further assess the generalization capabilities of our model and better understand its performance across diverse conditions. This limitation induce a great unknwon about the generalization of our work. Furthermore, it underlines how the hardware limitations curve research and constitute a major obstacle to the development of new technologies such as the building of a universal deep-learning tool to increase image resolution whatever the context.

In addition to data limitations, the incorporation of physics-based information in this project is not fully optimized. Currently, we only include the continuity equation (Equation A.8) as the physics loss term. However, this alone is insufficient for enhancing accuracy, as it provides limited information about the physical system under study. To improve the model's performance and accuracy, future research should explore incorporating a more comprehensive set of physics-based loss terms derived from relevant governing equations. Thus, while our work provides better outputs from a physics point of view, we cannot say that it increased significantly the state-of-the-art results from a mathematics point of view.

Introducing time-dependency to the model presents its own challenges, as adding time as an additional variable would significantly increase computational costs. Furthermore, the time derivative within the Navier-Stokes equation poses considerable difficulty, making it challenging to construct a comprehensive loss term. Researchers should consider these factors and develop innovative strategies to include essential physics information without incurring prohibitive computational expenses or compromising the model's performance.

Moreover, it would be promising to explore the development of a novel architecture tailored for handling spatio-temporal data. Currently, our model is restricted to increasing resolution solely in the spatial domain. It would be highly advantageous, however, to extend the model's capabilities to interpolate across time steps as well. One potential avenue for improvement could involve applying dynamic systems perspectives [2][9].

5.2 Conclusion

Although our setup has certain limitations, it is evident that we have added significant value to our area of research by substantially enhancing the physics-behavior of the model output. Our model, physics-aware Swin-IR, successfully produces physics consistent high resolution result with low resolution input.

Our achievement is the result of building upon state-of-the-art methods in the domain. We thoroughly reviewed the literature and selected the most suitable existing model, and then leveraged our expertise to modify a specific aspect of the model to tailor it to our specific needs. We incorporated a physics term in the loss to ensure that the output prediction adheres to physical constraints. Our results demonstrate substantial improvements in the selected metrics, thereby validating our success in providing increased-resolution data for fluid flows that are realistic from human, mathematical, and physical perspectives.

We believe that by following this path and continuously improving our model, we can expect that our projected applications may become a reality in the near or distant future. It is plausible that scientists will utilize such models to augment and improve their available data, which could have

far-reaching implications. Furthermore, our model has the potential to impact everyday life, such as in the field of movie-making, where low-resolution simulations could be improved by a similar model as ours, resulting in more realistic fluid simulations. Similarly, our model could contribute to advancements in weather forecasting.

A Appendix: Calculation of physics informed loss

A.1 Derivation of weights in the gradient free convolution filter

For a continuous function f, we can discrete this by using Taylor theorem

$$f(x+\Delta x) = f(x) + \Delta x f'(x) + \Delta x^2 \frac{f''(x)}{2!} + \Delta x^3 \frac{f'''(\xi_1)}{3!}, \quad \xi_1 \in (x, x+\Delta x)$$
$$f(x-\Delta x) = f(x) - \Delta x f'(x) + \Delta x^2 \frac{f''(x)}{2!} - \Delta x^3 \frac{f'''(\xi_2)}{3!}, \quad \xi_2 \in (x-\Delta x, x)$$

Rearrange the equations, we have

$$f(x+\Delta x) - f(x-\Delta x) = 2\Delta x f'(x) + \Delta x^3 \frac{(f'''(\xi_1) + f'''(\xi_2))}{6}$$

So that a second order central difference scheme for the first derivative of function f at location x is:

$$f'(x) = \frac{f(x+\Delta x) - f(x-\Delta x)}{2\Delta x} + \Delta x^2 \frac{(f'''(\xi_1) + f'''(\xi_2))}{12} \tag{A.1}$$

In 2D Cartesian coordinate, we obtain two kernels K_x and K_y for calculating the first derivative with respect to x and y as follows:

$$K_x = \begin{bmatrix} 0 & 1/2 & 0 \\ 0 & 0 & 0 \\ 0 & -1/2 & 0 \end{bmatrix} \tag{A.2}$$

$$K_y = \begin{bmatrix} 0 & 0 & 0 \\ -1/2 & 0 & 1/2 \\ 0 & 0 & 0 \end{bmatrix} \tag{A.3}$$

Similarly, the fourth order central difference scheme central difference scheme for the first derivative of function f at location x is:

$$f(x+\Delta x) = f(x) + \Delta x f'(x) + \Delta x^2 \frac{f''(x)}{2!} + \Delta x^3 \frac{f'''(x)}{3!} + \Delta x^4 \frac{f^{(4)}(x)}{4!} + \Delta x^5 \frac{f^{(5)}(\xi_1)}{5!}$$
$$f(x-\Delta x) = f(x) - \Delta x f'(x) + \Delta x^2 \frac{f''(x)}{2!} - \Delta x^3 \frac{f'''(x)}{3!} + \Delta x^4 \frac{f^{(4)}(x)}{4!} - \Delta x^5 \frac{f^{(5)}(\xi_2)}{5!}$$
$$f(x+2\Delta x) = f(x) + 2\Delta x f'(x) + 4\Delta x^2 \frac{f''(x)}{2!} + 8\Delta x^3 \frac{f'''(x)}{3!} + 16\Delta x^4 \frac{f^{(4)}(x)}{4!} + 32\Delta x^5 \frac{f^{(5)}(\xi_3)}{5!}$$
$$f(x-2\Delta x) = f(x) - 2\Delta x f'(x) + 4\Delta x^2 \frac{f''(x)}{2!} - 8\Delta x^3 \frac{f'''(x)}{3!} + 16\Delta x^4 \frac{f^{(4)}(x)}{4!} - 32\Delta x^5 \frac{f^{(5)}(\xi_4)}{5!}$$

$$f'(x) = \frac{-f(x+2\Delta x) + 8f(x+\Delta x) - 8f(x-\Delta x) + f(x-2\Delta x)}{12\Delta x} + \mathcal{O}(\Delta x^4) \tag{A.4}$$

Therefore, the forth order kernels K_x and K_y for calculating the first derivative with respect to x and y are:

$$K_x = \begin{bmatrix} 0 & 0 & 1/12 & 0 & 0 \\ 0 & 0 & -8/12 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 8/12 & 0 & 0 \\ 0 & 0 & -1/12 & 0 & 0 \end{bmatrix} \tag{A.5}$$

$$K_y = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1/12 & -8/12 & 0 & 8/12 & -1/12 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{A.6}$$

A.2 Governing equations and discretization

For the 2D incompressible fluid flow, the governing equation is:

$$(\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \boldsymbol{\nabla} \mathbf{u})\rho = -\boldsymbol{\nabla} P + \mu \nabla^2 \mathbf{u} + \mathbf{f} \tag{A.7}$$

$$\boldsymbol{\nabla} \cdot \mathbf{u} = 0 \tag{A.8}$$

In 2D, we can expand the continuity as follows where u and v are velocity field in the x and y direction.

$$\boldsymbol{\nabla} \cdot \mathbf{u} = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \tag{A.9}$$

We can discretely calculate the spatial derivatives in equation A.9 using kernel convolutions of the u and v components, where the kernels are derived in the previous section.

$$\frac{\partial u}{\partial x} = (u * K_x)[i, j] \tag{A.10}$$

$$\frac{\partial v}{\partial y} = (v * K_y)[i, j] \tag{A.11}$$

B Appendix: Modified optimization functions

The loss function can set up as follow:

$$\mathcal{L} = \mathcal{L}_{data} + \lambda \mathcal{L}_{phy} \tag{B.1}$$

, where our optimization objective is

$$\min_{\theta} \underbrace{||\hat{x}^{HR} - x^{HR}||_1}_{\mathcal{L}_{data}} + \lambda \underbrace{||\mathcal{R}(\hat{x}^{HR})||_2}_{\mathcal{L}_{physics}} \tag{B.2}$$

Consider fluid flow with divergence function as physics constraint in particular:

$$\min_{\theta}(||u^{HR} - \hat{u}^{HR}||_1 + \lambda||\boldsymbol{\nabla} \cdot \hat{u}^{HR}||_2) \tag{B.3}$$

In the discrete scheme, the physics loss function can be written as:

$$\mathcal{L}_{phy} = \sum_{j=1}^{N_x} \sum_{i=1}^{N_y} (\partial_x \mathbf{u}_{ij} + \partial_y \mathbf{v}_{ij}) \tag{B.4}$$

where

$$\hat{x}^{HR} = \mathcal{N}(x^{LR}; \theta) \tag{B.5}$$

References

[1] Saeed Anwar, Salman Khan, and Nick Barnes. A deep journey into super-resolution: A survey, 2019.

[2] Ricky T. Q. Chen, Yulia Rubanova, Jesse Bettencourt, and David Duvenaud. Neural ordinary differential equations, 2019.

[3] Chao Dong, Chen Change Loy, Kaiming He, and Xiaoou Tang. Image super-resolution using deep convolutional networks, 2015.

[4] Chao Dong, Chen Change Loy, and Xiaoou Tang. Accelerating the super-resolution convolutional neural network, 2016.

[5] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial networks, 2014.

[6] Wei Han, Shiyu Chang, Ding Liu, Mo Yu, Michael Witbrock, and Thomas S. Huang. Image super-resolution via dual-state recurrent networks, 2018.

[7] R. Keys. Cubic convolution interpolation for digital image processing. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 29(6):1153–1160, 1981.

[8] Jiwon Kim, Jung Kwon Lee, and Kyoung Mu Lee. Accurate image super-resolution using very deep convolutional networks, 2015.

[9] Aditi S. Krishnapriyan, Alejandro F. Queiruga, N. Benjamin Erichson, and Michael W. Mahoney. Learning continuous models for continuous physics, 2022.

[10] Christian Ledig, Lucas Theis, Ferenc Huszar, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, and Wenzhe Shi. Photo-realistic single image super-resolution using a generative adversarial network, 2016.

[11] Junyeop Lee, Jaihyun Park, Kanghyu Lee, Jeongki Min, Gwantae Kim, Bokyeung Lee, Bonhwa Ku, David K. Han, and Hanseok Ko. Fbrnn: feedback recurrent neural network for extreme image super-resolution. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 2021–2028, 2020.

[12] Jingyun Liang, Jiezhang Cao, Guolei Sun, Kai Zhang, Luc Van Gool, and Radu Timofte. Swinir: Image restoration using swin transformer, 2021.

[13] Jing Liu, Zongliang Gan, and Xiuchang Zhu. Directional bicubic interpolation — a new method of image super-resolution. In *International Conference on Model Transformation*, 2013.

[14] 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, 2016.

[15] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition, 2014.

[16] Zhendong Wang, Huangjie Zheng, Pengcheng He, Weizhu Chen, and Mingyuan Zhou. Diffusion-GAN: Training GANs with diffusion. In *The Eleventh International Conference on Learning Representations*, 2023.

[17] Zhihao Wang, Jian Chen, and Steven C. H. Hoi. Deep learning for image super-resolution: A survey, 2019.

[18] Yang Wenming, Xuechen Zhang, Yapeng Tian, Wei Wang, and Jing-Hao Xue. Deep learning for single image super-resolution: A brief review, 08 2018.

[19] Matthew D. Zeiler, Graham W. Taylor, and Rob Fergus. Adaptive deconvolutional networks for mid and high level feature learning. In *2011 International Conference on Computer Vision*, pages 2018–2025, 2011.