

High-Resolution Reconstruction Model for Spatial Temperature Distribution in Grain Storage Based on 3DSRCNN_Resnet

Yujie Cao^a, Jie Yuan^b, Junli Li^c, Bo Mao^{d,*}

Nanjing University of Finance & Economics, Nanjing, 210023, China

^acaoyj678@163.com, ^b620182913@qq.com, ^c2631846925@qq.com, ^dbo.mao@nufe.edu.cn

*Corresponding author

Abstract: Temperature data inside the granary is a crucial indicator for safe storage and quality preservation of grain. Currently, temperature measurement relies on temperature cables installed inside the grain piles. However, the sensors on the cables can only acquire temperature data at specific measurement points and cannot provide a complete overview of the entire grain pile. Most current temperature field generation methods based on conventional spatial interpolation techniques often suffer from significant errors. To achieve a more intuitive and accurate monitoring of temperature distribution in granary environments, this study proposes a high-resolution reconstruction model for grain temperature fields based on the 3DSRCNN_Resnet. The model firstly improves the SRCNN in 3D to obtain the 3DSRCNN model, so as to apply the SRCNN super-resolution reconstruction algorithm, which deals with 2D images, to the task of reconstructing the temperature field of a three-dimensional granary. Secondly, a nonlinear mapping module based on 3D residual blocks and a residual reconstruction module based on 3D deconvolution are proposed to adjust the network structure of the 3DSRCNN model. The experimental results show that compared with the traditional spatial interpolation method, the grain temperature data reconstructed by this model achieves better performance under MSE, MAE, PSNR and SSIM metrics. This model provides a novel approach for comprehensive and accurate understanding of temperature distribution within grain piles.

Keywords: Spatial interpolation, High-resolution reconstruction, Grain Storage, 3DSRCNN_Resnet

1. Introduction

Grain temperature is an important indicator for monitoring and analyzing the safety status of stored grain, and accurate temperature control is crucial for granary management^[1-3]. Currently, temperature measurement in granary mainly relies on temperature cables installed inside the grain piles^[4]. However, the sensors on the cables can only provide temperature data at specific measurement points, failing to capture the complete temperature distribution within the grain pile. Therefore, to visually and accurately monitor the temperature distribution in granary, specific three-dimensional visualization methods are required to obtain spatial temperature distribution information and variations within the grain pile^[5].

Spatial interpolation techniques^[6] play a key role in visualizing grain temperature distribution. These techniques utilize spatial correlations to interpolate temperature data, thereby obtaining denser temperature information and providing more detailed temperature spatial distribution. Currently, common spatial interpolation methods include Linear interpolation, Nearest-neighbor interpolation, Kriging interpolation, Radial basis function interpolation, Inverse distance weighting interpolation, and others. With the rise of neural networks, researchers have started to incorporate neural networks into the field of spatial interpolation. For example, in 2021, Wang et al.^[7] combined BP neural network with Kriging interpolation to develop a temperature field prediction model. Li^[8] proposed a grain temperature spatial interpolation algorithm using BP neural networks, ordinary Kriging interpolation, and inverse distance weighting interpolation.

Traditional high-resolution image reconstruction algorithms^[9] aim to establish the feature relationship between low-resolution (LR) images and high-resolution (HR) images based on known image information. Currently, most existing high-resolution reconstruction algorithms are based on two-dimensional (2D) images, with limited research on high-resolution reconstruction of three-dimensional (3D) spatial data. In this study, we propose a high-resolution reconstruction model for spatial temperature

distribution in granary based on 3DSRCNN_Resnet. Specifically, we abstract the temperature field in granary as a 3D volume, with each voxel representing a discrete temperature measurement point. The sparse temperature field is treated as 3DLR data, and the interpolated temperature field serves as 3DHR data. The 3DLR data is fed into the 3DSRCNN_Resnet model, which is trained to learn and generate corresponding 3DHR data, i.e. the reconstructed temperature field. This model leverages the sparsely measured temperature data from sensors inside the grain pile to generate a more detailed temperature field matrix, providing a new approach for visually and accurately monitoring the temperature distribution in granary. Compared to traditional linear interpolation, the trained model in this study demonstrates superior performance in terms of MSE, MAE, PSNR, and SSIM metrics, thereby showing potential for practical applications.

2. Overview of SRCNN High-Resolution Reconstruction Algorithm

The SRCNN network model is a pioneering work in deep learning-based image super-resolution reconstruction. In 2014, Dong et al.^[10] from The Chinese University of Hong Kong first used a three-layer convolutional neural network to model the nonlinear relationship between high-resolution images and low-resolution images, making it the first deep learning-based end-to-end image super-resolution algorithm. The SRCNN network model has proven that the application of deep learning in the super-resolution field can surpass traditional methods, such as interpolation methods and reconstruction methods, achieving better results and performance. The network structure of the SRCNN network model is shown in Figure 1.

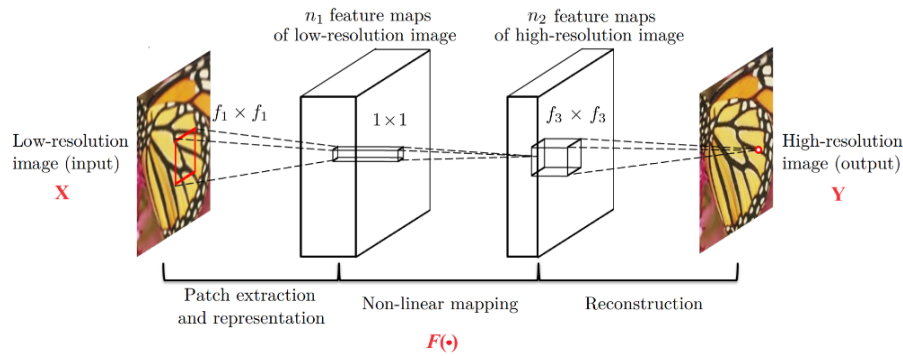


Figure 1: SRCNN Network Structure.

SRCNN preprocesses the input low-resolution images using Bicubic interpolation method. The interpolated low-resolution images are denoted as X , while the original high-resolution images are denoted as Y . The function representing the network mapping is denoted as $F(\cdot)$. The network mapping is divided into three modules: Patch extraction and representation, Non-linear mapping, and Reconstruction. The patch extraction and representation operation is achieved by obtaining the desired feature maps from the input image, i.e. extracting the image features through CNN and storing them into vectors. The non-linear mapping operation corresponds to the "convolution + activation" operation, i.e., further processing the feature map from the previous layer with non-linear mapping to increase the network depth, which is more conducive to learning. The reconstruction operation uses convolution without an activation function, borrowing from the pure interpolation method of traditional super-resolution, i.e. the idea of averaging the local image. The formula for the three modules are expressed as follows:

$$F_1(X) = \max(0, W_1 * X + B_1); F_2(X) = \max(0, W_2 * F_1(X) + B_2); F(X) = W_3 * F_2(X) + B_3 \quad (1)$$

$F_1(X)$ represents patch extraction, $F_2(X)$ represents non-linear mapping, and $F(X)$ represents reconstruction. In the above formulas, $W_1W_2W_3$ and $B_1B_2B_3$ respectively represent the weights and biases of the filter (convolution kernel), and the max function represents the ReLU activation function.

3. High-Resolution Reconstruction Model for Grain Temperature Based on 3DSRCNN_Resnet

3.1. Data Analysis and Processing

In granary, the layout of cable temperature measurement points is typically as shown in Figure 2. In this study, the temperature field of the granary is treated as a three-dimensional (3D) spatial dataset with dimensions (N, C, D, H, W), where N represents the number of samples, C represents the number of channels in the 3D spatial data. Since the sensors in this study specifically measure temperature, C=1. If the sensors are capable of simultaneously measuring multiple parameters such as temperature and humidity, the value of C would correspondingly be determined based on the number of measured indicators. Additionally, H, W, and D represent the total number of temperature measurement points in the height, width, and depth dimensions of the granary, respectively.

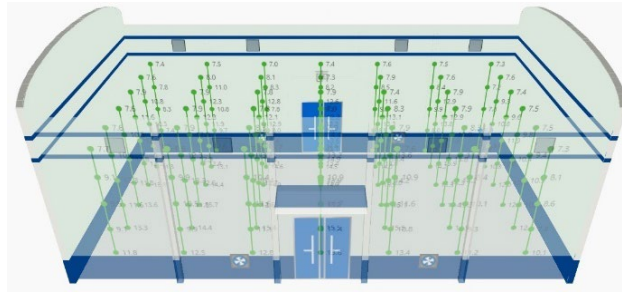


Figure 2: SRCNN Network Structure.

The dataset for this study comprises temperature measurements from various granaries sites in Hubei, China. Each data represents a three-dimensional spatial sample, which consists of all the temperature measurement point data of a grain depot in a granary company under one point in time. The data format of each temperature measurement point is (Temp, x, y, z), with Temp being the point's temperature and x, y, z its coordinates. The dataset combines data from multiple companies and grain depots, the total number of temperature measurement points in each sample and the number of them in the three dimensions of width, height and depth are not fixed. In addition, missing values may occur due to incomplete data. In order to better reconstruct the 3D data, the grain temperature dataset needs to be preprocessed first. The main steps are shown in Figure 3 below:

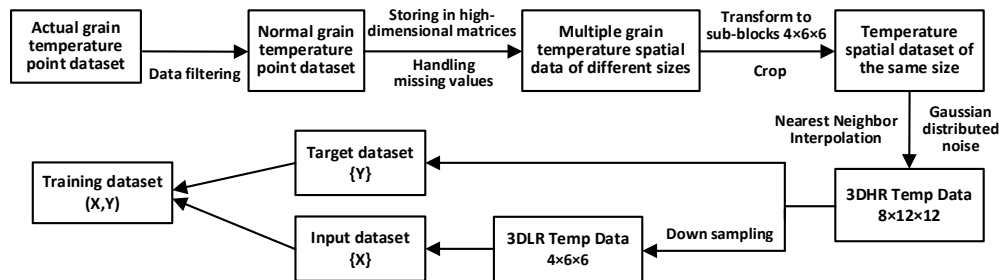


Figure 3: Data Processing Flowchart.

Step 1: Fixing the dimensions of the three-dimensional spatial dataset.

In the context of grain temperature interpolation and reconstruction in this study, the three-dimensional grain temperature spatial data before interpolation is referred to as three-dimensional low-resolution (3DLR) data, while the interpolated and reconstructed grain temperature data is referred to as three-dimensional high-resolution (3DHR) data. To facilitate the subsequent model training, the height, width, and depth of the pre-reconstructed samples are set to $4 \times 6 \times 6$, and the size of the reconstructed grain temperature spatial samples is set to $8 \times 12 \times 12$.

Step 2: Generating the three-dimensional temperature spatial dataset.

First, data filtering, delete the following kinds of data that do not conform to the temperature pattern of the grain store: 1) The overall fluctuation range of the temperature of all the temperature measurement points in a certain data is not more than 0.1°C or more than 40°C . 2) The temperature of some temperature measurement points in a data record is lower than -10°C or higher than 40°C . 3) The spatial dimensions of the data are smaller than $4 \times 6 \times 6$.

Second, the three-dimensional temperature spatial dataset is constructed based on the recorded temperature measurement points (Temp, x, y, z) in each data record. The temperature value (Temp) of each measurement point corresponds to the value of a spatial point in the temperature spatial dataset, which is determined by its coordinates (x, y, z) in the three-dimensional space. By filling each temperature value at the corresponding position, a three-dimensional temperature spatial dataset representing the temperature distribution inside the granary facility is formed. For missing temperature measurement points in the dataset, spatially weighted average interpolation is applied to generate a complete three-dimensional temperature spatial dataset.

Step 3: Generating the input dataset and the target dataset.

In practical applications, it is difficult for cable temperature measurement points to cover a range of $8 \times 12 \times 12$ due to factors such as cost constraints. To overcome this challenge, this study adopts a series of measures to address the data scarcity issue.

First, the target dataset is generated. The three-dimensional spatial dataset generated in Step 2 is divided into fixed-sized samples of $4 \times 6 \times 6$ with a step size of 3. These small samples form a new fixed-sized set of three-dimensional temperature spatial samples. They are then expanded to the scale of $8 \times 12 \times 12$ using the Nearest neighbor interpolation method. To simulate real-world conditions, Gaussian-distributed noise is added to each sample, resulting in a high-resolution reconstructed dataset, which is the target dataset.

Second, the input dataset is generated. A sampled dataset of $4 \times 6 \times 6$ size is obtained by downsampling against this newly generated $8 \times 12 \times 12$ target dataset, which is used as the input dataset. This data processing method based on interpolation and sampling aims to address the issue of data scarcity in practical scenarios. It effectively utilizes the original data to provide rich information for model training while ensuring data quality and reliability. This research approach not only effectively addresses challenges in real-world environments but also provides valuable insights and references for data processing and modeling in related fields.

3.2. Model Construction

In order to successfully apply the SRCNN model to the task of spatial high-resolution reconstruction of grain temperature, the temperature field in the grain depot is first treated as a three-dimensional spatial data. Secondly, the SRCNN model is adjusted and optimized to meet the requirements of grain temperature spatial interpolation tasks, and to ensure that the model can effectively handle the spatial features of three-dimensional data, thus achieving accurate reconstruction of 3DLR grain temperature data. The key step is to extend the SRCNN model from processing two-dimensional image data to processing three-dimensional volumetric data, and setting up the inner 3D convolution residual block and the outer 3D deconvolution residual block to adjust the network structure, enabling it to perform better when processing volumetric data. The structure of the 3DSRCNN_Resnet model proposed in this article is shown in Figure 4:

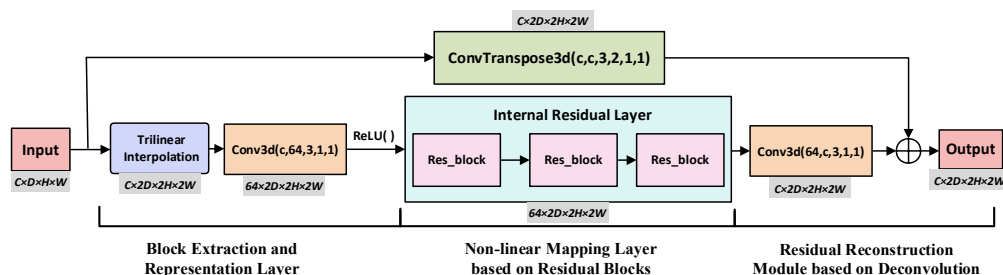


Figure 4: 3DSRCNN_Resnet Model.

The input and output data of this algorithm are both five-dimensional matrices, where the input data has dimensions of [B, C, D, H, W], and the output data has dimensions of [B, C, 2D, 2H, 2W]. Here, B represents the number of data batches and is set to 64; C represents the number of channels and is set to 1, specifically the temperature channel in this context; H represents the height and is set to 4; W represents the width and is set to 6; D represents the depth and is set to 6. The model structure mainly consists of three modules: Patch extraction and representation, Non-linear mapping based on internal residuals, and Reconstruction based on external residuals. The following is a description of each operation:

- (1) Patch extraction and representation layer.

This paper first preprocesses the data, that is, amplifies the spatial dimension of the 3DLR data through Trilinear interpolation method, and obtains three-dimensional low-resolution spatial data with the same spatial dimension as the 3DHR data. Secondly, with the transformation of image data into three-dimensional spatial voxel data, a three-dimensional convolutional neural network is introduced to replace the traditional convolutional neural network, capturing all features along the depth, width, and height directions to extract and represent the features of the data in three-dimensional space.

(2) Non-linear mapping based on 3D residual blocks.

The non-linear mapping layer is constructed by stacking three 3D residual convolutional blocks. The structure of each residual convolutional block is shown in Figure 5: The input data first goes through a 3D convolutional layer and a 3D batch normalization layer to obtain intermediate feature representations. It then undergoes a non-linear mapping through the ReLU activation function, followed by another round of 3D convolutional layer and 3D batch normalization layer. Finally, the processed result is added to the original input data, and the final output is obtained through the ReLU activation function again. The combination of these residual connections and multi-layer convolutional structures strengthens the model's spatial modeling ability for the three-dimensional data, allowing the network to more fully utilize the information in the input data and reduce the problem of vanishing gradients, further enhancing the feature representation capability.

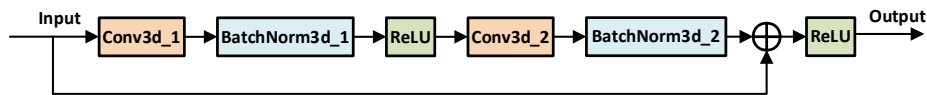


Figure 5: Residual Block.

(3) Residual reconstruction module based on 3D deconvolution.

First, the original input data undergoes deconvolution operations to extend its spatial dimensions, thereby increasing the spatial range of features. This helps better preserve the information from the original input during the reconstruction process. Second, the results obtained after the patch extraction and representation layer and the non-linear mapping layer are convolved to further extraction and dimensionality reduction of the spatial features of the data. Finally, the results of these two parts are added together to fuse the information from different stages, ensuring a comprehensive consideration of both global and local features of the original input data during the reconstruction process, thereby obtaining the reconstructed results of the grain temperature data.

3.3. Experimental Validation and Analysis

In the experiments, the dataset is divided into training and testing sets in a 7:3 ratio. The model is trained with a learning rate of 0.001, 100 epochs, Adam optimizer, and MSE loss function. In this paper, the SRCNN model, originally used for image super-resolution reconstruction tasks, is adapted to three-dimensional grain temperature high-resolution reconstruction tasks by making 3D improvements, two models are obtained: 3DSRCNN and 3DSRCNN_Resnet. The 3DSRCNN model removes both the inner and outer residuals of 3DSRCNN_Resnet, which means removing the external residuals based on 3D deconvolution and the internal residuals in the non-linear mapping layer based on 3D residual blocks. During the training of both models, it is observed that the initial few epochs have relatively high loss values. The MSE loss values for the first three epochs of 3DSRCNN are: 3.096, 1.156, and 0.945, while for 3DSRCNN_Resnet, they are: 9.162, 0.558, and 0.399. Therefore, to clearly show the loss trends in the subsequent training process, the loss curves for both models start from the fourth epoch, as shown in Figure 6. It can be observed that the 3DSRCNN_Resnet model has lower loss values compared to 3DSRCNN in the same epochs, indicating that the improvement of the internal and external residuals based on 3DSRCNN has a more significant impact on the model performance.

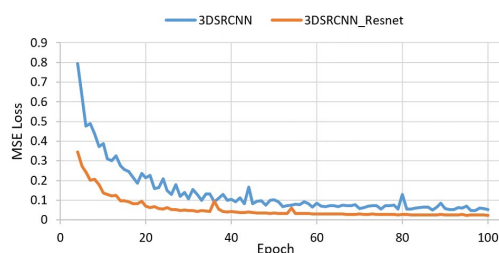


Figure 6: Loss Curve Graph.

To comprehensively evaluate the accuracy of the experimental results, this paper utilizes both traditional metrics and commonly used image reconstruction evaluation indices, namely Mean Absolute Error (MAE), Mean Square Error (MSE), Peak Signal-to-Noise Ratio (PSNR), and Structural Similarity Index (SSIM). In order to adapt to the spatial temperature data reconstruction task of this paper, the PSNR and SSIM formulas are optimized and adjusted, as shown in Formula (2) and (3):

$$PSNR = 10 \cdot \log_{10} \left(\frac{R^2}{MSE} \right) \in [m, n], \quad R = \max(Y_{pred}, Y_{true}) - \min(Y_{pred}, Y_{true}) \quad (2)$$

$$SSIM(Y_{pred}, Y_{true}) = \frac{(2\mu_{Y_{pred}}\mu_{Y_{true}} + C_1)(2\sigma_{Y_{pred}Y_{true}} + C_2)}{(\mu_{Y_{pred}}^2 + \mu_{Y_{true}}^2 + C_1)(\sigma_{Y_{pred}}^2 + \sigma_{Y_{true}}^2 + C_2)} \in [-1, 1] \quad (3)$$

In the above context, Y_{true} and Y_{pred} represent the true 3D high-resolution (3DHR) grain temperature data and the predicted 3DHR grain temperature data, respectively. m and n are constants, typically with $m > 20$ and $n > 50$. R refers to the dynamic range of the data. $\mu_{Y_{pred}}$ and $\mu_{Y_{true}}$ are the mean values of Y_{pred} and Y_{true} , respectively. $\sigma_{Y_{true}}^2$ and $\sigma_{Y_{pred}}^2$ are their variances, and $\sigma_{Y_{pred}Y_{true}}$ is their covariance. C_1 and C_2 are constants used to stabilize the terms in the denominator. MAE indicates the model's prediction accuracy, with a smaller value denoting higher accuracy. MSE assesses the error impact in predictions, with a lower value indicating greater accuracy. PSNR measures the overall error between the interpolation result and the actual temperature field, with a higher value indicating a closer match. SSIM evaluates structural similarity, including spatial distribution and continuity, with a value closer to 1 signifying a closer match to the real data in spatial terms.

To verify the effectiveness of the proposed 3DSRCNN-based inner-outer residual model, this study conducted a series of comparative experiments on the Testing Set with different interpolation methods, including Trilinear interpolation (Trilinear), Nearest Neighbor interpolation (Nearest), Ordinary Kriging interpolation (OK), and Radial Basis Function interpolation (RBF). Additionally, this study evaluated the performance of the 3DSRCNN model and the 3DSRCNN_Resnet model on the Testing Set separately. The experimental comparison results are shown in Table 1:

Table 1: Comparison of Model Results

Model	MSE	MAE	SSIM	PSNR
OK3D	2.815	0.976	0.483	28.855
Nearest	1.965	0.796	0.756	30.416
Trilinear	1.270	0.712	0.790	32.311
RBF	0.754	0.512	0.833	34.795
3DSRCNN	0.047	0.166	0.924	46.694
3DSRCNN_Resnet	0.023	0.112	0.934	49.787

To demonstrate the comparison of model performance, the results under the MSE scoring criterion are divided by 3, and the results under the PSNR scoring criterion are divided by 50. The graph is shown below in Figure 7:

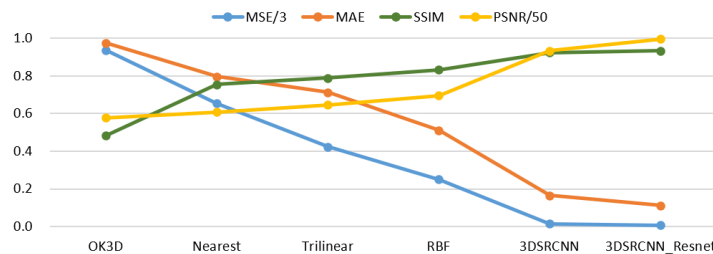


Figure 7: Model Results Comparison Chart

Based on the observations from Table 1 and Figure 7, the values of MSE and MAE show a clear downward trend, while the values of SSIM and PSNR show a clear upward trend. This indicates that the performance of the models, under all scoring criteria, ranks from worst to best as follows: OK3D < Nearest < Trilinear < RBF < 3DSRCNN < 3DSRCNN_Resnet. Particularly, compared to the worst interpolation algorithm OK3D, the 3DSRCNN_Resnet model has reduced MSE and MAE by approximately 99.18% and 88.52%, respectively, and has increased SSIM and PSNR by approximately

93.37% and 72.74%, respectively. This demonstrates that the model proposed in this paper successfully adapts the two-dimensional image super-resolution algorithm to the three-dimensional grain temperature spatial measurement point reconstruction task, and exhibits superior performance compared to traditional interpolation algorithms. Specifically, the 3DSRCNN_Resnet model performs the best under all scoring criteria, reducing MSE and MAE by approximately 51.06% and 32.53%, respectively, and increasing SSIM and PSNR by approximately 1.08% and 6.62%, respectively, compared to the 3DSRCNN model. Therefore, the improvements based on the Non-linear mapping based on 3D residual blocks and the outer residual layer based on 3D deconvolution for the reconstruction module significantly enhance the effectiveness of the 3DSRCNN model.

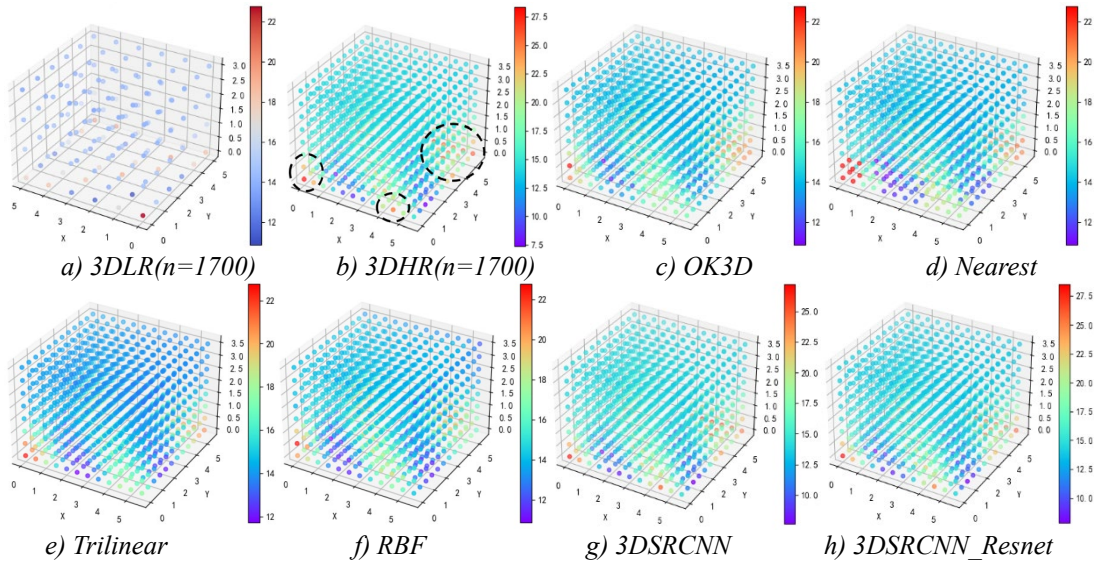


Figure 8: 3D Scatter Plot of the 17000th Sample in Three-Dimensional Space

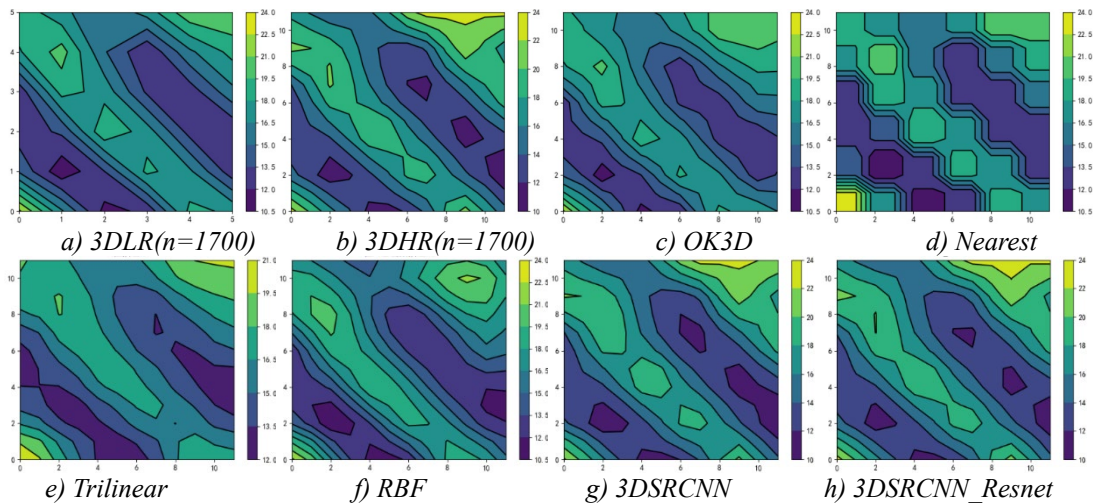


Figure 9: 2D Cross-Sectional View of the 17000th Sample

By analyzing the sub-figures in Figure 8, firstly, it can be seen from sub-figure b) that the temperature range of the 17000th sample is approximately between 7.5°C and 28°C. From sub-figures c), d), e), and f), it can be observed that the temperature range of the temperature field reconstructed using traditional spatial interpolation algorithms is approximately between 11°C and 23°C. From sub-figure g), it can be seen that the temperature range of the temperature field reconstructed using the 3DSRCNN model is approximately between 8°C and 27°C. From sub-figure h), it can be observed that the temperature range of the temperature field reconstructed using the 3DSRCNN_Resnet model is approximately between 8°C and 28°C. Based on the temperature range, the temperature field reconstructed using the 3DSRCNN_Resnet model is more consistent with the actual temperature field. By observing at the places circled with black dashed lines in Fig. b) of the actual temperature field, it can be seen that these positions belong to high-temperature blocks. Compared to other sub-figures, the deep learning-based 3DSRCNN and 3DSRCNN_Resnet models can better reconstruct these high-temperature blocks simultaneously. By

analyzing the temperature and contour lines in each sub-figure of Figure 9, it can be seen that the 3DSRCNN and 3DSRCNN_Resnet models can better reconstruct the temperature distribution of the original temperature field. Moreover, compared to 3DSRCNN, 3DSRCNN_Resnet can better reconstruct the low-temperature and medium-high temperature areas in the cross-section. Therefore, by observing the 3D scatter plot of the 17000th sample under different methods and the temperature change graph of a certain layer's 2D cross-section of the sample, it is once again verified that the proposed 3DSRCNN_Resnet model can achieve better results in the high-resolution reconstruction of storage grain temperature.

4. Summary

For the high-resolution reconstruction task of grain temperature in storage spaces addressed in this paper, the commonly used methods are spatial interpolation techniques such as Linear interpolation, Nearest neighbor interpolation, Kriging interpolation, Radial basis function interpolation and so on. However, these traditional methods struggle to mine the spatial distribution of temperature when the temperature measurement points are sparse in the grain storage. Therefore, this study introduces the SRCNN image super-resolution reconstruction algorithm and proposes a model based on 3DSRCNN_Resnet for high-resolution reconstruction of the grain temperature field in three-dimensional space. This model leverages scattered temperature data measured by temperature sensors within the grain storage to mine the internal temperature variations and generate a more comprehensive temperature field with additional temperature points. It provides a new approach for a more comprehensive and accurate understanding of the temperature distribution inside grain storages. Compared to traditional Linear interpolation, Nearest neighbor interpolation, Kriging interpolation, and Radial basis function interpolation, the trained model in this study exhibits superior performance in testing and demonstrates potential for practical applications.

Acknowledgements

This research has been supported by the National Key R&D Program of China (Grant No. 2022YFD2100200) and the Postgraduate Research & Practice Innovation Program of Jiangsu Province (KYCX22_1699).

References

- [1] FengPei L. *Design and implementation of an intelligent monitoring system for grain situation in the granary* [D]. *Journal of Tiangong University*, 2019.
- [2] Hongchao F. *Pre-warning model of stored grain safety risk based on temperature and humidity and its application* [D]. *Journal of Henan University of Technology*, 2019.
- [3] Mingyue Z. *Research on Spatial Interpolation of Grain Storage Temperature Field and Lost Data Interpolation Algorithm* [D]. *Journal of Beijing University of Posts and Telecommunications*, 2021.
- [4] Daosong Z. *Machine learning and numerical simulation based Grain Pile Temperature Field Prediction and Application* [D]. *Journal of Henan University of Technology*, 2023.
- [5] Yunhao C, Shanshan D, Weidong Y, et al. *A Spatial High-Resolution Reconstruction Model for Warehouse Grain Piles' Temperature and Humidity Fields*[J/OL]. *Journal of the Chinese Cereals and Oils Association*, 2024: 1-13.
- [6] Haitao L, Zedong S. *Review of Spatial Interpolation Analysis Algorithm*[J]. *Computer Systems & Applications*, 2019, 28(07): 1-8.
- [7] Chuanxu W, Kang W, Lin C, et al. *Research and Realization of Granary Temperature field Prediction Model Based on Kriging Interpolation and BP Neural Network*[J]. *Journal of Agricultural Science and Technology*, 2021, 23(09): 96-102.
- [8] Jicao L. *Research on 3D Visualization of grain storage temperature based on Neural Network*[D]. *Journal of Beijing University of Posts and Telecommunications*, 2021.
- [9] Caidong Y, Chengyang L, Zhongbo L, et al. *Review of Image Super-resolution Reconstruction Algorithms Based on Deep Learning*[J]. *Journal of Frontiers of Computer Science and Technology*, 2022, 16(9): 1990.
- [10] Dong C, Loy C C, He K, et al. *Image super-resolution using deep convolutional networks*[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2015, 38(2): 295-307.