

# Computational Spectrum Reconstruction Using the Self-Attention Method

Hao Wu<sup>1</sup>, Hui Wu<sup>1</sup>, Xinyu Su<sup>1</sup>, Fei Zhang<sup>2</sup>, Shuangli Liu<sup>1,\*</sup>

**Correspondence:** Shuangli Liu (liushuangli2301@gmail.com),

1.School of Information Engineering, Southwest University of Science and Technology, Mianyang 621010, China

2.CSIC (Wuhan) Lingjiui Electrical Technology Co.Ltd, Wuhan Digital Engineering Institute, Wuhan 430074, China

## **Abstract:**

Spectrometers can already give ultra-fine resolution and a broad spectrum range, but their systems are complex and bulky. And, as the spectroscopic application space expands fast, the requirement for portable or integrated spectroscopy equipment with decreased physical size, cost, or power consumption takes precedence over the necessity for excellent performance. Furthermore, the spectra of a newly emerged miniature computational spectrometer must be computationally recreated, however, the accuracy of spectral reconstruction based on existing algorithms is currently unsatisfactory. As a result, this work analyzes the reconstruction matrix's design and uses a self-

attentive mechanism algorithm to improve spectral reconstruction accuracy. The self-attentive mechanism algorithm is then applied to the spectral reconstruction, yielding  $R^2$  0.9780 and MSE 0.0019 when compared to a commercial spectrometer (USB2000, Ocean Optics Inc., America), which is an important reference for the improvement and optimization of the spectral reconstruction technique.

**Keywords: Spectral reconstruction, self-Attention, Encoding Matrix, Spectral Data**

### **Introduction:**

Computational spectral reconstruction is a novel sort of spectral reconstruction approach that differs from the standard grating spectrometer, which directly performs spectroscopy or filter design to obtain the spectral information of an item. Computational spectral system is more through the design of different material layers and coding matrix, it can use computer algorithms to process the optical signal, achieve the reconstruction of the spectrum, and acquire the spectral information of the object. When compared to traditional spectrometers, the computational spectral reconstruction type has the benefits of low cost, small size, and

simple operation, thus it has a broad application promise in various applications.[1]-[6]

Deep neural networks can learn the mapping between input and target output distribution by training a large number of examples [7], to which the addition of the self-attention mechanism can better learn more relevant features in the data, attention mechanism has been widely used in various applications in recent years with the emergence of deep learning, which can solve a series of complex problems [8], text classification [9][10], speech translation[11], natural language generation [12], image recognition [13][14], image generation [15][16], etc. However, most earlier computational spectral reconstructions were based on enhancements to various compressed perception algorithms [17]. For example, early spectroscopy reconstructions were developed utilizing the CS approach, with iterations leading to fully converged results. Recently, good results were obtained by introducing neural networks for the spectral reconstruction of computational spectral systems, demonstrating that deep learning can be very helpful for spectral reconstruction; however, the accuracy of the spectral reconstruction is not perfect because it only uses linear transformations of the fully connected layer itself and does not take into account the correlation between the spectrum and the spectra

[18][19][20]. As a result, the accuracy and stability of the spectral reconstruction are improved by introducing residual and self-attention mechanisms, where the residual mechanism can record and save the better results of the previous step, allowing the network to avoid overfitting when the number of layers is too deep, and the self-attention mechanism can better improve the correlation like spectral coding, allowing the features to be better extracted for recognition.

To improve spectral reconstruction accuracy, the simulated transmittance function library is inter-correlated, and the 100 transmittance functions with the weakest correlation and retained the coding effect are chosen to construct an encoding matrix, and the spectral data set is trained using the self-attentiveness mechanism with the addition of residuals.

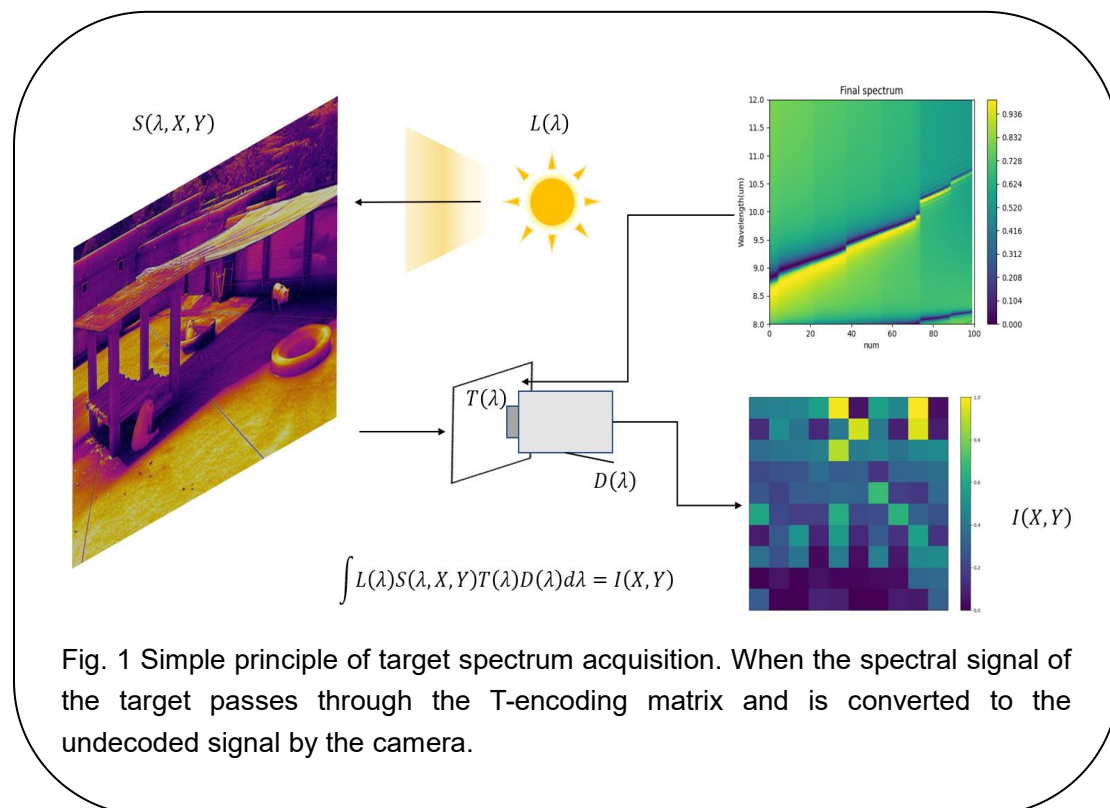
The study is organized as follows: first, the principle of how the spectral camera is used to obtain undecoded data after the encoding matrix is introduced, and then the principle of how this data is reconstructed by the designed neural network algorithm to obtain the object's spectral information. The design of the encoding matrix and data construction is then described, and different noise coefficients are added to simulate realistic environmental noise in the data set's construction, followed by a presentation of

the reconstruction results, which are divided into sampling results and overall reconstruction results. Finally, the work is summarized and it is discussed how better results might be obtained using the self-attentive mechanism.

## Result:

### Spectral camera diagram

The reflected spectrum of a target sample in sunlight is measured in this part, and the simplified schematic representation is illustrated in Figure 1. To begin, the designed coding matrix must



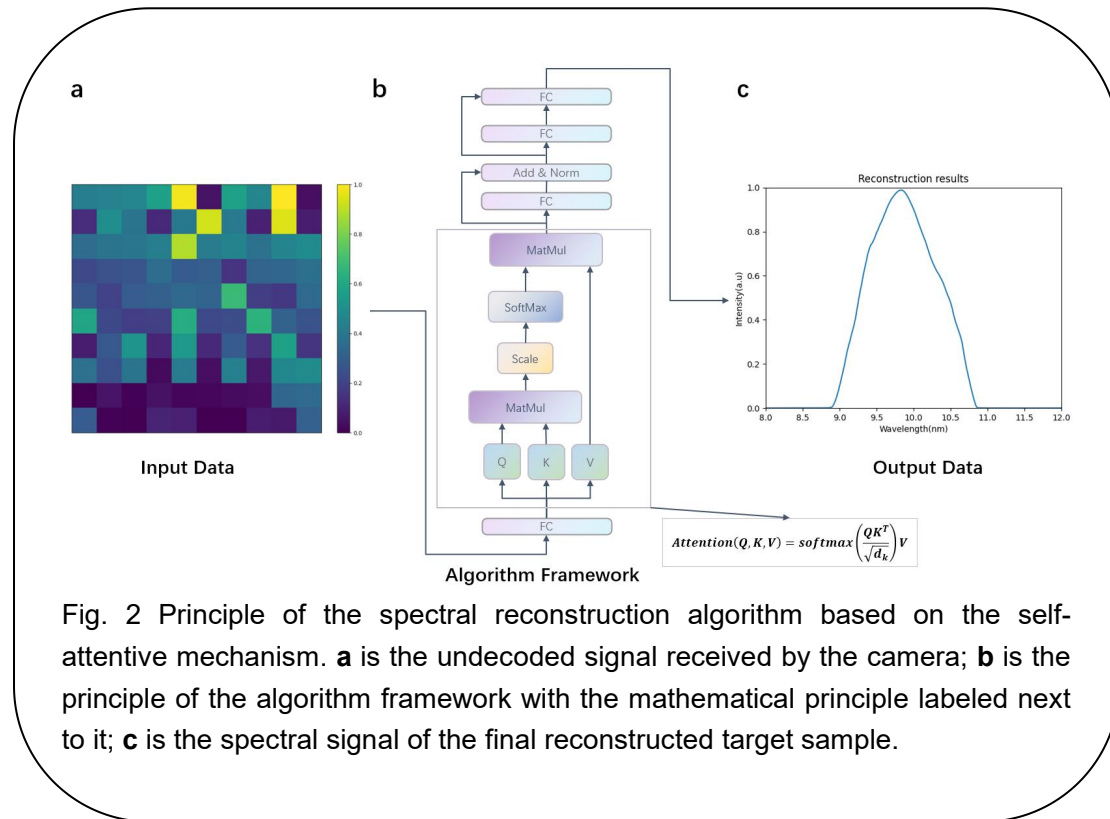
be attached in front of the camera to test the spectral information of the sample's reflected light  $S(\lambda, X, Y)$ , so the spectral intensity is

$L(\lambda)S(\lambda, X, Y)$  before entering the camera conversion, and then the undecoded intensity  $I(X, Y)$  is obtained after the coding matrix  $T(\lambda)$  and the camera response function  $D(\lambda)$ . To make spectrum measurement more convenient, the design encoding matrix can be turned into a chip that fits exactly with the camera.

### **Spectral reconstruction of the self-attentive mechanism framework**

Figure 2b depicts the algorithm model. The model employs a self-attentive method to enhance the relationship between characteristics to extract more valuable features from this randomly encoded input for better reconstruction. The self-attentiveness mechanism computes the similarity between each site and other locations and uses the similarity as a weight to weigh the features together, focusing attention on the key features to improve result accuracy. The residual mechanism involves adding cross-layer connections to the neural network layers, which can successfully avoid problems like gradient disappearance and gradient explosion while also improving model training [22]. The self-attention mechanism and residuals are used in this framework for training. First, the input incoming data is subjected to self-attentiveness operation to obtain more relevant features, and then

residuals are introduced in each neural layer to prevent parameter explosion and overfitting during training. On the other hand, denoising in standard CS algorithms is strongly based on a priori, and parameters are generally modified manually during the iterative process to offset the bias induced by noise, which is effective but does not provide convincing results when the noise level varies. Therefore, regularization parameters are added to the training model to improve the robustness of the model [23][24].



## Coding matrix design and dataset construction

In Fig. 3a, the filter function library produced from the FDTD simulation is shown, and in Fig. 3b, an initial screening yields a new filter function matrix. The initial screening is to calculate the

expectation and variance of each filter function to check that each filter function has an encoding effect, since if there is no light intensity change in the detection of the target, the spectrum cannot be decoded. Finally, the filter matrix is filtered again using the intercorrelation function [25] to yield the results shown in Fig. 3c.

The result of the generated spectral data is also displayed in Fig. 3, with varied amounts of Gaussian noise introduced to imitate genuine environmental conditions and improve the training model's durability. The following formula is applicable for converting simulated spectral data into undecoded input data:

$$I_i = \int_{\lambda_1}^{\lambda_2} T_i(\lambda) f(\lambda) + e_i \approx \sum_{j=1}^M T_i(\lambda_j) f(\lambda_j) + e_i, i = 1, \dots, N \quad (1)$$

The corresponding input undecoded data were obtained in this manner, where  $T_i(\lambda)$  is the encoding matrix,  $f(\lambda_j)$  the spectral data, and  $e_i$  is the added noise level, and the conversion was performed for each filter function, with the spectral data eventually all converted into the undecoded data, as shown in Fig. 2a. The 2W spectrum raw spectral data were generated with a Gaussian basis function, and then 5%, 10%, 15%, and 20% Gaussian noise were introduced to form a 10W trainable data set, which was divided 8:2 into training and validation sets, as shown in Figure 3e-f. The validation set is chosen at random from the 10W data to check that the distribution is appropriate, and the test set is real



1000 spectral data to evaluate the model's accuracy and robustness.

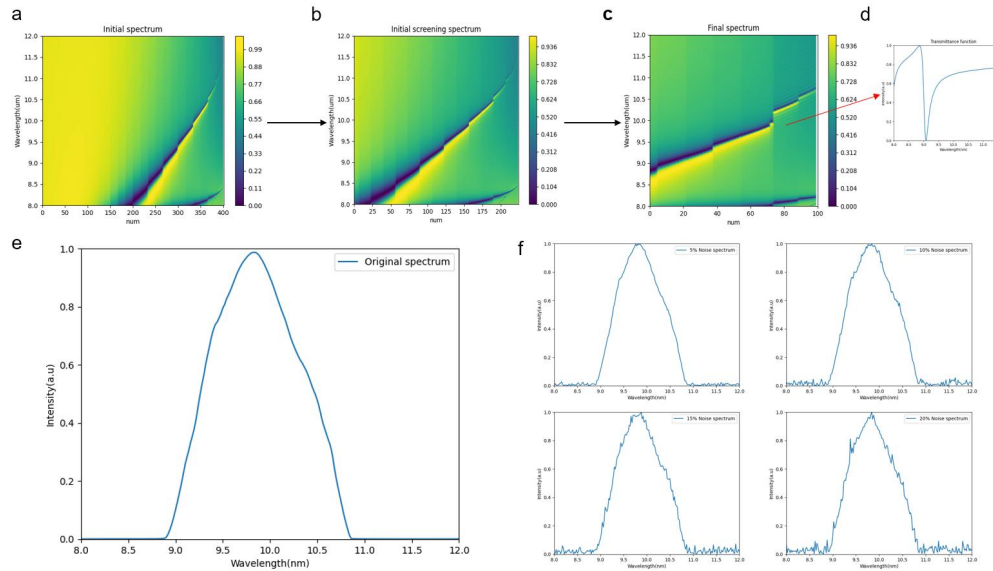
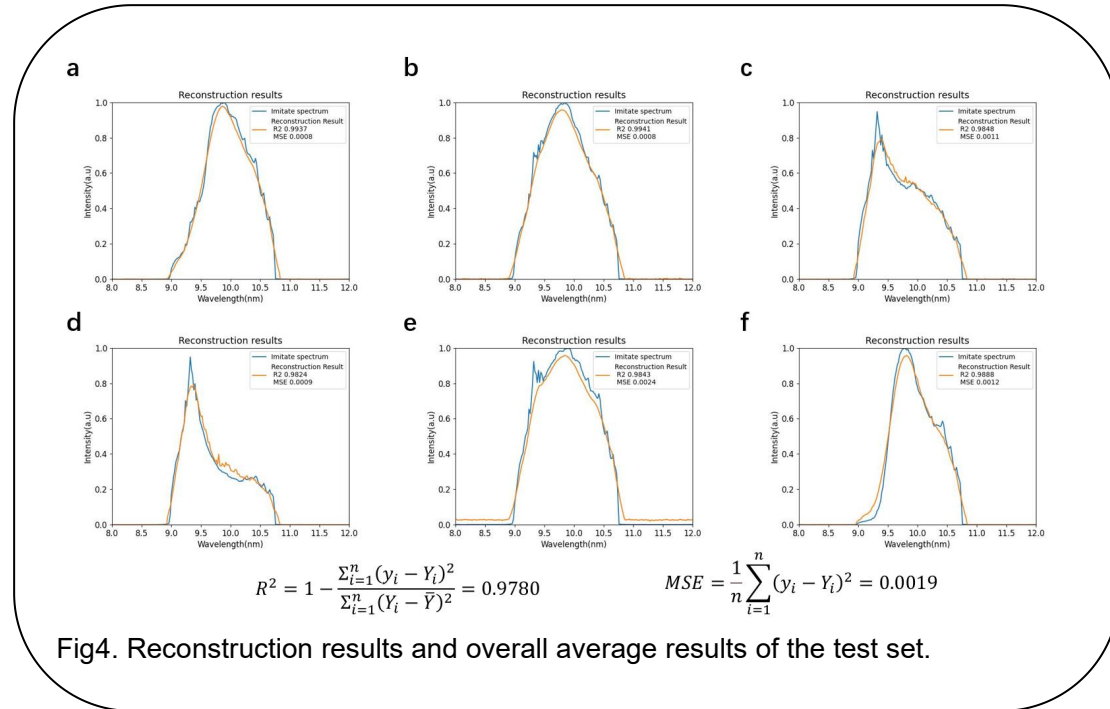


Fig3. coding matrix and simulated spectral data design. **a** is a library of filter functions obtained from FDTD simulation with 1 $\mu$ m height and 100 $\mu$ m-500 $\mu$ m length and width uniform silicon material; **b** is the result obtained after preliminary filtering of the library of filter functions; **c** is the result obtained by mutual correlation function, and **d** is one of the filter functions in **c**; **e** is a smooth spectral curve simulated by Gaussian basis function. **f** is a spectral curve with 5%, 10%, 15%, 20% Gaussian noise added respectively.

## Reconstruction results show

Figures 4a-f illustrate the results of the model reconstruction based on the attention algorithm, with the associated metric values annotated. The test set's overall R2 and MSE averages are also displayed. The results show that using the mutual correlation algorithm to filter the filter functions and combining the self-attentive mechanism network for spectral reconstruction yields

better reconstruction results, which is an important reference value for the practical application of spectral reconstruction.



## Discussion:

Computational spectrum reconstruction has several advantages, including low cost, compact size, ease of operation, and various applications. However, in practice, the accuracy of the spectral reconstruction is frequently reduced. In this study, we propose a new method to construct a filter function encoding matrix and perform spectral reconstruction to improve the accuracy of computational spectral reconstruction. While some previous studies used the method of randomly constructing the encoding matrix, the filtering method based on the mutual correlation

function can ensure that the constructed encoding matrix is sparse enough to better express the features and facilitate the subsequent spectral reconstruction data processing. Furthermore, a spectral reconstruction method based on a self-attentive mechanism is used, which can improve the correlation between features and fit the nonlinear data, thus improving the accuracy of the spectral reconstruction and preventing overfitting during training using an adding residuals mechanism. The experimental findings suggest that the approach has higher accuracy and better performance than other standard techniques, which can increase spectral reconstruction accuracy. Overall, the mutual correlation function and the self-attentive mechanism are useful methods in spectral reconstruction that can assist optimize the creation of the encoding matrix and feature extraction to improve the model's accuracy and performance. Our method extracts significant aspects of the spectral data and achieves more accurate spectral reconstruction by introducing the self-attentive mechanism and integrating it with a residual network. This discovery has significant implications for future spectrum reconstruction research and application.

Finally, while the experimental dataset in this study is limited to simulation tests, the feasibility and application prospects of the algorithm in a real-world context must be examined further. Future

studies might apply the algorithm to a broader range of real-world datasets and explore the method's performance under diverse scenarios to validate its resilience and usefulness. Furthermore, our study has several limitations, such as the size and variety of the dataset, which can be increased and optimized further. As a result, future research can further investigate the application of self-attentive mechanisms in spectral reconstruction, including how to optimize the network structure and select more effective feature extraction methods and try to use more datasets and methods to validate the algorithm's effectiveness and generalizability. The effectiveness of spectrum reconstruction approaches in practical applications can be increased further by ongoing optimization and refinement, providing better support and aid for research and applications in related domains.

## **Method:**

### **Filter function design and construction of data set**

This study's data construction is divided into two parts. To build an accurate spectrum reconstruction model, appropriate input and output data must be collected. The first part of the input data is a complete library of filter functions simulated by FDTD and set to 0.85 as the initial screening threshold condition, and then 100 filter

functions are finally filtered and arranged into a 10X10 coded array using the intercorrelation function. The second section simulates the spectral profile with a Gaussian function and adds varying levels of Gaussian noise, then encodes the matrix with the created filter function and converts it to unencoded input data using Equation 1. A total of 20,000 simulated spectral data were generated, with 5%, 10%, 15%, and 20% Gaussian noise added to create a total of 100,000 simulated data, which were then divided into a training set, and a validation set in an 8:2 ratio, with the validation set drawn at random from the overall data. Finally, 1000 spectral data points from realistic measurements are used as the test set to determine the model's accuracy.

### **The framework of the neural network model**

The neural network architecture for spectrum reconstruction can be expressed easily as "FC(100)->LR->Self-Attention(100)->FC(120)->LR->FC(150)->LR->FC(200)->LR". Each digit represents the number of cells in the relevant layer. The ReLU function is denoted by the symbol LR. fc means a fully connected layer, 100 random spectral filters correspond to the number of input cells, and 200 denotes the reconstructed spectral channels (8um-12um, 0.02um steps). and each layer has an added residual

method, as shown in Figure 2b.

### **Training and validation**

During training, the mean square error (MSE) is utilized as the loss function, and the batch normalization approach is employed to speed up the process. Adam is used as the optimization algorithm, and an acceptable learning rate is established. Furthermore, regularization constraints with L1 and dropout were used to minimize overfitting issues.

To ensure that the model has practical utility and can accurately recreate the results in the face of realistic environmental settings, we used actual spectrum data acquired by a marine optical spectrometer to validate its performance.

### **References:**

- [1] Z. Yang et al., Single-nanowire spectrometers. *Science* 365, 1017–1020 (2019). doi: 10.1126/science.aax8814; pmid: 31488686
- [2] L. P. Schuler, J. S. Milne, J. M. Dell, L. Faraone, MEMS-based microspectrometer technologies for NIR and MIR wavelengths. *J. Phys. D* 42, 133001 (2009). doi: 10.1088/0022-3727/42/13/133001

- [3] J. Malinen et al., Advances in miniature spectrometer and sensor development. Proc. SPIE 9101, 91010C (2014). doi: 10.1117/12.2053567
- [4] M.Ebermann et al., Tunable MEMS Fabry-Pérot filters for infrared microspectrometers: A review. Proc. SPIE 9760, 97600H (2016). doi: 10.1117/12.2209288
- [5] R. A. Crocombe, Portable Spectroscopy. Appl. Spectrosc. 72, 1701–1751 (2018). doi: .1177/0003702818809719; pmid: 30335465
- [6] R. F. Wolffenbuttel, MEMS-based optical mini- and microspectrometers for the visible and infrared spectral range. J. Micromech. Microeng. 15, S145–S152 (2005). doi: 10.1088/0960-1317/15/7/021.
- [7] Goodfellow, I., Bengio, Y. & Courville, A. Deep Learning 1 (MIT Press, 2016).
- [8] Jie Huang, Wengang Zhou, Qilin Zhang, Houqiang Li, and Weiping Li. Video-based sign language recognition without temporal segmentation. In Thirty-Second AAAI Conference on Artificial Intelligence, 2018.
- [9] Gang Liu and Jiabao Guo. Bidirectional lstm with attention mechanism and convolutional layer for text classification. Neurocomputing, 337:325–338, 2019.

- [10] Karim Ahmed, Nitish Shirish Keskar, and Richard Socher. Weighted transformer network for machine translation. arXiv preprint arXiv:1711.02132, 2017.
- [11] Matthias Sperber, Graham Neubig, Jan Niehues, and Alex Waibel. Attention-passing models for robust and data-efficient end-to-end speech translation. Transactions of the Association for Computational Linguistics, 7:313–325, 2019.
- [12] Kun Xu, Lingfei Wu, Zhiguo Wang, Yansong Feng, Michael Witbrock, and Vadim Sheinin. Graph2seq: Graph to sequence learning with attention-based neural networks. arXiv preprint arXiv:1804.00823, 2018.
- [13] Jianlong Fu, Heliang Zheng, and Tao Mei. Look closer to see better: Recurrent attention convolutional neural network for fine-grained image recognition. In 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 4476–4484, Honolulu, HI, July 2017. IEEE.
- [14] Kai Han, Jianyuan Guo, Chao Zhang, and Mingjian Zhu. Attribute-aware attention model for fine-grained representation learning. arXiv:1901.00392 [cs], January 2019. arXiv:1901.00392.
- [15] Dimitris Kastaniotis, Ioanna Ntinou, Dimitrios Tsourounis, George Economou, and Spiros Fotopoulos. Attention aware



generative adversarial networks (ata-gans). In 2018 IEEE 13th Image, Video, and Multidimensional Signal Processing Workshop (IVMSP), pages 1–5. IEEE, 2018.

- [16] Jiahui Yu, Zhe Lin, Jimei Yang, Xiaohui Shen, Xin Lu, and Thomas S. Huang. Generative image inpainting with contextual attention. arXiv:1801.07892 [cs], January 2018. arXiv: 1801.07892.
- [17] Baraniuk, R. G. Compressive sensing. IEEE Signal Process. Mag. 24, 118–121 (2007).
- [18] Yang J , Cui K , Cai X , et al. Ultraspectral Imaging Based on Metasurfaces with Freeform Shaped Meta-Atoms[J]. Laser & Photonics Reviews, 2022(7):16.
- [19] Song, H. Y. et al. Deep-learned broadband encoding stochastic filters for computational spectroscopic instruments. Adv. Theory Simul. 4, 2000299 (2021).
- [20] Nie, S. J. et al. Deeply learned filter response functions for hyperspectral reconstruction. In Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition 18–23 June 2018 (IEEE, Salt Lake City, UT, USA, 2018).
- [21] Leibe B , Matas J , Sebe N , et al. [Lecture Notes in Computer Science] Computer Vision – ECCV 2016 Volume 9908 || Identity Mappings in Deep Residual Networks[J]. 2016,

10.1007/978-3-319-46493-0(Chapter 38):630-645.

- [22] He K , Zhang X , Ren S , et al. Deep Residual Learning for Image Recognition[C]// IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2016.
- [23] Kukaka J , Golkov V , Cremers D . Regularization for Deep Learning: A Taxonomy[J]. 2017.
- [24] Wen F , Chu L , Liu P , et al. A Survey on Nonconvex Regularization Based Sparse and Low-Rank Recovery in Signal Processing, Statistics, and Machine Learning[J]. 2018.
- [25] Engle R . Dynamic Conditional Correlation[J]. Journal of Business & Economic Statistics, 2002, 20(3):339-350.