Transactions on Computer Science and Intelligent Systems Research
ISSN: 2960-1800, eISSN: 2960-2238 | Volume 1
AIEA 2023

Hyperspectral Image Study based on Deep Learning Model HSI-ConvNext

Zhonghui Bian, Xuying Zhao, Qian Wang, Zhibo Xing
Beijing Electronic Science and Technology Institute, Beijing 100000, China

Abstract. In recent years, deep learning techniques have made significant research progress in the field of hyperspectral image processing. Hyperspectral images are widely used in agriculture, environmental monitoring, geological exploration and other fields due to their richness in material and spectral information about objects. Deep learning models play an important role in hyperspectral image research and are able to learn feature representations from large-scale high-dimensional data, improving the performance of tasks such as classification, segmentation, and detection. In the study, researchers usually improve and optimize for the features of hyperspectral images by constructing deep learning models such as HSI-ConvNext. It is able to capture the correlation between different bands, thus realizing semantic segmentation and target detection for hyperspectral images. Data augmentation can effectively expand the limited hyperspectral dataset, and migration learning is able to migrate the knowledge of models trained in other domains to hyperspectral image tasks. Realistic hyperspectral images can be generated for data enhancement and model training. In conclusion, the study of hyperspectral images based on the deep learning model HSI-ConvNext provides new methods and techniques for the analysis and application of hyperspectral data. In the future, with the continuous development of deep learning technology, we can expect more innovations and breakthroughs in the field of hyperspectral image processing.

Keywords: HSI-ConvNext; Hyperspectral Image Processing; Hyperspectral Imaging; Hyperspectral Data Analysis; Spectral Characterization.

1. Introduction

Hyperspectral image classification, as one of the core technologies for hyperspectral remote sensing information extraction, has achieved important applications in the fields of precision agriculture, object detection, mineral exploration, urban planning and ecological science. However, the complexity of hyperspectral images also brings challenges. Hyperspectral images contain both two-dimensional geometric spatial information and one-dimensional spectral information, and direct classification may lead to the "Hughes" phenomenon and dimensionality catastrophe, and the classification results are far from the ideal state. In order to improve the classification accuracy, traditional methods usually extract the spatial and spectral information of hyperspectral images separately, and then perform spectral spatial classification, but the traditional methods ignore the spatial information in the original data, which limits the classification performance.

In recent years, deep learning models have developed rapidly in the field of image classification and become the most active research field in AI. Deep learning models have a multi-layered network structure that can better simulate human brain information processing and automatically learn features through training. In the field of hyperspectral images, deep learning models such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs) are widely used for image classification tasks. However, most of the current CNN-based hyperspectral image classification methods still do not fully integrate spectral and spatial information. To solve this problem, this paper proposes a deep learning model HSI-ConvNext for hyperspectral images, which improves on ConvNext by first introducing a spectral information processing module to better capture the spectral features of high-dimensional data. Secondly, a spectral downsampling layer is used to enhance the extraction of spatial and spectral information.The HSI-ConvNext model consists of an HSI-stem layer and multiple downsampling layers to reduce the spatial dimensions of the data and extract the initial features. After the downsampling process, the data is passed to the stage layer consisting of multiple convolutional groups. Each convolutional group uses depth-separable convolution to efficiently extract local features and capture features at different scales through multiple parallel convolutional kernels. Finally, the classification task is performed by feature normalization and linear layers to convert the extracted features into predictions.

The aim of this paper is to fully utilize the spatial and spectral information of hyperspectral images and propose the HSI-ConvNext model that combines deep learning features. Through experimental validation, the performance of the model in hyperspectral image classification task is explored to provide a more effective solution for hyperspectral remote sensing information extraction and application.

2. Related Work

Hyperspectral image classification is one of the core techniques for hyperspectral remote sensing information extraction, which is based on the spectral and spatial properties of image elements, and attribute determination and labeling of the feature classes contained in each image element or image element group. Hyperspectral images, as a kind of high-dimensional data consisting of images in hundreds of bands, contain a large amount of spectral and spatial information, which makes it possible to identify or detect materials at a fine-grained level, and thus have important applications in the fields of characterization such as precision agriculture, object detection, mineral exploration, urban planning and ecological science. Because hyperspectral image information has both 2-dimensional geometric spatial information and 1-dimensional spectral information, i.e., hyperspectral data has the form and structure of "image cube", and hyperspectral image has many bands, strong correlation between bands and high spatial resolution. These features of hyperspectral images make it difficult to apply the classification of pure spectral information or pure spatial information to all types of hyperspectral data, and if the hyperspectral images are directly classified, it is easy to produce the "Hughes" phenomenon and dimensionality catastrophe, which makes the classification results far from the ideal state. In order to improve the classification accuracy, it is a common practice to utilize both the spatial and spectral information of hyperspectral images, extract the spectral and spatial information contained in the dataset separately, and utilize different types of features for spectral spatial classification, which can effectively improve the classification accuracy. In recent years, researchers have proposed many joint spatial-spectral hyperspectral image classification methods.

In recent years, the ability of deep learning models in the field of image classification has grown rapidly and become the most active research area in AI. Different from traditional classification methods, deep learning networks often contain multiple layers of network structure, which well simulates the structure of the human brain to process information, and similar to the human learning mechanism, deep learning automatically learns from massive samples by means of training to obtain the features of these data. In recent years, some deep learning network models have been used in the field of hyperspectral images, such as convolutional neural network (CNN) and recurrent neural network (RNN)[1].

The above methods basically use a hierarchical structure to extract deep spectral features from the original spectral information to accomplish the classification of hyperspectral data, ignoring the spatial information contained in the original data[2]. To overcome this limitation, Convolutional Neural Networks (CNNs) have been widely introduced into the field of remote sensing due to their advantage of automatic local feature extraction, which is of great significance for HSI classification[3]. Typical examples include Siamese CNN, deep fusion residual networks, and multi-scale dense networks. However, most CNN-only based methods always extract features from the deepest convolutional or pooling layer, which results in a limited sensing field. But in this case, the model will focus more on local features and ignore the global information of the whole image[4]. This problem is greatly solved by the emergence of Transformer. The Transformer structure employs a self-attentive mechanism to better capture the global data correlation without relying on external information. In the field of computer vision, Transformer is further developed into visionTransform.

contrary to the trend of using Transformer to solve vision problems that has been popular in the past two years, ConvNeXt, recently proposed by FAIR, achieves ImageNet (dataset) Top-1 accuracy by relying on the convolutional structure alone.

2.1. ResNet

In traditional deep neural networks, as the number of network layers increases, the gradient decreases and may disappear, making the network difficult to train. To solve this problem, ResNet introduces the concept of residual connection. Residual connection adds input signals directly to the output by introducing jump connections in the network, allowing the network to learn the residual function instead of learning the original function directly. This design effectively mitigates the gradient vanishing problem and allows the network to be easily trained hundreds or even thousands of layers deep[5]. The basic unit of ResNet is the residual block, which consists of two or three convolutional layers. Each residual block has a jump connection that adds the input directly to the output. In addition, ResNet introduces commonly used network components such as batch normalization (batch normalization) and modified linear unit (ReLU) to enhance the expressive power and training stability of the network. Overall, ResNet is proposed based on an in-depth understanding of the gradient vanishing and degradation problem, and provides a solution through the introduction of residual connectivity, which allows deep neural networks to be better trained and optimized[6]. In hyperspectral image classification tasks, ResNet can be directly applied to hyperspectral data, where each pixel point has information from multiple spectral bands. By using a deep ResNet model, richer feature representations can be learned to improve the ability to capture complex features in hyperspectral data, thus improving classification performance[7].

2.2. ResNeXt

ResNeXt further extends the idea of ResNet, which focuses on residual connectivity and deep residual learning, while ResNeXt focuses on the design of parallel convolutional paths, which increases the width of the network by introducing parallel convolutional paths (cardinality)[8]. While traditional CNN models typically use a single convolutional path, ResNeXt learns features together by using multiple parallel convolutional paths that are somewhat similar to the idea of integrating multiple mini-models[9]. By combining features from different branches, it allows the network to better capture and utilize feature information at different scales. By taking full advantage of the feature learning capabilities of the parallel branches, it is able to better capture fine-grained features in the image, and is able to achieve the ability to better capture spatial and spectral information in hyperspectral data. By increasing the number of convolutional groupings, the expressive ability of the network can be improved and the learning ability of complex features can be enhanced, thus improving the performance of hyperspectral image classification.

2.3. ConvNeXt

ConvNeXt is further developed on the basis of ResNeXt, with the main focus on the design of convolutional kernels and the choice of scale. ConvNeXt captures feature information at different scales by connecting multiple convolutional kernels of different sizes in series, in order to improve the network's ability of recognizing objects at multiple scales[10]. This parallel convolutional kernel design gives ConvNeXt better flexibility and expressiveness. In addition to the convolutional kernel design, ConvNeXt also explores a variety of connectivity methods and network structures to further improve performance. For example, ConvNeXt can increase the depth of the network by stacking multiple parallel convolutional modules, or increase the width of the network by connecting multiple parallel convolutional modules in parallel. Hyperspectral images have a higher spectral resolution than ordinary RGB images and contain continuous spectral information from different wavelength bands. This makes hyperspectral images valuable for applications in the fields of feature classification, vegetation monitoring, and environmental monitoring[11]. The ConvNeXt network, however, can better capture the features from different bands through the design of parallel convolution kernel and multi-scale feature extraction, and shows excellent performance in the classification task.

In order to fully utilize the spatial and spectral information of hyperspectral remote sensing images, this paper proposes a convnext model (HSI-ConvNEXt) for hyperspectral by combining the features of hyperspectral data and deep learning.

2.4. HSI-ConvNext

The structure of ConvNeXt is shown in Figure 1:

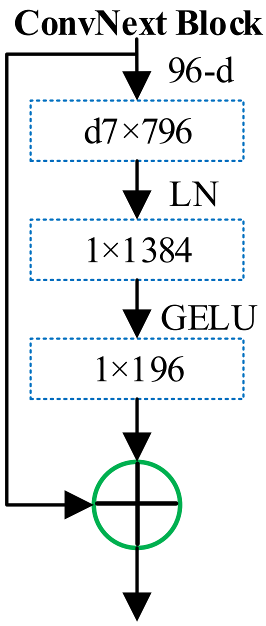

Figure 1. The structure of ConvNeXt.

The HSI-ConvNext model is obtained by improving ConvNext in many aspects. First, a processing module for hyperspectral information is added to the input layer to better capture the spectral features in the high-dimensional data. In the acquisition of spectral feature information, we introduced a spectral downsampling layer to improve the model's ability to extract spatial and spectral information from hyperspectral images[12].

The HSI-ConvNext model structure is specified as follows:

The following is the description of the HSI-ConvNext network model considering the dimensionality reduction processing of the STAGE layer after preprocessing:

The HSI-ConvNext network model first preprocesses the input data (both image and spectral information) through a large downsampling layer. This downsampling layer consists of an HSI-stem layer and three downsampling layers to reduce the spatial size of the data and extract initial features[13]. The HSI-stem layer consists of a 2x2x2 3D convolutional kernel for acquiring the spatial and spectral information, as well as a LayerNorm layer for feature normalization. The next three downsampling layers all consist of LayerNorm and a 2x2x2 3D convolution to further extract features and compress the spatial size of the data.

After processing by the downsampling layers, the data is passed to the stage layer consisting of four convolutional groups. At this stage, the data undergoes a dimensionality reduction process to reduce the computational complexity. These four convolutional groups consist of 3, 3, 9 and 3 Blocks respectively which are used to capture complex features and patterns at a higher level.

In each Block, Depth Separable Convolution (DwConv) is first applied to extract local features at a low computational cost. Then a Permute operation is performed on the resultant tensor to transform the Shape to (N, H, W, C)[14]. Next, LayerNorm is applied on the last axis (channels_last) for feature normalization, a linear transformation is performed to adjust the feature dimensions, the nonlinear expressiveness is increased by the GELU activation function, and another linear transformation is performed. Finally, the Shape Permute of the resultant tensor is returned to its original format for subsequent processing.

The information processed by the stage layer is passed to a LayerNorm layer for feature normalization. After LayerNorm processing, the data is passed to a linear layer for the classification task, where the extracted features are transformed into predictions.

3. Modeling

In this paper, we propose the HSI-ConvNext model, an improved network architecture for hyperspectral satellite image recognition. The HSI-ConvNext model is optimized with the following key components:

HSI-stem layer: use 3D convolutional kernel to obtain spatial and spectral information of input data and apply LayerNorm layer for feature normalization.

Downsampling layer: by connecting three downsampling layers with LayerNorm and 3D convolutional layers to reduce the data space size[15].

Stage layer: this stage contains 4 sets of convolutional operations, in each convolutional group, the data undergoes dimensionality reduction to reduce the computational complexity. The four convolutional groups include 3, 3, 9 and 3 Blocks respectively.

Block structure: each Block first applies depth separable convolution (DwConv) to extract local features. Then, a Permute operation is performed on the resultant tensor to transform the Shape to (N, H, W, C). Next, LayerNorm is applied on the last axis for feature normalization. A Linear transformation is performed to adjust the feature dimensions and increase the nonlinear expressiveness, and another Linear transformation is performed. Finally, the Shape of the resultant tensor is re-substituted to the initial format for subsequent processing[16].

LayerNorm layer: after the stage layer processing information is passed to a LayerNorm layer for feature normalization.

Linear classification layer: the data with the features processed by the LayerNorm layer is passed into a linear layer to complete the classification task, the extracted features are converted into predictions.

With these key steps and formulas, the HSI-ConvNext model achieves effective recognition of hyperspectral satellite images, reducing computational complexity while maintaining high accuracy.

4. Experimental Data

There are four sets of experimental data used in this paper, namely the Pavia University (PU), Pavia Center (PC), Indian Pines (IP), and Botswana datasets, with different datasets corresponding to different feature types and number of samples.

The Pavia University data band is 430-860 nm, with 103 available spectral channels and a spatial resolution of about 1.3 m. Each band contains 207,400 pixels.

The Indian Pines scene image has a size of 145 pixels × 145 pixels, a wavelength of 400-2,500 nm, a spectral resolution of 10 nm, and 224 original bands. The bad bands and the absorption bands of the water were removed, and the classification experiments were conducted with the remaining 200 bands.

The Pavia Center dataset is an image of downtown Pavia collected by the Airborne Reflectance Spectral Imager. The image size is 1096 pixels × 492 pixels with a spectral coverage of 0.43-0.86 μm and a spatial resolution of 1.3 m. The 102 spectral bands remaining for the study after removing the

13 noisy bands contain a total of 539,232 pixels, and only 7,456 pixels contain features, with a total of 9 ground objects[17].

The Botswana dataset is a series of data acquired by NASA's EO-1 satellite in the Okavango Delta of Botswana (Botswana) from 2001-2004. The Hyperion sensor on EO-1 acquired data at 30 km pixel resolution in the 242 km band over a 7.7 km strip, covering the 400-2500 nm portion of the 10 nm window in the 242 nm band. The UT Space Research Center preprocessed the data to mitigate the effects of bad detectors, inter-detector misalignment, and intermittent anomalies. Uncalibrated and noisy bands covering the absorption features were removed, and the remaining 145 bands were included as candidate features: [10-55, 82-97, 102-119, 134-164, 187-220]. The data analyzed in this study were obtained on May 31, 2001, and included observations from 14 identified categories representing seasonal marshes, ephemeral marshes, and dry woodland land cover types located at the end of the delta.

These four data sets were selected as experimental data primarily because of the different spatial resolution, different spectral resolution, and from different imaging sensors with different observational scenarios, allowing for a more comprehensive validation of the methodology. 4 Experimental Data

There are four sets of experimental data used in this paper, namely the Pavia University (PU), Pavia Center (PC), Indian Pines (IP), and Botswana datasets, with different datasets corresponding to different feature types and number of samples.

The Pavia University data band is 430-860 nm, with 103 available spectral channels and a spatial resolution of about 1.3 m. Each band contains 207,400 pixels.

The Indian Pines scene image has a size of 145 pixels × 145 pixels, a wavelength of 400-2,500 nm, a spectral resolution of 10 nm, and 224 original bands. The bad bands and the absorption bands of the water were removed, and the classification experiments were conducted with the remaining 200 bands.

The Pavia Center dataset is an image of downtown Pavia collected by the Airborne Reflectance Spectral Imager.... The image size is 1096 pixels × 492 pixels with a spectral coverage of 0.43-0.86 μm and a spatial resolution of 1.3 m. The 102 spectral bands remaining for the study after removing the 13 noisy bands contain a total of 539,232 pixels, and only 7,456 pixels contain features, with a total of 9 ground objects.

The Botswana dataset is a series of data acquired by NASA's EO-1 satellite in the Okavango Delta of Botswana (Botswana) from 2001-2004. The Hyperion sensor on EO-1 acquired data at 30 km pixel resolution in the 242 km band over a 7.7 km strip, covering the 400-2500 nm portion of the 10 nm window in the 242 nm band. The UT Space Research Center preprocessed the data to mitigate the effects of bad detectors, inter-detector misalignment, and intermittent anomalies. Uncalibrated and noisy bands covering the absorption features were removed, and the remaining 145 bands were included as candidate features: [10-55, 82-97, 102-119, 134-164, 187-220]. The data analyzed in this study were obtained on May 31, 2001, and included observations from 14 identified categories representing seasonal marshes, ephemeral marshes, and dry woodland land cover types located at the end of the delta.

These four data sets were selected as experimental data primarily because of the different spatial resolution, different spectral resolution, and from different imaging sensors with different observational scenarios, allowing for a more comprehensive validation of the methodology.

5. Experimental Results and Analysis

In order to evaluate the performance of the HSI-ConvNext model for hyperspectral satellite image recognition tasks, we can compare it with existing hyperspectral recognition methods. Here, an experimental scenario is designed in which HSI-ConvNext is compared with other typical hyperspectral recognition methods on several publicly available hyperspectral datasets, such as Indian Pines and Pavia University.

Table 1. Comparison based on hyperspectral dataset Pavia University.

PaviaU category	Sample size	Hsi-convnext	DL architecture	HSI-CNN
Asphalt	6631	0.960	0.927	0.732
Meadows	18649	0.917	0.907	0.828
Gravel	2099	0.932	0.843	0.000
Trees	3064	0.975	0.963	0.671
Painted metal sheets	1345	1.000	0.999	0.983
Bare Soil	5029	0.999	0.967	0.117
Bitumen	1330	0.986	0.885	0.000
Self-Blocking Bricks	3682	0.996	0.963	0.549
Shadows	947	0.999	0.996	0.982
total	42776			
Kappa		0.888	0.857	0.580
Average Accuracy		91.191%	88.797%	69.777%

Table 2. Comparison based on hyperspectral dataset Indian Pines.

IndianPines category	Sample size	Hsi-convnext	DL architecture	HSI-CNN
Alfalfa	46	1.000	0.686	0.000
Corn-notill	1428	0.980	0.044	0.000
Corn-mintill	830	0.819	0.415	0.000
Corn	237	0.871	0.482	0.000
Grass-pasture	483	0.845	0.760	0.000
Grass-trees	730	0.996	0.973	0.604
Grass-pasture-mowed	28	0.963	0.609	0.000
Hay-windrowed	478	0.923	0.948	0.694
Oats	20	1.000	0.571	0.000
Soybean-notill	972	0.924	0.604	0.255
Soybean-mintill	2455	0.956	0.740	0.000
Soybean-clean	593	0.898	0.605	0.000
Wheat	205	1.000	0.826	0.000
Woods	1265	0.971	0.334	0.831
Buildings-Grass-Trees-Drives	386	0.614	0.302	0.000
Stone-Steel-Towers	93	0.989	1.000	0.000
Total	10249			
Kappa		0.863	0.515	0.242
Average Accuracy		87.805%	56.644%	31.200%

The experimental platform is a high-performance computer with an Nvidia GeForce RTX 3060.

6. Conclusion

Research on hyperspectral imagery based on deep learning models has yielded significant results, and the following are some possible conclusions and summaries:

Effective Feature Learning and Representation: deep learning models perform well in hyperspectral image processing and are able to learn effective feature representations from large-scale, high-dimensional data. With architectures such as Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), and Transformer, deep learning models are able to capture features in both spatial and spectral dimensions, improving the performance of tasks such as classification, segmentation, and detection.

Semantic Segmentation and Target Detection: deep learning models have made breakthroughs in semantic segmentation and target detection in hyperspectral images. These models are able to accurately recognize different materials and targets in an image, and can achieve accurate localization and classification in complex backgrounds.

Data Enhancement and Generalization Capability: data enhancement techniques play an important role in hyperspectral image research by generating variant images to augment the limited training data. This helps to improve the generalization ability of the model so that it can perform well in different scenarios and conditions.

Cross-cutting domain applications: the success of deep learning models in hyperspectral image research is not just limited to the field of hyperspectral imaging. They have also been widely used in agriculture, environmental monitoring, geological exploration and other fields, providing new solutions to problems in related fields.

Model Optimization and Performance Improvement: Researchers are constantly working on improving the architecture and training techniques of deep learning models to further improve the performance of hyperspectral image processing tasks. From optimizing the loss function to introducing pre-trained models and migration learning, various methods have been explored and applied to obtain better results.

In conclusion, research on hyperspectral images based on deep learning models provides new perspectives and methods for the analysis and application of hyperspectral data. These studies not only provide strong support for scientific research, but also achieve important results in practical applications. With the continuous development of deep learning technology, we can expect to see more innovations and breakthroughs in the field of hyperspectral image processing.

Acknowledgments

This work was supported by “the Fundamental Research Funds for the Central Universities” (Grant Number: 328202234).

References

[1] Sherstinsky, A. (2020). Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network. Physical D: Nonlinear Phenomena, 404, 132306.

[2] Jiao, M., Wang, D., & Qiu, J. (2020). A GRU-RNN based momentum optimized algorithm for SOC estimation. Journal of Power Sources, 459, 228051.

[3] Kattenborn, T., Leitloff, J., Schiefer, F., & Hinz, S. (2021). Review on Convolutional Neural Networks (CNN) in vegetation remote sensing. ISPRS journal of photogrammetry and remote sensing, 173, 24-49.

[4] Bhatt, D., Patel, C., Talsania, H., Patel, J., Vaghela, R., Pandya, S., ... & Ghayvat, H. (2021). CNN variants for computer vision: History, architecture, application, challenges and future scope. Electronics, 10(20), 2470.

[5] Wen, L., Li, X., & Gao, L. (2020). A transfer convolutional neural network for fault diagnosis based on ResNet-50. Neural Computing and Applications, 32, 6111-6124.

[6] He, F., Liu, T., & Tao, D. (2020). Why resnet works? residuals generalize. IEEE transactions on neural networks and learning systems, 31(12), 5349-5362.

[7] Wu, Z., Shen, C., & Van Den Hengel, A. (2019). Wider or deeper: Revisiting the resnet model for visual recognition. Pattern Recognition, 90, 119-133.

[8] Zhou, T., Zhao, Y., & Wu, J. (2021, January). Resnext and res2net structures for speaker verification. In 2021 IEEE Spoken Language Technology Workshop (SLT) (pp. 301-307). IEEE.

[9] Pant, G., Yadav, D. P., & Gaur, A. (2020). ResNeXt convolution neural network topology-based deep learning model for identification and classification of Pediastrum. Algal research, 48, 101932.

[10] Zhu, Y., Yuan, K., Zhong, W., & Xu, L. (2023). Spatial-Spectral ConvNeXt for Hyperspectral Image Classification. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing.

[11] Liu, D., Lv, F., Wang, C., Wang, G., Zhang, H., & Guo, J. (2023). Classification of early mechanical damage over time in pears based on hyperspectral imaging and transfer learning. Journal of Food Science.

[12] Zhang, X., Yan, J., Tian, J., Li, W., Gu, X., & Tian, Q. (2023). Objective evaluation-based efficient learning framework for hyperspectral image classification. GIScience & Remote Sensing, 60(1), 2225273.

[13] Wu, H., Dai, S., Liu, C., Wang, A., & Iwahori, Y. (2023). A Novel Dual-Encoder Model for Hyperspectral and LiDAR Joint Classification via Contrastive Learning. Remote Sensing, 15(4), 924.

[14] Miao, J., Huang, Y., Wang, Z., Wu, Z., & Lv, J. (2023). Image recognition of traditional Chinese medicine based on deep learning. Frontiers in Bioengineering and Biotechnology, 11.

[15] Liu, Y., Yao, W., Qin, F., Zhou, L., & Zheng, Y. (2023). Spectral Classification of Large-Scale Blended (Micro) Plastics Using FT-IR Raw Spectra and Image-Based Machine Learning. Environmental Science & Technology, 57(16), 6656-6663.

[16] Feng, F., Zhang, Y., Zhang, J., & Liu, B. (2023). Low-Rank Constrained Attention-Enhanced Multiple Spatial–Spectral Feature Fusion for Small Sample Hyperspectral Image Classification. Remote Sensing, 15(2), 304.

[17] Hassanien, M. A., Singh, V. K., Puig, D., & Abdel-Nasser, M. (2022). Predicting breast tumor malignancy using deep ConvNeXt radiomics and quality-based score pooling in ultrasound sequences. Diagnostics, 12(5), 1053.