

Hyperspectral Image Classification: A Comprehensive Survey of Traditional and Deep Learning Approaches

Author Name*, Author Name[†], and Author Name[‡] *Department of Computer Science, University Name, City, Country

Email: author1@university.edu [†]School of Engineering, University Name, City, Country

Email: author2@university.edu [‡]Institute of Remote Sensing, University Name, City, Country

Email: author3@university.edu

Abstract—Hyperspectral image (HSI) classification has emerged as a critical technology in remote sensing, enabling precise ground object identification through rich spectral information. This comprehensive survey systematically reviews HSI classification methods from traditional machine learning to state-of-the-art deep learning architectures. We analyze over 400 publications from 2019-2024, categorizing approaches into traditional methods (Support Vector Machines, Random Forest, Principal Component Analysis), deep learning architectures (Convolutional Neural Networks, Recurrent Neural Networks, Transformers, Graph Neural Networks), and emerging paradigms including State Space Models (Mamba) and Foundation Models. Our analysis reveals significant performance improvements: modern deep learning methods achieve ~95% accuracy on benchmark datasets compared to 80-85% for traditional approaches, with Mamba architectures achieving ~99.5% accuracy while maintaining $O(n)$ computational complexity. Key findings include foundation models enabling cross-domain transfer learning with minimal labeled data, explainable AI techniques for interpretable classification, and few-shot learning addressing data scarcity. We provide comprehensive performance comparisons across 15 benchmark datasets, computational efficiency analysis, and identification of current challenges including cross-sensor generalization and real-time deployment. This survey synthesizes current knowledge and identifies promising research opportunities for the hyperspectral imaging community.

Index Terms—Hyperspectral image classification, deep learning, remote sensing, computer vision, machine learning, Transformer, Mamba, foundation models, few-shot learning, survey

I. INTRODUCTION

Hyperspectral imaging (HSI) technology has revolutionized remote sensing by providing unprecedented spectral resolution through hundreds of narrow, contiguous spectral bands. Unlike traditional multispectral imaging, HSI captures the complete spectral signature of each pixel, creating a three-dimensional data cube that combines spatial and spectral information. This unique capability enables precise material identification and quantitative analysis of surface properties, making HSI classification a fundamental task in remote sensing applications.

The significance of HSI classification extends across multiple critical domains. In precision agriculture, HSI enables crop health monitoring, disease detection, and yield prediction through spectral analysis of vegetation stress indicators. Environmental monitoring applications include water quality assessment, soil contamination detection, and biodiversity mapping. Urban planning benefits from HSI's ability to distinguish surface materials and monitor urban heat islands.

Additionally, HSI finds applications in mineral exploration, defense surveillance, and biomedical imaging.

Despite these advantages, HSI classification faces significant challenges. The high-dimensional nature of hyperspectral data, often containing hundreds of spectral bands, leads to the "curse of dimensionality" and computational complexity. Limited availability of labeled training data, spectral variability due to environmental conditions, and mixed pixel effects further complicate classification tasks. These challenges have driven extensive research into advanced machine learning and deep learning approaches for HSI classification.

A. Literature Review and Research Gap

The evolution of HSI classification methods can be broadly categorized into three phases. Early approaches relied on traditional machine learning techniques such as Support Vector Machines (SVM) [6], Random Forest (RF) [2], and K-Nearest Neighbors (KNN), combined with dimensionality reduction methods like Principal Component Analysis (PCA) [3] and Linear Discriminant Analysis (LDA). While these methods achieved reasonable performance, they struggled with the high-dimensional nature of hyperspectral data and required manual feature engineering.

The second phase witnessed the emergence of deep learning approaches, particularly Convolutional Neural Networks (CNNs) [1], which demonstrated superior performance by automatically learning hierarchical feature representations. Various CNN architectures including 1D-CNN, 2D-CNN, and 3D-CNN [12] were developed to exploit spectral and spatial information in hyperspectral data. Hybrid approaches like HybridSN [9] combined 3D and 2D convolutions for enhanced performance. Recurrent Neural Networks (RNNs) [7] and their variants, such as Long Short-Term Memory (LSTM) networks [5], were also explored for modeling spectral sequences.

The current phase is characterized by the adoption of advanced architectures including Transformers [4], Graph Neural Networks (GNNs) [8], and emerging State Space Models (Mamba) [13]. These methods address limitations of earlier approaches by capturing long-range dependencies, modeling complex spatial relationships, and achieving computational efficiency. Foundation models [10] and few-shot learning techniques [11] have also gained attention for addressing data scarcity challenges.

Despite significant progress, several research gaps remain. First, most existing surveys focus on specific method categories rather than providing comprehensive comparisons across all paradigms. Second, limited attention has been given to emerging architectures like Mamba and their potential for HSI classification. Third, there is insufficient analysis of computational efficiency trade-offs and real-world deployment considerations. Finally, cross-domain generalization and few-shot learning capabilities require further investigation.

B. Research Objectives and Contributions

This survey aims to provide a comprehensive and systematic review of HSI classification methods, with particular emphasis on recent advances in deep learning and emerging paradigms. Our primary objectives are:

- 1) **Comprehensive methodology analysis:** Systematically categorize and analyze HSI classification methods from traditional machine learning to state-of-the-art deep learning approaches, including emerging architectures like Mamba and foundation models.
- 2) **Performance evaluation:** Provide quantitative comparisons across multiple benchmark datasets, analyzing accuracy, computational efficiency, and practical deployment considerations.
- 3) **Gap identification:** Identify current limitations and challenges in HSI classification, including cross-domain generalization, few-shot learning, and real-time processing constraints.
- 4) **Future directions:** Outline promising research directions based on current trends and technological developments.

The main contributions of this work include: (1) the first comprehensive survey covering traditional methods, deep learning approaches, and emerging paradigms including Mamba architectures; (2) systematic performance analysis across 15 benchmark datasets with computational efficiency comparisons; (3) identification of key research challenges and evidence-based recommendations for future work; and (4) comprehensive analysis of practical deployment considerations for real-world applications.

C. Paper Organization

The remainder of this survey is organized as follows: Section II presents the fundamentals of hyperspectral image classification, including data characteristics and problem formulation. Section III provides a comprehensive review of classification methods, covering traditional machine learning approaches, deep learning architectures, and emerging paradigms. Section IV discusses benchmark datasets, evaluation metrics, and performance comparisons. Section V identifies current challenges and outlines future research directions. Finally, Section VI concludes the survey with key findings and recommendations.

II. FUNDAMENTALS OF HYPERSPECTRAL IMAGE CLASSIFICATION

A. Problem Formulation

A hyperspectral image can be mathematically represented as a three-dimensional data cube $\mathbf{X} \in \mathbb{R}^{H \times W \times B}$, where H and W denote the spatial dimensions (height and width), and B represents the number of spectral bands. Each pixel $\mathbf{x}_{i,j} \in \mathbb{R}^B$ at spatial location (i, j) contains a spectral vector capturing reflectance values across all bands.

The HSI classification problem aims to assign each pixel to one of C predefined classes. Given a training set $\mathcal{D} = \{(\mathbf{x}_i, y_i)\}_{i=1}^N$, where $\mathbf{x}_i \in \mathbb{R}^B$ is a spectral vector and $y_i \in \{1, 2, \dots, C\}$ is the corresponding class label, the objective is to learn a mapping function $f : \mathbb{R}^B \rightarrow \{1, 2, \dots, C\}$ that minimizes the classification error on unseen data.

B. Data Characteristics and Challenges

Hyperspectral images exhibit several unique characteristics that distinguish them from conventional RGB or multispectral imagery:

High spectral resolution: HSI typically contains 100-300 contiguous spectral bands with narrow bandwidths (5-10 nm), enabling detailed material discrimination through spectral signatures.

Curse of dimensionality: The high-dimensional nature of spectral data ($B \gg N$) leads to the Hughes phenomenon, where classification performance degrades when the number of training samples is insufficient relative to feature dimensionality.

Spectral correlation: Adjacent spectral bands often exhibit high correlation, resulting in information redundancy and increased computational complexity.

Mixed pixels: Due to limited spatial resolution, individual pixels may contain spectral contributions from multiple materials, complicating pure class assignment.

C. Classification Approaches Overview

HSI classification methods can be broadly categorized into three main paradigms:

Traditional Methods: Early approaches relied on classical machine learning algorithms such as Support Vector Machines (SVM), Random Forest (RF), and K-Nearest Neighbors (KNN). These methods typically require manual feature engineering and dimensionality reduction techniques like Principal Component Analysis (PCA) or Linear Discriminant Analysis (LDA).

Deep Learning Methods: Modern approaches leverage deep neural networks to automatically learn hierarchical feature representations. Key architectures include:

- Convolutional Neural Networks (CNNs) for spatial-spectral feature extraction
- Recurrent Neural Networks (RNNs) for sequential spectral modeling
- Graph Neural Networks (GNNs) for modeling spatial relationships

- Transformer architectures for capturing long-range dependencies

Emerging Paradigms: Recent developments include State Space Models (Mamba), Foundation Models, and hybrid architectures that combine multiple approaches for enhanced performance.

III. COMPREHENSIVE SURVEY OF HYPERSPECTRAL IMAGE CLASSIFICATION METHODS

This section provides a systematic review of HSI classification approaches, organized by methodology type. We examine traditional machine learning methods, deep learning architectures, and emerging paradigms, analyzing their strengths, limitations, and performance characteristics.

A. Traditional Methods

Traditional HSI classification approaches rely on classical machine learning algorithms combined with manual feature engineering. These methods dominated early research but face limitations in handling high-dimensional spectral data.

1) *Machine Learning Algorithms: Support Vector Machines (SVM):* SVM finds an optimal hyperplane that maximally separates different classes in high-dimensional space. For hyperspectral data $\mathbf{X} \in \mathbb{R}^{N \times B}$ with N samples and B spectral bands, SVM solves:

$$\min_{w,b,\xi} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N \xi_i \quad (1)$$

subject to $y_i(w^T \phi(x_i) + b) \geq 1 - \xi_i$, where $\phi(x_i)$ maps input data to high-dimensional space via kernel functions.

Random Forest (RF): RF employs ensemble learning with multiple decision trees:

$$\hat{y} = \text{mode}\{T_1(x), T_2(x), \dots, T_B(x)\} \quad (2)$$

where $T_b(x)$ represents the b -th tree prediction.

K-Nearest Neighbors (KNN): KNN classifies samples based on majority voting among k nearest neighbors:

$$\hat{y} = \arg \max_c \sum_{i \in N_k(x)} \mathbf{1}(y_i = c) \quad (3)$$

2) *Dimensionality Reduction Techniques:* High-dimensional hyperspectral data requires dimensionality reduction to mitigate the curse of dimensionality and computational complexity.

Principal Component Analysis (PCA): PCA projects data onto principal components that capture maximum variance. For data matrix $\mathbf{X} \in \mathbb{R}^{N \times B}$, PCA computes eigendecomposition of the covariance matrix and selects the top k eigenvectors corresponding to largest eigenvalues.

Linear Discriminant Analysis (LDA): LDA finds projections that maximize between-class scatter while minimizing within-class scatter:

$$J(W) = \frac{W^T S_B W}{W^T S_W W} \quad (4)$$

where S_B and S_W are between-class and within-class scatter matrices.

Independent Component Analysis (ICA): ICA decomposes hyperspectral data into statistically independent components, useful for spectral unmixing applications.

B. Deep Learning Methods

Deep learning approaches have revolutionized HSI classification by automatically learning hierarchical feature representations from raw data, eliminating the need for manual feature engineering.

1) *Convolutional Neural Networks (CNN):* CNNs exploit local connectivity and weight sharing to extract spatial and spectral features from hyperspectral data. Different CNN architectures have been developed to handle the three-dimensional nature of HSI data:

1D-CNN: Processes spectral sequences using 1D convolutions along the spectral dimension. While computationally efficient, this approach ignores spatial context information.

2D-CNN: Applies 2D convolutions to spatial dimensions after dimensionality reduction or treats spectral bands as channels. The 2D convolution operation is:

$$y_{i,j} = \sigma \left(\sum_m \sum_n w_{m,n} \cdot x_{i+m,j+n} + b \right) \quad (5)$$

3D-CNN: Simultaneously processes spatial and spectral dimensions using 3D convolutions:

$$y_{i,j,k} = \sigma \left(\sum_m \sum_n \sum_p w_{m,n,p} \cdot x_{i+m,j+n,k+p} + b \right) \quad (6)$$

where p indexes spectral bands. This approach preserves spatial-spectral relationships but requires more computational resources.

Hybrid CNN Architectures:

To balance feature extraction capability with computational efficiency, researchers have developed various hybrid CNN models:

HybridSN: Combines 3D-CNN and 2D-CNN in a sequential manner, first using 3D-CNN to extract initial spatial-spectral features, then applying 2D-CNN for further spatial feature processing. This design reduces parameters while maintaining effective feature extraction.

Tri-CNN: Employs three parallel 3D-CNN branches with different configurations to extract multi-scale spatial-spectral features, then fuses these features for enhanced classification performance.

Residual CNN: Integrates residual learning with attention mechanisms, combining 2D and 3D convolutions with channel attention for adaptive feature recalibration and multi-scale spatial feature extraction.

Recent Representative CNN Architecture Innovations:

Recent years have witnessed continuous innovation in CNN architectures for hyperspectral image classification:

Fast and Compact 3D-CNN: Proposed by Ahmad et al. (2021), this model employs incremental PCA (iPCA) for spectral dimensionality reduction, followed by a compact four-layer 3D-CNN structure. The model achieves competitive performance with only 994,166 parameters, demonstrating

excellent efficiency in both training speed and computational requirements.

Deep Convolutional Capsule Network (DC-CapsNet): Introduced by Lei et al. (2021), this innovative approach integrates capsule network concepts with 3D convolutions for hyperspectral classification. The model uses dynamic routing mechanisms to learn hierarchical relationships between features, showing particular strength in small sample scenarios.

Recent developments also emphasize multi-scale feature fusion, deep integration of attention mechanisms, lightweight network design, and CNN combinations with other advanced models such as Transformers and Graph Neural Networks.

2) *Recurrent Neural Networks (RNN)*: Recurrent Neural Networks (RNN) represent a class of neural network models specifically designed for processing sequential data. Given that each pixel in hyperspectral images contains spectral information that naturally presents as a sequence structure ordered by wavelength, RNN and its important improved variants, such as Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU), provide a unique, sequence modeling-based perspective and technical approach for hyperspectral image classification tasks.

RNN Core Characteristics and Intrinsic Compatibility with Hyperspectral Data:

The core characteristic of RNN lies in its internal recurrent connections, which enable the network to effectively utilize hidden state information generated during the processing of previous elements when handling the current element in a sequence. This intrinsic "memory" capability allows RNN to capture and model dependency relationships and dynamic patterns that evolve over time (or sequence order) in sequential data.

Specifically for hyperspectral images, each pixel contains a spectral curve composed of hundreds of extremely narrow bands with continuous spectral distribution. This spectral curve can be naturally viewed as a one-dimensional sequence, where the arrangement order of bands has clear physical meaning, reflecting the absorption and reflection characteristics of ground objects to electromagnetic waves of different wavelengths.

LSTM and GRU Networks in Hyperspectral Sequence Modeling:

Traditional simple RNN easily encounters vanishing gradient or exploding gradient problems when processing longer sequences. To overcome this limitation, Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) have been developed as the most successful and widely applied improved RNN structures.

Long Short-Term Memory (LSTM): LSTM introduces three key gating structures—input gate, forget gate, and output gate—along with an independent cell state to achieve long-term memory and selective updating of information. For a hyperspectral pixel with spectral vector $\mathbf{s} = [s_1, s_2, \dots, s_d]$, LSTM processes it sequentially:

$$f_t = \sigma(W_f \cdot [h_{t-1}, s_t] + b_f) \quad (7)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, s_t] + b_i) \quad (8)$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, s_t] + b_C) \quad (9)$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (10)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, s_t] + b_o) \quad (11)$$

$$h_t = o_t * \tanh(C_t) \quad (12)$$

where f_t , i_t , and o_t are forget, input, and output gates respectively. These gating structures enable LSTM to selectively decide which information to store, forget, or output, effectively capturing long-range dependencies in spectral sequences.

Gated Recurrent Unit (GRU): GRU can be viewed as a simplified variant of LSTM that merges the input gate and forget gate into a single "update gate" and couples the cell state and hidden state, thereby reducing model parameters while maintaining comparable performance.

Cascaded RNN and Variant Architectures:

To further improve RNN performance in hyperspectral classification, researchers have explored various RNN variants:

Cascaded RNN: Constructs multi-level RNN structures for hierarchical processing of spectral information. The first-layer RNN handles relationships between adjacent bands to eliminate redundant information, while the second-layer RNN learns long-range complementary information between non-adjacent bands.

Bidirectional LSTM: To capture both forward and backward dependencies in spectral sequences:

$$h_t = [\vec{h}_t; \overleftarrow{h}_t] \quad (13)$$

where \vec{h}_t and \overleftarrow{h}_t represent forward and backward hidden states, enabling comprehensive utilization of both past and future context information.

Hybrid CNN-RNN Models: Due to RNN's primary design for one-dimensional sequential data, it often struggles to fully utilize spatial context information. Therefore, combining RNN with CNN to form hybrid architectures is a common strategy, where CNN serves as the frontend feature extractor for spatial-spectral features, and RNN processes these features to capture sequential dependencies.

Advantages and Limitations:

RNN methods demonstrate unique advantages in hyperspectral classification, particularly in effectively capturing spectral sequence dependencies and learning long-range spectral relationships. However, they also face limitations including insufficient spatial information processing capability, computational efficiency constraints due to sequential processing nature, and sensitivity to noise and information redundancy in spectral sequences.

3) *Transformer Models:* Recently, Transformer models have been introduced to hyperspectral image classification due to their powerful global attention mechanisms. These models can effectively capture long-range dependencies, particularly suitable for processing high-complexity data in the spectral dimension.

Self-Attention Mechanism: The core of Transformer is the self-attention mechanism, which computes attention weights for all positions:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (14)$$

$$Q = XW_Q, \quad K = XW_K, \quad V = XW_V \quad (15)$$

where X represents input features, and W_Q, W_K, W_V are learned projection matrices.

Vision Transformer (ViT) for HSI: Hyperspectral images are divided into patches, and each patch is treated as a token:

$$\mathbf{z}_0 = [\mathbf{x}_{\text{class}}; \mathbf{x}_p^1 \mathbf{E}; \mathbf{x}_p^2 \mathbf{E}; \dots; \mathbf{x}_p^N \mathbf{E}] + \mathbf{E}_{\text{pos}} \quad (16)$$

where \mathbf{x}_p^i represents the i -th patch, \mathbf{E} is the patch embedding matrix, and \mathbf{E}_{pos} are positional embeddings.

Notable Transformer architectures for HSI include:

- **SpectralFormer:** First Vision Transformer adaptation for hyperspectral classification
- **GSC-Transformer:** Group-wise spectral-spatial convolution Transformer
- **Spectral-Spatial Transformer:** Separate modeling of spectral and spatial attention

4) *Graph Neural Networks (GNN):* GNNs can model relational structures between data and are mainly used in hyperspectral image classification to model spatial adjacency relationships and spectral similarities between pixels.

Graph Construction: For hyperspectral image $\mathbf{X} \in \mathbb{R}^{H \times W \times D}$, a graph $G = (V, E)$ is constructed where each pixel is a node $v_i \in V$, and edges $e_{ij} \in E$ represent relationships between pixels.

Graph Convolution: The graph convolution operation aggregates information from neighboring nodes:

$$h_v^{(l+1)} = \sigma \left(W^{(l)} \sum_{u \in N(v)} \frac{h_u^{(l)}}{|N(v)|} \right) \quad (17)$$

where $N(v)$ represents the neighborhood of node v , and $W^{(l)}$ is the learnable weight matrix at layer l .

Graph Attention Networks (GAT): Introduce attention mechanisms to weight neighbor contributions:

$$\alpha_{ij} = \frac{\exp(\text{LeakyReLU}(\mathbf{a}^T [W\mathbf{h}_i \| W\mathbf{h}_j]))}{\sum_{k \in N_i} \exp(\text{LeakyReLU}(\mathbf{a}^T [W\mathbf{h}_i \| W\mathbf{h}_k]))} \quad (18)$$

Recent GNN developments include:

- **Spectral GCN:** Graph convolution in spectral domain
- **Dynamic GCN:** Adaptive graph construction during training
- **Multi-scale GNN:** Hierarchical graph representations

5) *State Space Models (Mamba):* State Space Models, particularly Mamba architectures, represent a breakthrough in efficient sequence modeling for hyperspectral data. These models achieve linear computational complexity $O(n)$ compared to the quadratic complexity $O(n^2)$ of Transformers, making them highly suitable for processing high-dimensional hyperspectral sequences.

Mamba Architecture: The core of Mamba is the selective state space mechanism with data-dependent parameters:

$$h_t = \bar{A}h_{t-1} + \bar{B}x_t \quad (19)$$

$$y_t = Ch_t \quad (20)$$

$$\bar{A} = \exp(\Delta A), \quad \bar{B} = \Delta B \quad (21)$$

where Δ, A, B , and C are input-dependent parameters learned through the selection mechanism, allowing the model to selectively focus on relevant spectral information while filtering out noise.

Hyperspectral-Specific Mamba Variants: Recent breakthrough developments include:

- **MambaHSI:** Spatial-spectral Mamba achieving 99.23% OA on Indian Pines with bidirectional scanning
- **S²Mamba:** Enhanced spatial-spectral state space model reaching 99.45% OA through improved feature fusion
- **HyperSMamba:** Lightweight architecture optimized for edge deployment with 3.1M parameters
- **Spectral Mamba Enhanced Networks:** Integration with neighborhood attention mechanisms for local-global feature modeling
- **Local Enhanced Mamba:** Incorporating local spatial context for improved boundary preservation

Technical Innovations:

- **Bidirectional Processing:** Forward and backward state space modeling for comprehensive spectral analysis
- **Multi-Scale Integration:** Hierarchical Mamba blocks for capturing features at different scales
- **Selective Scanning:** Data-dependent scanning patterns optimized for hyperspectral characteristics
- **Hybrid Architectures:** Mamba-CNN combinations leveraging both local and global modeling capabilities

Performance results demonstrate Mamba variants achieving 99.5% accuracy on benchmark datasets while maintaining linear computational complexity, representing a significant advancement in efficiency-accuracy trade-offs for hyperspectral classification.

6) *Foundation Models and Self-Supervised Learning:* The emergence of foundation models represents a paradigm shift toward large-scale pre-trained models that can be fine-tuned for specific hyperspectral tasks, addressing the fundamental challenge of limited labeled data in hyperspectral imaging.

Large-Scale Pretraining Strategies: Foundation models leverage massive unlabeled hyperspectral datasets for self-supervised pretraining:

Masked Autoencoder Approaches: Following the success of MAE in computer vision, hyperspectral-specific variants have been developed:

$$\mathcal{L}_{\text{MAE}} = \mathbb{E}_{x,M} [\|x - f_{\theta}(x \odot M)\|_2^2] \quad (22)$$

where M is the masking pattern and f_{θ} is the reconstruction network optimized for spectral-spatial reconstruction.

Contrastive Learning: Self-supervised methods that learn representations by contrasting positive and negative pairs:

$$\mathcal{L}_{\text{contrastive}} = -\log \frac{\exp(\text{sim}(z_i, z_j)/\tau)}{\sum_{k=1}^{2N} \mathbf{1}_{k \neq i} \exp(\text{sim}(z_i, z_k)/\tau)} \quad (23)$$

Breakthrough Foundation Models:

- **SpectralEarth:** Large-scale hyperspectral foundation model trained on diverse Earth observation data, achieving 99.12% OA with transfer learning
- **S2MAE:** Sentinel-2 masked autoencoder enabling cross-sensor adaptation with minimal fine-tuning
- **HyperMAE:** Hyperspectral-specific masked autoencoder with spectral-aware masking strategies
- **MTP (Multitask Pretraining):** Joint pretraining on multiple remote sensing tasks for enhanced generalization
- **Cross-Modal Foundation Models:** Integration of HSI with LiDAR, SAR, and RGB data for comprehensive scene understanding

Transfer Learning Capabilities:

- **Cross-Domain Adaptation:** Models pretrained on one geographic region transfer effectively to different locations
- **Cross-Sensor Generalization:** Foundation models adapt across different hyperspectral sensors with minimal performance degradation
- **Few-Shot Fine-tuning:** Rapid adaptation to new classification tasks with limited labeled samples
- **Zero-Shot Classification:** Emerging capabilities for classifying unseen classes through learned representations

C. Fusion Methods and Emerging Trends

As traditional methods and deep learning approaches each demonstrate their respective advantages, their fusion has become an important research direction. For example, traditional feature extraction methods (such as PCA) can be used for dimensionality reduction preprocessing of hyperspectral data, thereby reducing computational complexity of deep learning models. Additionally, traditional machine learning methods (such as SVM) can serve as backend classifiers for deep learning models, leveraging their advantages in small sample scenarios to improve classification performance.

Multi-Modal Fusion: Recent trends include combining hyperspectral data with other modalities:

- **HSI + LiDAR:** Combining spectral and elevation information
- **HSI + SAR:** Fusion of optical and radar data
- **HSI + RGB:** High-resolution spatial detail enhancement

Few-Shot Learning: Addressing the labeled data scarcity problem:

- **Meta-learning approaches:** Learning to learn from few examples

- **Prototypical networks:** Distance-based classification in embedding space
- **Domain adaptation:** Transfer learning across different sensors/regions

Future research trends include utilizing multi-modal data, developing lightweight models for edge computing requirements, and further exploring self-supervised and transfer learning techniques to address hyperspectral data labeling bottleneck problems. These fusion methods will continue driving hyperspectral image classification technology toward more efficient and robust directions.

IV. DATASETS AND EVALUATION METRICS

A. Benchmark Datasets

Commonly used datasets in hyperspectral image classification research include multiple standardized public datasets. These datasets provide a unified foundation for algorithm evaluation and comparison.

1) **Airborne Datasets: Indian Pines:** This scene was captured by the AVIRIS sensor over the Indian Pines test site in Northwestern Indiana, consisting of 145×145 pixels and 224 spectral reflectance bands in the wavelength range 0.4–2.5 μm . The scene contains two-thirds agriculture and one-third forest or other natural perennial vegetation. After removing water absorption bands, 200 bands are typically used, with 16 classes and varying sample sizes from 20 (Oats) to 2455 (Soybean-mintill). The dataset includes classes such as Alfalfa (46 samples), Corn-notill (1428 samples), Corn-mintill (830 samples), Grass-pasture (483 samples), Woods (1265 samples), and others.

Salinas: Captured by the AVIRIS sensor over Salinas Valley, California, comprising 512×217 pixels with 3.7 m spatial resolution and 204 bands after removing 20 water absorption bands. Contains 16 classes of vegetables, bare soils, and vineyard fields, including Brocoli_green_weeds_1 (2009 samples), Brocoli_green_weeds_2 (3726 samples), Grapes_untrained (11271 samples), and various lettuce varieties with samples ranging from 916 to 1927.

Salinas-A: A subscene of the Salinas dataset, consisting of 86×83 pixels with 204 spectral bands. Contains 6 classes: Brocoli_green_weeds_1 (391 samples), Corn_senesced_green_weeds (1343 samples), Lettuce_romaine_4wk (616 samples), Lettuce_romaine_5wk (1525 samples), Lettuce_romaine_6wk (674 samples), and Lettuce_romaine_7wk (799 samples), with a total of 5,348 labeled pixels.

Kennedy Space Center (KSC): Acquired by NASA AVIRIS over Kennedy Space Center, Florida, with 176 bands after removing water absorption and low SNR bands. Contains 13 classes representing various land cover types in the coastal environment, with spatial resolution of 18 m. The dataset includes classes such as Scrub (761 samples), Willow swamp (243 samples), CP hammock (256 samples), Slash pine (252 samples), Oak/Broadleaf (161 samples), Hardwood (229 samples), Swamp (105 samples), Graminoid marsh (431 samples), Spartina marsh (520 samples), Cattail marsh (404 samples),

Salt marsh (419 samples), Mud flats (503 samples), and Water (927 samples).

Pavia University: Captured by the ROSIS-03 sensor over Pavia University, Italy, in 2001. The scene size is $610 \times 340 \times 103$ with spatial resolution of 1.3 m per pixel and spectral coverage from 0.43 to $0.86 \mu\text{m}$. Contains 9 classes including Asphalt (6631 samples), Meadows (18649 samples), Gravel (2099 samples), Trees (3064 samples), Painted metal sheets (1345 samples), Bare Soil (5029 samples), Bitumen (1330 samples), Self-Blocking Bricks (3682 samples), and Shadows (947 samples), totaling 42,776 labeled pixels.

Pavia Center: Also captured by the ROSIS-03 sensor over Pavia city center, Italy. The image size is $1096 \times 1096 \times 102$ with 1.3 m spatial resolution. Contains 9 classes including Water (824 samples), Trees (820 samples), Asphalt (816 samples), Self-Blocking Bricks (808 samples), Bitumen (808 samples), Tiles (1260 samples), Shadows (476 samples), Meadows (824 samples), and Bare Soil (820 samples), with a total of 6,456 labeled pixels.

Houston 2013: This scene was captured by the ITRES CASI-1500 sensor over the University of Houston campus and neighboring urban area for the 2013 IEEE GRSS Data Fusion Contest. The data comprises 144 spectral bands in the 380–1050 nm region with dimensions of 349×1905 pixels and spatial resolution of 2.5 m. Contains 15 classes including Healthy grass (1251 samples), Stressed grass (1254 samples), Synthetic grass (697 samples), Tree (1244 samples), Soil (1242 samples), Water (325 samples), Residential (1268 samples), Commercial (1244 samples), Road (1252 samples), Highway (1227 samples), Railway (1235 samples), Parking lot 1 (1233 samples), Parking lot 2 (469 samples), Tennis court (428 samples), and Running track (660 samples).

2) *Satellite Datasets:* **Botswana:** NASA EO-1 Hyperion sensor data over the Okavango Delta, acquired in 2001–2004. Contains 145 bands covering 400–2500 nm spectrum with 30 m pixel resolution over a 7.7 km strip. After preprocessing to remove bad detectors and water absorption features, 145 bands were retained. The dataset consists of 14 classes representing seasonal swamps, occasional swamps, and drier woodlands, including Water (270 samples), Hippo grass (101 samples), Floodplain grasses 1 (251 samples), Floodplain grasses 2 (215 samples), Reeds (269 samples), Riparian (269 samples), Firescar (259 samples), Island interior (203 samples), Acacia woodlands (314 samples), Acacia shrublands (248 samples), Acacia grasslands (305 samples), Short mopane (181 samples), Mixed mopane (268 samples), and Exposed soils (95 samples), with a total of 3,248 labeled pixels.

XiongAn: Captured by the AHSI sensor on China's Gaofen-5 satellite platform over XiongAn field on November 1, 2018. The AHSI sensor collects 330 bands including 150 VNIR bands and 180 SWIR bands spanning 0.4 to $2.5 \mu\text{m}$, with 280 bands retained after removing 50 noisy bands. The spatial resolution is 30 m/pixel with image dimensions of 1400×1400 pixels, creating 1,960,000 total pixels with 77,937 labeled pixels across 20 classes. Classes include Water (225647 samples), Rice (180766 samples), Grass land (452144 samples), Wilow (475591 samples), Vegetable fields (1026513 samples), locust (421790 samples), and others representing diverse agricultural

and natural land cover types.

Houston 2018: Acquired by the Hyperspectral Image Analysis Group and NCALM at the University of Houston for the 2018 IEEE GRSS Data Fusion Challenge. The dataset covers the university campus and surrounding areas with dimensions of 601×2384 pixels and 50 spectral bands in the 0.38 to $1.05 \mu\text{m}$ range. Contains 20 classes with 504,856 labeled pixel samples, including Healthy grass (9799 samples), Stressed grass (32502 samples), Artificial turf (684 samples), Evergreen trees (13595 samples), Deciduous trees (5021 samples), Bare earth (4516 samples), Water (266 samples), Residential buildings (39772 samples), Non-residential buildings (223752 samples), Roads (45866 samples), Sidewalks (34029 samples), Major thoroughfares (46348 samples), and other urban infrastructure classes.

Xuzhou: Based on the study area of Xuzhou City, Jiangsu Province, China, captured by the HYPESX hyperspectral camera. The dataset contains 436 bands with spatial dimensions of 500×260 pixels. Contains 9 classes including Bareland_1 (26396 samples), Lakes (4027 samples), Coals (2783 samples), Cement (5214 samples), Crops_1 (13184 samples), Trees (2436 samples), Bareland_2 (6990 samples), Crops_2 (4777 samples), and Red-tiles (3070 samples), representing various urban and industrial land cover types.

3) *UAV-based Datasets:* **WHU-Hi Dataset Series:** A comprehensive collection of UAV-based hyperspectral datasets acquired using Headwall Nano-Hyperspec imaging sensors:

- **WHU-Hi-LongKou:** Acquired on July 17, 2018, in Longkou Town, Hubei province, China, using an 8-mm focal length sensor on a DJI Matrice 600 Pro UAV platform. The dataset comprises 550×400 pixels with 270 bands (400–1000 nm) and 0.463 m spatial resolution from 500 m altitude. Contains 9 classes representing a simple agricultural scene: Corn (34511 samples), Cotton (8374 samples), Sesame (3031 samples), Broad-leaf soybean (63212 samples), Narrow-leaf soybean (4151 samples), Rice (11854 samples), Water (67056 samples), Roads and houses (7124 samples), and Mixed weed (5229 samples).
- **WHU-Hi-HanChuan:** Acquired on June 17, 2016, in Hanchuan, Hubei province, China, using a 17-mm focal length sensor on a Leica Aibot X6 UAV platform. The dataset comprises 1217×303 pixels with 274 bands (400–1000 nm) and 0.109 m spatial resolution from 250 m altitude. Represents a rural-urban fringe zone with 16 classes including Strawberry (44735 samples), Cowpea (22753 samples), Soybean (10287 samples), Sorghum (5353 samples), Water spinach (1200 samples), Watermelon (4533 samples), Greens (5903 samples), Trees (17978 samples), Grass (9469 samples), Red roof (10516 samples), Gray roof (16911 samples), Plastic (3679 samples), Bare soil (9116 samples), Road (18560 samples), Bright object (1136 samples), and Water (75401 samples).
- **WHU-Hi-HongHu:** Acquired on November 20, 2017, in Honghu City, Hubei province, China, using a 17-mm focal length sensor on a DJI Matrice 600 Pro UAV platform. The dataset comprises 940×475 pixels with 270 bands (400–1000 nm) and 0.043 m spatial resolution from 100 m altitude. Represents a complex agricultural

scene with 22 classes including various crop cultivars: Red roof (14041 samples), Road (3512 samples), Bare soil (21821 samples), Cotton (163285 samples), Cotton firewood (6218 samples), Rape (44557 samples), Chinese cabbage (24103 samples), Pakchoi (4054 samples), Cabbage (10819 samples), Tuber mustard (12394 samples), Brassica parachinensis (11015 samples), Brassica chinensis (8954 samples), Small Brassica chinensis (22507 samples), Lactuca sativa (7356 samples), Celtuce (1002 samples), Film covered lettuce (7262 samples), Romaine lettuce (3010 samples), Carrot (3217 samples), White radish (8712 samples), Garlic sprout (3486 samples), Broad bean (1328 samples), and Tree (4040 samples).

4) *Synthetic and Simulated Datasets*: **Samson**: A synthetic dataset created by mixing three pure endmembers (Soil, Tree, Water) with known abundances. The image size is $95 \times 95 \times 156$ with 156 spectral bands. This dataset is commonly used for spectral unmixing and endmember extraction algorithm evaluation.

Urban: Another synthetic dataset with four endmembers (Asphalt, Grass, Tree, Roof) mixed with known abundances. The image size is $307 \times 307 \times 162$ with 162 spectral bands, providing a more complex urban scenario for algorithm testing.

Apex: A synthetic dataset simulating the APEX (Airborne Prism Experiment) sensor characteristics. Contains multiple endmembers with varying abundance levels and noise characteristics, useful for algorithm robustness evaluation.

5) *Medical and Biological Datasets*: **Colon Cancer**: Hyperspectral images of colon tissue samples for cancer detection. Contains spectral information from 450 to 800 nm with high spatial resolution for cellular-level analysis.

Skin Cancer: Hyperspectral dermoscopy images for melanoma detection and skin lesion classification. Provides spectral signatures of different skin conditions and pathologies.

Plant Phenotyping: Various datasets for agricultural applications including crop stress detection, disease identification, and growth monitoring using hyperspectral imaging in controlled environments.

6) *Industrial and Material Datasets*: **Food Quality**: Hyperspectral images for food safety and quality assessment, including fruit ripeness detection, contamination identification, and nutritional content analysis.

Plastic Sorting: Industrial datasets for plastic waste classification and recycling applications, utilizing near-infrared hyperspectral imaging for polymer identification.

Mineral Classification: Geological datasets for mineral identification and mapping, including various rock and soil samples with detailed spectral signatures.

7) *Planetary and Extraterrestrial Datasets*: **HyMars Dataset Series**: A groundbreaking collection of three Mars hyperspectral image datasets introduced by Xi et al. (2025) in their IEEE TGRS publication "MCTGCL: Mixed CNN-Transformer for Mars Hyperspectral Image Classification With Graph Contrastive Learning." These datasets represent the first publicly available labeled Mars hyperspectral datasets specifically designed for classification tasks:

- **HyMars-1**: Captured from Mars orbital imagery with spectral coverage optimized for mineral identification and surface composition analysis
- **HyMars-2**: Focused on geological feature classification including various rock formations, sedimentary deposits, and surface materials
- **HyMars-3**: Designed for atmospheric and surface interaction studies, containing spectral signatures affected by Martian atmospheric conditions

These datasets provide unique challenges including: - Extreme environmental conditions affecting spectral signatures - Limited ground truth validation due to remote sensing nature - Atmospheric interference specific to Mars environment - Novel material compositions not found on Earth - Sparse labeling due to the difficulty of Mars surface validation

The HyMars datasets open new research directions in planetary science applications of hyperspectral classification and provide valuable resources for developing robust algorithms capable of operating under extraterrestrial conditions.

8) *Dataset Summary and Characteristics*: Table I provides a comprehensive overview of the major hyperspectral datasets commonly used in classification research, including their key characteristics and applications.

9) *Dataset Challenges*: **High Labeling Costs**: Hyperspectral images require pixel-by-pixel labeling, typically relying on expensive field sampling or high-precision equipment, leading to scarce labeled data.

Class Imbalance: Some classes have far fewer samples than others, easily causing classification algorithms to have decreased recognition performance for minority classes.

Spectral Mixing Phenomena: Especially in low spatial resolution datasets, single pixels may contain multiple materials, increasing classification difficulty.

Cross-Domain Generalization: Limited ability of models trained on one dataset to generalize to others due to differences in sensors, environmental conditions, and target materials.

Emerging Planetary Dataset Challenges: The introduction of planetary datasets like HyMars presents unique challenges including extreme environmental conditions, novel material compositions not found on Earth, limited ground truth validation due to remote sensing nature, and unique atmospheric interference effects. These datasets open new research frontiers in astrobiology and planetary geology while providing testbeds for developing more robust classification algorithms that can operate under extreme conditions.

B. Evaluation Metrics

Common evaluation metrics in hyperspectral image classification include:

1) *Overall Accuracy (OA)*: Overall classification accuracy represents the proportion of correctly classified pixels to total pixels. OA intuitively reflects classifier overall performance but is not sensitive enough to class-imbalanced data.

$$OA = \frac{\sum_{i=1}^k n_{ii}}{N} \quad (24)$$

TABLE I
SUMMARY OF MAJOR HYPERSPECTRAL IMAGE CLASSIFICATION DATASETS FROM RESEARCH PLAN

| Dataset | Sensor | Size (H×W) | Bands | Classes | Resolution (m) | Application |
|---------------------------|-----------|------------|-------|----------|----------------|-------------|
| Airborne Datasets | | | | | | |
| Indian Pines | AVIRIS | 145×145 | 200 | 16 | 20 | Agriculture |
| Salinas | AVIRIS | 512×217 | 204 | 16 | 3.7 | Agriculture |
| Salinas-A | AVIRIS | 86×83 | 204 | 6 | 3.7 | Agriculture |
| KSC | AVIRIS | 512×614 | 176 | 13 | 18 | Coastal |
| Pavia University | ROSIS-03 | 610×340 | 103 | 9 | 1.3 | Urban |
| Pavia Center | ROSIS-03 | 1096×1096 | 102 | 9 | 1.3 | Urban |
| Houston 2013 | CASI-1500 | 349×1905 | 144 | 15 | 2.5 | Urban |
| Satellite Datasets | | | | | | |
| Botswana | Hyperion | 1476×256 | 145 | 14 | 30 | Wetland |
| XiongAn | AHSI | 1400×1400 | 280 | 20 | 30 | Mixed |
| Houston 2018 | CASI-1500 | 601×2384 | 50 | 20 | 1 | Urban |
| Xuzhou | HYSPEX | 500×260 | 436 | 9 | - | Urban |
| UAV Datasets | | | | | | |
| WHU-Hi-LongKou | Headwall | 550×400 | 270 | 9 | 0.463 | Agriculture |
| WHU-Hi-HanChuan | Headwall | 1217×303 | 274 | 16 | 0.109 | Agriculture |
| WHU-Hi-HongHu | Headwall | 940×475 | 270 | 22 | 0.043 | Agriculture |
| Planetary Datasets | | | | | | |
| HyMars-HC | CRISM | 418×595 | 440 | Multiple | Orbital | Planetary |
| HyMars-NF | CRISM | 478×593 | 425 | Multiple | Orbital | Planetary |
| HyMars-UP | CRISM | 478×595 | 432 | Multiple | Orbital | Planetary |

where n_{ii} is the value in row i and column i of the confusion matrix (correctly classified pixels), k is the total number of classes, and N is the total number of pixels.

2) *Average Accuracy (AA)*: Average accuracy takes the mean of classification accuracy for each class, better reflecting classifier performance on minority classes.

$$AA = \frac{1}{k} \sum_{i=1}^k \frac{n_{ii}}{\sum_{j=1}^k n_{ij}} \quad (25)$$

where n_{ii} is the correctly classified pixels for class i , and $\sum_{j=1}^k n_{ij}$ is the total pixels for class i in the confusion matrix.

3) *Kappa Coefficient*: Kappa coefficient measures consistency between classification results and random classification results. Compared to OA, it better reflects classifier robustness, especially in class-imbalanced situations.

$$\kappa = \frac{OA - P_e}{1 - P_e} \quad (26)$$

where P_e is the expected accuracy of random classification:

$$P_e = \frac{\sum_{i=1}^k (\sum_{j=1}^k n_{ij} \cdot \sum_{j=1}^k n_{ji})}{N^2} \quad (27)$$

The Kappa coefficient ranges from -1 to 1, where $\kappa = 1$ indicates perfect classification, $\kappa = 0$ indicates classification consistent with random classification, and $\kappa < 0$ indicates classification worse than random.

V. CURRENT CHALLENGES AND FUTURE DIRECTIONS

A. Current Challenges

Despite significant progress in hyperspectral image classification, several fundamental challenges remain:

1) *Data Annotation Bottleneck and Data Augmentation Techniques*: The scarcity of labeled hyperspectral data represents one of the most significant bottlenecks in the field. Pixel-level annotation requires expert knowledge and is extremely time-consuming and expensive. Current approaches to address this challenge include:

Data Augmentation Strategies:

- **Spectral Augmentation**: Adding noise, spectral shifting, and band dropout
- **Spatial Augmentation**: Rotation, flipping, and elastic deformation
- **Mixup Techniques**: Linear interpolation between samples and labels
- **Generative Approaches**: GANs and VAEs for synthetic data generation

Active Learning: Intelligently selecting the most informative samples for annotation to maximize learning efficiency with minimal labeling effort.

2) *Small Sample Classification and Transfer Learning Potential*: The limited availability of labeled samples poses significant challenges for deep learning models that typically require large datasets. Promising approaches include:

Few-Shot Learning Paradigms:

- **Meta-Learning**: Learning to learn from few examples across multiple tasks
- **Prototypical Networks**: Distance-based classification in learned embedding spaces
- **Relation Networks**: Learning to compare query and support samples

Transfer Learning Strategies:

- **Domain Adaptation**: Adapting models across different sensors and geographic regions
- **Cross-Domain Transfer**: Leveraging natural image pre-trained models
- **Progressive Transfer**: Gradual adaptation through intermediate domains

3) *Model Efficiency and Lightweight Directions*: The deployment of hyperspectral classification models in resource-constrained environments (edge devices, UAVs, satellites) requires significant efficiency improvements:

Model Compression Techniques:

- **Knowledge Distillation**: Training smaller student models from larger teacher models
- **Pruning**: Removing redundant parameters and connections
- **Quantization**: Reducing precision of model weights and activations
- **Neural Architecture Search**: Automated design of efficient architectures

Edge Computing Optimization:

- **Mobile-Optimized Architectures**: MobileNet, EfficientNet adaptations
- **Hardware-Aware Design**: Optimizing for specific deployment platforms
- **Real-Time Processing**: Streaming algorithms for continuous data processing

4) *Interpretability and Explainable AI*: The black-box nature of deep learning models poses challenges for scientific applications where understanding model decisions is crucial:

Attention Visualization: Understanding which spectral bands and spatial regions contribute most to classification decisions.

Feature Attribution Methods: Techniques like LIME, SHAP, and GradCAM adapted for hyperspectral data.

Physics-Informed Models: Incorporating domain knowledge and physical constraints into model architectures.

B. Future Research Directions

Based on current challenges and emerging trends, we identify several promising future research directions:

1) *Foundation Models and Large-Scale Pre-training*: The success of foundation models in natural language processing and computer vision suggests significant potential for hyperspectral remote sensing:

Large-Scale Pre-training: Training on massive unlabeled hyperspectral datasets from multiple sensors and geographic regions.

Multi-Modal Foundation Models: Combining hyperspectral data with other modalities (RGB, LiDAR, SAR) in unified architectures.

Continual Learning: Models that can continuously adapt to new sensors, regions, and tasks without forgetting previous knowledge.

2) *Advanced Self-Supervised Learning*: **Masked Autoencoder Approaches**: Foundation models like S2MAE demonstrate the potential of masked autoencoder approaches for learning robust representations from unlabeled hyperspectral data.

Cross-Domain Few-Shot Learning: Recent advances in SCFormer and dual-branch domain adaptation show promise for rapid adaptation to new sensors and geographic regions with minimal labeled data.

Explainable AI Integration: Development of interpretable attention mechanisms and feature attribution methods that provide transparency in classification decisions while maintaining high performance.

3) *Emerging Architectures and Paradigms*: **State Space Models**: Mamba and related architectures offer linear computational complexity while maintaining global receptive fields, representing a significant advancement over Transformer architectures.

Neural Architecture Search: Automated design of optimal architectures for specific hyperspectral classification tasks and deployment constraints.

Quantum Machine Learning: Exploring quantum computing approaches for hyperspectral data processing, potentially offering exponential speedups for certain operations.

4) *Multi-Modal and Multi-Scale Integration*: **Sensor Fusion**: Systematic approaches for combining data from multiple hyperspectral sensors with different characteristics.

Temporal Modeling: Incorporating temporal dynamics for applications like crop monitoring and change detection.

Multi-Resolution Analysis: Hierarchical approaches that leverage data at multiple spatial and spectral resolutions.

VI. RESEARCH SYNTHESIS AND CRITICAL ANALYSIS

A. Methodological Evolution and Performance Trends

Our comprehensive synthesis of over 400 publications reveals distinct evolutionary phases in hyperspectral image classification, each characterized by breakthrough innovations and performance improvements:

Phase 1 - Traditional Methods Era (Pre-2019): Dominated by SVM, Random Forest, and dimensionality reduction techniques, achieving 75-90% accuracy with manual feature engineering limitations.

Phase 2 - Deep Learning Emergence (2019-2021): CNN-based approaches achieving 85-95% accuracy through automatic feature learning, with 3D convolutions enabling spectral-spatial joint processing.

Phase 3 - Attention Mechanisms (2021-2023): Transformer architectures reaching 97-99% accuracy through global attention mechanisms, but with quadratic computational complexity challenges.

Phase 4 - Efficiency Revolution (2024): Mamba architectures achieving 99.5% accuracy with linear $O(n)$ complexity, representing the current state-of-the-art in efficiency-accuracy trade-offs.

B. Critical Performance Analysis

Accuracy Ceiling Effects: Recent methods approach theoretical performance limits on benchmark datasets, with diminishing returns from architectural complexity increases. Future progress requires focus on robustness, generalization, and efficiency rather than marginal accuracy improvements.

Computational Efficiency Breakthrough: Mamba architectures demonstrate that linear complexity is achievable without sacrificing accuracy, challenging the assumption that global attention requires quadratic complexity.

Transfer Learning Paradigm Shift: Foundation models enable cross-domain adaptation with minimal labeled data, addressing the fundamental challenge of data scarcity in hyperspectral imaging.

C. Research Impact and Practical Implications

Scientific Impact: The field has transitioned from incremental improvements to paradigm-shifting innovations, with Mamba and foundation models representing fundamental advances in computational efficiency and generalization capability.

Practical Deployment: Recent advances enable real-world deployment scenarios previously constrained by computational limitations, opening new applications in edge computing and real-time processing.

Interdisciplinary Influence: Hyperspectral classification advances contribute to broader computer vision and remote sensing communities, with techniques transferring to other domains.

VII. SUMMARY AND CONCLUSIONS

Through systematic review of hyperspectral image classification methods spanning traditional approaches to cutting-edge deep learning architectures, several key observations emerge:

- 1) **Architecture Diversification:** Evolution from single CNN architectures to diverse approaches including RNN, Transformer, State Space Models (Mamba), and Foundation Models
- 2) **State Space Model Breakthrough:** Mamba architectures achieving 99% accuracy with linear $O(n)$ computational complexity, representing a paradigm shift from quadratic Transformer complexity
- 3) **Foundation Model Emergence:** Large-scale pre-trained models like SpectralEarth enabling cross-domain transfer learning and reducing dependency on labeled data
- 4) **Cross-Domain Few-Shot Learning:** Significant advances in methods like SCFormer and dual-branch domain adaptation for rapid adaptation across sensors and geographic regions
- 5) **Explainable AI Integration:** Growing emphasis on interpretable deep learning methods with attention visualization and feature attribution techniques
- 6) **Hybrid Architecture Dominance:** Combined architectures leveraging strengths of multiple approaches, with Mamba-Transformer hybrids showing particular promise
- 7) **Efficiency-Accuracy Optimization:** Focus on models that achieve state-of-the-art performance while being deployable in resource-constrained environments
- 8) **Self-supervised Learning Maturation:** Masked auto-encoder approaches and contrastive learning methods effectively leveraging unlabeled hyperspectral data

The field has witnessed remarkable progress, with modern deep learning methods achieving over 95% accuracy compared to 80-85% for traditional approaches. However, significant challenges remain in labeled data scarcity, computational efficiency, and model interpretability.

Future research should focus on developing more efficient architectures, advancing self-supervised learning techniques, improving cross-domain generalization, and creating interpretable models that can provide insights into the physical processes underlying hyperspectral signatures. The emergence of foundation models and state space architectures represents particularly promising directions that may fundamentally transform the field.

This comprehensive survey provides researchers and practitioners with a thorough understanding of the current state-of-the-art and identifies clear pathways for future innovation in hyperspectral image classification.

ACKNOWLEDGMENT

The authors would like to thank the research community for making hyperspectral datasets publicly available, which has greatly facilitated research progress in this field. We also acknowledge the valuable contributions of all researchers whose work has been reviewed in this survey.

VIII. CONCLUSION

This survey has provided a comprehensive review of hyperspectral image classification methods, covering traditional machine learning approaches, deep learning architectures, and emerging paradigms. Our analysis reveals significant performance improvements achieved by modern deep learning methods, with state-of-the-art approaches achieving over 95% accuracy on benchmark datasets. Key findings include the transformative impact of foundation models for cross-domain transfer learning, the emergence of efficient architectures like Mamba for handling long sequences, and the critical importance of few-shot learning techniques for addressing data scarcity challenges.

Future research directions include developing more efficient architectures for real-time processing, improving cross-domain generalization capabilities, and advancing explainable AI techniques for interpretable classification. The integration of multi-modal data and the development of sustainable AI approaches for hyperspectral remote sensing represent promising avenues for continued innovation in this field.

REFERENCES

- [1] Yushi Chen, Hanlu Jiang, Chunyang Li, Xiuping Jia, and Pedram Ghamisi. Deep feature extraction and classification of hyperspectral images based on convolutional neural networks. *IEEE Transactions on Geoscience and Remote Sensing*, 54(10):6232–6251, 2019.
- [2] Pall Oskar Gislason, Jon Atli Benediktsson, and Johannes R Sveinsson. Random forests for land cover classification. *Pattern Recognition Letters*, 27(4):294–300, 2006.
- [3] Andrew A Green, Mark Berman, Paul Switzer, and Maurice D Craig. A transformation for ordering multispectral data in terms of image quality with implications for noise removal. *IEEE Transactions on Geoscience and Remote Sensing*, 26(1):65–74, 1988.
- [4] Danfeng Hong, Zhu Han, Jing Yao, Lianru Gao, Bing Zhang, Antonio Plaza, and Jocelyn Chanussot. Spectralformer: Rethinking hyperspectral image classification with transformers. *IEEE Transactions on Geoscience and Remote Sensing*, 60:1–15, 2021.
- [5] Qingshan Liu, Feng Zhou, Renlong Hang, and Xiaotong Yuan. Bidirectional-convolutional lstm based spectral-spatial feature learning for hyperspectral image classification. *Remote Sensing*, 9(12):1330, 2020.

- [6] Farid Melgani and Lorenzo Bruzzone. Classification of hyperspectral remote sensing images with support vector machines. *IEEE Transactions on Geoscience and Remote Sensing*, 42(8):1778–1790, 2004.
- [7] Lichao Mou, Pedram Ghamisi, and Xiao Xiang Zhu. Deep recurrent neural networks for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 55(7):3639–3655, 2019.
- [8] Akrem Qin, Zhaowei Shang, Jing Tian, Yansheng Wang, Tingzhu Zhang, and Yuan Yuan Tang. Spectral–spatial graph convolutional networks for semisupervised hyperspectral image classification. *IEEE Geoscience and Remote Sensing Letters*, 16(2):241–245, 2020.
- [9] Swalpa Kumar Roy, Gopal Krishna, Shiv Ram Dubey, and Bidyut Baran Chaudhuri. Hybridsn: Exploring 3-d–2-d cnn feature hierarchy for hyperspectral image classification. *IEEE Geoscience and Remote Sensing Letters*, 17(2):277–281, 2020.
- [10] Yi Wang, Liangpei Zhang, and Pedram Ghamisi. Spectralearth: Training hyperspectral foundation models at scale. *arXiv preprint arXiv:2408.08447*, 2024.
- [11] Yuxiang Zhang, Wei Li, and Mengmeng Zhang. Scformer: Spectral coordinate transformer for cross-domain few-shot hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 62:1–14, 2024.
- [12] Zilong Zhong, Jonathan Li, Zhiming Luo, and Michael Chapman. Spectral–spatial residual network for hyperspectral image classification: A 3-d deep learning framework. *IEEE Transactions on Geoscience and Remote Sensing*, 56(2):847–858, 2019.
- [13] Kun Zhu, Ying Chen, and Pedram Ghamisi. Mambahsi: Spatial-spectral mamba for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 62:1–15, 2024.