Chapter 1

Introduction

This chapter has provided an introduction to digital imaging, gradually transitioning into the more advanced field of hyperspectral imaging (HS) and its broad range of applications. It first explored the different types of digital images, emphasizing the unique ability of hyperspectral image (HSI) to capture detailed spectral information across various wavelengths. Building on this foundation, the chapter then delved into HSI classification, outlining a generalized process that leverages machine learning (ML) and deep learning (DL) techniques. In addition, it also identified the key challenges inherent in HSI classification, offering a clear problem formulation and well-defined research objectives. Furthermore, the chapter highlighted the contributions of this thesis aimed at addressing these challenges and concluded by mapping out the structure of the thesis, setting the stage for the subsequent chapters. To ensure thoroughness, the chapter also included an overview of publicly available HSI datasets and evaluation metrics, providing essential details for further research and analysis.

1.1 Types of Digital Images

Digital images are representative of the visual information in matrix format, in which each element corresponds to a intensity of light called as pixel value. These images can be categorized based on the type of information stored and how it is represented. The preliminary image formats are binary, grayscale, and color images. This subsection delves into the primary types of digital images, which emphasize their structural composition and characteristics. Figure 1.1 illustrates the difference between binary, grayscale, RGB-color, multispectral image (MSI), and HSI.

1. **Binary Image:** Binary images are represented in two values, zeros and ones, for black and white, respectively. They are highly efficient in storing each pixel as a

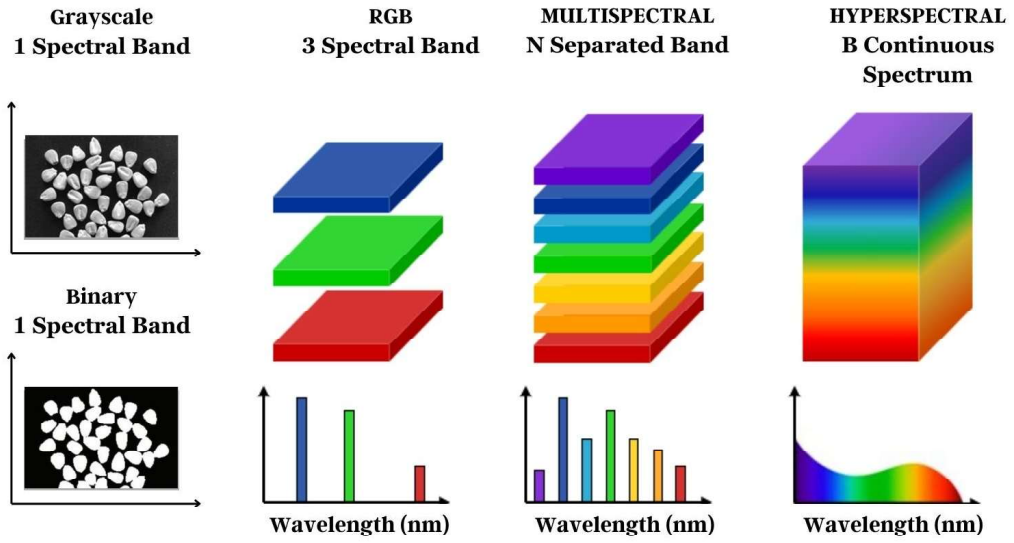

Figure 1.1: Visualization of different image types, including binary, grayscale, RGB color, multispectral, and HSI.

single bit. The minimal storage makes them ideal for applications like document scanning, optical character recognition, and barcode detection [1, 2]. However, their major limitation is the inability to depict detailed visual information or gradients, which restricts their use in complex imaging tasks.

2. **Grayscale Image:** Grayscale Image encodes intensity levels from 0 to 255 in a matrix format. It typically uses 8-bit storage, though higher bit depths like 16-bit or 32-bit offer greater precision. While requiring less storage than color images, they lack color information, limiting applications needing color analysis. Grayscale image is widely used in medical diagnostics, image processing, and computer vision [3, 4], where structural and intensity details are crucial.
3. **Color Image:** Color image represents visual information using multiple matrices, referred to as channels. These images are categorized into formats such as RGB (Red, Green, Blue), CMYK (Cyan, Magenta, Yellow, Key), and HSV (Hue, Saturation, Value) based on their channel composition. RGB employs three 8-bit channels per pixel, while other formats utilize different color models. Color image is widely applied in photography, video production, and multimedia [5, 6]. While offering realistic representation, it demands higher storage and computational resources than binary and grayscale images.
4. **Multispectral Image (MSI) :** MSI captures data across 3 to 10 discrete spectral bands, which include visible (RGB), near-infrared, and short-wave infrared.

Each pixel encodes reflectance or emission values at specific wavelengths that form a multidimensional array. MSI collects data across a limited number of discrete spectral bands characterized by relatively large differences (Δ) between adjacent wavelengths. This capability enables precise surface mapping by integrating both spatial and radiometric resolutions, leading to applications in remote sensing, environmental monitoring, geological exploration, and medical imaging [7,8]. However, despite providing valuable spectral insights, MSI requires significant storage, incurs high costs, and demands advanced algorithms for processing.

5. **Hyperspectral Images (HSI) :** HSI captures spectral data across hundreds of contiguous bands that extend beyond the visible spectrum to infrared, ultraviolet, and thermal regions. Each pixel encodes reflectance or emission values at specific wavelengths, forming a high-dimensional spectral cube. Unlike multispectral imaging, HSI records data in narrow, continuous intervals, enabling precise material characterization. It integrates spatial and radiometric resolutions, supporting applications in remote sensing, agriculture, and medical imaging [9, 10]. The "hyper" in HSI signifies its high spectral resolution, often exceeding 100 bands ($H \times W \times B, B > 100$). This fine spectral granularity enhances classification accuracy, anomaly detection, and vegetation analysis. However, its large data volume and computational complexity require advanced processing techniques. Dimensionality reduction methods improve efficiency while preserving spectral details. Despite challenges, HSI remains vital for extracting rich spectral information.

1.2 Hyperspectral Images (HSI): Background and Motivations

HS is an advanced technique to capture and analyze data on a wide range of electromagnetic wavelengths, which extend beyond the visible spectrum. It integrates multiple spectroscopic methods, including Raman spectroscopy, ultraviolet, visible light, NIR, SWIR, mid-wave infrared, long-wave infrared, and thermal infrared, that enable detailed spectral analysis. Unlike traditional imaging systems that use RGB or multispectral imaging that captures data across a few spectral bands, HSI captures hundreds of continuous spectral bands. This allows for a more detailed analysis of materials and processes. High-resolution images are produced using specialized hyperspectral sensors and cameras, which provide exceptional spectral accuracy.

The nature of the data captured by HS sensors varies based on several key factors:

1. **Spatial Resolution:** Determines the level of detail in the spatial domain, typically measured in terms of pixels per unit area. High spatial resolution enables

the precise identification of fine spatial details.

2. **Spectral Resolution:** Refers to the ability of the sensor to distinguish between adjacent wavelengths. Higher spectral resolution results in narrower wavelength intervals, offering greater detail for spectral analysis.
3. **Radiometric Resolution:** Relates to the sensitivity of the sensor in detecting differences in intensity levels, allowing for the differentiation of subtle variations in reflectance or emission.
4. **Temporal Resolution:** Indicates the frequency at which data is captured over time, which is crucial for applications requiring real-time or periodic monitoring.

HS sensors are utilized in various configurations, including push broom scanners, snapshot cameras, and whisk broom scanners, each tailored to specific applications. These sensors produce a three-dimensional data cube represented as $[H \times W \times B]$ or $[B \times H \times W]$, where $H \times W$ denotes the spatial dimensions and B indicates the number of spectral bands. This setup allows for the integration of both spatial and spectral information, facilitating tasks such as material identification, environmental monitoring, and medical diagnostics. However, the high dimensionality of HSI poses significant challenges in terms of storage, processing, and analysis. As a result, advanced algorithms for dimensionality reduction, feature extraction, and classification are required to utilize their vast potential for real-world applications effectively.

1.3 Hyperspectral Image Datasets

Publicly available HSI datasets are essential for developing and benchmarking classification models, providing standardized data for evaluating algorithm performance [11]. Captured using advanced sensors, these datasets offer high spectral and spatial resolution, enabling precise material identification and feature extraction.

Table 1.1 presents widely used HSI datasets, including Indian Pines (IP), University of Pavia (PU), Houston (HU), Salinas Valley (SA), Membrane Nephropathy (MN-HSI), Yellow RIver Estuary Coastal Wetland (YRE), WHU-Hi-Hanchuan (HC), and WHU-Hi-LongKou (LK) among others. These datasets vary in spatial dimensions, spectral bands, wavelength range, and class diversity, making them suitable for applications such as land cover classification, environmental monitoring, and precision agriculture. Their availability fosters reproducible research and facilitates the development of more accurate and efficient hyperspectral image analysis techniques.

Table 1.1: The table provides details on publicly available HSI datasets for classification methods. It includes various sensor-based HSI datasets with information such as mode, wavelength, spatial size, number of bands (NB), and Number of classes (NC).

Dataset	Sensor	Mode	Wavelength	Spatial Size	NB	NC
Indiana Pines (IP)	AVIRIS	Aerial	$0.4 - 2.5\mu m$	145×145	224	16
University of Pavia (PU)	ROSIS	Aerial	$0.43 - 0.85\mu m$	610×340	103	9
Pavia Center (PC)	ROSIS	Aerial	$0.43 - 0.85\mu m$	1096×715	102	9
Salinas Valley (SA)	AVIRIS	Aerial	$0.4 - 2.5\mu m$	512×217	227	16
Kennedy Space Center (KSC)	AVIRIS	Aerial	$0.4 - 2.5\mu m$	512×614	176	13
Botswana (BS)	HYPERION	Satellite	$0.4 - 2.5\mu m$	145×145	145	14
University of Houston (HU)	NCALM	Aerial	$0.38 - 1.05\mu m$	349×1905	144	15
AISA Eagle (AEG)	AISA	Aerial	$0.39 - 0.99\mu m$	–	272	9
PRISMA-SA-BAM-S1	PRISMA	Satellite	$0.4 - 2.5\mu m$	1168×1168	240	2
PRISMA-SA-BAM-S2	PRISMA	Satellite	$0.4 - 2.5\mu m$	1168×1168	240	2
Berlin (BR)	–	Satellite	$0.4 - 2.5\mu m$	797×220	244	8
Augsburg (AU)	HySpex	Satellite	$0.4 - 2.5\mu m$	332×485	180	7
C.indium	PikaXC2	Lab	$0.4 - 1.005\mu m$	265×106	75	3
WBC-3D	–	Lab	$0.45 - 0.78\mu m$	1280×1024	51	5
MN-HSI	–	Lab	$0.4 - 0.1\mu m$	696×520	128	2
FX-10	SPECIM	Lab	$0.4 - 0.1\mu m$	72×72	224	7
FX-17	SPECIM	Lab	$0.49 - 1.7\mu m$	72×72	224	7
YRE	Gaofen-5	Satellite	–	740×761	296	8
HyRank-Loukia(HR-L)	HYPERION	Satellite	$0.4 - 2.5\mu m$	249×945	176	14
WHU-Hi-HanChuan (HC)	Nano-Hyperspec	Aerial	$0.4 - 1\mu m$	1217×303	270	16
WHU-Hi-LongKou (LK)	Nano-Hyperspec	Aerial	$0.4 - 1\mu m$	550×400	270	9
DFC-2018	NCALM	Aerial	$0.38 - 1.05\mu m$	601×2385	50	20

1.4 Application of Hyperspectral Imaging

HS enables precise analysis and classification across diverse fields by capturing rich spectral and spatial data. Its applications range from agriculture, environmental monitoring, and industrial quality control to healthcare, mineral exploration, and remote sensing, highlighting its versatility and impact across various domains.

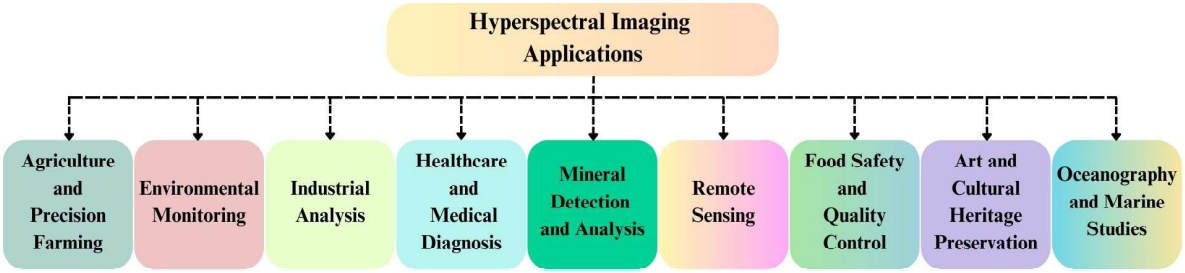

Figure 1.2: Applications of HSI techniques across multiple domains.

1. **Agriculture and Precision Farming:** HSI plays a critical role in modern agriculture by enabling crop health monitoring, disease detection, and yield estimation [12]. It aids in identifying stress factors such as nutrient deficiencies or pest infestations, ensuring timely interventions for precision farming.
2. **Environmental Monitoring:** HSI is extensively used for tracking environ-

mental changes, such as deforestation, urbanization, and water quality assessment [13]. It helps detect pollution, monitor biodiversity, and study the effects of climate change on ecosystems.

3. **Industrial Analysis:** HSI supports quality control and defect detection in industries such as manufacturing, food processing, and pharmaceuticals [14]. It provides detailed material composition analysis that ensure compliance with safety and quality standards.
4. **Healthcare and Medical Diagnostics:** In healthcare, HSI facilitates non-invasive diagnostics by detecting tissue abnormalities, identifying cancerous tissues, and monitoring wound healing [15] . It also aids in surgical guidance by differentiating healthy tissues from diseased ones.
5. **Mineral Detection and Analysis:** HSI is invaluable in geology for identifying and mapping mineral compositions [16]. It aids in resource exploration, mining, and studying the geological features of the Earth's surface.
6. **Remote Sensing:** In remote sensing, HSI is employed for land use mapping, vegetation analysis, and urban planning [17]. It provides detailed spatial and spectral information for disaster management and military reconnaissance.
7. **Food Safety and Quality Control:** HSI ensures food safety by detecting contaminants, adulterants, and microbial infections [18]. It is used to evaluate the freshness, texture, and composition of food products.
8. **Art and Cultural Heritage Preservation:** HSI is used in the analysis and preservation of artworks and historical artifacts [19]. It helps in identifying pigments, detecting restoration areas, and analyzing degradation processes without damaging the original material.
9. **Oceanography and Marine Studies :** HSI is applied in marine environments to monitor coral reef health, map seafloor habitats, and study the composition of water bodies [20].

1.5 Hyperspectral Image Classification

HSI data has been predominantly represented as a three-dimensional data cube $[H \times W \times B]$, where H and W denote the spatial dimensions, and B signifies the spectral dimension. This unique representation captures both spatial and spectral information across hundreds of contiguous wavelengths, enabling fine-grained material discrimination. This configuration establishes HSI as a powerful tool across diverse domains, including precision agriculture, environmental monitoring, and remote sensing,

by enabling detailed material characterization and analysis.

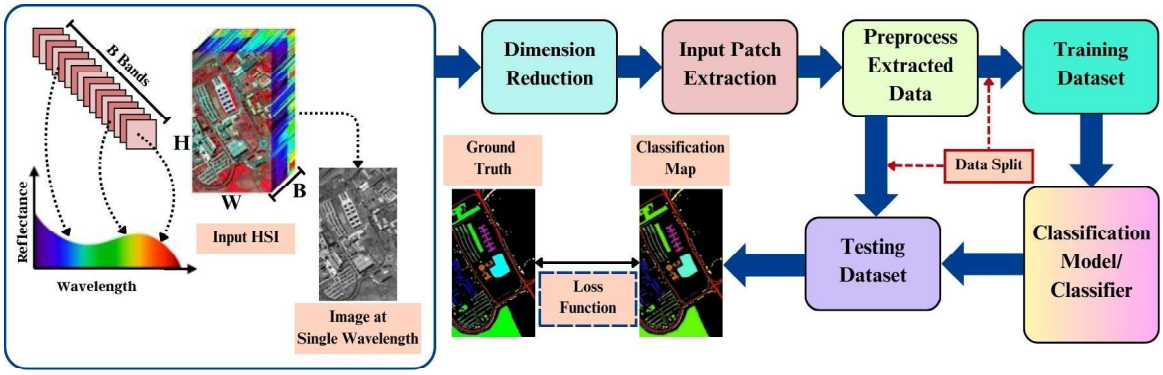

Figure 1.3: Workflow of HSI Classification: From Data Acquisition to Model Training and Classification Map Generation.

As shown in Figure 1.3, the HSI data cube captures spectral information across multiple wavelengths, with each slice representing a monochromatic spatial map of a specific band. This detailed spectral resolution enables precise material identification based on the unique reflectance properties of every pixel/object in the scene. Unlike RGB images, HSI provides hundreds of contiguous bands, enhancing classification accuracy for land cover mapping and anomaly detection applications. The HSI classification process begins with representing data as a 3D cube $[H \times W \times B]$, where each pixel holds a unique spectral signature. Due to high dimensionality, dimension reduction techniques minimize redundancy while preserving essential features. Input patches, including normalization and noise reduction, are extracted and preprocessed before being split into training and testing datasets. A classification model, such as a convolutional neural network (CNN) or vision transformer (ViT), is trained to learn spectral-spatial patterns, generating a classification map, which is refined using a loss function by comparing predictions with ground truth.

The classification of hyperspectral imagery has evolved as a sophisticated process of assigning individual pixels to specific classes based on their distinct spectral signatures. However, this task has posed significant challenges due to the high-dimensional nature of the data and the spectral redundancy among adjacent bands. To address these complexities, advanced ML and DL frameworks, such as CNN and ViT models, have been developed and successfully deployed. These algorithms leverage spectral-spatial feature extraction to enhance classification accuracy while reducing computational overhead.

1.6 Challenges in Hyperspectral Image Classification

HSI classification has become an important tool in many areas like agriculture, remote sensing, and environmental monitoring. However, it comes with many challenges because HSI data is complex and has many dimensions. Researchers face difficulties in choosing the right models, managing the large amount of data, and dealing with noise. These challenges need to be overcome to make the most of HSI. This section discusses the main obstacles that slow down progress in HSI classification.

1. **Lack of Comprehensive Survey:** The absence of a standardized evaluation framework for HSI classification models poses a significant challenge. Without a thorough survey, it becomes difficult to identify the most suitable models for specific datasets or applications, hindering progress and benchmarking efforts.
2. **Curse of Dimensionality:** The high dimensionality of HSI creates challenges in storage and processing. It can also lead to problems like overfitting, as there are often not enough labeled samples to train complex models properly. In addition, spectral redundancy and noise in nearby bands can reduce the performance.
3. **Data Noise and Quality Issues:** Noise in HSI arises from various sources, including sensor inaccuracies, atmospheric effects, and temporal drift. These factors compromise data quality and impact the reliability of classification results, making pre-processing and denoising essential yet complex tasks.
4. **Computational Complexity:** Advanced DL models, such as CNNs, Transformers, etc., require substantial computational resources. Their high memory and processing demands limit their scalability and applicability in resource-constrained environments. These issues pose challenges for real-time or edge-based deployments.
5. **Feature Extraction:** Effective extraction and integration of spectral and spatial features remain a persistent challenge. Spectral features capture material-specific information, while spatial features provide contextual relationships. To achieve accurate classification, these features need to be balanced and integrated, which requires advanced algorithms.
6. **Integration with Other Data Sources:** Combining HSI with other sources like light detection and ranging (LiDAR) or MSI enhances analysis but introduces complexities. Aligning datasets with differing resolutions and formats, as well as developing robust data fusion techniques, remains a challenging task.
7. **Scalability for Large-Scale Data:** The increasing availability of high-resolution HSI requires models that can handle large-scale datasets efficiently.

Existing models often face difficulties in scaling without compromising accuracy or computational efficiency, limiting their usability for real-world applications.

Addressing these challenges is crucial to advancing the field of HSI and ensuring its practical applicability in diverse domains.

1.7 Problems

In this thesis, we have addressed some of the key challenges mentioned above, which are outlined below:

1. **Limited Comprehensive Surveys for HSI Classification:** A major issue with HSI classification is the lack of systematic comparison of various feature extraction methods. Although feature extraction is a critical step, there is a noticeable limitation to detailed studies of different techniques in a consistent and comparative manner. As a result, researchers are left with limited guidance in selecting the best methods for different classification tasks.
2. **Inadequate Feature Extraction for HSI Classification:** HSI is a powerful tool for various applications, but effective classification is hindered by challenges in feature extraction, particularly concerning spatial-spectral relationships and long-term dependencies. The following points highlight the key issues related to inadequate feature extraction:
    - **Poor Correlated Spatial-Spectral Feature Extraction:** A major challenge in HSI classification is the failure to adequately capture the complex relationships between spectral and spatial features. Many methods do not effectively utilize the rich spatial-spectral information in HSI data, resulting in suboptimal classification accuracy. Developing improved methods to extract and integrate these features is essential for better outcomes.
    - **Inadequate Local and Global Feature Extraction with Lack of Long-term Dependencies Cues:** A significant challenge in HSI classification is the insufficient extraction of local and global features along with missing long-term dependency cues. HSI data contains both local (fine-scale) and global (coarse-scale) patterns vital for class differentiation. Local features emphasize small regions (e.g., texture, edges), while global features provide broader context (e.g., overall shape and distribution).
3. **High Computational Overhead in DL Models for Real-Time HSI Classification:** DL models have significantly enhanced HSI classification by capturing intricate spectral-spatial features. However, their complex architectures demand

extensive parameters, leading to high computational costs, prolonged training times, and substantial memory requirements. This limits real-time deployment in resource-constrained environments.

4. **Class Imbalance in HSI Datasets:** HSI datasets often have uneven class distributions, making some classes harder to classify due to limited training samples. Models tend to be biased toward frequent classes, reducing accuracy for underrepresented ones. Addressing this imbalance is essential for reliable performance across all classes.

1.8 Evaluation Metrics

This section delves into the essential evaluation metrics employed for HSI classification, highlighting their significance in gauging model performance and computational efficiency. These metrics will be systematically applied in Chapters 3 through 6 to analyze and compare the proposed methodologies.

The evaluation commences with foundational concepts: True Positives (TP), True Negatives (TN), False Negatives (FN), and False Positives (FP). TP refers to instances accurately classified within the target class, while TN represents cases correctly identified as not belonging to the target class. FN occurs when instances of the target class are mistakenly excluded, and FP arises when samples from non-target classes are incorrectly assigned to the target class. The key evaluation metrics are explored in subsequent subsections to ensure a detailed understanding of their relevance and application. These metrics collectively enable a comprehensive evaluation of model accuracy, computational efficiency, and practical applicability.

(a). Overall Accuracy (OA): It quantifies the ratio of correctly classified samples to the total number of samples. It serves as a general indicator of the accuracy of the classification model across all classes.

$$OA = \frac{\sum_{i=1}^{C} TP_i}{\sum_{i=1}^{C} (TP_i + TN_i + FP_i + FN_i)} \tag{1.1}$$

where TP_i, TN_i, FP_i, and FN_i, represent the true positives, true negatives, false positives, and false negatives for the i^{th} class, respectively.

(b). Average accuracy (AA): AA calculates the mean accuracy across all classes,

ensuring that performance is evaluated equally for each class, regardless of class size.

$$AA = \frac{1}{C}\sum_{i=1}^{C}\frac{TP_i}{TP_i + FN_i} \tag{1.2}$$

where AA addresses the limitations of OA by mitigating biases arising from class imbalances, providing a fairer evaluation for datasets with uneven class distributions.

(c). Kappa coefficient (κ): The Kappa Coefficient measures agreement between predicted and actual classifications while accounting for agreement occurring by chance.

$$\kappa = \frac{P_o - P_e}{1 - P_e}, \quad P_o = \frac{\sum_{i=1}^{C} TP_i}{\sum_{i=1}^{C}(TP_i + TN_i + FP_i + FN_i)}, \quad P_e = \sum_{i=1}^{C}\left[\frac{(TP_i + FP_i)(TP_i + FN_i)}{N^2}\right] \tag{1.3}$$

where N is the total number of samples. P_o is the observed agreement, and P_e is the expected agreement by chance, calculated as the sum of the expected probabilities for each class. The κ provides insight into the reliability of classifier beyond random chance, with values close to 1 indicating high reliability.

(d). Training Time (Tr): Tr is defined as the total duration required to optimize the model parameters over the given dataset. Mathematically, it can be expressed as:

$$Tr = \sum_{e=1}^{E}\sum_{b=1}^{B} T(e, b) \tag{1.4}$$

where E is the total number of epochs, B is the number of batches per epoch, and $T(e, b)$ represents the time taken to process batch b in epoch e.

(e). Inference/Testing Time (Te): Te represents the duration required for the trained model to generate predictions for a given input sample. Mathematically, it can be expressed as:

$$T_e = \frac{1}{N}\sum_{i=1}^{N} T(i) \tag{1.5}$$

where N is the total number of test samples, and $T(i)$ denotes the time taken to infer the i^{th} sample.

(f). Number of Parameters (P_M): The number of parameters represents the total trainable weights in a model, indicating its complexity and memory requirements. A higher number of parameters typically corresponds to a more complex model, which may have a greater capacity to learn intricate patterns, but it also demands more memory and computational resources.

The formulas serve as approximations to help understand the computational complexity

of different architectures and are expressed as:

$$P_M \approx \sum_{l=1}^{L_c} \left(C_{\text{in}}^l \times C_{\text{out}}^l \times K_v + B_l \right) + \sum_{m=1}^{L_t} \left(4d_m^2 + 2d_m N_m + B_m \right) + \sum_{n=1}^{L_m} \left(I_n \times O_n + B_n \right) \quad (1.6)$$

where L_c number of convolution layer,C_{in}^l and C_{out}^l are input and output channels of the l^{th} layer,K_v kernel size depends on the type of convolution 1D, 2D, or 3D convolution and ,B_l is the bias parameters. Further, for transformer layers, L_t is the number of transformer layers, d_m is the model's hidden dimension, N_m is the input sequence length, the term $4d_m^2$ accounts for the Multi-Head Self-Attention (MHSA) mechanism (query, key, value, and output projection), $2d_m N_m$ represents the feedforward network (FFN) parameters, and B_m denotes bias parameters. Finally for Multi-Layer Perceptron (MLP) layers, L_m is the number of MLP layers, I_n and O_n are input and output neurons for the n^{th} layer, and B_n represents bias parameters.

(g). Computational Overhead(C_{FLOPs}): As we know, the computational overhead is measured in Floating Point Operations (FLOPs), which quantifies the processing complexity of a model. It directly impacts the inference speed and hardware efficiency. It is a critical factor for real-time and resource-constrained applications. In our case, the total FLOPs of a model are estimated by summing the contributions of MLP layers, convolutional layers, and Transformer encoders. The approximate FLOP count is expressed as:

$$C_{\text{FLOPs}} \approx \sum_{l=1}^{L_c} \left(2 \times H_l \times W_l \times C_{\text{in}}^l \times C_{\text{out}}^l \times K_v \right) + \sum_{m=1}^{L_t} \left(8d_m^2 N_m + 4d_m N_m^2 \right) + \sum_{n=1}^{L_m} \left(2I_n \times O_n \right), \quad (1.7)$$

where for convolution layers H_l, W_l are the spatial dimensions of the input feature map of the l^{th} layer.

1.9 Frameworks and Tools

In this section, we discuss the tools and frameworks employed to implement the re-identification approaches outlined in Chapters 3 to 6. Several widely used frameworks are available for implementing DL models, among which we primarily utilize open-source libraries such as PyTorch, TensorFlow, and Keras. Google Colab is used as a Tool in our overall thesis work.

1. **TensorFlow:** It is developed by the Google Brain Team and released in 2015 [21]. It is a widely used open-source library for ML and DL. Over time, it has become a key framework for researchers and industry to enable projects such as

DeepDream [22].

2. **Keras:** It was introduced in 2015 and is an open-source Python library to implement DL models [23]. Keras is built on TensorFlow to create a user-friendly, modular, and extensible design.
3. **PyTorch:** It is an open-source ML library released in September 2016 [24]. It is developed by Facebook's AI Research (FAIR) lab [25]. It was built on the Torch framework, a scientific computing platform based on the Lua programming language. PyTorch has gained widespread adoption for tasks in computer vision and Natural Language Processing (NLP). The reason behind the adoption is due to its dynamic computation graph and ease of use.
4. **Google Colab:** It was launched in 2017 [26], a cloud-based platform by Google Research that enables Python execution via a Jupyter notebook interface. It provides a free environment for developing ML and DL models without requiring complex hardware. A key advantage is its free access to GPUs and TPUs, boosting computational performance [27].

1.10 Thesis Objectives

The primary aim of this thesis is to design accurate, efficient, and robust methods for HSI classification by addressing key limitations in existing approaches. The specific objectives are:

1. **Improve Spectral–Spatial Feature Discrimination:** Enhance the ability of models to capture meaningful spectral and spatial patterns through advanced feature learning strategies.
2. **Design an Effective Hybrid Framework:** Combine the strengths of CNNs for local feature extraction and Transformers for global context modeling in a balanced and lightweight manner.
3. **Achieve Computational Efficiency:** Develop architectures that reduce parameter count and attention complexity, ensuring scalability to large-scale HSI datasets.
4. **Handle Class Imbalance:** Incorporate tailored loss functions and training strategies to improve recognition of minority and difficult-to-classify classes.

1.11 Contribution of the Thesis

The thesis makes the following key contributions to HSI classification:

1. Firstly, we conducted an extensive survey that systematically evaluates the ML and DL approaches with respect to the improvement of feature extraction approaches for HSI classification. The survey encompasses classical methods of ML and DL (CNN, Recurrent Neural Network (RNN), Long-Short Term Memory (LSTM), Transformer, Graph Convolution Neural Network (GCN) and Generative Adversarial Network (GAN), hybrid models, and different learning paradigms (supervised, unsupervised, and semi-supervised). The survey has been divided into two sub-parts based on the types of features: (a) features that are based on spectral, spatial, and integrated spectral-spatial attributes, and (b) features that have been categorized as local, global, and integrated local-global features.
2. Secondly, we introduced the Morphologically Dilated Convolutional Neural Network (MDCNN) model, which combines mathematical morphological operations with DL for enhanced spectral-spatial feature extraction in HSI classification. The model uses morphological operations to extract spatial features, integrating them with HSI data to reduce the computational load of CNNs. MDCNN employs hybrid 3D-2D convolutions with dilation to expand the receptive field while minimizing parameters and overfitting. As spectral or spatial features alone are insufficient, thus we fused spectral-spatial features to improve accuracy. As a result, MDCNN effectively captures fine-grained local spectral–spatial details, which is highly beneficial for detecting subtle crop stress symptoms, nutrient deficiencies, and early-stage diseases that manifest in localized patterns.
3. Thirdly, our next chapter focuses on addressing key challenges, which are a) efficient extraction of local-global contextual features, b) integration and modeling of long-range dependency features, and c) high computational overheads of DL models. To resolve the challenges, we proposed a lightweight DL model that leverages logarithmic-based group convolutions and sine-cosine positional encoding for improved spectral-spatial feature extraction and capturing long-range dependency. LogGroupFormer achieved state-of-the-art results with reduced complexity. With its local–global feature learning, it is well suited for monitoring vegetation health, tracking forest degradation, and detecting wetland changes. It effectively captures both pixel-level details and the broader landscape context.
4. Second Lastly, we address the challenges of computation efficiency, long-range modeling, and robust feature extraction. In this case, we introduce a lightweight framework that optimizes the transformer model in terms of time complexity while maintaining a high classification accuracy. Basically, we designed a hybrid architecture called the Convolution-Kaiming-Gaussian Focused Linear Network

(CKGFLNet). A key component of this framework is the feature extraction module, which incorporates a Kaiming-Gaussian Focused Linear Attention (KGFLA) Transformer for effective global feature modeling. With its reduced computation and memory requirements, CKGFLNet enables scalable environmental monitoring across large geographical regions using satellite-based HSI data.

5. Lastly, we tackle the challenge of high computational complexity in CNN-Transformer models and the limitations of existing loss functions in addressing class imbalance. To overcome these issues, we propose the Hierarchical Clustering-Based Convolution with Flattened Kaiming-Gaussian Transformer Network (HieraKGTNet), a lightweight yet powerful framework for spectral-spatial learning in imbalanced HSI classification. This framework is designed to capture fine-grained local and global spectral-spatial features from imbalanced HSI data. It integrates several key components: the Local-Global Attentive Superpixel Segmentation (LGASS) module, Hierarchical Cluster-Based Convolution (HCBConv), a Kaiming Semantic Tokenizer, a Flattened Kaiming-Gaussian Transformer (FKGT), and a classifier. We also proposed Multiclass Poly-Focal Loss (MPF-Loss) to address class imbalance by dynamically adjusting class emphasis. HieraKGTNet, with its improved recognition of hard-to-classify classes, enables reliable discrimination of spectrally similar crop varieties (e.g., distinguishing between different grains or pulses), which is critical for precision agriculture.

These contributions collectively advance the field of HSI classification by addressing key challenges such as model suitability, high-dimensional data interpretation, and the development of cost-efficient, high-performance classification frameworks.

1.12 Thesis Organization

This thesis systematically addresses the challenges in HSI classification, with Chapters 2 to 6 building on published research contributions. The overall structure of the thesis is outlined in Figure 1.4, which provides a cohesive progression of concepts, methodologies, and findings. We have briefly illustrated the successive chapters as follows:

- **Chapter 2: Literature Survey**

  This chapter presents a comprehensive review of existing literature on HSI classification, covering classical ML techniques, advanced DL models, and hybrid DL approaches. Includes an in-depth survey of models based on spectral, spatial, and spectral-spatial feature extraction methods, categorizing them according to their scope. Furthermore, the chapter systematically classifies these techniques

into local, global, and integrated local-global feature representations, providing a structured analysis of their effectiveness in HSI classification.

– This chapter is an extension of the research paper: Kumar, V., Singh, R. S., Rambabu, M., & Dua, Y. (2024). "Deep learning for hyperspectral image classification: A survey". Computer Science Review, 53, 100658.

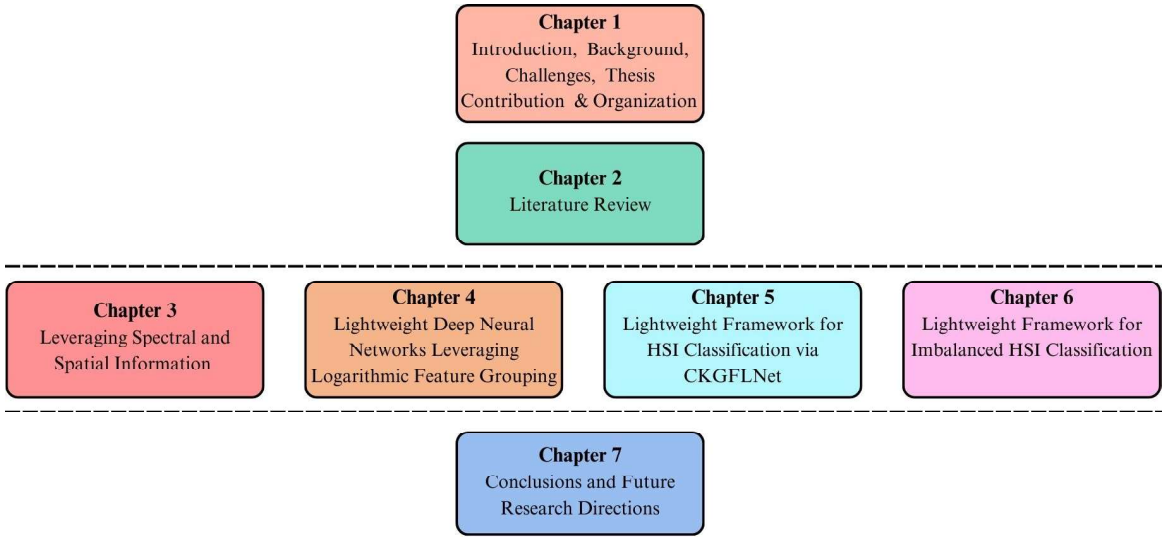

Figure 1.4: Structure of the thesis and outline of key chapters for HSI classification. Here, CKGFLNet is a Convolution-Kaiming-Gaussian Focused Linear Network.

- **Chapter 3: Leveraging Spectral and Spatial Information in HSI classification**

  In this chapter, we describe the proposed DL which is named as MDCNN. It addresses the challenges of enhancing spectral-spatial feature extraction in the HSI classification. The chapter elaborates on the usage of morphological operations to extract spatial features. Further, the chapter also explains the integration of spatial features with HSI data to reduce the computational load of CNNs. MDCNN also employs hybrid 3D-2D convolutions with dilation to expand the receptive field while minimizing parameters and overfitting.

  – This chapter is published: Kumar, V., Singh, R. S., & Dua, Y. (2022). "Morphologically dilated convolutional neural network for hyperspectral image classification". Signal Processing: Image Communication, 101, 116549.

- **Chapter 4: Lightweight Deep Neural Networks Leveraging Logarithmic Feature Grouping for HSI Classification**

  In this chapter, we focused on improving the extraction of local-global features with the integration of long-range dependencies features. The chapter also focuses

on the reduction of the high computational overheads of the proposed DL models. We described the proposed lightweight DL model via logarithmic-based group convolutions in detail. It also includes the concept of sine-cosine positional encoding for improvement of spectra-spatial feature extraction and capturing long-range dependency.

  - This chapter is published: Kumar, V., Singh, R. S., Nigam, N., Patel, K., & Jain, S. (2025). Enhancing Hyperspectral image classification through transformer-based contextualization and novel logarithmic convolutional techniques. Infrared Physics & Technology, 105826.

- **Chapter 5: Lightweight Framework for HSI Classification via Convolution-Kaiming-Gaussian Focused Linear Transformer Network**
  In this chapter, we introduce a streamlined framework that enhances the efficiency of the transformer model by reducing its time complexity while preserving high classification accuracy. Specifically, we propose a hybrid architecture, termed CKGFLNet. A central element of this framework is the feature extraction module, which integrates a KGFLA Transformer that enables robust global feature modeling.
  - This chapter is accepted at "6th International Conference on Machine Learning, Image Processing, Network Security, and Data Sciences (MIND2024)": Vinod Kumar, Ravi Singh, Nitika Nigam, Samujjal Choudhury and Rahul Kumar. "CKGFLNet: A Fast Hybrid Architecture for Hyperspectral Image Classification Leveraging Kaiming-Gaussian Attention".
- **Chapter 6: Lightweight Framework for Imbalanced HSI Classification**
  This chapter introduces the HieraKGTNet, a lightweight yet highly effective framework for spectral-spatial learning in imbalanced HSI classification. The proposed framework is specifically designed to capture both fine-grained local and global spectral-spatial features from imbalanced HSI data. Additionally, we introduce the MPF-Loss, which addresses class imbalance by dynamically adjusting the emphasis on different classes during training.
  - This chapter is under review in "IEEE Transactions on Instrumentation and Measurement": Vinod Kumar, Ravi Singh, Nitika Nigam, and Samujjal Choudhury. " HieraKGTNet: Hierarchical Kaiming-Gaussian Transformer for Imbalanced HSI Classification".
- **Chapter 7: Conclusions and Future Research Directions**
  This chapter consolidates the main outcomes of the thesis and emphasizes the novel contributions to the classification of HSI. It critically analyzes the effective-

ness of the proposed methodologies and their impact on advancing classification performance. Furthermore, this chapter outlines promising directions for future research, including strategies for scaling lightweight models to handle large-scale HSI datasets and the integration of emerging technologies such as DL optimization techniques, hybrid architectures, and advanced hardware accelerators to improve model efficiency and generalization further.

This organization ensures a systematic progression from problem identification to the development and evaluation of innovative solutions for HSI classification.