

Contents

| | |
|--|-------------|
| List of Figures | xi |
| List of Tables | xiv |
| List of Symbols | xv |
| List of Abbreviations | xvii |
| Abstract | xix |
| 1 Introduction | 1 |
| 1.1 Types of Digital Images | 1 |
| 1.2 Hyperspectral Images (HSI): Background and Motivations | 3 |
| 1.3 Hyperspectral Image Datasets | 4 |
| 1.4 Application of Hyperspectral Imaging | 5 |
| 1.5 Hyperspectral Image Classification | 6 |
| 1.6 Challenges in Hyperspectral Image Classification | 8 |
| 1.7 Problems | 9 |
| 1.8 Evaluation Metrics | 10 |
| 1.9 Frameworks and Tools | 12 |
| 1.10 Thesis Objectives | 13 |
| 1.11 Contribution of the Thesis | 13 |
| 1.12 Thesis Organization | 15 |
| 2 Literature Review | 19 |
| 2.1 HSI Fundamentals and Applications | 21 |
| 2.1.1 Challenges and Feature Extraction Strategies in HSI Classification | 22 |
| 2.2 Machine Learning Models | 23 |
| 2.2.1 Support Vector Machine (SVM) | 23 |
| 2.2.2 K-Nearest Neighbor (KNN) | 24 |

| | | |
|----------|--|-----------|
| 2.2.3 | Random Forest (RF) | 24 |
| 2.2.4 | K-means | 25 |
| 2.2.5 | Partial Least Square Discriminant Analysis (PLS-DA) | 26 |
| 2.2.6 | Machine Learning Model Analysis & Limitations | 26 |
| 2.3 | Transition to Deep Learning Models | 27 |
| 2.3.1 | Stacked Auto Encoder (SAE) | 27 |
| 2.3.2 | Deep Belief Network (DBN) | 28 |
| 2.3.3 | Convolution Neural Network (CNN) | 29 |
| 2.3.4 | Recurrent Neural Network (RNN) | 31 |
| 2.3.5 | Long Short Term Memory (LSTM) | 32 |
| 2.3.6 | Vision Transformer Network (ViT) | 33 |
| 2.3.7 | Generative Adversarial Network (GAN) | 35 |
| 2.3.8 | Graph Convolution Neural Network (GCN) | 36 |
| 2.4 | Categorization of DL Models for HSI Classification | 37 |
| 2.5 | DL Model Categorization Based on Feature Type | 37 |
| 2.5.1 | DL Models Leveraging Only Spectral Features | 37 |
| 2.5.2 | DL Model Leveraging Only Spatial Features | 38 |
| 2.5.3 | DL Model Leveraging Spectral-Spatial Features | 40 |
| 2.6 | Categorization Based on Contextual Scope | 45 |
| 2.6.1 | Local Contextual Features-Based Models | 46 |
| 2.6.2 | Global Contextual Features-Based Models | 47 |
| 2.6.3 | Combined Local-Global Contextual Feature-Based Models | 48 |
| 2.7 | Research Gaps and Objectives | 50 |
| 2.7.1 | Research Gaps: | 50 |
| 2.7.2 | Research Objectives: | 52 |
| 2.8 | Chapter Summary | 53 |
| 3 | Leveraging Spectral and Spatial Information in HSI classification | 55 |
| 3.1 | Introduction | 55 |
| 3.2 | Related Work | 57 |
| 3.3 | Problem Statement | 58 |
| 3.4 | Proposed Methodology | 58 |
| 3.4.1 | HSI Binarization Module | 60 |
| 3.4.2 | Mathematical Morphology Module | 60 |
| 3.4.3 | Fine-Grained Spectral-Spatial Feature Extraction Module | 61 |
| 3.4.4 | MDCNN Framework | 64 |
| 3.5 | Experiments: Implementations and Results | 65 |
| 3.5.1 | Dataset Description | 65 |

| | | |
|----------|---|------------|
| 3.5.2 | Training Details and Experimental Settings | 66 |
| 3.5.3 | Hyperparameter Sensitivity Analysis | 67 |
| 3.5.4 | Comparison with State-of-the-Art (SOTA) Methods | 68 |
| 3.6 | Chapter Summary | 73 |
| 4 | Lightweight Deep Neural Networks Leveraging Logarithmic Feature Grouping for HSI Classification | 75 |
| 4.1 | Introduction | 75 |
| 4.2 | Related Work | 77 |
| 4.2.1 | CNN-Based Local Contextual Feature Extraction | 77 |
| 4.2.2 | Vision Transformer (ViT) for Long-Range Dependencies | 77 |
| 4.2.3 | Group Convolution | 78 |
| 4.3 | Problem Statement | 78 |
| 4.4 | Proposed Methodology | 79 |
| 4.4.1 | Logarithmic 3D and 2D Convolution Module (LGConv) | 80 |
| 4.4.2 | Spatial-Context Patch Embedding (SCPE) module | 83 |
| 4.4.3 | Spatial Contextual-based Sine-Cosine Position Embedding Module | 83 |
| 4.4.4 | Transformer Encoder Module | 85 |
| 4.5 | Experiments: Implementations and Results | 87 |
| 4.5.1 | Dataset Description and Training Details | 87 |
| 4.5.2 | Experimental settings and Implementation Details | 87 |
| 4.5.3 | Hyperparameter Sensitivity Analysis | 88 |
| 4.5.4 | Ablation Studies | 91 |
| 4.5.5 | Comparison with State-of-the-Art (SOTA) Methods | 93 |
| 4.5.6 | Complexity Analysis | 99 |
| 4.6 | Chapter Summary | 100 |
| 5 | Lightweight Framework for HSI Classification via Convolution Kaiming-Gaussian Focused Linear Transformer Network | 101 |
| 5.1 | Introduction | 101 |
| 5.2 | Related Work | 102 |
| 5.2.1 | Convolution-Based Approaches | 102 |
| 5.2.2 | Vision Transformer-Based Approaches | 103 |
| 5.2.3 | Initialization Techniques | 103 |
| 5.3 | Problem Statement | 104 |
| 5.4 | Proposed Methodology | 104 |
| 5.4.1 | Hybrid Spectral-Spatial Feature Extractor and Tokenizer Module | 106 |
| 5.4.2 | Kaiming-Gaussian Focused Linear Attention Transformer Block | 106 |

| | | |
|----------|---|------------|
| 5.5 | Experiments: Implementations and Results | 108 |
| 5.5.1 | Dataset Description and Training Details | 108 |
| 5.5.2 | Hyperparameter Sensitivity Analysis | 109 |
| 5.5.3 | Comparison With State-of-the-Art (SOTA) Methods | 111 |
| 5.5.4 | Computational Cost Analysis | 113 |
| 5.6 | Chapter Summary | 114 |
| 6 | Lightweight Framework for Imbalanced HSI Classification | 115 |
| 6.1 | Introduction | 115 |
| 6.2 | Related Work | 116 |
| 6.2.1 | Heavy-weighted DNNs | 116 |
| 6.2.2 | Loss functions for Imbalance Sample Issues in HSI dataset . . . | 117 |
| 6.2.3 | Superpixel Segmentation | 118 |
| 6.3 | Problem Statement | 118 |
| 6.4 | Proposed Methodology | 119 |
| 6.4.1 | Local-Global Attentive Superpixel Segmentation (LGASS) . . . | 121 |
| 6.4.2 | Hierarchical-Cluster-Based 3D-2D Convolution Module(HCBCConv) | 122 |
| 6.4.3 | Kaiming Tokenizer | 124 |
| 6.4.4 | Flattened Kaiming-Gaussian Transformer (FKGT) Module . . . | 125 |
| 6.4.5 | Multiclass Poly-focal Loss | 126 |
| 6.5 | Experiments: Implementations and Results | 127 |
| 6.5.1 | Dataset Description and Training Details | 128 |
| 6.5.2 | Experimental Settings and Implementation Details | 128 |
| 6.5.3 | Hyperparameter Sensitivity Analysis | 129 |
| 6.5.4 | Ablation Studies | 130 |
| 6.5.5 | Comparison with State-of-the-Art Methods | 132 |
| 6.5.6 | Complexity Analysis | 137 |
| 6.6 | Chapter Summary | 138 |
| 7 | Conclusions and Future Research Directions | 141 |
| 7.1 | Contributions | 141 |
| 7.2 | Limitations | 144 |
| 7.3 | Future Research Directions | 145 |
| | References | 147 |
| | List of Publications | 173 |

List of Figures

| | | |
|------|--|----|
| 1.1 | Visualization of Types of Images | 2 |
| 1.2 | Applications of HSI techniques across multiple domains. | 5 |
| 1.3 | Workflow of HSI Classification | 7 |
| 1.4 | Organization of the thesis | 16 |
| 2.1 | HSI-classification framework categorizing DL models by type, features, extraction methods, and methodologies | 20 |
| 2.2 | (a) Distribution of HSI application articles (2014–2024); (b) Year-wise publications by DL method type | 21 |
| 2.3 | HSI classification taxonomy | 22 |
| 2.4 | ML-based HSI classification models | 23 |
| 2.5 | (a) Flow chart of Auto-encoder. (b) SAE-based HSI-classification Model. | 28 |
| 2.6 | (a) Flow chart of RBM. (b) DBN-based HSI-classification Model | 29 |
| 2.7 | CNN-based HSI-classification Model | 30 |
| 2.8 | RNN-LSTM frameworks in HSI classification | 31 |
| 2.9 | Attention-Transformer Framework in HSI classification | 33 |
| 2.10 | GAN framework in HSI classification | 35 |
| 2.11 | GCN-based HSI classification | 36 |
| 3.1 | Flowchart of the MDCNN Framework for HSI Classification. | 59 |
| 3.2 | Morphological Feature Maps erosion, gradient and closing. | 61 |
| 3.3 | Illustration of 3D and 2D convolutions | 62 |
| 3.4 | Dilated convolution with dilation rates: (a) one, (b) two and (c) three. | 62 |
| 3.5 | Visualization of classification outcome map for IP dataset | 69 |
| 3.6 | Visualization of classification outcome Map for PU dataset | 70 |
| 3.7 | Visualization of classification outcome Map for SA dataset | 71 |
| 3.8 | Visual Results based on Execution time of different models | 72 |
| 4.1 | Framework of proposed LogGroupformer | 79 |

| | | |
|------|---|-----|
| 4.2 | Illustration of The Logarithmic Group 3D and 2D Convolution (LGConv). | 80 |
| 4.3 | Illustration of the spatial-contextual patch Embedding (SCPE) module. | 83 |
| 4.4 | Illustration of proposed Spectral-Spatial Contextual-based Sine-Cosine Position Embedding Module | 85 |
| 4.5 | Transformer Attention Mechanism Visualization | 86 |
| 4.6 | Assessment of OA(%) for LogGroupFormer with Varying Patch Sizes. | 89 |
| 4.7 | Assessment of OA(%) with Varying Filter Counts in LGConv3D Layer. | 89 |
| 4.8 | Assessment of OA(%) with Varying Filter Counts in LGConv2D Layer. | 90 |
| 4.9 | Assessment of OA(%) for LogGroupFormer with Varying Number of Logarithmic Groups in the LGConv3D Layer. | 90 |
| 4.10 | Assessment of OA(%) for LogGroupFormer with Varying Number of Logarithmic Groups in the LGConv2D Layer. | 91 |
| 4.11 | Assessment of OA(%) for LogGroupFormer with Varying Learning Rates. | 91 |
| 4.12 | Comparison of Visual classification maps of models for IP dataset. | 94 |
| 4.13 | Comparison of Visual classification maps of models for PU dataset. | 95 |
| 4.14 | Comparison of Visual classification maps of models for HU dataset. | 97 |
| 4.15 | Comparison of Visual classification maps of models for LK dataset. | 98 |
| | | |
| 5.1 | Illustration of the CKGFLNet | 105 |
| 5.2 | Visualization of different transformers for HSI classification | 107 |
| 5.3 | Presents a comparative analysis of hyperparameter sensitivity of CKGFLNet’s performance across IP, PU, and SA datasets. | 110 |
| 5.4 | Visual Comparison Results of Initialization Strategies | 111 |
| 5.5 | Visual classification results for IP, PU, and SA datasets | 112 |
| | | |
| 6.1 | Illustration of proposed HieraKGTNet framework | 120 |
| 6.2 | LGASS incorporates two superpixel segmentation methods: local- and global-SLIC | 121 |
| 6.3 | Illustration of the HCBCConv Module | 122 |
| 6.4 | Illustration of (a) Kaiming Tokenizer, (b) FKGT Module, and (c) KGFL-MHSA Module for HSI classification. | 125 |
| 6.5 | Comparison of OA (%) for IP, PU, HU, and LK datasets across varying (a) Patch Sizes and (b) LR. | 129 |
| 6.6 | Comparison of OA (%) on numbers of filters and clusters of HBC3DConv and HBC2DConv | 130 |
| 6.7 | Ablation studies of different components in models | 131 |
| 6.8 | Visual classification results for IP, PU, and SA datasets for HieraKFGTNet. | 134 |

List of Tables

| | | |
|-----|---|----|
| 1.1 | Overview of Publicly Available Sensor-based HSI Datasets for Classification | 5 |
| 2.1 | Comparison of DL Models Using Spectral Features | 38 |
| 2.2 | Comparison of DL Models Using Spatial Features | 39 |
| 2.3 | Comparison of DL Models Using Both Spectral and Spatial Features in Pre-processing Stage | 41 |
| 2.4 | Comparison of DL Models Using Both Spectral and Spatial Features in Post-processing Stage | 42 |
| 2.5 | Comparison of DL Models Using Both Spectral and Spatial Features in Correlated Integrated Stage | 44 |
| 2.5 | Continued: Feature Fusion in the Correlated Integrated Stage. | 45 |
| 2.6 | Comparison of DL Models Using Local Contextual Features | 47 |
| 2.7 | Comparison of DL Models Using Global Contextual Features | 48 |
| 2.8 | Comparison of DL Models Using Both Local and Global Contextual Features | 49 |
| 2.9 | Comprehensive Analysis of HSI Classification DL Models and Feature Processing | 51 |
| 3.1 | Structure of MDCNN Framework for (a) IP (b), PU, and (c) SA datasets. | 65 |
| 3.2 | Overview of Class Distribution and Sample Allocation for IP, PU, and SA Datasets | 66 |
| 3.3 | MDCNN Hyperparameter Sensitivity Analysis | 67 |
| 3.4 | Impact of Dilated Convolution Layers on MDCNN Performance and Trainable Parameters | 68 |
| 3.5 | Classification Accuracy Comparison on IP Dataset with Best Results Highlighted in Terms of OA(%), AA (%) and Kappa Coefficient. | 69 |
| 3.6 | Classification Accuracy Comparison on PU Dataset with Best Results Highlighted in Terms of OA(%), AA (%) and Kappa Coefficient. | 70 |

| | | |
|-----|---|-----|
| 3.7 | Classification Accuracy Comparison on SA Dataset with Best Results Highlighted in Terms of OA(%), AA (%) and Kappa Coefficient. | 71 |
| 4.1 | Computation cost analysis of LGConv module. | 82 |
| 4.2 | IP, PU HU, and LK Dataset description with number of training, validation, and test samples per class and overall used within the model. | 88 |
| 4.3 | Ablation Study of the Components in the LogFormer Model | 92 |
| 4.4 | Ablation study of the proposed LogGroupFormer with its variant, which uses only conventional convolutional layers, named as $(SC)^2Former$. . . | 93 |
| 4.5 | Quantitative performance of various classification methods in terms of OA, AA, and $\kappa \times 100$, along with the accuracies for each class on the IP dataset | 94 |
| 4.6 | Quantitative performance of various classification methods in terms of OA (%), AA (%), and $\kappa \times 100$, along with the accuracies for each class on the PU dataset | 95 |
| 4.7 | Quantitative performance of various classification methods on HU dataset | 96 |
| 4.8 | Quantitative performance of various classification methods on LK dataset | 98 |
| 4.9 | Complexity Analysis of LGFormer vs. SOTA Models | 99 |
| 5.1 | Computational Analysis for Conventional and KGFLA Transformer. . . | 108 |
| 5.2 | IP, PU, and SA Dataset description with a number of training, and test samples per class and overall used within the model. | 109 |
| 5.3 | Classification Performance Comparison with SOTA in terms of OA (%), AA (%), and $\kappa \times 100$ | 112 |
| 5.4 | Comparative Computational Complexity Analysis of Models Based on OA, Training, and Testing Time | 113 |
| 6.1 | Computation cost analysis in the HCBCConv module | 124 |
| 6.2 | Time Complexity Analysis: Conventional vs FKGT Module. | 126 |
| 6.3 | Descriptions for IP, PU, HU, and LK, with color-code, class name, and the number of training, and test samples. | 128 |
| 6.4 | Ablation study of the HieraKGTNet model | 132 |
| 6.5 | Quantitative performance of SOTA methods in terms of OA (%), AA (%), and $\kappa \times 100$ | 133 |
| 6.6 | Complexity Analysis of SOTA Models on IP, PU, HU, and LK datasets for HieraKGTNet | 137 |

List of Symbols

| Symbol | Description |
|----------------------------|----------------------------------|
| AA | Average Accuracy |
| B | Spectral Bands in HSI |
| C_{FLOPs} | Computational Overheads |
| C_{in} | Number of Input Channels |
| C_{out} | Number of Output Channels |
| \hat{C} | Classification Accuracy |
| H | Height of the HSI |
| h | Number of Heads in ViT |
| I | HSI Input |
| K | Key of ViT |
| κ | Kappa Coefficient |
| K_v | Size of the Filter |
| $\hat{\mathcal{L}}_{comp}$ | Computational Burden |
| λ | Regularization Factor |
| OA | Overall Accuracy |
| P_M | Number of Parameters |
| Q | Query of ViT |
| S | Initial Input Patch Spatial Size |
| Te | Testing Time |
| Tr | Training Time |
| V | Value of ViT |
| W | Width of The HSI |
| μ | Micro meter |
| \hat{Y} | Predicted Class Probability |
| Y | Actual Class Probability |

Abbreviations

| Abbreviation | Description |
|---------------------|--|
| 1D | One Dimensional |
| 2D | Two Dimensional |
| 3D | Three Dimensional |
| CKGFLNet | Convolution-Kaiming-Gaussian Focus Linear Network |
| CNN | Convolutional Neural Network |
| DBN | Deep Belief Network |
| DL | Deep Learning |
| FC | Fully Connected Layer |
| FKGT | Flattened Kaiming-Gaussian Transformer |
| GAN | Generative Adversarial Network |
| GCN | Graph Convolution Network |
| GeLU | Gaussian Error Linear Unit |
| GRU | Gated Recurrent Unit |
| HBC3DConv | Hierarchical Band Clustering Based 3D Convolution |
| HCBCConv | Hierarchical Clustering-Based Convolution |
| HCC2DConv | Hierarchical Channel Clustering Based 2D Convolution |
| HieraKGTNet | Hierarchical Kaiming-Gaussian Transformer Network |
| HS | Hyperspectral Imaging |
| HSI | Hyperspectral Image |
| HU | Houston |
| IP | Indian Pines |
| KGFLA | Kaiming-Gaussian Focus Linear Attention |
| KNN | K-Nearest Neighbor |
| LGASS | Local-Global Attentive Superpixel Segmentation |
| LGConv | Logarithmic Group 3D and 2D Convolution |
| LGConv2D | Two Dimensional Logarithmic Group Convolution |
| LGConv3D | Three Dimensional Logarithmic Group Convolution |

| Abbreviation | Description |
|---------------------|---|
| LGSA-ViT | Light Self Gaussian Attention ViT |
| LiDAR | Light Detection and Ranging |
| LK | WHU-Hi-LongKou |
| LR | Learning Rate |
| LSTM | Long-Short Term Memory |
| MDCNN | Morphologically Dilated CNN |
| MHSA | Multi-Head Self-Attention |
| ML | Machine Learning |
| MLP | Multi-Layer Perceptron |
| MPF-Loss | Multiclass Poly-Focal Loss |
| MSI | Multispectral Image |
| PCA | Principal Component Analysis |
| PLS-DA | Partial Least Square Discriminant Analysis |
| PS | Patch Size |
| PU | University of Pavia |
| ReLU | Rectified Linear Unit |
| RF | Random Forest |
| RNN | Recurrent Neural Network |
| SA | Salinas Valley |
| SAE | Stacked Auto Encoder |
| SCPE | Spatial Contextual Patch Embedding |
| $(SC)^2PosEmbed$ | Spatial Contextual Sine-Cosine Positional Embedding |
| SF | Spectral Former |
| SOTA | State-of-the-Art |
| SSFTT | Spectral-Spatial Feature Transformer |
| SVM | Support Vector Machine |
| ViT | Vision Transformer |