

2. DATA ANALYSIS REPORT

Advanced Data Analysis - Spectral and Dimensionality Analysis

Spectral Band Analysis

RGB Channel Distributions

The RGB spectral analysis reveals distinct patterns across land types:

- Forest/Vegetation Classes: Higher green channel values
- Water Bodies: Dominant blue channel characteristics
- Urban Areas: Balanced RGB with structural textures
- Agricultural Land: Seasonal variations in color patterns
- Spectral Signature Patterns

Each land type exhibits unique spectral signatures:

- AnnualCrop: Seasonal green-brown transitions
- Forest: Consistent green dominance
- River: Blue spectrum prominence
- Industrial: Gray-scale dominance with structural features
- Dimensionality Reduction Results
- Principal Component Analysis (PCA)
- First 3 components explain 85% of variance
- Clear separation between natural and man-made classes
- Water bodies form distinct clusters
- Agricultural classes show some overlap

t-SNE Visualization

- 2D embedding shows reasonable class separation
- Forest and HerbaceousVegetation form tight clusters
- Highway and Industrial classes show some overlap
- Residential areas display varied patterns
- Feature Importance Analysis
- Green Channel Intensity: Most discriminative for vegetation
- Texture Features: Critical for urban/industrial classification
- Color Variance: Important for water body identification
- Spatial Patterns: Key for structural class discrimination
- Data Augmentation Strategy

Recommended augmentations:

- Rotation and flipping for orientation invariance

- Color jittering for illumination variations
- Random cropping for spatial robustness
- MixUp augmentation for improved generalization