Data Appendix

Analysis Data File 1: Training Dataset

Unit of Observation: Each row represents a 512x512 pixel image-mask pair from the training set. Each row represents a single 512x512 pixel satellite image from either the Amazon Forest biome with an associated TIFF mask (forest indicated in white and background in black color).

Variable: Image Bands

- Definition: Pixel intensity values for four Sentinel-2 spectral bands (Red, Green, Blue, Near-Infrared) for each image.
- Source: Derived from GeoTIFF files; converted to byte type (0-255).
- Missing Observations: 499(0)

	Band	mean	std_dev	min	25th_percentile	median
0	Band_1	622.730073	540.781477	1	226.0	375.0
1	Band_2	683.184264	368.232800	1	438.0	572.0
2	Band_3	436.802551	342.115868	1	225.0	352.0
3	Band_4	2950.608018	633.894052	0	2626.0	2962.0

75th_percentile	max
932.0	16280
893.0	15016
590.0	14520
3326.0	15505

Variable: Mask

- **Definition**: Binary classification of each pixel as forested (1/white) or non-forested (0/black).
- **Source**: Tiff files paired with each image in the Input Data Files.
- **Missing Observations**: 499(0)

	Band	mean	std_dev	min	25th_percentile	median	75th_percentile
0	Band_1	0.528634	0.499179	0	0.0	1.0	1.0

	max
0	1

Analysis Data File 2: Validation Dataset

Unit of Observation: Each row represents a 512x512 pixel image-mask pair from the validation set. Each row represents a single 512x512 pixel satellite image from either the Amazon Forest biome with an associated Tiff mask (forest indicated in white and background in black color).

Variable: Image Bands

- Definition: Pixel intensity values for four Sentinel-2 spectral bands (Red, Green, Blue, Near-Infrared) for each image.
- Source: Derived from GeoTIFF files; converted to byte type (0-255).
- Missing Observations: 100(0)

Band		mean	std_dev	min	25th_percentile	median
0	Band_1	644.288977	576.386682	1	232.0	389.0
1	Band_2	711.795556	429.313111	15	448.0	596.0
2	Band_3	463.694395	425.683430	1	240.0	372.0
3	Band_4	3013.682640	631.173541	0	2677.0	3009.0

	75th_percentile	max
0	955.0	9306
1	907.0	9423
2	590.0	27200
3	3389.0	9203

Variable: Mask

- **Definition**: Binary classification of each pixel as forested (1/white) or non-forested (0/black).
- **Source**: Tiff files paired with each image in the Input Data Files.
- **Missing Observations**: 100(0)

Band		mean	std_dev	min	25th_percentile	median	75th_percentile
0	Band_1	0.521894	0.49952	0	0.0	1.0	1.0

	max
0	1

Analysis Data File 3: Test Dataset

Unit of Observation: Each row represents a single 512x512 pixel satellite image from the Amazon Forest biomes. There are no masks included for this dataset.

Variable: Image Bands

- Definition: Pixel intensity values for four Sentinel-2 spectral bands (Red, Green, Blue, Near-Infrared) for each image.
- Source: Derived from GeoTIFF files; converted to byte type (0-255).
- Missing Observations: 20(0)

	Band	mean	std_dev	min	25th_percentile	median \
0	Band_1	700.459424	663.514171	1	225.0	385.0
1	Band_2	775.127870	564.288086	16	434.0	573.0
2	Band_3	538.301772	597.925915	1	227.0	338.0
3	Band_4	3019.951963	653.738088	0	2680.0	3020.0

	75th_percentile	max
0	1098.0	6216
1	952.0	6436
2	633.0	6600
3	3418.0	7264