

Supplementary Materials

Methodology

DTAug. As shown in Fig. 1, we present the detailed process of DTAug strategy.

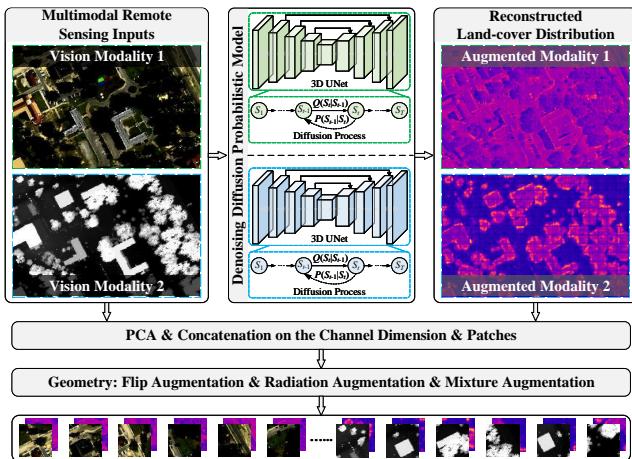


Figure 1: Diagram of DTAug strategy.

FRGCM. Detailed structure of FRGCM in SFIE is shown in Fig. 2. Specifically, the input information is first fed into a 1×1 Conv layer to adjust the number of channels, and then the channels are divided into two groups for local feature extraction within each group, respectively. Furthermore, to achieve progressive inter-group information interaction, output features of different 3×3 Conv layers in each group are performed progressive summed pixel-wise summation and then fed into inter-group interaction Conv branch for further feature extraction.

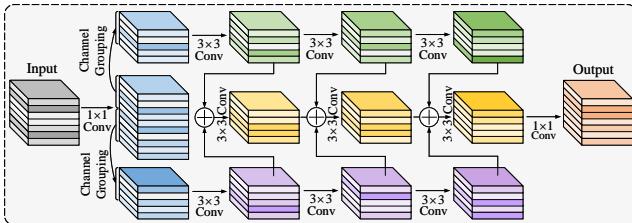


Figure 2: Structure of FRGCM.

Experiments

Datasets. We provide a detailed description of three datasets used in Table 1. For the MUUFL dataset, different class regions are large and densely distributed. In contrast, for the Trento dataset, different class regions are relatively sparsely distributed, and there are significant variations in region shapes. For the HU2013 dataset, different class regions show point-like patterns and have a wide distribution range. Therefore, there are significant distribution differences between the multimodality datasets, which pose higher demands on the model generalization ability. Additionally,

we present a overview of mutual RSMG in Fig. 3, namely SD \leftrightarrow TD: MUUFL \leftrightarrow Trento \leftrightarrow HU2013 (MU \leftrightarrow TR \leftrightarrow HU).

Table 1: Remote sensing dataset descriptions, including MUUFL, Trento, and HU2013 datasets.

| Dataset | Spatial Size | Data Type | Band No. | Class No. | Samples |
|---------|------------------|-------------|----------|-----------|---------|
| MUUFL | 325×220 | HS LiDAR | 64 2 | 3 | 36173 |
| Trento | 166×600 | HS LiDAR | 63 1 | 3 | 15200 |
| HU2013 | 349×960 | HS LiDAR | 144 1 | 3 | 2274 |

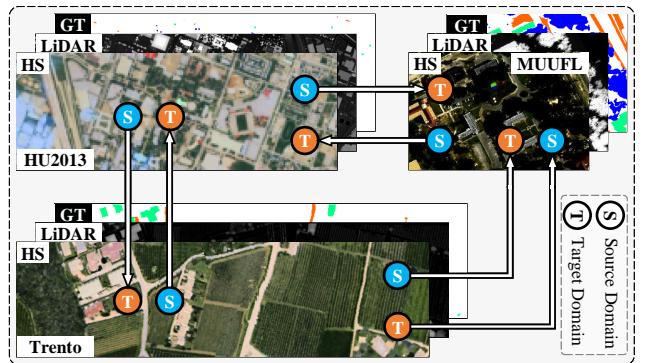


Figure 3: Mutual multimodality generalization description. MUUFL, Trento, and HU2013 datasets are used as source and target domains for mutual generalization, respectively.

Experimental environment. Python 3.9.18; PyTorch 2.0.1; Software VSCode; GPU NVIDIA GeForce RTX 4060; CPU Intel 12th Gen Core i7-12650H; Memory 16GB; Operation System Win 11.

Hyper-parameters. All hyper-parameters for the comparison methods are set to their default configurations.

Qualitative experimental results. To achieve a clearer comparison for classification performance in different classes, we expand the classification results presented in the paper, as shown in Table 3–8, where the symbol ‘—’ denotes no pixels correctly identified.

Modal extensibility. To preliminarily verify the classification performance of the proposed method on more multimodal remote sensing datasets, we conduct experiments on the MS-SAR Augsburg (Aug) and Berlin (Ber) datasets (Hong et al. 2023). The two datasets are divided into four sub-parts respectively for generalization, and the results are shown in Table 2. This dataset combination is more diverse and larger scene. We select 7 classes for the RSMG experiments. The results show that the proposed method has good modal extensibility.

Computational complexity. Computational complexity of different methods is shown in Table 9. Here, FVMGN-NW represents the FVMGN without the MWDIS module. Observation shows that the proposed method has a moderate computational complexity.

Table 2: Preliminary classification results (%) on the MS-SAR Augsburg and Berlin datasets.

| Class No. | Class Name | Aug1→Ber1 | Aug2→Ber2 | Aug3→Ber3 | Aug4→Ber4 |
|-----------|-----------------|-----------|-----------|-----------|-----------|
| 1 | Surface water | 32.14 | 26.88 | 29.21 | 26.96 |
| 2 | Urban fabrics | 80.46 | 78.05 | 80.25 | 45.73 |
| 3 | Industrial, ... | 58.12 | 60.37 | 56.26 | 54.84 |
| 4 | Mine, ... sites | 32.01 | 0.450 | 4.900 | 7.390 |
| 5 | Vegetation | 31.71 | 1.760 | 20.57 | 7.490 |
| 6 | Arable land | 50.68 | 53.82 | 94.99 | 86.54 |
| 7 | Forests | 68.42 | 84.29 | 50.77 | 79.75 |
| OA | / | 65.18 | 65.39 | 72.23 | 70.09 |
| AA | / | 50.50 | 43.66 | 48.13 | 44.10 |
| Kappa | / | 47.26 | 39.26 | 61.96 | 57.72 |

Visualization. We present the visual results of the proposed FVMGN and SOTA methods, as shown in Figs. 4–9. Overall, classification maps of FVMGN contain less noise and fewer contaminated areas, resulting in a good visual effect.

Manual fine-tuning for text descriptions: We provide GPT with modality-specific text generation prompts, and the generated texts may have redundancy such as repetition or excessive modifiers. Manual fine-tuning refers to removing irrelevant/repeated words to limit text length. Future work will explore more automated GPT post-fine-tuning process.

Label information: Different texts are generated from class labels, and they are available for training but are prohibited for inference, avoiding label information leakage.

Loss weight: Weights of different loss items are the same in this work. Future work will explore adaptive weighting, such as gradient norm or soft weighting methods.

Reproducibility. We will release the *source code* and the adjusted remote sensing *multimodality generalization datasets* in this work soon, which is beneficial for the development of the remote sensing community.

References

- Hong, D.; Zhang, B.; Li, H.; Li, Y.; Yao, J.; Li, C.; Werner, M.; Chanussot, J.; Zipf, A.; and Zhu, X. X. 2023. Cross-city matters: A multimodal remote sensing benchmark dataset for cross-city semantic segmentation using high-resolution domain adaptation networks. *Remote Sensing of Environment*, 299: 113856.

Table 3: Classification results (%) of different methods on the MU→TR dataset combination.

| Class No. | MFT | MsFE-IFN | CMFAEN | DKDMN | SDENet | LLURNet | FDGNet | TFTNet | ADNet | ISDGS | EHSNet | LDGNet | FVMGN |
|-----------|-------------------------|-------------------------|------------------|-------------------------|------------------|------------------|------------------|------------------|-------------------------|-------------------------|-------------------|-------------------|-------------------------|
| 1 | 99.80 \pm 0.33 | 98.03 \pm 2.40 | 98.94 \pm 1.89 | 99.81 \pm 0.51 | 97.98 \pm 1.32 | 98.76 \pm 1.29 | 99.64 \pm 0.24 | 98.49 \pm 0.81 | 99.64 \pm 0.27 | 98.79 \pm 0.96 | 99.71 \pm 0.49 | 98.76 \pm 1.86 | 98.08 \pm 2.11 |
| 2 | 73.98 \pm 8.81 | 86.75 \pm 5.72 | 90.39 \pm 6.08 | 55.82 \pm 13.5 | 85.28 \pm 10.7 | 90.75 \pm 6.87 | 97.16 \pm 1.11 | 77.81 \pm 10.3 | 98.49 \pm 0.61 | 97.10 \pm 2.48 | 35.39 \pm 40.65 | 72.92 \pm 13.98 | 86.89 \pm 6.40 |
| 3 | 10.59 \pm 6.91 | 36.01 \pm 8.31 | 22.94 \pm 5.42 | 37.81 \pm 10.1 | 8.000 \pm 9.59 | 0.010 \pm 0.02 | 0.070 \pm 0.22 | 5.030 \pm 3.62 | --- | 3.650 \pm 6.05 | 54.16 \pm 27.25 | 45.39 \pm 20.79 | 80.83 \pm 10.0 |
| OA | 77.59 \pm 1.01 | 83.98 \pm 1.86 | 82.82 \pm 1.06 | 78.95 \pm 2.38 | 78.14 \pm 1.40 | 78.23 \pm 1.42 | 80.11 \pm 0.26 | 76.32 \pm 1.67 | 80.36 \pm 0.16 | 80.27 \pm 0.67 | 77.58 \pm 4.17 | 83.17 \pm 1.63 | 92.44 \pm 1.24 |
| AA | 61.46 \pm 1.64 | 73.60 \pm 3.61 | 70.79 \pm 1.46 | 64.48 \pm 3.77 | 63.75 \pm 2.17 | 63.17 \pm 2.20 | 65.62 \pm 0.37 | 60.44 \pm 2.72 | 66.03 \pm 0.19 | 66.51 \pm 1.23 | 63.09 \pm 6.28 | 72.36 \pm 3.41 | 88.55 \pm 2.13 |
| Kappa | 53.54 \pm 2.70 | 69.25 \pm 4.30 | 66.81 \pm 2.02 | 58.08 \pm 4.82 | 60.22 \pm 2.84 | 60.49 \pm 2.94 | 64.14 \pm 0.51 | 56.15 \pm 3.77 | 64.67 \pm 0.28 | 64.59 \pm 1.18 | 58.49 \pm 8.52 | 69.65 \pm 3.15 | 86.59 \pm 2.16 |

Table 4: Classification results (%) of different methods on the MU→HU dataset combination.

| Class No. | MFT | MsFE-IFN | CMFAEN | DKDMN | SDENet | LLURNet | FDGNet | TFTNet | ADNet | ISDGS | EHSNet | LDGNet | FVMGN |
|-----------|------------------|------------------|-------------------------|------------------|------------------|------------------|-------------------------|------------------|-------------------------|-------------------------|------------------|-------------------|-------------------------|
| 1 | 2.530 \pm 1.41 | 7.540 \pm 0.56 | 3.220 \pm 2.33 | 7.560 \pm 4.18 | 98.64 \pm 0.96 | 98.53 \pm 0.04 | 99.72 \pm 0.39 | 97.40 \pm 2.33 | 100.0 \pm 0.00 | 99.51 \pm 0.66 | 84.39 \pm 14.9 | 98.44 \pm 1.41 | 94.25 \pm 3.29 |
| 2 | 51.02 \pm 18.4 | 9.720 \pm 7.65 | 6.430 \pm 6.10 | 19.37 \pm 9.51 | 64.76 \pm 11.2 | 70.14 \pm 13.1 | 65.11 \pm 6.90 | 69.12 \pm 14.9 | 88.13 \pm 10.2 | 89.63 \pm 6.11 | 7.11 \pm 11.12 | 16.53 \pm 13.71 | 93.11 \pm 2.85 |
| 3 | 29.35 \pm 20.4 | 28.57 \pm 5.15 | 48.90 \pm 11.1 | 13.37 \pm 4.51 | 33.33 \pm 7.69 | 26.13 \pm 11.5 | 19.39 \pm 2.96 | 46.44 \pm 8.07 | 37.26 \pm 9.28 | 24.23 \pm 5.21 | 34.06 \pm 26.1 | 38.30 \pm 11.5 | 92.50 \pm 3.28 |
| OA | 27.64 \pm 4.44 | 16.93 \pm 1.98 | 23.17 \pm 4.11 | 13.38 \pm 3.03 | 61.60 \pm 1.76 | 60.11 \pm 1.62 | 66.20 \pm 1.15 | 67.96 \pm 2.07 | 70.39 \pm 2.71 | 65.24 \pm 1.63 | 41.06 \pm 9.83 | 49.68 \pm 2.10 | 93.19 \pm 2.08 |
| AA | 27.63 \pm 3.03 | 15.28 \pm 2.57 | 19.51 \pm 3.32 | 13.43 \pm 3.19 | 65.58 \pm 2.12 | 64.93 \pm 1.88 | 61.41 \pm 1.57 | 70.98 \pm 2.68 | 75.13 \pm 2.47 | 71.12 \pm 1.54 | 41.85 \pm 7.83 | 51.09 \pm 2.09 | 93.28 \pm 1.99 |
| Kappa | 9.930 \pm 5.12 | 26.73 \pm 4.70 | 21.04 \pm 4.45 | 30.22 \pm 4.27 | 43.83 \pm 2.75 | 42.16 \pm 2.21 | 36.81 \pm 1.85 | 52.73 \pm 3.13 | 56.93 \pm 3.78 | 49.86 \pm 2.30 | 14.16 \pm 12.5 | 24.76 \pm 2.97 | 89.64 \pm 3.15 |

Table 5: Classification results (%) of different methods on the HU→TR dataset combination.

| Class No. | MFT | MsFE-IFN | CMFAEN | DKDMN | SDENet | LLURNet | FDGNet | TFTNet | ADNet | ISDGS | EHSNet | LDGNet | FVMGN |
|-----------|-------------------------|------------------|------------------|-------------------------|------------------|------------------|------------------|-------------------------|------------------|-------------------------|------------------|-------------------------|-------------------------|
| 1 | --- | 0.230 \pm 0.43 | 0.190 \pm 0.32 | 0.030 \pm 0.06 | 83.49 \pm 11.3 | 79.92 \pm 8.75 | 63.67 \pm 19.7 | 78.11 \pm 6.53 | 90.81 \pm 5.55 | 52.78 \pm 10.6 | 88.37 \pm 10.3 | 93.92 \pm 5.55 | 90.48 \pm 6.43 |
| 2 | 92.32 \pm 4.79 | 34.64 \pm 32.8 | 51.39 \pm 34.8 | 90.77 \pm 2.95 | 1.160 \pm 1.30 | 3.340 \pm 3.45 | 4.620 \pm 4.03 | 4.200 \pm 1.74 | 3.450 \pm 2.47 | 2.530 \pm 2.67 | 18.43 \pm 17.1 | 19.47 \pm 19.9 | 55.47 \pm 22.7 |
| 3 | 21.69 \pm 6.75 | 70.52 \pm 14.8 | 84.44 \pm 12.5 | 53.66 \pm 14.9 | 98.24 \pm 1.76 | 94.31 \pm 4.96 | 97.54 \pm 3.81 | 98.98 \pm 0.81 | 97.79 \pm 2.80 | 99.05 \pm 1.91 | 73.93 \pm 29.6 | 75.63 \pm 23.2 | 88.90 \pm 13.0 |
| OA | 23.32 \pm 0.86 | 20.66 \pm 4.94 | 26.75 \pm 6.53 | 29.05 \pm 2.64 | 69.11 \pm 6.76 | 66.68 \pm 5.34 | 57.80 \pm 12.1 | 66.66 \pm 3.95 | 73.90 \pm 3.45 | 51.12 \pm 6.21 | 71.01 \pm 7.27 | 74.88 \pm 4.94 | 82.87 \pm 3.77 |
| AA | 38.00 \pm 1.51 | 35.13 \pm 7.72 | 45.34 \pm 10.3 | 48.15 \pm 4.69 | 60.96 \pm 3.72 | 59.19 \pm 3.38 | 55.28 \pm 7.05 | 60.43 \pm 2.26 | 64.02 \pm 2.12 | 51.45 \pm 3.34 | 60.24 \pm 8.56 | 63.01 \pm 5.93 | 78.29 \pm 5.23 |
| Kappa | 1.530 \pm 1.53 | 4.120 \pm 8.09 | 6.300 \pm 9.89 | 10.10 \pm 3.58 | 49.01 \pm 8.64 | 45.42 \pm 7.05 | 36.06 \pm 13.9 | 45.81 \pm 4.97 | 55.49 \pm 4.88 | 28.18 \pm 6.35 | 47.78 \pm 13.9 | 55.55 \pm 8.48 | 69.93 \pm 6.40 |

Table 6: Classification results (%) of different methods on the HU→MU dataset combination.

| Class No. | MFT | MsFE-IFN | CMFAEN | DKDMN | SDENet | LLURNet | FDGNet | TFTNet | ADNet | ISDGS | EHSNet | LDGNet | FVMGN |
|-----------|-------------------------|------------------|-------------------------|------------------|------------------|------------------|-------------------------|------------------|-------------------|------------------|-------------------------|-------------------------|-------------------------|
| 1 | 0.230 \pm 0.18 | 71.65 \pm 24.2 | 0.260 \pm 0.47 | 45.07 \pm 39.3 | 83.00 \pm 11.1 | 55.87 \pm 24.0 | 7.640 \pm 22.9 | 38.55 \pm 5.58 | 62.96 \pm 33.61 | 41.11 \pm 31.2 | 96.02 \pm 8.43 | 87.06 \pm 4.47 | 88.49 \pm 5.55 |
| 2 | 94.80 \pm 1.47 | 0.220 \pm 0.62 | 99.28 \pm 1.55 | 59.92 \pm 37.6 | 22.35 \pm 26.6 | 10.15 \pm 18.4 | 20.22 \pm 31.1 | 17.66 \pm 15.2 | 5.36 \pm 10.50 | 30.56 \pm 19.7 | 7.510 \pm 21.0 | 80.23 \pm 12.6 | 77.32 \pm 13.4 |
| 3 | 0.830 \pm 0.77 | 12.76 \pm 15.3 | 0.350 \pm 0.85 | 36.23 \pm 42.1 | 27.74 \pm 18.6 | 61.97 \pm 25.3 | 78.89 \pm 34.1 | 73.11 \pm 14.0 | 57.75 \pm 31.3 | 54.02 \pm 20.1 | 9.480 \pm 24.1 | 55.84 \pm 15.8 | 64.74 \pm 24.0 |
| OA | 17.82 \pm 0.32 | 48.38 \pm 13.9 | 18.60 \pm 0.14 | 46.31 \pm 14.1 | 62.25 \pm 2.61 | 48.47 \pm 13.1 | 22.26 \pm 13.7 | 40.65 \pm 2.98 | 51.41 \pm 18.0 | 41.39 \pm 16.9 | 64.73 \pm 3.46 | 80.41 \pm 3.02 | 82.33 \pm 3.75 |
| AA | 31.95 \pm 0.62 | 28.21 \pm 5.86 | 33.29 \pm 0.14 | 47.07 \pm 13.0 | 44.36 \pm 5.13 | 42.66 \pm 2.26 | 35.58 \pm 5.53 | 43.11 \pm 2.49 | 42.03 \pm 8.23 | 41.90 \pm 5.58 | 37.67 \pm 12.4 | 74.38 \pm 4.01 | 76.88 \pm 7.05 |
| Kappa | 4.130 \pm 1.47 | 11.79 \pm 11.8 | 0.070 \pm 0.15 | 10.37 \pm 14.8 | 24.74 \pm 6.89 | 17.80 \pm 11.0 | 3.610 \pm 9.93 | 14.48 \pm 2.40 | 16.26 \pm 16.3 | 14.53 \pm 11.0 | 5.610 \pm 16.14 | 64.14 \pm 4.86 | 66.44 \pm 7.54 |

Table 7: Classification results (%) of different methods on the TR→HU dataset combination.

| Class No. | MFT | MsFE-IFN | CMFAEN | DKDMN | SDENet | LLURNet | FDGNet | TFTNet | ADNet | ISDGS | EHSNet | LDGNet | FVMGN |
|-----------|-------------------|------------------|------------------|------------------|------------------|------------------|------------------|-------------------------|------------------|-------------------------|------------------|-------------------------|-------------------------|
| 1 | 42.76 \pm 29.9 | 12.68 \pm 15.8 | 7.340 \pm 4.17 | 44.55 \pm 22.7 | 92.31 \pm 5.92 | 93.55 \pm 2.91 | 94.54 \pm 3.24 | 99.07 \pm 0.95 | 95.16 \pm 3.20 | 98.97 \pm 1.14 | 40.70 \pm 30.5 | 79.07 \pm 16.1 | 90.04 \pm 6.46 |
| 2 | 4.310 \pm 12.9 | 3.840 \pm 5.31 | --- | 26.83 \pm 18.1 | 45.68 \pm 11.8 | 50.15 \pm 9.06 | 49.48 \pm 8.68 | 80.18 \pm 4.97 | 67.40 \pm 4.15 | 74.60 \pm 4.12 | 51.52 \pm 31.6 | 62.30 \pm 20.0 | 98.72 \pm 1.85 |
| 3 | 36.20 \pm 18.7 | 15.02 \pm 10.3 | 50.88 \pm 5.04 | 9.990 \pm 9.24 | 32.42 \pm 6.37 | 39.63 \pm 4.09 | 33.79 \pm 5.36 | 61.39 \pm 4.72 | 52.95 \pm 5.75 | 53.01 \pm 7.63 | 40.54 \pm 25.8 | 63.52 \pm 12.8 | 82.64 \pm 6.30 |
| OA | 28.94 \pm 9.62 | 11.11 \pm 5.29 | 23.36 \pm 2.53 | 25.06 \pm 5.91 | 53.84 \pm 2.39 | 58.51 \pm 1.57 | 56.17 \pm 2.02 | 77.89 \pm 1.82 | 69.53 \pm 3.44 | 72.75 \pm 2.93 | 43.76 \pm 8.11 | 67.74 \pm 8.77 | 89.46 \pm 3.35 |
| AA | 27.76 \pm 9.43 | 10.51 \pm 5.24 | 19.41 \pm 2.27 | 27.12 \pm 0.86 | 56.80 \pm 2.74 | 61.11 \pm 2.24 | 59.27 \pm 2.03 | 80.21 \pm 1.70 | 71.84 \pm 3.16 | 75.53 \pm 2.34 | 44.25 \pm 8.73 | 68.30 \pm 9.61 | 90.47 \pm 3.09 |
| Kappa | 6.450 \pm 13.65 | 32.19 \pm 8.27 | 20.27 \pm 3.05 | 9.390 \pm 9.50 | 31.64 \pm 3.69 | 38.26 \pm 2.65 | 35.32 \pm 2.82 | 67.10 \pm 2.62 | 54.61 \pm 4.88 | 59.62 \pm 4.12 | 16.59 \pm 12.0 | 51.32 \pm 13.5 | 84.12 \pm 4.96 |

Table 8: Classification results (%) of different methods on the TR→MU dataset combination.

| Class No. | MFT | MsFE-IFN | CMFAEN | DKDMN | SDENet | LLURNet | FDGNet | TFTNet | ADNet | ISDGS</th |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |

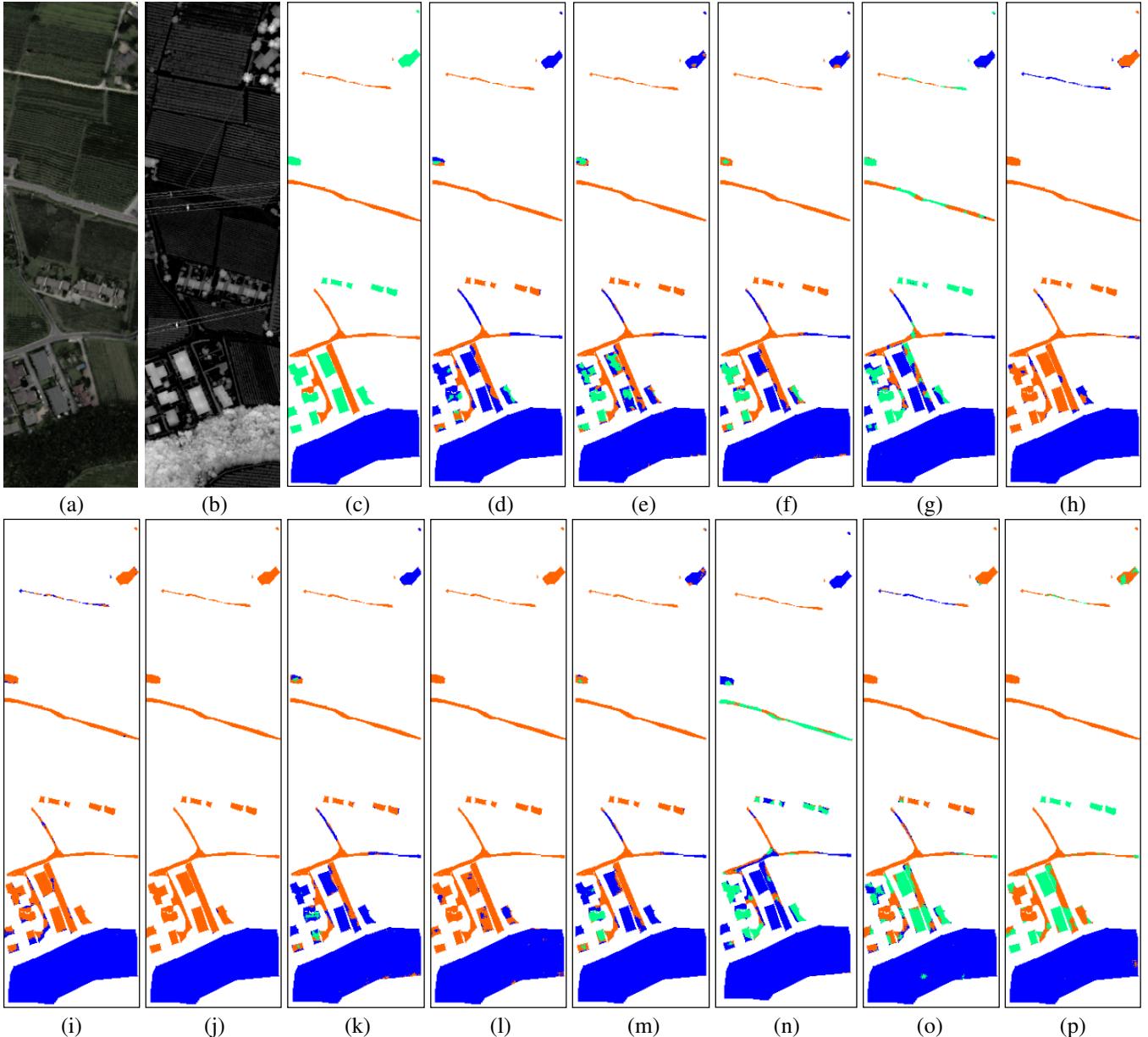


Figure 4: Classification maps (Blue/orange/green regions: Trees/Roads/Buildings) on the MU→TR dataset combination. (a) HSI. (b) LiDAR image. (c) Ground truth. (d) MFT. (e) MsFE. (f) CMFAEN. (g) DKDMN. (h) SDENet. (i) LLURNet. (j) FDGNet. (k) TFTNet. (l) ADNet. (m) ISDGS. (n) EHSNet. (o) LDGNet. (p) FVMGN.

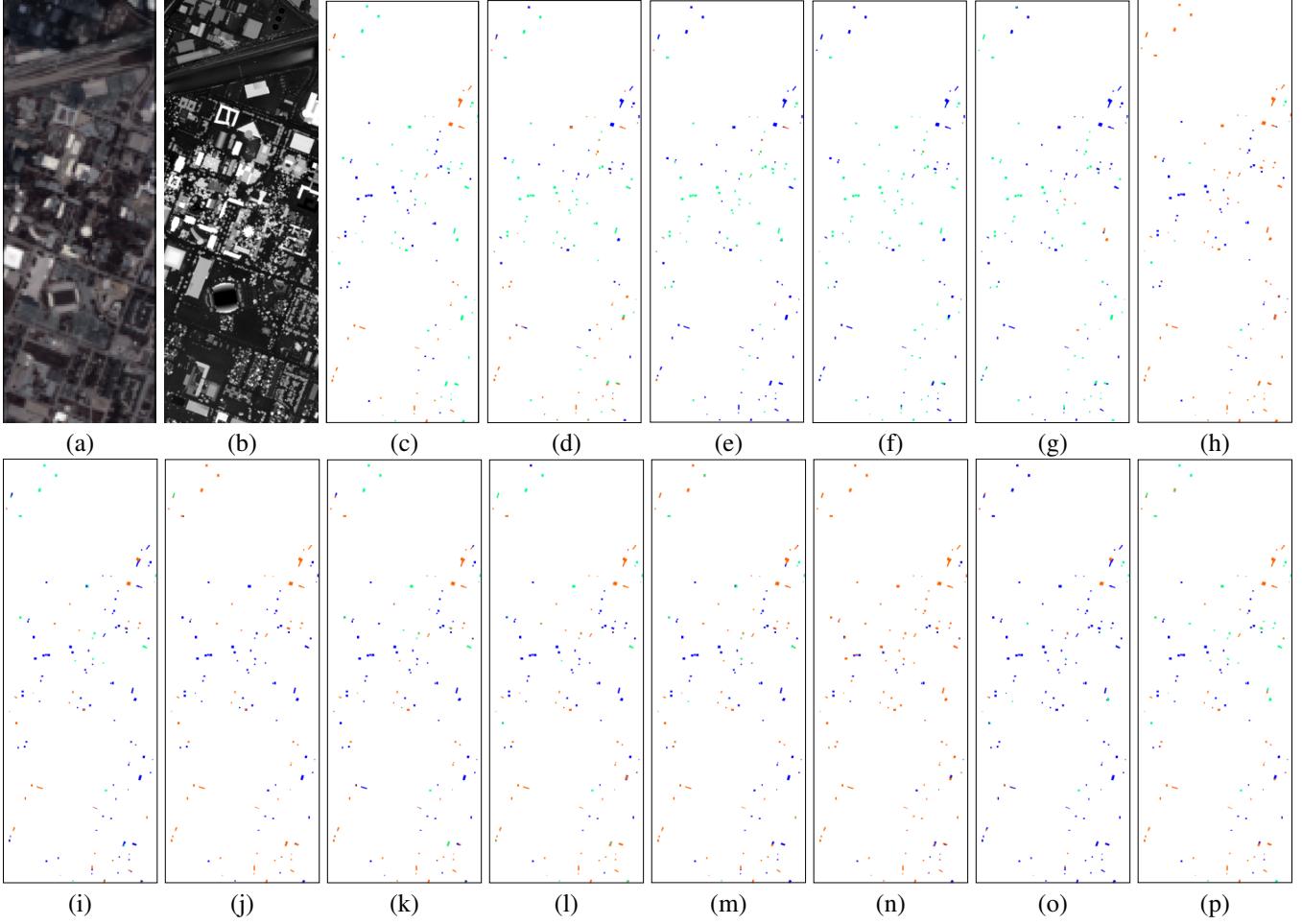


Figure 5: Classification maps (Blue/orange/green regions: Trees/Roads/Buildings) on the MU→HU dataset combination. (a) HSI. (b) LiDAR image. (c) Ground truth. (d) MFT. (e) MsFE. (f) CMFAEN. (g) DKDMN. (h) SDENet. (i) LLURNet. (j) FDGNet. (k) TFTNet. (l) ADNet. (m) ISDGS. (n) EHSNet. (o) LDGNet. (p) FVMGN.

Table 9: Computational complexity of different methods on six dataset combinations.

| SD→TD | Metric | MFT | MsFE-IFN | CMFAEN | DKDMN | SDENet | LLURNet | FDGNet | TFTNet | ADNet | ISDGS | EHSNet | LDGNet | FVMGN-NW | FVMGN |
|-------|-------------------|--------|-------------|--------|--------|---------|---------|---------|----------|---------|---------|----------|----------|----------|---------|
| MU→TR | FLOPs (G) | 0.44 | 66.23 | 3.09 | 6.26 | 1.32 | 0.33 | 1.32 | 2.67 | 0.33 | 1.32 | 8.62 | 7.14 | 2.81 | 2.84 |
| | Params (M) | 0.24 | 0.02 | 0.19 | 0.85 | 0.44 | 0.44 | 0.44 | 0.55 | 0.44 | 0.44 | 0.34 | 0.55 | 0.65 | 0.66 |
| | Training Time (s) | 345.16 | 171.79 | 475.24 | 402.73 | 6334.63 | 1411.16 | 3387.66 | 6865.30 | 1467.86 | 1832.94 | 15426.57 | 19725.71 | 5583.62 | 5650.54 |
| | Test Time (s) | 6.06 | 32.13 | 16.23 | 33.30 | 6.52 | 1.17 | 3.02 | 7.92 | 2.37 | 2.23 | 47.94 | 20.06 | 69.84 | 120.68 |
| MU→HU | FLOPs (G) | 0.44 | 66.23 | 3.09 | 6.26 | 1.32 | 0.33 | 1.32 | 2.67 | 0.33 | 1.32 | 8.62 | 7.14 | 2.81 | 2.83 |
| | Params (M) | 0.24 | 0.02 | 0.19 | 0.85 | 0.44 | 0.44 | 0.44 | 0.55 | 0.44 | 0.44 | 0.34 | 0.55 | 0.65 | 0.66 |
| | Training Time (s) | 340.47 | 183.10 | 526.25 | 393.82 | 3263.65 | 1379.43 | 3393.74 | 8020.18 | 1542.69 | 2074.61 | 8129.26 | 24887.24 | 605.17 | 641.41 |
| | Test Time (s) | 22.86 | 115.34 | 63.23 | 111.44 | 0.91 | 0.67 | 0.92 | 2.51 | 1.43 | 1.50 | 1.27 | 3.01 | 4.36 | 8.79 |
| TR→MU | FLOPs (G) | 0.44 | 66.23 | 3.09 | 6.26 | 1.32 | 0.33 | 1.32 | 2.67 | 0.33 | 1.32 | 8.62 | 6.98 | 2.81 | 2.84 |
| | Params (M) | 0.24 | 0.02 | 0.19 | 0.85 | 0.44 | 0.44 | 0.44 | 0.55 | 0.44 | 0.44 | 0.34 | 0.55 | 0.65 | 0.66 |
| | Training Time (s) | 138.24 | 72.47 | 206.19 | 177.90 | 6073.06 | 3163.14 | 7854.21 | 14877.36 | 835.30 | 3799.24 | 3392.07 | 9296.51 | 1207.13 | 1277.57 |
| | Test Time (s) | 4.19 | 21.22 | 11.49 | 24.16 | 7.23 | 3.02 | 7.24 | 18.01 | 6.09 | 5.21 | 24.37 | 58.51 | 80.65 | 144.57 |
| TR→HU | FLOPs (G) | 0.44 | 66.23 | 3.09 | 6.26 | 1.32 | 0.33 | 1.32 | 2.67 | 0.33 | 1.32 | 8.62 | 6.98 | 2.81 | 2.84 |
| | Params (M) | 0.24 | 0.02 | 0.19 | 0.85 | 0.44 | 0.44 | 0.44 | 0.55 | 0.44 | 0.44 | 0.34 | 0.55 | 0.65 | 0.66 |
| | Training Time (s) | 168.60 | 79.06 | 232.36 | 172.51 | 1093.65 | 554.18 | 1378.67 | 4976.93 | 649.74 | 848.53 | 3368.38 | 9157.70 | 1143.52 | 1314.23 |
| | Test Time (s) | 25.06 | 112.51 | 64.08 | 111.90 | 0.90 | 0.64 | 0.87 | 2.15 | 1.41 | 1.48 | 1.27 | 3.01 | 5.62 | 9.41 |
| HU→MU | FLOPs (G) | 0.44 | 66.23 | 3.09 | 6.26 | 1.32 | 0.33 | 1.32 | 2.67 | 0.33 | 1.17 | 8.62 | 6.82 | 2.81 | 2.84 |
| | Params (M) | 0.24 | 0.02 | 0.19 | 0.85 | 0.44 | 0.44 | 0.44 | 0.55 | 0.44 | 0.44 | 0.34 | 0.55 | 0.65 | 0.66 |
| | Training Time (s) | 23.27 | 12.18 | 35.48 | 25.96 | 801.25 | 366.68 | 839.26 | 2416.54 | 474.61 | 145.02 | 502.13 | 1343.84 | 301.27 | 363.44 |
| | Test Time (s) | 4.48 | 23.11 | 13.10 | 22.81 | 6.21 | 2.06 | 6.26 | 21.96 | 5.13 | 5.78 | 24.67 | 54.09 | 76.92 | 140.98 |
| HU→TR | FLOPs (G) | 0.44 | 66.23 | 1.94 | 6.26 | 1.32 | 0.33 | 1.32 | 2.67 | 0.33 | 1.32 | 8.62 | 6.98 | 2.81 | 2.84 |
| | Params (M) | 0.24 | 0.02 | 0.37 | 0.85 | 0.44 | 0.44 | 0.44 | 0.55 | 0.44 | 0.44 | 0.34 | 0.55 | 0.65 | 0.66 |
| | Training Time (s) | 24.57 | 12.28 | 9.47 | 26.19 | 808.52 | 362.88 | 846.25 | 2503.76 | 114.52 | 265.25 | 1105.90 | 1366.92 | 180.93 | 221.12 |
| | Test Time (s) | 6.81 | 32.59 | 5.81 | 31.71 | 2.49 | 1.05 | 2.54 | 9.88 | 2.30 | 2.20 | 59.26 | 24.21 | 37.41 | 63.32 |

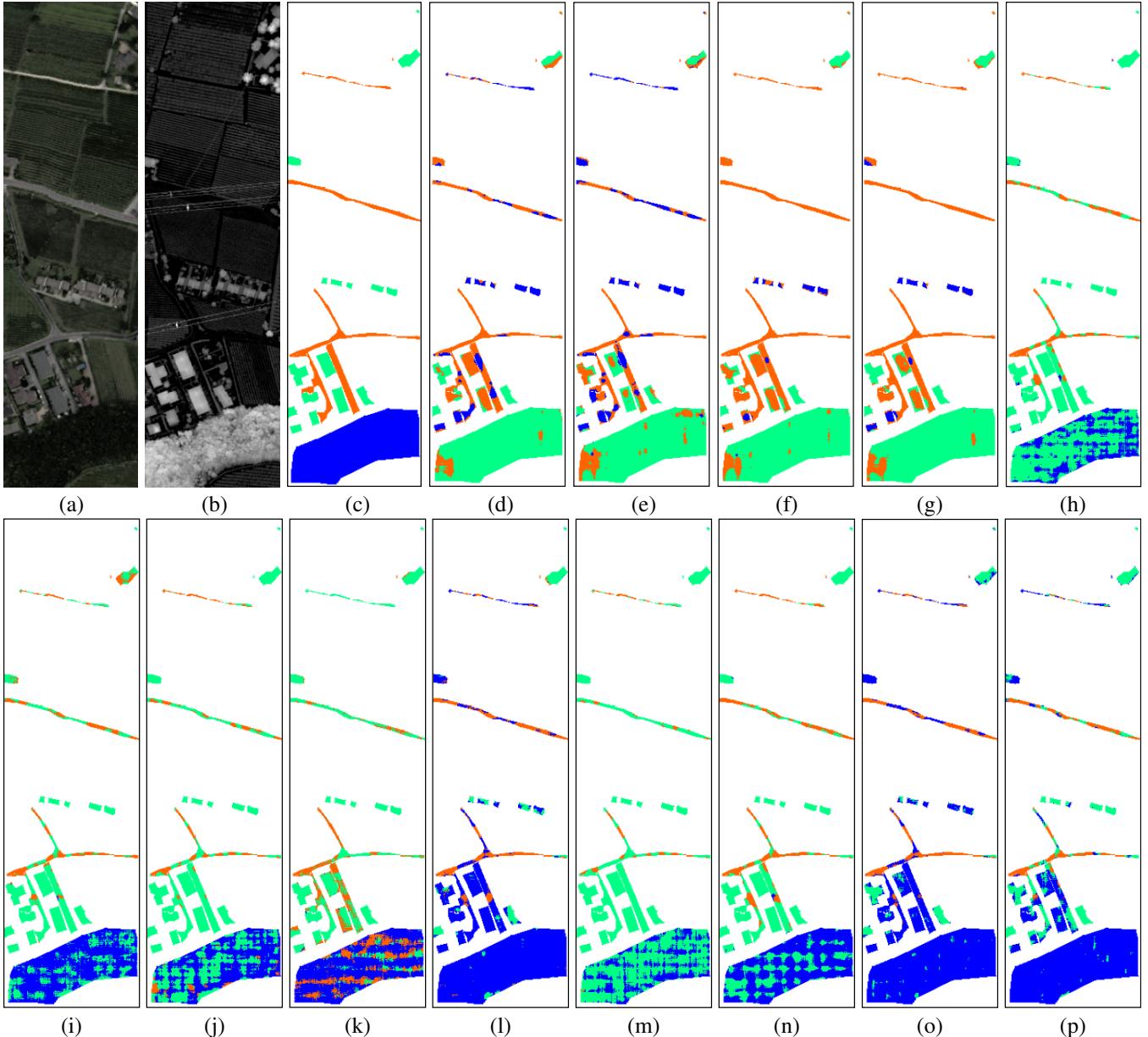


Figure 6: Classification maps (Blue/orange/green regions: Trees/Roads/Buildings) on the HU→TR dataset combination. (a) HSI. (b) LiDAR image. (c) Ground truth. (d) MFT. (e) MsFE. (f) CMFAEN. (g) DKDMN. (h) SDENet. (i) LLURNet. (j) FDGNet. (k) TFTNet. (l) ADNet. (m) ISDGS. (n) EHSNet. (o) LDGNet. (p) FVMGN.

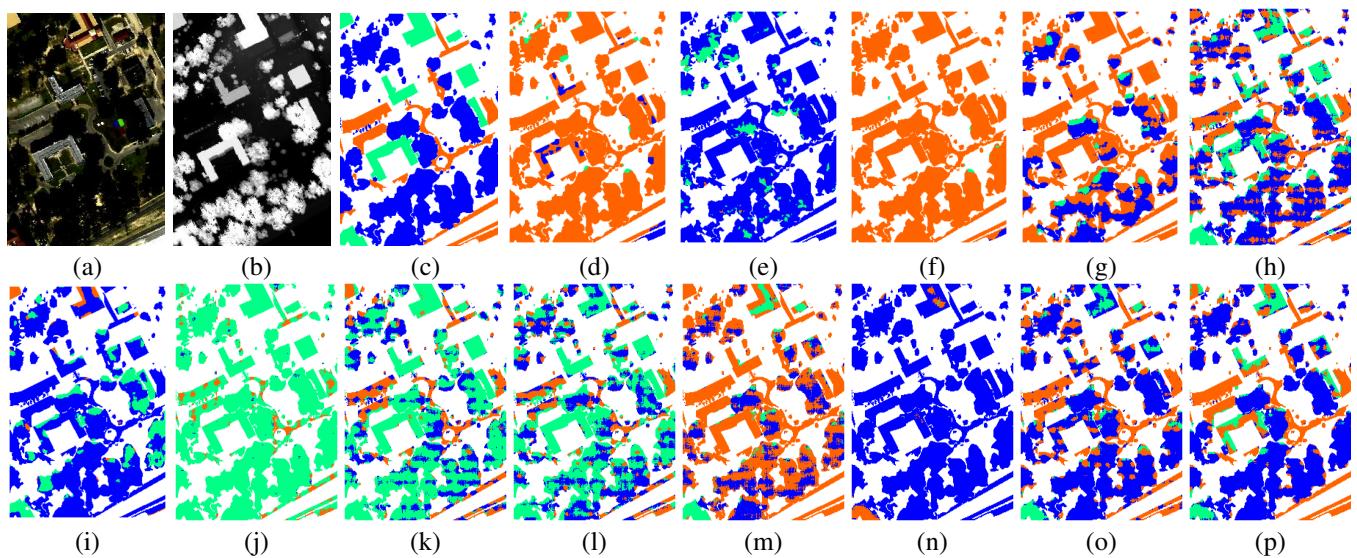


Figure 7: Classification maps (Blue/orange/green regions: Trees/Roads/Buildings) on the HU→MU dataset combination. (a) HSI. (b) LiDAR image. (c) Ground truth. (d) MFT. (e) MsFE. (f) CFEN. (g) DKDN. (h) SDEN. (i) LLURN. (j) FDGN. (k) TFTN. (l) ADNet. (m) ISDGGS. (n) EHSN. (o) LDGN. (p) FVMGN.

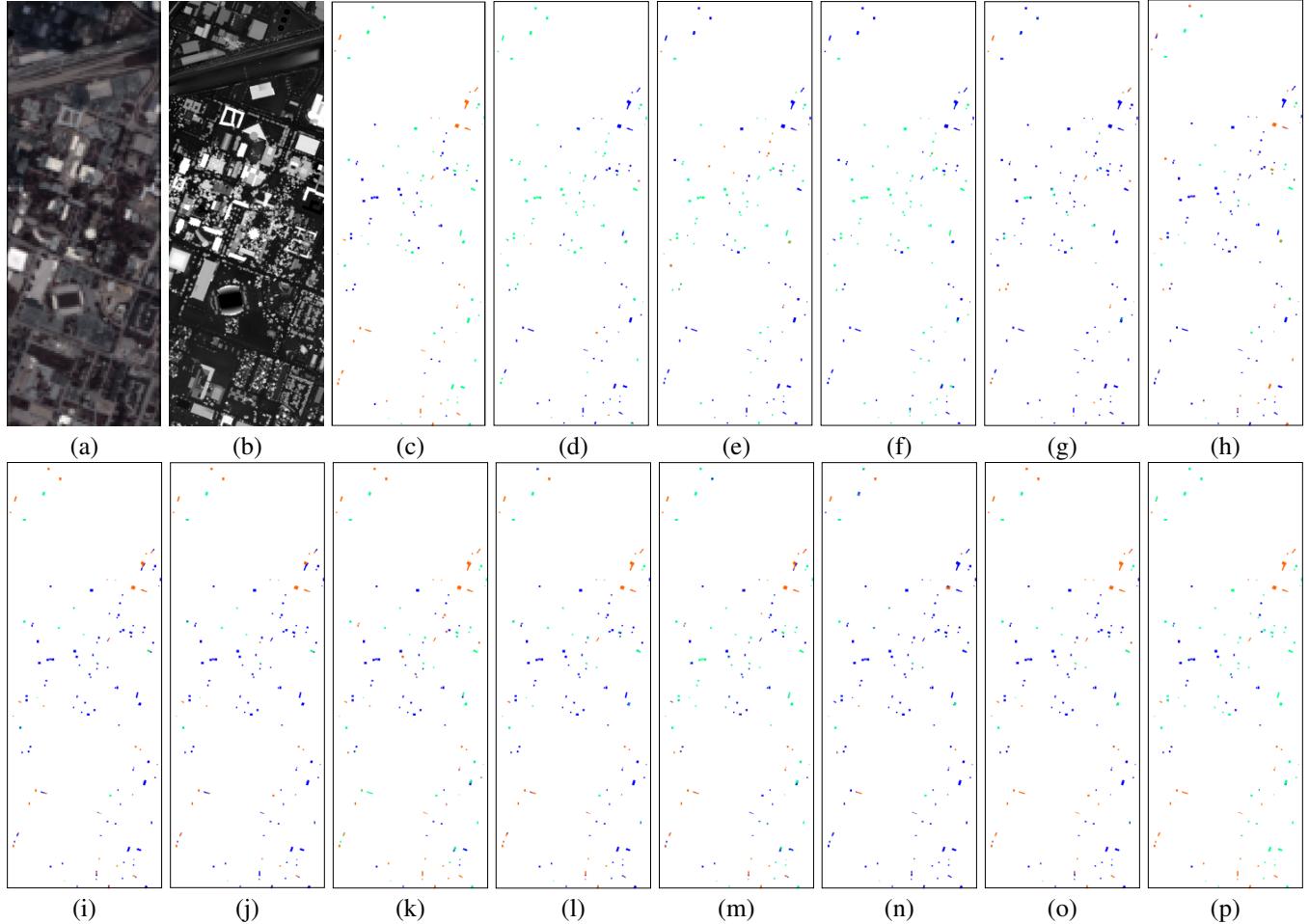


Figure 8: Classification maps (Blue/orange/green regions: Trees/Roads/Buildings) on the TR→HU dataset combination. (a) HSI. (b) LiDAR image. (c) Ground truth. (d) MFT. (e) MsFE. (f) CMFAEN. (g) DKDMN. (h) SDENet. (i) LLURNet. (j) FDGNet. (k) TFTNet. (l) ADNet. (m) ISDGs. (n) EHSNet. (o) LDGNet. (p) FVMGN.

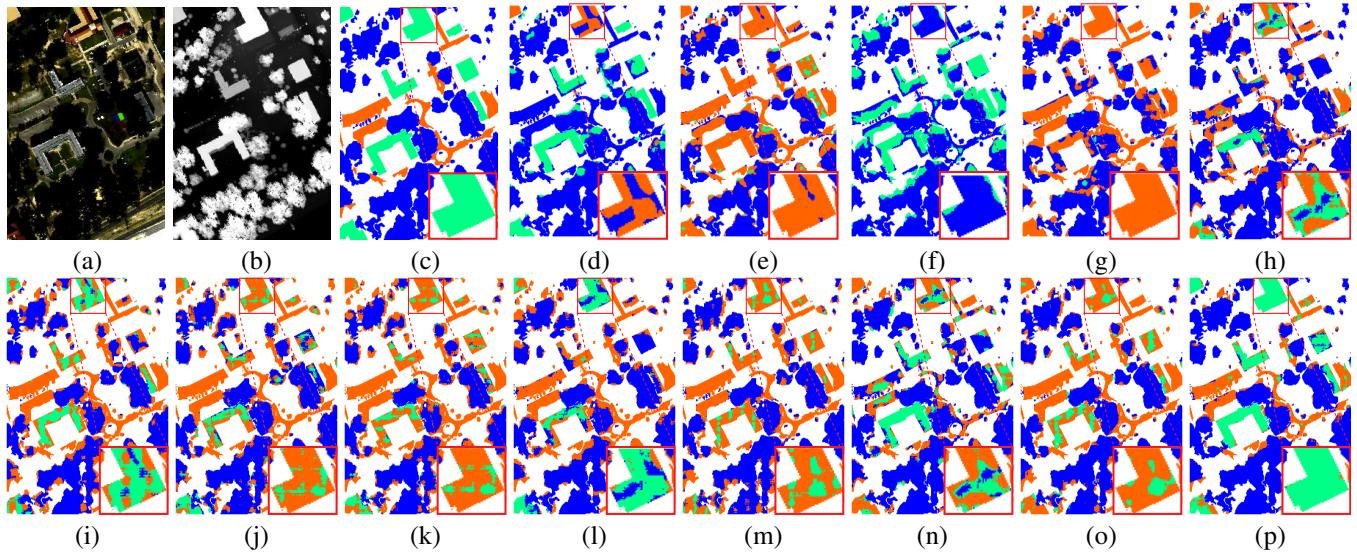


Figure 9: Classification maps (Blue/orange/green regions: Trees/Roads/Buildings) on the TR→MU dataset combination. (a) HSI. (b) LiDAR image. (c) Ground truth. (d) MFT. (e) MsFE. (f) CFEN. (g) DKDN. (h) SDEN. (i) LLURN. (j) FDGN. (k) TFTN. (l) ADNet. (m) ISDGGS. (n) EHSN. (o) LDGN. (p) FVMGN.