BAND-SELECTION STABILITY AND SEMANTIC SEGMENTATION PERFORMANCE: A STUDY ON HYPERSPECTRAL CITY

Jiarong Li [1,†], Imad Ali Shah [1,†], Enda Ward [2], Martin Glavin [1], Edward Jones [1] and Brian Deegan [1]

[1] School of Engineering and Ryan Institute, University of Galway, Ireland
[2] Valeo Vision Systems, Tuam, Ireland

ABSTRACT

Resource constraints make high-dimensional hyperspectral imaging challenging in autonomous perception, motivating the use of band selection methods. However, the sensitivity of band-selection methods to sampled data and their relationship to semantic segmentation models (SSMs) remain underexplored. This study evaluates six band selection methods on ten independently sampled, class-balanced region-of-interest (ROI) sets, yielding 60 top-25 band subsets from the Hyperspectral City V2 (128 bands: 450-950nm) dataset. Top-K bands ($K \in \{3, 5, ...13\}$) from the first three ROI sets are evaluated with three SSMs against the corresponding 128-band baseline. Experiments show that intra-method stability is method-dependent: Sim-LP shows the highest stability (pairwise Jaccard similarity) and, together with JMIM+CSNR, yields the best segmentation results. Top-K based SSMs remain competitive with baselines, with gains of up to 2.01 mIoU and 1.72 mF1 points, and 18–22x faster CPU inference for $K = 9$. However, performance does not improve monotonically with K, and stability shows no consistent association with SSM performance. These findings suggest that intra-method stability is informative but an unreliable indicator of downstream segmentation performance, highlighting the need to evaluate band-selection methods across repeated samples, subset sizes, and SSMs.

Index Terms— Autonomous driving, band selection, CMIM, Hyperspectral Segmentation, JMIM, mRMR.

1. INTRODUCTION

Hyperspectral imaging (HSI) captures dense spectral information that can provide material information beyond conventional RGB imaging. This capability is relevant to semantic segmentation in urban-driving scenes, where different materials may exhibit similar visible appearance [1]. However, HSI is high-dimensional, computationally demanding, and often contains spectral redundancy [2]. Band selection is therefore used to retain a small but informative set of wavelengths.

Band-selection criteria are typically estimated from sampled pixels or an image region of interest (ROI), and depend on the class distribution represented in the sample. This issue is particularly important in automotive segmentation datasets with severe class imbalance [3]. In the Hyperspectral City V2 (H-City) [4] dataset used in this study, Road accounts for 22.43% of labeled pixels versus 0.02% for Train. Band-selection criteria estimated directly on such data may therefore bias selection toward frequent classes. Moreover, a stable method that reliably selects the same bands when the sampled regions change does not necessarily imply that those bands are optimal for segmentation performance. Conversely, different subsets may provide similar performance due to inter-band spectral redundancy. This study therefore addresses two questions: how strongly do common band-selection methods vary under repeated class-balanced ROI sampling, and does greater intra-method stability correspond to better semantic segmentation models (SSMs) performance?

Six band-selection methods are applied to ten class-balanced ROI sets, producing 60 top-25 subsets from the H-City dataset. Top-K bands, where $K \in \{3, 5, ...13\}$, from three ROI sets are then evaluated with three SSMs against 128-band baselines. A frequency-consensus ranking derived from comparatively stable methods is included as an exploratory aggregate baseline. The study consequently contributes: (1) a repeated-sampling analysis of band-selection stability and spectral recurrence; (2) evaluation across K-subsets and three SSMs; and (3) analysis of the relationship between intra-method stability and downstream SSM performance, including the accuracy-inference-time trade-off.

2. RELATED WORK

2.1. HSI and Band-Selection in Autonomous Driving

HSI is increasingly investigated as a source of complementary spectral information for urban-scene understanding [1]. Huang et al. [5] used hyperspectral information to refine coarse urban-scene annotations under weak supervision. Subsequent works evaluated different baseline SSM performance [3] on the four publicly available annotated automotive HSI datasets considered in that study.

Band-selection has also been investigated for vulnerable-

This work was supported, in part, by Taighde Éireann – Research Ireland grants 13/RC/2094_P2 and 18/SP/5942, and Valeo Vision Systems.
† Equal Contribution and correspondence (i.shah2@universityofgalway.ie)

road-user perception [6] and compared with RGB [7]. In contrast to independently applied filter methods, embedded approaches learn spectral gates or selectors jointly with the downstream model [8, 9]. Embedded methods can optimize bands for a particular architecture and objective, but the learned selection may also depend on model initialization and optimization. They are therefore related to, but distinct from, the sampling-stability problem considered in this work.

2.2. Hyperspectral Band-Selection

Band-selection methods differ in class supervision and handling of inter-band redundancy. Minimum Redundancy Maximum Relevance (mRMR) [10] balances class relevance and inter-band redundancy, whereas Conditional Mutual Information Maximization (CMIM) [11] and Joint Mutual Information Maximization (JMIM) [12] rank candidates using conditional and joint mutual-information criteria, respectively. These supervised criteria can represent nonlinear statistical dependence, but their rankings depend on the labeled sample and mutual information estimation. Similarity-based Linear Prediction (Sim-LP) [13] is an unsupervised alternative that favors bands poorly reconstructed from selected bands, discouraging redundancy. Li et al. [6] combined JMIM with patch-based contrast signal-to-noise ratio (CSNR) and correlation-based duplicate control. JMIM+CSNR combines JMIM and CSNR, while JMIM+CSNR+Corr additionally applies Pearson-correlation-based duplicate removal.

2.3. Band-Selection Stability vs SSM Performance

Band-selection stability measures the degree of agreement among the selected variables when the data sample is perturbed [14]. It is commonly estimated by repeating the selection method and comparing the resulting subsets: Pairwise Jaccard similarity [15] provides a direct overlap measure for fixed-size subsets. Liu et al. [16] constructed hyperspectral perturbation datasets with controlled sample overlap and examined band-selection stability alongside classification accuracy. However, the relationship between intra-method stability under repeated class-balanced ROI sampling and downstream SSM performance remains underexplored, especially for the automotive HSI domain.

3. EXPERIMENTAL PROTOCOL

3.1. Dataset and Class-Balanced ROI Sampling

The experiments used H-City, whose 128-band, 450-950nm range and pixel-level annotations for 19 urban classes provide a comparatively large spectral search space for intra-method band-selection stability analysis. Due to the different annotation granularity level in the original H-City's validation and test sets, ROI sampling used the training split: SSMs were trained on the training and evaluated on the validation split. The class-balanced ROI-sampling procedure was repeated using ten different sampling seeds. Each sampling run produced one ROI subset, sampled under the same class-balancing rule.

3.2. Intra-Method Stability Characterization

Six band-selection methods were applied independently to each ROI set: JMIM, JMIM+CSNR, JMIM+CSNR+Corr, CMIM, mRMR, and Sim-LP. For method m and ROI set r, the top-L ($L = 25$) bands in the resulting ranking formed the subset $S^L_{m,r}$. Applying the six methods to the ten ROI sets produced 60 top-L subsets. Band b's recurrence under method m was defined as $R_m(b) = \sum_{r=1}^{10} \mathcal{I}\left(b \in S^L_{m,r}\right)$, where $\mathcal{I}(\cdot)$ is the indicator function, characterizing how consistently each band was selected by m across the ten ROI sets. Intra-method stability was quantified using pairwise Jaccard similarity among the ten top-L subsets produced by each method m. For ROI sets r_i and r_j:

$$J_m(r_i, r_j) = \frac{\left|S^{25}_{m,r_i} \cap S^{25}_{m,r_j}\right|}{\left|S^{25}_{m,r_i} \cup S^{25}_{m,r_j}\right|}$$

Each method yield $\binom{10}{2} = 45$ pairwise Jaccard similarities. The median and interquartile range summarized the central value and spread of this distribution. For reference, the expected Jaccard similarity of two random 25-of-128 subsets is approximately 0.11.

3.3. Frequency-Based Consensus Construction

As a secondary aggregation experiment, a frequency consensus $\mathcal{F}(\cdot)$ was constructed from the four most stable methods: JMIM, JMIM+CSNR, JMIM+CSNR+Corr, and Sim-LP. For each band b, its $\mathcal{F}(b) \in \{0, \ldots, 40\}$ was computed across the top-L subsets obtained from all ten ROI sets as:

$$\mathcal{F}(b) = \sum_{m \in \mathcal{M}_s} \sum_{r=1}^{10} \mathcal{I}\left(b \in S^L_{m,r}\right),$$

and the corresponding subsets: $S^K_{\text{Freq}} \in \arg\max_{\substack{S \subseteq \mathcal{B} \\ |S|=K}} \sum_{b \in S} \mathcal{F}(b)$.

Where $\mathcal{M}_s$ denotes the four retained methods, $\mathcal{B}$ is the set of all spectral bands, and $\mathcal{I}(\cdot)$ is the indicator function. Bands were then ranked in descending order of $\mathcal{F}(b)$, and the first K bands formed S^K_{Freq}, for $K \in \{3, 5, \ldots, 13\}$. These subsets were subsequently used for SSM training and evaluation.

3.4. Semantic-Segmentation Validation

For each method, rankings from the three ROI sets were truncated at K-bands for SSM evaluation. These sets were not selected based on segmentation performance and were limited to three for computational tractability, as using all ten would require 3.33× as many training runs. The subsets were evaluated with U-Net, DeepLabV3+, and SegFormer, representing convolutional encoder–decoder, atrous-convolutional, and transformer-based models, respectively. Each SSM–subset combination was trained with three training seeds. Thus, each SSM–method–K configuration comprised nine evaluations (three ROI sets × three training seeds). ROI sampling determined the band rankings, whereas the training seeds captured segmentation training variation.

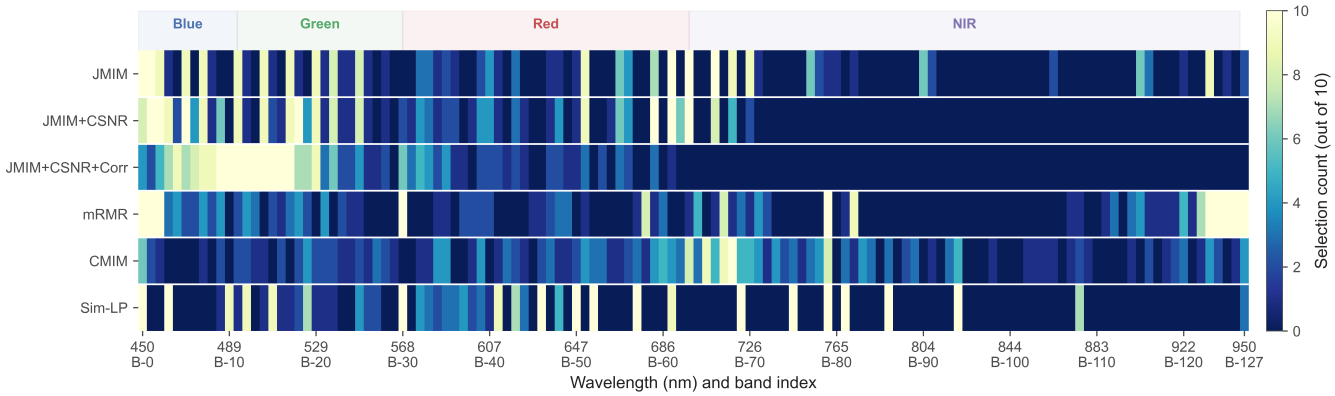


Fig. 1: Method-specific top-25 band recurrence across ten class-balanced ROI sets. Color indicates selection count.

The frequency consensus produced one ranking from all ten ROI sets, whose top-K subsets were evaluated with the same SSMs and three training seeds. The full 128-band input served as the SSM-specific baseline. SSM performance was measured using mean intersection over union (mIoU) and mean F1 score (mF1). Individual-method results are reported as mean $\pm$ standard deviation (SD) over nine evaluations, incorporating both ROI-set and SSM-training variation. Frequency-consensus and 128-band results are reported over three training seeds and therefore reflect only training variation. All configurations used the same preprocessing, training, and evaluation pipeline [1].

3.5. Method-Level Stability: Performance Comparison

For each method and subset size, the mean mIoU and mF1 obtained with each SSM were compared with the corresponding all-128-bands baseline: selected-band minus baseline, in percentage points. These differences were averaged across the three SSMs and six subset sizes for each method and metric, then compared with the method's mean over 45 pairwise Jaccard similarities. The frequency consensus was excluded because its single aggregated ranking has no intra-method stability value.

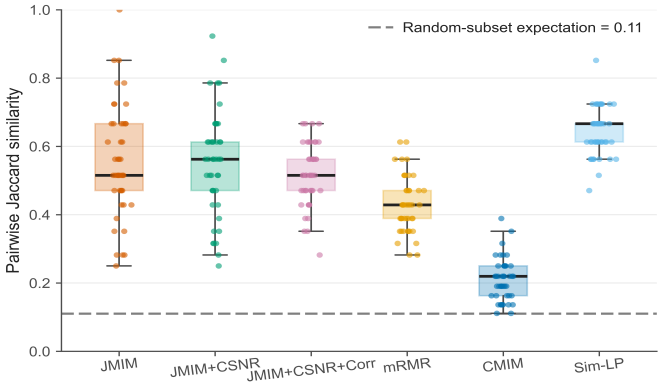


Fig. 2: Intra-method pairwise Jaccard similarities of top-25 subsets across ten ROI sets. The dashed line marks the random-subset expectation (0.11).

[1] https://github.com/imadalishah/multiseedHSI

4. RESULTS AND DISCUSSION

4.1. Intra-Method Band-Selection Stability

Figure 1 shows that the recurrence patterns of the selected bands across the spectrum differ between methods. Sim-LP selects a sparse set distributed across the visible and near-infrared (NIR) ranges, whereas CMIM's recurrent selections span much of the spectrum, with several of its highest selection counts in the NIR. mRMR concentrates recurrence near both spectral boundaries. Compared with JMIM, JMIM+CSNR shifts recurrent selections toward the visible range, while JMIM+CSNR+Corr concentrates them in the blue-green region with little NIR recurrence.

Figure 2 illustrates that the degree of agreement among the ten top-25 subsets also differs between methods. Sim-LP has the highest pairwise Jaccard similarity median and a relatively narrow interquartile range. JMIM, JMIM+CSNR, and JMIM+CSNR+Corr form a relatively high-agreement group below Sim-LP, although their distributions differ. mRMR has a lower median than this group, but remains above CMIM, which has the lowest median. The lowest CMIM values approach the random-subset reference of 0.11.

Overall, these results indicate that changing the ROI set affects the selected bands of each method differently, raising the question of whether methods with more consistent selections also achieve better segmentation performance.

4.2. Segmentation Performance of Band Subsets

Table 1 shows that Sim-LP and JMIM+CSNR achieve all six SSM-metric best results. For U-Net, JMIM+CSNR and Sim-LP at $K = 9$ improve mIoU and mF1 by 0.43 and 0.70 percentage points over the baseline, respectively. Sim-LP gives the best results for DeepLabV3+, improving mIoU by 1.02 points at $K = 3$ and mF1 by 1.28 points at $K = 9$. The largest gains occur with SegFormer: 2.01 mIoU points for JMIM+CSNR at $K = 7$ and 1.72 mF1 points for Sim-LP at $K = 5$.

$K = 9$ yields three of the six best results and the largest number of above-baseline results (25 of 36). However, performance is non-monotonic with K, and no subset size is optimal across all SSMs and metrics. The frequency con-

Table 1: Mean ± SD mIoU and mF1 (%) by SSM, method, and subset size. Individual methods use nine evaluations, while frequency-consensus and baseline entries use three. Cell color encodes the signed difference from SSM-specific 128-band baseline (green: positive; red: negative; white: parity), and bold marks each SSM-metric maximum.

Model	Set Size	CMIM		JMIM		JMIM+CSNR		JMIM+CSNR+Corr		mRMR		Sim-LP		Freq.		128-band	
		mIoU	mF1	mIoU	mF1	mIoU	mF1	mIoU	mF1	mIoU	mF1	mIoU	mF1	mIoU	mF1	mIoU	mF1
U-Net	3	70.73±1.23	79.68±1.32	68.43±0.32	77.95±0.25	68.79±0.39	78.01±0.31	67.76±1.75	77.72±1.23	68.62±0.48	77.63±0.69	69.88±1.01	79.31±0.73	69.66±1.27	78.94±0.88	70.67±0.61	79.56±0.47
	5	70.06±0.61	79.34±0.43	69.94±0.61	78.94±0.73	68.51±0.49	77.98±0.23	68.03±1.83	77.74±1.61	70.79±0.76	79.85±0.25	70.01±0.77	79.67±0.60	68.97±0.85	78.65±0.77		
	7	70.61±0.31	79.90±0.20	70.02±0.23	79.25±0.95	70.53±1.19	79.52±0.88	66.64±0.85	76.56±0.67	70.01±0.29	79.50±0.12	70.39±1.36	79.92±1.20	69.97±1.32	79.11±1.42		
	9	70.77±0.37	79.66±0.50	70.34±1.28	79.63±1.07	**71.10±1.09**	80.10±0.94	67.98±0.98	77.84±0.79	70.91±1.04	79.58±1.09	71.05±0.32	**80.26±0.50**	69.96±0.17	79.01±0.15		
	11	70.02±1.75	79.34±1.10	68.67±1.23	78.33±1.09	70.34±0.43	79.81±0.55	68.15±0.38	77.96±0.72	69.89±0.27	79.13±0.52	69.09±1.23	78.88±1.05	68.82±1.10	78.11±0.92		
	13	70.28±0.62	79.10±0.48	69.66±0.37	78.91±0.09	69.82±0.10	78.96±0.36	67.00±0.44	76.97±0.41	70.47±0.91	79.70±0.69	70.62±0.73	79.60±0.24	68.60±0.27	78.11±0.17		
DLV3+	3	70.89±0.98	79.90±0.82	69.13±0.87	78.38±0.65	69.67±0.66	79.05±0.39	68.02±1.07	77.54±0.95	68.33±0.42	77.81±0.66	**71.09±0.67**	80.39±0.83	69.47±0.66	78.85±0.71	70.07±0.66	79.34±0.24
	5	70.41±1.18	79.77±1.05	70.53±1.00	79.51±0.88	68.92±0.30	78.44±0.34	67.75±1.04	77.26±1.27	70.14±1.34	79.67±1.17	69.99±0.45	79.19±0.63	68.87±0.08	78.07±0.16		
	7	69.90±0.72	79.14±0.44	70.15±0.38	79.43±0.33	69.56±0.44	79.25±0.60	67.45±0.33	77.30±0.28	68.67±1.02	78.31±0.87	70.14±0.61	79.28±0.36	68.92±0.60	78.27±0.26		
	9	70.79±0.46	79.73±0.55	70.10±0.61	79.64±0.43	70.44±0.54	79.67±0.39	67.33±0.34	77.13±0.77	69.70±0.85	79.16±0.89	71.02±0.59	**80.62±0.48**	69.24±1.00	78.77±1.06		
	11	70.62±0.72	79.39±1.09	70.42±0.85	79.41±0.55	69.88±0.49	79.36±0.21	68.66±0.30	78.17±0.37	70.27±0.48	79.53±0.10	70.16±0.11	79.52±0.12	68.86±0.80	78.01±1.02		
	13	70.13±1.00	79.39±0.69	69.73±0.61	78.98±0.59	69.83±1.64	79.26±1.23	68.43±0.52	77.97±0.50	69.35±0.66	78.68±0.62	70.87±1.02	80.17±1.01	68.56±0.30	77.94±0.41		
SegFormer	3	67.99±0.81	77.50±0.50	65.84±0.93	75.86±1.02	67.08±0.87	76.59±0.37	65.10±0.81	75.33±1.04	65.64±1.34	75.43±0.56	68.76±0.79	78.21±0.92	66.24±2.03	75.91±1.78	67.50±1.05	77.06±0.60
	5	69.16±0.84	78.39±0.62	68.47±1.10	78.13±0.72	68.75±0.49	78.54±0.54	65.34±2.10	75.31±1.56	67.14±0.29	77.38±0.18	68.97±0.47	**78.78±0.52**	67.00±0.38	76.46±0.22		
	7	67.68±0.40	77.64±0.44	68.57±0.25	77.92±0.17	**69.51±0.27**	78.65±0.62	66.21±0.55	76.35±0.71	67.08±0.46	76.54±0.25	67.80±1.27	77.61±0.66	67.55±0.31	76.92±0.23		
	9	68.97±0.38	78.61±0.55	68.56±0.36	77.91±0.18	68.61±0.86	78.32±0.68	65.06±0.91	75.10±0.64	66.88±2.04	76.55±1.25	68.24±0.67	77.86±0.19	66.07±0.85	75.98±0.82		
	11	67.22±1.63	77.49±1.34	68.11±1.39	77.72±0.88	68.69±1.17	78.07±1.24	66.36±1.62	76.30±1.05	67.68±1.88	77.47±1.85	68.91±0.61	78.52±0.62	66.62±0.18	75.78±0.40		
	13	67.61±0.50	77.33±0.44	68.39±0.87	78.04±0.97	68.68±0.81	78.20±0.79	65.57±1.73	75.49±1.51	66.81±0.32	76.76±0.71	68.62±0.48	77.93±0.64	67.49±0.89	76.67±0.96		

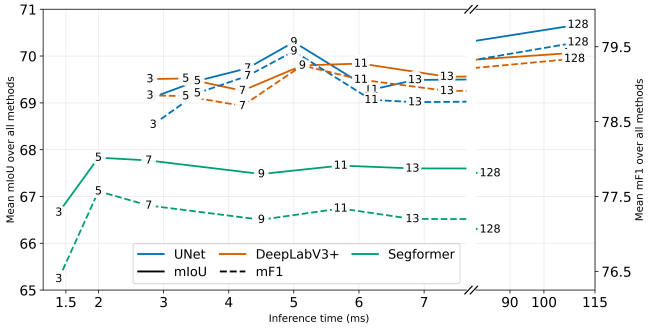


Fig. 3: CPU input-layer inference time and averaged mIoU/mF1 across methods in Table 1: Annotations indicate input and 128-band points. Inference used batch size 1, 10 warm-up iterations, excluded data loading, and was benchmarked on Intel Xeon Gold 6252, Debian 12.

sensus exceeds baseline in only one of 36 results and never surpasses the best individual method, indicating that selection frequency of stable methods alone does not identify the highest-performing subsets.

4.3. Inference-Time and Accuracy Trade-off

Figure 3 shows that spectral reduction substantially reduces CPU input-layer inference time while largely preserving segmentation performance. At $K = 9$, inference is approximately 18.7–22.0× faster than with 128 bands across the three SSMs. Averaged across the six band-selection methods, mean performance at $K = 9$ remains close to baseline (mIoU and mF1 differences within roughly ±0.1 percentage points). However, this aggregate masks substantial method-level variation: five of six methods stay within about 1 percentage point of baseline, while JMIM+CSNR+Corr trails baseline by 2.44–2.74 mIoU and 1.72–2.21 mF1 points. These results indicate that $K = 9$ provides a favorable aggregate accuracy–efficiency trade-off, although the optimal subset size remains method- and SSM-dependent.

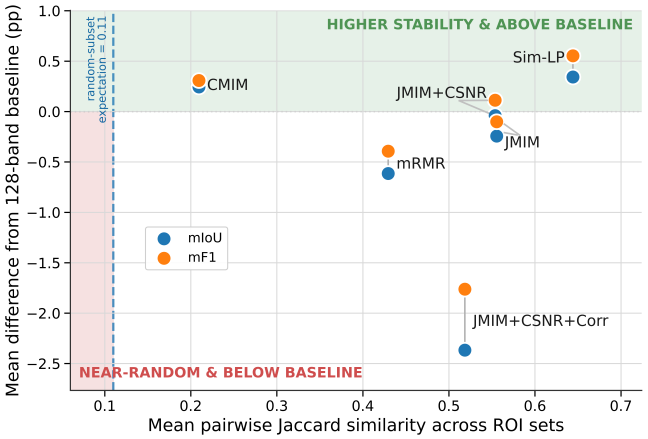


Fig. 4: Mean pairwise Jaccard similarity versus baseline-relative mIoU/mF1 differences, averaged across SSMs and subset sizes. Differences are reported in percentage points. Dashed lines mark the random-subset expectation (0.11) and baseline parity. Shading indicates positive performance differences and stability below the random-subset reference.

4.4. Method-Level Stability–Performance Association

Figure 4 shows no monotonic relationship between intra-method stability and baseline-relative segmentation performance. Sim-LP combines the highest stability with the largest positive method-level differences: approximately 0.34 percentage points in mIoU and 0.55 percentage points in mF1. JMIM+CSNR remains close to the corresponding baseline (-0.04 mIoU and +0.11 mF1), and JMIM+CSNR+Corr has the largest performance deficits (-2.37 mIoU and -1.76 mF1) despite relatively high stability. Whereas CMIM has the lowest stability but positive differences in both metrics. JMIM+CSNR remains close to baseline despite relatively high stability. The results indicate that greater stability neither guarantees better segmentation performance nor is required for it, and the two criteria should be assessed jointly for band selection performance studies.

5. CONCLUSION

This study evaluated the intra-method stability of six band-selection methods under repeated class-balanced ROI sampling. It examined whether stability was associated with downstream semantic-segmentation performance across subset sizes and SSMs. Repeated class-balanced ROI sampling showed that band-selection stability is method-dependent but does not reliably predict downstream segmentation performance. Sim-LP remained the most intra-method stable, while Sim-LP and JMIM+CSNR produced the six SSM-metric optimal results, including gains of up to 2.01 mIoU and 1.72 mF1 points over 128-band baselines. The relatively stable JMIM+CSNR+Corr showed the largest performance deficits, while the least stable CMIM method showed positive average baseline-relative differences in both metrics. Three of the best results occurred at $K = 9$, while the others occurred once each in $K \in \{3, 5, 7\}$. The frequency-consensus ranking derived from the more stable methods did not outperform individual methods, indicating that stability-driven consensus alone is insufficient for selecting high-performing subsets. Performance was non-monotonic with subset size, although $K = 9$ provided a favorable aggregate efficiency–accuracy trade-off, with 18.7–22.0$\times$ faster CPU input-layer inference while remaining close to baseline performance.

These results show that intra-method stability is useful for characterizing the reproducibility of a band-selection method, but is not a reliable indicator of downstream segmentation performance in these experiments. Band-selection methods should therefore be evaluated across repeated samples, subset sizes, and downstream SSMs rather than judged by stability or a single selected subset alone. These conclusions are limited to one urban HSI dataset and pairwise Jaccard similarity. Future work should assess whether the observed stability–performance relationship generalizes across additional HSI segmentation datasets and alternative stability measures.

6. REFERENCES

[1] I. A. Shah, J. Li, R. George, T. Brophy, E. Ward, M. Glavin, E. Jones, and B. Deegan, "Hyperspectral sensors and autonomous driving: Technologies, limitations, and opportunities," *IEEE Open Journal of Vehicular Technology*, vol. 7, pp. 124–143, 2026.

[2] Y. Yuan, G. Zhu, and Q. Wang, "Hyperspectral band selection by multitask sparsity pursuit," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 2, pp. 631–644, 2014.

[3] I. A. Shah, J. Li, M. Glavin, E. Jones, E. Ward, and B. Deegan, "Hyperspectral imaging-based perception in autonomous driving scenarios: Benchmarking baseline semantic segmentation models," in *2024 14th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, 2024, pp. 1–5.

[4] Q. Shen, Y. Huang, T. Ren, Y. Fu, and S. You, "Urban scene understanding via hyperspectral images: Dataset and benchmark," *Available at SSRN 4560035*, 2024.

[5] Y. Huang, Q. Shen, Y. Fu, and S. You, "Weakly-supervised semantic segmentation in cityscape via hyperspectral image," in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) Workshops*, October 2021, pp. 1117–1126.

[6] J. Li, I. A. Shah, D. Geever, F. Collins, E. Ward, M. Glavin, E. Jones, and B. Deegan, "CSNR and JMIM based spectral band selection for reducing metamerism in urban driving," *IEEE Open Journal of Intelligent Transportation Systems*, vol. 7, pp. 1021–1033, 2026.

[7] J. Li, I. A. Shah, D. Geever, E. Ward, M. Glavin, E. Jones, and B. Deegan, "Hyperspectral vs. RGB for pedestrian segmentation in urban driving scenes: A comparative study," in *2025 IEEE International Conference on Vehicular Electronics and Safety (ICVES)*, 2025, pp. 387–392.

[8] O. G. Yaniv Zimmer, Ofir Lindenbaum, "Supervised embedded methods for hyperspectral band selection," in *Volume 413: ECAI*, 2025, pp. 3154–3161.

[9] I. A. Shah, J. Li, E. Delaney, E. Ward, M. Glavin, E. Jones, and B. Deegan, "Learnable quantum efficiency filters for urban hyperspectral segmentation," *IEEE Open Journal of Vehicular Technology*, vol. 7, pp. 1311–1322, 2026.

[10] H. Peng, F. Long, and C. Ding, "Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 27, no. 8, pp. 1226–1238, 2005.

[11] F. Fleuret, "Fast binary feature selection with conditional mutual information," *Journal of Machine learning research*, vol. 5, no. Nov, pp. 1531–1555, 2004.

[12] M. Bennasar, Y. Hicks, and R. Setchi, "Feature selection using joint mutual information maximisation," *Expert systems with applications*, vol. 42, no. 22, pp. 8520–8532, 2015.

[13] Q. Du and H. Yang, "Similarity-based unsupervised band selection for hyperspectral image analysis," *IEEE geoscience and remote sensing letters*, vol. 5, no. 4, pp. 564–568, 2008.

[14] S. Nogueira, K. Sechidis, and G. Brown, "On the stability of feature selection algorithms," *Journal of Machine Learning Research*, 2018.

[15] P. Jaccard, "The distribution of the flora in the alpine zone. 1," *New phytologist*, vol. 11, pp. 37–50, 1912.

[16] Y. Liu, J. Yang, Y. Chen, K. Tan, L. Wang, and X. Yan, "Stability analysis of hyperspectral band selection algorithms based on neighborhood rough set theory for classification," *Chemometrics and Intelligent Laboratory Systems*, vol. 169, pp. 35–44, 2017.