

Assessing the Transferability of Road Network Extraction Methods to Slovenian Orthophotos for Power Distribution Network Planning

Janez Križaj, Klemen Knez, Vitomir Štruc

Faculty of Electrical Engineering, University of Ljubljana, Tržaška 25, 1000 Ljubljana, Slovenia
E-mail: janez.krizaj@fe.uni-lj.si

Abstract. In spatial planning of power distribution networks, road corridors provide an important basis for routing medium- and low-voltage power cables, as they often simplify permitting procedures, construction, and maintenance access. Automatic road extraction from aerial and satellite imagery is therefore not only a segmentation task, but also an input to subsequent optimization procedures, where connected and navigable road networks are required. The paper compares six recent methods for road network extraction: SAM-Road, D-LinkNet, CRESI, CU-dGCN, Sat2Graph, and U-Net-ResNet18. They are first evaluated on the SpaceNet dataset and then tested for transfer to Slovenian GURS orthophoto imagery. The results show that methods with an explicit graph output generally preserve road network connectivity better in the experiments. When transferred to Slovenian GURS orthophotos, their relative performance is largely retained. The methods that perform well on the SpaceNet also tend to perform well in this setting. A decline in their performance is observed mainly in rural areas, where roads are often narrow, partly occluded by vegetation or shadows, and visually less distinguishable from the surrounding terrain. Besides a comparative evaluation, the paper therefore provides practical guidance for selecting methods when constructing road networks for the optimization-supported power distribution network planning.

Keywords: road extraction, benchmarking, satellite imagery, SpaceNet, domain shift, power distribution spatial planning

Vrednotenje prenosljivosti metod za ekstrakcijo cestnih omrežij na slovenske ortofoto posnetke za potrebe načrtovanja distribucijskih omrežij

Pri prostorskem načrtovanju elektrodistribucijskih omrežij so cestni koridorji pomembna podlaga za izbiro tras srednje- in nizkonapetostnih vodov, saj pogosto omogočajo lažje pridobivanje dovoljenj, enostavnejšo gradnjo in boljši dostop za vzdrževanje. Zato avtomatska ekstrakcija cest iz letalskih in satelitskih posnetkov ni uporabna le kot segmentacijska naloga, temveč predvsem kot vhod v nadaljnje postopke optimizacije, kjer so potrebna povezana in usmerljiva cestna omrežja. V članku primerjamo šest sodobnih pristopov za ekstrakcijo cestnih omrežij: SAM-Road, D-LinkNet, CRESI, CU-dGCN, Sat2Graph in U-Net-ResNet18. Metode najprej ovrednotimo na zbirki SpaceNet, nato pa preverimo še njihov prenos na slovenske ortofoto posnetke GURS-a. Rezultati kažejo, da metode z eksplisitnim grafnim izhodom v eksperimentih praviloma bolje ohranjajo povezanost cestnega omrežja. Pri prenosu na slovenske ortofoto posnetke GURS-a se relativna uspešnost metod večinoma ohrani, saj se pristopi, ki dosegajo dobre rezultate na zbirki SpaceNet, praviloma dobro izkažejo tudi v tem okolju. Poslabšanje je opazno predvsem na podeželskih območjih, kjer so ceste pogosto ožje, delno

zakrite z vegetacijo ali sencami ter vizualno manj ločljive od okoliškega terena. Prispevek zato poleg primerjalnega vrednotenja podaja tudi praktične smernice za izbiro metod pri gradnji cestnega omrežja, uporabnega za optimizacijsko podprto načrtovanje elektrodistribucijskega omrežja.

Ključne besede: ekstrakcija cest, primerjalno vrednotenje, satelitski posnetki, SpaceNet, premik domene, prostorsko načrtovanje distribucijskega omrežja

1 INTRODUCTION

Accurate road network extraction from aerial and satellite imagery is a longstanding challenge in computer vision and remote sensing. It underpins crucial applications in navigation, autonomous driving, urban planning, and disaster response, and is also a key input for spatial planning of the power system infrastructure. In the power distribution network development, candidate medium- and low-voltage corridors are often derived from transport right-of-way: routing along roads can reduce permitting effort, limit excavations on private parcels, and guarantee a construction and maintenance access. In the Slovenian low-voltage power distribution context that motivates our work, the power distribution design rules go further and place cables along the road

Received 22 June 2026
Accepted 4 September 2026



Copyright: © 2026 by the authors.
Creative Commons Attribution 4.0
International License

edge wherever feasible, while the imposed planning criteria bound the admissible feeder length through the reference-impedance limit and cap the permissible voltage drop [1]. The road layer is therefore both a spatial substrate for routing and a determinant of whether the resulting feeder meets the imposed electrical limits. Consequently, the road layer quality and completeness can directly affect the corridor-aware routing, cost modelling, and feasibility screening of candidate alignments.

While early approaches relied on manual mapping or classical image processing, the advent of deep learning has led to a significant progress in the automated road extraction. Yet, producing routable road networks – continuous centerline graphs suitable for path planning – is still a highly demanding task. CNN-based segmentation often yields fragmented or noisy road masks, necessitating post-processing (e.g. thinning or connectivity analysis) to derive the vector road graphs. It should be noted that errors that are visually minor at the pixel level (small gaps, slightly misplaced intersections) can become a major issue once converted into graphs, by disconnecting otherwise feasible corridors or introducing spurious shortcuts.

Motivated by above challenges, our paper is framed as a topology-aware transferability study rather than considering the benchmark rankings as universally portable across image domains. It focuses on the state-of-the-art models covering the spectrum from a pure mask prediction to the approaches that explicitly model connectivity. The models are evaluated by a representation ultimately needed by corridor-routing workflows: a connected road graph.

Our paper tries to answer the following research questions:

- RQ1. When each output is converted to the centerline graph, do the graph-centric methods provide a more routing-useful topology than the segmentation-first baselines on the SpaceNet benchmark?
- RQ2. To what extent do conclusions from the SpaceNet transfer to the Slovenian GURS orthophotos when the OSM roads are used as a pseudo-reference rather than as manually-validated ground truth?
- RQ3. How do the rural/urban context and the input ground sample distance (0.25, 0.50, and 1.00 m/px) affect the road-graph quality under the GURS/OSM setup?
- RQ4. Which method-selection implications follow for the power distribution network corridor planning when the missing links remove the feasible corridors and the spurious links create infeasible shortcuts?

In our paper, the recent road extraction methods are classified as the segmentation-based, skeleton-based, and graph-based methods [2]–[5]. They are evaluated using

a common centerline-graph representation. Six representative pipelines are compared — D-LinkNet [6] and U-Net-ResNet18 [7] (segmentation baselines), CRESI [8] (segmentation+skeleton), CU-dGCN [9] and SAM-Road [10] (connectivity-aware hybrids), and Sat2Graph [11] (direct graph prediction) — on SpaceNet (RQ1) and, in a zero-shot transfer study, on the Slovenian GURS orthophotos with OSM as a pseudo-reference across rural/urban contexts and three resolutions (RQ2–RQ3). Finally, we relate the topology and the overlap metrics to the corridor-planning risks, such as disconnected feasible routes and spurious non-road shortcuts (RQ4).

2 RELATED WORK

Road extraction from a remote sensing imagery has been widely studied. Recent surveys show that there has been a clear shift from the classical image processing towards the deep learning pipelines [2]–[5]. Consistent with these surveys, we adopt a taxonomy that groups the approaches into: (i) segmentation-based methods that predict pixel-wise road masks, (ii) skeleton-based pipelines that derive centerlines/graphs from the predicted masks via thinning and pruning, and (iii) graph-based methods that model the connectivity explicitly, either by learning on graphs or by directly predicting the graph structures. The taxonomy is practically relevant whenever the downstream goal is a routable network, i.e. small gaps or spurious links that are visually minor at the pixel level can disconnect feasible routes or introduce unrealistic shortcuts once a mask is converted into a graph.

The segmentation-based methods commonly cast the road extraction as a semantic segmentation with encoder-decoder CNNs (e.g., FCN/U-Net/SegNet families) [12], [13] producing a dense road mask that can be thresholded and later vectorized. D-LinkNet [6] improves context aggregation (e.g., via dilated convolutions), while lightweight variants, such as U-Net-ResNet18 [7], remain strong baselines. A key limitation for optimisation-ready use is that visually minor mask defects (gaps, junction artifacts) can become major topological errors after skeletonization.

Skeleton-based pipelines address such limitation by augmenting segmentation with explicit mask cleanup, skeletonization, and pruning to improve routability. A representative example is CRESI [8]. While effective in practice, their final graph quality remains sensitive to the continuity of the predicted mask and to heuristic post-processing choices.

Graph-based methods target connectivity more directly. Some incorporate topology cues into learning, for example via multi-task designs that add structural supervision such as orientations [14], while hybrids inject graph reasoning modules to propagate connectivity information [9]. In parallel, graph-centric approaches

avoid heuristic vectorization by predicting road graphs explicitly: RoadTracer [15] pioneers on iterative graph construction, Sat2Graph [11] predicts vertices and edges in a single forward pass, and transformer-based detectors such as RNGDet [16] and RNGDet++ [17] extend this paradigm with detection-style graph decoding. Such trend towards graph outputs (rather than raster masks) is also emphasized in recent surveys [2], [4], [5].

Despite the progress, existing benchmarks and surveys evaluate road extraction almost exclusively on large public datasets, such as SpaceNet and DeepGlobe, and predominantly with pixel-overlap or generic topology metrics. To the best of our knowledge, no prior work has assessed how these methods transfer, in a strict zero-shot setting, to Slovenian national (GURS) orthophotos, nor has evaluated them under an explicit power-distribution corridor objective in which missing or spurious links translate into particular routing and electrical-feasibility risks. To solve the gap, our paper harmonizes six representative pipelines to a common centerline-graph representation and analyses both their SpaceNet behaviour and their domain-shifted behaviour on GURS imagery through the lens of the power-distribution network corridor planning.

3 METHODS

Our paper compares six representative approaches for the above three categories: segmentation-based mask predictors (D-LinkNet, UNet-ResNet18), a skeleton-based pipeline that derives a centerline graph from a predicted mask (CRESI), and graph-based methods that explicitly model connectivity, either by refining mask predictions with learned graph reasoning (CU-dGCN, SAM-Road) or by directly predicting a road graph (Sat2Graph). Each model architecture and output format are summarized below.

3.1 U-Net-ResNet18 (Segmentation baseline)

U-Net-ResNet18 [7] employs a standard U-Net encoder-decoder architecture with a ResNet-18 backbone, providing a lightweight yet effective road segmentation baseline. As a pure segmentation approach, it outputs a binary road mask after thresholding. Connectivity is only implicitly learned and not explicitly enforced.

3.2 D-LinkNet (Segmentation baseline)

D-LinkNet [6] is a CNN road segmenter based on an encoder-decoder architecture with skip connections. Dilated convolutions enlarge the receptive field while preserving spatial detail, which helps bridge gaps caused by occlusions or low contrast. In typical configurations, a pretrained ResNet backbone extracts multi-scale features and the decoder fuses them into a dense road-probability map. The output is a binary mask after thresholding. The connectivity is encouraged only implicitly through a context aggregation and is not explicitly enforced.

3.3 CRESI (Segmentation + skeleton)

CRESI [8] targets large-scale road graph extraction through a pragmatic multi-stage pipeline. A segmentation model predicts road likelihood on overlapping tiles which are then mosaicked into a global mask to reduce boundary artifacts. Post-processing cleans the mask and fills small gaps before skeletonization. The skeleton is converted into a graph, and pruning removes spurs and short branches to improve routability. This design emphasizes that the graph-construction stage can be as important as the segmentation backbone.

3.4 CU-dGCN (Segmentation + graph reasoning)

CU-dGCN [9] couples convolutional image features with graph neural network reasoning. A road centerline graph is constructed or predicted and then refined via message passing, allowing connectivity cues to propagate across gaps and complex junctions. The refined graph representation is mapped back to improve the road prediction, effectively combining local appearance with global structure. This hybrid design aims to deliver better routability without minimizing the efficiency and stability of mask-based segmentation.

3.5 SAM-Road (Segmentation + graph reasoning)

SAM-Road [10] adapts the Segment Anything Model (SAM) to road extraction using road-specific prompting strategies and a decoding pipeline that yields road masks and graph representations. Leveraging a foundation-model prior can help under domain shift and with limited labeled data. In practice, SAM-Road still depends on prompt design and decoding choices to avoid fragmented masks and to produce a routable network, but it reduces reliance on training a task-specific segmenter from scratch.

3.6 Sat2Graph (Direct graph prediction)

Sat2Graph [11] predicts a road graph in a single forward pass by learning structured representations of vertices and edges from the image. Compared to tracing, global prediction is less sensitive to local dead-ends and can better preserve long-range continuity. The method relies on carefully defined training targets and a robust graph decoding stage to avoid missing thin roads or introducing shortcuts. The output is a graph which can be rasterized to a mask for overlap-based evaluation when needed.

4 EXPERIMENTS

The experiments follow the research questions defined in Section 1. The SpaceNet benchmark addresses RQ1 by comparing graph-normalized outputs under topology-aware metrics. The GURS application-domain study addresses RQ2 and RQ3 by testing whether SpaceNet conclusions persist under Slovenian orthophoto imagery,

OSM pseudo-reference data, rural/urban variation, and three input resolutions. RQ4 is addressed by interpreting metrics in terms of the corridor-planning risks.

Each method is evaluated using its publicly released SpaceNet-pretrained weights. No fine-tuning, domain adaptation, or retraining on Slovenian imagery are performed. The GURS evaluation is therefore a strict zero-shot transfer test. It measures how SpaceNet-pretrained models behave under domain shift to national orthophotos, rather than the accuracy attainable after adaptation. This choice is deliberate. It isolates out-of-the-box robustness, i.e. the property that matters when such extractors are reused as off-the-shelf priors in an operational planning workflow, and it avoids confounding the method ranking with method-specific fine-tuning budgets. Accordingly, all differences reported below between SpaceNet and GURS reflect differences in zero-shot generalization of the pretrained models, not the effect of any Slovenia-specific training.

We first compare representative segmentation-first and graph-centric road extraction methods on the SpaceNet roads benchmark, which provides satellite imagery paired with ground-truth vector road graphs. As our target downstream use is a corridor-aware routing for spatial optimisation (including power system infrastructure planning), our focus is on the connectivity and routability rather than on the pixel overlap alone. To make our comparison consistent across the output types, all predictions are evaluated as centerline graphs. Mask predictions are thresholded and converted to the centerline graphs using a standard thinning-and-tracing pipeline with light cleanup, while native graph predictions are normalized to the same graph representation before scoring. We report TOPO* precision/recall/F1 and APLS (Average Path Length Similarity) following the SpaceNet evaluation definitions. The precision/recall/F1 values shown in Table 1 correspond to TOPO rather than pixel-level segmentation scores.

Table 1 summarizes quantitative results for RQ1. For corridor selection and cost modelling, APLS is particularly informative because it measures shortest-path similarity and is therefore sensitive to disconnections that would force unrealistic detours in an optimised route. TOPO complements APLS by capturing local structural agreement around intersections and branches. The high precision discourages spurious links (which can create infeasible shortcuts along non-existent roads) while high recall reduces the risk of missing feasible corridors. Compared to other methods, SAM-Road achieves the highest APLS and very high TOPO precision, indicating clean graph predictions with comparatively few false connections. However, its lower recall suggests that it can miss parts of the network. Sat2Graph provides the

Table 1. The SpaceNet results for the representative methods in our taxonomy. Prec./Rec./F1 are TOPO metrics. APLS measures the routing fidelity. The values are in %.

Method	SpaceNet			
	Prec. ↑	Rec. ↑	F1 ↑	APLS ↑
U-Net-ResNet18 [7]	68.96	66.32	67.61	53.77
D-LinkNet [6]	82.42	60.06	68.80	56.93
CRESI [8]	79.11	69.47	71.87	56.38
CU-dGCN [9]	62.13	61.66	62.32	58.19
Sat2Graph [11]	85.93	76.55	80.97	64.43
SAM-Road [10]	93.03	70.97	80.52	71.64

best balance in TOPO and also strong APLS. This can be advantageous when the planning goal is to enumerate alternative access corridors rather than only the most certain ones. CU-dGCN and CRESI are in the mid-range. The segmentation-first baselines (D-LinkNet and U-Net-ResNet18) exhibit lower APLS, consistently with the fact that small mask gaps and junction errors tend to be amplified during vectorization. The answer to RQ1 is therefore conditional rather than absolute. Thus, the strongest SpaceNet results come from the graph-centric methods, but graph reasoning alone is not sufficient. The specific decoder and post-processing strategy still determine whether topology is actually preserved.

Fig. 1 complements the quantitative SpaceNet results with representative urban and rural examples. It illustrates why graph-aware evaluation is needed: methods can recover visually plausible road evidence while still differing substantially in continuity, missing branches, and spurious fragments.

4.1 Application case: domain shift to GURS orthophotos

To address RQ2 and RQ3, each of the six pipelines on Slovenian national orthophotos from the Surveying and Mapping Authority of the Republic of Slovenia (GURS) [18] is reevaluated. The experiment covers three areas of interest (AOIs) in the Gorenjska region, including the road contexts relevant to distribution-grid corridor planning. One is urban AOI at the Šenčur suburb and two rural AOIs located around the localities of Grad Kamen and Križna Gora. They are grouped into a rural set (the two AOIs around Grad Kamen and Križna Gora) and an urban set (Šenčur). The two rural AOIs that together cover roughly 8 km² and the urban Šenčur AOI about 2 km², enable a combined evaluation footprint on the order of 10 km² at the native GURS resolution of 0.25 m/px.

In absence of manually validated AOIs road centerlines, are the OpenStreetMap (OSM) roads [19] used as a pseudo-reference layer. OSM is openly licensed and already encoded as a connected centerline topology that maps directly onto our graph-based scoring and onto the SpaceNet road graphs. This enables applying the same metrics without re-annotation and it mirrors

*TOPO evaluates the similarity of reachable subgraphs from randomly sampled vertices in the predicted road graph and ground truth.

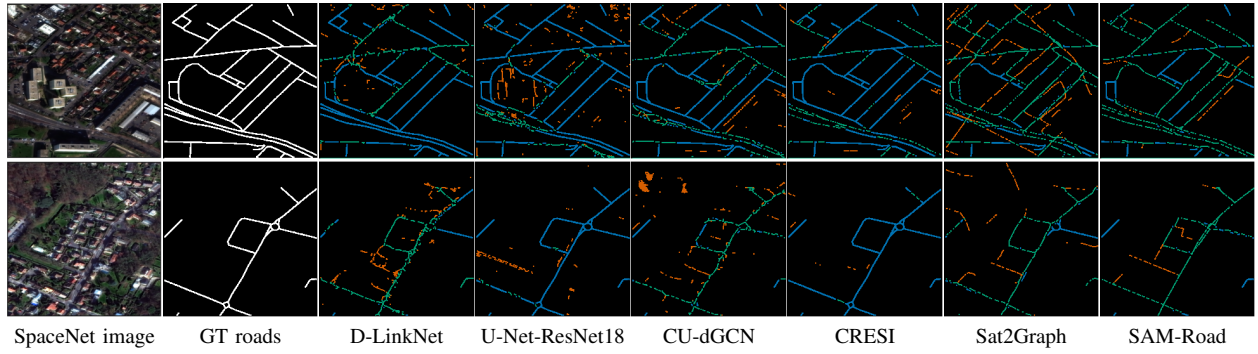


Figure 1. Qualitative comparison of road extraction outputs on SpaceNet examples. Top row: urban scene; bottom row: sub-urban/rural scene. Column 1 shows the SpaceNet image and column 2 the corresponding ground-truth road mask. The remaining columns show confusion-coded canonical centerline renderings: green denotes true positives, red/orange denotes false positives, and blue denotes false negatives. Native graph outputs are rasterized for display and mask outputs are skeletonized before rendering.

the open prior an operational planning workflow would realistically start from. This makes the experiment an application-domain robustness study rather than a fully controlled benchmark. The OSM incompleteness or small misalignments can penalize plausible predictions. As the study covers only three AOIs and reports micro-averaged group scores, the GURS numbers should be read as indicative robustness trends rather than statistically validated rankings and we emphasize consistent, large-margin effects over small differences between methods.

Each AOI is evaluated at three ground sample distances: 0.25, 0.50, and 1.00 m/px. Each method prediction is first converted to a canonical georeferenced centerline graph. Native graph outputs from CRESI, CU-dGCN, Sat2Graph, and SAM-Road are normalized to GeoJSON linework. The mask-only outputs from D-LinkNet and U-Net-ResNet18 are thresholded, skeletonized with a fixed 8-connected tracing rule, transformed to map coordinates and exported as the same GeoJSON linework. The canonical prediction graphs are then scored against the OSM pseudo-reference using the same TOPO precision, recall, and F1 defined for the SpaceNet evaluation above, together with an additional centerline-tolerance intersection-over-union (IoU_τ) that captures geometric agreement. For IoU_τ , the predicted and reference road geometries are expanded by the same spatial scoring tolerance $\tau = 2.0$ m and rasterized on the evaluation grid using a one-pixel base line width. If P_τ and R_τ denote the resulting prediction and reference masks, respectively, then

$$\text{IoU}_\tau = \frac{|P_\tau \cap R_\tau|}{|P_\tau \cup R_\tau|}, \quad (1)$$

where $|\cdot|$ denotes raster pixel count. This tolerance-based overlap reduces the penalty for small centerline misalignments, and the same $\tau = 2.0$ m is used throughout. Precision, recall, F1 and APLS thus share their definitions with the SpaceNet values in Table 1.

The two tables differ only in what they are measured against, since the GURS scores reflect agreement with the OSM pseudo-reference rather than validated ground truth, so cross-dataset magnitudes should thus be read with that caveat.

The grouped results answer RQ3 by showing a clear domain-structure effect. Urban AOIs are consistently easier for most pipelines because roads are wider, less frequently occluded by vegetation, and more visually separable from the background. Sat2Graph is the strongest method overall, obtaining the best F1 and IoU_τ in every group-resolution combination, and the highest APLS in all settings except the rural coarsest-resolution case, where D-LinkNet leads despite its much lower recall. This exception illustrates a limitation of APLS that is important for planning. As APLS is computed over paths between sampled nodes that exist in the prediction, a sparse but locally clean subgraph can score well even when most of the network is missing. A high APLS at near-trivial recall should therefore be read jointly with recall. It reflects faithful routing over the few roads recovered, not usable coverage of the corridor network - and we accordingly do not rank methods on APLS alone. In the rural set, Sat2Graph remains the most robust, but its agreement degrades markedly as resolution coarsens, reflecting the difficulty of narrow, partially occluded roads and forest tracks.

The answer to RQ2 is more cautious. Comparing SpaceNet and GURS supports transfer of a broad principle. Though the graph-centric outputs are preferable when the downstream task is routing, it does not support a direct transfer of the full method ranking. The shared APLS metric shows that the SpaceNet ordering does not transfer directly to Slovenian orthophotos. While Sat2Graph is still strong across both datasets, SAM-Road drops substantially under the GURS/OSM pseudo-reference setup despite having the best SpaceNet APLS. As no method is adapted to GURS, these gaps are a

Table 2. Application-domain results on GURS orthophotos using OSM roads as pseudo-reference after conversion to canonical centerline graphs. Precision, recall and F1 are the same TOPO metrics reported in Table 1; precision, recall, F1, and IoU_τ are micro-averaged within each AOI group; APLS is averaged over AOIs within each group. Values are in %. IoU_τ denotes centerline-tolerance IoU, computed by expanding prediction and reference road geometries with the same spatial scoring tolerance ($\tau = 2.0$ m) before rasterization and intersection-over-union calculation.

Group	Method	Prec. $\uparrow$			Rec. $\uparrow$			F1 $\uparrow$			$\text{IoU}_\tau \uparrow$			APLS $\uparrow$		
		0.25	0.50	1.00	0.25	0.50	1.00	0.25	0.50	1.00	0.25	0.50	1.00	0.25	0.50	1.00
Rural	U-Net-ResNet18 [7]	81.15	82.21	80.31	37.91	42.52	17.25	51.68	56.05	28.39	27.16	29.92	14.64	39.30	40.12	38.70
	D-LinkNet [6]	85.54	78.76	86.61	17.31	40.50	9.47	28.80	53.49	17.07	14.84	28.40	8.76	36.06	40.37	42.30
	CRESI [8]	30.51	42.20	18.95	18.78	14.74	7.80	23.25	21.85	11.05	11.35	10.66	5.36	40.56	18.28	11.46
	CU-dGCN [9]	74.54	56.13	8.76	17.17	26.20	7.20	27.91	35.73	7.91	14.19	18.89	3.75	49.94	31.69	8.60
	Sat2Graph [11]	54.97	56.47	23.22	71.91	71.18	75.12	62.31	62.98	35.47	35.83	36.71	17.87	62.64	48.37	31.25
	SAM-Road [10]	79.18	68.13	62.25	24.44	31.46	16.35	37.36	43.04	25.90	19.66	23.20	13.26	39.07	43.00	23.90
Urban	U-Net-ResNet18 [7]	66.41	69.95	71.97	55.06	72.88	54.60	60.20	71.39	62.09	31.55	38.51	33.86	30.27	37.30	33.81
	D-LinkNet [6]	70.04	70.35	71.35	38.33	70.69	47.98	49.54	70.52	57.38	26.22	38.25	30.94	40.35	47.44	41.10
	CRESI [8]	48.20	57.11	28.81	51.89	36.51	9.24	49.98	44.54	13.99	28.40	24.12	6.79	23.55	22.49	6.01
	CU-dGCN [9]	57.13	84.27	91.38	17.24	56.65	30.07	26.48	67.75	45.25	13.65	41.21	24.38	20.81	36.20	39.24
	Sat2Graph [11]	69.74	71.58	60.87	82.79	83.30	83.93	75.71	76.99	70.57	47.13	49.47	44.71	58.80	61.81	51.31
	SAM-Road [10]	68.29	78.77	80.17	62.23	72.87	34.09	65.11	75.71	47.84	40.28	48.13	26.06	37.34	53.27	33.36

property of the zero-shot generalization of the pretrained models rather than of any Slovenia-specific training. Robustness under domain shift does therefore not depend only on whether a method produces a graph, but also on the graph decoder, the SpaceNet training domain it inherits, thresholding behavior, and post-processing choices.

Resolution sensitivity also belongs to RQ3 and differs between AOI groups. In urban AOIs, 0.50 m/px is the strongest or near-strongest operating point for most methods, while 1.00 m/px still preserves enough structure for Sat2Graph and the segmentation baselines to remain useful. In rural AOIs, recall degrades more sharply at 1.00 m/px because small roads and junctions become poorly resolved. The native 0.25 m/px imagery does not automatically improve all methods. It increases texture detail and memory pressure, and segmentation-first models can become conservative or fragmented. For power-distribution spatial planning, this means that urban road layers can often be extracted with several methods after moderate cleanup, whereas rural planning workflows benefit more strongly from a graph prior and still require a manual verification in forested or weakly mapped corridors.

Fig. 2 provides a qualitative support for the grouped results and connects them to RQ4. The first row illustrates the urban Šenčur AOI. Here most methods follow the main mapped roads more reliably, but the skeletonized segmentation baselines can fragment around intersections and parking areas. The second and third row show two rural AOIs (near Grad Kamen and Križna Gora), where roads are narrower, more curved, and more frequently hidden by trees. Conservative methods therefore miss several OSM segments, while the aggressive methods risk hallucinating short disconnected fragments. Sat2Graph better preserves the main rural connectivity, although its lower precision shown in Table 2 indicates that some extra links still require checking. Since OSM is only a pseudo-reference, apparent false positives may

include unmapped local roads. Nevertheless, such deviations remain important in planning workflows because they identify segments that require GIS review before automated routing.

5 CONCLUSION

The paper examines topology-aware road extraction for the power-distribution network spatial planning in the light of four questions. On SpaceNet (RQ1), the strongest routing-oriented scores come from graph-centric methods, but no graph-reasoning method dominates universally. The topology quality still depends on the decoder and graph cleanup. The methodological takeaway is that road extraction for corridor planning should be evaluated as graph construction rather than as mask segmentation alone.

On Slovenian GURS orthophotos (RQ2–RQ3), where each model runs zero-shot from its SpaceNet-pretrained weights, transferability claims must remain cautious. The broad preference for graph-centric outputs carries over and Sat2Graph stays robust, but the exact SpaceNet ranking does not. SAM-Road, the SpaceNet APLS leader, turns conservative and loses recall under domain shift. Rural and forested AOIs are markedly harder than urban ones, and 0.50 m/px is the most balanced operating point. Since OSM serves only as a pseudo-reference, these values indicate agreement with an operational planning prior rather than absolute accuracy.

For planning (RQ4), the practical recommendation is to use graph-centric extractors as candidate-corridor generators feeding a Study 2400-compliant routing and load-flow stage [1], not as a source of final network geometry. Recall bounds the corridors the optimiser can consider, while edge-level geometric fidelity governs whether a proposed corridor lands on usable right-of-way, so missing or spurious links can translate into infeasible feeder lengths or non-existent routes. High-precision extractors suit single permit-ready connec-

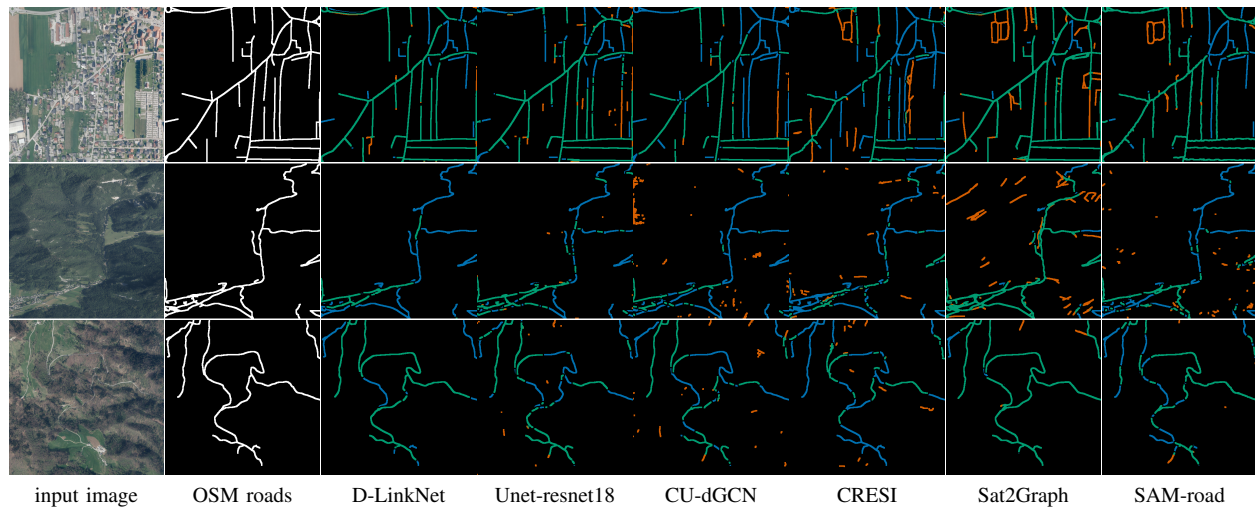


Figure 2. Visual comparison of road extraction outputs on GURS orthophotos [18]. Rows show the urban Šenčur AOI and the two rural AOIs near Grad Kamen and Križna Gora, respectively. Column 1: input orthophoto. Column 2: OSM road layer rasterized as a white skeleton. Remaining columns: confusion-coded canonical graph or skeleton predictions rendered with the same display width as the OSM layer: green denotes true positives, red/orange denotes false positives, and blue denotes false negatives.

tions, whereas higher-recall graph methods better support corridor enumeration, with GIS review retained in rural and weakly mapped areas. In the future work, these effects will be measured directly within routing, cost-modelling, and feeder-reinforcement workflows.

ACKNOWLEDGMENT

The authors gratefully acknowledge financial support from the fundamental research project Opti-AI (Optimal Integration of Power System Infrastructure through AI-based Spatial Planning, project no. J2-60030), funded by the Slovenian Research and Innovation Agency (ARIS).

DATA AVAILABILITY

This study relies on three publicly available data sources. The SpaceNet roads benchmark is publicly available under its published terms of use. The Slovenian national orthophoto imagery was obtained from the Surveying and Mapping Authority of the Republic of Slovenia (GURS) and is licensed under CC BY 4.0 [18]. The OpenStreetMap road layer used as a pseudo-reference is available under the Open Database License [19]. The six road-extraction models are evaluated using their publicly released, SpaceNet-pretrained weights without any fine-tuning. The original model repositories are referenced in [6]–[11].

REFERENCES

- [1] L. Valenčič, *Kriteriji načrtovanja NN omrežja: študija št. 2400*. Ljubljana: Elektroinštitut Milan Vidmar, 2018. Ov. nasl.
- [2] Z. Chen, L. Deng, Y. Luo, D. Li, J. Marcato Junior, W. Nunes Gonçalves, A. Awal Md Nurunnabi, J. Li, C. Wang, and D. Li, "Road extraction in remote sensing data: A survey," *International Journal of Applied Earth Observation and Geoinformation*, vol. 112, p. 102833, 2022.
- [3] P. Sharma, R. Kumar, M. Gupta, and A. Nayyar, "A critical analysis of road network extraction using remote sensing images with deep learning," *Spatial Information Research*, vol. 32, no. 4, pp. 485–495, 2024.
- [4] R. Liu, J. Wu, W. Lu, Q. Miao, H. Zhang, X. Liu, Z. Lu, and L. Li, "A review of deep learning-based methods for road extraction from high-resolution remote sensing images," *Remote Sensing*, vol. 16, no. 12, 2024.
- [5] X. Wang, X. Jin, Z. Dai, Y. Wu, and A. Chehri, "Deep learning-based methods for road extraction from remote sensing images: A vision, survey, and future directions," *IEEE Geoscience and Remote Sensing Magazine*, vol. 13, no. 1, pp. 55–78, 2025.
- [6] L. Zhou, C. Zhang, and M. Wu, "D-linknet: Linknet with pretrained encoder and dilated convolution for high resolution satellite imagery road extraction," in *Proc. CVPRW*, pp. 192–1924, 2018.
- [7] C. Robinson, I. Corley, A. Ortiz, R. Dodhia, J. M. L. Ferres, and P. Najafirad, "Seeing the roads through the trees: A benchmark for modeling spatial dependencies with aerial imagery," 2024.
- [8] A. V. Etten, "City-scale road extraction from satellite imagery v2: Road speeds and travel times," in *Proc. WACV*, pp. 1775–1784, 2020.
- [9] A. A. Vekinis, "Graph reasoned multi-scale road segmentation in remote sensing imagery," in *Proc. IGARSS*, pp. 6890–6893, 2023.
- [10] C. Hetang, H. Xue, C. Le, T. Yue, W. Wang, and Y. He, "Segment anything model for road network graph extraction," in *Proc. CVPR*, pp. 2556–2566, 2024.
- [11] S. He, F. Bastani, S. Jagwani, M. Alizadeh, H. Balakrishnan, S. Chawla, M. M. Elsharif, S. Madden, and M. A. Sadeghi, "Sat2graph: Road graph extraction through graph-tensor encoding," in *Proc. ECCV*, 2020.
- [12] V. Badrinarayanan, A. Kendall, and R. Cipolla, "Segnet: A deep convolutional encoder-decoder architecture for image seg-

- mentation,” *IEEE transactions on pattern analysis and machine intelligence*, vol. 39, no. 12, pp. 2481–2495, 2017.
- [13] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, “Encoder-decoder with atrous separable convolution for semantic image segmentation,” in *Proc. ECCV*, pp. 801–818, 2018.
 - [14] A. Batra, S. Singh, G. Pang, S. Basu, C. Jawahar, and M. Paluri, “Improved road connectivity by joint learning of orientation and segmentation,” in *Proc. CVPR*, pp. 10385–10393, 2019.
 - [15] F. Bastani, S. He, S. Abbar, M. Alizadeh, H. Balakrishnan, S. Chawla, S. Madden, and D. DeWitt, “Roadtracer: Automatic extraction of road networks from aerial images,” in *Proc. CVPR*, pp. 4720–4728, 2018.
 - [16] Z. Xu, Y. Liu, L. Gan, Y. Sun, X. Wu, M. Liu, and L. Wang, “Rngdet: Road network graph detection by transformer in aerial images,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–12, 2022.
 - [17] Z. Xu, Y. Liu, Y. Sun, M. Liu, and L. Wang, “Rngdet++: Road network graph detection by transformer with instance segmentation and multi-scale features enhancement,” *IEEE Robotics and Automation Letters*, vol. 8, no. 5, pp. 2991–2998, 2023.
 - [18] Ministry of Natural Resources and Spatial Planning and Surveying and Mapping Authority of the Republic of Slovenia, “Orthophoto.” <https://podatki.gov.si/dataset/ortofoto>, 2026. National orthophoto dataset. Licensed under CC BY 4.0. Accessed January 23, 2026.
 - [19] OpenStreetMap contributors, “Planet dump retrieved from <https://planet.osm.org>.” <https://www.openstreetmap.org>, 2017.

Janez Križaj is a researcher at the Faculty of Electrical Engineering, University of Ljubljana, and a member of the Laboratory for Machine Intelligence. His research focuses on artificial intelligence, machine learning, computer vision and biometrics. He has participated in several national and European research projects.

Klemen Knez is a researcher and teaching assistant at the Faculty of Electrical Engineering, University of Ljubljana, working in the field of electric power systems. His research focuses on distribution system modelling and analysis, state estimation, flexibility assessment, and the application of artificial intelligence and machine learning in power systems. He is actively involved in CIGRE and CIRED activities and serves as Chair of the Slovenian CIGRE Next Generation Network (NGN). Through his research and professional activities, he contributes to the development of smarter, more flexible, and sustainable power systems.

Vitimir Štruc is a full professor at the Faculty of Electrical Engineering, University of Ljubljana, and a researcher at the Laboratory for Machine Intelligence. His research interests include artificial intelligence, machine learning and deep learning, biometrics, computer vision, and signal and image analysis and processing. He has authored or co-authored more than 250 scientific publications and is actively involved in the editorial boards of leading international journals and in the organization of major conferences in computer vision and biometrics.