Proceedings of the Thirty-Fourth International Joint Conference on Artificial Intelligence (IJCAI-25)
Special Track on AI and Social Good

IGraSS: Learning to Identify Infrastructure Networks from Satellite Imagery by Iterative Graph-constrained Semantic Segmentation

Oishee Bintey Hoque[2], **Abhijin Adiga**[1], **Aniruddha Adiga**[1], **Siddharth Chaudhary**[4], **Madhav V. Marathe**[1,2], **S. S. Ravi**[1], **Kirti Rajagopalan**[3], **Amanda Wilson**[1] and **Samarth Swarup**[1]
[1]Biocomplexity Institute, University of Virginia,
[2]Department of Computer Science, University of Virginia
[3]Department of Biomedical Systems Engineering, Washington State University
[4]Earth System Science Center, University of Alabama in Huntsville
{gza5dr, abhijin, aa5dw, marathe, ssravi, alw4ey, swarup}@virginia.edu,
sc0248@uah.edu, kirtir@wsu.edu

Abstract

Accurate canal network mapping is essential for water management, including irrigation planning and infrastructure maintenance. State-of-the-art semantic segmentation models for infrastructure mapping, such as roads, rely on large, well-annotated remote sensing datasets. However, incomplete or inadequate ground truth can hinder these learning approaches. Many infrastructure networks have graph-level properties such as reachability to a source (like canals) or connectivity (roads) that can be leveraged to improve these existing ground truth. This paper develops a novel iterative framework *IGraSS*, combining a semantic segmentation module—incorporating RGB and additional modalities (NDWI, DEM)—with a graph-based ground-truth refinement module. The segmentation module processes satellite imagery patches, while the refinement module operates on the entire data viewing the infrastructure network as a graph. Experiments show that IGraSS reduces unreachable canal segments from 18% to 3%, and training with refined ground truth significantly improves canal identification. IGraSS serves as a robust framework for both refining noisy ground truth and mapping canal networks from remote sensing imagery. We also demonstrate the effectiveness and generalizability of IGraSS using road networks as an example, applying a different graph-theoretic constraint to complete road networks.

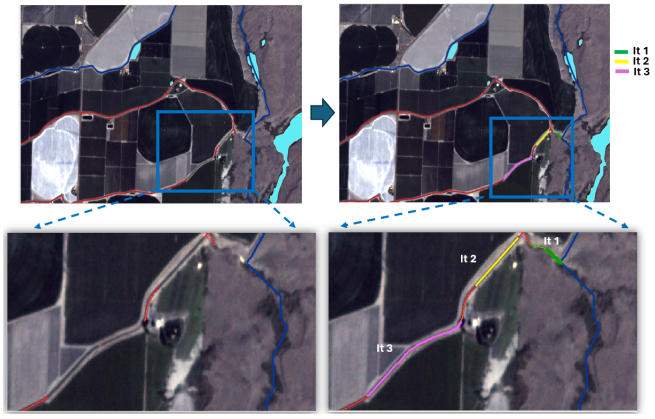


Figure 1: Visualization of Canal Network Completion via IGraSS: Blue lines represent reachable canal pixels, while red lines indicate unreachable canal pixels. The images demonstrate gaps in the red canal segments that are iteratively filled. Initially, the green segments connect one of the unreachable red segments with the blue reachable segment, making the upper red segment reachable in Iteration 1 (It #1). In the next iteration (It #2), yellow segments connect the smaller unreachable red segments. Finally (It #3), pink segments connect the remaining bottom segment by filling the gaps, thus making these canals reachable.

1 Introduction

Given the growing challenges of water conservation, modern irrigation patterns are shifting towards more cost-effective and water-efficient systems [Pérez-Blanco *et al.*, 2020; Fan *et al.*, 2023]. Furthermore, increasing pressures from drought, rising operational costs of canal infrastructure, and the decreasing cost of canal technology are driving many irrigation districts toward modernization [Belt and Smith, 2009; Creaco *et al.*, 2023; Fan *et al.*, 2023; Creaco *et al.*, 2023]. Modernizing irrigation infrastructure requires accurate knowledge of existing canal networks, but manual mapping is slow and outdated maps lack completeness [Hosseinzade *et al.*, 2017].We tackle this challenge by using high-resolution remote sensing to automatically extract canal networks, which will inform efforts on water management and infrastructure planning. This contributes to sustainable development and aligns with United Nations Sustainable Development Goals 12 (SDG 12) by promoting responsible resource use in agriculture and related sectors [United Nations, 2015].

Irrigation canal mapping by agencies typically relies on a labor-intensive GIS process, where experts manually annotate lines and polygons [Belt and Smith, 2009; Archuleta and Terziotti, 2023]. This approach often results in noisy, incom-

plete labels (See Fig 1) and requires updates over time as infrastructure evolves—for example, open canals being converted to closed pipe systems or covered with solar panels for efficiency [Loureiro *et al.*, 2024]. An automated remote-sensing approach can streamline updates and help assess water and energy efficiency improvements. While road network extraction is well-studied [Abdollahi *et al.*, 2020] with high-quality benchmarks [SpaceNet on Amazon Web Services (AWS), 2018; Demir et al., 2018], datasets for other infrastructure networks, such as irrigation canals, remain limited, often containing insufficient data and noisy, incomplete annotations.

Our approach leverages graph-theoretic properties (e.g., reachability, connectivity, planarity) to address noisy or incomplete infrastructure annotations. In irrigation networks, for instance, every canal segment should be reachable from a water source. Similar constraints apply to road and power networks. Unlike prior work using topological representations for road extraction, our method integrates graph constraints by aggregating segmentation outputs across multiple images. To effectively apply graph constraints such as reachability and connectivity, segmentation outputs must be aggregated across multiple images. Beyond segmentation, we propose a method to refine ground truth, which is often fragmented or disconnected. Using graph-theoretic properties, we correct these inconsistencies and show that improved annotations further boost model performance (See Figure 1).

Contributions. Our contributions in this work are: (i) a framework that allows the combination of learning with optimization/constraint satisfaction methods through the use of *pseudo-labels*, (ii) a refined set of metrics, called *r-neighborhood* metrics, which are more suitable for evaluating the performance of semantic segmentation problems like the ones studied here, (iii) a demonstration of improved performance in canal network identification (using reachability constraints), (iv) demonstration of generalizability using road network completion as an example, optimizing pairwise distances under a graph-theoretic constraint. (v) we provide the code and framework used to refine the canal network for the state of Washington, making it more connected and less noisy. The implementation is available for research purposes at: https://github.com/oishee-hoque/IGraSS.

Team. This work is the result of an interdisciplinary collaboration between computer scientists, an agro-ecosystems modeler with expertise in water and agricultural resource management, and an earth science and remote sensing expert.

2 Related Work

Liu *et al.* [2022] provide a comprehensive review of infrastructure network extraction using deep learning, particularly in the context of roads.) We use popular models from this literature as backbone networks for our semantic segmentation module. Similar approaches have also been applied to related problems such as crack detection, blood vessel segmentation, abnormality in anatomical structures, and extracting power systems [Ganaye *et al.*, 2018; Cheng *et al.*, 2021; Ren *et al.*, 2022].

Graph-based segmentation methods have been explored for road network inference and related tasks. Our work is most closely related to RoadTracer [Bastani *et al.*, 2018], which iteratively constructs road graphs using dynamic labels. However, unlike our approach, which refines labels based on global constraints (e.g., reachability), their method relies on a CNN-based decision function constrained to local patches.

Other relevant works include Sat2Graph [He *et al.*, 2020], which encodes road graphs as tensors for deep aggregation networks, and GA-Net [Chen *et al.*, 2022], which integrates segmentation with geometric road structures to enhance connectivity. Additionally, road boundaries detected via traditional filtering serve as inputs to a deep learner with D-LinkNet architecture [Zhou *et al.*, 2018]. Cira *et al.* [2022] use an inpainting approach as a postprocessing technique to link unconnected road segments. Zhang and Long [2023] utilize hypergraphs to capture high-order and long-range relationships among roads, incorporating various pretext tasks for optimization and demonstrating significant improvements across multiple datasets, tasks, and settings. None of these approaches use global graph constraints/optimization.

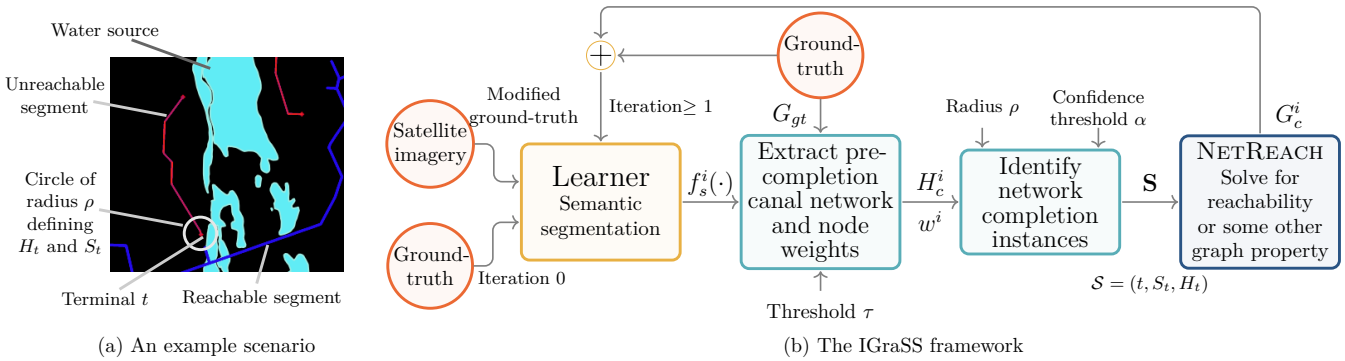
Mapping water bodies from remote-sensed data is an active area in remote sensing. Various unsupervised and supervised methods have been used. A prominent approach is the use of water indices such as normalized difference water index (NDWI). The use of deep learning methods in this context is an emerging area of research, with several papers using standard segmentation techniques. (See Nagaraj and Kumar [2024] for an extensive review.) Gharbia [2023] highlights challenges such as lack of quality data and variations in water body types. Li *et al.* [2022] consider the extraction of natural water bodies from binarized NDWI images. They apply a connected-component method followed by an analysis of shape and spectral characteristics to assign a confidence value to each water body. This is subsequently used to train peer networks. Yu *et al.* [2023] address the problem of fine-grained extraction of water bodies, where the challenge is to accurately detect the boundaries of water bodies. They propose a novel boundary-guided semantic context network in this regard. In our work, the emphasis is more on the accuracy of the network structure of the canal network.

3 Proposed Framework

3.1 Preliminaries

We use the problem of identifying irrigation canals from satellite images as the main running example in this paper. However, the framework can also be applied to other problems, such as identifying road networks.

We assume that the study region is overlaid with a grid where each grid cell corresponds to a pixel of a satellite image. Let G denote the graph induced by this grid where $V(G)$ denotes the set of grid cells. Two nodes $u, v \in V(G)$ are adjacent if and only if they share a side or a corner. This corresponds to the Moore neighborhood. The set of edges is denoted by $E(G)$. The *ground-truth canal network* G_{gt} is a subgraph of G induced by $V(G_{gt})$, the set of grid cells identified as canal pixels. Further, each $v \in V(G)$ is assigned a



(a) An example scenario

(b) The IGraSS framework

Figure 2: An example patch with gaps in data is shown in (a). In (b), an outline of our framework is shown.

label $\ell(v) \in \{0, 1\}$, where 0 denotes a non-canal node and 1 denotes a canal node. For all canal nodes ($\ell(v) = 1$), the label $\gamma(v) = 1$ implies that it is reachable from a water source, while 0 means otherwise. Henceforth, for brevity, we use the phrase "v is reachable" to denote that the node is reachable from a water source.

3.2 Problems

Inference goal. Under the assumption that the ground-truth canal network is incomplete (but not erroneous), the goal is to infer the canal network G_c, an induced subgraph of G such that all nodes in $V(G_c) \supseteq V(G_{gt})$ are reachable; i.e., $\forall v \in V(G_c)$, $\gamma(v) = 1$.

Prediction goal. Given the satellite imagery of a new region and the trained model, the objective is to identify all the edges of the canal network G_c.

3.3 Approach

Training data for neural network-based semantic segmentation approaches divide large satellite images into *patches*, where a patch is a small rectangular part of the large image. Typical patch sizes are 256×256 pixels or 512×512 pixels, whereas the image as a whole can be, say, ≈ 20,000×30,000 pixels in size. (See the Experimental Evaluation section for details regarding the data). While the classifier can be trained using each patch as a training input and the intersection of $V(G_{gt})$ with that patch as the corresponding expected output for training, it is important to note that the overall graph constraints (reachability, connectedness) are *global* constraints. These cannot be directly incorporated into the learning process by augmenting the objective function, as is customary in constrained learning problems.

Figure 2(a) shows a toy example where red-marked canals are not reachable from any water sources. The dark blue canals are reachable from the water source. Our goal is to identify these disconnected canal segments and connect them to nearby water sources or to reachable (blue) canals. We therefore develop a new, iterative approach composed of (i) a learner that infers a canal network given ground-truth data and satellite imagery after training for a number of intermediate epochs, and (ii) a *constraint solver* algorithm that modifies the training data after each iteration by adding positive *pseudo-labels* to chosen pixels. A positive pseudo-label refers to changing the label of a pixel from $\ell(v) = 0$ to $\ell(v) = 1$ to satisfy the constraint. In the canal network identification problem, the constraint solver is a network completion algorithm to satisfy the reachability constraint. See the outline in Figure 2. In each iteration, step (ii) modifies the expected outputs in the training data for the next round of training. The canal network is initially set to the ground-truth network, i.e., $G_c^0 = G_{gt}$, and is then modified in step (ii) of each iteration. The following are the steps in each iteration $i \geq 1$.

Learner. The learner's objective is to provide the likelihood that a node in the grid graph belongs to a canal. It is trained using modified ground-truth obtained from G_c^{i-1} for λ_i epochs. For the purpose of training in this iteration, for any node v, $\ell(v, i-1) = 1$ (i.e., v is a canal pixel in iteration $i-1$) if and only if $v \in V(G_c^{i-1})$. Let $f_s^i(v)$ denote the output of the learner. Note that the output depends on the learning methodology.

Pre-completion network. Given the learner's output, ground-truth network G_{gt}, and a user-specified threshold τ, a canal network graph H_c^i is computed. Firstly, a likelihood $w^i(v) \in [0, 1]$ is computed from $f_s^i(v)$. We construct a *pre-completion canal network* H_c^i as a graph induced by the nodes that satisfy the following condition: (i) $v \in V(G_{gt})$ (ground-truth) or (ii) $w^i(v) \geq \tau$.

Network-completion instances. Given the pre-completion network H_c^i, its node set is partitioned into reachable and unreachable nodes. A set of candidate instances to apply network completion, called *network completion instances*, are identified. Each such instance $\mathcal{S}$ consists of a tuple (t, S_t, H_t) where t is an unreachable *terminal*, $S_t = \{s_1, s_2, \ldots\}$ is a collection of reachable nodes called *sources*, and H_t, referred to as the *t-local graph*, is a subgraph of the grid graph G containing $\{t\} \cup S_t$, where each node $v \in V(H_t)$ has weight $1/w^i(v)$ if $1/w^i(v) > \alpha$ (a confidence threshold), 0 otherwise (See Algorithm 2). Only the nodes of V will be used for extending t to a reachable node. More implementation details are given in the Appendix. Let $\mathbf{S} = \{\mathcal{S}_1, \mathcal{S}_2, \ldots\}$ be the collection of such instances.

Algorithm 1: Directly Connected Canal Nodes

Input: M: A binary matrix of size $m \times n$
F: A set of water index pairs (i, j)
Output: C: Set of directly connected 1s
Create boolean mask B of size $m \times n$ with $B[i, j] = 1$ if and only if $(i, j) \in F$.
Define kernel K:

$$K = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

$V \leftarrow$ Convolve B with K;
Initialize empty set C;
foreach $(i, j) \in M$ **do**
  if $V[i, j] > 0$ ***and*** $M[i, j] = 1$ **then**
    Add (i, j) to C;
return C;

Identifying a network-completion instance. Given H_c^i, using morphological thinning [Fisher *et al.*, 2003] end points of canal segments are identified. An end point t is a terminal if it is not reachable. All pixels at distance ρ (user-specified) from t form the node set of its local graph H_v. The source set of v is the set of all end points in the local graph that are reachable.

Network reachability computation. For each instance $\mathcal{S} = (t, S_t, H_t) \in \mathbf{S}$, the objective is to find a minimum weighted shortest path from t to S_t in H_t. This involves converting the node-weighted H_t to an edge-weighted graph (See Algorithm 3), followed by the application of Dijkstra's [Dijkstra, 1959] shortest path algorithm. More implementation details are given in the Appendix B included in the Online Supplementary Material.

3.4 Framework Modules

Reachable and Unreachable Nodes. Given water source indices as F, Algorithm 1 identifies canal pixels directly connected to these sources. Given a binary matrix M representing the canal network (pre-completion H_c^i), the algorithm creates a boolean mask B, marking the locations of F. It then applies a convolution with an 8-connectivity kernel to determine the set of directly connected canal pixels, C, which are classified as reachable. A breadth-first search (BFS) expands from C to find all connected canal pixels, forming the set R of reachable pixels. The remaining canal pixels in M that are not in R are considered non-reachable U.

Terminals. Given the set of unreachable nodes U, we identify the terminal nodes. For each point $p \in U$, the algorithm examines its 8-connected neighborhood defined by the directions $\Delta = \{(0, \pm 1), (\pm 1, 0), (\pm 1, \pm 1)\}$. A point is classified as a terminal if it has one or fewer neighbors within the set U. The algorithm maintains a set V of visited points to avoid redundant computations and it returns the set of terminals E.

Source-Terminal Pairs. For each terminal $t \in E$, we identify source points S_t within a radius ρ. The sources include water source edge points (those with fewer than eight neighbors) and reachable canal nodes. Using an approach similar to Algorithm 5, we first detect water source edges and then determine source-terminal pairs by selecting all source points $p \in S$ that satisfy $\|p - t\|_2 \leq \rho$.

Algorithm 2: Edge Point Processing

Input: E: Set of terminal points
$w^i(v)$: Likelihood matrix
H_c^i: Pre-completion network
ρ: Sampling radius
α: Confidence threshold
Output: X_r: Resultant matrix
$X_r \leftarrow H_c^i$ `// Initialize with pre-completion network`
$(n, m) \leftarrow$ size of H_c^i `// Extract matrix dimensions`
foreach $p \in E$ **do**
  $N_p \leftarrow$ **GetNeighbors**$(p, \rho, (n, m))$;
  foreach $n \in N_p$ **do**
    if $w^i[n] > \alpha$ ***and*** $X_r[n] = 0$ **then**
      $X_r[n] \leftarrow \lfloor 1/w^i[n] \rfloor$;
return X_r;

Function `GetNeighbors` $(p, \rho, (n, m))$ **:**
  $(x, y) \leftarrow p$;
  $N \leftarrow \emptyset$;
  for $dx \in [-\rho, \rho]$ **do**
    for $dy \in [-\rho, \rho]$ **do**
      if $dx \neq 0$ ***or*** $dy \neq 0$ **then**
        $(nx, ny) \leftarrow (x + dx, y + dy)$;
        if $0 \leq nx < n$ ***and*** $0 \leq ny < m$ **then**
          $N \leftarrow N \cup \{(nx, ny)\}$;
  return N;

3.5 Metrics

Alongside conventional metrics—Precision (P), Recall (R), F1 Score (F1), and Intersection over Union (IoU)—*we introduce parameterized metrics* to address width inconsistencies in thin-structure segmentation. These metrics account for minor spatial misalignments in single-pixel annotations, such as canal networks represented by shapefile line segments, which may not perfectly align with the ground truth. Conventional metrics may underestimate performance in such cases, even when the model captures the overall structure. To address this, we define an r-neighborhood ($\mathcal{N}_r(i, j)$) around each pixel (i, j), allowing for small spatial deviations. Let $p_{i,j} \in \{0, 1\}$ be the predicted value at (i, j) and $y_{k,l} \in \{0, 1\}$ the ground truth at (k, l).

r-Neighborhood True Positives (rTP) is the number of predicted positive pixels that lie within the r-neighborhood of the actual positive pixels: $r\text{TP} = \sum_{i=1}^{N} \sum_{j=1}^{M} \max_{k,l \in \mathcal{N}_r(i,j)} (y_{k,l}) \cdot p_{i,j}$.

Algorithm 3: Directed Subgraph Around Terminal

Input: $M_{m\times n}$: Matrix representing the canal network
$t = (t_x, t_y)$: Terminal coordinates
r: Radius
Output: $G = (V, E)$: Directed graph
$R \leftarrow \{(i,j) \mid |i - t_x| \leq r, |j - t_y| \leq r, 0 \leq i < m, 0 \leq j < n\}$;
$\Delta \leftarrow \{(0, \pm 1), (\pm 1, 0), (\pm 1, \pm 1)\}$;
$V, E \leftarrow \emptyset$;
foreach $(i,j) \in R$ ***where*** $M[i,j] > 0$ **do**
  $V \leftarrow V \cup \{(i,j,1), (i,j,2)\}$;
  $E \leftarrow E \cup \{((i,j,1), (i,j,2), M[i,j])\}$;
  foreach $(d_x, d_y) \in \Delta$ **do**
    if $(i + d_x, j + d_y) \in R$ ***and*** $M[i + d_x, j + d_y] > 0$ **then**
      $E \leftarrow E \cup \{((i,j,2), (i + d_x, j + d_y, 1), 0), ((i + d_x, j + d_y, 2), (i,j,1), 0)\}$;
return $G = (V, E)$;

r-Neighborhood False Positives (rFP) are the number of predicted positive pixels that do not lie within the r-neighborhood of any actual positive pixels: $r\text{FP} = \sum_{i=1}^{N} \sum_{j=1}^{M} p_{i,j} \cdot \left(1 - \max_{k,l \in \mathcal{N}_r(i,j)} y_{k,l}\right)$.

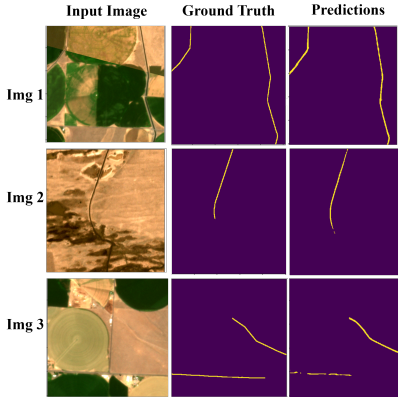
r-Neighborhood False Negatives (rFN) are the number of actual positive pixels for which there are no predicted positive pixels within the r-neighborhood: $r\text{FN} = \sum_{i=1}^{N} \sum_{j=1}^{M} y_{i,j} \cdot \left(1 - \max_{k,l \in \mathcal{N}_r(i,j)} p_{k,l}\right)$.

The r-IoU (rI), r-precision (rP), r-recall (rR), and r-F1 score (rF1) are similar to their conventional counterparts with TP, FP, and FN replaced by rTP, rFP, and rFN respectively. The usefulness of these metrics is demonstrated in Figure 3.

4 Experimental Setup

Canal Network Dataset. We used PlanetScope (2020–2023) [NASA, 2023] 3m-resolution RGB imagery to map irrigation canals in central Washington. NDWI was computed from Green and NIR channels, while USGS 3DEP Digital Elevation Map (DEM) (1m) [USGS, 2024] was resampled to 3m and used alongside NDWI and RGB. Canal waterway data came from the National Hydrography Dataset (NHD) [2020]. To prepare the dataset, we merged imagery tiles into a $\approx$ 20,000×30,000 tile per year and divided them into non-overlapping 512×512 patches. We filtered patches with over 30% black pixels and excluded mask patches with fewer than 0.5% canal pixels. We have total of 30,000 512x512x5 (RGB, NDWI, DEM) patches in our dataset.

For experiments, we created two distinct sets (Set 1 and Set 2) of spatially separated training, validation, and test data. The test set was 20% of the data, while 80% was used for training and validation. We employed 5-fold cross-validation and ran IGraSS to track average performance metrics. After tuning hyperparameters (iterations, epochs, ρ, etc.), we re-ran IGraSS on training data and evaluated on the test sets.



Model's Performance Evaluation

	Img 1	Img 2	Img 3
P	0.561	0.672	0.609
rP	**0.887**	**0.912**	**0.937**
R	0.808	0.808	0.505
rR	**0.982**	**0.982**	**0.734**
D	0.662	0.735	0.552
rD	**0.932**	**0.954**	**0.823**

Figure 3: Comparison of conventional and refined metrics in evaluating thin-structure segmentation. These examples illustrate how our refined metrics more accurately assess model performance by reducing reliance on pixel-level precision. While the model effectively captures overall structures, conventional metrics fail to fully reflect this capability. In contrast, our refined metrics provide a more nuanced evaluation, better representing the model's ability to identify key structural elements.

Road Network Data Following RoadTracer [Bastani *et al.*, 2018], we obtained 60 cm/pixel satellite imagery of New York City (24 sq km) from Google Maps and merged the tiles. We extracted the road network from OpenStreetMap, converted coordinates to match the imagery, and generated road masks. To create partial road maps, we iteratively removed α road segments of random length β from a predefined list (e.g., $\beta \in \{20, 30, 50, 100\}$). Further details, including train-test split, are in Appendix A.

5 Results

5.1 Irrigation canals

Our evaluation focuses on two main aspects: (1) comparing IGraSS's performance against the state-of-the-art models used as learners in our framework, and (2) assessing its ability to complete canal networks given reachability constraints. We also perform extensive experiments under various parameter settings to provide a comprehensive evaluation.

Segmentation Baseline Networks. For our experiments, we select three popular state-of-the-art models to serve as the Learner in our framework: DeepLabV3+ [Chen *et al.*, 2018], ResNet50 [He *et al.*, 2015], ResUNet [Diakogiannis *et al.*, 2019], and Swin Transformer [He *et al.*, 2022]. To the best of our knowledge, no work has been done on irrigation canal identification using deep learning and remote sensing images. Therefore, we use state-of-the-art models' performances as our baselines.

To assess the impact of iterative ground truth refinement on the model's overall performance, we conduct a systematic analysis. Intuitively, breaks or inconsistencies in the ground truth should negatively affect the model's performance. Conversely, as the quality of the ground truth improves through our iterative process, we expect to see a positive impact on the model's performance. This analysis aims to verify this hy-

Model	Test Set	w/ or w/o	P	rP	R	rR	F1	rF1	I	rI
ResUnet	1	w/o	0.591	0.838	0.540	0.612	0.564	0.708	0.531	0.601
		w	0.643	0.874	0.589	0.668	0.615	0.757	0.546	0.620
	2	w/o	0.549	0.802	0.521	0.599	0.535	0.686	0.560	0.630
		w	0.648	0.878	0.587	0.664	0.617	0.756	0.585	0.660
Deeplabv3+	1	w/o	0.613	0.805	0.580	0.765	0.596	0.780	0.420	0.600
		w	0.644	0.835	0.610	0.790	0.626	0.808	0.440	0.620
	2	w/o	0.605	0.798	0.570	0.758	0.587	0.773	0.415	0.590
		w	0.636	0.829	0.600	0.785	0.617	0.800	0.435	0.610
SwinTransformer	1	w/o	0.775	0.850	0.765	0.835	0.770	0.842	0.720	0.805
		w	**0.820**	**0.900**	**0.810**	**0.885**	**0.815**	**0.892**	**0.760**	**0.850**
	2	w/o	0.770	0.845	0.760	0.830	0.765	0.838	0.715	0.800
		w	**0.815**	**0.895**	**0.805**	**0.880**	**0.810**	**0.887**	**0.755**	**0.845**

Table 1: Performance comparison of IGraSS against baseline models on Canal Network datasets. Here 'w/' indicates the using IGraSS in conjunction with the baseline models, while 'w/o' represents training the baseline models without IGraSS for the same number of epochs.

pothesis and quantify the relationship between ground truth refinement and model's performance.

To establish a fair comparison, we train each baseline model independently for same number of epochs without implementing our framework. The training setup for these baselines is identical to the one used within our framework, ensuring consistency in our evaluation. The key distinction lies in the treatment of ground truth data. In our proposed method, the ground truth is updated after each iteration using the output from our framework. In contrast, the baseline models are trained using the original, unmodified ground truth throughout the entire process. We ran IGraSS framework for 5 iterations with a radius of 100 and an initial confidence threshold α of 0.2, which was later reduced to 0.01 for optimal result.

Performance evaluation by refining Ground Truth. Table 1 presents the performance evaluation on the two different test sets of the Canal Network Dataset across the 3 models used for training. The reported results in Table 1, demonstrate that training the model using refined ground truth from IGraSS significantly enhances the performance of all models across all metrics. Swin Transformer outperformed other models, as it breaks the image into small patches, computing self-attention locally while enabling cross-window connections. This helps capture the overall canal structure relationships within patches. For Swin Transformer, model trained with the IGraSS's refined ground truth improves precision from 0.775 to 0.820 (5.8% increase), recall from 0.765 to 0.810 (5.2% increase), F1-score from 0.770 to 0.815 (**5.8% increase**), and IoU from 0.720 to 0.760 (5.6% increase) on Test Set 1. Similar trends are observed in Test Set 2, with improvements of 5.8% in precision, 5.3% in recall, 5.9% in F1-score, and 5.6% in IoU. The refined ground truth also improves Deeplabv3+ performance by 5% and enhances ResUnet performance 10% across all metrics. These improvements are evident not only in our proposed metrics but also in conventional scores showing better ground truth refined by IGraSS, boost model's performance.

Network completion assessment. Figure 4 presents a quantitative analysis of the number of canal pixels that get

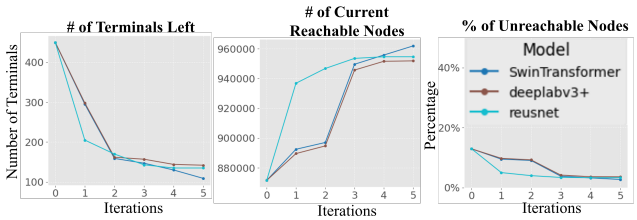


Figure 4: Network Completion Analysis with IGraSS.

connected in each iteration. All three models achieved similar results in connecting the terminals after 5 iterations. As we refine the main ground truth by connecting only the terminals to the nearest water sources, the number of reachable canals increases with each iteration, while the number of unreachable canal pixels and terminals decreases. In both cases IGraSS was able to reduce down the unreachable canals from $(18-15)\%$ to $(5-3)\%$. Upon manual analysis of the terminals that remained unreachable, we found that in most cases, there were either no water sources within the selected radius or the corresponding images did not contain any visible canals at all. This lack of visible canals might have led the neural network to fail in predicting any canal pixels.

5.2 Ablation Study

Effect of Framework Parameters. The IGraSS framework's performance is influenced by parameters such as intermediate epochs, radius ρ, threshold τ, and confidence threshold α. We experimented with $\rho \in \{20, 50, 100, 150\}$, $\alpha \in \{0.3, 0.2, 0.1, 0.01\}$, and epochs $\in \{10, 20, 30\}$ using SwinTransformer, DeepLabv3+, and ResUnet, evaluated over 5 iterations with K-Fold cross-validation. Lower α (0.01) with fewer epochs (10) reduces unreachable canals but introduces noise, while moderate α (0.1) with more epochs (20) yields cleaner results. Thresholds between 0.2 and 0.1 were generally effective. The number of epochs is crucial—too few with low α cause noise, while extended training risks learning noisy ground truth. An adaptive approach, lowering α after sufficient training, may be beneficial. Due to space

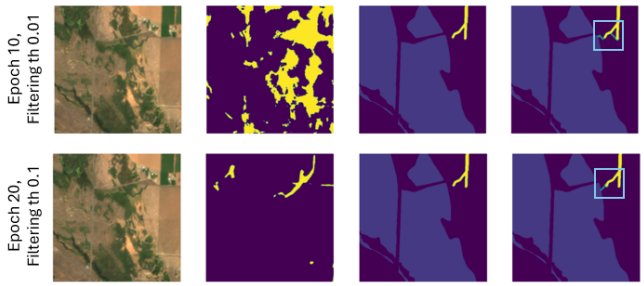


Figure 5: Error Results

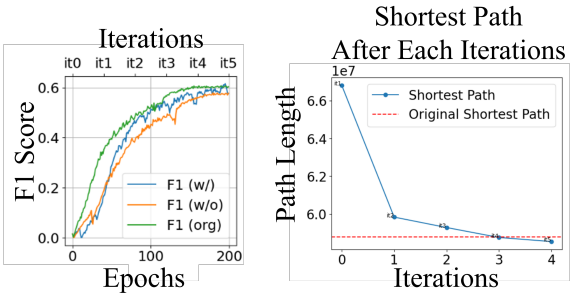


Figure 6: The plot illustrates how gaps in the ground truth (represented by the orange curve) impact model performance compared to the correct ground truth (shown by the green curve). The blue curve demonstrates improvement over time as the noisy ground truth is corrected using the IGraSS framework. The rightmost curve indicates how the shortest path length decreases as more breaks in the network are connected over time through the application of IGraSS.

limit, extensive parameter sensitivity analysis is presented in Appendix included in Online Supplementary Material (Tables S3 and S1; further details are in Section C).

Error Analysis As discussed in our parameter analysis, selecting appropriate parameters is crucial to avoid erroneous connections. IGraSS's focus on connecting points via the shortest path helps minimize errors when adding new data to the ground truth. Directly using the neural network output would have introduced significant noise to the ground truth, which our adaptive thresholding process helps mitigate. However, as illustrated in Figure 5, unwanted connections may still occur if the right parameters are not chosen.

5.3 Road Networks

To further demonstrate our framework's effectiveness and generalizability, we use it to complete road networks under a different graph-theoretic constraint; the objective in this case is to minimize the distance (on the network) between any pair of a user-specified set of points on the road network. We would like to emphasize here that our objective is to demonstrate the effectiveness of our IGraSS framework in satisfying graph-based constraints and not to show improved performance on these road networks which are already of high quality.

Problem Statement: Given a complete road network $G = (V, E)$, we introduce random gaps by removing a subset of edges $E_r \subset E$, resulting in an incomplete network $G' = (V, E \setminus E_r)$. For a set of N randomly sampled points $S = \{s_1, s_2, ..., s_N\} \subset V$, we compute the all-pairs shortest paths D_{GT} in the original network G and D_{pred} in the reconstructed network G^* produced by our model. The objective is to optimize G^* such that $\sum D_{pred} \leq \sum D_{GT}$, ensuring the total shortest path length in G^* is minimized while restoring connectivity. The process continues iteratively until convergence.

Results. Our results show that IGraSS effectively reduces shortest path lengths to match the ground truth by the third iteration (see Fig. S6). Since the network remains well-connected despite small gaps, reachability is not a suitable constraint; instead, shortest path minimization yields better results. Performance metrics exhibit similar trends with and without the framework, as observed in the canal network. Comparing final outputs with the ground truth confirms that our approach achieves comparable results across all metrics (see Table 2). Due to space constraints, extensive results are provided in the Appendix (Online Supplementary Materials).

Models	Method	rI	rI	rF1	rI
ResNet50	Original	0.782	0.852	0.816	0.698
	w/o	0.709	0.834	0.767	0.675
	w	0.752	0.846	0.797	0.691
ResUnet	Original	0.859	0.912	0.885	0.732
	w/o	0.825	0.899	0.859	0.699
	w	0.848	0.910	0.878	0.711
DeepLabV3+	Original	0.872	0.943	0.906	0.768
	w/o	0.846	0.929	0.885	0.709
	w	0.861	0.942	0.898	0.737

Table 2: Performance Analysis on Road Datasets.

6 Discussion

In summary, we developed a framework called IGraSS to address the challenges posed by weak and incomplete annotations by leveraging global constraints inherent in network-like infrastructures. We assessed IGraSS on canal networks with reachability constraints, successfully filling gaps in the ground truth and achieving improved results. IGraSS can be generalized to other network infrastructures with different constraints, as demonstrated with road networks. While we use DEM alongside RGB to enhance canal segmentation by providing elevation context, the refinement module currently relies solely on reachability. We acknowledge that performance may degrade in areas where canals are obscured by vegetation or terrain, and future work will explore integrating flow direction and additional multi-modal data to improve robustness. By enabling accurate and automated mapping of infrastructure networks, IGraSS supports sustainable water and land resource management—advancing SDG 12 by promoting more responsible and efficient agricultural planning and infrastructure development.

Ethical Statement

There are no ethical issues.

Acknowledgments

This material is based upon work supported by the AI Research Institutes program supported by NSF and USDA-NIFA under the AI Institute: Agricultural AI for Transforming Workforce and Decision Support (AgAID) award No. 2021-67021-35344. This work was partially supported by University of Virginia Strategic Investment Fund award number SIF160.

References

[Abdollahi *et al.*, 2020] Abolfazl Abdollahi, Biswajeet Pradhan, Nagesh Shukla, Subrata Chakraborty, and Abdullah Alamri. Deep learning approaches applied to remote sensing datasets for road extraction: A state-of-the-art review. *Remote Sensing*, 12(9):1444, 2020.

[Archuleta and Terziotti, 2023] Christy-Ann M. Archuleta and Silvia Terziotti. Elevation-derived hydrography—representation, extraction, attribution, and delineation rules. Technical report, U.S. Department of the Interior, U.S. Geological Survey, January 2023.

[Bastani *et al.*, 2018] Farzaneh Bastani, Songtao He, Miao Liu, Hamed Samet, and John Krumm. Roadtracer: Automatic extraction of road networks from aerial images. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4720–4728, 2018.

[Belt and Smith, 2009] Richard L. Belt and Stephen W. Smith. *Infrastructure Inventory and GIS Mapping for Canal Irrigation Delivery Systems*. U.S. Committee on Irrigation and Drainage, June 2009. Presented at Irrigation District Sustainability - Strategies to Meet the Challenges: USCID Irrigation District Specialty Conference, June 3-6, 2009, Reno, Nevada.

[Chen *et al.*, 2018] Liang-Chieh Chen, Yukun Zhu, George Papandreou, Florian Schroff, and Hartwig Adam. Encoder-decoder with atrous separable convolution for semantic image segmentation. *CoRR*, abs/1802.02611, 2018.

[Chen *et al.*, 2022] Xin Chen, Qun Sun, Wenyue Guo, Chunping Qiu, and Anzhu Yu. Ga-net: A geometry prior assisted neural network for road extraction. *International Journal of Applied Earth Observation and Geoinformation*, 114:103004, 2022.

[Cheng *et al.*, 2021] Mingfei Cheng, Kaili Zhao, Xuhong Guo, Yajing Xu, and Jun Guo. Joint topology-preserving and feature-refinement network for curvilinear structure segmentation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 7147–7156, 2021.

[Cira *et al.*, 2022] Calimanut-Ionut Cira, Martin Kada, Miguel-Ángel Manso-Callejo, Ramón Alcarria, and Borja Bordel Sanchez. Improving road surface area extraction via semantic segmentation with conditional generative learning for deep inpainting operations. *ISPRS International Journal of Geo-Information*, 11(1):43, 2022.

[Creaco *et al.*, 2023] E. Creaco, G. Barbero, A. Montanaro, et al. Effective optimization of irrigation networks with pressure-driven outflows at randomly selected installation nodes. *Scientific Reports*, 13:19218, 2023.

[Demir et al., 2018] Demir et al. Deepglobe 2018: A challenge to parse the earth through satellite images. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 172–181, 2018.

[Diakogiannis *et al.*, 2019] Foivos I. Diakogiannis, François Waldner, Peter Caccetta, and Chen Wu. Resunet-a: a deep learning framework for semantic segmentation of remotely sensed data. *CoRR*, abs/1904.00592, 2019.

[Dijkstra, 1959] Edsger W. Dijkstra. A note on two problems in connexion with graphs. *Numerische Mathematik*, 1(1):269–271, 1959.

[Fan *et al.*, 2023] Yu Fan, Haorui Chen, Zhanyi Gao, and Xiaomin Chang. Canal water distribution optimization model based on water supply conditions. *Computers and Electronics in Agriculture*, 205:107565, 2023.

[Fisher *et al.*, 2003] Robert Fisher, Simon Perkins, Ashley Walker, and Erik Wolfart. Hypermedia image processing reference (hipr2), 2003.

[Ganaye *et al.*, 2018] Pierre-Antoine Ganaye, Michaël Sdika, and Hugues Benoit-Cattin. Semi-supervised learning for segmentation under semantic constraint. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2018: 21st International Conference, Granada, Spain, September 16-20, 2018, Proceedings, Part III 11*, pages 595–602. Springer, 2018.

[Gharbia, 2023] Reham Gharbia. Deep learning for automatic extraction of water bodies using satellite imagery. *Journal of the Indian Society of Remote Sensing*, 51(7):1511–1521, 2023.

[He *et al.*, 2015] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition, 2015.

[He *et al.*, 2020] Songtao He, Favyen Bastani, Satvat Jagwani, Mohammad Alizadeh, Hari Balakrishnan, Sanjay Chawla, Mohamed M Elshrif, Samuel Madden, and Mohammad Amin Sadeghi. Sat2graph: Road graph extraction through graph-tensor encoding. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XXIV 16*, pages 51–67. Springer, 2020.

[He *et al.*, 2022] Xin He, Yong Zhou, Jiaqi Zhao, Di Zhang, Rui Yao, and Yong Xue. Swin transformer embedding unet for remote sensing image semantic segmentation. *IEEE Transactions on Geoscience and Remote Sensing*, 60:1–15, 2022.

[Hosseinzade *et al.*, 2017] Zeinab Hosseinzade, Sheree A. Pagsuyoin, Kumaraswamy Ponnambalam, and Mohammad J. Monem. Decision-making in irrigation networks:

Selecting appropriate canal structures using multi-attribute decision analysis. *Science of The Total Environment*, 601-602:177–185, 2017.

[Li *et al.*, 2022] Junjie Li, Yizhuo Meng, Yuanxi Li, Qian Cui, Xining Yang, Chongxin Tao, Zhe Wang, Linyi Li, and Wen Zhang. Accurate water extraction using remote sensing imagery based on normalized difference water index and unsupervised deep learning. *Journal of Hydrology*, 612:128202, 2022.

[Liu *et al.*, 2022] Pengfei Liu, Qing Wang, Gaochao Yang, Lu Li, and Huan Zhang. Survey of road extraction methods in remote sensing images based on deep learning. *Journal of Photogrammetry, Remote Sensing and Geoinformation Science*, 90(2):135–159, 2022.

[Loureiro *et al.*, 2024] D. Loureiro, P. Beceiro, E. Fernandes, et al. Energy efficiency assessment in collective irrigation systems using water and energy balances: methodology and application. *Irrigation Science*, 42:745–768, 2024.

[Nagaraj and Kumar, 2024] R Nagaraj and Lakshmi Sutha Kumar. Extraction of surface water bodies using optical remote sensing images: A review. *Earth Science Informatics*, 17(2):893–956, 2024.

[NASA, 2023] NASA. Commercial Smallsat Data Acquisition Program. https://earthdata.nasa.gov/esds/csdap, 2023. Accessed: 31 August, 2024.

[National Hydrography , 2020] National Hydrography . National Hydrography Dataset. https://www.usgs.gov/national-hydrography/national-hydrography-dataset, 2020. U.S. Department of the Interior.

[Pérez-Blanco *et al.*, 2020] C. Dionisio Pérez-Blanco, Arthur Hrast-Essenfelder, and Chris Perry. Irrigation technology and water conservation: A review of the theory and evidence. *Review of Environmental Economics and Policy*, 14(2):216–239, 2020.

[Ren *et al.*, 2022] Simiao Ren, Wayne Hu, Kyle Bradbury, Dylan Harrison-Atlas, Laura Malaguzzi Valeri, Brian Murray, and Jordan M Malof. Automated extraction of energy systems information from remotely sensed data: A review and analysis. *Applied Energy*, 326:119876, 2022.

[SpaceNet on Amazon Web Services (AWS), 2018] SpaceNet on Amazon Web Services (AWS). Datasets. https://spacenet.ai/datasets/, 2018. The SpaceNet Catalog. Last modified October 1st, 2018. Accessed on Sep 2, 2024.

[United Nations, 2015] United Nations. Transforming our world: The 2030 agenda for sustainable development. United Nations, 2015. Accessed: 2023-05-24.

[USGS, 2024] USGS. U.s. geological survey 3d elevation program (3dep), 1-meter digital elevation model (dem), 2024.

[Yu *et al.*, 2023] Jie Yu, Yang Cai, Xin Lyu, Zhennan Xu, Xinyuan Wang, Yiwei Fang, Wenxuan Jiang, and Xin Li. Boundary-guided semantic context network for water body extraction from remote sensing images. *Remote Sensing*, 15(17):4325, 2023.

[Zhang and Long, 2023] Liang Zhang and Cheng Long. Road network representation learning: A dual graph-based approach. *ACM Trans. Knowl. Discov. Data*, 17(9), jun 2023.

[Zhou *et al.*, 2018] Lichen Zhou, Chuang Zhang, and Ming Wu. D-linknet: Linknet with pretrained encoder and dilated convolution for high resolution satellite imagery road extraction. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 192–1924, 2018.