

GRAPH-BASED ACTIVE LEARNING FOR SURFACE WATER AND SEDIMENT DETECTION IN MULTISPECTRAL IMAGES

Bohan Chen^{1*}, Kevin Miller^{2†}, Andrea L. Bertozzi^{1‡}, Jon Schwenk^{3§}

¹ University of California, Los Angeles. Math Dept. 520 Portola Plaza, Los Angeles, CA 90095, USA.

² University of Texas, Austin. Oden Institute for Computational Engineering and Sciences.
201 E 24th St, Austin, TX 78712, USA.

³ Los Alamos National Laboratory. Earth & Env. Sci. Division. Los Alamos, NM 87545, USA.

ABSTRACT

We develop a graph active learning pipeline (GAP) to detect surface water and in-river sediment pixels in satellite images. The active learning approach is applied within the training process to optimally select specific pixels to generate a hand-labeled training set. Our method obtains higher accuracy with far fewer training pixels than both standard and deep learning models. According to our experiments, our GAP trained on a set of 3270 pixels reaches a better accuracy than the neural network method trained on 2.1 million pixels.

Index Terms— Rivers, Remote Sensing, Surface Water Detection, Graph Learning, Active Learning

1. INTRODUCTION

Surface water dynamics are critical to climate, flood monitoring and mitigation, freshwater resource management, water quality analyses, and geomorphology [1, 2, 3]. The importance of automated surface water detection in remote sensing is highlighted by the publication of global datasets or pre-trained surface water models, such as the Global Surface Water dataset [4], Surface Water Extent product [5], and Deep-WaterMap [6], in combination with different classification methods including support vector machine (SVM), random forest (RF), and convolutional neural networks (CNN).

Beyond just surface water, some river studies require mapping of in-channel sediment to identify their so-called “bankfull” state [7], which we define as the union of water

and active (unvegetated) in-channel sediment bars [8]. We build and employ our model on the small but high-quality, manually-labeled RiverPIXELS dataset [9] consisting of labeled water and in-river, unvegetated sediment from multi-spectral Landsat images.

Our main contribution is a graph-based active learning pipeline (GAP) to identify water and sediment pixels in multispectral images. The inclusion of the sediment class has largely been omitted in previous efforts or treated as an independent problem (e.g., [10]). Our model is based on the RiverPIXELS dataset to classify water and in-channel sediment pixels in multispectral images. The first challenge of this problem is the paucity of training data. The RiverPIXELS dataset does not have enough images to train an accurate CNN model to classify pixels. We address this insufficiency by graph learning, which is good at classification tasks with a low label rate [11]. The second challenge is designing an efficient graph learning approach. The RiverPIXELS dataset contains millions of labeled pixels. Using all available pixels would result in extreme computational demand with long run-times. We implement an active learning approach to condense the training set by selecting representative pixels accounting for approximately 0.1% of the total number of labeled pixels. Compared to methods in other products, our GAP achieves the highest accuracy with minimal training data (i.e., the number of labeled pixels). All our codes are available on GitHub¹.

2. GRAPH ACTIVE LEARNING PIPELINE

The graph active learning pipeline (GAP) is summarized in Figure 1). We extract a non-local means feature vector corresponding to a neighborhood of each pixel in a given image, then construct a similarity graph $G = (X, W)$ based on the feature vectors $X = \{x_1, x_2, \dots, x_N\}$, where W is the edge weight matrix generated according to KNN angular similarity. With the ground-truth labels on a subset $X_L \subset X$, we apply graph Laplace learning [12] to predict the classification

*Supported by the UC-National Lab In-Residence Graduate Fellowship Grant L21GF3606

†Supported by the Department of Defense (DOD) National Defense Science and Engineering Graduate (NDSEG) Fellowship

‡ALB is supported by the National Geospatial-Intelligence Agency under Award No. HM0476-21-1-0003. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Geospatial-Intelligence Agency. Approved for public release, NGA-U-2023-01028.

§Supported by the Laboratory Directed Research and Development program of Los Alamos National Laboratory under project numbers 20170668PRD1 and 20210213ER

¹<https://github.com/wispcarey/SurfWater-Graph-Active-Learning-GAP->

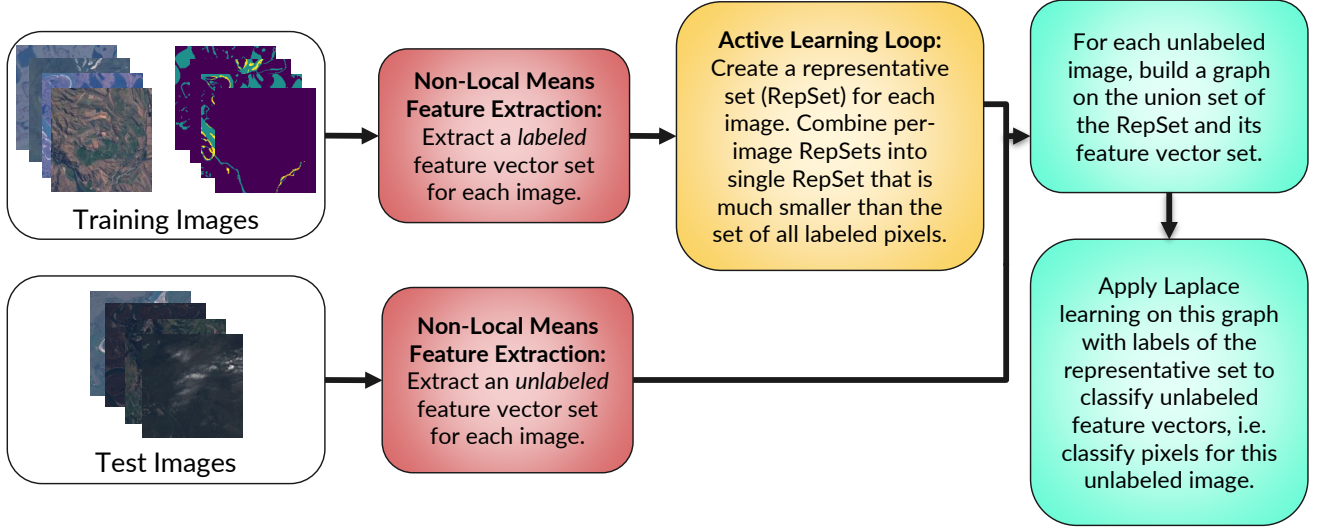


Fig. 1. The flowchart of our GAP: 1. (Red Boxes) Apply the non-local means method to extract feature vectors. 2. (Yellow Box) Condense the labeled feature vector set into a smaller representative set (RepSet) through active learning. 3. (Cyan Box) Build a graph based on the union of the RepSet and the unlabeled feature set. Then, apply graph learning approaches to predict labels for unlabeled feature vectors.

of the unlabeled nodes. Active learning approaches are applied to select the labeled set X_L based on the Model-Change (MC) acquisition function $\mathcal{A} : X - X_L \rightarrow \mathbb{R}$ [13, 14]. A similar method for an image segmentation task is introduced in [15].

The first step to classifying pixels is to generate a feature vector for each pixel. For pixel i , consider a $(2k+1) \times (2k+1)$ neighborhood patch centered at pixel i (for boundary pixels, use the reflection padding). Inspired by the non-local means method [16], we use the flattened Gaussian weighted neighborhood patch as the non-local means feature vector [15] of pixel i . Then we extract the non-local means feature vector for each image in the training image set \mathcal{I} and select a subset of labeled feature vectors as our training set.

Selecting a reasonable X_L is tricky. Each image in River-PIXELS is of the size 256×256 , and the whole dataset includes millions of labeled pixels. It is thus inefficient to build a graph using all of the labeled pixels. Instead, we apply an active learning approach to select $X_L = R$, where R is a subset of representative labeled feature vectors, called the representative set (RepSet). We find a representative set R_j for each image $I_j \in \mathcal{I}$ and combine them together to obtain $R = \cup_{i=1}^{l_n} R_i$. With the following four steps (Algorithm 1), we determine the representative set for a certain image $I_i \in \mathcal{I}$:

1. Preprocess: Apply the non-local means method to get the feature set $X_i = \{x_j^i\}$ of image I_i .

2. Initialization: Initialize the representative set R_i^0 by randomly selecting the same number of feature vectors of each class from the feature set X_i . This initialization requires all ground-truth labels of feature vectors in X_i and gives a class-balanced initial R_i^0 .

Algorithm 1 Create the Representative Set

INPUT: Image set $\mathcal{I} = \{I_1, I_2, \dots, I_l\}$.

OUTPUT: Representative set \mathcal{R} of the training image set \mathcal{I} .

$\mathcal{R} \leftarrow \emptyset; i \leftarrow 1$

FOR $i = 1$ to l , DO:

Preprocess image I_i to construct the feature vector of each pixel

Initialize the representative set R_i^0 for image I_i

Active learning loop until **Termination** to obtain final R_i

$\mathcal{R} \leftarrow \mathcal{R} \cup R_i$

Algorithm 2 Active learning loop

INPUT: The initialized representative set R^0 and feature vector set X .

OUTPUT: The representative set R of the feature vector set X .

Build a graph G on $R^0 \cup X; t \leftarrow 0$

WHILE **terminal condition** is NOT triggered:

Apply Laplace learning on G with labels of R^t

Get the predicted labels of X

Calculate the acquisition function $\mathcal{A}(x_k), \forall x_k \in X \setminus R^0$

$k^* \leftarrow \arg \max_{x_k \in X \setminus R^0} \mathcal{A}(x_k)$

$R^{t+1} \leftarrow R^t \cup \{x_{k^*}\}; t \leftarrow t + 1$

$R \leftarrow R^t$

3. Active learning loop: Using active learning approaches, add feature vectors from X_i and corresponding labels one by one to the representative set (Algorithm 2).

4. Termination: Stop the active learning loop when a certain termination condition is satisfied. The accuracy-based terminal condition is applied. In iteration t of the **active learning loop** step for image I_i , we apply Laplace learning based on labels of the current representative set R_i^t to predict other nodes

$X_i - R_i^t$. Denote the prediction accuracy on $X_i - R_i^t$ by a_t . Given a positive integer K_{\max} and a positive real number $\epsilon > 0$, the active learning loop terminates if $|a_t - a_{t-1}| < \epsilon$ or $t > K_{\max}$.

In summary, we extract non-local means feature vectors for the training images, then apply the active learning method to condense them into a small RepSet R . For each test image \tilde{I}_i (considered unlabeled), we combine its extracted feature vector set \tilde{X}_i with R to build a graph. Finally, we apply graph Laplace learning to predict labels on \tilde{X}_i . This node classification gives the segmentation of image \tilde{I}_i .

3. EXPERIMENTS AND RESULTS

We choose five rivers from the RiverPIXELS dataset: the Kolyma, Yana, Waitaki, Colville, and Ucayali Rivers. Our pipeline is trained on a small subset of pixels chosen from the first four rivers, while the performance of different methods is tested on the remaining pixels of all five rivers. There are 42 images belonging to the first four river regions, while the Ucayali River includes 54 images. We randomly sample 75% of the labeled data (32 images) for each region as set \mathcal{I} from which we develop the smaller training set (RepSet) R and use the remaining 25% (10 images) as the test set $\tilde{\mathcal{I}}$. In addition, we employ an extra test set $\tilde{\mathcal{I}}_{\text{ex}}$ formed by 54 images of the Ucayali river. Of note is that the Ucayali river is in the tropical Amazonian region, while the other four are within temperate or arctic regions, creating an additional challenge for the method.

Model performances are evaluated using metrics of the **Boundary Accuracy** and **Overall Accuracy**. We define a pixel's boundary distance D_B as the minimal distance to a pixel with a different ground-truth label. The **Boundary Accuracy** $BA(d)$ is defined as the accuracy on pixels with a boundary distance $D_B \leq d$. The **Overall Accuracy** (OA) is the accuracy on all pixels. BA is more indicative of the model performance than OA since a naive classifier that classifies all pixels into the land will have an OA of around 80%.

We compare the classification performance of our GAP to DeepWaterMap (DWM) [6], support vector machine (SVM) [17] and random forest (RF) [18] models. For our GAP, we extract 7×7 non-local means features of pixels in the training image set. The original extracted feature set $X \subset \mathbb{R}^{294}$ has over 2 million feature vectors from the training set consisting of 32 labeled images. DWM only provides the classification of water and land pixels, while RiverPIXELS images include water, bare sediment, and land. We provide two approaches to compare the performance between other methods and DWM. The first approach is to *retrain DeepWaterMap* (DWM_R). We train a new neural network with the same structure of DWM on our training set with 32 labeled images and labels of water, sediment, and land. The second approach is to *modify labels*. Inspection of the training set of the original DeepWaterMap (DWM_O) shows that nearly all sediment pixels are

labeled as land. We modify the labels of our training set and the ground-truth labels of our test set by changing sediment labels to land labels. Each method uses a different amount of training data chosen to optimize the performance of each method. For SVM and RF, we use a training set named T-NLM consisting of 42.6K labeled non-local means feature vectors randomly sampled and balanced in each class. The retrained DWM (DWM_R) is trained on \mathcal{I} while the original DWM (DWM_O) is trained on around 100K labeled images.

Table 1 shows our experiment results, including the training and test set information and the BA(3), BA(10), and OA accuracy metrics under both the retrained DWM method and the modified label (Sed→Land). Note that our GAP method is trained on the smallest training set yet has the highest boundary accuracies, BA(3) and BA(10), as well as the highest overall accuracy on both the test set $\tilde{\mathcal{I}}$ and the extra test set $\tilde{\mathcal{I}}_{\text{ex}}$. Note that our GAP method still performs the best across different regions (extra test set). Figure 2 displays results for a sampled image from the Ucayali river that includes cloud cover.

4. CONCLUSION

We develop a graph active learning pipeline (GAP) to classify pixels into the land, water, and bare sediment classes from the RiverPIXELS dataset. GAP outperforms the classical methods of support vector machine and random forest, and a cutting-edge CNN approach (DeepWaterMap) trained on a dataset hundreds of times larger than RiverPIXELS. With the help of active learning techniques, GAP can be trained on much smaller datasets than other methods. Furthermore, GAP demonstrates superior performance on images of rivers in different environments relative to other approaches. Our work demonstrates how graph-based active learning techniques can provide efficient labeling of and greater accuracy in detecting surface water and in-river sediment using significantly fewer training data than classic and deep approaches.

5. REFERENCES

- [1] Paul D Bates, Jefferey C Neal, Douglas Alsdorf, and Guy J-P Schumann, "Observing global surface water flood dynamics," *The Earth's Hydrological Cycle*, pp. 839–852, 2013.
- [2] Sarah W Cooley, Jonathan C Ryan, and Laurence C Smith, "Human alteration of global surface water storage variability," *Nature*, vol. 591, no. 7848, pp. 78–81, 2021.
- [3] Tan Chen, Chungqiao Song, Linghong Ke, Jida Wang, Kai Liu, and Qianhan Wu, "Estimating seasonal water budgets in global lakes by using multi-source remote sensing measurements," *Journal of Hydrology*, vol. 593, pp. 125781, 2021.
- [4] Jean-François Pekel, Andrew Cottam, Noel Gorelick, and Alan S Belward, "High-resolution mapping of global surface water and its long-term changes," *Nature*, vol. 540, no. 7633, pp. 418–422, 2016.
- [5] John W Jones, "Improved automated detection of subpixel-scale inundation—revised dynamic surface water extent (DSWE) partial surface water tests," *Remote Sensing*, vol. 11, no. 4, pp. 374, 2019.

Method	Training Set		Test Set		Retrain DWM(%)			Sed→Land(%)		
	Dataset	Size	Dataset	# of Images	BA(3)	BA(10)	OA	BA(3)	BA(10)	OA
GAP (ours)	RepSet	3.27K Pixels	$\tilde{\mathcal{I}}$	10	81.90	90.41	95.50	83.75	91.66	96.30
SVM	T-NLM	42.6K Pixels			79.28	88.05	93.55	82.77	91.08	95.41
RF	T-NLM	42.6K Pixels			77.40	86.84	93.31	81.50	89.63	95.07
DWM_R	\mathcal{I}	32 Images			72.56	84.56	92.86	N\A	N\A	N\A
DWM_O	\mathcal{I}_{DWM}	100K Images			N\A	N\A	N\A	78.02	88.81	95.11
GAP (ours)	RepSet	3.27K Pixels	$\tilde{\mathcal{I}}_{ex}$	54	83.07	92.93	97.48	85.08	94.17	98.21
SVM	T-NLM	42.6K Pixels			79.08	90.29	96.26	77.49	90.96	97.23
RF	T-NLM	42.6K Pixels			26.41	26.79	12.52	76.01	89.09	96.21
DWM_R	\mathcal{I}	32 Images			69.28	83.43	94.39	N\A	N\A	N\A
DWM_O	\mathcal{I}_{DWM}	100K Images			N\A	N\A	N\A	79.42	91.14	97.29

Table 1. Comparison of different methods (GAP, SVM, RF, DWM_R, and DWM_O), where each is evaluated by boundary accuracies BA(3), BA(10), and the overall accuracy (OA). The training set size of our GAP, SVM, and RF is the number of pixels used in the training set, while that of DWM_R and DWM_O is the number of images in the training set since DWM is a CNN-based method. Each image contains 65536 pixels. The performance of different methods is evaluated on both the test set $\tilde{\mathcal{I}}$ and the extra test set $\tilde{\mathcal{I}}_{ex}$, in a different region from the training set \mathcal{I} . Our GAP method has the best performance according to these accuracy values.

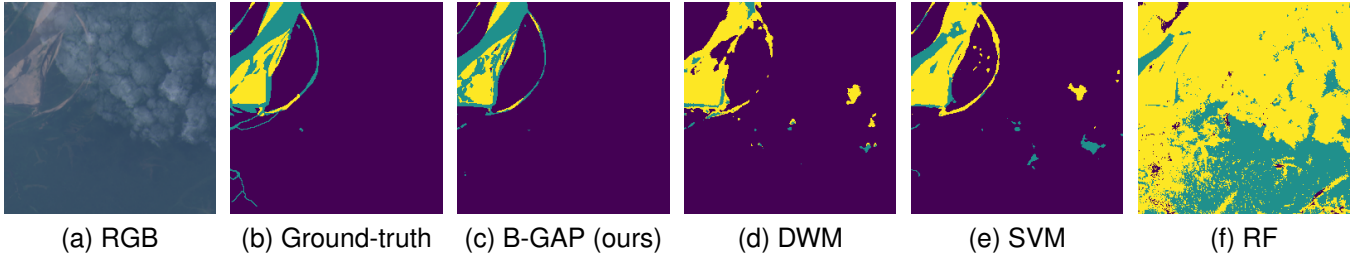


Fig. 2. Results for a Patch of Ucayali river. Original Patch name: *Ucayali_River_I 2018-09-11 006 066 L8 316 landsat*. This image has some clouds on the top. In panel (b)–(f), purple, blue, and yellow are land, water, and sediment respectively.

- [6] Leo F Isikdogan, Alan Bovik, and Paola Passalacqua, “Seeing through the clouds with deepwatermap,” *IEEE Geoscience and Remote Sensing Letters*, vol. 17, no. 10, pp. 1662–1666, 2019.
- [7] David M Bjerklie, “Estimating the bankfull velocity and discharge for rivers using remotely sensed river morphology information,” *Journal of hydrology*, vol. 341, no. 3-4, pp. 144–155, 2007.
- [8] Jon Schwenk, Ankush Khandelwal, Mulu Fratkin, Vipin Kumar, and Efi Foufoula-Georgiou, “High spatiotemporal resolution of river planform dynamics from Landsat: The RivMAP toolbox and results from the Ucayali river,” *Earth and Space Science*, vol. 4, no. 2, pp. 46–75, 2017.
- [9] Jon Schwenk and Joel C. Rowland, “RiverPIXELS: Paired Landsat images and expert-labeled sediment and water pixels for a selection of rivers v1.0. a global, high-resolution river network model for improved flood risk prediction,” Tech. Rep., ESS-DIVE Deep Insight for Earth Science Data, 2022, <https://data.ess-dive.lbl.gov/view/doi:10.15485/1865732>.
- [10] Federico Monegaglia, Guido Zolezzi, Inci Güneralp, Alexander J Henshaw, and Marco Tubino, “Automated extraction of meandering river morphodynamics from multitemporal remotely sensed data,” *Environmental Modelling & Software*, vol. 105, pp. 171–186, 2018.
- [11] Kevin Miller, John Mauro, Jason Setiadi, Xoaquin Baca, Zhan Shi, Jeff Calder, and Andrea L. Bertozzi, “Graph-based active learning for semi-supervised classification of SAR data,” 2022, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference on Defense and Commercial Sensing, <https://arxiv.org/abs/2204.00005>.
- [12] Xiaojin Zhu, Zoubin Ghahramani, and John D Lafferty, “Semi-supervised learning using Gaussian fields and harmonic functions,” in *Proceedings of the 20th International conference on Machine learning (ICML-03)*, 2003, pp. 912–919.
- [13] Kevin Miller, Hao Li, and Andrea L. Bertozzi, “Efficient graph-based active learning with probit likelihood via Gaussian approximations,” in *ICML Workshop on Experimental Design and Active Learning*. July 2020, International Conference on Machine Learning (ICML), arXiv: 2007.11126.
- [14] Kevin Miller and Andrea L Bertozzi, “Model-change active learning in graph-based semi-supervised learning,” *arXiv preprint arXiv:2110.07739*, 2021.
- [15] Bohan Chen, Kevin Miller, Andrea L. Bertozzi, and Jon Schwenk, “Batch active learning for multispectral and hyperspectral image segmentation using similarity graphs,” 2023.
- [16] Antoni Buades, Bartomeu Coll, and J-M Morel, “A non-local algorithm for image denoising,” in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’05)*. IEEE, 2005, vol. 2, pp. 60–65.
- [17] Corinna Cortes and Vladimir Vapnik, “Support-vector networks,” *Machine learning*, vol. 20, no. 3, pp. 273–297, 1995.
- [18] Tin Kam Ho, “Random decision forests,” in *Proceedings of 3rd international conference on document analysis and recognition*. IEEE, 1995, vol. 1, pp. 278–282.