Deep Learning in Sea Ice Remote Sensing Image Segmentation

Kangqi Fu
Department of Computer Science
AMEP (Applied Math, Physics, & Engineer)
University of Wisconsin Madison
kfu9@wisc.edu

Abstract

This paper works on sea ice remote sensing image segmentation using two widely used deep neural network, U-net and Seg Net. It also presents a new labeling methods for generate sufficient number of training images and corresponding masks. This new labeling methods can be further widely used for remote sensing deep learning application due to its ability in resolving training data scarcity problems. In the instruction section, the background of the deep learning in sea ice study will be introduced. The methodology section introduces the new labeling methods and give a brief review of the two neural networks. This section will also talk about the training process. The result section provides the image segmentation result of two models and then discuss the results both quantitatively using evaluation metric and qualitatively using output images. Information about the future work of this paper will be provided in the conclution and future work section.

1 INTRODUCTION

Sea ice is one of the most critical components of the whole Earth system. It floats on the ocean's surface and covers about seven percent of the Earth's surface and about 12 percent of the world's ocean [1]. It governs the high-latitude surface radiation balance and atmosphere-ocean exchanges of heat, moisture, and momentum [2]. Moreover, due to the effect of global warming, polar regions have become the most sensitive areas to climate change on earth [3]. Therefore, a scientific method capable of accurately simulating and predicting sea ice states is highly required.

Traditional model used for simulating ice floe includes continuum model. However this model has issues because it does not account for the small scale structure of sea ice. Another efforts are underway to discrete element methods, or discrete element models (DEM) [4][5], in order to represent sea ice as individual chunks of ice. While DEM for sea ice may represent better physics, it is also very computationally expensive.It is only in recent years that DEM for sea ice has computationally feasible due to the progress of supercomputers and developments of algorithms and code which gives good parallel performance [5]. All those traditional model are highly based on techniques from computational mathematics and physical dynamics. In other words, to analyze and predict ice floe dynamics or behaviors, researchers using traditional methods need to pre-program lots of known dynamics equations, such as translational momentum balance and angular momentum balance for individual ice floe, or develop many mathematically based equation solvers, such as those used in [6] and [7], then create virtual models instead of fully utilized nature based directly observations from remote sensing images or videos captured via satellite.

As the increasing popularity and maturity of deep learning and other machine learning techniques, traditional models may be refined by these new techniques. However, as mentioned in [8], although machine learning has achieved notable success in building compute vision and speed recognition model using sequential and spatial data, applications to remote sensing and other climate science problems are not very commonly developed or even seldom considered. The reasons are in many aspects includes the complexity of accurately modeling giant climate systems and the scarcity of accessible and reliable data for training models.

Some of the recent successful machine learning applications on sea ice modeling, such as [8] in developing an unsupervised learning model to track ice sheet and glacier dynamics, [9] in producing ice-describing semantic maps of the polar regions of the Earth, [10] in doing automatic SAR ice concentration estimation using deep CNN followed by a stochastic fully connected random field, and [11] in the use of deep learning in emulating sub-grid processes to simplify models of cloud at a part of the high resolution physics models computation cost, provides opportunities for doing further exploration in sea ice modeling using machine learning techniques.

In this work, my main goals are (1) introducing a new dataset created using satellite data and an algorithm that can create more accurate labels; (2) doing image segmentation using this new created dataset and two popular deep CNNs, U-net [12] and Seg-net [13]; and (3) introducing the future and next step of this project.

2 METHODOLOGY

2.1 Data Collection and Labeling

One contribution in this work here is the Data Collection and Labeling process for creating the training set. The images used here are from NASA WORLDVIEW [14]. The world view tool from NASA's Earth Observing System Data and Information System (EOSDIS) provides the capability to interactively browse over 900 global, full-resolution satellite imagery layers and then download the underlying data by simply crop the specific area that the users want to analyze. Many of the imagery layers are updated daily and are available within three hours of observation - essentially showing the entire Earth as it looks "right now". Since the focus here is ice floe, therefore the layer chosen here is arctic area. Then images of ice floe are obtained by taking a snapshot of the specific area. The whole processes of getting images can be learnt from the tutorial provided by [15]. In this work, 20 images with 250m resolution (per pixel) are taken from this Satellite front end.

The next thing to do for this dataset is to create the corresponding groundtruth label. As shown in Figure 1, the specific algorithm used for this labeling process is GVF snake algorithm [16]. This algorithm is a modified version of K-means method for doing image segmentation because of its strong ability in both ice edge detection and ice shape enhancement. The whole algorithm starts from doing ice edge detection by initializing the contours that are adapted to the floe size and then run on each contour of each ice floe to identify the floe boundary. Then in the ice shape enhancement step, because of the noise of the images, some floes may have holes or smaller ice pieces inside. To smoothen the shape of the ice floe in images, morphological cleaning [17], is chosen. The morphological cleaning process contains morphological closing, which tends to smooth the contours of objects, generally joins narrow breaks, fills long thing gulfs, and fills holes smaller than the structuring element, and morphological opening [18] removes complete regions of an object that cannot contain the structuring element, smooths object contours, break thin connections, and removes thin protrusions.

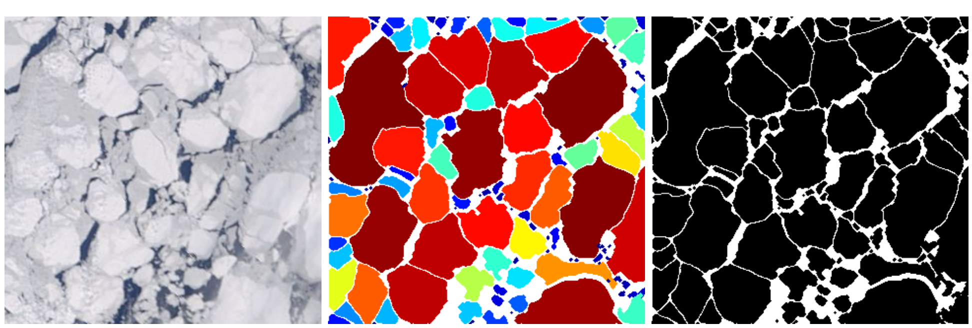

Figure 1: GVF snake labeling: From left to right: Original Image, RGB Label, Binary Label

To present a better visualization of the sizes, which is defined as the pixel number of each ice piece, each ice piece are labeled in different colors based on formula provided in [16]. For details of the algorithm for both ice floe boundary detection, ice shape enhancement and colorizing formulas, check attached supplementary material. Codes for this whole Data Collection, Labeling and the implementation of the algorithms are provided in the attached source code folder.

2.2 Image Segmentation

The training process is proceed by using two most widely cited models [12][13]. Because of the popularity and empirical nature of the these two models, the descriptions are in brief and high level.

The first model is U-Net[12] which was introduced for neuronal structure segmentation in electron microscopic images and won the 2015 ISBI challenge. The name comes from its U shape (Fig 2) and it combines a contraction path with a symmetric upsampling expansion path. U-net can be trained with few training samples because of its U shape fully convolutional architecture which can both utilize contextual information and achieve good localization. The difference between U-net and traditional FCN is that successive layers, where pooling operators are replaced by upsampling operators, are supplemented to the usual contracting network. These successive layers can increase the resolution of the output. Because remote sensing images from satellite are usually highly blurred due to clouds and other noises, the ability of enhancing resolution of images makes U-net a very capable model in doing image segmentation here.

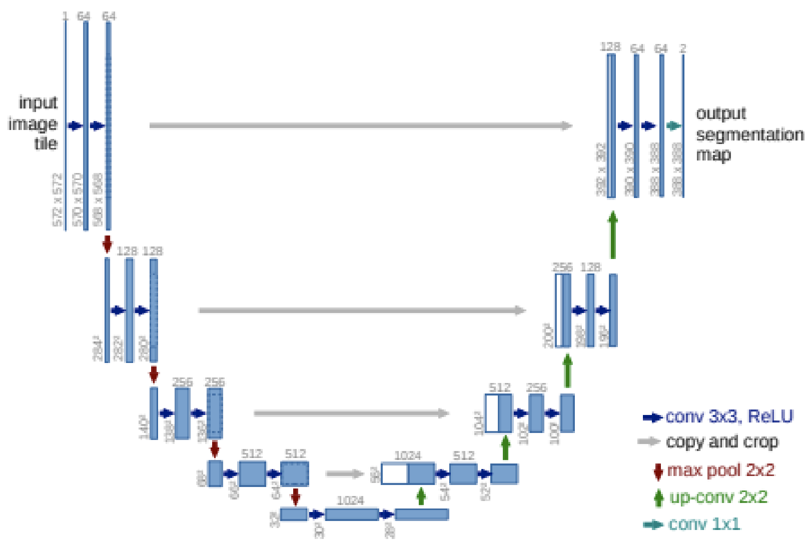

Figure 2: U net architecture [12]

The second model is Segnet [13] which was designed for segmenting natural images of outdoor and indoor scenes for scene understanding applications. This Segnet is in an auto-encoder structure and it's very similar to U-net however, the contracting and expanding parts are here termed as encoder and decoder, and the upsampling units in Segnet are not trainable. It instead shares weights with the corresponding max-pooling layers.

To prevent overfitting, dropout layers are also added in the models using here with drop out rate to be 0.1 or 0.2. The implementation of this two models are in python Tensorflow Keras framework and are both provided in the attached source code. The architecture detailing in the implementations are provided in the supplementary materials and here only provides the number of parameters of two models provided by the Keras model summary in Table 1.

2.3 Data Augmentation and Training

Data augmentation is used here to produce more training images. The method here is to use each image and it's corresponding label to do random rotation, skew, distortion, shear and flip to generate

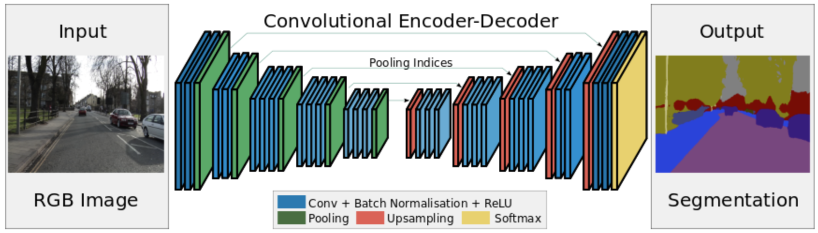

Figure 3: Segnet architecture [13]

Table 1: Number of training parameter counts for U net and Segnet

Model	trainable parameter / total parameter
UNet	31,031,745 / 31,031,745
SegNet	29,443,077 / 29,458,951

more images. Using this data augmentation techniques, the dataset is expanded from 20 manually obtained sea ice floe images from NASA WorldView to 1200 images. Here we randomly pick 4 of out 20 manually obtained images as test set and the rest of the 16 images are training set. Among 1200 images, 1000 are for training which means they are augmented using those 16 training images. The rest of the 200 images are for the test set. Figure 3-5 is an example of the results of the augmentation of one specific image.

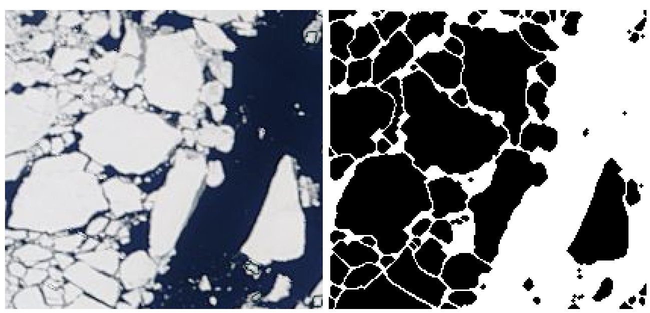

Figure 4: Original Image-Label Pair

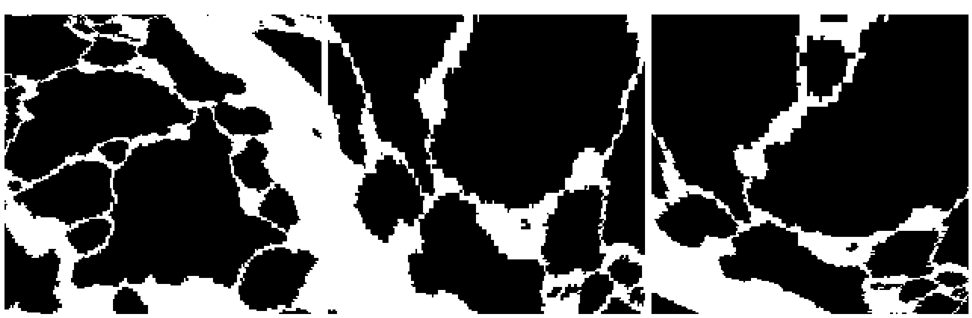

Figure 5: Augmented Corresponding Label

UNet and SegNet were both trained for 300 epochs but with early stopping to prevent overfitting and the training and validation accuracies and loss are evaluated after each epoch. About overfitting, the patience is set to be 40 and the monitor is set to be validation loss. The best training model will be stored during the training process and also the optimization approach using here is stochastic gradient descent which is the most commonly used for training deep convolutional neural network. Because the work here is just to demonstrate the idea, therefore, the label used here is the binary version which

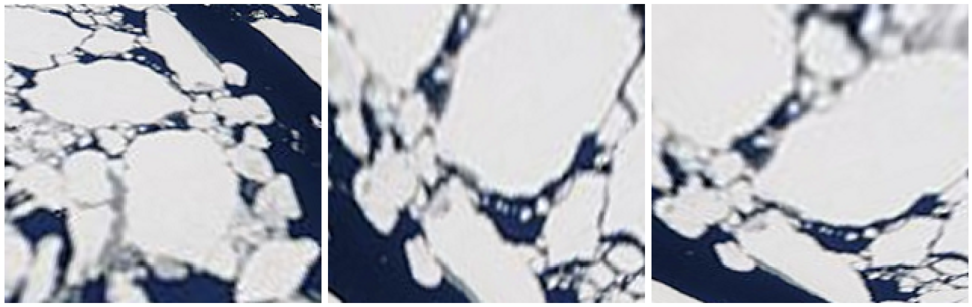

Figure 6: Augmented Image

just contains information about the difference between ocean(in white color) and ice floe (in black color). However, in the future work, the colored labels can also be used to train the model because the information then using colored labels will also include the identification of the size of each ice floe. This case will be furthered discussed in the later section of this paper. Because the binary label is chosen for training and testing here, the loss function that need to be optimized here is the binary cross-entropy loss and the accuracy is the corresponding accuracy. The parameters updating details during the training process are in the supplementary material.

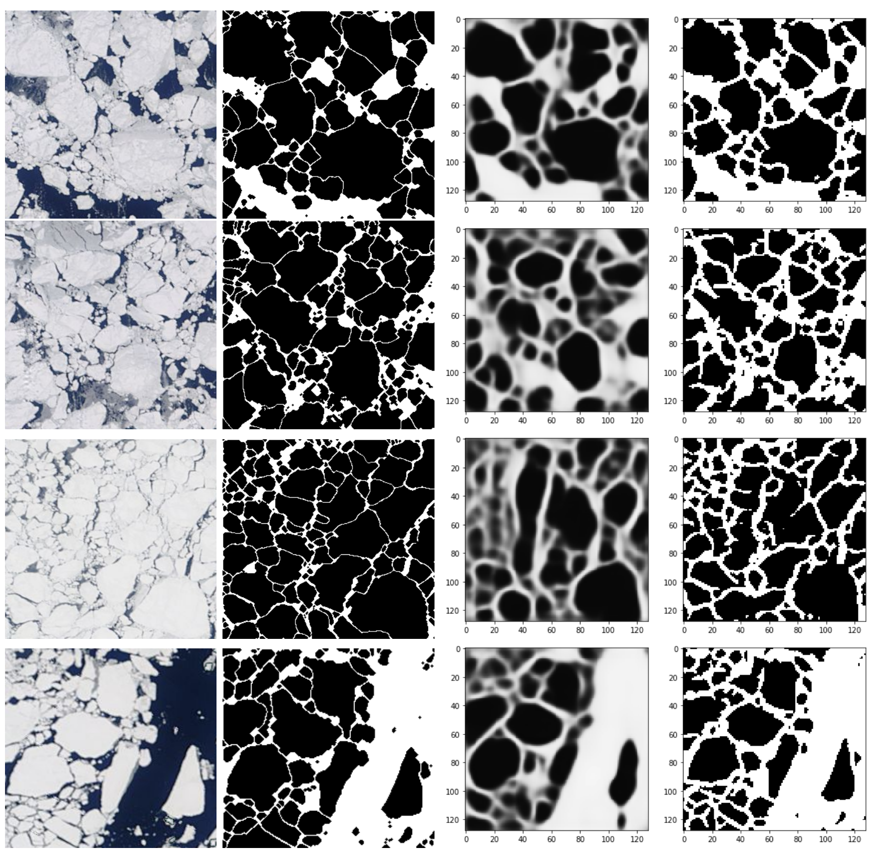

Figure 7: Test set Image Segmentation Result

3 Results

3.1 Evaluation Metrics

The evaluation metrics implemented here are four very commonly used in image segmentation [19]. They are Pixel accuracy, Mean accuracy, Mean IOU, and Frequency Weighted IOU:

(1) Pixel accuracy:

$$pa = \frac{\sum_i n_{ii}}{\sum_i t_i}$$

(2) Mean accuracy:

$$ma = \frac{1}{n} \sum_i \frac{n_{ii}}{t_i}$$

(3) Mean IOU

$$miou = \frac{1}{n} \sum_i \frac{n_{ii}}{t_i + \sum_j n_{ji} - n_{ii}}$$

(4) Frequency Weighted IOU

$$fwiou = \sum_k (t_k)^{-1} \sum_i \frac{t_i n_{ii}}{t_i + \sum_j n_{ji} - n_{ii}}$$

Here, n is the number of classes, n_{ij} is the number of pixels of class i predicted to belong to class j and t_i is the total number of pixels of class i in the ground truth. Furthermore, the accuracy measures only the true positive rate while IOU accounts for false positives. Therefore, accuracy and IOU are actually equivalent to recall and precision in image segmentation.

	Pixel Accuracy	Mean Accuracy	Mean IOU	Frequency Weighted IOU
Image 1	0.790	0.783	0.646	0.655
Image 2	0.727	0.726	0.566	0.574
Image 3	0.680	0.662	0.497	0.525
Image 4	0.794	0.784	0.651	0.657

Table 2: Evaluation metrics of SegNet

	Pixel Accuracy	Mean Accuracy	Mean IOU	Frequency Weighted IOU
Image 1	0.931	0.926	0.866	0.871
Image 2	0.871	0.864	0.765	0.772
Image 3	0.865	0.838	0.740	0.762
Image 4	0.898	0.895	0.812	0.815

Table 3: Evaluation metrics of UNet

3.2 Quantitative and Qualitative Results of the Two Models

Here we use the only four original images (not after augmentation, just the original 4) in the test set, the trained model, and the evaluation metrics shown above to get both Quantitative and Qualitative Results of U net model and Segnet model.
In Figure 7 is the test set image segmentation result. From Left to right is the real satellite image, binary groundtruth image, Segnet segmentation image, and U net segmentation image. From first row to the last row is test image 1 to test image 4.

In Table 2 and 3 are the evaluation metrics result of the four test images in two different models, and this qualitatively demonstrate that the U-net has a better performance in this ice floe image segmentation due to its special architecture and its ability in generate higher resolution outputs.

4 Conclusion

In this paper, the main work is to give a new idea of labeling ice floe dataset using real Satellite images and use two widely used deep neural network framework to work on and test this labeling remote sensing sea ice dataset. The overall performance indicates that the U net has a better image segmentation outcome than Segnet in this special cases and therefore the next step of this work will be implemented using U net framework.

5 Future Work

The future work includes using the data labeling techniques mentioned in this paper to work on sequential remote sensing sea ice images by either implement a LSTM CNN or develop another deep learning framework. From the real images provided in the past, this new work will predict the future behavior of the sea ice then generate a video based segmentation which is more suitable for further analyzing the sea ice dynamics system.

Reference

[1] Wadhams, P. (2006). How Does Arctic Sea Ice Form and Decay? *Pmel Arctic Zone*. Retrieved from National Oceanic and Atmospheric Administration Website:
`https://www.pmel.noaa.gov/arctic-zone/essay_wadhams.html`

[2] Roberts AF, Hunke EC, Allard R, Bailey DA, Craig AP, Lemieux J-F, Turner MD. 2018 Quality control for community-based sea-ice model development.Phil. Trans. R. Soc. A 376: 2017.0344.
`http://dx.doi.org/10.1098/rsta.2017.0344`

[3] NOAA. How does sea ice affect the global climate? National Ocean Service Website,
`https://oceanservice.noaa.gov/facts/sea-ice-climate.html`, 11/15/19

[4] U.S. Department of Energy Office of Science. (2017). *A new Discrete Element Sea-Ice Model for Earth System Modeling*. Retrieved from U.S. Department of Energy Office of Science Website:
`https://climatemodeling.science.energy.gov/projects/new-discrete-element-sea-ice-model-earth-system-modeling`

[5] DEMSI (2018). Distrcrete Element Model for Sea Ice [PowerPoint slides]. Retrieved from:
`https://www.orau.gov/scidac4pi2018/presentations/9-BER/05TurnerA_BER_Discrete_Element_Model_for_Sea_Ice.pdf`

[6] Herman, A.: Discrete-Element bonded-particle Sea Ice model DESIgn, version 1.3a – model description and implementation, Geosci. Model Dev., 9, 1219–1241, `https://doi.org/10.5194/gmd-9-1219-2016`, 2016.

[7] Damsgaard, A., Adcroft, A.,& Sergienko, O. (2018).Application of discrete elementmethods to approximate sea icedynamics. *Journal of Advances in Modeling Earth Systems*, 10, 2228–2244.
`https://doi.org/10.1029/2018MS001299`

[8] Min, Y., Mukkavilli, K.S. & Bengio, Y. (2019) Predicting ice flow using machine learning. In *33rd Conference on Neural Information Processing Systems (NeurIPS), Workshop on Tackling Climate Change with Machine Learning, Vancouver, Canada, 2019*. arXiv: 1910.08922

[9] Baumhoer, C.A., Dietz, A.J., Kneisel, C. & Kuenzer, C. (2019) Automated Extraction of Antarctic Glacier and Ice Shelf Fronts from Sentinel-1 Imagery Using Deep Learning. In *Remote Sens*. 2019, 11(21), 2529; `https://doi.org/10.3390/rs11212529`

[10] Wang, L., A. Wong, D. A. Clausi, A. K. Scott, L. Xu, M. J. Shafiee, &F. Li. (2015) Sea Ice concentration estimation from satellite SAR imagery using convolutional neural network and stochastic fully connected conditional random field. In *CVPR 2018 Earthvision Workshop*. Retrieved from: `https://uwaterloo.ca/vision-image-processing-lab/publications/sea-ice-concentration-estimation-satellite-sar-imagery-using`

[11] Rasp, Stephan, Pritchard, Michael S, & Gentine, Pierre. (2018) Deep learning to represent subgrid processes in climate models. Proceedings of the National Academy of Sciences, 115(39):1–6, 2018. doi: 10.1073/pnas.1810286115.

[12] Ronneberger, O., Fischer, P. & Brox, T. (2015) U-Net: Convolutional Networks for Biomedical Image Segmentation. In *MICCAI (3)*, volume 9351 of *Lecture Notes in Computer Science*, pages 234-241. arXiv:1505.04597

[13] Badrinarayanan, V.,Kendall, A. & Badrinarayanan, V. (2017) SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 12, pp. 2481-2495, 1 Dec. 2017, doi: 10.1109/TPAMI.2016.2644615.

[14] EOSDIS Worldview. Retrieved from: `https://worldview.earthdata.nasa.gov/`

[15] EARTHDATA. Retrieved from: `https://earthdata.nasa.gov/worldview`

[16] Zhang, Q & Skjetne, R. (2015) Image processing for identification of sea-ice floes and the floe size distributions. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 53, No. 5, pp. 2913-2924, 2015.

[17] L.-K. Soh, C. Tsatsoulis, and B. Holt, “Identifying ice floes and computing ice floe distributions in SAR images,” in Analysis of SAR Data of the Polar Oceans. Berlin, Germany: Springer-Verlag, 1998, pp. 9–34.

[18] R. C. Gonzalez, R. E. Woods, and S. L. Eddins, Digital Image Processing Using MATLAB. Upper Saddle River, NJ, USA: Prentice-Hall, 2003.

[19] E. Shelhamer, J. Long, and T. Darrell. Fully convolutional networks for semantic segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 39(4):640–651, 2017.