

# Semi-supervised Convolutional Neural Networks for Flood Mapping using Multi-modal Remote Sensing Data

Viet-Hung Luu<sup>1,3</sup>, Minh-Son Dao<sup>2</sup>, Thi Nhat-Thanh Nguyen<sup>1</sup>, Stuart Perry<sup>3</sup>, and Koji Zettsu<sup>2</sup>

<sup>1</sup>Vietnam National University of Engineering and Technology  
Hanoi, Vietnam

<sup>2</sup> Big Data Integration Research Center, National Institute of Information and Communications Technology  
Tokyo, Japan

<sup>3</sup> School of Electrical and Data Engineering, University of Technology Sydney  
New South Wales, Australia

**Abstract**—When floods hit populated areas, quick detection of flooded areas is crucial for initial response by local government, residents, and volunteers. Space-borne polarimetric synthetic aperture radar (PolSAR) is an authoritative data sources for flood mapping since it can be acquired immediately after a disaster even at night time or cloudy weather. Conventionally, a lot of domain-specific heuristic knowledge has been applied for PolSAR flood mapping, but their performance still suffers from confusing pixels caused by irregular reflections of radar waves. Optical images are another data source that can be used to detect flooded areas due to their high spectral correlation with the open water surface. However, they are often affected by day, night, or severe weather conditions (i.e., cloud). This paper presents a convolution neural network (CNN) based multi-modal approach utilizing the advantages of both PolSAR and optical images for flood mapping. First, reference training data is retrieved from optical images by manual annotation. Since clouds may appear in the optical image, only areas with a clear view of flooded or non-flooded are annotated. Then, a semi-supervised polarimetric-features-aided CNN is utilized for flood mapping using PolSAR data. The proposed model not only can handle the issue of learning with incomplete ground truth but also can leverage a large portion of unlabelled pixels for learning. Moreover, our model takes the advantages of expert knowledge on scattering interpretation to incorporate polarimetric-features as the input. Experiments results are given for the flood event that occurred in Sendai, Japan, on 12<sup>th</sup> March 2011. The experiments show that our framework can map flooded area with high accuracy ( $F1 = 96.12$ ) and outperform conventional flood mapping methods.

**Index Terms**—flood extent, polarimetric feature, semi-supervised, PolSAR, CNN.

## I. INTRODUCTION

The risk of flood disaster to economic damages and human casualties is an issue of mounting concern all over the world [1], [2]. Remote sensing systems have been widely used to rapidly but costly monitor a large area in case of natural disasters [3], [4]. Assessing flood areas can be done using space-borne optical and Synthetic Aperture Radar (SAR), among which SAR plays a vital role due to its ability to work independently of day-night and weather conditions [3], [5]. Flood extent mapping using SAR data has been widely

studied over the last decade thanks to the advantages of many newly launched satellites (i.e., German TerraSAR-X, Italian COSMO-SkyMed, ESA Sentinel-1) and space-borne platforms (i.e., Japan PiSAR, US AIRSAR). The preliminary mapping of flood extent is a pre-requisite for assessing several flood impacts such as inundation depth, flow velocity, debris factor [6]. Several algorithms have been proposed to flood mapping using SAR data and can be roughly divided into two main categories: unsupervised and supervised.

Among unsupervised algorithms, thresholding not only is an efficient and most widely used method but also probably the most straightforward approach of flood mapping, as the backscatter values of flat water on SAR images are relatively low [1], [7]–[11]. This approach based on the assumption that data follows bi-modal distribution (water vs. non-water), and all pixels whose values lower than the threshold are considered as a water class. Thresholding algorithm can be applied globally in the whole SAR image, resulting in a single threshold value or tile-based, resulting in adaptive threshold values for each sub-region of the image. Unfortunately, both strategies suffer from the same drawbacks as they do not consider the spatial context of the image pixels, and strongly depend on the bi-modal distribution assumption. On other hands, the algorithm will not work if the proportion of the flooded area is either too small or too large [7]. Another unsupervised approach that is popular in flood mapping domain leverages change detection, which involves the analysis of two images acquired over the same area [9], [12], [13]. One of these two images must be taken during the flood event, while the other can be a pre-flood or post-flood image. Based on these two images, a difference image is generated and used as an input for the classification process, which in the case is usually thresholding technique [7].

Supervised classification approaches can take a single-pixel (pixel-based) or cluster of pixels (object-based) as a unit of analysis. Pixel-based algorithms assign label for each pixel in the SAR image by calculating spectral, and texture features of grid blocks around target pixel [14]–[17]. Meanwhile, the

object-based algorithm first segment an image into constituent regions according to a similarity criterion, and then assign a label to the whole region based on spectral, texture, and shape features [18], [19]. Most of the works working on a supervised classification area are based on various conventional machine learning algorithms such as artificial neural network [16], random forest [15], [20], k-Nearest Neighbour [17] or late ensemble of these models [17].

Deep convolutions neural network (CNN) can be considered as a particular case of machine learning where a massive data is required to train a complex multi-layer neural network. In [3], the authors split TerraSAR-X image into non-overlapping patches with a size of  $32 \times 32$  pixels and then trained a supervised CNN model to classify them into three classes: flooded open areas, flooded built-up areas, and non-flooded areas. Flood extent mapping can be regarded as a pixel-wise semantic segmentation problem where only two classes are considered (water and non-water). While this method has been considered as the effective one among many tasks such as medial image segmentation [21], land cover classification [22], to the best of our knowledge, no existing works have been documented for flood extent mapping from SAR image. One of the significant challenges is that to train a CNN model require a large amount of data. Unfortunately, annotating flooded areas in optical and SAR images suffers difficulty. As flood events often take place during periods of extended cloud coverage, the exploitation of optical data may become questionable [20]. Meanwhile, SAR data contains a large number of confusing pixels caused by irregular reflections of radar waves, which make it challenging to annotate a pixel as water or non-water. Several works have been proposed to deal with the limitation of data annotation described above. In [23], the authors proposed a semi-supervised approach based on a teacher-student paradigm for general image classification problem. They can leverage billions of unlabelled images along with a relatively smaller set of task-specific labeled data. In [21], the authors proposed an active learning approach to address the issue of learning with incomplete ground truth in medical image segmentation. Their model is trained by ignoring non-annotated pixels during initial epochs and automatically relabelling missing annotations for use in the later epoch.

Inspired by these works, in this paper, we propose a CNN-based approach for rapidly flood extent mapping using polarimetric SAR (PolSAR) data with reference training data retrieved from the timely close optical-derived inundation maps (see Fig. 1). To accelerate the speed of data annotation process, as well as to reduce the affection of weather conditions, only areas in an optical image with high confidence (i.e., unobstructed view of flood or non-flood with no cloud cover or fog cover) are annotated. A semi-supervised training strategy based on teacher-student paradigm is proposed to handle the incomplete annotation issues and leverage a large portion of unlabelled pixels. Moreover, expert knowledge of target scattering mechanism interpretation is incorporated to enhance the performance of CNN model.

Overall, this paper makes the following contributions:

- We explore the multi-modal approach for flood mapping by utilizing both space-borne optical images and SAR images.
- We analyze semi-supervised deep learning in a flood mapping application and show that a simple semi-supervised strategy is valid even if the number of training samples is small.
- We explore PolSAR features based on heuristic knowledge of target scattering mechanism.
- We demonstrate the performance of our method and show that it outperforms traditional approaches. To the best of our knowledge, this is the first work utilizing the CNN model for flood mapping using multi-modal remote sensing data.

This paper is organized as follows. Our proposed method is described in Sect. II, followed by experimental results in Sect. III, and conclusions in Sect. IV.

## II. METHODOLOGY

Fig.1 illustrates the overall workflow of our proposed approach. In this section, we are going to describe our method from a general view to technical details in the following subsections.

### A. Annotation

Annotating of SAR data is done by using high-resolution optical images as a reference. Unfortunately, as the optical image is strongly affected by cloud cover and severe weather conditions, many pixels are ambiguous. In this paper, only pixels with high confidence are assigned a label of *water* or *non-water*, while the rest are left unlabelled. As a consequence, the collected annotations contain incomplete annotation as illustrated by **Original Set** in Fig.1.

Fig.2 shows an example of how the optical image captured one day after may help to identify flooded areas. Since the optical image captured on the same day (Fig.2.a) with PolSAR data (Fig.2.c) may be cloudy, the optical images of following days (Fig.2.b) are used to verify the flooded areas (Fig.2.d). We assume that if an area was flooded on the following days, it should be flooded on the target day.

### B. Polarimetric features

PolSAR data consists of three bands, including HH, VH, VV. Different from optical data, PolSAR data can be interpreted using specific knowledge of target scattering mechanism. Therefore, direct implementation of deep CNN using HH, VH, VV may limit the performance of the model [25]. In this section, we explore the polarimetric features to assist our model training.

1) *Polarimetric Scattering Power*: For PolSAR, the scattering matrix is defined as:

$$[S_{HV}] = \begin{bmatrix} S_{HH} & S_{HV} \\ S_{VH} & S_{VV} \end{bmatrix} \quad (1)$$

assuming  $S_{HV} = S_{VH}$ .

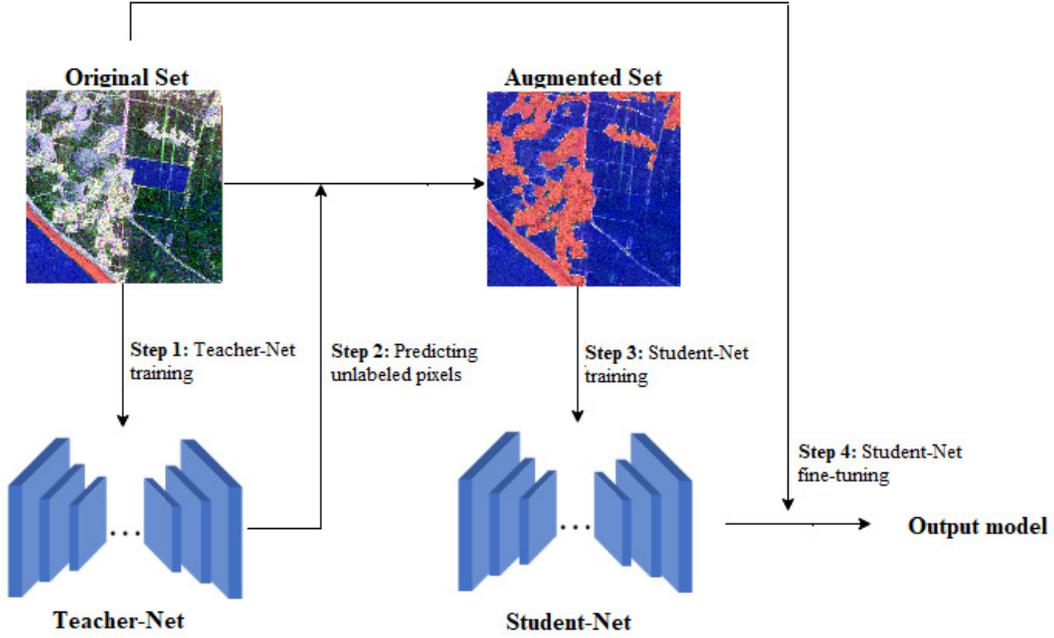


Fig. 1: The proposed approach. Teacher-Net is trained on Original Set manually labeled, then used to predict the class label for unlabelled pixels to form Augmented Set. Student-Net is trained on Augmented Set and fine-tuned with original Original Set. The label of the dataset is as follow: Red: *non-water*, Blue: *water*, no color: *unlabelled*.

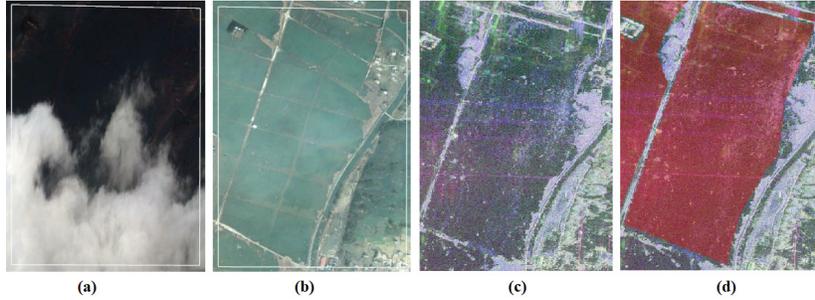


Fig. 2: The proposed annotation procedure. (a), (c) optical image and PolSAR image captured on 12<sup>th</sup> March 2011, respectively (b) optical image captured on 14<sup>th</sup> March 2011 (d) Annotation of flooded-areas

The polarimetric coherency matrix is formed as:

$$\langle [T^{HV}] \rangle = \langle k_P k_P^* \rangle = \begin{bmatrix} T_{11} & T_{12} & T_{13} \\ T_{21} & T_{22} & T_{23} \\ T_{31} & T_{32} & T_{33} \end{bmatrix} \quad (2)$$

where  $k_p = \frac{1}{\sqrt{2}} [ S_{HH} + S_{VV} \quad S_{HH} - S_{VV} \quad 2S_{HV} ]$  is the Pauli scattering vector, superscript \* denotes conjugate transpose.

We can expand the coherency matrix as sum of scattering coherency matrices:

$$\langle [T^{HV}] \rangle = f_s [T]_s^{hv} + f_d [T]_d^{hv} + f_v [T]_v^{hv} + f_c [T]_h^{hv} \quad (3)$$

where  $[T]_s^{hv}$ ,  $[T]_d^{hv}$ ,  $[T]_v^{hv}$ ,  $[T]_h^{hv}$  are surface, double, volume,

and helix scattering coherency matrix, respectively.

$$[T]_s^{hv} = \begin{bmatrix} 1 & \beta^* & 0 \\ \beta & |\beta|^2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (4)$$

$$[T]_d^{hv} = \begin{bmatrix} |\alpha|^2 & \alpha & 0 \\ \alpha^* & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (5)$$

$$[T]_v^{hv} = \frac{f_v}{4} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (6)$$

$$[T]_h^{hv} = \frac{f_c}{2} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & \pm j \\ 0 & \pm j & 1 \end{bmatrix} \quad (7)$$

where  $f_s, f_d, f_c, f_v, \alpha, \beta$  are six unknowns which can be found using four-component decomposition algorithm proposed in [24].

Finally, the scattering powers  $P_s, P_d, P_v,$  and  $P_c$  corresponding to surface, double, volume, and helix scattering, respectively, are determined by:

$$P_s = f_s (1 + |\beta|^2) \quad (8)$$

$$P_d = f_d (1 + |\alpha|^2) \quad (9)$$

$$P_v = f_v \quad (10)$$

$$P_c = f_c \quad (11)$$

2) *Roll-invariant Features*: The roll-invariant polarimetric features of entropy  $H$ , mean alpha angle  $\bar{\alpha}$  and anisotropy  $ani$  are proved to enhance the performance of PolSAR image classification [25]. In this work, we also adopt these features.

Similar to coherency matrix, the covariance matrix is defined as:

$$\langle [C^{HV}] \rangle = \langle k_P k_P^* \rangle = \begin{bmatrix} C_{11} & C_{12} & C_{13} \\ C_{21} & C_{22} & C_{23} \\ C_{31} & C_{32} & C_{33} \end{bmatrix} \quad (12)$$

where  $k_p = [S_{HH} \ S_{HV} \ S_{VV}]^T$ .

Because the Covariance matrix is a  $3 \times 3$  positive definite Hermitian matrix, it has real eigenvalues  $\lambda$ :

$$\langle [C] \rangle = [U] \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} [U]^* \quad (13)$$

where  $U$  is the unitary matrix.

The complexity of the scattering mechanism is called entropy  $H$ , and is defined as:

$$H = \sum_{i=1}^3 -P_i \log_3 P_i \quad (14)$$

where  $P_i = \frac{\lambda_i}{(\lambda_1 + \lambda_2 + \lambda_3)}$ .

The alpha angle is defined as:

$$\bar{\alpha} = \sum_{i=1}^3 P_i \alpha_i \quad (15)$$

where  $\alpha_1 = 0, \alpha_2 = \frac{\pi}{2}, \alpha_3 = \frac{\pi}{2}$ .

Finally, the anisotropy is defined as:

$$A = \frac{(\lambda_2 - \lambda_3)}{(\lambda_2 + \lambda_3)} \quad (16)$$

### C. Semi-supervised learning

To leverage a large portion of unlabelled pixels in our dataset, semi-supervised training pipeline is proposed, which consists of four-stage as follows:

- 1) The **Teacher-Net** is trained on *Original Set*.
- 2) The **Teacher-Net** is then used to automatically predict labels of remaining unlabelled pixels in the *Original Set*. Only pixels with their prediction confidence higher

than the predefined threshold  $T$  are used to form the *Augmented Set*.

- 3) A new model, namely **Student-Net**, is trained on the *Augmented Set* to take advantage of larger size of training data.
- 4) Finally, the **Student-Net** is fine-tuned on *Original Set*. It ensures that our final model is fine-tuned with clean labels.

1) *CNN Architecture*: The scattering powers  $P_s, P_d, P_v,$  and  $P_c$  and roll-invariant features  $H, \bar{\alpha},$  and  $A$  are normally preferred in PolSAR data interpretation and processing. In this work, these values are adopted as additional input bands together with the back-scattering matrix of  $HH, VV,$  and  $HV$ . The architecture of both Teacher-Net and Student-Net are identical to U-Net [26].

2) *Dice loss with missing annotations*: Based on Dice Similarity Coefficient, Dice loss [27] is widely used for semantic segmentation tasks. Considering the imbalance between the number of *water* and *non-water* samples, we introduce weighted Dice loss (WDL), defined as:

$$WDL = 1 - \frac{2 \sum_i^K w_i p_i g_i}{\sum_i^K p_i^2 + \sum_{k=1}^K g_i^2} \quad (17)$$

where  $K$  is the number of classes,  $p_i$  is the predicted probability of class  $i$ ,  $g_i$  is the binary ground truth (0 or 1), and  $w_i$  is the weight for class  $i$ .

Traditionally, Dice loss is computed by sum (or mean) of every pixels in the image. Since our training images are not fully annotated, unlabelled pixels are ignored during loss computation and do not contribute to the back propagation.

## III. EXPERIMENTS AND RESULTS

In this section, dataset, evaluation metrics, comparisons and discussion that are used to evaluate our method are introduced.

### A. Experiment setup

1) *Data Collection*: PolSAR image was collected by PiSAR-2 system over Sendai, Japan on 12<sup>th</sup> March 2011, one day after the tsunami happens at the area. Optical images for label reference are obtained from Google Earth on three days of 12<sup>th</sup>, 13<sup>th</sup>, and 14<sup>th</sup> March 2011.

10 PiSAR-2 images covering Sendai area are collected, in which 7 images are used for *training+validating*, and 3 images are used for *testing*. Images are divided into patches of size  $512 \times 512$  pixels. In order to increase the number of training data, *training+validating* patches are overlapped by 256 pixels.

2) *Dataset Statistics*: In total, the number of image patches for training, validating, and testing are 496, 29, and 66, respectively. Fig.3 shows the histogram of of annotation percentage for training set. The average proportion of annotated pixels is 27.37%. While some image patches are fully annotated, the minimum annotation proportion is 0.0015% which is equivalent to the area of  $\approx 40$  pixels in the image.

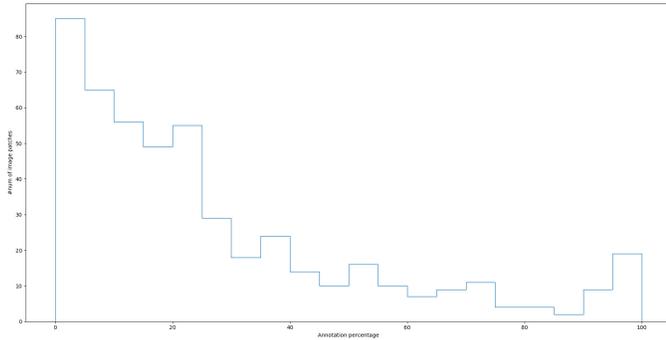


Fig. 3: Histogram of annotation percentage for training set

TABLE I: Flood mapping results comparison.

Method	F1 score
Semi-supervised CNN (proposed)	<b>96.12</b>
Supervised CNN (Teacher-Net)	94.03
Otsu on HH [8], [11], [29] (baseline)	90.33

3) *Training Details*: Both Teacher-Net and Student-Net are trained using Adam optimizer [28], weights are randomly initialized and updated with the learning rate set by 0.0001, momentum parameter set by 0.9, and weight decay set by 0.001. Learning rate is reduced by a factor of 0.1 when training accuracy has stopped improving for ten epochs. Class weights  $w_i$  for WDL in Eq.17 are set as 0.33 and 0.67 for *non-water* and *water*. During training, image patches are randomly flip horizontal and flip vertical.

### B. Flood Mapping Comparisons

We compare our proposed approach to the widely used conventional flood mapping algorithm, which is tile-based Otsu thresholding [29]. Otsu thresholding is applied on band HH as suggested by many studies [8], [11]. Moreover, to prove the effectiveness of the proposed semi-supervised strategy, we also compare our proposed model (Student-Net) with supervised CNN trained with incomplete annotations (Teacher-Net).

Quantitative comparisons are summarized in Table I. As expected, CNN-based approaches perform better than widely-used conventional approaches. The improvement is even more significance for the proposed semi-supervised approach, which result is 96.12% in term of F1 score. It is interesting to see that, our proposed model can learn a strong feature extractor network regarding that only a small proportion of data ground truth was provided.

Finally, we give in Fig. 4 and Fig. 5 the final flood mapping results for all methods in some test images. Fig. 4 represents the case where Otsu-based algorithm fails to work. In Fig. 4, only small top-right area is flooded (see ground truth map in Fig. 4.a), while other areas are non-flooded. Unfortunately, due to the SAR scattering mechanism, many areas in the image appear darker and are recognized as flooded by Otsu-based algorithms. In operation, in order to make Otsu work, the extensive pre-selection of areas with a high probability of flood

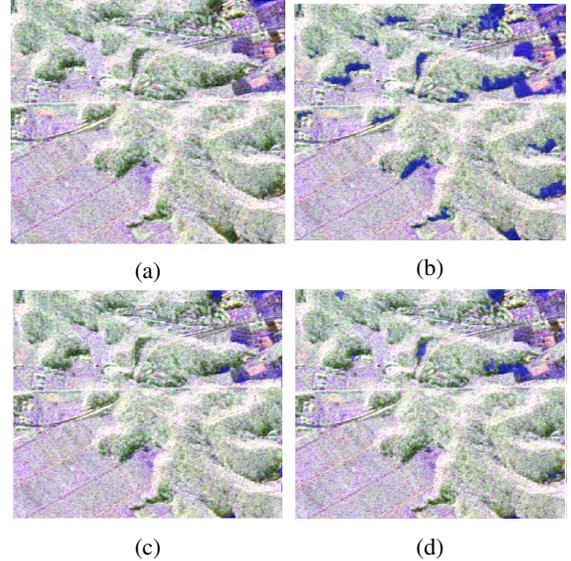


Fig. 4: Visualization comparison of generated flood map in mostly non-flooded region. (a) Ground truth (b) Otsu (c) Supervised CNN (Teacher-Net) (d) Proposed. Note that only area at top-right of the image is flooded.

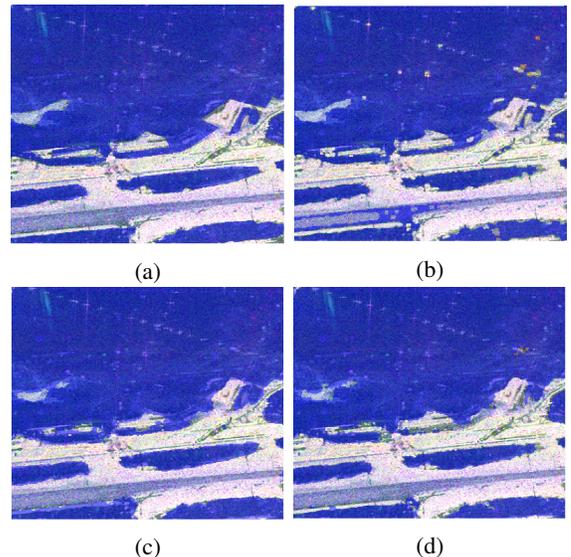


Fig. 5: Visualization comparison of generated flood map in semi-flooded region. (a) Ground truth (b) Otsu (c) Supervised CNN (Teacher-Net) (d) Proposed.

must be performed. Clearly, the results of CNN-based models are less affected.

Fig.5 represents the case where flooded and non-flooded area follows a bi-modal distribution, and the performance of the Otsu method is comparable to CNN-based models. While supervised CNN model fails to classify many pixels at the boundary of a flooded and non-flooded area, our proposed semi-supervised model can handle them. It should be noted that, since this image is captured at the urban area, many *ghost objects* might appear as small brighter areas [30]. Since the Otsu method does not take into account the neighborhood of target pixel, its generated flood map contains small holes at the location of *ghost objects*. Thus, the morphological closing operator is utilized to closing these small holes in flooded areas.

#### IV. CONCLUSIONS

We introduce a new CNN-based approach dedicated to flood mapping with incomplete ground truth in PolSAR and optical images data. Our approach utilizes the use of semi-supervised learning and multi-modal data to leverage a large portion of unlabelled pixels to improve the quality of the vanilla CNN model. Experiments show that our proposed model, trained with polarimetric-features, outperforms conventional approaches widely used in the literature. Our future work would focus on the augmentation of data using additional modalities including widely-available Social Network Services (SNS) photos.

#### REFERENCES

- [1] Wan, L., Liu, M., Wang, F., Zhang, T., You, H.J.: Automatic extraction of flood inundation areas from SAR images: a case study of Jilin, China during the 2017 flood disaster. *Int. J. Remote Sens.* 40, 50505077 (2019). 10.1080/01431161.2019.1577999
- [2] IPCC: Summary for policymakers. , Cambridge, UK (2014).
- [3] Li, Y., Martinis, S., Wieland, M.: Urban flood mapping with an active self-learning convolutional neural network based on TerraSAR-X intensity and interferometric coherence. *ISPRS J. Photogramm. Remote Sens.* 152, 178191 (2019). 10.1016/j.isprsjprs.2019.04.014
- [4] Gebrehiwot, A., Hashemi-Beni, L., Thompson, G., Kordjamshidi, P., Langan, T.E.: Deep convolutional neural network for flood extent mapping using unmanned aerial vehicles data. *Sensors (Switzerland)*. 19, (2019). 10.3390/s19071486
- [5] Ishikuro, T., Sato, R., Yamaguchi, Y., Yamada, H.: A Fundamental Study on Vehicle Detection in Flooded Urban Area Using Quad-Polarimetric SAR Data. *IEICE Trans. Electron.* E102.C, 3845 (2019). 10.1587/transele.e102.c.38
- [6] Cian, F., Marconcini, M., Ceccato, P., Giupponi, C.: Flood depth estimation by means of high-resolution SAR images and lidar data. *Nat. Hazards Earth Syst. Sci.* 18, 30633084 (2018). 10.5194/nhess-18-3063-2018
- [7] Landuyt, L., Van Wesemael, A., Schumann, G.J.P., Hostache, R., Verhoest, N.E.C., Van Coillie, F.M.B.: Flood Mapping Based on Synthetic Aperture Radar: An Assessment of Established Approaches. *IEEE Trans. Geosci. Remote Sens.* 57, 722739 (2019). 10.1109/TGRS.2018.2860054
- [8] Bolanos, S., Stiff, D., Brisco, B., Pietroniro, A.: Operational surface water detection and monitoring using Radarsat 2. *Remote Sens.* 8, (2016). 10.3390/rs8040285
- [9] Clement, M.A., Kilsby, C.G., Moore, P.: Multi-temporal synthetic aperture radar flood mapping using change detection. *J. Flood Risk Manag.* 11, 152168 (2018). 10.1111/jfr3.12303
- [10] Benoudjit, A., Guida, R.: A Web Application for the Automatic Mapping of the Flood Extent on Sar Images. In: 2018 IEEE 4th International Forum on Research and Technology for Society and Industry (RTSI). pp. 16. IEEE (2018). 10.1109/RTSI.2018.8548435
- [11] Jo, M.J., Osmanoglu, B., Zhang, B., Wdowinski, S.: Flood extent mapping using dual-polarimetric sentinel-1 synthetic aperture Radar imagery. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch.* 42, 711713 (2018). 10.5194/isprs-archives-XLII-3-711-2018
- [12] Amitrano, D., Di Martino, G., Iodice, A., Riccio, D., Ruello, G.: A Novel Tool for Unsupervised Flood Mapping Using Sentinel-1 Images. In: IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium. pp. 49094912. IEEE (2018). 10.1109/IGARSS.2018.8517957
- [13] Benoudjit, A., Guida, R.: A Novel Fully Automated Mapping of the Flood Extent on SAR Images Using a Supervised Classifier. *Remote Sens.* 11, 779 (2019). 10.3390/rs11070779
- [14] Insom, P., Cao, C., Boonsrimuang, P., Liu, D., Saokarn, A., Yomwan, P., Xu, Y.: A Support Vector Machine-Based Particle Filter Method for Improved Flooding Classification. *IEEE Geosci. Remote Sens. Lett.* 12, 19431947 (2015). 10.1109/LGRS.2015.2439575
- [15] Huang, W., DeVries, B., Huang, C., Lang, M.W., Jones, J.W., Creed, I.F., Carroll, M.L.: Automated extraction of surface water extent from Sentinel-1 data. *Remote Sens.* 10, 118 (2018). 10.3390/rs10050797
- [16] Pham-Duc, B., Prigent, C., Aires, F.: Surface water monitoring within cambodia and the Vietnamese Mekong Delta over a year, with Sentinel-1 SAR observations. *Water (Switzerland)*. 9, 121 (2017). 10.3390/w9060366
- [17] Gomathi, M., Greetha Priya, M., Krishnaveni, D.: Supervised classification for flood extent identification using sentinel-1 radar data. In: The 39th Asian Conference on Remote Sensing. pp. 32773284 (2018).
- [18] Heremans, R., Wiilekens, A., Borghys, D., Verbeeck, B., Valckenborgh, J., Achery, M., Perneel, C.: Automatic detection of flooded areas on ENVISAT/ASAR images using an object-oriented classification technique and an active contour algorithm. *RAST 2003 - Proc. Int. Conf. Recent Adv. Sp. Technol.* 311316 (2003). 10.1109/RAST.2003.1303926
- [19] Martinis, S., Twele, A., Voigt, S.: Towards operational near real-time flood detection using a split-based automatic thresholding procedure on high resolution TerraSAR-X data. *Nat. Hazards Earth Syst. Sci.* 9, 303314 (2009). 10.5194/nhess-9-303-2009
- [20] Manakos, I., Kordelas, G.A., Marini, K.: Fusion of Sentinel-1 data with Sentinel-2 products to overcome non-favourable atmospheric conditions for the delineation of inundation maps. *Eur. J. Remote Sens.* 0, 114 (2019). 10.1080/22797254.2019.1596757
- [21] Petit, O., Thome, N., Charnoz, A., Hostettler, A., Soler, L.: Handling missing annotations for semantic segmentation with deep convnets. *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*. 11045 LNCS, 2028 (2018). 10.1007/978-3-030-00889-5\_3
- [22] Marcos, D., Volpi, M., Kellenberger, B., Tuia, D.: Land cover mapping at very high resolution with rotation equivariant CNNs: Towards small yet accurate models. *ISPRS J. Photogramm. Remote Sens.* (2018). 10.1016/j.isprsjprs.2018.01.021
- [23] Yalniz, I.Z., Jgou, H., Chen, K., Paluri, M., Mahajan, D.: Billion-scale semi-supervised learning for image classification. (2019).
- [24] Yamaguchi, Y., Yajima, Y., Yamada, H.: A four-component decomposition of POLSAR images based on the coherency matrix. *IEEE Geosci. Remote Sens. Lett.* 3, 292296 (2006). 10.1109/LGRS.2006.869986
- [25] Chen, S., Tao, C.: PolSAR Image Classification Using Polarimetric-Feature-Driven Deep Convolutional Neural Network. *IEEE Geosci. Remote Sens. Lett.* 15, 627631 (2018). 10.1109/LGRS.2018.2799877
- [26] Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*. 9351, 234241 (2015). 10.1007/978-3-319-24574-4\_28
- [27] Milletari, F., Navab, N., Ahmadi, S.-A.: V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation. In: 2016 Fourth International Conference on 3D Vision (3DV). pp. 565571. IEEE (2016). 10.1109/3DV.2016.79
- [28] Kingma, D.P., Ba, J.: Adam: A Method for Stochastic Optimization. 115 (2014).
- [29] Otsu, N.: A Threshold Selection Method from Gray-Level Histograms. *IEEE Trans. Syst. Man. Cybern.* 9, 6266 (1979). 10.1109/TSMC.1979.4310076
- [30] Ishikuro, T., Sato, R., Yamaguchi, Y., Yamada, H.: A Fundamental Study on Vehicle Detection in Flooded Urban Area Using Quad-Polarimetric SAR Data. *IEICE Trans. Electron.* E102.C, 3845 (2019). 10.1587/transele.e102.c.38.