Neural-network fusion processing and inverse mapping to combine multi-sensor satellite data and to analyze significant features

This paper was downloaded from TechRxiv (<https://www.techrxiv.org>).

LICENSE

CC BY-NC-SA 4.0

SUBMISSION DATE / POSTED DATE

10-11-2021 / 18-11-2021

CITATION

Joshi, Gunjan; Natsuaki, Ryo; Hirose, Akira (2021): Neural-network fusion processing and inverse mapping to combine multi-sensor satellite data and to analyze significant features. TechRxiv. Preprint.
<https://doi.org/10.36227/techrxiv.16970446.v1>

DOI

[10.36227/techrxiv.16970446.v1](https://doi.org/10.36227/techrxiv.16970446.v1)

Neural-network fusion processing and inverse mapping to combine multi-sensor satellite data and to analyze significant features

Gunjan Joshi, Ryo Natsuaki, *Senior Member, IEEE* and Akira Hirose, *Fellow, IEEE*

Abstract—In the last decade, the increase in the number of active and passive earth observation satellites has provided us with more remote sensing data. This fact has led to increased interests in the field of fusion of the different satellite data since some of the satellites have properties complementary to one another. Fusion techniques can improve the estimation in areas of interest (AOIs) by using complementary information and inferring unknown parameters. However, when the observation area is large, extensive human labor and domain expertise are required for processing and analysis. Thus, we propose a neural network which combines and analyzes the data obtained from synthetic aperture radars (SAR) and optical sensors. The neural network employs a modified logarithmic activation function, unlike conventional networks, to realize inverse mapping for significant feature analysis based on dynamics consistent with its forward processing. In this paper, we focus on earthquake damage detection by dealing with the data of the 2018 Sulawesi earthquake in Indonesia. The fusion-based results show increased classification accuracy compared to the results of independent sensors. We further attempt to understand which input features are the significant contributors for which classification outputs by inverse-mapping in the data-fusion neural network. We observe that inverse mapping shows reasonable explanations in a consistent manner. It also indicates contributions of features different from straightforward counterparts, namely pre- and post-seismic features, in the detection of particular classes.

Index Terms—Synthetic aperture radar (SAR), optical satellite, multi-source fusion, neural network, inverse-mapping, significant features

I. INTRODUCTION

THE number of earth observation satellites have increased in recent years. Their sensors offer a wide range of spatial, spectral, and temporal resolutions [1]

and possess different peculiarities designed for their specific tasks.

Earthquake damage detection is a field where remote sensing satellites are extensively helpful. They can assess an inaccessible region after a disaster and can cover a wider scene, because of their synoptic imaging capability, in particular when the event is located in a remote area or a region where communication infrastructures have been damaged. In such cases, a damage map is required for the establishment of the first step in the evacuation and mitigation plan [2]. Remote sensing technologies have shown to be effective damage evaluation tools in such cases by providing high amount of critical information in spatially and temporally. Conventionally, optical satellites have been used for direct detection of earthquake damages [3]–[5]. Optical satellites with their ever increasing spatial resolutions can provide us with easy-to-understand general surface information. However, their viewing capacity is limited by cloud coverage.

On the other hand, SAR sensors are also beneficial for rapid disaster assessment, as they have the ability to penetrate clouds and view the earth's surface irrespective of the time of the day. Conventional methods for damage assessment using SAR include generating interferometric coherence changes from multitemporal SAR images for wide region mapping [6], [7], building level [8], [9] and rapid damage mapping [10]–[13]. Recent studies have focused on the fusion of SAR and optical remote sensing for forest mapping [14]–[17], land-use land-cover (LULC) applications [18]–[20] and earthquake damage detection [21], [22]. In such cases of fusion of these two sensor data, SAR can provide information that is not visible in the optical data and vice-versa. The two data sources are thus regarded as complementary, not competing. Multisensor fusion studies on earthquake damage detection [21], [22] focus on the detection of damage stricken urban areas as they tend to have a higher fatality rate.

A recent paper [23] reported multi-source data fusion by using ensemble learning classifiers along with feature importance analysis. They concluded that digital elevation model (DEM) and SAR-derived features contribute

Manuscript received xxxx; revised xxxx and xxx; accepted xxx.

A part of this work was presented in Asia-Pacific Conference on Synthetic Aperture Radar (APSAR) 2021, Bali, Indonesia, online.

This work was supported in part by JSPS KAKENHI under Grant 18H04105, and in part by the Cooperative Research Project Program of the Research Institute of Electrical Communication (RIEC), Tohoku University. (Corresponding author: Gunjan Joshi.)

The authors are with the Department of Electrical Engineering and Information Systems, The University of Tokyo, Tokyo 113-8656, Japan (e-mail: gunjanjoshi88@g.ecc.u-tokyo.ac.jp; natsuaki@ee.t.u-tokyo.ac.jp; ahirose@ee.t.u-tokyo.ac.jp)

the most to the overall damage classification. Their classification focuses primarily on building damage mapping and classifies the building as destroyed, damaged, possibly damaged, or no damage. In addition, the classification requires multiple stages of processing. Thus the importance analysis becomes complicated and indirect, making it harder to relate the observations with the analysis decision. In order to provide more information about the damaged areas, it is necessary to focus on the various other land type classes as well.

This paper focuses on land classification along with accurate damage estimation of various landforms, not just urban areas. We employ a fully connected neural network for the fusion of different features of various sensors to classify each pixel as one of the predefined land types.

Neural networks are very useful and effective in solving problems difficult for traditional feature extractors and classifiers [24]. To build trust in such intelligent systems and integrate them with our society, there should be clear transparency in the neural models that can explain why they predict what they predict. In the area of neural-network explainability, attribution value method and feature visualization have gained significant attention [25]. Important pixels related to a particular output are highlighted with the use of the attribution method. Feature visualization method, on the other hand, provides an understanding of a single kernel or convolutional layer [26]–[29]. However, convolutional neural networks, which are widely used nowadays, pay attention to spatial features such as textures and shapes.

In this paper, an investigation is performed to understand the pixel-by-pixel and class-by-class contributions of each feature in the classification results. We propose inverse-mapping dynamics, which pays key attention to the signal flow in the neural network. We find that the inverse mapping shows reasonable and consistent results and reveals show the contribution of non-counterpart features in class decision.

In the subsequent sections, we first focus on the proposed neural network and inverse mapping in Section II. Section III presents experiments, where we apply the proposed neural network to detect earthquake damage in Sulawesi and determine the significant factors. This paper concludes with relevant discussion in Section IV and conclusion in Section V.

II. PROPOSAL OF A NEURAL NETWORK HAVING MODIFIED LOGARITHMIC ACTIVATION FUNCTION FOR FUSION AND ANALYZIS

A. Neural network structure and forward processing

Fig. 1 shows the construction of the proposed fully-connected neural network, which consists of an input-terminal layer, a hidden-neuron layer, and an output-neuron layer.

Considering $\mathbf{x} = [x_1 \ x_2 \ \dots \ x_i \ \dots \ x_I]^T$, $\mathbf{y} = [y_1 \ y_2 \ \dots \ y_j \ \dots \ y_J]^T$, to be the input and hidden signals respectively, ${}^1\mathbf{b} = [{}^1b_1 \ {}^1b_2 \ \dots \ {}^1b_j \ \dots \ {}^1b_J]^T$, ${}^2\mathbf{b} = [{}^2b_1 \ {}^2b_2 \ \dots \ {}^2b_k \ \dots \ {}^2b_K]^T$, to be the biases for the hidden and output layers respectively, and ${}^1\mathbf{W}$, ${}^2\mathbf{W}$ to be the input - hidden weights and the hidden - output weights obtained after the learning phase. The values of $\mathbf{z}$, where $\mathbf{z} = [z_1 \ z_2 \ \dots \ z_k \ \dots \ z_K]^T$, at the output layer of the neural network can be calculated as

$$\begin{aligned} \mathbf{z} &= f({}^2\mathbf{W}\mathbf{y} + {}^2\mathbf{b}) \\ &= f({}^2\mathbf{W}f({}^1\mathbf{W}\mathbf{x} + {}^1\mathbf{b}) + {}^2\mathbf{b}) \end{aligned} \quad (1)$$

where f represents the proposed modified logarithmic activation function working component-wise and defined as

$$f(x) = \begin{cases} -\log(-(x-1)) & (x < 0) \\ 0 & (x = 0) \\ \log(x+1) & (x > 0) \end{cases} \quad (2)$$

Fig. 2 shows the curve. We employ this function instead of a conventional activation function such as tanh or Rectified Linear Unit (ReLU) because the use of these activation functions leads to a limited range of inversely input values in the inverse mapping, which can be avoided by using the modified logarithmic function defined in (2).

B. Inverse mapping

The weights of the forward processing neural network are helpful in understanding how the network determines a winner neuron. Inverse mapping is a dynamics that traces which input is significant for which output by paying attention to the signal flow of the network.

Fig. 3 illustrates the process. The value obtained at the winning node $\hat{k}$ in the forward processing is fed to the same node at the input layer of the inverse mapping, while all the other nodes are fed with zeros. This is done to suppress any influence that the non-significant classes could introduce. The inverse mapping is represented by

$$\tilde{\mathbf{x}} = {}^1\mathbf{W}^T f^{-1}({}^2\mathbf{W}^T f^{-1}(\tilde{\mathbf{z}})) \quad (3)$$

where $\tilde{\mathbf{z}} = [0 \ 0 \ \dots \ z_{\hat{k}} \ \dots \ 0]^T$ is the modified output values of the forward processing network fed as an input to the inverse-mapping network, $\tilde{\mathbf{x}}$ is the output of the inverse-mapping network and the inverse activation function f^{-1} is defined as

$$f^{-1}(x) = \begin{cases} -e^{-x} + 1 & (x < 0) \\ 0 & (x = 0) \\ e^x - 1 & (x > 0) \end{cases} \quad (4)$$

If $|z_{\hat{k}}|$ is small (e.g. $|z_{\hat{k}}| < 1$) indicating that the network is not confident of its estimation, we discard the pixel in the inverse mapping. The obtained value

Fig. 1: Forward processing neural network.

Fig. 2: Activation function proposed for the neural network.

$\tilde{x}$ suggests significant factors to determine the class of the input pixel data. A large absolute value $\tilde{x}_i$ indicates the i -th observation data contributing to the winner class determination for that pixel. For each pixel, all the elements of $\tilde{x}$ are normalized by the maximum absolute value among the elements.

Fig. 3: Inverse-mapping network.

III. EXPERIMENT

A. Study area

Sulawesi Island, in eastern Indonesia, is at the triple junction of the Sunda, Philippine, and Australian plates [30]. On September 28, 2018, an earthquake of Magnitude 7.5 struck Sulawesi. This earthquake was responsible for generating multiple landslides, soil liquefaction, and a tsunami. The locations of some of these slides can be seen in Fig 4.

The earthquake's epicenter was Donggala Regency (0.20 S, 119.89 E); 80km north of the provincial capital city Palu. The most significant contributors to the casualties were the landslides within the alluvial valley to the south of Palu that were responsible for over 4,340 people that went missing from the Petobo and Balaroa villages, which got buried completely [2], [31].

B. Dataset

We map the damaged areas by using L-band SAR sensor Advanced Land Observation Satellite-2 (ALOS-2) of Japan Aerospace Exploration Agency (JAXA) and Sentinel-2 of the European Space Agency (ESA). The SAR datasets and optical datasets that were required for the damage assessment in this study have been described in Table I.

C. Pre-processing and feature extraction

Preprocessing operations are meant to correct the images for sensor-specific and platform-specific radio-

TABLE I: Datasets used in this study. Dotted line denotes the occurrence of the earthquake.

Satellite Images	Sensors	Acquisition Dates	Resolutions	Image bands	Polarisations	
ALOS 2	SAR	11 May 2018	10m	L	HH & HV	Pre-Event
ALOS 2	SAR	17 August 2018	10m	L	HH & HV	Pre-Event
Sentinel 2	Optical	27 September 2018	10m,20m	Red,NIR & SWIR		Pre-Event
Sentinel 2	Optical	02 October 2018	10m,20m	Red,NIR & SWIR		Post-Event
ALOS 2	SAR	12 October 2018	10m	L	HH & HV	Post-Event

Fig. 4: Map depicting the Earthquake triggered landslides in Sulawesi, Indonesia.

metric and geometric distortions and convert them to intelligible form. Preprocessing steps involve geometric and radiometric calibration of the raw data. In the case of SAR data, speckle filtering is additionally required. To avoid mismatch in the spatial resolution, resampling operation is performed on the Sentinel-2 data, following which all the bands of Sentinel-2 are sampled to 10 m resolution. The preprocessing of the satellite data is majorly performed by using ESA's Sentinel Application Platform (SNAP) [32]. From the preprocessed SAR and the optical datasets, we choose features that have been studied to show sensitivity to the damage caused by disasters. Fig. 5 illustrates the preprocessing steps. From SAR data, we obtained the intensity (in dB) and the coherence values, where both these features are obtained

Fig. 5: Preprocessing for SAR and optical raw data.

in Horizontal-Horizontal (HH) and Horizontal-Vertical (HV) polarizations. From the optical data, we extract the Normalized Difference Vegetation Index (NDVI), Normalized Difference Building Index (NDBI), and the Build-Up Index (BUI). These features are defined as follows.

- Intensity refers to the backscattering received by the SAR sensor. Intensity values can provide useful information about the damage [33] as there is a

change in the backscattering coefficient before and after the disaster. It is obtained as the square of the amplitude of the SAR raw image as

$$\text{Intensity} = 20 \log_{10}(\text{Amplitude}) \quad (5)$$

- Coherence is a feature obtained from SAR data that depicts the stability of an area. It is based on both amplitude and phase of the backscattered signals. It shows where changes have occurred and detects even subtle changes with high levels of accuracy ranging from several millimetres to centimetres [22]. Coherence can be calculated as

$$\text{Coherence} = \frac{|\langle S^* P \rangle|}{\sqrt{\langle P^* P \rangle \langle S^* S \rangle}} \quad (6)$$

where P and S represent the complex amplitude obtained in primary and secondary observations, $(\cdot)^*$ above them represents the complex conjugate transpose, and the bracket denotes the spatial average for 3×3 window.

- The Normalized Difference Vegetation Index (NDVI) is an index obtained from the optical sensors. NDVI is often used to classify an area into vegetation or non-vegetation class [21]. This feature can be determined from the contribution of the satellite bands corresponding to visible Red and Near-Infrared (NIR) wavelengths [34] corresponding to Band-4 (B4) and Band-8 (B8) of Sentinel-2, respectively, and is given by their power values as

$$\text{NDVI} = \frac{\text{NIR}(B8) - \text{Red}(B4)}{\text{NIR}(B8) + \text{Red}(B4)} \quad (7)$$

- The Normalized Difference Build-up Index (NDBI) is an index that highlights the urban areas and is given by the normalization of the Short-Wave-Infra-Red (SWIR) and Near Infra-Red (NIR) wavelengths [35] corresponding to Band-11 (B11) and Band-4 (B4) of Sentinel-2 and can be calculated by their power values as

$$\text{NDBI} = \frac{\text{SWIR}(B11) - \text{NIR}(B8)}{\text{SWIR}(B11) + \text{NIR}(B8)} \quad (8)$$

- Build-Up Index (BUI) is the index for analysis of urban pattern using NDVI and NDBI. A higher value of a pixel in this index indicates a greater possibility that it represents a built-up area [36], which allows built-up areas to be mapped effectively. BUI is given their power values as

$$\text{BUI} = \text{NDBI} - \text{NDVI} \quad (9)$$

Tables II, III and IV illustrate the normalized features extracted from the multiple sensors. A total of fourteen features were extracted from the given datasets. Seven

features were extracted from the pre-event scenes, and seven features were extracted from the post-event scenes.

To ensure geometric accuracy and good spatial alignment between the fourteen extracted feature images, we perform collocation by using the SNAP toolbox. In the collocation operation, for a pixel in the primary product, the corresponding pixel is looked up (by geolocation) in the secondary product, and the primary and secondary images are mapped to the same geographical location. After which, the data sets are clipped very precisely to the coordinates of the three AOIs, namely Sidera, Petobo and Balaroa. All the feature values are then normalized between the range -1 and 1 .

When we observe intensity in Tables II, III and IV, we can see that in the areas where the landslide occurred, there was an increase in the brightness of the intensity in both HH and HV polarizations. However, in the coherence, it can be observed that post the landslide, there is a loss of coherence at the places where the landslide occurred. A similar change in the characteristic can be seen for the three optical features NDVI, NDBI and BUI. In the case of NDVI, there is a loss in the NDVI index due to the fact that the area had a loss of vegetation. However, for NDBI and BUI, the forest and urban areas were converted to barren lands, and thus there is an increase in these indices for the affected regions.

D. Neural network training

To understand the effect of the fusion of SAR and optical features on the neural decision accuracy, we perform three different types of neural network processing, where the inputs are:

1. Only SAR features
2. Only optical features
3. Both SAR and optical features.

The input layer is fed with pixel-by-pixel values of the feature images shown in Table II, Table III or Table IV obtained after the preprocessing steps illustrated in Section III-C.

At the output layer of the neural network, the network classifies each pixel as one of the following seven land-form classes: damaged forest, damaged irrigation field, damaged urban, undamaged forest, undamaged irrigation field, undamaged urban and undamaged dryland regions. Since the dryland region was not affected by the disaster, "damaged dryland class" was not included in our assessment.

Table V illustrates the use of the AOIs in this study. We adopt a pixel-by-pixel image analysis method. We train and validate the neural networks on randomly selected 3,970 pixels in the Sidera AOI (enclosed in the black boxes) shown in Fig. 6. The training and validation data set is split into 80:20 ratio. Table VI shows the neural

TABLE II: Features extracted for the Sidera AOI.

Feature	Pre-seismic image	Post-seismic image
RGB reference image
Intensity HH	0 50 100 150 200 250 300 350 0 100 200 300 400	0 50 100 150 200 250 300 350 0 100 200 300 400
Intensity HV	0 50 100 150 200 250 300 350 0 100 200 300 400	0 50 100 150 200 250 300 350 0 100 200 300 400
Coherence HH	0 50 100 150 200 250 300 350 0 100 200 300 400	0 50 100 150 200 250 300 350 0 100 200 300 400
Coherence HV	0 50 100 150 200 250 300 350 0 100 200 300 400	0 50 100 150 200 250 300 350 0 100 200 300 400
NDVI	0 50 100 150 200 250 300 350 0 100 200 300 400	0 50 100 150 200 250 300 350 0 100 200 300 400
NDBI	0 50 100 150 200 250 300 350 0 100 200 300 400	0 50 100 150 200 250 300 350 0 100 200 300 400
BUI	0 50 100 150 200 250 300 350 0 100 200 300 400	0 50 100 150 200 250 300 350 0 100 200 300 400

TABLE III: Features extracted for the Petobo AOI.

Feature	Pre-seismic image	Post-seismic image
RGB reference image
Intensity HH
Intensity HV
Coherence HH
Coherence HV
NDVI
NDBI
BUI

TABLE IV: Features extracted for the Balaroa AOI.

Feature	Pre-seismic image	Post-seismic image
RGB reference image
Intensity HH
Intensity HV
Coherence HH
Coherence HV
NDVI
NDBI
BUI

Fig. 6: This figure shows the pixels that are used for three training of the neural network (enclosed in black boxes).

TABLE V: Use of the AOIs in this study.

AOI	Total pixels	Pixels usage and use of area in this study
Sidera	185,942	3,970 (80% train 20% validation) Classification, Accuracy estimation
Petobo	163,747	Classification, Accuracy estimation
Balaroa	27,874	Classification

parameters determined empirically. The neural networks are also applied to other unseen areas in Sidera as well as Petobo and Balaroa AOIs.

E. Neural network training results

After the training on the selected pixels, the neural networks are applied to the unseen pixels of the various AOIs. They are expected to estimate which of the seven classes each pixel belongs to. The neural networks generates classification maps for the Sidera and the Petobo AOIs, which are compared with the ground truth maps shown in Table VII(d) and (h), respectively. These ground truth maps were derived from the land classification published in Ref. [31] and modified for the damaged region based on the damage mapping by the Copernicus emergency management service [37]. Moreover, since there is no ground truth for the Balaroa region, the results are compared visually with the RGB image given in Table VII(i).

Table VII(a), (b) and (c) are learning curves showing the training and validation accuracy for the neural networks trained on SAR-only features, optical-only features and both SAR and optical features, respectively. All of them illustrate smooth learning. In Fig. 7, we can observe an increase in the accuracy for the neural network trained on both SAR and optical features.

TABLE VI: Neural network training parameters.

	SAR only	Optical only	SAR & optical
Input layer terminals	8	6	14
Hidden layer neurons	128	64	64
Output layer neurons	7	7	7
Epochs	8,000	2,000	1,000

Fig. 7: Accuracy comparison for SAR-only, optical-only and SAR and optical neural networks.

Table VII(e), (i) and (m) show the classification results for Sidera, Petobo and Balaroa regions, respectively, for the neural network trained on SAR-only parameters. The classification is poor, and not much information can be obtained from this damage map. However, it can be seen that the undamaged dryland class and the urban class appear to be detected well.

Table VII(f), (j) and (n) illustrate the results for the neural network trained on only optical features. The classification is much better than the SAR-only classification, particularly the shape of the landslide is detected well. However, with a close visual inspection, it is observed that some classes, specifically the urban regions, are not detected at all.

Table VII(g), (k) and (o) show the classification for the neural network trained on both SAR and optical features. The classification appears similar to the optical analysis results. The urban class is detected, and the identification of the other classes is better. Thus, it can be stated that the results of this neural network are much closer to the ground truth data as compared to the classifications by the other neural networks.

F. Confusion matrix

The classification maps in Table VII(g), (k) and (o) generated by the fusion-based neural network are compared for Sidera and Petobo AOIs with the ground truth data shown as Table VII(d) and (h). Fig. 8 presents confusion matrices that compare the classification maps

TABLE VII: Neural-network learning curves and classification maps. (Dark green : Damaged forest, Orange : Damaged irrigation fields, Purple : Damaged urban, Light green : Undamaged forest, Yellow : Undamaged irrigation fields, Brown : Undamaged urban, Grey : Undamaged dryland)

Actual Class	DF	5823 47%	616 4%	237 3%	20 0%	861 2%	693 1%	66 1%
	DIF	4212 34%	8370 58%	2178 26%	69 0%	476 1%	2690 4%	378 5%
	DU	136 1%	556 4%	1003 12%	12 0%	80 0%	307 0%	410 5%
	UDF	26 0%	313 2%	214 3%	16935 80%	595 1%	4999 7%	356 4%
	UDIF	1146 9%	577 4%	572 7%	665 3%	39817 79%	18341 26%	708 9%
	UDU	917 7%	3234 22%	2230 27%	2608 12%	7694 15%	38444 54%	2084 25%
	UDDL	154 1%	819 6%	1853 22%	732 3%	737 1%	5706 8%	4273 52%
		DF	DIF	DU	UDF	UDIF	UDU	UDDL
		Predicted Class						

(a) Sidera - 7-class confusion matrix

Actual Class	D	23131 66%	6062 4%
	UD	12055 34%	144694 96%
		D	UD
		Predicted Class	

(b) Sidera - 2-class confusion matrix

Actual Class	DF	4031 48%	1551 17%	1682 8%	7 0%	152 1%	480 1%	75 0%
	DIF	1554 18%	1876 20%	1459 7%	11 0%	178 1%	598 1%	215 1%
	DU	266 3%	1321 14%	4994 24%	7 0%	142 0%	687 1%	1689 8%
	UDF	50 1%	263 3%	687 3%	6713 67%	565 2%	5962 9%	415 2%
	UDIF	1148 14%	786 8%	1433 7%	511 5%	19711 65%	15049 24%	1848 8%
	UDU	1007 12%	1916 21%	2323 11%	1620 16%	7819 26%	28487 45%	2813 13%
	UDDL	383 5%	1618 17%	7906 39%	1175 12%	1630 5%	11646 19%	15288 68%
		DF	DIF	DU	UDF	UDIF	UDU	UDDL
		Predicted Class						

(c) Petobo - 7-class confusion matrix

Actual Class	D	18734 49%	4241 3%
	UD	19520 51%	121252 97%
		D	UD
		Predicted Class	

(d) Petobo - 2-class confusion matrix

Fig. 8: Confusion matrices for Sidera and Petobo using both SAR and optical data. (D: Damaged, UD: Undamaged, DF: Damaged Forest, DIF: Damaged Irrigation Fields, DU: Damaged Urban, UDF: Undamaged Forest, UDIF: Undamaged Irrigation Fields, UDU: Undamaged Urban, UDDL: Undamaged Dryland.)

TABLE VIII: The correlation between the predicted class and the SAR and optical features.

Class	Significant Feature
Damaged Forest	Post-seismic NDVI, Pre-seismic NDBI, Pre-seismic NDVI
Damaged Irrigation Fields	Post-seismic NDVI, Post-seismic BUI
Damaged Urban	Pre-seismic BUI, Pre-seismic Coherence HV, Post-seismic NDBI
Undamaged Forest	Post-seismic NDBI
Undamaged Irrigation Fields	Post-seismic NDBI, Pre-seismic BUI
Undamaged Urban	Post-seismic Coherence HH, Pre-seismic NDVI
Undamaged Dryland	Post-seismic NDVI, Post-seismic BUI

Fig. 9: Overall confusion-matrix accuracy comparison.

pixel-by-pixel with the ground truth data. It shows what percentage of the AOI pixels are correctly identified in the resulting classification. A 2-class comparison has also been performed, in which we group all damaged pixels as a single damaged class and all undamaged pixels as a single undamaged class.

The confusion matrices for Sidera region are shown in Figs. 8(a) and (b) for the 7-class and 2-class classifications, respectively. Similarly, the confusion matrices for Petobo region are shown in Figs. 8(c) and (d). The classifications are quantitatively evaluated by using precision scores. The precision scores for each individual class are shown in confusion matrices themselves. The confusion matrices indicate a high precision score for most classes. However, the damaged urban class in Sidera AOI, and additionally damaged irrigation field class in Petobo AOI show low precision scores. The cause of the low precision scores for these classes is attributed to the mismatch in the spatial resolutions. That is, the ground-truth data has a higher spatial resolution compared to the datasets used in this study. Thus, there exists this precision issue.

Confusion matrices were generated for the results from the only-SAR and only-optically trained neural networks as well for evaluation using an overall accuracy score. Fig. 9 presents the comparison of the accuracy score for the confusion matrices of all three neural networks. The

7-class results illustrate that the classification accuracy increases with the fusion of the data. From the 2-class results, we can see that the classification accuracy is similar for optical-only and SAR and optical classifications. Thus, it can be postulated that, for optical-only classification, though the pixels were correctly identified as damaged or undamaged, the land type could not be accurately identified as compared to the fusion-based classification. This means that optical data detects less information compared to fusion-based data.

G. Inverse mapping result using both SAR and optical features

Section II-B explained the inverse mapping. Figs. 10 to 13 illustrate the inverse mapping results. The output $\tilde{x}$ are plotted in box-whisker plots for all the pixels. The density of the pixels closer to -1 or 1 in these plots shows the significant features. The results of the inverse mapping illustrate the impact of the 14 features on the classification results. Table VIII shows the inference based on the box and whisker plots in Figs. 10 to 13.

Fig. 10 and Fig. 12 show the inverse mapping results for the pixels evaluated in the damaged category for Sidera and Petobo AOIs, respectively. In both figures, we can observe that the damaged forest class was determined by post-seismic NDVI, pre-seismic NDBI and pre-seismic NDVI. The damaged irrigation field class was selected by post-seismic NDVI and also post-seismic BUI. Finally, the damaged urban class was chosen by pre-seismic BUI, pre-seismic coherence HV and post-seismic NDBI.

Fig. 11 and Fig. 13 illustrate the inverse mapping results for the undamaged pixels. Post-seismic NDBI is considered a significant factor for undamaged forest decision. Pixels labelled as undamaged irrigation fields are attributed to the post-seismic NDBI and pre-seismic BUI. Post-seismic coherence HH and pre-seismic NDVI influenced the undamaged urban class decision. Post-seismic NDVI and post-seismic BUI are prominent features for the undamaged dryland class. We can observe consistency in the results for the different AOIs. The results are reasonable for each class when we consider the land type and physical landslide dynamics. Further, it can also be noted that features that independently contribute

to the decision of a class worked collaboratively for the class, as observed for the damaged urban class.

significant information to generate effective post-disaster maps in an explainable manner.

IV. DISCUSSION

From the results illustrated in Sections III-E, III-F and III-G. This work focuses on the fusion of features from two satellites by using a fully connected neural network, which employs a modified logarithmic function. This study indicates that the accuracy of damage estimation increases with the fusion of data from the SAR and optical sensor. For the classification using only SAR parameters, it can be inferred that the selected parameters were insufficient, and thus, the SAR-only combinations were not that impactful. Though SAR features did not show significant results when used alone, it increased the accuracy and interpretability of the damage classification results when used synergetically with the optical data. Certain classes, such as urban areas which were not detected by the optical sensor, were detected by the SAR sensor. This suggests that the coherence features of SAR carry information that complements the optical data. We also focus on determining the prominent features for each class. Consistent results were obtained for the three AOIs, and the results were reasonable for the particular classes. It was also found that features that individually contributed to the detection of any class worked together to detect that class. For example, in the case of damaged urban class, NDBI, BUI and coherence HV were determined to be significant, and these features had been previously studied to be individually good indicators of detecting changes in urban regions. Traditionally, counterparts such as pre-seismic and post-seismic of any feature are used for damage detection. In our study, it was found that such counterparts do not necessarily complement each other when used with other features. Investigation of this observation can be the future scope of this work.

V. CONCLUSION

In this paper, we proposed a neural network with a modified logarithmic activation function that combines the capabilities of various damage indicating features obtained from SAR and optical sensors. The combination of two sensors has been found effective for land-type and damage detection and gave higher accuracy decision as well as consistent and reasonable significant features. Through this study, we were able to identify the relevance of these features to damage assessment. It was observed that features that independently contribute to damage detection of a particular class also determine collaboratively the specific class. The significant features belong to both SAR and optical sensors, highlighting the necessity of sensor integration. This study illustrated that the synergetic use of multiple sensors can provide us with

Fig. 10: Inverse-mapping results for Sidera region, for pixels classified in the damaged category.

Fig. 11: Inverse-mapping results for Sidera region, for pixels classified in the undamaged category.

Fig. 12: Inverse-mapping results for Petobo region, for pixels classified in the damaged category.

Fig. 13: Inverse-mapping results for Petobo region, for pixels classified in the undamaged category.

REFERENCES

- [1] M. Schmitt and X. X. Zhu, "Data Fusion and Remote Sensing: An ever-growing relationship," *IEEE Geoscience and Remote Sensing Magazine*, vol. 4, no. 4, pp. 6–23, 2016.
- [2] M. Syifa, P. R. Kadavi, and C. W. Lee, "An artificial intelligence application for post-earthquake damage mapping in Palu, central Sulawesi, Indonesia," *Sensors (Switzerland)*, vol. 19, no. 3, 2019.
- [3] A. J. Cooner, Y. Shao, and J. B. Campbell, "Detection of urban damage using remote sensing and machine learning algorithms: Revisiting the 2010 Haiti earthquake," *Remote Sensing*, vol. 8, no. 10, 2016.
- [4] A. Ajmar, P. Boccardo, and F. G. Tonolo, "Earthquake damage assessment based on remote sensing data. The Haiti case study," *Italian Journal of Remote Sensing / Rivista Italiana di Telerilevamento*, vol. 43, no. 2, pp. 123–128, 2011.
- [5] F. Yamazaki, Y. Yano, and M. Matsuoka, "Visual damage interpretation of buildings in Bam city using QuickBird images following the 2003 Bam, Iran, earthquake," *Earthquake Spectra*, vol. 21, no. SUPPL. 1, 2005.
- [6] J. Hoffmann, "Mapping damage during the bam (iran) earthquake using interferometric coherence," *International Journal of Remote Sensing*, vol. 28, no. 6, pp. 1199–1216, 2007. [Online]. Available: <https://doi.org/10.1080/01431160600928567>
- [7] G. A. Arciniegas, W. Bijker, N. Kerle, and V. A. Tolpekin, "Coherence- and amplitude-based analysis of seismogenic damage in bam, iran, using envisat asar data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 6, pp. 1571–1581, 2007.
- [8] W. Liu and F. Yamazaki, "Extraction of collapsed buildings in the 2016 Kumamoto earthquake using multi-temporal PALSAR-2 data," *Journal of Disaster Research*, vol. 12, no. 2, 2017.
- [9] R. Natsuaki, H. Nagai, N. Tomii, and T. Tadono, "Sensitivity and limitation in damage detection for individual buildings using InSAR coherence-A case study in 2016 Kumamoto earthquakes," *Remote Sensing*, vol. 10, no. 2, 2018.
- [10] P. Ge, H. Gokon, K. Meguro, and S. Koshimura, "Study on the intensity and coherence information of high-resolution ALOS-2 SAR images for rapid massive landslide mapping at a pixel level," *Remote Sensing*, vol. 11, no. 23, 2019.
- [11] L. Ge, A. H.-M. Ng, X. Li, Y. Liu, Z. Du, and Q. Liu, "Near real-time satellite mapping of the 2015 gorkha earthquake, nepal," *Annals of GIS*, vol. 21, no. 3, pp. 175–190, 2015. [Online]. Available: <https://doi.org/10.1080/19475683.2015.1068221>
- [12] S. H. Yun, K. Hudnut, S. Owen, F. Webb, M. Simons, P. Sacco, E. Gurrula, G. Manipon, C. Liang, E. Fielding, P. Milillo, H. Hua, and A. Coletta, "Rapid damage mapping for the 2015 Mw 7.8 Gorkha Earthquake Using synthetic aperture radar data from COSMO-SkyMed and ALOS-2 satellites," *Seismological Research Letters*, vol. 86, no. 6, pp. 1549–1556, 2015.
- [13] S. Plank, "Rapid Damage Assessment by Means of Multi-Temporal SAR — A Comprehensive Review and Outlook to Sentinel-1," *Remote Sensing*, vol. 6, no. 6, pp. 4870–4906, 2014. [Online]. Available: <https://www.mdpi.com/2072-4292/6/6/4870>
- [14] J. Reiche, J. Verbesselt, D. Hoekman, and M. Herold, "Fusing landsat and sar time series to detect deforestation in the tropics," *Remote Sensing of Environment*, vol. 156, pp. 276–293, 2015. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0034425714003885>
- [15] J. Reiche, C. M. Souza, D. H. Hoekman, J. Verbesselt, H. Persaud, and M. Herold, "Feature level fusion of multi-temporal alos palsar and landsat data for mapping and monitoring of tropical deforestation and forest degradation," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 6, no. 5, pp. 2159–2173, 2013.
- [16] D. Deus, "Integration of ALOS PALSAR and Landsat data for land cover and forest mapping in northern Tanzania," *Land*, vol. 5, no. 4, 2016.
- [17] G. Vaglio Laurin, V. Liesenberg, Q. Chen, L. Guerriero, F. Del Frate, A. Bartolini, D. Coomes, B. Wilebore, J. Lindsell, and R. Valentini, "Optical and sar sensor synergies for forest and land cover mapping in a tropical site in west africa," *International Journal of Applied Earth Observation and Geoinformation*, vol. 21, pp. 7–16, 2013. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0303243412001651>
- [18] W. Kou, X. Xiao, J. Dong, S. Gan, D. Zhai, G. Zhang, Y. Qin, and L. Li, "Mapping deciduous rubber plantation areas and stand ages with palsar and landsat images," *Remote Sensing*, vol. 7, no. 1, pp. 1048–1073, 2015. [Online]. Available: <https://www.mdpi.com/2072-4292/7/1/1048>
- [19] N. Torbick, L. Ledoux, W. Salas, and M. Zhao, "Regional mapping of plantation extent using multisensor imagery," *Remote Sensing*, vol. 8, no. 3, 2016. [Online]. Available: <https://www.mdpi.com/2072-4292/8/3/236>
- [20] R. Jhonnerie, V. P. Siregar, B. Nababan, L. B. Prasetyo, and S. Wouthuyzen, "Random forest classification for mangrove land cover mapping using landsat 5 tm and alos palsar imageries," *Procedia Environmental Sciences*, vol. 24, pp. 215–221, 2015, the 1st International Symposium on LAPAN-IPB Satellite (LISAT) for Food Security and Environmental Monitoring. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1878029615000961>
- [21] N. Tamkuan and M. Nagai, "Fusion of multi-temporal interferometric coherence and optical image data for the 2016 Kumamoto earthquake damage assessment," *ISPRS International Journal of Geo-Information*, vol. 6, no. 7, pp. 1–17, 2017.
- [22] S. Havivi, I. Schwartzman, S. Maman, S. R. Rotman, and D. G. Blumberg, "Combining TerraSAR-X and Landsat images for emergency response in urban environments," *Remote Sensing*, vol. 10, no. 5, 2018.
- [23] B. Adriano, J. Xia, G. Baier, N. Yokoya, and S. Koshimura, "Multi-source data fusion based on ensemble learning for rapid building damage mapping during the 2018 Sulawesi earthquake and Tsunami in Palu, Indonesia," *Remote Sensing*, vol. 11, no. 7, 2019.
- [24] A. Hirose, *Complex-Valued Neural Networks*, 2nd ed. Springer, 2012.
- [25] R. Guidotti, A. Monreale, S. Ruggieri, F. Turini, F. Giannotti, and D. Pedreschi, "A survey of methods for explaining black box models," *ACM Computing Surveys*, vol. 51, no. 5, 2018.
- [26] R. Shi, T. Li, and Y. Yamaguchi, "Group visualization of class-discriminative features," *Neural Networks*, vol. 129, pp. 75–90, 2020. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0893608020301969>
- [27] E. Protas, J. D. Bratti, J. F. Gaya, P. Drews, and S. S. Botelho, "Visualization Methods for Image Transformation Convolutional Neural Networks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 7, pp. 2231–2243, 2019.
- [28] F. Wang, H. Liu, and J. Cheng, "Visualizing deep neural network by alternately image blurring and deblurring," *Neural Networks*, vol. 97, pp. 162–172, 2018. [Online]. Available: <https://doi.org/10.1016/j.neunet.2017.09.007>
- [29] A. Mahendran and A. Vedaldi, "Visualizing deep convolutional neural networks using natural pre-images," *International Journal of Computer Vision*, vol. 120, pp. 233–255, 2016.
- [30] O. Bellier, L. Siame, T. Beaudouin, M. Villeneuve, and R. Braucher, "High slip rate for a low seismicity along the Palu-Koro active fault in Central Sulawesi (Indonesia)," *Terra Nova*, vol. 13, no. 6, pp. 463–470, 2001.
- [31] K. Bradley, R. Mallick, H. Andikagumi, J. Hubbard, E. Meilianda, A. Switzer, N. Du, G. Brocard, D. Alfian, B. Benazir, G. Feng, S. H. Yun, J. Majewski, S. Wei, and E. M. Hill, "Earthquake-triggered 2018 Palu Valley landslides enabled by wet rice cultivation," *Nature Geoscience*, vol. 12, no. 11, pp. 935–939, 2019. [Online]. Available: <http://dx.doi.org/10.1038/s41561-019-0444-1>
- [32] E. Agency. SNAP-ESA Sentinel Application Platform 2018. [Online]. Available: <http://step.esa.int/accessedon:2021June16>
- [33] P. Ge, H. Gokon, and K. Meguro, "Building damage assessment using intensity SAR data with different incidence angles and longtime interval," *Journal of Disaster Research*, vol. 14, no. 3, pp. 456–465, 2019.
- [34] G. Kuc and J. Chormański, "Sentinel-2 imagery for mapping and monitoring imperviousness in urban areas," *International Archives of the Photogrammetry, Remote Sensing and Spatial*

- Information Sciences - ISPRS Archives*, vol. 42, no. 1/W2, pp. 43–47, 2019.
- [35] P. E. Osgouei, S. Kaya, E. Sertel, and U. Alganci, “Separating built-up areas from bare land in mediterranean cities using Sentinel-2A imagery,” *Remote Sensing*, vol. 11, no. 3, pp. 1–24, 2019.
- [36] S. S. Bhatti and N. K. Tripathi, “Built-up area extraction using Landsat 8 OLI imagery,” *GIScience and Remote Sensing*, vol. 51, no. 4, pp. 445–467, 2014.
- [37] C. E. M. S. Mapping. Copernicus emergency management service - mapping. [Online]. Available: <https://emergency.copernicus.eu/mapping/list-of-components/EMSR317>(accessedon:2021June16)