International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

ISSN:2147-6799 www.ijisae.org Original Research Paper

Tsunami Building Damage Assessment using Multiclass Segmentation Model

Calvin Surya Widjaja[#1*], Alexander Agung Santoso Gunawan[)2], and Edy Irwansyah[)3]

Submitted: 10/09/2022 **Accepted:** 20/12/2022

Abstract: Natural Disaster is an event caused by environment, it has been concerned as it can caused casualties that makes manual damage assessment become inefficient. Automated damage assessment is one of field of study in Remote Sensing which already studied for several years, from using Traditional Machine Learning into Deep Learning. Recently, semantic segmentation with multitemporal fusion is a method used for Damage Assessment using Deep Learning. Multitemporal Fusion is a method fusing two features from Pre and Post Disaster Images as one using concatenation to get the feature of all two images. Semantic Segmentation is a method to classify each pixel in images into specified class given. This research creates Baseline Model (ResNet-50 + Panoptic FPN + Multitemporal Fusion) for comparison with our proposed method, called SCAMU-Net, which consists of U-Net (with different backbone, DenseNet 121, 169, and 201 layers) and followed by Spatial Channel Attention Module (SCAM) using xBD Dataset in Sunda and Palu Dataset. According to finding of the study, SCAMU-Net with DenseNet 121 shows biggest result in Macro F1 in Palu Dataset with 89.8% outperforms the Baseline Model about 3.6%. Sunda Dataset cannot perform for Training and Testing caused by destroyed class too few for Models to generalized. SCAMU-Net has 1,203,549 less parameters than baseline model. SCAMU-Net also good for detecting different class (No Damaged and Destroyed) that adjacent each other. Results shown that SCAMU-Net DenseNet 121 is enough for classify damage in this research, it shown that extending from DenseNet 121 provide no significant results.

Keywords: *Spatial Channel Attention, U-Net, Remote Sensing, Deep Learning, Semantic Segmentation*

1. Introduction

Natural Disaster is an event that caused by natural process without human intervention, one of the Natural Disaster is Tsunami. One of task in Disaster Recovery are damage assessment, which the task is how to identify the building and its classification based on damage levels. Damage assessment is suggested to be automated to minimize the risk because of post-disaster situation that can endanger the responder to enter the location, with automated, will reduce the cost, time and resources needed compared to manual damage assessment [1].

Navalgund et al. [2] stated that there are many applications in Remote Sensing, one of it is Urban Landuse which are Demography, Housing Quality, Traffic Modelling and Planning Utilities. Urban Landuse are used for detecting many things in urban area, such as buildings, park, school, cemetery, etc [3]. Remote Sensing data have contributed greatly in Landuse mapping, Remote Sensing utilize Satellite to capture Image Data for further processing, because it is utilizing Satellite then it is relatively safe to get the data for region impacted by disaster recently.

[#] *Computer Science Department, BINUS Graduate Program - Master of Computer Science, Bina Nusantara University, Jakarta, Indonesia, 11480*
[)] *Computer Science Department, School of Computer Science, Bina Nusantara University, Jakarta, Indonesia, 11530*
[1]*ORCID ID : 0000-0003-0831-2298*
[2]*ORCID ID : 0000-0002-1097-5173*
[3]*ORCID ID : 0000-0002-3876-1943*
**Corresponding Author Email: calvin.surya@binus.ac.id*

In the past few years, many studies use Machine Learning Methods to make classification of building damage with SVM [4]–[6], many Deep Neural Network (DNN) used for Damage Assessment. DNN as Nia et al. [7] stated outperforms Machine Learning (ML) Model (SVM as baseline) on ground-level image data. DNN approach for damage assessment mostly used in Literature Research are Classification [1], [7]–[15] and recently the studies are moved from Classification into Multiclass Segmentation [16]–[19]. Approach in this study is use Multiclass Segmentation, which are an approach to classifying each pixel on Image with collection of labels such as background, buildings, and roads.

Damage Assessment in Multiclass Segmentation also use Multitemporal Fusion, Weber et al. [17] use Pre-Disaster Images and Post Disaster Images extracted from one single model with shared weights, and its output features are stacked (concatenate) to get the features between pre and post. This method is significance, because it's increased F1 score about 0.283. Contribution in this paper is modifying state of the art from section 2 with Spatial Attention and Channel Attention Module after using Multitemporal Fusion module to increased damaged F1 Score. Firstly, this study will compare from the Benchmark models (Weber et al.) with additional Spatial and Channel Attention [20]. Second, this study also replaced the ResNet-50 + FPN model with U-Net + ASPP and replace the backbone from ResNet [21] into DenseNet [22], to increasing Macro Average F1 score.

The structure paper as follows: first section discusses about problem, motivation, approach of the paper, and contribution of this paper. Second section discuss about similar research in this study. Third section discuss about methodology about this paper, about datasets, pre-processing, augmentation, proposed method, and evaluation metrics. Fourth section discuss about results, comparison between model selected and benchmark selected. Fifth section discuss about conclusion, future works and limitations.

2. Related Works

In the early phase of this study, Cooner et al. [8] developed a comparison using three MLA (Machine Learning Algorithm), which are ANN (Artificial Neural Network), RBFNN (Radial Based Function Neural Network) and RF (Random Forest). Dataset used in that study are Haiti Earthquake 2010. The results are ANN outperforms the other two MLA with building omission error of 37.7%. Ji et al. [9] developed MLA, which is SVM (Support Vector Machine), dataset used in the study are Tokyo Tsunami. The dataset divided into 4 categories in block map which are SED (Slightly Damaged Area), MOD (Moderately Damaged Area), SLD (Slightly Damaged Area) and NOD (Undamaged Area), the study achieved accuracy of 88.81% to 92.28%.

Recently, the research of this study using Satellite Imagery to get the data, Bai et al. [23] created tsunami damage recognition with Deep Neural Networks. This study uses TerraSAR-X Satellite Imagery to get the data in Tohoku Earthquake, Japan, 2011. This study divided into 4 regions, which are non-build-up region, washed away region, collapsed region, and slightly damaged region. The study is divided into two phases, which are Built-Up Region Extraction (BRE) and Damage-level Recognition (DR). The model used in this study are SqueezeNet, WRN, ResNet-50 and AlexNet. The best result accuracy for BRE is 81.9% and DR is 74.9% which WRN best performed in this case.

In previous study above, the data is separated into region (block), more recently, the studies about Damage Assessment are directly classified the building to be more precise for damage rescue. Firstly, Gupta et al. [1] created benchmark dataset called xBD Dataset, which are consist of disaster images from 15 countries, dataset is captured using WorldView-2 satellite, the study developed two stage architecture with U-Net (Localization) and ResNet-50 (Classification) to tackle the problem. The model gets F1 0.2654 as Baseline Model. Baseline model suffers from Minor and Major damage class. (Liu et al., 2021) [15] developed two stages with SE-ResNeXt (Localization) and HRNet (Classification) with post only in stage two in xBD Dataset achieved F1 0.65.

More recently, the study of Damage Assessment moved from two phase (Localization and Classification) into one phase (Multiclass Segmentation). Firstly, Weber et al. [17] create two approaches for damage assessment, which are Instance Segmentation and Semantic Segmentation, with base model of Instance Segmentation of ResNet-50 + FPN. This study also used Multitemporal Fusion, which concatenate the output layer of each FPN from Pre-Disaster Feature and Post-Disaster Feature, and use Panoptic FPN as Segmentation Head, where the model is trained with same ResNet-50 + FPN (Shared Weights). The result of this study stated that Semantic Segmentation are suitable for Damage Assessment in xBD Dataset, because the buildings are too small to use Instance Segmentation. This study gets F1 score of 0.697 with Semantic Segmentation approach. Gupta et al. [24] create model called RescueNet, which consist of one model Dilated ResNet-50 + ASPP for feature extraction of Pre and Post Disaster Images and post disaster features are integrated with Encoder-Decoder style which concatenate feature between Post Disaster Features and Skip Connection from ResNet Block of Post Disaster to make classification. This study states the model gets F1 score 0.74. Although, the study says the model gets F1 0.74, when the model test with Tier 3 datasets, the F1 score are fall to 0.37 and Tier 1 validation dataset about 0.66 of F1 Score. The model is suffered significant degradation in Damage Assessment Task.

Based on the facts regarding good performance in Damage Assessment using xBD Dataset Benchmark, this study will use basic model from Weber et al. [17] as basic model architecture for this research, feature extraction will be replaced by model that perform the best in hyperparameter tuning to ensure the model achieved better F1 score than previous research without increasing trainable parameters exponentially.

3. Methodology

3.1. Dataset

In this study, Tsunami Sunda Dataset from xBD Dataset Tier 3 (https://xview2.org) and Tsunami Palu Dataset in Tier 1 are used for Training and Testing. Dataset Sunda consists of 148 images of Pre and Post Disaster Images. Dataset Palu consists of 69 images

Figure 1. Tsunami Sunda Strait and Palu 2018 Location

of Pre and Post Images. Dataset split into 2 parts, Train Dataset and Test Dataset with distribution in Table 1 and 2. Location of Tsunami Sunda and Tsunami Palu marked by red polygon in Fig 1.

Table 1. Distribution of Dataset Sunda

Dataset	Total
Train Dataset (90%)	133 Images
Test Dataset (10%)	15 Images

Table 2. Distribution of Dataset Palu

Dataset	Total
Train Dataset (88%)	54 Images
Test Dataset (22%)	15 Images

3.2. Data Preprocessing and Augmentation

Sunda Dataset extracted from original size of 1024x1024x3 into smaller patch of 256x256x3, with the stride of 128x128 when the Image does contain Damaged, and 256x256 when the Image does not contain Damaged (only Background). Dataset increased from 133 Images into 5191 Images for Train Dataset, and 15 Images into 528 Images for Test Dataset. Pre-Disaster Dataset have 2 class (Background and Building) and Post Disaster Dataset will be simplified into 3 class (Background, No Damaged and Destroyed). To balance the Minority Class (Destroyed), Oversampling will be used in this study. Oversampling can be seen in Fig 2.

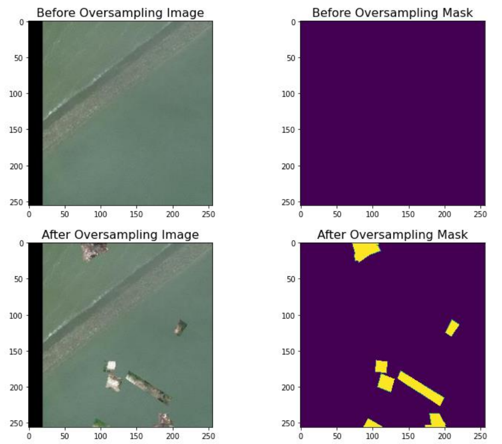

Figure 2. Oversampling Method To balance the Minority Class (Destroyed)

Dataset of Tsunami Palu extracted the same way as Sunda, but the stride instead using 128x128, the dataset cropped use stride of 64x64, when the Image does contain Destroyed category, it will use 128x128 if there is only no damaged class. Train Dataset increased from 54 Images into 3264 Images of Pre and Post Disaster Images. Under sampling and Oversampling are used in this study. 60% Images are used if there is not destroyed class (class 2) but there is no damaged / building (class 1) class. If no damaged class (class 1 and 2) is not present, the background image is discarded completely. In this study, the class will be simplified to 3 class, which are Background, No Damage and Destroyed. Minor Damaged and Major Damaged are converted into No Damaged. The dataset is reduced from 3264 into 1747 Images. Test Dataset increased from 15 Images (1024x1024) into 1024 Images (256x256) of Pre and Post Images. Firstly, Training Dataset are divided into 70% for Training (1222 Images) and Validation Dataset 30% (524 Images). After that, the dataset for training is rotated each of 30 degrees generated new patch. Steps per Epoch used in this study are 16000 images for Training, and 5000 for validation. Distribution of Sunda Dataset and Palu Dataset are displayed in figure 3 and 4.

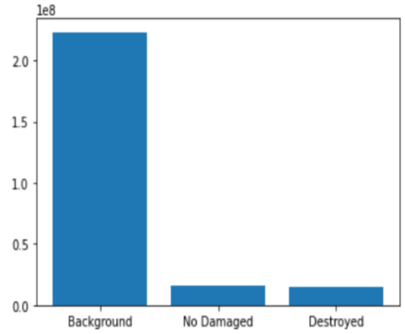

Figure 3. Distribution Sunda Dataset

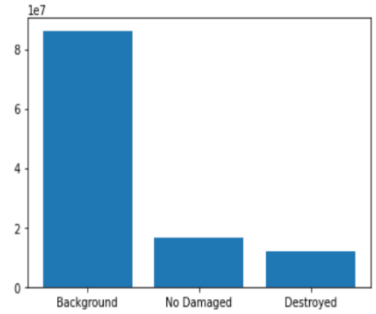

Figure 4. Distribution Palu Dataset

3.3. Proposed Method

This study proposed model named SCAMU-Net which consists of single segmentation model U-Net [25] with Hyperparameter Tuning included and extended with Spatial Channel Attention Module (SCAM) [20]. U-Net, in this study is a single U-Net with different backbone is used with DenseNet 121 Layers, 169 Layers and 201 Layers [22], because the recent studies in section 2 uses ResNet for feature extraction, this study will try to extend from ResNet with deeper backbone for optimal solution, that is DenseNet.

On top of that, U-Net will use ASPP (Atrous Spatial Pyramid Pooling) module in the end of encoder block to capture global context information. U-Net with ASPP can be seen in Figure 5. After ASPP used in U-Net Architecture, the module will use BottleNeck layer as shown in Figure 6. After that, U-Net Decoder (Expansion Path) configuration are used in this model shown in Figure 7. Concatenation in U-Net decoder consists of last layer of previous bottleneck, and concatenate with last concatenation layer in DenseNet block, to ensure all context information extracted by DenseNet can be used to classify the problem more correctly.

The following structure of U-Net Decoder BottleNeck Layer are Concatenation (from output of previous layer and output of ResNet block), Convolution 3x3 with N channels with BatchNormalization and ReLU activation function and followed

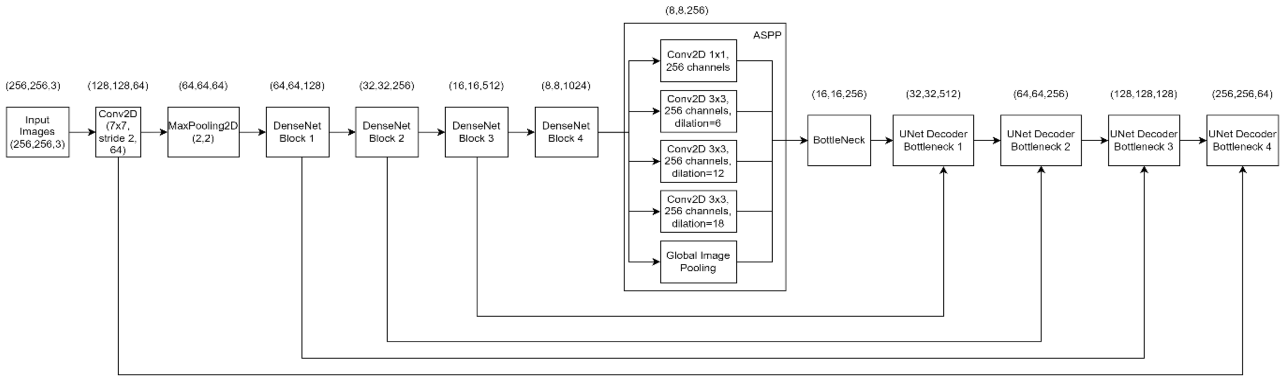

Figure 5. U-Net DenseNet Architecture with ASPP (Output Channels in DenseNet is 121 layers)

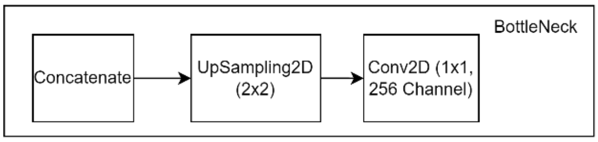

Figure 6. BottleNeck Layer

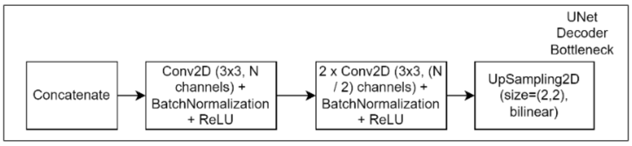

Fig 7. U-Net Decoder Bottleneck Layer

by two Convolution 3x3 with N/2 (half) channels (from first convolution layer in current bottleneck) with BatchNormalization and ReLU. N channels for each bottleneck are (1024,512,256, and 128) respectively.

This study proposed model named SCAMU-Net (Spatial and Channel Attention Module U-Net), which consist of Single U-Net with shared weights from configuration in Fig 5, for processing Pre-Disaster and Post-Disaster Images. Firstly, Pre-Disaster Images and Post-Disaster Images are trained with same model to get feature from each image (pre and post disaster image) and will be fusion into one with Concatenate Layers. After the Concatenation layers, instead the model predicts, the architecture extended by using Spatial Attention Module and Channel Attention Module inspired from Woo et al. [20] to learn what feature and where the location of feature is important to be selected.

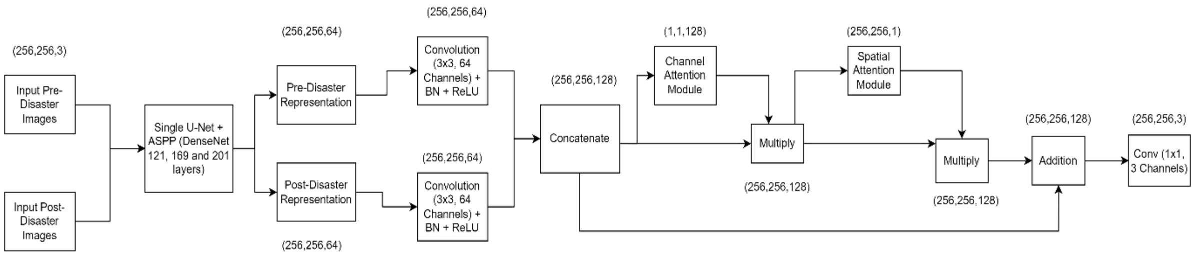

Figure 8. SCAMU-Net Architecture

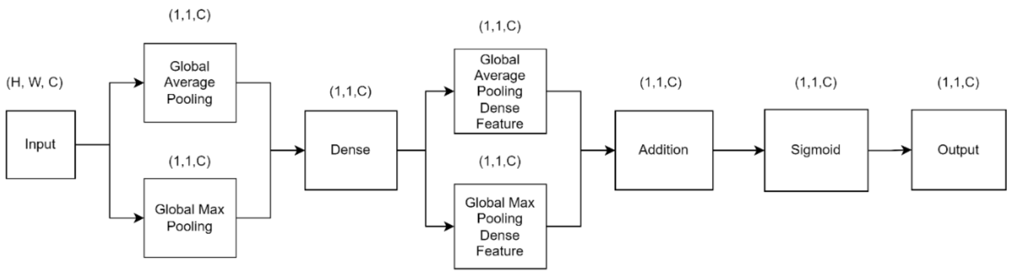

Figure 9. Channel Attention Module

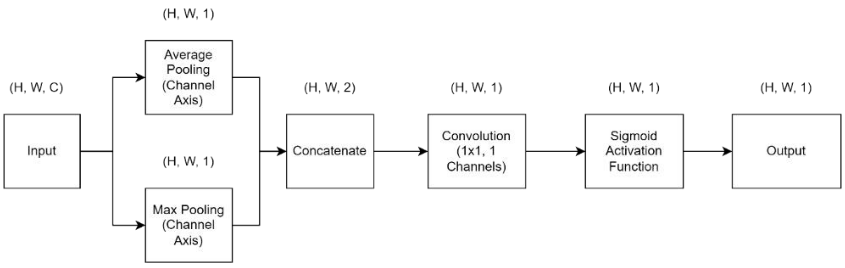

Figure 10. Spatial Attention Module

Base Architecture modified can be seen in Figure 8. In this study, Dense layers in Figure 9 for Channel Attention Module is modified, from single Dense Layer into double dense layer, with channels of [C * 2, C] respectively. Figure 9 and 10 denote detail of Channel and Spatial Attention Module used in the architecture. In this experiment, all models will be trained with same loss function which is categorical cross entropy, learning rate 1e-4, with Adam optimizer and same epochs of 10.

3.4. Evaluation Metrics

Evaluation Metrics for this study is Macro Average F1 Score will become the comparable metrics for baseline model and proposed model, which have formula as described below. Example for True Positive (TP), False Positive (FP), False Negative (FN) and True Negative (TN) for class 0 in Table 3.

Table 3. Confusion Matrix for Class 0

Label / Prediction	Class 0	Class 1	Class 2
Class 0	TP	FP	FP
Class 1	FN	TN	
Class 2	FN		TN

From Table 3, precision, recall and F1 Score formula for class i (0 to 2, in this study), described in equation (eq) 1, 2 and 3.

$$Precision(i) = \frac{TP_i}{TP_i + FP_i} \quad \text{(eq 1.)}$$

$$Recall(i) = \frac{TP_i}{TP_i + FN_i} \quad \text{(eq 2.)}$$

$$F1(i) = \frac{2 * Precision(i) * Recall(i)}{Precision(i) + Recall(i)} \quad \text{(eq 3.)}$$

Each of individual F1 score (class 0 to 2) will be summed and average, to get the final results as described in eq 4.

$$MacroF1 = \frac{\sum_{i=0}^{n} F1(i)}{(n+1)} \quad \text{(eq. 4)}$$

4. Results and Discussion

Firstly, SCAMU-Net is train using Sunda Dataset. The validation result and Test Result are shown in Fig 11 and Table 4. From the experiment, the results shown that, although the class 0 (background) and class 1 (no damaged) are retained, but class 2 are heavily impacted in the Test Set, the model cannot predict the Test Set that very different from Validation Set, feature from Train Dataset are too few so the model cannot generalize the feature in class 2 (destroyed). The result of oversampling in Training Dataset Sunda only makes the model memorize the pattern in Train Dataset, but worse in Test Set (very overfit). Hypothesis concluded that because too few Destroyed class in Sunda dataset, Sunda Dataset cannot use for training and testing. Alternatively, this study also using Palu Dataset to create damage assessment.

Fig 11. Macro F1 Validation Sunda Dataset SCAMU-Net DenseNet 169 Layers.

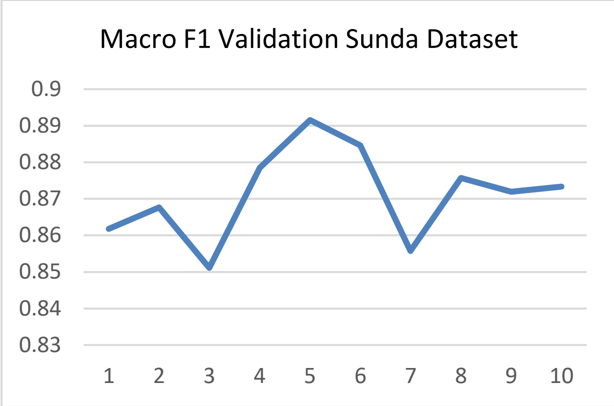

Table 4. Validation and Test Result of Dataset Sunda UNet DenseNet169

SCAMU-Net DenseNet169	F1 Background	F1 No Damage	F1 Destroyed	Macro Average F1
Best Validation Dataset Result	0.9881576	0.7552062	0.7552062	0.8915569
Test Dataset	0.9761409	0.7857033	0.0199892	0.5939445

Second, the models (baseline models and proposed method) are train with Palu Dataset, macro F1 of Validation in each model in Palu Dataset is described in Fig 12. The experiment shows in Validation dataset in each epoch, U-Net with each different DenseNet bottleneck (121, 169 and 201 layers) outperforms the Baseline Model.

Maximum Validation result are shown in Table 5 to assess individual Macro F1. Test results are shown in Table 5 to assess individual macro F1 of each model. From Table 5, Spatial and Channel Attention Module (SCAM) is sacrificing macro F1 by 0.1%, but it increases F1 of destroyed class (class 2) by 1% in Test Set, it means that SCAM focus in destroyed class more than without SCAM. From Table 5 and 6, SCAMU-Net with DenseNet is outperforms in this dataset from benchmark given, with whopping 2.3% to 2.8% in maximum validation results, and 3.2% to 3.6% in test results.

In this study, layers of DenseNet are extended from 121 to 201, although in Figure 14, validation results are increased slightly for different DenseNet, test results shown in Table 5 stated that the optimum layer of DenseNet is 121. SCAMU-Net also reducing total parameters from benchmark 43,653,251 to 42,449,702 which are 1,203,549 parameters less than the benchmark model.

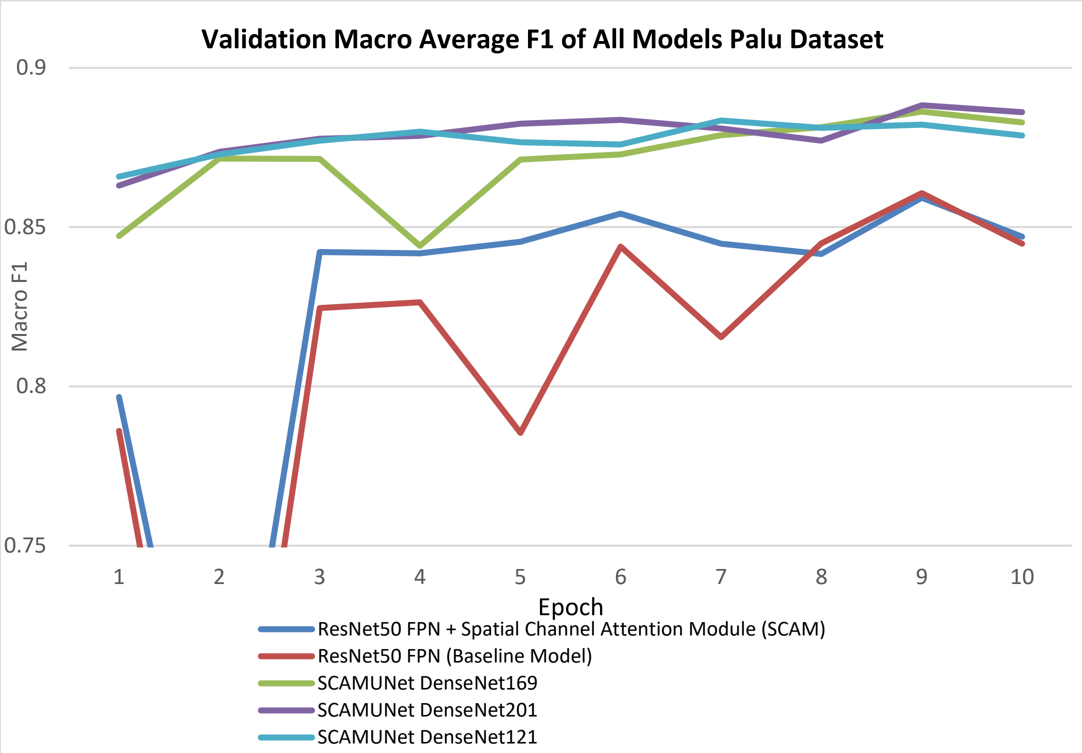

Figure 12. Validation Macro Average F1 of All Models in Palu Dataset

Table 5. Maximum Validation Result each models in Palu Dataset

Maximum Validation Dataset Result					
Model	F1 Background (Class 0)	F1 No Damage (Class 1)	F1 Destroyed (Class 2)	Macro Average F1	Total Number of Parameters
Baseline (Weber et al., 2020)	0,952955	0,791920	0,837206	0,86069	43,653,251
Baseline (Weber et al., 2020) + Spatial Channel Attention Module (SCAM)	0,952929	0,793140	0,831640	0,859236	49,161,132
SCAMU-Net (DenseNet-121)	0,962916	0,817383	0,870129	0,883476	**42,449,702**
SCAMU-Net (DenseNet-169)	**0,964829**	0,823015	0,870980	0,886275	54,870,822
SCAMU-Net (DenseNet-201)	0,964435	**0,824145**	**0,876207**	**0,888262**	62,450,470

Table 6. Test Result each model in Palu Dataset

Model	F1 Background (Class 0)	F1 No Damage (Class 1)	F1 Destroyed (Class 2)	Macro Average F1	Total Number of Parameters
Baseline (Weber et al., 2020)	0,981430	0,856837	0,750398	0,86289	43,653,251
Baseline (Weber et al., 2020) + Spatial Channel Attention	0,981293	0,843873	0,760476	0,861881	49,161,132
SCAMU-Net (DenseNet-121)	0,986257	**0,886659**	0,823825	**0,898914**	**42,449,702**
SCAMU-Net (DenseNet-169)	0,986096	0,885865	0,811915	0,894625	54,870,822
SCAMU-Net (DenseNet-201)	**0,986267**	0,882162	**0,824338**	0,897589	62,450,470

This study also random one of the images in Test Dataset to visualize the different between benchmark and proposed models perform, the pre disaster image and post disaster image can be seen in Fig 13,14. Ground Truth Label of post disaster image can be seen in Fig 15. The results between Benchmark model with SCAMU-Net DenseNet 121 (which the best test result in Table 5) can be seen in Fig 16 and 17.
Firstly, the image is cut with patches of size 256x256x3, from 1024x1024x3, making one image pre and post are converted into 16 different patches pre and post. The models are predicting 16 different patches, and then 16 predicted masks are stitches into one original mask.
From the results, Benchmark model misclassified some no damaged label into destroyed label, meanwhile proposed model is predicting almost every building correctly in fig 17, when comparing to Ground Truth in fig 15. SCAMU-Net also can detect destroyed building when in dense area, as seen in fig 17, compared to benchmark model which misclassified between No damaged class and destroyed class in fig 16.

Figure 13. Pre-Disaster Image

Figure 14. Post-Disaster Image

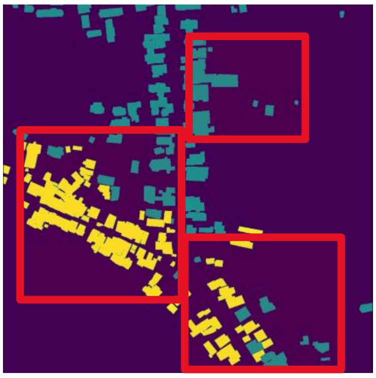

Figure 15. Ground Truth Label

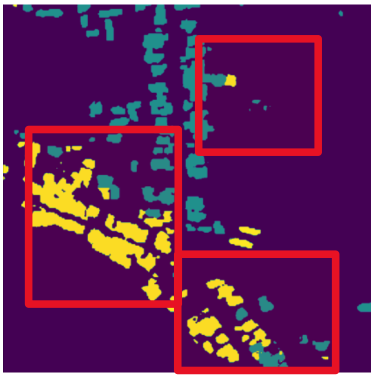

Figure 16. Prediction of Benchmark (ResNet50 + Panoptic FPN + Multitemporal Fusion)

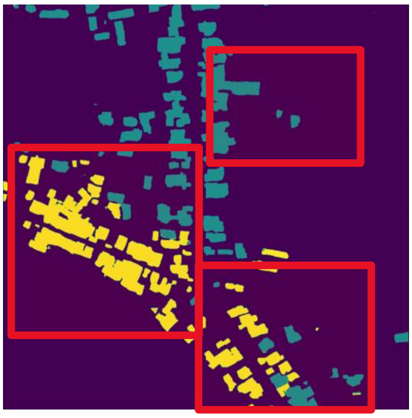

Figure 17. Prediction of SCAMU-Net (DenseNet-121)

5. Conclusion

In this research, the proposed method called SCAMU-Net classify damage with multiclass semantic segmentation approach method using satellite imagery, by comparing baseline model (ResNet-50 + Panoptic FPN + Multitemporal Fusion) with proposed model which are (UNet + Multitemporal Fusion + Spatial Channel Attention Module) with 3 different DenseNet backbone (DenseNet-121, 169 and 201 layers).

The dataset used in this study are Sunda Dataset and Palu Dataset in xBD Datasets. Sunda Dataset cannot give satisfying results in Destroyed class, because the class itself in Sunda Dataset are too few for deep learning model can generalize the feature, meanwhile in Palu Dataset, all model (including benchmark) is producing satisfying results.

In Palu Dataset, Spatial and Channel Attention Module (SCAM) are increasing F1 Score 1% on Destroyed class and sacrificing about 0.1% on Macro F1 Score, it means that SCAM is focusing on important features in Destroyed class which are minority class. SCAMU-Net (DenseNet-121) are outperformed baseline model by 3.6% in F1 Score and reducing its total parameters from benchmark by 2.75% (total of 1,203,549 less from benchmark model). Benchmark models are struggling classificate between no damage class and destroyed class when the building is mixed between two different class.

The suggestion for future works is, firstly, extend the model from only Tsunami models into all category's disasters model (such as volcano eruption, earthquake, hurricane, etc.) to ensure the model accurateness in damage assessment beside Tsunami disaster. Second, extend the model with different combination of State-of-the-Art model for Semantic Segmentation such as PSPNet [26], DeepLab V3[27]. Third, at the last concatenation layer on architecture, modificate the model by using Change Detection by subtracting the Pre-Disaster Image and Post Disaster Image and compare the result by using Concatenation and Change Detection.

Author contributions

Calvin Surya Widjaja: Methodology, Writing Original Draft Preparation, Implementation of Deep Learning Models. **Alexander Agung Santoso Gunawan:** Methodology, Verifying and Correcting Original draft, Validate Deep Learning Results, **Edy Irwansyah:** Verifying and Correcting Original Draft, Validate Deep Learning Results.

Conflicts of interest

The authors declare no conflicts of interest.

References

[1] R. Gupta *et al.*, "xBD: A dataset for assessing building damage from satellite imagery," *arXiv preprint arXiv:1911.09296*, 2019.

[2] R. R. Navalgund *et al.*, "Remote Sensing Data Acquisition, Platforms and Sensor Requirements," *Journal of the Indian Society of Remote Sensing 1996 24:4*, vol. 24, no. 4, pp. 207–237, Dec. 1996, doi: 10.1007/BF03026230.

[3] S. Srivastava, J. E. Vargas-Muñoz, and D. Tuia, "Understanding urban landuse from the above and ground perspectives: A deep learning, multimodal solution," *Remote Sens Environ*, vol. 228, pp. 129–143, Jul. 2019, doi: 10.1016/J.RSE.2019.04.014.

[4] P. Li, B. Song, and H. Xu, "Urban building damage detection from very high resolution imagery by One-Class SVM and shadow information," *International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 1409–1412, 2011, doi: 10.1109/IGARSS.2011.6049330.

[5] M. Shimada and A. Mita, "Damage assessment of bending structures using support vector machine," *https://doi.org/10.1117/12.600840*, vol. 5765, pp. 923–930, May 2005, doi: 10.1117/12.600840.

[6] Y. Bai, B. Adriano, E. Mas, and S. Koshimura, "Building Damage Assessment in the 2015 Gorkha, Nepal, Earthquake Using Only Post-Event Dual Polarization Synthetic Aperture Radar Imagery," *https://doi.org/10.1193/121516eqs232m*, vol. 33, no. Special issue 1, pp. S185–S195, Dec. 2017, doi: 10.1193/121516EQS232M.

[7] K. R. Nia and G. Mori, "Building damage assessment using deep learning and ground-level image data," *Proceedings - 2017 14th Conference on Computer and Robot Vision, CRV 2017*, vol. 2018-Janua, pp. 95–102, Feb. 2018, doi: 10.1109/CRV.2017.54.

[8] 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 Sens (Basel)*, vol. 8, no. 10, p. 868, Oct. 2016, doi: 10.3390/rs8100868.

[9] Y. Ji, J. T. S. Sumantyo, M. Y. Chua, and M. M. Waqar, "Earthquake/Tsunami Damage Assessment for Urban Areas Using Post-Event PolSAR Data," *Remote Sensing 2018, Vol. 10, Page 1088*, vol. 10, no. 7, p. 1088, Jul. 2018, doi: 10.3390/RS10071088.

[10] A. Calantropio, F. Chiabrando, M. Codastefano, and E. Bourke, "DEEP LEARNING for AUTOMATIC BUILDING DAMAGE ASSESSMENT: APPLICATION in POST-DISASTER SCENARIOS USING UAV DATA," *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 5, no. 1, pp. 113–120, Jun. 2021, doi: 10.5194/ISPRS-ANNALS-V-1-2021-113-2021.

[11] S. Ahadzadeh and M. R. Malek, "Earthquake Damage Assessment in Three Spatial Scale Using Naive Bayes, SVM, and Deep Learning Algorithms," *Applied Sciences 2021, Vol. 11, Page 9737*, vol. 11, no. 20, p. 9737, Oct. 2021, doi: 10.3390/APP11209737.

[12] D. Duarte, F. Nex, N. Kerle, and G. Vosselman, "SATELLITE IMAGE CLASSIFICATION OF BUILDING DAMAGES USING AIRBORNE AND SATELLITE IMAGE SAMPLES IN A DEEP LEARNING APPROACH," 2018, doi: 10.5194/isprs-annals-IV-2-89-2018.

[13] J. Z. Xu, W. Lu, Z. Li, P. Khaitan, and V. Zaytseva, "Building Damage Detection in Satellite Imagery Using Convolutional Neural Networks," Oct. 2019, doi: 10.48550/arxiv.1910.06444.

[14] C. S. Cheng, A. H. Behzadan, and A. Noshadravan, "Deep learning for post-hurricane aerial damage assessment of buildings," *Computer-Aided Civil and Infrastructure Engineering*, vol. 36, no. 6, pp. 695–710, Jun. 2021, doi: 10.1111/MICE.12658.

[15] C. Liu, L. Ge, and S. M. E. Sepasgozar, "Post-Disaster Classification of Building Damage Using Transfer Learning," pp. 2194–2197, Oct. 2021, doi: 10.1109/IGARSS47720.2021.9554795.

[16] Z. M. Hamdi, M. Brandmeier, and C. Straub, "Forest Damage Assessment Using Deep Learning on High Resolution Remote Sensing Data," *Remote Sensing 2019, Vol. 11, Page 1976*, vol. 11, no. 17, p. 1976, Aug. 2019, doi: 10.3390/RS11171976.

[17] E. Weber and H. Kané, "Building Disaster Damage Assessment in Satellite Imagery with Multi-Temporal Fusion," Apr. 2020, doi: 10.48550/arxiv.2004.05525.

[18] Y. Bai *et al.*, "Pyramid Pooling Module-Based Semi-Siamese Network: A Benchmark Model for Assessing Building Damage from xBD Satellite Imagery Datasets," *Remote Sensing 2020, Vol. 12, Page 4055*, vol. 12, no. 24, p. 4055, Dec. 2020, doi: 10.3390/RS12244055.

[19] Y. Bai, E. Mas, and S. Koshimura, "Towards operational satellite-based damage-mapping using U-net convolutional network: A case study of 2011 Tohoku Earthquake-Tsunami," *Remote Sens (Basel)*, vol. 10, no. 10, p. 1626, Oct. 2018, doi: 10.3390/rs10101626.

[20] S. Woo, J. Park, J. Y. Lee, and I. S. Kweon, "CBAM: Convolutional Block Attention Module," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11211 LNCS, pp. 3–19, Jul. 2018, doi: 10.48550/arxiv.1807.06521.

[21] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Dec. 2016, vol. 2016-Decem, pp. 770–778. doi: 10.1109/CVPR.2016.90.

[22] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, "Densely Connected Convolutional Networks," *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, vol. 2017-January, pp. 2261–2269, Aug. 2016, doi: 10.48550/arxiv.1608.06993.

[23] Y. Bai *et al.*, "A framework of rapid regional tsunami damage recognition from post-event terraSAR-X imagery using deep neural networks," *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 1, pp. 43–47, Jan. 2018, doi: 10.1109/LGRS.2017.2772349.

[24] R. Gupta and M. Shah, "RescueNet: Joint building segmentation and damage assessment from satellite imagery," *Proceedings - International Conference on Pattern Recognition*, pp. 4405–4411, 2020, doi: 10.1109/ICPR48806.2021.9412295.

[25] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," *ArXiv*, pp. 1–8, 2015, Accessed: Oct. 01, 2022. [Online]. Available: http://lmb.informatik.uni-freiburg.de/%0Aarxiv:1505.04597v1

[26] H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia, “Pyramid Scene Parsing Network,” 2017, pp. 2881–2890. Accessed: Dec. 20, 2021. [Online]. Available: https://openaccess.thecvf.com/content_cvpr_2017/html/Zhao_Pyramid_Scene_Parsing_CVPR_2017_paper.html

[27] L.-C. Chen, G. Papandreou, F. Schroff, and H. Adam, “Rethinking Atrous Convolution for Semantic Image Segmentation,” Jun. 2017, Accessed: Jun. 01, 2021. [Online]. Available: http://arxiv.org/abs/1706.05587