16th World Congress in Computational Mechanics (WCCM)
21-26 July 2024, Vancouver, Canada
A. Korobenko, M. Laforest, S. Prudhomme, R. Vaziri (Eds)

# DEVELOPMENT OF A LAND CLASSIFICATION MODEL BASED ON SEMANTIC SEGMENTATION USING AERIAL PHOTOGRAPHS AND ITS APPLICATION TO TSUNAMI SIMULATION

**YUTO HABUTSU [1], TOMOHIRO MIYAKE[2], HIROSHI OKAWA[3]**

**AND KAZUO KASHIYAMA[4]**

[1] Graduate School of Civil, Human and Environmental Engineering, Chuo University
1-13-27, Kasuga, Bunkyo-ku, Tokyo, Japan
a19.6t5r@g.chuo-u.ac.jp

[2] Graduate School of Civil, Human and Environmental Engineering, Chuo University
1-13-27, Kasuga, Bunkyo-ku, Tokyo, Japan
a20.bwew@g.chuo-u.ac.jp

[3] EJ Innovation Technology Center, Eight-Japan Engineering Consultants Inc.
5-33-11 Honcho, Nakano-ku, Tokyo 164-8601, Japan
ookawa-hi@ej-hds.co.jp

[4] Department of Civil and Environment Engineering, Chuo University
1-13-27, Kasuga, Bunkyo-ku, Tokyo, Japan
kaz@civil.chuo-u.ac.jp

**Key words:** Deep Learning, Land Use Classification, FEM, Tsunami Simulation.

**Abstract.** This paper presents a development of land use classification model based on semantic segmentation using aerial photographs and its contribution to the efficiency of 2D tsunami inundation simulations. The proposed method uses Artificial Intelligence(AI)-based image classification to generate a roughness coefficient mesh, which is then applied to a 2-D tsunami run-up simulation using the finite element method on real geometry. Numerical results are compared to evaluate the improvement in simulation efficiency, and the potential benefits of the proposed method are discussed by analyzing the differences in simulation results.

## 1 INTRODUCTION

Land use data is an important component of geospatial information and has been applied in various fields such as tsunami, storm surge, and flood inundation simulation and urban development planning. However, these data have been of limited use due to their low update frequency and the difficulty of obtaining high-resolution data. In previous work [1], the main focus was on building single-category segmentation models for roads and buildings, which limited the data set.

We have conducted numerical analysis [2] of tsunami, storm surge, and flood inundation. However, Manning's roughness coefficient for each land use category has been given semi-automatically by hand, which is a very time-consuming task.

In this study, we aim to generate high-resolution, multi-category land use data using aerial photographs and the deep learning model is Pyramid Scene Parsing Network (PSPNet) [3], and apply it to 2D tsunami run-up analysis. For deep learning, we focus on data augmentation, batch size, and Convolutional Neural Network (CNN) hierarchy to quantitatively evaluate the impact of these parameters on classification results. Then, to verify the performance of multiple classification models built with different parameters, the proposed model will be applied to aerial photographs of the target area. Finally, the contribution of the segmentation results to the efficiency of the 2D tsunami run-up simulation is verified by comparing simulation results with computational costs.

## 2 DEVELOPMENT OF DEEP LEARNING MODELS

This study is a development of deep learning models based on a supervised learning method. Figure 1 shows the flow of the development and validation procedure of the deep learning model. First, in the pre-process, aerial photographs and land use data are obtained and used to create training data to be input into the deep learning model. Next, in the main process, the deep learning model is set up, training conditions are set, and training is performed. Finally, in the post-process, Manning's roughness coefficient data are applied to the 2D tsunami run-up simulation to evaluate the simulation results. The following sections present the details of the aforementioned processes.

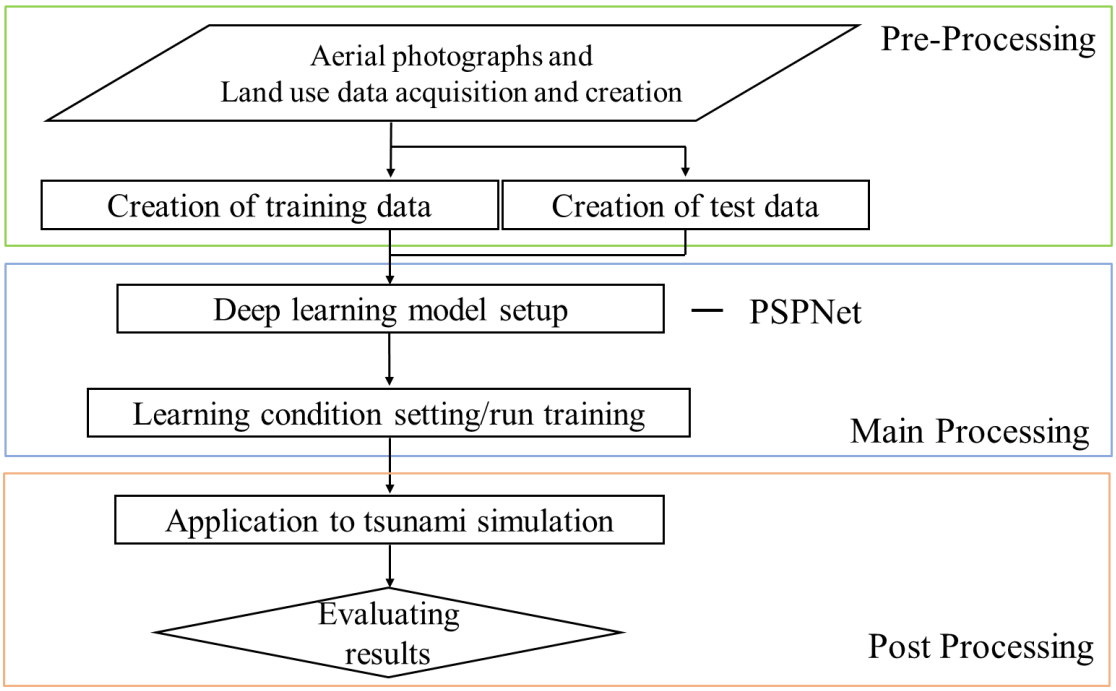


Figure1: Flowchart of proposed method

## 2.1 PREPARATION OF DATA FOR DEEP LEARNING

### 2.1.1 LAND USE DATA

In this study, five land use categories shown in Table 1 are used as label categories for deep learning. Five label categories were set: "Buildings," "Fields," "Forests," "Oceans/Rivers," and "Roads. Table 1 shows the relationship between Manning's roughness coefficients allocated from the land use.

**Table 1**: Relation between land use and Manning's roughness coefficient

| Land use categories | Manning's roughness coefficient |
|---|---|
| Buildings | 0.040 |
| Fields | 0.020 |
| Mountains/Forests | 0.030 |
| Oceans/Rivers | 0.025 |
| Roads | 0.025 |

### 2.1.2 DATA SET CREATION MEETHOD

In this study, training and test data were created using GIS software. The training data was created in the Kinki region, where the latest (2008) Numerical Map 5000 was available. Numeric Map 5000 is land use data. Test data were created in kure district, Nakatosa Town, Kochi Prefecture, and used for 2D tsunami run up simulation. Figure 2, 3, and 4 show how the data sets were created. As shown in Figure 2 and 3, the training data was created by superimposing aerial photographs, Numerical Map 5000, and 100m mesh on GIS software, and then processing the data to create original and label images((b) in Figure 2 and 3). The same procedure was used for the test data as shown in Figure 4. However, there is no Numerical Map 5000 available for the subject area, a land use data ((b) in Figure 4) was manually created using GIS software as an alternative.

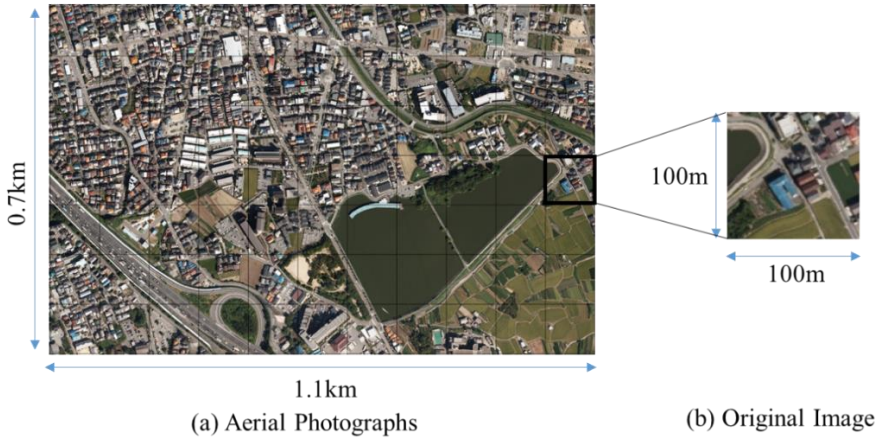


Figure 2: Creating training data (original image)

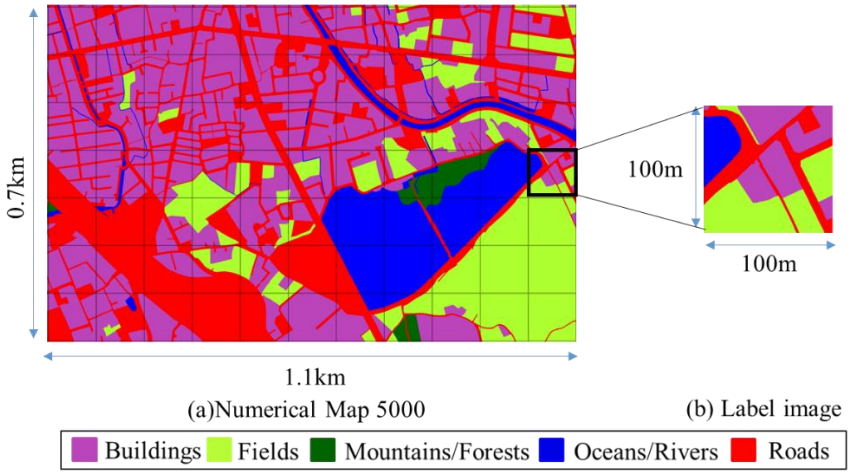


Figure 3: Creating training data (label image)

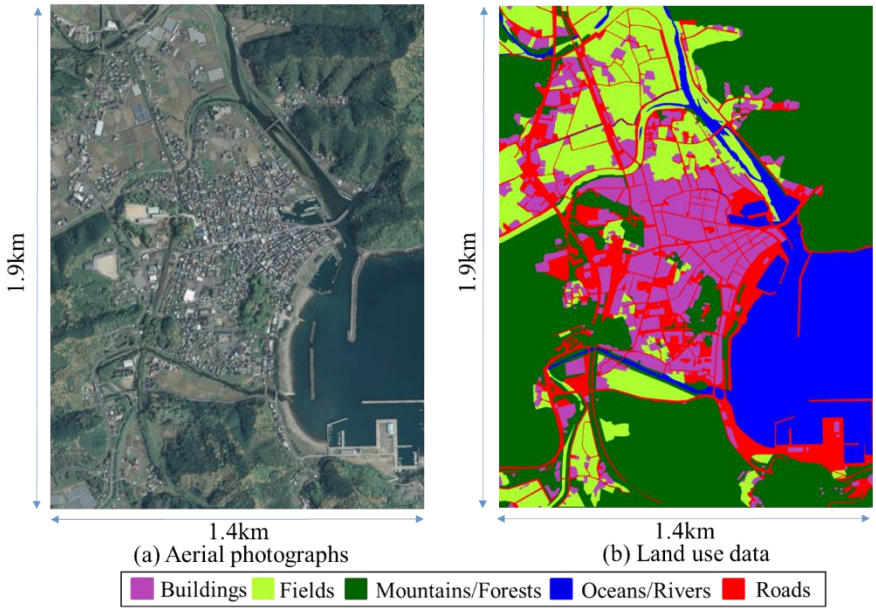


Figure 4: Creating test data (original and label image)

## 2.2 DEEP LEARNING MODELS

### 2.2.1 PYRAMID SCENE PARSING NETWORK

Deep learning is a machine learning technique that uses neural networks that mimic the details of the human brain.

In this study, a deep learning-based image classification method is used for land use classification. For deep learning, we used PSPNET, which is possible to capture both global and local fine-grained features. Figure 6 shows an overview of the PSPNET network model. Feature maps of different scales are obtained by Pyramid Pooling. The obtained feature maps of different scales are up-sampled to the same scale as the original feature maps. The up-sampled feature maps are concatenated by adding channels to the original feature maps.

This allows us to obtain not only global features but also local features. The concatenated feature maps are then convolved with a filter size of 1*1 to obtain the semantic segmentation result.

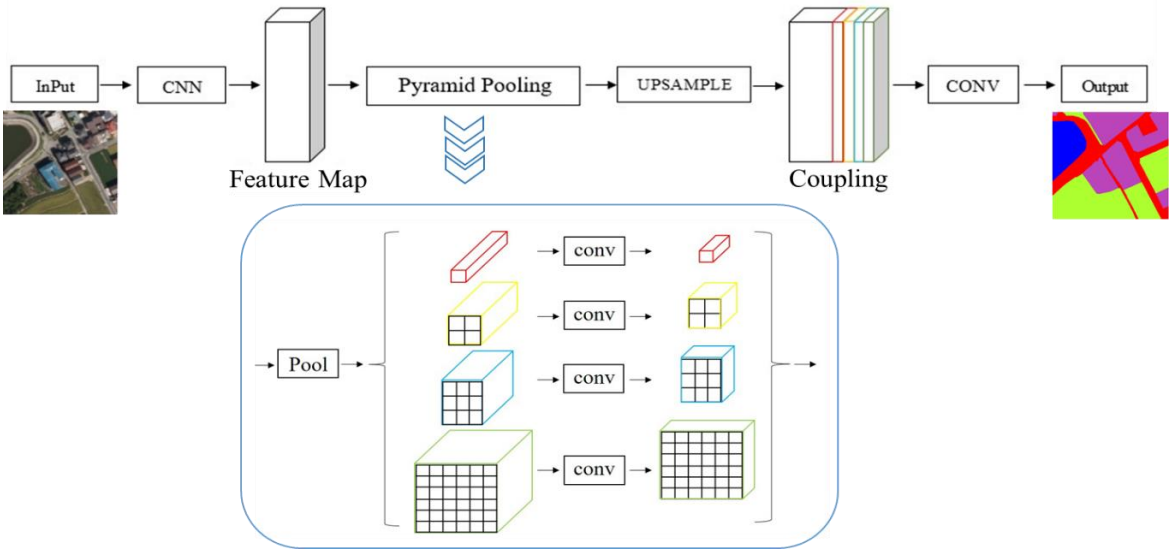


Figure 5: Overview of deep learning model

### 2.2.2 SETTING UP DEEP LEARNING CONDITIONS

We used the data set created from 30,040 pairs of training data and 266 pairs of test data for training, using the data creation method shown in Section 2.1.1. The training data was separated into training data and validation data at a ratio of 8:2. Table 2 shows the training conditions and Figure 6 shows an example of data augmentation. In this study, the best deep learning model is investigated by rotating the image by 60 degrees as an augmentation of the training data. The training epoch was set to 75 times. Other conditions were Residual Neural Networks (ResNet)-34 or 50[4] as the CNN model, Cross Entropy as the loss function, and Adam as the optimization algorithm. The ReLU function was used as the activation function. However, softmax function was used for the activation function before output. The learning results were evaluated using Pixel Accuracy, F1 Score, and Intersection over Union (IOU) to compare the accuracy and validate the deep learning models.

**Table 2**: Deep learning conditions

| Conditions | CNN model | Data Augmentation | Batch size |
|---|---|---|---|
| 1 | Resnet-34 | - | 32 |
| 2 | | | 64 |
| 3 | | Randam 60° rotation | 32 |
| 4 | | | 64 |
| 5 | Resnet-50 | - | 32 |
| 6 | | | 64 |
| 7 | | Randam 60° rotation | 32 |
| 8 | | | 64 |

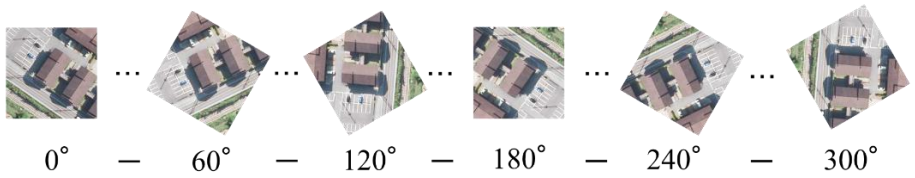


Figure 6: Example of Data Augmentation

## 2.3 DEEP LEARNING RESULTS

### 2.3.1 Training Results

We developed deep learning models as shown in Table 2 for the learning conditions in Table 2 and compared the accuracy of the models. Figure 7 shows the learning transition of the models that were trained under all conditions. The results show that learning is performed without difference in all conditions. In addition, a comparison of the F1 Score shown in Table 3 shows that the F1 Score was close to 0.8 in almost all conditions. This confirms that the model can be learned with high accuracy in all conditions for developing the model.

Next, in order to examine a more optimal model, transfer learning is performed under all conditions using test data. The results are shown in the next section.

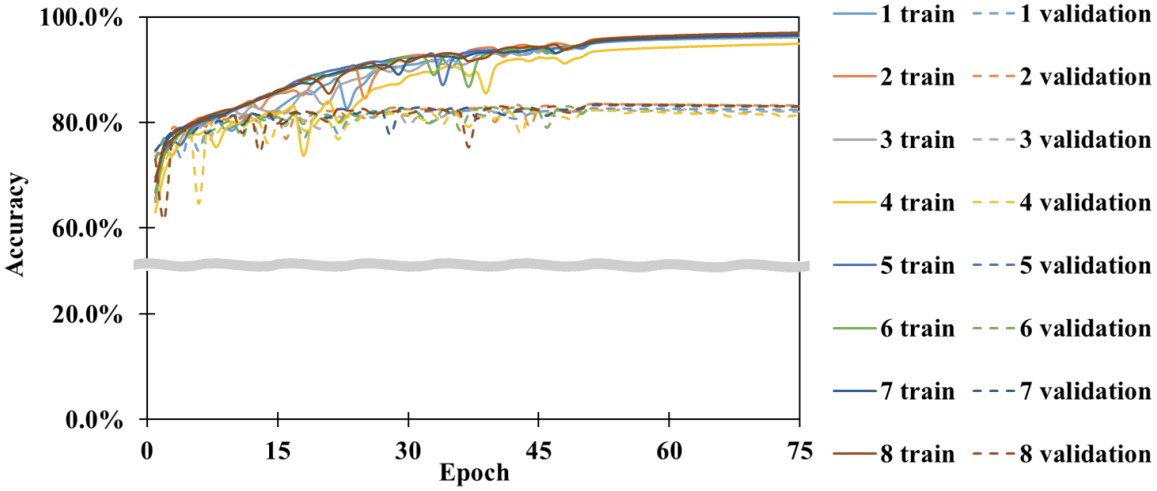


Figure 7: Learning transition of deep learning models

**Table 3**: Training learning results

| Conditions | CNN | Data Augmentation | Batch size | F1 |
|---|---|---|---|---|
| 1 | Resnet 34 | - | 32 | 0.83 |
| 2 | | | 64 | 0.68 |
| 3 | | 60° rotation | 32 | 0.82 |
| 4 | | | 64 | 0.79 |
| 5 | Resnet 50 | - | 32 | 0.82 |
| 6 | | | 64 | 0.72 |
| 7 | | 60° rotation | 32 | 0.78 |
| 8 | | | 64 | 0.82 |

### 2.3.2 Transfer Learning Result

We applied transfer learning to the developed models shown in the previous section using the test data. Table 4 shows the Accuracy, F1 Score, and Mean IoU of the transfer learning results. The results show that Condition 7 has the highest accuracy in the three evaluation metrics and is the most optimal model in this study. In addition, the classification accuracy tended to improve as the CNN hierarchy was deepened. Better classification accuracy results were obtained when the batch size was set to 32 rather than 64. For the data expansion of random 60° rotation, the results were effective when the CNN hierarchy was deeper. Based on these results, an image visualizing the prediction results for condition 7 is shown in Figure 8, and the IoU for all categories is shown in Table 5. Comparing the correct and predicted labels, it can be seen that they are generally in agreement ((a) and (b) in Figure 8).The IoU results in Figure 9 show that the IoU for fields and roads is below 0.5, and it is necessary to study the development of an optimal model that will improve the accuracy of these two categories. Furthermore, considering the application of these prediction results to tsunami simulation, it is also important to improve the classification accuracy of buildings with the largest Manning's roughness coefficient.

In the next section, we will apply the most accurate condition 7 transition learning results to the tsunami simulation.

**Table 4**: Transfer learning results

| Conditions | CNN | Data Augmentation | Batch size | Evaluation metrics | | |
|---|---|---|---|---|---|---|
| | | | | Accuracy | $F_1$ | Mean IoU |
| 1 | Resnet 34 | - | 32 | 0.64 | 0.66 | 0.47 |
| 2 | | | 64 | 0.65 | 0.66 | 0.49 |
| 3 | | 60° rotation | 32 | 0.65 | 0.68 | 0.48 |
| 4 | | | 64 | 0.59 | 0.62 | 0.43 |
| 5 | Resnet 50 | - | 32 | 0.67 | 0.70 | 0.51 |
| 6 | | | 64 | 0.59 | 0.63 | 0.44 |
| 7 | | 60° rotation | 32 | 0.69 | 0.71 | 0.52 |
| 8 | | | 64 | 0.66 | 0.68 | 0.50 |

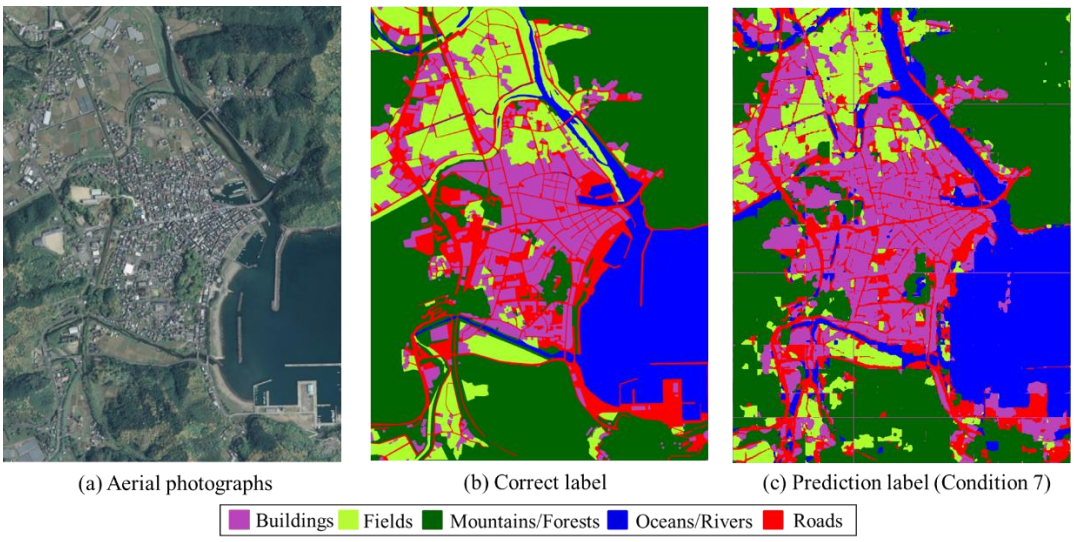


Figure 8: Visualized image of the prediction results

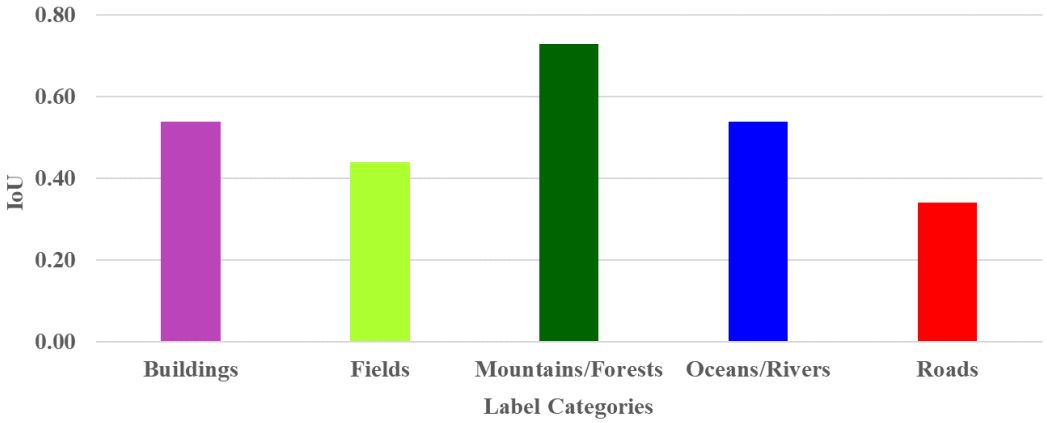


Figure 9: IoU of transition learning results for condition 7

## 3 APPLICATION TO 2D TSUNAMI RUN-UP SIMULATION

### 3.1 NUMERICAL METHODS

The governing equation for tsunami is the nonlinear shallow-water equations based on the nonlinear long-wave theory shown below.

$$\frac{\partial \mathbf{U}}{\partial t} + \mathbf{A}_i \frac{\partial \mathbf{U}}{\partial x_i} - \frac{\partial}{\partial x_i}\left(\mathbf{N}_{ij} \frac{\partial \mathbf{U}}{\partial x_i}\right) = \mathbf{R} - \mathbf{G}\mathbf{U} \tag{1}$$

where **U** is the unknown vector, **R** is the gradient vector, $\mathbf{A}_i$is the advection matrix, $\mathbf{N}_{ij}$ is the diffusion matrix, **G** is the friction matrix.

A stabilized finite element method based on the SUPG method is used for discretization in the spatial direction, the Crank-Nicolson method with second-order accuracy is used for discretization in the temporal direction, and the Bi-CG STAB method with Element-By-Element processing is used to solve simultaneous linear equations.

We applied a moving boundary condition based on the Eulerian method [5], which is excellent in adaptability to arbitrary shapes and relatively easy to algorithmically apply.

### 3.2 NUMERICAL CONDITIONS

The mesh shown in Figure 10 was used. As shown in Figure 10, the analytical mesh, correct and prediction labels were superimposed on the GIS software, and roughness coefficients were allocated to the triangular elements. The predictive labels were the results of Condition 7, which gave the most accurate results in the transfer learning shown in the previous section.

The method of allocation is as follows

1) Allocate Manning's roughness coefficients to each node of the triangular element from the Correct label and Prediction label based on the land use data（Figure 2）.
2) After allocating the coefficients, the average of the three roughness coefficients for the three nodes of the triangular element is calculated and reflected in the analysis mesh.

This created an analytical mesh possessing Manning's roughness coefficients.

In this study, as an initial condition, we referred to the fault parameters of the Nankai Trough

earthquake with a large slip zone off the coast of Shikoku (Case 4), calculated by the method of Manisinha\&Smylie[6] and used them as initial waveforms.

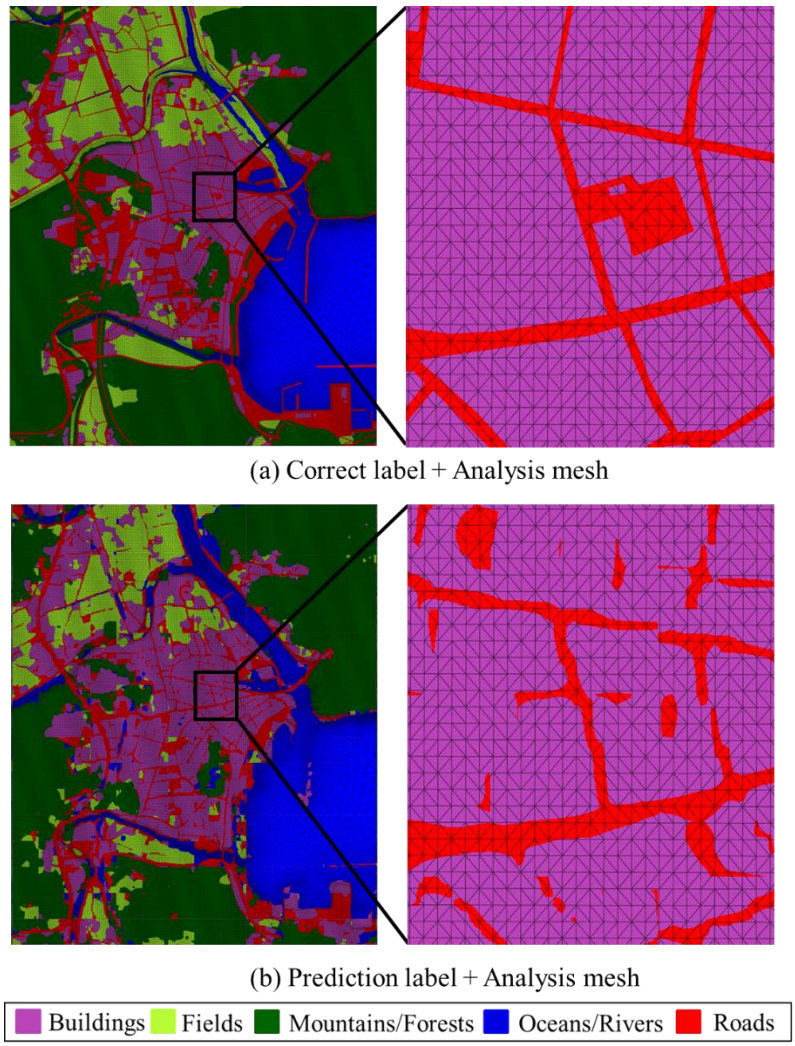


Figure 10: Comparison of the mesh of correct labels and the mesh of prediction labels

### 3.3 Application examples

We are conducted 2D tsunami run up simulation for the Kure area in Nakatosa town, Kochi Prefecture, Japan. The mesh shown in the previous section will be used to compare the run-up area and run-up speed. Details of the results will be presented at the presentation.

## 4 CONCLUSIONS

In this paper, we have developed a deep learning land use classification model using aerial photographs and quantitatively evaluated the effect of different parameters on the classification accuracy. To examine the effectiveness of the method, the segmentation results were applied to a tsunami simulation to verify its contribution to the efficiency of the simulation, and obtained the following conclusions.

- Deep learning can generate high-resolution and multi-category land use data, and the application of the segmentation results to tsunami simulation improves the efficiency of the simulation better than before.
- The deep learning model is deepened and the batch size is not increased, which is confirmed to be effective for this method.

As future work, we plan to apply the deep learning model to improve its accuracy and to evaluate its generalization performance in other domains.

## REFERENCES

[1] Hiroyuki Ohno, Ryo Endo, Takayuki Nakano, Masako Shinoda: Construction of dataset for feature extraction performance evaluation using aerial photographs, *The 33rd Annual Conference of the Japanese Society for Artificial Intelligence*, 2019.

[2] Daisuke Tonegawa, Kazuo Kashiyama: Development of a numerical method for tsunami runup and fluid force based on stabilized finite element method, *Journal of Applied Mechanics*, Vol.12, pp.127-134, 2009.

[3] Zhao, H., Shi, J., Qi, X., Wang, X., and Jia, J.: Pyramid scene parsing network, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.2881-2890, 2017.

[4] Kaiming He et al.: Deep residual learning for image recognition, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.

[5] Kawahara, M. and Umetsu, T. (1986). Finite element method for moving bound ary problems in river flow, *International Journal for Numerical Methods in Flu ids*, 6, pp. 365-386.

[6] Manshinha, L. and D. E. Smylie (1971): The Displacement field of inclined faults, *Bulletin of the Seismological Society of America*, Vol. 61, No.5, pp.1433‐1440.