XXVII FIG CONGRESS
11–15 SEPTEMBER 2022
Warsaw, Poland

Volunteering for the future – Geospatial excellence for a better living

Generating Cadastral Maps from Aerial Photos of Rural Areas

Wonho Song, South Korea
(LX Education Institute)

Purpose of Presentation

- To generate a cadastral maps from the aerial images.
- To check whether cGAN(Conditional Generative Adversarial Network) is worth to use or not.

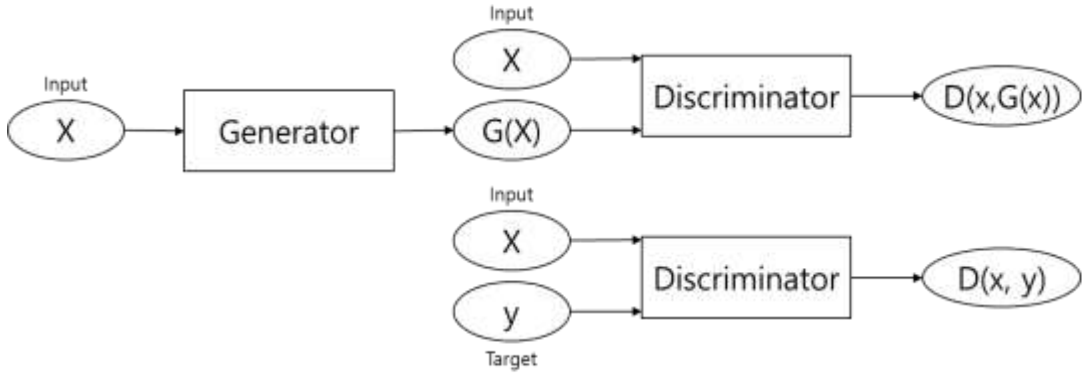
Introduction

- Artificial intelligence is a very effective tool in image processing fields such as segmentation and object detection.
- In fact, deep learning shows better results than existing image processing technology in many computer vision problems.
- CNNs and GANs have become commonly used tools for these various image processing and prediction problems.
- The cGAN model is one of a general solution of the image translation task.

- This study performed image conversion to recognize cadastral boundaries(roads) from aerial images using cGAN.

Theory-cGAN

- cGAN is a composed model of generator G and discriminator D. This model learns the mapping from input image x to output image y, G: x → y. Adversarially trained discriminator D discriminates authenticity of Generator G. Figure below is the illustration of cGAN architecture. As training continues by the G and D, the converted images gradually become closer to the original images.



Theory

- The cGAN learns a model using input of pair images and generates output images. As a generator, ResNet-9 model shows good results so far. ResNet-101 is a deeper version of ResNet.
- In order to test the effect of the loss function, test the BCE loss and the MSE loss.
- As a discriminator, PatchGAN has excellent performance and uses small resources. ImageGAN is a method of scoring the entire image.
- For overall quantitative evaluation of the translated image, it uses MSE (Mean Squared Error).

- The main goals of this study are:
  - Test the loss function on the cGAN model. (BCE vs MSE)
  - Test the generator ResNet-9 and ResNet-101.
  - Test discriminator PatchGAN and ImageGAN.
  - Overall performance is quantitatively evaluated by MSE indicators.

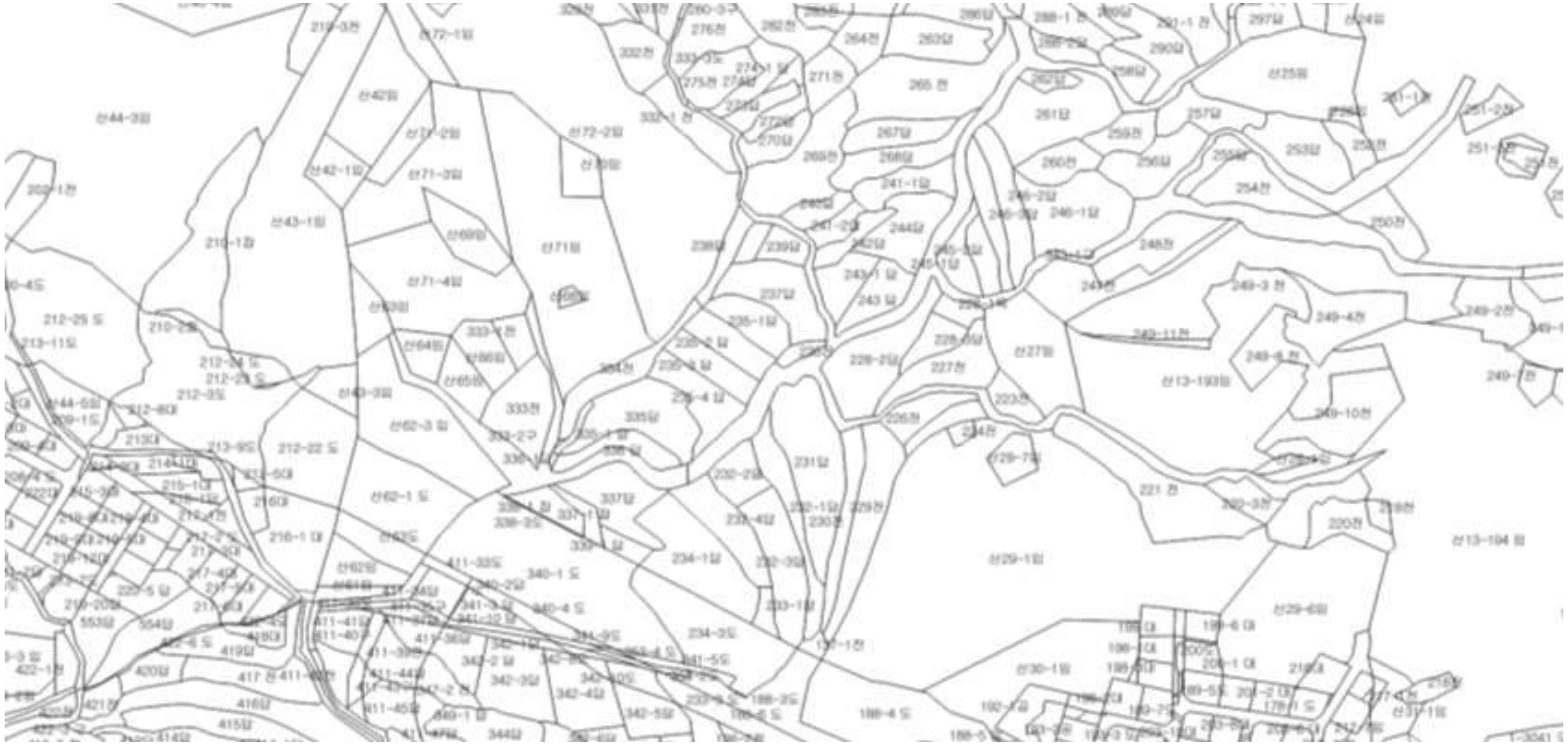
Methods-Preperation

- Data sets are half-meter-level aerial images of areas below and corresponding cadastral maps. The specifications of the aerial images are as follows.

Images	Y1	X1	Y2	X2	Band	X Resol	Y Resol
Gunpo	193046.79	525032.36	197579.29	530683.36	X: 9065 Y: 11302 Band: 3	0.500000	- 0.500000
Bucheon	177578.93	541691.20	184329.93	550132.20	X: 13502 Y: 16882 Band: 3	0.500000	-0.500000
Sungnam	210749.83	536136.98	215283.33	541793.98	X: 9067 Y: 11314 Band: 3	0.500000	- 0.500000
Yeoju	235179.13	403919.51	268051.39	437707.28	X: 65744 Y: 67575 Band: 3	0.500004	- 0.500004
Icheon	228694.65	391536.55	257781.88	430492.85	X: 58174 Y: 77912 Band: 3	0.500004	- 0.500004
PyoungTack	179845.23	377560.94	215563.02	405412.67	X: 71435 Y: 55703 Band: 3	0.500004	- 0.500004

- Divide it into 256 x 256 tiles and resize the input image to 286 x 286 for the random cropping and did normalization.

Methods-Example of cadastral map



Methods-Task environment

- The computer specification is as follows.
  - 1. CPU i7, Graphic GTX 1660, RAM 16G
  - 2. CPU i5, Graphic GTX 1060 3G, RAM 8G

- Software's used are Python, Tensorflow, PyTorch, etc. Each epoch generates a checkpoint to check the detail checkpoint status. The processing time using GPU was about 30 minutes per an epoch.

- Splits the dataset into 80% of training dataset and 20% of validation dataset, respectively. Number of images is about 9.5 thousand in rural areas. The images are the RGB channel images

Study Area	Epoch	Sum	80%	20%	Cities
Rural	100	9,515	7,612	1,903	Bucheon, Yeoju, Icheon ,PyoungTack

Methods-Training

- use Mini-batch SGD and ADAM solver optimizer. Hyper-parameters are below.
  - batch size to 1
  - dropout ratio to 0.5
  - learning rate of 0.0002
  - GAN loss to BCE or MSE
  - G model is ResNet-9 or ResNet-101
  - D model is PatchGAN or ImageGAN
  - Each model runs 100 epochs.

  - (Optional) Dropout 0.5 → 0.8
  - (Optional) Learning rate 0.0002 → 0.0004

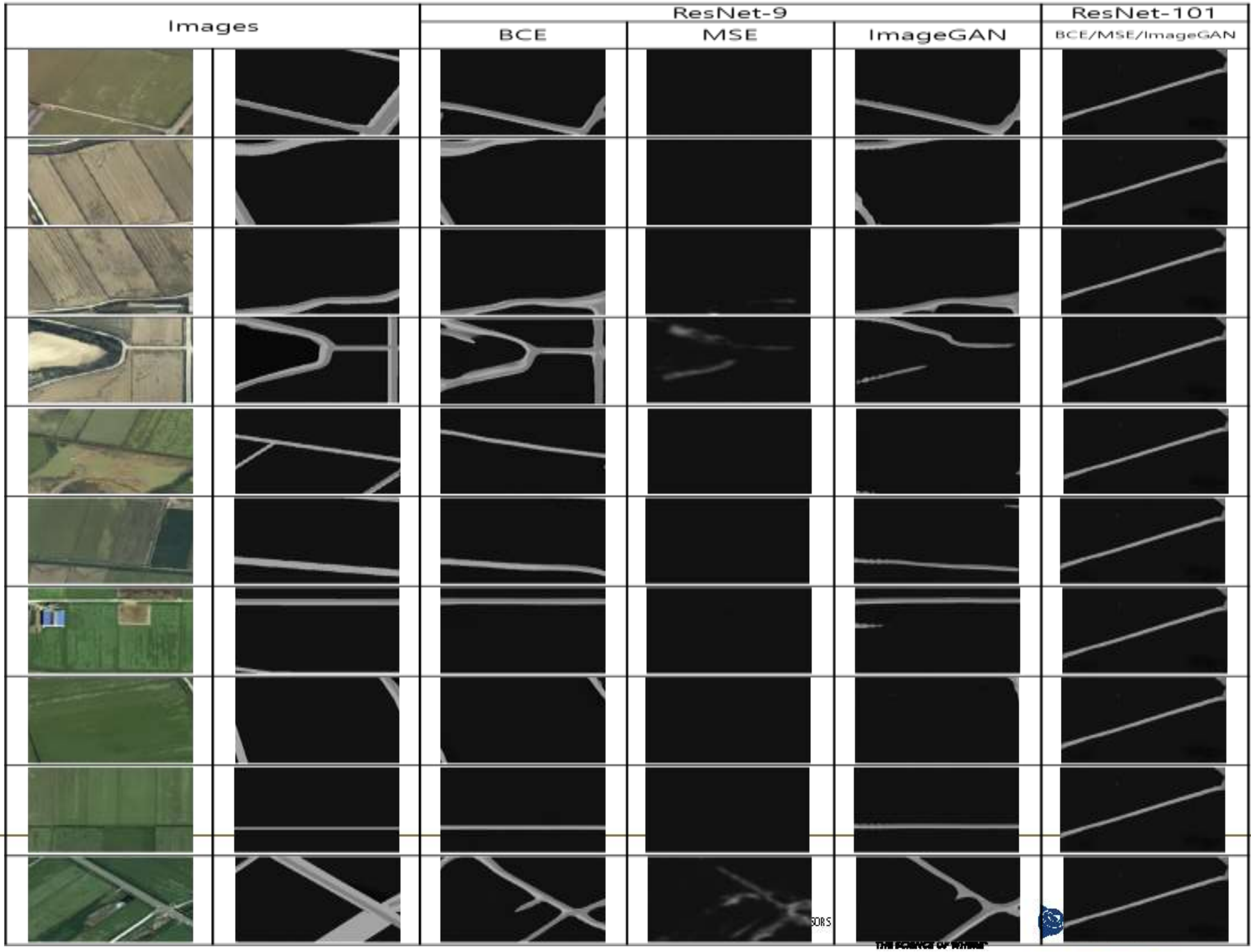
Results

- Experimental results show that ResNet-9 generator can capture common features of aerial images while ResNet-101 falls into mode collapse.
- ResNet-9 produces clearer road definitions.
  - expecting ResNet-101 with deeper network layers produces better results but it did not.

- Difficult to distinguish cadastral boundaries and common boundaries.

Results (this study)

- Right shows, 10 sample images. ResNet-9 with BCE(Binary Cross Entropy) loss is the best result.
- ResNet-9 with ImageGAN Discriminator has the second result because ImageGAN only judge the whole image distribution at once while PatchGAN check details with local distribution.
- Third is the MSE loss that uses a L2 loss. The exponential L2 value is more sensitive than the Euclidian L1 value.



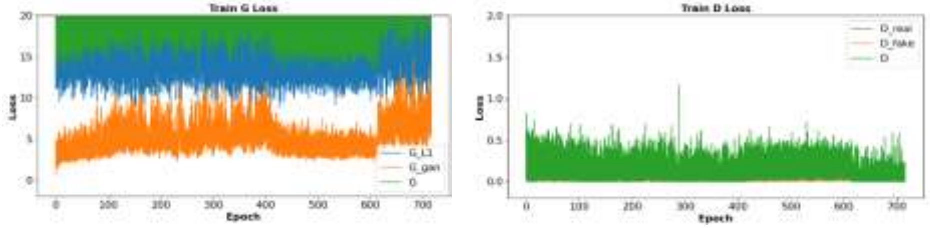
Results

- The evaluation uses the mean square error (MSE) per pixel between the generated image and the original cadastral map image.
- MSE is smaller, the better. The following table shows the mean square error (MSE) observed for two different generator networks. ResNet-9 with BCE has the lowest MSE, so we can assume that it is closest to the real thing.

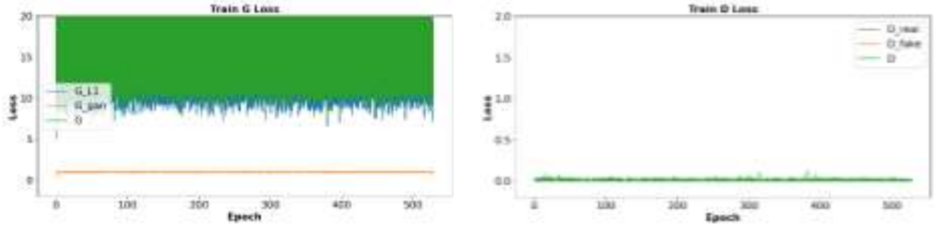
Image No	RenNet-9			ResNet-101
	BCE	MSE	ImageGAN	BCE/MSE/ImageGAN
Average	1263.855	1685.744	1352.507	2039.539
1	1343.072	2115.19	1433.135	2333.292
2	641.1027	1709.467	1143.234	1988.268
3	840.6652	1396.481	631.1481	1819.789
4	1546.08	2724.162	2503.745	3376.895
5	1974.294	1428.981	1599.517	1770.283
6	607.7222	1936.426	820.6284	1506.057
7	728.2318	1694.259	718.7419	2036.222
8	503.2178	697.509	1323.413	972.0905
9	1079.648	554.9541	634.3876	793.6936
10	503.0614	673.9441	681.1909	877.7407
11	4135.313	3611.812	3388.435	4960.598

Results

- It is difficult to train GAN well. Figure below shows the generator and discriminator loss of the ResNet-9 and Resnet-101
- Below shows the training loss of ResNet-9 with BCE.
  - G and D seem to converge well. Especially, G is stably adapted between epoch 70~80, and D also converges to 0.5 in epoch 70~80. Therefore, I use a checkpoint in epoch 70~80. After epoch 90, overfitting seems to have occurred.



- Below shows, the training loss of ResNet-9 with MSE, which did not trained at all.



Results-Summary

- The main results for the main goals are:
- Test the loss function on the cGAN model. (BCE vs MSE)
  - The BCE, which uses the L1 loss, is much better than the MSE.
- Test the generator ResNet-9 and ResNet-101.
  - The ResNet-9-based GAN model has much better performance in the result qualities and the visual analysis.
- Test discriminator PatchGAN and ImageGAN.
  - The PatchGAN discriminator is better than the ImageGAN discriminator.
- Performance of MSE evaluation indicators.
  - ResNet-9 with BCE model has the best value.

- (Optional-Dropout-Learning rate) Changes in small variables(parameters) do not significantly affect the results.

Conclusion

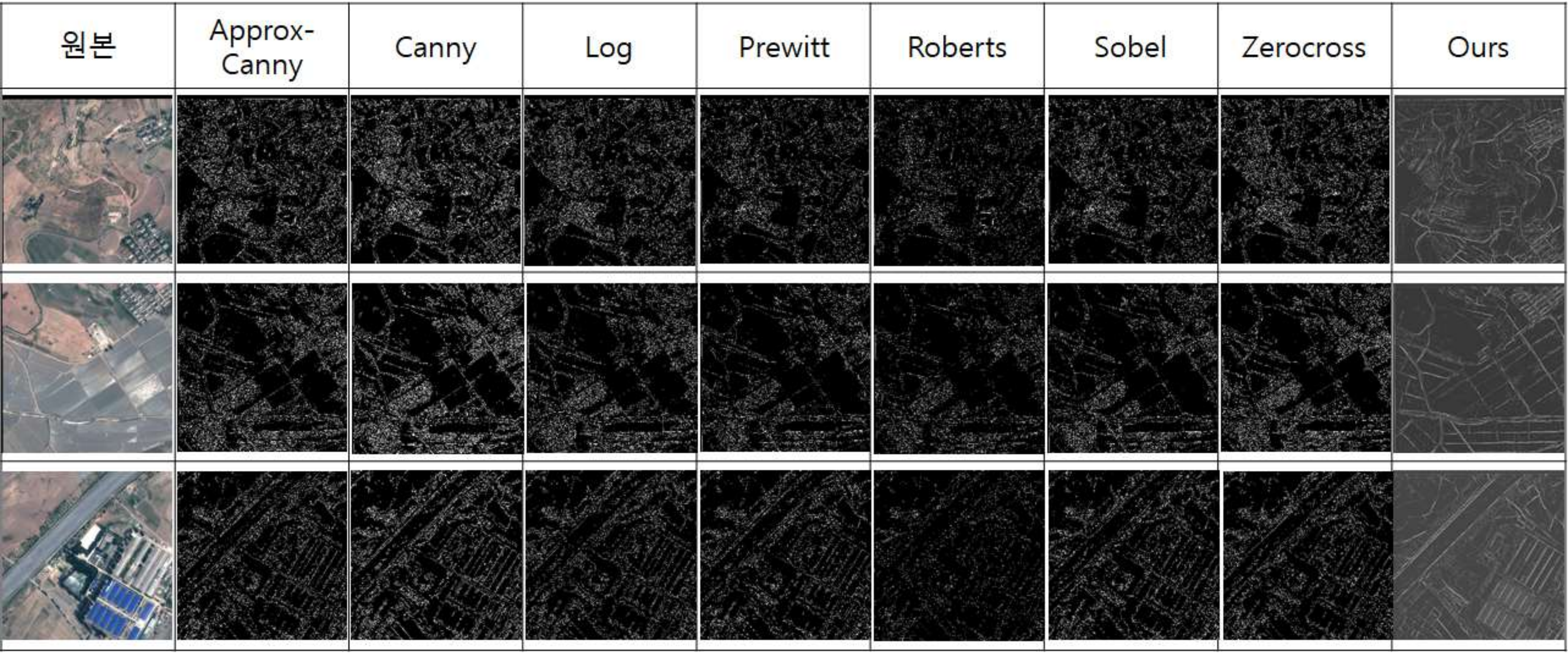
- cGAN is a promising approach for many image translation tasks from one visual domain to another(cadastre etc).
  - cGAN showed high performance in creating and manipulating image data.
  - cGAN can be used in various fields.

- Most important things are model and the dataset itself. To ensure better generalization of the network, refine and enriches the datasets.

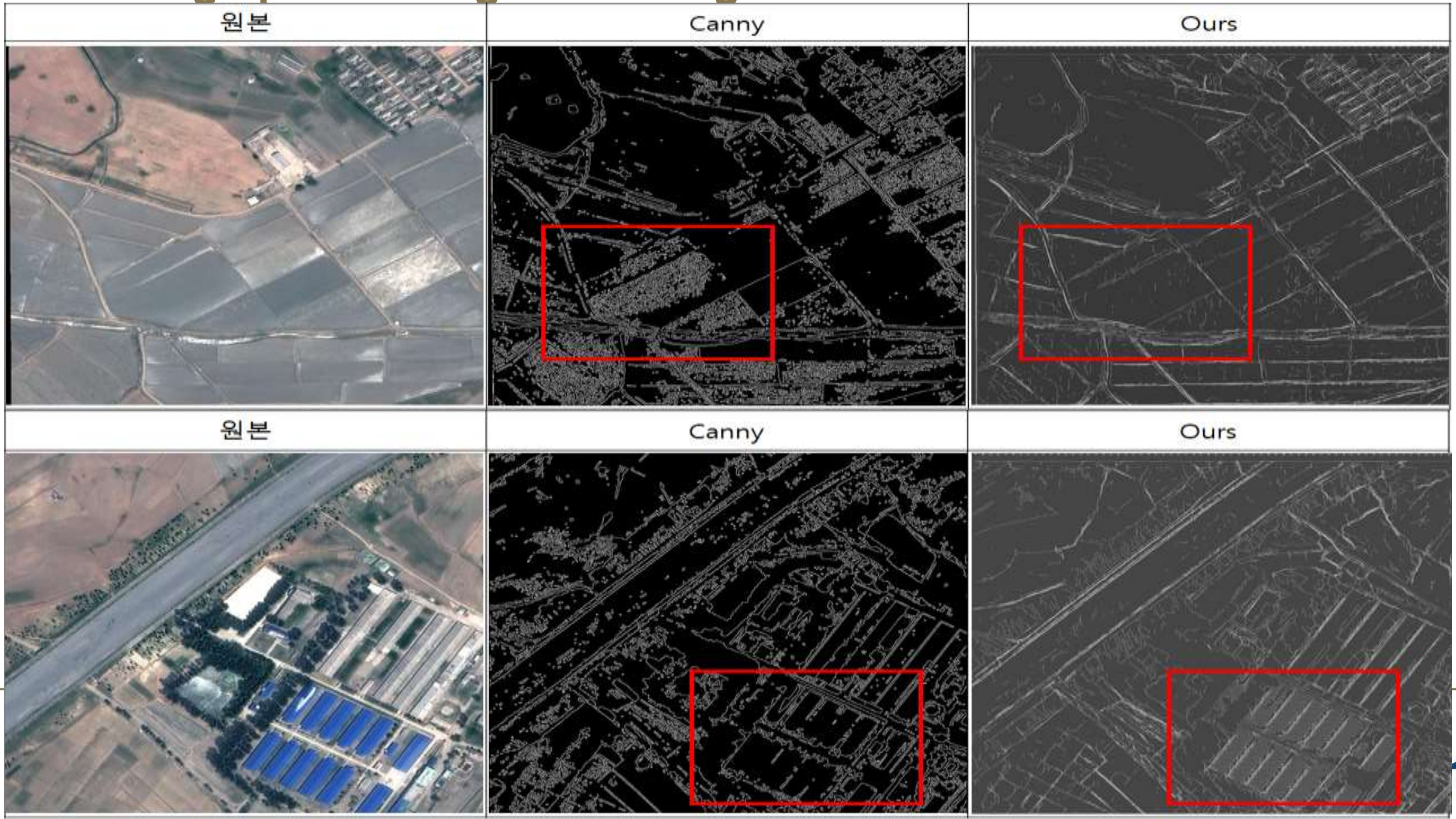
- In addition, the drone image datasets are the candidate to experiment.

Thank you

General results – image processing tech – edge detection



General results – image processing tech – edge detection



General results – image processing tech – edge detection by NN