Instance Segmentation of Water Body from Aerial Image using Mask Region-based Convolutional Neural Network

Sangdaow Noppitak,
Sarayut Gonwirat
and Olarik Surinta
Multi-agent Intelligent Simulation Laboratory,
Department of Information Technology,
Faculty of Informatics, Mahasarakham University, Thailand

Outline

- INTRODUCTION
- CONTRIBUTION
- MASK R-CNN ARCHITECTURE
  - Backbone Architecture
  - Head Architecture
- AERIAL IMAGE WATER RESOURCES DATASET
- EXPERIMENTAL RESULTS
  - Model evaluation
  - Result of instance segmentation of water body
- CONCLUSION

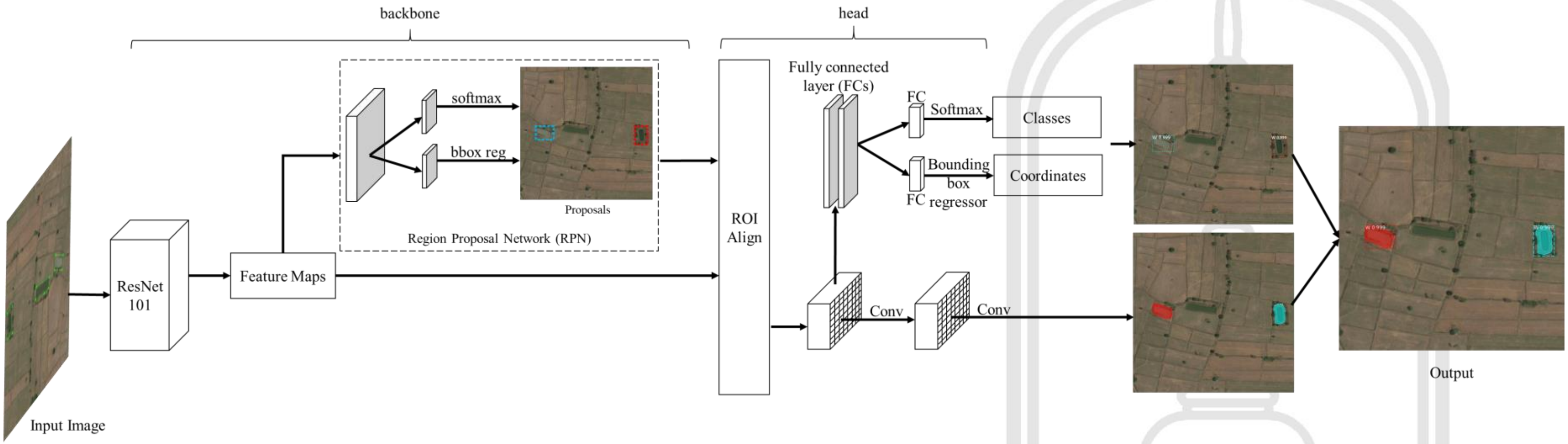
INTRODUCTION

- The two terms “land cover” and “land use” are typically used together
- Over the past ten years the difference between land cover and land use has attracted many researchers prompted by a change in land cover to accommodate changes in land use.
- As such, if land use data are accurate and up-to-date, we can apply that information to many objectives, such as city planning, environmental audit or evaluation, and national policy

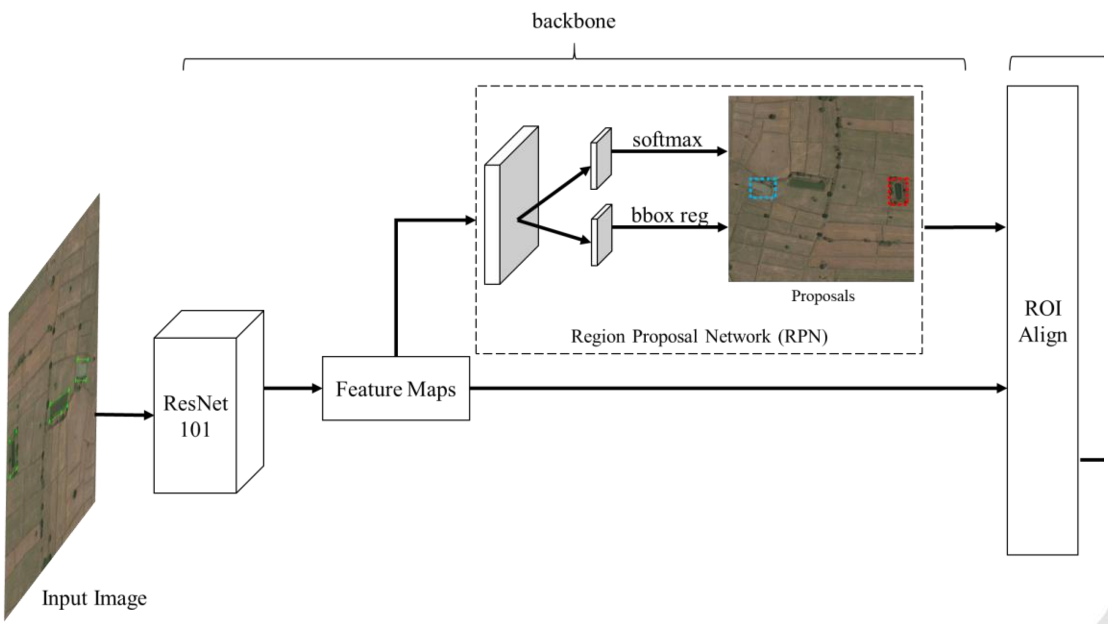
Contribution

- This article presents mask region-based convolutional neural network (Mask R-CNN) for water body segmentation from aerial images.
- Mask R-CNN architecture was tested with aerial image water resources dataset (AIWR).
- The AIWR is the images of agricultural areas in the northeast region of Thailand
- Water body data were collected from 2 types, natural water bodies (W1) and artificial water bodies (W2).
- The aerial images of water bodies were different in color, shape, size, and similarity.
- This dataset includes 800 images, so AIWR dataset challenges the instance segmentation process.

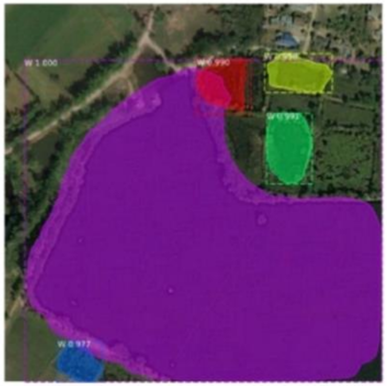
MASK R-CNN ARCHITECTURE



- Backbone Architecture



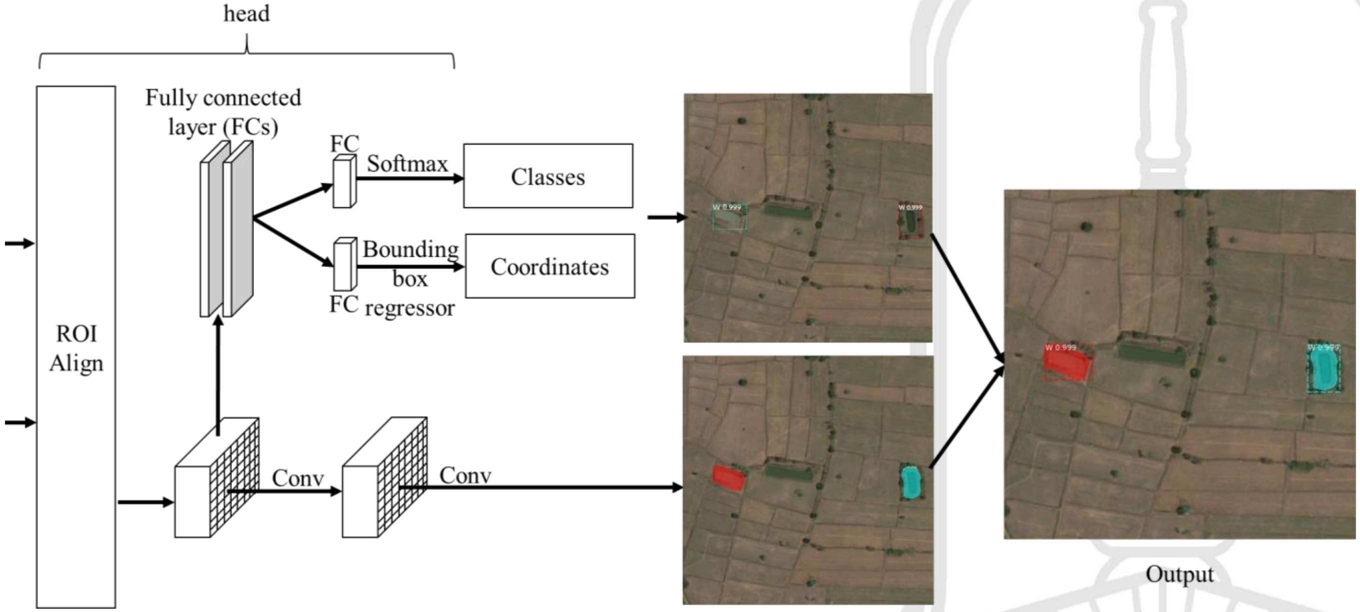
Faster R-CNN method



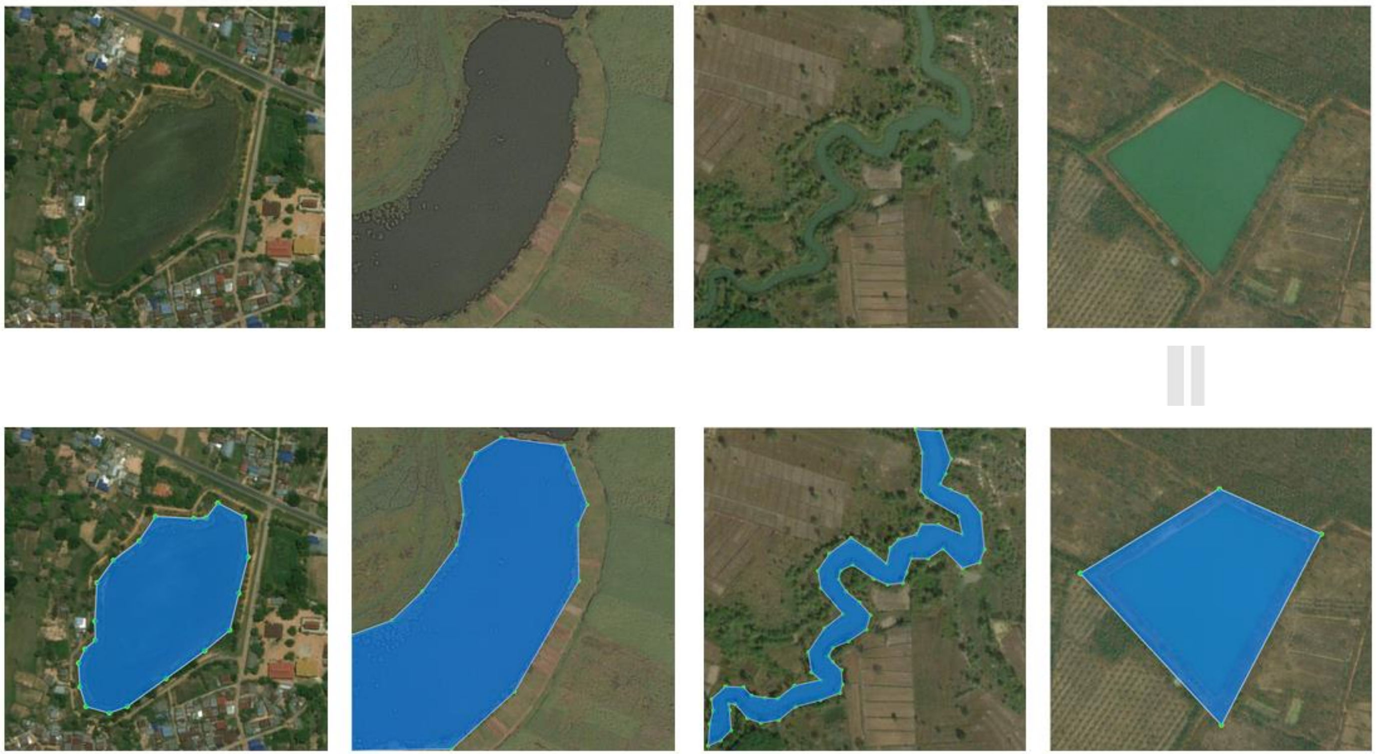
Mask R-CNN

ResNet-101 was used as backbone architecture.

- Head Architecture



AERIAL IMAGE WATER RESOURCES DATASET (AIWR)



aerial images

ground truth

- Aerial image water resources dataset, AIWR : 800 images

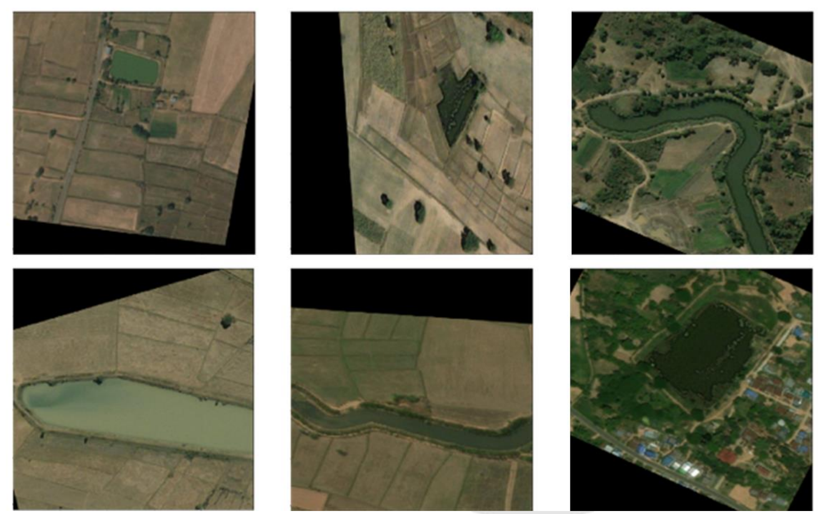
Train : 720	Test : 80

- There are 4 challenging objectives:
  - color,
  - shape,
  - size,
  - similarity

Color shape size similarity

EXPERIMENTAL RESULTS

- Those images were divided by the 10-fold cross-validation method.
- All aerial images were resized to 512x512 pixel.
- The parameters used for Fine-tune, for example
  - NUM_CLASSES=2
  - IMAGES_PER_GPU = 1
  - LEARNING_RATE=0.0001
- GPU GeForce GTX 1070 Ti, Intel(R) Core-i5, 7400CPU @ 3.00GHz, 8GB RAM, Linux Operating system.
- Data augmentation, which includes
  - Scale
  - translate_percent
  - Rotate
  - shear



- Model evaluation

To Evaluation Mask R-CNN algorithm, we use mean average precision (mAP)

$$mAP = \frac{1}{N}\sum_{i=1}^{N} \mathrm{AP_i}$$

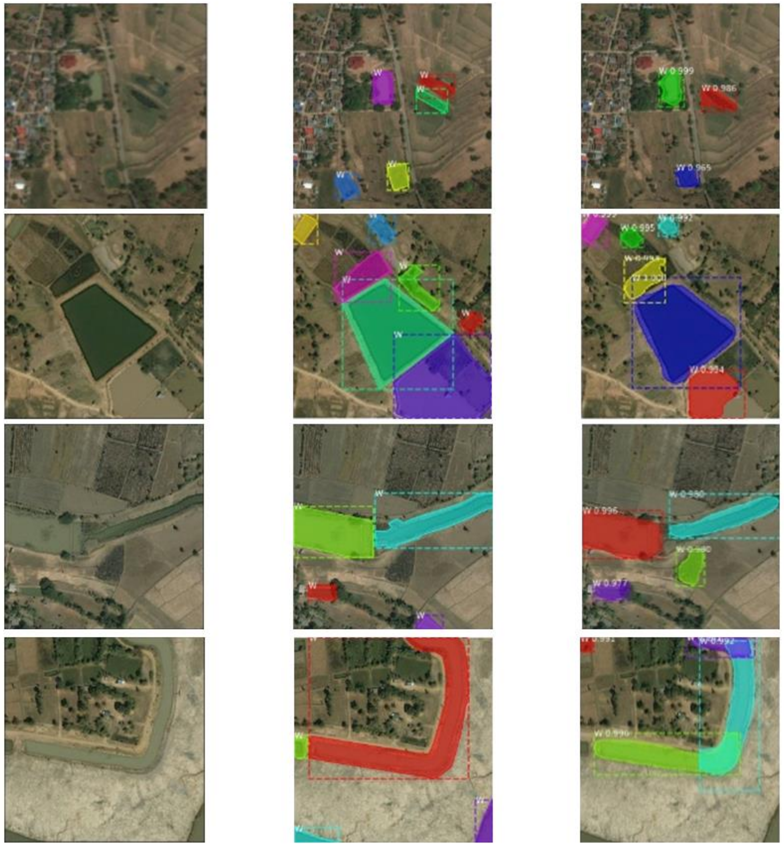
Result of instance segmentation of water body

The result of the experiment using mask R-CNN with the AIWR Dataset.

Augment	Validation loss	mAP	Training Time	Test Time /img
False	1.08	0.30	11d 15h 16min 27s	3 µs
True	0.41	0.59	12d 9h 48min 25s	4 µs

Result of instance segmentation using mask R-CNN with data augmentation.

A) Aerial images
B) images with ground truth
C) instance segmentation



a) b) c)

a) b) c)

Error results in segmentation.

A) Aerial images

B) images with ground truth and

C) error of instance segmentation.

CONCLUSION

- This research tested with aerial images of water resources dataset (AIWR).

- This research tested the accuracy of instance segmentation by Mask R-CNN together with data augmentation.

- This research has shown that the mask R-CNN architecture combined with data augmentation can identify the water surface using the mAP value for measurement. The mAP value was up to 0.59.

- In future work, because the data tested is aerial photography obtained from Bing map, only RGB colors can be tested.

- If other research can use data from satellites, such as Landsat, which has a band specifically for water analysis, the result of an analysis of water bodies with different color might give higher accuracy.

- Any new architecture suitable for water body analysis might be used to expect an even higher accuracy rate.

THANK YOU FOR YOUR ATTENTION