UDT
Undersea Defence Technology
13-15 May 2019
Stockholmsmässan, Sweden

Deep Learning based tonal detection for passive sonar signals

Dae-Jin Jung, Jihun Park, Sang Ho Lee and Taekyu Reu

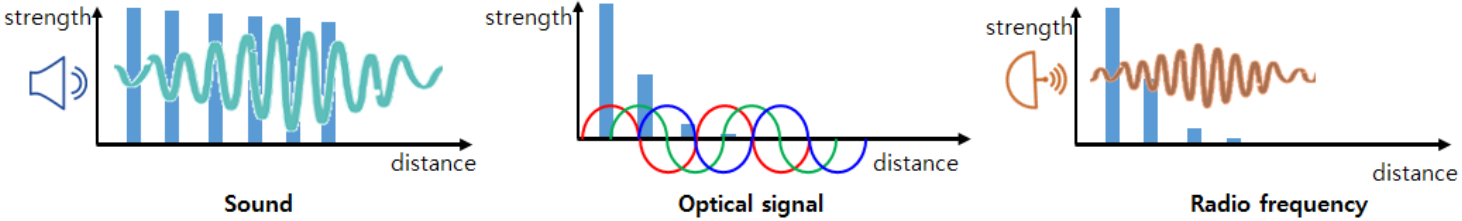
Intelligence & Information Technology Center,
Agency for Defense Development, South Korea

#UDT2019

Agency for Defense Development

Introduction

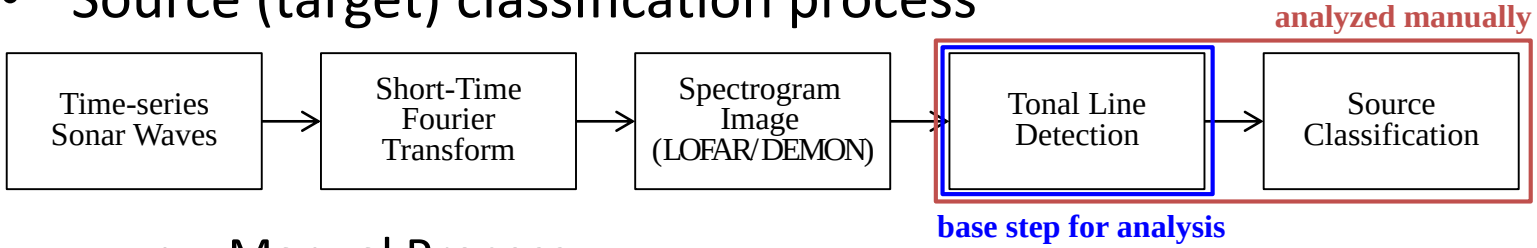
- Underwater target detection
  - SONAR
    - Acoustic signals : best propagation under the water
      (Others have severe attenuation under the water)



    - Sole technique for underwater target detection
    - Most military systems use **passive SONAR**
    - SONAR signal analysis : LOFAR / DEMON

Introduction

- Source (target) classification process

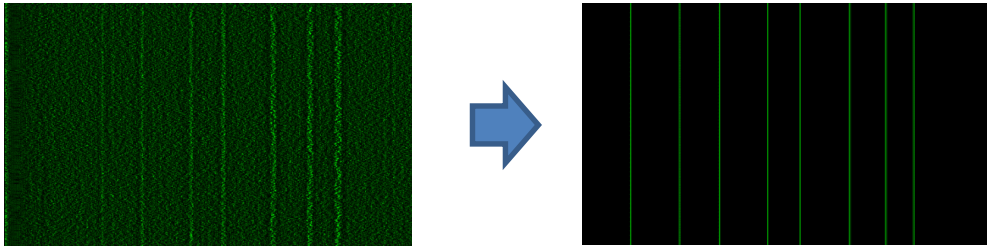


  - Manual Process
    - Dependent on personal ability
    - Time-consuming (analysis / training analysts)
    - Likely to make a few mistakes

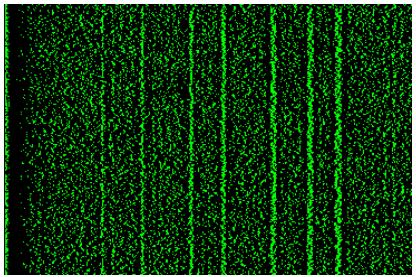
- Automated method is required for accurate/fast identification
  - Adopt deep learning to this problem

Introduction

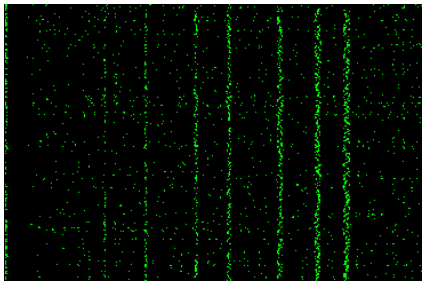
- Tonal line detection examples
  - Desired



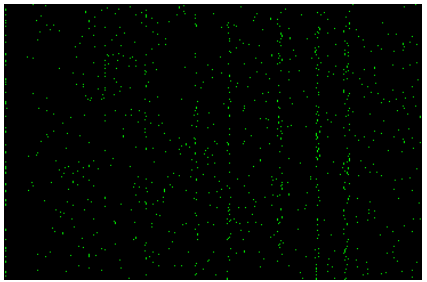
  - Conventional methods



< thresholding >

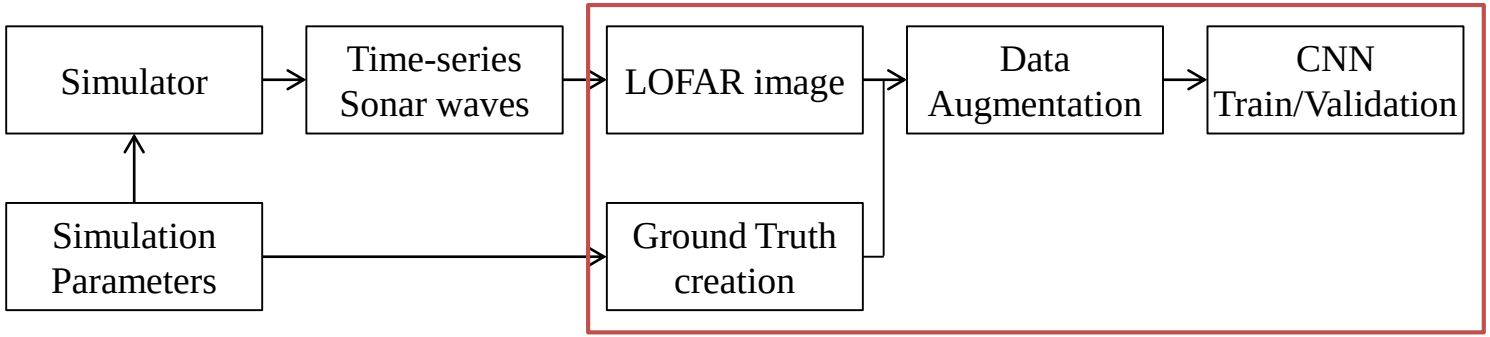


< adoptive thresholding – Otsu method >

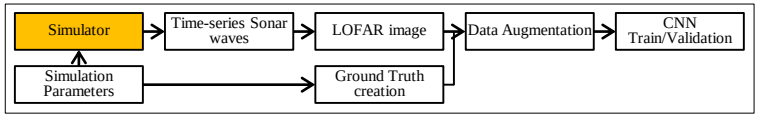


< peak detection >

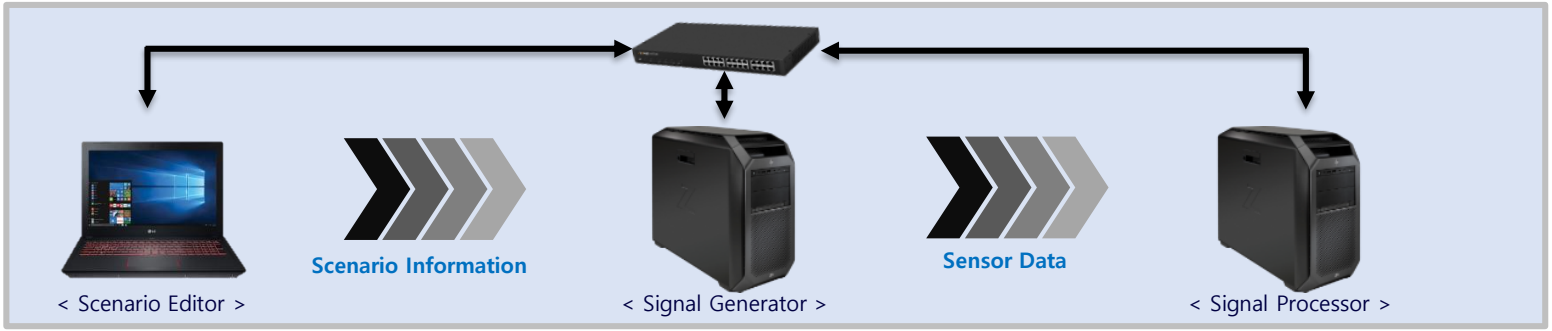
Proposed Method

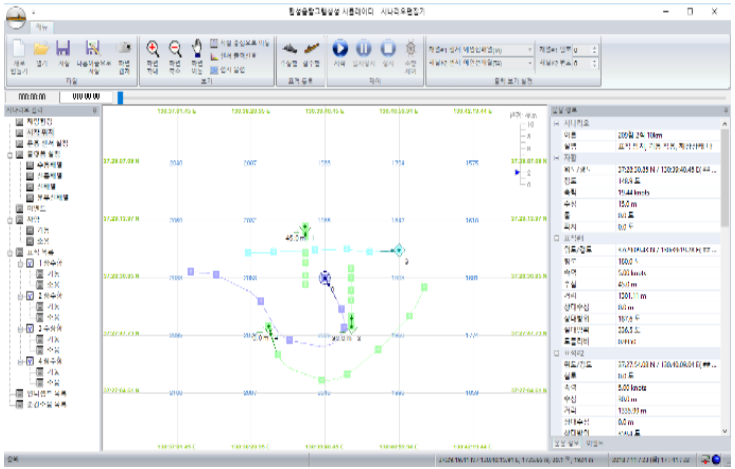


- Use of CNN (Convolutional Neural Network)
  - Great performance in pattern recognition
- Use of simulator
  - Difficulties in access to real SONAR data
  - Easiness of creating training data

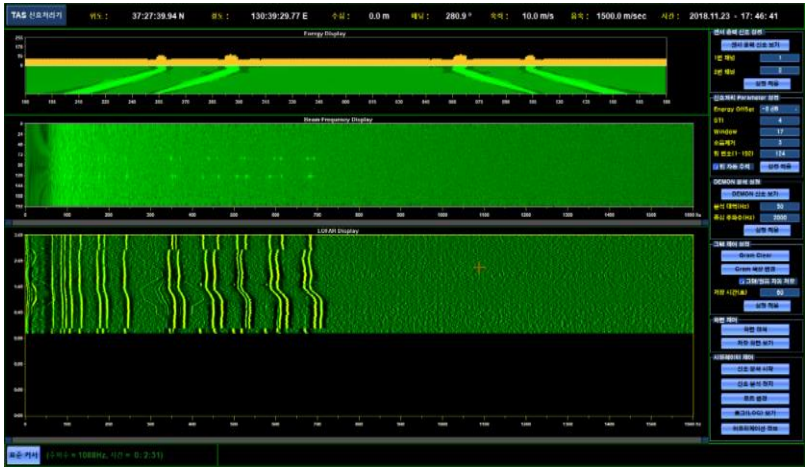


Proposed Method - Simulator

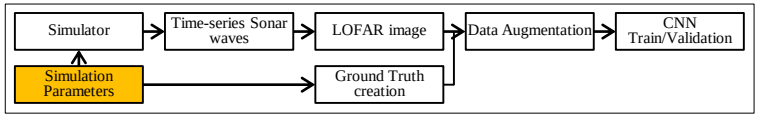




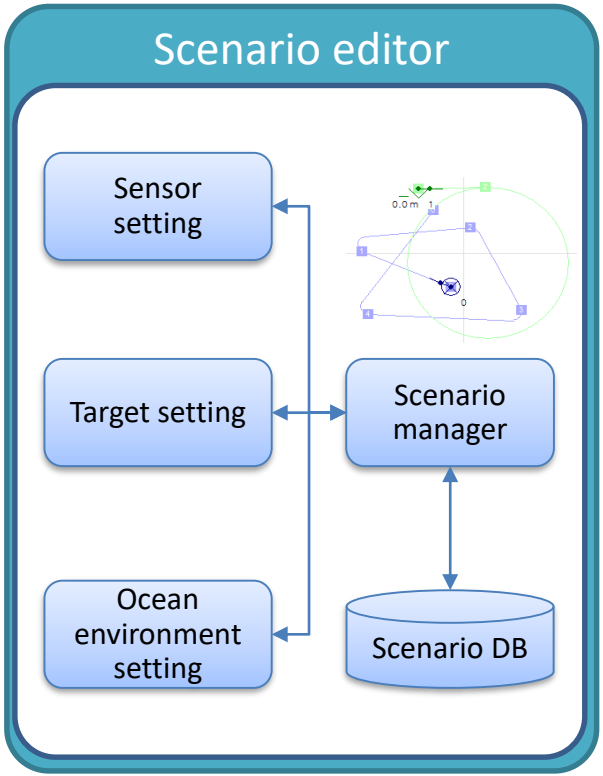
< Scenario Editor >

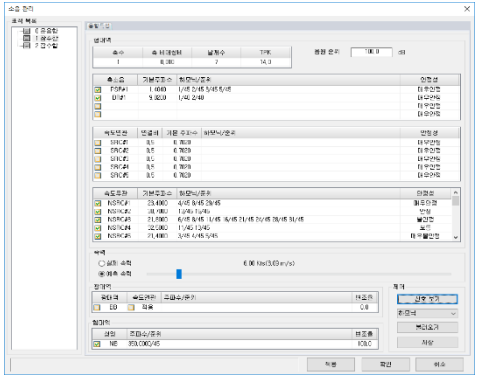


< Signal Processor >

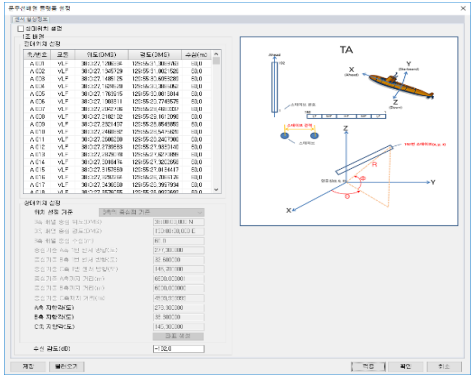


Proposed Method - Simulator Parameters

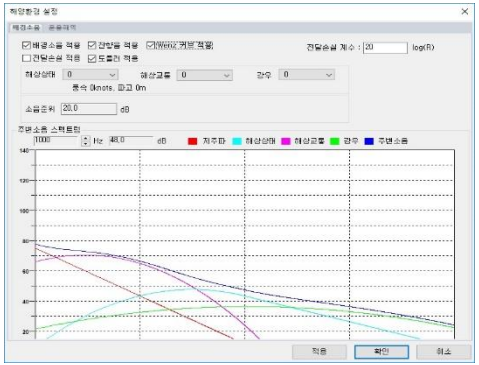


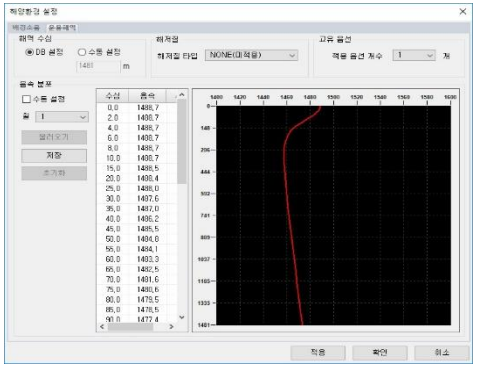


< Target settings >

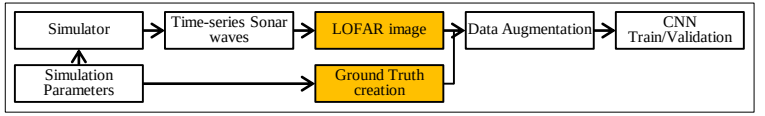


< Sensor settings >



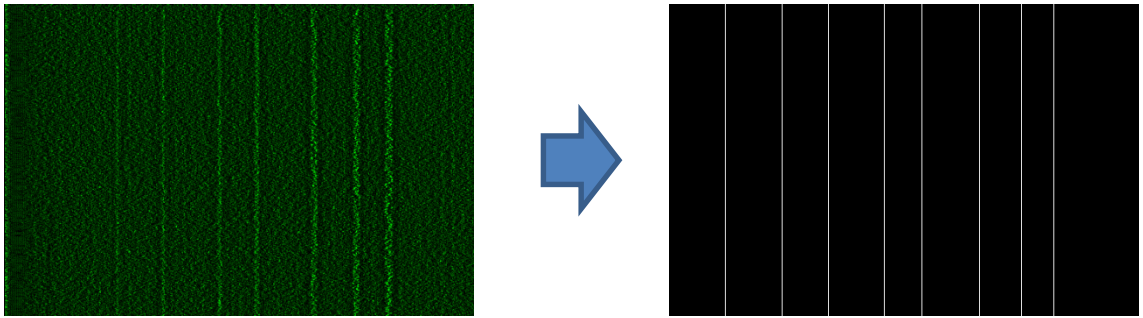


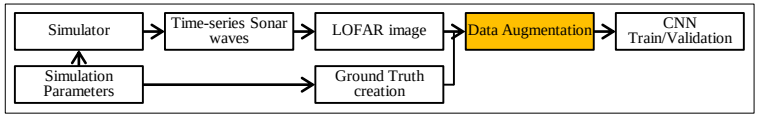
< Ocean environment settings >



Proposed Method - LOFAR image/Ground Truth

- LOFAR image
  - Single target scenario (target speed : 0)
  - S3PM : Normalization (window size : 17 / gap size : 3)
  - Frequency resolution : 0.5 Hz
  - Short-time integration : 2 seconds
- Ground Truth creation using scenario data

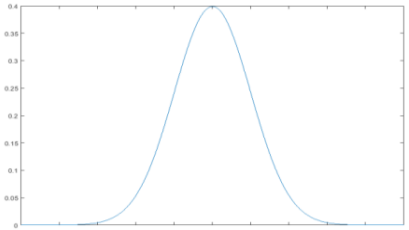




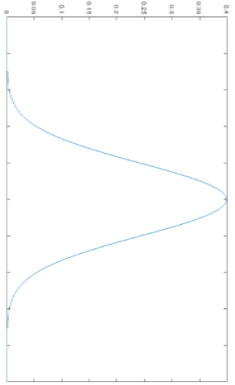
Proposed Method - Data augmentation

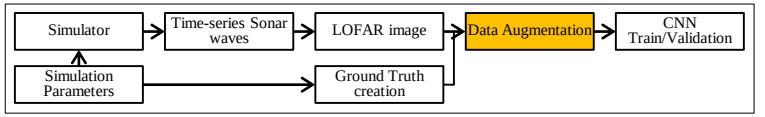
- Various data creation
- Target motion simulation using image manipulation
  (Image row (frequency) shifting)
  - No consideration for harmonics
  - Gaussian probability density function

$$g(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$



rotate



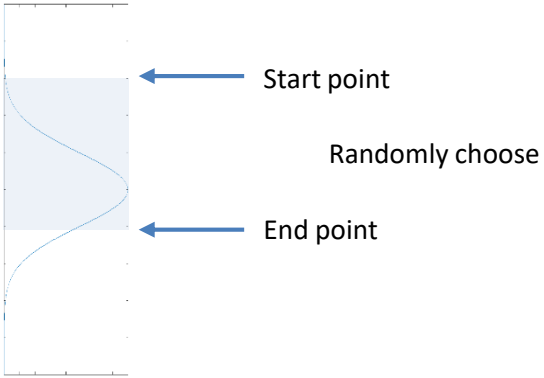


Proposed Method - Data augmentation

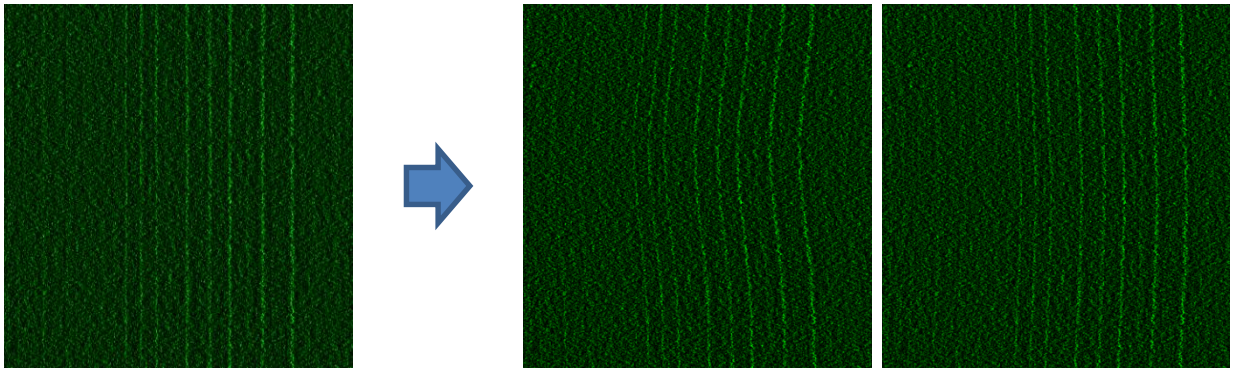
- Data augmentation
  - Frequency-shift function

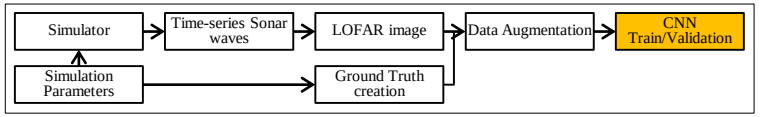
$$f(x) = d * m * g(x) + \mathrm{ms}$$

    - d ∈{-1, 1} : directivity
    - m=[20, 40] : magnitude
    - ms=[-10, 10] : magnitude-shift
    - Random selection of start-point



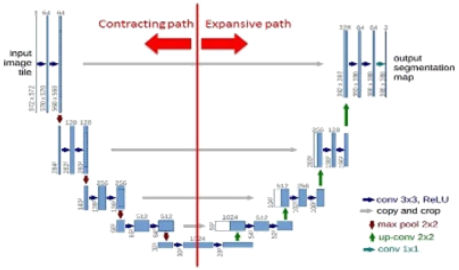
  - Example



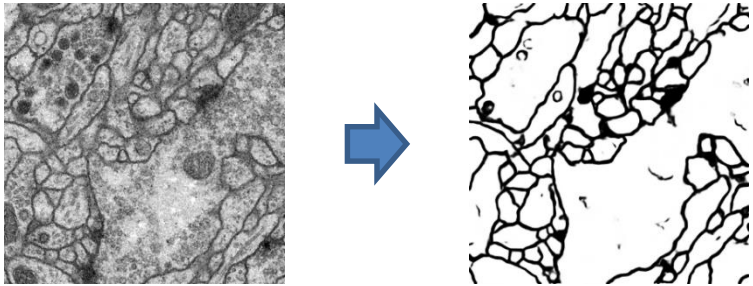


Proposed Method - CNN Train/Validation

- CNN (Convolutional Neural Network)
  - U-Net
    - Fully convolutional network for semantic segmentation
    - Tonal line detection can be considered as semantic segmentation



< prediction result >



< U-net prediction example >

    - Cons
      - Imbalanced training sets
        (# of tonal line pixels << # ambient noise pixels)
      - Dependency on image size (should be multiple of 16)

Test results - settings

- 312 scenario data

	Model 1(U-Net)
Scenario # (10 min/scenario)	312
Data augmentation	X20 (6,240 pairs)
Train/Validation Data #	2,640 / 3,600 images

- Workstation specification
  - 4 GPUs (NVidia Titan XP – 12GB)
  - 128 GB RAM

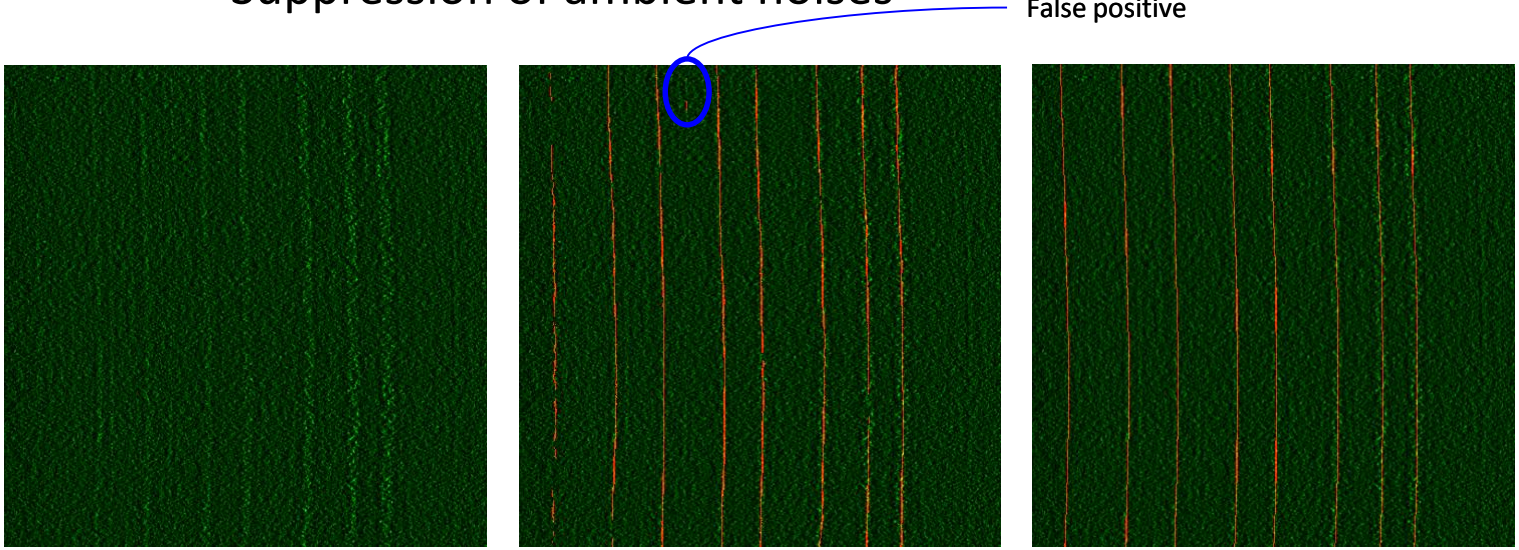
Test results - results

- Trained model test results
  - Detection result criteria
    - Tolerate 1 pixel (0.5Hz) displacement

	Train data	Validation data
Precision	0.9959	0.9618
Recall	0.9045	0.9206
Prediction time (sec)	0.3217 (10-minute LOFAR image)	

Test results - results

- Example
  - Tonal line detection
  - Suppression of ambient noises



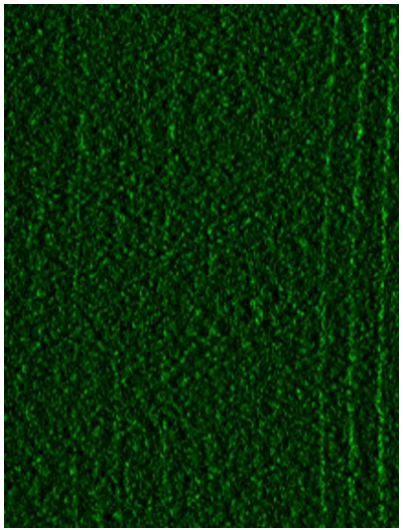
< LOFAR image >

< prediction result >

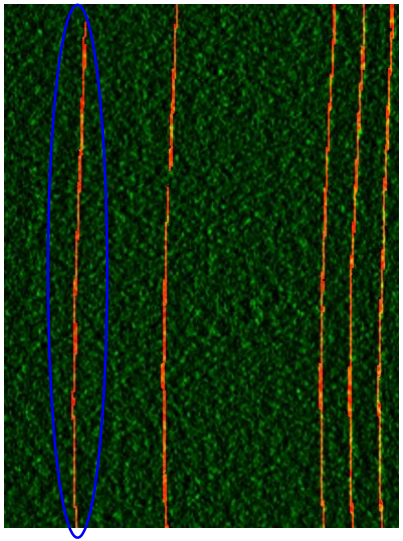
< ground truth >

Test results - results

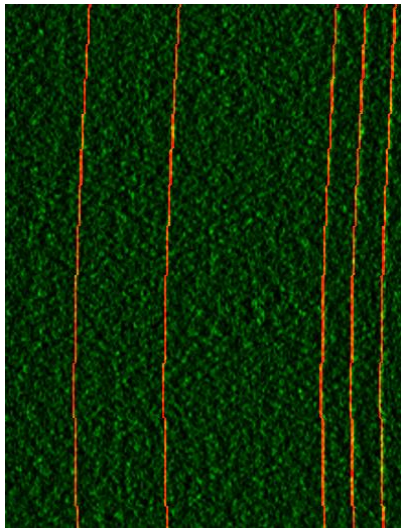
- Example
  - Tonal lines hardly detected by human eyes



< LOFAR image >



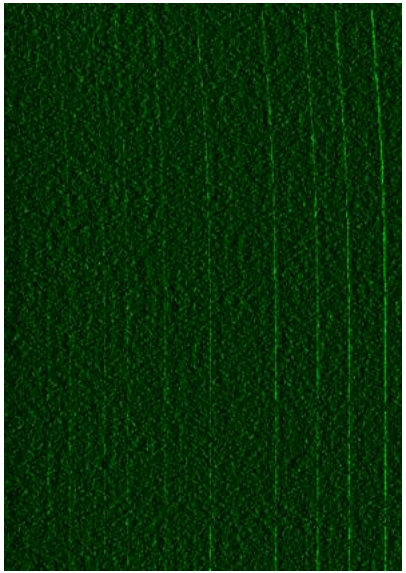
< prediction result >



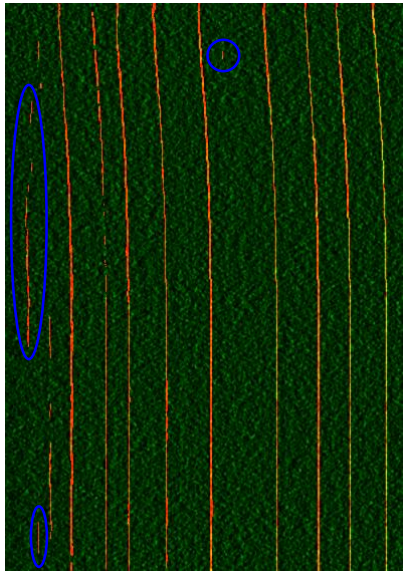
< ground truth >

Test results - results

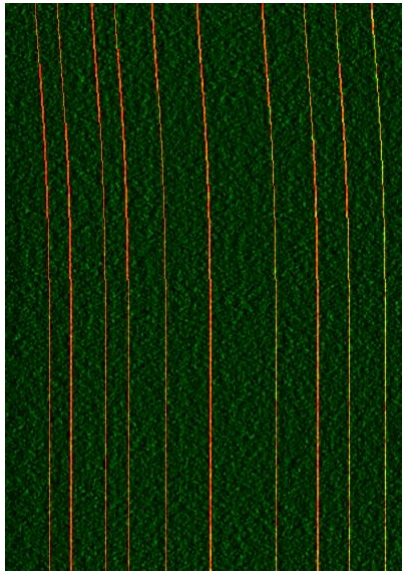
- Example
  - False positivies



< LOFARgram >



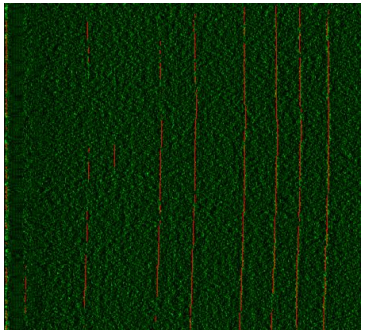
< prediction result >



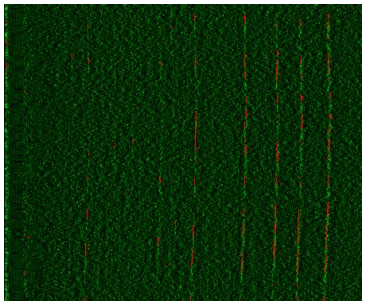
< ground truth >

Test results - results

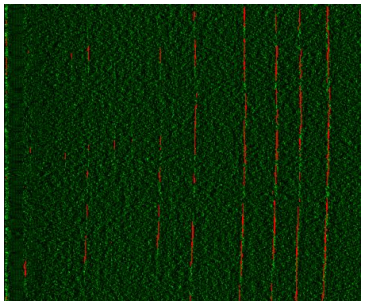
- Results in different time-unit executions (Model - 1)
  - 480 sec example (using Sliding window)



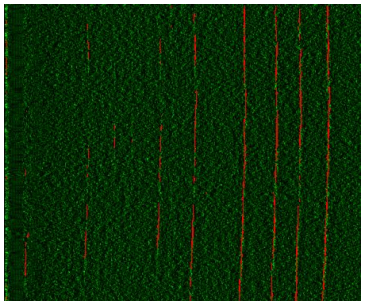
< Time Unit : 480 sec >



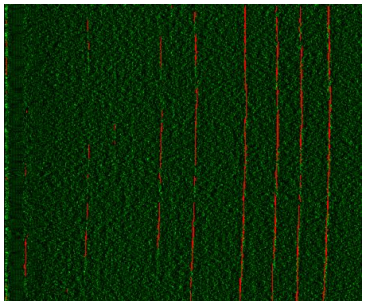
< Time Unit : 16 sec >



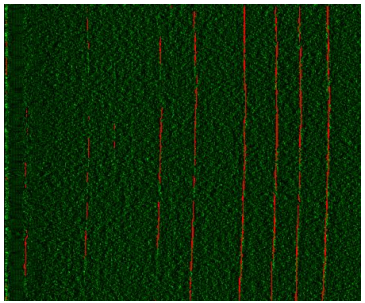
< Time Unit : 32 sec >



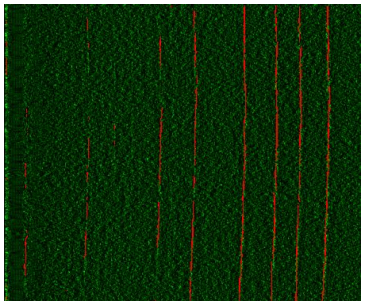
< Time Unit : 48 sec >



< Time Unit : 64 sec >



< Time Unit : 80 sec >



< Time Unit : 96 sec >

Conclusions

- Automated sonar tonal detection
- Good performance on synthetic simulation data
  - Image based training
  - Connection between disconnected lines
  - Accurate, Speedy
- Future work
  - Trying various CNN architectures
  - Extraction of various information for ship classification
  - Validation task on real data