Multi-task deep learning from Sentinel-1 SAR: ship detection, classification and length estimation

C. Dechesne [1]

S. Lefèvre [2], R. Vadaine [3], G. Hajduch [3], R. Fablet [1]

[1] *IMT Atlantique – Lab-STICC, UMR CNRS 6285, Brest, FR*
[2] *Univ. Bretagne Sud – IRISA, UMR CNRS 6074, Vannes, FR*
[3] *Collecte Localisation Satellites, Brest, FR*

BiDS 2019
19–21 February 2019

Outline

Context

The detection of inshore and offshore ships is an important issue:

- Monitoring fisheries,
- Managing maritime traffic,
- Ensuring safety of coast and sea.

In operational contexts, ship detection is traditionally performed by a human observer from **visual analysis** on remotely-sensed images. It is very **time consuming** and cannot be conducted at a very **large scale**.

SAR images

Sentinel-1 SAR data provides regular, worldwide coverage.

Almost all coastal zones and shipping routes are covered by Interferometric Wide Swath Mode (IW), while Extra-Wide Swath Mode (EW) acquires data over open oceans, providing a **global coverage** for sea-oriented applications.

AIS

The automatic identification system (AIS) is an automatic tracking system that uses transponders on ships.

AIS provides meaningful and relevant information about ships (such as position, **type**, **length**, rate of turn, speed over ground, etc.)

Deep learning

Deep learning is considered as one of the major breakthrough related to big data and computer vision.

A Deep Neural Network:

- consists of multiple layers (such as convolution, pooling, fully connected and normalization layers),
- transforms original data (raw input) into higher semantics representation,
- can learn very complex functions.

Objectives

Employ the synergy of SAR and AIS in order to detect and characterize ships:

- Detect ships on SAR images.
- **Estimate ship length from SAR images.**
  - Could not be directly retrieved from ship footprint from SAR images.
- **Classify ships from SAR images.**

Context
Dataset
Framework
Results
Conclusion

Dataset

The dataset is composed 18,894 SAR images of size 400×400 obtained by coupling AIS information and SAR data:

- Each image is accompanied with the incidence angle

→ cropping to reduce the size to 80×80 and preserve significant contextual information while reducing the amount of data to process.

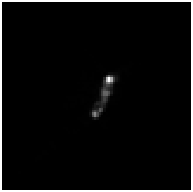
- 5 classes (Tanker - Cargo - Fishing - Passenger - Tug)
- Unbalanced database (10,430 instances of Tanker and only 1,071 instances of Passenger)

→ data augmentation with translations and rotations.

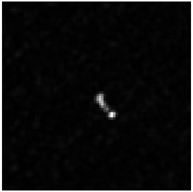
Final dataset:

- **Balanced dataset, 20,000 images 80×80.**

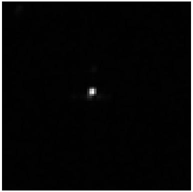
Dataset



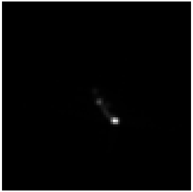
Tanker



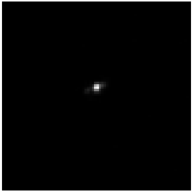
Cargo



Fishing



Passenger



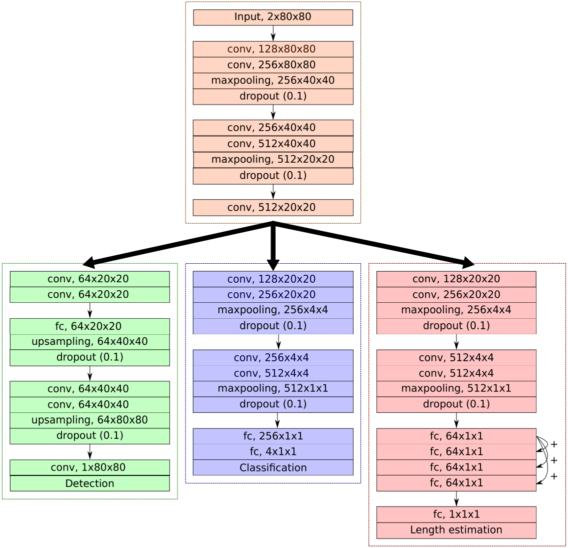
Tug

Framework

The proposed multi-task framework is based on two stages inspired from the VGG framework:

- a first common part,
- three task-oriented branches for ship detection, classification and length estimation.

Model


Input, 2x80x80
conv, 128x80x80
conv, 256x80x80
maxpooling, 256x40x40
dropout (0.1)
conv, 256x40x40
conv, 512x40x40
maxpooling, 512x20x20
dropout (0.1)
conv, 512x20x20
conv, 64x20x20
conv, 64x20x20
fc, 64x20x20
upsampling, 64x40x40
dropout (0.1)
conv, 64x40x40
conv, 64x40x40
upsampling, 64x80x80
dropout (0.1)
conv, 1x80x80
Detection
conv, 128x20x20
conv, 256x20x20
maxpooling, 256x4x4
dropout (0.1)
conv, 256x4x4
conv, 512x4x4
maxpooling, 512x1x1
dropout (0.1)
fc, 256x1x1
fc, 4x1x1
Classification
conv, 128x20x20
conv, 256x20x20
maxpooling, 256x4x4
dropout (0.1)
conv, 512x4x4
conv, 512x4x4
maxpooling, 512x1x1
dropout (0.1)
fc, 64x1x1
fc, 64x1x1
fc, 64x1x1
fc, 64x1x1
+
+
+
fc, 1x1x1
Length estimation

Loss function - detection

Detection loss (binary cross-entropy):

$$L_{det} = -\frac{1}{N}\sum_{n=1}^{N}\sum_{k\in I}(y_k \log(p(k)) + (1-y_k)\log(1-p(k)), \quad (1)$$

with y_k: ground truth of ship presence (0 or 1) of pixel k and $p(k)$ probability of ship presence of pixel k.

Allows to provide a probability map of ship presence.

Loss function - classification

Classification loss (categorical cross-entropy):

$$L_{class} = -\frac{1}{N} \sum_{n=1}^{N} \sum_{c=1}^{n_c} (y_{o,c} \log(p_{o,c})), \qquad (2)$$

with $y_{o,c}$ a binary indicator (0 or 1) if class label c is the correct classification for observation o and $p_{o,c}$ the probability for the observation o to belong to c.

Allows to provide the class probability for each input image.

Context

Dataset

Framework

Results

Conclusion

Loss function - length estimation

Length estimation loss (mean squared error):

$$L_{length} = \frac{1}{N} \sum_{n=1}^{N} (l_{pred} - l_{true})^2, \quad (3)$$

with l_{pred} the predicted length and l_{true} the true length.

Allows to minimize the error of ship length.

Context

Dataset

Framework

Results

Conclusion

Loss function

The loss function for the network is:

$$\mathsf{L} = \mathsf{L}_{det} + \mathsf{L}_{class} + \mathsf{L}_{length} \tag{4}$$

- Each specific loss employed to design the loss of the whole network could have been weighted.

$\rightarrow$ The range is not uniform among the different losses but it appears to have no effect on the optimization process.

- Each specific loss is equally weighted.

Context
Dataset
Framework
Results
Conclusion

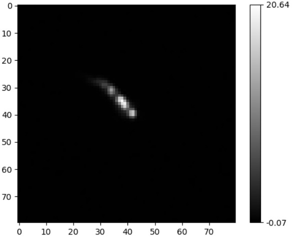
Training

Network trained on a PC with a single NVIDIA GTX 1080 Ti, an Intel Xeon W-2145 CPU 3.70GHz and 64GB RAM.

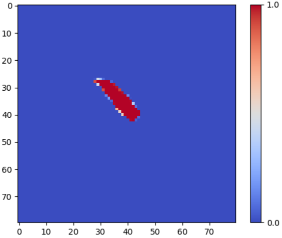
- Dataset splitted:
  - 16,000 samples for training,
  - 4,000 samples for testing.
- Training the model with 800 epochs and batches of 100 samples.
- Training takes about 9 hours.
- Testing takes less than a minute.

Detection results

In terms of detection, only a visual assessment has been performed.



SAR image



Detection

Context

Dataset

Framework

Results

Detection

Classification

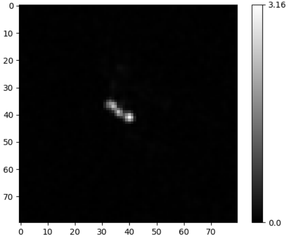
Length estimation

Comparison

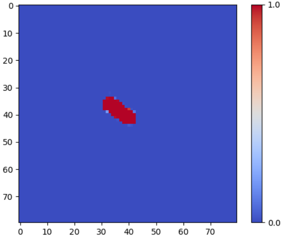
Conclusion

Detection results

In terms of detection, only a visual assessment has been performed.



SAR image



Detection

Classification results - 4 classes

Confusion matrix					
Label	Tanker	Cargo	Fishing	Passenger	Precision
Tanker	985	11	0	4	98.5
Cargo	65	907	12	16	90.7
Fishing	0	2	998	0	99.8
Passenger	0	0	0	1000	100.0
Recall	93.81	98.59	98.81	98.04	

Accuracy metrics					
Label	Tanker	Cargo	Fishing	Passenger	Overall
IoU	92.49	89.54	98.62	98.04	94.67
F-Score	96.1	94.48	99.3	99.01	97.22
Accuracy	98.0	97.35	99.65	99.5	97.25
κ	0.95	0.93	0.99	0.99	0.97

We only report some light confusion for the *Cargo* class.

Classification results - 5 classes

Confusion matrix						
Label	Tanker	Cargo	Fishing	Passenger	Tug	Precision
Tanker	771.0	28.0	0.0	1.0	0.0	96.38
Cargo	60.0	732.0	3.0	3.0	2.0	91.5
Fishing	0.0	1.0	799.0	0.0	0.0	99.88
Passenger	3.0	1.0	0.0	796.0	0.0	99.5
Tug	0.0	0.0	0.0	0.0	800.0	100.0
Recall	92.45	96.06	99.63	99.5	99.75	

Accuracy metrics						
Label	Tanker	Cargo	Fishing	Passenger	Tug	Overall
IoU	89.34	88.19	99.5	99.0	99.75	95.16
F-Score	94.37	93.73	99.75	99.5	99.88	97.44
Accuracy	97.7	97.55	99.9	99.8	99.95	97.45
κ	0.93	0.92	1.0	0.99	1.0	0.97

We only report some light confusion for the *Cargo* class.

Length estimation results

The length estimation is well performed by our network:

- 4 classes: mean error: **4.65 m ± 8.55 m**
- 5 classes: mean error: **1.93 m ± 8.8 m**

The mean error is in both cases lower than the resolution of the image and the standard deviation is also very low.

Institut Mines-Télécom
Université Bretagne Sud
CLS COLLECTE LOCALISATION SATELLITES

Comparison with other networks

Length estimation - comparison with a MLP with one hidden layer having 128 hidden units.

	Our network	MLP
Mean error	**4.65 m ± 8.55 m**	-7.5 m ± 128 m

Classification - comparison with the MLP and a RCNN.

	Our network	MLP	RCNN
Overall accuracy(%)	**97.25**	25.00	88.57

Institut Mines-Télécom
Université Bretagne Sud
CLS
COLLECTE LOCALISATION SATELLITES

Conclusion

The ship detection has been widely investigated, a ship probability presence map has been proposed using a deep neural network.

The proposed framework has shown **very good results in terms of classification**. Some light confusion with the *Cargo* class was reported.

The length estimation achieves a **sub-pixel error and standard deviation**.

Multi-task deep learning from Sentinel-1 SAR: ship detection, classification and length estimation

C. Dechesne[1]

S. Lefèvre [2], R. Vadaine [3], G. Hajduch [3], R. Fablet [1]

[1] *IMT Atlantique – Lab-STICC, UMR CNRS 6285, Brest, FR*
[2] *Univ. Bretagne Sud – IRISA, UMR CNRS 6074, Vannes, FR*
[3] *Collecte Localisation Satellites, Brest, FR*

Questions?

BiDS 2019
19–21 February 2019