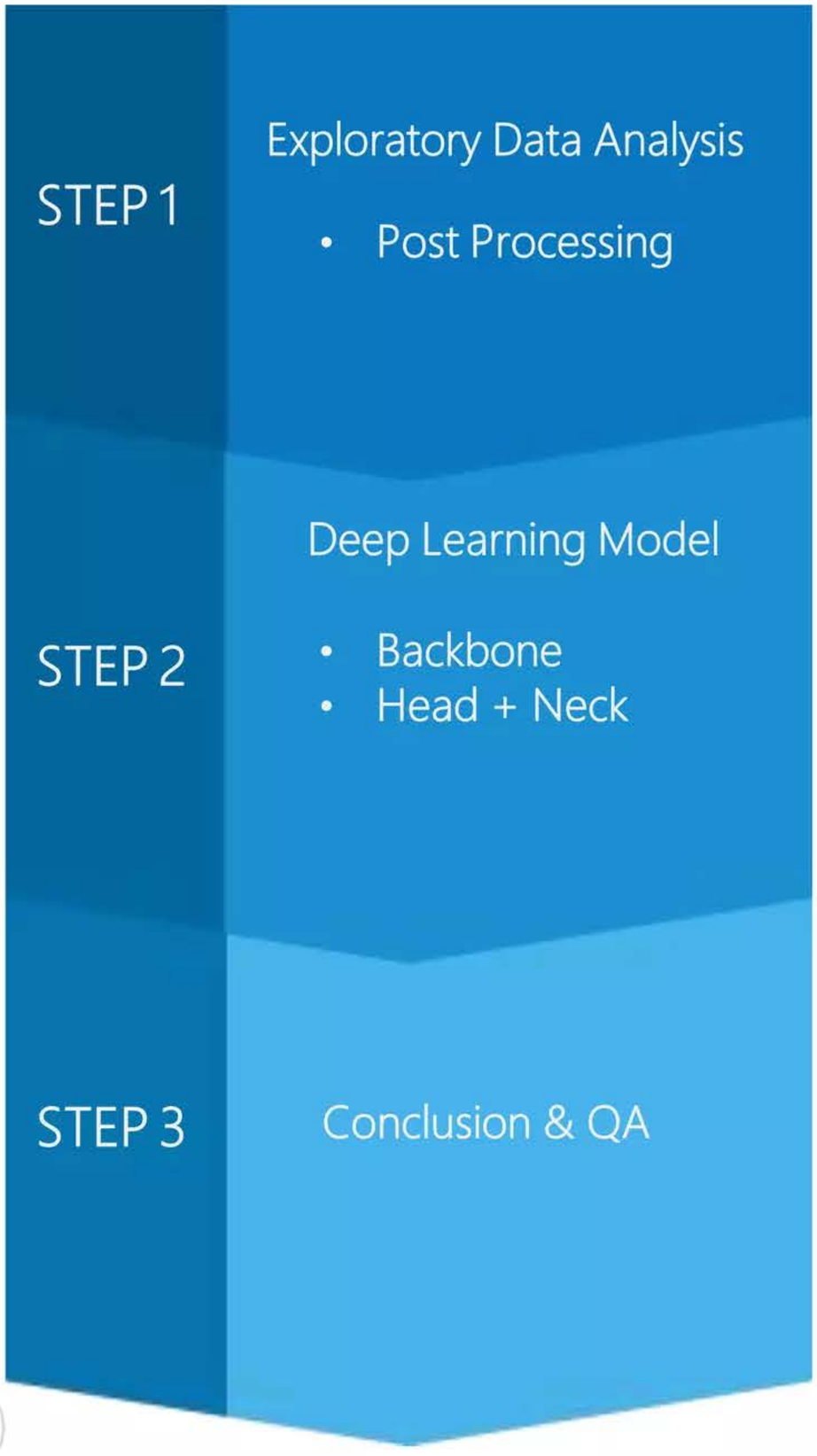
DACON

아리랑 위성영상 AI 객체 검출 경진대회

Nov 24 2020

Kyung Tae Kim (MTC Ai LAB & Seoul National University)

Index



Exploratory Data Analysis

1. Rotation Augmentation

Since there are many orientation variations in aerial images implement the online rotation augmentation. (e.g. 180, 270)

Fig.1 origin rotated

Fig.2 rotated

Fig.3 rotated

Exploratory Data Analysis

2. Mosaic Data Augmentation

This allows for the model to learn how to identify objects at a smaller scale than normal.
It also encourages the model to localize different types of images in different portions of the frame.

Fig.1 origin image

Fig.2 combines 4 images

2. Model

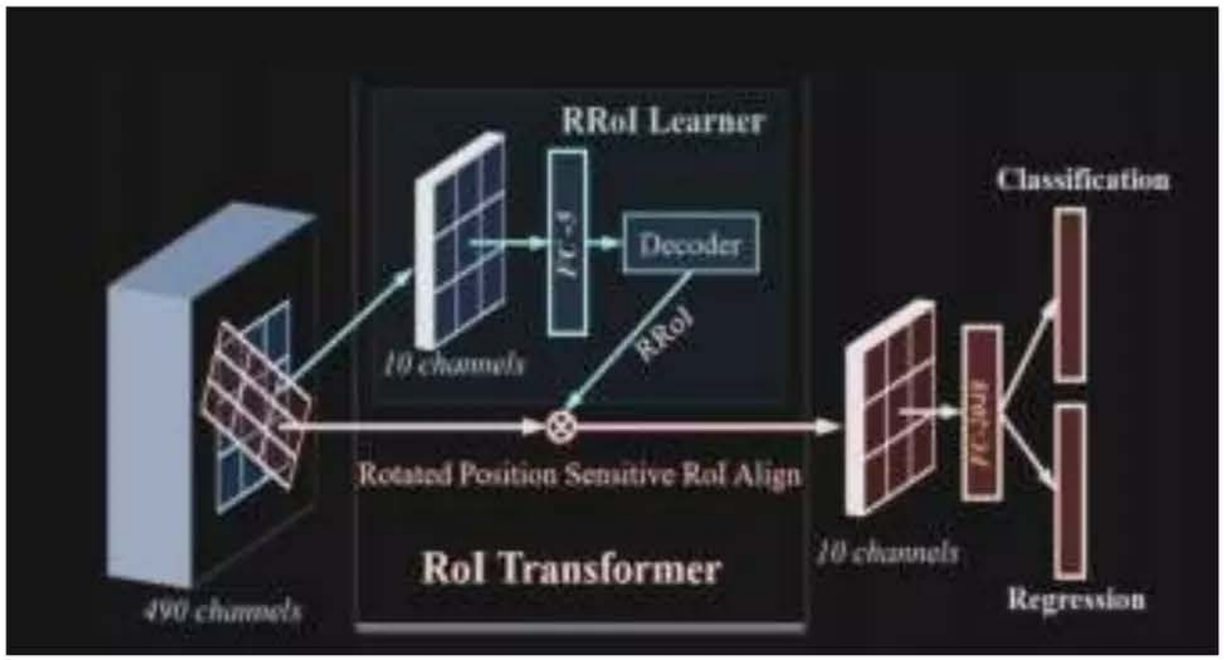


Fig.1 ROI Transform

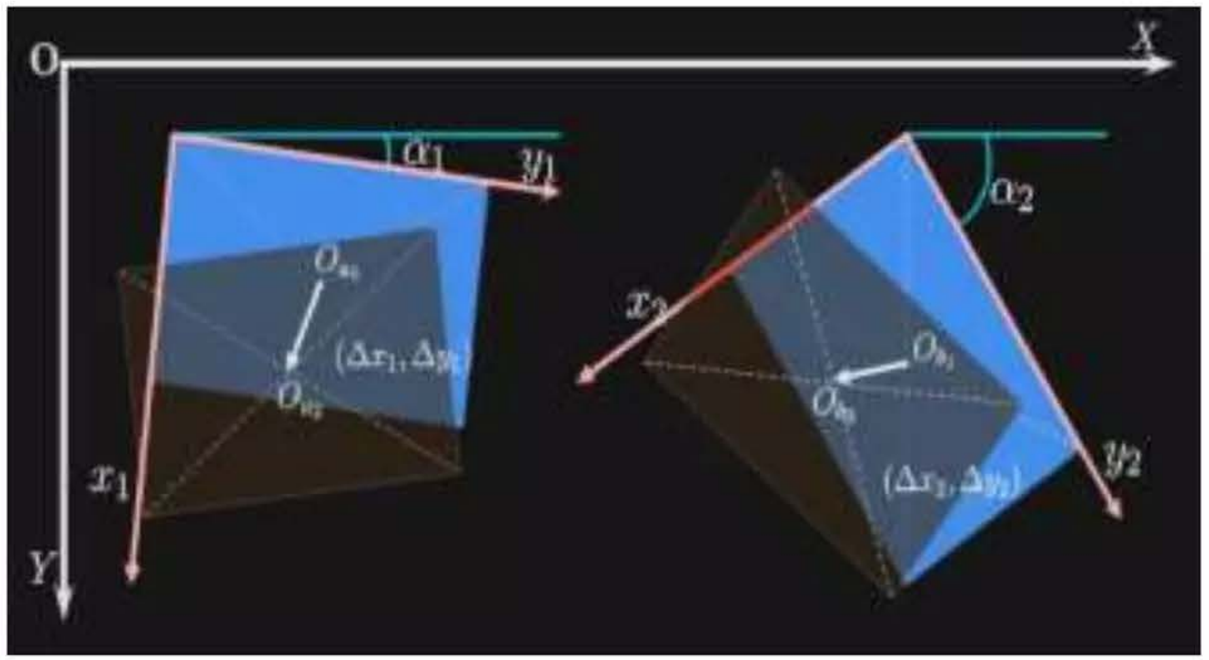


Fig.2 ROI Transform

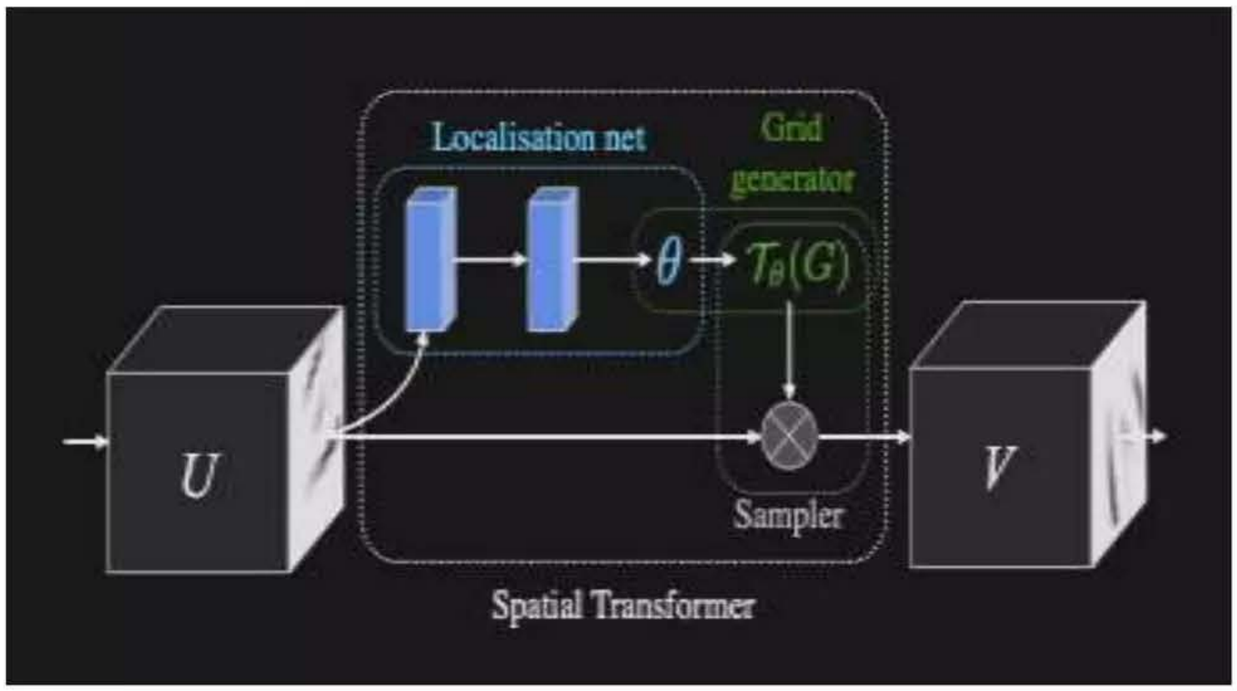


Fig.3 Spatial Transformer Network

Procedure 1 Routing algorithm.

1: **procedure** ROUTING($\hat{\mathbf{u}}_{j|i}, r, l$)
2: for all capsule i in layer l and capsule j in layer $(l+1)$: $b_{ij} \leftarrow 0$.
3: **for** r iterations **do**
4: for all capsule i in layer l: $\mathbf{c}_i \leftarrow \mathrm{softmax}(\mathbf{b}_i)$ ▷ softmax computes Eq. 3
5: for all capsule j in layer $(l+1)$: $\mathbf{s}_j \leftarrow \sum_i c_{ij} \hat{\mathbf{u}}_{j|i}$
6: for all capsule j in layer $(l+1)$: $\mathbf{v}_j \leftarrow \mathrm{squash}(\mathbf{s}_j)$ ▷ squash computes Eq. 1
7: for all capsule i in layer l and capsule j in layer $(l+1)$: $b_{ij} \leftarrow b_{ij} + \hat{\mathbf{u}}_{j|i} . \mathbf{v}_j$
return $\mathbf{v}_j$

Fig.4 Dynamic Routing Algorithm

2. Deformable Convolution Networks

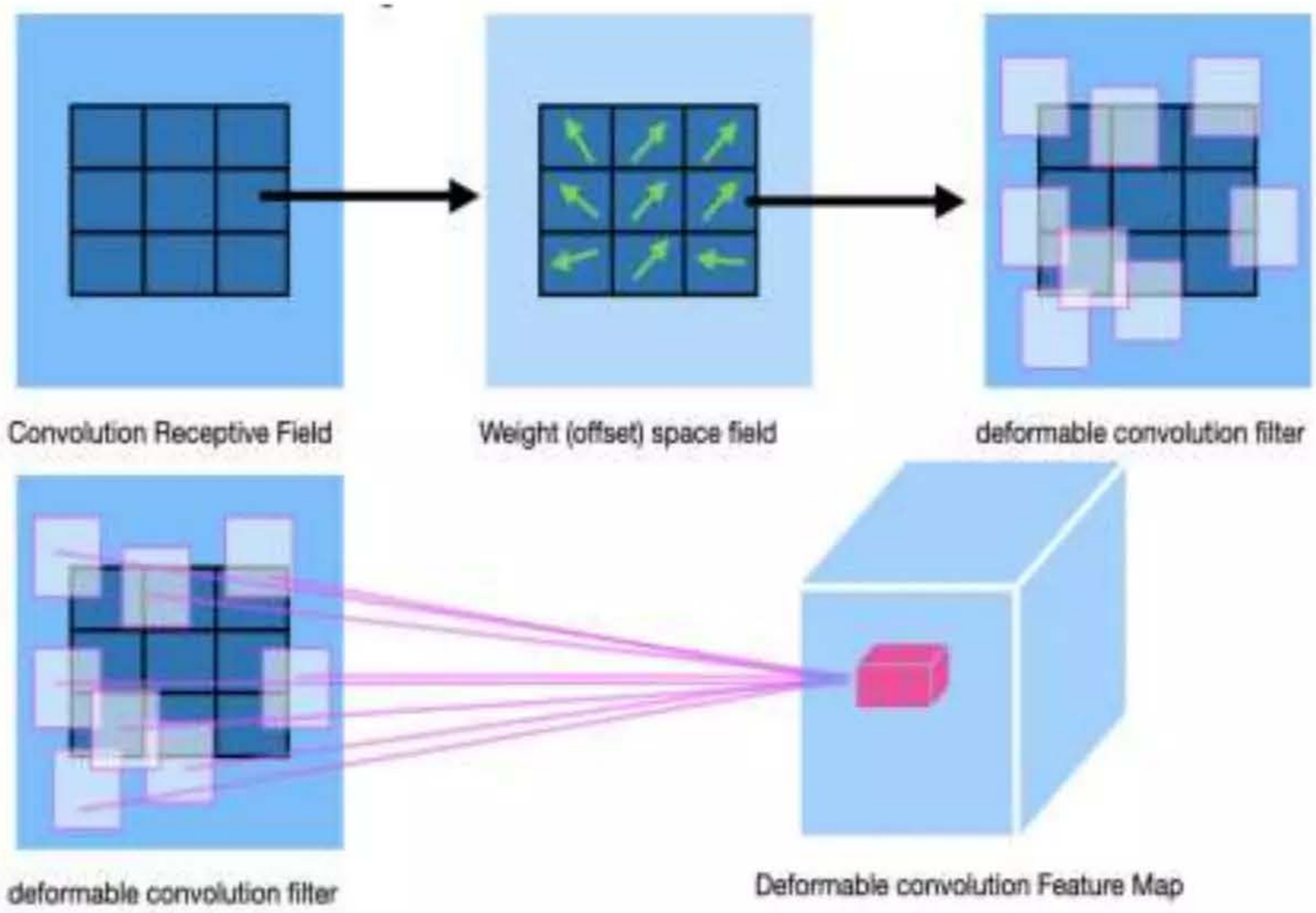


Fig.1 Training DCN Module

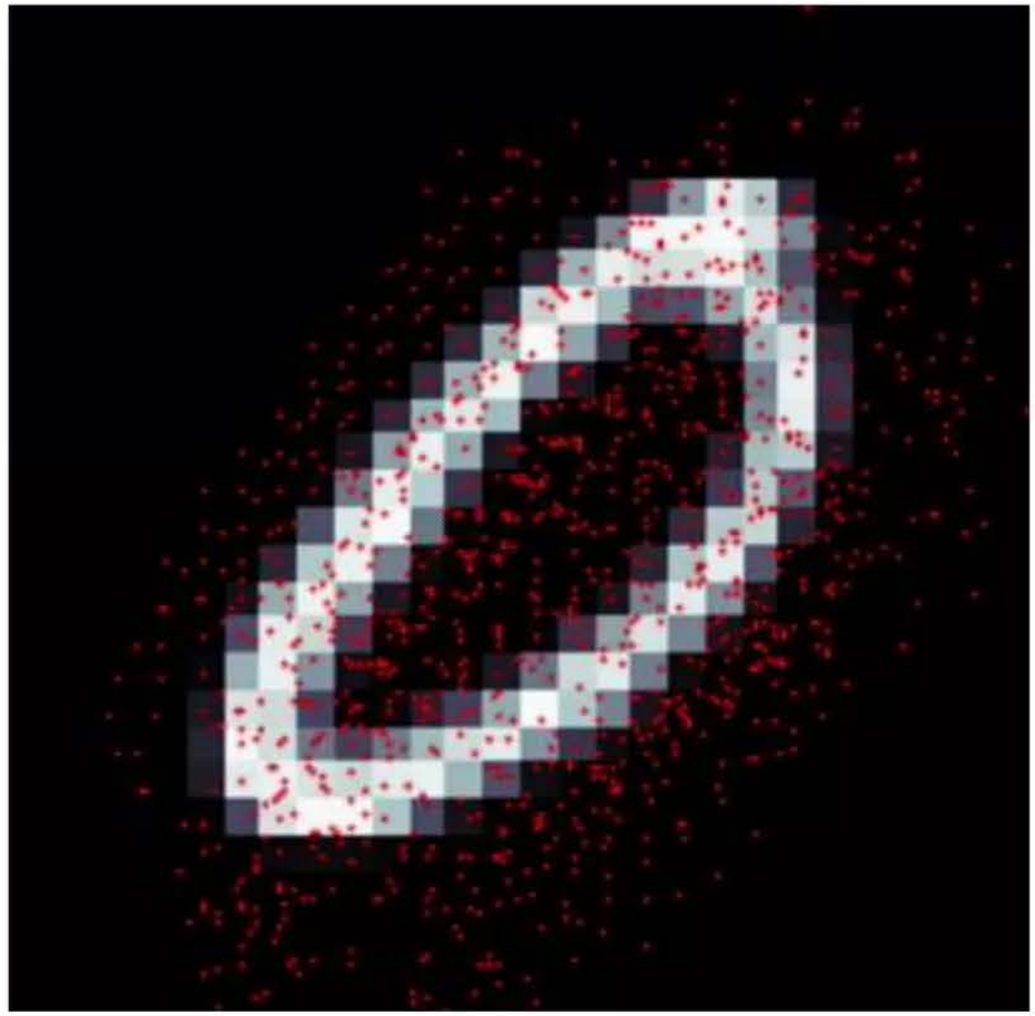


Fig.2 DCN sampling locations (from the learned offset)

1. **Enabling effective modeling of spatial transformation** in convolution neural network

2. **No additional supervision** for learning spatial transformation

3. **Significant accuracy improvements** on sophisticated vision tasks

2. Model

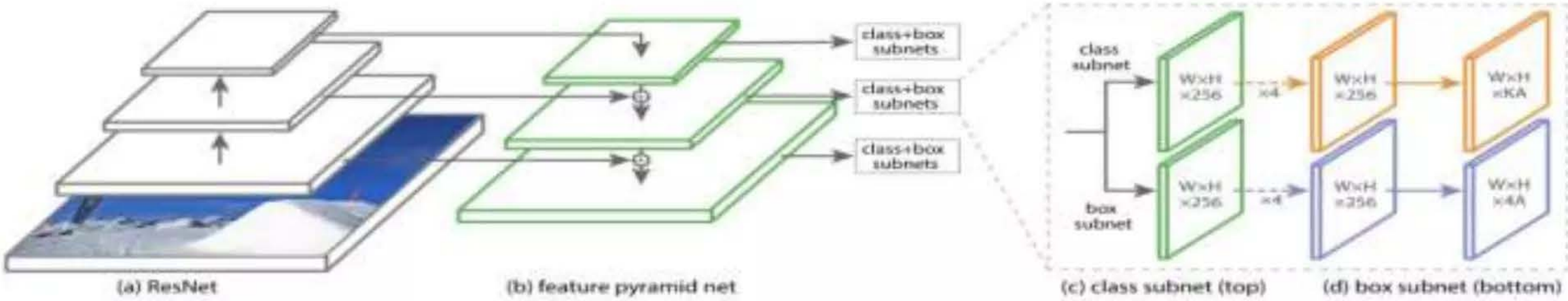


Fig.1 FPN

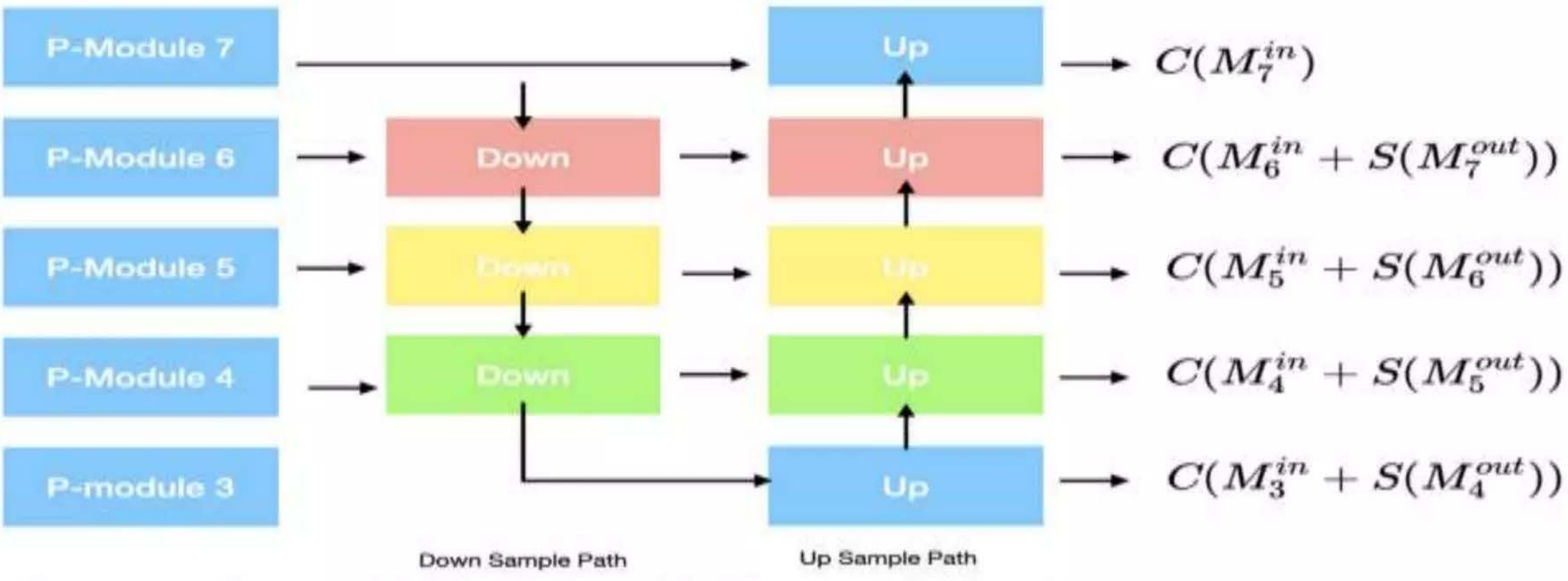
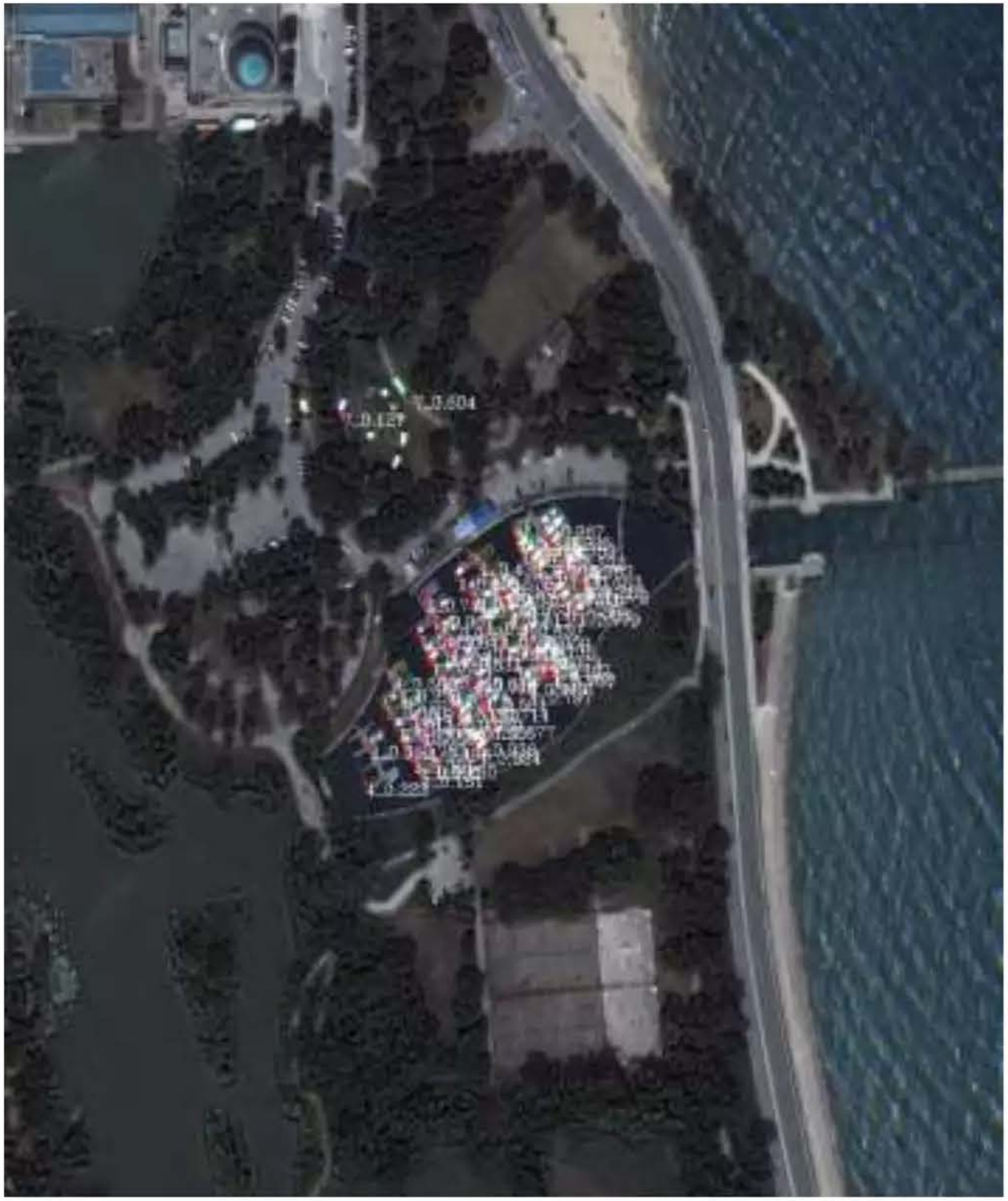
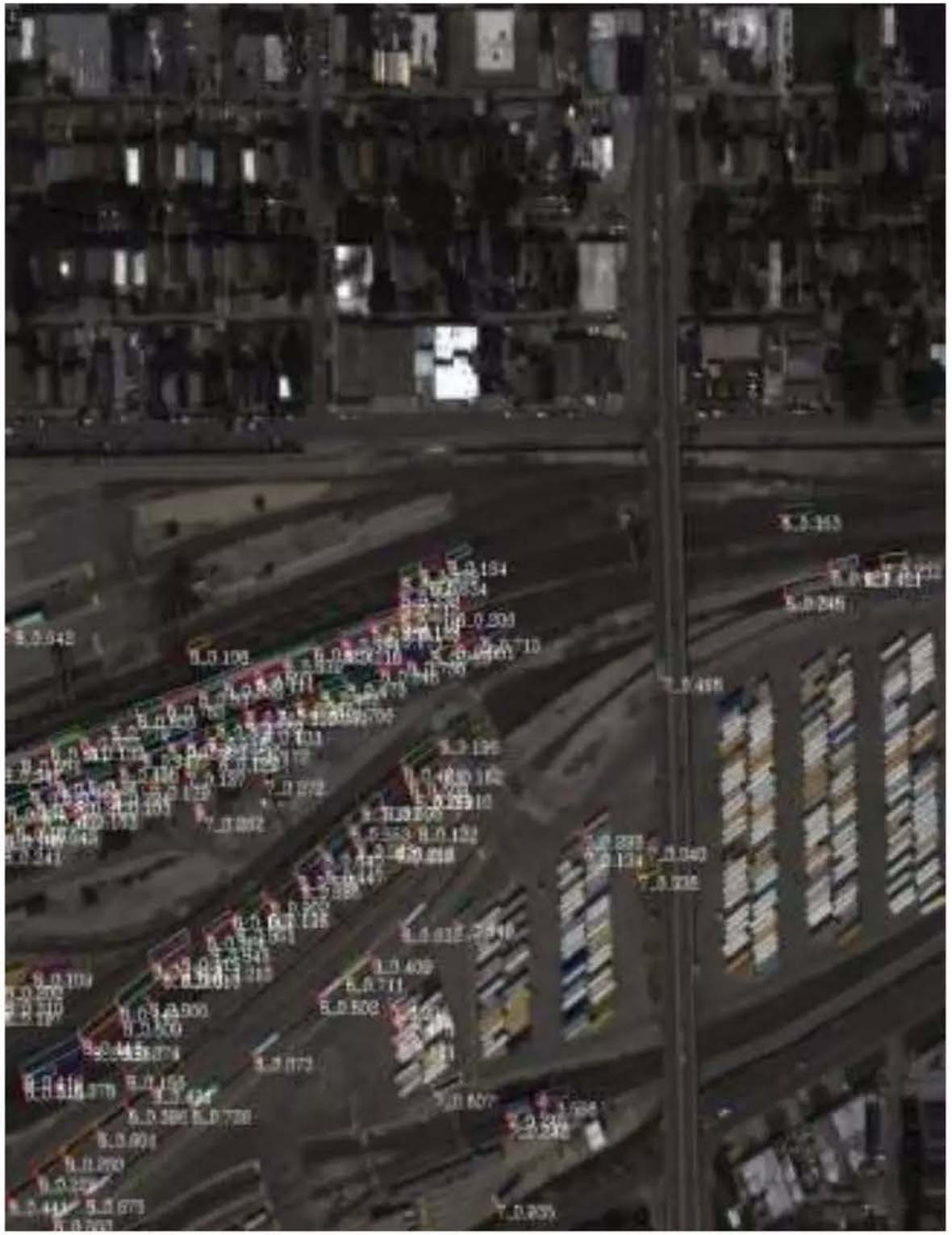


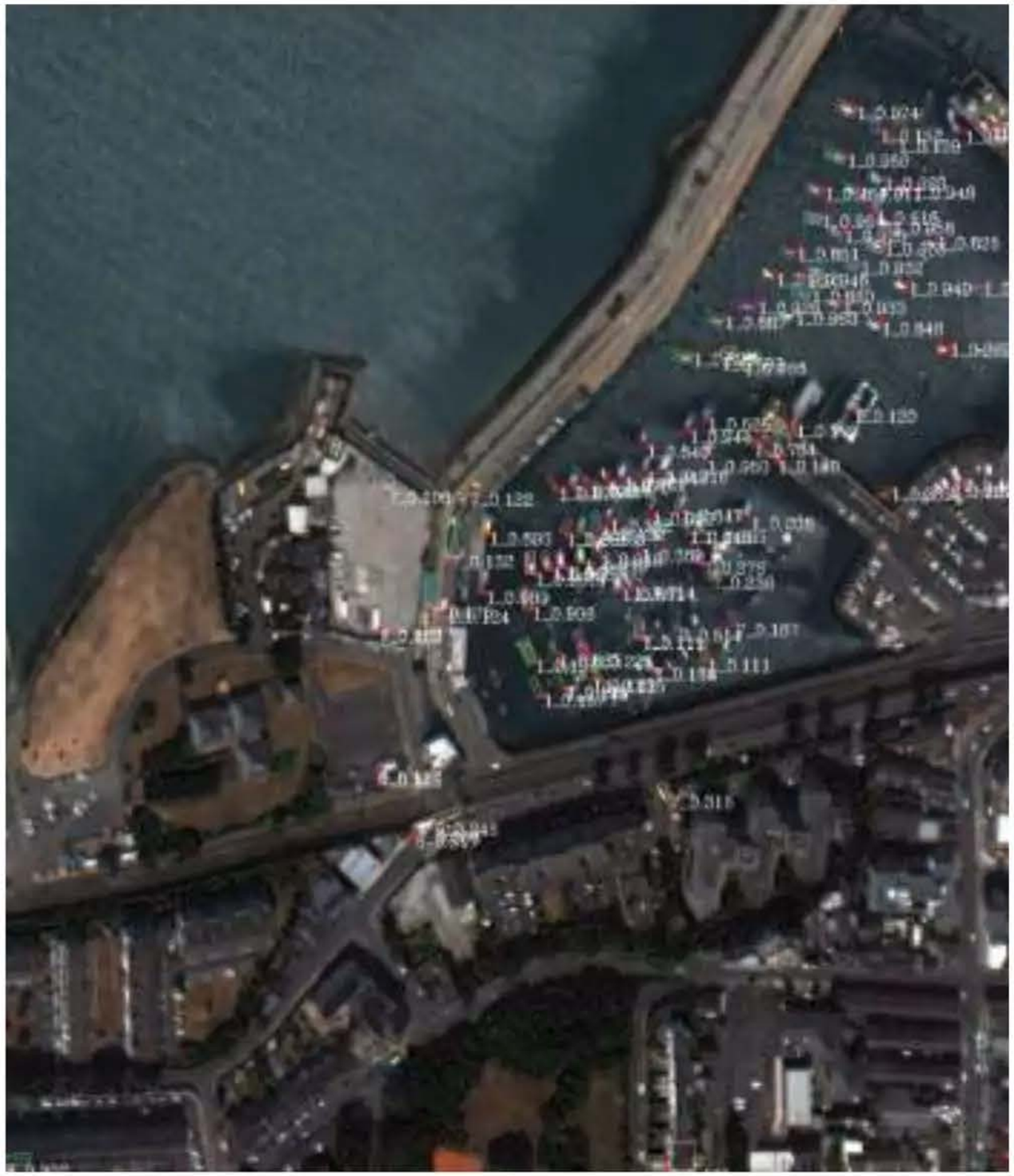
Fig.2 Expansion Pyramid Network

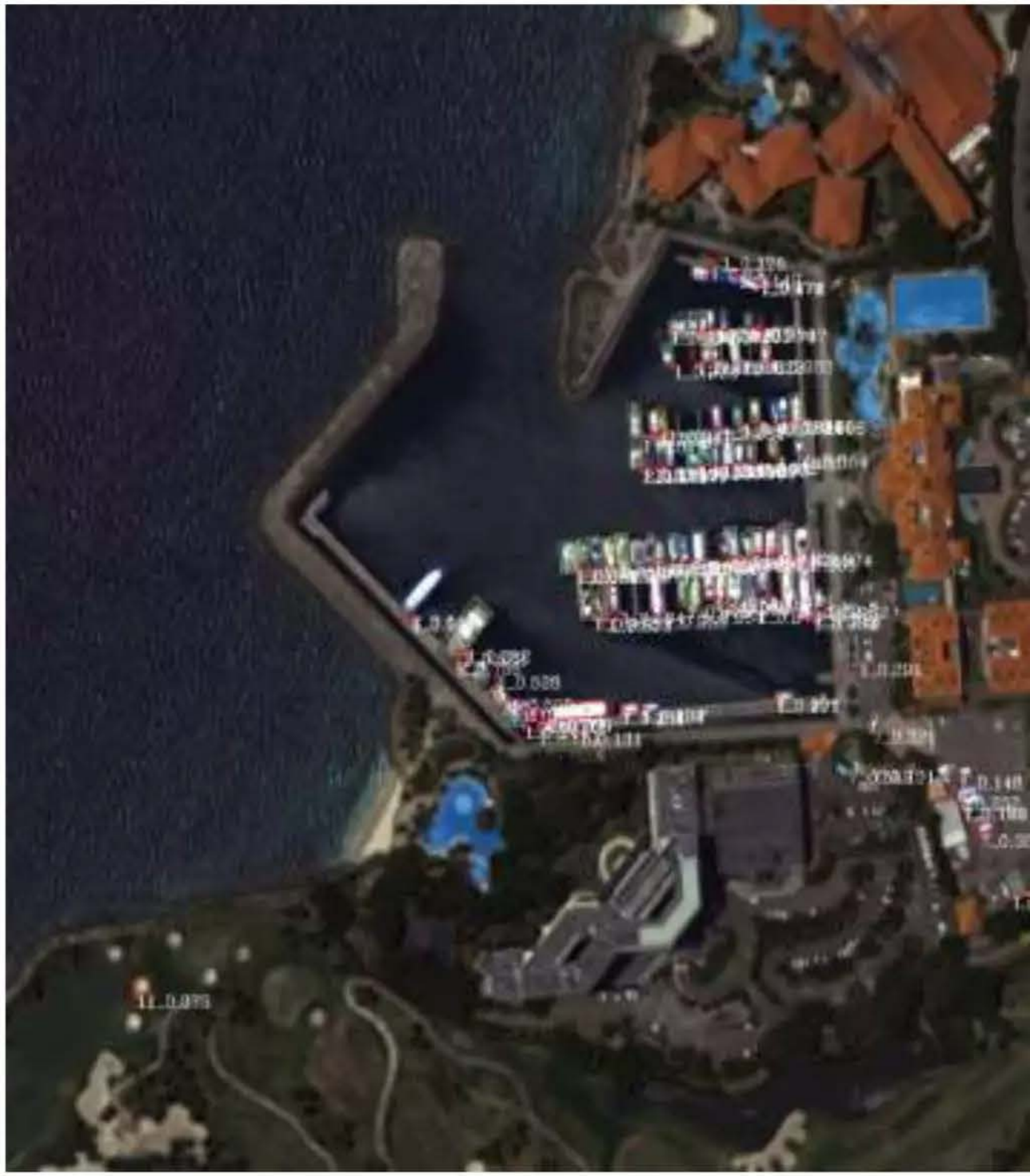
3. Result





3. Result





3. Attention Expansion Pyramid Network

3. Conclusion

1. Mean Average Precision
2. Align Deep Feature & Activation MAP
3. Explainable Neural Networks
4. Convolutional vs Transformers

3. Result

Need good system 1 functionality to make a system efficient

4. Reference

1. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks

2. Fully Convolutional Networks for Semantic Segmentation

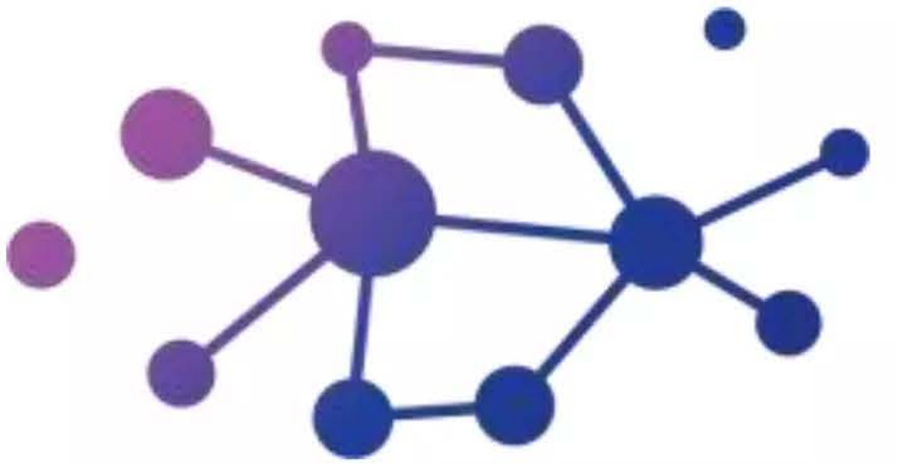
3. Focal Loss for Dense Object Detection

4. Deformable Convolutional Networks

5. Learning RoI Transformer for Detecting Oriented Objects in Aerial Images

6. YOLOv4: Optimal Speed and Accuracy of Object Detection

7. Arbitrary-Oriented Scene Text Detection via Rotation Proposals



THANK YOU