Digital Approaches to Cartographic Heritage
Thessaloniki 2019


# Detection of Pictorial Map Objects with Convolutional Neural Networks

Raimund Schnürer
René Sieber
Jost Schmid-Lanter
A. Cengiz Öztireli
Lorenz Hurni

ZB Zentralbibliothek Zürich

ETH zürich

Disney Research

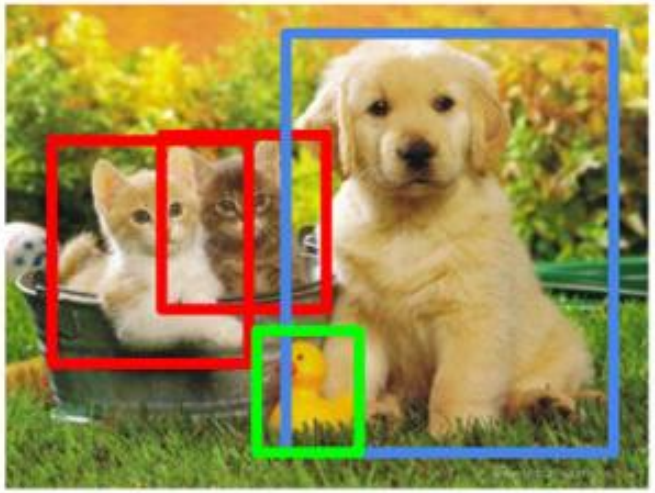

https://greekmythology.wikia.org/wiki/Talos

# Tasks

Classification

CAT

Object detection



CAT, DOG, DUCK

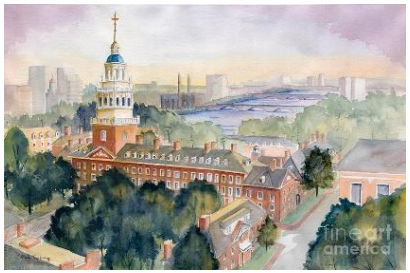
https://www.datacamp.com/community/tutorials/object-detection-guide

# Experiments

## Classification

## Object detection

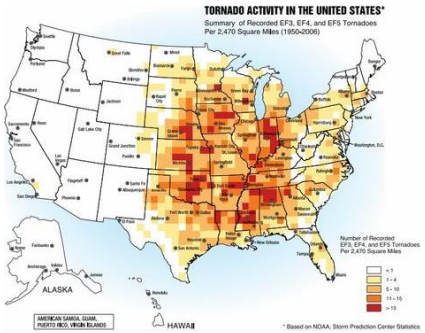
https://i.pinimg.com/736x/20/a5/3f/20a53fa9a346ddc6e722c86a77bd74e1--campus-map-state-university.jpg



https://fineartamerica.com/featured/harvard-university-melly-terpening.html

Maps vs. non-maps


https://i.pinimg.com/736x/74/64/eb/7464ebd207ec45992c590c29e53b4c17.jpg




https://i.pinimg.com/736x/0a/47/5e/0a475e6591388b74bb88bffaf35c0a6f.jpg

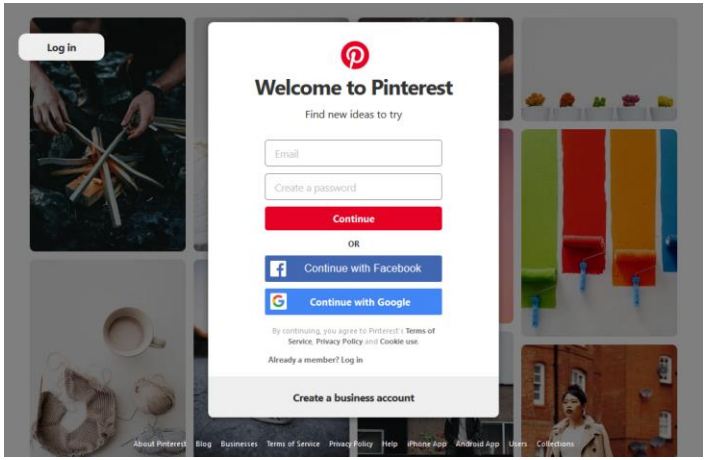
Pictorial maps vs. non-pictorial maps


https://gallica.bnf.fr/ark:/12148/btv1b84692644

Sailing ships

# Data

## Classification




https://www.pinterest.com/

Pinterest

3100 maps
(1500 pictorial, 1500 non-pictorial)
+
3100 non-maps

## Object detection


https://www.davidrumsey.com/

Digital map libraries

525 maps with 3200 ships

# Definitions

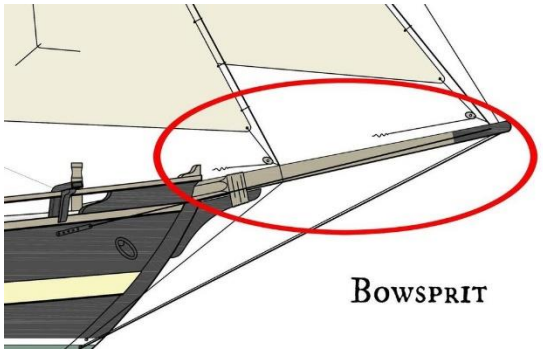
## Classification

A **map** is a
scaled-down
2D or 3D representation
– optionally animated and interactive –
of macroscopic spaces
– possibly with additional temporal and thematic information –
where features are symbolized and
relationships between them are mainly preserved.

A **pictorial map** is a
map with verisimilar and indexed representations,
which are rather individual than typified.

## Object detection

A **ship** is a
large sea-going vessel
with at least one mast
but not necessarily a bowsprit.

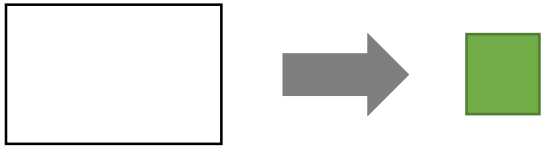



http://www.prominenceenergy.com.au/irm/content/what-s-a-bowsprit.aspx?RID=537

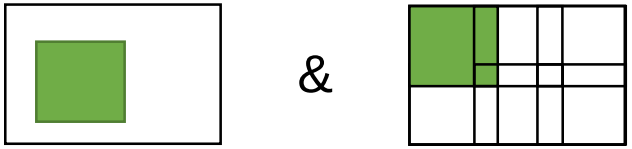
# Convolutional Neural Networks

## Classification

Xception vs. InceptionResNetV2
(both use 299px input images)



resizing
(train & eval)



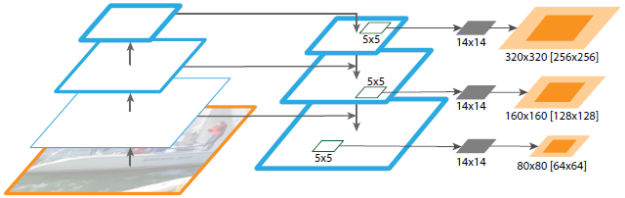
cropping
(train)

&

averaging
(eval)

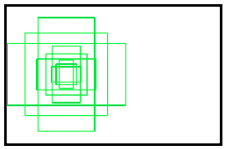
## Object detection

Faster R-CNN vs. RetinaNet
(both use ResNet50 classifier)




https://arxiv.org/abs/1612.03144

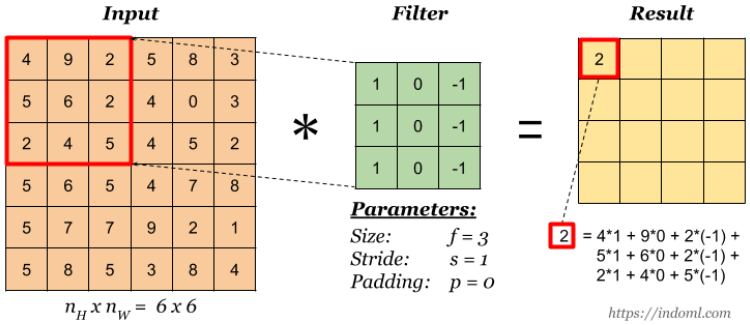
pyramiding



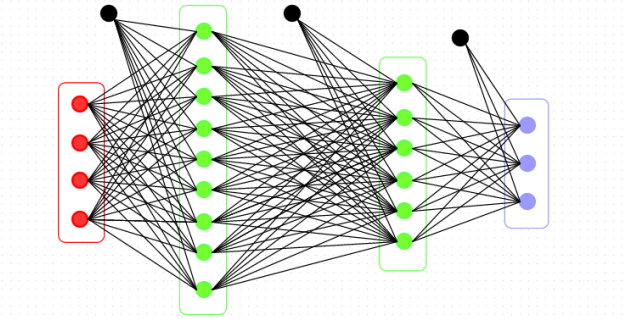
scaling of anchor boxes

TensorFlow and Keras (Python)
NVIDIA GTX 1080
Transfer learning
Training time: 1.5–3.0h

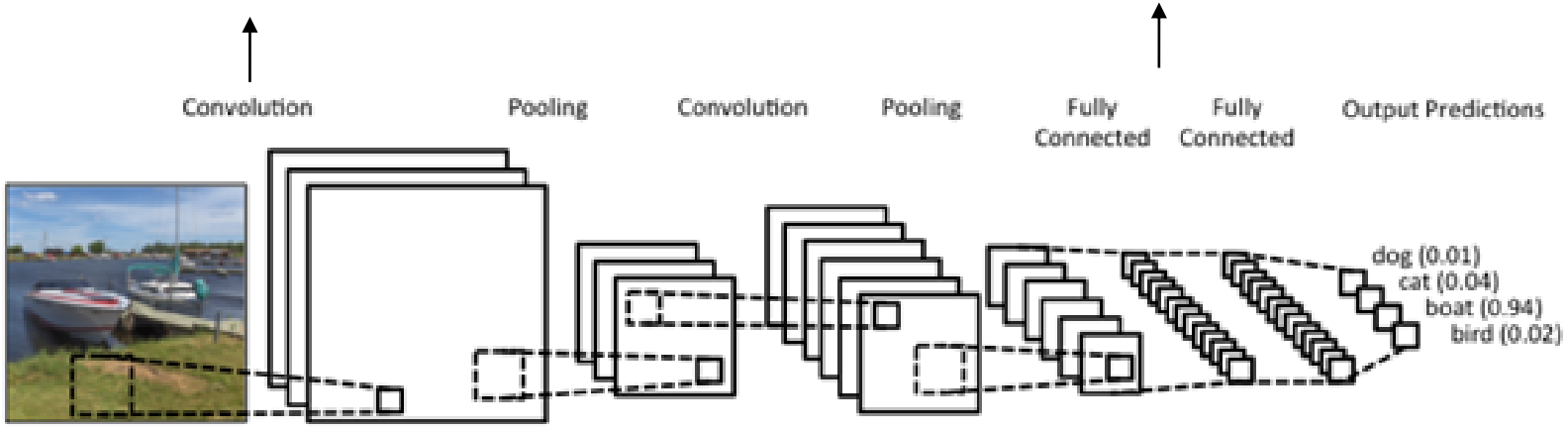
# Convolutional Neural Networks



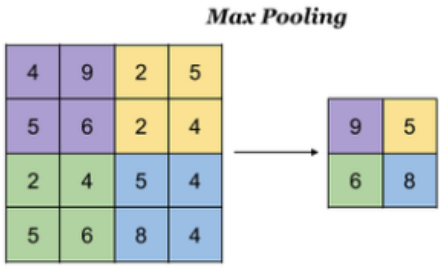

https://www.kaggle.com/manmohan291/57-cnn-basics



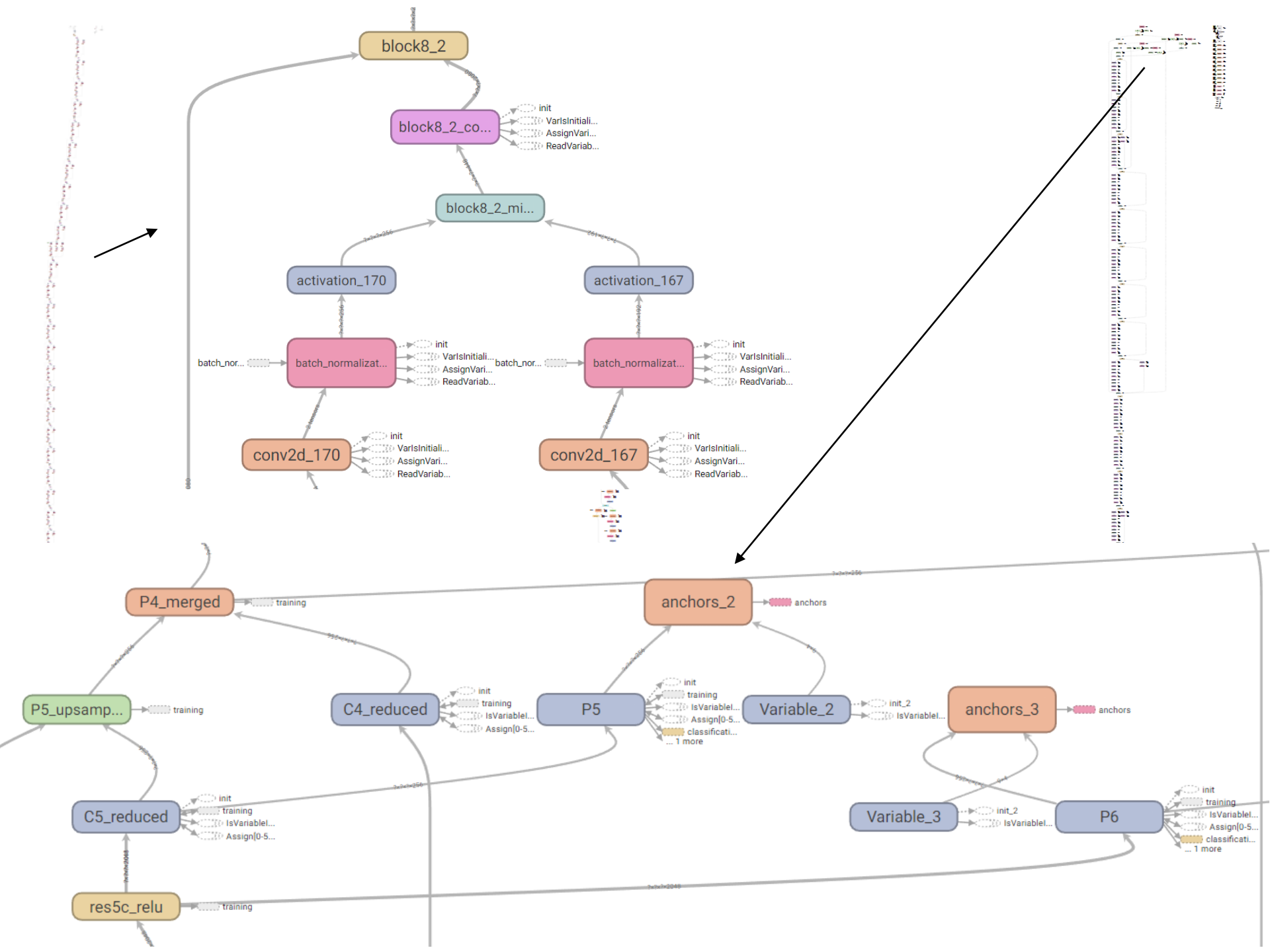
https://www.cyzne.com/are-there-several-types-of-deep-learning-models/




multiple sources




https://www.kaggle.com/manmohan291/57-cnn-basics


block8_2
block8_2_co...
init
VarIsInitiali...
AssignVari...
ReadVariab...
block8_2_mi...
activation_170
activation_167
batch_nor...
batch_normalizat...
conv2d_170
conv2d_167
P4_merged
training
anchors_2
anchors
P5_upsamp...
C4_reduced
P5
Variable_2
init_2
IsVariableI...
anchors_3
C5_reduced
IsVariableI...
Assign[0-5...
classificati...
... 1 more
Variable_3
P6
res5c_relu

# Results

## Classification accuracy

| | Xception | InceptionResNetV2 |
|---|---|---|
| resizing | 96.47% | 96.60% |
| averaging | 96.63% | **96.76%** |

Maps vs. non-maps

| | Xception | InceptionResNetV2 |
|---|---|---|
| resizing | 89.64% | 88.61% |
| averaging | **91.89%** | 90.83% |

Pictorial maps vs. non-pictorial maps

## Object detection average precision

| Scales (normal pyramid) | Faster R-CNN | RetinaNet |
|---|---|---|
| 0.5, 1.0, 1.5 | 20.24% | **35.37%** |
| 0.125, 0.25, 0.5, 1.0 | 24.93% | 32.62% |

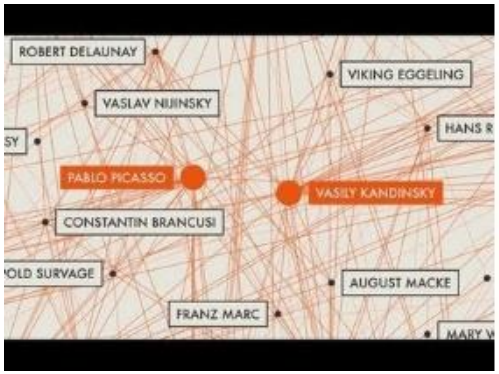
Sailing ships

| Scales (shifted pyramid) | Faster R-CNN | RetinaNet |
|---|---|---|
| 0.5, 1.0, 1.5 | 27.61% | **36.24%** |
| 0.125, 0.25, 0.5, 1.0 | 32.26% | 30.04% |

Sailing ships

# Results

CNNs misclassfied those as maps




https://i.pinimg.com/736x/bb/4c/52/bb4c5218917d2368937279be05deb528–moma-org-the-artist.jpg
https://i.pinimg.com/736x/bb/8d/1b/bb8d1b57149f189eaf3deed58e4a7482.jpg
https://i.pinimg.com/736x/c1/4c/0a/c14c0a9ec176a79addef054c1e134e95–heart-map-my-heart.jpg

CNNs misclassfied those as non-maps

https://i.pinimg.com/736x/9d/40/93/9d4093cebf375ef698c2022857b83de4–world-map-canvas-world-map-art.jpg
https://i.pinimg.com/736x/5e/94/a1/5e94a1054a88227364c82db48cbbd747–easter–food-design.jpg
https://i.pinimg.com/736x/f5/7d/38/f57d38e1acddef874dcb529fee617290–maps.jpg

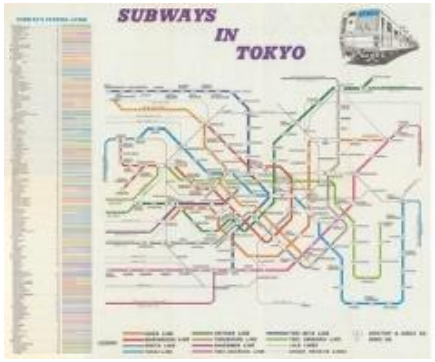
# Results

CNNs misclassfied those as pictorial maps



https://i.pinimg.com/736x/5c/0a/a1/5c0aa1e7bfd262b50d2b6e43c237b833.jpg
https://i.pinimg.com/736x/6f/34/73/6f34737d2cc282e23c5668217ddf3544–printable-maps-vintage-printable.jpg
https://i.pinimg.com/736x/5f/f7/a2/5ff7a22d4cdf35580c12a654ac208ca5–map-mind-illustrated-maps.jpg

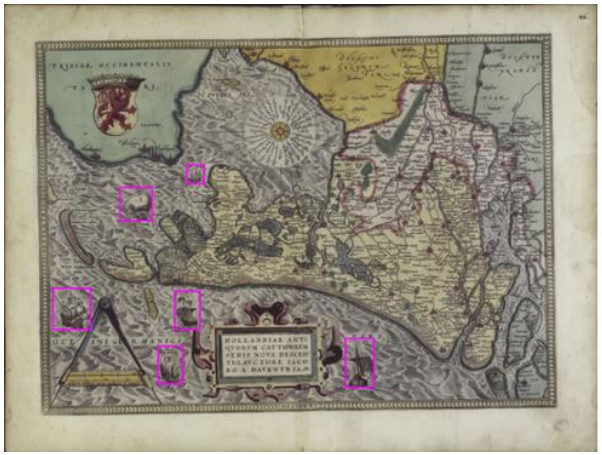
CNNs misclassfied those as non-pictorial maps

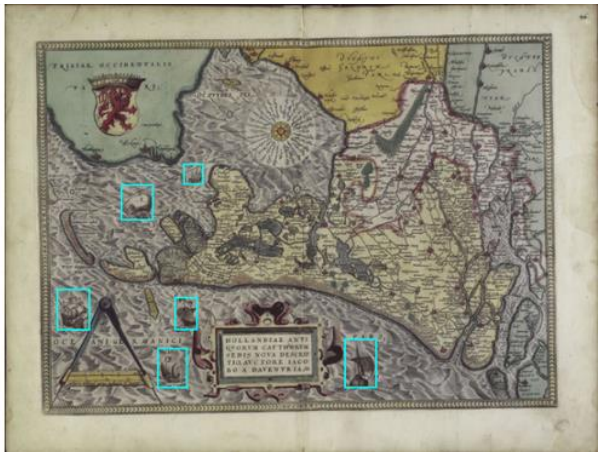



https://i.pinimg.com/736x/b6/d2/d1/b6d2d1375ac9808cc2998c814862e5d8.jpg
https://i.pinimg.com/736x/51/d7/8b/51d78ba2f5a0f8217a20dbb7fe8ca883–nice-map-beijing.jpg
https://i.pinimg.com/736x/50/a8/5a/50a85a6ea7ee74f36b5141138d47a281.jpg

# Results

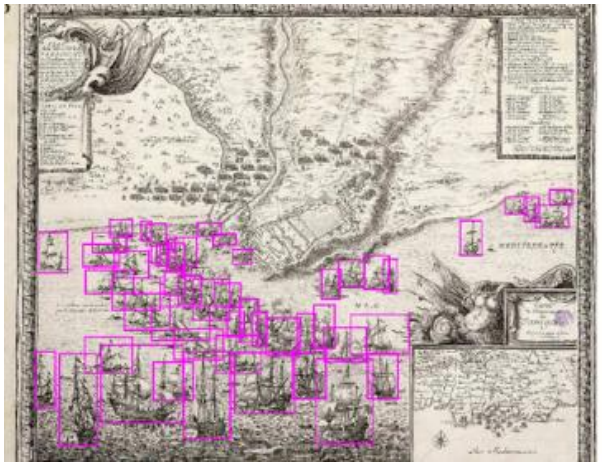
CNNs detect large, freestanding ships well

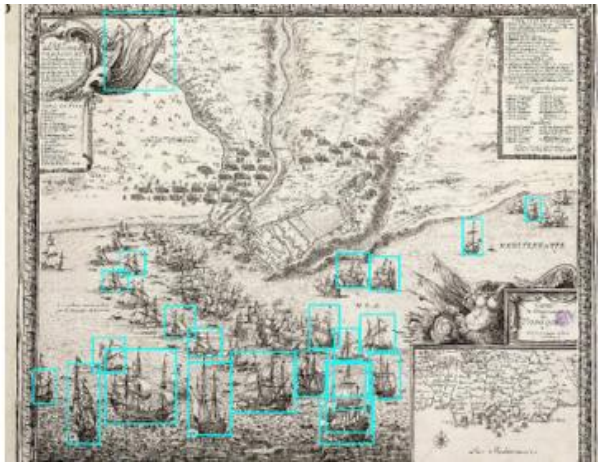




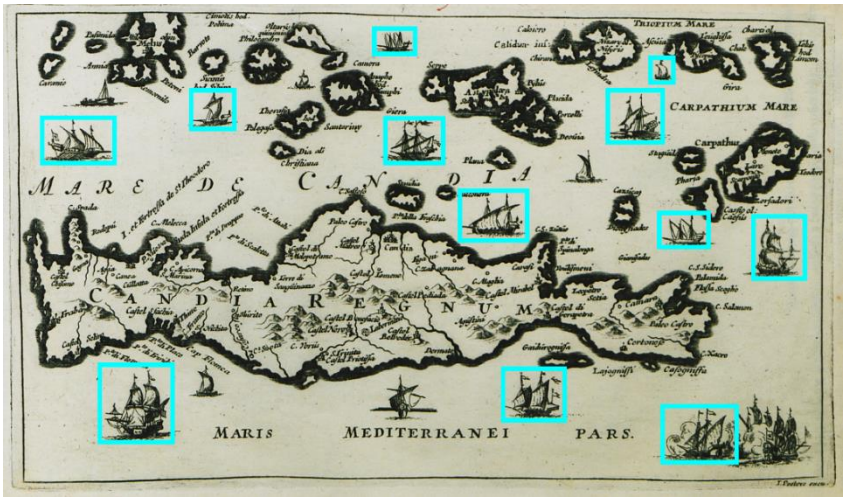
https://aleph.unibas.ch/F/?local_base=DSV01&con_lng=GER&func=find-b&find_code=SYS&request=001055619

CNNs often fail to detect occluded and small ships

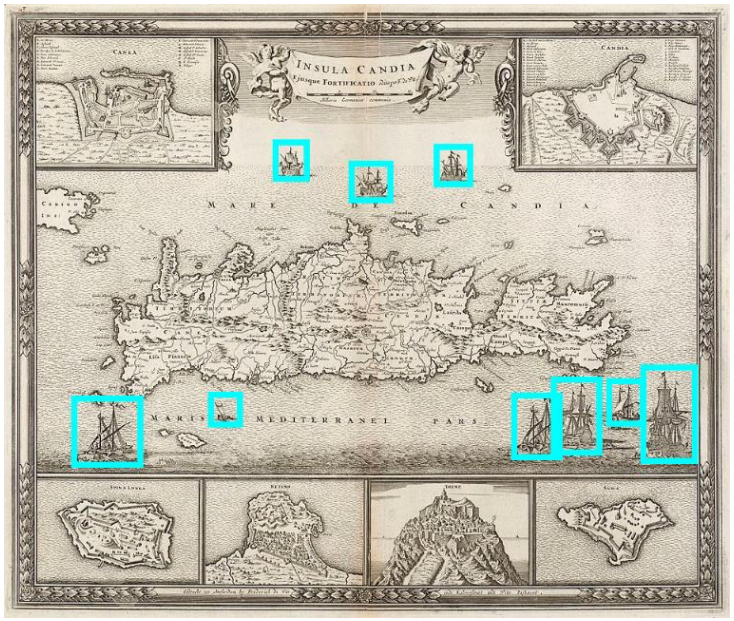




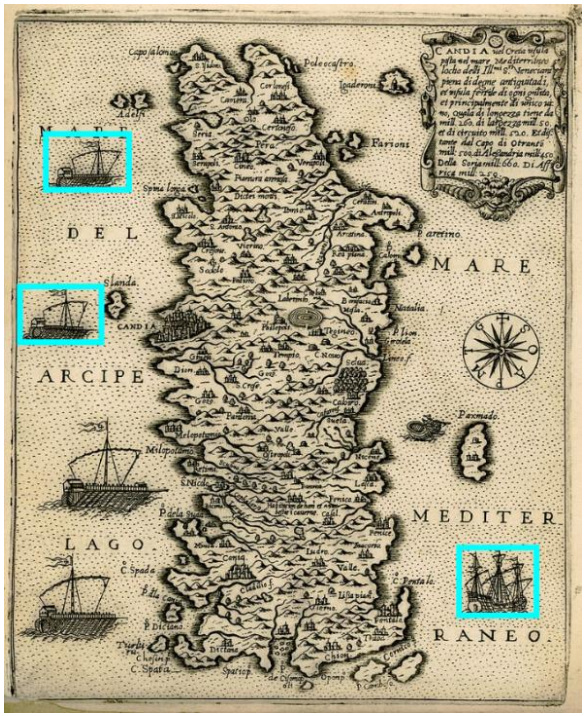
http://bdh.bne.es/bnesearch/detalle/bdh0000022147




https://commons.wikimedia.org/wiki/File:Candia_Regnum_-_Peeters_Jacob_-_1690.jpg




https://commons.wikimedia.org/wiki/File:Insula_Candia_(8343734844).jpg




https://commons.wikimedia.org/wiki/File:Candia_nel_Creta_insula_-_Camocio_Giovanni_Francesco_-_1574.jpg


ευχαριστώ
Training data,
code, and models:
http://narrat3d.ethz.ch/
This work is supported by ETH
Research Grant ETH-11 17-1.