

通用目标检测框架下的船只识别系统

“复杂海情和气象条件下的海上船只识别”

Ship Identification with General Object Detection Frameworks

2017 年 12 月

第 11 小组

李勰 徐易难 肖景匀 俞子舒 张远航

中国科学院大学



目录

组员介绍

问题概述

方法学

方案概述

总结



组员介绍



人员分工 (not final!)

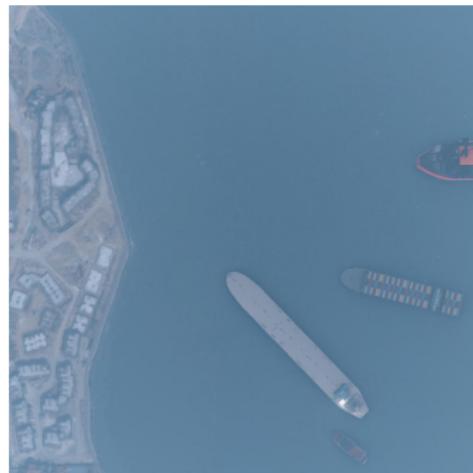
- Titan X: 独立完成了绝大部分工作
- 张远航: 训练网络, 文献调研 (关注网络结构)
- 俞子舒: 查找开源代码, fine-tune
- 徐易难: 数据集划分, 标签清理, 结果分析
- 肖景匀: 图片去雾, 分类筛选
- 李 魏: 数据集扩增, 文献调研

问题概述



题目要求

输入: 图片, 1024×1024



输出: 所有船只的位置, 大小及类别 (游轮、游艇、货船中的一种)

huochuan, 714 565 1009 692

huochuan, 460 562 810 918

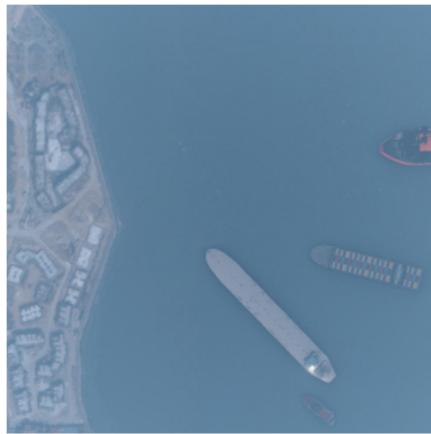
huochuan, 883 281 1024 398

youlun, 693 926 796 1009

目标所处的位置以及大小通常用矩形的界定框 (bounding box) 来表示: (x, y, w, h)

题目要求 (续)

输入：图片， 1024×1024

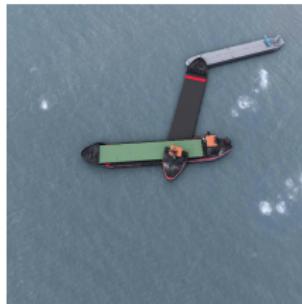


输出：所有船只的位置，大小及类别（游轮、游艇、货船中的一种）

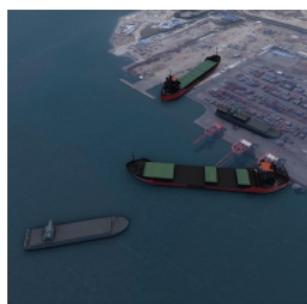


Are U Kidding Me

“以真实卫星影像为基础生成的模拟图像”



(a) 船身重叠



(b) 两栖攻击舰？

淡晴<raytong52@qq.com> 14:36:38

然后我模型精度越高分数越低 😂

少飞侯(2363405977) 14:38:51

有可能哟



图 1: 不合理样本示例

Bad ground truth labels hurt performance!

性能度量

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{Missed}}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$F_1 = \frac{2PR}{P + R}$$

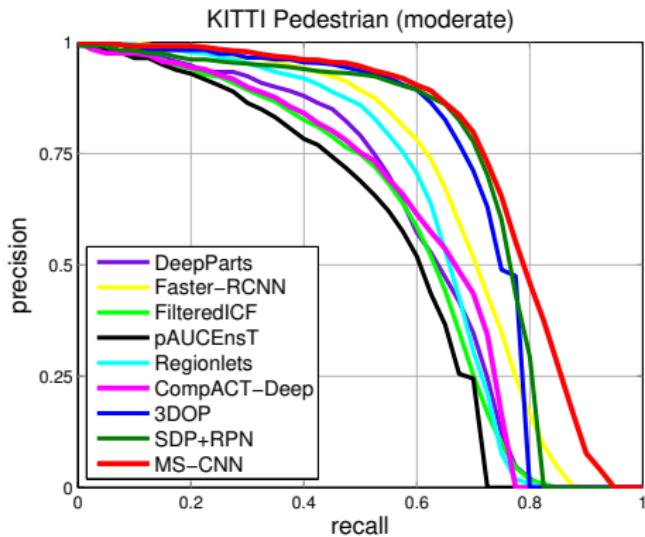


图 2: PR 精度评价曲线

方法学



常用的目标检测器

组员介绍

问题概述

方法学

基于区域候选框的 R-CNN 系列

方案概述

尺度差异

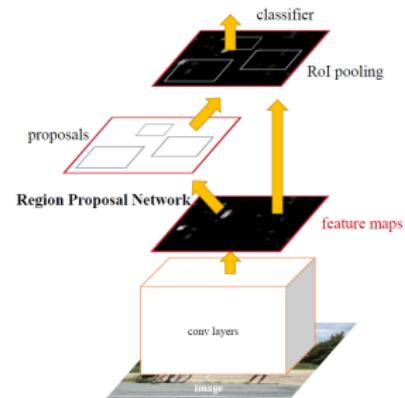
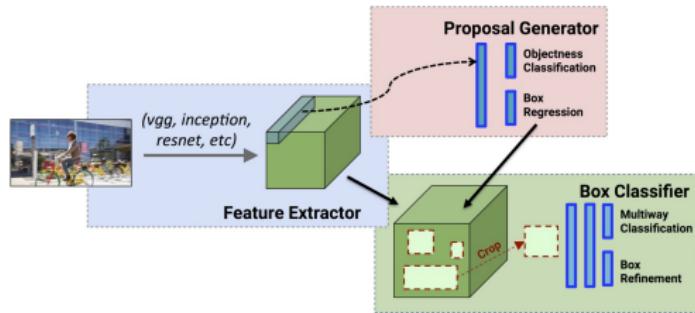
Race to the Top

总结



Faster R-CNN

Faster Region-based Convolutional Networks



- 两个 CNN 共享卷积层
 - 生成候选窗口：RPN
 - 检测：Fast R-CNN

方案概述



去雾

基于暗通道先验 (dark channel prior) 的单图去雾

- 雾成像模型: $I(x) = J(x)t(x) + A(1 - t(x))$



去雾

基于暗通道先验 (dark channel prior) 的单图去雾

- 雾成像模型: $I(x) = J(x)t(x) + A(1 - t(x))$
- 大气光成分 A (雾最浓区域) 的估计 ?

去雾

基于暗通道先验 (dark channel prior) 的单图去雾

- 雾成像模型: $I(x) = J(x)t(x) + A(1 - t(x))$
- 大气光成分 A (雾最浓区域) 的估计 ?
- 轻度



去雾 (续)

- 中度



- 按直方图分类

测试/训练集划分

- 留出法，训练集：测试集 = 4 : 1，随机打乱



测试/训练集划分

- 留出法，训练集：测试集 = 4 : 1，随机打乱
- 刚开始训练时的一段对话

XXX(6XXX8) 18:38:20

我只训了 3000 步

XXX(6XXX8) 18:38:35

基本上没有训练到游轮

XXX 19:34:02

不对呀，都打乱过吧

XXX(4XXX8) 19:34:24

最后给你们的压缩包应该都是乱的 除了分类

XXX(6XXX8) 20:38:52

我感觉是游轮太少了吧



测试/训练集划分

- 留出法，训练集：测试集 = 4 : 1，随机打乱
- 刚开始训练时的一段对话

XXX(6XXX8) 18:38:20

我只训了 3000 步

XXX(6XXX8) 18:38:35

基本上没有训练到游轮

XXX 19:34:02

不对呀，都打乱过吧

XXX(4XXX8) 19:34:24

最后给你们的压缩包应该都是乱的 除了分类

XXX(6XXX8) 20:38:52

我感觉是游轮太少了吧

- 类别严重不均衡

测试/训练集划分

- 留出法，训练集：测试集 = 4 : 1，随机打乱
- 刚开始训练时的一段对话

XXX(6XXX8) 18:38:20

我只训了 3000 步

XXX(6XXX8) 18:38:35

基本上没有训练到游轮

XXX 19:34:02

不对呀，都打乱过吧

XXX(4XXX8) 19:34:24

最后给你们的压缩包应该都是乱的 除了分类

XXX(6XXX8) 20:38:52

我感觉是游轮太少了吧

- 类别严重不均衡

- 初赛：货船：6299，游轮：2442，游艇：4473，渔船：239

测试/训练集划分

- 留出法，训练集：测试集 = 4 : 1，随机打乱
- 刚开始训练时的一段对话

XXX(6XXX8) 18:38:20

我只训了 3000 步

XXX(6XXX8) 18:38:35

基本上没有训练到游轮

XXX 19:34:02

不对呀，都打乱过吧

XXX(4XXX8) 19:34:24

最后给你们的压缩包应该都是乱的 除了分类

XXX(6XXX8) 20:38:52

我感觉是游轮太少了吧

- 类别严重不均衡

- 初赛：货船：6299，游轮：2442，游艇：4473，渔船：239
- 复赛：货船：10757，游轮：**2929**，游艇：**618**

测试/训练集划分

- 留出法，训练集：测试集 = 4 : 1，随机打乱
- 刚开始训练时的一段对话

XXX(6XXX8) 18:38:20

我只训了 3000 步

XXX(6XXX8) 18:38:35

基本上没有训练到游轮

XXX 19:34:02

不对呀，都打乱过吧

XXX(4XXX8) 19:34:24

最后给你们的压缩包应该都是乱的 除了分类

XXX(6XXX8) 20:38:52

我感觉是游轮太少了吧

- 类别严重不均衡

- 初赛：货船：6299，游轮：2442，游艇：4473，渔船：239
 - 复赛：货船：10757，游轮：**2929**，游艇：**618**

- 怎么办？

类别不均衡

解决方法：过采样！

“A SYSTEMATIC STUDY OF THE CLASS IMBALANCE PROBLEM IN CNN”

- The method that in most of the cases outperforms all others with respect to multiclass ROC AUC was **oversampling**.
- Oversampling should be applied to the level that totally eliminates the imbalance.
- As opposed to some classical machine learning models, **oversampling does not necessarily cause overfitting of convolutional neural networks**.



数据集扩增

- Random saturation



数据集扩增

- Random saturation
- Random cropping



数据集扩增

- Random saturation
- Random cropping
- Random contrast



数据集扩增

- Random saturation
- Random cropping
- Random contrast
- Random dehazing efforts



尺度问题

固定大小感受野不能很好处理不同尺度的目标

图片中可能出现各种不同尺度的物体（黄色的界定框）。RPN 的单一感受野（阴影区域）不能应对这种多变性。[Cai 2016]

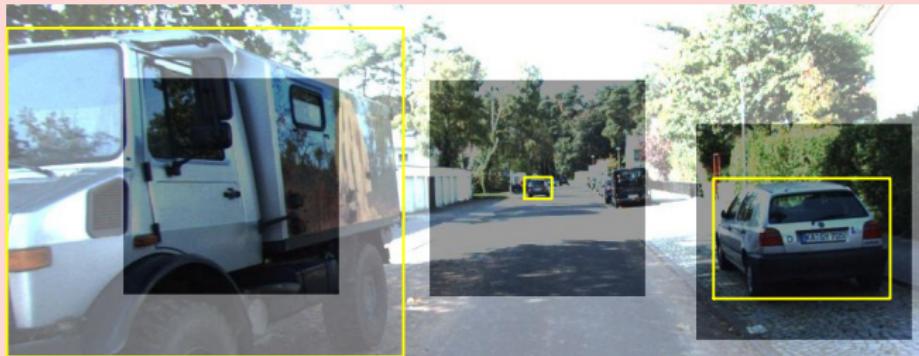


图 3: KITTI 数据集中的尺度差异

MSCNN (Multi-Scale CNN)

Cai et. al., 2016

候选框生成子网络中，检测在多个输出层完成，感受野便可匹配不同尺度的物体。

特征图上采样

- 传统多尺度策略：图像上采样得到图像金字塔（慢，占内存，小目标响应弱）
- MSCNN：
 - 只输入一个尺度的特征图
 - 引入反卷积层，特征图上采样（快，小目标可产生大而强的响应区域）

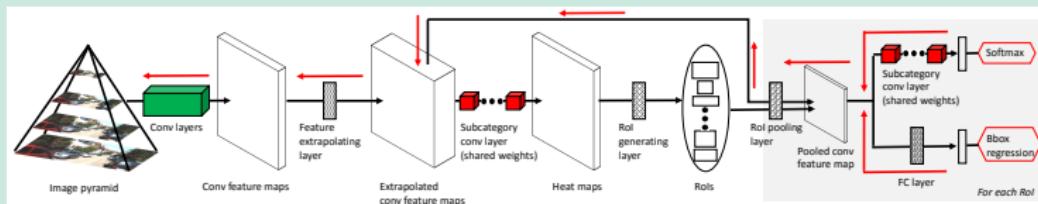
(* 感受野：卷积神经网络每一层输出的特征图 (feature map) 上的像素点在原始图像上映射的区域大小)

Sub-CNN (Subcategory-aware CNN)

Xiang et. al., 2016

特征外插层

- 图像上采样，特征图外插
- 输入 N (不会太大!) 个不同尺度下的特征图
- 每两个尺度之间双线性外插 M 个中间尺度，得到 $N' = (N - 1)M + N$ 个特征图



其他可能的性能提升来源 ?

- Multi-crop Inference



- Ensembling: diversity is key

总结



Takeaway Messages

- Ground truth matters



Takeaway Messages

- Ground truth matters
- Image dehazing works very well



Takeaway Messages

- Ground truth matters
- Image dehazing works very well
- Class imbalance: oversample

Takeaway Messages

- Ground truth matters
- Image dehazing works very well
- Class imbalance: oversample
- Accuracy: handling objects with scale variation



Takeaway Messages

- Ground truth matters
- Image dehazing works very well
- Class imbalance: oversample
- Accuracy: handling objects with scale variation
 - feature up-sampling by deconvolution / feature extrapolation

Takeaway Messages

- Ground truth matters
- Image dehazing works very well
- Class imbalance: oversample
- Accuracy: handling objects with scale variation
 - feature up-sampling by deconvolution / feature extrapolation
 - multi-scale feature fusion

参考文献

Dehazing and others

- Berman, Dana, and Shai Avidan. "Non-local image dehazing." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.
- He, Kaiming, Jian Sun, and Xiaoou Tang. "Single image haze removal using dark channel prior." IEEE transactions on pattern analysis and machine intelligence 33.12 (2011).
- Buda, Mateusz, Atsuto Maki, and Maciej A. Mazurowski. "A systematic study of the class imbalance problem in convolutional neural networks." arXiv preprint arXiv:1710.05381 (2017).

参考文献（续）

Object Detection

- Ren, Shaoqing, et al. "Faster R-CNN: Towards real-time object detection with region proposal networks." Advances in neural information processing systems. 2015.
- Huang, Jonathan. "Object Detection at Google". IEEE SIDAS 2017 Industry Keynote.
- “目标检测一唱三叹：更准，更快，更智能”. 邬书哲，2016.

参考文献（续）

State-of-the-art

- Cai, Zhaowei, et al. "A unified multi-scale deep convolutional neural network for fast object detection." European Conference on Computer Vision. Springer International Publishing, 2016.
- Hong, Sanghoon, et al. "PVANet: Lightweight Deep Neural Networks for Real-time Object Detection." arXiv preprint arXiv:1611.08588 (2016).
- Ren, Jimmy, et al. "Accurate Single Stage Detector Using Recurrent Rolling Convolution." arXiv preprint arXiv:1704.05776 (2017).
- Xiang, Yu, et al. "Subcategory-aware convolutional neural networks for object proposals and detection." IEEE Winter Conference on Applications of Computer Vision. IEEE, 2017.

Thank you!

Code and models will be available at:

<https://github.com/sailordiary/ship-identification>

The screenshot shows the GitHub repository page for 'sailordiary / ship-identification'. The repository name is at the top left. At the top right, there are buttons for 'Unwatch' (with a count of 3), 'Star' (0), 'Fork' (0), and 'Edit'. Below the header, there are tabs for 'Code' (selected), 'Issues' (0), 'Pull requests' (0), 'Projects' (0), 'Wiki', 'Insights', and 'Settings'. A description below the tabs reads 'Ship Identification with General Object Detection Frameworks (BDCI 2017)' with an 'Edit' button. There is also an 'Add topics' link. Below the description, there are summary statistics: '2 commits', '1 branch', '0 releases', and '1 contributor'. A dropdown menu shows 'Branch: master' and a 'New pull request' button. To the right are buttons for 'Create new file', 'Upload files', 'Find file', and a green 'Clone or download' button. The main content area shows a commit history. The first commit is by 'sailordiary' titled 'Update README.md', with a timestamp of 'Latest commit 976379e 7 hours ago'. Below it is another commit for 'README.md' with the same update message and timestamp. A third commit is listed as 'README.md' with a timestamp of '7 hours ago'. At the bottom of the commit list is a separator line consisting of 12 '=' characters. Below this line, the repository's purpose is summarized in bold text: 'Ship Identification with General Object Detection Frameworks'.

Questions?