Adaptive Brightness Learning For Active Object Recognition

Nuo Xu[1,2], Chunlei Huo[1,2], Chunhong Pan[1,2]

1. NLPR, Institute of Automation, Chinese Academy of Sciences ; 2. University of Chinese Academy of Sciences

INTRODUCTION

Traditional object recognition approaches are limited due to the passive nature:

- **In imaging procedure, images are acquired driven by visual inspection rather than object recognition performance.**
- **In object recognition procedure, images are directly used for training or testing without active adjustment.**

Adaptive brightness learning procedure is described as follows:

- **At step t, an image $I(t)$ is represented by the contextual feature F_c and the brightness histogram feature F_b.**
- **Agent selects action $A(t)$ based on current state $S(t)$. The image $I(t+1)$ is obtained by applying an action $A(t)$ on $I(t)$.**
- **Reward $R(t+1)$ is calculated based on the change of detection performance between $I(t+1)$ and $I(t)$.**
- **$I(T)$ can be obtained, after applying a series of actions to $I(0)$.**

MOTIVATION

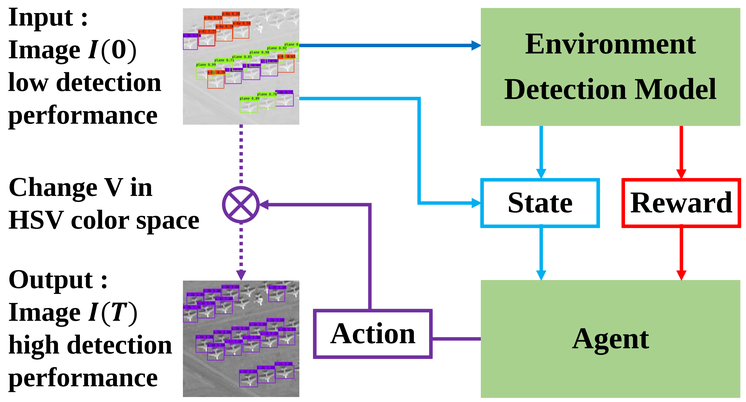

Fig. 1 Motivation of the proposed approach.

Environment
Detector YOLOv3.
Agent
Fully connected network
State $S(t)$
Features of the image
Reward $R(t)$
+1, if performance is improved. Otherwise, -1
Action $A(t)$
Brighten or darken

METHODOLOGY

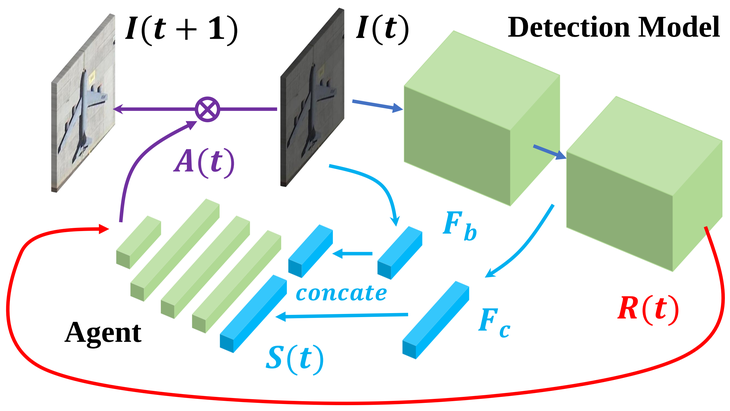

Fig. 2 Overview of the proposed approach.

Contextual feature F_c
F_c is the middle output of YOLOv3(13×13×1024). By taking the average along channel, F_c is straightened into a 169-dimension vector.

Brightness feature F_b
F_b is the histogram of the component V in HSV color space, of 64 dimensions.

EXPERIMENT

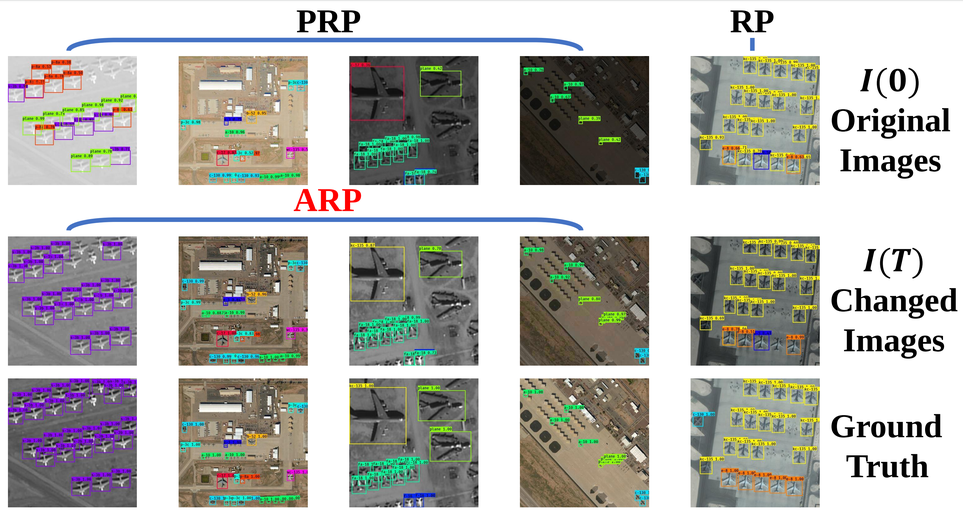

Fig. 3 Recognition performance comparison.

Tab. 1 Performance comparison of different methods.

Model+Backbone	RP	PRP	ARP
YOLOv3+Darknet53	0.807	0.714	0.798
Faster RCNN+VGG16	0.727	0.471	0.687
Faster RCNN+ResNet50	0.730	0.507	0.677
R-FCN+ResNet50	0.740	0.589	0.700
R-FCN+ResNet101	0.749	0.577	0.711
SSD300+VGG16	0.802	0.539	0.580
RetinaNet+VGG16	0.799	0.282	0.559
RetinaNet+ResNet50	0.795	0.401	0.643
RetinaNet+ResNet101	0.802	0.468	0.678

TIPS

RP:
Reference Performance (mAP 50) of normal brightness images

PRP:
Passive Recognition Performance (mAP 50) of bright or dark images

ARP:
Active Recognition Performance (mAP 50) of changed images