Small Object Detection by Generative and Discriminative Learning

Yi Gu, Jie Li, Chentao Wu, Weijia Jia and Jianping Chen

Department of Computer Science and Engineering, Shanghai Jiao Tong University, China
State Key Laboratory of Internet of Things for SmartCity, University of Macau, China
Institute of Electronics and Information Engineering, Suzhou University of Science and Technology, China
E-mail: {louisgu, lijiecs}@sjtu.edu.cn

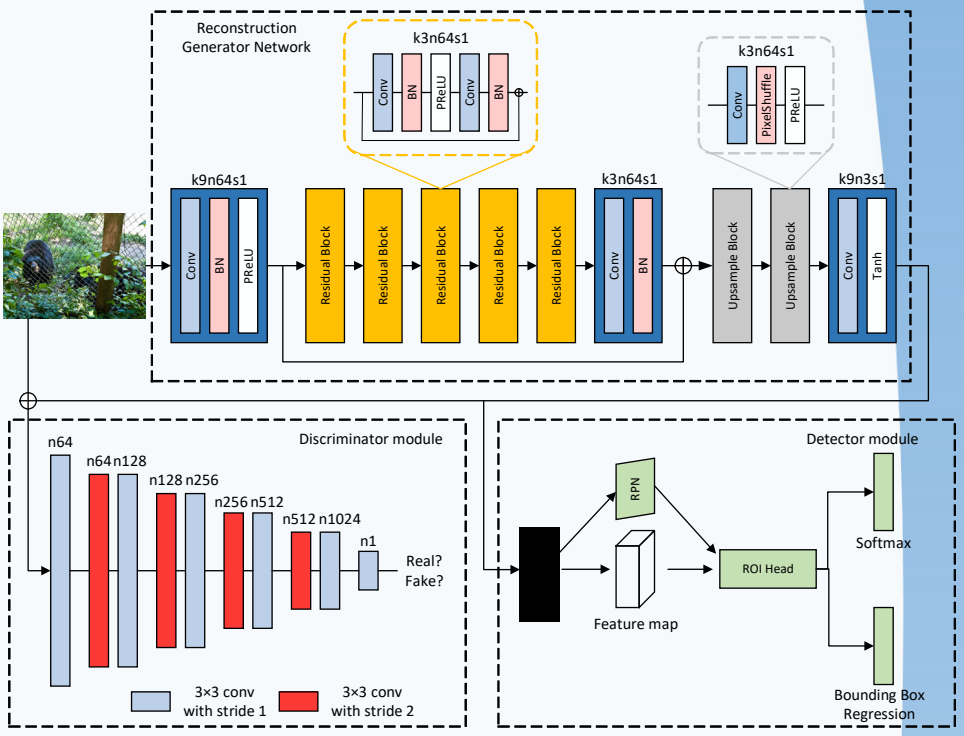
Problem

Small object detection is a challenging problem due to the limited information. Existing methods focus on improving classification accuracy but still suffer from the limitation of bounding box prediction.

Solution

We argue a preferable way to implement a reconstruction network on the detection module, which is different from the previous work that applies reconstruction to the discriminator module. The purpose of such a framework design is to deblur objects and produce more details for bounding box prediction so that the detection module can identify the anchor box of small object more precisely.

Result



Framework

We propose a detection framework By generative and discriminative learning.
A reconstruction generator network is designed to reconstruct the mapping from low frequency to high frequency for anchor box prediction.
A detector module extracts the regions of interest (ROIs) from generated results and implements a RoI-Head to predict object category and refine bounding box.

- $L_{adv} = \frac{1}{m}\sum_{i=1}^{m} \log(1 - D_{\theta_D}(G_{\theta_G}(z_{LR}^{(i)})))$
- $L_{MSE} = \frac{1}{m}\sum_{i=1}^{m} \left\| z_{HR}^{(i)} - G_{\theta_G}(z_{LR}^{(i)}) \right\|_2^2$
- $L_{TV} = \frac{1}{r^2WH}\sum_{i=1}^{rW}\sum_{j=1}^{rH} \left\| \nabla G_{\theta_G}(z_{LR}^{(i,j)}) \right\|$
- $L_{cls} = \frac{1}{m}\sum_{i=1}^{m}[-\left(\log\left(p^{(i)}p^{*(i)} + (1-p^{(i)})(1-p^{*(i)})\right) - \log(D_{cls}\left(G_{\theta_G}\left(z_{LR}^{(i)}\right)\right)\right) + \log\left(D_{cls}\left(z_{HR}^{(i)}\right)\right)]$
- $L_{reg} = \frac{1}{m}\sum_{i=1}^{m}\sum_{j\in(x,y,w,h)}[u_i \geq 1]\left(S_{L_1}\left(t_{SR}^{i,j} - t^{*(i,j)}\right)\right)$
- $\max_{\theta_D}\min_{\theta_G}(\frac{1}{m}\sum_{i=1}^{m}\log D_{\theta_D}(z_{HR}^{(i)})) + \alpha\, L_{adv} + \beta\, L_{cls} + \gamma\, L_{reg} + L_{TV} + L_{MSE}$